A Spitzer Transmission Spectrum for the Exoplanet Gj 436b, Evidence for Stellar Variability, and Constraints on Dayside Flux Variations

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:41397469

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

A *SPITZER* TRANSMISSION SPECTRUM FOR THE EXOPLANET GJ 436b, EVIDENCE FOR STELLAR VARIABILITY, AND CONSTRAINTS ON DAYSIDE FLUX VARIATIONS

HEATHER A. KNUTSON[1,13], NIKKU MADHUSUDHAN[2], NICOLAS B. COWAN[3], JESSIE L. CHRISTIANSEN[4], ERIC AGOL[5],
DRAKE DEMING[6], JEAN-MICHEL DÉSERT[7], DAVID CHARBONNEAU[7], GREGORY W. HENRY[8], DEREK HOMEIER[9],
JONATHAN LANGTON[10,11], GREGORY LAUGHLIN[9], AND SARA SEAGER[12]
[1] Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA; hknutson@berkeley.edu
[2] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[3] CIERA, Northwestern University, 2131 Tech Drive, Evanston, IL 60208, USA
[4] NASA Ames Research Center, Moffett Field, CA 94035, USA
[5] Department of Astronomy, Box 351580, University of Washington, Seattle, WA 98195, USA
[6] NASA Goddard Space Flight Center, Planetary Systems Laboratory, Code 693, Greenbelt, MD 20771, USA
[7] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[8] Center of Excellence in Information Systems, Tennessee State University, 3500 John A. Merritt Blvd., Box 9501, Nashville, TN 37209, USA
[9] Institut für Astrophysik, Georg-August-Universität, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany
[10] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA
[11] Department of Physics, Principia College, 1 Maybeck Place, Elsah, IL 62028, USA
[12] Department of Earth, Atmospheric, and Planetary Sciences, Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ABSTRACT

In this paper, we describe a uniform analysis of eight transits and eleven secondary eclipses of the extrasolar planet GJ 436b obtained in the 3.6, 4.5, and 8.0 μm bands using the IRAC instrument on the *Spitzer Space Telescope* between UT 2007 June 29 and UT 2009 February 4. We find that the best-fit transit depths for visits in the same bandpass can vary by as much as 8% of the total (4.7σ significance) from one epoch to the next. Although we cannot entirely rule out residual detector effects or a time-varying, high-altitude cloud layer in the planet's atmosphere as the cause of these variations, we consider the occultation of active regions on the star in a subset of the transit observations to be the most likely explanation. We find that for the deepest 3.6 μm transit the in-transit data have a higher standard deviation than the out-of-transit data, as would be expected if the planet occulted a star spot. We also compare all published transit observations for this object and find that transits observed in the infrared typically have smaller timing offsets than those observed in visible light. In this case, the three deepest *Spitzer* transits are all measured within a period of five days, consistent with a single epoch of increased stellar activity. We reconcile the presence of magnetically active regions with the lack of significant visible or infrared flux variations from the star by proposing that the star's spin axis is tilted with respect to our line of sight and that the planet's orbit is therefore likely to be misaligned. In contrast to the results reported by Beaulieu et al., we find no convincing evidence for methane absorption in the planet's transmission spectrum. If we exclude the transits that we believe to be most affected by stellar activity, we find that we prefer models with enhanced CO and reduced methane, consistent with GJ 436b's dayside composition from Stevenson et al. It is also possible that all transits are significantly affected by this activity, in which case it may not be feasible to characterize the planet's transmission spectrum using broadband photometry obtained over multiple epochs. These observations serve to illustrate the challenges associated with transmission spectroscopy of planets orbiting late-type stars; we expect that other systems, such as GJ 1214, may display comparably variable transit depths. We compare the limb-darkening coefficients predicted by PHOENIX and ATLAS stellar atmosphere models and discuss the effect that these coefficients have on the measured planet–star radius ratios given GJ 436b's near-grazing transit geometry. Our measured 8 μm secondary eclipse depths are consistent with a constant value, and we place a 1σ upper limit of 17% on changes in the planet's dayside flux in this band. These results are consistent with predictions from general circulation models for this planet, which find that the planet's dayside flux varies by a few percent or less in the 8 μm band. Averaging over the eleven visits gives us an improved estimate of $0.0452\% \pm 0.0027\%$ for the secondary eclipse depth; we also examine residuals from the eclipse ingress and egress and place an upper limit on deviations caused by a non-uniform surface brightness for GJ 436b. We combine timing information from our observations with previously published data to produce a refined orbital ephemeris and determine that the best-fit transit and eclipse times are consistent with a constant orbital period. We find that the secondary eclipse occurs at a phase of 0.58672 ± 0.00017, corresponding to $e\cos(\omega) = 0.13754 \pm 0.00027$, where e is the planet's orbital eccentricity and ω is the longitude of pericenter. We also present improved estimates for other system parameters, including the orbital inclination, $a/R_\star$, and the planet–star radius ratio.

Key words: binaries: eclipsing – planetary systems – stars: activity – techniques: photometric

Online-only material: color figures

1. INTRODUCTION

Transiting planet systems have proven to be a powerful tool for studying exoplanetary atmospheres. Observations of

---

[13] Miller Fellow.

transiting systems have been used to detect the signatures of atomic and molecular absorption features at wavelengths ranging from the ultraviolet to the infrared (e.g., Charbonneau et al. 2002; Vidal-Madjar et al. 2003; Swain et al. 2008; Désert et al. 2008; Pont et al. 2008a; Linsky et al. 2010), although sometimes the results have proven to be controversial (e.g., Gibson et al. 2011). They have enabled studies of the dayside emission spectra and pressure–temperature profiles of close-in planets (e.g., Charbonneau et al. 2005; Deming et al. 2005; Knutson et al. 2008; Grillmair et al. 2008), and they have informed us about their atmospheric circulation (e.g., Knutson et al. 2007, 2009a; Cowan et al. 2007; Crossfield et al. 2010). Although we currently know of 101 transiting planet systems, our knowledge of these planets (including a majority of the studies cited above) has so far been dominated by studies of the brightest and closest handful of systems, including HD 209458b and HD 189733b. Planets orbiting small stars offer additional advantages, as they produce proportionally deeper transits and secondary eclipses as a result of their favorable radius ratios and lower stellar effective temperatures. By this standard, GJ 436 (Butler et al. 2004; Maness et al. 2007; Gillon et al. 2007a, 2007b; Deming et al. 2007; Demory et al. 2007; Torres 2007) represents an ideal target, as the primary in this system is an early M star with a K-band magnitude of 6.1.

GJ 436b is currently one of the smallest known transiting planets, with a mass only 22 times that of the Earth (Torres 2007). Of the planets orbiting stars brighter than ninth magnitude in the K band, only GJ 1214b (Charbonneau et al. 2009), which also orbits a nearby M dwarf, is smaller. New discoveries of low-mass transiting planets from space-based surveys such as the *Kepler* and *CoRoT* missions are unlikely to change this picture significantly, as both include relatively few bright stars. GJ 436b is also one of the coolest known transiting planets, with a dayside effective temperature of only 800 K (Stevenson et al. 2010). Like GJ 1214b, GJ 436b has a high average density indicative of a massive rocky or icy core. In GJ 436b's case, models indicate that it must also maintain 1–3 $M_\oplus$ of its mass in the form of an H/He atmosphere (Adams et al. 2008; Figueira et al. 2009; Rogers & Seager 2010; Nettelmann et al. 2010) in order to match the observed radius.

By measuring the wavelength-dependent transit depth as GJ 436b passes in front of its host star, we can study its atmospheric composition at the day–night terminator, which should be dominated by methane, water, and carbon monoxide (Spiegel et al. 2010; Stevenson et al. 2010; Shabram et al. 2011; Madhusudhan & Seager 2011). Pont et al. (2008b) observed two transits of GJ 436b with NICMOS grism spectrograph on the *Hubble Space Telescope* and placed an upper limit on the amplitude of the predicted water absorption feature between 1 and 2 μm. More recently, Beaulieu et al. (2011) reported the detection of strong methane absorption in the 3.6, 4.5, and 8.0 μm *Spitzer* bands.

We can compare these results to observations of the planet's dayside emission spectrum, obtained by measuring the depth of the secondary eclipse when the planet passes behind the star. Stevenson et al. (2010) measured secondary eclipse depths for GJ 436b in the 3.6, 4.5, 5.8, 8.0, 16, and 24 μm *Spitzer* bands, from which they concluded that the planet's dayside atmosphere contained significantly less methane and more CO than the equilibrium chemistry predictions. In this work, we present an analysis of eight transits and eleven secondary eclipses of GJ 436b observed with *Spitzer*, including an independent analysis of the transit data described in Beaulieu et al., and

discuss the corresponding implications for GJ 436b's atmospheric composition.

Unlike most close-in planets, which typically have circular orbits, GJ 436b has an orbital eccentricity of approximately 0.15 (Maness et al. 2007; Deming et al. 2007; Demory et al. 2007; Madhusudhan & Winn 2009). Atmospheric circulation models for eccentric Jovian planets suggest that they may exhibit significant temperature variations from one orbit to the next (Langton & Laughlin 2008; Iro & Deming 2010), although Lewis et al. (2010) find little evidence for significant temporal variability in general circulation models for GJ 436b. The extensive nature of our data set, which includes 11 secondary eclipse observations in the same bandpass obtained between 2007 and 2009, allows us to test the predictions of these models by searching for changes in the planet's 8 μm dayside emission on timescales ranging from weeks to years.

It has also been suggested (Ribas et al. 2008) that GJ 436b's orbital parameters are changing in time, perhaps through perturbations by an unseen second planet in the system. Such a planet could serve to maintain GJ 436b's non-zero eccentricity despite ongoing orbital circularization and would not necessarily produce transit timing variations large enough to be detected by earlier, ground-based studies (Batygin et al. 2009). Although more recent studies (Alonso et al. 2008; Bean & Seifahrt 2008; Coughlin et al. 2008; Madhusudhan & Winn 2009; Cáceres et al. 2009; Shporer et al. 2009; Ballard et al. 2010a) have failed to find any evidence for either time-varying orbital parameters or a second transiting object in the system, *Spitzer*'s unparalleled sensitivity and stability allow us to extend the current baseline by nine months with new, high-precision transit observations.

2. OBSERVATIONS

We analyze 19 separate observations of GJ 436, including two 3.6 μm transits, two 4.5 μm transits, four 8 μm transits, and eleven 8 μm secondary eclipses, as listed in Table 1. All observations were obtained between 2007 and 2009 using the IRAC instrument (Fazio et al. 2004) on the *Spitzer Space Telescope* (Werner et al. 2004) in subarray mode. Some of these data were previously published by other groups, including a transit and secondary eclipse observed on UT 2007 June 29/30 (Deming et al. 2007; Gillon et al. 2007a; Demory et al. 2007) and six transits observed in 2009 (Beaulieu et al. 2011). Because the two shorter wavelength IRAC channels (3.6 and 4.5 μm) use InSb detectors and the two longer wavelength channels (5.8 and 8.0 μm) use Si: as detectors, each of which displays different detector effects, we describe our analysis for each type separately below.

We calculate the BJD_UTC values at mid-exposure for each image using the DATE_OBS keyword in the image headers and the position of *Spitzer*, which is in an Earth-trailing orbit, as determined using the JPL Horizons ephemeris. Each set of 64 images obtained in subarray mode comes as a single FITS file with a time stamp corresponding to the start of the first image; we calculate the time stamps for individual images assuming uniform spacing and using the difference between the AINTBEG and ATIMEEND headers, which record the start and end of the 64-image series. We then use the routines described in Eastman et al. (2010) to convert from *Spitzer* JD to BJD_UTC. Eastman et al. further advocate a conversion from UTC to TT timing standards, which provide a more consistent treatment of leap seconds. We note that for the dates spanned by these observations the conversion from BJD_UTC to BJD_TT simply

Table 1
Spitzer Observations of GJ 436b

UT Date	Event	λ (μm)	Duration (hr)	t_{int} (s)	$N_{exposures}$	Bkd (MJy sr^{-1})[a]	Flux (MJy sr^{-1})[a,b]	σ_{resid}^{c}
UT 2007 Jun 29	Transit	8.0	3.4	0.4	28,480	530.8	9,149.0	0.500%
UT 2007 Jun 30	Eclipse	8.0	5.9	0.4	49,920	544.2	9,148.2	0.498%
UT 2008 Jun 11	Eclipse	8.0	3.4	0.4	28,800	226.0	9,156.6	0.497%
UT 2008 Jun 13	Eclipse	8.0	3.4	0.4	28,800	255.0	9,161.3	0.502%
UT 2008 Jun 16	Eclipse	8.0	3.4	0.4	28,800	296.5	9,154.5	0.494%
UT 2008 Jun 19	Eclipse	8.0	3.4	0.4	28,800	306.4	9,151.7	0.493%
UT 2008 Jul 12[d]	Eclipse	8.0	70	0.4	588,480	669.2	9,159.9	0.506%
UT 2008 Jul 14[d]	Transit	8.0	70	0.4	588,480	695.6	9,160.4	0.509%
UT 2008 Jul 15[d]	Eclipse	8.0	70	0.4	588,480	714.4	9,158.1	0.515%
UT 2009 Jan 9	Transit	3.6	4.3	0.1	117,056	37.7	36,164.3	0.387%
UT 2009 Jan 17	Transit	4.5	4.3	0.1	117,056	61.6	24,382.6	0.561%
UT 2009 Jan 25	Transit	8.0	4.3	0.4	35,904	474.5	9,151.9	0.502%
UT 2009 Jan 27	Eclipse	8.0	3.4	0.4	28,800	455.1	9,164.5	0.495%
UT 2009 Jan 28	Transit	3.6	4.3	0.1	117,056	82.5	36,744.5	0.389%
UT 2009 Jan 29	Eclipse	8.0	3.4	0.4	28,800	411.8	9,161.7	0.496%
UT 2009 Jan 30	Transit	4.5	4.3	0.1	117,056	86.3	24,177.0	0.567%
UT 2009 Feb 1	Eclipse	8.0	3.4	0.4	28,800	395.9	9,163.9	0.499%
UT 2009 Feb 2	Transit	8.0	4.3	0.4	35,904	393.6	9,143.8	0.501%
UT 2009 Feb 4	Eclipse	8.0	3.4	0.4	28,800	376.6	9,154.5	0.499%

Notes.
[a] Average sky backgrounds and stellar fluxes estimated for a 5 pixel aperture.
[b] In order to minimize the effects of the detector ramp in the 8.0 μm observations, we estimate the out-of-transit flux using data after the end of the eclipse event where the ramp is generally smallest; for consistency we use the same region to estimate the fluxes at 3.6 and 4.5 μm. We use a 5.0 pixel aperture for the photometry at [3.6, 4.5, 8.0] μm and apply the appropriate aperture correction of [1.049, 1.050, 1.068] from Table 4.7 of the IRAC Instrument Handbook to determine the total flux from the star in each observation.
[c] Standard deviation of residuals after dividing out best-fit corrections for instrument effects and transit light curves.
[d] These events were observed as part of a single, continuous phase curve observation with a duration of 70 hr spanning two secondary eclipses and one transit.

Table 2
Individual Best-fit Transit Parameters

UT Date	λ (μm)	$R_p/R_\star$	Depth	Transit Center (BJD)	$O - C$ (s)[a]
UT 2007 Jun 29	8.0	0.08322 ± 0.00052	$0.6926\% \pm 0.0087\%$	$2454280.78193 \pm 0.00012$	12.5 ± 10.2
UT 2008 Jul 14	8.0	0.08247 ± 0.00061	$0.6801\% \pm 0.0101\%$	$2454661.50314 \pm 0.00017$	6.2 ± 14.4
UT 2009 Jan 9	3.6	0.08182 ± 0.00037	$0.6694\% \pm 0.0061\%$	$2454841.28821 \pm 0.00008$	6.9 ± 6.5
UT 2009 Jan 17	4.5	0.08286 ± 0.00047	$0.6865\% \pm 0.0078\%$	$2454849.21985 \pm 0.00012$	2.1 ± 10.5
UT 2009 Jan 25	8.0	0.08224 ± 0.00051	$0.6763\% \pm 0.0084\%$	$2454857.15155 \pm 0.00012$	3.2 ± 10.1
UT 2009 Jan 28	3.6	0.08495 ± 0.00056	$0.7216\% \pm 0.0095\%$	$2454859.79504 \pm 0.00012$	-31.9 ± 10.3
UT 2009 Jan 30	4.5	0.08502 ± 0.00057	$0.7227\% \pm 0.0097\%$	$2454862.43970 \pm 0.00011$	33.7 ± 9.6
UT 2009 Feb 2	8.0	0.08424 ± 0.00049	$0.7096\% \pm 0.0083\%$	$2454865.08345 \pm 0.00012$	20.8 ± 10.6

Note. [a] Observed minus calculated transit times. Predictions use the best-fit ephemeris of $T_c = 2454865.083208 \pm 0.000042$ BJD and $P = 2.6438979 \pm 0.0000003$ days from Table 5.

requires the addition of [65.184, 65.184, 66.184] s for data obtained in [2007, 2008, 2009], and we leave the dates listed in Table 2 in BJD_UTC for consistency with other studies.

2.1. 3.6 and 4.5 μm Photometry

GJ 436 has a *K*-band magnitude of 6.07, and as a result we elect to use short 0.1 s exposures at 3.6 and 4.5 μm in order to ensure that we remain well below saturation. Subarray images have dimensions of 32×32 pixels, making it challenging to estimate the sky background independent of contamination from the wings of the star's point-spread function. We choose to exclude pixels within a radius of 12 pixels of the star's position, as well as the 14th–17th rows, which contain a horizontal diffraction spike that extends close to the edges of the array. We also exclude the top (32nd) row of pixels, which have values that are consistently lower than those for the rest of the array. We then iteratively trim 3σ outliers from the remaining subset of approximately 600 pixels, create a histogram of the remaining

values, and then fit a Gaussian to this histogram to determine the sky background for each image. We find that the background is 0.1%–0.2% and 0.3%–0.4% of the total flux in a 5 pixel aperture for the 3.6 and 4.5 μm arrays, respectively.

We correct for transient hot pixels by taking a 10 pixel running median of the fluxes at a given pixel position within each set of 64 images and replacing outliers greater than 4σ with the median value. We found that using a wider median filter or tighter upper limit for discriminating outliers increased the scatter in the final time series while failing to significantly reduce the number of images that are ultimately discarded. We find that approximately 0.4%–0.8% of our images have one or more pixels flagged as outliers using this filter.

Several recent papers have investigated optimal methods for estimating the position of the star on the array for *Spitzer* photometry, with the most extensive discussions appearing in Stevenson et al. (2010) and Agol et al. (2010). These papers conclude that flux-weighted centroiding (e.g., Knutson et al.

