Draft version December 15, 2020
Preprint typeset using LaTeX style emulateapj v. 12/16/11

DRAFT!

TRACING MOLECULAR GAS MASS IN HIGH REDSHIFT ($z \simeq 6$) GALAXIES WITH [CII]

D. Vizgan[1], T.R. Greve[2], K. P. Olsen[3], A. Zanella[4], Georgios Magdis[5], D. Narayanan[6], G. Popping[7], R. Davè[8], and Francesco Valentino[9]

Draft version December 15, 2020

ABSTRACT

The canonical molecular gas mass tracer for galaxies in the local universe, carbon monoxide (^{12}CO) may have less utility at redshifts $z > 6$ owing to difficulties in accessing the ground state transition with current facilities, as well as potential drowning out by the warmer cosmic microwave background at these epochs. In this paper, we investigate the utility of the fine-structure line of singly-ionized carbon at 158 μm ([CII]) as a tracer of molecular gas mass and analyzes the relationship between molecular gas mass and [CII] line luminosity within 9,356 normal $z \simeq 6$ star-forming galaxies from the SIMBA simulations, with line emission modeled by SÍGAME (Simulator of Galaxy Millimeter/Submillimeter Emission). Using our simulations and observational samples from the literature, we derive an equation which describes the relationship of [CII]-to-H_2 . We analyze secondary effects of star formation rate and gas-phase metallicity and apply this equation to 19 galaxies at z > 1 to estimate their molecular gas mass.

Keywords: galaxies: formation — galaxies: evolution — galaxies: ISM — galaxies: high-redshift

1. INTRODUCTION

Molecular gas, H_2, and star formation are observed to be tightly correlated (e.g. Kennicutt 1998; Kennicutt and Evans 2012; Bolatto *et al.* 2013) in our Galaxy as well as in local and high-redshift galaxies. With the possible exception of star-formation in extremely metal-poor environments where stars may form directly from atomic gas inflows (Krumholz 2012; Michałowski *et al.* 2015), molecular gas overwhelmingly provides the fuel for the star-formation in galaxies. Molecular gas is, therefore, fundamental to galaxy formation and evolution and drives many of the important galaxy scaling relations (REFs).

Unfortunately, H_2 is practically invisible under the typical conditions prevailing in the molecular interstellar medium (ISM). Due to its lack of a permanent dipole moment, H_2 only emits from warm (> 100 K) gas (Fumagalli 2012), and its two lowest rotational transitions have upper energy levels of 510K and 1015 K. The low ($< 10-50$ K) temperatures typical of the giant molecular clouds (Cox 2005), which make up the bulk of the star formation fuel, are unable to excite H_2, and are therefore H_2-"dark" to observers. Our inability to directly trace the distribution and amount of molecular gas in galaxies is a long-standing challenge for extragalactic studies (REFS), and has led to the development of a number of indirect H_2-tracing methods.

The J=1→0 rotational transition line of carbon monoxide (CO) is the most rigorously calibrated and most regularly used tracer of molecular gas mass (e.g., Dickman *et al.* 1986; Solomon and Barrett 1991; Downes and Solomon 1998; Bolatto *et al.* 2013). The method relies on the fact that the lowest rotational lines of CO are easily excited in the molecular ISM conditions, are relatively bright, and that CO is present throughout the molecular phase. It makes use of a CO-to-H_2 conversion factor, $\alpha_{\rm CO}$, which converts the observed CO(1-0) line luminosity into an estimate of the underlying molecular gas mass (e.g., Bolatto *et al.* 2013). A downside of using CO(1-0) as a tracer of H_2 is the uncertainty associated with $\alpha_{\rm CO}$. The conversion factor is calibrated for giant molecular clouds (GMCs) in our Galaxy (REFS), and a galaxy-averaged value for $\alpha_{\rm CO}$ has been determined for local IR-luminous galaxies and found to be $\sim 4\times$ lower than the Galactic value (Downes and Solomon 1998). Beyond that, however, and especially for high-z galaxies, we do not have strong constraints on $\alpha_{\rm CO}$ (cf. REFS). Both observations and simulations suggest that $\alpha_{\rm CO}$ depends on the physical properties of the molecular gas, in particular, the temperature, velocity dispersion, and metallicity. In regions of high gas temperature and velocity dispersion, the optically thick CO(1-0) emission line emits more luminosity per unit H_2 column density, which leads to lower conversion factors as seen in the local IR-luminous galaxies (Downes and Solomon 1998; Narayanan *et al.* 2012). In low-metallicity galaxies, $\alpha_{\rm CO}$ scales inversely with metallicity since in such environments CO is photo-dissociated in a larger fraction of the molecular gas, which is thus referred to as 'CO-dark' gas. (e.g., Wilson 1995; Arimoto *et al.* 1996; Wolfire *et al.* 2010; Narayanan *et al.* 2012; Genzel *et al.* 2012). In galaxies undergoing intense star-formation, and where the cosmic-ray ionisation rates are high ($\sim 100-1000\times$ the Galactic value), CO can be destroyed deep within molecular clouds, which will degrade the H_2-tracing capabilities of CO and in turn imply a higher $\alpha_{\rm CO}$ value

[1] Department of Astronomy, Wesleyan University, Middletown, CT 06457, USA
[2] Department of Space, Danish Technical University, 2800 Kgs. Lyngby, Denmark
[3] Department of Astronomy and Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[4] Istituto Nazionale di Astrofisica, 5, 35122 Padova PD, Italy
[5]
[5] Department of Astronomy, University of Florida, Gainesville, Florida 32611
[6] European Southern Observatory, 85478 Garching bei München, Germany
[7] Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, U.K.
[8]

(Bisbas *et al.* 2015, 2017). All of the above are conditions we are likely to encounter in galaxies in the early Universe. Furthermore, at high redshifts it quickly becomes expensive in terms of telescope time to detect CO(1-0). A more fundamental problem, however, is that as the redshift increases, the cosmic microwave background (CMB) gradually becomes brighter than CO(1-0) emission from cold molecular gas; for example da Cunha *et al.* (2013) showed that at high redshifts, the CMB can drastically affect both CO(1-0) and CO(2-1) line fluxes and neglecting these effects can lead to significant underestimations of the derived luminosities, and ultimately, the derived masses (see also Tunnard and Greve 2016; Zhang *et al.* 2016).

The fine-structure transitions of neutral carbon – CI($^3P_1$–$^3P_0$) and CI($^3P_2$–$^3P_1$) (hereafter simply CI(1−0) and (2−1)) – were first proposed as a tracer of molecular gas in galaxies over 15 years ago (Gerin and Phillips 2000; Papadopoulos and Greve 2004). CI is abundant ($\sim 10^{-5}$ relative to H_2), yet the lines have low to moderate optical depths ($\tau_{\rm CI} \sim 0.1 - 1$). Both lines are easily excited in typical molecular gas conditions ($n_{\rm cr} \sim 500\,{\rm cm}^{-3}$, $h\nu/k_{\rm B} \sim 23.6\,{\rm K}$ for CI(1−0) and $n_{\rm cr} \sim 1000\,{\rm cm}^{-3}$, $h\nu/k_{\rm B} \sim 62.4\,{\rm K}$ for CI(2−1)) which, together with its simple 3-level partition function, ought to make it a straightforward tracer of H_2. CI surveys of molecular clouds in our galaxy (Ikeda *et al.* 2002) showed a remarkable spatial and kinematical similarity in the CI and CO emission, suggesting that CI does indeed trace the bulk H_2 reservoir. Recent state-of-the-art simulations of molecular clouds have shown that cloud inhomogeneity (resulting in deeper UV penetration and thus an increase in the CI layer; Offner *et al.* (2014)) and turbulence (smoothing any initial [CI]/[CO] abundance gradient; Glover *et al.* (2015)) are responsible for the observed CI-CO concomitant. Offner *et al.* (2014) further finds that over a wide range of column densities CI is a comparable or superior tracer of the molecular gas, especially when an intense UV field is impinging on it.

The far-IR/sub-millimeter dust continuum emission has been used as an indirect tracer of the total gas content (molecular and atomic) of the Galactic ISM Hildebrand (1983), of nearby galaxies Guelin *et al.* (1993, 1995); Eales *et al.* (2012), and of high-z galaxies Magdis *et al.* (2012); Scoville *et al.* (2014, 2016); Kaasinen *et al.* (2019). One approach has been to derive the dust mass from careful modelling of the far-IR/sub-millimeter spectral energy distribution (SED), and apply a dust-to-gas mass ratio. The problems of this method is not only the uncertainties associated with the SED fitting, but a number of other factors, such as the dust temperature, the grain emissivity, and the dust-to-gas mass ratio. The latter depends on metallicity, which implies a redshift dependence via the mass-metallicity relation. Furthermore, this method requires flux measurements over a range of far-IR/(sub-)mm bands. To circumvent these problems, a direct calibration between the gas mass and the dust continuum emission at a single, broad-band (sub-)mm wavelength has been sought (e.g., Scoville *et al.* 2014). Such calibrations have been established for optically thin continuum emission along the Rayleigh-Jeans tail, e.g., at 850 μm, and successfully applied to both local and high-z galaxies (Scoville *et al.* 2014, 2016). This method has the advantage over CO and [CI] that is is easier to detect the continuum emission in a single broad band than it is to detect a single line. At redshifts $\gtrsim 4-5$, however, this method too runs into the limitations set by the increasing CMB temperature, which will not only thermally couple to the cold dust but also 'drown out' a significant fraction of the dust continuum emission (e.g. da Cunha *et al.* 2013; Zhang *et al.* 2016).

