

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE VETERINARIA**



**TESIS DOCTORAL**

**Desarrollo de métodos para la detección de gluten en  
alimentos, a partir de repertorios de anticuerpos  
recombinantes derivados de la respuesta inmune celíaca**

**Development of methods for gluten detection in foods,  
applying recombinant antibody libraries derived from the  
celiac immune response**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR**

**PRESENTADA POR**

**Eduardo Rafael García Calvo**

**Directores**

**Teresa García Lacarra  
Rosario Martín de Santos  
Aina García García**

**Madrid**

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE VETERINARIA**



**TESIS DOCTORAL**

DESARROLLO DE MÉTODOS PARA LA DETECCIÓN DE GLUTEN EN ALIMENTOS, A PARTIR DE REPERTORIOS DE ANTICUERPOS RECOMBINANTES DERIVADOS DE LA RESPUESTA INMUNE CELÍACA

DEVELOPMENT OF METHODS FOR GLUTEN DETECTION IN FOODS, APPLYING RECOMBINANT ANTIBODY LIBRARIES DERIVED FROM THE CELIAC IMMUNE RESPONSE

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

EDUARDO RAFAEL GARCÍA CALVO

DIRECTORAS

Teresa García Lacarra  
Rosario Martín de Santos  
Aina García García









# FINANCIACIÓN

La presente tesis doctoral ha sido realizada gracias al contrato FPI PRE2018- 08427 del Ministerio de Ciencia e Innovación del Gobierno de España, que incluyó la financiación de la estancia en el centro de investigación en el extranjero VTT-Technical Research Center of Finland en Espoo, Finlandia por la que se opta a la Mención Internacional.

El trabajo experimental y la asistencia a congresos durante la realización de la presente tesis doctoral ha recibido financiación de los siguientes proyectos:

- Proyecto “Generación de conocimiento” Tipo B 2021. Título: “Generación de repertorios de anticuerpos recombinantes para el desarrollo de inmunoensayos destinados al control de alérgenos en la industria alimentaria” (PID2021-122925OB-I00). Ministerio de Ciencia e Innovación. 2022-2025. IP: Teresa García Lacarra.
- Proyecto “Programa de actividades de I+D entre grupos de investigación en Tecnologías” 2018. Título: “Teranóstica de las VNP para alergias alimentarias” (P2018/BAA-4547). 2019-2022. Comunidad de Madrid. IP (programa): Araceli Díaz Perales. IP (grupo): Teresa García Lacarra
- Proyecto “I+D+I (Retos)” Tipo B 2017. Título: “Producción y modificación funcional de anticuerpos recombinantes frente a nuez, pistacho, gluten y látex: trazabilidad y control de alérgenos en la industria alimentaria” (AGL2017-84316-R). 2017-2021. Ministerio de Ciencia y Universidades. IP: Teresa García Lacarra y María del Rosario Martín de Santos.



*A mis padres*



*“Si no has encontrado nada extraño durante el día,  
no ha sido un buen día”*

*John A. Wheeler (Físico teórico)*

*“Si es una buena idea, continuad y llevadla a término.*

*Es mucho más fácil pedir disculpas que conseguir el permiso necesario”*

*Grace Hopper (Oficial naval y una de las primeras programadoras de ordenadores)*



# AGRADECIMIENTOS

La sabiduría popular lo ha tenido muy claro desde siempre, “es de bien nacidos ser agradecidos”, y afortunadamente, tengo esta página para hacerlo. En cuanto a expresar sentimientos quizás no soy la persona más abierta del mundo, pero desde el fondo de mi alma, quiero recordar y agradecer una y mil veces a todas las personas que, de una manera u otra, me han ayudado en los últimos años.

Esta tesis doctoral puede parecer un compendio de datos, gráficas y explicaciones tediosas, y lo es, pero detrás hay un camino arduo y complejo en el que, al desarrollar una investigación para descubrir un proceso, una teoría o una molécula en el laboratorio acabas, inesperadamente, descubriéndote a ti mismo.

Tengo que empezar agradeciendo a las personas sin las cuales nada de esto hubiera sido posible, mis directoras. Gracias Teresa por haber estado ahí en los momentos de alegría y de tristeza, por haber sacado tiempo de dónde no lo había para atenderme y por haber apreciado y ayudado en aquellos proyectos laterales, más arriesgados y sin garantías, que se nos fueron ocurriendo. Gracias Charo por tu pragmatismo, por enseñarme a convertir las ideas en realidades y por el curso acelerado de atención al detalle que me has dado en los últimos meses, sin duda, lo necesitaba. Me dirijo a ambas cuando os digo, sé que ha sido difícil, tenemos ideas bastante diferentes en prácticamente todo, pero este trabajo es la prueba de que hemos conseguido tener una gran sinergia. Aina, gracias por tu apoyo, por compartir tu excelencia en el laboratorio como compañera, y ahora como directora, por el gran trabajo que has realizado para encauzar y terminar esta tesis.

A Eduardo, mi tocayo y jefe, e Isra, compañero del IQOG-CSIC, de cierta forma con vosotros empezó todo, en ese laboratorio me picó el gusanillo por la ciencia y, gracias a eso, ahora estoy aquí.

Santi, mi compañero de sitio, de poyata, de café, un grandísimo compañero, gracias por aguantarme día a día con toda mi intensidad, por las conversaciones científicas llenas de creatividad y por el “apoyo técnico” en los congresos. Raquel, aunque sólo coincidimos un tiempo al principio de mi tesis, te agradezco tener esa sonrisa y esa alegría contagiosa.

A mis chicas del despacho, Irene y Ester, gracias por estar ahí, mis terapeutas particulares, dispuestas a aguantar cualquier tontería o dramita que se me ocurra, todo ha sido más fácil desde que llegasteis.

Al resto de becarios-precarios, compañeros de café, de arriba y abajo, los que seguís por el departamento o los que os fuisteis a conocer mundo, gracias por los cafés y las charlas para desconectar, despotricar y quejarnos. Gracias al resto de profesores e investigadores del departamento, a Isabel por su simpatía, a Juan por las conversaciones sobre la novedad científica. Pablo, gracias por ser, directa o indirectamente, “el padre científico” de todos nosotros.

Mamá, no existe la palabra para poder expresar mi agradecimiento y mi amor por ti, te nos fuiste muy pronto, sólo puedo decir gracias, por darme la vida, por haberme apoyado y protegido siempre, por haberme enseñado a luchar hasta en los momentos que no tienes fuerzas, eres mi ejemplo, mi campeona, estarás conmigo, donde vaya y para siempre. Papá, qué decir, gracias por tu apoyo, a veces silencioso, pero incondicional y auténtico, eres el cimiento que siempre está ahí, sosteniendo sin estruendo, pero el más fundamental.

A mis *yayos* Jesús y Lorenzo, aunque ya no estén, y a mis *yayas* Teresa y Carmina, tengo que agradecer el apoyo que me habéis dado, a pesar de ser el nieto que estudia mucho pero no sabemos exactamente qué, ese apoyo brota de un amor incondicional que sólo lo pueden dar los abuelos. A mis tías, las *tatas* Almu, Palo, Chiqui y MariCruz, por estar ahí cuando os he necesitado. A Fernando del Campo, *Fernandito*, gracias por tu humor, compañía, conversación y apoyo a mí y a mi padre en los momentos más difíciles.

Ahora quiero avanzar, de la familia que se tiene a la que se escoge. ¡Espero no dejarme a nadie! A los Roceños, Miguel, Bea y Marco, casi media vida juntos, ¡que decir!, a los Santoñeses, Nacho y Elena, al Canario-Berlinés-Chipriota- Filo y a mi filósofo favorito, Iván, gracias a todos por soportarme (sobre todo, cuando me sale el profe en el momento más inesperado) y por todas las aventuras que hemos tenido por el mundo, ¡y las que quedan!

A mis parroquianos del CSIC (sede La Hiedra) Felipón, Cumella, Paco, Laura, Carol, Patrick, Quique, las nuevas generaciones, Antonio, Naroa y a todos los que pasan por allí de vez en cuando. A ese concilio de científicos, de reunión espontánea, gracias por la terapia de grupo, por vuestros consejos cuando estoy en crisis, por las conversaciones esperpénticas y *bizarrras*, en definitiva, por las risas.

No puedo olvidarme de la gran aventura del doctorado, la estancia en Finlandia. En primer lugar, a quién lo hizo posible, Kristiina (Lilli) Takkinen, gracias por acogerme en tu gran grupo de investigación, por tu apoyo y simpatía, por todo lo que aprendí en nuestras conversaciones, allí recuperé la pasión por la ciencia después de una fuerte crisis existencial. Al equipo del VTT, Tarja, Kristiina, Sirpa, Heidi y el resto, gracias a vuestro buen hacer esta tesis ha llegado a buen puerto. A mis compatriotas Rafa, Alba, Jorge, a los *pakolainen* del encuentro Erasmus 1922, a los *Suomalainen* y a todos con los que coincidí, *Kiitos paljon* por hacer mi estancia en un gran país, una experiencia inolvidable.

Por último, pero no menos importante, tengo que acordarme del que pasa más tiempo conmigo (literalmente pegado a mí), la última incorporación a la familia, alegría y fidelidad en formato peludo y de cuatro patas, Puchi, gracias por darme una vida mucho más alegre.



# PRESENTACIÓN DE LA TESIS DOCTORAL

La presente tesis doctoral titulada “Desarrollo de métodos para la detección de gluten en alimentos, a partir de repertorios de anticuerpos recombinantes derivados de la respuesta inmune celíaca” ha sido realizada en el Departamento de Nutrición y Ciencia de los Alimentos de la Facultad de Veterinaria de la Universidad Complutense de Madrid, y se presenta en formato de compendio de artículos científicos.

El objetivo principal de este trabajo de investigación ha consistido en el desarrollo de nuevos sistemas de inmunodetección de gluten en alimentos basados en anticuerpos recombinantes. Para la obtención de estas sondas de detección de gluten, se ha optado por el empleo de metodologías de evolución dirigida, concretamente la tecnología de presentación sobre fagos filamentosos o *phage display*. Con este fin, se describe la construcción de dos repertorios de fago-anticuerpos tipo Fab generados, completa o parcialmente, a partir de linfocitos de sangre periférica de pacientes celíacos. Se detallan, asimismo, los resultados obtenidos de la selección por afinidad de variantes de Fab con capacidad de unión a gliadina empleados para el desarrollo de diferentes formatos de inmunoensayos para la detección de gluten en alimentos, basados tanto en fago-anticuerpos como en los anticuerpos recombinantes obtenidos tras su producción soluble. Por último, se presenta la caracterización de los repertorios mediante tecnologías de secuenciación masiva, y aplicación de técnicas de modelización *in silico* para estudiar la interacción de los Fab obtenidos con el antígeno.

Esta Tesis Doctoral está constituida por tres artículos de investigación que han sido publicados en revistas recogidas dentro del campo de *Food Science and Technology* del *Journal Citation Reports* (JCR):

- 1. Garcia-Calvo Eduardo,** García-García Aina, Rodríguez Santiago, Farrais Sergio, Martín Rosario, García Teresa (2022). Construction of a Fab library merging chains from semisynthetic and immune origin, suitable for developing new tools for gluten immunodetection in food. *Foods*, 12(1), 149. <https://doi.org/10.3390/foods12010149>

2. **García-Calvo Eduardo**, García-García Aina, Rodríguez-Gómez Santiago, Farras Sergio, Martín Rosario, García Teresa (2023). Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods. *Current Research in Food Science*, 7, 100578. <https://doi.org/10.1016/j.crfs.2023.100578>
3. **García-Calvo, Eduardo**, García-García Aina, Rodríguez Santiago, Takkinen Kristiina, Martín Rosario, García Teresa. (2023). Production and Characterization of Novel Fabs Generated from Different Phage Display Libraries as Probes for Immunoassays for Gluten Detection in Food. *Foods*, 12(17). <https://doi.org/10.3390/foods12173274>

Por otra parte, se incluyen también dentro de la Tesis Doctoral, dos trabajos de investigación que serán enviados para su publicación en revistas científicas especializadas, así como un artículo de revisión ya publicado que se recoge dentro del anexo.

- **García-Calvo Eduardo**, García-García Aina, Rodríguez Santiago, Martín Rosario, García Teresa (2023). Exploring gluten assessment in marketed products through a sandwich-ELISA methodology based on novel recombinant antibodies. Comunicación breve en preparación.
- **García-Calvo Eduardo**, García-García Aina, Rodríguez Santiago, Martín Rosario, García Teresa (2023). Unveiling the properties of phage display Fab libraries and their use in gliadin-specific probes selection by high-throughput sequencing techniques. Artículo en preparación.
- **Anexo I: García-Calvo Eduardo**, García-García Aina, Madrid Raquel, Martín Rosario, García Teresa (2020). From Polyclonal Sera to Recombinant Antibodies: A Review of Immunological Detection of Gluten in Foodstuff. *Foods*, 10(1), 66. <https://doi.org/10.3390/foods10010066>

# ÍNDICE GENERAL

<b>RESUMEN</b> .....	<b>1</b>
<b>SUMMARY</b> .....	<b>5</b>
<b>EXTENDED SUMMARY</b> .....	<b>9</b>
<b>1. INTRODUCCIÓN</b> .....	<b>27</b>
<b>1.1 Características del gluten</b> .....	<b>29</b>
<b>1.2 Patologías relacionadas con el gluten</b> .....	<b>30</b>
<b>1.2.1</b> Patologías de etiología alérgica .....	30
<b>1.2.2</b> Patologías de etiología autoinmune .....	32
<b>1.2.3</b> Otras patologías.....	33
<b>1.2.4</b> Tratamiento farmacológico.....	33
<b>1.2.5</b> Tratamiento dietético .....	34
<b>1.3 Marco legislativo internacional sobre el contenido en gluten de los alimentos</b> .....	<b>35</b>
<b>1.4 Metodologías analíticas para la detección de gluten en alimentos</b> .....	<b>38</b>
<b>1.5 Métodos para el desarrollo de anticuerpos frente al gluten</b> .....	<b>39</b>
<b>1.5.1</b> Estructura y generación de anticuerpos .....	39
<b>1.5.2</b> Anticuerpos policlonales .....	42
<b>1.5.3</b> Anticuerpos monoclonales .....	42
<b>1.5.4</b> Anticuerpos recombinantes y evolución dirigida .....	48
<b>1.5.5</b> Producción biotecnológica de anticuerpos recombinantes.....	58
<b>1.6 Caracterización de repertorios de anticuerpos mediante técnicas de secuenciación masiva</b> .....	<b>61</b>
<b>1.6.1</b> Tecnologías de secuenciación masiva.....	61
<b>1.6.2</b> Análisis masivo de anticuerpos.....	63
<b>1.7 Elucidación <i>in silico</i> de la interacción antígeno y anticuerpo</b> .....	<b>69</b>
<b>1.7.1</b> Predicción de estructuras de proteínas y métodos clásicos utilizados .....	69
<b>1.7.2</b> La revolución de la inteligencia artificial.....	71
<b>1.7.3</b> Aplicaciones de la inteligencia artificial a la resolución de la estructura de los anticuerpos .....	75
<b>1.7.4</b> Estudio de la interacción antígeno-anticuerpo mediante la tecnología de <i>docking</i> .....	77
<b>2. JUSTIFICACIÓN, HIPÓTESIS Y OBJETIVOS</b> .....	<b>81</b>

<b>3. METODOLOGÍA RESUMIDA .....</b>	<b>87</b>
<b>3.1 Preparación de muestras y antígenos .....</b>	<b>89</b>
<b>3.2 Selección de pacientes con niveles elevados de biomarcadores de la respuesta humoral celíaca .....</b>	<b>90</b>
<b>3.3 Extracción de ARN y retrotranscripción del material biológico empleado para la construcción de repertorios de fragmentos de anticuerpo tipo Fab.....</b>	<b>90</b>
<b>3.4 Amplificación de las secuencias codificantes de cadenas pesadas y ligeras de anticuerpos .....</b>	<b>91</b>
<b>3.5 Construcción de los repertorios de anticuerpos recombinantes y selección por afinidad de variantes con capacidad de unión a gliadina.....</b>	<b>92</b>
<b>3.6 Seguimiento del proceso de panning. Selección y secuenciación de clones individuales.....</b>	<b>93</b>
<b>3.7 Evaluación de los diferentes candidatos de fago-Fab como sondas para la detección de gluten en alimentos .....</b>	<b>94</b>
<b>3.8 Producción soluble de Fab recombinantes frente al gluten .....</b>	<b>94</b>
<b>3.9 Evaluación de los Fab recombinantes como sondas para la detección de gluten .....</b>	<b>95</b>
<b>3.10 Desarrollo de un ELISA sándwich basado en los Fab recombinantes .....</b>	<b>96</b>
<b>3.11 Análisis de productos comerciales mediante los ELISA desarrollados con los Fab solubles.....</b>	<b>97</b>
<b>3.12 Caracterización de los repertorios obtenidos y estudio de los procesos de selección por afinidad mediante secuenciación masiva.....</b>	<b>97</b>
<b>3.13 Modelización <i>in silico</i> de la estructura de los Fab8E-4 y Fab-C y estudio de su interacción con la gliadina .....</b>	<b>98</b>
<b>4. RESULTADOS .....</b>	<b>99</b>
<b>Capítulo 4.1</b>	
Construcción de un repertorio de Fab mediante la combinación de cadenas de anticuerpos de origen semisintético e inmune, apropiada para el desarrollo de nuevas herramientas de inmunodetección de gluten en alimentos / Construction of a Fab library Merging Chains from Semisynthetic and Immune Origin, Suitable for Developing New Tools for Gluten Immunodetection in Food .....	101
<b>Capítulo 4.2</b>	
Desarrollo de un nuevo anticuerpo recombinante, seleccionado mediante la tecnología del phage display a partir de un repertorio derivado de pacientes celíacos, para la detección de gluten en alimentos / Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods.....	123

**Capítulo 4.3**

Producción y caracterización de nuevos Fab generados a partir de diferentes repertorios para phage display como sondas en inmunoensayos para detección de gluten en alimentos / Production and characterization of novel Fabs generated from different phage display libraries as probes for immunoassays for gluten detection in food ..... 141

**Capítulo 4.4**

Evaluación de la cuantificación de gluten en productos comerciales mediante una metodología de ELISA tipo sándwich basada en el empleo de nuevos anticuerpos recombinantes / Exploring gluten quantification in marketed products through a sandwich-ELISA methodology based on novel recombinant antibodies ..... 159

**Capítulo 4.5**

Caracterización de las propiedades de repertorios de Fab y de su uso para la selección de sondas con alta afinidad por la gliadina mediante técnicas de secuenciación masiva / Unveiling the properties of phage display Fab libraries and their use in gliadin-specific probes selection by high-throughput sequencing techniques ..... 175

**5. DISCUSIÓN INTEGRADORA..... 203****5.1 Obtención y preparación de muestras biológicas como material de partida para el desarrollo de repertorios de anticuerpos recombinantes..... 207**

**5.1.1** Proceso de selección de pacientes para su inclusión en el estudio..... 207

**5.1.2** Aislamiento de linfocitos B de sangre periférica para la extracción y caracterización de su RNA ..... 208

**5.2 Construcción de un repertorio semiinmune de Fab, mediante la combinación de cadenas de anticuerpos de origen semisintético e inmune, y desarrollo de un método ELISA para la detección de gluten, utilizando los fago-Fab seleccionados por *phage display*..... 209**

**5.2.1** Construcción del subrepertorio de cadenas pesadas semisintéticas ..... 209

**5.2.2** Construcción del subrepertorio de cadenas ligeras inmunes..... 210

**5.2.3** Clonación del repertorio de Fab y selección por afinidad frente a la gliadina (*panning*) ..... 210

**5.2.4** Selección de clones de fago-Fab individuales mediante ELISA monoclonal ..... 211

**5.2.5** Secuenciación y análisis de los clones seleccionados..... 212

**5.2.6** Caracterización del fago-Fab8E-4 como sonda para detección de gluten en alimentos mediante una técnica ELISA indirecto..... 213

**5.3 Construcción de un repertorio inmune de Fab y desarrollo de un método ELISA para la detección de gluten, utilizando los fago- Fab seleccionados por *phage display*..... 214**

**5.3.1** Construcción del repertorio inmune ..... 214

5.3.2	Selección por afinidad ( <i>panning</i> ) del repertorio inmune frente a la gliadina....	216
5.3.3	Secuenciación y análisis de los clones seleccionados.....	216
5.3.4	Caracterización de los fago-Fab seleccionados con mejores características como sondas para la detección de gluten .....	218
<b>5.4</b>	<b>Producción soluble de los Fab seleccionados y desarrollo de técnicas inmunoenzimáticas ELISA para detección de gluten en alimentos.....</b>	<b>220</b>
5.4.1	Producción recombinante a escala semiindustrial de los Fab seleccionados....	220
5.4.2	Validación de los Fab recombinantes como sondas para la detección de gluten.....	221
5.4.3	Desarrollo de una metodología de ELISA indirecto con Fab soluble para detección de gluten en alimentos.....	222
5.4.4	Desarrollo de una metodología de ELISA sándwich con Fab recombinantes para detección de gluten en alimentos .....	223
5.4.5	Comparación entre las dos metodologías desarrolladas.....	224
5.4.6	Análisis de muestras de alimentos con los sistemas de ELISA desarrollados. ...	225
<b>5.5</b>	<b>Modelización <i>in silico</i> de la estructura e interacción de algunos de los Fab seleccionados y su interacción con la gliadina .....</b>	<b>227</b>
5.5.1	Modelización de la estructura e interacción del Fab8E-4.....	228
5.5.2	Modelización de la estructura e interacción del Fab-C.....	230
<b>5.6</b>	<b>Caracterización de los repertorios y procesos de selección de Fab mediante técnicas de secuenciación masiva .....</b>	<b>232</b>
5.6.1	Análisis primario.....	232
5.6.2	Análisis secundario e interpretación biológica de los resultados .....	233
<b>6.</b>	<b>CONCLUSIONES.....</b>	<b>239</b>
<b>7.</b>	<b>TRABAJO FUTURO.....</b>	<b>245</b>
<b>8.</b>	<b>BIBLIOGRAFÍA .....</b>	<b>251</b>
<b>9.</b>	<b>ANEXOS .....</b>	<b>281</b>
	De los sueros policlonales a los anticuerpos recombinantes: una revisión sobre la detección inmunológica de gluten en alimentos / From polyclonal sera to recombinant antibodies: a review of immunological detection of gluten in foodstuff.....	283

# ÍNDICE DE FIGURAS

## INTRODUCCIÓN

- Figura 1.** Clasificación de las diferentes patologías relacionadas con el gluten según su etiología. \*WDEIA: Wheat-Dependent Exercise-Induced Anaphylaxis, anafilaxia inducida por el ejercicio dependiente de trigo. Fuente: elaboración propia..... 31
- Figura 2.** Representación esquemática de la estructura de anticuerpos y moléculas derivadas. Ig (Inmunoglobulinas); Fab (fragmento de unión a antígeno); Fc (fragmento cristalizante); VH (región variable de la cadena pesada); VL (región variable de la cadena ligera); CH (región constante de la cadena pesada); CL (región constante de la cadena ligera); Dab (anticuerpo de dominio único); scFv (fragmento de anticuerpo variable de cadena única); VHH (fragmento correspondiente al dominio variable de la cadena pesada de anticuerpo de camélido). Fuente: elaboración propia. .... 40
- Figura 3.** Representación esquemática de la producción de anticuerpos monoclonales. Fuente: elaboración propia..... 44
- Figura 4.** Construcción de un repertorio inmune y desarrollo del proceso de evolución molecular dirigida mediante la tecnología de phage display. Fuente: elaboración propia. .... 52
- Figura 5.** Representación esquemática de los métodos de evolución dirigida de anticuerpos. Fuente: elaboración propia..... 54
- Figura 6.** Representación de los dominios estructurales de la región VH del trastuzumab. A la izquierda aparece una representación plana de la estructura en forma de “collar de perlas”. Los HCDR1, HCDR2 y HCDR3 (correspondientes a los bucles BC, C’C” y FG, respectivamente) se representan en rojo, naranja y morado. Los círculos rayados corresponden a las posiciones que faltan con respecto a la secuencia de referencia en la base de datos IMGT. Las flechas indican la dirección de las láminas beta y sus designaciones en las estructuras 3D. Los puntos de anclaje se muestran en cuadrados. Los puentes de hidrógeno (líneas verdes) fueron añadidos a partir de los datos estructurales experimentales. El puente disulfuro (línea naranja) también fue incorporado a partir de los datos estructurales experimentales. A la derecha se muestra la representación tridimensional de la misma estructura con los CDRs coloreados con la misma gama de colores que en la figura a la izquierda. Figura adaptada de Lefranc & Lefranc (2020). .. 67
- Figura 7.** Representación del concepto de clonotipo en anticuerpos. Los círculos representan varios linfocitos B en diferentes etapas en el desarrollo y la maduración de su BCR. Los bloques representan los genes V(D)J y las barras negras, las mutaciones que ocurren tras la reorganización. Fuente: elaboración propia. .... 68
- Figura 8.** Ejemplo gráfico del uso de mutaciones correlacionadas para descubrir contactos entre aminoácidos como primer paso para la predicción de la estructura tridimensional de una proteína. Fuente: Marks et al. (2011). .... 70
- Figura 9.** Representación del funcionamiento de AlphaFold2. Figura modificada de <https://www.deepmind.com/blog> que actualmente no se encuentra accesible. .... 72
- Figura 10.** Representación esquemática del concepto triángulos de residuos. Fuente: elaboración propia..... 74

## METODOLOGÍA RESUMIDA

**Figura 11.** Representación esquemática de las técnicas de ELISA desarrolladas en este trabajo para la detección de gluten en alimentos. (A) ELISA indirecto basado en el uso de fago-Fab. (B) ELISA indirecto basado en el uso de Fab soluble. (C) ELISA sándwich directo basado en el uso de Fab biotinilados. (D) ELISA sándwich indirecto basado en el uso de Fab solubles. Abreviaturas empleadas: Ag (antígeno, gluten); Fab (fragmento de unión a antígeno); Anti-pVIII (anticuerpo secundario dirigido frente a la proteína VIII del bacteriófago VSC-M13); Fab-BT (Fab biotinilado); SA (estreptavidina); AP (fosfatasa alcalina); Anti-Fab (anticuerpo secundario dirigido frente a Fab de origen humano). .... 96

## RESULTADOS

### CAPÍTULO 4.1

**Figure 1.** Agarose gel containing coding fragments for antibody chains to be cloned for library construction. (A) Different constructs that built the semi-synthetic heavy chains. Ladder: 1 kb plus Invitrogen™ (#ref 10787018); lane A: dAb-8E amplified sequence from the first round of PCR (VH-392 bp); lane B: CH1 amplified sequence from the first round of PCR (CH-300 bp); lane C: construct with the heavy chains obtained from the second round of PCR (SOE-PCR) prepared for Fab library cloning (H-680 bp). (B) Electrophoretic analysis of the human kappa chain genes amplified by PCR (Ladder: 1 kb plus Invitrogen®)..... 109

**Figure 2.** Monoclonal indirect phage-ELISA of Fabs isolated from the panning process. (A) Indirect phage-ELISA analysis of individual induced clones isolated from the last round of affinity enrichment.(B) Comparison of the dose-response curves by means of the indirect phage-ELISA of the four selected binding clones performed with gliadin-PWG dilutions from 2.5 to 75 µg/mL. Mean values of three independent determinations and standard derivation of each data set are shown. (C) Indirect phage-ELISA of positive clones against ethanol protein extracts of different cereal flours (1:100). ..... 111

**Figure 3.** Sequencing results, chain families and features identification of the gliadin-binding selected clones. (A) Alignment of the sequences coding for the Fab8E-4, 6, 10 and 11. Main antibody features are shown in blue and orange charts for light and heavy chains, respectively. Alignment was computed using SnapGene® 5.0.8. (B) Identification of the re-arranged genes coding the light and heavy chains of the isolated gliadin-binding clones. The VJ and VDJ genes were elucidated using the IMGT®/V-QUEST tool, and the IgBLAST tool from NCBI was used for elucidating the CH1 subclass of the heavy chains. 112

**Figure 4.** Characterization of phage-Fab8E-4. (A) Comparative dose-response curve obtained against gliadin-PWG dilutions (0–6 µg/mL) by means of dAb-8E (grey) and Fab8E-4 (red) antibody fragments attached to VSCM13 helper phage ( $4 \times 10^{11}$  pfu/well). (B) Linear regression obtained against gliadin dilutions (0.01–2 µg/mL) by means of Fab8E-4 in order to determinate the LOD and LOQ. (C) Representative dose-response curve from the analysis of experimental rice-based binary mixtures with increasing presence of a proportional mixture of wheat, barley and rye in the range of 0.1 to 100 mg/g. Origin 8.0 software was employed to plot and analyze the experimental data. Mean values of three independent determinations and standard derivation of each data set are shown. (D)Western blot analysis of gliadin and flour proteins incubated with phage-Fab8E-4 as primary antibody. Lane 1 (1 mg per well) and Lane 2 (2 mg per well): gliadin (Merck®, CAS #9007–90-3) dissolved in 50% isopropanol-water; Isopropanol extractions of cereal flours: Lane 3: wheat; Lane 4: barley; Lane 5: rye. .... 113

- Figure 5.** Computational models of the Fab8E-4. (A) Predictive model of Fab8E-4 generated with Alpha-Fold2. The pLDDT scores are shown with the following scale of color red > white > blue. (B) Same Alpha-Fold2 predicted model rotated for better visualization of the different domains. (C) Fv predictive model generated by Antibody-builder (sAbpred server)..... 115
- Figure 6.** Epitope prediction (A,B) in  $\alpha/\beta$ -gliadin (A) and  $\gamma$ -hordein (B) facing Fab8E-4. The residues colored in red are more likely to be part of the conformational epitope. Notice that the images have been rotated to show the conformational epitope. Models generated with EpiPred (SAbPred server). Paratope prediction in Fab8E-4 (C,D). The residues are colored according to the likelihood of being part of the paratope following the order: red > white > blue. Notice that the images have been rotated to show the zones with a higher likelihood of being part of the paratope. The model was generated with antibody i-patch (SAbPred server). (C) In red the residues of the heavy chain identified to be more likely to interact with the antigen. CDR2 and neighboring sequences in a green silhouette. (D) In red the residues of the light chain identified to be more likely to interact with the antigen. .... 116
- Figure S1.** Transformants recovery after panning rounds. An inference was calculated from the colonies grown in after plating in SB-carbenicilin (100  $\mu\text{g}/\text{mL}$ ) infected E.coli XL-1 with the output phages from each selection round. Mean values of two independent determinations and standard derivation of each data set are shown. .... 121
- Figure S2.** Indirect phage-ELISA analysis of the phage-Fab library rescued after each round of panning. Results of absorbance value (450 nm) obtained against gliadin-PWG (objective antigen) and BSA (blocking solution). Mean values of three independent determinations and standard derivation of each data set are shown. .... 122
- ## CAPÍTULO 4.2
- Fig. 1.** Characterization of the panning progress for selection of gliadin-binding phage-Fabs. (A) Phage titration after each round of panning. (B) Polyclonal indirect phage-ELISA analysis of the phage-Fab li-brary isolated after each round of panning. Notice that round 4 is referring to the amplification of output phages from round 3. (C) Monoclonal indirect phage- ELISA analysis of individual clones isolated from the third and fourth rounds of affinity enrichment against gliadin-PWG. Origin 8.0 software was used to plot and analyze the experimental data. Mean values of three independent determinations and standard derivation of each data set are shown. .... 130
- Fig. 2.** Main features of the sequencing and classifi-cation of the selected Fabs. (A) Phylogenetic tree of isolated Fabs. An alignment of sequencing results was conducted with SnapGene® software, computing MUSCLE algorithm and then the phylogenetic tree was represented using the Simple Phylogeny tool (ClustalW2 package, EMBL-EBI). (B-H) Antibody se-quences were analyzed, and clustering was calcu-lated, using IgBLAST tool (NCBI) and Origin 8.0 software for data representation. The pie charts represent the frequency of the genes, and the depicted number refers for the clonotypes that presented each gene: (B) VH genes identification and frequency. (C) JH genes identification and frequency. (D) D genes identification and frequency. (E) Heavy chain CDR3 length distribution per clonotype. (F) VK genes identification and frequency. (G) JK genes identifi-cation and frequency. (H) Light chain CDR3 length distribution per clonotype. .... 131

- Fig. 3.** Characterization and selection of gliadin-binding Fabs by phage-ELISA. (A) Comparative dose-response curve of the fourteen leading candidates against gliadin-PWG dilutions (0–5  $\mu\text{g}/\text{mL}$ ). (B) Dose-response curve of the four strongest candidates (Fab-C, Fab-E, Fab-H and Fab-G) using an experimental rice-based binary mixture with increasing presence of a proportional mixture of wheat, barley and rye in the range of 0.1 to 100 mg/g. (C) Calibration curve obtained for gluten detection with the experimental mixture and the Fab-C in an indirect phage-ELISA. Origin 8.0 software was used to plot and analyze the experimental data. Mean values of three independent determinations and standard derivation of each data set are shown. .... 132
- Fig. 4.** Computational model of the structure and interactions of Fab-C. (A) Fv structure prediction using two modeling programs: IgFold (blue) and SAbPred (brown). Comparison by superimposing the models obtained using the Matchmaker tool from ChimeraX. HCDR3 is flagged with a black arrow. (B) Representation of the epitope prediction of  $\alpha/\beta$ -gliadin. The residues colored in red are more likely to be part of the conformational epitope. Representation of the main likely interaction points between Fab-C and depicted  $\alpha/\beta$ -gliadin (C to F). The colored residues are part of the CDRs (yellow and flagged with a black arrow) or the antigen epitopes (red) in the docking model: (C) HCDR1; (D) HCDR2; (E) HCDR3; (F) LCDR3. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.) ..... 133
- Figure 1S.** Representation of the normalized indexes depicting the humoral response against native gliadin per IgG isotype. .... 138
- Figure 2S.** Agarose gel containing coding fragments for antibody chains to be cloned for library construction. Ladder: 1 kb plus Invitrogen™ (#ref 10787018). (A) Amplicons coding for VH1, 3, and 5 gene families for all IgGs. Lane 1: IgG1. Lane 2: IgG2. Lane 3: IgG3. Lane 4: IgG4. (B) Amplicons coding for IgA VH genes. Lane 1: VH135. Lane 2: VH2. Lane 3: VH4. Lane 4: VH4b. (C) Amplicons coding for kappa light chains. Lanes 1 and 2: Vk1. Lanes 3 and 4: Vk2/4. Lanes 5 and 6: Vk3. Lanes 7 and 8: Vk5. Lane 9: Vk pool. (D) Amplicons coding for lambda light chains: Lane 1: VL1a. Lane 2: VL1b. Lane 3: VL2. Lane 4: VL3/9. Lane 5: VL7/8. (E) Comparison between first round (Lane 1) and second round (Lane 2) PCR amplicons. The second round amplicon is slightly bigger (39 bps) than the first one, due to the expansion of the neighboring sequences to the restriction points. . 139

### CAPÍTULO 4.3

- Figure 1.** Schematic workflow of the production and characterization of novel Fab as probes in ELISA methods for gluten detection in foodstuff. .... 145
- Figure 2.** Western blot analysis of transformed RV308 culture supernatants after induction for Figure 2. E. coli ; expression of recombinant Fabs. Lane 1: control Fab previously produced; Lanes 2 and 3: Fab-C Lanes 4 and 5: Fab-E; Lanes 6 and 7: Fab-H; Lanes 8 and 9: Fab8E-4. .... 150
- Figure 3.** Recombinant production, purification, and identification of Fabs. (A) Example of the immobilized metal affinity chromatography (IMAC) profile for the purification of the recombinant Fabs (Fab-C in this image). In green, monitorization of the protein quantity (absorbance 280 nm) going through the column. In red, conductivity measures. In black, percentage of elution buffer pumped to the system. In blue, measure of the system pressure. (B) Analysis of the purification steps by SDS-PAGE electrophoresis.

Lane 1: control Fab previously produced. Lanes 2–4: samples from different fractions representing the first peak of the chromatogram (\*). Lanes 5–9: samples from different fractions representing the second peak of the chromatogram (\*\*) containing the desired Fab. .... 151

**Figure 4.** Indirect ELISA results for characterization of the four recombinant Fabs produced. (A) Specificity results against ethanolic extracts of gluten-containing (wheat, barley, and rye) and gluten-free (oats, corn, and rye) flours (diluted 1:5 in PBS). (B) Sensitivity evaluation by comparative dose-response curves obtained against gliadin-PWG (0–20 µg/mL) diluted in PBS. (C) Comparison of dose–response curves for detection of gluten in an experimental mixture (rice flour spiked with growing concentrations of wheat, rye, and barley flours). (D) Linear regression results obtained from the analysis of the gliadin-PWG standard (0.025–0.625 µg/mL) by indirect ELISA using Fab-C as primary antibody..... 153

**CAPÍTULO 4.4**

**Figure 1.** Dose response curve obtained from the analysis of the gliadin-PWG standard (0.025-25 µg/mL) by sandwich ELISA using Fab-C as detection antibody and Fab8E-4 as capture antibody. This model was used to estimate the gluten content of the tested commercial food products. An HRP-conjugated anti- human Fab was used as the secondary antibody. Mean values of three independent determinations and standard deviation of each data set are shown..... 167

**CAPÍTULO 4.5**

**Figure 1.** Schematic representation of main workflow used for the primary analysis of the sequences read with the MinION device..... 182

**Figure 2.** Quality metrics for Oxford Nanopore sequencing of the home- made phage-displayed libraries. (A) Distribution of quality scores according to read position, in the boxplot in black lines represent medians, the box shows the lower and upper boundaries of the first and third quartiles, respectively, and the line shows the mean quality score. (B) Mean per-read quality score diagram. (C) Per base sequence content plot. (D) Read length distribution. All the analysis were performed using FastQC toolkit. The data shown correspond to one of the multiplexed samples, however the data recorded for all the samples presented similar profiles. .... 185

**Figure 3.** Pairwise analysis of VH genes composition of the unpanned immune library (red) and its fourth round of selection (blue). Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each VH subfamily. The color of the intervals depicts test interpretation: non-significant (red), significant for all p-values (blue), significant for at least 2 tests (pink) and significant for only Benjamini-Hochberg test (green). .... 187

**Figure 4.** Pairwise analysis of VH genes composition of round 1 (red) and round 2 (blue) of selection of the immune library. Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each VH subfamily. The color of the intervals depicts test interpretation: non-significant (red), significant for all p-values (dark blue), significant for one test (light blue). .... 188

- Figure 5.** Clustered imaged maps depicting the relation between immunoglobulins VDJ genes of the immune library. (A) VD genes; (B) DJ genes and (C) VJ genes. Darker colors in heatmaps indicate the strongest associations between genes..... 190
- Figure 6.** Pairwise analysis of  $V_k$  genes composition of the unpanned immune library (red) and its fourth round of selection (blue). Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each VH subfamily. The color of the intervals depicts test interpretation: non-significant (red), significant for all  $p$ -values (dark blue), significant for one test (light blue), and significant only Benjamini-Hochberg test (green). ..... 193
- Figure 7.** Clustered imaged maps depicting the relation between VJ ( $\kappa$  chains) genes of the immune library. Darker colors in heatmap indicate the strongest associations between genes. .... 194
- Figure 8.** Pairwise analysis of  $V_k$  genes composition of the unpanned semi-immune library (red) and its third round of selection (blue). Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each VH subfamily. The color of the intervals depicts test interpretation: non-significant (red). .. 195
- Figure 9.** Antibody subclass composition of the unpanned immune library (R0, orange) and its fourth round of selection (R4, green). Data were extracted from the alignment of the sequences of each round against the human genome, followed by counting sequences with the coverage function of the bedtools package..... 196
- Figure 10.** Antibody subclass composition of the unpanned semi-immune (R0, orange) and its third round of selection (R3, green). Data were extracted from the alignment of the sequences of each round against the human genome, followed by counting sequences with the coverage function of the bedtools package..... 197

# ÍNDICE DE TABLAS

## INTRODUCCIÓN

<b>Tabla 1.</b> Fracciones proteicas de cereales. ....	29
<b>Tabla 2.</b> Anticuerpos policlonales comerciales obtenidos frente al gluten. ....	43
<b>Tabla 3.</b> Patentes registradas basadas en anticuerpos monoclonales para la detección de gluten. ....	49
<b>Tabla 4.</b> Ejemplos de la aplicación de la tecnología de phage display a la investigación de la enfermedad celíaca. ....	55
<b>Tabla 5.</b> Aplicación de métodos de evolución dirigida de anticuerpos para la detección de gluten en alimentos. ....	57
<b>Tabla 6.</b> Clasificación de los 10 tipos de secuencias integradas en la base datos IMGT. ....	65

## RESULTADOS

### CAPÍTULO 4.1

<b>Table S1.</b> Primers used for heavy chain amplification .....	120
<b>Table S2.</b> Primers used for light chain amplification .....	120
<b>Table S3.</b> Sequencing primers.....	120
<b>Table S4.</b> ELISA results obtained for detection of gluten with the phage Fab8E-4 and R5 monoclonal antibodies from ethanolic extracts of 10 commercial food products.....	121

### CAPÍTULO 4.2

<b>Table 1.</b> Analysis of the humoral response of potential celiac patients to be included in this study. ....	129
<b>Table 2.</b> Results obtained for detection of gluten in ten commercial food products using the indirect phage-ELISA with Fab-C, and the sandwich ELISA with R5 mono-clonal antibody as reference method. ....	132
<b>Table 1S.</b> Primers for antibody chain amplification (Barbas, 2001). ....	136

**CAPÍTULO 4.3**

**Table 1.** List of heterologous species (common name and scientific name) that did not show cross-reactivity when tested in the indirect ELISA with Fab-C. .... 148

**Table 2.** Results obtained for the detection of gluten in commercial food products using the developed Fab-C-based indirect ELISA and the sandwich monoclonal R5 antibody for result confirmation. .... 155

**CAPÍTULO 4.4**

**Table 1.** Checkerboard compatibility test between pairs of recombinant Fab in a direct sandwich ELISA. The values represent average absorbance data (405 nm) in three independent assays. .... 165

**Table 2.** Results obtained for the detection of gluten in commercial food products using the developed indirect sandwich ELISA and the sandwich monoclonal R5 antibody for result confirmation. .... 170

**CAPÍTULO 4.5**

**Table 1.** Identification of mismatch mutations heavily repeated on the selected Fab variants from the immune library after the biopanning process against gliadin. .... 191

# RESUMEN

El gluten es la fracción proteica de granos de cereales como el trigo, cebada, centeno, espelta y kamut, que es insoluble en agua y soluble en disoluciones hidroalcohólicas. Su uso está muy extendido en la industria alimentaria, debido a que proporciona características funcionales como extensibilidad a las masas o consistencia a las salsas. Además, al ser termorresistente, se puede utilizar como ingrediente en productos que se someten a tratamiento térmico. El gluten no sólo se usa como ingrediente en alimentos, sino también como excipiente de medicamentos y cosméticos. Sin embargo, contiene péptidos que pueden desencadenar respuestas patológicas en personas sensibles, como alergias, enfermedades autoinmunes como la celiaquía o intolerancia al gluten no celíaca.

A pesar de que se han desarrollado muchas terapias farmacológicas para tratar las patologías relacionadas con el gluten, ninguna ha demostrado mayor efectividad que la adhesión estricta a una dieta sin gluten. Para ello, es necesario tomar medidas de prevención, control y trazabilidad en la industria, que garanticen que los productos comercializados como “sin gluten” contienen menos de 20 mg/kg de este compuesto, en cumplimiento la legislación vigente.

Las técnicas más utilizadas para la detección de gluten en alimentos son las inmunoquímicas, basadas en la interacción específica entre un anticuerpo y su correspondiente antígeno. Actualmente, el método de referencia avalado por el *Codex Alimentarius* para la detección de gluten en alimentos es una técnica inmunoenzimática de ELISA sándwich basada en el anticuerpo monoclonal R5. Sin embargo, la identificación y cuantificación precisa del gluten sigue siendo un gran desafío para la ciencia de los alimentos, debido a que el gluten no es una proteína bien definida, como ocurre con otros alérgenos alimentarios, sino que se define en términos de características químicas que abarcan una amplia gama de proteínas con ciertas similitudes, pero cuyas secuencias pueden variar. Por ello, diversos estudios han revelado notables discrepancias en la cuantificación de gluten cuando se comparan los resultados obtenidos con los kits comerciales disponibles en el mercado. En consecuencia, sigue siendo necesaria la investigación y el desarrollo de métodos de detección de gluten en alimentos, ya sea para abordar sus limitaciones o para complementar los métodos existentes.

La mayoría de los anticuerpos para la detección de gluten desarrollados hasta el momento se han obtenido a partir de hibridomas, que requieren de la inmunización de animales de experimentación. Con esta metodología se han obtenido anticuerpos de elevada afinidad frente al antígeno empleado en la inmunización, pero que no reproducen la respuesta inmune de las personas que padecen una enfermedad desencadenada por el gluten.

El objetivo principal de esta tesis doctoral ha consistido en el desarrollo de nuevas metodologías analíticas que permitan detectar en los alimentos componentes del gluten potencialmente peligrosos para una persona sensibilizada. Para ello, se han desarrollado nuevos anticuerpos recombinantes a partir de repertorios que derivan de la información genética expresada en pacientes con enfermedad celíaca.

La colaboración de los servicios de Medicina Digestiva e Inmunología del Hospital Universitario Fundación Jiménez Díaz de Madrid permitió seleccionar pacientes con sospecha de padecer celiaquía. Estas personas, tras firmar un consentimiento informado, donaron sangre de la que obtener linfocitos productores de anticuerpos específicos frente al gluten. Los pacientes seleccionados presentaban unos elevados niveles de biomarcadores utilizados en el diagnóstico de celiaquía, y no habían comenzado una dieta sin gluten (dado que inhibe muy rápidamente la respuesta inmunitaria). Además, se realizaron pruebas complementarias sin fines diagnósticos, para profundizar en la caracterización de la respuesta inmune (isotipado y subisotipado de los anticuerpos del suero de los pacientes que reaccionaban frente a la gliadina). Finalmente se seleccionaron dos pacientes idóneos, que donaron 350 mL de sangre periférica.

A partir de estas muestras biológicas se aislaron las células mononucleares, entre las que se incluyen los linfocitos B, de las que se extrajo su ARN total. Tras la retrotranscripción a ADNc, se amplificaron las secuencias codificantes para las cadenas ligeras y pesadas de los anticuerpos. Además, a partir de un dAb, de origen semisintético, desarrollado anteriormente por nuestro grupo de investigación, se generó un juego de cadenas pesadas de Fab. A partir de este material, se generaron dos repertorios de Fab recombinantes. Un repertorio se obtuvo mediante la clonación de las secuencias obtenidas de las muestras biológicas (cadenas pesadas y ligeras inmunes), y el segundo repertorio, que se denominó semiinmune, se obtuvo mediante la combinación de cadenas ligeras inmunes con cadenas pesadas semisintéticas.

Los repertorios generados permitieron la selección de diversos anticuerpos recombinantes mediante la tecnología del *phage display*, que posteriormente se caracterizaron para determinar su capacidad de detectar gluten de forma específica. Los cuatro anticuerpos recombinantes que presentaron mejores características de especificidad y afinidad frente al gluten (Fab-C, Fab-E y Fab-H a partir del repertorio inmune, y el Fab8E-4 del semiinmune) fueron producidos en formato soluble a escala semiindustrial a partir de *Escherichia coli* RV308 para su posterior caracterización.

El Fab-C demostró ser el anticuerpo recombinante con mejores características de sensibilidad y especificidad en diversos ensayos con muestras de gliadina en solución y con mezclas experimentales de harinas de cereales con gluten en una matriz de harina de arroz. Utilizando el Fab-C se pudo desarrollar una metodología de ELISA indirecto con un límite de detección de 11 mg/kg de gluten en muestras de alimentos. También se analizaron 50 muestras de productos comerciales. Adicionalmente, se desarrolló una metodología de ELISA sándwich indirecto, mediante la combinación del Fab8E-4 como anticuerpo de captura y el Fab-C como anticuerpo de detección, capaz de detectar hasta 10 mg/kg de gluten en muestras de alimentos, con mayor exactitud que el ELISA indirecto, y que permitió confirmar los resultados obtenidos al analizar los productos comerciales.

La caracterización de los anticuerpos recombinantes obtenidos se completó mediante el empleo de diversos sistemas de predicción del plegamiento de proteínas basados en la inteligencia artificial, para modelizar la estructura tridimensional del Fab-C y el Fab8E-4, así como su interacción con la gliadina. Los resultados obtenidos muestran que ambos Fab, sobre todo los CDR de su cadena pesada, interaccionan principalmente con epítomos presentes en la región C-terminal de la gliadina. Esta característica es novedosa, ya que otros anticuerpos monoclonales comerciales interaccionan con la región N-terminal.

Finalmente, utilizando técnicas de secuenciación masiva del ADN, se ha analizado la composición de los dos repertorios de Fab recombinantes generados en esta tesis doctoral. Los resultados han puesto de manifiesto que muchas características de la respuesta inmune celíaca se transfirieron al repertorio inmune y a los anticuerpos obtenidos del mismo, y que mediante la metodología de *phage display* se produjo una auténtica selección por afinidad de los mejores anticuerpos frente a la gliadina.



## SUMMARY

Gluten is the protein fraction of cereal grains such as wheat, barley, rye, spelt, and kamut, which is insoluble in water and soluble in hydroalcoholic solutions. Its use is widespread in the food industry because it provides functional characteristics such as extensibility to doughs or consistency to sauces. Moreover, being heat-resistant, it can be used as an ingredient in products that undergo heat treatment. Gluten is not only used as an ingredient in food, but also as an excipient in medicines and cosmetics. However, it contains peptides that can trigger pathological responses in sensitive individuals, such as allergies, autoimmune diseases like celiac disease, or the non-celiac gluten intolerance.

Although many pharmacological therapies have been developed to treat gluten-related pathologies, none have proven more effective than strict adherence to a gluten-free diet. To this end, it is necessary to take measures for prevention, control, and traceability in the industry to ensure that products marketed as “gluten-free” contain less than 20 mg/kg of this compound, in compliance with current legislation.

The most commonly used techniques for detecting gluten in foods are immunochemical methods, based on the specific interaction between an antibody and its antigen. Currently, the reference method endorsed by the *Codex Alimentarius* for the detection of gluten in food is an immunoenzymatic sandwich ELISA technique based on the monoclonal antibody R5. However, the accurate identification and quantification of gluten remain a significant challenge for food science because gluten is not a well-defined protein, unlike other food allergens. Instead, it is defined in terms of chemical characteristics that encompass a wide range of proteins with some similarities but different sequences. As a result, several studies have revealed notable discrepancies in gluten quantification when comparing results obtained with commercially available kits. Consequently, there is still a need for research and development of methods for gluten detection in foods, either to address their limitations or to complement existing methods.

Most of the antibodies for gluten detection developed so far have been obtained from hybridomas, which require the immunization of experimental animals. This methodology has produced antibodies with high affinity for the antigen used in immunization, but which do not reproduce the immune response of people suffering from a gluten-triggered disease.

The primary objective of this doctoral thesis has been the development of new analytical methodologies for the detection of potentially harmful gluten components for sensitized individuals in food samples. To this end, new recombinant antibodies have been developed by building libraries derived from the genetic information expressed in patients with celiac disease.

The collaboration of the Digestive Medicine and Immunology services of the Hospital Universitario Fundación Jiménez Díaz in Madrid made it possible to select patients suspected of suffering from celiac disease. After signing the informed consent form, these individuals donated blood from which lymphocytes producing specific antibodies against gluten were isolated. The selected patients had high levels of biomarkers used in the diagnosis of celiac disease and had not started a gluten-free diet (because its absence rapidly inhibits the pathological immune response). Additional non-diagnostic tests (isotyping and subisotyping of serum antibodies in patients reacting to gliadin) were performed to further characterize the immune response. Finally, two suitable patients were selected and donated 350 mL of peripheral blood.

Mononuclear cells, including B lymphocytes, were isolated from these biological samples and their total RNA was extracted. After retro-transcription into cDNA, the coding sequences for the light and heavy chains of the antibodies were amplified. In addition, a set of Fab heavy chains was generated from a semi-synthetic dAb, previously developed by our research group. Two recombinant Fab libraries were then produced. A library was obtained by cloning the sequences amplified from the biological samples (immune heavy and light chains), and the second library, called semi-immune, was obtained by combining immune light chains with the semi-synthetic heavy chains.

The libraries generated allowed the selection of different recombinant antibodies using phage display technology, which were subsequently characterized for their ability to specifically detect gluten. The four recombinant antibodies with the best specificity and affinity characteristics against gluten (Fab-C, Fab-E, and Fab-H from the immune repertoire, and Fab8E-4 from the semi-immune) were produced in soluble format on a semi-industrial scale from *Escherichia coli* RV308 for further characterization.

Fab-C proved to be the recombinant antibody with the best sensitivity and specificity characteristics in several assays with gliadin samples in solution and with experimental mixtures of gluten-containing cereal flours in a rice flour matrix. An indirect ELISA method with a detection limit of 11 mg/kg of gluten in food samples was developed using Fab-C. In addition, 50 samples

of commercial products were analyzed. Furthermore, an indirect sandwich ELISA method was developed using a combination of Fab8E-4 as the capture antibody and Fab-C as the detection antibody. This method was able to detect up to 10 mg/kg gluten in food samples with a higher accuracy than the indirect ELISA and allowed confirmation of the results obtained when analyzing commercial products.

The recombinant antibodies were further characterized by modelling the three-dimensional structure of Fab-C and Fab8E-4 and their interactions with gliadin, using various artificial intelligence-based protein folding prediction systems. The results showed that both Fabs, especially the CDRs of their heavy chains, interact mainly with epitopes located in the C-terminal region of the gliadin. This is a novel feature, as other commercially available monoclonal antibodies interact with the N-terminal region.

Finally, by applying high-throughput DNA sequencing techniques, the composition of the two recombinant Fab libraries generated within this doctoral thesis has been analyzed. The results have shown that many characteristics of the celiac immune response have been successfully transferred to the immune library and to the antibodies obtained from it, and that a true affinity selection of the best antibodies against gliadin was produced using phage display methodology.



# EXTENDED SUMMARY

## 1. INTRODUCTION

Gluten is the general term used to define the group of ethanol soluble proteins found in the endosperm of cereals such as wheat (including varieties like spelt and kamut), barley and rye (Fasano & Catassi, 2012). The Codex Alimentarius provides the following definition based on its physicochemical properties: “a protein fraction from wheat, rye, barley or their crossbred varieties and derivatives thereof, to which some persons are intolerant and that is insoluble in water and 0.5 M NaCl” (Codex Alimentarius, 2008). These cereals are widely used in the food industry due to the properties that gluten gives to doughs. Primarily, it plays an essential structural role in many bakery products, providing elasticity and helping dough rise, resulting in the desired texture and volume of bread, pastries, and cakes. Its cohesive properties also make it an ideal ingredient in pasta production, contributing to the formation of a firm and chewy dough. Moreover, gluten acts as a binding agent and stabilizer in processed foods, such as sausages and processed meats, helping to maintain their shape and consistency. Additionally, Gluten can also be incorporated into different foods to improve their nutritional profile by increasing their protein content.

However, this remarkable physical and chemical stability of gluten protein, leads to a gradual and prolonged digestion process within the gastrointestinal tract. This poor digestibility has been associated with the pathogenic effects of gluten ingestion. Gluten-related disorders can be classified into three main groups according to their pathological mechanism. Firstly, there are allergic conditions, which include diseases such as baker’s asthma (a respiratory allergy allergic caused mainly by inhalation of cereal flour), IgE-mediated food allergies (particularly against wheat proteins), contact urticaria and WDEIA (wheat-dependent exercise-induced anaphylaxis) (Sharma et al., 2020). The second group comprises autoimmune diseases, with celiac disease being the most widely recognized example. Other conditions in this category include gluten ataxia and dermatitis herpetiformis (Sapone et al., 2012). The third group includes the non-celiac gluten sensitivity (NCGS) and encompasses situations whose pathogenesis cannot be attributed to any of the previous categories. Although the specific molecular causes of NCGS are not fully elucidated, some research suggests a significant involvement of the innate immune system (Volta et al., 2019).

In this context, celiac disease is particularly relevant due to its high prevalence, which has been estimated at 1% of the population. It has been described that in celiac patients, the immune system mounts a strong response against self-transaminases involved in the intestinal gluten processing, resulting in severe damage of the intestinal tissue. Undigested gliadin peptides can increase intestinal permeability and exhibit specific binding to the CXCR3 chemokine receptor. Consequently, this interaction triggers the release of zonulin, a modulator of intestinal tight junctions. While this phenomenon occurs in all individuals, it can lead to inflammatory responses in celiac consumers, particularly in those individuals carrying the human leukocyte antigen (HLA)-DQ2 and DQ8 haplotype (Leonard et al., 2017).

Although there are different therapeutic alternatives for celiac disease under investigation (such as anti-inflammatory drugs, inhibitory monoclonal antibodies, and gluten chelators), the strict adherence to a gluten-free diet stands as the most effective treatment. This dietary approach not only halts the autoimmune destruction of intestinal tissue, but also promotes the healing of intestinal crypts and the recovery of a proper digestive function (Itzlinger et al., 2018). Based on the establishment of a safe gluten threshold for celiac patients (50 mg of gluten per day), current European legislation indicates that foods containing less than 20 mg/kg of gluten can be labeled as “gluten-free” (Official Journal of the European Union, 2014). In this scenario, taking into account the high prevalence of gluten-related disorders and the proven efficacy of a gluten-free diet, rapid and effective methods for gluten detection in foods are essential to strengthen traceability and allergen control systems in the food industry. Only in this way could it be guaranteed that accurate labeling is provided to allergic consumers, thus contributing to their safety and well-being.

Among the different analytical techniques available for the detection of gluten in foods (García-García et al., 2018), immunoassays based on the detection of proteins, specifically ELISA methodologies, are of particular interest due to their precision and simplicity. These methods rely on high-affinity antibodies specific to gluten. There are currently immunoassays on the market for the qualitative and quantitative detection of gluten based on monoclonal antibodies obtained by immunization of animals. Among them, the sandwich ELISA based on the monoclonal antibody R5 is considered as a Type I method by the Codex Alimentarius, and is therefore the gold standard for the detection of gluten in food products (Valdés et al., 2003). Despite significant legislative advances in food safety and the availability of a wide variety of gluten immunoassays, the analysis of this allergen presents several challenges. Multiple independent comparative studies have revealed considerable discrepancies in gluten quantification when using different commercial kits (Bugyi et al., 2013). The term gluten itself encompasses a wide variety of proteins, which consequently confers further complexity to the analysis process. It has been found that, despite sharing certain similarities, the sequences of gluten constituting proteins can differ significantly. Furthermore, the composition of gluten proteins can vary considerably even among cereals of the same species, and can be influenced by various factors, such as the variety of cereal, its origin and even the method of cultivation (Wieser et al., 2023). Due

to this large variability, further research is needed in the development of gluten detection methods to obtain more consistent results.

In this context, it is crucial to advance in the generation of new antibodies against gluten which exhibit different characteristics to the currently available antibodies. The implementation of these new affinity probes in a set of detection systems, either individually or in combination with other probes, could provide an essential strategy to address the limitations inherent to gluten quantification in food. The search for new antibodies using directed protein evolution techniques, such as phage display technology (Smith et al., 2019), is postulated as an interesting alternative to meet these requirements. Recombinant antibodies offer several advantages such as scalability, customizability, and the ability of working with easy and consistent production processes (Bradbury et al., 2018).

The last years of the 20th century ushered in a new era in antibody development, thanks to the pioneering work in phage display of Winter and Smith (Smith et al., 2019), who introduced the concept of directed molecular evolution. This innovative approach transcends the mere recombinant expression of antibodies, encompassing three key steps: the generation of diversity, the selection of high-affinity antibodies against a specific target antigen, and their subsequent amplification. The main prerequisite for the successful generation of recombinant antibodies using this technology relies on the availability of high quality antibody libraries from which guide the selection of specific probes. Thereafter, through 3 to 6 rounds of screening, one to ten variants with the desired affinity can be identified from the initial set of  $10^7$  antibodies. Following the selection process, these binders become dominant in the phage population, and their genetic characteristics (genotype) are linked to their observable features (phenotype).

Phage display has been applied to elucidate the molecular mechanisms of celiac and related diseases, both through immune and naïve antibody libraries (Not et al., 2011) as well as peptide libraries (Chen et al., 2011). Although clinical research has been the primary focus, the generation of antibodies against gluten for the analysis of food has also been described. For this purpose, a sandwich-ELISA was successfully developed based on VHH antibodies which were isolated from a phage library generated from immunized llamas (Doña et al., 2010). As in this work, libraries are frequently constructed from the genetic material of allergen immunized animals, since these antibodies have undergone a natural affinity maturation process. However, the response elicited by intramuscular allergen injection during animal immunization may significantly differ from the response triggered by gluten ingestion in a person with celiac disease. For this reason, it was hypothesized in this doctoral thesis that transferring the characteristics of the humoral response of celiac individuals to the antibody libraries could potentially lead to the production of probes that recognize the immunogenic gluten fractions for these consumers. Therefore, the methodologies developed in this work should be based on recombinant antibodies selected from

libraries generated, either partially or entirely from the genetic information of celiac patients. This innovative approach could not only contribute to improve detection accuracy, but would also allow a significant advance in the ability of analytical methods to adapt to the complexity and diversity of gluten proteins present in food.

## 2. OBJECTIVES

Taking into account these considerations, the main objective of this doctoral thesis was the development of new gluten immunodetection systems based on recombinant antibodies obtained by phage display. For this purpose, two Fab (antigen-binding fragment) libraries were constructed from peripheral blood lymphocytes (PBLs) of celiac donors. Firstly, a semi-immune library was generated through a protein engineering process in order to improve the performance of a previously developed single domain antibody against gluten (dAb8E). Alternatively, in order to find new affinity probes, it was proposed the construction of a second immune library in which the humoral response of celiac patients is transferred to both Fab chains. In this context, to achieve this global objective, the following partial objectives were established:

1. Selection of celiac patients and collection of biological samples to be use as starting material for the construction of the recombinant antibody libraries.
2. Construction of a semi-immune Fab library by fusion of a heavy chain based on a variable domain against gliadin (dAb8E) with light chains derived from the PBLs of celiac patients.
3. Construction of an immune Fab library by amplification of both heavy and light chains from PBLs of celiac patients.
4. Affinity selection of Fab variants with gliadin-binding capacity from both libraries and characterization of the selected phage-antibodies.
5. Semi-industrial soluble production of the best gliadin-binding Fab candidates and development of ELISA methods for gluten detection in food.
6. *In silico* elucidation of the three-dimensional structure of the two Fab with the best features for gluten detection and study of their interaction with gliadin.
7. Characterization of both libraries and their affinity selection processes through a massive DNA sequencing technique based on nanopores.

## 3. SUMMARIZED METHODOLOGY

To accomplish the objectives outlined in this doctoral thesis, diverse immunological, biochemical, genetic, genomic and proteomic methodologies were employed. A brief description of these methodologies is given in this section.

### 3.1. Sample and antigen preparation

In this work, gliadin-PWG (produced by the Prolamin Working Group) was used as a reference material for gluten detection in food. Following the supplier's recommendation, this material was dissolved at a concentration of 1 mg/mL in a 60% (v/v) ethanol/water solution.

All the analyzed samples (cereals with and without gluten, heterologous species, and commercial products) were individually ground in an analytical mill using 50 g of each sample. Ethanol extracts were prepared from several cereals (wheat, rye, barley, corn, rice, among others) from 250 mg of the finely ground sample with 10 mL of a 60% (v/v) ethanol/water solution, followed by shaking for 30 minutes at room temperature and centrifugation at 867 g for 15 minutes. For immunoblotting assays, extracts were performed in a similar manner as that described above but starting from 200 mg of sample with 1 mL of 50% (v/v) isopropanol/water solution.

To evaluate the developed ELISA techniques, binary mixtures prepared with increasing concentrations of wheat, barley, and rye flours in a gluten-free matrix based on rice flour were analyzed. Commercial food products subjected to diverse processing methods were also analyzed. In these cases, gluten was extracted from 250 mg of the ground samples with 2.5 mL of Ingezim gluten extraction solution. This mixture was homogenized and incubated at 50 °C for 40 minutes before adding 7.5 mL of an 80% (v/v) ethanol/water solution. Finally, it was incubated for 1 hour on a vertical shaker and then centrifuged at 2000 g for 10 minutes at room temperature to obtain the protein extract from the supernatant.

### 3.2. Identification of patients with high levels of biomarkers of the celiac humoral response

Patient selection was performed at the Fundación Jiménez Díaz University Hospital (FJD, Madrid, Spain) after approval of the study by the ethics committee. The physicians from the Department of Digestive Medicine and Immunology recruited donors among those patients with initial suspicion of celiac disease during their medical consultations. Patients identified as potential donors, who had not started a gluten-free diet (since the humoral response disappears rapidly when gluten is eliminated from the diet), underwent an analysis to evaluate their celiac humoral response. The biomarkers used in this evaluation were IgG and IgA antibody titers against tissue transglutaminase and gliadin deamidated peptides. Finally, patients who exhibited high levels of these humoral markers according to the clinical guidelines used at the FJD Hospital, and which provided the signed informed consent to participate in the study, made a donation of 350 mL of blood. Additionally, isotyping (general levels of IgA and IgG) and subisotyping (specific levels of IgG1, IgG2, IgG3, and IgG4) of their humoral response to native gliadin were performed using ELISA techniques.

### 3.3. RNA extraction and retro-transcription of the biological samples for Fab libraries building

Peripheral blood mononuclear cells (PBMCs) were isolated from the blood samples by a Ficoll-Paque® density gradient separation procedure. From the isolated PBMCs, total RNA was extracted using a standard protocol based on TriZol LS reagent. Before employing the RNA for the reverse transcription reaction (RT-PCR), a microarray systems was used to measure its concentration and quality. Afterwards, non-specific RT-PCR was performed using oligo-dT-20 primer. The complementary DNA (cDNA) generated was used as template in the following step for subsequent amplification of the specific genes encoding the antigen-binding site of immunoglobulins.

### 3.4. Amplification of antibody chain-encoding sequences

Throughout this research work, two distinct Fab repertoires were constructed. The first library resulted from the combination of semi-synthetic heavy chains with a sub-library of immune light chains obtained from the biological samples of the celiac donors, resulting in a semi-immune library. In contrast, the second library was constructed with an entirely immune nature, with both heavy and light chains amplified from the cDNA obtained from the celiac patients.

The semi-synthetic heavy chains of the semi-immune library derived from a parental dAb (a single domain antibody named as dAb-8E) that exhibited high affinity for gluten. This antibody fragment was isolated in previous works of our research group from a semi-synthetic library composed of mutated human VH3-23 genes. The dAb-8E served as a template for the VH domain of the Fab, and a set of CH1 domains (corresponding to each IgG sub-isotype) were added by overlap extension polymerase chain reaction (OE-PCR). This strategy allows the assembly of overlapping DNA fragments in which the 3' end of the first amplicon (VH domain) is completed with the 5' end of the second amplicon (CH1 domain).

To obtain the immunoglobulin chains of immune origin, the light and heavy chains were independently amplified using a collection of primer sets. Amplification was performed through two consecutive rounds of PCR. In the first round of PCR, primer pairs were used that hybridized at 5' with sequences specific to each V gene family and at 3' with sequences characteristic of each immunoglobulin subtype. Thus, the target chains were amplified adding a different restriction site to each end. Resulting PCR products were purified through agarose gel electrophoresis and subjected to a second round of PCR to introduce enhanced restriction sites by elongation of the DNA regions adjacent to the restriction sites in order to facilitate subsequent cloning.

### 3.5. Construction of the recombinant antibody libraries and affinity selection of gliadin binders

Final PCR products were cloned into the pComb3X vector, a phagemid that allows the expression of antibody fragments fused to the protein III of the helper phage capsid (VSC-M13). A sequential cloning approach was employed to construct the libraries. Initially, a sub-library containing the amplified light chains was generated and transformed into *E. coli* XL1-Blue via electroporation. The transformed cells were cultured in a carbenicillin-containing medium, from which the phagemids containing the sub-library of light chains were purified. Subsequently, the heavy chains obtained from two different sources were separately cloned, resulting in the generation of the two described Fab libraries (semi-immune and immune libraries).

Each library was transformed into electrocompetent *E. coli* XL1-Blue cells. The transformed cells were recovered in SOC medium, followed by plating of serial dilutions of the transformed bacteria to evaluate transformation efficiency and determine library size (RZ). Subsequently, the culture was expanded in Super-Broth media (SB: 30 g/L tryptone, 20 g/L yeast extract, 10 g/L MOPS, pH 7) with increasing concentrations of carbenicillin as bacterial growth progressed. Then, 10<sup>12</sup>-10<sup>13</sup> particles of the auxiliary phage VSC-M13 were added, and after 2 h of incubation, kanamycin was added to the medium for selection of those bacteria co-transformed with the phagemid and the helper phage. Finally, phage-Fab were expressed in the culture supernatant after incubation at 37 °C for 16 h.

The next day, phage-antibodies were isolated from the supernatant through PEG/NaCl precipitation. For library selection against the target antigen, an immunoplate was coated with gliadin-PWG standard and blocked with 3 % BSA in PBS (137 mM NaCl, 2.7 KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4). Once the phage-Fab library was exposed to the antigen, five washes with PBS-T (PBS with 0.05% of Tween20 detergent) were performed to remove phages bound with non-specific interactions. Thereafter, phage variants specifically bound to the antigen were eluted by a sharp pH shift generated with glycine-HCl (pH 2.2). The eluted phage-Fab were rapidly neutralized with Tris-base. The recovered phages from the first round of selection were used to infect a new *E. coli* XL1-Blue culture that was expanded in SB medium, followed by a process analogous to the previously explained procedure of helper phage infection and PEG/NaCl precipitation. This cyclic steps, commonly referred to as (bio) panning, was repeated over 3-4 successive rounds, increasing the stringency of selection by introducing additional PBS-T washes (five in the first round, ten in the second, and fifteen in the subsequent rounds).

### 3.6. Monitoring of the panning process. Selection and sequencing of individual clones

The panning process was monitored after each selection round by the number of transformed bacteria infected with the phage-Fab eluted after each round. The selection of phage-Fab variants against the antigen was considered positive when a significant increase in the number of transformed bacteria was observed between two consecutive rounds. In this assay, a negative control (selection rounds without antigen) was performed to ensure that phage-Fab were not selected against other components of the system (blocking protein, support, etc...).

To confirm the enrichment of the library in antigen-specific variants, an indirect polyclonal ELISA against gliadin-PWG was conducted using the set of phage-Fab precipitated after each panning round. The system was revealed using a conjugated antibody directed to the phage capsid protein VIII. If the enrichment process was successful, an increase in the absorbance signal should be observed in the phage populations from the later rounds, which is indicative of a higher proportion of high-affinity phage-Fab.

To conduct the study of individual clones, expression of phage-Fab was induced in the supernatant of isolated colonies from the last rounds of panning that exhibited enrichment in the polyclonal ELISA. The induced supernatants were used as primary antibody in a monoclonal indirect ELISA. Those clones that demonstrated proper recognition of the target antigen, without presenting cross-reactivity against the blocking agent used as a negative control, were selected for further characterization.

In addition, the phagemids encoding each individual Fab were purified and sequenced by Sanger employing two sets of primers (one for light chains sequencing and another for heavy chains). The obtained sequences were classified using online tools such as IMGT/V-QUEST and IgBLAST, considering their V(D)J gene composition, isotype, subisotype and clonotype, as well as their structural features (CDR loops).

### 3.7. Evaluation of the selected phage-Fab as probes for gluten detection

Before expressing the Fab as soluble molecules, a first characterization in phage-Fab format was performed with all the different clones that exhibited affinity for gliadin in the monoclonal ELISA. For this purpose, PEG/NaCl purified phages were obtained at larger scale (50-100 mL of culture) to provide sufficient phage-Fab for the characterization assays.

The phage-Fab were used as the primary antibody in an indirect ELISA format against different antigens. Firstly, a dose-response curve was obtained by analyzing different concentrations of the gliadin-PWG standard. Additionally, ethanolic extracts from different cereal flours (wheat, barley, rye, oats, corn, rice, etc.) and other widely used heterologous species in food production were

examined. An experimental mixture with increasing concentrations of a gluten-containing cereals homogenate (40 wheat varieties, 10 barley varieties and 10 rye varieties) in a gluten-free rice based matrix was analyzed in order to evaluate gluten detectability in real food samples. Finally, commercial products with different types of labeling regarding the gluten-containing cereals were analyzed.

### 3.8. Production of the recombinant Fab as soluble molecules

The four Fab that exhibited superior gluten detection properties in phage-Fab assays were expressed in its soluble format. Three of the selected Fab (Fab-C, Fab-H, and Fab-E) were isolated from the immune library, while only one (Fab8E-4) was from the semi-immune library. Synthetic genes encoding each Fab were cloned into the pKktac expression vector, which allowed the optimized expression of the Fab in the culture supernatant due to the pelB signal that precedes the Fab sequences. Besides, Fab were expressed with a C-terminal histidine tag which facilitated their subsequent detection and purification. After verifying that vectors with the appropriate inserts had been obtained, they were transformed into chemically competent *E. coli* K12 RV308.

To induce the expression of the recombinant antibodies, an individual colony was inoculated in Terrific Broth (TB) media (12 g/L tryptone, 24 g/L yeast extract, and 4 g/L glycerol in phosphate buffer (0.17 M KH<sub>2</sub>PO<sub>4</sub>, 0.72 M K<sub>2</sub>HPO<sub>4</sub>, pH 7.2)) supplemented with glucose and ampicillin. After scaling up the culture to 1.8 L, it was incubated at 37 °C at 220 rpm until reaching an optical density of 4. At that point, Fab expression was induced by adding 1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside) and further incubation at 30 °C and 170 rpm for 16 h. The next day, cell-free supernatant was obtained by centrifugation, and Fab expression was confirmed through protein electrophoresis and immunoblotting.

Once the proper production of the Fab was confirmed, the supernatant was clarified by treatment with DNase, and a tangential filtration unit was used to buffer exchange and sample concentration. The clarified supernatant was purified using an Immobilized Metal Affinity Chromatography (IMAC) column coupled to a Fast Protein Liquid Chromatography (FPLC) system. The purification process was monitored at 280 nm, and collected fractions were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and peptide fingerprinting (MALDI-TOF mass spectrometry) to confirm their identity.

### 3.9. Evaluation of the recombinant Fab as probes for gluten detection

The four Fab produced as soluble molecules were used as primary antibodies in an indirect ELISA methodology against the same antigens previously described (section 3.7). Among them, clone Fab-C exhibited the best recognition properties for gluten detection, prompting a more detailed characterization of this antibody fragment.

The specificity indirect ELISA based on the Fab-C was evaluated by analyzing 60 heterologous species that showed no cross-reactivity. Assay sensitivity was calculated from the dose-response curve (linear regression) generated from the analysis of decreasing concentrations of gliadin-PWG. The limits of detection (LOD) and quantification (LOQ) were calculated from the interpolation of three and ten times the standard deviation of 10 blank measurements, respectively. Additionally, a recovery assay was conducted by analyzing certified gluten-free rice flour spiked with solid gliadin-PWG. In parallel, the rice flour extract was spiked with an equivalent concentration of gliadin-PWG as control for the recovery assay.

### **3.10. Development of a sandwich ELISA based on the recombinant Fab**

In order to develop a sandwich immunoassay based on the obtained Fab fragments, compatibility between the different antibody fragments was first confirmed. First, the four Fab fragments were chemically biotinylated. Subsequently, the Fab fragments were immobilized on streptavidin-coated plates as capture antibodies. Similarly, the biotinylated Fab fragments were incubated with the enzyme alkaline phosphatase conjugated to streptavidin to be used in a direct sandwich ELISA without the need of secondary antibody.

With this methodology, it was determined the pair of recombinant antibodies (Fab8E-4 as capture antibody and Fab-C as detection antibody) that exhibited the best properties for gluten detection in the sandwich immunoassay.

Since a loss of sensitivity was observed compared to the expected detection levels, an indirect sandwich ELISA was designed. This assay was based on direct adsorption of the capture Fab on immunoplates and revelation of the detection Fab using an HRP-conjugated anti-Fab secondary antibody. This sandwich ELISA was evaluated against the same food samples as described for the other developed immunoassays.

### **3.11. Analysis of food products with the developed ELISA systems**

To verify the applicability of the indirect ELISA based on the Fab-C and the indirect sandwich ELISA developed in this work, a total of 50 commercial products obtained from various local stores were analyzed. Based on their labeling, food products were classified as follows according to their gluten-containing cereal content: 9 products declared gluten-containing cereals in their label, 6 products had no declaration regarding the presence of gluten-containing cereals, 21 products declared that may contain gluten, and 14 were certified as gluten-free. To validate the obtained results, all products were analyzed in parallel with the direct sandwich ELISA based on the monoclonal antibody R5, which has been approved as Type I method by the Codex Alimentarius for gluten detection in food and offers a limit of quantification (LOQ) of 3 mg/kg of gluten.

### 3.12. Characterization of the generated libraries and study of the affinity selection processes by high-throughput sequencing

The phagemids encoding the phage-antibody subpopulation obtained after each round of panning were extracted for analysis using high-throughput sequencing. Sequencing was performed using the nanopore technology on an Oxford Nanopore® MinION device. Since linear fragments of double-stranded DNA are needed for this sequencing system, the obtained phagemids were treated with the restriction enzyme *NheI*. Each sample coming from a different panning round was labeled with a unique barcode for sequence differentiation after the data acquisition process. Sequence data were collected for 30 h, and the raw data was recorded in FAST5 format files. The raw data generated was analyzed in two distinct phases: primary analysis and secondary analysis.

During the primary analysis, Megalodon v1.9.5 was used for basecalling. This procedure transformed the FAST5 files into high quality base sequences using human genome as reference. This initial analysis produced a demultiplexed FASTQ file, meaning that the information was classified into different files based on their respective barcodes. Subsequently, the adapter sequences were removed, and a quality control of the obtained sequences was carried out using the FASTQC program. Sequences of low quality or incorrect size were excluded from the analysis using the LUMC/FASTQ program.

The secondary analysis was performed with the IMGT/HighV-QUEST tool which is based on the identification of V, D, and J genes by sequence alignment against a reference database. The statistical analysis and graphical representations were generated with the R package called IMGT/StatClonotype.

### 3.13. *In silico* modeling of the structure of the Fab8E-4 and Fab-C, and elucidation of their interaction with gliadin

*In silico* modeling of the Fab8E-4 derived from the semi-immune library was performed using the sAb Pred server. Once the Fab structure (specifically its variable fragment or Fv fragment) was generated, paratopes and epitopes were identified using the structure of gliadin generated by AlphaFold2 as the antigen.

In the case of the Fab-C obtained from the immune library, the structure and antigen-antibody interactions were elucidated following the next process: 1) The nucleotide sequence was analyzed with the IMGT/V-QUEST server in order to detect and define the antibody domains; 2) the prediction of the Fab structure was obtained through the bioinformatics tools Abodybuilder2 and IgFold; 3) matches and discrepancies between both models were studied using the ChimeraX Matchmaker tool; 4) interactions between Fab-C and gliadin were elucidated by epitope and paratope predictions using the Epipred and Antibody iPatch tools; and 5) based on the epitope and paratope information obtained, a simulation of the interaction of Fab-C with the antigen (docking) was conducted using the HADDOCK 2.4 server.

## 4. RESULTS AND DISCUSSION

The emphasis in gluten detection assays has traditionally centered around monoclonal antibodies produced through hybridoma technology (such as R401, R5, and G12). Rarely have recombinant antibodies been employed to address this issue, and mainly they have been applied for studying clinical aspects of the gluten-related diseases rather than the direct detection of gluten in foodstuffs as is the case of the present research work.

**Chapter 4.1** describes the development of a library suitable for selecting antibody fragments for prolamins detection that involved a direct transformation of one antibody format into another (dAb to Fab). This work was the continuation of a previous study where a semi-synthetic library based on human VH3 germline, was confronted to a peptide proposed to be a consensus HLA-presented peptide of digested gluten. From this process, a suitable candidate was selected (dAb8E) and used as the specific reagent for the development of a phage-ELISA method for gluten detection in foodstuff. It was hypothesized that the transformation of the dAb8E into a Fab could enhance its properties, since constant domains also might affect to antigen binding due to their influence in segmental flexibility and inter-antibody associations. Therefore, a novel Fab antibody library merging chains of two different origins was generated: light-chains derived from plasmocytes isolated from celiac patients, and heavy chains derived from the dAb8E coding sequence transformed into human heavy-chains (IgG1, IgG2, IgG3 and IgG4). The light chain coding sequences were amplified from the cDNA obtained from lymphocyte-isolated RNA of two donors diagnosed of celiac disease and presenting high serum response against gluten proteins, since they had not yet adhered to a gluten-free diet. Unlike conventional methods, this approach incorporated a direct evolutionary process in order to enhance the properties of the parental dAb8E. Fab library was constructed by a two-step cloning process consisting of cloning firstly the immune-light chains and subsequently the semi-synthetic heavy chains. Firstly, double digested inserts coding for light chains were cloned into linearized pComb3X and transformed into electrocompetent *E. coli* XL-1 Blue, obtaining the light chain sublibrary ( $1.4 \times 10^7$  transformants). Subsequently, double digested heavy chain constructs were cloned into the relinearized plasmid, resulting in a Fab library with a repertoire size (RZ) of  $2.7 \times 10^7$ . The extensive diversity within this library facilitated the isolation of positive clones after three rounds of biopanning against gliadin-PWG. From the 24 individual clones tested, 4 clones (approximately 17 %) were identified as high-affinity gliadin-binding clones. Following the characterization of the different phage variants by monoclonal phage-ELISA, clone Fab8E-4 showed the most promising recognition properties, exhibiting the most sensitive dose-response curve against gliadin and specifically recognizing the extracts from the gluten-containing cereals. The dose response curve obtained against increasing concentrations of gliadin-PWG (0.5 – 6  $\mu\text{g}/\text{mL}$ ) with the phage-Fab8E-4 seemed to significantly improve the methodologies for gluten detection in food products based on the semi-synthetic dAb-8E, reaching a limit of detection (LOD) of 0.01  $\mu\text{g}/\text{mL}$ . Besides, a set of binary mixtures of

a rice-based matrix spiked with different levels of an un-treated wheat/barley/rye proportional mixture (WBRm) were analyzed. According to these experimental binary mixtures, the developed phage-Fab8E-4 was able to detect up to 10 ppm of gluten, fulfilling the legislation requirements that set the threshold of 20 ppm to classify the products as gluten-free. The field application functionality of the proposed method was further assessed through the analysis of ten food products (five labeled as gluten-free and five as containing gluten) being the Fab8E-4 indirect phage-ELISA able to differentiate both types of products. Consequently, Fab8E-4 exhibited enhanced properties compared to its parental dAb, underscoring the utility of this methodology for enhancing the performance of existing antibodies. Finally, an in-depth characterization of the generated antibody was conducted using computational tools. Based on the sequencing results of the Fab8E-4, the whole molecule 3D-structure prediction was generated using Alpha-Fold2 and the Fv (fragment variable of the antibody, composed of the VL and VH), was computed using the AbodyBuilder tool included in the sAbPred server. Once the 3D model of the Fab8E-4 was obtained, epitope and paratope prediction was assessed to investigate the antigen-antibody interaction surface. These predictions were performed with EpiPred and Antibody-iPatch algorithms, using as protein templates the 3D antibody structure generated by AbodyBuilder and the antigen 3D-structure of wheat  $\alpha/\beta$ -gliadin and barley  $\gamma$ -hordein solved by Alpha-Fold, since the structure of gliadins are not completely solved by X-ray crystallography or cryomicroscopy. Regardless the prolamin protein used in the analysis of the paratope models, the main antigen-antibody interactions remained in the heavy chain of Fab8E-4. However, reformatting of the dAb8E added new contacts between the antigen and light chain residues, which presumably could have improved the characteristics of the detection method. The epitope predictive models showed that the conformational epitopes were mainly found in the  $\alpha$ -helix motifs on the C-terminal region of both antigens. This fact suggests the potential of dAb8E and Fab8E-4 recombinant antibodies as new alternatives for gluten determination, providing likely contacts to the C-terminal region of the antigens, which could complement the already available antibodies that usually bound to the N-terminal region. In conclusion, the obtained Fab8E-4 presented enhanced properties regarding the parental dAb, indicating that the proposed methodology was a useful tool to improve the performance of existing antibodies. This is the first time that a directed evolved engineered antibody fragment has been applied for gluten detection in foodstuffs.

**Chapter 4.2** presents the first construction and application of a fully human immune library with the final goal of obtaining a novel tool for gluten detection in foods. It was proposed the isolation of new Fab recombinant antibodies against gluten through the construction of an immune library generated from the mRNA of two celiac patients with a high humoral response to gluten-related proteins. The objective of this study was to translate the humoral features found in the selected donors to build a platform for antibody discovery which could result in developing accurate systems for gluten detection in food. Amplification of the fragments coding for the antibody chains was performed in two sequential PCR rounds including as many VH genes as possible to generate

a wide diversity, in contrast to other studies that focused on certain VH families (such as VH4 or VH5). The Fab library was built through a two-step cloning procedure: firstly, the light chains were introduced into the pComb3X vector and transformed into *E. coli* XL1-Blue cells, followed by a similar procedure for cloning the Fab-heavy chains. The resulting Fab library (with a RZ of  $1.1 \times 10^7$ ) was subjected to selection against gliadin using phage display technology. Following three rounds of enrichment, forty-eight individual clones were assayed and the Fab-C variant was selected as the most promising candidate for gluten detection in foods. Fab-C belonged to the IgA2 immunoglobulin subclass (directly linked to celiac pathology), presented the VH3-15 locus (a scarcely expressed gene in the human repertory, but one of the most identified loci in gluten-reactive antibodies in celiac disease) and had an unusually long HCDR3 of 21-mer. The indirect ELISA based in phage-Fab-C demonstrated good functionality to specifically detect the gluten-containing cereals with a limit of detection of 15 mg/kg (obtained from the analysis of the rice-based mixtures spiked with increasing amounts of the WBRm). The developed methodology demonstrated its capability to discriminate gluten-free products in accordance with the existing legislation, showing a strong correlation with the results obtained by the standard method for gluten detection in foods (R5 sandwich ELISA). Finally, a quick, simple, and universal pipeline, based in novel prediction tools, has been developed. *In silico* elucidation of the Fab-gliadin interaction revealed that the antigen epitopes recognized by the Fab-C were likely to occur on the C-terminus. As expected, the main paratopes of the Fab-C were found in the heavy chain. The HCDR1 interacts with proline and glutamine residues (position 250–251), the HCDR2 mainly interacts with glutamines (positions 203, 207, 251 and 254), and the abnormal length of the HCDR3 enabled it to have double contact with the antigen. Interestingly, the PQ pattern form by proline and glutamine residues (which are likely to be detected with the Fab-C) have been identified in several peptides related to celiac toxicity. While the contribution of the light chain to the interaction appeared to be comparatively weaker, the model indicated a potential contact between the LCDR3 and the glycine 159 and arginine 160 residues, which could likely contribute to the overall stability of the union.

**Chapter 4.3** describes the semi-industrial scale production of the best phage-Fab candidates for gluten detection in food based on the pkktac expression vector and the *Escherichia coli* RV308 biofactory. The four selected gliadin-binding recombinant antibodies had been previously isolated by phage display from the semi-immune (Fab8E-4) and the immune (Fab-C, Fab-H, Fab-E) Fab libraries generated in this doctoral thesis. To enhance the recombinant soluble expression of Fab, the original isotypes found after library isolation were switched to IgG1. The recombinant Fab fragments were expressed in the culture supernatant, facilitated by the pelB secretion signal located upstream the antibody sequence and the weak cell wall of the bacterial strain used. Western blot analysis of the supernatants confirmed effective Fab expression (stained protein bands of approximately 50 kDa), prior to the purification process by IMAC (Immobilized Metal Affinity Chromatography). The recombinant production strategy yielded 6

to 9 milligrams of antibodies, a high yield for *E. coli* production using non-continuous cultures (1.8-liter flasks). This approach presents an economically viable and highly efficient alternative to conventional antibody production methods, which can be both time-intensive and costly. Furthermore, recombinant antibodies offer a more uniform and reproducible affinity probe, allowing the development of standardized methods without the need for experimental animals. Despite their close phylogenetic relationship, the soluble Fab antibodies did not exhibit any cross-reactivity with the non-gluten-containing cereals in the indirect ELISA developed. The sensitivity of the different assays was evaluated by analyzing gliadin-PWG reference material dilutions. The results showed that FabC exhibited the highest sensitivity (limit of detection of 28 ng/mL of gliadin, which corresponds to a gluten concentration of 11 mg/kg), followed by Fab8E-4. Therefore, the indirect ELISA based on Fab-C was further characterized. A gluten recovery analysis was performed by testing a gluten-free certified rice flour spiked with both solid and dissolved gliadin-PWG, resulting in adequate recovery values in the range of 75 – 125 %. Moreover, the analysis of 60 heterologous samples demonstrated that Fab-C exhibited no cross-reactivity with these matrices, indicating its high specificity for gluten detection. Finally, a wide variety of commercial food products (50 samples) were analyzed with the developed indirect Fab-C ELISA and the R5-based sandwich ELISA. Most of the samples yielded similar result with both methods and showed good agreement with product labeling following the guidelines of European legislation. However, three products tested positive by the R5 reference method but negative by the Fab-C assay. It is hypothesized that this result could be attributed to the absence of cross-reactivity of Fab-C with oats, since this product declared to contain this cereal as the main ingredient. It has been reported that certain oat cultivars produce cross-reactivity when analyzed with other detection antibodies like R5, reporting up to 100 mg/kg of gluten for some oats cultivars (Benoit et al., 2017).

**Chapter 4.4** describes the development of a sandwich ELISA methodology for gluten detection in foods based on the four recombinant Fab antibodies against gliadin previously isolated by phage display. Fab molecules underwent chemical biotinylation for immobilization on streptavidin-coated plates, serving as capture antibodies. Simultaneously, alkaline phosphatase-streptavidin was used to obtain conjugated Fab which could be used as detection probes. Although the implementation of this direct approach presented several advantages such as avoiding the need for secondary antibodies, a signal loss was observed in ELISA results. This fact was mainly attributed to the chemical biotinylation procedure which enabled the conjugation of biotin molecules to all available amino groups, (N-terminal and lysine residues). This could have led to significant alterations in the conformation of the Fabs, resulting in the observed reduced affinity for the target. From these experiments, it was concluded that the Fab-pair with the best compatibility in a sandwich ELISA format was Fab8E-4 as the capture probe and Fab-C for the detection reaction. Besides, it was demonstrated that in this case, the use of the same Fab in both positions (as capture and detection antibody) generated low absorbance values, probably

due to the competition for the same epitopes of gliadin. Subsequently, the unmodified Fab8E-4 (capture probe) and Fab-C (detection probe) were used to develop a sandwich immunoassay capable of detecting 26 ng/mL of gliadin, which corresponds to a limit of detection of 10 mg/kg of gluten in foods. A total of 60 different gluten-free species commonly used in the food industry were assayed with the developed ELISA showing no cross-reactivity. Finally, the analysis of 50 commercial products showed consistence with the results obtained with the standard method for gluten detection. Moreover, the developed immunoassay exhibited no cross-reactivity with oat products, which could be a potential advantage over other available detection systems that cross-react with this cereal. Compared to the results of the previously developed indirect Fab-C ELISA, this sandwich assay requires an additional step, which increases the time of the analysis. However, the sandwich immunoassay showed a slightly higher sensitivity, which could be attributed to the increased affinity due to the existence of two paratope-epitope interactions. In addition, this ELISA methodology yielded better results in the recovery analysis, obtaining an excellent recovery range close to 100 % (95 - 105 %), affirming the accuracy and reliability of the developed method.

**Chapter 4.5** presents the development of a rapid methodology to characterize the obtained recombinant antibody libraries (semi-immune and immune Fab libraries) by high-throughput nanopore sequencing. The use of Oxford Nanopore technology streamlines sequencing with portable equipment, yielding significantly longer reads that enable sequencing of the complete Fab without the need for assembly from fragmented data. Besides, since the phage populations recovered after each biopanning round were labeled with different barcodes, this method also allowed an in-depth understanding of the affinity selection process, capturing information often overlooked by traditional Sanger sequencing of individual colonies due to data size constraints. Comparison of the changes in library composition in the different antibody chains over the consecutive rounds of biopanning provided an overall picture of the performance of the gliadin-guided selection procedure. A total of 4.68 million of reads were collected in a FAST5 format. Through the primary analysis workflow, the reads were transformed into demultiplexed and trimmed FASTQ sequences grouped by barcode (0.5 million sequences per barcode). The primary analysis streamlined the raw data, encompassing steps like neural network-based basecalling, demultiplexing, quality assessment, low-quality sequence trimming, and reference genome alignment. After completing acquiring a set of adequate sequences in terms of quality and length (sequences with 4-6 kb in length and phred-scores greater than 15), secondary analysis was performed with high-quality sequences to study the antibody composition of the constructed libraries. Throughout the panning process of the immune library against gliadin, a pattern of selection was detected, since the VH3-15 variants expanded significantly within the Fab population after round 2. Although statistically significant changes in light chain composition occurred, light chains appeared to play a minimal role in the selection process. In contrast, the analysis of the semi-

immune library revealed that the conservation and expansion of certain features observed in the transformation of a pre-existing dAb into a Fab were not influenced by molecular selection. Suggesting that other factors, such as structural characteristics or stability, may have contributed to their preservation and amplification. Another crucial aspect under investigation was the identification of specific features associated with the celiac response within the immune library. analysis of the VH3-15 variants selected during panning shed light on specific immunological characteristics of the Fab. The lower mutation rate in the VH segment suggested that the generated Fab originated from plasma cells activated by extrafollicular T-cell dependent stimulation. This finding indicated the presence of crucial attributes found in individuals with celiac disease in the Fab derived from the immune library.

In summary, the methodology described in the present doctoral thesis for generating Fab (construction of celiac-derived Fab libraries for phage display selection) allowed the transference of some features from the humoral response of celiac patients to the obtained probes instead of the classical approach of using polyclonal or monoclonal antibodies, generated by a forced immune response of experimentation animals. This feature could be important, as the recombinant probes produced are prone to recognize those epitopes that may be harmful to gluten-sensitive patients. Despite the existence of other antibody-based tests in the market with lower LOD for gluten detection, the Fab developed in this doctoral thesis exhibit interesting attributes that warrant further exploration and advancement. Notably, both Fab-C and Fab8E-4 complied with the legal standards of detection, having a single interaction with the antigen, in contrast to monoclonal antibodies that typically require two interactions (which can be increased to four in a sandwich ELISA format). This distinctive feature of the Fab molecules results in highly efficient binding to gliadin, emphasizing their potential as valuable candidates for gluten detection. Another noteworthy aspect in the developed immunoassays is the complete absence of cross-reactivity with oats. The cross-reaction of the R5 monoclonal antibody has been attributed to the substantial similarity between certain peptides present in avenins (gluten-like proteins in oats) and gluten peptides (Hardy et al., 2015). The absence of this inadvertent detection in oat-derived products labeled as gluten-free may highlight the efficacy of Fab-C and Fab8E-4, making them valuable tools for assessing gluten content in these oats-based products.





# 1. INTRODUCCIÓN



# 1. INTRODUCCIÓN

## 1.1 CARACTERÍSTICAS DEL GLUTEN

Gluten es el término general que define al conjunto de proteínas solubles en etanol que se encuentran en diferentes endospermos de cereales como el trigo, cebada, centeno, espelta y kamut (Fasano & Catassi, 2012). La definición aportada por el *Codex Alimentarius* señala, además, que se compone de proteínas insolubles en agua y disoluciones 0,5 M de cloruro sódico (*Codex Alimentarius*, 2008). Debido a estas características, el gluten es digerido lentamente y presenta una elevada permanencia en el intestino.

En 1924, Osborne desarrolló un método, que todavía se sigue utilizando, de clasificación de las proteínas vegetales mediante su extracción con diferentes disolventes. Según este método, las proteínas de trigo se pueden dividir según su solubilidad en las siguientes fracciones (Tabla 1): globulinas (solubles en soluciones salinas diluidas), albúminas (solubles en agua), prolaminas (solubles en etanol) y glutelinas (solubles en ácido acético) (Goesaert et al., 2005).

**Tabla 1.** Fracciones proteicas de cereales.

Fracción	Trigo	Centeno	Cebada	Avena	Maíz	
<b>Globulinas</b>	Edestina					
<b>Albúminas</b>	Leucosina					
<b>Gluten</b>	<b>Prolaminas</b>	Gliadina	Secalina	Hordeína	Avenina	Zeína
	<b>Glutelinas</b>	Glutelina	Secalinina	Hordeinina	Avenina	Zeinina

Fuente: Martín et al. (2010).

Tradicionalmente, las proteínas de gluten se han dividido en dos fracciones según su solubilidad en soluciones hidroalcohólicas. Las proteínas solubles en disoluciones hidroalcohólicas (habitualmente 60-70 % de etanol) se denominan gliadinas y las insolubles, glutelinas. Estas fracciones contienen proteínas estructuralmente relacionadas, con diferente solubilidad, derivada de la presencia de monómeros que se encuentran unidos por fuerzas no covalentes (gliadinas) o puentes disulfuro intercatenarios (glutelinas). Las glutelinas se vuelven solubles en disoluciones hidroalcohólicas mediante reacciones de reducción (Shewry et al., 2003). Las glutelinas son muy heterogéneas y se pueden separar mediante técnicas electroforéticas en doce fracciones según su peso molecular (Cebolla et al., 2018). Las subunidades de glutelina se correlacionan con la calidad de los alimentos horneados (Shewry et al., 2003). Las gliadinas son polipéptidos monocatenarios y se clasifican por su movilidad electroforética a pH bajo en cuatro grupos:  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\omega$ -gliadinas (Wieser, 2007).

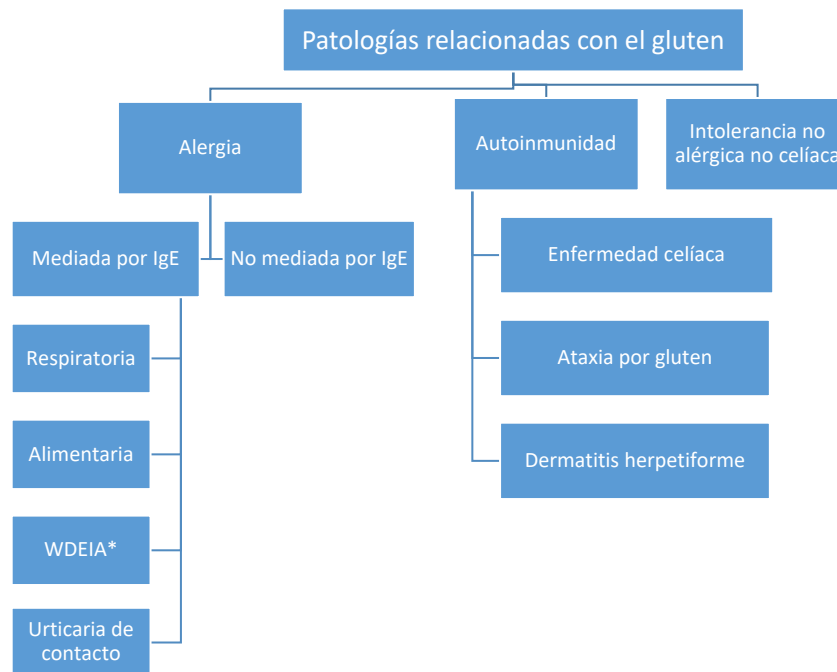
Las proteínas del gluten contienen muchos dominios repetitivos compuestos de secuencias homólogas de 6 a 8 aminoácidos, ricas en prolina (P) y glutamina (Q) (Shewry, 2019). Concretamente, en la estructura de la  $\alpha$ -gliadina se encuentra un dominio central que contiene penta (PQQPY) y heptapéptidos (PQPQFP), que forman un característico péptido inmunogénico de 33 aminoácidos (33-mer), compuesto por seis epítopos que se superponen entre sí, y son importantes en la enfermedad celíaca (Ozuna et al., 2015). Sin embargo, este péptido no aparece en todas las variedades de trigo (Schalk et al., 2017). Varios fragmentos de gliadinas y glutelinas están asociados con enfermedades como, por ejemplo, las  $\gamma$ -gliadinas con la enfermedad celíaca (Balakireva & Zamyatnin, 2016), mientras que las  $\alpha/\beta$ ,  $\gamma$ ,  $\omega 5$ ,  $\omega 1,2$ -gliadinas y las subunidades de alto y bajo peso molecular de glutelina están implicadas en alergias al gluten (Matsuo et al., 2015).

## 1.2 PATOLOGÍAS RELACIONADAS CON EL GLUTEN

Se han descrito varias enfermedades relacionadas con la exposición al gluten en personas sensibilizadas (Figura 1), que se pueden clasificar en tres grupos: alergias, enfermedades autoinmunes e intolerancia al gluten no celíaca (Caio et al., 2019).

### 1.2.1 Patologías de etiología alérgica

Las alergias relacionadas con el gluten también son conocidas como alergia al trigo, y presentan una prevalencia del 0,1 % en la población general (Sapone et al., 2012). Se desarrollan mediante un mecanismo patológico bien conocido con dos etapas: la fase de sensibilización y la fase efectora.



**Figura 1.** Clasificación de las diferentes patologías relacionadas con el gluten según su etiología. \*WDEIA: *Wheat-Dependent Exercise-Induced Anaphylaxis*, anafilaxia inducida por el ejercicio dependiente de trigo. Fuente: elaboración propia

En la fase de sensibilización, las células dendríticas mucosales capturan, procesan y transportan los alérgenos a los ganglios linfáticos más cercanos, para promover la generación de células Th2 específicas, en un proceso regulado parcialmente por la interleuquina 4 (IL-4). La expansión clonal y la activación de las células Th2 específicas que producen IL-4 e IL-13 favorecen el cambio de clase en los anticuerpos de IgM a IgE en las células B. Los anticuerpos IgE específicos de alérgenos se unen a la superficie de las células efectoras (mastocitos y basófilos) a través del receptor de alta afinidad FcεRI de las IgE, lo que lleva a la sensibilización del paciente. También se genera en esta fase una reserva de memoria de células Th2 y células B específicas de alérgenos. Los basófilos y las células NKT (*Natural killer T*) también podrían representar una fuente temprana de IL-4 (Palomares et al., 2017).

En la fase efectora, las nuevas exposiciones de los pacientes sensibilizados al alérgeno provocan su unión a las IgE ancladas a basófilos y mastocitos, lo que conduce a la liberación de mediadores proinflamatorios como la histamina, proteasas, prostaglandinas, leucotrienos y citoquinas responsables de los síntomas de la inflamación aguda. La acumulación de estos mediadores y la activación de células Th2 específicas, en cooperación con las células linfoides innatas de tipo 2 (Annunziato et al., 2015), la eosinofilia y el reclutamiento de células con capacidad proinflamatoria, así como la producción de moco, y el daño tisular es lo que produce las manifestaciones clínicas más graves y la inflamación crónica. Las células NKT, Th1 y/o Th17 también podrían contribuir a la infiltración de neutrófilos y al daño tisular (Akdis et al., 2012).

La alergia mediada por IgE incluye una serie de patologías clasificadas en función de la localización anatómica de sus síntomas: (a) alergia respiratoria, también conocida como el asma del panadero con sintomatología bronquial (Al Badri et al., 2020); (b) alergia alimentaria, con manifestación digestiva (Sharma et al., 2020); (c) anafilaxia inducida por el ejercicio dependiente de trigo (WDEIA, *Wheat-Dependant Exercise-Induced Anaphylaxis*), desencadenada por situaciones de estrés (Al Badri et al., 2020) y (d) urticaria de contacto, con sintomatología dérmica (Sapone et al., 2012).

El trigo puede causar otro tipo de alergia que no está mediada por IgE. En esta patología se produce una respuesta Th2 exacerbada, con procesos inflamatorios que suelen ser órgano-específicos como la esofagitis, gastritis, gastroenteritis y colitis, todas ellas con un pronunciado componente eosinofílico (Cianferoni, 2016).

### 1.2.2 Patologías de etiología autoinmune

El segundo grupo de enfermedades relacionadas con el gluten son las patologías autoinmunes. La principal patología de este grupo es la enfermedad celíaca, y tiene una prevalencia del 1 % en la población general (Sapone et al., 2012).

Aunque muchos mecanismos moleculares que producen el daño intestinal en la enfermedad celíaca son bien conocidos, todavía quedan otros por esclarecer. La ingestión de gluten por personas sensibilizadas da lugar a la digestión parcial de la gliadina, que interacciona con el receptor CXC3 (receptor de quimioquinas 3) que estimula la liberación de zonulina (Gao et al., 2020), produciéndose un incremento de la permeabilidad intestinal que facilita la translocación de péptidos de gliadina del lumen a la lámina propia, estimulándose la secreción de mediadores de la inmunidad innata (IL-15 e IL-18), asociada a un reclutamiento neutrofílico (Hollon et al., 2015). El debilitamiento de la barrera intestinal facilita el reclutamiento del complejo del receptor de tipo toll (TLR) 4-M2-CD14 por inhibidores de tripsina y  $\alpha$ -amilasa, estimulando la liberación de citoquinas proinflamatorias (Kim et al., 2015). Debido a la apoptosis de células intestinales mediada por la inmunidad innata, se produce la liberación de transglutaminasa tisular intracelular (tTG), que provoca la desaminación parcial de los péptidos de gliadina (Caio et al., 2019). Dichos péptidos desaminados son presentados a través del haplotipo DQ2/8 del Complejo Mayor de Histocompatibilidad (CMH) a los linfocitos T ayudantes que, a su vez, estimulan la activación, diferenciación y expansión de los linfocitos B productores de anticuerpos IgM, IgG e IgA dirigidos frente a la tTG (Stamnaes & Sollid, 2015). Además, los linfocitos T activados producen citoquinas proinflamatorias como el factor de necrosis tumoral o el interferón-gamma (Pagliari et al., 2015). La enteropatía se inicia debido a esta respuesta humoral sumada a la actividad de los linfocitos T citotóxicos. Los enterocitos dañados expresan fuertemente el transportador CD71 que estimula la retrotranscitos (Schumann et al., 2008). Se induce, a continuación, un incremento de la permeabilidad intestinal generando un ambiente proinflamatorio y estimulante del desarrollo celular, que da como resultado la aparición de criptas hiperplásicas que afectan a la absorción de nutrientes (Caio et al., 2019).

Existen otras enfermedades autoinmunes diferentes a la enfermedad celíaca y relacionadas con la ingestión de gluten. La ataxia por gluten es una enfermedad neurológica caracterizada por una pérdida irregular de células de Purkinje en todo el córtex cerebeloso, debido a la aparición de autoanticuerpos frente a estas células que pueden reconocer y reaccionar de forma cruzada con péptidos de gliadina en la circulación intestinal y, a través de un proceso de propagación de epítomos, generar anticuerpos antigliadina (Hadjivassiliou et al., 2008). También se han descrito patologías dermatológicas, como la dermatitis herpetiforme causada por la respuesta mediada por autoanticuerpos dirigidos frente a la tTG2 al igual que en la enfermedad celíaca (Salmi, 2019). Aunque esta última es la más común, también se han descrito otras patologías dérmicas relacionadas con el gluten como la psoriasis, pustulosis palmoplantar, estomatitis, vitíligo, angioedema hereditario, urticaria, vasculitis, dermatosis, dermatitis atópica, y alopecia (Muddasani et al., 2021).

### 1.2.3 Otras patologías

La intolerancia al gluten no alérgica y no celíaca es el tercer grupo de enfermedades relacionadas con el gluten. La prevalencia está estimada en un 7 % en la población general (Sapone et al., 2012). Los mecanismos patológicos son desconocidos pero la inmunidad innata juega un papel muy importante (Volta et al., 2019). Los signos y síntomas son muy similares a otras patologías relacionadas con el gluten, la enfermedad de Crohn o el síndrome de intestino irritable. El haplotipo DQ2/8 y la expresión de IgG e IgA antigliadina están presentes solo en la mitad de los pacientes, y se observa un menor daño intestinal que en la enfermedad celíaca (Roszkowska et al., 2019). Además de las fracciones del gluten, se han descrito otros componentes alimentarios que están relacionados con el desencadenamiento de esta patología, como son los inhibidores de amilasa y tripsina que son un tipo de albúmina presente en cereales y otras plantas, con una gran resistencia a la acción de las enzimas digestivas y con una alta permanencia en el intestino, que pueden activar receptores TLR, desencadenando respuestas inflamatorias (Biesiekierski, 2017). Debido a la falta de biomarcadores y criterios diagnósticos claros, esta patología se solía diagnosticar cuando se descartaban las anteriores. Actualmente se recomienda no seguir este criterio, sino llevar a cabo una exhaustiva observación de la aparición o desaparición de síntomas tras la eliminación del gluten en la dieta del paciente, seguida por una reintroducción paulatina del mismo (Catassi et al., 2015).

Algunos autores sugieren que el consumo de ciertas sustancias fermentables en el intestino como mono, di, oligosacáridos y polioles podría exacerbar los síntomas de la enfermedad celíaca y de la intolerancia no celíaca (Priyanka et al., 2018).

### 1.2.4 Tratamiento farmacológico

El desarrollo de un tratamiento farmacológico para las patologías relacionadas con el gluten ha sido un objetivo prioritario para grupos de investigación y el sector industrial. Una vía consiste en el uso de enzimas con capacidad hidrolítica del gluten, para eliminar su presencia en el intestino, de manera análoga al uso de lactasas en la intolerancia a la lactosa (Tack et al., 2013).

Otra estrategia trataría de prevenir la absorción intestinal de péptidos relacionados con estas patologías, mediante la utilización de un polímero no absorbible, con capacidad de unirse al gluten, evitando así su degradación y la aparición de péptidos inmunogénicos (McCarville et al., 2014). De forma análoga, en lugar de un polímero sintético, se pueden emplear anticuerpos aviares (IgY) con alta afinidad por el gluten, que han demostrado la capacidad de interrumpir su absorción en el tracto intestinal de animales de experimentación (Gujral et al., 2015). Otra propuesta consistiría en el uso de antagonistas de la zonulina, un modulador de las uniones celulares entre enterocitos, que evitaría la permeabilidad intestinal inducida por la gliadina (Leffler et al., 2015).

La catálisis de la desaminación de péptidos del gluten por parte de enzimas tTG2 endógenas es central en la patología celíaca. Por este motivo se han desarrollado inhibidores de estas enzimas, aunque aún no hay ensayos clínicos para su utilización terapéutica (Rauhavirta et al., 2013).

Una estrategia completamente diferente consistiría en la aplicación de inhibidores de la catepsina S, una cistein-proteasa que puede tener un importante papel en la degradación de antígenos a péptidos procesados para su presentación unidos al complejo CMH de clase II (Kurada et al., 2016).

En diversos estudios se ha demostrado que las patologías relacionadas con el gluten provocan cambios en la microbiota intestinal que favorecen la colonización de patógenos, por lo que se ha propuesto la utilización de antibióticos (Chang et al., 2011) y probióticos (Smecuol et al., 2013) para reestablecer un equilibrio saludable en la composición de la microbiota intestinal.

Se han propuesto también otros abordajes, como la infección con helmintos intestinales (debido al efecto inmunomodulador que provocan) o nanopartículas unidas a gliadina. Otras aproximaciones más tradicionales incluyen la suplementación con enzimas pancreáticas o terapias antiinflamatorias basadas en aminosalicilatos, corticoides, análogos de nucleósidos o inhibidores de calcineurina. También se han aplicado trasplantes autólogos de células madre para regenerar el intestino dañado.

Finalmente, también se ha planteado el uso de interleuquinas recombinantes (IL-10), anticuerpos monoclonales inhibidores de IL-15 y bloqueantes del receptor de quimioquinas CXCR2-B (Kurada et al., 2016). Existen estudios basados en la utilización de anticuerpos monoclonales como el alemtuzumab y el infliximab para el tratamiento de la inflamación intestinal causada por el gluten (Verbeek et al., 2006).

### 1.2.5 Tratamiento dietético

A pesar de todos los intentos efectuados para desarrollar un tratamiento para estas patologías, la mejor alternativa para frenar los síntomas sigue siendo que las personas sensibilizadas frente al gluten lleven una dieta libre de este compuesto. Además de la sintomatología típica

de estas enfermedades, cuando estos pacientes se exponen al gluten se puede producir un agravamiento de la patología, acompañada de malnutrición o de un aumento de la posibilidad de padecer adenocarcinoma, melanoma, neoplasias esofágicas o linfoma no- Hodgking (Aljada et al., 2021). El aumento de la incidencia de estas patologías ha impulsado a la industria alimentaria a comercializar una mayor cantidad de productos sin gluten. En los Estados Unidos, este mercado supuso un beneficio de 15.500 millones de dólares en el año 2016, más del doble que 5 años antes (Niland & Cash, 2018).

Una dieta sin gluten no solo inhibe los síntomas de estas patologías, si no que en la enfermedad celíaca se ha demostrado que permite la regeneración de los daños histológicos. Los niños que llevan una dieta estricta sin gluten son capaces de regenerar en solo dos años el tejido intestinal dañado (Wahab et al., 2002). Esta capacidad de recuperación va disminuyendo con la edad. En adultos, solo un tercio de los pacientes regeneran el tejido dañado en dos años, aunque a los 5 años, dos tercios presentan una recuperación de la mucosa intestinal (Rubio- Tapia et al., 2010). Sin embargo, esta recuperación no se observa en personas a partir de los 60 años con una dieta sin gluten (Tursi et al., 2006).

La regeneración tisular inducida por una dieta libre de gluten también se ha observado en pacientes sensibles, pero no celíacos (Tursi & Brandimarte, 2003). El empleo generalizado del gluten en la industria alimentaria debido a sus propiedades funcionales, agrega complejidad a este desafío (Day et al., 2006).

En los últimos años, se ha constatado que muchas personas no consumen productos con gluten siguiendo modas dietéticas. Diversos estudios han demostrado que seguir este tipo de restricciones dietéticas puede causar problemas de salud, ya que la mayoría de productos preparados sin gluten no se pueden comparar con sus análogos con gluten en características nutricionales (Demirkesen & Ozkaya, 2022). Además, es habitual que las personas que siguen una dieta sin gluten sin supervisión médica presenten un elevado consumo de grasas saturadas y carbohidratos, lo que puede conllevar un déficit de ácidos grasos esenciales, fibra, folatos, hierro, niacina, riboflavina y tiamina (Thompson et al., 2005).

### 1.3 MARCO LEGISLATIVO INTERNACIONAL SOBRE EL CONTENIDO EN GLUTEN DE LOS ALIMENTOS

A diferencia de otros alérgenos, y siguiendo las recomendaciones incluidas en la norma del *Codex Alimentarius* 118-1979 (*Codex Alimentarius*, 2008), existe una amplia legislación internacional sobre la presencia de gluten en los alimentos.

En la Unión Europea, el Reglamento de Ejecución (UE) N° 828/2014 de la Comisión del 30 de julio sobre los requisitos para la información al consumidor sobre la ausencia o reducción de gluten

en los alimentos (Diario Oficial de la Unión Europea, 2014) establece que “la declaración «sin gluten» solamente podrá utilizarse cuando los alimentos, tal como se venden al consumidor final, no contengan más de 20 mg/kg de gluten” y, “la declaración «muy bajo en gluten» solamente podrá utilizarse cuando alimentos que consistan en trigo, centeno, cebada, avena o sus variedades híbridas, o que contengan uno o más ingredientes hechos a partir de estos cereales, que se hayan procesado específicamente para reducir su contenido de gluten, no contengan más de 100 mg/kg de gluten en el alimento tal como se vende al consumidor final”.

La Administración de Alimentos y Medicamentos de los Estados Unidos (FDA) ha definido el término “sin gluten” para su uso voluntario en alimentos que carecen de gluten. Cualquier presencia de gluten en los alimentos etiquetados como sin gluten debe ser inferior a 20 mg/kg, según el *Food Allergen Labeling and Consumer Protection Act (FALCPA, 2004)*.

En Canadá, las autoridades sanitarias consideran que los alimentos sin gluten son aquellos que contienen niveles de gluten que no exceden los 20 mg/kg como resultado de una contaminación cruzada. Con respecto a la avena se permite el uso del etiquetado “sin gluten” para la avena certificada sin gluten (Scherf & Poms, 2016).

La legislación actual de Australia y Nueva Zelanda es la más estricta. El Código de Normas Alimentarias de Australia y Nueva Zelanda, establece que para que un alimento se etiquete como “sin gluten”, no debe contener: gluten detectable; avena o sus productos derivados; cereales que contengan gluten que puedan haber sido malteados, o sus productos derivados. Para la afirmación “gluten no detectable”, el límite se estableció en 3 mg/kg (Australia and New Zealand Food Standards Code, 2018).

En México, las órdenes ejecutivas NOM-247-SSA1-2008 (Diario Oficial de la Federación, 2008) y NOM-051-SCFI/SSA1-2010 (Diario Oficial de la Federación, 2020) establecen que los alimentos que contengan granos de trigo, cebada, centeno o derivados deben etiquetarse con la mención “este producto contiene gluten”.

En Argentina, existe una legislación específica para la enfermedad celíaca que declara esta patología como enfermedad de interés nacional, regulando no solo cuestiones de seguridad alimentaria sino también aspectos sociales. Esta norma estableció un límite de 10 mg/kg de gluten para que un producto se etiquete como “sin gluten”, incluyendo un logotipo específico.

En Brasil, la Ley Federal 10674/2003, estableció que todos los alimentos deben incluir en su etiqueta y envase la indicación “contiene gluten” o “no contiene gluten”. La legislación brasileña fijó el límite de 20 mg/kg del *Codex Alimentarius* para los alimentos que no lo contienen (Diario Oficial de la Unión, 2003).

