

# Imaging Orientation of a Single Molecular Hierarchical Self-Assembled Sheet: The Combined Power of a Vibrational Sum Frequency Generation Microscopy and Neural Network.

Jackson C. Wagner<sup>#,1</sup>, Zishan Wu<sup>#,1</sup>, and Wei Xiong<sup>1,2,3\*</sup>

<sup>1</sup>*Department of Chemistry and Biochemistry, UC San Diego, La Jolla, CA, 92093*

<sup>2</sup>*Materials Science and Engineering Program, UC San Diego, La Jolla, CA, 92093*

<sup>3</sup>*Department of Electrical and Computer Engineering, UC San Diego, La Jolla, CA 92093*

## Abstract

In this work, we determined the tilt angles of molecular units in hierarchical self-assembled materials on a single-sheet level, which were not available previously. This was achieved by developing a fast linescanning vibrational sum frequency generation (VSFG) hyperspectral imaging technique in combination with neural network analysis. Rapid VSFG imaging enables polarization resolved images on a single sheet level to be measured within a short time period, circumventing technical challenges due to long term optical setup instability. The polarization resolved hyperspectral images were then used to extract the supramolecular tilt angle of a self-assembly through a set of spectra-tilt angle relationships which were solved through neural network techniques. This unique combination of both novel techniques offers a new pathway to resolve molecular level structural knowledge of self-assembled materials. Understanding these properties can further drive self-assembly design from a bottom-up approach for applications in biomimetic and drug delivery researches.

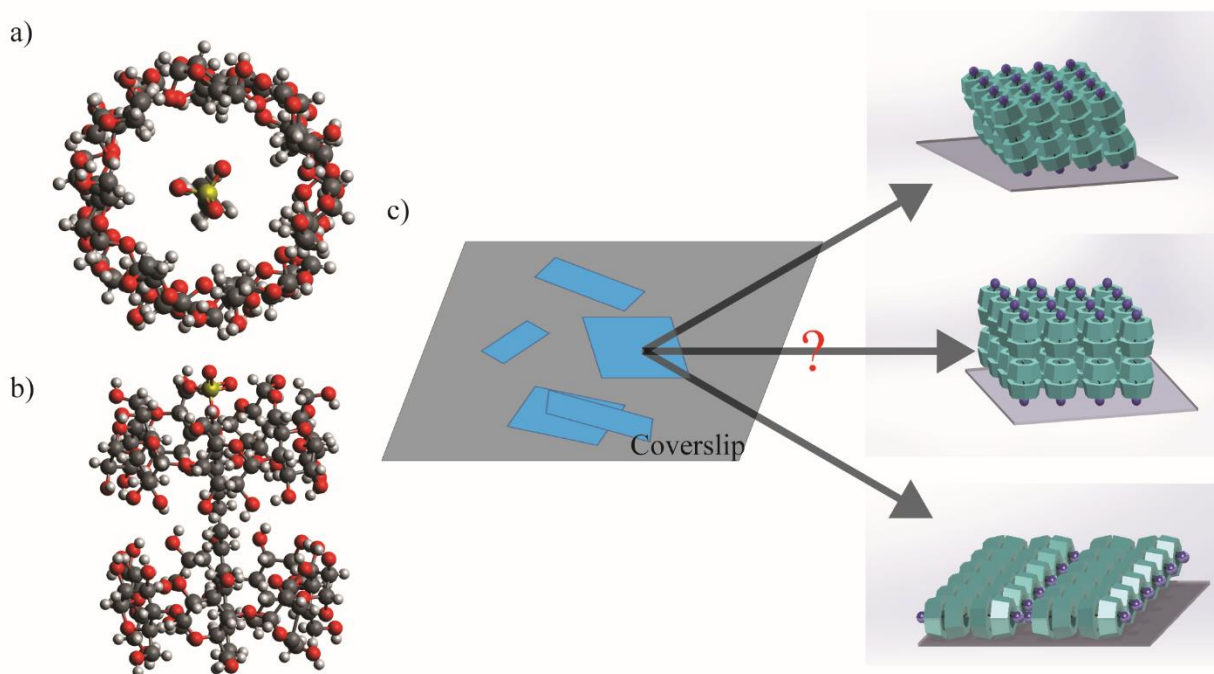
## Introduction

Molecular self-assemblies (MSAs) are a class of materials that spontaneously organize from individual molecular subunits into an ordered structure without templates or external guidance.<sup>1-6</sup> MSAs can have larger architectures that maintain the high ordering and orientation of the smaller structures, known as hierarchical organization.<sup>2,6-10</sup> In both natural and synthetic materials, even when formed from identical subunits, different hierarchical organization can lead to various material functions.<sup>11,12</sup> For example, the diverse structures of natural collagen enable them to assume different tissue functions such as bone, skin etc.<sup>13</sup> In particular, bone possesses a twelve-level hierarchical structure from collagen fibrils to the macroscopic fractal-like architectures, affording it both high stiffness and toughness, properties often considered mutually exclusive.<sup>14</sup> Moreover, liquid crystal displays show different optical properties based on the orientation of the building blocks.<sup>15</sup> Thus, it is feasible to design and manipulate materials functions through hierarchical organizations however, to do so, it is necessary to understand and control the relative positions and orientations of the subunits within the MSAs.<sup>16,17</sup>

Specifically, the orientation of building blocks within a MSA are important to its functions. Applications can be found in chemistry, such as alkanethiol self-assembled monolayers, which are used as an active layer in molecular electronics. The tilting of the alkanethiol chains with respect to surface normal can open additional tunneling pathways and change interfacial dipole properties,

altering the electron transport properties.<sup>18–20</sup> In biomaterials, mechanical properties often depend on the hierarchical subunit orientation. For example, the longitudinal modulus of nacre is higher than the transverse one because of oriented tiles in the hierarchical structures.<sup>21</sup> It has also been demonstrated that surface wettability is correlated with the tilting of subunits<sup>22</sup>, which could further affect protein adsorption and cell adhesion.<sup>23</sup>

An interesting, recent development in MSAs is a lattice self-assembly composed of beta-cyclodextrin ( $\beta$ -CD) and sodium dodecyl sulfate (SDS) in a 2:1 ratio, formed through intermolecular forces, especially hydrogen bonds.<sup>11</sup> This MSA assumes a variety of morphologies depending on the concentration of SDS and  $\beta$ -CD in water. We will refer to this MSA as SDS@2 $\beta$ -CD herein. The primary subunit of the SDS@2 $\beta$ -CD self-assembly is the supramolecule comprised of two  $\beta$ -CD molecules penetrated by one SDS molecule (Figure 1a and b). These subunits form highly ordered and oriented, rhombic nano-thickness sheets (referred to as a nanosheet herein) that can fold into larger mesoscopic architectures such as, lamella sheets, microtubules, rhombic dodecahedra, and micelles among others.<sup>11</sup> This MSA has drawn much attention because of the biomimetic nature of many of its mesoscopic architectures in addition to its broad application such as wastewater treatment<sup>24</sup>, drug delivery<sup>25</sup>, and optoelectronics.<sup>26</sup> However, the structural details of the nanosheets – the most basic formation of the self-assembly that folds into all other higher order molecular architectures, are not fully understood.<sup>11</sup> Through small angle X-ray scattering (SAXS) the inter-supramolecular distance was determined to be 1.52 nm but, open questions remain regarding the relative orientation of the supramolecule in SDS@2 $\beta$ -CD. Because molecular orientations often act as critical factors to MSA's functions, it is pertinent to understand whether and how the supramolecules in the self-assembled sheets are tilted (Figure 1c), which could potentially indicate the van der Waals contact, the structural symmetry or the materials macroscopic properties.<sup>18,22,27,28</sup>