2008; Charbonneau et al. 2008) and parabola-fitting routines (e.g., Deming et al. 2006, 2007) tend to produce less than optimal results, while Gaussian fits and least asymmetry methods appear to have fewer systematic biases and a lower overall scatter. We confirm that for all three wavelengths we obtain better results (defined as a lower scatter in the final trimmed light curve after correcting for detector effects) with Gaussian fits than with flux-weighted centroiding, with a total reduction of 2%–7% in the standard deviation of the final time series binned in sets of 64 images. We obtain the best results in both the 3.6 and 4.5 μm bands when we first subtract the best-fit background flux from each image, correct bad pixels as described above, and then fit a two-dimensional Gaussian function to a circular region with a radius of 4 pixels centered on the position of the star. Using smaller or larger fitting regions does not significantly alter the time series but does result in a slightly higher scatter in the normalized light curve. Although error arrays are available as part of the standard *Spitzer* pipeline, we find that in this case we obtain better results using uniform error weighting for individual pixels. We use a radially symmetric Gaussian function and run our position estimation routines twice, once where the width is allowed to vary freely in the fits and a second time where we fix the width to the median value over the time series. Reducing the degrees of freedom by fixing the width produces fits that converge more consistently, with a corresponding improvement in the standard deviation of the normalized time series and fewer large outliers. Stevenson et al. (2010) report that they obtain better position estimates when fitting Gaussians to images that have been interpolated to 5× higher resolution, but we find that using interpolated images for our position fits resulted in a slight increase in the scatter in our final light curves.

We perform aperture photometry on our images using the position estimates derived from our Gaussian fits; we expect that aperture photometry will produce the optimal results in light of the low background flux at these wavelengths. We use apertures with radii ranging between 2.5 and 7.0 pixels in half-pixel steps. We find that apertures smaller than 3.5 pixels show excess noise, likely connected to position-dependent flux losses, while apertures larger than 5 pixels are more likely to include transient hot pixels and higher background levels, resulting in a higher root-mean-square variance in the final light curve. We use a 3.5 pixel aperture for our final analysis, but we find consistent results for apertures between 3.5 and 5.0 pixels. We trim outliers from our final time series using a 50 point running median, where we discard outliers greater than 3σ, approximately 2% of the points in a typical light curve. We find that we trim fewer points when we use flux-weighted centroiding for our position estimates (typically 0.6%), but the uncertainties in our best-fit transit parameters are still larger than with the Gaussian fits due to the increased scatter in the final trimmed time series. We also trim the first 15 minutes in all observations except for the 4.5 μm transit on UT 2009 January 30, where we trim the first hour of data. Images taken at the start of a new observation tend to have larger pointing offsets, most likely due to the settling time of the telescope at the new pointing; we find that discarding these early data improves the quality of the fit to the subsequent points. For all visits other than the transit on UT 2009 January 30, we find that we achieve consistent results when we trim either the first 15, 30, or 60 minutes of data, and we therefore choose to trim the minimal 15 minute interval. For the UT 2009 January 30 observation, we find that the data display an additional time dependence that is not well described by the standard linear function of time in Equation (1), but is instead better described

by a linear function of $\ln(dt)$. This may be due to the fact that the star falls near the edge of a pixel in these observations, which could introduce additional time-dependent effects. Rather than changing the functional form used to fit these data, we instead opt to trim the first hour of observations, which removes the most steeply varying part of the time series and leaves a trend that is well described by the same linear function of time used in the other transit fits. We find that we obtain the same transit parameters for this visit when we either trim the first 15 minutes of data and fit with a linear function of $\ln(dt)$ or trim the first hour of data and fit with a linear function of time, so this choice does not affect our final conclusions.

Fluxes measured at these two wavelengths show a strong correlation with the changing position of the star on the array, at a level comparable to the depth of the secondary eclipse. This effect is due to a well-documented intrapixel sensitivity variation (e.g., Reach et al. 2005; Charbonneau et al. 2005, 2008; Morales-Calderon et al. 2006; Knutson et al. 2008), in which the sensitivity of an individual pixel differs by several percent between the center and the edge. The 3.6 μm array typically exhibits larger sensitivity variations than the 4.5 μm array, as demonstrated by the UT 2009 January 9 and 17 transits. The UT 2009 January 30 transit falls very near the edge of a pixel in the 4.5 μm subarray and thus displays a sensitivity variation comparable to that of the more centrally located 3.6 μm transit on UT 2009 January 28. We correct for these sensitivity variations by fitting a quadratic function of x and y position simultaneous with the transit light curve:

$$f = f_0 \times (a_1 + a_2(x - x_0) + a_3(x - x_0)^2 \\ + a_4(y - y_0) + a_5(y - y_0)^2 + a_6 t), \quad (1)$$

where f_0 is the original flux from the star, f is the measured flux, x and y denote the location of the star on the array, x_0 and y_0 are the median x and y positions, t is the time from the predicted eclipse center, and a_1–a_6 are free parameters in the fit. In both bandpasses, we find that quadratic terms in both x and y are necessary to achieve a satisfactory fit to the observed variations, although the χ^2 value for the fits is not improved by the addition of an xy term, or higher-order terms in x and y. We find that the fits are also improved by the addition of a linear term in time, consistent with previous observations at these wavelengths (e.g., Knutson et al. 2009b; Todorov et al. 2010; Fressin et al. 2010; O'Donovan et al. 2010; Deming et al. 2011).

2.2. 8.0 μm Photometry

We follow the same methods described in Section 2.1 to estimate the sky background in the 8.0 μm images, except in this case we include pixels at distances of more than 10 pixels from the position of the star in our estimate instead of the previous 12 pixel radius. The background in these images ranges between 2.6% and 7.7% of the total flux in a 5 pixel aperture, and we find that including pixels between 10 and 12 pixels from the star's position improves the accuracy of our background estimates without adding significant contamination from the star's point-spread function. In Agol et al. (2010), we find that using a slightly larger 4.5 pixel aperture instead of 3.5 pixels minimizes correlated noise in 8 μm *Spitzer* observations (albeit at the cost of slightly higher Gaussian noise), and we therefore elect to use a 4.5 pixel aperture for our 8 μm data. Our choice of aperture has a negligible effect on the best-fit eclipse depths and times, as we find consistent results for apertures between 3.5 and 5.0 pixels.

Spitzer fluxes for stars observed using the IRAC $8.0\,\mu$m array, the IRS $16\,\mu$m array, and the MIPS $24\,\mu$m array do not appear to have a significant position dependence, but do display a ramp-like behavior where higher-illumination pixels converge to a constant value within the first hour of observations while lower-illumination pixels show a continually increasing linear trend on the timescales of interest here. This effect is believed to be due to charge trapping in the array and is discussed in detail in Knutson et al. (2007, 2009c) and Agol et al. (2010), among others. We mitigate this effect in our data by staring either at a bright star (HD 107158 in the case of the $8\,\mu$m secondary eclipse observations between UT 2008 June 11 and June 19) or an H II region with bright diffuse emission at $8\,\mu$m (LBN 543 for the $8\,\mu$m transit observations and G111.612+0.374 for the $8\,\mu$m secondary eclipse observations between UT 2009 January 27 and February 4) for approximately 30 minutes prior to the start of our observations. The 2007 observations were obtained prior to the development of this preflash technique, but as discussed in Deming et al. (2007) the transit observation happened to follow an observation of another bright object and thus was effectively preflashed in the same manner as the 2008 and 2009 data. The secondary eclipse observed in 2007 was not preflashed and thus displays a much steeper ramp than the other observations. We examine the distribution of ramp slopes in our $8\,\mu$m data and find no correlation between the relative offsets in the positions of GJ 436 and the preflash star and the slope of the subsequent ramp; the preflash star is offset by 0.4, 0.2, 0.3, and 0.1 pixels in the UT 2008 June 11, 13, 16, and 19 observations, respectively, but the shallowest ramp occurs in the June 13 observation while the smallest offset occurs in the June 19 observation. We speculate that the UT 2008 June 13 observation may have been effectively preflashed by the preceding science observations in the same way as the UT 2007 June 29 transit observation. We find that all forms of preflash reduce the slope of the subsequent ramp as compared to the non-preflashed secondary eclipse on UT 2007 June 30, but the H II regions consistently produce a larger reduction in the ramp slope than preflashes using a bright star.

We can describe the ramps in our $8\,\mu$m science data with the following functional form:

$$f = c_1 \left(1 - c_2 \exp\left(-\delta t / c_3\right) - c_4 \exp(-\delta t / c_5)\right), \qquad (2)$$

where f is the measured flux, δt is the elapsed time from the start of the observations, and c_1–c_5 are free parameters in the fit. Previous studies have elected to use either a single exponential (e.g., Harrington et al. 2007), a linear + log function of δt (e.g., Deming et al. 2007), or a quadratic function in $\log(\delta t)$ (e.g., Charbonneau et al. 2008; Knutson et al. 2008; Désert et al. 2009). However, in Agol et al. (2010) we find that the functional forms involving $\log(\delta t)$ produce eclipse depths that are correlated with the slope of the observed ramp function, while the single exponential does not provide a good fit to data with a steep ramp. We conclude that a double exponential function has enough degrees of freedom to fit a range of ramp profiles, while still avoiding correlations between the measured eclipse depths and the slope of the detector ramp. Although we require a double exponential function in order to fit the steeper, non-preflashed 2007 secondary eclipse observation in this study, we obtain comparable results with a single exponential term for our preflashed $8\,\mu$m data. We therefore elect to use this simpler single exponential in our subsequent analysis for all $8\,\mu$m visits except the 2007 secondary eclipse.

For our fits to phase curve data obtained on UT 2008 July 12–15, we select a 4 hr subset of data centered on the position of the transit or eclipse and use that in our fits. The first eclipse takes place at the start of the observations, which exhibit a residual ramp, and we therefore fit this light curve with the same single exponential as our other data. We use a linear function of time to fit the out-of-eclipse trends in the transit, which occurs in the middle of the observations, as well as the secondary eclipse toward the end of the observations. We find that the scatter in the central region of the time series near the transit, when the star is closest to the edge of the pixel, is higher than for either secondary eclipse or for the other $8\,\mu$m transit observations. Stevenson et al. (2010) found that the fluxes measured with the $5.8\,\mu$m *Spitzer* array sometimes display a weak dependence on the position of the star, which may be due to either flat-fielding errors or intrapixel sensitivity variations similar to those observed in the 3.6 and $4.5\,\mu$m arrays, although no such effect has been definitively detected in the $8\,\mu$m array to date. We test for the presence of position-dependent flux variations in our data by adding linear functions of x and y positions to each of our $8\,\mu$m transit fits and find that the χ^2 value of the resulting fits is effectively unchanged in all cases except for the UT 2008 July 14 transit, where it decreased from 37,186.6 to 37,177.7 for 33,636 points and six degrees of freedom. Using the Bayesian Information Criterion described in Stevenson et al. (2010), we conclude that this reduction in χ^2 is not significant, and we exclude these position-dependent terms in our subsequent analysis of the $8\,\mu$m data.

2.3. Transit and Eclipse Fits

We carry out simultaneous fits to determine the best-fit transit functions and detector corrections using a nonlinear Levenberg–Marquardt minimization routine (Markwardt 2009). We calculate our eclipse curve using the equations from Mandel & Agol (2002) assuming a longitude of pericenter equal to $334° \pm 10°$ (update based on complete set of published and unpublished radial velocity data; A. Howard 2010, private communication). The orbital eccentricity determined from the updated radial velocity data is 0.145 ± 0.017, but we choose to set the orbital eccentricity equal to 0.152 in our fits, which we calculate using the above longitude of pericenter and the published value of $e\cos(\omega) = 0.1368 \pm 0.0004$ from Stevenson et al. (2010). We find that the uncertainty in the calculated eccentricity is dominated by the uncertainty in ω, but this has a minimal impact on our transit fits. Our best-fit parameters change by less than 1σ for eccentricity values between 0.142 and 0.169, corresponding to $\pm 10°$ in ω, where our best-fit inclination is most sensitive to the assumed eccentricity (0.9σ change), $a/R_\star$ is somewhat sensitive (0.5σ change), and the best-fit radius ratios and transit times for individual fits are minimally sensitive ($<0.3\sigma$ change). Our nominal values for the eccentricity and longitude of pericenter result in a transit length of 60.9 minutes, 0.5 minutes longer than the zero eccentricity case. Using the same parameters for the secondary eclipse, which occurs shortly before periastron passage, produces a length of 62.6 minutes.

We fit the eight transits simultaneously and assume that the inclination and the ratio of the orbital semimajor axis to the stellar radius $a/R_\star$ are the same for all transits, but allow the planet–star radius ratio $R_p/R_\star$ and transit times to vary individually. Figure 1 shows the final binned data from these fits with the best-fit normalizations for the detector effects and transit light curves in each channel overplotted, and Figure 2 shows the binned data once these trends are removed, with
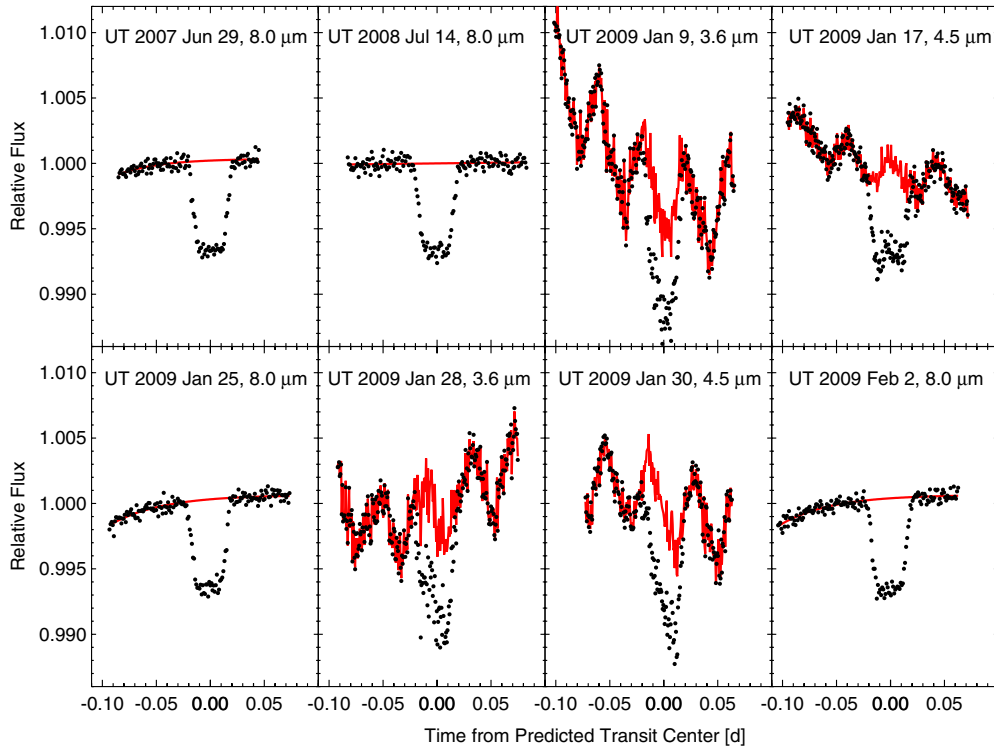
Figure 1. Raw photometry for the eight observed transits of GJ 436b, arranged in chronological order and with best-fit detector functions overplotted (solid red lines). Data have been binned in either 0.9 minute (3.6, 4.5 μm) or 1.5 minute (8.0 μm) bins.

(A color version of this figure is available in the online journal.)

best-fit transit curves overplotted. Figure 3 shows the root-n scaling for the residuals for these fits. Best-fit parameters are given in Tables 2 and 5.

2.3.1. A Comparison of ATLAS and PHOENIX Limb-darkening Models

We derive limb-darkening coefficients for the star using a Kurucz ATLAS stellar atmosphere model with $T_{\rm eff} = 3500$ K, $\log(g) = 5.0$, and [Fe/H] = 0 (Kurucz 1979, 1994, 2005), where we take the flux-weighted average of the intensity profile in each IRAC band and then fit this profile with four-parameter nonlinear limb-darkening coefficients (Claret 2000). We also derive limb-darkening coefficients for a PHOENIX atmosphere model (Hauschildt et al. 1999) with the same parameters and list both sets of coefficients in Table 3. We trim the maximum stellar radius in the PHOENIX models, which is set to an optical depth of 10^{-9}, to match the level of the $\tau = 1$ surface in each *Spitzer* band. We estimate the location of this surface by determining when the intensity relative to that at the center of the star first drops below e^{-1} and find that the new stellar radius is 0.09%–0.10% smaller than the old $\tau = 10^{-9}$ value. We find that we can achieve satisfactory four-parameter nonlinear fits to the PHOENIX intensity profiles only when we exclude points where $\mu < 0.025$, whereas the ATLAS models are well described by fits including this region.

As illustrated in Figure 4, the PHOENIX model predicts stronger limb darkening in all bands as compared to the ATLAS model, with the largest differences in the 3.6 μm band. When we compare our best-fit transit parameters using either the ATLAS or PHOENIX limb-darkening coefficients, we find that the best-fit planet–star radius ratios are 0.8σ–1.2σ (0.5%–0.6%) deeper in the 3.6 μm band, 0.06σ–0.07σ (0.04%–0.05%) smaller in the 4.5 μm band, and 0.3σ–0.4σ (0.2%–0.3%) larger in the

Table 3
Four-parameter Nonlinear Limb-darkening Coefficients[a]

Model	Band (μm)	c_1	c_2	c_3	c_4
ATLAS	3.6	1.122	-1.852	1.675	-0.582
ATLAS	4.5	0.749	-0.917	0.718	-0.230
ATLAS	5.8	0.815	-1.147	0.947	-0.310
ATLAS	8.0	0.770	-1.141	0.942	-0.304
PHOENIX	3.6	1.284	-1.751	1.433	-0.470
PHOENIX	4.5	1.203	-1.796	1.512	-0.500
PHOENIX	5.8	0.918	-1.264	1.064	-0.358
PHOENIX	8.0	0.619	-0.762	0.645	-0.220

Note. [a] Both models assume $T_{\rm eff} = 3500$ K and [Fe/H]=0. The ATLAS model uses $\log(g) = 5.0$ and the PHOENIX model uses $\log(g) = 4.76$ for better consistency with the radius and luminosity in Torres (2007), but empirical tests show that the assumed surface gravity has a negligible effect on the resulting limb-darkening profiles. For a definition of this limb-darkening law, see Claret (2000).

