

Driving forces for covalent assembly of porphyrins by selective C-H bond activation and intermolecular coupling on a copper surface

Andrea Floris^{1,4*}, Sam Haq², Mendel In't Veld², David B. Amabilino³, Rasmita Raval^{2*}, and Lev Kantorovich^{1*}

¹Department of Physics, King's College London, London, Strand WC2R 2LS, UK.

²Surface Science Research Center and Department of Chemistry, University of Liverpool, Liverpool L69 3BX, UK.

³School of Chemistry, The University of Nottingham, University Park, Nottingham NG7 2RD, UK

⁴School of Mathematics and Physics, University of Lincoln, Brayford Pool, Lincoln LN6 7TS, UK

*e-mail: an.floris@gmail.com; lev.kantorovitch@kcl.ac.uk; R.Raval@liverpool.ac.uk

ABSTRACT

Recent synthesis of covalent organic assemblies at surfaces has opened up the promise of producing robust nanostructures for functional interfaces. To uncover how this new chemistry works at surfaces and understand the underlying mechanism(s) that control bond-breaking and bond-making processes at specific positions of the participating molecules, we study here the coupling reaction of tetra(mesityl)porphyrin molecules, which creates covalently connected networks on the Cu(110) surface by utilising the 4-methyl groups as unique connection points. Using scanning tunneling microscopy (STM), state-of-the-art density functional theory (DFT) and Nudged Elastic Band (NEB) calculations, we show that the unique directionality of the covalent bonding is found to stem from a chain of highly selective C-H activation and de-hydrogenation processes, followed by specific intermolecular C-C coupling reactions that are facilitated by the surface, by steric constraints and by anisotropic molecular diffusion. These insights provide the first steps towards developing synthetic rules for complex two-dimensional covalent organic chemistry that can be enacted directly at a surface to deliver specific macromolecular structures designed for specific functions.

1. Introduction

Construction of molecular assemblies and networks at surfaces [1,2] provides a highly promising protocol for synthesizing new 2-D materials, delivering new functionalities for biological applications such as sensors and drug delivery [3] and advanced nanotechnology applications in energy harvesting, catalysis and nano-electronic devices [4,5,6]. For many applications, robust networks stabilized by *covalent* bonds between constituent molecules [1,2,7-12] are technologically more promising, compared to supramolecular networks stabilized by other types of inter-molecular interactions such as H-bonding or van der Waals [13-19]. One broadly applicable approach [2] for the on-surface synthesis of complex and diverse covalent structures is to exploit the prevalence of C-H bonds in organic entities and activate them at a surface to drive intermolecular coupling via C-C or C-Metal-C bond formation [2,12,20,21]. However, while this general approach provides broad applicability, there remains a real need to understand and control both the selectivity of C-H bond activation and the intermolecular coupling process so that specific final products are favored, as the first step towards delivering targeted and tailored structures.

Clearly, theory must play a central role in understanding the parameters that govern specific C-H bond activation and the subsequent intermolecular reactions mediated at the surface, hence providing the necessary insights for the experiment. Recently, calculations based on density functional theory (DFT) have started to address the mechanisms underlying the on-surface covalent bonding of molecules in simple cases [10,22-25]. However, for covalent structures involving large and complex organic molecules abundant with C-H bonds, mechanistic details are scarce; here, an important advance would be to *predict* why, at the given experimental conditions, only particular C-H bonds get activated leading to specific intermolecular connectivities. Such knowledge would underpin future strategies for steering the assembly in the desired direction. To our knowledge, the question of *selectivity* in the on-surface chemistry mediated *only* by C-H activation and dehydrogenated (de-H) reactions has not been addressed theoretically before.

In this work we analyse the general mechanisms underpinning such selectivity by using, as a prototype example, the covalent coupling of *tetra-(2,4,6-tri-methyl-phenyl)-porphyrin* (TMTPP) molecules on the Cu(110) surface, as reported by scanning tunnelling microscopy (STM) experiments [12]. We analyse how upon annealing to 500K, these rather large molecules, possessing a multitude of external C-H bonds, form uniquely oriented one-dimensional chains and small clusters via specific C-H bond activation. We demonstrate that this high selectivity results from a combination of the intrinsic chemistry of the molecule, the geometry adopted by the molecule at the surface, the catalytic effect of the surface and specific kinetics associated with underlying processes. All these effects combine to drive C-H activation, dehydrogenation (de-H) and C-C coupling reactions to occur only at *particular methyl groups*, explaining the unique molecular connectivity. Finally, we explain the role of annealing in forming the networks and identify the preferential diffusion patterns of TMTPP on this surface, which are paramount in determining the network growth.

2. Results and Discussion

2.1 Structure of Tri-methyl-tetra-phenyl-porphyrin (TMTPP) on Cu(110)

TMTPP is composed of a porphyrinic tetra-pyrrolic core functionalized at the meso positions by four phenyl rings, each having three methyl groups, two in the carbon atom positions adjacent to the porphyrin (the 2- and 6-positions) and one in the 4-position, as shown in Fig. 1(a). Our gas-phase DFT calculations show a stable structure with a flat porphyrin core and the planes of the phenyl rings oriented almost perpendicularly to it, Fig. 1(b).

The molecule is then placed on the Cu(110) surface, consisting of close-packed Cu rows running along the $[1\bar{1}0]$ direction, Fig. 1(c). Periodic DFT calculations show that the presence of the surface strongly modifies the molecular geometry, which undergoes a complex re-configuration into a structure, which is very different from the gas-phase one. The pyrrole rings incorporating the Lewis basic nitrogen atoms lie with their mean plane almost perpendicular to the substrate because of the formation of two N-Cu bonds with the Cu rows, as shown in Fig. 1(c). In addition, two C-Cu bonds are formed by the $-NH$ containing pyrrole rings, which are oriented almost horizontally to the surface plane, Fig. 1(c,d). After a comprehensive DFT analysis of other (~ 10) less stable conformations [Supporting Information (SI), Fig SI-2], we conclude that the N-Cu and C-Cu covalent links dominate the energetic landscape of the system. In addition, intra-molecular interactions are important in determining the final geometry. Each phenyl ring is rotated, with the 6-methyl groups located much closer to the surface than the 2- and the 4-methyl moieties, with the alternating orientations of the pyrrole rings and the planar phenyl group configurations due to the balance between the electrostatic repulsion of the 2- and 6-methyl groups with the core and the steric constraints imposed by the surface, Fig. 1(c,d). Indeed, our simulations show that in the absence of phenyl groups, the core lies completely flat on Cu(110) (Fig. SI-3), as reported before [26]. The interplay and optimization of intra-molecular interactions and covalent core-surface N-Cu and C-Cu bonds ensures a strong molecule-surface binding, with an adsorption energy of 5.78 eV and a characteristic geometry that is supported by STM data, as discussed below.

