Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **539,** 1381–1403 (2025) https://doi.org/10.1093/mnras/staf530
Advance Access publication 2025 April 1

BOWIE-ALIGN: substellar metallicity and carbon depletion in the aligned TrES-4b with *JWST* NIRSpec transmission spectroscopy

Annabella Meech [1★] Alastair B. Claringbold [2,3] Eva-Maria Ahrer [4] James Kirk [5] Mercedes López-Morales [6] Jake Taylor [7] Richard A. Booth [8] Anna B. T. Penzlin [5] Lili Alderson [9] Duncan A. Christie [4] Emma Esparza-Borges [10,11] Charlotte Fairman [12] Nathan J. Mayne [13] Mason McCormack [14] James E. Owen [5] Vatsal Panwar [2,3] Diana Powell [14] Denis E. Sergeev [12] Daniel Valentine [12] Hannah R. Wakeford [12] Peter J. Wheatley [2,3] and Maria Zamyatina [13]

[1]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA*
[2]*Centre for Exoplanets and Habitability, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*
[3]*Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*
[4]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[5]*Department of Physics, Imperial College London, Prince Consort Rd, London, SW7 2AZ, UK*
[6]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[7]*Department of Physics, University of Oxford, Parks Rd, Oxford, OX1 3PU, UK*
[8]*School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, UK*
[9]*Department of Astronomy, Cornell University, 122 Sciences Drive, Ithaca, NY 14853, USA*
[10]*Instituto de Astrofísica de Canarias, C. Vía Láctea, San Cristóbal de La Laguna, E-38205, Spain*
[11]*Department of Astrophysics, University of La Laguna, San Cristóbal de La Laguna, E-38200, Spain*
[12]*School of Physics, HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK*
[13]*Department of Physics and Astronomy, University of Exeter, Exeter, EX4 4QL, UK*
[14]*Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL 60637, USA*

Accepted 2025 March 26. Received 2025 March 21; in original form 2025 January 15

ABSTRACT

The formation and migration history of a planet is expected to be imprinted in its atmosphere, in particular its carbon-to-oxygen (C/O) ratio and metallicity. The BOWIE-ALIGN (Bristol, Oxford, Warwick, Imperial, Exeter – A spectral Light Investigation into gas Giant origiNs) programme is performing a comparative study of *JWST* spectra of four aligned and four misaligned hot Jupiters, with the aim of characterizing their atmospheres and corroborating the link between the observables and the formation history. In this work, we present the $2.8 - 5.2\,\mu$m transmission spectrum of TrES-4 b, a hot Jupiter with an orbit aligned with the rotation axis of its F-type host star. Using free chemistry atmospheric retrievals, we report a confident detection of H_2O at an abundance of $\log X_{H_2O} = -2.98^{+0.68}_{-0.73}$ at a significance of 8.4σ. We also find evidence for CO and small amounts of CO_2, retrieving abundances $\log X_{CO} = -3.76^{+0.89}_{-1.01}$ and $\log X_{CO_2} = -6.86^{+0.62}_{-0.65}$ (3.1σ and 4.0σ, respectively). The observations are consistent with the atmosphere being in chemical equilibrium; our retrievals yield C/O between $0.30 - 0.42$ and constrain the atmospheric metallicity to the range $0.4 - 0.7\times$ solar. The inferred substellar properties (C/O and metallicity) challenge traditional models, and could have arisen from an oxygen-rich gas accretion scenario, or a combination of low-metallicity gas and carbon-poor solid accretion.

Key words: techniques: spectroscopic – exoplanets – planets and satellites: atmospheres – planets and satellites: composition – planets and satellites: gaseous planets.

1 INTRODUCTION

In the field of exoplanet atmospheres, we are in the era of exquisite data sensitivity, sufficient to probe chemical abundances and derive precise carbon-to-oxygen (C/O) ratios and atmospheric metallicities, thanks to the resolution and near-infrared wavelength coverage of the instruments on *JWST* (Ahrer et al. 2023; Alderson et al. 2023; Feinstein et al. 2023; Rustamkulov et al. 2023; JWST Transiting Exoplanet Community Early Release Science Team 2023). The challenge is now to interpret these measurements, linking them to exoplanet formation and evolution history. Early ideas (e.g. Öberg, Murray-Clay & Bergin 2011; Madhusudhan 2012) suggested that atmospheric C/O can be used to inform where a planet formed in the protoplanetary disc with respect to snow lines of the dominant

★ E-mail: annabella.meech@cfa.harvard.edu

volatiles, such as H_2O, CO, and CO_2. However, more recent studies have challenged that idea. For example, the radial composition profile of the disc varies between systems (Law et al. 2021), with different initial metallicities and host stars. Also, the radial location of snow lines in protoplanetary discs evolve over time (Morbidelli et al. 2016; Eistrup, Walsh & van Dishoeck 2018; Owen 2020), and volatile-carrying solids drift through the discs (e.g. Booth et al. 2017; Schneider & Bitsch 2021). More recently, Penzlin et al. (2024) demonstrated that the highly unconstrained nature of key planet formation and disc parameters – e.g. temperature profile, dust-to-gas ratio, and chemical composition within the disc – make it difficult to precisely predict the atmospheric composition of individual exoplanets from formation models. However, the key insight from Penzlin et al. (2024) is that comparing populations of planets (as opposed to single systems) with different migration histories could constrain planet formation models. Specifically, they show that the C/O and metallicity of close-in exoplanets that evolve through migration in a disc should be different from close-in exoplanets that undergo disc-free migration, after formation at another location in the disc.

The BOWIE-ALIGN survey (Bristol, Oxford, Warwick, Imperial, Exeter – A spectral Light Investigation into gas Giant origiNs; *JWST* GO 3838; PIs: Kirk & Ahrer) seeks to characterize the atmospheres of eight hot Jupiters in order to test tracers of planet formation. The prudently curated sample consists of four planets believed to have migrated through a disc, and four believed to have undergone disc-free migration, based on the current alignments of their orbital planes with respect to the rotational plane of their host stars. The targets are considered 'aligned' if their sky-projected obliquity $|\lambda| < 30°$, and 'misaligned' if $|\lambda| > 45°$, based on the definition in Spalding & Winn (2022). The main objectives, methods, and target selection for this programme are outlined in the survey paper (Kirk et al. 2024a), while the theoretical basis is explored in Penzlin et al. (2024). To date, the transmission spectrum of one target from this program has been published, the misaligned hot Jupiter WASP-15 b. In Kirk et al. (2025), we found WASP-15 b to host a superstellar metallicity atmosphere, with a solar C/O, and evidence of SO_2 absorption, the combination of which points to late planetesimal accretion. Here, we present the transmission spectrum of the second target in the BOWIE-ALIGN sample, TrES-4 b.

The hot Jupiter TrES-4 b ($R_p = 1.838^{+0.081}_{-0.090}\,R_J$, $M_p = 0.494 \pm 0.035\,M_J$, and $T_{eq} = 1795$ K; Sozzetti et al. 2015), discovered by Mandushev et al. (2007), is one of the four aligned planets in the sample. It orbits an F-type host star (Mandushev et al. 2007), which is crucially above the Kraft break, as are all of the targets in the programme. The 'Kraft break', at $T_{eff} \simeq 6100$ K, is a distinct shift in the rotation rates of stars, thought to occur due to the magnetic braking contribution of the convective zone in cooler stars (Kraft 1967; Beyer & White 2024). It was later observed that the hotter stars above the Kraft break are seen to host hot Jupiters with various orbital obliquities, while the cooler stars below show predominantly low stellar obliquities (Winn et al. 2010; Albrecht et al. 2012). Regarding the latter, misaligned hot Jupiters are thought to invoke stellar tides which can act to realign the stellar spin axis with the planet orbital plane. This is more efficient for the cooler stars; with deeper surface convective zones and consequent stronger magnetic braking, these stars have lower angular momentum which can be overcome by the planet (Dawson 2014; Lin & Ogilvie 2017). On the reverse, stars above the Kraft break have thin surface convective zones, which do not produce efficient tides. There is a period in which stars above the Kraft break (with masses $1.2\,M_\odot < M_* \lesssim 5\,M_\odot$[1]) may tidally re-align, since they are fully convective during their pre-main-sequence phase (Spalding & Winn 2022). However, TrES-4 b would need to orbit within 0.02 au of the star to have tidally re-aligned within the ~ 20 Myr period during which its host star was convective (Spalding & Winn 2022). Since the TrES-4 b lies at 0.05 au, it is highly unlikely to have undergone tidal re-alignment post migration. In this work, we base conclusions upon the assumption that TrES-4 b exhibits primordial alignment; its obliquity was measured by Narita et al. (2010) to be $\lambda = 6.3 \pm 4.7°$. As outlined in Kirk et al. (2024a), the presumption is thus that TrES-4 b formed in the outer regions of its system and consequently migrated towards the star, through the disc.

In this work, we aim to characterize the atmosphere of this hot Jupiter. We focus on constraining its C/O and metallicity, towards the wider goal of identifying trends in planet formation scenarios within the wider context of the BOWIE-ALIGN survey. We provide a description of the observations in Section 2. The procedures used to reduce the data and the results, including the transmission spectrum, are presented in Section 3. We then present our retrieval analysis of the extracted transmission spectrum in Section 4. Finally, the discussion of our results and concluding remarks are outlined in Sections 5 and 6.

2 OBSERVATIONS

TrES-4 was observed by *JWST* as part of the GO 3838 programme. We observed the full transit of planet b, using the Near Infrared Spectrograph (NIRSpec; Jakobsen et al. 2022) in Bright Object Time Series mode, with the NRSRAPID readout pattern. We used the *G395H* grating, *F290LP* filter, the S1600A1 aperture, and SUB2048 subarray. This setting provides a wavelength coverage of $2.8 - 5.2$ μm, and an average resolution of $R = 2700$. There exists a physical gap between the NIRSpec detectors. The positioning of this gap in wavelength depends on the dispersion-filter combination used, and falls between $3.72 - 3.82$ μm for *G395H*/*F290LP*. For the target acquisition, we used the WATA mode of NIRSpec on a fainter, nearby star 2MASSJ17531241 + 3712396, with the SUB32 array and S1600A1 slit. The science observation started at 01:27:06.022 UTC on 2024 June 17, and consisted of a total 378 integrations over 7.65 h, which included 180 integrations (3.65 h) in-transit, 85 pre- and 113 post-transit baseline (1.73 and 2.27 h, respectively), with 80 groups per integration.

3 DATA REDUCTION

To reduce these data, we used two separate pipelines, TIBERIUS[2] (Kirk et al. 2017, 2021) and EUREKA![3] (v.1.1; Bell et al. 2022). We detail each in turn below. Multiple, independent reductions have been standard in the community, as recommended by the work of the Early Release Science programme (Ahrer et al. 2023; Alderson et al. 2023; Feinstein et al. 2023; Rustamkulov et al. 2023; JWST Transiting Exoplanet Community Early Release Science Team 2023), and allow us to check how robust the transmission spectrum is to reduction choices. Additionally, we present a second independent reduction using the TIBERIUS pipeline in Appendix A.

[1]Stars more massive than $4 - 5$ $M_\odot$ become radiative on the pre-main sequence (e.g. Palla & Stahler 1993), and are often still accreting when they reach the zero-age main sequence (Zinnecker & Yorke 2007).
[2]https://tiberius.readthedocs.io/en/latest/
[3]https://eurekadocs.readthedocs.io/en/latest/

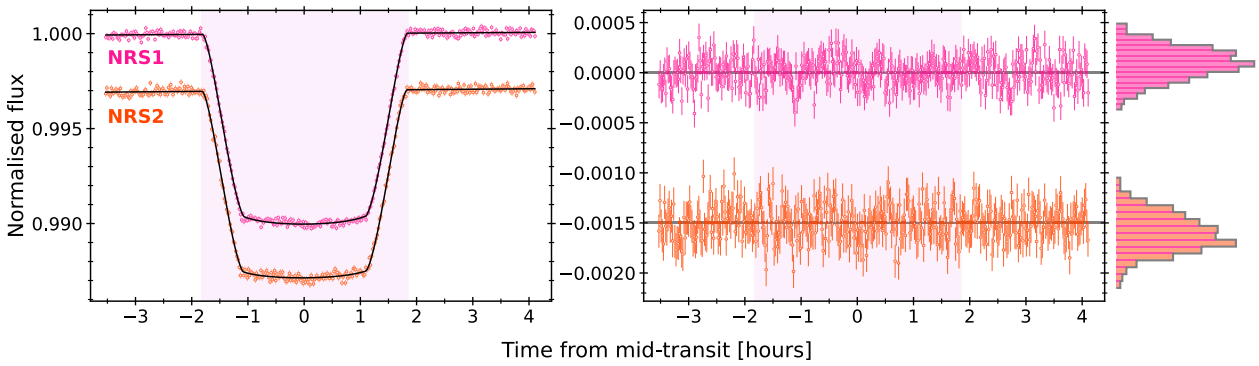

Figure 1. Left: extracted white light curves using the TIBERIUS pipeline, with best-fitting models overplotted. These have been plotted against time, having subtracted the best-fitting mid-transit time. Right: residuals and associated histograms between the light curves and their best-fitting models. The in-transit integrations are shaded for reference.

3.1 TIBERIUS

TIBERIUS is an open-source PYTHON package, written to reduce and fit exoplanet transit light curves (Kirk et al. 2017, 2021). The pipeline has been used by the community to reduce *JWST* exoplanet time-series data (see e.g. Esparza-Borges et al. 2023; Moran et al. 2023; Kirk et al. 2024b). Our implementation, described below, largely follows the standard routine, and is similar to that described in Kirk et al. (2025).

3.1.1 Light-curve extraction

We begin with the uncalibrated detector files [`uncal.fits`]. We initially process these using the stage 1 steps from the JWST pipeline[4] (v1.13.4) and apply a custom $1/f$ correction prior to the `ramp_fit` step, as in other applications of TIBERIUS (e.g. Kirk et al. 2024a). For this, we first fit the centre of the aperture trace with a fourth-order polynomial, and create a custom mask with a defined width centred on the fitted trace. In each column, we take the median of the remaining unmasked pixels in the detector image and subtract it from the column. We then run the `ramp_fit` and `gain_scale` steps, combining the 80 groups per integration.

To extract the light curves, we begin by applying the TIBERIUS cosmic ray correction to the `gainscalestep.fits` products; for each pixel we take a running median over every three integrations and flag any integrations in which the pixel value is more than four standard deviations away from that median. The flagged pixels are replaced by their median values. We take the cosmic ray-corrected `gainscalestep.fits` 2D images, and fit the stellar aperture trace. Using a user-defined initial guess row location and search width, first TIBERIUS fits each pixel column (in the cross-dispersion direction) with a Gaussian to find the peak flux. Then, we fit a fourth-order polynomial through the peaks of the fitted Gaussian for smoothing. To extract the stellar flux, we replace any bad pixels in the 2D image (the bad, saturated, hot, and dead pixels flagged in the `dq_init` step of stage 1, combined with the TIBERIUS-identified 5σ-clipped pixels) with the median of the pixels to the left and right. We define an aperture width of 6 pixels (chosen to minimize the noise in the white light curve), and perform aperture photometry, summing the flux centred on the fitted aperture trace for each integration. We extracted the stellar flux over pixels [600, 2040] and [5, 2040] inclusive (in the dispersion direction) for detectors NRS1 and NRS2, respectively. The result is a time-series of stellar spectra. For the wavelength solution, we make use of the `assign_wcs` product from the stage 1 JWST processing. The white light curves are shown in Fig. 1, having integrated over wavelengths 2.75–3.72 and 3.82–5.17 μm for NRS1 and NRS2, respectively.

We create the spectroscopic light curves by integrating over a defined pixel bin width; we use the binning schemes from Kirk et al. (2025), creating spectra at resolutions of $R \simeq 100$ (30/31, ∼60 pixel-wide wavelength bins for NRS1/2) and $R \simeq 400$ (120/121, ∼15 pixel-wide bins). In addition, we create 'high-resolution', pixel-level light curves which will be analysed in a future publication.

[4] https://jwst-pipeline.readthedocs.io/en/latest/

Table 1. Adopted system parameters.

Parameter [Unit]	Value
Star	
T_* [K]	6295 ± 65
M_* [$M_\odot$]	1.45 ± 0.05
R_* [$R_\odot$]	1.81 ± 0.08
$\log g_*$ [cgs]	4.09 ± 0.03
[Fe/H]	0.28 ± 0.09
Planet	
P [d]	$3.55392889 \pm 0.00000044^\dagger$
M_p[M_J]	0.494 ± 0.035
R_p[R_J]	$1.838^{+0.081}_{-0.090}$
T_{eq} [K]	1795^{+35}_{-39}
$e \cos \omega$	$0.0010^{+0.0022}_{-0.0017}$
$e \sin \omega$	$0^{+0.012}_{-0.022}$

Notes. † All values from Sozzetti et al. (2015), with the exception of the period, P, from Kokori et al. (2023).

3.1.2 Light-curve fitting

For the fitting of the light curves, we zero-centre the time arrays on the predicted mid-transit time, propagated from the latest published ephemeris (Tables 1 and 2; Kokori et al. 2023). Prior to fitting, we applied sigma clipping along the time axis with a 4σ threshold to remove outliers. The light-curve model consisted of a transit light-curve model (using the BATMAN package; Kreidberg 2015) multiplied by a systematics model. For the latter, we found a linear polynomial in time to sufficiently detrend the data, affording the lowest Bayesian Information Criterion compared to more complex models tested. This contributed two free parameters, the polynomial coefficients, c_n; we used wide, uniform priors, $|c_n| < 10$. For the transit model,

Table 2. Best-fitting planet parameters from the extracted white light curves. We only provide the weighted mean values for comparison; only the individual fit values highlighted in bold were used in this work.

