

MID-INFRARED SPECTROSCOPY OF LENSED GALAXIES AT $1 < z < 3$: THE NATURE OF SOURCES NEAR THE MIPS CONFUSION LIMIT

J. R. RIGBY,^{1,2,3} D. MARCILLAC,¹ E. EGAMI,¹ G. H. RIEKE,¹ J. RICHARD,⁴ J.-P. KNEIB,⁵
 D. FADDA,⁶ C. N. A. WILLMER,¹ C. BORYS,⁶ P. P. VAN DER WERF,⁶
 P. G. PÉREZ-GONZÁLEZ,⁷ K. K. KNUDSEN,⁸ AND C. PAPOVICH¹

Received 2007 July 9; accepted 2007 October 27

ABSTRACT

We present *Spitzer* IRS mid-infrared spectra for 15 gravitationally lensed, $24\ \mu\text{m}$ –selected galaxies, and combine the results with four additional very faint galaxies with IRS spectra in the literature. The median intrinsic $24\ \mu\text{m}$ flux density of the sample is $130\ \mu\text{Jy}$, enabling a systematic survey of the spectral properties of the very faint $24\ \mu\text{m}$ sources that dominate the number counts of *Spitzer* cosmological surveys. Six of the 19 galaxy spectra (32%) show the strong mid-IR continua expected of AGNs; X-ray detections confirm the presence of AGNs in three of these cases, and reveal AGNs in two other galaxies. These results suggest that nuclear accretion may contribute more flux to faint $24\ \mu\text{m}$ –selected samples than previously assumed. Almost all the spectra show some aromatic (PAH) emission features; the measured aromatic flux ratios do not show evolution from $z = 0$. In particular, the high signal-to-noise mid-IR spectrum of SMM J163554.2+661225 agrees remarkably well with low-redshift, lower luminosity templates. We compare the rest-frame $8\ \mu\text{m}$ and total infrared luminosities of star-forming galaxies, and find that the behavior of this ratio with total IR luminosity has evolved modestly from $z = 2$ to $z = 0$. Since the high aromatic-to-continuum flux ratios in these galaxies rule out a dominant contribution by AGNs, this finding implies systematic evolution in the structure and/or metallicity of infrared sources with redshift. It also has implications for the estimates of star-forming rates inferred from $24\ \mu\text{m}$ measurements, in the sense that at $z \sim 2$, a given observed frame $24\ \mu\text{m}$ luminosity corresponds to a lower bolometric luminosity than would be inferred from low-redshift templates of similar luminosity at the corresponding rest wavelength.

Subject headings: galaxies: evolution — galaxies: high-redshift — infrared: galaxies

1. INTRODUCTION

The *Spitzer Space Telescope* (Werner et al. 2004) has been tremendously successful at detecting star-forming galaxies and active galactic nuclei (AGNs) at $z > 1$ by their emission in the observed-frame $24\ \mu\text{m}$ band. Down to the confusion limit of $\sim 50\ \mu\text{Jy}$ at $24\ \mu\text{m}$ (Dole et al. 2004), the MIPS instrument (Rieke et al. 2004) can detect luminous infrared galaxies (LIRGs)⁹ out to $z \sim 2$, ultraluminous infrared galaxies (ULIRGs)¹⁰ out to $z \sim 3$, and hyperluminous infrared galaxies (HLIRGs)¹¹ out to even higher redshifts.

Detectability at $z > 0.7$ is increased by strong aromatic emission features, at rest-frame $6\text{--}12\ \mu\text{m}$, that pass through the $24\ \mu\text{m}$ band. Indeed, photometric redshifts place $\sim 25\%\text{--}30\%$ of the faint $24\ \mu\text{m}$ sources at redshifts above 1.4 (Le Floc'h et al.

2005; Pérez-González et al. 2005; Caputi et al. 2006; Wang 2006).

Spectra have been obtained with the Infrared Spectrograph (IRS; Houck et al. 2004) for extremely high-luminosity galaxies, (e.g., Houck et al. 2005; Yan et al. 2007; Menéndez-Delmestre et al. 2007). However, since spectral properties depend strongly on luminosity, results obtained for hyperluminous galaxies may have limited applicability for the bulk of the IR-detected population. Spectra of lower luminosity galaxies have been obtained only in a few cases (e.g., Lutz et al. 2005; Teplitz et al. 2007).

Numerous authors (e.g., Caputi et al. 2006; Choi et al. 2006; Reddy et al. 2006) have used the observed $24\ \mu\text{m}$ band as an estimator of star formation rates. Since most of the infrared power is radiated redward of the observed band, the $24\ \mu\text{m}$ diagnostic must be calibrated using low-redshift templates (e.g., Chary & Elbaz 2001; Dale & Helou 2002; Lagache et al. 2004; Brandl et al. 2006; Armus et al. 2007). The central assumption in these works is that the spectra and spectral energy distributions of high- z galaxies are matched accurately by the low-redshift templates. For $0.7 < z < 3$, the $24\ \mu\text{m}$ band fluxes are strongly influenced by aromatic band emission. Evolution over time either in the behavior of these aromatic bands, in metallicity, or in the geometry and radiative transfer within these infrared sources, could all undermine this central assumption.

To probe these possibilities, we have been obtaining IRS spectra for intrinsically faint $24\ \mu\text{m}$ –selected galaxies at $1 \lesssim z \lesssim 3$. Our targets are strongly lensed by the gravitational potential of foreground clusters of galaxies, such that their fluxes are amplified by factors of 3–25. Although these galaxies have observed $24\ \mu\text{m}$ flux densities of $\sim 1\ \text{mJy}$, sufficient for useful, efficient spectroscopy with IRS, their intrinsic luminosities are down in the LIRG to ULIRG range.

¹ Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721.

² Current address: Observatories, Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101; jrigby@ociw.edu.

³ Spitzer Fellow.

⁴ Department of Astronomy, Caltech, 1200 East California Boulevard, Pasadena, CA 91125.

⁵ Laboratoire d'Astrophysique de Marseille.

⁶ Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, Netherlands.

⁷ Departamento de Astrofísica y CC. de la Atmósfera, Facultad de CC. Físicas, Universidad Complutense de Madrid, 28040 Madrid, Spain.

⁸ Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany.

⁹ Defined as $11 < \log L(\text{TIR})/L_{\odot} < 12$, where $L(\text{TIR})$ is the total infrared luminosity between 8 and $1000\ \mu\text{m}$.

¹⁰ Defined as $12 < \log L(\text{TIR}) < 13\ L_{\odot}$.

¹¹ Defined as $\log L(\text{TIR}) > 13\ L_{\odot}$.

TABLE 1
THE IRS LENSED SAMPLE

Source (1)	R.A. (J2000) (2)	Decl. (J2000) (3)	Submillimeter Name (4)	Reference (5)	PID (6)	z (7)	z Reference (8)	Amplitude (9)
Our Sample								
MS0451a.....	04 54 07.12	−03 00 36.4			1	1.95	1	2.5
A851a.....	09 42 54.55	+46 58 44.3	SMM J09429+4658	A	2	2.38	1	1.3, 2.0
A1689a.....	13 11 28.04	−01 19 18.7			2	1.1537	2	8.5 ± 2.3
A1689b.....	13 11 29.14	−01 20 46.5	SMM J13115−01208	B	2	2.63	2	25–47
A1835a.....	14 01 04.96	+02 52 24.8	SMM J140105+025223.5	C	12	2.565	3	3.5 ± 0.5 ^a
A2218a.....	16 35 54.18	+66 12 24.8	SMM J163554.2+661225	D	1	2.516	4	22 ± 2
A2218b.....	16 35 55.16	+66 11 50.8	SMM J163555.2+661150	E	2	1.034	5	6.1
A2218c.....	16 35 59.12	+66 11 47.4			2	0.97	1	6.7
A2219a.....	16 40 19.50	+46 44 00.5	SMM J16403+46440	F	1	2.03	1	3.6
A2261a.....	17 22 21.00	+32 07 05.5	SMM J17223+3207	F	12	1 < z < 3	1	3.3
A2390a.....	21 53 33.24	+17 42 10.6			12	2.858	6	10.8
A2390b.....	21 53 34.43	+17 42 21.6			12	0.913	7	3.35
A2390c.....	21 53 34.540	+17 42 03.0	“Source 14”	G	2	0.91	8	10.4
AC 114a.....	22 58 49.79	−34 48 47.0			2	1.47	1	9.7
A2667a.....	23 51 40.00	−26 04 52.0			2	1.034	9	17
Additional Sources from the Literature								
Teplitz-1	3 32 44.00	−27 46 35.0			8	2.69	10	1.0
Teplitz-1-BzK.....	3 32 38.52	−27 46 33.5			8	2.55	10	1.0
Teplitz-2	3 32 34.85	−27 46 40.4			8	1.09	10	1.0
A370a.....	2 39 51.87	−1 35 58.78	SMM J02399-0136	A	9	2.81	11	2.45

NOTES.—Col. (1): Source name.

Cols. (2)–(3): Coordinates: units of right ascension are hours, minutes, and seconds, and units of declination are degrees, arcminutes, and arcseconds.

Col. (4): Submillimeter name, if source has a submillimeter detection in the literature.

Col. (5): References for detected submillimeter counterpart: (A) Smail et al. 2002; (B) Knudsen 2004 and 2008; (C) Smail et al. 2000 and Ivison et al. 2000; (D) Kneib et al. 2004; (E) Knudsen et al. 2006; (F) Chapman et al. 2002; (G) Cowie et al. 2002.

Col. (6): *Spitzer* program in which IRS spectra were obtained: (1) PID 82; (2) PID 30775; (8) PID 252; (9) PID 3241.

Cols. (7)–(8): Redshift and redshift reference: (1) redshift determined from our IRS spectra; (2) redshift based on Keck spectra (J. Richard et al. in prep.); (3) from Frayer et al. 1999; (4) Kneib et al. 2004; (5) Ebbels et al. 1998; (6) Keck spectra presented in the Appendix; (7) Pello et al. 1991; (8) Frye & Broadhurst 1998; (9) Sand et al. 2005; (10) Teplitz et al. 2007; (11) Frayer et al. 1998 and Lutz et al. 2005.

Col. (9): Amplification estimated from the lensing models. Values from the literature are footnoted; un-footnoted values are new.

^a The amplification for this source has been the subject of several papers; Downes & Solomon (2003) and Motohara et al. (2005) argue that an intervening cluster galaxy boosts the amplification to ~ 25 . Smail et al. (2005) and Smith et al. (2005) find that the intervening galaxy is a dwarf with small velocity dispersion and thus contributes only modestly to the total amplification, which they estimate as 3.5 ± 0.5 . We note the controversy and adopt the Smail et al. (2005) amplification.

In this paper we present IRS spectra for the first 15 galaxies in the sample. The median observed flux density is $f_{\nu}(24 \mu\text{m}) = 0.82 \text{ mJy}$, but the median intrinsic flux density is nearly an order of magnitude lower. We compare their aromatic feature flux ratios, and the ratios of their X-ray, aromatic, and total infrared luminosities to those of low-redshift LIRG and ULIRG samples. We examine in detail the high-quality spectrum of SMM J163554.2+661225, a LIRG at $z = 2.516$. We merge our sample with four additional objects from the literature (Lutz et al. 2005; Teplitz et al. 2007) to provide a first systematic look at the IRS-measured properties of the faint $24 \mu\text{m}$ galaxy population.

We assume $\Omega_m = 0.265$, $\Omega_\Lambda = 0.735$, and $H_0 = 72 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (Spergel et al. 2003; Freedman et al. 2001).

2. SAMPLE SELECTION

Our goal was to select sources with intrinsic flux densities close to, or even below, the MIPS $24 \mu\text{m}$ confusion limit. We first obtained deep $24 \mu\text{m}$ images of galaxy clusters, generally identified as being massive through high X-ray luminosity. From these $24 \mu\text{m}$ images, we selected sources within $1.5'$ of the cluster center that met the following criteria:

1. Had observed $24 \mu\text{m}$ flux densities above 0.4 mJy (to make efficient use of IRS).

2. Had faint, irregular, or arclike *HST* counterparts. This excluded galaxies at the cluster redshift.¹²

3. Had known spectroscopic or probable photometric redshifts ≥ 1.0 .

4. Had lensing amplification factors calculated to be greater than 3. This restricted selection to clusters with good lensing models. Since amplification depends on the source redshift, if the redshift was unknown, we calculated the amplification for $1 < z < 3$ and required amplifications factors greater than 3 over that range.

Since submillimeter sources in lensing clusters have been studied extensively, with published amplifications and redshifts, we prioritized sources in the target list with published submillimeter detections. Our sample contains 15 galaxies, targeted in GTO campaigns 82 and 30775. Eight of these galaxies have submillimeter detections in the literature, as referenced in Table 1. Figure 1 shows optical-band images of the sample galaxies.

We supplement the sample with four sources from the literature. We use three of the four sources from Teplitz et al. (2007)

¹² A2219a and A2261a lack *HST* imagery; we instead used ground-based F606W images from the Steward Observatory 90 inch telescope. Both galaxies are known submillimeter sources (see Table 1), and can thus be added to the sample despite the lack of high-resolution optical images.