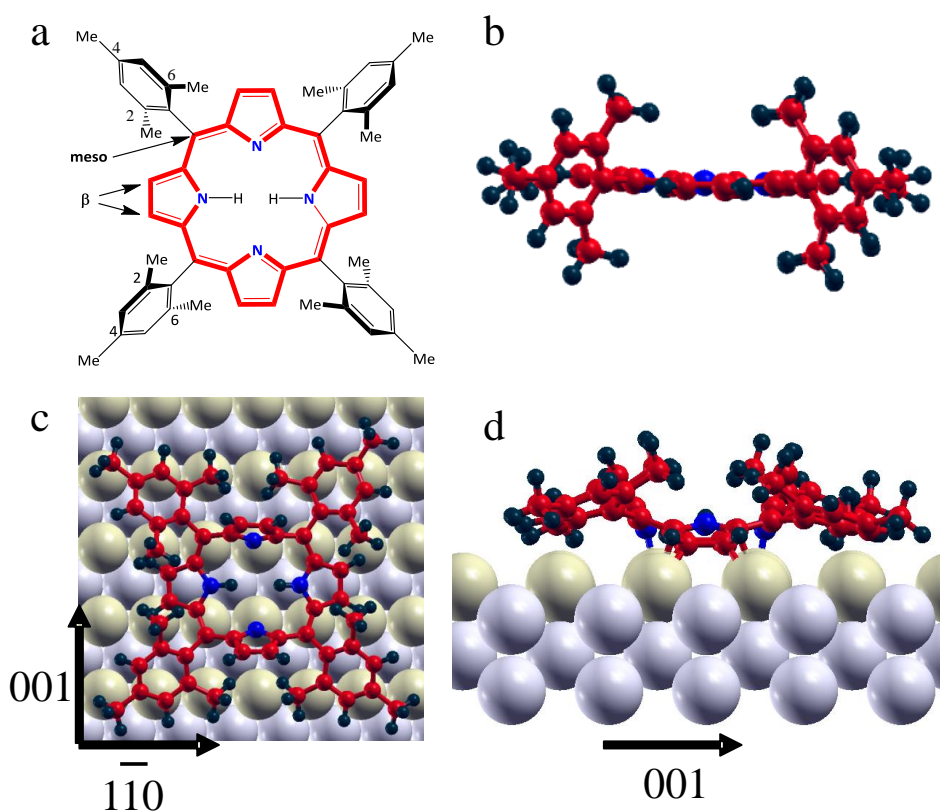


Figure 1. (a) Chemical structure of TMTPP. (b) DFT relaxed configuration in the gas-phase (side view) with a flat core and phenyl rings oriented almost perpendicular to it. Top (c) and side (d) views of the energetically most favorable TMTPP adsorption geometry on Cu(110) calculated by DFT showing the alternated orientation of the central pyrrolic groups and the configuration of the phenyl rings. Red, blue and black spheres correspond to C, N and H atoms, respectively. Cu atoms in the top-most close-packed rows are shown as green, while lower lying Cu atoms are grey.

Experimental STM images of isolated molecules obtained after adsorption on the Cu(110) surface at 300 K show a rectangular symmetry with two pronounced arc-shaped structures running along the $[1\bar{1}0]$ close-packed Cu row direction, Fig. 2(a). All observed molecules share the same sub-molecular features and a unique orientation relative to the substrate. The images in Fig. 2(a) show the molecules imaged with bright intense lobes around the center of the molecule, arising from the porphyrin core. There is an additional intensity associated with the methyl groups, which appears in the form of winged lobes (legs). Their orientation with respect to the $\langle 110 \rangle$ Cu row direction is found to be exclusively perpendicular (the other orientation was not observed), and suggests a unique orientation of the molecule on the surface. In order to verify the predicted adsorption geometry of Fig. 1(c,d), STM images were simulated for the most stable and a number of less energetically favorable relaxed geometries [27,28]. The simulated STM images were found to be very sensitive to the TMTPP orientation and structure (Figs. 2(b,c) and Figs. S1-4). Importantly, only the geometry corresponding to the most stable structure of Figs. 1(c,d),

with its alternate pyrrole geometries and non-planar phenyl ring orientation, provides good agreement with the experimental STM images, particularly with respect to the orientation of the winged lobes (legs) relative to the $\langle 110 \rangle$ Cu rows, as can be seen by comparing the right panel in Fig. 2(a) with those in Figs. 2(b,c).

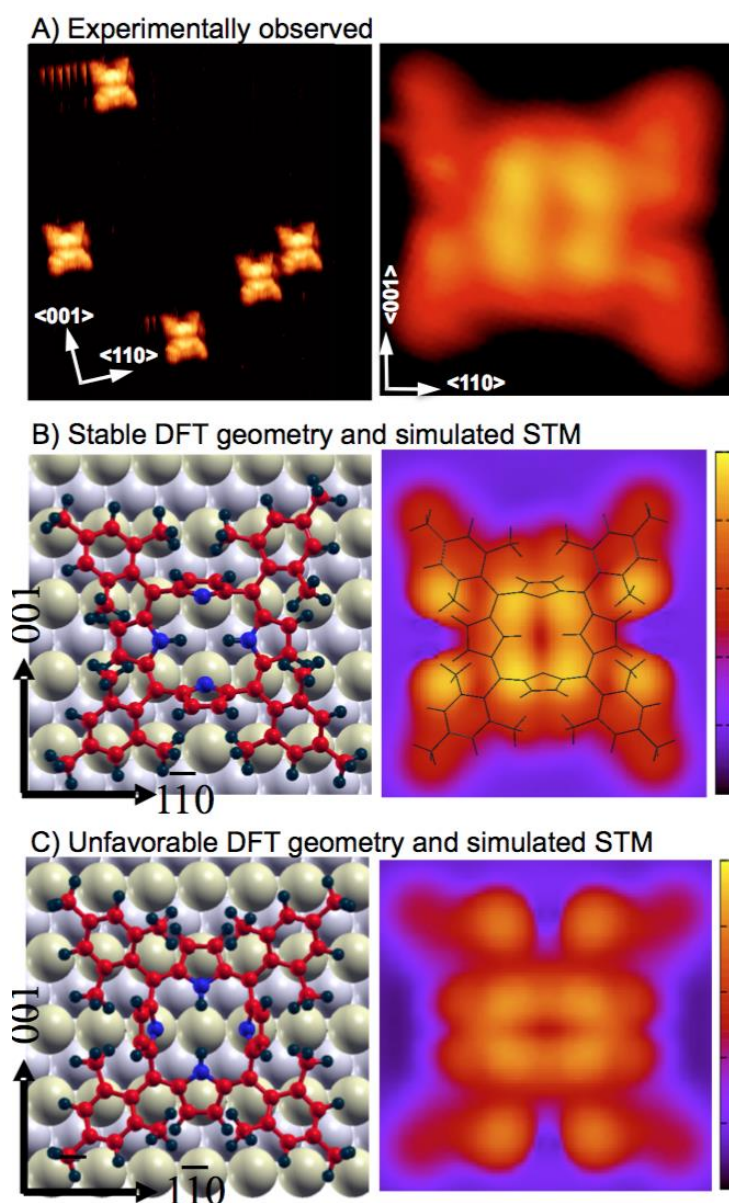


Figure 2. (a) TMTTP molecule on the Cu(110) surface observed in our STM experiments. Left panel: a large-scale image (area $94 \times 94 \text{ \AA}^2$, $V_t = 0.575 \text{ V}$, $I_t = 0.34 \text{ nA}$); right panel: a high-resolution image of a single adsorbed molecule (19 \AA^2 , -1.03 V , 0.1 nA). (b) The theoretically simulated geometry and the STM image in the most stable geometry. On the right panel, the molecular structure is superimposed on the image to guide the eye. The horizontal arc-shaped protrusions correspond to the vertical pyrroles and the 2-methyl groups, which are the highest molecular chemical groups in the adsorbed molecule. (c) Same as (b), but for an unfavourable geometry, with the STM simulated image in disagreement with the experiment.

2.2 Surface Driven Inter-molecular Coupling

Experimental STM data show that highly directional macromolecular patterns are formed when TMTPP is adsorbed on Cu(110) and the system annealed to 575 K [12]. Fig 3a shows high-resolution STM data obtained for part of the complex mixture of covalently linked structures formed by this system. The evolution of H₂ gas was observed between 450-520 K (Fig. 3b), concomitant with the pattern formation, indicating that de-H processes accompany the inter-molecular bonding associated with the creation of macromolecular structures. TMTPP contains a number of H-containing groups, so three main questions need to be addressed to understand the pattern of reactivity displayed by the system: i) which H atoms of the molecule are most prone to the de-H processes; ii) what are the diffusion patterns of a single molecule on the surface; and, (iii) what is the mechanism of the intermolecular bond formation at the surface. Each aspect is considered below.