			$t_{\rm mid}$ [BJD$_{\rm TDB}$]	a/R_*	i [deg]	$R_{\rm p}/R_*$
Literature value			60478.21264 ± 0.00045 *	6.04 ± 0.23 *	$83.1^{+0.5}_{-0.4}$ *	$0.10452^{+0.00066}_{-0.00072}$ †
TIBERIUS	Priors	–	$\lvert t_{\rm mid} - 60478.21\rvert < 0.1$	$a/R_* \geq 1$	$i \sim \mathcal{U}[0, 90]\, i - i_0 > -5^{\ddagger}$	$R_{\rm p}/R_* \sim \mathcal{U}[0, 0.5]$
	Individual	**NRS1**	$\mathbf{60478.212976^{+0.000039}_{-0.000040}}$	$\mathbf{5.941 \pm 0.016}$	$\mathbf{82.574^{+0.035}_{-0.034}}$	$\mathbf{0.100064^{+0.000069}_{-0.000067}}$
	Individual	**NRS2**	$\mathbf{60478.212839 \pm 0.000038}$	$\mathbf{5.945^{+0.017}_{-0.016}}$	$\mathbf{82.591 \pm 0.035}$	$\mathbf{0.099260 \pm 0.000066}$
	Weighted mean	–	$60478.212885 \pm 0.000023$	5.943 ± 0.011	82.582 ± 0.025	0.099650 ± 0.000047
EUREKA!	Priors	–	$t_{\rm mid} \sim \mathcal{N}[60478.21, 0.1]$	$a/R_* \sim \mathcal{N}[6.04, 1.0]$	$i \sim \mathcal{U}[80, 90]$	$R_{\rm p}/R_* \sim \mathcal{N}[\delta_0, 0.1]^{\ddagger}$
	Individual	NRS1	$60478.213010 \pm 0.000060$	5.925 ± 0.029	$82.528^{+0.067}_{-0.066}$	0.10005 ± 0.00021
	Individual	NRS2	$60478.213100^{+0.000077}_{-0.000079}$	$5.945^{+0.029}_{-0.028}$	$82.59^{+0.066}_{-0.063}$	0.09952 ± 0.00024
	Weighted mean	–	$60478.213043 \pm 0.000048$	5.935 ± 0.020	82.560 ± 0.046	0.09982 ± 0.00016

Notes. *Values from Kokori et al. (2023). The mid-transit time from Kokori et al. (2023) has been propagated to the epoch of our *JWST* observation, accounting for the uncertainties on the literature mid-transit time and period (provided in Table 1);
† Values from Sozzetti et al. (2015).
‡ i_0 is the initialized value, for which we use the literature value listed above. δ_0 is the literature value for $R_{\rm p}/R_*$.

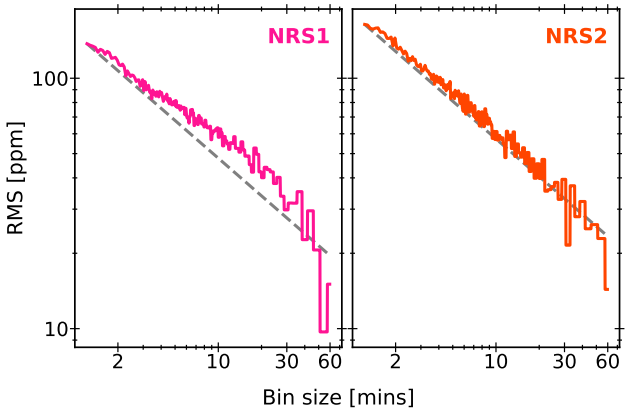

Figure 2. Allan variance plot for the individually fit TIBERIUS light curves presented in Fig. 1. The grey dashed lines show the expectation from white noise.

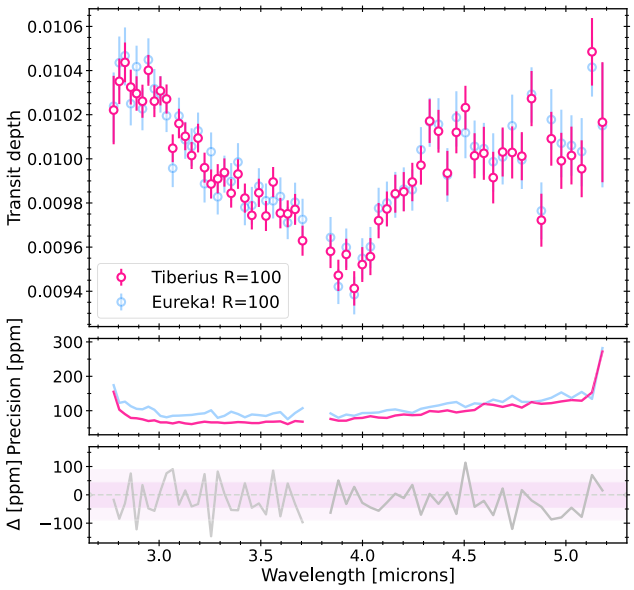

Figure 3. Top panel: the $R \simeq 100$ transmission spectrum of TrES-4 b from the TIBERIUS and EUREKA! reductions. Middle panel: the precision of each spectrum and bottom panel: the difference between the two reductions. The $1/2\sigma$ intervals are shaded for reference.

we leave $t_{\rm mid}$, a/R_*, i, and $R_{\rm p}/R_*$ as free parameters, again applying only loose, wide priors (given in Table 2). We fixed the orbital period to $P = 3.55392889$ d (Kokori et al. 2023). We also fix the eccentricity and longitude of periastron to $e = 0$ and $\omega = 90^\circ$, as consistent with literature values (see Table 1). Assuming a quadratic limb-darkening law, we used EXOTIC-LD with the 3D STAGGER stellar atmosphere model grid (Magic et al. 2015; Grant & Wakeford 2024) to generate limb-darkening coefficients for the published stellar parameters and fixed these in our light-curve fitting (NRS1: $u_1 = 0.0658$, $u_2 = 0.1117$; and NRS2: $u_1 = 0.0560$, $u_2 = 0.0928$).

We leveraged EMCEE (Foreman-Mackey et al. 2013) to explore the posterior distribution, with 30 walkers per free parameter. We run an initial burn-in with 10 000 steps which are discarded, rescale the photometric uncertainties to give a reduced chi-square $\chi^2_\nu = 1$, then run 10 000 production steps. In Fig. 1, we show the TIBERIUS-extracted white light curves, and their best-fitting models, having fit the detectors separately. The associated Allan (1966) variance plots are shown in Fig. 2. The parameters recovered from the transit light-curve models are given in Table 2, and the posterior distributions are displayed in Fig. B1.

We proceed to fit the wavelength-binned light curves, with linear systematics models for all wavelength bins. Again, we fixed the quadratic limb-darkening coefficients to that pre-computed with the same STAGGER grid, for the bin-centre wavelengths. Having fixed the $t_{\rm mid}$, a/R_*, and i to the fitted white light-curve values (using the individual detector parameters; Table 2), we fit $R_{\rm p}/R_*$ for each light curve. For the spectroscopic light-curve fits, we employ a Levenberg–Marquadt algorithm to optimize the model parameters (Moré 1978), scaling the uncertainties to give $\chi^2_\nu = 1$ and re-running the fit as done for the white light curves. The resulting TIBERIUS $R \simeq 100$ transmission spectrum is shown in pink in Fig. 3, with associated precision shown in the middle panel. The $R \simeq 400$ spectrum is shown in Fig. C2 of Appendix C.

To comment on the residuals from the white light-curve model fits (right panel, Fig. 1), we notice possible correlated noise at time -1.1 and $+0.4$ h from mid-transit. These features are more pronounced in the NRS1 light curve. Indeed, looking at the Allan variance in Fig. 2 there is some residual red noise in NRS1, while the residual noise closely resembles what we would expect of photon noise for NRS2. Such wavelength dependence may support a stellar origin, namely occultations of photospheric features, rather than a

systematic origin. To investigate this, we attempted to add a spot component to our transit model, with four additional free parameters: the (x, y) coordinates of the spot centre, the spot radius, and the spot contrast. We found no evidence for a spot-occultation in the white light curve. Further, we saw no evidence of a wavelength-dependent feature, as would be expected from a spot-crossing event, in the residuals of the spectroscopic light curves. We thus conclude these features are unlikely to be caused by occultations of star-spots.

3.2 EUREKA!

3.2.1 Light-curve extraction

We started our EUREKA! reduction with stages 1 and 2, which are wrapped around the default JWST pipeline[5] (v1.12.2). Within this wrapper and before the `ramp_fit` step, we applied a custom $1/f$ correction at the group level which uses the stage 3 column-by-column background subtraction setup. We subtracted a zero-order polynomial that was fitted to the background area (having masked the trace) and rejected outliers $> 3\times$ the median absolute deviation. We opted to increase the jump detection step at stage 1 to avoid false positives, as the default is often too low (set to 10.0 instead of 4.0) and we skipped the `photom_step` at stage 2.

Using EUREKA!'s stage 3 we extracted the stellar spectra. Within this stage we corrected for the curvature of the trace by identifying the central trace pixel in each column and moving the spectra such that the central trace pixel aligns. We performed column-by-column background subtraction: using the pixels more than 6 pixels away from the centre of the trace, we employ an outlier rejection threshold of $5\times$ the median in the spatial direction and a 7σ along the time axis following two iterations. Finally, we used optimal spectral extraction (Horne 1986) over a total aperture size of 9 pixels, to extract the time-series spectra over pixels [600, 2040] and [5, 2040] for detectors NRS1 and NRS2, respectively.

We bin our spectra into light curves using EUREKA!'s stage 4 at resolutions of $R \simeq 100$ and $\simeq 400$, equal to the binning scheme used in the TIBERIUS reduction and in previous BOWIE-ALIGN analyses. We utilize a 5σ-clipping from a rolling median of 25 pixels to mask outliers in the light curves (0–1 outliers per light curve in both unbinned and binned cases). In addition, we adopted manual masking prior to stage 4 to mask bad wavelength columns since we found some outliers from stage 3 were not masked sufficiently, and were propagating through to the light curves and transmission spectrum. For that we identified outliers in the spectra, which differed more than $20\times$ the uncertainty from the rolling mean of 25 pixels.

3.2.2 Light-curve fitting

Using EUREKA!'s stage 5 we fitted our light curves. First, we fit the white light curves of both NRS1 and NRS2 separately, freely fitting for the orbital parameters $t_{\rm mid}$, a/R_*, i, and $R_{\rm p}/R_*$ and using a simple linear-in-time systematics model. As with the TIBERIUS reduction, we fix the orbital period to $P = 3.55392889$ d (Kokori et al. 2023) and eccentricity to zero. For the limb-darkening, we used the quadratic limb-darkening law and left both u_1 and u_2 as free parameters, using a wide uniform prior. For all light-curve fitting within EUREKA!, we utilize the BATMAN package (Kreidberg 2015) for our transit model and the Markov Chain Monte Carlo (MCMC) package EMCEE (Foreman-Mackey et al. 2013) to retrieve our fitted parameters. We used 50 walkers per free parameter, 20 000 steps for the white light-curve fit (with 10 000 steps of that discarded as burn-in) and 1000 steps for the spectroscopic light curves (with 500 steps of that discarded as burn-in). The fitted values from the white light curves using EUREKA! can be found in Table 2.

For the purpose of comparing the transmission spectra directly between the TIBERIUS and EUREKA! reductions we fixed the system parameters (i, a/R_*, and $t_{\rm mid}$) to the TIBERIUS values for the spectroscopic light curves. This leaves six free parameters for each spectroscopic light curve: $R_{\rm p}/R_*$, limb-darkening coefficients u_1 and u_2, the two linear trend coefficients, and an error inflation term. The resulting EUREKA! $R \simeq 100$ transmission spectrum is shown in Fig. 3 alongside the equivalent TIBERIUS spectrum; the $R \simeq 400$ is shown in Appendix C (Fig. C2).

3.3 The transmission spectrum of TrES-4 b

We show the transmission spectrum of TrES-4 b from each pipeline, at a spectral resolution of $R \simeq 100$, in the top panel of Fig. 3. The $R \simeq 400$ transmission spectra are shown in Appendix C (Fig. C2). In Appendix A, we detail a third independent reduction, a second implementation of TIBERIUS (v2). The spectra from this reduction are also shown in Fig. C2, and the $R \simeq 100$ spectrum compared to the result from the primary TIBERIUS reduction in Fig. A1.

We see excellent agreement between the independent reductions. Indeed, the median differences between the $R \simeq 100$ spectra were 6 and 28 ppm for NRS1 and NRS2 (bottom panel Fig. 3), which were lower than the median precisions (79 and 101 ppm for TIBERIUS and EUREKA!, respectively; see the middle panel Fig. 3).

One key difference to highlight between the TIBERIUS and EUREKA! light-curve-fitting routines was the treatment of limb darkening. In the primary TIBERIUS reduction (Section 3.1.2), we used a quadratic limb-darkening law, with fixed limb-darkening coefficients. We also used a quadratic limb-darkening law in the fitting of the EUREKA! light curves, but the coefficients were left as free parameters. We obtained consistent transmission spectra, as described above, with larger uncertainties on the EUREKA! transit depths, which would be consistent with the higher model dimensionality. We found similar inflation of the transit depth uncertainties in the secondary TIBERIUS reduction, described in Appendix A, when fitting for the limb-darkening coefficients. In summary, our tests attest to the robustness of the transmission spectrum against choice of limb-darkening treatment. We proceed to analyse the primary TIBERIUS and the EUREKA! transmission spectra using atmospheric retrievals.

4 RETRIEVALS

In order to interpret our transmission spectra, we perform a series of atmospheric retrieval tests on both the TIBERIUS and EUREKA! transmission spectra, with the publicly available packages POSEIDON and PETITRADTRANS. The configuration and priors used for each retrieval are summarized in Tables D1 and D2. We detail each in turn below.

4.1 POSEIDON

We leverage the PYTHON package POSEIDON[6] (v1.2.1; MacDonald & Madhusudhan 2017; MacDonald 2023) to perform free chemistry and equilibrium chemistry atmospheric retrievals on the transmission spectrum of TrES-4 b. POSEIDON uses the forward-modelling tool `TRIDENT` (MacDonald & Lewis 2022) to compute the transmission spectra, integrated with a nested sampling framework to explore the parameter space. For the forward models, we use a resolution of

[5] https://jwst-pipeline.readthedocs.io/en/latest/

[6] https://poseidon-retrievals.readthedocs.io/en/latest/index.html

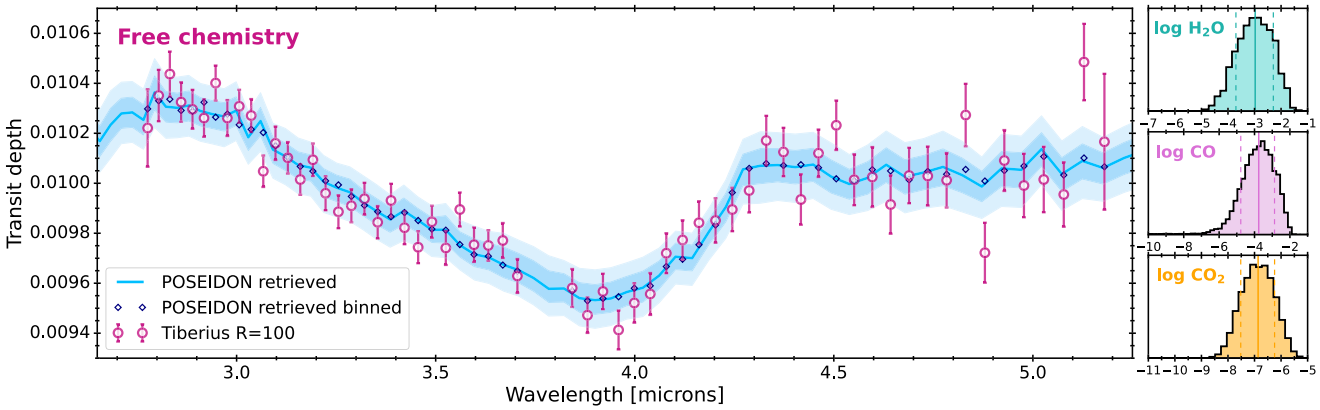

Figure 4. Main: median retrieved POSEIDON spectrum for the free chemistry retrieval, smoothed to a resolution of $R = 100$ (with $1/2\sigma$ intervals in successively lighter shades), for the TIBERIUS $R \simeq 100$ transmission spectrum. The median retrieved spectrum convolved with the instrument point spread and sensitivity function, then binned to the data resolution is also shown in the diamond data points. Right: posterior distributions for the log volume mixing ratios of H_2O (top), CO (middle), and CO_2 (bottom).

