Electronic Supplementary Information for

Two-Dimensional Self-Assembly of Benzotriazole on an Inert Substrate.

Federico Grillo[a,*], José A. Garrido Torres[a], Michael-John Treanor[a], Christian R. Larrea[a], Jan Philpp Götze[a], Paolo Lacovig[b], Herbert A. Früchtl[a], Renald Schaub[a], Neville V. Richardson[a]

[a]EaStCHEM and School of Chemistry, University of St. Andrews, St. Andrews, KY16 9ST, UK;

[b]Elettra Sincrotrone Trieste SCpA, SS 14 km 163.5, 34149 Basovizza - TS, Italy;

* federico.grillo@st-andrews.ac.uk

Contents:

ESI1. NEXAFS N K-edge comparison with the calculated N LDOS

Other than to obtain geometrical information, NEXAFS is used to obtain an electronic description of the empty electron energy levels of a system, by populating its empty states. The spectrum obtained can be compared with the calculated DOS above the Fermi level. As shown in **Figure S1**, by comparison with the spectra reported in figure 2, considering the energy region between 1.5 eV and 3.2 eV above the Fermi level, the best qualitative agreement is with the DOS calculated for type 1 interaction. The total DOS curves are in all cases made by the summation of p_z states over the N atoms; p_x and p_y states contributions are not so significant when compared to the p_z, to the extent that in the first approximation, for the energy range of interest, the total DOS can be well described as p_z, which is related to the high directionality of the π system (π* normal to the molecular plane). In this framework, the z direction is perpendicular to the substrate which lies on the (x, y)-plane.

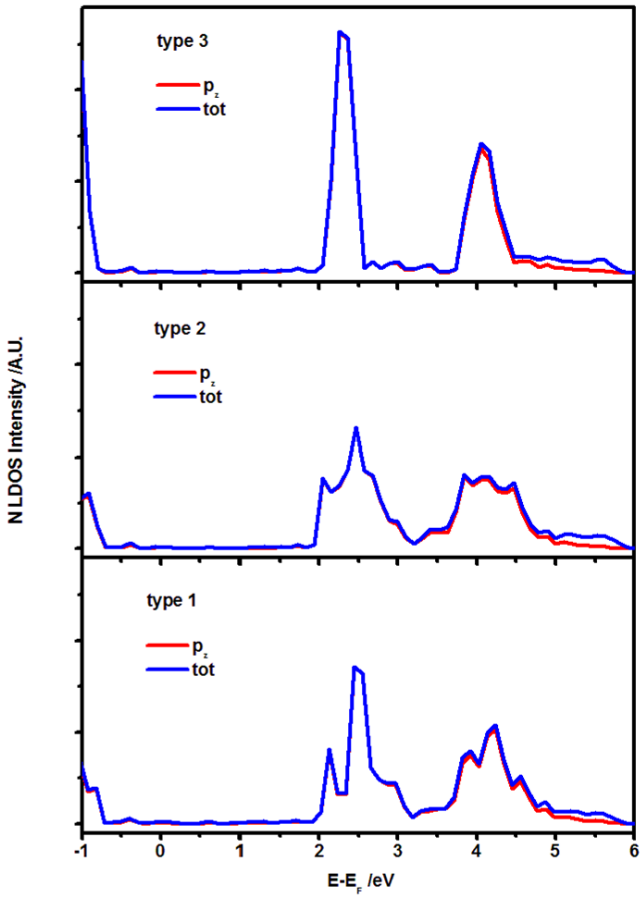


Figure S1. Calculated N DOS (total and p_z contribution) for the three configurations shown in **Fig. 4c-e.**