8.0 μm band for the PHOENIX models. The best-fit values for the inclination and $a/R_\star$ increase by 1.0σ (0.6%) and 0.9σ (0.04%), respectively, for the PHOENIX model fits; this is a product of the stronger limb-darkening profile, as GJ 436b's relatively high impact parameter creates a partial degeneracy between the limb-darkening profile and the other transit parameters.

We examine the relative importance of the assumed stellar parameters by comparing two PHOENIX models with effective temperatures of 3400 K and 3600 K. We find that for this 200 K range in effective temperature, the best-fit planet–star radius ratios change by 0.11σ–0.16σ at 3.6 μm, 0.07σ–0.09σ at 4.5 μm, and 0.10σ–0.12σ at 8.0 μm. The changes in the best-fit values for the inclination and $a/R_\star$ were similarly small, 0.04σ and 0.4σ, respectively. We therefore conclude that changes in
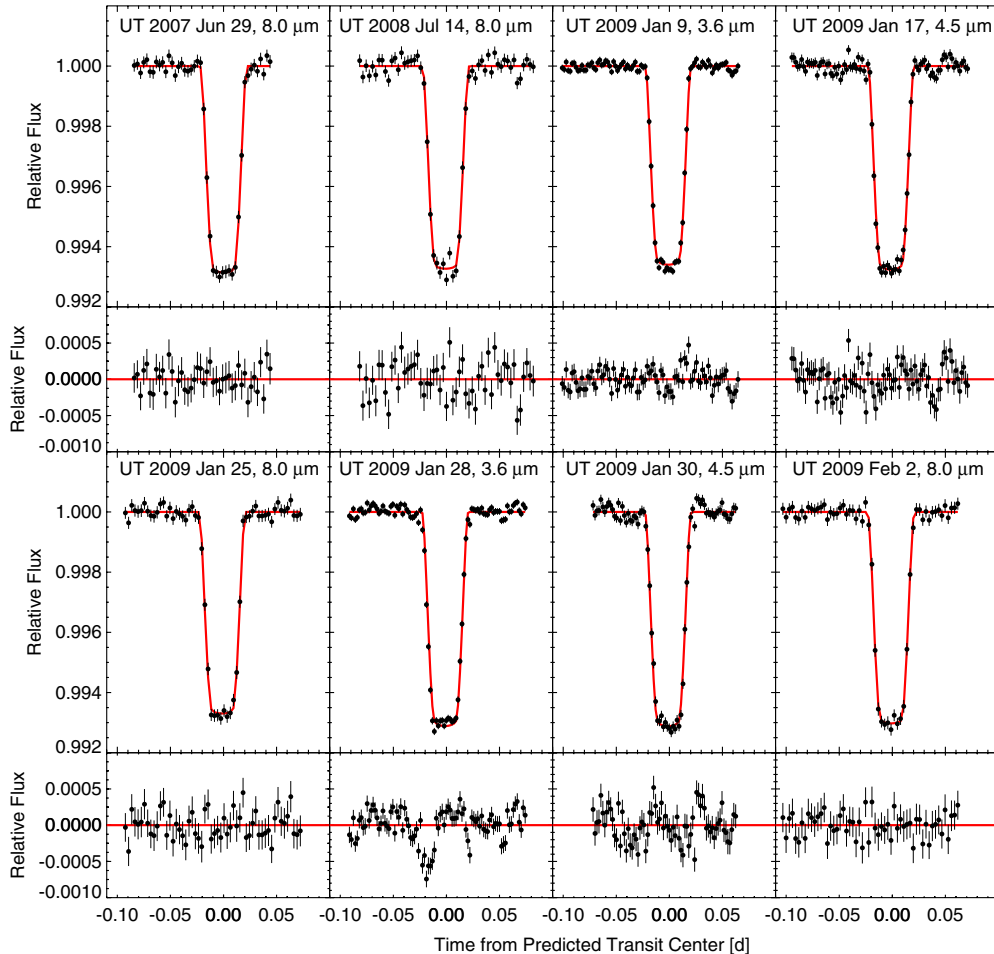
Figure 2. Photometry for the eight observed transits of GJ 436b after the best-fit corrections for instrument effects are removed, arranged in chronological order. Data have been binned in either 2.7 minute (3.6, 4.5 μm) or 4.3 minute (8.0 μm) bins. Best-fit transit curves are overplotted in red, and the residuals from each fit are shown in the lower panel. In this plot, we have assumed a constant ephemeris for the planet rather than using the best-fit transit times. Note that although the out-of-transit residuals for the second 3.6 μm observation on UT 2009 January 28 appear to be relatively Gaussian, there are additional variations during the transit that are not well accounted for by the best-fit transit light curve. These variations are likely due to occultations of spots or faculae by the planet. The residuals for the 4.5 μm transit observed on UT 2009 January 30 display excess correlated noise both in and out of transit, most likely due to an imperfect correction for the sharp flux variations caused by the star's location at the edge of a pixel.

(A color version of this figure is available in the online journal.)

the stellar effective temperature of less than 200 K are negligible for the purposes of our transit fits. We also compute PHOENIX model intensity profiles for $0.0 < $ [Fe/H] $ < +0.3$, but we find that the differences between models are much smaller than for our 200 K change in the effective temperature.

As there are currently few observational constraints on limb-darkening profiles for main-sequence stars (e.g., Claret 2008, 2009), and even fewer constraints for M stars in the mid-infrared, we also consider simultaneous transit fits in which we allow quadratic limb-darkening coefficients in each band to vary as free parameters. As a result of the planet's high impact parameter, our observations do not directly constrain the limb-darkened intensity for values of $\theta \leqslant 50°$, corresponding to $\mu \geqslant 0.64$, as the planet does not cross this region on the star. However, we can infer the limb-darkening profile in this region if we assume a simple quadratic limb-darkening law. We require the intensity profile computed from these coefficients to be always less than or equal to one (i.e., no limb brightening) and require the relative intensity at the edge of the star to be greater than or equal to the equivalent K-band limb-darkening from Claret (2000). The dotted lines in Figure 4 show the resulting

best-fit limb-darkening profiles in each band; these profiles show less contrast than either model, but the ATLAS models appear to provide the closest match.

This agreement is reflected in the χ^2 values for the simultaneous transit fits; the total χ^2 for the best-fit quadratic coefficients is 536,729.25, for the 3500 K ATLAS limb-darkening coefficients it is 536,733.98, and for the [3400, 3500, 3600] K PHOENIX models it is [536,740.69, 536,739.75, 536,738.81], for 536,798 points and either 53 (with fixed limb-darkening) or 59 (with freely varying quadratic limb-darkening coefficients) free parameters. We use the ATLAS limb-darkening coefficients in our subsequent analysis, as they produce a marginally better agreement with the best-fit profiles than the PHOENIX models. Although the χ^2 value for the best-fit quadratic limb-darkening coefficients is formally smaller than that of either model, this fit also contains six additional degrees of freedom, making the difference negligible.

As an additional test, we also repeat our fits with the limb-darkening coefficients fixed to zero in all bands. This produces planet–star radius ratios that are 1.6σ–2.4σ (1.1%) smaller in the 3.6 μm band, 1.7σ–2.0σ (1.1%) smaller in the 4.5 μm band,
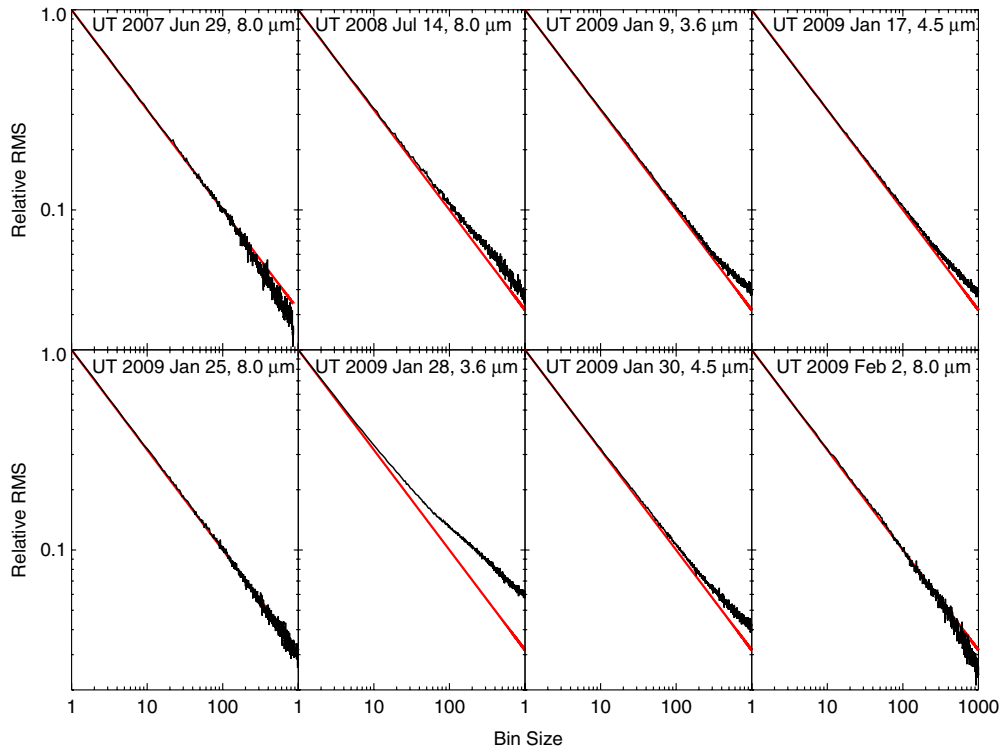
Figure 3. Standard deviation of residuals vs. bin size for the eight transits observed with *Spitzer*, arranged in chronological order. Observations were taken at 8.0, 8.0, 3.6, and 4.5 μm (top row, left to right), 8.0, 3.6, 4.5, and 8.0 μm (bottom row, left to right), respectively. The red curve shows the predicted root-n scaling expected for Gaussian noise.
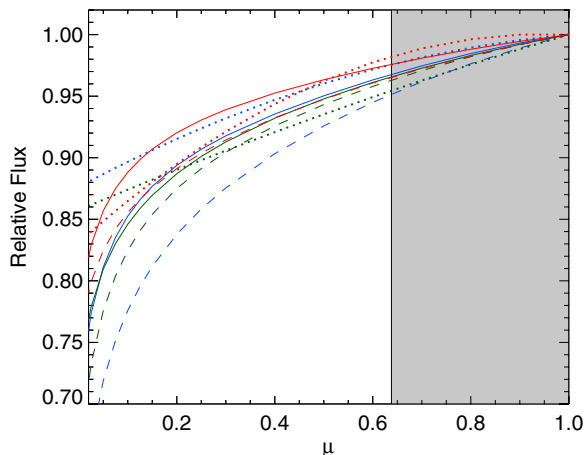

Figure 4. Comparison of the model limb darkening as a function of $\mu = \cos(\theta)$, where θ ranges from 0° at the center of the star to 90° at the edge. We show limb darkening from four-parameter nonlinear limb-darkening coefficients obtained using ATLAS (solid lines) or PHOENIX (dashed lines) stellar atmosphere models, as well as the best-fit quadratic limb darkening obtained by a fit to our data (dotted lines). The color indicates the bandpass, including 3.6 μm (blue), 4.5 μm (green), and 8.0 μm (red) *Spitzer* bands. The gray shaded region indicates values of μ for which we have no direct observational constraints, as the planet does not cross this part of the star during its transit.

and 0.9σ–1.1σ (0.6%–0.7%) smaller in the 8.0 μm band. The best-fit inclination and $a/R_\star$ are 2.7σ (1.5%) and 2.0σ (0.1%) smaller, respectively. The χ^2 value for this fit is 536,738.04, equivalent to the PHOENIX model fits and marginally worse than the ATLAS models or the fitted limb-darkening coefficients. This fit confirms the pattern suggested earlier, namely that stronger limb darkening leads to larger planet–star radius ratios and larger values for the inclination and $a/R_\star$. If we consider the

constraints imposed by the transit fits, stronger limb darkening means that for a grazing transit the planet must occult a relatively larger fraction of the star in order to produce the same apparent transit depth. This effect will be even larger in visible light, and we conclude that accurate limb-darkening coefficients are essential when calculating the planet–star radius ratio and corresponding transmission spectrum for near-grazing transits.

It is difficult to diagnose the origin of the disagreement between ATLAS and PHOENIX stellar atmosphere models for GJ 436; Kurucz (2005) notes that the ATLAS models should not be used for stellar effective temperatures below 3500 K, as they do not include important low-temperature opacity sources such as TiO and VO. However, these molecules primarily affect the star's visible and near-infrared spectra, and at 3500 K they are still relatively weak (Cushing et al. 2005). Both disk-integrated and intensity spectra for the ATLAS models in this temperature range are featureless longward of 2.4 μm, with the exception of the CO band between 4.3 and 5.0 μm, whereas PHOENIX spectra also show clear molecular band structures, mainly due to H_2O and OH, between 2.5–3.6 μm and 6.5–8.0 μm with corresponding increases in the amount of limb darkening in these bands. The presence of the CO band in both model sets would appear to explain the relatively good agreement in limb-darkening profiles for the 4.5 μm Spitzer bandpass, but we were unable to determine the reason behind the missing mid-IR H_2O absorption features in the ATLAS models, which incorporate the strongest water lines (Kurucz 1999) from the Ames list of Partridge & Schwenke (1997). It is possible that the spherical geometry used in the PHOENIX models (ATLAS models use a plane-parallel geometry) may also affect the resulting limb-darkening profiles (Orosz & Hauschildt 2000; Claret & Hauschildt 2003), but we find that PHOENIX models computed with a planet-parallel geometry show nearly identical
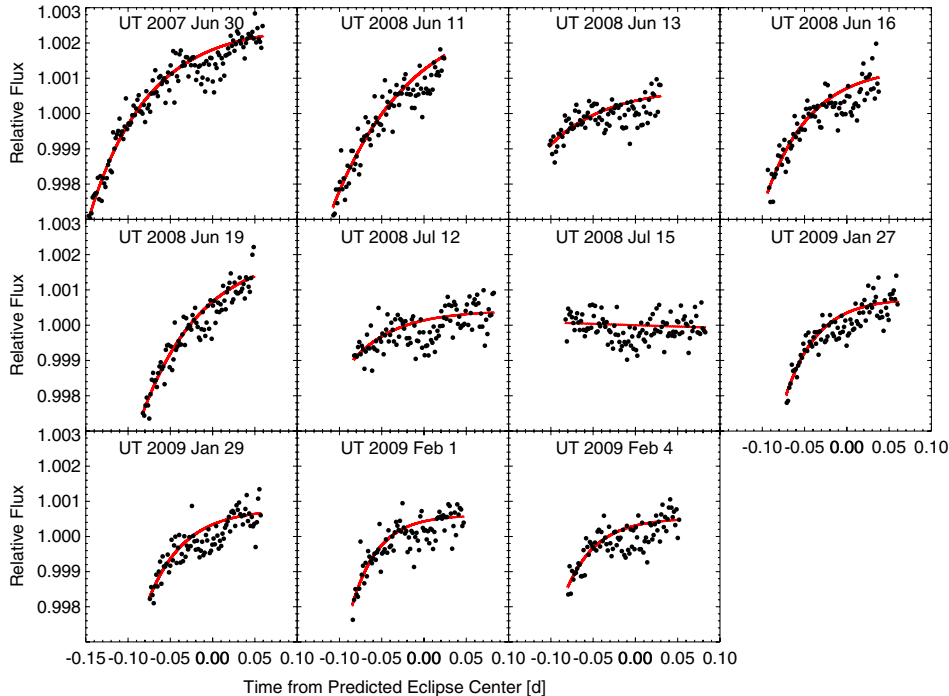
Figure 5. Raw photometry for eleven 8 μm secondary eclipses of GJ 436b, arranged in chronological order. Data have been binned in 2.2 minute bins, and the best-fit corrections for detector effects in each visit are overplotted in red.

(A color version of this figure is available in the online journal.)

limb darkening, with the exception of an exponential decline in the optically thin limb. We conclude that the differing opacities in the 3.6 and 8.0 μm bands appear to be the most likely explanation for the disagreement between the limb-darkening profiles at these wavelengths. In this case, the change in the χ^2 value indicates that the differences between the two models are not statistically significant for this data set; near-IR grism spectroscopy of transits of GJ 436b, such as those obtained by Pont et al. (2008b), might help to better distinguish between these models.

2.3.2. Error Analysis

We calculate uncertainties for our best-fit transit parameters using a Markov Chain Monte Carlo (MCMC) fit (see, for example, Ford 2005; Winn et al. 2007) with a total of 6×10^6 steps, 14 independent chains, and 53 free parameters. Free parameters in the fits include $a/R_\star$, i, eight individual estimates of $R_P/R_\star$, eight transit times, eight constants, a linear function of time and linear and quadratic terms in x and y for each of the 3.6 and 4.5 μm transits (20 variables total), the amplitude c_2 and decay time c_3 from Equation (2) for the exponential fits to three 8.0 μm transits (6 variables total), and a linear function of time for the other 8.0 μm visit. We assume a constant error for the points in each individual transit light curve, defined as the uncertainty needed to produce a reduced χ^2 equal to one for the best-fit transit solution.

We initialize each chain at a position determined by randomly perturbing the best-fit parameters from our Levenberg–Marquardt minimization. After running the chain, we search for the point where the χ^2 value first falls below the median of all the χ^2 values in the chain (i.e., where the code had first found the optimal fit) and discard all steps up to that point. We calculate the uncertainty in each parameter as the symmetric range about the median containing 68% of the points in the

distribution, except for the inclination and $a/R_\star$, which we allow to have asymmetric error bars spanning 34% of the points above and below the median, respectively. The distribution of values was very close to symmetric for all other parameters, and there did not appear to be any strong correlations between variables. As a check we also carried out a residual permutation error analysis (Gillon et al. 2007b; Winn et al. 2008), which is sensitive to correlated noise in the light curve, on each individual transit. At the start of each new permutation, we randomly drew values for the inclination and $a/R_\star$ from the simultaneous MCMC distribution and then fit for the corresponding best-fit values for the transit time and $R_p/R_\star$ in that step. This ensures that our resulting error distributions for individual transit times and $R_p/R_\star$ values also take into account the uncertainties in the best-fit values for the inclination and $a/R_\star$. In each case where both an MCMC and residual permutation uncertainty are available for a given parameter, we use the higher of the two values. We find that the MCMC fits generally produce larger uncertainties for the 8 μm observations, whereas for 3.6 and 4.5 μm data sets, which have higher levels of correlated noise, the residual permutation uncertainties are typically 50% larger than the MCMC errors.

