Cascaded emission of single photons from the biexciton in monolayered WSe_2

Yu-Ming He[1,4], Oliver Iff[1], Nils Lundt[1], Vasilij Baumann[1], Marcelo Davanco[2], Kartik Srinivasan[2], Sven Höfling[1,3] and Christian Schneider[1,*]

[1]*Technische Physik and Wilhelm-Conrad-Röntgen-Research Center for Complex Material Systems, Universität Würzburg, D-97074 Würzburg, Am Hubland, Germany.*

[2]*Center for Nanoscale Science and Technology, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA*

[3]*SUPA, School of Physics and Astronomy, University of St. Andrews, St. Andrews KY 16 9SS, United Kingdom*

[4]*Hefei National Laboratory for Physical Sciences at the Microscale and Department of Modern Physics, & CAS Center for Excellence and Synergetic Innovation Center in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

*Correspondence and requests for materials should be addressed to Christian Schneider (christian.schneider@physik.uni-wuerzburg.de)

Monolayers of transition metal dichalcogenide materials emerged as a new material class to study excitonic effects in solid state , since they benefit from enormous coulomb correlations between electrons and holes. Especially in WSe_2, sharp emission features have been observed at cryogenic temperatures, which act as single photon sources . Tight exciton localization has been assumed to induce an anharmonic excitation spectrum, however, the evidence of the hypothesis, namely the demonstration of a localized biexciton, is elusive. Here, we unambiguously demonstrate the existence of a localized biexciton in a monolayer of WSe_2, which triggers an emission cascade of single photons. The biexciton is identified by its time-resolved photoluminescence, superlinearity and distinct polarization in micro-photoluminescence experiments. We evidence the cascaded nature of the emission process in a cross-correlation experiment, which yields a strong bunching

behavior. Our work paves the way to a new generation of quantum optics experiments with two-dimensional semiconductors.

Introduction

Increasing interest and need for secure communications[1], precision measurements, metrology[2] and optical quantum emulation[3] explain the ever growing demand for non-classical sources of light. Ultra-compact implementations of such sources in solid state are particularly interesting, as they promise long-term stability, and outline possible ways for scalable integration, even in complex on chip quantum networks[4]. Regarding compactness, the ultimate limit is represented by a zero-dimensional emission center of non-classical light in atomic monolayers[5]. Transition metal dichalcogenides semiconductors have emerged as a new platform to study excitonic effects in two dimensions[6,7,8,9,10], in particular since they benefit from enormous coulomb correlations between electrons and holes as a result of reduced dielectric screening[11,12,13] and feature unique spinor properties[14]. Strong progress has been made in studying excitonic effects in two dimensional materials, including the exploration of the valley pseudospin dynamics[14,15], valley polarization control[16,17,18,19,20] and strong coupling effects[21,22]. While single photon emission from localized states in monolayers of transition metal dichalcogenides has recently been observed[23,24,25,26,27], the characteristic feature which determines an anharmonic excitation spectrum in an excitonic system, which is a well isolated biexcitonic state, has not been observed. In addition, strong spectral wandering and even blinking, induced by a supposedly noisy environment, made it debatable whether the full potential of such ultra-compact solid state single photon sources or pair sources based on mono-atomic layers can be harnessed. Recently, delocalized biexcitonic states in WSe_2 have been attributed to a superlinearly increasing, broad emission feature on the low energy side of the characteristic excitonic resonance[28]. However, these indications clearly outline the need for further unambiguous evidence of multiexcitonic complexes in monolayer materials.

Here, we demonstrate the existence of a localized biexciton in a monolayer of WSe_2 at cryogenic temperatures. The optical properties of the localized excitons are strongly improved by transferring them onto epitaxially grown semiconducting material. We routinely observe sharp emission features

with linewidth < 70 µeV by means of non-resonant photoluminescence spectroscopy. The localized exciton-biexciton pair is identified by its characteristic polarization, the characteristic power dependency, and most importantly, the emission cascade is observed in single photon correlation studies.

Results

Device description and optical characterization

Figure 1a depicts a graphic illustration of our investigated device: It consists of a 250 nm thick GaInP layer, which has been grown on a semi-insulating GaAs substrate via Gas-Source molecular beam epitaxy. A single layer of WSe_2, mechanically exfoliated via commercial adhesive tape (Tesa brand) from a bulk crystal was transferred onto the atomically smooth and chemically inert GaInP layer with a polymer stamp. We did not apply any capping technique to the monolayer, to take advantage of the light extraction from the surface allocated emitters. Figure 1b illustrates a two-dimensional scanning image of the photoluminescence emission in the energy range between 1.525eV and 1.734eV. The bulk shape could be reproduced by the suppressed photoluminescence from the substrate in the energy range 1.525-1.55eV and the localized shining spots show the sharp peaks emission, which are randomly found at the edge of the flake.

A series of non-resonant micro photoluminescence spectra recorded at varying pump powers is shown in Fig. 1c). The single quantum emitter, which was excited by a 532 nm continuous wave laser at a sample temperature of nominally 4.5K, exhibits two pronounced emission features with a distinctly different power dependency. We note, that the two emission features, which have an energetic separation of 4.6 meV, have a spectral linewidth which is limited by the resolution of our spectrometer (70 µeV) at lowest pump powers. This outlines the excellent optical properties of our localized emitters, which are not affected by obvious effects of long term spectral wandering, as commonly observed for WSe_2 emitters on insulating substrates[23,24,25,26,27]. More details concerning the optical

quality of our quantum emitters can be found in Supplementary Figure 1 and Supplementary Note 1. As we increase the pump power, the linewidth of our emission features steadily increases, which we attribute to power induced dephasing[29] (see Fig 1d). More importantly, as shown in Fig. 1e), we observe a distinctively different behavior of the integrated intensity of the two features, as we increase the excitation power. The intensity of the high energy signal (P2) increases less rapidly than the low energy signal (P1), and a characteristic saturation behavior of P1 is present as P2 still increases in intensity. We fit our data with a power law $I \sim P^x$, and find a powerlaw coefficient of $x = 0.84$ for P2 and a superlinear coefficient of $x = 1.42$ for P1. This pronounced sub- and superlinear behavior of the two emission features already outlines a possible emission cascade which is initialized by a biexcitonic state. With increasing sample temperature, P1 and P2 experience a characteristic broadening of the emission lines as well as a quenching of their intensity which is detailed in the Supplementary Figure 2 and Supplementary Note 2. A second pair of emission lines, which we found in our monolayer and which behaves qualitatively similar, is depicted in Supplementary Figure 3.

The exciton and biexciton could be further identified by the time-resolved photoluminescence of the localized emitter excited with a 3ps pulsed laser at 475nm (See Fig. 1f, and Supplementary Figure 4). Both P1 and P2 were fitted with a single exponential decay function with the time constant of $\tau = 0.793 \pm 0.017$ns for P1 and $\tau = 1.504 \pm 0.028$ns for P2. For the exciton and biexciton in semiconductor quantum dots, the time evolution of the probabilities for exciton (X) and biexciton (XX) could be modeled with the radiative transition functions[30] and the ratio τ_x / τ_{xx} would be dependent on their electron-hole spatial wave function separation. A ratio $\tau_x / \tau_{xx} = 2$ is expected for excitons in the presence of a confining potential, and if the biexciton size was much larger than the exciton size. Here the extracted ratio ~ 1.897 corroborates our assignment of the exciton and biexciton lines. We point out, that in the monolayer WSe_2, the free biexciton diameter was expected to be 4 times larger than the free exciton size[28], which would translate into a similar ratio.

Resulting from the inherent selection rules of tightly localized excitons, emission from the biexcitonic cascade features a very characteristic polarization behavior. In fact, if the exciton is subject to a fine structure splitting, induced by structural anisotropies, the biexcitonic emission feature is typically subject to the same splitting, yet with opposite sign. This is a natural consequence of the emission

cascade, which is sketched in Fig. 2a). We test the polarization features of our emission pair at low excitation powers. Indeed, for low pump powers, our two primary emission features of interest each split into doublets, which already indicates the presence of a fine structure in our system (Fig. 2b). We carry out polarization resolved spectroscopy by inserting a linear polarizer and a $\lambda/2$ wave-plate in our beam path, and study the intensity of each of these split peaks as a function of the polarizer orientation. Figure 2c depicts the intensity of each signal, which all feature a strong sinusoidal behavior suggesting a close-to perfect linear polarization of both the P1 and the P2 signal. This becomes even clearer as we normalize each peak, and plot the corresponding spectra in the contour graph in Fig. 2d. As expected from the emission cascade, the two possible branches of the cascade are separated by the characteristic fine structure splitting both in the X and the XX feature with an oscillation period shifted by a phase of π.

Single photon correlation spectroscopy

The key hypothesis, namely the presence of an emission cascade from the biexcitonic state to the crystal ground state, can be verified via photon correlation experiments. We prove the capability of our system to emit quantum light by measuring the second order autocorrelation function of the emission feature P2 under pulse-laser excitation, which we have attributed to the excitonic state. The corresponding autocorrelation histogram from the emission signal is shown in Fig 3a. The emission is spectrally filtered by a pair of band filters and then coupled into a fiber-based Hanbury-Brown and Twiss setup. Figure 3b shows the autocorrelation of the emission P1, which was attributed to the biexcitonic state and filtered out by a monochromator. The results around τ =0 reveal $g^2(0)=0.397\pm 0.04$ for P1 (XX) and $g^2(0)=0.218\pm0.06$ for P2 (X), which reach down well below 0.5 and therefore prove single photon emission from both states. Corresponding measurements under CW non-resonant excitation are shown in Supplementary Figure 5. Next, we measure the cross correlation between the emission signals P1 and P2. Figure 3c shows the photon-correlation measurements under the same pulse excitation between two peaks, from which we obtained the second-order correlation at zero delay of $g^2(0)=1.404\pm0.038$ (see Methods for details on the data analysis). The right inset of Fig. 3c shows the cross correlation between P1 and P2 under Continuous-Wave (CW) excitation. For $\tau<0$, we

observe a characteristic antibunching, which transits to a value significantly above unity at small positive delays. This bunching effect definitely proves the cascaded nature of the single photon emission[34,35], and thus strongly suggests the presence of an emission cascade from a biexciton in our monolayer system.

Discussion

The microscopic properties of the localized biexciton and its selection rules in monolayers of WSe_2 require more work for clarification. The free biexciton states in WSe_2 monolayers have been predicted by variational calculations and experimental indications were previously discussed[28]. Similarly to the expectations for free biexcitonic states in WSe_2, our lifetime measurements indicate that the localized biexcitons are probably larger than the exciton. Furthermore, a previous study of localized excitons in WSe_2 indicated the presence of a dark localized exciton as well as a free dark exciton state several meV below the bright exciton[37,38]. In our current study, we did not observe any significant contribution of a dark state, which would necessarily lead to a blinking in the defect emission, and thus to a long term bunching in the autocorrelation of the exciton emission. However, it will be interesting to see, if e.g. a temperature induced population of the dark exciton will lead to a modification of the observed quantum statistics in our cascaded system.