ESI2. C 1s and N 1s X-ray photoemission spectra

The photoemission studies were performed at the SuperESCA beamline [1, 2] of the Elettra third generation synchrotron radiation source in Trieste, Italy. The experimental chamber was equipped with a Phoibos hemispherical energy analyzer (SPECS GmbH) with a homemade delay-line detector [3], and has a background pressure of about 2×10^{-10} mbar. Core level spectra were recorded with the sample at room temperature, and at a photon energy of 320 eV (Au 4f), 400 eV (C 1s), or 500 eV (N 1s), with an overall energy resolution better than 100 - 250 meV. Core level spectra binding energies were referenced to the Fermi level. XP spectra were used to verify surface cleanliness. X-ray overlayer damage was minimized by probing the sample at different surface locations for each measurement.
Figure S2 shows the behaviour of the N 1s and C 1s regions after BTAH dosing at room temperature to yield saturation coverage, and with subsequent annealing to the stated temperatures.
After dosing, the N 1s region shows a signal which is best fitted using a doublet with maxima at 400.4 eV and 399.4 eV (FWHM 0.85 eV) and an area ratio of 1.7:1. The two components represent C-N and N-H environments respectively [4, 5]. Different works report the N 1s exhibiting a single peak at 399.3 - 400.5 eV for N coordinated with different metal ions [6 - 10]. Finšgar and co-workers [11] recorded a single N 1s peak at 400.5 eV for adsorption from solution (shifting to 398.5 eV after sputtering in vacuum, the second peak tentatively attributed to molecular decomposition due to Ar^+ damaging). The observation of a single peak was considered to indicate that all nitrogen atoms have the same chemical environment. A shift toward slightly higher binding energy is usually observed, implying that the N_1 hydrogen has been removed in favour of coordination with a metal ion. This consideration has been often used as an argument supporting polymerization of a $M(BTA)_n$ complex [9 - 12], where M is Cu, Ni, Zn, Fe. The presence of a doublet in the N 1s region in our case is strong evidence for the molecule not decomposing either during vacuum sublimation [13], or as a result of interaction with the gold substrate. In contrast, when dosed on reactive substrates [5 - 12], BTAH loses its hydrogen. In the case of Cu(111), the reaction occurs with freely diffusing copper atoms [14]. The very reactive displaced hydrogen reconstructs the Cu(111) surface, rather than recombining into molecular hydrogen and desorbing [15, 16].
After exposure to BTAH, the C 1s region is characterised by a signal which is best fitted using a doublet with maxima at 285.3 eV and 284.3 eV (FWHM 0.68 eV) and an area ratio of 1:2.2. The two components are associated with C-N and C-C environments respectively [4, 5]. For a different variety of systems prepared from an aqueous solution of BTAH, a complex C 1s signal involving components attributed to aliphatic carbon –C–C– or –C–H (285.0 eV), and carbon bonded to oxygen –C–O– (286.5 eV), carboxyl carbon –COO− (at 288.8 eV) species, is usually seen [9, 12]. In the present work, no signals related to carbon-oxygen species were detected, as also confirmed by a featureless signal in the O 1s region.
Both regions show a small decrease in intensity, and increase of FWHM (0.73 eV for C and 0.93 eV for N) for the fitted peaks, with annealing to 340 K. A much larger decrease in intensity and increase in FWHM (1.05 eV for C and 1.13 eV for N) are seen after annealing to 360 K. In particular, the components at higher binding energy are preferentially lost at lower temperatures.

