3c 220.3: A Radio Galaxy Lensing a Submillimeter Galaxy

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

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1088/0004-637X/790/1/46</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:30212124">http://nrs.harvard.edu/urn-3:HUL.InstRepos:30212124</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
3C 220.3: A RADIO GALAXY LENSING A SUBMILLIMETER GALAXY

Martin Haas1, Christian Leipski2, Peter Barthel3, Belinda J. Wilkes4, Simona Vegetti5, R. Shane Busmann4, S. P. Willner1, Christian Westhus1, Matthew L. N. Ashby4, Rolf Chini6,7, David L. Clements8, Christopher D. Fassnacht9, Assaf HoreSh9, Ulrich Klaas2, Léon V. E. Koopmans9, Joanna Kuraszkiewicz4, David J. LagattaT10,11, Klaus Meisenheimer2, Daniel Stern12, and Dominika Wylezalek12,13

1 Astronomisches Institut, Ruhr Universität, Bochum, Germany; haas@astro.ruw.de
2 Max-Planck-Institut für Astronomie, Heidelberg, Germany
3 Kapteyn Astronomical Institute, University of Groningen, The Netherlands
4 Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
5 Max-Planck-Institut für Astrophysik, Garching, Germany
6 Universidad Católica del Norte, Antofagasta, Chile
7 Imperial College, London, UK
8 University of California, Davis, CA, USA
9 Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA, USA
10 Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn, Australia
11 ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), 44 Rosehill Street Redfern, NSW 2016, Australia
12 Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
13 European Southern Observatory, Garching, Germany

Received 2013 August 23; accepted 2014 June 6; published 2014 July 2

ABSTRACT

Herschel Space Observatory photometry and extensive multiwavelength follow-up have revealed that the powerful radio galaxy (PRG) 3C 220.3 at z = 0.685 acts as a gravitational lens for a background submillimeter galaxy (SMG) at z = 2.221. At an observed wavelength of 1 mm, the SMG is lensed into three distinct images. In the observed near infrared, these images are connected by an arc of ~1′.8 radius forming an Einstein half-ring centered near the radio galaxy. In visible light, only the arc is apparent. 3C 220.3 is the only known instance of strong galaxy-scale lensing by a PRG not located in a galaxy cluster and therefore it offers the potential to probe the dark matter content of the radio galaxy host. Lens modeling rejects a single lens, but two lenses centered on the radio galaxy host A and a companion B, separated by 1″.5, provide a fit consistent with all data and reveal faint candidates for the predicted fourth and fifth images. The model does not require an extended common dark matter halo, consistent with the absence of extended bright X-ray emission on our Chandra image. The projected dark matter fractions within the Einstein radii of A (1″.02) and B (0″.61) are about 0.4 ± 0.3 and 0.55 ± 0.3. The mass to i-band light ratios of A and B, M/Li ~ 8 ± 4 M⊙/L⊙1, appear comparable to those of radio-quiet lensing galaxies at the same redshift in the CfA-Arizona Space Telescope LEns Survey, Lenses Structure and Dynamics, and Strong Lenses in the Legacy Survey samples. The lensed SMG is extremely bright with observed f(250 μm) = 440 mJy owing to a magnification factor μ ~ 10. The SMG spectrum shows luminous, narrow C iv λ1549 Å emission, revealing that the SMG houses a hidden quasar in addition to a violent starburst. Multicolor image reconstruction of the SMG indicates a bipolar morphology of the emitted ultraviolet (UV) light suggestive of cones through which UV light escapes a dust-enshrouded nucleus.

Key words: dark matter – galaxies: individual (3C 220.3) – gravitational lensing: strong – radio continuum: galaxies – submillimeter: galaxies

Online-only material: color figures

1. INTRODUCTION

Powerful radio galaxies (PRGs, with L_{178 MHz} > 10^{27} W Hz^{-1} sr^{-1} and Fanaroff–Riley morphology type FR II) mark places of exceptional energy transformation. Giant radio lobes far outside the massive host galaxies are powered by relativistic jets launched from accreting supermassive black holes in the galaxy centers. PRGs are among the most massive galaxies at their epoch, as traced by the K–z brightness–redshift diagram (Lilly & Longair 1984). In the cold dark matter paradigm of cosmic structure formation (White & Rees 1978), PRGs might be expected to reside in the most massive dark matter halos and in rich galaxy clusters. For a few PRGs in rich, X-ray-luminous clusters, massive amounts of dark matter have been revealed by giant luminous arcs, e.g., 3C 220.1 at z = 0.61 (Dickinson 1993) or by the weak shear of background galaxies, e.g., 3C 324 at z = 1.2 (Smail & Dickinson 1995), both caused by cluster-scale gravitational lensing. However, for PRGs not in a cluster, no galaxy-scale gravitational lensing measurements of their dark matter content have been achieved to date.

This paper reports the first case of a lensing PRG not located in a galaxy cluster. As a bonus, the lensed source is a submillimeter galaxy (SMG). SMGs are dust-enshrouded star-forming galaxies at high redshift (z > 1) and are thought to be the progenitors of today’s massive elliptical galaxies. Lensed examples of SMGs are not rare (Negrello et al. 2010), but a full picture of the population has yet to emerge.

