Lyman-alpha emission at the end of reionization: line strengths and profiles from MMT and JWST observations at $z \sim 5-6$

Gonzalo Prieto-Lyon, *1,2, Charlotte A. Mason^{1,2}, Victoria Strait^{1,2}, Gabriel Brammer^{1,2}, Rohan P. Naidu⁴, Romain A. Meyer³, Pascal Oesch^{1,3}, Sandro Tacchella^{5,6}, Alba Covelo-Paz³, Emma Giovinazzo³, and Mengyuan Xiao³

- Cosmic Dawn Center (DAWN)
- Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 Copenhagen N, Denmark
- Departement d'Astronomie, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland
- MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, 02139, Massachusetts, USA
- Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK 2
- Cavendish Laboratory, University of Cambridge, 19 JJ Thomson Avenue, Cambridge CB3 0HE, UK

ABSTRACT

Aims. With JWST, it is now possible to use Lyman-Alpha (Ly α) emission from galaxies beyond z > 8 to trace neutral hydrogen in the intergalactic medium (IGM) as the Universe became reionized. However, observed Ly α emission lines are affected by scattering by neutral hydrogen in the IGM and the interstellar and circum-galactic medium, necessitating 'baseline' models of Ly α properties in the ionized IGM to disentangle their impacts. In this work, we characterize Ly α properties at the end of reionization, $z \sim 5-6$,

Methods. We targeted GOODS-N with MMT/Binospec, obtaining R~4360 rest-frame UV spectra of 236 galaxies at $z \sim 5-6$, selected primarily from HST/CANDELS, finding 62 Ly α detections. We use JWST observations from JADES and FRESCO for a subset of our sources to characterize Ly α properties as a function of UV continuum and H α emission. We present the first statistical measurements of the Ly α FWHM distribution at $z \sim 5$ – 6, and produce empirical baseline models of Ly α equivalent width (EW_{Ly α}) and escape fraction ($f_{\rm esc}^{{
m Ly}\alpha}$) conditional on UV magnitude and slope.

Results. We find our EW_{Ly\alpha} and $f_{\rm esc}^{\rm Ly\alpha}$ models strongly depend on UV magnitude, and infer 45±5% and < 62 ± 8% of M_{uv} = -19.5 galaxies have EW_{Ly\alpha} > 25 Å and $f_{\rm esc}^{\rm Ly\alpha}$ > 0.2, respectively. We find a mean FWHM_{Ly\alpha} of 245 km s⁻¹ and median Ly\alpha velocity offset of 258 km s⁻¹, and that both correlate with increasing UV luminosity. Our median observed Ly α line profile is broader and has higher velocity offset compared to pre-JWST models based on $z \sim 2$ lines, which may reflect both resonant scattering by residual neutral hydrogen in the IGM at $z \sim 5 - 6$ and increasing ISM/CGM densities. Our median line profile thus predicts higher Ly α transmission in a fully neutral IGM, providing insight into recent detections of Ly α at z > 10.

Key words. Galaxies: emission lines – Galaxies: high-redshift – Galaxies: evolution

1 Cosmic Dawn Center (DAWN)
2 Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DR
3 Departement d'Astronomie, University of Geneva, Chemin Pege
4 MIT Kavli Institute for Astronomie, University of Geneva, Chemin Pege
5 Kavli Institute for Cosmology, University of Cambridge, Madin
6 Cavendish Laboratory, University of Cambridge, 19 JJ Thomsor
September 24, 2025

ABST

Aims. With JWST, it is now possible to use Lyman-Alpha (Lyα) the intergalactic medium (IGM) as the Universe became reionized by neutral hydrogen in the IGM and the interstellar and circum-gr in the ionized IGM to disentangle their impacts. In this work, we providing a baseline that can be applied to z > 6 observations.

Methods. We targeted GOODS-N with MMT/Binospec, obtaining selected primarily from HST/CANDELS, finding 62 Lyα detection subset of our sources to characterize Lyα properties as a function measurements of the Lyα FWHM distribution at z ~ 5 - 6, and properties and escape fraction (f_{esc} conditional on UV magnitude and slope Results. We find our EWLyα and f_{esc} models strongly depend on galaxies have EWLyα > 25Å and f_{esc} on 2.2, respectively. We find a 258 km s⁻¹, and that both correlate with increasing UV luminosity velocity offset compared to pre-JWST models based on z ~ 2 lin hydrogen in the IGM at z ~ 5 - 6 and increasing ISM/CGM densing in a fully neutral IGM, providing insight into recent detections of I Key words. Galaxies: emission lines – Galaxies: high-redshift – (1) in the intergalactic medium (IGM) became ionized due to the intergalactic medium (IGM) decame ionized due to the intergalactic medium (IGM) became ionized due to the intergalactic medium (IGM) decame ionized due to the intergalactic medium (IGM) decame ionized due to the intergalactic medium (IGM) decame ionized due to the intergalactic m dent studies have reached consensus that the Universe was fully reionized at $z \sim 5$ - 6 (e.g. Fan et al. 2006; Becker et al. 2015; Eilers et al. 2018; Yang et al. 2020; Qin et al. 2021; Bosman et al. 2022; Gaikwad et al. 2023), but likely predominantly neutral at $z \sim 7 - 8$ (e.g., Stark et al. 2010; Treu et al. 2012; Pentericci et al. 2014; Schenker et al. 2014; Mason et al. 2018b; Davies et al. 2018; Yang et al. 2020; Jung et al. 2020; Bolan et al. 2022; Umeda et al. 2024; Nakane et al. 2024; Tang et al. 2024b; Jones et al. 2024; Mason et al. 2025). The earliest stages of reionization at z > 9 are still very poorly constrained, but will be crucial for understanding the build up of the first galaxies and the higher

than expected UV luminosity density detected by JWST (Gelli et al. 2024; Muñoz et al. 2024; Whitler et al. 2025).

Lyman-alpha (Ly α) emission from galaxies is one of our best probes of the early stages of the EoR. Its high cross-section for damping wing scattering by neutral hydrogen means Ly α flux traces the IGM opacity (e.g., Miralda-Escude 1998; Mesinger & Furlanetto 2008) and we are now detecting 1000s of z > 8galaxies with JWST (e.g. Adams et al. 2022; Donnan et al. 2024; Whitler et al. 2025), providing large samples to comprehensively trace the IGM at high redshift. In contrast, the number density of quasars - the typical IGM probes at lower redshifts -(e.g., Jiang et al. 2016; Matsuoka et al. 2018; Wang et al. 2020; Schindler et al. 2023), decreases significantly at z > 7 (Euclid Collaboration et al. 2019). The visibility of Ly α from galaxies drops considerably at z > 6, implying an increasing neutral IGM (Seen in the declining fraction of galaxies detected with strong (EW>25 Å) Lyα, Stark et al. 2010; Pentericci et al. 2014; De Barros et al. 2017; Jung et al. 2019; Fuller et al. 2020; Laporte et al. 2021; Tang et al. 2024b). JWST has opened a new window on reionization studies – enabling Ly α observations in faint galaxies at $z \ge 6$ without the atmospheric and sensitivity limitations of ground-based telescopes. Early JWST results have

^{*} E-mail: gonzalo.p.lyon@gmail.com

confirmed the downturn in Ly α emission at $z \gtrsim 6$ seen from the ground and demonstrated strong Ly α is extremely rare at z > 8 in current samples (Nakane et al. 2024; Chen et al. 2024; Tang et al. 2024b; Jones et al. 2024; Kageura et al. 2025), and have also provided the first detections of Ly α at z > 10 (Bunker et al. 2023; Witstok et al. 2025a). These observations demonstrate JWST will be a key tool to understand the earliest stages of reionization.

However, the interpretation of Ly α observations during the Epoch of Reionization rests on understanding the transmission of Ly α on multiple spatial scales. Ly α photons are subject to scattering not only by neutral hydrogen in the IGM, but also in the interstellar and circumgalactic media (ISM and CGM). To use Ly α to track the reionization history, an understanding of Ly α emerging from the ISM/CGM is necessary. In particular, because of the strong wavelength dependence of the IGM Ly α damping wing (Miralda-Escude 1998), both an understanding of the Ly α line strength and *spectral shape* emerging from the ISM/CGM is important to disentangle the impact of the IGM (Dijkstra et al. 2011; Mason et al. 2018a; Endsley et al. 2022).

To this end, samples of galaxies at the end of the EoR, z ~ 5 - 6, have been commonly used as a baseline for interpreting higher redshift observations (Stark et al. 2011; De Barros et al. 2017). These galaxies are expected to be physically similar to those at earlier epochs, but suffer minimal damping wing attenuation by the neutral IGM. Therefore, these galaxies provide a window into the emergent Ly α emission from the ISM and CGM, and have been used to infer the additional impact of the IGM on z > 6 galaxies' Ly α (e.g., Mason et al. 2018b; Whitler et al. 2020; Jung et al. 2020; Bolan et al. 2022). However, these baseline samples had been limited to the most readily available observables pre-JWST – Ly α EW and UV photometry, and, with only a handful of systemic redshifts known from strong rest-UV or FIR lines (e.g., Stark et al. 2017; Bradač et al. 2017): our knowledge of the emerging Ly α line profile had been severely limited. These and other limitations, such as the unknown Ly α reionization analyses, have led previous models to rely on empirical relations based on $z \sim 0 - 2$ galaxy samples (e.g., Mason et al. 2018b, 2019; Hayes et al. 2023), where the rest-frame optical is accessible from the ground. However, it is unclear whether these samples are good analogs of reionization-era sources. For example, recent works have shown the Ly α velocity offsets of z~6 galaxies appear systematically higher than models derived from $z \sim 2$ galaxies (Prieto-Lyon et al. 2023b; Tang et al. 2024a).

JWST's ability to observe the rest-frame optical is transforming our ability to understand galaxies in this era. Thanks to $R \ge 1000$ -resolution observations, such as NIRCam grism data, we now have accurate systemic redshifts for 1000s of z > 5galaxies (e.g., Oesch et al. 2023; Covelo-Paz et al. 2025; Meyer et al. 2024) and measurements of Ly α escape fractions ($f_{\rm esc}^{{\rm Ly}\alpha}$), relative to Balmer lines, and velocity offsets $(\Delta v_{Ly\alpha})$ are now routinely being measured for $z \sim 5 - 9$ galaxies (e.g., Lin et al. 2024; Chen et al. 2024; Tang et al. 2023, 2024a; Saxena et al. 2024; Chen et al. 2025). Tang et al. (2024a) recently presented constraints on the emergent Ly α EW and escape fraction distributions in bins of M_{uv} , UV β slope and [OIII]+H β EW, and velocity offset distributions in $z \sim 5 - 6$ galaxies. These observations have provided a more robust baseline for predicting emergent Ly α at z > 6 and thus interpreting observed Ly α in the context of reionization (Tang et al. 2024b).

However, these works have mostly focused on measuring $Ly\alpha$ emission distributions in bins of galaxies properties, limiting the information used in the conditional distributions, and less attention has focused on the shape of the $Ly\alpha$ line profile at

 $z \sim 5-6$. Accurately measuring the emergent lineshape is critical for understanding the transmission of Ly α through the IGM (Dijkstra & Wyithe 2010; Mason et al. 2018b, 2019; Endsley et al. 2022; Tang et al. 2024c; Mukherjee et al. 2024). High resolution spectroscopy is required to accurately measure both the asymmetric shape of Ly α , and its offset from the systemic redshift $(\Delta v_{\rm Ly\alpha})$. Resolution below ~100 km s⁻¹ ($R \gtrsim 4000$) is optimal, as typical full-width half-max (FWHM) in Ly α -emitters in the ionized IGM (z < 5) range from $\sim 100 - 400 \,\mathrm{km \, s^{-1}}$ (Verhamme et al. 2018; Tang et al. 2024c) and values of $\Delta v_{Lv\alpha}$ range from $150 \,\mathrm{km}\,\mathrm{s}^{-1}$ to $500 \,\mathrm{km}\,\mathrm{s}^{-1}$ (e.g., Prieto-Lyon et al. 2023b; Tang et al. 2024c). Tang et al. (2024a) demonstrated that $z\sim 5-6$ Ly α line profiles appear more offset from systemic than profiles for sources with comparable EW_{Ly α} at $z \sim 2-3$ (Tang et al. 2024c), which they discuss may be due to residual neutral gas in the ionized IGM resonantly scattering Ly α close to line center and/or increased scattering in (relatively dust-poor) ISM and CGM. But we currently have no statistical information about the broadness and asymmetry of Ly α line profiles at these redshifts, which provides additional information to understand the evolution in line profiles, primarily due to the moderate to low resolution of Ly α spectroscopy in fields overlapping with deep JWST spectroscopy (i.e. VLT/MUSE R ~ 3000). Additionally, JWST spectroscopic observations using NIRSpec/PRISM have insufficient resolution (R \sim 100) to accurately measure Ly α velocity offsets. By contrast, NIRCam grism spectra, with a higher resolution ($R\sim1600$), is better suited for studies such as this one.

To that end, we have carried out a deep, high-resolution $(R \sim 4400)$ near-infrared spectroscopic survey of 236 z ~ 5 - 6.5 Lyman-break galaxies in the GOODS North field (Giavalisco et al. 2004), selected from deep HST imaging, with MMT/Binospec (Fabricant et al. 2019), with the aim of constraining the Ly α strength, and line profiles emerging from the ISM/CGM of galaxies in the post-reionization era. MMT/Binospec, with its high resolution (R~4360), high throughput, and wide on-sky coverage (two 8' x 15' fields of view), offers great advantage for Ly α studies, surpassing the resolution of most $z\sim5-6$ detections in the literature (e.g., LRIS, MUSE, VIMOS, JWST, R<3500). Our sample includes 62 high-resolution Ly α detections. Thanks to the excellent ancillary datasets in GOODS-N, we supplement our Ly α spectroscopy with JWST/NIRCam slitless spectra from FRESCO + CONGRESS (Oesch et al. 2023; Egami et al. 2023), and deep HST and JWST photometry from the CANDELS (Giavalisco et al. 2004; Grogin et al. 2011; Koekemoer et al. 2011) and JADES surveys (Eisenstein et al. 2023) to obtain precise measurements of galaxy properties (UV magnitudes, UV beta slopes, systemic velocities, Ly α escape fractions: $f_{\rm esc}^{{\rm Ly}\alpha}$). With these data we build two new empirical models for EW_{Ly α} and $f_{\rm esc}^{{\rm Ly}\alpha}$ conditional on UV observables, M_{UV} and β , which can be used as a baseline for constraining the reionization history. We also explore the properties of Ly α lineprofiles ($\Delta v_{Lv\alpha}$, FWHM, asymmetry) in our sample.

The paper is structured as follows. In Section 2 we describe our target selection, the MMT/Binospec spectroscopy, and ancillary datasets. In Section 3 we describe our Ly α and rest-optical emission lines measurements. We describe our Ly α measurements in the context of other galaxy properties in Section 4. We present our new empirical model for emergent EW_{Ly α} and $f_{\rm esc}^{\rm Ly}$ in Section 5. We discuss our results in 6 and present our conclusions in Section 7. We assume a flat Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_{\Lambda} = 0.7$, h = 0.7. All magnitudes are in the AB system.