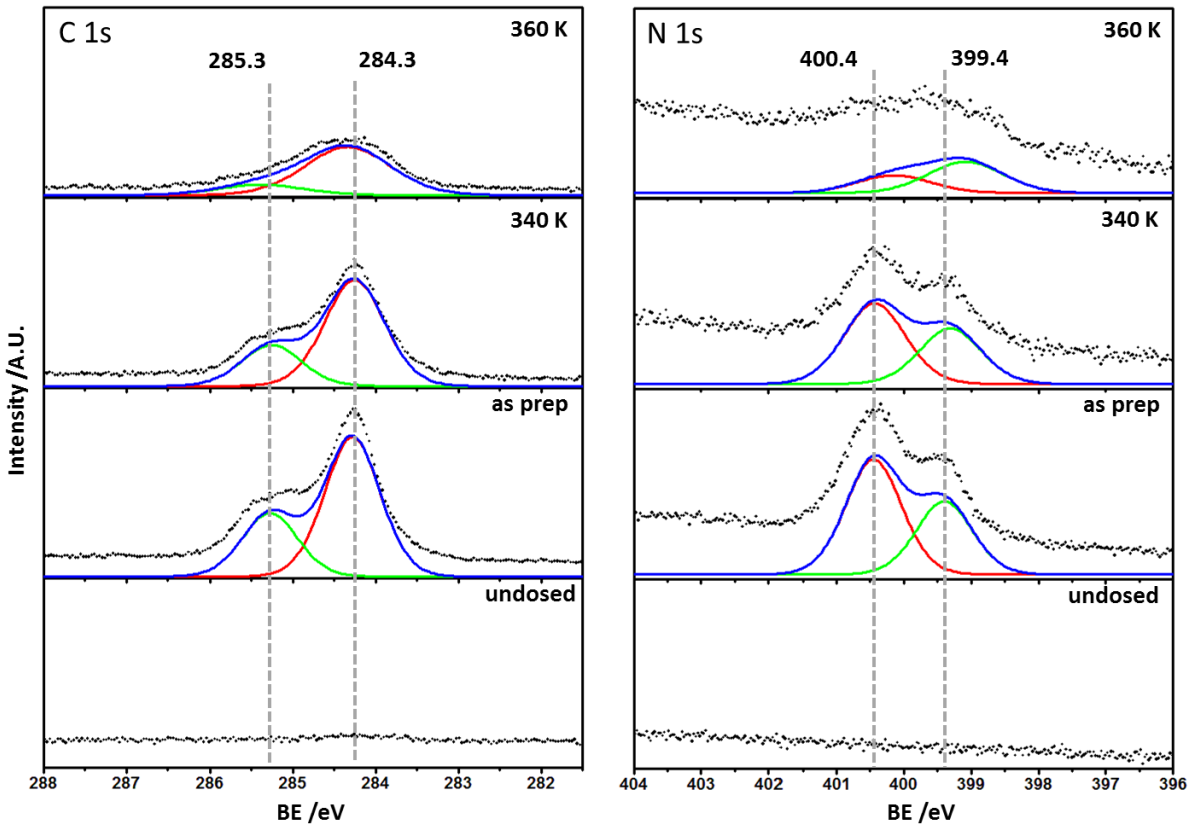


Figure S2. C 1s (left) and N 1s (right) regions recorded for the as clean Au(111) substrate, the as prepared saturated BTAH/Au(111) layer and after annealing to 340 K and 360 K.

ESI3. Additional STM images

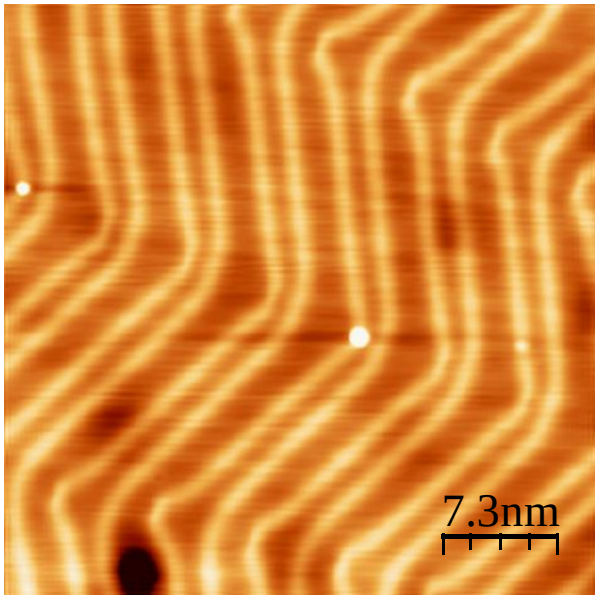


Figure S3.1. STM image of clean Au(111), prior to exposure to BTAH, obtained at 77 K. 36.5 × 36.5 nm^2, -1.643 V, 0.120 nA.

