S1. Additional AFM scratch test results of very thin PMMA layers

The additional AFM scratch test results for PMMA layer thicknesses down to $d_1=d_2=5$ and 2.5 nm are provided in Figs. S1 and S2, respectively. It is seen that in the whole field of view the PMMA film surface is quite flat (with a height fluctuation below 0.5 nm), and the PMMA film thicknesses in different positions of the scratch remain the designated values ($d_1=5$ nm and $d=10$ nm for Fig. S1, $d_1=2.5$ nm and $d=5$ nm for Fig. S2). These results further confirm that the formed PMMA film has a uniform and controlled thickness even for very thin film with thickness down to several nanometers.
Fig. S1 Scratch test results of the double PMMA layers with $d_1=5$ nm and $d=10$ nm. (a) AFM image of the scratched PMMA layer on the gold film after the first spin coating. Scale bar: 2 μm. The position of 0 height (PMMA surface) is marked with the black triangle. (b) Cross-section profiles along the two probing paths marked with blue and green lines in (a). (c)-(d) The same as (a)-(b) but for the scratched PMMA layers after the second spin coating. The values of $d_1$ and $d$ are determined by the difference of the two heights at positions marked by the red triangles in (b) and (d).

Fig. S2 The same as Fig. S1 but for PMMA layer thicknesses down to $d_1=2.5$ nm and $d=5$ nm.
**S2. Absorption cross-section spectrum of the gold NR**

![Normalized absorption cross-section](image)

**Fig. S3.** Normalized absorption cross-section of the gold NR as a function of the wavelength for different nano-gap sizes $d=$10 (blue solid curve), 30 (red) and 50 nm (black). The vertical dashed line marks the fluorescence excitation wavelength of 640 nm.

We calculate the normalized absorption cross-section of the gold NR versus the wavelength for different nano-gap sizes $d=$10, 30 and 50 nm, as shown in Fig. S3. The absorption cross-section is calculated with the COMSOL software for a normally incident $x$-polarized plane wave, and is normalized by the geometrical cross-section of the NR (i.e. the projection area of the NR along the propagation $z$-direction of the plane wave). The vertical dashed line in Fig. S3 marks the fluorescence excitation wavelength of 640 nm, which is close to the absorption peak of the gold NR.
S3. SEM images of the NRs dispersed on the top surface of the double PMMA layers

To achieve a good dispersity of gold NRs for the sample, the ultrapure water suspension of NRs is diluted to $1 \times 10^{-10}$ M and is then spin coated on the double PMMA layers. To check the density of NRs on the top surface of the double PMMA layers, we take SEM images of the sample (with PMMA layer thicknesses $d_1 = d_2 = 10$ nm) on a Sigma500 electron microscope. As shown in Fig. S4, the SEM images exhibit a good dispersity of individual NRs, and the density of NRs on the top surface of the double PMMA layers is about $10/(100 \ \mu \text{m}^2)$. In view that the size of the gold NRs is less than 0.1 $\mu$ m, the possibility of the fluorophores being located within the NR-substrate nanogap created by two or more aggregated NRs is extremely low.

Fig. S4 (a) SEM image of the gold NRs dispersed on the top surface of the double PMMA layers (with thicknesses $d_1 = d_2 = 10$ nm). Scale bar: 1 $\mu$ m. (b)-(d) Enlarged SEM images of the three blue squares 1, 2 and 3 in (a), respectively. The scale bars are 300, 150 and 300 nm, respectively.
S4. IRF of the TCSPC system

![Image of IRF measurement](image)

**Fig. S5.** Measured IRF with the ludox scatterer, shown with a logarithmic scale (a) and a linear scale (b), respectively.

The instrument response function (IRF) of the TCSPC system is obtained by measuring the scattered light of diluted Ludox AS-30 colloidal silica solution, as shown in Fig. S5a or S5b with a logarithmic or linear scale, respectively. The full width of half maximum (FWHM) of the IRF is shown to be 39 ps, which is shorter than the lifetimes measured in the present work and thus can ensure the measurement precision.