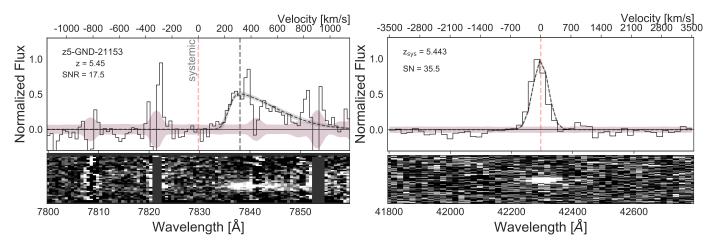


Fig. 1: Spectra of z5-GND-21153. Left: MMT/Binospec spectra of Ly α . Right: H α counterpart from JWST/NIRCam slitless spectra of the same source. We plot the emission line fit with its corresponding 1σ uncertainty as a gray shaded area. On the left, the vertical red line shows the systemic/H α redshift, in this example there is a 315 km s⁻¹ velocity offset between Ly α and H α . The upper axis shows velocity centered at the peak of the systemic line. In the bottom are the 2D spectra.

2. Data

Here we describe the observational datasets used in this work. First, we describe our target selection (Section 2.1) and the MMT/Binospec spectroscopic observations of z \sim 5 - 7 Lyman Break Galaxies in the GOODS-N field (Section 2.2). Second, we describe the JWST/NIRCam slitless spectroscopy from the FRESCO survey¹ which provides rest-frame optical spectroscopy of a subset of our targets (Section 2.3). Finally we describe the JWST and HST photometric datasets from the DJA² (The DAWN JWST Archive) photometric catalog of JADES³ (Eisenstein et al. 2023), the CANDELS/SHARDS (Giavalisco et al. 2004; Grogin et al. 2011; Koekemoer et al. 2011) catalog by Barro et al. (2019) and 3D-HST⁴ (Skelton et al. 2014; Brammer et al. 2012) imaging in Section 2.4. Our identification of emission lines is described in Section 3.

2.1. Target Selection

Our primary targets are 236 $z_{\rm phot} \sim 5-7$ candidates in GOODS-N (Giavalisco et al. 2004) selected by Lyman-Break color criteria from Bouwens et al. (2015). All candidates are detected with a magF160W<27.5. ACS/WFC & WFC3/IR filters from the deep GOODS imaging were used for tracing the Lyman-Break and the rest-frame UV continuum, and two low wavelength Spitzer/IRAC bands break the degeneracy between the Lyman-Break and 4000Å break. The 236 sources include 24 spectroscopically confirmed Ly α -detected galaxies previously reported by Hu et al. (2010), Stark et al. (2011), and Jung et al. (2018), in order to obtain higher resolution, higher S/N observations of their Ly α lineshapes.

2.2. MMT / Binospec

We performed observations of our targets in GOODS-N with MMT/Binospec (Fabricant et al. 2019) from 2019 to 2021. The observations are divided into 4 masks with two detectors of 8'x15' field of view, covering a total footprint of 0.27 deg², with

exposure times between 10 - 16 hours per mask. We used the 600 lines/mm grating, providing $R{\sim}4360$ and a wavelength range between 7255 - 9750 Å, enabling Ly α detections between z=4.9 to 7.0. Thanks to the high resolution per pixel, 0.61 Å/pix and instrumental resolution of $\approx 70~{\rm km\,s^{-1}}$, we are able to accurately retrieve the profile shape of the observed Ly α lines, improving on e.g., JWST NIRSpec and VLT/MUSE Ly α observations ($R \sim 3000$) by a factor of 45%. In Figure 1 we show an example of our Ly α spectra. Properties of each mask are described in Table 1. All MMT/Binospec spectra discussed in this work is publicly available 5 .

2.3. JWST / NIRCam - Slitless Spectra

We include NIRCAM/grism observations from FRESCO (Oesch et al. 2023) which targeted 62 arcmin² in GOODS-N, covering 82/236 of our MMT/Binospec targets. The wavelength range with the F444W filter covers $3.8-5\mu m$, allowing the observation of the most luminous optical lines, $H\alpha$ and the [OIII] doublet, from $z\sim4.8-6.6$ and $z\sim6.6-9$ respectively, with $R\sim1600$ $\approx 180 \, \text{km s}^{-1}$. This provides systemic redshift for a subset of our targets, with detected H α and [OIII] line in FRESCO, with an average 5σ flux limit $\gtrsim 2 \cdot 10^{-18} \text{erg s}^{-1} \text{ cm}^{-2}$ (see Section 3.2) for description of the line detections). H α flux measurements also provide an estimate of the intrinsic Ly α flux, enabling us to estimate Ly α escape fractions. The reduction and spectral extraction of FRESCO data is detailed in Oesch et al. (2023). The accuracy of the wavelength calibrations is crucial due to Ly α velocity offsets near the instrumental resolutions (See Figure 1). To verify the calibration, we compare the systemic redshift of FRESCO sources with known ground-based spectra, and find no substantial systematic differences, with a scatter in the order of the wavelength pixel resolution. We show the spectra of all galaxies with a Ly α and optical line detection in Appendix B.

2.4. Photometric Data

Deep HST and JWST photometry of our candidates is provided by CANDELS (Giavalisco et al. 2004; Grogin et al. 2011; Koekemoer et al. 2011), 3D-HST (Skelton et al. 2014), and

¹ https://jwst-fresco.astro.unige.ch

² https://dawn-cph.github.io/dja/

³ https://jades-survey.github.io

⁴ https://archive.stsci.edu/prepds/3d-hst/

⁵ https://sid.erda.dk/sharelink/FCZBDRW6O2

JADES (Eisenstein et al. 2023) imaging in GOODS-N. These data provide rest-frame UV coverage of our targets. CANDELS and 3D-HST provide imaging in ACS/WFC: F435W, F606W, F775W, F814W and F850LP; and WFC3/IR: F105W, F125W, F140W and F160W. JADES JWST/NIRCam imaging covers 128/236 of our targets in F090W, F115W, F150W, F200W, F277W, F335M, F356W, F410M, and F444W.

Inside the JADES FoV, we use the data reduction from DJA processed with *grizli* (Brammer 2023) which includes both legacy HST ACS/WFC & WFC3/IR bands and NIRCam imaging from JADES reduced self-consistently (Valentino et al. 2023) (hereafter, JADES-DJA). For sources outside of the FoV of JADES, we use ACS/WFC & WFC3/IR photometry catalogs from Barro et al. (2019), which provide consistent photometry to the Valentino et al. (2023) catalogs. In cases where we do not find a match in either catalog, we use the 3D-HST catalog (Skelton et al. 2014). For all catalogs we use aperture corrected fluxes.

2.5. Binospec Data Reduction

In the following section, we describe how we reduced our Binospec data applying a telluric calibration, a flux calibration and additional sky subtraction. The initial reduction and sky subtraction was done by the Binospec Pipeline ⁶. To optimize the reductions for our faint emission line targets we further reduced each individual exposure. None of our targets showed evidence of a UV continuum detection.

To reduce features due to atmospheric emission such as the O2 A& B bands, we perform a telluric calibration of our sources. We follow a similar method to Vacca et al. (2003). On each mask we observe two standard F-type stars. We compare the observed star in each exposure to F-star synthetic models using the $SEDPy^7$ library. We use the synthetic models that best fits the spectrum of our standard stars. We extract and normalize the spectra to obtain the telluric transmission $T(\lambda)$ by dividing the data by the model (both continuum normalized). We apply the transmission function to all spectra in each exposure:

$$f_{corr}(\lambda) = \frac{f_{obs}(\lambda)}{T(\lambda - \alpha)^{\beta}}$$
 (1)

The constant α accounts for any wavelength shifts between the spectra and transmission and β is used as a scaling factor of the transmission. To calculate α and β we minimize the RMS in the two most important telluric regions, O2 A& B bands in $f_{corr}(\lambda)$ for every exposure for every slit containing a high redshift target.

We perform flux calibration for each exposure by calibrating our standard stars with magnitudes obtained in the SDSS (Abdurro'uf et al. 2022). We use the i filter of Sloan since it covers a similar wavelength range as the MMT/Binospec data. We convolve our standard star spectra with the corresponding i_{sdss} filter using SEDPy.

We perform additional sky subtraction of the exposures for each mask, to improve on the pipeline sky subtraction. We make a global 1D sky spectrum for each individual exposure, $S(\lambda)$, using the standard deviation of flux in all slits containing z>5 targets with no clear emission lines (as our targets are faint, we do not expect to detect the continuum, so the 'empty' slits should measure the sky), using sigma clipping to remove any cosmic rays. We perform additional sky subtraction to each spectrum per exposure as:

$$f_{\text{skysub}}(\lambda) = f_{\text{obs}}(\lambda) - A \cdot S(\lambda)$$
 (2)

We find A by minimizing the RMS of $f_{\rm skysub}(\lambda)$ per slit's exposure given the $S(\lambda)$ measured in each mask, so that the distribution of signal in an empty region of the slit is centered around $\frac{S}{N} \sim 0$. For this we use LMFIT (non-linear least-squares minimization)⁸.

We finally merge the telluric corrected, flux calibrated and sky subtracted exposures into the final science spectra, using sigma clipping to remove cosmic rays. Exposures with seeing above 1.2" are excluded. Table 1 provides an overview of the exposure times, average seeing and targets for each of our four masks.

3. Lyman-Alpha and rest-frame optical line identification

In the following section we describe our search for Ly α emission in our MMT/Binospec spectra (Section 3.1), and H α and/ or [OIII] emission in the FRESCO JWST/NIRCam slitless spectra (Section 3.2).

3.1. Lyman-alpha emission line detection

For the case of spectroscopically confirmed galaxies (either from rest-optical lines or existing Ly α spectroscopy), extracting the Ly α profile is straightforward since the redshift is already known. For sources without spectroscopic confirmation we perform a systematic search for the emission line given the photometric redshift.

We first collapse the 2D spectra into 1D spectra for each target by extracting within the median spatial extension of our Ly α detections (= 1.7" = 7 pixels). We maximize the signal captured in the extraction, with more extended extractions the SNR plateaus adding mostly noise. Starting at the position of the UV continuum, we repeat this process for every spatial position of the slit within 2.4" of the UV coordinates to account for any spatial offset between the UV continuum and Ly α emission (Hoag et al. 2019; Lemaux et al. 2021). We then scan through the 1D spectra to search for high SNR peaks, scanning all wavelengths. We flag line candidates where SNR≥3 is found in 3 consecutive wavelength pixels, essentially an integrated SNR \geq 5 detection. Finally, we manually inspect the candidate lines in the 1D and 2D spectra to remove any false positives near the sky lines or from randomly high peaked background. For galaxies where we have a systemic redshift from the rest-frame optical lines, we do a manual search of Ly α within ~2000 km s⁻¹ of the systemic redshift. For detected emission lines we measure Ly α fluxes from a fitted asymmetric Gaussian (Section 3.1.3) as this allows for more accurate measurements in the case of strong skylines near the emission line than obtaining the flux directly from the 1D spectra.

3.1.1. Completeness

To evaluate the completeness of our emission line search we perform Monte Carlo simulations. We add simulated Gaussian $Ly\alpha$ profiles at random wavelengths into the slits of non-detected targets and then attempt to recover them with the method described above. We vary the line flux and FWHM as these are the most important factors in determining the detection of a line, and show the resulting completeness in Figure 2. The asymmetry of the profile is sub-dominant to the flux and FWHM so we use a Gaussian profile for simplicity. The presence of sky lines plays an

⁶ http://mingus.mmto.arizona.edu/ bjw/mmt/binospec_info.html

⁷ https://github.com/bd-j/sedpy

⁸ https://lmfit.github.io/lmfit-py/

	Exposure Time	Median Seeing [†]	Flux Limit (5σ)	Targets	Ly α detections
	[hr]		[erg s ⁻¹ cm ⁻²]		
Mask 1	15.5	0.90"	1.26×10^{-17}	93	30
Mask 2	13.0	0.85"	1.26×10^{-17}	80	17
Mask 3	13.3	0.92"	1.35×10^{-17}	60	12
Mask 4	7.8	0.97"	1.63×10^{-17}	50	7

Table 1: Overview of Binospec masks and $z \sim 5 - 7$ targets (Section 2.2). † After excluding exposures with seeing > 1.2".

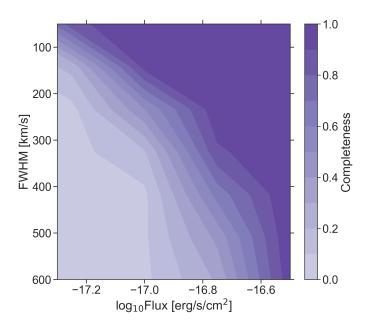


Fig. 2: Here we show the completeness of our Ly α search as a function of observed FWHM and flux. The fraction of recovered detections drops with decreasing flux and increased line broadness.

important role, retrieval is unlikely if an emission line falls at the same wavelength, leading to high redshift line recovery to be even harder due to the increased sky line emission at longer wavelengths.

Figure 2 shows that Ly α line flux and FWHM play a crucial role in completeness. We achieve > 50% completeness for lines with FWHM=100 km s⁻¹ and flux $\gtrsim 10^{-17.2}$ erg s cm⁻², while lines broader than 500 km s⁻¹ need a flux of $\gtrsim 10^{-16.7}$ erg s cm⁻² to reach the same completeness.

3.1.2. Slit Losses

As Ly α can be spatially extended beyond the width of our extraction window (1.7") we correct our Ly α fluxes for slit losses following the approach of Tang et al. (2024a). We first produce a 2D Ly α surface brightness profile following for typical z~5-6 galaxies as measured by MUSE (Leclercq et al. 2017): consisting of two declining exponential functions to describe a Ly α core and an extended Ly α halo, with 0.3 kpc and 3.8 kpc exponential decay constants respectively. We allocate 35% and 65% of the total flux to the core and halo respectively. We convolve the resulting model with the PSF of each mask, and quantify the flux that would be obtained following the Ly α extraction (7pix ~ 1.7").

For all sources, we estimate 65-70% of the total Ly α flux is captured by our extractions in the Binospec slits. We take this into account for any calculations involving these quantities. We calculate the slit loss directly for individual sources, and use the median of 67% slit loss for our non-detection upper limits. We note that we assume each galaxy has the same surface brightness profile, therefore the only difference between galaxies will be the redshift and following angular size conversion.

3.1.3. Ly α fluxes and upper limits

For galaxies with Ly α detections we obtain line fluxes by fitting the lines with an asymmetric Gaussian profile, this enables us to account for the presence of sky lines. Thanks to the $R \sim 4000$ resolution of our spectra, we are able to observe detailed Ly α line profiles for our detections. We fit a skewed Gaussian model to our Ly α lines:

$$f(x; A, \xi, \omega, \alpha) = \frac{A}{\omega \sqrt{2\pi}} e^{[-(x-\xi)^2/2\omega^2]} \left\{ 1 + \text{erf}\left[\frac{\alpha(x-\xi)}{\omega \sqrt{2}}\right] \right\}$$
(3)

The location is represented by ξ . The skewness is parameterised by α . The scale is ω and it describes the standard deviation with a factor of skewness. For example, a symmetrical Gaussian profile will have $\alpha=0$, while for a completely red-sided asymmetric profile $\alpha\to +\infty$. When fitting we fix $\alpha\geq 0$ as we expect only profiles which are truncated on the blue side due to the resonant scattering of Ly α photons (Laursen et al. 2011). The flux of the line is then obtained by integrating the resulting line profile. To be consistent with the literature, we extract the FWHM by its definition, by measuring the width of the profile at half its maximum and subtract the instrumental resolution in quadrature.

We fit the line profiles using an MCMC with the python library $emce^9$ (Foreman-Mackey et al. 2013). As the skewness, is most sensitive to $\alpha \sim 0-3$ we set a log uniform prior on the skewness between $\alpha \in [0,15]$, allowing perfectly Gaussian to completely asymmetric profiles. For all other parameters of the fit, amplitude,location and scale, we use uniform priors with positive values. By setting 30 walkers with 2000 steps, with a burnin of 750 steps, we obtain converged distributions for our parameters. We use the chains of the MCMC to measure the FWHM posterior distribution. In Appendix B we show some examples of the spectra and the resulting fits.