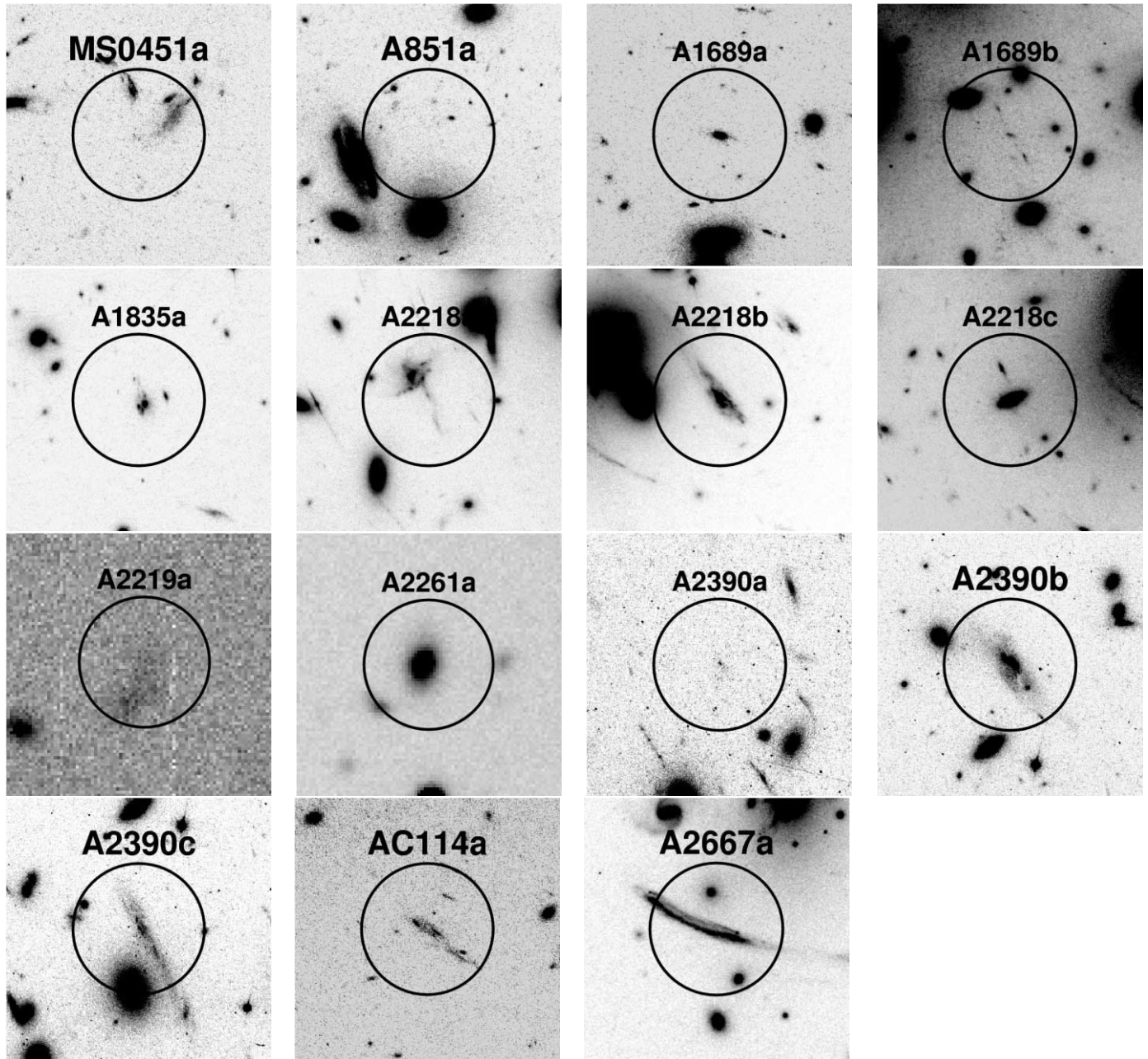


FIG. 1.—Postage stamps of the lensed galaxies. Images are from the *Hubble Space Telescope* Advanced Camera for Surveys, filter F850LP, with the following exceptions: the MS 0451a image is *HST* ACS F814W; the A2667a and A851a images are *HST* WFPC2 F814W; and for A2261a and A2219a, the images are from the Steward Observatory 90 inch telescope, F606W. Each postage stamp is $12''$ by $12''$. Circles of $R = 6''$ are drawn to illustrate the $24\ \mu\text{m}$ *Spitzer* beam out to the first Airy minimum.

with $f_\nu(24\ \mu\text{m}) = 0.13\text{--}0.2\ \text{mJy}$; these are unlensed galaxies at $1 < z < 3$ with very long (9–12 hr) IRS exposures, and are not known submillimeter sources. We also use the lensed $z = 2.81$ submillimeter source in A370 from Lutz et al. (2005)¹³ with $f_\nu(24\ \mu\text{m}) = 1.36\ \text{mJy}$.

Source coordinates, redshifts, amplifications, and flux densities of our sample and for the four literature sources are all listed in Tables 1 and 2.

3. OBSERVATIONS AND DATA REDUCTION

This paper is based on a combination of imaging with MIPS, spectroscopy with IRS, archival imaging from *Chandra*, and

submillimeter photometry from the literature. Exposure times are summarized in Table 3.

3.1. *Spitzer* IRS Spectroscopy

Low-resolution IRS spectra were obtained as part of *Spitzer* GTO programs 82 and 30775 (PI G. Rieke). Data reduction used a package developed by D. Fadda, which has already been successfully used to reduce low-resolution IRS spectra of faint high- z sources (Yan et al. 2007). Residual background, rogue pixels¹⁴ (pixels with dark current values abnormally high and variable with time), and cosmic-ray hits have to be corrected to obtain an optimal reduction of the data. The package considers all the bidimensional frames produced by the IRS/SSC pipeline. First,

¹³ The two other sources from those papers, “2-x” from Teplitz et al. (2007) and the Abell 2125 source from Lutz et al. (2005) have insufficient signal-to-noise for our analysis.

¹⁴ See <http://ssc.spitzer.caltech.edu/irs/roguepixels>.

TABLE 2
LONG-WAVELENGTH PHOTOMETRY AND TEMPLATE FITS

Source (1)	$f(3.5 \mu\text{m})$ (μJy) (2)	$f(4.5 \mu\text{m})$ (μJy) (3)	$f(5.7 \mu\text{m})$ (μJy) (4)	$f(8.0 \mu\text{m})$ (μJy) (5)	$f(24 \mu\text{m})$ (mJy) (6)	(PAH/tot) (7)	$f(70 \mu\text{m})$ (mJy) (8)	$L(\text{TIR})$ (L_{\odot}) (9)	Template (10)
MS0451a.....	32 ± 2	49 ± 2	52 ± 3	36 ± 3	1.32	6.3E-11	5.7 ± 1.6	< 2.2 × 10 ¹³	CE ^a
A851a.....	29 ± 3	44 ± 3	58 ± 3	45 ± 4	0.63	8.4E-11	...	2.7 (2.3–3.6) × 10 ^{13b}	DH
A1689a.....	36 ± 2	35 ± 2	28 ± 2	66 ± 4	0.58	ND	...	ND	
A1689b.....	CR	CR	CR	CR	0.32	2.4E-10	...	1.6 (0.43–2.4) × 10 ¹¹	DH
A1835a.....	88 ± 2	105 ± 2	136 ± 3	124 ± 5	0.99	1.4E-10	<13	6.8 (4.2–11) × 10 ¹²	DH
A2218a.....	67 ± 2	70 ± 2	92 ± 6	105 ± 10	1.16	2.0E-10	<7	7.6 (5.7–9.5) × 10 ¹¹	DH
A2218b.....	182 ± 4	147 ± 3	107 ± 3	107 ± 5	1.67	ND	7.4 ± 1.5	4.8 (2.8–5.3) × 10 ¹¹	DH
A2218c.....	99 ± 3	79 ± 3	67 ± 2	523 ± 11	0.50	ND	10 ± 1	1.3 (0.88–2.0) × 10 ¹¹	Mrk 231
A2219a.....	CR	CR	CR	CR	0.82	7.1E-11	...	2.9 (2.7–3.4) × 10 ¹²	DH
A2261a.....	142 ± 4	138 ± 4	117 ± 3	156 ± 6	0.58	ND	<3.2	^c	
A2390a.....	39 ± 3	59 ± 3	90 ± 3	196 ± 7	0.83	5.8E-11	<7.6	...	
A2390b.....	462 ± 8	334 ± 6	258 ± 5	199 ± 7	0.88	ND	5.3 ± 1.6	1.7 (1.1–2.5) × 10 ¹¹	Mrk 231
A2390c.....	CR	CR	CR	CR	0.60	ND	<8	8.2 (5.5–11) × 10 ¹⁰	DH
AC114a.....	91 ± 3	97 ± 3	78 ± 3	76 ± 4	0.41	1.4E-10	
A2667a.....	466 ± 6	349 ± 5	269 ± 4	320 ± 8	1.52	ND	
Literature Sample									
Teplitz-1.....					0.13	...			
Teplitz-1-BzK.....					0.20	...			
Teplitz-2.....					0.13	...			
A370a.....					1.36	...			

NOTES.—“ND” indicates a nondetection; ellipses indicate no data. Col. (1): Source name. Cols. (2)–(5): IRAC photometry, in μJy . Formal errors from the photometry are quoted, although field crowding probably and aperture corrections probably introduce additional 5%–10% errors. The photometry for A2218a is superior, since it is multiply imaged (with different crowding). “CR” indicates that crowding was too severe to quote photometry. Col. (6): Observed 24 μm flux density, in mJy, from DAOPHOT PSF fitting. Error bars from DAOPHOT are overly optimistic and not quoted; we use error bars of 0.1 mJy for template fitting. Col. (7): Aromatic–to–mid-IR flux ratio, as defined in § 4.2, for the $z > 1.4$ sample. Col. (8): Observed 70 μm flux density, in mJy, from aperture photometry. Error bars are dominated by the sky. Col. (9): Total infrared luminosity, our best–fit and the range of acceptable templates, in L_{\odot} . Col. (10): Best-fitting template: Chary & Elbaz (2001) or Dale & Helou (2002).

^a Mrk 231 template may also fit.

^b Using amplification factor of 1.3.

^c $L(\text{TIR})$ cannot be usefully constrained because the redshift is unknown.

TABLE 3
EXPOSURE TIMES

Cluster	$t(24 \mu\text{m})$ (ks)	$t(3.6 \mu\text{m})$ (ks)	$t(\text{IRS, LL})$ (ks)	$t(\text{Chandra})$ (ks)
MS 0451.....	3.68	2.40	3.66 ^b	53
A851.....	2.78	2.40	6.34 ^a	0
A1689.....	2.78	2.40		40
a.....			7.31 ^a	
b.....			14.63 ^b	
A1835.....	2.77	6.00	3.66 ^b ; 3.66 ^a	29
A2218.....	2.78	2.40		58
a.....			3.66 ^b	
b.....			3.66 ^a	
c.....			7.31 ^a	
A2219.....	2.78	2.40	3.66 ^b	41
A2261.....	2.77	2.40	5.85 ^b ; 7.31 ^a	33
A2390.....	2.77	2.40		109
a.....			3.66 ^b ; 5.61 ^a	
b.....			3.66 ^b ; 5.61 ^a	
c.....			7.31 ^a	
AC 114.....	2.77	2.40	7.31 ^a	70
A2667.....	2.77	2.40	1.95 ^a	9.2

NOTES.—Exposure times, in kiloseconds, for *Spitzer* MIPS 24 μm , IRAC 3.6 μm , and IRS Long-Low grating. The 24 μm exposure time is the median (per-pixel) exposure time within the central 4' × 4' box.

^a Both LL1 and LL2 were given this exposure time.

^b Only order LL1.

background and noise images are produced by co-adding frames after masking target and serendipitous spectra on each frame. (The background is better estimated for the PID 30775 observations where both LL1 and LL2 orders provide redundant background measurements, compared to the PID 82 observations using only LL1, where the background can only be determined by differencing nodes and thus potentially suffers from contamination by faint sources.)

A robust statistical estimator (biweight) is used to minimize the effects of deviant pixels on the co-added value. This part is iterative in order to allow manual identification of sources. Rogue pixels are then identified by computing the dispersion of the noise around every pixel and flagging pixels which are $>5 \sigma$ deviant from the mean local value. Spectral extraction is done optimally (e.g., by using the PSF profile to weight the spectrum), taking into account the spectral distortion, and also rejecting pixels affected by cosmic rays. Finally, the spectra obtained at each position are co-added.

We downloaded and reduced one archival IRS spectrum, from program 3453 (PI P. van der Werf). This is SMM J163555.2+661238, image “A” in Kneib et al. (2004), one of three images of the same triply-imaged submillimeter galaxy in Abell 2218. We obtained an IRS spectrum of the brightest image of this galaxy, SMM J163554.2+661225 (image “B” in Kneib et al. 2004); its 24 μm flux is 1.8 times brighter than the SMM J163555.2+661238. In the archival spectrum, the source was miscentered across the short axis of the slit, which reduced throughput. Because its signal-to-noise ratio is low, we use the archival spectrum only to confirm

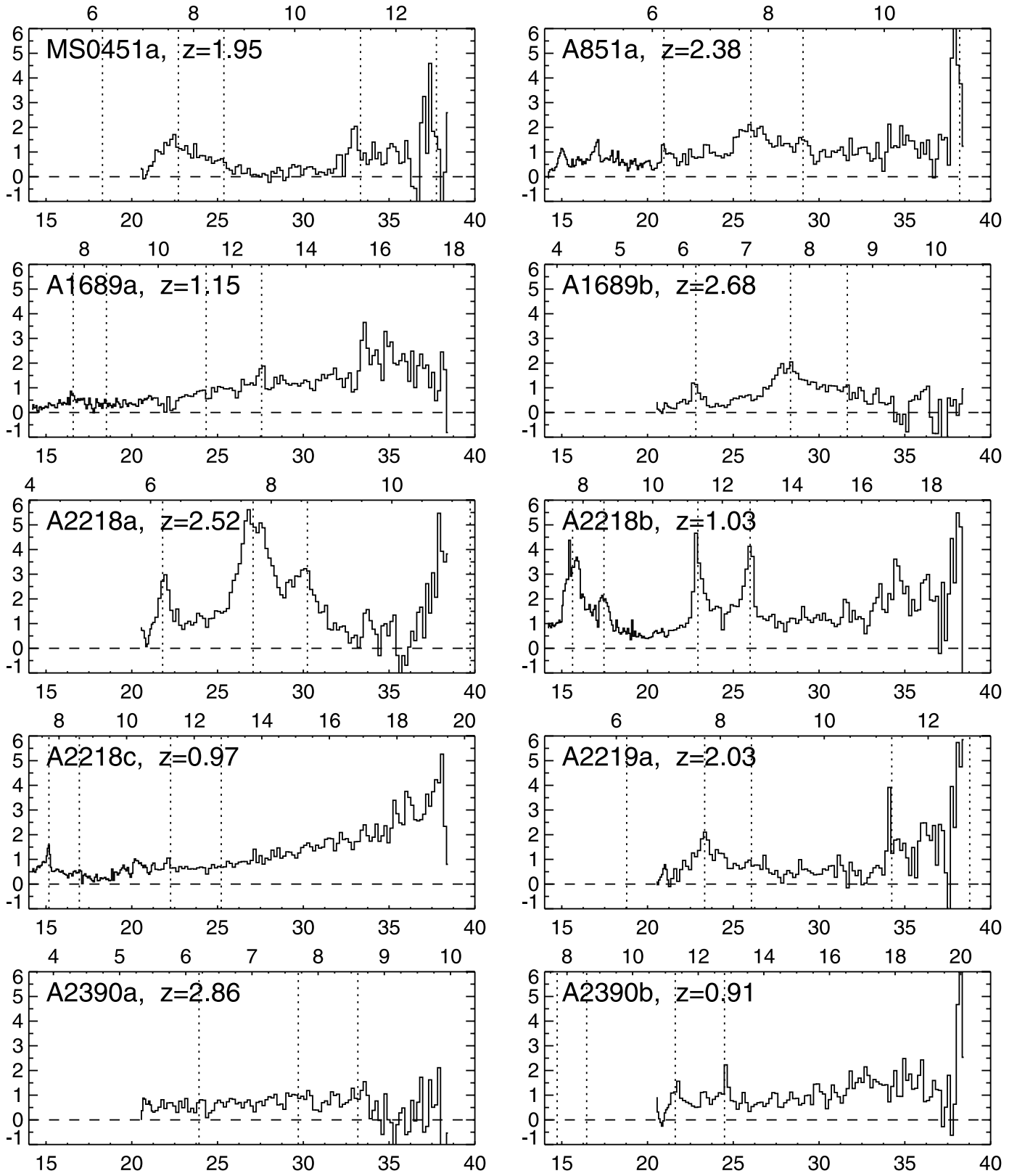


FIG. 2.—IRS spectra. Bottom x-axes show observed wavelength in μm ; top x-axes show rest wavelength (μm) if redshift is known; y-axes plot observed flux density (mJy). Spectra obtained for the same object in different programs are plotted separately, with program 82 plotted first. Vertical dotted lines show the positions of expected aromatic and fine structure emission features. Note that artifacts are present at $\lambda > 34 \mu\text{m}$.

