

# Evidence for chromium hydride in the atmosphere of hot Jupiter WASP-31b

Marrick Braam<sup>1,2,3</sup>, Floris F. S. van der Tak<sup>1,4</sup>, Katy L. Chubb<sup>5</sup>, and Michiel Min<sup>5</sup>

<sup>1</sup> Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands

<sup>2</sup> School of GeoSciences, University of Edinburgh, King's Buildings, Edinburgh EH9 3FF, UK  
 e-mail: [marrick.braam@ed.ac.uk](mailto:marrick.braam@ed.ac.uk)

<sup>3</sup> Centre for Exoplanet Science, University of Edinburgh, Edinburgh EH9 3FD, UK

<sup>4</sup> SRON Netherlands Institute for Space Research, Landleven 12, 9747 AD Groningen, The Netherlands

<sup>5</sup> SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands

Received 23 September 2020 / Accepted 24 November 2020

## ABSTRACT

**Context.** The characterisation of exoplanet atmospheres has shown a wide diversity of compositions. Hot Jupiters have the appropriate temperatures to host metallic compounds, which should be detectable through transmission spectroscopy.

**Aims.** We aim to detect exotic species in the transmission spectra of hot Jupiters, specifically WASP-31b, by testing a variety of chemical species to explain the spectrum.

**Methods.** We conduct a re-analysis of publicly available transmission data of WASP-31b using the Bayesian retrieval framework TAUREX II. We retrieve various combinations of the opacities of 25 atomic and molecular species to determine the minimum set that is needed to fit the observed spectrum.

**Results.** We report evidence for the spectroscopic signatures of chromium hydride (CrH), H<sub>2</sub>O, and K in WASP-31b. Compared to a flat model without any signatures, a CrH-only model is preferred with a statistical significance of  $\sim 3.9\sigma$ . A model consisting of both CrH and H<sub>2</sub>O is found with  $\sim 2.6$  and  $\sim 3\sigma$  confidence over a CrH-only model and an H<sub>2</sub>O-only model, respectively. Furthermore, weak evidence for the addition of K is found at  $\sim 2.2\sigma$  over the H<sub>2</sub>O+CrH model, although the fidelity of the data point associated with this signature was questioned in earlier studies. Finally, the inclusion of collision-induced absorption and a Rayleigh scattering slope (indicating the presence of aerosols) is found with  $\sim 3.5\sigma$  confidence over the flat model. This analysis presents the first evidence for signatures of CrH in a hot Jupiter atmosphere. At a retrieved temperature of  $1481^{+264}_{-355}$  K, the atmosphere of WASP-31b is hot enough to host gaseous Cr-bearing species, and the retrieved abundances agree well with predictions from thermal equilibrium chemistry. Furthermore, the retrieved abundance of CrH agrees with the abundance in an L-type brown dwarf atmosphere. However, additional retrievals using VLT FORS2 data lead to a non-detection of CrH. Future observations with *James Webb* Space Telescope have the potential to confirm the detection and/or discover other CrH features.

**Key words.** planets and satellites: atmospheres – planets and satellites: individual: WASP-31b – techniques: spectroscopic

## 1. Introduction

Soon after the discoveries of the first exoplanets (Wolszczan & Frail 1992; Mayor & Queloz 1995), their atmospheres became a curiosity (e.g. Seager & Sasselov 2000). Nowadays, the confirmed number of exoplanets has exceeded 4000<sup>1</sup> and this number is expected to increase significantly over the coming years. Even more remarkable than the large number of discoveries itself is the wide parameter space in which these planets are being found: Equilibrium temperatures range from  $\sim 100$ – $4050$  K and masses and radii are continuously found within ranges of  $0.1$ – $10^4 M_{\oplus}$  and  $0.3$ – $25 R_{\oplus}$ . Naturally, an enormous diversity in exoplanet atmospheres can be expected.

Currently, the main method for characterising these exoplanet atmospheres is through transmission spectroscopy (e.g. Crossfield 2015). A transmission spectrum measures the dip in the stellar light when a planet transits in front of its host star. If the planet has an atmosphere, the opacity and, consequently, the apparent planet size change as a function of wavelength. The atmospheric composition and physical structure can be inferred from these variations with wavelength (Seager & Sasselov 2000). Using the Space Telescope Imaging Spectrograph (STIS) on the *Hubble* Space Telescope (HST), the first

detection of an exoplanet atmosphere was the discovery of the sodium (Na) doublet during a transit of the hot Jupiter HD 209458b (Charbonneau et al. 2002). Since then, evidence for the features of a variety of other chemical species has been reported, such as H<sub>2</sub>O, CH<sub>4</sub>, CO, CO<sub>2</sub>, and K (see Madhusudhan (2019) for an overview). Furthermore, the existence of metallic compounds such as TiO, VO, and AlO has been found on several planets (e.g. Sedaghati et al. 2017; Evans et al. 2017; Chubb et al. 2020a).

The search for absorption signatures of metallic compounds is inspired by their detections in brown dwarfs (e.g. Kirkpatrick et al. 1999a; Kirkpatrick 2005; Lodders & Fegley 2006), and they are also predicted to be important species in the temperature ranges of hot exoplanets (e.g. Burrows & Sharp 1999; Woitke et al. 2018). Amongst these metallic compounds, chromium hydride (CrH) and iron hydride (FeH) are relevant in the brown dwarf classification scheme, notably in specifying the transition from L to T dwarfs (Kirkpatrick 2005). The detections of atomic metal species in ultra-hot Jupiters, such as Cr I, Fe I, Mg I, Na I, Ti I, and V I (Hoeijmakers et al. 2018, 2019; Ben-Yami et al. 2020), suggest that the hydrides CrH and FeH can also be expected in the atmospheres of hot exoplanets. Tentative detections of FeH have been reported for four planets: WASP-62b (Skaf et al. 2020), WASP-79b (Sotzen et al. 2020;

<sup>1</sup> Based on data in the [NASA Exoplanet Archive](https://exoplanetarchive.nasa.gov/)

**Table 1.** Planetary parameters.

$R_p(R_J)$	$M_p(M_J)$	SMA(AU)	$T_{eq}(K)^{(a)}$	$T_{range}^{(a)}$	$R_*(R_\odot)$	$T_*(K)$	Reference
$1.549 \pm 0.050$	$0.478 \pm 0.029$	$0.0466 \pm 0.0004$	1393	869–1882	$1.252 \pm 0.033$	$6302 \pm 102$	Anderson et al. (2011)

**Notes.** <sup>(a)</sup>Equilibrium temperatures were calculated under varying assumptions for the Bond albedo and redistribution factor (see text).

Skaf et al. 2020), WASP-121b (Evans et al. 2016), and WASP-127b (Skaf et al. 2020), whereas Kesseli et al. (2020) did not find statistically significant detections for 12 planets using high dispersion transmission spectroscopy. Furthermore, evidence for the presence of metal hydrides in the exo-Neptune HAT-P-26b was found by MacDonald & Madhusudhan (2019), who identified three possible candidates to explain these features in the optical part of the transmission spectrum: TiH, CrH, or ScH. Found as part of the Wide Angle Search for Planets (Pollacco et al. 2006), WASP-31b is thought to be in the right temperature range to host metal hydrides.

WASP-31b, which is in orbit around an F-type star, was discovered by Anderson et al. (2011). The planet has a mass of  $0.478 M_J$  and a radius of  $1.549 R_J$ , making it one of the lowest density exoplanets known to date. Orbiting at a distance of 0.047 AU from its host star, it has an equilibrium temperature of 1393 K (assuming Jupiter’s Bond albedo of 0.34). Its low density (surface gravity) and high temperature lead to a large scale height, making WASP-31b a suitable candidate for atmospheric characterisation using transmission spectroscopy. Its host star has an effective temperature of  $6300 \pm 100 K$  and a metallicity of  $-0.20 \pm 0.09$  dex. The system age is estimated to be  $1^{+3}_{-0.5}$  Gyr (Anderson et al. 2011). Using optical to mid-infrared transmission spectra to probe the atmosphere of WASP-31b, Sing et al. (2015) found a strong potassium (K) feature as well as evidence for the presence of aerosols, both in the form of clouds (grey scatter) and hazes (Rayleigh scatter). Evidence for a grey cloud deck was also obtained by Barstow et al. (2017), whereas other comparative studies found some weak  $H_2O$  (Pinhas et al. 2019; Welbanks et al. 2019) and  $NH_3$  features (MacDonald & Madhusudhan 2017; Min et al. 2020). The existence of the K signature has been called into question by recent observations using the FOCAL Reducer and low dispersion Spectrograph 2 (FORSS2) and the Ultraviolet and Visual Echelle Spectrograph (UVES) on the Very Large Telescope (VLT; Gibson et al. 2017, 2019) and the Inamori-Magellan Areal Camera and Spectrograph (IMACS) on the Magellan Baade Telescope (McGruder et al. 2020).

In this study, we conduct a re-analysis of the publicly available transmission spectrum of WASP-31b using the TAUREX retrieval framework (Waldmann et al. 2015). In Sect. 2, we describe the observations that were used in this analysis and provide the details of the retrieval setup. The retrieval results are presented in Sect. 3. In Sect. 4, we compare our findings with earlier detections and discuss the physical implications before providing the conclusions in Sect. 5.

## 2. Methodology

### 2.1. Observations

The optical and near-infrared transit light curves of WASP-31b were observed using HST and then analysed by Sing et al. (2015). Transits were observed using STIS with the G430L and G750L gratings, providing spectral coverage from 0.29 to  $1.027 \mu m$  at a resolution of 530–1040. These were supplemented

by observations from 1.1 to  $1.7 \mu m$  at  $R \sim 70$  using the G141 grism of the Wide Field Camera 3 (WFC3). Sing et al. (2015) combined their observations with photometric measurements in the 3.6 and  $4.5 \mu m$  channels obtained using *Spitzer*’s Infrared Array Camera (IRAC)<sup>2</sup>.

