Supporting Information

Predicting Molecular Self-Assembly at Surfaces: A Statistical Thermodynamics & Modeling Approach

Simone Conti and Marco Cecchini*

Laboratoire d'Ingénierie des Fonctions Moléculaires
ISIS, UMR 7006 CNRS, Université de Strasbourg, F-67083 Strasbourg Cedex, France

Contents

PREVIOUS WORKS ON SELF-ASSEMBLY

In the following we demonstrate how previous literature approaches to 2D self-assembly are contained in the theoretical framework presented in this work. The seminal work by Reuter and Scheffler (**RS**) [1] to model the surface stability of ruthenium oxide is analyzed first. In this approach, the most stable structure and composition is the one that minimizes the surface free energy (Eq. 2 in **RS**)

$$\gamma = \frac{1}{2A_{slab}} \left[G_{slab} - N_{Ru}\mu_{Ru} - N_O\mu_O \right] \tag{S1}$$

with A_{slab} and G_{slab} being the surface area and the free energy of the simulated supercell (DFT and periodic conditions are used), N_{Ru} and N_O are the number of ruthenium and oxygen atoms on the surface, and μ_{Ru} and μ_O their chemical potentials in the gas phase. Here, the area is multiplied by two as the oxide is modeled as a slab with two identical surfaces exposed to the gas phase. Using Eq. S1 for a general assembly reaction producing a SAM made of α molecules of A, the surface free energy of **RS** can be written as

$$\gamma = \frac{1}{A_{sam}} \left[G_{sam} - \alpha\mu_A \right] \tag{S2}$$

Assuming extensibility of the free energy of the supercell, which is a no-condition since a supercell approach assumes it by definition, the relations $A_{sam} = \alpha A'_{uc}$ and $G_{sam} = \alpha\mu'_{uc}$ hold, and yield

$$\gamma = \frac{1}{A'_{uc}} \left[\mu'_{uc} - \mu_A \right] \tag{S3}$$

which is Eq. 19 in the Main Text. Assuming that $\mu'_{uc} = E'_{uc}$, which implies that the translational, rotational and vibrational contributions the chemical potential of the SAM are negligible, **RS** gives

$$\gamma = \frac{1}{A'_{uc}} \left[E'_{uc} - \mu_A \right] \tag{S4}$$

providing the definition of γ by Kučera and Gross (**KG**) [2] as well as that of the ΔG of self-assembly in Meier, Ziener *et al.* (**MZ**) [3] (with $\rho = 1/A'_{uc}$). Although this approximation is reasonable for the chemisorption of gases as shown by **RS**, it is generally too strong for a physisorbed SAM at the solid-liquid interface. In fact, as illustrated in the Main Text for a series of cases, the vibrational contribution to the chemical potential of the SAM or the free monomer in solution can be sizable. However, since the molecular building blocks analyzed here are rather rigid, the vibrational free energy change on self-assembly per molecules is close to zero and can be safely ignored. This observation implies that instead of Eq. S4, a more correct expression of γ would be

$$\gamma = \frac{1}{A'_{uc}} \left[E'_{uc} - (\mu_A - \mu_{vib,A}) \right] \tag{S5}$$

which involves no *a priori* assumption on the magnitude of the vibrational contribution to the chemical potential. It follows that the plots in Fig. 6 of **KC** and Fig. 9 of **MZ** are actually shifted by the quantity $\mu_{vib,A}/A'_{uc}$ on the x axis. In addition and unlike in **RS**, in **KG** and **MZ** there is no link between μ_A and the monomer concentration in solution, which is the experimental quantity under control. Finally, in **KG** the energy of adsorption is the only contribution to the unit cell energy, which implies that E_{sam} and E_{strain} in Eq. 14 of the Main Text are neglected. The strain energy is also neglected in **MZ**.

A similar approach was developed by Loffreda, Delbecq and Sautet (**LDS**) [4], which uses an adaptation of **RS** to study the chemisorption of acrolein on platinum (Eq. 1 in **LDS**)

$$\Delta G = \frac{\Theta}{A_{uc}}\left(\Delta E_{ads} - 3kT + \Delta ZPE - N_A kT \ln\frac{1}{Z_{tr}Z_{rot}} - N_A kT \ln\frac{P}{P_0}\right) \tag{S6}$$

In this formulation the chemical potential of the free monomer is separated into an internal energy contribution arising from the rigid-body rotations and translations of the molecule ($3kT$), an electronic contribution including the zero point vibrational energy that enters into ΔE_{ads} and ΔZPE, an entropy contribution associated with the translational and rotational degrees of freedom, which is accessed through the configurational partition functions Z_{tr} and Z_{rot}, and a term that links directly to the pressure of acrolein in the gas phase. Although, it is not clear why the full vibrational contribution was omitted, Eq. S6 is essentially equivalent to Eq. S21, which can be derived from Eq. 19 in the Main Text by separating out the enthalpy versus entropy contributions, as we shall see.

The same expression for the surface free energy was obtained by Gutzler, Lackinger *et al.* (**GL**) [5] and Dienstmaier, Lackinger *et al.* (**DL**) [6] following a completely different approach. In these works, the free energy change on self-assembly is evaluated from the combination of a per unit of area, per molecule contribution as

$$\Delta g = \frac{\Delta G}{A} = \frac{\Delta H}{A} - \frac{T\Delta S}{A} \tag{S7}$$

with A being the surface area occupied by one molecule, ΔH the energy gain per molecule calculated by force field or empirical interaction energy models, and ΔS the entropy loss per molecule. Clearly, Δg is equivalent to the surface free energy γ presented in the Main Text. Interestingly, in **GL** and **DL** the entropy change in Eq. S7 is evaluated using a statistical mechanics approach. First, the total entropy change is decomposed in translational, rotational, vibrational, and conformational contributions as

$$\Delta S = \Delta S_{tr} + \Delta S_{rot} + \Delta S_{vib} + \Delta S_{conf} \tag{S8}$$

Then, assuming that the molecular building blocks are rigid and that upon self-assembly they completely lose all translational and rotational degrees of freedom, the vibrational and conformational contributions vanish and ΔS can be obtained from the translational and rotational entropy of the solution state as

$$\Delta S = -\left(S_{tr} + S_{rot}\right) \tag{S9}$$

where both S_{tr} and S_{rot} are evaluated analytically using statistical thermodynamics and a free volume correction for the translation entropy [7] (not used in the present work). Introducing this result in Eq. S7 yields

$$\gamma = \frac{1}{A'_{uc}}\left[E'_{uc} + T\left(S_{A,tr} + S_{A,rot}\right)\right] \tag{S10}$$

where the area per molecule A and the enthalpy variation ΔH have been replaced by the corresponding A'_{uc} and E'_{uc} used in the present work. Clearly, the result of Eq. S10 is equivalent to Eq. 19 of the Main Text when the vibrational contribution is neglected. Also, this comparison shows that in the derivation of **GL** and **DL** the $3RT$ corresponding to the internal energy of a rigid monomer is erroneously missing.

Finally, we show that the orthogonal approach by Lei, De Feyter *et al.* (**LDF**) [8], Bellec, Charra *et al.* (**BC**) [9], and Blunt, De Feyter *et al.* (**BDF**) [10] to study the 2D phase transition between a porous and a dense architectures as a function of monomer concentration or temperature can also be reconducted to our theory. In this case, the chemical reaction studied is

$$dD \rightleftharpoons pP + (d-p)\,A \tag{S11}$$

where d molecules in the dense SAM (D) are converted in p molecules in the porous SAM (P) thus releasing $d-p$ molecules in solution. The stoichiometric coefficients d and p are proportional to the surface area covered by one molecule in the two different SAMs, and thus can be expressed as $d = 1/A'_{uc,d}$ and $p = 1/A'_{uc,p}$. At chemical equilibrium, the difference in chemical potentials for the reaction above is zero, which yields (Eq. 1 in **LDF**)

$$d\mu_D = p\mu_P + (d-p)\,\mu_A \tag{S12}$$

with μ_D, μ_P and μ_A the chemical potentials per molecule of the dense, porous and monomeric state, respectively. Remarkably, the result of Eq. S12 can be obtained using the expression of γ in Eq. 22 of the Main Text and by imposing $\gamma_D = \gamma_P$. Based on **LDF**, **BC** developed an expression for the critical concentration of monomers corresponding to a 2D phase transition (Eq. 8 in **BC**)

$$C_0 = \exp\left(-\frac{S_0}{k}\right)\exp\left[-\frac{1}{kT}\left(H_D - \frac{H_P - H_D}{m-1}\right)\right] \tag{S13}$$

Interestingly, by rearranging this result into

$$C_0 = \exp\left[\frac{1}{kT}\left(\frac{H_D + TS_0}{A_D} - \frac{H_P + TS_0}{A_P}\right)\frac{A_P A_D}{A_P - A_D}\right] \tag{S14}$$

with H_i and A_i being the enthalpy gain and unit cell area of the two SAMs, and S_0 is the entropy associated to the confinement of one monomer to the surface, and noticing that $(H_i + TS_0)/A_i$ provides an approximated expression for the surface free energy $\gamma_i = (\mu_i - \mu_A)/A_i$, which is straightforward to see by separating the enthalpy versus entropy contributions, one recovers Eq. 25 of the Main Text. A more complete expression for the critical concentration C_0 was derived by **BDF** when considering the entropy change associated with solvent coadsorption.

Overall, this analysis demonstrates that the principal equations of previously established and conceptually orthogonal methods to predict 2D self-assembly can be rederived based on Eq. 19 of the Main Text.

DECOMPOSITION OF THE SURFACE FREE ENERGY IN ENERGETIC AND ENTROPIC CONTRIBUTIONS

As defined in the Main Text, the surface free energy is

$$\gamma = \frac{1}{A'_{uc}} \left(\mu'_{uc} - \mu_A \right) \tag{S15}$$

where the unit cell chemical potential μ'_{uc} is

$$\mu'_{uc} = E'_{uc} + \mu'_{uc,vib} \tag{S16}$$

and the chemical potential of the monomer in solution μ_A can be decomposed in translational, rotational and vibrational contributions:

$$\mu_A = \mu_{A,tr} + \mu_{A,rot} + \mu_{A,vib} \tag{S17}$$

No energetic contribution for the molecule in solution is present, since, from the definition of the unit cell energy, the energy of the molecule in solution is the zero of the energy scale. Substituting these expressions into the definition of γ one obtains

$$\gamma = \frac{1}{A'_{uc}} \left(E'_{uc} + \mu'_{uc,vib} - \mu_{A,tr} - \mu_{A,rot} - \mu_{A,vib} \right) \tag{S18}$$

where all chemical potential contributions can be expressed as a sum of an energetic and entropic term:

$$\gamma = \frac{1}{A'_{uc}} \left[\left(E'_{uc} + E'_{uc,vib} - E_{A,tr} - E_{A,rot} - E_{A,vib} \right) - T \left(S'_{uc,vib} - S_{A,tr} - S_{A,rot} - S_{A,vib} \right) \right] \tag{S19}$$

Noting that the energy associated with the translations and rotations is $kT/2$ for each degree of freedom, the previous expression can be recast as

$$\gamma = \frac{1}{A'_{uc}} \left[\left(E'_{uc} + \Delta E'_{vib} - 3kT \right) - T \left(\Delta S'_{vib} - S_{A,tr} - S_{A,rot} \right) \right] \tag{S20}$$

If the vibrational energy and entropy are neglected, the simplified expression for γ previously used is recovered [11, 12]:

$$\gamma = \frac{1}{A'_{uc}} \left[E'_{uc} - 3kT + T \left(S_{tr} + S_{rot} \right) \right] \tag{S21}$$

From Eq. S20, it is immediate to write the surface free energy γ as a sum of energetic γ_E and entropic γ_S contributions

$$\gamma = \gamma_E - T\gamma_S \tag{S22}$$

where

$$\gamma_E = \frac{1}{A'_{uc}} \left(E'_{uc} + \Delta E'_{vib} - 3kT \right) \tag{S23}$$

$$\gamma_S = \frac{1}{A'_{uc}} \left(\Delta S'_{vib} - S_{A,tr} - S_{A,rot} \right) \tag{S24}$$

ON THE VIBRATIONAL CONTRIBUTION

The thermodynamic analysis of 2D self-assembly into thirteen distinct SAMs (Table S3) shows that the vibrational contribution is not always negligible. In particular in some cases such as for coronene and perchlorocoronene self-assembly this accounts for more than 10% of the surface free energy; see Tab. S1. In the following, we analyze the vibrational contribution of γ and identify cases where it should be properly accounted for. Eq. S20 presents the complete deconvolution of the chemical potential difference on self-assembly, which includes a vibrational contribution. Following Reuter and Scheffler [1] this term is usually neglected, assuming $\Delta\mu'_{vib} \approx 0$ even on physisorption [3]. As discussed in the Main Text, this approximation is valid in most cases, but in some examples, i.e. coronene self-assembly, it is significant and corresponds to more than the 10% of the total chemical potential difference. To understand the nature of this contribution, it is useful to decompose it in energetic ($\Delta E'_{vib}$) and entropic ($\Delta S'_{vib}$) contributions, such that $\Delta\mu'_{vib} = \Delta E'_{vib} - T\Delta S'_{vib}$.

On the energy contribution, in the classical limit each vibrational degree of freedom contributes for kT kcal mol^{-1}. For one monomer in solution $E_{vib,A}$ is thus $(3N-6)kT$, where N is the number of atoms of the monomer, whereas for a SAM composed of α molecules, the vibrational energy is $(3N\alpha-3)kT$. Using Eq. 12 of the Main Text the vibrational energy contribution per unit cell is

$$E'_{vib,uc} = \lim_{\alpha\to+\infty} \frac{(3N\alpha-3)kT}{\alpha} = 3NkT \tag{S25}$$

and

$$\Delta E'_{vib} = E'_{vib,uc} - E_{vib,A} = 3NkT - (3N-6)kT = 6kT \tag{S26}$$

Therefore, at 300 K 2D self-assembly involves a constant vibrational energy cost of about 3.6 kcal mol^{-1}. This is expected as the number of vibrational degrees of freedom increases upon self-assembly, while the number of rotational and translational degrees of freedom decreases. Even when the correct quantum expression for the vibrational energy is used, which makes this contribution frequency-dependent, the vibrational energy cost remains in the range 3.7-5.2 kcal mol^{-1} (see Tab. S1), and is unable to account for the observed variations in $\Delta\mu'_{vib}$.

