The ALPINE–ALMA [CII] Survey: Exploring the Dark Side of Normal Galaxies at the End of Reionisation

Matthieu Béthermin¹, Miroslava Dessauges-Zavadsky², Andreas L. Faisst³, Michele Ginolfi², Carlotta Gruppioni⁴, Gareth C. Jones⁵,⁶, Yana Kushanova¹,⁷, Brian Lemaux⁸, Peter L. Capak⁵,⁹,¹⁰, Paolo Cassata¹¹,¹², Olivier Le Fèvre¹, Daniel Schaerer², John D. Silverman¹³,¹⁴, Lin Yan¹⁵ and the ALPINE collaboration

Cold gas and cosmic dust are the fuel of star formation. ALPINE is an ALMA Large Programme which has built the first statistically representative sample of star-forming galaxies in the adolescent Universe by targeting emission from singly ionised carbon [CII] at 158 µm, which traces both emission from star-forming regions and molecular hydrogen gas clouds, and the thermal continuum from dust at the end of the epoch of reionisation (4.4 < z < 5.9). Observations by the ALPINE team have revealed that a significant fraction of the star formation at this epoch is already hidden by dust clouds. ALPINE observations have also shown how unruly these young galaxies were by finding a large fraction of mergers and ubiquitous gas outflows.

Description of the survey

The aims of the ALPINE survey

Before the inception of the Atacama Large Millimeter/submillimeter Array (ALMA), the main constraints on the evolution of galaxies in the adolescent Universe (4 < z < 6) came from rest-frame ultraviolet (UV) light that was redshifted to optical and near-infrared wavelengths. Thanks to its unprecedented sensitivity in the infrared, ALMA has opened a new window through which to explore the cold and dusty Universe at these early times. Pioneering studies demonstrated that ALMA can detect both the dust continuum and the far-infrared (FIR) fine-structure line [CII] (see, for example, Capak et al., 2015). Continuum and line emission can be targeted simultaneously with ALMA and are both valuable tracers of dust-obscured star formation. ALPINE (the ALMA Large Programme to INvestigate [CII] at Early times; Le Fèvre et al., 2020; Faisst et al., 2020; Béthermin et al., 2020) builds the first comprehensive, statistically representative, multi-wavelength sample of normal (main-sequence) star-forming galaxies at the end of the epoch of reionisation, with observations of all galaxies in the sample from the rest-frame UV to the far-infrared. In total, 118 spectroscopically selected galaxies at redshifts of 4.4 < z < 5.9 were observed as part of the ALPINE survey.

The main goals of this survey are:
– To test the [CII] line as a star formation tracer at high redshift, since some theoretical models predict a deficit in [CII] per unit star formation in the low-metallicity galaxies that dominate at high redshift as compared to the local SFR–[CII] relation.
– To use both the dust continuum and [CII] to estimate the contribution of obscured star formation in known spectroscopically selected sources at 4 < z < 6, and to use this estimate to understand how this contribution modulates the total star formation rate density at these epochs.

Figure 1. UV-to-millimetre spectrum of a galaxy and illustration of the various features probed by ALPINE and ancillary data (figure from Faisst et al., 2020). The spectrum sketch is based on a typical z = 5 galaxy.
To estimate the precise relationship between the stellar mass of a galaxy and its star-formation activity at $4 < z < 6$ by combining UV (emission tracing exposed young stars) and infrared (emission tracing young stars hidden by dust) data to estimate the total star formation rate (SFR).

To understand the basic interstellar medium (ISM) properties of these systems from their dust and [CII] luminosities.

To measure the dynamical masses of these systems using [CII] to constrain their gas fractions.

To measure the merger rate using the velocity and position information of [CII].

To identify and quantify possible gas outflows from the [CII] line profiles.