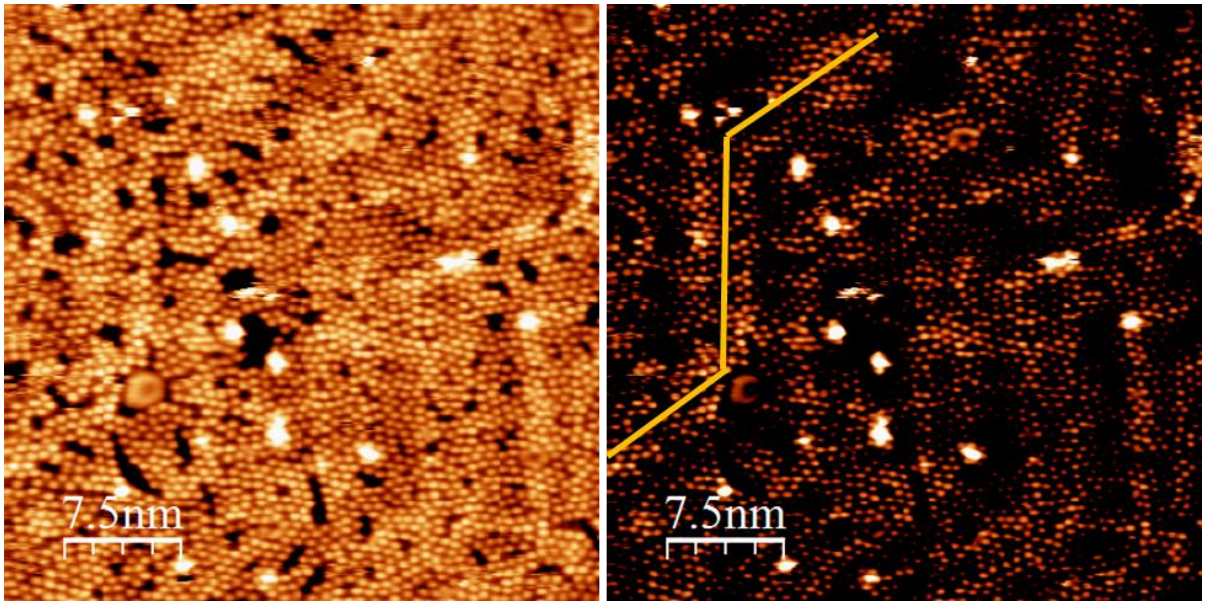


Figure S3.2. STM image acquired at 77 K (with enhanced contrast) highlighting the unperturbed 22 × √3 herringbone reconstruction present underneath the saturated BTAH overlayer prepared at room temperature. Molecules do not show any preferred orientation in relation to the substrate. 37.4 × 37.4 nm^2, -0.71 V, 0.07nA.

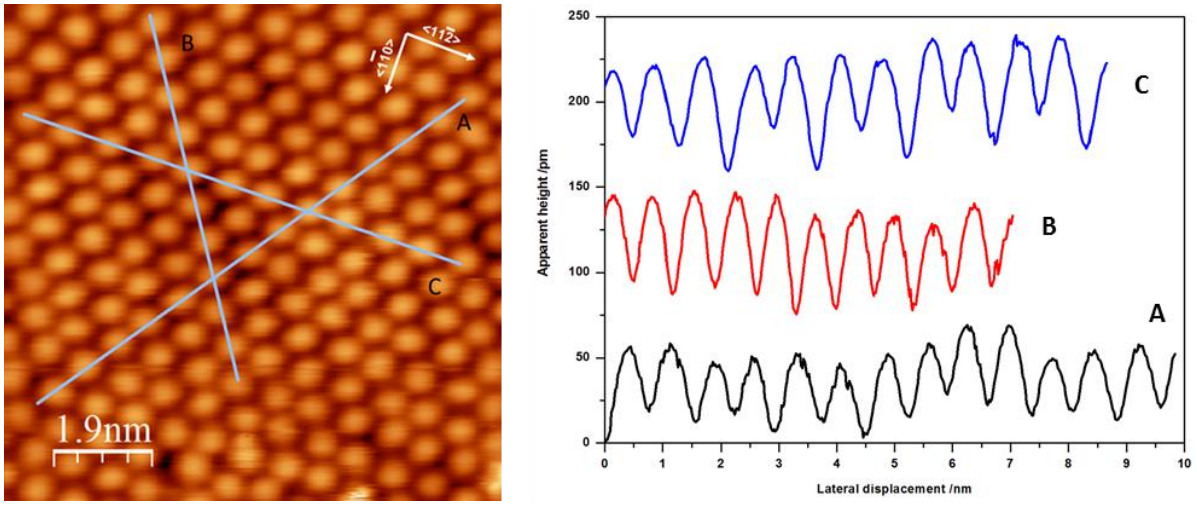


Figure S3.3. High magnification STM image obtained at 77 K of an as-prepared saturated surface (293 K), and line profiles taken along the three equivalent $[11\bar{2}]$-type directions showing different spacing. This points to a chain-like organisation. The average peak-to-peak separation along profile A is 0.87 nm; along B is 0.68 nm, and along C the shorter separation is 0.71 nm while the larger is 0.89 nm. 9.3 × 9.3 nm^2, 1.07 V, 0.1 nA.