2.3.3. Secondary Eclipse Fits

We fit the secondary eclipses individually using the best-fit values for inclination and $a/R_\star$ from our transit fits but allowing the eclipse depths and times to vary freely. Figure 5 shows the final binned data from these fits with the best-fit normalizations for the detector ramp in each channel overplotted, and Figure 6 shows the binned data once these trends are removed, with best-fit eclipse curves overplotted. Figure 7 shows the root-n scaling for the residuals from these fits. Best-fit parameters for individual eclipses are given in Table 4, and the error-weighted average (i.e., weights equal to $1/\sigma^2$) of these eclipse depths is
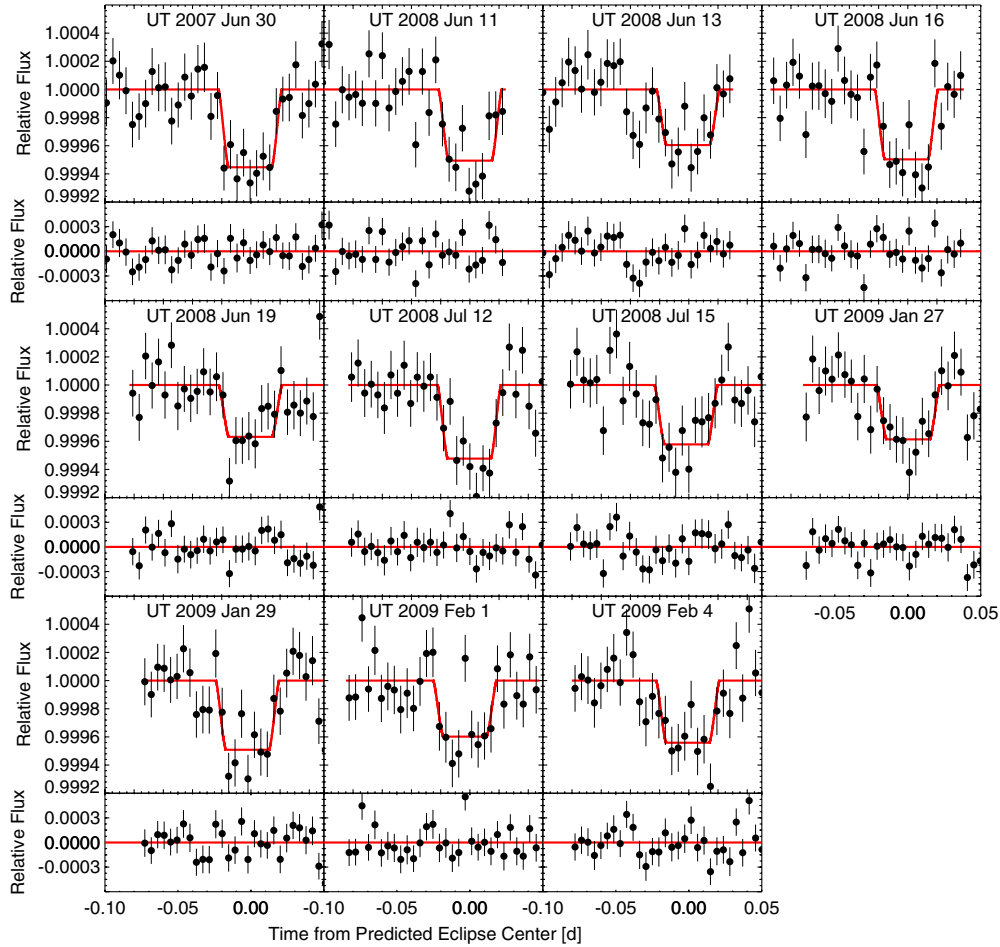
Figure 6. Photometry for eleven 8 μm secondary eclipses of GJ 436b, arranged in chronological order. Data have been binned in 6.4 minute bins, and the best-fit eclipse curve for each visit is overplotted in red. The residuals for each visit are shown in the panels below the eclipses.

(A color version of this figure is available in the online journal.)

Table 4
Individual Best-fit 8 μm Secondary Eclipse Parameters

UT Date	Depth (%)	Eclipse Center (BJD)	$O - C$ (minutes)[a]
UT 2007 Jun 30	0.0553 ± 0.0083	2454282.3329 ± 0.0016	−0.2 ± 2.3
UT 2008 Jun 11	0.0506 ± 0.0110	2454628.6850 ± 0.0017	2.0 ± 2.4
UT 2008 Jun 13	0.0395 ± 0.0097	2454631.3281 ± 0.0021	0.9 ± 3.0
UT 2008 Jun 16	0.0497 ± 0.0087	2454633.9716 ± 0.0013	0.3 ± 1.9
UT 2008 Jun 19	0.0368 ± 0.0089	2454636.6162 ± 0.0021	1.2 ± 3.0
UT 2008 Jul 12	0.0523 ± 0.0090	2454660.4112 ± 0.0019	1.2 ± 2.8
UT 2008 Jul 15	0.0422 ± 0.0078	2454663.0537 ± 0.0040	−0.9 ± 5.8
UT 2009 Jan 27	0.0386 ± 0.0087	2454858.7047 ± 0.0026	2.8 ± 3.8
UT 2009 Jan 29	0.0491 ± 0.0088	2454861.3460 ± 0.0015	−1.0 ± 2.2
UT 2009 Feb 1	0.0398 ± 0.0086	2454863.9889 ± 0.0017	−2.4 ± 2.4
UT 2009 Feb 4	0.0441 ± 0.0087	2454866.6355 ± 0.0023	1.4 ± 3.3

Note. [a] Observed minus calculated transit times. Predictions use the best-fit ephemeris of $T_c = 2454865.083208 \pm 0.000042$ BJD and $P = 2.6438979 \pm 0.0000003$ days, and an orbital phase of 0.58685 ± 0.00017 from Table 5.

reported in Table 5. We find that fixing the time of eclipse to a constant value, defined here as the best-fit orbital phase, does not significantly change our best-fit eclipse depths, nor does it reduce the uncertainties in our measurement of those depths. We calculate the uncertainties on individual eclipses using both an MCMC analysis and a residual permutation error analysis, again taking the higher of the two values as the final uncertainty for each parameter.

3. RESULTS

3.1. Orbital Ephemeris and Limits on Timing Variations

We fit the transit times given in Table 2, together with the transit times published in Pont et al. (2008a), Bean & Seifahrt (2008), Coughlin et al. (2008), Alonso et al. (2008), Shporer et al. (2009), Cáceres et al. (2009), and Ballard et al. (2010a),
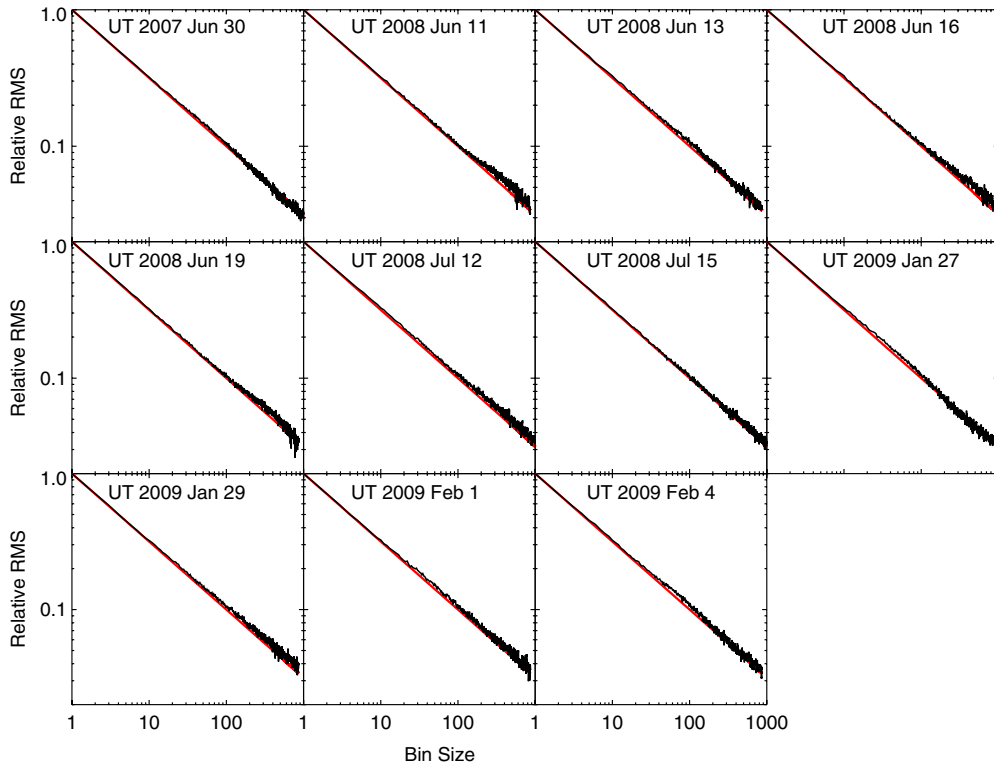
Figure 7. Standard deviation of residuals vs. bin size for the eleven secondary eclipses observed with *Spitzer*, arranged in chronological order. The red curve shows the predicted root-*n* scaling expected for Gaussian noise, and the lack of any excess noise for large bin sizes suggests that these light curves should be well described by the standard MCMC error analysis.

with the following equation:

$$T_c(n) = T_c(0) + n \times P, \qquad (3)$$

where T_c is the predicted transit time as a function of the number of transits elapsed since $T_c(0)$ and P is the orbital period. We find that $T_c = 2454865.083208 \pm 0.000042$ BJD and $P = 2.6438979 \pm 0.0000003$ days. As demonstrated by Figure 8, the 34 published transit times appear to be markedly inconsistent with a constant orbital period, with the most statistically significant outliers (6.2σ and 7.1σ, respectively), occurring during the sequence of eight transits observed by the EPOXI mission between UT 2008 May 5–29 (Ballard et al. 2010a). The most significant deviations in the *Spitzer* transit data presented here occur during the last three visits (UT 2009 January 28–February 2) and range between -3.1σ and $+3.5\sigma$ in significance. Given the size of these discrepancies, it is perhaps not surprising that the reduced χ^2 value for the linear fit to Equation (3) is 6.8 (total χ^2 of 216.4, 34 points, two degrees of freedom). It is unlikely that the observed deviations could be explained by perturbations from a previously unknown second planet in the system, as the measured transit times shift by as much as several minutes on timescales of only a few days (i.e., a single planet orbit). As we discuss in more detail in Section 4.1.3, we believe that the presence of occulted star spots in a subset of the transit light curves is the most likely explanation for the observed deviations.

We carry out a similar fit to the secondary eclipse times given in Table 4, along with the additional secondary eclipse times reported in Stevenson et al. (2010), and find that $T_c(0) = 2454866.63444 \pm 0.00082$ BJD and $P = 2.6438944 \pm 0.0000071$ days. This period is consistent with the best-fit transit period to better than 1σ, and we therefore conclude that there
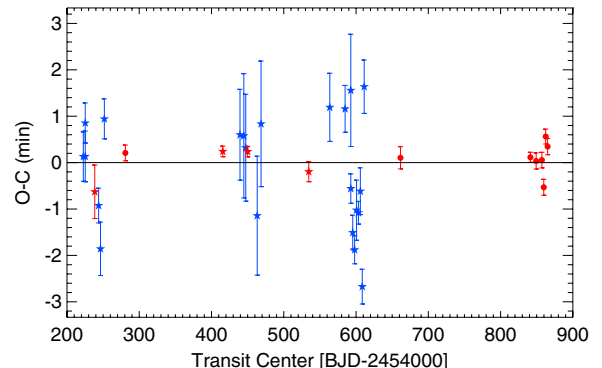

Figure 8. Observed minus calculated transit times using the new best-fit ephemeris. The dashed lines indicate the $\pm 1\sigma$ uncertainty in the predicted transit times, with a solid line at $O - C$ equal to zero. *Spitzer* measurements from this paper are plotted as filled circles, and previously published observations are shown as filled stars. The color of the points denotes the wavelength of the observations (blue for visible, red for IR). Moving from left to right, transits between 200 and 300 BJD $-2,454,000$ are from Shporer et al. (2009) and Cáceres et al. (2009), transits between 400 and 500 BJD $-2,454,000$ with small uncertainties are from Pont et al. (2008a), and those with large uncertainties are from Bean & Seifahrt (2008). Between 530 and 620 BJD $-2,454,000$, observations are from Coughlin et al. (2008), Alonso et al. (2008), and Ballard et al. (2010a). Visible-light transit observations typically show larger timing variations than the IR observations, indicating that spot occultations may be responsible for the apparent timing variations.

is no evidence for orbital precession in this system. We also see no evidence for statistically significant variations in the secondary eclipse times (see Figure 9), as would be expected if the shifted transit times were due to occulted spots, but our measurements are not precise enough to rule out timing variations of the same magnitude as those observed in the transit data. If we fix the orbital period to the value from the transit fits and

Table 5
Global System Parameters

Parameter	Value
Transit parameters	
$i(°)$	$86.699^{+0.034}_{-0.030}$
$a/R_\star$	$14.138^{+0.093}_{-0.104}$
$R_p/R_\star$ [a]	0.08311 ± 0.00026
Duration T_{14} (d)[b]	0.04227 ± 0.00016
T_{12} ($\approx T_{34}$)[b] (d)	0.01044 ± 0.00014
b	0.8521 ± 0.0021
a (AU)[c]	0.0287 ± 0.0003
$R_\star$ ($R_\odot$)[c]	0.437 ± 0.005
R_p ($R_\oplus$)[c]	3.96 ± 0.05
T_c (BJD)	$2454865.083208 \pm 0.000042$
P (d)	2.6438979 ± 0.0000003
Secondary eclipse parameters	
8 μm depth	$0.0452\% \pm 0.0027\%$
T^d_{bright}	$740 \pm$ m 1 K
Duration T_{14} (d)[e]	0.04347
T_{12} ($\approx T_{34}$)[e] (d)	0.00700
Orbital phase	0.58672 ± 0.00017
$e\cos(\omega)$	0.13775 ± 0.00027
$T_c(0)$ (BJD)	$2454866.63444 \pm 0.00082$
P (d)	2.6438944 ± 0.0000071

Notes.
[a] Calculated from the error-weighted average of the four 8 μm planet–star radius ratio; this value was used for secondary eclipse fits.
[b] The transit duration T_{14} is defined as the time from first to fourth contact (i.e., the start of ingress to the end of egress). T_{12} is the length of ingress, which is equal to the egress length in the limit of a circular orbit. Our best-fit transit ingress and egress lengths differ by less than 3 s.
[c] These parameters incorporate the stellar mass estimate of $0.452 \pm 0.013\ M_\odot$ from Torres (2007).
[d] Brightness temperature is defined as the temperature required to match the observed planet–star flux ratio in the 8 μm *Spitzer* band assuming that the planet radiates as a blackbody and using a Phoenix stellar atmosphere model ($T_{eff} = 3585$ K, $\log(g)$=4.843) for the star.
[e] The secondary eclipse duration and the length of ingress/egress were fixed in our fits.

subtract the 28 s light travel time delay for this system (Loeb 2005), we find that the secondary eclipses occur at an orbital phase of 0.58672 ± 0.00017, consistent with the best-fit phase from Stevenson et al. (2010).

We can use the offset in the best-fit secondary eclipse time to calculate a new estimate for $e\cos(\omega)$. We find that the secondary eclipse occurs 330.18 ± 0.67 minutes later on average than the predicted time for a circular orbit, including the correction for the light travel time. We can convert this to $e\cos(\omega)$ using the expression reported in Equation (19) of Pál et al. (2010). Note that this expression is more accurate than the commonly used approximation of $e\cos(\omega) \approx \frac{\pi\delta t}{2P}$ (e.g., Charbonneau et al. 2005; Deming et al. 2005), where δt is the delay in the measured secondary eclipse time and P is the planet's orbital period. We find that using the less accurate approximation gives $e\cos(\omega) = 0.13622 \pm 0.00026$, while the equation from Pál et al. yields $e\cos(\omega) = 0.13754 \pm 0.00027$, a 4$\sigma$ difference in this case (see also Sterne 1940; de Kort 1954). If we take the best-fit longitude of pericenter from the radial velocity fits, $334° \pm 10°$, we find an orbital eccentricity equal to 0.153 ± 0.014. This is consistent with the current best-fit orbital eccentricity from radial velocity data alone, $e = 0.145 \pm 0.017$ (A. Howard 2010, private communication).
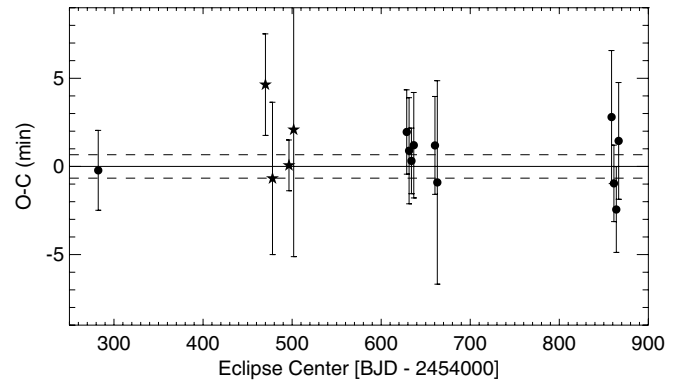
Figure 9. Observed minus calculated secondary eclipse times using the best-fit period from the transit fits and allowing the phase of the secondary eclipse to vary freely. Filled circles are eclipse times reported in this paper, and filled stars are additional 3.6 and 4.5 μm eclipse times from Stevenson et al. (2010). The solid line indicates the best-fit phase, with $\pm 1\sigma$ uncertainties plotted as dashed lines.

Table 6
$a/R_\star$ and Inclination Values from Independent Transit Fits

UT Date	λ (μm)	i (°)	$a/R_\star$
UT 2007 Jun 29	8.0	86.68 ± 0.12	14.11 ± 0.35
UT 2008 Jul 14	8.0	86.54 ± 0.12	13.67 ± 0.36
UT 2009 Jan 9	3.6	86.76 ± 0.07	14.40 ± 0.22
UT 2009 Jan 17	4.5	86.85 ± 0.10	14.60 ± 0.32
UT 2009 Jan 25	8.0	86.70 ± 0.14	14.19 ± 0.43
UT 2009 Jan 28	3.6	86.67 ± 0.07	13.96 ± 0.20
UT 2009 Jan 30	4.5	86.58 ± 0.10	13.76 ± 0.28
UT 2009 Feb 2	8.0	86.80 ± 0.14	14.49 ± 0.44

3.2. System Parameters from Transit Fits

In this work, we examine two transits obtained at 3.6 μm, two transits at 4.5 μm, and four transits at 8.0 μm. We carry out two sets of transit fits, one where the ratio of the orbital semimajor axis to the stellar radius $a/R_\star$ and the orbital inclination i are allowed to vary freely, and the other where they have a single common value for all visits. In all cases, we allow the planet–star radius ratio $R_p/R_\star$ and best-fit transit times to vary independently for each visit. In fits where $a/R_\star$ and i are allowed to vary individually, we find no evidence for statistically significant variations in either of these parameters (see Table 6) and therefore proceed assuming that these parameters have a single common value in our subsequent analysis. Our best-fit values for i, $a/R_\star$, and $R_p/R_\star$ are consistent with those reported by Ballard et al. (2010a) to better than 1σ, and the impact parameter b and transit duration $T = T_{14} - T_{12} = 0.0318 \pm 0.0007$ days that we derive from our fits are similarly consistent with the value reported by Pont et al. (2008b).

