The metal-poor atmosphere of a potential sub-Neptune progenitor

Saugata Barat [1] ✉, Jean-Michel Désert [1] ✉, Allona Vazan[2], Robin Baeyens [1], Michael R. Line [3], Jonathan J. Fortney[4], Trevor J. David[5], John H. Livingston [6,7,8], Bob Jacobs [1], Vatsal Panwar[1,9,10], Hinna Shivkumar [1], Kamen O. Todorov [1], Lorenzo Pino [11], Georgia Mraz[1] & Erik A. Petigura [12]

Young transiting exoplanets offer a unique opportunity to characterize the atmospheres of freshly formed and evolving planets. We present the transmission spectrum of V1298 Tau b, a 23-Myr-old warm Jupiter-sized ($0.91 \pm 0.05\,R_J$, where R_J is the radius of Jupiter) planet orbiting a pre-main-sequence star. We detect a mostly clear primordial atmosphere with an exceptionally large atmospheric scale height, and a water vapour absorption at a 5σ level of significance, from which we estimate a planetary mass upper limit (23 Earth masses, $0.12\,\mathrm{g\,cm^{-3}}$ at a 3σ level). This is one of the lowest-density planets discovered so far. We retrieve a low atmospheric metallicity ($\log Z = -0.7^{+0.8}_{-0.7}$ solar), consistent with solar/sub-solar values. Our findings challenge the expected mass–metallicity relation from core-accretion theory. Our observations can instead be explained by in situ formation via pebble accretion together with ongoing evolutionary mechanisms. We do not detect methane, which hints at a hotter-than-expected interior from just the formation entropy of this planet. Our observations suggest that V1298 Tau b is likely to evolve into a sub-Neptune.

Exoplanet population studies reveal the crucial impact of planet formation and early evolutionary mechanisms[1,2] on their demographic characteristics. However, evolutionary processes such as atmospheric mass loss driven by host star extreme ultraviolet (XUV) flux[3], interior cooling[4] and contraction[5] can substantially alter their thermal structure and composition within the first 100 Myr, thereby obscuring the imprints of planet formation. In this context, young transiting exoplanets represent a unique opportunity to probe the atmospheres of freshly formed planets and test formation and early evolution theories[6–8]. However, studying these young planets is challenging as most of them do not have well-constrained masses due to the large uncertainties

in radial velocity (RV) measurements from their highly variable host stars[9]. Young stars are known to have large spot coverage and frequent flaring activity[10], which can contaminate the measured transmission spectrum by the transit light source effect[11]. Most of the known young transiting planets[12,13] lie above the radius valley and are theoretically predicted to be Neptune or sub-Neptune/super-Earth progenitors[5].

V1298 Tau is one of the youngest transiting multi-planet systems known, consisting of three confirmed planets in a near 3:2:1 mean motion resonance and a fourth planet with an unconfirmed period[14–17]. The host is a 23-Myr-old weak-lined T Tauri star, which is a member of Group 29 (a young association in the foreground of the Taurus–Auriga

[1]Anton Pannekoek Institute, University of Amsterdam, Amsterdam, the Netherlands. [2]Astrophysics Research Center (ARCO), The Open University of Israel, Ra'anana, Israel. [3]School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA. [4]Department of Astronomy & Astrophysics, University of California, Santa Cruz, CA, USA. [5]Center for Computational Astrophysics, Flatiron Institute, New York, NY, USA. [6]Astrobiology Center, Mitaka, Japan. [7]National Astronomical Observatory of Japan, Mitaka, Japan. [8]Department of Astronomy, Graduate University for Advanced Studies (SOKENDAI), Mitaka, Japan. [9]Department of Physics, University of Warwick, Coventry, UK. [10]Centre for Exoplanets and Habitability, University of Warwick, Coventry, UK. [11]INAF–Osservatorio Astrofisico di Arcetri, Florence, Italy. [12]Department of Physics & Astronomy, University of California Los Angeles, Los Angeles, CA, USA. ✉e-mail: s.barat@uva.nl; j.m.l.b.desert@uva.nl

Table 1 | Best-fit transit depths, linear limb-darkening coefficients and root mean squared residual compared with expected photon noise for V1298 Tau b

Central wavelength (µm)	Transit depth (R_p^2/R_*^2)	u1	Residual (photon noise)
1.13	3,766±49	0.59±0.031	1.29
1.16	3,803±41	0.58±0.025	1.06
1.20	3,872±42	0.53±0.026	1.08
1.23	3,757±40	0.58±0.025	1.04
1.26	3,752±43	0.56±0.027	1.12
1.30	3,869±45	0.56±0.028	1.16
1.33	3,945±43	0.55±0.026	1.11
1.36	4,175±41	0.53±0.023	1.05
1.39	4,110±39	0.56±0.023	1.01
1.43	4,111±41	0.57±0.024	1.05
1.46	3,906±41	0.57±0.025	1.06
1.49	3,819±43	0.50±0.028	1.13
1.52	3,849±40	0.57±0.023	1.02
1.55	3,720±46	0.58±0.03	1.19
1.59	3,706±50	0.50±0.034	1.30
1.62	3,682±51	0.54±0.033	1.33
1.64	3,702±79	0.53±0.053	1.36

R_p is the radius of a planet at a reference pressure of 1bar. R_* is the radius of the star. u1 is the linear limb-darkening coefficient.

star-forming region)[18,19]. Several age estimates have been published for V1298 Tau: 23 ± 4 Myr (ref. [15]), 20 ± 10 Myr (ref. [20]) and 28 ± 4 Myr (ref. [21]). All of these estimates agree within 1σ and we adopt 23 ± 4 from ref. [15]. We observed one primary transit of V1298 Tau b using ten Hubble Space Telescope (HST) orbits with the Wide Field Camera 3 (WFC3) instrument in Grism-141 (G141) observing mode for General-Observer program (GO) 16083 (see 'Observations' in the Methods). V1298 Tau b is a warm (equilibrium temperature (T_{equ}) = 670 K)[14], Jovian-sized planet (0.8–0.9 R_J; where R_J is the radius of Jupiter)[14,16,17] orbiting its host star in 24.14 days (refs. [14–17]). Mass estimates made using RVs provide a Jovian mass (0.64 ± 0.19 M_J (ref. [20]); <0.5 M_J (ref. [16])), but the reliability of these constraints has recently been questioned[9]. The high stellar activity and uncertainty in the orbital period of planet e could result in unreliable RV mass estimates for all of the V1298 Tau planets[9]. Using RV mass measurements from ref. [20], ref. [22] concluded that V1298 Tau b would be stable to atmospheric mass loss due to its strong gravity.

The raw HST images were reduced using an in-house data-reduction pipeline[23] (see 'Data reduction' in the Methods for details). We extracted a broadband integrated 'white' light curve in the HST/WFC3 G141 bandpass (1.12 µm–1.65 µm) and used a divide-white common-mode approach to derive systematics-corrected spectroscopic light curves[24]. The extracted white and detrended spectroscopic light curves are shown in Extended Data Figs. 1 and 2, respectively. The detrended spectroscopic light curves were fitted with a batman planetary transit model, linear limb darkening and a linear stellar baseline (see 'Light curve analysis' in the Methods and Table 1 for details). We estimated the effect of unocculted star spots on the transmission spectrum using techniques outlined in ref. [11] (see 'Accounting for stellar activity' in the Methods).

Results

The transmission spectrum of V1298 Tau b (Fig. 1) shows a high-amplitude absorption feature around the 1.4 µm water band (~400 ppm) that is larger than those of known Neptune/super-Neptune-mass planets, such as HAT-P-26b (250 ppm)[25] and GJ 3470b (150 ppm)[26]. The water

absorption amplitude is also large in comparison with well-studied hot Jupiters, such as HD209458b (~200 ppm)[27]. The amplitude of the water feature is indicative of a large atmospheric scale height, revealing an extended H-rich atmosphere. Assuming a clear, cloud-free, H/He-rich and isothermal atmosphere, we determined the scale height of this planet (1,100 ± 200 km), from which we estimated the mass to be 24 ± 5 Earth masses ($M_⊕$) using a known method[28]. This mass estimate becomes an upper limit if the atmosphere is partly cloudy or hazy. We also derived an estimate of the planet mass from atmospheric retrievals, keeping the mass free (see 'Mass estimate' in the Methods and Extended Data Fig. 3). These fits yielded a 99.73% (3σ) upper limit of the mass of 23 $M_⊕$, consistent with the analytically derived mass upper limit (24 ± 5 $M_⊕$). RV measurements of this system provide Jovian/sub-Jovian mass (220 ± 70 $M_⊕$) for V1298 Tau b[16,20]. However, our observation rules out a 100 $M_⊕$ (-2σ lower limit from ref. [20]) transmission spectrum model (Fig. 1) at ~5σ confidence. We compared the derived mass and radius of planet b to the population of exoplanets (Fig. 2). With a density upper limit of 0.12 g cm^{-3}, V1298 Tau b is comparable to the lowest-density planets known (super-puffs)[29], but has a clear atmosphere compared with most super-puffs[30]. The estimated upper limit of the mass of V1298 Tau b is consistent with a Neptune/sub-Neptune-mass planet with a substantial H/He envelope (~40%; Fig. 2).

The observed transmission spectrum (without a stellar activity correction applied) was modelled using a one-dimensional (1D) radiative transfer code (PetitRADTRANS)[31]. We fixed the planetary mass to the estimated upper limit (24 $M_⊕$). We modelled the atmosphere using an analytic temperature–pressure ($T–P$) profile[32] and a grey cloud deck. The dominant carbon-bearing species at 670 K is expected to be methane[33]. However, we did not detect methane absorption around 1.6 µm (Fig. 1). The absence of methane has been reported for other warm planets[26], which may be explained by vertical mixing[33]. We simulated the effect of vertical mixing using a 'quench' pressure in our models (see 'Atmospheric models' in the Methods). The observed and modelled transmission spectra, and the retrieved atmospheric properties, are shown in Fig. 1. The retrievals converge to a low atmospheric metallicity ($\log Z = -0.1^{+0.66}_{-0.72}$ solar) compared with theoretical mass–metallicity relations for exoplanets[34] based on core-accretion theory[35] and known constraints for Neptune/sub-Neptune planets[36,37]. The observed spectrum can also be explained by lower planet masses (10–20 $M_⊕$) with even lower metallicities (0.1–0.01 solar). Cloudy models are marginally favoured over cloud-free models (Extended Data Fig. 4 and Table 2) for the fixed-mass retrievals. However, the retrieved cloud deck pressures are relatively high (≥10^{-2} bar), implying that clouds do not have a strong effect on the atmospheric absorption feature. Similarly, for free-mass retrievals, our posteriors showed an upper limit of ~10^{-1} bar for the cloud deck pressure.