For the galaxies which do not show a Ly α emission feature, we calculate flux upper limits using the following procedure. We add simulated emission lines to the 2D spectra of our sources, considering a grid of flux, FWHM and redshift. We take uniformly distributed values of flux between $0.1-3.2\times10^{-17}\,\mathrm{erg/s/cm^2}$, and FWHM values between $100-400\,\mathrm{km/s}$,

⁹ https://emcee.readthedocs.io/

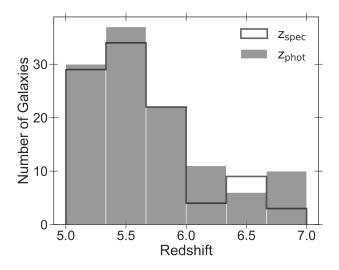
which span the range we find most of our detections (see Section 4). To account for the presence of skylines and the photometric redshift we take uniformly distributed values of redshift within $z_{phot} \pm 0.5$. We generate fully asymmetric Gaussian profiles and run the line selection code (Section 3.1) to detect SNR>5 simulated lines.

We simulated 500,000 lines for each galaxy across the above grid of parameters. We calculate the minimum detectable line flux for every sampled FWHM and redshift. We take the median of the resulting flux distribution as the $Ly\alpha$ flux limit for the galaxy.

3.2. Rest-frame optical emission line detection

82/236 of our Binospec targets fall into the footprint of JWST/NIRCam grism slitless spectroscopy from FRESCO (Section 2.3), providing rest-frame optical emission line coverage. We perform an initial search for emission lines based on the grizli extractions (Brammer & Matharu 2021) and refine the search around the Ly α redshift for Ly α detections or photometric redshift for Ly α non-detections. We require a total SNR>5 extracted over the median FWHM of the detections (385 km s⁻¹) as a detection limit for rest-frame optical lines. Due to the nature of the slitless spectra, contamination from emission lines of overlapping galaxies is present in some of our sources, this effect is mostly taken into account by the extraction as described by Oesch et al. (2023). In cases where there is a misidentification between the Ly α of our target and emission lines of a foreground source, we look for extra emission lines to discard the foreground source, as these do not match with the z_{spec} .

For the 22 galaxies with detected Ly α emission which fall in the FRESCO footprint, we do a manual search for SNR>5 restoptical emission lines within $\sim 2000 \,\mathrm{km \, s^{-1}}$ of the Ly α redshift 10 . We detect either H α or [OIII] for 14/22 galaxies. For the remaining 60 targets with no Ly α detection in Binospec we search over the complete FRESCO wavelength range 3.835 - 5.084 micron for any optical emission lines, focusing mostly within $z_{phot} \pm 0.5$. We detect H α or [OIII] for 56/82 galaxies. The M_{UV} distribution of unconfirmed galaxies is slightly fainter ($M_{uv} = -20.0$) than confirmed galaxies (M_{uv} =-20.2). Using the systemic redshifts obtained from JWST/FRESCO, we go back to the rest-frame UV data of MMT/Binospec. We replace the photometric redshifts for systemic redshifts and perform a new search for Ly α , and find 2 new detections that we failed to recover in our previous searches due as they are on the threshold of our SNR limit (\sim 5). We also discarded 3 low SNR (\sim 5) Ly α line candidates due highconfidence H α detections in FRESCO which made the Ly α redshift implausible. We compared our optical line identification to the catalogs by Covelo-Paz et al. (2025) using FRESCO (Oesch et al. 2023) and CONGRESS (Egami et al. 2023) NIRCam grism data. Using the Covelo-Paz et al. (2025) FRECSO catalog we confirmed an additional 3 H α detections for our sample. We did not find any emission lines for our sample in the Covelo-Paz et al. (2025) CONGRESS catalogs between 3.1 – 3.8 µm, finding no evidence to suggest lower photometric redshifts (zphot) for our Ly α non-detections, or an incorrect spectroscopic redshift (zspec) for our low-SNR Ly α detections without > 3.8 μ m optical line detections.



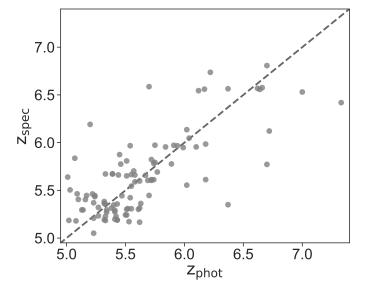


Fig. 3: Top: Redshift distribution. The final sample consists of 236 galaxies. 105 sources have a spectroscopic redshift, from either Ly α and/or H α or [OIII] (black line). 13 sources have both Ly α and H α , and 1 source has Ly α and [OIII]. 131 sources lack spectroscopic confirmation but have $z_{\rm phot} \sim 5-7$ and Ly α flux limits from Binospec (gray shaded histogram). Bottom: spectroscopic redshift ($z_{\rm Ly}$ α or $z_{\rm sys}$) and photometric redshift comparison.

In Figure 3 we show the redshift distribution of the sample. We have a final number of 236 galaxies. We are able to obtain a spectroscopic redshift for 105 galaxies via three approaches: Ly α , H α and [OIII]. Out of the 105 sources, 62 have Ly α emission, 50 have H α emission and 1 has [OIII]. Out of these, 13 have Ly α + H α and 1 has Ly α + [OIII]. For 131 targets no emission line was found. We find a median redshift of 5.6 in both the spectroscopic and photometric redshift samples. The median difference $|z_{\rm spec}-z_{\rm phot}|\approx 0.15$.

To obtain line fluxes, and centroids we fit a Gaussian line profile to the detected rest-optical emission lines. We show the resulting 1D and 2D spectra for a sub-set of rest-frame optical

 $^{^{10}}$ Ly α velocity offsets are typically $\sim 100-500\,\mathrm{km\,s^{-1}}$, with rare instances of offsets up to $1000-2000\,\mathrm{km\,s^{-1}}$ (e.g. Erb et al. 2014; Steidel et al. 2014; Tang et al. 2023; Bunker et al. 2023; Prieto-Lyon et al. 2023b)

lines and the fitted models in Appendix B. We calculate the flux limits for non-detections with a SNR=5 limit, given the 1D noise array with an extraction the size of the median $H\alpha$ FWHM (= 375 km s $^{-1}$) at $z_{Ly\alpha}$ or z_{phot} . The observed $H\alpha$ FWHM of our galaxies is 2-2.5 times higher than reported with JWST/NIRSpec in Prieto-Lyon et al. (2023b). These increased line broadening is due to JWST/NIRCam slitless spectra data having the spatial extension of the galaxy convolved with the broadness of the emission lines. Since the spatial extension dominates the observed FWHM of $H\alpha$ and [OIII], recovering the intrinsic broadness is challenging for resolved sources, and we leave this to future work.

4. Lyman Alpha and Galaxy Properties

The combination of high resolution Ly α spectroscopy, deep HST and JWST photometry, and rest-frame optical JWST spectroscopy allows us to characterize Ly α emission in our $z\sim 5-6$ sample via measurements of Ly α EW, line profiles and velocity offsets, and escape fractions. In the following section we describe these properties in our sample and investigate trends between Ly α properties and galaxy properties to better understand Ly α transmission in the ISM and CGM in our sample.

4.1. UV Continuum properties

We calculate the UV absolute magnitude (M_{uv}) and UV slope (β) for our targets using HST and JWST photometry as described in Section 2.4. We include only filters that have their effective width (W_{eff}) within the rest-frame UV between 1250-2600 Å so as to not include bands containing Ly α flux. We obtain all filter information from Rodrigo & Solano (2020). For the redshift, we use z_{sys} when available. If z_{sys} is unavailable but Ly α is detected, we use $z_{Ly\alpha}$. If neither are available, we use z_{phot} . We note the choice of redshift has negligible effect on our M_{uv} and β results.

We fit a power law, $f_{\lambda} \propto \lambda^{\beta}$, to the selected photometric filters that capture the UV continuum. We retrieve M_{UV} by evaluating the fit at 1500 Å. For sources with JWST NIRCam data, the M_{UV} and β estimates are consistent with the HST estimates, but we obtain lower uncertainties. We fit the continuum using an MCMC method assuming a Gaussian likelihood for the flux densities.

We plot the resulting M_{uv} and β for our sample in Figure 4, showing the median and 1σ uncertainties. Similarly to previously reported trends in the literature (e.g. Bouwens et al. 2012, 2014) we see β increase as galaxies are brighter in UV magnitude, but this correlation becomes unclear for fainter sources. We find our F160W selection is incomplete for $M_{uv} >$ -19.5 sources with blue (β <-2) UV slopes, which we discuss later in Section 6.1

Overall we see no significant difference in the M_{uv} distributions between galaxies with and without Ly α detection, with median values of $M_{uv} = -20.0$ and -20.1 respectively. We find a slight difference in the β slope distributions for galaxies that had a Ly α detection and galaxies that do not: galaxies with Ly α detection show bluer β slopes, with a median value of β = -2.0, compared to β = -1.8 for galaxies without Ly α detections. Performing a KS test shows a p-value>0.1, implying a weak difference between the β slope distribution of galaxies with and without detected Ly α .

4.2. Lyα Equivalent Widths

To calculate Ly α rest-frame equivalent widths (EW), we compare the Ly α flux, or upper limit, with the flux density of the UV continuum. We estimate the continuum flux density, $f_{\lambda,UV}$, at the Ly α wavelength, 1215.6 Å, using our power-law fit to the UV continuum (Section 4.1). The rest-frame EW is:

$$EW_{rest} = \frac{F_{Ly\alpha}}{f_{\lambda UV}(1+z)} \tag{4}$$

Where $F_{Ly\alpha}$ is the total flux of the $Ly\alpha$ line in erg s^{-1} cm⁻² (or upper limit) and $f_{\lambda,UV}$ the flux density [erg s^{-1} cm⁻² Å⁻¹] of the UV continuum at 1215.6Å. As our sources are faint and the continuum is not detected in the spectra, we do not perform continuum subtraction. For sources with only photometric redshifts we use z_{phot} as the redshift.

In Figure 5 we show our $EW_{Ly\alpha}$ measurements against the 2 UV observables. We mainly find enhanced $EW_{Ly\alpha}$ in UV faint galaxies, while a trend with UV slope is not evidently present. We also show the average 5σ flux limit of our observations. Since the line flux limit is related to $Ly\alpha$ broadness, it is possible to observe lines below this limit if they are narrower than the median $FWHM_{Ly\alpha}$ of our sample.

4.3. Ly α line profiles

Thanks to the $R \sim 4000 \ (\approx 70 \ \text{km s}^{-1})$ resolution of the Binospec spectra and high S/N of many of our detections, we are able to observe detailed Ly α line profile shapes, which offer insights into the scattering of Ly α in the ISM and CGM (e.g., Neufeld 1991; Verhamme et al. 2006; Gronke et al. 2015). At $z \sim 0.2 - 2$, the strongest Ly α emitters tend to show narrow Ly α emission lines with flux close to the systemic velocity, implying Ly α is scattering in low column density neutral gas and thus not diffusing significantly in frequency space (e.g., Erb et al. 2014; Hashimoto et al. 2015; Henry et al. 2015; Yang et al. 2017; Matthee et al. 2021; Naidu et al. 2022; Tang et al. 2024c). At low redshifts (z = 0.05 - 0.44), Ly α line widths and velocity offsets are seen to correlate with properties, e.g. M_{UV}, FWHM of rest-frame optical lines, which likely trace gas mass (Hayes et al. 2023). The sample of published spectra with sufficient S/N and resolution to measure Ly α lineshapes at $z \sim 5 - 6$ is small (e.g. MUSE resolution is $\sim 100 \, \mathrm{km \, s^{-1}}$). Our sample thus provides new insights into the lineprofiles at $z \sim 5-6$. We are particularly interested in understanding if Ly α lineshapes at $z \sim 5-6$ are significantly different from lower-z samples in ways which may impact the transmission of Ly α through the IGM.

As discussed in Section 3.1.3 we fit our 62 detected Ly α lines with an asymmetric Gaussian profile, enabling us to measure the broadness and skewness of the lines, which we discuss in more detail below. A strong asymmetric Ly α lineshape is expected at $z \gtrsim 5$ due to resonant scattering by residual neutral gas in the ionized IGM (e.g., Laursen et al. 2011; Mason & Gronke 2020). To explore the asymmetry in our data, we choose a binary classification of asymmetric and symmetric. We set profiles with $\alpha > 3$ in our asymmetric Gaussian model (see Equation 3) as asymmetric profiles and sample from our posteriors in our error analysis. We find that $69\pm7\%$ of our Ly α sample have asymmetric line profiles, suggesting the IGM is already playing a significant role in shaping Ly α emission at $z \sim 5-6$, consistent with previous analyses of stacked lines without and with systemic redshifts by Pentericci et al. (2018); Hayes et al. (2021) and Tang et al. (2024a) respectively.

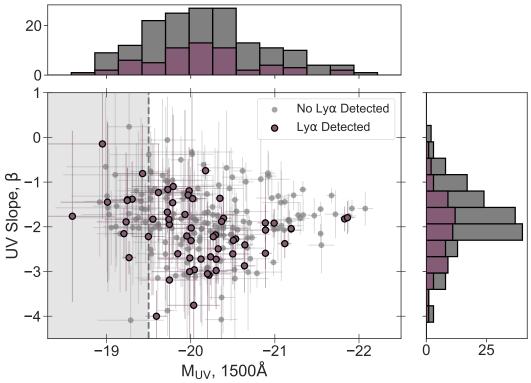


Fig. 4: M_{UV} and β slope. We show galaxies with Ly α detected (purple) and galaxies without detected Ly α (gray). These observations are the input into our two empirical models and other results. On top we show the M_{UV} distribution, there is no substantial difference between Ly α galaxies and Ly α non detections, with median M_{UV} of -20.0 and -20.1 respectively. On the right, the UV slope distribution, the two populations show a small difference, with median β of -2.0 and -1.8 respectively. The shaded area highlights the range where our sample becomes incomplete due to our target selection (see Section 6.1 for discussion). Extremely blue sources with β < 3 have large uncertainties due to being constrained by only three photometric data points.

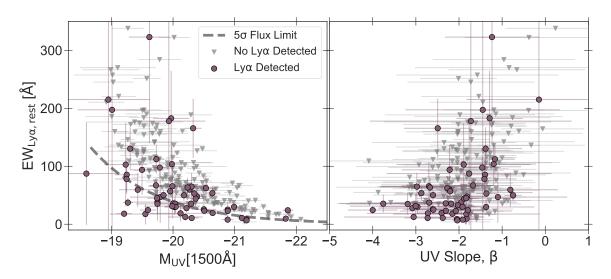


Fig. 5: Left: M_{UV} and rest-frame $EW_{\text{L}y\alpha}$. Right: β slope and rest-frame $EW_{\text{L}y\alpha}$. We show galaxies with $Ly\alpha$ detection (purple) and galaxies without $Ly\alpha$ detection (gray) as upper-limits. We find increased $EW_{\text{L}y\alpha}$ at fainter M_{UV} , similar to previous works (e.g. Shapley et al. 2003; Stark et al. 2011; Oyarzún et al. 2017; Tang et al. 2024a), and an unclear correlation with UV-slope. We show the median 5σ flux limit of the survey in dashed line.