En China, la Ley de Alimentos GB/T23779 de 2009 de la Administración General de Supervisión de Calidad, Inspección y Cuarentena (AQSIQ), establece un límite máximo de 20 mg/kg para poder utilizar la mención “sin gluten” en el etiquetado. Sin embargo, esta regulación no se aplica en alimentos destinados a la importación.

En la ley japonesa, se establece que cualquier alimento que contenga proteínas alergénicas superiores a 10 mg/kg debe ser etiquetado con la mención “contiene alérgenos” (Akiyama et al., 2011).

El Reglamento (CE) Nº 852/2004 (Diario Oficial de la Unión Europea, 2004), relativo a la higiene de los productos alimenticios, establece que en la gestión de la inocuidad de los alimentos producidos por las empresas alimentarias, deben aplicarse los sistemas de autocontrol basados en los principios del Análisis de Peligros y Puntos Críticos de Control (APPCC). Durante el análisis de peligros se puede incluir como peligro químico, y aplicable a todas las etapas de la elaboración de un producto, la contaminación cruzada por presencia de ingredientes alergénicos, como el gluten, con una gravedad calificada normalmente como media (Comunidad de Madrid, 2013).

Para garantizar que aquellos alimentos etiquetados como “sin gluten” cumplen con el límite legislativo de 20 mg/kg, los fabricantes deberán tener presente un código de Buenas Prácticas de Fabricación que implique utilizar equipos específicos para la fabricación de estos productos “sin gluten”, o proceder a la limpieza exhaustiva de los equipos previo a su uso. En caso de compartir en la misma línea de producción la elaboración de productos sin gluten con otros, es conveniente proceder a la elaboración de estos en primer lugar, para evitar contaminaciones cruzadas con harinas o materias primas que contengan gluten.

Se deben evitar las corrientes de aire dentro de la zona de fabricación y envasado, y se procederá a homologar a los proveedores. Asimismo, se recomienda no utilizar harinas de trigo, cebada, centeno, avena o similares. Por último, los productos finales se mantendrán totalmente separados de cualquier otro producto que pueda suponer una fuente de contaminación (CECAM, 2004).

Además, la Asociación de Sociedades de Celíacos de Europa permite el uso del logo de la Marca Registrada “Espiga Barrada” a aquellas empresas que superen el proceso de certificación gestionado por las asociaciones nacionales de celíacos FACE (Federación de Alérgicos y Celíacos de España) en España. La certificación se basa en el análisis de los niveles de gluten de productos tal y como se venden al consumidor, realizados por laboratorios certificados por la Entidad Nacional de Acreditación (ENAC) o instituciones europeas equivalentes, para las normas 17025:2005 y 17065:2005 (UNE-EN ISO/IEC) y, un control anual del sistema de APPCC de la empresa candidata por parte de FACE (FACE, 2023).

## 1.4 METODOLOGÍAS ANALÍTICAS PARA LA DETECCIÓN DE GLUTEN EN ALIMENTOS

En los últimos años, los requerimientos legislativos, y la creciente concienciación del consumidor sobre la seguridad alimentaria han alentado a la industria alimentaria y a los grupos de investigación a desarrollar métodos más precisos para garantizar la trazabilidad de componentes potencialmente dañinos como el gluten. En la actualidad, los métodos de detección de gluten son ampliamente utilizados en la industria alimentaria y, se pueden clasificar en dos grupos principales, dependiendo de las biomoléculas detectadas: proteínas o ADN.

Los métodos basados en la detección de proteínas se pueden dividir en técnicas inmunológicas y no inmunológicas. Las técnicas inmunológicas se basan en la interacción de alta afinidad entre anticuerpos y antígenos, lo que ha llevado al desarrollo de diferentes aplicaciones:

- *ELISA (Enzyme-Linked Immunosorbent Assay)*: se trata de un método rápido, económico, versátil y robusto. La tecnología ELISA presenta una alta sensibilidad y detección óptica. Sin embargo, dependiendo de los anticuerpos empleados, existe la posibilidad de falsos negativos debido a la desnaturalización de las proteínas, y de falsos positivos debido a reacciones cruzadas con proteínas similares a la diana a detectar (Panda & Garber, 2019a).
- *Ensayos inmunocromatográficos*: su interpretación de resultados es visual y simple, siendo muy fáciles de usar por parte del operador final. Su principal debilidad es que no son cuantitativos (Ferre-López et al., 2004).
- *Western blot*: es un método muy específico y sensible, con fortalezas adicionales al aportar valores confirmatorios (pesos moleculares) y una detección eficiente de proteínas insolubles. Sin embargo, es un método que requiere más tiempo y debe ser realizado por personal cualificado (Panda & Garber, 2019b).

Los métodos no inmunológicos también se basan en la detección de proteínas, pero utilizando un mecanismo analítico completamente diferente al de la detección inmunológica. Este grupo abarca:

- *Métodos cromatográficos*: se basan en la separación y detección de péptidos y proteínas con una sensibilidad muy alta. El principal inconveniente es que requieren instrumentación compleja y costosa (Li et al., 2019).
- *Espectrometría de masas*: es un método rápido, reproducible y muy preciso, que permite la detección de péptidos específicos. Sin embargo, no es un método cuantitativo y precisa también de instrumentación compleja y costosa (Boukid et al., 2019).

El segundo grupo de métodos de detección de gluten se basa en la amplificación y detección de ADN mediante la reacción en cadena de la polimerasa (PCR). Esta técnica es altamente sensible (5 - 50 pg de ADN) y puede llevar a la identificación de especies vegetales que contengan gluten. Los métodos de PCR en tiempo real también se pueden utilizar para la detección cuantitativa (García- García et al., 2019; Martín-Fernández et al., 2015; Martín-Fernández et al., 2016). Sin embargo, los métodos de PCR son indirectos al no permitir una detección de componentes de las proteínas del gluten, sino del ADN que lo codifica.

## 1.5 MÉTODOS PARA EL DESARROLLO DE ANTICUERPOS FRENTE AL GLUTEN

### 1.5.1 Estructura y generación de anticuerpos

Los anticuerpos son moléculas secretadas por los linfocitos B con la capacidad de unirse a antígenos específicos, y también presentan algunas funciones de señalización celular. La capacidad de estas moléculas para neutralizar antígenos deriva de su extraordinaria diversidad, de forma que el cuerpo humano es capaz de producir hasta  $10^{18}$  anticuerpos distintos (Joyce et al., 2020).

Los anticuerpos son proteínas formadas por cuatro cadenas unidas por puentes disulfuro. Dos cadenas presentan menor tamaño, por lo que se denominan cadenas ligeras y las otras, de mayor tamaño, cadenas pesadas. Las cadenas pesadas tienen cinco isotipos principales (IgM, IgD, IgG, IgE, IgA), con cuatro subisotipos para la IgG (IgG1, IgG2, IgG3 e IgG4) y dos para la IgA (IgA1 e IgA2). Cada cadena pesada puede presentar cuatro dominios (IgD, IgG e IgA) o cinco (IgM e IgE). Las cadenas ligeras tienen dos dominios y se dividen en dos tipos ( $\kappa$  y  $\lambda$ ) (Barbas, 2001).

Los genes estructurales (genes VDJ) que codifican los anticuerpos humanos se encuentran en tres cromosomas diferentes: el 2 para las cadenas  $\kappa$ , el 22 para las cadenas  $\lambda$  y el 14 para la cadena pesada. Se han encontrado 50 secuencias germinales para la VH (región variable de la cadena pesada) agrupadas en 7 familias, 40 para la V $\kappa$  (región variable de la cadena  $\kappa$ ) agrupadas en 6 familias, y 30 para la V $\lambda$  (cadena  $\lambda$ ) agrupadas en 10 familias. Los tres genes D se encuentran exclusivamente en el locus de la cadena pesada. Se han descrito 7 genes J para el locus de la cadena pesada, 9 para el locus  $\lambda$  y 6 para el locus  $\kappa$ .

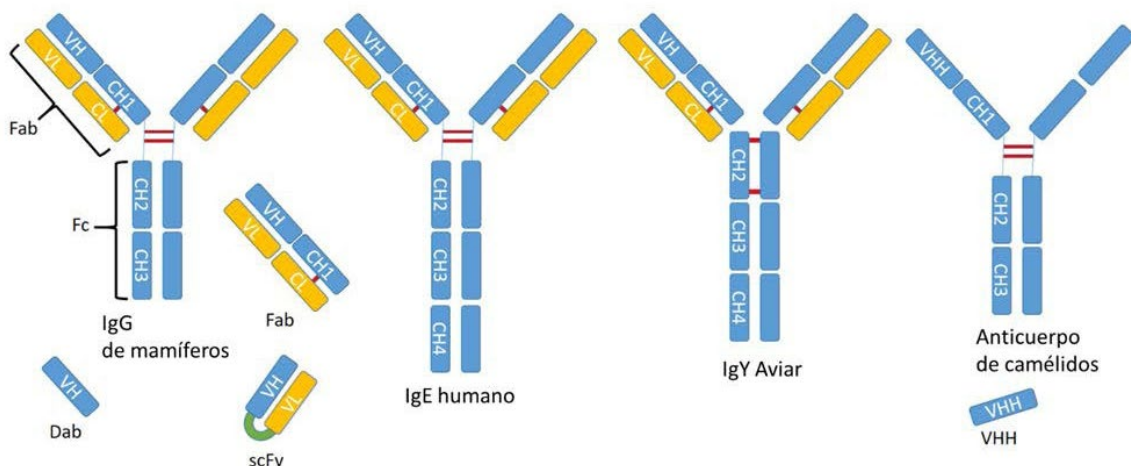
Al digerir un anticuerpo con enzimas como la papaína, se obtienen dos fragmentos: el fragmento de unión al antígeno (Fab) y el fragmento cristizable (Fc). El Fab está compuesto por 2 dominios variables (uno de la cadena pesada, VH y otro de la ligera, V $\kappa$  o V $\lambda$ ) y otros 2 constantes (uno de la cadena pesada CH1 y otro de la ligera, C $\kappa$  o C $\lambda$ ). En cambio, el Fc está formado por 4 o 6 dominios constantes de la cadena pesada.

Mediante ingeniería genética se han desarrollado muchas moléculas de fragmentos de anticuerpos que no se encuentran en la naturaleza y que permiten su aplicación en diferentes campos de la

biomedicina. Los Fab se pueden producir sin estar unidos al Fc (Frenzel et al., 2013). Los anticuerpos de fragmento variable de cadena única (scFv) son estructuras artificiales compuestas de las regiones variables de las cadenas ligeras y pesadas, unidas por un enlace flexible. Los camellos son mamíferos capaces de producir un tipo diferente de anticuerpo que carece de la cadena ligera (también conocido como anticuerpos de cadena pesada) que presentan una cadena pesada compuesta por una región de dominio único variable (VHH) y dos dominios constantes. Basándose en esta estructura, se desarrolló un nuevo tipo de fragmento recombinante que consiste solo en el dominio VHH (Figura 2) (Muyldermans, 2013). Algunos peces cartilagosos, como los tiburones, también presentan anticuerpos monocatenarios, cuya parte variable (VNAR, *Variable Domain of New Antigen Receptor*) se ha utilizado para el desarrollo de nuevas terapias (Gauhar et al., 2021).

El enorme repertorio de anticuerpos generado por los mamíferos deriva de eventos genéticos como la recombinación V(D)J, la hipermutación somática y el cambio de clase.

La recombinación V(D)J comienza cuando un complejo de dos enzimas recombinasas (RAG1-RAG2), un ión metálico divalente y las proteínas de alta movilidad (HMGB1) se unen a una única secuencia señal de recombinación (RSS) para formar un complejo de 12 o 23 nucleótidos. Posteriormente, se captura otra secuencia RSS formando un complejo emparejado. A continuación, el ADN se corta generando una horquilla donde los genes codificantes y los extremos RSS terminan en un grupo hidroxilo en 3' interactuando con un ión magnesio que estabiliza el complejo. La horquilla se libera, se abre y se une nuevamente mediante la unión de extremos no homólogos (NHEJ) (Chi et al., 2020).



**Figura 2.** Representación esquemática de la estructura de anticuerpos y moléculas derivadas. Ig (Inmunoglobulinas); Fab (fragmento de unión a antígeno); Fc (fragmento cristalizante); VH (región variable de la cadena pesada); VL (región variable de la cadena ligera); CH (región constante de la cadena pesada); CL (región constante de la cadena ligera); Dab (anticuerpo de dominio único); scFv (fragmento de anticuerpo variable de cadena única); VHH (fragmento correspondiente al dominio variable de la cadena pesada de anticuerpo de camélido). Fuente: elaboración propia.

Se han descrito dos mecanismos para la unión de extremos no homólogos (NHEJ), que ocurren principalmente mediante el mecanismo canónico, en el cual la proteína ku70/86 se une a los extremos escindidos de la doble cadena para evitar su digestión. A continuación, se reclutan y activan una quinasa de ADN que, a su vez, activa el complejo Artemis, capaz de eliminar los nucleótidos dañados o que no se pueden unir. La brecha que se crea es reparada mediante la actividad de las polimerasas de ADN  $\mu$  y  $\lambda$ , la ligasa de ADN IV y sus factores coayudantes (Pryor et al., 2015).

Se ha descrito un mecanismo alternativo que inicialmente se pensaba que era un mecanismo de respaldo de emergencia, pero se ha demostrado que está activo en células fisiológicas intactas (Truong et al., 2013). Este mecanismo alternativo implica la formación de complejos catalíticos que conducen a la generación de extremos de una sola hebra de ADN, que son extendidos por la polimerasa  $\theta$  y unidos nuevamente por el complejo de ligasas de ADN I o III (Sfeir & Symington, 2015). La recombinación ocurre de forma deletérea o inversional dependiendo de la posición de los genes V y J. Deletérea si se encuentran en la misma hebra e inversional si están en hebras opuestas (Chi et al., 2020). Los mecanismos descritos permiten 6300 recombinaciones posibles dentro de los genes VDJ para la cadena pesada (Alt & Baltimore, 1982). Durante los pasos de unión de la recombinación, ocurre otro evento de generación de diversidad consistente en la adición de nucleótidos aleatorios (segmentos N) catalizada por una transferasa en las uniones V-D y V-J (Motea & Berdis, 2010). Tras el proceso de recombinación V(D)J, las células B experimentan dos modificaciones genéticas, denominadas hipermutación somática (SHM) y recombinación de cambio de clase (CSR). Estas modificaciones ocurren principalmente en el centro germinal con el objetivo de mejorar la afinidad y modificar las características biológicas de la inmunoglobulina, lo que le permite otorgar especificidad hacia el antígeno diana. La SHM, la CSR y la conversión génica comienzan con la desaminación de la citidina catalizada por una enzima expresada exclusivamente en células B activadas (Muramatsu et al., 2000). La desaminación provoca la activación de mecanismos de reparación del ADN, como la reparación de bases desapareadas (MMR, *MisMatch Repair*) (Jiricny, 2013) y la reparación de bases escindidas (BER, *Base Excision Repair*) (Krokan & Bjoras, 2013).

La hipermutación somática (SHM) es el resultado de la reparación del ADN en las posiciones desaminadas mediante un proceso de reparación por MMR y BER, ambos propensos a errores para generar más diversidad. Esto ocurre debido a la intervención de polimerasas de baja fidelidad, como las de la familia  $\eta$  o  $\gamma$ , en un evento de recombinación caracterizado por la creación de rupturas de doble cadena en las regiones de cambio (*switch regions* o S), una región altamente repetitiva ubicada antes de cada gen de la región constante (CH). Este proceso finaliza mediante la unión de extremos no homólogos (NHEJ) entre las regiones S donante y aceptora, lo que conduce a la sustitución del conjunto expresado de exones CH. La CSR permite que una región variable reordenada y mutada se combine con diferentes estructuras que le proporcionan funciones efectoras diversas (Matthews et al., 2014).

### 1.5.2 Anticuerpos policlonales

A finales del siglo XIX, Von Behring y Kitasato descubrieron que el plasma sanguíneo producía sustancias capaces de neutralizar toxinas como la diftérica o la tetánica. Unos años más tarde, Ehrlich propuso que estas sustancias neutralizantes respondían a las leyes de la química y que las células sanguíneas podían producir unas sustancias que denominó “cadenas laterales” que se unían específicamente a las toxinas (García Merino, 2011). Estos descubrimientos encontraron rápidamente aplicaciones biomédicas.

Los primeros inmunoensayos para la detección de gluten en los alimentos emplearon anticuerpos policlonales obtenidos del suero de animales inmunizados. Aunque la gran mayoría de los anticuerpos policlonales se aíslan de mamíferos (lagomorfos, roedores, cabras y caballos), también se han obtenido en embriones de pollo (IgY). Las IgY frente al gluten se han utilizado no solo como reactivo para inmunoensayos, sino también en la terapia del daño intestinal de la enfermedad celíaca (Gujral et al., 2015). Algunos ejemplos de anticuerpos policlonales disponibles comercialmente, junto con la información sobre la forma en que se han obtenido, se muestran en la Tabla 2.

### 1.5.3 Anticuerpos monoclonales

En paralelo con la investigación de la función del sistema inmune, en los años 40 del siglo XX se descubrió el origen celular de los anticuerpos y se identificaron las cadenas ligeras y pesadas. Jerne desarrolló la teoría de que los anticuerpos preexistían en el organismo y que se seleccionaban los más adecuados en función del antígeno (Jerne, 1955). Años después, esta teoría se confirmó y se amplió por Burnet y Talmage, que propusieron la selección clonal de linfocitos B dirigidos frente a antígenos concretos (Burnet, 1976).

Estas teorías, junto con los avances en el cultivo de células de mamífero, permitieron que en 1975 Köhler y Milstein descubrieran la tecnología para la generación de hibridomas secretores de anticuerpos monoclonales. Su objetivo era estudiar la biología básica de los linfocitos B, cuyo cultivo *in vitro* era muy complejo. Para ello, fusionaron los linfocitos B a células de mieloma y cultivaron estas fusiones en un medio con timidina e hipoxantina, que solo permite el crecimiento de las fusiones buscadas (Figura 3). Los hibridomas pueden crecer indefinidamente debido a las características de inmortalidad de las células cancerígenas. Dado que la fusión proviene de un solo linfocito B, al anticuerpo se le denomina monoclonal en contraposición con los anticuerpos policlonales del plasma sanguíneo que surgen de una expansión clonal de varios linfocitos diferentes (Köhler & Milstein, 1975).

**Tabla 2.** Anticuerpos policlonales comerciales obtenidos frente al gluten.

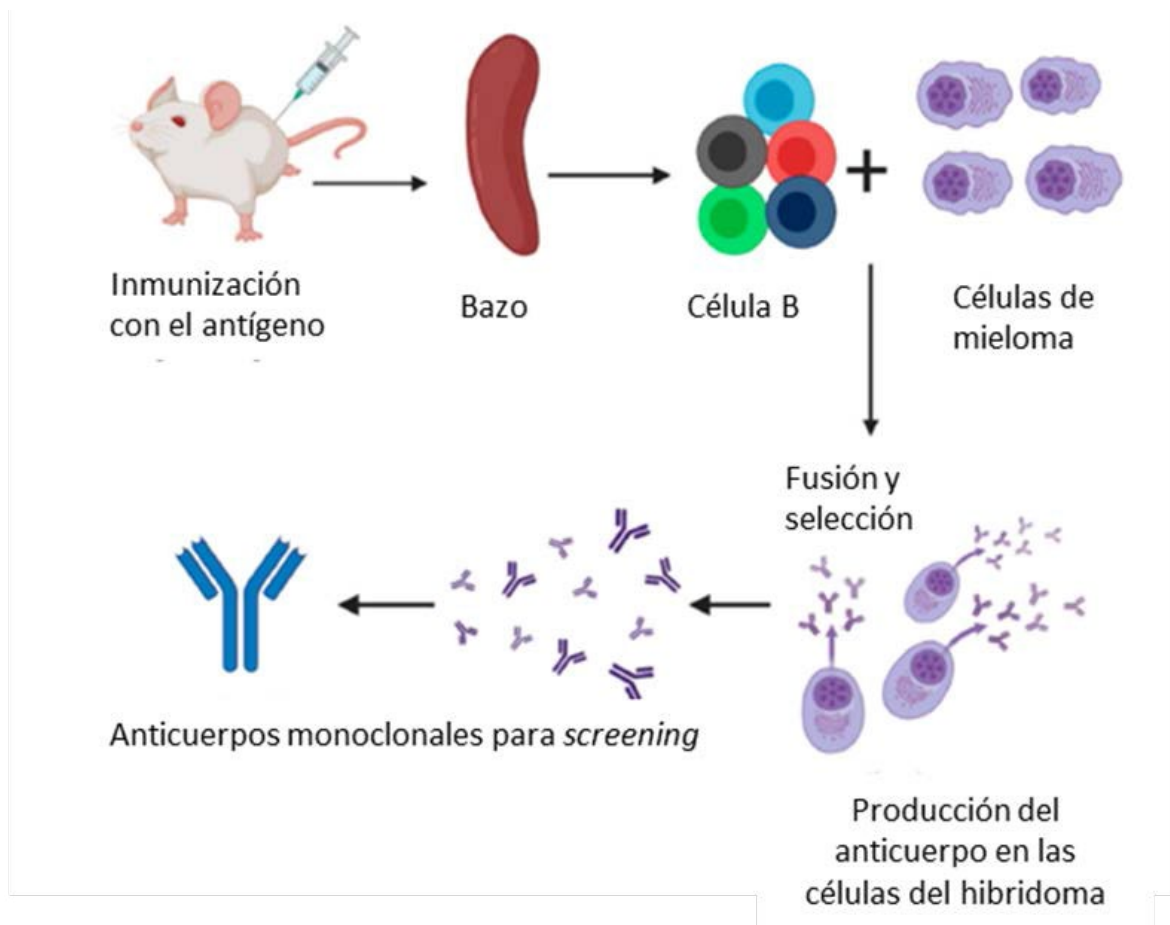
Anticuerpo	Empresa	Especie	Isotipo	Generado frente a
PAB29118	Abnova	Pollo	IgY	Extracto de harina de trigo
MBS617177	MyBioSource	Conejo	IgG	Gluten de trigo
MBS838918	MyBioSource	Conejo	IgG	Gluten de trigo
MBS625849	MyBioSource	Pollo	IgY	Gluten de trigo
LS-C66756	LifeSpanBiosciences	Conejo	IgG	Gluten de trigo
LS-C129350	LifeSpanBiosciences	Pollo	IgY	Gluten de trigo
LS-C750830	LifeSpanBiosciences	Pollo	IgY	Gluten de trigo
G8138-01	USBiological	Conejo	IgG	Gluten de trigo
G8138-02	USBiological	Pollo	IgY	Extracto de harina de trigo
AS09 571	Agrisera	Pollo	IgY	Extracto de harina de trigo
PA5-97536	Invitrogen	Conejo	IgG	Gliadina nativa de trigo

Fuente: elaboración propia

La tecnología de producción de anticuerpos monoclonales supuso una innovación revolucionaria en muchos campos biomédicos. Esta tecnología también se ha aplicado a la detección de gluten en alimentos. La principal fortaleza de los anticuerpos monoclonales en comparación con los policlonales es la uniformidad entre lotes (Bradbury et al., 2018). De hecho, el método de referencia para la detección de gluten en alimentos se basa actualmente en un anticuerpo monoclonal (Scherf & Poms, 2016).

Uno de los primeros anticuerpos monoclonales desarrollados frente al gluten fue desarrollado por Skerit y Underwood (1986). Estos autores emplearon una extracción de proteínas (prolaminas) de harina cruda de trigo blanco para la inmunización de ratones BALB/c.

Las fracciones de gliadina se separaron mediante cromatografía de intercambio iónico sobre sulfocelulosa etílica para obtener las fracciones  $\alpha\beta$  y  $\beta\gamma$  de la gliadina. También se obtuvo una fracción de glutelina reducida y alquilada por precipitación con etanol, dodecil sulfato sódico (SDS) y 2-mercaptoetanol. Los linfocitos esplénicos de los ratones BALB/c inmunizados con estas fracciones de gliadina y glutelina, se fusionaron con células de mieloma SP 2/0. Los hibridomas secretores de anticuerpos de elevada afinidad se seleccionaron utilizando métodos habituales de aislamiento (Galfre et al., 1977).



**Figura 3.** Representación esquemática de la producción de anticuerpos monoclonales. Fuente: elaboración propia.

Posteriormente, Skerritt y Hill (1990) obtuvieron el anticuerpo monoclonal IgG1 401.21, generado frente a  $\omega$ -gliadinas termoestables, que reconoce los epítomos PQQPFPQE y PQQPPFEE. Este anticuerpo monoclonal reacciona frente a las  $\omega$ -gliadinas y las prolaminas de centeno y cebada, así como con las subunidades de glutelina de alto peso molecular. Un ELISA sándwich basado en el anticuerpo 401.21 fue adoptado como Método Oficial 991.19 por la Asociación Internacional de Químicos Analíticos Oficiales (AOACI) (Scherf & Poms, 2016). Aunque este anticuerpo ya no se considera el de referencia internacional, sigue disponible en algunos kits como el ensayo Allertek™ Gluten ELISA (ELISA Technologies Inc.) o Veratox® para gliadina (Neogen®).

Sorell et al. (1998) desarrollaron un ELISA sándwich para el análisis de gluten en alimentos utilizando un cóctel de anticuerpos monoclonales. Se inmunizaron ratones hembra BALB/c con extractos etanólicos de trigo, centeno y avena. Los esplenocitos de los animales inmunizados se fusionaron con células de mieloma P3/X63-Ag.653. Los híbridos seleccionados se cultivaron induciendo ascitis en ratones BALB/c tratados con pristano, y los anticuerpos se purificaron del líquido ascítico por cromatografía de afinidad en una columna de proteína A. Posteriormente, se caracterizaron siete anticuerpos monoclonales (cinco frente a centeno, denominados R1 a R5,

uno frente a avena y otro frente a gliadina conocido como 13B4). La mayoría de los anticuerpos obtenidos mostraron un amplio espectro de reactividad cruzada con gliadinas, hordeínas y secalinas. Algunos anticuerpos también producían reacciones cruzadas con aveninas pero no reconocieron a las zeínas. El cocktail de R5 y 13B4 como anticuerpo de captura y R3 conjugado a la peroxidasa de rábano (HRP) como anticuerpo de detección, permitió el reconocimiento de gliadinas, secalinas y hordeínas en el intervalo de 3 a 200 ng/mL, consiguiendo mejores resultados en muestras de alimentos que otros test comerciales disponibles en aquel momento, y en muchos casos, superando los resultados obtenidos con técnicas de espectrometría de masas.

Posteriormente, Valdés et al. (2003) desarrollaron un ELISA sándwich usando un único anticuerpo monoclonal (R5) como anticuerpo de captura y como anticuerpo de detección conjugado con HRP. Para ello prepararon un patrón de gliadina mediante extracción con etanol para evaluar el sistema y, además, desarrollaron un cóctel acuoso de extracción que contenía agentes reductores. El ELISA R5 pudo identificar gliadinas, hordeínas y secalinas con una sensibilidad de los ensayos de 0,78, 0,39 y 0,39 ng/mL, respectivamente para cada prolamina. El límite de detección fue de 1,5 ng/mL de gliadinas (1,56 mg/kg de gliadinas y 3,2 mg/kg de gluten), inferior al umbral de los tests comercializados en ese momento, y con una buena reproducibilidad (8,7 %) y repetibilidad (7,7 %). Estos resultados situaron a este test como el mejor en el campo, por lo que el Grupo de Trabajo sobre Análisis y Toxicidad de Prolaminas (WGPAT) propuso que se incluyera en el *Codex Alimentarius*.

El ELISA sándwich R5 (Valdés et al., 2003), junto con el cóctel de extracción (García et al., 2005), fue validado (Köhler et al., 2013; Méndez et al., 2005) y adoptado como Método de referencia 38-50.01 aprobado por la AACCI (Asociación Internacional Americana de Químicos de Cereales) para la detección de gluten en harina de maíz y productos a base de maíz. También fue reconocido por la AOACI como Método Oficial 2012.01 para la detección de gluten en productos a base de arroz y maíz. Este método también es considerado por el *Codex Alimentarius* como un método de tipo 1 para el análisis de gluten intacto en matrices a base de maíz. También se desarrolló un ELISA competitivo R5 para la determinación de gluten parcialmente hidrolizado y se aceptó como Método 38-55.01 aprobado por la AACCI para la detección de gluten en productos a base de cereales fermentados (Koerner et al., 2013).

El método de ELISA competitivo está protegido por la patente WO2006051145A1, y el cóctel de extracción por la patente WO2007104825A1. Este ELISA y el método de extracción con el cóctel se han convertido en la metodología más utilizada para la detección de gluten en alimentos y ha sido comercializado por muchas empresas en todo el mundo. Aunque los métodos basados en el anticuerpo R5 son los predominantes y oficiales, se han desarrollado otras alternativas para cubrir la creciente demanda del mercado.

Teniendo en cuenta que los anticuerpos disponibles en el mercado reconocían péptidos del gluten, pero no específicamente relacionados con péptidos responsables de la enfermedad celíaca, Morón et al. (2008) propusieron una nueva estrategia. Identificaron los residuos 57 a 89 de la  $\alpha$ -2 gliadina, que comprenden el péptido de 33 aminoácidos LQLQPFQPQLPYPQPQLPYPQPQLPYPQPQPF, como un epítipo importante del gluten, altamente antigénico. Además, de este 33-mer, se han identificado hasta la fecha, varios péptidos antigénicos con regiones ricas en prolina (Shan et al., 2002). Para conseguir una correcta inmunización con este péptido, debido a su bajo peso molecular, se requería la fusión a una proteína transportadora para estimular la respuesta inmunitaria del hospedador. Se utilizaron con este propósito dos moléculas transportadoras: la proteína recombinante de choque térmico de *Trypanosoma cruzi* HSP70 y un fragmento específico de la misma (T-HSP70). Mediante la técnica de ELISA, se seleccionaron los hibridomas H-G12 (de linfocitos B de ratones inoculados con 33-mer-T-HSP70) y H-A1 (de los linfocitos B de ratones inmunizados con 33-mer-X2-HSP70) por su especificidad y afinidad de unión a este péptido. Los anticuerpos monoclonales purificados se analizaron frente a gliadinas (extraídas de muestras comerciales, materiales de referencia y digestiones con pepsina de diversos productos) y varias prolaminas extraídas con etanol (trigo, cebada, centeno, avena, maíz y arroz). Además, se diseñó un ELISA tipo sándwich usando el anticuerpo A1 como captura y el anticuerpo G12 conjugado con HRP como anticuerpo de detección, obteniéndose un límite de detección para trigo, cebada y centeno inferior a 1 mg/kg. Además, se diseñó un ELISA competitivo basado en el anticuerpo G12 conjugado con HRP para la detección del péptido inmunogénico en alimentos hidrolizados (con un límite de detección inferior a 0,5 mg/kg de gliadina). Las secuencias de los epítopos reconocidos por el anticuerpo fueron hexaméricas (QPQLPY) para el G12 y heptaméricas (QLPYPQP) para el A1. Aunque la afinidad del G12 para el 33-mer fue superior a la del A1, este último presentó una mayor capacidad de detección para el gluten (Morón et al., 2008). Este trabajo estableció un método válido en cuanto a precisión, exactitud y reproducibilidad para la cuantificación de fracciones del gluten que no pueden tolerar los pacientes con enfermedad celíaca. Además, se obtuvo una sensibilidad ligeramente mayor en comparación con otros anticuerpos comerciales como el R5. AgraQuant® (Romerlabs) es un método de ELISA tipo sándwich basado en el anticuerpo monoclonal G12 aprobado por la AACCI (Método 38-52.01) y por la AOAC (Método Oficial de Análisis, OMA 2014.03) como un método certificado para la detección de gluten en harina de arroz y otros productos a base de arroz.

En la actualidad, debido a que muchos estudios han demostrado la complejidad de la enfermedad celíaca (Anderson et al., 2000; Arentz-Hansen et al., 2000; Shan et al., 2002; Stern et al., 2001), se ha producido un cambio de paradigma en los objetivos de la detección de gluten en alimentos. La generación de anticuerpos monoclonales no solo debe lograr un límite de detección muy bajo de gluten, sino también apuntar a la determinación de péptidos que inducen enfermedades. Con este objetivo Mitea et al. (2008) inmunizaron ratones BALB/c con péptidos sintéticos correspondientes

a epítomos estimuladores de células T, que se acoplaron a la toxina tetánica para inducir una mayor respuesta inmunitaria en el animal de experimentación. Entre los péptidos objetivo se encontraron componentes de gliadinas: glia- $\alpha$ 9 (QFPQPQLPYP), la glia- $\alpha$ 20 (PFRPQQPYPQP), glia- $\gamma$ 1 (PQSFPPQQRPFIQPSL), LMW Glt-156 (PPFSQQQQSPFS) y de glutelinas de alto peso molecular HMW-Glt (PGQGQ(Q/P)GYYPST(L/Q)QQPQGQQGYYPSTPQQ(P/S)). Debido a que muchas proteínas del gluten comparten un alto grado de homología, estos autores intentaron demostrar si los anticuerpos monoclonales obtenidos reaccionaban específicamente solo con el péptido utilizado para la inmunización o detectaban las otras secuencias estimuladoras de células T. En este trabajo pusieron de manifiesto que el anticuerpo específico para la glia- $\alpha$ 20 también reaccionaba con los péptidos glia- $\alpha$ 9 y glia- $\gamma$ 1. Además, los anticuerpos monoclonales obtenidos frente a la glia- $\alpha$ 9, glia- $\alpha$ 20, Glt-156 y HMW-Glt reaccionaron con péptidos de gluten que se forman naturalmente durante la digestión, como resultado de la actividad de la pepsina y tripsina. Además, excepto los anticuerpos específicos de LMW Glt-156, también detectaron proteínas de almacenamiento en cebada y centeno, mientras que los anticuerpos específicos frente a la glia- $\gamma$ 1 también reconocieron proteínas de avena. Finalmente, en comparación con el kit Ridascreen® Gliadin, el ELISA desarrollado para el epítomo glia- $\alpha$ 9 detectó concentraciones de gluten más altas en todas las muestras de alimentos analizadas. La empresa EuroProxima comercializó un ELISA competitivo basado en el anticuerpo monoclonal antiglia- $\alpha$ 20, llamado Gluten-Tec®, que fue validado por un estudio interlaboratorial (Mujico et al., 2019), presentando un límite de cuantificación (LOQ) inferior al de los métodos basados en el anticuerpo R5 disponibles en el momento (3,6 frente a 5 mg/kg de gluten, respectivamente). Posteriormente se obtuvieron otros anticuerpos monoclonales frente a los mismos y diferentes epítomos de células T (glia- $\gamma$ 1, Glt-156, una variante de HMW-gly y ocho péptidos) y se seleccionaron cinco anticuerpos (uno anti- $\alpha$ 1-gliadina, uno anti- $\gamma$ 1-gliadina, dos anti-LMW y uno anti-HMW). Este método está protegido por la patente WO2006004394A2.

Tras la aparición, a principios de los años 2000, de alergias graves, relacionadas con la ingesta de alimentos que contienen incluso pequeñas cantidades de proteínas de trigo hidrolizadas, surgieron nuevos desafíos. Las principales patologías relacionadas fueron la anafilaxia inducida por ejercicio dependiente de trigo (WDEIA, *Wheat Dependent Exercise-Induced Anaphylaxis*), anafilaxia y urticaria. Se identificó que las proteínas de trigo hidrolizadas eran el ingrediente desencadenante. Concretamente, las condiciones de alta temperatura y bajo pH que se producen durante el procesamiento industrial provocan que las glutaminas se conviertan en ácido glutámico por reacciones de desaminación. La desaminación conduce a la aparición de nuevos epítomos para las IgE, y se encontró que los sueros de pacientes predispuestos presentaban una mayor reactividad frente a las proteínas de gluten desaminadas, especialmente en  $\alpha$ ,  $\gamma$ ,  $\omega$ 2 y  $\omega$ 5-gliadina (Tranquet et al., 2020).

Con relación a este nuevo tipo de alergenicidad, se utilizó un nuevo epítipo inmunodominante (QPEEPFPE, derivado de la desaminación de QPQQPFPQ) para producir un anticuerpo monoclonal a partir de ratones inmunizados con un péptido que incluye dicho epítipo (Tranquet et al., 2015). Posteriormente, las secuencias que codifican para los dominios VH y VL se clonaron en plásmidos de expresión modificados para expresar cadenas pesadas de IgE, generando una IgE quimérica recombinante (Tranquet et al., 2017). La expresión recombinante se realizó en células de mamíferos (HEK293). Los sueros de pacientes alérgicos y el anticuerpo recombinante se analizaron por ELISA frente a fracciones de gluten desaminado, y en ensayos funcionales de basófilos. Se demostró que las proteínas de trigo hidrolizadas en medio ácido presentaban mayores niveles de desaminación, y una mayor capacidad de unión a IgE, y de activación de basófilos. Además, el anticuerpo recombinante estimulaba la degranulación de basófilos en presencia de proteínas de trigo desaminadas, imitando las propiedades de las IgE de los pacientes. Este trabajo es un ejemplo de la transformación de un anticuerpo monoclonal clásico en uno recombinante.

Los anticuerpos monoclonales secretados por hibridomas han demostrado ser una buena herramienta para el diseño y mejora de sistemas de detección de gluten. Algunos de estos desarrollos han sido protegidos por patentes (Tabla 3). Hoy en día, la mayoría de los tests comerciales disponibles para la detección de gluten se basan en los anticuerpos monoclonales 401.21, R5, G12 y anti- $\alpha$ 20 (Scherf & Poms, 2016).

#### 1.5.4 Anticuerpos recombinantes y evolución dirigida

Aunque la introducción de la tecnología de hibridomas supuso un gran avance para el desarrollo de anticuerpos en muchos campos biomédicos, presentaba algunas limitaciones. Debido al uso generalizado de estos anticuerpos, se empezaron a encontrar problemas de inespecificidad que iban más allá de un reconocimiento cruzado de epítipos similares en antígenos distintos (mimotopos). Este problema derivaba de la coexpresión en los hibridomas de cadenas de anticuerpos productivas pero no específicas para el antígeno, que alteran la especificidad y la afinidad del anticuerpo monoclonal correspondiente (Bradbury et al., 2018). La coexpresión de cadenas deriva de la inestabilidad genómica de los hibridomas, fusiones generadas *in vitro*, tetraploides, que expresan hasta 50 cromosomas en ciertas fases de crecimiento (Kontsek et al., 1988).

Uno de los primeros clones utilizados para generar las líneas de mieloma para fusionar presentaba la secreción de una cadena  $\lambda$  adicional (Ding et al., 2010).

**Tabla 3.** Patentes registradas basadas en anticuerpos monoclonales para la detección de gluten.

Patente Nº	Resumen	Solicitante
WO2006004394A2	Un método para la selección de materias primas, ingredientes semielaborados y productos alimentarios destinados a ser utilizados en la dieta sin gluten, basado en anticuerpos desarrollados frente a péptidos estimuladores de células T.	Academisch Ziekenhuis Leiden (Países Bajos)
WO2006051145A1	ELISA competitivo para la detección de hidrolizados de gluten, basado en el anticuerpo monoclonal R5.	Consejo Superior de Investigaciones Científicas (España)
WO2007104825A1	Método para la extracción de gluten de alimentos procesados (por calor) y sin procesar, basado en el uso de detergentes iónicos y no iónicos, como paso previo a los ensayos ELISA.	Consejo Superior de Investigaciones Científicas (España)
WO2014132204A1	Anticuerpo monoclonal capaz de unirse a proteínas de gluten desaminadas (relacionados con la patogénesis de la enfermedad celíaca) y que no tienen reacción cruzada con las proteínas de gluten no desaminadas.	Institut National de la Recherche Agronomique (Francia)
WO2015164615A1	Anticuerpos monoclonales y fragmentos aislados que se unen a 11 péptidos que se pueden encontrar en las proteínas de gluten.	University of Chicago (EEUU) y University of Oslo (Noruega)
WO2018071718A1	Anticuerpos, fragmentos o polipéptidos para la detección de gliadina: secuencias variables de cadena pesada y cadena ligera, y secuencias asociadas de regiones determinantes de la complementariedad (CDR).	Nima Labs Inc. (EEUU)
WO2019154559A1	Métodos de inmunoensayo para la cuantificación del contenido total de gluten en muestras de alimentos	R. Biopharm AG. (Alemania)
ES2392412A1	Soluciones para la extracción y solubilización de gluten, compuestas por arginina y etanol.	Biomedal S.L (España)
GB2207921A	Línea de células híbridas ATCC HB9798 que produce anticuerpos monoclonales dirigidos frente a la gliadina del trigo y proteínas relacionadas con el centeno y cebada.	Commonwealth Scientific and Industrial Research Organization (Reino Unido)
CN101698832A	Anticuerpo monoclonal anti-gliadina y la línea de hibridomas para producirlo.	Quingdao Biomade Technology Company Ltd. (China)
CN107860918A	Método de producción de una tira para prueba de inmunocromatografía con inmovilización de anticuerpos por oro coloidal para la detección de gluten en alimentos.	Biofront Technology Company Ltd. (China)

Fuente: elaboración propia

También se ha notificado la aparición de genes de cadenas  $\kappa$  no productivos que no alteran el anticuerpo final pero dificultan su expresión (Carroll et al., 1988). La secreción por parte del hibridoma de una única cadena pesada con varias cadenas ligeras productivas pueden generar diferentes anticuerpos con menor afinidad. Este hecho puede producirse en el caso de que una sola cadena ligera correcta se combine con la pesada generando un solo parátopo, o anticuerpo monovalente también puede ocurrir que la cadena pesada se combine con dos cadenas ligeras erróneas, modificando completamente la especificidad del anticuerpo (Bradbury et al., 2018). Este último caso es el más común, pero también se puede dar la expresión de cadenas pesadas y ligeras erróneas en el mismo hibridoma, teniendo en cuenta que con la expresión de dos cadenas pesadas (una correcta y otra no) y dos ligeras (mismo caso) se podrían generar hasta 10 anticuerpos diferentes con cuatro parátomos diferentes (Schaefer et al., 2016). Asimismo, el cultivo de células de hibridoma a escala industrial es más complejo y costoso que otros sistemas biológicos, entre los que se incluyen: expresión *in vitro* de proteínas sin células, biofactorías bacterianas como *E. coli* modificadas para diferentes usos, levaduras, células de insecto o de mamífero, todas ellas convenientemente transfectadas para expresar el anticuerpo de interés (Trimmer, 2020).

Los avances ocurridos en la segunda mitad del siglo XX en la expresión de proteínas heterólogas e ingeniería genética, y el profundo conocimiento adquirido sobre la genética y la generación de anticuerpos, permitieron el desarrollo de nuevas metodologías para su desarrollo. En este contexto se generó el concepto de evolución molecular dirigida que implica la creación de anticuerpos adaptados y especialmente diseñados para fines humanos (Arnold, 2019; Smith, 2019). La poderosa maquinaria de generación y selección de anticuerpos de los animales puede ser replicada y guiada para obtener una aplicación específica, como la detección de gluten. El proceso natural del sistema inmunológico para la producción de anticuerpos puede resumirse en cinco pasos: (a) la reorganización de genes que codifican las regiones variables (Hozumi & Tonegawa, 1976); (b) la presentación de anticuerpos en la superficie de una célula B; (c) la selección impulsada por el antígeno; (d) la secreción de anticuerpos solubles por una célula plasmática y (e) la maduración por afinidad (Winter, 2019).

La evolución dirigida de los anticuerpos necesita superar dos condiciones limitantes: la generación de suficiente diversidad mediante la construcción de repertorios de genes codificantes para cadenas de anticuerpos y el desarrollo de un sistema adecuado para la selección y amplificación. Los repertorios de anticuerpos se pueden clasificar en cuatro grupos principales (Hust, 2004):

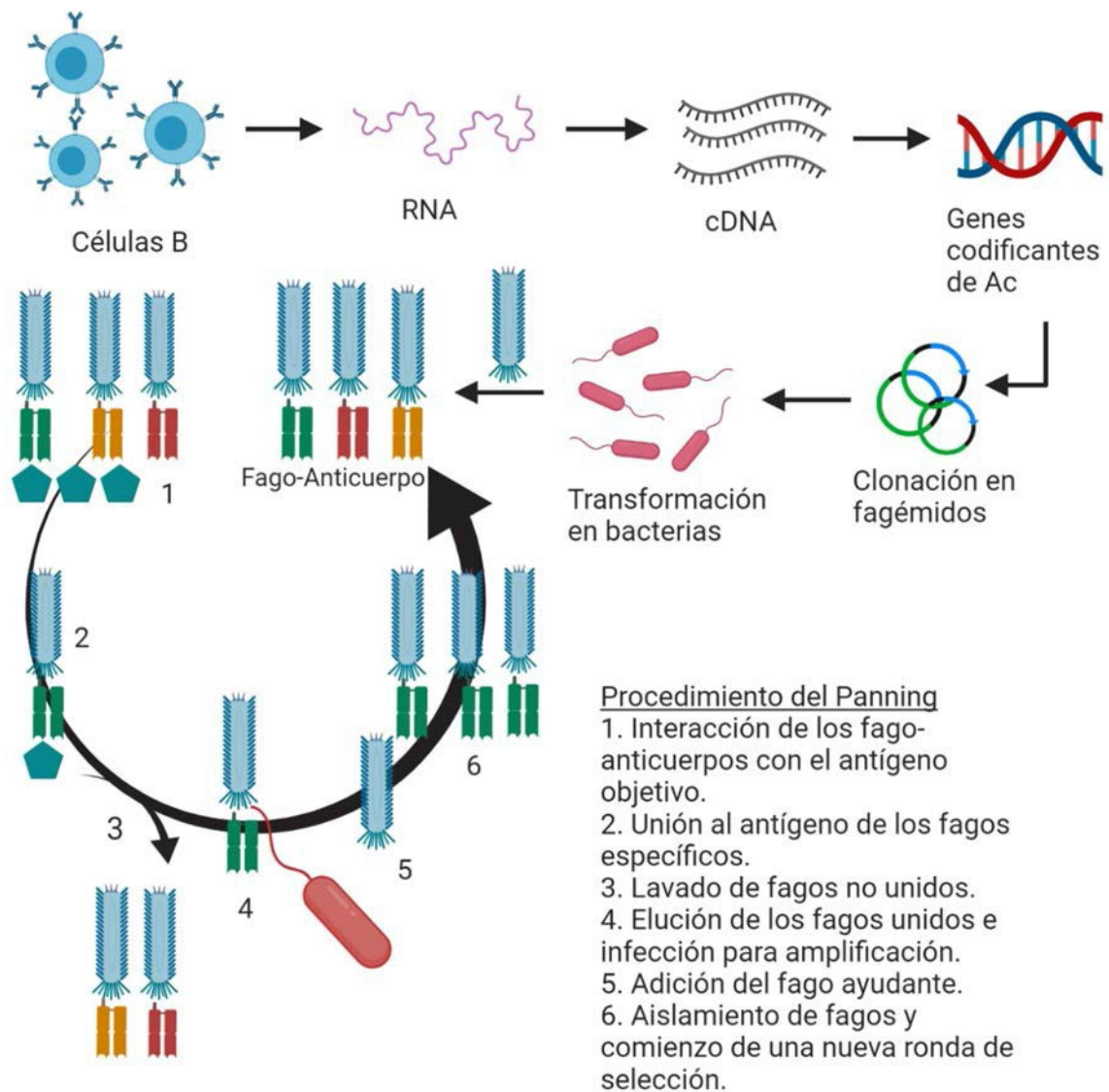
- **Inmunes:** contruidos a partir de amplificaciones de genes variables aislados de células plasmáticas secretoras de inmunoglobulinas de donantes inmunizados (Clackson et al., 1991).
- **Naïves:** basados en amplificaciones de genes aislados de células plasmáticas secretoras de inmunoglobulinas de donantes no inmunizados (De Haard et al., 1999).

- Semisintéticos: derivados de genes V no reorganizados procedentes de pre-células B (células germinales) o de un solo marco de un anticuerpo con regiones determinantes de complementariedad (CDR) aleatorizados genéticamente (Pini et al., 1998).
- Sintéticos: basados en un marco de anticuerpo humano en el que se integran casetes de CDR aleatorios (Wagner et al., 2018).

Los repertorios inmunes suelen construirse después del aislamiento de ARNm del tipo celular deseado y la preparación de ADNc para utilizarlo como molde en la amplificación de los fragmentos que codifican para las cadenas de anticuerpos (Hust, 2004). Los repertorios *naïve*, semisintéticos y sintéticos se consideran repertorios “*single-pot*”, lo que significa que se pueden utilizar para seleccionar anticuerpos frente a (casi) todos los antígenos que se pueden presentar (Böldicke, 2017). La afinidad de los anticuerpos desarrollados a partir de este tipo de repertorios depende directamente del tamaño del repertorio, RZ (número de anticuerpos diferentes en el repertorio). Para lograr una afinidad en el rango micromolar ( $\mu\text{M}$ ), se requiere un  $\text{RZ} = 10^7$ , y se necesitan valores de RZ de  $10^{10}$  para una afinidad en el rango nanomolar (nM). Sin embargo, los repertorios inmunes deben diseñarse y construirse específicamente para cada antígeno individual o grupo de antígenos relacionados. La afinidad del anticuerpo resultante está determinada por su inmunogenicidad, siendo posible alcanzar un rango de afinidad nanomolar con un antígeno muy inmunogénico incluso con un repertorio reducido (Hoogenboom, 1997).

Los genes codificantes de anticuerpos que conforman un repertorio deben clonarse en un vector apropiado para desarrollar un sistema adecuado de selección y amplificación. Una vez que se construye un repertorio con suficiente diversidad, se realiza un proceso que consiste en rondas sucesivas de selección por afinidad y amplificación llamado “*panning*” o “*biopanning*”. Para realizar el *panning*, es necesario expresar todos los anticuerpos o sus fragmentos incluidos en el repertorio (Leow et al., 2017). Se ha establecido que si antes de la selección, hay de 1 a 10 anticuerpos afines a la diana de interés en  $10^7$  clones, debería haber de 1 a 10 anticuerpos afines en 10 clones después de tres a cinco rondas de *panning*. Estos procesos permiten el enriquecimiento de un conjunto de anticuerpos de alta afinidad del repertorio, y deben presentar el emparejamiento de genotipos (un ácido nucleico codificante) con fenotipos (los anticuerpos expresados).

Diferentes péptidos, dominios proteicos y fragmentos de anticuerpos pueden ser expresados de diversas maneras para su selección y amplificación: a) unidos a fagos (Smith, 2019); b) mediante la expresión *in vitro* de ARNm quedando unidos a ribosomas (Plückthun, 2012) y c) por expresión indirecta de ADNc (Yamaguchi et al., 2009) y expresados en la superficie de células (Lee et al., 2003). La expresión de repertorios de fragmentos de anticuerpos en fagos es el método más ampliamente utilizado para la obtención de anticuerpos recombinantes (Figura 4).



**Figura 4.** Construcción de un repertorio inmune y desarrollo del proceso de evolución molecular dirigida mediante la tecnología de *phage display*. Fuente: elaboración propia.

La metodología de expresión de fago-anticuerpos (*phage display*) fue diseñada por Smith et al. (1985), quienes obtuvieron la prueba de concepto de la tecnología mediante la introducción del gen que codifica para la enzima bacteriana *EcoRI* junto al gen III de un bacteriófago. Así lograron la expresión heteróloga de *EcoRI* fusionada a la proteína III del fago. Posteriormente, se diseñó un vector llamado fUSE5 para mejorar la introducción de ADN exógeno en los fagos filamentosos y para la expresión de proteínas (Parmley & Smith, 1988). El siguiente paso fue el desarrollo del procedimiento de selección por afinidad a través de: (a) la inmovilización de un antígeno como selector; (b) la adición de viriones que expresan un repertorio de anticuerpos o sus fragmentos; (c) el lavado de viriones no unidos; (d) la liberación de viriones unidos al selector y (e) la amplificación de los viriones liberados mediante la infección de una bacteria compatible (Smith, 2019).

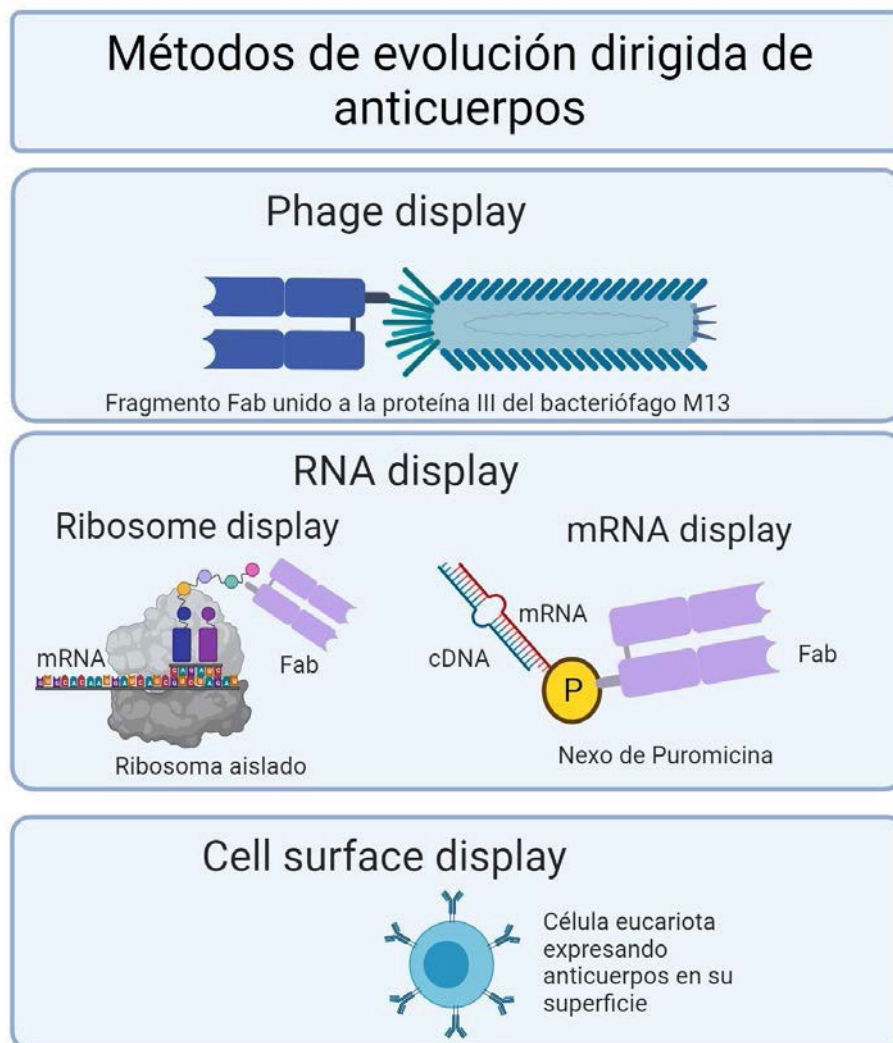
La primera aplicación de la técnica de *phage display* fue el mapeo de los epítomos a los que se une un determinado anticuerpo (de la Cruz et al., 1988). Se han producido grandes avances en este sistema con diferentes aplicaciones, aunque la base metodológica no ha cambiado significativamente desde entonces.

Se han desarrollado nuevos vectores con algunas características comunes: un doble origen de replicación (*f1* o similar, para la expresión en fagos y un *ori* de *Escherichia coli*), un marcador de selección (generalmente un gen de resistencia a un antibiótico) y genes codificantes de los ligandos (generalmente péptidos, dominios de proteínas o fragmentos de anticuerpos) fusionados a una proteína de los fagos (generalmente proteínas de la cápsida) (Smith, 1993).

Los vectores pueden clasificarse en varios tipos según su diseño. Si el ligando está fusionado al gen III de los fagos, se expresará fusionado a la proteína III (de la que se expresan hasta cinco copias), constituyendo el vector un fagémido de tipo 3. Sin embargo, si está fusionado al gen VIII, se expresará fusionado a la proteína VIII (hasta 2400 copias), constituyendo un fagémido de tipo 8. Los tipos 33 y 88 siguen el mismo concepto, pero tienen dos copias de los genes III o VIII. Los tipos 33 y 88 también se basan en fagémidos pero las partículas virales solo se forman en células que llevan el genoma del fagémido cuando están infectadas con el virus auxiliar (*helper phage*) (Armstrong et al., 1996). Aunque los vectores basados en pIII y pVIII son los más utilizados, ha habido algunos intentos de clonar ligandos en los genes VII y IX, que codifican para las proteínas menores de la cápsida fágica (Makowski, 1993).

Debido a la variabilidad que el cultivo de células puede introducir en la expresión, se han desarrollado nuevos métodos libres de células para la exposición de repertorios de anticuerpos: *ribosome display* (exposición en ribosomas) y *ARNm (o ADNc) display* (exposición de ARNm o ADNc) (Figura 5). El *ribosome display* se basa en un repertorio en forma de ADN que se transcribe *in vitro* a ARNm y que carece de codón stop, que se traduce *in vitro* para obtener una proteína nativa unida a un ribosoma para realizar la selección por afinidad. El ARNm se libera de los ribosomas y se retrotranscribe a ADN monocatenario que puede amplificarse y/o mutarse, convirtiéndose así en la base de otra ronda de selección (Lipovsek & Plückthun, 2004).

El *ARNm display* mejora el acoplamiento entre genotipo y fenotipo. El proceso comienza con la transcripción de un repertorio de ADN. A continuación, el ARNm se modifica en su extremo 3' con puromicina (un antibiótico cuya estructura emula a un ARNt aminoacilado). La traducción *in vitro* de este complejo ocurre gracias a una reacción de transferencia peptídica que da lugar a un complejo covalente ARNm-proteína, con puromicina uniendo ambas biomoléculas. Posteriormente, el ARNm es retrotranscrito para realizar la selección de afinidad. Finalmente, se eluyen los complejos de alta afinidad, liberando solo las hebras de ADN por hidrólisis del ARN. Este ADN puede mutarse y amplificarse y es el material de partida para la ronda posterior (Plückthun, 2012).



**Figura 5.** Representación esquemática de los métodos de evolución dirigida de anticuerpos. Fuente: elaboración propia.

La expresión en la superficie celular (*cell surface display*) permite que los péptidos y proteínas se expresen en la superficie de las células, fusionándolos con dominios de anclaje a la membrana celular (Lee et al., 2003) en lugar de utilizar bacteriófagos. La expresión de proteínas se mejora al utilizar organismos más complejos que los fagos, como bacterias Gramnegativas (Hoischen et al., 2002), bacterias Grampositivas (Mesnage et al., 1999), levaduras (Chao et al., 2006) y células de mamíferos (Bowers et al., 2014). Un ejemplo de este tipo de metodologías deriva del aprovechamiento de una serie de fenómenos relacionados con la reproducción sexual de las levaduras, que implican el apareamiento de ciertos receptores anclados a la membrana celular, que desencadenan una serie de cascadas de señalización que acaban provocando la fusión de dos levaduras. Aprovechando y modificando este mecanismo y, expresando antígenos y anticuerpos en la superficie de *Saccharomyces cerevisiae* se ha desarrollado una metodología para estudiar la interacción entre proteínas. Cuando la interacción entre los antígenos y anticuerpos expresados

es lo suficientemente fuerte, las dos células de levadura están lo suficientemente próximas para que se activen los mecanismos de fusión celular, generando una célula diploide de mayor tamaño. Este sistema ha servido para seleccionar y mejorar anticuerpos, para generar pegamentos moleculares y para estudiar de manera masiva, las interacciones moleculares que ocurren una en disolución con un repertorio de anticuerpos y antígenos (Younger et al., 2017).

Los métodos de evolución dirigida se han convertido en tecnologías innovadoras para el desarrollo de anticuerpos a un nivel similar al de los anticuerpos monoclonales derivados de hibridomas. En el campo de las enfermedades relacionadas con el gluten, se han aplicado con dos objetivos principales: la caracterización molecular de la enfermedad celíaca y la detección de gluten. La mayoría de los trabajos publicados se han centrado en el uso del *phage display* como herramienta para la caracterización molecular de la enfermedad celíaca (Tabla 4). En este contexto, la tecnología del *phage display* se ha utilizado principalmente para desarrollar anticuerpos de alta afinidad frente a la transglutaminasa tisular (Dieterich et al., 1997).

**Tabla 4.** Ejemplos de la aplicación de la tecnología de *phage display* a la investigación de la enfermedad celíaca.

Tipo de repertorio	Formato de anticuerpo	Isotipo/ familia de genes VH	Selección frente a	Referencia
Inmune (PBLs)	scFv	IgG VH4	$\alpha$ -gliadina tTG otros antígenos	(Sblattero et al., 2000)
Inmune (PBLs y IBLs)	scFv	VH5	$\alpha$ -gliadina tTG	(Marzari et al., 2001)
Inmune (IBLs)	scFv	VH5	tTG	(Sblattero et al., 2004)
Inmune (IBLs)	scFv	VH51-1	tTG	(Not et al., 2011)
<i>Naïve</i>	scFv	-	Péptidos de gliadina presentados por el CMH	(Høydahl et al., 2019)
Inmune	scFv	IgA, IgG e IgM	Gliadina	(Rhyner et al., 2003)
Péptidos	Péptidos	-	Células presentadoras de antígenos (gliadina)	(Chen et al., 2011)

Fuente: elaboración propia

Una primera estrategia consistió en estudiar la respuesta inmune frente a la transglutaminasa tisular (tTG) mediante la construcción de un repertorio inmune a partir de los linfocitos de sangre periférica (PBLs) de un paciente con enfermedad celíaca, realizando una selección frente a cuatro antígenos relacionados con la patología ( $\alpha$ -gliadina,  $\beta$ -lactoglobulina, tTG humana y de cobaya) (Sblattero et al., 2000). Se obtuvieron algunos anticuerpos monoclonales, pero eran polirreactivos frente a los dos primeros antígenos, y no reaccionaban frente a la tTG. La principal

familia de anticuerpos seleccionados frente a la  $\alpha$ -gliadina fue la VH4. Siguiendo este enfoque, se produjeron y analizaron seis repertorios inmunes de scFv a partir de linfocitos de sangre periférica e intestinales (IBLs) de tres pacientes celíacos, y concluyeron que los anticuerpos derivados de los IBLs de todos los pacientes reconocían la tTG con una selección preferente hacia la familia VH5, mientras que los anticuerpos derivados de la sangre periférica reconocían la  $\alpha$ -gliadina (Marzari et al., 2001). A partir de esta investigación, Sblattero et al. (2004) construyeron un repertorio de scFv basado en la amplificación de dos genes de la familia VH5 a partir de los IBLs, lo que permitió una rápida caracterización de la respuesta anti-tTG, que se podría aplicar en pacientes asintomáticos cuyos anticuerpos séricos son indetectables.

La técnica del *phage display* se ha revelado como un método muy útil para la caracterización de la inmunidad celíaca. Not et al. (2011) generaron un repertorio inmune mediante la amplificación del gen VH5-51 a partir de muestras de biopsia intestinal de 22 familiares de pacientes celíacos, analizando su interacción con la tTG. Descubrieron que los individuos genéticamente predispuestos a la enfermedad celíaca producían anticuerpos intestinales VH5 anti-tTG2 en ausencia de anticuerpos séricos anti-tTG2. Otro uso de los repertorios de *phage display* para la caracterización de la enfermedad celíaca fue el trabajo de Høydahl et al. (2019), quienes investigaron el proceso de presentación de antígenos durante la respuesta inmune mucosal. Aplicaron un gran repertorio *naïve* humano en formato scFv para el aislamiento de anticuerpos específicos frente al Complejo Mayor de Histocompatibilidad CMH, haplotipo DQ2.5 unido a gliadina- $\alpha$ 1a. A continuación, los anticuerpos obtenidos se utilizaron frente a células de biopsias intestinales de pacientes con enfermedad celíaca, lo que permitió la identificación de células que expresan péptidos de gluten unidos al complejo mayor de histocompatibilidad en tejidos intestinales inflamados.

Rhyner et al. (2003) construyeron tres repertorios scFv de isotipos únicos (IgA, IgG y IgM) de un paciente celíaco y de un individuo control sano. Estos autores demostraron que todos los repertorios del paciente celíaco, pero ninguno del donante control, se enriquecieron selectivamente en clones de fagos que se unían a la gliadina después de cuatro rondas de *biopanning*. Este método no solo resultó ser un enfoque adecuado para obtener anticuerpos de alta afinidad frente a la gliadina, sino que también permitió la caracterización iso-específica de las respuestas inmunitarias que ocurren en esta condición patológica.

La tecnología del *phage display* también se puede utilizar como un método de exploración para estudiar las interacciones biológicas mediante repertorios de péptidos. Un ejemplo de este concepto aplicado a la enfermedad celíaca fue desarrollado por Chen et al. (2011) utilizando un repertorio de péptidos aleatorios expresados en fagos, que se fue enriqueciendo en péptidos que se unían a la gliadina tras varias rondas de *panning*. Varios clones de estos fagopéptidos fueron capaces de inhibir la interacción entre la gliadina y los anticuerpos antigliadina. Además, los péptidos de 12 aminoácidos codificados por los clones seleccionados fueron sintetizados *in vitro* y analizados en experimentos de competición que revelaron su unión a diferentes lugares

de la gliadina. Los autores sugirieron el uso potencial de estos péptidos para reducir la toxicidad del gluten.

Todos los trabajos mencionados son ejemplos del potencial de la tecnología del *phage display* como herramienta para la caracterización molecular de la enfermedad celíaca. Un mayor conocimiento de la enfermedad celíaca ha contribuido a desarrollar mejores medidas de diagnóstico y prevención, como aquellas derivadas del descubrimiento del papel principal de la transglutaminasa (tTG) y los componentes inmunogénicos del gluten.

Aunque la mayoría de las aplicaciones de *phage display* relacionadas con el gluten se han centrado en el estudio de la respuesta de los anticuerpos en la enfermedad celíaca, hay un creciente interés en el desarrollo de anticuerpos recombinantes para la detección de gluten en alimentos como aplicación final (Tabla 5).

**Tabla 5.** Aplicación de métodos de evolución dirigida de anticuerpos para la detección de gluten en alimentos.

Tecnología	Tipo de repertorio	Formato de anticuerpo	Selección frente a	Método de detección de gluten	Referencia
<i>Phage display</i>	Inmune	VHH	Gliadina	ELISA de captura	(Doña et al., 2010)
<i>Phage display</i>	Semisintético	dAb	Péptido consenso de gluten	ELISA indirecto con fagos	(García- García et al., 2020)
<i>ADNc display</i>	<i>Naïve</i>	VHH	Gliadina	cD-IPCR	(Jayathilake et al., 2020)

Fuente: elaboración propia

Doña et al. (2010) desarrollaron un sistema para la detección de gluten en muestras de alimentos basado en anticuerpos VHH. Construyeron un repertorio de VHH a partir de linfocitos B aislados de llamas inmunizadas con gliadina, que seleccionaron en condiciones desnaturalizantes. Los VHH seleccionados permitieron el desarrollo de un sistema de ELISA sándwich utilizando un VHH frente a gliadina como anticuerpo de captura y un anticuerpo monoclonal murino dirigido frente a la gliadina como anticuerpo de detección, este sistema fue capaz de detectar gliadina en muestras de alimentos que dieron negativo con otros kits de ELISA. Sin embargo, el método solo fue aplicable a la detección de gluten de trigo, ya que los VHH seleccionados no reaccionaron con las prolaminas de cebada o centeno.

Nuestro equipo (García-García et al., 2020) desarrolló un método de fago-ELISA para el análisis de gluten en alimentos basado en fragmentos de anticuerpos de dominio único (o dAb) utilizando un repertorio semisintético (Lee et al., 2007). El repertorio se enriqueció en dAb de alta afinidad mediante rondas sucesivas de selección frente al péptido consenso



Debido al ambiente reductor del citoplasma bacteriano, los fragmentos de anticuerpos recombinantes no se pliegan correctamente y forman agregados disfuncionales. Sin embargo, a través de procedimientos de ingeniería genética se han conseguido biofactorías de *E. coli* capaces de producir proteínas heterólogas correctamente plegadas y con gran rendimiento, mediante la coexpresión de chaperonas como GroES/L o enzimas isomerasas o transferasas. Un ejemplo de estos sistemas es el CyDisCo (Matos et al., 2014), en el que un conjunto de plásmidos que permiten la expresión de diversas sulfidriloxidasas y/o disulfilisomerasas (patente WO 2010/139858), se cotransforman en cepas comerciales de *E. coli*, junto a los plásmidos para la expresión recombinante de uso común (como pET-28) y permiten la expresión citoplásmica de proteínas que presentan gran cantidad de puentes disulfuro (como los anticuerpos o sus fragmentos). Además de la modificación de esta bacteria mediante ingeniería genética, se han desarrollado diversos sistemas para optimizar su cultivo en biorreactores, consiguiendo rendimientos muy elevados (superando 1 g/L) de fragmentos de anticuerpos (Miethe et al., 2013). A pesar de que *E. coli* es la biofactoría más ampliamente utilizada, también existen varios trabajos dónde se utilizan otras bacterias como *Proteus mirabilis* (Rippmann et al., 1998) o *Pseudomonas putida* (Dammeyer et al., 2011) para la expresión recombinante de anticuerpos.

Las bacterias Grampositivas presentan la ventaja de que pueden secretar directamente los anticuerpos recombinantes al medio de cultivo, gracias a que carecen de una membrana externa. Se han utilizado diversas especies del género *Bacillus*, como *Bacillus brevis* (Shiroza et al., 2003), *Bacillus subtilis* (Wu et al., 2002), y *Bacillus megaterium* (David et al., 2010). También se han utilizado lactobacilos, como *Lactocaseibacillus casei* o *Lactocaseibacillus paracasei*, que debido a que se consideran microorganismos GRAS (*Generally Regarded as Safe*), podrían administrarse directamente al paciente y, una vez en la localización anatómica deseada (boca o intestino), producirían los anticuerpos recombinantes de interés (Krüger et al., 2002).

Las células eucariotas presentan una serie de ventajas para producir anticuerpos recombinantes sobre las células procariotas. Al ser células más evolucionadas, son capaces de llevar a cabo un plegamiento más eficaz de las proteínas, presentan aparatos de secreción complejos y pueden realizar una amplia variedad de modificaciones postraduccionales. Las levaduras son células eucariotas simples que reúnen ventajas de ambos tipos celulares (procariotas y eucariotas), como la facilidad para manipularlas genéticamente, y cultivarlas en medios más sencillos, además de que no producen endotoxinas, y presentan las capacidades mejoradas para síntesis, plegamiento y modificación de las proteínas de las células más complejas (Frenzel et al., 2013). La levadura más utilizada es *Pichia pastoris*, una biofactoría catalogada como GRAS. *P. pastoris* puede utilizar el metanol como única fuente de carbono, lo que facilita el control de la expresión de proteínas recombinantes, y se puede cultivar en biorreactores de volumen medio con rendimientos que superan 1 g/L (Potgieter et al., 2010). Se han utilizado también especies de levaduras como *Saccharomyces cerevisiae*, *Hansenula polymorpha*, *Schizosaccharomyces pombe*, *Schwanniomyces occidentalis*, *Kluyveromyces lactis*, y *Yarrowia lipolytica* (Frenzel et al., 2013).

Además de las levaduras unicelulares, también se han cultivado hongos filamentosos modificados genéticamente para producir anticuerpos recombinantes. Aprovechando su gran capacidad secretora, se han utilizado *Aspergillus niger* para obtener anticuerpos completos (Ward et al., 2004), *Trichoderma reesei* para Fab (Nyssönen et al., 1993) y *Aspergillus awamori* para producir dAb, scFv y sus fusiones con enzimas (Joosten et al., 2003).

Las células de insectos son fácilmente transfectables mediante infección con baculovirus. Además, son fáciles de cultivar y presentan un sistema de producción y secreción de proteínas avanzado, lo que facilita la purificación de anticuerpos y otras proteínas recombinantes. Sin embargo, producen anticuerpos recombinantes con patrones de glicosilación diferentes a los de las células de mamífero, por lo que pueden provocar inmunogenicidad no deseada si se aplican directamente con fines terapéuticos (Edelman et al., 1997).

Actualmente, entre el 60 y el 70 % de los principios activos consistentes en proteínas recombinantes y el 95 % de los anticuerpos terapéuticos aprobados se siguen produciendo en líneas celulares de mamífero, a pesar de sus costes de producción y la complejidad de los procesos industriales. Sin embargo, la capacidad avanzada de plegamiento, secreción y modificación postraduccional de proteínas, permite la obtención directa de anticuerpos prácticamente indistinguibles de los producidos por el ser humano, sin apenas problemas de inmunogenicidad no deseada. Además, la secreción de inmunoglobulinas complejas correctamente plegadas y modificadas, permite reducir los esfuerzos en procesos de purificación (Frenzel et al., 2013). Las líneas de células de ovario de hámster chino (CHO) y sus derivadas son las más utilizadas. Otras líneas celulares que han recibido la aprobación por parte de agencias reguladoras para la producción de anticuerpos recombinantes de uso terapéutico en humanos son: las células per.C6, células de mieloma de ratón (NS0), células de riñón de hámster y células embrionarias de riñón humano (HEK 293) (Butler, 2005).

Otra opción para la obtención de anticuerpos es la producción recombinante en plantas (*planticuerpos*) cuyo coste puede ser hasta 100 veces inferior a la producción en hibridomas (Daniell et al., 2001). Las células vegetales se pueden transfectar rápida y eficientemente gracias a la infección con *Agrobacterium tumefaciens*, que vehicula el gen recombinante para la producción de anticuerpos. A pesar de estas ventajas, las plantas también imprimen patrones de glicosilación inmunogénicos a las proteínas producidas, aunque esto se puede paliar gracias a los avances en la glicoingeniería. Este tipo de expresión recombinante permite la generación de productos vegetales que contienen antígenos o anticuerpos para inmunización, las denominadas vacunas comestibles (Fausther-Bovendo & Kobinger, 2021).

Por último, existen muchos estudios en los que se generan animales transgénicos que expresan cadenas de anticuerpos humanos en vez de las propias de la especie, sirviendo como sistemas de producción de anticuerpos y como modelos para estudiar el sistema inmune y las enfermedades

asociadas (Flisikowska et al., 2011). Un paso más en esta línea de trabajo, es generar mamíferos transgénicos que produzcan proteínas recombinantes en su leche, facilitando su aislamiento cuando sea necesario, o incluso el consumo directo a través de esa leche como alimento (Pollock et al., 1999).

## 1.6 CARACTERIZACIÓN DE REPERTORIOS DE ANTICUERPOS MEDIANTE TÉCNICAS DE SECUENCIACIÓN MASIVA

### 1.6.1 Tecnologías de secuenciación masiva

El análisis mediante Secuenciación de ADN de Próxima Generación (NGS, *Next-Generation Sequencing*) de los repertorios de anticuerpos humanos ha permitido una comprensión más profunda y una caracterización completa de los receptores de células B (BCR, *B-Cell Receptor*) con una resolución sin precedentes (Robinson, 2015). La tecnología NGS ha dado lugar a numerosas aplicaciones relacionadas con el desarrollo de anticuerpos. Se ha mejorado el *panning* con repertorios de anticuerpos expresados en fagos facilitando el proceso de su aislamiento (Vaisman-Mentesh & Wine, 2018). Además, cuando se combina con la modelización estructural computacional a gran escala, ha revelado correlaciones secuenciales y estructurales entre los repertorios de anticuerpos *naïve* y sus posibles antígenos (Dekosky et al., 2016). Otro ejemplo de su aplicación, fue el estudio del “anticuerpoma” de individuos infectados con el VIH, que ha aumentado la comprensión de la respuesta inmunológica, y ha ayudado al diseño de inmunógenos basados en los linajes de anticuerpos que podrían activar las células B *naïve* para dar lugar a clones neutralizantes de amplio espectro (Havenar-Daughton et al., 2018). En consecuencia, la NGS de los repertorios de anticuerpos humanos será una tecnología fundamental para el descubrimiento de nuevos anticuerpos (Choi et al., 2023) y el desarrollo de vacunas (Song & Zhang, 2022).

El método de secuenciación de Sanger automatizado se considera una tecnología de primera generación. A pesar del impacto que tuvo en los albores de la biología molecular, y de que se sigue utilizando para ciertas aplicaciones concretas, su rapidez no es suficiente para poder llevar a cabo análisis masivos. Las tecnologías NGS permiten la obtención de millones de secuencias en paralelo, con mayor rapidez y menor coste que las tecnologías de primera generación, ya que no es necesario separar las moléculas de ADN mediante técnicas de electroforesis. El principal avance ofrecido por las tecnologías de NGS es la capacidad de producir un enorme volumen de datos a bajo coste, ampliando el ámbito de experimentación más allá de simplemente determinar el orden de las bases. La capacidad de secuenciar el genoma completo de muchos organismos ha permitido realizar estudios comparativos y evolutivos a gran escala que eran inimaginables hace apenas unos años. La aplicación más amplia de la NGS puede ser la resecuenciación de genomas humanos para mejorar nuestra comprensión de cómo las diferencias genéticas afectan a la salud y a la enfermedad. En este campo, se ha completado recientemente la secuenciación

del pangenoma humano, una colección de ensamblajes del genoma completo de 47 personas, con la posibilidad de incluir hasta 350 genomas de personas diferentes en 2024, para obtener la mayor colección de variantes en genomas completos hasta la fecha (Liao et al., 2023). La variedad de características de la NGS hace posible que múltiples plataformas coexistan en el mercado.

Uno de los primeros sistemas comercializados en 2005 fue la tecnología Roche/454 (Roche®, Basilea, Suiza) basada en la pirosecuenciación paralela a gran escala, tras la amplificación de moléculas individuales de ADN en microrreactores formados por una emulsión de aceite-agua. Este sistema dejó rápidamente de ser competitivo y se retiró del mercado en 2013. Otros sistemas basados en la pirosecuenciación fueron el Life/APG y Polonator (Metzker, 2010).

La tecnología Illumina® (San Diego, California, EEUU) se basa en un mecanismo diferente de terminadores reversibles que son nucleótidos marcados químicamente mediante la adición de un fluoróforo. Las enzimas polimerasas expanden la cadena de ADN a partir de los cebadores de secuenciación añadiendo terminadores reversibles. Los terminadores reversibles provocan que la incorporación de más bases se detenga inmediatamente después del primer nucleótido. Las polimerasas y los nucleótidos no unidos son eliminados. Los fluoróforos son excitados con dos láseres separados (rojo: A, C y verde: G, T). Posteriormente, se eliminan los fluoróforos y terminadores, y la secuenciación se reanuda con la inclusión de la siguiente base (Kircher & Kelso, 2010). Este sistema pronto dominó el mercado mundial de la secuenciación, y aún hoy es el más utilizado gracias a que presenta varias ventajas como el nivel adecuado de velocidad, capacidad y coste para su empleo por diferentes laboratorios o requerimientos de secuenciación. Además, está asociado a herramientas bioinformáticas fáciles de usar y accesibles a través de la web, o en servidores externos (Frey et al., 2014). También presenta una serie de inconvenientes, principalmente la necesidad de un control estricto en la carga de las muestras, ya que una sobrecarga puede dar lugar a la superposición de clústeres con una baja calidad de secuenciación, lo que resulta en una tasa de error general de aproximadamente el 1 % (Nakamura et al., 2011). El tamaño máximo de secuencia registrado por este sistema es de 2x150 pares de bases, lo que hace necesario fragmentar, marcar con códigos de barras, secuenciar los fragmentos y reensamblarlos mediante herramientas bioinformáticas (Sedlazeck et al., 2018). Este inconveniente derivó en una gran competencia entre empresas para desarrollar una tecnología capaz de secuenciar moléculas largas (de miles o incluso millones de pares de bases). En este contexto se desarrollaron las tecnologías PacBio® (Pacific Biotech, Menlo Park, California, EEUU) y Oxford Nanopore® (Oxford, Reino Unido).

En el sistema PacBio® se añade al dispositivo una disolución que contiene cada uno de los cuatro nucleótidos marcados con una molécula fluorescente diferente. Cada vez que un nucleótido marcado se une al sitio activo de la polimerasa que está anclada en el fondo de los pocillos del dispositivo, el fluoróforo se excita, generando una señal específica para cada base incorporada. A continuación, el fluoróforo se escinde de la base nucleotídica, deja de emitir, y difunde fuera del

pocillo, de forma que se puede reiniciar el ciclo con la incorporación de un nuevo nucleótido. Este sistema permite generar lecturas de manera directa con la misma longitud que la capacidad de elongación de la polimerasa empleada, hasta 60 kb, con muy bajas tasas de error (Rhoads & Au, 2015). Este sistema ha demostrado ser especialmente útil para secuenciar repertorios de scFv, generando lecturas de más de 800 pb, y se ha empleado para caracterizar el proceso de selección de scFv de un repertorio de origen humano expresado en fagos que se unían a placas de ateroma en animales de experimentación (Hemadou et al., 2017).

El mecanismo de Oxford Nanopore es completamente diferente. En este caso, se trata de un dispositivo que mide variaciones de la corriente eléctrica en un líquido que inunda una cámara que contiene una membrana porosa. A los fragmentos de ADN que se van a secuenciar se les añaden unos adaptadores y unas proteínas motoras en los extremos, que permiten el direccionamiento de las hebras de ADN a través de los poros. El paso de cada una de las bases nitrogenadas produce una variación diferente en la corriente eléctrica, que permite registrar la secuencia de la hebra de ADN. La principal novedad de este sistema es que se han desarrollado dispositivos miniaturizados como el MinION, que permiten realizar experimentos de secuenciación en cualquier laboratorio, sin necesidad de grandes equipos, simplemente con un ordenador, el dispositivo miniaturizado, el kit para preparación de muestras y un juego de pipetas automáticas (Chen et al., 2023). A pesar de tener la mejor relación coste- efectividad del mercado, el sistema portátil presentaba la debilidad de una tasa de error en la secuenciación mayor que otros sistemas, por lo que se desarrollaron otros dispositivos, esta vez más grandes y complejos, que permiten otros mecanismos de secuenciación como el 1D2 que permite la lectura consecutiva de ambas hebras de una molécula de ADN, o la 2D, análoga a la anterior pero en este caso, las dos hebras de la molécula (paralela y antiparalela) se encuentran unidas físicamente una tras otra. Esto permite una doble lectura de información complementaria para reducir errores (Lin et al., 2021).

Otro sistema, no tan extendido, son los dispositivos Ion Torrent® (San Francisco, California, EEUU). Su mecanismo de secuenciación se produce también en microemulsiones de agua-aceite, en cuyo interior ocurre que cada tipo de base nucleotídica añadida por una polimerasa produce una cantidad diferente de protones, que son registrados por detectores de pH muy sensibles, permitiendo así la secuenciación (Lahens et al., 2017).

### 1.6.2 Análisis masivo de anticuerpos

La complejidad del análisis de datos de secuenciación masiva de repertorios de anticuerpos presenta dos retos principales. Por un lado, la complejidad intrínseca de analizar moléculas hipervariables como son los anticuerpos (y en concreto algunas de sus zonas como los HCDR3), y, por otro lado, la dificultad de analizar el volumen de datos generados por la secuenciación. Lo habitual en este tipo de experimentos es seguir un flujo de trabajo con dos etapas principales. La primera etapa o análisis primario, consiste en la organización de las secuencias brutas

generadas por el secuenciador, para eliminar secuencias de mala calidad que pueden generar errores e introducir ruido en el análisis posterior (Han et al., 2018). Seguidamente se realiza un análisis secundario con las secuencias ya depuradas con el fin de extraer la información de los anticuerpos. El primer análisis se basa en el alineamiento de las secuencias obtenidas frente a bases datos curadas con información de secuencias de anticuerpos (Lefranc et al., 2008). Este alineamiento permite la identificación de los genes V(D)J y C que han generado cada anticuerpo, su clasificación por familias, subfamilias y alelos para las regiones variables, la clase y la subclase para las regiones constantes y proporcionan información sobre los CDRs, como su longitud y composición en aminoácidos (Lefranc & Lefranc, 2019). Una vez anotada toda esta información, y si se tienen datos de varias etapas de un proceso (como por ejemplo las diferentes rondas en un proceso de *panning*) se puede ejecutar un segundo análisis con el objetivo de encontrar, principalmente cambios significativos en las diferentes etapas del proceso (Udoye et al., 2022), rastrear las secuencias de anticuerpos a lo largo del tiempo o en respuesta a diferentes condiciones, haciendo posible comprender cómo evolucionan y se adaptan los repertorios de anticuerpos en diferentes situaciones.

Este tipo de análisis es posible gracias a la existencia de bases datos, que se pueden consultar en internet, y que contienen las secuencias que codifican para las diferentes familias de anticuerpos de diversas especies. Existe un repertorio de bases de datos de anticuerpos y herramientas para su análisis universalmente utilizado, denominado sistema de información internacional IMGT (*Internacional ImMunoGeneTics Information System*) iniciado y mantenido por la Universidad de Montpellier y el Centro Nacional para la Investigación Científica (CNRS, *Centre National de la Recherche Scientifique*), ambos en Francia. En 1989, con la entrada de todos los genes del locus TRG de *Homo sapiens* en la base de datos HGM, los genes variables (V), de diversidad (D), unión (J) y constante (C) de inmunoglobulinas (IG) o receptores de células B (BCR), y de células T (TR) fueron reconocidos oficialmente como genes por primera vez, ya que hasta entonces eran considerados segmentos génicos o pseudogenes (Lefranc & Lefranc, 2020). Este importante avance permitió que los genes y alelos de IG y TR, que son parte de las respuestas inmunitarias adaptativas complejas y altamente diversificadas, fueran gestionados en bases de datos y herramientas genómicas. Los nombres de genes y alelos de IMGT se basan en la clasificación entorno a las categorías grupo, subgrupo, gen, y alelo (Lefranc, 2013). La categoría grupo permite clasificar un conjunto de genes que pertenecen a la misma familia de genes múltiples, dentro de la misma especie. Por ejemplo, existen 10 grupos para las IG de vertebrados superiores: IGHV, IGHD, IGHJ, IGHC, IGKV, IGKJ, IGKC, IGLV, IGLJ, e IGLC. La categoría subgrupo permite identificar un subconjunto de genes que pertenecen al mismo grupo y que, en una especie dada, comparten al menos un 75 % de identidad a nivel de nucleótidos. Por ejemplo, el subgrupo IGHV1 de *Homo sapiens*. Los subgrupos genes y alelos siempre están asociados con el nombre de la especie (Lefranc, 2008). Un alelo es una variante polimórfica de un gen, que se caracteriza por las mutaciones de su secuencia a nivel de nucleótidos,

identificadas en sus secuencias centrales y comparadas con la secuencia de referencia de cada alelo del gen. Para el locus IGH, los genes constantes se designan con la letra y número si fuera necesario correspondiente a los isotipos codificados (IGHM, IGHD, IGHG3...), en lugar de usar la letra C. El origen de las secuencias de esta base de datos son aquellas depositadas en la base de datos GENBANK del NCBI (Centro Nacional para Información de Biotecnología de los Estados Unidos). Debido a la enorme complejidad de la dinámica genómica que conduce a la generación de diversidad de los receptores de la inmunidad adaptativa, surgió la controversia sobre qué secuencias se debían utilizar como referencia teniendo en cuenta que todas las células del organismo contienen los genes codificantes para anticuerpos, pero solo en los linfocitos B se expresan las enzimas que llevan a cabo los mecanismos que generan anticuerpos funcionales (Hoolehan et al., 2022). Además, cuanto más maduro es un linfocito B su genoma está más mutado y, por lo tanto, alejado de las secuencias originales de la línea germinal (Borbet et al., 2021). En la Tabla 6 se describe el origen de las secuencias integradas en la base de datos IMGT para el análisis de anticuerpos.

Todas estas secuencias se ordenan por familia génica, obteniéndose una base de datos que contiene el repertorio potencial de genes que codifican para anticuerpos definidos (por genoma haploide y haplotipo más frecuente) por 46 genes funcionales V clasificados en 7 familias, 23 genes funcionales D, 6 genes funcionales J y 9 genes C (que definen las subclases de los anticuerpos) para la cadena pesada (Scaviner et al., 1999).

**Tabla 6.** Clasificación de los 10 tipos de secuencias integradas en la base datos IMGT.

Entidad molecular	Dominio/s	Origen de la secuencia	Tipo de molécula
Gen V	NA	Línea germinal	gDNA
Gen D	NA	Línea germinal	gDNA
Gen J	NA	Línea germinal	gDNA
Gen C	NA	Línea germinal	gDNA
Gen V-D-J	VH	Reorganizado	gDNA
Gen V-J	VL	Reorganizado	gDNA
Secuencia L-V-D-J	VH	Reorganizado	cDNA
Secuencia L-V-J	VL	Reorganizado	cDNA
Cadena V-D-J-C	Cadena Pesada	Reorganizado	Proteína
Cadena V-J-C	Cadena Ligera	Reorganizado	Proteína

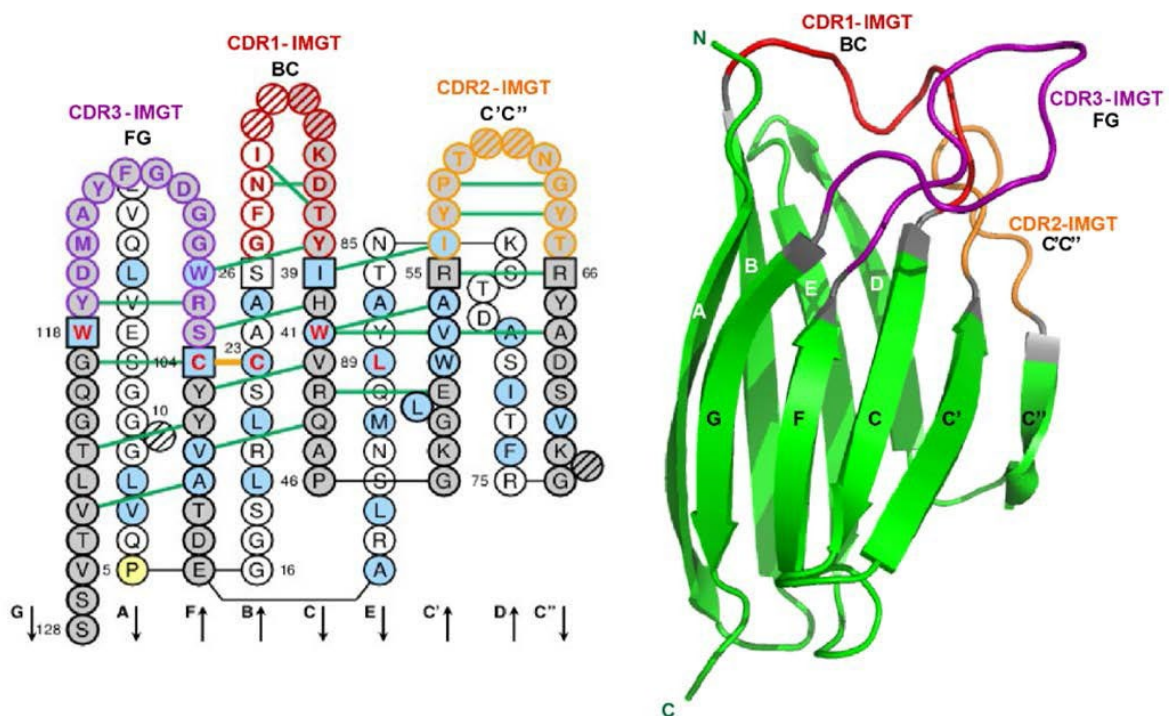
IMGT: *Internacional ImMunoGeneTics Information System* Fuente: Lefranc & Lefranc (2020).

Con respecto a la cadena ligera, la base de datos está organizada en 35 genes funcionales para las cadenas kappa, clasificadas en 5 familias, 5 genes funcionales para J y un único gen C (Barbié & Lefranc, 1998). La organización de los genes que codifican las cadenas lambda es más compleja porque no hay aún consenso sobre si varios de los genes observados son realmente genes o pseudogenes y su funcionalidad es dudosa. Por motivos prácticos, se clasifican en 10 familias para los genes V (algunos autores consideran 11) y hasta 7 familias para los J (4 han sido identificadas mediante análisis de *Southern blot* y los otros 3 difieren en su clasificación según el autor que se consulte, aunque la base de datos IMGT considera los 3) (Lefranc et al., 1999). La complejidad en la clasificación de los genes que codifican para las cadenas ligeras es aún mayor, teniendo en cuenta que se han encontrado numerosos orfones, que son genes localizados fuera del locus principal dónde se encuentran el resto de su familia génica, algunos de ellos en posiciones tan alejadas como el cromosoma 18 (Lefranc & Lefranc, 2020).

La generación de las bases de datos como IMGT permite la clasificación de los anticuerpos secuenciados según un repertorio potencial de referencia, pero debido a la enorme diversidad que se genera en ciertos dominios (CDRs), no es posible obtener un análisis completo de estas zonas mediante la comparación con los repertorios de referencia. El análisis de las zonas hipervariables se realiza *de novo* basado principalmente en el tamaño y composición de aminoácidos del HCDR3. Para ello es necesario delimitar estas zonas mediante la información estructural reflejada en las secuencias de aminoácidos (Lefranc et al., 2003). Un dominio V (H o L) consta de aproximadamente 100 aminoácidos, y procede de la reorganización VDJ o VJ, estando compuesto por nueve láminas beta antiparalelas (A, B, C, C', C'', D, E, F y G) unidas por láminas betas giradas (AB, CC', C''D, DE y EF) y tres bucles (BC, C'C'' y FG) que se corresponden con tres CDRs, formando un sándwich de dos láminas [ABED] [GFCC'C'']. Las láminas están empaquetadas estrechamente entre sí, formando un núcleo hidrofóbico, y se unen mediante un puente disulfuro entre una primera cisteína altamente conservada en la hebra B (en la primera lámina) y una segunda cisteína igualmente conservada en la hebra F (en la segunda lámina) (Lefranc, 2019). La delimitación de los CDRs se define mediante una serie de puntos de anclaje, que están bien definidos para los dos primeros CDRs, y son las posiciones 26 y 39 del bucle BC para el CDR1, y las posiciones 55 y 66 del bucle C'-C'' para el CDR2. Con respecto al CDR3, se consideran como punto de anclaje las únicas posiciones realmente conservadas, que son la segunda cisteína (hebra F) y la fenilalanina o el triptófano de la hebra G (Figura 6) (Lefranc, 2013).

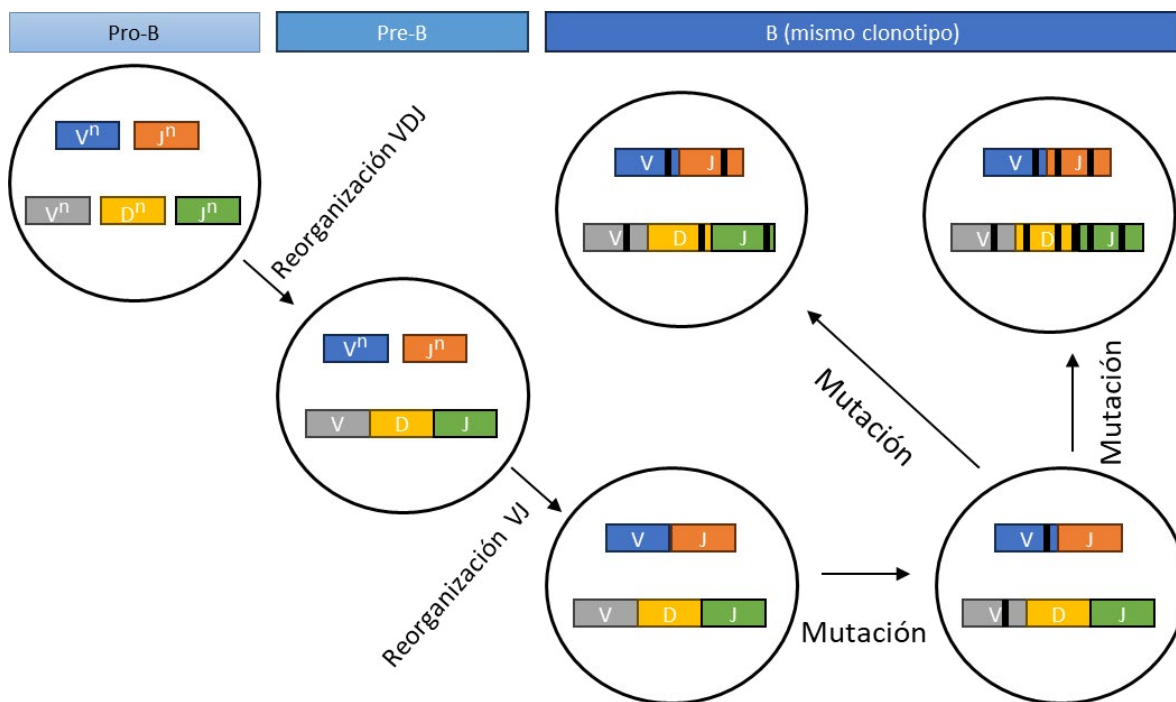
Una vez definidas las propiedades que permiten identificar un anticuerpo de un repertorio (familia, subfamilia y alelo al que pertenecen), y ciertas características estructurales se puede realizar el análisis de datos NGS de manera automática con la herramienta IMGT/HighV-QUEST, a través de la cual se pueden organizar las colecciones masivas de datos de la secuenciación de repertorios.

Esta herramienta realiza 6 funciones principales: 1) la introducción de huecos y/o reordenamiento de las posiciones de los aminoácidos en el anticuerpo para facilitar el alineamiento en la base de datos (esto ocurre especialmente en los CDRs, que presentan un número muy variable de aminoácidos) (Lefranc et al., 2003); 2) identificación de inserciones y deleciones (*indels*), y en su caso, su corrección; 3) identificación de familias y subfamilias génicas y alelos más similares a aquellos registrados en la base de datos; 4) delimitación de los diferentes dominios estructurales; 5) análisis pormenorizado de la zona de reorganización de los genes VDJ o VJ, que se encuentra en el CDR3 y zonas adyacentes y 6) anotación de mutaciones y cambios de aminoácidos (Alamyar et al., 2012).



**Figura 6.** Representación de los dominios estructurales de la región VH del trastuzumab. A la izquierda aparece una representación plana de la estructura en forma de “collar de perlas”. Los HCDR1, HCDR2 y HCDR3 (correspondientes a los bucles BC, C’C’’ y FG, respectivamente) se representan en rojo, naranja y morado. Los círculos rayados corresponden a las posiciones que faltan con respecto a la secuencia de referencia en la base de datos IMGT. Las flechas indican la dirección de las láminas beta y sus designaciones en las estructuras 3D. Los puntos de anclaje se muestran en cuadrados. Los puentes de hidrógeno (líneas verdes) fueron añadidos a partir de los datos estructurales experimentales. El puente disulfuro (línea naranja) también fue incorporado a partir de los datos estructurales experimentales. A la derecha se muestra la representación tridimensional de la misma estructura con los CDRs coloreados con la misma gama de colores que en la figura a la izquierda. Figura adaptada de Lefranc & Lefranc (2020).

Otro análisis muy útil para caracterizar un repertorio de anticuerpos es agrupar las diferentes secuencias en grupos o clústeres con características comunes, gracias al concepto de clonotipo. Un clonotipo se define biológicamente como un conjunto de anticuerpos que derivan de la misma línea clonal. Son la evolución de un BCR (*B-cell receptor* o receptor de célula B) que es un anticuerpo anclado a la membrana celular, a través del receptor de Fc parental con recombinación V(D)J completa, que va sufriendo mutaciones en las etapas finales de la maduración (Hong et al., 2022) (Figura 7).



**Figura 7.** Representación del concepto de clonotipo en anticuerpos. Los círculos representan varios linfocitos B en diferentes etapas en el desarrollo y la maduración de su BCR. Los bloques representan los genes V(D)J y las barras negras, las mutaciones que ocurren tras la reorganización. Fuente: elaboración propia.

Al secuenciar repertorios inmunes, el material biológico de partida proviene de linfocitos B que ya han sido madurados, por lo que las secuencias obtenidas de los anticuerpos que generan han sido ya intensamente mutadas. Esto dificulta la clasificación de diferentes secuencias en un clonotipo. Se han propuesto diferentes metodologías para organizar los componentes de un clonotipo, pero la más utilizada actualmente es la propuesta y utilizada por el IMGT, que define que varios anticuerpos diferentes pertenecen al mismo clonotipo cuando presentan una única reorganización V(D)J y el mismo HCDR3 en composición de aminoácidos y presentando sus puntos de anclaje (Li et al., 2013). La clusterización por clonotipos facilita enormemente el análisis de grandes repertorios y la búsqueda de significado biológico de los procesos estudiados.

Una vez conocida la distribución en grupos, familias, subfamilias y alelos de los genes V(D)J, sus dominios estructurales, y distribución por clonotipos, se puede proceder a comparar si hay

cambios significativos en estas características en diferentes puntos de un proceso. Por ejemplo, si se tienen datos de linfocitos que se encuentran en diferentes tejidos o de la evolución en un proceso de selección por afinidad mediado por fagos. Para ello se pueden elegir diferentes formas de hacer comparaciones entre los datos obtenidos, curados y clasificados, pero lo más habitual es comparar la diversidad o expresión de aminoácidos por genes V, D o J de un grupo de anticuerpos dados (Aouinti et al., 2016).

## 1.7 ELUCIDACIÓN *IN SILICO* DE LA INTERACCIÓN ANTÍGENO Y ANTICUERPO

### 1.7.1 Predicción de estructuras de proteínas y métodos clásicos utilizados

En 1953, Frederick Sanger publicó la primera secuencia de aminoácidos de una proteína (Insulina) (Stretton, 2002). En 1959, Max Perutz obtuvo la primera estructura tridimensional de una proteína (Mioglobina) mediante difracción de rayos X (Dickerson, 1992). A partir de ese momento, diversas investigaciones buscaron métodos que permitieran predecir la estructura de una proteína a partir de su secuencia lineal de aminoácidos.

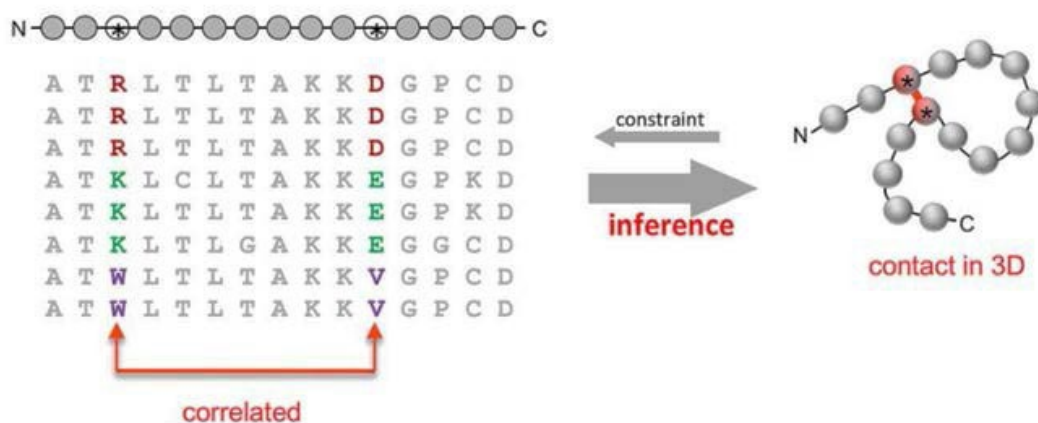
Hasta finales del siglo XX no se consiguió la suficiente potencia computacional para que los algoritmos implementados fueran capaces de obtener predicciones aceptables. Desde entonces se han propuesto diferentes metodologías que mejoraban paulatinamente la capacidad de predicción. Los métodos de predicción suelen dividirse en el modelado basado en estructuras de referencia (TBM, *Template-based modeling*) y en el modelado libre (FM, *Free modeling*) (de Almeida Paiva et al., 2022). Sin embargo, en ocasiones se entremezclan, ya que algunos métodos TBM utilizan metodologías de refinamiento de modelos guiados por energía libre, y parte de los FM utilizan enfoques de muestreo basados en fragmentos, extrayendo información de la base de datos de proteínas (PDB, *Protein Data Bank*) mediante aprendizaje automático. Dentro de estos dos tipos de métodos, los algoritmos desarrollados suelen clasificarse en tres grupos diferentes: *ab initio* (metodología FM), reconocimiento del plegamiento (metodología TBM), y por homología (metodología TBM) (Agnihotry et al., 2022).

Los enfoques *ab initio* se basan en la hipótesis termodinámica de que la estructura proteica nativa presenta la energía libre más baja posible (Hardin et al., 2002). La idea de estos métodos es predecir nuevos pliegues considerando propiedades fisicoquímicas del proceso de plegamiento de proteínas como la formación de puentes de hidrógeno, la energía potencial de contacto, las posibles propensiones de las estructuras secundarias, y el plegamiento que involucra tanto interacciones enlazadas como no enlazadas. El método *ab initio* puede tener en cuenta o no la identificación de determinados motivos en bases de datos mediante la división de la proteína problema en fragmentos pequeños. Algunas revisiones bibliográficas incluso dividen esta

categoría en dos metodologías distintas, una dependiente y otra independiente de la información de la base de datos (Bertoline et al., 2023). La ventaja principal de los métodos *ab initio* es la capacidad de obtener pliegues desconocidos. No obstante, la complejidad del problema y el alto número de posibilidades conformacionales requieren un alto poder computacional, lo que limita su uso en secuencias proteicas largas. Los métodos que utilizan fragmentos pueden ayudar a reducir el espacio conformacional, pero pueden ignorar el plegamiento de zonas importantes (Dorn et al., 2014).

Las metodologías de reconocimiento del plegamiento se basan en la idea de que la estructura está más conservada que la secuencia de aminoácidos y que existe un número limitado, aunque amplio de pliegues de estructuras proteicas en la naturaleza. Estos métodos consisten en elegir la mejor estructura tridimensional de referencia que se ajuste bien a la secuencia problema, considerando una función de puntuación construida en base a potenciales de interacción, comparación de estructuras secundarias y propiedades del solvente. Así, la secuencia objetivo se alinea con el modelo de estructura utilizando la función de puntuación óptima, reorganizando los átomos de la secuencia problema sobre el esqueleto alineado. Finalmente, se verifica la afinidad de la secuencia con el pliegue tridimensional, y se verifica manualmente el resultado final (Jones, 1999). Esta metodología presenta el inconveniente de que depende en gran medida de estructuras preexistentes y, en proteínas con amplia variabilidad, como los anticuerpos, pueden no existir estructuras lo suficientemente similares como para obtener un buen modelo predictivo.

Los modelos de homología fueron los más utilizados hasta la aparición de sistemas de inteligencia artificial. La metodología más utilizada era la búsqueda de mutaciones correlacionadas (Figura 8), que se basa en que la secuencia de la proteína sobre la cual se va a predecir la estructura tridimensional es parte de una familia de secuencias relacionadas evolutivamente que, se presume, tienen esencialmente el mismo plegamiento (familia iso-estructural).



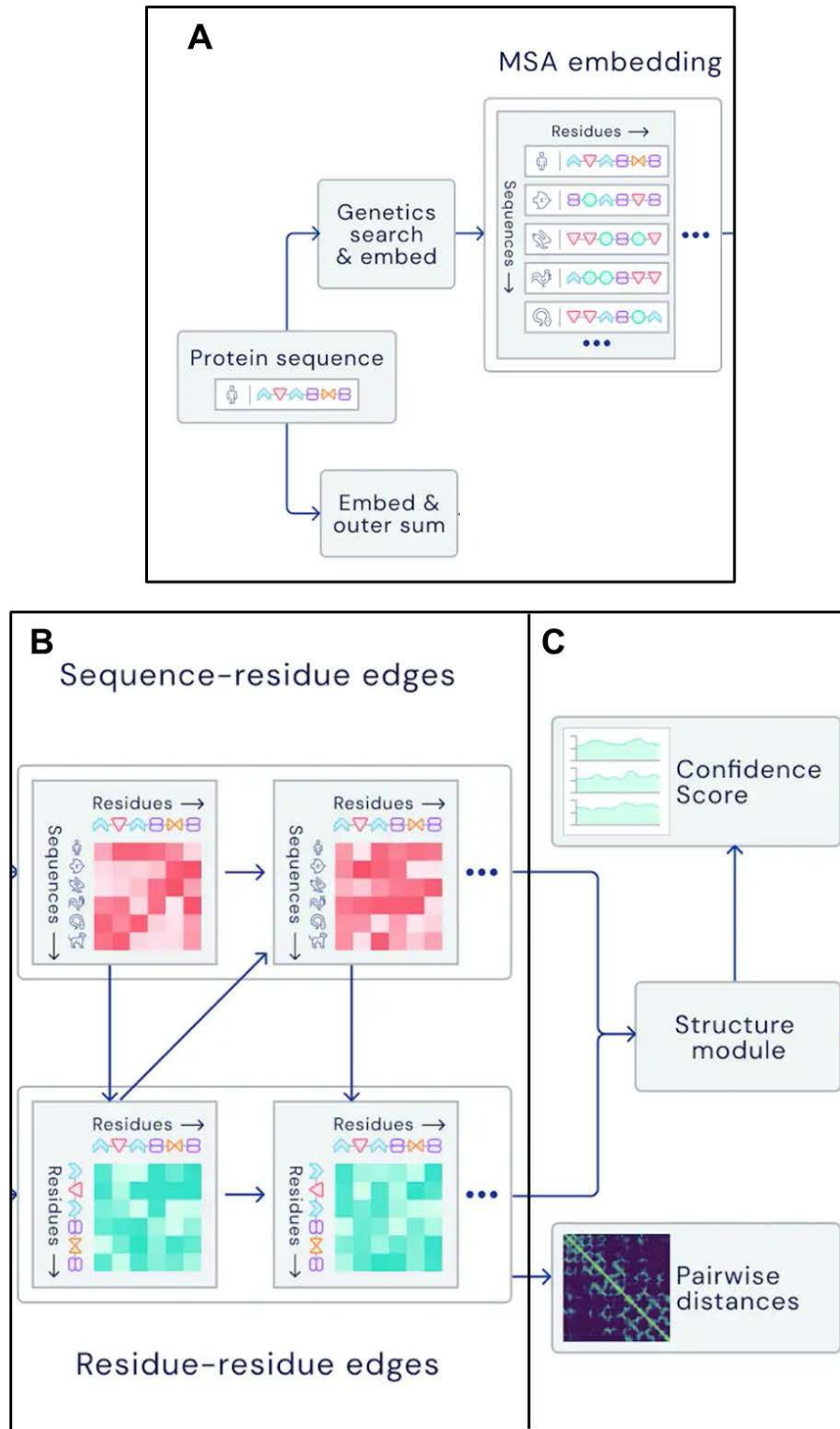
**Figura 8.** Ejemplo gráfico del uso de mutaciones correlacionadas para descubrir contactos entre aminoácidos como primer paso para la predicción de la estructura tridimensional de una proteína. Fuente: Marks et al. (2011).

La variación evolutiva en las secuencias está limitada por varios requisitos, incluida la conservación de interacciones favorables en contactos entre residuos. La predicción del plegamiento de proteínas a partir de su secuencia aprovecha las correlaciones entre pares en el alineamiento de múltiples secuencias para deducir qué pares de residuos probablemente estén cerca uno del otro en la estructura tridimensional. Un subconjunto de los pares de residuos en contacto predichos se utiliza posteriormente para plegar cualquier proteína de la familia en una forma tridimensional aproximada, que luego se refina utilizando técnicas estándar de física molecular (Marks et al., 2011). La calidad final de la estructura predicha por estos métodos depende de que los contactos encontrados sean realmente significativos y directos. Para ello, necesitan alineamientos lo más amplios posibles, por lo que el refinamiento de los métodos estadísticos permitió un incremento en la calidad de las predicciones, con un aumento de la capacidad computacional necesaria para realizar los cálculos (Morcos et al., 2011). Los principales inconvenientes de este tipo de predicciones están relacionados con que no son capaces de predecir cambios como inserciones o deleciones que provoquen cambios conformacionales. Además, tienen baja sensibilidad prediciendo las posiciones de las cadenas laterales (Pitman & Menz, 2006).

### 1.7.2 La revolución de la inteligencia artificial

A pesar de las múltiples metodologías propuestas, se seguían encontrando inconsistencias entre los modelos predichos y las estructuras obtenidas por técnicas físicas, como la difracción de rayos X o la criomicroscopía. A finales de la década de 2010, las grandes compañías tecnológicas empezaron a desarrollar sistemas de inteligencia artificial para resolver algunos problemas complejos (Bertoline et al., 2023). Deep Mind, la filial centrada en el desarrollo de inteligencia artificial de Google®, se interesó por el problema de la predicción de la estructura tridimensional de proteínas a partir de su secuencia. El primer sistema fue presentado en 2019, pero no obtuvo mejores resultados que sus competidores, que resolvían las predicciones con la metodología de homología, mientras que la inteligencia artificial los resolvía mediante el procesamiento de imágenes (Senior et al., 2019). En 2020 se produjo una auténtica revolución en este campo, a través de la presentación de AlphaFold2 (Jumper et al., 2021), un sistema de inteligencia artificial, basado en el procesamiento del lenguaje natural, que era capaz de resolver la estructura de proteínas con una precisión y rapidez nunca vistas hasta entonces, dejando prácticamente obsoletos los métodos anteriores. La metodología que utiliza AlphaFold2 para resolver la estructura de las proteínas se puede dividir en tres fases. En la primera (Figura 9A) utiliza la secuencia de aminoácidos de la proteína problema para consultar varias bases de datos de secuencias de proteínas y construir un alineamiento de secuencias múltiples (MSA, *Multiple Sequence Alignment*). Esto permite determinar las partes de la secuencia que son más probables que muten, además de detectar correlaciones entre ellas. También busca proteínas que puedan tener una estructura similar a la proteína problema y construye una representación inicial de la estructura, denominada representación de pares.

En esencia, crea un modelo de qué aminoácidos es probable que estén en contacto entre sí, de manera análoga a los sistemas de predicción por homología (Elofsson, 2023).



**Figura 9.** Representación del funcionamiento de AlphaFold2. Figura modificada de <https://www.deepmind.com/blog> que actualmente no se encuentra accesible.

En la segunda fase (Figura 9B), toma como base el alineamiento de secuencias múltiples y las estructuras similares, y las pasa a través de un transformador llamado *Evoformer*. El objetivo de esta fase es refinar las representaciones, tanto para el MSA como para las interacciones entre pares, pero también para intercambiar información entre ellos de manera iterativa. Un mejor modelo del MSA mejorará la caracterización de la geometría de la red, lo que a su vez ayudará a refinar el modelo del MSA. Este proceso está organizado en bloques que se repiten de manera iterativa hasta un número especificado de ciclos (48 bloques) (Jumper et al., 2021).

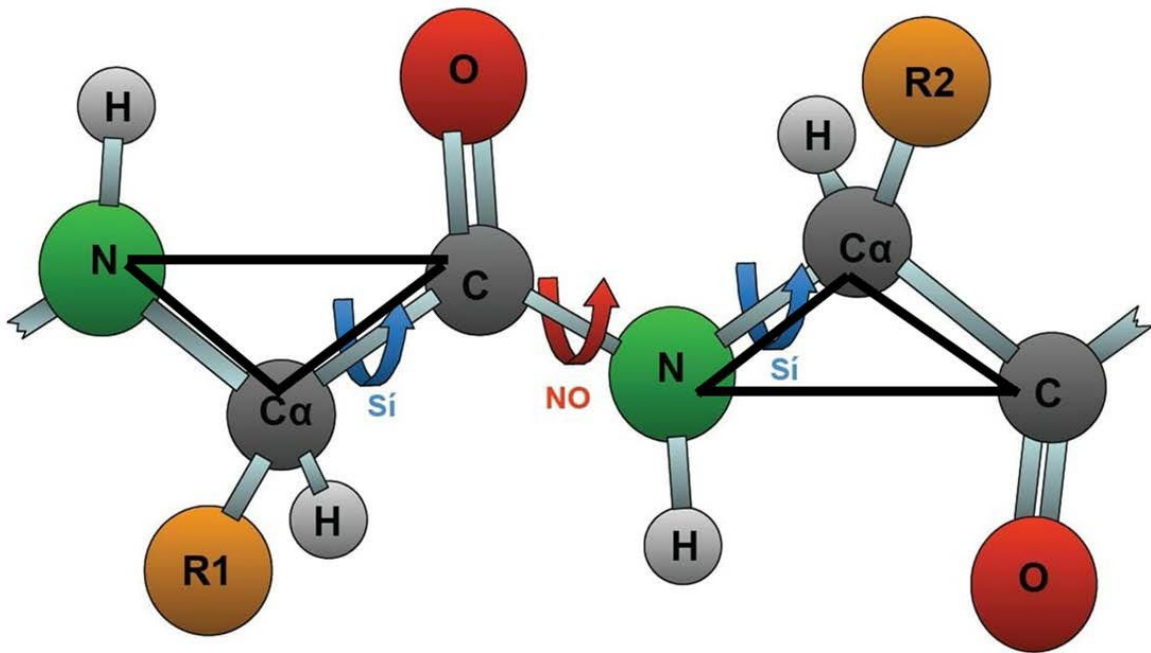
Terminadas las iteraciones, la información pasa a la tercera fase (Figura 9C), denominada, módulo estructural, el cual genera la estructura tridimensional. A diferencia de las metodologías anteriores, esta red no utiliza ningún algoritmo de optimización y genera una estructura estática y final en un solo paso. El resultado final es una larga lista de coordenadas cartesianas que representan la posición de cada átomo de la proteína, incluyendo las cadenas laterales.

Todos estos procesos se ejecutan gracias a una arquitectura de red neuronal en la que cuando se resuelve cada subtarea, se pasa a un nuevo nodo de la red. Además, cuando ha generado una estructura determinada, vuelve a lanzar esta información al *Evoformer*, lo que permite refinar la estructura de manera iterativa hasta obtener una estructura óptima.

