Electronic Supplementary Information (ESI) to "Diacetylene polymerization on a bulk insulator surface"

A. Richter,[a] V. Haapasilta,[b] C. Venturini,[c] R. Bechstein,[a] A. Gourdon,[c] A. S. Foster[b,d] and A. Kühnle[a]

[a] *Institute of Physical Chemistry, Johannes Gutenberg University Mainz, 55099 Mainz, Germany.*
[b] *COMP, Department of Applied Physics, Aalto University, PO Box 11100, FI-00076 Aalto, Finland.*
[c] *CNRS, CEMES, Nanoscience Group, BP 94347, 29 Rue J. Marvig, 31005 Toulouse, France.*

I. Synthesis Details

Solvents and Reagents

Solvents were purchased from Sigma Aldrich and Acros Organics. They were dried over molecular sieves 3Å, unless specified. NEt3 was dried over KOH then distilled under reduced pressure. All the reagents were purchased from Sigma Aldrich, Acros Organics or TCI. They were employed without additional purification.

Chromatographic purifications were carried out on silica gel (Fluka, pore size 6Å, 230-400Å mesh) or with Flash Chromatography SPOT II Armen Instrument (cartridge of silica Agela 60 Å; 20 g; 40 g). All the reactions were followed by thin layer chromatography (TLC) on silica gel (silica gel 60 F254, Merck or Fluka)

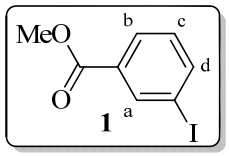
Nuclear magnetic resonance (NMR) spectroscopy

The NMR spectra were recorded on Bruker Advance 300 or 500 MHz spectrometers (Common service of University Paul Sabatier, Toulouse). The chemical shift (δ) reported in this manuscript are in ppm. The following abbreviations were employed: s = singlet, d = doublet, dd = doublet of doublets, t = triplet and m = multiplet.

Mass spectroscopy

The mass spectra and high resolution mass (HRMS) spectra were recorded at the common service of University Paul Sabatier (Toulouse), on different instrument in ESI, DCI (CH4 or NH3) and APCI modes.

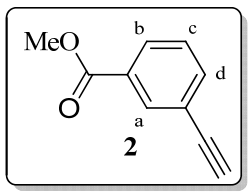
Synthesis of methyl 3-iodobenzoate [1]



In a round-bottomed flask were added 3-iodobenzoic acid (2g, 8.06 mmol), some drops of conc. H_2SO_4 and MeOH (50 mL). The solution was heated under reflux overnight. The mixture was then cooled at room temperature and diluted with diethyl ether (50 mL) and water (50 mL). The water phase was washed with diethyl ether (2 x 50 mL). All the organic phases were collected and washed with a saturated solution of NaHCO3 and brine. The organic phase was then dried over anhydrous $MgSO_4$ and evaporated under vacuum. The product **7** (1.879 g) is a white solid and was obtained with a yield of 89 %.

^{1}H-NMR (300 MHz, $CDCl_3$): δ_H 8.38 (1H, s , Ar-*H*$_a$), 8.02-7.99 (1H, m, Ar-*H*$_b$), 7.90-7.87 (1H, m, Ar-*H*$_d$), 7.19 (1H, t, *J* = 7.8 Hz, Ar-*H*$_c$), 3.92 (3H, s, -OC*H*$_3$).

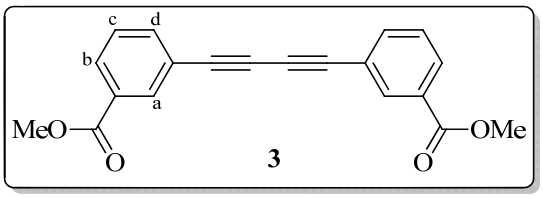
Synthesis of methyl 3-ethynylbenzoate [2-4]



Methyl 3-iodobenzoate **1** (1 g, 3.8 mmol, 1 eq), CuI (29 mg, 0.152 mmol, 0.04 eq), $Pd(PPh_3)_2Cl_2$ (53 mg, 75.5 µmol, 0.02 eq) were dissolved in degassed THF(10 mL) in a Schlenk flask. Then trimethylsilylacetylene (0.81 mL, 1.5 eq) and NEt_3 (1.06 mL) were added and the mixture was again degassed for some minutes. The solution was stirred at room temperature under argon atmosphere. After 3h (when all the compound **1** had reacted) were added TBAF in THF solution (7.65 mL) and some molecular sieves (3Å). The reaction was stirred under air atmosphere and at room temperature overnight. The mixture was then filtered on Celite®, washing with Et_2O. The residue was then purified by column chromatography (SiO_2, EtOAc/ hex 0.5 : 9.5) to afford the product **2** (0.392 g, 64 %).

^{1}H NMR (300 MHz, $CDCl_3$) δ_H 8.16 (1H, s, Ar-*H*$_a$), 8.03-8.00 (1H, m, Ar-*H*$_b$), 7.68-7.64 (1H, m, Ar-*H*$_d$) , 7.41 (1H, t, *J* = 7.8 Hz, Ar-*H*$_c$), 3.92 (3H, s, -OC*H*$_3$), 3.12 (1H, s, -C*H*).

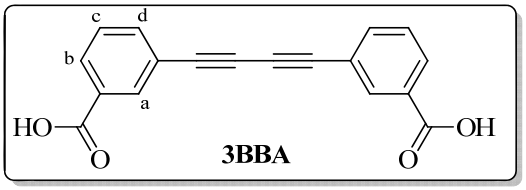
Synthesis of dimethyl 3,3'-(buta-1,3-diyine-1,4-diyl)dibenzoate [2, 5]



In a flask equipped with magnetic stirrer were added CuCl (7 mg, 68.7 µmol, 0.10 eq), TMEDA (0.031mL, 0.207 mmol, 0.30 eq) and CH_2Cl_2. Once that all the compounds had dissolved, molecular sieves (3Å, 1g) were added and after few minutes methyl 3-ethynylbenzoate **2** (0.111 g, 0.693 mmol, 1 eq). After completion (ca 2 hours, TLC), the reaction was quenched by addition of water (7 mL). The suspension was filtered and the precipitate was washed with CH_2Cl_2. After addition of water, the aqueous phase was extracted with CH_2Cl_2 and all the organic phases were collected together and filtered on SiO_2. The filtrate was evaporated to afford the product **3** (0.100 g, 91 %).

^{1}H NMR (300 MHz, $CDCl_3$) δ_H 8.20-8.22 (2H, m, 2 x Ar-H_a), 8.06-8.03 (2H, m, 2 x Ar-H_b), 7.72-7.69 (2H, m, 2 x Ar-H_d), 7.46-7.41 (2H, m, 2 x Ar-H_c), 3.94 (6H, s, 2 x -OCH_3).

