THE COMPLETE TRANSMISSION SPECTRUM OF WASP-39b WITH A PRECISE WATER CONSTRAINT

H.R. Wakeford[1,2]
1. Astrophysics Group, University of Exeter, Physics Building, Stocker Road, Devon, EX4 4QL UK and
2. Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

D.K. Sing[1], D. Deming[3], N.K. Lewis[2], J. Goyal[1], T.J. Wilson[1], J. Barstow[4], T. Kataria[5], B. Drummond[1], T.M. Evans[1], A.L. Carter[1], N. Nikolov[1], H.A. Knutson [6], G.E. Ballester[7], A.M. Mandell [8]
3. Department of Astronomy, University of Maryland, College Park, MD 20742, USA
4. University College London, London, UK
5. NASA Jet Propulsion Laboratory, 4800 Oak Grove Dr, Pasadena, CA 91109, USA
6. Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA
7. Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA and
8. NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

ABSTRACT

WASP-39b is a hot Saturn-mass exoplanet with a predicted clear atmosphere based on observations in the optical and infrared. Here we complete the transmission spectrum of the atmosphere with observations in the near-infrared (NIR) over three water absorption features with the Hubble Space Telescope (HST) Wide Field Camera 3 (WFC3) G102 (0.8–1.1 μm) and G141 (1.1–1.7 μm) spectroscopic grisms. We measure the predicted high amplitude H_2O feature centered at 1.4 μm, and the smaller amplitude features at 0.95 and 1.2 μm, with a maximum water absorption amplitude of 2.4 planetary scale heights. We incorporate these new NIR measurements into previously published observational measurements to complete the transmission spectrum from 0.3–5 μm. From these observed water features, combined with features in the optical and IR, we retrieve a well constrained temperature $T_{eq} = 1030^{+30}_{-20}$ K, and atmospheric metallicity $151^{+48}_{-46} \times$ solar which is relatively high with respect to the currently established mass-metallicity trends. This new measurement in the Saturn-mass range hints at further diversity in the planet formation process relative to our solar system giants.

Subject headings: techniques: spectroscopic, planets and satellites: atmospheres, planets and satellites: individual (WASP-39b)

1. INTRODUCTION

Exoplanets have greatly advanced our understanding of planetary systems, expanding theory and observations beyond our solar system (e.g. Seager & Deming 2010; Fortney et al. 2010; Marley et al. 2013; Deming et al. 2013; Stevenson et al. 2014; Sing et al. 2016; Kataria et al. 2016; Evans et al. 2016; Deming & Seager 2017). Due to remote observing techniques close-in giant planets dominate atmospheric characterization studies, in part because of their large planet-to-star radius ratios, but mostly due to large extended atmospheres which scatter and transmit a greater number of photons. Observations using transmission spectroscopy have measured atomic (Na and K) and molecular (H_2O) absorption in a range of exoplanet atmospheres with ground- and space-based telescopes (e.g. Charbonneau et al. 2002; Sing et al. 2011, 2015; Deming et al. 2013; Kreidberg et al. 2014; Stevenson et al. 2014; Nikolov et al. 2016; Sing et al. 2016; Evans et al. 2016). While water has been detected in a number of exoplanet atmospheres, precise and constraining measurements of H_2O absorption are still rare (e.g. Kreidberg et al. 2014; Wakeford et al. 2017).

Observational data analysis is a non-trivial exercise and there are an array of techniques being applied to reduce data (e.g. Berta et al. 2012; Gibson et al. 2012; Deming et al. 2013; Wakeford et al. 2016). In a com-

parative observational study of ten hot Jupiter exoplanets using the Hubble and Spitzer Space Telescopes with consistent data analysis techniques, Sing et al. (2016) present a startling diversity in atmospheric transmission spectral features, where clouds play a significant role muting and obscuring atomic and molecular absorption features. Sing et al. (2016) defined a transmission spectral index as a means to distinguish between cloudy and clear atmospheres using the amplitude of water absorption observed with Hubble Space Telescope (HST) Wide Field Camera 3 (WFC3) versus the altitude difference between optical wavelengths, measured with Space Telescope Imaging Spectrograph (STIS), and IR wavelengths, measured with Spitzer Space Telescope Infrared Array Camera (IRAC). The transmission spectral index effectively displayed trends in the 'clarity' of an exoplanet atmosphere with a continuum from clear to cloudy for the ten planets in the study. One of the planets in this study is the highly inflated Saturn-mass planet WASP-39b.

WASP-39b has a radius of 1.27 R_J and a mass of just 0.28 M_J, and is in orbit around a late G-type star with a period of 4.055 days (Faedi et al. 2011). Out of current well-characterized exoplanets, only the hot Jupiters WASP-31b and WASP-17b have lower bulk densities (Anderson et al. 2010; Anderson et al. 2011; Sing et al. 2016). However, both are more massive and more irradiated than WASP-39b, which has a lower equilibrium
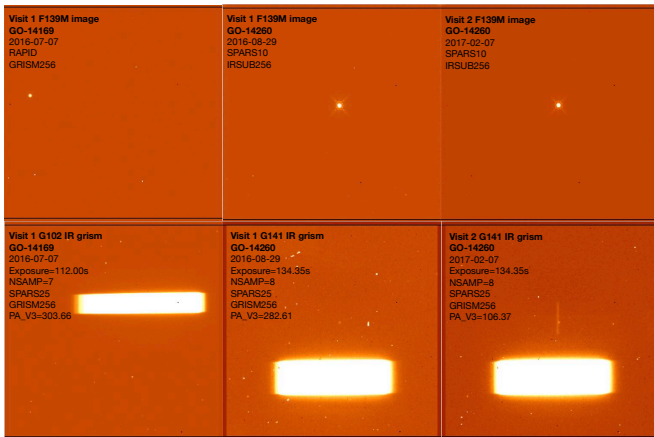
FIG. 1.— *"IMA"* image files of WASP-39 from each visit. The top row shows the direct image taken of the target star with the WFC3 F139M filter, which is used for wavelength calibration of the spectroscopic trace. The bottom row shows the first exposure of each visit: from left to right we show the G102 grism trace, G141 visit 1 trace, and G141 visit 2 trace. Each of the independent observational parameters are listed on the figures.

temperature of 1116 K. Analysis of STIS data in the optical and *Spitzer*/IRAC data in the IR shows a spectrum consistent with a predominately clear atmosphere based on both forward model and retrieval studies (Sing et al. 2016; Barstow et al. 2017). The strongest evidence of a clear atmosphere from the current observational data of WASP-39b is the presence of a strong Na feature in the transmission spectrum with pressure broadened line wings for both Na and K absorption lines (Fischer et al. 2016; Nikolov et al. 2016). These alkali features would be muted or obscured if clouds were present, and would not extend to the H/He continuum. Near-infrared (NIR) features, on the other hand, would be less sensitive to these clouds, given the inherent wavelength dependence of light scattered by small particles (Bohren & Huffman 2008). At the time of the Sing et al. (2016) study, no NIR HST/WFC3 observations existed for WASP-39b, and the atmospheric water absorption could not be measured. Therefore, the 'clarity' of the atmosphere (i.e, the transmission spectral index) could not be calculated. In this paper we present three transit observations of WASP-39b using WFC3 from 0.8 to 1.7 μm and place precise constraints on the water content in the atmosphere of the planet and calculate a definitive transmission spectral index. In section 2 we detail the analysis of the new observations in the NIR. We then present the complete transmission spectrum of WASP-39b from 0.3–5 μm in section 3 and introduce the theoretical models used to interpret this dataset. We detail the retrieved atmospheric parameters in section 4 and discuss the implications these have on the nature of WASP-39b with respect to temperature, metallicity, and chemical abundances, as well as planetary formation.

2. WFC3 OBSERVATIONS AND ANALYSIS

We observed WASP-39b during three transit events using HST WFC3 G102 grism as part of GO-14169 (PI. Wakeford) on 7 July 2016, and WFC3 G141 grism as part of GO-14260 (PI. Deming) on 29 August 2016 and 7 February 2017. Observations for both spectroscopic grisms were conducted in forward spatial scan mode.

Spatial scanning involves exposing the telescope during a forward slew in the cross-dispersion direction and resetting the telescope position to the top of the scan prior to conducting subsequent exposures. Scans with G102 were conducted at a scan rate of ∼0.26 pixels per second with a final scan covering ∼28 pixels in the cross-dispersion direction. For G141, we used a scan rate of ∼0.30 pixels per second with a final spatial scan covering ∼44 pixels in the cross-dispersion direction on the detector (Fig. 1).

We use the IMA output files from the CalWF3 pipeline which are calibrated using flat fielding and bias subtraction. We extract the spectrum from each exposure by taking the difference between successive non-destructive reads. A top-hat filter (Evans et al. 2016) is then applied around the target spectrum and all external pixels are set to zero to aid the removal of cosmic rays (Nikolov et al. 2015). The image is then reconstructed by adding the individual reads back together. We then extract the stellar spectrum from each exposure with an aperture of ±14 pixels for G102 and ±22 pixels for G141 around a centering profile, which was found to be consistent across the spectrum for each exposure for each observation.

We monitored each transit with HST over the course of five orbits with observations occurring before, during, and after transit. We discard the first orbit of each visit as it contains vastly different systematics to the subsequent orbits (e.g. Deming et al. 2013; Kreidberg et al. 2015; Wakeford et al. 2016; Sing et al. 2016; Wakeford et al. 2017). This is due to thermal settling required after target acquisition which results in a different thermal pattern for the telescope not replicated in subsequent orbits.

2.1. *White lightcurve*

We first analyse the band-integrated light curves of WASP-39, to obtain the broad-band planet-to-star radius ratio (R_p/R_*), by summing the flux between 0.8 and 1.13μm for G102, and between 1.12 and 1.66μm for G141. The uncertainties on each data point were initially set to pipeline values dominated by photon and readout noise. We account for stellar limb darkening using a 4-parameter limb-darkening law (Claret 2000; Sing 2010). Each light curve is fit using the IDL routine MPFIT which uses a Lavenberg-Markwardt (L-M) least-squares algorithm (Markwardt 2009), which has been shown to produce posterior distributions that are well described by a multi-variable Gaussian distribution (e.g. Sing et al. 2016; Wakeford et al. 2016; Wakeford et al. 2017). After an initial fit, the uncertainties on each time series exposure are rescaled based on the standard deviation of the residuals, taking into account any underestimated uncertainties calculated by the reduction pipeline in the data points. The final uncertainty for each point is then determined from the covariance matrix from the L-M algorithm. We fix the system parameters to the previously established values: inclination = 87.36°, a/R_* = 11.043, period = 4.055259 days (Nikolov et al. 2016). From the band-integrated lightcurve we fit for the center of transit time and fix it for subsequent spectroscopic lightcurves.