Using the aforementioned derived mass upper limit and radius measurement (0.91 ± 0.05 R_J) from ref. [14], we estimated a bulk metallicity of 6× solar for V1298 Tau b, applying the formalism from ref. [34]. This estimate should be interpreted cautiously as the models presented in ref. [34] do not account for high interior flux from the planet. Given that V1298 Tau b is young with a potentially hot interior, the true bulk metallicity could be higher. The estimated bulk metallicity is shown in Fig. 3. However, the posteriors from our atmospheric retrievals (Extended Data Fig. 5) ruled out 6× solar values at ~3σ. The atmospheric metallicity of V1298 Tau b is significantly lower than its estimated bulk metallicity, which indicates that the atmosphere and interior of this planet may not be well mixed[34].

In Fig. 3 we present V1298 Tau b in the context of the exoplanet population. V1298 Tau b has a mass consistent with a Neptune/sub-Neptune or even a super-Earth and a metallicity comparable to (or lower than) Jupiter. A high-metallicity atmosphere (100× solar) with the estimated mass upper limit of this planet can be ruled out at ~5σ confidence (see the orange dashed line in Fig. 1). Therefore, in spite of being a likely Neptune/sub-Neptune, or even a super-Earth progenitor, V1298 Tau b
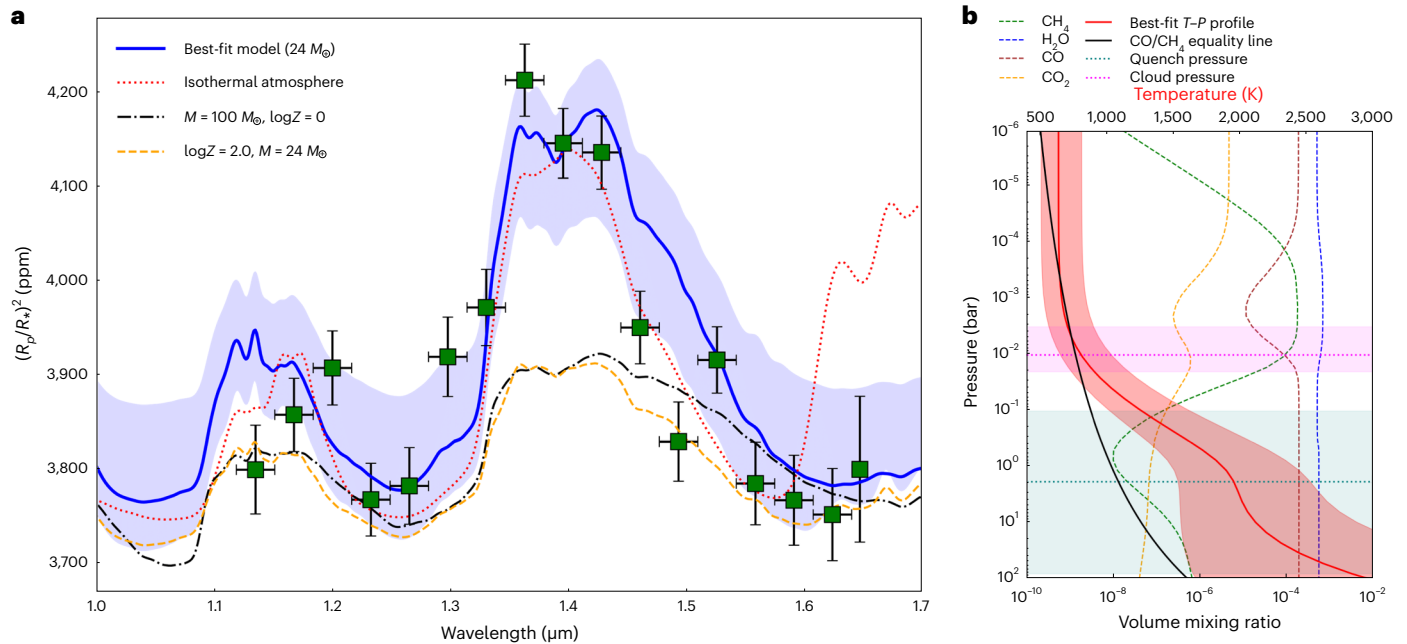
Fig. 1 | Observed HST/WFC3 transmission spectrum and retrieved *T*–*P* profile and molecular abundances for V1298 Tau b. a, Extracted transmission spectrum (stellar activity not corrected) of V1298 Tau b with 1σ uncertainties (green data points and error bars). The transit depths and uncertainties have been derived by fitting the spectroscopic light curves using an MCMC algorithm. The extraction of the spectrum is described in detail in 'Light curve analysis' in the Methods, and the transit depths and uncertainties are shown in Table 1. The solid blue line shows a best-fit atmospheric model from retrievals of the observed spectrum (see 'Atmospheric models' in the Methods). The dash-dotted black line shows a transmission spectrum for a 100 $M_\oplus$ solar metallicity model and the orange dashed line represents a

24 $M_\oplus$ 100× solar metallicity model, both of which are ruled out at 5σ. The red dotted line represents an isothermal model, which shows an absorption feature around 1.6 µm due to methane and is ruled out at 3σ. **b**, Best-fit *T*–*P* profile for 24 $M_\oplus$ retrieval with a 1σ confidence interval (red shaded region). The CH_4/CO transition[33] line, with the CH_4-dominated region on the left and the CO-dominated region on the right, is shown. The coloured dashed lines represent the equilibrium abundances for the chemical species included in our model (calculated for the red solid *T*–*P* profile). The retrieved grey cloud deck and quenching pressure and the 1σ confidence region from our retrieval analysis (Table 2) are shown.
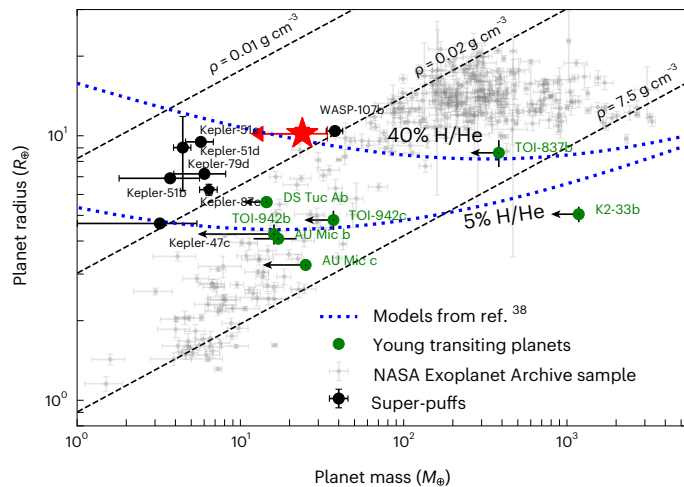

Fig. 2 | V1298 Tau b in the exoplanet mass–radius diagram. The mass upper limit of V1298 Tau b (red star) was calculated from the observed transmission spectrum using free-mass retrievals and the formalism presented in ref. 28 (see 'Mass estimate' in the Methods). Other known young planets and low-density super-puffs are shown for comparison. The population of mature transiting planets is shown for comparison (data and uncertainties obtained from the NASA Exoplanet Archive[80]). The arrows indicate mass upper limits. The blue dotted lines are theoretical models from ref. 38 and show that the measured mass and radius of V1298 Tau b are consistent with an atmosphere with a large H/He envelope (~40% mass fraction assuming a 10 $M_\oplus$ core). Constant density contours show the density (ρ) of V1298 Tau b (0.12 g cm^{-3}) is among the lowest discovered.

possesses an atmosphere that is 100–1,000× depleted in metal compared with Neptune and Uranus. The atmospheric metallicity in Fig. 3 has been shown with respect to stellar metallicity.

We also investigated the reason for the absence of expected methane. We ran chemical kinetics models that incorporated a self-consistent *T*–*P* profile, photochemistry and vertical mixing (see 'Atmospheric models' in the Methods). We found that photochemical processes did not play an important role in methane depletion (see 'Atmospheric models' in the Methods and Supplementary Figs. 3 and 4). These models, with different internal temperatures, demonstrate that it is possible to remove methane through deep quenching, although it requires a high interior temperature (Extended Data Fig. 6). At the highest intrinsic temperature we tested (T_{int} = 400 K), the quenched molar fraction of methane was still $10^{-4.7}$, which is close to the detectability limit ($10^{-5.5}$) with HST[33]. Retrievals using free chemistry (see 'Atmospheric models' in the Methods and Supplementary Fig. 1) put an upper limit of 10^{-6} on the methane volume mixing ratio. JWST observations of this planet will be more sensitive to such low methane abundances.

Discussion

The differences between the masses estimated from the atmospheric scale height and those from dynamical studies[16,20] could potentially originate from the treatment of the impact of stellar activity on RV signals, as recently demonstrated in ref. 9. There is also uncertainty in the orbital period of planet e (ref. 17) in this system, which could substantially impact the RV mass constraints. Complementary efforts to estimate the mass from transit timing variation measurements are ongoing (Livingston et al., manuscript in preparation).

Table 2 | Atmospheric parameters of V1298 Tau b retrieved from its transmission spectrum

Retrieval parameter	$24\,M_{\oplus}$	$15\,M_{\oplus}$	$10\,M_{\oplus}$	$24\,M_{\oplus}$ (corrected)
$\log[Z\,(Z_{\odot})]$	$-0.1^{+0.66}_{-0.72}$	$-0.9^{+0.9}_{-1.1}$	$-1.5^{+0.9}_{-1.1}$	$-0.73^{+0.81}_{-0.76}$
C/O ratio	$0.23^{+0.25}_{-0.13}$	$0.24^{+0.24}_{-0.13}$	$0.26^{+0.25}_{-0.15}$	$0.29^{+0.23}_{-0.17}$
$R_{p}\,(R_{J})$	$0.68^{+0.03}_{-0.04}$	$0.70^{+0.02}_{-0.02}$	$0.65^{+0.02}_{-0.01}$	$0.73^{+0.03}_{-0.04}$
T_{int} (K)	335^{+185}_{-210}	380^{+150}_{-250}	400^{+150}_{-250}	429^{+135}_{-253}
$\log[P_{quench}\,(\text{bar})]$	$0.26^{+1.68}_{-1.36}$	$0.42^{+1.10}_{-0.63}$	$0.46^{+1.30}_{-0.78}$	$0.81^{+1.16}_{-0.11}$
T_{equ} (K)	760^{+210}_{-160}	670 (fixed)	670 (fixed)	757^{+230}_{-180}
$\log[P_{cloud}\,(\text{bar})]$	$-2.1^{+0.41}_{-0.47}$	$-1.6^{+0.54}_{-0.57}$	$-2.0^{+0.5}_{-0.6}$	$-1.28^{+0.32}_{-0.6}$

See 'Atmospheric models' in the Methods for details of the atmospheric models used. $Z_{\odot}$ is solar metallicity, T_{int} is interior temperature, P_{quench} is quench pressure and P_{cloud} is grey cloud deck pressure.