From the planetary parameters in Table 1, it can be seen that WASP-31b is larger than Jupiter and orbits close to its host star. Furthermore, with only half of Jupiter’s mass, the planet is one of the lowest density planets known. We calculated the equilibrium temperature assuming Jupiter’s Bond albedo of 0.34 and a redistribution factor  $f=1$  defining isotropic re-emission; for the lower and upper boundaries, we assumed  $A = 0.9$ ,  $f = 1$  and  $A = 0.12$ ,  $f = 0.5$ , respectively.

### 2.2. TauREx II

The observed transmission spectra are a complex function of many underlying parameters, and acquiring information about these parameters is known as the inverse, or retrieval, problem. We need to determine, given the planetary spectrum that is observed, what the most likely composition and state of the planetary atmosphere are. The retrieval was conducted using TAUREX II (Waldmann et al. 2015), which is a Bayesian retrieval framework based on a forward model that computes 1D atmospheric radiative transfer (Hollis et al. 2013). The propagation of radiation through an atmosphere is strongly dependent on pressure-temperature structure and composition. TAUREX maps the correlations between atmospheric parameters and provides statistical estimates on their values.

In planetary atmospheres, there are two forms of interaction between stellar radiation and the gases in an atmosphere: scattering and absorption of radiation. For molecular opacities, TAUREX relies on the ExoMol (Tennyson et al. 2016), HITEMP (Rothman et al. 2010), and MoLLIST (Bernath 2020) databases. These databases contain line lists of many molecular species, providing their energy levels and transition probabilities, up to high temperatures. This allows us to compute the wavelength-dependent absorption of a particular species as a function of temperature and pressure. The line lists that are used in our analysis are shown in Table 2. There is a threefold motivation for the choice of chemical species. Firstly, thermal equilibrium chemistry can predict the presence and expected abundances of species as a function of temperature, pressure, and elemental composition (Woitke et al. 2018). Assuming initial solar composition, this predicts the main element-bearing species at the temperatures of hot Jupiters to be, for example,  $H_2O$  and CO for oxygen, CO,  $CO_2$ , and  $CH_4$  for carbon, and TiO for titanium. Secondly, a possible detection requires the species to have signatures in the observed spectral regime. Knowledge of the absorption signatures of metal hydrides and oxides is informed by their detections in brown dwarfs (e.g.

<sup>2</sup> See [https://pages.jh.edu/~dsing3/David\\_Sing/Spectral\\_Library.html](https://pages.jh.edu/~dsing3/David_Sing/Spectral_Library.html) and <https://stellarplanet.org/science/exoplanet-transmission-spectra/>

**Table 2.** Atomic and molecular data used in this analysis.

Molecule	Wavelength range	Number of lines	Database/reference
AlH	0.37–100 $\mu\text{m}$	36 000	ExoMol: Yurchenko et al. (2018)
AlO	0.29–100 $\mu\text{m}$	4 945 580	ExoMol: Patrascu et al. (2015)
C <sub>2</sub> H <sub>2</sub>	1.00–100 $\mu\text{m}$	4 347 381 911	ExoMol: Chubb et al. (2020b)
C <sub>2</sub> H <sub>4</sub>	1.41–100 $\mu\text{m}$	49 841 085 051	ExoMol: Mant et al. (2018)
CaH	0.45–100 $\mu\text{m}$	19 095	MoLLIST: Li et al. (2012); Bernath (2020)
CH <sub>4</sub>	0.83–100 $\mu\text{m}$	34 170 582 862	ExoMol: Yurchenko et al. (2017)
CN	0.23–100 $\mu\text{m}$	195,120	MoLLIST: Brooke et al. (2014)
CO	0.45–100 $\mu\text{m}$	752 976	Li et al. (2015)
CO <sub>2</sub>	1.04–100 $\mu\text{m}$	11 167 618	HITEMP: Rothman et al. (2010)
CP	0.67–100 $\mu\text{m}$	28 752	MoLLIST: Ram et al. (2014)
CrH	0.67–100 $\mu\text{m}$	13 824	MoLLIST: Burrows et al. (2002); Bernath (2020)
FeH	0.67–100 $\mu\text{m}$	93 040	MoLLIST: Wende et al. (2010)
H <sub>2</sub> CO	0.99–100 $\mu\text{m}$	12 688 112 669	ExoMol: Al-Refaie et al. (2015)
H <sub>2</sub> O	0.24–100 $\mu\text{m}$	5 745 071 340	ExoMol: Polyansky et al. (2018)
HCN	0.56–100 $\mu\text{m}$	34 418 408	ExoMol: Barber et al. (2014)
K	0.29–100 $\mu\text{m}$	186	NIST: Kramida et al. (2013); Allard et al. (2016)
MgH	0.34–100 $\mu\text{m}$	30 896	MoLLIST: Gharib-Nezhad et al. (2013)
MgO	0.27–100 $\mu\text{m}$	72 833 173	ExoMol: Li et al. (2019)
Na	0.24–100 $\mu\text{m}$	523	NIST: Kramida et al. (2013); Allard et al. (2019)
NH <sub>3</sub>	0.43–100 $\mu\text{m}$	16 941 637 250	ExoMol: Coles et al. (2019)
OH	0.23–100 $\mu\text{m}$	54 276	MoLLIST: Yousefi et al. (2018)
ScH	0.63–100 $\mu\text{m}$	1 152 826	LYT: Lodi et al. (2015)
TiH	0.42–100 $\mu\text{m}$	199 072	MoLLIST: Burrows et al. (2005)
TiO	0.33–100 $\mu\text{m}$	59 324 532	ExoMol: McKemmish et al. (2019)
VO	0.29–100 $\mu\text{m}$	277 131 624	ExoMol: McKemmish et al. (2016)

Kirkpatrick et al. 1999a; Lodders & Fegley 2006; Sharp & Burrows 2007), whereas the prominent signatures of Na (near 0.59  $\mu\text{m}$ ) and K (near 0.77  $\mu\text{m}$ ) are identifiable in the optical regime (e.g. Charbonneau et al. 2002; Welbanks et al. 2019). Most of the nitrogen is expected to be present in N<sub>2</sub>. Being a homonuclear diatomic molecule, N<sub>2</sub> has no prominent signatures in the infrared. NH<sub>3</sub>, especially prominent at cool temperatures, is the next main nitrogen-bearing species and does have signatures, motivating its inclusion. Lastly, species such as C<sub>2</sub>H<sub>2</sub>, HCN, and OH are expected to be related to photochemical processes (e.g. Line et al. 2010; Kawashima & Ikoma 2019).

The data in Table 2 were converted into cross-sections and k-tables by Chubb et al. (2021) in order to feed them into TAUREX as part of the ExoMolOP database<sup>3</sup>. In this study, we have used the k-tables with  $R = \frac{\Delta\lambda}{\lambda} = 300$ . A k-table provides the absorption coefficient as a function of wavelength for a certain temperature and pressure.

Furthermore, TAUREX includes the continuum opacity caused by collision-induced absorption (CIA) of H<sub>2</sub>–H<sub>2</sub> and H<sub>2</sub>–He pairs and a parametrisation for the opacity caused by particle scattering. This represents the interaction between radiation and aerosols, thus quantifying the influence of clouds and hazes. For atmospheric particles that are small relative to the wavelength of the incoming light, there is a strong  $\lambda^{-4}$  dependence of Rayleigh scattering. The opacity due to Rayleigh scattering is based on pre-computed cross-sections (Hollis et al. 2013). Besides that, an optically thick grey cloud cover would lead to a flat opacity as a function of  $\lambda$  and is modelled via the cloud top-pressure  $P_{\text{cl}}$ . TAUREX also contains a more complex cloud model that

parametrises the opacity due to the scattering of light by spherical particles, following the Mie theory (Lee et al. 2013). This parametrisation was tried but not found to be significant.

### 2.3. Retrieval

TAUREX searches the multi-dimensional parameter space for solutions through the MULTINEST algorithm (Feroz & Hobson 2008; Feroz et al. 2009, 2019). As an output, MULTINEST provides the global log-evidence, or simply the Bayesian evidence, which tests the adequacy of the model itself and can be used to compare models of varying complexity. In this comparison, Occam’s razor is applied: Adding a factor of complexity to an atmospheric model is only appropriate when this inclusion gives a significantly better fit to the data. When comparing two models,  $\mathcal{M}_2$  having an extra atmospheric parameter and thus more complexity than  $\mathcal{M}_1$ , their Bayesian evidence can be used to calculate the ratio of the model probabilities, or the Bayes factor (Kass & Raftery 1995; Waldmann et al. 2015),

$$\mathcal{B}_{21} = \frac{E_2}{E_1}, \quad (1)$$

or to define the detection significance (DS),

$$\text{DS} = \ln(\mathcal{B}_{21}) = \ln(E_2) - \ln(E_1). \quad (2)$$

Table 3 shows the empirically calibrated Jeffreys’ scale from Trotta (2008), which we used to quantify the preference for an additional atmospheric parameter: a DS greater than one provides evidence in favour of the more complex model. We refer to this as the “detection significance” since more complexity is usually represented by the addition of a particular chemical species to our model.

<sup>3</sup> <http://www.exomol.com/data/data-types/opacity/>



**Table 3.** Empirically calibrated Jeffreys’ scale (Jeffreys 1998) with translation to frequentist values specifying the odds in favour of the more complex model, adapted from Trotta (2008).

DS	$p$ -value	$\sigma$	Category
1.0	0.04	2.1	“Weak” at best
2.5	0.006	2.7	“Moderate” at best
5.0	0.0003	3.6	“Strong” at best
11.0	$6 \times 10^{-7}$	5.0	“Very strong”

#### 2.4. General setup

The atmospheric models consist of 100 isothermal layers with pressures ranging from  $10^6$  to  $10^{-5}$  Pa. We assumed a hydrogen-dominated atmosphere with a Jupiter-like  $\text{He}/\text{H}_2 = 0.157$  and used the prior values and temperature boundaries from Table 1. The planetary radius was fitted within ranges of  $0.1 R_p$  around the prior value, and the retrieved presence of a grey cloud cover was allowed in the full pressure range. Furthermore, the chemical abundances were retrieved with volume mixing ratios (VMRs) or abundances between  $10^{-10}$  and  $10^{-1}$ . From the retrieved abundances, the atmospheric molecular weight was then calculated. In the end, up to 28 free parameters can thus be retrieved in the procedure. However, given the limited number of data points and spectral coverage, only a small fraction of these parameters will statistically be required. Together with the feasibility of quantifying the importance of individual molecules, this is one of the main reasons for a bottom-up approach.