Arch.	$\Delta E'_{vib}$	$-T\Delta S'_{vib}$	$\Delta\mu'_{vib}$	$\Delta\mu'_{vib}/\Delta\mu^{\circ\prime}$ %
ISA	4.7	-6.2	-1.6	6.7
TRA	4.7	-6.1	-1.4	5.8
GUA	4.3	-6.5	-2.1	7.6
MEL	4.8	-6.2	-1.4	7.4
COR	3.7	-9.3	-5.6	19.6
CLC	4.3	-9.4	-5.1	12.0
C12	4.5	-4.3	0.2	-1.2
B12	4.6	-3.8	0.8	-4.1
A12	4.8	-3.3	1.5	-6.3
L12	4.4	-4.3	0.1	-0.5
TMA/CHK	5.1	-5.4	-0.2	0.7
TMA/FLW	5.2	-5.3	-0.1	0.3
TMA/SFW	5.2	-5.2	0.0	0.0

TABLE S1: Difference in vibrational energy, entropy and chemical potential for the self-assembly of the thirteen SAMs studied in this work. All values are reported in kcal mol^{-1}.

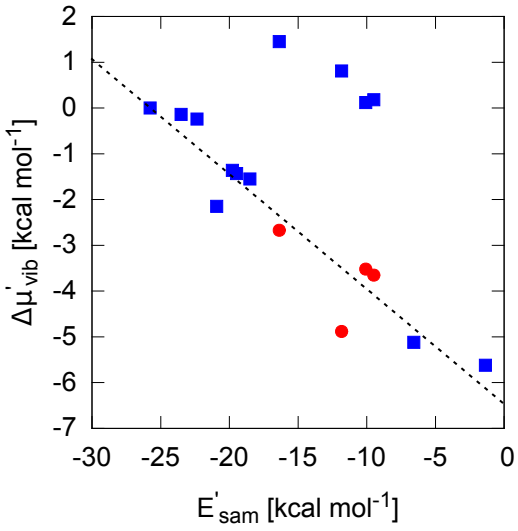


FIG. S1: Correlation between the difference in vibrational chemical potential and the energy of the SAM.

In sharp contrast, the entropy contribution, which is also reported in Tab. S1, spans a significantly wider range (i.e. from -3.3 kcal mol^{-1} to -9.4 kcal mol^{-1}) and is responsible for the variability of $\Delta\mu'_{vib}$ in Tab. S1. Since all considered molecules are rather rigid, we expect no significant change in the internal vibrations upon self-assembly and the vibrational entropy stabilization of the SAM ($-T\Delta S'_{vib} < 0$) arising from the intermolecular or external vibrations of the SAM. If so, one would expect weaker intermolecular interactions in the SAM, such as Van der Waals contacts in coronene self-assembly, to produce softer external vibrations than the strong hydrogen bonding in TMA self-assembly and corresponds to a larger vibrational entropy stabilization. To check for this, the correlation between $\Delta\mu'_{vib}$ and the energy of the SAM (E'_{sam}) has been explored. Fig. S1 shows that the majority of the blue points lies on the linear correlation (dashed line), which implies that, as expected, the weaker the interactions in the SAM, the larger the vibrational contribution to γ. The only exceptions to this rule are the four points at the top of the plot, which correspond to the linear alkane (C12) and its chemical functionalizations. Interestingly, those are the only molecules of the dataset that are flexible. Upon self-assembly, these molecules, which find themselves in a dense molecular environment in full interaction with the substrate, actually lose vibrational entropy through conformational confinement of the dihedral angles which may counterbalance the vibrational entropy stabilization resulting from the external vibration of the SAM. To check for this assumption, the dihedrals of these four molecules have been restrained by an external harmonic potential to reduce their internal flexibility, and the calculation of γ repeated. Strikingly, in the absence of internal flexibility all points lie on a linear correlation ($R^2 = 0.91$). Based on this analysis we can delineate two limiting scenarios. First, for rigid molecule the weaker the interactions in the SAM, the more favorable the vibrational entropy contribution will be. By contrast, very strong interactions make the vibrational contribution negligible. Second, for flexible molecules the stronger the intermolecular interactions in the SAM, the larger the entropy destabilization to self-assembly as the internal entropy loss is not compensated by an external entropy gain. Vice versa, for flexible molecules and week interactions, the vibrational contribution to γ is negligible. As a "rule of thumbs", the following scheme can be followed:

		Molecule	
		Flexible	Rigid
Interactions	Strong	Unfavorable	Negligible
	Weak	Negligible	Favorable

SOLVATION FREE ENERGY

Solvent effects on the probability of 2D self-assembly can be directly included in the definition of the surface free energy as

$$\gamma = \gamma_E - T\gamma_S + \gamma_{solv} \tag{S27}$$

where γ_E and γ_S correspond to the energetic and entropic contributions to the surface free energy (see above), and γ_{solv} is the contribution of the solvent which is defined as

$$\gamma_{solv} = \frac{1}{A'_{uc}} \Delta\mu'_{solv} \tag{S28}$$

with $\Delta\mu'_{solv}$ being the solvation free energy change per molecule upon 2D self-assembly, that is

$$\Delta\mu'_{solv} = \mu'_{solv,sam} - \mu_{solv,A} - \mu'_{solv,sub} \tag{S29}$$

where the three term in r.h.s. are the per-molecule contribution to the solvation free energy of the SAM, the monomer, and the portion of substrate covered by the SAM. A rigorous evaluation of $\Delta\mu'_{solv}$ is computationally challenging and would require intensive free energy calculations. Based on previous work [13, 14], the solvation free energy is approximated here as a linear function of the solvent accessible surface area (SASA)

$$\mu_{i,solv} = \alpha \cdot \mathrm{SASA}_i \tag{S30}$$

with α being an empirical parameter which depends on the solvent. Introducing this result in Eq. S29 for a SAM of N molecules and assuming the same α for the both SAM and the substrate, it yields

$$\Delta\mu'_{solv} = \frac{\alpha}{N} \left[\mathrm{SASA}_{sam} - N \cdot \mathrm{SASA}_A - \mathrm{SASA}_{sub}\right] = \alpha\, \Delta\mathrm{SASA}' \tag{S31}$$

which provides numerical access to the solvent contribution to the surface free energy (Eq. S28) from the only knowledge of the SASA of the SAM, monomer and substrate, and the value of α.

In this work, the solvation correction in Eq. S31 was included in the analysis of TMA self-assembly to quantify its impact on the critical concentrations predicted by the theory. In the calculations, the solvent accessible surface area was evaluated using the APBS software [15] with van der Waals radii taken from the GAFF force field [16, 17] and a probe radius of 1.4 Å. $\Delta\mathrm{SASA}'$ was evaluated using model architectures of about 200 molecules (200 for CHK, 216 for FLW and 196 for SFW). The value of α was obtained from the solvation free energy of model compounds and their SASA using Eq. S30. To this aim, the solvation free energy of TMA (the molecule studied) and coronene (as a model of graphene) have been determined by free energy perturbation (FEP) / molecular dynamics (MD) calculations in the apolar solvent toluene. For FEP, 10 and 21 windows of 1 ns each were used to evaluate the polar (electrostatic) and the non polar (van der Waals) contributions to the solvation free energy, respectively. All simulations were carried out using GROMACS 5.1.2 [18], and the FEP analysis was performed using the Alchemical Analysis tool [19]. The μ_{solv} calculated for TMA and coronene were $-17.12\ \mathrm{kcal\,mol^{-1}}$ and $-19.35\ \mathrm{kcal\,mol^{-1}}$, respectively. Their SASA were 3.90 $\mathrm{nm^2}$ and 5.11 $\mathrm{nm^2}$, respectively, which provides values for the parameter α of $-4.39\ \mathrm{kcal\,mol^{-1}\,nm^{-2}}$ and $-3.79\ \mathrm{kcal\,mol^{-1}\,nm^{-2}}$. Hence, an approximated value of $-4\ \mathrm{kcal\,mol^{-1}\,nm^{-2}}$ for α appears to be reasonable for modeling the solvation free energy of TMA self-assembly using Eq. S31. With this value of α and the $\Delta\mathrm{SASA}'$ for the CHK, FLW and

SFW architectures, i.e. -3.66, -3.72, and -3.85 nm^2, respectively, the solvent contribution to the free energy of self-assembly per molecule (Eq. S29) were 14.64, 14.88, and 15.40 $\mathrm{kcal\,mol}^{-1}$, for the different architectures. Introducing these results in Eq. S28 yields surface free energy corrections of 12.09, 15.15, and 20.10 $\mathrm{kcal\,mol}^{-1}\,\mathrm{nm}^{-2}$ for CHK, FLW, and SFW. As shown in the Main Text, this causes an almost systematic shift of the critical concentrations of about ten orders of magnitude.

ENERGY PER UNIT CELL LIMIT

Given a layer of $n \times n$ cells it will have $(n-2)^2$ internal cells and the remaining $4(n-1)$ will be border cells with missing interactions. The energy of such assembly assembly can be approximated as

$$E(n) \simeq (n-2)^2 E_{uc} + 4(n-1)E_{border} \tag{S32}$$

Given

$$\alpha = n_{cells} \cdot n_{uc} \tag{S33}$$

and

$$n_{cells} = n^2 \tag{S34}$$

the previous expression can be written as function of α as

$$E(\alpha) \simeq \left(\sqrt{\frac{\alpha}{n_{uc}}} - 2\right)^2 E_{uc} + 4\left(\sqrt{\frac{\alpha}{n_{uc}}} - 1\right) E_{border} = \tag{S35}$$

$$= \alpha E'_{uc} + 4\left(E_{uc} - E_{border}\right)\left(1 - \sqrt{\frac{\alpha}{n_{uc}}}\right) \tag{S36}$$

When the energy of the self-assembly divided by α is studied:

$$\frac{E(\alpha)}{\alpha} \simeq E'_{uc} - \frac{4\left(E'_{uc} - E'_{border}\right)}{\sqrt{\alpha/n_{uc}}} + \frac{4\left(E'_{uc} - E'_{border}\right)}{\alpha/n_{uc}} \tag{S37}$$

Discarding the third term, which is negligible if compared with the second, the equation reported in the Main Text is obtained

$$\frac{E(\alpha)}{\alpha} \simeq E'_{uc} - \frac{b}{\sqrt{\alpha/n_{uc}}} \tag{S38}$$

where $b = 4\left(E'_{uc} - E'_{border}\right)$.

HYDROGEN BOND CORRECTION FOR TRIMESIC ACID

From the force field energies obtained for the trimesic acid in the three assemblies (chickenwire, flower and superflower) it is clear that the GAFF force field is not adequate to describe the energetics involved in the carboxylic acid interaction. In fact, the superflower architecture is predicted to be the energetically most favored, while experimentally the chickenwire and the flower architectures are clearly more stable [20]. To understand the problem, the interaction energy for the two typical patterns of hydrogen bonds present in the chickenwire, flower and superflower architectures, see Fig. S2, were benchmarked. The first kind of interaction involves linear dimers of carboxylic acid groups and it is know to be highly stabilized by resonance, as a prototypical example of resonance-assisted hydrogen bond (RAHB) [21]. High level of theory *ab-initio* calculations on this dimer are available from the S66 database, where the dimerization energy of acetic acid dimer is -19.41 kcal mol^{-1} [22]. The second hydrogen bond pattern involves three carboxylic groups and, as opposite to the first pattern, no reference energy value was found present in literature.

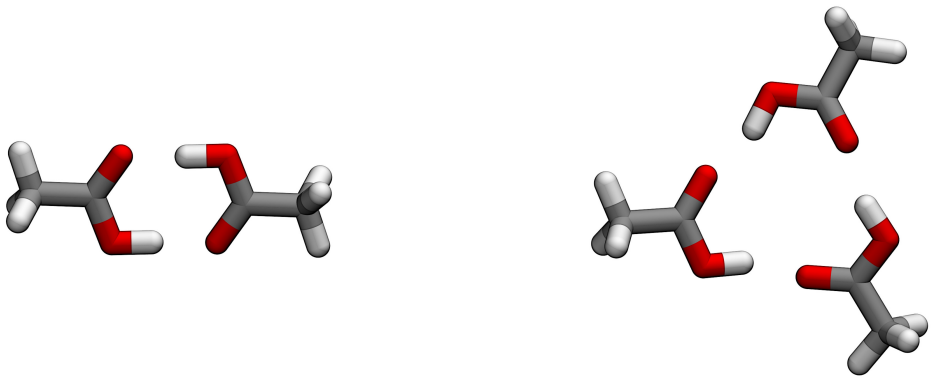


FIG. S2: Sketch representations of the linear (left) and trigonal (right) hydrogen bond patterns of carboxylic acid self-assembly.

Since no reference values for the trimer patter are present in literature, *ab-initio* calculation have been performed on both systems. The chosen level of theory is a density-fitted (DF) MP2 with an extended basis sets (the aug-cc-pVTZ and the cc-pVQZ have been used). The system was geometry optimized and then the interaction energy calculated using a counterpoise correction (CP) to minimize the basis set superimposition error (BSSE). The resulting dimerization energies are reported in Table S2. When comparing the energy for the acetic acid dimer with the CCSD(T)/CBS literature value, the DF-MP2/cc-pVQZ gives clearly optimal results with an error of only 0.26 kcal mol^{-1}. If the reference energy value of -16.41 kcal mol^{-1} is compared with the GAFF value of -14.80 kcal mol^{-1}, it is evident that this force field underestimates the interaction energy by -4.5 kcal mol^{-1}. In sharp contrast, for the trigonal pattern the recognition strength by GAFF matches the value predicted by DF-MP2/cc-pVQZ. A possible reason for this could be that while the nature of the interaction in the two cases are completely different (resonance effect in the first, mainly electrostatic in the second), the force field models both in a pure electrostatic way.