**Figure 1 Structure of SDS@2 $\beta$ -CD.** a) Top down view of the SDS@2 $\beta$ -CD MSA building blocks b) Side view of the MSA c) Microscopic formations of the SDS@2 $\beta$ -CD which hierarchically oriented to form rhombically shaped nanosheets. The orientation of the MSA relative to surface normal is unknown.

In this work, we determined the orientation of the supramolecules in SDS@2 $\beta$ -CD through the development of a fast line-scanning vibrational sum-frequency generation (VSFG) microscopy in combination with neural network data analysis.<sup>29,30</sup> Building on our previous efforts in ultrafast hyperspectral imaging techniques,<sup>31,32</sup> this new technical advancement enabled imaging single SDS@2 $\beta$ -CD sheets hyperspectrally with eight different polarization combinations. Then, to reveal structural information, we applied a neural network method to solve a set of equations which relate the supramolecular tilt angle to the 2<sup>nd</sup> order susceptibility of different polarizations. We found that the supramolecules were tilted relative to normal of the nanosheets by  $\sim 23^\circ$ . This provides insight into how to design future materials as well as offer details as to what role hierarchical orientation played in MSAs. This structural knowledge is revealed through the combination of rapid acquisition of hyperspectral imaging and neural networks. Both are crucial to extract these parameters with the former minimizing long term laser drift issues, and the latter offering a route to solve a complex set of equations that were otherwise difficult to be solved.

## Results and Discussion

### *Linescanning VSFG imaging*

The hyperspectral microscope was based on VSFG spectroscopy, a second-order nonlinear optical phenomenon. VSFG employs a broadband IR pulse (290 fs duration) centered at  $\sim 2850 \text{ cm}^{-1}$  to excite CH vibrational modes of the SDS@2 $\beta$ -CD system, which is followed by a narrowband

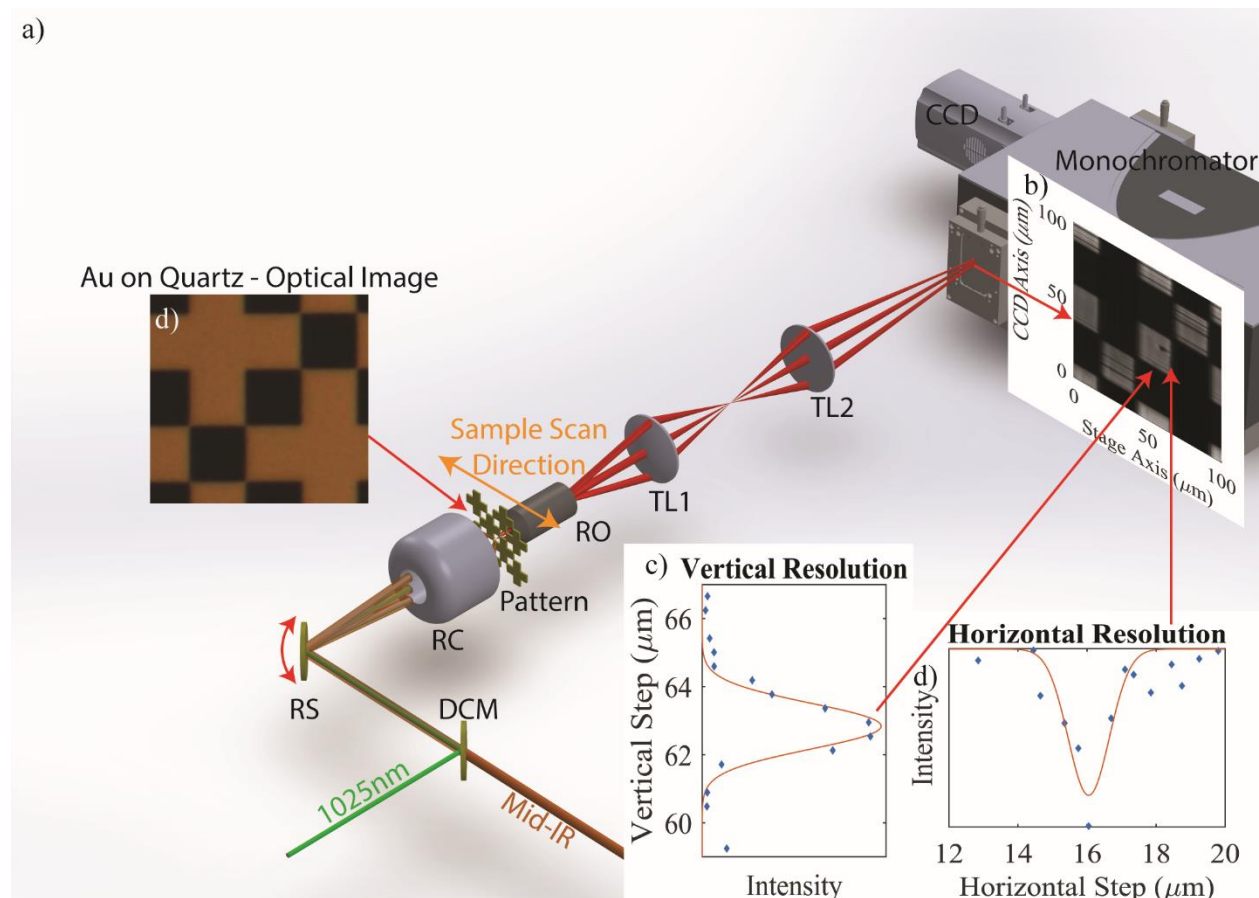
1025nm near IR pulse to upconvert the vibrational coherence to a higher frequency macroscopic polarization emitted as the VSFG signal. As an even-order nonlinear optical process, only non-centrosymmetric systems produce VSFG responses, such as the air/liquid, air/solid and solid/liquid interfaces<sup>29,31,40–42,32–39</sup> with a powerful demonstration of the techniques capabilities found in its application to the air/water interface.<sup>43–45</sup> Another class of systems that can be studied by VSFG that have not been paid much attention are materials without inversion centers.<sup>46–55</sup> Many MSAs are non-centrosymmetric systems, and widely exist in nature, such as collagen<sup>49,50</sup> and amyloid fibers<sup>56</sup>; as well as in artificial materials<sup>57</sup> used in drug delivery<sup>58,59</sup>, metal-organic-frameworks<sup>60,61,62</sup> and piezoelectric crystals.<sup>63</sup> Thus, VSFG, which probes the molecular vibrations of the subunits in the non-centrosymmetric MSAs, could similarly reveal MSA structures with molecular level detail.