$R = 20,000$; high resolution is recommended when using opacity sampling (see e.g. Garland & Irwin 2019). We instantiate a H_2/He-dominated (with fixed ratio He/H_2 = 0.17; Asplund et al. 2009) atmosphere with trace gases CH_4, CO_2, H_2O (ExoMol; Polyansky et al. 2018; Yurchenko et al. 2020, 2024), CO (Li et al. 2015), H_2S (Azzam et al. 2016), HCN (Barber et al. 2014), and SO_2 (ExoAmes; Underwood et al. 2016). We also adopt contributors to continuum absorption, namely collision-induced absorption (dominant for NIRSpec *G*395*H* wavelengths). We use H_2–H_2, H_2–He, H_2–CH_4, CO_2–H_2, CO_2–CO_2, and CO_2–CH_4 cross-sections from the high-resolution transmission molecular absorption (HITRAN) database (Karman et al. 2019; MacDonald & Lewis 2022).

The atmosphere of TrES-4 b is modelled with 100 layers evenly distributed in log-pressure, from $P_{\rm min} = 10^{-7}$ to $P_{\rm max} = 10^2$ bar. In all of these retrievals, we allow the reference radius $R_{\rm p,ref}$ (defined at a reference pressure of $P_{\rm ref} = 10$ bar) to vary freely, having applied a uniform prior with a bounded range of 20 per cent the average of the fitted white light-curve radius (Table 2). Where specified,

(i) we employ the MacDonald & Madhusudhan (2017) cloud prescription, selecting a cloud deck + haze layer as defined by parameters: opaque cloud deck pressure, $P_{\rm cloud}$, the Rayleigh-enhancement factor, a, and a scattering slope, γ. The haze opacity is parametrized as $\kappa_{\rm cloud} = a\sigma_0(\lambda/\lambda_0)^\gamma$ above the cloud deck pressure, $P_{\rm cloud}$ (below which the atmosphere is rendered opaque). The remaining parameters are constants: σ_0 is the H_2 Rayleigh scattering cross section at reference wavelength λ_0 (5.31×10^{-31} m^2 and 350 nm, respectively). This cloud is uniform across the terminators. This adds three free parameters to the retrieval;

(ii) we allow the log g of the planet to vary, having applied a Gaussian prior according to the solution given by the published mass ($M_{\rm p} = 0.494 \pm 0.035\ M_{\rm J}$; Sozzetti et al. 2015) and average of the radii measured from our TIBERIUS white light curves. This adds one free parameter to the retrieval;

(iii) we include an offset between detectors NRS1 and NRS2, adding one free parameter.

Regarding the star, we fix all parameters to the literature values, provided in Table 1. We employ the nested sampling algorithm PYMULTINEST (Feroz, Hobson & Bridges 2009; Buchner et al. 2014), with 1000 live points, to explore the parameter space.

4.1.1 *Free chemistry*

We first test atmospheric models wherein the abundances of the atmospheric species are allowed to vary freely. For these free chemistry tests, we implement an isothermal pressure–temperature (PT) profile. Our reference retrieval test:

(I) a free chemistry retrieval including the species mentioned in the preceding section, a cloud deck and haze parametrization, and an offset between detectors, totalling 14 free parameters.

For each of the trace species, we assume constant abundance with pressure, and apply a wide uniform prior on the volume mixing ratio of each ($-12 < \log_{10} X_i < -1$ dex; see Table D1).

Test (I). On the $R \simeq 100$ TIBERIUS spectrum, we retrieve abundances $\log X_{\rm H_2O} = -2.98^{+0.68}_{-0.73}$, $\log X_{\rm CO} = -3.76^{+0.89}_{-1.01}$, and $\log X_{\rm CO_2} = -6.86^{+0.62}_{-0.65}$. We plot the median retrieved spectrum in Fig. 4, together with the posterior distributions for H_2O, CO, and CO_2. We provide all reference retrieval results in Tables 3 and 4 and the associated posterior distributions are shown in Fig. D3 (Appendix D). The median retrieved spectrum from the reference retrieval (I) yields a reduced chi-square of $\chi^2_\nu = 1.27$, with 47 degrees of freedom. In Fig. 5, we plot the contributions from individual gas opacities to the overall median retrieved spectrum.

To quantify the contribution of each parameter in the model, we run series of 'nested model' retrievals, with one contribution removed at a time. We assign detection significance by Bayesian model comparison. Comparing the Bayesian evidence $\mathcal{Z}$ to that of the reference model, we compute a Bayes' factor which is then transformed to a frequentist detection 'sigma' significance (Trotta 2008; Benneke & Seager 2013). This significance nominally quantifies the level at which the added complexity contributes to a better model fit, and is a standard metric in the reporting of exoplanet atmospheric characteristics (see e.g. Benneke & Seager 2013; Welbanks & Madhusudhan 2021; Taylor et al. 2023).

For the TIBERIUS $R \simeq 100$ spectrum, we detect H_2O and CO_2 to 8.4σ and 4.0σ, respectively (12.3σ and 4.5σ for the $R \simeq 400$ transmission spectrum). We detect CO to 3.1σ (2.4σ) but find insufficient evidence in favour of CH_4, HCN, or SO_2. We see possible hints of H_2S; Bayesian model comparison yielded a 1.2σ model preference. This is insufficient to claim a detection, and the posterior was non-Gaussian (Fig. D3). Still, we place a 2σ upper limit $\log X_{\rm H_2S} < -4.51$, i.e. 31 ppm; for reference, the solar sulphur

Table 3. Retrieval results: mean and 1σ confidence intervals of chemical mixing ratios, $\log X_i$, from the free chemistry retrievals, as presented in Section 4. For non-detected species, we provide the 2σ upper limits. We also provide values for the chemical equilibrium and hybrid chemistry retrievals, listing the mean and 1σ uncertainties, at the pressure of maximum contribution (see e.g. Fig. D1).

Input spectrum	H_2O	CO	CO_2	CH_4	H_2S	SO_2	HCN
POSEIDON							
Free chemistry							
TIBERIUS $R = 100$	$-2.98^{+0.68}_{-0.73}$	$-3.76^{+0.89}_{-1.01}$	$-6.86^{+0.62}_{-0.65}$	< -7.07	< -4.51	< -6.97	< -6.97
TIBERIUS $R = 400$	-2.88 ± 0.61	$-3.97^{+0.95}_{-1.25}$	$-6.81^{+0.56}_{-0.54}$	< -7.01	< -4.98	< -7.25	< -6.82
EUREKA! $R = 100$	$-2.88^{+0.61}_{-0.72}$	$-3.93^{+0.89}_{-1.03}$	$-7.11^{+0.63}_{-0.67}$	< -7.01	< -5.18	< -7.38	< -7.06
EUREKA! $R = 400$	$-2.89^{+0.59}_{-0.53}$	$-3.71^{+0.89}_{-1.03}$	-6.94 ± 0.56	< -7.47	< -5.41	< -7.37	< -6.20
Equilibrium chemistry							
TIBERIUS $R = 100$	$-3.40^{+0.26}_{-0.23}$	$-3.46^{+0.24}_{-0.26}$	-7.04 ± 0.43	$-9.38^{+0.69}_{-0.68}$	$-4.77^{+0.23}_{-0.20}$	$-13.31^{+1.05}_{-1.02}$	$-10.48^{+0.26}_{-0.25}$
TIBERIUS $R = 400$	$-3.32^{+0.27}_{-0.24}$	$-3.56^{+0.26}_{-0.32}$	$-7.02^{+0.51}_{-0.49}$	$-9.20^{+0.64}_{-0.75}$	$-4.75^{+0.24}_{-0.25}$	$-13.48^{+1.16}_{-1.09}$	$-10.67^{+0.24}_{-0.22}$
PETITRADTRANS							
Free chemistry							
TIBERIUS $R = 100$	$-3.29^{+0.57}_{-0.59}$	$-3.71^{+0.85}_{-0.84}$	$-7.21^{+0.56}_{-0.51}$	< -7.54	< -4.87	< -7.24	< -7.03
TIBERIUS $R = 400$	$-3.26^{+0.51}_{-0.56}$	$-3.74^{+0.80}_{-0.83}$	$-7.13^{+0.49}_{-0.52}$	< -7.20	< -4.57	< -7.12	< -7.14
EUREKA! $R = 100$	$-3.18^{+0.57}_{-0.65}$	$-3.31^{+0.84}_{-0.95}$	$-7.19^{+0.59}_{-0.64}$	< -7.53	< -5.17	< -7.03	< -7.03
Hybrid chemistry							
TIBERIUS $R = 100$	$-3.49^{+0.24}_{-0.25}$	$-3.74^{+0.22}_{-0.26}$	$-7.30^{+0.48}_{-0.47}$	$-11.78^{+0.69}_{-0.76}$	< -4.70	< -6.90	$-12.28^{+0.26}_{-0.23}$
TIBERIUS $R = 400$	$-3.45^{+0.23}_{-0.25}$	$-3.75^{+0.21}_{-0.23}$	$-7.25^{+0.45}_{-0.48}$	$-11.68^{+0.70}_{-0.59}$	< -4.64	< -6.86	$-12.33^{+0.25}_{-0.22}$
EUREKA! $R = 100$	-3.61 ± 0.26	$-3.62^{+0.27}_{-0.25}$	$-7.34^{+0.51}_{-0.46}$	$-11.88^{+0.80}_{-0.61}$	< -4.96	< -6.74	$-11.99^{+0.30}_{-0.27}$
Equilibrium chemistry							
TIBERIUS $R = 100$	$-3.48^{+0.24}_{-0.25}$	$-3.63^{+0.22}_{-0.26}$	$-7.20^{+0.48}_{-0.47}$	-11.89 ± 0.69	$-4.95^{+0.21}_{-0.20}$	–	$-12.13^{+0.28}_{-0.27}$
TIBERIUS $R = 400$	$-3.45^{+0.29}_{-0.25}$	$-3.67^{+0.29}_{-0.34}$	$-7.18^{+0.50}_{-0.48}$	$-11.75^{+0.69}_{-0.76}$	$-4.98^{+0.23}_{-0.25}$	–	$-12.22^{+0.26}_{-0.23}$

abundance is approximately 13 ppm (Asplund et al. 2009). We also investigated the presence of clouds in the atmosphere of TrES-4 b. In the reference retrieval (I) on the TIBERIUS $R \simeq 100$ transmission spectrum, an opaque cloud deck level was retrieved at a pressure of $\log P_{\rm cloud} = 0.20^{+1.14}_{-1.13}$; looking at the pressure contribution function (see Fig. D1; Mullens, Lewis & MacDonald 2024), this cloud deck is below the pressures probed and not likely to provide a significant contribution to the observed spectrum. Indeed, the evidence for the baseline model is $\ln \mathcal{Z} = 468.87 \pm 0.12$, and then $\ln \mathcal{Z} = 469.96 \pm 0.12$ when we remove the cloud + haze layer from the model. There is therefore insufficient evidence to support the presence of a cloud deck or haze layer, and we place a conservative 2σ upper limit of $\log P_{\rm cloud}\,[\rm bar] > -1.70$. Further, the reference retrieval recovers an offset between detectors NRS1 and NRS2 that is consistent with zero ($\delta_{\rm rel} = -10^{+466}_{-476}$ ppm), though not well constrained. We note that when removing the offset parameter from the retrieval, we recover chemical abundances that are completely in agreement with the reference retrieval. We therefore do not see evidence for a detector offset in the transmission spectrum.

We run the same retrieval tests on both of the EUREKA! transmission spectra. The baseline model (I) from the reference retrieval results in a reduced chi-square of $\chi^2_\nu = 1.45$, with Bayesian evidence $\ln \mathcal{Z} = 452.54 \pm 0.12$. We retrieved H_2O, CO, and CO_2 at abundances comparable to those retrieved for the TIBERIUS spectra, at significances of 9.8σ, 2.9σ, and 3.0σ, respectively, for the $R \simeq 100$ EUREKA! spectrum (9.3σ, 2.8σ, and 3.6σ for the EUREKA! $R \simeq 400$ spectrum). In this case, we also find no evidence for a detector offset, clouds, CH_4, H_2S, HCN, or SO_2.

4.1.2 *Chemical equilibrium*

We independently test atmospheric compositions wherein the relative abundances of the atmospheric species are governed by equilibrium chemistry. POSEIDON interpolates over a pre-computed chemistry grid from `FastChem` (Stock et al. 2018; Stock, Kitzmann & Patzer 2022) for the equilibrium calculations. We employ the subset of chemical species indicated in Section 4, and a gradient PT profile, linearly varying with $\log P$ between two temperatures, $T_{\rm deep}$ and $T_{\rm high}$. We apply equal priors to the high and deep temperatures, with no user-defined prior on $\Delta T = T_{\rm deep} - T_{\rm high}$ or $\delta T/\delta P$ (Table D1). Our reference retrieval includes:

(II) an atmosphere under equilibrium chemistry, with a cloud deck and haze layer. We include an offset between detectors. Since the chemistry is entirely defined by the carbon-to-oxygen (C/O) ratio and the metallicity relative to solar metallicity ($Z/Z_\odot$)[7], we have a total of 10 free parameters.We used uniform priors for C/O and metallicity, across the full ranges supported by POSEIDON ($\rm C/O \sim \mathcal{U}[0.2, 2.0]$; $\rm [M/H]_{O/H} \sim \mathcal{U}[-1, 4]$).

Test (II). For the $R \simeq 100$ TIBERIUS transmission spectrum, we retrieve $\rm C/O = 0.35^{+0.12}_{-0.10}$ and $\log Z/Z_\odot = -0.15^{+0.27}_{-0.26}$ ($\rm C/O = 0.32^{+0.11}_{-0.08}$ and $\log Z/Z_\odot = -0.23^{+0.30}_{-0.28}$ for the $R \simeq 400$ spectrum).[8] We retrieve a reference radius and planetary $\log g$ comparable to those of the free retrieval. The median retrieved spectrum is shown in Fig. 6, along with posterior distributions for C/O, and metallicity (see Appendix D, Fig. D4 for the full corner plot), and yields a reduced chi-square of $\chi^2_\nu = 1.13$ with 51 degrees of freedom. In

[7]Throughout this work, the $Z/Z_\odot$ notation refers to metallicity set by C/H, i.e. $\rm [M/H]_{C/H}$.

[8]We note that we have converted the retrieved POSEIDON metallicities to the same scale as PETITRADTRANS, where the metallicity is set by C/H, and the O/H is then calculated from the metallicity and the C/O ratio, listed as $\rm [M/H]_{C/H}$, for ease of comparison. The metallicity posterior in Fig. D4 is a direct output of POSEIDON with the metallicity set by O/H. We discuss this further in Section 4.3.

Table 4. Retrieval results (mean and 1σ confidence intervals) as presented in Section 4. For the unconstrained parameters, we provide the 2σ upper limits.