Given these results, the energy per unit cell for the three architectures of trimesic acid have been *a posteriori* corrected: for all hydrogen bond patterns of the first type found in the three architectures, a stabilization energy of -4.5 kcal mol^{-1} has been added, while for the second type of hydrogen bond pattern no correction is added. Focusing on just the first pattern of hydrogen bonds, the chickenwire architecture has 1 occurrence inside the unit cell, plus 4 occurrences shared with the surrounding cells. It means a total of 3 occurrences per unit cell (1 intra plus 4/2 inter),

	Acetic acid dimer	Acetic acid trimer
df-MP2/aug-cc-pVTZ	-18.81	-24.47
df-MP2/cc-pVQZ	-19.15	-24.68
CCSD(T)/CBS [22]	-19.41	//
GAFF	-14.80	-24.78
Empirical Correction	-4.5	//

TABLE S2: Interaction energies for the two patterns of hydrogen bonds evaluated at ab-initio (DF-MP2) and force field (GAFF) levels of theory. All values are reported in kcal mol^{-1}.

which correspond to a total energy correction of 3 times the -4.5 kcal mol^{-1}. Once normalized per the number of molecules inside the unit cell (2 for the chickenwire architecture) a final correction on E'_{uc} of -6.75 kcal mol^{-1} is calculated. For the flower architecture the hydrogen bond pattern count give 2 occurrences inside the cell and 8 between neighbors, which give a total correction on E'_{uc} of -4.5 kcal mol^{-1}. In the superflower architecture, no occurrences of the first pattern are present, so no correction on the energy per unit cell is applied. In Table S3, both the uncorrected and corrected values, marked with an asterisk (*), are reported.

MODELED SELF-ASSEMBLED LAYERS

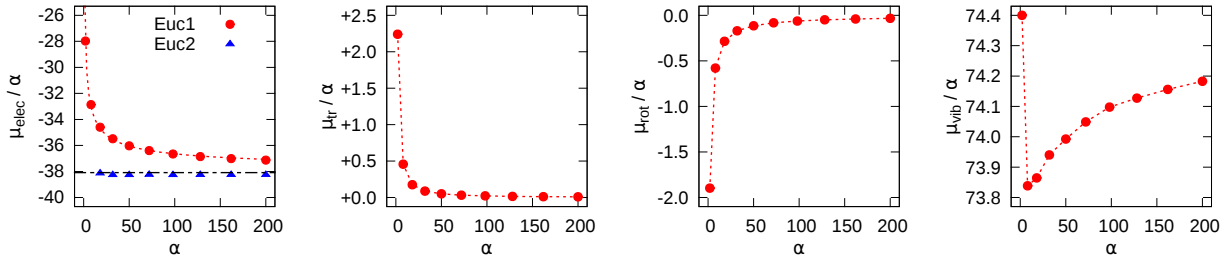
For each of the 13 self-assembled monolayers cited in the Main Text, atomistic molecular models have been built to access their thermodynamic stability. The general procedure used to build the atomistic models is as follow. Starting from the STM images, we first modeled the configuration of the molecules inside one unit cell. To do so, one tries to collect as much information as possible from the bright spots of the STM images. This includes: the most probable position of functional groups, or the orientation of the molecule inside the unit cell. Once a possible configuration of the unit cell was available, an extended supramolecular architecture was built by replicating it onto the xy plane, using the experimental unit cell parameters as initial guess for the translation vectors. The model architecture was then energy minimized by 1000 steps of steepest descent (SD) followed by an adopted basis Newton Raphson (ABNR) optimization until convergence to an energy gradient of 10^{-8} kcal mol^{-1}Å^{-1}. The self-assembled architectures were benchmarked against the STM images mainly comparing the unit cell parameters. In case of significant disagreement, the conformation of the molecule inside the initial unit cell was (manually) modified, optimized on the surface and used as seed to produce a new architecture to be compared with the STM images until a reasonable agreement with the experimental images was achieved. If more than one model structures were compatible with the low-resolution of the STM images, the one with the lowest surface free energy was chosen. Although a local optimization (energy minimization) is performed at every step of the conformational search, no global optimization strategy was available to identify automatically the most probable solution.

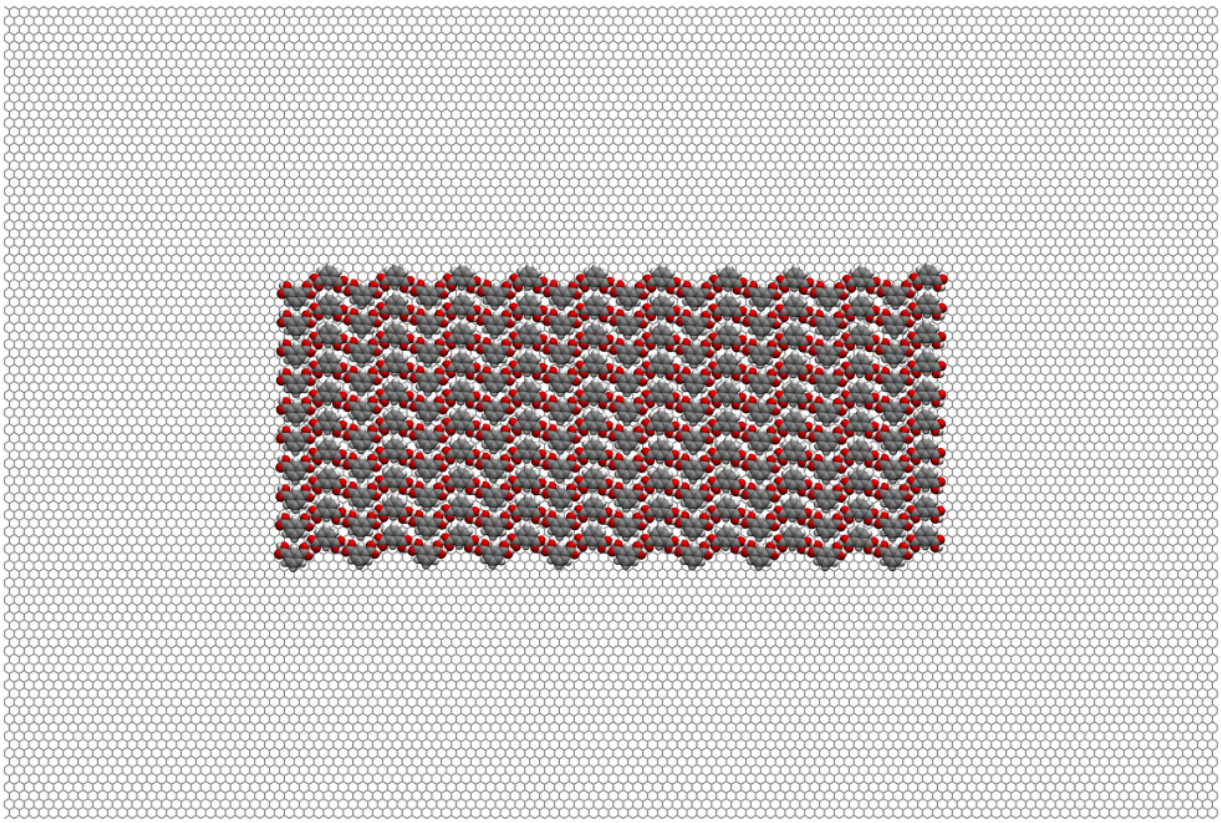
In the table below, the calculated thermodynamic quantities for all modeled architectures are reported. In the next pages, all architectures are showed, with the plots of the convergence of the unit cell chemical potential when the size of the modeled SAMs is increased.

Arch.	Monomer in solution				Self-Assembled Monolayers						Difference						
	μ°_{tr}	μ_{rot}	μ_{vib}	μ°_{tot}	E'_{sam}	E'_{ads}	E'_{strain}	E'_{uc}	μ'_{vib}	μ'_{uc}	$\Delta\mu'_{vib}$	$\Delta\mu^{\circ\prime}$	A'_{uc}	γ°	γ_E	$-T\gamma^{\circ}_S$	$\log C_{cac}$
ISA	-9.0	-8.4	75.7	58.3	-18.5	-20.8	1.0	-38.3	74.2	35.9	-1.5	-22.4	59.9	-37.4	-59.1	21.7	-16.3
TRA	-9.0	-7.5	75.7	59.2	-19.5	-20.8	0.9	-39.4	74.3	34.9	-1.4	-24.3	59.0	-41.2	-61.7	20.5	-17.7
GUA	-9.0	-8.4	101.1	83.7	-20.9	-23.5	1.5	-42.9	99.0	56.1	-2.1	-27.6	65.9	-41.9	-61.3	19.4	-20.1
MEL	-8.7	-7.7	66.0	49.6	-19.8	-16.3	2.3	-33.8	64.6	30.8	-1.4	-18.8	44.8	-42.0	-68.5	26.5	-13.7
COR	-9.5	-7.8	181.7	164.4	-1.4	-38.9	0.0	-40.3	176.1	135.8	-5.6	-28.6	114.2	-25.0	-33.5	8.5	-20.8
CLC	-10.3	-10.1	99.5	79.1	-6.6	-58.2	7.1	-57.7	94.4	36.7	-5.1	-42.4	163.9	-25.9	-33.7	7.8	-30.9
C12	-9.0	-8.3	214.4	197.1	-9.5	-25.2	0.1	-34.6	214.6	180.0	0.2	-17.1	77.5	-22.1	-41.3	19.2	-12.4
B12	-9.1	-8.9	215.2	197.2	-11.8	-26.6	0.2	-38.2	216.0	177.8	0.8	-19.4	81.2	-23.9	-43.6	19.7	-14.1
A12	-9.2	-9.0	202.8	184.6	-16.4	-28.0	0.9	-43.5	204.3	160.8	1.5	-23.8	81.8	-29.1	-49.6	20.5	-17.3
L12	-9.2	-9.0	207.7	189.5	-10.1	-27.7	0.1	-37.7	207.8	170.1	0.1	-19.4	82.8	-23.4	-42.4	19.0	-14.1
TMA/CHK	-9.2	-8.2	81.4	64.0	-22.4	-25.7	1.3	-46.8	81.2	34.4	-0.2	-29.6	121.1	-24.4	-35.8	11.4	-21.6
TMA/CHK*	-9.2	-8.2	81.4	64.0	-29.1	-25.7	1.3	-53.5	81.2	27.7	-0.2	-36.3	121.1	-30.0	-41.4	11.4	-26.5
TMA/FLW	-9.2	-8.2	81.4	64.0	-23.5	-25.7	1.3	-47.9	81.3	33.4	-0.1	-30.6	98.2	-31.2	-45.3	14.1	-22.3
TMA/FLW*	-9.2	-8.2	81.4	64.0	-28.0	-25.7	1.3	-52.4	81.3	28.9	-0.1	-35.1	98.2	-35.7	-49.8	14.1	-25.6
TMA/SFW	-9.2	-8.2	81.4	64.0	-25.8	-25.7	1.3	-50.2	81.4	31.2	0.0	-32.8	76.6	-42.8	-61.1	18.3	-23.9
TMA/STR	-9.2	-8.2	81.4	64.0	-15.9	-25.7	0.9	-40.7	79.9	39.2	-1.5	-24.8	117.0	-21.2	-32.3	11.1	-18.1

TABLE S3: All energies and chemical potentials are reported in kcal mol^{-1}, the unit cell area in Å^2, the values of γ in kcal mol^{-1} nm^{-2}, the $\log C_{cac}$ is the decimal logarithm of the molar concentration. The asterisk (*) indicates that the interaction energy in the SAM is corrected to fit quantum data (see previous section).