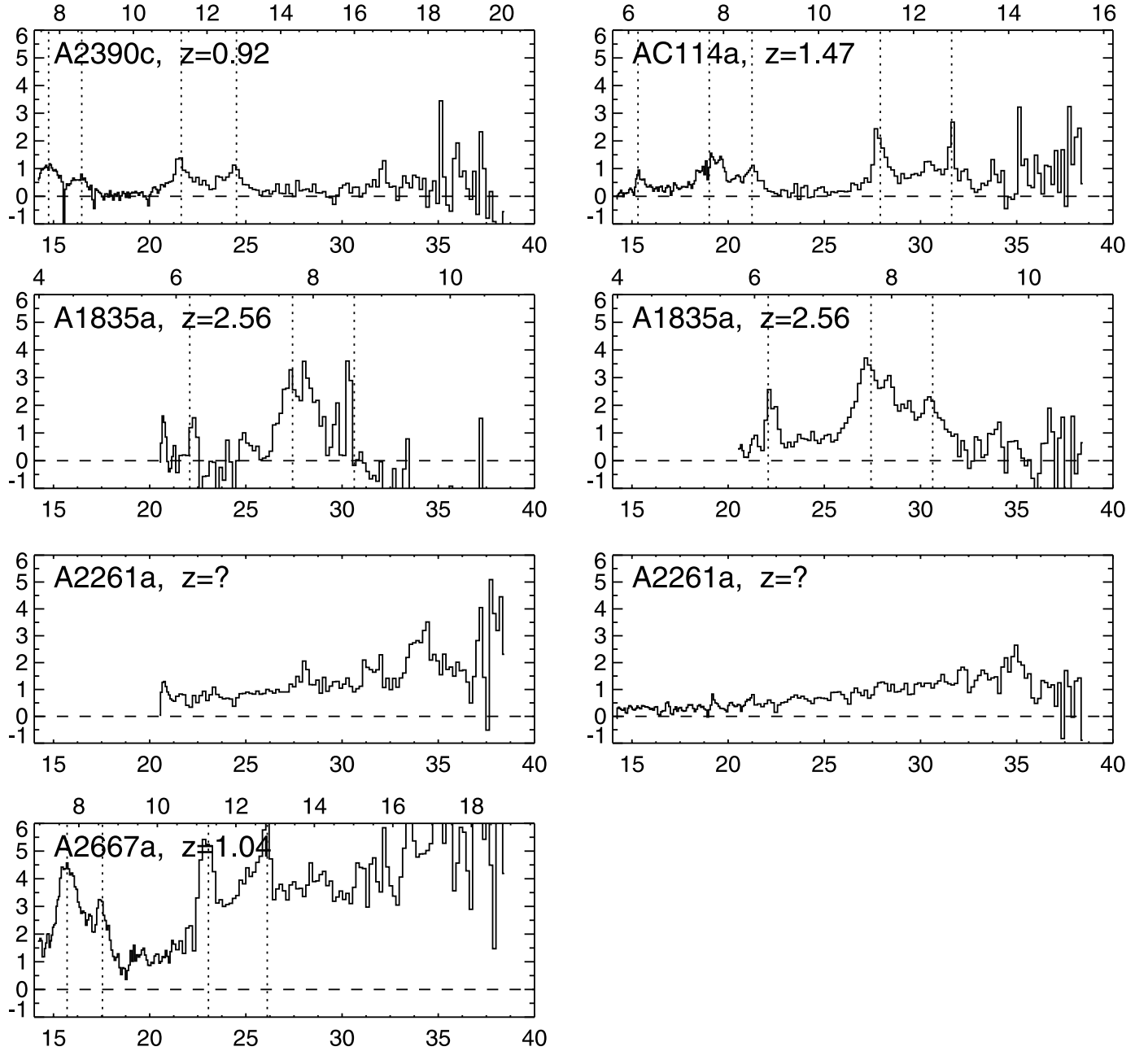


FIG. 2—Continued

the detection of features, and use the spectrum from our program for all analysis.

The IRS spectra are presented in Figure 2. Spectra obtained in programs 82 and 30775 for the same object are plotted separately; observations from the latter program have superior background subtraction. Flux densities are as observed, without any correction for gravitational amplification. When redshifts were known from the literature, or are evident from aromatic emission features in the IRS spectrum, then the upper x -axis shows rest wavelength. Only one source (A2261a) lacks a redshift. Redshifts are listed in Table 1, with a reference if the redshift is from the literature.

The rest-frame spectra were run through PAHFIT (Smith et al. 2007) to fit the continuum and aromatic (PAH) features simultaneously. Measured aromatic fluxes are reported in Table 4.

We allowed PAHFIT to fit the silicate absorption (e.g., we did not specify “NO_EXTINCTION”). Fixing the silicate optical depth at zero changes the best-fit $7.7 \mu\text{m}$ feature fluxes by 4%

(median), the best-fit $8.6 \mu\text{m}$ feature fluxes by 8% (median), and the best-fit $11.3 \mu\text{m}$ feature fluxes by 11% (median).

We report fluxes in Table 4 for features in which the measured flux is at least twice the uncertainty calculated for that line. We add back in a few features, mostly at $11.3 \mu\text{m}$, which are obviously detected even though the error bars are large.

3.2. Spitzer MIPS 24 and $70 \mu\text{m}$ Photometry

MIPS photometry-mode images at 24 and $70 \mu\text{m}$ were obtained through *Spitzer* GTO program 83 (PI G. Rieke). Photometry is given in Table 2, and exposure times are given in Table 3. The data were reduced and mosaicked using the Data Analysis Tool (Gordon et al. 2006) with a few additional processing steps (Egami et al. 2006b).

Photometry at $24 \mu\text{m}$ was obtained by PSF fitting, using the IRAF implementation of the DAOPHOT task `a11star`. The PSF was created empirically using all available images from program

TABLE 4
MEASURED AROMATIC FLUXES

Source	6.2	σ	7.7	σ	8.3	σ	8.6	σ	11.3	σ	12.0	σ	12.6	σ	13.6	σ	14.2	σ	16.4	σ	17	σ
MS 0451a.....	...		28	6	4.8	2			4.3	2			5.0	2	
A851a.....	2.8	0.6	15.6	0.8	2.4	1	2.7	0.8	
A1689a.....	...		12	5									
A1689b.....	6.5	1.1	21	3	3.8	1	4.0	2	
A1835a.....	13	2	39	10	7.3	2	7.8	2	
A2218a.....	18.3	0.9	66	3	8.2	1	13.5	1	
A2218b.....	...		65	3	4.8	1.7	16.3	2	20	3	3.8	1.2	10.8	1.3					1.7	0.8	14	5
A2218c.....	...		14.9	0.9			1.5	0.8	1.3	0.4												
A2219a.....	...		19.3	1.3			4.6	1	5.6	2			
A2390a.....	1.1	1	12.4	3			5.8	2	
A2390b.....		14	10			6.1	2.3								
A2390c.....	...		35	11			11	8	17	12			6.3	1.5								
AC 114a.....	6.5	0.5	23	2			7.1	1	9.2	1			4.2	2			
A2667a.....	...		102	11	12	3	22	4	29	7	9.1	3	16	3					5.6	1.9		

NOTES.—Aromatic dust emission feature fluxes and uncertainties, in units of 10^{-15} erg s $^{-1}$ cm $^{-2}$, measured with PAHFIT. Blank cells are nondetections. Column headers give the central wavelength, in μ m, of each feature. For features comprising multiple components (such as 7.7 μ m), total flux and uncertainty were calculated using PAHFIT’s “main feature” function. Ellipses indicate no coverage; blank indicates nondetections.

83. An aperture correction of 1.131 was applied, calculated from a Tiny Tim model of the 24 μm *Spitzer* PSF that extends to $r = 220''$ (C. Engelbracht, 2006, private communication).¹⁵

Aperture photometry at 70 μm was performed on the 1.0 pixel scale images. The aperture radius was 16'' (twice the HWHM), the background annulus had radii of 18'' and 39'' (the position of the first airy ring), and an aperture correction of 1.968 was applied.¹⁶ If the source was undetected, we took the upper limit as 2 times the 1 σ sky noise plus any positive flux in the source aperture.

3.3. *Spitzer* IRAC Photometry

IRAC images at 3.6, 4.5, 5.8, and 8 μm were obtained as part of GTO program 83 (PI G. Rieke) for five of the six clusters; IRAC data for Abell 1835 were obtained through GTO program 64 (PI G. Fazio). Exposure times are listed in Table 3. IRAC images were mosaicked as described by Huang et al. (2004).

Photometry was obtained by aperture photometry with sky subtraction. The high density of cluster galaxies in the cluster cores necessitated use of irregularly shaped polygons for both target and sky apertures. The flux density within these polygonal regions was determined using CIAO¹⁷ versions 3.2.1 and 3.4.1.1. IRAC photometry is listed in Table 2; the values do not include aperture corrections.

3.4. *HST* Photometry from *WFPC2* and *ACS*

From the *HST* archive, we downloaded all publicly available ACS images for the clusters. We used *multidrizzle* (Koekemoer et al. 2002) to distortion-correct each flat-fielded pipeline image, and then cross-correlated the images to measure the small ($<1''$) coordinate registration offsets with high precision. The measured offsets were then used by *multidrizzle* for the final mosaicking. The data were photometered using the same technique as for the IRAC bands.

We also used reduced *WFPC2* images kindly made available by D. Sand, which are described in Sand et al. (2005). ACS and *WFPC2* cutouts of the galaxies in our sample are shown in Figure 1.

3.5. *Chandra* ACIS Imagery

From the *Chandra* archive, we downloaded all publicly available ACIS observations of clusters with IRS targets in this sample: observation ID numbers 529 (obs/00529) and 902 (obs/00902) for MS 0451; 1663 (obs/01663), 5004 (obs/05004), and 540 (obs/00540) for A1689; 495 (obs/00495) and 496 (obs/00496) for A1835; 1454 (obs/01454), 1666 (obs/01666), and 553 (obs/00553) for A2218; 896 (obs/00896) for A2219; 5007 (obs/05007) and 550 (obs/00550) for A2261; 4193 (obs/04193), 500 (obs/00500), and 501 (obs/00501), for A2390; and 1562 (obs/01562) for AC114.

We used CIAO (3.3) to update the CTI correction when necessary and remove data obtained during periods of high solar flares. We mosaicked multiple overlapping observations following the CIAO thread “Reprojecting Images: Making an Exposure-corrected Mosaic.”

For X-ray detections, fluxes were determined from the individual (not mosaicked) observations as described in Marcillac et al. (2007). X-ray flux upper limits (3σ) were determined from the mosaicked images and mosaicked exposure maps as in Donley et al. (2005) except that the 90% encircled energy radius was

¹⁵ The aperture correction is specific to the DAOPHOT parameters used, in this case sky annulus radii of 31.3'' and 40'', and a PSF defined out to $r = 22.41''$.

¹⁶ See <http://ssc.spitzer.caltech.edu/mips/apercorr>.

¹⁷ See <http://xc.harvard.edu/ciao>.

TABLE 5
X-RAY FLUXES AND 3σ UPPER LIMITS

Source	$f(2-8\text{ keV})$ ($\text{erg s}^{-1}\text{ cm}^{-2}$)	$f(0.5-8\text{ keV})$ ($\text{erg s}^{-1}\text{ cm}^{-2}$)
MS 0451a.....	$<1.55 \times 10^{-15}$	$<1.30 \times 10^{-15}$
A851a.....
A1689a.....	$<7.9 \times 10^{-15}$	$<4.5 \times 10^{-15}$
A1689b.....	$<1.1 \times 10^{-14}$	$<7.2 \times 10^{-15}$
A1835a.....	$<5.8 \times 10^{-15}$	$<4.4 \times 10^{-15}$
A2218a.....	$<5.2 \times 10^{-15}$	$<2.6 \times 10^{-15}$
A2218b.....	$<2.8 \times 10^{-15}$	$<2.1 \times 10^{-15}$
A2218c.....	$<2.4 \times 10^{-15}$	$<1.6 \times 10^{-15}$
A2219a.....	$<3.4 \times 10^{-15}$	$<3.0 \times 10^{-15}$
A2261a.....	3.7×10^{-15}	1.8×10^{-15}
A2390a.....	5.1×10^{-14}	7.6×10^{-14}
A2390b.....	4.4×10^{-14}	2.5×10^{-14}
A2390c.....	$<7.5 \times 10^{-15}$	$<5.2 \times 10^{-15}$
AC 114a.....	$<2.8 \times 10^{-15}$	$<2.9 \times 10^{-15}$
A2667a.....	$<3.2 \times 10^{-14}$	$<4.1 \times 10^{-14}$

NOTES.—X-ray limits are the 3σ sky background plus any additional counts at the source position, as described in § 3.5.

used, and the Monte Carlo sky apertures were taken close to the source (within 30'') in order to sample a representative cluster background.

Chandra X-ray fluxes and upper limits for the 2–8 and 0.5–8 keV bands are reported in Table 5.

3.6. Long-Wavelength Photometry from the Literature

Eight galaxies in our sample, and one galaxy from the extended (literature) sample, have submillimeter detections in the literature; these are listed and referenced in Table 1. Four sources were detected at 15 μm by *Infrared Space Observatory* (ISO): A851a and A2218b (Barvainis et al. 1999), A2390b (Lemonon et al. 1998), and A2390c (Metcalf et al. 2003). We use this photometry to help determine the total infrared luminosities of the sources.

