

Efficient hierarchical models for reactivity of organic layers on semiconductor surfaces

Jan-Niclas Luy^{1,2}, Mahlet Molla¹, Lisa Pecher¹, and Ralf Tonner^{1,2*}

¹Jan-Niclas Luy, Mahlet Molla, Dr. Lisa Pecher, Prof. Dr. Ralf Tonner

Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Straße 4, D-35032 Marburg, Germany

²Current address: Wilhelm-Ostwald-Institut für Physikalische und Theoretische Chemie, Universität Leipzig, Linnéstraße 2, D-04103 Leipzig, Germany

e-Mail: ralf.tonner@chemie.uni-leipzig.de

Computational modeling of organic interface formation on semiconductors poses a challenge to a density functional theory-based description due to structural and chemical complexity. A hierarchical approach is presented, where parts of the interface are successively removed in order to increase computational efficiency while maintaining the necessary accuracy. First, a benchmark is performed to probe the validity of this approach for three model reactions and five dispersion corrected density functionals. Reaction energies are generally well reproduced by GGA-type functionals but accurate reaction barriers require the use of hybrid functionals. Best performance is found for the model system that does not explicitly consider the substrate but includes its templating effects. Finally, this efficient model is used to provide coverage dependent adsorption energies and suggest synthetic principles for the prevention of unwanted growth termination reactions for organic layers on semiconductor surfaces.

Introduction

The creation of hybrid organic-inorganic interfaces is an important aspect of many technological applications. Some prominent examples include the combination of established silicon-based technology with the emerging approach of molecular electronics [1-4], encapsulation of organic solar cells or organic light emitting diodes (OLED) [5, 6] and (biomedical) sensing [7]. The goal of advancing miniaturization leads to a relative increase in size of the surface-area with respect to the bulk. This increases the importance of surface properties. To analyze and understand these properties, interfaces have to be uniform and defect free. Such quality requirements can only be fulfilled with molecular-scale control during synthesis while remaining compatible with established clean-room manufacturing techniques. One promising approach is the build-up of layered

organic films on a solid inorganic semiconductor surface using well-tailored molecular building blocks [8]. Here, the most critical step is the attachment of the first layer since pristine semiconductor surfaces are generally highly reactive [9] resulting in low selectivity. In order to increase selectivity, attachment should be covalent which has the added benefit of being self-limiting to one layer once the substrate is fully covered. Subsequent deposition of a different building block on top of the first layer then offers fine control over thickness and composition of the organic material. This approach shares some ideas with molecular layer deposition (MLD) [10, 11] and layer-by-layer (LbL) [12] methods for polymer films [13]. However, the focus of this work is not on growth of bulk materials but the reactivity of the crucial interface region shown in Figure 1.

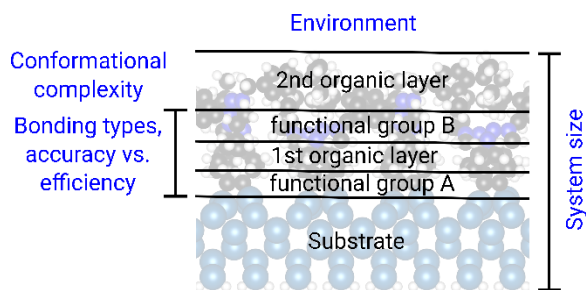


Figure 1. The interface model investigated in this work is comprised of a solid substrate in slab geometry and up to two layers of an organic material in contact to vacuum or a solvent. The organic layers are covalently linked to the substrate and each other by functional groups A and B. The computational challenges arising from this model are printed in blue.

Commonly, interface formation on solid semiconducting substrates is studied in ultra-high vacuum (UHV) [14]. However, synthetic routes towards organic layers are often first investigated and optimized in solution [15]. Computations can thus help to assess the question of transferability of wet-chemical schemes to UHV before performing time-consuming experiments. This can be done with modelling techniques that go beyond supplementing experimental research towards suggesting new avenues of investigation. Furthermore, interface characterization based on electronic structure methods can lead to a deeper understanding of the physicochemical origins of various material properties.

We note in passing that additional synthetic challenges such as double adsorption of bifunctional molecules are still not fully overcome. These issues are also under investigation but not further discussed in this study [39, 40]. However, based on the IEDDA reaction, attachment of molecules in the second organic layer could recently be demonstrated [41].

With this work we are aiming at deriving an accurate yet efficient computational model based on density functional theory (DFT) for studying hybrid interfaces and their growth processes. The proposed model is then validated with three experimentally significant reaction schemes.

Computational challenges

For the modelling of hybrid interfaces, some general computational challenges exist, which are sketched in the following (see Figure 1).

System size. The computational effort for solving the electronic Schrödinger equation scales approximately with the third power with regard to system size (i.e. number of basis functions) using generalized gradient approximation (GGA)-based density functionals. It is therefore evident that there exists a limit for the size of systems that can be solved on current computer hardware in acceptable time frames.

Bonding types. Different bonding types can be used to build up an organic-inorganic interface. In addition to the covalent connection in the first organic layer mentioned above, further layers and cross-links between molecular strands could be formed by hydrogen bonds or non-covalent interactions (e.g. halogen bonds) [16]. Moreover, dispersion attraction is a significant driving force behind the self-assembly into extended structures. Particularly, the latter is insufficiently captured by many density functionals, requiring dispersion correction schemes. In order to correctly model a hybrid interface, the different bonding mechanisms, including those in the substrate, must be described on an equal footing.

Accuracy versus efficiency. Closely linked to the first challenge is the task of finding a tradeoff between accuracy and efficiency for the determination of the total energy. Density functional approximations have partially been optimized to produce highly accurate results for specific research questions or systems (e.g., TB09 for optical properties of elemental semiconductors). Thus, general purpose functionals are a more appropriate choice for the complex electronic structure found in hybrid interfaces. More accurate methods such as hybrid functionals and wave-function based approaches have less favorable scaling with system size and are therefore often limited to benchmark studies or small systems.

Conformational sampling. The amorphous nature of organic layers leads to an exponential increase of the conformational degrees of

freedom with respect to the number of building blocks. Furthermore, torsional angle motions are usually soft, resulting in shallow potential energy surfaces (PES) and slow convergence of the structural optimization algorithms. Thus, ample conformational sampling has to be performed before predicting interface properties. In this study, simulated annealing [17] is used to improve on “naively” guessed minima. This is achieved by sampling the PES at elevated temperature using molecular dynamics approaches, thus allowing the system to escape from the local minimum and relax into a more favorable structure during the subsequent cooling process. However, solving the electronic structure problem self-consistently for every atomic displacement using DFT quickly becomes a computational bottleneck. Considerable speed-up can be achieved by resorting to more approximate electronic structure methods for structural optimization. The challenge then becomes finding a substitute method that well reproduces the ab initio results. We resort here to the semi-empirical density functional tight binding (DFTB) [18] method.

Chemical environment. A difference between conventional thin-film synthesis and deposition from gas phase is the solvent environment. Generally, two families of approaches are considered for treating solvent interactions. The explicit approach simply adds solvent molecules to the system while the implicit one uses a continuum plus cavity describing mostly the electrostatic solute-solvent interaction. While the explicit approach is more accurate and realistic, computational demand grows exponentially with the number of solvent molecules. Since our aim is to model large interface systems, computational demand is a concern. Therefore, implicit solvent corrections are the first choice for large interface models [19]. Their performance has been well tested for molecular solutes [20] and encouraging results have been obtained for liquid/solid interfaces [21]. In a later part of this study, we probe the influence of different solvents on calculated reaction energies in the framework of the polarizable continuum model (PCM) and discuss

trends with increasing interface complexity. Other environmental effects such as finite temperatures and pressures are outside the scope of the present study.

The model reactions

We validate our computational protocol for organic layers by selecting a suitable test set. Common synthetic approaches for organic layers are derived from the concept of “click-chemistry” [22]. Those were initially developed for biochemical applications but are also increasingly employed in material science [23-25]. Click reactions must be fast (i.e. low barriers) and produce excellent yields with high selectivity (i.e. highly negative reaction energies). In addition, reactions should proceed without a catalyst, produce as few side products as possible and be largely independent of solvent environments [10]. Application of click-chemistry in interface synthesis under UHV conditions constrains possible precursors further to be simultaneously volatile and thermally stable. The three reaction schemes chosen for investigation in this article are introduced briefly in the following (Figure 2).

The first reaction is a variant of the azide-alkyne 1,3-dipolar cycloaddition (AAC) [26]. This reaction can be considered a prototypical click reaction and has shown excellent performance when combined with a Cu(I) catalyst in organic solvents [27]. However, the catalyst cannot be applied in UHV conditions and is thus not considered here. The second model reaction shown in Figure 2 is an acyl-chloride mediated esterification (ACE) [28]. The reaction provides a low kinetic barrier even in the gas phase and could therefore be an alternative to the AAC when concerns about precursor stability limit the use of elevated temperatures. As a side-note, the by-product HCl could initiate undesired side-reactions with the substrate or other precursor molecules – a problem we are not further investigating here. Finally, an inverse electron demand Diels-Alder (IEDDA) reaction utilizing a tetrazine derivative as the diene [15] has been recently proposed. As part of our test

set we limit our investigation to the rate-determining first synthetic step [29]. An advantage of this reaction is the tolerance of the strain promoted AAC (SPAAC) [30] variant which allows for sequential execution in an interface building approach.