The low envelope metallicity and relatively large H/He content that we measured for V1298 Tau b are in agreement with early evolution models[38], yet this planet must have been on the verge of runaway gas accretion. We emphasize that the origin and early evolution of Neptunes/sub-Neptunes have been open questions: it is not known why these planets accreted only a small fraction of H/He and did not become gas giants[7,39]. Such systems probably formed in situ, either early with an enhanced atmospheric opacity due to dust grains[7] or with substantial disk–envelope interaction to replenish the proto-atmosphere with high-entropy gas[40]. Late formation in a depleting transitional disk in which the core does not have enough time to accrete a large H/He envelope[6] can also produce Neptune/sub-Neptune-mass planets.

The standard core-accretion picture of planet formation[35] predicts a mass–metallicity relationship that has been observed in the Solar System[41] and reported for exoplanets[42]. The relatively water-poor atmosphere of V1298 Tau b that we found in this work indicates that this planet must have spent most of its accretion phase within the water ice line, thereby failing to accrete volatile-rich pebbles[43]. The volatile content of the inner disk can be strongly affected by the growth of massive planets in the outer part of the disk[44]. In this scenario, a massive planet, formed beyond the water ice line, blocks the supply of volatile-rich pebbles to the inner part of the disk. RV constraints on the mass of V1298 Tau e put it in a Jupiter/sub-Jupiter range[16,17,20], with an orbital period that may be longer than 40 days. Pebble filtering could therefore play an important role in this system by producing volatile-poor atmospheres of the inner planets.

Alternatively, V1298 Tau b could have accreted volatile-rich material that ended being locked up in the interior of the planet. Young planets could also experience extreme mass loss driven by the intense XUV flux of their active host stars. Using our update mass upper limits ($24\,M_{\oplus}$) for this planet, we simulated the mass and radius evolution (Fig. 4; see also 'Atmospheric evolution models' in the Methods). We estimated the Jean's escape parameter[45] to be 27. Our calculations suggest that V1298 Tau b is susceptible to photoevaporation, in contrast to the conclusions of ref. 22 based on RV mass estimates. V1298 Tau b may lose up to a few Earth masses within first billion years of its life. The rocky pebble/planetesimal accretion theory of planet formation[46] predicts a gradually mixed interior structure, as observed for Jupiter[47]. We show two possible interior and evolution models for V1298 Tau b (Extended Data Fig. 7): a core–envelope structure and a diluted core structure. The observed mass, radius and metal-poor envelope can all be explained by both models, but in the diluted core scenario, the atmospheric metallicity is expected to evolve due to the removal of the upper layer of the atmosphere by mass loss, as well as convective mixing in the interior that could ultimately reconcile V1298 Tau b with the mature exoplanet population[48].
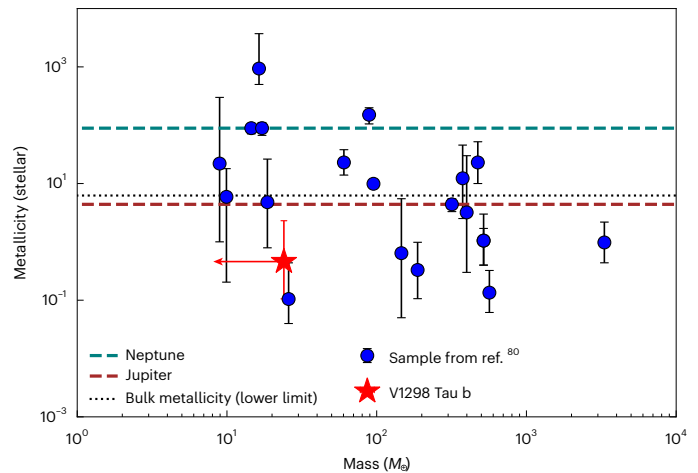
Fig. 3 | V1298 Tau b in a mass–metallicity diagram with a sample of exoplanets. The uncertainties (error bars) have been taken from individual references for each planet in the sample, which was compiled from ref. 81. The atmospheric metallicity of V1298 Tau b is derived from the retrieval analysis and has been converted into a multiple of stellar metallicity using the estimate of 0.1 ± 0.15 dex (ref. 20). The metallicity uncertainty is calculated by adding the atmospheric metallicity uncertainty from the retrievals (Table 2) and the stellar metallicity uncertainty in quadrature. The dashed blue and brown lines show the metallicities of Neptune and Jupiter, respectively. We note that the Solar System metallicity estimates are from methane abundance measurements[41], whereas the metallicity estimates for exoplanets are derived from oxygen abundance measurements. V1298 Tau b has a mass consistent with Neptune/sub-Neptunes or potentially even super-Earths, but its metallicity is comparable to giant planets such as Jupiter. From the estimated mass upper limit and radius of V1298 Tau b, we estimate a bulk metallicity following the formalism of ref. 34, shown by a black dotted line.

Self-consistent atmospheric modelling for V1298 Tau b required an extremely high internal temperature (~400 K) and strong vertical mixing to explain the non-detection of methane. The internal temperature (170 K) expected from the early evolution models (Extended Data Fig. 7) is consistent with previous theoretical expectations[33]. Internal temperatures as high as 300–400 K may require external heating mechanisms such as tidal heating[49]. Alternatively, photolytic destruction of methane could also potentially produce a methane-poor atmosphere[50], which may be feasible given the young age and high activity levels of V1298 Tau. We tested this hypothesis by re-running our self-consistent forward model using a published UV spectrum of V1298 Tau[51]. However, we found that photochemistry did not impact the methane abundance for V1298 Tau b for pressures higher than 10^{-4} bar, even for an extreme case (1,000× solar XUV flux; Supplementary Fig. 4).

The detection of a strong water vapour absorption feature in the near-infrared (NIR) spectrum of V1298 Tau b allowed us to put a stringent upper limit on its mass as determined analytically from the atmospheric scale height ($24 \pm 5\,M_{\oplus}$) and from retrievals ($23\,M_{\oplus}$ at a 99.73% level). The V1298 Tau system is likely to have formed either late, within the water ice line in a gas-poor, dry and depleting protoplanetary disk, or early in the inner region of the disk with an accretion rate that was probably moderated by disk gas replenishment or enhanced envelope opacity. V1298 Tau b is likely to undergo atmospheric mass loss and could end up as a Neptune, a low-density sub-Neptune or even a super-Earth (Fig. 4). The haze-free, low-metallicity atmosphere of V1298 Tau b, a potential Neptune/sub-Neptune progenitor, is distinct from its hazy/metal-rich mature counterparts. This could be a manifestation of the ongoing evolution of its atmospheric composition. We leveraged the absence of spectral signatures of methane and provide constraints on the internal temperature (>300 K) of the planet, which hints at external heating mechanisms such as tidal heating. Future JWST studies will be important in further investigation of the atmospheric chemistry of this planet.
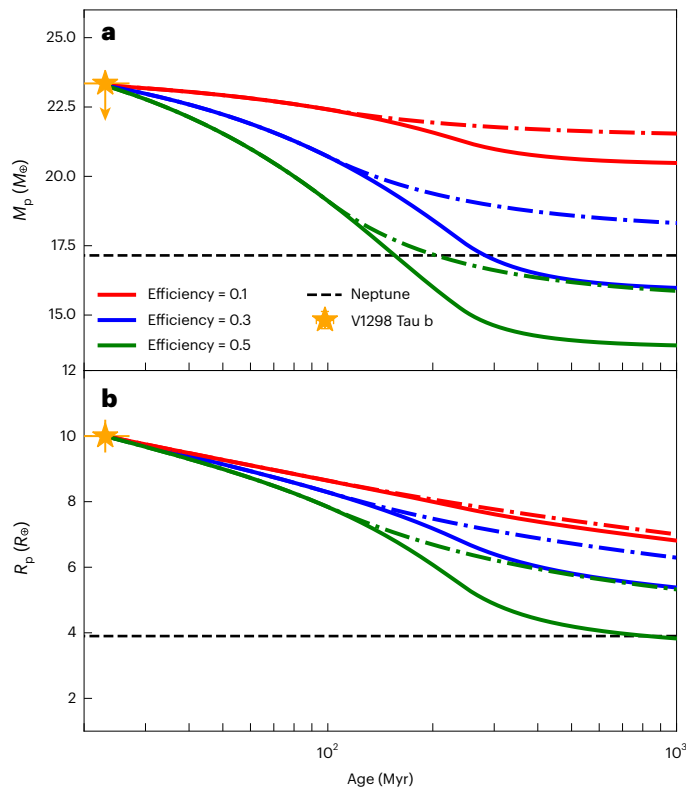
Fig. 4 | Evolutionary track for the mass and radius of V1298 Tau b using the mass upper limit derived in this work. a,b, The mass (**a**) and radius (**b**) evolution tracks have been calculated for the first billion years using energy-limited atmospheric evolution models presented in the platypos code[75] (see 'Atmospheric evolution models' in the Methods). The evolution of the radius is a combined effect of atmospheric contraction and mass loss. Simulations for different values of the mass loss efficiency parameter (0.1, 0.3, 0.5) are shown with different colours. The solid lines show a high stellar activity track (activity timescale 250 Myr) and the dash-dotted lines show a low stellar activity track (activity timescale 100 Myr; for details see ref. [75]). The current position of V1298 Tau b is marked by an orange star. The arrow on the orange star in **a** represents the mass upper limit we use from this work. The error bars on the age and radius of V1298 Tau b represent radius and age uncertainty obtained from ref. [14].

Methods

Observations

The observations were taken using the HST/WFC3 G141 grism in bidirectional spatial scanning mode, covering a range of 1.1–1.7 μm with a scan rate of 0.23 arcsec s^{-1}. This resulted in 180 exposures over ten HST orbits. The individual pixels reached a maximum flux level of 30,000 electrons, which is roughly 40% of the saturation level and well within the linear response regime of the detector. We used the 256 × 256 pixel subarray and SPARS25, NSAMP=5 readout mode, which resulted in 88.4 s exposures.

Data reduction

We used a data-reduction pipeline for our data analysis[23,52]. The WFC IR detectors were read multiple times non-destructively (without flushing out the accumulated charge) during an exposure. First, sub-exposures were formed for each exposure by subtracting consecutive non-destructive reads and each sub-exposure was reduced separately for improved background subtraction and cosmic ray rejection. We calculated a wavelength solution by matching the first exposure of the visit to a convolution of a PHOENIX stellar spectrum[53] for V1298 Tau (temperature = 4,920 K (ref. [15])) with the response function of G141.

We applied a wavelength-dependent flat-field correction and flagged bad pixels with data quality DQ = 4, 32 or 512 by calwf3 and

applied a local median filter to identify cosmic rays and clip pixels that deviated more than five median deviations. On average, we found that 0.53% of pixels were affected by cosmic rays for each sub-exposure. To account for the dispersion direction drift of the spectrum, we used the first exposure of a visit as a template and shifted the spectrum for each exposure along the dispersion direction to match the template. The maximum shift that we measured was 0.3 pixels. Finally, we applied an optimal extraction algorithm[54] on each sub-exposure to maximize signal-to-noise ratios. We shifted and shrank the spectra of each sub-exposure to match the wavelength grid of the first sub-exposure by a maximum of 1.05 pixels and 0.65%.