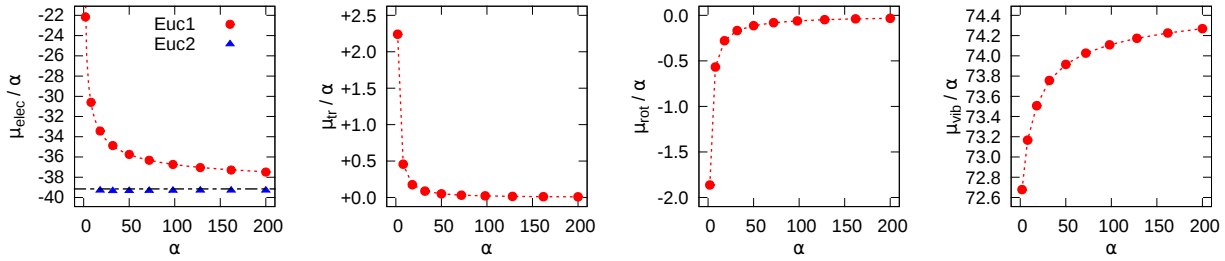
Isophthalic Acid

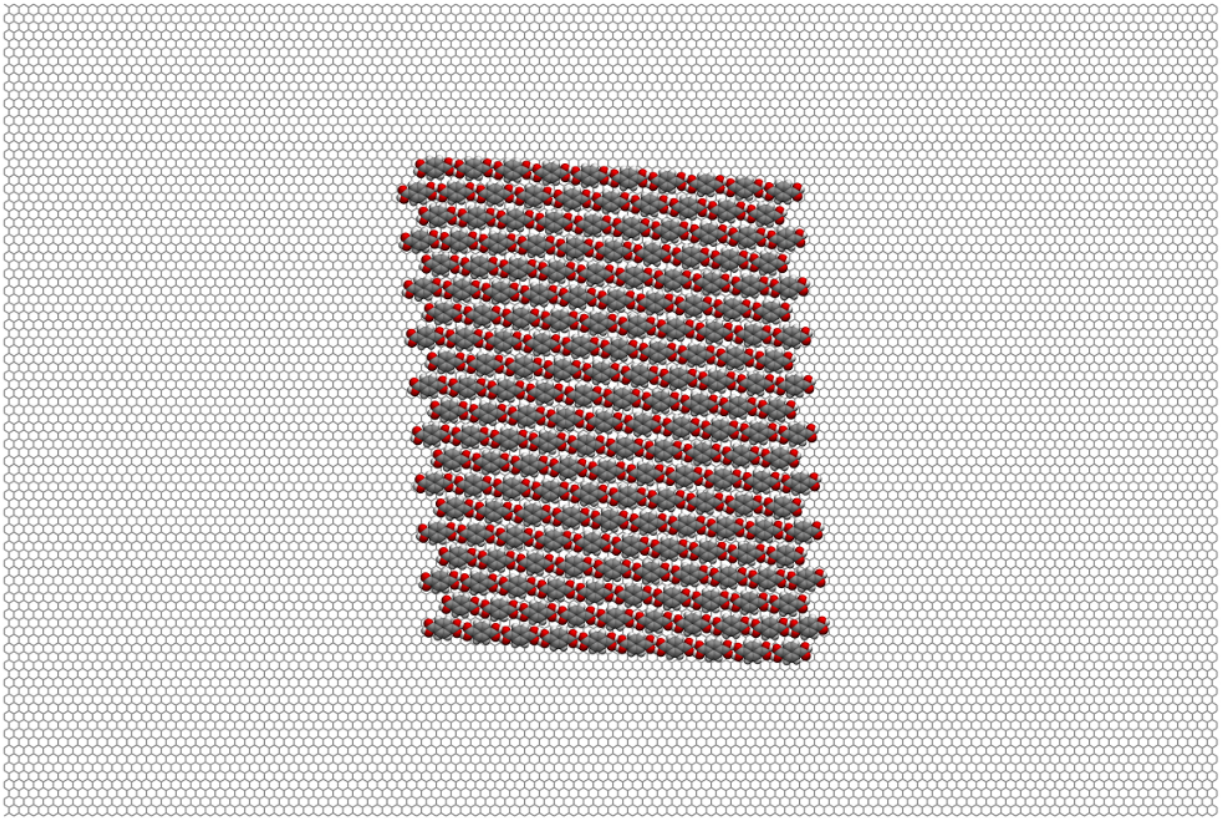




$\mu^{\circ}_{A,tr}$ = -9.0	E'_{ads} = -20.8	$\mu'_{vib,uc}$ = 74.2	A'_{uc} = 59.9
$\mu_{A,rot}$ = -8.4	E'_{sam} = -18.5	μ'_{uc} = 35.9	γ° = -37.4
$\mu_{A,vib}$ = 75.7	E'_{strain} = 1.0	$\Delta\mu'_{vib}$ = -1.5	γ_E = -59.1
$\mu^{\circ}_{A,tot}$ = 58.3	E'_{uc} = -38.3	$\Delta\mu^{\circ\prime}$ = -22.4	$-T\gamma^{\circ}_{S}$ = 21.7
			$C_{cac} = 10^{-16.3}$

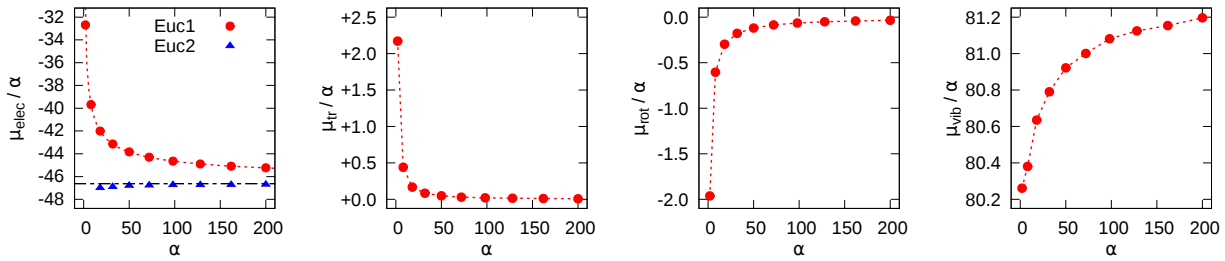
Terephthalic Acid

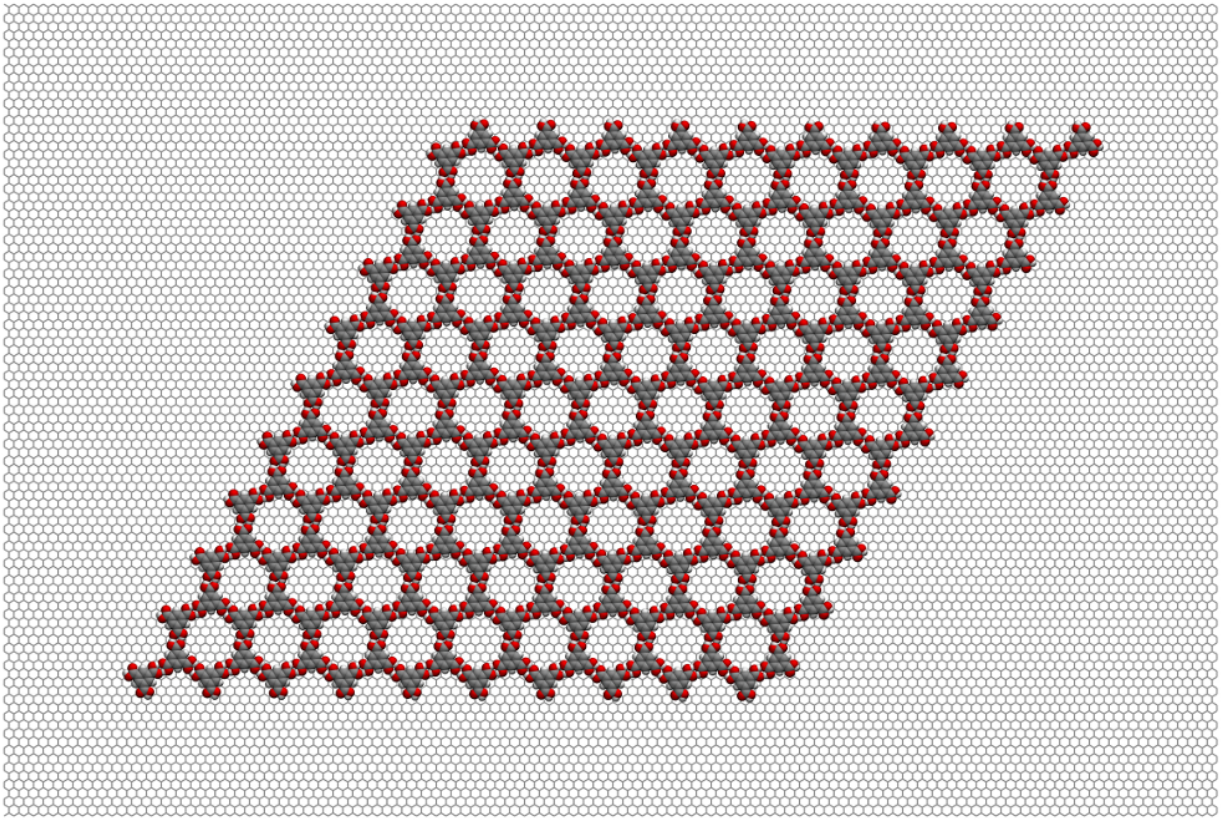




$\mu^{\ominus}_{A,tr}$ = -9.0	E'_{ads} = -20.8	$\mu'_{vib,uc}$ = 74.3	A'_{uc} = 59.0
$\mu_{A,rot}$ = -7.5	E'_{sam} = -19.5	μ'_{uc} = 34.9	$\gamma^{\ominus}$ = -41.2
$\mu_{A,vib}$ = 75.7	E'_{strain} = 0.9	$\Delta\mu'_{vib}$ = -1.4	γ_E = -61.7
$\mu^{\ominus}_{A,tot}$ = 59.2	E'_{uc} = -39.4	$\Delta\mu^{\ominus\prime}$ = -24.3	$-T\gamma^{\ominus}_S$ = 20.5
			$C_{cac} = 10^{-17.7}$

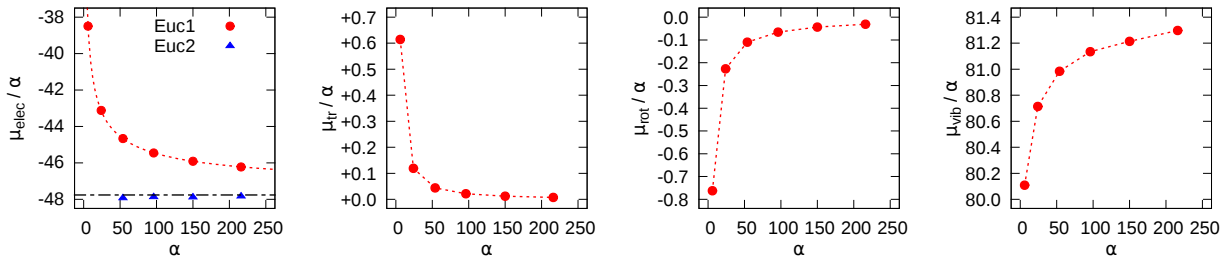
Trimesic Acid – Chickenwire Architecture

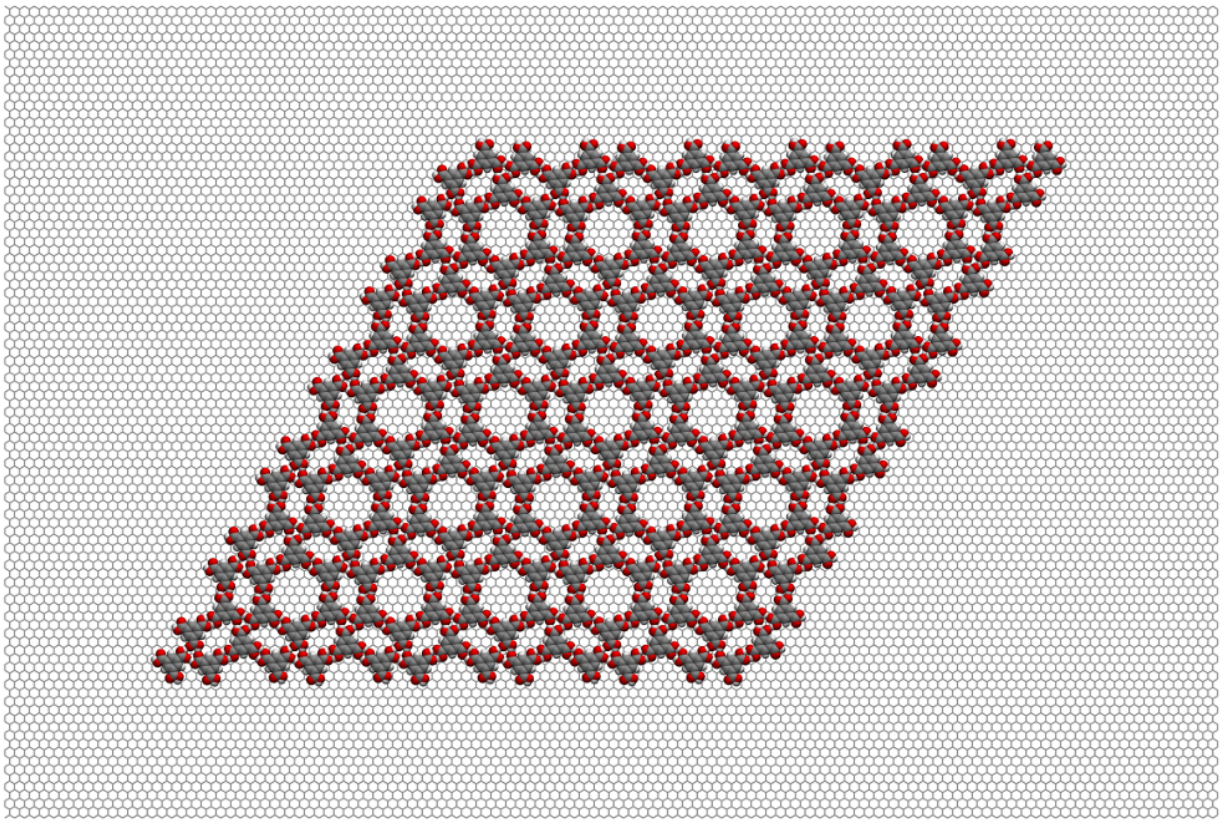




$\mu^{\ominus}_{A,tr} =$ -9.2	$E'_{ads} =$ -25.7	$\mu'_{vib,uc} =$ 81.2	$A'_{uc} =$ 121.1
$\mu_{A,rot} =$ -8.2	$E'_{sam} =$ -22.4	$\mu'_{uc} =$ 34.4	$\gamma^{\ominus} =$ -24.4
$\mu_{A,vib} =$ 81.4	$E'_{strain} =$ 1.3	$\Delta\mu'_{vib} =$ -0.2	$\gamma_E =$ -35.8
$\mu^{\ominus}_{A,tot} =$ 64.0	$E'_{uc} =$ -46.7	$\Delta\mu^{\ominus\prime} =$ -29.6	$-T\gamma^{\ominus}_S =$ 11.4
			$C_{cac} = 10^{-21.6}$

