Problems, successes, and challenges for the application of dispersion-corrected Density-Functional Theory combined with dispersion-based implicit solvent models to large scale hydrophobic self assembly and polymorphism

Jeffrey R. Reimers,[ab] Michael J. Ford,[b] and Lars Goerigk[c]

[a] *International Centre for Quantum and Molecular Structure, College of Sciences, Shanghai University, Shanghai, China*

[b] *Department of Physics and Advanced Materials, University of technology Sydney, Sydney, Australia*

[c] *School of Chemistry, The University of Melbourne, Melbourne, Australia*

Corresponding Author: Jeffrey R. Reimers email Jeffrey.reimers@uts.edu.au

Problems, successes, and challenges for the application of dispersion-corrected Density-Functional Theory combined with dispersion-based implicit solvent models to large scale hydrophobic self assembly and polymorphism

The recent advent of dispersion-corrected density-functional theory (DFT) methods allows for quantitative modelling of molecular self-assembly processes, and we consider what is required to develop applications to the formation of large self-assembled monolayers (SAMs) on hydrophobic surfaces from organic solution. Focus is on application of the D3 dispersion correction of Grimme combined with the solvent dispersion model of Floris, Tomasi, and Pascual-Ahuir to simulate observed scanning-tunnelling microscopy (STM) images of various polymorphs of tetraalkylporphyrin SAMs on highly oriented pyrolytic graphite (HOPG) surfaces. The most significant problem is identified as the need to treat SAM structures that are incommensurate with those of the substrate, providing a challenge to the use of traditional periodic-imaging boundary techniques. Using nearby commensurate lattices introduces non-systematic errors into calculated lattice constants and free-energies of SAM formation that are larger than experimental uncertainties and polymorph differences. Developing non-periodic methods for polymorph interface simulation also remains a challenge. Despite these problems, existing methods can be used to interpret STM images and SAM atomic structures, distinguishing between multiple feasible polymorph types. They also privide critical insight into the factors controlling polymorphism. All this stems from a delicate balance that the intermolecular D3 and solvent Floris, Tomasi, and Pascual-Ahuir corrections provides. Combined optimized treatments should yield fully quantitative approaches in the future.

Keywords: self-assembly, dispersion interactions, solvent dispersion, Gibbs free energy of formation, scanning-tunnelling microscopy

Introduction

Kohn-Sham density functional theory[1] (DFT) has become the by far most applied methodology in quantum chemistry, thanks to the ground-breaking development of general-gradient approximation[2,3,4,5,6] (GGA) and hybrid DFT[7,8,9] approaches in the 1980s and 1990s. However, initial hopes to have found a universal approach with a relatively low computational cost were somewhat dampened when first studies appeared in the mid-1990s showing that common DFT approximations, owing to their semi-local nature, incorrectly described the asymptotic decay of the electron density, and so did not properly describe the well-known R^{-6} distance behaviour of the London-dispersion[10] force.[11,12,13,14] This sparked vivid research aimed at solving this problem.

One of the first strategies used in this endeavour was to design empirical functionals whose parameters were optimised against non-covalent interaction energies.[15] Indeed, it at first appeared that some functionals such as PW91[16] seemed to be attractive near van der Waals equilibrium distances and they were subsequently used in various applications,[17] including the treatment of the adsorption of molecules onto surfaces. This method thus gave better results for non-covalent interactions than other methods and were initially proposed to do so by treating dispersion properly, but this was shown (by one of us) not to be the case.[18] Further, it has now been established that even highly-parameterised functionals[19,20] do not describe the asymptotes of the London-dispersion energy correctly and that they still suffer from the same underlying problem as all conventional DFT approximations.[21]

In the last decade, other promising approaches have been developed to tackle the dispersion problem, and they can be divided into three categories: van-der-Waals density functionals (vdW-DFs),[22,23,24,25] effective-core potentials (ECPs) that try to mimic dispersion effects,[26,27,28,29,30] and additive dispersion

corrections.[31,32,33,34,35,36,37,38,39,40,41,42] Herein, we briefly outline only the most important aspects of these corrections; for detailed reviews of the advantages and disadvantages of the different approaches we refer the reader to Refs.[43,44,45].

vdW-DFs are a combination of a conventional exchange-correlation functionals with a nonlocal kernel that describes the London-dispersion contribution to the total energy. The first functional of this type has been proposed by Langreth and co-workers.[46] The latest developments and applications in this field have been based on the VV10 kernel developed by Vydrov and van Voorhis.[47] For general thermochemistry, VV10-based methods were shown to overall have the same accuracy as some of the best additive dispersion corrections discussed further below.[45,48] Moreover, they showed smaller outliers and performed better in some systems that were more difficult to treat with additive corrections, e.g. tri- and tetrapeptide conformers.[45,48] However, the evaluation of the nonlocal kernel in vdw-DFs can be costly for larger systems, and we decided to not apply these approaches in the present work.

ECP methods are fitted to interaction energies of non-covalently bound dimers, and they have been developed for crystalline and molecular systems.[26,27,28,29,30] The technical advantage of these methods is that ECPs are implemented in all standard quantum-chemistry codes and that they can be readily used. A disadvantage is that the potentials have to be fitted for every single element, which limits the variety of systems that can be treated with such methods. Furthermore, while interaction energies in non-covalently bound complexes can be obtained with good accuracy,[29] questions have been raised about their performance for general thermochemistry.[45]

There are different variants of additive dispersion corrections. They all have in common that in principle they can be applied to any density functional. One subclass of

these approaches are density-dependent corrections, such as the exchange-hole dipole-moment (XDM) model by Becke and Johnson,[36,37,38,39,40,41] the DFT+vdW/vdW-TS method by Tkatchenko and Scheffler,[42] and the dDsC correction by Steinmann and Corminboeuf.[33] Applications of these methods are numerous and they all showed that these models can provide results with very good accuracy.[33,36,37,38,39,40,41,42,49,50,51,52,53] As our study is related to material science, we would like to particularly highlight recent reviews on the vdW-TS and dDsC methods that showcase their applicability to computational material science, including molecular crystals.[54,55]

Herein, we focus our study on the second subclass of additive corrections, the (usually) atomic pairwise approaches denoted as DFT-D type methods. These methods can be combined very easily with conventional DFT methods without any significant additional computational cost. They can simply be described as an additive energy correction E^D to the underlying DFT energy E^{DFT}, so that the total DFT-D energy is given simply by[31,32,33,34,35]

$$E^{DFT-D} = E^{DFT} + E^D. \quad (1)$$

Analytical gradients and second derivatives can also be derived from this scheme and therefore also dispersion-corrected DFT geometry optimisations and frequency calculations can be carried out.

In this *Ian Snook Memorial* issue of *Molecular Simulations*, we note the significant contributions that he made to this field. First, he focused on the most elementary systems: van der Waals dimers of noble gasses,[56] akin to what had been done previously in assessing density-functionals[11,12,13,14] including our assessment of the merits of the PW91 method for predicting weak non-bonded interactions.[18] In

this way he demonstrated the appropriateness of the DFT-D type approach. Then he performed a basic study, considering not just the requirements for a successful DFT-D approach to noble gas, π-stacking, and the other small-molecule phenomena that were being considered at the time, but also the requirements for a method that could treat large biological and nanosystems including systems close to his heart, metal clusters and graphene flakes.[57] However, DFT-D is only an approximate empirical scheme, and the leading developer of methods appropriate when DFT-D approaches intrinsically fail is Ian Snook's long-time colleague John Dobson. He has demonstrated that failure occurs fundamentally whenever the molecules or nanoparticles have degenerate ground states, leading to incorrect asymptotic behaviour of the potential. This situation does not apply in most chemical scenarios, leading to the widespread applicability of DFT-D.

Meanwhile, a tremendously successful additive dispersion correction for interacting non-degenerate ground-state systems that was developed is Grimme's DFT-D3 approach.[34,35] It has been presented in two variants, with the first often being called the "zero-damping" form[34] while the later improved method is called the "Becke-Johnson-damping"[36,38,39] (DFT-D3(BJ)) version.[35] Both versions have in common that they ensure the correct asymptotic R^{-6} long-range behaviour for non-degenerate ground-state interactions. Most of the parameters in these models are derived from first principles but either two parameters (zero-damping case) or three parameters (Becke-Johnson damping) need to be empirically adjusted for each density functional approximation. These parameters are independent of basis set, applied field, and all other parameters used in electronic-structure calculations. As the original implementation is no longer in common use, the later refinement is typically called just "DFT-D3" and we follow that approach herein. The number of successful DFT-D3 applications now goes into the hundreds. It suffices to say that this scheme has been

proven to be accurate and robust for non-covalently bound complexes,[21,34,35,53,58,59,60], for conformers of biologically relevant compounds,[17,35,61,62] in thermochemistry,[17,35,61,63,64,65,66] and also for the geometry optimisation of small,[64] medium.[67,68] and even very large systems such as proteins.[69]

It is interesting to note that the additive dispersion corrections were first introduce in the 1970's to correct Hartree-Fock theory.[70,71] Recent studies have gone back to this idea and it has turned out that HF-D3 can be of great value, too.[64,68,69,72,73,74] While this method is much more expensive than generalized-gradient-approximation (GGA) functionals when applied to periodic systems, when combined with D3, is also generally more accurate and robust.[69]

The D3 scheme has also been defined for calculations with periodic boundary conditions and so can be used in solid-state calculations[75] and, with some slight modifications, in surface calculations too. Although earlier forms of dispersion corrections had been made available in solid-state quantum-chemistry packages, the established standard procedure was to use dispersion-uncorrected GGA density functionals. This changed in 2010 when Moellmann and Grimme showed the importance of London-dispersion for stacking effects and structural properties in a crystal of a bis-thiophene-derivative.[75] Although this study was conducted with the original zero-damping version of DFT-D3, subsequent studies on organic crystals have shown that the Becke-Johnson damping scheme is indeed more reliable.

Later applications have shown that DFT-D3 leads to improved cohesive energies and lattice parameters for rare-gas crystals and 23 organic crystals,[76] allows an accurate description of crystals containing ethyl-acetate conformers,[77] and allows understanding crystal packing of organometallic-ruthenium(IV) [78] and

zirconocene[79] complexes. DFT-D3 has now also been successfully applied to inorganic systems, for instance, in the optimization of TiO_2 structures for which the inclusion of dispersion effects turned out to be important for the relaxation of the crystal surface.[80] A modified DFT-D3 version also allowed the description of adsorption processes to ionic surfaces.[81]

The D3 scheme has been successfully combined with other approaches and it allowed reliable prediction of organic molecular crystals at the HF wave-function and the density-functional-tight-binding (DFTB[82,83]) semi-empirical levels.[84,85] Thanks to these successes, the periodic DFT-D3 scheme is now available in the *Vienna Ab Initio Simulation Package* (VASP[86]) and it can therefore be used by a large community of researchers.