In this bottom-up approach, the retrieval was first performed assuming the simplest atmospheric forward model consisting of three free parameters ( $R_p$ ,  $T$ , and  $P_{\text{cl}}$ ), which is equivalent to an atmosphere completely lacking spectral signatures. Afterwards, retrievals were done by adding parameters to the atmospheric model in the form of the abundance of a chemical species. The first stage is to compare the models with a single chemical species to the flat model ( $R_p$ ,  $T$ , and  $P_{\text{cl}}$  only), using the DS (see Eq. (2)). As opposed to the flat model, the opacities caused by Rayleigh scattering and CIA are from now on also included. If the addition of a chemical species leads to an improved fit to the data, this results in stronger evidence, and the significance of such a detection is specified by Jeffreys’ scale in Table 3. As a second stage, we tested the inclusion of  $\text{H}_2\text{O}$  plus another species, mainly because  $\text{H}_2\text{O}$  has a prominent absorption feature in the relatively well-covered near-infrared (e.g. Sing et al. 2016; Tsiraras et al. 2018). Evidence levels from these models can be compared to the models containing a single species as well as to the flat model. Lastly, the model with the strongest evidence was expanded by adding the absorption features of the alkali metals Na and K in order to quantify their possible presence.

### 3. Results

Following the bottom-up approach for WASP-31b, its spectrum was retrieved assuming 53 different atmospheric models. The resulting evidence for each of these models can be seen in Fig. 1, and the exact values for the evidence and DSs (see Eq. (2)) are shown in Table 4. Starting from the lower left of Fig. 1, it can be seen that the flat model without any signatures (only  $R_p$ ,  $T$ , and  $P_{\text{cl}}$ ; represented by an orange dot) leads to a Bayesian evidence level of 399.69. The inclusion of Rayleigh scattering and CIA is labelled as “Ray+CIA” and is detected with a confidence

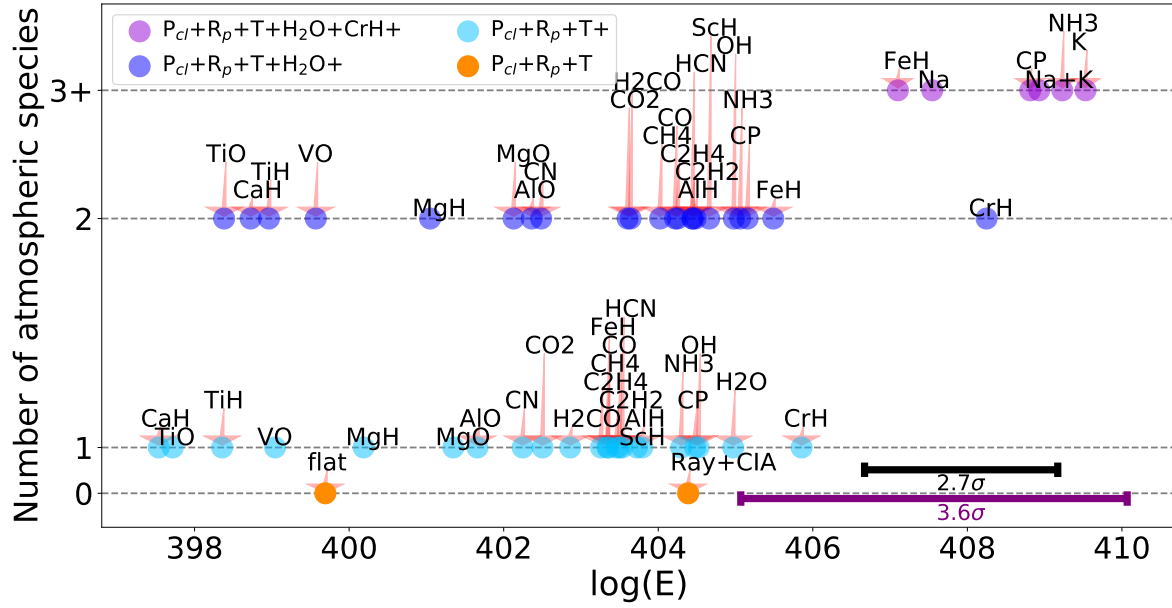
level of  $\sim 3.5\sigma$  over the flat model. Except for the “flat model”, all atmospheric models contain the opacity caused by Rayleigh scattering and CIA.

An increase in the complexity, by adding a single atmospheric species on top of the opacity from Rayleigh scattering and CIA, gives the models that are shown as the cyan dots and which are specified by their accompanying labels. It can be seen that the addition of only a small selection of chemical species leads to an increase in evidence levels. The decrease in evidence seen for several species (e.g. CaH and TiO) is caused by our choice of the lower boundary for the abundances in the retrievals. For example, TiO at  $\log(X_{\text{TiO}}) = -10$  would still cause absorption features. A negative DS then means that the abundance of the added species is lower than the retrieval boundary, signifying the abundance at which features are no longer seen.

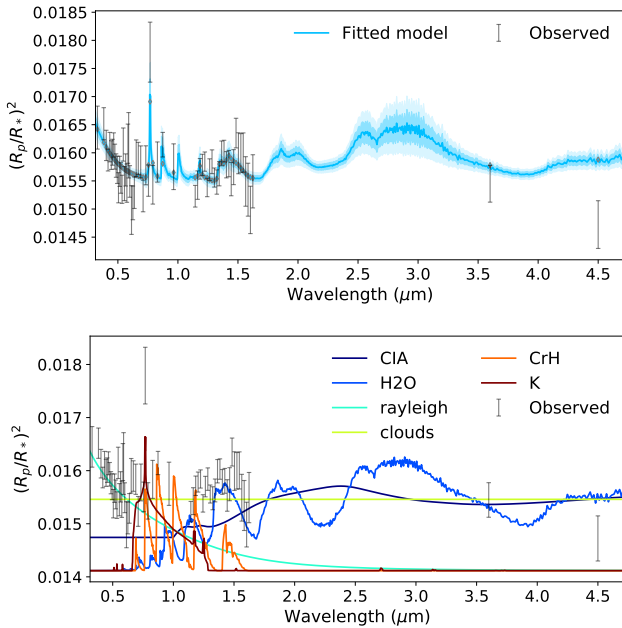
At this stage, the strongest preference is found for the inclusion of either CrH or  $\text{H}_2\text{O}$ , with DSs over the flat model of 6.16 and 5.28, respectively. Following Jeffreys’ scale (see Table 3), this corresponds to confidence levels of  $\sim 3.9\sigma$  and  $\sim 3.7\sigma$ , respectively. Compared to the model containing Rayleigh scattering and CIA, the CrH signature in this single-molecule model is detected at  $\sim 2.3\sigma$ . Ascending one stage in complexity, the blue dots show the models consisting of  $\text{H}_2\text{O}$  and another species. It can be seen that the combined inclusion of both  $\text{H}_2\text{O}$  and CrH is preferred, with a DS of 8.56 (or  $\sim 4.4\sigma$  confidence) over the flat model, or 3.86 ( $\sim 3.2\sigma$ ) over a model of only Rayleigh scattering and CIA. Compared to a CrH- or  $\text{H}_2\text{O}$ -only model, a model containing both species corresponds to confidence levels of  $\sim 2.6\sigma$  and  $\sim 3.0\sigma$ , respectively.

The final stage is given by the violet dots and represents the addition of further atmospheric species to the best models of previous stages, in this case the one containing  $\text{H}_2\text{O}$  and CrH. Regarding the alkali metals, the inclusion of K in the WASP-31b atmospheric model is preferred, with a DS of 1.28 as compared to the  $\text{H}_2\text{O}$ +CrH model, whereas the model that includes both Na and K leads to a DS of 0.68. Hence, statistical evidence is only found for the absorption signature of K. The K detection corresponds to weak evidence at a confidence level of  $\sim 2.2\sigma$  over the  $\text{H}_2\text{O}$ +CrH model. The fact that the detection of K is mainly based on a single strong absorption peak can explain this weak evidence since the signature is covered by just a single data point. Naturally, providing a better fit to only one out of 63 data points may correspond to such a weak increase in Bayesian evidence. In this final stage, we also examined the individual additions of FeH, CP, and  $\text{NH}_3$  to the  $\text{H}_2\text{O}$ +CrH model since these species correspond to the highest evidence levels in the previous stage, which included two atmospheric species. Compared to the  $\text{H}_2\text{O}$ +CrH model, the addition of  $\text{NH}_3$  or CP leads to a small DS (0.98 and 0.56, respectively) and the addition of FeH leads to a negative DS of  $-1.15$ . None of these are significant according to the Jeffreys’ scale.

We conclude that out of the models fitted in this study, the spectrum of WASP-31b is best represented by a model that includes  $\text{H}_2\text{O}$ , CrH, and K in addition to  $\text{H}_2$ , He, a grey cloud deck, and Rayleigh scattering. This atmospheric model and the observed transmission spectrum of WASP-31b are shown in the upper panel of Fig. 2. The lower panel shows the individual contributions of the atmospheric constituents to the opacity. The signatures of  $\text{H}_2\text{O}$  in the near-infrared (around 1.0, 1.2, and  $1.4 \mu\text{m}$ ; in blue) and K in the visible (around  $0.77 \mu\text{m}$ ; dark red) are easily recognised, whereas the inclusion of CrH leads to the six absorption signatures between 0.7 and  $1.5 \mu\text{m}$ , as shown by the orange line. On top of that, the navy line represents the continuum opacity provided by CIA, and the presence of aerosols



**Fig. 1.** Model comparison for WASP-31b using Bayesian evidence for different atmospheric models. The flat model is shown as the orange dot, whereas cyan dots indicate higher complexity in the form of a chemical species, as labelled. One stage higher, blue dots represent a model that includes  $\text{H}_2\text{O}$  and an additional parameter. For the final stage, the strongest evidence model from lower complexities is complemented by K and potentially Na, as shown by the violet dots. In this final stage, we also tested the inclusion of some of the more likely species of lower levels. The horizontal scale bars indicate the statistical preference for a more complex model, based on Table 3.