A big challenge in using VSFG spectroscopy to probe MSA is that most MSAs only form nano-to micrometer sizes domains, while the illumination area of VSFG spectroscopy is generally around 100 $\mu$ m-by-100 $\mu$ m. Thus, traditional VSFG spectroscopy will probe multiple MSAs in an ensemble-averaged manner which does not accurately reflect the molecular structure of an individual MSA.<sup>32</sup> The development of VSFG spectroscopy into a hyperspectral imaging technique<sup>32,48–50,52,53,64–67</sup> overcame this challenge with 1-micron or submicron resolution being obtained, which offered a platform that could resolve multiple micron-sized MSAs individually.

Additionally, to retrieve molecular orientations, it was necessary to measure VSFG images with multiple polarization combinations.<sup>49,50,65,68</sup> Though theoretically feasible, it was practically prohibited in our previous point-scanning VSFG microscope<sup>31,32</sup> since it took nearly 4 hours to scan a 100 $\mu$ m-by-100 $\mu$ m image and would take at least 30 hours to collect all eight polarization combinations. The long acquisition time would introduce fluctuations in opto-mechanics and laser output which further complicates data analysis. To decrease acquisition time, we hybridized the line-scanning technique with our existing VSFG microscope. A line scanning method was first implemented in VSFG microscopy by the Ge and Potma groups using a photomultiplier tube as a detector,<sup>49,50,64,65</sup> which required scanning the source IR frequency to gain spectral information. The integration of linescanning with a CCD detector reported here, enabled simultaneous measurement of a set of spectra of a vertical line, maximizing the information measured by the 2D detector.

Specifically, we combined the mid-IR and upconversion beam collinearly using a dichroic mirror and then steered the angle of the beam along the vertical axis using a fixed-frequency resonant scanner of 325Hz. The scanned beams were focused through a purely reflective Schwarzschild objective (Pike Technologies Inc 0.70NA) onto the sample area. At the sample plane, the incoming beams were transformed into a vertical line of illumination which generated a line of VSFG signal that was imaged by another objective lens (Zeiss 0.75NA) and further magnified by a home-built tube lens to match the vertical dimension of the CCD. The signal was frequency resolved finally by a spectrograph horizontally(Figure 2a). Thus, the CCD measured the spectra dispersed along the horizontal axis and the spatial profile along vertical axis. 2D images were acquired by scanning the sample in the horizontal direction with an automated mechanical stage. This improvement decreased image acquisition times by 10, compared to the point-by-point microscope, to 20 mins for a 100  $\mu$ m -by-100  $\mu$ m image.

The VSFG images obtained from the line-scanning microscope captured the same geometric features of the optical image of plated gold patterns on quartz substrates (Fig. 2a and b). Depending on the scan angle of the resonant mirror and magnification of the tube lens, we achieved a 100  $\mu\text{m}$  vertical field of view. The vertical and horizontal resolutions were 1.2 and 1.6  $\mu\text{m}$  respectively (Figure 2c and d), and the total magnification was 66.



**Figure 2 Linescanning VSFG microscope.** a) Schematic of the setup. b) VSFG image of quartz substrate target image. c) Vertical and horizontal resolution generated by taking the derivative of the corresponding cut. d) Optical image of gold target on quartz substrate in a similar area.

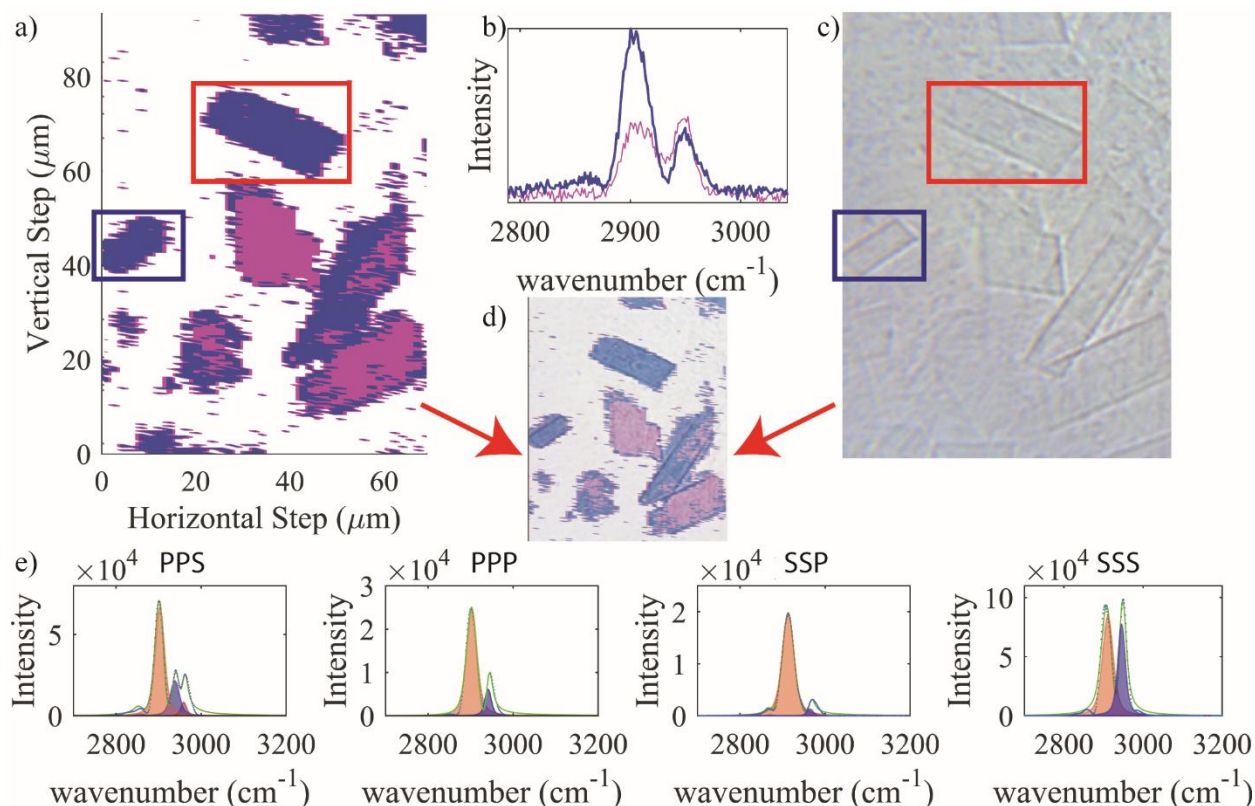
### ***Polarization resolved hyperspectral imaging on SDS@2 $\beta$ -CD***

