The star-formation rates of $1.5 < z < 2.5$ massive galaxies*

R. Nordon$^1$, D. Lutz$^1$, L. Shao$^1$, B. Magnelli$^1$, S. Berta$^1$, B. Altieri$^2$, P. Andreani$^3,11$, H. Aussel$^7$, A. Bongiovanni$^5,6$, A. Cava$^5,6$, J. Cepa$^5,6$, A. Cimatti$^3$, E. Daddi$^7$, H. Dominguez$^9$, D. Elbaz$^1$, N. M. Förster Schreiber$^1$, R. Genzel$^1$, A. Grazian$^{10}$, G. Magdis$^7$, R. Maiolino$^{10}$, A. M. Pérez García$^{5,6}$, A. Poglitsch$^1$, P. Popesso$^1$, F. Pozzi$^{10}$, L. Riguccini$^7$, G. Rodighiero$^8$, A. Santonge$^1$, M. Sanchez-Portal$^5$, P. Santini$^{10}$, E. Sturm$^1$, L. Tacconi$^3$, I. Valtchanov$^2$, M. Wetzstein$^1$, and E. Wieprecht$^1$

(Affiliations are available in the online edition)

Received 31 March 2010 / Accepted 28 April 2010

ABSTRACT

The star formation rate (SFR) is a key parameter in the study of galaxy evolution. The accuracy of SFR measurements at $z \sim 2$ has been questioned following a disagreement between observations and theoretical models. The latter predict SFRs at this redshift that are typically a factor 4 or more lower than the measurements. We present star-formation rates based on calorimetric measurements of the far-infrared (FIR) luminosities for massive $1.5 < z < 2.5$, normal star-forming galaxies (SFGs), which do not depend on extinction corrections and/or extrapolations of spectral energy distributions. The measurements are based on observations in GOODS-N with the Photodetector Array Camera and Spectrometer (PACS) onboard Herschel, as part of the PACS evolutionary probe (PEP) project, that resolve for the first time individual SFGs at these redshifts at FIR wavelengths. We compare FIR-based SFRs to the more commonly used 24 $\mu$m and UV SFRs. We find that SFRs from 24 $\mu$m alone are higher by a factor of $\sim$4–7.5 than the true SFRs. This overestimation depends on luminosity: gradually increasing for log $L(24 \mu m) > 12.2 L_\odot$. The SFGs and AGNs tend to exhibit the same $24 \mu m$ excess. The UV SFRs are in closer agreement with the FIR-based SFRs. Using a Calzetti UV extinction correction results in a mean excess of up to 0.3 dex and a scatter of 0.35 dex from the FIR SFRs. The previous UV SFRs are thus confirmed and the mean excess, while narrowing the gap, is insufficient to explain the discrepancy between the observed SFRs and simulation predictions.

Key words. galaxies: evolution – galaxies: starburst – galaxies: fundamental parameters – cosmology: observations – infrared: galaxies

1. Introduction

The evolution of massive galaxies around $z \sim 2$, when the cosmic star-formation rate density was at its peak, is currently the subject of intense study. A key parameter in constraining their nature and stellar-mass buildup is the instantaneous star-formation rate (SFR). While there are various methods for estimating SFRs (see Kennicutt 1998, for a review), it is notoriously difficult to constrain SFRs of $z \sim 2$ galaxies with an accuracy better than a factor of a few. One of the reasons is that calibrations between observed luminosities (rest-frame UV, mid-IR, and submm being the most accessible and widely used) and bolometric SFRs are based on local galaxies and their validity at high $z$ has not yet been firmly established.

Both UV (BX/BM) and optical/near-IR (BzK) selected samples of star-forming galaxies at $z \sim 2$ have been successful in producing large samples for galaxy evolution studies (Reddy et al. 2005; Pérez-González et al. 2005; Daddi et al. 2007; Santini et al. 2009). However, UV-based SFRs are very sensitive to dust obscuration and there are uncertainties about the applicable reddening laws and dust/source geometry at high $z$ (e.g., Reddy et al. 2010). Estimates from the 24 $\mu$m luminosities rely on template-based extrapolations to the total IR luminosity that are poorly constrained at high $z$ and may be affected by an AGN. Likewise, submm luminosities rely on assumptions about IR SEDs, which depend on poorly known dust temperatures, composition, and spatial distribution.