Since none of the classical tracers of molecular gas in galaxies are completely reliable, especially in the most distant galaxies, where the CO and [CI] lines become extremely faint and the CMB presents an insurmountable obstacle, there has been a growing effort, from both the observation and simulation sides, to explore the 158 μm (1900.5 GHz) fine-structure transition of singly-ionized carbon ([CII]) as a molecular gas mass tracer (Accurso *et al.* 2017; Hughes *et al.* 2017; Zanella *et al.* 2018; Dessauges-Zavadsky *et al.* 2020; Madden *et al.* 2020)

The [CII]158μm line is one of the strongest cooling lines of the ISM, and can carry up to a few percent of the total far-IR energy emitted from galaxies (e.g., Malhotra *et al.* 1997). Carbon is the fourth most abundant element, and has an ionisation potential of 11.3 eV, lower than that of neutral hydrogen. C^+, therefore, permeates much of the ISM. It is found in photodissociation regions (PDRs), diffuse ionised and atomic regions, and even in molecular gas (e.g., Kaufman *et al.* 1999). Because the [CII] line arises in most phases of the ISM, it is important to be able to disentangle the different contributions in order to correctly interpret galaxy-wide, integrated [CII] observations. How much of the [CII] emission comes from the molecular phase, and to what extent the line can be used as a molecular gas tracer is of particular interest. Observations of [CII] in the Milky Way also show that ∼75% of the [CII] emission comes from dense PDRs and CO-dark H_2 gas (Pineda *et al.* 2014).

Because of its association with PDRs, [CII] has been seen as a tracer of the star-formation rate (SFR) (e.g. De Looze *et al.* 2014; Herrera-Camus *et al.* 2015; Magdis *et al.* 2014). At low redshift, metal-poor dwarf galaxies and star-forming galaxies display slightly offset but tight log-linear relations between the [CII] luminosity and the sum of total obscured and unobscured SFR (De Looze *et al.* 2014). However, compact starbursts with high SFR surface densities and strong UV radiation fields tend to exhibit a "[CII]/ FIR deficit" by as much as an order of magnitude compared to the local [CII]-SFR relations (e.g. Magdis *et al.* 2014; Díaz-Santos *et al.* 2013; Narayanan and Krumholz 2017). The [CII]/ FIR deficit deficit is typically observed for local ultra luminous IR galaxies (ULIRGS; $L_{\rm IR} \geq 10^{12}\,{\rm L}_\odot$) and high-z starbursts/mergers, while normal star-forming galaxies at high-redshifts are in line with the local [CII]-SFR relationship (REFs). By extension, this explains why [CII] traces molecular gas, since the Kennicutt-Schmidt law shows a strong linear relationship between SFR and molecular gas mass. However, this relationship is not unique; from observations, normal galaxies seem to differ from starbursts/mergers in that the latter seem to deplete their gas reservoirs around 10 times faster. As pointed out by Zanella *et al.* (2018), if one assumes [CII] traces molecular gas mass, the [CII]/IR relationship be-

comes M_{mol}/SFR, which is the gas depletion timescale of a galaxy. In this case, the observed "FIR deficit" of starbursts reflects their 10× shorter gas depletion time-scales compared to normal star-forming galaxies (REFS).

Theoretically, a strong connection between [CII] luminosity and molecular gas mass is also expected from the fact that the low critical density of carbon (11.3 eV) allows for the excitation of [CII] in dense,molecular regions by collisions with H_2 molecules. Indeed, simulations of high-redshift normal galaxies (e.g. Vallini *et al.* 2015; Olsen *et al.* 2017) found that most of the [CII] emission comes from dense PDRs associated with regions of molecular gas.

The use of [CII] as a mass tracer has been implemented on a sample of 10 main-sequence galaxies at $z \sim 2$ by Zanella *et al.* (2018) (henceforth; Z18). The authors combine Band 9 ALMA (Atacama Large Millimeter/submillimeter Array) observations of [CII] emission with existing multiwavelength observations to derive a [CII] mass-to-luminosity conversion factor $\alpha = 31 M_\odot/L_\odot$. Molecular gas mass was derived from the integrated Schmidt-Kennicutt relation using SFR derived from pixel-by-pixel SED fitting. The resulting conversion factor was found to be mostly invariant with galaxies' redshift, depletion time, and gas phase metallicity. At lower redshifts [CII] has also been demonstrated to trace molecular gas mass (Madden *et al.* 2020). The aim of this research is to demonstrate that for lower-mass and lower-metallicity galaxies at earlier epochs, i.e. $z > 6$ [CII] is able to trace molecular gas mass.

Repeating the experiment on a much larger sample of simulated high-redshift galaxies, we aim to derive a conversion factor for molecular gas mass of high-redshift galaxies using [CII] line luminosity, as has been done in the nearby universe ($z \sim 0.03$-0.2) for CO (Hughes *et al.* 2017), and as done for galaxies in the local Herschel Dwarf Galaxy Survey for [CII] (see Madden *et al.* 2020).

In this paper we present new SÍGAME (Simulator of Galaxy Millimeter/Submillieter Emission) simulations of [CII] line emission from several thousands of $z \simeq 6$ galaxies in order to examine whether the line can be used to trace the gas content in normal star-forming galaxies at cosmic dawn. Section 2 of the paper describes the observation samples that we compare our simulations with in this work; Section 3 describes the simulation sample and the post-processing done by SÍGAME. Section 4 compares the integrated physical properties of the simulation sample in the context of observations. In Section 5 we present our [CII] simulation results, derive a [CII]-to-H_2 mass conversion factor, and examine [CII] as a universal tracer of gas mass. In Section 6 we discuss our findings, and apply our conversion [CII]-to-H_2factor to a sample of 19 [CII]- *and* CO-detected high-z galaxies and compare the molecular gas masses obtained with the two lines. Section 9 summarizes the main conclusions from our work.

2. OBSERVED COMPARISON SAMPLES

We employ three samples for comparison to simulations. The first is the observed galaxies from Z18, a sample of ten $z \sim 2$ main sequence galaxies with [CII] detections from ALMA (Atacama Large Millimetre/submillimetre Array) combined with other [CII] samples at different redshift. Their work determined that independent of a galaxy's main sequence or starburst nature, its [CII] luminosity and molecular gas mass are correlated by a linear relation, and that the ratio between these quantities does not vary with redshift, metallicity, or gas depletion time. We use their sample in combination with our simulations (see Section 3) to verify these statements.

We also use the more recent ALPINE-ALMA [CII] Survey (henceforth referred to as A2C2S; see Le Fèvre *et al.* (2019) and and Dessauges-Zavadsky *et al.* (2020)). A2C2S has preliminarily investigated 118 different galaxies for [CII] and FIR emission. In 64% of galaxies observed, [CII] was detected. The sample observed galaxies from a redshift range of $4 < z < 6$, which would include galaxies existing just after reionization where "rapid mass assembly" of galaxies was occurring in the universe. A2C2S finds a broad range of galaxy types across its survey, including 40% galaxy mergers.

Dessauges-Zavadsky *et al.* (2020) explicitly analyzes how molecular gas "budget" for star formation is traced by [CII], analyzing the correlation between [CII] and molecular gas mass as well as gas depletion timescales and molecular gas fractions for galaxies with [CII] measurements. We use their sample to calibrate [CII], both with and without SFR, in combination with our simulations and the Z18 galaxies as a tracer of molecular gas mass.

Lastly we include galaxies from the final data release of the APEX low-redshift legacy survey for molecular gas (ALLSMOG; see Cicone *et al.* (2017)), which consists of CO measurements of 88 galaxies from APEX-1 and 9 galaxies from IRAM. We consider this sample of low-redshift main-sequence galaxies to compare with the local main-sequence sources generated by our simulations. We do not, however, use these galaxies for our [CII] fit, as this was not measured by ALLSMOG.

3. SIMULATIONS

3.1. *Hydro-simulations*

This work builds on the analysis of snapshots taken from the SIMBA suite of cosmological galaxy formation simulations, which themselves were evolved using the meshless finite mass hydrodynamics technique of GIZMO (Hopkins 2015, 2017; Davé *et al.* 2019). A total of 1024^3 gas elements and 1024^3 dark matter particles are evolved from $z = 249$ to low redshift. Compared to its predecessor, MUFASA, new features in SIMBA include the growth and feedback of supermassive black holes as well as a subgrid model to form and destroy dust during the simulation run, and for details we refer to Davé *et al.* (2019) and Li *et al.* (2018) for these two processes, respectively. The galaxy properties in SIMBA have been compared to various observations across cosmic time and shown to provide reasonable agreement (.e.g Thomas *et al.* 2019; Appleby *et al.* 2020).

In total, 11,137 $z \simeq 6$ galaxies were selected with the YT-based package CAESAR. The sample, as well as the cuts made to the sample, are described in greater detail in Leung et al. (2020, in press). Our final sample consists of 9,364 galaxies between a [CII] luminosity range of $10^{3.82}$ to $10^{8.91}$ $L_\odot$ and a molecular gas mass range of $10^{6.25}$ to $10^{10.33}$ $M_\odot$ The galaxy properties derived with

CAESAR include star formation rate, star formation rate surface density, stellar mass ($M_\star$) and total gas mass. In addition, we derive a SFR-weighted gas-phase metallicity, Z. All of these properties will be used for the analysis in Section 4.

3.2. SÍGAME *post processing*

The sample of SIMBA galaxies are post-processed with the latest version of the SÍGAME module[10]. At its core, SÍGAME uses CLOUDY to model the line emission from all ISM within each model galaxy. Essentially, the only four parameters needed to calculate the line emission are; an interstellar radiation field (ISRF) spectrum, a cosmic ray (CR) intensity, gas density and gas metal abundances. The radiation of importance for the chemical state of the gas, and hence the line emission, is the far-ultraviolet (FUV). In SÍGAME, the FUV luminosity of each stellar particle in the SIMBA galaxies is calculated by integrating the spectrum from STARBURST99 modeling of the stellar population contained in the stellar particle, based on its mass, age and metallicity. For each gas particle, the FUV flux is derived as a sum of all nearby stars within the smoothing length of that gas particle. The CR intensity is assumed to scale linearly with the FUV field strength. As for the metal abundances, gas-phase metallicities are carried over from the simulation directly. Finally, the gas density requires a sub-gridding of the fluid elements which are themselves too large ($\geq 10^5\,M_\odot$) to follow the intricate processes taking place inside giant molecular clouds (GMCs) of mass 10^4–$10^6\,M_\odot$. This sub-gridding is carried out via subdivision of the gas into three ISM phases. First, the gas mass of each fluid element is divided into a dense and a diffuse part based on the H_2 gas mass fraction of the simulation itself. The diffuse gas is further subdivided into neutral and ionized diffuse gas, leading to three ISM phases; dense gas which is later on organized in giant molecular clouds (GMCs), diffuse neutral gas (DNG) and diffuse ionized gas (DIG). For a full description of the main methods and assumptions of SÍGAME, we refer to Olsen *et al.* (2017) and references therein. A few updates to the code were introduced for the creation of this sample of galaxies (refer to Daisy et al. in prep?), listed in the following. The version of SÍGAME used here has been equipped to run intensive tasks in parallel. The CLOUDY look-up tables used for calculating line emission are now based on a finer grid of in total 8^3 parameters for both GMCs and diffuse gas clouds. As shown in Figure **??** , the DNG accounts for a miniscule percentage of the total gas mass, and for this reason we combine the DNG and the DIG into a "diffuse" gas phase (which is $> 99\%$ ionized gas).