La innovación disruptiva de este sistema reside principalmente en el *Evoformer*, cuyo funcionamiento se detalla a continuación. La idea central es que la información fluye de ida y vuelta a lo largo de la red. Antes del AlphaFold2, la mayoría de los modelos de aprendizaje profundo tomaban un MSA y generaban alguna inferencia sobre la proximidad geométrica. Por lo tanto, la información geométrica era un producto de la red. En el *Evoformer*, la representación de pares es tanto un producto como una capa intermedia. En cada ciclo, el modelo utiliza la hipótesis estructural actual para mejorar la evaluación del alineamiento de secuencias múltiples, lo que a su vez conduce a una nueva hipótesis estructural, y así sucesivamente. Ambas representaciones de secuencia y estructura, intercambian información hasta que la red llega a una inferencia sólida. El primer paso en la red consiste en definir incrustaciones para el MSA y las estructuras similares, tomándolas como variables discretas. Las redes neuronales, por otro lado, son dispositivos intrínsecamente continuos que se basan en la diferenciación para aprender de su conjunto de algoritmos de entrenamiento. Una incrustación permite la transformación de una variable discreta a un espacio continuo para que la red pueda ser entrenada.

El flujo de información en la red no es lineal, sino que está dirigido por un mecanismo de atención que focaliza la potencia de la red en las subtarefas más importantes (Jumper et al., 2021). El *Evoformer* utiliza en realidad dos transformadores (arquitectura de dos torres), con un canal de comunicación entre ambos. Cada torre está especializada en el tipo particular de datos que está analizando, ya sea un MSA o una representación de interacciones entre pares de aminoácidos. También incorporan la información de la representación contigua, lo que permite un intercambio

regular de información y un refinamiento iterativo. El transformador del MSA calcula la atención sobre una matriz muy grande de símbolos de proteínas. Para reducir el costo computacional, que de otro modo sería infinito, la atención se factoriza en componentes por fila y por columna. Es decir, la red primero calcula la atención en la dirección horizontal, lo que le permite identificar qué pares de aminoácidos están más relacionados y después en la dirección vertical, determinando qué secuencias son más informativas. El otro transformador trabaja sobre la representación de pares. En esta red, la atención se organiza en términos de triángulos de residuos. Computa los aminoácidos como un triángulo flexible, en el que en uno de los vértices se encuentran el carbono en alfa, y en los otros dos el nitrógeno y el carbono del grupo amida del enlace peptídico (Figura 10) cumpliendo siempre la desigualdad triangular, uno de los axiomas de los espacios métricos. Tras las 48 iteraciones, la red ha generado el modelo de interacciones y puede construir la estructura.



**Figura 10.** Representación esquemática del concepto triángulos de residuos. Fuente: elaboración propia.

El módulo de estructura considera la proteína como un gas de residuos. Cada aminoácido se modela como un triángulo que representa los tres átomos del esqueleto del enlace peptídico. Estos triángulos flotan en el espacio y son movidos por la red neuronal para formar la estructura. Estas transformaciones están parametrizadas como matrices afines, que son una forma matemática de representar las traslaciones y rotaciones en una sola matriz  $4 \times 4$ . La parte  $3 \times 3$  corresponde a una matriz de rotación, y la columna  $1 \times 3$  es el vector de desplazamiento. La multiplicación de una matriz afín por un vector es equivalente a desplazar y posteriormente rotar dicho vector. Estos movimientos son resueltos gracias un mecanismo de atención diferente, denominado punto de atención invariante (Lau et al., 2023). Se basa en que la red sabe que, si cualquier rotación o

traslación posible de los datos conduce a la misma respuesta, necesitará mucha menos información para alejarlo de los modelos incorrectos y, por lo tanto, podrá aprender más. Además, este sistema es capaz de resolver la posición de las cadenas laterales con mucha mayor precisión que las metodologías anteriores. Esto lo consigue porque sus posiciones están parametrizadas por una lista de ángulos de torsión, que son predichos por la red, y luego implementados con subrutinas de análisis geométrico estándar (Skolnick et al., 2021). La red implementa un sistema de control de calidad sobre la predicción que realiza, fácilmente identificable porque sobreimpresiona un código de colores en la estructura de la proteína, permitiendo distinguir fácilmente las zonas resueltas con mayor confianza. Este control de calidad se resume en el cálculo de un nuevo parámetro diseñado por el equipo de Deep Mind (pLDDT), que es muy similar al RMSD (*Root Mean Square Deviation*) en el que la desviación cuadrática media de las raíces se refiere a la medida de la diferencia entre las posiciones tridimensionales de átomos o moléculas en dos estructuras comparadas (Kufareva & Abagyan, 2012), y es invariante a las torsiones de los enlaces químicos. En cambio, pLDDT lo tiene en cuenta, de manera que evita la predicción de proteínas con la quiralidad incorrecta (Roney & Ovchinnikov, 2022).

### 1.7.3 Aplicaciones de la inteligencia artificial a la resolución de la estructura de los anticuerpos

El análisis de los repertorios de anticuerpos puede ampliar nuestra comprensión de la respuesta inmunitaria adaptativa. Sin embargo, los datos de secuencia por sí solos ofrecen solo una visión parcial del repertorio inmunológico. Las interacciones que facilitan la unión de antígenos están determinadas por la estructura de un conjunto de seis bucles que conforman los CDRs. El modelado preciso de estos bucles CDR proporciona conocimientos sobre el mecanismo de unión y posibilita el diseño racional de anticuerpos específicos (Adolf-Bryfogle et al., 2018). Cinco de los bucles CDR tienden a adoptar estructuras canónicas que pueden predecirse eficazmente mediante similitud de secuencia (Almagro et al., 2014). Sin embargo, el moldeado del tercer bucle CDR de la cadena pesada (HCDR3) ha demostrado ser un desafío debido a su mayor diversidad, tanto en secuencia como en longitud (Ruffolo et al., 2020). Además, la posición HCDR3 en la interfaz entre las cadenas pesada y ligera hace que su conformación dependa de la orientación de las cadenas (Marze et al., 2016). Algunos autores han encontrado diversas inconsistencias entre las estructuras predichas de los anticuerpos, y los datos obtenidos con técnicas experimentales, como la difracción de rayos X. Por ejemplo, la colocación correcta de D-aminoácidos, enlaces amida en conformación *cis* y/o no planos, *cis*-prolinas y la predicción de interacciones por fuerzas de Van der Waals (Fernández-Quintero et al., 2023). Teniendo en cuenta su papel central en la unión con el antígeno, los avances en la predicción de las estructuras del HCDR3 son fundamentales para comprender las interacciones anticuerpo-antígeno.

La implementación de la inteligencia artificial basada en el procesamiento del lenguaje natural ha supuesto una auténtica revolución en el desarrollo de metodologías de predicción de la estructura

tridimensional de las proteínas a partir de su secuencia. Sin embargo, la resolución de las estructuras de los anticuerpos continúa siendo una de las principales limitaciones de AlphaFold2 (Xu et al., 2022).

Debido a estas limitaciones, se han desarrollado varios métodos de aprendizaje profundo específicos para anticuerpos como DeepAb (Ruffolo et al., 2022) y ABlooper (Abanades et al., 2022), que han mejorado significativamente la precisión en el modelado de los bucles CDR, incluido el desafiante HCDR3 (Alford et al., 2017). DeepAb predice un conjunto de restricciones geométricas entre residuos que se proporcionan a Rosetta (servidor de modelado de proteínas por métodos clásicos) para producir una estructura completa del FV. ABlooper predice las estructuras de los bucles CDR de manera integral, al tiempo que proporciona una estimación de la calidad del bucle. Otra herramienta, NanoNet (Cohen et al., 2022), ha sido entrenada específicamente para la predicción de anticuerpos de cadena única (*nanobodies*) y ofrece predicciones muy rápidas. Aunque son efectivos, ciertas decisiones de diseño limitan la utilidad de cada uno de los modelos. Las predicciones de DeepAb son relativamente lentas (10 minutos por secuencia), no pueden incorporar eficazmente datos de estructuras de referencia, y ofrecen poca información sobre la calidad esperada. ABlooper, aunque más rápido e informativo, depende de herramientas externas para el modelado, no puede incorporar referencias de bucles CDR y no admite *nanobodies* (Ruffolo et al., 2023).

La abundancia de datos de repertorios inmunes, proporcionados por experimentos de secuenciación masiva, ha permitido el desarrollo de lenguajes específicos para anticuerpos. Se ha demostrado que los modelos entrenados con lenguaje natural aprenden representaciones significativas de secuencias de repertorios inmunitarios (Olsen et al., 2022), e incluso se han utilizado para humanizar anticuerpos (Prihoda et al., 2022). También se han desarrollado modelos generativos entrenados con secuencias para generar repertorios *de novo* de anticuerpos de alta calidad (Shin et al., 2021).

Entre las metodologías de predicción de la estructura de anticuerpos mediante la implementación de inteligencias artificiales entrenadas específicamente para este tipo de moléculas destacan ABodyBuilder-ML e IgFold.

ABodyBuilder-ML (*Machine learning*) es la evolución de un programa anterior ABodyBuilder. Se trataba de una metodología automatizada cuyo algoritmo seguía los 4 pasos estándar: selección de estructuras de referencia, predicción de orientación VH-VL, modelado de CDR y predicción de cadenas laterales. A continuación, ABodyBuilder anotaba la calidad del modelo como una probabilidad de que un componente del anticuerpo (por ejemplo, el bucle LCDR3) se modelase dentro de un umbral determinado de RMSD (Leem et al., 2016). Este programa fue uno de los más utilizados, gracias a su rapidez y facilidad de uso a través de una página web. Sin embargo, al desarrollarse los sistemas de aprendizaje automático a partir del año 2020, pronto quedó

obsoleto, por lo que se actualizó el programa, reemplazando el modelado basado en estructuras de referencia por una red neuronal que computa gráficos equivariantes.

El modelo se entrena en la posición de los átomos en el esqueleto del enlace peptídico de los seis bucles CDR, entre dos residuos de anclaje en cada extremo. Las redes gráficas equivariantes requieren una geometría de inicio, por lo que se genera una geometría de entrada no descriptiva, espaciando uniformemente cada residuo del CDR en una línea recta entre los residuos de anclaje. Al modelo se le proporcionan cuatro tipos diferentes de características por nodo, lo que resulta en un vector de 41 dimensiones. Estos incluyen un vector codificado que describe el tipo de aminoácido, tipo de átomo y a qué bucle pertenece el residuo. Además, se proporcionan incrustaciones posicionales a cada residuo que describen cómo de cerca están de los anclajes. De esta manera, y tras varias iteraciones, resuelve la estructura del anticuerpo (Abanades et al., 2022).

Otra propuesta es IgFold, que trata de solucionar el problema que supone el número limitado de estructuras de anticuerpos determinadas experimentalmente del orden de miles (Dunbar et al., 2014) para el entrenamiento de un predictor efectivo. En ausencia de datos estructurales, los modelos de lenguaje autosupervisados brindan un marco ideal para extraer patrones de un número significativamente mayor (miles de millones) de secuencias naturales de anticuerpos, identificadas por estudios de secuenciación de repertorios inmunes (Kovaltsuk et al., 2018). Para ello, se desarrolló un transformador preentrenado con 558 millones de secuencias de anticuerpos naturales, para generar incrustaciones para la predicción de estructuras. Para predecir las coordenadas atómicas a partir de las incrustaciones de secuencia, se genera una representación gráfica de la estructura de anticuerpos, donde cada residuo es un nodo y la información se transmite entre todos los pares de residuos. Además, se implementan matrices de atención que codifican las vías de información residuo-residuo aprendidas por el modelo. A continuación, se integran las incrustaciones de secuencia de cada cadena. El modelo de predicción de estructuras está compuesto por cuatro capas de transformador de gráficos entrelazadas a través de la capa multiplicativa triangular propuesta para AlphaFold2 (Ruffolo et al., 2023).

#### 1.7.4 Estudio de la interacción antígeno-anticuerpo mediante la tecnología de *docking*

Las interacciones proteína-proteína son la base de la función celular. Para entender la célula y sus procesos, se requiere una comprensión de las interacciones biomoleculares a escala atómica. Importantes iniciativas genómicas y proteómicas están trabajando hacia este objetivo. Sin embargo, mientras se predice que el tamaño del proteoma humano es del orden de 20.000 proteoformas únicas, se estima que el interactoma, la red de todas las proteínas interaccionantes, está en el rango de aproximadamente 650.000 unidades (Aebersold et al., 2018), con niveles adicionales de complejidad, vinculados a la dinámica de las agrupaciones, la localización y el tiempo de expresión de sus componentes en la célula.

La cristalización de proteínas es muy compleja y costosa y tampoco es sencillo estudiarlas mediante Resonancia Magnética Nuclear (RMN). Para estudiar las interacciones y conocer su estructura, se han ideado enfoques computacionales complementarios a los métodos experimentales. Uno de estos métodos es el acoplamiento proteína-proteína mediante metodologías de *docking*, que tiene como objetivo predecir la estructura de un complejo a partir de modelos atómicos de las subunidades no enlazadas.

El objetivo del *docking* es lograr una conformación optimizada tanto para la proteína como para el ligando (anticuerpo y antígeno en este caso) determinando la posición entre la proteína y el ligando de manera que su energía libre se minimice (Kuntz et al., 1982). El reconocimiento molecular juega un papel clave en promover procesos bimoleculares elementales, como la interacción enzima- sustrato, fármaco-proteína y fármaco-ácido nucleico o proteína-proteína (DesJarlais et al., 1988). Esto se consigue mediante la comprensión detallada de los principios universales que rigen la naturaleza de las conexiones (fuerzas de van der Waals, enlaces de hidrógeno, y electrostáticos) entre dos moléculas que interactúan.

Las metodologías de *docking* se clasifican tradicionalmente en dos tipos: rígidos y flexibles. Los modelos rígidos obvian la flexibilidad de una molécula y buscan la conversión en el espacio 3D de una de las moléculas, con un ajuste más favorable con las otras moléculas en términos de una función de puntuación de energía mínima (Dnyandev et al., 2021). Estos métodos fueron los primeros en utilizarse y constituyeron la prueba de concepto del *docking*.

El *docking* flexible tiene en cuenta las diferentes conformaciones de receptor y ligando y como se adaptan entre sí al interactuar (Dar, 2017). Se han propuesto diversos métodos para resolver la interacción entre dos moléculas. El método más habitual es la combinación de la aproximación de Monte-Carlo combinada con el criterio metrópolis, que comienza generando una configuración preliminar del ligando en un lugar energéticamente favorable (como el centro activo de una enzima o el paratopo de un anticuerpo) determinado por una conformación aleatoria, con determinada conversión y rotación. A partir de este estado inicial, va desplazando el ligando y evaluando cada desplazamiento (Kurcinski et al., 2021). El sistema selecciona la mejor solución aplicando el criterio metrópolis que se basa en que si la puntuación de la nueva solución es mejor que la anterior, se acepta de inmediato. Si la configuración no es nueva, se aplica una función de probabilidad basada en una distribución de Boltzmann. Si la solución pasa la prueba de la función de probabilidad se acepta, en caso contrario, la configuración se descarta (Chen & Roux, 2015). Otra metodología también utilizada es el método basado en fragmentos, en el que el ligando se separa en varios fragmentos, se realiza el *docking* de cada uno de ellos, y se integran todos los datos para obtener el modelo de interacción completo. Este método es especialmente útil para trabajar con moléculas muy grandes (Sun et al., 2021).

Otra forma de clasificar los modelos de *docking*, se basa en la cantidad de información que se introduce en el modelo para realizar la predicción. El *docking* a ciegas se realiza cuando no se dispone de ninguna información sobre los sitios catalíticos o de unión en una proteína. Generalmente se hace con proteínas recién descubiertas. El *docking* específico de sitio se puede obtener cuando se conocen los sitios de unión de esa proteína, optimizando los recursos computacionales, y centrando el modelo de interacción en esas zonas de las proteínas. El *docking* a ciegas es una herramienta poderosa para descubrir interacciones desconocidas (Huang et al., 2023).

La principal función de los anticuerpos es el reconocimiento con elevada afinidad de antígenos, por lo que son moléculas perfectas para llevar a cabo estudios de *docking*. Mientras que para el anticuerpo se puede extraer fácilmente la interfaz de unión a partir de la secuencia y, en particular, de los CDRs (Sela-Culang et al., 2013), predecir el/los epítipo/s en el antígeno es un problema más desafiante. A pesar de muchos esfuerzos para desarrollar métodos fiables para predecir los epítopos específicos, la mayoría de los enfoques publicados hasta la fecha siguen siendo muy limitados en términos de precisión. Este problema deriva en que la interacción anticuerpo-antígeno es compleja y habitualmente ocurre tridimensionalmente (los epítopos conformacionales son mucho más comunes que los epítopos lineales) (Forsström et al., 2015).

Dada una secuencia o estructura de un antígeno, la predicción *in silico* de epítopos identificados por células B tiene como objetivo identificar un conjunto de residuos en el antígeno capaces de unirse a un anticuerpo. La mayoría de los métodos operan sin contar con la información de los anticuerpos, buscando identificar todos los posibles sitios de unión (Kringelum et al., 2012). Sin embargo, intentar mapear todos los epítopos podría no resultar óptimo, ya que en algunos antígenos, como la lisozima de huevo de gallina, se han encontrado complejos con muchos anticuerpos diferentes, que se unen a áreas distintas, lo que significa que la mayor parte de la superficie de la lisozima constituye parte de algún epítipo (Sela-Culang et al., 2013). Además, se ha demostrado que dos anticuerpos terapéuticos diferentes, Gevokizumab y Canakinumab, activan dos vías distintas al unirse a diferentes epítopos de IL-1 $\beta$  (Blech et al., 2013).

La predicción computacional de epítopos de células B proporciona información sobre las regiones del antígeno que se unen al anticuerpo, pero no contribuye directamente al conocimiento de los residuos específicos del anticuerpo que deben mutarse para modificar su función. Este problema puede abordarse mediante el *docking* anticuerpo-antígeno, que dada la estructura del anticuerpo y el antígeno, proporciona una lista de posibles orientaciones de las dos moléculas con respecto una de la otra. El *docking* anticuerpo-antígeno requiere una metodología diferente de la utilizada para otras moléculas. Los complejos enzima-inhibidor generalmente presentan una excelente complementariedad en su superficie, con el inhibidor convexo coincidiendo con el sitio de unión cóncavo de la enzima o viceversa. La mayoría de las interfases enzima-inhibidor nativas también

tienen interacciones hidrofóbicas y polares favorables, que facilitan el acoplamiento y la puntuación en los modelos de validación (Vajda, 2005). Por el contrario, las interfases en complejos anticuerpo-antígeno son en su mayoría planas y menos hidrofóbicas. La interfaz plana implica que apenas se puede buscar complementariedad espacial entre los componentes del acoplamiento (Kapingidza et al., 2020). Además, las interacciones polares son más sensibles a las posiciones atómicas que las hidrofóbicas, y, por lo tanto, la puntuación basada en funciones de energía mecánica se vuelve menos fiables debido a las inevitables diferencias conformacionales entre las estructuras de proteínas libres y unidas. Por lo tanto, se requiere un sistema de puntuación asimétrico que tenga en cuenta todas estas discrepancias (Brenke et al., 2012).

Para poder predecir epítomos exactamente es necesario implementar algoritmos que calculen qué zonas del antígeno presentan una elevada probabilidad de contacto con el anticuerpo, a partir de datos estructurales, y después evaluar cuál de estas zonas se ha predicho con mayor significancia. Los epítomos candidatos se evalúan mediante un ajuste geométrico y una puntuación específica de interacción anticuerpo-antígeno. El ajuste geométrico se puede calcular enumerando todos los posibles contactos entre el conjunto de residuos de epítomos putativos y las CDRs, y evaluando qué pares de contactos anticuerpo-antígeno pueden satisfacerse simultáneamente. La puntuación final del epítomo es una suma de todos los posibles contactos entre los átomos del epítomo dado y las CDRs, ponderados por el número de otros contactos que pueden producirse simultáneamente (Krawczyk et al., 2014).

Una vez encontrados los aminoácidos que más probablemente pertenezcan a un epítomo, y definidos correctamente los CDRs en la estructura del anticuerpo que se está estudiando, se puede dilucidar la interacción entre las dos moléculas mediante metodologías de *docking* dirigido y centrado en las zonas detectadas como de posible interacción. Esto permite obtener modelos de interacción que suelen presentar mejor resolución (medida mediante RMSD y/o *z-score*, el número de desviaciones típicas) que los modelos a ciegas, y con una mejor eficiencia computacional (Ambrosetti et al., 2020).



## 2. JUSTIFICACIÓN, HIPÓTESIS Y OBJETIVOS



## 2. JUSTIFICACIÓN, HIPÓTESIS Y OBJETIVOS

El gluten está compuesto por un conjunto de proteínas presentes de forma mayoritaria en el endospermo de los granos de trigo (incluidas variedades como la espelta y el kamut), cebada y centeno (Fasano & Catassi, 2012). Estos cereales son ampliamente utilizados en la industria alimentaria debido a las propiedades que el gluten aporta a la masa de harina (elasticidad, consistencia y esponjosidad). Sin embargo, los cereales con gluten son considerados como importantes alérgenos ya que contienen péptidos altamente inmunogénicos que con frecuencia causan reacciones adversas en personas sensibilizadas (Shewry, 2019).

Se han identificado diversas patologías relacionadas con el gluten, entre las que destaca por su elevada prevalencia (estimada en el 1 % de la población), la enfermedad celíaca (Sapone et al., 2012). Aunque se continúan investigando diferentes alternativas terapéuticas para tratar la enfermedad celiaca (Kurada et al., 2016), ninguna ha demostrado tener una efectividad equiparable a una dieta sin gluten (Wahab et al., 2002). Basándose en el establecimiento de un umbral de gluten seguro para pacientes con enfermedad celiaca (50 mg de gluten al día), la legislación europea actual indica que los alimentos que contengan menos de 20 mg/kg de este componente pueden etiquetarse como “libres de gluten” (Diario Oficial de la Unión Europea, 2014). En este escenario, resulta esencial disponer de métodos rápidos y eficaces de detección de gluten en alimentos con el propósito final de fortalecer los sistemas de trazabilidad y control de alérgenos de la industria alimentaria. Solo de esta manera se puede garantizar un etiquetado preciso que proporcione información fiable al consumidor alérgico, contribuyendo así a su seguridad y bienestar.

Entre las diferentes técnicas analíticas disponibles para detectar la presencia de gluten, destacan las que se basan en la detección de proteínas mediante inmunodetección, concretamente las metodologías ELISA. Actualmente existen en el mercado inmunoensayos para la detección cualitativa y cuantitativa de gluten basados en anticuerpos monoclonales obtenidos mediante inmunización de animales de experimentación. Entre todos ellos, el ELISA sándwich basado en

el anticuerpo monoclonal R5, está considerado como método de análisis Tipo I por el *Codex Alimentarius* para la detección en gluten en alimentos (Valdés et al., 2003).

A pesar de los grandes avances legislativos en materia de seguridad alimentaria y de la disponibilidad en el mercado de una amplia variedad de sistemas de detección, el análisis de gluten en alimentos sigue presentando diversos desafíos. Varios estudios comparativos independientes han revelado considerables discrepancias en la cuantificación de gluten al emplear diferentes kits de ELISA disponibles en el mercado (Bugyi et al., 2013). Por su propia definición, la palabra gluten abarca a una amplia variedad de proteínas, lo que añade mayor complejidad al proceso de análisis. Se ha comprobado que, a pesar de compartir ciertas similitudes, las secuencias de estas proteínas pueden diferir significativamente. Además, se ha demostrado que la composición de las proteínas del gluten puede variar considerablemente incluso entre cereales de la misma especie, y verse influenciada por diversos factores como la variedad del cereal, su procedencia e incluso la forma de cultivo (Wieser et al., 2023). Debido a esta gran variabilidad, es necesario continuar investigando en el desarrollo de métodos de detección de gluten que permitan obtener resultados más consistentes. En este contexto, resulta fundamental avanzar en la generación de nuevos anticuerpos frente al gluten con características diferentes a los anticuerpos actualmente disponibles. La implementación de estas nuevas sondas de afinidad en diversos sistemas de detección ya sea de manera individual o en combinación con otras sondas, se presenta como una estrategia esencial para abordar las limitaciones inherentes a la cuantificación de gluten en alimentos.

En ese sentido, la búsqueda de nuevos anticuerpos mediante técnicas de evolución dirigida de proteínas, como la tecnología de presentación en fagos o *phage display* (Smith et al., 2019) puede representar una alternativa eficaz para satisfacer estas necesidades. El principal prerrequisito para la generación exitosa de anticuerpos recombinantes mediante esta tecnología recae en la disponibilidad de repertorios de anticuerpos de alta calidad a partir de los cuales guiar la selección de sondas específicas frente al gluten. Estos repertorios se construyen frecuentemente a partir del material genético obtenido de animales inmunizados frente al antígeno de interés, ya que permiten obtener sondas de alta afinidad debido a que estos anticuerpos han sufrido un proceso natural de maduración de la afinidad. Sin embargo, la respuesta generada tras la inyección intramuscular del alérgeno en la inmunización animal puede diferir significativamente de la respuesta que la ingesta de gluten desencadena en un paciente celíaco. Por este motivo, se partió de la hipótesis de que trasladando a los repertorios las características presentes en la respuesta humoral de personas celíacas, sería factible obtener anticuerpos con capacidad de reconocer las fracciones del gluten inmunogénicas para estos individuos. Por tanto, las metodologías desarrolladas se debían basar en anticuerpos recombinantes seleccionados a partir de repertorios generados, parcial o totalmente, de la información genética obtenida de pacientes celíacos. Este innovador enfoque no solo podría contribuir a mejorar la precisión en la detección,

sino que además permitiría un avance significativo en la capacidad de los métodos analíticos para adaptarse a la complejidad y diversidad de las proteínas gluten presentes en los alimentos.

Teniendo en cuenta estas consideraciones, esta tesis doctoral tiene como objetivo principal el desarrollo de nuevos sistemas de inmunodetección de gluten basados en anticuerpos recombinantes obtenidos mediante *phage display*. Con este fin, se propone la generación de dos repertorios de fago-anticuerpos tipo Fab (fragmento de unión a antígeno) a partir de linfocitos de sangre periférica (PBLs) de pacientes celíacos, siguiendo dos estrategias diferentes. El primero de los repertorios, de naturaleza semiinmune, se generará a través de un proceso de ingeniería de proteínas con el fin de mejorar un anticuerpo de dominio único frente al gluten (dAb8E) previamente desarrollado. Por el contrario, para la búsqueda de nuevas sondas de afinidad se propone la construcción de un segundo repertorio completamente inmune donde se traslade la respuesta humoral de los pacientes celíacos a ambas cadenas del Fab.

En este contexto, para abordar este objetivo global, se establecieron los objetivos parciales que se enumeran a continuación:

1. Selección de pacientes celíacos y obtención de muestras biológicas para su empleo como material de partida en la construcción de repertorios de anticuerpos recombinantes.
2. Construcción de un repertorio de Fab semiinmune combinando una cadena pesada basada en un dominio variable frente a gliadina (dAb8E) con cadenas ligeras obtenidas de linfocitos de sangre periférica (PBLs) de pacientes celíacos.
3. Construcción de un repertorio de Fab inmune mediante amplificación tanto de cadenas pesadas como de cadenas ligeras de PBLs de pacientes celíacos.
4. Selección por afinidad de variantes de Fab con capacidad de unión a gliadina a partir de ambos repertorios y caracterización de los fago-anticuerpos seleccionados.
5. Producción soluble a escala semiindustrial de los mejores candidatos de anticuerpos recombinantes y desarrollo de metodologías ELISA para la detección de gluten en alimentos.
6. Elucidación *in silico* de la estructura tridimensional de los Fab con mejores características para la detección de gluten y estudio de su interacción con los antígenos de interés.
7. Caracterización de ambos repertorios y de los procesos de selección por afinidad llevados a cabo mediante secuenciación masiva de ADN.





# 3. METODOLOGÍA RESUMIDA



## 3. METODOLOGÍA RESUMIDA

Para el desarrollo y consecución de los objetivos propuestos, durante la realización de la presente tesis doctoral se emplearon numerosas metodologías inmunológicas, bioquímicas, genéticas, genómicas y proteómicas, las cuales se describen brevemente a continuación.

### 3.1 PREPARACIÓN DE MUESTRAS Y ANTÍGENOS

En este trabajo, se utilizó la gliadina-PWG (elaborada por el Grupo de Trabajo en Prolaminas) como material de referencia para la detección de gluten en alimentos. Siguiendo la recomendación del proveedor, este material se disolvió a 1 mg/mL en una solución etanol/agua al 60 % (v/v).

Todas las muestras alimenticias analizadas (diferentes cereales con y sin gluten, especies heterólogas y productos comerciales) se molieron individualmente en un molinillo analítico empleando 50 g de muestra. Se realizaron extractos etanólicos de diferentes cereales (trigo, centeno, cebada, maíz y arroz, entre otros), a partir de 250 mg de las muestras finamente molidas con 10 mL de una solución etanol/agua al 60 % (v/v), seguido de agitación durante 30 minutos a temperatura ambiente y centrifugación a 867 g durante 15 minutos. Para realizar los ensayos de inmunotransferencia (o *Western blot*) se realizaron los extractos de forma análoga a la descrita anteriormente, pero partiendo de 200 mg de muestra con 1 mL de isopropanol/agua al 50 % (v/v).

Para evaluar las técnicas de ELISA desarrolladas, se analizaron mezclas binarias con concentraciones crecientes de un homogeneizado de trigo, cebada y centeno en una matriz sin gluten a base de harina de arroz. También se incluyeron en los análisis diversos productos comerciales sometidos a diferentes tipos de procesado. En estos casos, las proteínas similares al gluten se extrajeron a partir de 250 mg de las muestras finamente molidas con 2,5 mL de la solución de extracción de gluten Ingezim. Esta mezcla se homogeneizó y se incubó a 50 °C durante 40 minutos antes de agregar 7,5 mL de una solución etanol/agua al 80 % (v/v). Finalmente, las muestras se incubaron durante 1 hora en un agitador vertical y se centrifugaron a 2000 g durante 10 minutos a temperatura ambiente para obtener los extractos del sobrenadante.

### 3.2 SELECCIÓN DE PACIENTES CON NIVELES ELEVADOS DE BIOMARCADORES DE LA RESPUESTA HUMORAL CELÍACA

La selección de pacientes se realizó en el Hospital Universitario Fundación Jiménez Díaz (FJD, Madrid, España) tras la aprobación del estudio por parte de su comité ético. Los facultativos colaboradores de los servicios de medicina digestiva e inmunología llevaron a cabo el reclutamiento de los donantes entre aquellos pacientes que acudieron a consulta con una sospecha inicial de padecer la enfermedad celíaca. Los pacientes identificados como potenciales donantes que no habían comenzado una dieta sin gluten (debido a que la respuesta humoral desaparece rápidamente cuando se elimina el gluten de la dieta), se sometieron a un análisis para evaluar la respuesta humoral celíaca que presentaban. Los biomarcadores utilizados fueron la titulación de anticuerpos IgG e IgA dirigidos frente a la transglutaminasa tisular y frente a péptidos desaminados de gliadina.

Finalmente, aquellos pacientes que presentaron altos niveles de estos marcadores humorales según las guías clínicas utilizadas en el Hospital FJD, y que firmaron el consentimiento informado para participar en el estudio, realizaron una donación de 350 mL de sangre. Además, se realizó el isotipado (niveles generales de IgA e IgG) y subisotipado (niveles específicos de IgG1, IgG2, IgG3 e IgG4) de su respuesta humoral frente a gliadina nativa mediante técnicas de ELISA (*Enzyme-Linked ImmunoSorbent Assay*, ensayo por inmunoadsorción ligado a enzimas).

### 3.3 EXTRACCIÓN DE ARN Y RETROTRANSCRIPCIÓN DEL MATERIAL BIOLÓGICO EMPLEADO PARA LA CONSTRUCCIÓN DE REPERTORIOS DE FRAGMENTOS DE ANTICUERPO TIPO FAB

Las células mononucleares de sangre periférica (PBMC, *Peripheral Blood Mononuclear Cells*) se aislaron de las muestras mediante el procedimiento basado en el medio Ficoll-Paque® que genera un gradiente de densidad. A partir de las PBMC aisladas, se extrajo el ARN total mediante un protocolo clásico de extracción con el reactivo TriZol LS. Antes de emplear el ARN como hebra molde para llevar a cabo la reacción de retrotranscripción (RT-PCR), se comprobó su concentración y calidad usando un *microarray*. La RT-PCR se realizó de manera inespecífica mediante el uso de oligo-dT-20 como cebador. De esta manera, el ADN complementario (ADNc) se utilizó como molde para las posteriores reacciones de amplificación específica de los genes que codifican las diferentes cadenas de anticuerpos.

### 3.4 AMPLIFICACIÓN DE LAS SECUENCIAS CODIFICANTES DE CADENAS PESADAS Y LIGERAS DE ANTICUERPOS

A lo largo de este trabajo de investigación se construyeron dos repertorios de Fab diferentes. El primero de los repertorios, surgió de la combinación de cadenas pesadas de origen semisintético con un subrepertorio de cadenas ligeras inmunes obtenidas del material biológico de los donantes celíacos, dando lugar a un repertorio semiinmune. El segundo repertorio, por el contrario, fue de naturaleza completamente inmune, dónde tanto las cadenas pesadas como las ligeras procedieron de la amplificación del ADNc obtenido de los donantes celíacos.

Las cadenas pesadas semisintéticas del repertorio semiinmune se generaron a partir de un dAb parental (un fragmento de dominio único denominado dAb8E) que presentaba una elevada afinidad frente al gluten. Este dAb fue aislado en trabajos previos de nuestro grupo de investigación a través del cribado frente a un péptido inmunogénico del gluten de un repertorio semisintético comercial basado en los genes humanos VH3-23. Por tanto, para la generación del repertorio semiinmune se empleó el dAb8E como dominio VH al que se añadió mediante PCR de extensión solapada (OE-PCR) diferentes dominios CH1 correspondientes a cada subisotipo de IgG. Esta estrategia permite la unión de fragmentos de ADN solapantes, de tal forma que el extremo 3' del primero (dominio VH) se complementa con el 5' del segundo (dominio CH1).

Para la obtención de las cadenas de inmunoglobulinas de origen inmune, se amplificaron de forma independiente las cadenas ligeras y las cadenas pesadas empleando una colección con diferentes parejas de cebadores. La amplificación se realizó mediante dos rondas de PCR consecutivas. En la primera ronda de PCR se utilizaron parejas de cebadores que hibridaban en 5' con secuencias específicas de cada familia de genes V y en 3' con secuencias característica de cada subisotipo de inmunoglobulina. De esta forma, se amplificaron las cadenas de interés añadiendo en cada extremo una diana de restricción diferente. Los productos de PCR obtenidos se purificaron mediante electroforesis en gel de agarosa, y se sometieron a una segunda ronda de PCR para incluir sitios de restricción mejorados, alargando las regiones de ADN adyacentes a los sitios de restricción para facilitar el posterior clonaje.

### 3.5 CONSTRUCCIÓN DE LOS REPERTORIOS DE ANTICUERPOS RECOMBINANTES Y SELECCIÓN POR AFINIDAD DE VARIANTES CON CAPACIDAD DE UNIÓN A GLIADINA

Los productos de PCR se clonaron en el vector pComb3X, un fagémido que permite la expresión de fragmentos de anticuerpos unidos a la proteína III de la cápside del fago auxiliar, en nuestro caso el bacteriófago VSC-M13. Para construir los repertorios, se realizó una clonación secuencial, generando primero un subrepertorio únicamente con las cadenas ligeras que se transformó en *E. coli* XL1-Blue mediante electroporación. Las células transformadas se cultivaron en un medio con carbenicilina, y a partir de este cultivo se purificaron los fagémidos que contenían el subrepertorio de cadenas ligeras. A continuación, se clonaron las cadenas pesadas que se habían obtenido a partir de dos orígenes diferentes para generar de este modo los dos repertorios de Fab propuestos (repertorio semiinmune e inmune).

Cada repertorio se transformó en células *E. coli* XL1-Blue electrocompetentes. Las células transformadas se recuperaron primeramente en medio SOC, sembrando posteriormente diluciones seriadas de las bacterias transformadas para conocer la eficiencia de la transformación y calcular el tamaño del repertorio (RZ). Posteriormente, el cultivo se expandió en medio *Super-Broth* (SB: 30 g/L triptona, 20 g/L extracto de levadura, 10 g/L MOPS, pH 7) empleando concentraciones crecientes de carbenicilina según progresaba el crecimiento bacteriano. Después, se añadieron  $10^{12}$ - $10^{13}$  partículas del fago auxiliar VSC-M13 y, tras 2 h de incubación, se incorporó kanamicina al medio para seleccionar aquellas células transformadas con el fagémido e infectadas con el fago auxiliar. Finalmente, la expresión en el sobrenadante de los fago-Fab se realizó incubando los cultivos a 37 °C durante 16 h.

Al día siguiente los fago-anticuerpos generados se aislaron del cultivo mediante precipitación con PEG/NaCl. Para la selección de los repertorios frente al antígeno diana, se empleó una placa inmunoabsorbente tapizada con el estándar de gliadina-PWG y, seguidamente, bloqueada con BSA al 3 % en PBS (137 mM NaCl, 2,7 KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1,8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7,4). Una vez enfrentado el repertorio de fago-Fab al antígeno, se realizaron cinco lavados con PBS-T (PBS con 0,05 % del detergente Tween20) para eliminar los fago-Fab que no se unieron de forma específica con el antígeno. A continuación, se eluyeron los clones unidos específicamente mediante un cambio brusco de pH con glicina-HCl (pH 2,2), tras el cual, los fago-Fab eluidos se neutralizaron rápidamente con base-Tris. Estos fago-Fab recuperados de la primera ronda de selección se infectaron en un nuevo cultivo de *E. coli* XL1-Blue, que tras su expansión en medio líquido SB- carbenicilina, siguió un proceso análogo al explicado previamente de infección con el fago auxiliar y precipitación PEG/NaCl. Este proceso cíclico, denominado generalmente como (*bio*)*panning*, se repitió a lo largo de 3-4 rondas sucesivas, aumentando la astringencia de la selección mediante un incremento en el número de lavados con PBS-T (cinco en la primera ronda, diez en la segunda, y quince en las siguientes).

### 3.6 SEGUIMIENTO DEL PROCESO DE PANNING. SELECCIÓN Y SECUENCIACIÓN DE CLONES INDIVIDUALES

El proceso de *panning* se monitorizó tras cada ronda de selección contabilizando el número de bacterias transformadas con los fago-Fab obtenidos tras el proceso de elución. La selección de variantes de fago-Fab frente al antígeno se considera positiva cuando entre una ronda y la siguiente se observa un aumento significativo de transformantes. En estos ensayos, es necesario incluir un control negativo (rondas de selección sin antígeno) para asegurar que no se seleccionen fago-Fab frente a otros elementos del sistema (proteína de bloqueo, soporte, etc...).

Para confirmar el enriquecimiento del repertorio en variantes frente al antígeno, se llevó a cabo un ELISA indirecto policlonal, enfrentando a la gliadina-PWG el conjunto de fago-Fab precipitados tras cada ronda de selección. El sistema se reveló empleando un anticuerpo conjugado dirigido contra la proteína VIII de la cápside del fago. Si el proceso de enriquecimiento ha sido efectivo, se debe observar un aumento de señal en las poblaciones de las últimas rondas de *panning*, representativo de una mayor proporción de fago-Fab de elevada afinidad.

Para llevar a cabo el estudio de clones individuales, se indujo la expresión en el sobrenadante de fago-Fab a partir de colonias aisladas de las últimas rondas de *panning* que presentaron enriquecimiento en fago-Fab frente a la gliadina. Los sobrenadantes inducidos se emplearon como anticuerpo primario en un ELISA indirecto monoclonal (Figura 11A). Aquellos clones que mostraron mayor señal de reconocimiento frente al antígeno diana, sin presentar reconocimiento del agente de bloqueo usado como control negativo, fueron seleccionados para una mejor caracterización de sus propiedades.

Adicionalmente, los fagémidos que codificaban cada Fab individual, se purificaron y se secuenciaron por el método Sanger empleando dos parejas de cebadores (una para obtener la secuencia de las cadenas ligeras y otra para las pesadas). Las secuencias se caracterizaron con las herramientas en línea IMGT/V-QUEST e IgBLAST atendiendo a su composición de genes V(D)J, isotipo, subisotipo y clonotipo, así como según sus características estructurales (bucles CDR).

### 3.7 EVALUACIÓN DE LOS DIFERENTES CANDIDATOS DE FAGO-FAB COMO SONDAS PARA LA DETECCIÓN DE GLUTEN EN ALIMENTOS

Antes de expresar los Fab de forma soluble, se realizó una primera caracterización en formato fago-Fab de los diferentes clones con afinidad por la gliadina identificados en el ELISA monoclonal. Para ello, se realizó una producción y precipitación PEG/NaCl a mayor escala (50 -100 mL de cultivo) para disponer de suficiente fago-Fab para los ensayos de caracterización.

Los fago-Fab se emplearon como anticuerpo primario en formato ELISA indirecto frente a diferentes antígenos. En primer lugar, se obtuvo un curva dosis- respuesta analizando diferentes concentraciones del estándar de gliadina-PWG. Además, se analizaron extractos etanólicos de diferentes harinas de cereales (trigo, cebada, centeno, avena, maíz, arroz, etc.) y de otras especies heterólogas ampliamente empleados en la elaboración de productos alimenticios. Para evaluar la capacidad de detección de gluten en muestras reales, se analizó una mezcla experimental preparada con concentraciones crecientes de un homogeneizado de diferentes cereales con gluten (40 variedades de trigo, 10 variedades de cebada y 10 variedades de centeno) en una matriz a base de harina de arroz certificado sin gluten. Por último, se analizaron productos comerciales con diferente tipo de etiquetado atendiendo al contenido de cereales con gluten.

### 3.8 PRODUCCIÓN SOLUBLE DE FAB RECOMBINANTES FRENTE AL GLUTEN

Los cuatro Fab que mostraron mejores propiedades para la detección de gluten en los ensayos con los fago-Fab se expresaron de forma heteróloga en su formato soluble. Tres de los Fab seleccionados (Fab-C, Fab-H y Fab-E) provenían del repertorio inmune y, solamente uno de ellos (Fab8E-4) del repertorio semiinmune. Los genes sintéticos que codificaban cada Fab se clonaron en el vector de expresión pKKtac, que permitió su expresión optimizada en el sobrenadante bacteriano gracias al péptido señal *pelB* que precede a las secuencias de los Fab. Además, con este vector de expresión, los Fab se expresan fusionados a una cola de histidinas en su extremo carboxilo-terminal que facilita su posterior detección y purificación. Una vez obtenidos los vectores con los insertos apropiados, se transformaron en *E. coli* K12 RV308 químicamente competentes.

Para inducir la expresión de los anticuerpos recombinantes, se inoculó una colonia seleccionada en medio *Terrific Broth* (TB: 12 g/L triptona, 24 g/L extracto de levadura y 4 g/L glicerol en tampón fosfato (0,17 M KH<sub>2</sub>PO<sub>4</sub>, 0,72 M K<sub>2</sub>HPO<sub>4</sub>, pH 7,2)) suplementado con glucosa y ampicilina. Una vez escalado el cultivo hasta 1.8 L, se incubó a 37 °C y 220 rpm hasta alcanzar una densidad óptica de 4. En ese momento, se indujo la expresión del Fab con la adición de

1 mM de IPTG (isopropil- $\beta$ -D-1-tiogalactopiranosido) e incubación a 30 °C y 170 rpm durante 16 h. Al día siguiente, se obtuvo el sobrenadante libre de células mediante centrifugación y se comprobó la expresión de los Fab mediante electroforesis de proteínas e inmunotransferencia (*Western blot*).

Una vez comprobada la correcta producción de los Fab se clarificó el sobrenadante mediante tratamiento con DNAsa, y se empleó un dispositivo de filtración tangencial para intercambiar la solución amortiguadora de pH y concentrar la muestra. El sobrenadante clarificado se purificó en un equipo de cromatografía líquida de proteínas a alta velocidad (FPLC, del inglés *fast protein liquid chromatography*) mediante cromatografía de afinidad con iones metálicos inmovilizados (IMAC, del inglés *immobilized metal affinity chromatography*). El proceso de purificación se monitorizó mediante espectrofotometría a 280 nm y las fracciones recolectadas se analizaron por electroforesis en gel de poliacrilamida con SDS y huella peptídica (espectrometría de masas MALDI-TOF) para confirmar su correcta identidad.

### 3.9 EVALUACIÓN DE LOS FAB RECOMBINANTES COMO SONDAS PARA LA DETECCIÓN DE GLUTEN

Los cuatro Fab producidos de forma soluble se emplearon como anticuerpos primarios en una metodología de ELISA indirecto (Figura 11B) frente a los mismos antígenos descritos previamente para los fago-Fab (apartado 7). Entre todos ellos, el Fab-C presentó las mejores características de reconocimiento de las proteínas del gluten, por lo que se decidió profundizar en la caracterización de este fragmento de anticuerpo.

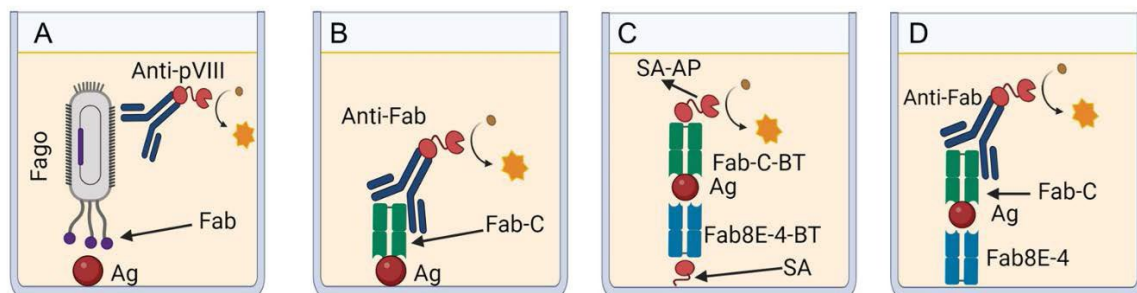
La especificidad del ELISA indirecto basado en el Fab-C se evaluó mediante el análisis de 60 especies heterólogas. La sensibilidad del ensayo se calculó a través de la curva dosis-respuesta (regresión lineal) obtenida al analizar diferentes concentraciones de gliadina-PWG. Los límites de detección (LOD) y de cuantificación (LOQ) se calcularon a partir de los valores obtenidos de interpolar en la curva tres y diez veces la desviación típica de 10 mediciones del blanco, respectivamente. Además, se realizó un ensayo de recuperación analizando harina de arroz certificada sin gluten previamente adulterada con gliadina-PWG sólida. Como control de este ensayo de recuperación, se utilizó un extracto de harina de arroz contaminado con concentraciones de antígeno equivalentes de una solución de gliadina-PWG.

### 3.10 DESARROLLO DE UN ELISA SÁNDWICH BASADO EN LOS FAB RECOMBINANTES

Para desarrollar un inmunoensayo en formato sándwich basado en los Fab obtenidos, se comprobó en primer lugar la compatibilidad entre los diferentes fragmentos de anticuerpo. Para ello, se sometieron los cuatro Fab a un proceso químico de biotilación. De esta forma, empleando placas tapizadas con estreptavidina se inmovilizaron en las placas los diferentes Fab empleados como anticuerpos de captura. De manera análoga, los Fab biotilados se incubaron con la enzima fosfatasa alcalina conjugada a estreptoavidina, para realizar un ELISA sándwich directo sin necesidad de emplear anticuerpo secundario (Figura 11C).

Con esta metodología, se discriminó el par de anticuerpos recombinantes (Fab8E-4 como anticuerpo de captura y Fab-C como anticuerpo de detección) que presentó las mejores propiedades para la detección de gluten en el inmunoensayo tipo sándwich.

Dado que se observó una pérdida de sensibilidad con respecto a los niveles de detección esperados, se diseñó una técnica de ELISA sándwich indirecto basada en la inmovilización del Fab de captura mediante adsorción directa en las inmunoplaacas, y visualización de la unión del Fab de detección con un anticuerpo anti-Fab conjugado a la enzima HRP (Figura 11D). Este ELISA sándwich se evaluó frente a las mismas muestras descritas para el resto de inmunoensayos.



**Figura 11.** Representación esquemática de las técnicas de ELISA desarrolladas en este trabajo para la detección de gluten en alimentos. (A) ELISA indirecto basado en el uso de fago-Fab. (B) ELISA indirecto basado en el uso de Fab soluble. (C) ELISA sándwich directo basado en el uso de Fab biotilados. (D) ELISA sándwich indirecto basado en el uso de Fab solubles. Abreviaturas empleadas: Ag (antígeno, gluten); Fab (fragmento de unión a antígeno); Anti-pVIII (anticuerpo secundario dirigido frente a la proteína VIII del bacteriófago VSC-M13); Fab-BT (Fab biotilado); SA (estreptavidina); AP (fosfatasa alcalina); Anti-Fab (anticuerpo secundario dirigido frente a Fab de origen humano).

### 3.11 ANÁLISIS DE PRODUCTOS COMERCIALES MEDIANTE LOS ELISA DESARROLLADOS CON LOS FAB SOLUBLES

Para verificar la aplicabilidad de los inmunoensayos ELISA indirecto basado en el Fab-C y el ELISA sándwich indirecto desarrollados en este trabajo, se analizaron un total de 50 productos comerciales obtenidos en diferentes comercios locales. Atendiendo a su etiquetado, se clasificaron de la siguiente manera según su contenido en cereales con gluten: 9 productos contenían cereales con gluten, 6 productos no tenían ninguna declaración respecto de los cereales con gluten, 21 declaraban que podían contener gluten, y 14 estaban certificados como libres de gluten. Para validar los resultados obtenidos, todos los productos se analizaron paralelamente con el ELISA sándwich directo basado en el anticuerpo monoclonal R5, que ha sido aprobado como el método de Tipo I por el Codex *Alimentarius* para la detección de gluten en alimentos y presenta un límite de cuantificación (LOQ) de 3 mg/kg de gluten.

### 3.12 CARACTERIZACIÓN DE LOS REPERTORIOS OBTENIDOS Y ESTUDIO DE LOS PROCESOS DE SELECCIÓN POR AFINIDAD MEDIANTE SECUENCIACIÓN MASIVA

A partir de los pellets bacterianos generados en la purificación PEG/NaCl de los fago-Fab tras cada ronda de *panning*, se extrajeron los fagémidos codificantes de cada subpoblación de fago-anticuerpos para analizarlos mediante secuenciación masiva. La secuenciación se realizó mediante la tecnología de nanoporos en un dispositivo MinION de Oxford Nanopore®. Dado que para este sistema es necesario utilizar fragmentos lineales de ADN de cadena doble, los fagémidos obtenidos se trataron con la enzima de restricción *NheI*. Cada muestra proveniente de una ronda de *panning* diferente se identificó con un código de barras único (*barcode*) para poder diferenciarlo tras el proceso de adquisición de secuencias. Los datos de secuenciación se recolectaron durante 30 h, registrándose en archivos crudos con formato FAST5. Con estos archivos, se realizó el análisis de datos en dos etapas diferenciadas: análisis primario y análisis secundario.

Durante el análisis primario, se empleó el programa Megalodon v1.9.5 para realizar la llamada de bases (*basecalling*). Este procedimiento transforma los archivos FAST5 en secuencias de bases de alta calidad tomando como referencia un genoma humano. Este primer análisis, devolvió un archivo FASTQ demultiplexado, es decir, la información se encuentra clasificada atendiendo a los distintos códigos de barras o *barcodes*. A continuación, se eliminaron las secuencias adaptadoras y se realizó un control de calidad con el programa FASTQC, eliminando del análisis mediante el programa LUMC/FASTQ aquellas secuencias con baja calidad o que presentaron un tamaño incorrecto.

El análisis secundario se realizó utilizando la herramienta IMGT/HighV-QUEST basada en la identificación de genes V, D y J mediante el alineamiento de las secuencias obtenidas frente a una base de datos de referencia. El análisis estadístico y representaciones gráficas realizados para estudio de estos datos se generaron con el paquete de R llamado IMGT/StatClonotype.

### 3.13 MODELIZACIÓN *IN SILICO* DE LA ESTRUCTURA DE LOS FAB8E-4 Y FAB-C Y ESTUDIO DE SU INTERACCIÓN CON LA GLIADINA

La modelización *in silico* del Fab8E-4 obtenido del repertorio semiinmune se realizó utilizando el servidor sAb Pred. Una vez generada la estructura del Fab (concretamente su fracción variable o fragmento Fv), el mismo servidor se empleó para identificar los parátomos y epítomos encontrados entre el Fab y la gliadina. Para este estudio se empleó la estructura de la gliadina generada por Alpha-fold2.

En el caso del Fab-C obtenido a partir del repertorio inmune, se llevó a cabo el siguiente proceso para la elucidación *in silico* de su estructura y de su interacción antígeno-anticuerpo: 1) la secuencia nucleotídica se analizó con el servidor IMGT/V-QUEST para detectar y definir los dominios del anticuerpo; 2) la predicción de la estructura del Fab se obtuvo a través de las herramientas bioinformáticas Abodybuilder2 e IgFold; 3) se estudiaron las concordancias y discrepancias entre ambos modelos utilizando la herramienta Matchmaker de ChimeraX; 4) se elucidaron las interacciones entre el Fab-C y la gliadina calculando las predicciones de epítomos y parátomos con las herramientas Epipred y Antibody i-Patch; y 5) basándose en la información sobre epítomos y parátomos obtenida, se realizó una simulación global de la interacción con el antígeno (*docking*) utilizando el servidor HADDOCK 2.4.



## 4. RESULTADOS



## CAPÍTULO 4.1

### Construcción de un repertorio de Fab mediante la combinación de cadenas de anticuerpos de origen semisintético e inmune, apropiada para el desarrollo de nuevas herramientas de inmunodetección de gluten en alimentos

---

### Construction of a Fab library Merging Chains from Semisynthetic and Immune Origin, Suitable for Developing New Tools for Gluten Immunodetection in Food

The observed increase in the prevalence of gluten-related disorders has prompted the development of novel immunological systems for gluten detection in foodstuff. The innovation on these methods relies on the generation of new antibodies, which might alternatively be obtained by molecular evolution methods such as phage display. This work presents a novel approach for the generation of a Fab library by merging semi-synthetic heavy chains built-up from a pre-existent recombinant antibody fragment (dAb8E) with an immune light chain set derived from celiac donors. From the initial phage population ( $10^7$  candidates) and after three rounds of selection and amplification, four different clones were isolated for further characterization. The phage Fab8E-4 presented the best features to be applied in an indirect ELISA for the detection of gluten in foods, resulting in improved specificity and sensitivity.

**García-Calvo, Eduardo; García-García, Aina; Rodríguez, Santiago; Farras, Sergio; Martín, Rosario; & García, Teresa.** (2022). Construction of a Fab library merging chains from semisynthetic and immune origin, suitable for developing new tools for gluten immunodetection in food. *Foods*, 12(1), 149.

<https://doi.org/10.3390/foods12010149>

Journal category: Food Science and Technology

Journal Impact Factor (2022): 5.2

Journal Rank: 34/142 (Q1)





Article

# Construction of a Fab Library Merging Chains from Semisynthetic and Immune Origin, Suitable for Developing New Tools for Gluten Immunodetection in Food

Eduardo Garcia-Calvo <sup>1</sup>, Aina García-García <sup>1,\*</sup>, Santiago Rodríguez <sup>1</sup>, Sergio Farrais <sup>2</sup>, Rosario Martín <sup>1</sup> and Teresa García <sup>1</sup>

<sup>1</sup> Departamento de Nutrición y Ciencia de los Alimentos, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040 Madrid, Spain

<sup>2</sup> Servicio de Medicina Digestiva, Hospital Universitario Fundación Jiménez Díaz, 28040 Madrid, Spain

\* Correspondence: ainagarcia@ucm.es; Tel.: +34-91-394-3749

**Abstract:** The observed increase in the prevalence of gluten-related disorders has prompted the development of novel immunological systems for gluten detection in foodstuff. The innovation on these methods relies on the generation of new antibodies, which might alternatively be obtained by molecular evolution methods such as phage display. This work presents a novel approach for the generation of a Fab library by merging semi-synthetic heavy chains built-up from a pre-existent recombinant antibody fragment (dAb8E) with an immune light chain set derived from celiac donors. From the initial phage population ( $10^7$  candidates) and after three rounds of selection and amplification, four different clones were isolated for further characterization. The phage Fab8E-4 presented the best features to be applied in an indirect ELISA for the detection of gluten in foods, resulting in improved specificity and sensitivity.

**Keywords:** gluten; prolamins; celiac disease; phage display; ELISA; immune library; semi-synthetic library



**Citation:** Garcia-Calvo, E.; García-García, A.; Rodríguez, S.; Farrais, S.; Martín, R.; García, T. Construction of a Fab Library Merging Chains from Semisynthetic and Immune Origin, Suitable for Developing New Tools for Gluten Immunodetection in Food. *Foods* **2023**, *12*, 149. <https://doi.org/10.3390/foods12010149>

Academic Editor: Qinchun Rao

Received: 4 November 2022

Revised: 20 December 2022

Accepted: 23 December 2022

Published: 28 December 2022



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Gluten is defined by Codex Alimentarius as ethanol-soluble but water and 0.5 M sodium chloride insoluble protein fractions, that are present in some cereal endosperms like wheat, rye, barley, spelt, and kamut [1]. Gluten proteins are divided into two fractions by their solubility in a 60% ethanol/water solution: soluble gliadins and insoluble glutelins. These chemical features result in different technological properties of dough, as gliadins are related to viscosity and extensibility, while glutelins contribute to strength and stability [2]. The extraordinary physical-chemical stability of gluten protein fractions (like gliadin) results in slow partial digestion and high permanence in the gut, that is related to its pathogenicity: undigested gliadin peptides increase gut permeability and have specific binding to CXCR3 chemokine receptor, with a subsequent release of zonulin (a modulator of tight junctions). This occurs in all individuals; however, it can lead to inflammatory processes in sensitized persons, like those presenting the human leukocyte antigen (HLA)-DQ2 and DQ8 haplotype [3].

There are several gluten-related diseases classified by etiopathology. Firstly, those classified as allergic diseases like baker's asthma (respiratory allergy), food allergy, contact urticaria and, WDEIA (wheat-dependent exercise-induced anaphylaxis) [4]. The second group clusters autoimmune diseases like celiac disease, the best known, caused by a strong immune response against gluten-processing self-transaminases, ending in a severe destruction of gut tissue. There are other diseases included in this group, as gluten ataxia and dermatitis herpetiformis [5]. The third group is called non-celiac gluten sensitivity for those situations that do not fit in the previous groups. It seems that they could be the most

prevalent diseases and their molecular etiology is not completely elucidated but, some studies show a core implication of the innate immune system [6].

Due to the health impact of gluten-related diseases, several countries have developed gluten labeling legislation, most of them following Codex Alimentarius limit of 20 parts per million (ppm) to be considered a gluten-free product [7].

Several analytical methods have been developed to detect and quantify gluten in foodstuffs, including PCR [8], chromatography and mass spectrometry [9]. However, the immunological techniques, including ELISA methods [10], are the most widely used. The development of novel and more reliable gluten immune-detection methodologies are coupled with antibody discovery. There are three main technologies: polyclonal antibodies from animal serum, monoclonal antibodies from hybridomas, and recombinant antibodies in the context of molecular evolution strategies [11].

The first methods were developed on the basis of polyclonal antibodies derived from rabbit IgGs [12] and chicken IgYs [13] and, nowadays they are still in use. The revolution of monoclonal antibodies also had a great impact on this field. Three hybridoma derived antibodies represent the main reactants for the most widely used gluten-detection methods: 401.21 [14], R5 [15] and G-12 [16]. A sandwich ELISA based on R5-mAb is the golden standard for gluten detection in food validated by AACI (American Association of Cereal Chemists International), AOAC International, and Codex Alimentarius [17].

Although polyclonal and monoclonal antibodies are widely applied for gluten detection in food, the development of new recombinant antibodies could bring new insights to the field. Recombinant antibodies can be developed in shorter, cheaper, and simpler processes than monoclonal “classical” antibodies. Unlike mono and polyclonal antibodies, recombinant antibody technology allows complete animal-free development and production of specific binders for immunoassays [18].

Antibodies are plasmocyte secreted proteins composed of four polypeptide chains, classified by their molecular weight in two heavy and two light chains. Light chains are composed by a constant and a variable domain and, in humans, they are coded by genes that reside in two different chromosomes (resulting in two kinds of light chains, kappa and lambda). Heavy chains are composed of a variable domain and three or four constant domains (the case of human IgE or avian antibodies). The variable domains have two topological domains; three juxtaposed loops called complementary determining regions or CDRs, that drive the main antigen-antibody interaction and, in between, rigid and conserved  $\beta$ -barrels, the framework regions (FRs) [19]. Camelids show a unique antibody structure, with two neighbouring single heavy chains [20]. Thanks to the development of recombinant expression of antibodies, it is possible to produce non-natural engineered antibody fragments and natural constructions, like Fabs (conformed by the light chain attached to the variable and first constant domain of heavy chain), dAbs (the variable domain of heavy chain) and many combinations of the elements mentioned above.

The final years of the 20th century meant a new revolution on antibody development thanks to the works of Winter and Smith [21], that introduced the concept of directed molecular evolution. This method goes beyond the recombinant expression of antibodies. It encloses three main steps: genesis of diversity, selection and amplification of high-affinity antibodies against a target antigen. One to ten binders can be chosen from initial  $10^7$  antibodies, after 3 to 6 rounds of selection. After selection, these binders become dominant within the population, and their genotype is coupled with their phenotype (thanks to phage, ribosome or cell surface display of molecules). There are several examples of the application of directed evolution for the elucidation of molecular mechanisms of celiac and related diseases, using phage display. They include immune libraries, but also naïve [22] and peptide libraries [23]. Although the clinical research has been the main application, there have been several projects aimed to develop antibodies for gluten detection in food using camelid antibodies. Immunized llamas were used for the construction of a phage-displayed VHH library, with the objective of obtaining a capture antibody for gluten detection by

sandwich-ELISA [24]. Moreover, a naïve llama library was used for the development of a novel gluten detection methodology (cDNA display mediated immuno-PCR) [25].

The present work is the continuation of a novel strategy for isolation of high affinity antibodies against gluten [26], where a semi-synthetic library based on human VH3 germline, was confronted to a peptide proposed to be a consensus HLA-presented peptide of digested gluten. From this process, a suitable candidate was selected (dAb-8E) and used as the specific reagent for the development of a phage-ELISA method for gluten detection in foodstuff.

Although the antibody specificity is mainly the result of the interaction between the V-region with the antigen, constant domains also affect to antigen binding, despite they are not participating on the physical antibody-antigen interaction [27]. The constant domains influence in segmental flexibility and inter-antibody associations, fixing the conformational signaling through the V<sub>H</sub>-C<sub>H1</sub> domains, so that, structural and kinetic constraints are imposed to the contact surface [28]. So, it was hypothesized that the transformation of the dAb-8E into a Fab could enhance its properties. Therefore, the main objective of this work was the generation of a new antibody library merging chains of two different origins: a light-chain library derived from plasmocytes isolated from celiac patients, and heavy chains derived from the dAb8E coding sequence transformed into human heavy-chains (IgG1, IgG2, IgG3 and IgG4).

## 2. Materials and Methods

### 2.1. Bacterial Strains and Growth Media

*Escherichia coli* XL1-Blue strain (recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F proAB lacIqZΔM15 Tn10 (Tetr)]) (Agilent<sup>®</sup>, Santa Clara, CA, USA, #ref 200150) was used for cloning, building the libraries and producing phage-displayed Fabs. Electrocompetent cells were in-house produced by growing the mentioned strain in super-broth medium (SB: 30 g/L tryptone, 20 g/L yeast extract, 10 g/L MOPS pH 7) and concentrated by sequential cycles of centrifugation (2218× *g*, 4 °C) and resuspension in 10% glycerol (mL/mL) (PanReac AppliChem<sup>®</sup>, Monza, Lombardy, Italy, CAS: 56-81-5) diluted in water. After electroporation, cells were grown in SOC medium (Invitrogen<sup>™</sup>-Thermo Fisher<sup>®</sup>, Waltham, MA, USA, #ref 15544-034). For DNA extraction before sequencing, bacteria were grown in Luria Broth (LB: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl). Agar plates were prepared with a 15 g/L agar concentration.

### 2.2. DNA Isolation, Quantification, and Restriction

All DNA products used for library construction were isolated by gel electrophoresis using UltraPure<sup>™</sup> Low Melting Point agarose (Thermo-fisher<sup>®</sup>, #ref 16520050), and purified with NucleoSpin<sup>®</sup> gel DNA clean-up columns (Machery-Nagel<sup>®</sup>, Allentown, PA, USA, #ref 740609). DNA was quantified in a Qubit<sup>®</sup> Fluorometer (Invitrogen Thermo-fisher<sup>®</sup>), and its quality was measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Montchanin, DE, USA). Restriction enzymes *SacI* #ref R0156, *XbaI* #ref R0145, *XhoI* #ref R0146 and *SpeI* #ref R0133 and T4 DNA ligase #ref M0202, were purchased from New England Biolabs<sup>®</sup> (Ipswich, MA, USA).

### 2.3. Semisynthetic Heavy Chain Sub-Library Construction

The original sequence of dAb-8E [26] had an intra-sequence *XhoI* restriction site that required removal for proper cloning by a single-base mutation with Agilent<sup>®</sup> QuickChange II kit (Santa Clara, CA, USA, ref #200523). Once the single-point mutation was verified by sequencing, the heavy chains for library construction were built by a two-step PCR process. Forward primer VH135 and reverse primers ovlp-conga and ovlp-cong4 (Table S1) were used in the first round of PCR to obtain the VH domain of the mutated dAb-8E. This first step incorporated a *XhoI* site at 5' end, and, the initial nucleotides of CH1 domains at the 8E-FR4 end (IgG1, IgG2 and IgG3 sequences hybridized with ovlp-conga primer, and IgG4 sequence with ovlp-cong4 primer). During the second round of PCR, the first-

round amplicons were fused to the remaining CH1 sequences by overlap extension PCR with the forward primer VH135ext and the reverse primers CG1Zext, CG2aext, CG3aext, and CG4aext (Table S1) [29], generating the coding sequences of the heavy chains to be incorporated into the final Fab-repertoire, and a *SpeI* site at 3' end for clonation.

#### 2.4. Immune Light Chain Sub-Library Construction

Two patients newly diagnosed with celiac disease were selected as donors for this work in the digestive medicine service of Fundación Jiménez Díaz Hospital (Madrid, Spain). The humoral response against gliadin was quantified using ELISA tests:  $\alpha$ -Gliatest S Chromo IgA (Eurospital, Trieste, Italy, ref #910796) and  $\alpha$ -Gliatest S Chromo IgG (Eurospital, ref #910896). After acceptance of the informed consent, 350 mL of peripheral blood was obtained from each donor and mononuclear cells were isolated using a Ficoll gradient (Merck, Darmstadt, Germany, CAS #26873-85-8). Total RNA was extracted by a standard method with TRIzol LS reagent (Invitrogen™-Thermo Fisher®, ref #10296010). Once the RNA quality and quantity was checked by Agilent® 2022 Bioanalyzer (ref #G2939BA), it was used as template in reverse transcription reactions (SuperScript IV®, Invitrogen™-Thermo Fisher®, ref #18091050). Double stranded DNA encoding the light chain repertoire was amplified using the collection of primers described by Barbas et al. [29] (Table S2). Finally, the purified amplicons were re-amplified in order to add the *SacI/XbaI* restriction sites needed for cloning in the pComb3X vector.

#### 2.5. Fab-Library Construction

Library construction was developed as described by Burton et al. [30] with minor modifications. Light chain amplicons (1.5  $\mu$ g) and pComb3X vector were *SacI/XbaI* double-digested and gel purified. Digested inserts and linearized vector were ligated (3:1 molar ratio) at 16 °C for 16 h with T4 DNA ligase. The resulting ligation was ethanol precipitated, re-suspended in 20  $\mu$ L of water and transformed into high-efficiency electrocompetent *E. coli* XL1-Blue. Subsequently, 3 mL of SOC medium was added, incubated, and shaken at 250 rpm for 1 h at 37 °C. At this point, a small aliquot of the culture was removed and titrated in LB agar plates containing 100  $\mu$ g/mL of carbenicillin. The rest of the culture was transferred to 12 mL of LB broth with 20  $\mu$ g/mL of carbenicillin and 10  $\mu$ g/mL of tetracycline and incubated for an additional hour. Afterwards, the culture was expanded to 100 mL of LB (carbenicillin at 50  $\mu$ g/mL and tetracycline at 10  $\mu$ g/mL) and grown overnight. To ensure the proper performance of the cloning process, the plasmid DNA from 30 individual clones was sequenced (Table S3).

Phagemid DNA containing the immune light chain repertoire was purified and *XhoI/SpeI* digested for heavy chain insertion. A total of 1.5  $\mu$ g of the previously obtained light chain sub-library were *XhoI/SpeI* digested and the linearized phagemids gel purified. Subsequently, the semi-synthetic heavy chain sub-library was cloned following the steps described above for the light chain. The resulting ligation was ethanol precipitated and transformed into *E. coli* XL1-Blue. After electroporation, 5 mL of recovering SOC medium was added, and shaken at 250 rpm at 37 °C for 1 h. A small aliquot was removed, ten-fold serially diluted and plated into SB plates (containing 100  $\mu$ g/mL of carbenicillin) in order to calculate the repertoire size of the library. Ten milliliters of SB (containing 20  $\mu$ g/mL of carbenicillin and 10  $\mu$ g/mL of tetracycline) was added to the remaining culture and incubated under the same conditions for 1 h. At this point, carbenicillin concentration was increased to 50  $\mu$ g/mL and the culture was incubated for one additional hour. Then, 2 mL of VSCM13 helper phage ( $10^{12}$ – $10^{13}$  pfu/mL) was added, and the culture was transferred to 200 mL of SB (containing 50  $\mu$ g/mL of carbenicillin and 10  $\mu$ g/mL of tetracycline) and shaken at 300 rpm for 2 h. Kanamycin was added to a concentration of 70  $\mu$ g/mL and the culture incubated overnight with shaking (300 rpm) at 37 °C.

The following day, the culture was centrifuged at  $2218\times g$  for 20 min at 4 °C. The phagemid DNA obtained from the pelleted cells constituted the DNA library, and the phage-Fab library was isolated from the culture supernatant by PEG-NaCl precipitation.

The supernatant (200 mL) was mixed with 8 g of PEG-8000 (Sigma<sup>®</sup>, Burlington, MA, USA, #ref 89510) and 6 g of NaCl (Sigma<sup>®</sup>, #ref S3014), incubated on ice for 1 h and centrifuged at  $10,016 \times g$  for 20 min at 4 °C. The supernatant was discarded and the phage-antibodies were resuspended in 2 mL of 1% BSA (NZYtech<sup>®</sup>, Lisbon, Portugal, #ref MB04602) in PBS (137 mM NaCl, 2.7 KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4).

#### 2.6. Selection of Gliadin-Binding Phage-Fab Variants by Panning

Four wells from a 96-well immunoplate were coated for 1 h at 37 °C with 100 µL of gliadin-PWG at 10 µg/mL in 0.1 M carbonate/bicarbonate buffer pH 9.6. The coating solution was removed, and after three washing steps with PBS, 400 µL of blocking solution (3% BSA in PBS) was added to each well. After incubating the plate at 37 °C for 1 h and performing ten washes with PBS, 50 µL of the phage-antibody library was added per well and the plate was incubated at 37 °C for 2 h. Non-binding phages were washed-away with five washing steps with PBS-T (PBS containing 0.05% of Tween20). Gliadin binding phage-antibodies were eluted with 50 µL of 0.1 M glycine-HCl-1% BSA (pH 2.2) by scratching the bottom of the well with a cut pipette tip for 10 min. Finally, the recovered phages were neutralized with Tris-Base 2 M (Sigma<sup>®</sup>, #ref T0440) to pH 7.

The eluted phage-antibodies were used to infect a 2 mL culture of *E. coli* XL1-Blue grown in SB (10 µg/mL of tetracycline) at  $OD_{600} = 0.8-0.9$ . After 15 min of infection, a small aliquot was ten-fold diluted and plated in order to calculate the output titering. Afterwards, 6 mL of SB (containing 20 µg/mL of carbenicillin and 10 µg/mL of tetracycline) was added to the infected culture and further incubated at 37 °C for 1 h (250 rpm). Carbenicillin concentration was increased to 50 µg/mL and the culture incubated for 1 h at 37 °C with shaking at 270 rpm. Then 1 mL VSCM13 helper phage was added, and the culture expanded to 100 mL of SB (containing 50 µg/mL of carbenicillin and 10 µg/mL of tetracycline) and shaken for 2 h at 37 °C (300 rpm). Kanamycin was added to a concentration of 70 µg/mL and the culture was incubated overnight at 37 °C (300 rpm). This amplified phage-antibody population obtained from the first round of panning was titered and used as input for the second round of selection against gliadin. A total of three rounds of selection were performed following the protocol described above, but with an increase in the number of washes for the removal of non-specific phages: 5, 10 and 15 for the first, second, and third round of panning, respectively. In addition, four wells were used in the first round, then the number was reduced to two in following rounds.

#### 2.7. Preparation of Antigens

Gliadin-PWG was used as reference material for gluten detection in foods (Arbeitsgemeinschaft Getreideforschung (AGF) e.V. DIGeFa GmbH, Detmold, Germany). Following the supplier's recommendation, it was dissolved in a 60% ethanol/water solution at 100 µg/mL and stored at room temperature protected from light. Gluten-like proteins were also extracted from different cereal matrices such as wheat, rye, barley, corn, and rice, as well as a binary mixture of wheat, rye, and barley kernels in a rice-based gluten-free matrix. Ten food products were tested (Table S4) following the protocols for extraction, analysis and result confirmation as previously described in García-García et al. [24]. Gluten-like proteins were extracted from 250 mg of the finely ground sample in 10 mL of a 60% ethanol (or 200 mg in 1 mL of 50% isopropanol solution in the case of Western blot analysis), followed by shaking for 30 min at room temperature and centrifugation at  $867 \times g$  for 15 min at room temperature. The ethanol extracts were transferred to a clean vial and stored at room temperature protected from light until analysis.

#### 2.8. Small and Medium Scale Individual Clone Induction for Analysis

Individual colonies isolated from the titering output plates were grown in 4 mL of SB medium (50 µg/mL of carbenicillin) and incubated at 37 °C until  $OD_{600}$  reached 0.9–1. At this moment, 40 µL of VSCM13 helper phage preparation ( $10^{12}$ – $10^{13}$  pfu/mL) was added and incubated a 37 °C for 2 h with shaking. After helper infection, 70 µg/mL of kanamycin

was added and the phage-antibodies produced in the supernatant during an overnight incubation at 37 °C. Selection of gliadin binding phage-antibody candidates was performed by indirect ELISA of supernatants. The phage-antibodies from the selected clones were produced using the same protocol but on a larger scale (100 mL of SB-carbenicillin) and the phage-Fabs were PEG-NaCl purified as described in the panning protocol. The phage preparation was tittered using the classical top-agar methodology, presenting  $4\text{--}6 \times 10^{12}$  phage particles per milliliter.

#### 2.9. Indirect Phage-Fab ELISA

Gluten proteins extracted from flour samples or gliadin-PWG were diluted in coating buffer (0.1 M carbonate/bicarbonate buffer, pH 9.6). One hundred microliters of the prepared extracts were added per well and the plate was sealed and incubated for 1 h at 37 °C. The coating solution was shaken out and 200 µL blocking solution (3% BSA in PBS) was added per well and, the plate was incubated for 1 h at 37 °C. Following 10 washing steps with PBS, 100 µL of diluted phages were added and plates incubated 2 h at 37 °C. The non-binding phages were washed away 10 times with PBS. Then, 100 µL of the secondary antibody (anti-phage protein VIII HRP conjugated from Sinobiological<sup>®</sup>, Beijing, China, #ref 11973) diluted 1:5000 in blocking buffer was added. The plate was incubated for 1 h at 37 °C and PBS washed 10 times. One hundred microliters of TMB (Sigma<sup>®</sup>, #ref T0440) was added, and the reaction was stopped after 15–20 min with 50 µL of a diluted sulphuric acid solution. Absorbance readings were performed at 450 nm wavelength (FluoStar Optima<sup>™</sup> from BMG labtech<sup>®</sup>, Ortenberg, Germany).

#### 2.10. Western Blot Analysis of Gliadin and Gluten-Containing Cereals

The different alcohol-soluble fractions of the commercial gliadin and the gluten-containing cereals were separated in two identical acrylamide gels (AnyKD precast-gel from Bio-Rad<sup>®</sup>, Hercules, CA, USA, #ref 4569033) under denaturing conditions at 90 V using NZYcolour Protein Marker II (NZYtech<sup>®</sup>, #ref M090). After electrophoresis, one of the gels was stained with Coomassie blue and the protein bands of the twin gel were transferred to a PVDF membrane at 100 V for 1.5 h. The PVDF membrane was blocked overnight with 3% BSA in TBS (50 mM Tris HCl, 150 mM NaCl, pH 7.5) before incubation with a 1:45 dilution (TBS-BSA 1%) of the PEG-NaCl precipitated phage-Fab8E-4 for 2 h at 37 °C in a rocking platform. Membrane was rinsed 5 times with TBS and incubated for 1 h at 37 °C with the anti-M13proteinVIII-HRP secondary antibody diluted 1:5000 in blocking solution. The image was revealed after the addition of a chemo-luminescent reactive (Clarity<sup>™</sup> ECL Western blotting substrate from Bio-Rad<sup>®</sup>, #ref 1705061) in a ChemiDoc Imaging System (Bio-Rad<sup>®</sup>).

#### 2.11. Sequencing and Molecular Modeling of Fab8E-4

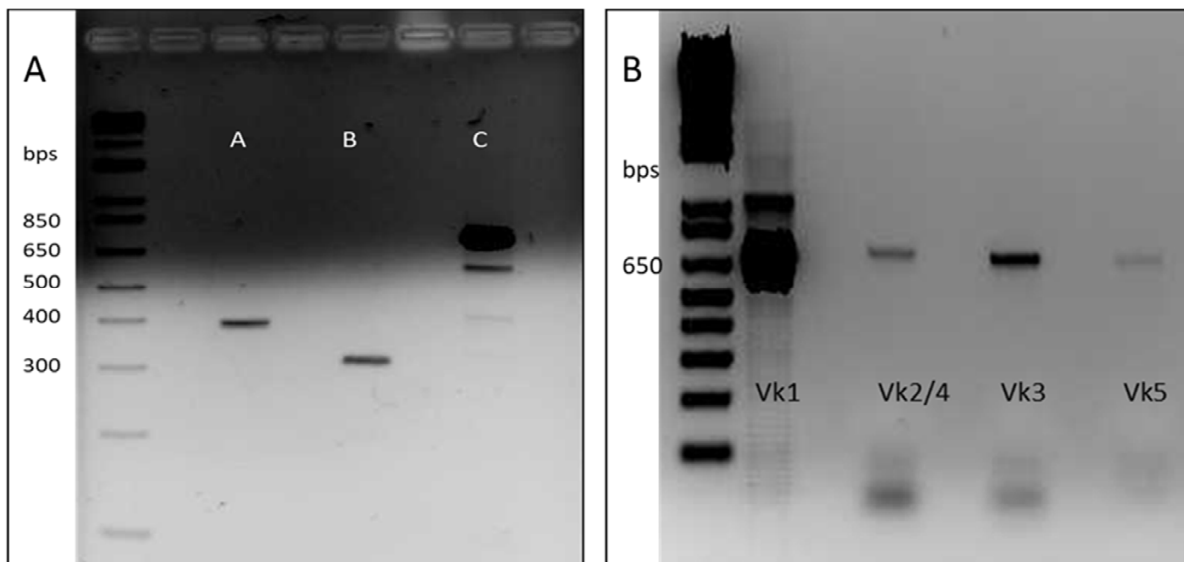
DNA from the isolated gliadin-binding phage-Fab clones was extracted with the plasmid mini-prep (GenElute<sup>™</sup> plasmid miniprep kit from Sigma<sup>®</sup> #PLN70), and the phagemids sequenced by Sanger method (Eurofins genomics<sup>®</sup>, Luxemburg, Luxemburg) using the primers ompAseq and g-back (Table S3) for light and heavy chain reading, respectively [29].

The antibody coding sequences were searched using the IMGT<sup>®</sup>/V-QUEST alignment tool [31] and the IgBLAST tool from NCBI [32]. Furthermore, amino acid sequence from Fab8E-4 was used to obtain a predictive model of the Fab structures using the sAb Pred server [33]. Once the Fab structure was generated, paratope and epitope prediction were calculated using the above mentioned server for which antigen structures were modeled with Alpha-fold [34].

### 3. Results and Discussion

#### 3.1. Semi-Synthetic Heavy Chain Sub-Library Construction

The main objective of this work was the development of a Fab library constructed by a two-step cloning process consisting of cloning firstly the immune-light chains and subsequently the semi-synthetic heavy chains. For Fab heavy chain construction, the sequence of a single domain antibody (named dAb-8E) previously obtained by our group against gluten was used [25]. Before cloning, the dAb-8E sequence was mutated in order to eliminate an intra-sequence restriction site that could originate the formation of wrong constructs. Since dAb sequence lacks constant heavy chain region, the Fab sequences were built-up from the dAb-8E sequence conforming the variable region, and the CH1 constant region was added by SOE-PCR. A first round of PCR was performed to obtain the dAb and CH1 sequences originating the intermediate constructs of 392 and 300 bp, respectively (Figure 1A-lanes A and B). The gel purified DNA fragments were used as templates for the second round of PCR that allowed the overlap of the dAb-8E with the CH1 coding sequences, obtaining the heavy-chain coding constructs (Figure 1A-lane C) of approximately 680 bp, prepared for next step.



**Figure 1.** Agarose gel containing coding fragments for antibody chains to be cloned for library construction. (A) Different constructs that built the semi-synthetic heavy chains. Ladder: 1 kb plus Invitrogen™ (#ref 10787018); lane A: dAb-8E amplified sequence from the first round of PCR (VH-392 bp); lane B: CH1 amplified sequence from the first round of PCR (CH-300 bp); lane C: construct with the heavy chains obtained from the second round of PCR (SOE-PCR) prepared for Fab library cloning (H-680 bp). (B) Electrophoretic analysis of the human kappa chain genes amplified by PCR (Ladder: 1 kb plus Invitrogen©).

#### 3.2. Immune Light Chain Sub-Library Construction

The light chain coding sequences were amplified from the cDNA template obtained by RT-PCR from lymphocyte-isolated RNA. The B-cells were obtained from two donors diagnosed of celiac disease during this study and presenting high serum response against gluten proteins, since they had not yet adhered to a gluten-free diet. The first patient showed a 168 UA/mL of IgA response against native gliadin (11 times the baseline of 15 UA/mL) and 550 UA/mL of IgG response against native gliadin (11 times the baseline of 50 UA/mL). The second donor presented 96.5 and 170 UA/mL for IgA and IgG responses, respectively (6.5 and 3.4 times the baselines mentioned above). From the 350 mL of peripheral blood collected from each patient, approximately  $1.8 \times 10^8$  B-cells (counted with a Neubauer

chamber after Trypan Blue staining) were isolated from each sample using a Ficoll gradient. Total RNA was isolated (100 µg) showing high integrity (RIN (RNA integrity number) value of 8.75–9.5 over 10) and used to produce the cDNA of interest by RT-PCR.

The cDNA obtained was used as template for a first round of PCR to amplify the complete light chain coding genes, resulting in amplified fragments of approximately 650 bp. Due to the limited quantity of RNA obtained from the biological samples, only the human kappa chains were amplified in this study, showing a higher expression of the Vk1 set in the electrophoretic analysis (Figure 1B). After gel DNA purification, a second round of PCR was performed to include enhanced restriction sites on the obtained kappa chains for proper cloning [28].

### 3.3. Fab Library Cloning and Panning

For Fab library construction, a two-step cloning process was performed. Firstly, double digested inserts coding for light chains were cloned into linearized pComb3X and transformed into electrocompetent *E. coli* XL-1 Blue, obtaining the light chain sub-library, formed by  $1.4 \times 10^7$  total transformants. Plasmid DNA was isolated and re-linearized to clone the double digested heavy chain constructs. The final ligation products were precipitated and transformed (elapsed time of 4 ms), resulting in a Fab library with a repertoire size (RZ) of  $2.7 \times 10^7$  (autoligation factor of 2%), fulfilling the standards considered acceptable for this kind of libraries of  $RZ > 10^7$  and autoligation factor of 5% [27].

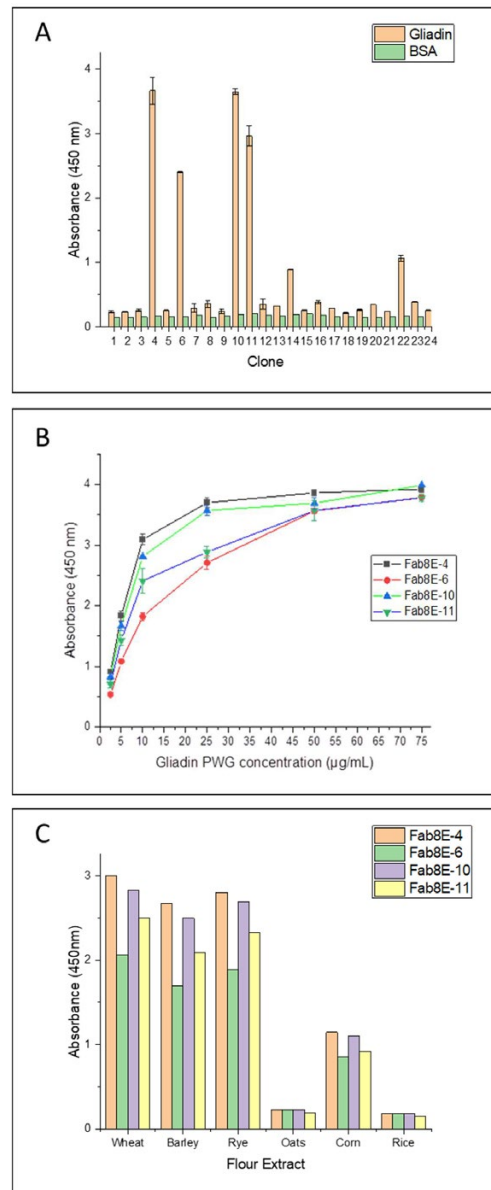
Following helper phage infection, the phage-antibody repertoire was isolated by PEG purification and used for biopanning (successive rounds of selection and amplification) against gliadin-PWG. The repertoire enrichment process was tracked by titrating the output phages after each of the three rounds of selection performed, which corresponds to the infection of the eluted phage-antibodies after antigen exposition. Phage recovery increased 58 times after the third round of panning (Figure S1), indicating the likely selection of phage variants against the target molecule. To further confirm this result, ELISA experiments were developed to evaluate if there was an effective enrichment of the phage population in high-affinity gliadin binders.

### 3.4. Polyclonal Phage-ELISA

The suitability of the panning methodology was checked by ELISA analysis of gliadin-PWG coated plates with the PEG-purified phage repertoires (Figure S2). As expected, gliadin was barely detected by the un-selected library (R0), followed by a 7-fold signal increase after the first interaction with the antigen of interest, and maximized absorbance values were obtained with the purified phage populations after the second and third rounds of selection. Very low signals (<0.2 AU) were recorded for the negative control (BSA used as blocking solution) when using the phages obtained after each round of enrichment, indicating that there was no non-specific selection of phage variants against other components of the antigen binding support. Based on these results, a fourth round of selection was considered unnecessary.

### 3.5. Selection of Positive Binders by Monoclonal Phage ELISA

A total of 24 individual clones were picked from the titrating plates of the output phage population from round 3. Induced phage supernatants were 1:2 diluted and subjected to monoclonal phage-ELISA in immunoplates coated with gliadin-PWG at 1 µg/well. Predictably, binding and non-binding phage variants were observed, resulting in 4 out of the 24 clones (approximately 17%) selected as high-affinity gliadin-binding clones (absorbance values greater than 2 AU and negative control signals under 0.3 AU) (Figure 2A).



**Figure 2.** Monoclonal indirect phage-ELISA of Fabs isolated from the panning process. (A) Indirect phage-ELISA analysis of individual induced clones isolated from the last round of affinity enrichment. (B) Comparison of the dose-response curves by means of the indirect phage-ELISA of the four selected binding clones performed with gliadin-PWG dilutions from 2.5 to 75 µg/mL. Mean values of three independent determinations and standard derivation of each data set are shown. (C) Indirect phage-ELISA of positive clones against ethanol protein extracts of different cereal flours (1:100).

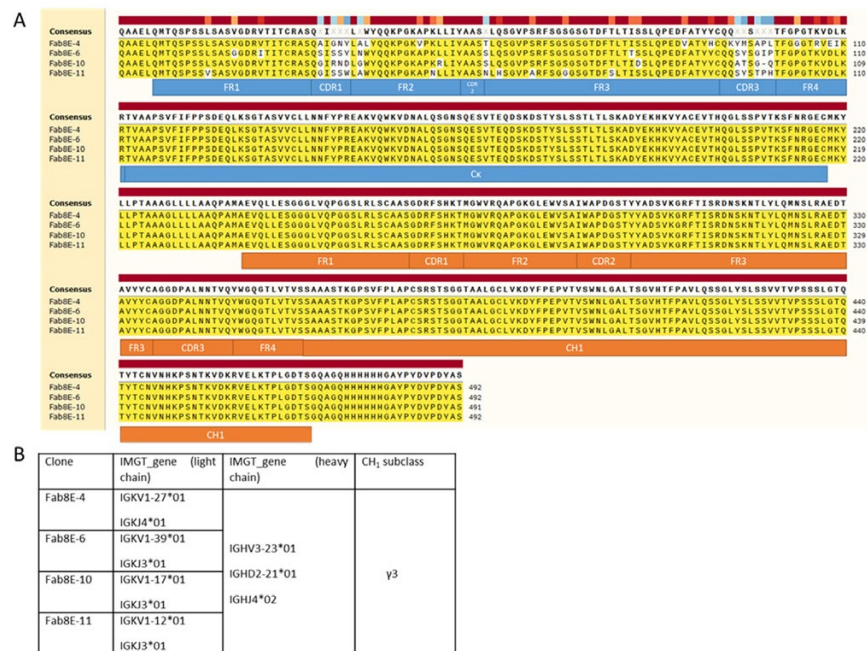
### 3.6. Affinity and Specificity Analysis of the Selected Positive Clones

Monoclonal phage-ELISA experiments were performed to evaluate the affinity and specificity of the four gliadin-binding phage-Fabs. A dose-response curve against increasing concentrations of gliadin-PWG (2.5, 5, 10, 25 and 75 µg/mL) was obtained for each phage variant. Fab8E-4 showed the most promising recognition properties, exhibiting the most sensitive dose-response curve (higher absorbance values at all the gliadin concentrations tested) (Figure 2B).

In addition, a specificity test was conducted with different cereal flours to evaluate the performance of the selected clones in a real food matrix and to discriminate the type of prolamin recognized by the phage-Fabs. As in the case of the gliadin curve analysis, Fab8E-4 showed highest detectability for wheat, barley, and rye prolamins (Figure 2C). Considering these results, among the isolated phage-Fabs, Fab8E-4 presented the best performance for prolamin detection, both in the detection of gliadin dilutions and ethanol extracts from natural gluten sources. Consequently, further characterization experiments were only performed with the phage-Fab8E-4.

3.7. Sequencing of Leading Candidates

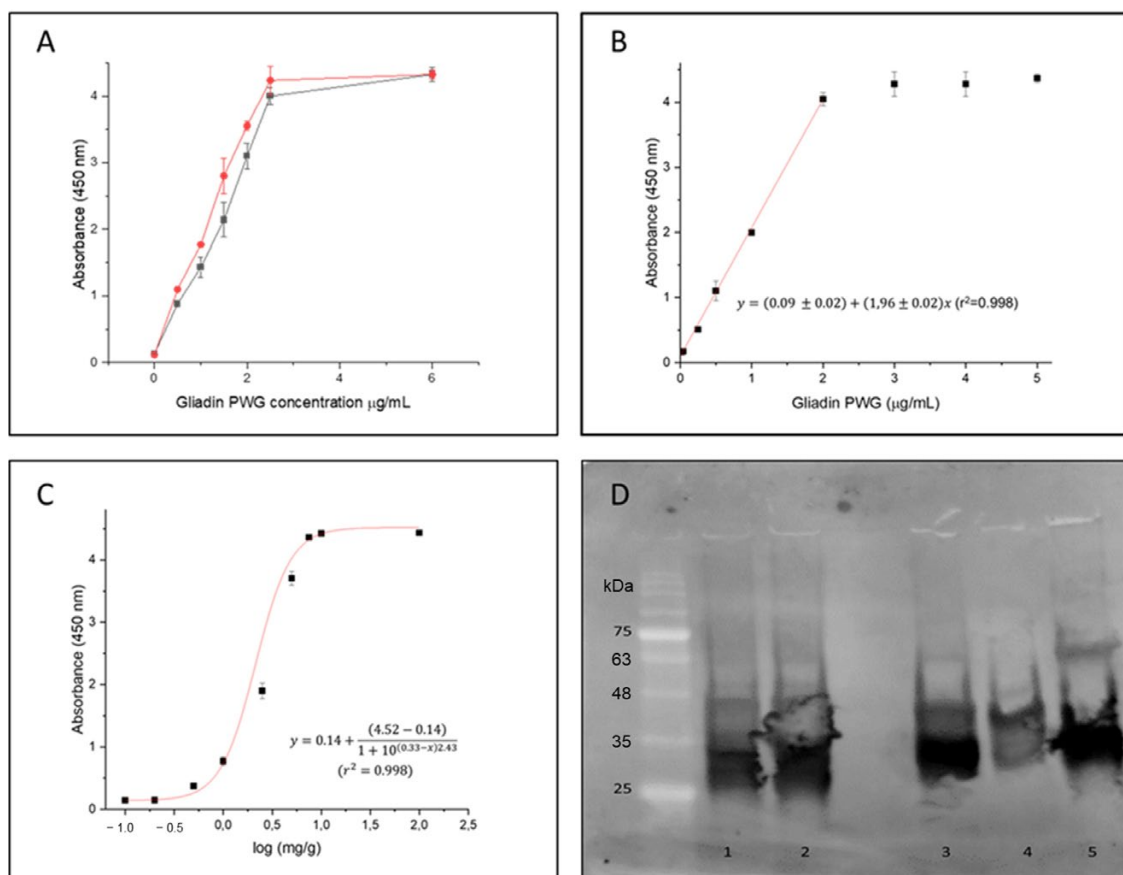
The results of plasmid DNA sequencing of several clones after bacterial transformation showed that the two-step cloning process was successful and that the presence of self-ligation events or other artifacts was not noticeable. Therefore, the constructed library consisted of a Fab repertoire constituted by a collection of light chains expressed in celiac disease patients and the sequence of the dAb-8E transformed into four sets of heavy chains (IgG1, IgG2, IgG3 and IgG4 subclasses). From the DNA sequence analysis of the clones selected as positives after biopanning (Figure 3A), it was inferred that all belonged to different clonotypes derived from the combinatorial arrangement of four different light chains and the same heavy chain, dAb-8E-IgG3 (Figure 3B). The clones selected after three rounds of panning presented light chain diversity (different V and J genes) but only one heavy chain (with the same subclass), this fact could be meaning that the heavy and not the light chains is the principal agent in the molecular evolution process that is selecting the high-affinity clones.



**Figure 3.** Sequencing results, chain families and features identification of the gliadin-binding selected clones. (A) Alignment of the sequences coding for the Fab8E-4, 6, 10 and 11. Main antibody features are shown in blue and orange charts for light and heavy chains, respectively. Alignment was computed using SnapGene® 5.0.8. (B) Identification of the re-arranged genes coding the light and heavy chains of the isolated gliadin-binding clones. The VJ and VDJ genes were elucidated using the IMGT®/V-QUEST tool, and the IgBLAST tool from NCBI was used for elucidating the CH1 subclass of the heavy chains.

### 3.8. Characterization of Phage Fab8E-4

In order to compare the sensitivity of the designed phage-ELISA based on the Fab8E-4 with that of the parental phage-dAb8E, both phage-antibody formats were expressed using the same helper phage. For this purpose, it was necessary to produce the dAb-8E attached to the VSCM13 pIII minor capsid protein, instead of the KM13 bacteriophage used before [25]. The comparative dose response curve obtained against increasing concentrations of gliadin-PWG (0.5–6 µg/mL) with both phage-antibody formats is shown in Figure 4A. Following the Eurachem guides for method validation [34], the limit of detection (LOD) was calculated from the interpolation of three times the standard deviation of ten blank determinations in the linear regression in the range 0.01–2 µg/mL (Figure 4B). The LOD for the method using the Fab8E-4 were set at 0.01 µg/mL.



**Figure 4.** Characterization of phage-Fab8E-4. (A) Comparative dose-response curve obtained against gliadin-PWG dilutions (0–6 µg/mL) by means of dAb-8E (grey) and Fab8E-4 (red) antibody fragments attached to VSCM13 helper phage ( $4 \times 10^{11}$  pfu/well). (B) Linear regression obtained against gliadin dilutions (0.01–2 µg/mL) by means of Fab8E-4 in order to determinate the LOD and LOQ. (C) Representative dose-response curve from the analysis of experimental rice-based binary mixtures with increasing presence of a proportional mixture of wheat, barley and rye in the range of 0.1 to 100 mg/g. Origin 8.0 software was employed to plot and analyze the experimental data. Mean values of three independent determinations and standard derivation of each data set are shown. (D) Western blot analysis of gliadin and flour proteins incubated with phage-Fab8E-4 as primary antibody. Lane 1 (1 mg per well) and Lane 2 (2 mg per well): gliadin (Merck©, CAS #9007–90-3) dissolved in 50% isopropanol-water; Isopropanol extractions of cereal flours: Lane 3: wheat; Lane 4: barley; Lane 5: rye.

Merging the VL immune chains with the semi-synthetic dAb-8E seemed to slightly improve the LOD for the gliadin reference material solution. However, since immunoassays could be frequently influenced by matrix effects, a closer-to-reality approach was applied. Therefore, a set of binary mixtures was prepared including different spiked levels (0, 0.1, 0.2, 0.5, 1, 2.5, 5, 7.5, 10 and 100 mg/g) of an un-treated wheat/barley/rye proportional mixture (WBRm) in a rice flour matrix [25]. Rice flour was selected as the matrix ingredient of the binary mixture due to the absence of background signal interference when tested in the indirect phage-Fab8E-4 ELISA method and since it is frequently used in gluten-free formulations. Previous studies have adopted a theoretical approach to estimate the amount of gluten as about 10% of the gluten-containing cereal weight [35]. Therefore, the binary mixtures used to assess the suitability of the developed immunoassay to detect low concentrations of gluten in a real food matrix could be interpreted as samples containing approximately 10, 20, 50, 100, 250, 500, 750, 1000 and 10,000 mg of gluten per kilogram (particles per million or ppm of gluten in a food sample). According to these experimental binary mixtures used for in-house validation and to derive the model features, antigen dose-response curve was fitted to a dose response model (Figure 4C). The LOD and LOQ obtained from the analysis of the gluten-spiked mixtures with the developed phage Fab8E-4 assay were 10 and 88 ppm of gluten, respectively. The test seemed to fulfill the legislation requirements, with the ability to distinguish the threshold of 20 ppm used to classify the products as gluten-free or not. The field application functionality of the proposed method was tested with ten food products (five labeled as gluten-free and five as containing gluten) after ethanolic extraction. The indirect phage-ELISA using Fab8E-4 was able to differentiate both types of products. These results were confirmed with a commercial sandwich ELISA based on the R5 monoclonal antibody, the gold standard for gluten analysis (Table S3).

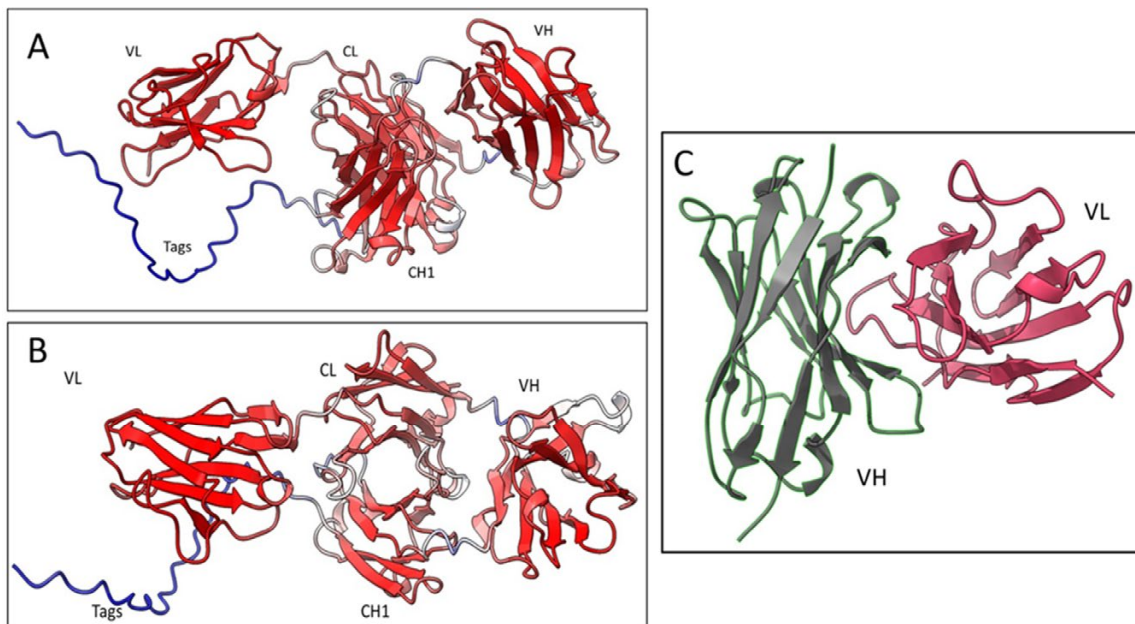
A Western blot analysis was performed to elucidate which prolamin fractions from different gluten-containing cereals were recognized by the developed recombinant antibody (Figure 4D). A commercial preparation of wheat gliadin was used as control because it contains all the different gliadin subunits. The results evidenced that the phage-Fab8E-4 was able to significantly recognize the protein bands between 25 and 48 kDa, that could correspond to the LMW glutenin subunits and alpha-, beta- and gamma-gliadins [35].

### 3.9. Fab8E-4 Sequence Analysis and Structure Prediction

In silico predictions of the Fab8E-4 structure and antibody-antigen interaction were performed to further characterize the antibody fragment obtained.

The library obtained in this work presents half the CDR diversity than a usual immune library, as it applies only to the light chains. Nevertheless, the overall affinity of the library is improved, because many of the phages of a full immune library might not bind properly to the target antigen, while in this library the heavy chains are constituted by dAb-8E gliadin-binding chains. Although this situation could lead to miss a “better” new Fab that could be found, or not, in a classical immune library, the selection of a minor expressed class, like IgG3, shows that the artificial rearrangement could generate functional structures that are very underrepresented in nature. This improvement is paired with an increase of Fab8E-4 affinity and avidity to gliadin, which derives from the new structure.

Based on the sequencing results of the Fab8E-4, the whole molecule 3D-structure prediction was generated using Alpha-fold2, an artificial intelligence program that performs predictions of protein structure with high accuracy [36]. Five different models were generated, and each model was recycled three times. The models were ranked by pLDDT score (a per-residue confidence estimation on a scale of 0–100). The fifth model (Figure 5A and 5B) was selected as the best, with an overall pLDDT score of 82,3 which was considered a good backbone prediction [37]. Besides, taking into account that the worst resolved area was the histidine tag attached to the CH1 domain which lowered the overall score, the Fab8E-4 domains were fairly well resolved with pLDDT scores >85.



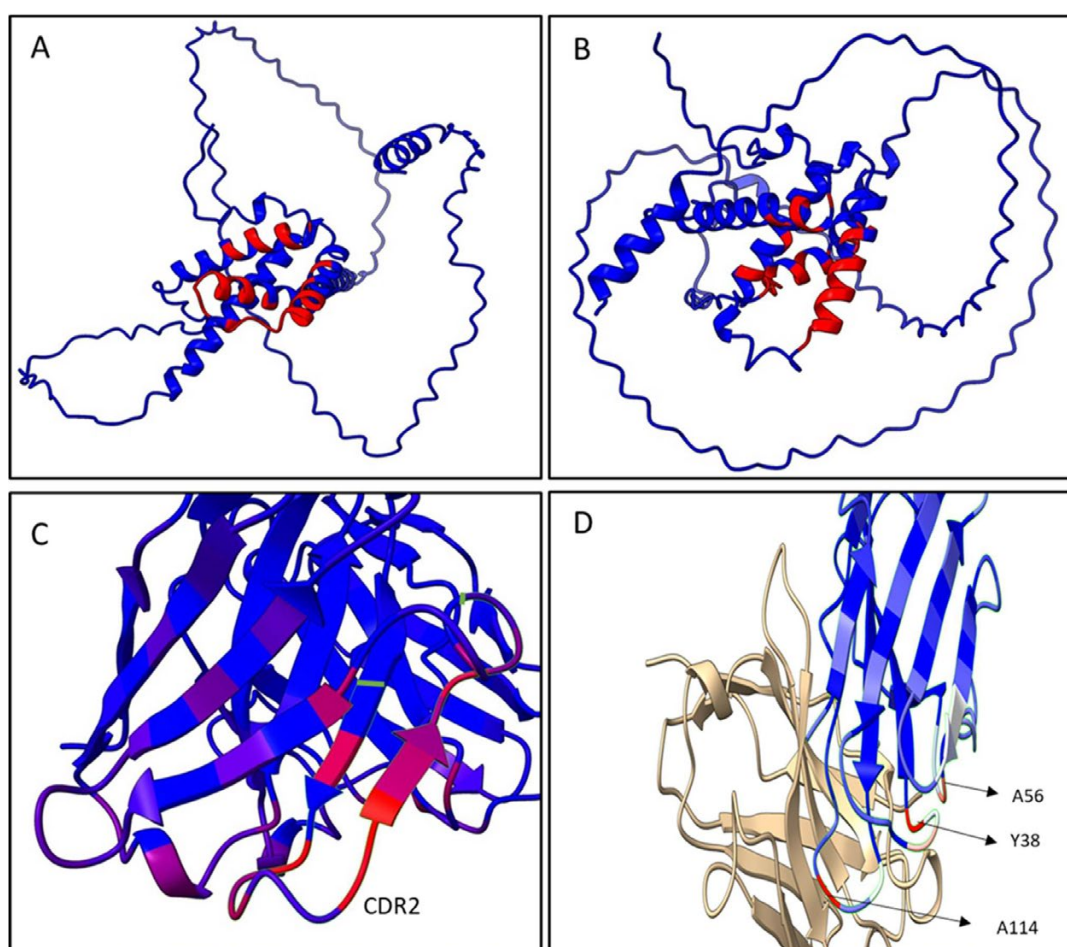
**Figure 5.** Computational models of the Fab8E-4. (A) Predictive model of Fab8E-4 generated with Alpha-Fold2. The pLDDT scores are shown with the following scale of color red > white > blue. (B) Same Alpha-Fold2 predicted model rotated for better visualization of the different domains. (C) Fv predictive model generated by Antibody-builder (sAbpred server).

In addition to the whole molecule model, a second model of the Fab was generated focused on the Fv (fragment variable of the antibody, composed of the VL and VH), the main antigen-antibody interaction surface. This model was computed using the AbodyBuilder tool included in the sAbPred server (Figure 5C). This bioinformatic tool encompasses up to seven different algorithms for the determination of the most likely antibody structure, such as identification of the best templates for VH and VL domains, prediction of the VH-VL orientation or modeling of CDR-loops and side chains. The background structures that were chosen as framework templates were (RCSB PDB identifiers) 6ghg (VH) and 6co3 (VL) as they showed the highest sequence identity for each chain (86.2% sequence identity for VH and 93.4% for VL). The VH-VL orientation was predicted using the angles inherited from the background structure 2uzi.

Once the 3D model of the Fab8E-4 was obtained, epitope and paratope prediction was assessed to investigate the antigen-antibody interaction surface. To identify the main amino acids involved in the binding region, knowledge of the 3D structure of both, the antigen and the Fv of the Fab, is required. However, crystal structures of the different gluten proteins were not available in databases such as Uniprot or RCSB PDB (which only include structures from some linear peptides presented by HLA that were not useful for the present work). Despite the lack of experimental data on crystal gluten structures, the structures of some gliadins and other gluten-related proteins have been recently solved with AlphaFold, and uploaded in Uniprot database. Therefore, in this study epitope and paratope predictions were performed with Epipred and Antibody-ipatch algorithms, respectively, using as protein templates the 3D antibody structure generated by AbodyBuilder and the antigen 3D-structure of wheat  $\alpha/\beta$ -gliadin and barley  $\gamma$ -hordein solved by AlphaFold (P18573 and P17990 Uniprot entries). These entries were chosen as they showed the highest quality of information recorded in the database. Until the structure of gliadins are not completely solved by X-ray crystallography or cryomicroscopy, this approach was chosen as the most adequate approximation for epitope and paratope identification of Fab8E-4.

The epitope predictive models depicted in Figure 6A,B show that the conformational epitopes were mainly found in the  $\alpha$ -helix motifs on the C-terminal region of both antigens,

the  $\alpha/\beta$ -gliadin and the  $\gamma$ -hordein. Within the gliadins studied, the main interaction zone of the dAb-8E and Fab8E-4 is located downstream one of the most pathogenic reported zones for celiac disease [38]. Besides, it presented more dispersed and shorter repeats of glutamines than those found in the N-terminal half. This fact could be explained because the heavy chain was originally derived from a semi-synthetic library instead of an immune one, which presumably would have resulted in an antibody binding the pathogenic N-terminal segment. In this sense, Fab8E-4 and their parental antibody fragment dAb-8E constitute a new alternative for gluten determination, providing likely contacts to the C-terminal region of the antigens, which could complement the already available antibodies that usually bound to the N-terminal region. This opens the possibility of further developments of sandwich and competitive ELISAs with antibodies binding different epitopes.



**Figure 6.** Epitope prediction (A,B) in  $\alpha/\beta$ -gliadin (A) and  $\gamma$ -hordein (B) facing Fab8E-4. The residues colored in red are more likely to be part of the conformational epitope. Notice that the images have been rotated to show the conformational epitope. Models generated with EpiPred (SAbPred server). Paratope prediction in Fab8E-4 (C,D). The residues are colored according to the likelihood of being part of the paratope following the order: red >white >blue. Notice that the images have been rotated to show the zones with a higher likelihood of being part of the paratope. The model was generated with antibody i-patch (SAbPred server). (C) In red the residues of the heavy chain identified to be more likely to interact with the antigen. CDR2 and neighboring sequences in a green silhouette. (D) In red the residues of the light chain identified to be more likely to interact with the antigen.

As shown in the paratope models obtained in Figure 6C,D, regardless the prolamin used in the analysis of the paratope models, the main antigen-antibody interaction remained in the heavy chain of Fab8E-4, which was derived from dAb-8E. The residues most likely to be in contact with the antigen were in the CDR-2 and neighboring heavy chain sequences (Figure 6C). However, reformatting of the dAb-8E added new strong contacts with the antigen through several light chain residues (A56, Y38 and A114), which presumably could have improved the characteristics of the detection method. In addition, the disulfide bond between the constant and the variable domains fixed the structure, reducing the conformational freedom of the protein [27].

#### 4. Conclusions

This work presents a novel approach to generate a suitable library for selection of antibody fragments for prolamin detection. This method goes beyond the direct transformation of an antibody format to another (dAb into Fab), implying a direct evolution process in order to improve the features of the previously obtained dAb-8E. A library composed of light chains and heavy chains from different origin was built. The designed SOE-PCR reactions allowed a quick conversion of the dAb sequence into a Fab heavy chain and, once cloned into the subset of light chains obtained from celiac donors, the library was obtained. The high diversity of the constructed library resulted in the successful isolation of positive clones. After characterization of different phage variants, the Fab8E-4 was selected and tested in an immunoassay for the specific and sensitive detection of gluten. Besides, a deeper study of the generated antibody was developed using *in silico* tools. The proposed ELISA method demonstrated that transformation of the dAb-8E into a Fab structure resulted in an improvement of the characteristics of the developed methodologies for gluten detection in foodstuff. Therefore, the obtained Fab8E-4 presented enhanced properties regarding the parental dAb, indicating that the proposed methodology was a useful tool to improve the performance of existing antibodies.

The phage-Fab8E-4 obtained and characterized in this work has demonstrated ability to detect gliadin in solution, an also in complex mixtures. Thus, it can be directly applied for gluten detection in foodstuff. Moreover, thanks to the *in silico* model, that showed contacts with the C-terminal gliadin portion, it could be used as a complementary or alternative reagent to the other antibodies in the market for gluten detection, that bind to the N-terminal portion of the antigen.

The development of gluten detection assays has been historically focused on hybridoma generated monoclonal antibodies (R401, R5, G12). Seldom recombinant antibodies have been used for solving this problem, and mainly they were applied for studying clinical features of gluten-related diseases rather than gluten detection in foods. The present work is a novel rational approach that goes beyond the construction of an immune library from celiac patients. The construction of a novel Fab library is provided, beginning with the scaffold of a pre-existing antibody selected against likely pathogenic gluten peptides, transformed into a Fab heavy chain, and merged with a light chain subset derived from celiac patients. This is the first time a directed evolved engineered antibody fragment has been applied for gluten detection in foodstuffs.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/foods12010149/s1>, Figure S1: Transformants recovery after panning rounds; Figure S2: Indirect phage-ELISA analysis of the phage-Fab library rescued after each round of panning; Table S1: Primers used for heavy chain amplification; Table S2: Primers used for light chain amplification; Table S3: Sequencing primers; Table S4: ELISA results obtained for detection of gluten with the phage Fab8E-4 and R5 monoclonal antibodies from ethanolic extracts of 10 commercial food products.

**Author Contributions:** E.G.-C.: methodology, investigation, experimentation, validation, formal analysis, writing original draft. A.G.-G.: investigation, formal analysis, and writing and revision. S.R.: investigation and formal analysis. S.F.: celiac patient selection. R.M.: funding acquisition and project

management. T.G.: supervision, writing and revision, funding acquisition and project management. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by Ministerio de Ciencia e Innovación Grant: PID2021-122925OB-I00. Comunidad de Madrid and Fondo Social Europeo (FSE) S2018/BAA-4574. PRE2018-08427 fellowship of the Ministerio de Ciencia e Innovación granted for Eduardo Garcia-Calvo.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Hospital Universitario Fundación Jiménez Díaz (protocol code October 2018 and date of approval, 24 September 2019).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data is contained within the article.