Although the best-fit orbital inclination and $a/R_\star$ appear to be consistent with a constant value over the approximately two-year period spanned by our observations, we do see evidence for statistically significant differences in the transit depths within the same *Spitzer* bandpass (see Figure 10). We would expect to see the transit depth vary with wavelength due to absorption from the planet's atmosphere, but this signal should remain constant from epoch to epoch for observations in the same bandpass. If we compare individual visits in a given bandpass, we find that the two 3.6 μm radius ratios, measured on UT 2009 January 9 and 28, are inconsistent at the 4.7σ level. The two 4.5 μm radius ratios, measured on UT 2009 January 17 and 30, differ by 2.9σ. The four 8 μm transits, measured on
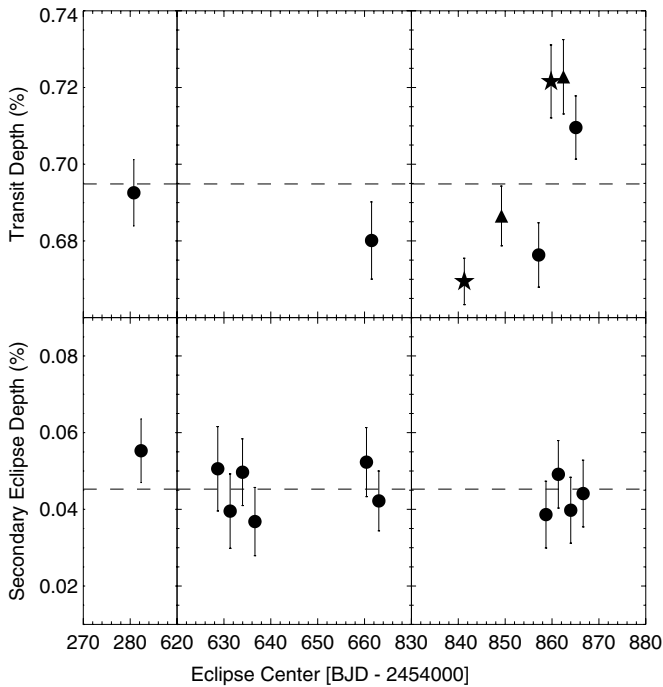
Figure 10. Best-fit transit (upper panel) and secondary eclipse (lower panel) depths as a function of time. Average transit/eclipse depths are shown as a dashed line in each plot. 3.6 μm observations are denoted with stars, 4.5 μm observations are denoted with triangles, and 8.0 μm observations are marked with solid circles. The most recent three transits are systematically high when compared to earlier transit visits; this appears to coincide with a decrease in the total visible-light flux from the star (Figure 11), suggesting that the fractional spot coverage on the star's visible face was increasing in time during the later part of our observations.

UT 2007 June 29, UT 2008 July 14, UT 2009 January 28, and UT 2009 February 2, differ from the error-weighted average by 0.2σ, 1.0σ, 1.5σ, and 2.0σ, respectively. These offsets are still present in the fits where the inclination and $a/R_\star$ are allowed to vary individually, indicating that the discrepancy cannot be due to a change in these two parameters.

4. DISCUSSION

4.1. Transit Depth Variations

In the sections below, we consider three possible explanations for the observed depth variations: first, that the effective radius of the planet is varying in time, second, that residual correlated noise in the data affected the best-fit transit solutions, and third, that spots or other stellar activity produced apparent depth variations.

4.1.1. A Time-varying Radius for the Planet

We first consider the possibility that the radius of the planet is changing in time, either due to thermal expansion of the atmosphere or the presence of a variable cloud layer at sub-mbar pressures. We require a change in radius of approximately 4% in order to match both of the measured 3.6 μm transit depths; if this change is due to thermal expansion, we can estimate the energy input required using simple scale arguments.

The effective change in the planet's radius due to heating of the atmosphere depends on both the amount of heating and the range in pressures over which this heating takes place. We use the secondary eclipse depths in Section 4.3 to place an upper limit on the allowed change in temperature at the level

of the mid-IR photosphere and then calculate the corresponding range in pressure that must be heated by this amount in order to increase the radius of the planet by 4%. If we assume a hydrogen atmosphere with a baseline temperature of 700 K, we find a corresponding scale height of approximately 240 km, where the scale height is defined as $H = \frac{kT}{\mu g}$, T is the temperature of the atmosphere, μ is the mean molecular weight, and g is the surface gravity. We know from the secondary eclipse observations described in Section 4.3 that the temperature of the planet's dayside atmosphere must change by less than 30%, which would correspond to an upper limit of 100 km on corresponding changes in the planet's scale height.

In order to calculate the required energy input to produce the observed change in radius, we must first determine the range of pressures affected by this heating. We model the planet as an interior region with a constant temperature, surrounded by an outer envelope that expands and contracts freely with changing temperature. We set the upper boundary on this region equal to 50 mbar, corresponding to the approximate location of the $\tau = 1$ surface in the mid-infrared. As illustrated in Figure 14, opaque clouds at this pressure suppress but do not entirely remove absorption features in the planet's transmission spectrum at these wavelengths, making this a reasonable estimate for the location of the $\tau = 1$ surface. We assume that when the planet is heated the scale height changes by 100 km, which requires the lower boundary of the heated region to be located at a pressure of approximately 1 bar in order to produce a 1% expansion in radius. If we then calculate the change in the planet's gravitational energy corresponding to this expansion, we find that an energy input of approximately 10^{26} J is required. Repeating this calculation for a 4% increase in radius, we find a lower boundary at 8000 bars and a corresponding energy input of 10^{30} J. The insolation received by the planet is 10^{20} W, which gives an energy budget of 10^{25} J per orbit. When we examine Figure 10 we find that the observed change in radius occurs primarily between the third and fourth visits (UT 2009 January 25–28). This would require an energy input as much as 10^5 times higher than the total insolation over this epoch, which is clearly unphysical.

One alternative explanation for the observed change in radius would be to invoke the presence of intermittent, high-altitude clouds. Such clouds could produce a change in the apparent radius of the planet across multiple bands without requiring any actual heating or cooling of the atmosphere. In this picture, smaller radii for the planet would correspond to the cloud-free state, while larger radii would require the presence of an additional cloud layer. A change of 4% in apparent radius would require the clouds to form at a pressure approximately 100 times lower than the location of the nominal cloud-free radius. In Section 4.2, we find that the average pressure of the $\tau = 1$ surface for the nominal methane-poor (green) model between 3 and 10 μm is 40 mbar, indicating that the clouds would have to extend to 0.4 mbar to explain the largest measured 3.6 μm radius for the planet. This conclusion is reasonably independent of our assumed composition, as the average $\tau = 1$ surface for the methane-rich (blue) model is located at 30 mbar. Gravitational settling would presumably pose a challenge for cloud layers at sub-mbar levels, but vigorous updrafting of condensate particles might compensate for this effect. The broadband nature of the data presented here makes it difficult to directly test this hypothesis; we therefore recommend the acquisition of high signal-to-noise, near-infrared grism spectroscopy over multiple transits in order to resolve this issue. A 0.5 mbar cloud

layer would lead to a near-featureless transmission spectrum whereas a lower cloud layer would still exhibit many of the same absorption features as a cloud-free atmosphere. Such a data set would also allow us to test the theory, outlined in Section 4.1.3, that the observed transit depth variations are due to the occultation of regions of non-uniform brightness on the surface of the star, as these regions should also produce a wavelength-dependent effect.

4.1.2. Poorly Corrected Systematics

It is possible that poorly corrected instrument effects, such as the intrapixel sensitivity variations at 3.6 and 4.5 μm, or the detector ramp at 8.0 μm, might lead to variations in the measured transit depth. Because there is complete overlap between the positions spanned by the star in the in-eclipse and out-of-eclipse data for all 3.6 and 4.5 μm visits, fits that inadequately describe the pixel response as a function of position should fail equally for both sections of the light curve. The UT 2009 January 30 transit serves as an example of imperfectly removed detector effects, as the residuals display a sawtooth signal with a shape and timescale similar to the original intrapixel sensitivity variations (see Section 2.1 for a more detailed discussion of this light curve). Conversely, it is much more difficult to explain the 3.6 μm transit on UT 2009 January 28 with this scenario, as there appears to be a large dip in residuals during ingress, but when the star spans the same pixels in the out-of-eclipse data we see no comparable deviations.

In this section, we consider an alternate decorrelation function that better accounts for small-scale variations in the intrapixel sensitivity function as discussed in Ballard et al. (2010b). Following the discussion in Ballard et al., we describe the intrapixel sensitivity variations using a position-weighted average of the time series after the best-fit transit function and linear function of time from the position fits described above has been divided out. Unlike Equation (1), this formalism does not assume a functional form for the intrapixel sensitivity variations and therefore should in principle produce an unbiased correction for these variations. We calculate the weighting function as follows:

$$W(x_i, y_i) = \frac{\sum_{i \neq j} \exp\left(-\frac{(x_j - x_i)^2}{2\sigma_x^2}\right) \exp\left(-\frac{(y_j - y_i)^2}{2\sigma_y^2}\right) f_j}{\sum_{i \neq j} \exp\left(-\frac{(x_j - x_i)^2}{2\sigma_x^2}\right) \exp\left(-\frac{(y_j - y_i)^2}{2\sigma_y^2}\right)}, \quad (4)$$

where x_i and y_i are the x and y positions of the ith frame, x_j and y_j are the x and y positions for the rest of the time series. We optimize our choice of σ_x and σ_y to produce the smallest possible scatter in the final time series when we fix the transit light curve to the best-fit solutions listed in Table 2. We find that the preferred values range between 0.0053–0.0120 pixels in σ_x and 0.0024–0.0045 pixels in σ_y for the four 3.6 and 4.5 μm transits examined here. For ease of computation, we bin our time series in intervals corresponding to one point per original set of 64 images (in some instances there are less than 64 images in a given bin after removing outliers) and iteratively calculate the weighting function and the linear function of time plus transit fits until we converge to a consistent solution.

Once we have a final solution we calculate the weighting function for the unbinned data and carry out a final fit for the transit function to determine our best-fit transit depth. In this case, we fix the inclination and $a/R_\star$ to their best-fit values from the simultaneous fits to all transits described in Section 2.3, which allows us to fit each transit individually using the weighting function while still preserving the constraints

imposed in a simultaneous fit. We find that in all cases we obtain transit depths and times that are consistent with the values from our fits using Equation (1), with a standard deviation that is comparable or slightly worse than that achieved with our polynomial fits.

We also carried out a second set of fits in which we derived our corrections for the intrapixel sensitivity variations using only the out-of-transit data and found that our best-fit planet–star radius ratios changed by less than 0.4σ in all cases. Because the star samples the same regions of the pixel in both the in-transit and out-of-transit data, it is possible to obtain an equivalently good correction for the intrapixel sensitivity variations using only the out-of-transit points. Conversely, this means that poor corrections for this effect should produce equally large deviations in both the in-transit and out-of-transit regions of the light curve. As we will discuss in the following section, we find that the residuals for the deepest transits in these two bands have a significantly higher rms in transit than out of transit. This behavior is inconsistent with our expectations for poorly corrected instrument effects, and we therefore conclude that it is unlikely that these effects are responsible for the discrepant transit depths measured at 3.6 and 4.5 μm.

At 8.0 μm we fit the data with a single or double exponential function to describe the smoothly varying detector ramp. In Agol et al. (2010), we conclude that this functional form avoids correlations between the slope of the ramp and the measured transit or eclipse depth; however, we check this assertion using our 8 μm data as well. For our 8 μm transit fits, we find that the exponential term has a coefficient of [0.00156, 0.00000, 0.00288, 0.00299], corresponding to planet–star radius ratios of [0.08234, 0.08162, 0.08138, 0.08336], where we have set the amplitude of the exponential term to zero for the transit occurring in the middle of our 70 hr phase curve observation. For the 11 secondary eclipse observations, we find coefficients of [0.00645, 0.00627, 0.00194, 0.00433, 0.00534, 0.00140, 0.00000, 0.00359, 0.00321, 0.00353, 0.00262], corresponding to eclipse depths of [0.0552, 0.0507, 0.0395, 0.0495, 0.0367, 0.0523, 0.0421, 0.0386, 0.0491, 0.0397, 0.0441], respectively, where we have set the exponential coefficient to zero for the secondary eclipse at the end of the phase curve observation. We find no evidence for any correlation between the slope of the exponential function and the measured transit or secondary eclipse depths. As an additional check we also confirm that there is no correlation between these depths and either the measured sky background or the total stellar flux given in Table 1.

4.1.3. Stellar Variability

The presence of spots or faculae on the visible face of the star can have two distinct effects on the measured light curve for a transiting planet. Non-occulted spots on the visible face of the star reduce the star's total flux, increasing the measured transit depth, while spots occulted by the planet cause a small positive deviation in the light curve with a timescale proportional to the physical size of the occulted spot (e.g., Rabus et al. 2009) and occulted faculae would have the opposite effect. The early K dwarf HD 189733 ($T_{\rm eff} = 5100$ K) is perhaps the best-studied example of an active star with a transiting hot Jupiter (e.g., Bakos et al. 2006; Pont et al. 2008a; Désert et al. 2011a), but the late G dwarf CoRoT-2 ($T_{\rm eff} = 5600$ K) also exhibits a high level of spot activity that may have resulted in early overestimates of its planet's inflated radius (Guillot & Havel 2011). This problem is likely to be even more common for M dwarfs, and in fact several
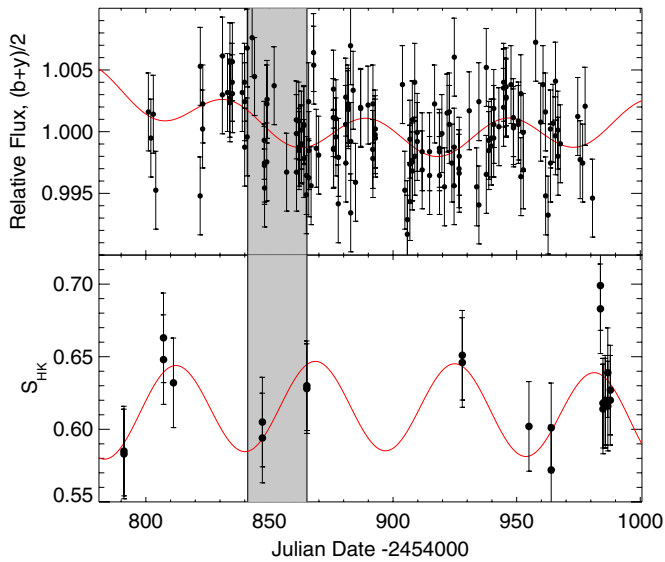
Figure 11. Upper panel plots APT measurements of variation in the averaged Strömgren *b+y* band fluxes (filled circles) obtained for GJ 436 between UT 2008 November 30 and UT 2009 May 29, where the epoch of our 2009 *Spitzer* observations is denoted by the gray shaded region. These bandpasses are sensitive to the rotation-modulated flux of the star, which we find has a best-fit period of 57 days during this epoch. We overplot a red curve showing our best sine-curve fit for the spot modulation, together with a quadratic function of time to describe the evolution of the spot coverage on longer timescales. The lower panel shows the measured S_{HK} values for GJ 436 from the Keck HIRES instrument during this same period (Isaacson & Fischer 2010), with a best-fit sine + quadratic function overplotted in red. Error bars for both panels are set equal to the standard deviation of the residuals. Although the best-fit period of 57 days for the S_{HK} data during this epoch is only marginally significant, the S_{HK} values appear to be anti-correlated with the flux variations. This is consistent with a model in which increased magnetic activity is associated with the presence of spots or other dark regions on the surface of the star.

(A color version of this figure is available in the online journal.)

instances of occulted spots were reported in transit light curves for the super-Earth GJ 1214b, which orbits a 3000 K primary (Carter et al. 2011; Berta et al. 2011; Kundurthy et al. 2011).

Although it is important to correct for these effects in any transit fit, it is particularly crucial when comparing non-simultaneous, multi-wavelength transit observations such as the ones described in this paper, which require a relative precision of better than a part in 10^{-4} in the measured transit depths. We evaluate the likely impact of GJ 436's activity on the measured transit depths using several complementary approaches. First, we estimate the average activity level on GJ 436 by measuring the amount of emission in the cores of the Ca II H&K lines; in Knutson et al. (2010), we determined that GJ 436 had an average S_{HK} of 0.620. Isaacson & Fischer (2010) found that other stars in the California Planet Search database with similar $B - V$ colors have S_{HK} values ranging between 0.5 and 2.0, indicating that GJ 436b is relatively quiet for its spectral type. Demory et al. (2007) report that this star's rotation period is greater than 40 days, consistent with upper limits on $v \sin i$ of 1 km s^{-1} (Jenkins et al. 2009), also suggesting that it is likely to be relatively old and correspondingly quiet. The upper limit of 3 km s^{-1} on $v \sin i$ from spectroscopy (Butler et al. 2004) is also consistent with an inclined or pole-on viewing geometry, although it is not required as long as the star's rotation period is longer than 7 days.