Synthesis of 3,3'-(buta-1,3-diyne-1,4-diyl)dibenzoic acid (3BBA) [6, 7]

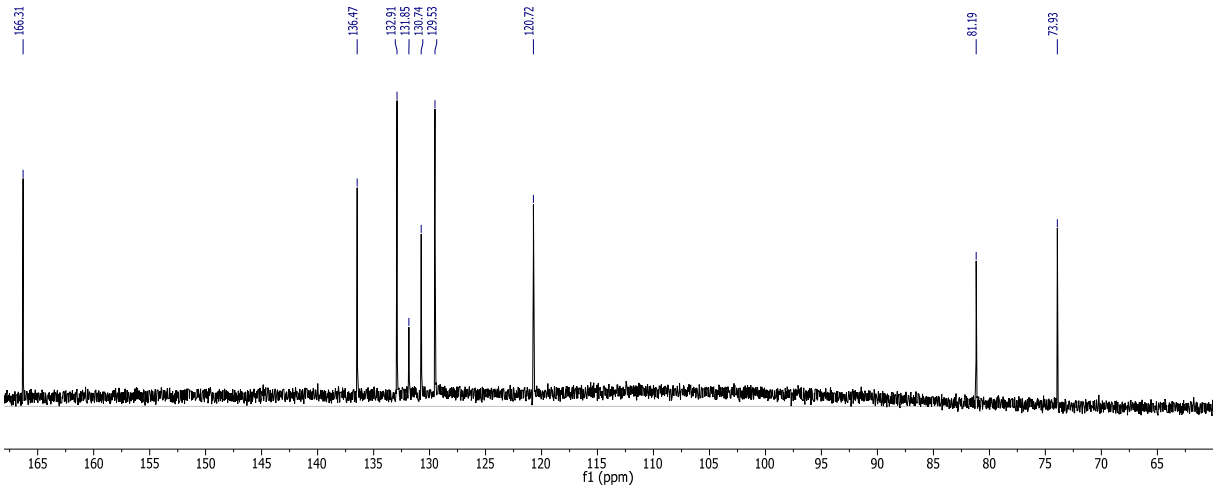


In a small flask the reactant **3** (60 mg, 0.188 mmol, 1 eq) was dissolved in THF (7.2 mL). Then a solution of LiOH (0.092 g, 3.84 mmol, 20.38 eq) in water (4mL) was added and stirred at room temperature overnight. HCl 3M (1.2 mL) was then added drop by drop under stirring to the cooled solution (ice bath) and the *p*H was adjusted to 2~3, with formation of a precipitate. This precipitate was then filtered on funnel, washed with Et_2O and cold water, and dried under vacuum (**3BBA**, yield 54 %).

^{1}H NMR (500 MHz, DMSO) δ_H 8.09 (2H, s, 2 x Ar-H_a), 8.04-8.01 (2H, m, 2 x Ar-H_b), 7.88-7.86 (2H, m, 2 x Ar-H_d), 7.62-7.56 (2H, m, 2 x Ar-H_c).

^{13}C NMR (500 MHz, DMSO) δ_C 166.3, 136.4, 132.9,131.8, 130.7,129.5, 120.72, 81.1, 73.9.

HRMS (TOF MS ES-) calculated for [$C_{18}H_9O_4$] 289.0501, found 289.0498



^{13}C NMR of molecule **3BBA**.

	C_a	C_b	C_c	C_d	C_e	C_f	C_g	C_h	C_i
δ_C **(ppm)**	132.9	130.7	129.5	136.5	*131.9*	*120.7*	81.2	73.9	166.3

II. Computational Details

To better understand molecular-level details, we performed density functional theory (DFT) calculations (T = 0 K) and first-principles molecular dynamics (FPMD) simulations (T ≠ 0 K) using the CP2K program suite.[8] In all computations the PBE functional[9] was used together with the D3 dispersion correction scheme.[10] The Gaussian Plane Wave Method[11] was used to solve the electronic structure. Core electronic states were modelled with GTH pseudopotentials[12] and the valence electrons were described with molecularly optimized polarized double-ζ basis functions.[13] Electronic density was described with a plane wave basis with an energy cutoff of 500 Ry. The internal convergence criterion in the self-consistent field iteration step was 10^{-7}.
In the geometry optimization calculations, we used the conjugate gradient algorithm. Four-layers calcite slabs were used in all the computations, with the atomic coordinates of the two bottom-most layers always kept frozen. Periodic boundary conditions were used in all computations. In the island and stripe pattern calculations with three or six (explicit) molecules, the lateral size of the calcite slab was (3×3) surface unit cells (15.1575 Å × 24.5595 Å). While optimizing the isolated molecule and benchmarking, a slab of (6×3) surface unit cells was used (30.315 Å × 24.5595 Å). In the FPMD simulations with two (explicit) molecules, a slab of

(2×3) surface unit cells (10.105 Å × 24.5595 Å) was used. In all computations, there were 25 Å room in the direction perpendicular to the surface.

In the case of large organic molecules on calcite (10.4) surface, the optimizations are far from being trivial: the flexibility of the calcite ions together with the floppy molecules poses a serious challenge to even the best of software. For example, in the calculations with three molecules on surface, there are 822 cartesian degrees of freedom to optimize (32 atomic nuclei per molecule and two layers of calcite with 90 atomic nuclei per layer). Nevertheless, we managed to converge most of the geometries to 0.03 eV/Å or better.

In addition to the optimizations, we ran FPMD simulations at the elevated temperature of 600 K to investigate the stability of the obtained optimized geometries and to explore the extent of thermal fluctuations around the optimized bonding patterns. To ensure effective thermalization even for relatively short simulation runs (~4 ps for the single island rows and ~2 ps for the complete island pattern and for the polymer rows; see below), we used canonical sampling through velocity rescaling method with a time constant of 50 fs.[14]

Figure ESI_01 (a) shows a calcite slab with (6×3) surface unit cells from three angles. The zigzag nature of the protruding carbonate oxygen atoms is very characteristic of calcite (10.4) and is essential in the anchoring mechanism with the 3BBA molecule studied here. Figure ESI_01 (b) shows all the six possible flat gas-phase configurations of 3BBA molecule. The geometries differ by the relative angle of the carbon ring and the orientation of the COOH group. Energetically the configurations are in practice identical.

The 3BBA molecule interacts with calcite surface using the COOH groups. Taking only this constraint for the bonding, there would be several possible adsorption geometries to consider. However, the experiments suggest that the angle the backbone of the molecule forms with respect [-4-21] and [010] surface directions is roughly 45° (see Figure 2 in the main text). In addition, both calculations and simulations suggest that the isolated molecule prefers to be flat on the surface. In figure ESI_02 the isolated molecule is shown in three different “tilt angles” with corresponding adsorption energies. Although the flat orientation yields the lowest total energy, the energetic difference between flat and slightly tilted geometry is very small. The calculations reveal that there are no signs of proton transfer in none of the geometries with an isolated 3BBA molecule on calcite – the molecule is controlling both of the COOH group

protons. It is worth noting that in the experimentally observed self-assembled island patterns the molecules are so close to each other that they must be tilted upwards.

In figure ESI_03, we compare isolated monomer, dimer and trimer energetics in the tilted orientation on a large slab. The geometries of the individual molecules are practically identical in all cases. However, the molecules do gain some energy by interacting with each other, 0.09 eV/molecule and 0.08 eV/molecule in the case of dimer and trimer, respectively.