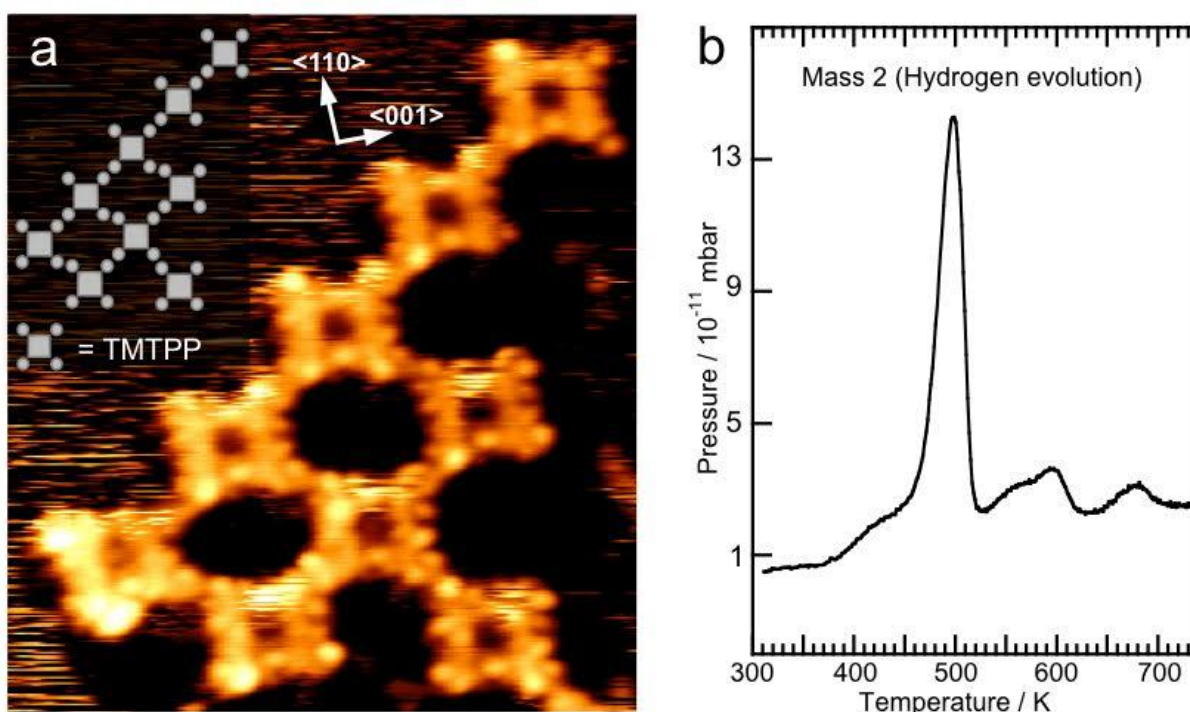


Figure 3. a) An experimental STM image of a covalently bonded network of TMTPP molecules following heating to 600 K, ($65 \times 73 \text{ \AA}^2$, 0.236 V, 0.35 nA). The inset shows a pictorial representation of the networked structure imaged. b) Temperature Programmed Desorption data showing the evolution of H₂ during the coupling process. STM images of larger areas are given in the SI (Fig. SI-11).

Selective De-hydrogenation Processes: Using periodic DFT calculations, we identified the de-H processes that are energetically most favorable, and, thereby, essentially decide the molecular positions that become available for intermolecular covalent bond formation. We start by evaluating the removal energies of each H from the molecule in the gas-phase,

$\Delta E_{rem}(gas)$, Fig. 4(a). Each $\Delta E_{rem}(gas)$ is calculated as the energy difference [29] between the de-H gas phase molecule and the fully hydrogenated (f-H) gas phase configuration of Fig. 1(b). The reaction is endothermic, and a hierarchy of C-H bond breaking energies is obtained as shown in Fig. 4(a). Specifically, the most favorable hydrogen atoms to remove belong to the 4-, 2- and 6-methyl groups and to the N-H groups in the central core [30]. These energies reflect the bonding properties of each hydrogen atom within the gas phase molecule.

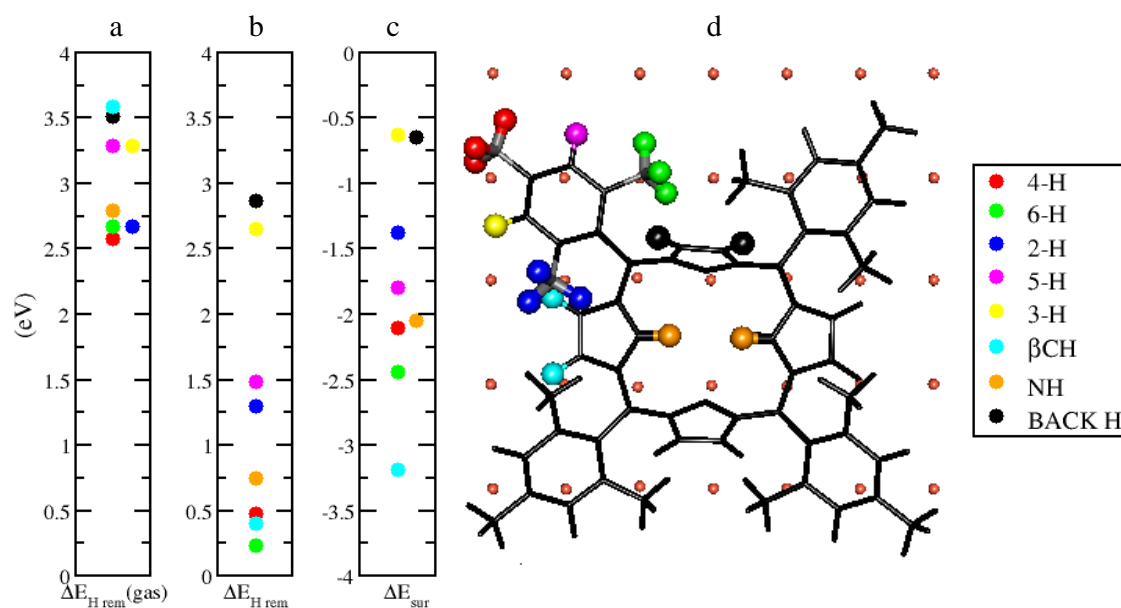


Figure 4. Removal energies of H atoms from non-equivalent positions in TMTPP: (a) in the gas-phase, $\Delta E_{rem}(gas)$, and (b) adsorbed on Cu(110), ΔE_{rem} ; (c) the influence of the in each de-H process is quantified by $\Delta E_{sur} = \Delta E_{rem} - \Delta E_{rem}(gas)$. The different types of H atoms are shown in (d) in the molecular structure by the corresponding color code. The removal energies shown are equivalent to dissociation energies; note, the final position of the removed H atoms is on the surface next to the molecule. Although energy barriers for the de-H reaction provide a better indicator of the ability to de-hydrogenate, a systematic calculation of the energy barriers for all non-equivalent H atoms was performed only for the hydrogen atoms from the 4-methyl groups (see Fig. 6a). However, it is obvious that the barriers cannot be smaller than the energy difference ΔE_{rem} between the initial and final states, hence ΔE_{rem} serves as a realistic guide of the ease with which H atoms can be removed from various positions in the molecule. Note that H atoms belonging to a specific group (e.g. labeled red or green), are inequivalent on the surface. However, their removal energies were found to be the same due to rotational flexibility of the corresponding group, enabling them to rearrange to the same final configuration.

To understand the role played by the surface, we calculated the corresponding energy differences ΔE_{rem} , between the fully relaxed de-H and f-H configurations on the surface, for all non-equivalent H atoms, Fig. 4 (b). For the 4-methyl group, we also calculated the de-hydrogenation energy barrier (shown in Fig. 8, left panel). The removed hydrogen atoms are adsorbed nearby on the surface in the most stable positions bridging two Cu atoms in a row. First, we note that ΔE_{rem} values are significantly lower than the equivalent $\Delta E_{rem}(gas)$ values. Second, the hierarchy of C-H bond breaking is altered significantly from the gas phase system and the "easiest" H atoms to remove belong to the 6- and 4-methyl groups and to the horizontal pyrrole C-H groups (β CH), all of which lie very close to the Cu surface. This is due to a strong and selective reactivity effect of the surface on these specific de-H reactions. Thus, the 6- and 4-methyl C-H bonds remain the easiest to break *both* due to their specific chemistry in the molecule *and* the activating effect of the surface. Additionally, the horizontal pyrrole C-H groups transform from being the hardest to dissociate in the gas phase to one of the easiest in the adsorbed state due to their proximity to the surface. In contrast, the 2-methyl group becomes relatively harder to break at the surface, as intra-molecular interactions orientate the C-H bonds away from the surface plane. We quantify this surface effect in Fig. 4(c) by computing $\Delta E_{sur} = \Delta E_{rem} - \Delta E_{rem}(gas)$, where the lowering of the energy barrier to de-hydrogenation is greater for more negative values of ΔE_{sur} .