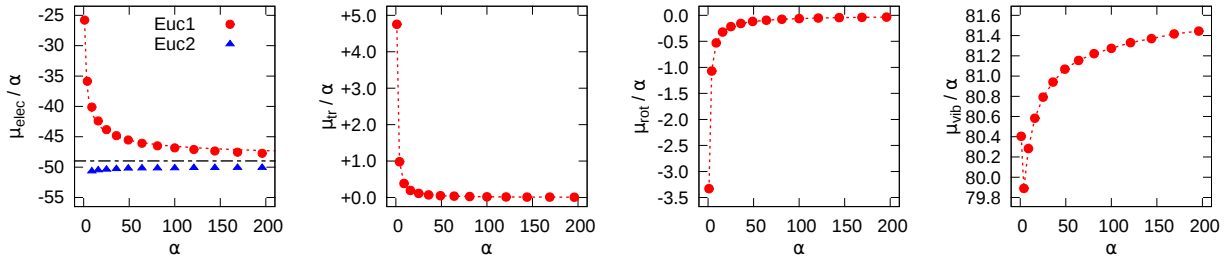
Trimesic Acid – Flower Architecture

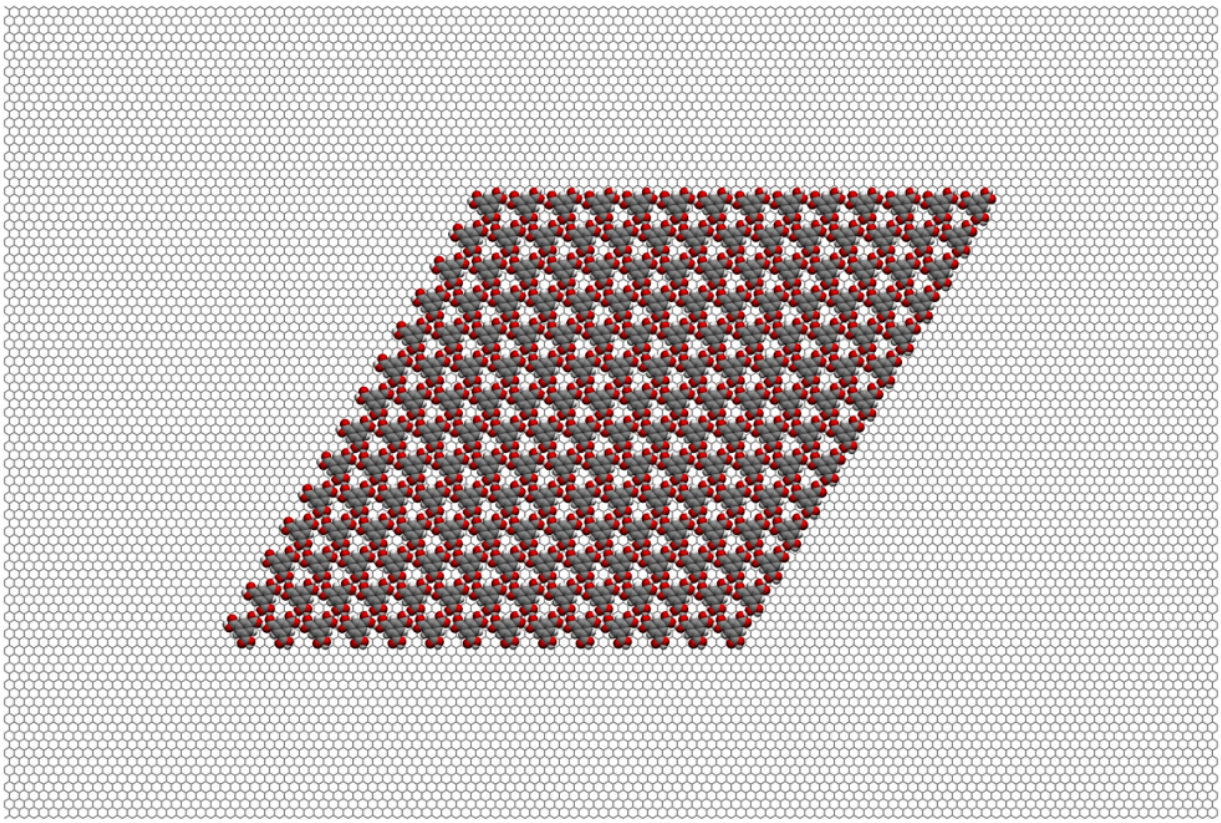




$\mu^{\ominus}_{A,tr} =$ -9.2	$E'_{ads} =$ -25.7	$\mu'_{vib,uc} =$ 81.3	$A'_{uc} =$ 98.2
$\mu_{A,rot} =$ -8.2	$E'_{sam} =$ -23.5	$\mu'_{uc} =$ 33.4	$\gamma^{\ominus} =$ -31.2
$\mu_{A,vib} =$ 81.4	$E'_{strain} =$ 1.3	$\Delta\mu'_{vib} =$ -0.1	$\gamma_E =$ -45.3
$\mu^{\ominus}_{A,tot} =$ 64.0	$E'_{uc} =$ -47.9	$\Delta\mu^{\ominus\prime} =$ -30.6	$-T\gamma^{\ominus}_S =$ 14.1
			$C_{cac} = 10^{-22.3}$

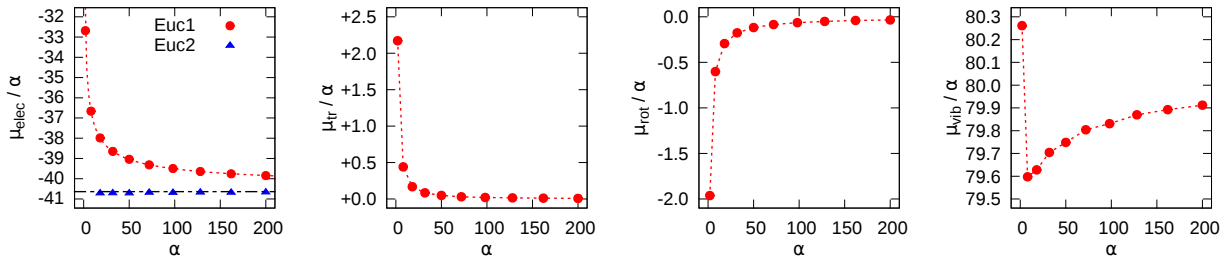
Trimesic Acid – Superflower Architecture

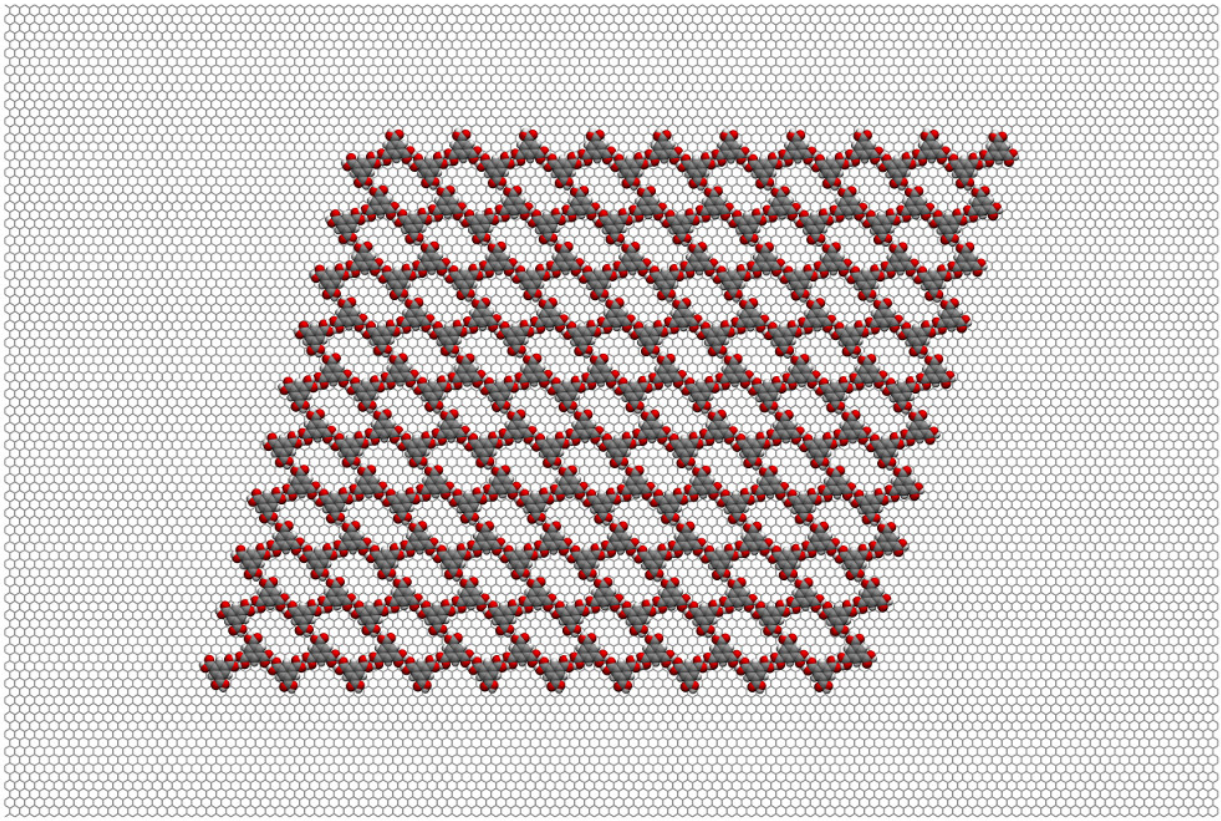




$\mu^{\circ}_{A,tr} =$ -9.2	$E'_{ads} =$ -25.8	$\mu'_{vib,uc} =$ 81.4	$A'_{uc} =$ 76.6
$\mu_{A,rot} =$ -8.2	$E'_{sam} =$ -25.7	$\mu'_{uc} =$ 31.2	$\gamma^{\circ} =$ -42.8
$\mu_{A,vib} =$ 81.4	$E'_{strain} =$ 1.3	$\Delta\mu'_{vib} =$ 0.0	$\gamma_E =$ -61.1
$\mu^{\circ}_{A,tot} =$ 64.0	$E'_{uc} =$ -50.2	$\Delta\mu^{\circ\prime} =$ -32.8	$-T\gamma^{\circ}_{S} =$ 18.3
			$C_{cac} = 10^{-23.9}$

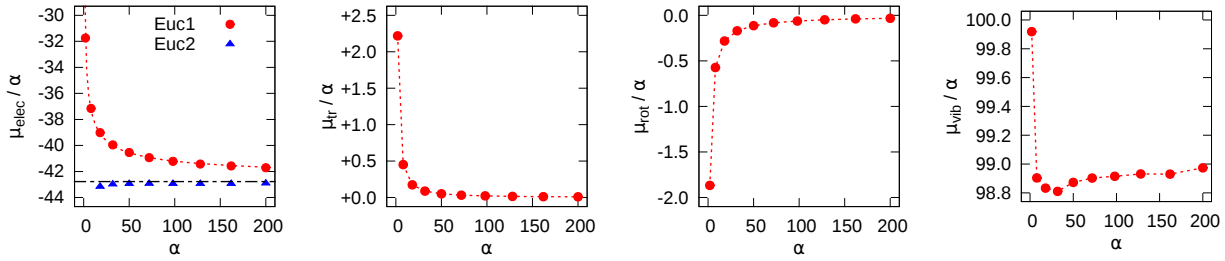
Trimesic Acid – Stripe (hypotetical) Architecture

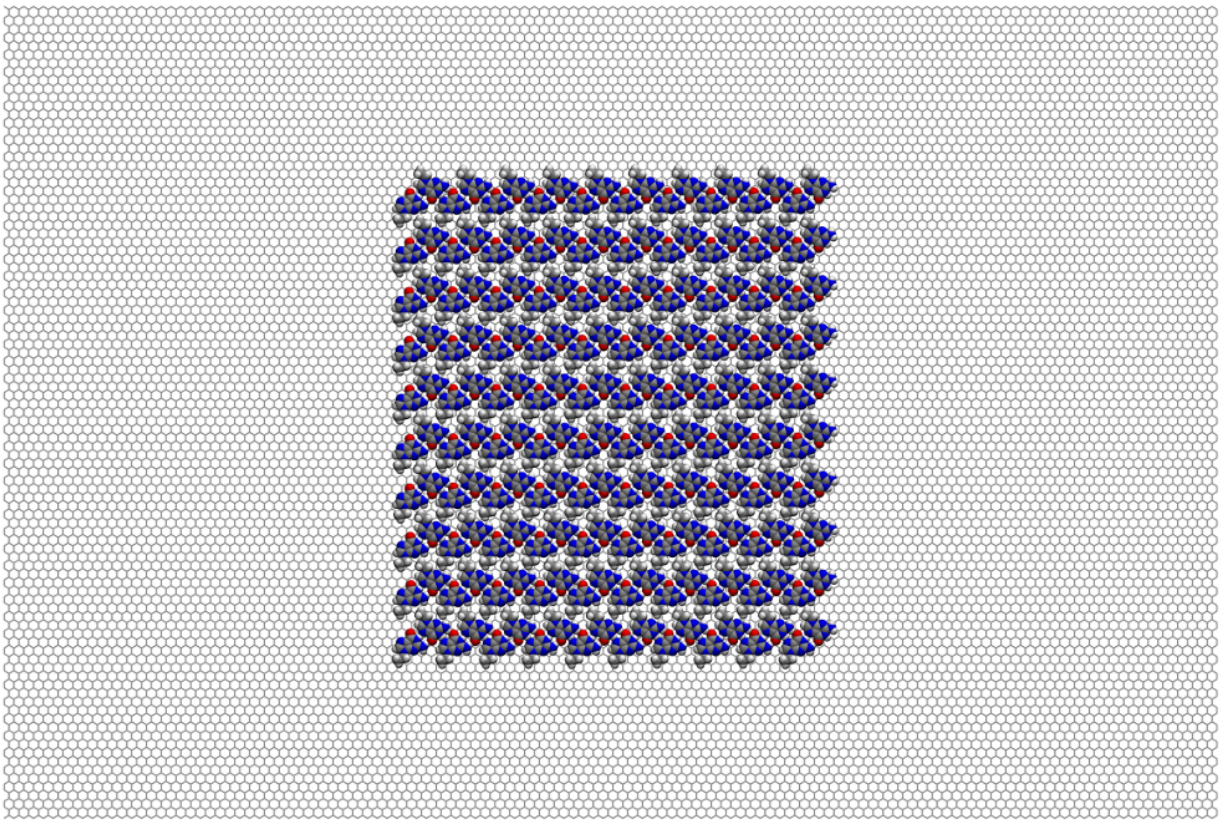




$\mu^{\ominus}_{A,tr} =$ -9.2	$E'_{ads} =$ -25.7	$\mu'_{vib,uc} =$ 79.9	$A'_{uc} =$ 117.0
$\mu_{A,rot} =$ -8.2	$E'_{sam} =$ -15.9	$\mu'_{uc} =$ 39.2	$\gamma^{\ominus} =$ -21.2
$\mu_{A,vib} =$ 81.4	$E'_{strain} =$ 0.9	$\Delta\mu'_{vib} =$ -1.5	$\gamma_E =$ -32.3
$\mu^{\ominus}_{A,tot} =$ 64.0	$E'_{uc} =$ -40.7	$\Delta\mu^{\ominus\prime} =$ -24.8	$-T\gamma^{\ominus}_{S} =$ 11.1
			$C_{cac} = 10^{-18.1}$