We use marginalization across a series of systematic models to account for observatory- and instrument-based systematics (Gibson 2014). As detailed in Wakeford et al. (2016), we use a grid of polynomial models, which approximates stochastic models, to account for estab-
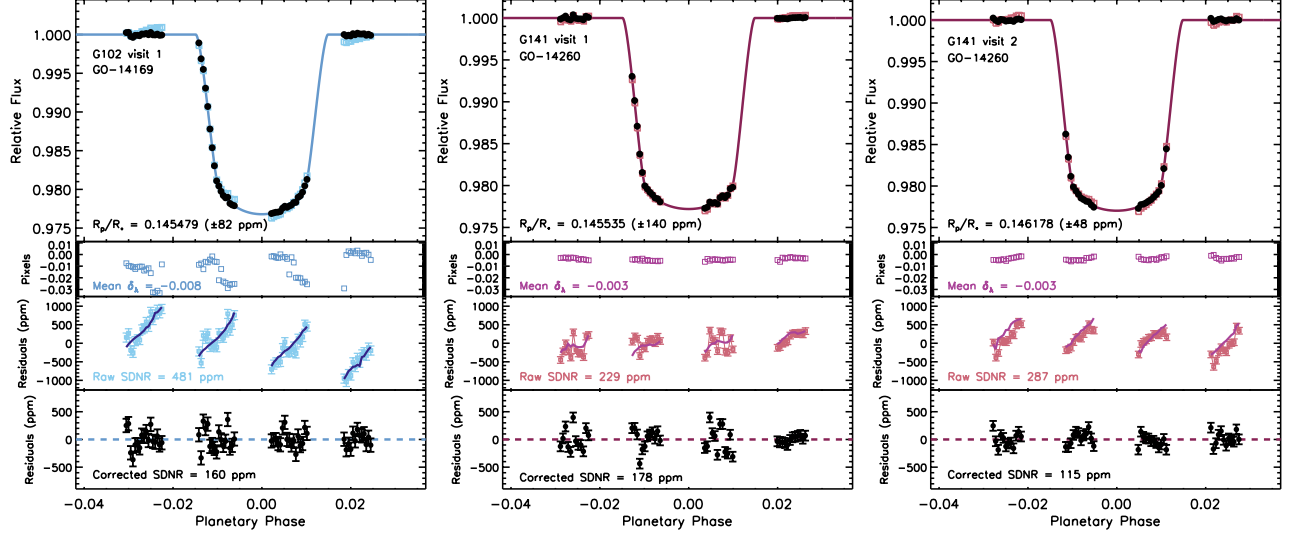
FIG. 2.— White light curves of the WASP-39b transit in each visit. Top: The raw (colored) and corrected (black) lightcurves with the best fit model (solid line). Middle top: δ_λ which is the positional shift on the detector between each exposure spectrum. Middle bottom: Raw residuals and uncertainties (colored points), with the best fit systematic model (solid lines). Bottom: Corrected lightcurve residuals (black points). a) G102, b) G141 visit 1, c) G141 visit 2.
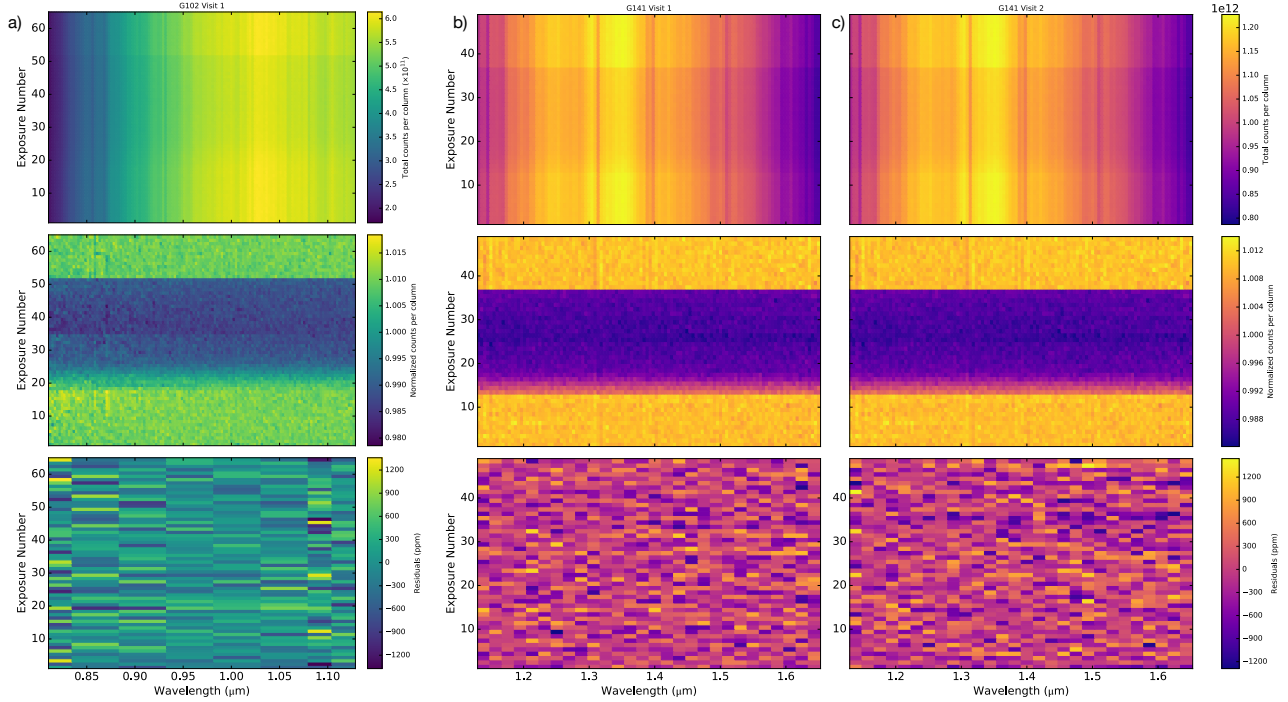

FIG. 3.— Pixel maps for all three observations of WASP-39 with WFC3. Columns a) G102, b) G141 visit 1, and c) G141 visit2. Top row: shows the stellar spectra for each visit following extraction and cosmic ray removal, but prior to systematic corrections, in units of total counts per pixel column. Middle row: shows the stellar spectra normalized by the average counts per wavelength column which enhances the contrast to the in transit exposures. Bottom row: shows the residuals for each spectroscopic light curve after systematic correction.

lished systematic trends. Each systematic models accounts for up to three detrending parameters; θ, ϕ, and δ_λ, where θ accounts for a linear increase/decrease of flux in time across the whole observation, ϕ is the HST orbital phase trends based on the thermal 'breathing' of the telescope over the course of each orbit around the Earth, and δ_λ is the shift in wavelength position of the spectrum caused by telescope pointing. We use combinations with and without θ, and with or without ϕ and δ_λ, up to the

4th order. This results in a grid of 50 systematic models which we test against our data, where the most complex has the following function:

$$S(t, \lambda) = T_1 \, \theta \times \sum_{i=1}^{n} p_i \phi^i \times \sum_{j=1}^{n} l_j \delta_\lambda^j. \qquad (1)$$

Here, T_1, p_i, and l_j are fit variables or fixed to zero if unfit in a specific model (see Wakeford et al. 2016 for
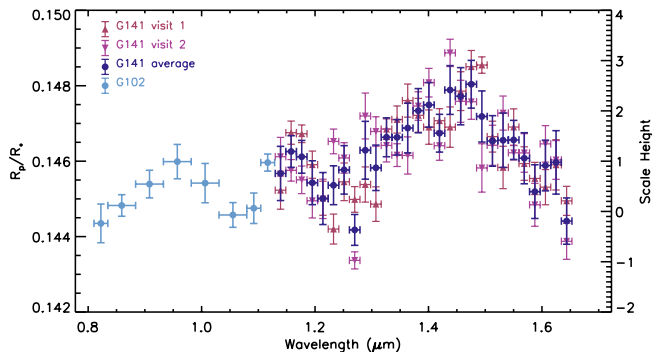
FIG. 4.— WFC3 transmission spectrum measured over three visits with G102 and G141 grisms. The measured transmission spectrum from the G102 visit from 0.8–1.12 μm is shown in light blue circles. We analyzed the two visits with G141 separately and show each transmission spectrum (visit 1 triangles, visit 2 downward triangles), and the transmission spectrum computed from the average of the separate spectra (dark red circles). The G141 spectrum is computed from 1.12–1.66 μm with a constant bin width of $\Delta\lambda = 0.0186\,\mu$m.
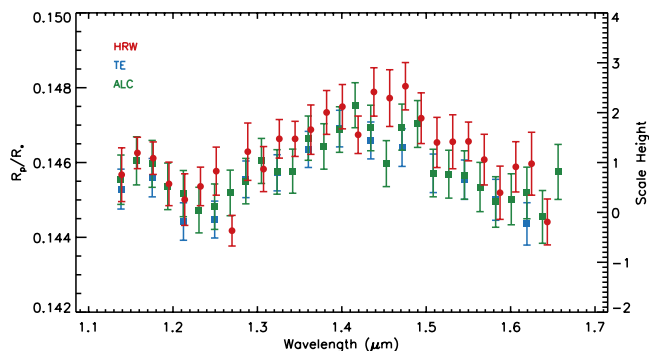
FIG. 5.— Comparison of the average transmission spectrum computed for the G141 visits based on three different analyses.

a table all parameter combinations). It is important to note that marginalization relies on the fact that at least one of the models being marginalized over is a good representation of the systematics in the data. Marginalizing over a series of models then allows for greater flexibility in satisfying this condition, over the use of a single systematic model. We use the maximum likelihood estimation (MLE) based on the Akaike Information Criterion (AIC) to approximate the evidence-based weight for each systematic model (Burnham & Anderson 2004). While the AIC is an approximation to the evidence, it allows for more flexible models to be folded into the likelihood, as compared to the Bayesian Information Criterion (BIC), which typically leads to more conservative error estimates on the marginalized values (Gibson 2014; Wakeford et al. 2016). We marginalize across all systematic models to compute the desired light curve parameters. Using marginalization across a large grid of models allows us to account for all tested combinations of systematics and obtain robust center of transit times from the band-integrated light curve, and transit depths for each spectroscopic light curve. We measure a center of transit time of 2457577.425837±0.000044 (JD$_{UTC}$) for the G102 visit, 2457630.149461±0.000044 (JD$_{UTC}$) for visit 1 with G141, and 2457792.356338±0.000041 (JD$_{UTC}$) for visit 2 with G141. We show the white light curves for each visit in Fig.2 along with the highest weighted model from our

analysis and the marginalized broadband transit depth in terms of R_p/R_*. To demonstrate the noise properties of the data, we show the pixel maps of the stellar spectra following spectral extraction, cosmic ray corrections, and wavelength alignment (Fig.3) in both total counts (top) and in normalized counts (middle). The bottom panel of Fig.3 shows the residuals from each spectroscopic light curve, the analysis of which is outlined in the following section. These plots demonstrate that there are no obvious bad pixels in these data, and no wavelength-dependent trends in the spectroscopic lightcurves.

2.2. *Spectroscopic light curves*

We divide the WFC3 wavelength range of each grism into a series of bins and measure the R_p/R_* from each spectroscopic light curve (Table 1) following the same procedure as detailed for the band-integrated light curve. For each visit we test a range of bin widths and wavelength ranges. For G102, we test bin widths from $\Delta\lambda = 0.0146$–$0.0490\,\mu$m over wavelength ranges of 0.8–1.12 μm and 0.81–1.13 μm. For G141, we test bins of $\Delta\lambda = 0.0186$, 0.0128, 0.0373, and 0.0467 μm, over two wavelength ranges (1.12–1.65 μm 1.13–1.66 μm). For each bin and wavelength range we find that the shape of the transmission spectrum is robust and consistent.