This leads to the conclusion that in terms of de-H processes, there are three candidate positions on the molecule, namely the 6- and 4-methyl and the β CH groups of the horizontal pyrrole, that are important to consider as potential intermolecular linking points for a surface-bound system. Therefore, the following intermolecular connections need to be considered: 4-methyl-4-methyl; 4-methyl-6-methyl; 4-methyl- β CH; β CH- β CH; 6-methyl-6-methyl and 6-methyl- β CH. Of these, only the first two are sterically allowed, with the experimentally observed final product showing a clear preference for 4-methyl-4-methyl connections where the linked molecules have a diagonal juxtaposition. In order to understand this clear preference, we need to consider other factors such as molecular diffusion and intermolecular bond formation that are important in guiding the covalent assembly. The molecule's mobility on the surface is also an essential ingredient for understanding their assembly, as the most probable diffusion patterns may dictate the most likely relative arrangements of connecting molecules and hence bonding topographies. Using the Nudged Elastic Band (NEB) method [31], we calculated the energy barriers for single molecule diffusion along and across the Cu rows. We find that TMTTP diffuses on the surface by sliding preferentially along the close-packed rows, Fig. 5(a), where the energy barrier of ~ 1.3 eV is almost half of that across the rows, ~ 2.5 eV, Fig. 5(b). This anisotropy in the diffusion pattern reflects the dissymmetry in surface corrugation in the two main directions of the surface.

Since covalent bonding between molecules may proceed after the de-H reactions, it is important to understand the mobility of de-H molecules as well. Hence, we simulated the diffusion of a de-H molecule with one H atom removed at the 4-methyl group of the phenyl ring, which corresponds to the experimentally observed connection. Interestingly, we find that the anisotropy of diffusion is enhanced upon de-hydrogenation, with the barrier along the rows remaining essentially unchanged, but increasing substantially by ~ 0.5 eV across the rows. This effect is attributed to the fact the de-H molecule forms a C-Cu bond between the de-hydrogenated C atom and the nearest Cu atom on the surface, which can easily translate from one Cu atom to the next when diffusion occurs along the close-packed rows, while this

is more difficult across the rows due to the larger Cu-Cu distance requiring the C-Cu bond to be completely broken in the transition. We believe that the asymmetry in diffusion we find is general and does not depend on which particular H atom was removed; moreover, we expect that values of the barriers must be not very sensitive to the position of the removed H atom.

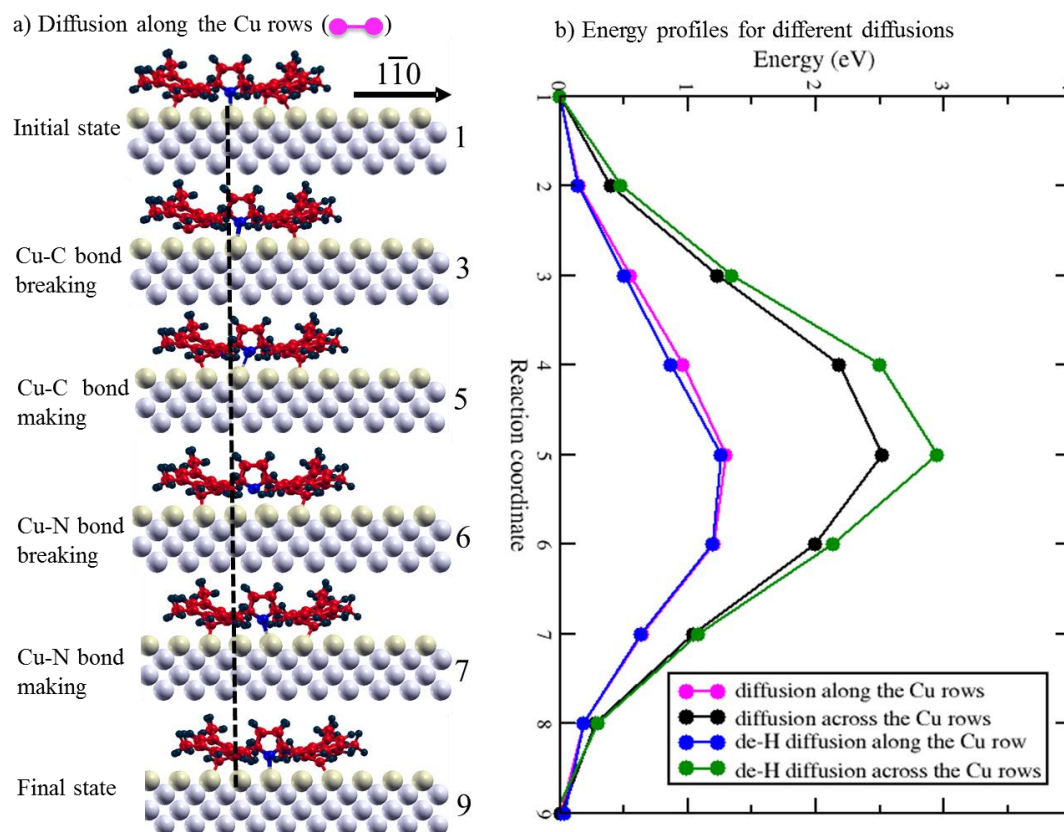


Figure 5. Minimum diffusion paths of a single molecule diffusing on Cu(110) between two stable equivalent configurations (reaction coordinates 1 and 9). (a) A selection of atomic geometries along the path for the fully hydrogenated molecule along the close-packed Cu rows. The black dashed line is a guide for the eye to highlight the diffusion step. (b) Energy profiles along (magenta) and across (black) the Cu rows. The very different energy barriers highlight a strongly anisotropic diffusion, clearly favorable along the rows. Blue and green curves are associated with the diffusion of a de-hydrogenated molecule along and across the rows, respectively, showing an increased anisotropy (for detail see Supporting Information).

iii) Inter-molecular Bonding Configurations: We shall now consider two closely positioned molecules on the surface in a number of geometries that are compatible with the favorable de-H processes, diffusion along the rows and sterically allowed covalent products as identified above. Fig. 6(a) shows the relaxed configurations of well separated and un-

reacted de-H molecules, with the removed H atoms bonded to the free surface in their most stable position some distance away. The relaxed configuration of two separated fully hydrogenated molecules was also computed (Fig. SI-9). We now consider the 4-methyl-4-methyl (4Me-4Me) coupling product, which arises when the de-H 4-methyl groups at the corresponding molecular corners come into contact, forming a covalent C-C bond. This product can actually be accommodated at the surface in a number of ways. The two most stable and essentially degenerate diagonal arrangements are shown in Fig. 6(b,c), which differ by a small change in relative positions of the TMTTP components as indicated by the core-to-core surface vectors of (5,4) and (6,3). In both bonded structures, the TMTTP molecules have a configuration and orientation similar to that of the most stable geometry of a single TMTTP molecule on the surface (Fig. 1c). Both products are more stable than two unreacted de-H molecules by 0.58 eV, which means that upon de-hydrogenation two (or more) approaching molecules are energetically driven to bond. The covalently linked (5,4) and (6,3) accommodated products exhibit a *trans* conformation of the interconnecting 1,2-ethylene group with an inter-core distance of ~ 19.4 Å and 19.0 Å, respectively. These connections also lead to slight offsets between the diagonals of the two molecules [32]. Several 4Me-4Me products accommodated in a (4,4) configuration were also calculated (Fig. SI-10) with only one energetically driven to bond, with one molecule having a slightly rotated configuration (Fig. 6e). However, this geometry is 0.34 eV less energetically favorable than the (5,4) and (6,3) accommodated products.

We also investigated the role of Cu adatoms as possible mediators in organometallic C-Cu-C bonds, as has been observed for other porphyrins at the Cu(110) surface [2,20,21]. A single Cu adatom was placed between the de-H molecules, starting from the relevant configurations in Figs. 6(b,c,d). The relaxed structures obtained (Figs. 6(g,h)) were between 0.79 eV and 1.18 eV higher in energy than the ones in Figs. 6(b,c), hence deemed to be substantially less favorable.