The shape of the profile emerging from the ISM/CGM affects the fraction of Ly α transmitted through the IGM (e.g., Dijkstra et al. 2007, 2011; Mason et al. 2018a; Yuan et al. 2024), but can also reveal the impact of the IGM at $z \gtrsim 6$ as the IGM damping wing should produce smooth attenuation as a function of wavelength (Miralda-Escude 1998), making profiles more symmetric

at higher redshifts. To test this, we explore the redshift evolution of asymmetry splitting our samples in two bins. We find the asymmetric fraction is $72\pm6\%$ between z=5.2-5.8, and $60\pm9\%$ at z=5.8-6.4. With the large uncertainties, we cannot conclude if profiles become more symmetric at z>5.8, with a

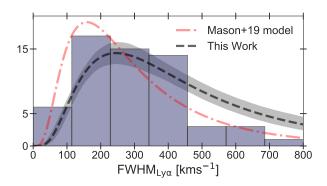


Fig. 6: FWHM_{Ly\alpha} distribution of our sample. We also plot our fitted log-normal distribution: median as a black dashed line, and the 16th and 84th percentile range in the shaded region. Our lognormal distribution is similar in shape but peaked at higher values than the prediction of Mason et al. (2019) (red dashed line) evaluated at the same median M_{UV} and redshift as our sample.

stronger impact of the IGM damping wing. Larger samples of Ly α line profiles would provide more insight into this effect.

In additional to the asymmetry, the distribution of FWHM_{Lyα} is a key uncertainty in reionization inferences (Mason et al. 2019), as it is an important factor in understanding non-detections: i.e. are Lyα detection rates lower at z>6 because Lyα is intrinsically broader and thus more difficult to detect? (Section 3.1.1). We can address this in our sample by inferring the distribution of FWHM_{Lyα} at $z\sim 5-6$. In Figure 6 we show the distribution of FWHM_{Lyα} our sample. As the observed distribution looks log-normal, we fit the FWHM_{Lyα} sample with a log-normal distribution. The resulting distribution is shown in Figure 6, and is described by:

$$P(\log_{10} \text{FWHM}_{\text{Ly}\alpha}) = \mathcal{N}(\mu = 2.39^{+0.04}_{-0.04}, \sigma = 0.30^{+0.04}_{-0.03})$$
 (5)

where μ is the mean and σ the standard deviation of the normal distribution of $\log_{10}(\text{FWHM}_{\text{Ly}\alpha})$. The mean of the log-normal distribution is $245\,\text{km}\,\text{s}^{-1}$ on a linear scale. In Figure 6 we show this distribution evaluated at the median M_{UV} and redshift of our $\text{Ly}\alpha$ -detected galaxies. For comparison, we plot the model FWHM distribution predicted by Mason et al. (2019), based on a model for $\text{Ly}\alpha$ velocity offsets as a function of M_{UV} , and an empirical relation between FWHM $_{\text{Ly}\alpha}$ and $\Delta \nu_{\text{Ly}\alpha}$ derived by Verhamme et al. (2018). Our inferred distribution peaks at higher FWHM $_{\text{Ly}\alpha}$ than the Mason et al. (2019) model, which peaks at 155 km s $^{-1}$. We discuss the implications of this in Section 6.

To better understand what drives $Ly\alpha$ line shapes in our sample we compare $FWHM_{Ly\alpha}$ with galaxy properties. In Figure 7, we plot $FWHM_{Ly\alpha}$ versus M_{uv} . Using the Pearson correlation coefficient we find a mild correlation (p-value ~ 0.1) of increasing $FWHM_{Ly\alpha}$ with M_{uv} , which we expect if M_{uv} traces galaxy mass and/or size (e.g., Shibuya et al. 2015; Roper et al. 2022; Allen et al. 2024; Morishita et al. 2024). Higher galaxy mass and size may indicate enhanced resonant scattering of $Ly\alpha$ due to higher N_{HI} .

4.4. Ly α velocity offsets

For the 14 galaxies in our sample with both Ly α and systemic redshifts we can measure the velocity offset of Ly α from systemic, $\Delta v_{\text{Ly}\alpha}$, which provides additional insight into the scattering of Ly α in the ISM and CGM (Neufeld 1991; Verhamme et al.

2008; Erb et al. 2014; Yang et al. 2017):

$$\Delta v_{\rm Ly\alpha} = c \left(\frac{z_{\rm Ly\alpha} - z_{\rm sys}}{1 + z_{\rm sys}} \right) \tag{6}$$

We measure z_{lya} from the peak of the Skewed Gaussian profile. To measure z_{sys} , we follow the procedure described in Section 4.5. The uncertainty in $\Delta \nu_{Ly\alpha}$ is mostly dominated by the uncertainties of the $Ly\alpha$ profile caused by the presence of strong sky-lines, though a few cases of weak $H\alpha$ also add to the $\Delta \nu_{Ly\alpha}$ uncertainty. For most measurements the error bars remain small, but nonetheless, we are limited by the spectral resolution of both Binospec and the NIRCam grism ($\sim 100 \, \text{km s}^{-1}$).

We show $\Delta v_{Ly\alpha}$ as a function of UV magnitude for our sample in Figure 7. We compare our results to those in the reionization era (i.e., Bunker et al. 2023; Tang et al. 2023, 2024b, z>7.5), mid redshift (i.e., Bradač et al. 2017; Stark et al. 2015; Willott et al. 2015; Inoue et al. 2016; Pentericci et al. 2016; Stark 2016; Mainali et al. 2018; Cassata et al. 2020; Endsley et al. 2022; Prieto-Lyon et al. 2023b, 4<z<7.5), and low redshift (i.e., Steidel et al. 2014; Erb et al. 2014, 2<z<3). We find a median and standard deviation (σ) $\Delta v_{\rm Ly\alpha}$ of 258±144 km s⁻¹. By excluding $\Delta\nu_{{\rm Ly}\alpha}$ <50 km s $^{-1}$ results, we find a median and $\sigma,~\Delta\nu_{{\rm Ly}\alpha}$ of $315\pm125~{\rm km\,s^{-1}}$. Both results are higher than the 205 ± 75 measured in $M_{uv} > -19$ galaxies in Prieto-Lyon et al. (2023b) at $z \sim 3-5$. Our results are consistent with similar measurements at $z \sim 5 - 6$ by Tang et al. (2024a), who find a median and $\sigma, \Delta v_{Lv\alpha}$ =250±156 km s⁻¹ for galaxies in the same M_{uv} range as our sample ($M_{UV} < -19.5$). Our new data spans a relatively narrow M_{UV} range, but combining with other samples at z > 3 referenced in the previous paragraph, we see a strong correlation of increasing $\Delta v_{Ly\alpha}$ for higher UV luminosities (p-value «0.01). We also compare with the semi-empirical model from Mason et al. (2018b). We see most of the z > 5 data lie above the model, which was derived from a sample of $M_{\rm uv} < -19$ and z~2 galaxies (Steidel et al. 2014; Erb et al. 2014).

Our sample includes two galaxies with extremely low $\Delta \nu_{\rm Ly\alpha}$ <50 km s⁻¹, z5-GND-17752 (z=5.77) and z5-GND-39445 (z=5.50), implying only minimal scattering in the ISM/CGM. We find that z5-GND-17752, our lowest $\Delta \nu_{\rm Ly\alpha}$ (20 km s⁻¹) measurement, also shows low FWHM_{H α} (120 km s⁻¹). Looking at their Ly α line profiles (Appendix B) we find significant emission bluewards of the Ly α peak, implying a highly ionized and/or low density sightline through the IGM enabling Ly α transmission around systemic velocity (Mason & Gronke 2020). However, such low $\Delta \nu_{\rm Ly\alpha}$ is at the limit of the spectral resolution, making them susceptible to systematic errors. We further discuss implications of our $\Delta \nu_{\rm Ly\alpha}$ results on Ly α transmission in Section 6.

4.5. Ly α escape fraction

Ly\$\alpha\$ and \$H\$\alpha\$ are both emitted predominantly in photoionized nebulae in a series of hydrogen line transitions, or recombination cascades, therefore linking the production of both types of photons. The ratio of products of the recombination cascade depends on the electron temperature (\$T_e\$) and electron number density (\$n_e\$), but most importantly the optical depth of the medium (\$Dijkstra 2014\$). In this work, we assume a Case-B recombination scenario, where the surrounding medium is optical thick to the Lyman series, as expected of HII regions, with \$n_e = 250 cm^{-3}\$, \$T_e = 10^4\$ (\$Dijkstra 2014\$). This yields an intrinsic Ly\$\alpha\$ flux of 8.7× the dust-corrected \$H\$\alpha\$ flux.

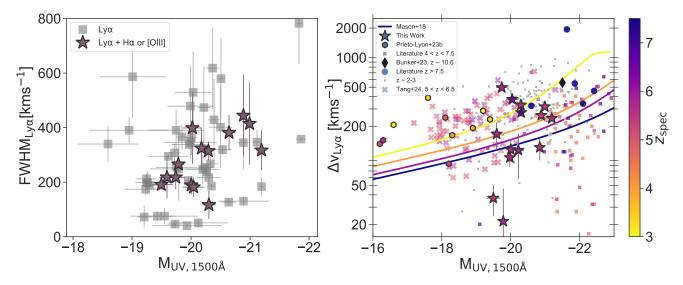


Fig. 7: Left: FWHM_{Lyα} and M_{uv}. We show galaxies with only a Lyα emission (gray) and with Lyα +Hα emission (purple). We find a mild (p-value ~ 0.1) correlation between FWHM_{Lyα} and M_{uv}. Right: Lyα velocity offset versus UV magnitude, with redshift as a color bar. We show the data from this work as stars, and the UV-faint (z~3-5) sample of Prieto-Lyon et al. (2023b) as circles. For reference we show a high redshift sample (i.e., Bunker et al. 2023; Tang et al. 2023, z>7.5), mid redshift (i.e. (i.e., Bradač et al. 2017; Stark et al. 2015; Willott et al. 2015; Inoue et al. 2016; Pentericci et al. 2016; Stark 2016; Mainali et al. 2018; Cassata et al. 2020; Endsley et al. 2022; Tang et al. 2024a, 4<z<7.5), and low redshift (i.e., Steidel et al. 2014; Erb et al. 2014, 2<z<3). We add the median semi-empirical model from Mason et al. (2018b). Combining our our data with the literature samples above z > 3 we find a strong correlation of $\Delta v_{\rm Ly\alpha}$ with increasing M_{uv} (p-value<<0.01).

The Ly α escape fraction is the ratio of the observed to intrinsic Ly α flux:

$$f_{\rm esc,ly\alpha} = \frac{F_{\rm Ly\alpha}}{8.7F_{\rm dust\ corrected,H\alpha}} \tag{7}$$

With both terms referring to the total flux of the emission lines (observed in the case of Ly α , dust-corrected in the case of H α), and the 8.7 is the factor derived from case-B recombination. Recovering $f_{\rm esc}^{\rm Ly}{}^{\alpha}$ < 1 implies not all Ly α escapes from the galaxy, due to dust absorption or scattering by high column densities of neutral hydrogen, effectively removing Ly α from the line-of-sight.

We correct all H α fluxes for dust attenuation following Lam et al. (2019). We assume an SMC (Small Magellanic Cloud) dust curve (Prevot et al. 1984), which is expected to be similar to that in $z \sim 4-6$ galaxies based on the infrared excess (IRx) - β slope relationship (Bouwens et al. 2016). The dust correction increases H α fluxes by $\sim 1-10\%$. We propagate the uncertainty of β into the dust correction, though we find the error bars of $f_{\rm esc}^{\rm Ly}$ are dominated by the uncertainty of the line fluxes in all cases.

In Figure 8 we show the resulting $Ly\alpha$ escape fractions as a function of M_{uv} and β . We have three classes of measurements, galaxies with $H\alpha$ emission only, galaxies with $Ly\alpha$ emission only, and galaxies with both $Ly\alpha$ and $H\alpha$. These classes produce upper limits, lower limits and measurements of $f_{\rm esc}^{\rm Ly\alpha}$, respectively. For galaxies with only $H\alpha$, we use the $Ly\alpha$ flux limit as presented in Section 3.1.3 to obtain an upper limit on $f_{\rm esc}^{\rm Ly\alpha}$. When only $Ly\alpha$ is present, we calculate $H\alpha$ upper limits as the 5σ flux limit at the expected location of $H\alpha$, assuming $z_{\rm Ly\alpha}$, integrated in a width randomly drawn from our FWHM $_{\rm H\alpha}$ distribution. We use the median flux limit of 10,000 realizations for this method. Uncertainties in $f_{\rm esc}^{\rm Ly\alpha}$ resulting from unknown $Ly\alpha$ velocity offsets are negligible. Our results in Figure 8 are dominated by upper-limits and lower-limits, to measure any

correlations, we use a Pearson correlation coefficient, and do random draws from a uniform distribution between [0,upper-limit] and [lower-limit,1]. With 1000 iterations we find a strong increasing correlation between $f_{\rm esc}^{{\rm L}y\alpha}$ and ${\rm M}_{\rm uv}$ (p-value<0.01), while $f_{\rm esc}^{{\rm L}y\alpha}$ and UV slope show no trend.

In Figure 9 we compare $f_{\rm esc}^{{\rm L}y\alpha}$ with ${\rm EW_{Ly\alpha}}$, and ${\rm FWHM_{Ly\alpha}}$, using the same classification explained in the previous paragraph. Applying the random sampling method outlined in the previous paragraph, we identify mild correlations and anticorrelations (p-value < 0.05): $f_{\rm esc}^{{\rm L}y\alpha}$ increases with higher ${\rm EW_{Ly\alpha}}$ but decreases with larger ${\rm FWHM_{Ly\alpha}}$. Beyond the positive correlation between $f_{\rm esc}^{{\rm L}y\alpha}$ and ${\rm EW_{Ly\alpha}}$, we find a small population of sources with weak ${\rm EW_{Ly\alpha}}$ <10Å and $f_{\rm esc}^{{\rm L}y\alpha}$ >0.1, indicating a possible ISM configuration that allows weak ${\rm Ly\alpha}$ to still escape from the galaxy. For ${\rm Ly\alpha}$ broadness, we find a mild correlation of high $f_{\rm esc}^{{\rm L}y\alpha}$ in sources with narrow ${\rm Ly\alpha}$ profiles, including two of the narrowest profiles in our sample, which exhibit $f_{\rm esc}^{{\rm L}y\alpha} \sim 1$. We further discuss the implications of these correlations and anti-correlations in Section 6.

One important caveat is that our dataset is inherently biased towards detecting sources with high Ly α luminosity. As a result, our sample exhibits a Malmquist bias towards the UV-faint regime, where detecting weaker Ly α becomes increasingly challenging (see Fig. 5 and Fig. 8). This bias may enhance the strength of the apparent correlations between EW_{Ly α} and $f_{\rm esc}^{\rm Ly}\alpha$ with M_{UV}, but we note that we account for upper limits in our empirical models in the following section.

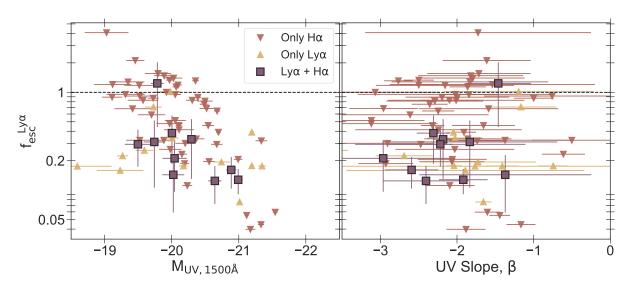


Fig. 8: Left: M_{uv} and $f_{esc}^{Ly\alpha}$. Right: β and $f_{esc}^{Ly\alpha}$. We separate the sample into three types of measurements; galaxies with $Ly\alpha$ and $H\alpha$ detections (purple), galaxies with only $H\alpha$ detection as upper-limits (red) and galaxies with only $Ly\alpha$ detection as lower-limits (yellow). We find a strong trend (p-value<0.01) where fainter galaxies have increased $f_{esc}^{Ly\alpha}$. For UV-slope, as in Figure 5, we do not find any clear trends.