For the presented transmission spectrum we divide the G102 spectroscopic range into 8 bins between 0.81 and 1.13 μm with variable bin sizes of $\Delta\lambda = 0.0244$ or 0.0490 μm (Table 1). For each spectroscopic lightcurve we test all 50 models and marginalize over the transit parameters to compute the R_p/R_*. We find that there are few wavelength-dependent systematics as evidenced by the similarity between the highest weighted systematic model for each spectroscopic lightcurve fit.

For the G141 grism observations we extract and analyse each visit separately. As with the band-integrated analysis, we marginalize over all systematic models to compute the transit parameters. For both G141 visits, we use 28 equal bin sizes of $\Delta\lambda = 0.0186\mu$m between 1.13 and 1.66$\mu$m (Table 1). We found both transmission spectra using the G141 grism to be consistent in absolute depth (Fig. 4). We then combine them into a single spectrum by taking the weighted mean of each spectroscopic measurement (Table 1).

To determine if the analysis method used is justified we compute the transmission spectrum from each visit using a different analysis pipeline (Evans et al. 2016) over multiple bin sizes (T. Evans (TE): $\Delta\lambda$=0.0500 μm, A.L. Carter (ALC): $\Delta\lambda$=0.0186 μm) and compute the weighted mean spectrum. We find that the different methods result in the same transmission spectrum in shape and absolute depth within the 1σ uncertainties (Fig. 5). The additional analyses performed used similar techniques to extract the stellar spectra, however, differ in their lightcurve analysis. TE and ALC used Gaussian process (GP) analysis with a Matern v=3/2 kernel (see Evans et al. 2017) on the white light curve to obtain the common-mode systematics. For each spectroscopic light curve, TE used the white ligthcurve systematics and divided each spectroscopic lightcurve by the residuals with the addition of simple linear corrections in time to fit to each wavelength bin (e.g. Deming et al. 2013; Evans et al. 2017). Following GP analysis of the white lightcurve, ALC used the stretch and shift method to re-

TABLE 1
Marginalized transmission spectrum of WASP-39b measured with HST WFC3 G102 and G141 grism. λ marks the center of the bin with $\Delta\lambda$ representing the total width of the bin.

λ	$\Delta\lambda$	R_p/R_*	Uncertainty
μm	μm		(ppm)
	– G102 –		
0.8225	0.0244	0.14435	310
0.8592	0.0490	0.14482	190
0.9082	0.0490	0.14539	140
0.9572	0.0490	0.14598	150
1.0062	0.0490	0.14541	130
1.0552	0.0490	0.14457	150
1.0920	0.0244	0.14475	220
1.1165	0.0244	0.14596	240
	– G141 –		
1.1391	0.0186	0.14567	710
1.1578	0.0186	0.14625	410
1.1765	0.0186	0.14611	430
1.1951	0.0186	0.14542	580
1.2138	0.0186	0.14500	680
1.2325	0.0186	0.14536	510
1.2512	0.0186	0.14576	640
1.2699	0.0186	0.14417	400
1.2885	0.0186	0.14628	760
1.3072	0.0186	0.14582	602
1.3259	0.0186	0.14663	515
1.3446	0.0186	0.14663	469
1.3633	0.0186	0.14687	660
1.3819	0.0186	0.14733	580
1.4006	0.0186	0.14749	592
1.4193	0.0186	0.14674	500
1.4380	0.0186	0.14788	650
1.4567	0.0186	0.14772	758
1.4753	0.0186	0.14803	633
1.4940	0.0186	0.14718	677
1.5127	0.0186	0.14653	670
1.5314	0.0186	0.14655	721
1.5501	0.0186	0.14656	518
1.5687	0.0186	0.14607	677
1.5874	0.0186	0.14519	711
1.6061	0.0186	0.14588	671
1.6248	0.0186	0.14596	840
1.6435	0.0186	0.14441	614

move common-mode systematics from each spectroscopic lightcurve (see Deming et al. 2013; Wakeford et al. 2016). Both additional methods were computed at different bin sizes and wavelength positions to further test the robust nature of the computed transmission spectrum (Fig. 5).

We use these new transmission spectral measurements from G102 and G141 between 0.8–1.66 μm (Table 1) to complete the transmission spectrum of WASP-39b from the optical to the IR.

3. WASP-39b's COMPLETE TRANSMISSION SPECTRUM

We detect distinct H_2O absorption in three bands centered at 0.9, 1.15, and 1.4 μm in the new HST WFC3 transmission spectral data (Fig. 6). We combine these directly with the previously published HST STIS and *Spitzer* IRAC data (Sing et al. 2016), and VLT FORS2 data (Nikolov et al. 2016), without the need of an offset in absolute depth. As the VLT FORS2 measurements match in wavelength position and depth to the HST STIS data between 0.41–0.81 μm (Nikolov et al. 2016) we take the weighted mean between these measurements, and again do not apply an offset in the absolute depth measured. An offset between datasets can be caused by stellar or planetary variability. However, for an inactive star like WASP-39A (R_{HK} = -4.994), it is more likely caused by differences in the reduction and analysis method used. In our method we apply a consistent analysis across all datasets using marginalization across systematic model grids. The STIS and FORS2 analyses include a common-mode correction; however, both match exactly in depth and wavelength position with overlapping regions with the new WFC3 data, suggesting that any common mode corrections have no significant impact on the absolute transit depth measured. Additionally, each analysis applied the same system parameters as detailed in section 2.1.

We use the complete transmission spectrum of WASP-39b to interpret this inflated Saturn's atmosphere through forward models, retrievals, and 3D GCM simulations. In the following subsections we outline each of these theoretical models and discuss the implications of the results in §4.

3.1. *Transmission spectral index*

We can now compare WASP-39b to the Sing et al. (2016) sample by computing the transmission spectral index from the H_2O amplitude and the $\Delta Z_{J-LM}/H_{eq}$ altitude difference (see methods section of Sing et al. 2016). The H_2O amplitude is calculated by assuming a clear solar atmospheric model, with no cloud or haze opacities, and scaling it in amplitude to fit the observed feature. The $\Delta Z_{J-LM}/H_{eq}$ altitude difference compares the relative strength of the continuum in the near-IR (1.22–1.33 μm) to the mid-IR absorption calculated from the average of the two *Spitzer* points. We measure an H_2O amplitude of 37±7% with $\Delta Z_{J-LM}/H_{eq}$ = -0.09±0.35. A near zero altitude difference and relatively low water amplitude places WASP-39b further away from the clear solar abundance model, although this does not change its position in the clear to cloudy continuum displayed in Sing et al. (2016) as compared to the other nine exoplanets in the study. While this index shows that the water amplitude is less than expected for a completely clear solar abundance atmosphere, each of these model tracks varies only a single parameter, with all others held fixed; in reality, each of these parameters will vary simultaneously. Additionally, previous observations and theoretical models of extrasolar gas giant planets show that these planets may have metallicities greater than 1× solar (e.g. Kreidberg et al. 2014; Thorngren et al. 2016; Wakeford et al. 2017). We comment further on this planet's metallicity in §4.5.
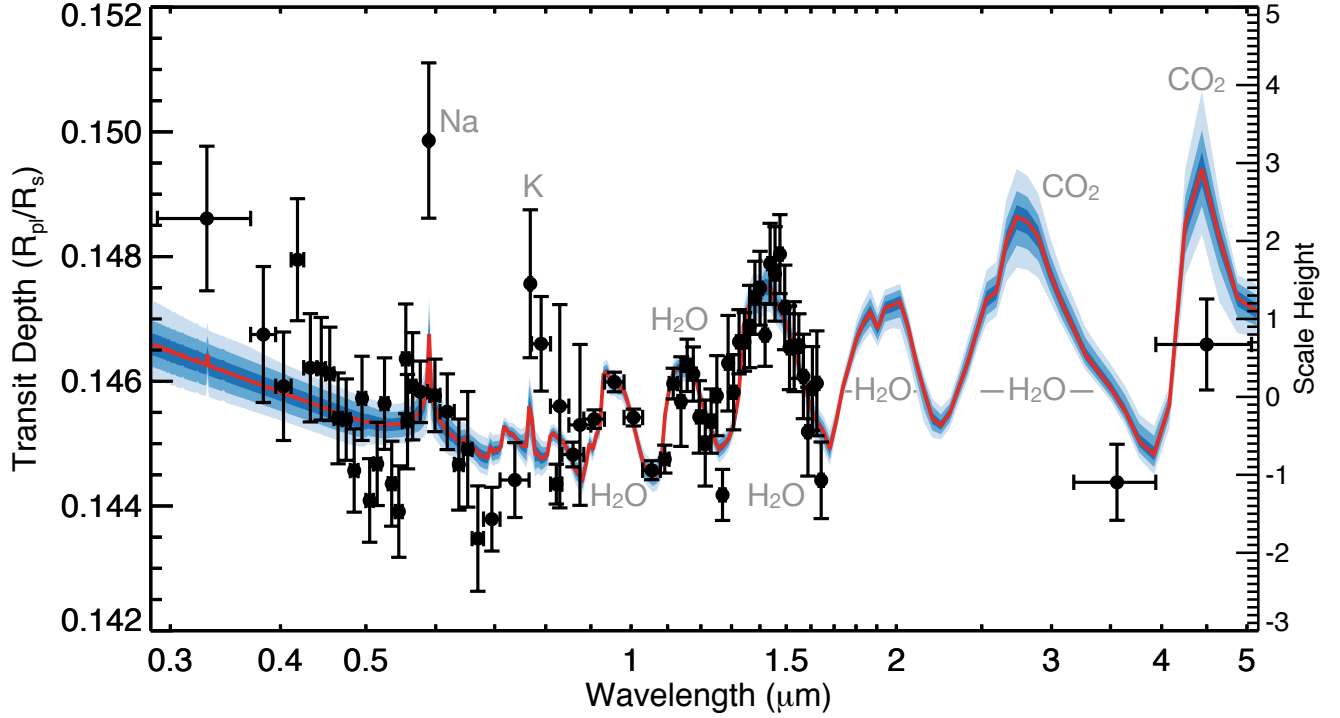
FIG. 6.— The complete transmission spectrum of WASP-39b (black points). This transmission spectrum incorporates data from HST STIS and WFC3, *Spitzer* IRAC, and VLT FORS2 completing the spectrum from 0.3–5.0 μm with currently available instruments. Using the ATMO retrieval code, which implements an isothermal profile and equilibrium chemistry, we determine the best fit atmospheric model (red) and show the 1, 2, and 3σ confidence regions (dark to light blue) based on the retrieved parameters.
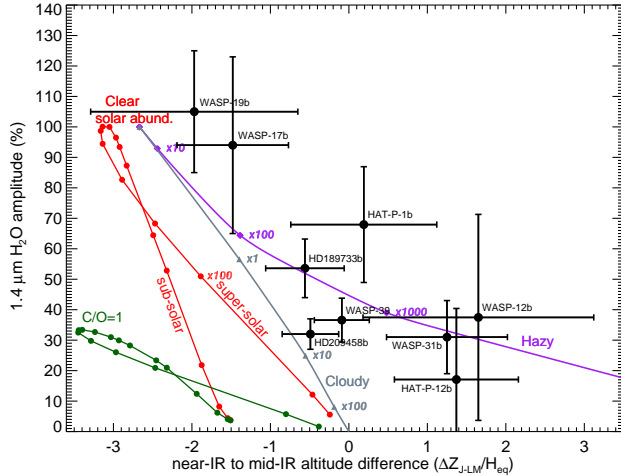

FIG. 7.— Transmission spectral index diagram of ΔZ_{J-LM} versus H_2O amplitude as defined in Sing et al. (2016). Black points show the altitude difference between the NIR and IR spectral features (ΔZ_{J-LM}) versus the H_2O amplitude measured at 1.4 μm, each with 1σ errorbars. Purple and grey lines show the model trends for haze and clouds respectively. Red shows the trend for sub- and super-solar metallicities. Green shows the model trend for different C/O.