To measure the tilt angle of supramolecules in SDS@2 $\beta$ -CD, we applied the linescanning VSFG microscope to image single SDS@2 $\beta$ -CD sheets. Single SDS@2 $\beta$ -CD sheets were carefully prepared by spin-coating 5  $\mu\text{L}$  of the SDS@2 $\beta$ -CD suspension in water, synthesis details in ESI, at 10,000 rpm spin rate onto glass coverslips, to minimize sheet stacking which would complicate analysis later as well as affect the image quality. An example of hyperspectral images of SPS polarization combination (left to right were polarization of VSFG, upconversion and IR beams) was shown in Fig. 3a. Edges of a single SDS@2 $\beta$ -CD sheet could be clearly seen as well as the rotational orientation of the rhombic shape inherent to the self-assembly, and the VSFG and optical images agreed well. (Fig. 3c and d) We noted that there was a significant improvement in the

quality of VSFG images and their agreement with the optical images of the SDS@2 $\beta$ -CD, comparing to the images in our previous publications.<sup>31,32</sup> This improvement is achieved by (1) an improved sample deposition method to prepare single sheets instead of multiple sheets stacking on top of each other; (2) that the fast linescanning VSFG microscope allowed optimizations of image quality within a short image acquisition time and large field of view, which was not available before.

To further disentangle spectral features of the VSFG image, spectral maps were generated using the MatLab hyperspectral imaging toolbox. Two unique spectra for the SPS polarization combination were identified, highlighted in blue and magenta in Fig. 3a with corresponding spectra shown in Fig. 3b. Clearly, there were two types of sheets and spectra with the sheets with magenta color coding is due to sheet stacking. In the following, we only analyzes the areas highlighted with blue spectra to extract tilt angles, which were single sheets.

To extract molecular orientations, all eight lab frame VSFG polarization combinations (SSS, SSP, SPS, PSS, SPP, PSP, PPS, PPP) were collected. Four representative polarization resolved VSFG spectra of a single sheet (blue rectangle labeled area in Figure 3a and c) area shown in Figure 3e (additional spectra shown in ESI). Each spectrum was fitted with multiple Voigt functions (shaded area, Fig. 3e, fitting methods see Methods, detailed fitting parameters are summarized in ESI). In all spectra, we identified a clear peak at the 2910  $\text{cm}^{-1}$  position, the  $-\text{CH}_3$  stretching mode, which we used for the orientation analysis below.



**Figure 3 VSFG hyperspectral image and spectral analysis of the SDS@2 $\beta$ -CD.** a) Polarization resolved hyperspectral VSFG image of SDS@2 $\beta$ -CD. Purple and pink colors represent areas

where different spectra reside and the corresponding spectra are plotted in **b**). The sheets in the blue and red boxes are analyzed explicitly below to extract the supramolecule tilt angles. **c**) optical image of the same area as that in **a**. **d**) VSFG hyperspectral image overlayed with optical image of identical area. **e**) from left to right: PPS, PPP, SSP, SSS polarization resolved spectra. All spectra had a dominant feature centered at approximately 2910 cm<sup>-1</sup>. The spectra were fitted with multiple Voigt functions, which were represented by the shaded areas.

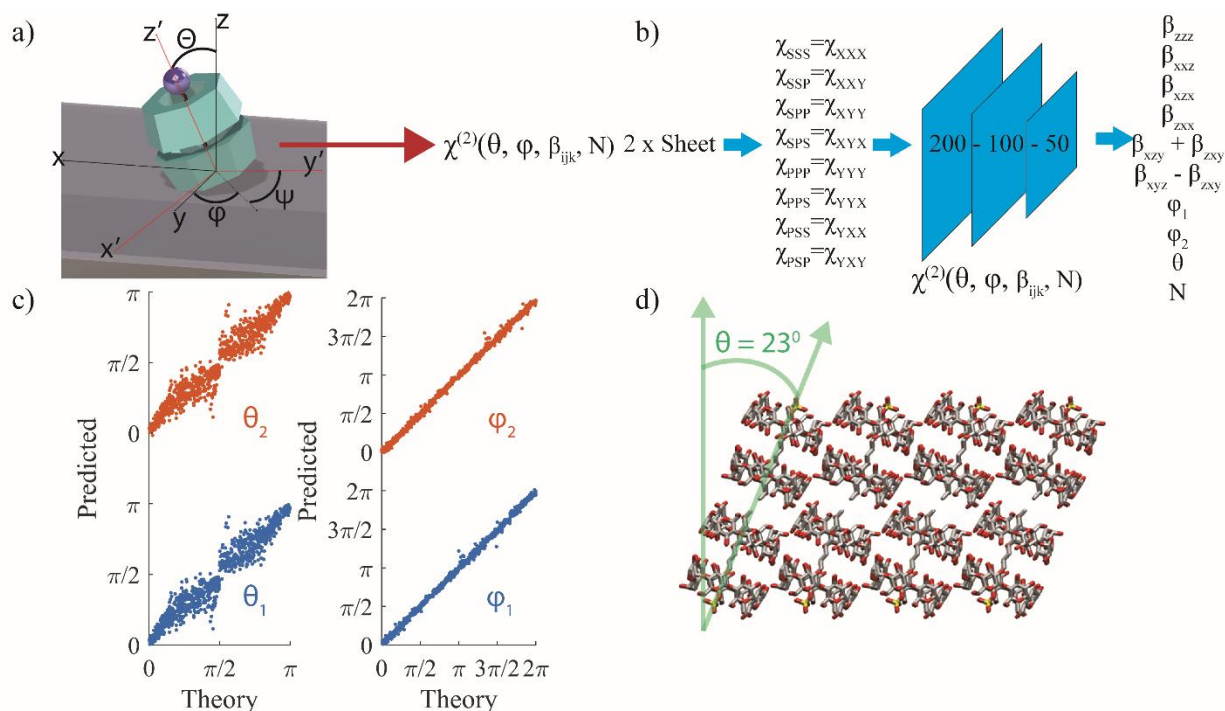
### ***Theoretical basis of orientation analysis of a single MSA sheet***

Theoretically, VSFG spectra with different polarization combinations, which were related to the lab frame 2<sup>nd</sup> order susceptibility,  $\chi^{(2)}_{\text{ijk}}$ , could be expressed in terms of the molecular orientation (such as tilt angle and in-plane rotation) in the lab frame and molecular frame hyperpolarizability tensor,  $\beta_{\text{ijk}}$ , through an Euler rotation.<sup>49,65,69,70</sup> In our measurement, the z-axis of the lab frame and the MSA frame were identical (i.e. sheets lying flat), and the x-y axis of the MSA was only rotated away from their counter parts in the lab frame. Thus, in principle, we could extract the relative supramolecular orientation in the MSA using the lab frame VSFG spectral intensity.