Light curve analysis

WFC3 light curves are known to exhibit strong time-dependent ramp-like (charge-trapping) and visit-long systematics[24,27,55]. It is known that the first orbit of each visit has stronger systematics than the rest of the visit. Following common practice[27], we excluded this orbit from the rest of the analysis. We modelled the white light curve instrumental systematics using the charge-trapping model RECTE[56]. The out-of-transit baseline is a combination of instrumental visit-long slopes, well known for HST/WFC3 time-series observations[27], and rotational variability from the active young host star. We modelled the instrumental systematics, stellar rotational variability and transit shape simultaneously. Visit-long slopes have been modelled using linear functions in time[27,57], but temporal nonlinearity exhibited by the baseline highlights the effects of stellar variability. We tested polynomial functions of first order to fourth order, as well as a sinusoidal function, to model the baseline. A third-order polynomial provided the best fit (lowest Bayesian information criteria (BIC) value) to the observations (Supplementary Table 2). Therefore, we modelled the stellar baseline using a third-order polynomial and the stellar disk using a linear limb-darkening model. The best-fit polynomial function is shown in Extended Data Fig. 1 and shows ~0.3% variability during the entire visit. The planetary transit signal was modelled using batman[58]; we fixed the orbital parameters to known literature values[15–17]. We ran a Markov chain Monte Carlo (MCMC) using emcee[59] to estimate model parameter uncertainties (Supplementary Fig. 2). We found the ninth exposure of the seventh orbit to be affected by a satellite crossing event and excluded this exposure[60].

We generated 7-pixel-bin spectroscopic light curves from the reduced 1D stellar spectra across 17 wavelength channels. We detrended the spectroscopic light curves using a common-mode approach given the deviations from the standard HST instrument systematics (possibly due to stellar activity). The common-mode divide-white has been used previously for WFC3 analysis[24]; it adopts an agnostic approach to the exact mathematical form of the instrument systematics assuming it is independent of wavelength. We modelled the spectroscopic light curves using a batman model and a linear stellar baseline. We fitted for the linear limb-darkening coefficient. The observed white light curve, best-fit transit model and the derived systematics function are shown in Extended Data Fig. 1. The systematics detrended spectroscopic light curves, along with the residuals, are shown in Extended Data Fig. 2. We also derived the transmission spectrum by fitting each spectroscopic light curve using a RECTE and polynomial stellar baseline models, and the derived spectrum agrees within 1σ with the common-mode spectrum. However, the quality of the fits in the common-mode approach was superior. The residual noise in all the spectroscopic channels was less than 1.3× the expected photon noise and the average precision on the extracted transit depths was 47 ppm. The fitted transit depths and linear limb-darkening coefficients are shown in Table 1. The root mean-squared noise is relatively high[61]; however, this could be a combination of stellar variability, spot crossings and high measured horizontal (*x*) shifts.

We noted a possible bright spot occultation in the third orbit and also a potential flaring event affecting the latter half of the seventh

orbit (Extended Data Fig. 1). To estimate the effect of these exposures on the derived transmission spectrum, we fitted the spectroscopic light curves with and without these exposures. We did not find any change in the derived transmission spectrum and the average residuals decreased by 3 ppm when these exposures were excluded. We concluded that the removal of these exposures did not have a significant manifestation on the spectrum. We also tested the effect of the large horizontal drift of the telescope. We incorporated a linear function of x shifts as a correction factor for the white light curve fits, following the approach used in ref. 55. We found $\Delta BIC = 3$ when we included horizontal drift into the fitting algorithm, and hence we concluded that including the effect of horizontal drifts was not statistically significant.

Accounting for stellar activity

V1298 Tau is a young pre-main-sequence star, known to exhibit 2–3% variability in Kepler and Transiting Exoplanet Survey Satellite light curves[15–17]. Variability in such young stars can be attributed photospheric inhomogeneity (star spots and faculae) and fast stellar rotation. Unoccculted star spots can contaminate the observed transmission spectrum[62]. We estimated the effect of stellar contamination on the transmission spectrum of V1298 Tau b following the prescription of ref. 11. We adopted a surface inhomogeneity model (20% spot coverage) for V1298 Tau from ref. 63. Photospheric temperature contrasts have been studied for T Tauri stars[64]; stars with photospheric temperatures similar to V1298 Tau can have spot temperature contrasts of up to 1,000 K. We estimated an extreme case contamination spectrum for V1298 Tau assuming 20% spot coverage and a 1,000 K spot temperature contrast. The stellar-contamination correction function is shown in Supplementary Fig. 5. As V1298 Tau is a K-type star, even extreme temperature contrasts do not result in the appearance of water absorption features in the contamination spectrum, which can be the case for cooler M dwarfs[62]. Therefore, the mass measurement, which depends on transit depth difference between different spectroscopic channels, is relatively unaffected. A comparison between the corrected and uncorrected spectra is shown in Supplementary Fig. 5. We re-ran retrievals on the contamination-corrected transmission spectrum. The retrievals were identical in set-up to the uncorrected case (see 'Atmospheric models'). The posterior distributions of the parameters are shown in Extended Data Fig. 5 with the posteriors from the uncorrected spectrum. All of the parameters agreed for both cases within 1σ. The retrieved atmospheric metallicity in the corrected case favoured more sub-solar values compared with the uncorrected case, thereby confirming the robustness of the conclusions drawn in this work. The retrieved parameters are shown in Table 2. The effect of stellar absorption has been seen in the limb-darkening coefficients[24]. We set the limb-darkening coefficients as free parameters while fitting for the spectroscopic light curves; the results are provided in Table 1. The limb-darkening coefficients did not show any effect of stellar absorption. To further confirm that the water absorption feature we found in the spectrum of V1298 Tau b around 1.4 μm is of planetary origin, we defined a quantity B as the ratio of the flux observed in two wavelength bands:

$$B = \frac{\int_{\lambda_1}^{\lambda_2} F(\lambda)\, \mathrm{d}\lambda}{\int_{\lambda_3}^{\lambda_4} F(\lambda)\, \mathrm{d}\lambda} \qquad (1)$$

where F is electrons per unit wavelength in the 1D extracted spectra of our reduced exposures, λ_1 and λ_2 give us lower and upper limits of the first wavelength band and λ_3 and λ_4 give us lower and upper limits of the second wavelength band. We calculated B for all of the exposures, first using the wavelengths 1.25–1.35 μm (left end of the water feature) and 1.45–1.55 μm (right end of water feature) (top panel in Supplementary Fig. 6) and subsequently using 1.35–1.45 μm (centre of water feature) and 1.45–1.55 μm (bottom panel in Supplementary Fig. 6). For the latter

case we found an excess absorption during the transit of the planet, which indicates that the water absorption is of planetary origin.

Atmospheric models

We used the publicly available 1D radiative transfer code PetitRADTRANS to retrieve the atmospheric properties of V1298 Tau b from its observed transmission spectrum. The transmission spectrum did not show a methane absorption signature around 1.6 μm as would be expected for a warm planet like V1298 Tau b based on equilibrium chemistry. The lack of methane can be explained by disequilibrium processes such as vertical mixing[33,65] dredging up methane-poor gas from the hot interior parts of the atmosphere. In our retrieval framework, we modelled this effect using a quenching pressure, where the volume mixing ratios of C-, H-, O- and N-bearing molecules were calculated using PetitRADTRANS. However, the molecular concentrations were held constant above the quench point. We modelled the atmospheric thermal structure with a Guillot T–P profile[32] shown in equation (2), where T_{equ} and T_{int} are the equilibrium and internal temperature of the planet. κ_{IR} is the average IR atmospheric opacity and γ is the ratio between optical and IR opacity. τ is the optical depth and g is acceleration due to gravity. We constrained the models by fixing the values of both κ_{IR} (to 0.01 cm^2 g^{-1}) and γ (to 0.01), assuming the atmospheric opacity at the observed band pass to be water dominated. We included H_2O, CH_4, CO_2 and CO opacities in our retrieval framework, as these molecular species have absorption features in the NIR[66]. We did not include HCN or NH_3 opacities in our retrievals as we did not find evidence of these species in free retrievals (Supplementary Fig. 1). We assumed a grey cloud deck opacity model to simulate cloud absorption.

$$T^4 = \frac{3T_{int}^4}{4}\left(\frac{2}{3} + \tau\right) + \frac{3T_{equ}^4}{4}\left[\frac{2}{3} + \frac{1}{\gamma\sqrt{3}} + \left(\frac{\gamma}{\sqrt{3}} - \frac{1}{\gamma\sqrt{3}}\right)\mathrm{e}^{-\gamma\tau\sqrt{3}}\right], \qquad (2)$$
$$\tau = P\kappa_{IR}/g$$

We fixed the mass of the planet to 24 $M_\oplus$ based on the mass upper limit estimated analytically from the scale height. The free parameters in our models were atmospheric metallicity, the C/O ratio, R_p, T_{equ}, T_{int}, P_{quench} and P_{cloud}. We ran an MCMC with 3,000 burn-in steps and 30,000 post-burn-in steps with 50 walkers. We imposed uniform priors on the fitting parameters. The posterior distribution of the fitted parameters is shown in Extended Data Fig. 5. We retrieved a sub-solar/solar metallicity. The retrieved equilibrium temperature was consistent with the expected equilibrium temperature of the planet. The retrieved parameters are summarized in Table 2.

We tested the importance of the internal temperature by fixing the internal temperature to 0 K (that is, fitting for an isothermal atmosphere). Models with high internal temperatures were statistically favoured by a $\Delta BIC = 50$. Comparing 0 K interior temperature models with 300 K interior temperature models (Fig. 1), we concluded that high internal temperature models are preferred at 3σ. We also performed free-chemistry retrievals using an isothermal atmosphere (Supplementary Fig. 1). This yielded an upper limit to the methane volume mixing ratio in the atmosphere (~10^{-6}) that is lower than the detection threshold for HST[33], thereby independently confirming the non-detection of methane. The free retrieval did not yield evidence for other molecular species such as HCN or NH_3, putting upper limits of 10^{-6} on their volume mixing ratios. We explored the effect of fixing the planet's mass in Extended Data Fig. 4 and the following section.