Input spectrum	Number	ln $\mathcal{Z}$	$R_{\mathrm{p,ref}}$ [R_{J}]	log g [cgs]	T [K]†		log P_{cloud} [bar]	log a	γ	C/O	log $Z/Z_{\odot}$ ‡	Offset [ppm]
POSEIDON												
Free chemistry												
TIBERIUS $R = 100$	(I)	468.9 ± 0.1	$1.63^{0.01}_{-0.02}$	2.57 ± 0.04	1180^{+139}_{-138}	–	> -1.70	< 7.41	< -0.97	–	–	< 939
TIBERIUS $R = 400$		1727.7 ± 0.1	1.64 ± 0.01	$2.57^{+0.04}_{-0.03}$	1089^{+132}_{-94}	–	> -1.44	< 7.56	< -1.05	–	–	< 939
EUREKA! $R = 100$		452.5 ± 0.1	1.65 ± 0.01	2.55 ± 0.05	1005^{+193}_{-144}	–	> -1.32	< 7.43	< -1.22	–	–	< 907
EUREKA! $R = 400$		1710.3 ± 0.1	1.65 ± 0.01	$2.57^{+0.04}_{-0.05}$	981^{+135}_{-110}	–	> -1.17	< 7.43	< -1.26	–	–	< 934
Equilibrium chemistry												
TIBERIUS $R = 100$	(II)	473.3 ± 0.1	$1.61^{+0.03}_{-0.04}$	$2.57^{0.04}_{-0.05}$	1352^{+149}_{-184}	1235^{+420}_{-350}	$0.21^{+1.14}_{-1.22}$	$1.08^{+4.14}_{-3.35}$	$-11.6^{+6.0}_{-5.4}$	$0.35^{+0.12}_{-0.10}$	$-0.15^{+0.27}_{-0.26}$	-6^{+472}_{-459}
TIBERIUS $R = 400$		1730.2 ± 0.1	$1.61^{+0.02}_{-0.04}$	$2.57^{+0.05}_{-0.06}$	1283^{+153}_{-254}	1287^{+412}_{-358}	$0.23^{+1.18}_{-1.16}$	$1.16^{+4.23}_{-3.42}$	$-11.6^{+6.1}_{-5.4}$	$0.32^{+0.11}_{-0.08}$	$-0.23^{+0.30}_{-0.28}$	-27^{+457}_{-452}
EUREKA! $R = 100$		455.7 ± 0.1	1.61 ± 0.03	$2.56^{+0.06}_{-0.07}$	1279^{+241}_{-295}	1298^{+388}_{-364}	$0.23^{+1.16}_{-1.13}$	$1.05^{+4.13}_{-3.32}$	$-11.7^{+6.1}_{-5.2}$	$0.40^{+0.15}_{-0.12}$	$-0.26^{+0.37}_{-0.33}$	$+20^{+461}_{-465}$
EUREKA! $R = 400$		1714.0 ± 0.1	$1.62^{+0.02}_{-0.03}$	$2.62^{+0.05}_{-0.07}$	1365^{+195}_{-313}	1329^{+375}_{-385}	$0.22^{+1.16}_{-1.19}$	$1.08^{+4.12}_{-3.38}$	$-11.3^{+5.8}_{-5.4}$	$0.40^{+0.13}_{-0.11}$	-0.04 ± 0.36	$+13^{+478}_{-464}$
TIBERIUS $R = 100$	(III)	473.2 ± 0.1	1.61 ± 0.03	$2.56^{+0.04}_{-0.05}$	1324^{+168}_{-205}	1301^{+385}_{-359}	$0.14^{+1.14}_{-1.15}$	$1.06^{+3.77}_{-3.20}$	$-11.9^{+5.8}_{-4.9}$	$0.37^{+0.13}_{-0.10}$	$-0.14^{+0.27}_{-0.26}$	$+30^{+417}_{-430}$
PETITRADTRANS												
Free chemistry												
TIBERIUS $R = 100$	(IV)	-51.4 ± 0.1	1.80 ± 0.02	$2.55^{+0.03}_{-0.04}$	1157^{+94}_{-123}	–	$0.36^{+1.08}_{-1.21}$	–	–	–	–	–
TIBERIUS $R = 400$		-143.9 ± 0.1	1.80 ± 0.02	2.54 ± 0.03	1132^{+88}_{-101}	–	$0.28^{+1.09}_{-1.16}$	–	–	–	–	–
EUREKA! $R = 100$		-49.3 ± 0.1	1.80 ± 0.02	2.56 ± -0.04	1138^{+128}_{-133}	–	$0.28^{+1.12}_{-1.17}$	–	–	–	–	–
Equilibrium chemistry												
TIBERIUS $R = 100$	(V)	-44.5 ± 0.1	1.80 ± 0.01	2.52 ± 0.03	1223^{+116}_{-65}	–	$0.29^{+1.18}_{-1.29}$	–	–	$0.36^{+0.16}_{-0.12}$	$-0.29^{+0.30}_{-0.23}$	–
TIBERIUS $R = 400$		-139.1 ± 0.1	1.80 ± 0.01	2.52 ± 0.03	1198^{+78}_{-67}	–	$0.31^{+1.14}_{-1.23}$	–	–	$0.33^{+0.13}_{-0.11}$	$-0.33^{+0.24}_{-0.19}$	–
EUREKA! $R = 100$		-44.9 ± 0.1	1.80 ± 0.02	2.54 ± 0.04	1241^{153}_{-81}	–	$0.33^{+1.09}_{-1.27}$	–	–	$0.42^{+0.20}_{-0.15}$	$-0.28^{+0.35}_{-0.27}$	–
Hybrid chemistry												
TIBERIUS $R = 100$	(VI)	-45.7 ± 0.1	1.80 ± 0.01	2.55 ± 0.03	1198^{+70}_{-63}	–	$0.24^{+1.17}_{-1.22}$	–	–	$0.32^{+0.13}_{-0.10}$	$-0.40^{+0.24}_{-0.20}$	–
TIBERIUS $R = 400$		-139.1 ± 0.1	1.80 ± 0.01	$2.54^{+0.03}_{-0.04}$	1180^{+47}_{-59}	–	$0.33^{+1.16}_{-1.25}$	–	–	$0.30^{+0.11}_{-0.09}$	-0.41 ± 0.18	–
EUREKA! $R = 100$		-46.4 ± 0.1	1.81 ± 0.01	2.56 ± 0.03	1203^{+90}_{-67}	–	$0.28^{+1.15}_{-1.19}$	–	–	$0.38^{+0.14}_{-0.13}$	-0.39 ± 0.25	–

Notes. †If one value given, an isothermal temperature profile was used. If two values given, a gradient PT profile was used, and the values are ordered as T_{high} and T_{deep}.
‡The metallicities reported are set by [C/H] (native for PETITRADTRANSand converted for POSEIDON).
(I) and (II) are the POSEIDON reference free and chemical equilibrium retrievals presented in Sections 4.1.1 and 4.1.2, respectively;
(III) is the POSEIDON stellar contamination retrieval presented in Section 4.1.3; (IV), (V), and (VI) are the PETITRADTRANS free, equilibrium, and hybrid chemistry retrievals respectively, presented in Section 4.2.

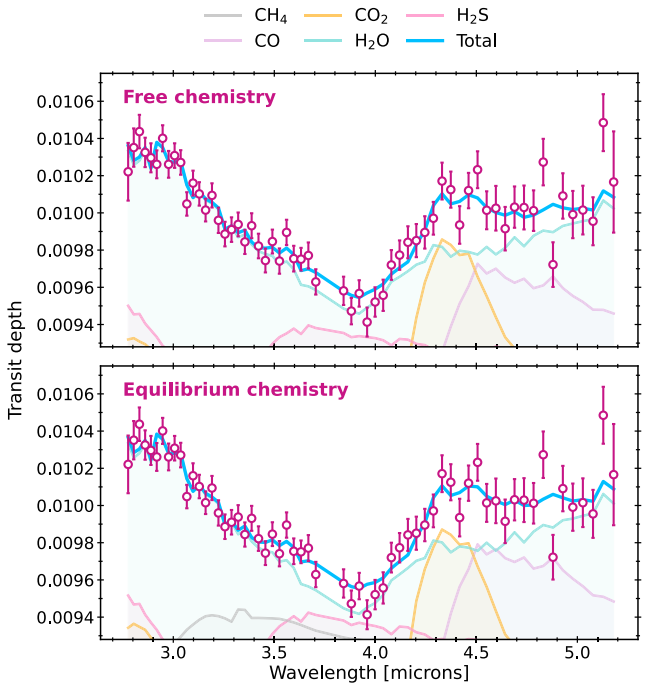

Figure 5. Spectral contributions, convolved to the instrument PSF and binned to $R = 100$, from the (top) free chemistry and (bottom) equilibrium chemistry POSEIDON retrievals on the $R \simeq 100$ TIBERIUS spectrum. The total spectrum is the median retrieved as shown in Figs 4 and 6. Though HCN and SO_2 were included in the equilibrium chemistry retrieval, their constrained contributions are not within the ranges plotted here.

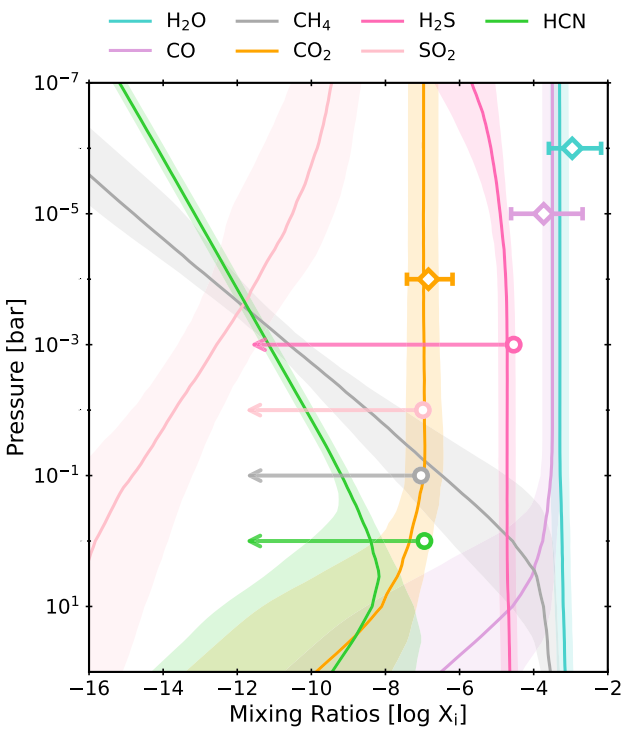

Figure 7. Vertical chemical profiles (median and 1σ confidence intervals) of the trace species in chemical equilibrium, from the best-fitting POSEIDON retrieval on the $R \simeq 100$ TIBERIUS transmission spectrum (see Section 4.1.2). We retrieve a $C/O = 0.35^{+0.12}_{-0.10}$ and $\log Z/Z_\odot = -0.15^{+0.27}_{-0.26}$. Abundance constraints from the free chemistry retrieval on the same transmission spectrum in Section 4.1.1 are overplotted (at arbitrary pressures): we retrieve bounded constraints on H_2O, CO, and CO_2 (diamond markers). We display the 2σ upper limits on the other, non-detected species (circle markers).

Fig. 7, we show the vertical mixing profiles (mean and 1σ intervals) for the model in Fig. 6, and provide the abundances at the maximum pressure contribution (see Fig. D1) in Table 3. The low abundances of (e.g.) SO_2 and HCN are not actually observed (see Fig. 5), but rather constrained by the equilibrium model abundances.

We note that we did investigate a more complex PT profile, namely the Madhusudhan & Seager (2009) profile. This parametrization divides the atmospheres into three layers, provisionally allowing for a temperature inversion within the middle layer, and has six free parameters. We show the median retrieved PT profile along with the transmission spectrum contribution function, and the simpler gradient profile in Fig. D1. We saw no evidence for a temperature inversion, and notably the retrieved limb temperature is in agreement with that retrieved using the simpler gradient PT profile ($T_{\rm ref} = 1313^{+104}_{-89}$ K at $P_{\rm ref} = 10$ bar for the Madhusudhan & Seager 2009 PT profile). We retrieved completely consistent atmospheric parameters, including $C/O = 0.35^{+0.12}_{-0.09}$ and $\log Z/Z_\odot = -0.17^{+0.27}_{-0.25}$. We thus favour the model with the less complex gradient PT profile presented in the preceding paragraph, as recommended by Schleich et al. (2024).

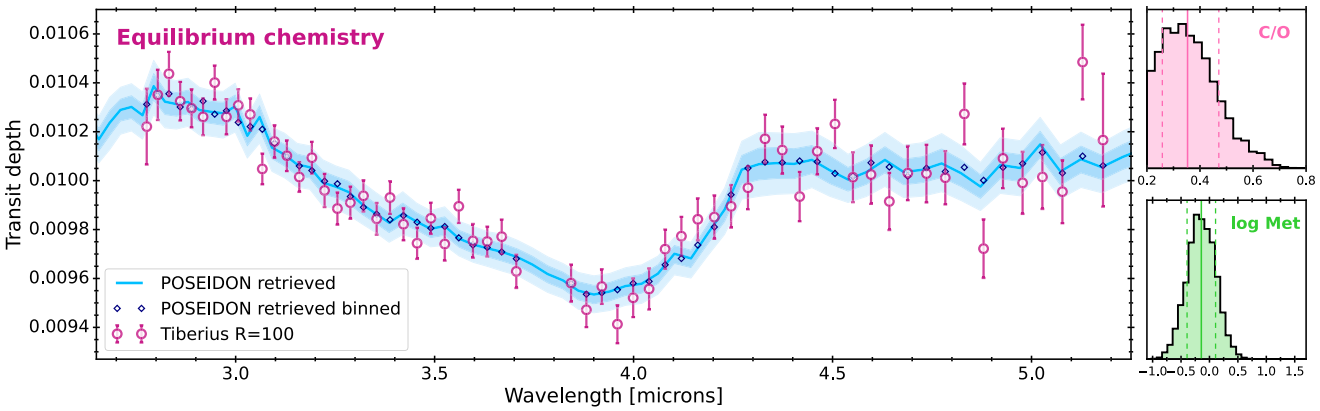

Figure 6. Main: median retrieved equilibrium chemistry spectrum with POSEIDON, in the same format as Fig. 4, again for the TIBERIUS $R \simeq 100$ transmission spectrum. Right: posterior distributions for C/O (top), and metallicity (bottom).

We run the same reference retrieval tests on the EUREKA! spectra. The reference retrieval on the $R \simeq 100$ spectrum yields a C/O comparable to that retrieved for the equivalent TIBERIUS spectrum, $\mathrm{C/O} = 0.40^{+0.15}_{-0.12}$ ($\mathrm{C/O} = 0.40^{+0.13}_{-0.11}$ for the $R \simeq 400$ spectrum), with the corresponding model $\chi^2_\nu = 1.33$. However, we recover a slightly varied metallicities for the EUREKA! spectra: $\log Z/Z_\odot = -0.26^{+0.37}_{-0.33}$ and -0.04 ± 0.36 for the $R \simeq 100$ and 400 spectra, respectively. These are still consistent with those retrieved for the

TIBERIUS spectra within 1σ. Again, all the retrieved parameters are given in Table 4.

4.1.3 *Stellar heterogeneities*

Stellar surface heterogeneities have the potential to affect the observed transmission spectrum of a transiting planet. Unocculted dark spots and bright regions introduce a wavelength-dependent flux variation, and since the transit depth is inferred assuming the planet atmosphere is entirely illuminated by the homogeneous photospheric spectrum, this can leave artefacts in the transmission spectrum. This is a known phenomenon, dubbed the 'transit light source effect' (TLSE; Rackham, Apai & Giampapa 2018, 2019). This has been shown to introduce degeneracies and even entirely account for transmission spectra in the near-infared, including with NIRSpec *G395H* observations (see e.g. Moran et al. 2023). Though one might not expect significant contributions from heterogeneities on the F-type host, we nevertheless run a test retrieval with a contamination factor included, considering the residual variability seen in the white light curves (Section 3.1). To test for stellar contamination:

(III) we model one photospheric heterogeneity feature with an assumed PHOENIX model stellar spectrum defined by a characteristic temperature $T_{\rm het}$, and gravitational field strength $\log g_{\rm het}$. The planet atmosphere is assumed to be under equilibrium chemistry and is parametrized as described in Section 4.1.2.Effectively it is a heterogeneity-photosphere temperature contrast over which we marginalize, since we also fit for the equivalent photospheric characteristics, $T_{\rm phot}$ and $\log g_{\rm phot}$. Therefore, along with the heterogeneity filling factor $f_{\rm het}$, this introduces five additional free parameters to the retrieval. Using POSEIDON's in-built functionality, we compute the transmission spectrum contamination factor defined in Rackham, Apai & Giampapa (2019). We run a retrieval, assuming chemical equilibrium, on the $R \simeq 100$ TIBERIUS spectrum. We apply heterogeneity priors $T_{\rm het} \sim \mathcal{U}[0.8T_*, 1.2T_*]$, and $\log g_{\rm het} \sim \mathcal{U}[3.0, 5.0]$ (see Table D2 for the full list of priors), where we use the literature value for the stellar effective temperature, T_* (Table 1). The definitions and priors of the other parameters explained above remain unchanged compared to the reference retrieval without stellar contamination (Section 4.1.2). We thus have 15 free parameters.

We retrieve a possible bright feature at $T_{\rm het} = 6389^{+388}_{-313}$ K, with a filling factor of $f_{\rm het} = 0.22^{+0.21}_{-0.13}$. The heterogeneity temperature is consistent with the retrieved photospheric temperature ($T_{\rm phot} = 6306^{+56}_{-60}$ K).

Moreover, the retrieved filling factor is relatively large considering the host stellar type; Rackham et al. (2019) derive $f_{\rm fac} = 0.01^{+0.02}_{-0.01}$ for F6V stars. It is also close to zero at the 1σ lower limit. That said, there are other studies that have retrieved similarly high filling factors for other F stars (see Section 5.3 for further discussion). This retrieval yields a reduced chi-square of $\chi^2_\nu = 1.21$, and the Bayesian evidence was $\ln \mathcal{Z} = 473.16 \pm 0.11$. The reference retrieval (II) ($\ln \mathcal{Z} = 473.35 \pm 0.11$) without stellar contamination was thus marginally preferred to 1.4σ. We find that though a combination of a planetary atmosphere as well as stellar heterogeneities could explain the transmission spectrum of TrES-4 b, the reference model (II) without stellar contamination is preferred due to the reduced model complexity. We note that even when including the contamination from photospheric heterogeneities, the atmospheric parameters do not change; the C/O and metallicity are consistent with those from the chemical equilibrium reference retrieval (II) (see Table 4).

4.2 PETITRADTRANS

We also perform atmospheric retrievals on the TrES-4 b transmission spectrum using the PETITRADTRANS[9] (v2.7.7; Mollière et al. 2019; Nasedkin, Mollière & Blain 2024) package, with a near identical setup to that used to analyse the spectrum of the first BOWIE-ALIGN target, WASP-15 b (Kirk et al. 2025). We use free chemistry, equilibrium chemistry, and hybrid chemistry retrieval setups assuming an H_2/He-dominated atmosphere, with absorption from trace gases computed from $R = 1000$ correlated-k opacity tables of the following atmospheric species: CH_4 (Yurchenko et al. 2017), H_2O and CO (Rothman et al. 2010), CO_2 (Yurchenko et al. 2020), H_2S (Azzam et al. 2016), SO_2 (Underwood et al. 2016), and HCN (Barber et al. 2014), plus collision-induced absorption from H_2–H_2 and H_2–He, and Rayleigh scattering from H_2 and He.

For each retrieval we model the atmosphere using 100 equal-log-spaced pressure layers from 10^{-6} to 10^2 bar and assume an isothermal pressure-temperature profile, with a wide uniform prior on the isothermal temperature from 500–3000 K. We use a wide uniform prior for the planetary reference radius (0.8–1.8 $R_{\rm J}$) and a Gaussian prior for gravity based on the published mass and radius (Sozzetti et al. 2015), both of which are defined at a fixed reference pressure of 10^{-3} bar. We determined the appropriate reference pressure by running a retrieval with a Gaussian prior on the radius, based on the mean transit depth across both detectors, and the reference pressure as a free parameter. We also include an opaque grey cloud deck, with a log-uniform prior on the cloud-top pressure, from 10^{-6} to 10^2 bar.

