# Supplementary Information

## Removing Orientation-Induced Localization Biases in Single Molecule Microscopy Using a Broadband Metasurface Mask

Mikael P. Backlund, Amir Arbabi, Petar N. Petrov, Ehsan Arbabi, Saumya Saurabh, Andrei Faraon, and W. E. Moerner

## Green’s tensor and symmetry

The complex amplitude of the electric field at the BFP due to an oscillating electric dipole can be written in terms of the Green’s tensor $\mathbf{G}_{\mathrm{BFP}}$ [1, 2]:

$$\mathbf{E}_{\mathrm{BFP}} = Ae^{inkz_D(1-\rho^2)^{1/2}}\mathbf{G}_{\mathrm{BFP}}\hat{\boldsymbol{\mu}} \tag{S1}$$

where $A$ is a constant, $k$ is wave number of the collected light, $z_{\mathrm{D}}$ is the axial position of the molecular dipole relative to the focal plane, and $n$ is the index of refraction of the immersion medium. For simplicity, in this section we assume that the index of refraction of the sample is matched to that of the objective immersion medium (on the other hand, our simulations below allow for mismatched media and layered samples[1]). We assume scaled units such that $\rho = 1$ corresponds to a distance $f_{\mathrm{OL}}$ (the focal length of the objective lens) from the optical axis. Due to the finite numerical aperture (NA) of the objective, $\mathbf{E}_{\mathrm{BFP}}$ is supported only on the region $\rho < 1$. Note that equation (S1) applies exactly for a molecule with lateral position along the optical axis, i.e. $(x_{\mathrm{D}} = 0, y_{\mathrm{D}} = 0)$; the important results below still apply for any $(x_{\mathrm{D}}, y_{\mathrm{D}})$ since we assume a shift-invariant imaging system. The Green’s tensor is defined in the Cartesian basis as follows[1, 2]:

$$\mathbf{G}_{\mathrm{BFP}} = \frac{e^{inkf_{\mathrm{OL}}}}{4\pi f_{\mathrm{OL}}}\sqrt{\frac{n}{\left(1-\rho^2\right)^{1/2}}}\begin{bmatrix} \sin^2\varphi + \sqrt{1-\rho^2}\cos^2\varphi & \left(\sqrt{1-\rho^2}-1\right)\sin 2\varphi & -\rho\cos\varphi \\ \left(\sqrt{1-\rho^2}-1\right)\sin 2\varphi & \cos^2\varphi + \sqrt{1-\rho^2}\sin^2\varphi & -\rho\sin\varphi \\ 0 & 0 & 0 \end{bmatrix} \tag{S2}$$

Expanding equations (S1) and (S2) gives:

$$\mathbf{E}_{\mathrm{BFP}} = B\frac{e^{inkz_{\mathrm{D}}(1-\rho^2)^{1/2}}}{\left(1-\rho^2\right)^{1/4}}\begin{bmatrix} \mu_x\left(y_{\mathrm{BFP}}^2 + x_{\mathrm{BFP}}^2\sqrt{1-\rho^2}\right) + \mu_y x_{\mathrm{BFP}} y_{\mathrm{BFP}}\left(\sqrt{1-\rho^2}-1\right) - \mu_z \rho x_{\mathrm{BFP}} \\ \mu_x x_{\mathrm{BFP}} y_{\mathrm{BFP}}\left(\sqrt{1-\rho^2}-1\right) + \mu_y\left(y_{\mathrm{BFP}}^2 + x_{\mathrm{BFP}}^2\sqrt{1-\rho^2}\right) - \mu_z \rho y_{\mathrm{BFP}} \end{bmatrix} \tag{S3}$$

where $x_{\mathrm{BFP}} = \rho\cos\varphi$, $y_{\mathrm{BFP}} = \rho\sin\varphi$, and we have absorbed all constants into *B*. In equation (S3) and all subsequent equations we omit the *z* component of the field since it is 0. In both components of $\mathbf{E}_{\mathrm{BFP}}$ the first two terms are symmetric with respect to inversion, i.e. under the substitution $(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}) \mapsto (-x_{\mathrm{BFP}}, -y_{\mathrm{BFP}})$, while the last term is antisymmetric with respect to inversion. Therefore $\mathbf{E}_{\mathrm{BFP}}$ is neither symmetric nor antisymmetric with respect to inversion except in the special cases $\mu_z = 0$ or $\mu_x = \mu_y = 0$, and in turn $I_{\mathrm{BFP}} = \left|\mathbf{E}_{\mathrm{BFP}}\right|^2$ is not symmetric with respect to inversion under the same conditions.

The field in the image plane of the microscope is related to that in the BFP via a scaled Fourier transform[2]:

$$\mathbf{E}_{\mathrm{IP}} = C\iint \mathbf{E}_{\mathrm{BFP}}\left(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}\right)\exp\left\{\frac{ik}{f_{\mathrm{TL}}}\left(x_{\mathrm{BFP}}x + y_{\mathrm{BFP}}y\right)\right\}dx_{\mathrm{BFP}}dy_{\mathrm{BFP}} \tag{S4}$$

where *x* and *y* are the image plane coordinates, *C* is a constant, and $f_{\mathrm{TL}}$ is the focal length of the tube lens. From the properties of Fourier transforms[3] it follows that if $\mathbf{E}_{\mathrm{BFP}}\left(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}\right) = \mathbf{E}_{\mathrm{BFP}}\left(-x_{\mathrm{BFP}}, -y_{\mathrm{BFP}}\right)$ then $\mathbf{E}_{\mathrm{IP}}(x, y) = \mathbf{E}_{\mathrm{IP}}(-x, -y)$; and if $\mathbf{E}_{\mathrm{BFP}}\left(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}\right) = -\mathbf{E}_{\mathrm{BFP}}\left(-x_{\mathrm{BFP}}, -y_{\mathrm{BFP}}\right)$ then $\mathbf{E}_{\mathrm{IP}}(x, y) = -\mathbf{E}_{\mathrm{IP}}(-x, -y)$. Thus if $I_{\mathrm{BFP}}\left(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}\right)$ is symmetric upon inversion, then so too is $I_{\mathrm{IP}}(x, y) = \left|\mathbf{E}_{\mathrm{IP}}(x, y)\right|^2$. Hence, to ensure that

$I_{\mathrm{IP}}(x, y)$ is symmetric for all $z_{\mathrm{D}}$ and avoid orientation-induced mislocalization, it is sufficient to seek a symmetric or antisymmetric $\mathbf{E}_{\mathrm{BFP}}\left(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}\right)$.

An azimuthal polarization filter located at the BFP would provide such an automatic correction[4]. This can be understood mathematically by left-multiplying equation (S1) by the azimuthal polarizer operator $\hat{\varphi}\hat{\varphi}^{\mathrm{T}}$, with $\hat{\varphi} = \left[-\sin\varphi, \cos\varphi, 0\right]^{\mathrm{T}}$. The resulting field $\mathbf{E}_{\mathrm{BFP}}^{(\varphi)}$ is given by:

$$\mathbf{E}_{\mathrm{BFP}}^{(\varphi)} = B\frac{e^{inkz_{\mathrm{D}}(1-\rho^2)^{1/2}}}{\left(1-\rho^2\right)^{1/4}}\begin{bmatrix} \mu_x y_{\mathrm{BFP}}^2 - \mu_y x_{\mathrm{BFP}} y_{\mathrm{BFP}} \\ -\mu_x x_{\mathrm{BFP}} y_{\mathrm{BFP}} + \mu_y x_{\mathrm{BFP}}^2 \end{bmatrix} \tag{S5}$$

From equation (S5) it is evident that $\mathbf{E}_{\mathrm{BFP}}^{(\varphi)}\left(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}\right) = \mathbf{E}_{\mathrm{BFP}}^{(\varphi)}\left(-x_{\mathrm{BFP}}, -y_{\mathrm{BFP}}\right)$ for any $\boldsymbol{\mu}$, and hence the resulting phi-PSF is symmetric.