Our discovery of the lensing nature of 3C 220.3 is based on far infrared and submillimeter photometry with the Herschel Space Observatory. The source appeared abnormally bright, suggesting contamination of a background source in the Herschel beam (5″–35″). Reprocessed archival Hubble Space Telescope (HST) visible (702 nm) images revealed an arc surrounding two objects, indicating that gravitational lensing plays a role. To identify the radio source among the two lensing candidate galaxies, we obtained a deep 3.3 cm radio map with the Karl
G. Jansky Very Large Array (JVLA, 0.′4 beam width). We performed high resolution imaging at 1.0 mm with the Submillimeter Array (SMA), near-infrared (NIR, K′-band, 2.124 μm) adaptive optics (AO) imaging with the 10 m Keck II telescope, 3.6 and 4.5 μm imaging with the Spitzer Space Telescope, 360–880 nm spectroscopy with the 5 m Palomar telescope, 320–1000 nm spectroscopy with the 10 m Keck I telescope, and obtained a 10 ks Chandra X-ray image.

This paper reports results of the imaging and spectroscopy campaigns (Section 2) and of lens modeling (Section 3) to fit the resulting data. Section 4 treats the wavelength dependent morphology of the SMG. Section 5 compares the derived mass to light ratio (M/L) for the PRG to ratios for galaxies at similar redshift, and Section 6 summarizes the results. We adopt a standard ΛCDM cosmology (H₀ = 72 km s⁻¹ Mpc⁻¹, Ω₀ = 0.73, and Ωₚ = 0.27).

2. OBSERVATIONS AND DATA

2.1. Herschel Space Observatory

Photometric observations of 3C 220.3 were carried out with the Herschel Space Observatory (Pilbratt et al. 2010) using the 70, 100, and 160 μm channels of the instrument PACS (Poglitsch et al. 2010) on UT 2011 May 28 in scan-map mode with on-source times of ~80 s at 70 and 100 μm and ~160 s at 160 μm. SPIRE (Griffin et al. 2010) observations at 250, 350, and 500 μm were made on UT 2012 November 5 in small-map mode with ~111 s on-source time. Data reduction was performed within the Herschel Interactive Processing Environment (HIPE; Ott 2010) following standard procedures including source masking and high-pass filtering. Flux densities were measured for PACS with aperture photometry including corrections for aperture size. SPIRE photometry utilized the timeline source extractor implemented in HIPE. Because the source is bright in all filters, the photometric uncertainties are small, typically a few percent.

3C 220.3 was observed with the MIPS (Rieke et al. 2004) instrument on the Spitzer Space Telescope (Werner et al. 2004) with reported 70 and 160 μm flux densities (Cleary et al. 2007) significantly lower than those from Herschel. To understand this discrepancy, we retrieved newly calibrated MIPS maps from the Spitzer archive and performed photometry with aperture corrections from the MIPS Instrument Handbook. The new 70 and 160 μm Spitzer photometry yielded higher flux densities which now are consistent within 15% with the Herschel results. The Spitzer MIPS 24 μm photometry agrees to within 4% with 22 μm photometry from the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010).

For comparison with 3C 220.3, a sample of 12 PRGs was taken from the 3CR catalog. The sample (C. Westhus et al., in preparation) consists of galaxies with 0.6 < z < 0.8 and radio and mid-infrared properties similar to those of 3C 220.3. The MIPS 24 μm photometry of galaxies in the comparison sample is consistent with the WISE data.

2.2. Karl G. Jansky Very Large Array

The deepest radio image of 3C 220.3 to date, at λ = 6 cm by Mullin et al. (2006), shows the double-lobed morphology but not the unresolved core. With the aim of detecting the core, we observed 3C 220.3 with the Karl G. JVLA in its long-baseline A configuration on UT 2012 November 30 employing 25 antennas.

The bandwidth was 2048 MHz centered at 9 GHz (X band). Prime flux calibrators were 3C 48 and 3C 147, and the nearby radio source J1010+8250 served as phase calibrator. The total on-source integration time was 37 minutes.

The resulting image, which was made using CASA and employed natural weighting of the visibility data, has an angular resolution of ~0′.4. It shows the ~9′ double-lobed morphology reported previously (Mullin et al. 2006) and an unresolved core with a flux density of 0.8 mJy (Figure 1).

2.3. Hubble Space Telescope

The HST archive contains two images of 3C 220.3 with WFPC2 in the broadband filter F702W taken on 1994 March 11 and 1995 May 19, each with an exposure time of 300 s. The reprocessed preview images now in the archive clear show an arc, which was not visible on the early map published by McCarthy et al. (1997). To improve the contrast further, we have retrieved, aligned, and co-added the two F702W images (Figure 1), adjusting the positions of A and B to match those from the Keck AO (see Section 2.4).

Photometric calibration was derived by measuring the total 702 nm flux density in a circular aperture of 5′/6 diameter, containing sources A and B and the entire arc, and scaling it to the total r- and i-band photometry of 3C 220.3 listed by the Sloan Digital Sky Survey (SDSS, Data Release 9). The total magnitudes (in the AB system) are r = 20.96 ± 0.13 at 623 nm and i = 20.09 ± 0.10 at 764 nm, yielding the total 702 nm flux density listed in Table 1. The 702 nm flux densities of A and B separately are listed in Table 2.

2.4. Near-infrared Imaging with Keck Adaptive Optics

The 3C 220.3 system was observed with the Keck II telescope on UT 2012 December 24 utilizing NIR2 behind the AO bench. The laser guide star system was used in conjunction with an
Figure 1. Top: negative gray-scale Keck AO (K'-band, 2.124 μm) image of 3C 220.3 with linear contours at other wavelengths: blue = HST at 702 nm; red = SMA at 1 mm; green = JVLA at 3.3 cm. The position of the radio core is at R.A. = 09:39:23.837, decl. = +83:15:25.90 (J2000). Bottom: high resolution HST, Keck AO, and SMA images of 3C 220.3 with key components labeled. Image sizes are 12′′ square with north up and east to the left.