4. SAMPLE PROPERTIES

4.1. *Integrated properties*

In this section we perform a series of sanity checks of our simulated galaxies by examining their star formation rates, masses (stellar and molecular), and metallicities in the context of known empirical correlations between these quantities for our local and high-z comparison samples (§2). A disclaimer is important at this point: in principle, we ought to compare observations with mock observations of our simulated galaxies. However, given the heterogeneous nature of the observed samples, this is not straightforward. Furthermore, we stress that observed integrated properties such as stellar and molecular gas masses are likely to be uncertain by at least a factor

[10] `https://kpolsen.github.io/SIGAME/index.html`

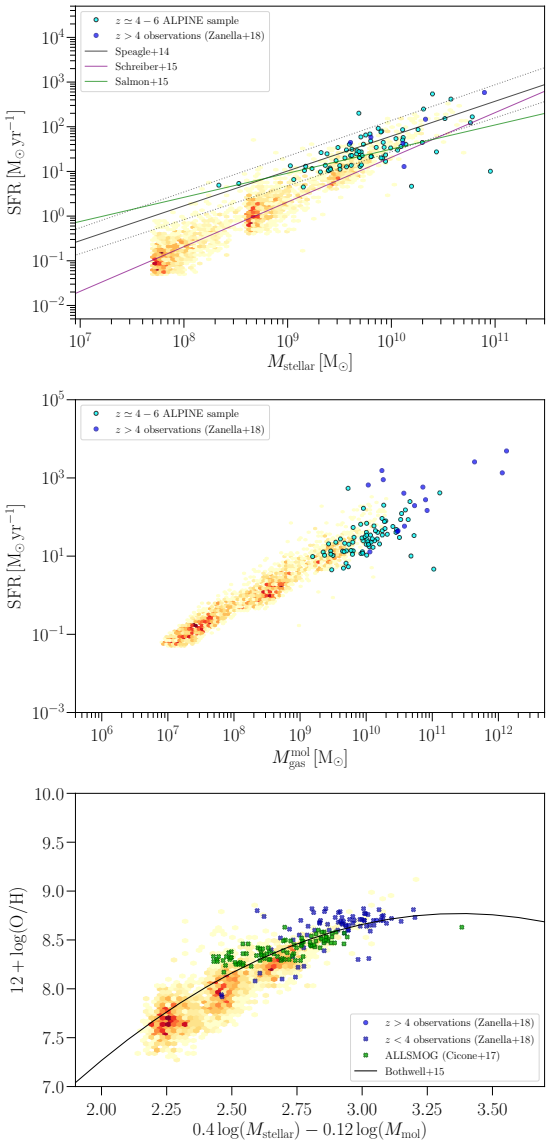

Figure 1. Top: The SFR $-$ $M_{\rm stellar}$ plane with the distribution of our simulated galaxies (yellow-red hexbins) shown. Our simulations fall on this main sequence where there is overlap between the two, but drop below it when extrapolated to lower masses. The ALPINE sample (cyan filled circles) lie on the main sequence, as do the $z > 4$ galaxies from Z18. The $z \simeq 0$ galaxies from Z18 along with the ALLSMOG sample (Cicone *et al.* 2017) form a separate main sequence with a lower normalisation as expected at lower redshifts. **Middle:** The integrated Kennicutt-Schmidt relation (SFR vs $M_{\rm mol}$) for our simulated galaxies. The high-z samples are seen to lie on the relation. **Bottom:** Metallicity ($12+\log({\rm O/H})$) vs the optimum linear combination of $\log M_{\rm stellar}$ and $\log M_{\rm H_2}$ as determined from a PCA analysis of the ALLSMOG sample (Bothwell *et al.* 2016). The solid curve represents the quadratic fit to the ALLSMOG data as given in Bothwell *et al.* (2016). Here, the ALLSMOG sample is shown using the revised data values from Cicone *et al.* (2017). The revised data (green symbols) are clearly offset from the curve and are in much better agreement with our simulations (yellow-red hexbins) and the observations by Z18 (blue symbols).

of two (Lower *et al.* 2020).

The top panel of Fig. 1 shows the location of our simulated galaxies in the SFR $-$ $M_{\rm stellar}$ plane, where they define a galaxy main sequence that extends over more than three orders of magnitude in stellar mass. The galaxy main sequence is usually modelled by a log-linear relation of the form: $\log {\rm SFR} = \alpha \log M_{\rm stellar} + \beta$, where the parameters α and β are functions of cosmic time, or redshift. The normalisation, β, in particular, is found to increase at earlier cosmic epochs. Observational constraints on the galaxy main sequence at $z \simeq 5-6$ come from Speagle *et al.* (2014), who inferred $\alpha = 0.78 \pm 0.02$ and $\beta = -6.01 \pm 0.15$ over the stellar mass range $10^{9.5}\,{\rm M}_\odot \lesssim M_{\rm stellar} \lesssim 10^{10.8}\,{\rm M}_\odot$ (thick solid line segment in Fig. 1). In this stellar mass range our simulations, as well as the $z > 4$ sources from Z18 and the ALPINE sample, agree with the observed main sequence relation from Speagle *et al.* (2014) both in terms of slope and normalisation. At lower stellar masses our simulations not only fall below the extrapolated main sequence relation by Speagle *et al.* (2014), but they also seem to follow a steeper slope. In contrast, the ALPINE sources at the low mass end are in line with the extrapolated relation. At face value, this would suggest that our simulations fail to reproduce the main sequence at stellar masses $\lesssim 10^{9.5}\,{\rm M}_\odot$, and that the simulations become increasingly worse at increasingly lower stellar masses. On the other hand we do not have good constraints on the low-mass end of the galaxy main sequence at $z > 4$, and any sampling of the main sequence at these epochs is incomplete below the mass range probed by Speagle *et al.* (2014). In this context, it is worth pointing out that Lee *et al.* (2015), who studied the evolution of the main sequence out to $z \sim 1.3$, found a turnover in the sequence above a stellar mass of $\sim 10^{10}\,{\rm M}_\odot$ at all redshifts. In Fig. 1 this is illustrated by the dot-dashed line, which is the parametrisation of the $z = 0.02$ galaxy main sequence by Lee *et al.* (2015). For comparison we have overplotted the $z \sim 0.02$ ALLSMOG sample, and the local galaxies from Z18. An extrapolation of this parametrisation to $z = 6$, using the prescription given in Lee *et al.* (2015), fails to match our simulations as well as the observations. This is not surprising since, as stated in Lee *et al.* (2015), there are large uncertainties on the prescription and its evolution with redshift, uncertainties which will be amplified when extrapolated to large redshifts. Nonetheless a turn-over in the $z = 6$ main sequence cannot be ruled out but will have to await upcoming surveys with the James Webb Space Telescope.

The middle panel of Fig. 1 shows SFR vs $M_{\rm mol}$ for our simulated galaxies, along with our comparison samples. Our simulated galaxies trace a tight correlation, which is essentially a galaxy-integrated version of the Kennicutt-Schmidt (KS) law (Kennicutt 1998), and is shared by the $z > 4$ comparison galaxies. The comparison galaxies at lower redshifts tend to lie below this relation, which is due to a longer global gas depletion time scales in normal galaxies at later epoch (REFs).

Lastly, we examine the relationship between gas-phase metallicity, stellar mass and molecular gas mass. Bothwell *et al.* (2016) showed that the SFR-metallicity relation at a given stellar mass (the so-called fundamental metallicity relation) is in fact a result of an underlying relation between metallicity and molecular gas mass. Using the ALLSMOG sample, they derived a relation between metallicity ($12 + \log({\rm O/H})$), stellar mass, and molecular gas mass. This relation (corrected to the cosmology adopted in this paper) is shown as the solid line in Fig. 1 (bottom panel). It represents the linear combination in $\log M_{\rm stellar}$ and $\log M_{\rm H_2}$, which yields the least scatter in metallicity according to a principal component analysis performed in Bothwell *et al.* (2016). More recently, however, Cicone *et al.* (2017) has reanalysed the ALLSMOG sample and derived improved metallicities, as well as stellar and molecular gas masses. Using these revised and final values, we plot the ALLSMOG sample in Fig. 1 (green symbols). They are clearly offset from the original relation by Bothwell *et al.* (2016), which highlights the importance of using revised values but at the same time verifies the existence of a relation of the kind suggested by Bothwell *et al.* (2016). Interestingly, we find that the galaxies from Z18 are consistent with the ALLSMOG sample, independent of their redshifts, and even extends the relation to the point where it turns over. To overlay the SIMBA simulations, we use the calibration $12 + \log({\rm O/H}) = \log(\langle Z \rangle_{\rm mass}) + 9$ derived by Ma *et al.* (2016) using high-resolution cosmological simulations over the redshift range $z = 0-6$. Ma *et al.* (2016) uses the mass-weighted average gas-phase metallicity, $\langle Z \rangle_{\rm mass}$, and for that reason we have used this quantity from our simulations, and not the SFR-weighted average gas-phase metallicities (see §3.1), to infer $12 + \log({\rm O/H})$. Our simulations are seen to be in good agreement with the observations in the region where there is overlap, and they show a continuing declining trend towards lower metallicities.

Based on the above analysis, we hold that the simulated SIMBA galaxies are in overall agreement with key scaling relations observed for $z > 4$ main sequence galaxies, and that the simulations therefore are representative, even at the low stellar mass end that extends well below the observed range, of the bulk of the normal star forming galaxies at these epochs.