When all the molecules have a neighbor on both sides, the interaction energy is 0.11 eV/molecule. This geometry, corresponding to a molecular row, is shown in figure ESI_04 (a). Here we have three explicit molecules in the calculation. Figure ESI_04 (b), (c) and (d) show the other three possibilities for the molecular row structure, which are consistent with the experimentally observed 0.5 nm periodic separation along the [010] surface direction and have COOH group anchoring mechanism. The structures differ in the tilt direction (towards positive or negative [010] direction) and in the relative angle of the carbon rings. Besides geometries, figure ESI_04 contains also the total and formation energies. Out of the four structures, (a) and (d) are clearly lower in total energy than the remaining two, which are almost 2 eV less stable. The same trend in the stabilities was also observed in the FPMD simulations of molecular rows at 600 K. In all the simulations proton transfer reactions were observed from the 3BBA molecules to the calcite carbonate oxygen atoms and back. Furthermore, the carbon rings of the 3BBA molecules were able to exhibit rotations with respect to each other while still retaining the general bonding motif of the anchoring (COOH group hydrogen bonding with carbonate group oxygen atoms). A simulation starting from the structure in figure ESI_04 (a) showed tendency towards the geometries (a) and (d), and a simulation starting from the (d) structure remained mostly in this geometry with occasional frames towards the geometry (a). Snapshots from these FPMD trajectories are shown in figure ESI_05 (a) and (d), respectively. The structures (b) and (c) in figure ESI_04, that were less stable at T = 0 K, were also less stable at T = 600 K. A simulation starting from the geometry (b) transformed into geometry (a) – in other words, the direction of the carbon ring tilt angle was reversed. Snapshots of this simulation are shown in figure ESI_05 (b) demonstrating a close resemblance towards figure ESI_05 (a) trajectory. Similarly, a simulation starting from the geometry of figure ESI_04 (c) was not stable but transformed into (b) geometry of figure ESI_04. Snapshots of this trajectory are shown in figure ESI_05 (c). To summarize the results of the stability simulations: presumably after longer simulations all the trajectories would converge so that also the latter two cases (structures (b) and (c) in figure ESI_04) would end up in a geometry oscillating between the structures (a) and (d), which seems to be the stable state of the 3BBA molecular row on calcite.

However, as argued in the main text, it is likely that individual molecules cause the bright features observed in the self-assembled island patterns. On molecular-level this means that the structures shown in figure ESI_04 correspond to exactly half of the experimentally observed density. Basically, they are isolated molecular rows: these structures explain only every second of the observed bright molecular rows. To obtain molecular understanding of the higher density structures and to ascertain the experimental interpretation, we took the most stable "half-density" configurations ((a) and (d) in figure ESI_04) and increased the number of molecules in the (1×3) supercell from three to six. The additional three molecules were placed at 1.5 unit cell lengths (1.2 nm) along the [-4-21] direction. The resulting structures are shown in figure ESI_06. The structure (b) in figure ESI_06, derived from the geometry (d) in figure ESI_04, turned out to be the most stable molecular structure obtained in this study with formation energy of 2.7 eV per molecule. What is more, in this structure the carbon rings of the molecules form an angle with each other with the effect of notably protruding one end of the molecule higher from the surface while flattening the other end (see figure ESI_06 (b)). These characteristics match very well with the experimental images, as is shown in the main text Figure 3d. Closer look of the structure is shown in figure ESI_08 (a) with the average OH-O distances, where it can be seen that each of the molecules transfer one proton to a nearby carbonate group. Furthermore, figure ESI_08 (c) shows snapshots of FPMD simulation of this structure at 600 K, together with the average proton control percentages. The simulation reveals dynamic hydrogen bonding between the molecules and the surface. The protons that are transferred to the carbonate groups in static optimizations are more dominantly controlled by the surface also in the FPMD simulation. Similarly, the protons that remain with the molecules in the T = 0 K optimizations are almost completely controlled by the molecules also in the FPMD simulation. Four optimized structures for the observed polymerized stripes are shown in figure ESI_07: structures (a) to (d) are derived from structures (a) to (d) shown in figure ESI_04, respectively, by introducing covalent bonds between the molecules. Figure ESI_07 reports also the total and formation energies. All the considered polymerized structures are more stable than their corresponding non-covalent versions. The stabilities of the polymerized structures are consistent with the corresponding molecular row structure stabilities, although the energetic differences are even larger. The polymerized structures derived from the more stable island geometries shown in figure ESI_04 (a) and (d) are clearly more stable than the remaining two structures. The most stable structure, shown in figure ESI_07 (a), is almost one eV more stable than the second best geometry. As the polymerization is achieved by annealing or by radiation, that is, heavily disturbing the system, we assume that the most stable polymerized structure is

always able to form. Thus, the energy gain in the polymerization from the most stable self-assembly pattern (figure ESI_06 (d)) to the most stable polymer stripe structure (figure ESI_07 (a)) is 1.18 eV per diacetylene unit. The most stable polymer geometry is shown with average OH-H distances in ESI_08 (b). Finally, we also studied the stabilities of the polymerized structures using FPMD simulations at 600 K. Snapshots of the simulation trajectories of the most stable structure are shown in figure ESI_08 (d), together with proton control percentages. In the simulations, the structures exhibited thermal movements around their optimized bonding patterns, similarly to the island patterns. The simulation results suggest that the optimized structures provide good approximations for the room temperature molecular geometries.

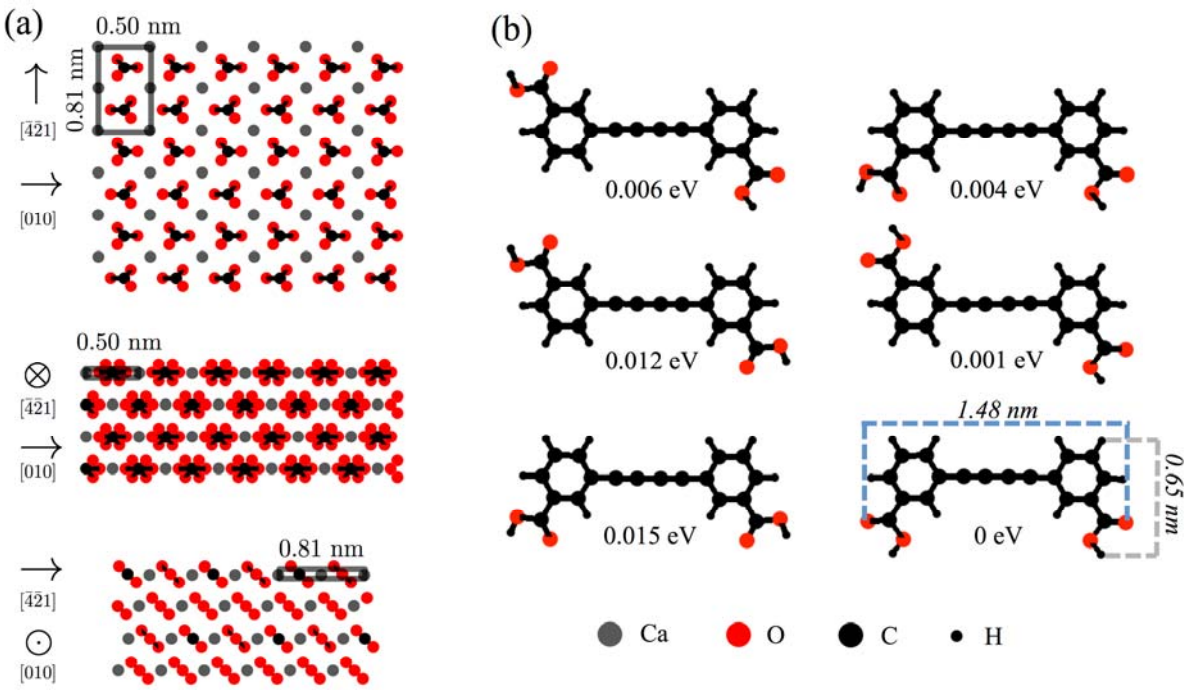


Fig. ESI_01. Calcite slab of (6×3) surface unit cells with (10.4) surface is shown in (a) from three angles. In (b) six flat gas-phase configurations of the 3,3'-(1,3-butadiyne-1,4-diyl)bisbenzoic acid (3BBA) are shown, with total energies with respect to the most stable one. The energetic differences are very small. All geometries are converged better than 0.003 eV/Å. Please note that here the surface and the molecules are not in scale.