The y-phi mask used in this study was designed to rotate the local polarization of the transmitted light without any additional spatially dependent phase delay. The effective Jones matrix of the device is:

$$\mathbf{J} = \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix} \tag{S6}$$

which yields the relations $\mathbf{J}\hat{\rho} = \hat{x}$ and $\mathbf{J}\hat{\varphi} = -\hat{y}$. Thus in order to effectively filter out the azimuthally polarized light at the Fourier plane we can place the y-phi mask at the BFP followed immediately by a linear polarizer passing y-polarized light. The expression for the field $\mathbf{E}_{\mathrm{BFP}}^{(y-\varphi)}$ is related to that given for $\mathbf{E}_{\mathrm{BFP}}^{(\varphi)}$, but not the same. It can be obtained by left-multiplying equation (S1) by $\hat{y}\hat{y}^{\mathrm{T}}\mathbf{J}$ to give:

$$\mathbf{E}_{\mathrm{BFP}}^{(y-\varphi)} = B\,\frac{e^{inkz_{\mathrm{D}}(1-\rho^2)^{1/2}}}{\left(1-\rho^2\right)^{1/4}}\,\rho \begin{bmatrix} 0 \\ \mu_x y_{\mathrm{BFP}} - \mu_y x_{\mathrm{BFP}} \end{bmatrix} \tag{S7}$$

Hence $\mathbf{E}_{\mathrm{BFP}}^{(y-\varphi)}$ is antisymmetric with respect to inversion, and so the final image (i.e. the y-phi-PSF) is symmetric.

## Phi-to-y conversion efficiency

The ideal y-phi mask is an ideal half-wave plate (HWP) whose optical axis orientation gradually varies across the mask (as shown in Fig. 2c). An ideal HWP with optical axis along the *x* axis has unity transmission for both *x*- and *y*-polarized optical waves, and imposes a relative phase shift of $\pi$ on the two polarizations. For a realistic HWP, such as the one realized using an array of nanoposts as shown in Supplementary Fig. 7a, both the transmission efficiencies and the relative phase deviate from their ideal values. To optimize the metasurface mask design for maximum phi-to-y conversion efficiency, it is important to relate the conversion efficiency to the HWP parameters. To this end, we assume that a y-phi mask is made using a non-ideal HWP and find the phi-to-y conversion efficiency. Assuming the orientation of the principal axis of the HWP varies across the mask as shown in Fig. 2c, the Jones matrix at a location with polar coordinates of $(\rho,\varphi)$ is given by:

$$\begin{aligned} \mathbf{J}' &= \begin{bmatrix} \cos(\varphi/2) & -\sin(\varphi/2) \\ \sin(\varphi/2) & \cos(\varphi/2) \end{bmatrix} \begin{bmatrix} t_x & 0 \\ 0 & t_y \end{bmatrix} \begin{bmatrix} \cos(\varphi/2) & \sin(\varphi/2) \\ -\sin(\varphi/2) & \cos(\varphi/2) \end{bmatrix} \\ &= \begin{bmatrix} \epsilon + \Delta\cos\varphi & \Delta\sin\varphi \\ \Delta\sin\varphi & \epsilon - \Delta\cos\varphi \end{bmatrix} \end{aligned} \tag{S8}$$

where $t_x$ and $t_y$ are the transmission coefficients of the HWP for the light polarized along its two principal axes, and for simplicity, we have defined $\Delta = \left(t_x - t_y\right)/2$ and $\epsilon = \left(t_x + t_y\right)/2$. Note that for an ideal HWP $\Delta = 1$, $\epsilon = 0$, and equation (S8) is reduced to equation (S6). For a general input beam with an electric field of $\mathbf{E}^{\text{in}} = f(\rho,\varphi)\hat{\varphi} + g(\rho,\varphi)\hat{\rho}$, we obtain the electric field of the light transmitted through the mask using $\mathbf{E}^{\text{out}} = \mathbf{J}' \cdot \mathbf{E}^{\text{in}}$:

$$\mathbf{E}^{\text{out}} = \begin{bmatrix} -\epsilon \sin\varphi \\ -\Delta + \epsilon \cos\varphi \end{bmatrix} f(\rho,\varphi) + \begin{bmatrix} \Delta + \epsilon \cos\varphi \\ \epsilon \sin\varphi \end{bmatrix} g(\rho,\varphi) \tag{S9}$$

The *y* component of $\mathbf{E}^{\text{out}}$ that passes through the polarizer (LP in Fig. 3) is given by:

$$E_y^{\text{out}} = -\Delta f(\rho,\varphi) + \epsilon\left[\cos\varphi\, f(\rho,\varphi) + \sin\varphi\, g(\rho,\varphi)\right] \tag{S10}$$

The first term on the right hand side of equation (S10) is the desired term representing the phi-to-y conversion, while the second term is undesired and should be minimized for the proper device operation. Thus, we define the phi-to-y conversion efficiency as:

$$\eta = |\Delta|^2 = \frac{1}{4}\left|t_y - t_x\right|^2 \tag{S11}$$

Note that $|\epsilon|^2 + |\Delta|^2 = \left(|t_x|^2 + |t_y|^2\right)/2 \le 1$, and therefore $|\epsilon|^2 \le \left(1-\eta\right)$, which sets an upper limit on the amplitude of the undesired term in equation (S10). As a result, we optimized the metasurface design by maximizing $\eta$ over the emission bandwidth of the fluorescent dye.

**Characterization of the y-phi mask**

The metasurface y-phi mask was characterized using the measurement setup schematically shown in Supplementary Fig. 8a. The mask was illuminated by a linearly polarized Gaussian beam, and the polarization state of the transmitted beam was analyzed for two different polarization states of the input beam. Light from a fiber-coupled external cavity semiconductor laser (Newport Velocity, tunable from 668 nm to 678 nm) was collimated using a fiber collimation package (Thorlabs F220APC-780), and polarized using a fiber polarization controller and the polarizer $P_1$ (Thorlabs LPNIR050-MP). The light transmitted through the y-phi mask was imaged using the objective lens $L_1$ (Thorlabs AC254-050-B-ML) and the tube lens $L_2$ (Thorlabs AC254-200-B-ML) to a camera (CoolSNAP K4, Photometrics). The polarization distribution of the transmitted light was examined by placing the polarizer $P_2$ between the objective lens and the tube lens, and recording the intensity profiles of the transmitted light for four different orientations of the $P_2$ transmission axis, and with $P_2$ removed from the setup. The measurements were performed at three different wavelengths: 668 nm, 673 nm, and 678 nm. Supplementary Fig. 8b shows the intensity images captured by the camera at $\lambda = 678$ nm for the *x*- and *y*-polarized input beams and for four polarization projection directions of $\varphi = 0°$, 45°, -45°, and 90°, which were set by rotating the transmission axis of the polarizer $P_2$. As the projected intensity profiles show, for the *x*-polarized input light the transmitted light is radially polarized, and for the *y*-polarized input beam the transmitted light is azimuthally polarized. Similar results were obtained at the other two wavelengths. The transmission efficiency of the y-phi mask was obtained as the ratio of the optical power of the transmitted beam to that of the input beam. The transmitted efficiency averaged over the three wavelengths was found to be 86.5%.