5. RESULTS

5.1. *Does [CII] trace molecular gas mass or star formation rate?*

In this section we examine whether there is a correlation between the [CII] emission and the molecular gas masses of our simulated $z \simeq 6$ galaxies, and to what extent this correlation agrees with the empirical relation found by Z18. We also seek to infer a [CII] luminosity to molecular gas mass conversion factor, $\alpha_{\rm [CII]}$.

Fig. 2(top) shows the [CII] luminosity vs the total molecular gas mass for our simulated galaxies (yellow and red hexbins), along with observed low- and high-z galaxies (pink and blue symbols; Z18) and $z = 4-6$ galaxies from the ALPINE survey (cyan symbols; Faisst, private communication). The $\log L_{\rm [CII]} - \log M_{\rm mol}$ relation exhibited by our simulated galaxies is significant, with a Pearson correlation coefficient of $R \simeq 0.42$ between $\log L_{\rm [CII]}$ and $\log M_{\rm mol}$, which correspond to a p-value of $\sim 10^{-23}$ that the observed correlation is random. A Bayesian log-linear regression model of the form $\log L_{\rm [CII]} = a \log M_{\rm mol} + b$ fitted to simulations only yields:

$$\log L_{\rm [CII]} = (0.92 \pm 0.01) \log M_{\rm mol} - (0.99 \pm 0.10), \quad (1)$$

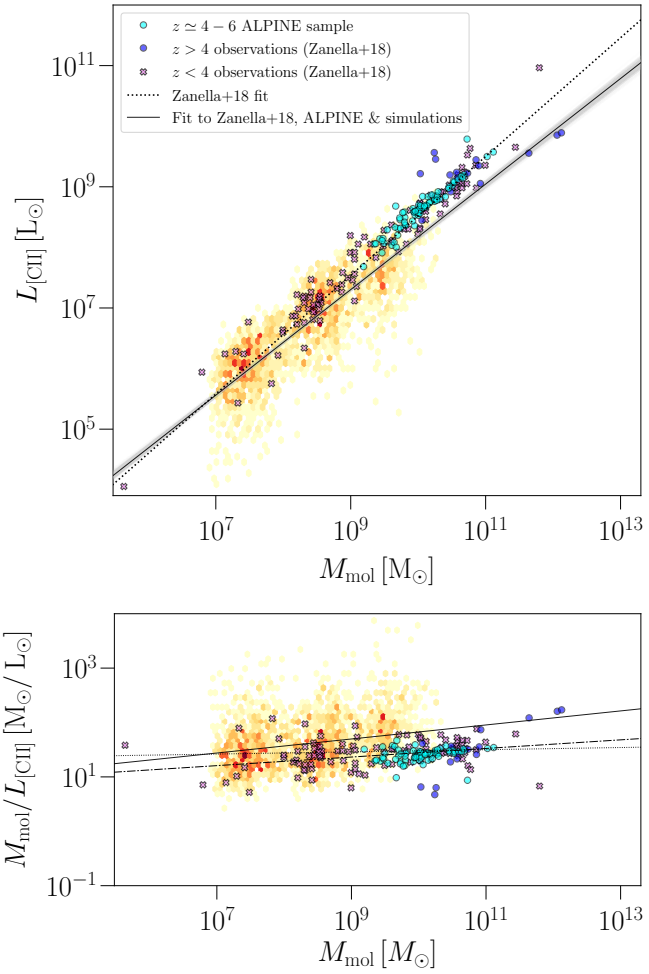

Figure 2. Top: Molecular gas masses vs [CII] luminosities for our simulated galaxies and observed comparison samples. Our fit in is excellent agreement with the fit parametrized by Z18.

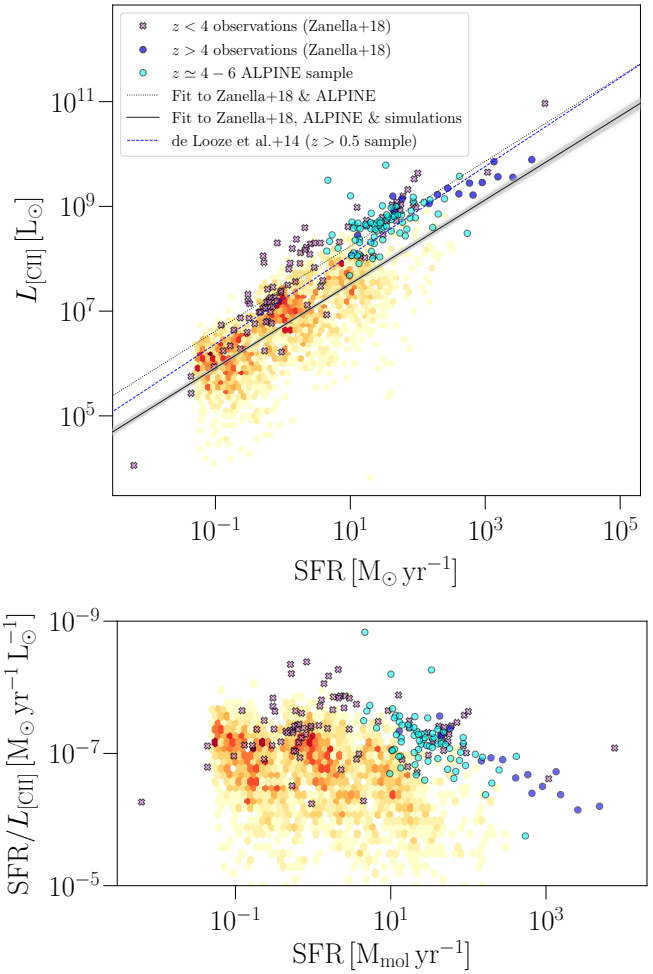

Figure 3. Top: [CII] luminosity vs star formation rate for our simulated galaxies, and the samples from ALPINE and Z18.

shown as the solid line in Fig. 2 (top). The mean absolute deviation (MAD) of the simulations around the fitted line is 0.3 dex. This relation is slightly discrepant with the relation derived by Z18, i.e., $\log L_{[CII]} = (0.98 \pm 0.02) \log M_{mol} - (1.28 \pm 0.01)$, which is shown as the dotted line in Fig. 2. They reported a scatter of 0.2 dex. The discrepancy with our simulations is because the latter contains a number of galaxies, which exhibit very low [CII] luminosities given their gas masses. The effect of this is a shallower slope and lower intercept in the simulations only relation. Fitting to the combined data from the simulations *and* the observations, we get:

$$\log L_{[CII]} = (0.94 \pm 0.01) \log M_{mol} - (1.18 \pm 0.10), \quad (2)$$

shown as the grey solid line in Fig. 2(top). The Z18 relation is seen to go through the bulk of the simulated data as well as the ALPINE data. We find the median absolute deviation of this relation to be 0.3 dex, where we have included our simulations and the ALPINE data, in addition to the Z18 data. This is only slightly larger than the 0.2 dex reported by Z18.

Fig. 2(bottom) shows $M_{mol}/L_{[CII]}$ vs $M_{mol}/L_{[CII]}$ for our simulations and the data samples. The overall Pearson correlation coefficient is 0.018, which tells us that $M_{mol}/L_{[CII]}$ is consistent with being constant across the molecular gas mass range $10^7 - 10^{12}$ M$_\odot$. This is also true for the simulations only. Overall, therefore, our simulations are in agreement with the original linear relation (i.e., slope of unity) proposed by Z18

The traditional interpretation of the [CII] line has been as a diagnostic of the physical conditions in photon-dominated regions, and as a tracer of the star formation activity of normal star forming galaxies (Stacey *et al.* 1991; De Looze *et al.* 2014). In Fig. 3 we show $L_{[CII]}$ vs SFR for our simulations, the ALPINE, and the Z18 sample (henceforth referred to as Z18). Also shown are log-linear fits to the combined Z18 and ALPINE samples, and to the two samples combined with our simulations. The two fits have similar slopes (~ 0.8) but the latter has a lower normalisation due to the tail of low-$L_{[CII]}$ galaxies at a given SFR. For comparison the global $L_{[CII]} - $ SFR relation inferred for $z > 0.5$ star forming galaxies by De Looze *et al.* (2014) is also shown. It matches quite well the fit to the combined Z18 and ALPINE samples. Both tend to overshoot the simulations. The relationship between SFR and $L_{[CII]}$ clearly exhibit more scatter than the $M_{mol} - L_{[CII]}$ relation. This is particularly evident in the bottom panel of Fig. 3, which shows significant scatter in SFR/$L_{[CII]}$.

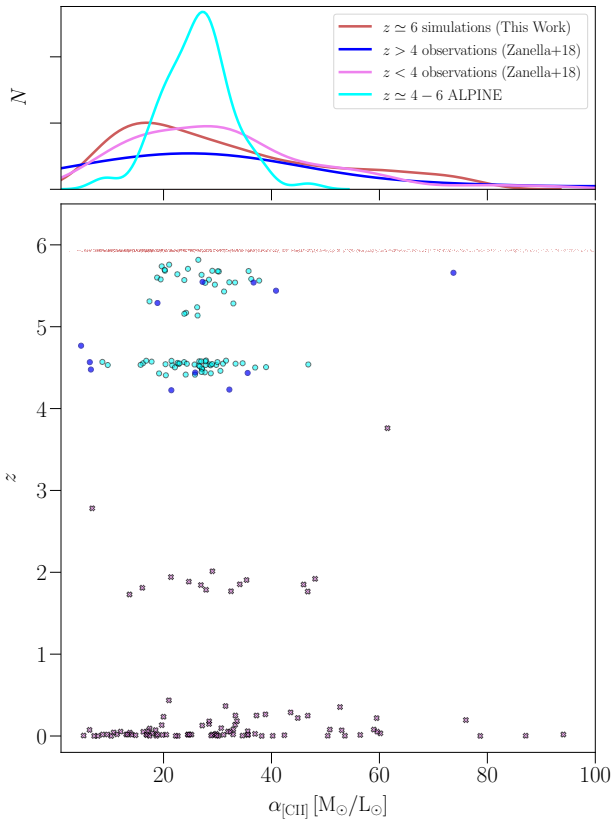

Figure 4. $\alpha_{[CII]}$ vs z for our simulations and comparison samples (bottom panel) and their $\alpha_{[CII]}$-distributions integrated over redshift (top panel).