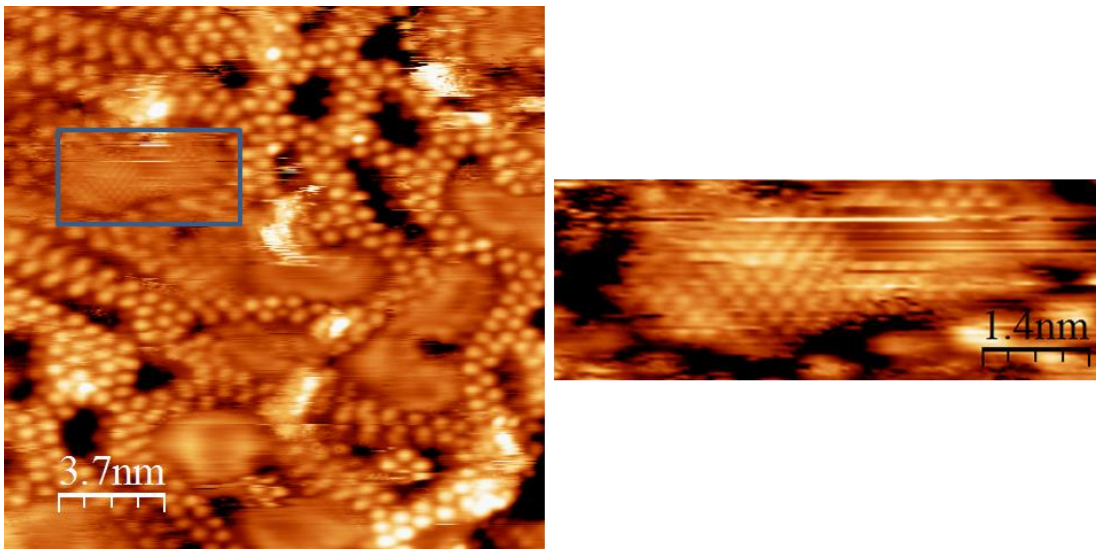


Figure S3.4. STM images obtained at 77 K after annealing a BTAH saturated surface for *ca.* 5 minutes to 313 K; 18.7 × 18.7 nm^2, 1.076 V, 0.14 nA; and magnification of the area in the blue rectangle.

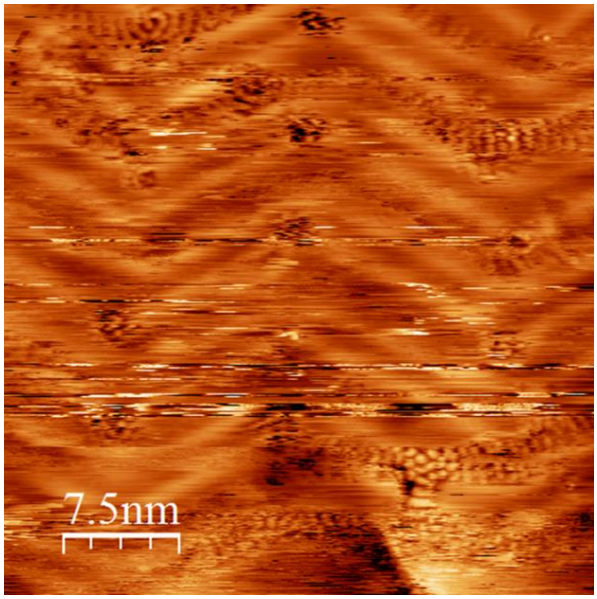


Figure S3.5. STM image obtained at 77 K after annealing a BTAH saturated surface for *ca.* 5 minutes to 323 K; the Au(111) reconstruction is visible over the entire image, some molecular arrangement can still be seen along with streakiness due to highly mobile fast diffusing species; 37.4 × 37.4 nm^2, 1.075 V, 0.079 nA.