### Explanation for inclusion of half wave plate

As mentioned in the main text, we had to insert a phase compensating element in the setup (Fig. 3) in order to effectively compensate for the reflection from the glass prism within our Olympus IX71 microscope and the birefringence of the dichroic mirror. For our DCDHF-A-

6 imaging configuration this was achieved with a half wave plate. The reflection from the glass prism affects $\mathbf{E}_{\mathrm{BFP}}\left(x_{\mathrm{BFP}}, y_{\mathrm{BFP}}\right)$ in two ways. First, there is a geometrical transformation captured by the substitution $y_{\mathrm{BFP}} \mapsto -y_{\mathrm{BFP}}$. Second, there is relative phase introduced between the *x* and *y* components of $\mathbf{E}_{\mathrm{BFP}}$. If the reflection were from a perfect metal mirror then $E_{\mathrm{BFP},y}$ would carry a phase factor $e^{i\pi} = -1$. One can work through the math in order to convince oneself that this negative sign would essentially remove the effect of the $y_{\mathrm{BFP}} \mapsto -y_{\mathrm{BFP}}$ substitution and ultimately yield the ideal y-phi-PSF. However, the glass prism affects the relative phase differently than a metal mirror would. Light incident on a glass-air interface from the glass side at an angle of 45° undergoes TIR with phase delays on each of the *x* and *y* components that can be determined from the appropriate Fresnel reflection coefficients. The resulting relative phase delay is captured by multiplication by the following Jones matrix:

$$\mathbf{J}_{\mathrm{GP}} = \begin{bmatrix} 1 & 0 \\ 0 & \exp(i\delta_{\mathrm{GP}}) \end{bmatrix} \tag{S12}$$

where $\delta_{\mathrm{GP}} = -0.628$ rad under our conditions. Using an arrangement of polarizers and a quarter wave plate, we experimentally confirmed that the effective Jones matrix of the microscope body without the dichroic mirror is consistent with Equation (S12). However, with the dichroic mirror in place the effective phase delay of the microscope body was measured to be $\delta_{\mathrm{scope}} \approx -0.35$ rad. Such dichroic mirrors are known to be birefringent in both transmission and reflection[5], so this is not wholly surprising.

Clearly $\delta_{\mathrm{scope}}$ is much closer to 0 than it is to $\pi$. If $\delta_{\mathrm{scope}}$ were exactly 0 then one would need to apply an additional $\pi$ phase delay to effectively undo the substitution $y_{\mathrm{BFP}} \mapsto -y_{\mathrm{BFP}}$; a half wave plate with fast axis aligned with *x* would do exactly that. Thus the half wave plate inserted in our setup compensates for nearly all of the effect of reflection from the prism, except the residual 0.35 rad (≈ $\lambda/18$) phase delay. Simulations confirm that this compensation is enough to nearly recover the ideal y-phi-PSF. We were faced with similar but distinct phase errors with two different dL5 imaging configurations, as described in the Methods. For one such configuration we compensated by replacing the half wave plate with a quarter wave plate. For the other configuration we used a second dichroic as the phase compensating element. In the future, more precise correction can be applied with a Soleil-Babinet compensator.

## Gaussian estimators

Simple Gaussian-based estimators were used to fit to simulated and experimental images of the standard PSF, phi-PSF, and y-phi-PSF. For the standard PSF we used the image model:

$$\tilde{\mathrm{I}}(x,y) = \alpha_1 \exp\left[-\frac{(x-\alpha_2)^2}{2\alpha_3^2} - \frac{(y-\alpha_4)^2}{2\alpha_5^2}\right] + \alpha_6 \qquad \text{(S13)}$$

with free parameters $\alpha_1, \ldots \alpha_6$. For the phi-PSF we fit to:

$$\tilde{\mathrm{I}}^{(\mathrm{phi})}(x,y) = \beta_1 \exp\left[-\chi_1(x-\beta_5)^2 + 2\chi_2(x-\beta_5)(y-\beta_6) - \chi_3(y-\beta_6)^2\right] + \beta_7 \qquad \text{(S14)}$$

with $\chi_1 = \frac{\sin^2\beta_2}{2\beta_3^2} + \frac{\cos^2\beta_2}{2\beta_4^2}$, $\chi_2 = -\frac{\sin 2\beta_2}{4\beta_3^2} + \frac{\sin 2\beta_2}{4\beta_4^2}$, and $\chi_3 = \frac{\sin^2\beta_2}{2\beta_3^2} + \frac{\cos^2\beta_2}{2\beta_4^2}$, where $\beta_1, \ldots, \beta_7$ are the free parameters. This formulation expresses the angle of the major axis of the ellipse explicitly ($\beta_2$). For the y-phi-PSF we fit to:

$$\tilde{\mathrm{I}}^{(\text{y-phi})}(x,y) = \gamma_1 \exp\left[-\frac{(x-\gamma_2)^2}{2\gamma_3^2} - \frac{(y-\gamma_4)^2}{2\gamma_5^2}\right] + \gamma_6 \exp\left[-\frac{(x-\gamma_7)^2}{2\gamma_3^2} - \frac{(y-\gamma_8)^2}{2\gamma_5^2}\right] + \gamma_9 \qquad \text{(S15)}$$

with free parameters $\gamma_1,\cdots\gamma_9$. This is certainly not the only reasonable choice for Gaussian estimator, but it was chosen since it is the one most commonly used for fitting the Double-Helix PSF (DH-PSF)[6], which very closely resembles the y-phi-PSF except that it rotates with $z_\mathrm{D}$ rather than in-pane orientation. For simulated images we fixed $\gamma_6 = \gamma_1$ due to symmetry considerations, but for experimental images they were allowed to be unequal in order to compensate for residual aberrations (see Supplementary Fig. 10 and the section below on residual lobe asymmetry).

## Correcting sample drift

A small amount of sample drift was detected and corrected for in the localizations depicted in Fig. 4 and Supplementary Fig. 4. To correct this we identified 10-20 molecules in each shared FOV that did not appear to shift along the direction of its orientation by a significant amount (i.e. <10-20 nm). These molecules presumably have $\theta_\mathrm{D} \approx 90^\circ$ (an abundance of such molecules were found in the data, but in this paper we selected more interesting molecules with $\theta_\mathrm{D} < 90^\circ$). Any residual motion of these in-plane molecules was assumed to be due to lateral drift of the stage. We averaged this motion across all selected fiducial molecules in each FOV, then subtracted this drift from the localizations of the molecules in Fig. 4 and Supplementary Fig. 4. For dL5 complexes 1 and 2 shown in Fig. 5 we determined the stage drift was small enough to ignore.