Sample selection and ancillary data

The ALPINE sample is primarily rest-frame UV-limited at $M_{1500} < -20.2$, a limit which corresponds to galaxies about 2.5 times fainter than those typical at this epoch (i.e., $L_{*UV}$). This limit naturally leads to a sample selected to an SFR limit of $\geq 10 M_{\odot}$ yr$^{-1}$. We found that this cut maximises the sample size while simultaneously minimising the amount of observing time with ALMA. To set the expectations of the ALMA observations, the [CII] emission fluxes were conservatively predicted using the relation between the observed UV luminosity and [CII] line emission based on a pilot sample by Capak et al. (2015). Because of ALMA’s narrow frequency bands, the redshifts of all targeted galaxies were already determined precisely by one of several large spectroscopic surveys on the targeted legacy fields (COSMOS and ECFDS).

Most of the spectroscopic redshifts were measured from observations taken as part of the VUDS survey (Le Fèvre et al., 2014) and the Keck/DEIMOS 10k survey (Hasinger et al., 2018). To mitigate potential biases associated with spectroscopic selections, the target sample was selected by several different methods, for example, via colours (Lyman-break dropout technique), narrow bands (Ly$\alpha$ emission selection), photometric redshifts, and serendipitous detections. Furthermore, the spectroscopic redshifts are derived from UV absorption lines as well as the Ly$\alpha$ emission feature. In total, 13 target galaxies are located in the ECDFS field (from the VISTA Deep Extragalactic Observations survey, VIDEO) and 105 in the COSMOS field from Cosmic Evolution Survey. As shown in Faisst et al. (2020), the ALPINE galaxies represent the average population of galaxies at these redshifts well in terms of the mass of their stellar content and SFRs. As such, ALPINE observations enable, for the first time, the study of the panchromatic properties of the average galaxy at these cosmic times with a high degree of statistical certainty.

The ALPINE team combined ALMA observations in the far-infrared with a wealth of exquisite ancillary imaging and spectroscopy products at rest-frame UV and optical wavelengths, all of which constitute the first large multi-wavelength survey of galaxies at these redshifts (Figure 1). All galaxies have deep photometry from UV to near-infrared (NIR) from ground-based telescopes, the Hubble Space Telescope, and the Spitzer Space Telescope. These data allow accurate constraints of galaxies’ physical properties (stellar mass, SFR, age) from high-quality spectral energy distribution (SED) fitting. In addition, the Spitzer [3.6 μm]–[4.5 μm] colours provide estimates of their H$\alpha$ emission strength, hence additional constraints on their rates of star formation. The deep rest-frame UV spectroscopy available for all galaxies provides valuable insights into the metallicity and stellar wind properties via UV absorption lines and the Ly$\alpha$ emission line. Furthermore, the UV continuum slope is a good measure of dust opacity along the line of sight and can, together with the ratio of far-infrared to UV luminosity, constrain the dust properties of these galaxies. The ancillary imaging and spectroscopic data, as well as various measurements of physical parameters, are detailed in Faisst et al. (2020).

ALMA observations and achieved performance

The ALPINE targets were selected to have a redshift in which [CII], if present, was observable by ALMA’s band 7. Owing to low atmospheric transmission, redshifts of $4.6 < z < 5.1$ were excluded. We used the two most compact configurations to maximise the sensitivity of the integrated flux and in order to detect extended [CII] and far-infrared continuum emission. To minimise the calibration overheads, we grouped our sources by pairs of galaxies with similar redshifts in order to observe them with the same correlator tuning.

ALPINE was mainly observed during Cycle 5, between May and August 2018.
Dust-obscured star formation in the adolescent Universe

Massive galaxies already contain a lot of dust

Observations of high-redshift galaxies are traditionally limited to rest-frame UV wavelengths and, if dust is present, are therefore strongly affected by obscuration. Generally, it is assumed that galaxies in the early Universe contain very little amounts of dust as the dust content must be built up over time. With ALPINE, we can directly measure the dust content and fraction of obscured UV light in a large sample of 4 < z < 6 galaxies (Fudamoto et al., in preparation) that span over an order of magnitude in stellar mass and SFRs. Overall, the ALPINE data confirm the drop in dust attenuation at high redshifts relative to similar galaxies at lower redshifts and show that galaxies at high redshifts can, on average, be characterised by a dust attenuation parameterisation similar to that found in the metal-poor Small Magellanic Cloud.