Because the SDS@2 $\beta$ -CD supramolecule has C<sub>7</sub> symmetry, it has 13  $\beta_{\text{ijk}}$  elements, of which only 7 are nondegenerate,  $\beta_{\text{zzz}}, \beta_{\text{xxz}}, \beta_{\text{xzx}}, \beta_{\text{zxx}}, \beta_{\text{xzy}}, \beta_{\text{zyx}}, \beta_{\text{xyz}}$  (ESI).<sup>32</sup> Here, we also do not assume Kleinman symmetry, which has been previously reported as non-universal.<sup>71,72</sup> Then, for the lab frame, with the NA of the condenser objective being 0.7, the axial z-component can be neglected<sup>73</sup>, which renders 8 independent 2<sup>nd</sup> order measurement  $\chi^{(2)}_{\text{xxx}} = \chi^{(2)}_{\text{sss}}, \chi^{(2)}_{\text{xxz}} = \chi^{(2)}_{\text{ssp}}, \chi^{(2)}_{\text{xyx}} = \chi^{(2)}_{\text{spz}}, \chi^{(2)}_{\text{yxx}} = \chi^{(2)}_{\text{pss}}, \chi^{(2)}_{\text{xyy}} = \chi^{(2)}_{\text{spp}}, \chi^{(2)}_{\text{yyx}} = \chi^{(2)}_{\text{pss}}, \chi^{(2)}_{\text{yyy}} = \chi^{(2)}_{\text{ppp}}$ . Through an Euler rotation (A graphical depiction of the rotation is shown in Fig. 4a.) and assuming the twist angle,  $\psi$ , is arbitrary, the lab frame  $\chi^{(2)}$  could be expressed as a function of  $\beta_{\text{ijk}}$  and solid angles, resulting in a set of 8 equations. One of the output equations of the Euler rotation is provided, **Eq. 1**, as an example and the rest were listed in the ESI. As evident from Eq. 1 and ESI Eq. 5 - 12, three hyperpolarizability elements were not completely independent as they appeared as  $\beta_{\text{xzy}} + \beta_{\text{zyx}}$  and  $\beta_{\text{xyz}} - \beta_{\text{zxx}}$  grouped terms in all equations. Therefore, 7 nondegenerate hyperpolarizability elements were grouped down to 6 independent terms simplifying our set of eight equations with 8 inputs to 8 unknowns (6  $\beta_{\text{ijk}}$  grouped terms, and two solid angles).

$$\chi_{\text{PPS}}^{(2)} = \frac{1}{16} [ (4 \cos(3\varphi) \sin^3(\theta) + \cos(\varphi) (13 \sin(\theta) + \sin(3\theta)) \beta_{\text{xxz}} - 8 \sin(\varphi) (\sin(2\theta) (\beta_{\text{xzy}} + \beta_{\text{zyx}}) + \sin^3(\theta) \sin(2\varphi) (\beta_{\text{xzx}} + \beta_{\text{zxx}} - \beta_{\text{zzz}})) ] \quad \text{Eq. 1}$$





**Figure 4** Euler transformation, neural network and extracted tilt angles of the supramolecules in MSAs. **a)** The Euler transformation between the laboratory coordinates (XYZ) and the molecular coordinates (xyz). z-y'-z'' Euler rotation is performed on the molecular coordinates, with  $\phi$  as the in-plane rotation angle,  $\theta$  as the tilt angle, and  $\psi$  as the twist angle. **b)** Schematic illustration of the neural network employed to extract orientation information of the SDS@2 $\beta$ -CD. **c)** Neural network results for the tilt (left) and in-plane rotation (right) angles. **d)** Visualization of the tilted supramolecule subunits forming a sheet determined by the neural network results.

To extract molecular orientations in the MSA frames, we need to solve the equations based on the experimentally measured  $\chi^{(2)}$  to extract  $\theta$ ,  $\phi$  and  $\beta_{ijk}$ . To enhance the consistency of the result, we analyzed the signal on two single sheets and assumed that the supramolecular subunit in all single sheets had the same tilt angle. Two single sheets, i.e., without stacking and with visibly different orientations, were selected as a simplified scenario (identified by red and blue rectangles in Fig. 3c and a. See details in SI Fig. 2). The  $\chi^{(2)}$  of the 2910  $\text{cm}^{-1}$  peak of two sheets were extracted by the fitting results illustrated in Figure 3e, for the orientation analysis. Based on the optical images, (Fig. 3 and S2), we determined that the two single sheets are rotated approximately  $60^\circ$  in the XY plane relative to one another. Therefore, we further restrained our model with  $\phi_2 = \phi_1 + 60$  degrees. As a result, for two sheets, we have 16 input values (2 x 8 different polarization combinations) and 9 outputs: 6 independent variables consisting of  $\beta_{ijk}$ , the in-plane rotation  $\phi_1$ , tilt angle  $\theta$  and the relative coverage ratio  $N$  between sheets.

### Neural networks to train the solvers and extract tilt angles

This equation set was solvable however, traditional solvers required higher computational load (such as memory) for numerous iterations to obtain solution sets and re-calculation, and had low



tolerance to noise, which was always present experimental data.<sup>74,75</sup> These limitations prohibited us from obtaining reasonable results so, we turned to a neural network approach.<sup>76,77</sup> The training set was created by randomly generating angles and hyperpolarizabilities (training output) and then supplying the values to the polarization-dependent equations (ESI Eq. 5-12) to calculate the corresponding susceptibilities (training input) (Fig.4b). The training set was used to train the relationships between the training input and outputs. The trained model was then tested with another data set generated via a similar random process, which investigated how well the model could solve for molecular orientations when supplied with susceptibility values that were not part of the training input. Fig. 4c displayed the correlation between Euler angles, i.e., molecular orientations, predicted by our model which captured the true values in the test data set well, resembling a  $y=x$  relationship. Thus, our neural network based orientation solver was appropriate in predicting the molecular orientations.

