ChemComm

Self-Assembly Controlled at the Level of Individual Functional Groups

Journal:	*ChemComm*
Manuscript ID	CC-FEA-08-2022-004537.R1
Article Type:	Feature Article

PCCP

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Self-Assembly Controlled at the Level of Individual Functional Groups

Benjamin R. Heiner, Alexander M. Pittsford, and S. Alex Kandel[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Molecular self-assembly is driven by intermolecular interactions between the functional groups on the component molecules. Small changes in molecular structure can make large differences in extended structure, and understanding this connection will lead to predictive power and control of the self-assembly process. Scanning tunneling microscopy is used to study self-assembly in two-dimensional clusters and monolayers, and the experimental approach is to study "families" of molecules where one or more functional groups is varied in a methodical way. Studied families include indole carboxylic acids, isatin derivatives (which have the indole backbone), quinaldic acid, thioethers, and fluorenone derivatives. In these systems, a variety of intermolecular interactions drive the assembly of the molecular monolayer, including hydrogen bonds, van der Waals forces, zwitterionic interactions, surface interactions, and halogen interactions.

1 Introduction

The central topic in the study of self-assembly is the correspondence between the chemical structure of a molecule and the two- or three-dimensional structure into which it assembles. An improved understanding of this correspondence will impact fields including crystal structure prediction,[1,2,3,4,5,6] supramolecular chemistry,[7,8,9] and protein folding.[10,11,12,13] Because assembly is driven by intermolecular interactions (or additionally, in the case of two-dimensional assembly on surfaces, molecule-surface interactions), each functional group in a molecule influences the final structure produced, potentially in synergistic ways and with unexpected results.[14]

Our approach is to use low-temperature scanning tunneling microscopy (STM) to study self-assembly in two-dimensional clusters and monolayers. STM allows sample surfaces to be imaged at sub-angstrom resolution, making it ideal for the study of surface-assembled molecules.[15] In order to explore the connection between molecular structure and assembly, we choose 'families' of molecules to study. By making iterative changes to the molecule while keeping the basic structure, we better understand how different functional groups affect two-dimensional crystal structure formation.

Polymorphism is the propensity of molecules to assemble into multiple different crystal structures with similar free energies.[5,6] It is a key consideration in crystal structure prediction, and a potential confounding factor in self-assembly. We study polymorphism through the intentional creation of metastable structures during the molecular adsorption process. This is achieved through pulse deposition, in which a pulsed-solenoid valve is used to spray microliter droplets directly into vacuum and onto the surface.[16] The solvent rapidly evaporates and the molecules are kinetically trapped in energy minima. Even for small and simple molecules, pulse deposition typically results in several co-existing structures. Controlled low-temperature annealing can then be used to guide the system towards configurations with lower free energy, potentially forming new structures and eventually achieving the most stable molecular monolayer.

2 Tiling

Some molecules, when deposited on a surface, exhibit extreme structural flexibility in self-assembled superstructure. We have investigated this comparing the assembly of ferrocenes according to a Penrose tiling,[17] while Blunt *et al.* uses a similar technique with tetracarboxylic acids, resulting in a rhombus-based random tiling.[18] In the case of ferrocenes, the images in fig. 1a show that the addition of a methylene group completely changes the surface-adsorbed configurations. Note in fig. 1b, the ferroceneacetic acid ($FcCH_2COOH$) forms regular, periodic rows of dimers. Ferrocenecarbocylic acid, fig. 1c,d, predominately forms aperiodic pentamers on the surface.

The addition of a CH_2 group to the ferrocene dramatically changes the outcome of molecular self assembly. By calculating the energy of FcCOOH dimers as a function of the COH bond angle, we can gain insight into the forces driving the formation of pentamers. As seen in fig. 2a, the lowest energy configuration of the dimer increases as the angle increases, until a well is formed between around 200° and 230°. Within this well is the necessary angle for pentamer formation, 224.8°. While the dimer is formed

[a] *Department of Chemistry, University of Notre Dame, 2002 Cavanaugh Dr, Notre Dame, IN 46556, USA. E-mail: skandel@nd.edu*

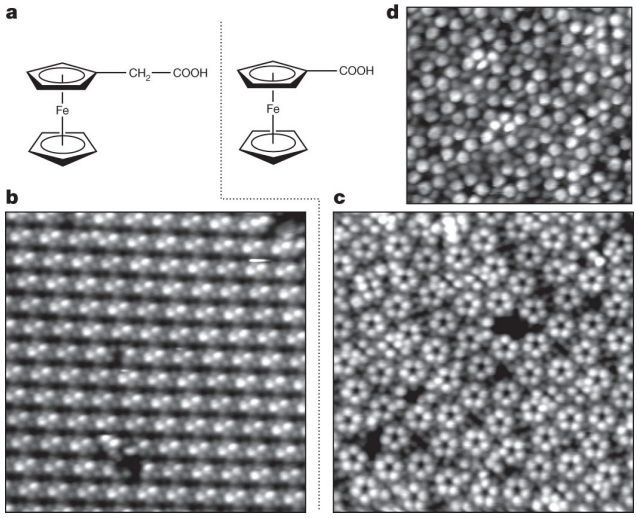

Fig. 1 (a) Molecular structures of $FcCH_2COOH$ and FcCOOH. (b) STM image (205 × 205 Å) of $FcCH_2COOH$. The surface is characterized by periodic dimer rows. (c) STM image (205 × 205 Å) of FcCOOH. This surface is predominately pentamers. (d) STM image (110 × 110 Å) of FcCOOH. Figure 1 from reference 17.