**Acknowledgments:** We thank the Immunology and Citogenetics Services of Fundación Jiménez Díaz Hospital for the management of the clinical samples used in this work. Biorender was used for drawing and editing the graphical abstract.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Codex-Alimentarius. Standard for foods for special dietary use for persons intolerant to gluten. *CSX* **2018**, *118*, 1979.
2. Cebolla, Á.; Moreno, M.; Coto, L.; Sousa, C. Gluten Immunogenic Peptides as Standard for the Evaluation of Potential Harmful Prolamin Content in Food and Human Specimen. *Nutrients* **2018**, *10*, 1927. [[CrossRef](#)] [[PubMed](#)]
3. Leonard, M.M.; Sapone, A.; Catassi, C.; Fasano, A. Celiac Disease and Nonceliac Gluten Sensitivity: A Review. *JAMA* **2017**, *318*, 647–656. [[CrossRef](#)]
4. Sharma, N.; Bhatia, S.; Chunduri, V.; Kaur, S.; Sharma, S.; Kapoor, P.; Kumari, A.; Garg, M. Pathogenesis of Celiac Disease and Other Gluten Related Disorders in Wheat and Strategies for Mitigating Them. *Front. Nutr.* **2020**, *7*, 6. [[CrossRef](#)] [[PubMed](#)]
5. Sapone, A.; Bai, J.C.; Ciacci, C.; Dolinsek, J.; Green, P.H.; Hadjivassiliou, M.; Kaukinen, K.; Rostami, K.; Sanders, D.S.; Schumann, M.; et al. Spectrum of gluten-related disorders: Consensus on new nomenclature and classification. *BMC Med.* **2012**, *10*, 13. [[CrossRef](#)]
6. Volta, U.; De Giorgio, R.; Caio, G.; Uhde, M.; Manfredini, R.; Alaedini, A. Nonceliac Wheat Sensitivity: An Immune-Mediated Condition with Systemic Manifestations. *Gastroenterol. Clin. N. Am.* **2019**, *48*, 165–182. [[CrossRef](#)]
7. Scherf, K.A.; Poms, R.E. Recent developments in analytical methods for tracing gluten. *J. Cereal Sci.* **2016**, *67*, 112–122. [[CrossRef](#)]
8. García-García, A.; Madrid, R.; Sohrabi, H.; Cruz, S.; García, T.; Martín, R.; Gonzalez Alonso, I. A sensitive and specific real-time PCR targeting DNA from wheat, barley and rye to track gluten contamination in marketed foods. *LWT* **2019**, *114*, 108378. [[CrossRef](#)]
9. Lexhaller, B.; Colgrave, M.L.; Scherf, K.A. Characterization and Relative Quantitation of Wheat, Rye, and Barley Gluten Protein Types by Liquid Chromatography–Tandem Mass Spectrometry. *Front. Plant Sci.* **2019**, *10*, 01530. [[CrossRef](#)]
10. Panda, R.; Garber, E.A.E. Detection and Quantitation of Gluten in Fermented-Hydrolyzed Foods by Antibody-Based Methods: Challenges, Progress, and a Potential Path Forward. *Front. Nutr.* **2019**, *6*, 97. [[CrossRef](#)]
11. Garcia-Calvo, E.; García-García, A.; Madrid, R.; Martín, R.; García, T. From Polyclonal Sera to Recombinant Antibodies: A Review of Immunological Detection of Gluten in Foodstuff. *Foods* **2020**, *10*, 66. [[CrossRef](#)]
12. Manohar, R.S. Development of Enzyme-Linked Immunosorbent Assay for Evaluation of Chapati-Making Quality of Wheat Varieties. *J. Agric. Food Chem.* **2002**, *50*, 7455–7460. [[CrossRef](#)]
13. Gujral, N.; Suh, J.W.; Sunwoo, H.H. Effect of anti-gliadin IgY antibody on epithelial intestinal integrity and inflammatory response induced by gliadin. *BMC Immunol.* **2015**, *16*, 41. [[CrossRef](#)]
14. Skerritt, J.H.; Hill, A.S. Monoclonal antibody sandwich enzyme immunoassays for determination of gluten in foods. *J. Agric. Food Chem.* **1990**, *38*, 1771–1778. [[CrossRef](#)]
15. Sorell, L.; López, J.A.; Valdés, I.; Alfonso, P.; Camafeita, E.; Acevedo, B.; Chirido, F.; Gabilondo, J.; Méndez, E. An innovative sandwich ELISA system based on an antibody cocktail for gluten analysis. *Fed. Eur. Biochem. Soc. Lett.* **1998**, *439*, 46–50. [[CrossRef](#)]
16. Morón, B.; Bethune, M.T.; Comino, I.; Manyani, H.; Ferragud, M.; López, M.C.; Cebolla, Á.; Khosla, C.; Sousa, C. Toward the Assessment of Food Toxicity for Celiac Patients: Characterization of Monoclonal Antibodies to a Main Immunogenic Gluten Peptide. *PLoS ONE* **2008**, *3*, e2294. [[CrossRef](#)]
17. Méndez, E.; Vela, C.; Immer, U.; Janssen, F.W. Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *Eur. J. Gastroenterol. Hepatol.* **2005**, *17*, 1053–1063. [[CrossRef](#)]
18. Peltomaa, R.; Barderas, R.; Benito-Peña, E.; Moreno-Bondi, M.C. Recombinant antibodies and their use for food immunoanalysis. *Anal. Bioanal. Chem.* **2022**, *414*, 193–217. [[CrossRef](#)]
19. Schroeder, H.W.; Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **2010**, *125*, S41–S52. [[CrossRef](#)]

20. Arbabi-Ghahroudi, M. Camelid Single-Domain Antibodies: Historical Perspective and Future Outlook. *Front. Immunol.* **2017**, *8*, 1589. [[CrossRef](#)]
21. Smith, G.P. Phage Display: Simple Evolution in a Petri Dish (Nobel Lecture). *Angew. Chem. Int. Ed.* **2019**, *58*, 14428–14437. [[CrossRef](#)] [[PubMed](#)]
22. Not, T.; Ziberna, F.; Vatta, S.; Quaglia, S.; Martellosi, S.; Villanacci, V.; Marzari, R.; Florian, F.; Vecchiet, M.; Sulic, A.-M.; et al. Cryptic genetic gluten intolerance revealed by intestinal antitransglutaminase antibodies and response to gluten-free diet. *Gut* **2011**, *60*, 1487–1493. [[CrossRef](#)] [[PubMed](#)]
23. Chen, T.; Hoffmann, K.; Östman, S.; Sandberg, A.-S.; Olsson, O. Identification of gliadin-binding peptides by phage display. *BMC Biotechnology* **2011**, *11*, 16. [[CrossRef](#)] [[PubMed](#)]
24. Doña, V.; Urrutia, M.; Bayardo, M.; Alzogaray, V.; Goldbaum, F.A.; Chirido, F.G. Single Domain Antibodies Are Specially Suited for Quantitative Determination of Gliadins under Denaturing Conditions. *J. Agric. Food Chem.* **2010**, *58*, 918–926. [[CrossRef](#)] [[PubMed](#)]
25. Jayathilake, C.; Kumachi, S.; Arai, H.; Motohashi, M.; Terai, T.; Murakami, A.; Nemoto, N. In vitro selection of anti-gliadin single-domain antibodies from a naïve library for cDNA-display mediated immuno-PCR. *Anal. Biochem.* **2020**, *589*, 113490. [[CrossRef](#)]
26. García-García, A.; Madrid, R.; González, I.; García, T.; Martín, R. A novel approach to produce phage single domain antibody fragments for the detection of gluten in foods. *Food Chem.* **2020**, *321*, 126685. [[CrossRef](#)]
27. Torres, M.; Casadevall, A. The immunoglobulin constant region contributes to affinity and specificity. *Trends Immunol.* **2008**, *29*, 91–97. [[CrossRef](#)]
28. Pritsch, O.; Magnac, C.; Dumas, G.; Bouvet, J.-P.; Alzari, P.; Dighiero, G. Can isotype switch modulate antigen-binding affinity and influence clonal selection? *Eur. J. Immunol.* **2000**, *30*, 3387–3395. [[CrossRef](#)]
29. Barbas, C.F. *Phage Display: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2001.
30. Burton, D.R.; Barbas, C.F.; Persson, M.A.; Koenig, S.; Chanock, R.M.; Lerner, R.A. A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 10134–10137. [[CrossRef](#)]
31. Brochet, X.; Lefranc, M.P.; Giudicelli, V. IMGT/V-QUEST: The highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* **2008**, *36*, W503–W508. [[CrossRef](#)]
32. Ye, J.; Ma, N.; Madden, T.L.; Ostell, J.M. IgBLAST: An immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res.* **2013**, *41*, W34–W40. [[CrossRef](#)]
33. Dunbar, J.; Krawczyk, K.; Leem, J.; Marks, C.; Nowak, J.; Regep, C.; Georges, G.; Kelm, S.; Popovic, B.; Deane, C.M. SAbPred: A structure-based antibody prediction server. *Nucleic Acids Res.* **2016**, *44*, W474–W478. [[CrossRef](#)]
34. David, A.; Islam, S.; Tankhilevich, E.; Sternberg, M.J.E. The AlphaFold Database of Protein Structures: A Biologist's Guide. *J. Mol. Biol.* **2022**, *434*, 167336. [[CrossRef](#)]
35. Wan, Y.; Gritsch, C.S.; Hawkesford, M.J.; Shewry, P.R. Effects of nitrogen nutrition on the synthesis and deposition of the  $\omega$ -gliadins of wheat. *Ann. Bot.* **2014**, *113*, 607–615. [[CrossRef](#)]
36. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
37. Monzon, V.; Haft, D.H.; Bateman, A. Folding the unfoldable: Using AlphaFold to explore spurious proteins. *Bioinform. Adv.* **2022**, *2*, vbab043. [[CrossRef](#)]
38. Shan, L.; Molberg, Ø.; Parrot, I.; Hausch, F.; Filiz, F.; Gray, G.M.; Sollid, L.M.; Khosla, C. Structural basis for gluten intolerance in celiac sprue. *Science* **2002**, *297*, 2275–2279. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

## Supplementary Material

**Table S1.** Primers used for heavy chain amplification

Name	Sequence
<b>First PCR round</b>	
VH135 fw	AGG TGC AGC TGC TCG AGT CTG G
ovlp-conga rev	AAG ACC GAT GGG CCC TTG GTG GAT GCG GCC GCG CTC GAT ACG G
ovlp-cong4 rev	ACG GAT GGG CCC TTG GTG GAA GCT GCG GCC GCG CTC GAT ACG G
<b>Second PCR round</b>	
VH135ext fw	(GTC) <sup>7</sup> G TAG GTG CAG CTG CTC GAG TC
CG1Zext rev	(GA) <sup>10</sup> G GCA TGT ACT AGT TTT GTC A C
CG2a ext rev	(GA) <sup>10</sup> CTC GAC ACT AGT TTT GCG
CG3a ext rev	(GA) <sup>10</sup> TGT GTG ACT AGT GTC ACC
CG4a ext rev	(GA) <sup>10</sup> GCA TGA ACT AGT TGG GGG

**Table S2.** Primers used for light chain amplification

Name	Sequence
<b>First PCR round</b>	
Vk1 fw	GAG CCG CAC GAG CCC GAG CTC CAG ATG ACC CAG TCT CC
Vk2/4 fw	GAG CCG CAC GAG CCC GAG CTC GTG ATG ACY CAG TCT CC
Vk3 fw	GAG CCG CAC GAG CCC GAG CTC GTG WTG ACR CAG TCT CC
Vk5 fw	GAG CCG CAC GAG CCC GAG CTC ACA CTC ACG CAG TCT CC
Ck1d rev	GCG CCG TCT AGA ATT AAC ACT CTC CCC TGT TGA AGC TCT TTG TGA CGG GCG AAC TCA G
<b>Second PCR round</b>	
Vk ext fw	GCG CCG TCT AGA ATT AAC ACT CTC CCC TGT TGA AGC TCT TTG TGA CGG GCG AAC TCA G
Ck ext rev	GAG GAG GAG GAG GAG GAG GCG CCG TCT AGA ATT AAC ACT CTC

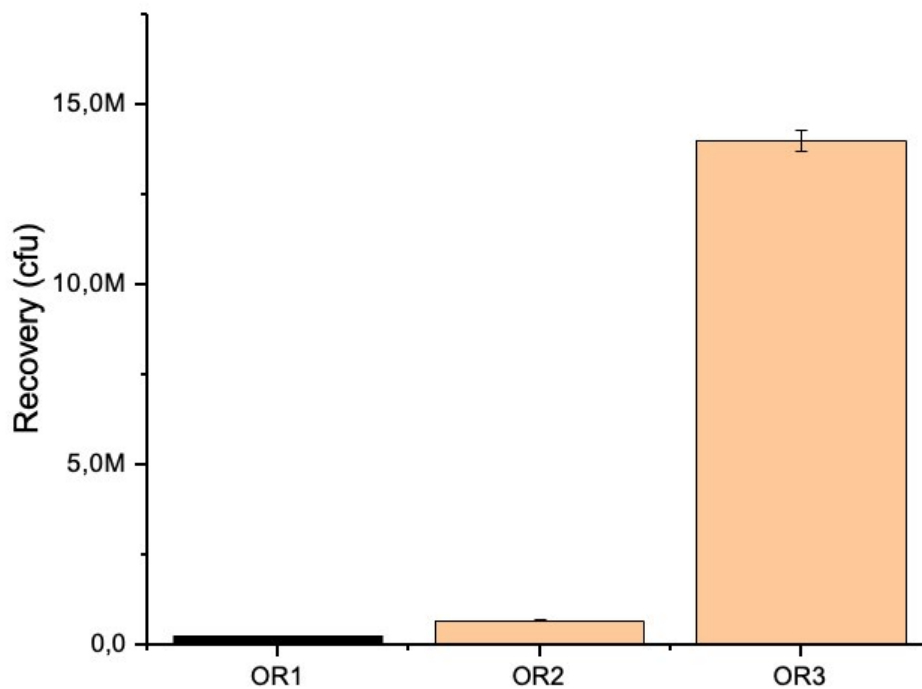
**Table S3.** Sequencing primers

Name	Sequence
ompAseq	AAG ACA GCT ATC GCG ATT GCA G
g-back	GCC CCC TTA TTA GCG TTT GCC ATC

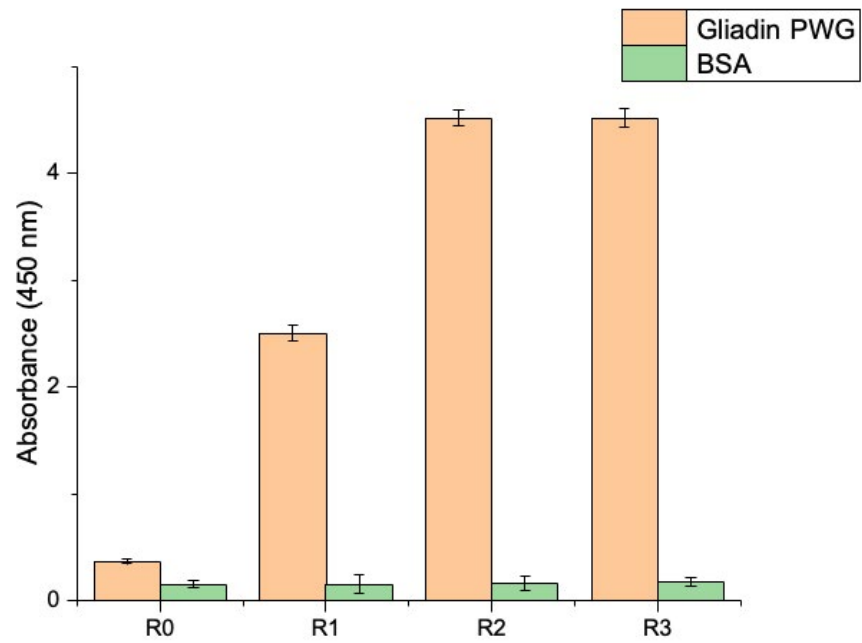
Degenerated basZes: W=A,T; R=A, G.

**Table S4.** ELISA results obtained for detection of gluten with the phage Fab8E-4 and R5 monoclonal antibodies from ethanolic extracts of 10 commercial food products.

Nº	Sample	Gluten Label	Fab8E-4	R5
1	Macarrons (Bakery)	+	+	+
2	Apple cake	+	+	+
3	Free-sugar biscuits	+	+	+
4	Caramelized biscuits	+	+	+
5	Cake	+	+	+
6	Sausages	-	-	-
7	Vegan sausages	-	-	-
8	Surimi	-	-	-
9	Nuts mixture	-	-	-
10	Sunflower seeds	-	-	-



**Figure S1.** Transformants recovery after panning rounds. An inference was calculated from the colonies grown in after plating in SB-carbenicilin (100 µg/mL) infected E.coli XL-1 with the output phages from each selection round. Mean values of two independent determinations and standard derivation of each data set are shown.



**Figure S2.** Indirect phage-ELISA analysis of the phage-Fab library rescued after each round of panning. Results of absorbance value (450 nm) obtained against gliadin-PWG (objective antigen) and BSA (blocking solution). Mean values of three independent determinations and standard derivation of each data set are shown.

## CAPÍTULO 4.2

### Desarrollo de un nuevo anticuerpo recombinante, seleccionado mediante la tecnología del *phage display* a partir de un repertorio derivado de pacientes celíacos, para la detección de gluten en alimentos

---

### Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods

Gluten, a group of ethanol-soluble proteins present in the endosperm of cereals, is extensively used in the food industry due to its ability to improve dough properties. However, gluten is also associated with a range of gluten-related diseases (GRDs), such as wheat allergies, celiac disease, and gluten intolerance. The recommended treatment for GRDs patients is a gluten-free diet. To monitor adherence to this diet, it is necessary to develop gluten-detection systems in food products. Among the available methods, immunodetection systems are the most popular due to their simplicity, reproducibility, and accuracy. The aim of this study was to generate novel high-affinity antibodies against gluten to be used as the primary reactant in an enzyme-linked immunosorbent assay (ELISA) test. These antibodies were developed by constructing an immune library from mRNA obtained from two celiac patients with a high humoral response to gluten-related proteins. The resulting library (composed by  $1.1 \times 10^7$ ) was subjected to selection against gliadin using phage display technology. Following several rounds of selection, the Fab-C was selected, and demonstrated good functionality in ELISA tests, presenting a limit of detection of 15 mg/kg for detection of gluten in spiked mixtures and food products. The methodology can discriminate gluten-free products according to the current legislation.

**García-Calvo, Eduardo; García-García, Aina; Rodríguez-Gómez, Santiago; Farras, Sergio; Martín, Rosario; & García, Teresa.** (2023). Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods. *Current Research in Food Science*, 7, 100578

<https://doi.org/https://doi.org/10.1016/j.crfs.2023.100578>

Journal category: Food Science and Technology

Journal Impact Factor (2022): 6.3

Journal Rank: 21/142 (Q1)







ELSEVIER

Contents lists available at ScienceDirect

## Current Research in Food Science

journal homepage: [www.sciencedirect.com/journal/current-research-in-food-science](http://www.sciencedirect.com/journal/current-research-in-food-science)

## Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods

Eduardo García-Calvo<sup>a</sup>, Aina García-García<sup>a,\*</sup>, Santiago Rodríguez-Gómez<sup>a</sup>, Sergio Farras<sup>b</sup>, Rosario Martín<sup>a,1</sup>, Teresa García<sup>a,1</sup>

<sup>a</sup> Departamento de Nutrición y Ciencia de los Alimentos, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040, Madrid, Spain

<sup>b</sup> Servicio de Medicina Digestiva, Hospital Universitario Fundación Jiménez Díaz, 28040, Madrid, Spain

## ARTICLE INFO

Handling Editor: Dr. Yeonhwa Park

## Keywords:

Gluten  
Celiac disease  
Phage display  
ELISA  
Immune library  
Recombinant Fab

## ABSTRACT

Gluten, a group of ethanol-soluble proteins present in the endosperm of cereals, is extensively used in the food industry due to its ability to improve dough properties. However, gluten is also associated with a range of gluten-related diseases (GRDs), such as wheat allergies, celiac disease, and gluten intolerance. The recommended treatment for GRDs patients is a gluten-free diet. To monitor adherence to this diet, it is necessary to develop gluten-detection systems in food products. Among the available methods, immunodetection systems are the most popular due to their simplicity, reproducibility, and accuracy. The aim of this study was to generate novel high-affinity antibodies against gluten to be used as the primary reactant in an enzyme-linked immunosorbent assay (ELISA) test. These antibodies were developed by constructing an immune library from mRNA obtained from two celiac patients with a high humoral response to gluten-related proteins. The resulting library (composed by  $1.1 \times 10^7$ ) was subjected to selection against gliadin using phage display technology. Following several rounds of selection, the Fab-C was selected, and demonstrated good functionality in ELISA tests, presenting a limit of detection of 15 mg/kg for detection of gluten in spiked mixtures and food products. The methodology can discriminate gluten-free products according to the current legislation.

### 1. Introduction

Gluten-related disorders (GRDs) have an estimated global prevalence of 5%. Nowadays, GRDs are considered a significant global health issue, mainly in Western countries (Taraghihah et al., 2020). GRDs are classified by their etiopathology into three main groups: Allergy to gluten or its components, autoimmune diseases, and non-celiac gluten sensitivity.

Celiac disease is the most prevalent within the second group, presenting a 1- 2% prevalence in the Caucasian population (Peña and Rodrigo, 2015). The molecular mechanisms of celiac pathology are well described (Caio et al., 2019). The diagnosis of celiac disease is closely related to the autoimmune response described, and usually implies several tests like: excluding selective IgA deficiency, high serum levels of IgA and IgG anti-tTG, serum IgA and IgG anti-deaminated gliadin peptide (DGP), serum IgA endomysial antibodies, and HLA-DQ typing (genetic screening). In addition to the serological and genetic tests, the diagnosis is usually confirmed by a gut biopsy (Raiteri et al., 2022).

Despite several therapies have been proposed to mitigate celiac disease (anti-inflammatory drugs, inhibitory monoclonal antibodies, gluten chelators ...), even nowadays, a gluten-free diet is the best therapy that not only inhibits the autoimmune gut tissue destruction, but also conducts to a correct restoration of the intestinal crypts and the digestive function (Itzlinger et al., 2018). Due to the high prevalence of GRDs and the efficacy of a gluten-free diet to treat them, it was necessary to develop systems for gluten detection in foods. Several methodologies have been developed (García-García et al., 2018), but immunoassays are the most used, due to their accuracy and simplicity. The main reactants of these methods are high-affinity antibodies to gluten. Polyclonal and monoclonal approaches have been used, outstanding the monoclonal antibody R5, obtained using a classical hybridoma strategy after mice immunization with gluten-containing flour extracts. A sandwich ELISA using the R5 antibody is the golden standard method for gluten detection in foodstuffs (Valdés et al., 2003). Despite monoclonal antibodies are widely used nowadays, new developments are needed, based on

\* Corresponding author.

E-mail address: [ainagarcia@ucm.es](mailto:ainagarcia@ucm.es) (A. García-García).

<sup>1</sup> These authors share senior authorship.

<https://doi.org/10.1016/j.crfs.2023.100578>

Received 18 June 2023; Received in revised form 28 July 2023; Accepted 24 August 2023

Available online 25 August 2023

2665-9271/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

novel technologies like recombinant antibodies, that present several advantages like scalability, customizability, fast and consistent production, and avoidance of the expression of unwanted chains that occurs with hybridomas (Bradbury et al., 2018).

The Fab (fragment antigen binding region) is a recombinant antibody fragment composed of two chains: the complete light chain (variable and constant regions) and a heavy chain, that consists in the variable domain and only one constant region (CH<sub>1</sub>) of the mammal immunoglobulins, but Fabs lack the Fc (fragment crystallizable) region. This antibody format was chosen because it replicates the region that binds the antigens in natural antibodies, providing the complete paratope pocket. Also, Fabs are more stable than single chain formats due to the disulfide bond that joins the heavy and the light chains, and their smaller size (compared to complete immunoglobulins) enhances their diffusive properties and facilitates the recombinant expression in prokaryotic cells (Kang and Seong, 2020).

The main aim of this work was to obtain new alternatives to gluten immunodetection in foods, based on the development of recombinant Fabs presenting high-affinity against gluten.

## 2. Materials and methods

### 2.1. Bacterial strains, plasmids, and growth media

*Escherichia coli* XL1-Blue strain (*recA1, endA1, gyrA96, thi-1, hsdR17, supE44 relA1, lac [F proAB, lacIqZΔM15, Tn10 (Tetr)]*) (Agilent©, Santa Clara, CA, USA, ref #200150) was used for cloning, building the libraries, and for production of phage-displayed Fabs. Electrocompetent cells were in-house produced by growing *E. coli* XL1-Blue in Super-Broth medium (SB: 30 g/L tryptone, 20 g/L yeast extract, 10 g/L MOPS, pH 7) and concentrating by sequential cycles of centrifugation (3200 g, 4 °C) and resuspension in 10% glycerol (v/v) (PanReac AppliChem© Monza, Lombardy, Italy, CAS: 56-81-5) in water. After electroporation, cells were grown in SOC medium (Invitrogen™-Thermo Fisher©, Waltham, MA, USA, ref #15544-034). For DNA extraction before sequencing, bacteria were grown in Luria Broth (LB: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7). Agar plates were prepared with a 15 g/L agar concentration.

The antibody chain coding fragments were cloned into the plasmid pComb3X (Andris-Widhopf et al., 2000).

### 2.2. Antigen preparations

Glutadin-PWG (Arbeitsgemeinschaft Getreideforschung (AGF) e.V. DIGeFa GmbH, Detmold, Germany) was used as reference material for gluten detection in foods. Following the supplier's recommendation, glutadin-PWG was dissolved in a 60% ethanol/water solution at 1 mg/mL and stored at room temperature protected from light. Gluten-like proteins were also extracted from different flour cereal matrices such as wheat, rye, barley, corn, and rice, as well from a binary mixture of wheat, rye, and barley kernels prepared in a rice-based gluten-free matrix as previously described (García-García et al., 2020). Gluten-like proteins were extracted by dissolving 0.25 g of the finely ground sample in 10 mL of 60% ethanol, followed by shaking for 20 min at room temperature and centrifugation at 2000g for 10 min at room temperature.

### 2.3. Selection of celiac patients as suitable lymphocytes donors

Prospective donors were patients attending the digestive medicine service of the Fundación Jiménez Díaz Hospital (Madrid, Spain), with symptoms of a prototypical celiac disease humoral response. The following diagnostic parameters were used to assess the patient eligibility: IgG and IgA tissular anti-transglutaminase (anti-tTG), IgG and IgA anti-deaminated gliadin peptide (DGP).

In addition, patients with selective IgA deficiency were not

considered potential donors. It was essential that the patients would not have started a free-gluten diet to assure the maximum of gluten-reactive B-cells in their peripheral blood (Agardh et al., 2006).

Once a preliminary search was done based on the diagnosis of celiac disease of the prospective donors, the humoral response profile to complete gliadin was characterized. The ELISA tests α-Gliatest S Chromo IgA (Eurospital, Trieste, Italy, ref #910796) and α-Gliatest S Chromo IgG (Eurospital, ref #910896) were used for that purpose, according to the manufacturers' protocols.

An in-house ELISA test was implemented to quantify the amount of each immunoglobulin sub-class expressed in the patients. Immunosorb plates (96-well Maxisorp™ plates, Thermo Fisher©, Waltham, MA, USA, ref #44-2404-21) were coated for 1 h at 37 °C with 100 μL of 3.3 μg/mL of gliadin PWG in 0.1 M carbonate/bicarbonate buffer pH 9.6. The coating solution was removed and, after three washing steps with PBS, 200 μL of blocking solution (3% BSA in PBS) was added to each well. The plate was incubated at room temperature for 1 h and washed three times with PBS-T (PBS containing 0.05% of Tween (Sigma©, Burlington, MA, USA, ref #P1379)). Then, 100 μL of patients' serum dilution (1:200 for IgG1 and IgG2 tests and 1:50 for IgG3 and IgG4) in sterile PBS was added per well. After incubating the plate at room temperature for 2 h, three washing steps were performed with PBS-T. One hundred microliters of anti-human sub-isotypes (Cygnus Technologies LLC, Southport, USA, ref #IM151 to #IM154) diluted 1:1000 in blocking solution were added. The plate was incubated for 1 h at room temperature and washed three times with PBS-T. One hundred microliters of TMB (Sigma©, ref #T0440) was then added, and the reaction was stopped after 15–20 min with 50 μL of a diluted sulphuric acid solution. Absorbance readings were performed at 450 nm (FluoStar Optima™ from BMG labtech©, Ortenberg, Germany). The serum of a non-celiac person was used as negative control.

The results obtained were normalized to consider the different dilution of the serum used for each isotype. The normalization index (*i*) was calculated as follows:

$$i = \left( \frac{AU_1}{AU_2} \right) \times \text{serum dilution}$$

Where: AU1 are the absorbance units of the donors' serum samples and AU2 are the absorbance units of the non-reactive serum sample used as negative control.

### 2.4. Lymphocyte RNA preparation and retro-transcription

Blood samples were collected from selected donors following the protocols approved by the human research ethics committee of Fundación Jiménez-Díaz (Madrid). After signing the informed consent, 350 mL of peripheral blood were extracted from each donor, and cells were counted in a Neubauer chamber after Trypan Blue staining. Total RNA was extracted by a standard method with TriZol LS reagent (Invitrogen™-Thermo Fisher©, ref #10296010). Once the RNA quality and quantity were checked by Agilent© 2100 Bioanalyzer (ref #G2939BA), it was used as a template in reverse transcription reactions (SuperScript IV®, Invitrogen™- Thermo Fisher©, ref #18091050), following the manufacturer protocol, but the reaction was performed at 52.5 °C for 1 h instead of 10 min.

### 2.5. Amplification and isolation of antibody chain coding fragments

The antibody chain coding fragments were amplified in two rounds of PCR. The first reaction was performed for specific amplification of the genes, and the second just to add sequences flanking the restriction sites for cloning improvement. Light chains and Fab heavy chains (VH and CH1 domains) were amplified separately using the primer set developed by Barbas et al. (Barbas, 2001) and Table 1S. For the first round, the reactions were set up with the following proportions: 1.5 μL of cDNA; 60

pmol of each primer (synthesized by Eurofins genomics®, Luxemburg, Luxemburg); 8 µL of 2.5 mM dNTPs set (Thermo Fisher®, ref #AB0196); 10 µL of 10X AmpliTaqGold® buffer; 0.5 µL AmpliTaqGold® (Thermo Fisher®, ref #4311806) and molecular biology grade water (Sigma®, CAS: 7732-18-5) to a final volume of 100 µL. Amplification was performed under the following cycling conditions: Enzyme activation at 94 °C for 5 min; 35 cycles with denaturation at 94 °C for 15 s, primer annealing at 56 °C for 15 s, primer extension at 72 °C for 90 s; and a final extension step at 72 °C for 10 min. The amplified DNA fragments were isolated by gel electrophoresis using UltraPure™ Low Melting Point agarose (Thermo-fisher®, ref #16520050), and purified with NucleoSpin® gel DNA clean-up columns (Machery-Nagel®, Allentown, PA, USA, ref #740609). DNA was quantified in a Qubit® Fluorometer (Invitrogen Thermo-fisher®), and its quality was measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Montchanin, DE, USA). For the second round of PCRs (Table 1S), the reactions were set up with the same proportions of reactants as in the first round, but using 100 ng of purified first-round PCR product as a template, and the following cycling conditions: 5 min at 94 °C for enzyme activation; 15 cycles with denaturation at 94 °C for 15 s, primer annealing at 54 °C for 15 s, primer extension at 72 °C for 2 min and, a final extension step at 72 °C for 10 min. Then, the product purification was performed as in the previous round.

## 2.6. Library construction (clonation and transformation)

The purified PCR products and the plasmid pComb3X were digested with the corresponding enzymes and then, ligated in a two-step process. In the first step, the light chain coding fragments and the plasmid were digested for 1 h at 37 °C with *SacI* and *XbaI* (New England Biolabs® Ipswich, MA, USA ref #R0156 and ref #R0145), following manufacturers' set-up reactions, and then purified by agarose electrophoresis as previously described. Digested inserts and linearized vector were ligated (3:1 M ratio) overnight at 16 °C with T4 DNA ligase. The resulting ligation was ethanol precipitated, re-suspended in 20 µL of water, and transformed into high-efficiency electrocompetent *Escherichia coli* XL1-Blue. SOC recovery medium (3 mL) was then added to the transformed cells before incubation at 250 rpm for 1 h at 37 °C. At this point, 10 µL of the culture was removed, diluted, and titrated in LB agar plates containing 100 µg/mL of carbenicillin. The rest of the culture was transferred to 12 mL of LB broth with 20 µg/mL of carbenicillin and 10 µg/mL of tetracycline and incubated for an additional hour. The culture was then expanded to 100 mL of LB (carbenicillin at 50 µg/mL and tetracycline at 10 µg/mL) and grown overnight at 37 °C. Phagemid DNA containing the immune light chain repertoire was purified by midi-prep (PureLink™ HiPure plasmid Miniprep Kit Invitrogen™-Thermo Fisher® ref #K210004) and digested alongside the heavy chain coding fragments with *XhoI* and *SpeI* (New England Biolabs® ref #R0146 and ref #R0133). After gel purification, the resulting fragments were cloned following the steps described above for the light chain. The ligation reaction (containing the Fab library) was ethanol precipitated and transformed into *E. coli* XL1-Blue. Following electroporation, 5 mL of recovering SOC medium was added, and the culture was shaken at 250 rpm for 1 h at 37 °C. Then, 10 µL of the culture was removed, and ten-fold dilutions were plated into SB plates (containing 100 µg/mL of carbenicillin) to calculate the repertoire size of the library. The following steps of culture amplification, phage induction and isolation were done using the protocol depicted (Barbas, 2001).

## 2.7. Selection of gliadin-binding phage-Fabs

The panning procedure was performed as previously described (Garcia-Calvo et al., 2022). Briefly, 1 µg of antigen (gliadin PWG) diluted in coating solution (0.1 M carbonate/bicarbonate buffer, pH 9.6) was added per well in a microtiter plate. The plates were washed and blocked with PBS-BSA 3%, incubated for 1 h at 37 °C and washed three

times with PBS. Then, 100 µL of isolated phages were added per well, and incubated at 37 °C for 2 h. Non-binding phages were washed out with PBS-T and binding ones were eluted with acidic conditions. The eluted phage-Fabs were amplified by infecting a culture of *E. coli* XL1-Blue, the helper phage VSCM13 was added and grown overnight. Phage-Fabs were precipitated with PEG8000 (polyethylene-glycol) and sodium chloride. The cycle re-started by adding the resulting precipitated phages to another round of selection against the antigen. Three rounds of selection and amplification were performed for the present experiment, and an extra amplification of the third round was considered the fourth.

## 2.8. Characterization individual gliadin-binding phage-Fabs

Individual colonies isolated after each round of panning from the titrating output plates were grown at a small scale of 4 mL of initial culture, infected with helper phage, and induced for a quick screening of clones by analyzing the supernatants containing the phage-Fabs. The clones with better features were produced at a bigger scale of 100 mL and PEG-precipitated.

## 2.9. Indirect phage-ELISA

The specificity and affinity of the clones obtained were analyzed by indirect-phage ELISA. One hundred microliters of the suitable antigenic extracts were added per well, diluted in coating buffer (0.1 M carbonate/bicarbonate buffer, pH 9.6), and the plate was sealed and incubated for 1 h at 37 °C. The coating solution was shaken out and 200 µL blocking solution (3% BSA in PBS) was added per well, and the plate was incubated for 1 h at 37 °C. Following 10 washing steps, 100 µL of diluted phage was added and plates were incubated for 2 h at 37 °C. The non-binding phages were washed away 10 times with PBS. Then, 100 µL of the secondary antibody (anti-phage protein VIII HRP conjugated from Sinobiological®, Beijing, China, ref #11973) diluted 1:5000 in blocking buffer was added. The plate was incubated for 1 h at 37 °C, and PBS washed 10 times. One hundred microliters of TMB (Sigma®, ref #T0440) was added, and the reaction was stopped after 15–20 min with 50 µL of a diluted sulphuric acid solution. Absorbance readings were performed at 450 nm (FluoStar Optima™ from BMG labtech®). This methodology was applied for process evaluation, using the polyclonal mixture of phage-antibodies obtained from each panning round, and for characterization of monoclonal phage-Fabs.

## 2.10. DNA isolation and sequencing

Individual colonies from the titration plates of panning rounds three and four were picked up and grown overnight in 5 mL of LB medium supplemented with 100 µg/mL of carbenicillin. The plasmid DNA was isolated from the pellet using a mini-prep kit (GenElute™ plasmid miniprep kit from Sigma® ref #PLN70). The phagemids were sequenced by the Sanger method (Eurofins genomics®, Luxemburg, Luxemburg) using the primers ompAseq and g-back (Barbas, 2001) for light and heavy chain reading, respectively. The sequences were visualized, and preliminary analysis was done using Snapgene software (Dotmatrix®, Boston, MA, USA).

## 2.11. Searching and characterization of the gluten-binding Fabs

The following discovery pipeline was proposed, using indirect phage-ELISA for searching of strong gliadin binding candidates: 1) for the initial screening, 48 individual clones picked from panning outputs plates were grown, induced for production of phage-Fabs, and the supernatants were tested in a phage-ELISA against 2 µg of gliadin PWG per well; 2) the clones showing the highest signals against gliadin were sequenced to discard repeated events; 3) specificity of the clones was characterized (by their response to gluten-containing flours like wheat,

rye, and barley, and gluten-free flours like oats, corn, and rice), and their affinity against the objective antigen (gliadin-PWG) was analyzed in the range of 0-5  $\mu\text{g}/\text{mL}$ ; and 4) the top performance clones were also characterized for their response to the experimental binary mixture of wheat, rye, and barley kernels in a rice-based gluten-free matrix. Origin 8.0 software was employed to plot and analyze all the experimental data. The mean values of three independent determinations and standard derivation of each data set are shown in the figures.

### 2.12. Analysis of food products

The applicability of the leading phage-Fab candidate to detect gluten in complex matrixes, such as food samples, was analyzed by indirect-phage ELISA. Gluten-like proteins from ten commercial products were extracted in a two-steps procedure: 0.25 g of the sample was dissolved in 2.5 mL of extraction buffer SENSISpec INgezim (Ingenasa®, Madrid, Spain ref #30.GLU.K.2) and shaken for 20 min in a rotating shaker. Then, 7.5 mL of 80% ethanol/water solution was added, and the mixture was shaken for 20 additional minutes. The mixture was centrifuged at 2000 g, 10 min at room temperature. The analysis was performed by indirect-phage ELISA using the same protocol previously explained.

### 2.13. *In silico* elucidation of Fab-C structure and its interaction with the target antigen

*In silico* elucidation of the structure of the leading candidate (Fab-C) and the antibody-antigen interaction was done to have a deeper characterization of the antibody. The primary data input was the DNA sequence from Fab-C. From this sequence, the following pipeline was implemented: 1) the sequence was launched to IMGT/V-QUEST, an alignment-based tool for antibody domain definition and classification (Brochet et al., 2008); 2) the structure of the Fab was predicted using the bioinformatics tools Abodybuilder2 (Dunbar et al., 2016) and IgFold (Ruffolo et al., 2022); 3) concordances and discrepancies between both models were discerned using the Matchmaker tool from ChimeraX (Pettersen et al., 2021); 4) once the Fab models were obtained, the interactions with the target antigen (gliadin) were elucidated by computing epitope and paratope predictions with the tools Epipred (Krawczyk et al., 2014) and Antibody i-Patch (Krawczyk et al., 2013); and 5) based on the information about the antigen-antibody obtained from the previous analysis and, using both models of the Fab, a global simulation of the interaction with the antigen was calculated by docking, using HADDOCK 2.4 server (van Zundert et al., 2016).

## 3. Results and discussion

### 3.1. Antibody library design strategy

Several approaches have been proposed in the last decades to generate high-affinity antibodies against gluten. Most of them used classical strategies, like immunizing an animal with cereal or recombinant proteins and obtaining monoclonal antibodies secreted by hybridoma cells. Some recent approaches were based in developing recombinant antibodies with a variety of objectives. On one hand, immune libraries derived from patients aimed to profile their humoral response in celiac disease. On the other hand, camelid (Doña et al., 2010) or semisynthetic antibody libraries have been used as source of recombinant probes for detection of gluten (Garcia-Calvo et al., 2020).

A novel approach is proposed in this work, the construction of a new, fully human immune library from peripheral blood of celiac patients, with the objective of translating the humoral features found in the selected donors to build a platform for antibody discovery, with the final goal of developing accurate systems for gluten detection in food.

### 3.2. Clinical features and evaluation of the humoral response of selected patients

Building an immune library with enough diversity to allow isolation from of high-affinity Fabs against gluten proteins that are toxic to celiac patients required the capture of coding sequences for antibodies that are involved in the humoral response against the target antigen. To fulfill this requirement, potential suitable donors with high humoral response to several antigens related to celiac pathology that usually correlates with a high response to gliadin were searched at the digestive medicine service of the Fundación Jiménez-Díaz Hospital (Madrid). Initial parameters for selection of donors included anti-tTG and anti-DGP IgG and IgA titers. These serological markers are widely used because their presence is correlated with the intestinal damage, and their sensitivity and specificity in detecting untreated celiac disease are close to or above 95% (Leffler and Schuppan, 2010). The utilization of these markers, particularly circulating anti-tTG, for predictive purposes, encounters challenges in establishing a standardized and universally applicable cutoff point. The intricacy arises from the lack of harmonization among the diverse assay methods available, compounded by the preponderance of research focused on the pediatric population, as evidenced in many published works (Beltran et al., 2014). Consequently, in the selection process for this work, our clinical collaborators adhered to their hospital's clinical guidelines, which dictated the cut-offs adapted to their population's context.

In contemporary clinical practice, the detection of specific antibodies against complete gliadin in peripheral blood, has waned as a biomarker due to its diminished sensitivity in comparison to the currently employed biomarkers (Singh et al., 2022). Nonetheless, for the present work, this detection method assumed critical significance, given the role of gliadin as the primary antigen for the selection process. Thoroughly characterizing the humoral response of potential donors against gliadin was paramount, as it enabled the translation of these immune characteristics into the library.

Also, it was very important to assure that the patients had not started a gluten-free diet, because with the removal of gluten from their diet the inhibition of the abnormal celiac humoral response is very strong and sharp (Sharma et al., 2020).

Four patients were included in the study as potential donors (Table 1). Patient 2 was discarded because the levels of IgG Anti-tTG were low, and IgA anti-gliadin was bordering the levels for positive consideration. Patient 3 presented high humoral response levels and a clear celiac diagnosis but could not be included because started a gluten-free diet before the blood donation. Finally, patients 1 and 4 fulfilled all the requirements and donated blood samples.

In addition to these clinical features, a serum sample from the donors (patients 1 and 4) was tested to quantify their response against native gliadin by IgG isotype, with the aim to translate the heterogeneous antibody expression of the patients to the library, instead of inserting equimolar constructs of the possible isotypes and sub-isotypes (Fig. 1S).

Patients 1 and 4 presented a different IgG isotype response against gliadin, with the usual dominant expression of IgG1 in patient 1 (Nahm et al., 1987; Napodano et al., 2021). Nevertheless, patient 4 presented also a very high expression of IgG2 isotype, an unusual feature in celiac disease (Husby et al., 1986) but was considered appropriate for the library construction.

To transfer the affinity of the humoral response of celiac patients to gluten proteins, the Fab sequences included in the library were amplified from the mRNA extracted from lymphocytes, which encode the rearranged sequences from the recombination of the V, D, and J genes. In addition, obtaining the lymphocytes from peripheral blood assures the amplification of matured antibodies against the objective antigen (Victoria and Nussenzweig, 2022).

**Table 1**  
Analysis of the humoral response of potential celiac patients to be included in this study.

Patient	IgA Anti-tTG (U/mL)	IgG Anti-tTG (U/mL)	Selective IgA deficiency (negative/positive)	IgA Anti-DGP (U/mL)	IgG Anti-DGP (U/mL)	IgA Anti-native Gliadin (U/mL)	IgG Anti-native Gliadin (U/mL)
1	5654	<0.8	negative	1934	22.6	168	550
2	97.6	<0.8	negative	44.8	37.2	19	96
3	6994	49.6	negative	2165.1	537.6	35.37	200.70
4	3238	2	negative	1971	228.50	96.56	169.27
Cut off	>15	>15		>15	>15	>15	>50

### 3.3. Lymphocyte RNA preparation and retro-transcription

Mononuclear cells isolated from patients' blood were  $1.2 \times 10^8$  cells/mL from patient 1 and  $1.8 \times 10^8$  cells/mL from patient 4. Total RNA was isolated with TriZol LS reagent, and tested for integrity, showing high RIN (RNA integrity number) values of 9.5/10 for patient 1 and 8.75/10 for patient 4. The total RNA from each patient was retro-transcribed using general oligodT-20 primers, obtaining the cDNA scaffold for PCR amplification.

### 3.4. Antibody chain coding fragments amplification and library construction

Amplification of the fragments coding for the antibody chains was performed in two sequential rounds. Due to the amplification bias during the PCR, some chains can be overrepresented in comparison with those expressed in the donors. To avoid this issue, the same PCR conditions were used for all chains, so different quantities of each chain (gene family and isotype) were obtained for cloning instead of an equimolar pool of chains. Separate amplifications were done for each combination of family genes and isotypes (e.g., VH135-IgG1; VH4-IgA ...).

This approach is novel and unique, first because of the use of the Fab format while most of the previous works were ScFvs (Marzari et al., 2001; Rhyner et al., 2003), and also because of the amplification of as many VH genes as possible to generate a wide diversity, in contrast to previous studies that focused on certain VH families like VH4 (Sblattero et al., 2000) and VH5 (Not et al., 2011).

Fab Heavy chain amplicons (Fig. 2S A-B) were amplified with different terminal restriction sites than the light chains (Fig. 2S C-D) for sequential cloning. The second round of PCRs was only used to expand the neighboring sequences to the restriction point (addition of 39 bp) to improve the cloning steps. The amplification was performed with fewer cycles than the first step in order to avoid alteration of the first-round chain composition (Fig. 2S E).

This cloning strategy was used to avoid the more classical protocol of building the Fabs by OE-PCR (overlap-PCR) that implied three rounds of amplification (first of the V and C genes from cDNA, second an overlap of these genes, and third the fusion of the light and heavy chains by a second overlap) (Barbas, 2001). The methodology proposed for this work avoids one amplification step and overlaps by directly amplifying the light and heavy chains of the Fabs and cloning them into the plasmid. This strategy avoids PCR errors, misrepresentation of minority chains, and truncated overlaps (Omar and Lim, 2018).

The Fab library was built through a two-step cloning procedure; firstly, the light chains were introduced into the pComb3X vector and transformed into *E. coli* XL1-Blue cells. Plasmid DNA containing the light-chain sub-library was isolated from the culture for the subsequent cloning of the Fab-heavy chains, resulting in the Fab library that was transformed into *E. coli*. After the electroporation and recuperation in SOC media, repertoire size (RZ) of the library was calculated, being  $1.1 \times 10^7$  total clones. Some random clones were sequenced to assure that the Fab were correctly cloned, showing a self-ligation of 4%. The RZ >  $10^7$  and autoligation level (<5%) were good enough to continue with the panning protocol (Barbas, 2001).

### 3.5. Panning (selection and amplification) of gliadin binding phage-Fab

The Fab-library was screened against gliadin by phage-display. Three rounds of panning were conducted with increasing stringency by incrementing the number of washing steps. The output titers were measured by plating dilutions of the *E. coli* XL1-Blue culture infected with the eluted phages in SB-agar supplemented with 100 µg/mL carbenicillin (Fig. 1A). A 19-fold increase was observed in round two and then, it doubled in round three.

Although measuring the output titers was an adequate way to follow panning performance, partial characterization of the overall affinity of the selected, amplified, and isolated phage-Fabs from the library was measured by indirect polyclonal phage-ELISA (Fig. 1B). A three-fold signal increase was observed after round one, and the maximum signal of the test was obtained after round three.

Despite the insignificant increase in the transformants recovered in the first round, the polyclonal ELISA showed a clear increase in the overall affinity. On the contrary, the higher number of transformants recovered in the second round was not impaired with a noticeable increment of the affinity, which could be explained because a wide population of antibodies has been amplified with a mixture of affinities.

Focusing on the third round, a strong increase of the colonies recovered, and also in the signal of the ELISA, means that the high-affinity phage-Fabs against gliadin dominate in the population, and a monoclonal binder could be easily isolated from the pool.

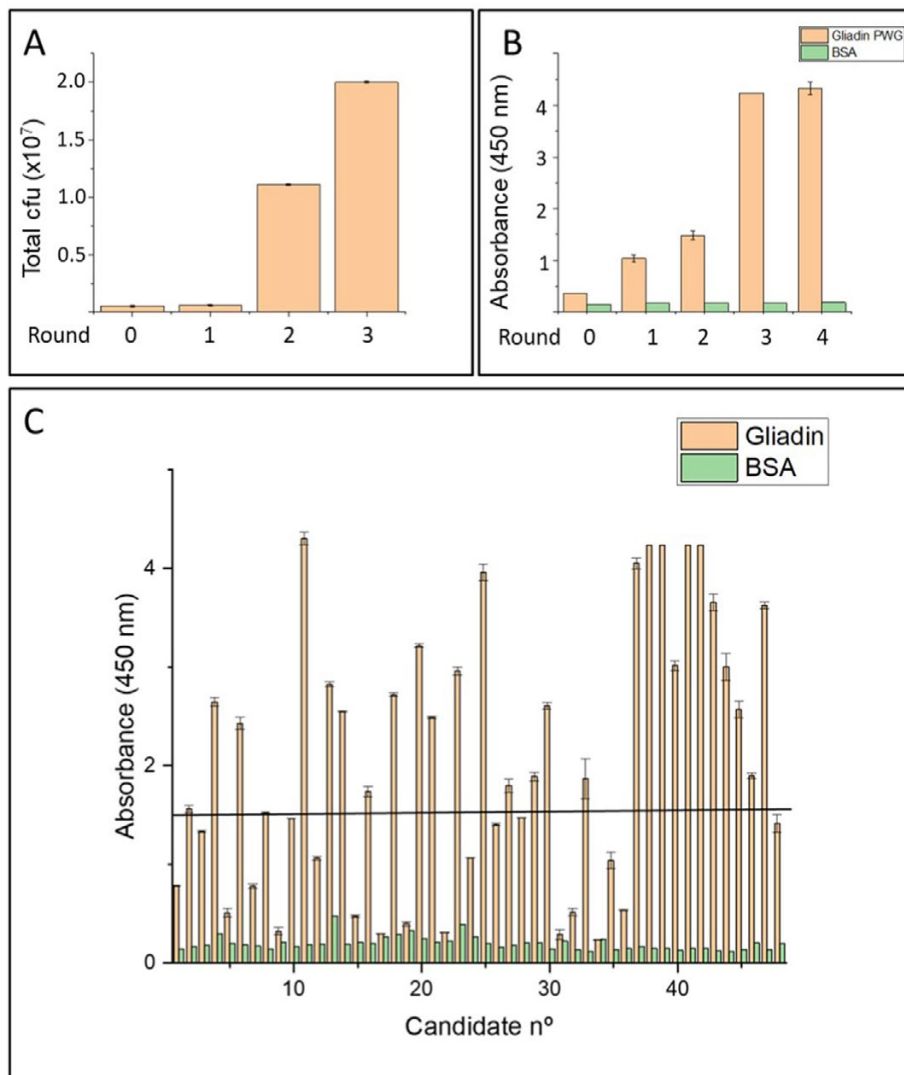
### 3.6. Selection of gliadin binders by monoclonal phage ELISA

Forty-eight individual clones were picked from the titrating plates of the output phage population from last panning rounds (36 from R3 and 12 from R4). Induced culture supernatants (non-precipitated but containing phage-Fabs) were 1:2 diluted and subjected to monoclonal phage-ELISA in immunoplates coated with gliadin-PWG at a concentration of 2 µg/well. Twenty-nine out of the forty-eight clones produced absorbance signals higher than 1.5 and did not bind the negative control (BSA). Those clones were selected as candidates for further analysis (Fig. 1C).

### 3.7. Sequencing and classification of the selected positive clones

The 29 clones identified in the monoclonal phage-ELISA as gliadin PWG binders, were sequenced to detect repetitive clones, to identify clonotypes and for their classification by gene family, class, and subclass. From the 29 sequenced clones, there were only 14 unique sequences, hereinafter named A-N (Fig. 2A).

One of the unique sequences (Fab-C) was heavily repeated in round four (with 9 clones (31%) presenting that sequence; Fab-H was found in 4 (14%) of the sequenced clones; Fab-N in 3 (10%); Fab-G and Fab-K two times each (6.89% each) and the remaining ones were found once (3.34% each). Two clones belonged to the IgA2 subclass (C and H), five to IgG1 (A, D, G, L and N), four to IgG2 (B, E, J and I), and three to IgA1 (F, K and M) for the heavy chain. Regarding the light chains, there was a total dominance of kappa chains. These unique sequences were organized in clonotypes (the clones belonging to the same clonotype presented a rearrangement with the same germlines V(D)J gene segments)



**Fig. 1.** Characterization of the panning progress for selection of gliadin-binding phage-Fabs. (A) Phage titration after each round of panning. (B) Polyclonal indirect phage-ELISA analysis of the phage-Fab library isolated after each round of panning. Notice that round 4 is referring to the amplification of output phages from round 3. (C) Monoclonal indirect phage-ELISA analysis of individual clones isolated from the third and fourth rounds of affinity enrichment against gliadin-PWG. Origin 8.0 software was used to plot and analyze the experimental data. Mean values of three independent determinations and standard derivation of each data set are shown.

using IgBLAST.

The 14 different clones were classified into 8 clonotypes for the heavy chains: clonotype 1H (B, E, I and J), clonotype 2H (F, K and M), clonotype 3H (A and G), clonotype 4H (C), clonotype 5H (H), clonotype 6H (D), clonotype 7H (L) and clonotype 8H (N). Three clonotypes (3H, 6H and 7H) presented VH1 genes, four VH3 (clonotypes 1H, 2H, 4H and 5H), and the remaining (clonotype 8H) VH4 genes for the heavy chains (Fig. 2B).

The variability of the V, D, J genes whose rearrangement produced the Fabs is shown in Fig. 2B, 2C and 2D. On one hand, the vast majority of the gliadin-reactive Fabs presented VH3 genes (64.3% of the clones) even surpassing the frequency found in a normal human antibody repertoire (approximately 50%).

Surprisingly, the most expressed family in human antibodies, VH3-23, (Sun et al., 2022), was only represented by one clonotype (2H), including by three Fabs.

A special remark must be noticed on the selection of Fab-C, that presents the VH3-15 locus, a scarcely expressed gene (3% of the human repertoire) but one of the most identified loci in gluten-reactive antibodies in celiac disease (Steinsbø et al., 2014). On the other hand, although VH4 antibodies represent almost 30% of the human repertoire, only one selected Fab presented this gene (Tiller et al., 2013).

The light chains could be classified into eight clonotypes: clonotype 1L (A, C and E); clonotype 2L (J and K); clonotype 3L (F and G);

clonotype 4L (B and M); clonotype 5L (H and I), clonotype 6L (D), clonotype 7L (L) and clonotype 8L (N) (Fig. 3G). Six of the clonotypes for the light chain presented Vk1 genes (clonotypes 1L, 3L, 4L, 5L, 6L and 7L), one Vk2 (clonotype 8L) and one Vk3 (clonotype 2L).

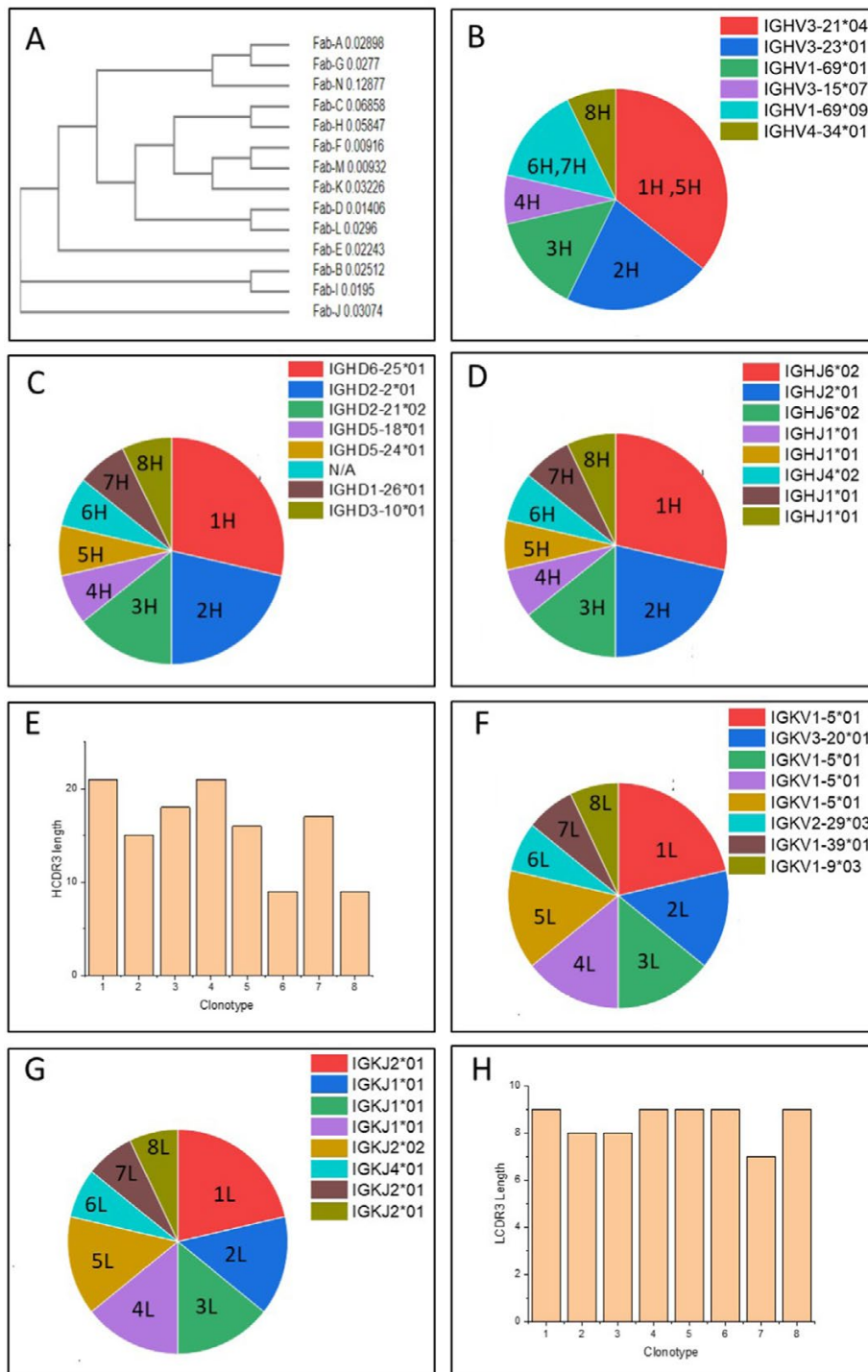
The frequency of the V and J genes whose rearrangement produced the light chain of the selected Fab are depicted in Fig. 2F and Fig. 2G. Most of the clonotypes in the selected Fabs presented Vk1 (78.6%), that is also prevalent in the human repertoire (approximately the 45% of kappa chains) (Tiller et al., 2013).

The length of HCDR3 of the selected Fabs was remarkably high ( $16 \pm 4.7$  amino acids), especially in two clonotypes (1H and 4H) that presented more than 20 amino acids, an unusual feature (Fig. 2E), when the average in human antibodies is only 14.8 amino acids (Joyce et al., 2020). A high variability was found in the length and composition of HCDR3, in contrast to the shorter and less diverse LCDR3 ( $8.5 \pm 0.76$ ) (Fig. 2H).

### 3.8. Characterization of gliadin-binding clones

Affinity and specificity of individual gliadin-binding clones were assessed by indirect-phage ELISA. The optimal phage-Fab concentration to be used as primary antibody in the phage-ELISA was  $8 \times 10^9$  phage particles per well.

A dose-response curve against increasing concentrations of gliadin-



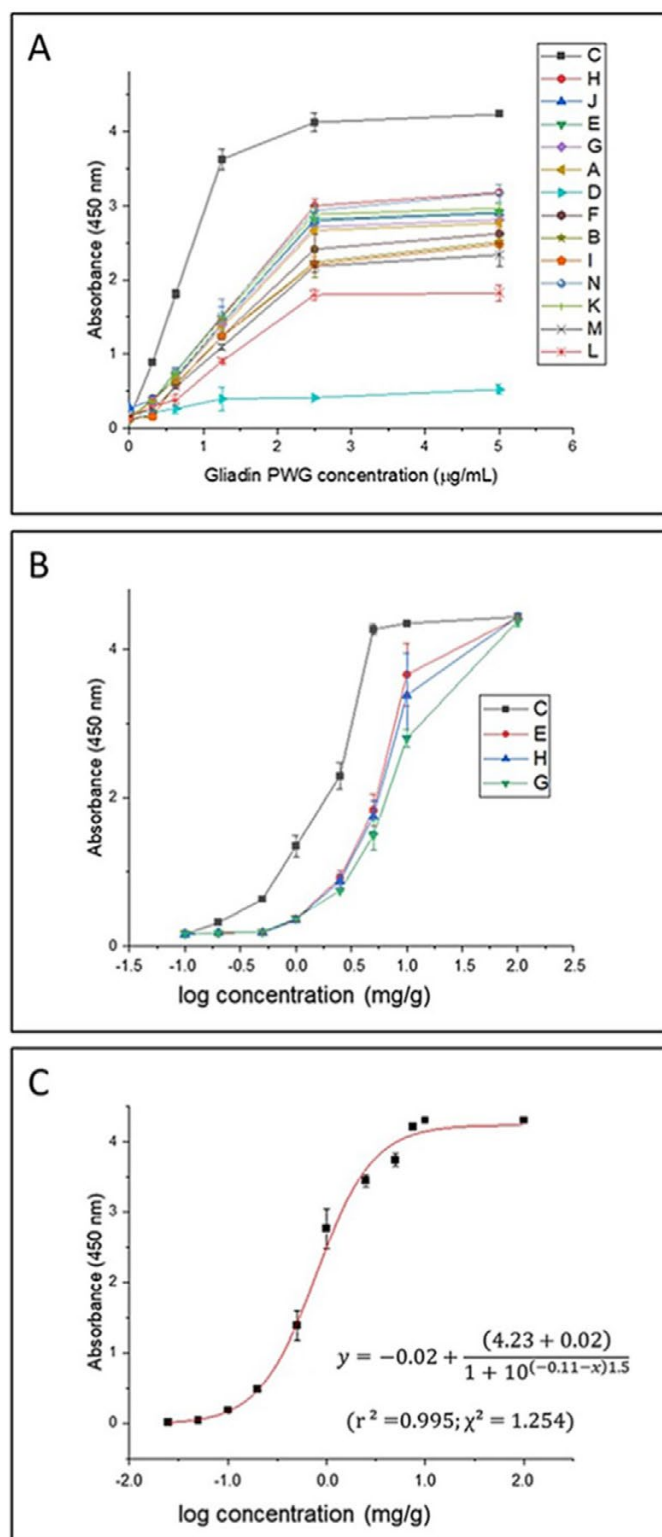
**Fig. 2.** Main features of the sequencing and classification of the selected Fabs. (A) Phylogenetic tree of isolated Fabs. An alignment of sequencing results was conducted with SnapGene® software, computing MUSCLE algorithm and then the phylogenetic tree was represented using the Simple Phylogeny tool (ClustalW2 package, EMBL-EBI). (B-H) Antibody sequences were analyzed, and clustering was calculated, using IgBLAST tool (NCBI) and Origin 8.0 software for data representation. The pie charts represent the frequency of the genes, and the depicted number refers for the clonotypes that presented each gene: (B) VH genes identification and frequency. (C) JH genes identification and frequency. (D) D genes identification and frequency. (E) Heavy chain CDR3 length distribution per clonotype. (F) VK genes identification and frequency. (G) JK genes identification and frequency. (H) Light chain CDR3 length distribution per clonotype.

PWG (0–5 µg/mL) was obtained for each phage variant (Fig. 3A), Fab-C stands out from the other Fabs, with a linear range between 0 and 1.25 µg/mL, reaching a saturation plateau at 2.5 µg/mL. In addition, a specificity test was conducted with different cereal flours to evaluate the performance of the selected clones in an actual food matrix, and to discriminate the cereal species recognized by the phage-Fabs. Fab-J, Fab-F and Fab-K showed cross-reactivity for oats; Fab-J, Fab-F and Fab-L for corn; Fab-J and Fab-F for rice extracts, these clones were not considered for further characterization because of their cross-reactivity with cereals tolerated by people suffering from GRDs.

The four strongest candidates (Fab-C, Fab-E, Fab-H and Fab-G), were tested against the gluten-like proteins extracted from a binary mixture

containing wheat, rye, and barley kernels (0; 0.1; 0.2; 0.5; 1; 2.5; 5; 10 and 100 mg/g) in a rice-based gluten-free matrix, equivalent to 0; 10; 20; 50; 100; 250, 500, 1000 and 10000 mg/kg of gluten (Fig. 3B) (García-García et al., 2020).

From the results obtained, the clone Fab-C clearly showed the best features: high response to gliadin-PWG in solution (purified antigen), and selectivity against gluten-containing flours (wheat, rye, and barley). Fab-C also showed a high performance for detection of gluten in the experimental binary mixture, with a LOD of 15 mg/kg of gluten (Fig. 3C), that fulfills the sorting capacity between gluten-free and gluten-containing foods set by current legislation (where gluten-free are considered when the product is containing less than 20 mg/kg of gluten)



**Fig. 3.** Characterization and selection of gliadin-binding Fabs by phage-ELISA. (A) Comparative dose-response curve of the fourteen leading candidates against gliadin-PWG dilutions (0–5 µg/mL). (B) Dose-response curve of the four strongest candidates (Fab-C, Fab-E, Fab-H and Fab-G) using an experimental rice-based binary mixture with increasing presence of a proportional mixture of wheat, barley and rye in the range of 0.1 to 100 mg/g. (C) Calibration curve obtained for gluten detection with the experimental mixture and the Fab-C in an indirect phage-ELISA. Origin 8.0 software was used to plot and analyze the experimental data. Mean values of three independent determinations and standard derivation of each data set are shown.

(Codex-Alimentarius, 2018).

Despite the existence of other antibody-based tests in the market with lower LOD for gluten detection (Scherf and Poms, 2016), Fab-C exhibits interesting attributes that warrant further exploration and advancement. Notably, Fab-C complies with the legal standards of detection, having a single paratopic interaction with the antigen, in contrast to monoclonal antibodies that typically require two interactions (which can be increased to four in a sandwich ELISA format). This distinctive feature of Fab-C results in highly efficient binding with gliadin, emphasizing its potential as a valuable candidate for gluten detection.

### 3.9. Analysis of commercial food products

Once the Fab-C was selected as the best gluten-binding probe, the next step was to demonstrate that it could be applied in an indirect-phage ELISA to detect gluten-like proteins extracted from market products, using the experimental binary mixture as standard curve (Fig. 3C).

Ten commercial products were analyzed, five of them were labeled as gluten-free (like nuts or plant based products) and other five declared to contain gluten (like bakery products) (Table 2). After calculating gluten concentration, the products were sorted in two categories according to the current legislation: negative or gluten-free (less than 20 mg/kg of gluten) and positive or gluten-containing product (with more than 20 mg/kg). All products tested were labeled according to the current legislation, and the results were confirmed by an R5 sandwich ELISA (the standard method for gluten detection in food) (Scherf and Poms, 2016).

These results confirm that phage-Fab-C can be used as the main reactant of an indirect phage-ELISA to differentiate gluten-free products.

### 3.10. In silico elucidation of the structure of Fab-C and its interaction with gliadin

An *in-silico* approach was considered the best way to study the molecular interaction between the Fab-C and gliadin, because there are no experimental data available of high-resolution structures (obtained by X-ray crystallography or cryo-microscopy) for either the antigen (gliadin) or the antibody (Fab-C).

In addition, Alpha-Fold is a high-accuracy tool for predicting the 3D structure of single-chain proteins (David et al., 2022).

The first step for modeling the interaction was to obtain their molecular structures. The structure of a wheat alpha/beta gliadin has been recently resolved with Alpha-Fold and uploaded to the UNIPROT database (IDS: P18573 or GDA9\_WHEAT, presenting 307 aminoacids and a molecular weight of 35 KDa). These proteins are considered intrinsically disordered (IDPs) (Markgren et al., 2020) and the *in silico* approach is, to the date, the best solution of their structure.

Characterization of the Fab structure faced several challenges. A first

**Table 2**

Results obtained for detection of gluten in ten commercial food products using the indirect phage-ELISA with Fab-C, and the sandwich ELISA with R5 monoclonal antibody as reference method.

Sample	Product type	Labeled as "gluten free"	R5	Phage-Fab-C
1	Vegan sausages	no	+	+
2	Vegan Deli	no	+	+
3	Surimi	no	+	+
4	Nuts	no	+	+
5	Gluten-free pasta	no	+	+
6	Macarrons	yes	-	-
7	Apple cake	yes	-	-
8	Free-sugar biscuits	yes	-	-
9	Caramel biscuits	yes	-	-
10	Sponge cake	yes	-	-

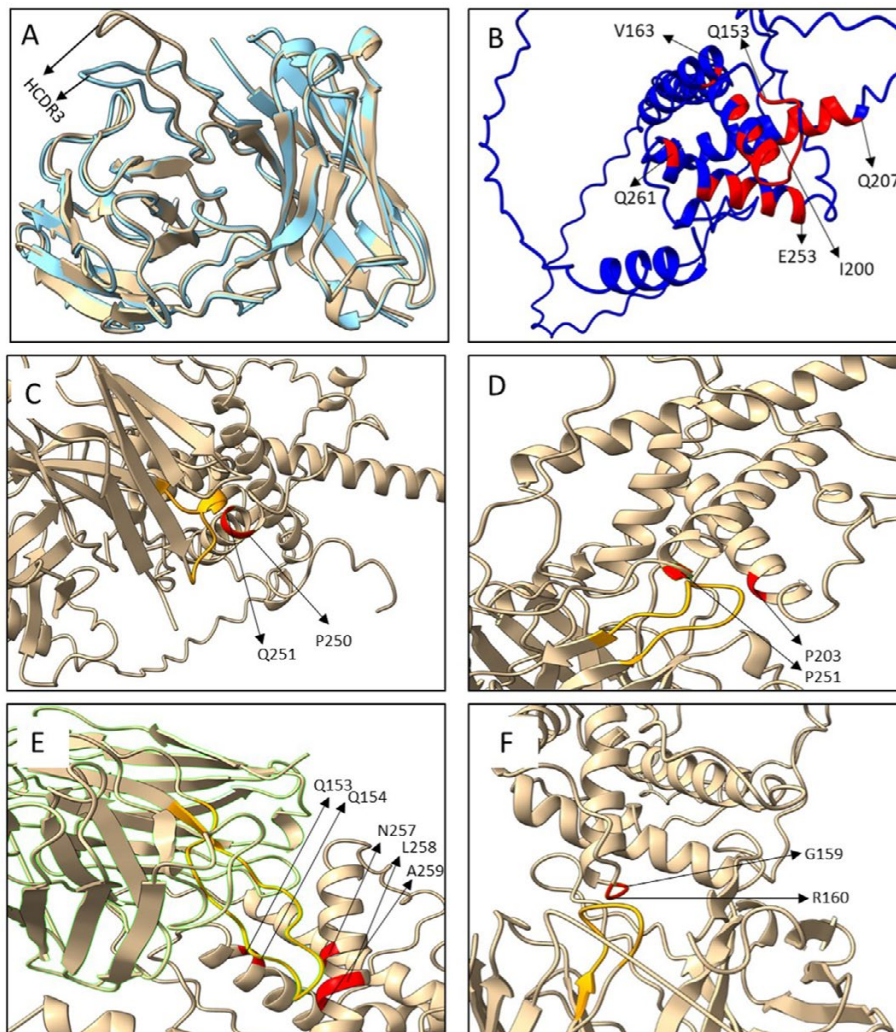
attempt to solve the structure using Alpha-fold program showed many inconsistencies, which could be caused because the Fab is a multi-chain protein. Then, alternative machine-learning antibody training tools were used, like IgFold and AbodyBuilder2. To compare both results, the structures were over-imposed with the ChimeraX alignment-based tool Matchmaker, visualizing a moderate discrepancy between the models regarding the spatial distribution of the HCDR3 (Fig. 4A). In an attempt to check which model could be closer to an actual experimental structure, molecules containing HCDR3 with similar length (20 to 23-mer) and amino acid sequences were searched using the SabDab database (Schneider et al., 2021), but they were not found. Thus, both structures were considered for further work.

To identify the epitopes recognized by the Fab-C, the zones of the gliadin molecule with higher probability to produce contacts between de Fab and the antigen were assessed to assure better docking. The EpiPred tool was used for the identification of contacts between each of the Fab-C models and the antigen, obtaining the same results regardless the Fab-C model used. The main contacts are more likely to occur on the C-terminus, with three “hot zones” of several amino acids that are more likely to be part of the epitope: Q153 to V163; I200 to Q207 and E253 to Q261 (Fig. 4B).

Once reliable structural models have been obtained (from a database for the antigen and generated *de novo* for the Fab) and a predicted epitope identified, a model of interaction by docking was computed using HADDOCK 2.4 tool, adjusting the contact zones to the CDRs for both Fab models and the predicted epitopes in the antigen.

Several docking models were computed, and the best are described by their quality scores: for the IgFold model (RMSD =  $0.6 \pm 0.3$  Å; Z-score = -1.2) and for the Abodybuilder2 model (RMSD =  $5.8 \pm 0.5$  Å; Z-score = -2) (Kufareva and Abagyan, 2011). Root Mean Square Deviation (RMSD) represents the average distance between atoms in superimposed protein structures, and is a crucial quality parameter for docking models. The model produced by IgFold demonstrates a high level of resolution, which proves valuable in both docking (where an RMSD <2 indicates successful docking), and even for crystallography modeling, despite being a computational model in this case (López-Camacho et al., 2016).

Based on these scores, the antigen-Fab interaction is better modeled using the IgFold Fab model. Thanks to this high-quality model, several likely interaction points between the antigen and the antibody were identified. As expected, the main paratopes are found in the heavy chain. The HCDR1 (Fig. 4C) interacts with proline and glutamine residues (position 250–251), and this PQ pattern has been identified in several peptides related to celiac toxicity (Morón et al., 2008). The HCDR2 (Fig. 4D) mainly interacts with glutamines in several positions within helices (positions 203, 207, 251 and 254). This could mean a strength of this Fab, as gluten-related proteins are very rich in this amino acid. The abnormal length of the HCDR3 (Fig. 4E) enables it to have double contact with the antigen. The proximal side of the HCDR3 loop (positions 109–114) could interact with 153–157 positions of the antigen (with a couple of glutamines) and the distal side (positions 105–109) with the residues asparagine 257, leucine 258 and glutamine 261 in a parallel helix regarding the previous one. Although the contribution of



**Fig. 4.** Computational model of the structure and interactions of Fab-C. (A) Fv structure prediction using two modeling programs: IgFold (blue) and SAbPred (brown). Comparison by superimposing the models obtained using the Matchmaker tool from ChimeraX. HCDR3 is flagged with a black arrow. (B) Representation of the epitope prediction of  $\alpha/\beta$ -gliadin. The residues colored in red are more likely to be part of the conformational epitope. Representation of the main likely interaction points between Fab-C and depicted  $\alpha/\beta$ -gliadin (C to F). The colored residues are part of the CDRs (yellow and flagged with a black arrow) or the antigen epitopes (red) in the docking model: (C) HCDR1; (D) HCDR2; (E) HCDR3; (F) LCDR3. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

the light chain to the interaction seems to be much weaker, there is a possible contact between the LCDR3 and the glycine 159 and arginine 160 (Fig. 4F), which probably help the overall stability of the union. Although this results are simulations, thanks to the quality of the docking procedure, it could be concluded that Fab-C is a good candidate as binder to the gliadin C-terminal domain, something unusual compared to other gluten detection antibodies, like R5 or G12 that bind in the N-terminal region. In that case, the antibody-antigen interactions were elucidated *in vitro*, using peptide arrays composed of overlapping 15-mer gluten peptide sequences (Röckendorf et al., 2017). The different binding profiles could be useful for development of immunoassays, as it is highly likely that Fab-C and other available antibodies could be combined in an immunoassay with enhanced features due to a multiple epitope detection.

#### 4. Conclusions

This work presents the first construction and application (by phage display) of a fully human immune library with the final goal of obtaining a novel tool for gluten detection in foods.

The new Fab library has been built from peripheral blood lymphocytes from celiac patients and, using phage-display technology, the Fab-C was selected as suitable probe for detection of gluten. The Fab-C is an antibody fragment with some uncommon features: it belongs to IgA2 subclass, which is directly correlated to celiac pathology; it also belongs to VH3-15 subfamily (not highly expressed in the human repertoire, but highly represented after the selection process), and it has a lengthy 21-mer HCDR3. The Fab-C has demonstrated to detect gluten in solution, and also in an experimental matrix and in food samples. A phage-ELISA functional test has been developed, with a LOD of 15 mg/kg, which can discriminate gluten-free products according to the current legislation.

A quick, simple, and universal pipeline, based in novel prediction tools, has been developed. Its application resulted in the generation of the structure of Fab-C that was suitable to be used to calculate the features of a proposed antibody-antigen interaction. These results showed that Fab-C is able to interact with the C-terminus of gliadin, another unexpected feature compared to other characterized antibodies (like R5 or G12).

In essence, this work introduces an innovative methodology for crafting antibody fragments tailored for application in food analysis systems. By harnessing the distinctive immune response features observed in patients with celiac disease, a comprehensive and original approach emerges, integrating principles from healthcare, molecular evolution, and food safety domains.

#### Funding

This work was funded by Ministerio de Ciencia e Innovación Grant: PID2021-122925OB-I00. PRE2018-08427 fellowship of the Ministerio de Ciencia e Innovación granted for Eduardo Garcia-Calvo.

#### Institutional Review Board statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Hospital Universitario Fundación Jiménez Díaz (protocol code October 2018 and date of approval, 24 September 2019, Madrid, Spain). Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Data Availability Statement: Data is contained within the article.

#### CRedit authorship contribution statement

**Eduardo Garcia-Calvo:** Methodology, Investigation, experimentation, Validation, Formal analysis, Writing – original draft, All authors have read and agreed to the published version of the manuscript. **Aina**

**García-García:** Investigation, Formal analysis, and writing and revision, All authors have read and agreed to the published version of the manuscript. **Santiago Rodríguez-Gómez:** Investigation, and, Formal analysis, All authors have read and agreed to the published version of the manuscript. **Sergio Farras:** celiac patient selection, All authors have read and agreed to the published version of the manuscript. **Rosario Martín:** Funding acquisition, and project management, Project administration, All authors have read and agreed to the published version of the manuscript. **Teresa García:** Supervision, writing and revision, Funding acquisition, and project management, Project administration, All authors have read and agreed to the published version of the manuscript.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

No data was used for the research described in the article.

#### Acknowledgments

We thank the Immunology and Cytogenetic Services of Fundación Jiménez Díaz Hospital (Madrid) for the management of the clinical samples used in this work.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.crfs.2023.100578>.

#### References

- Agardh, D., Lynch, K., Brundin, C., Ivarsson, S.A., Lernmark, Å., Cilio, C.M., 2006. Reduction of tissue transglutaminase autoantibody levels by gluten-free diet is associated with changes in subsets of peripheral blood lymphocytes in children with newly diagnosed coeliac disease. *Clin. Exp. Immunol.* 144 (1), 67–75. <https://doi.org/10.1111/j.1365-2249.2006.03036.x>.
- Andris-Widhopf, J., Rader, C., Steinberger, P., Fuller, R., Barbas Iii, C.F., 2000. Methods for the generation of chicken monoclonal antibody fragments by phage display. *J. Immunol. Methods* 242 (1–2), 159–181. [https://doi.org/10.1016/S0022-1759\(00\)00221-0](https://doi.org/10.1016/S0022-1759(00)00221-0).
- Barbas, C.F., 2001. *Phage Display : a Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- Beltran, L., Koenig, M., Egner, W., Howard, M., Butt, A., Austin, M.R., Patel, D., Sanderson, R.R., Goubet, S., Saleh, F., Lavender, J., Stainer, E., Tarzi, M.D., 2014. High-titre circulating tissue transglutaminase-2 antibodies predict small bowel villous atrophy, but decision cut-off limits must be locally validated. *Clin. Exp. Immunol.* 176 (2), 190–198. <https://doi.org/10.1111/cei.12249>.
- Bradbury, A.R.M., Trinklein, N.D., Thie, H., Wilkinson, I.C., Tandon, A.K., Anderson, S., Bladen, C.L., Jones, B., Aldred, S.F., Bestagno, M., Burrone, O., Maynard, J., Ferrara, F., Trimmer, J.S., Görmemann, J., Glanville, J., Wolf, P., Frenzel, A., Wong, J., Koh, X.Y., Eng, H.Y., Lane, D., Lefranc, M.P., Clark, M., Dübel, S., 2018. When monoclonal antibodies are not monospecific: hybridomas frequently express additional functional variable regions. *mAbs* 10 (4), 539–546. <https://doi.org/10.1080/19420862.2018.1445456>.
- Brochet, X., Lefranc, M.P., Giudicelli, V., 2008. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* 36, W503–W508. <https://doi.org/10.1093/nar/gkn316>.
- Caio, G., Volta, U., Sapone, A., Leffler, D.A., De Giorgio, R., Catassi, C., Fasano, A., 2019. Celiac disease: a comprehensive current review. *BMC Med.* 17 (1) <https://doi.org/10.1186/s12916-019-1380-z>.
- Codex-Alimentarius, 2018. Standard for Foods for Special Dietary Use for Persons Intolerant to Gluten. CSX, pp. 118–1979. <https://www.fao.org/fao-who-codexalimentarius/codex-texts/list-standards/es/>.
- David, A., Islam, S., Tankhilevich, E., Sternberg, M.J.E., 2022. The AlphaFold database of protein structures: a biologist's guide. *J. Mol. Biol.* 434 (2), 167336 <https://doi.org/10.1016/j.jmb.2021.167336>.
- Dunbar, J., Krawczyk, K., Leem, J., Marks, C., Nowak, J., Regep, C., Georges, G., Kelm, S., Popovic, B., Deane, C.M., 2016. SAbPred: a structure-based antibody prediction server. *Nucleic Acids Res.* 44 (W1), W474–W478. <https://doi.org/10.1093/nar/gkw361>.

- García-Calvo, E., García-García, A., Madrid, R., Martín, R., García, T., 2020. From polyclonal sera to recombinant antibodies: a review of immunological detection of gluten in foodstuff. *Foods* 10 (1), 66. <https://doi.org/10.3390/foods10010066>.
- García-Calvo, E., García-García, A., Rodríguez, S., Farras, S., Martín, R., García, T., 2022. Construction of a Fab library merging chains from semisynthetic and immune origin, suitable for developing new tools for gluten immunodetection in food. *Foods* 12 (1), 149. <https://doi.org/10.3390/foods12010149>.
- García-García, A., Madrid, R., García, T., Martín, R., González, I., 2018. Use of multiplex ligation-dependent probe amplification (MLPA) for screening of wheat, barley, rye and oats in foods. *Food Control* 84, 268–277. <https://doi.org/10.1016/j.foodcont.2017.07.037>.
- García-García, A., Madrid, R., González, I., García, T., Martín, R., 2020. A novel approach to produce phage single domain antibody fragments for the detection of gluten in foods. *Food Chem.* 321, 126685 <https://doi.org/10.1016/j.foodchem.2020.126685>.
- Husby, S., Foged, N., Oxelius, V.A., Svehag, S.E., 1986. Serum IgG subclass antibodies to gliadin and other dietary antigens in children with coeliac disease. *Clin. Exp. Immunol.* 64 (3), 526–535.
- Itzlinger, A., Branchi, F., Elli, L., Schumann, M., 2018. Gluten-free diet in celiac disease—forever and for all? *Nutrients* 10 (11), 1796. <https://doi.org/10.3390/nu10111796>.
- Joyce, C., Burton, D.R., Briney, B., 2020. Comparisons of the antibody repertoires of a humanized rodent and humans by high throughput sequencing. *Sci. Rep.* 10 (1) <https://doi.org/10.1038/s41598-020-57764-7>.
- Kang, T.H., Seong, B.L., 2020. Solubility, stability, and avidity of recombinant antibody fragments expressed in microorganisms. *Front. Microbiol.* 11 <https://doi.org/10.3389/fmicb.2020.01927>.
- Krawczyk, K., Baker, T., Shi, J., Deane, C.M., 2013. Antibody i-Patch prediction of the antibody binding site improves rigid local antibody–antigen docking. *Protein Eng. Des. Sel.* 26 (10), 621–629. <https://doi.org/10.1093/protein/gzt043>.
- Krawczyk, K., Liu, X., Baker, T., Shi, J., Deane, C.M., 2014. Improving B-cell epitope prediction and its application to global antibody–antigen docking. *Bioinformatics* 30 (16), 2288–2294. <https://doi.org/10.1093/bioinformatics/btu190>.
- Kufareva, I., Abagyan, R., 2011. Methods of protein structure comparison. *Methods Mol. Biol.* 857, 231–257. [https://doi.org/10.1007/978-1-61779-588-6\\_10](https://doi.org/10.1007/978-1-61779-588-6_10).
- Leffler, D.A., Schuppan, D., 2010. Update on serologic testing in celiac disease. *Am. J. Gastroenterol.* 105 (12), 2520–2524. <https://doi.org/10.1038/ajg.2010.276>.
- López-Camacho, E., García-Godoy, M.J., García-Nieto, J., Nebro, A.J., Aldana-Montes, J. F., 2016. In: A New Multi-Objective Approach for Molecular Docking Based on RMSD and Binding Energy. Springer International Publishing, pp. 65–77. [https://doi.org/10.1007/978-3-319-38827-4\\_6](https://doi.org/10.1007/978-3-319-38827-4_6).
- Markgren, J., Hedenqvist, M., Rasheed, F., Skepö, M., Johansson, E., 2020. Glutenin and gliadin, a piece in the puzzle of their structural properties in the cell described through monte carlo simulations. *Biomolecular* 10 (8), 1095. <https://doi.org/10.3390/biom10081095>.
- Marzari, R., Sblattero, D., Florian, F., Tongiorgi, E., Not, T., Tommasini, A., Ventura, A., Bradbury, A., 2001. Molecular dissection of the tissue transglutaminase autoantibody response in celiac disease. *J. Immunol.* 166 (6), 4170–4176. <https://doi.org/10.4049/jimmunol.166.6.4170>.
- Morón, B., Bethune, M.T., Comino, I., Manyani, H., Ferragud, M., López, M.C., Cebolla, A., Khosla, C., Sousa, C., 2008. Toward the assessment of food toxicity for celiac patients: characterization of monoclonal antibodies to a main immunogenic gluten peptide. *PLoS One* 3 (5), e2294. <https://doi.org/10.1371/journal.pone.0002294>.
- Nahm, M.H., Scott, M.G., Shackelford, P.G., 1987. Expression of human IgG subclasses. *Ann. Clin. Lab. Sci.* 17 (3), 183–196.
- Napodano, C., Marino, M., Stefanile, A., Pocino, K., Scatena, R., Gulli, F., Rapaccini, G.L., Delli Noci, S., Capozio, G., Rigante, D., Basile, U., 2021. Immunological role of IgG subclasses. *Immunology Investigations* 50 (4), 427–444. <https://doi.org/10.1080/08820139.2020.1775643>.
- Not, T., Ziberna, F., Vatta, S., Quaglia, S., Martellosi, S., Villanacci, V., Marzari, R., Florian, F., Vecchiet, M., Sulic, A.-M., Ferrara, F., Bradbury, A., Sblattero, D., Ventura, A., 2011. Cryptic genetic gluten intolerance revealed by intestinal antitransglutaminase antibodies and response to gluten-free diet. *Gut* 60, 1487–1493. <https://doi.org/10.1136/gut.2010.232900>.
- Omar, N., Lim, T.S., 2018. Construction of naive and immune human Fab phage-display library. *Methods Mol. Biol.* 1701, 25–44. [https://doi.org/10.1007/978-1-4939-7447-4\\_2](https://doi.org/10.1007/978-1-4939-7447-4_2).
- Peña, A.S., Rodrigo, L., 2015. Epidemiology of celiac disease and non-celiac gluten-related disorders. In: *Advances in the Understanding of Gluten Related Pathology and the Evolution of Gluten-free Foods*. Omniscience.
- Petersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J. H., Ferrin, T.E., 2021. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 30 (1), 70–82. <https://doi.org/10.1002/pro.3943>.
- Raiteri, A., Granito, A., Giamperoli, A., Catenaro, T., Negrini, G., Tovoli, F., 2022. Current guidelines for the management of celiac disease: a systematic review with comparative analysis. *World J. Gastroenterol.* 28 (1), 154–176. <https://doi.org/10.3748/wjg.v28.i1.154>.
- Rhyner, C., Weichel, M., Hübner, P., Achatz, G., Blaser, K., Cramer, R., 2003. Phage display of human antibodies from a patient suffering from celiac disease and selection of isotype-specific scFv against gliadin. *Immunology* 110 (2), 269–274. <https://doi.org/10.1046/j.1365-2567.2003.01728.x>.
- Röckendorf, N., Meckelein, B., Scherf, K.A., Schalk, K., Koehler, P., Frey, A., 2017. Identification of novel antibody-reactive detection sites for comprehensive gluten monitoring. *PLoS One* 12 (7), e0181566. <https://doi.org/10.1371/journal.pone.0181566>.
- Ruffolo, J.A., Chu, L.-S., Mahajan, S.P., Gray, J.J., 2022. Fast, Accurate Antibody Structure Prediction from Deep Learning on Massive Set of Natural Antibodies. Cold Spring Harbor Laboratory. <https://doi.org/10.1101/2022.04.20.488972>.
- Sblattero, D., Florian, F., Not, T., Ventura, A., Bradbury, A., Marzari, R., 2000. Analyzing the peripheral blood antibody repertoire of a celiac disease patient using phage antibody libraries. *Hum. Antibodies* 9, 199–205. <https://doi.org/10.3233/HAB-2000-9402>.
- Scherf, K.A., Poms, R.E., 2016. Recent developments in analytical methods for tracing gluten. *J. Cereal. Sci.* 67, 112–122. <https://doi.org/10.1016/j.jcs.2015.08.006>.
- Schneider, C., Raybould, M.I.J., Deane, C.M., 2021. SABDab in the age of biotherapeutics: updates including SABDab-nano, the nanobody structure tracker. *Nucleic Acids Res.* 50 (D1), D1368–D1372. <https://doi.org/10.1093/nar/gkab1050>.
- Sharma, N., Bhatia, S., Chunduri, V., Kaur, S., Sharma, S., Kapoor, P., Kumari, A., Garg, M., 2020. Pathogenesis of celiac disease and other gluten related disorders in wheat and strategies for mitigating them. *Front. Nutr.* 7 <https://doi.org/10.3389/fnut.2020.00006>.
- Singh, P., Singh, A.D., Ahuja, V., Makharia, G.K., 2022. Who to screen and how to screen for celiac disease. *World J. Gastroenterol.* 28 (32), 4493–4507. <https://doi.org/10.3748/wjg.v28.i32.4493>.
- Steinsbø, Ø., Dunand, C.J.H., Huang, M., Mesin, L., Salgado-Ferrer, M., Lundin, K.E.A., Jahnsen, J., Wilson, P.C., Sollid, L.M., 2014. Restricted VH/VL usage and limited mutations in gluten-specific IgA of celiac disease lesion plasma cells. *Nat. Commun.* 5 (1) <https://doi.org/10.1038/ncomms5041>.
- Sun, Z., Li, W., Mellors, J.W., Orentas, R., Dimitrov, D.S., 2022. Construction of a large size human immunoglobulin heavy chain variable (VH) domain library, isolation and characterization of novel human antibody VH domains targeting PD-L1 and CD22. *Front. Immunol.* 13, 869825 <https://doi.org/10.3389/fimmu.2022.869825>.
- Taraghihkah, N., Ashtari, S., Asri, N., Shahbazkhani, B., Al-Dulaimi, D., Rostami-Nejad, M., Rezaei-Tavirani, M., Razzaghi, M.R., Zali, M.R., 2020. An updated overview of spectrum of gluten-related disorders: clinical and diagnostic aspects. *BMC Gastroenterol.* 20 (1) <https://doi.org/10.1186/s12876-020-01390-0>.
- Tiller, T., Schuster, I., Deppe, D., Siegers, K., Strohner, R., Herrmann, T., Berenguer, M., Poujol, D., Stehle, J., Stark, Y., Heßling, M., Daubert, D., Felderer, K., Kaden, S., Kölln, J., Enzelberger, M., Urlinger, S., 2013. A fully synthetic human Fab antibody library based on fixed VH/VL framework pairings with favorable biophysical properties. *mAbs* 5 (3), 445–470. <https://doi.org/10.4161/mabs.24218>.
- Valdés, I., García, E., Llorente, M., Méndez, E., 2003. Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *Eur. J. Gastroenterol. Hepatol.* 15 (5), 465–477. <https://doi.org/10.1097/01.meg.0000059119.41030.df>.
- van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastrius, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., Bonvin, A.M.J.J., 2016. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* 428 (4), 720–725. <https://doi.org/10.1016/j.jmb.2015.09.014>.
- Victora, G.D., Nussenzweig, M.C., 2022. Germinal centers. *Annu. Rev. Immunol.* 40 (1), 413–442. <https://doi.org/10.1146/annurev-immunol-120419-022408>.

## Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods.

Eduardo GARCIA-CALVO<sup>1</sup>, Aina GARCÍA-GARCÍA<sup>\*1</sup>, Santiago RODRÍGUEZ GÓMEZ<sup>1</sup>, Sergio FARRAIS<sup>2</sup>, Rosario MARTÍN<sup>1</sup>, Teresa GARCÍA<sup>1</sup>.

<sup>1</sup>Departamento de Nutrición y Ciencia de los Alimentos, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040 Madrid, Spain

<sup>2</sup>Servicio de Medicina Digestiva, Hospital Universitario Fundación Jiménez Díaz, 28040 Madrid, Spain

‡ These authors share senior authorship

\* Correspondence: ainagarcia@ucm.es; Tel.: +34-91-3943749

### Supplementary Material

**Table 1S.** Primers for antibody chain amplification (Barbas, 2001).

#### First PCR round VH 5' primers

Primer	Sequence (5'→3')
VH135	AGG TGC AGC TGC TCG AGT CTG G
VH2	CAG RTC ACC TTG CTC GAG TCT GG
VH4	CAG GTG CAG CTG CTC GAG TCG GG
VH4b	CAG GTG CAG CTA CTC GAG TGG GG
VH4gs	CAG GTG CAG CTA CTC GAG TGG GGC
VH	CAG GTA CAG CTG CTC GAG TCA GG

#### CH1 3' primers

Primer	Sequence (5'→3')
CG1z (IgG1)	GCA TGT ACT AGT TTT GTC ACA AGA TTT GGG
CG2a (IgG2)	CTC GAC ACT AGT TTT GCG CTC AAC TGT CTT
CA1 (IgA)	AGT TGA ACT AGT TGG GCA GGG CAC AGT CAC

#### Vκ 5' primers

Primer	Sequence (5'→3')
Vκ1	GAG CCG CAC GAG CCC GAG CTC CAG ATG ACC CAG TCT CC
Vκ2/4	GAG CCG CAC GAG CCC GAG CTC GTG ATG ACY CAG TCT CC
Vκ3	GAG CCG CAC GAG CCC GAG CTC GTG WTG ACR CAG TCT CC
Vκ5	GAG CCG CAC GAG CCC GAG CTC ACA CTC ACG CAG TCT CC

**C $\kappa$  3' primer**

Primer	Sequence (5'→3')
CK1d	GCG CCG TCT AGA ATT AAC ACT CTC CCC TGT TGA AGC TCT TTG TGA CGG GCG AAC TCA G

**VL 5' primers**

Primer	Sequence (5'→3')
VL1a	GAG CCG CAC GAG CCC GAG CTC GTG TTG ACG CAG CCG CCC TC
VL1b	GAG CCG CAC GAG CCC GAG CTC GTG CTC ACT CAG CCA CCC TC
VL2	GAG CCG CAC GAG CCC GAG CTC GCC CTG ACT CAG CCT SCC TCC GT
VL3/9	GAG CCG CAC GAG CCC GAG CTC GWG CTG ACT CAG CCA CCY TC
VL7/8	GAG CCG CAC GAG CCC GAG CTC GTG GTG ACY CAG GAG CCM TC

**CL 3' primer**

Primer	Sequence (5'→3')
CL2d	GCG CCG TCT AGA ATT ATG AAC ATT CTG TAG G

**Second PCR round VH 5' primers**

Primers	Sequence (5'→3')
VH-ext	GTC GTC GTC GTC GTC GTC GTC GTA GGT GCA GCT GCT CGA GTC
VH4-ext	GTC GTC GTC GTC GTC GTC GTC GTA GGT GCA GCT ACT CGA GTG

**CH1 3' primers**

Primer	Sequence (5'→3')
CG1z-ext (IgG1)	GAG AGA GAG AGA GAG AGA GAG GCA TGT ACT AGT TTT GTC AC
CG2a-ext (IgG2)	GAG AGA GAG AGA GAG AGA GAC TCG ACA CTA GTT TTG CG
CA1-ext (IgA)	GAG AGA GAG AGA GAG AGA GAA GTT GAA CTA GTT GGG CAG GGC ACA GTC AC

**V $\kappa$  5' primer**

Primer	Sequence (5'→3')
V $\kappa$ -ext	GAG GAG GAG GAG GAG GAG CCG CAC GAG CCC GA

**C $\kappa$  3' primer**

Primer	Sequence (5'→3')
C $\kappa$ -ext	GAG GAG GAG GAG GAG GAG GCG CCG TCT AGA ATT AAC ACT CTC

## VL 5' primer

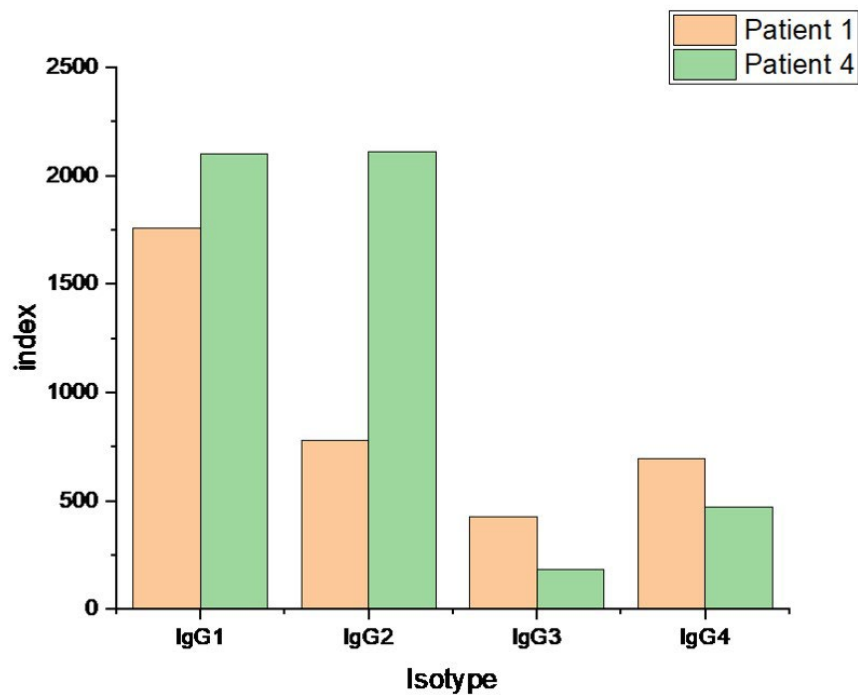
Same as V $\kappa$ -ext

Primer	Sequence (5'→3')
CL-ext	AGA GAG AGA GAG AGA GAG AGC GCC GTC TAG AAT TAT GAA CAT TCT GTA GG

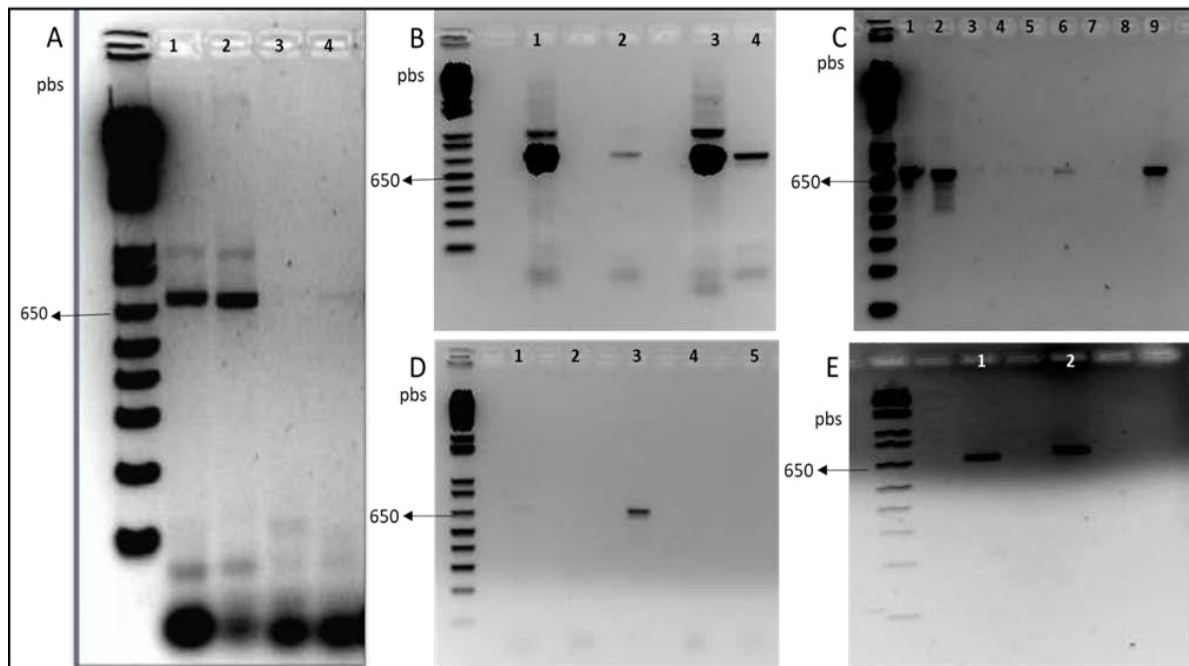
## Sequencing primers

Primer	Sequence (5'→3')
ompAseq	AAG ACA GCT ATC GCG ATT GCA G
g-back	GCC CCC TTA TTA GCG TTT GCC ATC

The degenerated bases stand for: R = A or G; Y = C or T; W = A or T; S = G or C; M = A or C



**Figure 1S.** Representation of the normalized indexes depicting the humoral response against native gliadin per IgG isotype.



**Figure 2S.** Agarose gel containing coding fragments for antibody chains to be cloned for library construction. Ladder: 1 kb plus Invitrogen™ (#ref 10787018). (A) Amplicons coding for VH1, 3, and 5 gene families for all IgGs. Lane 1: IgG1. Lane 2: IgG2. Lane 3: IgG3. Lane 4: IgG4. (B) Amplicons coding for IgA VH genes. Lane 1: VH135. Lane 2: VH2. Lane 3: VH4. Lane 4: VH4b. (C) Amplicons coding for kappa light chains. Lanes 1 and 2: Vk1. Lanes 3 and 4: Vk2/4. Lanes 5 and 6: Vk3. Lanes 7 and 8: Vk5. Lane 9: Vk pool. (D) Amplicons coding for lambda light chains: Lane 1: VL1a. Lane 2: VL1b. Lane 3: VL2. Lane 4: VL3/9. Lane 5: VL7/8. (E) Comparison between first round (Lane 1) and second round (Lane 2) PCR amplicons. The second round amplicon is slightly bigger (39 bps) than the first one, due to the expansion of the neighboring sequences to the restriction points.



## CAPÍTULO 4.3

### Producción y caracterización de nuevos Fab generados a partir de diferentes repertorios para *phage display* como sondas en inmunoensayos para detección de gluten en alimentos

---

### Production and characterization of novel Fabs generated from different phage display libraries as probes for immunoassays for gluten detection in food

Gluten is the main fraction of wheat proteins. It is widely used in the food industry because of the properties that are generated in the dough, but it is also able to trigger diseases like allergies, autoimmunity processes (such as celiac disease), and intolerances in sensitized persons. The most effective therapy for these diseases is the total avoidance of gluten in the diet because it not only prevents damage but also enhances tissue healing. To ensure the absence of gluten in food products labeled as gluten-free, accurate detection systems, like immunoassays, are required. In this work, four recombinant Fab antibody fragments, selected by phage display technology, were produced and tested for specificity and accuracy against gluten in experimental flour mixtures and commercial food products. A high-affinity probe (Fab-C) was identified and characterized. An indirect ELISA test was developed based on Fab-C that complied with the legal detection limits and could be applied in the assessment of gluten-free diets.

**García-Calvo, Eduardo; García-García, Aina; Rodríguez, Santiago; Takkinen, Kristiina; Martín, Rosario; & García, Teresa.** (2023). Production and Characterization of Novel Fabs Generated from Different Phage Display Libraries as Probes for Immunoassays for Gluten Detection in Food. *Foods*, 12(17).

<https://doi.org/10.3390/foods12173274>

Journal category: Food Science and Technology

Journal Impact Factor (2022): 5.2






Journal Rank: 34/142 (Q1)





Article

# Production and Characterization of Novel Fabs Generated from Different Phage Display Libraries as Probes for Immunoassays for Gluten Detection in Food

Eduardo Garcia-Calvo <sup>1</sup>, Aina García-García <sup>1,\*</sup>, Santiago Rodríguez <sup>1</sup>, Kristiina Takkinen <sup>2</sup>,  
Rosario Martín <sup>1,†</sup> and Teresa García <sup>1,†</sup>

<sup>1</sup> Departamento de Nutrición y Ciencia de los Alimentos, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040 Madrid, Spain; edugar01@ucm.es (E.G.-C.); santro03@ucm.es (S.R.); rmartins@ucm.es (R.M.); tgarcia@ucm.es (T.G.)

<sup>2</sup> Biosensors Team, VTT Technical Research Center of Finland Ltd., P.O. Box 1000, FI-02044 Espoo, Finland; kristiina.takkinen@vtt.fi

\* Correspondence: ainagarcia@ucm.es; Tel.: +34-91-394-3749

† These authors share senior authorship.

**Abstract:** Gluten is the main fraction of wheat proteins. It is widely used in the food industry because of the properties that are generated in the dough, but it is also able to trigger diseases like allergies, autoimmunity processes (such as celiac disease), and intolerances in sensitized persons. The most effective therapy for these diseases is the total avoidance of gluten in the diet because it not only prevents damage but also enhances tissue healing. To ensure the absence of gluten in food products labeled as gluten-free, accurate detection systems, like immunoassays, are required. In this work, four recombinant Fab antibody fragments, selected by phage display technology, were produced and tested for specificity and accuracy against gluten in experimental flour mixtures and commercial food products. A high-affinity probe (Fab-C) was identified and characterized. An indirect ELISA test was developed based on Fab-C that complied with the legal detection limits and could be applied in the assessment of gluten-free diets.

**Keywords:** gluten; recombinant Fab; celiac disease; phage display; ELISA



**Citation:** Garcia-Calvo, E.; García-García, A.; Rodríguez, S.; Takkinen, K.; Martín, R.; García, T. Production and Characterization of Novel Fabs Generated from Different Phage Display Libraries as Probes for Immunoassays for Gluten Detection in Food. *Foods* **2023**, *12*, 3274. <https://doi.org/10.3390/foods12173274>

Received: 3 August 2023

Revised: 23 August 2023

Accepted: 27 August 2023

Published: 31 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Grains are widely recognized as an essential component of a nutritious diet. However, the ingestion of gluten, which accounts for 80–90% of wheat proteins [1], has been linked to several adverse reactions that affect specific and growing population groups. Gluten-related diseases (GRDs) can be classified into three groups based on their etiology: allergic diseases, autoimmune diseases, and a third group that encompasses non-allergic and non-autoimmune gluten intolerance.

In the Western general population, GRDs that involve an allergic component have a prevalence rate of 0.1% [2]. This condition is characterized by a pathogenic immune response mediated by IgE antibodies. The second GRD group corresponds to autoimmune diseases [3]. Celiac disease is the most common of these disorders and is characterized by damage to the lining of the small intestine, which can cause various gastrointestinal and non-gastrointestinal symptoms [4]. The third group corresponds to non-celiac gluten intolerance, a disease that cannot be fit in the previous groups. It has an unknown molecular mechanism, but there is a strong implication of innate immunity [5].

Even though several pharmacological approaches have been developed to mitigate these diseases (especially celiac disease), none has demonstrated a better cost-efficiency than eliminating gluten from the diet [6]. Individuals who are affected by GRD often find it difficult to adhere to a gluten-free diet. Not only do they need to learn which products to avoid, but they also depend heavily on accurate product labeling [7]. The widespread

use of gluten in the food industry due to its functional properties adds complexity to this challenge [8]. Without an effective allergen management plan integrated into the Hazard Analysis Critical Control Points (HACCP) system, unintentional contamination of products during production can occur, leading to mislabeling and posing a threat to individuals who are sensitive to gluten.

In the European Union (EU), if a food product contains gluten, it must be clearly labeled as such on the packaging. The labeling must indicate the presence of gluten in the product and include the words “contains gluten” or “contains cereals containing gluten” in the ingredients list or in a separate allergen statement. Additionally, the gluten content of a food product must not exceed 20 mg/kg to be labeled as “gluten-free”, and 100 mg/kg to be labeled as “very low gluten” [9].

To fulfill these requirements, several analytical methodologies have been developed that can be classified into genomic, proteomic, and immunological methods [10]. Genomic methods are based on the detection of the specific genes that allow the identification of gluten-containing species. These genomic methods are widely based on real-time polymerase chain reactions (PCR) [11]. Proteomic methods are based on the detection and quantification of gluten proteins and their peptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), which allows reliable determination of protein levels as low as 0.01 mg/mL [12] and can be coupled to HPLC (high-performance liquid chromatography) systems [13] building up a highly reliable methodology. Despite the power of these techniques, they are expensive and applicable only for semi-quantitative measurements [14]. Immunodetection methods are based on antigen–antibody-specific interactions, and they are the most used for gluten detection in food due to their specificity, sensitivity, speed, and accessibility [15]. Among these immunodetection techniques, the application of ELISA (enzyme-linked immunosorbent assay) for gluten detection in food offers a highly sensitive and specific method to quantitatively detect even trace amounts of gluten proteins in food samples. This is essential for regulatory compliance and accurate labeling of gluten-free products. ELISA’s ability to differentiate between potentially harmful gluten components allows for comprehensive analysis, aiding in the identification of potential sources of gluten contamination. By providing a reliable means to assess gluten content, ELISA techniques contribute to safeguarding the health of individuals with gluten sensitivities, enabling them to make informed dietary choices and mitigating the risk of adverse reactions [16].

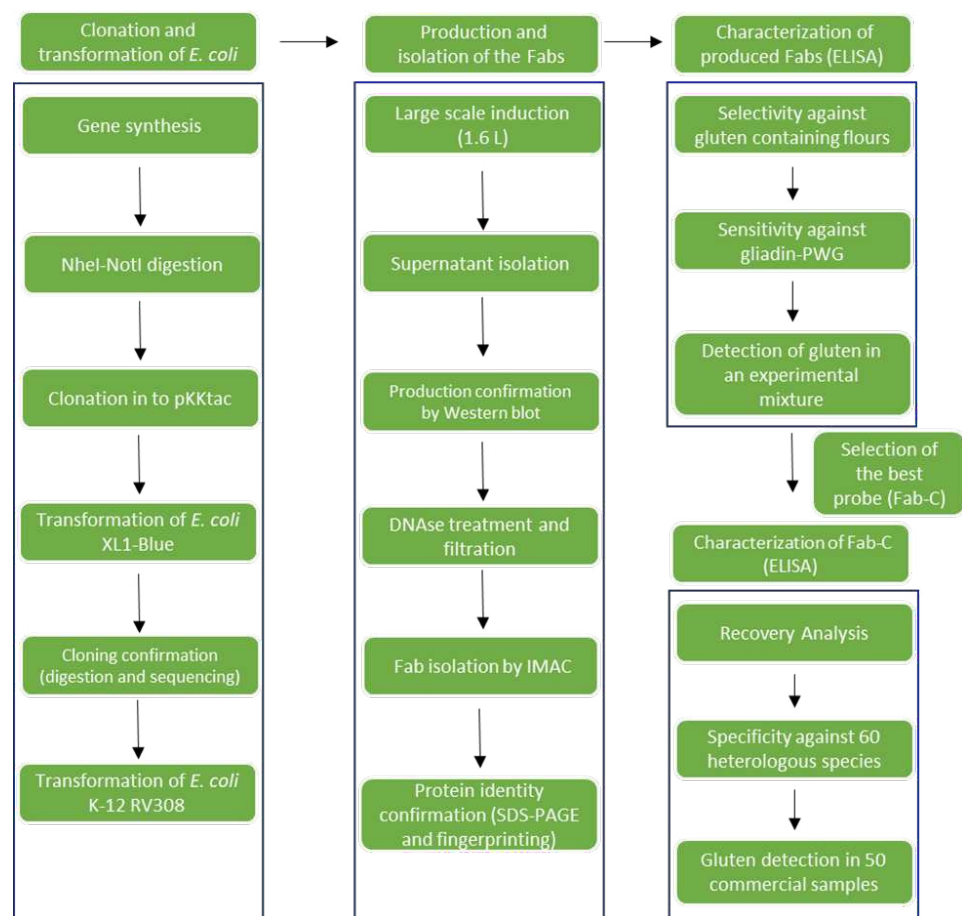
Immunodetection systems depend on high-affinity antibodies against gluten or its components. The most widely used is the R5 mouse monoclonal antibody [17], and a direct sandwich ELISA based on R5 is considered the gold standard technique for gluten detection in foodstuff [18].

Recombinant antibodies offer numerous advantages over both classical monoclonal and polyclonal antibodies. Firstly, recombinant antibodies are produced using *in vitro* techniques, avoiding the need for animal immunization, and can be generated in large quantities with consistent batches, ensuring a stable and scalable supply. They also exhibit reduced batch-to-batch variability, increasing the reproducibility and reliability of detection systems [19]. In addition, the recombinant production of antibodies avoids the presence of unwanted light chains generated by the hybridoma cells [20]. Overall, recombinant antibodies stand out as versatile and customizable tools that offer improved performance and sustainability.

Phage display is a vital technique for antibody discovery and engineering. It involves genetically modifying bacteriophages to display antibody fragments on their surface. By fusing genes encoding diverse antibody fragments with the gene codifying a phage coat protein, a library of phages with millions of different displayed antibodies is generated [21]. When exposed to a specific target antigen, phages with high binding affinity for the target are selected. These phages can be isolated, and the corresponding antibody fragments can be further developed into fully functional antibodies. This method enables the efficient

identification and isolation of antibodies that have the desired binding characteristics, making phage display a cornerstone in antibody development [22].

Novel single-domain (dAb) recombinant antibodies [23], suitable for the development of gluten detection ELISA systems, have been recently obtained from directed evolution processes based on phage display technology [24]. In this work, the use of Fabs is proposed instead of dAbs because of their enhanced stability and a bigger paratope [25]. The first step for this workflow was building Fab libraries, derived from two strategies: a Fab library merging antibody chains of different origins [26] and an immune library by cloning the genes expressed by peripheral blood lymphocytes from celiac patients [27]. Thanks to the phage display technology, four Fabs were identified from affinity selection processes. The aim of this work was the production of the selected Fabs in a soluble format and studying their feasibility as immunoassay probes for gluten detection in food (Figure 1).



**Figure 1.** Schematic workflow of the production and characterization of novel Fab as probes in ELISA methods for gluten detection in foodstuff.

## 2. Materials and Methods

### 2.1. Bacterial Strains and Growth Media

*Escherichia coli* XL1-Blue strain (*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, *lac* (F, *proAB*, *lacIqZΔM15*, *Tn10*, (Tetr))) (Agilent©, Santa Clara, CA, USA, ref #200150) was used for subcloning. *E. coli* K-12 strain RV308 substrain ATCC 31608 (*lacIq*, *su-*, *ΔlacX74*, *gal*, IS II::OP308, *strA*, (DE3)) was used for recombinant Fab production. Luria Broth (LB: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7) was used for growth before the preparation of chemical competent cells. The transformed cells were grown

in Super-Broth medium (SB: 30 g/L tryptone, 20 g/L yeast extract, 10 g/L MOPS, pH 7) supplemented with 1% glucose for DNA extraction. The RV308 cells were grown in Terrific Broth medium (TB: 12 g/L tryptone, 24 g/L yeast extract, 4 g/L glycerol, and phosphate buffer (0.17 M  $\text{KH}_2\text{PO}_4$ , 0.72 M  $\text{K}_2\text{HPO}_4$ , pH 7.2)) for recombinant protein production.

### 2.2. Recombinant Antibodies Selection and Bacterial Transformation

Four gliadin-binding Fab recombinant antibodies (named Fab-C, Fab-H, Fab-E, and Fab8E-4) were previously selected using phage display from Fab libraries cloned into the phagemid pComb3X [26]. The genes coding for the four Fabs were synthesized, changing their original antibody isotype to IgG1 and flanked by restriction sites (*NheI* and *NotI*) for cloning (Eurofins genomics©, Luxemburg, Luxemburg).

The synthetic genes were cloned into the pKktac expression vector [28] by digestion of the plasmids and inserts with *NheI* and *NotI* restriction enzymes (New England Biolabs© Ipswich, MA, USA ref #R3131 and ref #R3189), purification by gel electrophoresis, and ligation with T4 ligase (Promega©, Madison, WI, USA ref #M1801) overnight at 16 °C.

The resulting ligations were transformed by chemical transformation into *E. coli* XL1-Blue for storage, and into *E. coli* RV308 for protein production.

Chemical-competent cells were in-house produced by growing in 3 mL of SB for 3 h at 37 °C on a shaker (220 rpm). Cells were centrifuged at  $4700 \times g$  for 2 min at room temperature and gently resuspended into 100  $\mu\text{L}$  of ice-cold sterile  $\text{CaCl}_2$  (1 M). The cells were incubated on ice for 30 min. One hundred nanograms of plasmid DNA were added to the chemically competent cells, which were then incubated for 30 min on ice, followed by a 2 min heat shock (37 °C), and incubated again on ice for 2 min. The bacterial cells were recovered by adding 400  $\mu\text{L}$  of SOC medium (Invitrogen™-Thermo Fisher©, Waltham, MA, USA, ref #15544-034) and were grown for 1 h at 37 °C, with shaking at 220 rpm. Then, the culture was plated on LB-ampicillin (100  $\mu\text{g}/\text{mL}$ ) medium and incubated at 37 °C overnight.

Plasmid DNA was isolated from several transformed *E. coli* RV308 colonies using NucleoSpin® Mini kit for plasmid DNA (Machery-Nagel©, Allentown, PA, USA, ref # 740588.50). Plasmids were digested with *NheI* and *NotI* and analyzed by agarose gel electrophoresis (Thermo-fisher©, ref #16520050) and by Sanger sequencing to ensure that the clonation process was correct.

### 2.3. Production of Recombinant Antibodies

A single colony of *E. coli* RV308, transformed with the plasmid codifying the Fab of interest, was inoculated into 50 mL of TB medium supplemented with 1% glucose and 100  $\mu\text{g}/\text{mL}$  ampicillin and incubated overnight at 37 °C with shaking at 220 rpm. Then, this culture was expanded to 1.8 L of TB medium supplemented with 100  $\mu\text{g}/\text{mL}$  ampicillin, and divided into 6 flasks that were shaken (220 rpm) at 37 °C until the  $\text{OD}_{600}$  reached 4. The recombinant production was induced using 1 mM of IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside), lowering the temperature to 30 °C, and shaking at 170 rpm overnight. Finally, the recombinant Fabs were recovered from the overnight culture by centrifugation at  $4700 \times g$  for 15 min.

Production of the recombinant Fabs was assessed by Western blot analysis. Supernatant proteins were separated by electrophoresis in a 15% polyacrylamide gel at 100 V, using Biorad dual color as the molecular weight marker (BioRad, Hercules, CA, USA ref #1610364). The separated proteins were transferred to a PVDF (Polyvinylidene Difluoride) membrane at 100 V for 1 h. The PVDF membrane was blocked with TBST (50 mM Tris HCl, 150 mM NaCl, Tween20 0.05%, pH 7.5) for 30 min at 37 °C. Goat anti-human IgG F(ab)<sub>2</sub> alkaline phosphatase conjugate (Rockland©, Philadelphia, PA, USA ref #709-1518) diluted 1:1000 in TBST was added and incubated for 1 h at room temperature. The membrane was rinsed 3 times with TBST before the addition of 5-bromo-4-chloro-3-indolyl-phosphate and Nitroblue-Tetrazolium (BCIP/NBT) to visualize the bands containing Fab.