4. DISCUSSION

4.1. Aromatic Feature Flux Ratios in High-Redshift Galaxies

We now test whether the mid-IR spectra of star-forming galaxies, with their prominent aromatic emission features, show evidence for evolution from $z = 0$ to high redshift.

Given its high quality, we first examine the IRS spectrum for SMM J163554.2+661225 in detail. The image is amplified by a factor of 22 ± 2 (Kneib et al. 2004), such that without lensing, its 24 μm flux density would have been 53 μJy , at the MIPS confusion limit. The galaxy redshift, from $\text{H}\alpha$, is $z = 2.5165 \pm 0.0015$ (Kneib et al. 2004). Its IRS spectrum, presented in Figure 3, shows highly significant detections of aromatic features at 6.22, 7.7, 8.33, and 8.61 μm . (The lower quality spectrum from program 3453 confirms the detection of these aromatic features, although the two longer wavelength features are detected at low significance.) Overplotted with arbitrary normalization are two spectral templates: (1) the starburst galaxy NGC 2798 from the SINGS survey (Dale et al. 2006); and (2) the average spectrum of 13 nearby starburst galaxies with little apparent AGN contribution (Brandl et al. 2006), whose mean IR luminosity is $4.9 \times 10^{10} L_{\odot}$. Coincidentally, this is also the $L(\text{TIR})$ of NGC 2798.

The $z = 2.516$ spectrum in Figure 3 closely resembles the low-redshift, lower luminosity (by a factor of 15) templates. Aromatic flux ratios are reported in Table 6, along with those we measure

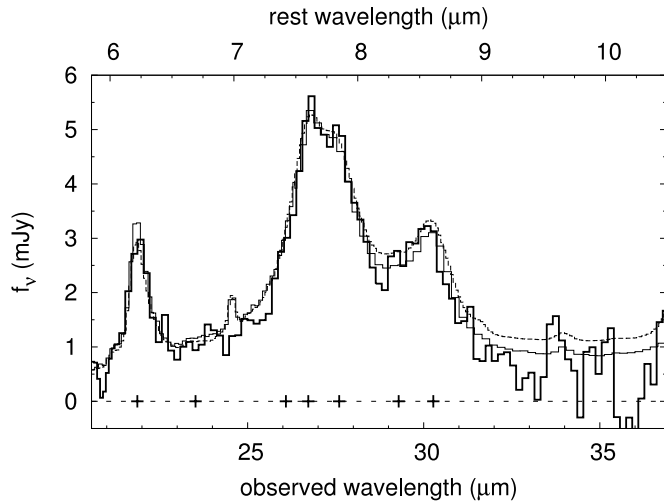


FIG. 3.—IRS spectrum of $z = 2.516$ lensed source SMM J163554.2+661225 behind Abell 2218 (solid thick line). Overplotted are two $z \sim 0$ spectral templates: the starburst galaxy NGC 2798 from Dale et al. (2006; thin solid line); and the average of 13 starburst galaxies from Brandl et al. (2006; thin dashed line). Crosses show the wavelengths of known aromatic components from Smith et al. (2007).

for the Brandl template, and those reported for the SINGS sample. The strength of the $8.6 \mu\text{m}$ feature in SMM J163554.2+661225 relative to those at 6.2 and $7.7 \mu\text{m}$, is within the 10%–90% variation within the SINGS sample. In addition, the flux ratios of SMM J163554.2+661225 are a close match to those of the Brandl et al. (2006) average starburst template; while $7.7 \mu\text{m}$ is relatively stronger than the template, the difference is comparable to the standard deviation of line ratios for individual galaxies that were averaged to create the template. *Thus, we conclude that this $z = 2.5$ galaxy has aromatic feature flux ratios that are consistent with those observed for lower luminosity, $z = 0$ starbursting galaxies.*

While the other spectra in the sample have lower signal-to-noise ratios, they are still sufficient, in the aggregate, to test whether high-redshift galaxies have markedly different aromatic flux ratios. Figure 4 compares, for four aromatic feature flux ratios, the observed ratios in our sample with those measured for the low-redshift (Brandl et al. 2006) template. Although the error bars in individual measurements are sizable, Figure 4 confirms the result seen for SMM J163554.2+661225: star-forming galaxies at $1 < z < 3$ have aromatic feature ratios consistent with those of $z = 0$ star-forming galaxies.

Variations in aromatic feature ratios have been reported in certain H II regions—this may be attributable to variations in size, composition, and ionization state of the carriers (Draine & Li 2007 and references therein). In particular, large PAHs emit more strongly at $11.3 \mu\text{m}$, and neutral PAHs emit much more strongly at 3.3 and $11.3 \mu\text{m}$ than charged PAHs (Draine & Li 2007; Galliano 2006; Allamandola et al. 1989). That our measured feature ratios at $z > 1$ are consistent with those measured at $z \sim 0$ suggests that the dust grain size and ionization distributions are not strongly evolving.

4.2. Evidence for Compact Source Accretion

Active galactic nuclei are detectable in multiple ways—by the dust they heat that radiates in the mid-IR; by X-rays emitted from the accretion disk coronae; and by high-excitation or broad emission lines. We now examine the evidence for AGN activity in our extended sample.

TABLE 6
AROMATIC RATIOS FOR SMM J163554.2+661225

Lines (1)	Ratio (2)	SINGS (3)	Starburst (4)
$7.7 \mu\text{m}/6.2 \mu\text{m}$	3.6 ± 0.2	3.6 (1.3–4.8)	3.9
$7.7 \mu\text{m}/8.6 \mu\text{m}$	4.9 ± 0.4	5.7 (4.7–9.0)	5.1
$6.2 \mu\text{m}/8.6 \mu\text{m}$	1.36 ± 0.1	1.5 (1.2–3.0)	1.3

NOTES.—Col. (1): Line wavelengths (μm); Col. (2): measured flux ratios; Col. (3): median flux ratios and 10%–90% range of variation from the SINGS sample (Smith et al. 2007); Col. (4): flux ratios of the average starburst template of Brandl et al. (2006).

The first AGN diagnostic we examine is X-ray luminosity, plotted in Figure 5 against aromatic luminosity. Most of the sample are X-ray nondetections, at limiting luminosities that rule out X-ray-loud QSOs or bright Seyfert galaxies, although an X-ray-weak or highly obscured AGN could still be present (see Donley et al. 2005; Alonso-Herrero et al. 2006a).

Two of the four galaxies from the literature, and three of the galaxies in our program are detected in X-rays: two have no apparent aromatic features (sources A2261a and A2390a); and the other, source A2390b, is interesting in that it contains a luminous X-ray-emitting AGN, yet its spectrum still shows aromatic features. That source emits roughly equal power in the $7.7 \mu\text{m}$ aromatic feature (scaled from the $11 \mu\text{m}$ feature) and at 10 – 30 keV, reminiscent of the more luminous $z = 1.15$ source CXO GWS J141741.9+522823 discussed by Le Floc’h et al. (2007).

We now consider a second AGN diagnostic, the relative contribution of aromatic versus continuum emission to the mid-IR flux. At low redshift, low aromatic feature equivalent widths have been demonstrated as an effective AGN diagnostic (Armus et al. 2007; Brandl et al. 2006). However, it is extremely difficult to measure equivalent widths accurately for the high-redshift IRS sources, due to the limited wavelength baseline, the fact that the aromatic features have broad wings, and most importantly, imperfect sky subtraction which adds a (positive or negative) pedestal to each spectrum. For example, for SMM J163554.2+661225 (our highest quality spectrum), we cannot measure an accurate equivalent width due to the difficulty in determining the true continuum level. We are able to set a lower limit of $>0.4 \mu\text{m}$ (rest-frame, conservatively assuming a high continuum level), which is a typical value for low-redshift star-forming galaxies. For the other spectra in the sample, equivalent widths are even harder to measure.

Therefore, we instead create an aromatic-to-mid-IR flux ratio, which we define as the ratio of the flux in the $7.7 \mu\text{m}$ PAH feature (in $\text{ergs s}^{-1} \text{cm}^{-2}$, as fit by PAHFIT) to the MIPS photometry at $24 \mu\text{m}$ (in Jy, corrected for bandwidth compression to $z = 0$.) This metric does not suffer as strongly from the difficulty in determining the continuum level. In the low-redshift, star-forming comparison sample we construct in § 4.4, this ratio ranges from $(7\text{--}32) \times 10^{-11}$. We thus take 7×10^{-11} as a dividing line between spectra dominated by star formation (above the value), and spectra with substantial AGN contribution (below the value). This discriminant selects the following mid-IR spectra¹⁸ as having substantial AGN contribution: A370a, MS 0451a, A1689a, A2218c, A2261a, and A2390a. Thus, of 19 galaxies with adequate spectra, six show strong indications of AGN contribution to their mid-IR outputs.

¹⁸ For A2390b, we scale from the flux in the $11 \mu\text{m}$ aromatic feature, using the average $7.7/11 \mu\text{m}$ flux ratio of 3.6 from Smith et al. (2007).

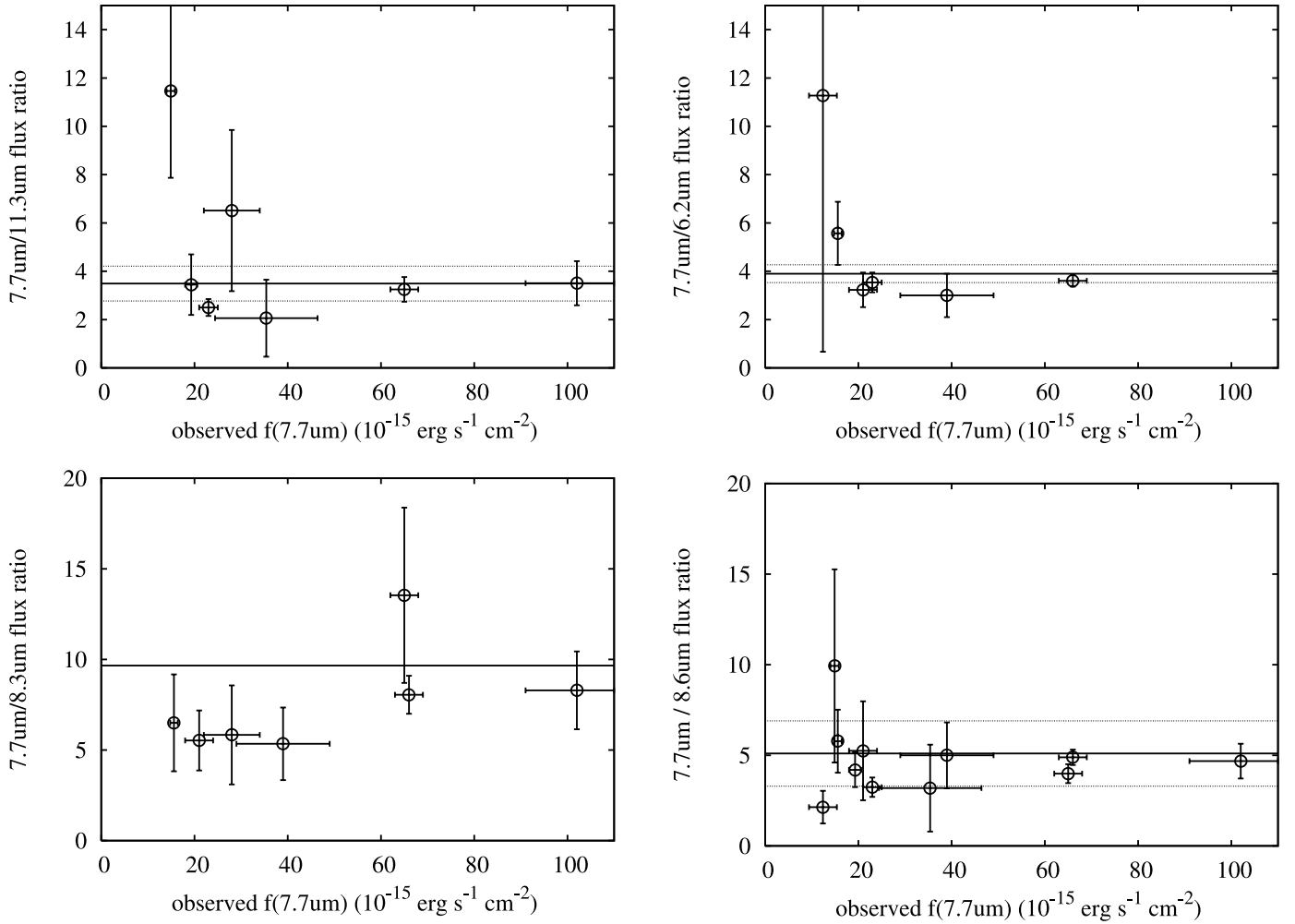


FIG. 4.—Aromatic flux ratios. The aromatic flux ratios are plotted against observed flux in the 7.7 μm feature. Fluxes were fit using PAHFIT. Amplifications have not been divided out. Also plotted are the flux ratios of the low-redshift Brandl et al. (2006) starburst template (*solid horizontal lines*), with the ± 1 standard deviations from that sample (*dotted horizontal lines*).

A third indication of AGN activity is a mid-IR spectrum that rises steeply with increasing wavelength. Four spectra (A2261a, A2218c, A1689a, and A370a) show this behavior; none has detected aromatic emission.

A fourth indication of AGN activity is the presence of broad or highly excited lines in an optical spectrum. Of the spectroscopy that has been published, or is in preparation, for our sample (see references in Table 1), three show evidence for AGNs: A1689a has a Keck spectrum showing highly ionized neon (J. Richard et al., in preparation); A2390a has a Keck spectrum (see Appendix) which shows Ly α with FWHM $\sim 945 \text{ km s}^{-1}$, typical of an AGN narrow-line region; and the spectrum of A370a is reported to contain AGN lines (Lutz et al. 2005).

A fifth indication of AGN activity is a high ratio of [Ne III] 15.5 μm to [Ne II] 12.8 μm . Solar-metallicity starbursts are not observed to excite this line ratio above unity (Thornley et al. 2000; Rigby & Rieke 2004), whereas the harder continuum of AGNs do. Eight spectra cover rest-frame 12.8 μm ; [Ne II] is detected in six of these. Six spectra cover rest-frame 15.5 μm ; [Ne III] is detected in four. Line fluxes are added in Table 7. Of the four galaxies where both lines are detected (A1689a, A2218b, A2390b, and A2667a), only in A1689a does the [Ne III]/[Ne II] ratio exceed unity: 2.15 ± 0.76 . This high line ratio offers additional evidence that A1689a hosts a luminous AGN. The other, lower line ratios are more difficult to interpret.