In conclusion, we have observed clear evidence for the presence of a localized biexcitonic complex in a monolayer of WSe_2. The successful observation of the localized biexciton is a consequence of the very clean emission spectra, which in turn are a result of our hybrid semiconductor-monolayer heterostructure . The emission signal features a superlinear increase of its intensity with excitation power, its fine structure splitting is of the same magnitude as the corresponding exciton, yet of opposite linear polarization, and most importantly, the characteristic emission cascade is unambiguously verified in a cross correlation experiment. Time resolved spectroscopy on the exciton-biexciton pair suggests that the biexcitonic Bohr radius is approximately four times the size of the exciton's spatial extension. We believe that this demonstration of a multi-particle complex in a monolayer material will give rise to a new plethora of interesting effects, such as the demonstration of valley entanglement. Furthermore, the high optical quality of our emitter on the surface of a sample it is a promising step towards the implementation of ultra compact, bright sources of single,

indistinguishable[31] photons and even pairs of polarization and time-bin entangled photons[32]. Since the emitting dipole is allocated on a surface, rather than embedded in a high index medium, we believe that simple photonic architectures can be developed to obtain broadband quantum light sources with very large photon extraction[36].

Methods

Sample design and fabrication.

The WSe_2 monolayer was mechanically exfoliated from bulk WSe_2. We followed the method described in ref 33. The flake was exfoliated onto a polymer gel film (polydimethylsiloxane) and was then transferred onto a GaAs substrate with 250nm GaInP layer on top, where the surface was cleaned with oxygen plasma before the exfoliation.

The GaInP/GaAs substrate was fabricated by gas-source molecular beam epitaxy (GS-MBE) on epi-ready GaAs (001) substrates. After a thermal oxide-desorption step, a 150nm thick GaAs buffer layer was deposited, followed by growth of a 250nm thick layer of $Ga_{0.51}In_{0.49}P$ lattice matched to GaAs. The sample was then cooled down under supply of thermally cracked PH3 gas and removed from the MBE system.

Optical measurements and data analysis

The diagram of the experimental setup is shown in Supplementary Figure 6 and described in great detail in Supplementary Note 3. The sample was optically excited non-resonantly using a CW 532nm laser and a mode locked, up-converted Ti:Sapphire (475nm) with a pulse frequency of 82 MHz and a pulse length of 3 ps. A short pass (670nm) filter was inserted into the excitation arm to filter out the unwanted scattering laser light from the fiber. The collected light was passing through a long-pass (690nm) filter. A spectrometer with a CCD (charge-coupled device) cooled by liquid nitrogen was used to measure the PL. The spectrum was defined by 1500mm grating with the spectral resolution~70 μeV. Photon autocorrelation measurements were performed after collecting the signal into a multimode fiber beam splitter, where the counts were recorded by two single photon avalanche

photodiodes. The correlation function for the pulse experiments was described in the following formula:

$$g^2(t) = y0 + \sum_{n=1}^{n=7} A_n * (\exp(\frac{t-t_n}{\tau_a})\Theta(-t) + \exp(-\frac{t-t_n}{\tau_b})\Theta(t)) \quad (1)$$

where $\Theta(t)$ is the step function, $y0$ is the offset, A_n is the peak height of the *n*th peak, t_n is the peak position of the *n*th peak. $\tau_a = \tau_b$ in the exciton (X) autocorrelation and the biexciton (XX) autocorrelation measurement, which are related to characteristic radiation time of X and XX.

The CW cross correlation data in the inset of Fig.3c was fitted with multiexcitonic model similar to that in semiconductor QDs[34]:

$$g^2(t) = g^2(t, neg)\Theta(-t) + g^2(t, pos)\Theta(t) \quad (2)$$

where $g^2(t,c) = 1 - A_{1c} * \exp(-|\frac{t}{t_1}|) - A_{2c} * \exp(-|\frac{t}{t_2}|)$. (3)

For $g^2(t<0)$, we will see the anti-bunching while the bunching for $g^2(t>0)$.

In order to account for the finite timing resolution of our setup, we fitted the pulse correlation data in Fig.3 with the correlation function convolved with a Gaussian distribution function:

$$f_{\mathrm{Det}}(t) = \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{1}{2}\left(\frac{t-t_0}{\sigma}\right)^2}, \quad (4)$$

where the $2\sqrt{2\ln 2}\sigma$ represents the 350ps APD time resolution. So the real fitting function would be:

$$g^2_{\mathrm{real}}(\tau) = (g^2(t) * f_{\mathrm{Det}})(\tau) = \int_{-\infty}^{\infty} g^2(t) f_{\mathrm{Det}}(\tau - t)dt \quad (5)$$

Photon cross-correlation measurements were carried out after passing the signal into a 50:50 beam splitter. One beam was filtered out at P2 by a pair of 1nm bandwidth filters centered at 720nm and the other beam was leading to the monochromator with the filtered spectrum at P1. The pulsed $g^2(0)$ value was calculated from the fitted area in the zero-time delay peak divided by the average of the adjacent six peaks.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

References:

1. Pan, J. W. *et al*. Multiphoton entanglement and interferometry. *Rev. Mod. Phys*. **84,** 777-838 (2012).
2. Giovannetti, V. *et al.* Quantum Metrology. *Phys. Rev. Lett.* **96,** 010401 (2006)
3. Georgescu, I. M. *et al.* Quantum simulation. *Rev. Mod. Phys.* **86,** 153 (2014)
4. Kimble, H. J. The quantum internet. *Nature* **453,** 1023–1030 (2008).
5. Wang, Q. H., Kalantar-Zadeh, K., Kis, A., Coleman, J. N. & Strano, M. S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nat. Nanotechnol.* **7,** 699–712 (2012).
6. Splendiani, A. *et al.* Emerging photoluminescence in monolayer MoS_2. *Nano Lett.* **10,** 1271–1275 (2010).
7. Mak, K. F., Lee, C., Hone, J., Shan, J. & Heinz, T. F. Atomically thin MoS_2: A new direct-gap semiconductor. *Phys. Rev. Lett.* **105,** 2–5 (2010).
8. Mak, K. F. *et al.* Tightly bound trions in monolayer MoS_2. *Nat. Mater.* **12,** 207–211 (2013).
9. He, K. *et al*. Tightly bound excitons in monolayer WSe_2. *Phys. Rev. Lett*. **113,** 26803 (2014).
10. Ross, J. S. *et al*. Electrical control of neutral and charged excitons in a monolayer semiconductor. *Nat. Commun.* **4,** 1474 (2013).
11. Steinhoff, A., Kim, J. H., Jahnke, F., Rösner, M., Kim, D. S., Lee, C., Han, G. H., Jeong, M. S., Wehling, T. O. and Gies, C. Efficient Excitonic Photoluminescence in Direct and indirect Band Gap Monolayer MoS_2. *Nano. Lett.* **15,** 6841-6847 (2015).
12. Chernikov, A. *et al.* Exciton Binding Energy and Nonhydrogenic Rydberg Series in Monolayer WS_2. *Phys. Rev. Lett.* **113,** 076802 (2014).
13. Ugeda, M. M. *et al.* Giant bandgap renormalization and excitonic effects in a monolayer transition metal dichalcogenide semiconductor. *Nat. Mater.* **13,** 1091–1095 (2014).
14. Xu, X., Yao, W., Xiao, D. & Heinz, T. F. Spin and pseudospins in layered transition metal dichalcogenides. *Nat. Phys.* **10,** 343–350 (2014).
15. Wang, G. *et al.* Valley dynamics probed through charged and neutral exciton emission in monolayer WSe_2. *Phys. Rev. B* **90,** 075413 (2014).

16. Xiao, D., Liu, G.-B., Feng, W., Xu, X. & Yao, W. Coupled spin and valley physics in monolayers of MoS_2 and other Group-VI dichalcogenides. *Phys. Rev. Lett.* **108,** 196802 (2012).
17. Jones, A. M. *et al.* Optical generation of excitonic valley coherence in monolayer WSe_2. *Nat. Nanotechnol.* **8,** 634–638 (2013).
18. Zeng, H., Dai, J., Yao, W., Xiao, D. & Cui, X. Valley polarization in MoS_2 monolayers by optical pumping. *Nature Nanotechnol*. **7,** 490–493 (2012)
19. Mak, K. F., He, K., Shan, J. & Heinz, T. F. Control of valley polarization in monolayer MoS_2 by optical helicity. *Nat. Nanotechnol.* **7,** 494–498 (2012).
20. Cao, T. *et al.* Valley-selective circular dichroism of monolayer molybdenum disulphide. *Nat. Commun.* **3,** 887 (2012).
21. Dufferwiel, S. *et al.* Exciton-polaritons in van der Waals heterostructures embedded in tunable microcavities. *Nat. Commun.* **6,** 8579 (2015).
22. Lundt, N. *et al.* Room temperature Tamm-Plasmon Exciton-Polaritons with a WSe_2 monolayer. Preprint at http://arxiv.org/abs/1604.03916 (2016).
23. He, Y. M., Clark, G., Schaibley, J. R., He, Y., Chen, M. C., Wei, Y. J., Ding, X., Zhang, Q., Yao, W., Xu, X., Lu, C. Y. and Pan, J. W. Single quantum emitters in monolayer semiconductors. *Nature. Nanotechnol.* **10,** 497–502 (2015).
24. Koperski, M., Nogajewski, K. *et al.* Single photon emitters in exfoliated WSe_2 structures. *Nature. Nanotechnol.* **10,** 503–506 (2015).
25. Chakraborty, C., Kinnischtzke, L. *et al.* Voltage-controlled quan-tum light from an atomically thin semiconductor. *Nature. Nanotechnol.* **10,** 507–511 (2015).
26. Srivastava, A., Sidler, M. *et al.* Optically active quantum dots in monolayer WSe_2. *Nature. Nanotechnol.* **10,** 491–496 (2015).
27. Tonndorf, P. Schmidt, R. *et al.* Single-photon emission from localized excitons in an atomically thin semiconductor. *Optica.* **2,** 347–352 (2015).
28. You, Y., Zhang, X.X. *et al.* Observation of biexcitons in monolayer WSe_2. *Nature. Physics.* **11,** 477–481 (2015).
29. Moody, G. *et al.* Intrinsic homogeneous linewidth and broadening mechanisms of excitons in monolayer transition metal dichalcogenides. *Nat. Commun.* **6,** 8315 (2015)
30. Bacher, G., Wiegand, R., Seufert, J. *et al.* Biexciton versus Exciton Lifetime in a Single Semiconductor Quantum Dot. *Phys. Rev. Lett.* **83,** 4417-4420 (1999).
31. Vamivakas, A. N., Zhao, Y. *et al*. Spin-resolved quantum-dot resonance fluorescence. *Nature physics* **5,** 198-202 (2009)
32. Jayakumar, H. *et al.* Time-bin entangled photons from a quantum dot. *Nat. Commun.* **5,** 4251 (2014)
33. Castellanos- Gomez, A. *et al* . Deterministic transfer of two-dimensional materials by all-dry viscoelastic stamping. *2D Mater*. **1,** 011002 (2014)
34. Moreau, E., Robert, I., Manin, L. *et al.* Quantum Cascade of Photons in Semiconductor Quantum dots. *Phys. Rev. Lett.* **87,** 183601 (2001).
35. Aspect, A., Roger, G. *et al.* Time Correlations between the Two Sidebands of the Resonance Fluorescence Triplet. *Phys. Rev. Lett.* **45,** 617-620 (1980).
36. Lien, D. H., *et al.* Engineering Light Outcoupling in 2D Materials. *Nano. Lett.* **15,** 1356-1361 (2015).