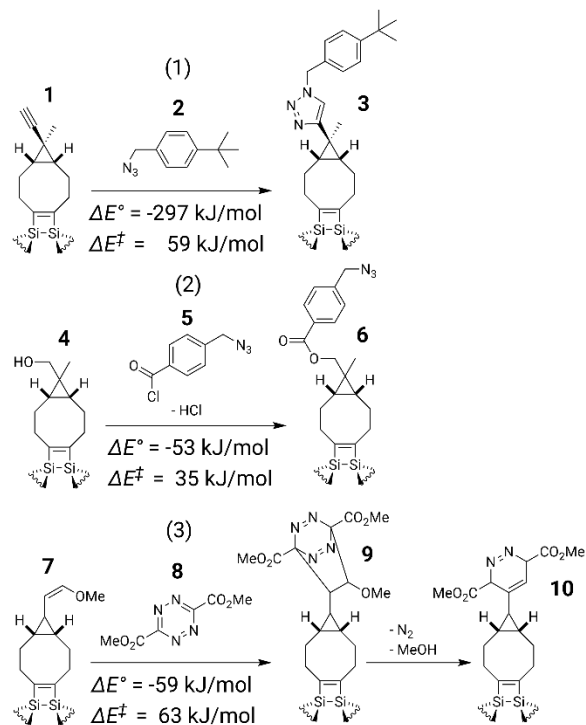


Figure 2. Selection of interface building schemes chosen as model systems in this work. (1) Cu-free azide-alkyne cycloaddition (AAC). The mono-functional variant of alkyne **2** is investigated here. The bi-functional form is discussed at the end of this article. (2) Acyl chloride mediated esterification (ACE). The azide moiety indicates the possibility of attaching a third layer via AAC. (3) Inverse electron demand Diels-Alder (IEDDA) reaction. We focus on the first reaction step leading to **9** here.

Defining the system

The substrate

As a model for interface formation between a semiconductor surface and a layered organic material (Figure 1), silicon and its (001) facet are an obvious choice due to their technological prevalence and versatility in forming covalent bonds [31]. Even though novel materials such as III/V compound semiconductors or transition metal dichalcogenides are increasingly

incorporated in devices, the “post-silicon” age is only just appearing at the horizon [4]. Si-based technology is very mature and will, for the foreseeable future, remain an important part of nanoelectronic devices. However, the methods used in this study are built upon first principles and insights are therefore transferable to other substrates.

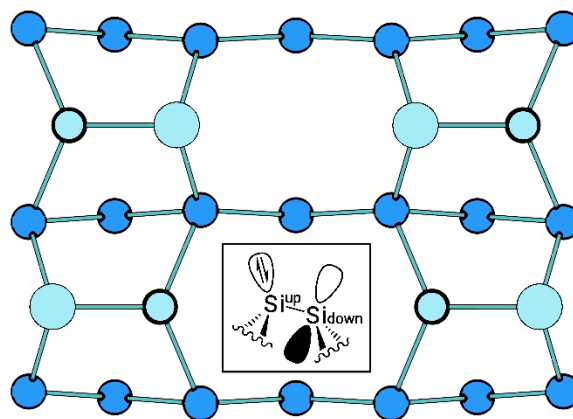


Figure 3. Top view of the $c(4 \times 2)$ reconstruction of the Si(001) surface. Dark blue: sub-surface atoms. Light blue: buckled surface dimers with Si_{down} (small circle) and Si_{up} atoms (large circle). Inset: side view of a buckled dimer with simplified frontier orbital sketch.

A model of the Si(001) surface is given in Figure 3. Upon cleavage from the single crystal, each surface atom has two unpaired electrons (dangling bonds). The surface energy is lowered by those atoms reconstructing into rows of tilted dimers which share a single bond. Furthermore, the Si_{down} atom transfers its remaining electron to Si_{up} resulting in an unoccupied p -orbital and a non-bonding electron pair respectively (Figure 3 inset). In order to reduce repulsion between lone pairs, dimers tilt alternately along a row. Between rows, trenches are formed across which the tilt also alternates. The resulting $c(4 \times 2)$ reconstruction is the most stable under UHV conditions at low temperatures [31].

The organic adsorbate

Cyclooctyne solves many questions of low selectivity at the Si(001) surface. Due to the release of about 80 kJ/mol of ring strain energy (Figure S1) the triple bond reacts barrierless with

a single surface dimer via a 2+2 cycloaddition [32]. This markedly improves selectivity with respect to other functional groups [33]. Thus, cyclooctyne is an ideal platform molecule for interface build-up [34]. After some optimization of the molecular design, bicyclic derivatives of cyclooctyne (Figure 4) have emerged as preferred building block candidates due to increased rigidity, restriction of conformational space and limiting of side reactions [34-36].

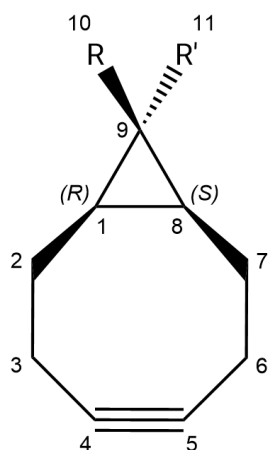


Figure 4. [(1R,8S)-bicyclo[6.1.0]non-4-yne] (CCO, $R = R' = H$) platform molecule with chiral centers at carbons one and eight. **1:** $R = -C\equiv CH$, $R' = -CH_3$ **4:** $R = -CH_2-OH$, $R' = -CH_3$ **7:** $R = -CH=CH-OMe$, $R' = H$.

All organic layers investigated in this work and shown in Figure 2, are thus based on [(1R,8S)-bicyclo[6.1.0]non-4-yne] (CCO) [27]. Apart from the strained triple bond, the molecule contains two additional substituents R (functional group) and R' (aliphatic group or H). As a consequence of its specific molecular structure, group R is forced to point in the direction of growth where the next layer can be attached.

The first layer

In previous work, parent cyclooctyne was shown to block two surface dimers due to tilting of the backbone [36]. The first adsorbed molecule steers the next one through dispersion attraction towards adjacent sites [36]. Hence, an ordered first layer is formed on Si(001) [14] with a maximum coverage of $\Theta = 0.5$ monolayers (ML)

CCO with a complete ML being defined as one adsorbate per surface dimer.

Here, we model the experimental system by considering an idealized CCO/Si(001) interface that is uniform and defect free. During the creation of our interface model we took both the steric demand and the inherent symmetry of the molecule into consideration. Since CCO has two chiral centers, several possibilities exist for the adsorption of more than adsorbate. Two neighboring molecules can either face the same direction or be positioned 'back-to-back'. Assuming a random distribution of N molecules, 2^N possible arrangements emerge. In order to calculate the relative energies of different arrangements, a 4×4 supercell containing eight dimers and four CCO derivatives **1** ($R = -C\equiv CH$ and $R' = -CH_3$) was chosen (see Figure 5).

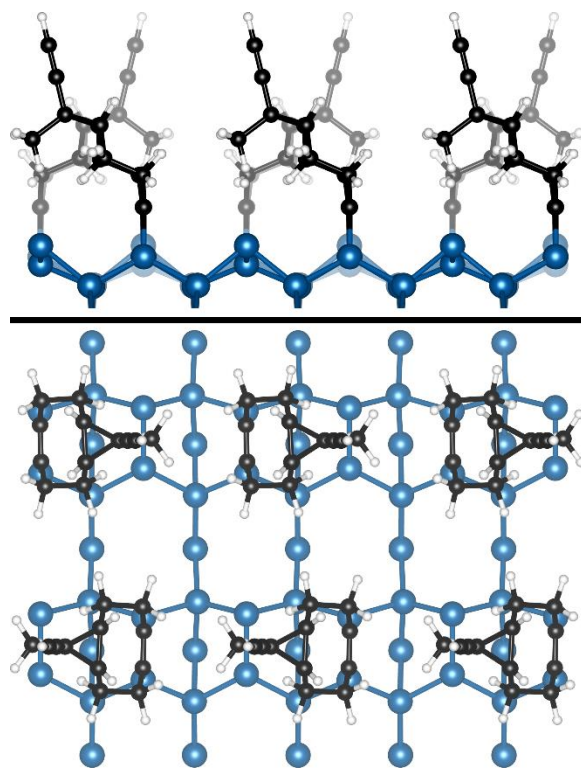


Figure 5. Structure of the most favorable arrangement of CCO **1** on Si(001)c(4x2) at full coverage. Molecules within a row point in the same direction while molecules in adjacent rows point in the opposing direction. Alternating rows are offset by one dimer (see Figure S2 for discussion of alternative configurations).

Consequently, there exist $2^4=16$ possible configurations of four CCO molecules in the 4×4

supercell. This number is doubled when a shift by one dimer is applied to every second row. Through the exploitation of symmetry operations, the number of unique structures is reduced to nine (see Figure S3). In this setup, the computational demand is still feasible while all possible nearest neighbor interactions between adsorbate molecules are sampled.