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

Table 2

<table>
<thead>
<tr>
<th>Lens</th>
<th>702 nm</th>
<th>2.124 μm</th>
<th>Sérsic Index $n$</th>
<th>$r_{\text{eff}}$ (arcsec)</th>
<th>Axis Ratio $b/a$</th>
<th>P.A. (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>14.8 ± 3.5</td>
<td>76.7 ± 15.0</td>
<td>0.9 ± 0.09</td>
<td>0.29 ± 0.04</td>
<td>0.7 ± 0.07</td>
<td>29 ± 12</td>
</tr>
<tr>
<td>B</td>
<td>4.45 ± 1.1</td>
<td>20.0 ± 3.0</td>
<td>1.0 ± 0.12</td>
<td>0.20 ± 0.04</td>
<td>0.6 ± 0.11</td>
<td>31 ± 25</td>
</tr>
</tbody>
</table>

Notes. The flux densities in μJy were determined using SExtractor with auto aperture option. The resulting elliptical apertures have an average circular diameter of about twice the Einstein radii of A and B. For surface profile modeling, the Keck AO data were fitted by PSF-convolved Sérsic profiles.

$R = 15.5$ mag tip-tilt star located 23′′ from 3C 220.3. The wide camera was used, which provides a pixel scale of 0′′.0397 and a field of view of 40″ on a side. We obtained nine dithered exposures in the K' band (2.124 μm), each exposure consisting of four co-added frames of 30 s each.

The data were reduced using a Python pipeline that flat-fields the data, subtracts sky emission, aligns the images, and drizzles to create the final image. The last step corrects for distortions that create a non-uniform pixel scale. The final image (Figure 1) used eight of the nine input images, giving a total exposure time of 960 s. The image center position was derived from three Two Micron All Sky Survey (2MASS) point sources in the frame. With this astrometry, A is within 0′.4 of the radio core position, and the final image coordinates were adjusted by this amount to match A to the radio core position (Figure 1).

Photometric calibration was derived using an object (probably a compact galaxy) located on the Keck AO mosaic at R.A. = 09:39:34.18, decl. = +83:14:51.4 (J2000), and $K_s = 15.04 ± 0.124$ mag listed by 2MASS.

2.5. Spitzer/IRAC Imaging

We observed 3C 220.3 on UT 2013 June 14 using the IRAC instrument (Fazio et al. 2004) of the Spitzer Space Telescope (Werner et al. 2004). Nine dithered exposures of 30 s each were secured at 3.6 μm and at 4.5 μm. The dithered images were combined and resampled to a pixel size of 0′.3 using IRACproc (Schuster et al. 2006) and MOPEX. 3C 220.3 is clearly resolved at both 3.6 and 4.5 μm. Approximate deconvolution gives full widths at half maximum surface brightness of 3″ at 3.6 μm and 3″ at 4.5 μm. These are comparable to the diameter of the Einstein ring (3″). However, the limited angular resolution makes it not straightforward to distinguish the components seen on the KECK AO image, i.e.,
sources A and B and the knots and arc. The IRAC photometry of 3C 220.3 listed in Table 1 is consistent with that from WISE.

2.6. Submillimeter Array

SMA imaging of 3C 220.3 was obtained with Director’s Discretionary Time in the 2012B semester as part of program 2012B-S084. Observations were conducted in extended array configuration (maximum baseline length of 226 m) on UT 2013 January 11 in good weather conditions (\(T_{225\text{GHz}} \sim 0.08\)) for \(t_{\text{int}} = 9.6\) hr on-source integration time. Compact array observations (maximum baseline length of 76 m) were obtained on UT 2013 February 10, again in good weather conditions (\(T_{225\text{GHz}} \sim 0.08\)) for \(t_{\text{int}} = 6.7\) hr. Phase stability for both observing nights was good (phase errors between 10’ and 20’ rms). The SMA receivers provided 8 GHz of instantaneous bandwidth (considering both sidebands) and were tuned such that the upper sideband was centered on 302.927 GHz.

Calibration of the \(uv\) visibilities was performed using the Interactive Data Language MIR package. The blazar 3C 84 was used as the primary bandpass calibrator and Titan for absolute flux calibration. The nearby quasars 0721+713 (\(S_{303\text{GHz}} = 5.5\) Jy, 13:6 from target) and 1048+717 (\(S_{303\text{GHz}} = 0.8\) Jy 12:0 from target) were used for phase and gain calibration. We checked the reliability of the phase-gain solutions using the quasar 1044+809 (\(S_{303\text{GHz}} = 0.3\) Jy, 3:2 from target).

The Multichannel Image Reconstruction, Image Analysis, and Display (MIRIAD) software package (Sault et al. 1995) was used to invert the \(uv\) visibilities and deconvolve the dirty map. Natural weighting was chosen to obtain maximum sensitivity and resulted in an elliptical Gaussian beam with a full-width half-maximum (FWHM) of 1.40’ × 1.17’ and major axis position angle 146° east of north.

2.7. Visible Spectroscopy

2.7.1. Palomar

We observed 3C 220.3 for 1 hr using the Double Spectrograph on the Palomar 5 m Hale Telescope on UT 2013 February 12. The observations were obtained with three 1200 s exposures at a position angle of 31° and the 5500 Å dichroic, yielding a wavelength coverage of 3200–8800 Å. Conditions were photometric, though the seeing was poor (\(\sim 2’’\)). The observations used the 1’’ wide longslit, the 600/4000 blue grating (resolving power \(R \sim 1000\)), and the 600/10000 red grating (\(R \sim 2400\)). We processed the data using standard techniques and flux calibrated the spectra using standard stars from Massey & Gronwall (1990) observed during the same night.