**Fig. 2.** TAUREX retrieval results for WASP-31b, with the transmission spectrum and the best-fitting atmospheric model (top) and the individual contributions of each molecule to the opacity (bottom). The vertical error bars indicate the observed transit depths, and the different shadings in the upper panel represent 1 and  $2\sigma$  regions.

results in two distinct signatures: the grey scattering opacity of a low altitude cloud deck and Rayleigh scattering at short wavelengths due to haze.

In Fig. 2, a discrepancy can be seen between the transit depth resulting from our atmospheric model and the measured transit depth at  $4.5\mu\text{m}$  as observed by *Spitzer*. Uncertainties exist in cross-calibrating measurements by different instruments, and

the usefulness of *Spitzer*'s broadband photometry in inferring atmospheric compositions has been called into question (see e.g. Hansen et al. 2014). To test our findings, we conducted the same analysis for a spectrum that excludes the *Spitzer* measurements. The spectrum and its best-fitting atmospheric model can be seen in Fig. 3. Excluding the *Spitzer* measurements leads to a significant preference for the  $\text{CrH}+\text{H}_2\text{O}$  model, with confidence levels of  $\sim 4.8\sigma$  over a flat model and  $\sim 3.9\sigma$  over a model containing Rayleigh scattering and CIA. Besides that, the retrieved values are consistent with our earlier findings within  $1\sigma$ . Therefore, we conclude that removing the *Spitzer* measurements does not change our results significantly, illustrating that mid-infrared coverage is not required to detect CrH. The discrepancy between the data and our model at these wavelengths hints at the influence of species that show spectral activity in these regions (such as CO and  $\text{CO}_2$ ).

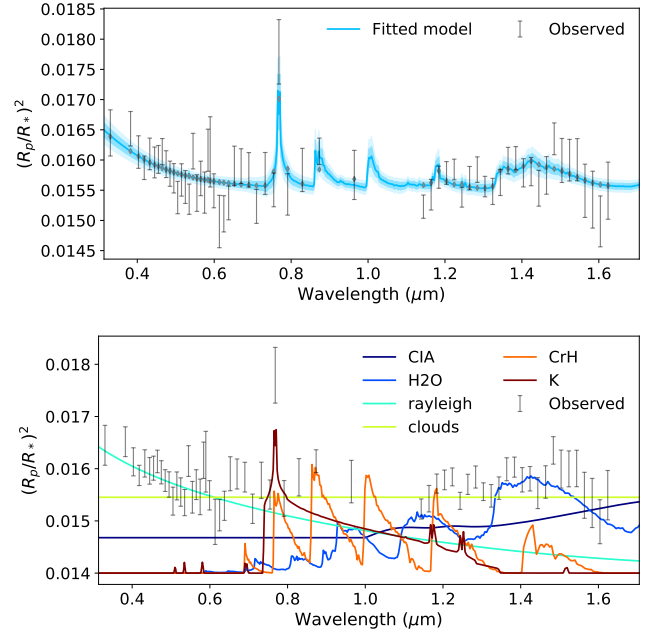
## 4. Discussion

Before discussing the retrieved parameters, it is important to emphasise that the resulting atmospheric parameters are based on the models that turned out to be the best fit to the spectral data of WASP-31b. A bottom-up approach is valuable in inferring the presence of chemical species in an atmosphere but may lead to biases in the derived constraints on retrieved parameters. Excluding a particular chemical species from the atmospheric model means that the spectroscopic signature of the species has not been detected on the basis of statistics. This does not necessarily mean that a chemical species is completely absent from the atmosphere that is probed. Instead, the signatures of a species can fall outside of the observed spectral range, be too weak to be detected, or be affected by an overlap with other spectral signatures. The omission of a particular species can then influence the retrieval outcomes since its signatures (even if they are statistically insignificant) have to be explained by the absorption of other species. This may result in unreasonably

**Table 4.** Resulting Bayesian evidence levels and DSs for WASP-31b.

Model parameters	$\log(E)$	DS
Flat model	399.69	
Rayleigh + CIA	404.39	4.70
Compared to flat model		
CaH	397.54	-2.16
TiO	397.72	-1.97
TiH	398.36	-1.34
VO	399.04	-0.65
MgH	400.19	0.50
MgO	401.35	1.66
AlO	401.66	1.97
CN	402.25	2.55
CO <sub>2</sub>	402.50	2.81
H <sub>2</sub> CO	402.86	3.17
C <sub>2</sub> H <sub>4</sub>	403.26	3.57
FeH	403.34	3.65
CH <sub>4</sub>	403.35	3.66
C <sub>2</sub> H <sub>2</sub>	403.46	3.77
CO	403.50	3.80
HCN	403.53	3.84
ScH	403.72	4.03
AlH	403.79	4.10
NH <sub>3</sub>	404.30	4.60
CP	404.48	4.78
OH	404.52	4.83
H <sub>2</sub> O	404.97	5.28
CrH	405.86	6.16
Compared to H <sub>2</sub> O-only model		
H <sub>2</sub> O + TiO	398.38	-6.59
H <sub>2</sub> O + CaH	398.73	-6.24
H <sub>2</sub> O + TiH	398.96	-6.00
H <sub>2</sub> O + VO	399.57	-5.40
H <sub>2</sub> O + MgH	401.05	-3.92
H <sub>2</sub> O + MgO	402.13	-2.84
H <sub>2</sub> O + AlO	402.36	-2.61
H <sub>2</sub> O + CN	402.48	-2.48
H <sub>2</sub> O + CO <sub>2</sub>	403.60	-1.37
H <sub>2</sub> O + H <sub>2</sub> CO	403.64	-1.33
H <sub>2</sub> O + CH <sub>4</sub>	404.02	-0.95
H <sub>2</sub> O + CO	404.22	-0.75
H <sub>2</sub> O + C <sub>2</sub> H <sub>4</sub>	404.25	-0.72
H <sub>2</sub> O + HCN	404.44	-0.53
H <sub>2</sub> O + C <sub>2</sub> H <sub>2</sub>	404.44	-0.52
H <sub>2</sub> O + AlH	404.48	-0.49
H <sub>2</sub> O + ScH	404.66	-0.31
H <sub>2</sub> O + OH	404.98	0.01
H <sub>2</sub> O + NH <sub>3</sub>	405.06	0.09
H <sub>2</sub> O + CP	405.16	0.19
H <sub>2</sub> O + FeH	405.49	0.52
H <sub>2</sub> O + CrH	408.25	3.28
Compared to H <sub>2</sub> O+CrH model		
H <sub>2</sub> O + CrH + FeH	407.10	-1.15
H <sub>2</sub> O + CrH + Na	407.55	-0.70
H <sub>2</sub> O + CrH + CP	408.81	0.56
H <sub>2</sub> O + CrH + Na + K	408.93	0.68
H <sub>2</sub> O + CrH + NH <sub>3</sub>	409.23	0.98
H <sub>2</sub> O + CrH + K	409.53	1.28

**Notes.** Except for the flat model, every model includes the opacity caused by Rayleigh scattering and CIA.

**Fig. 3.** TAUREX retrieval results for WASP-31b without the *Spitzer* measurements.**Table 5.** Retrieved atmospheric parameters using TAUREX for the model shown in Fig. 2.

Parameter	Retrieved value
$T_{\text{atm}}$ (K)	$1481^{+264}_{-355}$
$R_{\text{pl}}$ ( $R_J$ )	$1.48^{+0.02}_{-0.01}$
$\log(P_{\text{clouds}})$ (Pa)	$3.87^{+0.20}_{-0.20}$
$\log(X_{\text{H}_2\text{O}})$	$-5.40^{+0.37}_{-0.43}$
$\log(X_{\text{CrH}})$	$-8.51^{+0.62}_{-0.60}$
$\log(X_K)$	$-7.59^{+0.66}_{-0.94}$

**Notes.** The extended posterior distributions can be found in Fig. A.1.

tight constraints as well as unrealistic values for retrieved abundances. In this way, the retrievals may introduce biases in, for example, the abundances of species that are included in the model. The retrieved parameters for our best-fitting model are shown in Table 5, and the extended posterior distributions can be found in Fig. A.1. To test whether the omission of atmospheric species leads to biases in the retrieved parameters, we retrieved the same spectrum assuming a model with a variety of opacity sources (H<sub>2</sub>O, CrH, K, CO<sub>2</sub>, CO, CH<sub>4</sub>, NH<sub>3</sub>, and Na). This retrieval results in similar abundances of  $\log(X_{\text{H}_2\text{O}}) = -5.39^{+0.42}_{-0.73}$ ,  $\log(X_{\text{CrH}}) = -8.19^{+0.75}_{-0.77}$ , and  $\log(X_K) = -7.92^{+0.82}_{-1.59}$  (see the second row of Table A.1). This shows that including the other chemical species is not essential when retrieving abundances from this spectrum of WASP-31b. Increasing the number of opacity sources leads to a larger error on the retrieved parameters. This is as expected: widening the allowed parameter space increases the number of possible solutions in the retrieval procedure.