Relative energies of these nine structures are listed in Table S1. Evidently, even the least favorable configuration is only 7 kJ/mol less stable than the most favorable one depicted in Figure 5. Energy differences of this magnitude are smaller than the overall accuracy of DFT [37] and all configurations can therefore be considered isoenergetic. Thus, a random arrangement of molecules on the macroscopic scale appears likely.

The computational models – a hierarchical approach

The central question of this study is how to create an accurate structural model of hybrid organic-inorganic interfaces that can be efficiently computed. As mentioned in the section on computational challenges, accurate reaction barriers and energies based on highly accurate approaches quickly become unattainable for such systems. Thus it is imperative to develop means to minimize computational complexity while maintaining acceptable accuracy for properties of interest. We propose to reduce the complexity of an interface model based on the assumption that bond cleavage and formation are predominantly local phenomena. Thus, it stands to reason that only the immediate chemical environment exerts a noticeable influence on the shape of a PES. In the following, we test this hypothesis by comparing the energetic signatures of the reactions shown in Figure 2 with those obtained from structurally increasingly simplified model systems. For this purpose, a hierarchical system (Figure 6) is introduced and benchmarked with respect to the full interface model which is assigned rank 1.

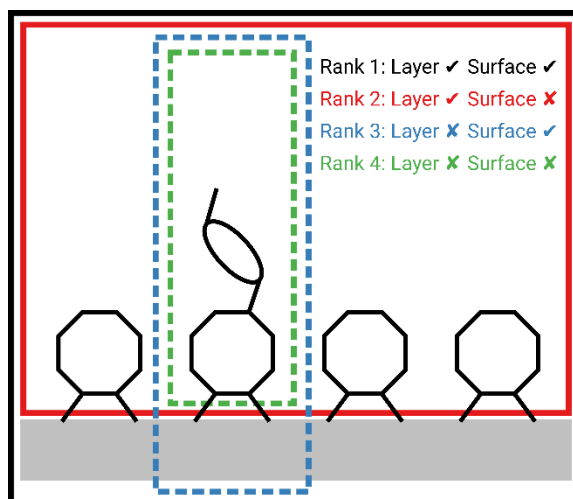


Figure 6. Hierarchically designed model systems of an interface between a semiconductor surface and an organic layer. Rank 1 (black): Full system containing the substrate and an organic layer. Rank 2 (red): The substrate is removed while the organic layer is structurally constrained to mimic the substrate template. Rank 3 (blue): Substrate same as rank 1 with a single adsorbed molecule. Rank 4 (green): A single molecule confined to surface-accessible geometry.

Rank 1. The full system contains the surface slab, which is constructed as described in the methods section, and the desired number of organic layers at complete coverage. No structural approximations besides the supercell size and slab thickness are made.

Rank 2. At the first rank of simplification, the substrate is removed. Whenever a covalent bond must be cut during system size reduction, dangling bonds are saturated with hydrogen atoms. Thus, rank 2 models explicitly capture intra-layer effects in the organic layer but disregard influence of the interface. However, template effects of the inorganic substrate are included by freezing positions of some atoms in the first organic layer. One advantage of rank 2 systems is the option of using cluster approaches instead of periodic boundary conditions (see methods section for details).

Rank 3. The opposite route to rank 2 is taken with rank 3 approximations. Here, effects at the interface are explicitly included. A reduction in computational effort is achieved by diluting the surface coverage to a single adsorbate.

Consequently, all coverage and intra-layer effects cannot be captured with rank 3.

Rank 4. The simplest possible system is just the individual molecule without surface. There are, however, two possible ways in which one can design this model. Without prior knowledge of the surface structure, all molecular degrees of freedom have to be explored in order to find the minimum energy path connecting reactants and products. This model is assigned the rank 4b. On the other hand, some degrees of freedom are constrained when one of the reactants is immobilized on the substrate. Here, the gas phase minimum energy path might be inaccessible. Some mechanism-altering effects of the surface are therefore already contained in a rank 4 model. For instance, the cycloaddition of the AAC reaction cannot be initiated from below when the alkyne is attached to the surface.

7

Methods

Rank 1-4. Structure optimizations of all systems were performed with the Vienna Ab initio Simulation Package (VASP 5.4.4 [42-44]) compiled with the transition state tools by Henkelman et al.[45, 46]. The conjugate gradient algorithm for structural optimization was considered converged with forces smaller than 10^{-2} eV/Å. For gas phase calculations, stationary points found were verified by calculation of the Hessian matrix. The GGA-based exchange-correlation density functional proposed by Perdew, Burke and Ernzerhof (PBE) [47] was used to calculate the electronic energy and forces.

The inclusion of dispersion interaction corrections is important for the performance of semi-local density functionals (DFs) [48]. Benchmarks for low-density solids show good accuracy for bond lengths, angles, volumes and lattice constants.[49] In this study, dispersion effects were captured with the DFT-D3 scheme including the improved Becke-Johnson damping function [50, 51] for PBE and the rPBE [52]

functional. Furthermore, the range-separated hybrid functional HSE06 [53] which is based on PBE and includes 25% Hartree-Fock exchange at short distances was also paired with D3 in order to quantify improvements on reaction barriers. For comparison, the dispersion correction scheme by Tkatchenko and Scheffler (TS) [54] was applied to the PBE functional. With optB88 [55] a third approach to capture dispersion effects was investigated.

The planewave energy cutoff was set to 400 eV for all calculations and “standard” pseudo potentials for all atoms (version PBE5.4) [56] were applied in conjunction with the projector-augmented wave (PAW) method. The precision tag was set to accurate and a total energy difference of at most 10^{-5} eV was applied for self-consistent field (SCF) convergence. The Si(001) surface was described with a six-layered slab-supercell model with periodic boundary conditions (PBC). The theoretically optimized lattice parameter ($a=5.418$ Å) was determined previously [35]. The surface opposite to the organic layer was saturated with two hydrogen atoms per Si atom using the silane (SiH_4) bond length (1.480 Å) [57] while the bottom two layers were frozen at their bulk positions during relaxation. The thickness of the vacuum layer between periodically repeated units was at least 10 Å. Electronic levels around the Fermi-energy are smeared out by $\sigma=0.05$ eV with Gaussian functions to accelerate SCF convergence. Momentum space was sampled by a 2×2 Γ -centered k-point mesh for calculations including the substrate and Γ -only otherwise. This computational setup has yielded accurate results for molecular adsorbates in the past [35]. Organic layers were constructed by removal of the Si slab from the optimized system. Cleaved Si-C bonds were capped with H atoms on relaxed atomic positions while keeping all other atoms fixed. In all subsequent optimizations, the capping H atoms and directly attached C atoms were frozen. In order to calculate reaction barriers, transition state structures were optimized with the dimer method [45] after pre-determination with the nudged elastic band (NEB) [46] method.

Table 1. Benchmark of several density functionals and wavefunction-based methods in comparison to PNO-CCSD(T) data. Gas phase structures (rank 4) were optimized at the PBE-D3 level. Reaction energies for the three model reactions AAC, ACE and IEDDA are listed in the first part. From these three data points mean absolute errors (MAE) and mean signed errors (MSE) are calculated. In the second half of the table reaction barriers and their errors are listed.

Rkt.	PNO-CCSD(T)[a]	PNO-CCSD[a]	PNO-MP2[a]	PBE-D3[b]	PBE-D3[c]	PBE-TS[c]	rPBE-D3[c]	HSE06-D3[c]	optB88[c]
$\Delta E^0(\text{AAC})$	-289	-291	-292	-291	-297	-293	-295	-334	-286
$\Delta E^0(\text{ACE})$	-88	-95	-80	-66	-65	-59	-75	-80	-63
$\Delta E^0(\text{IEDDA})$	-64	-66	-60	-48	-51	-54	-39	-59	-54
MAE(0)	/	3.6	4.9	13.4	14.8	14.3	14.9	19.2	13.0
MSE(0)	/	-3.6	3.1	12.4	9.6	12.2	11.0	-10.4	13.0
$\Delta E^{\ddagger}(\text{AAC})$	96	117	72	65	63	69	60	83	61
$\Delta E^{\ddagger}(\text{ACE})$	55	74	15	27	26	34	17	39	25
$\Delta E^{\ddagger}(\text{IEDDA})$	99	112	86	71	69	73	73	90	66
MAE(‡)	/	17.7	26.0	28.9	30.5	24.8	33.2	13.0	32.8
MSE(‡)	/	17.7	-26.0	-28.9	-30.5	-24.8	-33.2	-13.0	-32.8

Energies and errors in kJ/mol; [a] basis set: cc-pVTZ; [b] basis set: def2-TZVPP; [c] planewave basis set ($E_{\text{cutoff}} = 400$ eV).

Effects of solvation were modelled by using VASPsol [21] that implements a PCM-like [58] model. Here, the electrostatic solute-solvent interaction can be efficiently described through a generalized Poisson equation. The only free parameter is the relative permittivity ϵ_r which is solvent-dependent and quantifies the rate at which a medium screens charges with respect to the vacuum ($\epsilon_r=1$ per definition). The dielectric constant is applied to the continuum region outside the solute cavity depending on an isodensity cutoff n_c . In order to prevent numerical instabilities due to a sharp change of the electrostatic potential around the cavity a diffuse region of width σ is introduced in which the relative permittivity is smoothly varied from 1 to ϵ . The solvation energy can be corrected for the destabilization due to cavitation by scaling the cavity shape function with the surface tension τ of the solvent. Default values for n_c and σ were applied with all other settings kept identical to the non-solvated case.