Our present study deals with adsorption of molecular assemblies to solid surfaces. It presents an analysis of calculations preformed previously to understand polymorphism in tetraalkylporphyrin (M-C_mP, see Fig. 1) self-assembled monolayer (SAM) polymorphs on a highly-ordered pyrolytic graphite (HOPG) surface from organic solutions like 1-phenyloctane, 1-octanoic acid, and *n*-tetradecane.[87,88] Two generically different types of polymorphs are considered, a low density form L as well as variants of a medium density form M. These studies:

- Introduced a variety of methods, some including DFT-D3, for calculating the Gibbs free energies for SAM formation from solution using implicit solvent models.[87]
- Showed that cancellation of very large, dramatically different contributions led to polymorphs of similar free energy.[87]
- Showed that D3 contributions can be very large, sometimes in excess of 300 kcal mol^{-1}.[87,88]

- Demonstrated the factors controlling polymorph properties and SAM formation as a function of the length of the alkyl chains.[88]
- Demonstrated that D3 was accurate to at least 15% based on analysis of observed scanning-tunnelling microscopy (STM) images alone.[88]

These outcomes are significant as experimental research into the thermodynamical and kinetic effects controlling the production of hydrophobic SAMs from solution is in its infancy. This field is also a precursor to other self-assembly processes of organic molecules including polymer structuring and protein folding. As recently reviewed,[89] only in the last few years have measurements of thermodynamic quantities for hydrophobic SAM production been made,[90,91,92,93] but what is now apparent is that once formed alkylporphyrin SAMs become kinetically trapped. A critical issue to be solved is the interplay between kinetic and thermodynamic effects in controlling structure and polymorphism.[90,91,92,94,95,96,97,98,99,100,101,102,103,104,105,106]
Calculations can contribute to solving these challenges in a number of ways:

- Through interpretation of observed STM images.
- Through the prediction of the structures of SAM polymorphs.
- Through the calculation of SAM enthalpy changes, identified as a major aid to experimental studies of SAM kinetics and thermodynamics.[89,106]
- Through the prediction of the free energies of SAM formation and polymorphism.
- Through the prediction of the rates of dynamical SAM processes.
- By providing insight into the chemical features that control thermodynamic and kinetic properties and hence polymorph structure and stability.

Our studies[87,88] have addressed all structural and thermodynamic aspects of these contributions but not those related to kinetics. Kinetics studies usually involve what in this case would be long and laborious searches for manifold transition states for polymorph interconversion processes, or else extremely computationally intensive molecular dynamics (MD) simulations. While MD simulations have been applied to SAM formation problems and provided insight,[92,103,104,107] including even DFT MD simulations,[108,109] conventional quantum-chemistry approaches[110,111,112] offer direct evaluation of the critical quantities for minimal human and computational effort, allowing for large numbers of calculations using the best available methods such as DFT-D3. This is the approach that we have taken.

Herein, following Snook,[56,57] by considering aspects of the previous studies,[87,88] we identify the level of achievement that a calculation must reach so that it can contribute to understanding the current challenges concerning hydrophobic SAM formation and other types of organic self assembly. These systems pose an interesting set of challenges and intricacies. Calculations must be able to:

- Automatically include all chemical features in a systematic a-priori fashion.
- Quickly and accurately model dispersion interactions both between different molecules and between molecules and nanoscopic objects such as graphene or HOPG substrates.
- Quickly and accurately treat the strong dispersive interactions that occur between these species and the organic solvent.
- Quickly and accurately determine zero-point energy, entropy, and other thermal corrections that define the Gibbs free energy at finite system temperature.
- Deal with substrate lattices that are incommensurate with the SAM lattice.

We consider how modern computational methods can address these issues to deal with the challenges facing today's self-assembly research. Four computational approaches are utilized to do this:[87,88,113]

(1) A mixed quantum-mechanical/molecular (QM/MM) mechanical method utilizing the classic B3LYP hybrid density functional in conjunction with force fields for treating the intermolecular dispersion interactions. An advantage of this approach is that each individual aspect of the method has been independently verified for accuracy.
(2) A quality GGA functional PBE[5] used with D3, PBE-D3.
(3) The traditional GGA density-functional preferred when intermolecular interactions are important, PW91,[16] used with D3, PW91-D3.
(4) The PBE functional used with the older DRSLL[22] DFT-D-type dispersion treatment, PBE-DRSLL.

Throughout, our implementation of Floris, Tomasi, and Pascual-Ahuir[114] is used to determine the solvent interactions between the substrate, SAM, and fully solvated adsorbate molecule. This treats the dispersion contribution to the solvation energy as being proportional to the available surface area. It is like the algorithm recently implemented[109] in VASP except that the solvent-accessible surface is used, a critical feature for the type of SAM investigated. Also, our computationally efficient QM/MM approach is used throughout to evaluate the zero-point energy, entropy and enthalpy thermodynamic corrections.

One other significant issue needs to be solved, the fact that organic SAMs on HOPG tend to be incommensurate with the substrate lattice.[87,88,113,115,116,117,118,119] This poses a serious technical challenge as

electronic-structure computational methods treat surfaces using periodic boundary conditions, therefore demanding that the two lattices (or multiples thereof) be commensurate. The QM/MM calculations overcome this problem as the MM force field only is used for the intermolecular interactions and these are so efficient that a large real-space sum can be performed, including interactions up to very large distances.[113] However, this approach cannot be used for the DFT calculations and so we introduced a technique whereby all structures are optimized at 10-15 nearby commensurate lattices and the result of lowest free-energy taken. The effects of this approximation, in particular, need to be well understood.

Methods

All methods are described in detail elsewhere.[87] Briefly, the QM/MM calculations are performed by GAUSSIAN[120] using the EXTERNAL command to supply a specialized molecule-HOPG interaction potential[113] fitted to experimental enthalpy data.[121] This is then combined with the standard AMBER force field[122] describing intermolecular interactions and the B3LYP/6-31G*[8,123] method describing the porphyrin.[8,123] The PBE-D3 and PW91-D3 calculations were performed using VASP[86,124] while the PBE-DRSLL calculations were performed using SIESTA.[125] The zero-point energy, enthalpy, and entropy corrections were evaluated in usual fashion[110] except that a real-space sum over a 5×5 grid of unit cells is used to include phonon modes. The solvation energies are expressed using an interpolating scheme fitted to solution-phase data evaluated by GAUSSIAN as $\Delta G = -\alpha A$ where A is the solvent-excluded surface area and we fit α = 0.0866 kcal mol^{-1} Å^{-2}. All STM images are processed using our internal calibration and statistical methods to provide accurate lattice dimensions and image shapes.[113]

Results

Incommensurate SAM and substrate lattices

The HOPG substrate lattice is triangular with a vector length of 2.46 Å (only every second surface carbon is visible owing to interactions between the top two graphene layers) and so half of this length is about the maximum amount by which a SAM lattice could be incommensurate with the surface. Compressing or expanding SAMs by this amount could induce significant energy penalties, however, especially for SAMs of smaller molecules. A way to reduce the effect is to find the best-matching substrate lattice parameters without regard to the alignment of the HOPG and SAM vectors. Alkane or tetraalkylporphyrin SAMs align with the alkyl chain directions parallel to a $\langle 11\bar{2}0 \rangle$ substrate vector.[87,88,113,115,116,117,118,119] Relaxing this constraint allows much better fits to be made to the vector lengths, resulting in an overall lower energy penalty for commensurate-lattice production.

Table 1 gives the QM/MM optimized lattice parameters defined in Fig. 2 for the L polymorphs of C_mP (see Fig. 1) SAMs with $6 \le m \le 24$. This computation method does not require commensurate lattices and predicts lattice vectors in excellent agreement with those observed.[87,88] Also listed in the table are the changes in these lattice parameters resulting from SAM optimization using PBE-D3, restricted to commensurate lattices only. While these differences could be systematic and result from simply the use of different computational methods, they appear much more random in nature, suggesting that the treatment of incommensurability provides a more significant contribution.

The total unit-cell surface area *A* provides one of the best indicators of this as it reflects the calculated SAM densities directly rather than just the way the molecules are fitted together. For C_6P, the incommensurate lattice has 1 $Å^2$ less surface area, that for $C_{12}P$ has 36 $Å^2$ more, while that for $C_{24}P$ has only 18 $Å^2$ more, so there is no correlation between cell size and chain length as there would be if the differences were owing to a systematic change in computation method. Also, none of the differences in the unit-cell lengths *a*, *b*, or *c* exceed the maximum expected for the incommensurability effect.

Variations in the orientation angle α (see Fig. 2) reflect the relative alignment of the SAM and substrate vectors. This angle undergoes a rapid change as the chain length becomes shorter than 8 as the chains no longer need to align tightly in order to form a SAM,[88] but change slowly and continuously as the chain length increases above 8. However, the PBE-D3 orientation angles vary widely as the found commensurate lattice of lowest free energy is driven more by the need to get the cell lengths and areas realistic than the need to properly orient the SAM and substrate lattices. This effect applies to alkane chains on HOPG but is not expected to be a general property of SAMS on surfaces. Note that the differences in area shown in Table 1 are almost always positive. These reflect the fact that SAMs are easier to stretch than compress owing to the hard-wall repulsions that occur when molecules are squashed together.

Table 2 shows specific results for the optimization of the $C_{13}P$ L lattice using PBE-D3. There, the calculated free energies of SAM formation ΔG are shown evaluated at 14 different commensurate lattices nearby to the QM/MM lattice. Note that these free energies are shown per mole of porphyrin as $\Delta G/n$, as relevant for chemical reactions in the gas phase and in solution, but is also shown as $\Delta G/A$, as is pertinent to SAMs on surfaces. The SAM that forms will be the one of lowest $\Delta G/A$,

but the relative ordering of the two quantities is largely preserved. Also included in the table are the components contributing to the free energy,

$$\Delta G = \Delta E + \Delta G_s + \Delta G_{corr} \qquad (2)$$

where ΔE is the electronic interaction energy evaluated using PBE-D3, ΔG_s is the differential change in solvation free energy on SAM formation, and ΔG_{corr} is the thermal correction, which is dominated by the entropy term $-T\Delta S$ at 298 K. The thermal correction is not evaluated using PBE-D3 owing to the expense, so instead the QM/MM value of 37 kcal mol^{-1} is used throughout.

We see that when the commensurate lattice is compressed to low surface area, the free energy can increase alarmingly, of order 20-30 kcal mol^{-1}. This again indicates the feature of the calculations that the cell area is almost always overestimated at the lowest-energy commensurate lattice.

In principle, it is possible to interpolate the free-energy data shown in Table 2 as a function of the 3 key lattice parameters *a*, *b*, and θ. However, this assumes that the relative orientations of the SAM and substrate lattice do not affect the free energy at all. Attempts to fit this data using simple means and thence estimate optimized incommensurate lattice parameters and energies failed to yield physically meaningful results, however. This is attributed to the hidden effects of the neglected orientational variable. It therefore appears that the best estimate that can be made of the unit cell structure and formation free energy is that for the lowest-energy commensurate lattice.