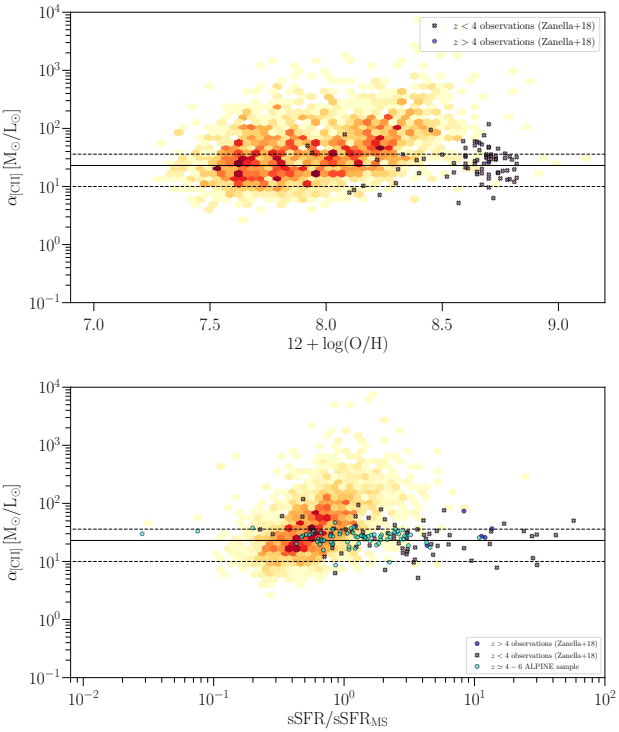

Figure 5. Recreation of Figures 9 and 10 from Z18 using our simulations and the Z18 galaxies. The top plot shows α as a function of distance from the main sequence. Our galaxies generally lie to the left of the Z18 galaxies, for the most part. The bottom plot shows metallicity as a function of α. Again our galaxies seem to lie to the left of most of the galaxies in the Z18 sample, albeit with some overlap.

5.2. *The [CII] conversion factor*

The fact that the $L_{[CII]} - M_{mol}$ relation is log-linear with a slope of unity allowed Z18 to define a [CII] luminosity to molecular gas mass conversion factor, $\alpha_{[CII]} = M_{mol}/L_{[CII]}$. They inferred a value of $\alpha_{[CII]} \sim 30\,M_\odot\,/L_\odot$ averaged across their sample (see Fig. 4 (bottom)). Z18 argued that $\alpha_{[CII]}$ was effectively invariant with redshift, metallicity, and galaxy type (e.g., merger vs main sequence galaxy). Our simulated galaxies are all main sequence galaxies, and cover a similar range in terms of masses and star formation rate as the local galaxies in the Z18 sample, but are at $z \sim 6$ and probe $\sim 10\times$ lower metallicities than the observations (Fig. 1). Averaging over our sample of simulated galaxies, we find a median value of $\alpha_{[CII]} \sim 38\,M_\odot\,/L_\odot$, with a median absolute deviation (MAD) of $23\,M_\odot/L_\odot$, i.e., in excellent agreement with Z18. We also find agreement with the ALPINE sample, which has a median conversion factor ofm $\alpha_{[CII]} \sim 27\,M_\odot\,/L_\odot$ and a MAD of $4\,M_\odot/L_\odot$.

In Fig. 4 we plot $\alpha_{[CII]}$ vs z for our simulations and the Z18 and ALPINE samples, along with kernel density estimations of their overall $\alpha_{[CII]}$ distributions (top panel). The $\alpha_{[CII]}$-distributions for our simulated galaxies peaks at $15\,M_\odot/L_\odot$ with an extended tail towards higher values. The $\alpha_{[CII]}$-distribution for the Z18 sample, and in particular the $z < 4$ sub-sample, exhibit a similar behaviour. The $\alpha_{[CII]}$-values of ALPINE sample are more narrowly distributed around their peak, which might be due to them being a relatively homogeneous sample of main sequence galaxies. In contrast, the Z18 sample, as well as our simulations span a wider range in star formation rate and masses. Within the uncertainties of our simulations, as well as the observations, we conclude that $\alpha_{[CII]}$ is indeed invariant with redshift.

In Fig. 5 (top panel) we show $\alpha_{[CII]}$ vs metallicity (i.e., 12 + log(O/H)) for our simulations along with the Z18 sample. Starting at the lowest metallicities ($12 + \log O/H \sim 7.3$) and until $12 + \log O/H \sim 8.1$ the simulated $\alpha_{[CII]}$ values are constant and consistent with the Z18 values (note, we do not have metallicity estimates for the ALPINE sample). At higher metallicities, however, which is where the simulations start to overlap with the observations in terms of metallicity, the former exhibit an upturn in $\alpha_{[CII]}$ that is not seen in the Z18 galaxies.

Naively, one would expect $L_{[CII]}$ to increase with metallicity, as this implies a higher abundance of carbon. This in turn would imply a decreasing trend in $\alpha_{[CII]}$ with metallicity, which is not what is seen in Fig. 5 (top). However, an increase in metallicity also implies a higher dust content, and thus fewer $h\nu > 11.3\,eV$ photons available to ionise neutral carbon. A higher dust content also implies more UV-shielded regions, as well as a larger overall dust grain surface where H_2 can form. In such a scenario, one would therefore expect an increase in $\alpha_{[CII]}$ with metallicity. Overall, our simulations confirm that there does not seem to be a strong dependency in the

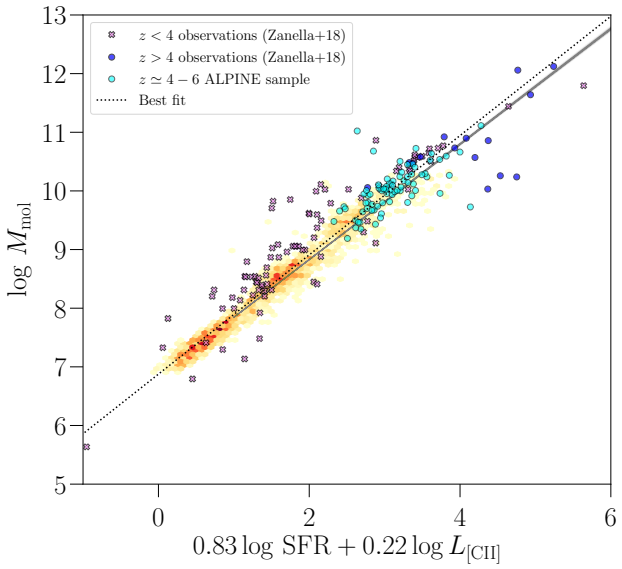

Figure 6. Equation 7 for our simulated galaxies, ALPINE galaxies, and Z18 galaxies.

conversion factor with metallicity, and that a constant conversion can be applied to galaxies in the epoch of cosmic reionisation to infer the molecular gas masses.

In the bottom panel of Fig. 5 we show $\alpha_{\rm [CII]}$ vs offset from the main sequence, which for a given galaxy is defined as the ratio of its specific star formation rate (sSFR = SFR/$M_{\rm stellar}$) and the specific star formation of a galaxy with the same stellar mass on the main sequence relation (at the given redshift). We adopt the prescription for the main sequence at $z = 6$ as given by Schreiber *et al.* (2015), since it agrees very well with our simulations (§4.1). As for our simulations, most of which are consistent with the main sequence defined by Schreiber *et al.* (2015), we see a trend of increasing $\alpha_{\rm [CII]}$ with main-sequence offset. In contrast, both the Z18 and ALPINE samples show constant $\alpha_{\rm [CII]}$. This is true even for the sources with large offsets from the main-sequence.

5.3. *Star formation rate and metallicity*

(Working on this section) So far we have only considered ways of deriving the molecular gas mass from the [CII] luminosity. This is because the latter is a direct observable, which can be measured in a relatively straightforward manner in (sufficiently bright) $z \sim 6$ galaxies with ALMA (REFs). In §5.1 we saw that a log-linear model to the $M_{\rm gas}^{\rm mol}$ relation predicts the [CII] luminosity, albeit with considerable scatter. It is important to consider second-order effects from other physical quantities which may play a role in shaping the $M_{\rm gas}^{\rm mol}$ relation and which may constrain the scatter further. Given the above, it is to be expected that other physical properties of the galaxies and their ISM will affect the $M_{\rm gas}^{\rm mol}$ relation. In this section we shall explore how additional observables, if available, might be used in conjunction with the [CII] luminosity to derive more accurate molecular gas masses. The observables we shall focus on are the star formation rate and the metallicity. The star formation rate is one of the observables that can be measured with some confidence towards $z \simeq 6$ galaxies. The unobscured star formation rate can be measured in a number of ways, from the strength of the rest-frame continuum at 1500 Å, for example, or from the Lyα emission strength. Obscured star formation rates are typically inferred from (sub-)millimeter or far-IR continuum measurements. We consider here the total star formation rates of our simulated galaxies and shall not distinguish between obscured or unobscured star formation rates.

$$\mathrm{PC1} = 0.52 \log {\rm SFR} + 0.54 \log M_{\rm mol} + 0.48 \log L_{\rm [CII]} + 0.46 Z_{\rm OH} \quad (3)$$

$$\mathrm{PC2} = -0.01 \log {\rm SFR} - 0.14 \log M_{\rm mol} - 0.58 \log L_{\rm [CII]} + 0.80 Z_{\rm OH} \quad (4)$$

$$\mathrm{PC3} = 0.57 \log {\rm SFR} + 0.34 \log M_{\rm mol} - 0.63 \log L_{\rm [CII]} - 0.40 Z_{\rm OH} \quad (5)$$

$$\mathrm{PC4} = 0.63 \log {\rm SFR} - 0.76 \log M_{\rm mol} + 0.17 \log L_{\rm [CII]} - 0.01 Z_{\rm OH} \quad (6)$$

Of the four principal components, PC1 is responsible for 83% of the variance, and PC2 and PC3 are responsible for 11 and 6%, respectively. Given its negligible contribution to the variance, we can set PC4 to zero, which in turn allows us to derive the following expression for the molecular gas:

$$\log M_{\rm mol} = 0.83 \log {\rm SFR} + 0.22 \log L_{\rm [CII]} - 0.01 Z_{\rm OH}. \quad (7)$$