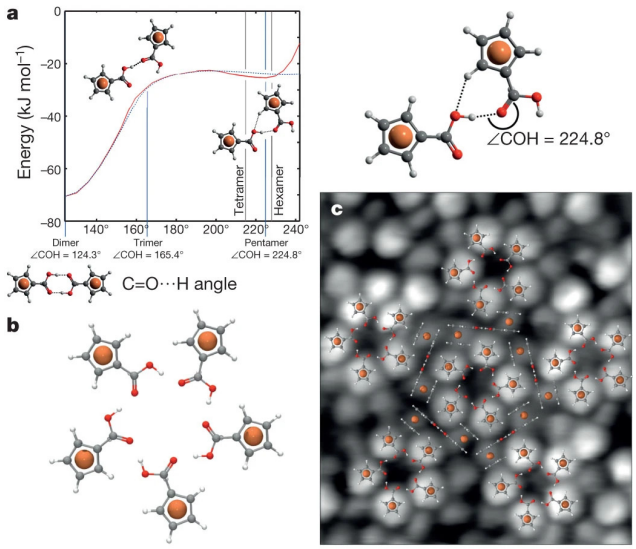

Fig. 2 (a) Calculated energies of bond angles of FcCOOH dimers. Angles and energies for trimers, tetramers, pentamers, and hexamers are noted, with trimers being near a peak (unstable), and tetramers, pentamers, and hexamers being near a well (stable). To the right, a dimer at the pentamer angle of 224.8 degrees is shown. The angle is formed by the primary O–H···O hydrogen bond being stabalized by a secondary C–H···O hydrogen bond. (b) A DFT minimized structure for a pentamer. (c) Calculated pentamers superimposed onto a representative STM image of ferrocene carboxylic acid. Dim features are understood to be Cp rings lying perpendicular to the surface. Figure 2 from reference 17.

because of reciprocal OH···O hydrogen bonding between the carboxylic acid groups, the potential well promoting pentamer formation comes from a secondary CH···O hydrogen bond between a Cp hydrogen and the carboxylic acid oxygen on the neighboring molecule.

$FcCH_2COOH$ molecules are therefore unable to form pentamers because the additional length added to the side chain prevents the requisite interaction between the carboxyl oxygen and the Cp hydrogens. If other factors were responsible for driving the pentamer formation, such as steric packing, we would observe pentamers forming for both molecules.

Another example of the intricacies that result when multiple intermolecular interactions drive the formation of a two-dimensional structure is from the work of Blunt *et al.*, who observed a rhombus-based random tiling for the molecule p-terphenyl–3,5,3',5'–tetracarboxylic acid (TPTC), assembled and imaged at the graphite/nonanoic acid interface.[18] The resulting structure (fig. 3A) is complex and non-periodic.ccEach of the four COOH groups on a TPTC hydrogen-bonds reciprocally with the COOH on another molecule; however different cycles involving from 3 to 6 molecules (fig. 3B–G are possible given the number and relative orientation of the COOH groups. Because molecules can come together in multiple different ways with equal or very similar interaction energies, the resulting two-dimensional tiling is random, and lacks any long-range periodicity.

3 Indole Carboxylic Acid

The structures of the indole isomers used are shown in Fig. 4. Fig. 5 shows indole-2-carboxylic acid after pulse deposition onto Au(111).[19] In the figure, three different clusters are highlighted: pentamers, hexamers, and catemer chains. The 3-D crystal structure is catemeric,[20] and only the catemer is observed after vapor deposition of this molecule.[21] This is a strong indication that pentamers and hexamers are metastable structures produced by pulsed deposition.

Pentamers for indole-2-carboxylic acid likely form because of the same combination of hydrogen-bonding interactions that result in pentamers of FcCOOH. Like FcCOOH, indole-2-carboxylic acid has a five-membered ring, an attached carboxylic acid group, and a hydrogen-bond donor immediately adjacent to the carboxylic acid. Pentamer formation was attributed to two separate hydrogen bonds forming a cyclic catemer, with the N-H acting as a donor for a secondary hydrogen bond. DFT calculations of cyclic catemers found that the N–H···O bond was maximally stable for the pentamer, being too long for smaller cyclic catemers and too short for larger.

Indole-3-carboxylic acid was investigated in order to study the importance of the nitrogen as a secondary hydrogen bond donor. With the carboxylic acid in the 3-position, the N–H group is isolated from the primary hydrogen bond, forcing a C–H group to donate in order to achieve a pentamer. In a crystal, indole-3-carboxylic acid forms dimer sheets.[22] Indeed, dimers were found when indole-3-carboxylic acid was imaged. As seen in fig. 6, however, the surface of the sample was mainly disordered, though with several pentamers visible. Thus, the atoms adjacent to the carboxylic acid group act as secondary hydrogen-bond donors, stabilizing the formation of cyclic clusters; the stronger N–H donor does this more effectively than C–H.

4 Isatin

1H-indole-2,3-dione (isatin, fig. 7) presents an opportunity to extend the results on indole-type molecules.[23] While there is no

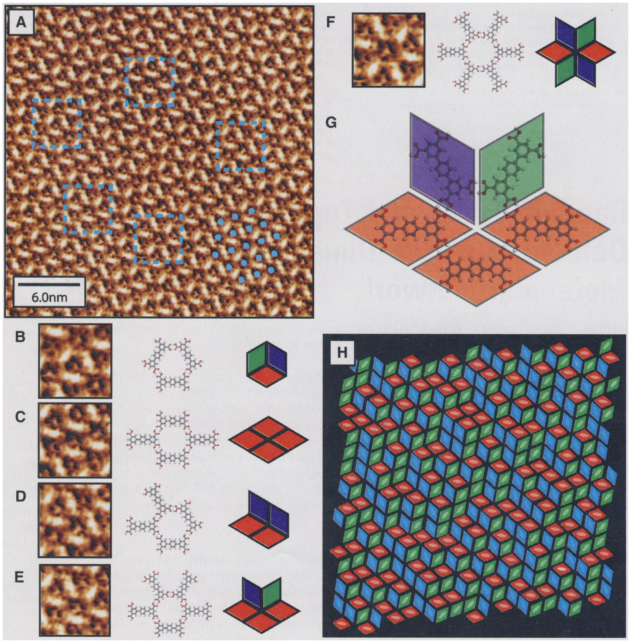

Fig. 3 (A) STM image of the molecule TPTC on graphite. (B–F) small excerpted images, shown alongside the arrangement of hydrogen bonds necessary to produce the feature, and a colored rhombus diagram characteristic of the assembly; a close-up of such a diagram is shown in (G). The full tessellation of rhombuses is shown in (H). Figure 2 from reference 18.

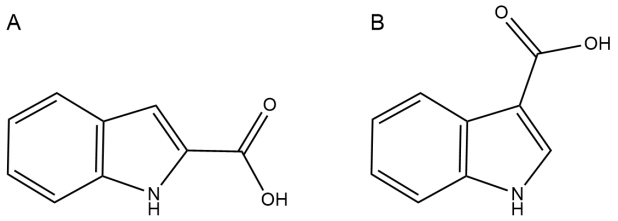

Fig. 4 Molecular structures of (a) indole-2-carboxylic acid and (b) indole-3-carboxylic acid.