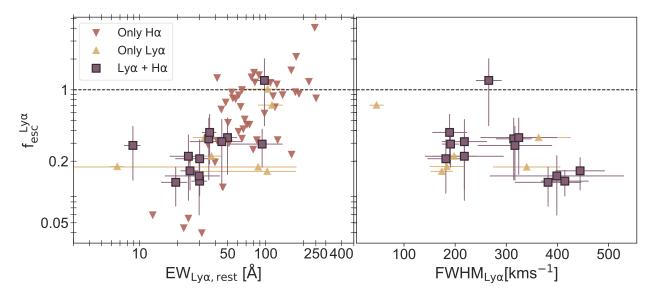


Fig. 9: Escape fraction of Ly α against other Ly α properties. Left: $f_{\rm esc}^{{\rm Ly}\alpha}$ and EW_{Ly α}. Right: $f_{\rm esc}^{{\rm Ly}\alpha}$ and FWHM_{Ly α}. The data shows an increase in EW_{Ly α} with higher $f_{\rm esc}^{{\rm Ly}\alpha}$, and an interesting population of weak Ly α with relatively high leakage ($f_{\rm esc}^{{\rm Ly}\alpha}$ >0.2). We find a mild correlation (p-value<0.05) of higher $f_{\rm esc}^{{\rm Ly}\alpha}$ with low FWHM_{Ly α}.

5. An empirical model for emergent Lylpha

We now seek to build an empirical model to predict emergent $\text{Ly}\alpha$ properties based on easily observable galaxy properties, M_{UV} and UV slope β . In the following section we describe the approach we used to produce model distributions for both $\text{Ly}\alpha$ EW and $f_{\text{esc}}^{\text{Ly}\alpha}$.

To estimate the Ly α EW and $f_{\rm esc}^{{\rm Ly}\alpha}$ distributions we use a Bayesian approach (e.g., Treu et al. 2012; Schenker et al. 2014; Oyarzún et al. 2017; Mason et al. 2018b; Tang et al. 2024a) which allows to obtain the posterior distribution of the parameters that describe the EW_{Ly α} and $f_{\rm esc}^{{\rm Ly}\alpha}$ distributions. We follow

Mason et al. (2018b) and assume the distributions for EW and $f_{\rm esc}^{{\rm Ly}\alpha}$ are described by an exponential with a potential peak at zero for non-emitters, with two parameters $\theta=(A,X_0)$. We note that previous work has shown the exact form of the distributions (exponential, log-normal etc.) does not significantly impact results (Schenker et al. 2014; Oyarzún et al. 2017; Tang et al. 2024a). In the following we use X to denote the 'data' EW or $f_{\rm esc}^{{\rm Ly}\alpha}$ as we model both distributions in the same way. The model distribution is:

$$p(X|\theta = \{A, X_0\}) = B\frac{A}{X_0}e^{-X/X_0}H(X) + (1 - A)\delta(X)$$
 (8)

The parameter X_0 represents the scale length of the exponential, and A is a normalization term which represents the fraction of Ly α emitters (i.e. the fraction of galaxies which have EW or $f_{\rm esc}^{\rm Ly}$ α). B is a normalization term for the distribution given the limits of EW and $f_{\rm esc}^{\rm Ly}$ α (B=1 for EW, and $B=1/(1-\exp(-1/X_0))$) for $f_{\rm esc}^{\rm Ly}$ α). H(X) is the Heaviside function and $\delta(X)$ is the Dirac-delta function.

We assume Gaussian measurement uncertainties with standard deviation σ on our data X_{obs} for each galaxy such that:

$$p(X_{\text{obs}}|X) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(X - X_{\text{obs}})^2}{2\sigma^2}\right]$$
(9)

Thus the likelihood for an individual observation is:

$$p(X_{\text{obs}}|\theta) = \int dX \, p(X_{\text{obs}}|X) p(X|\theta) \tag{10}$$

Where the limits of integration are $[0, \infty]$ for EW_{Ly\alpha} and [0,1] for $f_{\rm esc}^{\rm Ly\alpha}$.

Convolving our likelihood with the Gaussian distribution and solving the integral, we obtain the likelihood for our observa-

$$p(X_{\text{obs}}|\theta) = B \frac{A}{2X_0} e^{\frac{\sigma^2 - 2X_0 X_{\text{obs}}}{2X_0^2}} \left[\text{erf} \left(Z + \frac{X_{\text{max}}}{\sqrt{2}\sigma} \right) - \text{erf}(Z) \right]$$

$$+ \frac{(1 - A)}{\sqrt{2\pi}\sigma} e^{-\frac{X_{\text{obs}}^2}{2\sigma^2}}.$$
Where $Z = \frac{1}{\sqrt{2}\sigma} \left(\frac{\sigma^2}{X_0} - X_{\text{obs}} \right)$ (11)

Where *erf* is the error function, and $X_{\text{max}} = \infty$ for EW and 1 for $f_{\text{esc}}^{\text{Ly}\alpha}$.

For galaxies with Ly α non-detections (i.e. upper limits in EW_{Ly α} and $f_{\rm esc}^{{\rm Ly}\alpha}$) the likelihood is:

$$p(X_{i,\text{obs}}|\theta) = p(X_{i,\text{obs}} < X_{\text{uplim}}|\theta) + p(X_{i,\text{obs}} > X_{\text{uplim}}|\theta) \cdot (1 - C)$$
 (12)

Where the cumulative distribution function is obtained by integrating $p(X_{i,\text{obs}}|\theta)$ (Equation 10) from $x_{i,\text{obs}} = -\infty$ to x_{uplim} . We incorporate the second term using the methods from Schenker et al. (2014), accounting for the probability of an emission line above the 5σ flux limit being undetected due to a skyline. The completeness (C) is calculated for individual slits over $\Delta z_{\text{phot}} = 0.5$. If $H\alpha$ or [OIII] are available, we determine completeness for z_{sys} within a range of typical $\Delta v_{\text{Ly}\alpha}$ from 0 to 500 km s⁻¹.

For galaxies with no H α detection we have only a lower limit on $f_{\rm esc}^{{\rm Ly}\alpha}$, thus in those cases the likelihood is:

$$p(X_{i,\text{obs}}|\theta) = p(X_{i,\text{obs}} > X_{\text{lowlim}}|\theta)$$
 (13)

Where we integrate $p(X_{i,\text{obs}}|\theta)$ from X_{lowlim} to 1.

The final posterior is the product of all individual galaxy posteriors, $p(\theta|\{X_{\text{obs}}\} \propto p(\theta) \prod_i p(X_{i,\text{obs}}|\theta)$. We assume flat priors on A and X_0 : $A \in [0,1]$, $W_0 \in [0,500]$ and $f_{\text{esc},0}^{\text{Ly}\alpha} \in [0,1]$.

Article number, page 12 of 21

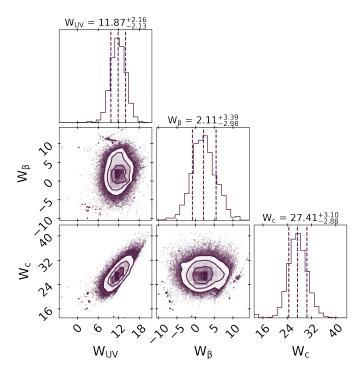


Fig. 10: Posteriors for the Ly α EW distribution model (Section 5.2). The posteriors show the coefficients for the linear model for EW scale length, W_0 as a function of $M_{\rm UV}$ ($W_{\rm UV}$), UV-slope (W_{β}), and a constant ($W_{\rm c}$). We see a strong correlation between $W_{\rm UV}$ and $W_{\rm c}$, and that there is no significant dependence on UV slope ($W_{\beta} \approx 0$).

5.1. Dependence of the distributions on galaxy properties

To better understand the dependence of Ly α emission on galaxy properties we parameterise our model as a function of M_{UV} and β . Previous work, and our results in Section 4, have demonstrated that Ly α EW and escape may be enhanced in UV-faint, bluer, galaxies (e.g., Tang et al. 2024a; Oyarzún et al. 2017).

We follow the approach done by Oyarzún et al. (2017) and parameterise our model parameters A and X_0 as a linear combination of M_{uv} and β .

$$A = A_{M_{UV}} M_{UV} + A_{\beta} \beta + A_{c}$$

$$X_{0} = X_{M_{UV}} M_{UV} + X_{\beta} \beta + X_{c}$$
(14)

To compute the posterior distributions, we follow the same method as before, assuming the same priors on A and X_0 .

We perform tests to mock datasets to assess whether our sample size is sufficient to robustly recover these parameters. We find that our Ly\$\alpha\$ EW sample (236 galaxies, including 62 Ly\$\alpha\$ detections) is large enough that we expect to recover the parameters in Equation 14 within 5-30% of the true parameters. We find our $f_{\rm esc}^{\rm Ly}$ sample (72 galaxies, including 10 Ly\$\alpha\$ +H\$\alpha\$ detections) is not large enough to fit the linear model robustly, so we assume constant A and X0, and find we can recover these robustly when splitting our sample into two bins. Both A and X0 are recovered with similar accuracy and precision in our tests.

5.2. Lyman Alpha Equivalent Width Model

As described above we will fit for the Ly α EW distribution assuming an exponential model with parameters described by a linear model in M_{uv} and β (Equation 14). We use the

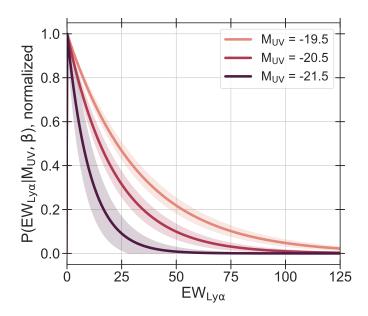


Fig. 11: Probability distribution of $EW_{Lv\alpha}$. We plot 3 distributions for different M_{UV} within our data range, showing UV faint galaxies are more likely to be stronger Ly α emitters.

measured EW, uncertainties and 5σ upper limits, as described in Section 4.2, in our likelihood (Equation 11), accounting for uncertainties in M_{uv} and β for each galaxy by sampling from their posteriors (Section 4.1). We perform the fitting using an MCMC with emcee (Foreman-Mackey et al. 2013) with 10,000 steps and 25 walkers, enough to reach convergence. We find that our model clearly returns $A \approx 1$, i.e. a purely exponential distribution with no peak at EW= 0 Å. Thus we fit only for the scale-length $X_0 \equiv EW_0$.

We show the resulting corner plots for the EW_0 parameters in Figure 10. We find that our model can be completely described by M_{uv} , with no significant dependence on the UV slope – as the linear coefficient, W_{β} , is consistent with zero within 1σ . This lack of correlation with the UV slope is likely due to our source selection, as discussed further in 6.1. Consequently, we opt to rerun the model excluding the UV slope. The inferred scale-length

$$EW_0 = 11.2^{+2.0}_{-2.0} \cdot (M_{\text{uv}} + 20) + 27.2^{+3.1}_{-2.8} \tag{15}$$

Where we give the median and 68% credibile intervals for each parameter given their posterior distributions.

In Figure 11 we visualize the normalized $EW_{Ly\alpha}$ distribution. We show the resulting distribution for 3 examples with $M_{uv} = -$ 19.5, -20.5, -21.5. We plot the median and 68% credible interval of the distributions obtained by sampling from the posterior predictive distributions for the EW_0 parameters. We see that faint galaxies have a broader EW distribution than bright galaxies, and therefore much higher likelihood of having a strong EW_{Lv α} as discussed in the next paragraph.

Using our EW distribution model we can also calculate the 'Ly α fraction': the fraction of Lyman-break galaxies with EW> 25 Å (e.g., Stark et al. 2010). We obtain this by integrating Equation 8 given our inferred parameters in Equation 15. For the three cases shown in Figure 11 the Ly α fraction ranges from 8 to 45% (\pm 5%) for $M_{\rm uv} = -21.5$ to $M_{\rm uv} = -19.5$. Assuming the most commonly used UV magnitude range in the literature $(-18.75 < M_{UV} < -20.25$, i.e. median $M_{UV} = -19.5$) we find

Sub-sample	$f_{ m esc~0}^{ m Ly}$	A
All	$0.27^{+0.16}_{-0.08}$	> 0.99*
$M_{uv} > -20.1$	$0.52^{+0.23}_{-0.22}$	> 0.98*
$M_{\rm uv} < -20.1$	$0.20^{+0.09}_{-0.05}$	> 0.98*
$\beta > -2.0$	$0.40^{+0.27}_{-0.17}$	> 0.97*
β < -2.0	$0.38^{+0.13}_{-0.08}$	> 0.99*

Table 2: Results for exponential scale length $(f_{\rm esc}^{\rm Ly\alpha})$ and delta (A) parameters for the four $f_{\rm esc}^{\rm Ly\alpha}$ bins. We state the median, 16th and 84th percentile values from the posterior for each parameter. *For lower limits we give the 68% limit. We note that due to the large number of upper limits, the recovered $f_{\rm esc}^{{\rm Ly}\alpha}$ should be viewed as upper limits.

a Ly α fraction of 45%±5%, consistent with recent findings by Tang et al. (2024a) of $35\pm7\%$. Based on our distributions, we expect 21±3% of M_{uv} = -19.5 galaxies to show $EW_{Ly\alpha} > 50\text{Å}$, fully consistent with Tang et al. (2024a) who found $22 \pm 3\%$, and 5±2% to show EW_{Ly α} >100Å, $\approx 2\sigma$ lower than Tang et al. (2024a). Our sample is $\sim 3\times$ smaller than that of Tang et al. (2024a), and less complete at $M_{\rm uv} \sim -19.5$, so it is likely that we are missing some of the high EW tail of the distribution, which are likely more common towards the fainter end of the M_{UV} range. Overall, we should expect a non-negligible fraction of very strong emitters as we move deep into the reionization era where $M_{\text{UV}} > -19.5$ galaxies become more common.

5.3. Lyman Alpha Escape Fraction Model

As our sample of sources with $f_{\rm esc}^{{\rm Ly}\alpha}$ measurements is not large enough to fit the linear model (Equation 14), we fit the $f_{\rm esc}^{{\rm Ly}\alpha}$ distribution assuming constant parameters A and X_0 in two bins. We create two UV magnitude bins: UV-faint ($M_{UV} > -20.1$), UVbright (M_{uv} <-20.1); and two UV slope bins: red (β >-2.0) and blue (β <2.0). These bins are defined so each contains the same number of sources.

We infer the parameters A and X_0 in our 4 sub-samples and give the resulting inferred parameters in Table 2. We show the inferred distributions in Figure 12. We recover median $f_{\rm esc}^{\rm Ly\alpha} =$

10.18 ± 0.06 from our distribution fit to the full sample. Similarly to our EW fits, we find a delta function at $f_{\rm esc}^{\rm Ly\alpha} = 0$ is disfavored. We find the probability of high $f_{\rm esc}^{\rm Ly\alpha}$ is much higher in our fainter $M_{\rm uv}$ bin. We find median $f_{\rm esc}^{\rm Ly\alpha} = 0.13 \pm 0.05$ in our UV-bright bin compared to 0.28 ± 0.08 in our UV-faint bin. From the first this case of the following that $f_{\rm esc}^{\rm Ly\alpha} = f_{\rm esc}^{\rm Ly\alpha}$ and $f_{\rm esc}^{\rm Ly\alpha} = f_{\rm esc}^{\rm Ly\alpha}$. our UV-faint bin we obtain that $62\pm8\%$ of $M_{uv} >-20.1$ galaxies should have $f_{\rm esc}^{\rm Ly\alpha} > 0.2$. This is considerably higher than the 30±6% found by Tang et al. (2024a) for $M_{\rm uv} \sim -19.5$ galaxies. We attribute this difference to our high upper limits on $f_{\rm esc}^{{
m Ly}lpha}$ which do not significantly constrain our models and the use of a uniform prior on $f_{\rm esc}^{{\rm Ly}\alpha}$. If we decrease our Ly α flux limits by a factor 2(4), we obtain 50±12% (35±10%) of $M_{\rm uv}$ >-20.1 galaxies have $f_{\rm esc}^{{\rm Ly}\alpha} > 0.2$, demonstrating the sensitivity of the inference to upper limits. Thus we consider our reported distribution an upper limit on the underlying distribution.