Cosmological simulations predict a tight correlation between stellar mass and SFR with a slope similar to that observed (e.g., Finlator et al. 2006), but with SFRs that are a factor $\sim$4 lower than those inferred for $z \sim 2$ massive galaxies from a combination of UV and 24 $\mu$m luminosities (Daddi et al. 2007; Wilkins et al. 2008; Damen et al. 2009). Comparisons between the derivative of the mass density distribution and direct measurements of SFR density also lead to a discrepancy, in particular at $z > 2$ (e.g. Pérez-González et al. 2008). Suggestions on how to resolve this discrepancy include changes to the adopted initial mass functions (IMF, Davé 2008) and star formation efficiency (Khochfar & Silk 2009), but these cannot be properly tested given the uncertainties associated with the common SFR indicators themselves.

The photodetector array camera and spectrometer (PACS, Poglitsch et al. 2010) onboard the Herschel Space Observatory (Pilbratt et al. 2010) offers a new opportunity to infer the SFR of $z \sim 2$ galaxies from their far infrared (FIR) emission. PACS observes at 160 $\mu$m (45–65 $\mu$m rest frame for the sample used here), close to the peak of the emission from dust heated by young stars, away from AGN-heated dust emission and with unprecedented spatial resolution that reduces confusion noise. It
allows the most reliable and least biased SFR measurements to date, which avoid large SED extrapolation errors and avoid the uncertainties of attenuation corrections. In this letter, we report on the first measurements of FIR-based calorimetric SFRs of optically selected galaxies at $1.5 < z < 2.5$ and compare them to results from previous UV and 24 $\mu$m measurements. We adopt $\Omega_m = 0.3$, $\Omega_{\Lambda} = 0.7$, and $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$ cosmology throughout this letter.

## 2. Data, sample and SFR(160 $\mu$m)

The PACS data consists entirely of PACS evolutionary probe (PEP) guaranteed-time observations in the GOODS-N field, taken during the science demonstration phase of the Herschel mission. For details of the observations and data reduction process, we refer to Berta et al. (2010). Fluxes were extracted using Spitzer-MIPS 24 $\mu$m sources with $3\sigma$ detections or better as priors (Magnelli et al. 2009). At the current PACS depth (5.7 mJy, 3$\sigma$, at 160 $\mu$m), testing showed that practically all PACS sources are detected at 24 $\mu$m. Hence, using 24 $\mu$m sources as priors allows a complete PACS source extraction. The priors are matched to bviz (ACS), JHK (FLAMINGO), and Spitzer-IRAC photometry. Spectroscopic redshifts are from Barger et al. (2008) and photometric redshifts were derived using the code EAZY (Brammer et al. 2008). The Chandra 2Ms catalog (Alexander et al. 2003) is matched to the PACS sources and is used to flag X-ray sources. Using the Ranalli et al. (2003) $L_\nu/L_{\text{FIR}}$ relation for SFG, the expected X-ray flux of a log$(L_{\text{FIR}}/L_\odot) = 12$ galaxy at $z = 2$ is almost an order of magnitude lower than the $1.4 \times 10^{-16}$ erg cm$^{-2}$ s$^{-1}$ hard-band (observed 2–8 keV) sensitivity limit of the 2 Ms catalog. We therefore treat every X-ray source with a hard-band detection as an AGN. One source without an X-ray detection but which exhibits a clear power-law SED in the IRAC bands, was also flagged as an AGN.

In this study, we are interested in $1.5 < z < 2.5$ range and require redshift information. Hence, only sources for which we have spectroscopic redshifts, and/or were successfully fitted by EAZY for a photometric redshift are considered in the following selections. In addition, we require that in a 10” radius around a 24 $\mu$m prior there is no more than one other 24 $\mu$m source and it may have no more than 50% of the prior flux. This condition ensures that the sample does not experience significant flux confusion between MIR priors and/or FIR fluxes due to neighbors. These criteria result in a sample of 23 detected sources in total, 17 classified as SFGs, and 4 of which have a spectroscopic redshift. 306 24 $\mu$m priors without a PACS detection are available for stacking analysis (see below).