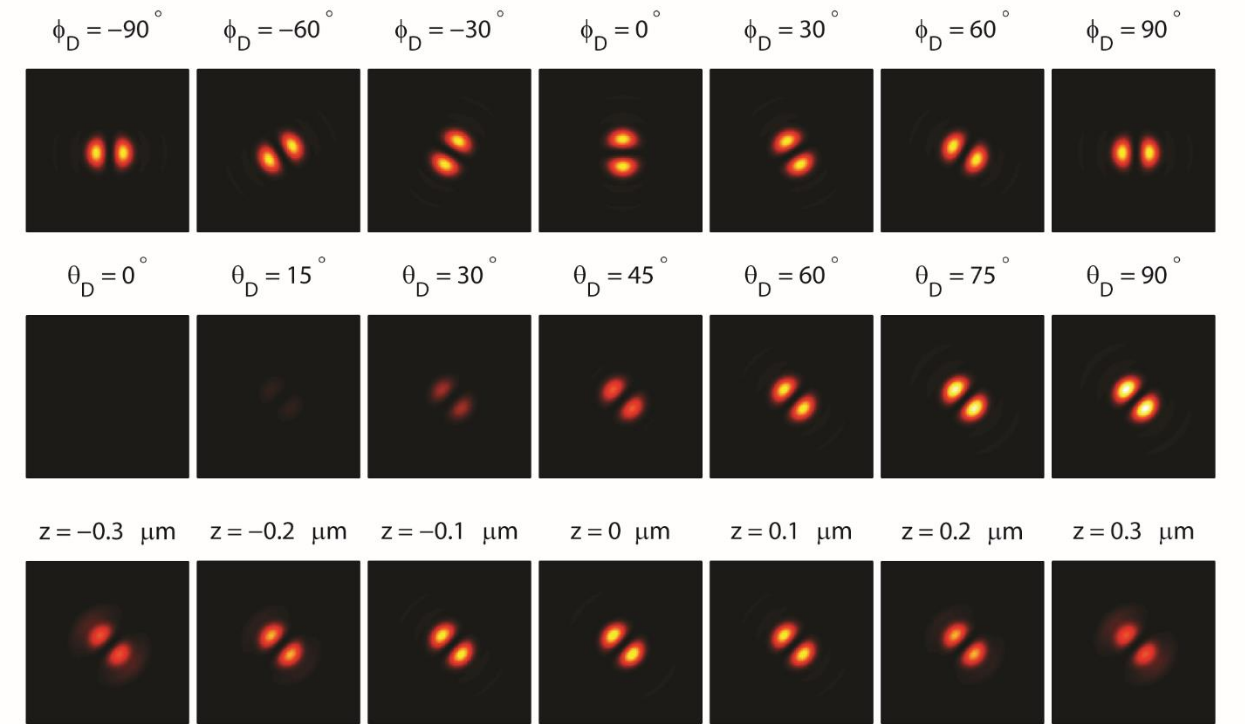
## Residual lobe asymmetry

The y-phi-PSFs realized in our experimental data tended to show a slight lobe asymmetry, wherein the top lobe was somewhat brighter than the bottom lobe (e.g. see Fig. 4a). This might be a result of residual phase and/or amplitude aberrations present in the microscope itself. In particular, even high-quality objectives are known to often have significant aberrations[7-9], which can be problematic for some applications. Supplementary Fig. 10 shows that a small amount of vertical coma phase aberration produces simulated images that appear qualitatively consistent with this asymmetry. This effect was simulated by multiplying $\mathbf{E}_{\mathrm{BFP}}(\rho,\varphi)$ by the factor $\exp\left[ic\,\mathrm{Z}_3^{-1}(\rho,\varphi)\right]$, where

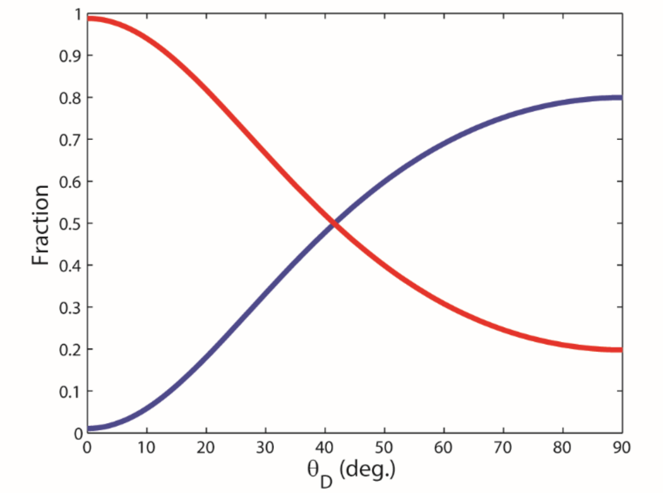
$$\mathrm{Z}_3^{-1}(\rho,\varphi)=\sqrt{8}\left(3\rho^3-2\rho\right)\sin\varphi \qquad \text{(S16)}$$

is the normalized Zernike polynomial corresponding to vertical coma[10], and *c* is a constant. For Supplementary Fig. 10 we heuristically tuned to $c = 0.2$.

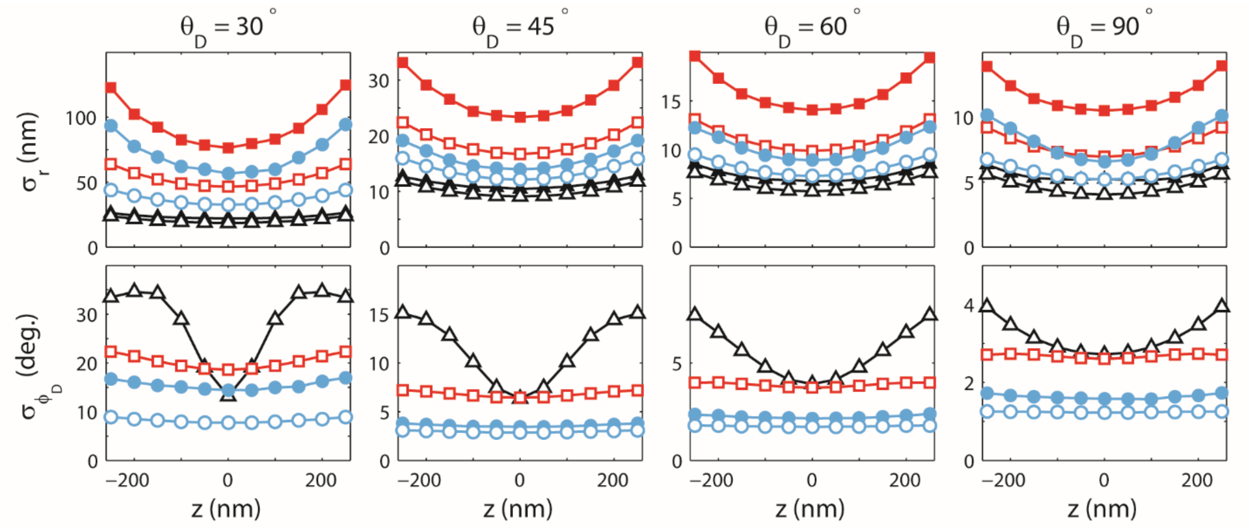
By fitting to the sum of two Gaussians with unequal intensity rather than localizing with the centroid we avoided introducing strong additional localization errors, as evidenced by the successful results in Fig. 4c. However, it is possible that future applications will require more careful consideration and correction of these aberrations.