The supernatants containing the Fabs were treated with DNase I (Merck© Darmstadt, Germany ref #11284932001) and filtered through Whatman® glass filters GF/B and GF/C (Merck© ref #182110 and ref #182202) in a suction bottle to reduce viscosity. Then, the supernatant was nine-fold concentrated, and the buffer was exchanged for PBS-imidazole (137 mM NaCl, 2.7 KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, 5 mM imidazole, pH 7.4) with a prep/scale tangential flow filtration (TFF) 1 ft<sup>2</sup> cartridge (Merck©, ref #CDUF001LG). The concentrated supernatant was centrifuged at 15,000 × g for 10 min at room temperature and filtered through Whatman® glass filters GF/F and GF/D (Merck© ref #1825090 and ref #182390) in a suction bottle.

#### 2.4. IMAC-Purification and Analysis

The recombinant Fabs were purified using the NGC discovery chromatography system (BioRad, ref #7880009) and a HisTrap FF crude 5 mL affinity column (Cytiva, Marlborough, MA, USA, ref #11-0004-58) charged with Nickel ions from a NiSO<sub>4</sub> 100 mM solution. Two buffers were used for the purification: a binding buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>, 500 mM NaCl, 5 mM imidazole, pH 7.4) and an elution buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>, 500 mM NaCl, 500 mM imidazole, pH 7.4).

Once the column was equilibrated, the treated supernatant was fed to the system with a 5 mL/min flow rate. Then, the unbound proteins were flushed away by a washing program (5 mL/min flow rate and 5 column volumes of binding buffer). The proteins bound to the column were eluted by a gradient of elution (2 mL/min flow rate from 5 to 500 mM of imidazole in 75 min). A final wash with elution buffer was done to strip off every remaining protein.

The process was followed by spectrophotometry, measuring the output flow at 280 nm, and samples from each fraction of interest were analyzed by SDS-PAGE (sodium dodecyl-sulfate polyacrylamide gel electrophoresis). The fractions that contained the recombinant Fabs were dialyzed against 100 volumes of PBS (137 mM NaCl, 2.7 KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) overnight at 4 °C, for imidazole elimination.

To confirm the presence of Fabs, the bands from the SDS-PAGE, which matched with the Fab size, were analyzed by mass fingerprint. The corresponding bands were excised using a sterile scalpel. Following the methodology described by Sechi and Chait [29], the samples underwent in-gel reduction, alkylation, and trypsin digestion. After overnight trypsinization, 1 µL of the resulting supernatant was allowed to dry and then spotted onto a MALDI plate. The plate was prepared using α-cyano-4-hydroxy-cinnamic acid matrix (Sigma) in 50% acetonitrile. Subsequently, peptide analysis was conducted at the Proteomics Unit of Complutense University of Madrid (Spain) using a 4800 Plus Proteomics Analyzer MALDI-TOF/TOF instrument (Applied Biosystems, MDS Sciex, Toronto, ON, Canada). Protein identification via peptide mass fingerprinting (PMF) was performed using the MASCOT v2.6.2 search engine through Global Protein Server (GPS) v.3.6 (ABSCIEX, Toronto, ON, Canada), employing the following search parameters: carbamidomethylcysteine as a fixed modification, oxidized methionine as a variable modification, allowing one missed trypsin cleavage site, and with peptide mass tolerance set to 80 mg/kg and MS/MS fragment tolerance to 0.3 Da.

#### 2.5. Reference Materials and Food Samples Used for Analysis

Gliadin-PWG (prolamin working group) was used as reference material. Gliadin-PWG was obtained from the ethanolic extraction of a mixture of 28 European wheat cultivars after the elimination of albumins and globulins using a 0.4 M NaCl solution and then were concentrated, desalted by ultrafiltration, freeze-dried, and homogenized [30].

To test the ability of soluble Fabs obtained to detect gluten in complex matrixes, an experimental mixture of gluten-free rice flour spiked with flour made of kernels of gluten-containing cereals was prepared, as previously described [31]. The kernels, kindly provided by the national seed repository (Instituto Nacional de Investigaciones Agrarias, INIA, Spain), were ten different cultivars of each of the following cereals: common wheat

(*Triticum aestivum vulgare*), spelt wheat (*Triticum aestivum spelta*), rivet wheat (*Triticum turgidum turgidum*), durum wheat (*Triticum turgidum durum*), barley (*Hordeum vulgare*), rye (*Secale cereale*), and two cultivars of the hybrid crop triticale ( $\times$  *Triticosecale*). All the mixtures were ground and mixed in an IKA A11 analytical mill (IKA<sup>®</sup>, Staufen, Germany).

To test specificity for gluten, 60 heterologous species not containing gluten were tested to ensure there was no cross-reactivity (Table 1).

**Table 1.** List of heterologous species (common name and scientific name) that did not show cross-reactivity when tested in the indirect ELISA with Fab-C.

Okra ( <i>Abelmoschus esculentus</i> )	Lentil ( <i>Lens culinaris</i> )
Button mushroom ( <i>Agaricus bisporus</i> )	Cassava ( <i>Manihot esculenta</i> )
Onion ( <i>Allium cepa</i> )	Mango ( <i>Mangifera indica</i> )
Leek ( <i>Allium ampeloprasum</i> )	Red banana ( <i>Musa acuminata</i> )
Cashew ( <i>Anacardium occidentale</i> )	Myrtle ( <i>Myrtus communis</i> )
Peanut ( <i>Arachis hypogaea</i> )	Olive ( <i>Olea europaea</i> )
Strawberry tree ( <i>Arbutus unedo</i> )	Rice ( <i>Oryza sativa</i> )
Beetroot ( <i>Beta vulgaris</i> )	Passion fruit ( <i>Passiflora edulis</i> )
Cabbage ( <i>Brassica oleracea</i> )	Pepperomia ( <i>Peperomia pellucida</i> )
Chinese cabbage ( <i>Brassica rapa subsp. pekinensis</i> )	Avocado ( <i>Persea americana</i> )
Pigeon pea ( <i>Cajanus cajan</i> )	Common bean ( <i>Phaseolus vulgaris</i> )
Chili pepper ( <i>Capsicum annuum</i> )	Runner bean ( <i>Phaseolus coccineus</i> )
Tabasco pepper ( <i>Capsicum frutescens</i> )	Cape gooseberry ( <i>Physalis peruviana</i> )
Scotch bonnet pepper ( <i>Capsicum chinense</i> )	Pea ( <i>Pisum sativum</i> )
Papaya ( <i>Carica papaya</i> )	Almond ( <i>Prunus dulcis</i> )
Chickpea ( <i>Cicer arietinum</i> )	Sweet almond ( <i>Prunus dulcis var. dulcis</i> )
Lemon ( <i>Citrus limon</i> )	Plum ( <i>Prunus domestica</i> )
Mandarin orange ( <i>Citrus reticulata</i> )	Raspberry ( <i>Rubus idaeus</i> )
Watermelon ( <i>Citrullus lanatus</i> )	Atlantic salmon ( <i>Salmo salar</i> )
Cantaloupe ( <i>Cucumis melo var. cantalupensis</i> )	Castor bean ( <i>Ricinus communis</i> )
Cucumber ( <i>Cucumis sativus</i> )	Blackberry ( <i>Rubus fruticosus</i> )
Quince ( <i>Cydonia oblonga</i> )	Sesame ( <i>Sesamum indicum</i> )
Persimmon ( <i>Diospyros kaki</i> )	Tomato ( <i>Solanum lycopersicum</i> )
Teff ( <i>Eragrostis tef</i> )	Cherry tomato ( <i>S. lycopersicum cerasiforme</i> )
Arugula ( <i>Eruca vesicaria subsp. sativa</i> )	Spinach ( <i>Spinacia oleracea</i> )
Strawberry ( <i>Fragaria x ananassa</i> )	Vanilla ( <i>Vanilla planifolia</i> )
European strawberry ( <i>Fragaria vesca</i> )	Adzuki bean ( <i>Vigna angularis</i> )
Soybean ( <i>Glycine max</i> )	Mung bean ( <i>Vigna radiata</i> )
Sunflower ( <i>Helianthus annuus</i> )	Grape ( <i>Vitis vinifera</i> )
Walnut ( <i>Juglans regia</i> )	Ginger ( <i>Zingiber officinale</i> )

In addition, 50 commercial food products were purchased from several local stores (Spain). According to their labels, they were classified into 9 gluten-containing products, 6 not declaring gluten, 21 declaring that they could contain gluten, and 14 labeled as gluten-free.

### 2.6. Extraction of Gluten from Samples

Firstly, 50 g of each tested sample was individually ground in the analytical mill, ensuring thorough cleaning between samples, and stored at  $-20\text{ }^{\circ}\text{C}$ . The gluten-like proteins were extracted from 250 mg of finely ground samples (either experimental mixtures, spiked samples, or food products) using 2.5 mL of the Ingezim Gluten extraction solution (Ingenasa©, Madrid, Spain). The mixture was homogenized, incubated at  $50\text{ }^{\circ}\text{C}$  for 40 min, and cooled to room temperature before addition of 7.5 mL of 80% ethanol/water solution and shaking for 1 h in a vertical rotator. The extract was centrifuged at  $2000\times g$  for 10 min at room temperature, and the supernatant containing gluten proteins was transferred to a glass vial and stored in darkness at  $23\text{ }^{\circ}\text{C}$  until used.

### 2.7. Indirect Enzyme-Linked Immunosorbent Assay (ELISA) Method Based on the Recombinant Fabs

One hundred microliters of the ethanolic sample extracts obtained as described above (diluted 1/5 in PBS), or a solution of the appropriate gliadin-PWG concentration, was added per well. The immune-sorbent plate (Thermo©, ref #163320) was sealed and incubated for 1 h at  $37\text{ }^{\circ}\text{C}$ . The coating solution was shaken out, 200  $\mu\text{L}$  blocking solution (3% BSA in PBS) was added per well, and the plate was incubated for 1 h at room temperature. Following 10 washing steps with PBS, 0.5  $\mu\text{g}$  of recombinant Fab was added per well, diluted in 100  $\mu\text{L}$  of blocking solution, and incubated for 1 h at room temperature in a plate shaker. The plate was washed 10 times with PBS, and the secondary antibody (rabbit anti-human H + L HRP conjugated, Abcam©, Cambridge, UK ref #6759), diluted in 100  $\mu\text{L}$  of blocking solution was added, and incubated for 1 h at room temperature in a plate shaker. After washing 10 times with PBS, 100  $\mu\text{L}$  of TMB (Sigma©, ref #T0440) was added. The reaction was stopped after 20 min with 50  $\mu\text{L}$  of a diluted sulphuric acid solution, and the signal detected at 450 nm.

### 2.8. Assay Validation

The mean values of three independent determinations and standard derivation of each data set are shown in the figures. The data were plotted, and fitting models were determined using Origin 8.0 software (OriginLab Corp., Wellesley Hills, MA, USA). The limits of detection (LOD) and quantification (LOQ) were calculated as three and ten times the standard deviation of ten blank replicates, respectively [32].

The recovery of gliadin from food samples was analyzed by using a gluten-free certified rice flour spiked with solid gliadin-PWG equivalent to theoretical 20 and 100 mg/kg of gluten. As a control, a rice-flour ethanol extract was spiked with the equivalent quantities of ethanol-dissolved gliadin-PWG (0.05  $\mu\text{g}/\text{mL}$  and 0.25  $\mu\text{g}/\text{mL}$ ).

ELISA results obtained from 50 commercial food products were compared with those obtained with R5 monoclonal antibody-based sandwich ELISA (Ingenasa©, ref #30.GL2.K.2), which has a LOQ of 3 mg/kg of gluten and has been approved as a Type I method by *Codex Alimentarius*, following the manufacturer protocol.

## 3. Results and Discussion

The phage display technology has witnessed several applications concerning gluten, mainly in clinical contexts but also for its detection in food. Immunized llamas [33] and a semi-synthetic library [34] have been used for these purposes. However, this study marks the inaugural production and characterization of antibody fragments derived from immune libraries sourced from celiac patients' peripheral blood. These fragments, whether originating from total or partial immune libraries, have been deployed for gluten detection in food.

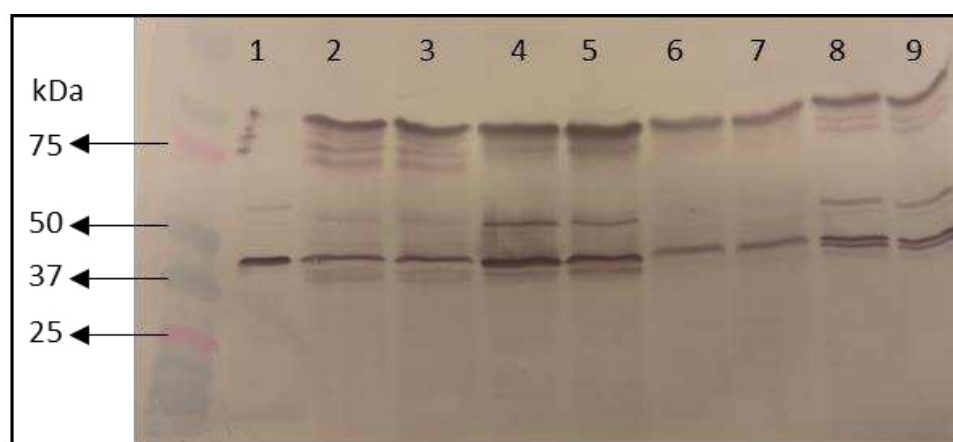
### 3.1. Expression and Purification of Recombinant Fabs

Four gliadin-binding recombinant phage-Fabs were selected by phage-display for their soluble expression and characterization from two libraries of different origins previously obtained by our group. One of the Fabs (Fab8E-4) was isolated from a library that resulted

from merging semi-synthetic heavy chains and immune light chains [26], and the other three (Fab-C, Fab-E, and Fab-H) were isolated from a fully immune phage-display library [27].

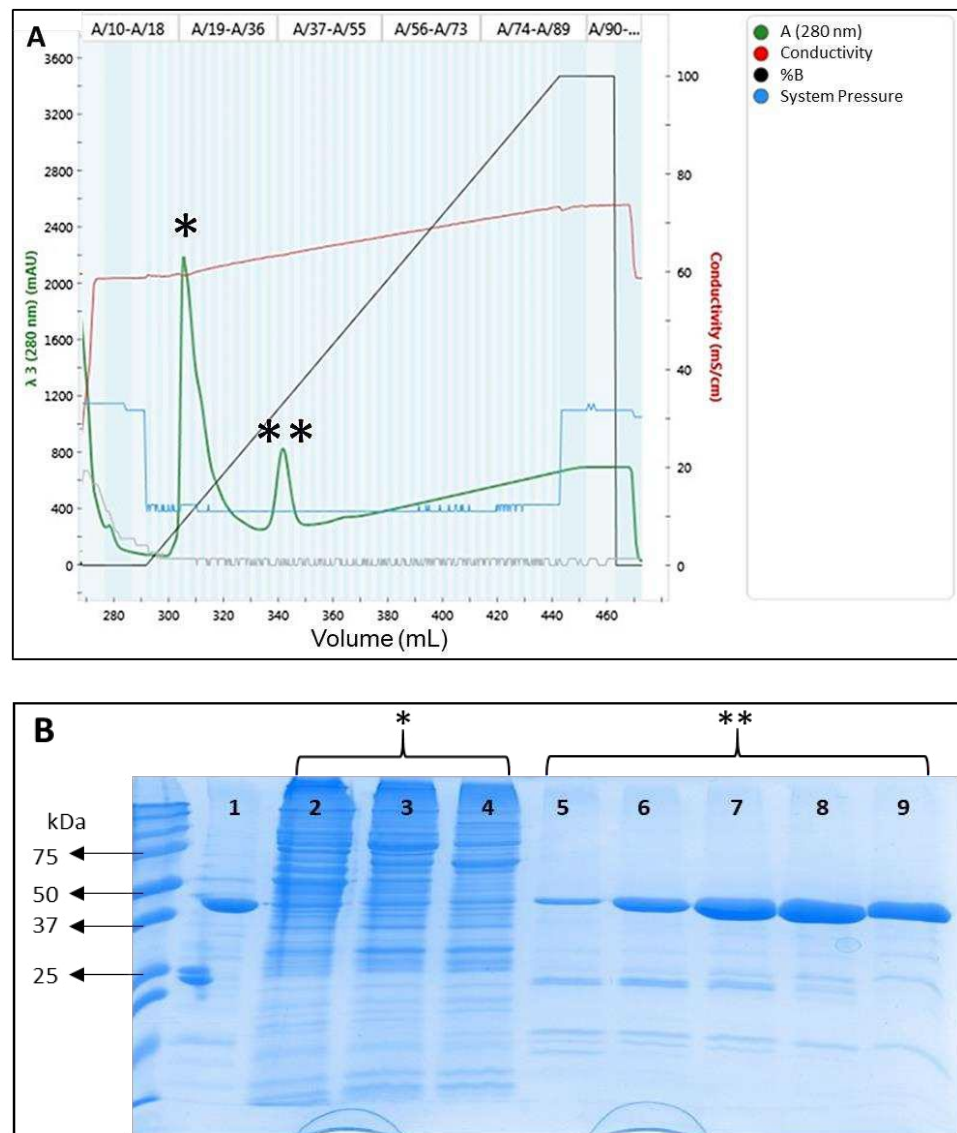
To enhance the recombinant expression of Fabs, the original isotypes found after their isolation from the library were switched to IgG1. The isotype change allowed the use of the same secondary antibody for different primary Fabs, increasing test uniformity. It should be noticed that most of the antibodies selected from the immune library presented an IgA isotype, an antibody class that presents an extra disulfide bond, which challenges the production and folding of the recombinant protein [35]. However, it was not surprising to find IgA antibodies because the library they were isolated from was produced from peripheral blood lymphocytes from celiac disease patients, and celiac disease is a mucosal autoimmune condition where gluten-directed IgA antibodies are generated [36].

The recombinant Fabs were expressed in the supernatant thanks to the combination of the clonation of the antibody chains after a *pelB* secretion signal and the weak cell wall of the bacterial strain used. After growing and inducing the transformed colonies for recombinant production, a Western blot was performed to confirm the effective expression of the Fabs in the supernatant prior to purification. All the supernatants contained bands coincident with the Fab size (approx 50 kDa), and the size of the Fab used as a control. Apart from the Fab, other bands in the gel are proteins secreted to the medium by *E. coli*. These results demonstrated the suitability of the strategy used for the heterologous expression of the selected Fabs and allowed the progression of the purification process (Figure 2).



**Figure 2.** Western blot analysis of transformed *E. coli* RV308 culture supernatants after induction for expression of recombinant Fabs. Lane 1: control Fab previously produced; Lanes 2 and 3: Fab-C; Lanes 4 and 5: Fab-E; Lanes 6 and 7: Fab-H; Lanes 8 and 9: Fab8E-4.

The antibody fragments, produced as histidine-tagged proteins, were purified from the cleared supernatant through IMAC (Immobilized Metal Affinity Chromatography). The isolation process was monitored by spectrometry (Figure 3A). When the imidazole gradient was activated, the bound proteins were eluted, producing two significant protein signal peaks that were analyzed by SDS-PAGE (Figure 3B). The proteins eluted in the first peak were *E. coli* proteins bound to the column and those from the second peak, the Fabs. The isolation of the target Fabs was confirmed by mass fingerprinting of the band with the appropriate molecular weight (approx 50 kDa) of a Fab. For Fab-C, the hit analysis of the mass fingerprint against the sequence of the Fab showed a protein score (a measure of the probability that the observed match is a random event) of 84, when the statistical analysis indicated that scores greater than 13 are significant ( $p < 0.05$ ). Among the tryptic peptides identified, there were some containing part of the constant regions and the HCDR3 (third Complementarity-Determining Region of the Heavy Chain) that allowed a clear identification of the Fab-C.



**Figure 3.** Recombinant production, purification, and identification of Fabs. **(A)** Example of the immobilized metal affinity chromatography (IMAC) profile for the purification of the recombinant Fabs (Fab-C in this image). In green, monitorization of the protein quantity (absorbance 280 nm) going through the column. In red, conductivity measures. In black, percentage of elution buffer pumped to the system. In blue, measure of the system pressure. **(B)** Analysis of the purification steps by SDS-PAGE electrophoresis. Lane 1: control Fab previously produced. Lanes 2–4: samples from different fractions representing the first peak of the chromatogram (\*). Lanes 5–9: samples from different fractions representing the second peak of the chromatogram (\*\*) containing the desired Fab.

This data confirmed that the Fabs were found in fractions eluted at a concentration of approximately 100 mM of imidazole.

To ensure that the desired antibody fragment was obtained with an acceptable level of purity and homogeneity, the purified fraction was dialyzed to remove the imidazole. The recombinant production strategy resulted in the obtention of between 6 and 9 milligrams per antibody, a high yield for *E. coli* production using non-continuous cultures (flasks) [37]. Moreover, the proposed methodology has been proven as a fast, accurate, and reproducible

option to produce recombinant Fabs, allowing high uniformity between batches of the same and also different antibodies [38].

### 3.2. Indirect ELISA Development and Validation for Gluten Detection

The specificity of the recombinant Fabs for gluten was thoroughly evaluated by an indirect ELISA, using the purified recombinant Fabs as probes and analyzing immunoplates coated with wheat, barley, and rye as the target species and maize, oats, and rice flours as non-target species (Figure 4A). As expected, the soluble Fabs antibodies did not exhibit any cross-reactivity with the cereals not containing gluten despite their close phylogenetic relationship.

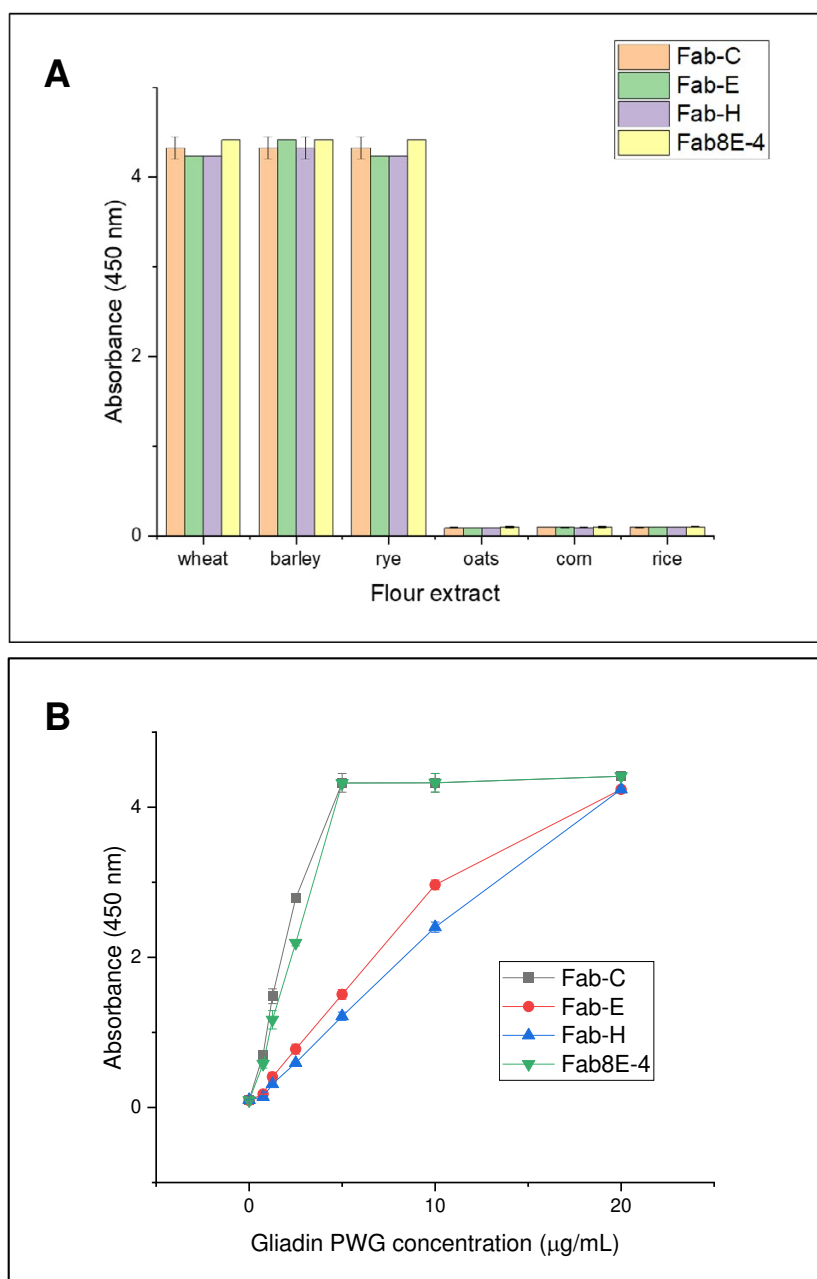
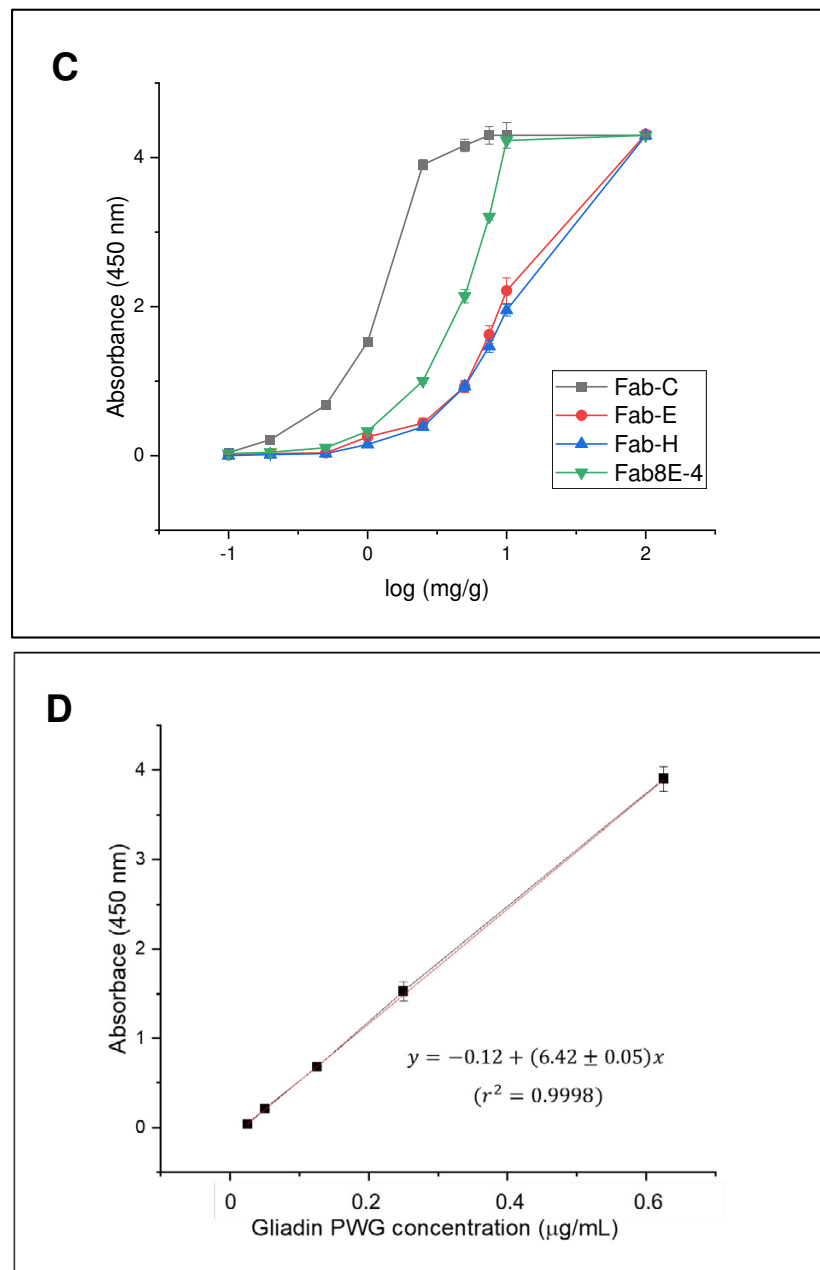


Figure 4. Cont.



**Figure 4.** Indirect ELISA results for characterization of the four recombinant Fabs produced. (A) Specificity results against ethanolic extracts of gluten-containing (wheat, barley, and rye) and gluten-free (oats, corn, and rye) flours (diluted 1:5 in PBS). (B) Sensitivity evaluation by comparative dose-response curves obtained against gliadin-PWG (0–20 µg/mL) diluted in PBS. (C) Comparison of dose-response curves for detection of gluten in an experimental mixture (rice flour spiked with growing concentrations of wheat, rye, and barley flours). (D) Linear regression results obtained from the analysis of the gliadin-PWG standard (0.025–0.625 µg/mL) by indirect ELISA using Fab-C as primary antibody.

Once the specificity of the four Fabs was demonstrated, the sensitivity of the indirect ELISA using the different purified recombinant Fabs was evaluated by analyzing gliadin-PWG reference material (Figure 4B). The results showed that Fab-C exhibited the highest

sensitivity, followed by Fab8E-4. Conversely, Fab-E and Fab-H demonstrated comparatively lower sensitivity.

The performance of the indirect ELISA test with soluble Fabs was not only evaluated against purified gliadin but also in a matrix that resembles the intended application for gluten detection in food. To this end, an ELISA test was performed to detect gluten in an experimental flour mixture consisting of increasing concentrations of gluten containing wheat, barley, and rye flours blended with a rice flour matrix (non-containing gluten) (Figure 4C). These results reinforced those obtained from the gliadin-PWG curves, and the Fabs demonstrated consistent performance in different tests.

A gluten recovery analysis was performed by testing the gluten-like proteins extracted from a gluten-free certified rice flour spiked with solid gliadin-PWG (equivalent to theoretical 20 and 100 mg/kg of gluten) with the indirect Fab-C ELISA. The actual gluten content of these mixtures was assessed by analysis with the R5 sandwich ELISA. The results of the recovery test for the methodology proposed were  $76.71 \pm 1.72\%$  for the 20 mg/kg mixture and  $90.89 \pm 0.36\%$  for the 100 mg/kg mixture. As a control, a rice flour extract was spiked with the equivalent quantities of dissolved gliadin-PWG (0.05  $\mu\text{g}/\text{mL}$ , equivalent to 20 mg/kg in food samples, and 0.25  $\mu\text{g}/\text{mL}$ , equivalent to 100 mg/kg). The results for these controls were  $79.93 \pm 3.35\%$  for the 0.05  $\mu\text{g}/\text{mL}$  solution and  $92.11 \pm 1.07\%$  for the 0.25  $\mu\text{g}/\text{mL}$  solution. All the results obtained presented an appropriate recovery limit in the 75–125% range [39].

In summary, the results demonstrated that Fab-C exhibited the strongest response against gluten, using purified gliadin-PWG and gluten-like proteins extracted from an experimental flour mixture.

Moreover, 60 heterologous samples were analyzed by an indirect ELISA to assess any potential cross-reactivity. The results demonstrated that Fab-C exhibited no cross-reactivity with these matrices, indicating its high specificity for gluten detection. These findings suggest that Fab-C may be a reliable and accurate tool for gluten immunodetection in food products, as it can effectively differentiate between gluten-containing and non-gluten-containing samples.

The sensitivity of the Fab-C-based indirect ELISA was evaluated using a linear model based on increasing concentrations of gliadin-PWG, ranging from 0.025 to 0.625  $\mu\text{g}/\text{mL}$  (Figure 4D). The limit of detection (LOD) was calculated as 0.028  $\mu\text{g}/\text{mL}$  by interpolating three times the standard deviation of ten blanks. This corresponds to a gluten concentration of 11 mg/kg in the samples (considering a dilution factor of 5, an extraction factor of 40 (0.25 g of sample extracted with 10 mL of buffer), and the conversion factor of gliadin to gluten, 2). The limit of quantification (LOQ) was determined by interpolating ten times the standard deviation of the blank, being 0.05  $\mu\text{g}/\text{mL}$  of gliadin-PWG, equivalent to 19.8 mg/kg of gluten. These results confirm that the indirect ELISA using Fab-C recombinant antibody fulfills the limits established by the current legislation, and it will differentiate gluten-free products (less than 20 mg/kg of gluten) from gluten-containing ones.

The sensitivity of the assay is slightly lower than that of the R5 sandwich ELISA, which could be explained by the different types of assay and antibodies involved. The developed indirect ELISA shows a single molecular interaction per antigen molecule because the Fab possesses only one paratope, thereby allowing for a single binding event per Fab. Conversely, the R5 method involves two molecular interactions with the antigen due to the presence of two paratopes per whole antibody molecule and the utilization of a direct sandwich format, enabling dual interactions involving the capture and detection of antibodies [40]. Considering these factors, it is demonstrated that the single-paratope molecule (Fab-C) exhibits a high affinity towards the target antigen.

### 3.3. Detection of Gluten in Commercial Food Products by Indirect ELISA Based on the Recombinant Fab-C

Once the proposed indirect ELISA with Fab-C demonstrated good specificity, sensitivity, and gluten recovery features, the methodology was applied to analyze gluten-like

proteins extracted from a wide variety of commercial food products, including cereal, dairy, meat, and drinks, processed and raw products.

Samples were sorted according to their labeling in four different categories: (A) products that declared to contain gluten (9 samples); (B) products with “may contain gluten” precautionary labeling (6 samples); (C) products that did not declare gluten or that did not specifically warn of the presence of gluten (21 samples); and (D) products with “gluten-free” labeling and/or certification (14 samples).

To assure the accuracy of the proposed test, the food samples were also analyzed by the gold standard technique for gluten detection in food, a direct sandwich ELISA based on the monoclonal antibody R5. The results are summarized in Table 2. Most of the samples analyzed produced the same result with both methods and showed good agreement with labeling regarding the classification of the food samples following the guidelines of European legislation.

**Table 2.** Results obtained for the detection of gluten in commercial food products using the developed Fab-C-based indirect ELISA and the sandwich monoclonal R5 antibody for result confirmation.

Products	No. of Products	Indirect Fab-C	Sandwich R5
(A) Products declared to contain gluten in the labeling (9)			
Pasta	2	+(2)	+(2)
Soup or rice plates	3	+(3)	+(3)
Oats flakes	1	+(1)	+(1)
Cereal bars	2	+(2)	+(2)
Dairy products	1	+(1)	+(1)
(B) Products with “may contain” precautionary labeling (6)			
Cereal	4	+(3)/−(1)	+(3)/−(1)
Flakes	2	−(2)	−(2)
(C) Products that did not declare gluten or that did not specifically warn of the presence of gluten (21)			
Cereal products	9	+(7)/−(2)	+(7)/−(2)
Oat drinks	3	+(1)/−(2)	+(1)/−(2)
Oat flakes	8	+(2)/−(6)	+(5)/−(3)
Meat	1	−(1)	−(1)
(D) Products with “gluten free” labeling and/or certification (14)			
Cookies/cakes	6	−(6)	−(6)
Pasta	1	−(1)	−(1)
Baby food	2	−(2)	−(2)
Corn flakes	1	−(1)	−(1)
Oat flakes	1	−(1)	−(1)
Meat products	3	−(3)	−(3)

Minus sign (−) indicates gluten content lower than the legal limit of 20 mg/kg of gluten, which allows the labeling with “gluten free” statement, and plus sign (+) indicates values above the mentioned legal limit. The gliadin PWG standard curve was used as reference for the indirect Fab-C ELISA and the R5 Sandwich ELISA.

From the 50 products analyzed with the Fab-C indirect ELISA, the results of 47 matched with the validation technique used (R5-ELISA). However, three products tested positive by the R5 reference assay but negative by the Fab-C indirect ELISA. It is remarkable that the samples that did not conform with the R5 standard test contained oats as the main ingredient. It is hypothesized that this result can be attributed to the absence of cross-reactivity exhibited by Fab-C with oats. It has been demonstrated that certain oat cultivars produce cross-reactivity when analyzed with other detection antibodies like R5, reporting up to 100 mg/kg of gluten in some oats cultivars [41]. The cross-reaction of the R5 monoclonal antibody has been explained due to the high similarity of certain peptides found in avenins (gluten-like proteins in oats) with gluten peptides [42]. The absence of

this unintended detection in gluten-free oat-derived products could mean a stronghold of Fab-C, and it could help in the assessment of gluten in these types of products, especially those containing cultivars that are cross-reactive with other antibodies in the market. In light of these results, the proposed methodology demonstrated a reliable capacity for discriminating gluten-free products in the samples analyzed.

Detection of gluten in foodstuff is still a major challenge in food science because the concept of gluten is defined by chemical features (ethanol solubility), encompassing a very big range of proteins (unlike other food allergens that are represented by one or few proteins) that present some common characteristics but can differ in their sequences [43]. This fact makes it very hard to discover a universal probe for gluten detection. Several comparative studies unveiled that the commercial kits available in the market presented important discrepancies between them for the quantification of gluten [40,44]. These differences were explained by the different epitopes of gluten that the antibodies were binding to. For example, the R5 antibody recognized epitopes that were more expressed in rye, barley, or triticale than in wheat, in contrast to the G12 monoclonal antibody, which presented a better detection of wheat epitopes [45]. The recognition of wheat, rye, and barley was very similar in an indirect ELISA with the recombinant Fabs produced in this work (Figure 4A). Moreover, although R5 and G12 monoclonal antibodies each detected different epitopes, all were found in the N-terminal zone of the gliadins. On the contrary, computational studies with some of the antibodies developed in this work presented interactions with the gliadin C-terminal portions like Fab8E-4 [26] and Fab-C [27]. The methodology applied in this work for generating Fabs (construction of celiac-derived Fab libraries for phage display selection) allowed the transference of some features from the humoral response of celiac patients to the obtained probes instead of the classical approach of using polyclonal or monoclonal antibodies, generated by a forced immune response of experimentation animals [13]. This feature could be important, as the recombinant probes produced are prone to recognize those epitopes that may be harmful to gluten-sensitive patients.

#### 4. Conclusions

Using the pKKtac plasmid and *E. coli* RV308 expression system, four soluble recombinant Fabs obtained from two different phage display libraries (immune and merged semi-synthetic) were produced. This recombinant production methodology provides a cost-effective and efficient alternative to traditional antibody production methods, which can be time-consuming and expensive. Moreover, recombinant antibodies also offer a more consistent and reproducible approach to antibody production, allowing for the availability of standardized affinity probes. All these benefits are achieved without the need to use animals.

The four recombinant Fabs produced selective reactivity to gluten, but Fab-C demonstrated the best detection features in indirect ELISA methodology. Fab-C was capable of detecting gliadin traces at very low concentrations, as low as 28 ng/mL equivalent to 11 mg/kg in food samples, not showing cross-reactions with all the analyzed heterologous species that do not contain gluten. In addition, the ELISA test developed meets the legislative requirements for identifying gluten-free products, and it could have a desirable feature not covered by some commercial tests, as it does not cross-react with oat proteins.

**Author Contributions:** E.G.-C.: methodology, investigation, experimentation, validation, formal analysis, writing original draft. A.G.-G.: investigation, formal analysis, and writing and revision. S.R.: investigation and formal analysis. K.T.: assessment and support on recombinant antibody production. R.M.: funding acquisition and project management. T.G.: supervision, writing and revision, funding acquisition, and project management. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by Ministerio de Ciencia e Innovación Grant: PID2021-122925OB-I00. PRE2018-08427 fellowship of the Ministerio de Ciencia e Innovación granted to Eduardo Garcia-Calvo, which funded his visit to VTT.

**Data Availability Statement:** Data is contained within the article.

**Acknowledgments:** We thank the biosensor team from Espoo VTT premises for their support in the production of the recombinant proteins here presented. In addition, we would like to thank the proteomic unit of Universidad Complutense de Madrid for their support with the mass fingerprint analysis.

**Conflicts of Interest:** Author Kristiina Takkinen was employed by VTT Technical Research Centre of Finland Ltd. VTT Technical Research Centre of Finland Ltd. is a state-owned and controlled non-profit limited liability company established by law and operating under the ownership steering of the Finnish Ministry of Employment and the Economy. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Hochegger, R.; Mayer, W.; Prochaska, M. Comparison of R5 and G12 antibody-based ELISA used for the determination of the gluten content in official food samples. *Foods* **2015**, *4*, 654–664. [[CrossRef](#)] [[PubMed](#)]
2. Lammers, K.; Herrera, M.; Dodero, V. Translational chemistry meets gluten-related disorders. *Chem. Open* **2018**, *7*, 217–232. [[CrossRef](#)] [[PubMed](#)]
3. Denham, J.M.; Hill, I.D. Celiac disease and autoimmunity: Review and controversies. *Curr. Allergy Asthma Rep.* **2013**, *13*, 347–353. [[CrossRef](#)] [[PubMed](#)]
4. Caio, G.; Volta, U.; Sapone, A.; Leffler, D.A.; De Giorgio, R.; Catassi, C.; Fasano, A. Celiac disease: A comprehensive current review. *BMC Med.* **2019**, *17*, 142. [[CrossRef](#)]
5. Rotondi Aufiero, V.; Fasano, A.; Mazzarella, G. Non-celiac gluten sensitivity: How It's gut immune activation and potential dietary management differ from celiac disease. *Mol. Nutr. Food Res.* **2018**, *62*, 1700854. [[CrossRef](#)]
6. Itzlinger, A.; Branchi, F.; Elli, L.; Schumann, M. Gluten-free diet in celiac disease—forever and for all? *Nutrients* **2018**, *10*, 1796. [[CrossRef](#)]
7. Leffler, D.A.; Edwards-George, J.; Dennis, M.; Schuppan, D.; Cook, F.; Franko, D.L.; Blom-Hoffman, J.; Kelly, C.P. Factors that influence adherence to a gluten-free diet in adults with celiac disease. *Dig. Dis. Sci.* **2008**, *53*, 1573–1581. [[CrossRef](#)]
8. Day, L.; Augustin, M.A.; Batey, I.L.; Wrigley, C.W. Wheat-gluten uses and industry needs. *Trends Food Sci. Technol.* **2006**, *17*, 82–90. [[CrossRef](#)]
9. EU. Commission Implementing Regulation (EU) No. 828/2014 of 30 July on the Requirements for the Provision of Information to Consumers on the Absence or Reduced Presence of Gluten in Food OJEU (Official Journal of the European Union) L 228, 31.7.2014; EU: Luxembourg, 2014; pp. 5–8.
10. Osorio, C.E.; Mejías, J.H.; Rustgi, S. Gluten detection methods and their critical role in assuring safe diets for celiac patients. *Nutrients* **2019**, *11*, 2920. [[CrossRef](#)]
11. Garcia-García, A.; Madrid, R.; Sohrabi, H.; Cruz, S.; García, T.; Martín, R.; González Alonso, I. A sensitive and specific real-time PCR targeting DNA from wheat, barley and rye to track gluten contamination in marketed foods. *LWT-Food Sci. Technol.* **2019**, *114*, 108378. [[CrossRef](#)]
12. Mejías, J.; Lu, X.; Osorio, C.; Ullman, J.; Von Wettstein, D.; Rustgi, S. Analysis of wheat prolamins, the causative agents of celiac sprue, using reversed phase high performance liquid chromatography (RP-HPLC) and matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS). *Nutrients* **2014**, *6*, 1578–1597. [[CrossRef](#)]
13. Scherf, K.A.; Poms, R.E. Recent developments in analytical methods for tracing gluten. *J. Cereal Sci.* **2016**, *67*, 112–122. [[CrossRef](#)]
14. Ferranti, P.; Mamone, G.; Picariello, G.; Addeo, F. Mass spectrometry analysis of gliadins in celiac disease. *J. Mass Spectrom.* **2007**, *42*, 1531–1548. [[CrossRef](#)] [[PubMed](#)]
15. Lerner, B.A.; Phan Vo, L.T.; Yates, S.; Rundle, A.G.; Green, P.H.R.; Lebowhl, B. Detection of gluten in gluten-free labeled restaurant food: Analysis of crowd-sourced data. *Am. J. Gastroenterol.* **2019**, *114*, 792–797. [[CrossRef](#)]
16. Rzychon, M.; Brohée, M.; Cordeiro, F.; Haraszi, R.; Ulberth, F.; O'Connor, G. The feasibility of harmonizing gluten ELISA measurements. *Food Chem.* **2017**, *234*, 144–154. [[CrossRef](#)]
17. Sorell, L.; López, J.A.; Valdés, I.; Alfonso, P.; Camafeita, E.; Acevedo, B.; Chirido, F.; Gavalondo, J.; Méndez, E. An innovative sandwich ELISA system based on an antibody cocktail for gluten analysis. *FEBS Lett.* **1998**, *439*, 46–50. [[CrossRef](#)]
18. Méndez, E.; Vela, C.; Immer, U.; Janssen, F.W. Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *Eur. J. Gastroenterol. Hepatol.* **2005**, *17*, 1053–1063. [[CrossRef](#)]
19. Kunert, R.; Reinhart, D. Advances in recombinant antibody manufacturing. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 3451–3461. [[CrossRef](#)]
20. Bradbury, A.R.M.; Trinklein, N.D.; Thie, H.; Wilkinson, I.C.; Tandon, A.K.; Anderson, S.; Bladen, C.L.; Jones, B.; Aldred, S.F.; Bestagno, M.; et al. When monoclonal antibodies are not monospecific: Hybridomas frequently express additional functional variable regions. *MAbs* **2018**, *10*, 539–546. [[CrossRef](#)]
21. Winter, G. Harnessing Evolution to Make Medicines (Nobel Lecture). *Angew. Chem. Int. Ed.* **2019**, *58*, 14438–14445. [[CrossRef](#)]

22. Jaroszewicz, W.; Morcinek-Orłowska, J.; Pierzynowska, K.; Gaffke, L.; Węgrzyn, G. Phage display and other peptide display technologies. *FEMS Microbiol. Rev.* **2021**, *46*, fuab052. [[CrossRef](#)] [[PubMed](#)]
23. Holt, L.J.; Herring, C.; Jespers, L.S.; Woolven, B.P.; Tomlinson, I.M. Domain antibodies: Proteins for therapy. *Trends Biotechnol.* **2003**, *21*, 484–490. [[CrossRef](#)] [[PubMed](#)]
24. García-García, A.; Madrid, R.; Garcia-Calvo, E.; Mendoza-Chamizo, B.; García, T.; Martín, R. Production of a recombinant single-domain antibody for gluten detection in foods using the *Pichia pastoris* expression system. *Foods* **2020**, *9*, 1838. [[CrossRef](#)] [[PubMed](#)]
25. Bates, A.; Power, C.A. David vs. Goliath: The structure, function, and clinical prospects of antibody fragments. *Antibodies* **2019**, *8*, 28. [[CrossRef](#)] [[PubMed](#)]
26. Garcia-Calvo, E.; García-García, A.; Rodríguez, S.; Farrais, S.; Martín, R.; García, T. Construction of a Fab library merging chains from semisynthetic and immune origin, suitable for developing new tools for gluten immunodetection in food. *Foods* **2022**, *12*, 149. [[CrossRef](#)]
27. Garcia-Calvo, E.; García-García, A.; Rodríguez Gómez, S.; Farrais, S.; Martín, R.; García, T. Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods. *Curr. Res. Food Sci.* **2023**, *7*, 100578. [[CrossRef](#)]
28. Takkinen, K.; Laukkanen, M.L.; Sizmann, D.; Alfthan, K.; Immonen, T.; Vanne, L.; Kaartinen, M.; Knowles, J.K.; Teeri, T.T. An active single-chain antibody containing a cellulase linker domain is secreted by *Escherichia coli*. *Protein Engineering* **1991**, *4*, 837–841. [[CrossRef](#)]
29. Sechi, S.; Chait, B.T. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* **1998**, *70*, 5150–5158. [[CrossRef](#)]
30. Van Eckert, R.; Berghofer, E.; Ciclitira, P.J.; Chirido, F.; Denery-Papini, S.; Ellis, H.J.; Ferranti, P.; Goodwin, P.; Immer, U.; Mamone, G.; et al. Towards a new gliadin reference material— isolation and characterisation. *J. Cereal Sci.* **2006**, *43*, 331–341. [[CrossRef](#)]
31. García-García, A.; Madrid, R.; González, I.; García, T.; Martín, R. A novel approach to produce phage single domain antibody fragments for the detection of gluten in foods. *Food Chem.* **2020**, *321*, 126685. [[CrossRef](#)]
32. Thompson, M.; Ellison, S.L.R.; Wood, R. Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC Technical Report). *Pure Appl. Chem.* **2002**, *74*, 835–855. [[CrossRef](#)]
33. Doña, V.; Urrutia, M.; Bayardo, M.; Alzogaray, V.; Goldbaum, F.A.; Chirido, F.G. Single domain antibodies are specially suited for quantitative determination of gliadins under denaturing conditions. *J. Agric. Food Chem.* **2010**, *58*, 918–926. [[CrossRef](#)]
34. Garcia-Calvo, E.; García-García, A.; Madrid, R.; Martín, R.; García, T. From polyclonal sera to recombinant antibodies: A review of immunological detection of gluten in foodstuff. *Foods* **2020**, *10*, 66. [[CrossRef](#)]
35. Gaciarz, A.; Veijola, J.; Uchida, Y.; Saaranen, M.J.; Wang, C.; Hörkkö, S.; Ruddock, L.W. Systematic screening of soluble expression of antibody fragments in the cytoplasm of *E. coli*. *Microb. Cell Factories* **2016**, *15*, 22. [[CrossRef](#)]
36. Koninckx, C.R.; Giliams, J.P.; Polanco, I.; Peña, A.S. IgA anti-gliadin antibodies in celiac and inflammatory bowel disease. *J. Pediatr. Gastroenterol. Nutr.* **1984**, *3*, 676–682. [[CrossRef](#)]
37. Frenzel, A.; Hust, M.; Schirrmann, T. Expression of Recombinant Antibodies. *Front. Immunol.* **2013**, *4*, 217. [[CrossRef](#)]
38. De Marco, A. Recombinant antibody production evolves into multiple options aimed at yielding reagents suitable for application-specific needs. *Microb. Cell Factories* **2015**, *14*, 125. [[CrossRef](#)]
39. Thompson, M.; Ellison, S.L.R.; Fajgelj, A.; Willetts, P.; Wood, R. Harmonized guidelines for the use of recovery information in analytical measurement. *Pure Appl. Chem.* **1999**, *71*, 337–348. [[CrossRef](#)]
40. Amnuaycheewa, P.; Niemann, L.; Goodman, R.E.; Baumert, J.L.; Taylor, S.L. Challenges in gluten analysis: A comparison of four commercial sandwich ELISA kits. *Foods* **2022**, *11*, 706. [[CrossRef](#)]
41. Benoit, L.; Masiri, J.; Del Blanco, I.A.; Meshgi, M.; Gendel, S.M.; Samadpour, M. Assessment of avenins from different oat varieties using R5-based sandwich ELISA. *J. Agric. Food Chem.* **2017**, *65*, 1467–1472. [[CrossRef](#)]
42. Hardy, M.Y.; Tye-Din, J.A.; Stewart, J.A.; Schmitz, F.; Dudek, N.L.; Hanchapola, I.; Purcell, A.W.; Anderson, R.P. Ingestion of oats and barley in patients with celiac disease mobilizes cross-reactive T cells activated by avenin peptides and immuno-dominant hordein peptides. *J. Autoimmun.* **2015**, *56*, 56–65. [[CrossRef](#)]
43. Wieser, H.; Koehler, P.; Scherf, K.A. Chemistry of wheat gluten proteins: Qualitative composition. *Cereal Chem.* **2023**, *100*, 23–35. [[CrossRef](#)]
44. Bugyi, Z.; Török, K.; Hajas, L.; Adonyi, Z.; Popping, B.; Tömösközi, S. Comparative study of commercially available gluten ELISA kits using an incurred reference material. *Qual. Assur. Saf. Crops Amp. Foods* **2013**, *5*, 79–87. [[CrossRef](#)]
45. Lexhaller, B.; Tompos, C.; Scherf, K.A. Fundamental study on reactivities of gluten protein types from wheat, rye and barley with five sandwich ELISA test kits. *Food Chem.* **2017**, *237*, 320–330. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

## CAPÍTULO 4.4

**Evaluación de la cuantificación de gluten en productos comerciales mediante una metodología de ELISA tipo sándwich basada en el empleo de nuevos anticuerpos recombinantes**

---

**Exploring gluten quantification in marketed products through a sandwich-ELISA methodology based on novel recombinant antibodies**

**Comunicación breve en preparación**



## Exploring gluten assessment in marketed products through a sandwich-ELISA methodology based on novel recombinant antibodies.

### ABSTRACT

Based on four recombinant Fab antibodies against gliadin previously isolated by phage display, this study describes the development of a sandwich ELISA methodology for gluten detection in foods. Firstly, Fab molecules were chemically biotinylated to be used immobilized on streptavidin-coated plates, as capture antibodies and, alkaline phosphatase-conjugated Fab as detection antibodies. Although the use of this direct approach eliminated the need for secondary antibodies, a loss of signal was observed in the ELISA results that was attributed to the chemical biotinylation procedure. From these experiments, it was revealed that the Fab-pair with the best compatibility was Fab8E-4 as the capture probe and Fab-C for the detection reaction. Subsequently, the unmodified antibodies were used to develop a sandwich immunoassay capable of detecting 26 ng/mL of gliadin, which corresponds to a limit of detection of 10 mg/kg of gluten in foods. A total of 60 different gluten-free species commonly used in the food industry were assayed with the developed ELISA showing no cross-reactivity. Finally, the results of the analysis of 50 commercial products showed consistent results when compared with the standard method for gluten detection. Moreover, the developed immunoassay exhibited no cross-reactivity with oat products, which could be a potential advantage over other available detection systems that cross-react with this cereal.

## 1 INTRODUCTION

Gluten, an ethanol-soluble protein fraction found in grains, constitutes a substantial portion, often up to 90%, of wheat proteins (Shewry, 2019). Specific components of gluten, present in various grain species such as wheat, rye, and barley, can trigger adverse effects when consumed by individuals with particular sensitivities. Gluten consumption has been associated with several diseases, classified based on their etiopathology (Cabanillas, 2020). The first category encompasses allergic reactions to gluten components (Inomata, 2009), while the second includes autoimmune processes, notably celiac disease, the most extensively studied gluten-related disorder with a prevalence of 1% in Western countries (Tye-Din et al., 2018). The third category refers to non-celiac gluten intolerance, a condition that does not align with the previous groups. Its molecular mechanism remains elusive, but there is a compelling implication of innate immunity in its pathogenesis (Cárdenas-Torres et al., 2021).

The most efficient and economically viable treatment for gluten-related disorders is a gluten-free diet. It not only arrests disease progression but also fosters tissue healing (Itzlinger et al., 2018). This situation has prompted the establishment of stringent labeling regulations aimed at ensuring the provision of gluten-free diets for individuals allergic or intolerant. Within the European Union, any food product containing gluten must bear clear labeling on its packaging. Such labeling should clearly indicate the presence of gluten in the product, and it must include the precise terms “contains gluten” or “contains cereals containing gluten” within the ingredients list or in a dedicated allergen statement. Furthermore, to qualify for the designation of “gluten-free,” a food product must not exceed 20 parts per million (ppm) of gluten content, while “very low gluten” labeling is permissible when the gluten content remains below 100 ppm” (OJEU, 2014).

Immunoassays offer unparalleled relevance in the context of gluten detection in food products when compared to alternative analytical methodologies. Their sensitivity, specificity, and efficiency in detecting gluten, even at very low concentrations, underscore their importance. Immunoassays, including enzyme-linked immunosorbent assays (ELISAs) and lateral flow devices, employ specialized antibodies designed to specifically target (gluten proteins). Furthermore, immunoassays yield rapid results and are often more cost-effective than alternative methods like polymerase chain reaction (PCR) or mass spectrometry, justifying their widespread adoption across the food industry (Garcia-Calvo et al., 2020).

Sandwich-ELISA offers several distinct advantages in the realm of immunological and biochemical research. One key advantage is its exceptional sensitivity and specificity. By utilizing two antibodies that bind to different epitopes on the target molecule, sandwich ELISA minimizes the potential for false-positive or false-negative results, making it highly reliable for detecting and quantifying specific proteins or antigens (Wieser et al., 2021). Additionally, this assay format allows for the measurement of proteins in complex mixtures, without interference from

other molecules. It provides quantitative data, enabling precise concentration measurements, robustness, reproducibility, and adaptability that have made it an indispensable tool in various fields, including gluten detection in food, where the most used methodology is a sandwich-ELISA based in a classical monoclonal antibody named R5 (Méndez et al., 2005).

In earlier studies, novel approaches were introduced to develop recombinant antibody fragments, Fab (Fragment antigen binding), for gluten detection in foodstuffs. These approaches involved the construction of two distinct antibody libraries. The first library was a comprehensive immune library derived from lymphocytic RNA (Garcia-Calvo et al., 2023a), while the second library was a hybrid of semi-synthetic heavy chain (Garcia-Calvo et al., 2022), based on a pre-existing dAb template (an antibody fragment consisting of the variable region of the heavy chain), and immune light chains obtained from the aforementioned immune library. Several Fab were successfully isolated against gluten through phage display technology, and subsequently produced as recombinant proteins. This process resulted in the obtention of four high-affinity Fab (Fab-C, Fab-H, and Fab-E from the first library, and Fab8E-4 from the second one), which were characterized for their suitability as probes in ELISA assays, and one of them was tested for its ability to detect gluten in marketed products (Garcia-Calvo et al., 2023b). The aim of this work was the development of a sandwich ELISA methodology based on those recombinant Fab, to be used as an alternative method for gluten detection in commercial food products.

## 2 MATERIALS AND METHODS

### 2.1 Recombinant Fab used as sandwich-ELISA probes

The Fab (Fab-C, Fab-E, Fab-H and Fab8E-4) were selected by phage display, expressed in the supernatant of *E. coli* K12 RV308, and purified by IMAC (Immobilized Metal Affinity Chromatography) as previously described (Garcia-Calvo et al., 2023b).

### 2.2 Compatibility study between the selected Fab

The compatibility between the Fab in a sandwich disposition was checked in a direct sandwich ELISA, by using *in vitro* biotinylated recombinant Fab, immobilized in streptavidin plates, as capture antibodies and, alkaline phosphatase-conjugated Fab as detection antibodies.

The biotinylation process started with each Fab dissolved in PBS, that was changed to carbonate/bicarbonate buffer pH 8.3 using an Amicon® ultra-15 centrifugal device (30 kDa MWCO, Merk® ref #UFC903024). The chemical biotinylation was carried out using the EZ-link® kit (Thermo®, ref #21363), by adding 37 µg of reactant per Fab milligram and incubating the reaction for 30 min at room temperature. The excess of free biotin was washed away using an Econopac 10 DG column (Biorad®, ref #7322010). The eluted fractions containing protein were identified by

their absorbance at 280 nm. The biotinylated Fab were immobilized to streptavidin-coated plates (Thermo ref #15122) to be used as capture antibodies.

One milligram of each biotinylated Fab was conjugated with streptavidin-alkaline phosphatase following manufacturers' specifications (Merck® ref #S2890) to be used as the detection antibody.

Once the capture antibody was immobilized in the plate, it was blocked with PBS- BSA 3 %. Different antigen concentrations were added, incubated, and washed away. The detection antibody (alkaline phosphatase-conjugated Fab) was then added (0.5 µg/well in the blocking solution), and incubated for 1 h at room temperature. Unbound conjugated Fab were washed away and, the revealing solution (2 mg/mL 4-nitrophenyl phosphate disodium salt hexahydrate pNPP (Sigma©, ref #N4645) dissolved in diethanolamine-MgCl<sub>2</sub> buffer (Reagena©, Toivala, Finland, ref #170057)) was added and measured in a spectrophotometer at 405 nm.

### 2.3 Sandwich ELISA protocol

The capture antibodies were directly adsorbed to the immunosorbent plates (0.5 µg/well of recombinant Fab in PBS) and incubated for 1 h at 37 °C.

Then, the coating solution was shaken out, 200 µL blocking solution (3% BSA in PBS) was added per well, and the plate was incubated for 1 h at room temperature. Following 10 washing steps with PBS, the gliadin-PWG or gluten proteins extracted from samples were diluted in PBS, and 100 µL of antigen dilution were added per well. The plate was incubated for 1 h at 37 °C. After 10 washing steps, 0.5 µg of unmodified detection Fab was added per well, diluted in 100 µL of blocking solution, and incubated for 1 h at room temperature in a plate shaker. The plate was washed 15 times with PBS, and the secondary antibody (rabbit anti-human H + L HRP conjugated, Abcam©, Cambridge, UK ref #6759), diluted in 100 µL of blocking solution was added, and incubated for 1 h at room temperature in a plate shaker. After washing 10 times with PBS, 100 µL of TMB (Sigma©, ref #T0440) was added. The reaction was stopped after 20 min with 50 µL of a diluted sulphuric acid solution, and the signal was detected at 450 nm.

### 2.4 Assay validation and analysis of commercial products

The assays used for validating and characterizing the key analytical features of the proposed methodology were essentially as performed with the indirect ELISA based on Fab-C (García-Calvo et al., 2023b). The limits of detection (LOD) and quantification (LOQ) were calculated by performing a dose-response analysis against a reference material, the gliadin-PWG (prolamin working group) (Van Eckert et al., 2006). To test specificity for gluten, 60 heterologous species that did not contain gluten were tested to ensure there was no cross-reactivity, and a recovery analysis was also implemented. Once the analytical features were characterized, 50 commercial

food products were analyzed to demonstrate the functionality of the assay in its intended actual application.

### 3 RESULTS AND DISCUSSION

#### 3.1 Selection of capture and detection recombinant Fab

The compatibility of different Fab pairs was evaluated using 10 µg/mL of gliadin- PWG as antigen for all combinations. Four gliadin-binding recombinant Fab (Fab8E-4, Fab-E, Fab-H and Fab-C) obtained in a previous work (García-Calvo et al., 2023b), were evaluated as candidates for development of a sandwich ELISA for detection of gluten. To assess compatibility between the four Fab used to detect gliadin, a checkerboard test was performed, using all biotinylated Fab as capture antibodies, and all the alkaline-phosphatase conjugated Fab as detection antibodies (Table 1).

**Table 1.** Checkerboard compatibility test between pairs of recombinant Fab in a direct sandwich ELISA. The values represent average absorbance data (405 nm) in three independent assays.

		Detection Fab			
		Fab8E-4	Fab-E	Fab-H	Fab-C
Capture Fab	Fab8E-4	1.254	0.612	0.412	4.741
	Fab-E	0.841	0.523	0.341	0.745
	Fab-H	0.647	0.389	0.415	0.714
	Fab-C	2.751	0.748	0.477	1.784

The combination of Fab8E-4 as a capture antibody and Fab-C conjugated with alkaline phosphatase as the detection antibody produced the best results. The same Fab could be used as the capture and the detection antibody because the absorbance values were lower than as expected for the saturating concentration of antigen added. This could be explained due to the competition for the same epitopes in the gliadin molecule when using the same antibodies for capture and detection. In contrast, the R5 monoclonal antibody is used both as capture and detection antibody in the sandwich ELISA assay used as reference for gluten detection. That may be due to its ability to recognize repetitive motifs dispersed across gluten proteins, predominantly within the more conserved N-terminal segments of gliadins (Lexhaller et al., 2017).

Computational models (García-Calvo et al., 2022; García-Calvo et al., 2023) indicated that Fab-C and Fab8E-4 exhibit prominent interactions with the gliadin C-terminal regions, albeit in distinct positions (a less conserved region characterized by a reduced occurrence of repetitive motifs).

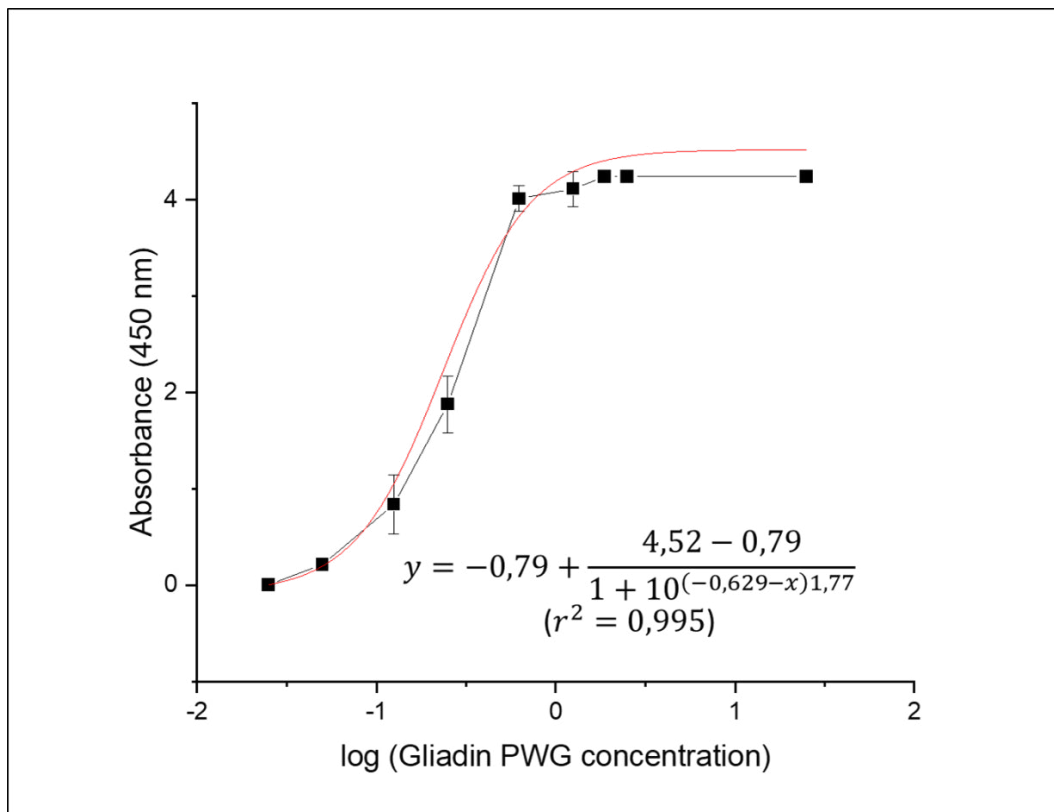
This may explain why pairs of identical Fab compete for the same epitopes, while combinations of dissimilar Fab synergistically enhance detection efficiency.

The *in vitro* biotinylation procedure enabled the conjugation of biotin molecules with up to all available free amino groups. This led to significant alterations in the conformation of the Fab, resulting in a reduced affinity towards the target (Fairhead & Howarth, 2015). While the direct ELISA protocol adequately assessed Fab compatibility, the biotinylation process introduced a discernible decline in the detection signal. Consequently, the ultimate assay configuration adopted was an indirect sandwich ELISA, employing the Fab pair that exhibited superior detection performance.

### 3.2 INDIRECT SANDWICH ELISA DEVELOPMENT AND VALIDATION FOR GLUTEN DETECTION

In order to ascertain the selectivity of the sandwich method, an indirect sandwich ELISA methodology was set up and tested with a panel of 60 diverse samples to evaluate its potential cross-reactivity with matrices devoid of gluten. The outcomes unequivocally showcased that the sandwich ELISA displayed an absence of cross-reactivity with these matrices, underscoring its remarkable specificity in detecting gluten. These results strongly endorse the method as a valuable and precise instrument for gluten immunodetection in food products, as it clearly distinguishes between samples containing gluten and those lacking these proteins.

The sensitivity of the indirect sandwich ELISA was assessed through a dose- response model of increasing concentrations of gliadin-PWG, ranging from 0.025 to 25  $\mu\text{g}/\text{mL}$  (Figure 1). The limit of detection (LOD) was calculated as 0.026  $\mu\text{g}/\text{mL}$  by interpolating three times the standard deviation of ten blank samples. This translates to a gluten concentration of 10.4 mg/kg in the samples, considering a dilution factor of 5, an extraction factor of 40 (0.25 g of sample extracted with 10 mL of buffer), and the conversion factor of gliadin to gluten (2).



**Figure 1.** Dose response curve obtained from the analysis of the gliadin-PWG standard (0.025-25  $\mu\text{g}/\text{mL}$ ) by sandwich ELISA using Fab-C as detection antibody and Fab8E-4 as capture antibody. This model was used to estimate the gluten content of the tested commercial food products. An HRP-conjugated anti-human Fab was used as the secondary antibody. Mean values of three independent determinations and standard deviation of each data set are shown.

Furthermore, the limit of quantification (LOQ) was established by interpolating ten times the standard deviation of the blank samples, resulting in a value of 0.044  $\mu\text{g}/\text{mL}$  of gliadin-PWG, equivalent to 17.4 mg/kg of gluten. These outcomes affirm the method's adherence to the regulatory limits currently in force.

The method developed, effectively distinguishes gluten-free products (with less than 20 mg/kg of gluten) from those containing gluten. It is noteworthy that the sensitivity of this assay is slightly lower than that of the R5 sandwich ELISA, which may be attributed to variations in assay types and the antibodies employed. The newly developed indirect sandwich ELISA relies on a single molecular interaction per antigen molecule due to the Fab's composition of only one paratope (two in total), facilitating a singular binding event per Fab. In contrast, the R5 method entails two molecular interactions with the antigen, owing to the presence of two paratopes per whole antibody molecule and the adoption of a direct sandwich format, allowing dual interactions, that involve four paratopes from the capture and detection antibodies (Amnuaycheewa et al., 2022). The results obtained thus demonstrate that the single-paratope molecules (Fab) exhibit a high affinity for the target antigen.

A gluten recovery analysis using the indirect sandwich ELISA was conducted by testing gluten-like proteins (prolamins) extracted from gluten-free certified rice flour that had been spiked with solid gliadin-PWG, equivalent to theoretical concentrations of 20 and 100 mg/kg of gluten. The actual gluten content of these mixtures was concurrently assessed using the R5 sandwich ELISA. The results obtained yielded recovery percentages of  $104.14 \pm 2.2$  % for the 20 mg/kg mixture and  $95.44 \pm 6.36$  % for the 100 mg/kg mixture. To establish a control baseline, a rice flour extract spiked with equivalent quantities of dissolved gliadin- PWG (0.05  $\mu\text{g/mL}$ , corresponding to 20 mg/kg in food samples, and 0.25  $\mu\text{g/mL}$ , equivalent to 100 mg/kg) was prepared. The results for these control samples demonstrated recovery percentages of  $100.08 \pm 3.67$  % for the 0.05  $\mu\text{g/mL}$  solution and  $95.44 \pm 6.36$  % for the 0.25  $\mu\text{g/mL}$  solution. Accordingly, all the results demonstrated an excellent recovery range close to 100 %, affirming the accuracy and reliability of the method (Thompson et al., 1999).

### 3.3 DETECTION OF GLUTEN IN COMMERCIAL SAMPLES

The developed methodology was applied for the analysis of gluten-like proteins extracted from a wide variety of commercial food products, encompassing categories such as cereals, dairy, meat, and beverages, spanning both processed and raw items. The collected samples were categorized based on their labeling into four distinct groups: (A) products explicitly indicating the presence of gluten (9 samples); (B) products bearing precautionary labeling stating “may contain gluten” (6 samples); (C) products lacking any declaration of gluten presence or specific gluten warnings (21 samples); and (D) products featuring “gluten-free” labeling and/or certification (14 samples).

To ensure the precision and reliability of the proposed test, these food samples underwent parallel analysis using the gold standard technique for gluten detection in food, a direct sandwich ELISA with the monoclonal antibody R5. The summarized results are presented in Table 2. Notably, most of the analyzed samples yielded congruent outcomes with both methods, exhibiting a strong alignment with their respective labeling claims. This alignment underscores compliance with the guidelines stipulated by European legislation for classifying food products in terms of gluten content.

The results of 47 out of the 50 products subjected to analysis using the indirect sandwich ELISA, were consistent with those obtained from the validation technique, the R5-ELISA. Notably, three products yielded positive outcomes when assessed using the R5 reference assay but registered as negative when evaluated with the proposed methodology. A noteworthy observation was that these particular samples, which deviated from the results of the R5 standard test, contained oats as their primary ingredient. It is postulated that this outcome can be attributed to the lack of cross-reactivity of the recombinant Fab with oats, potentially explaining the discordance in these instances.

### 3.4 COMPARISON WITH OTHER DEVELOPED METHODS

A direct comparison between the previously developed indirect ELISA, which relies solely on Fab-C (Garcia-Calvo et al., 2023b), and the newly introduced indirect sandwich ELISA featuring the Fab-C and Fab-8E pair, is of significant relevance. On one hand, the indirect approach offers a speed advantage, as it involves one step less than the sandwich method. Conversely, the sandwich methodology has demonstrated greater sensitivity, proving a slightly lower limit of detection, as well as a reduced limit of quantification (17 ppm for the sandwich method as opposed to 20 ppm for the indirect method). This can be explained by considering the number of molecular interactions. While both formats allow for a minimum of a single interaction per antigen molecule (avidity), the sandwich format exhibits stronger paratopic affinity due to the provision of two interactions paratope-epitope, in contrast to the singular paratope in the indirect format. A similar occurrence takes place in the sandwich-R5 method, but because R5 is a complete antibody molecule, a double interaction per antigen occurs, involving up to four paratopes within the two participating antibodies.

In terms of accuracy, the recovery analysis unveiled notable enhancements with the sandwich format, where the retrieved gluten quantities consistently fell within a narrow range of 95 - 105 % compared to the theoretical concentration. In contrast, the indirect ELISA exhibited a broader variability, with recovery results spanning from 76 % to 92 %.

Another noticeable feature in both indirect and sandwich ELISA tests developed is the complete absence of cross-reactivity with oats.

Some authors have found that specific varieties of oats can exhibit cross-reactivity when tested with other detection antibodies such as R5, revealing gluten levels of up to 100 mg/kg in certain oat cultivars (Benoit et al., 2017). The R5 monoclonal antibody's cross-reaction can be attributed to the substantial similarity between certain peptides found in avenins (which are gluten-like proteins in oats) and gluten peptides (Hardy et al., 2015). The absence of this unintentional detection in oat-derived products labeled as gluten-free may indicate the effectiveness of Fab-C and Fab8E-4, making it a valuable tool for evaluating gluten content in such products, especially those containing oat varieties that trigger cross-reactivity with other antibodies available in the market. The proposed methodology has demonstrated a consistent ability to distinguish gluten-free products among the samples examined.

**Table 2.** Results obtained for the detection of gluten in commercial food products using the developed indirect sandwich ELISA and the sandwich monoclonal R5 antibody for result confirmation.

Products	No. of products	Sandwich Fab	Sandwich R5
<b>(A) Products declaring to contain gluten in the labeling (9)</b>			
Pasta	2	+(2)	+(2)
Soup or rice plates	3	+(3)	+(3)
Oats flakes	1	+(1)	+(1)
Cereal bars	2	+(2)	+(2)
Dairy products	1	+(1)	+(1)
<b>(B) Products with “may contain” precautionary labeling (6)</b>			
Cereal	4	+(3)/-(1)	+(3)/-(1)
Flakes	2	-(2)	-(2)
<b>(C) Products that did not declare gluten or that did not specifically warn of the presence of gluten (21)</b>			
Cereal products	9	+(7)/-(2)	+(7)/-(2)
Oat drinks	3	+(1)/-(2)	+(1)/-(2)
Oat flakes	8	+(2)/-(6)	+(5)/-(3)
Meat	1	-(1)	-(1)
<b>(D) Products with “gluten free” labeling and/or certification (14)</b>			
Cookies/cakes	6	-(6)	-(6)
Pasta	1	-(1)	-(1)
Baby food	2	-(2)	-(2)
Corn flakes	1	-(1)	-(1)
Oat flakes	1	-(1)	-(1)
Meat products	3	-(3)	-(3)

Identifying gluten in food products remains a significant hurdle in the field of food science. This challenge arises from the fact that gluten is characterized based on its chemical properties, such as solubility in ethanol. Unlike other food allergens, which are typically represented by one or just a few proteins, gluten encompasses a wide variety of proteins. These proteins share certain common characteristics but can vary significantly in their amino acid sequences (Biesiekierski, 2017). This circumstance renders the quest for a universal gluten detection method exceptionally challenging. Numerous comparative investigations have revealed significant disparities among the commercially available kits in the market when it comes to quantifying gluten (Bugyi et al., 2013). These discrepancies can be attributed to variations in the gluten epitopes to which the antibodies

bind. For instance, the R5 antibody predominantly identifies epitopes that are more prevalent in rye, barley, or triticale compared to wheat, while the G12 monoclonal antibody demonstrates superior detection capabilities for wheat epitopes (Lexhaller et al., 2017). In addition, it has been demonstrated that the quantification of gluten depends on the species and cultivars analyzed, showing significant differences when analyzing the same wheat cultivar with ELISA methods based on polyclonal and monoclonal antibodies such as R5 and G12 (Schopf & Scherf, 2018).

## 4 CONCLUSIONS

The use of recombinant antibodies for gluten detection systems provides a more uniform source for standardized affinity probes compared to polyclonal and hybridoma-secreted monoclonal antibodies. Importantly, these advantages are attained without the necessity of animal involvement. Sandwich ELISA techniques have enhanced the sensitivity and accuracy of single-probe indirect ELISA approaches. Among the four recombinant Fab exhibiting selective reactivity to gluten, Fab-C and Fab8E-4 have proven to be not only compatible but also showing superior detection capabilities when employed in the direct sandwich ELISA method. Utilizing the Fab-C and Fab8E-4 pair in an indirect sandwich ELISA enables the detection of gliadin concentrations, reaching levels as low as 26 ng/mL, equivalent to 10 mg/kg in food samples. Importantly, this method exhibits no cross-reactivity with any of the analyzed non-gluten-containing species. Furthermore, the developed ELISA test complies with legislative requirements for identifying gluten-free products and offers an advantageous feature not present in some commercial tests, as it does not exhibit cross-reactivity with oat proteins.

## 5 BIBLIOGRAPHY

- Amnuaycheewa, P., Niemann, L., Goodman, R. E., Baumert, J. L., & Taylor, S. L. (2022). Challenges in Gluten Analysis: A Comparison of Four Commercial Sandwich ELISA Kits. *Foods*, *11*(5), 706. <https://www.mdpi.com/2304-8158/11/5/706>
- Benoit, L., Masiri, J., Del Blanco, I. A., Meshgi, M., Gendel, S. M., & Samadpour, M. (2017). Assessment of avenins from different oat varieties using R5-based sandwich ELISA. *Journal of Agricultural and Food Chemistry*, *65*(8), 1467-1472.
- Biesiekierski, J. R. (2017). What is gluten? *Journal of gastroenterology and hepatology*, *32*, 78- 81. <https://doi.org/10.1111/jgh.13703>
- Bugyi, Z., Török, K., Hajas, L., Adonyi, Z., Popping, B., & Tömösközi, S. (2013). Comparative study of commercially available gluten ELISA kits using an incurred reference material. *Quality Assurance and Safety of Crops & Foods*, *5*(1), 79-87. <https://doi.org/10.3920/qas2012.0174>
- Cabanillas, B. (2020). Gluten-related disorders: Celiac disease, wheat allergy, and nonceliac gluten sensitivity. *Critical Reviews in Food Science and Nutrition*, *60*(15), 2606-2621. <https://doi.org/10.1080/10408398.2019.1651689>
- Cárdenas-Torres, F. I., Cabrera-Chávez, F., Figueroa-Salcido, O. G., & Ontiveros, N. (2021). Non-Celiac Gluten Sensitivity: An Update. *Medicina (Kaunas)*, *57*(6). <https://doi.org/10.3390/medicina57060526>
- Commission Implementing Regulation (EU) N° 828/2014 of 30 July on the requirements for the provision of information to consumers on the absence or reduced presence of gluten in food OJEU L 228, 31.7.2014, p. 5–8.
- Fairhead, M., & Howarth, M. (2015). Site-specific biotinylation of purified proteins using BirA. *Methods Molecular Biology*, *1266*, 171-184. [https://doi.org/10.1007/978-1-4939-2272-7\\_12](https://doi.org/10.1007/978-1-4939-2272-7_12)
- García-Calvo, E., García-García, A., Madrid, R., Martín, R., & García, T. (2020). From polyclonal sera to recombinant antibodies: a review of immunological detection of gluten in foodstuff. *Foods*, *10*(1), 66. <https://doi.org/10.3390/foods10010066>
- García-Calvo, E., García-García, A., Rodríguez Gómez, S., Farrais, S., Martín, R., & García, T. (2023). Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods. *Current Research in Food Science*, *7*, 100578. <https://doi.org/https://doi.org/10.1016/j.crfs.2023.100578>
- García-Calvo, E., García-García, A., Rodríguez, S., Farrais, S., Martín, R., & García, T. (2022). Construction of a Fab library merging chains from semisynthetic and immune origin, suitable for developing new tools for gluten immunodetection in food. *Foods*, *12*(1), 149. <https://doi.org/10.3390/foods12010149>
- García-Calvo, E., García-García, A., Rodríguez, S., Takkinen, K., Martín, R., & García, T. (2023). Production and Characterization of Novel Fab Generated from Different Phage Display Libraries as Probes for Immunoassays for Gluten Detection in Food. *Foods*, *12*(17), 3274. <https://doi.org/10.3390/foods12173274>

- Hardy, M. Y., Tye-Din, J. A., Stewart, J. A., Schmitz, F., Dudek, N. L., Hanchapola, I., Purcell, A. W., & Anderson, R. P. (2015). Ingestion of oats and barley in patients with celiac disease mobilizes cross-reactive T cells activated by avenin peptides and immuno-dominant hordein peptides. *Journal of Autoimmunity*, *56*, 56-65.
- Inomata, N. (2009). Wheat allergy. *Current Opinion in Allergy and Clinical Immunology*, *9*(3), 238-243. <https://doi.org/10.1097/ACI.0b013e32832aa5bc>
- Itzlinger, A., Branchi, F., Elli, L., & Schumann, M. (2018). Gluten-free diet in celiac disease— forever and for all? *Nutrients*, *10*(11), 1796. <https://doi.org/10.3390/nu10111796>
- Lexhaller, B., Tompos, C., & Scherf, K. A. (2017). Fundamental study on reactivities of gluten protein types from wheat, rye and barley with five sandwich ELISA test kits. *Food Chemistry*, *237*, 320-330. <https://doi.org/https://doi.org/10.1016/j.foodchem.2017.05.121>
- Méndez, E., Vela, C., Immer, U., & Janssen, F. W. (2005). Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *European Journal of Gastroenterology & Hepatology*, *17*(10), 1053- 1063. <https://doi.org/10.1097/00042737-200510000-00008>
- Schopf, M., & Scherf, K. A. (2018). Wheat cultivar and species influence variability of gluten ELISA analyses based on polyclonal and monoclonal antibodies R5 and G12. *Journal of Cereal Science*, *83*, 32-41. <https://doi.org/https://doi.org/10.1016/j.jcs.2018.07.005>
- Shewry, P. (2019). What Is Gluten-Why Is It Special? *Frontiers in Nutrition*, *6*, 101. <https://doi.org/10.3389/fnut.2019.00101>
- Thompson, M., Ellison, S. L., Fajgelj, A., Willetts, P., & Wood, R. (1999). Harmonized guidelines for the use of recovery information in analytical measurement. *Pure and applied chemistry*, *71*(2), 337-348.
- Tye-Din, J. A., Galipeau, H. J., & Agardh, D. (2018). Celiac Disease: A Review of Current Concepts in Pathogenesis, Prevention, and Novel Therapies. *Frontiers in Pediatrics*, *6*, 350. <https://doi.org/10.3389/fped.2018.00350>
- Van Eckert, R., Berghofer, E., Ciclitira, P. J., Chirido, F., Denery-Papini, S., Ellis, H. J., Ferranti, P., Goodwin, P., Immer, U., Mamone, G., Méndez, E., Mothes, T., Novalin, S., Osman, A., Rumbo, M., Stern, M., Thorell, L., Whim, A., & Wieser, H. (2006). Towards a new gliadin reference material— isolation and characterisation. *Journal of Cereal Science*, *43*(3), 331- 341. <https://doi.org/10.1016/j.jcs.2005.12.009>
- Wieser, H., Segura, V., Ruiz-Carnicer, Á., Sousa, C., & Comino, I. (2021). Food Safety and Cross-Contamination of Gluten-Free Products: A Narrative Review. *Nutrients*, *13*(7), 2244. <https://doi.org/10.3390/nu13072244>



## CAPÍTULO 4.5

**Caracterización de las propiedades de repertorios de Fab y de su uso para la selección de sondas con alta afinidad por la gliadina mediante técnicas de secuenciación masiva**

---

**Unveiling the properties of phage display Fab libraries and their use in gliadin-specific probes selection by high-throughput sequencing techniques**

**Artículo en preparación**



## Unveiling the properties of phage display Fab libraries and their use in gliadin-specific probes selection by high-throughput sequencing techniques

**Eduardo GARCIA-CALVO, Aina GARCÍA-GARCÍA\*, Santiago RODRÍGUEZ, Rosario MARTÍN‡, Teresa GARCÍA‡.**

*Departamento de Nutrición y Ciencia de los Alimentos, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040 Madrid, Spain*

‡ *These authors share senior authorship*

\* *Correspondence: ainagarcia@ucm.es; Tel.: +34-91-3943749*

### ABSTRACT

Directed evolution has emerged as a crucial approach for the discovery of novel antibodies. Recent advances in high-throughput DNA sequencing techniques have revolutionized research in different fields such as genomics, epigenomics, and transcriptomics. For instance, these techniques have become essential tools for the characterization of phage antibody libraries and have facilitated the comprehension of the selection processes aimed at generating molecular binders for specific applications. In previous studies, two antibody libraries were constructed to select high-affinity Fabs against gliadin. One of the libraries was entirely derived from mRNA isolated from the peripheral blood of celiac patients (immune library), while the second library was generated through a protein engineering strategy merging the immune light chains derived from celiac patients with heavy chains from semisynthetic origin (semi-immune library). Both libraries underwent phage-mediated selection processes against gliadin. In this study, a compact in-lab sequencing device from Oxford Nanopore was used to comprehensively characterize the composition of the constructed libraries and analyze the performed selection processes against gliadin. A customized analysis pipeline was employed to select high-quality reads, perform sequence analysis by annotating chain distribution and other features, and conduct statistical comparisons between the different selection rounds. Furthermore, an additional analysis was conducted to infer some immunological attributes of the most representative phage variants after the selection process. Sequencing results of the immune library evidenced that many features of the celiac immune response had been successfully transferred to the phage repertoire and to the antibodies derived from it, indicating that these attributes might have been essential in guiding the selection of the best recombinant Fabs against gliadin.

## 1 INTRODUCTION

The paradigm-shifting development of hybridoma technology revolutionized the field of antibody production with the discovery of monoclonal antibodies (Kohler & Milstein, 1975). In 2022, six out of the fifteen best-selling pharmaceuticals were monoclonal antibodies (even more if COVID-19 vaccines are not considered), these figures illustrate the magnitude of the impact that these molecules have made in research and medicine (Urquhart, 2023). Subsequently, the advent of recombinant protein production techniques and the application of directed evolution methodologies further transformed the antibody discovery landscape (Packer & Liu, 2015).

Phage display technology stands as the most prominent methodology for the development of recombinant antibodies in the context of directed evolution of proteins. This method is based on the construction of antibody libraries derived from diverse sources with high variability (Winter, 2019). A characteristic feature of these libraries is that the antibodies are expressed fused to the coat proteins of bacteriophages, which facilitates their affinity selection through successive rounds against the target antigen (biopanning) (Smith, 2019). Thus, antibodies with higher affinities undergo expansion within the phage population, culminating in the isolation of antibodies endowed with the desired characteristics and for which their genetic coding information is also available (Jaroszewicz et al., 2021).

Traditionally, monitoring of the panning procedure has been performed by colony picking and Sanger sequencing since the 1990s. Although this method could be automated for the screening of thousands of different colonies, it would only cover a small amount of the clones implied in the process. This problem could be solved by using next generation sequencing (NGS) techniques (Glanville et al., 2015). The in-depth analysis of phage libraries by NGS has proven to be an exceptionally effective approach for clone identification, including low-frequency rare variants that may not amplify adequately within the phage population (D'Angelo et al., 2014). Moreover, NGS has facilitated the identification and reconstruction of complementarity determining region (CDR) combinations of antibody fragments from phage selection outputs (Barreto et al., 2019). Beyond antibody libraries, this transformative technique extends its applicability to the study of peptide libraries, essential for epitope mapping of antibodies (Matochko & Derda, 2015).

Oxford Nanopore Technology offers many advantages that have revolutionized the field of DNA sequencing and genomic analysis. Firstly, the portable and compact sequencing devices provide unprecedented flexibility, enabling real-time, *in situ*, and rapid DNA sequencing in diverse environments, including field research and clinical settings. Secondly, the sequencing technology can read exceptionally long DNA fragments, allowing for improved genome assembly and structural variant detection. This feature have been particularly valuable for

studying complex genomes or uncovering elusive genetic mutations. Additionally, Nanopore's procedures eliminates the need for time-consuming library preparation, making it a faster and cost-effective option for researchers. Furthermore, it allows both DNA and RNA analysis, broadening its applicability in different research areas, from genomics to transcriptomics. Overall, the Oxford Nanopore sequencing technology represents a powerful tools for more comprehensive and efficient nucleic acid analysis (Chen et al., 2023).

Antibodies are lymphocyte-secreted molecules with the ability to get attached to specific antigens and with several cell signaling functions. Their antigen-counteracting capacity derives from their extraordinary diversity, for instance humans may be capable of producing as many as  $10^{18}$  distinct antibodies (Joyce et al., 2020). This huge antibody repertoire that can be found in mammals derives from diverse genetic events like V(D)J recombination (Roth, 2014), somatic hypermutation and class switch recombination (Hwang et al., 2015).

In humans, the genes encoding antibodies are located on different locus found on three chromosomes (chr): heavy chain genes found on chr14 (14q32.33) and kappa and lambda light chain genes found on chr2 (2p11.2) and chr22 (22q11.2), respectively. There have been found 50 germinal sequences for VH (heavy chain variable region) grouped in 7 families, 40 for Vk (kappa chain variable region) grouped in 6 families and 30 for Vλ grouped in 10 families. The three major D gene fragments are located in the heavy chain locus. For the J genes have been described 7 genes for the heavy chain locus, 9 for lambda and 6 for the kappa locus. Moreover, IgH (immunoglobulin heavy chains) have five major isotypes (IgM, IgD, IgG, IgE, IgA), with four isotypes for IgG (IgG1, IgG2, IgG3 and IgG4) and two for IgA (IgA1 and IgA2) (Barbas, 2001).

Several diseases related to the ingestion of gluten, the ethanol-soluble protein fraction present in the endosperm of some cereals like wheat, barley and rye, have been described. The most studied is celiac disease, an autoimmune condition with a reported prevalence of 0.5 - 1 % of the general population, although it might be underestimated due to undiagnosed or misdiagnosed milder cases (Caio et al., 2019). This disease is characterized by the production of autoantibodies against the tissue transglutaminase (tTG) in the context of a loosen intestinal barrier (Sapone et al., 2012). Despite ongoing research into numerous therapeutic alternatives for celiac patients, the gluten-free diet remains the most effective treatment, not only to halt intestinal damage but also to promote tissue healing (Itzlinger et al., 2018).

In this work, two previously constructed antibody libraries which were subjected to a biopanning process for the selection of gliadin-binding antibodies were characterized by sequencing with Oxford Nanopore Technologies. Comparison of the changes in the composition of the different antibody chains over the consecutive rounds of panning provided an overall picture of the performance of the gliadin-guided selection procedure.

## 2 MATERIAL AND METHODS

### 2.1 Source material: DNA isolation and purification

In our previous works, high affinity antibody fragments against gliadin were developed as novel probes for immunodetection systems to ensure the absence of gluten in foods intended for people suffering from gluten-related disorders. These antibodies consisted of antigen-binding regions (Fab) generated by phage display technology. For this purpose, two Fab-libraries were constructed into the pComb3X vector (Barbas, 2001): an immune library derived from the retrotranscribed RNA of the peripheral blood lymphocytes of celiac patients (Garcia-Calvo et al., 2023) and a second library obtained by merging semi-synthetic heavy chains (derived from a pre-existent single domain antibody) with the immune light chain repertoire, that was named semi-immune library (Garcia-Calvo et al., 2022). Both libraries were used in biopanning processes to select high-affinity Fab against gliadin. After each round of selection, the eluted phages were used to infect *Escherichia coli* XL1-Blue (*recA1 endA1 gyrA96 thi- 1 hsdR17 supE44 relA1 lac [F proAB lacIqΔM15 Tn10 (Tetr)]*) (Agilent®, Santa Clara, CA, USA, ref #200150). Overnight cultures were grown in Super-Broth medium (SB: 30 g/L tryptone, 20 g/L yeast extract, 10 g/L MOPS, pH 7). After centrifugation at 4000 xg for 20 min at room temperature, the supernatant was subjected to phage precipitation to continue the selection process and phagemid DNA was isolated from the cell pellets by a midi-prep procedure (PureLink™ HiPure plasmid Midiprep Kit Invitrogen™, Thermo-Fisher®, Waltham, MA, USA, ref #K210004) following manufacturers' protocol.

In this study, the unsearched libraries as well as the sub-libraries obtained after each round of panning were characterized by NGS with Oxford Nanopore technology. For this purpose, DNA in a double strand and linear format is required. Plasmids were digested with *NheI* for 1 h at 37 °C (New England Biolabs® Ipswich, MA, USA ref #R0156), since this enzyme presented a single restriction site with an appropriate distance from the antibody construct, to avoid interfering with the sequencing results.

The DNA was purified by AMPure XP reagent (Beckman Coulter®, Brea, CA, USA, ref #A63880) following manufacturers' instructions for the protocol based on the use of magnetic beads. Finally, DNA was quantified in a Qubit® Fluorometer (Invitrogen, Thermo-Fisher®), and its quality was measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Montchanin, DE, USA).

### 2.2 DNA preparation for sequencing protocol

A total of nine samples were barcoded with the PCR barcoding expansion 1-12 kit (Oxford Nanopore®, Oxford, UK, ref #EXP-PBC001) using a magnetic bead-based protocol. Barcodes (BC) were assigned as follows: BC01-06 for the identification of the un-panned immune library (BC01)

and its four panning rounds (BC02-05) and BC06-09 for the identification of the un-panned semi-immune (BC06) and its three panning rounds (BC07-09).

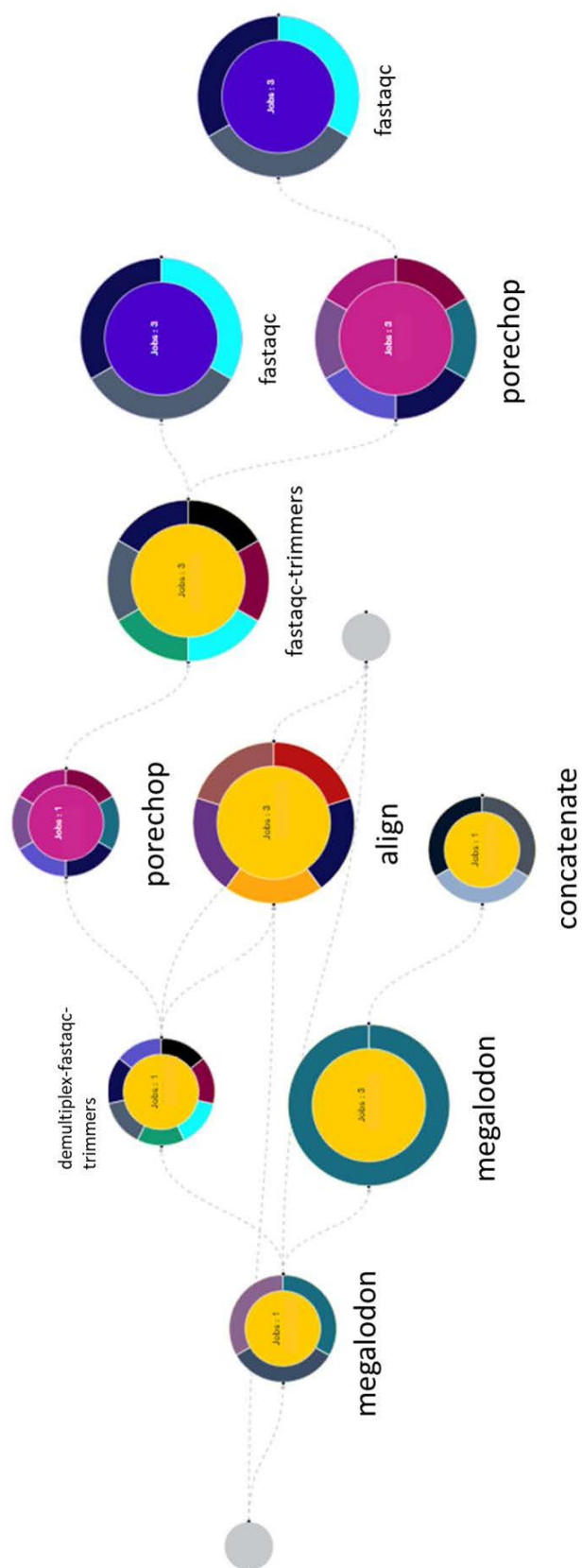
Finally, the magnetic bead system was used to subject the barcoded sample to end-repair and dA-tailing using the NEBNext End Repair/dA-tailing module, and subsequent ligation of Nanopore sequencing adapters onto the prepared DNA ends (Oxford Nanopore®, ref #SQK-LSK110). Samples were loaded into the Nanopore's flow-cell (Oxford Nanopore®, ref #R10.4.1) and NGS sequencing was performed on a MinION® device collecting sequencing data for 30 h.

### 2.3 Primary sequence analysis

NGS sequence analysis usually follows a workflow with two main stages. The primary analysis is based on the organization of the raw sequences generated by the sequencer and the elimination of poor-quality sequences that could generate errors and introduce noise into the subsequent analysis (Han et al., 2018).

MinKNOW software (Oxford Nanopore®) transformed the registered electrical signals into raw FAST5 files that were used as base for the primary analysis which was carried out using a workflow encompassing three different steps (Figure 1). In the first step, megalodon tool (v1.9.5, <https://github.com/nanoporetech/megalodon>) was used to extract high accuracy modified base calls and sequence variant calls from the raw reads by anchoring the information of the basecalling neural network output to the reference genome of *Homo sapiens* (Genome Reference Consortium Human Build 38, NCBI RefSeq assembly GCF\_000001405.26). A sub-pipeline appropriately divided the work and merged the results, yielding a single FASTQ file containing the basecalled reads. In the second step, this file was demultiplexed into new FASTQ files for as many barcodes as the analysis included, in order to cluster the sequences according to their barcodes. Subsequently, the files provided were submitted to quality control using FastQC and analyzed with the Porechop tool (<https://github.com/rrwick/Porechop>) to trim off the adapters found at the ends of the reads. The trimmed sequences were again subjected to quality control, before being analyzed with Trimmomatic, a tool which can identify low quality reads. Lastly, the trimmed sequences were subjected to a final quality control using FastQC. In the third step, a sorted BAM file was obtained with the alignment of the sequences against the reference genome of *Homo sapiens* using minimap2 tool (<https://github.com/lh3/minimap2>).

The tools used in the secondary analysis required sequence packages containing less than one million sequences. For this purpose, the tool LUMC/fastq filter (<https://github.com/LUMC/fastq-filter>) was used to optimize the FASTQ packages filtered by read size between 4-6 kb and sequence quality (phred scores higher than 15), removing low-quality sequences from the packages.



**Figure 1.** Schematic representation of main workflow used for the primary analysis of the sequences read with the MinION device.

## 2.4 Secondary analysis of antibody sequences

The filtered FASTQ data was uploaded to the online tool IMGT/HighV-Quest that orchestrated the analysis of the Fab sequences (Brochet et al., 2008). Each identified Fab was characterized by 1) IMGT unique numbering for the V-domain; 2) closest variable (V), diversity (D) and junction (J) genes and their respective alleles; 3) junction analysis and 3) mutation and amino-acid changes resulting from somatic hypermutations by comparison with the IMGT/V-QUEST reference directories set (Lefranc & Lefranc, 2020).

Subsequently, IMGT/StatClonotype tool was used to perform pairwise comparison of datasets from the IMGT/HighV-QUEST output. In this study, the different datasets corresponded with the consecutive rounds of selection of the Fab phage libraries. From this analysis V(D)J gene associations by multivariant phylogenetic heatmap plots and CDRs analysis by length and amino-acid composition were obtained (Aouinti et al., 2016).

Minimap2 alignments were used to identify classes and subclasses of the antibody sequences that could not be analyzed by IMGT tools. The number of sequences that aligned at the positions of the genome where the genes of each subclass are located was counted using the coverage function of the bedtools package (<https://bedtools.readthedocs.io/en/latest/content/tools/coverage.html>).

# 3 RESULTS AND DISCUSSION

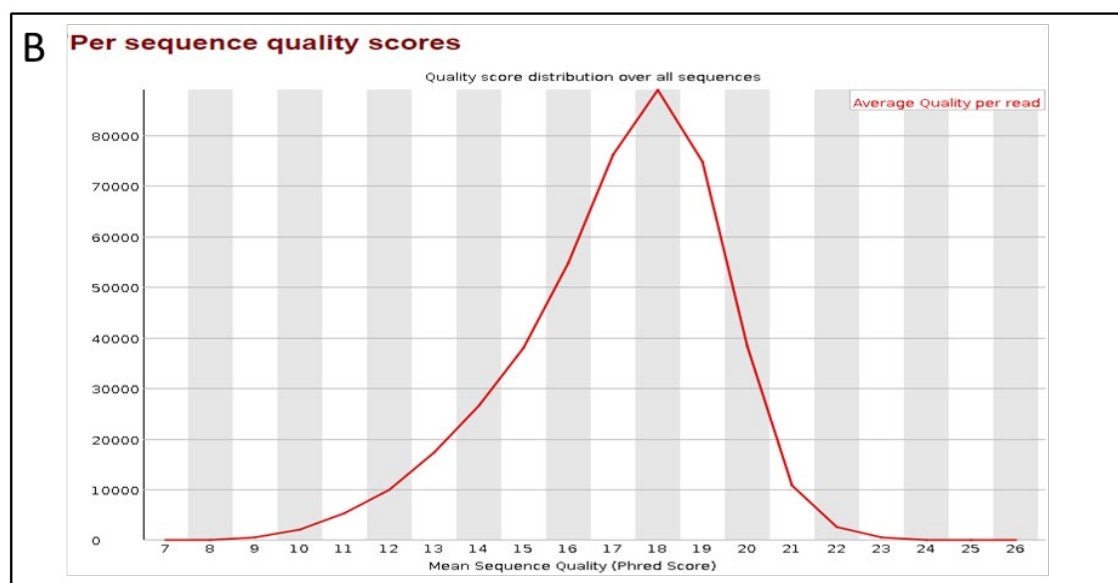
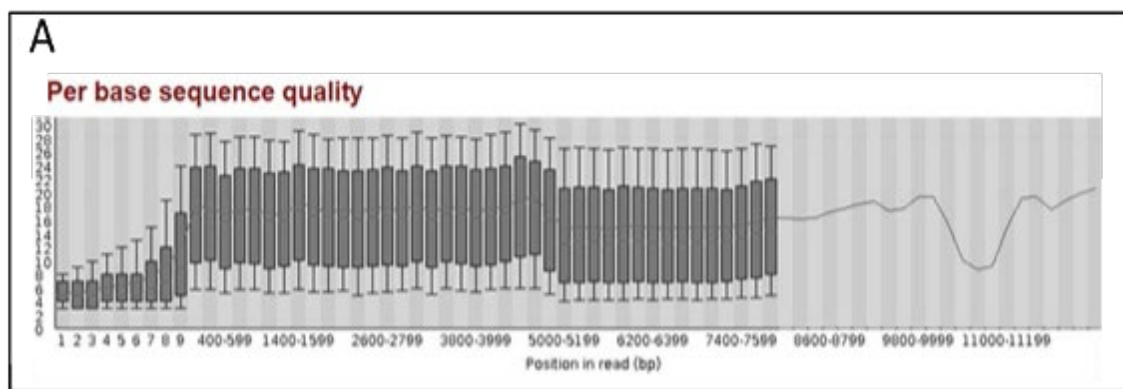
## 3.1 DNA preparation

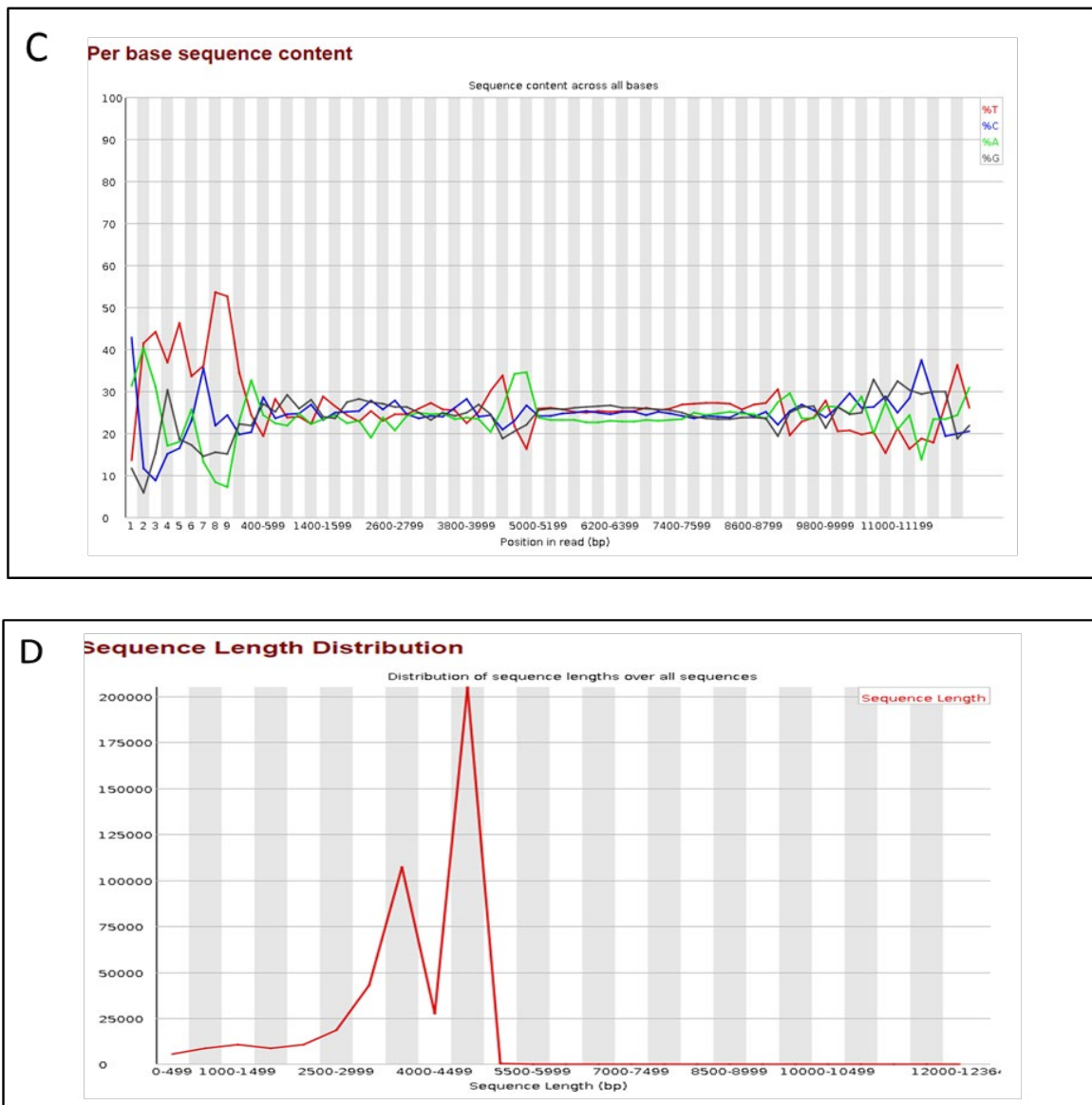
Phagemid DNA isolated from all the library and sub-library samples resulted in high yields (DNA concentrations above 390 ng/ $\mu$ L) with appropriate purity (OD 260/280 > 1.8 and OD 230/260 > 2). Therefore, the obtained DNA samples showed suitable parameters to be subjected to the sequencing process.

## 3.2 Primary sequences analysis

After the transformation of the recorded electrical signal into nucleotide data (basecalling algorithm), a total of 4.68 million of reads were collected in a FAST5 format. Through the primary analysis workflow, the reads were transformed into demultiplexed and trimmed FASTQ sequences grouped by barcode. An average of 0.5 million sequences per barcode were collected, ranging in length from 120 to 12,400 bases, something common in this methodology (Wang et al., 2021). This broad spectrum of reads underscored the imperative need to implement a filtering process. The quality of sequences was assessed and showed a relatively uniform distribution across the reads (Figure 2A). The analysis of quality scores distribution per sequence revealed that the majority of sequences exhibited a phred-score of 18 (Figure 2B), demonstrating an improvement in quality over those reported by previous similar studies were most of the sequences presented a phred-score between 10 and 15 (Lowden & Henry,

2018). The nucleotide content analysis per base sequence revealed that the proportion of each of the four bases was remain relatively constant over the lengths between 0.4 and 9 kb of the reads, except for an enrichment in T and A residues in the 4-5 kb range that are found also in the original vector (Figure 2C). Notably, this bias in nucleotide composition did not affect the antibody coding segments. The examination of the sequence length distribution revealed the presence of two prominent populations. The larger population, obtained in lengths ranging from 4.5 to 5.5 kb, corresponded to the linearized phagemid of the library (Figure 2D). On the contrary, the smaller population centered on 4 kb, could potentially denote truncated plasmids which might result from the existence of unforeseen restriction points. Based on the quality data described above, sequence filtering was performed obtaining a collection of cleansed FASTQ sequences of 4-6 kb in length and phred-scores greater than 15. Remarkably, the majority of the antibody coding sequences were found to meet these parameters and were therefore included in the secondary analysis; other sequences with varying lengths were considered noise or outliers and were excluded from the analysis.





**Figure 2.** Quality metrics for Oxford Nanopore sequencing of the home-made phage-displayed libraries. (A) Distribution of quality scores according to read position, in the boxplot in black lines represent medians, the box shows the lower and upper boundaries of the first and third quartiles, respectively, and the line shows the mean quality score. (B) Mean per-read quality score diagram. (C) Per base sequence content plot. (D) Read length distribution. All the analysis were performed using FastQC toolkit. The data shown correspond to one of the multiplexed samples, however the data recorded for all the samples presented similar profiles.

### 3.3 Secondary analysis of antibody sequences and biological interpretation of the results

After completing the primary analysis and acquiring a set of adequate sequences in terms of quality and length, subsequent analysis delved into the antibody composition of the constructed libraries.

Firstly, all the sequences were aligned to the V(D)J germlines using IMGT/HighV- Quest. The immune library) data were analyzed for both the heavy and light chains, whereas for the semi-immune library sequences only the light chains were analysed, since in this case all heavy chains derived from the same dAb (VH3-23 subfamily) and their study was not relevant.

The results obtained in this study have led to the conclusion that we successfully captured 31 out of the 50 VH subfamilies documented in the human population, 16 Vk subfamilies within the immune repertoire, and 22 V $\kappa$  subfamilies within the semi-immune repertoire, from the total of 40 V $\kappa$  subfamilies known in humans (Barbas et al., 2001). It should be noted that not all antibody chains were included in our analysis, primarily due to the PCR amplification conditions. Our primary objective was to focus on including antibody chains specifically involved in the immune response associated with celiac disease.

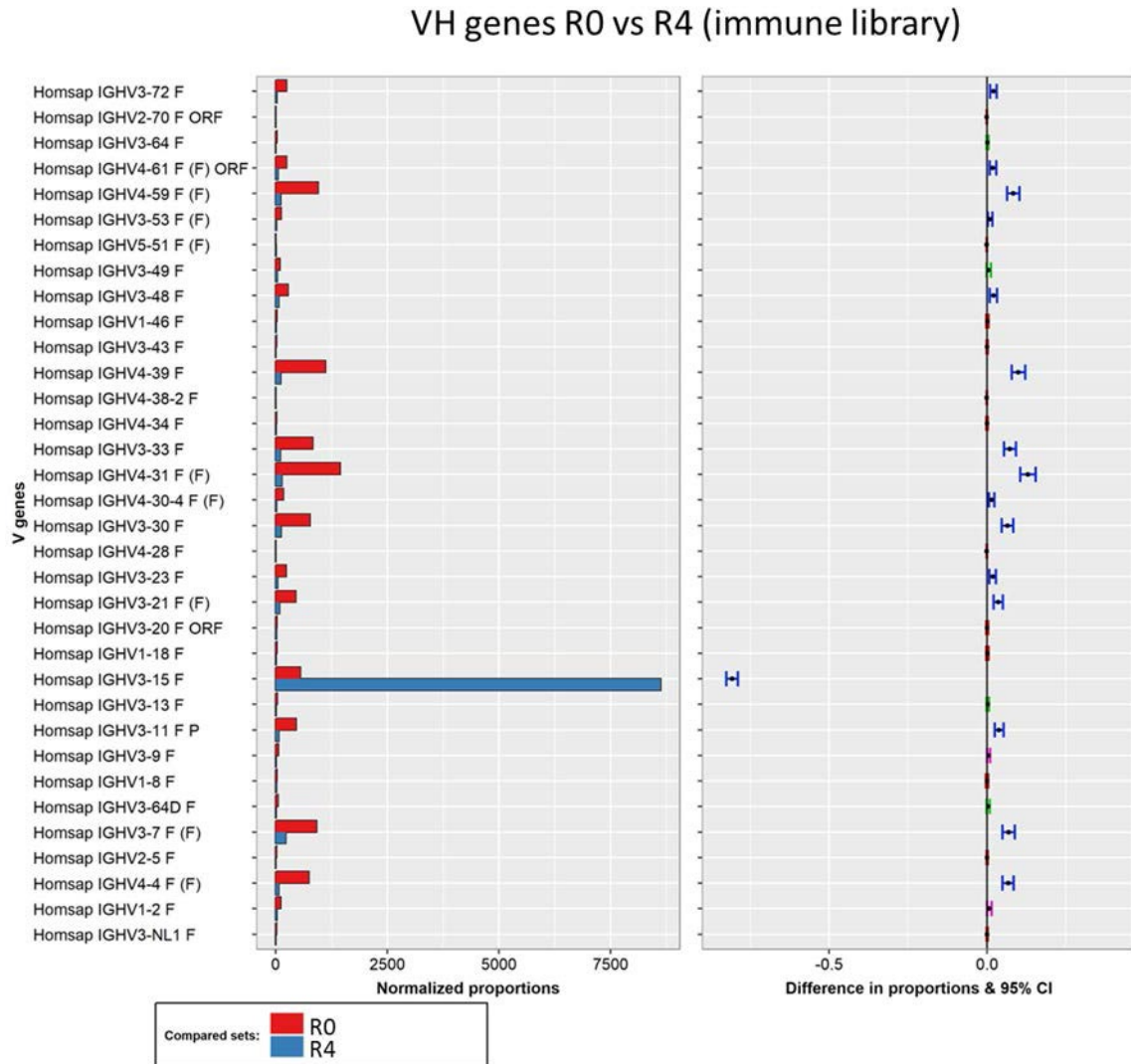
Moreover, it is worth noting that our decision not to include all antibody chains aligns with findings from previous studies on repertoire construction for phage display. These studies have indicated that including many minor subfamilies in repertoires does not necessarily add practical value, as such subfamilies are often compromised and do not significantly contribute to the functionality of the repertoires. Therefore, a more strategic approach involves concentrating on the amplification of subfamilies known to exhibit reactivity, which enhances the overall utility of the repertoire in our research context.

#### 3.3.1 Analysis of the variable domains

Regarding VH composition of the immune, two different profiles can be observed during the panning rounds. First, the profile of the 31 VH subfamilies between the unpanned library and the phage population recovered from the first round of selection remained practically unaltered. Then a slight increment of VH3-15 subfamily was observed after the second round, followed by a substantial dominance of this subfamily in rounds three and four. Interestingly, this subfamily Fab exhibited the highest affinity to the antigen.

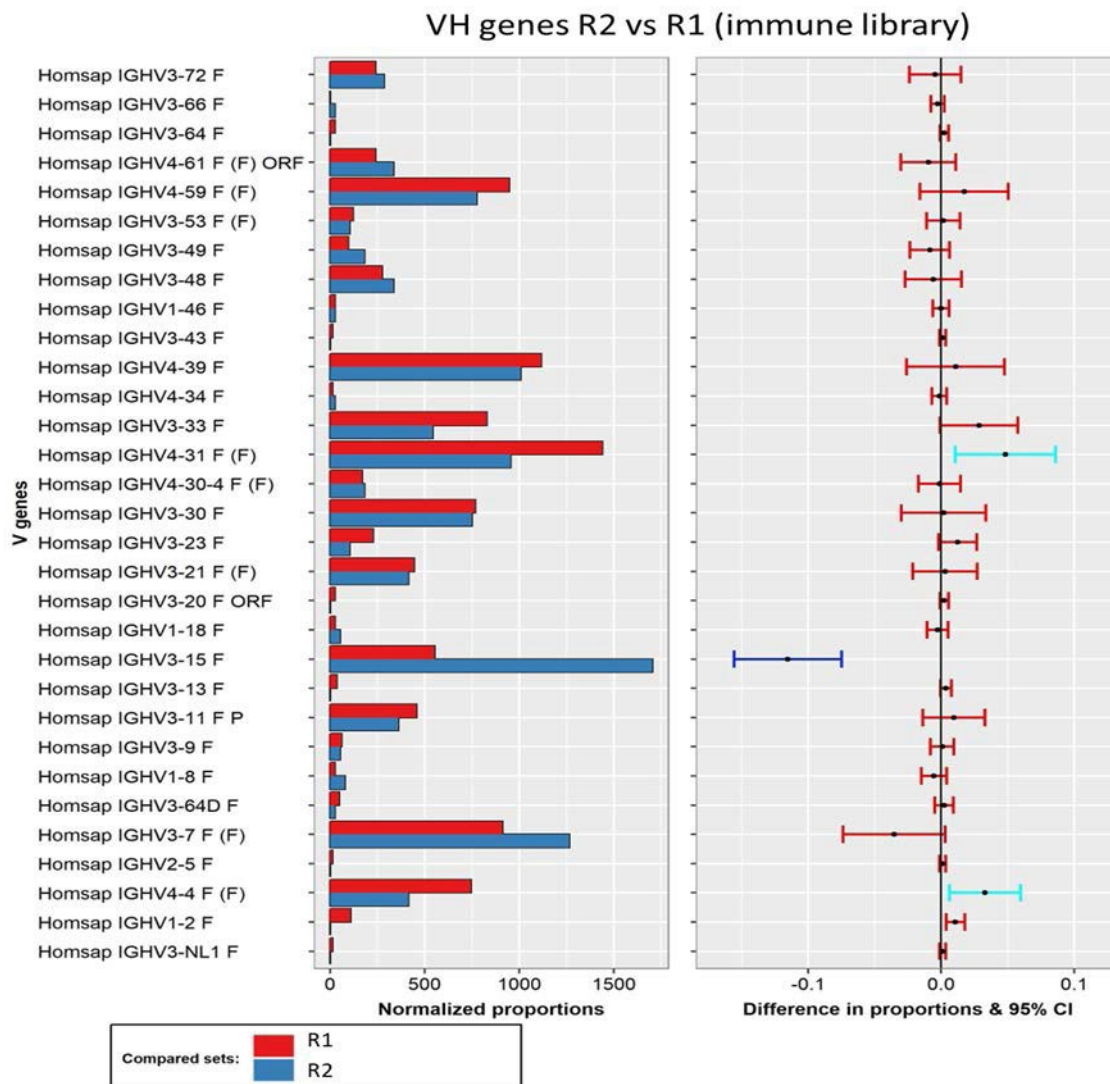
In order to test this hypothesis, we conducted pairwise comparisons of the heavy chain composition between the different rounds using IMGT/StatClonotype. To study the selection process, the unpanned immune library and its fourth round of selection were compared to observe any changes in distribution of the V gene. As expected, a significant increase (statistically significant

for all p-values) in the VH3-15 subfamily, which evolved from a minority to a predominant family (Figure 3). Although less pronounced, there was also a statistically significant decrease in the presence of 15 subfamilies, with the remaining subfamilies showing no notable differences.