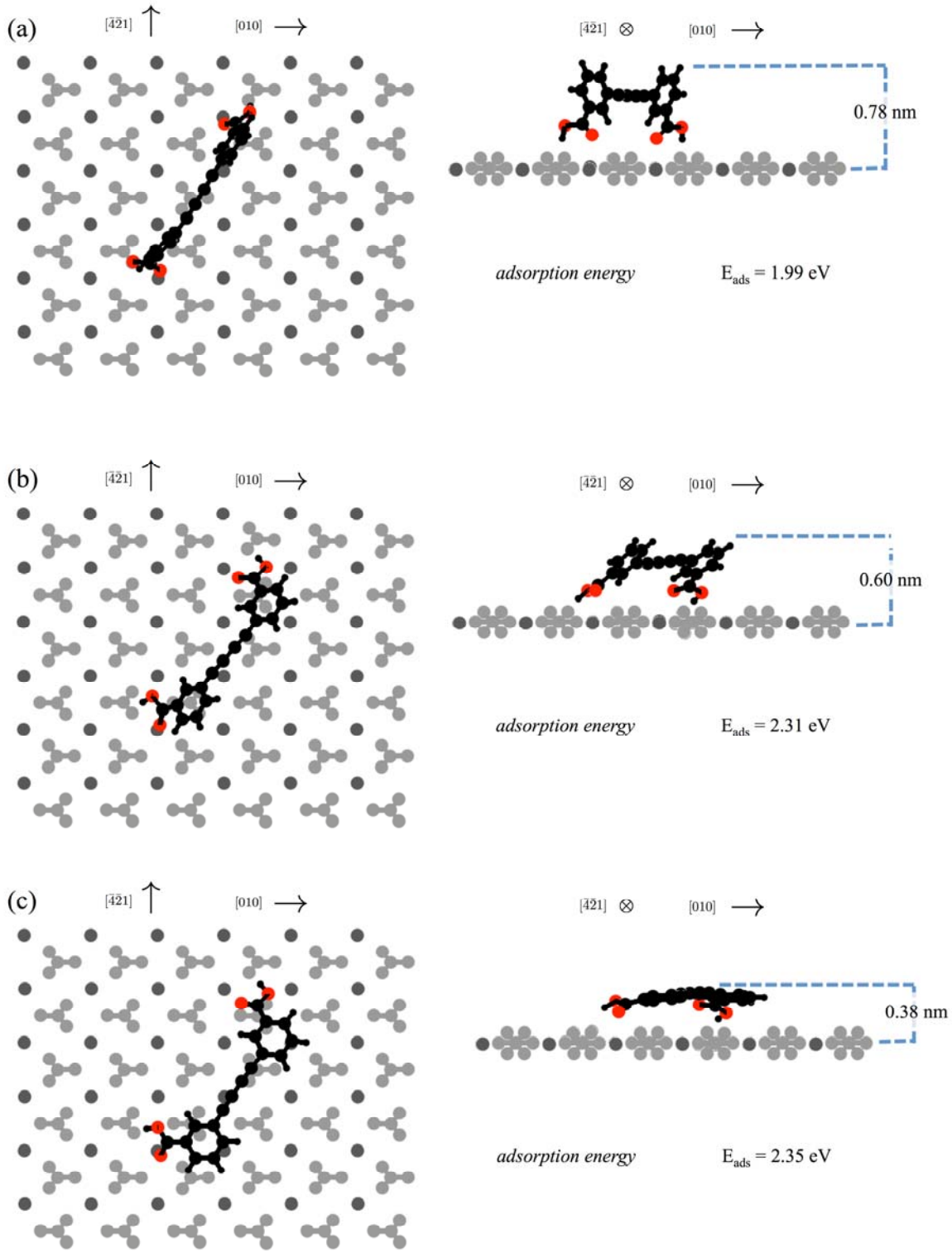


Fig. ESI_02. An isolated 3BBA molecule on calcite surface with different “tilt angles”. The adsorption energies are reported, together with heights. The energetic difference between the flat and slightly tilted configurations is very small.