We do not find significant difference between the $f_{\rm esc}^{{\rm Ly}\alpha}$ distributions binned by UV slope. Previous works at z~5-6 (Tang et al. 2024a; Chen et al. 2024) have shown that $f_{\rm esc}^{{\rm Ly}\alpha}$ strongly cor-

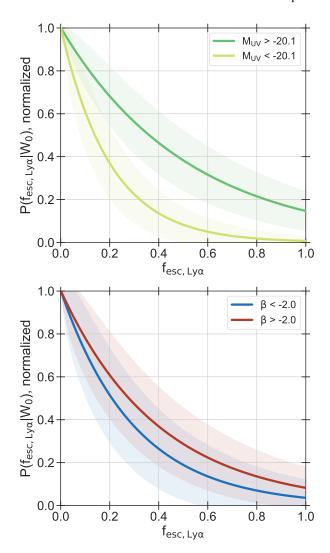


Fig. 12: Probability distributions of observed $f_{\rm esc}^{{\rm Ly}\alpha}$. Results are obtained by binning the dataset in $M_{\rm uv}$ and UV slope separately. As in our ${\rm EW}_{{\rm Ly}\alpha}$ results, we find that the model is dominated by $M_{\rm uv}$, while the two UV slope scenarios are not statistically different.

relates with UV slope. However, similar to $z\sim2$ Ly α surveys (i.e. Naidu et al. 2022), we do not see this correlation in our sample (see also Figure 8). The lack of trends with UV slope might also be driven by our target selection which is incomplete for blue $(\beta \lesssim -2)$ M_{UV} <-19.5 galaxies (see discussion in Section 6.1).

6. Discussion

In this paper we have presented new measurements and models of Ly α for Lyman-Break selected galaxies at $z \sim 5-6$. In Sections 4 and 5 we have connected the shape and strength of the Ly α line with other physical properties, such as M_{uv}. In the following discussion, we explore the implications of our findings for predicting Ly α emission emerging from the ISM/CGM (Section 6.1), as well as the impact of this transmission on its observability during the epoch of reionization (Section 6.2),.

6.1. Predicting Ly
$$\alpha$$
 transmission from the ISM/CGM at $z\sim 5-6$

The main goal of this paper was to build a predictive model for Ly α at the end of reionization, providing a basis to better interpret Ly α observations at z>6. These baseline observations are important due to the degeneracy between scattering of Ly α by neutral hydrogen in the IGM and ISM/CGM during the EoR. In particular, we focused on Ly α line shape and strength properties, including FWHM_{Ly α}, $\Delta \nu_{Ly\alpha}$, EW_{Ly α} and $f_{\rm esc}^{\rm Ly}$. In this section we discuss their trends with physical galaxy properties, M_{UV}, UV slope. Finally, we explore the possible physical drivers behind these trends and how they can be used to better predict Ly α at z>6.

In Section 4.3 we showed that Ly α lineshape properties: $\Delta v_{Ly\alpha}$ and FWHM_{Ly\alpha}, have tentative correlations with M_{uv}. In Figure 7, combining our $\Delta v_{Ly\alpha}$ results with other samples from the literature works at z > 3, we see a strong correlation of increasing $\Delta v_{Ly\alpha}$ with higher UV luminosity. We also found a mild correlation of increased FWHM $_{{\rm Ly}\alpha}$ in UV bright galaxies. It has been shown in theoretical works (e.g. Neufeld 1991; Verhamme et al. 2006) that broader Ly α profiles and higher $\Delta v_{\text{Ly}\alpha}$ are byproducts of scattering in high column density HI gas, increasing $Ly\alpha$ resonant scattering events. The trends we find of increasing $\Delta v_{Ly\alpha}$ and FWHM_{Ly\alpha} with UV luminosity are thus consistent with the physical picture that UV bright galaxies generally trace more massive and spatially extended galaxies, favoring higher HI column densities. This scenario is supported by observations showing a strong correlation between galaxy size and UV luminosity (e.g., Shibuya et al. 2015; Morishita et al. 2024). Furthermore, results from hydrodynamical simulations predict high HI column densities are more prevalent in UV bright galaxies, where supernova feedback is less efficient in disrupting the neutral ISM and opening channels of low HI column density (e.g., Rosdahl et al. 2022, Gelli et al. in prep, albeit with large sightline variance).

A key goal of this work was to better understand what shapes Ly α line profiles at $z \sim 5 - 6$, to help predict emergent Ly α during the EoR at z > 6, as pre-JWST reionization inferences have primarily been based on line profiles from $z \sim 2$ galaxies (e.g., Mason et al. 2018b). Recently, Tang et al. (2024a) demonstrated the median Ly α line profile at $z \sim 5 - 6$ is shifted to higher velocity offsets relative to a $z \sim 2$ comparison sample with similar Ly α EW (Tang et al. 2024c), and discussed this may reflect enhanced scattering in the ISM/CGM at these redshifts and/or resonant scattering by infalling residual neutral gas in the IGM with mean neutral fraction $\bar{x}_{\rm HI} \gtrsim 10^{-4}$ (Dijkstra et al. 2007; Laursen et al. 2011). Resonant scattering by infalling IGM gas should produce a sharp cut-off in the line profile, at the velocity of the infalling gas, resulting in observed Ly α lines which are narrower and with more redshifted peaks compared to lines transmitted through the highly ionized IGM (Santos 2004; Mason et al. 2018b; Park et al. 2021; Tang et al. 2024c). We find that $\approx 70\%$ of Ly α lines in our $z \sim 5-6$ sample are highly asymmetric (see also, Pentericci et al. 2018; Hayes et al. 2021), and our median FWHM_{Ly α} =245 km s⁻¹ is lower than that of the $z \sim 2 - 3$ sample presented by Tang et al. (2024c) (=290 km/s), which were selected as reionization-era analogs, with $[OIII] + H\beta$ EW comparable to $z \ge 6$ galaxies. These observations are qualitatively consistent with the predicted sharp flux cut-off due to resonant scattering by the IGM.

However, in Section 4.3 we also demonstrated our $z \sim 5-6$ FWHM_{Ly\alpha} distribution peaks at higher values than the model predictions by Mason et al. (2018b, 2019), which were based

on $z \sim 2$ Lyman-break galaxies (Erb et al. 2014; Steidel et al. 2014). We also found the majority of our velocity offset measurements exceed this model (also seen among other z > 5 sources in the literature). In the Mason et al. (2018b) model, Ly α velocity offsets and $FWHM_{Ly\alpha}$ are empirically estimated to scale with halo mass, $\sim M_h^{1/3}$, with redshift evolution incorporated by a shift in star formation to lower mass halos at fixed $M_{\mbox{\tiny UV}}$ at higher redshifts (due to more rapidly rising star formation histories, e.g., Behroozi et al. 2013; Mason et al. 2015). One possible factor missing from the Mason et al. (2018b) model is cosmological density evolution, which would increase the HI column density in the ISM/CGM at fixed halo mass with increasing redshift. This would be expected to produce a redshift scaling of both $\Delta v_{\rm Ly\alpha}$ and FWHM_{Ly\alpha} $\sim (1+z)^{2/3}$ at fixed halo mass, implying $\sim 1.6 \times$ higher $\Delta v_{\rm Ly\alpha}$ and FWHM_{Ly\alpha} at $z \sim 5$ compared to the original model. Such a scaling appears in good agreement with the ratio between our median observed line profile and the original model predictions: median observed $\Delta \nu_{Ly\alpha}$ =258(FWHM_{Ly\alpha} =245) km s⁻¹, compared to $\Delta \nu_{Ly\alpha}$ =150(FWHM_{Ly\alpha} =160) km s⁻¹ predicted by the Mason et al. (2018b) model. Current samples of sources with both Ly α velocity offsets and FWHM_{Ly α} measurements at $z \sim 5-6$ are still small, but larger samples will be important for better determining the impact of IGM infall relative to ISM/CGM density evolution on Ly α line profiles at these redshifts.

In Section 5.2, we showed that $EW_{Ly\alpha}$ and $f_{esc}^{Ly\alpha}$ generally increase with decreasing UV luminosity. Additionally, we identified a strong correlation between $f_{esc}^{Ly\alpha}$ and $EW_{Ly\alpha}$. We also showed find that $f_{\rm esc}^{{\rm Ly}\alpha}$ generally increases as ${\rm Ly}\alpha$ lines become narrower. These findings are consistent with previous studies, where $EW_{Ly\alpha}$ and $f_{esc}^{Ly\alpha}$ have an enhanced probability of being stronger in galaxies that are UV faint and bluer (Oyarzún et al. 2017; Tang et al. 2024a). However we find no strong trends with UV slope in our sample, as we discuss below. As discussed by Tang et al. (2024a), the strong correlation of $EW_{Ly\alpha}$ and $f_{esc}^{Ly\alpha}$ suggests that strong $EW_{Ly\alpha}$ requires an environment capable of high $f_{\rm esc}^{{\rm Ly}\alpha}$. This is consistent with the scenario discussed above, where reduced resonant scattering events are more common in UV faint galaxies, leading to lower probability of Ly α being destroyed by dust or scattered out the line of sight (Dijkstra 2014), and therefore increased $f_{\rm esc}^{{\rm Ly}\alpha}$. The link between ${\rm Ly}\alpha$ strength and the properties of the ISM/CGM suggest that EW_{Ly α} and $f_{\rm esc}^{\rm Ly}$ can be predicted statistically using non-Ly α observables. This is supported by previous work at low redshifts linking Ly α emission to rest-frame optical properties (e.g., Trainor et al. 2019; Hayes et al. 2023), and at z > 6 by our results, which, consistent with other recent studies using JWST observations of Ly α emitting galaxies at $z \sim 5-6$ (Tang et al. 2024a; Chen et al. 2024; Lin et al. 2024), show that Ly α strength can be predicted statistically from UV observables at these redshifts. Empirical baselines mapping from UV observables to Ly α are crucial for interpreting Ly α detections in z > 10 spectra, where we lose rest-frame optical wavelength coverage with NIRSpec.

Our measurements in Figures 5 and 8 show no clear trends of $EW_{Ly\alpha}$ and $f_{esc}^{Ly\alpha}$ with UV slope. The same is true of our $EW_{Ly\alpha}$ and $f_{esc}^{Ly\alpha}$ empirical models in Section 5.2 and 5.3: we find no significant dependencies of the two models on UV slope. Nonetheless, previous works have found that $Ly\alpha$ is strongly enhanced in galaxies with bluer UV continuum (Oyarzún et al. 2017; Tang et al. 2024a; Lin et al. 2024). We attribute this missing trend in our sample to our target selection, magF160W < 27.5 (see Sec-

tion 2.1) which results in a lack of $M_{uv} > -19.5$, blue ($\beta < -2$) galaxies. We tested this by comparing our sample in M_{uv} and UV slope against the JADES-DJA photometric catalogs. We find that imposing a magF160W < 27.5 cut for galaxies within $5 < z_{\text{phot}} < 6.5$ in the JADES-DJA catalog biases against blue ($\beta < -2$) UV slopes for $M_{uv} \ge -19.5$ galaxies. At $M_{uv} > -19.5$ we find median $\beta = -1.6$ in our data, while the full JADES-DJA sample finds $\beta = -2.2$. If we apply the magF160W < 27.5 cut to JADES-DJA, the median UV slope also becomes $\beta = -1.6$. As the faintest, bluest galaxies typically have strongest Ly α emission (Oyarzún et al. 2017; Tang et al. 2024a; Lin et al. 2024) we attribute our lack of trends of $EW_{Ly\alpha}$ and $f_{esc}^{Ly\alpha}$ with UV slope to our sample incompleteness. However, omitting sources with $M_{uv} > -19.5$ from our $EW_{Ly\alpha}$ and $f_{esc}^{Ly\alpha}$ models does not significantly impact our conclusions.

6.2. Implications for Ly α observability during reionization

Our primary goal was to better understand the typical emergent Ly α emission as it leaves the ISM/CGM in high-redshift galaxies. As the effects of HI in the IGM and ISM/CGM on Ly α are entangled during the EoR, we must separate the emergence of Ly α from galaxies and its absorption caused by the damping wing of the IGM (Miralda-Escude 1998; Dijkstra et al. 2007, 2011; Mason et al. 2018b). With this goal, we have produced baseline measurements with observations at $z \sim 5-6$. Our observations at the end of the EoR, allow us to isolate the effects of the ISM/CGM over Ly α , and to better infer the reionization timeline in the future. In particular, with $R \sim 4000$ resolution and observations of up to 15 hours, we have made the first measurement of the FWHM_{Ly α} distribution at $z \sim 5-6$, previously limited to bright sources and stacked profiles with lower resolution (e.g., Pentericci et al. 2018; Hayes et al. 2021).

The Ly α lineshape plays a key role in the Ly α detection rate once the IGM becomes increasingly neutral at $z \ge 6$, as the damping wing profile of the IGM is strongly wavelength dependent (Miralda-Escude 1998; Mesinger et al. 2015). Therefore, we must understand to what extent the Ly α transmission drop past $z \gtrsim 6$ is purely due to the IGM damping wing, or to the emergent line properties of Ly α . The FWHM_{Ly α} has two opposing effects on the observability of Ly α : First, a broader line profile will more easily transmit through the IGM as it increases the fraction of flux emitted at redder wavelengths, where it experiences lower damping wing optical depths (Prieto-Lyon et al. 2023b; Mukherjee et al. 2024; Yuan et al. 2024). However, a broader profile will lead to the same flux being spread over a wider wavelength range, hindering its detection in flux-limited observations. Increased $\Delta v_{L \nu \alpha}$ will always boost Ly α transmission through the IGM by shifting Ly α to redder wavelengths with lower damping wing optical depths.

To assess the impact of the lineshape on reionization analyses we compare our lineshape with the assumptions by Mason et al. (2018b, 2019) in their neutral fraction inference. We compare the IGM Ly α transmission for an $M_{\rm uv}=-20$ galaxy at z=9: comparing our median line profile and the model by Mason et al. (2018b). We assume a completely neutral IGM, in a region without ionized bubbles (using the damping wing approximation by Dijkstra 2014). The transmitted Ly α using our line shape median values (Gaussian, with FWHMLy α = 245 km s⁻¹, $\Delta v_{\rm Ly}\alpha$ = 258 km s⁻¹) is 9× higher than that predicted by Mason et al. (2019) (Gaussian FWHMLy α = 160 km s⁻¹ $\Delta v_{\rm Ly}\alpha$ = 150 km s⁻¹, and truncated at 80 km/s due to infalling IGM) – finding 0.9% vs 0.1% of Ly α flux is transmitted through the IGM

respectively. We note that for bubbles $R \gtrsim 1 \, \mathrm{pMpc}$ (roughly the expected mean bubble size when the IGM is $\sim 70\%$ neutral, Lu et al. 2023) the difference in transmission becomes negligible, implying the lineshape plays an important role in Ly α visibility in the earliest stages of reionization.