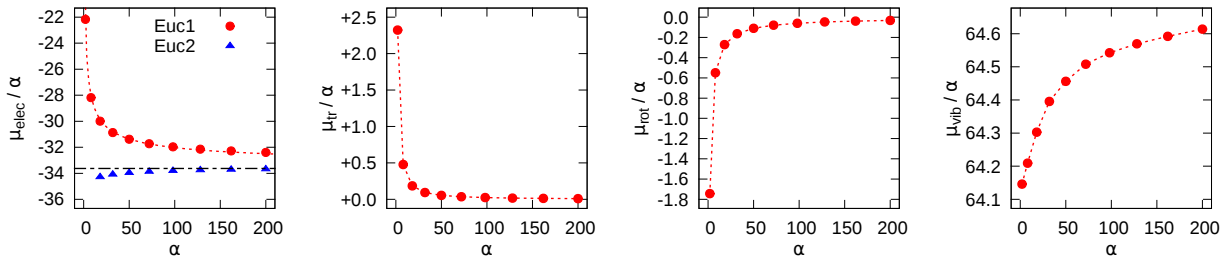
N^9-ethyl Guanine

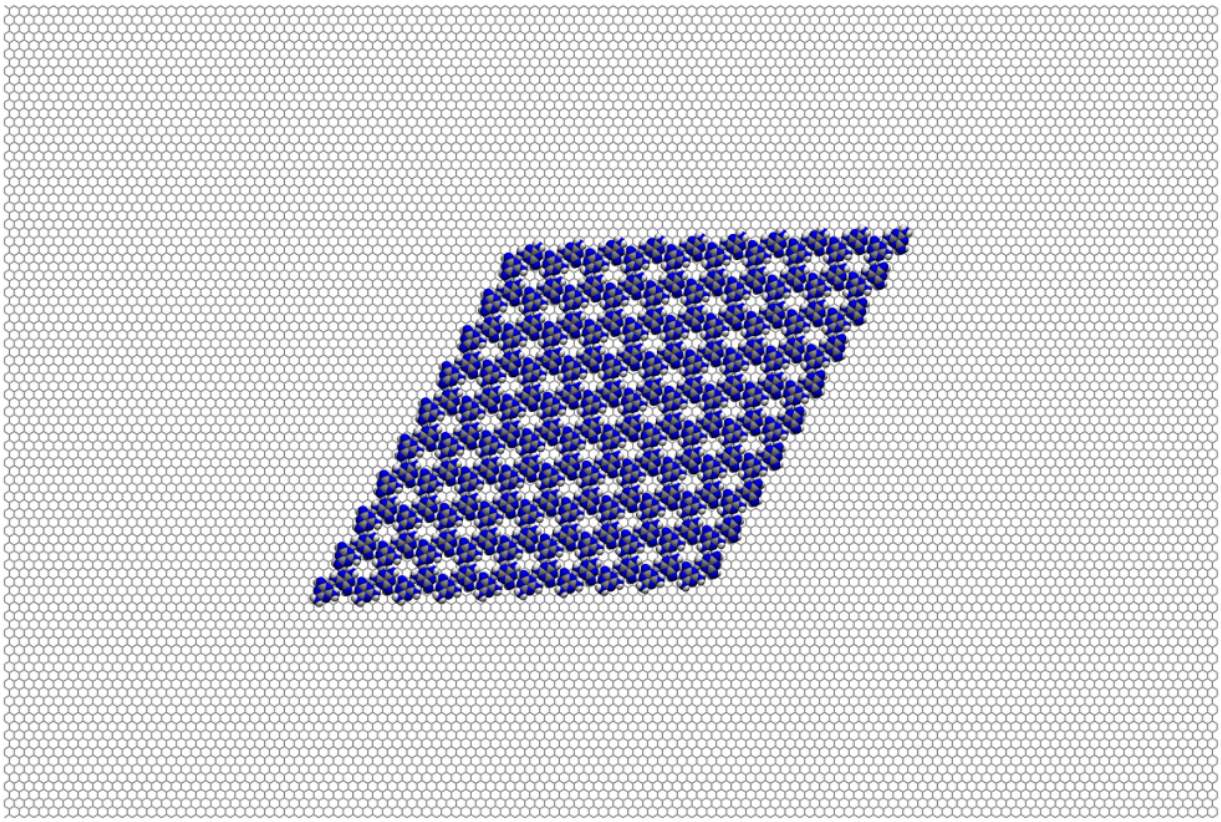




$\mu^{\circ}_{A,tr} =$ -9.0	$E'_{ads} =$ -23.5	$\mu'_{vib,uc} =$ 99.0	$A'_{uc} =$ 65.9
$\mu_{A,rot} =$ -8.4	$E'_{sam} =$ -20.9	$\mu'_{uc} =$ 56.1	$\gamma^{\circ} =$ -41.9
$\mu_{A,vib} =$ 101.1	$E'_{strain} =$ 1.5	$\Delta\mu'_{vib} =$ -2.1	$\gamma_E =$ -61.3
$\mu^{\circ}_{A,tot} =$ 83.7	$E'_{uc} =$ -42.9	$\Delta\mu^{\circ\prime} =$ -27.6	$-T\gamma^{\circ}_{S} =$ 19.4
			$C_{cac} = 10^{-20.1}$

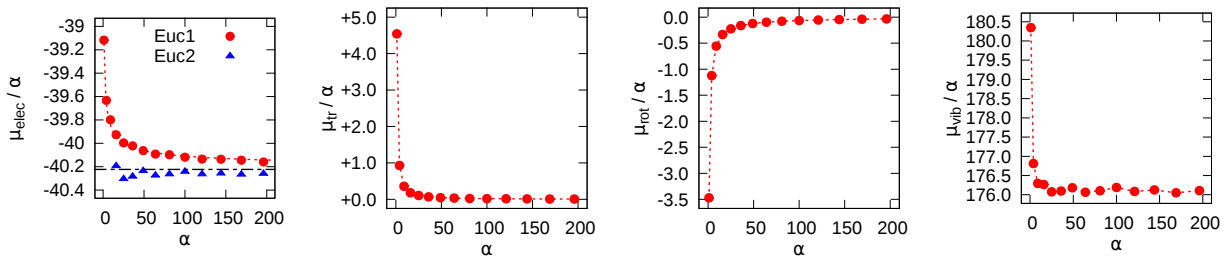
Melamine

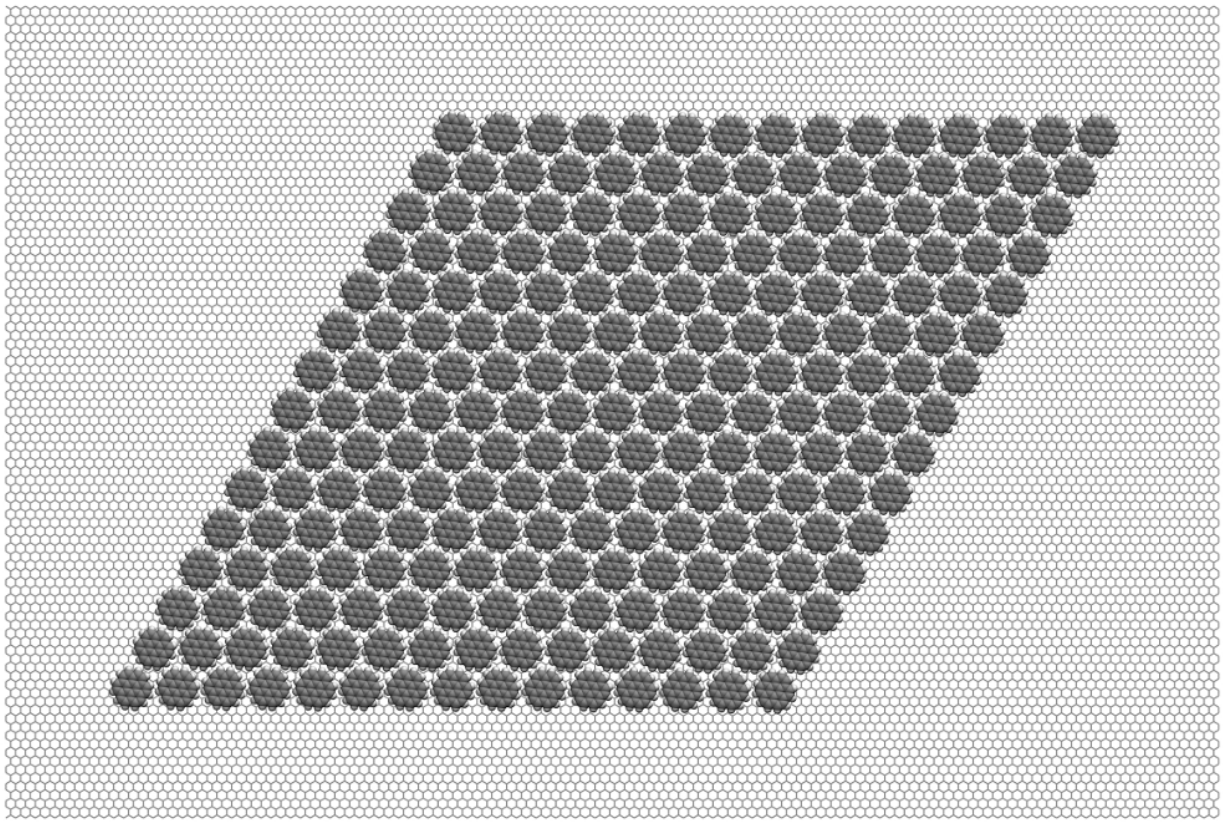




$\mu^{\circ}_{A,tr} =$ -8.7	$E'_{ads} =$ -16.3	$\mu'_{vib,uc} =$ 64.6	$A'_{uc} =$ 44.8
$\mu_{A,rot} =$ -7.7	$E'_{sam} =$ -19.8	$\mu'_{uc} =$ 30.8	$\gamma^{\circ} =$ -42.0
$\mu_{A,vib} =$ 66.0	$E'_{strain} =$ 2.3	$\Delta\mu'_{vib} =$ -1.4	$\gamma_E =$ -68.5
$\mu^{\circ}_{A,tot} =$ 49.6	$E'_{uc} =$ -33.8	$\Delta\mu^{\circ\prime} =$ -18.8	$-T\gamma^{\circ}_{S} =$ 26.5
			$C_{cac} = 10^{-13.7}$

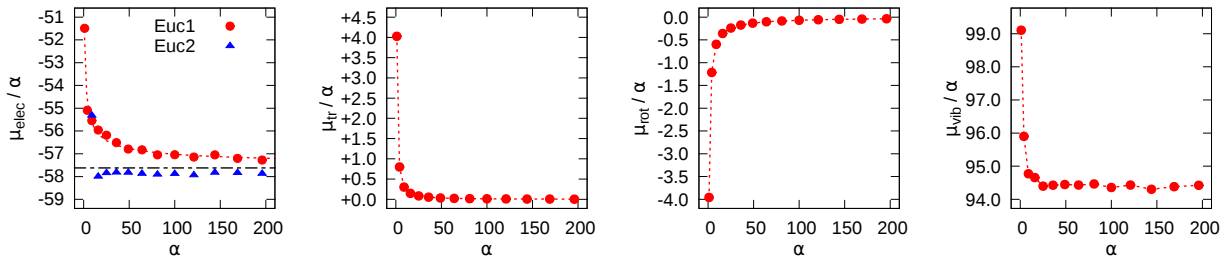
Coronene

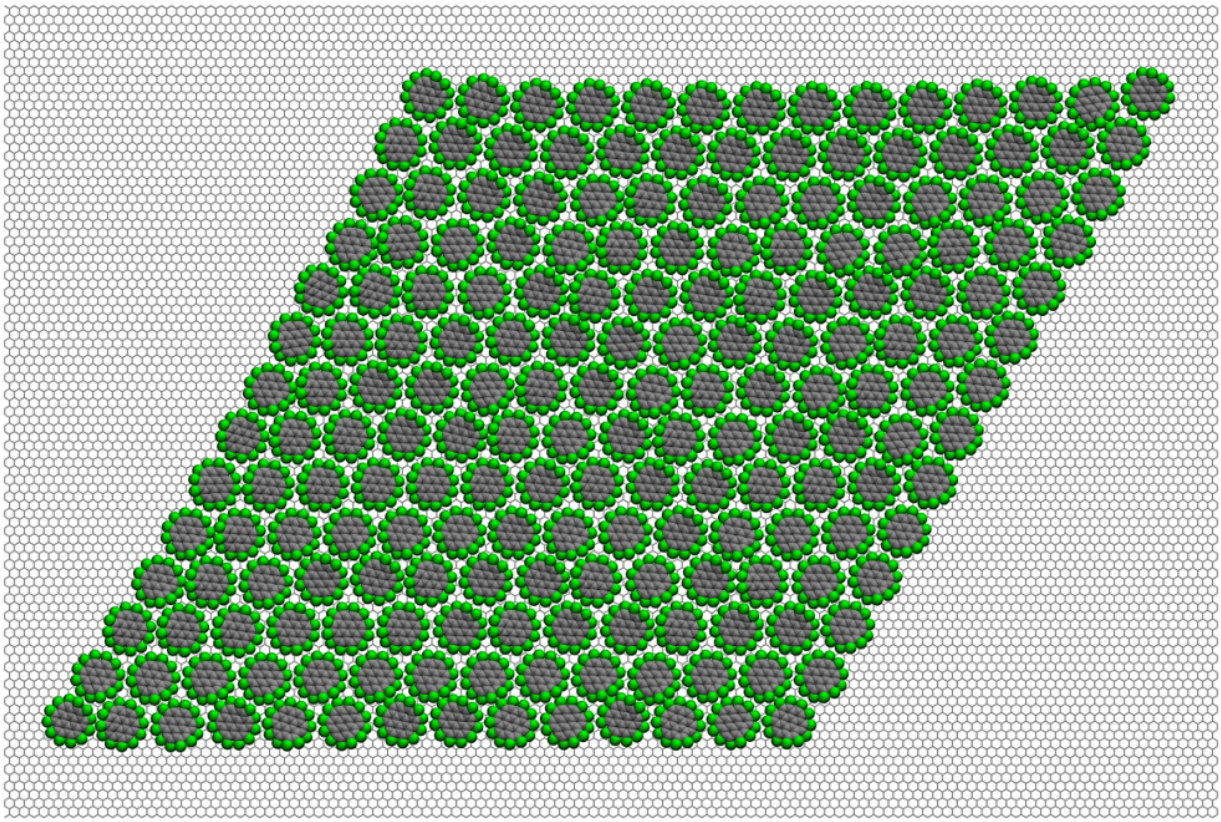




$\mu^{\circ}_{A,tr} =$ -9.5	$E'_{ads} =$ -38.9	$\mu'_{vib,uc} =$ 176.1	$A'_{uc} =$ 114.2
$\mu_{A,rot} =$ -7.8	$E'_{sam} =$ -1.4	$\mu'_{uc} =$ 135.8	$\gamma^{\circ} =$ -25.0
$\mu_{A,vib} =$ 181.7	$E'_{strain} =$ 0.0	$\Delta\mu'_{vib} =$ -5.6	$\gamma_E =$ -33.5
$\mu^{\circ}_{A,tot} =$ 164.4	$E'_{uc} =$ -40.3	$\Delta\mu^{\circ\prime} =$ -28.6	$-T\gamma^{\circ}_{S} =$ 8.5
			$C_{cac} = 10^{-20.8}$