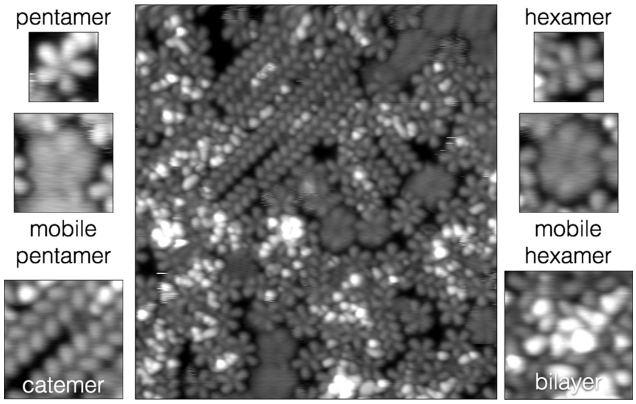

Fig. 5 Pulse deposited indole-2-carboxylic acid, 250 × 275 Å. Pulse deposition creates an environment where multiple crystal structures are kinetically trapped. In this image, we can see catemer chains, pentamers, and hexamers. Secondary hydrogen bonds drive the formation of the pentamers and hexamers. Figure 2 from reference 19.

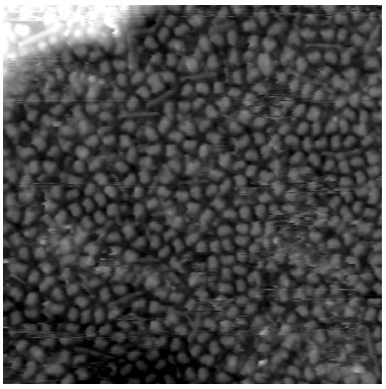

Fig. 6 Indole-3-carboxylic acid, 200 × 200 Å. While a pentamer can be seen, there are very few on the surface, most of which is disordered, though dimers can be found. Modified from figure 5 from reference 19.

carboxylic acid present, it has the indole backbone, two carbonyls to act as hydrogen-bond acceptors, and both N–H and C–H hydrogens that can act as potential donors. Figure 7 shows three closely related molecules: Isatin, 3-methyl 2-oxindole (3M2O), and 7-fluoroisatin. Each of the molecules have been pulse deposited onto Au(111) and imaged. Isatin predominantly forms pentamers with a structure shown in figure 8. Once again this is a cyclic catemer formed using N–H···O hydrogen bonds to the 2-position carbonyl group. However, the structure also brings the 7-position C–H into position to form a secondary C–H···O hydrogen bond with the 3-position carbonyl.

It is this bond that we investigate through comparison to 3M2O, which lacks the 3-position carbonyl, and 7-fluoroisatin, which no longer has a 7-position hydrogen. DFT calculations (figure 8) show substantially decreased binding energies for these molecules; note that the binding energy of the dimer (which is the primary unit of the 3-D crystal structure[24,25]) remains relatively unaffected. Interestingly, the 3M2O pentamer binding energy is similar to that of the dimer, and indeed 3M2O pentamers are observed. Relative to isatin, however, they are qualitatively changed. Additional bright spots appear in a wider, more spread-out pentamer. This is likely due to the methyl group in the 3-position leaving the plane of the rest of the molecule, thus making a brighter spot on the image. Pentamer formation of 7-fluoroisatin, on the other hand, is completely disrupted when fluorine is added to the fused benzene ring, instead creating asymmetric hexamers.

The secondary C–H···O hydrogen is thus the primary driver of pentamer stabilization, but we investigated additional molecules to determine the specific role of the N–H···O bond. Phthalimide is a structural isomer of isatin, where the carbonyl and nitrogen have been swapped, separating the two carbonyl groups; the structure can be seen in fig. 9. .

The N–H···O and C–H···O interactions are similar in phthalimide to isatin, and so should form an analogous pentamer structure, and DFT calculations of per-molecule binding energies suggest stability in the pentamer structure, though not as energetically favorable as the isatin. Experimentally, no pentamers were found. STM images of phthalimide, fig. 10, instead show

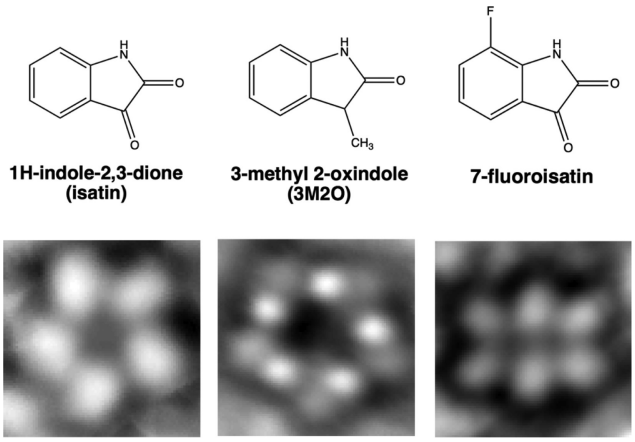

Fig. 7 Assemblies of isatin, 3M2O, and 7-fluoroisatin 20 × 20 Å. Isatin and 3M2O form pentamers while the 7-fluoroisatin forms hexamers. The pentamer formation is driven by the C–H···O hydrogen bonding in the 7 position. When the 7-position hydrogen is replaced by a fluorine, hexamers form instead of pentamers. From abstract figure from reference 23.

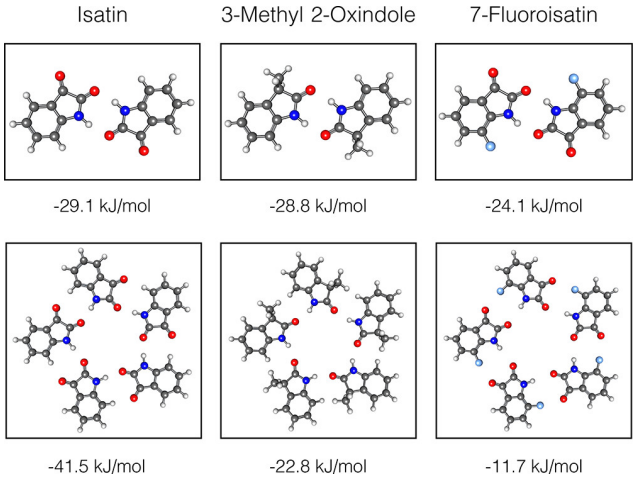

Fig. 8 Optimized geometries of dimers and C5 symmetric pentamers for isatin, 3M2O, and 7-fluoroisatin. The dimers and pentamers of isatin and 3M2O are found to be competitively energetic, with the pentamer in isatin to be more stable than the dimer. The dimer of 7-fluoroisatin is found to be significantly more stable than a theoretical pentamer. Figure 5 from reference 23