We also examine how the strongly asymmetric shape of Ly α should facilitate Ly α transmission in a mostly neutral IGM. For the 14 sources in our sample with measured velocity offsets we calculate the fraction of flux at $\Delta v_{\rm Ly}\alpha > 400\,{\rm km/s}$, which can be transmitted even in a neutral IGM at $z\sim 13$ (Dijkstra et al. 2011; Mason et al. 2018a; Yuan et al. 2024). We obtain a median and standard deviation of 22±18% from this sample, agreeing closely with recent results from hydrodynamical radiative transfer simulations by Yuan et al. (2024). This suggests the red tail of Ly α emission be important for allowing Ly α to be visible even at extremely high redshifts (e.g., Witstok et al. 2025a).

The higher transmission for Ly α which is both more offset from systemic and has extended red wings implies Ly α can be visible even in a very neutral IGM (Dijkstra et al. 2011; Witstok et al. 2025b). Our results suggest we may require a higher neutral fraction at $z \gtrsim 8$ to explain the Ly α decline compared to analyses which use $z \sim 2$ lineshapes as a baseline. This highlights the importance of including $z \sim 5-6$ samples, such as those used here, when making reionization inferences (see also, Yuan et al. 2024). For instance, Tang et al. (2024b) recently adopted the $z \sim 5-6$ Ly α line profile from Tang et al. (2024a) as a baseline when constraining the IGM neutral fraction at $z \sim 6-13$ from JWST observations.

7. Conclusions

A detailed understanding of Ly α as it emerges from galaxies in the first billion years is critical for interpreting Ly α observations during the reionization era. Our main goal was to produce empirical baseline models for Ly α to enable better constraints on the reionization history. We have presented an analysis of Lymanbreak galaxies at $z \sim 5-6$, when the IGM is expected to be mostly ionized. With our new ground-based Ly α spectra, supplemented by JWST imaging and spectra, we have studied the correlations between Ly α shape and strength with UV luminosity. We conclude the following:

- 1. We present high-resolution MMT/Binospec restframe-UV spectroscopy for 236 Lyman-break galaxies. We detect Ly α at S/N> 5 for 62 of these, with a median 5σ flux limit of $1.34\times10^{-17} {\rm erg~s^{-1}~cm^{-2}}$. Overlapping ancilliary JWST/NIRCam spectra from FRESCO allows us to measure $z_{\rm sys}$ for 56 of our targets, including 14 galaxies detected with Ly α .
- 2. With $R \sim 4600$ spectroscopy we measure $\text{Ly}\alpha$ line profiles with a high-resolution of $\sim 70\,\text{km s}^{-1}$. We measured for the first time the FWHM_{Ly\alpha} distribution at $z\sim 5$ 6, a key unknown quantity in interpreting the decline of $\text{Ly}\alpha$ at z>6. We find mean FWHM_{Ly\alpha} =245 km s⁻¹, higher than models based on $z\sim 2$ Ly\alpha observations (Mason et al. 2018b, 2019), but lower than $z\sim 2$ 3 observations of reionization-era analogs by Tang et al. (2024c). We discussed that this may be due to resonant scattering by infalling IGM with $\overline{x}_{\text{HI}}\gtrsim 10^{-4}$ at $z\sim 5$ 6, as well as increasing ISM/CGM densities at higher redshifts at fixed halo mass the latter an effect which was not included in the Mason et al. (2018b) model.

- 3. We obtain 14 new measurements of Ly α velocity offsets at z>5 through systemic redshifts from H α or [OIII] from FRESCO NIRCam slitless spectra. We reported correlations between physical properties and Ly α lineshape: $\Delta \nu_{\rm Ly}\alpha$ and FWHM_{Ly}\alpha}. We find that the majority of galaxies show $\Delta \nu_{\rm Ly}\alpha \gtrsim 100 \, {\rm km \, s^{-1}}$, with a median 258 km s⁻¹. Combined with z>3 data from the literature, we find a strong correlation for increased $\Delta \nu_{\rm Ly}\alpha$ in UV bright galaxies. We found mild correlations of higher FWHM_{Ly}\alpha} in galaxies that are UV-bright. Our lineshape trends are consistent with a scenario of higher Ly α optical depth, with numerous scattering events in the ISM/CGM, for increasingly UV bright galaxies.
- 4. We measured the line strength properties $EW_{Ly\alpha}$ and $f_{\rm esc}^{Ly\alpha}$. We found strong correlations of $EW_{Ly\alpha}$ and $f_{\rm esc}^{Ly\alpha}$ with increasing UV luminosity, and mild correlations of increasing $f_{\rm esc}^{Ly\alpha}$ with lower FWHM_{Ly\alpha}. The decline of $f_{\rm esc}^{Ly\alpha}$ with line broadness and UV luminosity is consistent with a scenario of higher Ly\alpha optical depth in UV bright galaxies.
- 5. We created an empirical model for the probability distribution of EW_{Ly\alpha} and $f_{\rm esc}^{{\rm Ly}\alpha}$ at the end of the EoR, as a function of M_{uv} and UV slope. Our EW_{Ly\alpha} and $f_{\rm esc}^{{\rm Ly}\alpha}$ models strongly depend on M_{uv}. We find strong EW_{Ly\alpha} (> 25 Å) and $f_{\rm esc}^{{\rm Ly}\alpha}$ (>0.2) are common at $z \sim 5$ 6: Our models predict that 45±5% and < 62 ± 8% of M_{uv} =-19.5 galaxies have EW_{Ly\alpha} > 25 Å and $f_{\rm esc}^{{\rm Ly}\alpha}$ > 0.2.

The NIR capabilities of JWST are now allowing us to unlock new information about emergent Ly α line profiles and escape fractions through rest-frame optical emission lines. As $R \gtrsim 1000$ optical line samples at $z \sim 5-9$ grow steadily by the thousands, in large part thanks to NIRCam grism surveys (e.g. Kashino et al. 2023; Naidu et al. 2024; Covelo-Paz et al. 2025; Meyer et al. 2024), we can increasingly refine our understanding of emergent Ly α at $z \sim 5-6$. The unprecedented detections of Ly α at $z \sim 10-13$ (Bunker et al. 2023; Witstok et al. 2025b) have opened a new window on studying the earliest stages of reionization. Baseline Ly α empirical correlations and models, such as those we have presented (see also Chen et al. 2024; Tang et al. 2024a), are crucial for interpreting the growing JWST observations of Ly α at $z \sim 6-13$ to constrain the reionization process (Tang et al. 2024b; Mason et al. 2025).

Data access: All MMT/Binospec spectra and catalogs used in this work are publicly available in the Electronic Research Data Archive at University of Copenhagen at the link: https://sid.erda.dk/sharelink/FCZBDRW602.

Acknowledgements. We thank Hakim Atek, Anne Verhamme and Johan Fynbo for useful discussions. We thank Igor Chilingarian, Sean Moran for their initial reductions and advice regarding the Binospec pipeline, and Dan Fabricant, Ben Weiner and Ben Johnson for their advice in planning and reducing Binospec observations. We thank Daniel Eisenstein for helpful discussions on the Binospec survey design and comments on a draft of this manuscript. GPL and CAM acknowledge support by the VILLUM FONDEN under grant 37459. CAM acknowledges support from the Carlsberg Foundation under grant CF22-1322. The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant DNRF140.

References

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35

```
Adams, N. J., Conselice, C. J., Ferreira, L., et al. 2022, arXiv e-prints,
   arXiv:2207.11217
Allen, N., Oesch, P. A., Toft, S., et al. 2024, arXiv e-prints, arXiv:2410.16354
Barro, G., Pérez-González, P. G., Cava, A., et al. 2019, ApJS, 243, 22
Becker, G. D., Bolton, J. S., Madau, P., et al. 2015, MNRAS, 447, 3402
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Bolan, P., Lemaux, B. C., Mason, C., et al. 2022, MNRAS, 517, 3263
Bosman, S. E. I., Davies, F. B., Becker, G. D., et al. 2022, MNRAS, 514, 55
Bouwens, R. J., Illingworth, G. D., a. Oesch, P., et al. 2015, ApJ, 811, 140
Bouwens, R. J., Illingworth, G. D., Oesch, P., et al. 2012, ApJ, 754, 83
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, ApJ, 793, 115
Bouwens, R. J., Smit, R., Labbé, I., et al. 2016, ApJ, 831, 176
Bradač, M., Garcia-Appadoo, D., Huang, K.-H., et al. 2017, ApJ, 836, L2
Brammer, G. 2023, grizli
Brammer, G. & Matharu, J. 2021, gbrammer/grizli: Release 2021
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Bunker, A. J., Saxena, A., Cameron, A. J., et al. 2023, A&A, 677, A88
Cassata, P., Morselli, L., Faisst, A., et al. 2020, A&A, 643, A6
Chen, Z., Stark, D. P., Mason, C., et al. 2024, MNRAS, 528, 7052
Chen, Z., Stark, D. P., Mason, C. A., et al. 2025, arXiv e-prints,
   arXiv:2505.24080
Covelo-Paz, A., Giovinazzo, E., Oesch, P. A., et al. 2025, A&A, 694, A178
Davies, F. B., Hennawi, J. F., Bañados, E., et al. 2018, ApJ, 864, 142
De Barros, S., Pentericci, L., Vanzella, E., et al. 2017, A&A, 608, A123 Dijkstra, M. 2014, PASA, 31, e040
Dijkstra, M., Lidz, A., & Wyithe, J. S. B. 2007, MNRAS, 377, 1175
Dijkstra, M., Mesinger, A., & Wyithe, J. S. B. 2011, MNRAS, 414, 2139
Dijkstra, M. & Wyithe, S. 2010, MNRAS, 408, 352
Donnan, C. T., McLure, R. J., Dunlop, J. S., et al. 2024, MNRAS, 533, 3222
Egami, E., Sun, F., Alberts, S., et al. 2023, Complete NIRCam Grism Redshift
   Survey (CONGRESS), JWST Proposal. Cycle 2, ID. #3577
Eilers, A.-C., Davies, F. B., & Hennawi, J. F. 2018, ApJ, 864, 53
Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023, arXiv e-prints,
    arXiv:2306.02465
Endsley, R., Stark, D. P., Bouwens, R. J., et al. 2022, MNRAS, 517, 5642
Erb, D. K., Steidel, C. C., Trainor, R. F., et al. 2014, ApJ, 795, 33
Euclid Collaboration, Barnett, R., Warren, S. J., et al. 2019, A&A, 631, A85
Fabricant, D., Fata, R., Epps, H., et al. 2019, PASP, 131, 075004
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, Astron. J., 132, 117
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publ. As-
   tron. Soc. Pacific, 125, 306
Fuller, S., Lemaux, B. C., Bradač, M., et al. 2020, ApJ, 896, 156
Gaikwad, P., Haehnelt, M. G., Davies, F. B., et al. 2023, MNRAS, 525, 4093
Gelli, V., Mason, C., & Hayward, C. C. 2024, ApJ, 975, 192
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93 Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Gronke, M., Bull, P., & Dijkstra, M. 2015, ApJ, 812, 123
Hashimoto, T., Verhamme, A., Ouchi, M., et al. 2015, ApJ, 812, 157
Hayes, M. J., Runnholm, A., Gronke, M., & Scarlata, C. 2021, ApJ, 908, 36
Hayes, M. J., Runnholm, A., Scarlata, C., Gronke, M., & Rivera-Thorsen, T. E.
   2023, MNRAS, 520, 5903
Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, ApJ, 809, 19
Hoag, A., Treu, T., Pentericci, L., et al. 2019, MNRAS, 488, 706
Hu, E. M., Cowie, L. L., Barger, A. J., et al. 2010, ApJ, 725, 394
Inoue, A. K., Tamura, Y., Matsuo, H., et al. 2016, Science, 352, 1559
Jiang, L., McGreer, I. D., Fan, X., et al. 2016, ApJ, 833, 222
Jones, G. C., Bunker, A. J., Saxena, A., et al. 2024, A&A, 683, A238
Jung, I., Finkelstein, S. L., Dickinson, M., et al. 2019, ApJ, 877, 146
Jung, I., Finkelstein, S. L., Dickinson, M., et al. 2020, ApJ, 904, 144
Jung, I., Finkelstein, S. L., Livermore, R. C., et al. 2018, ApJ, 864, 103
Kageura, Y., Ouchi, M., Nakane, M., et al. 2025, arXiv e-prints,
   arXiv:2501.05834
Kashino, D., Lilly, S. J., Matthee, J., et al. 2023, ApJ, 950, 66
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Lam, D., Bouwens, R. J., Coe, D., et al. 2019, arXiv e-prints, arXiv:1903.08177
Laporte, N., Meyer, R. A., Ellis, R. S., et al. 2021, MNRAS, 505, 3336
Laursen, P., Sommer-Larsen, J., & Razoumov, A. O. 2011, ApJ, 728, 52
Leclercq, F., Bacon, R., Wisotzki, L., et al. 2017, A&A, 608, A8
Lemaux, B. C., Fuller, S., Bradač, M., et al. 2021, MNRAS, 504, 3662
Lin, X., Cai, Z., Wu, Y., et al. 2024, ApJS, 272, 33
Lu, T.-Y., Mason, C., Hutter, A., et al. 2023, arXiv e-prints, arXiv:2304.11192
Mainali, R., Zitrin, A., Stark, D. P., et al. 2018, MNRAS, 479, 1180
Mason, C. A., Chen, Z., Stark, D. P., et al. 2025, arXiv e-prints,
   arXiv:2501.11702
Mason, C. A., Fontana, A., Treu, T., et al. 2019, MNRAS, 485, 3947
Mason, C. A. & Gronke, M. 2020, MNRAS, 499, 1395
Mason, C. A., Trenti, M., & Treu, T. 2015, ApJ, 813, 21
Mason, C. A., Treu, T., de Barros, S., et al. 2018a, ApJ, 857, L11
```