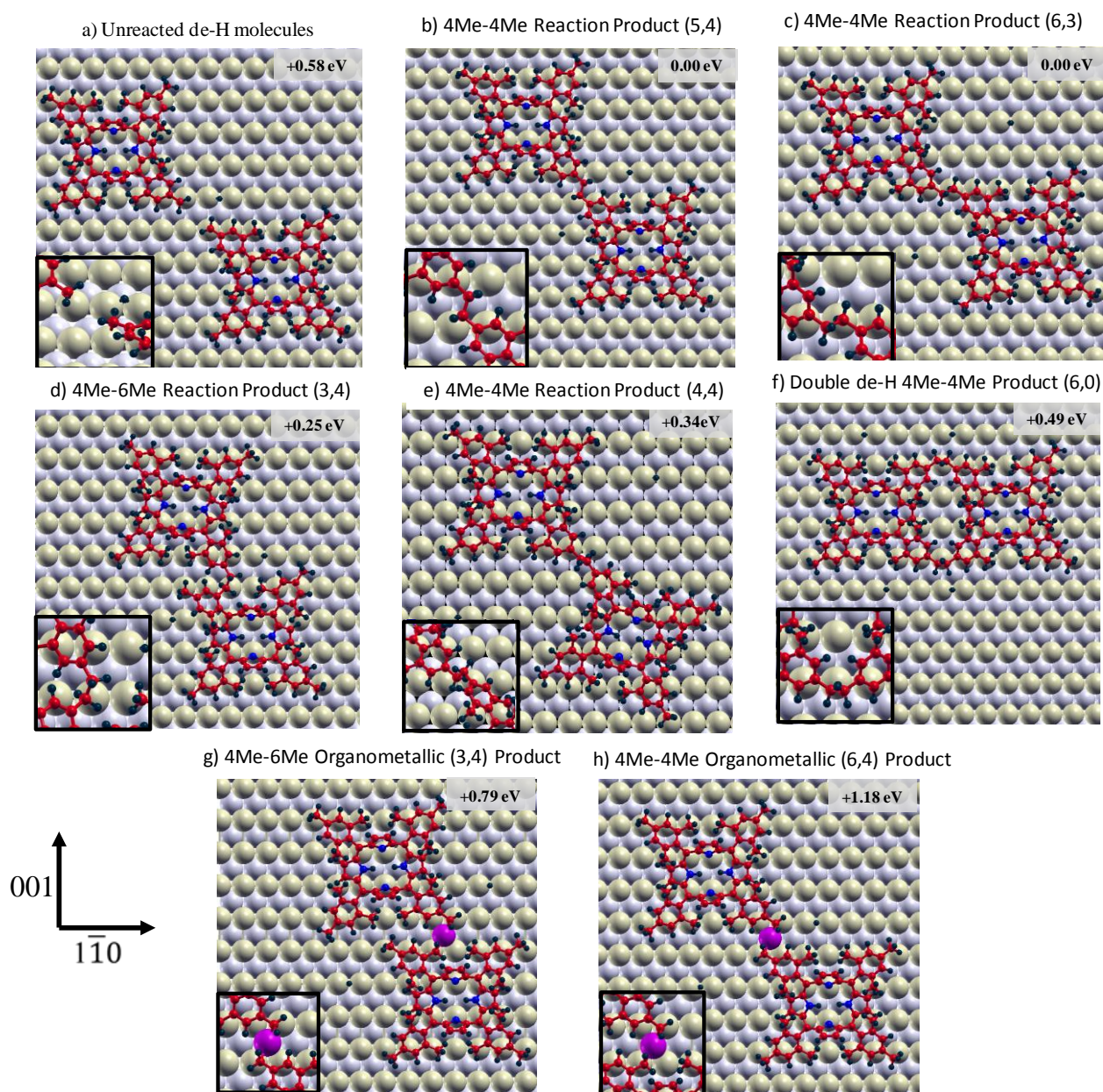


Figure 6. Calculated two molecule structures with their energies given relative to those in structures (b,c). (a) Separate and un-reacted de-H molecules on the surface. The removed hydrogens are visible some distance away from the molecules in their most stable adsorption geometry on the free surface at $T=0$ K. (b,c) The most stable geometries of de-H and 4Me-4Me bonded TMTTP molecules on Cu(110), connected via the peripheral 4-methyl groups in their functional phenyl groups. (e) Another configuration having the same intermolecular coupling as in (b,c), but with a different arrangement on the surface. (d,f) Differently bonded configurations compatible with steric constraints and an ability of a horizontal diffusion, but energetically less stable than the ones in panels (b,c). (g,h) Organometallic-coupled structures, with a Cu adatom (in magenta) mediating the interaction. The insets show the intermolecular bonds in each case.

We note that in our DFT calculations the fully hydrogenated structure of two molecules lies lower in energy than the de-H one by 1.1 eV. Furthermore, the energy of a single H₂ molecule above the surface is less favorable by ~0.3 eV than that of two well separated single H atoms adsorbed on it (Fig. SI-8). These results imply that both the de-H process and the recombination of H atoms in the gas phase subsequent to their removal from the surface are not feasible at T=0 K. This is in full agreement with the experiment where the system needs to be heated to over 400 K in order to initiate the de-H process and observe hydrogen gas evaporation from the surface.

In order to rationalize these DFT results, one has to compare the *free energies* of hydrogenated and de-H molecules on the surface as a function of temperature (see SI, Section 4). In the de-H case, it is essential to take into account the presence of the H₂ gas above the surface, assumed here to be in thermodynamic equilibrium with the H atoms either adsorbed on the substrate or attached to the molecules. Assuming that the vibrational contribution to the free energy due to H atoms on the surface in all relevant systems is approximately the same, the main contribution to the free energy difference ΔF will come from the difference in DFT energies *and* from the hydrogen gas free energy. The latter provides an important entropic contribution to ΔF , making the total free energy of de-H processes favorable at elevated temperatures. Indeed, for the relevant range of T=500-600 K we estimate that approximately between -1.8 and -2.2 eV contribution comes from the H₂ gas evaporated into the ultra-high vacuum chamber at a pressure of 10⁻⁷ Pa, which is sufficient to reverse the trend of total energies calculated by DFT for the hydrogenated and de-H molecules (see SI for details).

The ΔF gain due to H₂ gas is so significant that it can also facilitate de-hydrogenation processes from *other* molecular sites with relatively low removal energies - like the 6-methyl groups and the β CH groups - and hence these events cannot be completely excluded. Assuming diffusion along the Cu rows and a single de-H per molecule, the only sterically possible bond resulting from these processes is the connection between 4- and 6- methyl groups (4Me-6Me), Fig. 6(d), accommodated with a core-to-core surface vector of (3,4). It is also worth considering bonded configurations involving *two* de-Hs per molecule and still compatible with diffusion along the rows. One such possibility leads to a horizontal molecular chain, with two adjacent molecules connected by two bonds involving four 4-methyl groups, which has a *cis* conformation of the interconnecting 1,2-ethylene group and a core-to-core surface vector of (6, 0), Fig. 6(f). Additionally, the organometallic products shown in Figs. 6(g,h) might become accessible. However, although ΔF considerations should allow these structures to form, in principle, they are less stable than the 4-methyl-4-methyl products in Figs. 6(b,c) [by 0.25 eV (d), 0.49 eV (f), 0.79 eV (g) and 1.18 eV (h)] and the diffusion barriers to reach them are much higher, as discussed in the next Section. We would, therefore, expect these to be minority products.

In order to establish the types of products created in the experiments, high-resolution STM data were obtained, which allowed both the macromolecular products and the underlying Cu surface atoms to be imaged and core-to-core surface vectors established. Figure 7(c,d) show examples of the 4Me-4Me (5,4) and (6,3) reaction products at the surface. A histogram of the distribution of products with the core-to-core vectors measured

from our high-resolution data is shown in Fig. 7 (right panel). The data are obtained only when advantageous imaging conditions arise, hence, this represents a small subset of all data collected, and detailed statistical analysis is not possible. Nevertheless, it can be seen that almost 75% of the products possess the 4-methyl-4-methyl (5,4) and (6,3) configurations, which are predicted by theory to be the most favoured, and 10% possess the 4-methyl-4-methyl (4,4) configuration (predicted by DFT to be less stable), with good agreement between the measured and calculated inter-molecular distances. The minority structures observed correspond to the energetically less favoured calculated configurations, with the horizontal chains along the Cu rows observed extremely rarely.