The results for these AGN diagnostics are compiled in Table 8. Combining these diagnostics, 8 galaxies out of the sample of 19 have at least one indication of AGN activity (X-ray detection, low aromatic contribution to the mid-IR flux, rising mid-IR spectrum, or optical AGN lines). Five galaxies have two different diagnostics that indicate AGNs. Thus, we find that 42% of the faint 24 μm sources in this sample have some nuclear activity; the 95% confidence interval is 23%–63%. Further, in 32% of the sample, the AGN strongly affects or dominates the mid-IR spectrum; the 95% confidence interval is 15%–53%. (Confidence intervals were calculated using eq. [26] of Gehrels 1986.)

The high proportion of AGNs in our sample of high-redshift LIRGs and ULIRGs mirrors the general behavior of local IR-luminous galaxies, where AGNs become substantially more common with increasing luminosity (e.g., Lutz et al. 1998). Although the error bars are large given the small sample size, the high proportion (32%) of AGNs sufficiently powerful to dominate the mid-IR spectra may be inconsistent with the frequent assumption that $\sim 100 \mu\text{Jy}$, $z > 1$ sources are dominated by star formation, with AGNs a modest contaminant. Our result, if confirmed with better statistics, has implications for 24 μm -derived star formation rates in deep surveys, where it is often assumed that the mid-IR outputs of high-redshift, faint galaxies are dominated by strong aromatic features from star formation. It also implies that a higher fraction of the integrated

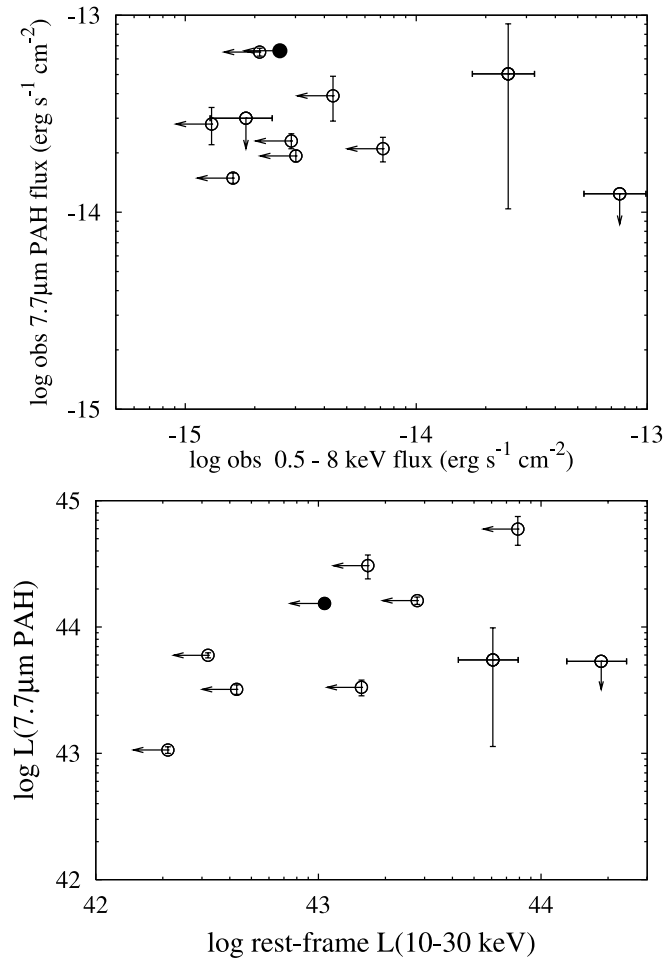


FIG. 5.— Comparison of aromatic and X-ray fluxes and luminosities. SMM J163554.2+661225 is represented by filled symbols.

TABLE 7
[Ne III] 15.5 μm AND [Ne II] 12.8 μm

Source	[Ne II]	σ	[Ne III]	σ	[Ne III]/[Ne II]	σ
MS 0451	ND		...			
A851			
A1689a	0.93	0.26	2.0	0.43	2.15	0.76
A1689b			
A2218a			
A2218b	2.9	0.6	0.91	0.31	0.31	0.13
A2218c	ND		ND			
A2219			
A2390a			
A2390b	2.2	0.57	0.90	0.69	0.41	0.33
A2390c	0.83	0.47	ND			
AC 114a	1.6	0.29	...			
A1835a			
A2261a			
A2667a	2.5	0.49	0.92	0.51	0.37	0.22
Tep 1			
T1-bzk			
Tep 2			
A370a			

NOTES.—Measured fine structure line fluxes, from PAHFIT, in units of $10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$. Ellipses indicate that the wavelength regime was not covered; “ND” indicates that the line was covered but not detected.

TABLE 8
AGN DIAGNOSTICS

Source (1)	X-Ray (2)	Weak Aromatic (3)	Rising (4)	Optical (5)
MS0451a		X
A851a
A1689a		X	X	X
A1689b	
A1835a	
A2218a	
A2218b				
A2218c		X	X	...
A2219a
A2261a	X	X	X	...
A2390a	X	X	...	X
A2390b	X			
A2390c				
AC 114a
A2667a				
Tep-1	X		...	
Tep-1-BzK
Tep-2				
A370a	X	X	X	X

NOTES.—“X” indicates detection; blank indicates nondetection; ellipses indicate no data. Col. (1): source name; Col. (2): X-ray detection; Col. (3): low aromatic flux contribution, as defined in § 4.2; Col. (4): mid-IR spectrum rises strongly toward the red; Col. (5): AGN lines in the optical spectrum.

24 μm flux on the sky may arise from accretion than previously thought.

4.3. Total Infrared Luminosity

We now determine the total infrared (TIR) luminosities of the lensed sources, so that we may examine the dependence of aromatic feature emission on total infrared luminosity, and how this behavior may evolve with redshift.

4.3.1. SMM J163554.2+661225

The submillimeter galaxy behind Abell 2218 has exceptionally good photometric coverage from 0.4 to 850 μm . Since it is triply-lensed, we increase the signal-to-noise ratio of its multi-band photometry by using fluxes from all three images, weighed by the amplifications given in Kneib et al. (2004). (Bright neighboring galaxies forced us to use only image A and B for the optical and IRAC photometry.) Figure 6 plots this photometry.

We fit redshifted Chary & Elbaz (2001) and Dale & Helou (2002) templates to the 24, 70, 450, and 850 μm photometry, corrected for lensing amplification. For each template family, we determine the best-fit template and the range of acceptable templates, then determine $L(\text{TIR})$ for those templates in a method consistent with the definition given in Table 1 of Sanders & Mirabel (1996) using 12, 25, 60, and 100 μm photometry. We estimate the intrinsic total infrared luminosity of SMM J163554.2+661225 as follows:

1. $7.6 \pm 1.9 \times 10^{11} L_{\odot}$ for Dale & Helou (2002) templates.
2. $1.3^{+1.0}_{-0.2} \times 10^{12} L_{\odot}$ for Chary & Elbaz (2001) templates.

As Figure 6 shows, the best-fit Chary & Elbaz (2001) template overpredicts all $>24 \mu\text{m}$ points, especially at 70 μm . The best-fit Dale & Helou (2002) template is a better match to the shape of the overall infrared spectral energy distribution (SED) and also has high aromatic feature equivalent widths, consistent with the spectrum of the source. Thus, this galaxy’s intrinsic luminosity is that of a LIRG.

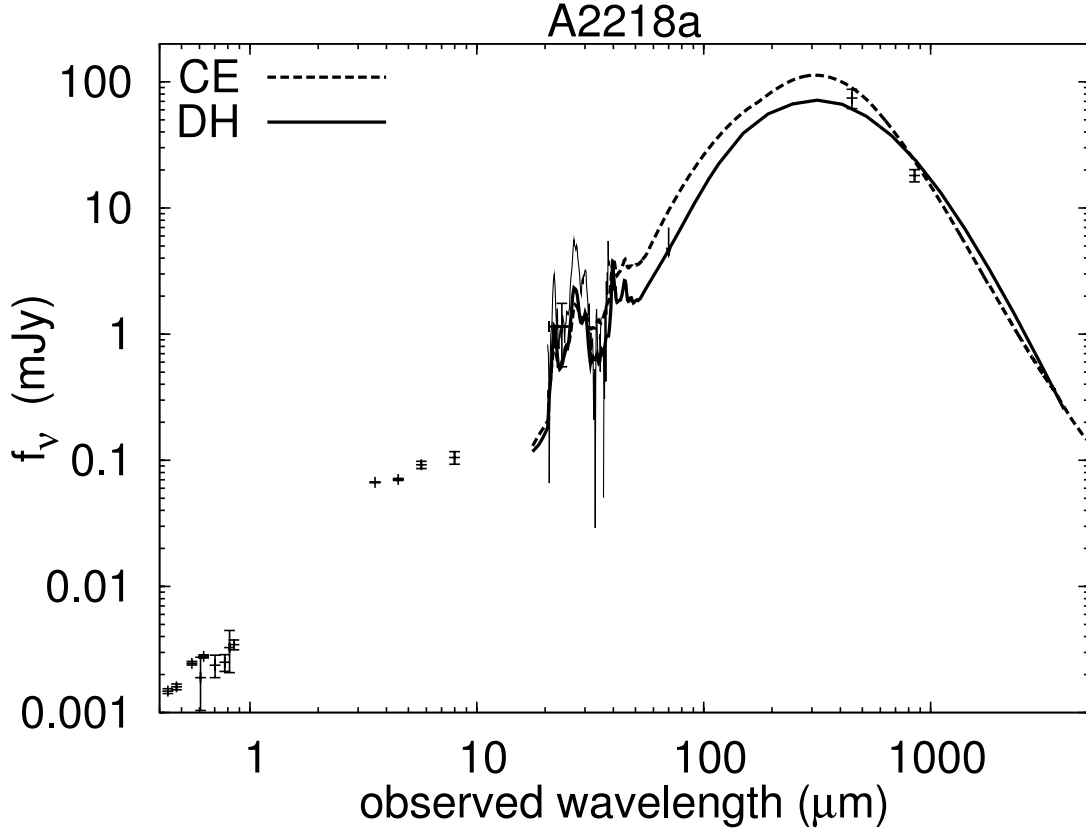


FIG. 6.—Photometry and $L(\text{TIR})$ fits. Photometry and IRS spectra (*thin lines*) are plotted, with flux densities and wavelengths in the observed frame. Overplotted are best-fit templates: Dale & Helou (2002; *thick solid line*), Chary & Elbaz (2001; *thick dashed line*), and Mrk 231 (*dotted line*). IRAC and *HST* photometry are plotted for SMM J163554.2+661225 to demonstrate that its SED is dominated by star formation.

4.3.2. Other Galaxies

We adopt a similar approach for the other galaxies with sufficient far-infrared and submillimeter data to provide reasonable constraints. We tried both Dale & Helou (2002) and Chary & Elbaz (2001) templates, fitted to fall just below any upper limits so long as they were compatible with the other constraints, i.e., to give a maximal luminosity. Two galaxies are fit poorly by a star-forming template, and much better by an AGN template, Mrk 231; MS 0451a may be, as well. The ranges of acceptable luminosities and the preferred templates are listed in Table 3. Dale & Helou (2002) templates provided generally better fits to our sample than Chary & Elbaz (2001) templates, as expected given the results of Marcillac et al. (2006) for $0 < z < 1$ galaxies, fitting *ISO*, *Spitzer*, and VLA photometry.

4.4. $8 \mu\text{m}$ (Rest)-to-Total Luminosity Ratio

We now consider the aromatic and total infrared (TIR) luminosities of the star-forming galaxies in the lensed sample, and how this ratio behaves as a function $L(\text{TIR})$. We compare to higher luminosity star-forming galaxies at $z \sim 2$, as well as comparable-luminosity star-forming galaxies at $z \sim 0.8$ and $z \sim 0$. We interpret the properties of the high-redshift galaxies strictly in terms of properties that are well determined for local analogs, so we can test whether inconsistencies emerge.

4.4.1. The $z \sim 0$ Comparison Sample

To probe rest-frame $8 \mu\text{m}$ and total infrared (TIR) luminosities for local galaxies, we use the SINGS sample (photometry from Dale et al. 2005) supplemented in the ULIRG range by IRAS 09111-1007, 10565+2448, 12112+0305, 14348-1447, 17208-0014,

22491-1808, and Arp 220 (Farrah et al. 2003; Armus et al. 2007), all selected because they appear to be dominated by young stars rather than AGNs.¹⁹ We also include IRAS 00262+4251, 01388-4618, 02364-4751, 16474+3430, 23128-5919, and 23365+3604 from Rigopoulou et al. (1999). The strength of the aromatic features relative to the continuum indicates that all these galaxies are also dominated by star formation. For the galaxies from Farrah et al. (2003) and Armus et al. (2007) we determined $8 \mu\text{m}$ photometry from archival IRAC images. In our reductions, we included a correction for the extended source calibration. For those galaxies in Rigopoulou et al. (1999) we converted the tabulated continuum and aromatic $7.7 \mu\text{m}$ peak fluxes into $8 \mu\text{m}$ photometric values by comparing with templates for galaxies in common between Armus et al. (2007): IRAS 12112+0305, 14348-1447, 1525+3609, and 22491-1808, plus Arp 220. The peak-to-peak scatter was a factor of 2 (in agreement with the finding of Chary & Elbaz 2001). We add a variety of LIRGs (Alonso-Herrero et al. 2006b), again avoiding any galaxies with indications of significant levels of AGN power. We took *Infrared Astronomical Satellite* (IRAS) measurements from Sanders et al. (2003) and computed $L(\text{TIR})$ to be consistent with the approach taken in that paper. We define $L(8 \mu\text{m})$ as being proportional to νf_ν , and use similar definitions for all luminosities at specific wavelengths. Figure 7 shows the relation between rest-frame $L(8 \mu\text{m})$ and $L(\text{TIR})$ for these $z \sim 0$ galaxies.

4.4.2. Results from the Literature at $z \sim 0.85$

We add a sample of infrared galaxies at $z \sim 0.85$ from Marcillac et al. (2006). This paper utilizes $15 \mu\text{m}$ data from *ISO*, as well

¹⁹ IRAS 12112+0305 and 14348-447 are double nuclei sources.