We constructed self-consistent atmospheric models with varying internal temperatures to study the quenching of methane and CO in the deep atmosphere. We computed the T–P profile using petitCODE[67,68], assuming radiative–convective equilibrium. Irradiation onto the planet was computed assuming a planetary-wide energy redistribution, with a host star effective temperature and radius of 4,970 K and 1.31 $R_\odot$, semi-major axis of 0.1688 au and planetary intrinsic temperatures of

100–400 K. Using our retrievals as a guidance, a solar metallicity was adopted with a slightly sub-solar C/O of 0.3. We achieved this C/O by reducing the carbon abundance from its solar value. The resulting temperature profiles are shown in Supplementary Fig. 7. Subsequently, we used a 1D chemical kinetics model[69] in combination with a photochemical network[70] to calculate self-consistent vertical quenching pressures for the main atmospheric species. We performed our calculations with a constant eddy diffusion coefficient (K_{zz}) of 10^{10} cm^2 s^{-1}. This value, although high, is in line with the expected values for convective mixing in giant planets and brown dwarfs[71–73]. We included photochemistry in our models, but we found that it did not substantially affect the molecular abundances at pressures typically probed by transmission spectroscopy. We tested the effect of a higher XUV flux of the host star by computing models for scaled solar spectra (10–1,000×; Supplementary Figs. 3 and 4). The resulting chemical disequilibrium abundances for methane, CO and water are shown in Supplementary Fig. 4. We found that the planet should have a high internal temperature (~300–400 K) for the carbon chemistry to be dominated by CO. This is consistent with the high internal temperature and deep quenching concluded from the retrieval analysis.

Mass estimate

We ran an atmospheric retrieval, keeping mass as a free parameter. The posterior distribution is shown in Extended Data Fig. 3. We found planet mass with 14, 19 and 23 $M_{\oplus}$ as the 68%, 95% and 99.73% confidence limits. The metallicity in this case yielded an upper limit of solar value at 95% confidence. The $3 - \sigma$ mass upper limit derived from the free-mass fit is consistent with the analytical calculation; however, the free-mass fit favoured lower masses (10 $M_{\oplus}$ from the posterior distribution peak; Extended Data Fig. 3). Overall, even for planet masses lower than 24 $M_{\oplus}$, given the Jovian size of the planet, the atmosphere is expected to be H/He dominated and the atmospheric chemistry to remain unaffected. Therefore, the conclusions of mass lower than 24 $M_{\oplus}$ and solar/sub-solar atmospheric metallicity are further supported by this test.

We estimated the mass of V1298 Tau b (M_p) from the transmission spectrum using an analytical approach described in ref. 28.

$$M_p = \frac{kTR_p^2}{\mu GH} \qquad (3)$$

We used the radius measurement from Kepler[14] ($0.91 \pm 0.05\,R_J$) and an equilibrium temperature of 670 K for the calculation. k is the Boltzmann constant, G is universal gravitational constant, μ is mean molecular weight which is assumed to be 2.33 for a primordial H/He atmosphere and H is the scale height of the planet. We estimate the scale height from the observed spectrum of V1298 Tau b. The height of an atmosphere can be estimated using equation (1) of ref. 74:

$$z(\lambda_2) - z(\lambda_1) = H \ln\left(\frac{\sigma(\lambda_2)}{\sigma(\lambda_1)}\right) \qquad (4)$$

In equation (4), σ is the absorption cross-section at a given wavelength and z is the measured radius of the planet at a given wavelength. We estimated 2.7 scale heights to account for the 1.4 μm water absorption feature, assuming a water-dominated atmospheric opacity and a cloud-free atmosphere. Given the young age and inflated size, we assumed a primordial H/He-rich atmosphere and fixed the mean molecular mass to 2.33. We found a large atmospheric scale height for V1298 Tau b of $1{,}000 \pm 200$ km and a mass estimate of $24 \pm 5\,M_{\oplus}$. The reported radius for V1298 Tau b differs between K2 ($0.91 \pm 0.05\,R_J$; ref. 14) and Transiting Exoplanet Survey Satellite ($0.85 \pm 0.03\,R_J$; ref. 17), which is explained by the stellar variability. From the present HST white light curve, we measured a radius of $0.84 \pm 0.003\,R_J$ for this planet. We repeated the analytical mass calculation using the HST radius measurement ($18 \pm 4\,M_{\oplus}$). The mass estimates for different radii of this planet

are shown in Supplementary Table 1. The mass estimates are consistent within 1σ for K2 and HST radii. As V1298 Tau is a variable star, absolute transit depths are expected to vary between epochs, biasing radius measurements. We therefore adopted the radius from K2 observations, which was obtained from multiple transits (three) and had a long baseline coverage (80 days) for effective modelling of the stellar variability. However, the scale height measurement of V1298 Tau b for our observation should be unaffected, as it is a relative measurement and independent of the continuum. Given that V1298 tau is a K-type star, the atmospheric water absorption is relatively unattenuated due to spot contamination (see 'Accounting for stellar activity' for further details). We tested this by analytically estimating the planet mass from an extreme-case stellar-contamination-corrected spectrum (for a comparison between uncorrected and stellar-contamination-corrected spectra, see Supplementary Fig. 6). This yielded a mass estimate of $23 \pm 5\,M_{\oplus}$, consistent with our previous estimate.

This estimate can be interpreted as an upper limit, given the assumption of a clear atmosphere. In the presence of aerosols (for example, clouds, hazes) in the atmosphere, the measured scale height from the spectrum would be underestimated, therefore leading to an overestimation of the mass. Given the observed spectrum, a cloud-free case would therefore yield the maximum possible mass for this planet.

To estimate the impact of mass uncertainty on atmospheric parameters, we ran retrievals on the observed transmission spectrum by fixing the temperature to 670 K for different masses (24, 15, 10 and 5 $M_{\oplus}$) and both cloud-free and cloudy cases (Extended Data Fig. 4). We included the same molecules as the 24 $M_{\oplus}$ case, as molecular opacities do not depend on the planet's gravity. We were able to fit the observations with 24, 15 and 10 $M_{\oplus}$ models. The 24 $M_{\oplus}$ (mass upper limit) model converged at solar atmospheric metallicity; for lower-mass models (10–20 $M_{\oplus}$) our retrievals converged at even lower (0.1–0.01 solar) metallicities to fit the water absorption feature. However, for the 5 $M_{\oplus}$ case, our retrievals did not converge as it could not reproduce the water absorption signal. We tested the robustness of the estimated mass upper limit by running a retrieval with 40 $M_{\oplus}$. This model failed to reproduce the observed water feature and could be rejected at high confidence.

Thus, from the transmission spectrum we could estimate a robust mass upper limit and conclude that V1298 Tau b is likely to be a Neptune, a low-density sub-Neptune or even a super-Earth progenitor[75].

Atmospheric evolution models

The atmospheric evolution models shown in Fig. 4 were simulated using the open-source platypos code[75]. The code calculated the mass loss rate at a given point in time using the energy-limited mass loss formalism[76,77] and evolved the planet's physical properties (mass and radius) at every step of the calculation. The radius evolution was a combined effect of atmospheric contraction and mass loss, and the updated sizes of the planets were calculated from the scaling relation given in ref. 38. We adopted the stellar luminosity from ref. 75. These simulations were performed considering the estimated mass upper limit (24 $M_{\oplus}$). For lower masses, we can expect higher mass loss rates.

Comparison with a previous study

V1298 Tau b was included in a sample of 70 transiting exoplanets whose spectra have been shown in ref. 78. Edwards et al. used a different pipeline (Iraclis[55]) for the data reduction. They also used a common-mode approach to derive the spectrum of this planet. The transmission spectrum obtained in this work is consistent within 1σ with the results of ref. 78, except for a constant offset of ~500 ppm. The constant offset is a result of ref. 78 using the third orbit in their white light curve fits, which we chose to exclude because of a potential spot crossing in that orbit. We tested the effect of including the third orbit in the white light curve fits. We found ΔBIC = 170 in favour of excluding the third orbit from the fits. The transmission spectrum obtained in ref. 78 and this work have been shown together for comparison in Supplementary Fig. 6.

Data availability

The data used in this study can be obtained from the Mikulski Archive for Space Telescopes (MAST; https://mast.stsci.edu/) and are associated with HST GO 16083. The reduced 1D stellar spectra obtained in this work are available via Zenodo at https://doi.org/10.48550/arXiv.2312.16924 (ref. 79).

Code availability

This research made use of public software including astropy, lmfit, emcee, platypos, petitCODE and PetitRADTRANS. Additional data analysis notebooks used for the figures and calculations are available via Zenodo at https://doi.org/10.48550/arXiv.2312.16924 (ref. 79).