Mason, C. A., Treu, T., Dijkstra, M., et al. 2018b, ApJ, 856, 2 Matsuoka, Y., Strauss, M. A., Kashikawa, N., et al. 2018, ApJ, 869, 150

```
Matthee, J., Sobral, D., Hayes, M., et al. 2021, MNRAS, 505, 1382
Mesinger, A., Aykutalp, A., Vanzella, E., et al. 2015, MNRAS, 446, 566
Mesinger, A. & Furlanetto, S. R. 2008, MNRAS, 385, 1348
Meyer, R. A., Oesch, P. A., Giovinazzo, E., et al. 2024, MNRAS, 535, 1067
Miralda-Escude, J. 1998, ApJ, 501, 15
Morishita, T., Stiavelli, M., Chary, R.-R., et al. 2024, ApJ, 963, 9
Muñoz, J. B., Mirocha, J., Chisholm, J., Furlanetto, S. R., & Mason, C. 2024,
   MNRAS, 535, L37
Mukherjee, T., Zafar, T., Nanayakkara, T., et al. 2024, PASA, 41, e105
Naidu, R. P., Matthee, J., Kramarenko, I., et al. 2024, arXiv e-prints,
   arXiv:2410.01874
Naidu, R. P., Matthee, J., Oesch, P. A., et al. 2022, MNRAS, 510, 4582
Nakane, M., Ouchi, M., Nakajima, K., et al. 2024, ApJ, 967, 28
Neufeld, D. A. 1991, ApJ, 370, L85
Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023, MNRAS, 525, 2864
Oyarzún, G. A., Blanc, G. A., González, V., Mateo, M., & Bailey, J. I. 2017, ApJ,
   843, 133
Park, H., Jung, I., Song, H., et al. 2021, ApJ, 922, 263
Pentericci, L., Carniani, S., Castellano, M., et al. 2016, ApJ, 829, L11
Pentericci, L., Vanzella, E., Castellano, M., et al. 2018, Astronomy & Astro-
   physics, 619, A147
Pentericci, L., Vanzella, E., Castellano, M., et al. 2018, A&A, 619, A147
Pentericci, L., Vanzella, E., Fontana, A., et al. 2014, ApJ, 793, 113
Prevot, M. L., Lequeux, J., Maurice, E., Prevot, L., & Rocca-Volmerange, B.
   1984, A&A, 132, 389
Prieto-Lyon, G., Mason, C., Mascia, S., et al. 2023b, ApJ, 956, 136
Qin, Y., Mesinger, A., Bosman, S. E. I., & Viel, M. 2021, MNRAS, 506, 2390
Rodrigo, C. & Solano, E. 2020, in XIV.0 Scientific Meeting (virtual) of the Span-
   ish Astronomical Society, 182
Roper, W. J., Lovell, C. C., Vijayan, A. P., et al. 2022, MNRAS, 514, 1921
Rosdahl, J., Blaizot, J., Katz, H., et al. 2022, MNRAS, 515, 2386
Santos, M. R. 2004, MNRAS, 349, 1137
Saxena, A., Bunker, A. J., Jones, G. C., et al. 2024, A&A, 684, A84
Schenker, M. A., Ellis, R. S., Konidaris, N. P., & Stark, D. P. 2014, ApJ, 795, 20
Schindler, J.-T., Bañados, E., Connor, T., et al. 2023, ApJ, 943, 67
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588,
Shibuya, T., Ouchi, M., & Harikane, Y. 2015, ApJS, 219, 15
Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24
Stark, D. P. 2016, ARA&A, 54, 761
Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, MNRAS, 464, 469
Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, MNRAS, 408,
   1628
Stark, D. P., Ellis, R. S., & Ouchi, M. 2011, ApJ, 728, L2
Stark, D. P., Walth, G., Charlot, S., et al. 2015, MNRAS, 454, 1393
Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165
Tang, M., Stark, D. P., Chen, Z., et al. 2023, MNRAS, 526, 1657
Tang, M., Stark, D. P., Ellis, R. S., et al. 2024a, MNRAS, 531, 2701
Tang, M., Stark, D. P., Ellis, R. S., et al. 2024c, ApJ, 972, 56
Tang, M., Stark, D. P., Topping, M. W., Mason, C., & Ellis, R. S. 2024b, ApJ,
   975, 208
Trainor, R. F., Strom, A. L., Steidel, C. C., et al. 2019, arXiv:1908.04794 [astro-
   ph] [arXiv:1908.04794]
Treu, T., Trenti, M., Stiavelli, M., Auger, M. W., & Bradley, L. D. 2012, ApJ,
   747, 27
Umeda, H., Ouchi, M., Nakajima, K., et al. 2024, ApJ, 971, 124
Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389
Valentino, F., Brammer, G., Gould, K. M. L., et al. 2023, ApJ, 947, 20
Verhamme, A., Garel, T., Ventou, E., et al. 2018, MNRAS, 478, L60
Verhamme, A., Schaerer, D., Atek, H., & Tapken, C. 2008, Astronomy and As-
   trophysics, 491, 89
Verhamme, A., Schaerer, D., & Maselli, A. 2006, Astronomy & Astrophysics,
   460, 397
Wang, F., Davies, F. B., Yang, J., et al. 2020, ApJ, 896, 23
Whitler, L., Stark, D. P., Topping, M. W., et al. 2025, arXiv e-prints,
   arXiv:2501.00984
Whitler, L. R., Mason, C. A., Ren, K., et al. 2020, MNRAS, 495, 3602
Willott, C. J., Carilli, C. L., Wagg, J., & Wang, R. 2015, ApJ, 807, 180
Witstok, J., Jakobsen, P., Maiolino, R., et al. 2025a, Nature, 639, 897
Witstok, J., Maiolino, R., Smit, R., et al. 2025b, MNRAS, 536, 27
Yang, H., Malhotra, S., Gronke, M., et al. 2017, ApJ, 844, 171
Yang, J., Wang, F., Fan, X., et al. 2020, ApJ, 897, L14
Yuan, Y., Martin-Alvarez, S., Haehnelt, M. G., et al. 2024, arXiv e-prints,
   arXiv:2412.07970
```

Appendix A: Lylpha Detections

ID	RA	DEC	z _{Lyα}	Z _{SYS}	M_{uv}	β	$EW_{Ly\alpha}$	FWHM _{Lyα}	$f_{\rm esc}^{{ m Ly}lpha}$	$\Delta v_{\mathrm{Ly}\alpha}$
			-Lyu	-sys		7	[Å]	[km s ⁻¹]	[km s ⁻¹]	[km s ⁻¹]
z5-GNW-1503	189.139303	62.111230	5.0508	_	-20.9 ± 0.1	-1.9 ± 0.4	8 ± 3	345 ± 97	_	_
z5-GNW-12024	189.142960	62.179790	5.1661	_	-18.6 ± 0.5	-1.8 ± 1.9	87 ± 89	340 ± 66	> 0.18	_
z5-GND-3052	189.288742	62.173836	5.1796	_	-20.3 ± 0.1	-1.4 ± 0.2	53 ± 6	185 ± 10	_	_
z5-GNW-11071	188.972424	62.173125	5.1848	_	-19.3 ± 0.1	-1.4 ± 0.8	130 ± 37	189 ± 8	_	_
z5-GND-32413	189.357854	62.211445	5.2326	_	-19.2 ± 0.1	-2.2 ± 0.6	18 ± 5	72 ± 41	_	_
z5-GNW-20906	189.437681	62.318150	5.2911	_	-21.8 ± 0.0	-1.8 ± 0.3	10 ± 1	782 ± 253	_	_
z5-GNW-11014	189.073073	62.172883	5.2946	_	-20.4 ± 0.1	-1.9 ± 0.3	44 ± 7	618 ± 96	_	_
z5-GNW-21219	189.310135	62.330258	5.2946	_	-20.3 ± 0.1	-2.2 ± 0.3	60 ± 7	237 ± 11	_	_
z5-GNW-22490	189.394847	62.308059	5.3211	_	-20.2 ± 0.1	-3.1 ± 0.8	28 ± 7	474 ± 128	_	_
z6-GND-35647	189.329865	62.200844	5.3496	_	-19.4 ± 0.3	-0.8 ± 1.1	59 ± 30	75 ± 18	_	_
z5-GND-27819	189.110337	62.225441	5.3510	5.3465	-20.9 ± 0.1	-2.6 ± 0.3	25 ± 4	444 ± 48	0.16 ± 0.06	258 ± 32
z5-GNW-12482	189.079256	62.182873	5.3543	_	-20.5 ± 0.1	-2.3 ± 0.2	23 ± 3	580 ± 113	_	_
z5-GNW-2438	189.106533	62.118483	5.3569	_	-19.8 ± 0.1	-2.6 ± 0.4	50 ± 9	366 ± 46	_	_
z5-GNW-1663	189.147548	62.112511	5.3697	_	-19.7 ± 0.1	-1.7 ± 0.5	67 ± 13	308 ± 34	_	_
z5-GNW-32760	189.451265	62.241037	5.4033	_	-19.7 ± 0.3	-3.2 ± 1.4	35 ± 19	254 ± 27	_	_
z5-GNW-21283	189.403559	62.315688	5.4046	_	-20.0 ± 0.3	-3.8 ± 1.0	35 ± 15	529 ± 58	_	_
z5-GNW-7098	189.214160	62.149048	5.4338	_	-20.5 ± 0.2	-2.3 ± 0.8	62 ± 21	319 ± 17	_	_
z5-GNW-29609	189.436412	62.265038	5.4455	_	-20.4 ± 0.1	-1.8 ± 0.5	48 ± 8	429 ± 30	_	_
z5-GND-21153	189.049628	62.244033	5.4483	5.4430	-21.0 ± 0.1	-1.9 ± 0.3	30 ± 4	414 ± 47	0.13 ± 0.04	316 ± 20
z5-GNW-13514	188.967311	62.189120	5.4623	_	-20.3 ± 0.2	-3.0 ± 0.9	65 ± 24	217 ± 30	_	_
z6-GNW-23350	189.414474	62.333523	5.4633	_	_	_	_	325 ± 37	_	_
z5-GNW-23042	189.415009	62.333538	5.4635	_	-20.0 ± 0.3	-1.3 ± 0.8	183 ± 82	338 ± 18	_	_
z5-GND-39445	189.178643	62.187234	5.5048	5.5043	-19.5 ± 0.2	-2.2 ± 1.1	94 ± 43	191 ± 16	0.29 ± 0.12	37 ± 13
z6-GNW-14478	189.001450	62.194984	5.5546	_	-19.0 ± 0.7	-0.1 ± 1.7	215 ± 447	390 ± 48	_	_
z5-GND-7766	189.139984	62.291809	5.5913	5.5888	-20.2 ± 0.1	-0.7 ± 0.2	50 ± 10	324 ± 75	0.34 ± 0.19	96 ± 19
z6-GND-36100	189.191274	62.199519	5.6029	5.5975	-20.3 ± 0.0	-2.2 ± 0.1	35 ± 3	315 ± 35	0.33 ± 0.20	329 ± 37
z6-GND-36553	189.156387	62.197773	5.6109	5.6052	-21.2 ± 0.0	-2.0 ± 0.1	9 ± 1	317 ± 73	0.29 ± 0.16	241 ± 18
Stark11-13066	189.156399	62.197716	5.6115	_	-21.2 ± 0.0	-2.0 ± 0.1	7 ± 1	184 ± 34	> 0.18	_
z6-GNW-14511	189.100538	62.195344	5.6155	5.6117	-20.6 ± 0.1	-2.4 ± 0.6	19 ± 5	382 ± 65	0.12 ± 0.05	122 ± 32
Hu10-3	189.055939	62.129990	5.6328	_	_	_	_	307 ± 14	_	_
Hu10-7	189.032780	62.143962	5.6389	_	-19.6 ± 0.3	-1.2 ± 1.1	323 ± 183	295 ± 8	_	_
z6-GND-43125	189.189137	62.300659	5.6585	5.6569	-20.0 ± 0.1	-2.3 ± 0.2	36 ± 5	189 ± 34	0.38 ± 0.20	116 ± 28
Hu10-6	189.324563	62.299734	5.6611	_	-20.2 ± 0.1	-2.7 ± 0.6	63 ± 17	253 ± 16	_	_
z6-GNW-10822	188.995306	62.171460	5.6657	_	-20.5 ± 0.2	-2.6 ± 0.7	26 ± 9	402 ± 47	_	_
Hu10-5	189.399713	62.239490	5.6713	_	-21.1 ± 0.1	-2.4 ± 0.3	11 ± 2	347 ± 46	_	-
Hu10-11	189.366007	62.196189	5.6719	-	-19.2 ± 0.2	-1.4 ± 0.7	78 ± 26	212 ± 20	_	-
z5-GND-37006	189.366007	62.196189	5.6720	-	-19.2 ± 0.2	-1.4 ± 0.7	85 ± 30	199 ± 20	_	_
Stark11-22381	189.255389	62.357754	5.6920	_	-20.2 ± 0.3	-3.1 ± 1.6	12 ± 6	247 ± 54	_	_
Stark11-3982	189.039261	62.247654	5.7080	5.7030	-19.6 ± 0.1	-4.0 ± 0.6	24 ± 7	218 ± 77	0.22 ± 0.14	167 ± 52
z7-GNW-22375	189.342596	62.308518	5.7718	_	-20.1 ± 0.5	-2.7 ± 1.4	19 ± 14	51 ± 16	-	_
z5-GND-17752	189.091798	62.253735	5.7744	5.7750	-19.8 ± 0.1	-1.5 ± 0.4	98 ± 19	266 ± 25	1.23 ± 0.79	22 ± 25
Stark11-17705	189.208207	62.232128	5.8027	5.7925	-19.7 ± 0.2	-1.8 ± 1.1	45 ± 20	218 ± 44	0.31 ± 0.21	496 ± 31
z6-GND-30340	189.388906	62.217840	5.8079	_	-20.0 ± 0.4	-2.7 ± 1.1	66 ± 42	479 ± 53	_	_
z6-GNW-21823	189.355591	62.312592	5.8228	_	-20.9 ± 0.4	-2.1 ± 1.3	23 ± 16	130 ± 17	-	_
z5-GNW-11701	189.055409	62.177498	5.8364	_	-19.6 ± 0.6	-1.8 ± 1.7	18 ± 18	75 ± 29	_	_
z5-GND-464	189.270721	62.148424	5.8745	-	-20.0 ± 0.1	-2.2 ± 0.2	57 ± 6	206 ± 18	-	_
Stark11-24923	189.284538	62.287491	5.9498	5.9483	-20.0 ± 0.2	-1.4 ± 1.1	30 ± 14	399 ± 131	0.14 ± 0.08	114 ± 56
Stark11-26902	189.307610	62.323443	5.9547	-	-19.0 ± 0.6	-1.4 ± 1.8	197 ± 215	587 ± 63	_	-
z6-GNW-25971	189.334732	62.286125	5.9549	-	-19.7 ± 0.5	-1.9 ± 1.3	38 ± 32	390 ± 75	_	_
z6-GND-14309	189.334047	62.263058	5.9676	-	-20.0 ± 0.2	-1.2 ± 1.0	104 ± 38	-	> 1.01	-
Stark11-6706	189.079712	62.141884	5.9719	-	-20.0 ± 0.5	-3.0 ± 1.5	31 ± 21	346 ± 62	_	_
Stark11-16773	189.197796	62.199983	5.9731	-	-19.3 ± 0.1	-2.7 ± 0.4	37 ± 7	198 ± 42	> 0.22	-
z6-GND-19165	189.347733	62.249866	6.0471	-	-20.0 ± 0.1	-2.0 ± 0.4	33 ± 6	363 ± 63	> 0.34	_
z7-GND-43678	189.235201	62.295597	6.1212	-	-19.7 ± 0.1	-1.2 ± 0.5	112 ± 25	46 ± 14	> 0.71	_
z6-GNW-2993	189.127197	62.122612	6.1357	-	-19.8 ± 0.5	-1.1 ± 1.6	47 ± 45	249 ± 117	_	-
Jung18-28438	189.178023	62.223718	6.5518	6.5442	-20.0 ± 0.2	-3.0 ± 0.9	30 ± 11	182 ± 36	0.21 ± 0.12	376 ± 25

ID	RA	DEC	z _{Lyα}	z _{sys}	M _{UV,1500}	β	$EW_{Ly\alpha}$	$FWHM_{Ly\alpha}$	$f_{\rm esc}^{{ m Ly}lpha}$	$\Delta v_{{ m Ly}lpha}$
							[Å]	$[{\rm km}{\rm s}^{-1}]$	[km s ⁻¹]	[km s ⁻¹]
Hu10-1	189.358260	62.207637	6.5604	_	-20.3 ± 0.1	-2.5 ± 0.8	166 ± 51	267 ± 25	_	_
Hu10-2	189.356888	62.295321	6.5760	_	-21.9 ± 0.1	-1.8 ± 0.4	24 ± 4	358 ± 6	_	_
Jung18-5752	189.199585	62.320965	6.5867	_	-19.2 ± 0.3	-1.9 ± 1.6	103 ± 72	174 ± 22	> 0.16	_
z6-GND-44831	189.175135	62.282267	6.7365	_	-20.6 ± 0.1	-2.9 ± 0.6	54 ± 13	127 ± 16	_	_
z7-GND-8358	189.155310	62.286461	6.8135	6.8072	-20.3 ± 0.1	-2.7 ± 0.5	11 ± 5	117 ± 52	_	276 ± 26
z8-GND-35384	189.231989	62.202333	6.8743	-	-19.9 ± 0.6	-1.7 ± 2.3	178 ± 198	41 ± 8	_	_

^{*:} Not sufficient photometric coverage in catalogs described in Section 2.4.

Appendix B: Spectra of Ly α -emitting galaxies: MMT/Binospec and NIRCam/JWST