We run our retrievals on the TIBERIUS $R \simeq 100$ and $\simeq 400$ spectra, as well as the EUREKA! spectrum at $R \simeq 100$. Across these three retrieval tests, we retrieve consistent posteriors for radius, gravity, isotherm temperatures, and cloud-top pressure. We find a much lower limb temperature of $T = 1200 \pm 100$ K compared to the planet's equilibrium temperature of $T_{\rm eq} = 1800$ K. This is a commonly observed feature in 1D isothermal retrievals (see e.g. Welbanks et al. 2019; Kirk et al. 2025), which is postulated to arise from model choices such as the PT profile parametrization (Welbanks & Madhusudhan 2022) or unresolved limb asymmetries (MacDonald, Goyal & Lewis 2020). The former was ruled out in Section 4.1.2, wherein we tested a more complex PT profile. We also retrieve a relatively cloud-free atmosphere in all cases, placing a 3σ upper limit on the grey cloud-top pressure of 10^{-3} bar.

(IV) For the free chemistry retrievals, the abundance of each species is a free parameter, with a wide uniform prior in log mass fraction from -12 to -0.5. The free chemistry retrievals between all reductions and resolutions were consistent (see Table 3), revealing absorption due to H_2O, CO, and CO_2, with retrieved abundances of $\log X_{H_2O} = -3.29^{+0.57}_{-0.59}$, $\log X_{CO} = -3.71^{+0.85}_{-0.84}$, and $\log X_{CO_2} = -7.21^{+0.56}_{-0.51}$. We also see hints of H_2S abundances of 10 ppm, but with a less than 2σ confidence interval. We place 3σ upper limits on the SO_2 and HCN abundances of 4 and 2 ppm, respectively. The posterior distributions are presented in Appendix D (Fig. D5).

[9] https://petitradtrans.readthedocs.io/en/latest/

(V) For the equilibrium chemistry retrievals, the abundances of CH_4, H_2O, CO, CO_2, HCN, and H_2S are calculated at each pressure layer by interpolating a pre-computed equilibrium chemistry table with temperature, C/O ratio, and metallicity ([M/H]). We then use a wide uniform prior in both C/O ratio ($C/O \sim \mathcal{U}[0.1, 1.5]$) and $[M/H] \sim \mathcal{U}[-2, 3]$). In PETITRADTRANS, [M/H] is defined in terms of the solar values from Asplund et al. (2009), fixing the C/H value, with the O/H value then further modified by adjusting the C/O ratio (solar $C/O \sim 0.55$). We retrieve a subsolar C/O ratio posterior of $0.33^{+0.13}_{-0.11}$ and a subsolar metallicity of $-0.33^{+0.24}_{-0.19}$ for the TIBERIUS $R \simeq 400$ spectrum, and consistent values at $R \simeq 100$ with slightly reduced precision (see Table 4). We retrieve a similar but slightly higher C/O ratio of $0.43^{+0.17}_{-0.14}$ from the EUREKA! reduction. We show the resulting posterior distributions from these retrievals in Appendix D (Fig. D6).

(VI) For the hybrid chemistry retrievals, we use the same equilibrium chemistry method to compute the abundances of CO_2, CO, H_2O, CH_4, and HCN, but allow the sulphur species H_2S and SO_2 to vary freely, with broad uniform priors. This is to allow the abundances to be in disequilibrium due to photochemical production of SO_2 and destruction of H_2S, and to permit measurement of sulphur enrichment or depletion. As with our free chemistry retrieval, we see no evidence for SO_2, and hints of H_2S. As seen with WASP-15 b (Kirk et al. 2025), moving to a hybrid retrieval results in tighter constraints on the C/O ratio; we retrieve a posterior of $0.30^{+0.11}_{-0.09}$. We also retrieve a somewhat lower metallicity of -0.41 ± 0.18. The hybrid retrieval posteriors are presented in Appendix D (Fig. D7). The Bayesian evidence for the equilibrium and hybrid retrievals are identical for the $R \simeq 400$ retrievals, giving us no cause to favour one parametrization over the other, likely due to the absence of SO_2 in the spectrum and the poor constraints on H_2S. At $R \simeq 100$, the hybrid retrieval is slighly disfavoured, this difference is likely due the increased sensitivity to H_2S at higher resolutions.

4.3 Comparison of retrieval results

As standard in the community, we used two independent retrieval analyses to assess the atmospheric composition of TrES-4 b. This enables a robustness check of our inferences. For ease of comparison, we plot the retrieved spectra from POSEIDON and PETITRADTRANS together in Fig. D2. We saw good agreement between the retrieved chemical abundances from the free chemistry retrievals (Table 3). At face value, the chemical equilibrium retrievals of POSEIDON and PETITRADTRANS yielded slightly discrepant atmospheric metallicity outputs. This is due to the difference in metallicity prescriptions. In POSEIDON, the abundances of all species except carbon are set by the metallicity, which is scaled from the solar value from Asplund et al. (2009). The [C/H] is then set according to the [O/H] and C/O ratio. Conversely, PETITRADTRANS sets all but the oxygen abundance based on the metallicity, and the [O/H] is derived. To enable a direct comparison of the retrieved metallicities from both frameworks, we converted the POSEIDON output $[M/H]_{O/H}$ to $[M/H]_{C/H}$. Such a scaling is deemed reasonable for this work, since we only detect carbon and oxygen-bearing species, and the main observables (CO, CO_2, and H_2O) are not sensitive to the abundances of the other species.

Our PETITRADTRANS retrievals all used isothermal temperature profiles, previously thought to be sufficient for transmission spectroscopy. Recently, Schleich et al. (2024) showed that an isothermal prescription can lead to incorrect inferences and may be insufficient in the era of *JWST* exoplanet spectra. With our POSEIDON chemical equilibrium retrievals, we followed their recommendation of a two-point gradient temperature profile (sufficient except in the case of a temperature inversion, which we ruled out in Section 4.1.2). However, we found only a small difference between the upper and deeper temperatures, maximally 347 K which is comparable to the 1σ constraints on T_{deep}.

Ultimately, all retrievals yielded metallicities (i.e. $[M/H]_{C/H}$) consistent within 1σ. To summarize our collection of retrieval results, the metallicity of TrES-4 b is consistent with being subsolar, and certainly lower than the metallicity of the host star ($[Fe/H] = 0.28$). The retrieved subsolar C/O was also consistent across all retrievals, varying between $C/O = 0.30 - 0.42$.

5 DISCUSSION

5.1 The atmosphere of TrES-4 b

We have analysed the atmosphere of TrES-4 b with *JWST* NIRSpec *G395H* data, two independent data reduction pipelines (Section 3), together with two retrieval packages (Section 4). Both retrieval frameworks yield consistent results, and our tests in Section 4 present clear evidence for gaseous H_2O and CO_2 in the atmosphere of TrES-4 b. Additionally, there is strong evidence for the presence of CO. Considering the Bayesian evidence, the favoured POSEIDON atmospheric model is that inferred by the chemical equilibrium retrieval (II). In the PETITRADTRANS retrievals, the chemical equilibrium atmosphere was preferred by the $R \simeq 100$ spectrum, while at $R \simeq 400$ the equilibrium and hybrid retrievals yielded equal evidence. Neither POSEIDON nor PETITRADTRANS found evidence of clouds or hazes. We also tested for sulphur-bearing molecules H_2S and SO_2, finding weak evidence of H_2S, a non-detection of SO_2. This is consistent with the lower retrieved metallicity (see e.g. Polman et al. 2023; Tsai et al. 2023).

TrES-4 b is the second planet from the BOWIE-ALIGN sample; the first, WASP-15 b, was published in Kirk et al. (2025). WASP-15 b is a misaligned hot Jupiter, and was shown to host a supersolar metallicity atmosphere with C/O consistent with the solar value. In Fig. 8, we show the NIRSpec *G395H* TIBERIUS $R \simeq 100$ transmission spectrum of WASP-15 b from Kirk et al. (2025), against the equivalent spectrum of TrES-4 b from this work (Section 3.1.2, shown in Fig. 3). For ease of comparison, we have scaled the transit depth by that for one atmospheric scale height (139 and 200 ppm for WASP-15 b and TrES-4 b, respectively), and then normalized to the 4.33 μm bin (the peak of the CO_2 bandhead). The comparatively reduced metallicity for TrES-4 b is apparent, considering the amplitude of the CO_2 feature relative to the blueward H_2O slope. Further, Kirk et al. (2025) presented evidence of sulphur chemistry, through absorption features at 4.0 and 4.9 μm. The SO_2-induced feature seen at 4.0 μm for WASP-15 b is clearly absent in our target TrES-4 b. Our retrievals yielded little evidence towards sulphur-bearing species; allowing the abundances of SO_2 and H_2S to vary freely in the PETITRADTRANS hybrid retrievals (to allow for depletion or enrichment) did not result in increased model evidence. Finally, the peak at 4.9 μm seen in WASP-15 b's transmission spectrum is apparently inverted in that of TrES-4 b. This inverted feature is consistent across reductions, but has a weaker structure at higher resolution (Fig. C2). Kirk et al. (2025) explored the possibility of OCS as the contributing absorber for the feature in WASP-15 b. In the case of TrES-4 b, the two outlier points just shy of 4.9 μm are still consistent with the best-fitting chemical equilibrium model within 3σ (Fig. 6). Given the lack of other sulphur species, we choose not to investigate the presence of OCS further for TrES-4 b.

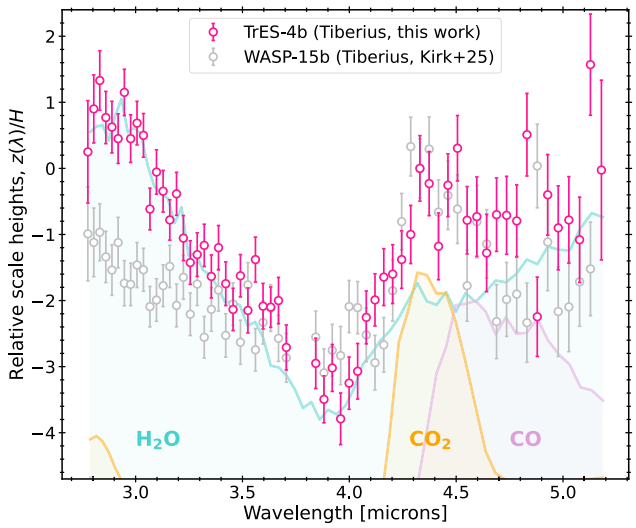

Figure 8. The transmission spectrum of TrES-4 b from this work (TIBERIUS $R \simeq 100$) compared to the equivalent for WASP-15 b, published in Kirk et al. (2025, grey). The spectral contributions from the detected gas opacities are overplotted for TrES-4 b.

5.2 Implications for formation scenarios

Inferring formation histories for a single planet such as TrES-4 b is challenging, as results are typically degenerate and generally dependent on unknown factors such as the chemical composition of the disc (see e.g. Penzlin et al. 2024). Nevertheless, the substellar metallicity and subsolar C/O ratio place TrES-4 b in an interesting region of parameter space, which we consider here. In the typical picture of giant planet formation, gas accretion of the carbon-rich gases leads to low metallicities but high C/O, while strong solid accretion increases the metallicity through adding oxygen-rich ices and refractories, leading to low C/O and high metallicity. In the TrES-4 b retrievals in Section 4, we retrieve a planet atmosphere metallicity between $\log Z/Z_\odot = -0.41$ to -0.04, which differs from the higher stellar metallicity at the 0.9σ to 3.4σ level. Meanwhile, we find a C/O ratio between $0.30 - 0.42$ across all independent analyses. Considering the host star metallicity, a stellar C/O ratio in the range 0.4–0.5 is likely, based on observations of similar stars (Brewer & Fischer 2016). This would be compatible with the observed C/O ratio of TrES-4 b at $0.5 - 1.5\sigma$. We therefore cannot completely rule out the possibility that the planet's composition is close to the stellar composition, in which case no special circumstances are needed to explain it.

Conversely, a substellar metallicity combined with substellar C/O is challenging to explain within the standard picture of planet formation, assuming typical disc compositions. A particular challenge is that the gas composition in discs is usually thought to have superstellar C/O ratios, but substellar metallicities due to the preferential condensation of oxygen-rich ices (e.g. Öberg et al. 2011; Bosman et al. 2021; Bergin et al. 2024). As a result, planets with substellar metallicities are generally expected to have superstellar C/O ratios and vice versa (e.g. Madhusudhan, Amin & Kennedy 2014), because adding more solids decreases the C/O at the cost of higher metallicity. This makes the preferred composition of TrES-4 b somewhat unexpected. To explain the preferred composition, we must either invoke conditions in the disc where the gas has substellar C/O and metallicity (reversing the trend) or invoke transport processes.

If the planet did form in a location where the gas composition has low C/O, then this points to one of two possible scenarios. First, the planet could have accreted the majority of its observable atmosphere inside the water ice line but outside the 'soot line', inside which carbon-rich grains are destroyed (Chachan, Lee & Knutson 2021; Bergin et al. 2023). We depict this scenario on the left side of Fig. 9. In this region, the solids can have high C/O ratios if carbon grains are sufficiently abundant. Here, most oxygen-rich species apart from silicates are in the gas phase, resulting in lower C/O ratios in the gas. Observations of the interstellar medium suggest that roughly half of the interstellar carbon is contained in refractory grains (Mishra & Li 2015), which is sufficient to explain the composition of TrES-4 b. This explanation would point to the planet forming in the inner disc, where temperatures are $200 - 500$ K. Alternatively, chemical kinetics models of protoplanetary discs show that after ~ 5 Myr the gas phase C/O ratio can become low in the inner $\sim 1 - 10$ au region of a disc, as CO and CO_2 are destroyed (Eistrup et al. 2018). Hence, the planet could have formed further out, but later in the stellar lifetime. A slight wrinkle in the second scenario is that the models producing low C/O ratio gas in the disc only include chemical kinetics, and accretion/pebble migration may overwhelm chemical processes that require such long time-scales to operate (e.g. Booth & Ilee 2019).

Alternatively, the stalling of solid drift could help explain the low metallicity and C/O ratio (right side, Fig. 9). If the disc has a pebble trap in the outer system where all volatile species are frozen out in grains (including H_2O, CO_2, CO, and CH_4), the disc inside the trap is metal depleted. Under such conditions, the planet would need to accrete fewer solids to acquire the observed C/O ratio and thus could have a lower metallicity. This mechanism is sensitive to the location and onset of the trapping feature in the disc (Mah, Savvidou & Bitsch 2024). Enhancing transport processes would not help explain

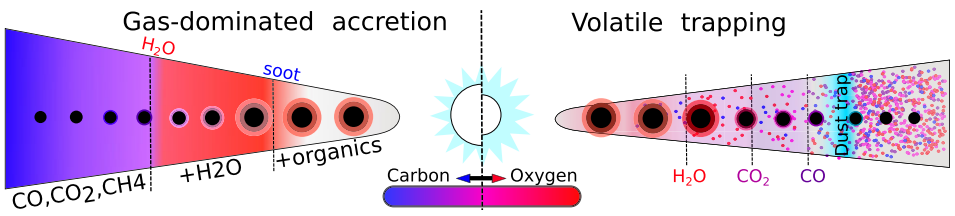

Figure 9. Schematic of two distinct protoplanetary disc environment scenarios, in which TrES-4 b may have formed, relative to the host star in the centre. Left: the gas-dominated accretion scenario whereby the planet forms between the water ice line and the soot line (labelled dashed lines). Right: the volatile trapping scenario, whereby the solids are trapped in the outer disc, rendering the inner disc lower in metallicity. The relative abundances of carbon and oxygen are expressed by the colour bar.

the proposed composition. For example, although efficient pebble migration can produce discs with low C/O ratios in the inner region, they do this by depositing large amounts of water, resulting in high metallicities (Booth et al. 2017; Schneider & Bitsch 2021). Similarly, although planet migration can modify the planet's composition, the effects are not large enough to change the broad conclusion that planets with low metallicity would ordinarily be expected to have high C/O ratios (Madhusudhan et al. 2014; Penzlin et al. 2024). For all these processes. the amount of carbon in solid 'soot' is crucial. Through depleting carbon from molecules that can sublimate at very low temperatures (e.g. CO) to long-chain complex molecules that require temperatures of several hundred Kelvin to sublimate, the distribution of carbon in solids within the disc can shift, affecting the C/O ratio of the close-in final planet. If carbon was predominately in solids, late accretion would significantly increase C/O. If C/O was predominately in gases and volatiles, gas-dominated accretion lead to higher C/O ratios and vice versa. With all the planets in the BOWIE-ALIGN programme (Kirk et al. 2024a), we will have a sample to investigate the carbon evolution and distribution during planet formation.

A strong constraint on the atmospheric sulphur abundance would help distinguish these scenarios because most sulphur is likely in refractory form in both protoplanetary discs and the Solar system (Kama et al. 2019; Chachan et al. 2021; Turrini et al. 2021; Crossfield 2023). However, our retrievals (Section 4) find no traces of sulphur species, except for hints of H_2S, for which the upper limit is too high to provide a meaningful constraint on the formation history.

Although the formation history of neither TrES-4 b nor WASP-15 b is certain, it is clear that there should be differences in their histories. WASP-15 b has a supersolar metallicity, a C/O to ratio close to solar and a presumed high sulphur abundance. In contrast, TrES-4 b has a substellar metallicity and C/O ratio. Whether or not these differences are due to their aligned vs misaligned natures remains to be seen, a point that may be elucidated once the full BOWIE-ALIGN sample (Kirk et al. 2024a) has been analysed.