3.2. *Goyal Forward Model Grid*

We use the newly developed open source grid of forward model transmission spectra produced using the 1D radiative-convective equilibrium model ATMO (Amundsen et al. 2014; Tremblin et al. 2015, 2016; Drummond et al. 2016) outlined in Goyal et al. (2017), to compare to our measured transmission spectrum. Each model assumes isothermal pressure-temperature (P-T) profiles

and equilibrium chemistry with rainout condensation. It includes multi-gas Rayleigh scattering and high temperature opacities due to H_2O, CO_2, CO, CH_4, NH_3, Na, K, Li, Rb, Cs, TiO, VO, FeH, PH_3, H_2S, HCN, C_2H_2, SO_2, as well as H_2-H_2, H_2-He collision-induced absorption (CIA). The grid consists of 6,272 model transmission spectra specifically for WASP-39b (i.e. with gravity, R_p, etc. of WASP-39b), which explores a combination of eight temperatures (516 K, 666 K, 741 K, 816 K, 966 K, 1116 K, 1266 K, and 1416 K), seven metallicities (0.005, 0.1, 1, 10, 50, 100, 200× solar), seven C/O values (0.15, 0.35, 0.56, 0.70, 0.75, 1.0, and 1.5), four "haze" parameters (1, 10, 150, and 1100), and four cloud parameters (0, 0.06, 0.2, and 1). In the Goyal grid the "haze" parameter defines an enhanced Rayleigh-like scattering profile, which increases the hydrogen cross section with a wavelength dependent profile. The "cloud" parameter defines a grey uniform scattering profile across all wavelengths between 0 and 100% cloud opacity (see Goyal et al. 2017 for more details).

We fit each model to the transmission spectrum by only allowing them to move in absolute altitude; we therefore have one free parameter for each model fit. We use the L-M routine MPFIT to determine the best fit altitude and calculate the χ^2. We transform from χ^2 to probability likelihood via the expression

$$p(\mathbf{x}) = A \exp\left(-\frac{1}{2}\chi^2(\mathbf{x})\right), \qquad (2)$$

where $\mathbf{x} \equiv \{T, [M/H], C/O, H, C\}$ are the temperature, metallicity, carbon-to-oxygen ratio, haze, and cloud parameters respectively and A is a normalization constant.
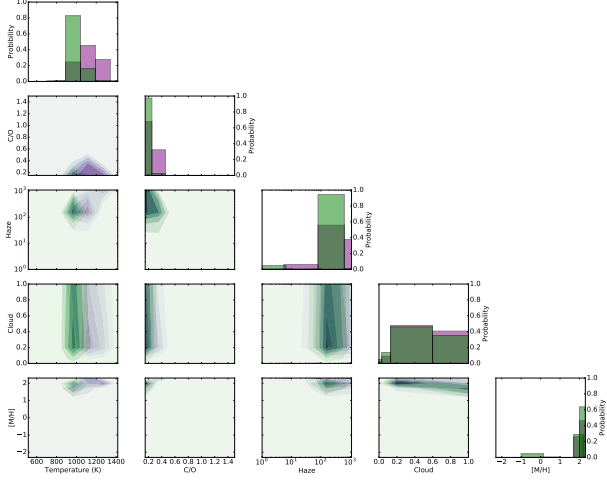
FIG. 8.— Probability distribution of the forward model grid fit to the data for the new WFC3 results only (purple) and the full transmission spectrum including STIS, VLT, and *Spitzer* points (green). These pairs plots show the correlations between the five parameters explored in the model grid with the respective probability histograms for each parameter.
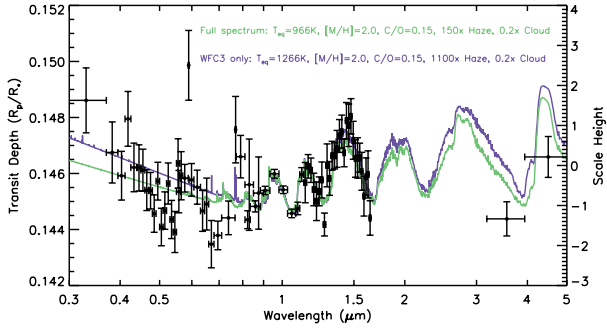

FIG. 9.— Highest probability grid models fit to the the transmission spectrum for, the new WFC3 results only (purple) and the full transmission spectrum (green) including STIS, VLT, and *Spitzer* points.

This constant is set by integrating the likelihood over all model parameters, such that

$$\int \cdots \int p(\mathbf{x}) \, d\mathbf{x} = 1. \qquad (3)$$

The Goyal grid is specifically generated for each planet and the range of each of the parameters are reasonable assumptions for a planetary atmosphere. As such we assume a weakly informative prior on each model, which is uniform to the edge of the grid space where it is truncated.

We fit the model grid to both the full transmission spectrum (0.3–5 μm) and a subset containing only the new WFC3 data (0.8–1.7 μm). The probability distribution of the model grid is shown in Fig. 8 for both scenarios. We then select the model with the highest probability for each dataset and plot them in Fig. 9 against the complete transmission spectrum of WASP-39b. From these plots we see a distinct difference between the best fit temperature and haze parameters with just the H$_2$O features, and with the full optical and NIR transmission spectrum. These differences predominantly impact the optical portion of the spectrum which anchors the haze

parameter, but is largely degenerate with the temperature. Additionally, the haze and cloud parameters have a widespread influence on the C/O and metallicity, which is further highlighted by the inclusion of the optical and IR data. We discuss this further in §4.

3.3. *ATMO Retrieval*

We use the ATMO Retrieval Code (ARC) to fully explore the parameter space covered by the transmission spectral measurements. Using ARC we calculate the posterior distribution of the data to the models, and determine the fit confidence intervals marginalizing over parameter space (see Wakeford et al. 2017 and Evans et al. 2017 for further details). ARC couples the ATMO model to a L-M least-squares minimizer and a Differential Evolution Chain Monte Carlo (DECMC) analysis (Eastman et al. 2013). We ran DECMCs with 12 chains each with 30,000 steps, with convergence typically occurring around 15,000 steps, where convergence is monitored using the Gelman-Rubin statistic. We considered two atmospheric retrieval models: a chemically consistent retrieval scheme which assumes chemical equilibrium, and a more flexible free-chemistry retrievals. In the equilibrium chemistry retrieval, the chemical abundances are consistent with the pressure-temperature profile. In the free-chemistry retrieval, the abundances are assumed to be constant with pressure and are independently fit. For the opacity sources, we include H$_2$-H$_2$, H$_2$-He CIA, and molecular opacities from H$_2$O, CO, CO$_2$, CH$_4$, NH$_3$, Na, and K. For both models we assume isothermal P-T profiles. Parameterized P-T profiles were explored, although found to be unnecessary to fit the data, as the retrieved P-T profiles were found to be isothermal over the altitudes probed by the data with no significant structure in the P-T profile indicated by the fit parameters.

In each retrieval we fit for a haze and cloud parameter, each of which are scaling factors applied to the hydrogen cross-section with either a wavelength dependent profile (haze) or uniform profile (cloud) - these two parameters are represented in ARC (Fig. 10) as a natural log ratio for numerical reasons only. Using the equilibrium retrieval we also run a fit to only the new WFC3 data, as with the Goyal grid, to determine the impact of the optical and IR transmission spectrum in the retrieved parameters.

In Table 2 we list the main results from the ARC analysis, with each model having the following free parameters, respectively:
Equilibrium: planetary radius (R$_p$), T$_{eq}$, Haze, Cloud, [M/H], C/O.
Free-chemistry: R$_p$, T$_{eq}$, Haze, Cloud, H$_2$O volume mixing ratio (VMR), CO$_2$ VMR, CO VMR, CH$_4$ VMR, Na VMR, and K VMR.
From this retrieval we are able to place constraints on the equilibrium temperature of the observed portion of the planetary atmosphere, as well as the C/O and metallicity by fitting the absorption features. In the equilibrium chemistry case the abundance of elemental carbon and oxygen species is set via the C/O and metallicity. Specifically in ATMO, the carbon abundance is set as a multiple of the solar carbon (C) abundance:

$$A(\text{C}) = A(\text{C}, solar) \times 10^{[\text{M/H}]}, \qquad (4)$$

TABLE 2
RESULTS FROM EACH RETRIEVAL MODEL

| Data	Model Chemistry	k	DOF	χ^2	BIC	T_{eq} (K)	M/H [$\log_{10}$] ×solar	C/O	H_2O [VMR] ×solar	Na [VMR] ×solar	K [VMR] ×solar
WFC3	Equilib.	5	31	31.9	49.8	950^{+130}_{-150}	$1.88^{+0.41}_{-0.65}$ 76^{+117}_{-59}	$0.19^{+0.29}_{-0.16}$			
Full	Equilib.	5	65	88.1	109.0	1030^{+30}_{-20}	$2.18^{+0.12}_{-0.16}$ 151^{+48}_{-46}	$0.31^{+0.08}_{-0.05}$			
Full	Free-Chem.	10	60	85.8	128.3	920^{+70}_{-60}	$2.07^{+0.05}_{-0.13}$ 117^{+14}_{-30}	$0.44^{+0.11}_{-0.44}$	$-1.37^{+0.05}_{-0.13}$ 117^{+14}_{-30}	$-3.7^{+0.6}_{-0.6}$ 120^{+330}_{-90}	$-4.9^{+0.8}_{-1.3}$ 105^{+620}_{-100}

[VMR] is $\log_{10}$(Volume Mixing Ratio)
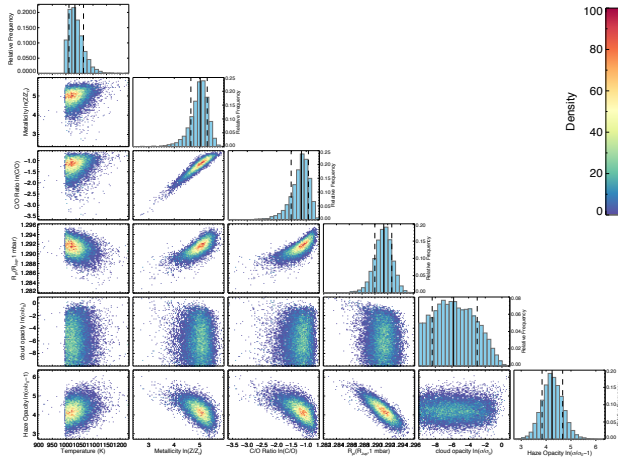
FIG. 10.— Probability density map of the ARC fit to the data for the the full transmission spectrum, using equilibrium chemistry and an isothermal P-T profile. These pairs plots show the correlations between the six parameters explored in the ARC DECMC with the respective probability density histograms for each parameter.

and then the oxygen (O) abundance is set via the carbon abundance and the CO ratio:

$$A(O) = A(C)/CO_{ratio}. \qquad (5)$$

In the free chemistry model each species abundance is fit for independently, and [M/H] is estimated based on the H_2O abundance. The C/O ratio is then estimated from the four fit molecules containing carbon and/or oxygen (CO, CO_2, CH_4, H_2O) and does not take into account any other species. The Na and K line profiles are modeled using an Allard profile (Allard et al. 2007); alternate profiles were tested but did not improve upon the statistical fit to the data.