References

1. Petigura, E. A., Howard, A. W. & Marcy, G. W. Prevalence of Earth-size planets orbiting Sun-like stars. *Proc. Natl Acad. Sci. USA* **110**, 19273–19278 (2013).
2. Fulton, B. J. et al. The California-Kepler Survey. III. A gap in the radius distribution of small planets. *Astron. J.* **154**, 109 (2017).
3. Owen, J. E. & Wu, Y. Kepler planets: a tale of evaporation. *Astrophys. J.* **775**, 105 (2013).
4. Gupta, A. & Schlichting, H. E. Sculpting the valley in the radius distribution of small exoplanets as a by-product of planet formation: the core-powered mass-loss mechanism. *Mon. Not. R. Astron. Soc.* **487**, 24–33 (2019).
5. Kubyshkina, D., Vidotto, A. A., Fossati, L. & Farrell, E. Coupling thermal evolution of planets and hydrodynamic atmospheric escape in MESA. *Mon. Not. R. Astron. Soc.* **499**, 77–88 (2020).
6. Lee, E. J. & Chiang, E. Breeding super-Earths and birthing super-puffs in transitional disks. *Astrophys. J.* **817**, 90 (2016).
7. Lee, E. J., Chiang, E. & Ormel, C. W. Make super-Earths, not Jupiters: accreting nebular gas onto solid cores at 0.1 AU and beyond. *Astrophys. J.* **797**, 95 (2014).
8. Owen, J. E. Constraining the entropy of formation from young transiting planet. *Mon. Not. R. Astron. Soc.* **498**, 5030–5040 (2020).
9. Blunt, S. et al. Overfitting affects the reliability of radial velocity mass estimates of the V1298 Tau planets. *Astron. J.* **166**, 62 (2023).
10. Feinstein, A. D. et al. Flare statistics for young stars from a convolutional neural network analysis of TESS data. *Astron. J.* **160**, 219 (2020).
11. Rackham, B. V., Apai, D. & Giampapa, M. S. The transit light source effect. II. The impact of stellar heterogeneity on transmission spectra of planets orbiting broadly Sun-like stars. *Astron. J.* **157**, 96 (2019).
12. David, T. J. et al. A Neptune-sized transiting planet closely orbiting a 5–10-million-year-old star. *Nature* **534**, 658–661 (2016).
13. Mann, A. W. et al. TESS Hunt for Young and Maturing Exoplanets (THYME). VI. An 11 Myr giant planet transiting a very-low-mass star in Lower Centaurus Crux. *Astron. J.* **163**, 156 (2022).
14. David, T. J. et al. Four newborn planets transiting the young solar analog V1298 Tau. *Astrophys. J. Lett.* **885**, L12 (2019).
15. David, T. J. et al. A warm Jupiter-sized planet transiting the pre-main-sequence star V1298 Tau. *Astron. J.* **158**, 79 (2019).
16. Sikora, J. et al. Updated planetary mass constraints of the young V1298 Tau system using MAROON-X. *Astron. J.* **165**, 250 (2023).
17. Feinstein, A. D. et al. V1298 Tau with TESS: updated ephemerides, radii, and period constraints from a second transit of V1298 Tau e. *Astrophys. J. Lett.* **925**, L2 (2022).
18. Oh, S., Price-Whelan, A. M., Hogg, D. W., Morton, T. D. & Spergel, D. N. *Comoving Stars in Gaia DR1 (Oh+ 2017)* no. J/AJ/153/257 (VizieR Online Data Catalog, 2017).
19. Luhman, K. L. The stellar membership of the Taurus star-forming region. *Astron. J.* **156**, 271 (2018).
20. Suárez Mascareño, A. et al. Rapid contraction of giant planets orbiting the 20-million-year-old star V1298 Tau. *Nat. Astron.* https://doi.org/10.1038/s41550-021-01533-7 (2021).
21. Johnson, M. C. et al. An aligned orbit for the young planet V1298 Tau b. *Astron. J.* **163**, 247 (2022).
22. Maggio, A. et al. New constraints on the future evaporation of the young exoplanets in the V1298 Tau system. *Astrophys. J.* **925**, 172 (2022).
23. Jacobs, B. et al. A strong H^- opacity signal in the near-infrared emission spectrum of the ultra-hot Jupiter KELT-9b. *Astron. Astrophys.* **668**, L1 (2022).
24. Kreidberg, L. et al. Clouds in the atmosphere of the super-Earth exoplanet GJ1214b. *Nature* **505**, 69–72 (2014).
25. Wakeford, H. R. et al. HAT-P-26b: a Neptune-mass exoplanet with a well-constrained heavy element abundance. *Science* **356**, 628–631 (2017).
26. Benneke, B. et al. A sub-Neptune exoplanet with a low-metallicity methane-depleted atmosphere and Mie-scattering clouds. *Nat. Astron.* **3**, 813–821 (2019).
27. Deming, D. et al. Infrared transmission spectroscopy of the exoplanets HD 209458b and XO-1b using the Wide Field Camera-3 on the Hubble Space Telescope. *Astrophys. J.* **774**, 95 (2013).
28. de Wit, J. & Seager, S. Constraining exoplanet mass from transmission spectroscopy. *Science* **342**, 1473–1477 (2013).
29. Jontof-Hutter, D., Lissauer, J. J., Rowe, J. F. & Fabrycky, D. C. Kepler-79's low density planets. *Astrophys. J.* **785**, 15 (2014).
30. Libby-Roberts, J. E. et al. The featureless transmission spectra of two super-puff planets. *Astron. J.* **159**, 57 (2020).
31. Mollière, P. et al. petitRADTRANS. A Python radiative transfer package for exoplanet characterization and retrieval. *Astron. Astrophys.* **627**, A67 (2019).
32. Guillot, T. On the radiative equilibrium of irradiated planetary atmospheres. *Astron. Astrophys.* **520**, A27 (2010).
33. Fortney, J. J. et al. Beyond equilibrium temperature: how the atmosphere/interior connection affects the onset of methane, ammonia, and clouds in warm transiting giant planets. *Astron. J.* **160**, 288 (2020).
34. Thorngren, D. & Fortney, J. J. Connecting giant planet atmosphere and interior modeling: constraints on atmospheric metal enrichment. *Astrophys. J. Lett.* **874**, L31 (2019).
35. Pollack, J. B. et al. Formation of the giant planets by concurrent accretion of solids and gas. *Icarus* **124**, 62–85 (1996).
36. Moses, J. I. et al. Compositional diversity in the atmospheres of hot Neptunes, with application to GJ 436b. *Astrophys. J.* **777**, 34 (2013).
37. Gao, P. et al. The hazy and metal-rich atmosphere of GJ 1214 b constrained by near- and mid-infrared transmission spectroscopy. *Astrophys. J.* **951**, 96 (2023).
38. Lopez, E. D. & Fortney, J. J. Understanding the mass-radius relation for sub-Neptunes: radius as a proxy for composition. *Astrophys. J.* **792**, 1 (2014).
39. Lee, E. J. The boundary between gas-rich and gas-poor planets. *Astrophys. J.* **878**, 36 (2019).
40. Ormel, C. W., Shi, J.-M. & Kuiper, R. Hydrodynamics of embedded planets' first atmospheres – II. A rapid recycling of atmospheric gas. *Mon. Not. R. Astron. Soc.* **447**, 3512–3525 (2015).
41. Atreya, S. K. et al. in *Saturn in the 21st Century* (eds Baines, K. H. et al.) 5–43 (Cambridge Univ. Press, 2018).
42. Welbanks, L. et al. Mass-metallicity trends in transiting exoplanets from atmospheric abundances of H_2O, Na, and K. *Astrophys. J. Lett.* **887**, L20 (2019).
43. Öberg, K. I., Murray-Clay, R. & Bergin, E. A. The effects of snowlines on C/O in planetary atmospheres. *Astrophys. J. Lett.* **743**, L16 (2011).

44. Bitsch, B. et al. Dry or water world? How the water contents of inner sub-Neptunes constrain giant planet formation and the location of the water ice line. *Astron. Astrophys.* **649**, L5 (2021).

45. Fossati, L. et al. Aeronomical constraints to the minimum mass and maximum radius of hot low-mass planets. *Astron. Astrophys.* **598**, A90 (2017).

46. Ormel, C. W., Vazan, A. & Brouwers, M. G. How planets grow by pebble accretion. III. Emergence of an interior composition gradient. *Astron. Astrophys.* **647**, A175 (2021).

47. Wahl, S. M. et al. Comparing Jupiter interior structure models to Juno gravity measurements and the role of a dilute core. *Geophys. Res. Lett.* **44**, 4649–4659 (2017).

48. Vazan, A., Sari, R. & Kessel, R. A new perspective on the interiors of ice-rich planets: ice-rock mixture instead of ice on top of rock. *Astrophys. J.* **926**, 150 (2022).

49. Millholland, S., Petigura, E. & Batygin, K. Tidal inflation reconciles low-density sub-Saturns with core accretion. *Astrophys. J.* **897**, 7 (2020).

50. Hu, R. Photochemistry and spectral characterization of temperate and gas-rich exoplanets. *Astrophys. J.* **921**, 27 (2021).

51. Maggio, A. et al. X-ray and ultraviolet emission of the young planet-hosting star V1298 Tau from coordinated observations with XMM-Newton and Hubble Space Telescope. *Astrophys. J.* **951**, 18 (2023).

52. Arcangeli, J. et al. H-opacity and water dissociation in the dayside atmosphere of the very hot gas giant WASP-18b. *Astrophys. J.* **855**, L30 (2018).

53. Husser, T. O. et al. A new extensive library of PHOENIX stellar atmospheres and synthetic spectra. *Astron. Astrophys.* **553**, A6 (2013).

54. Horne, K. An optimal extraction algorithm for CCD spectroscopy. *Publ. Astron. Soc. Pac.* **98**, 609 (1986).

55. Tsiaras, A. et al. A new approach to analyzing HST spatial scans: the transmission spectrum of HD 209458 b. *Astrophys. J.* **832**, 202 (2016).

56. Zhou, Y., Apai, D., Lew, B. W. P. & Schneider, G. A physical model-based correction for charge traps in the Hubble Space Telescope's wide field camera 3 near-IR detector and its applications to transiting exoplanets and brown dwarfs. *Astrophys. J.* **153**, 243 (2017).

57. Wakeford, H. R., Sing, D. K., Evans, T., Deming, D. & Mandell, A. Marginalizing instrument systematics in HST WFC3 transit light curves. *Astrophys. J.* **819**, 10 (2016).

58. Kreidberg, L. batman: basic transit model calculation in Python. *Publ. Astron. Soc. Pac.* **127**, 1161 (2015).

59. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: the MCMC hammer. *Publ. Astron. Soc. Pac.* **125**, 306 (2013).

60. Fu, G. et al. The Hubble PanCET program: transit and eclipse spectroscopy of the strongly irradiated giant exoplanet WASP-76b. *Astron. J.* **162**, 108 (2021).

61. Stevenson, K. B. & Fowler, J. Analyzing eight years of transiting exoplanet observations using WFC3's spatial scan monitor. Preprint at https://arxiv.org/abs/1910.02073 (2019).

62. Barclay, T. et al. Stellar surface inhomogeneities as a potential source of the atmospheric signal detected in the K2-18b transmission spectrum. *Astron. J.* **162**, 300 (2021).

63. Feinstein, A. D. et al. H-alpha and Ca II infrared triplet variations during a transit of the 23 Myr planet V1298 Tau c. *Astron. J.* **162**, 213 (2021).

64. Koen, C. Exploring a simple method to estimate spot temperatures in weak-lined T Tauri stars. *Mon. Not. R. Astron. Soc.* **463**, 4383–4395 (2016).

65. Baxter, C. et al. Evidence for disequilibrium chemistry from vertical mixing in hot Jupiter atmospheres. A comprehensive survey of transiting close-in gas giant exoplanets with warm-Spitzer/IRAC. *Astron. Astrophys.* **648**, A127 (2021).

66. Madhusudhan, N. Exoplanetary atmospheres: key insights, challenges, and prospects. *Annu. Rev. Astron. Astr.* **57**, 617–663 (2019).

67. Mollière, P., van Boekel, R., Dullemond, C., Henning, T. & Mordasini, C. Model atmospheres of irradiated exoplanets: the influence of stellar parameters, metallicity, and the C/O ratio. *Astrophys. J.* **813**, 47 (2015).

68. Mollière, P. et al. Observing transiting planets with JWST. Prime targets and their synthetic spectral observations. *Astron. Astrophys.* **600**, A10 (2017).

69. Agúndez, M., Parmentier, V., Venot, O., Hersant, F. & Selsis, F. Pseudo 2D chemical model of hot-Jupiter atmospheres: application to HD 209458b and HD 189733b. *Astron. Astrophys.* **564**, A73 (2014).

70. Venot, O. et al. New chemical scheme for giant planet thermochemistry. Update of the methanol chemistry and new reduced chemical scheme. *Astron. Astrophys.* **634**, A78 (2020).

71. Freytag, B., Allard, F., Ludwig, H. G., Homeier, D. & Steffen, M. The role of convection, overshoot, and gravity waves for the transport of dust in M dwarf and brown dwarf atmospheres. *Astron. Astrophys.* **513**, A19 (2010).

72. Moses, J. I. et al. Disequilibrium carbon, oxygen, and nitrogen chemistry in the atmospheres of HD 189733b and HD 209458b. *Astrophys. J.* **737**, 15 (2011).

73. Zhang, X. Atmospheric regimes and trends on exoplanets and brown dwarfs. *Res. Astron. Astrophys.* **20**, 099 (2020).

74. Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A. & Sing, D. Rayleigh scattering in the transit spectrum of HD 189733b. *Astron. Astrophys.* **481**, L83–L86 (2008).