S5. Measurement showing the cut off of Rayleigh scattering light and Raman scattering light

We measure the time trace curves of the structure in Fig. 3 ($d_1=d_2=10$ nm) under four different conditions: a) At the position of NR with NB molecules (i.e. the red curve in Fig. 3c); b) At the position of NR but without NB molecules; c) In the absence of both NR and NB molecules; d) Without the excitation of the 640 nm wavelength pulsed laser (i.e. the dark count $I_{\text{dark}}$). Fig. S6 shows that the time-averaged intensities under conditions b), c) and d) are 0.112, 0.109 and 0.110 counts/ms, respectively, which are quite close to each other. Then we can conclude that the Rayleigh scattering light is totally cut off in our measurement of fluorescence.
Fig. S6. Measured time traces of the structure in Fig. 3 ($d_1=d_2=10\text{nm}$) under four different conditions: (a) At the position of NR with NB molecules; (b) At the position of NR but without NB molecules; (c) In the absence of both NR and NB molecules; (d) Without the excitation of the 640 nm wavelength pulsed laser (dark counts). The traces in (b)-(d) are multiplied by 5.

Fig. S7. (a) Raman scattering spectroscopy of NB molecules measured with an excitation of 633 nm wavelength laser. Blue curve: result for the sample in Fig. S6a. Red curve: enhanced Raman scattering spectroscopy with gold nano-spheres. The two black dashed lines mark the transmission wavelength range of the emission beam filter in our TCSPC system. (b) The same as the red curve in (a) but shown as a function of the Raman shift.
We also measure the Raman scattering spectroscopy of the sample of Fig. S6a, as shown by the blue curve in Fig. S7a. The excitation wavelength is 633 nm that is close to the excitation wavelength 640 nm of Fig. S6a. One can observe some tiny peaks between 650 and 675 nm on the strong background of the fluorescence. To confirm that the tiny peaks are Raman scattering signal, we measure the Raman scattering spectroscopy of NB molecules enhanced by gold nano-spheres, as shown by the red curve in Fig. S7a. For the measurement the gold nano-spheres (with an average diameter of 180 nm) are distributed on a silicon wafer by spin coating of an ultrapure water suspension of gold nano-spheres. Then a droplet of ethanol dissolved NB molecules (concentration: $1 \times 10^{-6}$ M) is dropped on the silicon wafer and dried naturally. It is seen that the enhanced Raman peaks at 657.7 and 660.8 nm in the red curve are exactly located at the two tiny peaks of the blue curve, confirming that the tiny peaks are due to the Raman scattering. Fig. S7b plots the enhanced Raman spectrum as a function of Raman shift, showing that the two Raman peaks correspond to the Nile Blue Raman shifts 590 cm$^{-1}$ and 662 cm$^{-1}$, which is consistent with literatures.$^{1,2}$ From the above measurement, we finally conclude that the Raman scattering signal is much weaker than the fluorescence signal and thus can be neglected. Furthermore, the transmission wavelength range of the emission beam filter in our TCSPC system is 665-732 nm (as marked in Fig. S7a by the two black dashed lines), so that the major Raman signals at 657.7 and 660.8 nm cannot pass through the emission beam filter to reach the SPAD detector.
S6. Fluorescence decay curves and fitted parameters for the representative values of \( d \) and \( d_1 \) in Tables 1 and 2

For the representative values of \( d \) and \( d_1 \) in Table 1 (\( d=5, 10, 30, 50 \) nm and \( d_1=d_2 \)) and in Table 2 (\( d_1=2.5, 5, 20, 35, 37.5 \) nm and \( d=40 \) nm), all the fluorescence decay curves are shown in Fig. S8. The corresponding values of fitted parameters \( C, C_1, C_2, \gamma_0, \gamma_1, \gamma_2 \) are also provided in Table S1.

Fig. S8. (a1-4) Fluorescence decay curves for representative values of \( d=5, 10, 30, 50 \) nm and \( d_1=d_2 \) (corresponding to the data in Table 1). (b1-5) Fluorescence decay curves for representative values of \( d_1=2.5, 5, 20, 35, 37.5 \) nm and \( d=40 \) nm (corresponding to the data in Table 2). Five different NRs are measured for each set of values of \( d \) and \( d_1 \). For a clear display, all the decay curves are normalized to be unity at zero time, and then are vertically shifted in the logarithmic scale after multiplied with \( 5^0 \) (red curve), \( 5^1 \) (green), \( 5^2 \) (purple), \( 5^3 \) (black) and \( 5^4 \) (yellow) from bottom to top, respectively. The five curves from top to bottom sequentially correspond to the five measured data of \( f_F \) in Tables 1 and 2. The blue curves are the results in the absence of the NR, multiplied by \( 5^4 \). The fitted slow and fast exponential decay rates \( \gamma_1 \) and \( \gamma_2 \) of the red curve are
shown with black dashed lines. The fitted rate $\gamma_0$ in the absence of the NR is shown with the black solid line.