The Palomar observations failed to detect the arc but confirmed the redshift of 3C 220.3 at \(z = 0.685\) based on the detection of narrow [O iii] \(\lambda 5007\) emission lines, typical of PRGs (McCarthy 1993; Stern et al. 1999). The only previously published redshift of 3C 220.3 to our knowledge was based on [O iii] \(\lambda 3727\) and [Ne v] \(\lambda 3346\) emission with no figure available (Spinrad et al. 1985).\(^{18}\) [Ne v] (redshifted to \(\sim 5560\) Å) was not detected in our Palomar spectrum, likely owing to its proximity to the dichroic crossover.

\(^{18}\) Jasper Wall kindly showed us a plot of another spectrum of 3C 220.3 taken with the 4 m William Herschel Telescope, La Palma, yielding \(z = 0.685\) based on a single emission line adopted to be [O iii] \(\lambda 3727\). This spectrum is mentioned at http://www-astro.physics.ox.ac.uk/~sr/grimes.html.
atmospheric A

the blue and red arms of LRIS. Absorption at 7600 Å is due to the uncorrected
from both sources. The gap at 5500 Å corresponds to the dichroic separating
SMG at 

source labeled “ser 1” in Figure 2. The 3C 220.3 stripe has a width of ∼4 ″ , the lower part originating from source B, the central part from source A, and the upper part from the arc. The dark knots in the stripe are due to emission lines as labeled. Mg II at 2798 from the radio galaxy (source A) is centrally located in the continuum, while the C II and He II lines are offset upward, indicating an origin in the arc. The line on the left side (at 3920 Å) seems more central, suggesting it is primarily C II from the radio galaxy rather than Lyα from the lensed SMG.

3C 220.3 was observed with Chandra ACIS-S on UT 2013 January 21 for 10 ks. The total, broadband (0.3– 10 keV) X-ray flux density of the system is ∼1.14 × 10−14 erg cm−2 s−1, which yields a luminosity of \( L_X = 2.4 \times 10^{43} \text{ erg s}^{-1} \) if all photons originate at \( z = 0.685 \). The detection includes 16 photons (Figure 5) with 1.8 expected from a smooth background. Photons are distributed across the system. It is not generally possible to identify the X-ray photons with specific structure(s), although two photons are aligned with the radio core (implying \( L_X \sim 3 \times 10^{42} \text{ erg s}^{-1} \)). The observational upper limit of \( \sim 10^{43} \text{ erg s}^{-1} \) for extended, hot X-ray emitting gas is about a factor of 10 below typical values for X-ray clusters (Pratt et al. 2009).

2.8. Chandra X-Ray Observations

3C 220.3 was observed with Chandra ACIS-S on UT 2013 January 21 for 10 ks. The total, broadband (0.3– 10 keV) X-ray flux density of the system is \( \sim 1.14 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \), which yields a luminosity of \( L_X = 2.4 \times 10^{43} \text{ erg s}^{-1} \) if all photons originate at \( z = 0.685 \). The gap at 5500 Å corresponds to the dichroic separating

The horizontal dark stripe in the middle of the frame corresponds to the 3C 220.3 system, and the upper dark stripe originates from the faint extended source labeled “ser 1” in Figure 2. The 3C 220.3 stripe has a width of ∼4 ″ , the lower part originating from source B, the central part from source A, and the upper part from the arc. The dark knots in the stripe are due to emission lines as labeled. Mg II at 2798 from the radio galaxy (source A) is centrally located in the continuum, while the C II and He II lines are offset upward, indicating an origin in the arc. The line on the left side (at 3920 Å) seems more central, suggesting it is primarily C II from the radio galaxy rather than Lyα from the lensed SMG.

2.9. Spectral Energy Distribution

Figure 6 and Table 1 give the spectral energy distribution (SED) of the combined objects. The Herschel beam includes all flux from the SMG knots and arc and from the two lenses A and B, but the contributions of A and B are likely to be negligible. In the comparison sample of 12 Herschel observed 3C objects, only three were detected with Herschel. Their SEDs peak at \( \lambda_{\text{obs}} \sim 100 \text{ m} \) with flux densities below 100 mJy and decline toward shorter and longer wavelengths. The remaining nine 3C radio galaxies are undetected at 70–160 m with upper limits well below the 3C 220.3 flux densities. Lens galaxy B is even less likely than A to contribute to the FIR flux density because it is ∼3 times fainter than A in the NIR and has a slightly bluer visible-NIR color (Table 2, \( f(2.124 \mu m)/f(702 \text{ nm}) \sim 5 \) and 4 for A and B, respectively). The small 70–160 m flux densities of the 12 comparison 3C objects and the non-detection of A and B at 1 mm imply that most of the \( \lambda > 200 \text{ m} \) flux comes from the SMG at \( z = 2.221 \).

The visible-to-850 m flux density ratio of the 3C 220.3 SMG is consistent with the median of 73 SMGs at \( z = 2.2 \) (Chapman et al. 2005; Figure 6). The best-fit dust temperature \( T \sim 36 \text{ K} \) of the 250–1000 m dust emission (for a dust emissivity index \( \beta = 1.5 \) and \( z = 2.221 \)) is typical for unlensed SMGs (Magnelli et al. 2012).19 If the dust were associated with the radio source at \( z = 0.685 \), its temperature would be < 20 K and its dust mass, \( M_d > 10^{11} M_\odot \), would be similar to the stellar mass.