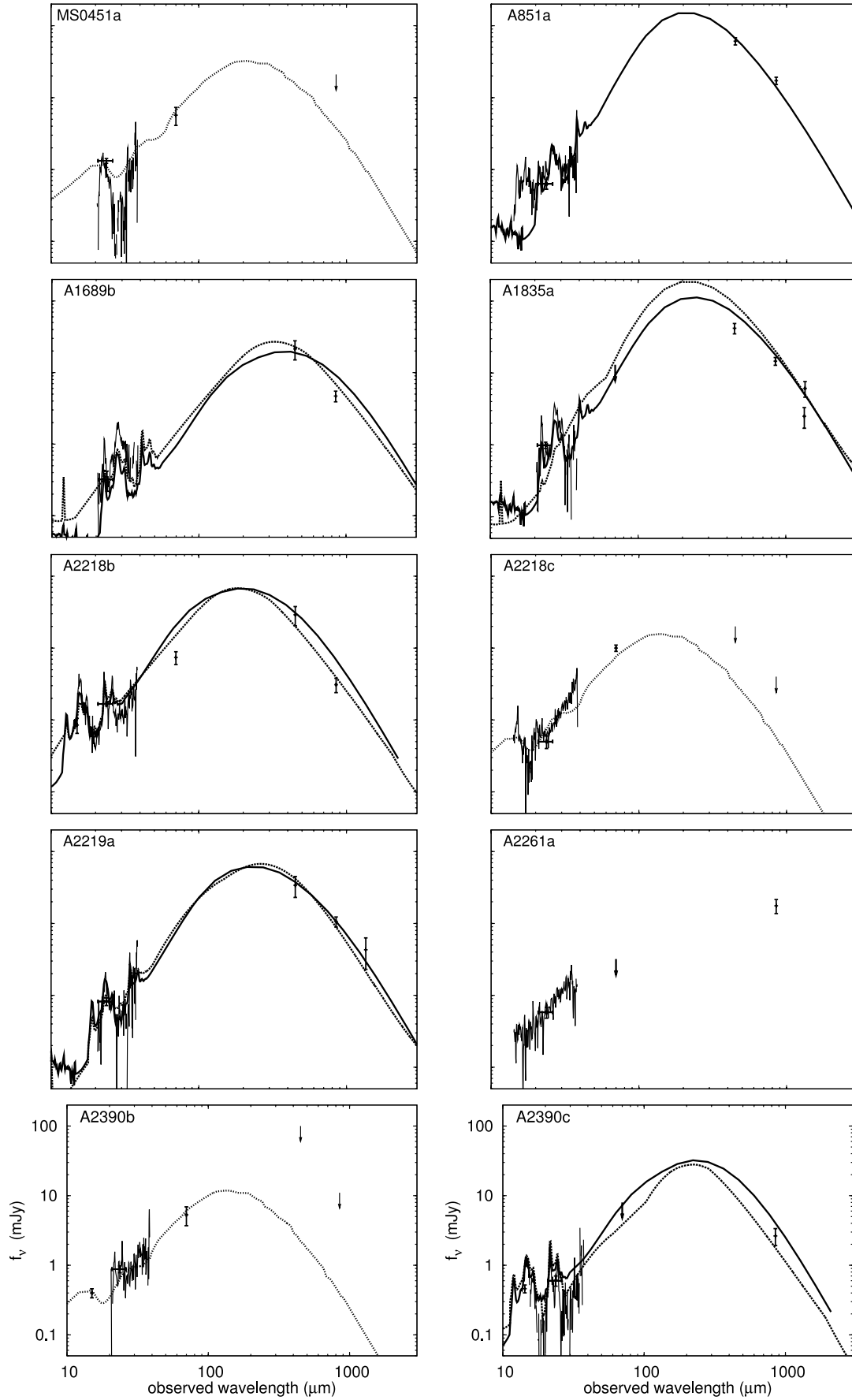


FIG. 6—*Continued*

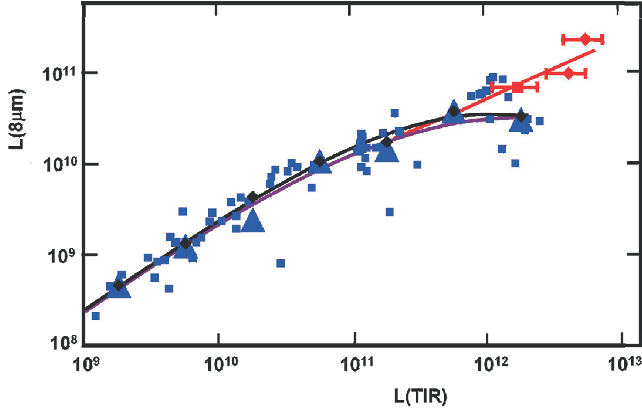


FIG. 7.—Comparison of $L(8\ \mu\text{m})$ and $L(\text{TIR})$ locally and for stacked measurements at $z \sim 2$. The individual local galaxies are indicated as small blue squares, while the values averaged over $L(\text{TIR})$ intervals of 0.5 dex are shown as triangles. The sizes of triangles indicate the uncertainties in the average values for the bins, estimated from the scatter in the values. A fit to the averages is shown as a blue line. The black diamonds and line show a similar fit after rejecting high and low outliers in the luminosity bins (see text). The stacked results at $z \sim 2$ are shown as a red square (Daddi et al. 2005) and two red diamonds (Papovich et al. 2007). The errors in these results have been consolidated into errors in $L(\text{TIR})$, since the conversion to this parameter is the dominant uncertainty. The error bars show uncertainties by a factor of 1.5 for Daddi et al. (2005) and 1.4 for Papovich et al. (2007). The red line is a fit to the $z \sim 2$ stacked points, constrained to agree with the local fit at $10^{11}\ L_{\odot}$.

as $24\ \mu\text{m}$ data from *Spitzer*, and, for a subset, radio data from the VLA; the authors found that the galaxies had rest-frame $L(8\ \mu\text{m})/L(12\ \mu\text{m})$ ratios at the high end of the local galaxy dispersion. We make the assumption that $L(12\ \mu\text{m})$ correlates with $L(\text{TIR})$ in that sample, as it is observed to do well at $z = 0$. We computed rest $8\ \mu\text{m}$ luminosities for the members of the sample with $0.55 < z < 1.2$ and adequate $15\ \mu\text{m}$ photometry, and corrected them to the IRAC $8\ \mu\text{m}$ band by comparison with the IRS spectrum of IRAS 2249, a local star-formation-dominated ULIRG. We took $L(\text{TIR})$ from the Dale & Helou (2002) template fits (Marcillac et al. 2006). The results are shown in green in Figure 8.

There is no mid-IR spectroscopy for these galaxies, but if they had strong mid-IR contributions from AGNs, we would expect substantial dispersion in the $L(8\ \mu\text{m})$ to $L(\text{TIR})$ ratio. Such a large dispersion is not seen.

4.4.3. Results from the Literature at $z \sim 2$

We now compare the behavior of the $8\ \mu\text{m}$ luminosity versus $L(\text{TIR})$ at $z \sim 2$ with that seen in local luminous galaxies. We first discuss three results from the literature, and then show how our new measurements complement them.

1. First, we took the stacked $z \approx 1.9$ galaxy template of Daddi et al. (2005) and compared the ratio of observed 850 to $24\ \mu\text{m}$ flux densities, referred to the appropriate rest wavelengths for each of the local star-forming galaxies where there is sufficient data to construct an accurate SED template (IRAS 1211, 1434, 17208, 2249, and Arp 220). Excluding Arp 220, for which the ratio is 54, the average value for the other four galaxies is 14, whereas the Daddi et al. (2005) template indicates a value of 6.5. Thus, the template indicates stronger emission at rest $8\ \mu\text{m}$ relative to the submillimeter than is typical of local star-forming ULIRGs.

We also compare with total infrared luminosity derived from submillimeter measurements. Since the rest wavelength for the observed $850\ \mu\text{m}$ point in the Daddi et al. template is $293\ \mu\text{m}$, we

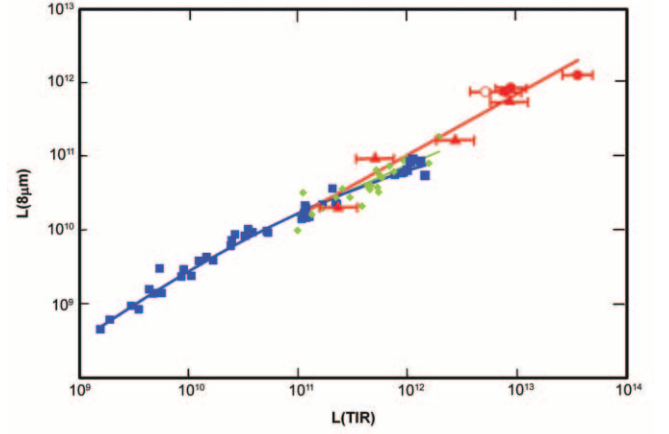


FIG. 8.—Comparison of $L(8\ \mu\text{m})$ and $L(\text{TIR})$ locally and for galaxies at $z \sim 2$ measured individually. The individual local galaxies with above average $L(8\ \mu\text{m})/L(\text{TIR})$ are indicated as blue squares. The blue line is a fit to them. The high- z galaxies from Yan et al. (2007) are shown as red circles, filled for those detected at $1.4\ \text{GHz}$ and open for the undetected example. The filled triangles are the lensed galaxies from this work. The errors have been consolidated into $L(\text{TIR})$. They are shown as a factor of 1.4 for the Yan et al. (2007) galaxies, equivalent to the ratio of flux densities at $60\ \mu\text{m}$ to $1.4\ \text{GHz}$ varying from 100 to 200. For the lensed galaxies, we show errors by a factor of 1.5, based on the range of luminosities indicated by our template fits. The red line is a fit to the $z \sim 2$ points, constrained to agree with the local fit at $10^{11}\ L_{\odot}$. The green points are individual galaxies at $z \sim 0.85$, from Marcillac et al. (2006), and the green line is a fit to those points constrained to agree with the local fit at $10^{11}\ L_{\odot}$.

construct the average ratio of $L(\text{TIR})/L(293\ \mu\text{m})$ for local LIRGs and ULIRGs with $4 \times 10^{11} < L(\text{TIR}) < 2.5 \times 10^{12}\ L_{\odot}$ and high-quality measurements at $350\text{--}450\ \mu\text{m}$ (IRAS 09111, 1211, 1434, and 17208, Arp 220, and NGC 1614). We correct the flux densities for the differing effective filter bandwidths using the IRAS 2259 template. We correct to $293\ \mu\text{m}$ assuming a spectral slope of -3.3 , determined from a larger sample of local ULIRG templates. We find

$$L(\text{TIR}) = (98 \pm 40) L(293\ \mu\text{m})$$

and then plot the Daddi et al. point as a red square in Figure 7. We did not fit the observations of lower luminosity galaxies, because the scatter at lower luminosities is very large, probably due to varying amounts of cold and warm dust in the infrared-emitting objects.

2. Second, we make use of the stacking analysis carried out by Papovich et al. (2007). They report observed-frame 24 and $70\ \mu\text{m}$ stacked photometry for galaxies that are both undetected in X-rays and have non-power-law IRAC SEDs, and are therefore assumed to be powered predominantly by star formation. We computed the rest 8 and $24\ \mu\text{m}$ flux densities for this category (since they quote values for $z = 2$, the rest wavelengths are represented directly by the measurements). We corrected the $8\ \mu\text{m}$ flux density (derived from that observed at $24\ \mu\text{m}$) to the equivalent for the IRAC band, again using the SED of IRAS 2249. Finally, we compute $L(\text{TIR})$ from the relation derived from our local galaxy sample:

$$L(\text{TIR}) = 86.5 \times L(24\ \mu\text{m}, \text{rest})^{\alpha},$$

where $\alpha = 0.9758$.

The rms scatter around this relation is only 0.14 dex. The resulting values are shown in Figure 7 as red diamonds (we omitted the lowest flux density bin from Papovich et al. because the signal-to-noise ratio at $70\ \mu\text{m}$ is too low).

3. Third, we use the results of Yan et al. (2007). We select those galaxies that have $z > 1.5$, and rest-frame $6.2 \mu\text{m}$ equivalent widths (measured by Sajina et al. 2007) greater than $0.3 \mu\text{m}$. From Armus et al. (2007) this equivalent width threshold should select galaxies dominated by star formation. The galaxies selected (MIPS 289, 8493, 15928, 16144, 22530) are all categorized as “type 1” by Yan et al. (2007), meaning that their spectra contain strong aromatic features. The median rest-frame EW($7.7 \mu\text{m}$) is $2.5 \mu\text{m}$, corroborating that star formation dominates the mid-IR spectra. We converted the tabulated $L(5.8 \mu\text{m})$ for these galaxies into $L(8 \mu\text{m})$ from an average template for our local ULIRG sample (IRAS 2249).

We estimate $L(\text{TIR})$ for these sources using the radio-infrared relation. We describe our adjustment of this relation for use at high redshift in detail in a forthcoming publication (Seymour et al. 2007) but give a short summary here. Yun et al. (2001) have quantified the relation for a large number of local galaxies, finding that the average ratio of flux density at $60 \mu\text{m}$ to 1.4 GHz is 110 independent of luminosity. The results of Yun et al. (2001) have not been K -corrected; since the infrared SEDs drop rapidly toward the blue, while the radio spectrum drops slowly, these corrections are a fairly steep function of redshift. They enter at a significant level only for the high-luminosity galaxies, because they tend to be at higher redshift than the lower luminosity ones. Including them increases the indicated ratio to ~ 170 at a $L(\text{TIR}) = 10^{12} L_{\odot}$. A second issue is that the radio outputs of LIRGs and ULIRGs often show signs of free-free absorption, which can decrease the outputs at 1.4 GHz . To estimate this effect, we took the 19 star-forming galaxies with adequate data from Condon et al. (1991) and computed an average slope for them of -0.58 between 1.49 and 8.44 GHz . Taking the 8.44 GHz value to represent the unabsorbed radio output and assuming the intrinsic slope is -0.7 (Condon et al. 1991), we revise the IR/radio ratio downward to 140 to take account of absorption; this value is calculated by extrapolating from 8.44 to 1.4 GHz and comparing results for slopes of -0.58 and -0.7 . The second concern is that the radio-infrared ratio may be a function of redshift (Kovács et al. 2006; Vlahakis et al. 2007). We tested this suggestion by determining a K -correction between 1.4 GHz and $850 \mu\text{m}$ for an average of local LIRG and ULIRG templates (A. Alonso-Herrero et al., in preparation) and applying it to a large sample of high-redshift sources. We find that the K -correction accurately reproduces the observed behavior with no offset, indicating there is no significant evolution of the relation. We have therefore used a value of 140 for the intrinsic flux density ratio at rest wavelengths to convert the radio measurements to TIR ones. We extrapolated the radio fluxes observed to a rest frequency of 1.4 GHz with a slope of -0.7 . The redshift usefully makes the observed rest frequency significantly higher than 1.4 GHz , reducing the effects of free-free absorption for the high- z galaxies. To convert to $L(\text{TIR})$, we use the relationship from our local sample, $L(\text{TIR}) = 740 \times L(60)^{\alpha}$, where $\alpha = 0.9427$ with an rms scatter of 0.08 dex .