Interpretation of STM images

A critical role to be played by calculations for the understanding of SAM properties is that of interpreting observed STM images. These images are measured manipulating many control parameters such as bias voltage, set-point current, type of image to

display, feedback control settings, scan rate, scan direction, and image processing controls such as brightness and contrast. The bias voltage is most critical as it selects the range of surface energy levels that can contribute to the tunnelling. Setting the bias voltage low allows only states of the HOPG surface to contribute to the tunnelling. In this way, SAM layers on the surface can be seen through to reveal the underlying substrate lattice. This procedure is critical to the determination of the absolute SAM orientation angle α shown in Fig. 2. It is also essential for our procedure that allows for accurate experimental measurements of lattice properties through internal calibration.[113] This internal calibration procedure eliminates the effects of scan direction, providing another important contribution.

Most STM images are measured in "constant current" mode in which the tunnelling current is held fixed by a feedback control loop that changes the tip height dynamically as the scan proceeds. Normally presented in images (e.g., like Fig. 2) are such constant-current images drawn by displaying the tip height at each pixel, known as a topology image. Alternatively, the actual current may be displayed instead. While this is supposed to be held fixed by the feedback loop, the electronics settings used always result in an error current (the difference between the actual current and the set-point value), and this can be displayed, known as a current image. The current image corresponds to the derivative of the topology image when typical constant-current feedback control settings are used. This information can be extremely useful, especially if surface pitting is involved.[126] Changing the set-point current changes the average height that the tip scans above the surface, with lower currents moving the tip further away. Taking the tip away from the surface can reduce image quality as the signal to noise ratio decreases and as molecular features become less distinct. However, it can also improve image quality as it causes only the long-range asymptotic region of the

molecular electron density sampled, reducing contributions to the tunnelling current. A consequence of this effect is that only hydrogen atoms can be seen on alkyl chains as the hydrogen electrons sit in orbitals with much lower exponents than do say carbon orbitals, meaning that these penetrate much further from the surface.

Obtaining images in which clear features are resolved is usually a difficult challenge involving optimization of all of these effects. Interpreting them is also difficult as the observed images embody at least three main electronic/nuclear structural effects: the selectivity of the bias voltage for particular types of electronic structure, the relative height of different atoms above the surface, and the contributions different atoms make to the electron density at long range. This leads to a critical role for calculations to play in experimental investigations of SAM properties.

Simulating STM images requires significant approximations as the actual control functions used in the experiments are not easy to model. Instead, crude images are simulated usually using the Tersoff-Hamann approximation,[127] displaying surfaces of constant electron density arising from electrons within an appropriate energy window. Choice of the value of the electron density used parallels the choice of set-point current in the experiments, whilst the energy window mimics the experimental bias voltage. There is no commonly applied methodology for mapping the electron-density level onto the actual set-point current, and so this key parameter is chosen empirically so as to make the best-looking simulated STM images, just as the chosen set-point current and feedback parameters are determined during experiments. The utilised energy window in principle directly corresponds to the bias voltage, except that DFT calculations tend to underestimate band gaps and hence this parameter is usually allowed to vary over some range when interpreting experimental data.

How these factors work out in practice can be assessed by examining the comparisons between calculated and observed STM images for the intrinsically different tetraalkylporphyrin SAMs on HOPG shown in Fig. 3. There the observed images measured over a range of bias voltages are compared to calculated ones in each case constructed for bias voltages of ±700 mV. Both the observed and calculated images are overlayed with the atomic structures pertaining to the calculated images, showing how image features correlate with structural ones.

Both the observed and calculated images depict tip heights furthest from the surface (light shadings) when the tip is above the central porphyrin macrocycle ring or the kinks in the alkyl chains near these rings. Kinks in the alkyl chains are necessary as the chains initially head off into solution perpendicular to the macrocycle plane and need to turned around to lie flat on the surface using gauche conformations around the chain β and γ carbon atoms. These atoms therefore sit above the plane of the porphyrin and the remainder of the alkyl chains and so are the closest atoms physically to the STM tip. The simulated images show that tunnelling from both the kink and flat-lying regions of the alkyl chains is dominated by the hydrogen atoms, though the CH bonds can be made more apparent under certain conditions. Observed images often highlight only one of the two uppermost hydrogen atoms per chain carbon, however, owing to slight distortions of the chain and to substrate structure.[116] Exposed ring π orbitals also generate strong tunnelling, as also can say *d* orbitals in metallated porphyrins, e.g., from the low-lying half-filled d_{z^2} orbitals of Co(II) porphyrins.

Three of the SAMs displayed in Fig. 3 are for M polymorphs and have just two alkyl chains lying flat on the surface. The other two orient into the solution where they can interact in many different ways with similarly oriented chains or with the base-level SAM structure. Tunnelling through the molecular layer is allowed because the substrate

orbitals delocalize into the molecular orbitals, providing a connection pathway. This connection utilizes both direct through-space interactions and superexchange intramolecular processes.[128] For the alkyl chains dangling into solution to be apparent in the STM images, superexchange couplings must carry the current. Such couplings are very sensitive to the bias voltage and decay exponentially with distance.[129,130] As a result, the simulated STM images indicate that these chains are not visualized. This is a critical feature required for the qualitative interpretation of the STM images of the M polymorphs.

With the aid of the calculations, it becomes clear that the calculated structures do account for the observed internal structure of the STM images. Many alternate structures for these SAMs have also been generated but fail to account for observed features such as the details of the alkyl chain locations and orientation, the internal structure of the porphyrin rings (i.e., where about on the image the meso position and β-pyrrolic positions sit), and the locations of the chain kinks. Combined with sufficient experimental accuracy in the determination of SAM lattice parameters so as to be able to discriminate between various calculated structures, calculations are able to provide a critical role in the determination of atomic structures from STM images.

Rationalizing interface structures

Atomically smooth interface boundaries between regions containing different SAM polymorphs can often be observed in experiments[87,115] and provide extremely useful structural information. To date, methods such as DFT-D3 have not been applied to optimize such boundaries, owing perhaps to their considerable physical size and the unwanted demand for periodicity in the perpendicular direction. Development of such a technique would be a very powerful addition, however. Meanwhile, it is possible to

postulate realistic interface structures simply by merging and averaging calculated structures for the two polymorphs on each side of the interface. Results from this procedure are shown in Figs. 4 and 5.

Specifically, Fig. 4 shows an observed boundary interface structure involving the L and M polymorphs of Cu-C_{11}P[115] with overlaid atomic structure.[87] A single row of molecules assembled in the L polymorph (akin to that shown in Fig. 3 for C_{13}P except the chains are shorter) structure is shown in the middle of the figure, surrounded on each side by two interface layers. Between these two interface layers, the SAM adopts the structure of the pure M polymorph shown in Fig. 3. While the L polymorph has all 4 alkyl chains lying on the surface, the M polymorph has only 2 and the interface molecules (I) have 3. Highlighted at the bottom of Fig. 4 are 3D molecular structures for the L, M, and I molecules, as well as an overlay of the L and M structures. This overlay reveals that the atomic structures of the pure L and M polymorphs are very similar, except for the varying chain orientations. The shown approximate interface structures are obtained by selecting the appropriate differing chain configurations and by averaging the coordinates of the macrocycle rings and the two chains that lie flat on the surface. A convincing starting structure is thus produced. Figure 5 highlights a related situation for the junction between two M domains of C_{13}P smoothly adjoined by 5 intermediate molecular rows of which the middle one adopts the structure of the L polymorph.

The challenge to be able to perform full free-energy calculations and optimize the structure using 1D periodic boundary conditions is an important one as the molecular configurations can be flexible in certain specific ways, facilitating smooth interfaces forming between differently aligned SAM domains. These domain boundaries provide a lot of structural information and their analysis greatly aids the

determination of atomic structures representing STM images. Initially,[115] the central L phase of the Cu-C_{11}P polymorph shown in Fig. 4 was attributed to a different polymorph then label "B" as its lattice parameters differed appreciably from those of the large domains of L. However, PBE-D3 and QM/MM calculations revealed that the potential-energy surface is very flat in the direction that causes this shift. This occurs as the porphyrin macrocycle can rotate about its centre, with associated chain realignment, at minimal energy cost. The differences made manifest by the overlay of the coordinates of the L and M phase molecules shown in Fig. 4 highlights this motion. Hence to take this type of flexibility into account in a regular and systematic fashion, full optimization of the interface structures is required.

Understanding the forces driving SAM formation and polymorphism

The ultimate goal of calculations is to aid in the understanding of the basic chemical process, and the calculated free energies of formation of the SAMs provide a basis for the understanding of the thermodynamic factors controlling polymorphism. Figure 6 shows calculated free energies for the formation of the L and M polymorphs of C_{13}P, evaluated using all 4 computation methods. The total free energy change is decomposed into contributing binding, solvation, and thermal contributions in Fig. 6a, and then Figs. 6b analyse the origin of the solvation energy contribution whilst Fig. 6c analyses the contributions to the DFT-D binding. Detailed breakdowns of the energy components of the L polymorphs for chain lengths $6 \le m \le 28$ and for the M polymorphs with $11 \le m \le 13$ are given in Table 3, focusing on the PBE-D3 energies only.

As highlighted previously,[87] the cancellation of large binding, solvation, and entropy terms (these dominate the thermal corrections[87,88]) results in rather small

free energies of SAM formation, with the magnitudes of all the terms differing dramatically between the L and M SAMs. Here we focus on the results from the three DFT-D methods considered, PBE-D3, PW91-D3, and PBE-DRSLL. The PBE-D3 results for both the L and M polymorphs are very similar to the QM/MM ones, with the final free energy changes being comparable to those observed. PBE-DRSLL gives a similar value for the M polymorph but is overly attractive by 30 kcal mol^{-1} for L. These calculations were performed with the atomic-basis set code SIESTA using a double-zeta plus polarization basis and involve basis-set superposition error (BSSE) corrections. These corrections are worth hundreds of kcal mol^{-1}, however, owing to the very wide contact area between the HOPG and the closely spaced macrocycle plane and alkyl arms. It would appear that the erratic results produced by this method originate from the need to account for such large BSSE corrections. As the other PBE-D3 calculations were performed using the plane-wave basis set code VASP, no BSSE corrections are involved. While our primary aim here was to investigate the accuracy of D3 compared to DRSLL, it is not clear that the calculations manifest this difference properly.