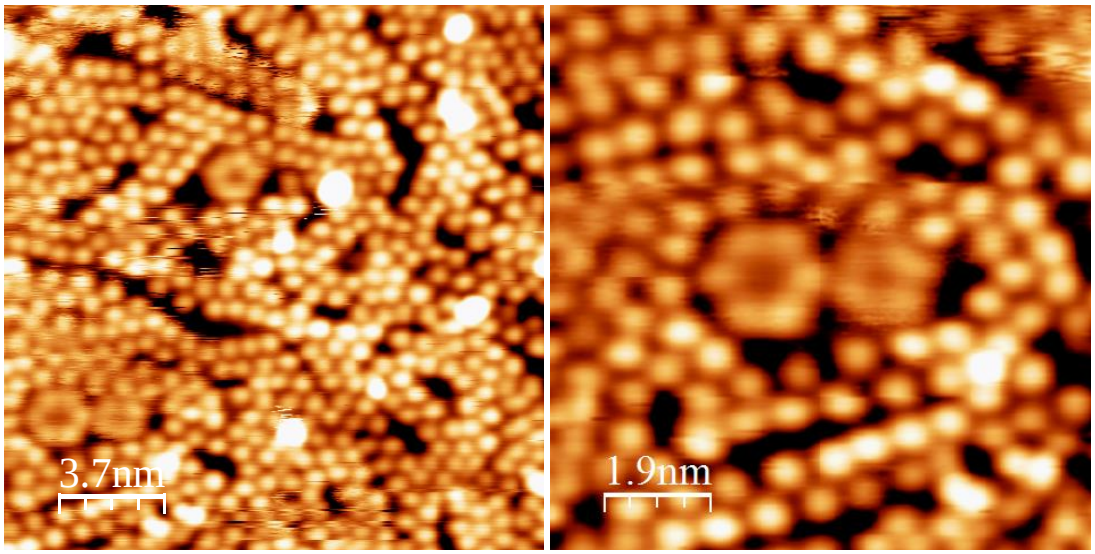


Figure S3.6. Six-fold and five-fold features observed after annealing a BTAH saturated surface for *ca.* 5 minutes to 298 K. STM images obtained at 77 K; 18.7 × 18.9 nm^2,-1.503 V, 0.037 nA; 9.3 × 9.3 nm^2, -1.501 V, 0.037 nA.