It is important to point out that both of the corrections we have derived have the effect of *reducing* the change we find in $L(8 \mu\text{m})/L(\text{TIR})$. That is, using the relation derived by Yun et al. (2001) and applying the proposed change at high redshift (Kovács et al. 2006; Vlahakis et al. 2007) would have both resulted in a much larger change than reported. Therefore, we are taking a conservative approach.

We take radio flux densities for the galaxies observed by Yan et al. (2007) from Condon et al. (2003) as tabulated on the SSC World Wide Web site.²⁰ Two of the remaining galaxies are de-

tected at low significance as indicated by inspection of the gray-scale images for them (MIPS 429 and 506), so we enter rough values of 0.1 mJy for them. For MIPS 289, which is undetected, we assigned an upper limit of 0.1 mJy . We reject the source MIPS 8493 because of fringing in the VLA image at its position. The other four (MIPS 289, 15928, 16144, and 22530) are in regions where the radio images are clean. We correct their observed radio measurements to the rest frame, assuming a radio K -correction of $(1+z)^{0.7}$, since a typical slope for a starburst radio spectrum is -0.7 (Condon et al. 1991). We plot the resulting $L(\text{TIR})$ and $L(8 \mu\text{m})$ values in Figure 8 as circles (filled for the detections, open for the upper limit.)

4.4.4. $8 \mu\text{m}/\text{TIR}$ Ratio for the $z \sim 2$ Lensed Galaxies

The lensed galaxies described in this paper bridge the luminosity divide between the $z \sim 2$ extremely high-luminosity galaxies in the literature, and the comparison LIRG and ULIRG samples at $z \sim 0$ and $z = 0.85$. For the lensed sample we can determine which galaxies have mid-IR spectra dominated by star formation, using the aromatic feature strengths. For the subset of lensed galaxies with submillimeter measurements and good upper limits at $70 \mu\text{m}$, the SEDs are well enough constrained for a direct determination of $L(\text{TIR})$. Therefore, the lensed galaxies can cleanly test for evolution with redshift in the behavior of the $L(8 \mu\text{m})/L(\text{TIR})$ ratio.

We use the aromatic-to-mid-IR flux ratio introduced in § 4.2 to select galaxies whose mid-IR spectra are dominated by star formation. The ratio selects A2219a, A1689b, A2218a, and A1835a as all having strong aromatic emission; their ratios range from $(7-24) \times 10^{-11}$, with an average value of 16×10^{-11} . By comparison, our sample of local star-forming LIRGs and ULIRGs has a range in ratio of $(7-32) \times 10^{-11}$, with an average of 21×10^{-11} . That is, the relative aromatic feature strength is virtually the same for the two samples, with no evidence that AGNs have significantly augmented the $8 \mu\text{m}$ fluxes of our star-forming high- z subsample. The EW($6.2 \mu\text{m}$) criterion we used for the Yan et al. (2007) sample would have returned the same sample of star-forming lensed galaxies, except that A2219a lacks $6.2 \mu\text{m}$ coverage and thus would not have been selected.

We plot the four lensed galaxies as red triangles in Figure 8.

4.5. Comparison of $8 \mu\text{m}/\text{TIR}$ Ratio at High and Low Redshift

The $L(8 \mu\text{m})/L(\text{TIR})$ ratio is important for two reasons: (1) it is one of the most accessible measures of the physical behavior of infrared emission as a function of redshift; and (2) many studies rely on fluxes at observed-frame 15 or $24 \mu\text{m}$ to deduce $L(\text{TIR})$ and related properties of galaxies at $z \sim 1$ or 2 , and are thus relying on the behavior of $L(8 \mu\text{m})/L(\text{TIR})$. Figures 7 and 8 illustrate the important result that this ratio is well behaved for star-forming galaxies out to $z \sim 2.5$. In fact, this good behavior is a reasonably demanding test, since the various samples at $z \sim 2$ are subject to different selection effects and have $L(\text{TIR})$ derived in different ways, yet they all follow a single trend. To recapitulate the different methods: the lensed galaxies are selected largely on amplification factor, and we measured $L(\text{TIR})$ directly from template fits to rest frame mid and far-IR and submillimeter photometry. The four galaxies from Yan et al. (2007) were selected on the basis of colors and brightness at $24 \mu\text{m}$, and we estimated $L(\text{TIR})$ from the radio-infrared relation. The stacked galaxies from Papovich et al. (2007) were selected based on faint K -band and $24 \mu\text{m}$ detections; we estimated $L(\text{TIR})$ from rest $24 \mu\text{m}$ measurements. The stacked galaxies from Daddi et al. (2005) were originally selected from BzK colors indicating they are at $z \sim 2$

²⁰ See <http://ssc.spitzer.caltech.edu/fls/extragal/vla.html>.

and were stacked based on selections at a variety of wavelengths, and we used rest submillimeter measurements to estimate $L(\text{TIR})$.

Having noted the overall uniform behavior, and the general similarity of this behavior independent of redshift, we now evaluate the evidence from Figures 7 and 8 for modest evolution with redshift. Specifically, there is a tendency for the ratio $L(8\ \mu\text{m})/L(\text{TIR})$ to decrease with increasing luminosity for the local galaxies for $L > 10^{11}\ L_{\odot}$, but this tendency seems reduced for the galaxies at $z \sim 2$. In fact, the high-redshift galaxies appear to have SED behavior similar to normal local galaxies of lower luminosity in this regard.

We have made two tests of the significance of this difference. The first is based on the stacking results of Daddi et al. (2005) and Papovich et al. (2007). Both studies select the galaxies to stack on the basis of K -band detections, and therefore the mid- and far-infrared characteristics of the galaxies should be representative of the entire population of IR-active galaxies (i.e., there is no bias toward a specific $L(8\ \mu\text{m})/L(\text{TIR})$). To have an analogous set of data for the local galaxies, we have averaged their $L(8\ \mu\text{m})/L(\text{TIR})$ ratios in luminosity intervals of $10^{0.5}$; these averages are shown as the large blue triangles in Figure 7. We have then fitted the averaged data points, shown as a blue line. To test whether the high-redshift stacked results were compatible with the same fit, we carried out a second fit in which we included the three stacked points. For the galaxies above $10^{11}\ L_{\odot}$, we then compared χ^2 for the low-redshift galaxies only in the two cases. We found that this χ^2 is 4.5 times larger if the three stacked points at $z \sim 2$ are included in the fit; that is, there is a strong indication that a single functional fit cannot be used for both the local galaxies and for those stacked at $z \sim 2$.²¹

To determine whether a few anomalous local galaxies could affect this conclusion, we repeated the procedure but using a more robust average for the luminosity-binned ratios. For each bin with 10 or more items, we discarded the two highest and two lowest ratios and then averaged the remainder. Where there were less than 10 items, we retained the straight average. The highest luminosity bin, which is responsible for most of the change in slope of the relation between the two luminosities, has 10 objects and hence its average was recomputed. The results were virtually identical to those with the straight averages, as shown in Figure 7.

A second test was applied to the high-redshift galaxies measured individually, either by Yan et al. (2007) or by us. In the case of the Yan et al. sample, we expect a strong bias toward objects with a large ratio of $L(8\ \mu\text{m})/L(\text{TIR})$, because they selected relatively bright objects for spectroscopy, given the sensitivity limitations of the IRS. Our sample was selected on a combination of $24\ \mu\text{m}$ flux density and lensing amplification, so a bias is likely, although perhaps less severe.

To compare with these two studies, we selected a local sample with relatively high $8\ \mu\text{m}$ output by removing all the galaxies that are below the fit made to the averaged values (above). This sample is shown in Figure 8. It is noteworthy that the scatter in $L(8\ \mu\text{m})/L(\text{TIR})$ is small, with no outliers toward high $L(8\ \mu\text{m})$, so this is an appropriate comparison sample for cases where a bias is suspected. As before, we carried out two fits, one just to the local galaxies, and another to them plus the eight individually measured high-redshift ones. We found that χ^2 for the low-redshift sample with

$L > 10^{11}\ L_{\odot}$ increased by a factor of 2.5 if we included the high-redshift galaxies in the fit. Comparing Figure 8 with Figure 7, the discrepancy would be substantially larger if we had included all the local galaxies in our fits. This result confirms that found for the stacked samples, namely, that a single functional fit is not a good fit to the data over all redshifts.

Another comparison can be made with the sample of infrared galaxies at $z \sim 0.85$ from Marcillac et al. (2006). The results fall in a region occupied by both low- and high-redshift fits in Figure 8, which confirms the relatively modest change in $L(8\ \mu\text{m})/L(\text{TIR})$ with redshift, but (not surprisingly given the small difference between $z \sim 0$ and 2) does not definitively deviate from either fit. In making this comparison, we have assumed that the galaxies studied by Marcillac et al. (2006) have a bias toward bright $15\ \mu\text{m}$ observed flux densities. If instead we treat them as an unbiased sample and average them in luminosity intervals, they fall along (or slightly above) the fit for the $z \sim 2$ sample in Figure 7.

The tendency at high redshift for relatively strong $8\ \mu\text{m}$ emission could arise either through a contribution of an extra, AGN-powered component at $8\ \mu\text{m}$ (e.g., the extra $5.8\ \mu\text{m}$ continuum claimed by Sajina et al. 2007 and the mid-IR excess galaxies with hard X-ray emission discussed by Daddi et al. 2007), or to a change in the properties of the stellar-powered infrared-emitting regions. The AGN explanation runs counter to our selection of individual galaxies with strong aromatic features that correspond to equivalent widths typical of purely star-forming local galaxies. It is also counter to the precautions taken in the stacking analyses to exclude AGNs (e.g., Papovich et al. 2007). In addition, AGN contributions would come in a broad range, so it would be hard to explain the consistent behavior of the five high-redshift samples in Figures 7 and 8 if AGNs were causing the departure from the behavior from local star-forming galaxies. Finally, the behavior we see is similar to that found by Zheng et al. (2007) in a stacking analysis at $z \sim 0.7$; specifically, they report that the far-IR SEDs appear to resemble those of local galaxies of lower luminosity. Thus, while AGNs may boost $8\ \mu\text{m}$ emission in other sources, it is unlikely that AGN contamination is responsible for the behavior seen in Figures 7 and 8.

We conclude that star-forming galaxies have a mild trend toward increasing $L(8\ \mu\text{m})/L(\text{TIR})$ ratio with redshift that reaches up to a factor of ~ 2 at $z = 2$. In other words, $L(8\ \mu\text{m})$ does not turn over with increasing $L(\text{TIR})$ for $z \sim 2$ galaxies as it does at low redshift. This probably helps explain why some high-redshift galaxies have higher $6.2\ \mu\text{m}$ aromatic equivalent widths than are observed for local ULIRGs (Desai et al. 2007; Menéndez-Delmestre et al. 2007; Sajina et al. 2007), although direct comparisons cannot be made because there are not local galaxies with such extremely high luminosities. Because the lensed galaxies and stacked samples directly probe the $10^{11} - 10^{12.5}\ L_{\odot}$ range at $z \sim 2$, we can directly compare to $z \sim 0$ galaxies at similar luminosity rather than extrapolating.

That aromatic luminosities do not turn over with $L(\text{TIR})$ in high-redshift galaxies may be related to lower densities and optical depths of the infrared-emitting regions in the high- z galaxies. Lower optical depths would result in part if the metallicity were lower, resulting in a higher gas-to-dust ratio. Draine et al. (2007) show that the dust content of galaxies has a very steep dependence on their metallicity. The densities and optical depths of the star-forming regions would also be lower if the high luminosities are associated with star formation on large spatial scales in the high-redshift galaxies, as opposed to strongly concentrated, major-merger-associated nuclear star formation in local ULIRGs.

²¹ This analysis makes the assumption that low-luminosity galaxies have similar behavior at low and high redshift.

Such very dense regions in local ULIRGs appear to account for the scattering of points below the $z = 0$ trend line.

Use of local templates to estimate $L(\text{TIR})$ for high- z galaxies detected at $24\ \mu\text{m}$ has indicated the existence of very high luminosities at $z \sim 2$ and beyond (e.g., Yan et al. 2007.) The change in $L(8\ \mu\text{m})/L(\text{TIR})$ suggested here would reduce $L(\text{TIR})$ for a given $L(8\ \mu\text{m})$ by a factor of ~ 4 . As shown in the figure, in our sample there is only one galaxy selected to be dominated by star formation that has a $L(\text{TIR})$ significantly exceeding $10^{13}\ L_{\odot}$.²² A luminosity of $10^{13}\ L_{\odot}$ corresponds to a star formation rate of $\sim 2000\ M_{\odot}\ \text{yr}^{-1}$ if we use the local calibration (Kennicutt 1998), and is suggestively close to the theoretical maximum luminosity that a starburst is allowed by causality arguments—that star formation throughout a galaxy should not transpire faster than the dynamical timescale (Lehnert & Heckman 1996; Elmegreen 1999). (The actual maximum value depends on efficiency and dynamical mass.) Thus, our lower calibration for $L(8\ \mu\text{m})/L(\text{TIR})$ may remove the discrepancy between $L(\text{TIR})$ inferred for high-redshift galaxies and plausible upper limits on what a starburst can produce. Spatially resolved imaging of high-redshift starbursts (e.g., Nesvadba et al. 2007) can test whether they are in fact forming stars at the dynamical limit.

5. SUMMARY AND IMPLICATIONS

The $24\ \mu\text{m}$ photometric band of *Spitzer* efficiently detects galaxies out to $z \sim 3$, and is being used as an extinction-robust measure of star formation rates in the distant universe. We present new IRS spectra and MIPS photometry of gravitationally amplified galaxies, and use them in concert with archival *Chandra* images and published submillimeter photometry to investigate the nature of the very faint $24\ \mu\text{m}$ sources.

The aromatic feature flux ratios observed in our spectra agree well with those of star-forming template spectra at $z \sim 0$. This result implies that dust size and ionization distribution are not strongly evolving with redshift. Thus, for $z \gtrsim 1$ star-forming galaxies with little AGN contribution, the mid-IR spectra can be approximated by $z \sim 0$ star-forming templates.

²² This galaxy was selected as star-forming based on its $\text{EW}(6.2\ \mu\text{m})$ (Sajina et al. 2007), but its $\text{EW}(7.7\ \mu\text{m})$ is unusually low, which may indicate a substantial AGN contribution.