However, the PW91-D3 results are systematically too attractive by 40 kcal mol^{-1}. We see that while PW91 empirically became the method of choice when GGA functionals were used in environments involving significant dispersion interactions, its qualities that gave it that feature result in it being very difficult to properly include the dispersion interaction. This effect was also found when the D3 parameters were fitted for this functional.[87] PW91-D3 is thus not a recommended functional, though it has been shown to deliver reliable results for the structures of sulfur SAMs on Au(111).[131,132]

That PBE-D3 gives results comparable to the QM/MM method is encouraging. The later approach is constructed including as one aspect an empirical force field fitted

to experimental thermodynamic data for the alkane-HOPG interaction and therefore is expected to yield realistic results for this critical interaction. To investigate whether or not the good agreement apparent in Fig. 6 is accidental or not, the QM/MM binding energies for the L polymorphs are compared to the PBE-D3 ones in Table 3. The PBE-D3 binding energies are less attractive by 9 kcal mol^{-1} for C_6P but more attractive by 1 kcal mol^{-1} for $C_{22}P$, implying a differential binding of 0.16 kcal mol^{-1} per CH_2 unit. This is significant as here D3 is being compared directly to the empirically fitted MM part of the QM/MM model. However, the difference is small and may be compensated for by other aspects of the calculations.

Solvation energies calculated for the PBE-D3 and QM/MM structure are very similar, with an average difference of 0.1±0.3 kcal mol^{-1}. The large net solvation energy changes reported in Table 3 of 60 – 200 kcal mol^{-1} arise because the SAM is solvated much more poorly than is the free molecule in solution and the bare HOPG surface, as shown in Fig. 6b. We use the same thermal corrections for both methods owing to the cost of evaluating them using DFT-D. Nevertheless, it is possible that the treatment of the solvent effects does compensate for the 0.16 kcal mol^{-1} interaction energy difference previously noted.

Table 3 and Fig. 6c highlight the importance of the dispersion correction to the DFT-D total binding energies, ranging from -107 to -324 kcal mol^{-1} for PBE-D3. Hence it is critical that the dispersion correction is represented as accurately as possible for the understanding of hydrophobic SAM formation and polymorphism.

Conclusions

Problems, successes, and challenges for dispersion-corrected DFT combined with implicit treatment of dispersion-dominated solvent interactions have been investigated through examination of the results from recent[87,88] free-energy

simulations of hydrophobic SAM polymorphism. The major technical challenge hampering this type of application is the propensity for hydrophobic SAMs to form in lattices that are oriented to say a HOPG substrate lattice but yet remain incommensurate with it. Because large DFT calculations required application of periodic imaging rather than say cluster models, SAM lattices must be distorted to commensurate ones prior to calculation. This is found to introduce non-systematic energy penalties of order a few kcal mol^{-1} into calculations. While this error is similar in magnitude to many other error terms contributing to the overall accuracy, its non-systematic nature makes trends difficult to identify and its magnitude considerably exceeds variations critical to understanding observed properties. Currently feasible approaches are thus realistic and informative but lack desired precision. This applies not just to calculated free energies but also to calculated lattice parameters as the uncertainties associated with the incommensurability exceed modern experimental error bars.

Independent of these problems, current DFT-D methods are shown to be very useful in predicting the internal structure of STM images, allowing key qualitative features of the atomic structure to be identified directly from the images. This is a significant result owing to the complexity of the experimental measurement and data presentation process, combined with the operation of different structural and electronic mechanisms contributing to the observed images. A significant way in which calculations can further contribute to interpreting STM data is by quantitative modelling of the interface boundaries between SAM domains and polymorphs, but this will require modified boundary conditions. Once again, the periodic imaging aspects of the calculations pose significant technical challenges.

At the most basic level, these calculations involve DFT to describe the often subtle chemical effects that control the conformational and ligation structures of the

adsorbate molecules, dispersion corrections like D3 and DRSLL to describe intermolecular interactions and molecule-substrate interactions, as well as dispersion-dominated interactions between the free molecule, the bare HOPG surface, and the SAM with the solvent. Two of these, the intermolecular dispersion interaction and the solvent interaction, involve empirical parameterization, whilst the other involves choice between a wide range of available density functionals embodying more and more chemical features. For hydrophobic self assembly from solution, we find that the balance between the two types of empirical dispersion corrections is critical to the calculations, especially for the calculated free energies. Minor differences between the representations of either of these terms could lead to large changes in calculated free energy changes for SAM polymorphism. Here we show that D3 combined with the Floris, Tomasi, and Pascual-Ahuir dispersion solvation relation yields excellent results. It is desirable that these two types of terms that historically have been considered separately become unified to give an optimal approach for considering the free energy of self assembly of systems from organic solutions or environments. Also, in future studies it is desirable to test the influence of many-body effects – for instance three-body terms in a DFT-D3 or related model – on similar systems, as they have been shown to become important in the thermochemical description of larger host-guest complexes.[53,65]

Acknowledgements

We thank the Australian Research Council for funding this research (grants LP0455238, DP12010259, and DE140100550) and National Computational Infrastructure (NCI) and INTERSECT for provision of computing resources; L.G. also thanks the Selby Scientific Foundation for receipt of a Research Award.

References:

1. Kohn W, Sham LJ. SELF-CONSISTENT EQUATIONS INCLUDING EXCHANGE AND CORRELATION EFFECTS. Phys Rev. 1965;140:A1133-A8.
2. Langreth DC, Mehl MJ. Beyond the local-density approximation in calculations of ground-state electronic properties. Phys Rev B. 1983;28:1809-34.
3. Perdew JP, Yue W. ACCURATE AND SIMPLE DENSITY FUNCTIONAL FOR THE ELECTRONIC EXCHANGE ENERGY - GENERALIZED GRADIENT APPROXIMATION. Phys Rev B. 1986;33:8800-2.
4. Perdew JP, Chevary JA, Vosko SH, Jackson KA, Pederson MR, Singh DJ, Fiolhais C. ATOMS, MOLECULES, SOLIDS, AND SURFACES - APPLICATIONS OF THE GENERALIZED GRADIENT APPROXIMATION FOR EXCHANGE AND CORRELATION. Phys Rev B. 1992;46:6671-87.
5. Perdew JP, Burke W, Ernzerhof M. Generalized gradient approximation made simple. Phys Rev Lett. 1996;77:3865-8.
6. Becke AD. Density-functional exchange-energy approximation with correct asymptotic behavior. Phys Rev A. 1988;38:3098-100.
7. Becke AD. A NEW MIXING OF HARTREE-FOCK AND LOCAL DENSITY-FUNCTIONAL THEORIES. J Chem Phys. 1993;98:1372-7.
8. Becke AD. Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys. 1993;98:5648-52.
9. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. J Phys Chem. 1994;98:11623.
10. London F. On the Theory and Systematic of Molecular Forces. Z Phys. 1930;63:245-79.
11. Kristyan S, Pulay P. CAN (SEMI)LOCAL DENSITY-FUNCTIONAL THEORY ACCOUNT FOR THE LONDON DISPERSION FORCES. Chem Phys Lett. 1994;229:175-80.
12. Perez-Jorda JM, Becke AD. A DENSITY-FUNCTIONAL STUDY OF VAN-DER-WAALS FORCES - RARE-GAS DIATOMICS. Chem Phys Lett. 1995;233:134-7.
13. Hobza P, Sponer J, Reschel T. DENSITY-FUNCTIONAL THEORY AND MOLECULAR CLUSTERS. J Comput Chem. 1995;16:1315-25.
14. Sponer J, Leszczynski J, Hobza P. Base stacking in cytosine dimer. A comparison of correlated ab initio calculations with three empirical potential models and density functional theory calculations. J Comput Chem. 1996;17:841-50.
15. Adamo C, Barone V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models. J Chem Phys. 1998;108:664-75.
16. Perdew JP, Wang Y. Accurate and simple analytic representation of the electron-gas correlation energy. Phys Rev B. 1992;45:13244-9.
17. Wesolowski TA, Parisel O, Ellinger Y, Weber J. Comparative Study of Benzene···X (X = O2, N2, CO) Complexes Using Density Functional Theory: The Importance of an Accurate Exchange−Correlation Energy Density at High Reduced Density Gradients. J Phys Chem A. 1997;101:7818-25.
18. Bilic A, Reimers JR, Hush NS, Hafner J. Adsorption of ammonia on the gold (111) surface. J Chem Phys. 2002;116:8981-7.

19. Zhao Y, Schultz NE, Truhlar DG. Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. J Chem Phys. 2005;123:161103.
20. Zhao Y, Truhlar DG. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor Chem Acc. 2008;120:215-41.
21. Goerigk L, Kruse H, Grimme S. Benchmarking Density Functional Methods against the S66 and S66x8 Datasets for Non-Covalent Interactions. ChemPhysChem. 2011;12:3421-33.
22. Dion X, Rydberg H, Schröder E, Langreth DC, B. I. Lundqvist. Phys Rev Lett. 2004;92:246401.
23. Lee K, Murray ED, Kong L, Lundqvist BI, Langreth DC. Higher-accuracy van der Waals density functional. Phys Rev B. 2010;82.
24. Vydrov OA, Van Voorhis T. Nonlocal van der Waals Density Functional Made Simple. Phys Rev Lett. 2009;103:063004.
25. Vydrov OA, Van Voorhis T. Nonlocal van der Waals density functional: The simpler the better. J Chem Phys. 2010;133:244103.
26. von Lilienfeld OA, Tavernelli I, Rothlisberger U, Sebastiani D. Optimization of effective atom centered potentials for London dispersion forces in density functional theory. Phys Rev Lett. 2004;93:153004.
27. von Lilienfeld OA, Tavernelli I, Rothlisberger U, Sebastiani D. Performance of optimized atom-centered potentials for weakly bonded systems using density functional theory. Physical Review B. 2005;71:195119.
28. DiLabio GA. Accurate treatment of van der Waals interactions using standard density functional theory methods with effective core-type potentials: Application to carbon-containing dimers. Chem Phys Lett. 2008;455:348-53.
29. Torres E, DiLabio GA. A (Nearly) Universally Applicable Method for Modeling Noncovalent Interactions Using B3LYP. J Phys Chem Lett. 2012;3:1738-44.
30. DiLabio GA, Koleini M, Torres E. Extension of the B3LYP-dispersion-correcting potential approach to the accurate treatment of both inter- and intra-molecular interactions. Theor Chem Acc. 2013;132:1389.
31. Grimme S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. J Comput Chem. 2004;25:1463-73.
32. Grimme S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. J Comp Chem. 2006:1787-99.
33. Steinmann SN, Corminboeuf C. Comprehensive Benchmarking of a Density-Dependent Dispersion Correction. J Chem Theory Comput. 2011;7:3567-77.
34. Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys. 2010;132:154104.
35. Grimme S, Ehrlich S, Goerigk L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. J Comput Chem. 2011;32:1456-65.
36. Becke AD, Johnson ER. A density-functional model of the dispersion interaction. J Chem Phys. 2005;123:154101.
37. Becke AD, Johnson ER. Exchange-hole dipole moment and the dispersion interaction. J Chem Phys. 2005;122:154104.
38. Johnson ER, Becke AD. A post-Hartree-Fock model of intermolecular interactions. J Chem Phys. 2005;123:024101.