Due to repetition of the unit cell perpendicular to the slab plane, finite size effects caused by the presence of a net dipole moment can occur. The resulting errors to the energy and forces can be corrected by applying a linear potential counteracting the dipole moment [59]. In our test, using this dipole correction did not result in any total energy changes of a H-saturated, bare slab and was thus not applied to the interfaces investigated.

The simulated annealing (SA) molecular dynamics approach can be used to find minimum structures when the initial guess is poor. We

used this approach to find minima when manually generated structures were poor initial guesses. Calculations were performed with CP2K 5.1 [60] and the solution to the electronic SCF was given by the semi-empirical SCC-DFTB method with the “mio-1-1” parameter set [61] for elements H, C, N. Dispersion interactions were here included with a simplified D3 correction. We employ a three-step procedure where the system is first equilibrated in an NVT ensemble at $T = 1500$ K using a Nosé-Hoover thermostat for 1.25 ps with a time step of 0.5 fs. The high temperature does not lead to bond cleavage in a tight binding approach but ensures that the system explores a larger subsection of the configuration space. In the second step, the temperature is decreased slowly to 300 K with a cooling rate of 0.1% per step. Lastly, the structure found is crudely optimized (conjugate gradient algorithm, $10^{-3} E_h/a_0$) to the closest local minimum before the final optimization is performed with VASP as described above.

Rank 2 & 4. Additional single-point and gas phase calculations with a cluster approach were performed with TURBOMOLE 7.2 [62] using PBE-D3. The basis set def2-TZVPP [63] was used together with a fine integration grid ($m4$) and the SCF energy convergence criteria set to $10^{-8} E_h$. Clusters are constructed by simply removing PBC from a rank 2 supercell leading to saturated clusters. The number of atoms is thereby retained.

For the benchmark of gas phase energies, the orbital-specific virtual pair natural orbital coupled-cluster (OSV-PNO-CCSD(T)) method

implemented in TURBOMOLE 7.3 was employed [64]. Here, the correlation-consistent cc-pVTZ basis set including auxiliary functions was used while the energetic convergence criterion was lowered to 10^{-7} E_h in order to speed up the calculation.

Results and Discussion

Functional benchmark

Before quantifying the model error introduced by the ranking approach, it needs to be disentangled from the systematic error of the functional. In Table 1, we assess the performance of five density functionals (DF) including dispersion corrected by benchmarking against the wavefunction based “gold standard” OSV-PNO-CCSD(T) [64]. However, due to the size of the investigated molecules and current unavailability of coupled cluster methods for periodic systems, only rank 4 is benchmarked. Due to the varying accuracy of DFs with respect to minimum structures and transition states, reaction energies ($\Delta E^0 = \Delta E_{\text{prod}} - \Delta E_{\text{reac}}$) and barriers ($\Delta E^\ddagger = \Delta E_{\text{TS}} - \Delta E_{\text{reac}}$) are discussed separately. All reported energies were determined from single point calculations on the PBE-D3 optimized structure.

Furthermore, the mean signed error (MSE) is calculated in addition to the mean absolute error (MAE) to evaluate systematic trends of over- and underestimation of energies.

We begin with the comparison of different dispersion corrected GGA functionals for ΔE^0 . As shown in Table 1, all DFs give rather similar errors with optB88 showing the smallest MAE, followed closely by PBE-D3, PBE-TS and rPBE-D3. While the deviations between the DFs are small for the AAC reaction, larger deviations are found for ACE and IEDDA. For the barriers, ΔE^\ddagger , MAE and MSE values are identical in absolute terms suggesting a systematic error of the chosen DFs. While for DFT and MP2 too small barriers are found (negative MSEs), CCSD overestimates them (positive MSE). The PBE-TS functional gives

smaller errors than the other GGA functionals by 6-8 kJ/mol but errors are about two times larger than for the reaction energies.

Notably, range-separated hybrid DF HSE06-D3 shows the largest MAE for ΔE^0 which is caused by a failure to correctly describe the product of the AAC reaction. This behavior is a systematic error since similar reaction energies are also yielded for ranks 1-3 (see Table S6). A likely explanation is the overestimation of Hartree-Fock exchange for the triazole-ring as the similarly constructed PBE0-D3 functional yields a reaction energy of -336 kJ/mol (compared to the -334 kJ/mol of HSE06-D3). Concerning the barriers HSE06-D3 shows the best agreement with the CCSD(T) reference value as is often found for hybrid functionals that typically suffer less from delocalization errors of stretched bonds [65].

Thus, the conclusion drawn from this small benchmark fits the general notion that GGAs perform reasonably well for reaction energies while accurate reaction barriers require hybrid functionals. Overall, mean errors introduced by the choice of the DF are below 15 kJ/mol but larger deviations of more than 30 kJ/mol are found for some reactions. Notably, errors caused by the change from an atom-centered to a plane wave basis set together with the application of periodic boundary conditions are much smaller with at most 6 kJ/mol (compare columns 7 and 8 of Table 1). Unless noted otherwise, PBE-D3 is employed for all further calculations since it is the least parametrized functional while performing comparable to others and being computationally much more efficient than HSE06-D3.

Geometry changes with respect to model ranks

Systematic reduction of the system size has some influence on the geometry of the organic molecules which in turn leads to energetic differences. Removal of the substrate in rank 2 causes the organic layer to be slightly compressed due to missing steric repulsion between it and the topmost silicon atoms (Figure

S4). Steric repulsion is also present between molecules of the first and second organic layers. Removal of all surrounding molecules in rank 3 causes the single azide molecule (**2**) to stick out less towards the vacuum. This effect is somewhat compensated by the absence of dispersion attraction between substrate and adsorbate. The single molecule therefore stands more upright than in rank 1 (see Figure S5). Lastly, comparison of optimized geometries of rank 3 and 4 leads to the same local minimum and reveals no discernible differences.

Model error

Next, we investigate the influence of the ranking approach on barriers and reaction energies as well as on reaction paths. For this purpose, the dataset of the AAC reaction, found in Table 2 and plotted in Figure 7, is discussed in detail.

Table 2. Energies of stationary points on the AAC reaction path with respect to the isolated reactants as well as reaction energies and barriers. Model systems of ranks 1-4 are compared.					
	1	2	3	4	4b
Precomplex	-39	-48	-7	-8	-32
TS	20	9	51	55	31
Product	-336	-340	-311	-306	-310
ΔE^0	-297	-292	-304	-297	-278
ΔE^\ddagger	59	57	58	63	64
PBE-D3 energies in kJ/mol.					

Afterwards, comparison to the other two reactions, ACE (Figure 8) and IEDDA (Figure 9), are made.

At rank 1, the reaction energy and barrier of the AAC reaction are found to be -297 and 59 kJ/mol respectively. These results are not significantly changed at rank 2 since the precomplex and product as well as the TS are shifted slightly downwards in energy with respect to the isolated reactants. Further analyzing the relative energies without dispersion correction (Table S13) reveals that purely electronic effects are captured with cluster models just as well as with slab models. Thus, cluster calculations can be used in cases where no PBC implementation of a method is available. For rank 3, relative energies

of the stationary points are shifted upwards by 30-40 kJ/mol, resulting in a ΔE^0 of -304 kJ/mol. A similar shift upwards in relative energies is also determined for rank 4 which presents the surface-anchored structure. Here, the reaction energy agrees with rank 1. Relaxing all structural constraints imposed by the layer or substrate leads to additional angles of attack becoming available for the reaction at rank 4b. Thus, considering otherwise inaccessible interface structures in the gas phase model results in a less good agreement of rank 4b with the more elaborate models. We will thus not consider this approach further.

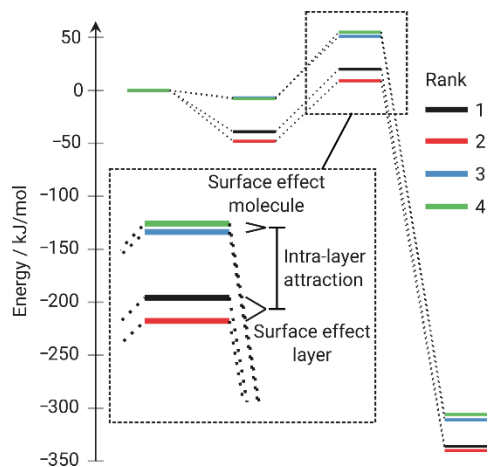


Figure 7. Energies (PBE-D3) of stationary points on the AAC reaction path from Table 2. Inset: The model systems show clustering of ranks 1 & 2 and 3 & 4 due to intra-layer attraction. Further splitting of ranks is caused by surface effects on layer and molecule.