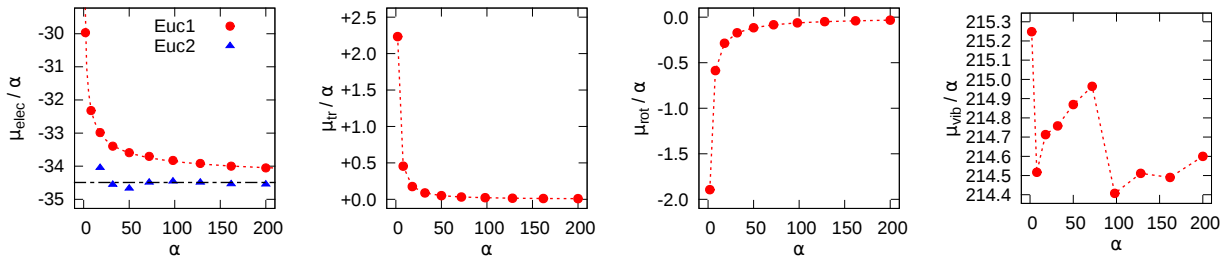
Perchlorocoronene

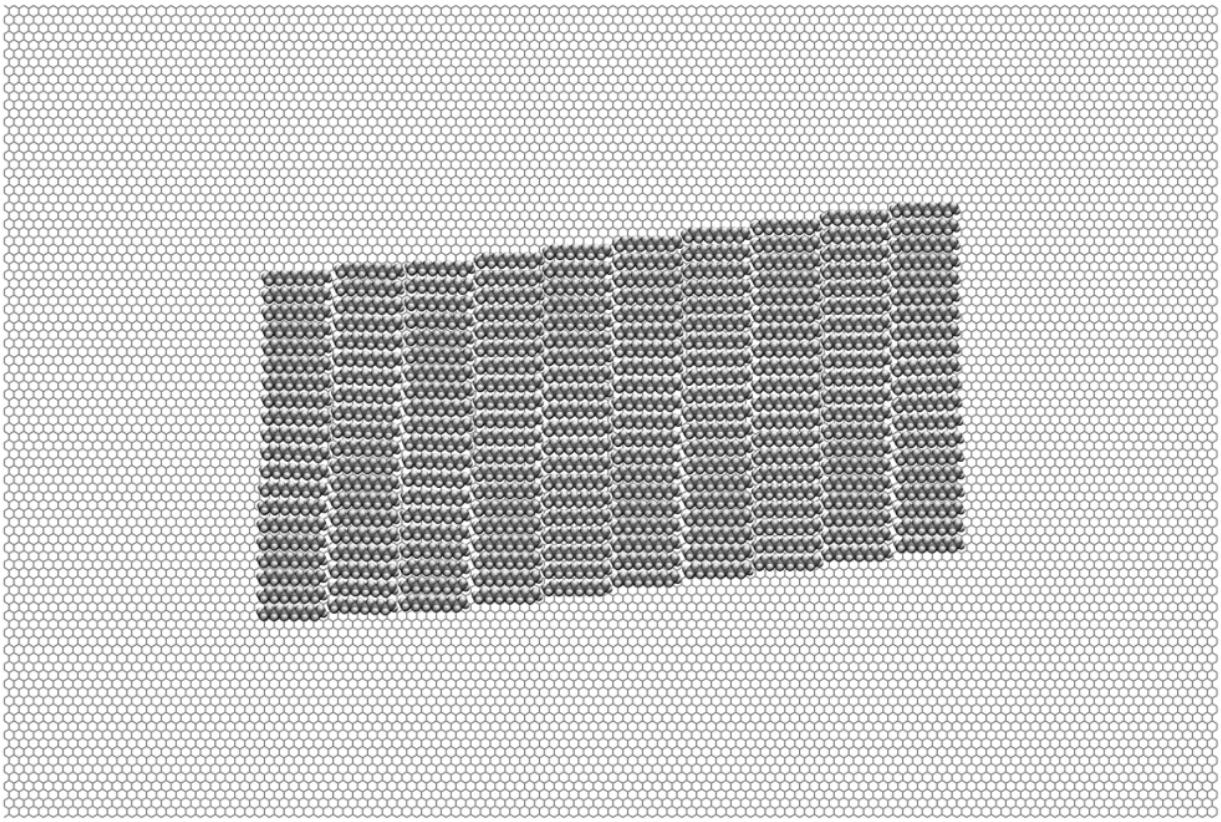




$\mu^{\ominus}_{A,tr}$ = -10.3	E'_{ads} = -58.2	$\mu'_{vib,uc}$ = 94.4	A'_{uc} = 163.9
$\mu_{A,rot}$ = -10.1	E'_{sam} = -6.6	μ'_{uc} = 36.7	$\gamma^{\ominus}$ = -25.8
$\mu_{A,vib}$ = 99.5	E'_{strain} = 7.1	$\Delta\mu'_{vib}$ = -5.1	γ_E = -33.7
$\mu^{\ominus}_{A,tot}$ = 79.1	E'_{uc} = -57.7	$\Delta\mu^{\ominus\prime}$ = -42.4	$-T\gamma^{\ominus}_S$ = 7.8
			$C_{cac} = 10^{-30.9}$

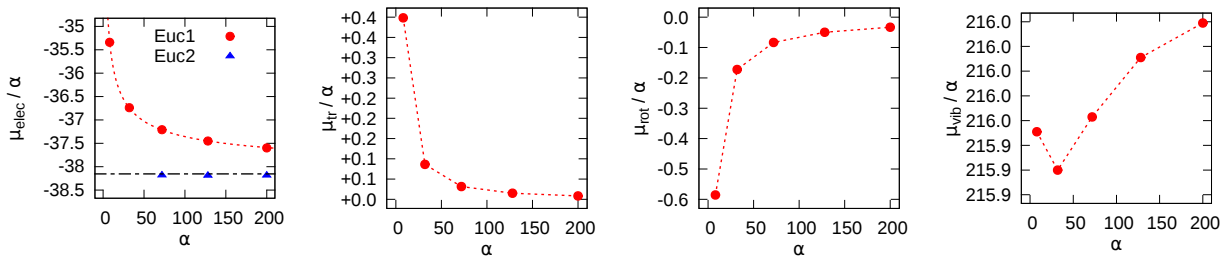
Dodecane

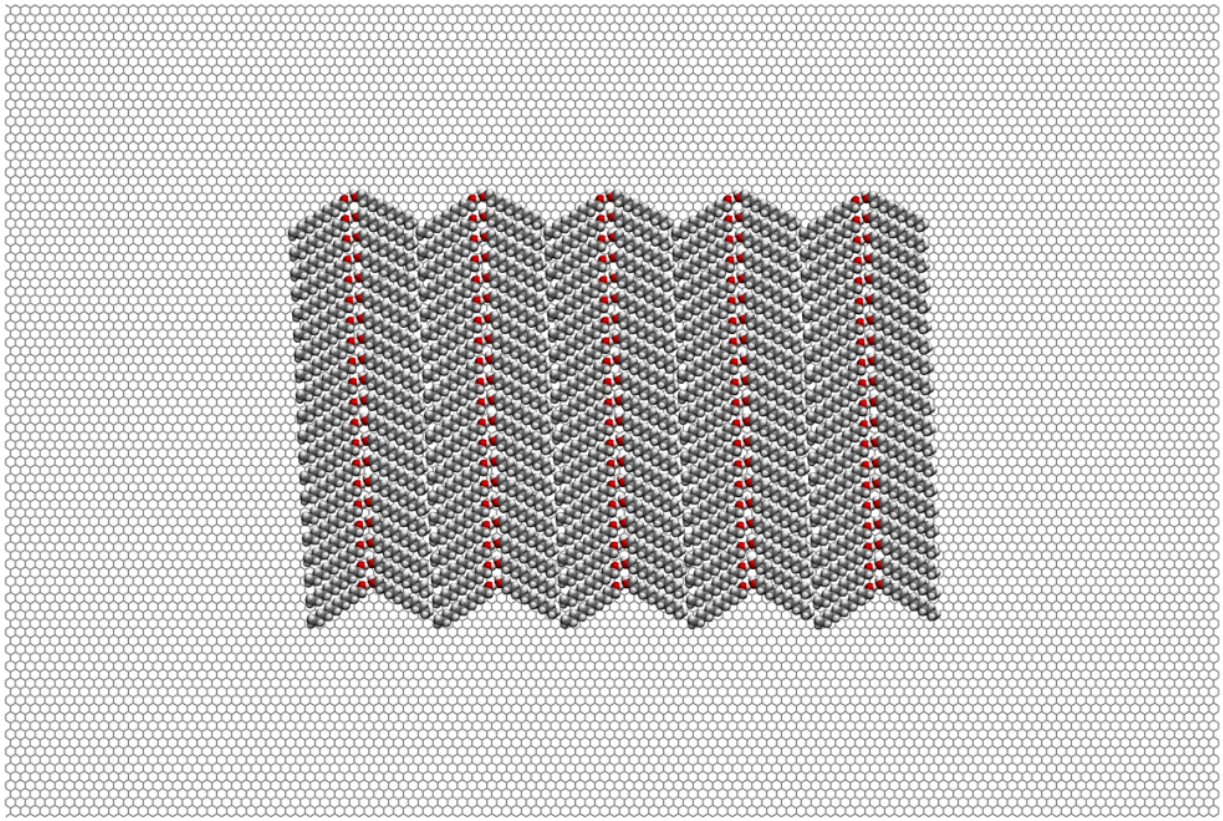




$\mu^{\ominus}_{A,tr} =$ -9.0	$E'_{ads} =$ -25.2	$\mu'_{vib,uc} =$ 214.6	$A'_{uc} =$ 77.5
$\mu_{A,rot} =$ -8.3	$E'_{sam} =$ -9.5	$\mu'_{uc} =$ 180.0	$\gamma^{\ominus} =$ -22.1
$\mu_{A,vib} =$ 214.4	$E'_{strain} =$ 0.1	$\Delta\mu'_{vib} =$ 0.2	$\gamma_E =$ -41.3
$\mu^{\ominus}_{A,tot} =$ 197.1	$E'_{uc} =$ -34.6	$\Delta\mu^{\ominus\prime} =$ -17.1	$-T\gamma^{\ominus}_S =$ 19.2
			$C_{cac} = 10^{-12.4}$