39. Johnson ER, Becke AD. A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. J Chem Phys. 2006;124:174104.
40. Becke AD, Johnson ER. Exchange-hole dipole moment and the dispersion interaction revisited. J Chem Phys. 2007;127.
41. Johnson ER, Mori-Sanchez P, Cohen AJ, Yang W. Delocalization errors in density functionals and implications for main-group thermochemistry. J Chem Phys. 2008;129:204112.
42. Tkatchenko A, Scheffler M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. Phys Rev Lett. 2009;102.
43. Grimme S. Density functional theory with London dispersion corrections. Wiley Interdiscip Rev: Comput Mol Sci. 2011;1:211-28.
44. Klimeš J, Michaelides A. Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory. J Chem Phys. 2012;137:120901.
45. Goerigk L. How Do DFT-DCP, DFT-NL, and DFT-D3 Compare for the Description of London-Dispersion Effects in Conformers and General Thermochemistry? J Chem Theory Comput. 2014;10:968-80.
46. Dion M, Rydberg H, Schroder E, Langreth DC, Lundqvist BI. Van der Waals density functional for general geometries. Phys Rev Lett. 2004;92.
47. Vydrov OA, Van Voorhis T. Nonlocal van der Waals density functional: The simpler the better. J Chem Phys. 2010;133.
48. Hujo W, Grimme S. Performance of the van der Waals Density Functional VV10 and (hybrid)GGA Variants for Thermochemistry and Noncovalent Interactions. J Chem Theory Comput. 2011;7:3866-71.
49. Tkatchenko A, Rossi M, Blum V, Ireta J, Scheffler M. Unraveling the Stability of Polypeptide Helices: Critical Role of van derWaals Interactions. Phys Rev Lett. 2011;106.
50. Marom N, Tkatchenko A, Rossi M, Gobre VV, Hod O, Scheffler M, Kronik L. Dispersion Interactions with Density-Functional Theory: Benchmarking Semiempirical and Interatomic Pairwise Corrected Density Functionals. J Chem Theory Comput. 2011;7:3944-51.
51. Santra B, Klimes J, Alfe D, Tkatchenko A, Slater B, Michaelides A, Car R, Scheffler M. Hydrogen Bonds and van der Waals Forces in Ice at Ambient and High Pressures. Phys Rev Lett. 2011;107.
52. Zhang G-X, Tkatchenko A, Paier J, Appel H, Scheffler M. van der Waals Interactions in Ionic and Semiconductor Solids. Phys Rev Lett. 2011;107.
53. Risthaus T, Grimme S. Benchmarking of London Dispersion-Accounting Density Functional Theory Methods on Very Large Molecular Complexes. J Chem Theory Comput. 2013;9:1580-91.
54. Kronik L, Tkatchenko A. Understanding Molecular Crystals with Dispersion-Inclusive Density Functional Theory: Pairwise Corrections and Beyond. Acc Chem Res. 2014;47:3208-16.
55. Corminboeuf C. Minimizing Density Functional Failures for Non-Covalent Interactions Beyond van der Waals Complexes. Acc Chem Res. 2014;47:3217-24.
56. Snook I, Per MC, Russo SP. An analysis of the correlation energy contribution to the interaction energy of inert gas dimers. J Chem Phys. 2008;129:164109/1-/5.
57. Snook IK, Per MC, Seyed-Razavi A, Russo SP. Some comments on the DFT+D method. Chem Phys Lett. 2009;480:327-9.

58. Ehrlich S, Moellmann J, Grimme S. Dispersion-Corrected Density Functional Theory for Aromatic Interactions in Complex Systems. Acc Chem Res. 2013;46:916-26.
59. Sure R, Antony J, Grimme S. Blind Prediction of Binding Affinities for Charged Supramolecular Host-Guest Systems: Achievements and Shortcomings of DFT-D3. J Phys Chem B. 2014;118:3431-40.
60. Antony J, Sure R, Grimme S. Using dispersion-corrected density functional theory to understand supramolecular binding thermodynamics. Chem Commun. 2015;51:1764-74.
61. Goerigk L. How Do DFT-DCP, DFT-NL, and DFT-D3 Compare for the Description of London-Dispersion Effects in Conformers and General Thermochemistry? J Chem Theory Comput 2014;10:968-80.
62. Goerigk L, Karton A, Martin JML, Radom L. Accurate quantum chemical energies for tetrapeptide conformations: why MP2 data with an insufficient basis set should be handled with caution. Phys Chem Chem Phys. 2013;15:7028-31.
63. Goerigk L, Grimme S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. Phys Chem Chem Phys. 2011;13:6670-88.
64. Grimme S, Huenerbein R, Ehrlich S. On the Importance of the Dispersion Energy for the Thermodynamic Stability of Molecules. Chem Phys Chem. 2011;12:1258-61.
65. Steinmetz M, Hansen A, Ehrlich S, Risthaus T, Grimme S. Accurate Thermochemistry for Large Molecules with Modern Density Functionals. Topics in Current Chemistry: Springer Berlin Heidelberg; 2014. p. 1-23.
66. Karton A, Goerigk L. Accurate reaction barrier heights of pericyclic reactions: Surprisingly large deviations for the CBS-QB3 composite method and their consequences in DFT benchmark studies. J Comput Chem. 2015;36:622-32.
67. Grimme S, Steinmetz M. Effects of London dispersion correction in density functional theory on the structures of organic molecules in the gas phase. Phys Chem Chem Phys. 2013;15:16031-42.
68. Goerigk L, Reimers JR. Efficient Methods for the Quantum Chemical Treatment of Protein Structures: The Effects of London-Dispersion and Basis-Set Incompleteness on Peptide and Water-Cluster Geometries. J Chem Theory Comput. 2013;9:3240-51.
69. Goerigk L, Collyer CA, Reimers JR. Recommending Hartree–Fock Theory with London-Dispersion and Basis-Set-Superposition Corrections for the Optimization or Quantum Refinement of Protein Structures. J Phys Chem B. 2014;118:14612-26.
70. Hepburn J, Scoles G, Penco R. A simple but reliable method for the prediction of intermolecular potentials. Chem Phys Lett. 1975;36:451-6.
71. Ahlrichs R. Intermolecular forces in simple systems. Chem Phys. 1977;19:119-30.
72. Goerigk L, Falklöf O, Collyer C, Reimers J. First Steps Towards Quantum Refinement of Protein X-Ray Structures. In: Zeng J, Zhang R-Q, Treutlein HR, editors. Quantum Simulations of Materials and Biological Systems: Springer Netherlands; 2012. p. 87-120.
73. Sure R, Grimme S. Corrected small basis set Hartree-Fock method for large systems. J Comput Chem. 2013;34:1672-85.
74. Conrad JA, Gordon MS. Modeling Systems with π–π Interactions Using the Hartree–Fock Method with an Empirical Dispersion Correction. J Phys Chem A. 2014:DOI: 10.1021/jp510288k.

75. Moellmann J, Grimme S. Importance of London dispersion effects for the packing of molecular crystals: a case study for intramolecular stacking in a bis-thiophene derivative. Phys Chem Chem Phys. 2010;12:8500-4.
76. Moellmann J, Grimme S. DFT-D3 Study of Some Molecular Crystals. J Phys Chem C. 2014;118:7615-21.
77. Brandenburg JG, Grimme S. A dispersion-corrected density functional theory case study on ethyl acetate conformers, dimer, and molecular crystal. Theor Chem Acc. 2013;132:1399.
78. Moellmann J, Grimme S. Influence of Crystal Packing on an Organometallic Ruthenium(IV) Complex Structure: The Right Distance for the Right Reason. Organometallics. 2013;32:3784-7.
79. Brandenburg JG, Bender G, Ren J, Hansen A, Grimme S, Eckert H, Daniliuc CG, Kehr G, Erker G. Crystal Packing Induced-Carbon-Carbon Double-Triple Bond Isomerization in a Zirconocene Complex. Organometallics. 2014;33:5358-64.
80. Moellmann J, Ehrlich S, Tonner R, Grimme S. A DFT-D study of structural and energetic properties of TiO2 modifications. Journal of Physics-Condensed Matter. 2012;24:424206.
81. Ehrlich S, Moellmann J, Reckien W, Bredow T, Grimme S. System-Dependent Dispersion Coefficients for the DFT-D3 Treatment of Adsorption Processes on Ionic Surfaces. Chem Phys Chem. 2011;12:3414-20.
82. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. Phys Rev B. 1998;58:7260.
83. Aradi B, Hourahine B, Frauenheim T. DFTB+, a Sparse Matrix-Based Implementation of the DFTB Method†. J Phys Chem A. 2007;111:5678-84.
84. Brandenburg JG, Grimme S. Dispersion corrected hartree-fock and density functional theory for organic crystal structure prediction. Top Curr Chem. 2014;345:1-23.
85. Brandenburg JG, Grimme S. Accurate Modeling of Organic Molecular Crystals by Dispersion-Corrected Density Functional Tight Binding (DFTB). J Phys Chem Lett. 2014;5:1785-9.
86. Kresse G, Furthmüller J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. Comput Mat Sci. 1996;6:15-50.
87. Reimers JR, Panduwinata D, Visser J, Chin Y, Tang C, Goerigk L, Ford MJ, Sintic M, Sum TJ, Coenen MJJ, Hendriksen BAL, Elemans JAAW, Hush NS, Crossley MJ. The Free Energy of Formation for Hydrophoboc Self-Assembled Monolayers: Factors Controlling Polymorphism. Nature Comm submitted. 2015.
88. Reimers JR, Panduwinata D, Visser J, Chin Y, Tang C, Goerigk L, Ford MJ, Sintic M, Sum TJ, Coenen MJJ, Hendriksen BLM, Elemans JAAW, Hush NS, Crossley MJ. From chaos to Order: Chain-Length Dependence of the Free Energy of Formation of Tetraalkylporphyrin Self-Assembled Monolayer Polymorphs JACS submitted. 2015.
89. Mazur U, Hipps KW. Kinetic and thermodynamic processes of organic species at the solution-solid interface: the view through an STM. Chemical Communications. 2015;51:4737-49.
90. Bhattarai A, Mazur U, Hipps KW. A Single Molecule Level Study of the Temperature-Dependent Kinetics for the Formation of Metal Porphyrin Monolayers on Au(111) from Solution. J Am Chem Soc. 2014;136:2142-8.
91. Hu F, Gong Y, Zhang X, Xue J, Liu B, Lu T, Deng K, Duan W, Zeng Q, Wang C. Temperature-induced transitions of self-assembled phthalocyanine molecular