#### 4.1. Comparison to previous work

Earlier investigations of the transmission spectrum of WASP-31b found signatures of K, a grey cloud deck, and Rayleigh scattering



(Sing et al. 2015, 2016). On top of that, a weak water absorption feature was found, which was also reported by later studies that used the same spectrum (Barstow et al. 2017; Tsiaras et al. 2018). Tsiaras et al. (2018) found weak evidence for a water VMR of  $\log(X_{\text{H}_2\text{O}}) = -3.84 \pm 1.90$ . Other retrieved water abundances are equal to  $-3.97^{+1.01}_{-2.27}$  (Pinhas et al. 2019) and  $-4.55^{+1.77}_{-4.33}$  (Welbanks et al. 2019) with confidence levels of  $\sim 2\sigma$  and  $\sim 2.2\sigma$ , respectively, and an abundance of  $-3.27^{+1.44}_{-2.18}$  is found using the classical retrieval method within the ARCS code (Min et al. 2020). Our retrieved water abundance (see Table 5) is lower but falls inside the error bars of the other investigations. Sing et al. (2015) reported a  $4.2\sigma$  significant detection of K, but the fidelity of this data point was later questioned by individual searches using the ground-based facilities FORS2 and UVES (Gibson et al. 2017, 2019) and IMACS (McGruder et al. 2020). Our detection agrees with the  $2.2\sigma$  detection that was made using combined FORS2 and STIS optical data (Gibson et al. 2017). MacDonald & Madhusudhan (2017) also found a weak detection ( $2.2\sigma$ ) of  $\text{NH}_3$ , which was found at a similar abundance by Min et al. (2020). As can be seen in Table 4, the addition of  $\text{NH}_3$  to our  $\text{H}_2\text{O}+\text{CrH}$  model also leads to an increase in the evidence level ( $\text{DS} = 0.98$ ). As shown in Table 3, this was not seen as significant in our analysis (just below  $2.1\sigma$ ). This small difference can be explained by the overlap in the CrH and  $\text{NH}_3$  features around  $1.5\mu\text{m}$ . Covering the features at longer wavelengths (e.g. at  $\sim 2.2\mu\text{m}$ ) will greatly improve our ability to detect  $\text{NH}_3$ , as illustrated by MacDonald & Madhusudhan (2017).

Retrieved atmospheric temperatures for WASP-31b vary from  $738^{+345}_{-231}$  K (Min et al. 2020) to  $1088.35 \pm 220.16$  K (Tsiaras et al. 2018) and  $1043^{+267}_{-172}$  K (Pinhas et al. 2019), all of which are exceeded by our retrieved temperature. Moreover, MacDonald et al. (2020) showed that hot Jupiter temperatures are generally underestimated by 1D retrievals. An explanation for this discrepancy in temperature might be an insufficient cloud model since Sing et al. (2016) reported that the spectrum of this planet is not well explained by a single cloud model. A more complex model is included in TAUREX, which parametrises the opacity caused by particle scattering following the Mie theory (Lee et al. 2013). Adding this parametrisation to our best-fitting model also results in a relatively high  $T_{\text{atm}} = 1507^{+239}_{-308}$  K and low  $\log(X_{\text{H}_2\text{O}}) = -5.73^{+0.51}_{-3.77}$  and is not found to be statistically significant. It is possible that this parametrisation is still not sufficiently complex. Another explanation may be the difference in opacity sources that are included: Our retrieval only includes  $\text{H}_2\text{O}$ , CrH, and K, whereas the others generally include  $\text{H}_2\text{O}$ ,  $\text{CH}_4$ , CO,  $\text{CO}_2$ , and  $\text{NH}_3$ . Additionally, Na and K (Pinhas et al. 2019; Min et al. 2020) and HCN (Pinhas et al. 2019) were also included in the retrievals. Increasing the number of opacity sources does indeed lead to a lower temperature of  $T_{\text{atm}} = 1172^{+435}_{-226}$  K (see the second row of Table A.1), consistent with the majority of earlier findings. With a Bayesian evidence level of 407.61, the addition of these opacity sources is not found to be statistically significant.

Associated with the lower temperature of this retrieval is an increase in the retrieved radius to  $1.50 R_J$ . As shown in Table 5, the radius of WASP-31b is equal to  $1.48 R_J$  in our best-fit model. This degeneracy between  $R_{\text{pl}}$  and T can also be seen from the posterior distributions in Fig. A.1: specifically, the plot in row 3 (from the bottom) and column 4 (from the left) shows that an increase in radius is degenerate with a decrease in temperature (due to its influence on the scale height). This degeneracy is another explanation for the difference with Min et al. (2020), who retrieve  $R_{\text{pl}} = 1.51^{+0.02}_{-0.03} R_J$ . To test this suspicion, a retrieval was conducted assuming a fixed radius of  $R_{\text{pl}} = 1.549$ , and it did

indeed result in a lower temperature of  $1267^{+351}_{-288}$  K. However, this result is accompanied by unphysically high abundances of CrH and K ( $\log(X) = -3.61$  and  $-0.18$ , respectively). Using equilibrium temperatures, we can predict the maximum temperature of the terminator region to be  $\sim 1550$  K for full redistribution ( $f = 1$ ) and perfect absorption of radiation ( $A = 0$ ). Hence, our retrieved temperature would be reasonable. The degeneracy that exists between temperature, radius, and abundances (e.g. Griffith 2014; Heng & Kitzmann 2017) may offer an explanation for the relatively low water abundances that we retrieve. A higher temperature leads to an overall increase in the scale height and thus the transit depth, dampening absorption features and leading to lower chemical abundances in the retrieval.

Previous identifications of the signatures of CrH have been reported for brown dwarfs (e.g. Kirkpatrick et al. 1999a,b). Specifically, Burrows et al. (2002) found an abundance of  $\text{CrH}/\text{H}_2 \sim 2 - 4 \times 10^{-9}$  for the L5 dwarf 2MASS J1507038-151648, which is in excellent agreement with the abundance that we retrieve for WASP-31b. MacDonald & Madhusudhan (2019) report a  $4.1\sigma$  detection of metal hydrides in the transmission spectrum of exo-Neptune HAT-P-26b, and they identify three possible candidates: TiH ( $4.1\sigma$ ), CrH ( $2.1\sigma$ ), or ScH ( $1.8\sigma$ ). As a possible candidate, CrH is retrieved at an abundance of  $-5.72^{+0.89}_{-1.37}$ , which exceeds our value by almost three orders of magnitude. They propose vertical transport or secular contaminations by planetesimals as possible explanations for this high abundance and the fact that Cr is expected to have condensed out at the temperature of HAT-P-26b ( $T_{\text{eq}} \sim 1000$  K).

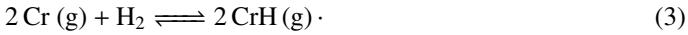
The retrieved abundances can also be compared to the predictions from equilibrium chemistry, for example by using the GGChem code (Woitke et al. 2018). Around our retrieved temperature of WASP-31b, GGChem predicts CrH to be present at  $\log(X_{\text{CrH}}) \sim -9$  for  $P = 1$  bar and solar composition, with lower abundances for lower pressures (Woitke et al. 2018). Hence, the retrieved CrH abundance of  $-8.51^{+0.62}_{-0.60}$  is higher than predicted but still consistent.  $\text{H}_2\text{O}$  is expected at  $\log(X_{\text{H}_2\text{O}}) \sim -3.3$  for the same temperature, about 100 times higher than the retrieved abundance. As previously stated, this might be related to degeneracies between different retrieval parameters. The fact that this large difference is not retrieved for the CrH abundance might also hint at an actual depletion of  $\text{H}_2\text{O}$ . Further observations can help in disclosing this.

#### 4.2. Chemistry

From the first-row transition metals, Cr is the third most abundant element after Fe and Ni in the Sun (Asplund et al. 2009). At the temperatures of close-in exoplanets, CrH is predicted to be an important Cr-bearing species. However, gaseous atomic Cr is expected to be the main Cr bearer, whereas significant fractions are also expected to be present in CrO or CrS (Woitke et al. 2018). These calculations were made assuming solar abundances and the corresponding solar abundance ratio  $\log(\text{Cr}/\text{O}) = -3.05$  (Asplund et al. 2009). If we make the simplifying assumption that for WASP-31b most of the Cr is in CrH, the planetary Cr/O abundance ratio can be calculated. At the temperature of WASP-31b, about half of the oxygen is expected in  $\text{H}_2\text{O}$  and the other half in CO (Madhusudhan 2012; Woitke et al. 2018). To correct for this, the retrieved  $\text{H}_2\text{O}$  abundance is multiplied by two, leading to a ratio of  $\log(\text{Cr}/\text{O}) = -3.41$ . The abundance ratio is lower than the solar value, but additional Cr is probably present in other species. Including the opacity data of these species in retrievals can lead to better constraints on the ratios, and the detectability of atomic Cr has recently been shown by its signatures in the

ultra-hot Jupiter WASP-121b (Ben-Yami et al. 2020). Of course, abundance ratios may differ per star. For WASP-31, ratios are measured to be  $O/H = +0.06$  dex and  $Cr/H = -0.08$  dex, relative to the Sun (Brewer et al. 2016). This leads to a lower stellar abundance ratio of  $\log(Cr/O) = -3.19$  for WASP-31, which may also partly explain the lower planetary ratio.

Monatomic Cr, the major gas-phase bearer at a wide range of temperatures (300–3000 K), is a refractory species. At the relevant pressures, it condenses into Cr metal between 1400 and 1520 K and into  $Cr_2O_3$  at lower pressures of  $\sim 10^{-3}$  bar (e.g. Burrows & Sharp 1999; Lodders & Fegley 2006; Morley et al. 2012). The CrH abundance is related to the monatomic gas according to the equilibrium (Lodders & Fegley 2006):



The condensation of Cr metal reduces the abundances of monatomic Cr and CrH, depleting the species from the atmosphere. Vertical mixing from lower, hotter layers is unlikely since Cr destruction reactions are highly exothermic at  $T < 1400$  K, resulting in chemical lifetimes much shorter than the timescales of vertical mixing (Lodders & Fegley 2006). For the WASP-31b  $T_{\text{atm}} = 1481^{+264}_{-355}$  K, the appearance of CrH would be reasonable since the gaseous Cr would not yet be fully depleted.

The discovery of Cr bearers can have implications for cloud formation in exoplanet atmospheres since they may cause the formation of Cr[s] clouds (Lodders & Fegley 2006; Morley et al. 2012). Moreover, Lee et al. (2018) suggested the possibility of Cr[s] being seed particles that provide condensation surfaces for other cloud layers (e.g. sulphide or KCl cloud layers).

Lastly, the finding of Cr-bearing species in an atmosphere may give clues about the formation conditions of a planet. Because Cr is a refractory species, it is expected to be in the solid phase throughout most of the protoplanetary disk (e.g. Lodders 2010). Consequently, its presence on an exoplanet hints at the accretion of solid material during its formation. Determining the planetary Cr abundance (also in other Cr bearers) can then provide clues about the amount of solid accretion.