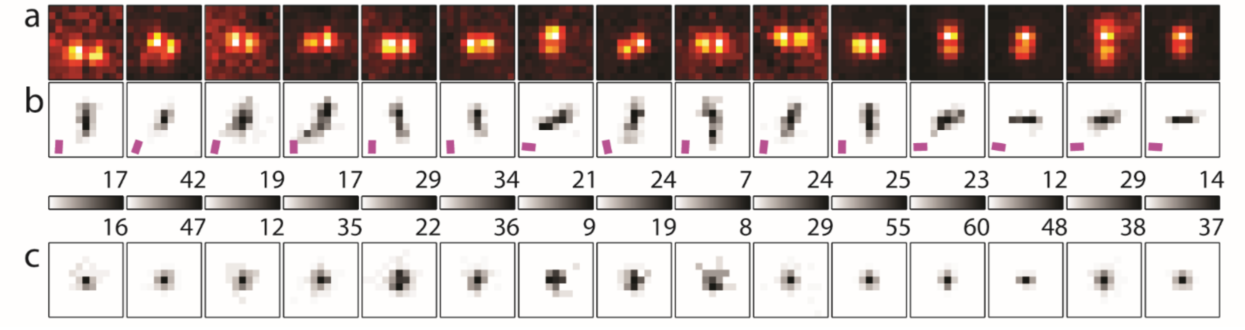
**Supplementary Figure 1 | Simulated behavior of y-phi-PSF of a molecule in matched media.** Index of refraction of sample and immersion medium taken to match immersion oil (*n* = 1.518). High-resolution images show behavior of the y-phi-PSF as a function of $\phi_{\mathrm{D}}$ (top row), $\theta_{\mathrm{D}}$ (middle), and $z_{\mathrm{D}}$ (bottom). In the top row $\theta_{\mathrm{D}} = 60^{\circ}$ and $z_{\mathrm{D}} = 100$ nm; in the middle $\phi_{\mathrm{D}} = 45^{\circ}$ and $z_{\mathrm{D}} = 100$ nm; in the bottom row $\phi_{\mathrm{D}} = 45^{\circ}$ and $\theta_{\mathrm{D}} = 60^{\circ}$. Length of each panel side = 2.24 μm.



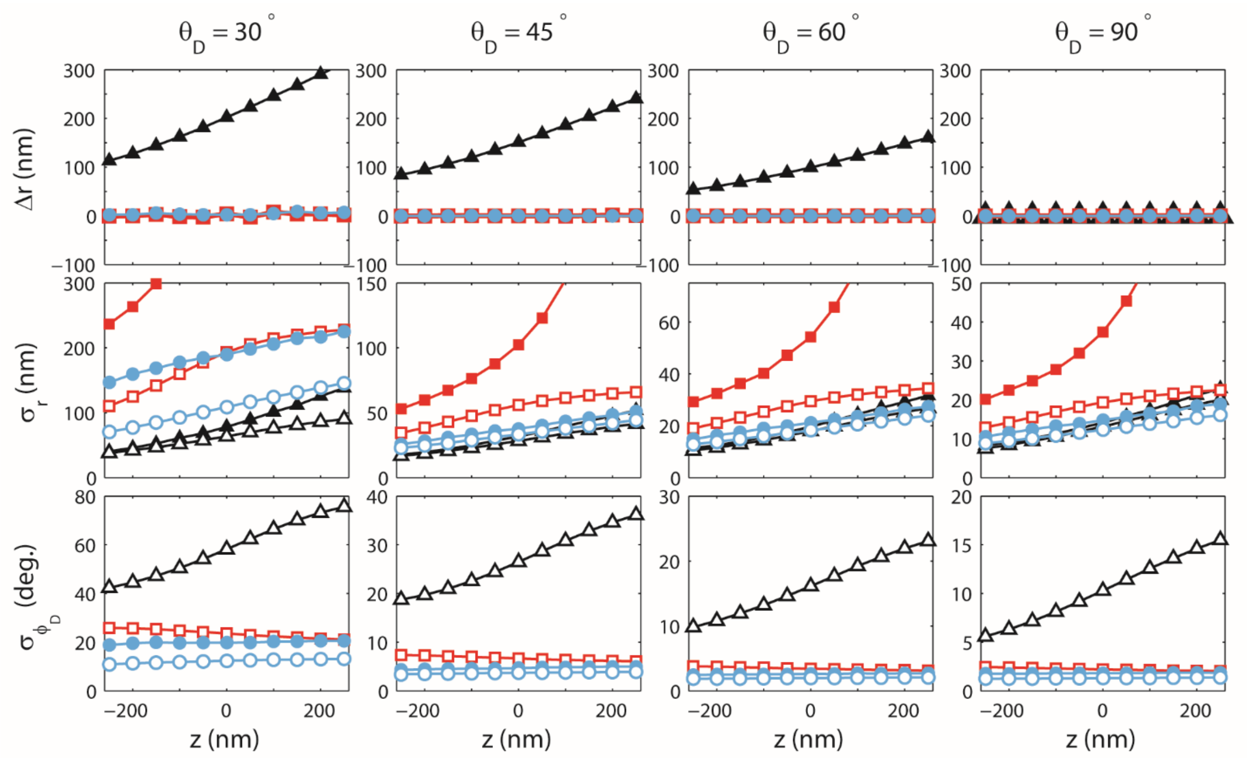
**Supplementary Figure 2 | Simulated fraction of photons in radial (red) and azimuthal (blue) polarization channels, as a function of polar dipole angle.** Simulations based on NA = 1.4, oil immersion, matched media.



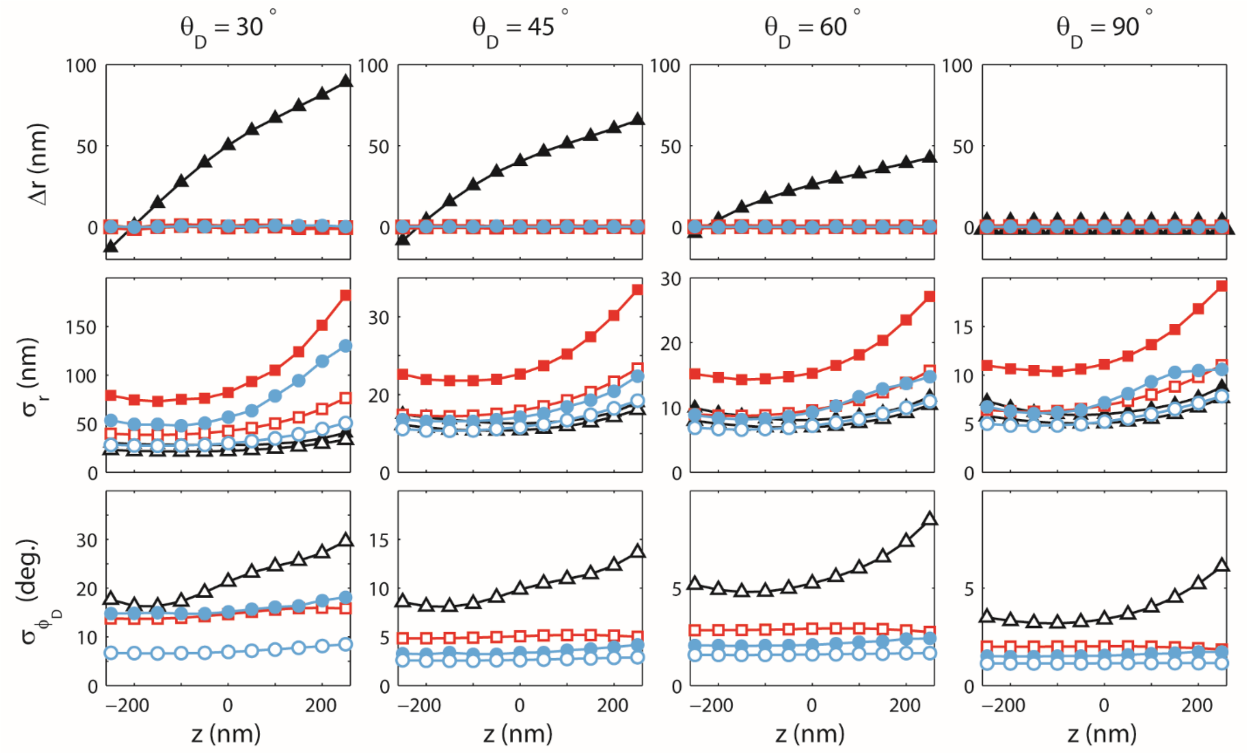
**Supplementary Figure 3 | Simulated precision and CRLB calculation.** Signal photon levels were chosen such that standard PSF image of a molecule with $\theta_{\mathrm{D}} = 90^{\circ}$ gives an average of 2000 detected photons; signal level was then scaled for pumping/detection efficiencies and polarization. Mean background = 10 photons/pixel for unpolarized (5 for polarized). Pixel size in images = 160 nm. Filled markers correspond to standard deviation from fitting $10^4$ noisy images with Gaussian-based estimators. (Top row) Lateral localization precision $\sigma_r = \sqrt{\sigma_x^2 + \sigma_y^2}$ . (Bottom row) Precision in estimation of in-plane orientation $\sigma_{\phi_{\mathrm{D}}}$ . Only the y-phi-PSF gave reliable Gaussian-based estimations of $\phi_{\mathrm{D}}$ and so the other two cases are not plotted. In all panels the unfilled markers correspond to the CRLB.