37. Zhang, X. X., *et al.* Experimental Evidence for Dark excitons in Monolayer WSe_2. *Phys. Rev. Lett.* **115,** 257403 (2015).
38. He, Y. M., *et al.* Phonon induced line broadening and population of the dark exciton in a deeply trapped localized emitter in monolayer WSe_2. *Optical Express.* **24,** 8066–8073 (2016).

End Notes

Acknowledgement: This work has been supported by the State of Bavaria and the European Research Council (Project UnLiMIt-2D). We further thank S. Stoll for assistance in the exfoliation and pre-characterization of the high quality monolayer material.

Author contributions: C.S and S.H. initiated the study and guided the work. Y.M.H and O.I. carried out spectroscopic investigations. N.L. and O.I. exfoliated, identified and transferred the monolayer. Y.M.H and C.S. analyzed and interpreted the experimental data, supported by all co-authors. M.D. and K.S. provided the theory. C.S. and Y.M.H. wrote the manuscript, with input from all coauthors.

Competing financial interest: The authors declare no competing financial interest.

Figures

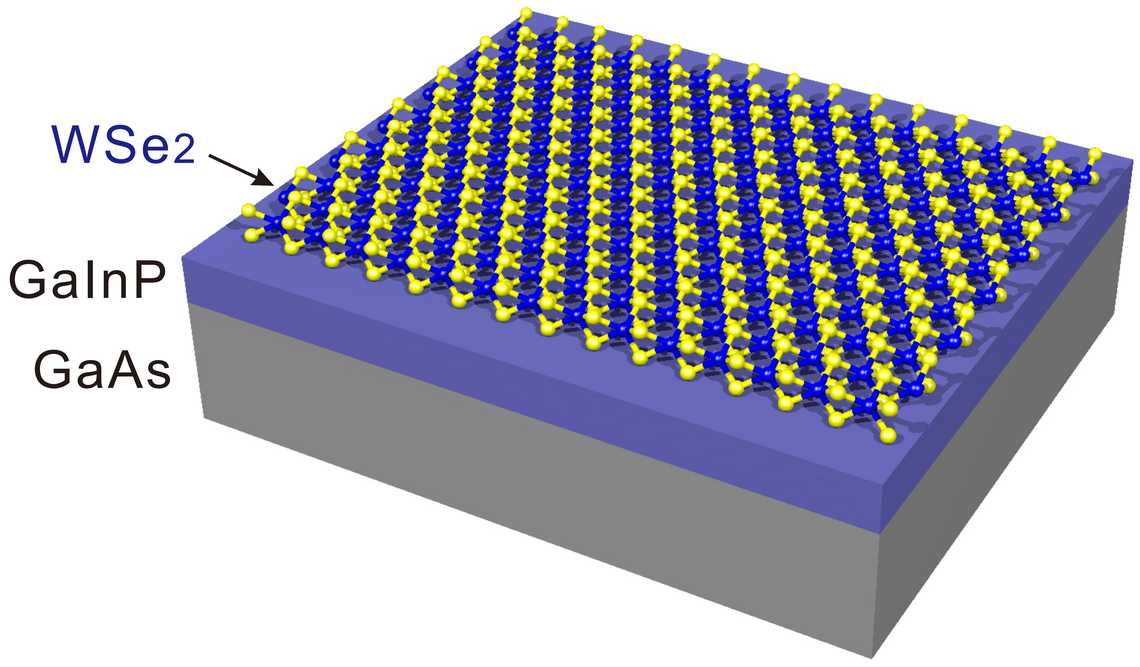
Figure 1 | Narrow spectral lines in monolayer WSe_2. a, Schematic illustration of the device with the deposited WSe_2 monolayer. **b**, Contour plot of photoluminescence intensity within the energy range between 1.525eV and 1.734eV, over 56μm×56μm. The dashed white line marks the potential monolayer area. **c**, Photoluminescence spectrum of the localized emitter in the WSe_2 monolayer at 4.5K, showing different emission behaviors with increasing laser power dominated by peaks at 1.7167eV (P1) and 1.7206eV (P2). **d**, Extracted linewidth of P1 and P2, plotted as a function of excitation power. The spectrum for low excitation power shows resolution limited linewidth for both P1 and P2. **e,** The integrated counts of the photon emission from P1 and P2 shows super-linear and sub-linear behavior with increasing laser power in Log-Log plot. The red line is the power law fitting $I \propto P^x$, with the extracted x=0.84 for P2 and x=1.42 for P1. **f,** Time resolved photoluminescence of P1 and P2 show the single exponential decay with the time constant of 0.793±0.017ns for P1 and 1.504±

0.028ns for P2. Uncertainties are one standard deviation values based on a least square fit to the data.

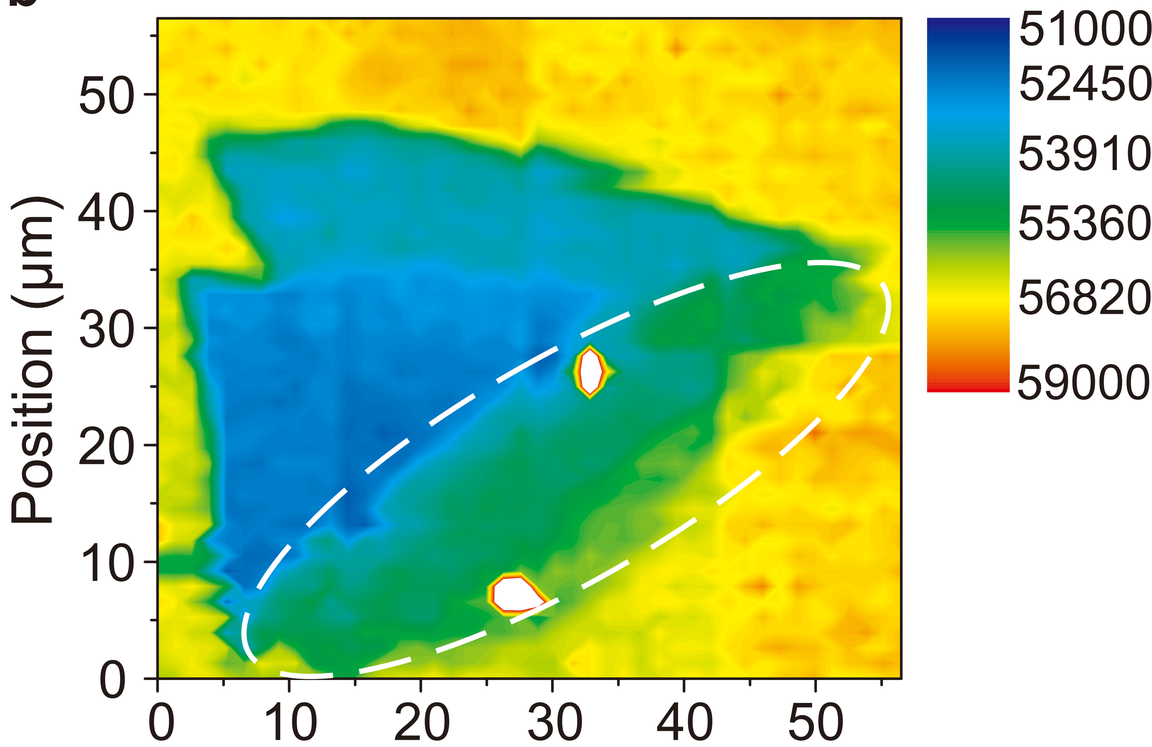
Figure 2 | Polarization-resolved photoluminescence. a, Schematic representation of biexcitonic emission cascade. The fine-structure splitting is expected for the electron-hole exchange interaction in presence of in-plane anisotropy. **b**, Polarization resolved spectrum at linear polarization H, V. Two pairs of spectral doublets are observed at 1.7167 (P1)-1.7171eV (P3) and 1.7202 (P4)-1.7206eV (P2). Four peaks are indicated by the dashed lines. **c,** The integrated counts of the photon emission from P1, P2, P3, P4 as a function of the polarization detection angle. The red lines are the sinusoidal fits, showing 2 pairs of cross linear-polarized doublets. **d,** Contour representation of the four peaks, after normalizing to the maximum peak intensity, yielding a fine structure splitting ~0.4meV.

Figure 3 | Photon correlation measurements. a,b Second-order autocorrelation of measurement of the P2 and P1 under 13.6nW pulsed excitation at 475nm with the repetition rate of 82MHz and the pulse length of 3ps. We could extract $g^2(0)=0.397\pm0.06$ for P1 and $g^2(0)=0.286\pm0.04$ for P2. **c,** Cross-correlation measurement between P1 and P2 under pulsed excitation. The obtained $g^2(0)=1.404\pm0.038$, which obviously shows the bunching effect of the emission. The red lines are fitted with two exponential decay convolved with the response function. The inset is the cross correlation measurement of P1 and P2 under a 70nw CW laser excitation at 532nm. The red line in the inset is a fit with multiexcitonic model convolved with the response function. The blue line is the deconvoluted curve, which shows $g^2(0)=0.038\pm0.004$ for negative delay and $g^2(0)=3.434\pm0.129$ for positive delay. All uncertainties correspond to one s.d. More details about fitting could see Methods.