In conclusion, reaction energies and barriers are only moderately affected by the rank of the model system. However, this observation only holds true provided structural information is carried over from previous ranks. Otherwise, changes in mechanism are possible in the simplified models as shown in column 5 of Table 2. From the data at hand, two dominant structural effects can be identified. The larger one, called intra-layer effect (Figure 7), causes a clustering of values for ranks 1 & 2 as well as 3 & 4. Due to the more crowded environment of the complete first organic layer, rank 1 & 2 structures are stabilized by dispersion interactions with respect to 3 & 4. The smaller

surface effect (Fig. 7) caused by restricting reaction possibilities through the surface as template is responsible for the remaining difference between ranks 1 & 2 and to a lesser extent between 3 & 4. Its influence can mostly be attributed to the presence of steric repulsion between the silicon surface and the first layer of molecules. This conclusion is supported by analyzing the energy diagrams of the ACE (Figure 8) and the IEDDA (Figure 9) reactions with respect to the clustering of ranks. However, the impact of the different ranks on ΔE^o and ΔE^\ddagger is not uniform, as shown in Table S5. While the reaction energy increases for ACE it decreases for IEDDA and stays approximately constant for AAC.

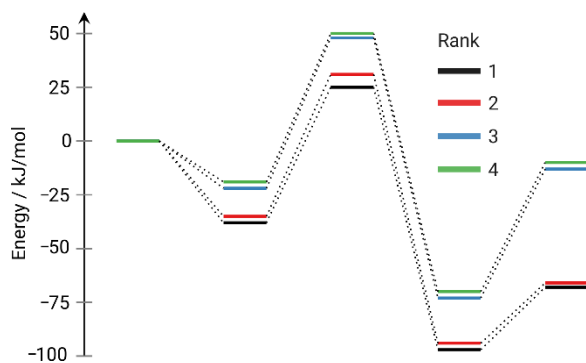


Figure 8. Energies of stationary points on the ACE reaction path from Table S3 at the PBE-D3 level.

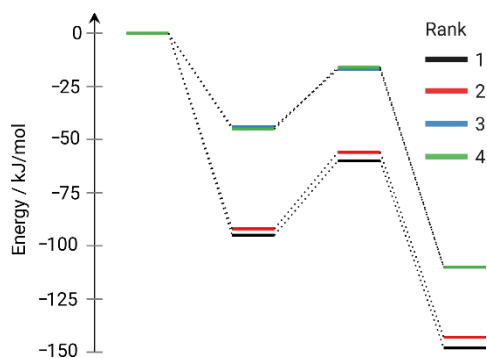


Figure 9. Energies of stationary points on the IEDDA reaction path from Table S4 at the PBE-D3 level.

Quantifying the model error introduced as a consequence of increasing computational simplicity is an important guideline for the decision whether an approximation is

acceptable. We therefore calculated the reaction energies and barriers of ranks 1-4 for all three reactions with the same five functionals benchmarked in Table 1. The full dataset is found in Tables S2-S9 of the supporting information. In Figure 10, the model error of ranks 2-4 for reaction barriers and energies is plotted as the MAE with respect to rank 1.

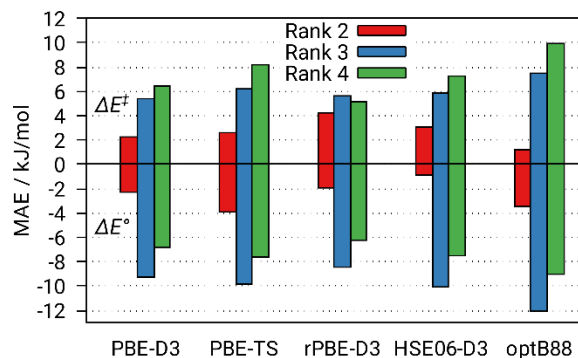


Figure 10. Model errors of different DFs with respect to the total system (rank 1). Mean absolute errors (MAE) of the barriers (ΔE^\ddagger) are plotted in the upper part while reaction energy errors (ΔE^o) are shown in the bottom part.

Independently of functional, models of rank 2 (Fig. 10, red bars) perform very well for reaction energies with mean errors relative to rank 1 consistently below 5 kJ/mol. In contrast, reaction energy MAEs for ranks 3 & 4 (Fig. 10, blue and green bars) are found to be significantly larger even surpassing 10 kJ/mol for optB88. A similar trend is present for the reaction barriers where rank 2 significantly outperforms 3 & 4 with the exception of rPBE. Taking a closer look at the functionals investigated clearly shows differing responses to the reduction of model complexity. The performance of rPBE is somewhat erratic as the functional has the lowest MAEs for ranks 3 & 4 but the highest for the barrier of rank 2. In light of these findings the usefulness of rPBE for the ranking approach is put into question. In terms of the three investigated dispersion correction schemes, DFT-D3 is slightly superior to the DFT-TS scheme while optB88 shows the largest errors. Surprisingly, the hybrid functional HSE06-D3 does not lead to significant improvements w.r.t.

PBE-D3.

However, in rank 2, HSE06 and optB88 show lower MAEs for reaction energies and barriers respectively.

We note that the statistical analysis performed here can only be understood as preliminary since a set of three reactions is subject to strong perturbation by outliers. An example is the low accuracy of rank 3 reaction energies over rank 4, which can be attributed to an outlier in the AAC reaction. Nonetheless we will discuss some implications in the following since the reactions are highly representative for interface chemistry.

Interplay of XC and model error. Model errors for reaction energies are smaller than those caused by the choice of the electronic structure method. We thus prioritize identifying a dispersion-corrected functional that well describes the chemistry at hand across different models. When calculating reaction rates which depend exponentially on the barrier model rank 2 should be considered. Interestingly, HSE06-D3 performs best for the reaction energy at rank 2 while having the largest error for the same quantity in the functional benchmark. The same is true for the barrier calculated with optB88. At this point we are unable to determine if these findings are correlated or by chance.

On the other hand, when the property of interest is not exponentially dependent on the error, gas phase models (rank 4) are suitable alternatives to rank 2. Their performance is only slightly worse than that of single molecule adsorbates (rank 3), at drastically reduced computational demand. Here, PBE-D3 convinces with a good overall performance and consistent error statistics across different model ranks. It is necessary though to compute the complete interface at rank 1 before resorting to simplified models.

Solvation effects

In the last part of this benchmark study, we use the PCM approach to investigate the influence of increasingly polar solvents on each model rank.

The ACE reaction is discussed here since the observed effects are most pronounced for its polar groups.

Relative energies of TS, pre- and postcomplex at rank 1 are listed in Table 3. The cavitation correction is neglected for the comparison of solvents since we aim to decrease the number of parameters by ignoring surface tension ($\tau=0$).

Table 3. Stabilization of the ACE reaction path (rank 1) by different solvents. The solvation is treated implicitly (ϵ in brackets). Positive values indicate destabilization w.r.t. the computations without solvent. All values at rank 1.

	DCM (9.1)	Acetone (20.7)	CH ₃ OH (32.7)	H ₂ O (80.1)
Precomplex	15	19	21	24
TS	11	15	17	21
Postcomplex	20	25	27	31
Σ Products [a]	9	11	12	13
ΔE^o	5	6	6	7
ΔE^{\ddagger}	-4	-4	-4	-3
Energies in kJ/mol without the cavitation energy. [a] Sum of the individual product energies after separation of the postcomplex.				

Inclusion of the solvent correction results in the destabilization of adsorbed structures with respect to the isolated reactants. This effect becomes stronger as solvent polarity increases. The postcomplex is destabilized more than the precomplex which is in turn less destabilized than the TS. However, the effect on reaction energies and barriers with varying polarity from 9.1 to 80.1 is small with energy changes of 3-7 kJ/mol.

Next, the behavior of implicit solvent modelling when applied to different model ranks is investigated (Table 4). For this part of the study the cavitation energy is included since its shape will be different for each rank. The cavitation energy is a destabilizing contribution to the solvation energy. Its effect is most pronounced for isolated reactants and products since the combined cavity volume is largest. Reduction of the volume upon formation of a precomplex has therefore a stabilizing effect on relative energies (Table S10).

Table 4. Stabilization of the ACE reaction path by implicit solvation (H ₂ O). Model system ranks 1-4 are compared. Positive values indicate destabilization w.r.t. the computations without solvent.				
	1	2	3	4
Precomplex	20	17	11	11
TS	16	13	-1	-1
Postcomplex	25	24	13	13
Σ Products [a]	10	10	-5	-6
ΔE^o	5	7	2	2
ΔE^{\ddagger}	-4	-4	-12	-11
Energies in kJ/mol including the cavitation energy.				
[a] Sum of the individual product energies after separation of the postcomplex.				

The previously noted clustering of results for model ranks 1 & 2 and 3 & 4 is also found for the solvation correction. This suggests a rather weak influence of the substrate but strong influence of the organic layer. This behavior is expected since the organic layer reduces the solvent accessible surface by blocking the substrate-facing hemisphere. Thus, adsorbing molecules need to shed a larger part of their solvent sphere before being able to react, resulting in an energetic penalty. This observation is generalizable as shown in Tables S11 & S12.

We note that the accuracy of the PCM model for interfaces cannot be judged with great confidence at this point since experimental data is scarce. Some computational parameters, especially concerning cavity construction might need to be reevaluated in the future.