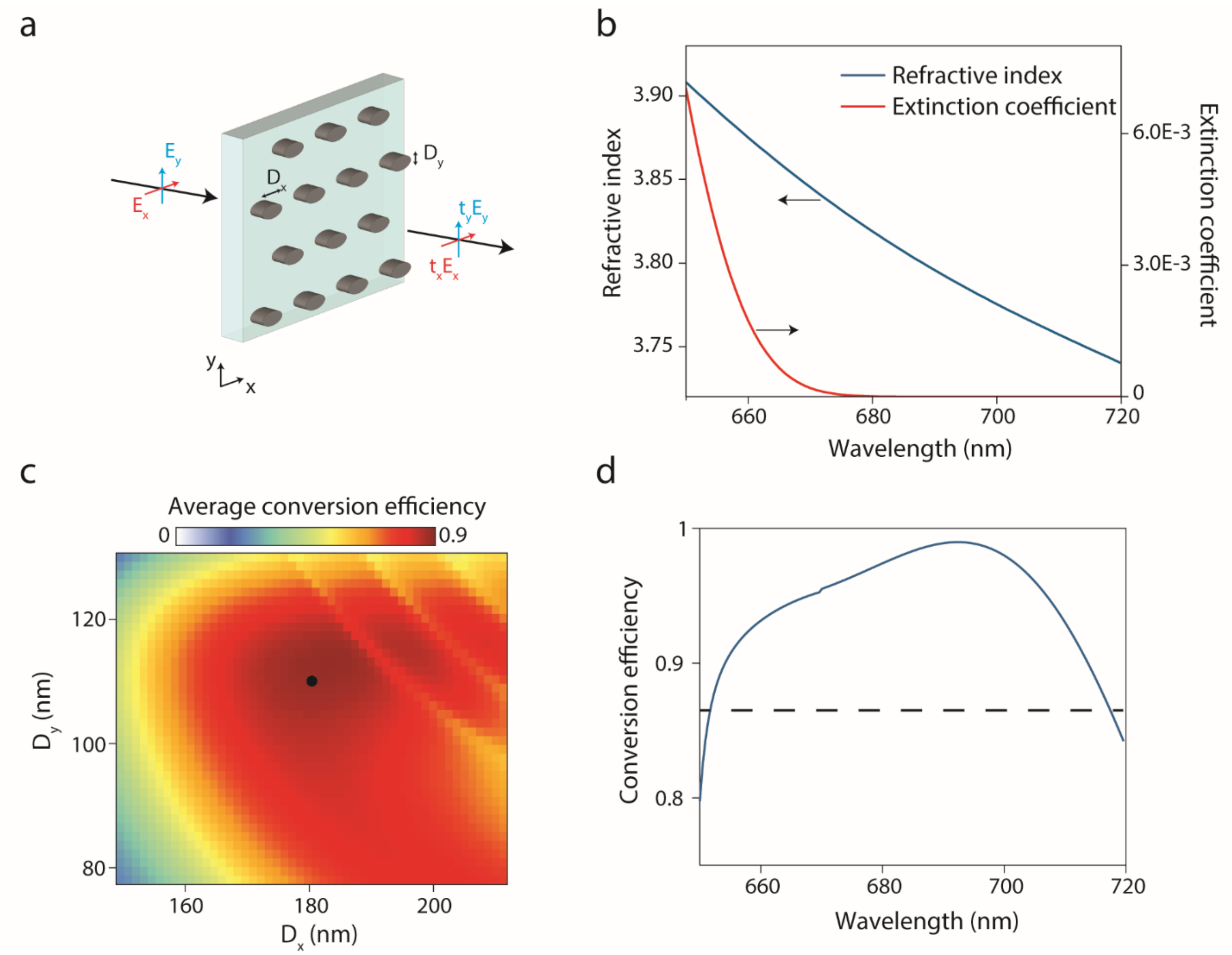
**Supplementary Figure 4 | Additional molecules demonstrating correction of orientation-induced mislocalization.** Panels are analogous to those in Fig. 4a-c. Here we show 15 more DCDHF-A-6 molecules (in addition to the 10 depicted in Fig. 4) that demonstrated our ability to effectively remove orientation-induced localization bias. **a,** in-focus y-phi images of these 15 molecules (pixel size = 160 nm). **b,** lateral localization histograms obtained with standard PSF (bin size = 20 nm). Magenta bar indicates mean $\phi_D$ as estimated from y-phi images. **c,** lateral localization histograms obtained with y-phi-PSF (bin size = 20 nm).



**Supplementary Figure 5 | Simulation results for molecule in water, 1 μm above water-glass interface.** (Top row) panels are the same as the ones in Fig. 1l-o except here we simulated molecules at a depth $d = 1$ μm above a water-glass interface. Standard PSF (black triangles), phi-PSF (red squares), and y-phi-PSF (blue circles). The index mismatch causes mislocalization for the standard PSF even at $z = 0$. Here the horizontal axis corresponds to moving the objective rather than moving the molecule. We define $z = 0$ as the objective position at which the nominal focal plane (i.e. the focal plane when the sample index is equal to the immersion index) is $^{1}\!/_{0.7}$ μm above the interface. The factor 0.7 is consistent with the focal shift that is known to occur when imaging into a water-glass mismatch[11]. Middle and bottom rows depict precision as determined by simulation (filled) and CRLB calculation (unfilled). Markers not visible in the plot are out of the relevant range.