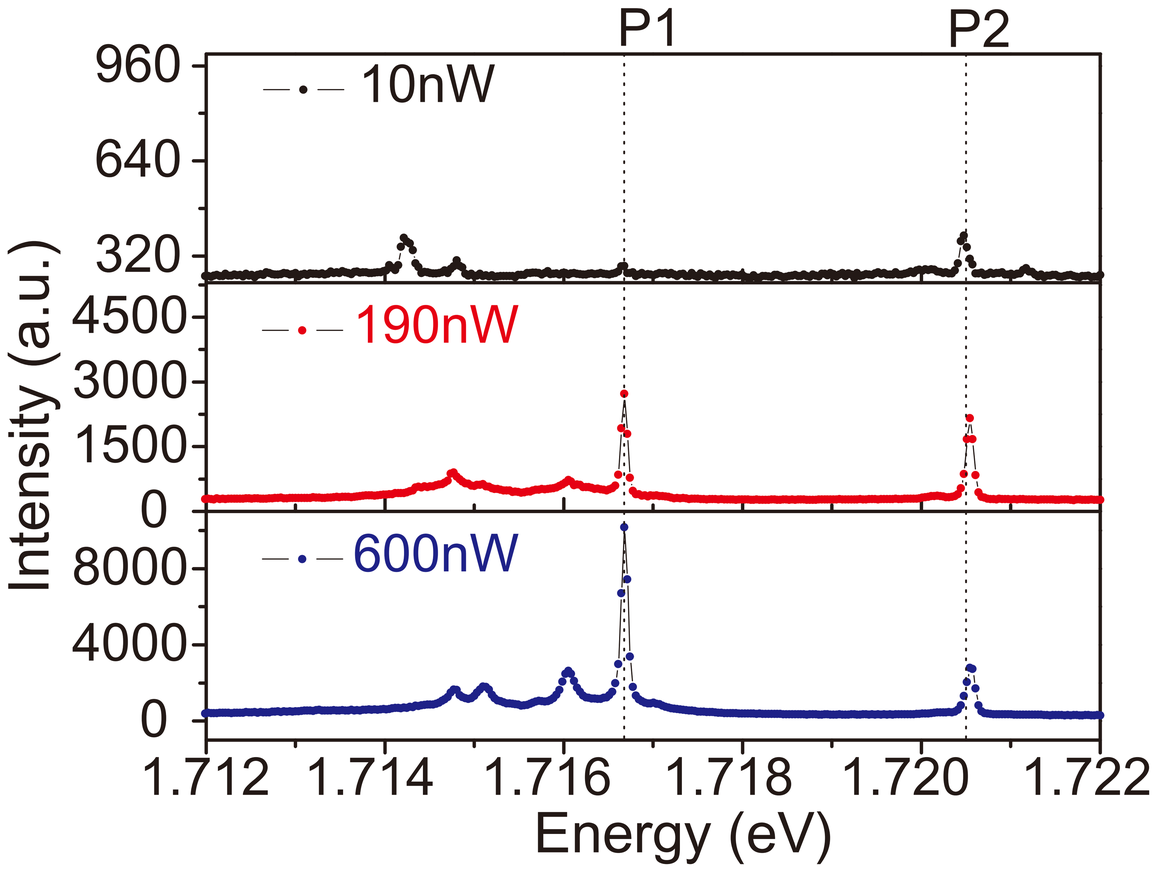
a


WSe2
GaInP
GaAs

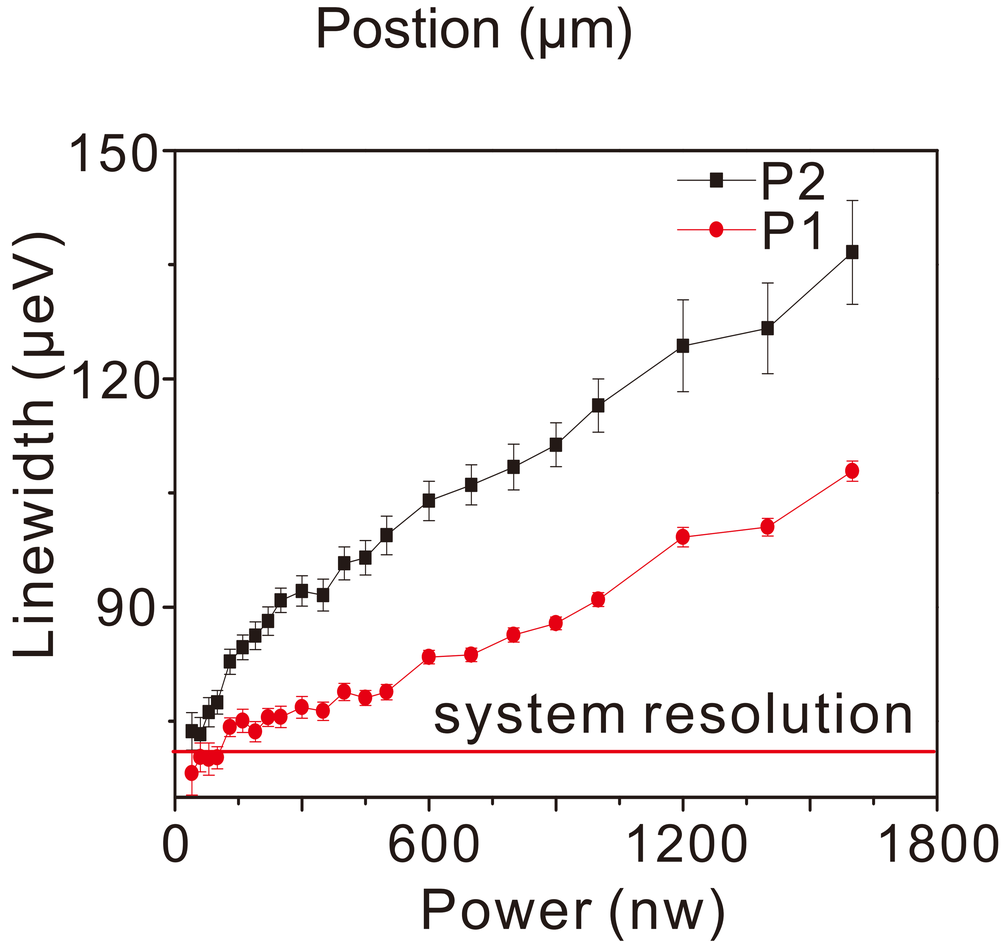
b


Intensity(cps)
51000
52450
53910
55360
56820
59000
Position (μm)
0
10
20
30
40
50

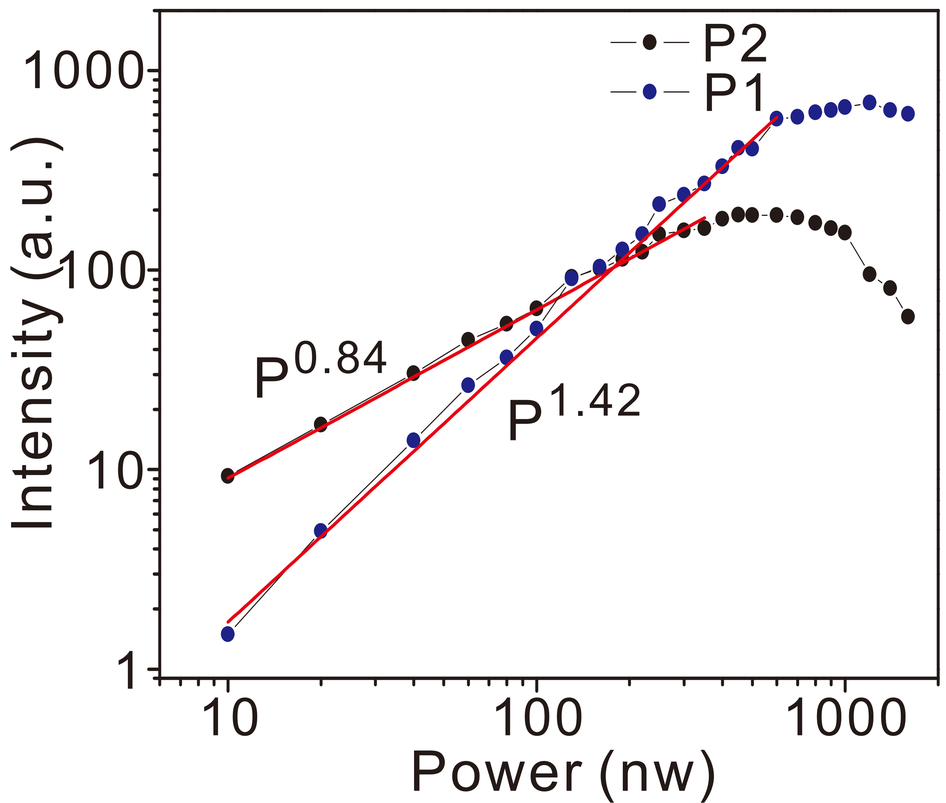
c


P1
P2
10nW
190nW
600nW
Intensity (a.u.)
960
640
320
4500
3000
1500
0
8000
4000
0
1.712
1.714
1.716
1.718
1.720
1.722
Energy (eV)

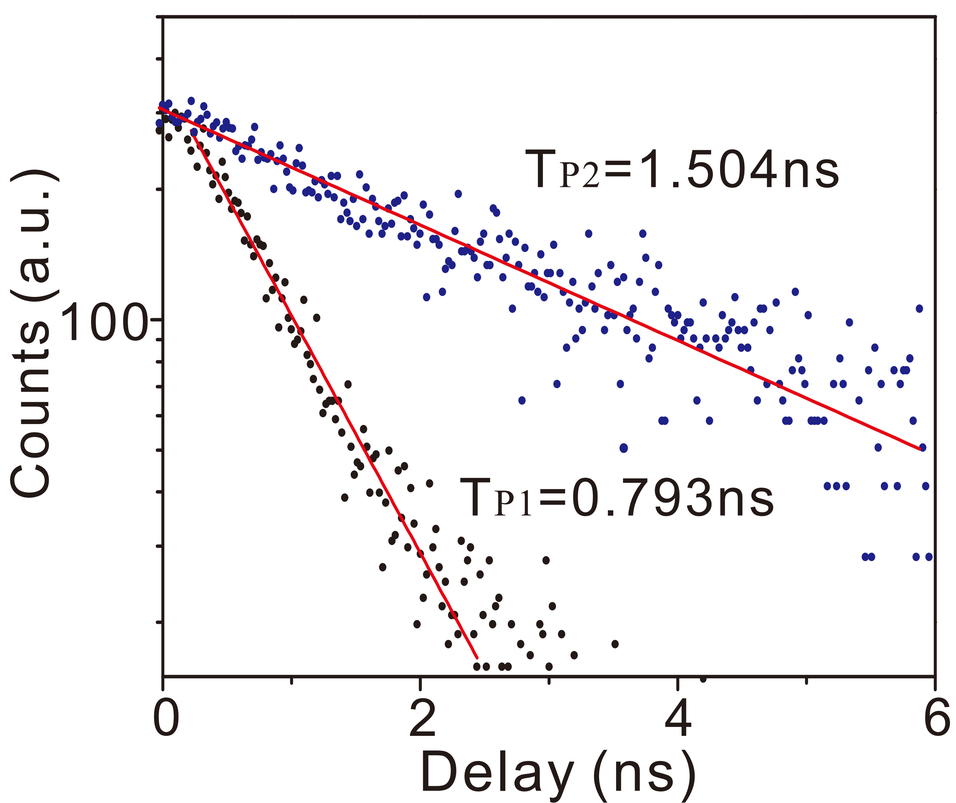
d


Postion (μm)
P2
P1
Linewidth (μeV)
150
120
90
system resolution
0
600
1200
1800
Power (nw)

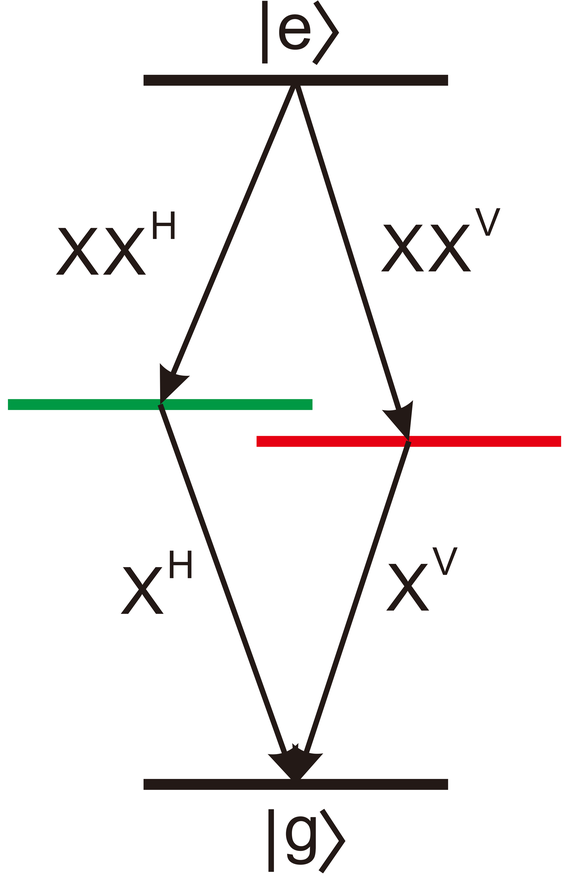
e


P2
P1
Intensity (a.u.)
1000
100
10
1
P0.84
P1.42
10
100
1000
Power (nw)