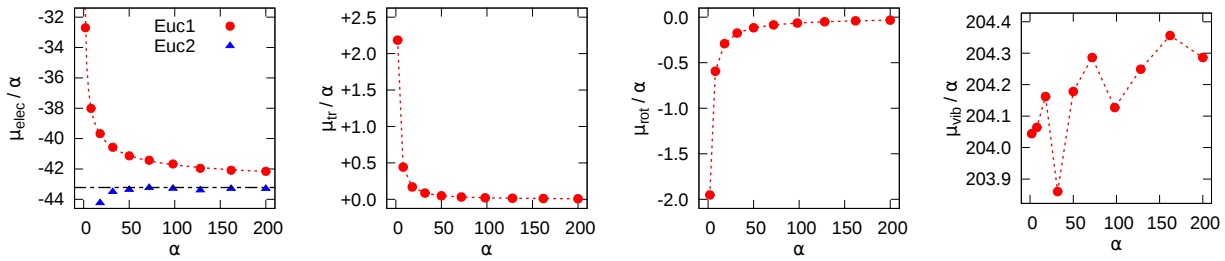
1-dodecanole

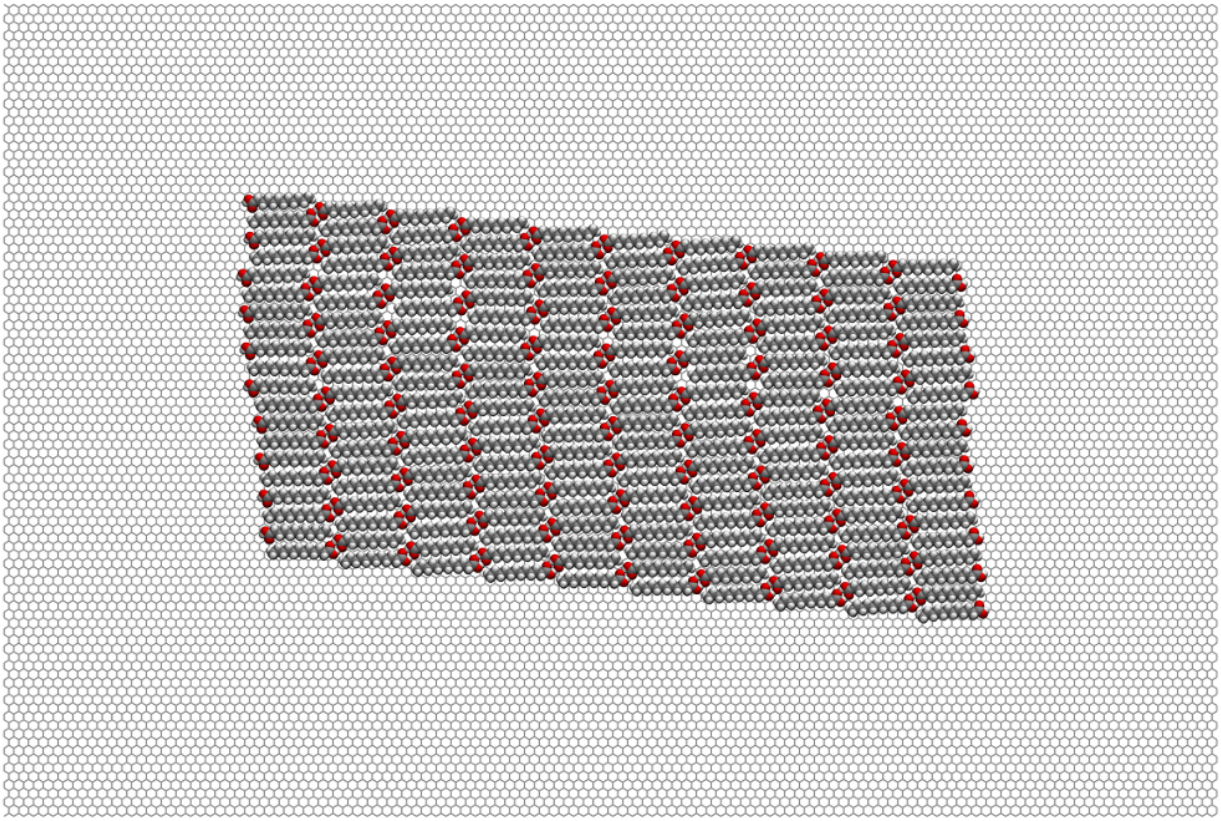




$\mu^{\ominus}_{A,tr} =$ -9.1	$E'_{ads} =$ -26.6	$\mu'_{vib,uc} =$ 216.0	$A'_{uc} =$ 81.2
$\mu_{A,rot} =$ -8.9	$E'_{sam} =$ -11.8	$\mu'_{uc} =$ 177.8	$\gamma^{\ominus} =$ -23.9
$\mu_{A,vib} =$ 215.2	$E'_{strain} =$ 0.2	$\Delta\mu'_{vib} =$ 0.8	$\gamma_E =$ -43.6
$\mu^{\ominus}_{A,tot} =$ 197.2	$E'_{uc} =$ -38.2	$\Delta\mu^{\ominus\prime} =$ -19.4	$-T\gamma^{\ominus}_S =$ 19.7
			$C_{cac} = 10^{-14.1}$

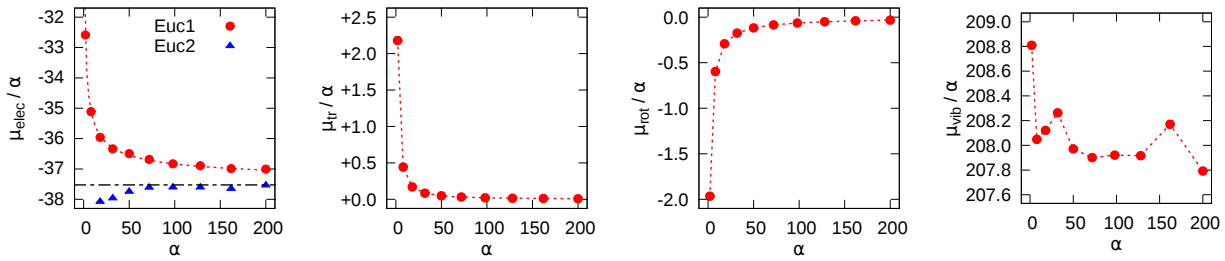
Dodecanoic Acid

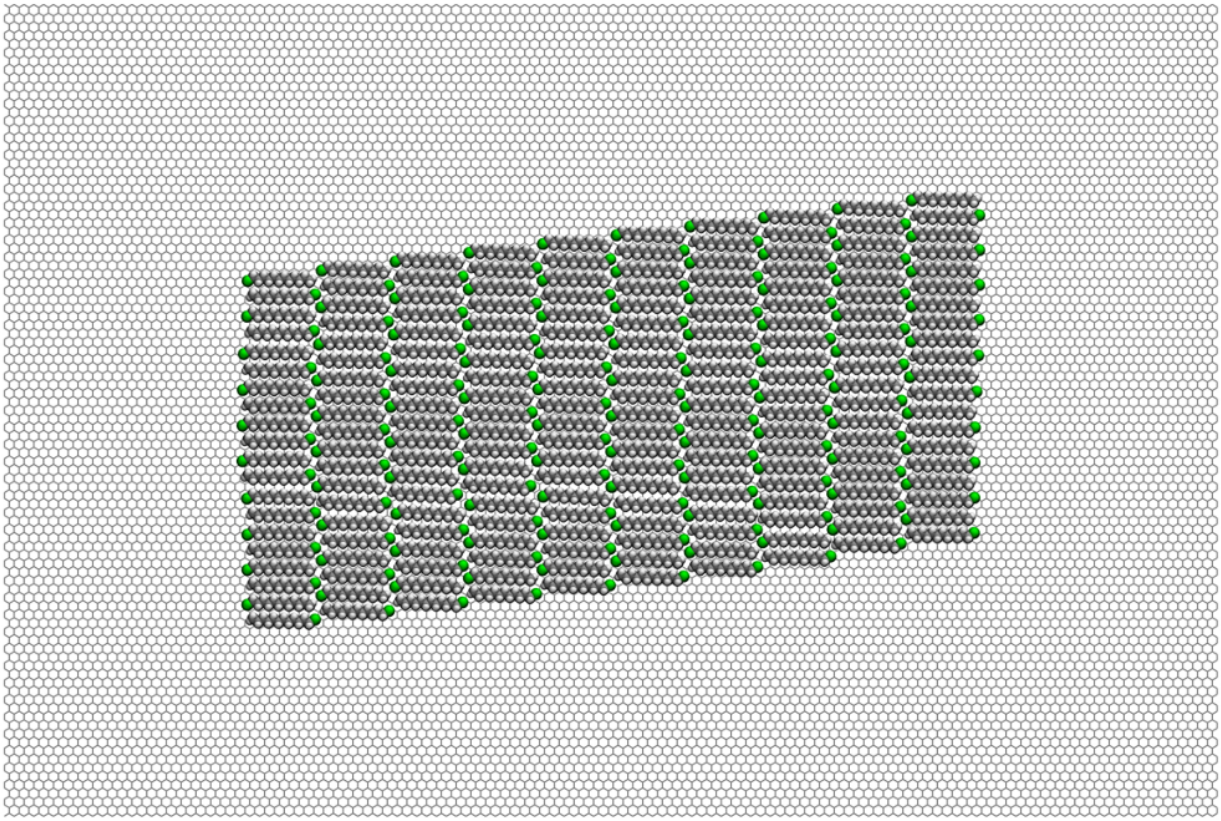




$\mu^{\ominus}_{A,tr} =$ -9.2	$E'_{ads} =$ -28.0	$\mu'_{vib,uc} =$ 204.3	$A'_{uc} =$ 81.8
$\mu_{A,rot} =$ -9.0	$E'_{sam} =$ -16.4	$\mu'_{uc} =$ 160.8	$\gamma^{\ominus} =$ -29.1
$\mu_{A,vib} =$ 202.8	$E'_{strain} =$ 0.9	$\Delta\mu'_{vib} =$ 1.5	$\gamma_E =$ -49.6
$\mu^{\ominus}_{A,tot} =$ 184.6	$E'_{uc} =$ -43.5	$\Delta\mu^{\ominus\prime} =$ -23.8	$-T\gamma^{\ominus}_S =$ 20.5
			$C_{cac} = 10^{-17.3}$

1-chlorododecane





$\mu^{\ominus}_{A,tr} =$ -9.2	$E'_{ads} =$ -27.7	$\mu'_{vib,uc} =$ 207.8	$A'_{uc} =$ 82.8
$\mu_{A,rot} =$ -9.0	$E'_{sam} =$ -10.1	$\mu'_{uc} =$ 170.1	$\gamma^{\ominus} =$ -23.4
$\mu_{A,vib} =$ 207.7	$E'_{strain} =$ 0.1	$\Delta\mu'_{vib} =$ 0.1	$\gamma_E =$ -42.4
$\mu^{\ominus}_{A,tot} =$ 189.5	$E'_{uc} =$ -37.7	$\Delta\mu^{\ominus\prime} =$ -19.4	$-T\gamma^{\ominus}_S =$ 19.0
			$C_{cac} = 10^{-14.1}$

* Electronic address: mcecchini@unistra.fr

[1] K. Reuter and M. Scheffler. "Composition, structure, and stability of RuO_2 (110) as a function of oxygen pressure". *Physical Review B*, 65(3), p. 035406, 2001.

[2] J. Kučera and A. Gross. "Adsorption of 4-mercaptopyridine on Au (111): a periodic DFT study". *Langmuir*, 24(24), pp. 13985–13992, 2008.

[3] C. Meier, M. Roos, D. Künzel, A. Breitruck, H. E. Hoster, K. Landfester, A. Gross, R. J. Behm, and U. Ziener. "Concentration and coverage dependent adlayer structures: from two-dimensional networks to rotation in a bearing". *The Journal of Physical Chemistry C*, 114(2), pp. 1268–1277, 2010.

[4] D. Loffreda, F. Delbecq, and P. Sautet. "Adsorption thermodynamics of acrolein on Pt(111) in realistic temperature and pressure from first-principle calculations". *Chemical Physics Letters*, 405(4), pp. 434–439, 2005.

[5] R. Gutzler, T. Sirtl, J. F. Dienstmaier, K. Mahata, W. M. Heckl, M. Schmittel, and M. Lackinger. "Reversible phase transitions in self-assembled monolayers at the liquid-solid interface: Temperature-controlled opening and closing of nanopores". *The Journal of the American Chemical Society*, 132(14), pp. 5084–5090, 2010.

[6] J. F. Dienstmaier, K. Mahata, H. Walch, W. M. Heckl, M. Schmittel, and M. Lackinger. "On the scalability of supramolecular networks: High packing density vs optimized hydrogen bonds in tricarboxylic acid monolayers". *Langmuir*, 26(13), pp. 10708–10716, 2010.

[7] M. Mammen, E. I. Shakhnovich, J. M. Deutch, and G. M. Whitesides. "Estimating the entropic cost of self-assembly of multiparticle hydrogen-bonded aggregates based on the cyanuric acid – melamine lattice". *The Journal of Organic Chemistry*, 63(12), pp. 3821–3830, 1998.

[8] S. Lei, K. Tahara, F. C. De Schryver, M. Van der Auweraer, Y. Tobe, and S. De Feyter. "One building block, two different supramolecular surface-confined patterns: Concentration in control at the solid–liquid interface". *Angewandte Chemie*, 120(16), pp. 3006–3010, 2008.

[9] A. Bellec, C. Arrigoni, G. Schull, L. Douillard, C. Fiorini-Debuisschert, F. Mathevet, D. Kreher, A.-J. Attias, and F. Charra. "Solution-growth kinetics and thermodynamics of nanoporous self-assembled molecular monolayers". *The Journal of Chemical Physics*, 134(12), p. 124702, 2011.

[10] M. O. Blunt, J. Adisoejoso, K. Tahara, K. Katayama, M. Van der Auweraer, Y. Tobe, and S. De Feyter. "Temperature-induced structural phase transitions in a two-dimensional self-assembled network". *The Journal of the American Chemical Society*, 135(32), pp. 12068–12075, 2013.

[11] S. Bonacchi, M. El Garah, A. Ciesielski, M. Herder, S. Conti, M. Cecchini, S. Hecht, and P. Samorì. "Surface-induced selection during in situ photoswitching at the solid/liquid interface". *Angewandte Chemie International Edition*, 54, pp. 4865–4869, 2015.

[12] S. Conti, M. G. del Rosso, A. Ciesielski, J. Weippert, A. Böttcher, Y. Shin, G. Melinte, O. Ersen, C. Casiraghi, X. Feng, K. Müllen, M. M. Kappes, P. Samorì, and M. Cecchini. "Perchlorination of coronene enhances its propensity to self-assembly on graphene". *ChemPhysChem*, 17, pp. 352–357, 2016.

[13] B. Roux and T. Simonson. "Implicit solvent models". *Biophysical Chemistry*, 78(1), pp. 1–20, 1999.

[14] J. R. Reimers, D. Panduwinata, J. Visser, Y. Chin, C. Tang, L. Goerigk, M. J. Ford, M. Sintic, T. Sum, M. J. J. Coenen, B. L. M. Hendriksen, J. A. A. W. Elemans, N. S. Hush, and M. J. Crossley. "A priori calculations of the free energy of formation from solution of polymorphic self-assembled monolayers". *Proceedings of the National Academy of Sciences*, 112(45), pp. E6101–E6110, 2015.

[15] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. "Electrostatics of nanosystems: application to microtubules and the ribosome". *Proceedings of the National Academy of Sciences*, 98(18), pp. 10037–10041, 2001.

[16] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. "Development and testing of a general amber force field". *Journal of Computational Chemistry*, 25(9), pp. 1157–1174, 2004.

[17] J. Wang, W. Wang, P. A. Kollman, and D. A. Case. "Automatic atom type and bond type perception in molecular mechanical calculations". *Journal of Molecular Graphics and Modelling*, 25(2), pp. 247–260, 2006.

[18] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl. "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". *SoftwareX*, 1–2, pp. 19–25, 2015.

[19] P. V. Klimovich, M. R. Shirts, and D. L. Mobley. "Guidelines for the analysis of free energy calculations". *Journal of Computer-Aided Molecular Design*, 29(5), pp. 397–411, 2015.

[20] M. Lackinger, S. Griessl, W. M. Heckl, M. Hietschold, and G. W. Flynn. "Self-assembly of trimesic acid at the liquid-solid interface a study of solvent-induced polymorphism". *Langmuir*, 21(11), pp. 4984–4988, 2005.

[21] G. Gilli and P. Gilli. "Towards an unified hydrogen-bond theory". *Journal of Molecular Structure*, 552, pp. 1–15, 2000.

[22] J. Řezáč, K. E. Riley, and P. Hobza. "Extensions of the S66 data set: More accurate interaction energies and angular-displaced nonequilibrium geometries". *Journal of Chemical Theory and Computation*, 7, pp. 3466–3470, 2011.