Compared to lower redshifts, where the dust attenuation in galaxies is typically found to be characterised well by those properties of local starburst galaxies, the different parameterisation seen in high-redshift galaxy populations means that, for a given line-of-sight dust extinction measured by the UV continuum slope, less far-IR emission is emitted (resulting in a lower UV to far-IR luminosity ratio). However, ALPINE data also show that massive galaxies, i.e., log(M_*/M☉) > 10, have already established a considerable dust content in a short time, less than 1 billion years after the Big Bang. Specifi-
upper limits on the total SFR, our galaxies show a [CII]-SFR relation comparable to the local one, and hence do not suggest a significant evolution of this relation over the last 13 Gyr of cosmic time. This is an important verification that [CII] can be used to trace the total SFR up to $z = 6$.

A non-negligible fraction of light from star formation is already hidden in dust clouds

With ALPINE, we are able to estimate the evolution of the dust-hidden star formation rate density (SFRD) beyond a redshift of 4 with two complementary methods. First, we stacked the far-infrared continuum data (of detected and non-detected ALPINE galaxies) to estimate the dust-hidden SFRD contributed by a UV-selected sample. We find that, even at a redshift of $z = 5.5$, the dust-obscured SFR is significant and contributes up to 50% of the total SFRD for this population alone (Figure 4 and Khusanova et al., in preparation). However, such an analysis does not speak to the total amount of dust-shrouded star formation at these epochs. The second contribution comes from galaxies that are faint or undetected in the UV, but which shine primarily at far-infrared wavelengths. In order to quantify the contribution from this population, we searched for galaxies within our ALMA observations that were detected in the far-infrared continuum (Béthermin et al., 2020; Gruppioni et al., in preparation) and/or the [CII] line (Loiacono et al., in preparation). These populations allowed us to study the contribution of the SFRD by dusty UV-faint galaxies at $z > 4$.

As shown in Figure 4, the SFRD from these ALPINE non-target continuum detections shows a flat distribution over the whole 0.5–6 redshift range, with no significant decrease beyond the Cosmic Noon ($1 < z < 3$). The high value of the SFRD at $z = 5$ is also confirmed by the results obtained for the [CII] serendipitous detections. By combining the results of the two methods, along with the UV-based redshift evolution of the SFRD, it is possible to map the total evolution of the SFRD at these redshifts.

The total SFRD as measured by ALPINE data is in agreement with those from previous far-IR and radio surveys, but higher than those estimated from optical/UV surveys at $z > 3$. The measured difference from the UV results is observed to increase with redshift and peaks at $z = 6$, where the observed excess is about a factor of 10. This discrepancy implies that a considerable amount of star-formation activity at high redshifts is still missed by surveys sampling the UV rest frame. This significant and increasing contribution of dust-obscured activity cannot be recovered, even correcting the UV data for dust-extinction, and rather requires the addition of UV-faint galaxies selected in the IR. Similarly, current galaxy formation models and simulations are not able to predict such a high SFR in dusty galaxies as is observed in the ALPINE sample at early times.

Figure 4. Redshift evolution of the comoving star formation rate density (SFRD), obtained from ALPINE: as in the legend, the orange boxes and black filled circles are the derivations from continuum non-target sources (Gruppioni et al., in preparation), the blue box and blue open triangle are from serendipitous [CII] detections (Loiacono et al., in preparation), while the yellow filled hexagons are the UV-IR SFR estimates based on the ALPINE targets (Khusanova et al., in preparation). For comparison, estimates from other surveys are shown as grey shaded areas and open or filled symbols, as described in the legend. The model from Madau & Dickinson (2014) is shown as a black dashed curve, while the prediction of the IllustrisTNG simulation is shown as a dark-green solid curve obscured at $z > 4$.