We find that the data are described best with an isothermal equilibrium model which has a χ^2_ν of 1.32 compared to 1.45 for the free chemistry fit. From this we retrieve $T_{eq}=1030^{+30}_{-20}$ and an atmospheric metallicity $151^{+48}_{-46}\times$ solar, as defined by the posterior distributions (Fig.10).

3.4. 3D GCM

To model the three-dimensional (3D) temperature structure of WASP-39b we use the SPARC/MITgcm (Showman et al. 2009). The SPARC/MITgcm couples the MITgcm, a finite-volume code that solves the 3D primitive equations on a staggered Arakawa C grid (Adcroft et al. 2004) with a two-stream adaptation of a multi-stream radiative transfer code for solar system planets (Marley & McKay 1999). The radiative transfer code employs the correlated-k method with 11 bands optimized for accuracy and computational efficiency. The opacities are calculated assuming local thermodynamic and chemical equilibrium. This code has been used extensively to model the atmospheric circulation of exoplanets (e.g., Lewis et al. 2010; Kataria et al. 2015, 2016; Wakeford et al. 2017; Lewis et al. 2017).

Here we show the P-T profiles from Kataria et al. (2016) for WASP-39b averaged over different regions of the atmosphere, demonstrating the impact of 3D circulation on the planetary P-T profiles (Fig. 11a). We also plot the retrieved temperature based on isothermal equilibrium models at the pressure probed by these observations. Also shown are the condensation curves of potential cloud species in the atmosphere of WASP-39b, as well as the molecular transition region from CO-dominated carbon chemistry to CH_4-dominated carbon chemistry. From the 3D GCM we derive a condensation map showing the average temperature as a function of longitude and pressure indicating where and what condensates might form in the atmosphere (Fig. 11b). The GCM results suggest that the two limbs of the planet likely have different cloud properties due to the recirculation of heat around the planet from the dayside hemisphere. The nightside trailing limb (west) is approximately 100 K colder than the sun trailing limb (east) at pressures of 0.1 bar, and up to 200 K colder at pressures less than 1 mbar. These differences in the limbs not only influence the average temperature observed in transmission, but also in this case will impact the condensate cloud species likely to form in the atmosphere on the colder western limb. This becomes important when considering the impact differences in the limbs will have on the transmission spectrum which measures the average absorption profile of the limb annulus around the planet (e.g. Line et al. 2016).

We note that this model is for a 1 ×solar composition case and with an increased atmospheric metallicity the temperature is likely to increase as well as the position of the condensation curves (Wakeford et al. 2017). Addi-
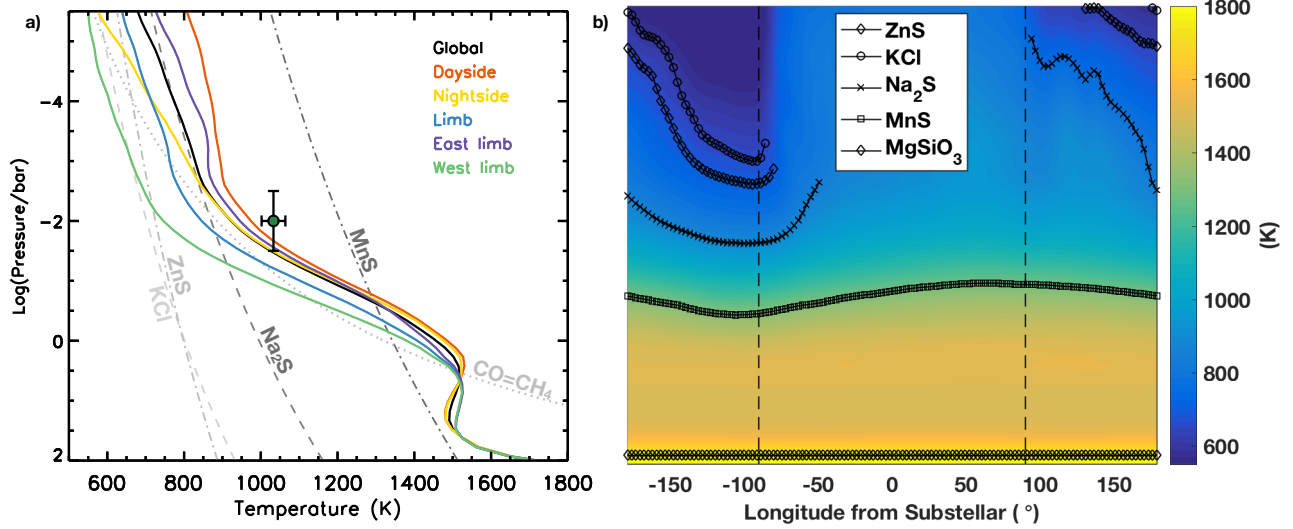
FIG. 11.— a) 1× solar 3D P-T profiles of WASP-39b. The global, dayside, nightside, limb, east, and west limb averaged profiles computed from 3D GCM models of WASP-39b (see Kataria et al. 2016). Also plotted are the condensation curves of the affecting species, as well as the CO/CH₄ abundance line. These P-T profiles indicate a difference in temperature expected on opposite limbs of the planet, resulting in the potential for one limb to be observationally clear while the other contains more significant cloud opacities. The point indicates the temperature retrieved for the atmosphere of WASP-39b based on an isothermal equilibrium model (Table 2). b) Average temperature of WASP-39b as a function of longitude and pressure. Temperatures are weighted by the cosine of the latitude, equivalent to weighting each grid point by its projection angle toward an observer at the equator. Low-temperature condensates are overplotted in this 2D space.

tionally, there is evidence that the day-night temperature contrast will also change (likely increase) with metallicity (Kataria et al. 2014; Charnay et al. 2015, Drummond et al. in prep), which would have consequences on the east-west limb cloud formation.

4. DISCUSSION

We present the complete transmission spectrum of WASP-39b from 0.3–5.0 μm combining data from HST STIS and WFC3, VLT FORS2, and *Spitzer* IRAC. Figure 6 shows the absorption features of WASP-39b's atmosphere which includes broad sodium line wings and three distinct water absorption peaks. There is also tentative evidence of absorption by potassium in the optical and absorption due to carbon-based species in the *Spitzer* 4.5 μm channel.

In our analysis we find that the data are described best with an isothermal equilibrium model with $T_{eq} = 1030^{+30}_{-20}$ and $[M/H] = 151^{+48}_{-46} \times$ solar. At the 1σ level WASP-39b has some of the most constrained atmospheric parameters to date. In this section we discuss the implications from the interpretive methods used and the importance of the complete transmission spectrum.

4.1. *Forward models and retrievals*

To interpret the transmission spectrum of WASP-39b we used a grid of 6,272 forward models that were specifically calculated for WASP-39b (Goyal et al. 2017). The Goyal grid samples five different atmospheric parameters commonly explored in a retrieval code; temperature, metallicity, C/O, scattering haze, and uniform clouds. Based on results from the Goyal grid we find that the data are best described by a high metallicity (100×solar), low C/O (0.15), hazy (150×) atmosphere (Fig. 9). The large sampling of models in the grid allows for more detailed interpretation compared to taking a handful of non-specific exoplanet models, with the addition that

the probability distributions can be explored. The grid probability distributions suggest that uniform clouds are not playing a significant role in the observed transmission spectrum, where any cloud parameter applied to the data is equally probable given a distribution of the other four parameters (Fig. 8). As there are no carbon-bearing species in the transmission spectrum of WASP-39b, the C/O ratio is constrained to the lowest value in the grid. This makes it more difficult to statistically infer precise atmospheric values as the tail end of the distributions are being cut (c.f. Goyal et al. 2017).

Using the ATMO retrieval code (ARC) we implement an isothermal equilibrium chemistry model and expand upon the parameter space covered by the Goyal grid. While the grid contains sparse sampling of all five parameters, the ARC uses DECMC to throughly explore tens of thousands of models. From both the Goyal grid and ARC we find hard cut-offs in the temperature space; this is likely due to the inclusion of the optical data and the presence of sodium absorption in the observed spectrum, which condenses in the models for lower temperatures (see section 4.3). In the Goyal grid, the temperature cut-off appears at 966 K, while the finer sampling used in the ARC leads to a more accurate cut-off of ∼1000 K. Below this temperature, significant amounts of rainout removes Na from the gas phase, which condenses into Na₂S and/or NaAlSi₃O₈.

Using ARC, we can further explore the distribution of C/O and metallicity. While these parameters are highly correlated, ARC is able to explore the limits of the distribution with C/O < 0.15, and [M/H] > 2.7. Interestingly, increased sampling on the cloud parameter yields no further interpretation. One advantage of using the Goyal grid over ARC is speed – the fit to the entire Goyal grid can be run in minutes, while a full retrieval with ARC can take many days. From our analysis, we can see that the Goyal grid is able to constrain planetary parameters

within reasonable limits as described by the current data.

4.2. Interpreting the chemistry

Using ARC we run both an equilibrium chemistry model and a free-chemistry model to fit for chemical species and abundances in the measured atmosphere (see §3.3). Both ATMO chemistry models retrieve similar parameters for temperature, metallicity, and C/O (Table 2). The high resolution and precision of data over the 0.9, 1.2, and $1.4\,\mu$m water absorption features place good constraints on the temperature and metallicity retrieved by each model. These are further reinforced in the free-chemistry fit by the individual abundances retrieved for the Na and K absorption features in the optical, which have similar volume mixing ratios (VMR) to the H_2O, and therefore inform the overall atmospheric metallicity. The presence of Na in the transmission spectrum also informs the temperature of the atmosphere, as at low temperatures ($<1000\,$K) Na would likely be depleted by rainout, confining it deeper in the atmosphere. The thermosphere of an exoplanet is most sensitively probed by the Na line core (e.g. Huitson et al. 2012). However, the line core of Na is not resolved in the transmission spectrum of WASP-39b, which prohibits a more detailed investigation of the line core and the inclusion of the thermosphere in the models detailed. In the case that the Na line core temperature is underestimated, the model will not be able to match the data higher in the atmosphere (Figs. 6, 9). To further constrain the temperature and structure of WASP-39bs thermosphere, high resolution observations of Na will be required.