nanoarrays at the solid-liquid interface: from randomness to order. Nanoscale. 2014;6:4243-9.
92. Blunt MO, Adisoejoso J, Tahara K, Katayama K, Van der Auweraer M, Tobe Y, De Feyter S. Temperature-Induced Structural Phase Transitions in a Two-Dimensional Self-Assembled Network. J Am Chem Soc. 2013;135:12068-75.
93. Song W, Martsinovich N, Heckl WM, Lackinger M. Born-Haber Cycle for Monolayer Self-Assembly at the Liquid-Solid Interface: Assessing the Enthalpic Driving Force. J Am Chem Soc. 2013;135:14854-62.
94. Jahanbekam A, Vorpahl S, Mazur U, Hipps KW. Temperature Stability of Three Commensurate Surface Structures of Coronene Adsorbed on Au(111) from Heptanoic Acid in the 0 to 60°C Range. J Phys Chem C. 2013;117:2914-9.
95. Friesen BA, Bhattarai A, Mazur U, Hipps KW. Single Molecule Imaging of Oxygenation of Cobalt Octaethylporphyrin at the Solution/Solid Interface: Thermodynamics from Microscopy. J Am Chem Soc. 2012;134:14897-904.
96. English WA, Hipps KW. Stability of a Surface Adlayer at Elevated Temperature: Coronene and Heptanoic Acid on Au(111). J Phys Chem C. 2008;112:2026-31.
97. Miyake Y, Nagata T, Tanaka H, Yamazaki M, Ohta M, Kokawa R, Ogawa T. Entropy-Controlled 2D Supramolecular Structures of N,N '-Bis(n-alkyl)-naphthalenediimides on a HOPG Surface. Acs Nano. 2012;6:3876-87.
98. Ikeda T, Asakawa M, Miyake K, Goto M, Shimizu T. Scanning Tunneling Microscopy Observation of Self-Assembled Monolayers of Strapped Porphyrins. Langmuir. 2008;24:12877-82.
99. Otsuki J, Kawaguchi S, Yamakawa T, Asakawa M, Miyake K. Arrays of Double-Decker Porphyrins on Highly Oriented Pyrolytic Graphite. Langmuir. 2006;22:5708-15.
100. Ikeda T, Asakawa M, Goto M, Miyake K, Ishida T, Shimizu T. STM Observation of Alkyl-Chain-Assisted Self-Assembled Monolayers of Pyridine-Coordinated Porphyrin Rhodium Chlorides. Langmuir. 2004;20:5454-9.
101. Gutzler R, Sirtl T, Dienstmaier JF, Mahata K, Heckl WM, Schmittel M, Lackinger M. Reversible Phase Transitions in Self-Assembled Monolayers at the Liquid-Solid Interface: Temperature-Controlled Opening and Closing of Nanopores. J Am Chem Soc. 2010;132:5084-90.
102. Sirtl T, Song W, Eder G, Neogi S, Schmittel M, Heckl WM, Lackinger M. Solvent-Dependent Stabilization of Metastable Monolayer Polymorphs at the Liquid-Solid Interface. ACS Nano. 2013;7:6711-8.
103. Künzel D, Groß A. Influence of the solvent on the stability of bis(terpyridine) structures on graphite. Beilstein Journal of Nanotechnology. 2013;4:269-77.
104. Meier C, Roos M, Künzel D, Breitruck A, Hoster HE, Landfester K, Gross A, Behm RJ, Ziener U. Concentration and Coverage Dependent Adlayer Structures: From Two-Dimensional Networks to Rotation in a Bearing. J Phys Chem C. 2010;114:1268-77.
105. Song W, Martsinovich N, Heckl WM, Lackinger M. Thermodynamics of halogen bonded monolayer self-assembly at the liquid-solid interface. Chemical Communications. 2014;50:13465-8.
106. Song W, Martsinovich N, Heckl WM, Lackinger M. Thermodynamics of 4,4[prime or minute]-stilbenedicarboxylic acid monolayer self-assembly at the nonanoic acid-graphite interface. Phys Chem Chem Phys. 2014;16:13239-47.

107. Killian BJ, Kravitz JY, Gilson MK. Extraction of configurational entropy from molecular simulations via an expansion approximation. J Chem Phys. 2007;127:024107.
108. Zhang J, Bilic A, Reimers JR, Hush NS, Ulstrup J. Coexistence of Multiple Conformations in Cysteamine Monolayers on Au(111). J Phys Chem B. 2005;109:15355-67.
109. Rodziewicz P, Meyer B. Interplay between molecule-molecule and molecule-substrate interactions: first-principles study of fluoroform aggregates on a hexagonal ice (0001) surface. Phys Chem Chem Phys. 2014;16:940-54.
110. Ochterski JW. Thermochemistry in GAUSSIAN: GAUSSIAN Inc. Available from: http://www.gaussian.com/g_whitepap/thermo/thermo.pdf.
111. Reuter K, Scheffler M. Composition, structure, and stability of RuO2 (110) as a function of oxygen pressure. Physical Review B. 2001;65:035406.
112. Curtiss LA, Redfern PC, Raghavachari K. G n theory. WIREs Comput Mol Sci. 2011;1:810-25.
113. Chin Y, Panduwinata D, Sintic M, Sum TJ, Hush NS, Crossley MJ, Reimers JR. Atomic-Resolution Kinked Structure of an Alkylporphyrin on Highly Ordered Pyrolytic Graphite. J Phys Chem Lett. 2011;2:62-6.
114. Floris FM, Tomasi J, Pascual Ahuir JL. Dispersion and Repulsion Contributions to the Solvation Energy: Refinements to a Simple Computational Model in the Continuum Approximation. J Computat Chem. 1991;12:784-91.
115. Coenen MJJ, den Boer D, van den Bruele FJ, Habets T, Timmers KAAM, van der Maas M, Khoury T, Panduwinata D, Crossley MJ, Reimers JR, van Enckevort WJP, Hendriksen BLM, Elemans JAAW, Speller S. Polymorphism in porphyrin monolayers: the relation between adsorption configuration and molecular conformation. Phys Chem Chem Phys. 2013;15:12451-8.
116. Ilan B, Florio GM, Hybertsen MS, Berne BJ, Flynn GW. Scanning Tunneling Microscopy Images of Alkane Derivatives on Graphite: Role of Electronic Effects. Nano Lett. 2008;8:3160-5.
117. Yang T, Berber S, Liu J-F, Miller GP, Tománek D. Self-assembly of long chain alkanes and their derivatives on graphite. J Chem Phys. 2008;128:124709.
118. Hentschke R, Schuermann BL, Rabe JP. Molecular-dynamics simulations of ordered alkane chains physisorbed on graphite. J Chem Phys. 1992;96:6213-21.
119. Yin S, Wang C, Qiu X, Xu B, Bai C. Theoretical study of the effects of intermolecular interactions in self-assembled long-chain alkanes adsorbed on graphite surface. Surf Interface Anal. 2001;32:248-52.
120. Frisch MJ, Trucks GW, Schlegel HB, et al. Gaussian 09, Revision C.09: Gaussian, Inc., Pittsburgh PA; 2009.
121. Battezzati L, Pisani C, Ricca F. Equilibrium conformation and surface motion of hydrocarbon molecules physisorbed on graphite. J Chem Soc, Faraday Trans 2. 1975;71:1629-39.
122. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, S. Profeta J, Weiner P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. J Am Chem Soc. 1984;106:765-84.
123. Hehre WJ, Ditchfield R, Pople JA. Self-consistent molecular orbital methods. XII. Further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules J Chem Phys. 1972;56:2257-61.
124. Kresse G, Hafner J. Ab initio molecular dynamics for liquid metals. Phys Rev B. 1993;47:558-61.

125. Soler JM, Artacho E, Gale JD, Garcia A, Junquera J, Ordejon P, Sanchez-Portal D. The SIESTA method for ab initio order-N materials simulation. J Phys: Condens Matter. 2002;14:2745.
126. Cafe PF, Larsen AG, Yang W, Bilic A, Blake IM, Crossley MJ, Zhang J, Wackerbarth H, Ulstrup J, Reimers JR. Chemisorbed and Physisorbed Structures for 1,10-Phenanthroline and Dipyrido[3,2-a:2',3'-c]phenazine on Au(111). J Phys Chem C. 2007;111:17285-96.
127. Tersoff J, Hamann DR. Theory and application for the scanning tunneling microscope Phys Rev Lett. 1983;50:1998-2001.
128. McConnell HM. Intramolecular charge transfer in aromatic free radicals. J Chem Phys. 1961;35:508.
129. Reimers JR, Hush NS. Electron and energy transfer through bridged systems. III. Tight binding linkages with finite asymmetric band gap. J Photochem Photobiol A. 1994;82:31.
130. Reimers JR, Hush NS. Electron and energy transfer through bridged systems. II. Tight binding linkages with zero asymmetric band gap. Chem Phys. 1990;146:89.
131. Ouyang R, Yan J, Jensen PS, Ascic E, Gan S, Tanner DA, Mao B, Ni L, Zhang J, Tang C, Hush NS, Reimers JR, Ulstrup J. Intermixed adatom and surface bound adsorbates in regular self-assembled monolayers of racemic 2-butanethiol on Au(111) ChemPhysChem. 2015;16: 928–32.
132. Yan J, Ouyang R, Jensen PS, Ascic E, Gan S, Tanner DA, Mao B, Ni L, Zhang J, Tang C, Hush NS, Ulstrup J, Reimers JR. Controlling the stereochemistry and regularity of butanethiol self-assembled monolayers on Au(111) J Amer Chem Soc. 2014;136:17087–94.

Table 1. QM/MM calculated SAM lattice parameters for the L polymorph of C_mP tetraalkylporphyrin SAMs, as defined in Fig. 2 (lengths *a*, *b*, *c* in Å, angles θ,ϕ,α in °, area *A* in Å^2), and the differences to those at the PBE-D3 minimum free-energy commensurate geometry.

	QM/MM							Commensurate Geometry Difference						
m	*a*	*b*	*c*	θ	ϕ	α	*A*	Δ*a*	Δ*b*	Δ*c*	Δθ	Δϕ	Δα	Δ*A*
6	22.1	17.4	14.8	41.9	51.8	42.4	257	-0.3	-0.2	0.1	1.1	-0.1	0.6	-1.0
7	20.3	18.9	15.0	45.0	62.6	20.8	271	1.6	0.3	0.3	-1.7	-3.5	-3.8	16.0
8	21.2	20.2	15.0	42.4	65.3	19.6	289	0.1	-0.1	-0.1	-0.2	-0.6	-7.3	-2.0
9	20.4	21.4	16.6	46.6	69.8	17.7	317	-0.3	0.0	0.6	2.2	-0.3	54.6	7.0
10	21.3	22.6	16.7	44.6	72.0	17.4	338	0.2	-0.1	1.0	2.9	-2.5	19.2	17.0
11	20.4	23.9	18.3	48.0	76.1	15.8	362	-0.3	-0.2	0.9	3.2	-2.0	5.3	9.0
12	21.1	25.0	18.5	46.4	77.9	15.7	382	1.4	0.6	0.7	0.0	-2.5	19.8	36.0
13	20.4	26.4	20.3	49.3	81.0	14.3	408	0.6	0.2	1.1	2.6	-3.4	51.6	31.0
14	21.1	27.4	20.4	47.5	82.7	14.3	428	0.7	0.9	0.6	-0.2	0.1	28.7	26.0
15	20.4	28.9	22.4	50.5	84.9	13.0	453	0.6	0.1	1.1	2.6	-3.7	52.8	33.0
16	21.0	29.8	22.6	49.1	86.3	13.1	473	0.8	0.0	0.8	2.1	-4.1	-4.8	34.0
17	20.4	31.4	24.5	51.4	88.1	12.0	498	0.6	0.0	1.0	2.7	-4.0	53.9	35.0
18	20.9	32.2	24.8	50.3	89.3	12.1	518	1.0	0.0	0.1	0.3	-2.5	30.9	25.0
19	20.4	33.8	26.6	51.9	88.9	11.1	543	0.6	-0.1	1.0	3.0	-2.2	-0.2	37.0
20	20.0	34.2	27.7	54.3	89.9	11.1	555	1.0	0.2	0.2	-0.1	-1.6	-71.1	30.0
21	20.4	36.3	28.9	52.6	86.7	10.4	588	0.6	-0.2	1.0	3.1	2.2	-20.5	39.0
22	19.9	36.6	30.2	55.5	88.4	10.3	600	1.5	0.6	-0.4	-2.3	-0.1	6.3	37.0
23	20.4	38.8	31.1	52.8	84.4	9.7	633	0.5	0.5	0.9	1.4	2.0	-9.7	36.0
24	19.9	39.1	32.5	56.0	86.5	9.7	645	0.2	0.4	0.6	0.6	0.6	62.5	18.0