**Figure 3.** Pairwise analysis of VH genes composition of the unpanned immune library (red) and its fourth round of selection (blue). Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each VH subfamily. The color of the intervals depicts test interpretation: non-significant (red), significant for all p-values (blue), significant for at least 2 tests (pink) and significant for only Benjamini-Hochberg test (green).

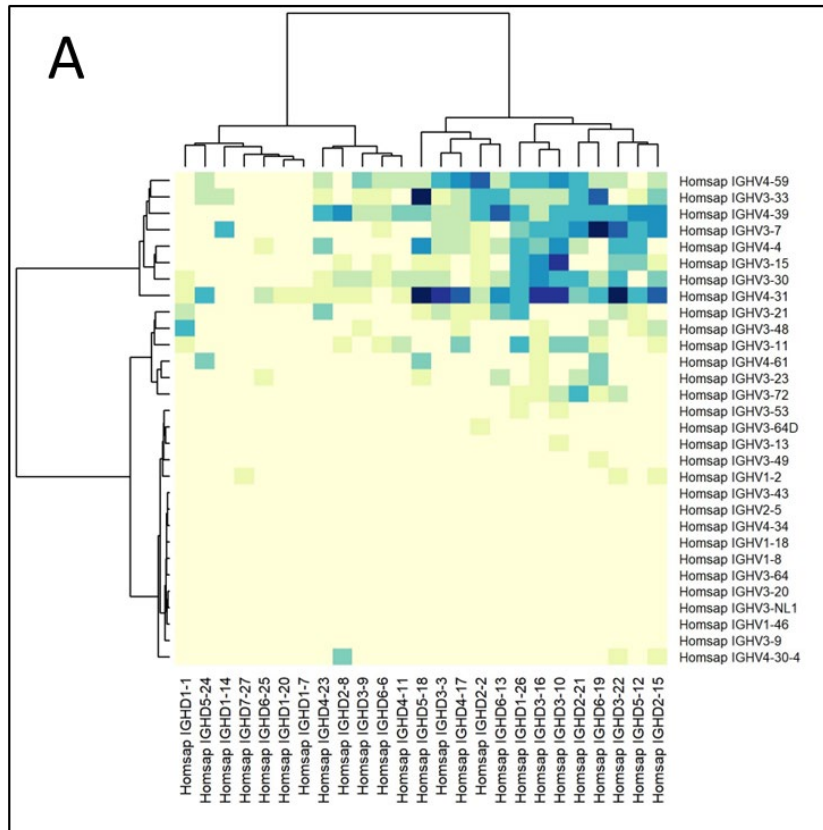
Additionally, it was also determined through pairwise analysis if the initial increase of the VH3-15 subfamily between rounds 1 and 2 was significant (Figure 4). The test confirmed that VH3-15 was the subfamily that exhibited the most substantial increase, and this increase was statistically significant for all p-values analyzed. Additionally, two VH4 subfamilies showed a significant increase, although this significance was observed for only one of the p-values. Although VH4-4 and VH4-31 subfamilies also presented a significant increase between the studied rounds, they were not highly represented in the last round.

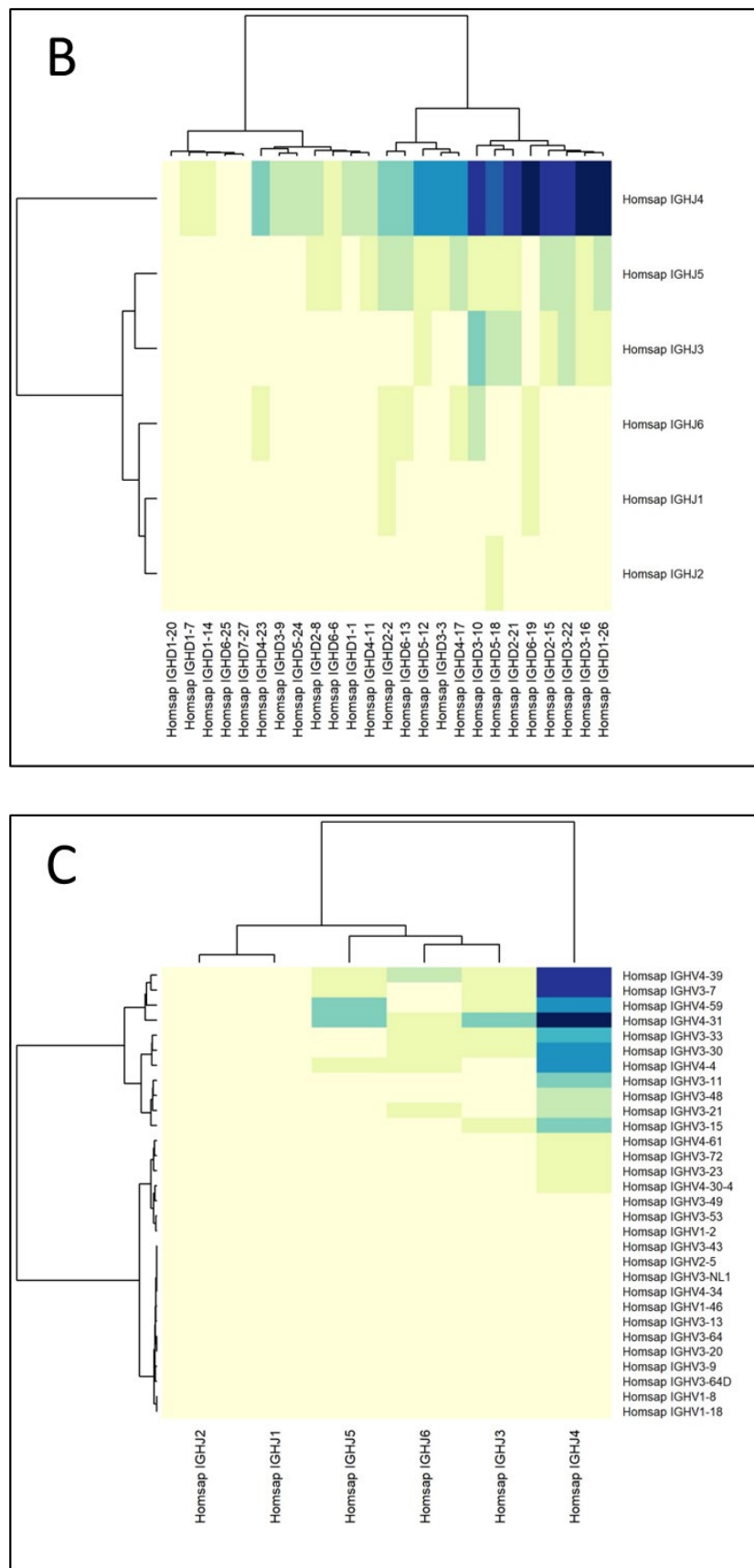


**Figure 4.** Pairwise analysis of VH genes composition of round 1 (red) and round 2 (blue) of selection of the immune library. Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each VH subfamily. The color of the intervals depicts test interpretation: non-significant (red), significant for all p-values (dark blue), significant for one test (light blue).

When examining the selection process of the immune library, a striking observation was made. Instead of the anticipated gradual increase in selected binders across the rounds, two punctual selection events were identified. The first event was characterized by a slight yet significant increase in VH3-15 Fab during the second round. This was followed by their substantial expansion in rounds 3 and 4. Notably, this event coincided with an increase in the stringency of the selection, achieved by incorporating additional washing steps. Following the expansion observed in round 3, the VH3-15 variants surpassed the abundance of other variants to such an extent that the likelihood of other selections became highly improbable. Although there were several notable changes in the composition of the light chain families throughout the depicted process, it appears that the role of the light chain as a selection agent may not be decisive.

The relationship between the VDJ genes of the heavy chains of the immune library was depicted by generating clustered image maps facing the identified VD, DJ and VJ genes (Figure 5). Considering that the rearrangement of the chains does not change through the selection rounds, the analysis was performed using the data from the unpanned library. The main gene associations found were as follows: VH3-33 with DH5-18 and JH4. VH4-31 with DH5-18, DH3-22 and JH4, VH3-7 with DH6-19 and JH4, VH3-15 with DH3-10, and JH4 with DH6-19, DH3-16, and DH1-26. Notably, the usage of the joining gene JH4 was very extended, a normal pattern in the human species were JH genes presents an asymmetric usage of more than the 40 % for JH4. This situation is derived from rearrangements events, in which conserved nonamers, heptamers and 23-bp spacers are involved (Shi et al., 2020).





**Figure 5.** Clustered imaged maps depicting the relation between immunoglobulins VDJ genes of the immune library. (A) VD genes; (B) DJ genes and (C) VJ genes. Darker colors in heatmaps indicate the strongest associations between genes.

Another crucial aspect under investigation pertains to the identification of specific characteristics associated with the celiac response within the immune library. Notably, it was observed that the selected chain variants (VH3-15) constituted approximately 3.5 % of the total reads during the initial round. This prevalence aligns with previous reports on directly sequenced biological samples obtained from healthy controls (Bowers et al., 2014) However, it should be highlighted that the observed expression level was notably higher than that reported in the human Naïve phage display library (Kim et al., 2017). These variants were selected until the total dominance of the population in last round (86.2 %).

Previous studies have also emphasized the preferential selection of VH3-15-IgA variants as highly reactive antibodies against gliadin. However, it is worth noting that the observed VH/VL pairing was not consistently maintained, as this phenomenon appeared to occur stochastically during the construction of the library (Steinsbø et al., 2014).

An additional analysis was performed focused on the gliadin-selected Fab variants from the immune library: the VH3-15 chains of these Fab were alignment against the germline gene (IMGT accession number M99406), finding several mismatches that could constitute biologically significant mutations in the gliadin-binding Fab (Table 1).

**Table 1.** Identification of mismatch mutations heavily repeated on the selected Fab variants from the immune library after the biopanning process against gliadin.

Position	Germline	Library	% of reads with the mismatch
13	G	C	92
92	A	G	91
98	G	A	95
99	G	T	93
150	T	G	96
176	C	A	91
188	C	G	93
193	C	T	95
252	A	G	96

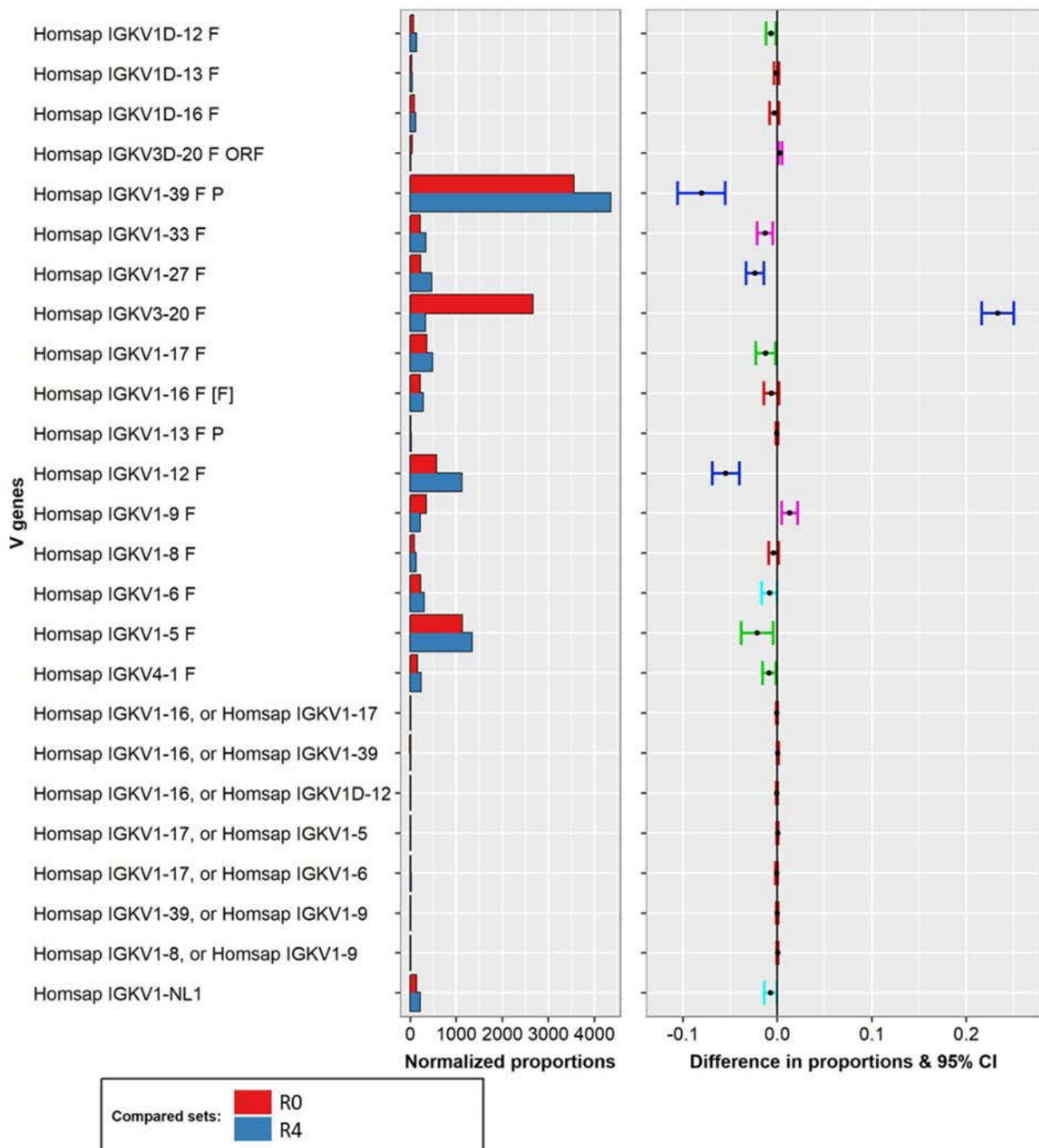
Upon analyzing the selected variants, a comprehensive examination revealed the presence of somatic mutations within the VH regions. Specifically, an average of  $9 \pm 2$  mismatched bases were identified across the VH3-15 variants from the final selection round. Remarkably, these mismatches were distributed extensively throughout the reads, indicating a strong likelihood of SMH as the underlying cause. Furthermore, the observed mutation rates were consistent with those found in other pertinent studies. In our study, we also observed an average of 9 mutations

in IgA serum antibodies reactive to gliadin, while other antibodies exhibited an average of 15 mutations compared to the germline (Steinsbø et al., 2014). This finding suggests that the heavy chain coding genes captured and subsequently selected within the library may originate from short-lived plasmatic cells that were activated through an extrafollicular immune response (Ho et al., 1986). Furthermore, it is important to note that the selection of celiac patients who had not yet initiated a gluten-free diet logically led to the inclusion of genes derived from short-lived plasma cells in the library. In the absence of gluten-related response-triggering peptides, it is highly probable that these short-lived plasma cells would have been inhibited and the patients to exhibit symptoms. (Sharma et al., 2020).

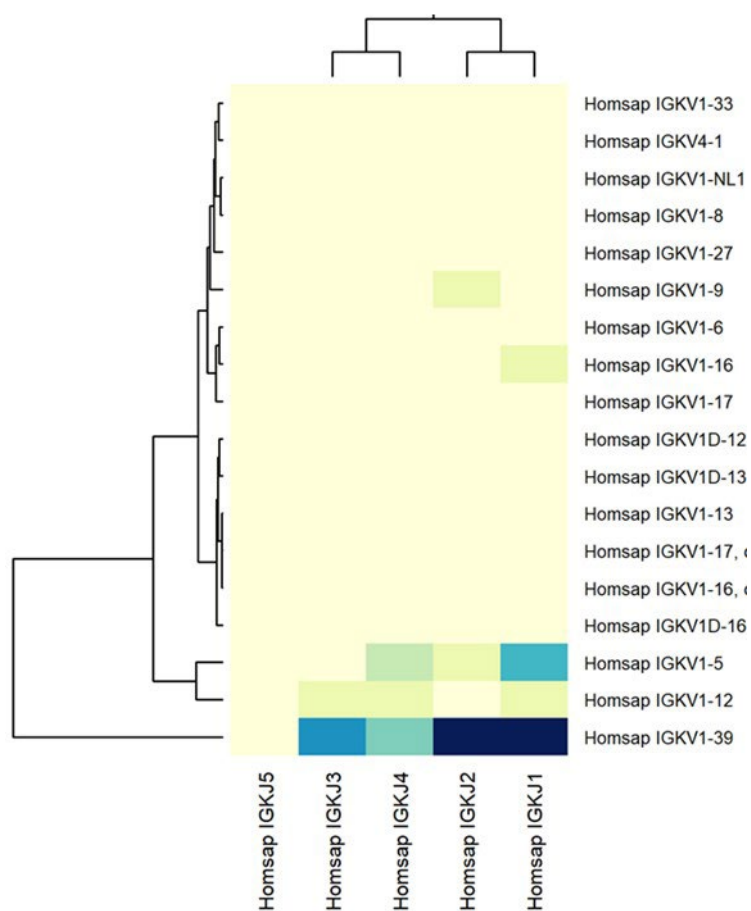
In addition, the development of B cells in the extrafollicular context is influenced by T cells. Previous reports have indicated that T cells present at the T-B border within lymph nodes, distinct in phenotype from germinal center T cells, play a crucial role in the initiation of B-cell priming for the generation of extrafollicular antibody responses (Lee et al., 2011). This lower mutation occurrence in gliadin reactive antibodies it was also found in transglutaminase 2 (TG2) autoantibodies suggesting that the presence of gluten antigens or gluten-specific T cells may potentially contribute to the observed low mutation rates in the specific IgA antibodies (Steinsbø et al., 2014).

When analyzing the evolution of the immune library kappa-light chains throughout the selection process (unpanned library against round 4), four statistically significant changes were found (Figure 6): the increase of the V $\kappa$ 1-39, V $\kappa$ 1-27 and V $\kappa$ 1-12 and a decrease of V $\kappa$ 3-20 subfamilies. In this sense, V $\kappa$ 1-39 has been identified as the most expressed light chain gene in other studies with naïve libraries (Kim et al., 2017). In the present study, this light chain was found to be the dominant subfamily in the unpanned immune library and also between the selected phages after the panning process (the most significant increase within the light chains) (Figure 6).

Regarding the re-arrangement process for the kappa-light chains, several VJ association were found between the most expressed V $\kappa$  gene (V $\kappa$ 1-39) with J $\kappa$ -1 to 4, with a stronger occurrence of J $\kappa$ -1 and J $\kappa$ -2 associations (Figure 7). Although there were several notable changes in the composition of light chain families throughout the immune repertoire selection process, it seems that the role of the light chain as a selection agent has not proven decisive.

V $\kappa$  genes R0 vs R4 (immune library)

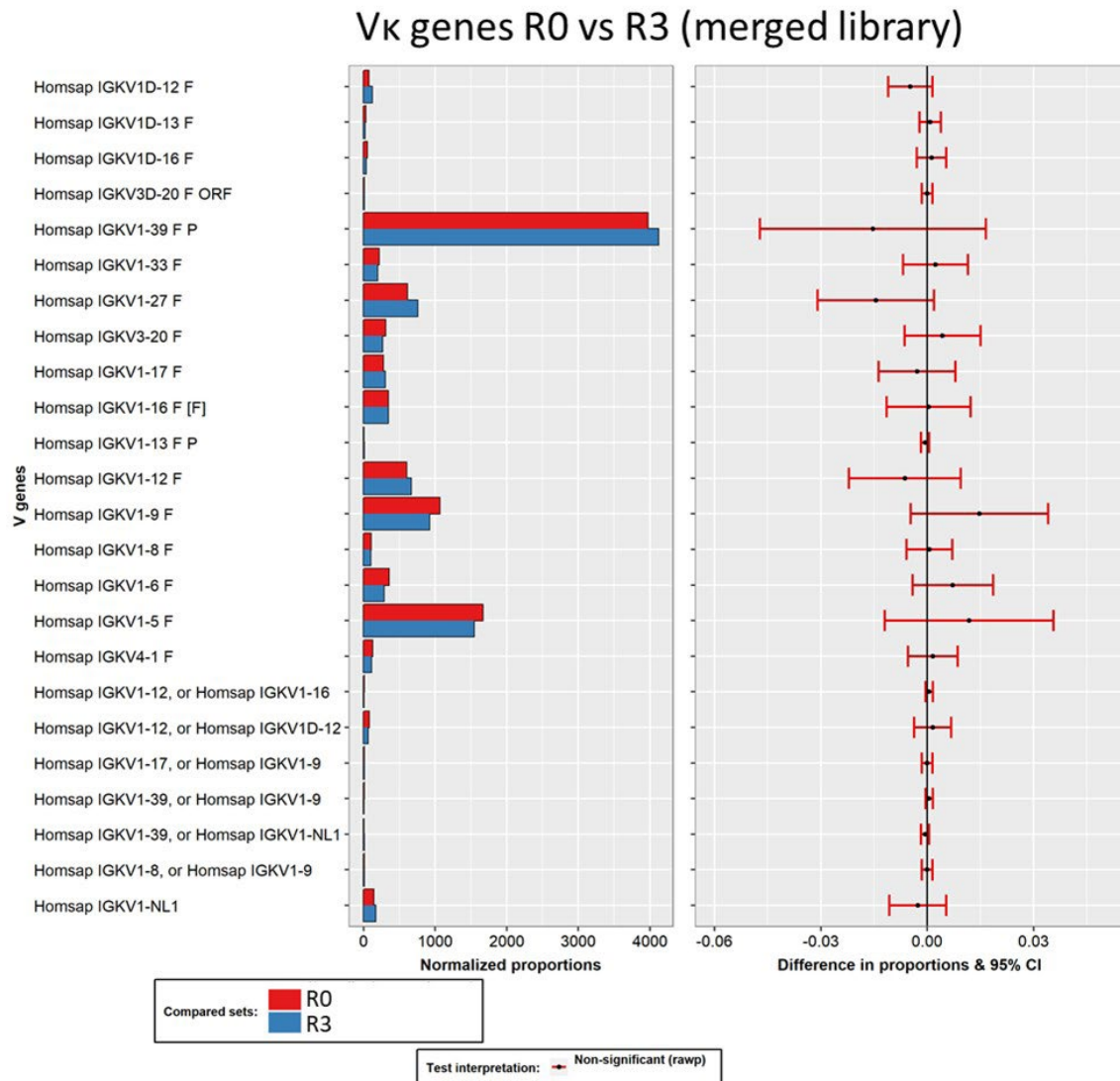
**Figure 6.** Pairwise analysis of V $\kappa$  genes composition of the unpanned immune library (red) and its fourth round of selection (blue). Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each VH subfamily. The color of the intervals depicts test interpretation: non-significant (red), significant for all p-values (dark blue), significant for one test (light blue), and significant only Benjamini-Hochberg test (green).



**Figure 7.** Clustered imaged maps depicting the relation between VJ ( $\kappa$  chains) genes of the immune library. Darker colors in heatmap indicate the strongest associations between genes.

In summary, the immune library successfully captured and preserved characteristics associated with celiac disease, as identified by multiple researchers. These features were utilized to generate novel high-affinity Fab targeting gluten. The resulting Fab possess unique attributes, incorporating key elements of the humoral response observed in celiac patients. Consequently, they serve as valuable probes for gluten immune-detection in food products. This innovative approach offers distinct advantages over the traditional monoclonal antibody-based methods that rely on animal immunization, which are currently available in the market.

Similar analyses were performed with the data obtained from the sequencing of the library which was constructed by merging immune and semi-synthetic chains, but mainly focus on the light chains. In this case, no significant changes in the composition of the light chains were found through the selection process (Figure 8). Since this library was built based on a preexistent dAb (previously isolated for its affinity against gliadin) which was present in all the sequences, such result demonstrated that the immune light chains were not a determinant selector of gliadin-binding Fab.

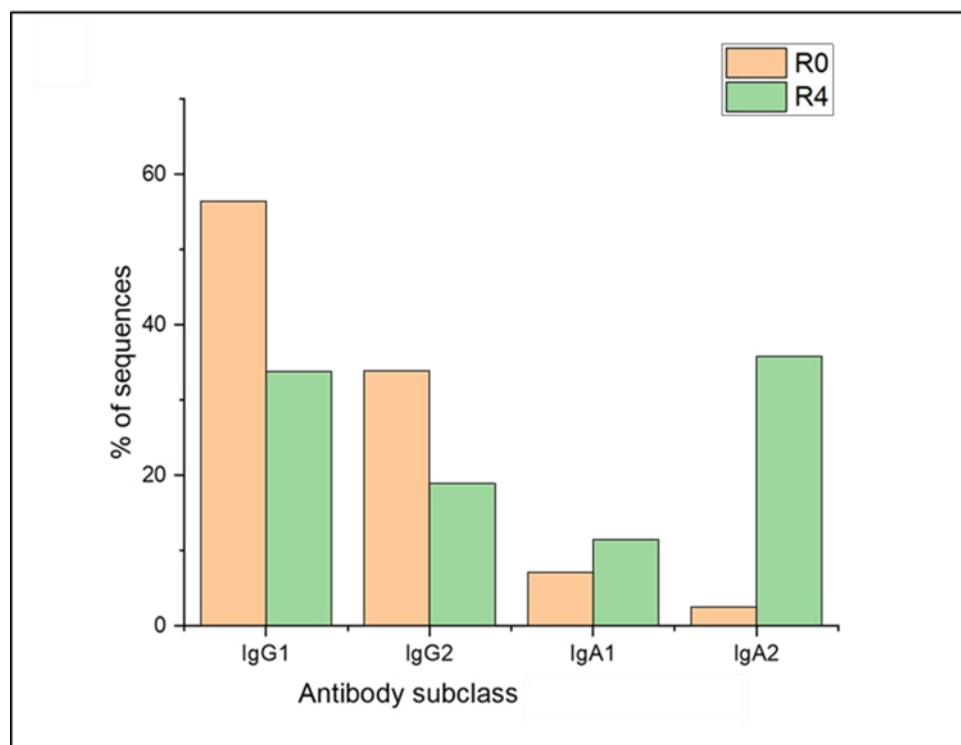


**Figure 8.** Pairwise analysis of V<sub>k</sub> genes composition of the unpanned semi-immune library (red) and its third round of selection (blue). Normalized proportions (left chart) and intervals at 95 % confidence of the differences in proportions (right chart) for each V<sub>H</sub> subfamily. The color of the intervals depicts test interpretation: non-significant (red).

### 3.3.2 Analysis of the constant domains

An additional analysis, not commonly implemented in this kind of studies, involved the identification of changes in antibody classes and subclasses that might occur during the selection process. This analysis is of particular importance in the characterization of the immune library, since it could contribute to demonstrate the translocation of the humoral response found in the celiac donors, in the phagemid DNA that constituted the library. This study revealed a clear change on the antibody subclass when comparing the unpanned library and its fourth round of selection (Figure 9). Before de selection process, the vast majority of chains presented an IgG1 subclass (56 %), followed by a 33 % of IgG2 subclass, and a the 10 % of IgA chains. On the other hand, after the selection process, the most represented chain was IgA2 (36 %) which resulted the minority subclass in the unpanned

library (14-fold increase in abundance). As observed, the composition in antibody classes and subclasses completely changed during the selection process against gliadin. This finding further correlates with the fact that IgA2 was the subclass of the best gliadin-binding Fab isolated from the library (Garcia-Calvo et al., 2023).

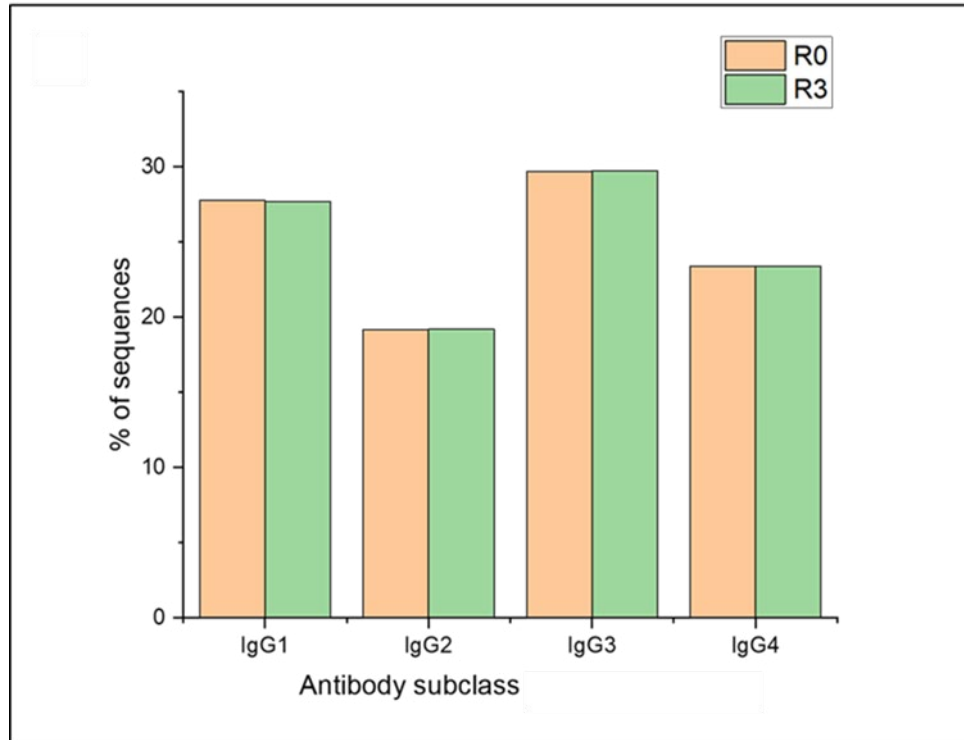


**Figure 9.** Antibody subclass composition of the unpanned immune library (R0, orange) and its fourth round of selection (R4, green). Data were extracted from the alignment of the sequences of each round against the human genome, followed by counting sequences with the coverage function of the bedtools package.

The library composition by subclass, prior to the selection process, closely resembled that of IgG in human serum, where IgG represents over 80 % of immunoglobulins. Furthermore, the customary 2:1 ratio in the quantity of IgG1 and IgG2 was maintained (Vidarsson et al., 2014). The proportion of IgA antibodies in the repertoire also mirrored the typical range found in healthy human serum, which is approximately 10 – 15 % (Loh et al., 2013). However, it is important to note that in celiac disease, this proportion is often increased, as reported in previous works (Caio et al., 2019), a phenomenon that was not captured in our repertoire.

Notably, there was a clear selection of antibodies from the IgA2 subclass. IgA2 is minimally represented in the bloodstream under normal circumstances, except in cases of intense celiac-type responses, where antibodies from this subclass (typically expressed in the intestinal mucosa) appear in the plasma (Volta et al., 2008). Therefore, it is not surprising that the most reactive antibodies against gliadin that were selected predominantly belonged to the IgA2 subclass.

Besides, no significant changes were observed in the comparison of antibody subclasses between the unpanned semi-immune library and its last round of selection (Figure 10).



**Figure 10.** Antibody subclass composition of the unpanned semi-immune (R0, orange) and its third round of selection (R3, green). Data were extracted from the alignment of the sequences of each round against the human genome, followed by counting sequences with the coverage function of the bedtools package.

Although equimolar PCR products were used in the library construction procedure with the intention of achieving equal representation of the heavy chain subclasses (25 % each), sequencing data indicated that certain chains had been amplified more efficiently than others, resulting in a slight imbalance.

In the present study, the investigation of the merged library revealed an absence of discernible evolutionary dynamics. Notably, no substantial alterations were observed between the initial and final rounds of selection for the light chains. This trend was similarly observed for the heavy chains, where the IgG3 Fab exhibited a high level of recurrence with nearly identical frequencies at both endpoints. The preservation of dAb functionality while transitioning into Fab can be attributed to primarily structural factors, specifically the transformation of the heavy chain and the incorporation of immune light chains. Consequently, it can be postulated that the acquisition of diverse antibody variants through evolution and selection processes played a lesser role in this particular context.

## 4 CONCLUSIONS

This study presents an illustrative demonstration of a rapid and precise methodology for sequencing and analyzing antibody repertoires (and their usage) generated through phage-mediated affinity-selection processes.

The utilization of Oxford Nanopore technology facilitated a user-friendly and expeditious sequencing approach using portable equipment. This compact device enables the acquisition of significantly longer reads, thereby allowing simultaneous sequencing of entire Fab, obviating the need for assembly from fragmented reads.

The proposed workflow analysis enables the generation of valuable data for a comprehensive characterization of the selection process, revealing features that may be overlooked by other analytical methods, such as traditional Sanger sequencing of individual colonies. This approach provides a deeper insight into the selection process, capturing additional information that would otherwise be inaccessible due to the size of the data generated.

The primary analysis approach facilitated a rapid and seamless handling of the raw data produced by the sequencing device. This encompassed key steps such as basecalling utilizing neural networks, demultiplexing, quality analysis, trimming of low-quality sequences, and alignment against the reference genome. These steps effectively prepared the data for secondary analysis focused on antibody data generation, utilizing only high-quality sequences.

The secondary analysis involved the utilization of bioinformatics tools, specifically IMGT/HighV-Quest, for the classification of reads based on VDJ genes, CDR features, and clonotype organization. Additionally, IMGT/StatClonotype was employed for statistical analysis and alignment-dependent analysis was conducted for isotyping and mutation analysis. These bioinformatics tools provided essential functionalities to enable comprehensive exploration and interpretation of the obtained data in an immunological context.

The progression of the panning process of the immune library against gliadin revealed a discernible selection pattern, as evidenced by the substantial expansion of VH3-15 variants within the Fab population starting from round 2, despite the initial presence of over 30 heavy chain families. While there were notable alterations observed in the composition of light chains, the data suggest that these chains did not play a significant role in the depicted selection process.

The analysis specifically targeting the merged chains library revealed that the conservation and expansion of features observed in the transformation of a pre-existing dAb into a Fab were not dependent on molecular selection. Significantly, no substantial changes were detected in the immune light chains cloned or in the transformation of the dAb into a heavy chain. This suggests that other factors, like structural characteristics, might have contributed to the preservation and amplification of these features.

Lastly, the analysis of the VH3-15 variants that were selected during the panning process provided insights into certain immunological characteristics of the Fab. The lower observed mutation rate in the VH segment, consistent with previous findings, suggests that the generated Fab originated from plasma cells activated through T-cell dependent extrafollicular stimulation. These identified features validate that the immune library generated encompasses crucial attributes transferred from the humoral response of individuals with celiac disease.

## 5 BIBLIOGRAPHY

- Aouinti, S., Giudicelli, V., Duroux, P., Malouche, D., Kossida, S., & Lefranc, M. P. (2016). IMGT/StatClonotype for pairwise evaluation and visualization of NGS IG and TR IMGT Clonotype (AA) diversity or expression from IMGT/HighV-QUEST. *Frontiers in Immunology*, 7, 339. <https://doi.org/10.3389/fimmu.2016.00339>
- Barbas, C. F. (2001). *Phage display : a laboratory manual*. Cold Spring Harbor Laboratory Press.
- Barreto, K., Maruthachalam, B. V., Hill, W., Hogan, D., Sutherland, A. R., Kusalik, A., Fonge, H., DeCoteau, J. F., & Geyer, C. R. (2019). Next-generation sequencing- guided identification and reconstruction of antibody CDR combinations from phage selection outputs. *Nucleic Acids Research*, 47(9), e50. <https://doi.org/10.1093/nar/gkz131>
- Bowers, E., Scamurra, R. W., Asrani, A., Beniguel, L., Mawhinney, S., Keays, K. M., Thurn, J. R., & Janoff, E. N. (2014). Decreased mutation frequencies among immunoglobulin G variable region genes during viremic HIV-1 infection. *PLoS ONE*, 9(1), e81913. <https://doi.org/10.1371/journal.pone.0081913>
- Brochet, X., Lefranc, M. P., & Giudicelli, V. (2008). IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Research*, 36, W503-W508. <https://doi.org/10.1093/nar/gkn316>
- Caio, G., Volta, U., Sapone, A., Leffler, D., De Giorgio, R., Catassi, C., & Fasano, A. (2019). Celiac disease: a comprehensive current review. *BMC Medicine*, 17(1), 142. <https://doi.org/10.1186/s12916-019-1380-z>
- Chen, P., Sun, Z., Wang, J., Liu, X., Bai, Y., Chen, J., Liu, A., Qiao, F., Chen, Y., Yuan, C., Sha, J., Zhang, J., Xu, L. Q., & Li, J. (2023). Portable nanopore-sequencing technology: Trends in development and applications. *Frontiers in Microbiology*, 14, 1043967. <https://doi.org/10.3389/fmicb.2023.1043967>
- D'Angelo, S., Kumar, S., Naranjo, L., Ferrara, F., Kiss, C., & Bradbury, A. R. (2014). From deep sequencing to actual clones. *Protein Engineering Design and Selection*, 27(10), 301-307. <https://doi.org/10.1093/protein/gzu032>
- García-Calvo, E., García-García, A., Rodríguez-Gómez, S., Farrais, S., Martín, R., & García, T. (2023). Development of a new recombinant antibody, selected by phage-display technology from a celiac patient library, for detection of gluten in foods. *Current Research in Food Science*, 7, 100578. <https://doi.org/https://doi.org/10.1016/j.crfs.2023.100578>
- García-Calvo, E., García-García, A., Rodríguez, S., Farrais, S., Martín, R., & García, T. (2022). Construction of a Fab library merging chains from semisynthetic and immune origin, suitable for developing new tools for gluten immunodetection in food. *Foods*, 12(1), 149. <https://doi.org/10.3390/foods12010149>

- Glanville, J., D'Angelo, S., Khan, T. A., Reddy, S. T., Naranjo, L., Ferrara, F., & Bradbury, A. R. M. (2015). Deep sequencing in library selection projects: what insight does it bring? *Current opinion in structural biology*, *33*, 146-160. <https://doi.org/10.1016/j.sbi.2015.09.001>
- Ho, F., Lortan, J. E., MacLennan, I. C., & Khan, M. (1986). Distinct short-lived and long-lived antibody-producing cell populations. *European Journal of Immunology*, *16*(10), 1297-1301. <https://doi.org/10.1002/eji.1830161018>
- Hwang, J. K., Alt, F. W., & Yeap, L. S. (2015). Related mechanisms of antibody somatic hypermutation and class switch recombination. *Microbiology Spectrum*, *3*(1), Mdna3-0037-2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0037-2014>
- Itzlinger, A., Branchi, F., Elli, L., & Schumann, M. (2018). Gluten-free diet in celiac disease—forever and for all? *Nutrients*, *10*(11), 1796. <https://doi.org/10.3390/nu10111796>
- Jaroszewicz, W., Morcinek-Orłowska, J., Pierzynowska, K., Gaffke, L., & Węgrzyn, G. (2021). Phage display and other peptide display technologies. *FEMS Microbiology Reviews*, *46*(2). <https://doi.org/10.1093/femsre/fuab052>
- Joyce, C., Burton, D. R., & Briney, B. (2020). Comparisons of the antibody repertoires of a humanized rodent and humans by high throughput sequencing. *Scientific Reports*, *10*(1), 1120. <https://doi.org/10.1038/s41598-020-57764-7>
- Kim, S., Park, I., Park, S. G., Cho, S., Kim, J. H., S. Ipper, N., Choi, S. S., Lee, E. S., & Hong, H. J. (2017). Generation, diversity determination, and application to antibody selection of a human naïve Fab library. *Molecules and Cells*. <https://doi.org/10.14348/molcells.2017.0106>
- Köhler, G., & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, *256*(5517), 495-497. <https://doi.org/10.1038/256495a0>
- Lee, S. K., Rigby, R. J., Zotos, D., Tsai, L. M., Kawamoto, S., Marshall, J. L., Ramiscal, R. R., Chan, T. D., Gatto, D., Brink, R., Yu, D., Fagarasan, S., Tarlinton, D. M., Cunningham, A. F., & Vinuesa, C. G. (2011). B cell priming for extrafollicular antibody responses requires Bcl-6 expression by T cells. *Journal of Experimental Medicine*, *208*(7), 1377-1388. <https://doi.org/10.1084/jem.20102065>
- Lefranc, M.-P., & Lefranc, G. (2020). Immunoglobulins or antibodies: IMGT® bridging genes, structures and functions. *Biomedicines*, *8*(9), 319. <https://doi.org/10.3390/biomedicines8090319>
- Loh, R. K., Vale, S., & McLean-Tooke, A. (2013). Quantitative serum immunoglobulin tests. *Australian Family Physician*, *42*(4), 195-198.
- Lowden, M. J., & Henry, K. A. (2018). Oxford nanopore sequencing enables rapid discovery of single-domain antibodies from phage display libraries. *BioTechniques*, *65*(6), 351-356. <https://doi.org/10.2144/btn-2018-0123>
- Matochko, W. L., & Derda, R. (2015). Next-generation sequencing of phage-displayed peptide libraries. *Methods in Molecular Biology*, *1248*, 249-266. [https://doi.org/10.1007/978-1-4939-2020-4\\_17](https://doi.org/10.1007/978-1-4939-2020-4_17)
- Packer, M. S., & Liu, D. R. (2015). Methods for the directed evolution of proteins. *Nature Reviews Genetics*, *16*(7), 379-394. <https://doi.org/10.1038/nrg3927>
- Roth, D. B. (2014). V(D)J Recombination: mechanism, errors, and fidelity. *Microbiology Spectrum*, *2*(6). <https://doi.org/10.1128/microbiolspec.MDNA3-0041-2014>

- Sapone, A., Bai, J. C., Ciacci, C., Dolinsek, J., Green, P. H., Hadjivassiliou, M., Kaukinen, K., Rostami, K., Sanders, D. S., Schumann, M., Ullrich, R., Villalta, D., Volta, U., Catassi, C., & Fasano, A. (2012). Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Medicine*, *10*, 13. <https://doi.org/10.1186/1741-7015-10-13>
- Sharma, N., Bhatia, S., Chunduri, V., Kaur, S., Sharma, S., Kapoor, P., Kumari, A., & Garg, M. (2020). Pathogenesis of celiac disease and other gluten related disorders in wheat and strategies for mitigating them. *Frontiers in Nutrition*, *7*. <https://doi.org/10.3389/fnut.2020.00006>
- Shi, B., Dong, X., Ma, Q., Sun, S., Ma, L., Yu, J., Wang, X., Pan, J., He, X., Su, D., & Yao, X. (2020). The usage of human IGHJ genes follows a particular non-random selection: The recombination signal sequence may affect the usage of human IGHJ genes. *Frontiers in Genetics*, *11*, 524413. <https://doi.org/10.3389/fgene.2020.524413>
- Smith, G. P. (2019). Phage Display: simple evolution in a Petri dish (Nobel Lecture). *Angewandte Chemie International Edition*, *58*(41), 14428-14437. <https://doi.org/10.1002/anie.201908308>
- Steinsbø, Ø., Dunand, C. J. H., Huang, M., Mesin, L., Salgado-Ferrer, M., Lundin, K. E. A., Jahnsen, J., Wilson, P. C., & Sollid, L. M. (2014). Restricted VH/VL usage and limited mutations in gluten-specific IgA of coeliac disease lesion plasma cells. *Nature Communications*, *5*(1). <https://doi.org/10.1038/ncomms5041>
- Urquhart, L. (2023). Top companies and drugs by sales in 2022. *Nature Reviews Drug Discovery*, *22*(260). <https://doi.org/10.1038/d41573-023-00039-3>
- Vidarsson, G., Dekkers, G., & Rispen, T. (2014). IgG subclasses and allotypes: from structure to effector functions. *Frontiers in Immunology*, *5*, 520. <https://doi.org/10.3389/fimmu.2014.00520>
- Volta, U., Molinaro, N., Fratangelo, D., & Bianchi, F. B. (2008). IgA subclass antibodies to gliadin in serum and intestinal juice of patients with coeliac disease. *Clinical and Experimental Immunology*, *80*(2), 192-195. <https://doi.org/10.1111/j.1365-2249.1990.tb05232.x>
- Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*, *39*(11), 1348-1365. <https://doi.org/10.1038/s41587-021-01108-x>
- Winter, G. (2019). Harnessing evolution to make medicines (Nobel Lecture). *Angewandte Chemie International Edition*, *58*(41), 14438-14445. <https://doi.org/10.1002/anie.201909343>





## 5. DISCUSIÓN INTEGRADORA



## 5 DISCUSIÓN INTEGRADORA

Existen diversas metodologías para la detección de gluten en alimentos, entre las que destacan las técnicas de inmunodetección, que se basan en la especificidad de la unión de los anticuerpos con sus correspondientes antígenos. Debido a la elevada prevalencia de las patologías relacionadas con el gluten (Taraghikhah et al., 2020), en las últimas décadas se han desarrollado varios anticuerpos con alta afinidad para su detección, según la tecnología disponible en cada momento. Las primeras aproximaciones se llevaron a cabo mediante el aislamiento de anticuerpos policlonales a partir del suero de animales inmunizados (Manohar, 2002). La revolución tecnológica que supuso la generación de hibridomas productores de anticuerpos monoclonales también se trasladó a este campo. Entre los anticuerpos monoclonales generados frente al gluten destaca el R5, a partir del cual se desarrolló un sistema de detección de ELISA sándwich directo que, a día de hoy, es el método de referencia del *Codex Alimentarius* para la detección de gluten en alimentos (Méndez et al., 2005).

Actualmente, los anticuerpos monoclonales siguen siendo muy utilizados, pero se han obtenido nuevos tipos de moléculas, como los anticuerpos recombinantes, que presentan varias ventajas, como una producción rápida y constante de diversos tipos de fragmentos de anticuerpos sin necesidad de utilizar animales de experimentación, la posibilidad de introducir modificaciones funcionales en la molécula y que, a diferencia de los anticuerpos secretados por hibridomas, no presentan cadenas no deseadas (Bradbury et al., 2018).

A finales del siglo XX se desarrolló el concepto de evolución dirigida de proteínas, que abrió nuevas posibilidades en la producción de anticuerpos. Se basa en generar amplios repertorios de proteínas (anticuerpos en este caso) que se someten a sucesivas rondas de selección frente a un antígeno, seguidas de la amplificación de los candidatos seleccionados. Esto permite generar una población de anticuerpos en la que están representados mayoritariamente aquellos con mayor afinidad frente al antígeno objetivo. La tecnología más utilizada es la de *phage display*, basada en el emparejamiento del fenotipo (anticuerpo) con el genotipo (la secuencia de ADN que lo codifica) en fagos filamentosos (M-13) que contienen un fagémido (genotipo) que codifica un anticuerpo unido a proteínas propias del fago (fenotipo).

El objetivo de este trabajo ha consistido en trasladar la respuesta humoral de pacientes celíacos a dos repertorios de anticuerpos recombinantes, para obtener moléculas que presentaran una elevada afinidad por el gluten.

Gracias a los avances en ingeniería de proteínas, existe una gran cantidad de fragmentos de anticuerpos que no se encuentran en la naturaleza y que están diseñados específicamente para un uso concreto. El tipo de fragmento de anticuerpo recombinante elegido para la construcción de repertorios durante la presente tesis doctoral fue el Fab (fragmento de unión al antígeno), que está compuesto por dos cadenas: la cadena ligera completa (con sus regiones variable y constante), y una cadena pesada que consiste en el dominio variable y la primera región constante (CH1) de las inmunoglobulinas de mamíferos, pero que carece de la región Fc (fragmento cristalizante). Este formato replica la región que se une a los antígenos en los anticuerpos naturales, proporcionando el bolsillo paratópico completo. Además, los Fab son más estables que los formatos de cadena única, debido al puente disulfuro que une las cadenas pesada y ligera. Su menor tamaño (en comparación con las inmunoglobulinas completas) mejora sus propiedades de difusión y facilita la expresión recombinante en células procariotas (Kang & Seong, 2020).

Para conseguir este objetivo general fue necesario desarrollar los siguientes objetivos parciales:

1. Obtención y preparación de muestras biológicas como material de partida para el desarrollo de repertorios de anticuerpos recombinantes.
2. Construcción de un repertorio semiinmune de Fab para seleccionar anticuerpos recombinantes de alta afinidad frente al gluten.
3. Construcción de un repertorio inmune de Fab para seleccionar anticuerpos recombinantes de alta afinidad frente al gluten.
4. Producción soluble de los Fab seleccionados a partir de los repertorios generados y desarrollo de técnicas inmunoenzimáticas ELISA para la detección de gluten en alimentos.
5. Modelización *in silico* de la estructura e interacción de los anticuerpos recombinantes obtenidos que presentaban mejores características de detección.
6. Caracterización de los repertorios mediante técnicas de secuenciación masiva.

## 5.1 OBTENCIÓN Y PREPARACIÓN DE MUESTRAS BIOLÓGICAS COMO MATERIAL DE PARTIDA PARA EL DESARROLLO DE REPERTORIOS DE ANTICUERPOS RECOMBINANTES

A través de una colaboración con el Hospital Universitario Fundación Jiménez Díaz (Madrid), se seleccionaron varios pacientes recién diagnosticados de enfermedad celíaca. Una vez firmado el preceptivo consentimiento informado, los pacientes seleccionados donaron su sangre, de la que se aislaron linfocitos B. El ARN aislado de los linfocitos se retrotranscribió a ADNc y fue utilizado como molde para amplificar los genes que codifican para las cadenas ligeras y pesadas de los anticuerpos.

### 5.1.1 Proceso de selección de pacientes para su inclusión en el estudio

La construcción de un repertorio inmune con suficiente diversidad como para permitir el aislamiento de Fab de alta afinidad dirigidos frente al gluten, requería capturar las secuencias codificantes de los anticuerpos involucrados en la respuesta humoral frente al antígeno objetivo. Para cumplir con este requisito, se buscaron donantes con una alta respuesta humoral frente a varios de los antígenos relacionados con la patología celíaca, que generalmente se correlacionan con una respuesta elevada frente a la gliadina (antígeno objetivo). Los parámetros iniciales para la selección de donantes incluyeron los títulos de anticuerpos IgG e IgA dirigidos frente a transglutaminasa tisular (anti-tTG) y a péptidos deaminados de gliadina (anti-DGP). Estos marcadores serológicos son muy utilizados en la práctica clínica, debido a que su presencia se correlaciona con el daño intestinal, y a que su sensibilidad y especificidad para el diagnóstico de la enfermedad celíaca no tratada son cercanas o superiores al 95 % (Leffler & Schuppan, 2010). El establecimiento de un punto de corte estandarizado y universalmente aplicable con fines predictivos para estos biomarcadores, particularmente el título de anti-tTG circulante, es complejo debido a la falta de armonización entre los diversos métodos de ensayo disponibles, y a que la mayoría de las investigaciones se dirigen a la población pediátrica (Beltran et al., 2014). En consecuencia, en el proceso de selección de pacientes para este trabajo, nuestros colaboradores clínicos se basaron en las pautas clínicas de su hospital, que dictaban los umbrales de las pruebas analíticas, adaptados al contexto de su población.

La utilización del título de anticuerpos específicos frente a la gliadina completa en sangre periférica como biomarcador de enfermedad celíaca no es muy común en la práctica clínica contemporánea, debido a que es menos sensible que otros biomarcadores (Singh et al., 2022). No obstante, en esta tesis doctoral se consideró importante incluirlo, debido a que se ha utilizado la gliadina como antígeno principal para el proceso de selección. Caracterizar minuciosamente la respuesta humoral de los posibles donantes frente a la gliadina fue fundamental, ya que permitió trasladar estas características inmunológicas a los repertorios desarrollados.

También fue muy importante asegurarse de que los pacientes no hubieran iniciado una dieta sin gluten, ya que su eliminación inhibe de manera significativa y abrupta la respuesta humoral anormal de la enfermedad celíaca (Sharma et al., 2020).

En el estudio se incluyeron cuatro pacientes como posibles donantes (Tabla 4.2.1). El paciente 2 fue descartado porque los niveles de IgG anti-tTG eran bajos y la IgA anti-gliadina rozaba los niveles límites para su consideración como positiva. El paciente 3 presentaba niveles elevados de respuesta humoral y un claro diagnóstico de celiaquía, pero no pudo ser incluido porque comenzó unilateralmente una dieta sin gluten antes de la donación de sangre. Finalmente, los pacientes 1 y 4 cumplieron con todos los requisitos, firmaron el consentimiento informado, y donaron muestras de sangre.

Con el fin de conocer y trasladar al repertorio la respuesta humoral de los pacientes seleccionados con mayor facilidad, se analizó una muestra de suero de cada donante (pacientes 1 y 4) para cuantificar la expresión de los distintos subisotipos IgG frente a la gliadina nativa (Figura 4.2.1S). Los pacientes 1 y 4 presentaron diferentes respuestas de subisotipo IgG frente a la gliadina, con la expresión dominante habitual de IgG1 en el paciente 1 (Nahm et al., 1987; Napodano et al., 2021). Sin embargo, el paciente 4 también presentó una expresión muy alta del subisotipo IgG2, una característica inusual en la enfermedad celíaca (Husby et al., 1986), pero se consideró de interés para la construcción del repertorio.

A pesar de que el número de donantes idóneos era reducido, los altos niveles de repuesta frente a la gliadina nativa indicaban que era un antígeno altamente inmunogénico para ellos, y esto aseguraba (al menos teóricamente) una elevada probabilidad de aislar anticuerpos de alta afinidad (Hoogenboom, 1997). Además, investigaciones recientes han demostrado que no es estrictamente necesario generar repertorios a partir de gran cantidad de muestras biológicas, si se amplifican las cadenas de anticuerpos correctamente, dado que el sistema inmune de un solo individuo genera la suficiente diversidad molecular (Erasmus et al., 2021).

### **5.1.2 Aislamiento de linfocitos B de sangre periférica para la extracción y caracterización de su RNA**

El material de partida para la amplificación de los genes codificantes para las cadenas de anticuerpos que se incluyeron en el repertorio de Fab fue el ARNm extraído de linfocitos B de sangre periférica de los pacientes celíacos seleccionados. Como estos linfocitos producen los anticuerpos maduros frente al antígeno objetivo de esta forma se aseguraba la amplificación de las moléculas de ARNm que codifican las secuencias ya reorganizadas a partir de la recombinación de los genes V, D y J (Victora & Nussenzweig, 2012).

Otros investigadores han desarrollado repertorios a partir de linfocitos B residentes en tejidos intestinales, cuyos anticuerpos están menos madurados, pero son más diversos, además de proceder de una mucosa continuamente estimulada por los antígenos (Marzari et al., 2001). Esta opción no se consideró adecuada ni posible para la presente tesis doctoral, porque hubiera requerido obtener muestras de biopsias intestinales de gran cantidad de pacientes, con las dificultades técnicas, logísticas y éticas que ello conllevaría.

Se aislaron células mononucleares de la sangre de los pacientes mediante separación con gradiente Ficoll®. A partir de estas células, se aisló el ARN total utilizando el reactivo TriZol LS, obteniendo 100 µg de alta calidad (con valores de RIN (*RNA integrity number*) superiores a 8/10). Con el ARN total de cada paciente se realizó RT-PCR utilizando cebadores oligo-dT-20 inespecíficos, y se obtuvo el ADNc que se empleó como molde para las amplificaciones específicas de los genes que codifican para anticuerpos. Teniendo en cuenta la cantidad limitada de ARN aislado, se consideró que esta era la mejor opción, aunque otros investigadores también proponen amplificar directamente del molde de ARN las secuencias que codifican para las cadenas de anticuerpos (Lee et al., 2021).

## 5.2 CONSTRUCCIÓN DE UN REPERTORIO SEMIINMUNE DE FAB, MEDIANTE LA COMBINACIÓN DE CADENAS DE ANTICUERPOS DE ORIGEN SEMISINTÉTICO E INMUNE, Y DESARROLLO DE UN MÉTODO ELISA PARA LA DETECCIÓN DE GLUTEN, UTILIZANDO LOS FAGO-FAB SELECCIONADOS POR PHAGE DISPLAY

Para la construcción de este repertorio se partió de un dAb preexistente (García- García et al., 2020) que se transformó en una cadena pesada de Fab (añadiendo las regiones constantes para IgG) y, a continuación, se clonó en un subrepertorio que contenía los genes que codifican para las cadenas ligeras de los anticuerpos de los pacientes celíacos. Se trata, por tanto, de un repertorio semiinmune.

### 5.2.1 Construcción del subrepertorio de cadenas pesadas semisintéticas

Para la construcción de la cadena pesada de los Fab, se utilizó como base la secuencia de un anticuerpo de dominio único (denominado dAb-8E) dirigido frente al gluten, y obtenido previamente por nuestro grupo (García-García et al., 2020). Previamente a la clonación, la secuencia del dAb-8E se sometió a una mutación dirigida para eliminar un sitio de restricción intrasecuencia no deseado. Dado que la secuencia del dAb carece de la región constante de la cadena pesada, las secuencias codificantes para los Fab se construyeron con la secuencia dAb-8E como VH, y la región CH1 se agregó mediante OE-PCR (*Overlap-PCR*). Se realizó una primera ronda de PCR para obtener las secuencias del dAb-8E y las secuencias codificantes para los CH1-IgG, que dieron origen a constructos intermedios de 392 y 300 pb, respectivamente, con extremos

modificados y solapantes. Los fragmentos de ADN, purificados a partir de la electroforesis en gel de agarosa, se utilizaron como ADN molde para la segunda ronda de PCR (OE-PCR), lo que permitió la superposición y fusión de las secuencias anteriores (dAb-8E y CH1-IgG), para obtener los constructos que codifican para la cadena pesada del Fab (Figura 4.1.1A). La metodología de ingeniería de anticuerpos desarrollada en esta tesis doctoral para transformar una molécula pequeña (el dAb) en una de mayor tamaño (Fab), es novedosa y única en el campo de la detección de alérgenos. Esta aproximación se aplica habitualmente en el diseño de receptores de células T quiméricos (CAR-T cell, *Chimeric Antigen Receptor- T cell*) una de las terapias oncológicas más eficaces actualmente (Schoonjans et al., 2000). También se suele utilizar la estrategia contraria, consistente en transformar una molécula completa en un fragmento más pequeño y de cadena única, como los scFv (de Aguiar et al., 2021).

### 5.2.2 Construcción del subrepertorio de cadenas ligeras inmunes

El ADNc obtenido de las muestras biológicas de los pacientes se utilizó como molde para una primera ronda de PCR con el fin de amplificar los genes codificantes para las cadenas ligeras completas, lo que dio lugar a fragmentos de aproximadamente 650 pb. Teniendo en cuenta la cantidad limitada de ARN obtenido de las muestras biológicas, se priorizó la amplificación de las cadenas kappa, observándose una mayor expresión de la familia Vk1 (Figura 4.1.1B), que es el tipo de cadena ligera más expresada por el ser humano (Guigou et al., 1990). Después de la purificación del ADN, se realizó una segunda ronda de PCR para incluir sitios de restricción mejorados (colas adyacentes a los sitios de restricción para facilitar el acceso de las enzimas) en las secuencias codificantes para cadenas kappa.

### 5.2.3 Clonación del repertorio de Fab y selección por afinidad frente a la gliadina (*panning*)

Para la construcción del repertorio de Fab, se realizó un proceso de clonación en dos pasos. Se clonaron primero las cadenas ligeras de origen inmune y posteriormente las cadenas pesadas de origen semisintético. Era necesario mantener este orden, debido a que las cadenas amplificadas con el cebador VH4gs (Tabla 4.2.1S) contienen secuencias muy similares a las dianas de restricción introducidas para la clonación de las cadenas ligeras (Barbas et al., 2001).

En primer lugar, se clonaron los insertos que codifican para las cadenas ligeras en el fagémido pComb3X linealizado, y se transformaron en células electrocompetentes *E. coli* XL-1 Blue, obteniéndose el subrepertorio de cadenas ligeras, compuesto por  $1,4 \times 10^7$  transformantes totales. Se aisló el ADN plasmídico y se linealizó nuevamente para clonar los constructos codificantes para las cadenas pesadas. Los productos de ligación finales se precipitaron y se transformaron (con un *elapsed time* 4 ms), lo que dio lugar a un repertorio de Fab con un RZ de  $2,7 \times 10^7$  (factor de autoligación del 2 %), cumpliendo con los estándares considerados aceptables

para este tipo de repertorios ( $RZ > 10^7$  y un factor de autoligación menor al 5 %). Estos criterios de calidad, desarrollados por los primeros investigadores en *phage display* (Barbas, 2001), son muy utilizados. Sin embargo, más recientemente se ha descubierto que el tamaño aparente del repertorio (determinado como transformantes totales) no se corresponde con el tamaño real (que se puede medir por técnicas de secuenciación masiva), y que puede llegar a ser bastante menor del esperado (del orden de  $10^4$  en lugar de  $10^7$  secuencias distintas), aunque el repertorio siga siendo funcional (Glanville et al., 2015). Por lo tanto, estos criterios de calidad más antiguos se siguen empleando por su fácil aplicación e inmediatez, pero no son una medida absoluta del tamaño y calidad de los repertorios.

Después de la infección con el fago auxiliar (VSC-M13), se aisló el repertorio de fago-anticuerpos mediante purificación con polietilenglicol (PEG) y se utilizó para el *panning* (rondas sucesivas de selección y amplificación) frente a gliadina- PWG. El proceso de enriquecimiento del repertorio se caracterizó mediante la titulación de los fagos eluidos después de cada una de las tres rondas de selección realizadas. La recuperación de fagos aumentó 58 veces después de la tercera ronda de *panning*, lo que indica una probable selección de variantes de fago-anticuerpos frente a la molécula objetivo (Figura 4.1.1S). Para confirmar este resultado, se realizaron ensayos de ELISA policlonal para evaluar si hubo un enriquecimiento efectivo en fago-Fab con alta afinidad por la gliadina. Como era de esperar, la gliadina apenas fue detectada por el repertorio no seleccionado, pero después de la primera interacción con el antígeno de interés, se produjo un aumento de señal de 7 veces, y se obtuvieron valores máximos de absorbancia con las poblaciones de fago-Fab de la segunda y tercera rondas de selección. Se registraron señales muy bajas ( $< 0.2$  UA) para el control negativo (BSA utilizado en la solución de bloqueo), lo que indica que no hubo una selección no específica de variantes de fagos frente a otros componentes del soporte o el bloqueante (Figura 4.1.2S). Con estos resultados, se consideró innecesaria una cuarta ronda de selección.

#### 5.2.4 Selección de clones de fago-Fab individuales mediante ELISA monoclonal

A partir de las placas de titulación de la población de fagos eluidos de la tercera ronda de *panning*, se seleccionaron aleatoriamente un total de 24 clones individuales. Los sobrenadantes (que contenían los fago-anticuerpos) se diluyeron y se sometieron a un ensayo de ELISA indirecto con fago-Fab monoclonales en placas recubiertas con gliadina-PWG (Figura 4.1.2A). Se identificaron los pocillos con fago-anticuerpos que reconocían el antígeno, lo que dio lugar a la selección de 4 de los 24 clones (aproximadamente el 17 %) como clones de alta afinidad frente a gliadina (aquellos que presentaban valores de absorbancia superiores a 2 UA y señales de control negativo inferiores a 0,3 UA, para asegurar la ausencia de reactividad cruzada con el bloqueante). La técnica de ELISA indirecto con fago-Fab monoclonales también se empleó para evaluar la afinidad y especificidad de los cuatro fago-Fab seleccionados, como una forma de caracterización parcial de los Fab aislados, antes de su producción como moléculas solubles no unidos a fagos, ya que esta estrategia demostró ser muy útil en trabajos anteriores (García-García et al., 2020; Madrid et al., 2020).

Para analizar la afinidad, se obtuvo una curva de dosis-respuesta para cada variante de fago-Fab frente a concentraciones crecientes (2,5 - 75  $\mu\text{g}/\text{mL}$ ) de gliadina-PWG (Figura 4.1.2B). El Fab8E-4 mostró las propiedades de reconocimiento más prometedoras, produciendo la curva de dosis-respuesta más sensible. Además, se realizó una prueba de especificidad con harinas de diferentes cereales (Figura 4.1.2C) para evaluar el rendimiento de los clones seleccionados en una matriz alimentaria real, y para comprobar los tipos de prolaminas reconocidas por los fago-Fab. El Fab8E-4 mostró mayor capacidad de detección de prolaminas de trigo, cebada y centeno que el resto de los clones analizados. Teniendo en cuenta estos resultados, entre los fago-Fab aislados, el Fab8E-4 presentó mejores características para la detección de prolaminas, tanto en diluciones de gliadina pura como en extractos etanólicos de fuentes naturales de gluten. En consecuencia, se seleccionó el fago-Fab8E-4 para ensayos posteriores.

### 5.2.5 Secuenciación y análisis de los clones seleccionados

La secuenciación del ADN plasmídico de varios clones aislados después de la transformación bacteriana mostró que el proceso de clonación en dos pasos fue exitoso, dado que la gran mayoría de clones secuenciados eran Fab clonados en el marco de lectura correcto, y que no se observaron ni eventos generalizados de autoligación ni otros artefactos. Por lo tanto, el repertorio construido consistió en una colección de cadenas ligeras expresadas en pacientes con enfermedad celíaca y cuatro conjuntos de cadenas pesadas (subclases IgG1, IgG2, IgG3 e IgG4) que comparten como región variable la secuencia del dAb-8E. A partir del análisis de la secuencia de ADN de los clones seleccionados como positivos después del *biopanning* (Figura 4.1.3A), se observó que todos derivaban de la disposición combinatoria de cuatro cadenas ligeras diferentes y la misma cadena pesada, dAb-8E-IgG3. (Figura 4.1.3B). Este hecho podría significar que las cadenas pesadas, y no las cadenas ligeras, son los principales agentes en el proceso de evolución molecular que selecciona los clones de alta afinidad; lo que no es extraño, puesto que en los anticuerpos humanos, la cadena con mayor participación en la interacción con el antígeno es la cadena pesada (Schoonbroodt et al., 2008). La avidéz general del Fab puede verse incrementada gracias a la adición de cadenas ligeras que también generen contactos adicionales con el antígeno, reforzando la afinidad de la cadena pesada dAb-8E-IgG3.

El repertorio derivado del dAb presenta la mitad de la diversidad de CDRs en comparación con un repertorio inmune completo, ya que la diversidad sólo aparece en las cadenas ligeras. Sin embargo, la afinidad total del repertorio se mejora, ya que muchos de los fago-Fab de un repertorio inmune podrían no unirse correctamente al antígeno diana, mientras que en este repertorio todas las cadenas pesadas se unen a la gliadina. Aunque esta situación podría conducir a no detectar un nuevo Fab “aún mejor” que podría encontrarse (o no) en un repertorio inmune clásico, la selección de una subclase de anticuerpo expresada minoritariamente en el ser humano, como IgG3 (Damelang et al., 2019), muestra que el reordenamiento artificial podría generar estructuras funcionales muy subrepresentadas en la naturaleza.

### 5.2.6 Caracterización del fago-Fab8E-4 como sonda para detección de gluten en alimentos mediante una técnica ELISA indirecto

A partir de los resultados anteriores, quedó demostrado que el clon Fab8E-4 presentaba mejores características que el resto, por lo que, se continuó con su caracterización. Para comparar la sensibilidad del ELISA basado en el fago- Fab8E-4 con el del fago-dAb8E parental, ambos formatos de anticuerpo (Fab y dAb) fueron expresados unidos a la proteína pIII del mismo fago auxiliar VSC- M13. Una vez obtenidos los dos formatos unidos al mismo fago, se compararon sus curvas dosis-respuesta frente a concentraciones crecientes de gliadina- PWG (0,5 - 6 µg/mL) (Figura 4.1.4A). El fago-Fab presentó una capacidad ligeramente superior al fago-dAb para detectar la gliadina-PWG en solución. Siguiendo las guías de Eurachem para la validación de métodos analíticos (Magnusson, 2014), el límite de detección (LOD) utilizando el fago-Fab8E-4 se estableció en 0,01 µg/mL (Figura 4.1.4B).

Considerando que los inmunoensayos podrían verse influenciados por efectos de la matriz alimentaria, se preparó un conjunto de mezclas binarias que incluían diferentes concentraciones (0; 0,1; 0,2; 0,5; 1; 2,5; 5; 7,5; 10 y 100 mg/g) de una mezcla proporcional de harinas de trigo, cebada y centeno sin tratar en una base de harina de arroz. Se seleccionó la harina de arroz como la base de las mezclas debido a que no produjo interferencias cuando se analizó mediante ELISA indirecto con el fago-Fab8E-4, y porque se utiliza con frecuencia en formulaciones de alimentos sin gluten. Algunos investigadores han propuesto adoptar un enfoque teórico para estimar la cantidad de gluten como aproximadamente el 10 % del peso del cereal (Martín-Fernández et al., 2015). Por lo tanto, las mezclas binarias utilizadas en este trabajo para evaluar la idoneidad del inmunoensayo en una matriz alimentaria real podrían interpretarse como muestras que contienen aproximadamente 10, 20, 50, 100, 250, 500, 750, 1000 y 10.000 mg de gluten por kg de producto (partes por millón o mg/kg de gluten en una muestra de alimento). Estas mezclas binarias experimentales fueron utilizadas para la validación interna del ensayo (Figura 4.1.4C).

Los límites de detección (LOD) y cuantificación (LOQ) obtenidos a partir del análisis de las mezclas binarias con el fago Fab8E-4 fueron de 10 y 88 mg/kg de gluten, respectivamente. En consecuencia, este ELISA indirecto permitiría detectar los alimentos que superen el umbral de 20 mg/kg establecido para clasificar los productos como “sin gluten” conforme a la legislación vigente (Diario Oficial de la Unión Europea, 2014). La aplicabilidad del método propuesto se probó con diez productos alimentarios comerciales (cinco etiquetados como sin gluten y cinco como con gluten) (Tabla 4.1.S4). El ELISA indirecto con el fago- Fab8E-4 fue capaz de diferenciar ambos tipos de productos. Estos resultados se confirmaron con un ELISA comercial de tipo sándwich basado en el anticuerpo monoclonal R5, el método más utilizado para el análisis de gluten.

Finalmente, se realizó un análisis de *Western blot* para elucidar qué fracciones de prolaminas de diferentes cereales que contienen gluten eran reconocidas por el anticuerpo recombinante

desarrollado (Figura 4.1.4D). Como control se utilizó una preparación comercial de gliadina de trigo que contenía todas las subunidades de gliadina. Los resultados mostraron que el fago-Fab8E-4 era capaz de reconocer significativamente las bandas de proteínas entre 25 y 48 kDa, que podrían corresponder a las subunidades de glutenina de bajo peso molecular y las  $\alpha$ ,  $\beta$  y  $\gamma$ -gliadinas (Wan et al., 2014).

### 5.3 CONSTRUCCIÓN DE UN REPERTORIO INMUNE DE FAB Y DESARROLLO DE UN MÉTODO ELISA PARA LA DETECCIÓN DE GLUTEN, UTILIZANDO LOS FAGO- FAB SELECCIONADOS POR PHAGE DISPLAY

En la construcción de este segundo repertorio, completamente inmune, se incluyeron cadenas ligeras y pesadas generadas directamente a partir de los linfocitos B de los pacientes incluidos en el estudio.

#### 5.3.1 Construcción del repertorio inmune

Para la construcción del repertorio inmune, el ADNc derivado de las muestras obtenidas de los pacientes seleccionados se empleó para amplificar las secuencias que codifican tanto las cadenas ligeras como las cadenas pesadas de los anticuerpos.

Al igual que en el subrepertorio de cadenas ligeras, la amplificación de los fragmentos que codifican las cadenas pesadas de anticuerpos se realizó en dos rondas de PCR secuenciales. Debido al posible sesgo de amplificación durante la PCR, algunas cadenas pueden acabar sobrerrepresentadas en comparación con las expresadas en los pacientes (Lee et al., 2021). Para evitar este problema, se utilizaron las mismas condiciones de PCR para todas las cadenas, de modo que se obtuvieron cantidades diferentes de ADN amplificado de cada cadena, según familia de genes V e isotipo (por ejemplo, VH135-IgG1; VH4-IgA...), en lugar de un conjunto equimolar de cada tipo de cadena. Además, los amplicones codificantes de las cadenas pesadas de los Fab (Figuras 4.2.2SA y 4.2.2SB), se generaron con sitios de restricción terminal diferentes que los de las cadenas ligeras (Figuras 4.2.2SC y 4.2.2SD) para permitir una clonación secuencial.

Este enfoque es novedoso y único debido al uso del formato Fab, mientras que la mayoría de los trabajos previos se centraron en el formato scFv (Marzari et al., 2001; Rhyner et al., 2003), y también debido a la amplificación de la mayor cantidad posible de genes VH para generar una amplia diversidad, a diferencia de estudios anteriores que se enfocaron en ciertas familias VH como VH4 (Sblattero et al., 2000) o VH5 (Not et al., 2011).

Los resultados obtenidos mostraron una mayor expresión de genes VH135-IgG1 y Vk1, lo que confirma que la amplificación de cadenas se ha producido respetando la expresión que se suele

observar en la respuesta fisiológica humana, dado que el primer VH135 amplifica de manera conjunta la primera (VH3), tercera (VH1) y cuarta (VH5) familias de genes V de cadena pesada (Guigou et al., 1990) y la subclase (IgG1) más habituales en el suero sanguíneo (Nahm et al., 1987). La elevada expresión de anticuerpos de la subclase IgG2 (Figura 4.2.2SA) es una característica poco común, y sin explicación fisiológica o relacionada con la enfermedad celíaca, pero que se observa tanto en el suero del paciente 4 como en la amplificación, por lo que también se incluyó en el repertorio.

La elevada expresión de cadenas pesadas de clase IgA (Figura 4.2.2SA) no es habitual en el suero sanguíneo, pero es característica de la enfermedad celíaca, por lo que resulta fundamental trasladar estas cadenas al repertorio. Además, se ha descubierto que estos anticuerpos IgA séricos están directamente relacionados con los que se producen en el intestino (Iversen et al., 2017). Con respecto a las cadenas ligeras, también se amplifican con niveles de expresión similares a los observados por otros investigadores, destacando la presencia mayoritaria de Vk1 (Figura 4.2.2SC) (Guigou et al., 1990) y de VL2 (Figura 4.2.2SD) (Popov et al., 1999).

La segunda ronda de PCR solo se utilizó para expandir las secuencias adyacentes al punto de restricción (adición de 39 pb) con el fin de mejorar la eficiencia de la clonación (Figura 4.2.2SE). La amplificación se realizó con menos ciclos que la primera etapa (20 en vez de 35) para evitar la alteración de la composición en las cadenas.

Esta estrategia de amplificación y clonación se utilizó para evitar el protocolo más clásico de construcción de repertorios de Fab, mediante OE-PCR, que implicaba tres rondas de amplificación: la primera para amplificar los genes V y C a partir del ADNc, la segunda para el solapamiento de estos genes, y la tercera para la fusión de las cadenas ligeras y pesadas mediante un segundo solapamiento (Barbas, 2001). La metodología propuesta para este trabajo evita los solapamientos, al amplificar directamente las cadenas ligeras y pesadas de los Fab y clonarlas en el plásmido. Esta estrategia evita errores de PCR, representación incorrecta de las cadenas minoritarias, y solapamientos truncados (Omar & Lim, 2018).

Tras la clonación de los amplicones codificantes para cadenas pesadas y ligeras en el vector pComb3X, y la transformación de células de *E. coli* XL-1 Blue, utilizando la misma metodología que en el repertorio semiinmune, se obtuvo un RZ de  $1,1 \times 10^7$  transformantes. Se secuenciaron algunos clones al azar para asegurar que los Fab se clonaron correctamente, y se encontró una autoligación del 4 %. El valor de  $RZ > 10^7$  y el nivel de autoligación (< 5 %) se consideraron adecuados como para continuar con el protocolo de selección (Barbas et al., 2001). Otros investigadores han publicado repertorios funcionales con una gran variabilidad de tamaño, con valores de RZ de  $10^6$  a  $10^{11}$  (Bashir & Paeshuyse, 2020), por lo que los repertorios desarrollados durante la presente tesis doctoral presentan valores de RZ dentro de estos límites.

### 5.3.2 Selección por afinidad (*panning*) del repertorio inmune frente a la gliadina

El antígeno diana utilizado como selector de Fab del repertorio fue la gliadina- PWG, inmovilizada mediante adsorción en pocillos de placas de poliestireno para inmunoensayo.

Se realizaron tres rondas de *panning*, con astringencia creciente mediante un mayor número de lavados con PBS-T. El título de fago-anticuerpos eluidos aumentó 19 veces entre la primera y la segunda ronda, y luego se duplicó en la tercera ronda de selección (Figura 4.2.1A). Además, se realizó una caracterización parcial de la afinidad general de los fago-Fab seleccionados mediante un ELISA indirecto policlonal (Figura 4.2.1B), en el que se observó que la señal se triplicó después de la primera ronda, y alcanzó la máxima absorbancia después de la tercera ronda.

Integrando ambos resultados se observó que, a pesar de que no hubo un aumento significativo en el número de transformantes recuperados en la primera ronda, el ELISA policlonal mostró un claro aumento en la afinidad general. Por otro lado, el mayor número de transformantes recuperados en la segunda ronda no se vio afectado por un incremento notable en la afinidad, lo que podría explicarse porque se amplificó una amplia población de fago-anticuerpos con una mezcla de afinidades. En la tercera ronda, un fuerte aumento en las colonias recuperadas, así como en la señal del ELISA, significó que los fago-Fab de alta afinidad frente a la gliadina dominan en la población y que se podría aislar fácilmente del conjunto un ligando monoclonal. Es fundamental observar este aumento en alguna de las rondas, si no, la selección no está siendo efectiva (Barbas et al., 2001). Dependiendo del repertorio utilizado y el antígeno objetivo, este aumento puede aparecer en diferentes rondas (García-García et al., 2020; Madrid et al., 2020).

Las colonias aisladas de las placas de titulación de las últimas rondas se cultivaron para inducir la producción de fago-anticuerpos monoclonales. Con los fagos aislados se realizó un ELISA indirecto para seleccionar aquellos fago-Fab que presentaban una elevada afinidad por la gliadina-PWG (Figura 4.2.1C). De los 48 clones analizados, 29 produjeron señales de absorbancia superiores a 1,5 UA y no detectaron la solución de bloqueo (control negativo). Estos clones fueron seleccionados para continuar con su caracterización.

### 5.3.3 Secuenciación y análisis de los clones seleccionados

El ADN plasmídico de los 29 clones identificados como ligandos de la gliadina- PWG en el ensayo ELISA monoclonal con fago-Fab, se secuenció para detectar clones repetitivos, identificar clonotipos y clasificarlos por familia génica, clase y subclase. De los 29 clones secuenciados, solo se encontraron 14 secuencias únicas (Fab-A a Fab-M) (Figura 4.2.2A).

Una de las secuencias únicas (Fab-C) se repitió en 9 clones de la última ronda de selección, lo que representa el 31 % de las secuencias; Fab-H se encontró en 4 clones (14 %); Fab-N en 3 (10 %);

Fab-G y Fab-K dos veces cada una (6,89 % cada una) y las restantes secuencias se encontraron sólo una vez (3,34 % cada una). De los 14 clones diferentes, dos pertenecían a la subclase IgA2 (C y H), cinco a la IgG1 (A, D, G, L y N), cuatro a la IgG2 (B, E, J e I), y tres a la IgA1 (F, K y M) para la cadena pesada. En cuanto a las cadenas ligeras, hubo un dominio total de las cadenas kappa. Estas secuencias únicas se organizaron en clonotipos utilizando la herramienta IgBLAST. Los clones pertenecientes al mismo clonotipo son aquellos que presentan una reorganización con los mismos genes germinales V(D)J. Los 14 clones diferentes se clasificaron en 8 clonotipos para las cadenas pesadas: clonotipo 1H (B, E, I y J), clonotipo 2H (F, K y M), clonotipo 3H (A y G), clonotipo 4H (C), clonotipo 5H (H), clonotipo 6H (D), clonotipo 7H (L) y clonotipo 8H (N). Tres clonotipos (3H, 6H y 7H) presentaron genes VH1, cuatro VH3 (clonotipos 1H, 2H, 4H y 5H), y el clonotipo 8H presentó genes VH4 (Figura 4.2.2B).

La variabilidad de los genes V, D y J cuya reorganización produjo los Fab, se muestra en las Figuras 4.2.2B, 4.2.2C y 4.2.2D. Por un lado, la gran mayoría de los Fab reactivos a la gliadina presentaron genes VH3 (64,3 % de los clones), incluso superando la frecuencia encontrada en un repertorio normal de anticuerpos humanos (donde representan aproximadamente el 50 % de los genes secuenciados). Sorprendentemente, la familia más expresada en los anticuerpos humanos, VH3-23, solo estuvo representada por un clonotipo (2H), que incluyó tres Fab.

Debe destacarse especialmente la selección del Fab-C, que presenta el gen VH3-15, escasamente expresado (3 %) en el repertorio humano, pero que es uno de los genes más frecuentemente identificados en los anticuerpos reactivos al gluten en la enfermedad celíaca (Steinsbø et al., 2014). Por otro lado, aunque los anticuerpos VH4 representan casi el 30 % del repertorio humano, y también se ha publicado que son reactivos a la gliadina en pacientes celíacos (Sblattero et al., 2000), solo uno de los Fab seleccionados de este repertorio presentó esta subfamilia génica (Tiller et al., 2013).

La frecuencia de los genes V y J cuya reorganización produjo la cadena ligera de los Fab seleccionados, se representa en las Figuras 4.2.2F y 4.2.2G. La mayoría de los clonotipos en los Fab seleccionados presentaron V $\kappa$ 1 (78,6 %), que también es la variante más prevalente en el repertorio humano (aproximadamente el 45 % de las cadenas kappa) (Tiller et al., 2013).

La longitud de los HCDR3 de los Fab seleccionados fue notablemente alta ( $16 \pm 4,7$  aminoácidos), especialmente en dos clonotipos (1H y 4H) que presentaron más de 20 aminoácidos (Figura 4.2.2E). Esta característica es inusual, ya que el promedio en los anticuerpos humanos es de 14,8 aminoácidos (Joyce et al., 2020). Se encontró una alta variabilidad en la longitud y composición de los HCDR3, en contraste con los LCDR3, que son más cortos ( $8,5 \pm 0,76$ ) y menos diversos (Figura 4.2.2H).

Hasta dónde llega nuestra información, este es el primer repertorio de Fab generado a partir

de linfocitos B de pacientes celíacos con el fin de obtener anticuerpos para sistemas de detección de gluten en alimentos. El proceso de selección por afinidad realizado en este trabajo ha seguido los protocolos clásicos aplicados en los comienzos de la tecnología del *phage display* (Barbas et al., 2001), que han resultado ser efectivos sin necesidad de aplicar métodos más modernos y complejos, que implican secuenciación masiva o citometría de flujo (D'Angelo et al., 2014; Younger et al., 2017). Siguiendo este procedimiento, se seleccionaron anticuerpos claramente relacionados con la enfermedad celíaca (clase IgA) a partir del repertorio inmune de Fab.

### 5.3.4 Caracterización de los fago-Fab seleccionados con mejores características como sondas para la detección de gluten

La sensibilidad y especificidad de los clones individuales que presentaban afinidad por la gliadina se evaluaron mediante ELISA indirecto con fago-Fab. Se obtuvo una curva dosis-respuesta con concentraciones crecientes de gliadina- PWG (0 - 5  $\mu\text{g}/\text{mL}$ ) para cada variante de fago-Fab producida (Figura 4.3.3A). El Fab-C destacó sobre los demás Fab, con un intervalo de respuesta lineal entre 0 - 1,25  $\mu\text{g}/\text{mL}$ , alcanzando una meseta de saturación a 2,5  $\mu\text{g}/\text{mL}$ . Además, se realizó una prueba de especificidad con harinas de diferentes cereales para evaluar el rendimiento de los clones seleccionados en una matriz alimentaria real, y para discriminar las especies de cereales reconocidas por los diferentes fago-Fab. Tres Fab (Fab-J, Fab-F y Fab-K) mostraron reactividad cruzada con la avena, otros tres (Fab-J, Fab-F y Fab-L) con el maíz, y otros dos (Fab-J y Fab-F) con extractos de arroz. Estos clones no fueron considerados para una caracterización adicional, debido a su reactividad cruzada con cereales tolerados por personas que padecen patologías relacionadas con el gluten.

Los cuatro fago-Fab que presentaron mejores características (Fab-C, Fab-E, Fab-H y Fab-G) se utilizaron para analizar muestras más complejas. Con este fin, se analizaron las proteínas similares al gluten extraídas de una mezcla binaria experimental que contenía trigo, centeno y cebada (0; 0,1; 0,2; 0,5; 1; 2,5; 5; 10 y 100 mg/g) en una matriz libre de gluten a base de arroz, equivalente a 0; 10; 20; 50; 100; 250; 500; 1000 y 10.000 mg/kg de gluten (Figura 4.2.3B) (García- García et al., 2020). A partir de los resultados obtenidos, el clon Fab-C mostró claramente las mejores características: una elevada respuesta frente a la gliadina-PWG en solución (antígeno purificado) y selectividad frente a harinas que contienen gluten (trigo, centeno y cebada).

El Fab-C también mostró una alta capacidad para detectar el gluten en la mezcla binaria experimental, con un límite de detección (LOD) de 15 mg/kg de gluten (Figura 4.2.3C), por lo que permite la discriminación entre alimentos libres de gluten y aquellos que contienen gluten, según la legislación actual (Diario Oficial de la Unión Europea, 2014).

En el mercado se pueden encontrar otros métodos analíticos para la detección de gluten basados en anticuerpos que tienen un LOD más bajo que el Fab-C (Scherf & Poms, 2016). Sin embargo, se debe tener en cuenta que el fago-Fab-C es capaz de cumplir con los estándares legales de detección de gluten con una sola interacción paratópica, a diferencia de los otros anticuerpos monoclonales clásicos, en los que se producen dos interacciones, que pueden aumentar a cuatro en un formato de ELISA tipo sándwich, como en el método basado en el monoclonal R5 (Méndez et al., 2005).

Una vez seleccionado el Fab-C como la mejor sonda para la detección de gluten, el siguiente paso fue demostrar que podía aplicarse en un ELISA indirecto con fagos para cuantificar proteínas similares al gluten extraídas de productos comerciales, utilizando la mezcla binaria experimental como curva patrón (Figura 4.2.3C). Se analizaron diez productos comerciales, cinco de los cuales estaban etiquetados como libres de gluten y otros cinco declaraban contenerlo (Tabla 4.2.3). La media de tres determinaciones realizadas con el fago-Fab-C, se interpoló en la curva patrón, y se calculó la concentración final de gluten teniendo en cuenta todas las diluciones realizadas. Se consideraron como negativas (sin gluten) todas las muestras que presentaron valores inferiores al LOD o que se encontraban entre 15 (el LOD) y 20 mg/kg de gluten, ya que la legislación actual, establece que para etiquetar un producto como “sin gluten” debe contener menos de 20 mg/kg de este componente (Diario Oficial de la Unión Europea, 2014). Se consideraron positivos aquellos productos que presentaban más de 20 mg/kg determinados con la metodología propuesta. Todos los productos analizados estaban etiquetados de acuerdo con la legislación vigente, y los resultados fueron confirmados mediante un ELISA sándwich con el anticuerpo R5 (el método de referencia, según el *Codex Alimentarius*, para la detección de gluten en alimentos). La metodología propuesta de ELISA indirecto con fago- anticuerpos, solo se utilizó como prueba de concepto para determinar parcialmente las características analíticas (como sensibilidad y especificidad) de los Fab aislados de los repertorios. Esta metodología, aunque resulta útil, no es definitiva, porque los anticuerpos secundarios (anti-pVIII) presentan muchas reacciones cruzadas, y se puede sobreestimar la señal obtenida, porque estos anti-pVIII se pueden unir simultáneamente hasta 2400 veces a un fago (Hertveldt et al., 2009). Además, no todas las partículas de fagos producidas en el sobrenadante expresan el anticuerpo, por lo que también se podría sobreestimar el título y dilución de los fago-anticuerpos utilizados (Kiguchi et al., 2018). Por todas estas razones, fue necesario continuar con el desarrollo de algunos de los Fab seleccionados, en este caso como moléculas solubles, no unidas al fago.

## 5.4 PRODUCCIÓN SOLUBLE DE LOS FAB SELECCIONADOS Y DESARROLLO DE TÉCNICAS INMUNOENZIMÁTICAS ELISA PARA DETECCIÓN DE GLUTEN EN ALIMENTOS.

### 5.4.1 Producción recombinante a escala semiindustrial de los Fab seleccionados

Los cuatro fago-Fab con mayor afinidad por la gliadina, obtenidos mediante la tecnología de *phage display* a partir de los dos repertorios desarrollados en esta tesis doctoral, se seleccionaron para su expresión soluble y caracterización. Uno de los Fab (Fab8E-4) se aisló del repertorio semiinmune, y los otros tres (Fab-C, Fab-E y Fab-H) se seleccionaron del repertorio inmune.

Los anticuerpos seleccionados a partir del repertorio inmune presentaron la clase IgA, hecho que no fue sorprendente, ya que el repertorio se produjo a partir de linfocitos de sangre periférica de pacientes con enfermedad celíaca, que es una afección autoinmune mucosal en la que se generan anticuerpos IgA dirigidos frente al gluten (Koninckx et al., 1984). Esta situación dificulta la expresión soluble de los anticuerpos recombinantes obtenidos, dado que la clase IgA presenta un enlace disulfuro adicional, lo que plantea desafíos en la producción y plegamiento de la proteína recombinante (Gaciarz et al., 2016). El Fab8E-4 presentaba el isotipo IgG3, que también tiene una estructura que dificulta su expresión, en este caso, una región bisagra de gran tamaño (Spiteri et al., 2021). Para mejorar la expresión recombinante de los Fab, se cambió el isotipo original (encontrado después de su aislamiento de cada repertorio) a IgG1. Este cambio de isotipo permitió el uso del mismo anticuerpo secundario para diferentes Fab primarios, lo que aumentó la uniformidad de los ensayos.

Los Fab recombinantes se introdujeron en el vector de expresión pKKtac, que dirige las moléculas recombinantes al sobrenadante del cultivo, gracias a la clonación de las secuencias codificantes para cadenas de anticuerpos después de una señal de secreción *pelB*. Además, se eligió como biofactoría la cepa de *E. coli* RV308, que facilita la salida de los Fab al sobrenadante porque presenta una pared celular débil. Después de cultivar e inducir las colonias transformadas para la producción del Fab recombinante, se realizó un *Western blot* (Figura 4.3.2) para confirmar la expresión efectiva de los Fab en el sobrenadante antes de su purificación. Todos los sobrenadantes contenían bandas coincidentes con el tamaño de los Fab (aproximadamente 50 kDa) y con el tamaño del Fab utilizado como control. Además del Fab, en el gel se observaron otras bandas que correspondían a proteínas secretadas al medio por *E. coli*. Este método no permite la identificación absoluta de la presencia de la proteína recombinante buscada, pero se consideró adecuado para tomar la decisión sobre continuar la purificación.

Los fragmentos de anticuerpos, producidos como proteínas recombinantes con colas de histidina, se purificaron mediante cromatografía de afinidad con metal inmovilizado (IMAC) a partir del sobrenadante clarificado. El proceso de aislamiento se monitorizó mediante espectrometría

a 280 nm (Figura 4.3.3A). Cuando se activó el gradiente de imidazol, las proteínas unidas a la columna se eluyeron, produciendo dos picos de señal significativos que se analizaron mediante SDS-PAGE (Figura 4.3.3B). Las proteínas eluidas en el primer pico de absorbancia eran proteínas de *E. coli* unidas a la columna y las del segundo pico coincidían con el tamaño de los Fab. La identificación de los Fab se confirmó mediante un análisis de huella peptídica a partir de la banda del gel de poliacrilamida coincidente con el peso molecular de un Fab (aproximadamente 50 kDa). Por ejemplo, para el Fab-C, el análisis mostró una puntuación (una medida de la probabilidad de que el evento observado sea aleatorio) de 84, y el análisis estadístico indicó que las puntuaciones mayores de 13 son significativas ( $p < 0,05$ ) y confirman la identificación. Entre los péptidos detectados, había algunos que contenían parte de las regiones constantes y el HCDR3, lo que permitió una clara identificación del Fab-C. Estos datos confirmaron que los Fab se encontraron en fracciones eluidas a una concentración de aproximadamente 100 mM de imidazol. Para garantizar que se había obtenido el fragmento de anticuerpo deseado con un nivel de pureza y homogeneidad aceptable, la fracción se dializó para eliminar el imidazol. La estrategia de producción recombinante dio lugar a la obtención de entre 6 y 9 miligramos de cada anticuerpo, un rendimiento adecuado para la producción en *E. coli* utilizando cultivos discontinuos (matraces) (Frenzel et al., 2013). Además, la metodología propuesta ha demostrado ser una opción rápida (cuatro días desde el inicio del cultivo hasta la recuperación del Fab dializado), precisa (una vez optimizadas las condiciones, todos los lotes producidos fueron correctos) y reproducible (la producción ha sido satisfactoria en dos laboratorios diferentes) para obtener Fab recombinantes, lo que permite una alta uniformidad entre lotes y diferentes anticuerpos (De Marco, 2015).

#### 5.4.2 Validación de los Fab recombinantes como sondas para la detección de gluten.

La especificidad de los Fab frente al gluten se evaluó exhaustivamente mediante el análisis de extractos etanólicos de una amplia variedad de especies heterólogas. Como primer paso, se realizó un ELISA indirecto utilizando los Fab recombinantes purificados como anticuerpo primario, sobre placas de inmunoadsorción tapizadas con extractos de harinas de trigo, cebada y centeno como especies objetivo (por contener gluten), y harinas de maíz, avena y arroz como especies heterólogas (sin gluten) (Figura 4.3.4A). Los anticuerpos solubles no mostraron ninguna reactividad cruzada con las prolaminas de cereales de las especies heterólogas, a pesar de su estrecha relación filogenética. Posteriormente, se evaluó la sensibilidad del ELISA indirecto mediante el análisis de un material de referencia, la gliadina-PWG (Figura 4.3.4B). Los resultados mostraron que el Fab-C exhibió la mayor sensibilidad, seguido del Fab8E-4. Por el contrario, el Fab-E y el Fab-H demostraron una menor sensibilidad.

Además de verificar la capacidad de unión a la gliadina purificada, era imprescindible evaluar la funcionalidad de los Fab para detectar el gluten en una matriz más compleja, que se asemeje a la aplicación final, la detección de gluten en alimentos. Con este fin, se realizó un ensayo de ELISA para detectar gluten en una mezcla compuesta por concentraciones crecientes de harinas de

trigo, cebada y centeno, mezcladas en una matriz de harina de arroz (que no contiene gluten). Los resultados coincidieron con las respuestas obtenidas de la curva de gliadina-PWG (Figura 4.3.4C).

### 5.4.3 Desarrollo de una metodología de ELISA indirecto con Fab soluble para detección de gluten en alimentos

Los resultados anteriores demostraron que el Fab-C exhibió la mejor respuesta frente al gluten en ensayos tanto con la gliadina-PWG purificada, como con proteínas similares al gluten extraídas de una mezcla experimental de harinas, por lo que se eligió como sonda para desarrollar un ELISA indirecto y se realizó una evaluación de la especificidad, sensibilidad y exactitud de la prueba.

En primer lugar, se realizó un análisis más exhaustivo de la especificidad, evaluando posibles reacciones cruzadas frente a 60 especies heterólogas (Tabla 4.3.1). Los resultados demostraron que el Fab-C no presentaba reactividad cruzada con estas matrices, lo que indica su alta especificidad para la detección de gluten. Estos hallazgos sugieren que el Fab-C podría ser una herramienta fiable y precisa para la inmunodetección de gluten en productos alimenticios, ya que puede diferenciar de manera efectiva entre muestras que contienen gluten y muestras sin gluten.

En segundo lugar, la sensibilidad del ELISA indirecto basado en el Fab-C se evaluó utilizando un modelo de regresión lineal con concentraciones crecientes de gliadina-PWG, que iban desde 0,025 hasta 0,625  $\mu\text{g}/\text{mL}$  (Figura 4.3.4D). El límite de detección (LOD) se calculó como 0,028  $\mu\text{g}/\text{mL}$  mediante interpolación en la ecuación del triple de la desviación estándar de diez blancos. Esto corresponde a una concentración de gluten de 11 mg/kg en las muestras (considerando un factor de dilución de 5, un factor de extracción de 40 (0,25 g de muestra extraída con 10 mL de tampón) y un factor de conversión de gliadina a gluten de 2). El límite de cuantificación (LOQ) se determinó interpolando diez veces la desviación estándar del blanco, siendo 0,05  $\mu\text{g}/\text{mL}$  de gliadina-PWG, equivalente a 19,8 mg/kg de gluten. Estos resultados confirman que el ELISA indirecto utilizando el anticuerpo recombinante Fab-C puede diferenciar claramente los productos sin gluten (menos de 20 mg/kg de gluten) de los que contienen gluten, según los límites establecidos por la legislación actual (Diario Oficial de la Unión Europea, 2014).

La sensibilidad del ensayo es ligeramente inferior que la del ELISA sándwich R5, utilizando el mismo patrón (gliadina-PWG), aunque se debe tener en cuenta que las condiciones y el formato del ensayo son diferentes. El ELISA indirecto con el Fab-C muestra una sola interacción por molécula de antígeno, porque el Fab posee solo un parátopo, lo que permite un solo evento de unión por Fab. Por otro lado, el método de ELISA sándwich R5 implica dos interacciones moleculares con el antígeno, debido a la presencia de dos parátomos por molécula de anticuerpo completo, y a la utilización de un formato de tipo sándwich directo, que permite interacciones duales que involucran la captura y la detección del antígeno (Amnuaycheewa et al., 2022). Teniendo en

cuenta estos factores, se demuestra que la molécula de un solo parátipo (Fab-C) exhibe una alta afinidad hacia el antígeno objetivo.

Se realizó un análisis de recuperación de gluten mediante un ELISA indirecto con el Fab-C, utilizando proteínas similares al gluten (prolaminas) que fueron extraídas de harina de arroz certificada sin gluten, enriquecida con gliadina-PWG sólida (equivalente a 20 y 100 mg/kg de gluten teórico). El contenido en gluten de estas mezclas se evaluó tomando como referencia los resultados obtenidos con el ELISA de tipo sándwich R5. Los resultados de la prueba de recuperación para la metodología propuesta fueron del  $76,71 \pm 1,72$  % para la mezcla de 20 mg/kg y del  $90,89 \pm 0,36$  % para la mezcla de 100 mg/kg. Como control positivo, se enriqueció un extracto de harina de arroz con las cantidades equivalentes de gliadina-PWG disuelta ( $0,05 \mu\text{g/mL}$ , equivalente a 20 mg/kg en muestras de alimentos, y  $0,25 \mu\text{g/mL}$ , equivalente a 100 mg/kg). Los resultados para estos controles fueron del  $79,93 \pm 3,35$  % para la solución de  $0,05 \mu\text{g/mL}$  y del  $92,11 \pm 1,07$  % para la solución de  $0,25 \mu\text{g/mL}$ . Todos los resultados obtenidos presentaron un límite de recuperación adecuado, en el intervalo del 75 al 125 % siguiendo los estándares de la Unión Internacional de Química Aplicada y Pura (IUPAC) (Thompson et al., 1999).

#### 5.4.4 Desarrollo de una metodología de ELISA sándwich con Fab recombinantes para detección de gluten en alimentos

Para evaluar el potencial de los Fab recombinantes generados para su aplicación en un ELISA de tipo sándwich, los anticuerpos se modificaron químicamente mediante una unión covalente a biotina. La biotinylación tenía un doble propósito: la inmovilización en placas recubiertas con estreptavidina para utilizarlos como anticuerpos de captura, y facilitar la conjugación con la enzima fosfatasa alcalina para su uso como anticuerpos de detección. Después de la modificación química, se evaluó la compatibilidad de diferentes pares de Fab, utilizando  $10 \mu\text{g/mL}$  de gliadina-PWG como antígeno, probando todas las combinaciones de anticuerpos. La combinación del Fab8E-4 como anticuerpo de captura y el Fab-C conjugado con fosfatasa alcalina como anticuerpo de detección produjo los mejores resultados (Tabla 4.4.1). La utilización del mismo Fab como anticuerpo de captura y de detección no resultó adecuada, ya que los valores de absorbancia no fueron tan altos como se esperaba para la concentración de antígeno aplicada (saturante con Fab no conjugado). Esto podría explicarse debido a que un mismo anticuerpo con diferentes usos (captura y detección) compite por los mismos epítopos, disminuyendo la respuesta global. Sin embargo, el anticuerpo monoclonal R5 sí se utiliza como captura y detección en el mismo ensayo de tipo sándwich, debido a que se une a regiones pentaméricas altamente repetitivas en la molécula del gluten (Kahlenberg et al., 2006), permitiendo una cuádruple interacción con el antígeno en el formato ELISA sándwich (dos parátipos por anticuerpo, de captura y detección) que aumenta su sensibilidad.

Conviene señalar que se observó una disminución en la sensibilidad de los Fab biotinylados con relación a los no modificados. Se teorizó que la fuerte modificación química causó cambios

estructurales en los anticuerpos, teniendo en cuenta que la biotinylación se podía producir en todos los grupos amino libres (de media hasta 24 en un Fab). Lo ideal sería una única biotinylación en la región constante del Fab, dado que otras uniones podrían modificar la estructura y plegamiento del Fab, resultando en una pérdida de funcionalidad (Fairhead & Howarth, 2015). Para abordar este problema, se propuso un ELISA de tipo sándwich indirecto sin modificación química, mediante la adsorción del anticuerpo de captura (Fab8E-4) en la placa inmunoabsorbente, el Fab-C como anticuerpo primario, y un anticuerpo secundario anti-Fab humano conjugado con peroxidasa de rábano (el mismo utilizado para el ELISA indirecto basado en el Fab-C).

La metodología de ELISA sándwich indirecto fue validada con los mismos ensayos que el ELISA indirecto basado en el Fab-C: no se observó reactividad cruzada con ninguna de las 60 especies heterólogas analizadas. La sensibilidad se evaluó a través de un modelo de dosis-respuesta utilizando concentraciones crecientes de gliadina-PWG, que iban desde 0,025 hasta 25  $\mu\text{g}/\text{mL}$  (Figura 4.4.1). El límite de detección (LOD) se calculó como 0,026  $\mu\text{g}/\text{mL}$  mediante interpolación del triple de la desviación estándar de diez blancos, que se traduce en una concentración de gluten de 10,4 mg/kg. El límite de cuantificación (LOQ) se estableció interpolando diez veces la desviación estándar de diez blancos, lo que resultó en un valor de 0,044  $\mu\text{g}/\text{mL}$  de gliadina-PWG, equivalente a 17,4 mg/kg de gluten. Los resultados del análisis de recuperación fueron de  $104,14 \pm 2,2 \%$  para la mezcla de 20 mg/kg y de  $95,44 \pm 6,36 \%$  para la mezcla de 100 mg/kg. Para establecer una línea base de control, se preparó un extracto de harina de arroz enriquecido con cantidades equivalentes de gliadina-PWG disuelta (0,05  $\mu\text{g}/\text{mL}$ , correspondiente a 20 mg/kg en muestras de alimentos, y 0,25  $\mu\text{g}/\text{mL}$ , equivalente a 100 mg/kg). Los resultados para estas muestras de control demostraron porcentajes de recuperación de  $100,08 \pm 3,67 \%$  para la solución de 0,05  $\mu\text{g}/\text{mL}$  y de  $95,44 \pm 6,36 \%$  para la solución de 0,25  $\mu\text{g}/\text{mL}$ .

#### 5.4.5 Comparación entre las dos metodologías desarrolladas

Por un lado, el ELISA indirecto ofrece una ventaja en términos de velocidad, ya que involucra un paso menos que el método de tipo sándwich. Por otro lado, la metodología de tipo sándwich ha demostrado una mayor sensibilidad, con un límite de detección ligeramente más bajo, así como un límite de cuantificación más reducido (17 mg/kg para el método de tipo sándwich en comparación con 20 mg/kg para el método indirecto). Este fenómeno se puede explicar considerando el número de interacciones moleculares. Si bien ambos formatos permiten, como mínimo, una interacción por molécula de antígeno (en términos de avidéz), el formato de tipo sándwich exhibe una mayor avidéz global debido a la provisión de dos interacciones parátipo-epítipo, en contraste con el único parátipo en el formato indirecto. Se produce un fenómeno similar en el método sándwich-R5, pero en este caso hay una doble interacción por antígeno, que implica hasta cuatro parátipos entre los dos anticuerpos participantes (Scherf & Poms, 2016).

En términos de exactitud, el análisis de recuperación reveló mejoras notables con el formato de tipo sándwich, donde las cantidades de gluten recuperadas se encontraron dentro de un intervalo estrecho del 95 % al 105 % en comparación con la concentración teórica. En contraste, el ELISA indirecto mostró una variabilidad más amplia, con resultados de recuperación que oscilaron entre el 76 % y el 92 %.

#### 5.4.6 Análisis de muestras de alimentos con los sistemas de ELISA desarrollados.

La metodología empleada en este estudio tuvo como objetivo determinar la especificidad y sensibilidad de los ensayos ELISA para la detección de gluten extraído de productos alimentarios, lo que representa el objetivo final de esta tesis doctoral.

Las muestras se clasificaron en cuatro categorías diferentes: A) productos que declararon contener gluten en el etiquetado (9 muestras); B) productos con etiquetas de precaución que indican “puede contener” gluten (6 muestras); C) productos que no declararon gluten o que no advirtieron específicamente sobre la presencia de gluten (21 muestras); y D) productos con etiquetas y/o certificación de “libre de gluten” (14 muestras).

La media de tres determinaciones de cada muestra, realizadas con las técnicas ELISA desarrolladas (indirecto y sándwich), se interpoló en la curva patrón, y se calculó la concentración final de gluten teniendo en cuenta todas las diluciones realizadas. Se consideraron como negativas (sin gluten) todas las muestras que presentaron valores inferiores al LOD o que se encontraban entre el LOD y 20 mg/kg de gluten, de acuerdo con la legislación actual, que establece que para etiquetar un producto como “sin gluten” debe contener menos de 20 mg/kg de este componente (Diario Oficial de la Unión Europea, 2014). Se consideraron positivos aquellos productos que presentaban más de 20 mg/kg determinados con la metodología propuesta.

Las muestras de alimentos también se analizaron utilizando la técnica estándar para la detección de gluten en alimentos incluida en el *Codex Alimentarius*, el ELISA sándwich directo basado en el anticuerpo monoclonal R5.

Los resultados se agrupan en la Tabla 4.3.2 para el indirecto y en la Tabla 4.4.2 para el sándwich. Los tres métodos mencionados mostraron una buena concordancia entre ellos, que permitió la clasificación de las muestras de alimentos estudiadas siguiendo las directrices de la legislación europea. De los 50 productos analizados, los resultados de 47 concuerdan en ambos formatos de ELISA (indirecto y sándwich) y con la técnica basada en R5. Sin embargo, tres productos a base de avena fueron considerados positivos con el ELISA R5 (resultando en más de 20 mg/kg de gluten), pero no se observó reactividad en ninguno de los ELISAs desarrollados con los Fab recombinantes, a pesar de que ambos formatos de ELISA presentan una capacidad fiable para discriminar productos sin gluten según la legislación en vigor.

La completa ausencia de reactividad cruzada con la avena es una característica destacable de los dos formatos ELISA (indirecto y sándwich) desarrollados en este trabajo. Otras investigaciones han demostrado que algunas variedades específicas de avena muestran reactividad cruzada cuando se analizan con otros anticuerpos de detección como el R5, revelando niveles aparentes de gluten de hasta 100 mg/kg en ciertas avenas (Benoit et al., 2017). La reacción cruzada del anticuerpo monoclonal R5 se debe a la sustancial similitud entre ciertos péptidos encontrados en la avenina y en la gliadina (Hardy et al., 2015). La ausencia de esta detección involuntaria en productos derivados de la avena etiquetados como sin gluten puede indicar la eficacia de los Fab-C y Fab8E-4, convirtiéndolos en herramientas valiosas para evaluar el contenido de gluten en tales productos, especialmente aquellos que contienen variedades de avena que provocan reactividad cruzada con otros anticuerpos disponibles en el mercado.

La identificación del gluten en productos alimenticios sigue siendo un gran desafío en el campo de la ciencia alimentaria, debido al hecho de que el gluten se define en función de sus propiedades químicas, como la solubilidad en etanol. A diferencia de otros alérgenos alimentarios, que suelen estar representados por una o unas pocas proteínas bien definidas, el gluten abarca una amplia variedad de moléculas. Estas proteínas comparten ciertas características comunes, pero pueden variar significativamente en sus secuencias de aminoácidos (Biesiekierski, 2017). Esta circunstancia hace que la búsqueda de un método universal de detección de gluten sea excepcionalmente difícil. Numerosas investigaciones comparativas han revelado disparidades significativas entre los kits disponibles comercialmente cuando se trata de cuantificar el gluten (Bugyi et al., 2013). Estas discrepancias se pueden atribuir a variaciones en los epítomos del gluten reconocidos por cada anticuerpo. Por ejemplo, el anticuerpo R5 identifica predominantemente epítomos que son más comunes en el centeno, la cebada o el triticale, en comparación con el trigo, mientras que el anticuerpo monoclonal G12 tiene una mayor capacidad de detección de epítomos del trigo (Lexhaller et al., 2017). Los resultados obtenidos con los Fab desarrollados para la presente tesis doctoral han demostrado que tienen capacidad de detectar proteínas de gluten tanto de trigo, como de cebada y de centeno, a diferencia de los monoclonales clásicos que presentan más sensibilidad por unas especies que por otras.

Además, otros trabajos también han demostrado que la cuantificación del gluten se ve influenciada por las especies y variedades de cereal analizadas, mostrando diferencias significativas al detectar la misma variedad de trigo con diferentes anticuerpos, como policlonales o R5 y G12 (Schopf & Scherf, 2018). Por lo tanto, aunque el método validado por el *Codex Alimentarius* (el ELISA sándwich R5) es ampliamente utilizado, no se puede considerar una forma de cuantificar absoluta y definitiva del gluten en alimentos. En consecuencia, es necesario considerar métodos alternativos y complementarios, basados en anticuerpos diferentes, como G12,  $\alpha$ -20, el dab-8E o los desarrollados durante la presente tesis doctoral, para garantizar la seguridad alimentaria de las personas sensibles al gluten.

Además de la complejidad de la detección derivada de la variabilidad de las proteínas del gluten y las sondas de detección disponibles, el establecimiento de límites de gluten en alimentos es igualmente problemática. El límite de 20 mg/kg de gluten se considera seguro por la legislación alimentaria mundial, salvo por la australiana y neozelandesa, que consideran un producto sin gluten cuando este no se puede detectar por el método R5, que ya se ha comentado que presenta limitaciones. A partir del descubrimiento de que el daño intestinal en la enfermedad celíaca es dependiente de la dosis de gluten, se intentó marcar un umbral de consumo seguro (Catassi et al., 1993). El límite de 20 mg/kg se estableció a partir de un extenso estudio clínico, en el que se descubrió que los pacientes que consumían menos de 50 mg de gluten al día no padecían daños intestinales, y que esto se traducía en el límite de 20 mg/kg (Cohen et al., 2019). Este estudio presenta algunas debilidades, ya que, a pesar de que se incluyeron pacientes de los tres grupos etiopatológicos de enfermedades relacionadas con el gluten, no se cubrieron todas las patologías concretas, además de que se centraba en la medición de daño intestinal, aunque se sabe que existen muchas manifestaciones extraintestinales (Losurdo et al., 2018; Therrien et al., 2020), y sólo tuvo en cuenta población de los Estados Unidos. Otros estudios poblacionales en otros lugares arrojaron datos diferentes. Por ejemplo, un estudio en la región de Carelia (dividida entre la Federación Rusa y la República de Finlandia, con población de etnicidad similar, pero con grandes diferencias económicas y sociales) demostró que además de la cantidad de gluten consumida hay otros factores determinantes en la prevalencia y daño de estas patologías (Kondrashova et al., 2008). Otras publicaciones, también basadas en estudios poblacionales, critican el umbral establecido, dado que encontraron que ingestas de tan sólo 1,5 mg de gluten podían causar daño intestinal (Chartrand et al., 1997).

Todas estas razones hacen necesario generar un debate profundo sobre la seguridad alimentaria para las personas sensibles al gluten, revisar los límites umbrales para considerar un producto sin gluten (idealmente para cada población), redefinir el concepto de gluten mediante la gran cantidad de información sobre sus efectos que se conoce actualmente, realizar análisis más completos de alimentos con sondas más variadas, y establecer sistemas de control y trazabilidad a través de toda la cadena de valor, desde los productores primarios hasta el consumidor final.

## 5.5 MODELIZACIÓN *IN SILICO* DE LA ESTRUCTURA E INTERACCIÓN DE ALGUNOS DE LOS FAB SELECCIONADOS Y SU INTERACCIÓN CON LA GLIADINA

Conocer la forma en la que interaccionan los anticuerpos y sus antígenos correspondientes, aporta una información valiosa a nivel molecular sobre el proceso de detección llevado a cabo en las metodologías ELISA desarrolladas, lo que permite una mejor caracterización de los procesos a su nivel más básico, y la posibilidad de introducir mejoras. Elucidar la interacción entre antígenos y anticuerpos no es una tarea trivial, principalmente porque es necesario conocer las estructuras

tridimensionales de ambas moléculas, algo complejo en sí mismo. Existen varias técnicas físicas como la cristalografía de rayos X o la resonancia magnética nuclear, muy útiles para determinar la estructura de proteínas, pero se trata de técnicas complejas que necesitan disponer de grandes cantidades de las proteínas con elevada pureza (Canduri & de Azevedo, 2008). Esta tarea se ha facilitado enormemente, gracias a la mejora continua de los sistemas de predicción de estructura tridimensional de proteínas a partir de su secuencia de aminoácidos, especialmente desde la implementación de predictores mediados por inteligencia artificial, que han permitido aproximarse a este problema con una rapidez y precisión nunca vistas hasta ahora.

### 5.5.1 Modelización de la estructura e interacción del Fab8E-4

Se realizaron predicciones *in silico* de la estructura del Fab8E-4, y de la interacción anticuerpo-antígeno para caracterizar mejor el fragmento de anticuerpo obtenido.

A partir de los resultados de secuenciación de Fab8E-4, se generó una predicción de la estructura tridimensional de la molécula completa utilizando AlphaFold2, un programa de inteligencia artificial que realiza predicciones de estructura de proteínas con alta precisión (Jumper et al., 2021). Se generaron cinco modelos diferentes, y cada modelo se recicló tres veces. Los modelos se clasificaron según la puntuación pLDDT (una estimación de confianza por residuo en una escala de 0 a 100). El quinto modelo (Figuras 4.1.5A y 4.1.5B) se seleccionó como el más fiable, con una puntuación pLDDT general de 82,3, que se consideró una buena predicción (Monzon et al., 2022). Además, hay que tener en cuenta que el área peor resuelta fue la cola de histidinas unida al dominio CH1, que disminuyó la puntuación general, mientras que los dominios del Fab8E-4 se resolvieron bastante bien, con una puntuación pLDDT > 85.