75. Poppenhaeger, K., Ketzer, L. & Mallonn, M. X-ray irradiation and evaporation of the four young planets around V1298 Tau. *Mon. Not. R. Astron. Soc.* **500**, 4560–4572 (2021).

76. Owen, J. E. & Jackson, A. P. Planetary evaporation by UV and X-ray radiation: basic hydrodynamics. *Mon. Not. R. Astron. Soc.* **425**, 2931–2947 (2012).

77. Salz, M., Schneider, P. C., Czesla, S. & Schmitt, J. H. M. M. Energy-limited escape revised. *Astron. Astrophys.* **585**, L2 (2015).

78. Edwards, B. et al. Exploring the ability of Hubble Space Telescope WFC3 G141 to uncover trends in populations of exoplanet atmospheres through a homogeneous transmission survey of 70 gaseous planets. *Astrophys. J.* **269**, 31 (2023).

79. Barat, S. The metal-poor atmosphere of a Neptune or sub-Neptune progenitor. *Zenodo* https://doi.org/10.48550/arXiv.2312.16924 (2024).

80. Akeson, R. L. et al. The NASA Exoplanet Archive: data and tools for exoplanet research. *Publ. Astron. Soc. Pac.* **125**, 989 (2013).

81. Wakeford, H. R. & Dalba, P. A. The exoplanet perspective on future ice giant exploration. *Phil. Trans. R. Soc. A* **378**, 20200054 (2020).

Acknowledgements

Author contributions

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41550-024-02257-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-024-02257-0.

Correspondence and requests for materials should be addressed to Saugata Barat or Jean-Michel Désert.

**Reprints and permissions information** is available at www.nature.com/reprints.

Extended Data Fig. 1 | Observed white light curve of V1298 Tau b using HST/ WFC3 G141. a: Observed white light curve (1.1 − 1.65μm) of a primary transit of V1298 Tau b and best fit model (black solid lines). Red and blue markers show forward and reverse scanned data respectively. b: Best fit planetary transit light curve model c: Systematics function model estimated by dividing the observed light curve in the upper panel by the best fit transit model shown in the second panel, following the prescription of[24]. Black dashed line shows the best fit baseline model. d: Residual from the white light curve fits. The residuals in the third orbit indicate a possible bright spot crossing. The residuals at the end of the seventh orbit rise sharply. This could potentially be due to a flare event. Assuming wavelength independent instrumental and stellar systematics, Z(t) is used to de-trend the spectroscopically binned light curves. For further details, See Methods Light curve analysis).
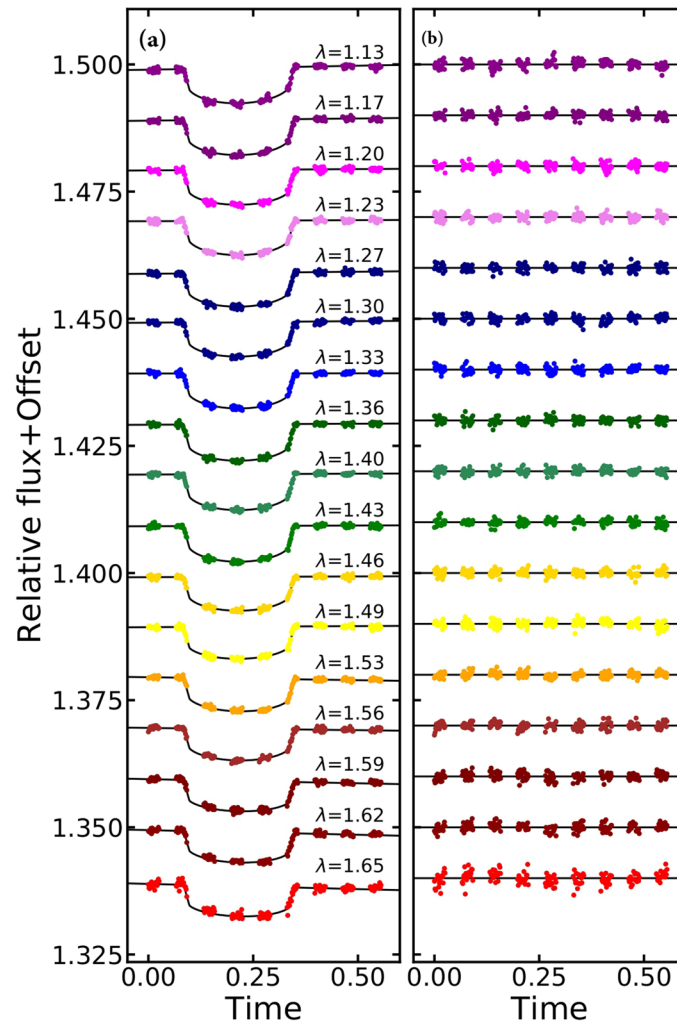
Extended Data Fig. 2 | Systematics de-trended spectroscopic light curves and residuals. a: Systematics de-trended spectroscopic light curves along with the best fit transit model. The light curves have been offset vertically for visual clarity. b: Residuals from the light curve fits in the left panel have been shown. For visual clarity, the residuals have been multiplied by three. The residuals are less than 1.3 times the expected photon noise in all spectroscopic channels (except the last one). The systematics de-trending prescription has been described in Methods Light curve analysis.
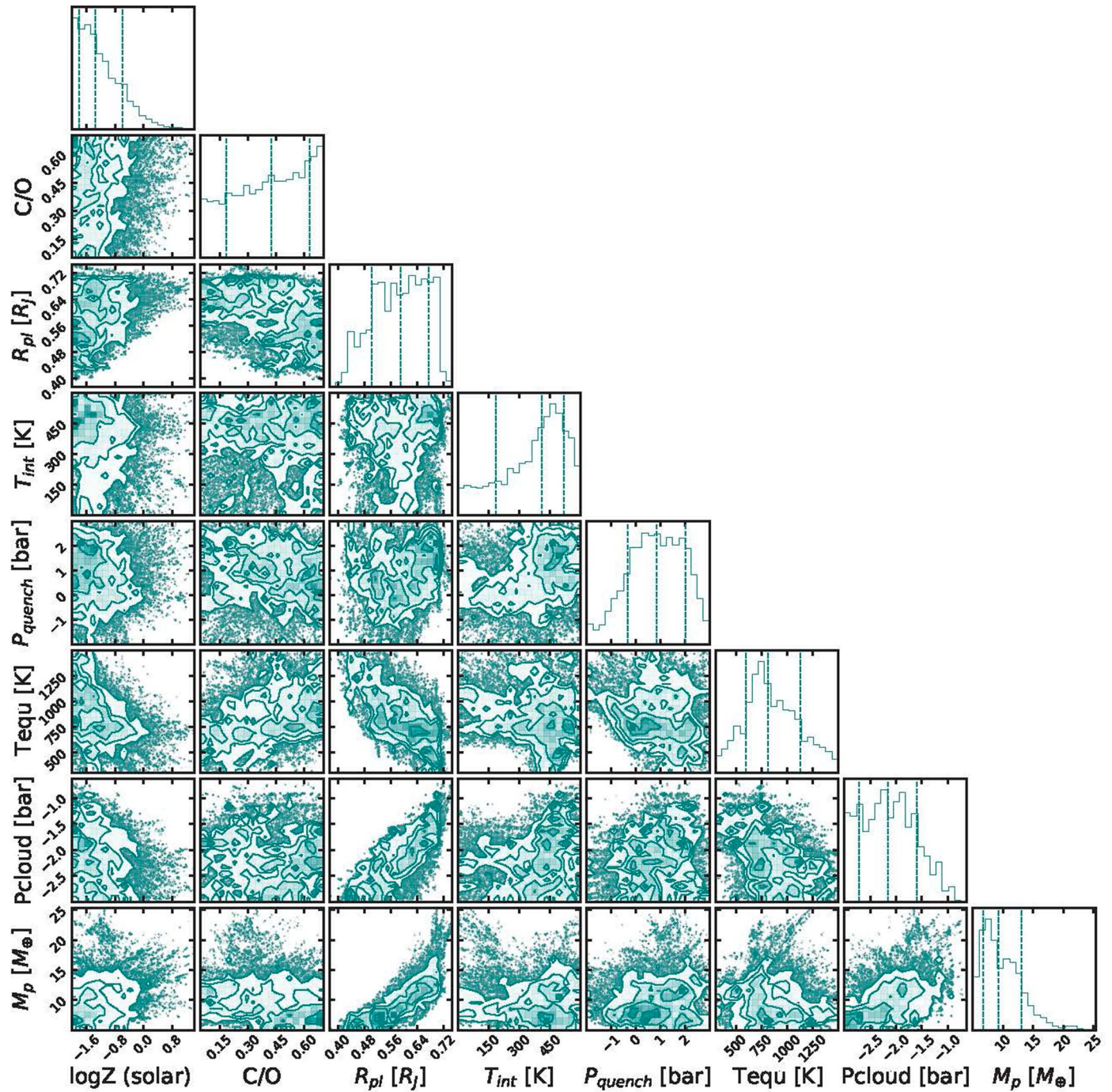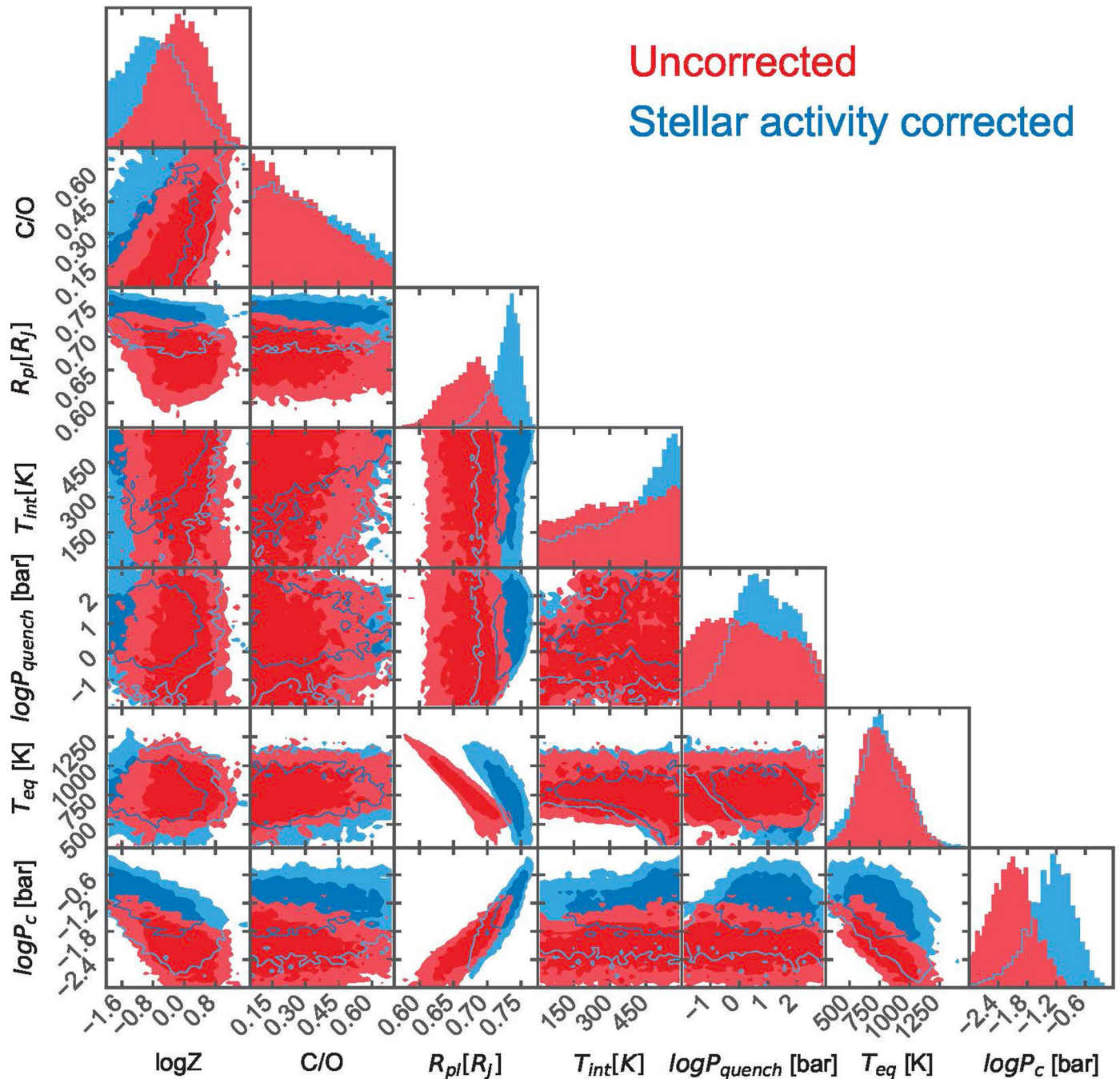
Extended Data Fig. 3 | Posterior distribution of atmospheric retrieval on the uncorrected transmission spectrum of V1298 Tau b, with mass left free. The mass posterior distribution peaks around $10 M_\oplus$ with an atmospheric metallicity 2σ upper limit at solar metallicity. The posterior distribution of the planet mass is in agreement with the upper limit quoted in this paper.
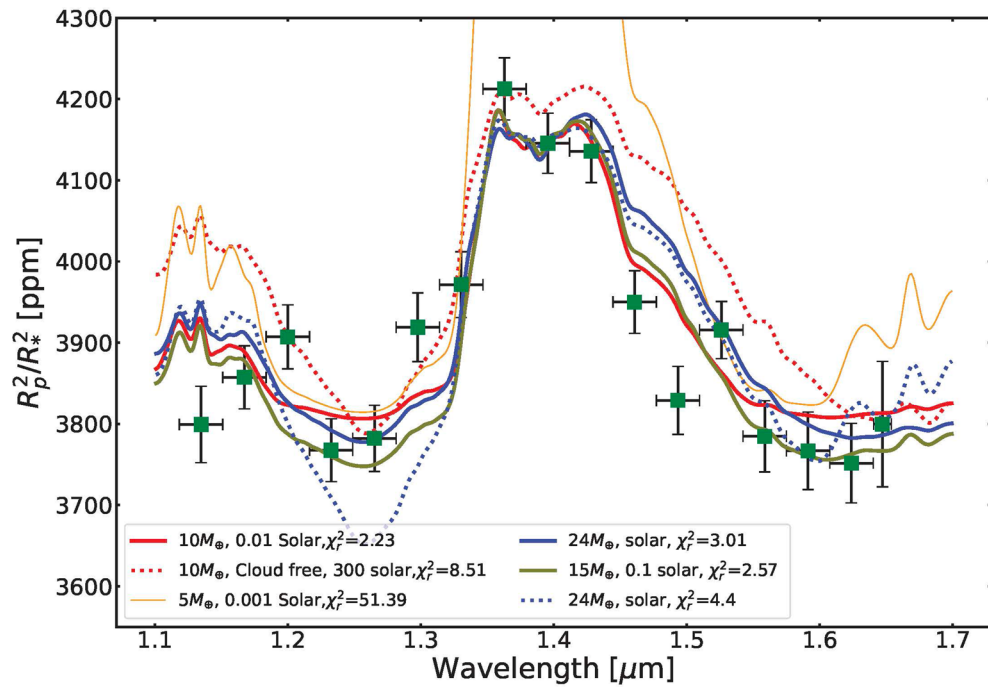
Extended Data Fig. 4 | Comparison between the observed transmission spectrum of V1298 Tau b and retrievals with different planet masses. The transmission spectrum (green points) have been derived using the spectroscopic light curve fitting formalism presented in Methods Light curve Analysis, and the data are shown in Table 1. The red, green and blue models represent retrieved median models for 10, 15 and 24$M_\oplus$ models with a grey cloud opacity respectively. The corresponding dotted line show cloud free models at the same mass. The orange model represents a 5$M_\oplus$ model; our retrievals failed to converge for this case. The observation can be fitted with 24, 15 and 10$M_\oplus$ models. For the lower mass cases our retrievals converged on extremely low atmospheric abundances (0.1-0.01 solar) to fit the water absorption feature. Our retrievals statistically favour cloudy models. The retrievals have been done using PetitRADTRANS[31].