#### 4.3. Other observations

Since the fidelity of the data point responsible for the K detection has been questioned (Gibson et al. 2017, 2019), a few retrievals were conducted excluding the observed transit depth at  $\sim 0.77 \mu\text{m}$ . CrH has absorption features around this wavelength, and these retrievals were done to make sure that the tentative CrH detection is not based on a disputed observation. A similar approach to what was described in Sect. 2 was followed. The resulting Bayesian evidence levels can be seen in Table 6 and agree with our earlier findings, albeit with slightly lower DSs: The inclusion of both  $H_2O$  and CrH is preferred, with a DS of 7.97 (or  $\sim 4.3\sigma$  confidence) over the flat model, or 2.98 ( $\sim 2.9\sigma$ ) over a model of only Rayleigh scattering and CIA. The influence of the measured transit depth at  $\sim 0.77 \mu\text{m}$  is explained by this decrease in DS, but, even without the measurement, statistically significant evidence for CrH is still found. In this case, a slightly lower  $T_{\text{atm}} = 1339^{+332}_{-321}$  K and higher  $R_{\text{pl}} = 1.49^{+0.02}_{-0.02} R_J$  are retrieved.

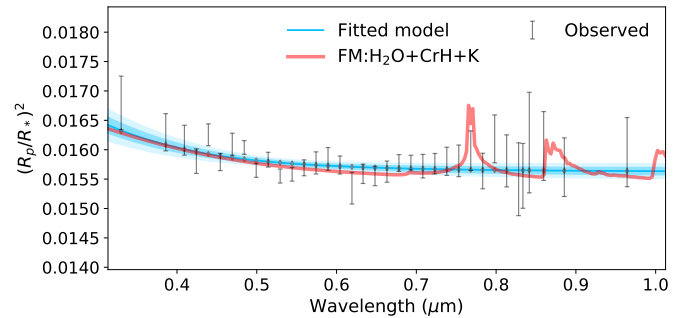
Ground-based optical data by the FORS2 at the VLT are also available for this planet (Gibson et al. 2017) and offer coverage from 0.4 to  $0.84 \mu\text{m}$ . The combined FORS2/STIS data that are presented by Gibson et al. (2017) were also analysed using TAUREX. With the same general setup, we retrieved the spectrum assuming five different models to test whether the CrH features are also found in the ground-based data. The resulting

**Table 6.** Resulting Bayesian evidence levels and DSs for the Sing et al. (2015) spectrum without  $0.77 \mu\text{m}$  observation.

Model parameters	$\log(E)$	DS
Flat model	400.44	
Rayleigh+CIA	405.43	4.99
Compared to flat model		
CrH	406.01	5.58
CrH+ $H_2O$	408.40	7.97

**Table 7.** Resulting Bayesian evidence levels and DSs for the combined FORS2/STIS data.

Model parameters	$\log(E)$	DS
Flat model	241.49	
Rayleigh+CIA	245.36	3.87
Compared to flat model		
CrH+K	244.56	3.08
CrH	244.72	3.24
CrH+ $H_2O$	245.50	4.01



**Fig. 4.** Transmission spectrum as observed by FORS2/STIS with the best-fitting atmospheric model. The vertical error bars indicate the observed transit depths, and the different shadings in the upper panel represent 1 and  $2\sigma$  regions. A forward model based on the retrieval results in Table 5 is shown by the red line.

evidence levels and accompanying DSs are shown in Table 7. Using these data, a model containing CrH is not found to be statistically significant over a model only containing Rayleigh scattering and CIA. Hence, in this case, the best-fitting atmospheric model was found to consist only of Rayleigh scattering and CIA and to be without any chemical species, indicating a cloudy atmosphere at a retrieved temperature of  $T_{\text{atm}} = 1503^{+267}_{-369}$  K. The retrieved planetary radius agrees with the value we previously found at  $1.48^{+0.02}_{-0.01} R_J$ . The spectrum and its lack of absorption features can be seen in Fig. 4. While the measured transit depths between 0.7 and  $1.0 \mu\text{m}$  seem to show some signatures, they are not found to be statistically significant. For comparison, the forward model based on the retrieved parameters in Table 5 is shown by the red line, indicating the signatures that should be visible when CrH is present in the atmosphere.

Taking another look at the bottom image of Fig. 3, the orange line shows that the presence of CrH results in six prominent absorption peaks. Only three of these peaks (at 0.69, 0.77, and



**Table 8.** Resulting Bayesian evidence levels and DSs for a combined FORS2/STIS and WFC3 spectrum.

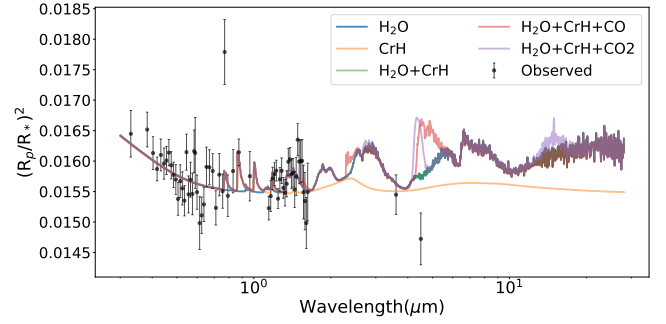
Model parameters	$\log(E)$	DS
Flat model	408.37	
Rayleigh+CIA	412.69	4.32
Compared to flat model		
FeH	410.81	2.45
CrH	411.60	3.23
H <sub>2</sub> O	413.91	5.55
Compared to H <sub>2</sub> O-only model		
H <sub>2</sub> O + FeH	412.60	-1.31
H <sub>2</sub> O + AlH	412.86	-1.05
H <sub>2</sub> O + CP	413.23	-0.68
H <sub>2</sub> O + CrH	413.29	-0.62
H <sub>2</sub> O + ScH	413.37	-0.55
H <sub>2</sub> O + NH <sub>3</sub>	414.19	0.28

0.88  $\mu\text{m}$ ) fall (partially) inside the range probed by this combined FORS2/STIS spectrum, which can also be seen from the red line in Fig. 4. To test whether the peaks in the WFC3 regime (at 1.18 and 1.43  $\mu\text{m}$ ) are driving the evidence for CrH, we added the WFC3 data to the combined FORS2/STIS data and conducted some additional retrievals on this spectrum. From the resulting evidence levels, as shown in Table 8, it can be seen that this does not lead to a significant detection of CrH and only results in a preference for H<sub>2</sub>O at  $\sim 3.7\sigma$  over a flat model. Using this spectral coverage, the temperature is retrieved to be  $T_{\text{atm}} = 1476^{+277}_{-376}$  K and the planetary radius agrees with our earlier findings at  $1.48^{+0.02}_{-0.02} R_J$ . We can conclude that the evidence for CrH is driven by the observations made with STIS, specifically the observed transit depths around 0.88 and 0.77  $\mu\text{m}$ . However, compatibility between different instruments cannot be taken for granted, and caution should thus be exercised when combining data, as illustrated by Hou Yip et al. (2020) for the case of WASP-96b. This uncertainty, the non-detection when including ground-based data, and the broad continuous wavelength coverage that will be offered by future facilities further motivate the characterisation of WASP-31b.

#### 4.4. Near-future coverage

Figure 5 shows the transmission spectrum for the wider wavelength range offered by the *James Webb* Space Telescope (JWST), scheduled for launch in October 2021. Using a variety of instruments, JWST will offer a spectral range from 0.6 to 28  $\mu\text{m}$  (Beichman et al. 2014). The spectra in Fig. 5 correspond to different atmospheric compositions and are based on forward models that assume the retrieved best-fit values for atmospheric parameters; because of the discrepancy in the K detection, the alkali metal is not included in the models.

The blue model is the only forward model that does not include CrH, and, by comparing it with the other models, it can be seen that the presence of CrH is purely based on the absorption signatures between 0.69 and 1.43  $\mu\text{m}$ . Next to that, the presence of water can clearly be inferred from the familiar feature around 1.4  $\mu\text{m}$  as well as the feature at 1.9  $\mu\text{m}$ , showing a clear distinction from the CrH-only model. Although many

**Fig. 5.** Forward models for the transmission spectrum of WASP-31b in the spectral regime of JWST, based on the retrieved values for atmospheric parameters in Table 5.

additional water signatures can be found at longer wavelengths, the fact that JWST's Near Infrared Imager and Slitless Spectrograph (NIRISS) can provide simultaneous coverage at  $R \sim 700$  from 1 to 2.5  $\mu\text{m}$  makes these H<sub>2</sub>O features and the CrH features in this regime interesting prospects for further characterisation. Another important goal is to derive improved C/O ratios, and combined observations from NIRISS and the Near Infrared Camera (NIRCam) are expected to deliver this (Stevenson et al. 2016), providing coverage from 1 to 5  $\mu\text{m}$ . This is also evident from Fig. 5, where a variation in transit depth can be seen around wavelengths of 4.5  $\mu\text{m}$ , depending on whether or not CO or CO<sub>2</sub> are included in the forward model. Hence, WASP-31b would be an interesting target for further characterisation, whereas the presence of CrH (and other metal hydrides) in exoplanets of similar temperatures is expected to be detectable with JWST. Because CrH rapidly condenses out for lower temperatures and the spectral signatures of more refractory materials such as TiO and VO start taking over at higher temperatures, its presence is probably detectable for planets with temperatures ranging from  $\sim 1300$ –2000 K.

## 5. Conclusions

In this study, a re-analysis of publicly available transmission data of the hot exoplanet WASP-31b has been conducted using the TAUREX II retrieval framework. Transmission data from STIS, WFC3, and *Spitzer* provide spectral coverage between 0.3 and 4.5  $\mu\text{m}$ . Assuming the simplified atmospheric representation in TAUREX and out of the models that were fitted in this analysis, it was found that the spectrum is best explained by a model containing H<sub>2</sub>O, CrH, and K in addition to H<sub>2</sub>, He, a grey cloud deck, and Rayleigh scattering. As compared to a flat model without any spectral features, the H<sub>2</sub>O-only model is statistically preferred at  $\sim 3.7\sigma$  and a CrH-only model at  $\sim 3.9\sigma$ . A model with both H<sub>2</sub>O and CrH was found at  $\sim 4.4\sigma$  and  $\sim 3.2\sigma$  over the flat model and a CIA+Rayleigh scattering model, respectively. Hence, we report the first statistical evidence for the signatures of CrH in an exoplanet atmosphere. Weak evidence for the addition of K to the atmospheric model was found at  $\sim 2.2\sigma$  confidence over the H<sub>2</sub>O+CrH model. As compared to earlier studies of WASP-31b, a relatively high temperature was retrieved, which can be explained by a combined influence of fewer opacity sources and degeneracies between temperature, radius, and chemical abundances.