19 SED fitting with several dust components covering a range of temperatures does not result in a unique stable solution, but the temperature \( T \) of the FIR-millimeter decline is a quite stable, widely used characteristic parameter.
and more massive than typically observed in AGN (>30 K, \( M_d < 0.01 M_{\odot,\text{stars}} \)). Accounting for a total magnification factor \( \mu \sim 10 \) derived from lens modeling (Section 3), the intrinsic (demagnified) FIR luminosity of the SMG is \( L(\text{FIR}) \sim 1 \times 10^{13} L_\odot \).

3. Lens Modeling

To derive the lens model of the 3C 220.3 system, we used the Keck AO image because it has higher signal-to-noise ratio (S/N) and shows more details than the \( HST \) and SMA images (Figure 1). The lens parameters derived from the Keck image were then used to model the \( HST \) and SMA images.

We first tried a model with a single lens centered on galaxy A, assuming object B is part of the lensed source. This failed because it leaves bright residuals at the B position. In other words, B is too bright relative to the other lensed features in the Keck image to be fully explained as part of the lensed image. Moreover, B is not seen in the SMA image. The (hypothetical) case that B is a low-luminosity foreground galaxy with mass too low to influence the lensing leaves bright residuals at the B position. The remaining possibility is that B is part of the lens. Therefore, our preferred model has two lenses centered on galaxies A and B.

3.1. Keck AO Image

Lens models were derived from a fully Bayesian, grid-based modeling technique (Koopmans 2005; Vegetti & Koopmans 2009) that simultaneously solves for the parameters \( \eta \) of the lensing galaxies and the surface brightness distribution of the background source \( s \). The Bayesian prior was expressed as the source regularization form \( R \) with regularization level \( \lambda \) (Vegetti & Koopmans 2009). Because this prior is flat in \( \log(\lambda) \) and \( \eta \), the maximum of \( P(d|\lambda, \eta, M, R) \), that is, the probability of the data \( d \) given a lens mass model \( M \), determines the most probable a posteriori fit (see Equation (35) of Vegetti & Koopmans 2009). The Bayesian approach takes into account the smoothness of the source via the regularization \( R \) at the level \( \lambda \) (Suyu et al. 2006).

For model fitting we have used the uncertainty map computed as the quadrature sum of the image read out noise and Poisson noise using the number of detected photons. Notably, the fitting results are very similar (within a few percent) to those obtained when using a uniform uncertainty determined from clean background areas of the image.

To avoid that the model is affected by image pixels unrelated to the source, all model fitting was restricted to the “smallest-possible” mask as shown in Figure 7. The mask was set by smoothing the images, then making an S/N cut. This ensures that structure, that is present but below the noise level, is not cut off from the data. Thus, the mask is a compromise between being large enough to allow identification of possible extended residuals and being small enough to cover only the arc and the area around B, where the fourth counter-image is expected.

We here consider three scenarios. In case 1, object B is part of the lensed source; in case 2, B is a low-luminosity foreground galaxy with mass too small to influence the lensing; and in case 3, B is part of the lens. The lens modeling began with case 1, using a postage stamp image centered on the lens system but with the lensing galaxy (A) subtracted using analytic models. For cases 2 and 3 both galaxies (A and B) were subtracted. The surface brightness of the galaxies A and B is best fitted by point-spread function (PSF)-convolved Sérsic (\( n \sim 1 \)) profiles (and not by de Vaucouleurs profiles). Their parameters are listed in Table 3.

The lens mass distribution was modeled as an ellipsoid with a free power-law slope for the density profile. In general, such approximations provide good fits to describe the combined luminous and dark mass profiles (e.g., Treu & Koopmans 2004; Auger et al. 2009). The profiles were centered on the light distribution of A (or A and B), allowing for a small positional deviation. The parameters of the best-fit single and double lens models are listed in Table 3.

Case 1. Figure 7 shows that the positions of the three knots connected by the arc can roughly be fitted using the single lens A alone. However, the image residual (data minus model) contains not only a bright relic at the position of B but also negative features of the three knots and the arc (seen as blue arcs in the bottom left panel of Figure 7). In other words, B is too bright relative to the other lensed features in the Keck image to be fully explained as part of the lensed image. In addition, the light profile of B is well fit by a smooth Sérsic profile that is unlike the complex structure of the three images along the arc, and finally the model requires the lensed source galaxy to be highly distorted. Overall case 1 is a poor fit to the data, and this is confirmed by the Bayesian evidence (Table 3).

Case 2. Figure 8 shows that the three knots connected by the arc can be fitted, but there is a bright negative residual just at the B position because this model requires a counter-image of the arc having about 2/3 the flux of the three bright knots. Therefore, B should be seen in the SMA image, but it is not (Figures 1 and 10). These discrepancies lead us to reject the single-lens model and consider B as part of the lens, consistent with its probable redshift (Section 2.7.2).

Case 3. Using two lenses significantly improves the modeling results.\(^{20}\) Figure 9 illustrates the consistency of the observed and

---

\(^{20}\) The two-lens model includes the possibility that B is a lens but not necessarily at the same redshift as A, because any two-lens-plane system can be modeled to first order as a single lens plane with a rescaled Einstein radius (Keeton 2001). The Einstein radii are free parameters of the model, and therefore our results are independent of the true redshift of B. If its redshift is lower than assumed, its mass has to be higher, but all other model results are exactly the same.
Figure 7. Results from modeling the Keck AO image with a single lens centered on galaxy A, located at (0, 0). The regions outside the Einstein half ring and source B are masked and not used for the model fit. While the total map has $200 \times 200 = 40,000$ pixels, the area used for modeling has 10,912 pixels. North is up and east is to the left. Top left: data after subtraction of A. Top right: modeled image. The cross marks the position of A. Bottom left: residual image, data minus model image. Bottom right: reconstructed image of the source in the source plane ($50 \times 50 = 2500$ source pixels). (A color version of this figure is available in the online journal.)