Applying the model

After having determined that a rank 2 model, containing just the organic layer, well reproduces the full interface system while significantly reducing computational demand, we now apply it to two typical questions in interface research (Figure 11).

The first question concerns the coverage dependence of adsorption energies. With increasing coverage, attraction between adsorbates rises, thus increasing the adsorption energy. This effect can also lead to a steering of subsequent molecules to adsorb close to those already present on the surface [36]. At higher

coverages, the adsorption energy falls again due to steric repulsion experienced in more crowded environments [66]. Knowledge of the adsorption energy profile with respect to coverage is important to monitor the deposition progress.

The second question treated concerns unwanted termination reactions of bifunctional molecules which negatively impact growth rates and interface uniformity. In a termination reaction, a single molecule reacts twice with the previous layer, effectively decreasing the number of reactive sites by two [69]. After a certain number of deposition cycles no reactive sites remain and growth comes to a halt. This problem is primarily observed for molecules carrying two identical functional groups. However, those molecules are still preferred for layer-by-layer synthesis since only two steps per growth cycle are required there. Furthermore, every additional different functional group in a molecule increases the risk of side reactions. Therefore, an approach to increase the selectivity of bifunctional molecules proposed earlier [68] is computationally examined here. The idea is to increase the steric demand of one functional group by attaching bulky carbon moieties in its vicinity (i.e. linkers **11** and **15** in Figure 12). Extending this idea, we present additional functional groups to suppress these termination reactions.

Both of these research questions cannot be fully addressed with gas phase or single adsorbate models as they cannot capture coverage effects. Hence, they lend themselves as showcases for the periodic, organic layer model (rank 2).

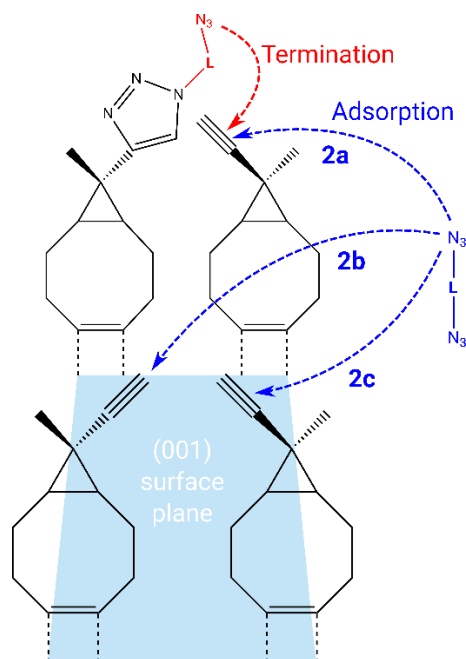


Figure 11. Lewis structure representation of the complete first organic layer after AAC reaction with a single molecule attached in the second layer. This molecule can perform a termination reaction. In the 4×4 supercell, three possible adsorption modes for a second molecule (N_3 -L- N_3) exist (2a, 2b, 2c). L is a linker group (see Figure 12).

Coverage dependent adsorption energy

Based on the idealized interface model of one CCO molecule per two surface dimers as introduced above in Figure 5, the smallest unit cell allowing the study of coverage contains four reactive sites in the first organic layer. This cell can accommodate four possible coverages: $\Theta = 0.25, 0.5, 0.75, 1$ ML. Representative structures were generated from simulated annealing molecular dynamics simulations as described in the methods section. The results are given in Table 5.

The adsorption energy of the first molecule (0.25 ML, 340 kJ/mol) has been reported in column 3 of Table 2. For a coverage of 0.5 ML, the average adsorption energy per molecule falls off considerably (-313 to -319 kJ/mol) before rising again for values of 0.75 and 1 ML. When increasing the coverage from 0.25 to 0.5 ML three adsorption modes are possible (**2a-2c**, Figure 11). A molecule adsorbing at the neighboring position across the Si-dimer trench (**2a**) has an adsorption energy of -286 kJ/mol.

Within a dimer-row (**2b**) adsorption is slightly more favorable at -295 kJ/mol.

Θ / ML	$\Delta E_{\text{ads}}(\Theta)$	$\Delta \Delta E_{\text{ads}}(\Theta)$	$\Delta \Delta E_{\text{ads}}(\text{disp})$
0.25	-340	-340	-63
0.5 [a]	-626 (-313)	-286	-64
0.5 [b]	-635 (-317)	-295	-26
0.5 [c]	-638 (-319)	-298	-55
0.75	-972 (-324)	-334	-101
1	-1320 (-330)	-348	-70

PBE-D3 energies in kJ/mol for rank 2 models.
 [a] Configuration **2a** (Figure 11).
 [b] Configuration **2b** (Figure 11).
 [c] Configuration **2c** (Figure 11).

Notably, the dispersion contribution for case **2a** (-63 kJ/mol) is almost the same as for the first molecule (-64 kJ/mol), while it is less for the intra-row case **2b** (-26 kJ/mol). This is due to space constraints disfavoring relaxation into a flattened structure of two neighboring molecules within a row (see Figure S6). In the case of two molecules being adsorbed along the diagonal (**2c**), a similar behavior to **2a** is found with a relatively high dispersion contribution to the adsorption energy.

For the third incoming adsorbate (0.75 ML), only one unique configuration exists. Its adsorption energy rises again to -334 kJ/mol (relative to **2c**) with dispersion attraction contributing strongly (30%).

The fourth molecule (1 ML) has the highest adsorption energy (-348 kJ/mol). Here, the dispersion contribution amounts to 20%.

In summary, adsorption to directly neighboring sites is electronically less favorable while dispersion attraction stabilizes higher coverages. Due to the small number of reactive sites per area in the second layer, additional molecules always react exothermically until full coverage is reached. In order to validate the performance of the rank 2 model, the value for the adsorption of the fourth molecule is re-calculated at rank 1. Here the adsorption energy is on average -342 instead of -330 kJ/mol. Two thirds of this increase can be attributed to additional

dispersion attraction of the substrate. An absolute error of 12 kJ/mol is acceptable for applications where correct trends are the primary concern.

Prevention of terminating reactions

In the following we assess the success of different modifications to a molecular building block for reducing the thermodynamic driving force of a termination reaction shown in Figure 11. As a test case the AAC reaction is used where the two functional groups of the azide precursor are differentiated. In Figure 12a two neighboring CCO derived molecules are terminated by a bisazide molecule. Through variation of the linker group **L** the thermodynamic termination energy is calculated with respect to reaction at a single site [$\Delta E_{\text{term}} = \Delta E_{\text{ads}}(2) - \Delta E_{\text{ads}}(1)$] and listed in Table 6. We focus here on links across the trench, however, the same approach could also be applied to termination reactions within a row.

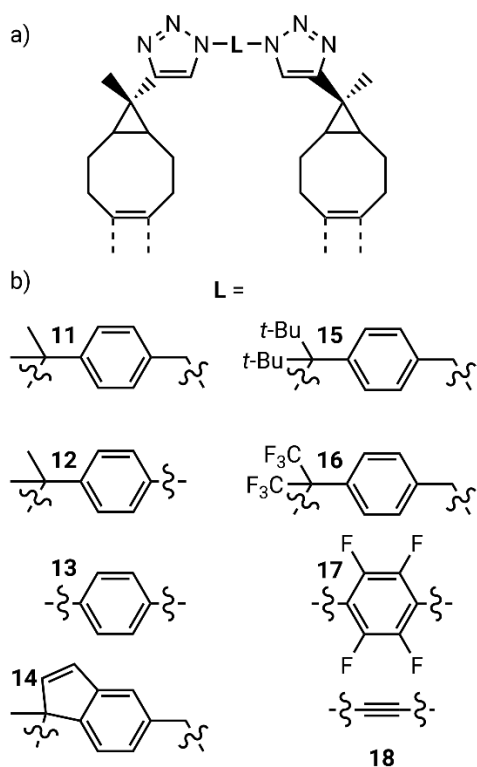


Figure 12. a) Termination of two alkyne reactive sites by a bisazide molecule in the AAC scheme. b) Linker groups **L** altering the reactivity of the bisazide through steric demand, strain and electron withdrawing groups.

Table 6. Differences in growth termination reaction energies caused by the chemical modification of linker groups L in bisazide molecules. Relative reaction energies in column three are calculated with respect to linker 11 .		
Linker (L)	$\Delta E_{\text{term}} / \text{kJ/mol}$	$\Delta\Delta E_{\text{term}} / \text{kJ/mol}$
11	-274	0
12	-263	10
13	-171	102
14	-275	-1
15	-256	18
16	-258	16
17	-167	107
18	-177	97

The modifications made to the bisazide are rationalized by one or more of the following design principles. Molecule **11** is structurally similar to reactant **2** of the AAC reaction but contains two azide moieties, one of which is sterically encumbered by methyl groups. In order to further reduce the termination energy, steric demand is increased by replacing both methyl groups with *tert*-butyl groups in linker **15**. An alternative to steric demand is strain engineering, with successively shorter linkers (**11**>**12**>**13**>**18**) or more rigid backbones (**14**). Lastly, the option of chemically tuning the linker through adding electron withdrawing groups (EWG) on the dienophile is explored. This last approach might be combined with the other ideas by exchanging -CH₃ with -CF₃ groups (**16**) or fluorine substitutions at the benzene ring (**17**).