We calculate SFRs by assuming that they are proportional to the integrated infrared luminosity (LIR, typically 8–1000 $\mu$mKennicutt 1998). We convert PACS 160 $\mu$m fluxes into LIR by fitting the flux using the Chary & Elbaz (2001, hereafter: CE01) SED library, for which a given redshift and band flux, a unique solution to LIR exists. The 100 $\mu$m filter is not used because at $z \sim 2$, emission at rest-frame wavelengths may already be affected by hot dust and suffer the same systematic effects as the 24 $\mu$m based SFR. In addition, at these redshifts 100 $\mu$m is on the same side of the SED peak and too close to 160 $\mu$m to provide meaningful temperature constraints, without longer wavelength data. Combined PACS and SPIRE sub-mm data show that fitting CE01 SEDs to the 160 $\mu$m flux provides a reliable estimate of the total 8–1000 $\mu$m luminosity (LIR) of the large-grain dust for $z \sim 2$ galaxies (Elbaz et al. 2010).

To explore the relationship between direct FIR luminosities, 24 $\mu$m extrapolated luminosities, and UV-SFRs to lower luminosities and SFRs, we also perform a stacking analysis. We stack the 160 $\mu$m image centered on the prior positions and measure a mean flux for the stack. The mean luminosity of the stack, LIR(160 $\mu$m)stack, is defined as the luminosity that will result in the 160 $\mu$m stacked flux, for the redshift distribution of the sources in the stack and the adopted CE01 SEDs. In this way, both k-corrections and luminosity distances are accounted for without bias. The error in the mean luminosity is calculated using a bootstrap method. The same number of sources are resampled with replacements from the original stack to produce a new stacked image. The new mean luminosity is then calculated. The error is finally calculated from the distribution of luminosities in the repeated resampling.

## 3. 24 $\mu$m based SFR

Spitzer-MIPS 24 $\mu$m fluxes are converted to LIR by fitting the 24 $\mu$m flux to CE01 SEDs using the same method we use for the PACS 160 $\mu$m fluxes (Sect. 2). This SED library is very commonly used in 24 $\mu$m (and other) studies where LIR is often derived from a single photometric point. Many other SED libraries...
exist, but the $\nu L_\alpha(8 \mu m)/L_{IR}$ ratio that is relevant to this study is typically calibrated to match local galaxy observations and vary by factors smaller than those discussed below. At $z \sim 2$, the rest-frame wavelengths are shorter than 10 $\mu m$ and 24 $\mu m$ fluxes probe the IR emission at the edge of the relevant range, far from the SED peak, making them sensitive to extrapolation errors. In addition, at these wavelengths polycyclic aromatic hydrocarbon (PAH) emissions contribute significantly and the ratio of their fluxes to LIR may create significant scatter. PACS 160 $\mu m$ fluxes probe the emission much closer to the SED peak and are unaffected by the above issues.

In Fig. 1, we plot LIR as derived from 160 $\mu m$ versus that derived from 24 $\mu m$ for 1.5 < $z$ < 2.5. Individual sources are detected starting at LIR(160 $\mu m$) > $6 \times 10^{15}$ $L_\odot$ at redshifts close to $z \sim 1.5$. The 24 $\mu m$ sources without a PACS detection (and no X-ray counterpart) are stacked in luminosity bins as indicated by the bars in the figure. Stacked luminosities represent the mean luminosity as defined in Sect. 2 and the stack error bars indicate the error in the mean LIR(160 $\mu m$) of the stack. Since we stack in luminosity and not flux bins, the typical redshifts of sources in a given stack are higher than those of detections with similar LIR(160 $\mu m$), which is one reason why not all sources in the luminosity bin are detected. It is clear that while there is some scatter, for most sources LIR(24 $\mu m$) is higher than LIR(160 $\mu m$).

Figure 2 plots the LIR(24 $\mu m$) to LIR(160 $\mu m$) ratio as a function of LIR(24 $\mu m$). The detections appear to show a linear trend, but this is mostly due to a selection effect. The combined mean for the detection and non-detections (detected in stacks) is much higher and the individual detections represent a fraction of the distribution for which at a given LIR(24 $\mu m$), LIR(160 $\mu m$) is high enough to be detected. There is a weak 24 $\mu m$ luminosity dependence and the mean of the excess increases with LIR(24 $\mu m$), at least for log LIR(24 $\mu m$) > 12.2. Below this luminosity, it is harder to determine the slope, but the stack indicates that it may flatten out. The dashed and dotted lines in Fig. 2 are ad hoc fits to the binned combined-means slope above and a constant offset below LIR(24) = 12.2

$$0.38 \log LIR(24) - 4.0 ; \log LIR(24) > 12.2$$

$$0.6 ; \log LIR(24) \leq 12.2$$

(1)

Though a large fraction of the sample relies on photometric redshifts, errors in redshifts will move the points in LIR(24 $\mu m$), but will have a much smaller effect on the LIR ratio. The red arrows at the top-left of Fig. 2 demonstrate how the typical phot-z errors, propagated into the quantities plotted in the figure, will affect the values. The photo-z induced dispersion is directed along the diagonal trend of detected sources and probably contributes to this pattern with the selection effects.