Finally, we extracted molecular orientation using this method, by supplying it with the experimentally determined susceptibility values. Since phase information was not retrieved in our homodyne experiment, we enumerated the signs of all 16 susceptibility values when supplying them to the model and selected the predicted results with the smallest mean squared error of susceptibilities. It was predicted that the SDS@2 $\beta$ -CD subunit was tilted away from the lab frame z-axis (axis of light propagation) by 23 degrees (Fig. 4d). From the polarization-dependent equations we could see that if the tilt angle was 0°, all susceptibility terms on the left side of the equations would be zero, which did not agree with the strong SFG signal, implying non-zero susceptibilities of the SDS@2 $\beta$ -CD system. On the other hand, using the literature reported in-plane unit cell parameters of the SDS@2 $\beta$ -CD system<sup>11</sup>, we could visually demonstrate that when the tilt angle was 30 degrees, the space between subunits was tight and when the tilt angle was 45-60 degrees, SDS@2 $\beta$ -CD subunits would collide with each other (Fig. S6). Hence, the SDS@2 $\beta$ -CD subunits were tilted slightly at 23 degrees retrieved with the neural network model was appropriate and consistent with exiting structural knowledge of the nanosheets. Comparable to the well-studied system of self-assembled monolayer on metal substrates, we believed the driving force of the tilting could be the interplay of intermolecular interactions (such as hydrogen bonds) among subunits and binding behaviors between subunits and substrate.<sup>22,27,78,79</sup>

It had been widely studied that the tilt of molecules within monolayers commonly existed in self-assembled materials and potentially adjusted the conduction<sup>80,81</sup>, wetting<sup>82,83</sup>, or mechanical properties<sup>80,84,85</sup>. With SDS@2 $\beta$ -CD as an important biomimetic motif, the tilt angle resolved in our neural network approach sheds light on how the subunits pack within the self-assembly and provides a protocol to study other MSA systems structure-properties correlations. Moreover, as the tilting is influenced by the interaction of the subunits and hydration level, future works on humidity dependent packing of the system might unravel how the chemical environments affect the self-assembly process and could be very interesting and significant for the drug delivery field<sup>86</sup> as the release of target molecules highly relies on the biological environment.

## Conclusion

The molecular self-assembly formed from a mixture of  $\beta$ -CD and SDS in water was analyzed using a linescanning hyperspectral VSFG microscope and neural network. A 1-D resonant scanner

coupled to a CCD/spectrograph increases image collection speed 10 fold with simultaneous spectral information. This development enables polarization resolved VSFG images of single MSA sheets, which were analyzed by a neural network approach. The analysis revealed that the supramolecular building blocks are tilted at around  $23^\circ$  in the SDS@ $2\beta$ -CD MSA frame. Such information could help us further understanding the structure and intermolecular interactions of other biomimetic morphologies that the nanosheets construct. This information became available because of the power of VSFG microscopy to extract spatially resolved, spectral information of the nanosheets.

## Experimental Section

### VSFG Line Scanning Microscope

Our VSFG linescanning microscope is shown in Fig 2. Laser pulses for the microscope is provided by an 100 kHz Yb based cavity femtosecond laser (Light Conversion, Carbide) centered at 1025nm. The output from the Carbide is used to pump an optical parametric amplifier (OPA) (Light Conversion, Orpheus-HP) centered at 3500 nm which covers the CH stretching vibration region of interest. The residual 1025 nm beam is used as the up-conversion and is first conditioned spectrally by directing it through a folded 4f pulse shaper. The frequency narrowed 1025 nm beam is then focused through an 8  $\mu$ m spatial filter followed by a  $\lambda/2$  waveplate. The mid IR (MIR) light is steered towards a delay stage filter followed by a  $\lambda/2$  waveplate and spatially overlapped with the up-conversion with a customized dichroic mirror that is transparent to MIR and reflective to near IR (NIR). The overlapped beams are then guided to a 1-D resonant scanner (EOPC) and focused onto the sample mounted to a 2D piezo stage (MadCity Labs) by a purely reflective 20x Schwarzschild objective (PIKE Technologies Inc. PN 891-0001) which acts as a condenser. The emitted nonlinear VSFG signal is collected by an infinity corrected, 20x refractive microscope objective (Zeiss, Fluar 20x/0.75 WD=0.6) and passes through a linear polarizer (ThorLabs). The polarization resolved signal is then analyzed using a Shamrock 500i spectrograph (Andor) coupled to a Newton CCD (Andor).

### 2B-CD@SDS Synthesis

B-CD sheets are synthesized by adding B-CD and SDS at a molar ratio of 2:1 in DI water until the percent concentration is 10% m/m. The suspension is then heated to clarity and cooled to room temperature overnight.  $\text{CuCl}_2$  is added with sheets fully forming approximately 3-5 days later. Isolated sheet samples with linear dimensions on the 10s of micron scale are produced by drop casting 5  $\mu$ L of the sheet suspension onto a 15 mm \* 15 mm \* 0.170 mm microscope slide spinning at 10000 rpm. The sheets are transparent, but the silhouette can be observed with a standard optical microscope (Fig3).

### Neural network model

Keras in Python is employed to set up the neural network model. A layered neural network modified from Github repository<sup>77</sup> is built with a 200-100-50 node structure and a hyperbolic tangent activation function between layers (Fig. 4a). Training data set contains 100,000 input vectors and the corresponding output vectors correlated via the polarization-dependent equations. Similar to simple trigonometric function where  $\theta = 0$  and  $\theta = \pi$  both satisfy  $\sin(\theta) = 0$ , one input vector (susceptibility values) in our model could also have multiple output vectors (different in-

plane rotation, tilt angle, hyperpolarizabilities combinations) at the same time. Hence, in-plane rotation angle is divided into  $[0, \pi)$  and  $[\pi, 2\pi)$  intervals and the tilt angle is divided into  $[0, \pi/2)$  and  $[\pi/2, \pi)$  intervals to differentiate these output vectors. Training is run with an epoch size of 1000 and a batch size of 100. Mean squared error of output vector is used to monitor the deviation of prediction from true values. During the training, only 90% of the dataset is used for actual modeling and the rest 10% of the training data is separated out to testify whether the model can generalize to unseen data. In such way one can mitigate the overfitting, as can be seen in the small and close loss values obtained in both training and test process (SI Figure 7).

After validating that our model is capable of predicting the tilt angle (Fig. 4c), we supply the model with experimentally determined susceptibilities (see ESI for details on spectral fitting) to extract the tilt angle. Since there is no phase information in our measurement, we enumerate the sign of each susceptibility values (total of  $2^{16}$  combinations) and provide all these as input vectors to our neural network. By calculating the mean squared errors of the predicted normalized susceptibility values and the experimentally determined ones, the smallest mean squared error we can obtain is 0.02, whose corresponding predicted tilt angle is  $23^\circ$  (SI Table 2).

## Acknowledgement

Z.W. and W. X. thank Prof. Sichun Gao's insightful discussion about neural network. W. X. thanks Prof. Hong-fei Wang sharing his knowledge on the Klienmann symmetry. The instrument development is supported by NSF CHE-1828666. Z.W. and J.C.W. are supported by National Institute of Health, National Institute of General Medical Sciences, 1R35GM138092-01.