The evidence for CrH naturally follows from its presence in brown dwarfs and is expected to be limited to planets with temperatures between 1300 and 2000 K. Cr-bearing species may

play a role in the formation of clouds in exoplanet atmospheres, and their detection is also an indication of the accretion of solids during the formation of a planet.

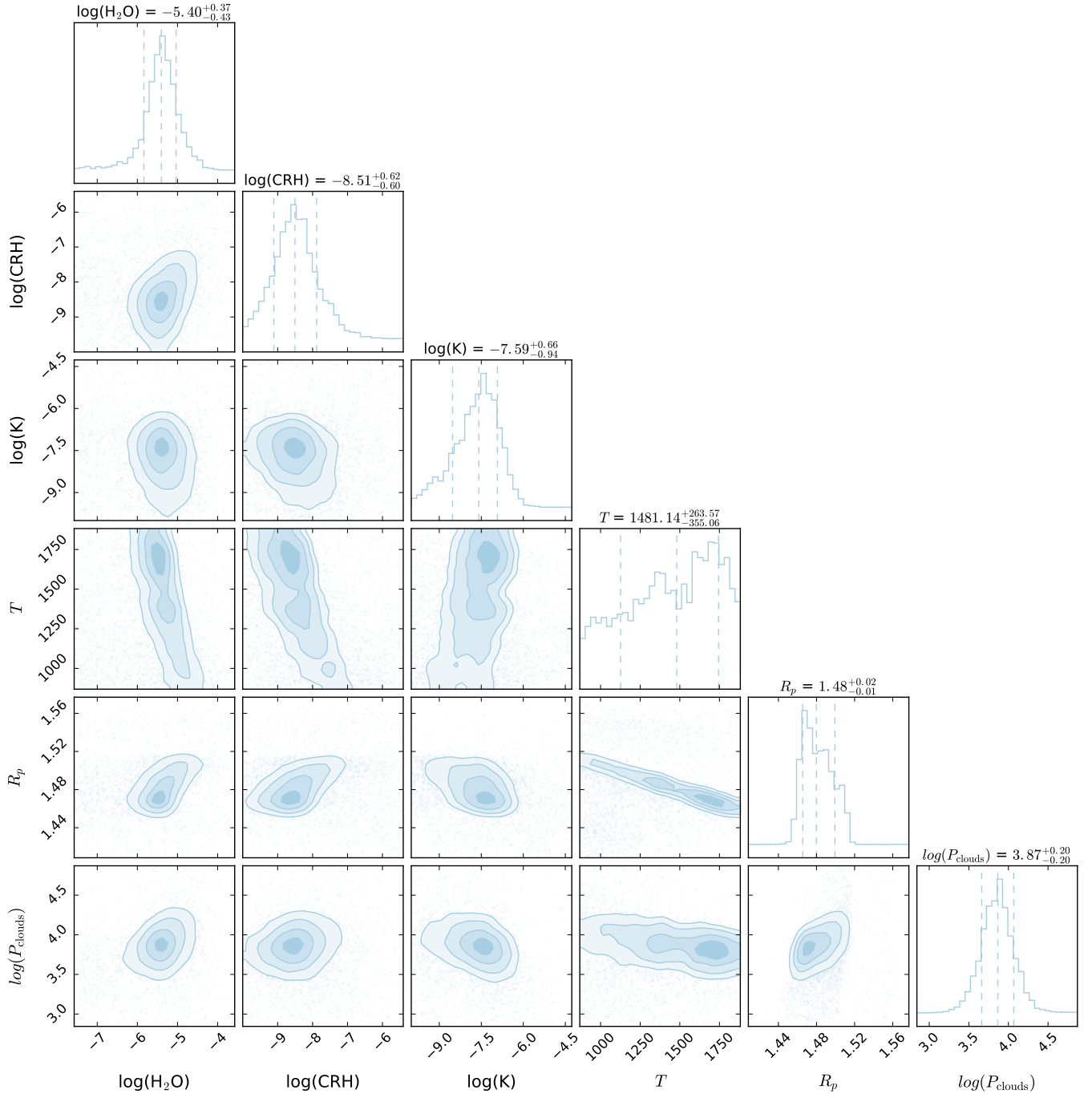
In additional retrievals, the disputed data point at  $\sim 0.77\ \mu\text{m}$  was excluded, but evidence for the  $\text{H}_2\text{O}+\text{CrH}$  model was still found at a confidence level of  $\sim 4.3\sigma$  over the flat model. A combined FORS2/STIS spectrum was also available and tests were performed to confirm the CrH detection, but in this case no statistically significant CrH feature was found. By analysing the retrieval outcomes for different combinations of spectral coverage, it was found that the evidence for CrH is mostly based on the observed transit depths around  $0.77$  and  $0.88\ \mu\text{m}$ . Inspired by the non-agreement between different instruments, and using the best-fit atmospheric model for WASP-31b, it was shown that the spectral regime of JWST has the potential to confirm the CrH features.

**Acknowledgements.** We kindly thank Joanna Barstow for an excellent review which was valuable in improving the manuscript. We greatly appreciate the developers of TauREx (I. P. Waldmann, Q. Changeat, A. F. Al-Refaie, G. Tinetti, M. Rocchetto, E. J. Barton, S. N. Yurchenko and J. Tennyson), for making the TAUREX II retrieval framework available and for their help in any inquiries. We would like to express our gratitude to the team who led the observations and made the data which was used in this work available (PI: D. K. Sing) (See [https://pages.jh.edu/~dsing3/David\\_Sing/Spectral\\_Library.html](https://pages.jh.edu/~dsing3/David_Sing/Spectral_Library.html) and <https://stellarplanet.org/science/exoplanet-transmission-spectra/>).