This tells us that the molecular gas mass primarily correlates with the star formation rate and, to a lesser extent, with the [CII] luminosity. The dependency on metallicity is negligible. In Fig. 6 we have plotted eq. 7 for our simulations and comparison data, without including the $Z_{\rm OH}$-term in eq. 7. Adopting a linear model of the form $\log M_{\rm mol} = mx + n$, where $x = 0.83 \log {\rm SFR} + 0.22 \log L_{\rm [CII]}$, and m and n are the slope and intercept, respectively, and fitting it to our data, we obtain $m = 1.0 \pm 0.1$ and $n = 6.9 \pm 0.1$. Thus, based on our simulations, we derive the following scaling relation between molecular gas mass, star formation rate and [CII] luminosity:

$$\log M_{\rm mol} = 0.83 \log {\rm SFR} + 0.22 \log L_{\rm [CII]} + 6.9 \quad (8)$$

This relation is shown as the dotted line in Fig. 6. **scatter?**

6. DISCUSSION

6.1. *Discussion of results*

6.2. *Applying $\alpha_{\rm [CII]}$ to [CII] observations at $z \simeq 1-7$*

In this section we shall apply our derived expression for [CII]-derived molecular gas mass to actual observations of the [CII] line at $z \simeq 1-7$ in order to derive their molecular gas masses. To this end we extract a sample of $z = 1-7$ galaxies observed in [CII] using the DÍGAME data-base (`https://www.digame-db.online`). We restrict our observed sample to include sources which in addition to the [CII] line also have observations of CO(1-0) or CO(2-1). Because some sources only report CO(2-1) luminosities we multiply CO(2-1) luminosities by a factor of 0.8. In total, nineteen galaxies in the DÍGAME database fulfill these criteria. Of these galaxies, three are quasars, two are Lyman-break galaxies, and the rest are submillimeter galaxies. The galaxies have [CII] luminosities which range from $\sim 10^{7.5}$ to $10^{10.5}$ $L_\odot$ and CO luminosities ranging between $10^{5.9}$ to $10^{7.4}$ $L_\odot$. We provide references to these galaxies in Appendix A. For galaxies which are gravitationally lensed we divide their luminosities by their reported lensing factor μ.

We first derive gas masses from CO luminosities using the canonical relation between molecular gas mass and CO luminosity:

$$M_{\rm mol} = \alpha_{\rm CO} L_{\rm CO} \quad (9)$$

Where $L_{\rm CO}$ is measured in units of K km/s pc^2. M_{mol} for [CII] luminosities were calculated using the linear relation in Eq. 2, taken from Z18. In Figure 7 we plot the logarithm of the ratio of the luminosity-derived gas masses against redshift of the DIGAME sources.

Uncertainties on CO and [CII] luminosities were calculated via Bayesian posterior sampling. We assume for each observation that a Gaussian describes the range and likelihoods of measurement, with the reported luminosity as the mean and the reported error as the full-width half-maximum (FWHM). Using Equation 2 we assume a fixed slope and intercept (i.e. no uncertainties) and calculate a gas mass for each point in the distribution in CO and [CII]; finally we take the ratio of the two and recalculate the standard deviation for the Gaussian distribution of gas mass ratios.

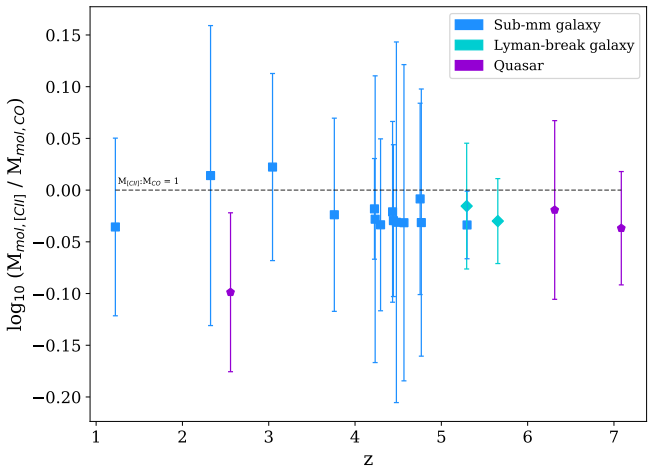

Figure 7. The ratios of [CII]-derived and CO-derived gas masses for our observational sample are plotted as a function of their redshift. The dashed line represents a 1:1 ratio. The observational effects that come into play with CO measurements naturally make this comparison difficult, but it is evident that for most of our galaxies [CII]-derived gas masses yield similar results to molecular gas masses derived from CO, the canonical gas mass tracer.

Figure 7 demonstrates an excellent agreement between [CII] and CO-derived molecular gas masses using Equation 2. We assume for the CO-derived gas mass that $\alpha_{CO} = 4.5$, the coefficient used for deriving molecular gas mass in normal, star-forming galaxies. However this number can be higher or lower (from about 0.5 to as high as 10) depending on galaxy type.

DV: what's left...

1. going to higher redshifts? Do observations agree?
2. spatially resolved observations: compare sigame with those and see how results compare

Combining ALMA observations of [CII] towards a sample of 10 $z \sim 2$ main sequence galaxies with [CII] data of local and high-z galaxies, main sequence galaxies as well as starburst galaxies, Z18 derive a [CII]-to-H_2 conversion factor of $\alpha_{[CII]} = 31\,{\rm M}_\odot\,/{\rm L}_\odot$. We find the median $\alpha_{[CII]}$ for our simulations to be in agreement with Z18. We also find that the conversion factor is invariant with gas phase metallicity and gas depletion time. However, we cannot conclusively confirm that the conversion factor is invariant with redshift. The median $\alpha_{[CII]}$ value of our simulations was lower than that of z > 4 observations we used, which itself was lower than the median value for the z < 4 observations. This could suggest that the [CII] conversion factor decreases as a function of redshift.

7. CONCLUSION

WORK IN PROGRESS Using SÍGAME postprocessing of galaxies from SIMBA cosmological simulations, we have simulated and analysed the [CII] emission from 9,364 $z = 6$ main sequence galaxies over the stellar mass range $5 \times 10^8 - 10^{11}\,{\rm M}_\odot$. Inspired by empirical calibrations based on galaxy samples at lower redshifts Z18, we have investigated whether [CII] – the brightest emission line in the far-IR to millimeter part of the spectra of galaxies – can be extended and understood as a tracer of the molecular gas in galaxies at the epoch of reionization. This would have important positive consequences for studying galaxy and star-formation at these early cosmic epochs, where the warm CMB can 'drown out' already faint traditional molecular gas tracers such as CO and cool dust emission.

Our simulated galaxies delineate a log-linear relationship between their molecular gas masses and their [CII] luminosities, that is found to be in excellent agreement with the empirical relationship derived by Z18, a well as with the [CII] luminosities and molecular gas masses inferred from massive main sequence galaxies at $z \sim 4-6$ from the ALPINE survey (Dessauges-Zavadsky *et al.* 2020). A log-linear fit to the simulations and observational data combined yields a slightly sub-linear relation: $\log L_{[CII]} = (0.87 \pm 0.01) \log M_{\rm mol} - (0.54 \pm 0.10)$. The median absolute deviation around this relation is 0.3 dex, which is largely drive by the scatter of the simulated galaxies, and in particular by the fraction ($\sim X\%$) of simulations that have very low [CII] luminosities compared to their molecular gas masses. **(ZZZ: explain why this is - ISM phases)**. Such [CII]-faint systems are likely to have been missed by observations, and would therefore not have been taken into account by the empirical relation.

The [CII] to molecular gas conversion factors for our simulated galaxies is consistent with the distributions for $z \geq 4$ and $z < 4$ sub-samples of the Z18 sources **(ZZZ: make statistical test)**.] All three distributions are broad and show a tail towards large $\alpha_{[CII]}$-values. In contrast, the $\alpha_{[CII]}$-distribution for the ALPINE sources is narrower. This may reflect the relative uniform and narrow stellar mass and star formation rate selection of the ALPINE sources, whereas the Z18, as well as our simulations, span a wider range in these quantities as well as in metallicity. The median $\alpha_{[CII]}$-value from the simulations is 38 ${\rm M}_\odot\,/{\rm L}_\odot$, in good agreement with the 30 ${\rm M}_\odot\,/{\rm L}_\odot$ derived by Z18.

As SFR has been shown to correlate with [CII] luminosity (e.g. Magdis *et al.* 2014; De Looze *et al.* 2014; Herrera-Camus *et al.* 2015), we use principal compo-

nent analysis (PCA) to derive an expression for molecular gas mass as a function of both [CII] luminosity and star-formation rate: $\log M_{\rm mol} = 0.83 \log {\rm SFR} + 0.22 \log L_{\rm [CII]} + 6.9$. **mention scatter and compare to other conversions/cii-sfr relation**

We lastly apply our log-linear relation to 19 high-redshift ($1 < z < 7$) galaxies with both [CII] and CO measurements, finding that [CII]-derived and CO-derived molecular gas masses are in agreement.

7.1. *Acknowledgements*

We thank Robert Thompson for developing CAESAR, and the YT team for development and support of YT. This research was made possible by the National Science Foundation (NSF)-funded DAWN-IRES program and would not be possible without support from the Cosmic Dawn Centre (DAWN) in Copenhagen, Denmark. The Cosmic Dawn Centre is funded by the Danish National Research Foundation. DV is grateful for the encouragement and support of his friends, teachers, and colleagues at the Wesleyan University Department of Astronomy.