Rather than relying on these indirect measures of activity, we can also directly measure the amplitude of the star's rotation-modulated flux variations using visible-light ground-based observations. We obtained observations of GJ 436 in Strömgren *b*
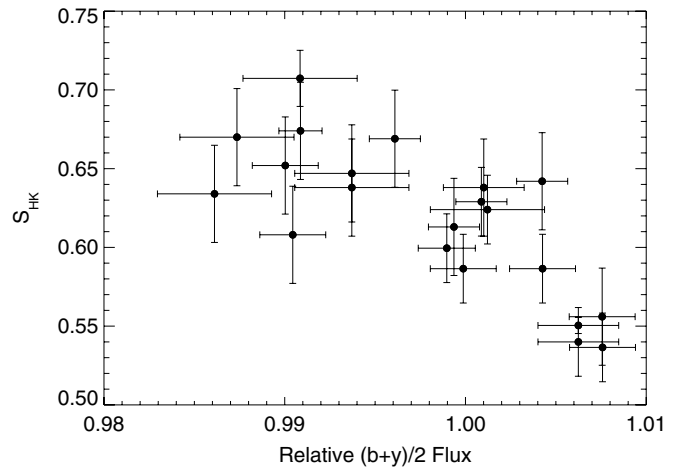

Figure 12. Averaged Strömgren *b* band fluxes vs. S_{HK} for GJ 436 over a period of seven years. Because the sampling for the S_{HK} measurements is much lower than for the flux measurements in most seasons, for the purposes of this plot the *b* fluxes are defined as the average of all *b* band measurements taken within one day of the S_{HK} measurement. We also average S_{HK} values taken on the same night, scaling the flux and S_{HK} error bars from Figure 11 by the square root of the number of measurements in each bin. We find that using *b* photometry instead of the averaged *b+y* fluxes results in increased scatter but also strengthens the observed correlation.

and *y* filters over a span of approximately six months surrounding our 2009 *Spitzer* transit and secondary eclipse observations from an ongoing monitoring program carried out with the T12 0.8 m Automatic Photoelectric Telescope (APT) at Fairborn Observatory in southern Arizona (Henry 1999; Eaton et al. 2003; Henry & Winn 2008). In these observations, the telescope nodded between GJ 436 and three comparison stars of comparable or greater brightness, which were then used to correct for the effects of variable seeing and airmass. We find that during the period between UT 2009 January 9 and February 4, when a majority of our transit data were obtained, the star varied in flux by less than a few mmag in visible light (Figure 11). We carry out a similar check for variability in the infrared using the fifteen 8 μm flux estimates listed in Table 1, which we find have a standard deviation of 0.07%. Both of these measurements indicate that the star is very nearly constant in flux in both visible and infrared light, and we can therefore rule out non-occulted spots as the cause of the observed transit depth variations.

We also use these same data to search for periodicities corresponding to GJ 436b's rotation period. If we fit the combined *b* and *y* band fluxes with a sine function plus a quadratic function of time as shown in Figure 11, we find a best-fit period of 56.5 days. We calculate a Lomb–Scargle periodogram (Lomb 1976; Scargle 1982) for these data and find that this period has a false alarm probability of only 2%, which we determine using a bootstrap Monte Carlo analysis. We find a nearly identical best-fit period of 56.6 days in the S_{HK} values measured with Keck HIRES during this epoch (Isaacson & Fischer 2010), but with a much higher false alarm probability of approximately 20%. We also examine the correlation between the measured *b* fluxes and S_{HK} values over the six-year period in which both were available (Figure 12) and find that these parameters are negatively correlated. Taken together, these data indicate that the small observed variations in GJ 436's visible-light fluxes are likely connected with the presence of regions of increased magnetic activity on the visible face of the star.

Although such low-amplitude flux variations generally indicate that a star has relatively few spots, there are two important

Table 7
A Comparison of the In-transit versus Out-of-transit Standard Deviations

UT Date	λ (μm)	N_{in}^a	N_{out}^a	σ_{rel}	$P(\sigma_{rel})^b$
		Unbinned data			
UT 2007 Jun 29	8.0	7,924	17,956	−1.3%	0.92
UT 2008 Jul 14	8.0	7,895	25,012	+1.4%	0.059
UT 2009 Jan 9	3.6	25,482	81,848	+0.2%	0.36
UT 2009 Jan 17	4.5	25,954	82,212	−0.3%	0.70
UT 2009 Jan 25	8.0	7,910	25,334	+0.1%	0.44
UT 2009 Jan 28	3.6	25,536	82,238	+1.4%	0.0023
UT 2009 Jan 30	4.5	25,955	62,334	+1.1%	0.018
UT 2009 Feb 2	8.0	7,890	25,318	+0.1%	0.45
		Binned data			
UT 2007 Jun 29	8.0	126	287	−13.6%	0.97
UT 2008 Jul 14	8.0	126	399	−4.9%	0.75
UT 2009 Jan 9	3.6	411	1311	+3.3%	0.21
UT 2009 Jan 17	4.5	412	1310	−3.4%	0.80
UT 2009 Jan 25	8.0	126	403	+9.7%	0.093
UT 2009 Jan 28	3.6	411	1311	+37.5%	1×10^{-6}
UT 2009 Jan 30	4.5	412	989	−1.4%	0.63
UT 2009 Feb 2	8.0	126	403	+2.9%	0.34

Notes.
[a] Number of in-transit and out-of-transit points.
[b] Probability that the standard deviation of the in-transit data would be greater than the standard deviation of the out-of-transit data by an amount σ_{rel} if both data sets are drawn from the same underlying Gaussian distribution.

exceptions. First, if the spots are uniformly distributed in longitude, it is theoretically possible to have a star with significant spot coverage and an effectively constant flux. It would not be surprising if the occurrence rate and distribution of spots was different for M stars than for G or K stars, but in GJ 436b's case the lack of any flux variations larger than a few mmag would seem to place a strong limit on the allowed spot distributions. We can quantify this limit if we assume that the deviation of approximately 0.08% in the first part of the 3.6 μm transit light curve from UT 2009 January 28 shown in Figure 2 is due to the occultation of a bright region on the star. This region must have a surface intensity that is 12% brighter than the rest of the star in order to produce the observed deviation. If we compare PHOENIX models with varying effective temperatures integrated over this band, we find that the star's temperature must increase by approximately 200 K in the affected region in order to match this surface intensity. We know that the total rotational modulation in the star's visible-light flux must remain below 0.1%, and we estimate that an increase of 12% in the 3.6 μm surface intensity should produce an increase of approximately 65% in the Strömgren $(b+y)/2$ band. In this case, the fractional area covered by active regions on the star must vary by less than 0.15% from the most active to the least active hemisphere. Of course, it is possible that the stellar atmosphere models do not provide an accurate match for the spectra of these active regions; if we instead use the measured 3.6 μm flux contrast of 12%, we find a more conservative limit of 1% on variations in the area affected by stellar activity.

A second, more plausible scenario involves tilting the rotation axis of star so that we are viewing it closer to pole-on, which would effectively suppress the amplitude of rotational flux variations regardless of spot coverage. If we assume that the star's spin axis is randomly oriented with respect to our line of sight, the probability that it will fall within 45° of a pole-on view is 30%. In this scenario, the star could be highly spotted, allowing for frequent occultations of spots by the planet, while still displaying a small rotational flux modulation. This scenario would require the planet's orbit to be misaligned with respect to the star's rotation axis, but such misalignments are commonly seen in other transiting planet systems (Winn et al. 2010a). Although the Rossiter–McLaughlin effect has never been successfully measured for GJ 436b, Winn et al. (2010b) find that the Neptune-mass planet HAT-P-11b, which is perhaps the best analogue to the GJ 436 system, has a sky-projected obliquity of $103°^{+26°}_{-10°}$ indicating that this system is significantly misaligned. If most close-in planets start out misaligned and are then gradually brought into alignment through tidal interactions with their host star as proposed by Winn et al. (2010a), the fact that HAT-P-11b still maintains both a non-zero orbital eccentricity and a significant misalignment would seem to suggest that the same could also be true for GJ 436b.

If we proceed with the hypothesis that GJ 436 is both spotty and tilted with respect to our line of sight, we can then search for evidence of occulted spots in the light curves with discrepant transit depths. We first compare the relative standard deviations of the in-transit (σ_{in}) and out-of-transit (σ_{in}) residuals plotted in Figure 2:

$$\sigma_{rel} = \frac{\sigma_{in} - \sigma_{out}}{\sigma_{out}}. \qquad (5)$$

We list the measured values of σ_{rel} for all eight transit observations in Table 7. Both the 3.6 μm transit on UT 2009 January 28 and the 4.5 μm transit on UT 2009 January 30

appear to have inflated values of σ_{rel}, as would be expected if the planet occulted active regions on the star during these visits. We can quantify the statistical significance of the measured σ_{rel} values if we assume that both the in-transit and out-of-transit points are drawn from the same underlying Gaussian distribution and then ask how many times in a sample of 100,000 random trials we measure a value of σ_{rel} greater than or equal to the value calculated directly from our observations. In each trial, we generate two synthetic data sets, each with the appropriate length corresponding to either the in-transit or out-of-transit measurements, and then calculate the standard deviation of each distribution and the corresponding value of σ_{rel}. In the 3.6 μm transit observation on UT 2009 January 9, there are 81,848 out-of-transit flux measurements and 25,482 in-transit flux measurements, and we find that over 100,000 trials, we obtain a value of σ_{rel} greater than or equal to the measured value of 0.2% approximately 36% of the time. Repeating the same calculation for the 3.6 μm transit observed on UT 2009 January 28, which has 82,238 out-of-transit points and 25,530 in-transit points, we obtain σ_{rel} greater than or equal to the measured value of 1.4% only 0.23% of the time. We list the corresponding probabilities for all eight transits in Table 7.

We also repeat this same test with data that have been binned in sets of 64 images, corresponding to 10 s bins at 3.6 and 4.5 μm and 30 s bins at 8 μm. This allows us to evaluate the relative contribution that correlated noise makes to the in-transit and out-of-transit variances, as the photon noise should be reduced by a factor of eight in these bins (also see Figure 3). In this case, we carry out 1,000,000 random trials for each visit, as each simulated data set is much smaller and the computations are correspondingly fast. We find that for the binned January 9 light curve there are 1311 points out of eclipse and 411 points in eclipse. In this case, σ_{rel} is 3.3%, and we obtain values greater than or equal to this number in 21% of our random trials. Repeating this calculation for the UT 2009 January 28 visit, we find that the measured value of σ_{rel} is 37% (i.e., a standard
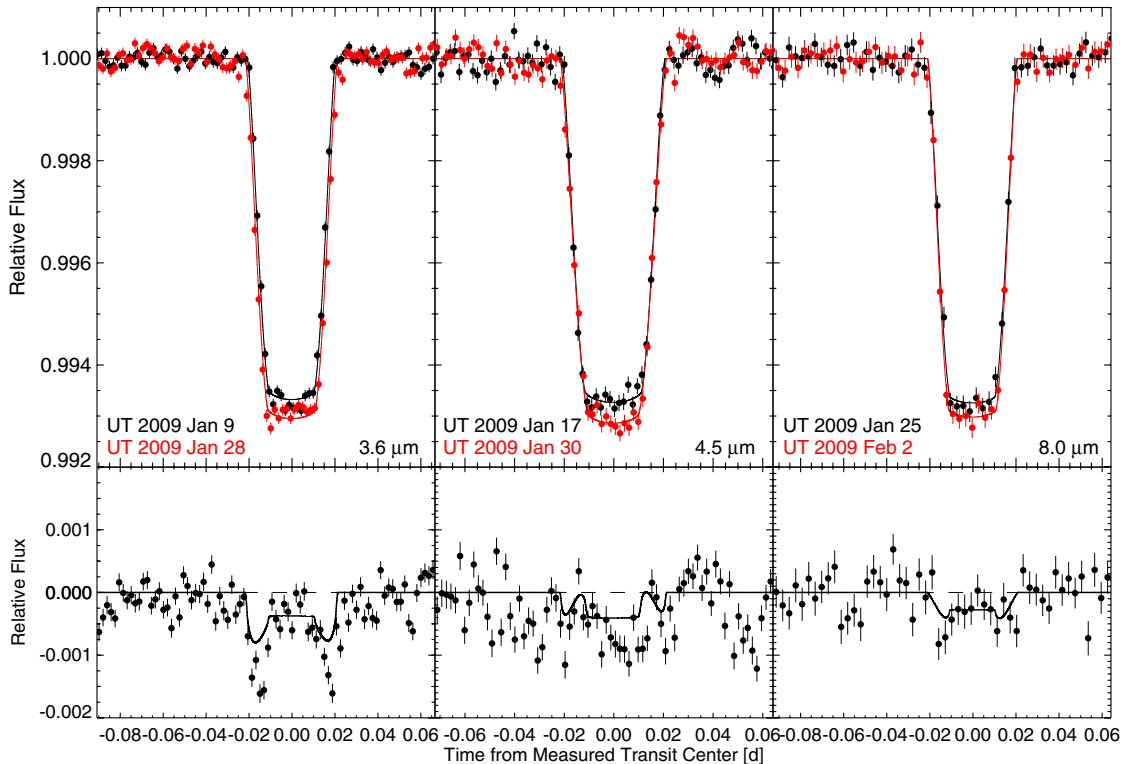
Figure 13. This plot shows the six transits observed in 2009 January/February. The left panel shows both 3.6 μm transits, the middle panel shows both 4.5 μm transits, and the right panel shows both 8.0 μm transits. The upper part of each panel overplots the normalized photometry for each visit (filled circles), with the first visit in black and the second visit in red, along with the best-fit transit light curves (solid lines) where the orbital inclination, $a/R_\star$, and transit time have been allowed to vary freely for each individual transit. The lower panel takes the difference between the two light curves (black filled circles) and compares it to the difference between the best-fit transit solutions (solid black line). Note that even when all transit parameters are allowed to vary freely, it is not possible to reproduce the sharp features visible during ingress and egress in the lower left panel.

deviation that is 37% higher in eclipse than it is out of eclipse), with 1311 points out of eclipse and 411 points in eclipse. In our simulations, assuming a single Gaussian probability distribution for both segments, this level of disagreement occurred only once in 10^6 trials. We find that in all other visits, including the 4.5 μm transit observed on UT 2009 January 30, the binned data in and out of eclipse are consistent with a single distribution.

One consequence of a misalignment between the star's rotation axis and the planet's orbit is that the planet will not necessarily occult the same spot on successive transits, as would be expected for a well-aligned system; we therefore consider each transit individually. Our analysis above indicates that the 3.6 μm transit on UT 2009 January 28 displays a statistically significant increase in the standard deviation of the in-transit data that is dominated by contributions from correlated noise on timescales greater than 30 s, as would be expected if the planet occulted an active region on the surface of the star. Although the 4.5 μm transit from UT 2009 January 30 does not appear to display a similar increase, our imperfect correction for the intrapixel sensitivity variations in this visit means that we are less sensitive to variations in σ_{rel}. We argue that even if the star's rotation axis and the planet's orbit are misaligned, it is still likely that the planet would occult the same active region during both the UT 2009 January 28 and January 30 visits, as the interval between these visits is much shorter than the star's approximately 50 day rotation period. As we discuss later in this section, the fact that both visits display increased transit depths and shifted transit times provides additional support for this hypothesis.

We also consider the possibility that the increased scatter in the in-transit residuals might be due to a change in the transit parameters, including the planet's radius, orbital inclination, transit time, or $a/R_\star$, from one visit to the next. We test this hypothesis by taking the difference of the first and second visits in each bandpass from 2009 and comparing the shape of the residual light curve to the differences we would expect due to changes in these parameters, which should be distinct from the deviations created by occulted star spots (Figure 13). Because we are directly differencing the two light curves, our results are independent of any assumptions about the shape of the transit light curve or the stellar limb darkening. We inspect the deviations in the residuals plotted in Figure 13 and conclude that they do not appear to be well matched by changes in the best-fit transit parameters, leaving occultations of active regions on the surface of the star as the most likely hypothesis.

If the planet occults a spot it can also cause a shift in the best-fit transit times, particularly when the spot is near the edge of the star and is occulted during ingress or egress. Indeed, we see that the UT 2009 January 28 3.6 μm appears to occur 31.4 $\pm$ 9.5 s early, while the 4.5 μm January 30 visit occurs 34.4 $\pm$ 9.4 s late (see Figure 8) in the fits where we fix $a/R_\star$ and i to a single common value. As a test we repeated our fit to the 3.6 μm transit excluding the first 1/3 of the transit light curve and found that the best-fit transit time shifted forward by approximately 30 s. We would also expect that transits observed in visible light, where the contrast between the spots and the star is more pronounced, would show proportionally larger timing deviations when the planet crosses a spot. As noted in Section 3.1, the scatter in the measured visible-light transit times is inconsistent with a

constant period, and the amplitude of the visible-light deviations is on average larger than the deviations in the infrared. We should also see this same wavelength dependence in the measured transit depths in Figure 10, and indeed we find that the 3.6 μm transit depth changes by 7.8%, the 4.5 μm transit depth changes by 5.3%, and the 8.0 μm transit depth changes by 4.9% during the period between UT 2009 January 9 and February 2. Lastly, we can examine the visible-light flux measurements for GJ 436 in Figure 11 and see that these two transits were obtained near a minimum in the star's flux, consistent with a relative increase in the fractional spot coverage as compared to earlier epochs. The measured values for S_{HK}, a common activity indicator, appear to be anti-correlated with the observed flux variations and reach a local maximum near this point.

4.2. Atmospheric Transmission Spectrum

In principle, the broadband transmission photometry of GJ 436b allows us to constrain the chemical composition and temperature structure near the limb of the planetary atmosphere (e.g., Madhusudhan & Seager 2009). However, time variability in either the properties of the star or of the planet poses a significant challenge to an analysis in which we are comparing transit observations at different wavelengths obtained days or weeks apart. As discussed in Section 4.1.1, we consider it unlikely that the discrepancies in the measured transit depths are due to changes in the properties of the planet, but instead conclude in Section 4.1.3 that the occultations of regions of non-uniform brightness in a subset of the transits appear to be responsible for the observed depth variations.

If we set aside those transits which we believe to be most strongly affected by stellar activity, including the UT 2009 January 28 and 30 visits, we may attempt to estimate the shape of the planet's transmission spectrum using the remaining transits. Although the evidence for spots in the final 8.0 μm transit on UT 2009 February 2 is somewhat weaker, we choose to exclude it on the grounds that it displays some of the same behaviors (increased depth, larger than usual timing offset) as the more strongly affected 3.6 and 4.5 μm transits immediately preceding it. If we then average the remaining three 8.0 μm depths, we find depths of 0.6694% ± 0.0061%, 0.6865% ± 0.0078%, and 0.6831% ± 0.0052% at 3.6, 4.5, and 8.0 μm, respectively. These three values are consistent with the near-IR transit depth from Pont et al. (2008b) of 0.6906% ± 0.0083% (1.1–1.9 μm), as well as the best-fit visible-light transit depth from Ballard et al. (2010b), 0.663% ± 0.014% (0.35–1.0 μm). Ground-based data provide additional constraints in the near-IR, including an H-band transit depth of 0.707% ± 0.019% from Alonso et al. (2008) and a Ks transit depth of 0.64% ± 0.03% from Cáceres et al. (2009),[14] both from individual transit observations.