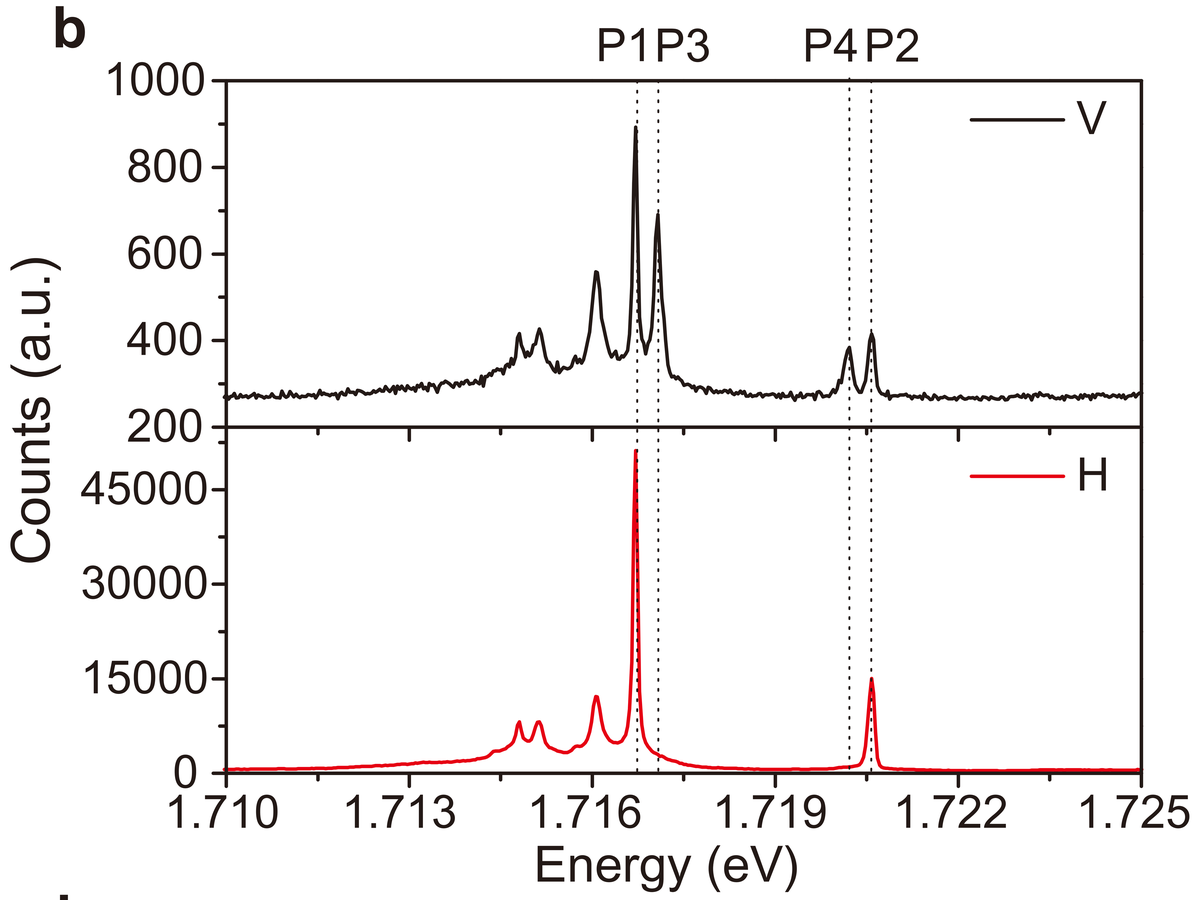
f


Counts (a.u.)
100
TP2=1.504ns
TP1=0.793ns
0
2
4
6
Delay (ns)

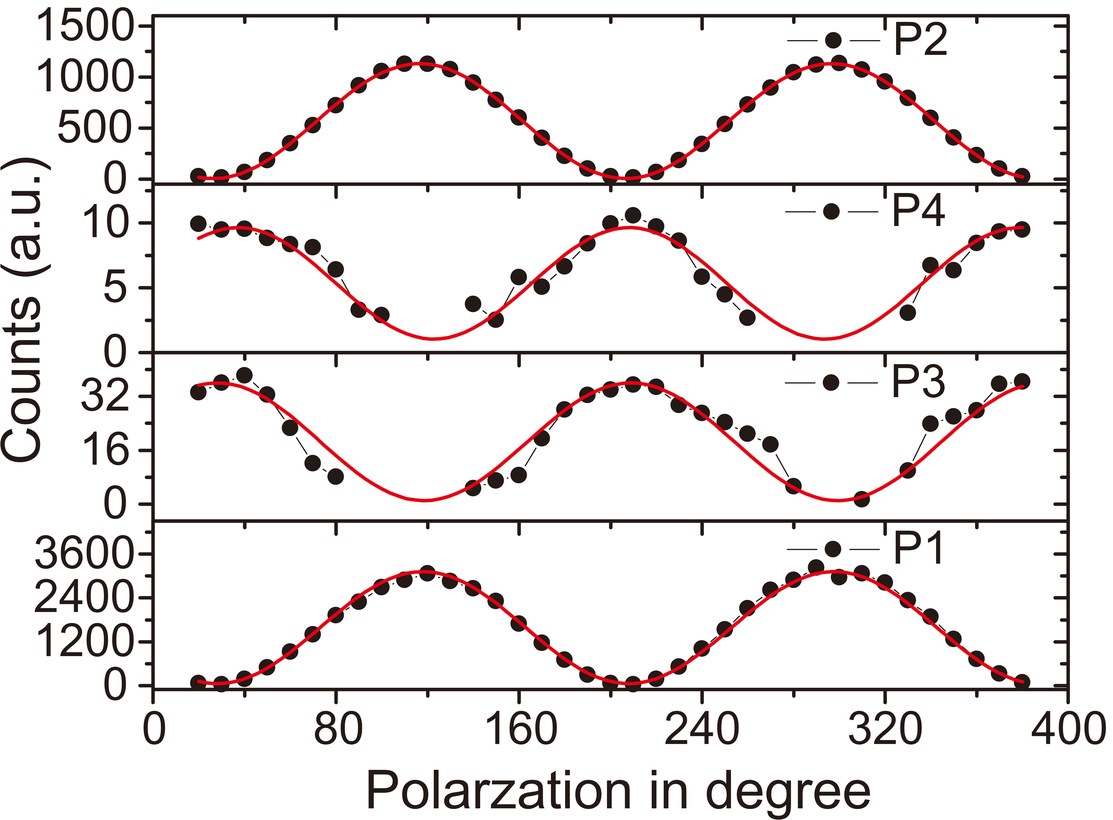
a


|e⟩
XX^H
XX^V
X^H
X^V
|g⟩

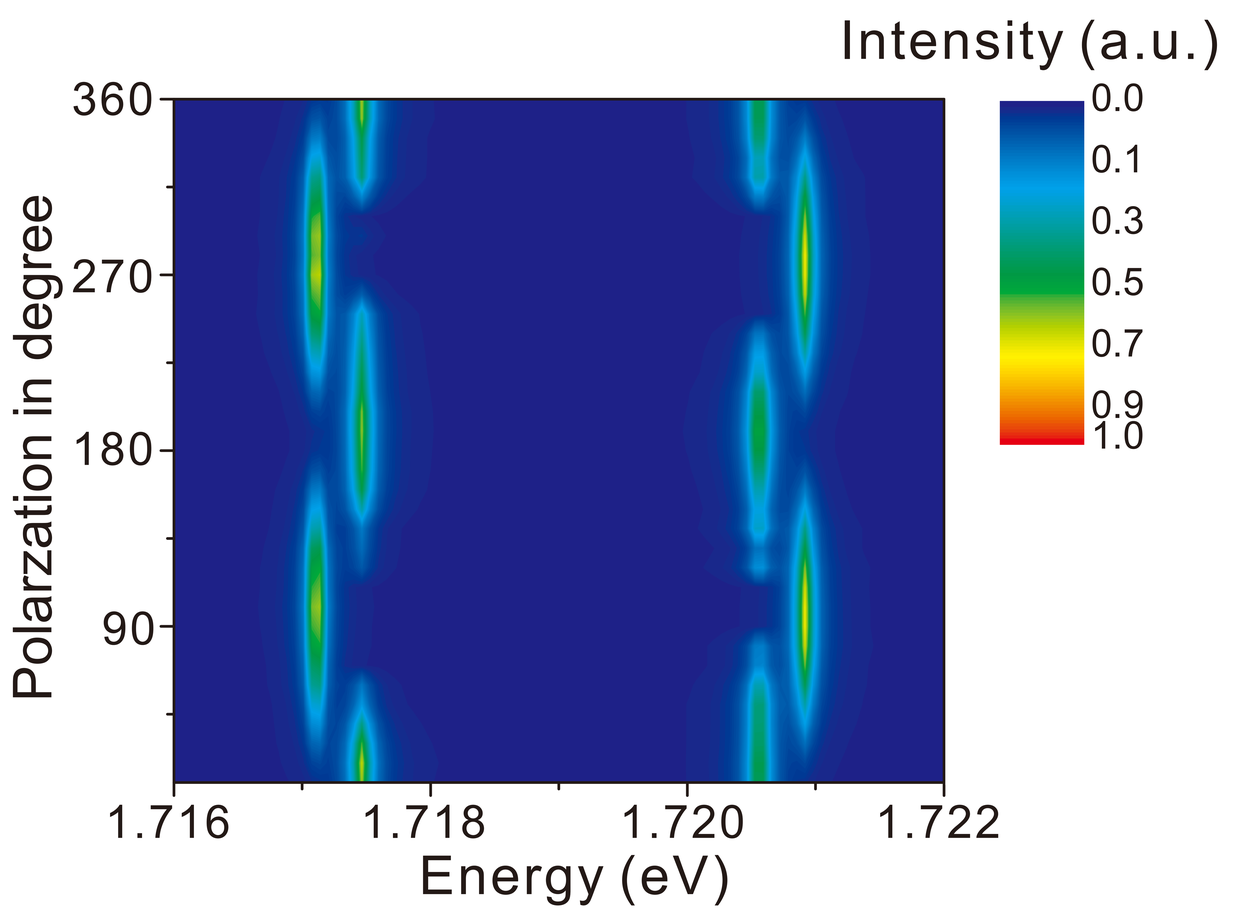
b


P1
P3
P4
P2
V
H
Counts (a.u.)
1000
800
600
400
200
45000
30000
15000
0
1.710
1.713
1.716
1.719
1.722
1.725
Energy (eV)