APPENDIX

ID	z	L_{CO} ($L_\odot$)	$L_{[CII]}$ ($L_\odot$)	μ	References	
AzTECJ100020.71+023518.2	5.298	$2.26 \pm 0.29 \times 10^7$	$6.33 \pm 0.32 \times 10^9$		Riechers+10, Pavesi+16	
GN26	1.223	$2.81 \pm 0.76 \times 10^7$	$6.88 \pm 0.95 \times 10^9$		Frayer+08, Stacey+10	
H1413+117	2.558	$2.13 \pm 0.39 \times 10^7$	$3.49 \pm 0.43 \times 10^8$	11	Reichers+11f, Reichers+07b	
H-ATLASJ090311.6+003906	3.046	$2.60 \pm 0.41 \times 10^7$	$2.86 \pm 0.41 \times 10^{11}$	11.1	Valtchanov+11	
HZ10	5.654	$1.08 \pm 0.22 \times 10^7$	$4.02 \pm 0.27 \times 10^9$		Pavesi+18	
LBG-1	5.294	$1.77 \pm 0.00 \times 10^6$	$1.70 \pm 0.16 \times 10^9$		Pavesi+18	
LESS65	4.445	$8.44 \pm 2.64 \times 10^6$	$3.29 \pm 0.37 \times 10^9$		Huynh+14, Swinbank+12	
LESSJ033229.4-275619	4.753	$7.52 \pm 1.67 \times 10^6$	$1.01 \pm 0.15 \times 10^{10}$		Coppin+10, De Breuck+11	
SDSSJ010013.02+280225.8	6.323	$4.83 \pm 1.65 \times 10^6$	$3.54 \pm 0.48 \times 10^9$		Wang+16	
SMMJ2135-0102	2.326	$2.81 \pm 0.16 \times 10^7$	$1.84 \pm 0.39 \times 10^{11}$	32.5	Ivison+10b	
SPT0113-46	4.233	$1.19 \pm 0.09 \times 10^8$	$5.24 \pm 1.09 \times 10^{10}$	23.9	Aravena+16, Gullberg+15	
SPT0345-47	4.296	$1.29 \pm 0.14 \times 10^8$	$3.76 \pm 0.49 \times 10^{10}$	8	Aravena+16, Gullberg+15	
SPT0418-47	4.225	$9.06 \pm 0.84 \times 10^7$	$7.30 \pm 0.59 \times 10^{10}$	32.7	Aravena+16, Gullberg+15	
SPT0441-46	4.477	$7.25 \pm 1.07 \times 10^7$	$2.67 \pm 0.67 \times 10^{10}$	12.7	Aravena+16, Gullberg+15	
SPT2103-60	4.436	$1.20 \pm 0.19 \times 10^8$	$7.99 \pm 1.12 \times 10^{10}$	27.8	Aravena+16, Bothwell+17	
SPT2132-58	4.768	$7.13 \pm 0.59 \times 10^7$	$2.41 \pm 0.48 \times 10^{10}$	5.7	Aravena+16, Bothwell+17	
SPT2146-55	4.567	$7.47 \pm 1.26 \times 10^7$	$2.53 \pm 0.58 \times 10^{10}$	6.7	Aravena+16, Gullberg+15	
SPT2147-50	3.760	$7.26 \pm 1.45 \times 10^7$	$3.85 \pm 0.56 \times 10^{10}$	6.6	Aravena+16, Gullberg+15	
ULASJ1120+0641	7.085	$5.11 \pm 0.00 \times 10^6$	$1.38 \pm 0.12 \times 10^9$		Venemans+17	

Table 1
The 19 observed galaxies used for molecular gas mass comparison in Section 6.2. The redshifts provided are from the [CII] observations. Carbon monoxide (CO) observations are either CO(2-1) or CO(1-0) measurements (see footnotes). μ, the gravitational lensing factor, is 1 unless otherwise specified. The ratio $M_{[CII]}/M_{CO}$ derived from the listed luminosities is plotted against redshift (z) in Figure 7. **help clean up the table? also any more information needed?**

REFERENCES

J. Kennicutt, Robert C., ARA&A **36**, 189 (1998), arXiv:astro-ph/9807187 [astro-ph].
R. C. Kennicutt and N. J. Evans, ARA&A **50**, 531 (2012), arXiv:1204.3552 [astro-ph.GA].
A. D. Bolatto, M. Wolfire, and A. K. Leroy, ARA&A **51**, 207 (2013), arXiv:1301.3498 [astro-ph.GA].
M. R. Krumholz, ApJ **759**, 9 (2012), arXiv:1208.1504 [astro-ph.CO].
M. J. Michałowski, G. Gentile, J. Hjorth, M. R. Krumholz, N. R. Tanvir, P. Kamphuis, D. Burlon, M. Baes, S. Basa, S. Berta, J. M. Castro Cerón, D. Crosby, V. D'Elia, J. Elliott, J. Greiner, L. K. Hunt, S. Klose, M. P. Koprowski, E. Le Floc'h, D. Malesani, T. Murphy, A. Nicuesa Guelbenzu, E. Palazzi, J. Rasmussen, A. Rossi, S. Savaglio, P. Schady, J. Sollerman, A. de Ugarte Postigo, D. Watson, P. van der Werf, S. D. Vergani, and D. Xu, A&A **582**, A78 (2015), arXiv:1508.03094 [astro-ph.GA].
M. Fumagalli, *Food for stars: The role of hydrogen in the formation and evolution of galaxies*, Ph.D. thesis, University of California, Santa Cruz (2012).
D. P. Cox, ARA&A **43**, 337 (2005).
R. L. Dickman, R. L. Snell, and F. P. Schloerb, ApJ **309**, 326 (1986).
P. M. Solomon and J. W. Barrett, in *Dynamics of Galaxies and Their Molecular Cloud Distributions*, Vol. 146, edited by F. Combes and F. Casoli (1991) p. 235.
D. Downes and P. M. Solomon, ApJ **507**, 615 (1998), arXiv:astro-ph/9806377 [astro-ph].
D. Narayanan, M. R. Krumholz, E. C. Ostriker, and L. Hernquist, MNRAS **421**, 3127 (2012), arXiv:1110.3791 [astro-ph.GA].
C. D. Wilson, ApJ **448**, L97 (1995), arXiv:astro-ph/9506103 [astro-ph].
N. Arimoto, Y. Sofue, and T. Tsujimoto, PASJ **48**, 275 (1996).
M. G. Wolfire, D. Hollenbach, and C. F. McKee, ApJ **716**, 1191 (2010), arXiv:1004.5401 [astro-ph.GA].
R. Genzel, L. J. Tacconi, F. Combes, A. Bolatto, R. Neri, A. Sternberg, M. C. Cooper, N. Bouché, F. Bournaud, A. Burkert, J. Comerford, P. Cox, M. Davis, N. M. Förster Schreiber, S. Garcia-Burillo, J. Gracia-Carpio, D. Lutz, T. Naab, S. Newman, A. Saintonge, K. Shapiro, A. Shapley, and B. Weiner, ApJ **746**, 69 (2012), arXiv:1106.2098 [astro-ph.CO].
T. G. Bisbas, P. P. Papadopoulos, and S. Viti, ApJ **803**, 37 (2015), arXiv:1502.04198 [astro-ph.GA].
T. G. Bisbas, E. F. van Dishoeck, P. P. Papadopoulos, L. Szűcs, S. Bialy, and Z.-Y. Zhang, ApJ **839**, 90 (2017), arXiv:1703.08598 [astro-ph.GA].
E. da Cunha, B. Groves, F. Walter, R. Decarli, A. Weiss, F. Bertoldi, C. Carilli, E. Daddi, D. Elbaz, R. Ivison, R. Maiolino, D. Riechers, H.-W. Rix, M. Sargent, and I. Smail, ApJ **766**, 13 (2013), arXiv:1302.0844 [astro-ph.CO].
R. Tunnard and T. R. Greve, ApJ **819**, 161 (2016), arXiv:1602.01095 [astro-ph.GA].