At $z \sim 2$, the $24\ \mu\text{m}$ data probe the $8\ \mu\text{m}$ rest wavelength region. We find that the ratio of rest-frame $8\ \mu\text{m}$ luminosity to total infrared luminosity is approximately the same as for local galaxies of similar $L(\text{TIR})$. However, the $z \sim 0$ relationship does not appear to hold perfectly. Combining data from the literature with this sample, we demonstrate that $z \approx 2$ star-forming galaxies have higher $L(8\ \mu\text{m})$ luminosities for a given $L(\text{TIR})$ than low-redshift analogs. Functionally, the break in the $L(8\ \mu\text{m})$ versus $L(\text{TIR})$ relationship occurs at a 10-times-higher $L(\text{TIR})$ at $z \sim 2$ than it does at $z \sim 0$. We speculate that this is due to lower optical depths in the IR-emitting regions of the high- z galaxies, perhaps because of lower metallicity or because star formation is more spatially extended than in local ULIRGs (e.g., across a disk rather than in a compact nucleus). This result has important implications for mid-IR-derived star formation rates at $z \sim 2$, and may indicate a maximum star-forming luminosity of a few $\times 10^{13}\ L_{\odot}$.

The results just summarized were found for sources where AGNs do not contribute substantially to the mid-IR spectra, as determined by the aromatic-to-mid-IR flux ratio. But 6 of the 19 mid-IR spectra in our extended sample *do* show a strong AGN. Including other AGN diagnostics (X-ray detections and optical emission lines), 8 of the 19 galaxies show evidence for nuclear activity. The prevalence of AGNs in this sample may contradict the frequent assumption that the $24\ \mu\text{m}$ flux density in $100\ \mu\text{Jy}$ sources at $1 < z < 3$ is dominated by star formation, with AGNs a modest contaminant.

We thank J. Donley for computing X-ray flux upper limits for the nondetections, and D. Sand for making available his reduction of the WFPC2 images. We thank D. Lutz and H. Teplitz for sending electronic versions of their published spectra. This work is based in part on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. Support for this work was provided in part by NASA through contract 1255094 issued to the University of Arizona by JPL/Caltech. J. Rigby was supported in part by NASA through the *Spitzer Space Telescope* Fellowship Program. This work has made use of archival data from the *Chandra X-Ray Observatory*, the *Hubble Space Telescope*, and the *Spitzer Space Telescope*. This research has made use of NASA's Astrophysics Data System Bibliographic Services and the NASA/IPAC Extragalactic Database.

APPENDIX

NOTES ON INDIVIDUAL SOURCES

A1. SOURCE A2390a

J. Richard obtained a Keck spectrum for this source, shown in Figure 9. The strong $\text{Ly}\alpha$ emission is redshifted by $z = 2.858$, and the line velocity spread ($\text{FWHM} \sim 945\ \text{km s}^{-1}$) indicates that the source is an AGN.

A2. SOURCE A2261a

This submillimeter galaxy (Chapman et al. 2002) lacks a spectroscopic redshift; submillimeter and radio-based photometric estimates place it at $z > 1.6$ Chapman et al. (2002) and $2.5 < z < 4$ (Aretxaga et al. 2003). It is the only source without a redshift in our sample. Our spectrum should be sensitive to the $11\ \mu\text{m}$ complex for $0.9 < z < 1.8$, and to the $7.7\ \mu\text{m}$ aromatic feature for $1.8 < z < 3.1$. Thus, for the aromatic features to fall outside the IRS spectral coverage, the source would have to have very low ($z < 1$) or very high ($z > 3.1$) redshift. If the source were at $z > 3.1$, it must be ~ 1000 times more luminous than M82 to have produced the observed $24\ \mu\text{m}$ flux density. It seems more probable that this source lies at $1 < z < 3$ and lacks aromatic emission.

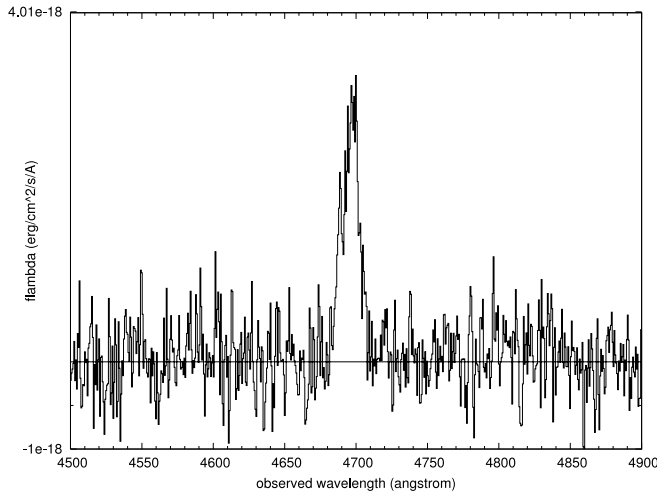


FIG. 9.—Keck/LRIS spectrum of A2390a. The strong feature at $\lambda = 4690 \text{ \AA}$ is presumably Ly α at $z = 2.858$.

REFERENCES

- Allamandola, L. J., Tielens, G. G. M., & Barker, J. R. 1989, *ApJS*, 71, 733
- Alonso-Herrero, A., Rieke, G. H., Rieke, M. J., Colina, L., Pérez-González, P. G., & Ryder, S. D. 2006a, *ApJ*, 640, 167
- Alonso-Herrero, A., et al. 2006b, *ApJ*, 640, 167
- Aretxaga, I., Hughes, D. H., Chapin, E. L., Gaztañaga, E., Dunlop, J. S., & Ivison, R. J. 2003, *MNRAS*, 342, 759
- Armus, L., et al. 2007, *ApJ*, 656, 148
- Barvainis, R., Antonucci, R., & Helou, G. 1999, *AJ*, 118, 645
- Brandl, B. R., et al. 2006, *ApJ*, 653, 1129
- Caputi, K. I., et al. 2006, *ApJ*, 637, 727
- Chapman, S. C., Scott, D., Borys, C., & Fahlman, G. G. 2002, *MNRAS*, 330, 92
- Chary, R., & Elbaz, D. 2001, *ApJ*, 556, 562
- Choi, P. I., et al. 2006, *ApJ*, 637, 227
- Condon, J. J., Cotton, W. D., Yin, Q. F., Shupe, D. L., Storrie-Lombardi, L. J., Helou, G., Soifer, B. T., & Werner, M. W. 2003, *AJ*, 125, 2411
- Condon, J. J., Huang, A.-P., Yin, Q. F., & Thuan, T. X. 1991, *ApJ*, 378, 65
- Cowie, L. L., Barger, A. J., & Kneib, J.-P. 2002, *AJ*, 123, 2197
- Daddi, E., et al. 2005, *ApJ*, 631, L13
- . 2007, *ApJ*, 660, 173
- Dale, D. A., & Helou, G. 2002, *ApJ*, 576, 159
- Dale, D. A., Smith, J. D., Armus, L., et al. 2006, *ApJ*, 646, 161
- Dale, D. A., et al. 2005, *ApJ*, 633, 857
- Desai, V., et al. 2007, *ApJ*, 669, 810
- Dole, H., et al. 2004, *ApJS*, 154, 93
- Donley, J. L., Rieke, G. H., Rigby, J. R., & Pérez-González, P. G. 2005, *ApJ*, 634, 169
- Downes, D., & Solomon, P. M. 2003, *ApJ*, 582, 37
- Draine, B. T., & Li, A. 2007, *ApJ*, 657, 810
- Draine, B. T., et al. 2007, *ApJ*, 663, 866
- Ebbels, T., Ellis, R., Kneib, J.-P., Leborgne, J.-F., Pello, R., Smail, I., & Sanahuja, B. 1998, *MNRAS*, 295, 75
- Egami, E., et al. 2006b, *ApJ*, 647, 922
- Elmegreen, B. G. 1999, *ApJ*, 517, 103
- Farrah, D., Afonso, J., Efstathiou, A., Rowan-Robinson, M., Fox, M., & Clements, D. 2003, *MNRAS*, 343, 585
- Frazer, D. T., Ivison, R. J., Scoville, N. Z., Yun, M., Evans, A. S., Smail, I., Blain, A. W., & Kneib, J.-P. 1998, *ApJ*, 506, L7
- Frazer, D. T., et al. 1999, *ApJ*, 514, L13
- Freedman, W. L., et al. 2001, *ApJ*, 553, 47
- Frye, B., & Broadhurst, T. 1998, *ApJ*, 499, L115
- Galliano, F. 2006, preprint (astro-ph/0610852)
- Gehrels, N. 1986, *ApJ*, 303, 336
- Gordon, K. D., et al. 2005, *PASP*, 117, 503
- Houck, J. R., et al. 2004, *ApJS*, 154, 18
- . 2005, *ApJ*, 622, L105
- Huang, J.-S., et al. 2004, *ApJS*, 154, 44
- Ivison, R. J., Smail, I., Barger, A. J., Kneib, J.-P., Blain, A. W., Owen, F. N., Kerr, T. H., & Cowie, L. L. 2000, *MNRAS*, 315, 209
- Kennicutt, R. C., Jr. 1998, *ARA&A*, 36, 189
- Kneib, J.-P., van der Werf, P. P., Kraiberg Knudsen, K., Smail, I., Blain, A., Frayer, D., Barnard, V., & Ivison, R. 2004, *MNRAS*, 349, 1211
- Knudsen, K. K. 2004, Ph.D. thesis, Leiden Univ.
- Knudsen, K. K., van der Werf, P. P., Kneib, J.-P. 2008, *MNRAS*, in press (arXiv: 0712.1904)
- Knudsen, K. K., et al. 2006, *MNRAS*, 368, 487
- Koekemoer, A., Fruchter, A., Hook, R., & Hack, W. 2002, in *HST Calibration Workshop, Hubble after the Installation of the ACS and the NICMOS Cooling System*, ed. S. Arribas, A. Koekemoer, & B. Whitmore (Baltimore: STScI), 337
- Kovács, A., Chapman, S. C., Dowell, C. D., Blain, A. W., Ivison, R. J., Smail, I., & Phillips, T. G. 2006, *ApJ*, 650, 592
- Lagache, G., et al. 2004, *ApJS*, 154, 112
- Le Floc'h, E., et al. 2005, *ApJ*, 632, 169
- . 2007, *ApJ*, 660, L65
- Lehnert, M. D., & Heckman, T. M. 1996, *ApJ*, 472, 546
- Lemonon, L., Pierre, M., Cesarsky, C. J., Elbaz, D., Pello, R., Soucail, G., & Vigroux, L. 1998, *A&A*, 334, L21
- Lutz, D., Spoon, H. W. W., Rigopoulou, D., Moorwood, A. F. M., & Genzel, R. 1998, *ApJ*, 505, L103
- Lutz, D., Valiante, E., Sturm, E., Genzel, R., Tacconi, L. J., Lehnert, M. D., Sternberg, A., & Baker, A. J. 2005, *ApJ*, 625, L83
- Marcillac, D., Elbaz, D., Chary, R. R., Dickinson, M. E., Galliano, F., & Morrison, G. 2006, *A&A*, 451, 57
- Marcillac, D., Rigby, J. R., Rieke, G. H., & Kelly, D. M. 2007, *ApJ*, 654, 825
- Menéndez-Delmestre, K., et al. 2007, *ApJ*, 655, L65
- Metcalfe, L., et al. 2003, *A&A*, 407, 791
- Motohara, K., et al. 2005, *AJ*, 129, 53
- Nesvadba, N. P. H., et al. 2007, *ApJ*, 657, 725
- Papovich, C., et al. 2007, *ApJ*, 668, 45
- Pello, R., Sanahuja, B., Le Borgne, J.-F., Soucail, G., & Mellier, Y. 1991, *ApJ*, 366, 405
- Pérez-González, P. G., et al. 2005, *ApJ*, 630, 82
- Reddy, N. A., Steidel, C. C., Fadda, D., Yan, L., Pettini, M., Shapley, A. E., Erb, D. K., & Adelberger, K. L. 2006, *ApJ*, 644, 792
- Rieke, G. H., et al. 2004, *ApJS*, 154, 25
- Rigby, J. R., & Rieke, G. H. 2004, *ApJ*, 606, 237
- Rigopoulou, D., Spoon, H. W. W., Genzel, R., Lutz, D., Moorwood, A. F. M., & Tran, Q. D. 1999, *AJ*, 118, 2625
- Sand, D. J., Treu, T., Ellis, R. S., & Smith, G. P. 2005, *ApJ*, 627, 32
- Sanders, D. B., Mazzarella, J. M., Kim, D.-C., Surace, J. A., & Soifer, B. T. 2003, *AJ*, 126, 1607
- Sanders, D. B., & Mirabel, I. F. 1996, *ARA&A*, 34, 749
- Sajina, A. Yan, L., Armus, L., Choi, P., Fadda, D., Helou, G., & Spoon, H. 2007, *ApJ*, 664, 713
- Seymour, N., et al. 2007, *MNRAS*, submitted
- Smail, I., Ivison, R. J., Blain, A. W., & Kneib, J.-P. 2002, *MNRAS*, 331, 495
- Smail, I., Ivison, R. J., Owen, F. N., Blain, A. W., & Kneib, J.-P. 2000, *ApJ*, 528, 612
- Smail, I., Smith, G. P., & Ivison, R. J. 2005, *ApJ*, 631, 121
- Smith, G. P., Kneib, J.-P., Smail, I., Mazzotta, P., Ebeling, H., & Czoske, O. 2005, *MNRAS*, 359, 417
- Smith, J. D. T., Draine, B. T., Dale, D. A., et al. 2007, *ApJ*, 656, 770
- Spergel, D. N., et al. 2003, *ApJS*, 148, 175
- Teplitz, H. I., et al. 2007, *ApJ*, 659, 941

- Thornley, M. D., Schreiber, N. M. F., Lutz, D., Genzel, R., Spoon, H. W. W.,
Kunze, D., & Sternberg, A. 2000, *ApJ*, 539, 641
Vlahakis, C., Eales, S., & Dunne, L. 2007, *MNRAS*, 379, 1042
Wang, Y. P. 2006, *ApJ*, 638, 138
Werner, M. W., et al. 2004, *ApJS*, 154, 1
Yan, L., et al. 2007, *ApJ*, 658, 778
Yun, M. S., Reddy, N. S., & Condon, J. J. 2001, *ApJ*, 554, 803
Zheng, X. Z., Dole, H., Bell, E. F., Le Floch, E., Rieke, G. H., Rix, H.-W., &
Schiminovich, D. 2007, *ApJ*, 670, 301