We fit these data using the retrieval technique described in Madhusudhan & Seager (2009), which explores the parameter space of a one-dimensional, hydrogen-rich model atmosphere. We compute line-by-line radiative transfer with the assumption of hydrostatic equilibrium and use parametric prescriptions for the relative abundances of H_2O, CH_4, CO, and CO_2. We also include other dominant visible-light and infrared opacity sources, including Na, K, H_2–H_2 collision-induced absorption, and Rayleigh scattering. Our molecular line data are from

---

[14] The best-fit planet–star radius ratio reported by these authors is inconsistent with their best-fit depth. We re-fit their data with an equivalent model and conclude that this discrepancy is most likely the result of a mistake in the reported value for the radius ratio, as our best-fit depth is a good match for the value stated in the paper.
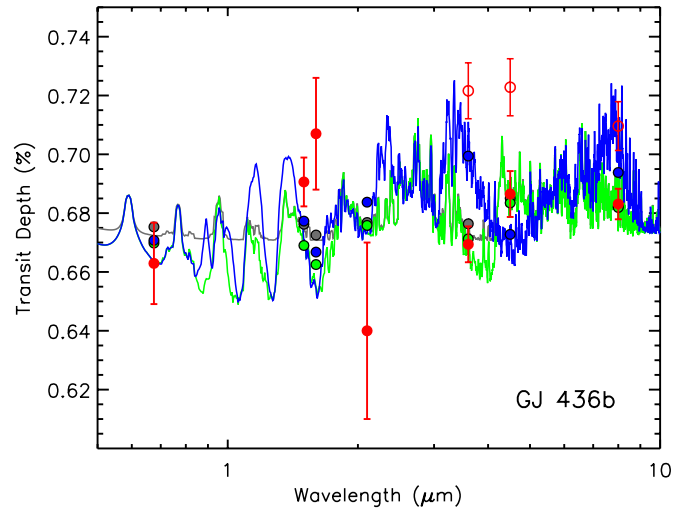

Figure 14. Comparison between measured transit depths (red circles) and model transmission spectra, where the transit depth is defined as the square of the best-fit planet–star radius ratio in each band. We include previously published visible and near-IR transit depths from (in order of increasing wavelength) Ballard et al. (2010b), Pont et al. (2008b), Alonso et al. (2008), and Cáceres et al. (2009) along with the *Spitzer* transit depths at 3.6, 4.5, and 8.0 μm, respectively. Open circles indicate *Spitzer* observations in which the planet appears to transit regions of non-uniform brightness on the star, as discussed in Section 4.1.3. Model transmission spectra include an atmosphere with reduced methane and enhanced CO abundances (Stevenson et al. 2010) in green, a methane-rich model similar to that described in Beaulieu et al. (2011) in blue, and a methane-poor model with an opaque cloud deck at pressures below 50 mbar in gray, which provides a better match to the visible-light transit depth from Ballard et al. (2010b). All models are calculated using the methods described in Madhusudhan & Seager (2009), where we fix the dayside P–T profiles to the nominal best-fit profile from Stevenson et al. (2010). We find that allowing the P–T profile to vary freely in our fits has a negligible effect on the agreement between the data and the green best-fit model. Colored green, blue, and gray circles indicate the predicted values for these models integrated over the bandpasses of the observations.

Rothman et al. (2005), Freedman et al. (2008), R. S. Freedman (2009, private communication), Karkoschka & Tomasko (2010), and E. Karkoschka (2011, private communication). The H_2–H_2 opacities are from Borysow et al. (1997) and Borysow (2002). We fix the pressure–temperature (P–T) profile to the best-fit dayside profile from Stevenson et al. (2010) and Madhusudhan & Seager (2011); it is possible to obtain a marginally improved fit to these data if we allow the P–T profile to vary freely in the fit, but the differences are not significant. We find that the observations can be explained to within the 1σ uncertainties by a methane-poor model (green line in Figure 14) that contains mixing ratios of $H_2O = 1.0 \times 10^{-3}$, $CO = 1.0 \times 10^{-3}$, and $CH_4 = 1.0 \times 10^{-6}$; the data used in this fit appear to be inconsistent with methane abundances $\geqslant 10^{-5}$. This model also includes $CO_2 = 1.0 \times 10^{-5}$, but the concentration of this molecule is less well constrained, as it is degenerate with the CO abundance in the 4.5 μm band. We do not expect strong absorption due to atomic Na and K in this temperature regime (Sharp & Burrows 2007), and we therefore adopt Na and K mixing ratios of $0.1\times$ solar abundances. If we compare the visible-light transit depth of 0.650% from this model to the value reported by Ballard et al. (2010b), we find that it is consistent at the 0.5σ level. Model transmission spectra for GJ 436b from Shabram et al. (2011), such as the rescaled model including higher-order hydrocarbons (model "g" in Shabram et al.), also provide a reasonably good match to these data.

We can reduce the disagreement between the measured transit depths and the green model in the 1–2 μm wavelength range by
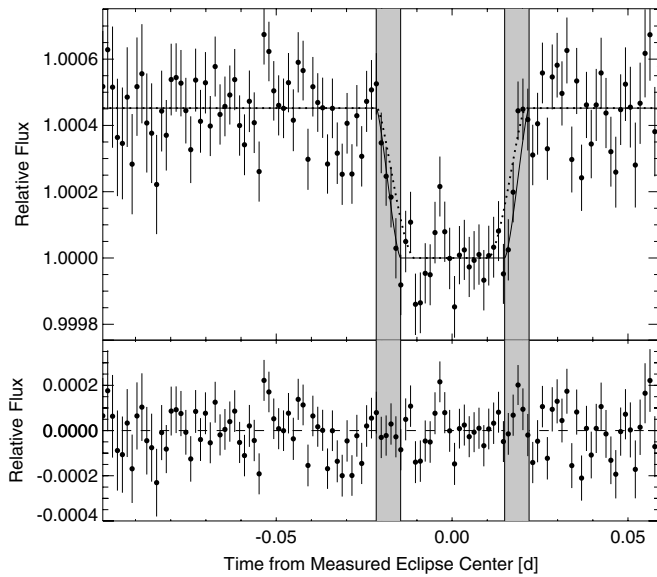
Figure 15. Photometry for eleven $8\,\mu$m secondary eclipses (filled circles), with detector effects removed. Individual visits have been aligned using the best-fit transit ephemeris and assuming a constant offset for the secondary eclipse. The best-fit secondary eclipse light curve is overplotted (solid line), and residuals from this curve are shown in the lower panel. The period spanned by ingress and egress is denoted as a gray shaded region; we find no significant deviations from a model in which the planet has a uniform surface brightness. The dotted line shows the best-fit transit light curve, rescaled to match the depth of the secondary eclipse shown here. The longer ingress and egress for the transit are due to the increased planet–star distance and correspondingly higher impact parameter during this event, which occurs close to apastron.
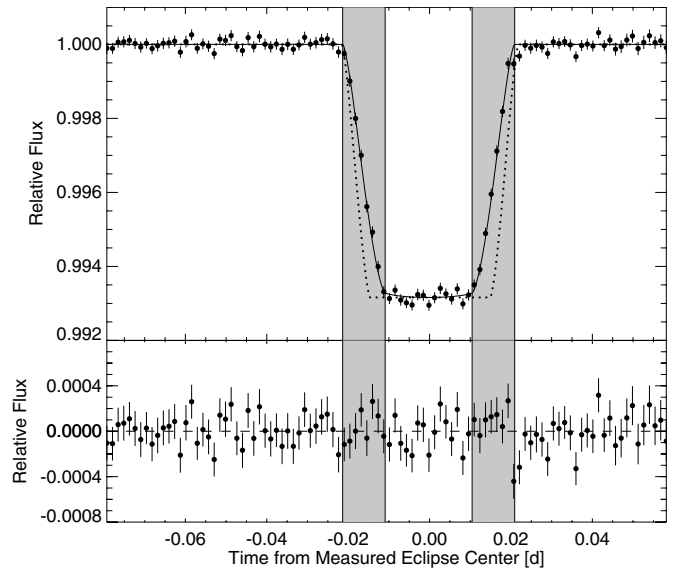
Figure 16. Photometry for four $8\,\mu$m transits (filled circles), with detector effects removed and light curves aligned using the best-fit transit ephemeris. The best-fit transit light curve is overplotted (solid line), and residuals from this curve are shown in the lower panel. The period spanned by ingress and egress is denoted as a gray shaded region; we find no significant deviations from the expected spherical planet model, as might be expected if the planet was significantly oblate. The dotted line shows the best-fit secondary eclipse light curve, rescaled to match the depth of the transit light curve where the limb darkening is set to zero. The shorter ingress and egress for the secondary eclipse are due to the reduced planet–star distance and correspondingly lower impact parameter during this event, which occurs shortly before periastron.

introducing an opaque cloud layer at 50 mbar (gray model in Figure 14). However, such a cloud layer would be inconsistent with the dayside emission spectrum measured by Stevenson et al. (2010) unless it was optically thin in the center of the dayside hemisphere or only intermittently present as discussed in Section 4.1.1. We also note that occultations of spots and other features on the star will have a stronger effect on the measured transit depth at shorter wavelengths, and it is therefore possible that these measurements (several of which were derived from individual transit observations) are unreliable for our purposes here.

Returning to the *Spitzer* data, we find that our conclusions about the atmospheric composition are strongly dependent on our choice of which transit depths to include in our analysis. We illustrate this with a blue model in Figure 14, which contains H_2O and CH_4 mixing ratios of 5.0×10^{-4} each and no CO or CO_2, and is comparable to the model presented in Beaulieu et al. (2011). Beaulieu et al. (2011) excluded the shallower $3.6\,\mu$m transit on UT 2009 January 9 and kept the deeper $3.6\,\mu$m UT 2009 January 28 and $8.0\,\mu$m UT 2009 February 2 visits in their analysis, and as a result they concluded that the planet's transmission spectrum contained strong methane features, as illustrated by this blue model. They argue that the correction for the intrapixel effect is degenerate with the transit depth for the UT 2009 January 9 visit and that this visit is therefore unreliable, but we find that there is good overlap between the x and y positions spanned by the in-transit and out-of-transit data. We obtain transit depths that are consistent at the 0.1σ level when we fit for our intrapixel sensitivity correction using either the entire light curve or the out-of-transit data alone. Although our 3.6 and $8.0\,\mu$m transit depths are in good agreement with the values obtained by Beaulieu et al., our best-fit transit depth for the $4.5\,\mu$m UT 2009 January 17 is 2.5σ larger. We note

that Beaulieu et al. allow $a/R_\star$ and b to vary individually for each transit and that their values for these parameters from the January 17 transit fit are outliers when compared to other visits; we conclude that this is likely the cause of their shallower best-fit radius ratio. Despite this disagreement, we find that if we include the same transits as Beaulieu et al. in our analysis, we also produce a transmission spectrum that is consistent with strong methane absorption.

If, as we propose, occulted regions of non-uniform brightness on the surface of the star are responsible for the discrepancies in the 3.6 and $4.5\,\mu$m transit depths, it will be difficult to provide a definitive characterization of GJ 436b's transmission spectrum with broadband *Spitzer* photometry. Our analysis suggests that the atmosphere of GJ 436b is likely underabundant in methane and overabundant in CO, consistent with the conclusions of Stevenson et al. (2010) and Madhusudhan & Seager (2011), but in order to reach these conclusions we have assumed that we have correctly identified and excluded all transits in which the planet occults active regions on the star. However, if the fractional spot coverage on the star is sufficiently high, it is possible that *all* transits are affected by these regions, in which case we cannot draw any robust conclusions about the shape of the planet's transmission spectrum.

4.3. Dayside Emission Spectrum and Limits on Variability

We can use the 11 secondary eclipse depths listed in Table 3 to study the properties of the planet's dayside atmosphere. We take the error-weighted average of the eclipse depths and find a combined value of $0.0452\% \pm 0.0027\%$, consistent with the value of $0.054\% \pm 0.008\%$ reported by Stevenson et al. (2010). Next, we construct a combined light curve incorporating all 11 secondary eclipse observations, shown in Figure 15. Figure 16 shows the equivalent combined $8\,\mu$m transit light curve for

comparison. As a check we fit these combined data with a secondary eclipse light curve and find that the best-fit eclipse depth agrees exactly with this error-weighted average from the individual eclipse fits. Because the strongest constraints on the relative abundances of methane and CO come from the 3.6 and 4.5 μm eclipse measurements, we do not expect the reduced 8 μm error bar to affect the conclusions reached by Stevenson et al. regarding these molecules. If we compare our results to the two models plotted in Figure 2 of Stevenson et al., we find that the revised 8 μm eclipse depth is best described by a cooler model with an effective blackbody temperature of 790 K (defined as the temperature needed to match the total integrated flux at all wavelengths) and a modestly enhanced (30 $\times$ higher) water abundance, rather than the hotter 860 K model with weaker water absorption. We also calculate a revised brightness temperature for the planet in the 8 μm band, defined as the temperature required to match the observed planet–star flux ratio in this bandpass assuming that the planet radiates as a blackbody. We use the parameters in Table 5 and assume a Phoenix atmosphere model with an effective temperature of 3585 K and log(g) equal to 4.843 (Torres 2007) for the star, and find that the planet has a best-fit brightness temperature of 740 $\pm$ 16 K.

Returning to Figure 15, we examine the residuals from our best-fit eclipse solution to search for evidence of deviations during ingress and egress caused by a non-uniform dayside surface brightness (Williams et al. 2006; Rauscher et al. 2007). The primary effect of a non-uniform brightness distribution is to shift the best-fit eclipse time (e.g., Agol et al. 2010), but in this case uncertainties in estimates for GJ 436b's orbital eccentricity and longitude of periastron prevent us from detecting the small (<1 minute) timing offsets expected from this effect. This timing offset will also display a small wavelength dependence, due to variations in the brightness distribution as seen in different bandpasses, but this signal is likely to be too weak to detect by comparing to the existing 3.6 and 5.8 μm eclipse observations from Stevenson et al. (2010). Instead, we seek to determine if the shape of the 8 μm eclipse ingress and egress can be used to constrain the planet's dayside brightness distribution. We compare the eclipse light curves for a uniform surface brightness disk to that of a local equilibrium model (i.e., one with the radiative time set to zero so that each region of the planet is at its local equilibrium temperature; Hansen 2008; Burrows et al. 2008) and find that the peak-to-trough residuals between these light curves are only 0.002%, if the eclipse depth is a free parameter. This is approximately a factor of 10 smaller than our measurement errors, as demonstrated by the binned residuals in Figure 15. As we increase the amount of energy advected to the planet's nightside using the models described in Cowan & Agol (2011), the location of the hot spot on the planet's dayside shifts away from the substellar point and the overall temperature contrast decreases. Because we are not sensitive to the timing offset caused by the shifted hot spot, the only effect of this increased advection is to homogenize the planet's temperatures, producing light curves increasingly similar to the uniform disk light curves.

4.3.1. A Variability Study for GJ 436b

Tidal dissipation is expected to have driven GJ 436b into a pseudosynchronous rotation state in which the planet's spin frequency is nearly commensurate with the planet's instantaneous orbital frequency at periastron. There are several competing theories of the pseudosynchronization process (see, e.g., Ivanov &
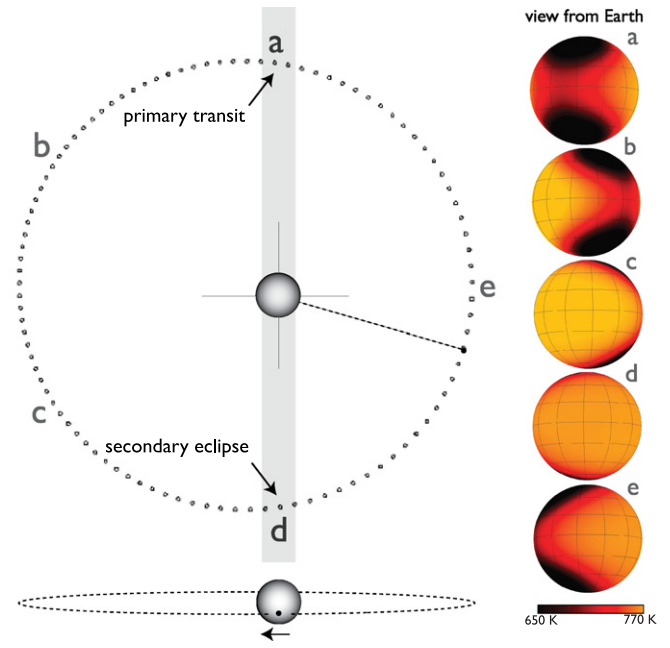

Figure 17. Left panel shows an orbital diagram for the GJ 436 system. Distances and radii are drawn to scale, and the location of periastron is marked by a dotted line. Gray shaded regions indicate the locations of the planet during the transit and secondary eclipse, where the viewer is assumed to be at the top of the plot. The right-hand panel shows five snapshots from a general circulation model for this planet as seen at different orbital phases by an observer on the Earth.

(A color version of this figure is available in the online journal.)

Papaloizou 2007). We adopt the expression given by Hut (1981):

$$\frac{\Omega_{\rm spin}}{\Omega_{\rm orbit}} = \frac{1 + \frac{15}{2}e^2 + \frac{45}{8}e^4 + \frac{5}{16}e^6}{\left(1 + 3e^2 + \frac{3}{8}e^4\right)(1 - e^2)^{3/2}}. \quad (6)$$

For GJ 436b, this relation gives $P_{\rm spin} = 2.32$ days, which yields a 19 day synodic period for the star as viewed from a fixed longitude on the planet. GJ 436b also experiences an 83% increase in incident flux during the 1.3 day interval between apastron and periastron.

We have computed simple hydrodynamical models to assess whether the asynchronous rotation and time-varying insolation are likely to generate atmospheric flows that are sufficiently chaotic to produce observable orbit-to-orbit variability in the secondary eclipse depths. Our two-dimensional hydrodynamical model contains three free parameters. The first, $p_{8\,\mu\rm m}$, is the atmospheric pressure at the 8 μm photosphere; the second, X, corresponds to the fraction of the incoming optical flux that is absorbed at or above the 8 μm photosphere; and the third, p_b, corresponds to the pressure at the base of our modeled layer. We adopt parameter values of $p_{8\,\mu\rm m} = 100$ mbar, $p_b = 4.0$ bar, and $X = 1.0$ for these models, which put our model's light curve in good agreement with GJ 436b's average 8 μm secondary eclipse depth. The full details of the computational scheme are the same as those adopted in Langton & Laughlin (2008), with updates as described in Laughlin et al. (2009). A model photometric light curve is then obtained by integrating at each time step over the planetary hemisphere visible from Earth, where we assume that each patch of the planet radiates with a blackbody spectrum corresponding to the local temperature.