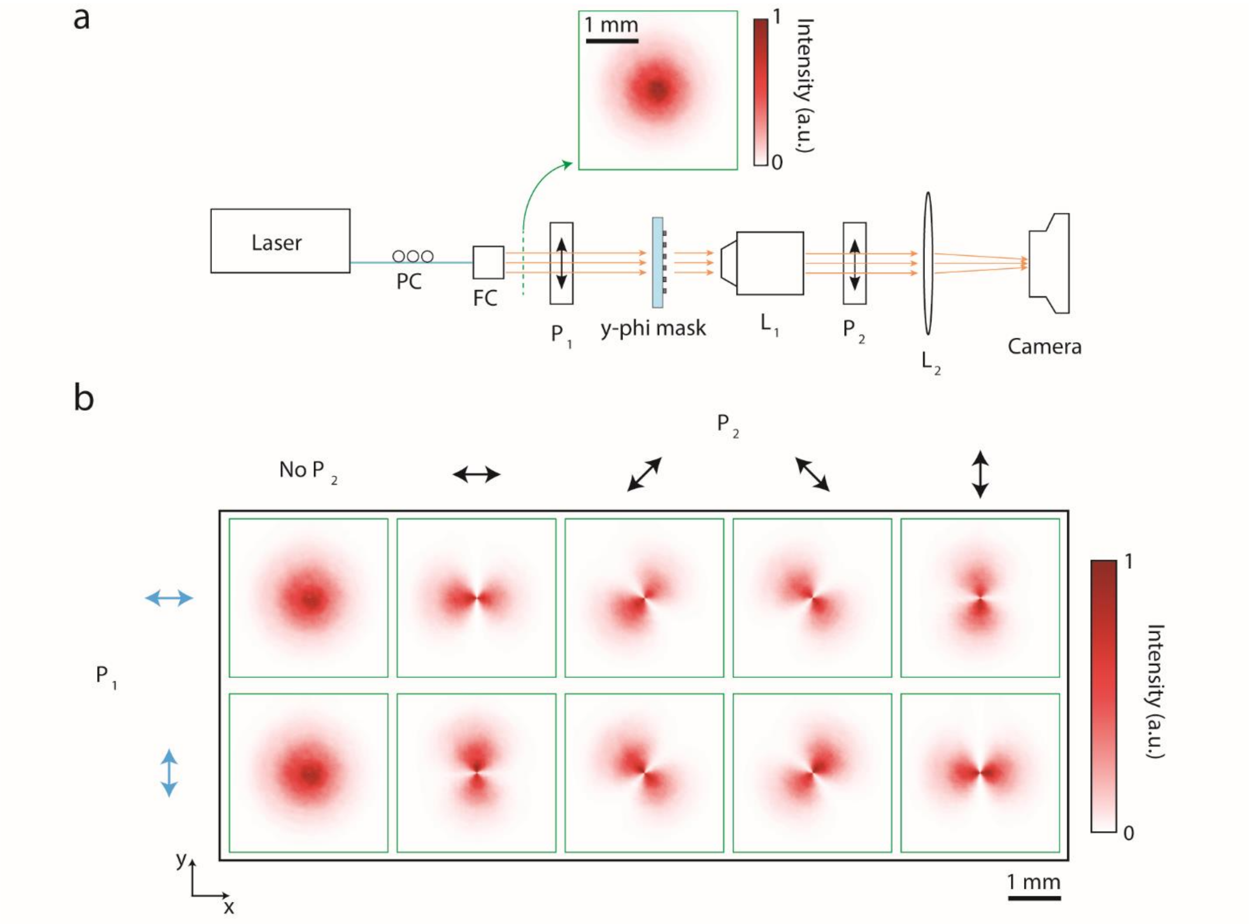


**Supplementary Figure 6 | Analog of Supplementary Fig. 5 simulated for molecule beneath thin layer of PMMA.** Panels are the same as the ones in Supplementary Fig. 5, except here we simulated molecules at the interface between glass (*n* = 1.518) and PMMA (*n* = 1.49), 30 nm beneath a PMMA-air interface. Standard PSF (black triangles), phi-PSF (red squares), and y-phi-PSF (blue circles). Note that the index mismatch causes mislocalization for the standard PSF even at *z* = 0. Here the horizontal axis corresponds to moving the objective rather than moving the molecule. As before, middle and bottom rows depict precision as determined by simulation (filled) and CRLB calculation (unfilled). Markers not visible in the plot are out of the relevant range.

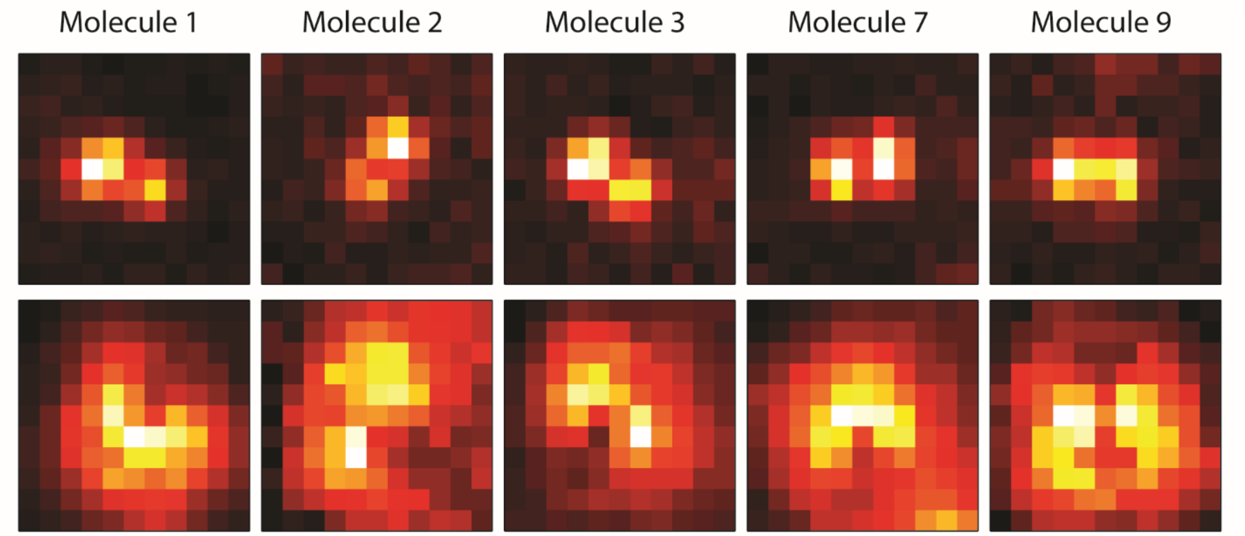


**Supplementary Figure 7 | Optimum design of the y-phi mask. a**, Schematic illustration of a uniform array of elliptical-cross-section a-Si nanoposts on glass. The nanoposts are arranged on a hexagonal lattice, the ellipse axes are along the $x$ and $y$ directions, and the array transmission coefficients are represented by $t_x$ and $t_y$ for the $x$- and $y$-polarized light, respectively. **b**, Measured wavelength dependent values of the a-Si refractive index and extinction coefficient. **c**, Simulated weighted average of the phi-to-y conversion efficiency for the array shown in **a** as a function of the diameters of the elliptical posts ($D_x$ and $D_y$). The post height of $h$ = 365 nm and lattice constant of $l$ = 325 nm are assumed. The black dot represents the diameter values for the selected design. **d**, Simulated phi-to-y conversion efficiency of the selected design ($D_x$ = 180 nm,