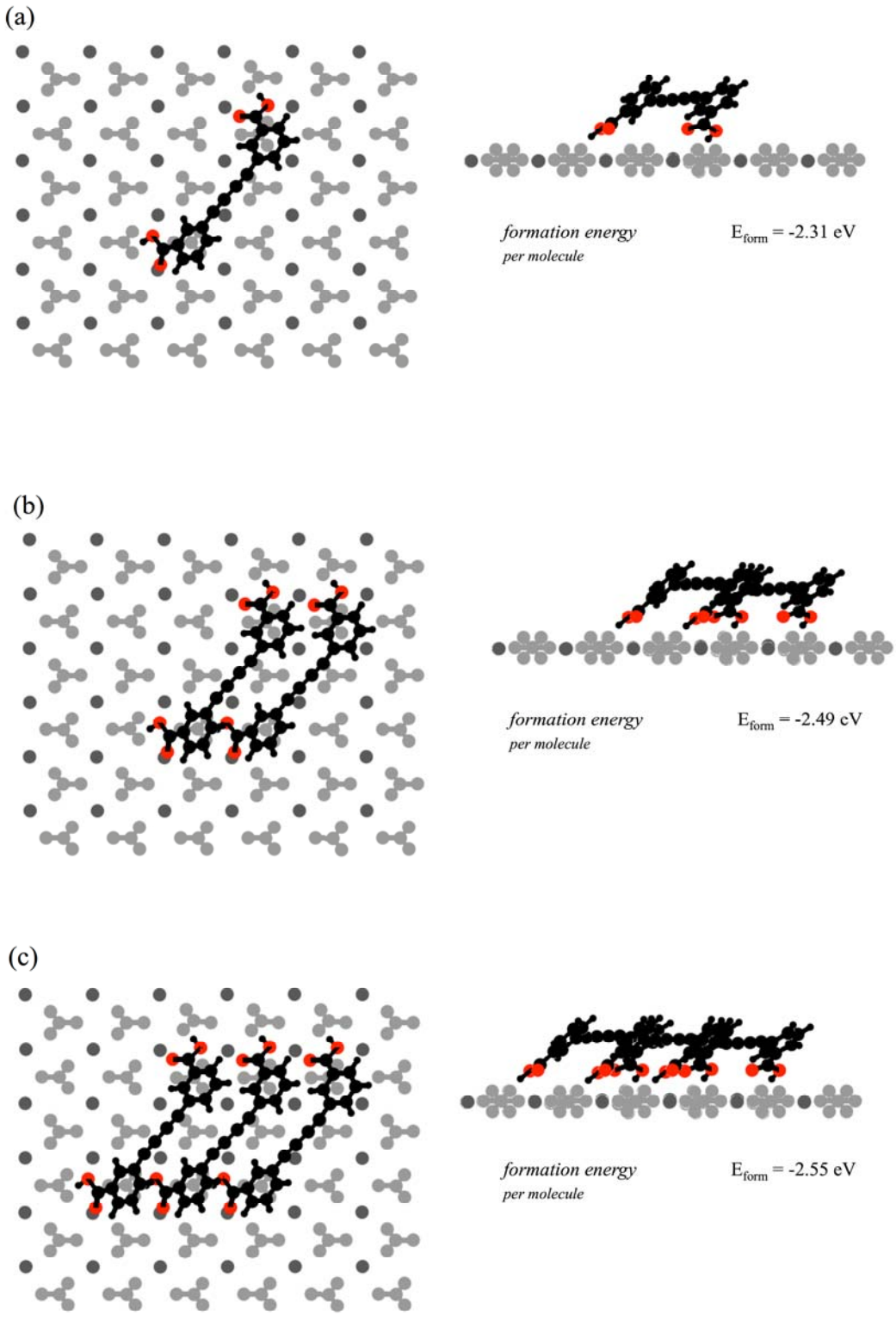


Fig. ESI_03. 3BBA monomer (a), dimer (b) and trimer (c) on calcite surface. The formation energies are shown, indicating small but attractive interaction between the molecules.

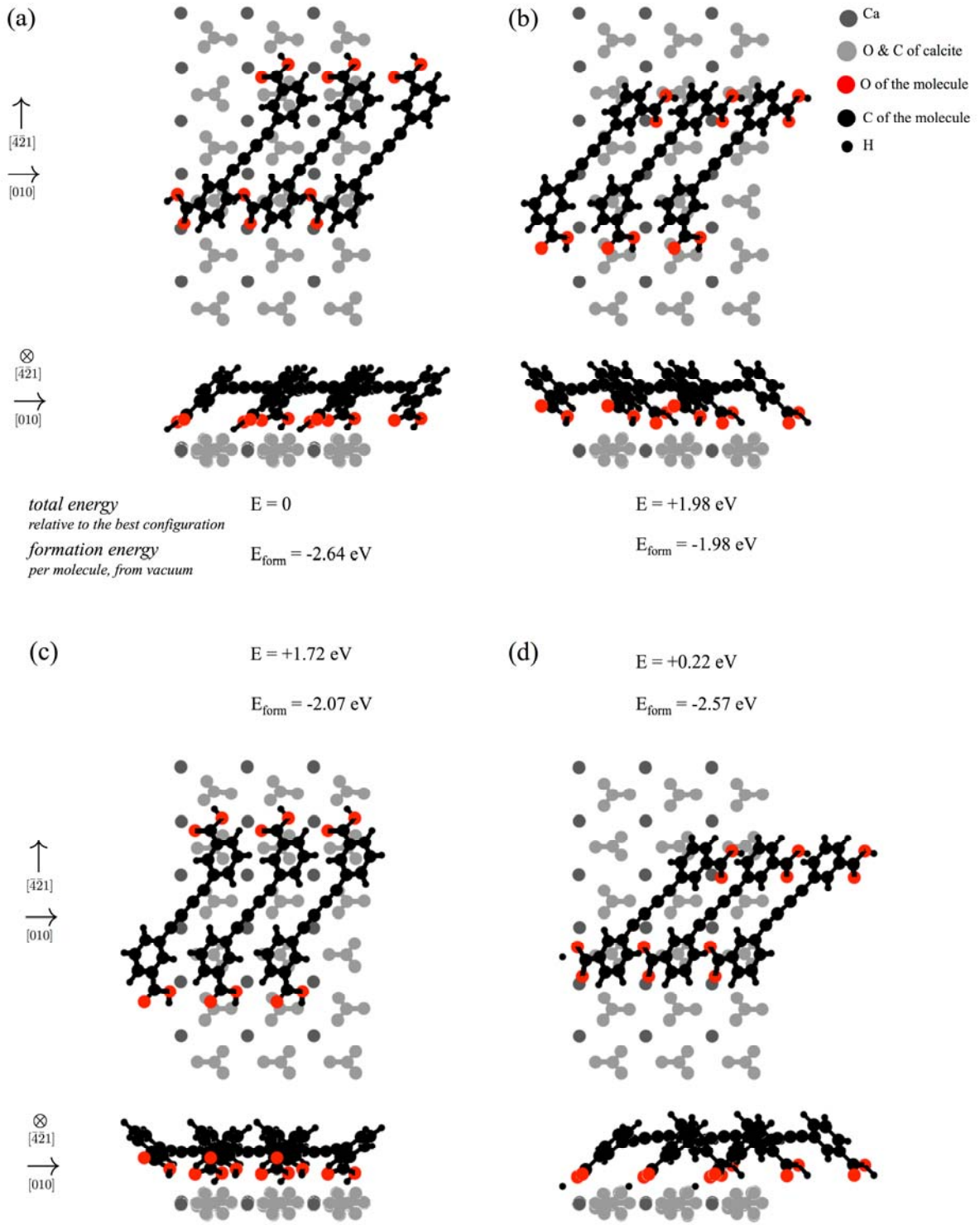


Fig. ESI_04. Four optimized configurations for the (isolated) molecular rows. The total and formation energies are shown.