5.3 Stellar contamination

We have assessed the presence of stellar contamination in both the light curves and the transmission spectrum of TrES-4 b. In Section 3, we saw possible correlated noise in the white light curves, more prevalent in that of NRS1 than NRS2. This was seen in the Allan variance curves (Fig. 2), and 'bumps' were seen in the residuals of the detrended light curves (Fig. 1). Though such features may be induced by occultations of stellar surface heterogeneities, within the transit chord, we saw no evidence for spot-crossing events when fitting the light curve. At the transmission spectrum level, in Section 4.1 we used the in-built functionality of POSEIDON to simultaneously fit for a stellar contamination factor when performing atmospheric retrievals (Rackham et al. 2019). This test was limited to the scenario whereby the observed transmission spectrum is a combination of atmospheric chemical signal as well as contribution from unocculted stellar surface heterogeneities. While neither stellar contamination models were statistically preferred, it would be prudent to examine the plausibility of such a scenario.

Though TLSE contribution from F-type stars is predicted to be marginal (Rackham et al. 2019), there has been a range of activity signals reported in the literature. For the F5-type star WASP-79, Rathcke et al. (2021) saw evidence for faculae in their *Hubble Space Telescope*/ *Space Telescope Imaging Spectrograph* (*HST*/*STIS*) transmission spectrum, $\Delta T \sim 500$ K hotter than the photosphere, with a filling factor of 15 per cent, despite observed low photometric variability (Sotzen et al. 2019).

Other examples include stars WASP-121 (F6-type star) and WASP-103 (F8), for which activity signals have varied between data and epochs. For WASP-121, a faculae filling factor $f_{\rm fac} = 0.08$ was derived from radial velocity jitter (Delrez et al. 2016), while the *HST*/STIS transmission spectrum showed no signs of contamination (Evans et al. 2018). WASP-103 displays little photometric variability (Gillon et al. 2014), but has a relatively high chromospheric emission $\log R'_{\rm HK} = -4.47$, which was noted to be higher than expected given the stellar age (Staab et al. 2017). Kirk et al. (2021) then measured $f_{\rm fac} = 0.22^{+0.12}_{-0.09}$ with a temperature contrast $\Delta T = 250$ K from their optical + IR transmission spectrum of WASP-103 b.

TrES-4 is thought to be a reasonably quiet, main-sequence F6V star. Narita et al. (2010) measured an RV variability of 20 m s^{-1}, but this was only tentatively postulated to be due to stellar jitter (see e.g. Wright 2005). TrES-4 has been observed to display little chromospheric emission, with $\log R'_{\rm HK} = -5.11 \pm 0.15$ (Sozzetti et al. 2009), and there have been no reports of activity-induced photometric variability (e.g. Knutson et al. 2009; Maciejewski et al. 2023).

Though the temperature $T_{\rm het}$ we retrieved (Section 4.1.3; see Table D3) for TrES-4 is consistent with theory (e.g. Gondoin 2008; Rackham et al. 2019), in comparison, $f_{\rm het}$ is uncharacteristically high, whilst also being poorly constrained, especially considering our target TrES-4 has a rotation rate around half that of WASP-79. Indeed, assuming Sun-like heterogeneity location distributions, Rackham et al. (2019) predicted a filling factor of $f_{\rm fac} = 0.01^{+0.02}_{-0.02}$ for a F6V star. On the other hand, though in theory there exists a degeneracy between the temperature of the heterogeneity and its filling factor (Pinhas et al. 2018), the stellar contamination is easier to disentangle at the bluer wavelengths of the *HST*/STIS observations of Rathcke et al. (2021) (shorter than ~ 1 μm).

As explained in Section 4.1, since there is no additional evidence to support the models with stellar contamination included, the preferred models are those without. Further, the key parameters presented in this work, namely the C/O and metallicity of the atmosphere of TrES-4 b, are not dependent on the choice of model. From these *JWST* NIRSpec *G*395*H* observations, we thus conclude the atmospheric measurements of TrES-4 b regardless of the presence of stellar surface heterogeneities. Future observations, for example at bluer wavelengths, will be required to constrain the level of stellar heterogeneity.

6 CONCLUSIONS

In this work, we have presented the *JWST*/NIRSpec *G*395*H* 2.8 – 5.2 μm transmission spectrum of TrES-4 b. Having implemented two independent reduction routines, we recover consistent transmission spectra. We use two independent atmospheric retrieval analyses to characterize the atmosphere of TrES-4 b, presenting clear evidence for detections of H_2O, CO_2, and CO. We do not find sufficient evidence to suggest that cloud opacity impacts the transmission spectrum within the *G*395*H* bandpass, nor the need to account for stellar contamination, or an offset between detectors. Though we do not detect other species at significance, we are able to place meaningful constraints on their abundances. We place upper limits on the abundances of sulphur-bearing molecules H_2S and SO_2; while we see possible hints of H_2S, we rule out large abundances of SO_2 (with a 2σ upper limit of $\log X_{SO_2} < -6.97$).

Under chemical equilibrium, we retrieve consistent C/O across the reductions and retrieval setups. The C/O is subsolar and likely

lower than that of the host star, TrES-4. The retrieved metallicity is also likely to be substellar, within the range $0.4 - 0.7\times$ solar. The expectation from theoretical models of formation is either high C/O and low metallicity, or vice-versa. Both values being substellar presents an interesting case with regard to formation scenarios. We proposed two scenarios by which substellar C/O and metallicity could be induced: the accretion of low C/O gas, or a combination of low-metallicity gas and low C/O solids accretion. These scenarios could be distinguished with a tighter constraint on the sulphur abundance.

TrES-4 b is the second target in the BOWIE-ALIGN programme, in which there are four aligned and four misaligned targets. We await the full sample before more significant links between atmospheric C/O, metallicity, and formation can be inferred.

ACKNOWLEDGEMENTS

This work is based on observations made with the NASA/ESA/CSA *JWST*. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127 for *JWST*. These observations are associated with program JWST-GO-3838. Support for program JWST-GO-3838 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127.

We thank the reviewer for their helpful comments, which helped to improve the quality of the manuscript. RAB and AP thank the Royal Society for their support in the form of a University Research Fellowship. JK acknowledges financial support from Imperial College London through an Imperial College Research Fellowship grant. HRW was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant no. EP/Y006313/1]. PJW acknowledges support from the UK Science and Technology Facilities Council (STFC) through consolidated grant ST/X001121/1. NJM, DES, and MZ acknowledge support from a UK Reseach and Innovation (UKRI) Future Leaders Fellowship (grant MR/T040866/1), a Science and Technology Facilities Funding Council Nucleus Award (grant ST/T000082/1), and the Leverhulme Trust through a research project grant (RPG-2020-82).

DATA AVAILABILITY

The data products associated with this manuscript can be found online at Zenodo at https://doi.org/10.5281/zenodo.15097695. We describe the data products resulting from our survey in Kirk et al. (2024a).

REFERENCES

Ahrer E.-M. et al., 2023, Nature, 614, 653
Albrecht S. et al., 2012, ApJ, 757, 18
Alderson L. et al., 2023, Nature, 614, 664
Allan D. W., 1966, IEEE Proc., 54, 221
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Azzam A. A. A., Tennyson J., Yurchenko S. N., Naumenko O. V., 2016, MNRAS, 460, 4063
Barber R. J., Strange J. K., Hill C., Polyansky O. L., Mellau G. Ch., Yurchenko S. N., Tennyson J., 2014, MNRAS, 437, 1828
Bell T. J. et al., 2022, J. Open Source Softw., 7, 4503
Benneke B., Seager S., 2013, ApJ, 778, 153
Bergin E. A., Kempton E., Hirschmann M., Bastelberger S. T., Teal D. J., Blake G. A., Ciesla F., Li J., 2023, ApJ, 949, L17
Bergin E. A., Booth R. A., Colmenares M. J., Ilee J. D., 2024, ApJ, 969, L21
Beyer A., White R., 2024, American Astronomical Society Meeting Abstracts, 243, 205.08
Booth R. A., Ilee J. D., 2019, MNRAS, 487, 3998
Booth R. A., Clarke C. J., Madhusudhan N., Ilee J. D., 2017, MNRAS, 469, 3994
Bosman A. D. et al., 2021, ApJS, 257, 7
Brewer J. M., Fischer D. A., 2016, ApJ, 831, 20
Buchner J. et al., 2014, A&A, 564, A125
Chachan Y., Lee E. J., Knutson H. A., 2021, ApJ, 919, 63
Crossfield I. J. M., 2023, ApJ, 952, L18
Dawson R. I., 2014, ApJ, 790, L31
Delrez L. et al., 2016, MNRAS, 458, 4025
Eistrup C., Walsh C., van Dishoeck E. F., 2018, A&A, 613, A14
Esparza-Borges E. et al., 2023, ApJ, 955, L19
Evans T. M. et al., 2018, AJ, 156, 283
Feinstein A. D. et al., 2023, Nature, 614, 670
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Garland R., Irwin P. G. J., 2019, preprint (arXiv:1903.03997)
Gillon M. et al., 2014, A&A, 562, L3
Gondoin P., 2008, A&A, 478, 883
Grant D., Wakeford H. R., 2024, J. Open Source Softw., 9, 6816
Horne K., 1986, PASP, 98, 609
Jakobsen P. et al., 2022, A&A, 661, A80
JWST Transiting Exoplanet Community Early Release Science Team, 2023, Nature, 614, 649
Kama M., Shorttle O., Jermyn A. S., Folsom C. P., Furuya K., Bergin E. A., Walsh C., Keller L., 2019, ApJ, 885, 114
Karman T. et al., 2019, Icarus, 328, 160
Kirk J., Wheatley P. J., Louden T., Doyle A. P., Skillen I., McCormac J., Irwin P. G. J., Karjalainen R., 2017, MNRAS, 468, 3907
Kirk J. et al., 2021, AJ, 162, 34
Kirk J. et al., 2024a, RAS Tech. Instrum., 3, 691
Kirk J. et al., 2024b, AJ, 167, 90
Kirk J. et al., 2025, MNRAS, 537, 3027
Knutson H. A., Charbonneau D., Burrows A., O'Donovan F. T., Mandushev G., 2009, ApJ, 691, 866
Kokori A. et al., 2023, ApJS, 265, 4
Kraft R. P., 1967, ApJ, 150, 551
Kreidberg L., 2015, PASP, 127, 1161
Law C. J. et al., 2021, ApJS, 257, 4
Li G., Gordon I. E., Rothman L. S., Tan Y., Hu S.-M., Kassi S., Campargue A., Medvedev E. S., 2015, ApJS, 216, 15
Lin Y., Ogilvie G. I., 2017, MNRAS, 468, 1387
MacDonald R. J., 2023, J. Open Source Softw., 8, 4873
MacDonald R. J., Lewis N. K., 2022, ApJ, 929, 20
MacDonald R. J., Madhusudhan N., 2017, MNRAS, 469, 1979
MacDonald R. J., Goyal J. M., Lewis N. K., 2020, ApJ, 893, L43
Maciejewski G. et al., 2023, Acta Astron., 73, 57
Madhusudhan N., 2012, ApJ, 758, 36
Madhusudhan N., Seager S., 2009, ApJ, 707, 24
Madhusudhan N., Amin M. A., Kennedy G. M., 2014, ApJ, 794, L12
Magic Z., Chiavassa A., Collet R., Asplund M., 2015, A&A, 573, A90
Mah J., Savvidou S., Bitsch B., 2024, A&A, 686, L17
Mandushev G. et al., 2007, ApJ, 667, L195
Mishra A., Li A., 2015, ApJ, 809, 120
Mollière P., Wardenier J. P., van Boekel R., Henning T., Molaverdikhani K., Snellen I. A. G., 2019, A&A, 627, A67
Moran S. E. et al., 2023, ApJ, 948, L11
Morbidelli A. et al., 2016, Icarus, 267, 368
Moré J. J., 1978, in Watson G. A., ed., Numerical Analysis. Springer, Berlin, Heidelberg, p. 105
Mullens E., Lewis N. K., MacDonald R. J., 2024, ApJ, 977, 105
Narita N., Sato B., Hirano T., Winn J. N., Aoki W., Tamura M., 2010, PASJ, 62, 653
Nasedkin E., Mollière P., Blain D., 2024, J. Open Source Softw., 9, 5875

Öberg K. I., Murray-Clay R., Bergin E. A., 2011, ApJ, 743, L16
Owen J. E., 2020, MNRAS, 495, 3160
Palla F., Stahler S. W., 1993, ApJ, 418, 414
Penzlin A. B. T. et al., 2024, MNRAS, 535, 171
Pinhas A., Rackham B. V., Madhusudhan N., Apai D., 2018, MNRAS, 480, 5314
Polman J., Waters L. B. F. M., Min M., Miguel Y., Khorshid N., 2023, Nature, 617, 483
Polyansky O. L., Kyuberis A. A., Zobov N. F., Tennyson J., Yurchenko S. N., Lodi L., 2018, MNRAS, 480, 2597
Rackham B. V., Apai D., Giampapa M. S., 2018, ApJ, 853, 122
Rackham B. V., Apai D., Giampapa M. S., 2019, AJ, 157, 96
Rathcke A. D. et al., 2021, AJ, 162, 138
Rothman L. S. et al., 2010, J. Quant. Spectr. Radiat. Transf., 111, 2139
Rustamkulov Z. et al., 2023, Nature, 614, 659
Schleich S., Saikia S. B., Changeat Q., Güdel M., Voigt A., Waldmann I., 2024, A&A, 690, A336
Schneider A. D., Bitsch B., 2021, A&A, 654, A72
Sotzen K. S. et al., 2019, AJ, 159, 5
Sozzetti A. et al., 2009, ApJ, 691, 1145
Sozzetti A. et al., 2015, A&A, 575, L15
Spalding C., Winn J. N., 2022, ApJ, 927, 22
Staab D., Haswell C. A., Smith G. D., Fossati L., Barnes J. R., Busuttil R., Jenkins J. S., 2017, MNRAS, 466, 738
Stock J. W., Kitzmann D., Patzer A. B. C., Sedlmayr E., 2018, MNRAS, 479, 865
Stock J. W., Kitzmann D., Patzer A. B. C., 2022, MNRAS, 517, 4070
Taylor J. et al., 2023, MNRAS, 524, 817
Trotta R., 2008, Contemp. Phys., 49, 71
Tsai S.-M. et al., 2023, Nature, 617, 483
Turrini D. et al., 2021, ApJ, 909, 40
Underwood D. S., Tennyson J., Yurchenko S. N., Huang X., Schwenke D. W., Lee T. J., Clausen S., Fateev A., 2016, MNRAS, 459, 3890
Welbanks L., Madhusudhan N., 2021, ApJ, 913, 114
Welbanks L., Madhusudhan N., 2022, ApJ, 933, 79
Welbanks L., Madhusudhan N., Allard N. F., Hubeny I., Spiegelman F., Leininger T., 2019, ApJ, 887, L20
Winn J. N., Fabrycky D., Albrecht S., Johnson J. A., 2010, ApJ, 718, L145
Wright J. T., 2005, PASP, 117, 657
Yurchenko S. N., Amundsen D. S., Tennyson J., Waldmann I. P., 2017, A&A, 605, A95
Yurchenko S. N., Mellor T. M., Freedman R. S., Tennyson J., 2020, MNRAS, 496, 5282
Yurchenko S. N., Owens A., Kefala K., Tennyson J., 2024, MNRAS, 528, 3719
Zinnecker H., Yorke H. W., 2007, ARA&A, 45, 481

APPENDIX A: SECOND TIBERIUS REDUCTION

As an additional independent check, we ran a second data reduction using the TIBERIUS package. This reduction was performed using the same `extraction_input.txt` file, JWST pipeline version (1.8.2) and calibration reference files as used in the BOWIE-ALIGN WASP-15b study (Kirk et al. 2025). We provide the list of used calibration reference files in our associated Zenodo repository. We performed this reduction to ensure one homogeneous reduction throughout the BOWIE-ALIGN programme.

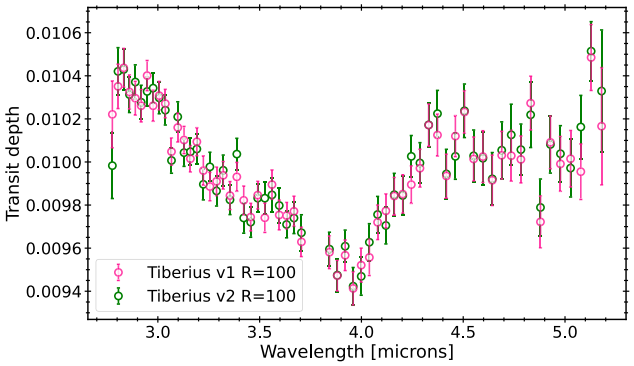

Figure A1. Comparison of the $R \simeq 100$ transmission spectra from the two independent implementations of TIBERIUS. The v1 spectrum is the primary transmission spectrum analysed in this work (Section 3.1.2).

The key differences to the principal TIBERIUS reduction presented in Section 3 are the choice of aperture width (6 pixels in the principal versus 8 pixels here) and the choice of JWST pipeline version (v1.13.4 versus 1.8.2). Given the good agreement between both TIBERIUS reductions (Fig. A1), this demonstrates that TrES-4 b's spectrum is insensitive to these differences in pipeline versions and calibration reference files. The light-curve model was the same for this second reduction as for the principal reduction. However, for this second reduction, all light curves were fitted with Levenberg–Marquadt optimization, rather than MCMC.