Table S1. Fitted amplitudes ($C, C_1, C_2$) and decay rates ($\gamma_0, \gamma_1, \gamma_2$) for the data shown in Fig. S8. As stated in the main text, $C \exp(-\gamma_0 t)$ and $C_1 \exp(-\gamma_1 t) + C_2 \exp(-\gamma_2 t)$ are used to fit the fluorescence decay curves without and with the presence of the NR, respectively, where $C_1 \exp(-\gamma_1 t)$ and $C_2 \exp(-\gamma_2 t)$ describe the slow and fast decay components, respectively. The five fitted values of $C_1$, $C_2$, $\gamma_1$ and $\gamma_2$ sequentially correspond to the five curves from top to bottom in each individual figure of Fig. S8.

<table>
<thead>
<tr>
<th>Gap size</th>
<th>C</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$1/\gamma_0$ (ns)</th>
<th>$1/\gamma_1$ (ns)</th>
<th>$1/\gamma_2$ (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d=5\text{nm}$, $d_1=d_2$</td>
<td>33.2</td>
<td>32.8</td>
<td>47.9</td>
<td>354.1, 333.0</td>
<td>2.55</td>
<td>2.74, 2.84, 0.0522, 0.0496,</td>
</tr>
<tr>
<td>$d=10\text{nm}$, $d_1=d_2$</td>
<td>52.7</td>
<td>55.7</td>
<td>61.1</td>
<td>521.3, 576.1</td>
<td>2.43</td>
<td>2.52, 2.42, 0.0587, 0.0556,</td>
</tr>
<tr>
<td>$d=30\text{nm}$, $d_1=d_2$</td>
<td>72.1</td>
<td>80.1</td>
<td>82.5</td>
<td>417.6, 416.6</td>
<td>2.67</td>
<td>2.74, 2.67, 0.329, 0.351,</td>
</tr>
<tr>
<td>$d=50\text{nm}$, $d_1=d_2$</td>
<td>65.4</td>
<td>99.9</td>
<td>85.2</td>
<td>300.4, 323.7</td>
<td>2.88</td>
<td>2.92, 2.94, 0.951, 1.00,</td>
</tr>
<tr>
<td>$d=40\text{nm}$, $d_1=2.5\text{nm}$</td>
<td>35.7</td>
<td>52.2</td>
<td>41.9</td>
<td>233.2, 256.3</td>
<td>2.67</td>
<td>2.85, 2.63, 0.0603, 0.074,</td>
</tr>
<tr>
<td>$d=40\text{nm}$, $d_1=5\text{nm}$</td>
<td>70.2</td>
<td>122.5</td>
<td>68.2</td>
<td>897.2, 556.9</td>
<td>2.45</td>
<td>2.50, 2.54, 0.103, 0.124,</td>
</tr>
<tr>
<td>$d=40\text{nm}$, $d_1=20\text{nm}$</td>
<td>71.0</td>
<td>94.6</td>
<td>82.5</td>
<td>423.6, 303.3</td>
<td>2.70</td>
<td>2.70, 2.68, 0.651, 0.619,</td>
</tr>
<tr>
<td>$d=40\text{nm}$, $d_1=35\text{nm}$</td>
<td>53.3</td>
<td>55.6</td>
<td>49.9</td>
<td>558.3, 478.9</td>
<td>2.47</td>
<td>2.47, 2.52, 0.0750, 0.0820,</td>
</tr>
<tr>
<td>$d=40\text{nm}$, $d_1=37.5\text{nm}$</td>
<td>29.9</td>
<td>28.9</td>
<td>35.5</td>
<td>250.4, 267.8</td>
<td>2.61</td>
<td>2.52, 2.47, 0.0651, 0.0549,</td>
</tr>
<tr>
<td>$d=40\text{nm}$, $d_1=40\text{nm}$</td>
<td>33.3</td>
<td>30.2</td>
<td>200.1, 222.2</td>
<td>2.52, 2.51, 0.0581, 0.0560,</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
S7. Experimental results for the cases that the emitters contact directly with the gold NRs or the gold film

We also investigate the three cases that the emitters contact directly with the gold NRs or the gold film: 1) $d_1=d_2=0$ (no PMMA layer, emitters contacting with both the NR and the gold film); 2) $d_1=0$, $d=40\text{nm}$ (single PMMA layer, emitters contacting with the gold film); 3) $d_2=0$, $d=40\text{nm}$ (single PMMA layer, emitters contacting with the gold NR). For these cases, the measured fluorescence decay curves (Fig. S10c) exhibit no fast decaying component and no enhanced fluorescence intensity is detected, which is attributed to a complete quenching of the fluorescence signal at the gold NR.