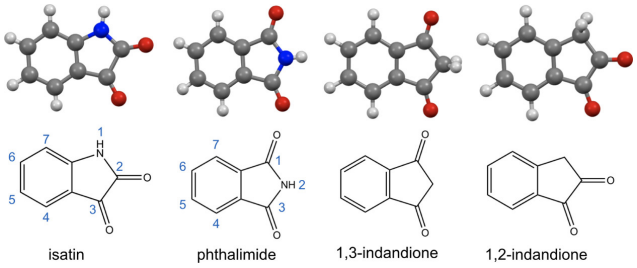

Fig. 9 Chemical structures and three dimensional rendering of isatin, phthalimide, 1,2-indandione, and 1,3-indandione. Figure 1 of reference 26.

a close-packed phase and a tetramer phase. The close-packed phase is unlikely to be the result of hydrogen bonding, as a phthalimide molecule simply does not have enough hydrogen-bonding groups to make equivalent contacts in a hexagonally packed region. Close-packing suggests an increased importance of molecule-surface interactions. The calculated tetramer structure, however, is substantially larger than the imaged feature.

Further information about the tetramer structure comes from the study of 1,3-indandione and 1,2-indandione, the analogues to isatin and phthalimide formed by replacing the NH groups with CH_2 groups as seen in fig. 9. Figure 11 shows images of 1,3-indandione, revealing close-packed networks of tetramers, seemingly identical to tetramers of phthalimide in fig. 10.

There are no strong hydrogen-bond donors in the structure of 1,3-indandione in fig. 9. However, keto-enol tautomerization to the enol form of the molecule moves one of the CH_2 hydrogens moves to the adjacent C=O, forming an OH functional group that can then act as a hydrogen-bond donor. It is then possible that the 1,3-indandione tetramers are cyclic structures driven by O–H···O hydrogen bonds. The similarities in the 1,3-indandione and phthalimide structures strongly suggests that this explanation could be extended to phthalimide. Evidence of the tautomerization of the phthalimide into the enol form is the clear mismatch between the DFT calculated structure (the keto tetramer) and the STM image it is superimposed upon in fig. 10d. Tautomerization of these molecules explains the observation that N–H···O hydrogen bonding appears less important for phthalimide and also explains the dramatic differences between phthalimide and isatin assemblies.

Images of 1,2-indandione, fig. 12, show domains of disordered clusters and areas of close-packed molecules. Because the model for isatin pentamer formation involves primary N–H···O and secondary C–H···O hydrogen bonds, it is unsurprising that the removal of the primary hydrogen-bond donor prevents the formation of pentamers. This confirms the importance of the N–H···O hydrogen bond in the isatin pentamer. Because removal of the N–H···O hydrogen bond results in no change in the structures of phthalimide and 1,3-indandione, we can conclude that the phthalimide tetramer structure is not dependent on the N–H···O hydrogen bond for formation.

The isatin pentamer structure shown in fig. 8 relies upon hydrogen bonds donated by the 1- and 7-position atoms to the 2- and 3-position carbonyls, respectively. The C–H hydrogens at the 4, 5, and 6 positions are remote from these intermolecular interactions, suggesting that functionalization of isatin at these positions should not impact pentamer formation. Investigation of 5- and 6-methylisatin (fig. 13) tests this hypothesis.

Images of 5- and 6-methylisatin after vapor deposition are shown in fig. 13; both molecules form pentamers. As with isatin, DFT calculations predict that pentamers are more favorable than 2-, 3-, or 4-mers.

Images acquired after pulse deposition are shown in fig. 14; pentamers are still seen, though domains of dimer chains and a close packed phase present themselves as well, consistent with the observation that pulse deposition usually produces a greater range of structures than vapor deposition. Pulse-deposited 5-methylisatin results in a monolayer structure differing from both isatin and 6-methylisatin, with figure 14 showing that monolay-

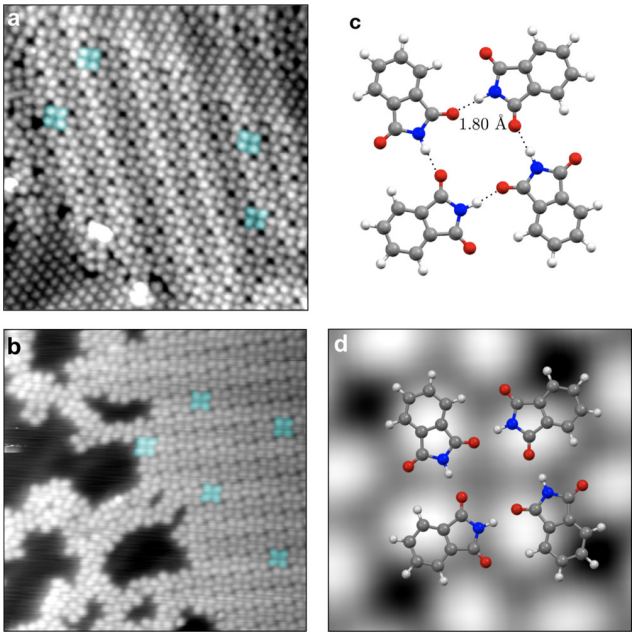

Fig. 10 (a) STM image (200 × 200 Å) of phthalimide on Au(111) showing a close-packed phase and a tetramer phase. (b) Lower coverage image of phthalimide on Au(111) (220 × 220 Å. (c) DFT-optimized geometry of a phthalimide tetramer, N–H···O distances shown. (d) Composite image of all tetramers in (b) (22 × 22 Å). Geometry in (c) is overlaid; poor agreement is evidence that the phthalimide self-assembles in the enol rather than the keto form. Figure 6 from reference 26.

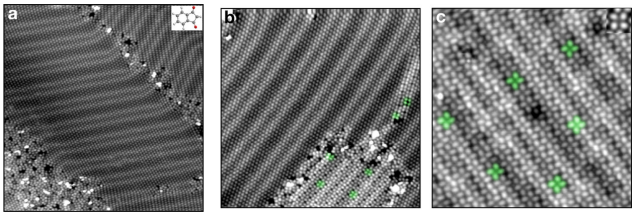

Fig. 11 STM images of 1,3-indandione on Au(111). (a) 500 × 500 Å image of close-packed molecules and disordered clustering. (b) 400 × 400 Å area of close-packed domain and tetramer networks. (c) An extended tetramer network (200 × 200 Å). Figure 7 from reference 26.

ers of 5-methylisatin consist of ordered, packed hexamers.