Table 3

<table>
<thead>
<tr>
<th>Lens</th>
<th>Einstein Radius (′)</th>
<th>Lens Position x, y (′)</th>
<th>b/a</th>
<th>P.A. (deg)</th>
<th>Slope (s)</th>
<th>Shear Strength</th>
<th>Lens Mass $10^{11} M_\odot$</th>
<th>ln Ev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single lens, case 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1.67</td>
<td>0.043, −0.183</td>
<td>0.61</td>
<td>38.7</td>
<td>2.071</td>
<td>0.093</td>
<td>5.4</td>
<td>8.9</td>
</tr>
<tr>
<td>Unc A</td>
<td>0.09</td>
<td>0.007, 0.010</td>
<td>0.10</td>
<td>8</td>
<td>0.140</td>
<td>0.031</td>
<td>20</td>
<td>1.1</td>
</tr>
<tr>
<td>Single lens, case 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1.68</td>
<td>0.113, −0.324</td>
<td>0.27</td>
<td>30.8</td>
<td>2.310</td>
<td>0.146</td>
<td>13.3</td>
<td>9.0</td>
</tr>
<tr>
<td>Unc A</td>
<td>0.08</td>
<td>0.013, 0.028</td>
<td>0.16</td>
<td>12</td>
<td>0.230</td>
<td>0.029</td>
<td>24</td>
<td>1.0</td>
</tr>
<tr>
<td>Double lens, case 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1.02</td>
<td>−0.174, 0.223</td>
<td>0.61</td>
<td>35.7</td>
<td>1.837</td>
<td>0.098</td>
<td>5.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Unc A</td>
<td>0.07</td>
<td>0.010, 0.010</td>
<td>0.13</td>
<td>10</td>
<td>0.090</td>
<td>0.030</td>
<td>23</td>
<td>0.5</td>
</tr>
<tr>
<td>B</td>
<td>0.61</td>
<td>0.861, −1.415</td>
<td>0.74</td>
<td>42.1</td>
<td>1.964</td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Unc B</td>
<td>0.05</td>
<td>0.007, 0.008</td>
<td>0.12</td>
<td>15</td>
<td>0.150</td>
<td></td>
<td></td>
<td>0.2</td>
</tr>
</tbody>
</table>

Notes. The uncertainties (labeled “unc”) are listed in the line underneath the parameter values. The lens positions (in Column 3) are relative to the axis from the observer to the source. Columns 4 and 5 give the axis ratio ($b/a = \text{minor/major axis}$) and position angle of the lenses. Radial mass surface density $\rho \propto r^{-s}$ (in Column 6), a slope $s = 2.0$ corresponds to a singular isothermal ellipsoid. Column 10 gives the natural logarithm of the Bayesian evidence (ln Ev) of cases 1–3, respectively.

modeled images and the featureless image residual all along the arc. The two-lens model predicts two additional faint images: a fourth image on the opposite side of the Einstein half-ring and a fifth (central) image. Their positions agree with two faint blobs in the Keck data after subtraction of a galaxy model for A and B and with faint features in the SMA map (Figure 10). However, both images are close to the noise level and need to be confirmed with deeper optical/NIR data and a higher spatial resolution 1 mm map. The magnification factor of the two-lens model at 2.124 $\mu$m is $\mu \sim 7 \pm 2$. 
Figure 8. Results from modeling the Keck AO image with a single lens centered on galaxy A, as for Figure 7, but using data after subtraction of both galaxies A and B. (A color version of this figure is available in the online journal.)

Figure 9. Results from modeling the Keck AO image with a double lens. The regions outside the Einstein half ring and source B are masked and not used for the model fit. North is up and east is to the left. Top left: data after subtraction of A and B. Top right: modeled image. The crosses mark the positions of A and B. There are two very faint objects, one between A and B and one southwest of B; they may be the central image and the counter-image of the arc. Bottom left: residual image, data minus model image. Bottom right: reconstructed image of the source in the source plane. (A color version of this figure is available in the online journal.)

Figure 10. Modeled Keck AO image (color representing surface brightness) based on the double lens. Frame size is 7.5′. Superposed contours show the SMA map with the three bright knots labeled. The contours start at 2.5σ and increase in steps of 1σ. Source image 4 coincides with a faint 1 mm contour. Source image 5 could be indicated by the faint tail of knot 3. Galaxy B lies between knots 4 and 5. (A color version of this figure is available in the online journal.)
underestimated, thus overestimating the reduced $\chi^2$ source pixels in the computation of number of degrees of freedom. Source plane pixels are regularized, it is not straightforward to calculate the Bayesian evidence to determine the best-fit lens model, which has been underestimated, or that the noise map is underestimated by about 5%–10%, which could quite easily be the case.

Visual inspection of the residual maps (data – model, Figures 7–9) and the absence of B in the SMA image clearly favors the two-lens model. The statistical quantification (Table 3) agrees. The two-lens model has six extra free parameters allowing us to fit the data better. This is not an artifact of the higher number of free parameters, rather the Bayesian evidence automatically encodes Occam’s razor, and models with more free parameters are automatically penalized unless they are really required by the data. In essence, adding six extra parameters to the mass model improves the Bayesian evidence by $\Delta \log E^d = 300$, which corresponds to about $25\sigma$ difference, and produces a source galaxy that looks “normal” (Figure 9). If B were not important in the lensing, the model could have set its mass to zero, but this was not the result.