For more details, the gold NRs can no longer be located with the fluorescence images as complete quenching occurs. For such cases, the fluorescence images (Fig. S10b) exhibit a homogeneous fluorescence background and no bright spots at NRs (like Fig. 3a in the main text) can be observed. To locate the positions of the gold NRs as complete quenching occurs, we take the dark-field microscope images of the NRs under a wide-field oblique-incidence illumination with the pulsed laser (see Fig. S9 for the optical setup). Then the positions of the NRs with bright scattered-light spots can be determined, as marked with the red squares in Fig. S10a. At last, we shelter the obliquely incident laser beam such that the NB molecules are excited by the normally incident focused laser spot of the objective, and measure the fluorescence decay curves at the predetermined positions of the NRs (the red squares in Figs. S10a-b).

Here note that the wide-field oblique-incidence illumination is also used for the fluorescence imaging, for which the DM and the emission filter in Fig. S9 are used. While for the dark-field imaging, the DM is replaced by the BS and the emission filter is removed. The lens that focuses the obliquely incident laser beam at the sample has a much longer focal length than the objective, which thus can generate a much larger focal spot than the objective to ensure a wide-field illumination.

![Fig. S9 Schematic illustration of the optical setup for the dark-field microscope imaging or the wide-field fluorescence imaging under oblique-incidence illumination through the lens with a long focal length, or for fluorescence lifetime and time trace measurements under normal-incidence illumination through the objective. DM: dichroic mirror. BS: beam splitter.](image)
Fig. S10 Experimental results for the three cases that the emitters contact directly with the gold NRs or the gold film: 1) \( d_1 = d_2 = 0 \) (top row); 2) \( d_1 = 0, \ d = 40 \text{ nm} \) (central row); 3) \( d_2 = 0, \ d = 40 \text{ nm} \) (bottom row). (a1-3) Dark-field images of gold NRs taken by EMCCD with a range of 10\( \mu \text{m} \times 10\mu \text{m} \). The red squares mark the positions of gold NRs. (b1-3) Images of NB fluorescence taken by EMCCD with a range of 10\( \mu \text{m} \times 10\mu \text{m} \). (c) Fluorescence decay curves measured at the pre-determined NR positions labelled with red squares in (a) and (b). All the decay curves are normalized to be unity at zero time, and then are vertically shifted in the logarithmic scale after multiplied by \(10^0, 10^1\) and \(10^2\) from bottom to top, respectively. The fitted slow exponential decay rates \( \gamma_1 \) for the three curves are shown with black dashed lines.
S8. Typical energy level diagram for the explanation of $\Gamma_A$ and $\Gamma_B$

For the calculation of $f_F$ and $f_p$, we have defined two quantities $\Gamma_A$ (the total molecular intrinsic loss rate, see Section 3.2A) and $\Gamma_B$ (the emission-related molecular intrinsic loss rate, see Section 3.2B), respectively. For a clear explanation of $\Gamma_A$ and $\Gamma_B$ in the main text (Section 3.2B), here a typical energy level diagram (Jablonski diagram) is provided, as shown Fig. S11.

Figure S11. Typical energy level diagram (Jablonski diagram) for absorption (i.e. excitation, denoted by A in the figure), fluorescence (F), and phosphorescence (P) of molecules. Absorptions of light of higher and lower frequencies are shown with blue and orange up-pointing vertical arrows, respectively. The radiative transitions by means of fluorescence (from singlet excited state $S_2$ to singlet ground state $S_0$) and phosphorescence (from triplet excited state $T_1$ to $S_0$) are shown with down-pointing vertical arrows of red and brown colors, respectively. The non-radiative transitions include three parts: vibrational relaxation (VR, shown with purple down-pointing vertical arrows), internal conversion (IC, shown with gray wave arrows), and inter-system crossing (ISC, shown with a gray wave arrow).

References