5-methylisatin hexamers reorganize into dimer rows when annealed, which indicates that the hexamer is a metastable intermediate structure. DFT calculations reveal similar trends in per-molecule cluster binding energies, so it is surprising that the 5-methylisatin and 6-methylisatin molecules have such different monolayer structures. The movement of alkyl groups have affecting 2D morphology is a known phenomenon, but this is ascribed to the changes in van der Waals forces from bulky alkyl chains. A methyl group would not be expected to change self-assembly results so drastically as the size of the group would not be expected to add significantly to van der Waals interactions.

5 Quinaldic Acid

Another family of molecules that can further our understanding of the forces driving self assembly is that of 2-naphthoic acid, quinaldic acid, and 3-quinoline carboxylic acid, structures in fig.

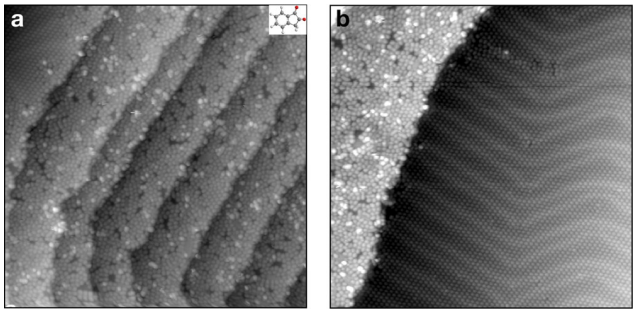

Fig. 12 STM images (500 × 500 Å) of 1,2-indandione. (a) disordered clustering on step edges. (b) close-packed molecules. Figure 8 from reference 26.

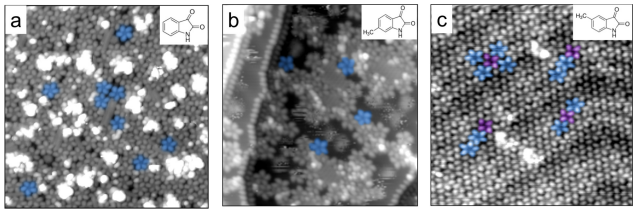

Fig. 13 Representative images of vapor deposited (a) Isatin (250 × 250 Å), (b) 6-methylisatin (250 × 250 Å), and (c) 5-methylisatin pentamers with intercalated tetramers colored in purple (200 × 200 Å). Chemical structures shown in insets. Figure 6 from reference 27.

15. These molecules differ from previous molecules in that they do not have an indole backbone and that quinaldic acid and 3-quinoline carboxylic acid have the ability to form a zwitterion. 2-naphthoic acid does not, so these molecules, which are very similar in structure, will be able to inform us on the role of zwitterions in the formation of two-dimensional crystals.

The scan of the pulse deposition of 2-naphthoic acid on Au(111) seen in fig. 16 revealed two structures: pentamers and dimer chains, with pentamers being more common. We have only observed C–H stabilization of a hydrogen-bonded ring structure from beta H-atom donors on five-membered rings with the indole backbone. Here, C–H stabilization is observed on a six-membered ring.

2-Naphthoic has two chemically equivalent C–H groups available for hydrogen bonding in the beta position relative to the carboxylic acid group. Each molecule, therefore, has two available conformations in the assembly, resulting in four being distinguishable pentamers. It is likely that the pentamers seen in fig. 16 are assembled into a mixture of the four conformations.

Quinaldic acid and 3-quinoline carboxylic acid are similar to 2-naphthoic acid; each of these molecules have fused six-membered rings, with the same position of the carboxylic acid group. For quinaldic acid, the carbon adjacent to the carboxylic acid is replaced with a nitrogen. 3-quinoline does the same, though the nitrogen and the carboxylic acid group are separated by another carbon. Our model thus far predicts that both molecules could form pentamer strucutures when pulse depostited as they have strong primary and secondary hydrogen bonds in the vacinity of each other.

The scans of the pulse depositions are seen in fig. 16.

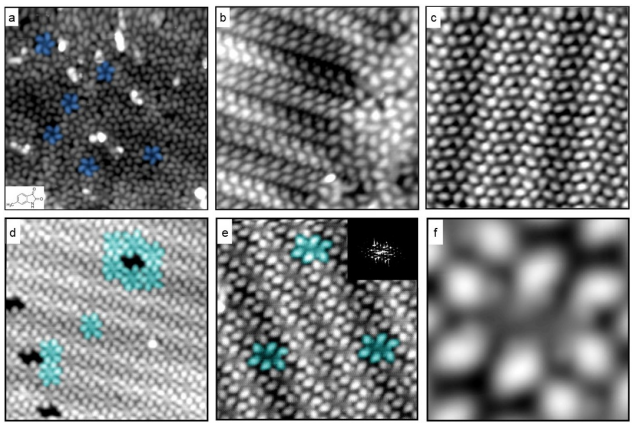

Fig. 14 Pulse deposition of 6 (a-c) and 5 methylisatin (d-e). (a) shows pentamers, (b) shows dimer chains, and (c) shows a close packed phase. The addition of a methyl group in the six position is distant from the hydrogen bonding and therefore does not affect pentamer formation. (c) A close-packed hexamer field (180 × 180 Å) arranging around vacancies (d) Close-packed hexamers (120 × 120 Å) are periodic. Inset is a fast Fourier transform of the area. (e) Composite image of 67 hexamers (23 × 23 Å). Modified from figures 5 and 6 from reference 27.

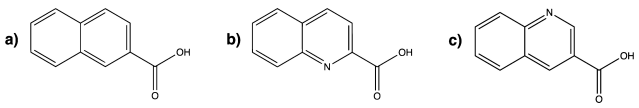

Fig. 15 Molecular structures of 2-napthoic acid, quinaldic acid, and 3-quinoline carboxylic acid. Figure 1 from reference 28.

Clearly, no pentamers are visable. Instead, quinaldic acid formed tetramers and dimers, while 3-quinoline formed a few hexamers and dimers on a largely disordered serface. This disparity between our prediction and the results are attributed to the molecules being able to form zwitterions and a hydrogen being transfered from the carboxylic acid groups to the nitrogen. In quinaldic acid, at least, the solid state consists of an even mixture of the carboxylic acid and zwitterionic species.[29]. The zwitterionic species may be stablized by the metal surface as well, producing the structures we see.