X-ray AGNs do not appear to differ from the rest of the sample in Figs. 1 and 2, except perhaps for the one source with the highest luminosity. This source also exhibits a very clear power-law SED in the mid-IR. For sources completely dominated by the AGN, fitting a CE01 SED to the 24 $\mu m$ emission is clearly inappropriate. Even though we can only produce one AGN stack that results in a good detection (Fig. 1), the mean luminosity is similar to that of the SFG stacks. This can be interpreted in two ways: either the flux at 6.8 < $\lambda$ < 9.6 $\mu m$ rest frame is dominated by the starburst emission at all the luminosities considered here, or the flux excess in all sources is due to an AGN, which is obscured in most. The latter claim and the relatively tight LIR(24 $\mu m$)/LIR(160 $\mu m$) ratio would imply a tight relation between the AGN luminosity and the galaxy’s starburst component FIR luminosity, which is not observed (Lutz et al. 2010; Shao et al. 2010). We conclude that while a hidden AGN may contribute to the scatter, it is not likely to be the main cause of the general 24 $\mu m$ excess. Enhanced emission (relative to local galaxy SEDs) in PAH features between 6.2–8.6 $\mu m$, that enter the 24 $\mu m$ filter at $z = 1.5$ is a more plausible explanation.

These new results corroborate the main findings of Murphy et al. (2009), based on Spitzer 16, 24, 70 $\mu m$, and SCUBA 850 $\mu m$ observations. For $z \sim 2$ galaxies, they find that SFRs using only 24 $\mu m$ are overestimated by a factor of 5 and conclude that a hidden AGN can only account for a fraction of this excess. Papovich et al. (2007), using Spitzer 24, 70 and 160 $\mu m$ stacked fluxes also find that LIR(24 $\mu m$) is in excess. At the high end, they find 24 $\mu m$ overestimates LIR by a factor of 2–10. They too conclude that X-ray sources have a 24 $\mu m$ excess similar to SFGs without X-ray counterparts.

4. UV-based SFR

The UV continuum at wavelengths longer than the Lyman edge can be used as a SFR indicator. The UV emission is dominated by young, massive, and short-lived stars and a SFR that is required to sustain their population can be calculated. The SFR(UV) is sensitive to the attenuation correction used, which varies significantly in the UV between different extinction curves, all derived from local galaxies. Testing the accuracy of SFR(UV) is essentially a test of the attenuation correction. Here we use $B$ and $z$ band photometry (probing rest frame ~1500 $\AA$ and ~3500 $\AA$ at $z \sim 2$) to estimate their SFR and dust...
attenuation. For the optical extinction, we apply the calibration given by Daddi et al. (2004)
\[ E(B-V) = 0.25(B-z + 0.1)_{AB}. \]  
(2)
The effective attenuation at 1500 Å is obtained using the Calzetti et al. (2000) extinction law \( A_{1500} = 10E(B-V), \) which was derived for local starburst galaxies.

The BzK sample is selected according to the Daddi et al. (2004) color criterion. We require at least 3\( \sigma \) detection in all 3 bands, a cut at \( K_s < 22 \) AB, and a redshift (spectroscopic or photometric) 1.5 < \( z < 2.5. \) PACS 160 \( \mu m \) fluxes were extracted using 24 \( \mu m \) priors, which implies a 24 \( \mu m \) detection for this selection. It also means that the BzK sample is a subsample of that used in Sect. 3. The sample includes 15 sources detected at 160 \( \mu m, \) 11 classified as SFG. Seventy-five sources without 160 \( \mu m \) detections are used for stacking. Strictly speaking, the requirement for a 24 \( \mu m \) prior is irrelevant to both SFR(UV) and SFR(160) and may even bias the selection against sources with lower obscuration. To improve the stacked signal, we also include BzK sources without a 24 \( \mu m \) detection, but verify that introducing this requirement gives consistent results. In converting from luminosities to SFRs, we follow the Kennicutt (1998) relations, but lower the SFR by a factor 1.7 to convert from a Salpeter initial mass function (IMF) to that of Chabrier (2003), though for the results discussed here the choice of IMF makes little difference:
\[ \text{SFR}_{\text{LIR}} \left[ M_\odot/\text{yr}^{-1}\right] = \frac{\text{L}_{\text{IR}}}{9.85 \times 10^9 L_\odot}; \]  
(3)
\[ \text{SFR}_{\text{UV}} \left[ M_\odot/\text{yr}^{-1}\right] = \frac{L_{\nu}(1500 \, \text{Å})}{1.5 \times 10^{28} \text{erg s}^{-1}\text{Hz}^{-1}}. \]  
(4)