Within this second reduction, we also explored the implications of choosing to fit the transit light curves with fixed quadratic limb-darkening coefficients. First, we refitted the data with a four parameter, non-linear law with coefficients fixed using the same stellar parameters, 3D STAGGER models and EXOTIC-LD software (Magic et al. 2015; Grant & Wakeford 2024) as done in the principal TIBERIUS reduction. This led to an almost identical spectrum as the fixed quadratic law (median difference of 3 ppm). In a further test, we refit the light curves with both quadratic coefficients as free parameters and sampled the parameter space with MCMC. In this case, each wavelength bin in the transmission spectrum agreed to within 1σ of the fixed quadratic spectrum but with spectral uncertainties $1.53\times$ larger. Given these tests, we are confident that our transmission spectra are not biased by our choice of quadratic limb darkening with fixed coefficients.

APPENDIX B: ASTROPHYSICAL PARAMETERS FROM TIBERIUS LIGHT CURVES

In this appendix, we provide the posterior distributions for the individual TIBERIUS white light-curve fits (Fig. B1), described in Section 3.1.2. The mean and 1σ values are provided in Table 2. We present the $t_{\rm mid}$ as an offset in seconds from the mid-transit time predicted by the propagated literature value (Table 2). We note that there were two other free parameters in each of these fits: coefficients for the linear-in-time systematics model.

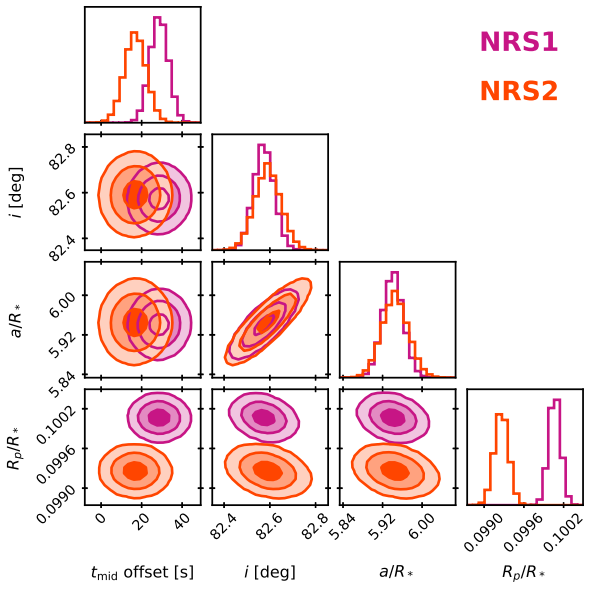

Figure B1. Posterior distribution (1σ, 2σ, and 3σ contours) of astrophysical parameters fitted in the TIBERIUS white light curves (each detector was fitted independently).

APPENDIX C: SPECTROSCOPIC LIGHT-CURVE ANALYSIS

In Section 3, we presented the $R \simeq 100$ transmission spectrum of TrES-4 b. The associated Allan variance plots for each of the detrended spectroscopic light curves are shown in Fig. C1; we found that a linear-in-time polynomial was sufficient to detrend these light curves. We also presented results for the $R \simeq 400$ transmission spectra, which are shown in Fig. C2 for each of the three reductions.

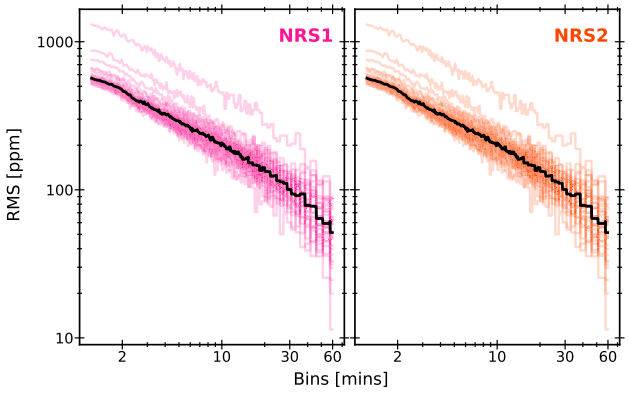

Figure C1. The Allan variance of the TIBERIUS $R \simeq 100$, detrended spectroscopic light curves. The medians across all spectroscopic bins are outlined in black.

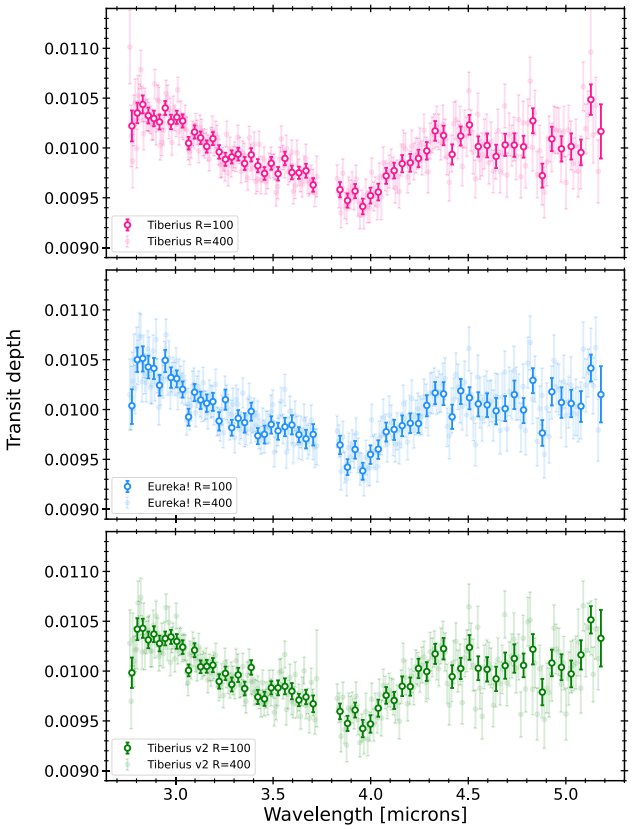

Figure C2. The $R \simeq 100$ and $\simeq 400$ transmission spectra from the three independent reductions using top panel: TIBERIUS, middle: EUREKA!, and bottom: TIBERIUS second reduction as detailed in Appendix A.

APPENDIX D: RETRIEVAL RESULTS

Here, we provide Supporting Information for the retrieval tests outlined in Section 4. A summary of the various priors used for the parameters of the retrievals are given in Tables D1 and D2. The resulting posterior distributions from the free and equilibrium chemistry POSEIDON retrievals on the TIBERIUS $R \simeq 100$ spectrum are shown in Figs D3 and D4, respectively. We provide the posterior distributions and retrieved spectra from the PETITRADTRANS retrievals on the same transmission spectrum in Figs D5–D7, and plot the best-fitting POSEIDON and PETITRADTRANS spectra together in Fig. D2.

In Section 4.1.2, we decided to use a gradient PT profile for our POSEIDON equilibrium chemistry retrievals. Here, we present in Fig. D1 the best-fitting PT profile. We compare the complex PT profile of Madhusudhan & Seager (2009) to our simplified gradient PT profile. We also plot the contribution function, using the functionality in POSEIDON introduced by Mullens et al. (2024). It can be seen that, in the region where the contribution function peaks (between $P = 10^{-2}$ to 10^{-3} bar) we are probing similar temperatures. Given that the PT

Table D1. Priors used for the retrievals in Section 4. Wide uniform priors are used for all mixing ratios: $\log \mathbf{X}_i \sim \mathcal{U}[-12, -1]$.

Retrieval	No.	$R_{\mathrm{p,ref}}$ [R_J]	$\log g$ [cgs]	T [K]†	P_{cloud} [bar]	$\log a$	γ	C/O	$Z/Z_\odot$	Offset [ppm]
POSEIDON										
Free chemistry	(I)	$\mathcal{U}[1.4, 2.11]$	$\mathcal{N}(2.60, 0.15)$	$\mathcal{U}[600, 2000]$	–	‡	‡	–	–	$\mathcal{N}(0, 500)$
Equilibrium chemistry	(II)	$\mathcal{U}[1.4, 2.11]$	$\mathcal{N}(2.60, 0.15)$	$\mathcal{U}[600, 2000]$	$\mathcal{U}[600, 2000]$	$\mathcal{U}[10^{-7}, 10^2]$	‡	‡	$\mathcal{U}[0.2, 2.0]$	$\mathcal{U}[10^{-1}, 10^4]$	$\mathcal{N}(0, 500)$
PETITRADTRANS										
Free chemistry	(IV)	$\mathcal{U}[0.8, 2.2]$	$\mathcal{N}(2.56, 0.05)$	$\mathcal{U}[850, 2850]$	–	$\mathcal{U}[10^{-6}, 10^2]$	–	–	–	–	–
Equilibrium chemistry	(V)	$\mathcal{U}[0.8, 2.2]$	$\mathcal{N}(2.56, 0.05)$	$\mathcal{U}[500, 3000]$	–	$\mathcal{U}[10^{-6}, 10^2]$	–	–	$\mathcal{U}[0.1, 1.5]$	$\mathcal{U}[10^{-2}, 10^3]$	–	–
Hybrid chemistry	(VI)	$\mathcal{U}[0.8, 2.2]$	$\mathcal{N}(2.56, 0.05)$	$\mathcal{U}[500, 3000]$	–	$\mathcal{U}[10^{-6}, 10^2]$	–	–	$\mathcal{U}[0.1, 1.5]$	$\mathcal{U}[10^{-2}, 10^3]$	–	–

Notes. †If one value given, an isothermal temperature profile was used. If two values given, a gradient profile was used, and the values are ordered as T_{high} and T_{deep}.
‡No user-specified prior, in which case POSEIDON applies a wide uniform prior.

profile of Madhusudhan & Seager (2009) is relatively isothermal throughout the region we are probing, and throughout the pressure ranges modelled, this justifies that a simplified PT profile is adequate to explain the observations. Furthermore, the location of the peak of the contribution function being higher than any of the cloud locations retrieved (see Section 4.1), aligns with the non-detections of any clouds within our observations.

Finally, in Section 4.1.3 we also explored the possibility of stellar heterogeneities, including a parametrization in our POSEIDON retrievals. The results from the stellar contamination test in Section 4.1.3 are provided in Table D3; the results from the other retrievals are given in Tables 3 and 4 in Section 4.

Table D2. Priors used for the stellar activity retrieval in Section 4.1.3. Other parameters follow the priors given in Table D1.

Input spectrum	No.	$f_{\rm het}$	$T_{\rm het}$	$\log_{g_{\rm het}}$	$T_{\rm phot}$	$\log_{g_{\rm phot}}$
POSEIDON						
TIBERIUS $R = 100$	(III)	$\mathcal{U}[0, 1]$	$\mathcal{U}[0.8T_*, 1.2T_*]$	$\mathcal{U}[3, 5]$	$\mathcal{N}(T_*, \sigma_{T_*})$	$\mathcal{N}(\log g_*, \sigma_{\log_{g_*}})$

Table D3. Results from the stellar contamination retrieval in Section 4.1.3. Detailed of priors provided in Table D2.

Input spectrum	No.	$f_{\rm het}$	$T_{\rm het}$	$\log_{g_{\rm het}}$	$T_{\rm phot}$	$\log_{g_{\rm phot}}$
POSEIDON						
TIBERIUS $R = 100$	(III)	$0.22^{+0.21}_{-0.13}$	6389^{+388}_{-313}	$3.97^{+0.63}_{-0.58}$	6306^{+56}_{-60}	4.09 ± 0.03

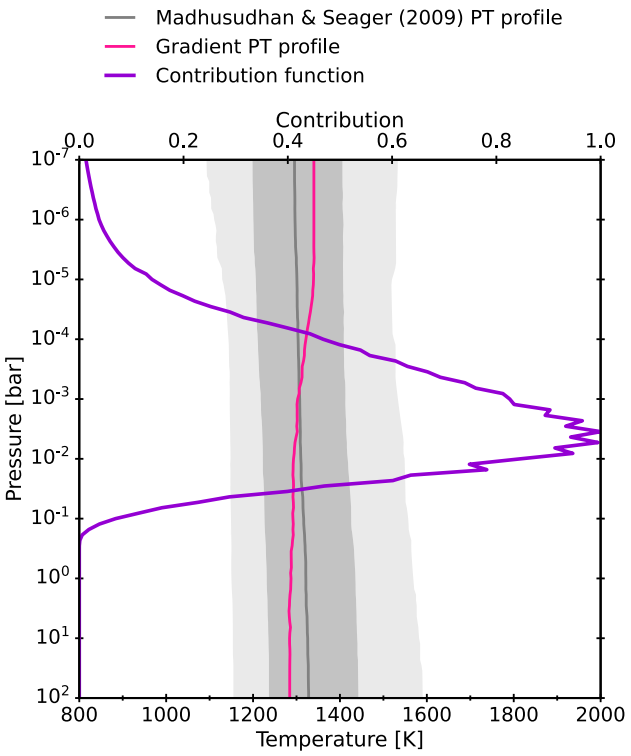

Figure D1. The median retrieved PT profile ($1/2\sigma$ regions shaded) of the chemical equilibrium reference retrieval on the TIBERIUS $R \simeq 100$ transmission spectrum, having implemented the more complex PT profile of Madhusudhan & Seager (2009). Overplotted is the median retrieved gradient PT profile on the same spectrum (preferred), and corresponding to the top x-axis is the photometric contribution function across the full NIRSpec *G*395*H* bandpass.

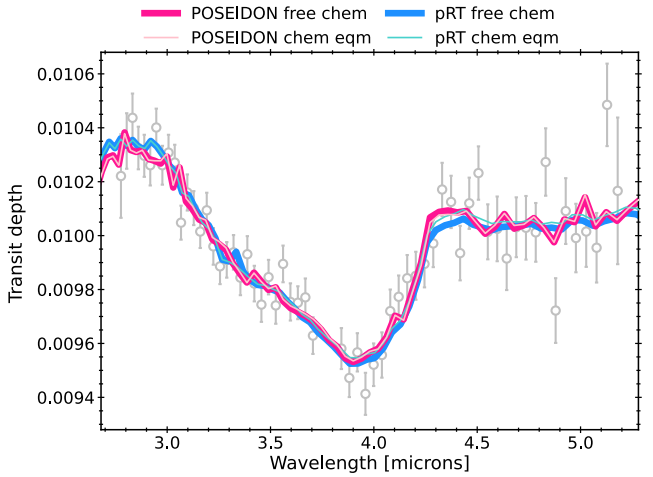

Figure D2. The best-fitting retrieved spectra from the POSEIDON free and equilibrium chemistry retrievals on the TIBERIUS $R \simeq 100$ transmission spectrum (data points in grey). Also plotted are the best-fitting spectra from PETITRADTRANS on the same spectrum. The forward models have been binned to $R = 100$.

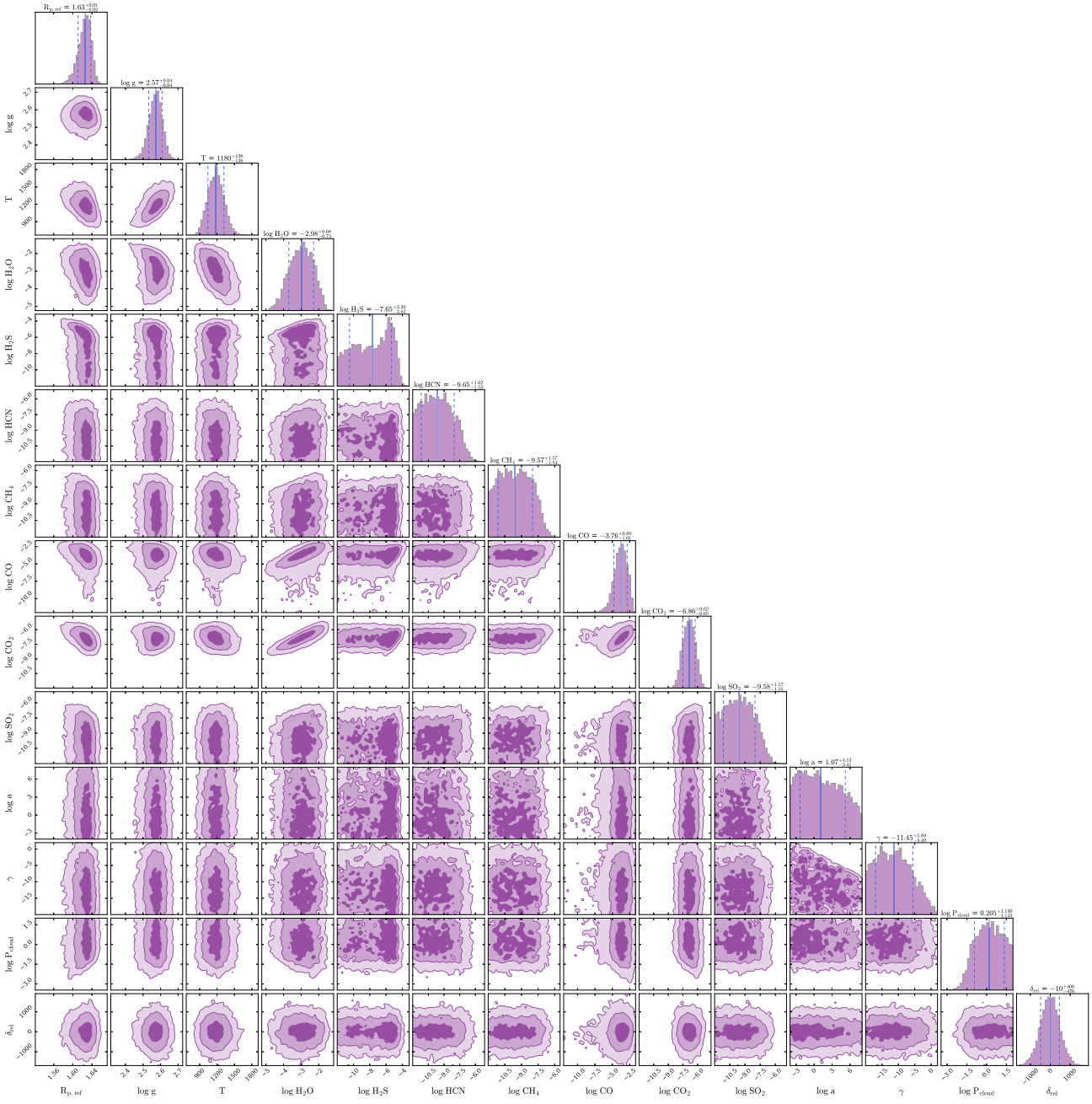

Figure D3. Posterior distributions from the POSEIDON free retrieval on the TIBERIUS $R \simeq 100$ transmission spectrum (Section 4.1.1).