Following the results from the solar-metallicity GCM of WASP-39b (Fig. 11), we might expect to observe absorption by CH_4 in the atmosphere of WASP-39b, as a number of the P-T profiles cross the $CO = CH_4$ boundary. However, we do not see any evidence for methane absorption in the transmission spectrum, which would be present at the red end of the G141 bandpass ($>1.55\,\mu$m) and in the $3.6\,\mu$m *Spitzer* bandpass. Both the equilibrium chemistry and free-chemistry retrievals suggest that the atmosphere of WASP-39b likely has a metallicity greater than $100\times$solar. The high metallicity will impact not only the P-T profiles by pushing them to higher temperatures but would also push the $CO = CH_4$ transition to lower temperatures, particularly at high pressures (Agúndez et al. 2014). This likely places the atmosphere of WASP-39b firmly above the $CO = CH_4$ boundary, in the CO dominated regime, for all pressures and all locations horizontally. It is therefore not likely that the CH_4 abundance is enhanced relative to equilibrium abundances due to horizontal or vertical quenching, as has been postulated for many hot Jupiter atmospheres (e.g. Cooper & Showman 2006; Agúndez et al. 2014).

The only current evidence of carbon-based species in the transmission spectrum of WASP-39b is in the $4.5\,\mu$m *Spitzer* bandpass, which covers potential absorption by CO and/or CO_2. In Fig. 12 we show the difference in retrieved models from the equilibrium and free-chemistry cases. From this we can see that the equilibrium chemistry model struggles to fit the data, as the fit is likely dominated by the high precision transmission spectrum below $1.7\,\mu$m. In the free-chemistry model, because all the molecules are individually fit, the CO/CO_2 can be balanced such that each molecules is fit by the model.
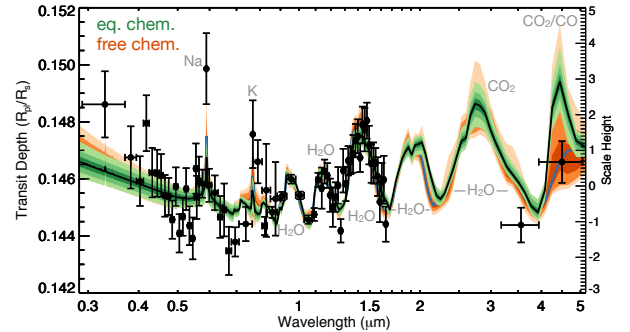

FIG. 12.— The transmission spectrum of WASP-39b (black points). We show the result from both ATMO retrievals using equilibrium chemistry (green) and free chemistry (orange) with the 1, 2, and 3σ bounds as in Fig.6.

However, this results in much larger uncertainties on the data as wider distributions are invoked to account for low resolution data. The CO feature also extends slightly beyond the photometric wavelength range covered by the *Spitzer* point, which also extends the uncertainty of the model. Due to the high metallicity of the atmosphere, the $4.5\,\mu$m feature is likely dominated by CO_2 absorption (Moses et al. 2013). Future observations with JWST will be able to better distinguish between carbon-based species in the atmosphere and it is then likely the free-chemistry model will provide a more informative retrieval. However, with the current data and based on the χ_ν^2 and BIC, the equilibrium model has the greatest statistical significance; as such we use this model and results to further interpret the atmosphere of WASP-39b.

4.3. The importance of the optical data in constraining molecular abundances

We next explored the impact different wavelength regions have on the overall metallicity constraint by performing an identical equilibrium retrieval, but excluding the optical HST and VLT data below $0.8\,\mu$m. Optical transmission spectra have been theoretically shown to be important when measuring the volume mixing ratios of species (Benneke & Seager 2012; Griffith 2014; Line & Parmentier 2016; Heng & Kitzmann 2017), as an infrared-only transmission spectra provides only the relative abundances between molecules. However, few demonstrations are available using optical spectral data. While similar data are available for other exoplanets (Sing et al. 2016), this test is particularly enlightening with WASP-39b, as it has continuous wavelength coverage from 0.3 to 1.7 μm thanks to the WFC3/G102 data. This complete coverage includes pressure-broadened alkali lines (Fischer et al. 2016), multiple well-resolved water features, and overall higher quality data (precision and resolution) than many other exoplanetary spectra.

We use equilibrium chemistry models to compare the retrieved results for the full and partial WASP-39b data, as it provided the best overall fit to the data as measured by the BIC. The results can be seen in Fig. 13. Despite adopting chemical equilibrium and including the Spitzer IRAC data in the retrieval, the infrared-only data was unable to constrain the lower end of the C/O ratio, and a very strong degeneracy was observed between the C/O and [M/H]. Without the optical data, the 3-σ range of the transmission spectra at 0.3μm encompasses 3.4 pres-
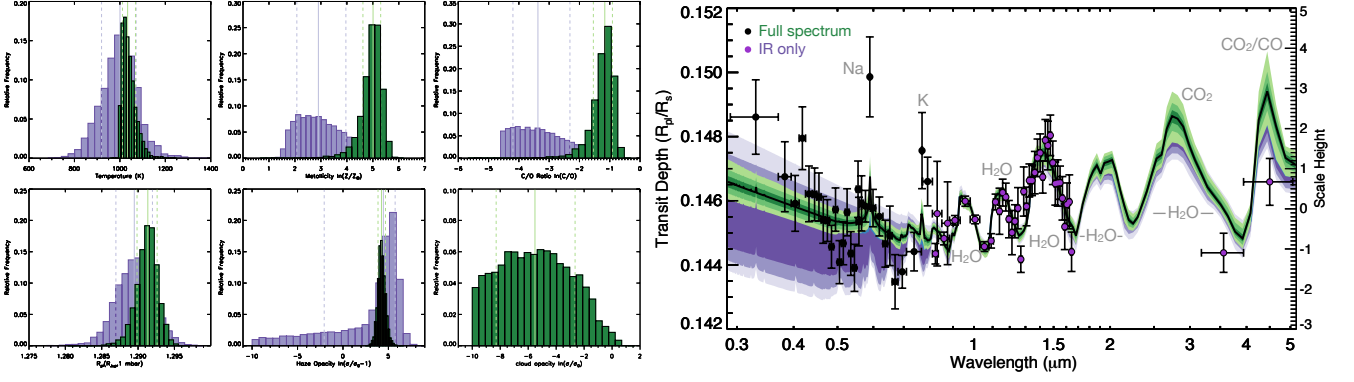
FIG. 13.— Left: histograms showing the relative frequency distributions for each parameter (Temperature, [M/H], C/O, R_p/R_*, Haze, and Cloud) in the equilibrium retrieval for the full dataset (green) and just the IR portion of the dataset (purple). Right: the complete transmission spectrum of WASP-39b (black points), where the IR data is indicated by purple points. The two models show equilibrium models fit to the full transmission spectrum (green), and just the IR points (purple), with the 1, 2, and 3σ bounds. These clearly shows the uncertainty associated with limited wavelength coverage and the importance of the optical.

sure scale heights (Fig. 13 purple range) and contains a wide range of models, including ones with a strong haze, completely cloud- or haze-free models, and models with both near solar and highly super-solar metallicity (see Fig. 13). The posterior of the infrared-only retrieval shows the [M/H] is measured to 0.64 dex, however, a lower prior value of 0.01 was enforced on the C/O ratio which, in turn, also limits the lower range of [M/H]. By including the optical HST and VLT data into the retrieval, pressure information from the alkali lines and constraints on the near-UV scattering slope help to limit the parameter space and [M/H], which in turn helps constrain the C/O ratio as well. Even though the near-UV transmission spectral features can not uniquely be pinned to Rayleigh scattering of the bulk H_2 gas, as a scattering aerosol contribution is also present and included in the retrieval model, the optical data does exclude models with either very high or very low haze components, which helps to constrain the fitted parameters. With the optical data, the resulting [M/H] is measured to 0.14 dex, which is a 0.5 dex improvement over excluding the optical data. The abundance constraints of water (as determined by the [M/H] and C/O ratio see §3.3) are also significantly improved from precisions of $\pm 86\times$ to $\pm 46\times$ solar.

From this test, we demonstrate that the optical spectra with the data quality as provided by HST and VLT can provide an important contribution in constraining the abundances of molecules identified from infrared transmission spectra. In the absence of optical data, breaking the [M/H] - C/O degeneracy will likely require complete near-infrared spectral coverage to identify all the major molecular components (CO, CO_2, CH_4, H_2O) such that the C/O ratio can be directly measured and in turn [M/H] constrained (Greene et al. 2016). For JWST, coverage between 0.6 and 5 μm for targets brighter than J∼10 will require at least two transit observations (e.g. NIRISS/SOSS and NIRSpec/G395H), while fainter targets can cover the range at low resolution in one transit with the NIRSpec prism. Given that significant O and C could be locked up in other species such as condensates (Greene et al. 2016), the inclusion of optical data may also prove useful to identify potential biases when estimating the C/O ratio solely from the major molecular components. Thus, significant leverage can be

gained by combining JWST spectra with that of HST or other ground-based facilities, though care must always be taken when interpreting non-simultaneously gathered data, especially if the planet orbits an active star.

4.4. The impact of limb differences

Overall the retrieved temperature matches well with the 3D GCM model, where increases in the metallicity would shift the P-T profiles hotter (Fig. 11). At higher temperatures it might then be expected that the difference in temperature at the two limbs will also increase (Kataria et al. 2016). Differences in limb temperatures may result in different conditions at the east and west limbs of the annulus such that cloud condensates can form at one and not the other (see Fig.11).

The presence of clouds on one limb of the planet but not the other can potentially mimic high metallicity signatures in the transmission spectrum (Line & Parmentier 2016). Using the simplistic toy model presented in Line & Parmentier (2016), which uses a linear combination of two forward models to approximate limb differences, we test the scenario that the two limbs are different with a series of models on the full transmission spectrum of WASP-39b. We use a range of 1D isothermal models from the Goyal grid to represent different atmospheric scenarios on the planetary limbs. Using Fig. 11 as a guide we select two models separated by 200 K, where the cooler model (741 K) represents the fully cloudy limb (i.e. a uniform optically thick cloud across all wavelengths), and the hotter model (966 K) represents the clear, cloud-free limb. For each model set we keep the C/O, [M/H], and haze parameters the same. We test six sets with varied [M/H] values of 0.0, 1.7, and 2.0 dex (1×, 50×, and 100× solar). For each of the three metallicities we then test two different haze values, 10× and 150× (Table 3). We fit each separate model to the WASP-39b transmission spectrum by allowing it to move in altitude only, with all other parameters considered fixed. In effect, this only introduces one free parameter to the fit, however, it should be noted that the models themselves while fixed are based on a number of variables which when modeled more comprehensively to determine limb-differences may impact the statistics of fit to the data. The ΔBIC values presented in Table 3 represent the difference in BIC between the best fit model and all other models in

TABLE 3
Test cases for the impact of limb differences on the observed transmission spectrum. In each pair the two models contribute 50% each to the combined model. Each model and the combined model are fit to the transmission spectrum and the χ^2, BIC, and ΔBIC (compared to the lowest BIC model) are calculated (see §4.4)

Group	T_{eq} (K)	[M/H]	Haze	Cloud	χ^2	BIC	ΔBIC
a)	741	0.0	10.0	1.0	196	200	50
	966	0.0	10.0	0.0	327	331	181
50/50					179	187	37
b)	741	0.0	150.0	1.0	360	364	214
	966	0.0	150.0	0.0	757	762	612
50/50					514	522	372
c)	741	1.7	10.0	1.0	149	153	3
	966	1.7	10.0	0.0	326	330	180
50/50					166	174	24
d)	741	1.7	150.0	1.0	180	184	34
	966	1.7	150.0	0.0	224	228	78
50/50					174	182	32
e)	741	2.0	10.0	1.0	157	161	11
	966	2.0	10.0	0.0	224	228	78
50/50					142	150	0
f)	741	2.0	150.0	1.0	175	178	28
	966	2.0	150.0	0.0	158	162	12
50/50					146	154	2

the table. From this test we find that the higher metallicity models statistically fit the data better than low metallicity cases, even when 50/50 clear/cloudy models are considered. However, it should be noted that this is an oversimplified model and the parameter space being explored. Future observations of WASP-39b, which will likely have complete wavelength coverage, and much higher precision and resolution data, will likely require sophisticated 3D modeling to accurately infer anything further from the data when considering partly cloudy scenarios.