Table 2. PBE-D3 free-energy ($\Delta G/n$ in kcal mol^{-1}, or $N_A\Delta G/A$ in cal mol^{-1} Å^{-2} where N_A is Avogadro's number) for $C_{13}P$ L polymorph at 14 commensurate lattices (lengths *a, b* in Å, angle θ in °, area *A* in Å^2) near the observed one, as well as its binding energy ($\Delta E^{PBE\text{-}D3}/n$ in kcal mol^{-1}) and differential solvation energy ($\Delta G_s/n$ in kcal mol^{-1}) contributions.

a	*b*	θ	*A*	$\Delta G/n$	$N_A\Delta G/A$	$\Delta E^{PBE\text{-}D3}/n$	$\Delta G_s/n$
21.4	26.6	50.5	439	-0.6	-1.4	-143.8	106.3
21.0	26.6	51.9	439	-4.4	-10.0	-147.5	106.1
19.6	26.6	46.1	376	23.6	62.7	-120.0	106.6
19.2	26.6	47.6	376	28.6	75.9	-115.0	106.7
21.8	27.3	52.0	470	0.0	-0.1	-141.3	104.3
21.3	27.9	52.4	470	1.2	2.6	-138.4	102.7
21.4	27.7	50.9	460	-1.0	-2.2	-142.3	104.4
22.1	27.3	51.1	470	-2.0	-4.2	-143.3	104.4
21.0	25.9	49.5	413	-2.2	-5.2	-145.8	106.7
21.3	26.0	49.1	418	-2.5	-5.9	-146.0	106.6
21.0	28.3	49.9	455	-2.8	-6.2	-142.9	103.2
21.0	27.3	56.9	481	-2.6	-5.4	-142.2	102.6
21.4	25.6	47.9	408	6.5	16.0	-137.2	106.8
21.8	25.9	47.7	418	0.6	1.5	-143.0	106.7

Table 3. Calculated BPE-D3 free-energy $\Delta G^{\text{PBE-D3}}$ per mol of porphyrin for the L and M polymorphs of tetraalkylporphyrin C_mP SAMs on HOPG, along with their component components (PBE-only binding energy ΔE^{PBE}, D3 correction ΔE^{D3}, net binding energy $\Delta E^{\text{PBE-D3}} = \Delta E^{\text{PBE}} + \Delta E^{\text{D3}}$, the thermal correction evaluated QM/MM ΔE_{corr}, and the differential solvation energy ΔE_{s}, evaluated using the PBE-D3 geometries), all in kcal mol^{-1}. The net QM/MM binding energy $\Delta E^{\text{QM/MM}}$ is included for comparison to $\Delta E^{\text{PBE-D3}}$.

	m	$\Delta E^{\text{PBE}}/n$	$\Delta E^{\text{D3}}/n$	$\Delta E^{\text{PBE-D3}}/n$	$\Delta G_{\text{corr}}/n$	$\Delta G_{\text{s}}/n$	$\Delta G^{\text{PBE-D3}}/n$	$\Delta E^{\text{QM/MM}}/n$
L	6	17.8	-107.1	-83.2	20.4	60.1	-2.7	-92
	7	16.8	-115.3	-92.4	22.6	66.1	-3.7	-96.9
	8	22.0	-131.1	-103	24.7	73.6	-4.7	-105.8
	9	22.2	-139.5	-111.2	27.4	80.0	-3.8	-113.5
	10	21.4	-146.2	-118.7	29.8	86.7	-2.2	-122.1
	11	25.4	-160.5	-129	32.2	93.3	-3.5	-130.9
	12	21.8	-165.1	-137.2	34.6	99.7	-2.9	-140.1
	13	27.0	-180.5	-147.4	36.9	106.1	-4.4	-149.2
	14	27.3	-188.9	-155.5	39.2	112.8	-3.5	-158
	15	29.7	-203.2	-167.4	41.6	119.4	-6.4	-167.5
	16	28.0	-207.9	-173.8	44	126.5	-3.3	-176
	17	34.2	-225.7	-185.4	46.2	132.7	-6.5	-185.9
	18	34.6	-172.9	-192.8	48.4	139.9	1.6	-193.9
	19	35.2	-245.6	-204.3	51.1	146.0	-7.2	-204.2
	20	38.6	-256.7	-212	53.5	153.4	-5.1	-212.5
	21	40.0	-271.4	-225.3	55.7	159.1	-10.5	-224.4
	22	39.6	-276.9	-231.2	57.8	166.2	-7.2	-230.8
	23	40.6	-305.1	-258.4	60.3	185.8	-12.3	-242.7
	24	52.7	-324.1	-265.3	62.2	194.9	-8.2	-249.1
	25				64.5			-261.0
	26				66.7			-267.3
	27				68.8			-279.1
	28				71.2			-285.4
M	11	22.6	-121.4	-98.8	24.2	69.6	-5.0	-101.6
	12	15.7	-118.5	-109.8	24.8	78.9	-6.1	-115.7
	13	24.9	-114.6	-102.7	24.6	72.0	-6.1	-103.1

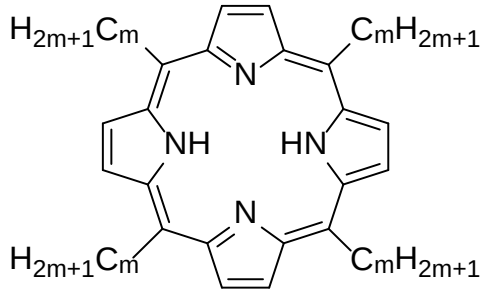

Figure 1. Tetraalklyporphyrins C_mP considered in SAM studies on HOPG from organic solvents; calculations for $6 \le m \le 28$, experiment for $11 \le m \le 19$ but sometimes also for metal-substituted porphyrins M-C_mP.

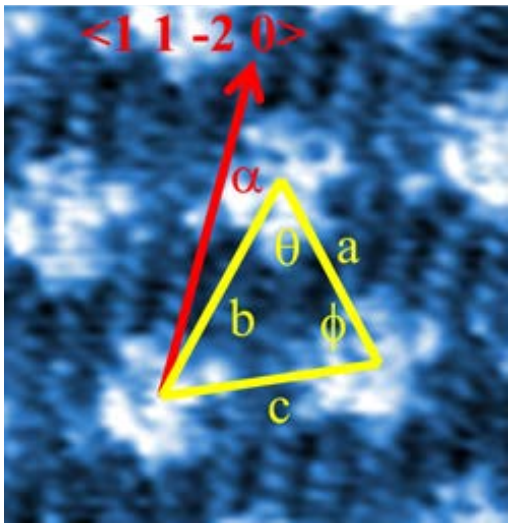

Figure 2. Observed STM image of the L polymorph of $C_{13}P$, overlaid with the definitions of the SAM lattice parameters and the direction of orientation of an HOPG substrate lattice vector.

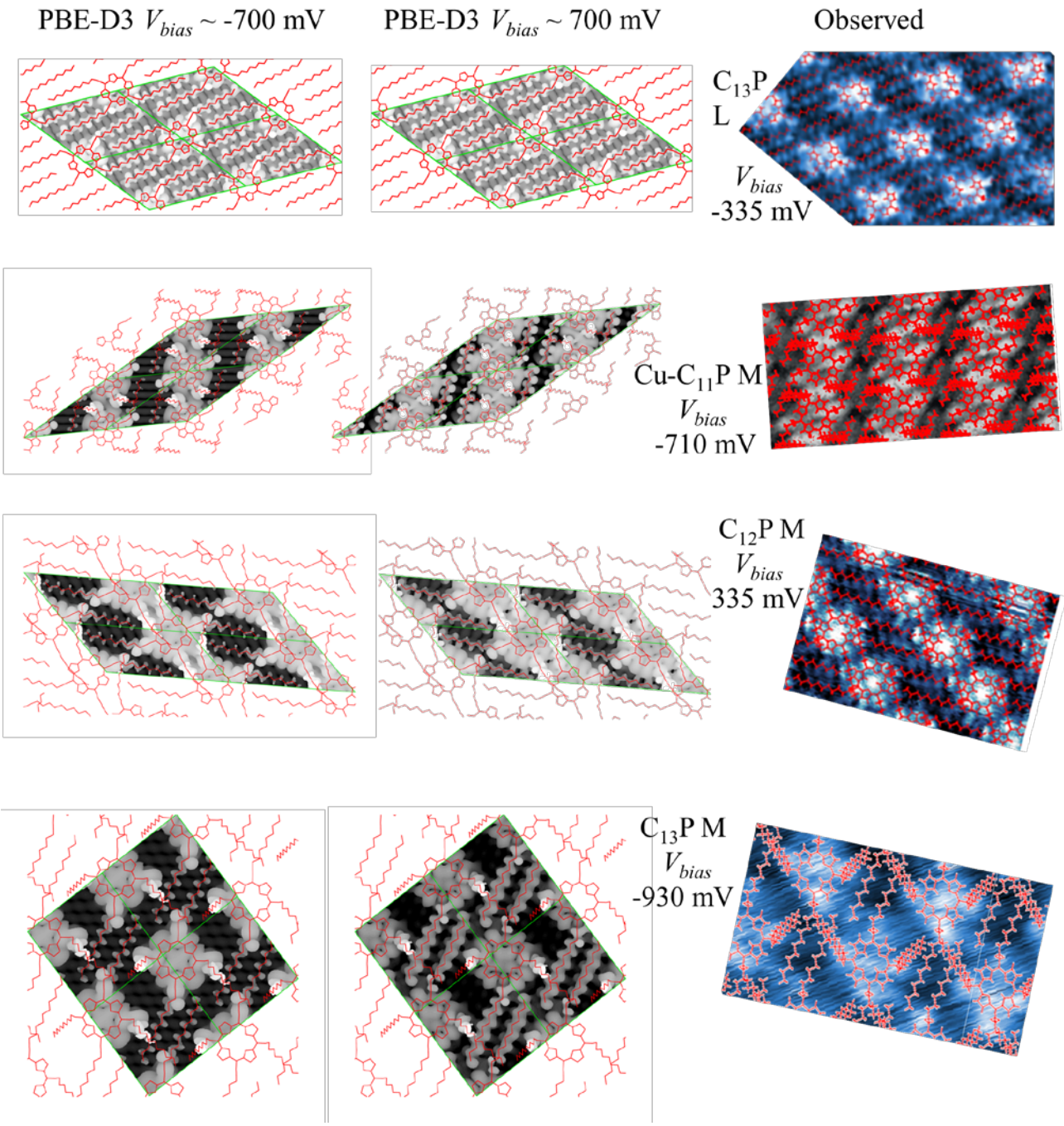

Figure 3. Observed STM image of some L and M polymorphs of Cu-C_{11}P, C_{12}P, and C_{13}P, overlaid with optimized SAM atomic structures (right), compared to simulated STM images at two different bias voltages (left, centre).