Both the single and double lens models require an external shear of $\sim 0.1$ for lens A (Table 3), suggesting the presence of additional lensing components. While the marginally visible knot 5 (Figure 10) could be the fifth central image, alternatively it could be a third lensing galaxy C. We checked for this possibility, running a lens model with three lenses centered on A, B, and C. While C is still produced by the modeling as an extra image, a bit of mass is removed from B to C (Einstein radius of $C = 0′030$), and the shear of lens A remains ($\gamma = 0.092$). Also, the Bayesian evidence for the three lens model is poorer than for the two-lens model, and we do not use the three lens model further. Section 5 discusses the implications of the shear.

3.2. HST Image

Figure 11 shows the results of applying the double lens model to the HST image. The image residuals are featureless, supporting the adequacy of the model.

3.3. SMA Image

For interferometers such as the SMA, proper lens modeling requires the use of visibilities in the Fourier domain rather than surface brightness maps. We followed the technique outlined by Bussmann et al. (2012). Using the Gravlens software, we ray-traced the emission in the source plane to the image plane for a given lensing mass distribution and source morphology. This image plane surface brightness map was then used as input to MIRIAD’s UVMODEL task, which computes the Fourier transform of an image and matches the resulting visibilities to the sampling of the observed SMA visibility data set. We used the emcee software package (for details, see Foreman-Mackey et al. 2013) to perform a Markov chain Monte Carlo sampling of the posterior probability density function of our model parameters.

We modeled the SMA data using the lens parameters of the two-lens model derived from the Keck AO data (Table 3) and a Sérsic profile to represent the morphology of the SMG in the source plane. The SMA data (having lower spatial resolution than the Keck AO data) do not justify a more complex model. The best-fit source has a Sérsic index $n_s = 1.38 \pm 0.46$, half-light radius $r_s = 1.26 \pm 0.38$ kpc, axis ratio $b/a = 0.52 \pm 0.09$, and P.A. = $20° \pm 17°$.

The best-fit result is shown in Figure 12. Because the model fitting is performed in the Fourier domain and then transformed

---

21 We have used the Bayesian evidence to determine the best-fit lens model, but for comparison we also give here the more commonly used $\chi^2$ values. $\chi^2$ values were determined from the residual maps (data minus model) divided by the uncertainty map inside an area of 10912 pixels (Figure 7). However, when source plane pixels are regularized, it is not straightforward to calculate the number of degrees of freedom $N$(dof). The calculation here excludes the $50 \times 50$ source pixels in the computation of $N$(dof), which is therefore underestimated, thus overstating the reduced $\chi^2$. With 8 and 14 free parameters for the single- and two-lens models $N$(dof) = 10,904 and 10,898, respectively. The $\chi^2$ values of cases 1, 2, and 3 are 14,393, 12,976, and 12,642, respectively. With these underestimated $N$(dof), the $\chi^2$ values of cases 1, 2, and 3 are 3.12, 1.19, and 1.16, respectively. In the limit of a large number of degrees of freedom, $N$(dof) = 10,000, $\chi^2$ has a standard deviation of $\sqrt{2/(N$(dof))}$ = 0.014, hence the $\chi^2$ values for case 2 and case 3 differ at the $2\sigma$ level and are about $10\sigma$ away from unity. The $\chi^2$ statistic suggests that either a more complex lens model may be needed, or more likely that $N$(dof) has been underestimated, or that the noise map is underestimated by about 5%–10%, which could quite easily be the case.

22 http://redfive.rutgers.edu/~keeton/gravlens/
A luminous AGN breaking up its dust cocoon is one interpretation of the so-called alignment effect in PRGs at high redshift (Chambers et al. 1987; McCarthy et al. 1987). In those cases, the breakup is apparently caused by the high-energy particles that create the radio lobes. In contrast to such radio galaxies, the JVLA radio map of 3C 220.3 shows no indication for powerful radio activity in the SMG, either at the known position of lensed features (the arc and the three knots) or from extended lobes, which would be expected to be aligned with the UV emission (approximately north–south). While starburst winds may be able to blow holes into the dust cocoon of the SMG and to expel part of the gas (Tenorio-Tagle & Munoz-Tunon 1998), a wind from the hidden QSO may be more likely to be forming the open channels. Regardless of the mechanism, the in-progress breakup of the dust cocoon makes the 3C 220.3 SMG an excellent lensed laboratory to explore, with brightness magnification and at higher spatial resolution than possible for unlensed systems, the properties of a dust-enshrouded AGN at high redshift and its feedback on the star forming activity in the host.

5. THE MASS-TO-LIGHT RATIO OF 3C 220.3

The double lens model (Table 3) gives total—baryonic and dark—masses for the lensing galaxies A and B. The calculated masses are \( M_A \sim 3.5 \times 10^{11} M_\odot \) inside a cylinder of Einstein radius \( r_E = 1.02 \) (7.2 kpc at \( z = 0.685 \)), and \( M_B \sim 1.2 \times 10^{11} M_\odot \) (\( r_E = 0.61 \)) kpc. The statistical uncertainties (Table 3), as determined from the model residuals, do not account for possible systematic effects.