The surprisingly complex life of galaxies at the end of reionisation

Star-formation-driven outflows and circumgalactic enrichment

Current models of galaxy formation broadly agree on the key importance of star formation feedback in regulating the evolution of galaxies across cosmic time. Intense episodes of star formation can induce supernova-driven winds that efficiently accelerate interstellar gas. Such processes eventually result in the expulsion of material from the disc, thus inhibiting future star formation and enriching the circumgalactic/intergalactic medium with heavy elements. However, to date, observational evidence for star-formation-driven outflows is still limited to galaxies observed in the local Universe and at intermediate redshifts.

Using the ALMA observations taken as part of the ALPINE survey, which were complemented by the previously described wealth of ancillary photometric/spectroscopic multi-wavelength data, we...
Figure 5. (a) Stacked [CII] spectrum of high-SFR (> 25 $M_\odot$ yr$^{-1}$) ALPINE galaxies, showing broad wings in the high-velocity tails, extending up to $|v| < ~500$ km s$^{-1}$. (b) Circularly averaged radial profile of the stacked [CII] line-core emission in highly star-forming galaxies (blue), extending to diameter scales > 20 kpc, compared with the analogue emission arising from lower-SFR galaxies (green) and the stacked far-infrared continuum. (c) Example of spatial distribution (map on the left and radial profile on the right) of [CII] (red contours and red squares) and rest-frame UV emission (grey-scale image and blue squares) in a galaxy showing a [CII] halo component.

Conducted a number of pioneering studies on the efficiency of galactic feedback and the cycle of gas, dust, and stars in the early Universe.

By stacking the ALPINE [CII] spectra, we discovered signatures of outflows driven by feedback from massive stars. In ALPINE galaxies with high star formation, high-velocity tails of the stacked [CII] line profile revealed gas that was outflowing at velocities of a few hundred km s$^{-1}$, and, in extreme cases, up to about 500 km s$^{-1}$. These winds typically excise gas at a rate similar to the SFR of these galaxies. The relatively small amount of outflowing material suggests that feedback by star formation, while capable of regulating stellar age to some extent, is not the dominant factor in the rapid quenching of high-z galaxies (Ginolfi et al., 2020a and Figure 5a), which necessarily must occur to explain the populations of passive galaxies with relatively old stellar ages observed at $z = 1$–3.

The [CII] line traces the velocity of cold gas in galaxies and is thus a good estimator of the velocity of the galaxy as a whole (also called the systemic redshift). We exploit this property and the available optical spectroscopy to explore the velocity offsets of UV rest-frame ISM lines. We find velocity blueshifts of the ISM lines in the range −500 < $\Delta v$[ISM−[CII]] < 0 km s$^{-1}$, values which are consistent with the [CII] spectral stacking analysis discussed above, and which we interpret as further evidence of ubiquitous galactic outflows (Cassata et al., 2020; Faisst et al., 2020).

In the stacked ALPINE data cubes, we also detected [CII] halos extending on physical scales of > 20 kpc around galaxies with high star formation. Such halos likely originate from metal-polluted circumgalactic gas that was enriched by past outflows (Ginolfi et al., 2020a). We followed up this result by exploring sizes and extended halo structures in and around a subset of individual ALPINE galaxies. We show that sizes of galaxies in [CII] (i.e., their effective radii) are generally larger than those measured in the UV by up to a factor of 4. About one third of galaxies without an observed companion are surrounded by a [CII] halo component that is extended over a scale of 10 or more kpc. Consistent with the stacking analysis described above, ALPINE observations revealed that galaxies that host larger [CII] halos generally have higher stellar masses and star formation rates (Fujimoto et al., 2020 and Figure 6).

Altogether, our findings suggest that star-formation-driven outflows that deposit metal-enriched gas into the circumgalactic medium are highly prevalent in galaxies with higher SFRs at higher redshift, and that such processes already regulate star formation and the cycling of gas, dust and star formation in normal galaxies at very early epochs in the Universe.