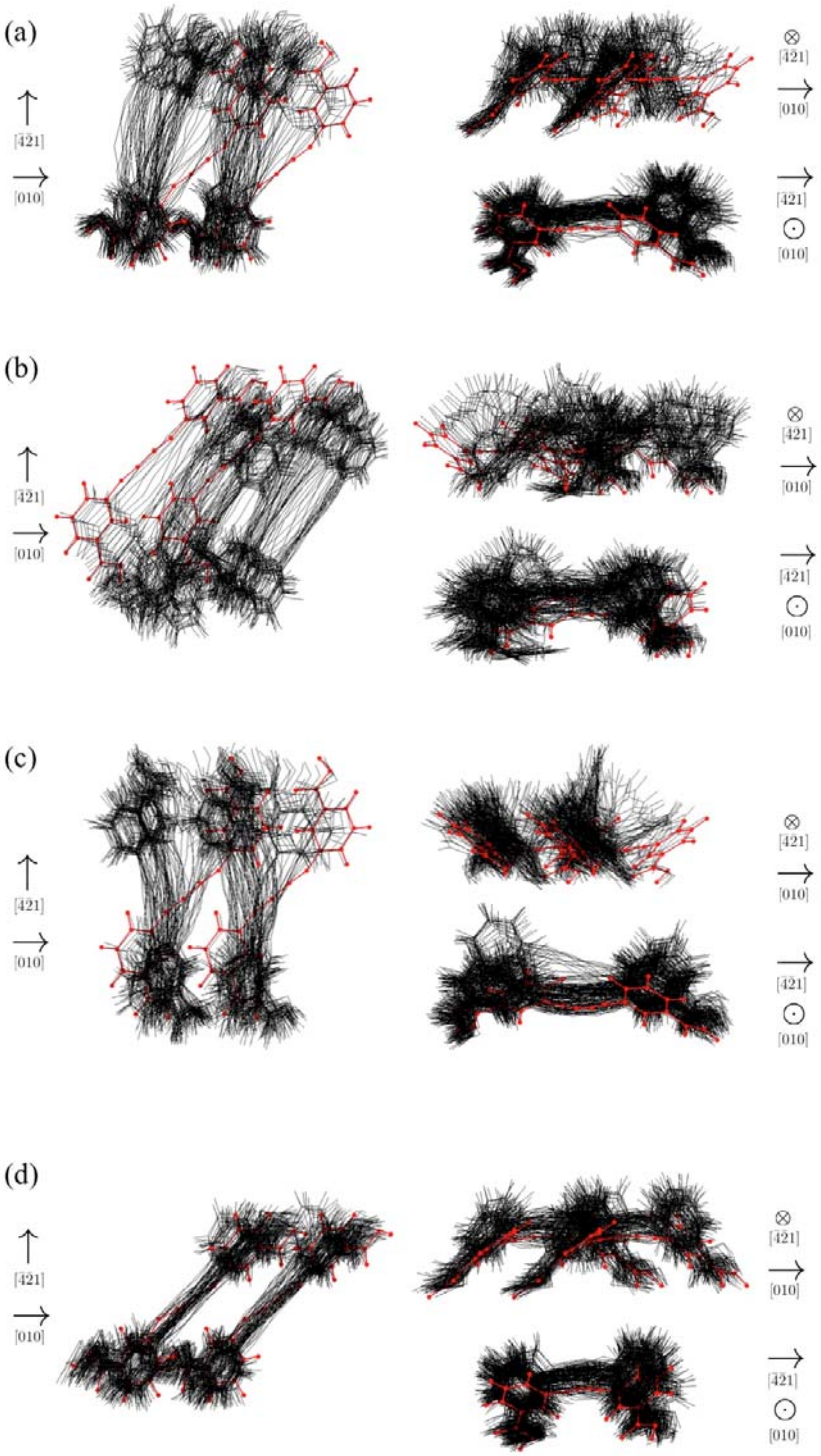
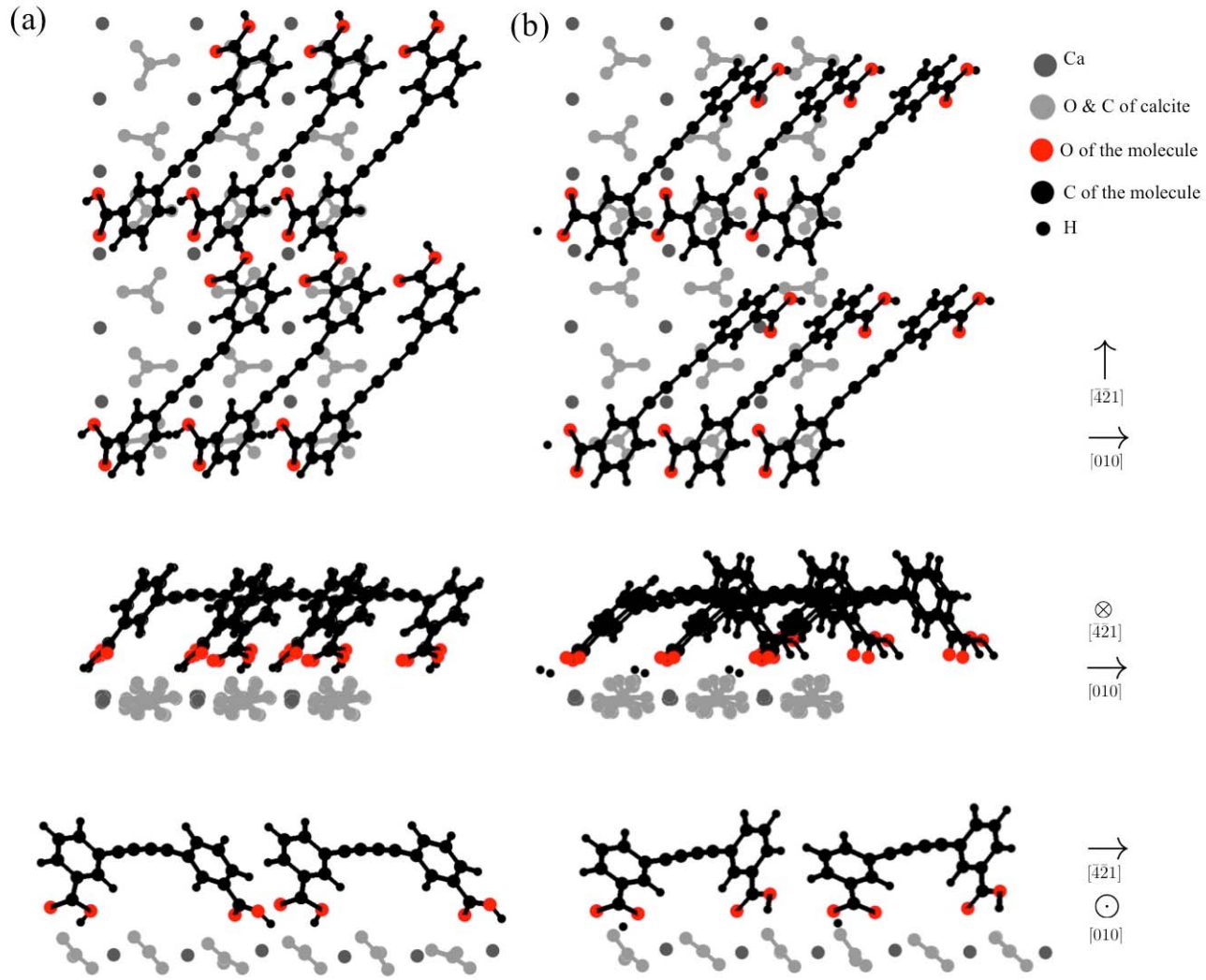


Fig. ESI_05. Combined snapshots of first-principles molecular dynamics simulation runs at 600 K. The optimized, starting geometries are shown in bold red. The combined FPMD trajectory snapshots (with 50 fs intervals for 4 ps) give an overview of the trajectories in one image. This smears out some details but clearly demonstrates the extent of thermal motion. For simplicity, the surface is not shown here.



total energy *relative to the best configuration*	E = +6.0 eV	E = 0
formation energy *per molecule, from vacuum*	E_{form} = -1.74 eV	E_{form} = -2.74 eV

Fig. ESI_06. Two optimized configurations for the (1×3) superstructure derived from the most stable molecular rows shown in Fig. ESI_04 ((a) from (a) and (b) from (d)). The total and formation energies are given.