One systematic effect for the mass estimates is that the best-fit model requires an external shear of \( \sim 0.1 \) for lens A (Table 3), suggesting the presence of additional lensing components. This may be consistent with the presence of the 2\( \sigma \) residuals of modeling the SMA image (Figure 12). In principle, such a component could be a dark matter halo encompassing A and B. On the other hand, our Chandra image shows no extended X-ray cluster gas (Figure 5), arguing against explaining the shear by a common dark matter halo. More likely is that 3C 220.3 is located in a small group with one or more faint galaxies having escaped detection on our images so far (possibly masked by the lensed structures). A similar situation was reported for the gravitational lens CLASS B2045+265, where the initially inferred high \( M/L \) was reduced by a group of faint lensing satellites (Fassnacht et al. 1999; McKeen et al. 2007). Therefore, we consider the masses \( M_A \) and \( M_B \) as upper limits.

To estimate the luminous mass of galaxies A and B inside the Einstein radius, only two filters (Table 2) provide adequate data. (The spectra show no stellar features.) Using the Starburst99 tool version 6.0.4 (Leitherer et al. 1999)\(^{24}\) and adopting a dominant old stellar population typical for radio galaxy hosts (age \( \sim 3 \) Gyr and Salpeter initial mass function), we obtain stellar masses \( M_A^{*} \sim 2.2 \times 10^{11} M_\odot \) and \( M_B^{*} \sim 5.5 \times 10^{10} M_\odot \). The uncertainty of these estimates is about a factor of two, dominated by uncertainties in the photometry of A and B inside the Einstein radius (especially from the noisy \( HST \) image) and the unknown age of the stellar population. With these numbers, the ratio of dark-to-luminous mass (projected within the Einstein radius) lies in the range between 0.7 and 1.2, i.e.,

\(^{23}\) The single lens model of 3C 220.3 yields \( M_{lt} \sim 9 \times 10^{11} M_\odot \) which is a factor of two higher than the sum of \( M_A \) and \( M_B \) in the two-component lens model. If the shear of the two-component lens model is due to further not yet identified lensing galaxies, the mass sum may be reduced further.

\(^{24}\) http://www.stsci.edu/science/starburst99/docs/default.htm
dark matter fractions $f_{DM}$ of about 40\% and 55\% with large uncertainties ($\sim$30\%).

To check whether the radio galaxy 3C 220.3 is associated with a usually high amount of dark matter, suitable comparison data are required. A proper comparison sample should contain field galaxies of stellar mass and redshift similar to 3C 220.3. The best existing comparison samples are the CfA-Arizona Space Telescope LEns Survey (CASTLES)\textsuperscript{25} of gravitational lenses (e.g., Lehar et al. 2000) and a few $z \sim 0.6$ objects in the Lenses Structure and Dynamics (LSD) survey (Treu & Koopmans 2004) and the Strong Lenses in the Legacy Survey (SL2S; Ruff et al. 2011). Figure 14 shows the Einstein radius (converted to kpc) versus the $i$-band luminosity of the CASTLES, LSD, and SL2S lensing galaxies in the redshift range $0.5 < z < 0.8$, comparable to the redshift of 3C 220.3, and having lensed source redshift $z > 1.3$ to ensure that it is sufficiently far from the lens. The $i$-band luminosities of 3C 220.3 A and B were derived by interpolation of the HST and Keck photometry (Table 2). Assuming the Einstein radius traces total mass, the $M/L$ of 3C 220.3 A and B, $M/L(i) \sim 8 \pm 4$, is comparable to that of the CASTLES, LSD, and SL2S lenses, perhaps at the high end of their distribution. LSD and SL2S list dark matter fractions in the range $0.25 < f_{DM} < 0.75$ (projected within the Einstein radius), encompassing the value $f_{DM} \sim 0.4$ of 3C 220.3.

6. CONCLUSIONS

The exceptional 3C 220.3 gravitational lensing system consists of an SMG (hosting a dust-enshrouded type-2 QSO) which is lensed by a PRG, i.e., a radio-loud type-2 quasar. The system permits determination of the total galaxy-scale mass for a powerful double-lobed radio galaxy that is not in a rich galaxy cluster. Comparison with available galaxy-scale lenses indicates an average to moderately high dark matter fraction, but the comparison with radio-quiet galaxies needs to be confirmed with future lens samples near the redshift of 3C 220.3. Further observations of 3C 220.3 with increased S/N will allow a more accurate lens model, give a better stellar mass measurement for the lenses, and allow a more detailed reconstruction of the SMG morphology. They should also clarify whether other galaxies important to the lensing, as for example in a small group, exist.

The SMG fits the paradigm of a composite AGN/starburst galaxy in the early universe. The available data indicate the breakup of the dust cocoon, which represents a fundamental transition phase in a violently star-forming galaxy. Deeper X-ray observations and (sub-)millimeter spectroscopy, both possible at high spatial resolution, will allow exploration of the break-up process and the kinematics of any related outflows in a high-redshift starburst-AGN.

We thank Shri R. Kulkarni, who enabled us to obtain the Keck LRIS spectrum, and Dan Perley, who assisted with these observations. We also thank Dale Frail for granting Director’s Discretionary time at the JVLA, Adam Deller for helping with the JVLA data reduction, and Ray Blundell for granting Director’s Discretionary time at the SMA. We thank the referee for comments that improved the paper.

REFERENCES


\textsuperscript{25} http://www.cfa.harvard.edu/castles

---

Figure 14. Einstein radius vs. $i$-band luminosity of 3C 220.3 A and B (large open circles) and the CASTLES, LSD, and SL2S lensing galaxies at comparable redshift (filled symbols).

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