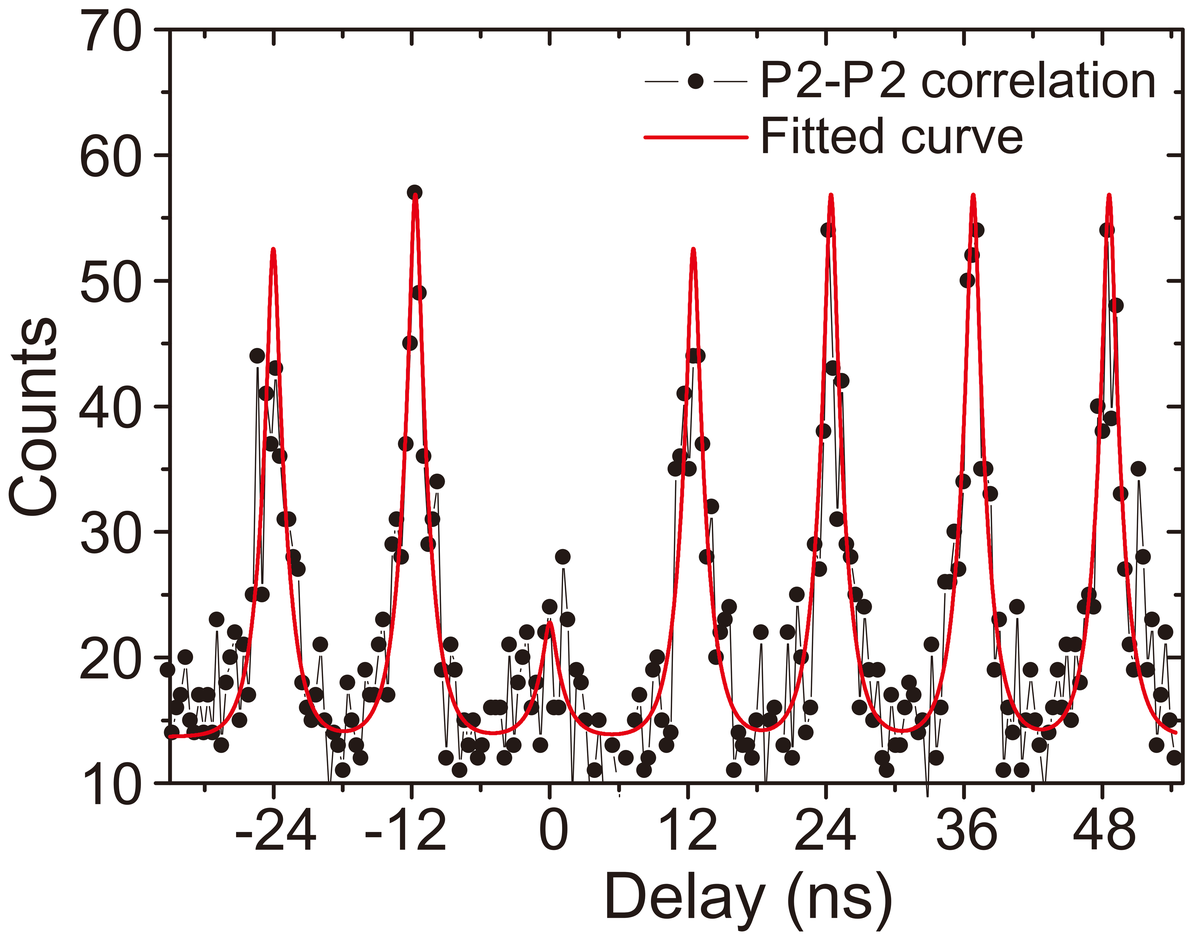
c


P2
P4
P3
P1
Counts (a.u.)
1500
1000
500
0
10
5
0
32
16
0
3600
2400
1200
0
0
80
160
240
320
400
Polarzation in degree

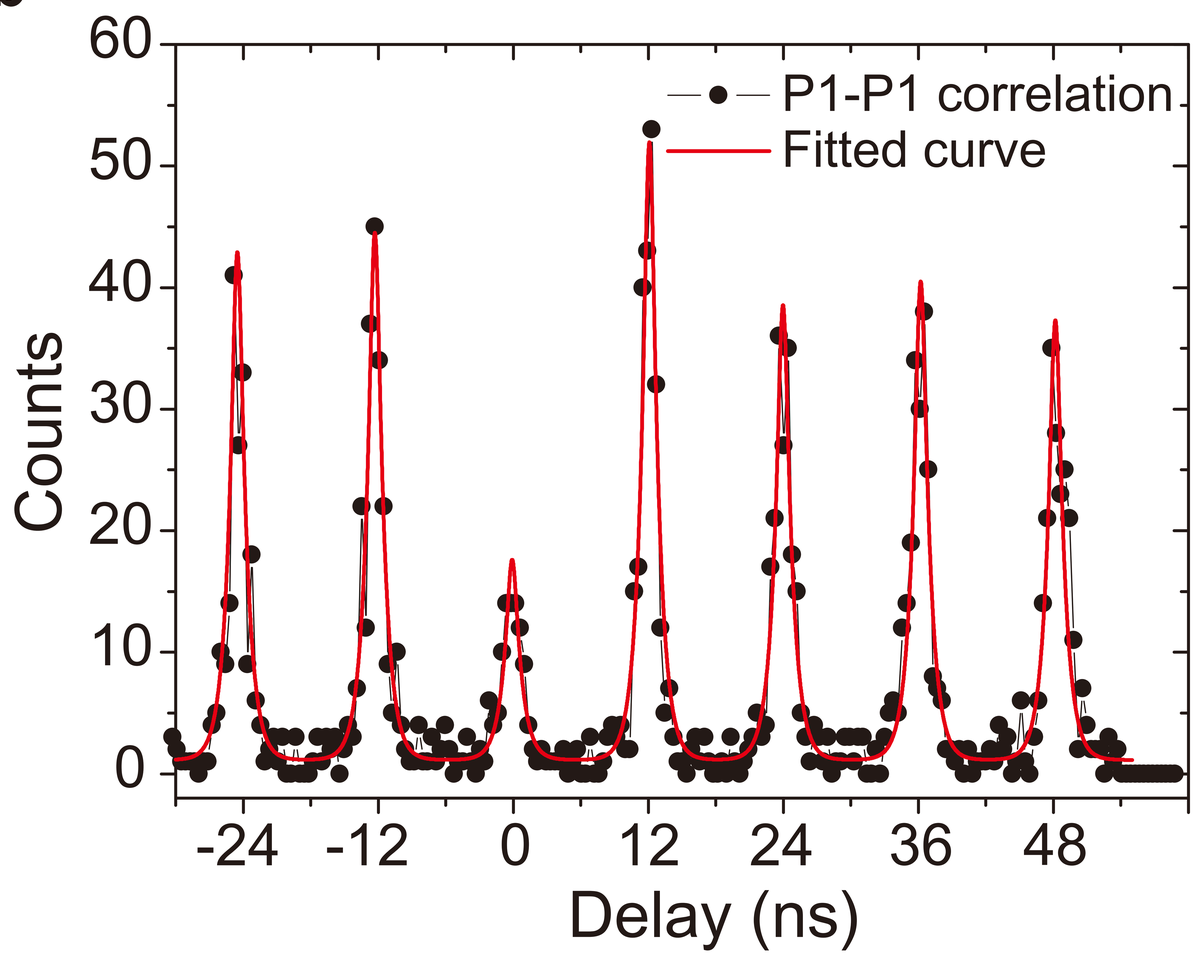
d


Intensity (a.u.)
0.0
0.1
0.3
0.5
0.7
0.9
1.0
Polarzation in degree
360
270
180
90
1.716
1.718
1.720
1.722
Energy (eV)