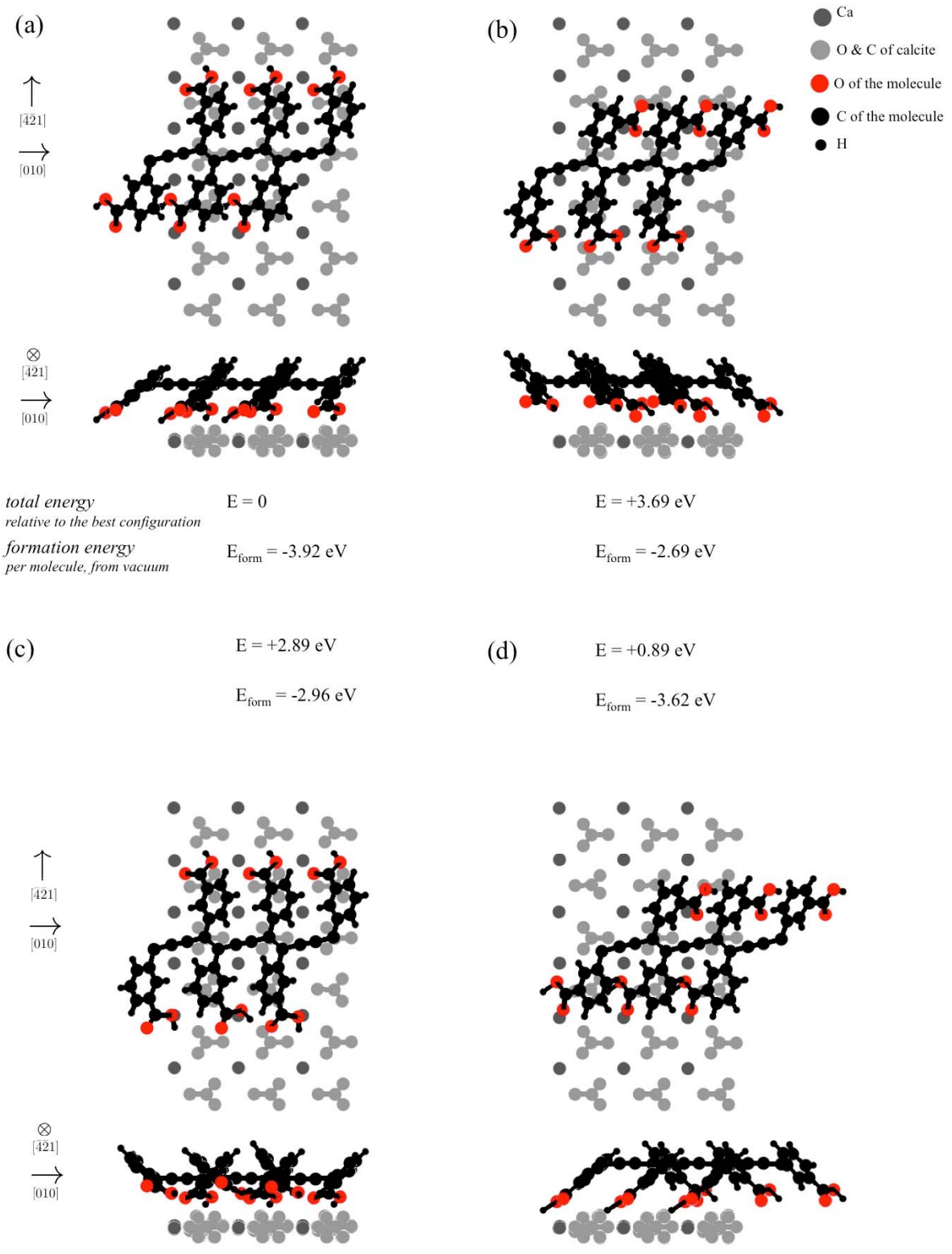


Fig. ESI_07. Four optimized configurations for the polymerized stripes. The structures (a) to (d) are derived from the structures (a) to (d) shown in figure ESI_04, respectively. The total and formation energies are shown.

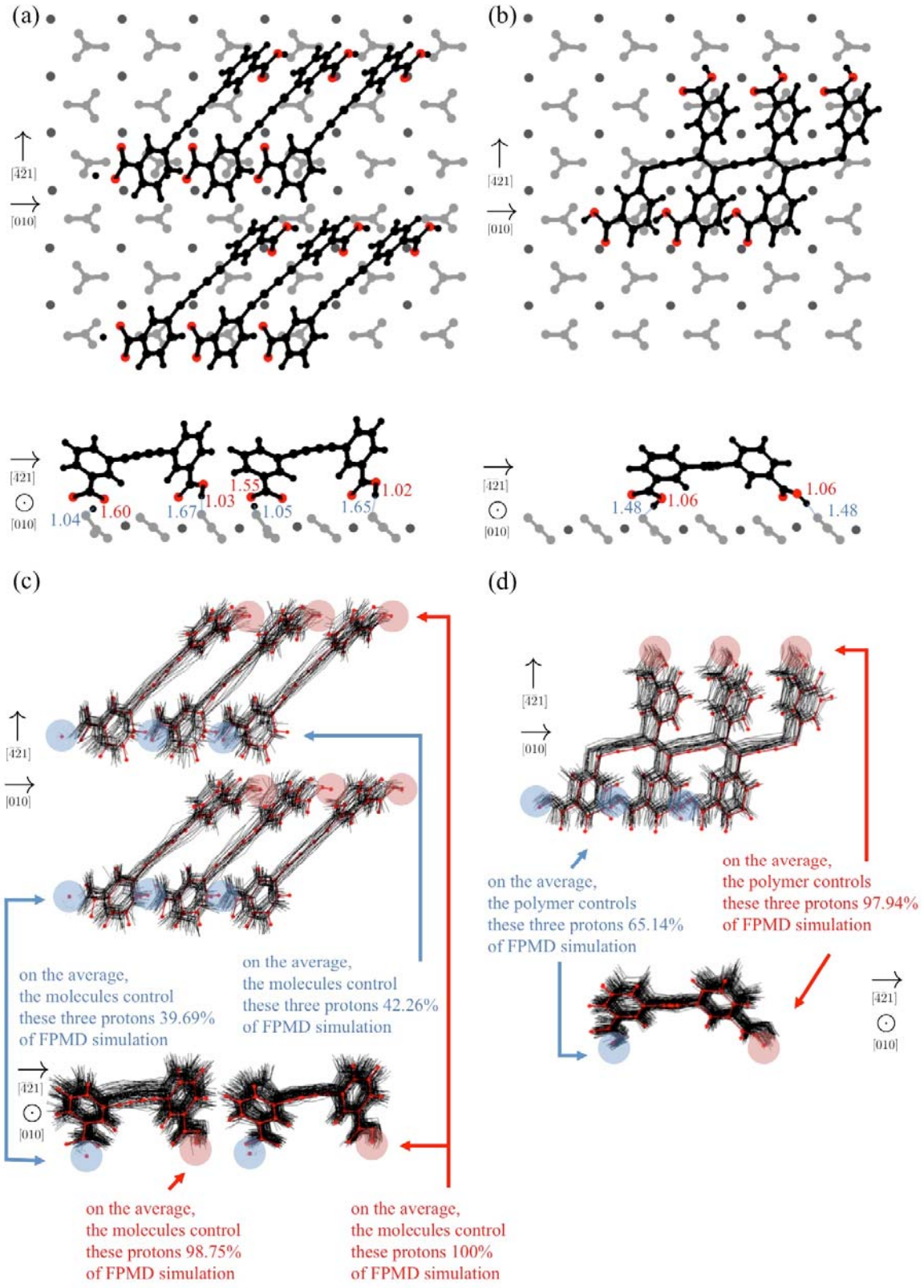


Fig. ESI_08. The most stable optimized structures for the self-assembly pattern (a) and for the polymer (b). Average OH-O distances are shown for both structures. In (c) and (d) snapshots from FPMD trajectories (at 15 fs

interval for 1.5 ps) are shown for the self-assembly pattern and the polymer, respectively. The optimized structures are shown in bold red. Simulations reveal dynamic hydrogen bonding around the optimized geometries. Average proton control percentages are reported.