The model is run for a large number of orbits, and a quasi-steady state surface flow emerges. The temperature structure of this flow as seen from an observer in the direction of Earth at five equally spaced intervals in the orbit is shown in Figure 17,
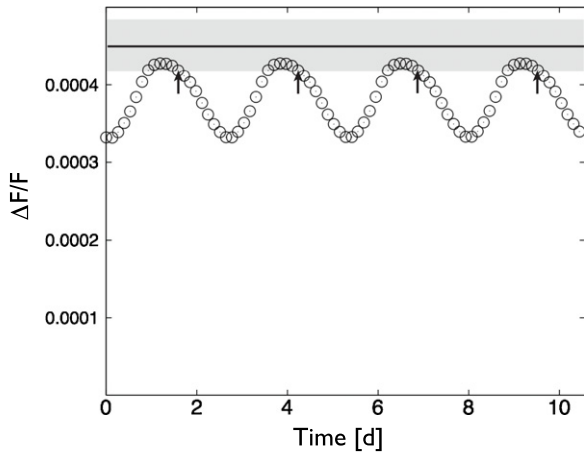
Figure 18. Predicted 8 μm emission for GJ 436b (open circles) as a function of time, from general circulation models described in Langton & Laughlin (2008). The periodic modulation in flux is primarily due to the changing orbital geometry as we watch the hotter dayside rotate in and out of view, with a secondary effect caused by the heating and cooling of the atmosphere as the planet moves from periastron to apastron and back. The predicted fluxes during the secondary eclipse, as indicated by the black arrows, are nearly constant in time. The horizontal black line and gray shaded region indicate the average secondary eclipse depth and corresponding 1σ uncertainty.

and the model light curve over these five orbits is shown in Figure 18. Over the course of a single orbit, the 8 μm planet-to-star flux ratio varies nearly sinusoidally from $\Delta F/F = 0.033\%$ to 0.043%. The model's flux at secondary eclipse agrees well with the observed value and varies by only 0.5% peak to peak from one orbit to the next. We note that more sophisticated three-dimensional general circulation models for GJ 436b from Lewis et al. (2010) also predict very low (1.3%–1.5%) levels of variability in the 8 μm band for a range of atmospheric metallicities.

Although these models indicate that GJ 436b's modest orbital eccentricity is likely not sufficient to induce significant variability, they also do not include many processes such as clouds, photochemistry, and small-scale turbulence that are known to contribute to temporal variability in planetary atmospheres. We therefore place empirical limits on GJ 436b's dayside variability using the eleven 8 μm secondary eclipse observations. We assume that the intrinsic dayside fluxes are drawn from either a Gaussian distribution with a standard deviation δ or from a boxcar distribution with a width equal to 2δ. In both cases, we set the mean of the distribution equal to the error-weighted mean of the measured secondary eclipse depths given in Table 5. We then conduct 10,000 random trials, where we draw 11 measurements from each distribution and calculate the reduced χ^2 of these values as compared to the measured secondary eclipse depths in Table 3. We then determine the fraction of the 10,000 random trials in which the reduced χ^2 is less than or equal to one, which should correspond to the probability that the underlying distribution is consistent with the measured eclipse depths. We repeat this calculation for a range of values for δ and plot the resulting probability distribution as a function of δ for both boxcar and Gaussian distributions. We find that for a boxcar distribution we can place [1σ, 2σ, 3σ] limits on the intrinsic variability of [29%, 42%, 58%], and for a Gaussian distribution our corresponding upper limits are [17%, 27%, 42%]. These limits are consistent with the predictions from general circulation models for this planet, but they are not low enough to provide meaningful constraints on these models.

5. CONCLUSIONS

In this paper, we present *Spitzer* observations of eight transits and eleven secondary eclipses of GJ 436b at 3.6, 4.5, and 8.0 μm, which allow us to derive improved values for the planet's orbital ephemeris, eccentricity, inclination, radius, and other system parameters. We discuss the effects that our assumptions about the longitude of periastron and stellar limb-darkening profiles have on our best-fit transit parameters and find that our best-fit parameters vary by 1σ or less in all cases. We find that all parameters are consistent with a constant value over the two-year period spanned by our observations, with the exception of the measured transit depths and times in the 3.6 and 4.5 μm bands. We find that the 3.6 μm radius ratio measured on UT 2009 January 28 is 4.7σ deeper than the value measured on UT 2009 January 9 in this same band, and the 4.5 μm radius ratio from UT 2009 January 30 is 2.9σ deeper than the value measured on UT 2009 January 17. The level of significance for these changing radius ratios remains high even after accounting for the effects of residual correlated noise in the data.

We also present an improved estimate for GJ 436b's 8 μm secondary eclipse depth, based on 11 eclipse observations in this bandpass. We find that the new depth is consistent with previous models described in Stevenson et al. (2010) and Madhusudhan & Seager (2011), although we prefer solutions with modestly lower effective temperatures (790 K instead of 860 K). We use the shape of the eclipse ingress and egress to search for the presence of a non-uniform temperature distribution in the planet's dayside atmosphere, but uncertainties in the predicted time of secondary eclipse ultimately limit our ability to place meaningful constraints on this quantity. Our eclipse depths in this band are consistent with a constant value, and we place a 1σ upper limit of 17% on variability in the planet's dayside atmosphere. This limit is in good agreement with the predictions of general circulation models for this planet, which are typically variable at the level of a few percent or less in this bandpass.

Although it is possible that such residual noise or a time-varying cloud layer at sub-mbar pressures could explain the apparent transit depth variations, the features observed in the transit light curves appear to be most consistent with the presence of occulted spots or other areas of non-uniform brightness on the surface of the star in the UT 2009 January 28 and 30 transits. We find that for the UT 2009 January 28 transit the in-transit data have a higher rms than the out-of-transit data, as would be expected for occulted spots; we would expect poorly corrected systematics to produce an equivalently large rms in both the in-transit and out-of-transit data, as the star spans the same region of the pixel in both segments. Although we are not as sensitive to such effects in the UT 2009 January 30 visit, which has higher levels of correlated noise due to an imperfect correction for intrapixel sensitivity variations, the short separation between these two observations relative to the star's approximately 50 day rotation period means that the planet is likely to have occulted the same feature in both visits. We also see statistically significant variations in the measured transit times, where the amplitude of the variations is typically smaller for infrared observations than for those obtained in visible light, also suggesting the presence of occulted spots. We note that the anomalously deep transits observed on UT 2009 January 28 and 30 also have best-fit transit times that are offset by 30 s (3.1σ–3.5σ significance) from the predicted values. The fact that the three deepest transits are all measured within the same five-day period is also consistent with a single epoch of increased stellar activity. We reconcile this conclusion with

the absence of any variations larger than a few mmag in the star's visible and infrared fluxes by proposing that the star's spin axis is likely inclined with respect to our line of sight, which has the effect of reducing the amplitude of any flux variations independent of spot coverage. If this is in fact the case, GJ 436b's orbit will be misaligned with respect to the star's spin axis.

If we examine the wavelength-dependent transit depths for the subset of visits that appear to be least affected by spots, we find that the resulting transmission spectrum is consistent with the same reduced methane and enhanced CO abundances used by Stevenson et al. (2010) to fit the planet's dayside emission spectrum. These same transit data are also consistent with models including an opaque cloud layer at a pressure of approximately 50 mbar or less in the planet's atmosphere, which reduces the amplitude of the absorption features in the model spectra. We find no convincing evidence for the strong methane absorption reported by Beaulieu et al. (2011), although we note that our conclusions vary significantly depending on which transits we include in our analysis. It is possible that all measured transit depths are affected to varying degrees by stellar activity, in which case it may not be feasible to characterize the planet's transmission spectrum using broadband photometry obtained over multiple epochs. Because active regions occulted by the planet display a characteristic wavelength dependence and also alter the local shape of the transit light curve, high signal-to-noise grism spectroscopy of the transit over multiple epochs would help to resolve this issue. Such observations would also provide an independent test of the reliability of the *Spitzer* transit data; if similar apparent depth variations were observed in other data sets, it would provide a strong argument against the hypothesis that the apparent depth variations in these data might be the result of poorly corrected instrument effects. Lastly, grism spectroscopy could also be used to search for time-varying clouds at sub-mbar pressures, which should produce a featureless transmission spectrum with a uniformly increased depth when present, as compared to the standard cloud-free transmission spectrum.

As indicated by its rotation rate and Ca II H&K emission, GJ 436 is an old and relatively quiet early M star. If the apparent transit depth variations we describe here are indeed due to the occultation of active regions on the star, as appears likely, we would expect similar features to occur frequently in the transit light curves of other planets orbiting M dwarfs at all activity levels. GJ 1214 is currently the only other M star known to host a transiting planet and has a similar 53 day rotation period and a modestly lower 3000 K effective temperature as compared to GJ 436 (Charbonneau et al. 2009; Berta et al. 2011). A majority of the published data on this system are in the visible and near-infrared wavelengths where star spots should be prominent, and several recent papers report the presence of occulted spots in a subset of transit observations (Berta et al. 2011; Carter et al. 2011; Kundurthy et al. 2011). Such spots might also account for the apparent disagreement in measurements of the planet's infrared transmission spectrum, which some authors find to be featureless (Bean et al. 2010; Désert et al. 2011b), while others detect absorption features (Croll et al. 2011). HD 189733b is currently the only other exoplanet with repeated *Spitzer* transit observations in the same band; although this planet orbits a relatively active K star (e.g., Knutson et al. 2010), it exhibits much smaller variations in the measured transit depths and times as compared to GJ 436b (Agol et al. 2010; Désert et al. 2011a). This is perhaps not surprising, as the relative fractional spot coverage, spot sizes, and spot temperatures may well be qualitatively different on K stars and M stars.

REFERENCES

Adams, E., Seager, S., & Elkins-Tanton, L. 2008, ApJ, 673, 1160
Agol, E., et al. 2010, ApJ, 721, 1861
Alonso, R., et al. 2008, A&A, 487, L5
Bakos, G. Á., et al. 2006, ApJ, 650, 1160
Ballard, S., et al. 2010a, ApJ, 716, 1047
Ballard, S., et al. 2010b, PASP, 122, 1341
Batygin, K., et al. 2009, ApJ, 699, 23
Bean, J. L., Kempton, E. M.-R., & Homeier, D. 2010, Nature, 468, 669
Bean, J. L., & Seifahrt, A. 2008, A&A, 487, L25
Beaulieu, J.-P., et al. 2011, ApJ, 731, 16
Berta, Z. K., et al. 2011, ApJ, submitted (arXiv:1012.0518)
Borysow, A. 2002, A&A, 390, 779
Borysow, A., Jorgensen, U. G., & Zheng, C. 1997, A&A, 324, 185
Burrows, A., Budaj, J., & Hubeny, I. 2008, ApJ, 678, 1436
Butler, R. P., et al. 2004, ApJ, 617, 580
Cáceres, C., et al. 2009, A&A, 507, 481
Carter, J. A., et al. 2011, ApJ, 730, 82
Charbonneau, D., et al. 2002, ApJ, 568, 377
Charbonneau, D., et al. 2005, ApJ, 626, 523
Charbonneau, D., et al. 2008, ApJ, 686, 1341
Charbonneau, D., et al. 2009, Nature, 462, 891
Claret, A. 2000, A&A, 363, 1081
Claret, A. 2008, A&A, 482, 259
Claret, A. 2009, A&A, 506, 1335
Claret, A., & Hauschildt, P. H. 2003, A&A, 412, 241
Coughlin, J. L., et al. 2008, ApJ, 689, L149
Cowan, N. B., & Agol, E. 2011, ApJ, 726, 82
Cowan, N. B., Agol, E., & Charbonneau, D. 2007, MNRAS, 379, 641
Croll, B., et al. 2011, ApJ, in press (arXiv:1104.1011)
Crossfield, I. J. M., et al. 2010, ApJ, 723, 1436
Cushing, M. C., Rayner, J. T., & Vacca, W. D. 2005, ApJ, 623, 1115
de Kort, J. J. M. A. 1954, Ricerche Astronomiche, Vol. 3 (Citta del Vaticano: Specola Vaticana), 109
Deming, D., Harrington, J., Seager, S., & Richardson, L. J. 2006, ApJ, 644, 560
Deming, D., Seager, S., & Richardson, L. J. 2005, Nature, 434, 740
Deming, D., et al. 2007, ApJ, 667, L199
Deming, D., et al. 2011, ApJ, 726, 95
Demory, B.-O., et al. 2007, A&A, 475, 1125
Désert, J.-M., et al. 2008, A&A, 492, 585
Désert, J.-M., et al. 2009, ApJ, 699, 478
Désert, J.-M., et al. 2011a, A&A, 526, A12
Désert, J.-M., et al. 2011b, ApJ, 731, L40
Eastman, J., Siverd, R., & Gaudi, B. S. 2010, PASP, 122, 894
Eaton, J. A., Henry, G. W., & Fekel, F. C. 2003, in The Future of Small Telescopes in the New Millennium, Volume II—The Telescopes We Use, ed. T. D. Oswalt (Dordrecht: Kluwer), 189
Fazio, G. G., et al. 2004, ApJS, 154, 10

Figueira, P., et al. 2009, A&A, 493, 671
Ford, E. 2005, ApJ, 129, 1706
Freedman, R. S., Marley, M. S., & Lodders, K. 2008, ApJS, 174, 504
Fressin, F., et al. 2010, ApJ, 711, 374
Gibson, N. P., Pont, F., & Aigrain, S. 2011, MNRAS, 411, 2199
Gillon, M., et al. 2007a, A&A, 471, L51
Gillon, M., et al. 2007b, A&A, 472, L13
Grillmair, C. J., et al. 2008, Nature, 456, 767
Guillot, T., & Havel, M. 2011, A&A, 527, A20
Hansen, B. M. S. 2008, ApJ, 179, 484
Harrington, J., et al. 2007, Nature, 447, 691
Hauschildt, P., Allard, F., Ferguson, J., Baron, E., & Alexander, D. R. 1999, ApJ, 525, 871
Henry, G. W. 1999, PASP, 111, 845
Henry, G. W., & Winn, J. N. 2008, AJ, 135, 68
Hut, P. 1981, A&A, 99, 126
Iro, N., & Deming, L. D. 2010, ApJ, 712, 218
Isaacson, H., & Fischer, D. 2010, ApJ, 725, 875
Ivanov, P. B., & Papaloizou, J. C. B. 2007, MNRAS, 376, 682
Jenkins, J. S., et al. 2009, ApJ, 704, 975
Karkoschka, E., & Tomasko, M. 2010, Icarus, 205, 674
Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, ApJ, 720, 1569
Knutson, H. A., et al. 2007, Nature, 447, 183
Knutson, H. A., et al. 2008, ApJ, 673, 526
Knutson, H. A., et al. 2009a, ApJ, 690, 822
Knutson, H. A., et al. 2009b, ApJ, 691, 866
Knutson, H. A., et al. 2009c, ApJ, 703, 769
Kundurthy, P., et al. 2011, ApJ, 731, 123
Kurucz, R. 1979, ApJS, 40, 1
Kurucz, R. 1994, Solar Abundance Model Atmospheres for 0, 1, 2, 4, and 8 km/s, CD-ROM No. 19 (Cambridge, MA: Smithsonian Astrophysical Observatory)
Kurucz, R. 1999, H$_2$O Linelist from Partridge & Schwenke (1997), part 2 of 2, CD-ROM No. 26 (Cambridge, MA: Smithsonian Astrophysical Observatory)
Kurucz, R. 2005, Mem. Soc. Astron. Ital. Suppl., 8, 14
Langton, J., & Laughlin, G. 2008, ApJ, 674, 1106
Laughlin, G., et al. 2009, Nature, 457, 562
Lewis, N. K., et al. 2010, ApJ, 720, 344
Linsky, J. L., et al. 2010, ApJ, 717, 1291
Loeb, A. 2005, ApJ, 623, L45
Lomb, N. R. 1976, Ap&SS, 39, 447
Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24
Madhusudhan, N., & Seager, S. 2011, ApJ, 729, 41
Madhusudhan, N., & Winn, J. N. 2009, ApJ, 693, 784
Mandel, K., & Agol, E. 2002, ApJ, 580, L171
Maness, H. L., et al. 2007, PASP, 119, 90
Markwardt, C. B. 2009, in ASP Conf. Ser. 411, Astronomical Data Analysis Software and Systems XVIII, ed. A. Bohlender, D. Durand, & P. Dowler (San Francisco, CA: ASP), 251
Morales-Calderon, M., et al. 2006, ApJ, 653, 1454
Nettelmann, N., Kramm, U., Redmer, R., & Neuhäuser, R. 2010, A&A, 523, A26
O'Donovan, F. T., et al. 2010, ApJ, 710, 1551
Orosz, J. A., & Hauschuldt, P. H. 2000, A&A, 364, 265
Pál, A., et al. 2010, MNRAS, 401, 2665
Partridge, H., & Schwenke, D. W. 1997, J. Chem. Phys., 106, 4618
Pont, F., et al. 2008a, MNRAS, 385, 109
Pont, F., et al. 2008b, MNRAS, 393, L6
Rabus, M., et al. 2009, A&A, 494, 391
Rauscher, E., et al. 2007, ApJ, 664, 1199
Reach, W. T., et al. 2005, PASP, 117, 978
Ribas, I., Font-Ribera, A., & Beaulieu, J.-P. 2008, ApJ, 677, L59
Rogers, L. A., & Seager, S. 2010, ApJ, 712, 974
Rothman, L. S., et al. 2005, J. Quant. Spectrosc. Radiat. Transfer, 96, 139
Scargle, J. D. 1982, ApJ, 263, 835
Shabram, M., Fortney, J. J., Green, T. P., & Freedman, R. S. 2011, ApJ, 727, 65
Sharp, C. M., & Burrows, A. 2007, ApJS, 168, 140
Shporer, A., et al. 2009, ApJ, 694, 1559
Spiegel, D. S., Burrows, A., Ibgui, L., Hubeny, I., & Milsom, J. A. 2010, ApJ, 709, 149
Sterne, T. E. 1940, Proc. Natl Acad. Sci., 26, 36
Stevenson, K. B., et al. 2010, Nature, 464, 1161
Swain, M. R., Vasisht, G., & Tinetti, G. 2008, Nature, 452, 329
Todorov, K., et al. 2010, ApJ, 708, 498
Torres, G. 2007, ApJ, 671, L65
Vidal-Madjar, A., et al. 2003, Nature, 422, 143
Werner, M. W., et al. 2004, ApJS, 154, 1
Williams, P. K., Charbonneau, D., Cooper, C. S., Showman, A. P., & Fortney, J. J. 2006, ApJ, 649, 1020
Winn, J. N., Fabrycky, D., Albrecht, S., & Johnson, J. A. 2010a, ApJ, 718, L145
Winn, J. N., et al. 2007, AJ, 134, 1707
Winn, J. N., et al. 2008, ApJ, 683, 1076
Winn, J. N., et al. 2010b, ApJ, 723, L223