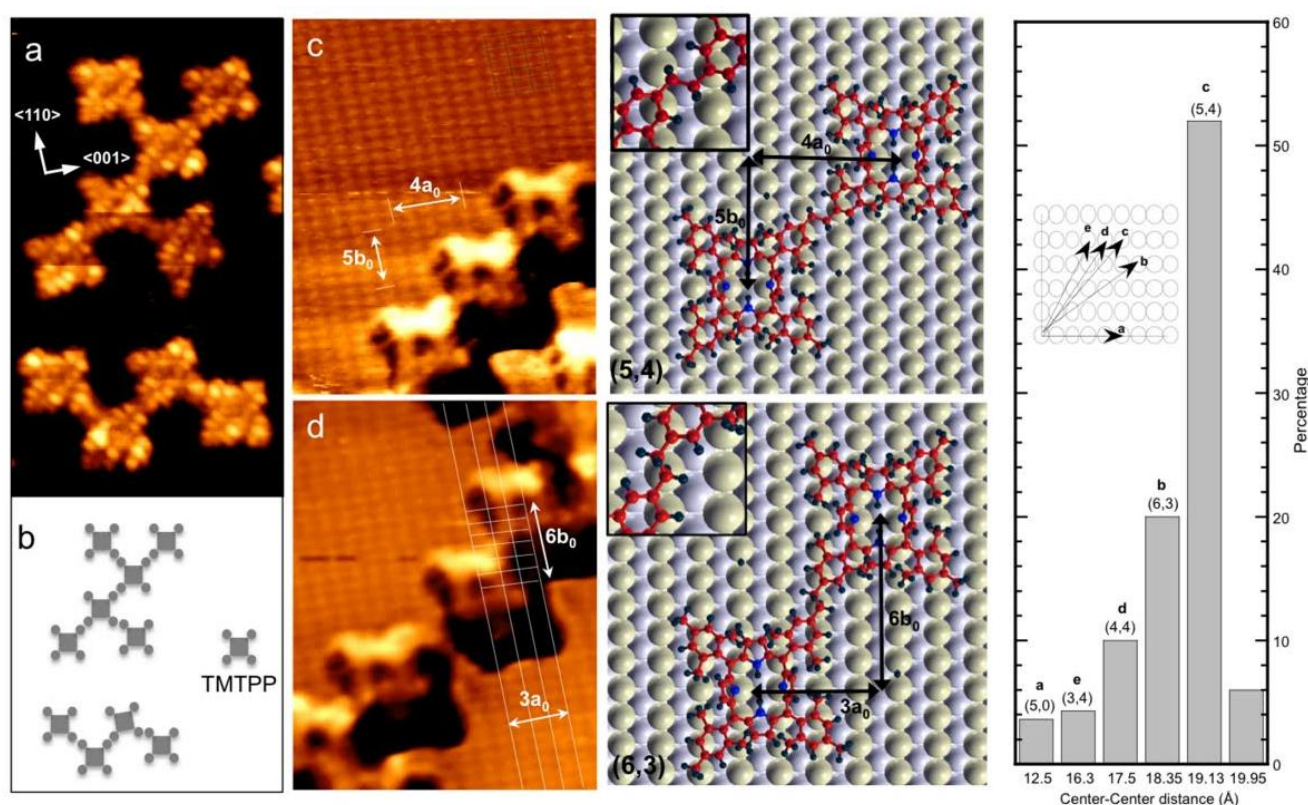


Figure 7. High resolution STM images showing the connectivity of the reacted TMTPP described in the DFT calculation. (a) The imaged orbital structure is shown ($60 \times 100 \text{ \AA}^2$, 0.284 V, 0.34 nA). (b) A pictorial representation of the networked structure imaged in (a). (c) and (d) STM images where both the Cu surface atoms and reacted products are imaged, from which the relative locations of the central cores can be measured (figures on the right of the STM images show the DFT calculated models of these). (c) Three reacted molecules with a (5,4) accommodated configuration are observed ($49 \times 60 \text{ \AA}^2$, 0.311 V, 0.32 nA). (d) Shows molecules arranged in (6,3) accommodated configuration ($58 \times 68 \text{ \AA}^2$ ($I_t = 0.37 \text{ nA}$, $V = 0.311 \text{ V}$)). The histogram on the right shows center-center molecular distances measured from STM data, compiled from 140 separate connection counts, where both the substrate atoms and the molecular reaction products could be simultaneously imaged. From the data we calculate an experimental error of $\pm 2\%$ in the measured bond lengths from the exact values expected for each structure.

2.3 Simulating the Overall mechanism of inter-molecular coupling

We now have all the ingredients needed for simulating the bonding process itself. We analyze this by means of a sequence of NEB calculations involving two molecules diffusing on the substrate towards each other, with subsequent de-hydrogenation at the facing corner sites and then bonding together. In all our simulations the molecules are initially fully hydrogenated and placed reasonably far apart in their stable configurations. As the final geometry, we considered one of the two most favorable de-H bonded pair shown in Fig. 6(b), with the two removed H atoms placed well away from the molecules on the free surface.

As described above, single molecule NEB calculations indicate that TMTTP will diffuse mainly along the $[1\bar{1}0]$ rows. Still, several different scenarios are conceivable depending on the order in which de-H and diffusion processes happen prior to the bonding. In the simplest case, all elementary processes happen “independently”: de-H of the first molecule (M1), de-H of the second (M2), their diffusion along the rows and, finally, bonding. Other more exotic mechanisms, in which some of the events are concomitant, may also be envisaged. For instance, the de-H of M1 may facilitate the de-H of M2, i.e. the first could catalyze the second. We, therefore, also calculated “simultaneous” and “asynchronous” de-H reactions concomitant with the diffusion (SI, Section 5). The comparison of the calculated energy barriers predicts that the “independent” scenario described above is by far the most favorable, while the one with simultaneous de-H events is the least. Other conceivable processes, e.g. where the second de-H happens slightly after the first one, lie in between. In Fig. 8, we show the main steps for the “independent” scenario, namely the de-H reaction for one molecule (a), and the combined diffusion-bonding process of the two de-H molecules (b). This scenario has the effect of splitting the whole process into several elementary steps, each having a low energy barrier. Obviously, when two C-H bonds are broken at the same (or at slightly different) time(s) during the diffusion, the corresponding barriers simply add up, which significantly decreases the rate of the whole process as compared with the “independent” mechanism. The same line of reasoning also explains the reduction in diffusion of oligomeric structures as observed by time-resolved STM experiments [12]; this is due to an increased number of N-Cu and C-Cu molecule-surface bonds that have to be broken concomitantly in order for these larger structures to become mobile on the surface. This also has implications for the growth of the covalent networks in that once a few molecules are bonded together, the ensemble becomes largely stationary and it is the monomer species that have to diffuse to react with it.