## Contact Information

w2xiong@ucsd.edu

## References

- 1 S. Manchineella and T. Govindaraju, *Chempluschem*, 2017, **82**, 88–106.
- 2 C. Yuan, W. Ji, R. Xing, J. Li, E. Gazit and X. Yan, *Nat. Rev. Chem.*, 2019, **3**, 567–588.
- 3 G. M. Whitesides, J. P. Mathias and C. T. Seto, *Science (80-. )*, 1991, **254**, 1312–1319.
- 4 S. Zhang, *Nat. Biotechnol.*, 2003, **21**, 1171–1178.
- 5 R. Yi, Y. Mao, Y. Shen and L. Chen, *J. Am. Chem. Soc.*, 2021, **143**, 12897–12912.
- 6 Z. Li, B. Cai, W. Yang and C.-L. Chen, *Chem. Rev.*, 2021, **121**, 14031–14087.
- 7 C. V. Kulkarni, *Nanoscale*, 2012, **4**, 5779.
- 8 B. B. Gerbelli, S. V. Vassiliades, J. E. U. Rojas, J. N. B. D. Pelin, R. S. N. Mancini, W. S. G. Pereira, A. M. Aguilar, M. Venanzi, F. Cavalieri, F. Giuntini and W. A. Alves, *Macromol. Chem. Phys.*, 2019, **220**, 1900085.
- 9 O. Ikkala and G. ten Brinke, *Chem. Commun.*, 2004, 2131.
- 10 M. S. Ganewatta, Z. Wang and C. Tang, *Nat. Rev. Chem.*, 2021, **5**, 753–772.

- 11 S. Yang, Y. Yan, J. Huang, A. V. Petukhov, L. M. J. Kroon-Batenburg, M. Drechsler, C. Zhou, M. Tu, S. Granick and L. Jiang, *Nat. Commun.*, 2017, **8**, 1–7.
- 12 A. Gautieri, S. Vesentini, A. Redaelli and M. J. Buehler, *Nano Lett.*, 2011, **11**, 757–766.
- 13 D. E. Birk and P. Brückner, in *The Extracellular Matrix: an Overview*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 77–115.
- 14 N. Reznikov, M. Bilton, L. Lari, M. M. Stevens and R. Kröger, *Science (80-. )*, , DOI:10.1126/science.aao2189.
- 15 J. A. Castellano, *Liquid Gold*, WORLD SCIENTIFIC, 2005.
- 16 L. Feng, K.-Y. Wang, G. S. Day and H.-C. Zhou, *Chem. Soc. Rev.*, 2019, **48**, 4823–4853.
- 17 D. L. Caulder and K. N. Raymond, *Acc. Chem. Res.*, 1999, **32**, 975–982.
- 18 T. Frederiksen, C. Munuera, C. Ocal, M. Brandbyge, M. Paulsson, D. Sanchez-Portal and A. Arnau, *ACS Nano*, 2009, **3**, 2073–2080.
- 19 Y. Qi, I. Ratera, J. Y. Park, P. D. Ashby, S. Y. Quek, J. B. Neaton and M. Salmeron, *Langmuir*, 2008, **24**, 2219–2223.
- 20 F. W. DelRio, K. L. Steffens, C. Jaye, D. A. Fischer and R. F. Cook, *Langmuir*, 2010, **26**, 1688–1699.
- 21 M. A. Meyers, J. McKittrick and P.-Y. Chen, *Science (80-. )*, 2013, **339**, 773–779.
- 22 Y. T. Tao, M. T. Lee and S. C. Chang, *J. Am. Chem. Soc.*, 1993, **115**, 9547–9555.
- 23 Y. Arima and H. Iwata, *Biomaterials*, 2007, **28**, 3074–3082.
- 24 N. Morin-Crini, M. Fourmentin, S. Fourmentin, G. Torri and G. Crini, *Environ. Chem. Lett.*, 2019, **17**, 683–696.
- 25 R. Liao, P. Lv, Q. Wang, J. Zheng, B. Feng and B. Yang, *Biomater. Sci.*, 2017, **5**, 1736–1745.
- 26 X. Zhang, X. Ma, K. Wang, S. Lin, S. Zhu, Y. Dai and F. Xia, *Macromol. Rapid Commun.*, 2018, **39**, 1800142.
- 27 A. Ulman, *Chem. Rev.*, 1996, **96**, 1533–1554.
- 28 S. A. Hughes, F. Wang, S. Wang, M. A. B. Kreutzberger, T. Osinski, A. Orlova, J. S. Wall, X. Zuo, E. H. Egelman and V. P. Conticello, *Proc. Natl. Acad. Sci.*, 2019, **116**, 14456–14464.
- 29 H. Wang and W. Xiong, *Annu. Rev. Phys. Chem.*, 2021, **72**, 279–306.
- 30 M. J. Blake, B. A. Colon and T. R. Calhoun, *J. Phys. Chem. C*, 2020, **124**, 24555–24565.
- 31 H. Wang, J. C. Wagner, W. Chen, C. Wang and W. Xiong, *Proc. Natl. Acad. Sci.*, 2020, **117**, 23385–23392.
- 32 H. Wang, W. Chen, J. C. Wagner and W. Xiong, *J. Phys. Chem. B*, 2019, **123**, 6212–6221.
- 33 K. E. Watts, T. J. Blackburn and J. E. Pemberton, *Anal. Chem.*, 2019, **91**, 4235–4265.
- 34 N. Takeshita, M. Okuno and T. Ishibashi, *Phys. Chem. Chem. Phys.*, 2017, **19**, 2060–2066.