DFT calculations predict that, in the presence of hydrogen bonds, pentamer clusters would be seen instead of the assemblies that are observed. Also, the distances between the molecules in the STM scans are much closer together than would be expected for a hydrogen bond. This is evidence that the clusters formed by quinaldic acid and 3-quinoline are not stabalized by hydrogen bonds but are rather driven by the formation of the zwitterionic species.

6 Thioether

Murphy *et al.* imaged a family of thioethers, attempting to understand the effect the alkyl portions of the molecules have on molecular self assembly.[30] The structures of this family can be seen in fig. 17, and include butyl methyl sulfide (BMS), tert-butyl methyl sulfide (TBMS), dibutyl sulfide (DBS), and butyl sec-butyl sulfide (BSBS). These samples were imaged on Au(111) after an annealing step in an effort to reach the most stable assemblies for these molecules.

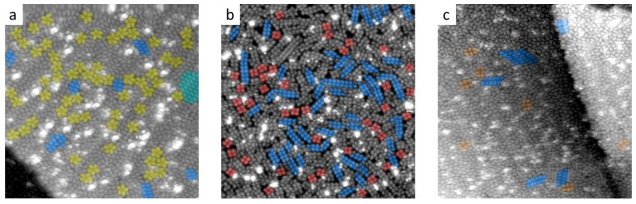

Fig. 16 (a) Pulse deposition of 2-napthoic acid, which forms pentamers. Pulse deposition of (b) quinaldic acid and (c) 3-quinoline. Neither molecule forms pentamers after pulse deposition, which differs from our prediction. This is attributed to the formation of zwitterions in the molecules. Modified from figures 2 and 3 from reference 28.

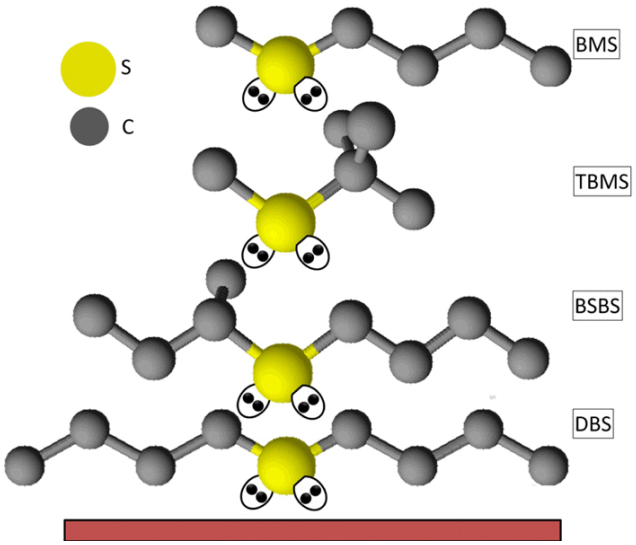

Fig. 17 Structures of the thioethers studied in reference 30. From top to bottom, BMS, TBMS, BSBS, and DBS. Figure 1 from reference 30.

BMS images, shown in fig. 18, are described as asymmetric "V-shape" protrusions in STM images. The methyl group can be seen in the image as the shorter side of the V. Their explanation for this behavior is that in order to maximize van der Waals interactions, BMS molecules assemble into long homochiral rows with adjacent alkyl tails of the same size. The behavior of BMS in a monolayer suggests that the butyl-butyl/methyl-methyl interactions lower the energy of the system relative to butyl-methyl/methyl butyl interactions.

The TBMS images, fig. 18, appear to be close-packed structures. The "V-shape" that the BMS showed during imaging completely disappears for TBMS. The change in the alkyl branching had a significant effect on the molecular reconstruction between BMS and TBMS.

Similar to BMS, DBS, fig 19, assembles into rows with their alkyl tails adjacent to one another in order to maximize attractive van der Waals interactions. Interestingly, at low surface coverage, DBS domains are restricted to the fcc regions of the gold surface and align with the underlying herringbone. As coverage increases, the molecules assemble into the gaps, creating a large, highly ordered two dimensional crystal. This behavior suggests that at low coverage, molecule-surface interactions are the driving force in the assembly, while at higher coverage, the molecule-

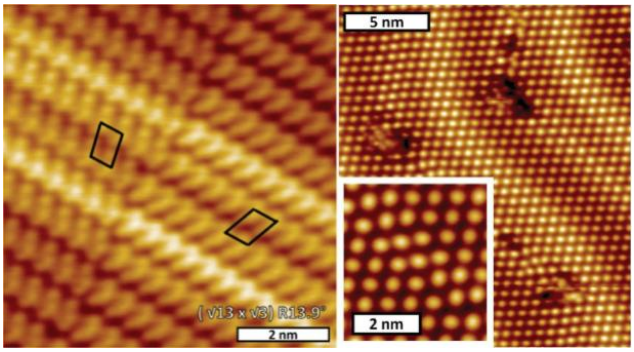

Fig. 18 BMS assembles into a "V-shape" on the surface of Au(111). TBMS images show close-packed structures with long-range order, typical of long-chain thioethers.Modified from figures 6 and 7 from reference 30.

molecule interactions overwhelm all others.

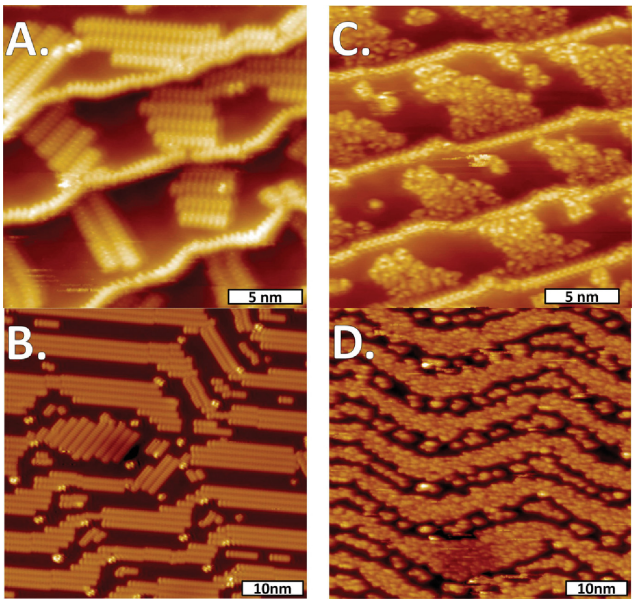

Fig. 19 DBS ((a) and (b)) and BSBS ((c) and (d)) on Au(111). The DBS assembles into highly ordered domains while the BSBS is disordered. Figure 9 from reference 30.