III. Further Details to AFM Measurements

Changes induced by annealing

Annealing of the molecule-covered substrate at 485K for 1 h results in three structural changes as can be seen in Fig. ESI_09 (a). The number of islands has decreased and clusters have formed on the surface. Most importantly, single stripes have also formed. When annealing instead at 520 K for 1 h the described structural changes advance further. Then, fewer islands are found while more clusters and even longer stripes have formed (Fig. ESI_09 (b)). After annealing at 555 K for 1 h, the islands have disappeared completely and only clusters and stripes are present on the surface (Fig. ESI_09 (c)).

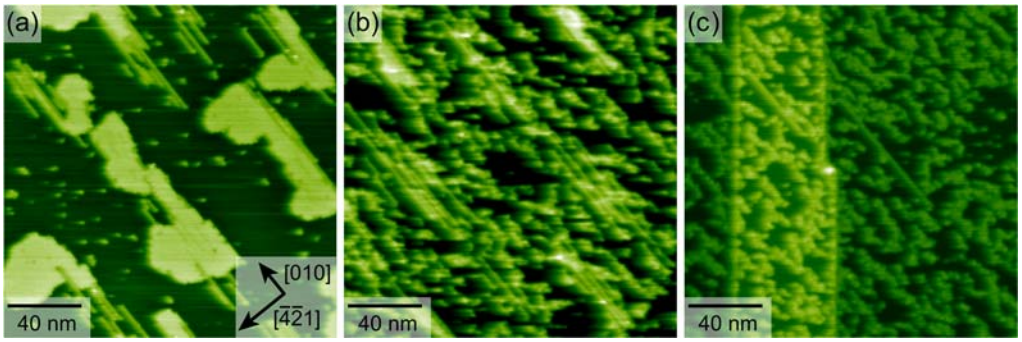


Fig. ESI_09. Molecular structures of 3BBA on calcite (10.4) after annealing. (a) Upon annealing at 485 K for 1 h, three distinct changes are observed: The number of islands has decreased. Clusters are formed. Most importantly molecular rows exist on the surface. (b) Upon annealing to 520 K for 1 h, structural changes advance further. (c) Upon annealing to 555 K for 1 h, the islands have disappeared completely and only clusters and stripes are present on the surface. Scale bar and directions in (a) apply for all images.

Changes induced by irradiation

Changes could also be induced by using irradiation with light from a mercury lamp. In both cases the uniformity of the islands vanish. The full lamp spectrum was tested and also only a part of the spectrum which was obtained by using an appropriate filter (302 ± 23) nm. Upon

irradiation with light of the full mercury lamp spectrum overnight, three distinct changes are observed: Clusters are formed. On the edges of the islands a few molecular rows have formed with a difference of the apparent height of 0.1 nm in contrast to the island (indicated by a white arrow in Fig. ESI_10 (a)). Furthermore, a corrugation of the apparent height inside the island is visible (indicated by yellow arrows). The observed clusters can be interpreted as by-products. The molecular rows with 0.1 nm apparent height difference can be compared to the separated stripes obtained via annealing. The corrugation of the apparent height inside the islands can be interpreted as reaction in between the island.

After irradiation overnight with a part of the spectrum (Fig. ESI_10 (b)), which was obtained by using an appropriate filter (302 ± 23) nm, only very few clusters are visible, while a clear corrugation of the apparent height inside the island is found. The darker stripes in between the island can be interpreted as reacted polymer structure.

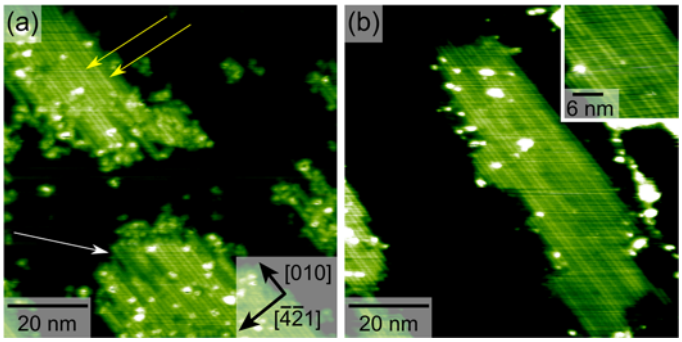


Fig. ESI_10. Molecular structures 3BBA on calcite (10.4) after irradiation. (a) Upon irradiation with full mercury lamp spectrum, three distinct changes are observed: Clusters are formed. On the edges of the island a few molecular rows have formed with a difference of the apparent of 0.1 nm in contrast to the island (indicated by a white arrow). Furthermore, a corrugation of the apparent height in the island is visible (indicated by yellow arrows). (b) Upon irradiation with an appropriate filter (302 ± 23) nm, clusters are forming and structural changes concerning the corrugation of the apparent height inside the island are visible. A zoom in the island is shown in the upper right corner. Scale bar and directions in (a) apply for all images.

References

1 D. A. Offermann, J. E. McKendrick, J. J. P. Sejberg, B. L. Mo, M. D. Holdom, B. A. Helm, R. J. Leatherbarrow, A. J. Beavil, B. J. Sutton and A. C. Spivey, *Journal of Organic Chemistry*, 2012, **77**, 3197-3214.

2 E. Merkul, D. Urselmann and T. J. J. Müller, *European Journal of Organic Chemistry*, 2011, 238-242.

3 W. B. Austin, N. Bilow, W. J. Kelleghan and K. S. Y. Lau, *Journal of Organic Chemistry*, 1981, **46**, 2280-2286.
4 Y. S. Feng, C. Q. Xie, W. L. Qiao and H. J. Xu, *Organic Letters*, 2013, **15**, 936-939.
5 M. H. Vilhelmsen, J. Jensen, C. G. Tortzen and M. B. Nielsen, *European Journal of Organic Chemistry*, 2013, 701-711.
6 S. B. Han, Z. B. Ma, Y. H. Wei, V. C. Kravtsov, B. S. Luisi, I. Kulaots and B. Moulton, *Crystengcomm*, 2011, **13**, 4838-4840.
7 H. Matsubara, T. Shimura, A. Hasegawa, M. Semba, K. Asano and K. Yamamoto, *Chemistry Letters*, 1998, 1099-1100.
8 J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Computer Physics Communications*, 2005, **167**, 103-128.
9 J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1996, **77**, 3865-3868.
10 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *Journal of Chemical Physics*, 2010, **132**, 15410.
11 G. Lippert, J. Hutter and M. Parrinello, *Molecular Physics*, 1997, **92**, 477-487.
12 S. Goedecker, M. Teter and J. Hutter, *Physical Review B*, 1996, **54**, 1703-1710.
13 J. VandeVondele and J. Hutter, *Journal of Chemical Physics*, 2007, **127**, 114105.
14 G. Bussi, D. Donadio and M. Parrinello, *The Journal of Chemical Physics*, 2007, **126**, 014101