## References

- Al-Refaie, A. F., Yurchenko, S. N., Yachmenev, A., & Tennyson, J. 2015, *MNRAS*, **448**, 1704
- Allard, N. F., Spiegelman, F., & Kielkopf, J. F. 2016, *A&A*, **589**, A21
- Allard, N. F., Spiegelman, F., Leininger, T., & Mollière, P. 2019, *A&A*, **628**, A120
- Anderson, D. R., Collier Cameron, A., Hellier, C., et al. 2011, *A&A*, **531**, A60
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Barber, R. J., Strange, J. K., Hill, C., et al. 2014, *MNRAS*, **437**, 1828
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., & Sing, D. K. 2017, *ApJ*, **834**, 50
- Beichman, C., Benneke, B., Knutson, H., et al. 2014, *PASP*, **126**, 1134
- Ben-Yami, M., Madhusudhan, N., Cabot, S. H. C., et al. 2020, *ApJ*, **897**, L5
- Bernath, P. F. 2020, *J. Quant. Spectr. Rad. Transf.*, **240**, 106687
- Brewer, J. M., Fischer, D. A., Valenti, J. A., & Piskunov, N. 2016, *ApJS*, **225**, 32
- Brooke, J. S. A., Ram, R. S., Western, C. M., et al. 2014, *ApJS*, **210**, 23
- Burrows, A., & Sharp, C. M. 1999, *ApJ*, **512**, 843
- Burrows, A., Ram, R. S., Bernath, P., Sharp, C. M., & Milsom, J. A. 2002, *ApJ*, **577**, 986
- Burrows, A., Dulick, M., Bauschlicher, C. W., et al. 2005, *ApJ*, **624**, 988
- Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, *ApJ*, **568**, 377
- Chubb, K. L., Min, M., Kawashima, Y., Helling, C., & Waldmann, I. 2020a, *A&A*, **639**, A3
- Chubb, K. L., Tennyson, J., & Yurchenko, S. N. 2020b, *MNRAS*, **493**, 1531
- Chubb, K. L., Rocchetto, M., Yurchenko, S. N., et al. 2021, *A&A*, in press <https://doi.org/10.1051/0004-6361/202038350>
- Coles, P. A., Yurchenko, S. N., & Tennyson, J. 2019, *MNRAS*, **490**, 4638
- Crossfield, I. J. M. 2015, *PASP*, **127**, 941
- Evans, T. M., Sing, D. K., Wakeford, H. R., et al. 2016, *ApJ*, **822**, L4
- Evans, T. M., Sing, D. K., Kataria, T., et al. 2017, *Nature*, **548**, 58
- Feroz, F., & Hobson, M. P. 2008, *MNRAS*, **384**, 449
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, **398**, 1601
- Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019, *Open J. Astrophys.*, **2**, 10
- Gharib-Nezhad, E., Shayesteh, A., & Bernath, P. F. 2013, *MNRAS*, **432**, 2043
- Gibson, N. P., Nikolov, N., Sing, D. K., et al. 2017, *MNRAS*, **467**, 4591
- Gibson, N. P., de Mooij, E. J. W., Evans, T. M., et al. 2019, *MNRAS*, **482**, 606
- Griffith, C. A. 2014, *Phil. Trans. R. Soc. London, Ser. A*, **372**, 20130086
- Hansen, C. J., Schwartz, J. C., & Cowan, N. B. 2014, *MNRAS*, **444**, 3632
- Heng, K., & Kitzmann, D. 2017, *MNRAS*, **470**, 2972
- Hoeijmakers, H. J., Ehrenreich, D., Heng, K., et al. 2018, *Nature*, **560**, 453
- Hoeijmakers, H. J., Ehrenreich, D., Kitzmann, D., et al. 2019, *A&A*, **627**, A165
- Hollis, M. D. J., Tesseney, M., & Tinetti, G. 2013, *Comput. Phys. Commun.*, **184**, 2351
- Hou Yip, K., Changeat, Q., Edwards, B., et al. 2020, *ApJ*, submitted [arXiv:2009.10438]
- Jeffreys, H. 1998, *The Theory of Probability* (Oxford: Oxford University Press)
- Kass, R., & Raftery, A. 1995, *J. Am. Stat. Assoc.*, **90**, 430
- Kawashima, Y., & Ikoma, M. 2019, *ApJ*, **877**, 109
- Kesseli, A., Snellen, I. A. G., Alonso-Floriano, F. J., Mollière, P., & Serindag, D. B. 2020, *AJ*, **160**, 228
- Kirkpatrick, J. D. 2005, *ARA&A*, **43**, 195
- Kirkpatrick, J. D., Allard, F., Bida, T., et al. 1999a, *ApJ*, **519**, 834
- Kirkpatrick, J. D., Reid, I. N., Liebert, J., et al. 1999b, *ApJ*, **519**, 802
- Kramida, A., Ralchenko, Y., & Reader, J. 2013, NIST Atomic Spectra Database – Version 5, <http://www.nist.gov/pml/data/asd.cfm>
- Lee, J.-M., Heng, K., & Irwin, P. G. J. 2013, *ApJ*, **778**, 97
- Lee, G. K. H., Blecic, J., & Helling, C. 2018, *A&A*, **614**, A126
- Li, G., Harrison, J. J., Ram, R. S., Western, C. M., & Bernath, P. F. 2012, *J. Quant. Spectr. Rad. Transf.*, **113**, 67
- Li, G., Gordon, I. E., Rothman, L. S., et al. 2015, *ApJS*, **216**, 15
- Li, H. Y., Tennyson, J., & Yurchenko, S. N. 2019, *MNRAS*, **486**, 2351
- Line, M. R., Liang, M. C., & Yung, Y. L. 2010, *ApJ*, **717**, 496
- Lodders, K. 2010, in *Formation and Evolution of Exoplanets*, ed. R. Barnes (Hoboken: John Wiley & Sons, Ltd), 157
- Lodders, K., & Fegley, B., J. 2006, *Chemistry of Low Mass Substellar Objects*, ed. J. W. Mason (Berlin: Springer), 1
- Lodi, L., Yurchenko, S. N., & Tennyson, J. 2015, *Mol. Phys.*, **113**, 1559
- MacDonald, R. J., & Madhusudhan, N. 2017, *ApJ*, **850**, L15
- MacDonald, R. J., & Madhusudhan, N. 2019, *MNRAS*, **486**, 1292
- MacDonald, R. J., Goyal, J. M., & Lewis, N. K. 2020, *ApJ*, **893**, L43
- Madhusudhan, N. 2012, *ApJ*, **758**, 36
- Madhusudhan, N. 2019, *ARA&A*, **57**, 617
- Mant, B. P., Yachmenev, A., Tennyson, J., & Yurchenko, S. N. 2018, *MNRAS*, **478**, 3220
- Mayor, M., & Queloz, D. 1995, *Nature*, **378**, 355
- McGruder, C. D., López-Morales, M., Espinoza, N., et al. 2020, *AJ*, **160**, 230
- McKemmish, L. K., Masseron, T., Hoeijmakers, H. J., et al. 2019, *MNRAS*, **488**, 2836
- McKemmish, L. K., Yurchenko, S. N., & Tennyson, J. 2016, *MNRAS*, **463**, 771
- Min, M., Ormel, C. W., Chubb, K., Helling, C., & Kawashima, Y. 2020, *A&A*, **642**, A28
- Morley, C. V., Fortney, J. J., Marley, M. S., et al. 2012, *ApJ*, **756**, 172
- Patrascu, A. T., Tennyson, J., & Yurchenko, S. N. 2015, *MNRAS*, **449**, 3613
- Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, *MNRAS*, **482**, 1485
- Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, *PASP*, **118**, 1407
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, *MNRAS*, **480**, 2597
- Ram, R. S., Brooke, J. S. A., Western, C. M., & Bernath, P. F. 2014, *J. Quant. Spectr. Rad. Transf.*, **138**, 107
- Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, *J. Quant. Spectr. Rad. Transf.*, **111**, 2139
- Seager, S., & Sasselov, D. D. 2000, *ApJ*, **537**, 916
- Sedaghati, E., Boffin, H. M. J., MacDonald, R. J., et al. 2017, *Nature*, **549**, 238
- Sharp, C. M., & Burrows, A. 2007, *ApJS*, **168**, 140
- Sing, D. K., Wakeford, H. R., Showman, A. P., et al. 2015, *MNRAS*, **446**, 2428
- Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, *Nature*, **529**, 59
- Skaf, N., Bieger, M. F., Edwards, B., et al. 2020, *AJ*, **160**, 109
- Sotzen, K. S., Stevenson, K. B., Sing, D. K., et al. 2020, *AJ*, **159**, 5
- Stevenson, K. B., Lewis, N. K., Bean, J. L., et al. 2016, *PASP*, **128**, 094401
- Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2016, *J. Mol. Spectr.*, **327**, 73
- Trotta, R. 2008, *Contemp. Phys.*, **49**, 71
- Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, *AJ*, **155**, 156
- Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, *ApJ*, **802**, 107
- Welbanks, L., Madhusudhan, N., Allard, N. F., et al. 2019, *ApJ*, **887**, L20
- Wende, S., Reiners, A., Seifahrt, A., & Bernath, P. F. 2010, *A&A*, **523**, A58
- Woitke, P., Helling, C., Hunter, G. H., et al. 2018, *A&A*, **614**, A1
- Wolszczan, A., & Frail, D. A. 1992, *Nature*, **355**, 145
- Yousefi, M., Bernath, P. F., Hodges, J., & Masseron, T. 2018, *J. Quant. Spectr. Rad. Transf.*, **217**, 416
- Yurchenko, S. N., Amundsen, D. S., Tennyson, J., & Waldmann, I. P. 2017, *A&A*, **605**, A95
- Yurchenko, S. N., Williams, H., Leyland, P. C., Lodi, L., & Tennyson, J. 2018, *MNRAS*, **479**, 1401

## Appendix A: Retrieval outputs



**Fig. A.1.** TAUREX posterior distributions of highest evidence retrieval for WASP-31b.



**Table A.1.** Retrieved atmospheric parameters using TAUREX for different retrieval setups and spectral coverage.

Coverage and retrieval setup	$T_{\text{atm}}$ (K)	$R_{\text{pl}}$ ( $R_{\text{J}}$ )	$\log(P_{\text{clouds}})$ (Pa)	$\log(X_{\text{H}_2\text{O}})$	$\log(X_{\text{CrH}})$	$\log(X_{\text{K}})$
<b>STIS + WFC3 + <i>Spitzer</i></b>						
H <sub>2</sub> O+CrH+K (highest Evidence model)	1481 <sup>+264</sup> <sub>-355</sub>	1.48 <sup>+0.02</sup> <sub>-0.01</sub>	3.87 <sup>+0.20</sup> <sub>-0.20</sub>	-5.40 <sup>+0.37</sup> <sub>-0.43</sub>	-8.51 <sup>+0.62</sup> <sub>-0.60</sub>	-7.59 <sup>+0.66</sup> <sub>-0.94</sub>
H <sub>2</sub> O+CrH+K+CO <sub>2</sub> +CO+CH <sub>4</sub> +NH <sub>3</sub> +Na	1172 <sup>+435</sup> <sub>-226</sub>	1.50 <sup>+0.01</sup> <sub>-0.02</sub>	4.03 <sup>+1.07</sup> <sub>-0.29</sub>	-5.39 <sup>+0.42</sup> <sub>-0.73</sub>	-8.19 <sup>+0.75</sup> <sub>-0.77</sub>	-7.92 <sup>+0.82</sup> <sub>-1.59</sub>
H <sub>2</sub> O+CrH+K (Complex cloud)	1507 <sup>+239</sup> <sub>-308</sub>	1.48 <sup>+0.02</sup> <sub>-0.01</sub>	3.86 <sup>+0.18</sup> <sub>-0.19</sub>	-5.41 <sup>+0.33</sup> <sub>-0.37</sub>	-8.53 <sup>+0.57</sup> <sub>-0.58</sub>	-7.63 <sup>+0.67</sup> <sub>-0.87</sub>
H <sub>2</sub> O+CrH+K (Radius fixed at 1.549 $R_{\text{J}}$ )	1267 <sup>+351</sup> <sub>-288</sub>	1.549 <sup>+0.00</sup> <sub>-0.00</sub>	0.93 <sup>+0.98</sup> <sub>-1.27</sub>	-5.25 <sup>+3.46</sup> <sub>-3.20</sub>	-3.61 <sup>+2.40</sup> <sub>-3.26</sub>	-0.18 <sup>+0.11</sup> <sub>-0.18</sub>
H <sub>2</sub> O+CrH	1339 <sup>+332</sup> <sub>-321</sub>	1.49 <sup>+0.02</sup> <sub>-0.02</sub>	3.89 <sup>+0.22</sup> <sub>-0.21</sub>	-5.33 <sup>+0.44</sup> <sub>-0.44</sub>	-8.45 <sup>+0.66</sup> <sub>-0.63</sub>	n/a
<b>STIS + WFC3</b>						
H <sub>2</sub> O+CrH+K	1614 <sup>+184</sup> <sub>-309</sub>	1.47 <sup>+0.02</sup> <sub>-0.01</sub>	3.86 <sup>+0.18</sup> <sub>-0.19</sub>	-5.35 <sup>+0.29</sup> <sub>-0.28</sub>	-8.65 <sup>+0.56</sup> <sub>-0.56</sub>	-7.54 <sup>+0.64</sup> <sub>-0.84</sub>
<b>FORS2/STIS</b>						
No species	1503 <sup>+267</sup> <sub>-369</sub>	1.48 <sup>+0.02</sup> <sub>-0.01</sub>	3.66 <sup>+0.18</sup> <sub>-0.23</sub>	n/a	n/a	n/a
<b>FORS2/STIS + WFC3</b>						
H <sub>2</sub> O	1476 <sup>+277</sup> <sub>-376</sub>	1.48 <sup>+0.02</sup> <sub>-0.01</sub>	3.76 <sup>+0.17</sup> <sub>-0.21</sub>	-5.51 <sup>+0.38</sup> <sub>-0.75</sub>	n/a	n/a

**Notes.** We note that, besides the setups described here, the models also include H<sub>2</sub>, He, a grey cloud deck, and Rayleigh scattering.