These simple toy model results, along with the abundance measurements from equilibrium and free chemistry retrievals, further suggest that WASP-39b has a high metallicity atmosphere. We detail the implications of this in the following section.

4.5. *The atmospheric metallicity of WASP-39b*

There are two scenarios generally considered for the formation of giant planets prior to the assumed migration of close-in giant exoplanets: gravitational instability and core accretion. Planetary atmospheres will exhibit different atmospheric properties under formation via these two formation pathways. Gravitational instability theory suggests that planets will have the same atmospheric metallicity as the central star, while with core accretion theory, lower mass planets will have higher atmospheric metallicity (e.g. Mordasini et al. 2012; Fortney et al. 2013). Studies of our solar system giant planets fit with the predicted trend of core accretion when the atmospheric methane abundance is used as a proxy for overall metallicity (c.f Kreidberg et al. 2014). More recently, atmospheric water absorption features in exoplanet atmospheres have been used with retrieval modeling to constrain the overall atmospheric metallicity using oxygen as a proxy for the heavy element abundance (Kreidberg et al. 2014; Fraine et al. 2014; Kreidberg et al.

2015; Wakeford et al. 2017). The first measurement of this type was conducted on WASP-43b (Kreidberg et al. 2014) which directly fit with the trend established by the solar system supporting core accretion theory prior to inward migration. However, in a more recent study of the Neptune-mass exoplanet HAT-P-26b (Wakeford et al. 2017), we showed a deviation from this trend in the low mass regime, hinting at diversity in formation location and/or time. This is consistent with the envelope accretion models by Lee & Chiang (2016), which argue that most hot Neptunes accrete their envelopes in-situ shortly before the disk dissipates resulting in lower heavy element contamination in the atmosphere.

To better approximate the correlation in mass-metallicity space, we separately calculate linear fits to the methane and water abundance measurements for the four solar system giant planets, and the four published exoplanet measurements respectively. To qualitatively assess the significance of fit to the data we use the ρ^2 statistic (McFadded 1974), defining it here as,

$$\rho^2 = 1 - \frac{\sum \frac{(x - line)^2}{\sigma_x^2}}{\sum \frac{(x - mean)^2}{\sigma_x^2}} = 1 - \frac{\chi^2_{line}}{\chi^2_{mean}} \qquad (6)$$

where, x is the data, $line$ is the linear fit to the data, $mean$ is the mean of the data, and σ_x is the uncertainty on the data assuming a Gaussian distribution with symmetric uncertainties in log-metallicity space. The ρ^2 statistic evaluates the improvement that the more complex model has to the fit, compared with a more simplistic model. In this case, it balances the likelihoods of the data being drawn from a model where there is a correlation between mass and metallicity and the data being drawn from a model with no correlation, fixed at the average metallicities of the data. From this statistic we find that ∼93% of the scatter observed in the solar system mass-metallicity relation can be explained by a linear model, even when the uncertainties are taken into account. For the previously published exoplanet data (Fig.14, blue circles) we find that 60% of the variance can be explained by a linear fit to the data.

Using the water abundance as a proxy for overall atmospheric metallicity, we constrain the atmospheric metallicity of WASP-39b to be $151^{+48}_{-46} \times$solar, at the 68% confidence interval, from a retrieval using equilibrium chemistry. We show the metallicty of WASP-39b relative to other giant planets in Fig. 14. We also include the 95% and 99.7% bounds of the retrieved metallicity from the equilibrium and free-chemistry fits to better demonstrate the similarities and bounds of each retrieval. When the new WASP-39b results are incorporated into the exoplanet fit, the ρ^2 statistic drops significantly to just 24% statistical association with a linear fit. This does not rule out a linear fit to the exoplanet data, and indeed the exoplanet linear fits both have log Bayes factors on the order of 5, suggesting a tentative positive relationship (Kass & Raftery 1995). It merely suggests that more high precision data are required to determine a trend in mass-metallicity space. As shown by the ρ^2 statistic, with each new exoplanet metallicity measurement, it should be expected that the linear fit of the mass-metallicity relation will evolve in both variance and constraint. These future observations may also show that systems with multiple
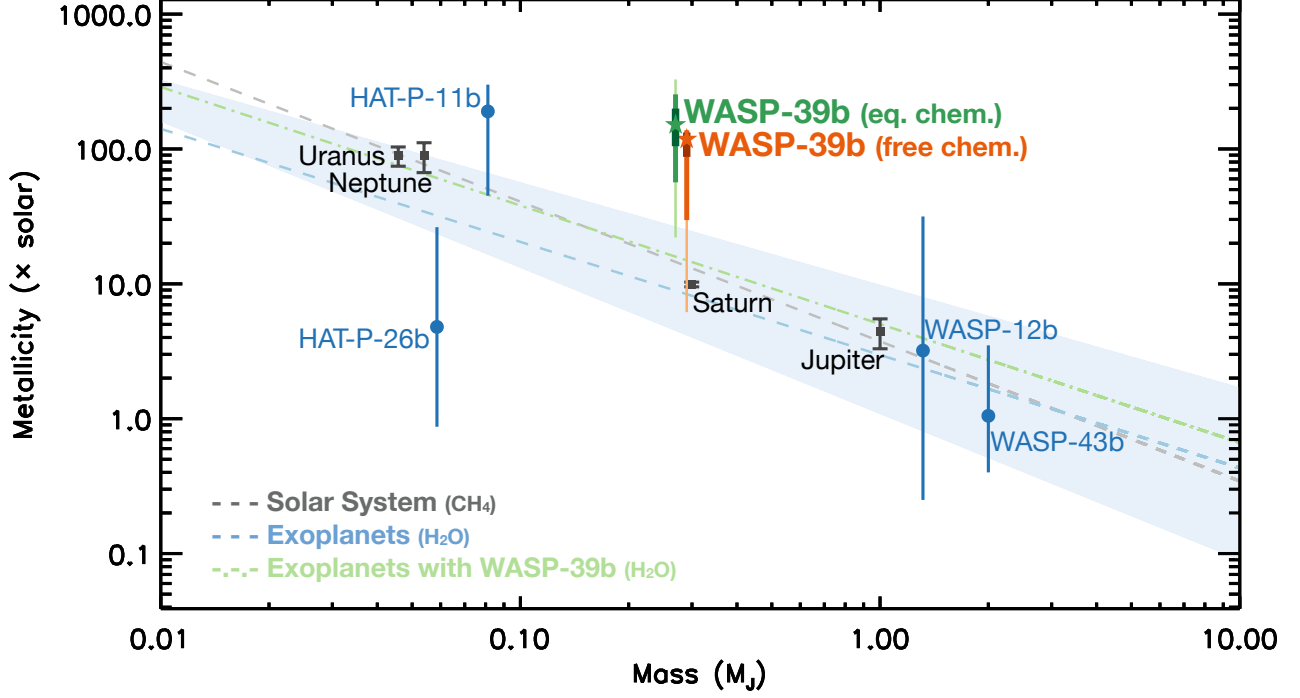
FIG. 14.— Mass-metallicity relation for the solar system and exoplanets. We show the measured metallicities of the four giant planets in our solar system (black squares) fit for the methane abundance (grey dashed line), and four previously published exoplanets (blue circles) fit for the water abundance (blue dashed line); all plotted metallicities show the 68% confidence interval. The shaded region represents the 1σ diversity from all eight measured [M/H] and uncertainties. We show the retrieved metallicity of WASP-39b from the equilibrium chemistry fit (green left) and the free-chemistry fit (orange right) based on the water abundance of the atmosphere, with the 68%, 95%, and , 99.7% confidence intervals (dark, medium, and light errorbars respectively). We also compute the fit to the exoplanet data including the WASP-39b equilibrium chemistry model using the 68% confidence interval measurement (green dash-dot line). WASP-39b metallicity results from each model are offset in mass by $0.01\,\mathrm{M}_J$ for clarity.

TABLE 4
STATISTICAL SIGNIFICANCE OF THE MASS-METALLICITY PLOT
FOLLOWING EQN. 6.

Data	ρ^2
Solar System	0.93
Exoplanets (no WASP-39b)	0.60
Exoplanets (with WASP-39b)	0.24

giant planets provide a more telling comparison to the solar system, as multi-planet systems may have entirely different metallicities than single-planet systems. Simulations presented in Thorngren et al. (2016) show that the scatter in the heavy element fraction relative to the mass of gas giant planets is expected to be large. This suggests that the metallicity measurement for WASP-39b retrieved here is not entirely unexpected, although possibly at the extreme top edge of the scatter. Following the core accretion theory, this suggests that WASP-39b formed in a region rich with heavy element planetesimals, likely in the form of ices, which were accreted by the planet during formation. The potential presence of ices implies formation beyond the ice lines of the planet-forming disk, more akin to Neptune and Uranus's orbital distance than that of the similar mass planet Saturn in our own solar system.

5. CONCLUSION

We present the complete transmission spectrum of the Saturn-mass exoplanet WASP-39b by introducing new

measurements between 0.8–1.7 μm using HST WFC3. We measure distinct water absorption over three bands with a maximum base to peak amplitude of 2.4 planetary scale heights (H) and an average amplitude of 1.7 H. Using the ATMO Retrieval Code (ARC) we constrain the temperature to 1030^{+30}_{-20} K, and the atmospheric metallicity at $151^{+48}_{-46} \times$ solar, based predominantly on the water abundance. At a 1σ confidence this still represents significant diversity from the current mass-metallicity trends based on either atmospheric methane or water abundance for giant planets. This suggests that WASP-39b formed beyond the snow line in the planet-forming disk of the host star, where it likely accumulated metal-rich ices and planetesimals prior to later inward migrations to its current orbital position. However, overall more precise exoplanet abundances are be needed before definitive conclusions can be made with regards to the exoplanet mass-metallicity relation and planetary formation pathways.

WASP-39b is an ideal target for follow-up studies with the James Webb Space Telescope (JWST) to precisely measure the atmospheric carbon species and abundance already hinted at in these early investigations. We predict that due to the high metallicity of WASP-39b's atmosphere, CO_2 will be the dominant carbon species. This can be measured at 4.5μm with JWST NIRSpec G395H, allowing further constraint to be placed on the C/O and atmospheric metallicity.