Z.-Y. Zhang, P. P. Papadopoulos, R. J. Ivison, M. Galametz, M. W. L. Smith, and E. M. Xilouris, Royal Society Open Science **3**, 160025 (2016), arXiv:1605.03885 [astro-ph.GA].
M. Gerin and T. G. Phillips, ApJ **537**, 644 (2000), arXiv:astro-ph/0003252 [astro-ph].
P. P. Papadopoulos and T. R. Greve, ApJ **615**, L29 (2004), arXiv:astro-ph/0409559 [astro-ph].
M. Ikeda, T. Oka, K. Tatematsu, Y. Sekimoto, and S. Yamamoto, ApJS **139**, 467 (2002).
S. S. R. Offner, T. G. Bisbas, T. A. Bell, and S. Viti, MNRAS **440**, L81 (2014), arXiv:1401.5072 [astro-ph.SR].
S. C. O. Glover, P. C. Clark, M. Micic, and F. Molina, MNRAS **448**, 1607 (2015), arXiv:1403.3530 [astro-ph.GA].
R. H. Hildebrand, QJRAS **24**, 267 (1983).
M. Guelin, R. Zylka, P. G. Mezger, C. G. T. Haslam, E. Kreysa, R. Lemke, and A. W. Sievers, A&A **279**, L37 (1993).
M. Guelin, R. Zylka, P. G. Mezger, C. G. T. Haslam, and E. Kreysa, A&A **298**, L29 (1995).
S. Eales, M. W. L. Smith, R. Auld, M. Baes, G. J. Bendo, S. Bianchi, A. Boselli, L. Ciesla, D. Clements, A. Cooray, L. Cortese, J. Davies, I. De Looze, M. Galametz, W. Gear, G. Gentile, H. Gomez, J. Fritz, T. Hughes, S. Madden, L. Magrini, M. Pohlen, L. Spinoglio, J. Verstappen, C. Vlahakis, and C. D. Wilson, ApJ **761**, 168 (2012), arXiv:1202.0547 [astro-ph.CO].
G. E. Magdis, E. Daddi, M. Béthermin, M. Sargent, D. Elbaz, M. Pannella, M. Dickinson, H. Dannerbauer, E. da Cunha, F. Walter, D. Rigopoulou, V. Charmandaris, H. S. Hwang, and J. Kartaltepe, ApJ **760**, 6 (2012), arXiv:1210.1035 [astro-ph.CO].
N. Scoville, H. Aussel, K. Sheth, K. S. Scott, D. Sanders, R. Ivison, A. Pope, P. Capak, P. Vand en Bout, S. Manohar, J. Kartaltepe, B. Robertson, and S. Lilly, ApJ **783**, 84 (2014), arXiv:1401.2987 [astro-ph.GA].
N. Scoville, K. Sheth, H. Aussel, P. Vanden Bout, P. Capak, A. Bongiorno, C. M. Casey, L. Murchikova, J. Koda, J. Álvarez-Márquez, N. Lee, C. Laigle, H. J. McCracken, O. Ilbert, A. Pope, D. Sanders, J. Chu, S. Toft, R. J. Ivison, and S. Manohar, ApJ **820**, 83 (2016), arXiv:1511.05149 [astro-ph.GA].
M. Kaasinen, N. Scoville, F. Walter, E. Da Cunha, G. Popping, R. Pavesi, B. Darvish, C. M. Casey, D. A. Riechers, and S. Glover, ApJ **880**, 15 (2019), arXiv:1905.11417 [astro-ph.GA].
G. Accurso, A. Saintonge, T. G. Bisbas, and S. Viti, MNRAS **464**, 3315 (2017), arXiv:1607.03488 [astro-ph.GA].
T. M. Hughes, E. Ibar, V. Villanueva, M. Aravena, M. Baes, N. Bourne, A. Cooray, L. Dunne, S. Dye, S. Eales, C. Furlanetto, R. Herrera-Camus, R. J. Ivison, E. van Kampen, M. A. Lara-López, S. J. Maddox, M. J. Michałowski, M. W. L. Smith, E. Valiante, P. van der Werf, and Y. Q. Xue, A&A **602**, A49 (2017), arXiv:1611.05867 [astro-ph.GA].
A. Zanella, E. Daddi, G. Magdis, T. Diaz Santos, D. Cormier, D. Liu, A. Cibinel, R. Gobat, M. Dickinson, M. Sargent, G. Popping, S. C. Madden, M. Bethermin, T. M. Hughes, F. Valentino, W. Rujopakarn, M. Pannella, F. Bournaud, F. Walter, T. Wang, D. Elbaz, and R. T. Coogan, MNRAS **481**, 1976 (2018), arXiv:1808.10331 [astro-ph.GA].
M. Dessauges-Zavadsky, M. Ginolfi, F. Pozzi, M. Béthermin, O. Le Fèvre, S. Fujimoto, J. D. Silverman, G. C. Jones, D. Schaerer, A. L. Faisst, Y. Khusanova, Y. Fudamoto, P. Cassata, F. Loiacono, P. L. Capak, L. Yan, R. Amorin, S. Bardelli, M. Boquien, A. Cimatti, C. Gruppioni, N. P. Hathi, E. Ibar, A. M. Koekemoer, B. C. Lemaux, D. Narayanan, P. A. Oesch, G. Rodighiero, M. Romano, M. Talia, S. Toft, L. Vallini, D. Vergani, G. Zamorani, and E. Zucca, arXiv e-prints , arXiv:2004.10771 (2020), arXiv:2004.10771 [astro-ph.GA].
S. C. Madden, D. Cormier, S. Hony, V. Lebouteiller, N. Abel, M. Galametz, I. De Looze, M. Chevance, F. L. Polles, M. Y. Lee, F. Galliano, A. Lambert-Huyghe, D. Hu, and L. Ramambason, arXiv e-prints , arXiv:2009.00649 (2020), arXiv:2009.00649 [astro-ph.GA].
S. Malhotra, G. Helou, G. Stacey, D. Hollenbach, S. Lord, C. A. Beichman, H. Dinerstein, D. A. Hunter, K. Y. Lo, N. Y. Lu, R. H. Rubin, N. Silbermann, J. Thronson, H. A., and M. W. Werner, ApJ **491**, L27 (1997).
M. J. Kaufman, M. G. Wolfire, D. J. Hollenbach, and M. L. Luhman, ApJ **527**, 795 (1999), arXiv:astro-ph/9907255 [astro-ph].
J. L. Pineda, W. D. Langer, and P. F. Goldsmith, A&A **570**, A121 (2014), arXiv:1409.0537 [astro-ph.GA].
I. De Looze, D. Cormier, V. Lebouteiller, S. Madden, M. Baes, G. J. Bendo, M. Boquien, A. Boselli, D. L. Clements, L. Cortese, A. Cooray, M. Galametz, F. Galliano, J. Graciá-Carpio, K. Isaak, O. Ł. Karczewski, T. J. Parkin, E. W. Pellegrini, A. Rémy-Ruyer, L. Spinoglio, M. W. L. Smith, and E. Sturm, A&A **568**, A62 (2014), arXiv:1402.4075 [astro-ph.GA].
R. Herrera-Camus, A. D. Bolatto, M. G. Wolfire, J. D. Smith, K. V. Croxall, R. C. Kennicutt, D. Calzetti, G. Helou, F. Walter, A. K. Leroy, B. Draine, B. R. Brandl, L. Armus, K. M. Sand strom, D. A. Dale, G. Aniano, S. E. Meidt, M. Boquien, L. K. Hunt, M. Galametz, F. S. Tabatabaei, E. J. Murphy, P. Appleton, H. Roussel, C. Engelbracht, and P. Beirao, ApJ **800**, 1 (2015), arXiv:1409.7123 [astro-ph.GA].
G. E. Magdis, D. Rigopoulou, R. Hopwood, J. S. Huang, D. Farrah, C. Pearson, A. Alonso-Herrero, J. J. Bock, D. Clements, A. Cooray, M. J. Griffin, S. Oliver, I. Perez Fournon, D. Riechers, B. M. Swinyard, D. Scott, N. Thatte, I. Valtchanov, and M. Vaccari, ApJ **796**, 63 (2014), arXiv:1409.5605 [astro-ph.GA].
T. Díaz-Santos, L. Armus, V. Charmandaris, S. Stierwalt, E. J. Murphy, S. Haan, H. Inami, S. Malhotra, R. Meijerink, G. Stacey, A. O. Petric, A. S. Evans, S. Veilleux, P. P. van der Werf, S. Lord, N. Lu, J. H. Howell, P. Appleton, J. M. Mazzarella, J. A. Surace, C. K. Xu, B. Schulz, D. B. Sanders, C. Bridge, B. H. P. Chan, D. T. Frayer, K. Iwasawa, J. Melbourne, and E. Sturm, ApJ **774**, 68 (2013), arXiv:1307.2635 [astro-ph.CO].
D. Narayanan and M. R. Krumholz, MNRAS **467**, 50 (2017), arXiv:1601.05803 [astro-ph.GA].
L. Vallini, S. Gallerani, A. Ferrara, A. Pallottini, and B. Yue, ApJ **813**, 36 (2015), arXiv:1507.00340 [astro-ph.GA].
K. Olsen, T. R. Greve, D. Narayanan, R. Thompson, R. Davé, L. Niebla Rios, and S. Stawinski, ApJ **846**, 105 (2017), arXiv:1708.04936 [astro-ph.GA].
O. Le Fèvre, M. Béthermin, A. Faisst, P. Capak, P. Cassata, J. D. Silverman, D. Schaerer, and L. Yan, arXiv e-prints , arXiv:1910.09517 (2019), arXiv:1910.09517 [astro-ph.CO].
C. Cicone, M. Bothwell, J. Wagg, P. Møller, C. De Breuck, Z. Zhang, S. Martín, R. Maiolino, P. Severgnini, M. Aravena, F. Belfiore, D. Espada, A. Flütsch, V. Impellizzeri, Y. Peng, M. A. Raj, N. Ramírez-Olivencia, D. Riechers, and K. Schawinski, A&A **604**, A53 (2017), arXiv:1705.05851 [astro-ph.GA].
P. F. Hopkins, MNRAS **450**, 53 (2015), arXiv:1409.7395 [astro-ph.CO].
P. F. Hopkins, arXiv e-prints , arXiv:1712.01294 (2017), arXiv:1712.01294 [astro-ph.IM].
R. Davé, D. Anglés-Alcázar, D. Narayanan, Q. Li, M. H. Rafieferantsoa, and S. Appleby, MNRAS **486**, 2827 (2019), arXiv:1901.10203 [astro-ph.GA].
Q. Li, D. Narayanan, R. Davè, and M. R. Krumholz, ApJ **869**, 73 (2018), arXiv:1810.12315 [astro-ph.GA].
N. Thomas, R. Davé, D. Anglés-Alcázar, and M. Jarvis, MNRAS **487**, 5764 (2019), arXiv:1905.02741 [astro-ph.GA].
S. Appleby, R. Davé, K. Kraljic, D. Anglés-Alcázar, and D. Narayanan, MNRAS **494**, 6053 (2020), arXiv:1911.02041 [astro-ph.GA].
M. S. Bothwell, R. Maiolino, C. Cicone, Y. Peng, and J. Wagg, A&A **595**, A48 (2016), arXiv:1606.04102 [astro-ph.GA].
S. Lower, D. Narayanan, J. Leja, B. D. Johnson, C. Conroy, and R. Davé, arXiv/2006.03599 , arXiv:2006.03599 (2020), arXiv:2006.03599 [astro-ph.GA].
J. S. Speagle, C. L. Steinhardt, P. L. Capak, and J. D. Silverman, ApJS **214**, 15 (2014), arXiv:1405.2041 [astro-ph.GA].
N. Lee, D. B. Sanders, C. M. Casey, S. Toft, N. Z. Scoville, C.-L. Hung, E. Le Floc'h, O. Ilbert, H. J. Zahid, H. Aussel, P. Capak, J. S. Kartaltepe, L. J. Kewley, Y. Li, K. Schawinski, K. Sheth, and Q. Xiao, ApJ **801**, 80 (2015), arXiv:1501.01080 [astro-ph.GA].

X. Ma, P. F. Hopkins, C.-A. Faucher-Giguère, N. Zolman, A. L. Muratov, D. Kereš, and E. Quataert, MNRAS **456**, 2140 (2016), arXiv:1504.02097 [astro-ph.GA].
G. J. Stacey, N. Geis, R. Genzel, J. B. Lugten, A. Poglitsch, A. Sternberg, and C. H. Townes, ApJ **373**, 423 (1991).
C. Schreiber, M. Pannella, D. Elbaz, M. Béthermin, H. Inami, M. Dickinson, B. Magnelli, T. Wang, H. Aussel, E. Daddi, S. Juneau, X. Shu, M. T. Sargent, V. Buat, S. M. Faber, H. C. Ferguson, M. Giavalisco, A. M. Koekemoer, G. Magdis, G. E. Morrison, C. Papovich, P. Santini, and D. Scott, A&A **575**, A74 (2015), arXiv:1409.5433 [astro-ph.GA]