Además del modelo de la molécula completa, se generó un segundo modelo del Fab, centrado en el Fv (fragmento variable del anticuerpo, compuesto por los dominios VL y VH), que contiene la principal superficie de interacción antígeno- anticuerpo. Este modelo se calculó utilizando la herramienta AbodyBuilder (Leem et al., 2016) incluida en el servidor sAbPred (Figura 4.1.5C). Esta herramienta bioinformática abarca hasta siete algoritmos diferentes para determinar la estructura más probable del anticuerpo, como la identificación de los mejores modelos para los dominios VH y VL, la predicción de la orientación VH-VL o el modelado de los bucles de los CDR y sus residuos laterales. Las estructuras (extraídas de la base de datos RCSB PDB) que se eligieron como referencia fueron 6ghg (VH) y 6co3 (VL), ya que mostraron la mayor identidad de secuencia para cada cadena (86.2 % de identidad de secuencia para VH y 93.4 % para VL). La orientación VH-VL se predijo utilizando los ángulos calculados a partir de la estructura de referencia 2uzi. Una vez obtenido el modelo 3D del Fab8E-4, se procedió a la predicción de epítomos y parátomos para investigar la superficie de interacción antígeno-anticuerpo. Para identificar los principales aminoácidos involucrados en la región de unión, se requiere conocimiento de la estructura 3D tanto del antígeno como del Fv del Fab. Sin embargo, no se disponía de estructuras generadas por

técnicas físicas de las diferentes proteínas del gluten en las bases de datos como Uniprot o RCSB PDB (que solo incluyen estructuras de algunos péptidos lineales presentados por el Complejo Mayor de Histocompatibilidad (CMH) que no eran útiles para el presente trabajo). A pesar de la falta de datos experimentales sobre las estructuras del gluten, recientemente se han resuelto las estructuras de algunas gliadinas y otras proteínas relacionadas con el gluten utilizando el sistema de inteligencia artificial AlphaFold2 y se han cargado a la base de datos Uniprot. Por lo tanto, en este estudio se realizaron predicciones de epítomos con un algoritmo de *docking* modificado, denominado Epipred (Krawczyk et al., 2014). En cambio, para predecir los parátomos se utilizó el programa Antibody-ipatch (Krawczyk et al., 2013). En ambos casos, (predicción de epítomos y parátomos) se tomó como base la estructura 3D del anticuerpo generada por AbodyBuilder y la estructura 3D del antígeno de la  $\alpha/\beta$ -gliadina de trigo y la  $\gamma$ -hordeína de cebada resueltas por AlphaFold2 (entradas P18573 y P17990 de Uniprot). Hasta que la estructura de las gliadinas no se resuelva completamente mediante cristalografía de rayos X o criomicroscopía, consideramos que este enfoque es la aproximación más adecuada para la identificación de epítomos y parátomos del Fab8E-4.

Los modelos predictivos de epítomos representados en las Figuras 4.1.6A y 4.1.6B muestran que los epítomos conformacionales se encuentran principalmente en una  $\alpha$ -hélice en la región C-terminal de ambos antígenos ( $\alpha/\beta$ - gliadina y  $\gamma$ -hordeína). Dentro de las gliadinas estudiadas, la zona principal de interacción del dAb-8E y el Fab8E-4 se encuentra después de una de las zonas más patogénicas relacionadas con la enfermedad celíaca, el péptido inmunogénico 33-mer (Shan et al., 2002). Además, esta región presenta repeticiones de glutaminas más dispersas y más cortas que las encontradas en la mitad N-terminal. En este sentido, el Fab8E-4 y su fragmento de anticuerpo parental, el dAb-8E, constituyen una nueva alternativa para la determinación del gluten, proporcionando contactos con la región C-terminal de los antígenos, lo que podría complementar a los anticuerpos disponibles en el mercado, que generalmente se unen a la región N-terminal. Esto abre la posibilidad de nuevos desarrollos de ELISA de tipo sándwich o competitivos con anticuerpos que se unen a diferentes epítomos. La gliadina también presenta un epítomo muy largo en las posiciones 179 a 190, incluye 5 glutaminas. Este epítomo de 11 aminoácidos se encuentra en diferentes especies, variedades y cruces del género *Triticum*, haciendo al Fab8E-4 especialmente reactivo al trigo y sus derivados, como le ocurre al monoclonal G12 y, a diferencia del R5, que se une más fuertemente a epítomos del centeno (Lexhaller et al., 2017).

Como se observa en los modelos de los parátomos mostrados en la Figuras 4.1.6C y 4.1.6D, independientemente de la prolamina utilizada en el análisis, la interacción principal antígeno-anticuerpo se mantuvo en la cadena pesada del Fab8E-4, que se deriva directamente del dAb-8E. Los residuos más propensos a estar en contacto con el antígeno se encontrarían en el HCDR2 y las secuencias adyacentes (Figura 4.1.6C). Sin embargo, la reformulación del dAb- 8E agregó nuevos contactos fuertes con el antígeno a través de varios residuos de la cadena ligera (A56, Y38 y A114).

Además, el puente disulfuro entre los dominios constante y variable fijó la estructura, reduciendo la libertad conformacional de la proteína (Torres & Casadevall, 2008), lo que presumiblemente podría haber mejorado las características del anticuerpo.

### 5.5.2 Modelización de la estructura e interacción del Fab-C

Para estudiar la interacción molecular entre el Fab-C y la gliadina, se consideró que un enfoque *in-silico* era el método más adecuado, debido a que, al igual que con el Fab8E-4, no hay datos experimentales disponibles de estructuras de alta resolución (obtenidas por cristalografía de rayos X o criomicroscopía) ni para el antígeno (gliadina) ni para el anticuerpo (Fab-C).

El primer paso para modelar la interacción fue obtener las estructuras de los antígenos. Se empleó el sistema de predicción de estructura de proteínas, AlphaFold2 porque es el que presenta una mejor precisión para obtener la estructura tridimensional de proteínas de cadena única, como la gliadina (David et al., 2022). Como ya se ha comentado, recientemente se resolvió la estructura de la  $\alpha/\beta$ -gliadina de trigo, que tiene 307 aminoácidos y un peso molecular de 35 kDa (Uniprot ID: P18573 o GDA9\_WHEAT), utilizando Alpha-Fold2. Estas proteínas se consideran intrínsecamente desordenadas (Markgren et al., 2020) y el enfoque *in silico* es, hasta la fecha, la mejor solución para conocer su estructura.

La caracterización de la estructura del Fab enfrentó varios desafíos. Un primer intento de resolver la estructura utilizando el programa AlphaFold2 mostró muchas inconsistencias, lo cual podría deberse a que el Fab es una proteína multcadena. Por ello, se utilizaron herramientas alternativas basadas en sistemas de aprendizaje profundo entrenadas con estructuras de anticuerpos, como IgFold (Ruffolo & Gray, 2022) y AbodyBuilder2 (Abanades et al., 2023). Para comparar ambos resultados, las estructuras se superpusieron con una herramienta de alineación de ChimeraX (Pettersen et al., 2021) llamada Matchmaker, visualizando una discrepancia moderada entre los modelos con respecto a la distribución espacial del HCDR3 (Figura 4.2.4A). En un intento por verificar qué modelo podría estar más cercano a una estructura experimental real, se realizaron búsquedas en la base de datos SabDab (Schneider et al., 2021) de moléculas que contuvieran un HCDR3 con una longitud (de 20 a 23 aminoácidos) y secuencias de aminoácidos similares al Fab-C, pero no se encontraron. Por lo tanto, ambos modelos se consideraron inicialmente válidos para continuar trabajando con ellos.

Para identificar los epítomos reconocidos por el Fab-C, se evaluaron las zonas de la molécula de gliadina con mayor probabilidad de establecer contactos con el Fab. Se utilizó la herramienta EpiPred para la identificación de contactos entre cada uno de los modelos del Fab-C y el antígeno, obteniendo los mismos resultados, independientemente del modelo de Fab-C utilizado. Los contactos que tienen más probabilidades de ocurrir están ubicados en el extremo C- terminal, con tres “zonas calientes” de varios aminoácidos que probablemente formen parte del epítomo: 153 a 163; 200 a 207 y 253 a 261 (Figura 4.2.4B).

Una vez obtenidos los modelos estructurales adecuados (a partir de una base de datos para el antígeno y generados *de novo* para el Fab) e identificados los epítomos más probables, se calculó un modelo de interacción mediante acoplamiento utilizando la herramienta HADDOCK 2.4 (Van Zundert et al., 2016), ajustando las zonas de contacto entre los CDRs de ambos modelos de Fab y los epítomos predichos en el antígeno. Se calcularon varios modelos de acoplamiento y los mejores se identificaron por sus puntuaciones de calidad: para el modelo IgFold (RMSD =  $0,6 \pm 0,3$  Å; Z-score = -1,2) y para el modelo Abodybuilder2 (RMSD =  $5,8 \pm 0,5$  Å; Z-score = -2) (Kufareva & Abagyan, 2011).

Basándose en estas puntuaciones, la interacción antígeno-Fab se modela mejor utilizando el modelo de Fab de IgFold. Gracias a este modelo de alta calidad, se identificaron varios puntos de interacción probables entre el antígeno y el anticuerpo. Como era de esperar, los principales parátomos se encuentran en la cadena pesada. El HCDR1 (Figura 4.2.4C) interactúa con residuos de prolina y glutamina (posición 250-251). Este patrón PQ se ha identificado en varios péptidos relacionados con la toxicidad celíaca (Morón et al., 2008). El HCDR2 (Figura 4.2.4D) interactúa principalmente con glutaminas en varias posiciones dentro de  $\alpha$ -hélices (posiciones 203, 207, 251 y 254). Esto podría significar una fortaleza de este Fab, ya que las proteínas relacionadas con el gluten son muy ricas en este aminoácido. La anormal longitud del HCDR3 (Figura 4.2.4E) le permite tener un doble contacto con el antígeno. El lado proximal del bucle de HCDR3 (posiciones 109-114) podría interactuar con las posiciones 153-157 del antígeno (con un par de glutaminas) y el lado distal (posiciones 105-109) con los residuos asparagina 257, leucina 258 y glutamina 261 en una hélice paralela con respecto a la anterior. Aunque la contribución de la cadena ligera a la interacción parece ser mucho más débil, existe un posible contacto entre el LCDR3 y la glicina 159 y arginina 160 (Figura 4.2.4F), que probablemente participe en la estabilidad general de la unión. Aunque estos resultados son simulaciones, gracias a la calidad del procedimiento de acoplamiento (*docking*), se puede concluir que el Fab-C es un buen candidato para la unión al dominio C-terminal de la gliadina, algo inusual en comparación con otros anticuerpos de detección de gluten, como el R5 o el G12 (Röckendorf et al., 2017), que se unen en la región N-terminal. Esto podría ser útil para el desarrollo de inmunoensayos, ya que es muy probable que el Fab-C y otros anticuerpos disponibles podrían combinarse en un inmunoensayo con características mejoradas debido a una detección de múltiples epítomos, pudiéndose utilizar los Fab reactivos frente al extremo C-terminal desarrollados en la presente tesis doctoral, como anticuerpos de captura, y alguno de los anticuerpos comerciales que se unen al extremo N-terminal, como el R5 o el G12, como elementos de detección en un ELISA sándwich, o viceversa.

Los anticuerpos obtenidos en este trabajo que muestran mejores características para la detección de gluten (Fab-C y Fab8E-4) se unen (según sus modelos *in silico*) al extremo C-terminal de la gliadina, pero en posiciones distintas. El Fab-C, al presentar un HCDR3 mucho más largo que el Fab8E-4, interacciona con más aminoácidos y estas diferencias hacen que no sólo sean compatibles en un ELISA sándwich, sino sinérgicos.

## 5.6 CARACTERIZACIÓN DE LOS REPERTORIOS Y PROCESOS DE SELECCIÓN DE FAB MEDIANTE TÉCNICAS DE SECUENCIACIÓN MASIVA

La secuenciación por el método automático de Sanger permitió conocer las secuencias de los Fab seleccionados a partir de los repertorios generados, iniciando su caracterización. Sin embargo, era conveniente llevar a cabo un análisis más profundo del proceso de selección, que ayudara también a conocer qué características de la respuesta inmune celíaca se habían trasladado al repertorio inmune. Todo esto fue posible gracias a la aparición en el mercado de los dispositivos de secuenciación MinION de Oxford Nanopore®, que permitieron llevar a cabo experimentos de secuenciación masiva, a un coste razonable, en nuestro propio laboratorio.

### 5.6.1 Análisis primario

A partir del sistema de secuenciación Nanopore se recopilaban 4,68 millones de lecturas en formato FAST5 y se transformaron en archivos FASTQ demultiplexados y agrupados por un código de barras para diferenciar las secuencias de cada repertorio y ronda de selección. Se recopilaban un promedio de 0,5 millones de secuencias por cada código de barras, con una gama de longitudes desde 120 hasta 12.400 bases, algo habitual con esta técnica (Wang et al., 2021). Este amplio espectro de lecturas hacía necesario implementar un proceso de filtrado. Se evaluó la calidad de las secuencias mostrando una distribución relativamente uniforme en todas las lecturas (Figura 4.5.2A). El análisis de la distribución de las puntuaciones de calidad por secuencia (Figura 4.5.2B) reveló que la mayoría de las secuencias exhibían un *Phred score* de 18 (que equivale a una probabilidad del 1,5 % de obtener una base incorrecta), lo que demuestra una mejora con respecto a estudios similares anteriores, donde la mayoría de secuencias presentaban un *Phred score* entre 10 y 15 (Lowden & Henry, 2018). La distribución promedio del contenido de nucleótidos mostró un enriquecimiento de timina y adenina en el rango de 4-5 kb. Es importante destacar que este sesgo en la composición de nucleótidos no afectó a los segmentos codificantes de los anticuerpos (Figura 4.5.2C). Además, la distribución de la longitud de las secuencias reveló el predominio de dos poblaciones. La población mayor, que abarcaba desde 4,5 hasta 5,5 kb, correspondía al plásmido linealizado. Por otro lado, la población más pequeña, centrada en 4 kb, podría estar formada por plásmidos truncados, como resultado de puntos de restricción no previstos. Es importante destacar que, a pesar de la presencia de cortes no deseados, la identificación de las secuencias codificantes de anticuerpos no se vio afectada (se identificaban como cadenas de anticuerpos al alinearlas contra la base de datos de referencia) y, por lo tanto, se incluyeron, mientras que otras secuencias de diferentes longitudes se consideraron ruido o valores atípicos (Figura 4.5.2D). Además, se realizó un filtrado de secuencias en función de la calidad de los datos previamente representados, obteniendo una colección de archivos FASTQ depurados que contenían secuencias de 4-6 kb de longitud y *Phred score* mayor de 15, cerca de medio millón de secuencias. Estas secuencias se consideraron como el material de partida del análisis secundario.

### 5.6.2 Análisis secundario e interpretación biológica de los resultados

Después de completar el análisis primario y obtener un conjunto de secuencias de alta calidad y longitud adecuada, el análisis posterior se centró en la composición de anticuerpos de los repertorios generados.

Como primera aproximación, todas las secuencias resultantes del análisis primario se alinearon, sin un filtro de calidad, con las líneas germinales V(D)J utilizando IMGT/HighV-Quest. Los datos del repertorio inmune se analizaron para ambas cadenas y los del repertorio semiinmune para la cadena ligera (debido a que todas las cadenas pesadas derivaron del mismo dAb de la subfamilia VH3-23, por lo que el análisis del dominio V de la cadena pesada no fue significativo).

Los resultados obtenidos permitieron concluir que se capturaron 31 subfamilias VH de las 50 documentadas, y de las 40 subfamilias Vk encontradas en el ser humano se capturaron 16 en el repertorio inmune y 22 en el semiinmune (Barbas et al., 2001). No fue posible incluir todas las cadenas porque las que son muy minoritarias no se amplificaron con las condiciones de PCR empleadas, y el objetivo principal era incluir cadenas de anticuerpos implicados en la respuesta inmune que ocurre en la patología celíaca. Estudios sobre la construcción de repertorios para *phage display* han demostrado que muchas subfamilias minoritarias que se incluían en los repertorios no aportaban utilidad real, por lo que es mejor centrarse en la amplificación de subfamilias que se sabe que son reactivas (Teixeira et al., 2022).

A continuación, se discuten los resultados obtenidos mediante el análisis secundario. En primer lugar, se analizan las regiones variables de las cadenas pesadas y ligeras de los anticuerpos de los repertorios secuenciados. En segundo lugar, se realiza el análisis de las regiones constantes.

Con respecto al análisis de la evolución de las regiones variables de las cadenas pesadas del repertorio inmune, la presencia de las 31 subfamilias originales se mantuvo estática en las rondas 0 y 1, luego se observó un ligero aumento de la subfamilia VH3-15 en la segunda ronda, seguido de una amplificación vertiginosa que condujo al dominio de esta subfamilia en las rondas tres y cuatro. Para probar esta hipótesis, se realizaron comparaciones por pares de las composiciones de cadenas pesadas entre diferentes rondas utilizando IMGT/StatClonotype. Como resumen del proceso de *panning*, se comparó la ronda 0 con la ronda 4 para observar cualquier cambio en la distribución de genes V. Se encontró un aumento significativo en la subfamilia VH3-15, que pasó de ser un componente minoritario a uno dominante. Este aumento fue estadísticamente significativo para todos los p-valores. Además, hubo una disminución estadísticamente significativa en la presencia de 15 subfamilias. Sin embargo, no hubo una diferencia notable en las proporciones de las subfamilias restantes (Figura 4.5.3). Cabe destacar que este evento coincidió con un aumento en la exigencia de la selección, mediante la incorporación de pasos de lavado adicionales. Después de la expansión observada en la tercera ronda, las variantes VH3-15 superaron en abundancia a otras variantes, hasta tal punto que la probabilidad de otras selecciones se volvió muy baja.

También, se investigó si el aumento inicial de la subfamilia VH3-15 era significativo. Esta subfamilia emergió como dominante en las rondas finales y mostró la mayor afinidad al antígeno. Esta comprobación se llevó a cabo mediante un análisis del cambio en la composición de la población entre la ronda 1 y la ronda 2, donde se observó un ligero aumento en la mencionada subfamilia (Figura 4.5.4). La prueba confirmó que VH3-15 fue la subfamilia que experimentó un mayor aumento, que además fue estadísticamente significativo para todos p- valores analizados. Otras dos subfamilias VH4 presentaron un aumento significativo (sólo para un p-valor) entre las rondas estudiadas, pero no estuvieron muy representadas en la última ronda (Figura 4.5.4). También se produjeron aumentos significativos en la presencia de Fab VH3-15 en la tercera y cuarta rondas, pero no entre las rondas 0 y 1.

Teniendo en cuenta que la reorganización de las cadenas no cambia a lo largo de las rondas de selección, el análisis de asociaciones entre genes se realizó utilizando los datos de la ronda 0 (sin selección, ya que en esta ronda se encuentran directamente las cadenas de anticuerpos amplificadas). Se encontraron (Figura 4.5.5) varias asociaciones significativas de genes, como VH3-33 con D5-18 y JH4, VH4-31 con D5-18 o D3-22 y JH4, VH3-7 con D6-9 y JH4, y VH3-15, D3-10 y JH4. El uso del gen de unión JH4 estaba muy extendido, lo que es común en la especie humana, en la que los genes JH presentan un uso preferencial (de más del 40 %) de JH4, frente al resto de genes JH. Esta situación se deriva de la reorganización génica, en la que están implicados no números conservados, heptámeros y espaciadores de 23 pares de bases con algunas composiciones determinadas que favorecen la implicación de JH4 (Shi et al., 2020).

Se realizó un análisis adicional centrado en las variantes de Fab seleccionadas del repertorio inmune mediante su alineamiento (VH3-15) con la secuencia germinal (región VH; número de acceso IMGT M99406), encontrándose varias discrepancias que pueden constituir mutaciones biológicamente significativas en el Fab de mayor afinidad (Tabla 4.5.1).

Otro aspecto investigado es la identificación de características específicas asociadas a la respuesta celíaca dentro del repertorio inmune. Es importante destacar que se observó que las variantes de cadenas seleccionadas (VH3-15) constituían aproximadamente el 3,5 % de las lecturas totales durante la ronda inicial. Esta prevalencia concuerda otros trabajos de secuenciación de muestras biológicas obtenidas directamente de controles sanos (Bowers et al., 2014). Sin embargo, el nivel de expresión observado fue muy superior al descrito en repertorios *naïve* humanos (Kim et al., 2017). Estas variantes se seleccionaron durante el proceso hasta dominar totalmente de la población en la última ronda (86.2 %).

La selección preferencial de variantes VH3-15-IgA como anticuerpos altamente reactivos frente a la gliadina también se ha descrito en otros estudios. En cambio, el emparejamiento VH/VL observado en los pacientes no se mantuvo en el repertorio analizado, ya que este fenómeno se produce de manera estocástica durante su construcción (Steinsbø et al., 2014).

Al analizar las variantes seleccionadas, un examen exhaustivo reveló la presencia de mutaciones somáticas dentro de las regiones VH. Específicamente, se identificaron en promedio  $9 \pm 2$  bases no coincidentes en las variantes VH3- 15 de la última ronda de selección. Sorprendentemente, estas discrepancias estaban distribuidas de manera extensa a lo largo de las lecturas, lo que indica una fuerte probabilidad de que las mutaciones somáticas sean la causa subyacente. Las tasas de mutación observadas fueron comparables a las descritas en otros estudios, dónde se observaron también 9 mutaciones de media en anticuerpos séricos IgA reactivos frente a gliadina, en cambio otros anticuerpos presentaron 15 mutaciones de media, con respecto a la línea germinal (Steinsbø et al., 2014). Este hallazgo sugiere que durante la construcción del repertorio inmune, los genes que codifican las cadenas pesadas capturados y posteriormente seleccionados, pudieron originarse en células plasmáticas de vida corta que fueron activadas a través de una respuesta extrafolicular (Ho et al., 1986). Además, es importante señalar que el hecho de seleccionar pacientes celíacos que aún no habían iniciado una dieta libre de gluten favorecía la inclusión en el repertorio de genes derivados de células plasmáticas de vida corta. En ausencia de péptidos desencadenantes de una respuesta relacionada con el gluten, es altamente probable que estas células plasmáticas de vida corta hubieran sido inhibidas y que los pacientes no hubieran sido elegibles (Sharma et al., 2020).

Además, el desarrollo de las células B en el contexto extrafolicular está influenciado por las células T. Algunos investigadores han indicado que las células T presentes en el límite T-B dentro de los ganglios linfáticos, que presentan un fenotipo distinto de las células T del centro germinal, desempeñan un papel crucial en la iniciación de la activación de las células B para la generación de respuestas de anticuerpos extrafoliculares (Lee et al., 2011). La menor presencia de mutaciones en los anticuerpos reactivos frente a la gliadina también se encontró en los autoanticuerpos frente a la transglutaminasa 2 (tTG2), lo que sugiere que la presencia de antígenos del gluten, o células T específicamente activadas por el gluten, pueden contribuir potencialmente a las bajas tasas de mutación observadas en los anticuerpos IgA específicos (Steinsbø et al., 2014).

En el análisis de la evolución de las cadenas ligeras del repertorio inmune (Figura 4.5.6), se encontraron cuatro cambios estadísticamente significativos para todos los p-valores: un aumento destacado de Vk1-39, un ligero aumento de Vk1-27 y de Vk1-12, y se observó un descenso de Vk3-20. Vk1-39 se ha identificado como el gen de cadena ligera más expresado tanto en este repertorio como en otros de origen *naïve* (Kim et al., 2017), y estas cadenas también fueron las más seleccionadas a través de las diferentes rondas de panning (el aumento más significativo dentro de las cadenas ligeras).

Respecto al proceso de reorganización de las cadenas ligeras, se encontraron varias asociaciones VJ (Figura 4.5.7). El gen Vk más expresado presentó varias asociaciones con genes Jk 1 a 4, siendo más habituales las asociaciones con Jk 1 y 2. Aunque hubo varios cambios notables en la composición de las familias de cadenas ligeras a lo largo del proceso de selección del repertorio

inmune, parece que el papel de la cadena ligera como agente de selección no ha resultado determinante.

El análisis realizado de los datos obtenidos de la secuenciación masiva del repertorio semiinmune, se centró en las cadenas ligeras. En este caso no se observaron cambios significativos a lo largo del proceso de selección, lo que podría significar que las cadenas ligeras no fueron un factor determinante en la selección (Figura 4.5.8), mientras que las cadenas pesadas sí podían haberlo sido, por lo que se comprobó este hecho.

Un análisis adicional, que no es frecuente encontrar en estudios similares, implica identificar cualquier cambio de las clases y subclases de anticuerpos que se haya producido durante el proceso de selección. Este análisis tiene una especial importancia en la caracterización del repertorio inmune, ya que ayuda a demostrar si se trasladaron características de la respuesta humoral de los pacientes seleccionados al repertorio. Para el repertorio inmune (Figura 4.5.9), se observó un cambio claro en la subclase de anticuerpos al comparar las primeras y últimas rondas de selección. Por un lado, en el repertorio original (R0), la gran mayoría de las cadenas presentaron la subclase IgG1 (56 %), mientras que un tercio del total de cadenas identificadas eran IgG2, y las cadenas IgA representaban solo el 10 %. Por otro lado, la composición cambió por completo durante la selección, de forma que en la última ronda la cadena más representada fue IgA2 (36 %, un aumento de 14 veces). Además, este cambio se correlaciona con que los Fab más reactivos presentaron esta subclase. La composición del repertorio inmune antes de la selección es muy similar, con respecto a las subclases de IgG, a la que se encuentra en el suero humano, donde las IgGs representan más del 80 % de las inmunoglobulinas. Además, se mantiene la proporción habitual 2:1 en la cantidad de IgG1 e IgG2 (Vidarsson et al., 2014). La proporción de anticuerpos IgA en el repertorio también es la habitual en el suero humano fisiológico, de un 10-15 % (Loh et al., 2013), pero suele estar aumentada en la patología celíaca (Caio et al., 2019), y este hecho no se capturó en el repertorio. Durante el *panning* se produjo una clara selección de anticuerpos de la subclase IgA2. Esta subclase apenas está representada en el plasma sanguíneo, salvo en respuestas de tipo celíaco intensas, en las que los anticuerpos de esta subclase (de habitual expresión intestinal) aparecen en la circulación (Volta et al., 2008). Por este motivo, no es sorprendente que los anticuerpos seleccionados como más reactivos frente a la gliadina presentasen la subclase IgA2.

No se observaron cambios significativos al comparar la primera y la última ronda de selección en el repertorio semiinmune con respecto a los subisotipos de cadenas pesadas (Figura 4.5.10). La intención era crear un repertorio con una representación equivalente de las subclases de cadenas pesadas (25 % cada una). Para lograr este objetivo, se agregaron productos de PCR equimolares durante la ligación. Sin embargo, los datos indican que, durante el proceso, ciertas cadenas se amplificaron de manera más efectiva que otras, lo que dio lugar a un ligero desequilibrio (por ejemplo, un 30 % de IgG3 frente a un 20 % de IgG2).

En resumen, el repertorio inmune capturó con éxito y preservó las características asociadas a la enfermedad celíaca que han identificado múltiples investigadores. Estas características se utilizaron para generar nuevos Fab de alta afinidad dirigidos frente al gluten. Los Fab resultantes poseen atributos únicos, incorporando elementos clave de la respuesta humoral observada en pacientes celíacos. En consecuencia, sirven como sondas para la detección inmunológica del gluten en productos alimenticios. Este enfoque innovador ofrece ventajas sobre los métodos tradicionales basados en anticuerpos monoclonales que dependen de la inmunización animal, y que son los que actualmente están disponibles en el mercado.





## 6. CONCLUSIONES



## 6. CONCLUSIONES

**Primera.** La evolución dirigida de proteínas ha permitido la mejora del anticuerpo de dominio único dAb8E para la detección de gluten en alimentos. Para ello, se generó un repertorio semiinmune de fago-Fab mediante la combinación de cadenas pesadas semisintéticas construidas a partir del dAb8E, con cadenas ligeras obtenidas de linfocitos de sangre periférica de pacientes celíacos. Mediante la tecnología de *phage display*, se aisló de este repertorio el fago-anticuerpo Fab8E-4, con capacidad de detectar hasta 10 mg/kg de gluten en la técnica de ELISA indirecto desarrollada.

**Segunda.** En este trabajo se describe, por primera vez, el cribado de un repertorio inmune de fago-Fab generado íntegramente a partir de linfocitos de sangre periférica de pacientes celíacos con una elevada respuesta humoral, para el aislamiento de sondas de detección de gluten en alimentos. Empleando el fago-anticuerpo con mayor afinidad por la gliadina (fago-Fab-C), se desarrolló un ELISA indirecto con capacidad de detectar hasta 15 mg/kg de gluten procedente de trigo, cebada y centeno. El fragmento Fab-C pertenece a una subclase de anticuerpos estrechamente relacionada con la patología de la enfermedad celiaca (IgA2), lo que indica que la respuesta humoral presente en la sangre de los donantes se ha trasladado con éxito al repertorio inmune de Fab.

**Tercera.** La producción a escala semiindustrial basada en el vector de expresión pKktac y la biofactoría *Escherichia coli* RV308, ha permitido la producción soluble de los mejores candidatos de fago-Fab para la detección de gluten en alimentos.

**Cuarta.** El ELISA indirecto basado en el anticuerpo recombinante Fab-C presentó una alta especificidad y sensibilidad, pudiendo detectar hasta 11 mg/kg de gluten en alimentos. Asimismo, se ha desarrollado un ELISA sándwich indirecto basado en el Fab8E-4 como anticuerpo de captura y el Fab-C como anticuerpo de detección, que presentó mejores niveles de sensibilidad y exactitud que el ensayo indirecto.

**Quinta.** El análisis de una amplia variedad de productos comerciales con los dos inmunoensayos desarrollados, y la confirmación de los resultados con la técnica de ELISA sándwich directo basado en el anticuerpo monoclonal R5, puso de manifiesto la capacidad de los ensayos implementados para discriminar aquellos productos que superaban el límite de 20 mg/kg de gluten establecido por la legislación europea. Los ensayos basados en los anticuerpos recombinantes Fab-C y Fab8E-4 no presentaron reactividad cruzada con las prolaminas de la avena.

**Sexta.** Los recientes avances de los sistemas de predicción del plegamiento de proteínas basados en inteligencia artificial han permitido la modelización *in silico* de la estructura tridimensional de los anticuerpos recombinantes Fab8E-4 y Fab-C, así como el estudio de sus posibles regiones de interacción con la gliadina. La interacción principal de ambos Fab ocurre entre los CDRs de su cadena pesada y la región C-terminal de la gliadina. Esta es una característica distintiva de las sondas de afinidad generadas en este trabajo con respecto a otros anticuerpos disponibles frente al gluten que interactúan con epítopos de la región N-terminal.

**Séptima.** Se ha desarrollado una metodología rápida para caracterizar los repertorios de anticuerpos recombinantes obtenidos y estudiar el proceso de selección por afinidad mediante la secuenciación por nanoporos. Esta técnica ha permitido comprobar que se habían transferido al repertorio inmune las características típicas de la respuesta humoral de la enfermedad celíaca.

## 6. CONCLUSIONS

**First.** The directed evolution of proteins has resulted in enhancing the ability of the single domain antibody dAb8E to detect gluten in food. To achieve this goal, a semi-immune phage-Fab library was created by combining semi-synthetic heavy chains engineered from dAb8E with light chains derived from the peripheral blood lymphocytes of celiac patients. Using phage display technology, the Fab8E-4 phage-antibody was isolated from this library and was able to detect up to 10 mg/kg of gluten in the indirect ELISA technique developed.

**Second.** This work describes, for the first time, the screening of an immune phage-Fab library generated entirely from peripheral blood lymphocytes of celiac patients with a high humoral response, for the isolation of probes to detect gluten in food. Using the phage-antibody with the highest affinity for gliadin (phage-Fab-C), an indirect ELISA was developed with the ability to detect up to 15 mg/kg of gluten from wheat, barley and rye. The Fab-C fragment belongs to an antibody subclass closely related to the pathology of celiac disease (IgA2), indicating that the humoral response present in the blood of the donors has been successfully transferred to the Fab immune library.

**Third.** Semi-industrial scale production based on the pKKtac expression vector and the *Escherichia coli* RV308 biofactory has enabled the soluble production of the best phage-Fab candidates for gluten detection in food.

**Fourth.** The indirect ELISA based on the recombinant antibody Fab-C showed high specificity and sensitivity, being able to detect up to 11 mg/kg of gluten in food. Moreover, a sandwich ELISA was developed with Fab8E-4 as the capture antibody and Fab-C as the detection antibody, demonstrating improved levels of sensitivity and accuracy compared to the indirect assay due to the interaction between two paratope-epitope pairs.

- Fifth.** The analysis of a wide variety of commercial products with the two immunoassays developed, and the confirmation of the results with the direct sandwich ELISA technique based on the monoclonal antibody R5, demonstrated the ability of the assays implemented to discriminate those products that exceeded the 20 mg/kg gluten limit established by European legislation. The assays based on the recombinant antibodies Fab-C and Fab8E-4 showed no cross-reactivity with oat prolamins.
- Sixth.** Recent developments in three-dimensional protein structure prediction systems have enabled *in silico* modelling of the recombinant antibodies Fab8E-4 and Fab-C, along with a study of their potential gliadin-interacting regions. The primary interaction of both Fabs is between the CDRs of their heavy chains and the C-terminal region of gliadin. This is a feature that distinguishes the affinity probes generated in this work from other available gluten antibodies, which interact with epitopes in the N-terminal region of gliadin.
- Seventh.** A rapid methodology has been developed to characterise the recombinant antibody libraries obtained and to study the affinity selection process using nanopore sequencing. This technique has shown that the characteristic features of the humoral response in celiac disease have been incorporated into the immune library.

A 3D molecular structure of a protein, rendered in a dark blue color, with a white text overlay. The protein structure is complex, showing various loops and strands. The text "7. TRABAJO FUTURO" is centered over the structure.

## 7. TRABAJO FUTURO



## 7. TRABAJO FUTURO

Teniendo en cuenta los resultados obtenidos en la presente tesis doctoral, se proponen en este capítulo diferentes líneas de trabajo que podrán abordarse para continuar con el desarrollo de los sistemas de detección de gluten descritos, así como la apertura de nuevas líneas de investigación basadas en los repertorios y anticuerpos recombinantes generados.

Una primera aproximación será utilizar los anticuerpos recombinantes obtenidos frente al gluten en inmunoensayos con una base analítica diferente a los ELISA propuestos. En este sentido, se está trabajando en el desarrollo de un biosensor enzimático amperométrico basado en un electrodo serigrafiado de carbono desechable (SPCE, *screen-printed carbon electrode*) que emplea hidroquinona como mediador de la transferencia de electrones y peróxido de hidrógeno como sustrato (Benedé et al., 2018). Alternativamente, los Fab obtenidos podrían conjugarse a partículas de oro coloidal para implementar dispositivos inmunocromatográficos de fácil uso e interpretación por parte del consumidor final (Huang et al., 2016).

El sistema de conjugación química empleado en este trabajo para biotinar las moléculas de Fab consistió en la unión de biotina a los distintos grupos amino libres presentes en los anticuerpos (residuos de lisina y extremos N-terminales). Este proceso puede provocar cambios relevantes en el plegamiento del Fab que hagan que se reduzca su afinidad por el antígeno. Por este motivo, con el fin de mejorar la sensibilidad de los ELISA indirectos y sándwich desarrollados, se propone realizar una estrategia para la biotinilación selectiva de los Fab. Por este motivo, se propone realizar una reacción de conjugación mediada por la enzima BirA (biotin-ligasa) que dirige el proceso de biotinilación a la región específica denominada dominio aceptor de biotina (DAB) (Fairhead & Howarth, 2015). Una de las ventajas más destacables de la tecnología de *phage display* empleada, radica en su gran reproducibilidad debido a que las secuencias codificantes de los anticuerpos recombinantes están bien definidas. Al disponer de dichas secuencias, el DAB se fusionará genéticamente al extremo C-terminal de los Fab (preferiblemente en la cadena ligera) con el objetivo de obtener anticuerpos biotinilados en una única posición. De esta forma, se podrán inmovilizar los Fab de forma orientada en un soporte tratado con estreptavidina o conjugarse a diferentes sistemas de detección sin alterar los dominios variables del anticuerpo.

Al disponer de las sondas de afinidad en su formato fago-Fab, otra alternativa para mejorar los Fab desarrollados consistirá en recurrir a técnicas de maduración *in vitro* de la afinidad. Para ello, se introducen pequeñas modificaciones en las secuencias de los CDRs (regiones determinantes de la complementariedad), como por ejemplo a través de la generación de mutaciones aleatorias mediante *error-prone PCR*. Esta técnica generará un repertorio de mutantes derivado del Fab original que habrá que seleccionar por afinidad frente al antígeno (Deng et al., 1993). En este sentido, se ha demostrado que un mayor número de mutaciones no tiene por qué resultar en mejores anticuerpos, si no que un menor número, pero en posiciones concretas puede aumentar la probabilidad de encontrar variantes mejoradas (Root et al., 2021). En este sentido, gracias a que disponemos de las predicciones de la interacción Fab-gliadina, se hace posible la realización de mutaciones dirigidas a posiciones relevantes para la unión. De esta forma, incluyendo modificaciones que se prevea que pueden mejorar la afinidad según los modelos predictivos y aplicando sistemas de inteligencia artificial, se ha demostrado que es posible obtener anticuerpos de mayor afinidad a partir del análisis de unas pocas variantes (Hie et al., 2023).

Otra propuesta para mejorar el rendimiento de los inmunoensayos desarrollados consistirá en la transformación de los fragmentos recombinantes obtenidos en una molécula de anticuerpo completo. De esta forma, al presentar la molécula dos zonas de interacción con el antígeno se espera una mejora en los límites de detección del ensayo. En esta aproximación, se pueden obtener anticuerpos cuyos extremos Fab sean iguales (dando lugar a una molécula capaz de interactuar con dos epítomos iguales) o combinar dos Fab diferentes (como el Fab-C y el Fab8E-4) en una molécula de anticuerpo bifuncional con capacidad de reconocer dos epítomos diferentes (Wang et al., 2019). Para llevar a cabo esta estrategia, se clonarán los dominios variables de los Fab en vectores de expresión para su producción soluble en células de mamífero.

Por otra parte, tras la realización de esta tesis doctoral, el grupo de investigación dispone de dos repertorios de fago-Fab (semiinmune e inmune) que podrían emplearse para otros fines. Dichos repertorios han sido cribados en este trabajo para la búsqueda de variantes con capacidad de unión a gliadina. Sin embargo, es posible que en la respuesta humoral de los donantes celíacos se generen anticuerpos frente a otros antígenos implicados en la enfermedad celíaca. Por ejemplo, se ha descubierto que la transglutaminasa (TG) microbiana, una enzima ampliamente utilizada como aditivo alimentario para mejorar la textura de los alimentos proteicos, provoca reacciones adversas en pacientes celíacos. Esto es debido a que los autoanticuerpos generados típicamente en la enfermedad celíaca contra la TG tisular humana, presenta reacción cruzada con la TG microbiana (Aaron & Torsten, 2019). Por tanto, con la finalidad de desarrollar técnicas de detección de este aditivo en alimentos, se procederá a la búsqueda de anticuerpos Fab dirigidos frente a la TG microbiana mediante una nueva selección por afinidad de los repertorios construidos.

Cabe mencionar que el desarrollo de esta Tesis Doctoral ha permitido establecer nuevos protocolos de trabajo, ampliando el *know-how* del grupo de investigación en la construcción de repertorios de anticuerpos. Estos protocolos, han permitido continuar trabajando en la línea de investigación basada en el uso de anticuerpos recombinantes para la detección de alérgenos alimentarios y se están empleando actualmente para la construcción de otros repertorios de scFv.





## 8. BIBLIOGRAFÍA



## 8. BIBLIOGRAFÍA

### A

- Aaron, L., & Torsten, M. (2019). Microbial transglutaminase: A new potential player in celiac disease. *Clinical Immunology*, *199*, 37-43. <https://doi.org/https://doi.org/10.1016/j.clim.2018.12.008>
- Abanades, B., Georges, G., Bujotzek, A., & Deane, C. M. (2022). ABlooper: fast accurate antibody CDR loop structure prediction with accuracy estimation. *Bioinformatics*, *38*(7), 1877- 1880.
- Abanades, B., Wong, W. K., Boyles, F., Georges, G., Bujotzek, A., & Deane, C. M. (2023). ImmuneBuilder: Deep-Learning models for predicting the structures of immune proteins. *Communications Biology*, *6*(1), 575. <https://doi.org/10.1038/s42003-023-04927-7>
- Adolf-Bryfogle, J., Kalyuzhniy, O., Kubitz, M., Weitzner, B. D., Hu, X., Adachi, Y., Schief, W. R., & Dunbrack Jr, R. L. (2018). RosettaAntibodyDesign (RAbD): A general framework for computational antibody design. *PLoS computational biology*, *14*(4), e1006112.
- Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., Costello, C. E., Cravatt, B. F., Fenselau, C., Garcia, B. A., Ge, Y., Gunawardena, J., Hendrickson, R. C., Hergenrother, P. J., Huber, C. G., Ivanov, A. R., Jensen, O. N., Jewett, M. C., Kelleher, N. L., ... Zhang, B. (2018). How many human proteoforms are there? *Nature Chemical Biology*, *14*(3), 206-214. <https://doi.org/10.1038/nchembio.2576>
- Agnihotry, S., Pathak, R. K., Singh, D. B., Tiwari, A., & Hussain, I. (2022). Protein structure prediction. In D. B. Singh & R. K. Pathak (Eds.), *Bioinformatics* (pp. 177-188). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-323-89775-4.00023-7>
- Akdis, M., Palomares, O., van de Veen, W., van Splunter, M., & Akdis, C. A. (2012). TH17 and TH22 cells: a confusion of antimicrobial response with tissue inflammation versus protection. *Journal of Allergy and Clinical Immunology*, *129*(6), 1438-1449. <https://doi.org/10.1016/j.jaci.2012.05.003>
- Akiyama, H., Imai, T., & Ebisawa, M. (2011). Japan food allergen labeling regulation- history and evaluation. In (pp. 139-171). Elsevier. <https://doi.org/10.1016/b978-0-12-385989-1.00004-1>
- Al Badri, F. M., Baatjies, R., & Jeebhay, M. F. (2020). Assessing the health impact of interventions for baker's allergy and asthma in supermarket bakeries: a group randomised trial. *International Archives of Occupational and Environmental Health*, *93*(5), 589-599. <https://doi.org/10.1007/s00420-019-01511-7>

- Alamyar, E., Duroux, P., Lefranc, M.-P., & Giudicelli, V. (2012). IMGT® Tools for the nucleotide analysis of immunoglobulin (IG) and T Cell Receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. In F. T. Christiansen & B. D. Tait (Eds.), *Immunogenetics: Methods and Applications in Clinical Practice* (pp. 569-604). Humana Press. [https://doi.org/10.1007/978-1-61779-842-9\\_32](https://doi.org/10.1007/978-1-61779-842-9_32)
- Alford, R. F., Leaver-Fay, A., Jeliaskov, J. R., O'Meara, M. J., DiMaio, F. P., Park, H., Shapovalov, M. V., Renfrew, P. D., Mulligan, V. K., & Kappel, K. (2017). The Rosetta all-atom energy function for macromolecular modeling and design. *Journal of chemical theory and computation*, *13*(6), 3031-3048.
- Aljada, B., Zohni, A., & El-Matary, W. (2021). The gluten-free diet for celiac disease and beyond. *Nutrients*, *13*(11), 3993. <https://doi.org/10.3390/nu13113993>
- Almagro, J. C., Teplyakov, A., Luo, J., Sweet, R. W., Kodangattil, S., Hernandez-Guzman, F., & Gilliland, G. L. (2014). Second antibody modeling assessment (AMA-II). *Proteins*, *82*(8), 1553-1562. <https://doi.org/10.1002/prot.24567>
- Alt, F. W., & Baltimore, D. (1982). Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions. *Proceedings of the National Academy of Sciences*, *79*(13), 4118-4122. <https://doi.org/10.1073/pnas.79.13.4118>
- Ambrosetti, F., Jiménez-García, B., Roel-Touris, J., & Bonvin, A. (2020). Modeling antibody-antigen complexes by information-driven docking. *Structure*, *28*(1), 119-129.e112. <https://doi.org/10.1016/j.str.2019.10.011>
- Amnuaycheewa, P., Niemann, L., Goodman, R. E., Baumert, J. L., & Taylor, S. L. (2022). Challenges in gluten analysis: a comparison of four commercial sandwich ELISA kits. *Foods*, *11*(5), 706.
- Anderson, R. P., Degano, P., Godkin, A. J., Jewell, D. P., & Hill, A. V. S. (2000). In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nature Medicine*, *6*(3), 337-342. <https://doi.org/10.1038/73200>
- Annunziato, F., Romagnani, C., & Romagnani, S. (2015). The 3 major types of innate and adaptive cell-mediated effector immunity. *Journal of Allergy and Clinical Immunology*, *135*(3), 626-635. <https://doi.org/10.1016/j.jaci.2014.11.001>
- Anzai, H., Terai, T., Jayathilake, C., Suzuki, T., & Nemoto, N. (2019). A novel immuno-PCR method using cDNA display. *Analytical Biochemistry*, *578*, 1-6. <https://doi.org/10.1016/j.ab.2019.04.017>
- Aouinti, S., Giudicelli, V., Duroux, P., Malouche, D., Kossida, S., & Lefranc, M. P. (2016). IMGT/StatClonotype for pairwise evaluation and visualization of NGS IG and TR IMGT clonotype (AA) diversity or expression from IMGT/HighV-QUEST. *Frontiers in Immunology*, *7*, 339. <https://doi.org/10.3389/fimmu.2016.00339>
- Arentz-Hansen, H., Körner, R., Molberg, Ø., Quarsten, H., Vader, W., Kooy, Y. M. C., Lundin, K. E. A., Koning, F., Roepstorff, P., Sollid, L. M., & McAdam, S. N. (2000). The intestinal T Cell response to  $\alpha$ -Gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *The Journal of Experimental Medicine*, *191*(4), 603-612. <https://doi.org/10.1084/jem.191.4.603>
- Armstrong, N., Adey, N. B., McConnell, S. J., & Kay, B. K. (1996). Vectors for phage display. In B. K. Kay, J. Winter, & J. McCafferty (Eds.), *Phage Display of Peptides and Proteins* (pp. 35-53). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-012402380-2/50005-8>

Arnold, F. H. (2019). Innovation by evolution: bringing new chemistry to life (Nobel Lecture). *Angewandte Chemie International Edition*, 58(41), 14420-14426. <https://doi.org/10.1002/anie.201907729>

Australia and New Zealand Food Standards Code. Standard 1.2.7. Nutrition, health and related claims F2018C00942. <https://www.mpi.govt.nz/food-business/labelling-composition-food-drinks/health-and-nutrition-content-claims-for-food-and-drink/health-and-nutrition-claims-regulations-standard-1-2-7/>

## B

Balakireva, A., & Zamyatnin, A. (2016). Properties of gluten intolerance: gluten structure, evolution, pathogenicity and detoxification capabilities. *Nutrients*, 8(10), 644. <https://doi.org/10.3390/nu8100644>

Barbas, C. F. (2001). *Phage display: a laboratory manual*. Cold Spring Harbor Laboratory Press.

Barbié, V., & Lefranc, M.-P. (1998). The human immunoglobulin Kappa variable (IGKV) genes and joining (IGKJ) segments. *Experimental and Clinical Immunogenetics*, 15(3), 171-183. <https://doi.org/10.1159/000019068>

Bashir, S., & Paeshuyse, J. (2020). Construction of antibody phage libraries and their application in veterinary immunovirology. *Antibodies* 9(2). <https://doi.org/10.3390/antib9020021>

Beckwith, J. (2013). The sec-dependent pathway. *Research in Microbiology*, 164(6), 497-504. <https://doi.org/10.1016/j.resmic.2013.03.007>

Beltran, L., Koenig, M., Egner, W., Howard, M., Butt, A., Austin, M., Patel, D., Sanderson, R., Goubet, S., & Saleh, F. (2014). High-titre circulating tissue transglutaminase-2 antibodies predict small bowel villous atrophy, but decision cut-off limits must be locally validated. *Clinical & Experimental Immunology*, 176(2), 190-198.

Benedé, S., Ruiz-Valdepeñas Montiel, V., Povedano, E., Villalba, M., Mata, L., Galán-Malo, P., Torrente-Rodríguez, R. M., Vargas, E., Reviejo, A. J., Campuzano, S., & Pingarrón, J. M. (2018). Fast amperometric immunoplatfrom for ovomucoid traces determination in fresh and baked foods. *Sensors and Actuators B: Chemical*, 265, 421-428. <https://doi.org/https://doi.org/10.1016/j.snb.2018.03.075>

Benoit, L., Masiri, J., Del Blanco, I. A., Meshgi, M., Gendel, S. M., & Samadpour, M. (2017). Assessment of avenins from different oat varieties using R5-based sandwich ELISA. *Journal of Agricultural and Food Chemistry*, 65(8), 1467-1472. <https://doi.org/10.1021/acs.jafc.6b05105>

Bertoline, L. M. F., Lima, A. N., Krieger, J. E., & Teixeira, S. K. (2023). Before and after AlphaFold2: An overview of protein structure prediction. *Frontiers in Bioinformatics*, 3, 1120370. <https://doi.org/10.3389/fbinf.2023.1120370>

Better, M., Chang, C. P., Robinson, R. R., & Horwitz, A. H. (1988). Escherichia coli secretion of an active chimeric antibody fragment. *Science*, 240(4855), 1041-1043. <https://doi.org/doi:10.1126/science.3285471>

Biesiekierski, J. R. (2017). What is gluten? *Journal of Gastroenterology and Hepatology*, 32, 78-81. <https://doi.org/10.1111/jgh.13703>

- Blech, M., Peter, D., Fischer, P., Bauer, M. M., Hafner, M., Zeeb, M., & Nar, H. (2013). One target—two different binding modes: structural insights into gevokizumab and canakinumab interactions to interleukin-1 $\beta$ . *Journal of molecular biology*, *425*(1), 94-111.
- Böldicke, T. (2017). Single domain antibodies for the knockdown of cytosolic and nuclear proteins. *Protein Science*, *26*(5), 925-945. <https://doi.org/10.1002/pro.3154>
- Borbet, T. C., Hines, M. J., & Koralov, S. B. (2021). MicroRNA regulation of B cell receptor signaling. *Immunology Reviews*, *304*(1), 111-125. <https://doi.org/10.1111/imr.13024>
- Boukid, F., Prandi, B., Faccini, A., & Sforza, S. (2019). A complete mass spectrometry (MS)- based peptidomic description of gluten peptides generated during in vitro gastrointestinal digestion of durum wheat: implication for celiac disease. *Journal of The American Society for Mass Spectrometry*, *30*(8), 1481-1490. <https://doi.org/10.1007/s13361-019-02212-8>
- Bowers, E., Scamurra, R. W., Asrani, A., Beniguel, L., Mawhinney, S., Keays, K. M., Thurn, J. R., & Janoff, E. N. (2014). Decreased mutation frequencies among immunoglobulin G variable region genes during viremic HIV-1 infection. *PLoS ONE*, *9*(1), e81913. <https://doi.org/10.1371/journal.pone.0081913>
- Bowers, P. M., Horlick, R. A., Kehry, M. R., Neben, T. Y., Tomlinson, G. L., Altobelli, L., Zhang, X., Macomber, J. L., Krapf, I. P., Wu, B. F., McConnell, A. D., Chau, B., Berkebile, A. D., Hare, E., Verdino, P., & King, D. J. (2014). Mammalian cells display for the discovery and optimization of antibody therapeutics. *Methods*, *65*(1), 44-56. <https://doi.org/10.1016/j.ymeth.2013.06.010>
- Bradbury, A. R. M., Trinklein, N. D., Thie, H., Wilkinson, I. C., Tandon, A. K., Anderson, S., Bladen, C. L., Jones, B., Aldred, S. F., Bestagno, M., Burrone, O., Maynard, J., Ferrara, F., Trimmer, J. S., Görnemann, J., Glanville, J., Wolf, P., Frenzel, A., Wong, J., . . . Dübel, S. (2018). When monoclonal antibodies are not monospecific: Hybridomas frequently express additional functional variable regions. *mAbs*, *10*(4), 539-546. <https://doi.org/10.1080/19420862.2018.1445456>
- Brenke, R., Hall, D. R., Chuang, G.-Y., Comeau, S. R., Bohnuud, T., Beglov, D., Schueler-Furman, O., Vajda, S., & Kozakov, D. (2012). Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics*, *28*(20), 2608-2614.
- Burnet, F. M. (1976). A modification of Jerne's theory of antibody production using the concept of clonal selection. *CA: A Cancer Journal for Clinicians*, *26*(2), 119-121. <https://doi.org/10.3322/canjclin.26.2.119>
- Butler, M. (2005). Animal cell cultures: recent achievements and perspectives in the production of biopharmaceuticals. *Applied Microbiology and Biotechnology*, *68*(3), 283-291. <https://doi.org/10.1007/s00253-005-1980-8>

## C

- Caio, G., Volta, U., Sapone, A., Leffler, D. A., De Giorgio, R., Catassi, C., & Fasano, A. (2019). Celiac disease: a comprehensive current review. *BMC Medicine*, *17*(1), 142. <https://doi.org/10.1186/s12916-019-1380-z>
- Canduri, F., & de Azevedo, W. F. (2008). Protein crystallography in drug discovery. *Current Drug Targets*, *9*(12), 1048-1053. <https://doi.org/10.2174/138945008786949423>

- Carroll, W. L., Mendel, E., & Levy, S. (1988). Hybridoma fusion cell lines contain an aberrant kappa transcript. *Molecular Immunology*, 25(10), 991-995. [https://doi.org/10.1016/0161-5890\(88\)90005-3](https://doi.org/10.1016/0161-5890(88)90005-3)
- Catassi, C., Elli, L., Bonaz, B., Bouma, G., Carroccio, A., Castillejo, G., Cellier, C., Cristofori, F., De Magistris, L., Dolinsek, J., Dieterich, W., Francavilla, R., Hadjivassiliou, M., Holtmeier, W., Körner, U., Leffler, D., Lundin, K., Mazzarella, G., Mulder, C., . . . Fasano, A. (2015). Diagnosis of non-celiac gluten sensitivity (NCGS): The salerno experts' criteria. *Nutrients*, 7(6), 4966-4977. <https://doi.org/10.3390/nu7064966>
- Catassi, C., Rossini, M., Rättsch, I. M., Bearzi, I., Santinelli, A., Castagnani, R., Pisani, E., Coppa, G. V., & Giorgi, P. L. (1993). Dose dependent effects of protracted ingestion of small amounts of gliadin in coeliac disease children: a clinical and jejunal morphometric study. *Gut*, 34(11), 1515-1519. <https://doi.org/10.1136/gut.34.11.1515>
- Cebolla, Á., Moreno, M., Coto, L., & Sousa, C. (2018). Gluten Immunogenic peptides as standard for the evaluation of potential harmful prolamins content in food and human specimen. *Nutrients*, 10(12), 1927. <https://doi.org/10.3390/nu10121927>
- Chang, M. S., Minaya, M. T., Cheng, J., Connor, B. A., Lewis, S. K., & Green, P. H. R. (2011). Double-blind randomized controlled trial of rifaximin for persistent symptoms in patients with celiac disease. *Digestive Diseases and Sciences*, 56(10), 2939-2946. <https://doi.org/10.1007/s10620-011-1719-6>
- Chao, G., Lau, W. L., Hackel, B. J., Sazinsky, S. L., Lippow, S. M., & Wittrup, K. D. (2006). Isolating and engineering human antibodies using yeast surface display. *Nature Protocols*, 1(2), 755-768. <https://doi.org/10.1038/nprot.2006.94>
- Chartrand, L. J., Russo, P. A., Duhaime, A. G., & Seidman, E. G. (1997). Wheat starch intolerance in patients with celiac disease. *Journal of the American Diet Association*, 97(6), 612-618. [https://doi.org/10.1016/s0002-8223\(97\)00156-9](https://doi.org/10.1016/s0002-8223(97)00156-9)
- Chen, P., Sun, Z., Wang, J., Liu, X., Bai, Y., Chen, J., Liu, A., Qiao, F., Chen, Y., Yuan, C., Sha, J., Zhang, J., Xu, L. Q., & Li, J. (2023). Portable nanopore-sequencing technology: Trends in development and applications. *Frontiers in Microbiology*, 14, 1043967. <https://doi.org/10.3389/fmicb.2023.1043967>
- Chen, T., Hoffmann, K., Östman, S., Sandberg, A.-S., & Olsson, O. (2011). Identification of gliadin-binding peptides by phage display. *BMC Biotechnology*, 11(1), 16. <https://doi.org/10.1186/1472-6750-11-16>
- Chen, Y., & Roux, B. (2015). Generalized Metropolis acceptance criterion for hybrid non-equilibrium molecular dynamics-Monte Carlo simulations. *Journal of Chemical Physics*, 142(2), 024101. <https://doi.org/10.1063/1.4904889>
- Chi, X., Li, Y., & Qiu, X. (2020). V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. *Immunology*, 160(3), 233-247. <https://doi.org/10.1111/imm.13176>
- Choi, H. L., Yang, H. R., Shin, H. G., Hwang, K., Kim, J. W., Lee, J. H., Ryu, T., Jung, Y., & Lee, S. (2023). Generation and Next-Generation Sequencing-based characterization of a large human combinatorial antibody library. *International Journal of Molecular Sciences*, 24(6), 6011. <https://doi.org/10.3390/ijms24066011>

- Cianferoni, A. (2016). Wheat allergy: diagnosis and management. *Journal of Asthma and Allergy*, 13. <https://doi.org/10.2147/jaa.s81550>
- Clackson, T., Hoogenboom, H. R., Griffiths, A. D., & Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature*, 352(6336), 624-628. <https://doi.org/10.1038/352624a0>
- Codex Standard 118-1979 for foods for special dietary use for persons intolerant to gluten. *Codex Alimentarius*. <https://www.fao.org/fao-who-codexalimentarius/1979FCXS.pdf>
- Cohen, I. S., Day, A. S., & Shaoul, R. (2019). Gluten in celiac disease more or less? *Rambam Maimonides Medical Journal*, 10(1). <https://doi.org/10.5041/rmmj.10360>
- Cohen, T., Halfon, M., & Schneidman-Duhovny, D. (2022). NanoNet: Rapid and accurate end-to-end nanobody modeling by deep learning. *Frontiers in Immunology*, 13, 958584.

## D

- D'Angelo, S., Kumar, S., Naranjo, L., Ferrara, F., Kiss, C., & Bradbury, A. R. (2014). From deep sequencing to actual clones. *Protein Engineering, Design and Selection*, 27(10), 301-307. <https://doi.org/10.1093/protein/gzu032>
- Damelang, T., Rogerson, S. J., Kent, S. J., & Chung, A. W. (2019). Role of IgG3 in infectious diseases. *Trends in immunology*, 40(3), 197-211.
- Dammeyer, T., Steinwand, M., Krüger, S.-C., Dübel, S., Hust, M., & Timmis, K. N. (2011). Efficient production of soluble recombinant single chain Fv fragments by a *Pseudomonas putida* strain KT2440 cell factory. *Microbial Cell Factories*, 10(1), 11. <https://doi.org/10.1186/1475-2859-10-11>
- Daniell, H., Streatfield, S. J., & Wycoff, K. (2001). Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends in Plant Science*, 6(5), 219-226. [https://doi.org/10.1016/s1360-1385\(01\)01922-7](https://doi.org/10.1016/s1360-1385(01)01922-7)
- Dar, A. (2017). Molecular docking: approaches, types, applications and basic challenges. *Journal of Analytical & Bioanalytical Techniques*, 08. <https://doi.org/10.4172/2155-9872.1000356>
- David, A., Islam, S., Tankhilevich, E., & Sternberg, M. J. E. (2022). The AlphaFold database of protein structures: a biologist's guide. *Journal of molecular biology*, 434(2), 167336. <https://doi.org/https://doi.org/10.1016/j.jmb.2021.167336>
- David, F., Westphal, R., Bunk, B., Jahn, D., & Franco-Lara, E. (2010). Optimization of antibody fragment production in *Bacillus megaterium*: The role of metal ions on protein secretion. *Journal of Biotechnology*, 150(1), 115-124. <https://doi.org/https://doi.org/10.1016/j.jbiotec.2010.07.023>
- Day, L., Augustin, M. A., Batey, I. L., & Wrigley, C. W. (2006). Wheat-gluten uses and industry needs. *Trends in Food Science & Technology*, 17(2), 82-90. <https://doi.org/https://doi.org/10.1016/j.tifs.2005.10.003>
- de Aguiar, R. B., da Silva, T. d. A., Costa, B. A., Machado, M. F. M., Yamada, R. Y., Braggion, C., Perez, K. R., Mori, M. A. S., Oliveira, V., & de Moraes, J. Z. (2021). Generation and functional characterization of a single-chain variable fragment (scFv) of the anti-FGF2 3F12E7 monoclonal antibody. *Scientific Reports*, 11(1), 1432. <https://doi.org/10.1038/s41598-020-80746-8>

- de Almeida Paiva, V., de Souza Gomes, I., Monteiro, C. R., Mendonça, M. V., Martins, P. M., Santana, C. A., Gonçalves-Almeida, V., Izidoro, S. C., de Melo-Minardi, R. C., & de Azevedo Silveira, S. (2022). Protein structural bioinformatics: An overview. *Computers in Biology and Medicine*, *147*, 105695.
- De Haard, H. J., Van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderikx, P., De Bruijn, A. P., Arends, J.-W., & Hoogenboom, H. R. (1999). A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *Journal of Biological Chemistry*, *274*(26), 18218-18230. <https://doi.org/10.1074/jbc.274.26.18218>
- de la Cruz, V. F., Lal, A. A., & McCutchan, T. F. (1988). Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. *Journal of Biological Chemistry*, *263*(9), 4318-4322.
- De Marco, A. (2015). Recombinant antibody production evolves into multiple options aimed at yielding reagents suitable for application-specific needs. *Microbial Cell Factories*, *14*(1). <https://doi.org/10.1186/s12934-015-0320-7>
- Dekosky, B. J., Lungu, O. I., Park, D., Johnson, E. L., Charab, W., Chrysostomou, C., Kuroda, D., Ellington, A. D., Ippolito, G. C., Gray, J. J., & Georgiou, G. (2016). Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. *Proceedings of the National Academy of Sciences*, *113*(19), E2636-E2645. <https://doi.org/10.1073/pnas.1525510113>
- Demirkesen, I., & Ozkaya, B. (2022). Recent strategies for tackling the problems in gluten-free diet and products. *Critical Reviews in Food Science and Nutrition*, *62*(3), 571-597. <https://doi.org/10.1080/10408398.2020.1823814>
- Deng, S. J., MacKenzie, C. R., & Narang, S. A. (1993). Simultaneous randomization of antibody CDRs by a synthetic ligase chain reaction strategy. *Nucleic Acids Research*, *21*(18), 4418-4419. <https://doi.org/10.1093/nar/21.18.4418>
- DesJarlais, R. L., Sheridan, R. P., Seibel, G. L., Dixon, J. S., Kuntz, I. D., & Venkataraghavan, R. (1988). Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *Journal of Medical Chemistry*, *31*(4), 722-729. <https://doi.org/10.1021/jm00399a006>
- Dickerson, R. E. (1992). A little ancient history. *Protein Science*, *1*(1), 182-186. <https://doi.org/10.1002/pro.5560010118>
- Dieterich, W., Ehnis, T., Bauer, M., Donner, P., Volta, U., Riecken, E. O., & Schuppan, D. (1997). Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nature Medicine*, *3*(7), 797-801. <https://doi.org/10.1038/nm0797-797>
- Ding, G., Chen, X., Zhu, J., & Cao, B. (2010). Identification of two aberrant transcripts derived from a hybridoma with amplification of functional immunoglobulin variable genes. *Molecular Immunology*, *7*(5), 349-354. <https://doi.org/10.1038/cmi.2010.33>
- Directrices para el desarrollo de un sistema APPCC en el sector de la pastelería (2013). *Comunidad de Madrid*. <https://www.comunidad.madrid/servicios/salud/appcc-empresa-alimentaria>
- Dnyandev, K., Galave, V., Kulkarni, V., Chandrakant, M., & Otari, K. (2021). A review on molecular docking. *International Research Journal of Pure and Applied Chemistry*, 60-68. <https://doi.org/10.9734/irjpac/2021/v22i330396>

- Doña, V., Urrutia, M., Bayardo, M., Alzogaray, V., Goldbaum, F. A., & Chirido, F. G. (2010). Single domain antibodies are especially suited for quantitative determination of gliadins under denaturing conditions. *Journal of Agricultural and Food Chemistry*, *58*(2), 918-926. <https://doi.org/10.1021/jf902973c>
- Dorn, M., e Silva, M. B., Buriol, L. S., & Lamb, L. C. (2014). Three-dimensional protein structure prediction: Methods and computational strategies. *Computational biology and chemistry*, *53*, 251-276.
- Dunbar, J., Krawczyk, K., Leem, J., Baker, T., Fuchs, A., Georges, G., Shi, J., & Deane, C. M. (2014). SAbDab: the structural antibody database. *Nucleic Acids Research*, *42*(D1), D1140-D1146.

## E

- Elofsson, A. (2023). Progress at protein structure prediction, as seen in CASP15. *Current opinion in structural biology*, *80*, 102594. <https://doi.org/https://doi.org/10.1016/j.sbi.2023.102594>
- Erasmus, M. F., D'Angelo, S., Ferrara, F., Naranjo, L., Teixeira, A. A., Buonpane, R., Stewart, S. M., Nastri, H. G., & Bradbury, A. R. M. (2021). A single donor is sufficient to produce a highly functional in vitro antibody library. *Communications Biology*, *4*(1), 350. <https://doi.org/10.1038/s42003-021-01881-0>

## F

- Fairhead, M., & Howarth, M. (2015). Site-specific biotinylation of purified proteins using BirA. *Methods in Molecular Biology*, *1266*, 171-184. [https://doi.org/10.1007/978-1-4939-2272-7\\_12](https://doi.org/10.1007/978-1-4939-2272-7_12)
- Fasano, A., & Catassi, C. (2012). Clinical practice. Celiac disease. *New England Journal of Medicine*, *367*(25), 2419-2426. <https://doi.org/10.1056/NEJMcp1113994>
- Fausther-Bovendo, H., & Kobinger, G. (2021). Plant-made vaccines and therapeutics. *Science*, *373*(6556), 740-741. <https://doi.org/10.1126/science.abf5375>
- Fernández-Quintero, M. L., Kokot, J., Waibl, F., Fischer, A. M., Quoika, P. K., Deane, C. M., & Liedl, K. R. (2023). Challenges in antibody structure prediction. *mAbs*, *15*(1), 2175319. <https://doi.org/10.1080/19420862.2023.2175319>
- Ferre-López, S., Ribes-koninckx, C., Genzor, C., Gamen, S., Peña, L., Ortigosa, L., & Méndez, E. (2004). Immunochromatographic sticks for tissue transglutaminase and antigliadin antibody screening in celiac disease. *Clinical Gastroenterology and Hepatology*, *2*(6), 480-484. [https://doi.org/https://doi.org/10.1016/S1542-3565\(04\)00166-1](https://doi.org/https://doi.org/10.1016/S1542-3565(04)00166-1)
- Flisikowska, T., Thorey, I. S., Offner, S., Ros, F., Lifke, V., Zeitler, B., Rottmann, O., Vincent, A., Zhang, L., Jenkins, S., Niersbach, H., Kind, A. J., Gregory, P. D., Schnieke, A. E., & Platzer, J. (2011). Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. *PLoS ONE*, *6*(6), e21045. <https://doi.org/10.1371/journal.pone.0021045>
- Food Allergen Labeling and Consumer Protection Act. Sec. 206 Regulations 21 USC 343. Title II of Public Law 108-282, enacted on August 2, 2004. <https://www.fda.gov/food/food-allergens/gluten-free-guidance-documents/regulatory-information/food-allergen-labeling-and-consumer-protection-act-2004-falcpa>

- Forsström, B., Bisławska Axnäs, B., Rockberg, J., Danielsson, H., Bohlin, A., & Uhlen, M. (2015). Dissecting antibodies with regards to linear and conformational epitopes. *PLoS ONE*, *10*(3), e0121673.
- Frenzel, A., Hust, M., & Schirrmann, T. (2013). Expression of recombinant antibodies. *Frontiers in Immunology*, *4*. <https://doi.org/10.3389/fimmu.2013.00217>
- Frey, K. G., Herrera-Galeano, J. E., Redden, C. L., Luu, T. V., Servetas, S. L., Mateczun, A. J., Mokashi, V. P., & Bishop-Lilly, K. A. (2014). Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. *BMC Genomics*, *15*(1), 96. <https://doi.org/10.1186/1471-2164-15-96>

## G

- Gaciarz, A., Veijola, J., Uchida, Y., Saaranen, M. J., Wang, C., Hörkkö, S., & Ruddock, L. W. (2016). Systematic screening of soluble expression of antibody fragments in the cytoplasm of *E. coli*. *Microbial Cell Factories*, *15*, 1-10.
- Galfre, G., Howe, S. C., Milstein, C., Butcher, G. W., & Howard, J. C. (1977). Antibodies to major histocompatibility antigens produced by hybrid cell lines. *Nature*, *266*(5602), 550-552. <https://doi.org/10.1038/266550a0>
- Gao, X., Wen, L., Li, H., Wang, R., & Yin, J. (2020). Genetic variation at the interleukin-18 locus is associated with wheat-dependent exercise-induced anaphylaxis in the Han Chinese population. *Gene*, *737*, 144462. <https://doi.org/10.1016/j.gene.2020.144462>
- García-García, A., Madrid, R., González, I., García, T., & Martín, R. (2020). A novel approach to produce phage single domain antibody fragments for the detection of gluten in foods. *Food Chemistry*, *321*, 126685. <https://doi.org/10.1016/j.foodchem.2020.126685>
- García-García, A., Madrid, R., Sohrabi, H., de la Cruz, S., García, T., Martín, R., & González, I. (2019). A sensitive and specific real-time PCR targeting DNA from wheat, barley and rye to track gluten contamination in marketed foods. *LWT*, *114*, 108378. <https://doi.org/https://doi.org/10.1016/j.lwt.2019.108378>
- Garcia, E., Llorente, M., Hernando, A., Kieffer, R., Wieser, H., & Mendez, E. (2005). Development of a general procedure for complete extraction of gliadins for heat processed and unheated foods. *European Journal of Gastroenterology & Hepatology* *17*(5), 529-539. <https://doi.org/10.1097/00042737-200505000-00010>
- GarcíaMerino, A. (2011). Anticuerpos monoclonales. Aspectos básicos [10.1016/j.nrl.2010.10.005]. *Neurología*, *26*(5), 301-306. <https://doi.org/10.1016/j.nrl.2010.10.005>
- Gauhar, A., Privezentzev, C. V., Demydchuk, M., Gerlza, T., Rieger, J., Kungl, A. J., Walsh, F. S., Rutkowski, J. L., & Stocki, P. (2021). Single domain shark VNAR antibodies neutralize SARS-CoV-2 infection in vitro. *Faseb Journal*, *35*(11), e21970. <https://doi.org/10.1096/fj.202100986RR>
- Ge, L., Knappik, A., Pack, P., Freund, C., & Plückthun, A. (1995). Expressing antibodies in *Escherichia coli*. In (pp. 229-266).
- Glanville, J., D'Angelo, S., Khan, T. A., Reddy, S. T., Naranjo, L., Ferrara, F., & Bradbury, A. R. (2015). Deep sequencing in library selection projects: what insight does it bring? *Current opinion in structural biology*, *33*, 146-160. <https://doi.org/10.1016/j.sbi.2015.09.001>

- Goesaert, H., Brijs, K., Veraverbeke, W. S., Courtin, C. M., Gebruers, K., & Delcour, J. A. (2005). Wheat flour constituents: how they impact bread quality, and how to impact their functionality. *Trends in Food Science & Technology*, *16*(1-3), 12-30. <https://doi.org/10.1016/j.tifs.2004.02.011>
- Guigou, V., Cuisinier, A. M., Tonnelle, C., Moinier, D., Fougereau, M., & Fumoux, F. (1990). Human immunoglobulin VH and VK repertoire revealed by in situ hybridization. *Molecular Immunology*, *27*(9), 935-940. [https://doi.org/10.1016/0161-5890\(90\)90161-r](https://doi.org/10.1016/0161-5890(90)90161-r)
- Gujral, N., Suh, J. W., & Sunwoo, H. H. (2015). Effect of anti-gliadin IgY antibody on epithelial intestinal integrity and inflammatory response induced by gliadin. *BMC Immunology*, *16*(1). <https://doi.org/10.1186/s12865-015-0104-1>
- ## H
- Hadjivassiliou, M., Sanders, D. S., Woodroffe, N., Williamson, C., & Grünewald, R. A. (2008). Gluten ataxia. *The Cerebellum*, *7*(3), 494-498. <https://doi.org/10.1007/s12311-008-0052-x>
- Han, S. Y., Antoine, A., Howard, D., Chang, B., Chang, W. S., Slein, M., Deikus, G., Kossida, S., Duroux, P., Lefranc, M. P., Sebra, R. P., Smith, M. L., & Fofana, I. B. F. (2018). Coupling of single molecules, long read sequencing with IMGT/HighV-QUEST analysis expedites identification of SIV gp140-specific antibodies from scFv phage display libraries. *Frontiers in Immunology*, *9*, 329. <https://doi.org/10.3389/fimmu.2018.00329>
- Hardin, C., Pogorelov, T. V., & Luthey-Schulten, Z. (2002). Ab initio protein structure prediction. *Current opinion in structural biology*, *12*(2), 176-181.
- Havenar-Daughton, C., Sarkar, A., Kulp, D. W., Toy, L., Hu, X., Deresa, I., Kalyuzhnyi, O., Kaushik, K., Upadhyay, A. A., Menis, S., Landais, E., Cao, L., Diedrich, J. K., Kumar, S.,
- Schiffner, T., Reiss, S. M., Seumois, G., Yates, J. R., Paulson, J. C., . . . Crotty, S. (2018). The human naive B cell repertoire contains distinct subclasses for a germline-targeting HIV-1 vaccine immunogen. *Science Translational Medicine*, *10*(448), eaat0381. <https://doi.org/doi:10.1126/scitranslmed.aat0381>
- Hemadou, A., Giudicelli, V., Smith, M. L., Lefranc, M. P., Duroux, P., Kossida, S., Heiner, C., Hepler, N. L., Kuijpers, J., Groppi, A., Korch, J., Mondon, P., Ottones, F., Jacobin-Valat, M. J., Laroche-Traineau, J., & Clofent-Sanchez, G. (2017). Pacific Biosciences sequencing and IMGT/HighV-QUEST analysis of full-length single chain fragment variable from an in vivo selected phage-display combinatorial library. *Frontiers in Immunology*, *8*, 1796. <https://doi.org/10.3389/fimmu.2017.01796>
- Hertveldt, K., Beliën, T., & Volckaert, G. (2009). General M13 phage display: M13 phage display in identification and characterization of protein-protein interactions. *Methods in Molecular Biology*, *502*, 321-339. [https://doi.org/10.1007/978-1-60327-565-1\\_19](https://doi.org/10.1007/978-1-60327-565-1_19)
- Hie, B. L., Shanker, V. R., Xu, D., Bruun, T. U. J., Weidenbacher, P. A., Tang, S., Wu, W., Pak, J. E., & Kim, P. S. (2023). Efficient evolution of human antibodies from general protein language models. *Nature Biotechnology*. <https://doi.org/10.1038/s41587-023-01763-2>
- Ho, F., Lortan, J. E., MacLennan, I. C., & Khan, M. (1986). Distinct short-lived and long-lived antibody-producing cell populations. *European Journal of Immunology*, *16*(10), 1297-1301. <https://doi.org/10.1002/eji.1830161018>

- Hoischen, C., Fritsche, C., Gumpert, J., Westermann, M., Gura, K., & Fahnert, B. (2002). Novel bacterial membrane surface display system using cell wall-less L-forms of *Proteus mirabilis* and *Escherichia coli*. *Applied and Environmental Microbiology*, *68*(2), 525-531. <https://doi.org/10.1128/aem.68.2.525-531.2002>
- Hollon, J., Puppa, E., Greenwald, B., Goldberg, E., Guerrero, A., & Fasano, A. (2015). Effect of gliadin on permeability of intestinal biopsy explants from celiac disease patients and patients with non-celiac gluten sensitivity. *Nutrients*, *7*(3), 1565-1576. <https://doi.org/10.3390/nu7031565>
- Hong, S. B., Shin, Y. W., Hong, J. B., Lee, S. K., & Han, B. (2022). Exploration of shared features of B cell receptor and T cell receptor repertoires reveals distinct clonotype clusters. *Frontiers in Immunology*, *13*, 1006136. <https://doi.org/10.3389/fimmu.2022.1006136>
- Hoogenboom, H. (1997). Designing and optimizing library selection strategies for generating high-affinity antibodies. *Trends in Biotechnology*, *15*(2), 62-70. [https://doi.org/10.1016/s0167-7799\(97\)84205-9](https://doi.org/10.1016/s0167-7799(97)84205-9)
- Hoolehan, W., Harris, J. C., Byrum, J. N., Simpson, D. A., & Rodgers, K. K. (2022). An updated definition of V(D)J recombination signal sequences revealed by high-throughput recombination assays. *Nucleic Acids Research*, *50*(20), 11696-11711. <https://doi.org/10.1093/nar/gkac1038>
- Høydahl, L. S., Richter, L., Frick, R., Snir, O., Gunnarsen, K. S., Landsverk, O. J. B., Iversen, R., Jeliaskov, J. R., Gray, J. J., Bergseng, E., Foss, S., Qiao, S.-W., Lundin, K. E. A., Jahnsen, J., Jahnsen, F. L., Sandlie, I., Sollid, L. M., & Løset, G. Å. (2019). Plasma cells are the most abundant gluten peptide MHC-expressing cells in inflamed intestinal tissues from patients with celiac disease. *Gastroenterology*, *156*(5), 1428-1439.e1410. <https://doi.org/10.1053/j.gastro.2018.12.013>
- Hozumi, N., & Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proceedings of the National Academy of Sciences*, *73*(10), 3628-3632. <https://doi.org/10.1073/pnas.73.10.3628>
- Huang, X., Aguilar, Z. P., Xu, H., Lai, W., & Xiong, Y. (2016). Membrane-based lateral flow immunochromatographic strip with nanoparticles as reporters for detection: A review. *Biosensors and Bioelectronics*, *75*, 166-180. <https://doi.org/10.1016/j.bios.2015.08.032>
- Huang, Y., Zhang, H., Jiang, S., Yue, D., Lin, X., Zhang, J., & Gao, Y. Q. (2023). DSDP: a blind docking strategy accelerated by GPUs. *Journal of Chemical Information and Modeling*, *63*(14), 4355-4363. <https://doi.org/10.1021/acs.jcim.3c00519>
- Husby, S., Foged, N., Oxelius, V., & Svehag, S. (1986). Serum IgG subclass antibodies to gliadin and other dietary antigens in children with coeliac disease. *Clinical and Experimental Immunology*, *64*(3), 526.
- Hust, M. (2004). Mating antibody phage display with proteomics. *Trends in Biotechnology*, *22*(1), 8-14. <https://doi.org/10.1016/j.tibtech.2003.10.011>
- Iversen, R., Snir, O., Stensland, M., Kroll, J. E., Steinsbø, Ø., Korponay-Szabó, I. R., Lundin, K. E. A., de Souza, G. A., & Sollid, L. M. (2017). Strong clonal relatedness between serum and gut IgA despite different plasma cell origins. *Cell reports*, *20*(10), 2357-2367. <https://doi.org/10.1016/j.celrep.2017.08.036>

## J

- Jayathilake, C., Kumachi, S., Arai, H., Motohashi, M., Terai, T., Murakami, A., & Nemoto, N. (2020). In vitro selection of anti-gliadin single-domain antibodies from a naïve library for cDNA-display mediated immuno-PCR. *Analytical Biochemistry*, *589*, 113490. <https://doi.org/10.1016/j.ab.2019.113490>
- Jerne, N. K. (1955). The natural-selection theory of antibody formation. *Proceedings of the National Academy of Sciences*, *41*(11), 849-857. <https://doi.org/10.1073/pnas.41.11.849>
- Jiricny, J. (2013). Postreplicative mismatch repair. *Cold Spring Harbor Perspectives in Biology*, *5*(4), a012633-a012633. <https://doi.org/10.1101/cshperspect.a012633>
- Jones, D. T. (1999). GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. *Journal of molecular biology*, *287*(4), 797-815.
- Joosten, V., Lokman, C., Van Den Hondel, C. A., & Punt, P. J. (2003). The production of antibody fragments and antibody fusion proteins by yeasts and filamentous fungi. *Microbial Cell Factories*, *2*(1), 1. <https://doi.org/10.1186/1475-2859-2-1>
- Joyce, C., Burton, D. R., & Briney, B. (2020). Comparisons of the antibody repertoires of a humanized rodent and humans by high throughput sequencing. *Science Reports*, *10*(1), 1120. <https://doi.org/10.1038/s41598-020-57764-7>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., . . . Hassabis, D. (2021). Applying and improving AlphaFold at CASP14. *Proteins*, *89*(12), 1711-1721. <https://doi.org/10.1002/prot.26257>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583-589. <https://doi.org/10.1038/s41586-021-03819-2>

## K

- Kahlenberg, F., Sanchez, D., Lachmann, I., Tuckova, L., Tlaskalova, H., Méndez, E., & Mothes, T. (2006). Monoclonal antibody R5 for detection of putatively coeliac-toxic gliadin peptides. *European Food Research and Technology*, *222*(1), 78-82. <https://doi.org/10.1007/s00217-005-0100-4>
- Kang, T. H., & Seong, B. L. (2020). Solubility, stability, and avidity of recombinant antibody fragments expressed in microorganisms. *Frontiers in Microbiology*, *11*, 1927.
- Kapingidza, A. B., Kowal, K., & Chruszcz, M. (2020). Antigen-Antibody Complexes. *Subcellular Biochemistry*, *94*, 465-497. [https://doi.org/10.1007/978-3-030-41769-7\\_19](https://doi.org/10.1007/978-3-030-41769-7_19)
- Kiguchi, Y., Oyama, H., Morita, I., Katayama, E., Fujita, M., Narasaki, M., Yokoyama, A., & Kobayashi, N. (2018). Antibodies and engineered antibody fragments against M13 filamentous phage to facilitate phage display-based molecular breeding. *Biological and Pharmaceutical Bulletin*, *41*(7), 1062-1070. <https://doi.org/10.1248/bpb.b18-00162>

- Kim, S., Park, I., Park, S. G., Cho, S., Kim, J. H., S. Ipper, N., Choi, S. S., Lee, E. S., & Hong, H. J. (2017). Generation, diversity determination, and application to antibody selection of a human naïve Fab library. *Molecules and Cells*. <https://doi.org/10.14348/molcells.2017.0106>
- Kim, S. M., Mayassi, T., & Jabri, B. (2015). Innate immunity: Actuating the gears of celiac disease pathogenesis. *Best Practice & Research Clinical Gastroenterology*, 29(3), 425-435. <https://doi.org/10.1016/j.bpg.2015.05.001>
- Kircher, M., & Kelso, J. (2010). High-throughput DNA sequencing - concepts and limitations. *BioEssays*, 32(6), 524-536. <https://doi.org/10.1002/bies.200900181>
- Koerner, T. B., Abbott, M., Godefroy, S. B., Popping, B., Yeung, J. M., Diaz-Amigo, C., Roberts, J., Taylor, S. L., Baumert, J. L., Ulberth, F., Wehling, P., & Koehler, P. (2013). Validation procedures for quantitative gluten ELISA methods: AOAC allergen community guidance and best practices. *Journal of AOAC International*, 96(5), 1033-1040. <https://doi.org/10.5740/jaoacint.13-043>
- Köhler, G., & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256(5517), 495-497. <https://doi.org/10.1038/256495a0>
- Köhler, P., Schwalb, T., Immer, U., Lacorn, M., Wehling, P., & Don, C. (2013). AACCI approved methods technical committee report: collaborative study on the immunochemical determination of partially hydrolyzed gluten using an R5 competitive ELISA. *Cereal Foods World*, 58, 113-117. <https://doi.org/10.1094/CFW-58-3-0402>
- Kondrashova, A., Mustalahti, K., Kaukinen, K., Viskari, H., Volodicheva, V., Haapala, A. M., Ilonen, J., Knip, M., Mäki, M., & Hyöty, H. (2008). Lower economic status and inferior hygienic environment may protect against celiac disease. *Annals of Medicine*, 40(3), 223- 231. <https://doi.org/10.1080/07853890701678689>
- Koninckx, C. R., Giliams, J., Polanco, I., & Pena, A. (1984). IgA antigliadin antibodies in celiac and inflammatory bowel disease. *Journal of Pediatric Gastroenterology and Nutrition*, 3(5), 676-682.
- Kontsek, P., Novák, M., & Kontseková, E. (1988). Karyotype analysis of hybridomas producing monoclonal antibodies against different antigens. *Folia Biology*, 34(2), 99-104.
- Kovaltsuk, A., Leem, J., Kelm, S., Snowden, J., Deane, C. M., & Krawczyk, K. (2018). Observed antibody space: a resource for data mining next-generation sequencing of antibody repertoires. *The Journal of Immunology*, 201(8), 2502-2509.
- Krawczyk, K., Baker, T., Shi, J., & Deane, C. M. (2013). Antibody i-Patch prediction of the antibody binding site improves rigid local antibody-antigen docking. *Protein Engineering, Design and Selection*, 26(10), 621-629. <https://doi.org/10.1093/protein/gzt043>
- Krawczyk, K., Liu, X., Baker, T., Shi, J., & Deane, C. M. (2014). Improving B-cell epitope prediction and its application to global antibody-antigen docking. *Bioinformatics*, 30(16), 2288-2294. <https://doi.org/10.1093/bioinformatics/btu190>
- Kringelum, J. V., Lundegaard, C., Lund, O., & Nielsen, M. (2012). Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS computational biology*, 8(12), e1002829. <https://doi.org/10.1371/journal.pcbi.1002829>

- Krokan, H. E., & Bjoras, M. (2013). Base excision repair. *Cold Spring Harbor Perspectives in Biology*, 5(4), a012583-a012583. <https://doi.org/10.1101/cshperspect.a012583>
- Krüger, C., Hu, Y., Pan, Q., Marcotte, H., Hultberg, A., Delwar, D., Van Dalen, P. J., Pouwels, P. H., Leer, R. J., Kelly, C. G., Van Dollenweerd, C., Ma, J. K., & Hammarström, L. (2002). In situ delivery of passive immunity by lactobacilli producing single-chain antibodies. *Nature Biotechnology*, 20(7), 702-706. <https://doi.org/10.1038/nbt0702-702>
- Kufareva, I., & Abagyan, R. (2011). Methods of Protein Structure Comparison. *Methods in Molecular Biology*, (857), 231-257. [https://doi.org/10.1007/978-1-61779-588-6\\_10](https://doi.org/10.1007/978-1-61779-588-6_10)
- Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., & Ferrin, T. E. (1982). A geometric approach to macromolecule-ligand interactions. *Journal of molecular biology*, 161(2), 269-288. [https://doi.org/10.1016/0022-2836\(82\)90153-x](https://doi.org/10.1016/0022-2836(82)90153-x)
- Kurada, S., Yadav, A., & Leffler, D. A. (2016). Current and novel therapeutic strategies in celiac disease. *Expert Review in Clinical Pharmacology*, 9(9), 1211-1223. <https://doi.org/10.1080/17512433.2016.1200463>
- Kurcinski, M., Kmiecik, S., Zalewski, M., & Kolinski, A. (2021). Protein-protein docking with large-scale backbone flexibility using coarse-grained Monte-Carlo simulations. *International Journal of Molecular Sciences*, 22(14). <https://doi.org/10.3390/ijms22147341>
- L**
- Lahens, N. F., Ricciotti, E., Smirnova, O., Toorens, E., Kim, E. J., Baruzzo, G., Hayer, K. E., Ganguly, T., Schug, J., & Grant, G. R. (2017). A comparison of Illumina and Ion Torrent sequencing platforms in the context of differential gene expression. *BMC Genomics*, 18(1). <https://doi.org/10.1186/s12864-017-4011-0>
- Lau, A., Kandathil, S., & Jones, D. (2023). *Merizo: A rapid and accurate domain segmentation method using invariant point attention*. <https://doi.org/10.1101/2023.02.19.529114>
- Lee, C. M. Y., Iorno, N., Sierro, F., & Christ, D. (2007). Selection of human antibody fragments by phage display. *Nature Protocols*, 2(11), 3001-3008. <https://doi.org/10.1038/nprot.2007.448>
- Lee, S. K., Rigby, R. J., Zotos, D., Tsai, L. M., Kawamoto, S., Marshall, J. L., Ramiscal, R. R., Chan, T. D., Gatto, D., Brink, R., Yu, D., Fagarasan, S., Tarlinton, D. M., Cunningham, A. F., & Vinuesa, C. G. (2011). B cell priming for extrafollicular antibody responses requires Bcl-6 expression by T cells. *Journal of Experimental Medicine*, 208(7), 1377- 1388. <https://doi.org/10.1084/jem.20102065>
- Lee, S. Y., Choi, J. H., & Xu, Z. (2003). Microbial cell-surface display. *Trends in Biotechnology*, 21(1), 45-52. [https://doi.org/10.1016/s0167-7799\(02\)00006-9](https://doi.org/10.1016/s0167-7799(02)00006-9)
- Lee, Y., Yoo, D. K., Noh, J., Ju, S., Lee, E., Lee, H., Kwon, S., & Chung, J. (2021). Amplification of a minimally biased antibody repertoire for in vitro display using a universal primer- based amplification method. *Journal of Immunological Methods*, 496, 113089. <https://doi.org/https://doi.org/10.1016/j.jim.2021.113089>
- Leem, J., Dunbar, J., Georges, G., Shi, J., & Deane, C. M. (2016). ABodyBuilder: Automated antibody structure prediction with data-driven accuracy estimation. *mAbs*, 8(7), 1259- 1268. <https://doi.org/10.1080/19420862.2016.1205773>

- Leffler, D. A., Kelly, C. P., Green, P. H. R., Fedorak, R. N., Dimarino, A., Perrow, W., Rasmussen, H., Wang, C., Bercik, P., Bachir, N. M., & Murray, J. A. (2015). Larazotide acetate for persistent symptoms of celiac disease despite a gluten-free diet: a randomized controlled trial. *Gastroenterology*, *148*(7), 1311-1319. e1316. <https://doi.org/10.1053/j.gastro.2015.02.008>
- Leffler, D. A., & Schuppan, D. (2010). Update on serologic testing in celiac disease. *Official journal of the American College of Gastroenterology*, *105*(12), 2520-2524.
- Lefranc, G. (2019). IMGT® and 30 Years of Immunoinformatics Insight in Antibody V and C Domain Structure and Function. *Antibodies*, *8*(2), 29. <https://www.mdpi.com/2073-4468/8/2/29>
- Lefranc, M.-P. (2008). WHO-IUIS nomenclature subcommittee for immunoglobulins and T cell receptors report: August 2007, 13th International Congress of Immunology, Rio de Janeiro, Brazil. *Developmental & Comparative Immunology*, *32*(5), 461-463. <https://doi.org/https://doi.org/10.1016/j.dci.2007.09.008>
- Lefranc, M.-P. (2013). IMGLEMT-ONTOLOGY, CLASSIFICATION Axiom. In W. Dubitzky, O. Wolkenhauer, K.-H. Cho, & H. Yokota (Eds.), *Encyclopedia of Systems Biology* (pp. 973-974). Springer New York. [https://doi.org/10.1007/978-1-4419-9863-7\\_682](https://doi.org/10.1007/978-1-4419-9863-7_682)
- Lefranc, M.-P., & Lefranc, G. (2020). Immunoglobulins or antibodies: IMGT® bridging genes, structures and functions. *Biomedicines*, *8*(9), 319. <https://doi.org/10.3390/biomedicines8090319>
- Lefranc, M.-P., Pallarès, N., & Fripiat, J.-P. (1999). Allelic polymorphisms and RFLP in the human immunoglobulin lambda light chain locus. *Human Genetics*, *104*(5), 361-369. <https://doi.org/10.1007/s004390050969>
- Lefranc, M.-P., Pommié, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., Thouvenin-Contet, V., & Lefranc, G. (2003). IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Developmental & Comparative Immunology*, *27*(1), 55-77. [https://doi.org/https://doi.org/10.1016/S0145-305X\(02\)00039-3](https://doi.org/https://doi.org/10.1016/S0145-305X(02)00039-3)
- Lefranc, M. P., Giudicelli, V., Regnier, L., & Duroux, P. (2008). IMGT, a system and an ontology that bridge biological and computational spheres in bioinformatics. *Brief Bioinformatics*, *9*(4), 263-275. <https://doi.org/10.1093/bib/bbn014>
- Lefranc, M. P., & Lefranc, G. (2019). IMGT(®) and 30 years of immunoinformatics insight in antibody V and C domain structure and function. *Antibodies* *8*(2). <https://doi.org/10.3390/antib8020029>
- Leow, C. H., Fischer, K., Leow, C. Y., Cheng, Q., Chuah, C., & McCarthy, J. (2017). Single domain antibodies as new biomarker detectors. *Diagnostics*, *7*(4), 52. <https://doi.org/10.3390/diagnostics7040052>
- Lei Nº 10674 de 16 de maio de 2003 obriga a que os produtos alimentícios comercializados informen sobre a presença de gluten, como medida preventiva e de controle da doença celíaca. *Diário Oficial da União*, de 19 de maio de 2003. [https://www.planalto.gov.br/ccivil\\_03/leis/2003/l10.674.htm](https://www.planalto.gov.br/ccivil_03/leis/2003/l10.674.htm)
- Li, H., Bose, U., Stockwell, S., Howitt, C., Colgrave, M. (2019). Assessing the utility of multiplexed liquid chromatography-mass spectrometry for gluten detection in Australian breakfast food products. *Molecules*, *24*(20), 3665. <https://doi.org/10.3390/molecules24203665>

- Li, S., Lefranc, M.-P., Miles, J. J., Alamyar, E., Giudicelli, V., Duroux, P., Freeman, J. D., Corbin, V. D. A., Scheerlinck, J.-P., Frohman, M. A., Cameron, P. U., Plebanski, M., Loveland, B., Burrows, S. R., Papenfuss, A. T., & Gowans, E. J. (2013). IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. *Nature Communications*, 4(1). <https://doi.org/10.1038/ncomms3333>
- Liao, W.-W., Asri, M., Ebler, J., Doerr, D., Haukness, M., Hickey, G., Lu, S., Lucas, J. K., Monlong, J., Abel, H. J., Buonaiuto, S., Chang, X. H., Cheng, H., Chu, J., Colonna, V., Eizenga, J. M., Feng, X., Fischer, C., Fulton, R. S., . . . Paten, B. (2023). A draft human pangenome reference. *Nature*, 617(7960), 312-324. <https://doi.org/10.1038/s41586-023-05896-x>
- Lin, B., Hui, J., & Mao, H. (2021). Nanopore technology and its applications in gene sequencing. *Biosensors*, 11(7), 214. <https://doi.org/10.3390/bios11070214>
- Lipovsek, D., & Plückthun, A. (2004). In-vitro protein evolution by ribosome display and mRNA display. *Journal of Immunological Methods*, 290 (1-2), 51-67. <https://doi.org/10.1016/j.jim.2004.04.008>
- Loh, R. K., Vale, S., & McLean-Tooke, A. (2013). Quantitative serum immunoglobulin tests. *Australian Family Physician*, 42(4), 195-198.
- Losurdo, G., Principi, M., Iannone, A., Amoruso, A., Ierardi, E., Di Leo, A., & Barone, M. (2018). Extra-intestinal manifestations of non-celiac gluten sensitivity: An expanding paradigm. *World Journal of Gastroenterology*, 24(14), 1521-1530. <https://doi.org/10.3748/wjg.v24.i14.1521>
- Lowden, M. J., & Henry, K. A. (2018). Oxford nanopore sequencing enables rapid discovery of single-domain antibodies from phage display libraries. *BioTechniques*, 65(6), 351-356. <https://doi.org/10.2144/btn-2018-0123>

## M

- Madrid, R., García-García, A., González, I., Martín, R., & García, T. (2020). Phage displayed domain antibodies (dAb) for detection of allergenic pistachio proteins in foods. *Foods*, 9(9). <https://doi.org/10.3390/foods9091230>
- Magnusson, B. Ö., Ulf (2014). *Eurachem Guide: The fitness for purpose of analytical methods – A laboratory guide to method validation and related topics* (2nd ed.). [www.eurachem.org](http://www.eurachem.org)
- Makowski, L. (1993). Structural constraints on the display of foreign peptides on filamentous bacteriophages. *Gene*, 128(1), 5-11. [https://doi.org/https://doi.org/10.1016/0378-1119\(93\)90146-T](https://doi.org/https://doi.org/10.1016/0378-1119(93)90146-T)
- Manohar, R. S. (2002). Development of enzyme-linked immunosorbent assay for evaluation of chapati-making quality of wheat varieties. *Journal of Agricultural and Food Chemistry*, 50(25), 7455-7460. <https://doi.org/10.1021/jf0258151>
- Manual de Aplicación del Sistema APPCC en Industrias de Confitería-Pastelería, Bollería y Repostería de Castilla-La Mancha (2004). *Confederación Regional de Empresarios de Castilla-La Mancha y Junta de Castilla-La Mancha*. <https://www.castillalamancha.es/content/manual-de-aplicaci%C3%B3n-del-sistema-appcc-en-industrias-de-confiter%C3%ADa-pasteler%C3%ADa-boller%C3%ADa-y-reposter%C3%ADa>
- Markgren, J., Hedenqvist, M., Rasheed, F., Skepö, M., & Johansson, E. (2020). Glutenin and Gliadin, a piece in the puzzle of their structural properties in the cell described through Monte Carlo simulations. *Biomolecules*, 10(8), 1095. <https://doi.org/10.3390/biom10081095>

- Marks, D. S., Colwell, L. J., Sheridan, R., Hopf, T. A., Pagnani, A., Zecchina, R., & Sander, C. (2011). Protein 3D structure computed from evolutionary sequence variation. *PLoS ONE*, *6*(12), e28766. <https://doi.org/10.1371/journal.pone.0028766>
- Martín-Fernández, B., Costa, J., Oliveira, M. B. P. P., López-Ruiz, B., & Mafra, I. (2015). Screening new gene markers for gluten detection in foods. *Food Control*, *56*, 57-63. <https://doi.org/10.1016/j.foodcont.2015.02.047>
- Martín-Fernández, B., Costa, J., Oliveira, M. B. P. P., López-Ruiz, B., & Mafra, I. (2016). Combined effects of matrix and gene marker on the real-time PCR detection of wheat. *International Journal of Food Science & Technology*, *51*(7), 1680-1688. <https://doi.org/10.1111/ijfs.13141>
- Martín, M., Cacho, J. F., Cepeda, A., Martín, F., & Prieto, I. (2010). *Informe del Comité Científico de la Agencia Española de Seguridad Alimentaria y Nutrición (AESAN) en relación con la enfermedad celíaca y los problemas que plantean las técnicas analíticas para el control del contenido de gluten en los alimentos* retrieved from [http://www.aecosan.msssi.gob.es/AECOSAN/docs/documentos/seguridad\\_alimentaria/evaluacion\\_riesgos/informes\\_comite/TECNICAS\\_ANALITICAS\\_GLUTEN.pdf](http://www.aecosan.msssi.gob.es/AECOSAN/docs/documentos/seguridad_alimentaria/evaluacion_riesgos/informes_comite/TECNICAS_ANALITICAS_GLUTEN.pdf)
- Marzari, R., Sblattero, D., Florian, F., Tongiorgi, E., Not, T., Tommasini, A., Ventura, A., & Bradbury, A. (2001). Molecular dissection of the tissue transglutaminase autoantibody response in celiac disease. *The Journal of Immunology*, *166*(6), 4170-4176. <https://doi.org/10.4049/jimmunol.166.6.4170>
- Marze, N. A., Lyskov, S., & Gray, J. J. (2016). Improved prediction of antibody VL–VH orientation. *Protein Engineering, Design and Selection*, *29*(10), 409-418.
- Matos, C. F. R. O., Robinson, C., Alanen, H. I., Prus, P., Uchida, Y., Ruddock, L. W., Freedman, R. B., & Keshavarz-Moore, E. (2014). Efficient export of prefolded, disulfide-bonded recombinant proteins to the periplasm by the Tat pathway in *Escherichia coli* CyDisCo strains. *Biotechnology Progress*, *30*(2), 281-290. <https://doi.org/10.1002/btpr.1858>
- Matsuo, H., Yokooji, T., & Taogoshi, T. (2015). Common food allergens and their IgE-binding epitopes. *Allergy International*, *64*(4), 332-343. <https://doi.org/10.1016/j.alit.2015.06.009>
- Matthews, A. J., Zheng, S., DiMenna, L. J., & Chaudhuri, J. (2014). Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. *Advance Immunology*, *122*, 1-57. <https://doi.org/10.1016/b978-0-12-800267-4.00001-8>
- McCarville, J. L., Nisemblat, Y., Galipeau, H. J., Jury, J., Tabakman, R., Cohen, A., Naftali, E., Neiman, B., Halbfinger, E., Murray, J. A., Anbazhagan, A. N., Dudeja, P. K., Varvak, A., Leroux, J.-C., & Verdu, E. F. (2014). BL-7010 demonstrates specific binding to gliadin and reduces gluten-associated pathology in a chronic mouse model of gliadin sensitivity. *PLoS ONE*, *9*(11), e109972. <https://doi.org/10.1371/journal.pone.0109972>
- Méndez, E., Vela, C., Immer, U., & Janssen, F. W. (2005). Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *European Journal of Gastroenterology & Hepatology*, *17*(10), 1053-1063. <https://doi.org/10.1097/00042737-200510000-00008>
- Mesnager, S. P., Weber-Levy, M., Haustant, M., Mock, M. L., & Fouet, A. S. (1999). Cell surface-exposed tetanus toxin fragment C produced by recombinant bacillus anthracis protects against tetanus toxin. *Infection and Immunity*, *67*(9), 4847-4850. <https://doi.org/10.1128/iai.67.9.4847-4850.1999>

- Metzker, M. L. (2010). Sequencing technologies — the next generation. *Nature Reviews Genetics*, *11*(1), 31-46. <https://doi.org/10.1038/nrg2626>
- Miethe, S., Meyer, T., Wöhl-Bruhn, S., Frenzel, A., Schirrmann, T., Dübel, S., & Hust, M. (2013). Production of single chain fragment variable (scFv) antibodies in *Escherichia coli* using the LEX™ bioreactor. *Journal of Biotechnology*, *163*(2), 105-111. <https://doi.org/https://doi.org/10.1016/j.jbiotec.2012.07.011>
- Mitea, C., Kooy-Winkelaar, Y., Van Veelen, P., De Ru, A., Drijfhout, J. W., Koning, F., & Dekking, L. (2008). Fine specificity of monoclonal antibodies against celiac disease-inducing peptides in the gluteome. *The American Journal of Clinical Nutrition*, *88*(4), 1057-1066. <https://doi.org/10.1093/ajcn/88.4.1057>
- Monzon, V., Haft, D. H., & Bateman, A. (2022). Folding the unfoldable: using AlphaFold to explore spurious proteins. *Bioinformatics Advances*, *2*(1). <https://doi.org/10.1093/bioadv/vbab043>
- Morcos, F., Pagnani, A., Lunt, B., Bertolino, A., Marks, D. S., Sander, C., Zecchina, R., Onuchic, J. N., Hwa, T., & Weigt, M. (2011). Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proceedings of the National Academy of Sciences*, *108*(49), E1293-E1301. <https://doi.org/10.1073/pnas.1111471108>
- Morón, B., Cebolla, Á., Manyani, H., Álvarez-Maqueda, M., Megías, M., Thomas, M. D. C., López, M. C., & Sousa, C. (2008). Sensitive detection of cereal fractions that are toxic to celiac disease patients by using monoclonal antibodies to a main immunogenic wheat peptide. *The American Journal of Clinical Nutrition*, *87*(2), 405-414. <https://doi.org/10.1093/ajcn/87.2.405>
- Motea, E. A., & Berdis, A. J. (2010). Terminal deoxynucleotidyl transferase: The story of a misguided DNA polymerase. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1804*(5), 1151-1166. <https://doi.org/10.1016/j.bbapap.2009.06.030>
- Muddasani, S., Rusk, A. M., & Baquerizo Nole, K. L. (2021). Gluten and skin disease beyond dermatitis herpetiformis: a review. *International Journal of Dermatology*, *60*(3), 281-288. <https://doi.org/10.1111/ijd.15098>
- Mujico, J. R., Dekking, L., Kooy-Winkelaar, Y., Verheijen, R., Wichen, P. v., Streppel, L., Sajic, N., Drijfhout, J. W., & Koning, F. (2019). Validation of a new enzyme-linked immunosorbent assay to detect the triggering proteins and peptides for celiac disease: interlaboratory study. *Journal of AOAC*, *95*(1), 206-215. <https://doi.org/10.5740/jaoacint.11-042>
- Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., & Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, *102*(5), 553-563. [https://doi.org/10.1016/S0092-8674\(00\)00078-7](https://doi.org/10.1016/S0092-8674(00)00078-7)
- Muyldermans, S. (2013). Nanobodies: natural single-domain antibodies. *Annual Review of Biochemistry*, *82*(1), 775-797. <https://doi.org/10.1146/annurev-biochem-063011-092449>
- N**
- Nahm, M., Scott, M., & Shackelford, P. (1987). Expression of human IgG subclasses. *Annals of Clinical & Laboratory Science*, *17*(3), 183-196.
- Nakamura, K., Oshima, T., Morimoto, T., Ikeda, S., Yoshikawa, H., Shiwa, Y., Ishikawa, S., Linak, M. C., Hirai, A., Takahashi, H., Altaf-Ul-Amin, M., Ogasawara, N., & Kanaya, S. (2011). Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Research*, *39*(13), e90. <https://doi.org/10.1093/nar/gkr344>

- Napodano, C., Marino, M., Stefanile, A., Pocino, K., Scatena, R., Gulli, F., Rapaccini, G. L., Delli Noci, S., Capozio, G., & Rigante, D. (2021). Immunological role of IgG subclasses. *Immunological Investigations*, 50(4), 427-444.
- Niland, B., & Cash, B. D. (2018). Health benefits and adverse effects of a gluten-free diet in non-celiac disease patients. *Gastroenterology and Hepatology*, 14(2), 82-91.
- Norma Oficial Mexicana NOM-247-SSA1-2008 de 27 de Julio de 2009, relativo a productos y servicios. Cereales y sus productos. Cereales, harinas de cereales, sémolas o semolinas. Alimentos a base de: cereales, semillas comestibles, de harinas, sémolas o semolinas o sus mezclas. Productos de panificación. Disposiciones y especificaciones sanitarias y nutrimentales. Métodos de prueba. *Diario Oficial de la Federación*. [https://dof.gob.mx/nota\\_detalle.php?codigo=5100356&fecha=27/07/2009#gsc.tab=0](https://dof.gob.mx/nota_detalle.php?codigo=5100356&fecha=27/07/2009#gsc.tab=0)
- Norma Oficial Mexicana NOM-051-SCFI/SSA1-2010 de 27 de marzo de 2010, relativo a especificaciones generales de etiquetado para alimentos y bebidas no alcohólicas preenvasados. Información comercial y sanitaria. *Diario Oficial de la Federación*. [https://www.dof.gob.mx/normasOficiales/4010/seeco11\\_C/seeco11\\_C.htm](https://www.dof.gob.mx/normasOficiales/4010/seeco11_C/seeco11_C.htm)
- Not, T., Ziberna, F., Vatta, S., Quaglia, S., Martelossi, S., Villanacci, V., Marzari, R., Florian, F., Vecchiet, M., & Sulic, A.-M. (2011). Cryptic genetic gluten intolerance revealed by intestinal antitransglutaminase antibodies and response to gluten-free diet. *Gut*, 60(11), 1487-1493. <https://doi.org/10.1136/gut.2010.232900>
- Nyysönen, E., Penttilä, M., Harkki, A., Saloheimo, A., Knowles, J. K. C., & Keränen, S. (1993). Efficient production of antibody fragments by the filamentous fungus *Trichoderma reesei*. *Bio/Technology*, 11(5), 591-595. <https://doi.org/10.1038/nbt0593-591>

## O

- Olsen, T. H., Moal, I. H., & Deane, C. M. (2022). AbLang: an antibody language model for completing antibody sequences. *Bioinformatics Advances*, 2(1), vbac046.
- Omar, N., & Lim, T. S. (2018). Construction of naive and immune human Fab phage-display library. *Phage Display: Methods and Protocols*, 25-44.
- Ozuna, C. V., Iehisa, J. C. M., Giménez, M. J., Alvarez, J. B., Sousa, C., & Barro, F. (2015). Diversification of the celiac disease  $\alpha$ -gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *The Plant Journal*, 82(5), 794- 805. <https://doi.org/10.1111/tpj.12851>

## P

- Pagliari, D., Urgesi, R., Frosali, S., Riccioni, M. E., Newton, E. E., Landolfi, R., Pandolfi, F., & Cianci, R. (2015). The interaction among microbiota, immunity, and genetic and dietary factors is the condicio sine qua non celiac disease can develop. *Journal of Immunology Research*, 2015, 1-10. <https://doi.org/10.1155/2015/123653>
- Palomares, O., Akdis, M., Martín-Fontecha, M., & Akdis, C. A. (2017). Mechanisms of immune regulation in allergic diseases: the role of regulatory T and B cells. *Immunological Reviews*, 278(1), 219-236. <https://doi.org/10.1111/imr.12555>

- Panda, R., & Garber, E. A. E. (2019a). Detection and quantitation of gluten in fermented- hydrolyzed foods by antibody-based methods: challenges, progress, and a potential path forward. *Frontiers in Nutrition*, 6. <https://doi.org/10.3389/fnut.2019.00097>
- Panda, R., & Garber, E. A. E. (2019b). Western blot analysis of fermented-hydrolyzed foods utilizing gluten-specific antibodies employed in a novel multiplex competitive ELISA. *Analytical and Bioanalytical Chemistry*, 411(20), 5159-5174. <https://doi.org/10.1007/s00216-019-01893-0>
- Parmley, S. F., & Smith, G. P. (1988). Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene*, 73(2), 305-318. [https://doi.org/10.1016/0378-1119\(88\)90495-7](https://doi.org/10.1016/0378-1119(88)90495-7)
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, 30(1), 70-82.
- Pini, A., Viti, F., Santucci, A., Carnemolla, B., Zardi, L., Neri, P., & Neri, D. (1998). Design and use of a phage display library. *Journal of Biological Chemistry*, 273(34), 21769-21776. <https://doi.org/10.1074/jbc.273.34.21769>
- Pitman, M. R., & Menz, R. I. (2006). Methods for protein homology modelling. In D. K. Arora, R. M. Berka, & G. B. Singh (Eds.), *Applied Mycology and Biotechnology* (Vol. 6, pp. 37-59). Elsevier. [https://doi.org/https://doi.org/10.1016/S1874-5334\(06\)80005-5](https://doi.org/https://doi.org/10.1016/S1874-5334(06)80005-5)
- Plückthun, A. (2012). Ribosome display: a perspective. In (pp. 3-28). Springer New York. [https://doi.org/10.1007/978-1-61779-379-0\\_1](https://doi.org/10.1007/978-1-61779-379-0_1)
- Pollock, D. P., Kutzko, J. P., Birck-Wilson, E., Williams, J. L., Echelard, Y., & Meade, H. M. (1999). Transgenic milk as a method for the production of recombinant antibodies. *Journal of Immunological Methods*, 231(1-2), 147-157. [https://doi.org/10.1016/s0022-1759\(99\)00151-9](https://doi.org/10.1016/s0022-1759(99)00151-9)
- Popov, A. V., Zou, X., Xian, J., Nicholson, I. C., & Brüggemann, M. (1999). A human immunoglobulin lambda locus is similarly well expressed in mice and humans. *Journal of Experimental Medicine*, 189(10), 1611-1620. <https://doi.org/10.1084/jem.189.10.1611>
- Potgieter, T. I., Kersey, S. D., Mallem, M. R., Nysten, A. C., & D'Anjou, M. (2010). Antibody expression kinetics in glycoengineered *Pichia pastoris*. *Biotechnology and Bioengineering*, 106(6), 918-927. <https://doi.org/10.1002/bit.22756>
- Prihoda, D., Maamary, J., Waight, A., Juan, V., Fayadat-Dilman, L., Svozil, D., & Bitton, D. A. (2022). BioPhi: A platform for antibody design, humanization, and humanness evaluation based on natural antibody repertoires and deep learning. *mAbs*, 14(1), 2020203. <https://doi.org/10.1080/19420862.2021.2020203>
- Priyanka, P., Gayam, S., & Kupec, J. T. (2018). The role of a low fermentable oligosaccharides, disaccharides, monosaccharides, and polyol diet in nonceliac gluten sensitivity. *Gastroenterology Research and Practice*, 2018, 1-8. <https://doi.org/10.1155/2018/1561476>
- Pryor, J. M., Waters, C. A., Aza, A., Asagoshi, K., Strom, C., Mieczkowski, P. A., Blanco, L., & Ramsden, D. A. (2015). Essential role for polymerase specialization in cellular nonhomologous end joining. *Proceedings of the National Academy of Sciences*, 112(33), E4537-E4545. <https://doi.org/10.1073/pnas.1505805112>

## R

- Rauhavirta, T., Oittinen, M., Kivistö, R., Männistö, P. T., Garcia-Horsman, J. A., Wang, Z., Griffin, M., Mäki, M., Kaukinen, K., & Lindfors, K. (2013). Are Transglutaminase 2 inhibitors able to reduce gliadin-induced toxicity related to celiac disease? a proof-of-concept study. *Journal of Clinical Immunology*, 33(1), 134-142. <https://doi.org/10.1007/s10875-012-9745-5>
- Reglamento (CE) nº 852/2004 del Parlamento Europeo y del Consejo, de 29 de abril de 2004, relativo a la higiene de los productos alimenticios. *Diario Oficial de la Unión Europea*, L139, de 30 de abril de 2004. <https://www.boe.es/buscar/doc.php?id=DOUE-L-2004-81035>
- Reglamento de ejecución (UE) nº 828/2014 de la Comisión, de 30 de julio relativo a los requisitos para la transmisión de información a los consumidores sobre la ausencia o la presencia reducida de gluten en alimentos. *Diario Oficial de la Unión Europea*, L228, de 31 de julio de 2014. <https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX%3A32014R0828>
- Rhoads, A., & Au, K. F. (2015). PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics*, 13(5), 278-289. <https://doi.org/10.1016/j.gpb.2015.08.002>
- Rhyner, C., Weichel, M., Hubner, P., Achatz, G., Blaser, K., & Cramer, R. (2003). Phage display of human antibodies from a patient suffering from coeliac disease and selection of isotype-specific scFv against gliadin. *Immunology*, 110(2), 269-274. <https://doi.org/10.1046/j.1365-2567.2003.01728.x>
- Rippmann, J. R. F., Klein, M., Hoischen, C., Brocks, B., Rettig, W. J., Gumpert, J., Pfizenmaier, K., Mattes, R., & Moosmayer, D. (1998). Prokaryotic expression of single-chain variable-fragment (scFv) antibodies: secretion from cells of *Proteus mirabilis* leads to active product and overcomes the limitations of periplasmic expression. *Applied and Environmental Microbiology*, 64(12), 4862-4869. <https://doi.org/10.1128/aem.64.12.4862-4869.1998>
- Robinson, W. H. (2015). Sequencing the functional antibody repertoire—diagnostic and therapeutic discovery. *Nature Reviews Rheumatology*, 11(3), 171-182. <https://doi.org/10.1038/nrrheum.2014.220>
- Röckendorf, N., Meckelein, B., Scherf, K. A., Schalk, K., Koehler, P., & Frey, A. (2017). Identification of novel antibody-reactive detection sites for comprehensive gluten monitoring. *PLoS ONE*, 12(7), e0181566. <https://doi.org/10.1371/journal.pone.0181566>
- Roney, J. P., & Ovchinnikov, S. (2022). State-of-the-art estimation of protein model accuracy using AlphaFold. *Physical Review Letters*, 129(23), 238101. <https://doi.org/10.1103/PhysRevLett.129.238101>
- Root, A. R., Guntas, G., Katragadda, M., Apgar, J. R., Narula, J., Chang, C. S., Hanscom, S., McKenna, M., Wade, J., Meade, C., Ma, W., Guo, Y., Liu, Y., Duan, W., Hendershot, C., King, A. C., Zhang, Y., Sousa, E., Tam, A., . . . Bloom, L. (2021). Discovery and optimization of a novel anti-GUCY2c x CD3 bispecific antibody for the treatment of solid tumors. *mAbs*, 13(1), 1850395. <https://doi.org/10.1080/19420862.2020.1850395>
- Roszkowska, A., Pawlicka, M., Mroczek, A., Bałabuszek, K., & Nieradko-Iwanicka, B. (2019). Non-celiac gluten sensitivity: a review. *Medicina*, 55(6), 222. <https://doi.org/10.3390/medicina55060222>
- Rubio-Tapia, A., Rahim, M. W., See, J. A., Lahr, B. D., Wu, T.-T., & Murray, J. A. (2010). Mucosal recovery and mortality in adults with celiac disease after treatment with a gluten-free diet. *American Journal of Gastroenterology*, 105(6), 1412-1420. <https://doi.org/10.1038/ajg.2010.10>

- Ruffolo, J. A., Chu, L.-S., Mahajan, S. P., & Gray, J. J. (2023). Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. *Nature Communications*, *14*(1), 2389. <https://doi.org/10.1038/s41467-023-38063-x>
- Ruffolo, J. A., & Gray, J. J. (2022). Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. *Biophysical Journal*, *121*(3), 155a-156a.
- Ruffolo, J. A., Guerra, C., Mahajan, S. P., Sulam, J., & Gray, J. J. (2020). Geometric potentials from deep learning improve prediction of CDR H3 loop structures. *Bioinformatics*, *36*(Suppl\_1), i268-i275. <https://doi.org/10.1093/bioinformatics/btaa457>
- Ruffolo, J. A., Sulam, J., & Gray, J. J. (2021). Antibody structure prediction using interpretable deep learning. *Patterns*, *3*(2), 100406. <https://doi.org/10.1016/j.patter.2021.100406>
- S**
- Salmi, T. T. (2019). Dermatitis herpetiformis. *Clinical and Experimental Dermatology*, *44*(7), 728-731. <https://doi.org/10.1111/ced.13992>
- Sapone, A., Bai, J. C., Ciacci, C., Dolinsek, J., Green, P. H., Hadjivassiliou, M., Kaukinen, K., Rostami, K., Sanders, D. S., Schumann, M., Ullrich, R., Villalta, D., Volta, U., Catassi, C., & Fasano, A. (2012). Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Medicine*, *10*, 13. <https://doi.org/10.1186/1741-7015-10-13>
- Sblattero, D., Florian, F., Azzoni, E., Ziberna, F., Tommasini, A., Not, T., Ventura, A., Bradbury, A., & Marzari, R. (2004). One-step cloning of anti tissue transglutaminase scFv from subjects with celiac disease. *Journal of Autoimmunity*, *22*(1), 65-72. <https://doi.org/10.1016/j.jaut.2003.09.004>
- Sblattero, D., Florian, F., Not, T., Ventura, A., Bradbury, A., & Marzari, R. (2000). Analyzing the peripheral blood antibody repertoire of a celiac disease patient using phage antibody libraries. *Human Antibodies*, *9*, 199-205. <https://doi.org/10.3233/HAB-2000-9402>
- Scaviner, D., Barbié, V., Ruiz, M., & Lefranc, M.-P. (1999). Protein displays of the human immunoglobulin heavy, kappa and lambda variable and joining regions. *Experimental and Clinical Immunogenetics*, *16*(4), 234-240. <https://doi.org/10.1159/000019115>
- Schaefer, W., Völger, H. R., Lorenz, S., Imhof-Jung, S., Regula, J. T., Klein, C., & Mølhøj, M. (2016). Heavy and light chain pairing of bivalent quadroma and knobs-into-holes antibodies analyzed by UHR-ESI-QTOF mass spectrometry. *mAbs*, *8*(1), 49-55. <https://doi.org/10.1080/19420862.2015.1111498>
- Schalk, K., Lang, C., Wieser, H., Koehler, P., & Scherf, K. A. (2017). Quantitation of the immunodominant 33-mer peptide from  $\alpha$ -gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Scientific Reports*, *7*(1), 45092. <https://doi.org/10.1038/srep45092>
- Scherf, K. A., & Poms, R. E. (2016). Recent developments in analytical methods for tracing gluten. *Journal of Cereal Science*, *67*, 112-122. <https://doi.org/10.1016/j.jcs.2015.08.006>
- Schneider, C., Raybould, M. I. J., & Deane, C. M. (2021). SAbDab in the age of biotherapeutics: updates including SAbDab-nano, the nanobody structure tracker. *Nucleic Acids Research*, *50*(D1), D1368-D1372. <https://doi.org/10.1093/nar/gkab1050>

- Schoonbroodt, S., Steukers, M., Viswanathan, M., Frans, N., Timmermans, M., Wehnert, A., Nguyen, M., Ladner, R. C., & Hoet, R. M. (2008). Engineering antibody heavy chain CDR3 to create a phage display Fab library rich in antibodies that bind charged carbohydrates. *The Journal of Immunology*, *181*(9), 6213-6221. <https://doi.org/10.4049/jimmunol.181.9.6213>
- Schoonjans, R., Willems, A., Schoonooghe, S., Fiers, W., Grooten, J., & Mertens, N. (2000). Fab chains as an efficient heterodimerization scaffold for the production of recombinant bispecific and trispecific antibody derivatives. *The Journal of Immunology*, *165*(12), 7050-7057. <https://doi.org/10.4049/jimmunol.165.12.7050>
- Schumann, M., Richter, J. F., Wedell, I., Moos, V., Zimmermann-Kordmann, M., Schneider, T., Daum, S., Zeitz, M., Fromm, M., & Schulzke, J. D. (2008). Mechanisms of epithelial translocation of the alpha(2)-gliadin-33mer in coeliac sprue. *Gut*, *57*(6), 747-754. <https://doi.org/10.1136/gut.2007.136366>
- Sedlazeck, F. J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., Von Haeseler, A., & Schatz, M. C. (2018). Accurate detection of complex structural variations using single-molecule sequencing. *Nature Methods*, *15*(6), 461-468. <https://doi.org/10.1038/s41592-018-0001-7>
- Sela-Culang, I., Kunik, V., & Ofran, Y. (2013). The structural basis of antibody-antigen recognition. *Frontiers in Immunology*, *4*, 302.
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Židek, A., Nelson, A. W., & Bridgland, A. (2019). Protein structure prediction using multiple deep neural networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13). *Proteins: structure, function, and bioinformatics*, *87*(12), 1141-1148.
- Sfeir, A., & Symington, L. S. (2015). Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway? *Trends in Biochemical Sciences*, *40*(11), 701-714. <https://doi.org/10.1016/j.tibs.2015.08.006>
- Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G., Sollid, L., & Khosla, C. (2002). Structural basis for gluten intolerance in celiac sprue. *Science*, *297*, 2275-2279. <https://doi.org/10.1126/science.1074129>
- Sharma, N., Bhatia, S., Chunduri, V., Kaur, S., Sharma, S., Kapoor, P., Kumari, A., & Garg, M. (2020). Pathogenesis of celiac disease and other gluten related disorders in wheat and strategies for mitigating them. *Frontiers in Nutrition*, *7*. <https://doi.org/10.3389/fnut.2020.00006>
- Shewry, P. (2019). What is gluten -why is it special? *Frontiers in Nutrition*, *6*, 101. <https://doi.org/10.3389/fnut.2019.00101>
- Shewry, P. R., Halford, N. G., Tatham, A. S., Popineau, Y., Lafiandra, D., & Belton, P. S. (2003). The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Advances in Food and Nutrition Research*, *45*, 219-302. [https://doi.org/10.1016/s1043-4526\(03\)45006-7](https://doi.org/10.1016/s1043-4526(03)45006-7)
- Shi, B., Dong, X., Ma, Q., Sun, S., Ma, L., Yu, J., Wang, X., Pan, J., He, X., Su, D., & Yao, X. (2020). The usage of human IGHJ genes follows a particular non-random selection: The recombination signal sequence may affect the usage of human IGHJ genes. *Frontiers in Genetics*, *11*, 524413. <https://doi.org/10.3389/fgene.2020.524413>
- Shin, J.-E., Riesselman, A. J., Kollasch, A. W., McMahon, C., Simon, E., Sander, C., Manglik, A., Kruse, A. C., & Marks, D. S. (2021). Protein design and variant prediction using autoregressive generative models. *Nature Communications*, *12*(1), 2403.