Steric effect. The adsorption energy of **11** with only the less sterically hindered azide group (-346 kJ/mol) is similar to that of **2** (-340 kJ/mol). Termination of a neighboring reactive site is energetically less favorable by 72 kJ/mol, however a thermodynamic driving force of -274 kJ/mol remains, indicating the insufficiency of two methyl groups to increase selectivity. Comparing this value at rank 2 to rank 1 (-280 kJ/mol) shows a small deviation of -6 kJ/mol.

The thermodynamic driving force is decreased by just 18 kJ/mol when using linker **15**. This suggests that steric demand can only slightly inhibit termination and other avenues should be pursued instead.

Strain effect. Removal of a single -CH₂- unit leading to **12** has only a small effect of 10 kJ/mol

while removal of two (**13**) results in a large destabilization of 102 kJ/mol. Replacing the phenyl ring with an alkyne **18** has a similar influence on the termination energy (97 kJ/mol). However, such a molecule is only a hypothetical option due to cross reactivity between the azide and alkyne moieties within the same molecule. Therefore, we conclude that the length of **L** is a successful indicator for the thermodynamic stability of termination products.

Simply increasing the rigidity by introducing a second ring (e.g. **14**) appears to not decrease reactivity, however higher growth rates per cycle can be expected [69].

Electronic effect. This effect is quantifiable but rather small with 6 kJ/mol for CF₃ groups in **16** and 5 kJ/mol decreased reactivity for **17**.

In summary, strain is the most promising approach for the prevention of termination reactions using thermodynamic control. This assessment can be expanded to low temperature kinetic control by invoking the Bell-Evans-Polany principle stating a linear correlation between reaction energy and barrier. The optimal length of linker groups is system-dependent due to templating effects of the underlying substrate. In the case of Si(001), inter- and intra-row dimers distances must therefore be taken into consideration. This is possible within the efficient rank 2 interface model.

Conclusions

We investigated the modeling of organic layers on semiconductor surfaces with a hierarchy of structural models on the density functional theory level. The assumption that chemical reactions involving only localized bonds can be accurately described by a limited model that mimics structural constraints imposed by the substrate could be supported by a small benchmark set of three different reactions. It was furthermore shown that intra-layer interactions such as dispersion attraction and steric repulsion are crucial to obtain accurate barriers and reaction energies. For the latter

case the least parametrized GGA functional PBE performs the most consistently while a hybrid functional description improves on barrier heights. We note however that due to the small size of our test set, statistical evaluation should be taken with caution. Additional calculations should be performed in the future to supplement existing data.

We have also expanded the theoretical description of attachment processes on organic layers to the wet-chemical domain. Here, implicit models like PCM offer an efficient means of including electrostatic solvent interactions. Again, models abandoning the substrate in favor of computational simplicity perform well since the shape of the solvent shell is mostly determined by the geometry of the previous organic layer.

Application of the model was then showcased with two typical questions of interface research. Both the variation of adsorption energy with coverage and the thermodynamics of termination reactions require a periodic model of the previous layer. For the latter question, a correlation between the length of the linker molecule and strain induced decrease of reaction energy was found. Steric demand and chemical tuning seem to be inferior means of inhibiting termination.

Lastly, we offer recommendations for preferred use cases of the four introduced model ranks (summarized in Table 7). The full system of rank 1 originally shown in Figure 1 is best employed to generate reference data to benchmark the other models. Furthermore, it is used to determine template effects such as lattice spacings and coverages. Finally, one may be forced to include the substrate when actual interface properties are of interest. A rank 2 model should be considered as the main method to study the chemistry within the organic layers. The rank 3 model is a common choice to investigate adsorption energies and reactivity on the substrate. Finally, the gas phase model (rank 4) can be used for benchmarking against more accurate approaches or quickly screening molecular building blocks before employing rank 2.

Table 7. Recommendations for model systems addressing common research questions in the study of organic-inorganic interfaces.

Model	Description	Research question
Rank 1	Full system	Benchmarks, template effects, interface properties
Rank 2	Organic layer	Energies and barriers, layer effects, bonding analysis
Rank 3	Substrate + 1 molecule	Adsorption energies, adsorption modes
Rank 4	Gas phase	High level methods, quick screening

It remains open for investigation how fast template effects decay with the number of layers. Disorder increases in the second layer as molecules become free to explore the configurational space. It is therefore highly likely that structural properties of a layered organic system change quickly from an ordered interface to an (amorphous) bulk. In that case computational tools developed for polymer research [70] including (semi-)classical force fields and potential of mean force approaches can be applied. This opens up an opportunity to implement a multi-scale approach where the inorganic semiconductor and the interface is treated with high level methods while very efficient semi-empirical methods are applied to the organic bulk. Hence, the computation of macroscopic properties such as the elastic modulus and thermal stability may become feasible.

Acknowledgments

We thank Prof. Ulrich Koert and Jannick Meinecke for the suggestion of model reactions and Fabian Pieck for sharing preliminary computational results on the adsorption behavior of CCO on Si(001). Financial support was given by the German Research Foundation (DFG) via SFB 1083. Computational resources provided by HRZ Marburg, RZ Regensburg, GOETHE-CSC Frankfurt and HLR Stuttgart are gratefully acknowledged.

Keywords: Density functional theory, model hierarchy, interfaces, hybrid organic-inorganic materials, semiconductor functionalization.

References and Notes

1. Wolkow, R. A. Controlled molecular adsorption on silicon: Laying a Foundation for Molecular Devices. *Annu. Rev. Phys. Chem* **1999**, *50*, 413–441.
2. Yates, J. T. A New Opportunity in Silicon-Based Microelectronics. *Science* **1998**, *279*, 335–336.
3. Aradhya, S. V.; Venkataraman, L. Single-molecule junctions beyond electronic transport. *Nat. Nanotechnol.* **2013**, *8*, 399–410.
4. Hills, G.; Lau, C.; Wright, A.; Fuller, S.; Bishop, M. D.; Srimani, T.; Kanhaiya, P.; Ho, R.; Amer, A.; Stein, Y.; Murphy, D.; Arvind,; Chandrakasan, A.; Shulaker, M. M. Modern microprocessor built from complementary carbon nanotube transistors. *Nature* **2019**, *572*, 595–602.
5. Miozzo, L.; Yassar, A.; Horowitz, G. Surface engineering for high performance organic electronic devices: the chemical approach. *J. Mater. Chem.* **2010**, *20*, 2513–2538.
6. Zhou, H.; Bent, S. F. Fabrication of organic interfacial layers by molecular layer deposition: Present status and future opportunities. *J. Vac. Sci. Technol* **2013**, *31*, 040801.
7. Teplyakov, A. V.; Bent, S. F. Semiconductor surface functionalization for advances in electronics, energy conversion, and dynamic systems. *J. Vac. Sci. Technol* **2013**, *31*, 050810.
8. George, S. M.; Yoon, B.; Dameron, A. A. Surface Chemistry for Molecular Layer Deposition of Organic and Hybrid Organic-Inorganic Polymers. *Acc. Chem. Res.* **2009**, *42*, 498–508.
9. Yates, J. T. Surface chemistry of silicon-the behaviour of dangling bonds. *J. Phys. Condens. Matter* **1991**, *3*, S143–S156.