- 35 C. Wang, Y. Li and W. Xiong, *J. Mater. Chem. C*, 2020, **8**, 12062–12067.
- 36 W. Sung, C. Müller, S. Hietzschold, R. Lovrinčić, N. P. Gallop, A. A. Bakulin, S. Nihonyanagi and T. Tahara, *Mater. Horizons*, 2020, **7**, 1348–1357.
- 37 D. Elsenbeck, S. K. Das and L. Velarde, *Phys. Chem. Chem. Phys.*, 2017, **19**, 18519–18528.
- 38 P. K. Gupta and M. Meuwly, *Faraday Discuss.*, 2013, **167**, 329.
- 39 Y. Rao, N. J. Turro and K. B. Eisenthal, *J. Phys. Chem. C*, 2009, **113**, 14384–14389.
- 40 Y. Rao, D. Song, N. J. Turro and K. B. Eisenthal, *J. Phys. Chem. B*, 2008, **112**, 13572–13576.
- 41 Y. Wu, W. Li, B. Xu, X. Li, H. Wang, V. F. McNeill, Y. Rao and H.-L. Dai, *J. Phys. Chem. Lett.*, 2016, **7**, 2294–2297.
- 42 G.-H. Deng, Y. Qian and Y. Rao, *J. Chem. Phys.*, 2019, **150**, 024708.
- 43 X. D. Zhu, H. Suhr and Y. R. Shen, *Phys. Rev. B*, 1987, **35**, 3047–3050.
- 44 Y. R. Shen, *Nature*, 1989, **337**, 519–525.
- 45 R. Kusaka, T. Ishiyama, S. Nihonyanagi, A. Morita and T. Tahara, *Phys. Chem. Chem. Phys.*, 2018, **20**, 3002–3009.
- 46 D. Wanapun, U. S. Kestur, D. J. Kissick, G. J. Simpson and L. S. Taylor, *Anal. Chem.*, 2010, **82**, 5425–5432.
- 47 Z. Song, S. Sarkar, A. D. Vogt, G. D. Danzer, C. J. Smith, E. J. Gualtieri and G. J. Simpson, *Anal. Chem.*, 2018, **90**, 4406–4413.
- 48 V. Raghunathan, Y. Han, O. Korth, N.-H. Ge and E. O. Potma, *Opt. Lett.*, 2011, **36**, 3891.
- 49 Y. Han, V. Raghunathan, R. R. Feng, H. Maekawa, C.-Y. Y. Chung, Y. Feng, E. O. Potma and N.-H. H. Ge, *J. Phys. Chem. B*, 2013, **117**, 6149–6156.
- 50 Y. Han, J. Hsu, N.-H. Ge and E. O. Potma, *J. Phys. Chem. B*, 2015, **119**, 3356–3365.
- 51 H. Maekawa, S. K. K. Kumar, S. S. Mukherjee and N.-H. Ge, *J. Phys. Chem. B*, 2021, **125**, 9507–9516.
- 52 J. Hsu, A. Haninnen, N.-H. Ge and E. O. Potma, in *Optics in the Life Sciences*, OSA, Washington, D.C., 2015, p. NT4C.4.
- 53 E. O. Potma, in *Optics in the Life Sciences Congress*, OSA, Washington, D.C., 2017, p. NM4C.2.
- 54 C. Leng, H.-C. Hung, S. Sun, D. Wang, Y. Li, S. Jiang and Z. Chen, *ACS Appl. Mater. Interfaces*, 2015, **7**, 16881–16888.
- 55 D. J. Kissick, D. Wanapun and G. J. Simpson, *Annu. Rev. Anal. Chem.*, 2011, **4**, 419–437.
- 56 P. K. Johansson and P. Koelsch, *Biomed. Opt. Express*, 2017, **8**, 743.
- 57 M. Zhao, *Biopolymers*, , DOI:10.1002/bip.23469.
- 58 Q. Lin, C. Owh, J. Y. C. Lim, P. L. Chee, M. P. Y. Yew, E. T. Y. Hor and X. J. Loh, *Accounts Mater. Res.*, 2021, **2**, 881–894.

- 59 M. He, A. Potuck, J. C. Kohn, K. Fung, C. A. Reinhart-King and C.-C. Chu, *Biomacromolecules*, 2016, **17**, 523–537.
- 60 Y. Guo, L. Zhang, K. Wang, A. Yu, S. Zhang and G. Ouyang, *Chem. Commun.*, 2021, **57**, 10343–10346.
- 61 W. Gong, Y. Liu and Y. Cui, *Faraday Discuss.*, 2021, **231**, 168–180.
- 62 J. C. Wagner, K. M. Hunter, F. Paesani and W. Xiong, *J. Am. Chem. Soc.*, 2021, **143**, 21189–21194.
- 63 J. Yang, M. Chen, H. Lee, Z. Xu, Z. Zhou, S.-P. Feng and J. T. Kim, *ACS Appl. Mater. Interfaces*, 2021, **13**, 20573–20580.
- 64 C.-Y. Chung and E. O. Potma, *Annu. Rev. Phys. Chem.*, 2013, **64**, 77–99.
- 65 A. Hanninen, M. W. Shu and E. O. Potma, *Biomed. Opt. Express*, 2017, **8**, 4230.
- 66 C. Katagiri, T. Miyamae, H. Li, F. Yang and S. Baldelli, *Phys. Chem. Chem. Phys.*, 2021, **23**, 4944–4950.
- 67 H. Li and S. Baldelli, *J. Phys. Chem. C*, 2021, **125**, 13560–13571.
- 68 B. M. de Lima, P. L. Hayes and P. M. Wood-Adams, *Langmuir*, 2021, **37**, 10036–10045.
- 69 H.-F. Wang, L. Velarde, W. Gan and L. Fu, *Annu. Rev. Phys. Chem.*, 2015, **66**, 189–216.
- 70 Y. Rao, M. Comstock and K. B. Eisenthal, *J. Phys. Chem. B*, 2006, **110**, 1727–1732.
- 71 C. A. Dailey, B. J. Burke and G. J. Simpson, *Chem. Phys. Lett.*, 2004, **390**, 8–13.
- 72 W. Zhang, D. Zheng, Y. Xu, H. Bian, Y. Guo and H. Wang, *J. Chem. Phys.*, 2005, **123**, 224713.
- 73 G. J. Simpson, *Nonlinear Optical Polarization Analysis in Chemistry and Biology*, Cambridge University Press, Cambridge, 2017.
- 74 P. G. Breen, C. N. Foley, T. Boekholt and S. P. Zwart, *Mon. Not. R. Astron. Soc.*, 2020, **494**, 2465–2470.
- 75 L. Lu, P. Jin, G. Pang, Z. Zhang and G. E. Karniadakis, *Nat. Mach. Intell.*, 2021, **3**, 218–229.
- 76 F. Chollet, Keras Neural Network Library, <https://github.com/fchollet/keras>.
- 77 TheElectricFlock, Neural Network Trigonometric Function, <https://github.com/TheElectricFlock/Neural-Network-Trigonometric-Function>.
- 78 E. Barrena, C. Ocal and M. Salmeron, *J. Chem. Phys.*, 2001, **114**, 4210–4214.
- 79 H.-T. Rong, S. Frey, Y.-J. Yang, M. Zharnikov, M. Buck, M. Wühn, C. Wöll and G. Helmchen, *Langmuir*, 2001, **17**, 1582–1593.
- 80 G. A. Buxton and A. C. Balazs, *Mol. Simul.*, 2004, **30**, 249–257.
- 81 Z. Liu, B. X. Dong, M. Misra, Y. Sun, J. Strzalka, S. N. Patel, F. A. Escobedo, P. F. Nealey and C. K. Ober, *Adv. Funct. Mater.*, 2019, **29**, 1805220.
- 82 Y. Niu, Y. Pei, Y. Ma, M. Li and X. Zhang, *Mater. Lett.*, 2020, **275**, 128158.



- 83 P. Chinwangso, H. J. Lee, A. C. Jamison, M. D. Marquez, C. S. Park and T. R. Lee, *Langmuir*, 2017, **33**, 1751–1762.
- 84 Y. Niu, J. Zhao, X. Zhang, X. Wang, J. Wu, Y. Li and Y. Li, *Appl. Phys. Lett.*, 2012, **101**, 181903.
- 85 M. Aizawa, H. Nakamura, K. Matsumoto, T. Oguma and A. Shishido, *Mater. Adv.*, 2021, **2**, 6984–6987.
- 86 Y. Fu and W. J. Kao, *Expert Opin. Drug Deliv.*, 2010, **7**, 429–444.