Similar to DBS, BSBS is restrained to the fcc region of the Au surface, which can be seen in fig. 19. However, even after annealing steps, which drive molecular assemblies to the thermodynamic minimum for the system, BSBS does not form an ordered two-dimensional crystal. Instead, disordered clusters appear in every image taken. This suggests that there is no energy preference to any assembly. These experiments were performed with a racemic mixture of BSBS, however, and they hypothesise that a enantiopure sample might assemble into an ordered structure, as the stereo center could be causing the disorder.

7 Fluorenone

Dong *et al.* recently investigated the effect halogen bonds have on molecular self assembly, comparing 2-bromo-7-((10-ethoxycarbonyl) decyloxy)-9-fluorenone (BEDF), 2-bromo-7-((10-carboxyl)decyloxy)-9-fluorenone (BCDF), and 2-bromo-7-((10-hydroxyl)decyloxy)-9-fluorenone (BHDF), the structures seen in fig. 20.[31] These molecules were deposited on HOPG and imaged in solution.

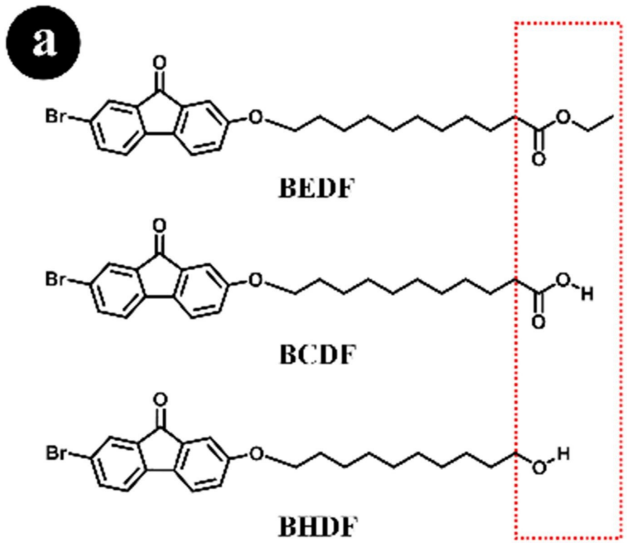

Fig. 20 Chemical structures of BEDF, BCDF, and BHDF. Modified from figure 1 from reference 31.

The assemblies of BEDF, BCDF, and BHDF are shown in fig. 21. Interestingly, none of the molecules formed a similar configuration on the surface of the graphite. the BEDF, in panel (a), forms a dimer. Deng and co-workers attributed this to the terminal $COOC_2H_5$ group of BEDF acting as a hydrogen bond acceptor and forming C-H···O=C hydrogen bonds. BCDF, panel (b), forms tetramers due to the COOH-terminated group of and the OH-terminated group serving as both electrophile and nucleophile to form strong hydrogen bonds. The secondary hydrogen bond formed from the carboxylic acid group further stabilizes the tetramer. BCDF forms into the tetramer pattern with Br···Br type-I bonds, while BHDF forms the lamellar pattern with Br···Br type-II halogen bonds taking the stronger secondary hydrogen bond, C-H···OH, and the fully interdigitation of the tetrameric aggregates into consideration. The octamer pattern of BHDF molecules forms at low concentrations with Br···O=C halogen bonds.

8 Conclusions

STM studies of "families"—molecules with similar structure but with methodical variations in one or more functional groups—helps build an understanding of how extended structure arises from the interplay of different intermolecular interactions. For indole carboxylic acids, it is the combination of strong O–H···O and supplementary N–H···O or C–H···O bonds that result in the assembly of cyclic pentamer structures. Similar structures are formed by isatin, based upon N–H···O and C–H···O interactions. Intramolecular interactions between functional groups, such as keto-enol tautomerization and zwitterion formation, must also be considered for their impact on extended structure. For example, isatin assembles in its keto form, with strong N–H···O bonds resulting in pentamers, while its isomer, phthalimide, forms O–H···O-bonded tetramers in its enol form. Further control over

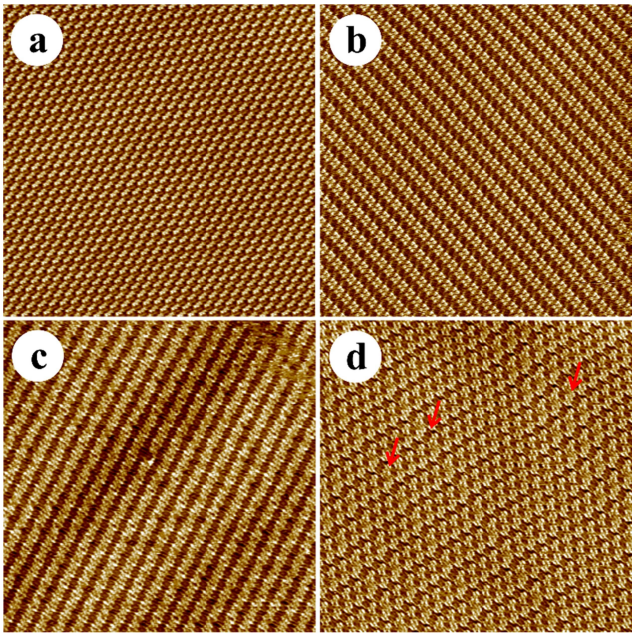

Fig. 21 STM images of the self-assembly networks at the 1-phenyloctane/HOPG interface. (a) The dimer pattern of BEDF. (b) The tetramer pattern of BCDF. (c) The lamellar pattern of BHDF. (d) The octamer pattern of BHDF. Figure 2 from reference 31.

the crystallization of a system could be achieved through alkyl placement. 5- and 6-methylisatin have different monolayer structures, suggesting that even movement of a methyl group on a non-interacting portion of a molecule can influence the superstructure. Longer chain alkyl chains, including branching alkyl chains, lead to further manipulation of the van der Waals forces, such as seen when comparing the different assemblies of BMS, TBMS, DBS, and BSBS.