10. Sundberg, P.; Karppinen, M. Organic and inorganic-organic thin film structures by molecular layer deposition: A review. *Beilstein J. Nanotechnol.* **2014**, *5*, 1104–1136.
11. Meng, X. An overview of molecular layer deposition for organic and organic-inorganic hybrid materials: mechanisms, growth characteristics, and promising applications. *J. Mater. Chem. A* **2017**, *5*, 18326–18378.
12. Zhou, H.; Bent, S. F. Molecular Layer Deposition of Functional Thin Films for Advanced Lithographic Patterning. *Appl. Mater. Interfaces* **2011**, *3*, 505–511.
13. Palomaki, P. K. B.; Dinolfo, P. H. A Versatile Molecular Layer-by-Layer Thin Film Fabrication Technique Utilizing Copper(I)-Catalyzed Azide-Alkyne Cycloaddition. *Langmuir* **2010**, *26*, 9677–9685.
14. Länger, C.; Heep, J.; Nikodemiak, P.; Bohamud, T.; Kirsten, P.; Höfer, U.; Koert, U.; Dürr, M. Formation of Si/organic interfaces using alkyne-functionalized cyclooctynes – precursor-mediated adsorption of linear alkynes versus direct adsorption of cyclooctyne on Si(001). *J. Phys. Condens. Matter* **2018**, *31*, 034001.
15. Meinecke, J.; Koert, U. Copper-Free Click Reaction Sequence: A Chemoselective Layer-by-Layer Approach. *Org. Lett.* **2019**, *21*, 7609–7612.
16. Mulliken, R. S. Structures of Complexes Formed by Halogen Molecules with Aromatic and with Oxygenated Solvents. *J. Am. Chem. Soc.* **1950**, *72*, 600–608.
17. Biswas, R.; Hamann, D. R. Simulated annealing of silicon atom clusters in Langevin molecular dynamics. *Phys. Rev. B* **1986**, *34*, 895–901.
18. Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.
19. Klamt, A.; Moya, C.; Palomar, J. A Comprehensive Comparison of the IEFPCM and SS(V)PE Continuum Solvation Methods with the COSMO Approach. *J. Chem. Theory Comput.* **2015**, *11*, 4220–4225.
20. McCann, B. W.; McFarland, S.; Acevedo, O. Benchmarking Continuum Solvent Models for Keto-Enol Tautomerizations. *J. Phys. Chem. A* **2015**, *119*, 8724–8733.
21. Mathew, K.; Sundararaman, R.; Letchworth-Weaver, K.; Arias, T. A.; Hennig, R. G. Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways. *The Journal of Chemical Physics* **2014**, *140*, 084106.
22. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
23. Oakdale, J. S.; Kwisnek, L.; Fokin, V. V. Selective and Orthogonal Post-Polymerization Modification using Sulfur(VI) Fluoride Exchange (SuFEx) and Copper-Catalyzed Azide-Alkyne Cycloaddition (CuAAC) Reactions. *Macromolecules* **2016**, *49*, 4473–4479.
24. Fan, H.; Ji, Y.; Xu, Q.; Zhou, F.; Wu, B.; Wang, L.; Li, Y.; Lu, J. Sulfur (VI) Fluoride Exchange Polymerization for Large Conjugate Chromophores and Functional Main-Chain Polysulfates with Nonvolatile Memory Performance. *ChemPlusChem* **2018**, *83*, 407–413.
25. Binder, W. H.; Kluger, C. Azide/Alkyne-Click Reactions: Applications in Material Science and Organic Synthesis. *Curr. Org. Chem.* **2006**, *10*, 1791–1815.
26. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.
27. Münster, N.; Nikodemiak, P.; Koert, U. Chemoselective Layer-by-Layer Approach Utilizing Click Reactions with Ethynylcyclooctynes and Diazides. *Org. Lett.* **2016**, *18*, 4296–4299.
28. Ruff, F.; Farkas, Ö. Concerted SN2 mechanism for the hydrolysis of acid chlorides: comparisons of reactivities calculated by the density functional theory with experimental data. *J. Phys. Org. Chem.* **2011**, *24*, 480–491.
29. Domingo, L. R.; Picher, M. T.; Sáez, J. A. Toward an Understanding of the Unexpected Regioselective Hetero-Diels-Alder Reactions of

Asymmetric Tetrazines with Electron-Rich Ethylenes: A DFT Study. *J. Org. Chem.* **2009**, *74*, 2726–2735.

30. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide-Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

31. Yoshinobu, J. Physical properties and chemical reactivity of the buckled dimer on Si(100). *Prog. Surf. Sci.* **2004**, *77*, 37–70.

32. Pecher, L.; Schmidt, S.; Tonner, R. Modeling the Complex Adsorption Dynamics of Large Organic Molecules: Cyclooctyne on Si(001). *J. Phys. Chem. C* **2017**, *121*, 26840–26850.

33. Reutzel, M.; Münster, N.; Lipponer, M. A.; Länger, C.; Höfer, U.; Koert, U.; Dürr, M. Chemoselective Reactivity of Bifunctional Cyclooctynes on Si(001). *J. Phys. Chem. C* **2016**, *120*, 26284–26289.

34. Pecher, L.; Tonner, R. Computational analysis of the competitive bonding and reactivity pattern of a bifunctional cyclooctyne on Si(001). *Theor. Chem. Acc.* **2018**, *137*, 48.

35. Pecher, L.; Schober, C.; Tonner, R. Chemisorption of a Strained but Flexible Molecule: Cyclooctyne on Si(001). *Chem. Eur. J.* **2017**, *23*, 5459–5466.

36. Pecher, L.; Schmidt, S.; Tonner, R. Dispersion-mediated steering of organic adsorbates on a precovered silicon surface. *Beilstein J. Org. Chem.* **2018**, *14*, 2715–2721.

37. Wodrich, M. D.; Corminboeuf, C.; Schreiner, P. R.; Fokin, A. A.; Schleyer, P. v. R. How Accurate Are DFT Treatments of Organic Energies? *Org. Lett.* **2007**, *9*, 1851–1854.

38. Alkorta, I.; Elguero, J. The carbon-carbon bond dissociation energy as a function of the chain length. *Chem. Phys. Lett.* **2006**, *425*, 221–224.

39. Heep, J.; Luy, J.-N.; Länger, C.; Meinecke, J.; Koert, U.; Tonner, R.; Dürr, M. Adsorption of Methyl-Substituted Benzylazide on Si(001): Reaction Channels and Final Configurations. *J. Phys. Chem. C* **2020**, *124*, 9940–9946.

40. Glaser, T.; Meinecke, J.; Länger, C.; Luy, J.-N.; Tonner, R.; Koert, U.; Dürr, M. Combined XPS and DFT investigation of the adsorption modes of methyl enol ether functionalized cyclooctyne on Si(001). Submitted.

41. Glaser, T.; Meinecke, J.; Freund, L.; Länger, C.; Luy, J.-N.; Tonner, R.; Koert, U.; Dürr, M. Click chemistry in ultra-high vacuum - tetrazine coupling with methyl enol ether covalently linked to Si(001). In preparation.

42. Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **1993**, *47*, 558.

43. Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15.

44. Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

45. Henkelman, G.; Jónsson, H. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.* **1999**, *111*, 7010–7022.

46. Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901–9904.

47. Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

48. Marom, N.; Tkatchenko, A.; Rossi, M.; Gobre, V. V.; 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*, 3944–3951.

49. Nazarian, D.; Ganesh, P.; Sholl, D. S. Benchmarking density functional theory predictions of framework structures and properties in a chemically diverse test set of metal-organic frameworks. *J. Mater. Chem. A* **2015**, *3*, 22432–22440.

50. 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.
51. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
52. Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **1999**, *59*, 7413–7421.
53. Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *J. Chem. Phys.* **2006**, *125*, 224106.
54. 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*, 073005.
55. Klimeš, J.; Bowler, D. R.; Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys. Condens. Matter* **2009**, *22*, 022201.
56. Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59*, 1758–1775.
57. Boyd, D. R. J. Infrared Spectrum of Trideuterosilane and the Structure of the Silane Molecule. *J. Chem. Phys.* **1955**, *23*, 922–926.
58. Gunceler, D.; Letchworth-Weaver, K.; Sundararaman, R.; Schwarz, K. A.; Arias, T. A. The importance of nonlinear fluid response in joint density-functional theory studies of battery systems. *Model. Simul. Mat. Sci. Eng.* **2013**, *21*, 074005.
59. Neugebauer, J.; Scheffler, M. Adsorbate-substrate and adsorbate-adsorbate interactions of Na and K adlayers on Al(111). *Phys. Rev. B* **1992**, *46*, 16067–16080.
60. Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. cp2k: atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.
61. 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*, 7260–7268.
62. TURBOMOLE V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
63. Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
64. Schmitz, G.; Hättig, C.; Tew, D. P. Explicitly correlated PNO-MP2 and PNO-CCSD and their application to the S66 set and large molecular systems. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22167–22178.
65. Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **2003**, *118*, 8207–8215.
66. Kachel, S. R.; Klein, B. P.; Morbec, J. M.; Schöniger, M.; Hutter, M.; Schmid, M.; Kratzer, P.; Meyer, B.; Tonner, R.; Gottfried, J. M. Chemisorption and Physisorption at the Metal/Organic Interface: Bond Energies of Naphthalene and Azulene on Coinage Metal Surfaces. *J. Phys. Chem. C* **2020**, *124*, 8257–8268.
67. Bergsman, D. S.; Closser, R. G.; Bent, S. F. Mechanistic Studies of Chain Termination and Monomer Absorption in Molecular Layer Deposition. *Chem. Mater.* **2018**, *30*, 5087–5097.
68. Golas, P. L.; Tsarevsky, N. V.; Matyjaszewski, K. Structure-Reactivity Correlation in "Click" Chemistry: Substituent Effect on Azide Reactivity. *Macromol. Rapid Commun.* **2008**, *29*, 1167–1171.
69. Bergsman, D. S.; Closser, R. G.; Tassone, C. J.; Clemens, B. M.; Nordlund, D.; Bent, S. F. Effect of Backbone Chemistry on the Structure of Polyurea Films Deposited by Molecular Layer Deposition. *Chem. Mater.* **2017**, *29*, 1192–1203.

70. Abbott, L. J.; Hart, K. E.; Colina, C. M. Polymatic: a generalized simulated polymerization algorithm for amorphous polymers. *Theor. Chem. Acc.* **2013**, 132, 1334.

GRAPHICAL ABSTRACT

Jan-Niclas Luy, Mahlet Molla, Lisa Pecher, and Ralf Tonner

Efficient hierarchical models for reactivity of organic layers on semiconductor surfaces

A hierarchical approach for modeling organic interface formation on semiconductors is presented. The largest model system (rank 1) constituting the full interface is computationally very demanding. By removing parts of the interface such as substrate or neighboring molecules from computations, properties of interest can be more efficiently calculated while retaining acceptable accuracy.