The dynamic life of high-redshift systems
Since [CII] is emitted from multiple phases of the ISM, it makes an ideal tracer of the underlying kinematics of a galactic system. As a first step towards characterising the dynamics of the ALPINE galaxies, a group of astronomers within the collaboration examined a set of diagnostic plots (for example, integrated intensity, velocity field, position-velocity diagrams) created from the [CII] data cube of each source. From these diagnostic plots, galaxies were classified as rotators, mergers, extended dispersion-dominated systems (i.e., large galaxies whose kinematics exhibited primarily random motions), or compact dispersion-dominated galaxies. Of the 63 galaxies whose [CII] emission was bright enough to allow for classification, 9 (14%) were classified as rotators, 31 (49%) as mergers, 15 (24%) as extended dispersion-dominated, and 8 (13%) as compact dispersion-dominated (Le Fèvre et al., 2020). This kinematic diversity suggests that galaxies in this early epoch were not a dynamically homogeneous group but had already travelled various evolutionary routes. The large number of merging systems within the ALPINE sample suggests that mass assembly in high-redshift, star-forming galaxies is not only driven by in situ star-formation activity, but aided considerably by ex situ contributions.

For a few of the ALPINE galaxies in which merging activity is detected, the [CII] data allow us to probe the detailed kinematics of the system. In one such
system, we find evidence for an ongoing major merger between two galaxies, with a third minor galaxy nearby (Jones et al., 2020). This configuration suggests that this system is the progenitor of the massive galaxies observed by $z \approx 2$. A different system reveals two major components and two minor components, surrounded by a significant [CII] halo (Ginolfi et al., 2020b), directly revealing the role of mergers in enriching the IGM of massive galaxies. We are currently improving upon this initial qualitative classification method using more systematic and homogeneous criteria (Jones et al., in preparation; Romano et al., in preparation).

Large gas reservoirs of high-redshift systems

For 18 ALPINE galaxies classified as non-mergers and with size measurements, we were able to infer dynamical mass estimates, which probe the sum of the various mass components integrated within the radius at which the dynamics are measured. These dynamical masses, in conjunction with independent measures of the mass of the stellar content from our ancillary data, were used to derive gas masses under the assumption that the relative contribution of dark matter in the internal regions of galaxies is low. The resultant estimated gas masses were then compared to the [CII] luminosities in these ALPINE galaxies to explore whether [CII] is indeed a reliable tracer of gas. We assumed the [CII] derived gas masses to be reliable for all 44 ALPINE non-merger [CII]-detected galaxies and extend the gas mass estimates to this sample to explore the redshift evolution of the molecular gas depletion timescale and the fraction of gas to total baryonic mass for $z > 4.5$. The ALPINE galaxies showed depletion timescales as long as $4.8 \times 10^{9}$ yr at $z \approx 5$, corresponding to star formation efficiencies only twice as high as in present-day galaxies, and a flattening of the gas fraction at $z > 4$, reaching a mean value of 65% at $z \approx 5$. This behaviour of the gas fraction is fully in line with the redshift evolution of the average star formation rate per unit stellar mass at the same redshifts (Khusanova et al., 2020).

Conclusion

The ALPINE surveys have demonstrated the importance of millimeter observations to understanding the nature of normal star-forming galaxies at $z > 4$. This first [CII] sample at $z > 4$ is that statistically representative of normal star-forming galaxies at these redshifts has provided hints of the nature of the complex and dynamic life of such galaxies and raises important new questions. The ALPINE survey will have an important legacy value for both observers and modellers and its products are publicly available to the global astronomical community.1,2

Acknowledgements

This paper is based on data obtained with the ALMA Observatory under Large Programmes 2017.1.00428.L. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

This article is dedicated to the memory of Olivier Le Fèvre.

References

Capak, P. L. et al. 2015, Nature, 522, 455

1 The data products and catalogues will be released on this website: https://cosm.lam.fr/als2c2s/

2 Outreach material is provided on this website: http://alpine.ipac.caltech.edu