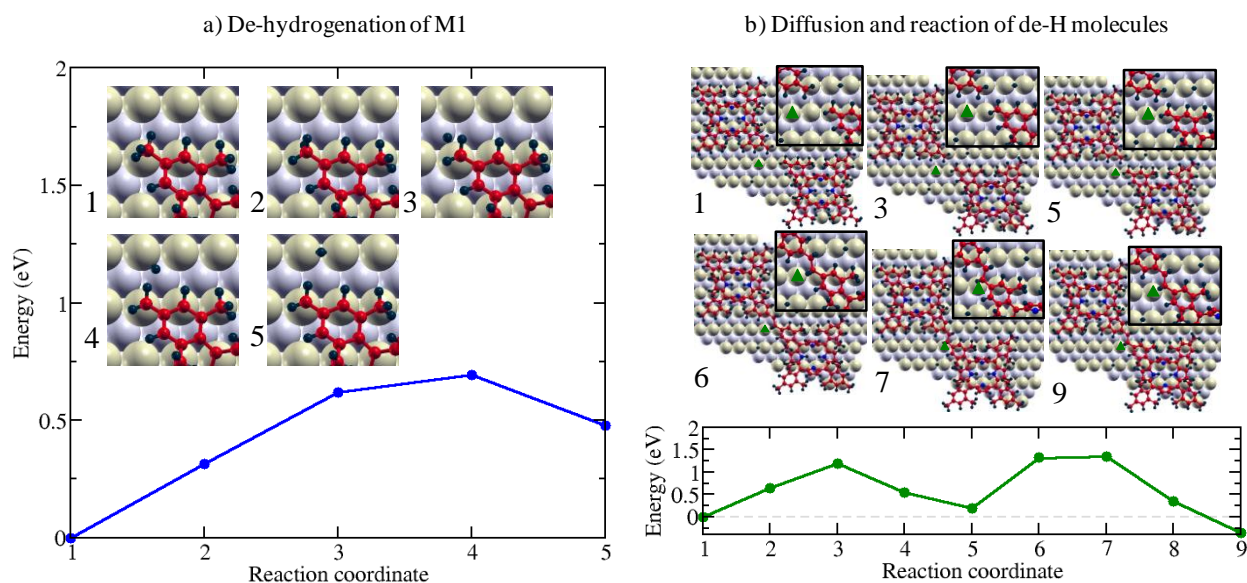


Figure 8. Minimum energy profile and reaction path bringing two hydrogenated molecules, M1 and M2, placed initially far apart on the surface as in Fig. 6(a), to the bonded configuration as in Fig. 6(b) via the mechanism whereby the two de-hydrogenations, diffusion and bonding happen independently one after another as described in the text. (a) De-hydrogenation reaction involving a 4-methyl hydrogen atom. The initial state (reaction coordinate $R_c=1$) is the stable configuration of the intact, fully hydrogenated molecule (reference energy 0 eV). In the final state at $R_c=5$, M1 is de-hydrogenated, with the removed hydrogen atom placed nearby on the free surface. The reaction path shows how the H atom avoids passing through a higher energy barrier in the central hollow position between 4 top-most Cu atoms. Note that ΔE_{rem} is the difference between the final and initial states. The minimum energy profile for de-H of M2 is the same. (b) Minimum energy profile and reaction path corresponding to the diffusion of M1 along the Cu rows towards M2 followed by the diagonal bonding between two 4-methyl groups of both molecules. The insets highlight the bonding region. At $R_c=1$ the molecules are non-bonded and de-hydrogenated. The peak at $R_c=3$ stems from the bond breaking with the surface during the M1 diffusion along the row, which at the minimum ($R_c=5$) reaches the next equilibrium position. After a further diffusion and initial interaction with M2 ($R_c=6,7$), the two molecules eventually connect ($R_c=7-9$). The energy gain of the final bonded configuration at $R_c=9$ relative to the initial state ($R_c=1$) of 0.58 eV indicates that two previously de-hydrogenated molecules are driven to connect.

Thus, we are now able to give a detailed characterization of the bonding mechanism steering the observed covalent assembly of TMTTP molecules on the copper surface. At a high enough temperature de-hydrogenation processes become favorable at specific H sites of the molecules, and these reactions are strongly activated by the substrate. These processes happen independently for different molecules. Hydrogenated and de-hydrogenated molecules diffuse mainly along the rows with comparable mobilities. When de-hydrogenated molecules approach each other along the rows with the phenyl groups capable of making a contact, they are energetically and sterically driven to “connect”

through the dehydrogenated $-CH_2$ 4-methyl groups at the corners giving a unique orientation to the molecular chains thus formed. The stable bonded configuration in each connected pair has a *trans* 1,2-ethylene unit between the porphyrins, with the coupled molecules retaining a similar orientation as for the unreacted molecule. If two de-H methyl groups are formed at the diagonally opposite sides of the same molecule, 1D chain formation may be initiated; if at least two of such groups are formed in the perpendicular or horizontal directions, this leads to zig-zag structures and the growth of 2D islands [12]. Other covalent assemblies are significantly disfavoured on the basis of energy, steric and diffusion grounds, and are only rarely observed. For instance, for molecules to be arranged along the Cu rows, four de-H processes need to happen for *two* 4-methyl groups of *each* molecule, which is a low probability event. Even if we assume that the two molecules are already appropriately de-hydrogenated, the estimated energy barrier to connect two doubly de-H molecules approaching horizontally along the same rows is found to be more than 1.5 eV higher (Fig. SI-7) than for the most favourable diagonal connection, where molecules diffuse along adjoining rows and, furthermore, yields a less stable structure.

3. Conclusions

In this work we employed *ab initio* theory and STM experiments to study the coupling reactions of tetra(mesityl)porphyrins (TMTPP) adsorbed on the Cu(110) surface. Upon annealing, diagonally oriented covalently bound nanostructures are formed with unique bond directionality. The covalent bonds between molecules are initiated by activation and scission of *selected* C-H bonds, which leads to the formation of specific and strong C-C intermolecular connections. Functionalized organic molecules of large size have an abundant number of peripheral C-H bonds, all in principle available for activation, thus providing an attractive 'synthon' for coupling strategies. The main and generic question we have addressed in this work is why only specific C-H bonds are at play in the TMTPP/Cu(110) system leading to highly selective molecular patterns. Using density functional theory, nudged elastic band methods and appropriate entropic considerations, we provide a detailed explanation of this *bond selectivity* and of the bonding mechanisms. The selection of the corner 4-methyl groups as activation and binding sites is the result of the interplay of several factors including intrinsic molecular chemistry, adsorption energetics, the selective catalytic effect of the surface, steric effects and asymmetric diffusion of the molecules on the surface. Entropic effects are also an essential driving force in leading to covalently bound structures at high annealing temperatures.

Growing complex, covalent surface networks in a controlled manner from large organic building blocks represents a real challenge in surface molecular nanoscience. Using selective C-H bond activation is a very promising route in this direction, allowing a diverse range of organic building blocks to be used directly. Our study, albeit on a specific system, provides important insights on the various factors and the underlying driving mechanisms at a surface that affect selective C-H bond scission and specific C-C intermolecular bonding. In particular, we have established that: (i) the *reactivity* of the molecule to the surface is *at the selected site(s)* and how it accommodates giving specific geometry of the corresponding group(s) are essential, since binding of the latter to the surface reduces the energy of the system and hence of the H removal energy barrier; (ii) there must be a correlation between

the anisotropic molecular diffusion on the surface and the more labile hydrogen atoms of the molecule selected for the H removal, otherwise the molecules may be sterically hindered to binding. Such knowledge is imperative for establishing a more complete set of future design rules for controlled covalent assembly at surfaces.

Methods

Computational Details: Density functional theory (DFT) calculations were performed with the Quickstep code [33] within the CP2K package [34], using a mixed Gaussian and plane waves basis set, the Goedecker, Teter and Hutter (GTH) pseudo-potentials [35] and a GGA-PBE [36] + rVV10 [37] exchange-correlation functional including self-consistently the van der Waals (vdW) interaction. Preliminary calculations made use also of the Grimme D2 functional [39]. We used a plane-wave basis energy cut-off of 400 Ry and the Γ point to sample the Brillouin-zone. The Cu(110) substrate was modeled with a periodically repeated slab of four layers, allowing a vacuum gap between the adsorbed molecule and the bottom layer of the slab above it of ~ 7 Å. Relaxations of two molecule system were performed with two layers and were considered completed when atomic forces reached 0.02 eV/Å. Only forces acting on atoms belonging to the two (or one) uppermost top layers and the molecule were used. Nudged Elastic Band (NEB) [31] calculations for single molecule diffusion in Fig. 5 were performed using nine replicas, including initial and final states. Fig. 8 included five replicas per each individual process. When calculating H removal energies, H atoms were considered for the energy balance as a part of the final systems. Calculated STM images were obtained by calculating the integrated local density of states (ILDOS) within the Tersoff-Hamann method [27] using the plane-wave-pseudo-potential package Quantum-ESPRESSO [38]. The constant current STM images were simulated using the LEV00 package [28].

Experimental Details: STM experiments were performed under ultra-high vacuum conditions using a Specs STM 150 Aarhus instrument. The STM was calibrated by measuring the atomic distances of the clean Cu(110) surface, All measurements were taken in constant current mode, using a tungsten tip and at a base pressure of 1.5×10^{-10} mbar. Bias voltages are measured at the sample. The Cu(110) surface was prepared in a UHV chamber using Argon ion sputtering and annealing cycles, and atomic flatness and cleanliness were checked by STM prior to dosing the molecule. Tetra- (2,4,6-tri-methyl-phenyl)-porphyrin (TMTPP) (Frontier Scientific) was used as purchased and sublimed onto the Cu(110) surface, which was held at room temperature during initial deposition.

Acknowledgments

Via our membership of the UK's HPC Materials Chemistry Consortium, which is funded by EPSRC (EP/F067496, EP/L000202), this work made use of the facilities of HECTOR and ARCHER, the UK's national high-performance computing service, which is funded by the Office of Science and Technology through EPSRC's High End Computing Programme. AF and LK thankfully acknowledge the computer resources, technical expertise and assistance provided by the Red Española de Supercomputación. AF thanks Matthew Dyer for

discussions. DA thanks EPSRC for funding. SH and RR thank the UK EPSRC (EP/J019844/1 and EP/F00981X/1) for funding.