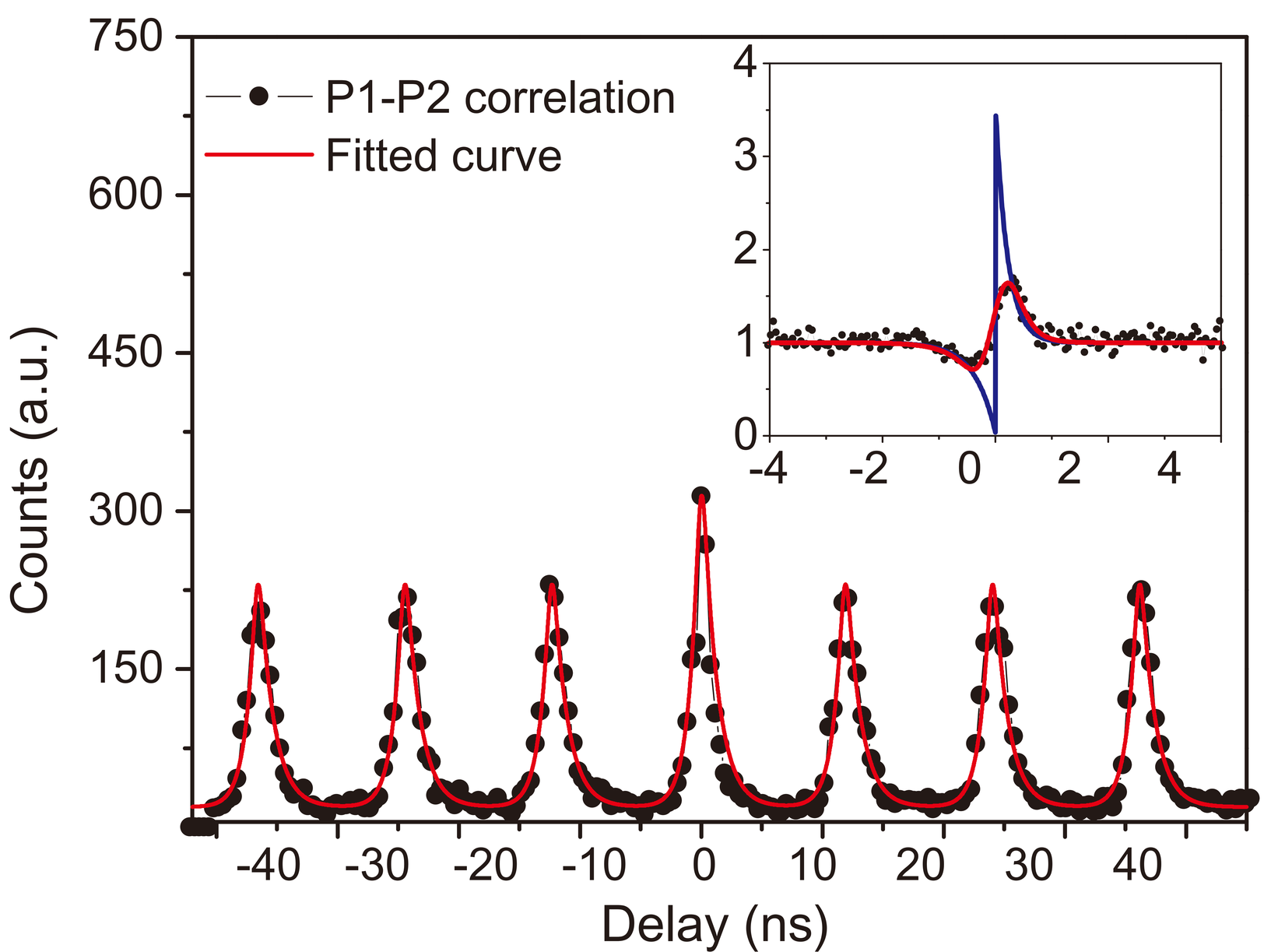
a

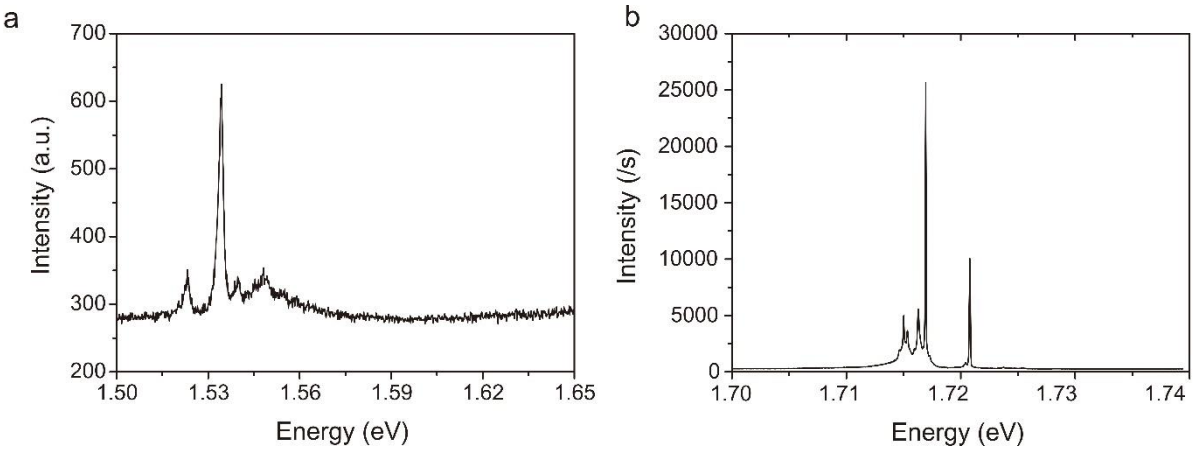

P2-P2 correlation
Fitted curve
Counts
Delay (ns)

b


P1-P1 correlation
Fitted curve
Counts
Delay (ns)

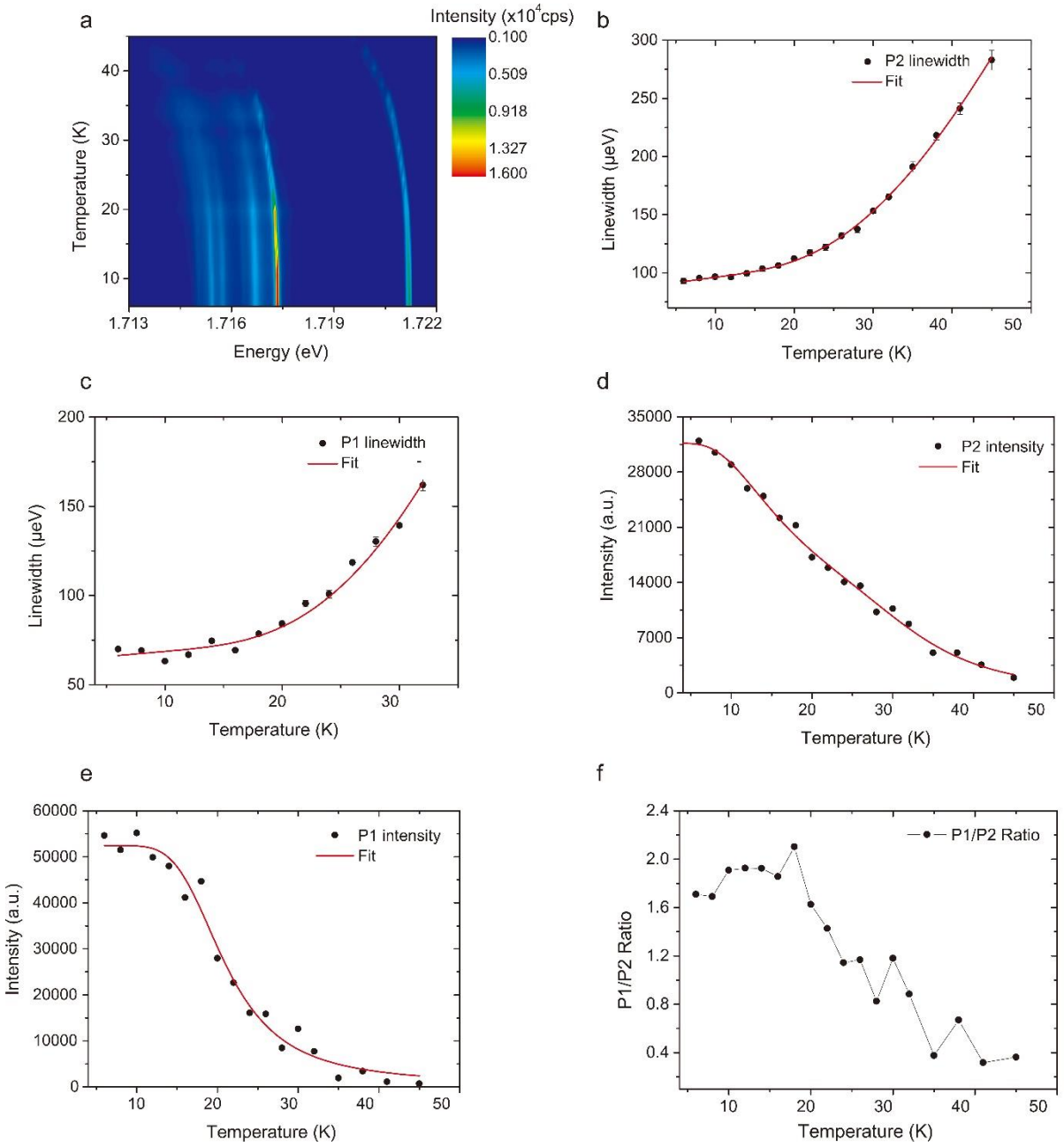
c


P1-P2 correlation
Fitted curve
Counts (a.u.)
Delay (ns)

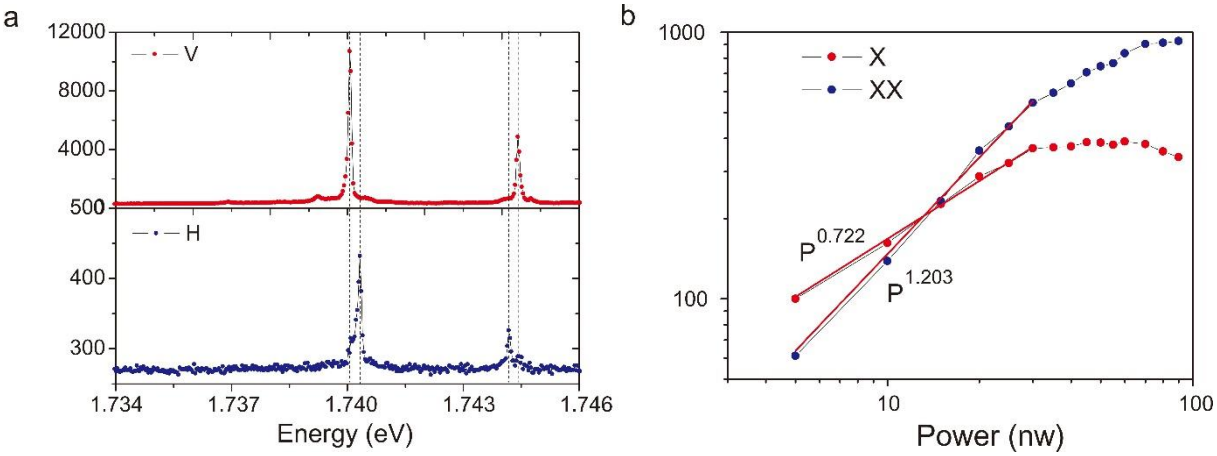


Supplementary Figure 1| Comparison of single quantum emitters on two type of substrates:

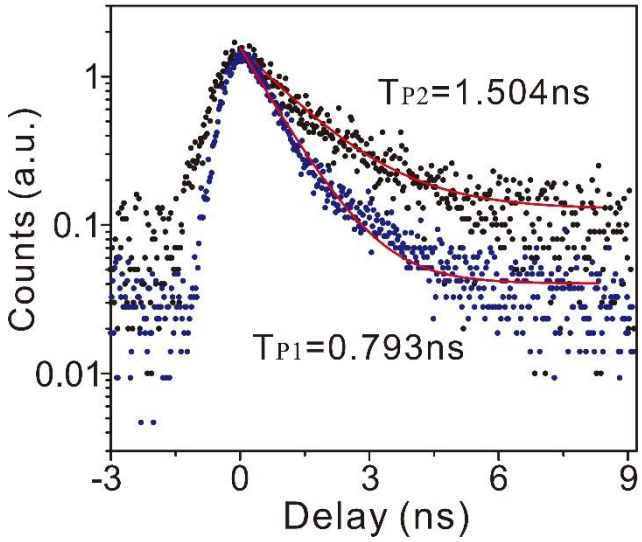
a, Photoluminescence (PL) spectrum of localized excitons in a WSe_2 monolayer, exfoliated onto a SiO_2/Si substrate (sample temperature is nominally 4.5K) . **b,** PL spectrum of localized excitons in a WSe_2 monolayer on a GaInP/GaAs heterostructure (sample temperature 4.5K). Here the linewidth is one order of magnitude narrower than that on the conventional SiO_2/Si heterostructure.