In Fig. 3, we plot the combined SFR from LIR(160 \( \mu m \)) and the observed (uncorrected) UV flux versus the attenuation-corrected UV SFR, SFR(UV). The LIR is assumed to be caused by the reprocessing of radiation from young stars by the dust and the addition of the SFR from the observed UV luminosity accounts for the escaped part of the UV radiation. In our sample, the latter makes only a small contribution with a mean of 2.4% and maximum of 8% of the total SFR for PACS sources. Our detection limit with PACS is at about 70 \( M_\odot \text{ yr}^{-1}, \) below which we can only study sources by means of stacking. However, for a cut level \( K_s < 22 \) the small number of BzK galaxies below the PACS threshold limits the stacking possibilities. The detections exhibit good agreement between SFR(UV) and the combined SFR(LIR+UV). The mean log-ratio of the detections is \(-0.02 \text{ dex and standard deviation of 0.25 dex}. \) We stack PACS non-detections by SFR(UV) bins. For the low SFR stack, the two SFR estimators are in very good agreement, while for the other, the value of SFR(UV) is slightly too high. The overall mean log-ratio of stacks and detections is 0.3 dex and the standard deviation is 0.35 dex. We therefore conclude that when using the Calzetti UV extinction law for SFR(UV) \( \geq 40 M_\odot \text{ yr}^{-1} \) the SFR is overestimated by a mean factor \( \sim 2 \) with a scatter of a similar magnitude.

The scatter in Fig. 3 exhibits a systematic pattern. For SFR(UV) \( > 100 M_\odot \text{ yr}^{-1}, \) this estimator tends to overpredict the SFR of the PACS detections by a factor \( \sim 1.5 \) and below this rate to underpredict by roughly the same factor. This trend is also seen in Fig 4, where the SFR indicator ratio is plotted vs. the \( K_s \) magnitude. Since the pattern appears to be also a function of \( K_s \) magnitude (and of both \( B \) and \( Z \) mag to a lesser degree), it is unlikely to be caused by phot-\( \gamma \) errors. It is not preserved when plotting against either redshift or extinction. Both stacks for the \( K_s \) magnitude bins lie above the zero-line, indicating a spread in the full population and shift the overall mean upwards. The difference between stacks and detections is likely a selection effect: for a given SFR(UV) (or \( K_s \)), one will generally tend to detect in 160 \( \mu m \) galaxies for which the attenuation is underpredicted. However, the clear dependency on both SFR and \( K_s \) magnitude and the observed trend for the stacks might suggest there is more to it. The geometry and metallicity of the galaxies may play a role in this. Reddy et al. (2010) suggested that some young galaxies may have a different UV reddening law, although these galaxies represent only 13% of their sample. The current sample is too small to draw conclusions about these effects.

Four X-ray AGNs that are selected by BzK were detected with PACS at 160 \( \mu m \). Three are indistinguishable from the SFG in Figs. 3 and 4 and both their UV and FIR emission must be dominated by the star formation. The fourth (log-ratio = -0.75) is an extremely bright AGN and both SFR indicators are probably not applicable to this case. This is the only 160 \( \mu m \) source that does not appear in Figs. 1 and 2 as it is too bright in 24 \( \mu m \) to be fit by CE01 SEDs. There were not enough non-detected AGN-BzKs to produce a meaningful stack. The presence of an AGN, except in the extreme cases, does not significantly affect the SFR(UV).

Acknowledgements. PACS has been developed by a consortium of institutes led by MPE (Germany) and including UVIE (Austria); KU Leuven, CSL, IMEC (Belgium); CEA, LAM (France); MPIA (Germany); INAF-OA/OAP/OAT, LENS, SISSA (Italy); IAC (Spain). This development has been supported by the funding agencies BMVIT (Austria), ESA-PRODEX (Belgium), CEA/CNES (France), DLR (Germany), ASI/INAF (Italy), and CICYT/MCYT (Spain).

References

Page 5 is available in the electronic edition of the journal at http://www.aanda.org