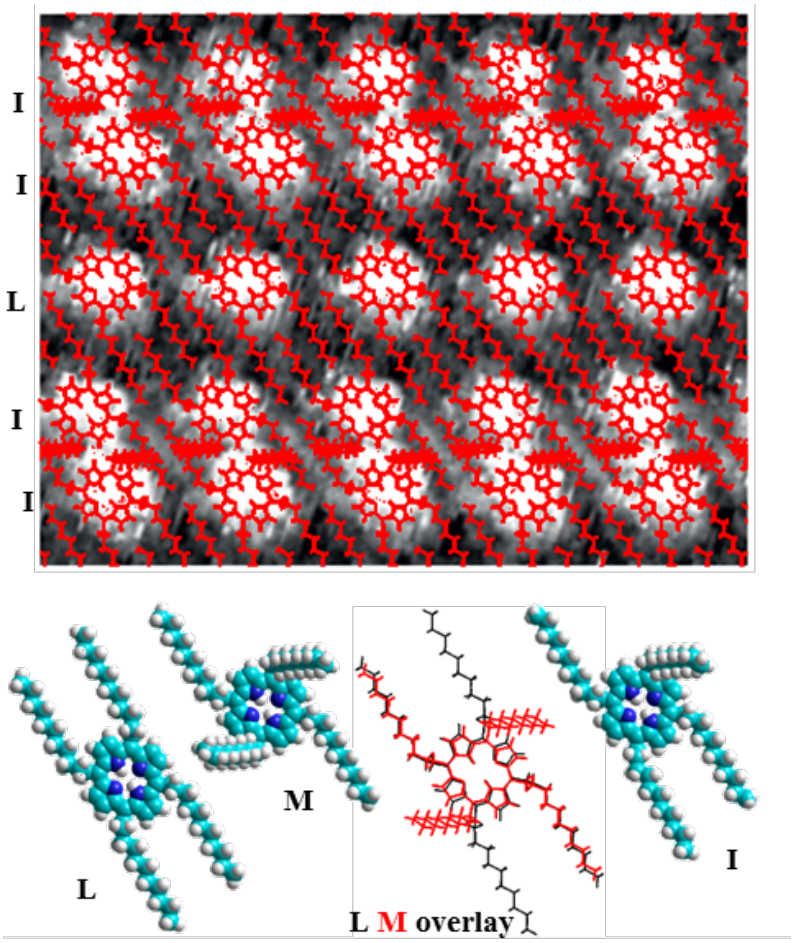

Figure 4. Top: observed STM image of a region of the SAM of Cu-C_{11}P throughout which isolated lines of molecules assemble as in the L polymorph (all 4 chains on surface) adjacent to in lines of interface molecules (3 chains on surface) that connect with each other using the assembly pattern of the M polymorph (2 chains on surface), with overlaid molecules taken from the calculated structure of L (middle) and that as merged with M (outer). Bottom: the construction of an approximate I structure by overlaying fully-optimized L and M structures, selecting the required unique chain conformations, and averaging the remaining coordinates.

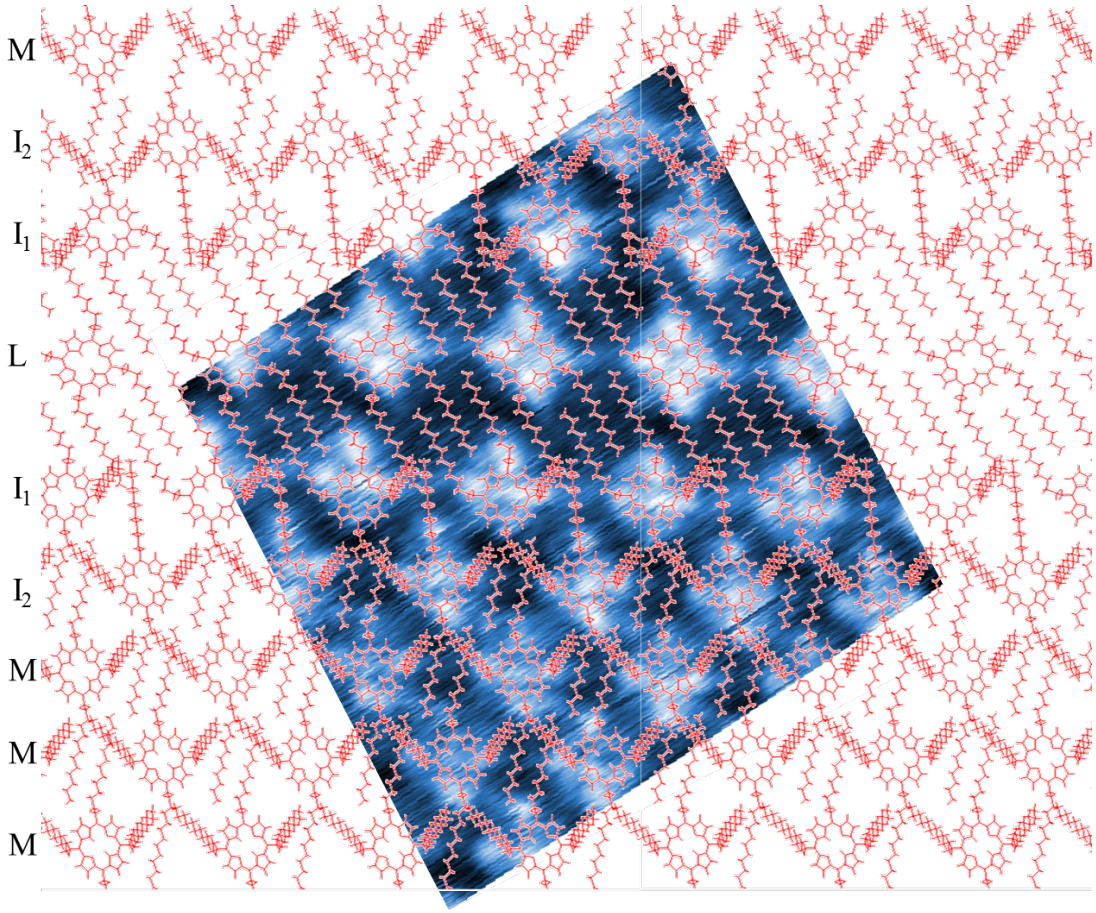

Figure 5. Observed STM image of an interface region of the SAM of $C_{13}P$ throughout which isolated lines of molecules assemble as in the L polymorph (all 4 chains on surface) adjacent to lines of two different types of interface molecules (2 or 3 chains on surface) that link to regions of the M polymorph (2 chains on surface), with overlaid molecules taken from the calculated structure of L, M, and the merged sections of each.

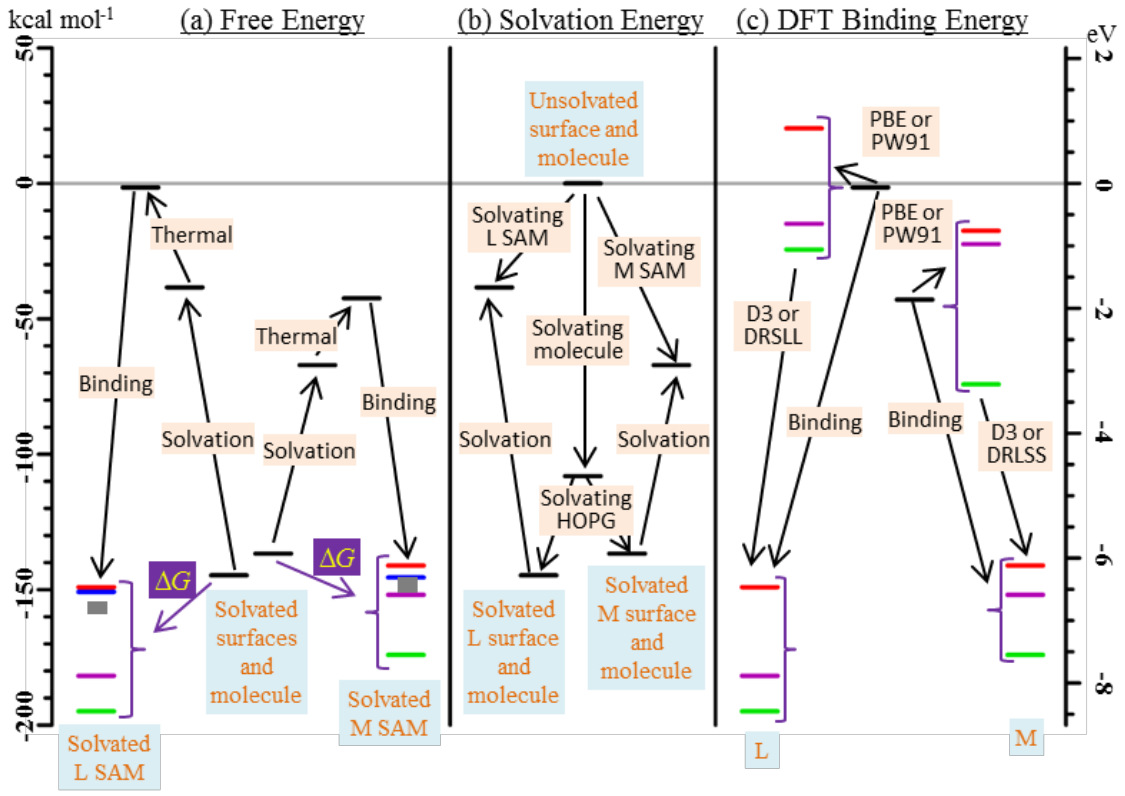

Figure 6. Calculated contributions to the Gibbs free-energy of formation ΔG per mole of porphyrin of the L and M polymorphs of the SAM of $C_{13}P$ on HOPG in 1-phenyloctane solution. Colour code: grey- crudely estimated observed free energies,[87] red- PBE-D3/PAW, blue- QM/MM, purple- PBE-DRSLL/DZP, and green- PW91-S3/USPP.