The understanding of self-assembly at the level of individual functional groups has promising applications in related fields, including crystal structure prediction, crystal engineering, and supramolecular chemistry. For these applications to be realized, greater understanding of the driving forces needs to be achieved. The study of more families of molecules are necessary to further probe the impact of different atoms, functional groups, geometries, and chemical phenomena have on molecular self assembly. We suggest that the use of a variety of deposition techniques will allow for a more in depth probe into how the molecules are interacting with each other, both on a surface and in a solution.

We believe also that the comparison of experiment with theoretical calculations has significant far-reaching implications. Characterization of 2D clustering and crystallization for families of molecules, each member of which may have several polymorphic structures, provides a rich data set for comparison with theory. Through these comparisons, details of the theoretical approach can be refined to produce the best agreement with observed behavior: the data will be able to guide the choice of density functional, for example, or the choice of force-field. With this refinement of choices, the predictive power of theoretical methods will improve, both for 2-D systems and in applications in 3-D crystal structure prediction.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Support for this work has been provided by the US National Science Foundation (NSF Grants No. CHE-1807313 and CHE-2108186).

Notes and references

1 G. R. Desiraju, *Journal of the American Chemical Society*, 2013, **135**, 9952–9967.
2 S. L. Price, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2018, **474**, 20180351.
3 R. Gutzler, L. Cardenas and F. Rosei, *Chem. Sci.*, 2011, **2**, 2290–2300.
4 M. K. Corpinot, S. A. Stratford, M. Arhangelskis, J. Anka-Lufford, I. Halasz, N. Judaš, W. Jones and D.-K. Bučar, *CrystEngComm*, 2016, **18**, 5434–5439.
5 D. Gentili, M. Gazzano, M. Melucci, D. Jones and M. Cavallini, *Chem. Soc. Rev.*, 2019, **48**, 2502–2517.
6 H. Chung, S. Chen, N. Sengar, D. W. Davies, G. Garbay, Y. H. Geerts, P. Clancy and Y. Diao, *Chemistry of Materials*, 2019, **31**, 9115–9126.
7 J. MacLeod, *Journal of Physics D: Applied Physics*, 2019, **53**, 043002.
8 J. V. Barth, *Annual Review of Physical Chemistry*, 2007, **58**, 375–407.
9 J. K. Gimzewski and C. Joachim, *Science*, 1999, **283**, 1683–1688.
10 B. Rodríguez-Spong, C. P. Price, A. Jayasankar, A. J. Matzger and N. Rodríguez-Hornedo, *Advanced Drug Delivery Reviews*, 2004, **56**, 241–274.
11 C. J. Drummond and C. Fong, *Current Opinion in Colloid & Interface Science*, 1999, **4**, 449–456.
12 E. D. Głowacki, M. Irimia-Vladu, S. Bauer and N. S. Sariciftci, *J. Mater. Chem. B*, 2013, **1**, 3742–3753.
13 S. Zhang, *Nature Biotechnology*, 2003, **21**, 1171–1178.
14 R. D. Brown, R. C. Quardokus, N. A. Wasio, J. P. Petersen, A. M. Silski, S. A. Corcelli and S. A. Kandel, *Beilstein Journal of Nanotechnology*, 2017, **8**, 1801–1807.
15 T. Ohshiro and M. Taniguchi, *Analytical Biochemistry*, 2022, 114645.
16 M. Wolf, V. Hayes, C. R. Gerber, P. G. Quardokus, J. J. Ortiz-Garcia, C. Plummer and R. C. Quardokus, *Journal of Vacuum Science & Technology A*, 2020, **38**, 022413.
17 N. A. Wasio, R. C. Quardokus, R. P. Forrest, C. S. Lent, S. A. Corcelli, J. A. Christie, K. W. Henderson and S. A. Kandel, *Nature*, 2014, **507**, 86–89.
18 M. O. Blunt, J. C. Russell, M. d. C. Gimenez-Lopez, J. P. Garrahan, X. Lin, M. Schroder, N. R. Champness and P. H. Beton, *Science*, 2008, **322**, 1077–1081.
19 N. A. Wasio, R. C. Quardokus, R. D. Brown, R. P. Forrest, C. S. Lent, S. A. Corcelli, J. A. Christie, K. W. Henderson and

S. A. Kandel, *The Journal of Physical Chemistry C*, 2015, **119**, 21011–21017.

20 B. Morzyk-Ociepa, D. Michalska and A. Pietraszko, *Journal of Molecular Structure*, 2004, **688**, 79–86.

21 F. De Marchi, D. Cui, J. Lipton-Duffin, C. Santato, J. M. MacLeod and F. Rosei, *The Journal of Chemical Physics*, 2015, **142**, 101923.

22 G. Smith, U. D. Wermuth and P. C. Healy, *Acta Crystallographica Section E*, 2003, **59**, o1766–o1767.

23 A. M. Silski, R. D. Brown, J. P. Petersen, J. M. Coman, D. A. Turner, Z. M. Smith, S. A. Corcelli, J. C. Poutsma and S. A. Kandel, *The Journal of Physical Chemistry C*, 2017, **121**, 21520–21526.

24 B. Hachuła, H. T. Flakus and A. Polasz, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2014, **120**, 287–296.

25 J. Lipkowski, R. Luboradzki, L. Stefaniak and J. Wójcik, *Journal of Chemical Crystallography*, 1995, **25**, 299–308.

26 A. M. Silski, J. P. Petersen, R. D. Brown, S. A. Corcelli and S. A. Kandel, *The Journal of Physical Chemistry C*, 2018, **122**, 25467–25474.

27 A. M. Silski-Devlin, J. P. Petersen, J. Liu, S. A. Corcelli and S. A. Kandel, *The Journal of Physical Chemistry C*, 2020, **124**, 17717–17725.

28 J. P. Petersen, R. D. Brown, A. M. Silski, S. A. Corcelli and S. A. Kandel, *The Journal of Physical Chemistry C*, 2019, **123**, 13610–13614.

29 D. Dobrzyńska and L. B. Jerzykiewicz, *Journal of Chemical Crystallography*, 2004, **34**, 51–55.

30 C. J. Murphy, X. Shi, A. D. Jewell, A. F. McGuire, D. O. Bellisario, A. E. Baber, H. L. Tierney, E. A. Lewis, D. S. Sholl and E. C. H. Sykes, *The Journal of Chemical Physics*, 2015, **142**, 101915.

31 M. Dong, T. Hu, Y. Wang, P. Pang, Y. Wang, X. Miao, B. Li and W. Deng, *Applied Surface Science*, 2020, **515**, 145983.