6. ACKNOWLEDGEMENTS

Facilities: HST (WFC3).

REFERENCES

Adcroft A., Campin J.-M., Hill C., Marshall J., 2004, Monthly Weather Review, 132, 2845

Agúndez M., Parmentier V., Venot O., Hersant F., Selsis F., 2014, A&A, 564, A73

Allard N. F., Spiegelman F., Kielkopf J. F., 2007, A&A, 465, 1085

Amundsen D. S., Baraffe I., Tremblin P., Manners J., Hayek W., Mayne N. J., Acreman D. M., 2014, A&A, 564, A59

Anderson D., Hellier C., Gillon M., Triaud A., Smalley B., Hebb L., Cameron A. C., Maxted P., Queloz D., West R., et al., 2010, ApJ, 709, 159

Anderson D. R., Collier Cameron A., Hellier C., Lendl M., Lister T. A., Maxted P. F. L., Queloz D., Smalley B., Smith A. M. S., Triaud A. H. M. J., West R. G., Brown D. J. A., Gillon M., Pepe F., Pollacco D., Ségransan D., Street R. A., Udry S., 2011, A&A, 531, A60

Barstow J. K., Aigrain S., Irwin P. G. J., Sing D. K., 2017, ApJ, 834, 50

Benneke B., Seager S., 2012, ApJ, 753, 100

Berta Z. K., Charbonneau D., Désert J.-M., Kempton E. M.-R., McCullough P. R., Burke C. J., Fortney J. J., Irwin J., Nutzman P., Homeier D., 2012, ApJ, 747, 35

Bohren C. F., Huffman D. R., 2008, Absorption and scattering of light by small particles. Wiley. com

Burnham K. P., Anderson D. R., 2004, Sociological methods & research, 33, 261

Charbonneau D., Brown T. M., Noyes R. W., Gilliland R. L., 2002, ApJ, 568, 377

Charnay B., Meadows V., Misra A., Leconte J., Arney G., 2015, ApJ, 813, L1

Claret A., 2000, A&A, 363, 1081

Cooper C. S., Showman A. P., 2006, ApJ, 649, 1048

Deming D., Seager S., 2017, ArXiv e-prints

Deming D., Wilkins A., McCullough P., Burrows A., Fortney J. J., Agol E., Dobbs-Dixon I., Madhusudhan N., Crouzet N., Desert J.-M., et al., 2013, ApJ, 774, 95

Drummond B., Tremblin P., Baraffe I., Amundsen D. S., Mayne N. J., Venot O., Goyal J., 2016, A&A, 594, A69

Eastman J., Gaudi B. S., Agol E., 2013, PASP, 125, 83

Evans T. M., Sing D. K., Kataria T., Goyal J., Nikolov N., Wakeford H. R., Deming D., Marley M. S., Amundsen D. S., Ballester e. a., 2017, Nature, 548, 58

Evans T. M., Sing D. K., Wakeford H. R., Nikolov N., Ballester G. E., Drummond B., Kataria T., Gibson N. P., Amundsen D. S., Spake J., 2016, ApJ, 822, L4

Faedi F., Barros S. C. C., Anderson D. R., Brown D. J. A., Collier Cameron A., Pollacco D., Boisse I., Hébrard G., Lendl M., Lister T. A., Smalley B., Street R. A., Triaud e. a., 2011, A&A, 531, A40

Fischer P. D., Knutson H. A., Sing D. K., Henry G. W., Williamson M. W., Fortney J. J., Burrows A. S., Kataria T., Nikolov N., Showman A. P., Ballester G. E., Désert J.-M., Aigrain S., Deming D., Lecavelier des Etangs A., Vidal-Madjar A., 2016, ApJ, 827, 19

Fortney J. J., Mordasini C., Nettelmann N., Kempton E. M.-R., Greene T. P., Zahnle K., 2013, ApJ, 775, 80

Fortney J. J., Shabram M., Showman A. P., Lian Y., Freedman R. S., Marley M. S., Lewis N. K., 2010, ApJ, 709, 1396

Fraine J., Deming D., Benneke B., Knutson H., Jordán A., Espinoza N., Madhusudhan N., Wilkins A., Todorov K., 2014, Nature, 513, 526

Gibson N. P., 2014, MNRAS, 445, 3401

Gibson N. P., Aigrain S., Roberts S., Evans T. M., Osborne M., Pont F., 2012, MNRAS, 419, 2683

Goyal J. M., Mayne N., Sing D. K., Drummond B., Tremblin P., Amundsen D. S., Evans T., Carter A. L., Spake J., Baraffe I., Nikolov N., Manners J., Chabrier G., Hebrard E., 2017, ArXiv e-prints

Greene T. P., Line M. R., Montero C., Fortney J. J., Lustig-Yaeger J., Luther K., 2016, ApJ, 817, 17

Griffith C. A., 2014, Philosophical Transactions of the Royal Society of London Series A, 372, 20130086

Heng K., Kitzmann D., 2017, MNRAS, 470, 2972

Huitson C. M., Sing D. K., Vidal-Madjar A., Ballester G. E., Lecavelier des Etangs A., Désert J.-M., Pont F., 2012, MNRAS, 422, 2477

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 773

Kataria T., Showman A. P., Fortney J. J., Marley M. S., Freedman R. S., 2014, ApJ, 785, 92

Kataria T., Showman A. P., Fortney J. J., Stevenson K. B., Line M. R., Kreidberg L., Bean J. L., Désert J.-M., 2015, ApJ, 801, 86

Kataria T., Sing D. K., Lewis N. K., Visscher C., Showman A. P., Fortney J. J., Marley M. S., 2016, ApJ, 821, 9

Kreidberg L., Bean J. L., Désert J.-M., Line M. R., Fortney J. J., Madhusudhan N., Stevenson K. B., Showman A. P., Charbonneau D., McCullough P. R., Seager S., Burrows A., Henry G. W., Williamson M., Kataria T., Homeier D., 2014, ApJ, 793, L27

Kreidberg L., Line M. R., Bean J. L., Stevenson K. B., Désert J.-M., Madhusudhan N., Fortney J. J., Barstow J. K., Henry G. W., Williamson M. H., Showman A. P., 2015, ApJ, 814, 66

Lee E. J., Chiang E., 2016, ApJ, 817, 90

Lewis N., Parmentier V., Kataria T., de Wit J., Showman A., Fortney J., Marley M., 2017, arXiv preprint arXiv:1706.00466

Lewis N. K., Showman A. P., Fortney J. J., Marley M. S., Freedman R. S., Lodders K., 2010, ApJ, 720, 344

Line M. R., Parmentier V., 2016, ApJ, 820, 78

Line M. R., Stevenson K. B., Bean J., Desert J.-M., Fortney J. J., Kreidberg L., Madhusudhan N., Showman A. P., Diamond-Lowe H., 2016, AJ, 152, 203

McFadden D., 1974, Conditional logit analysis of qualitative choice behavior. Academic Press

Markwardt C. B., 2009, in Bohlender D. A., Durand D., Dowler P., eds, Astronomical Data Analysis Software and Systems XVIII Vol. 411 of Astronomical Society of the Pacific Conference Series, Non-linear Least-squares Fitting in IDL with MPFIT. p. 251

Marley M. S., Ackerman A. S., Cuzzi J. N., Kitzmann D., 2013, Clouds and Hazes in Exoplanet Atmospheres. Univ. of Arizona, pp 367–391

Marley M. S., McKay C. P., 1999, Icarus, 138, 268

Mordasini C., Alibert Y., Benz W., Klahr H., Henning T., 2012, A&A, 541, A97

Moses J. I., Madhusudhan N., Visscher C., Freedman R. S., 2013, ApJ, 763, 25

Nikolov N., Sing D. K., Burrows A. S., Fortney J. J., Henry
    G. W., Pont F., Ballester G. E., Aigrain S., Wilson P. A.,
    Huitson C. M., Gibson N. P., Désert J.-M., Etangs A. L. d.,
    Showman A. P., Vidal-Madjar A., Wakeford H. R., Zahnle K.,
    2015, MNRAS, 447, 463
Nikolov N., Sing D. K., Gibson N. P., Fortney J. J., Evans T. M.,
    Barstow J. K., Kataria T., Wilson P. A., 2016, ArXiv e-prints
Seager S., Deming D., 2010, ARA&A, 48, 631
Showman A. P., Fortney J. J., Lian Y., Marley M. S., Freedman
    R. S., Knutson H. A., Charbonneau D., 2009, ApJ, 699, 564
Sing D. K., 2010, A&A, 510, A21
Sing D. K., Désert J.-M., Fortney J. J., Lecavelier Des Etangs A.,
    Ballester G. E., Cepa J., Ehrenreich D., López-Morales M.,
    Pont F., Shabram M., Vidal-Madjar A., 2011, A&A, 527, A73
Sing D. K., Fortney J. J., Nikolov N., Wakeford H. R., Kataria
    T., Evans T. M., Aigrain S., Ballester G. E., Burrows A. S.,
    Deming D., Désert J.-M., Gibson N. P., Henry G. W., Huitson
    C. M., Knutson H. A., Etangs A. L. d., Pont F., Showman
    A. P., Vidal-Madjar A., Williamson M. H., Wilson P. A., 2016,
    Nature, 529, 59
Sing D. K., Wakeford H. R., Showman A. P., Nikolov N., Fortney
    J. J., Burrows A. S., Ballester G. E., Deming D., et al. 2015,
    MNRAS, 446, 2428
Stevenson K. B., Désert J.-M., Line M. R., Bean J. L., Fortney
    J. J., Showman A. P., Kataria T., Kreidberg L., McCullough
    P. R., Henry G. W., Charbonneau D., Burrows A., Seager S.,
    Madhusudhan N., Williamson M. H., Homeier D., 2014,
    Science, 346, 838
Thorngren D. P., Fortney J. J., Murray-Clay R. A., Lopez E. D.,
    2016, ApJ, 831, 64
Tremblin P., Amundsen D. S., Chabrier G., Baraffe I., Drummond
    B., Hinkley S., Mourier P., Venot O., 2016, ApJ, 817, L19
Tremblin P., Amundsen D. S., Mourier P., Baraffe I., Chabrier G.,
    Drummond B., Homeier D., Venot O., 2015, ApJ, 804, L17
Wakeford H. R., Sing D. K., Evans T., Deming D., Mandell A.,
    2016, The Astrophysical Journal, 819, 10
Wakeford H. R., Sing D. K., Kataria T., Deming D., Nikolov N.,
    Lopez E. D., Tremblin P., Amundsen D. S., Lewis N. K.,
    Mandell A. M., Fortney J. J., Knutson H., Benneke B., Evans
    T. M., 2017, ArXiv e-prints
Wakeford H. R., Stevenson K. B., Lewis N. K., Sing D. K.,
    López-Morales M., Marley M., Kataria T., Mandell A.,
    Ballester G. E., Barstow J., Ben-Jaffel L., Bourrier V.,
    Buchhave L. A., Ehrenreich D. e. a., 2017, ApJ, 835, L12
Wakeford H. R., Visscher C., Lewis N. K., Kataria T., Marley
    M. S., Fortney J. J., Mandell A. M., 2017, MNRAS, 464, 4247