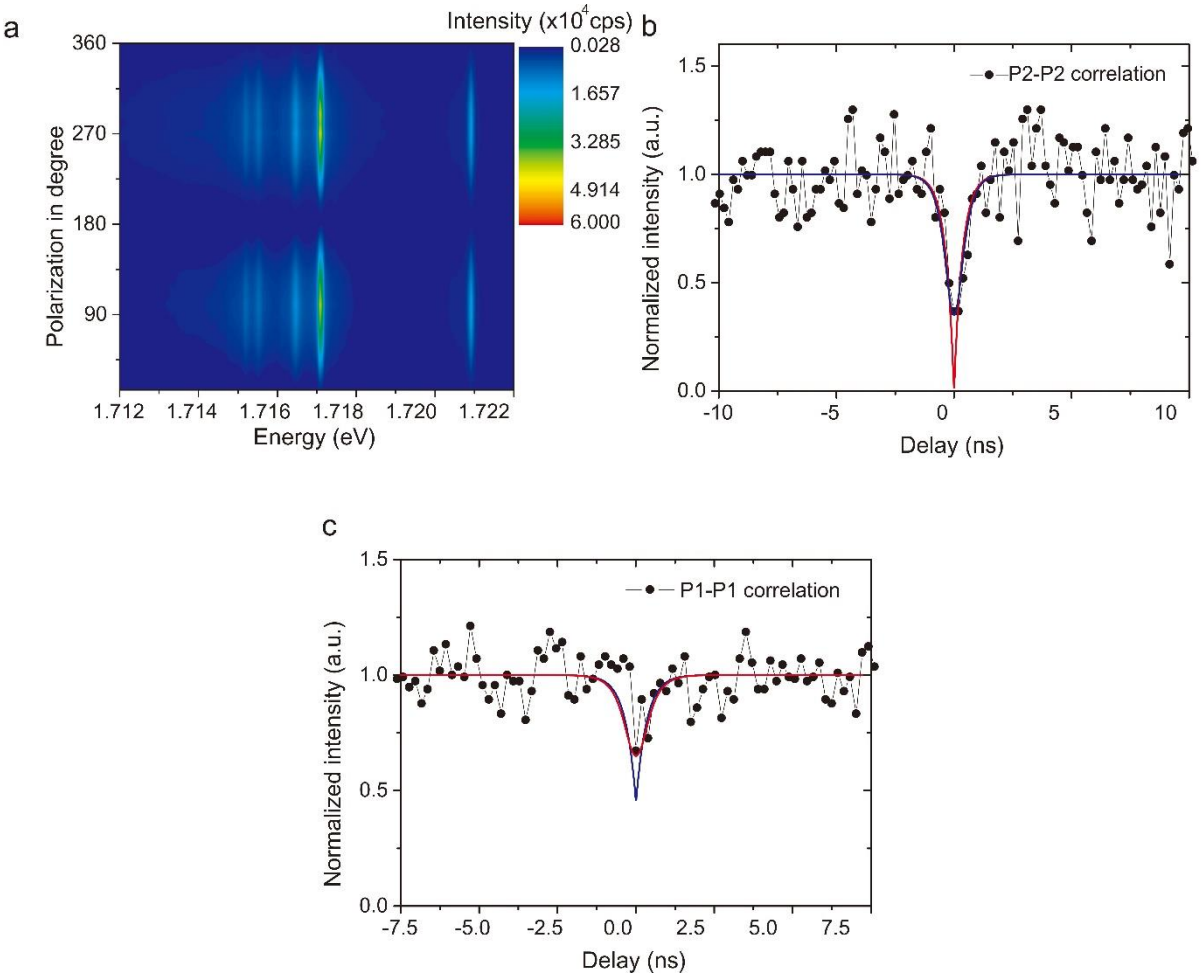
Supplementary Figure 2 | Temperature dependence of exciton and biexciton. a, Photoluminescence intensity plot as a function of temperature and photon energy. **b, c** linewidth of P1 and P2 in Fig. 1b in the main text. **d, e I**ntegrated counts as a function of temperature for P2 and P1. **f,** Intensity ratio between P1 and P2 as a function of temperature. The P1 intensity decreases much faster than P2 above 20K.



Supplementary Figure 3 | Identification of a second correlated pair of emission lines. a, PL spectrum of biexciton #2 is studied in the linear-polarization basis V and H. Two pairs of cross polarized spectral doublets are observed. The four dashed lines indicate the four energy positions of the X and XX attributed signal. **b,** The integrated counts of the photon emission from XX and X shows super-linear and sub-linear behavior with increasing laser power. The power law fitting yields coefficients of 0.722 and 1.203 for X and XX.

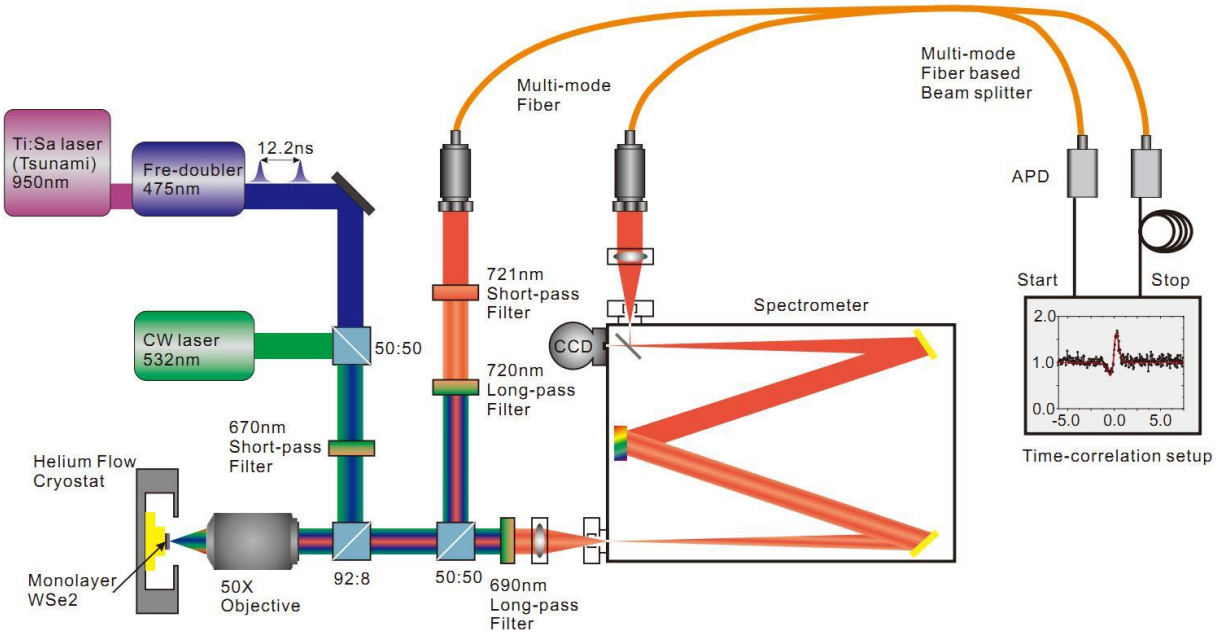


Supplementary Figure 4| Time-resolved Photoluminescence of P1 and P2 in one laser period (12.2ns). Compared to the XX, the X line is weak under higher excitation power and consequently it took longer time to perform the lifetime measurements. Thus the contribution of the APD (Avalanche Photon Diode) dark counts for the time resolved measurement on the X line is higher and results in an intensity offset.



Supplementary Figure 5 | Polarization-resolved PL and photon antibunching.

a, Photoluminescence intensity plot as a function of polarization detection angle and photon energy. Compared to P1 and P2, P3 and P4 described in Fig. 2b of the main text are too weak to be directly observed. The large fine structure splitting (~0.4meV) and unbalanced emission intensity in the doublets show a significantly high level of in-plane anisotropy. **b, c** Second-order auto correlation measurement of the P2 (1.7206eV) and P1 (1.7167eV) under 70nW CW excitation at 532nm. The red lines in the pictures are the fit with according to Eqn. 5 in the main text. The blue line is the deconvoluted curve, which yields $g^2(0)=0.014\pm0.002$ for P2 and $g^2(0)=0.457\pm0.12$ for P1. Obvious photon antibunching for both lines are observed.



Supplementary Figure 6 | Detailed sketch of the optical spectroscopy setup.

Supplementary Note 1: Comparison of single quantum emitters on two types of substrates

We believe that the successful observation of the localized biexciton in our work is a consequence of the very clean luminescence spectra, enabled by our hybrid semiconductor-monolayer heterostructure. To investigate the emission properties of the localized excitons in WSe_2 monolayers on different substrates, Supplementary figure 1(a) and 1(b) depict the photoluminescence from two WSe_2 monolayers that were transferred onto a SiO_2/Si and a GaInP/GaAs substrate. The observed linewidth of the monolayer on the GaInP substrate is more than one order of magnitude narrower than that on SiO_2 substrate (70 µeV vs. 1.9 meV). Our hypothesis is, that the semiconducting substrate allows to transfer trapped charges away from the interface, which leads to smaller spectral jitter. In addition, recent localized exciton studies on the influence of the substrate have revealed that the localized exciton linewidth and spectral wandering would be significantly enhanced after removing the SiO_2 substrate[1].

Supplementary Note 2: Temperature dependent photoluminescence

The defect's photoluminescence intensity is investigated as a function of the sample temperature (Supplementary figure 2a). In Supplementary figure 2b and 2c we plot the extracted linewidths of P1 and P2 in Fig. 1b in the main text. The linewidth broadens with increasing temperature. We extract the phonon activation energies via fitting to the data in Supplementary figure 2b, 2c via

$\gamma(T) = \gamma_0 + \gamma_{ac} * T + \frac{\gamma_{LO}}{e^{\frac{E_{LO}}{k_B T}} - 1}$ [2]. We yield $E_{LO} = 10.47 \pm 0.56\, meV$ for P2 and $E_{LO} = 10.50 \pm 0.13\, meV$ for P1, which confirms previous results about the phonon activation energy of localized emitters in WSe_2. Supplementary figure 2d and 2e depict the integrated counts as a function of temperature for P2 and P1. The red lines are the fit by the function: $I(T) = \frac{I_0}{1 + A_1 e^{\frac{E_1}{k_b T}} + A_2 e^{\frac{E_2}{k_b T}}}$. We could reproduce our experimental data with the characteristic energy $E_1 = 3.67 \pm 0.52\text{meV}$, $E_2 = 20.35 \pm 4.48\text{meV}$ for P2 and $E = 10.402 \pm 1.077\text{meV}$ by taking into account one loss channel for P1. All the extracted energies coincide with former measured results[2]. Supplementary fig. 2f shows the intensity ratio between P1 and P2 as a function of temperature. The intensity of P1 decreases much faster than P2 above 20K.

Supplementary Note 3: Details of the optical setup

The sample comprisi3g the WSe_2 monolayer is attached to the cold-finger of a liquid Helium flow cryostat (Janis ST-500). A 50x objective (NA=0.42) is used to excite and collect the photoluminescence from the flake. The WSe_2 monolayer flake is excited by a Continuous-Wave (CW) 532nm green laser, or a mode-locked pulsed laser, coupled into the beam by a 92:8 pellicle beam splitter. A short-pass filter (670nm) is inserted into the excitation arm. In the collection arm, a 690nm long pass filter was used to cut off the laser and lead the luminescence into the spectrometer (Princeton-Instrument SP2750i), which is equipped with a liquid nitrogen cooled charge coupled detector. To enable the pulse excitation, we used a mode-locked Ti: sapphire laser with the pulse duration of ~3ps, repetition rate of 82MHz and wavelength of ~950nm. The long wavelength pulse is up-converted to high frequency (~475nm) after passing through the frequency converter. A 50:50 beam splitter is used to combine both the green laser and pulse laser beam. For the photon cross-correlation measurements, the emitted photons are passed through a 50:50 beam splitter, where the transmitted photoluminescence is used to select the biexciton emission (~722.3nm). Thanks to the pure spectrum we could directly collect the pure exciton emission after passing the reflected light through a band-pass filter (center 720.6nm) composed of a long pass (720nm) filter and short pass (721nm) filter. The collected photons are coupled into two Silicon-based avalanche photon diodes (APD) with a timing resolution of $t_{\mathrm{Res}} \approx 350\,ps$ and the coincidence events are measured on the time-correlation measurement setup (SPC130). For the autocorrelation measurement, the collected photons (exciton or biexciton) are directly guided to a multi-mode fiber based 50:50 beam splitter and then the second order autocorrelation function $g^2(\tau)$ is measured.

Supplementary References:

1. Tonndorf, P. Schmidt, R. *et al.* Single-photon emission from localized excitons in an atomically thin semiconductor. *Optica.* **2,** 347–352 (2015).

2. He, Y. M., *et al.* Phonon induced line broadening and population of the dark exciton in a deeply trapped localized emitter in monolayer WSe_2. *Optics Express.* **24,** 8066–8073 (2016)