Extended Data Fig. 5 | The posterior distribution from retrieval done on the uncorrected (red) and stellar activity corrected (blue) transmission spectrum of V1298 Tau b assuming a mass of 24$M_\oplus$. We used 1D atmosphere model, Guillot T-P profile[32], equilibrium chemistry with atmospheric quenching (See Methods Atmospheric models for details). We retrieve an atmospheric metallicity consistent with sub-solar/solar for both cases. The retrieved parameters with their 1σ confidence intervals are shown in Table 2. The stellar activity corrected spectrum is shown in comparison with the uncorrected spectrum in Supplementary Fig. 5. In the same figure, the contamination function[11] used for correcting the spectrum is also shown.
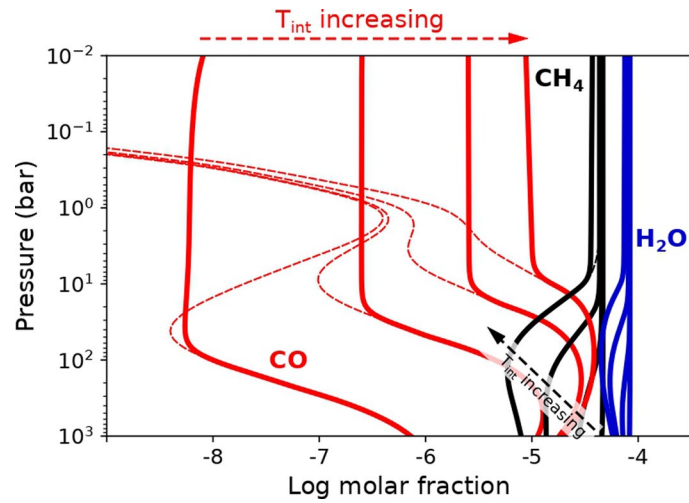
Extended Data Fig. 6 | Chemical abundances of methane, CO, and water in the deep atmosphere, for T_{int} = 100 K, 200 K, 300 K, and 400 K. Chemical abundances have been calculated using a self-consistent framework with petitCODE and a chemical kinetics model in combination with a photochemical network[72,73]. Arrows denote increasing T_{int}. Molecular abundances are shown for a composition in chemical equilibrium (dashed) and for chemical kinetics, that is when vertical quenching is included (solid). We include photochemistry in our models, however, it does not impact the molecular abundances at observable pressures.
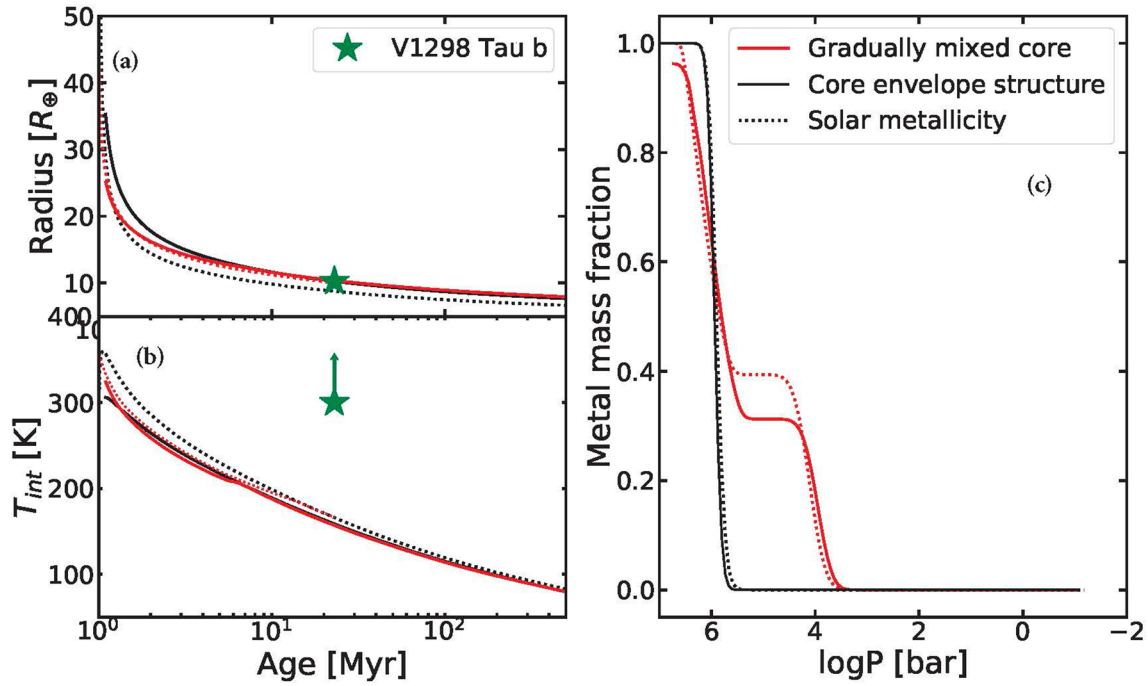
Extended Data Fig. 7 | Theoretical thermal and structural evolution models for different possible interior structures of V1298 Tau b. Evolution of radius (a) and internal temperature (b) of two possible formation-evolution tracks; core-envelope structure (black line) and diluted core structure (red line) of V1298 Tau b. Solid and dotted lines represent simulations with sub-solar (0.1 solar) and solar envelope metallicity. c: the metal distribution in the interior as a function of pressure for the two models at its current age (23 Myr). Models are calculated for in-situ formation of planets with 35-45% H/He (in mass), starting from Hill sphere radius. Evolution model is based on[51]. Both core-envelope models and diluted core models can explain the current size, mass and low metallicity envelope of V1298 Tau. (See Discussion for more details).