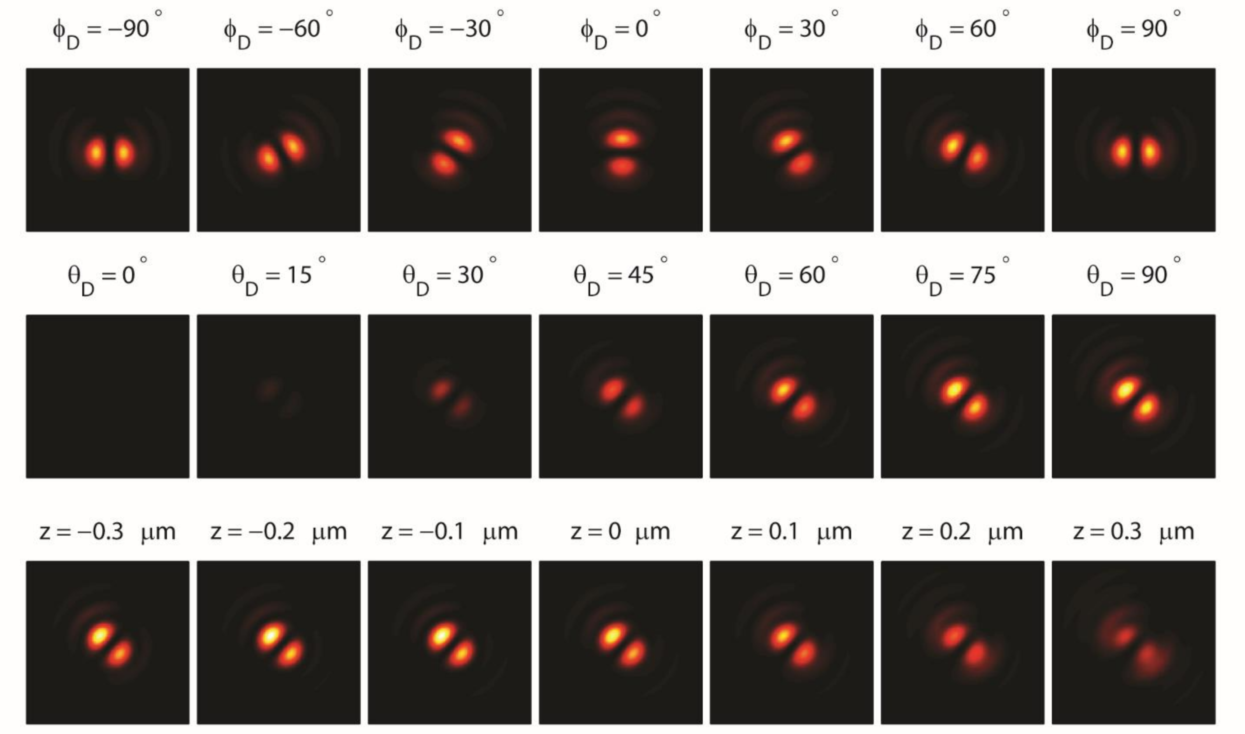
$D_y$ = 110 nm, $h$ = 365 nm, and $l$ = 325 nm) as a function of wavelength (solid blue). Average experimental transmission efficiency (black dashed).



**Supplementary Figure 8 | Characterization of the y-phi mask. a**, Schematic of the experimental setup used for characterization of the y-phi mask. The inset shows the intensity profile of the collimated beam incident on the y-phi mask. PC: polarization controller, P: polarizer, L: lens. **b**, Normalized intensity profiles captured by the camera for the *x* and *y* input polarizations which are represented by the blue double headed arrows which indicate the transmission axis of polarizer $P_1$, and four different polarization projection directions represented by the black double headed arrows which indicate the transmission axis of polarizer $P_2$. The intensity profiles of the transmitted light with $P_2$ removed from the setup are also shown. The projected intensity profiles confirm the radial and azimuthal polarizations of the transmitted beams when the input beams are *x*- and *y*-polarized, respectively.



**Supplementary Figure 9 | Visual comparison of y-phi images to defocused clear-aperture.** Top row shows in-focus y-phi images of 5 of the 10 molecules depicted in Fig. 4, while the bottom row shows the images obtained by removing the y-phi mask, linear polarizer, and half wave plate and defocusing the objective toward the sample by 1 μm. Defocused images are the average of 70 300-ms acquisitions, while y-phi images are the average of 10 300-ms acquisitions. Notably, the defocused and y-phi images share a line of symmetry along the direction of $\phi_D$ , as expected from simulations. The defocused images have an additional asymmetry due to the out-of-plane component of the orientation.



**Supplementary Figure 10 | Simulation of y-phi-PSF of molecule beneath thin layer of PMMA, with comatic phase aberration.** Vertical coma was simulated as described in the text of the Supplementary Information. High-resolution images show behavior of the y-phi-PSF as a function of $\phi_D$ (top row), $\theta_D$ (middle), and *z* (bottom). In the top row $\theta_D = 60^\circ$ and *z* = 100 nm; in the middle $\phi_D = 45^\circ$ and *z* = 100 nm; in the bottom row $\phi_D = 45^\circ$ and $\theta_D = 60^\circ$. Length of each panel side = 2.24 µm.

| Molecule | Clear Aperture | | y-phi | |
|---|---|---|---|---|
| | Signal | Background | Signal | Background |
| 1 | 17352 | 253 | 3999 | 79 |
| 2 | 1670 | 178 | 317 | 60 |
| 3 | 5585 | 196 | 2220 | 63 |
| 4 | 2104 | 275 | 1547 | 92 |
| 5 | 4468 | 246 | 1790 | 81 |
| 6 | 3129 | 354 | 676 | 123 |
| 7 | 5618 | 378 | 1623 | 134 |
| 8 | 5034 | 197 | 1954 | 69 |
| 9 | 8542 | 320 | 1766 | 115 |
| 10 | 1409 | 198 | 642 | 69 |

**Supplementary Table 1 | Mean photon numbers for molecules depicted in Fig. 4.** Numbers are mean photons detected per frame as averaged over each set of acquisitions of the molecules. Signal photons are per frame and background photons are per frame per pixel. Note that clear-aperture images were acquired with an EM gain setting of 100, while y-phi images were acquired with EM gain of 300 due to generally weaker signals, but the table still shows detected photons. Background was estimated by computing the mean in a hand-selected rectangle near each molecule. For these ten molecules, dividing the y-phi signal photons by the clear aperture signal photons, then dividing again by the mask efficiency 0.865 gives azimuthally polarized photon fractions that range between 0.22 and 0.85. These values are consistent with the range plotted in Supplementary Fig. 2 for various $\theta_D$. Note that the distribution here is skewed below 0.5 because we selected for molecules with larger shifts and thus smaller $\theta_D$. In addition, there are losses associated with the linear polarizer and half wave plate, which were present for the y-phi measurement but not the clear aperture.

## References

1. Novotny, L. & Hecht, B. in *Principles of Nano-Optics* (Cambridge University Press, New York, 2007).

2. Backer, A. S. & Moerner, W. E. Extending Single-Molecule Microscopy Using Optical Fourier Processing. *J. Phys. Chem. B* **118**, 8313-8329 (2014).

3. Boggess, A. & Narcowich, F. J. in *A first course in wavelets with Fourier analysis* (John Wiley & Sons, Hoboken, New Jersey, 2009).

4. Lew, M. D. & Moerner, W. E. Azimuthal Polarization Filtering for Accurate, Precise, and Robust Single-Molecule Localization Microscopy. *Nano Lett.* **14**, 6407-6413 (2014).

5. Tang, Y., Cook, T. A. & Cohen, A. E. Limits on fluorescence detected circular dichroism of single helicene molecules. *The Journal of Physical Chemistry A* **113**, 6213-6216 (2009).

6. Pavani, S. R. P. *et al*. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 2995-2999 (2009).

7. Backlund, M. P. *et al*. Simultaneous, accurate measurement of the 3D position and orientation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 19087-19092 (2012).

8. Blythe, K. L., Titus, E. J. & Willets, K. A. Objective-induced Point Spread Function Aberrations and their Impact on Super-resolution Microscopy. *Anal. Chem.* **87**, 6419-6424 (2015).

9. Kurvits, J. A., Jiang, M. & Zia, R. Comparative analysis of imaging configurations and objectives for Fourier microscopy. *JOSA A* **32**, 2082-2092 (2015).

10. Mahajan, V. N. in *Optical Shop Testing* (ed Malacara, D.) 498-546 (Wiley, Online edition, 2007).

11. Hell, S. W., Reiner, G., Cremer, C. & Stelzer, E. H. K. Aberrations in confocal fluorescence microscopy induced by mismatches in refractive index. *J. Microsc.* **169**, 391-405 (1993).