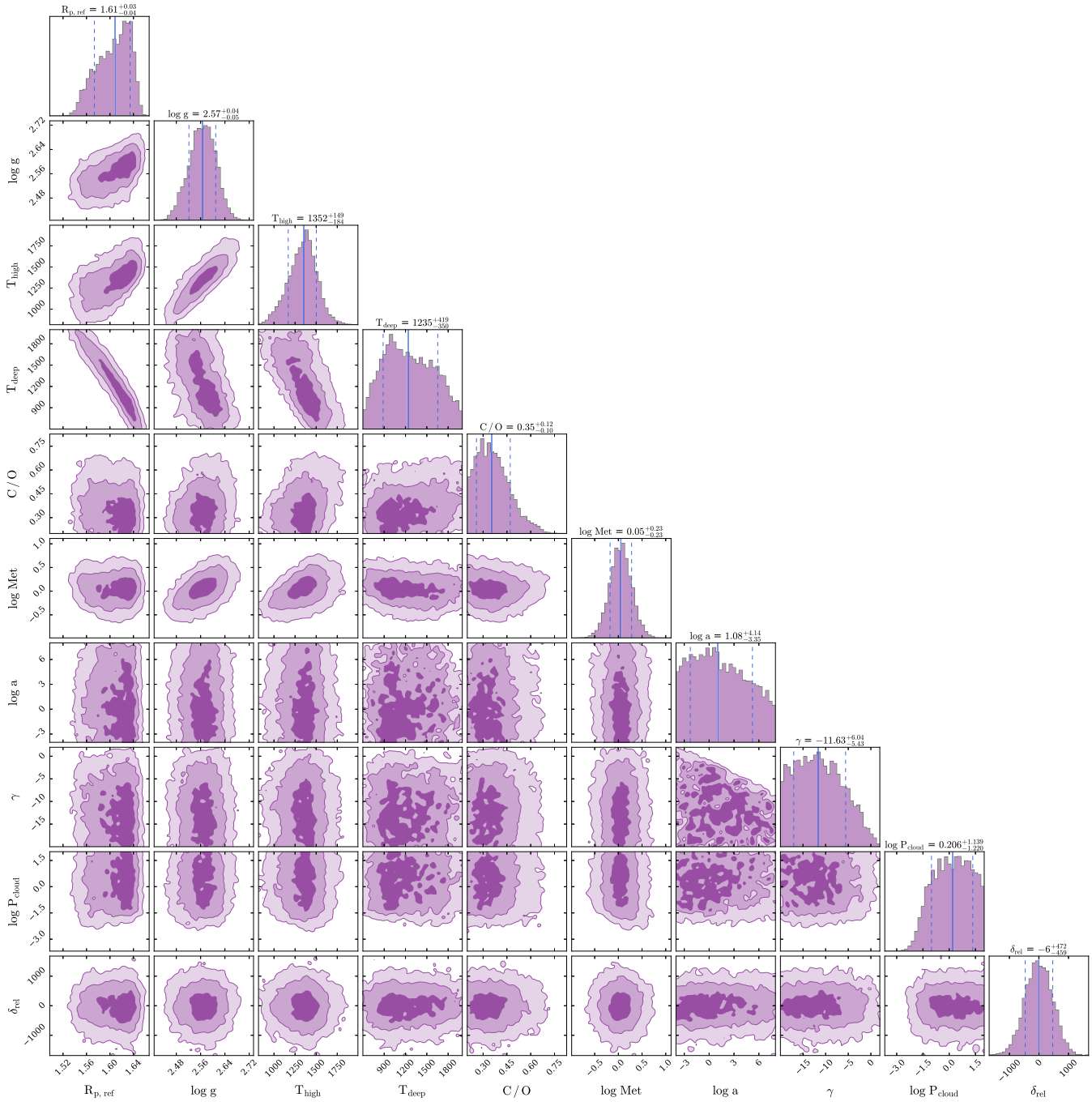

Figure D4. Posterior distributions from the POSEIDON equilibrium chemistry retrieval on the TIBERIUS $R \simeq 100$ transmission spectrum (Section 4.1.2).

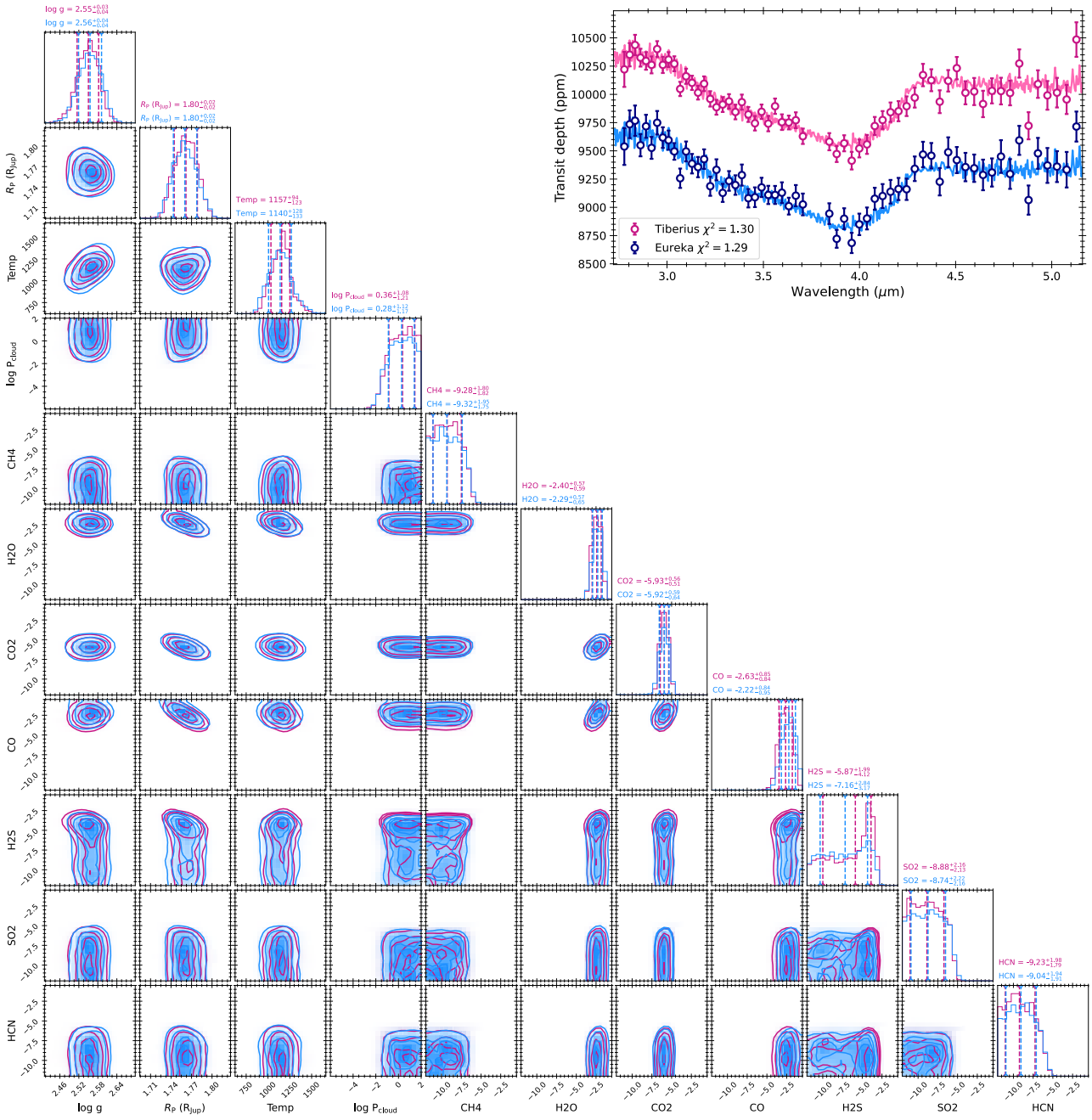

Figure D5. Posterior distributions from the PETITRADTRANS free chemistry retrieval (Section 4.2) on the $R \simeq 100$ transmission spectra from TIBERIUS and EUREKA!. Species abundances are quoted in mass fraction, rather than mixing ratio. The best-fitting models are shown in the top right.

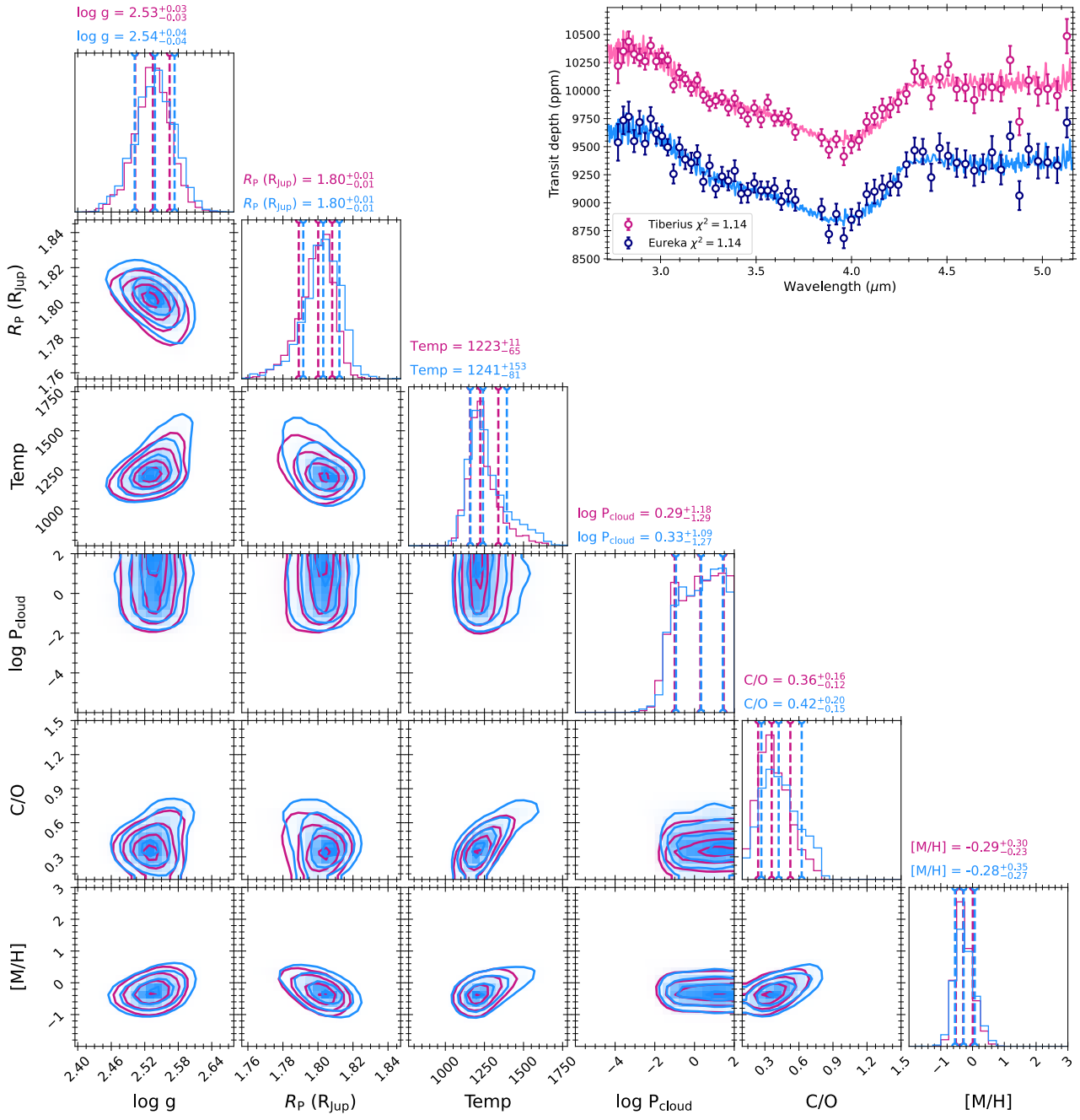

Figure D6. Posterior distributions from the PETITRADTRANS equilibrium chemistry retrieval (Section 4.2) on the $R \simeq 100$ transmission spectra from TIBERIUS and EUREKA!. The best-fitting models are shown in the top right.

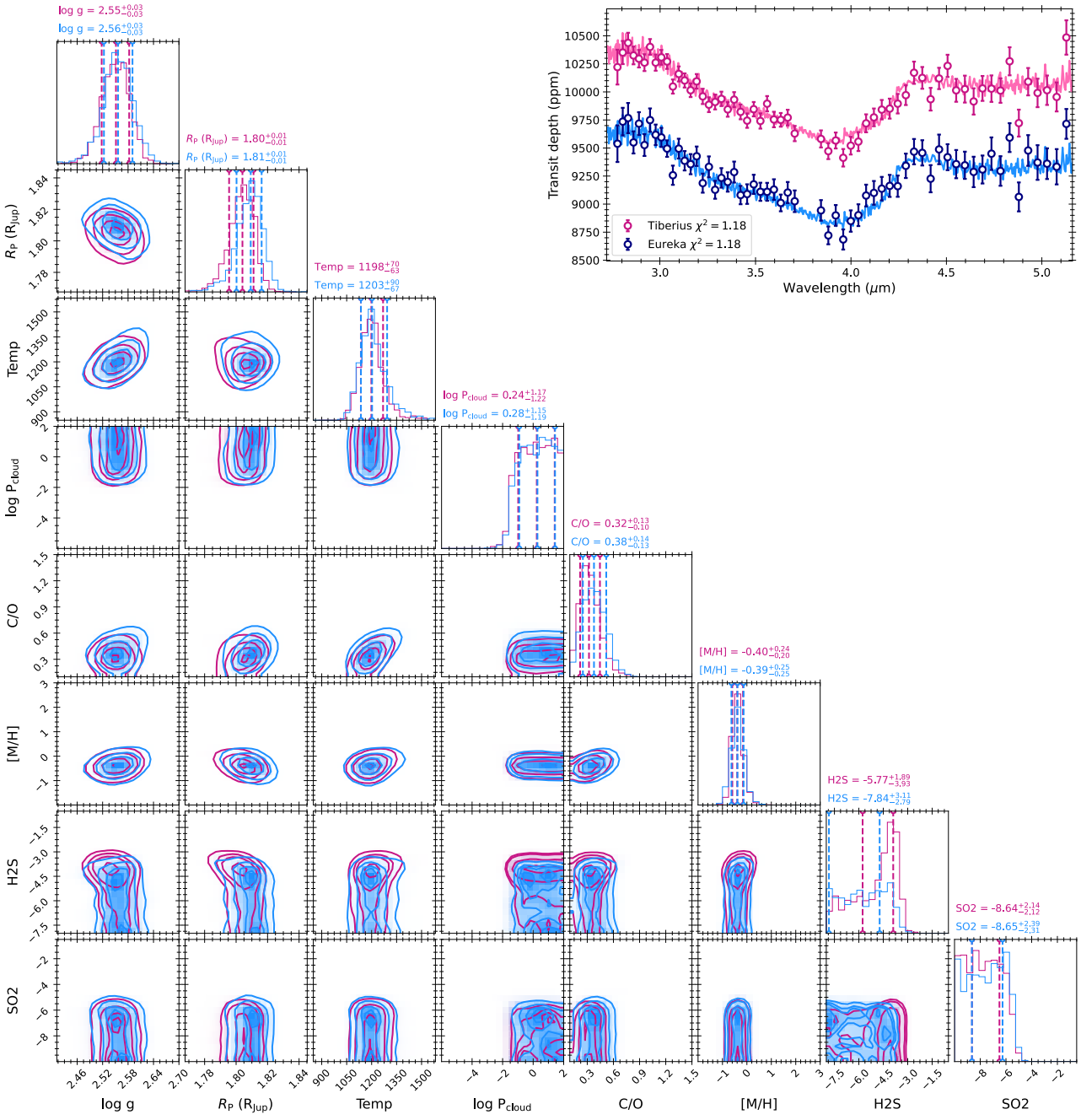

Figure D7. Posterior distributions from the PETITRADTRANS hybrid chemistry retrieval (Section 4.2) on the $R \simeq 100$ transmission spectra from TIBERIUS and EUREKA!. Species abundances are quoted in mass fraction, rather than mixing ratio. The best-fitting models are shown in the top right.

This paper has been typeset from a TEX/LATEX file prepared by the author.