Supporting Information available: Molecular structures on the Cu(110) surface. Role of the functional groups in the TMTTP stability on the Cu(110) surface. Simulated STM images of a single molecule in several configurations on the surface. Free energy comparison of hydrogenated and de-hydrogenated structures at finite temperatures. De-hydrogenation mechanisms: simultaneous, asynchronous and independent processes. Horizontal bonding of two molecules leading to higher energy barriers. Stability of hydrogen on Cu(110) at T=0K. Unreacted, fully hydrogenated molecules. Extra 4-methyl-4-methyl coupled structures.

References

- [1] L. Grill; M. Dyer; L. Lafferentz; M. Persson; M. V. Peters; and S. Hecht. *Nature Nanotech.* 2007, 2, 687.
- [2] S. Haq; F. Hanke; J. Sharp; M. Persson; D. B. Amabilino and R Raval. *ACS Nano* 2014, 8, 8856.
- [3] Ludlow; R. F.; Otto; S. Systems Chemistry. *Chem. Soc. Rev.* 2008, 37, 101–108.
- [4] E. C. P. Smits; S. G. J. Mathijssen; P. A. van Hal; S. Setayesh; T. C. T. Geuns; K. A. H. A. Mutsaers; E. Cantatore; H. J. Wondergem; O. Werzer; R. Resel; M. Kemerink; S. Kirchmeyer; A. M. Muzafarov; S. A. Ponomarenko; B. de Boer; P. W. M. Blom; D. M. de Leeuw. *Nature* 2008, 455, 956–959.
- [5] Aida T.; Meijer E. W.; Stupp S. I.. *Science* 2012, 335, 813–817.
- [6] Koepf M.; Cherioux F.; Wytko J. A. Weiss J.. *Coord. Chem. Rev.* 2012, 256, 2872–2892.
- [7] L. Lafferentz; V. Eberhardt; C. Dri; C. Africh; G. Comelli; F. Esch; S. Hecht; L. Grill. *Nature Chem.* 2012, 4, 215.
- [8] J Cai; P Ruffieux; R Jaafar; M Bieri; T Braun; S Blankenburg; M Muoth; Ari P. Seitsonen; M Saleh; X Feng; K Müllen; R Fasel. *Nature* 2010, 466, 470–473.
- [9] Lee J; Dougherty DB; Yates JT Jr.. *J Am Chem Soc.* 2006, 128, 6008-9.
- [10] M Treier; C A Pignedoli; T Laino; R Rieger; K Müllen; D. Passerone ; R. Fasel. *Nature Chem.* 2010, 3, 61–67.
- [11] Matena M.; Riehm T.; Stöhr M.; Jung, Thomas A.; Gade Lutz H.. *Angew. Chem. Int. Ed.* 2008, 47, 2414–2417.
- [12] In't Veld M.; Iavicoli P.; Haq S.; Amabilino D. B. and Raval R.. *Chem. Commun.* 2008, 1536-1538.

- [13] M.M.S. Abdel-Mottaleb; E. Gomar-Nadal; M. Surin; H. Uji-I; W. Mamdouh; J. Veciana; V. Lemaury; C. Rovira; J. Cornil; R. Lazzaroni; D.B. Amabilino; S. De Feyter; F.C. De Schryver. *J. Mater. Chem.* 2005, 15, 4601-4615.
- [14] J. Puigmartí-Luis; A. Minoia; H. Uji-i; C. Rovira; J. Cornil; S. De Feyter; R. Lazzaroni; D.B. Amabilino. *J. Am. Chem. Soc.* 2006, 128, 12602-12603.
- [15] Abdurakhmanova N.; Floris A.; Tseng, T.C.; Comisso, A.; Stepanow, S.; De Vita A.; Kern K. *Nature Comm.* 2012, 3, 940.
- [16] Floris A; Comisso A; De Vita A. *ACS Nano* 2013, 7 (9), 8059.
- [17] Blüm M.-C.; Cavar E.; Pivetta M; Patthey F; Schneider W.-D.. *Angew. Chem.* **2005**, 117, 5468–5471.
- [18] Tomba, G.; Stengel, M.; Schneider, W.-D.; Baldereschi, A.; De Vita, A. *ACS Nano* 2010 4, 7545-7551.
- [19] A. Della Pia; M. Riello; A. Floris; D. Stassen; D. Bonifazi; Tim S. Johnes; A. De Vita; G. Costantini. *ACS Nano* 2014, 8 (12), 12356–12364.
- [20] S. Haq; F. Hanke; M. S Dyer; M. Persson; P. Iavicoli; D. B. Amabilino; R. Raval. *J. Am. Chem. Soc.* 2011, 133, 12031.
- [21] F. Hanke; S. Haq; R. Raval; M. Persson. *ACS Nano* 2011 5, 9093.
- [22] Björk, J; Stafström S.; Hanke F. *J. Am. Chem. Soc.* 2011, 133, 14884–14887.
- [23] Björk, J; Hanke, F; Stafström; S.. *J. Am. Chem. Soc.* 2013 135, 5768–5775.
- [24] Nguyen, M-T.; Pignedoli C. A.; Passerone D.. *Phys. Chem. Chem. Phys.* 2011 13, 154-160.
- [25] Blankenburg, S.; Rauls, E.; Schmidt, W. G.. *J. Phys. Chem. Lett.* 2010 1, 3266–3270.
- [26] M. S. Dyer; A. Robin; S. Haq; R. Raval; M. Persson and J. Klimeš. *ACS Nano* 2011 5, 1831.
- [27] Tersoff J. and Hamann D. R.. *Phys. Rev. B* 1985, 31, 805.
- [28] LEV00 package, URL <http://www.mth.kcl.ac.uk/~lev/codes/lev00/index.html>.
- [29] We did not relax the molecule after the H removal, to avoid intra-molecular bonds between the de-H radical and parts of the molecule as that may spuriously alter the hierarchy of the removal energies.
- [30] For a more consistent comparison with the de-H energies of the molecule on Cu(110), where the removed H atom was placed on the bare substrate and its calculated removal energy from the bare surface is 3.3 eV, the $\Delta E_{rem}(gas)$ energies were shifted down by 3.3 eV.
- [31] G. Henkelman; B.P. Uberuaga; H. Jónsson. *J. Chem. Phys.* 2000 113, 9901.

[32] Two different 4-methyl-4-methyl coupling (5,4) and (6,3) structures with a different intermolecular bond orientation were also found having slightly higher energies in comparison with the most stable ones quoted in the text (by +0.1 eV and +0.35 eV, respectively, Fig. SI-9).

[33] VandeVondele J.; Krack M.; Mohamed F.; Parrinello M.; Chassaing T. and Hutter J.. *Comp. Phys. Comm.* 2005, 167, 103.

[34] URL <http://www.cp2k.org/>.

[35] Goedecker S.; Teter M.; Hutter J.. *Phys. Rev. B* 1996, 54, 1703.

[36] Perdew; J.P.; Burke; K.; Ernzerhof; M.. *Phys. Rev. Lett.* 1996, 77, 3865–3868.

[37] Sabatini, R; Gorni, T; de Gironcoli, S.. *Phys. Rev. B* 2013, 87, 041108(R).

[38] P. Giannozzi; S. Baroni; N. Bonini; M. Calandra; R. Car; C. Cavazzoni; D. Ceresoli; G. L. Chiarotti; M. Cococcioni; I. Dabo; A. Dal Corso; S. Fabris; G. Fratesi; S. de Gironcoli; R. Gebauer; U. Gerstmann; C. Gougoussis; A. Kokalj; M. Lazzeri; L. Martin-Samos; N. Marzari; F. Mauri; R. Mazzarello; S. Paolini; A. Pasquarello; L. Paulatto; C. Sbraccia; S. Scandolo; G. Sclauzero; A. P. Seitsonen; A. Smogunov; P. Umari; R. M. Wentzcovitch. *J. Phys.:Condens. Matter* 2009 21, 395502; URL <http://www.quantum-espresso.org>.

[39] Grimme, S., *J. Comput. Chem.* 2006, 27, 1787–1799.