- Shiroza, T., Shinozaki-Kuwahara, N., Hayakawa, M., Shibata, Y., Hashizume, T., Fukushima, K., Udaka, S., & Abiko, Y. (2003). Production of a single-chain variable fraction capable of inhibiting the *Streptococcus mutans* glucosyltransferase in *Bacillus brevis*: construction of a chimeric shuttle plasmid secreting its gene product. *Biochimica et Biophysica Acta (BBA) – Gene Structure and Expression*, *1626*(1), 57-64. [https://doi.org/https://doi.org/10.1016/S0167-4781\(03\)00038-1](https://doi.org/https://doi.org/10.1016/S0167-4781(03)00038-1)
- Singh, P., Sharma, L., Kulothungan, S. R., Adkar, B. V., Prajapati, R. S., Ali, P. S. S., Krishnan, B., & Varadarajan, R. (2013). Effect of signal peptide on stability and folding of *Escherichia coli* thioredoxin. *PLoS ONE*, *8*(5), e63442. <https://doi.org/10.1371/journal.pone.0063442>
- Singh, P., Singh, A. D., Ahuja, V., & Makharia, G. K. (2022). Who to screen and how to screen for celiac disease. *World Journal of Gastroenterology*, *28*(32), 4493.
- Skerritt, J., & Underwood, P. A. (1986). Specificity characteristics of monoclonal antibodies to wheat grain storage proteins. *Biochimica Et Biophysica Acta-protein Structure and Molecular Enzymology*, *874*, 245-254. [https://doi.org/10.1016/0167-4838\(86\)90023-3](https://doi.org/10.1016/0167-4838(86)90023-3)
- Skerritt, J. H., & Hill, A. S. (1990). Monoclonal antibody sandwich enzyme immunoassays for determination of gluten in foods. *Journal of Agricultural and Food Chemistry*, *38*(8), 1771-1778. <https://doi.org/10.1021/jf00098a029>
- Skolnick, J., Gao, M., Zhou, H., & Singh, S. (2021). AlphaFold 2: why it works and its implications for understanding the relationships of protein sequence, structure, and function. *Journal of Chemical Information and Modeling*, *61*(10), 4827-4831. <https://doi.org/10.1021/acs.jcim.1c01114>
- Smecuol, E., Hwang, H. J., Sugai, E., Corso, L., Cherňavsky, A. C., Bellavite, F. P., González, A., Vodánovich, F., Moreno, M. L., Vázquez, H., Lozano, G., Niveloni, S., Mazure, R., Meddings, J., Mauriño, E., & Bai, J. C. (2013). Exploratory, randomized, double-blind, placebo-controlled study on the effects of *Bifidobacterium infantis* natrene life start strain super strain in active celiac disease. *Journal of Clinical Gastroenterology*, *47*(2), 139-147. <https://doi.org/10.1097/MCG.0b013e31827759ac>
- Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, *228*(4705), 1315-1317. <https://doi.org/10.1126/science.4001944>
- Smith, G. P. (1993). Surface display and peptide libraries. *Gene*, *128*(1), 1-2. [https://doi.org/https://doi.org/10.1016/0378-1119\(93\)90145-S](https://doi.org/https://doi.org/10.1016/0378-1119(93)90145-S)
- Smith, G. P. (2019). Phage display: simple evolution in a Petri dish (Nobel Lecture). *Angewandte Chemie International Edition*, *58*(41), 14428-14437. <https://doi.org/10.1002/anie.201908308>
- Song, Y., & Zhang, Y. (2022). Research progress of neoantigens in gynecologic cancers. *International Immunopharmacology*, *112*, 109236. <https://doi.org/10.1016/j.intimp.2022.109236>
- Sorell, L., López, J. A., Valdés, I., Alfonso, P., Camafeita, E., Acevedo, B., Chirido, F., Gavilondo, J., & Méndez, E. (1998). An innovative sandwich ELISA system based on an antibody cocktail for gluten analysis. *Federation of European Biochemical Societies Letters*, *439*(1- 2), 46-50. [https://doi.org/10.1016/s0014-5793\(98\)01336-2](https://doi.org/10.1016/s0014-5793(98)01336-2)

- Spiteri, V. A., Goodall, M., Douch, J., Rambo, R. P., Gor, J., & Perkins, S. J. (2021). Solution structures of human myeloma IgG3 antibody reveal extended Fab and Fc regions relative to the other IgG subclasses. *Journal of Biological Chemistry*, 297(3), 100995. <https://doi.org/10.1016/j.jbc.2021.100995>
- Stamnaes, J., & Sollid, L. M. (2015). Celiac disease: autoimmunity in response to food antigen. *Seminars in Immunology*, 27(5), 343-352. <https://doi.org/10.1016/j.smim.2015.11.001>
- Steinsbø, Ø., Dunand, C. J. H., Huang, M., Mesin, L., Salgado-Ferrer, M., Lundin, K. E. A., Jahnsen, J., Wilson, P. C., & Sollid, L. M. (2014). Restricted VH/VL usage and limited mutations in gluten-specific IgA of coeliac disease lesion plasma cells. *Nature Communications*, 5(1). <https://doi.org/10.1038/ncomms5041>
- Stern, M., Ciclitira, P. J., van Eckert, R., Feighery, C., Janssen, F. W., Méndez, E., Mothes, T., Troncone, R., & Wieser, H. (2001). Analysis and clinical effects of gluten in coeliac disease. *European Journal of Gastroenterology and Hepatology*, 13(6), 741-747. <https://doi.org/10.1097/00042737-200106000-00023>
- Stretton, A. O. (2002). The first sequence. Fred Sanger and insulin. *Genetics*, 162(2), 527-532. <https://doi.org/10.1093/genetics/162.2.527>
- Sun, L., Fu, T., Zhao, D., Fan, H., & Zhong, S. (2021). Divide-and-link peptide docking: a fragment-based peptide docking protocol. *Phys Chem Chem Phys*, 23(39), 22647-22660. <https://doi.org/10.1039/d1cp02098f>

## T

- Tack, G. J., van de Water, J. M., Bruins, M. J., Kooy-Winkelaar, E. M., van Bergen, J., Bonnet, P., Vreugdenhil, A. C., Korponay-Szabo, I., Edens, L., von Blomberg, B. M., Schreurs, M. W., Mulder, C. J., & Koning, F. (2013). Consumption of gluten with gluten-degrading enzyme by celiac patients: a pilot-study. *World Journal of Gastroenterology*, 19(35), 5837-5847. <https://doi.org/10.3748/wjg.v19.i35.5837>
- Taraghikhah, N., Ashtari, S., Asri, N., Shahbazkhani, B., Al-Dulaimi, D., Rostami-Nejad, M., Rezaei-Tavirani, M., Razzaghi, M. R., & Zali, M. R. (2020). An updated overview of spectrum of gluten-related disorders: clinical and diagnostic aspects. *BMC Gastroenterology*, 20(1). <https://doi.org/10.1186/s12876-020-01390-0>
- Teixeira, A. A. R., D'Angelo, S., Erasmus, M. F., Leal-Lopes, C., Ferrara, F., Spector, L. P., Naranjo, L., Molina, E., Max, T., DeAgüero, A., Perea, K., Stewart, S., Buonpane, R. A., Nastri, H. G., & Bradbury, A. R. M. (2022). Simultaneous affinity maturation and developability enhancement using natural liability-free CDRs. *mAbs*, 14(1), 2115200. <https://doi.org/10.1080/19420862.2022.2115200>
- Therrien, A., Kelly, C. P., & Silvester, J. A. (2020). Celiac disease: extraintestinal manifestations and associated conditions. *Journal of Clinical Gastroenterology*, 54(1), 8-21. <https://doi.org/10.1097/mcg.0000000000001267>
- Thiel, M. A., Coster, D. J., Standfield, S. D., Brereton, H. M., Mavrangelos, C., Zola, H., Taylor, S., Yusim, A., & Williams, K. A. (2002). Penetration of engineered antibody fragments into the eye. *Clinical and Experimental Immunology*, 128(1), 67-74. <https://doi.org/10.1046/j.1365-2249.2002.01808.x>

- Thompson, M., Ellison, S. L. R., Fajgelj, A., Willetts, P., & Wood, R. (1999). Harmonized guidelines for the use of recovery information in analytical measurement. *Pure and Applied Chemistry*, *71*(2), 337-348. <https://doi.org/doi:10.1351/pac199971020337>
- Thompson, T., Dennis, M., Higgins, L. A., Lee, A. R., & Sharrett, M. K. (2005). Gluten-free diet survey: are Americans with coeliac disease consuming recommended amounts of fibre, iron, calcium and grain foods? *Journal of Human Nutrition and Dietetics*, *18*(3), 163-169. <https://doi.org/10.1111/j.1365-277x.2005.00607.x>
- Tiller, T., Schuster, I., Deppe, D., Siegers, K., Strohner, R., Herrmann, T., Berenguer, M., Poujol, D., Stehle, J., Stark, Y., Heßling, M., Daubert, D., Felderer, K., Kaden, S., Kölln, J., Enzelberger, M., & Urlinger, S. (2013). A fully synthetic human Fab antibody library based on fixed VH/VL framework pairings with favorable biophysical properties. *mAbs*, *5*(3), 445-470. <https://doi.org/10.4161/mabs.24218>
- Torres, M., & Casadevall, A. (2008). The immunoglobulin constant region contributes to affinity and specificities. *Trends in immunology*, *29*(2), 91-97. <https://doi.org/10.1016/j.it.2007.11.004>
- Tranquet, O., Gaudin, J.-C., Patil, S., Steinbrecher, J., Matsunaga, K., Teshima, R., Sakai, S., Larré, C., & Denery-Papini, S. (2017). A chimeric IgE that mimics IgE from patients allergic to acid-hydrolyzed wheat proteins is a novel tool for in vitro allergenicity assessment of functionalized glutens. *PLoS ONE*, *12*(11), e0187415. <https://doi.org/10.1371/journal.pone.0187415>
- Tranquet, O., Larré, C., & Denery-Papini, S. (2020). Allergic reactions to hydrolysed wheat proteins: clinical aspects and molecular structures of the allergens involved. *Critical Reviews in Food Science and Nutrition*, *60*(1), 147-156. <https://doi.org/10.1080/10408398.2018.1516622>
- Tranquet, O., Lupi, R., Echasserieau-Laporte, V., Pietri, M., Larré, C., & Denery-Papini, S. (2015). Characterization of antibodies and development of an Indirect competitive immunoassay for detection of deamidated gluten. *Journal of Agricultural and Food Chemistry*, *63*(22), 5403-5409. <https://doi.org/10.1021/acs.jafc.5b00923>
- Trimmer, J. S. (2020). Recombinant antibodies in basic neuroscience research. *Current Protocols Neuroscience*, *94*(1), e106. <https://doi.org/10.1002/cpns.106>
- Truong, L. N., Li, Y., Shi, L. Z., Hwang, P. Y.-H., He, J., Wang, H., Razavian, N., Berns, M. W., & Wu, X. (2013). Microhomology-mediated end joining and homologous recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proceedings of the National Academy of Sciences*, *110*(19), 7720-7725. <https://doi.org/10.1073/pnas.1213431110>
- Tursi, A., & Brandimarte, G. (2003). The symptomatic and histologic response to a gluten-free diet in patients with borderline enteropathy. *Journal of Clinical Gastroenterology*, *36*(1), 13-17. <https://doi.org/10.1097/00004836-200301000-00006>
- Tursi, A., Brandimarte, G., Giorgetti, G. M., Elisei, W., Inchingolo, C. D., Monardo, E., & Aiello, F. (2006). Endoscopic and histological findings in the duodenum of adults with celiac disease before and after changing to a gluten-free diet: a 2-year prospective study. *Endoscopy*, *38*(7), 702-707. <https://doi.org/10.1055/s-2006-925178>

## U

Udoye, C. C., Rau, C. N., Freye, S. M., Almeida, L. N., Vera-Cruz, S., Othmer, K., Korkmaz, R. Ü., Clauder, A.-K., Lindemann, T., Niebuhr, M., Ott, F., Kalies, K., Recke, A., Busch, H., Fähnrich, A., Finkelman, F. D., & Manz, R. A. (2022). B-cell receptor physical properties affect relative IgG1 and IgE responses in mouse egg allergy. *Mucosal Immunology*, *15*(6), 1375-1388. <https://doi.org/10.1038/s41385-022-00567-y>

## V

Vaisman-Mentesh, A., & Wine, Y. (2018). Monitoring phage biopanning by Next-Generation Sequencing. In M. Hust & T. S. Lim (Eds.), *Phage Display: Methods and Protocols* (pp. 463-473). Springer New York. [https://doi.org/10.1007/978-1-4939-7447-4\\_26](https://doi.org/10.1007/978-1-4939-7447-4_26)

Vajda, S. (2005). Classification of protein complexes based on docking difficulty. *Proteins: structure, function, and bioinformatics*, *60*(2), 176-180.

Valdés, I., García, E., Llorente, M., & Méndez, E. (2003). Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *European Journal of Gastroenterology & Hepatology*, *15*(5), 465-474. <https://doi.org/10.1097/01.meg.0000059119.41030.df>

Van Zundert, G., Rodrigues, J., Trellet, M., Schmitz, C., Kastritis, P., Karaca, E., Melquiond, A., van Dijk, M., De Vries, S., & Bonvin, A. (2016). The HADDOCK2. 2 web server: user-friendly integrative modeling of biomolecular complexes. *Journal of molecular biology*, *428*(4), 720-725.

Verbeek, W. H., Mulder, C. J., & Zweegman, S. (2006). Alemtuzumab for refractory celiac disease. *New England Journal of Medicine*, *355*(13), 1396-1397; author reply 1397. <https://doi.org/10.1056/NEJMc061784>

Victora, G. D., & Nussenzweig, M. C. (2012). Germinal centers. *Annual Review of Immunology*, *30*, 429-457.

Vidarsson, G., Dekkers, G., & Rispen, T. (2014). IgG subclasses and allotypes: from structure to effector functions. *Frontiers in Immunology*, *5*, 520. <https://doi.org/10.3389/fimmu.2014.00520>

Volta, U., De Giorgio, R., Caio, G., Uhde, M., Manfredini, R., & Alaedini, A. (2019). Nonceliac wheat sensitivity: an immune-mediated condition with systemic manifestations. *Gastroenterology Clinics of North America*, *48*(1), 165-182. <https://doi.org/10.1016/j.gtc.2018.09.012>

Volta, U., Molinaro, N., Fratangelo, D., & Bianchi, F. B. (2008). IgA subclass antibodies to gliadin in serum and intestinal juice of patients with coeliac disease. *Clinical and Experimental Immunology*, *80*(2), 192-195. <https://doi.org/10.1111/j.1365-2249.1990.tb05232.x>

## W

Wagner, H., Wehrle, S., Weiss, E., Cavallari, M., & Weber, W. (2018). A two-step approach for the design and generation of nanobodies. *International Journal of Molecular Sciences*, *19*(11), 3444. <https://doi.org/10.3390/ijms19113444>

Wahab, P. J., Meijer, J. W., & Mulder, C. J. (2002). Histologic follow-up of people with celiac disease on a gluten-free diet: slow and incomplete recovery. *American Journal of Clinical Pathology*, *118*(3), 459-463. <https://doi.org/10.1309/evxt-851x-whlc-rlx9>

- Wan, Y., Gritsch, C. S., Hawkesford, M. J., & Shewry, P. R. (2014). Effects of nitrogen nutrition on the synthesis and deposition of the  $\omega$ -gliadins of wheat. *Annals of Botany*, *113*(4), 607-615. <https://doi.org/10.1093/aob/mct291>
- Wang, Q., Chen, Y., Park, J., Liu, X., Hu, Y., Wang, T., McFarland, K., & Betenbaugh, M. J. (2019). Design and production of bispecific antibodies. *Antibodies*, *8*(3). <https://doi.org/10.3390/antib8030043>
- Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*, *39*(11), 1348-1365. <https://doi.org/10.1038/s41587-021-01108-x>
- Ward, M., Lin, C., Victoria, D. C., Fox, B. P., Fox, J. A., Wong, D. L., Meerman, H. J., Pucci, J. P., Fong, R. B., Heng, M. H., Tsurushita, N., Gieswein, C., Park, M., & Wang, H. (2004). Characterization of humanized antibodies secreted by *Aspergillus niger*. *Applied and Environmental Microbiology*, *70*(5), 2567-2576. <https://doi.org/doi:10.1128/AEM.70.5.2567-2576.2004>
- Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiology*, *24*(2), 115-119. <https://doi.org/10.1016/j.fm.2006.07.004>
- Winter, G. (2019). Harnessing evolution to make medicines (Nobel Lecture). *Angewandte Chemie International Edition*, *58*(41), 14438-14445. <https://doi.org/10.1002/anie.201909343>
- Wu, S.-C., Yeung, J. C., Duan, Y., Ye, R., Szarka, S. J., Habibi, H. R., & Wong, S.-L. (2002). Functional production and characterization of a fibrin-specific single-chain antibody fragment from *Bacillus subtilis*: effects of molecular chaperones and a wall-bound protease on antibody fragment production. *Applied and Environmental Microbiology*, *68*(7), 3261-3269. <https://doi.org/doi:10.1128/AEM.68.7.3261-3269.2002>

## X

- Xu, Z., Davila, A., Wilamowski, J., Teraguchi, S., & Standley, D. M. (2022). Improved antibody-specific epitope prediction using AlphaFold and AbAdapt. *ChemBiochem*, *23*(18), e202200303. <https://doi.org/10.1002/cbic.202200303>

## Y

- Yamaguchi, J., Naimuddin, M., Biyani, M., Sasaki, T., Machida, M., Kubo, T., Funatsu, T., Husimi, Y., & Nemoto, N. (2009). cDNA display: a novel screening method for functional disulfide-rich peptides by solid-phase synthesis and stabilization of mRNA-protein fusions. *Nucleic Acids Research*, *37*(16), e108-e108. <https://doi.org/10.1093/nar/gkp514>
- Younger, D., Berger, S., Baker, D., & Klavins, E. (2017). High-throughput characterization of protein-protein interactions by reprogramming yeast mating. *Proceedings of the National Academy of Sciences*, *114*(46), 12166-12171. <https://doi.org/10.1073/pnas.1705867114>



## 9. ANEXOS



# ANEXO I

## DE LOS SUEROS POLICLONALES A LOS ANTICUERPOS RECOMBINANTES: UNA REVISIÓN SOBRE LA DETECCIÓN INMUNOLÓGICA DE GLUTEN EN ALIMENTOS

---

### FROM POLYCLONAL SERA TO RECOMBINANT ANTIBODIES: A REVIEW OF IMMUNOLOGICAL DETECTION OF GLUTEN IN FOODSTUFF

Gluten is the ethanol-soluble protein fraction of cereal endosperms like wheat, rye, and barley. It is widely used in the food industry because of the physical- chemical properties it gives to dough. Nevertheless, there are some gluten- related diseases that are presenting increasing prevalences, e.g., celiac disease, for which a strict gluten-free diet is the best treatment. Due to this situation, gluten labeling legislation has been developed in several countries around the world. This article reviews the gluten immune detection systems that have been applied to comply with such regulations. These systems have followed the development of antibody biotechnology, which comprise three major methodologies: polyclonal antibodies, monoclonal antibodies (mAbs) derived from hybridoma cells (some examples are 401.21, R5, G12, and  $\alpha$ -20 antibodies), and the most recent methodology of recombinant antibodies. Initially, the main objective was the consecution of new high-affinity antibodies, resulting in low detection and quantification limits that are mainly achieved with the R5 mAb (the gold standard for gluten detection). Increasing knowledge about the causes of gluten-related diseases has increased the complexity of research in this field, with current efforts not only focusing on the development of more specific and sensitive systems for gluten but also the detection of protein motifs related to pathogenicity. New tools based on recombinant antibodies will provide adequate safety and traceability methodologies to meet the increasing market demand for gluten-free products.

**García-Calvo Eduardo**, García-García Aina, Madrid Raquel, Martín Rosario, García Teresa (2020). *From Polyclonal Sera to Recombinant Antibodies: A Review of Immunological Detection of Gluten in Foodstuff*.

*Foods*, 10(1), 66. <https://doi.org/10.3390/foods10010066>

Publicado en *Foods*

Índice de impacto (2020): 4,35 Posición JCR: 37/143 (Q1)



*foods*





Review

# From Polyclonal Sera to Recombinant Antibodies: A Review of Immunological Detection of Gluten in Foodstuff

Eduardo Garcia-Calvo , Aina García-García , Raquel Madrid , Rosario Martin and Teresa García \*

Departamento de Nutrición y Ciencia de los Alimentos, Facultad de Veterinaria, Universidad Complutense de Madrid, 28040 Madrid, Spain; edugar01@ucm.es (E.G.-C.); ainagarcia@ucm.es (A.G.-G.); raqmad01@ucm.es (R.M.); rmartins@ucm.es (R.M.)

\* Correspondence: tgarcia@ucm.es; Tel.: +34-91-3943747

**Abstract:** Gluten is the ethanol-soluble protein fraction of cereal endosperms like wheat, rye, and barley. It is widely used in the food industry because of the physical–chemical properties it gives to dough. Nevertheless, there are some gluten-related diseases that are presenting increasing prevalences, e.g., celiac disease, for which a strict gluten-free diet is the best treatment. Due to this situation, gluten labeling legislation has been developed in several countries around the world. This article reviews the gluten immune detection systems that have been applied to comply with such regulations. These systems have followed the development of antibody biotechnology, which comprise three major methodologies: polyclonal antibodies, monoclonal antibodies (mAbs) derived from hybridoma cells (some examples are 401.21, R5, G12, and  $\alpha$ -20 antibodies), and the most recent methodology of recombinant antibodies. Initially, the main objective was the consecution of new high-affinity antibodies, resulting in low detection and quantification limits that are mainly achieved with the R5 mAb (the gold standard for gluten detection). Increasing knowledge about the causes of gluten-related diseases has increased the complexity of research in this field, with current efforts not only focusing on the development of more specific and sensitive systems for gluten but also the detection of protein motifs related to pathogenicity. New tools based on recombinant antibodies will provide adequate safety and traceability methodologies to meet the increasing market demand for gluten-free products.

**Keywords:** antibodies; gluten; detection; immunoassays; celiac disease



**Citation:** Garcia-Calvo, E.; García-García, A.; Madrid, R.; Martin, R.; García, T. From Polyclonal Sera to Recombinant Antibodies: A Review of Immunological Detection of Gluten in Foodstuff. *Foods* **2021**, *10*, 66. <https://doi.org/10.3390/foods10010066>

Received: 26 November 2020

Accepted: 24 December 2020

Published: 30 December 2020

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Gluten is the general term for the ethanol-soluble proteins present in various cereal endosperms, including wheat, rye, barley, spelt, and kamut [1]. The definition by Codex Alimentarius also introduces some physical–chemical concepts: insoluble in water and 0.5 M sodium chloride solution [2]. Currently, this substance is slowly digested and presents a high permanence in the gut.

In 1924, Osborne introduced a classification method for plant proteins by extraction with different solvents that is still in use. After applying this classification (Table 1), wheat proteins are divided by their solubility behavior into the following fractions: globulins (soluble in a diluted salt solution), albumins (water soluble), gliadins (ethanol soluble), and glutelins (soluble in diluted acetic acid) [3].

Traditionally, gluten proteins have been separated into two fractions that are either soluble or insoluble in alcohol. This division, with some modifications, has remained in use to the present day, with the gluten proteins that are readily soluble in alcohol–water mixtures (e.g., 60–70% ethanol) being called gliadins and those that are insoluble being called glutenins. However, it is now known that the two fractions contain proteins that are structurally related, with the differences in solubility resulting from their presence as monomers that interact by non-covalent forces (gliadins) or as high molecular mass polymers stabilized by interchain disulphide bonds. When present as reduced subunits, the

glutenin proteins are also soluble in alcohol–water mixtures and can therefore be defined together with gliadins as prolamins [5]. Glutelins are heterogeneous and can be separated using electrophoresis into over a dozen fractions that are categorized into high molecular weight (HMW) and low molecular weight (LMW) groups [6]. Glutelin subunits have been found to correlate with gluten properties that are related to baking quality [5]. Gliadins are represented as single chain polypeptides, and it is accepted that gliadins are divided, according to their electrophoretic mobility in Polyacrylamide gel Electrophoresis (PAGE) at low pH (lactate-PAGE), into four major groups ( $\alpha$ -,  $\beta$ -,  $\gamma$ -, and  $\omega$ -gliadins, from fastest mobility to slowest) [7]. Gluten proteins contain large repeat domains composed of homologous and repetitive sequences of six-to-eight amino acids rich in proline and glutamine [8]. In addition, when considering the alpha-gliadin structure, their central domain contains the proline- (P) and glutamine-rich (Q) heptapeptide PQPQPPF and pentapeptide PQQPY. This domain contains the most characteristic immunogenic fragment, a 33-mer peptide comprising six overlapping epitopes significant for celiac disease pathogenesis [9], although this peptide is not present in every wheat cultivar [10].

**Table 1.** Protein fractions from cereal grains [4].

Osborne Fraction	Wheat	Rye	Oats	Barley	Corn	
Globulin	Edestin					
Albumin	Leucosin					
Gluten	Prolamin	Gliadin	Secalin	Avenin	Hordein	Zein
	Glutelin	Glutenin	Secalinin	Aveninin	Hordeinin	Zeinin

Several fragments of gliadins and glutelins are associated with different types of gluten-related diseases, e.g.,  $\alpha$  and  $\gamma$ -gliadins in celiac disease [11];  $\gamma$ -,  $\alpha/\beta$ -,  $\omega 5$ -, and  $\omega 1,2$ -gliadins, as well as HMW and LMW subunits of glutenin, are involved in wheat allergies [12].

## 2. Gluten-Related Diseases

Several diseases related to the exposure to gluten of prone persons that can be classified with etiology into the three main groups of allergy, autoimmunity, and non-celiac gluten sensitivity have been described [13].

Gluten-related allergies, also known as wheat allergies, have a prevalence of 0.1% in the general population [14], and they have developed a well-known two-step pathological mechanism: the sensitization and effector phases [15]. Within this last phase, the onset of the main reactions occurs in minutes to hours after gluten exposure driven by an IgE response. This group includes the following pathologies classified by symptomatology: (a) a respiratory allergy, also known as baker's asthma, with bronchial symptoms as severe clinical presentation [16]; (b) a food allergy with major digestive presentation [17]; (c) wheat-dependent exercise-induced anaphylaxis (WDEIA), an inflammatory situation triggered by stress [18]; and (d) contact urticaria, with major dermatologic symptoms [14].

The second group of gluten-related diseases is associated with an autoimmune etiology. Celiac disease, with a prevalence of 1% in general population, is the most important [14].

Many molecular mechanisms leading to intestinal damage in celiac disease have been described, although not all have been discovered yet. The ingestion of gluten by sensitized people results in the partial digestion of gliadin (wheat prolamin), which interacts with CXCR3 (chemokine receptor 3) and stimulates the liberation of zonulin [19]. This leads to an increased intestinal permeability, facilitating the translocation of gliadin peptides from lumen to lamina propria. Then, the secretion of innate immunity mediators (Interleukins IL15 and IL8) is triggered, with consequent neutrophil recruitment [20]. The loosened gut barrier facilitates the engagement of toll-like receptor complex 4-M2-CD14 by trypsin and alpha-amylase inhibitors, thus releasing pro-inflammatory cytokines [21]. Following the innate immune-mediated apoptosis of intestinal cells with the subsequent release

of intracellular tissue transglutaminase, gliadin peptides are partially deamidated [13]. These deaminated peptides are presented by DQ 2/8 (a class II Major Histocompatibility Complex or HLA cell surface receptor) antigen-presenting cells (APCs) to helper T cells that trigger the maturation and activation of B-cells producing IgM, IgG, and IgA against tissue transglutaminase [22] (for this reason, it is considered an autoimmune disease). Additionally, helper T cells produce pro-inflammatory cytokines like interferon-gamma and tumoral necrosis factor-alpha (TNF- $\alpha$ ) [23]. This immune response, together with the function of killer T cells, initiates the enteropathy. Damaged enterocytes express the CD71 transporter to facilitate retrotranscytosis events [24] and further increase intestinal permeability; this spurs a pro-inflammatory and pro-growth environment, resulting in the development of hyperplastic crypts and affecting the absorption of nutrients [13].

In addition to celiac disease, gluten ataxia (a neurological disease [25]) and dermatitis herpetiformis [14] are considered gluten-related autoimmunity diseases.

Non-celiac gluten sensitivity (NCGS, also denominated non-celiac wheat sensitivity and, sometimes, gluten intolerance) is the third group of gluten-related diseases by etiological classification (non-autoimmune and non-allergic), with a prevalence of up to 7% in the general population [14]. Pathogenic mechanisms are still uncertain, but it seems that innate immunity plays a major role [26]. The signs and symptoms are very similar to other gluten-related diseases, irritable bowel syndrome, and Crohn's disease. DQ2/8 haplotypes and IgG/IgA anti-gliadin antibodies are present only in 50% of cases. Intestinal damage in this disease is lower than that observed in celiac disease [27].

There has been an intense research into the pharmacological treatment of these diseases, (especially celiac disease) including gluten neutralization agents, disruptors of mucosal transportation or antigen processing enzymes, modifications of the microbiome, immunomodulators, and anti-inflammatory drugs [28]. Notwithstanding, a gluten-free diet is the most recommended and has long been considered the only effective treatment [29]. When gluten consumption is eliminated, the exacerbated immune response is inhibited, leading to the partial (if not complete) healing of the duodenal mucosa along with the resolution of symptoms and signs of malabsorption [30].

### 3. Gluten Content Labeling Legislation in Different Countries

In contrast to other allergens, and following the recommendations included in Codex Standard 118-1979 [2], there is well-developed legislation about gluten presence in food in several countries.

In Europe, the Commission Implementing Regulation (EU) No. 828/2014 of 30 July on the requirements for the provision of information to consumers on the absence or reduced presence of gluten in food [31] rules that "The statement 'gluten-free' may only be made where the food as sold to the final consumer contains no more than 20 mg/kg of gluten" and "The statement 'very low gluten' may only be made where the food, consisting of or containing one or more ingredients made from wheat, rye, barley, oats, or their crossbred varieties which have been specially processed to reduce the gluten content, contains no more than 100 mg/kg of gluten in the food as sold to the final consumer".

The U.S. Food and Drug Administration (FDA) has defined the term "gluten-free" for voluntary use in foods that are inherently gluten-free, or when they are not composed of gluten-containing grains—either raw or processed to remove gluten. Any unavoidable presence of gluten in the food must be less than 20 ppm (mg/kg), Food Allergen Labeling and Consumer Protection Act (FALCPA) [32].

Health Canada considers that gluten-free foods are those that contain levels of gluten not exceeding 20 mg/kg as a result of cross-contamination, and they must meet the health and safety intent of B.24.018 (2012). Regarding oats, on 19 May 20, Health Canada registered a marketing authorization 15 that permits the use of gluten-free claims for gluten-free oats [33].

The current legislation in Australia and New Zealand is the strictest. Australia and New Zealand Food Standard Code, standard 1.2.7, states that "for the food to be labeled as

gluten-free, the food must not contain: detectable gluten; or oats or their products; or cereals containing gluten that may have been malted, or their products." For the "not contain detectable gluten" statement, the limit was set at 3 ppm (mg/kg) [34].

In Mexico, the executive orders NOM-051-SCFI/SSA1-2010 and NOM-247-SSA1-2008 state that those foods that may produce any kind of allergy and intolerance must be labeled, and those containing grains and derivatives must be labeled with "this product contains gluten" statement [35,36]. In Argentina, there is specific legislation for celiac disease (celiac law, passed by the Congress in 2009 (26.588), modified in 2015 (27.196)), claiming this pathology as "disease of national interest," regulating not only food security issues but also social aspects. A gluten limit of 10 ppm (mg/kg) was set by this law for a product to be labeled as "gluten-free," including a specific logo. The legislation also applies to medicines [37].

In Brazil, Federal Law 8543/1992, mandates that all industrialized foods that contain gluten must carry a warning that they contain gluten. It was updated by Federal Law 10674/2003, determining that all industrialized foods must indicate on their labels and packaging the phrases "Contains Gluten" or "Does Not Contain Gluten." Brazilian legislation follows the 20 ppm (mg/kg) limit included in Codex Alimentarius [38].

In China, Food Law GB/T23779 issued in 2009 by General Administration of Quality Supervision, Inspection and Quarantine (AQSIQ) "Allergens in prepackaged foods" includes gluten-containing grains and related products amongst the substances that may induce allergic reactions. A 2015 standard specifically applicable to the inspection of gluten allergen ingredients in prepackaged food for export made a clear reference to Codex standard STAN 118-1979 in order to verify the compliance of gluten-free claims. This regulation set a maximum limit for a gluten-free claim of 20 mg/kg [39]. However, this regulation does not apply for import or domestic trade.

By Japanese law, the labeling of allergens is designated as mandatory or recommended based on the number of cases of actual illness and the degree of seriousness. To standardize official methods, the Japanese government described the validation protocol criteria in the 2006 official guidelines and stated that any food containing allergen proteins at greater than 10 ppm (mg/kg) must be labeled under the current law [40].

#### 4. Methods for Gluten Detection in Food Samples

In recent years, some issues like ongoing legislation, growing general public awareness in food security, and the commercialization of new products focused on a particular population segment (like gluten-free products for celiac patients) has encouraged food industry and research groups to develop more accurate and applicable methods for detecting the traceability of potentially harmful components like gluten.

Nowadays, the gluten detection methods that are widely used in food industry can be classified in two main groups, depending on their target biomolecules: proteins or DNA.

Methods based on protein detection can be divided into immunological and non-immunological techniques.

Immunological techniques are based on the high affinity interaction of antibodies and antigens, which has led to the development of different vastly used applications, including:

- ELISA [41]: this is a quick, economic, versatile and robust method. ELISA presents a high sensitivity (in the low ppm range) and optical detection. However, there is a possibility for some false negatives, due to protein denaturalization, and there is some risk of false positives due to cross-reactions with similar but non-target proteins.
- Immunochromatographic assays [42]: these have a visual and simple result interpretation, and they are very simple to use by final operator. Their main weakness is that they cannot quantify.
- Western Blot [43]: this is a highly specific and sensitive method (low ppm), with additional strengths, like confirmatory values (molecular weights) and highly efficient insoluble protein detection. Nevertheless, it is a time-consuming method that must be performed by qualified personnel.

Non-immunological methods are also based on the detection of proteins, but using a mechanism that is completely different to immunological detection. This group encompasses:

- Chromatography methods [44]: these are based on the separation and detection of peptides with a very high sensitivity. The main drawback is that they require complex and expensive instrumentation.
- Mass spectrometry [45]: this is a quick, reproducible, and very precise method of analysis that allows for species detection. However, it is not a quantitative method, and it requires complex and expensive instrumentation.

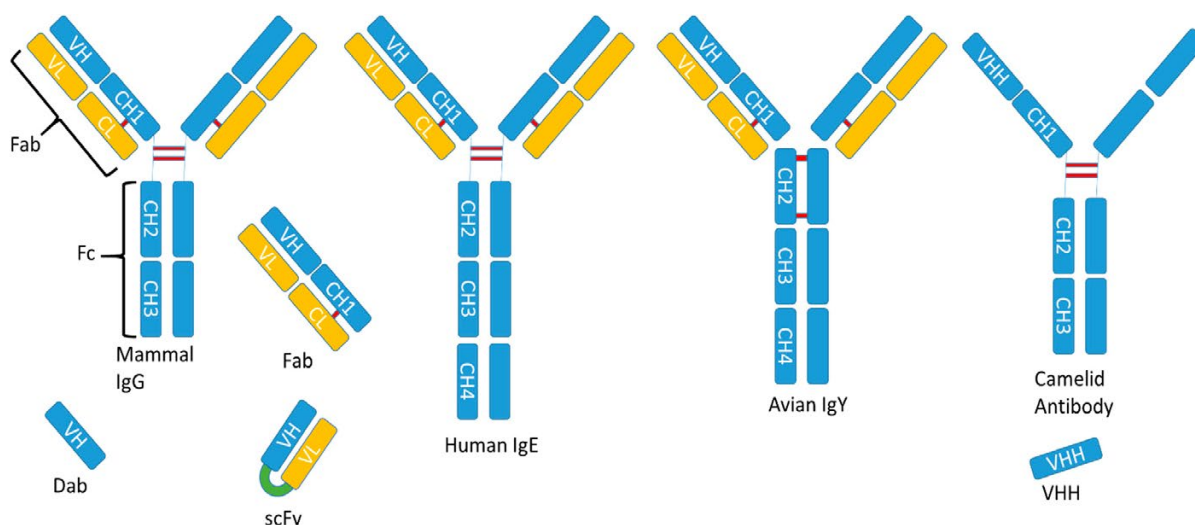
The other main group of detection methods for gluten is based on the amplification and detection of DNA by the PCR. This technique is highly sensitive (5–50 pg of DNA) and can lead to species identification (useful for cross-contamination studies). Real time PCR methods can be also used for the quantitative detection of gluten when using model mixtures [46,47]. Moreover, PCR methods perform indirect detection (not pathogenic compounds of gluten—rather the DNA that codifies for it) [48].

### 5. Immunological-Based Techniques for Gluten Detection in Food Samples

Immunological techniques are some of the most useful tools for gluten detection. The design of these methods is based in the obtainment of high affinity antibodies or their fragments, guided against noxious parts of gluten. However, though several immunogenic peptides in gluten have been identified, not all have been completely discovered and characterized.

Antibodies, also called immunoglobulins, are protective proteins produced by the immune system in response to the presence of a foreign substance, called an antigen. Most antibodies used as immune reactants are mammalian IgG. They have a four-chain structure (Figure 1), with two identical heavy chains and two light chains, and they are organized in three functional fractions (two antigen-binding arms, or Fabs, and the fragment crystallizable, or Fc, involved in cell effector systems). The light chain is composed of a constant and a variable region, whilst the heavy chain presents a variable region and a three-domain constant region [49]. Avian antibodies differ from mammals IgG molecules because they have an additional (CH4) domain, like human IgE, while lacking the hinge region that is observed in human IgGs [50]. In addition to these natural structures, many novel antibody fragment molecules have been developed by genetic engineering means, adapting these proteins to different uses. Fabs can be produced alone without attachment to the Fc [51]. Single-chain fragment variable antibodies (ScFvs) are unnatural structures composed of variable regions of light and heavy chains bonded by a flexible linker. Camels are mammals capable to produce a different kind of antibody that lack the light chain (are also known as heavy chain antibodies) and have a heavy chain composed of a variable single domain region (VHH) and two constant domains. Based on this structure, a new type of recombinant fragment consisting only of the VHH domain was developed (Figure 1) [52].

Immunoassays have been widely used since mid-20th century. The used antibodies can be classified in two types: polyclonal antibodies, which are obtained by animal immunization, and monoclonal antibodies, which can be obtained by hybridoma-based techniques [53,54] and by a recombinant protein approach.



**Figure 1.** Schematic representation of the structure of antibodies and derived molecules: Ig (Immunoglobulins); Fab (Fragment for antigen binding); Fc (Fragment crystallizable); VH (Variable region of Heavy chain); VL (Variable region of Light chain); CH (Constant regions of Heavy chain); CL (Constant region of Light chain); Dab (single Domain antibody); ScFV (Single chain Fragment Variable); VHH (Variable domain of Heavy chain of Heavy chain antibody).

## 6. Polyclonal Antibodies

The first immunoassays were developed thanks to polyclonal antibodies obtained from immunized animals' sera. The immunization process is well-known (Figure 2), and polyclonal antibodies used to be widely applied. As an example, to obtain polyclonal sera against gluten components in New Zealand white rabbits, emulsified gliadin and glutelin fragments were injected in Freund's complete adjuvant in the first shot, followed by two further shots of the antigen in Freund's incomplete adjuvant at two and four weeks later. Good responders (high serum titers) were intradermally given a booster of gliadin, the animals were bled, and antibodies were purified from antiserum by using the ammonium sulfate precipitation method [55]. Though a vast majority of polyclonal antibodies are isolated from mammals (mainly lagomorphs and rodents but also goats and horses), they have also been obtained in chicken embryos (IgY) (Figure 2). These IgYs against gluten have been used not only as reactive for immunoassays but also in prospective therapy for celiac disease intestinal damage [56].

Some examples of commercially available polyclonal antibodies, together with information about the way they have been obtained, can be found in Table 2.

**Table 2.** Some examples of polyclonal antibodies against gluten or its fragments.

Antibody	Company	Host	Isotype	Raised against
PAB29118	Abnova	Chicken	IgY	Wheat flour protein extract
MBS617177	MyBioSource	Rabbit	IgG	Wheat gluten
MBS838918	MyBioSource	Rabbit	IgG	Wheat gliadin
MBS625849	MyBioSource	Chicken	IgY	Wheat gluten
LS-C66756	LifeSpanBiosciences	Rabbit	IgG	Wheat gluten
LS-C129350	LifeSpanBiosciences	Chicken	IgY	Wheat gluten
LS-C750830	LifeSpanBiosciences	Chicken	IgY	Wheat gluten

Table 2. Cont.

Antibody	Company	Host	Isotype	Raised against
G8138-01	USBiological	Rabbit	IgG	Wheat gluten
G8138-02	USBiological	Chicken	IgY	Wheat flour protein extract
AS09 571	Agrisera	Chicken	IgY	Wheat flour protein extract
PA5-97536	Invitrogen	Rabbit	IgG	Wheat gliadin native protein
G9144	Sigma-Aldrich (Merck)	Rabbit	IgG	Native and heat-treated wheat gliadin

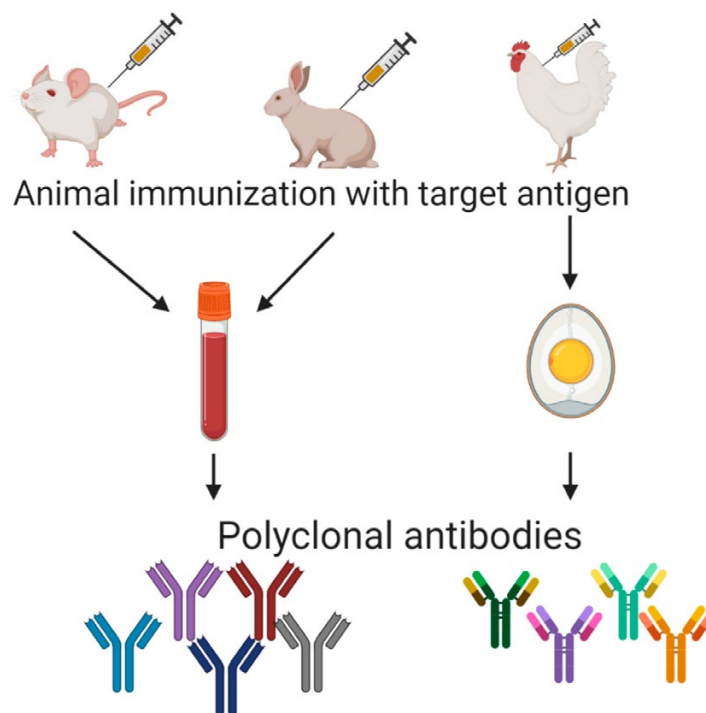


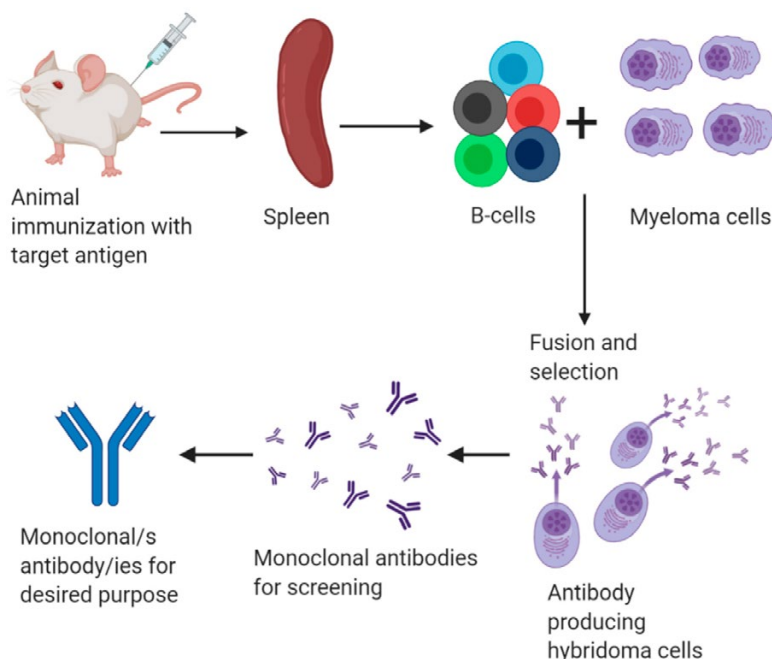
Figure 2. Representation of polyclonal antibodies production.

### 7. Hybridoma Secreted Monoclonal Antibodies

The research of George Kohler and César Milstein led to the production of monoclonal antibodies secreted by hybridoma cells (Figure 3). This was considered a groundbreaking innovation in many fields like therapeutics and diagnosis [53]. This technology has been also applied to the detection of gluten in foods. The main strength of monoclonal antibodies compared with polyclonal molecules raised in animals is their inter-batch evenness. Additionally, the gold standard method for gluten traceability in food is nowadays based on a monoclonal antibody [33].

One of the earliest developments was proposed by Skerrit and Underwood [57]. A protein fractionation from raw white wheat flour was performed for the immunization of BALB/c mice. Gliadin fractions were prepared by ion-exchange chromatography on sulfo-ethylcellulose to obtain  $\alpha\beta$  and  $\beta\gamma$  gliadin pools. Ethanol-precipitated, reduced alkylated glutenin for immunizations was prepared via the extraction of flour with sodium dodecyl sulphate (SDS) and 2-mercapthoethanol. Following the immunization of BALB/c mice with these gliadin and glutenin fractions, spleens were removed and used for fusion with SP 2/0 myeloma cells, and hybridoma cells were prepared and selected using well-established methods [58]. Supernatants from growing cultures were tested for antibodies to either gliadin or glutenins. Positive hybridomas were expanded and re-cloned by limiting dilution. Heavy-chain antibody-isotypes were determined, ultimately obtaining

six positive clones from mice immunized with gliadin (five IgG1 isotypes and one IgM), and 11 positive clones from mice immunized with glutenin (nine IgM and two IgG3) [57].



**Figure 3.** Schematic representation of common monoclonal antibody production.

A few years later, Skerrit and Hill [59] obtained the monoclonal antibody (IgG1) 401.21, raised against heat-stable  $\omega$ -gliadins, that recognizes the epitopes PQQPFQPE and PQQPPFPEE. This monoclonal antibody reacts with  $\omega$ -gliadins and the corresponding prolamins from rye and barley, as well as with high-molecular-weight glutenin subunits. A sandwich ELISA using the 401.21 antibody was adopted as Official Method 991.19 by the Association of Official Analytical Chemists International (AOACI) [33]. Though this antibody is no longer the basis of the gold-standard, it is currently available in some kits, e.g., the Aller-Tek™ Gluten ELISA assay (ELISA Technologies Inc.) or Veratox© for gliadin (Neogen).

Sorell et al. [60] developed a sandwich ELISA for gluten analysis in foods using a cocktail of monoclonal antibodies. BALB/c female mice were immunized with wheat, rye, and oat ethanol extracts. Splenocytes of the immunized animals were fused with P3/X63-Ag.653 myeloma cells. Selected hybridomas were grown as ascites in pristane-primed BALB/c mice, and antibodies were purified from ascites by affinity chromatography in a protein A-Sepharose column. Then, seven monoclonal antibodies were characterized (five raised against rye, named R1 to R5, 1 against oats, and one against gliadin, named 13B4). Most of the obtained antibodies displayed a wide cross-reactivity spectrum (R3 showed the highest) with gliadins, hordeins, and secalins. Some of them also cross-reacted with avenins but failed to recognize zeins. Six mAbs were assayed as coating antibodies in the sandwich ELISA using R3 conjugated to Horseradish Peroxidase (HRP) as the labeled antibody. The best results were obtained using 13B4, which allowed for the selective recognition of gliadins, and R5, which recognized secalins and hordeins. The R5 and 13B4 cocktail, as capture antibodies, and R3-HRP, as detection antibodies, permitted the recognition of gliadins, secalins, and hordeins to the same extent in the range of 3–200 ng/mL, thus improving the results obtained in food samples compared previously available commercial tests and, in many samples, better than mass spectrometry techniques.

Continuing this work, Valdés et al. [61] developed a novel sandwich ELISA using a single monoclonal antibody (R5) as both the capture (adsorbed) and detection (conjugated



Many studies demonstrated the complexity of celiac disease [67,71–73], thus leading to a paradigm shift in the field. The generation of mAbs should not only achieve a very low limit of detection of gluten but also target the determination of disease-inducing peptides. This was the aim of the work by Mitea et al. [74]; they immunized BALB/c mice with synthetic peptides corresponding to known T cell stimulatory epitopes that were coupled to a tetanus toxoid. The target peptides were found in the gliadins  $\text{glia-}\alpha 9$  (QPFQPQLPYP),  $\text{glia-}\alpha 20$  (PFRPQQPYQP),  $\text{glia-}\gamma 1$  (PQQSFPQQRPFIQPSL), LMW Glt-156 (PPFSQQQQSPFS), and HMW-Glt (PGQGQ(Q/P)GYYPST(L/Q)QQPQGQQGYYPSTPQQ(P/S)). Because many gluten proteins share a high degree of homology, the authors aimed to prove whether the obtained monoclonal antibodies reacted specifically with only the peptide used for immunization or they detected the other T cell stimulatory sequences. They found that the  $\text{glia-}\alpha 20$ -specific antibody also reacted with the  $\text{glia-}\alpha 9$  and  $\text{glia-}\gamma 1$  peptides. Moreover, the obtained monoclonal antibodies against  $\text{glia-}\alpha 9$ ,  $\text{glia-}\alpha 20$ , Glt-156, and HMW-Glt reacted with gluten peptides that are naturally formed during digestion in the gastrointestinal tract, as resulting from the activity of pepsin and trypsin. In addition, all except the LMW-specific antibodies also detected storage proteins in barley and rye, whereas the  $\text{glia-}\gamma 1$  specific antibodies also recognized oat proteins. Finally, compared to Ridascreen<sup>®</sup> Gliadin kit (R5-based), the in-house ELISA for the  $\text{glia-}\alpha 9$  epitope detected higher gluten concentrations in all analyzed food samples. A competitive ELISA based on the monoclonal antibody anti- $\text{glia-}\alpha 20$ , called Gluten-Tec<sup>®</sup>, was commercialized by EuroProxima and validated by an interlaboratory study [75], presenting a Limit of Quantification (LOQ) lower than that of the R5 methods (3.6 vs. 5 ppm of gluten, respectively). Monoclonal antibodies were later obtained against the same and different T cell epitopes ( $\text{glia-}\gamma 1$ , Glt-156, a variant of HMW-gly, and eight peptides). Five antibodies were selected (one anti- $\alpha 1$ -gliadin, one anti- $\gamma 1$ -gliadin, two anti-LMW, and one anti-HMW). This method is protected under patent WO2006004394A2.

New challenges appeared after the early 2000s-emergence of severe allergies linked to the ingestion of food products containing even small amounts of hydrolyzed wheat proteins. The main signs and symptoms were found to be WDEIA, anaphylaxis, and urticaria. Sensitization was often related to exposure through cosmetic products. It was described that the triggering ingredient of the mentioned cases of food allergies was the hydrolyzed wheat proteins, where glutamines are converted to glutamic acid by deamidation reactions occurring at high temperature and low pH conditions during industrial processing. Deamidation leads to the appearance of novel IgE epitopes, and it was found that prone patients' sera presented a higher reactivity to deaminated gluten proteins, especially in  $\alpha$ ,  $\gamma$ ,  $\omega 2$ ,  $\omega 5$ , and LMW [76].

Regarding this new kind of allergenicity, Tranquet et al. used a new immunodominant neo-epitope (QPPEPFPE derivate of the deamidation of QPQQPFQ) to produce a mAb raised in mice against a peptide that includes the neoepitope [77]. Then, VH and VL were cloned into expression plasmids modified to express IgE-heavy chains to produce a recombinant chimeric IgE [78]. Expression was performed in mammal cells (HEK293). Sera from allergic patients and the recombinant antibody were analyzed by ELISA (with deaminated gluten fractions) and functional basophil assays. It was demonstrated that acid-hydrolyzed wheat proteins that presented higher deamidation levels displayed a stronger IgE binding ability and a higher basophil activation capacity. Moreover, the recombinant antibody allowed for basophil degranulation in the presence of deaminated wheat proteins, mimicking patients' IgE properties. This work is an example of another available antibody methodology: the transformation of a "classical" mAb to a recombinant one.

Monoclonal antibodies secreted by hybridoma cells have been proven as a very reliable tool for designing and improving antibody-based gluten detection systems. Some of these developments have been protected by patents, and they are summarized up in Table 3. This research for patent protected monoclonal antibodies and tests for gluten detection was conducted with worldwide.europacenet.com, an international patent database powered

by the European Patent Office. Nowadays, most commercially available tests for gluten detection are based on the monoclonal antibodies 401.21, R5, G12, and anti- $\alpha$ 20 [33].

**Table 3.** Several examples of registered patents claiming monoclonal antibodies and antibody-based methods related to gluten detection.

Patent No.	Summarized Patents	Applicant
WO2006004394A2	A method for the screening of basic ingredients, semi-manufactured ingredients, and food products that are intended to be used in a gluten-free diet, based on antibodies raised against T cell stimulatory peptides.	Academisch Ziekenhuis Leiden (The Netherlands)
WO2006051145A1	Competitive ELISA for the detection of gluten hydrolysate based on the R5 monoclonal antibody.	Consejo Superior de Investigaciones Científicas (Spain)
WO2007104825A1	Method for extracting gluten from processed (by heat) and unprocessed foods based on the use of ionic and non-ionic detergents as prior step for ELISA tests.	Consejo Superior de Investigaciones Científicas (Spain)
WO2014132204A1	Monoclonal antibody that is capable of bonding to deamidated gluten proteins (related with celiac disease pathogenesis) and has no cross-reaction with non-deamidated gluten proteins.	Institut national de la recherche agronomique (France)
WO2015164615A1	Isolated monoclonal antibodies and fragments that bind to 11 peptides that can be found in gluten proteins.	University of Chicago (USA) and University of Oslo (Norway)
WO2018071718A1	Antibodies, fragments, or polypeptides in the detection of gliadin: heavy chain and light chain variable sequences, and associated sequences of complementarity-determining regions (CDRs).	Nima Labs Inc. (USA)
WO2019154559A1	Immunoassay methods for the quantification of the total gluten content of grains in food samples.	R-Biopharm AG. (Germany)
ES2392412A1	Solutions for the extraction and solubilization of gluten, composed of arginine and ethanol.	Biomedal S.L (Spain)
GB2207921A	Hybridoma cell line ATCC HB9798 that produces monoclonal antibodies directed against omega gliadin protein of wheat and related proteins in rye and barley.	Commonwealth Scientific and Industrial Research Organization (United Kingdom)
CN101698832A	Anti-gliadin monoclonal antibody and the hybridoma cell line obtained from it.	Quingdao Biomade Technology Company Ltd. (China)
CN107860918A	Colloidal gold immunochromatography test strip for the gluten allergen in food and the preparation method of colloidal gold immunochromatography test strip.	Biofront Technology Company Ltd. (China)

## 8. A New Era: Directed Evolution of Recombinant Antibodies

Though the introduction of hybridoma technology led to a great leap forward for the development of antibodies in many fields and forever changed therapeutics and diagnostics, there are still some weaknesses mainly related to the use of animal antibodies for human applications to be solved [79]. In the development of gluten-detection methods, the reliable generation and production of high-affinity antibodies had been achieved thanks to hybridoma technology, and detection limits have been lowered enough to make these antibodies widely used. However, antibodies raised in animals have not presented direct correlation to the human response in many cases (e.g., the humoral response is quite different in a pathological situation, like celiac disease, compared to that produced during animal immunization). Thus, new needs separate from detecting gluten in food samples, like the detection of potentially disease-inductor protein motifs, have appeared. Bio-ethic issues of animal use and sacrifice are also important, and the development of alternatives to the use of animal experimentation is still required [80].

In this context, the concept of “directed molecular evolution,” implies making antibodies tailored to, and uniquely suited for, human purposes [81,82]. This concept and its applications rely on the development of genetic engineering and heterologous protein expression. Nevertheless, the molecular evolution of antibodies happens in nature, including human beings, where billions of different antibodies can be produced with three unlinked loci containing the immunoglobulin gene segment [83]. This powerful machine can be replicated and guided to obtain a specific application for gluten detection. The natural machinery of the immune system for making antibodies can be summarized up in five steps: (a) the rearrangement of variable (V) gene segments [84], (b) the surface display of an antibody on a B cell, (c) antigen-driven selection, (d) the secretion of soluble antibody from a plasma cell, and (e) affinity maturation [79].

The directed evolution of antibodies needs to overcome two limiting conditions: the generation of enough diversity (referred to as building a library of coding genes for antibody chains), and development of an adequate system for selection and amplification.

Antibody libraries can be classified into four main groups [85]:

- Immune: constructed based on amplifications of variable (V) genes isolated from immunoglobulin-secreting plasma cells from immunized donors [86].
- Naïve: based on amplifications of V genes isolated from immunoglobulin-secreting plasma cells from non-immunized donors [87].
- Semi-synthetic: derived from unrearranged V genes from pre-B cells (germline cells) or a single antibody framework with genetically randomized complementarity determining regions (CDRs) [88].
- Synthetic: based on a human antibody framework with randomly integrated CDR cassettes [89].

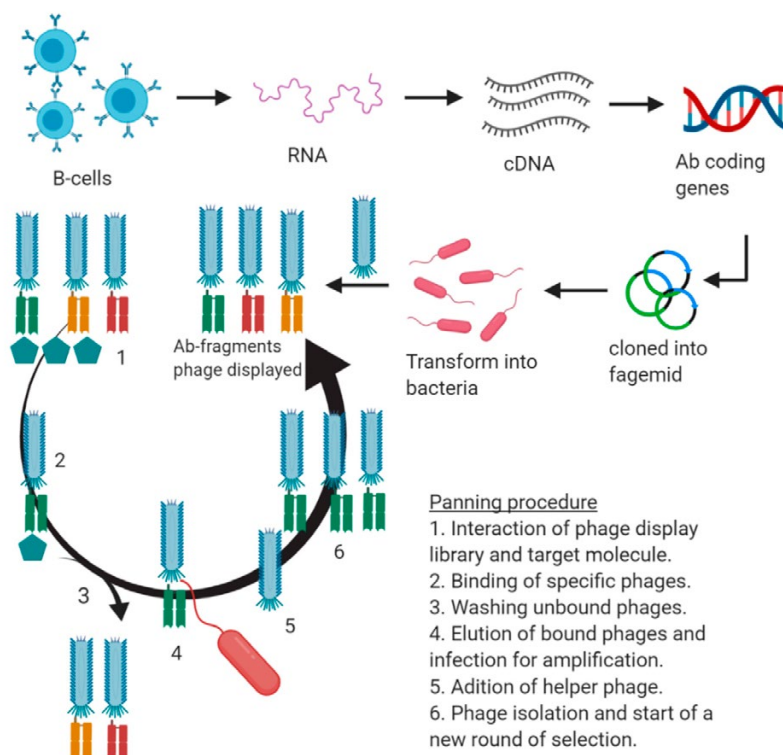
Immune libraries are usually constructed by a two-step cloning (heavy and light chains) assembly PCR method, after mRNA isolation and cDNA preparation from the desired cell type [85]. Naïve, semisynthetic, and synthetic libraries are considered “single-pot” libraries, which means that they can be used for picking an antibody that binds with (almost) every antigen that can be presented [90]. The affinity of antibodies developed from this kind of libraries directly depends on the repertory size RZ (number of different binders within the library). To achieve an affinity in the micromolar ( $\mu\text{M}$ ) range, an RZ =  $10^7$  is required, but an RZ value of  $10^{10}$  is necessary for an affinity of the nanomolar (nM) range. However, immune libraries must be designed and constructed specifically for every single antigen or group of related antigens, and the resulting antibody affinity is driven by their immunogenicity (nanomolar scale is possible with a very immunogenic antigen) [91].

The antibody coding genes that conform a library must be cloned in the appropriate vector to develop an adequate system of selection and amplification [85]. Once a library is properly constructed (with enough diversity and cloned in a suitable vector), a process called panning or biopanning (successive rounds of selection and amplification) is performed. For biopanning, it is necessary to express all the antibodies or their fragments included in a library [92]. It has been established that if there are 1–10 binders in  $10^7$  clones before selection, there should be 1–10 binders in 10 clones after three-to-five rounds of panning. This process allows for the enrichment of a set of high affinity binders from the library, pairing genotypes (single-strand phagemids that can be “rescued” with a helper phage) with phenotypes (phage-expressing antibodies).

Peptides, protein domains, and antibody fragments can be displayed for selection and amplification in various ways: phage display [93], in vitro RNA display (divided into ribosome display [94] and cDNA display [95]), and cell surface display [96].

The phage display of antibody fragments libraries is the most widely used method for the production of recombinant antibodies (Figure 4). The phage display methodology was designed by Smith et al. [97], who introduced gene coding for the bacterial enzyme EcoRI into the gene III of a native bacteriophage and achieved the heterologous expression of EcoRI fused to the phage protein III. Then, a vector called fUSE5 was designed to improve the introduction of exogenous DNA to the filamentous phages and for protein

expression [98]. The next step was the development of the affinity selection procedure: (a) the immobilization of a selector, (b) the addition of input virions (with the displayed peptides that bind the selector), (c) the washing away unbound virions, (d) the release of bound virions, and (e) the amplification of the released virions by infecting a proper bacteria [93].



**Figure 4.** Construction and performance of a directed evolution process mediated of an immune library by phage display.

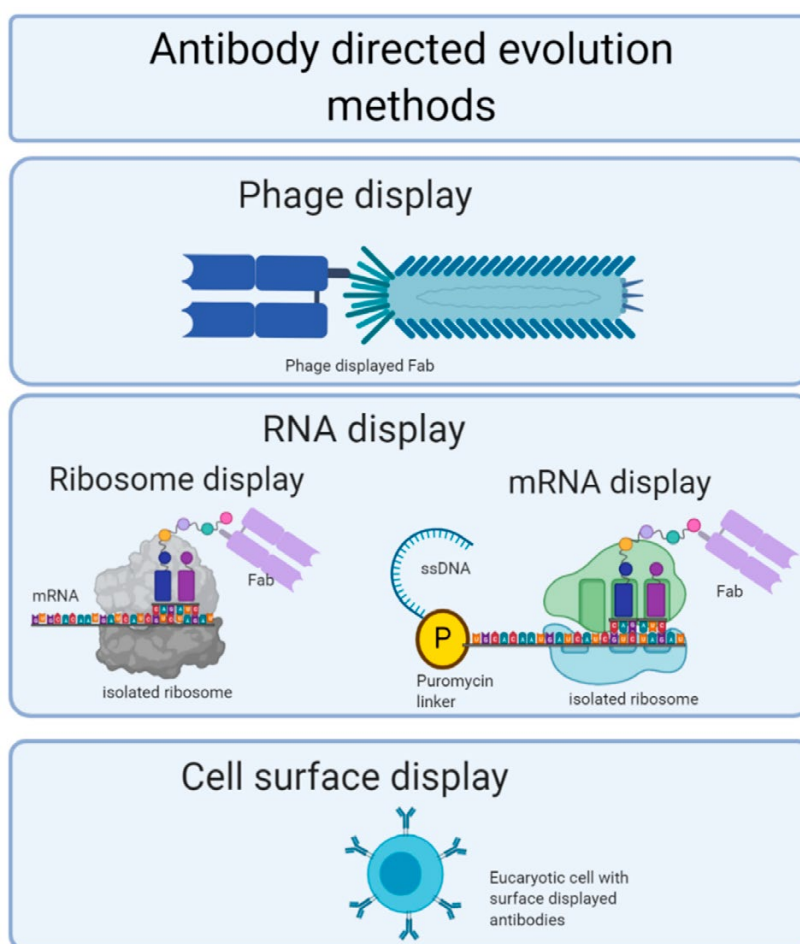
The first application of phage display was the epitope mapping of a given antibody [93,99]. There have been huge developments of this system with different applications, although the methodological basis has not significantly changed since Smith and collaborators achieved the proof of concept [93].

New vectors have been developed with some common features: a double replication origin (f1 or similar for protein phage fusion and an *Escherichia coli* ori), a selection marker (usually an antibiotic resistance gene), and binder-coding genes (usually peptides, protein domains, or antibody fragments) fused to a phage protein (usually capsid proteins) [100].

Vectors can be classified to several types according to their design. If the binder is fused to phage gene III, it will be expressed fused to protein III (up to five copies), constituting a type 3 phagemid. However, if it is fused to gene VIII, it will be expressed fused to protein VIII (up to 2400 copies), constituting a type 8 phagemid. Types 33 and 88 follow the same concept, but they have two copies of either gene III or VIII. Types 33 and 88 are also phagemid-based, but the virus particles are only formed in cells carrying the phagemid genome when they are infected with the helper phage virus [101]. Though pIII and pVIII-based vectors are the most frequently used, there have been some attempts to clone binders in the VII and IX genes, which are minor capsid proteins [102].

Due to the variability of culturing cells, new cell-free methods have been developed with two main variants: ribosome and mRNA (or cDNA) displays (Figure 5). The ribosome display is performed as follows: a DNA encoding library is in vitro transcribed to a single

strand mRNA lacking stop codon, then it is *in vitro* translated to obtain a native protein attached to a ribosome, and affinity selection is performed. Then, mRNA is released from the ribosomes and reverse transcribed to single strand DNA that can be replicated and/or mutated, thus becoming the base of another round of selection [103]. mRNA display improves the coupling between genotypes and phenotypes. The process starts with the transcription of a DNA library; then, mRNA is ligated to a DNA linker connected to puromycin. The *in vitro* translation of this complex allows for a peptidyl transferase reaction that results in a covalently linked mRNA–protein complex that has the puromycin–linker–mRNA. Then, the mRNA is retrotranscribed, and affinity selection is performed. Finally, high-affinity complexes are eluted, releasing only DNA strands by RNA hydrolysis. This DNA can be mutated and is the starting material for the subsequent round [94]. A cell-surface display (Figure 5) allows peptides and proteins to be displayed on the surface of cells by fusing them with anchoring motifs [96] instead of using bacteriophages. Protein expression improves by using more complex organisms like Gram-negative bacteria [104], Gram-positive bacteria [105], yeasts [106], and mammalian cells [107].



**Figure 5.** Schematic representation of antibody-directed evolution methods.

Directed-evolution methods have become groundbreaking technologies for antibody development at a similar level to hybridoma-derived monoclonal antibodies. In the field of gluten-related diseases, they have been applied with two main objectives. Most of the published works have been focused on using a phage display as a tool for celiac disease molecular characterization. The other significant approach is to apply recombinant antibody-

ies to develop gluten-detection systems, because hybridoma antibodies were previously used. In the context of the molecular characterization of celiac disease, phage display technology has mainly been used to develop high-affinity antibodies against human tissue transglutaminase, an autoantigen with a major role in celiac disease [108].

An early strategy by Sblattero et al. [109] was to study the immune response against tissue transglutaminase by building an immune library from the peripheral blood lymphocytes (PBLs) of a patient with celiac disease and performing a selection against four related antigens ( $\alpha$ -gliadin,  $\beta$ -lactoglobulin, human tissue transglutaminase, and guinea pig transglutaminase). Some polyreactive and monoreactive antibodies were obtained against the first two antigens but not to transglutaminase. VH4 was the main family of the anti- $\alpha$ -gliadin selected antibodies. Following this approach, Mazari et al. [110] produced and analyzed six immune ScFv libraries from peripheral and intestinal lymphocytes (IBLs) collected from three celiac patients, and they concluded that intestine-derived antibodies from all selected patients recognized human transglutaminase (with a bias toward the use of the VH5 family), whereas peripheral blood-derived antibodies recognized  $\alpha$ -gliadin. Following this research, Sblattero et al. [111] built up an ScFv library based on the amplification of the two VH5 family genes from intestinal lymphocytes, resulting in a rapid characterization of the anti-transglutaminase response that could be applied in asymptomatic patients whose serum antibodies may be undetectable.

Phage display has been revealed as a very reliable method for celiac immunity characterization. Not et al. [112] generated an immune library by amplifying the VH5-51 gene from bowel biopsy specimens of 22 relatives of celiac patients and analyzing its interaction with human transglutaminase. They found that genetically predisposed individuals to celiac disease produce VH5 anti-transglutaminase intestinal antibodies (anti-TG2) in the absence of serum anti-TG2 antibodies.

Another use of phage display libraries for the characterization of celiac disease was the work of Hoydal et al. [113], who investigated the antigen presentation process during mucosal immune response. They applied a large naïve human ScFv library for the isolation of specific binders against the complex HLA-DQ2.5:DQ2.5-gliadin- $\alpha$ 1a. Then, the obtained antibodies were applied to cells from intestinal biopsies from patients with celiac disease, which allowed for the identification of plasma cells as the most abundant gluten peptide-MHC-expressing cells in inflamed intestinal tissues from celiac patients.

Rhyner et al. [114] reported the construction of three unique isotype ScFv libraries (IgA, IgG, and IgM) from a celiac patient and a healthy control individual, and they demonstrated that all libraries from the celiac patient, but none from the control donor, were selectively enriched in gliadin-binding phage clones after four rounds of biopanning. This method not only resulted in a suitable approach to obtain high-affinity antibodies against gliadin but also allowed for the isotype-specific characterization of the immune responses occurring in a pathological condition.

Phage display technology can be also used as a scanning method to study biological interactions when conducted with peptides libraries. An example of this concept in celiac disease was developed by Chen et al. [115]; a random peptide phage display library was enriched in gliadin-binding peptide-phages by several panning cycles against immobilized gliadin proteins. Several peptide-expressing phage clones were able to inhibit the interaction between gliadin and anti-gliadin antibodies. Moreover, the 12-mer peptides encoded by selected clones were synthesized *in vitro* and analyzed in competition experiments that revealed the binding of different peptides to different sites of the gliadin. The authors suggested the potential use of these peptides to detoxify gluten.

All mentioned works (summarized up in Table 4) are examples of the potential of phage display as a tool for celiac disease molecular characterization. Increased understanding of celiac disease has led to better diagnosis and prevention measures, like those derived from the discovery of the major role of tissue transglutaminase and toxic components of gluten.

**Table 4.** Several examples of phage display technology applied to celiac disease research.

Type of Library	Antibody Format	Isotype VH Family Gene/s	Selection Driven by	References
Immune (peripheral blood lymphocytes (PBLs) from celiac patients)	Single-chain fragment variable antibody (ScFv)	IgG VH4	$\alpha$ -gliadin human transglutaminase Other antigens	[109]
Immune (PBLs and intestinal lymphocytes (IBLs) from celiac patients)	ScFv	VH5	Human transglutaminase $\alpha$ -gliadin	[110]
Immune (IBLs from celiac patients)	ScFv	VH5	Human transglutaminase	[111]
Immune (IBLs from celiac patient relatives)	ScFv	VH51-1	Human transglutaminase	[112]
Naïve	ScFv	NA	Gliadin HLA-presenting peptides	[113]
Immune	ScFv	IgA, IgG, and IgM	Gliadin	[114]
Random peptides	Random peptides	NA	Gliadin HLA-presenting cells	[115]

Though most of the published phage-display applications related to gluten have been focused on studying celiac disease antibody response, there is a growing interest in the development of recombinant antibodies for the detection of gluten in foodstuffs as a final application (Table 5).

**Table 5.** Several examples of directed evolution methods for the obtention of antibodies for detection of gluten in foods.

Technology Used	Type of Library	Antibody Format	Selection Driven by	Gluten Detection Method	References
Phage display	Immune	VHH	Gliadin	Capture ELISA	[116]
Phage display	Naïve	single-domain antibody (dAb)	Consensus gluten peptide	Phage ELISA	[117]
cDNA display	Naïve	VHH	Gliadin	cDNA display mediated immuno-PCR (cD-IPCR)	[118]

Doña et al. [116] developed a system for gluten detection in food samples based on VHH antibodies. A VHH phage display library was constructed from PBLs isolated from gliadin-immunized llamas, and the selection of gliadin-binding VHH was performed under denaturing conditions. The selected VHH allowed for the development of a capture ELISA system (using gliadin-binding VHH as capture and anti-gliadin mouse-derived monoclonal antibody for detection) able to detect gliadin in samples that tested negative with other ELISA kits. However, the method was only applicable to wheat gluten detection, as the selected VHH did not react to barley or rye prolamins.

García-García et al. [117] designed and developed a phage-ELISA method for gluten analysis in foodstuffs based on single-domain antibody (or dAb) fragments using a semi-synthetic library developed by Christ et al. [119]. The library was enriched in high-affinity dAbs by successive rounds of selection against the consensus peptide CPFPPQQPFPPQQPF-PQQQPFQQQPFQQPFQQQPFQQQP [120] that includes epitopes that can be found repeatedly in gluten proteins. A phage ELISA method was used to screen 50 commercially available food products, with a limit of detection of 20 ppm of gluten. The method was able to ascertain compliance with the labelling of gluten-free products. Moreover, this work completed an animal-free antibody developing process applicable to the detection of gluten in foods.

Phage display is not the only directed evolution technique that has been applied in gluten research. Jayathilake et al. [118] demonstrated the application of a cell-free display method in this field. A VHH cDNA library was developed from alpaca lymphocytes to pair genotypes (cDNA) and phenotypes (VHH) thanks to a puromycin linker. Several rounds of selection were performed against immobilized gliadin to obtain three high-affinity binders that were used as the basis of a novel gluten detection technique called cDNA display

mediated immuno-PCR (cD-IPCR) [121]. The cD-IPCR method was able to detect very low gliadin concentrations (0.001–10 µg/mL) in food. This method can be improved with the objective of obtaining effective binders for toxic and complex proteins, which is difficult to achieve using conventional methods.

VHH-, dAb- and cD-IPCR-based methodologies have become pioneers for a new generation of gluten traceability systems.

### 9. Final Remarks

Nowadays, the detection and traceability of allergens and related substances (like gluten) has become one of the hottest topics in the food and drink industry due to the following issues: (a) consumer demand for accessible, clear, and accurate label information; (b) growing health and nutrition social concerns; (c) an increasing prevalence of food-related diseases (as a result of better diagnosis and knowledge on their pathogenesis); and (d) the consumption of new foods and ingredients derived from a competitive and globalized market.

This situation is a continuous challenge for food science that has demanded the design and development of new gluten traceability systems that can work with changing and stricter regulated limits of detection and quantification. New developments will be focused not only on lowering gluten detection limits but also on identifying those components that are able to trigger gluten-related diseases.

Improving the immunodetection of gluten in foods depends on the obtention of better antibodies. Chronologically, the strategies used for developing these antibodies can be summarized as follows: (1) polyclonal antibodies raised by animal immunization comprised the first method that is still currently in use, with two main variants: mammal-derived IgG and chicken-derived IgY; (2) monoclonal antibodies like 401.21 and R5 were produced from hybridomas raised against different cereal protein extracts; (3) monoclonal antibodies like G12, produced from hybridomas, were raised against recombinant proteins that were implied in celiac pathogenesis; and (4) recombinant antibodies were obtained by directed molecular evolution with phage or cDNA display technology. Each of the technologies outlined in this review (polyclonal, monoclonal, and recombinant antibodies) has advantages and disadvantages (Table 6). The selection of the appropriate methodology will depend on the intended use and resources available.

**Table 6.** Comparison of technologies available for production of antibodies for gluten detection in foods.

Technology	Market Available Test for Gluten Detection Based on These Technologies	Validated Tests for Gluten Detection Based on These Antibodies	Clinical Applications Developments with These Antibodies	Inter-Batch Evenness	Need Animal Experimentation for Its Development	Technical Readiness
Polyclonal antibodies	Yes	Yes	Yes	Variable	Yes	Mature
Monoclonal antibodies	Yes	Yes	Yes	Yes	Yes	Mature
Recombinant antibodies	Not yet	Not yet	Yes	Yes	Not all	Recent

Future perspectives for this field include the discovery of new antibodies, with an increasing affinity for disease triggering gluten peptides. For such a purpose, it is necessary to continue to unveil the molecular mechanisms of gluten-related pathologies, like celiac disease (the most widely studied) and others, particularly non-celiac gluten disease, that present the highest prevalence.

An integrative strategy is currently being applied. Building antibody libraries from gluten-sensitive patients should allow for the study of immune response against gluten and the production of recombinant antibodies with the same “behavior” as those found in patients, thus making it possible to detect those protein motifs that are dangerous for prone people in foods.

Additionally, there has been intense research into new devices that unify gluten extraction and antibody-mediated detection, as well as the development of easier and faster analysis methodologies that can move from laboratory applications to consumer-friendly devices to be used in homes or restaurants.

Preventing gluten-derived food risks requires of a multi-disciplinary approach that implies basic and clinical research, biotechnological and engineering developments, and a farm-to-table philosophy; these goals will only be reachable with the help from all value-chain members, from raw material producers to final consumers.

**Author Contributions:** Conceptualization, E.G.-C., A.G.-G., and R.M. (Raquel Madrid); writing—original draft preparation, E.G.-C., A.G.-G., and R.M. (Raquel Madrid); writing—review and editing, E.G.-C., T.G., and R.M. (Rosario Martin); funding acquisition, T.G. and R.M. (Rosario Martin). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by MINISTERIO DE CIENCIA E INNOVACION, grant number AGL2017-84346-R and COMUNIDAD DE MADRID grant number FOODAL-CM S2018/BAA-4574

**Acknowledgments:** Figures 2–5 were generated with Biorender.com.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fasano, A.; Catassi, C. Clinical practice. Celiac disease. *N. Engl. J. Med.* **2012**, *367*, 2419–2426. [[CrossRef](#)]
2. Codex Alimentarius. Codex Standard for Foods for Special Dietary Use for Persons Intolerant to Gluten. Revised 2008. CODEX Standard 118-1979. Available online: [http://www.fao.org/fao-who-codexalimentarius/sh-proxy/es/?lnk=1&url=https%2F3A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252Fstandards%252FCXS%2B118-1979%252FCXS\\_118e\\_2015.pdf](http://www.fao.org/fao-who-codexalimentarius/sh-proxy/es/?lnk=1&url=https%2F3A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252Fstandards%252FCXS%2B118-1979%252FCXS_118e_2015.pdf) (accessed on 29 December 2020).
3. Goesaert, H.; Brijs, K.; Veraverbeke, W.S.; Courtin, C.M.; Gebruers, K.; Delcour, J.A. Wheat flour constituents: How they impact bread quality, and how to impact their functionality. *Trends Food Sci. Technol.* **2005**, *16*, 12–30. [[CrossRef](#)]
4. Martín, M.; Cacho, J.F.; Cepeda, A.; Martín, F.; Prieto, I. *Informe del Comité Científico de la Agencia Española de Seguridad Alimentaria y Nutrición (AESAN) en Relación con la Enfermedad Celíaca y los Problemas que Plantean las Técnicas Analíticas para el Control del Contenido de Gluten en los Alimentos*; AESAN: Madrid, Spain, 2010.
5. Shewry, P.R.; Halford, N.G.; Tatham, A.S.; Popineau, Y.; Lafandra, D.; Belton, P.S. The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Adv. Food Nutr. Res.* **2003**, *45*, 219–302. [[CrossRef](#)] [[PubMed](#)]
6. Cebolla, Á.; Moreno, M.; Coto, L.; Sousa, C. Gluten Immunogenic Peptides as Standard for the Evaluation of Potential Harmful Prolamin Content in Food and Human Specimen. *Nutrients* **2018**, *10*, 1927. [[CrossRef](#)] [[PubMed](#)]
7. Wieser, H. Chemistry of gluten proteins. *Food Microbiol.* **2007**, *24*, 115–119. [[CrossRef](#)] [[PubMed](#)]
8. Shewry, P. What Is Gluten—Why Is It Special? *Front. Nutr.* **2019**, *6*, 101. [[CrossRef](#)] [[PubMed](#)]
9. Ozuna, C.V.; Iehisa, J.C.M.; Giménez, M.J.; Alvarez, J.B.; Sousa, C.; Barro, F. Diversification of the celiac disease  $\alpha$ -gliadin complex in wheat: A 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant J.* **2015**, *82*, 794–805. [[CrossRef](#)] [[PubMed](#)]
10. Schalk, K.; Lang, C.; Wieser, H.; Koehler, P.; Scherf, K.A. Quantitation of the immunodominant 33-mer peptide from  $\alpha$ -gliadin in wheat flours by liquid chromatography tandem mass spectrometry. *Sci. Rep.* **2017**, *7*, 45092. [[CrossRef](#)]
11. Balakireva, A.; Zamyatnin, A. Properties of Gluten Intolerance: Gluten Structure, Evolution, Pathogenicity and Detoxification Capabilities. *Nutrients* **2016**, *8*, 644. [[CrossRef](#)]
12. Matsuo, H.; Yokooji, T.; Taogoshi, T. Common food allergens and their IgE-binding epitopes. *Allergol. Int.* **2015**, *64*, 332–343. [[CrossRef](#)]
13. Caio, G.; Volta, U.; Sapone, A.; Leffler, D.A.; De Giorgio, R.; Catassi, C.; Fasano, A. Celiac disease: A comprehensive current review. *BMC Med.* **2019**, *17*, 142. [[CrossRef](#)] [[PubMed](#)]
14. Sapone, A.; Bai, J.C.; Ciacci, C.; Dolinsek, J.; Green, P.H.; Hadjivassiliou, M.; Kaukinen, K.; Rostami, K.; Sanders, D.S.; Schumann, M.; et al. Spectrum of gluten-related disorders: Consensus on new nomenclature and classification. *BMC Med.* **2012**, *10*, 13. [[CrossRef](#)] [[PubMed](#)]
15. Palomares, O.; Akdis, M.; Martín-Fontecha, M.; Akdis, C.A. Mechanisms of immune regulation in allergic diseases: The role of regulatory T and B cells. *Immunol. Rev.* **2017**, *278*, 219–236. [[CrossRef](#)] [[PubMed](#)]
16. Al Badri, F.M.; Baatjies, R.; Jeebhay, M.F. Assessing the health impact of interventions for baker's allergy and asthma in supermarket bakeries: A group randomised trial. *Int. Arch. Occup. Environ. Health* **2020**, *93*, 589–599. [[CrossRef](#)]
17. Sharma, N.; Bhatia, S.; Chunduri, V.; Kaur, S.; Sharma, S.; Kapoor, P.; Kumari, A.; Garg, M. Pathogenesis of Celiac Disease and Other Gluten Related Disorders in Wheat and Strategies for Mitigating Them. *Front. Nutr.* **2020**, *7*, 6. [[CrossRef](#)]
18. Gao, X.; Wen, L.; Li, H.; Wang, R.; Yin, J. Genetic variation at the interleukin-18 locus is associated with wheat-dependent exercise-induced anaphylaxis in the Han Chinese population. *Gene* **2020**, *737*, 144462. [[CrossRef](#)]

19. Hollon, J.; Puppa, E.; Greenwald, B.; Goldberg, E.; Guerrerio, A.; Fasano, A. Effect of Gliadin on Permeability of Intestinal Biopsy Explants from Celiac Disease Patients and Patients with Non-Celiac Gluten Sensitivity. *Nutrients* **2015**, *7*, 1565–1576. [CrossRef]
20. Kim, S.M.; Mayassi, T.; Jabri, B. Innate immunity: Actuating the gears of celiac disease pathogenesis. *Best Pract. Res. Clin. Gastroenterol.* **2015**, *29*, 425–435. [CrossRef]
21. Junker, Y.; Zeissig, S.; Kim, S.-J.; Barisani, D.; Wieser, H.; Leffler, D.A.; Zevallos, V.; Libermann, T.A.; Dillon, S.; Freitag, T.L.; et al. Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *J. Exp. Med.* **2012**, *209*, 2395–2408. [CrossRef]
22. Stamnaes, J.; Sollid, L.M. Celiac disease: Autoimmunity in response to food antigen. *Semin. Immunol.* **2015**, *27*, 343–352. [CrossRef]
23. Pagliari, D.; Urgesi, R.; Frosali, S.; Riccioni, M.E.; Newton, E.E.; Landolfi, R.; Pandolfi, F.; Cianci, R. The Interaction among Microbiota, Immunity, and Genetic and Dietary Factors Is the Condicio Sine Qua Non Celiac Disease Can Develop. *J. Immunol. Res.* **2015**, *2015*, 123653. [CrossRef] [PubMed]
24. Schumann, M.; Richter, J.F.; Wedell, I.; Moos, V.; Zimmermann-Kordmann, M.; Schneider, T.; Daum, S.; Zeitz, M.; Fromm, M.; Schulzke, J.D. Mechanisms of epithelial translocation of the alpha(2)-gliadin-33mer in coeliac sprue. *Gut* **2008**, *57*, 747–754. [CrossRef] [PubMed]
25. Hadjivassiliou, M.; Sanders, D.S.; Woodroffe, N.; Williamson, C.; Grünewald, R.A. Gluten ataxia. *Cerebellum* **2008**, *7*, 494–498. [CrossRef] [PubMed]
26. Volta, U.; De Giorgio, R.; Caio, G.; Uhde, M.; Manfredini, R.; Alaedini, A. Nonceliac Wheat Sensitivity: An Immune-Mediated Condition with Systemic Manifestations. *Gastroenterol. Clin. N. Am.* **2019**, *48*, 165–182. [CrossRef]
27. Roszkowska, A.; Pawlicka, M.; Mroczek, A.; Bałabuszek, K.; Nieradko-Iwanicka, B. Non-Celiac Gluten Sensitivity: A Review. *Medicina* **2019**, *55*, 222. [CrossRef]
28. Kurada, S.; Yadav, A.; Leffler, D.A. Current and novel therapeutic strategies in celiac disease. *Expert Rev. Clin. Pharmacol.* **2016**, *9*, 1211–1223. [CrossRef]
29. Itzlinger, A.; Branchi, F.; Elli, L.; Schumann, M. Gluten-Free Diet in Celiac Disease—Forever and for All? *Nutrients* **2018**, *10*, 1796. [CrossRef]
30. Ludvigsson, J.F.; Bai, J.C.; Biagi, F.; Card, T.R.; Ciacci, C.; Ciclitira, P.J.; Green, P.H.R.; Hadjivassiliou, M.; Holdaway, A.; Van Heel, D.A.; et al. Diagnosis and management of adult coeliac disease: Guidelines from the British Society of Gastroenterology. *Gut* **2014**, *63*, 1210–1228. [CrossRef]
31. Commission Implementing Regulation (EU) No. 828/2014 of 30 July on the Requirements for the Provision of Information to Consumers on the Absence or Reduced Presence of Gluten in Food OJ L 228, 31.7.2014. pp. 5–8. Available online: [https://eur-lex.europa.eu/eli/reg\\_impl/2014/828/oj](https://eur-lex.europa.eu/eli/reg_impl/2014/828/oj) (accessed on 29 December 2020).
32. Food Allergen Labeling and Consumer Protection Act. Sec. 206 Regulations 21 USC 343. Title II of Public Law 108-282, Enacted on August 2, 2004. Available online: <https://www.fda.gov/food/food-allergens/gluten-free-guidance-documents-regulatory-information/food-allergen-labeling-and-consumer-protection-act-2004-falcpa> (accessed on 29 December 2020).
33. Scherf, K.A.; Poms, R.E. Recent developments in analytical methods for tracing gluten. *J. Cereal Sci.* **2016**, *67*, 112–122. [CrossRef]
34. Australia New Zealand Food Standards Code—Standard 1.2.7—Nutrition, Health and Related Claims. F2018C00942. Available online: <https://www.legislation.gov.au/Details/F2018C00942> (accessed on 29 December 2020).
35. MODIFICACIÓN a la Norma Oficial Mexicana NOM-051-SCFI/SSA1-2010, Especificaciones Generales de Etiquetado para Alimentos y Bebidas no Alcohólicas Preenvasados-Información Comercial y Sanitaria, Publicada el 5 de abril de 2010. Diario Oficial de la Federación, 27 de Marzo de 2020. Available online: [http://www.dof.gob.mx/normasOficiales/8150/seeco11\\_C/seeco11\\_C.html](http://www.dof.gob.mx/normasOficiales/8150/seeco11_C/seeco11_C.html) (accessed on 29 December 2020).
36. Norma Oficial Mexicana NOM-247-SSA1-2008, Productos y Servicios. Cereales y sus Productos. Cereales, Harinas de Cereales, Sémolas o Semolinas. Alimentos a base de: Cereales, Semillas Comestibles, de Harinas, Sémolas o Semolinas o sus Mezclas. Productos de Panificación. Disposiciones y Especificaciones Sanitarias y Nutrimientales. Métodos de Prueba. Diario Oficial de la Federación, 27 de julio de 2009. Available online: [http://dof.gob.mx/nota\\_detalle.php?codigo=5188649&fecha=10/05/2011](http://dof.gob.mx/nota_detalle.php?codigo=5188649&fecha=10/05/2011) (accessed on 29 December 2020).
37. Ley 27196. Enfermedad Celíaca Modificación Ley 26588. Boletín Oficial de la República Argentina No. 33259. Available online: <https://www.boletinoficial.gob.ar/detalleAviso/primera/136317/20151118> (accessed on 29 December 2020).
38. Lei No. 10674 de 16 de maio de 2003 Obriga a que os Produtos Alimentícios Comercializados Informem sobre a Presença de Gluten, como Medida Preventiva e de Controle da Doença Celíaca. Diário Oficial da União 19/5/2003. Available online: [https://www.camara.leg.br/proposicoesWeb/prop\\_mostrarintegra?codteor=440852&filename=Legislacao](https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=440852&filename=Legislacao) (accessed on 29 December 2020).
39. The China Food Law Blog. Available online: <https://chinafoodlaw.blog/2018/10/09/gluten-free-in-china/> (accessed on 16 July 2020).
40. Akiyama, H.; Imai, T.; Ebisawa, M. *Japan Food Allergen Labeling Regulation—History and Evaluation*; Elsevier: Amsterdam, The Netherlands, 2011; pp. 139–171. [CrossRef]
41. Panda, R.; Garber, E.A.E. Detection and Quantitation of Gluten in Fermented-Hydrolyzed Foods by Antibody-Based Methods: Challenges, Progress, and a Potential Path Forward. *Front. Nutr.* **2019**, *6*, 97. [CrossRef]
42. Ferre-López, S.; Ribes-koninckx, C.; Genzor, C.; Gamén, S.; Peña, L.; Ortigosa, L.; Méndez, E. Immunochromatographic sticks for tissue transglutaminase and antigliadin antibody screening in celiac disease. *Clin. Gastroenterol. Hepatol.* **2004**, *2*, 480–484. [CrossRef]

43. Panda, R.; Garber, E.A.E. Western blot analysis of fermented-hydrolyzed foods utilizing gluten-specific antibodies employed in a novel multiplex competitive ELISA. *Anal. Bioanal. Chem.* **2019**, *411*, 5159–5174. [[CrossRef](#)] [[PubMed](#)]
44. Li, H.; Bose, U.; Stockwell, S.; Howitt, C.A.; Colgrave, M. Assessing the Utility of Multiplexed Liquid Chromatography-Mass Spectrometry for Gluten Detection in Australian Breakfast Food Products. *Molecules* **2019**, *24*, 3665. [[CrossRef](#)] [[PubMed](#)]
45. Boukid, F.; Prandi, B.; Faccini, A.; Sforza, S. A Complete Mass Spectrometry (MS)-Based Peptidomic Description of Gluten Peptides Generated During In Vitro Gastrointestinal Digestion of Durum Wheat: Implication for Celiac Disease. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1481–1490. [[CrossRef](#)]
46. Martín-Fernández, B.; Costa, J.; Oliveira, M.B.P.P.; López-Ruiz, B.; Mafra, I. Screening new gene markers for gluten detection in foods. *Food Control* **2015**, *56*, 57–63. [[CrossRef](#)]
47. Martín-Fernández, B.; Costa, J.; Oliveira, M.B.P.P.; López-Ruiz, B.; Mafra, I. Combined effects of matrix and gene marker on the real-time PCR detection of wheat. *Int. J. Food Sci. Technol.* **2016**, *51*, 1680–1688. [[CrossRef](#)]
48. García-García, A.; Madrid, R.; Sohrabi, H.; De La Cruz, S.; García, T.; Martín, R.; González, I. A sensitive and specific real-time PCR targeting DNA from wheat, barley and rye to track gluten contamination in marketed foods. *LWT Food Sci. Technol.* **2019**, *114*, 108378. [[CrossRef](#)]
49. Burton, D.R. Structure and function of antibodies. In *New Comprehensive Biochemistry*; Calabi, F., Neuberger, M.S., Eds.; Elsevier: Amsterdam, The Netherlands, 1987; Volume 17, pp. 1–50.
50. Lee, W.; Syed Atif, A.; Tan, S.C.; Leow, C.H. Insights into the chicken IgY with emphasis on the generation and applications of chicken recombinant monoclonal antibodies. *J. Immunol. Methods* **2017**, *447*, 71–85. [[CrossRef](#)]
51. Frenzel, A.; Hust, M.; Schirrmann, T. Expression of Recombinant Antibodies. *Front. Immunol.* **2013**, *4*, 217. [[CrossRef](#)]
52. Muyldermans, S. Nanobodies: Natural Single-Domain Antibodies. *Annu. Rev. Biochem.* **2013**, *82*, 775–797. [[CrossRef](#)]
53. Kohler, G.; Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **1975**, *256*, 495–497. [[CrossRef](#)] [[PubMed](#)]
54. Köhler, G.; Milstein, C. Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. *Eur. J. Immunol.* **1976**, *6*, 511–519. [[CrossRef](#)] [[PubMed](#)]
55. Manohar, R.S. Development of Enzyme-Linked Immunosorbent Assay for Evaluation of Chapati-Making Quality of Wheat Varieties. *J. Agric. Food Chem.* **2002**, *50*, 7455–7460. [[CrossRef](#)] [[PubMed](#)]
56. Gujral, N.; Suh, J.W.; Sunwoo, H.H. Effect of anti-gliadin IgY antibody on epithelial intestinal integrity and inflammatory response induced by gliadin. *BMC Immunol.* **2015**, *16*, 41. [[CrossRef](#)] [[PubMed](#)]
57. Skerritt, J.; Underwood, P.A. Specificity characteristics of monoclonal antibodies to wheat grain storage proteins. *Biochim. Biophys. Acta-Protein Struct. Mol. Enzymol.* **1986**, *874*, 245–254. [[CrossRef](#)]
58. Galfre, G.; Howe, S.C.; Milstein, C.; Butcher, G.W.; Howard, J.C. Antibodies to major histocompatibility antigens produced by hybrid cell lines. *Nature* **1977**, *266*, 550–552. [[CrossRef](#)]
59. Skerritt, J.H.; Hill, A.S. Monoclonal antibody sandwich enzyme immunoassays for determination of gluten in foods. *J. Agric. Food Chem.* **1990**, *38*, 1771–1778. [[CrossRef](#)]
60. Sorell, L.; López, J.A.; Valdés, I.; Alfonso, P.; Camafeita, E.; Acevedo, B.; Chirido, F.; Gavilondo, J.; Méndez, E. An innovative sandwich ELISA system based on an antibody cocktail for gluten analysis. *Fed. Eur. Biochem. Soc. Lett.* **1998**, *439*, 46–50. [[CrossRef](#)]
61. Valdés, I.; García, E.; Llorente, M.; Méndez, E. Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *Eur. J. Gastroenterol. Hepatol.* **2003**, *15*, 465–474. [[CrossRef](#)]
62. Garcia, E.; Llorente, M.; Hernando, A.; Kieffer, R.; Wieser, H.; Mendez, E. Development of a general procedure for complete extraction of gliadins for heat processed and unheated foods. *Eur. J. Gastroenterol. Hepatol.* **2005**, *17*, 529–539. [[CrossRef](#)]
63. Köhler, P.; Schwalb, T.; Immer, U.; Lacorn, M.; Wehling, P.; Don, C. AACCI Approved Methods Technical Committee Report: Collaborative Study on the Immunochemical Determination of Partially Hydrolyzed Gluten Using an R5 Competitive ELISA. *Cereal Foods World* **2013**, *58*, 113–117. [[CrossRef](#)]
64. Méndez, E.; Vela, C.; Immer, U.; Janssen, F.W. Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *Eur. J. Gastroenterol. Hepatol.* **2005**, *17*, 1053–1063. [[CrossRef](#)] [[PubMed](#)]
65. Koerner, T.B.; Abbott, M.; Godefroy, S.B.; Popping, B.; Yeung, J.M.; Diaz-Amigo, C.; Roberts, J.; Taylor, S.L.; Baumert, J.L.; Ulberth, F.; et al. Validation Procedures for Quantitative Gluten ELISA Methods: AOAC Allergen Community Guidance and Best Practices. *J. AOAC Int.* **2013**, *96*, 1033–1040. [[CrossRef](#)] [[PubMed](#)]
66. Morón, B.; Cebolla, Á.; Manyani, H.; Álvarez-Maqueda, M.; Megías, M.; Thomas, M.D.C.; López, M.C.; Sousa, C. Sensitive detection of cereal fractions that are toxic to celiac disease patients by using monoclonal antibodies to a main immunogenic wheat peptide. *Am. J. Clin. Nutr.* **2008**, *87*, 405–414. [[CrossRef](#)]
67. Shan, L.; Molberg, Ø.; Parrot, I.; Hausch, F.; Filiz, F.; Gray, G.; Sollid, L.; Khosla, C. Structural Basis for Gluten Intolerance in Celiac Sprue. *Science* **2002**, *297*, 2275–2279. [[CrossRef](#)]
68. Morón, B.; Bethune, M.T.; Comino, I.; Manyani, H.; Ferragud, M.; López, M.C.; Cebolla, Á.; Khosla, C.; Sousa, C. Toward the Assessment of Food Toxicity for Celiac Patients: Characterization of Monoclonal Antibodies to a Main Immunogenic Gluten Peptide. *PLoS ONE* **2008**, *3*, e2294. [[CrossRef](#)]
69. Cereals & Grains Association. Available online: <http://methods.aacnet.org/summaries/38-52-01.aspx> (accessed on 18 October 2020).

70. Official Methods of Analysis AOAC International. Available online: <http://www.eoma.aoac.org/methods/info.asp?ID=50740> (accessed on 17 October 2020).
71. Anderson, R.P.; Degano, P.; Godkin, A.J.; Jewell, D.P.; Hill, A.V.S. In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat. Med.* **2000**, *6*, 337–342. [[CrossRef](#)]
72. Arentz-Hansen, H.; Körner, R.; Molberg, Ø.; Quarsten, H.; Vader, W.; Kooy, Y.M.C.; Lundin, K.E.A.; Koning, F.; Roepstorff, P.; Sollid, L.M.; et al. The Intestinal T Cell Response to  $\alpha$ -Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase. *J. Exp. Med.* **2000**, *191*, 603–612. [[CrossRef](#)]
73. Stern, M.; Ciclitira, P.J.; van Eckert, R.; Feighery, C.; Janssen, F.W.; Méndez, E.; Mothes, T.; Troncone, R.; Wieser, H. Analysis and clinical effects of gluten in coeliac disease. *Eur. J. Gastroenterol. Hepatol.* **2001**, *13*, 741–747. [[CrossRef](#)]
74. Mitea, C.; Kooy-Winkelaar, Y.; Van Veelen, P.; De Ru, A.; Drijfhout, J.W.; Koning, F.; Dekking, L. Fine specificity of monoclonal antibodies against celiac disease-inducing peptides in the gluteome. *Am. J. Clin. Nutr.* **2008**, *88*, 1057–1066. [[CrossRef](#)]
75. Mujico, J.R.; Dekking, L.; Kooy-Winkelaar, Y.; Verheijen, R.; Wichen, P.V.; Streppel, L.; Sajic, N.; Drijfhout, J.W.; Koning, F. Validation of a New Enzyme-Linked Immunosorbent Assay to Detect the Triggering Proteins and Peptides for Celiac Disease: Interlaboratory Study. *J. AOAC Int.* **2019**, *95*, 206–215. [[CrossRef](#)] [[PubMed](#)]
76. Tranquet, O.; Larré, C.; Denery-Papini, S. Allergic reactions to hydrolysed wheat proteins: Clinical aspects and molecular structures of the allergens involved. *Crit. Rev. Food Sci. Nutr.* **2020**, *60*, 147–156. [[CrossRef](#)] [[PubMed](#)]
77. Tranquet, O.; Lupi, R.; Echasserieau-Laporte, V.; Pietri, M.; Larré, C.; Denery-Papini, S. Characterization of Antibodies and Development of an Indirect Competitive Immunoassay for Detection of Deamidated Gluten. *J. Agric. Food Chem.* **2015**, *63*, 5403–5409. [[CrossRef](#)]
78. Tranquet, O.; Gaudin, J.-C.; Patil, S.; Steinbrecher, J.; Matsunaga, K.; Teshima, R.; Sakai, S.; Larré, C.; Denery-Papini, S. A chimeric IgE that mimics IgE from patients allergic to acid-hydrolyzed wheat proteins is a novel tool for in vitro allergenicity assessment of functionalized glutens. *PLoS ONE* **2017**, *12*, e0187415. [[CrossRef](#)]
79. Winter, G. Harnessing Evolution to Make Medicines (Nobel Lecture). *Angew. Chem. Int. Ed.* **2019**, *58*, 14438–14445. [[CrossRef](#)]
80. Meigs, L. Animal testing and its alternatives—the most important omics is economics. *Altern. Anim. Exp.* **2018**, *35*, 275–305. [[CrossRef](#)]
81. Smith, G.P.; Petrenko, V.A. Phage Display. *Chem. Rev.* **1997**, *97*, 391–410. [[CrossRef](#)] [[PubMed](#)]
82. Arnold, F.H. Innovation by Evolution: Bringing New Chemistry to Life (Nobel Lecture). *Angew. Chem. Int. Ed.* **2019**, *58*, 14420–14426. [[CrossRef](#)] [[PubMed](#)]
83. Joyce, C.; Burton, D.R.; Briney, B. Comparisons of the antibody repertoires of a humanized rodent and humans by high throughput sequencing. *Sci. Rep.* **2020**, *10*, 1120. [[CrossRef](#)]
84. Hozumi, N.; Tonegawa, S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 3628–3632. [[CrossRef](#)]
85. Hust, M. Mating antibody phage display with proteomics. *Trends Biotechnol.* **2004**, *22*, 8–14. [[CrossRef](#)] [[PubMed](#)]
86. Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G. Making antibody fragments using phage display libraries. *Nature* **1991**, *352*, 624–628. [[CrossRef](#)] [[PubMed](#)]
87. De Haard, H.J.; Van Neer, N.; Reurs, A.; Hufton, S.E.; Roovers, R.C.; Henderikx, P.; De Bruijne, A.P.; Arends, J.-W.; Hoogenboom, H.R. A Large Non-immunized Human Fab Fragment Phage Library That Permits Rapid Isolation and Kinetic Analysis of High Affinity Antibodies. *J. Biol. Chem.* **1999**, *274*, 18218–18230. [[CrossRef](#)] [[PubMed](#)]
88. Pini, A.; Viti, F.; Santucci, A.; Carnemolla, B.; Zardi, L.; Neri, P.; Neri, D. Design and Use of a Phage Display Library. *J. Biol. Chem.* **1998**, *273*, 21769–21776. [[CrossRef](#)]
89. Wagner, H.; Wehrle, S.; Weiss, E.; Cavallari, M.; Weber, W. A Two-Step Approach for the Design and Generation of Nanobodies. *Int. J. Mol. Sci.* **2018**, *19*, 3444. [[CrossRef](#)]
90. Böldicke, T. Single domain antibodies for the knockdown of cytosolic and nuclear proteins. *Protein Sci.* **2017**, *26*, 925–945. [[CrossRef](#)]
91. Hoogenboom, H. Designing and optimizing library selection strategies for generating high-affinity antibodies. *Trends Biotechnol.* **1997**, *15*, 62–70. [[CrossRef](#)]
92. Leow, C.H.; Fischer, K.; Leow, C.Y.; Cheng, Q.; Chuah, C.; McCarthy, J. Single Domain Antibodies as New Biomarker Detectors. *Diagnostics* **2017**, *7*, 52. [[CrossRef](#)]
93. Smith, G.P. Phage Display: Simple Evolution in a Petri Dish (Nobel Lecture). *Angew. Chem. Int. Ed.* **2019**, *58*, 14428–14437. [[CrossRef](#)]
94. Plückthun, A. *Ribosome Display: A Perspective*; Springer: New York, NY, USA, 2012; pp. 3–28. [[CrossRef](#)]
95. Yamaguchi, J.; Naimuddin, M.; Biyani, M.; Sasaki, T.; Machida, M.; Kubo, T.; Funatsu, T.; Husimi, Y.; Nemoto, N. cDNA display: A novel screening method for functional disulfide-rich peptides by solid-phase synthesis and stabilization of mRNA-protein fusions. *Nucleic Acids Res.* **2009**, *37*, e108. [[CrossRef](#)]
96. Lee, S.Y.; Choi, J.H.; Xu, Z. Microbial cell-surface display. *Trends Biotechnol.* **2003**, *21*, 45–52. [[CrossRef](#)]
97. Smith, G.P. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science* **1985**, *228*, 1315–1317. [[CrossRef](#)] [[PubMed](#)]
98. Parmley, S.F.; Smith, G.P. Antibody-selectable filamentous fd phage vectors: Affinity purification of target genes. *Gene* **1988**, *73*, 305–318. [[CrossRef](#)]

99. De la Cruz, V.F.; Lal, A.A.; McCutchan, T.F. Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. *J. Biol. Chem.* **1988**, *263*, 4318–4322.
100. Smith, G.P. Surface display and peptide libraries. *Gene* **1993**, *128*, 1–2. [[CrossRef](#)]
101. Armstrong, N.; Adey, N.B.; McConnell, S.J.; Kay, B.K. CHAPTER 3-Vectors for Phage Display. In *Phage Display of Peptides and Proteins*; Kay, B.K., Winter, J., McCafferty, J., Eds.; Academic Press: Burlington, MA, USA, 1996; pp. 35–53. [[CrossRef](#)]
102. Makowski, L. Structural constraints on the display of foreign peptides on filamentous bacteriophages. *Gene* **1993**, *128*, 5–11. [[CrossRef](#)]
103. Lipovsek, D.; Plückthun, A. In-vitro protein evolution by ribosome display and mRNA display. *J. Immunol. Methods* **2004**, *290*, 51–67. [[CrossRef](#)]
104. Hoischen, C.; Fritsche, C.; Gumpert, J.; Westermann, M.; Gura, K.; Fahnert, B. Novel Bacterial Membrane Surface Display System Using Cell Wall-Less L-Forms of *Proteus mirabilis* and *Escherichia coli*. *Appl. Environ. Microbiol.* **2002**, *68*, 525–531. [[CrossRef](#)]
105. Mesnage, S.P.; Weber-Levy, M.; Haustant, M.; Mock, M.L.; Fouet, A.S. Cell Surface-Exposed Tetanus Toxin Fragment C Produced by Recombinant *Bacillus anthracis* Protects against Tetanus Toxin. *Infect. Immun.* **1999**, *67*, 4847–4850. [[CrossRef](#)]
106. Chao, G.; Lau, W.L.; Hackel, B.J.; Sazinsky, S.L.; Lippow, S.M.; Wittrup, K.D. Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **2006**, *1*, 755–768. [[CrossRef](#)]
107. Bowers, P.M.; Horlick, R.A.; Kehry, M.R.; Neben, T.Y.; Tomlinson, G.L.; Altobelli, L.; Zhang, X.; Macomber, J.L.; Krapf, I.P.; Wu, B.F.; et al. Mammalian cell display for the discovery and optimization of antibody therapeutics. *Methods* **2014**, *65*, 44–56. [[CrossRef](#)] [[PubMed](#)]
108. Dieterich, W.; Ehnis, T.; Bauer, M.; Donner, P.; Volta, U.; Riecken, E.O.; Schuppan, D. Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat. Med.* **1997**, *3*, 797–801. [[CrossRef](#)] [[PubMed](#)]
109. Sblattero, D.; Florian, F.; Not, T.; Ventura, A.; Bradbury, A.; Marzari, R. Analyzing the peripheral blood antibody repertoire of a celiac disease patient using phage antibody libraries. *Hum. Antibodies* **2000**, *9*, 199–205. [[CrossRef](#)] [[PubMed](#)]
110. Marzari, R.; Sblattero, D.; Florian, F.; Tongiorgi, E.; Not, T.; Tommasini, A.; Ventura, A.; Bradbury, A. Molecular Dissection of the Tissue Transglutaminase Autoantibody Response in Celiac Disease. *J. Immunol.* **2001**, *166*, 4170–4176. [[CrossRef](#)]
111. Sblattero, D.; Florian, F.; Azzoni, E.; Ziberna, F.; Tommasini, A.; Not, T.; Ventura, A.; Bradbury, A.; Marzari, R. One-step cloning of anti tissue transglutaminase scFv from subjects with celiac disease. *J. Autoimmun.* **2004**, *22*, 65–72. [[CrossRef](#)]
112. Not, T.; Ziberna, F.; Vatta, S.; Quaglia, S.; Martellosi, S.; Villanacci, V.; Marzari, R.; Florian, F.; Vecchiet, M.; Sulic, A.-M.; et al. Cryptic genetic gluten intolerance revealed by intestinal antitransglutaminase antibodies and response to gluten-free diet. *Gut* **2011**, *60*, 1487–1493. [[CrossRef](#)]
113. Høydahl, L.S.; Richter, L.; Frick, R.; Snir, O.; Gunnarsen, K.S.; Landsverk, O.J.B.; Iversen, R.; Jeliakov, J.R.; Gray, J.J.; Bergseng, E.; et al. Plasma Cells Are the Most Abundant Gluten Peptide MHC-expressing Cells in Inflamed Intestinal Tissues From Patients With Celiac Disease. *Gastroenterology* **2019**, *156*, 1428–1439.e10. [[CrossRef](#)]
114. Rhyner, C.; Weichel, M.; Hubner, P.; Achatz, G.; Blaser, K.; Cramer, R. Phage display of human antibodies from a patient suffering from coeliac disease and selection of isotype-specific scFv against gliadin. *Immunology* **2003**, *110*, 269–274. [[CrossRef](#)]
115. Chen, T.; Hoffmann, K.; Östman, S.; Sandberg, A.-S.; Olsson, O. Identification of gliadin-binding peptides by phage display. *BMC Biotechnol.* **2011**, *11*, 16. [[CrossRef](#)]
116. Doña, V.; Urrutia, M.; Bayardo, M.; Alzogaray, V.; Goldbaum, F.A.; Chirido, F.G. Single Domain Antibodies Are Specially Suited for Quantitative Determination of Gliadins under Denaturing Conditions. *J. Agric. Food Chem.* **2010**, *58*, 918–926. [[CrossRef](#)]
117. García-García, A.; Madrid, R.; González, I.; García, T.; Martín, R. A novel approach to produce phage single domain antibody fragments for the detection of gluten in foods. *Food Chem.* **2020**, *321*, 126685. [[CrossRef](#)] [[PubMed](#)]
118. Jayathilake, C.; Kumachi, S.; Arai, H.; Motohashi, M.; Terai, T.; Murakami, A.; Nemoto, N. In vitro selection of anti-gliadin single-domain antibodies from a naïve library for cDNA-display mediated immuno-PCR. *Anal. Biochem.* **2020**, *589*, 113490. [[CrossRef](#)] [[PubMed](#)]
119. Lee, C.M.Y.; Iorno, N.; Sierro, F.; Christ, D. Selection of human antibody fragments by phage display. *Nat. Protoc.* **2007**, *2*, 3001–3008. [[CrossRef](#)] [[PubMed](#)]
120. Sharma, G.M.; Rallabhandi, P.; Williams, K.M.; Pahlavan, A. Characterization of Antibodies for Grain-Specific Gluten Detection. *J. Food Sci.* **2016**, *81*, T810–T816. [[CrossRef](#)] [[PubMed](#)]
121. Anzai, H.; Terai, T.; Jayathilake, C.; Suzuki, T.; Nemoto, N. A novel immuno-PCR method using cDNA display. *Anal. Biochem.* **2019**, *578*, 1–6. [[CrossRef](#)] [[PubMed](#)]