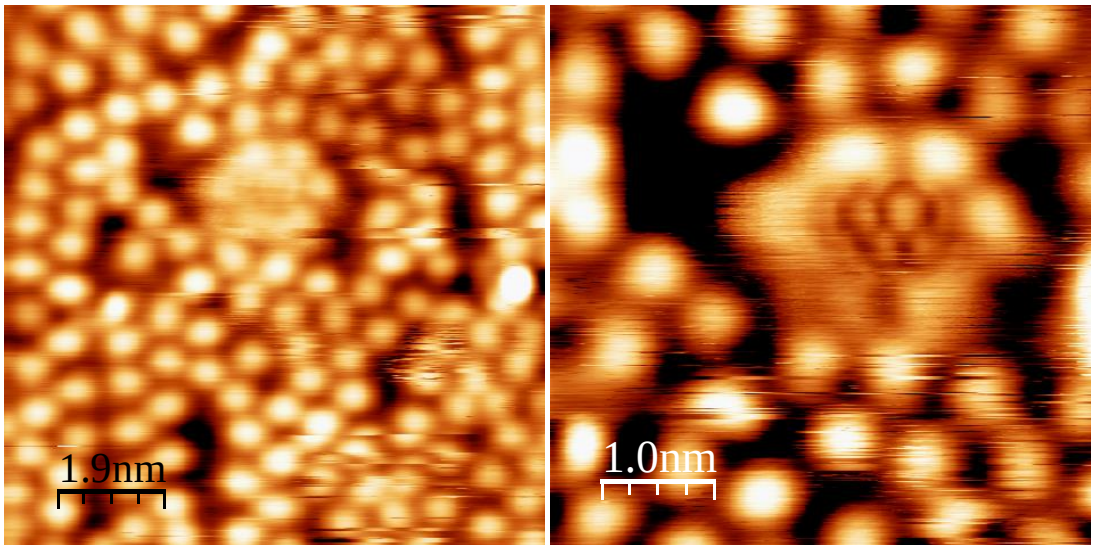


Figure S3.7. STM images obtained at 77 K after annealing a BTAH saturated surface for *ca.* 5 minutes to 298 K; 9.3 × 9.3 nm^2, 0.52 V, 0.02 nA; 4.7 × 4.7 nm^2, -0.1 V, 0.02 nA. The interference pattern inside the 6-fold feature is tentatively interpreted as a molecule trapped inside the hexagonal hole, hopping between equivalent adsorptions sites.

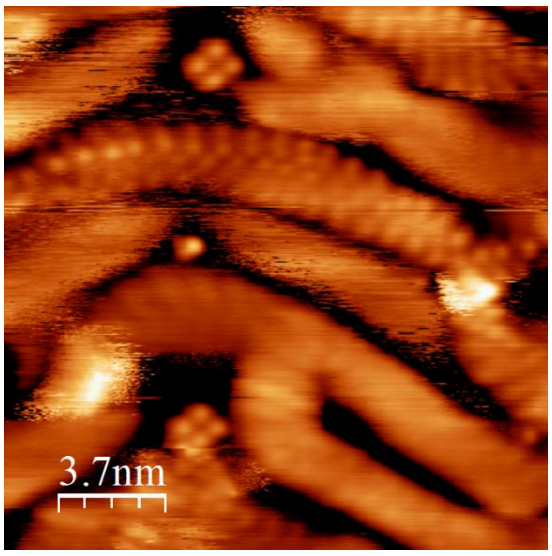
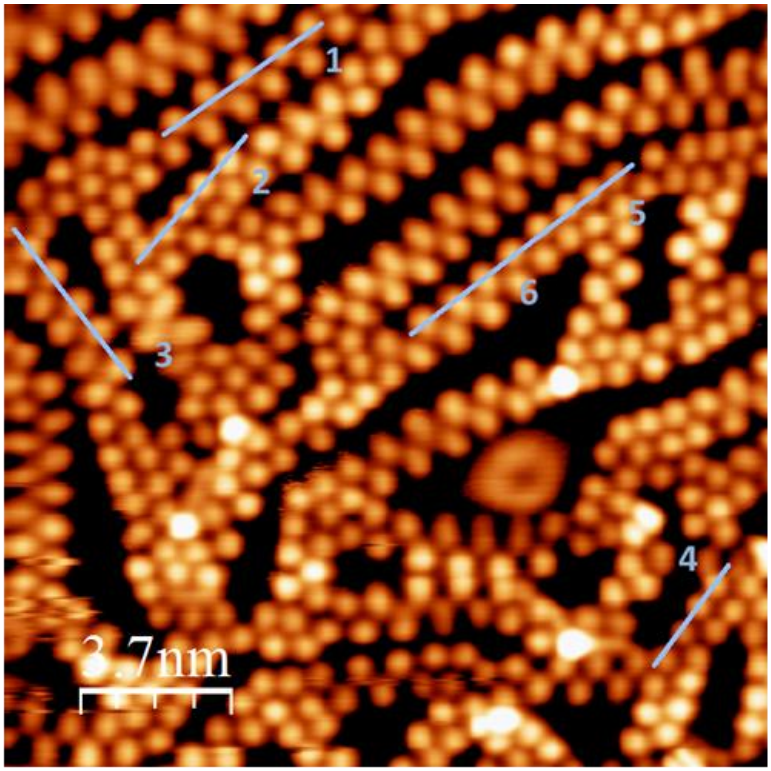


Figure S3.8. STM image obtained at 77 K after annealing a BTAH saturated surface for *ca.* 5 minutes to 318 K; molecular ribbons confined within the Au(111) herringbone elbows and 4-fold features at elbows. 18.7 × 18.7 nm^2, 1.052 V, 0.02 nA.



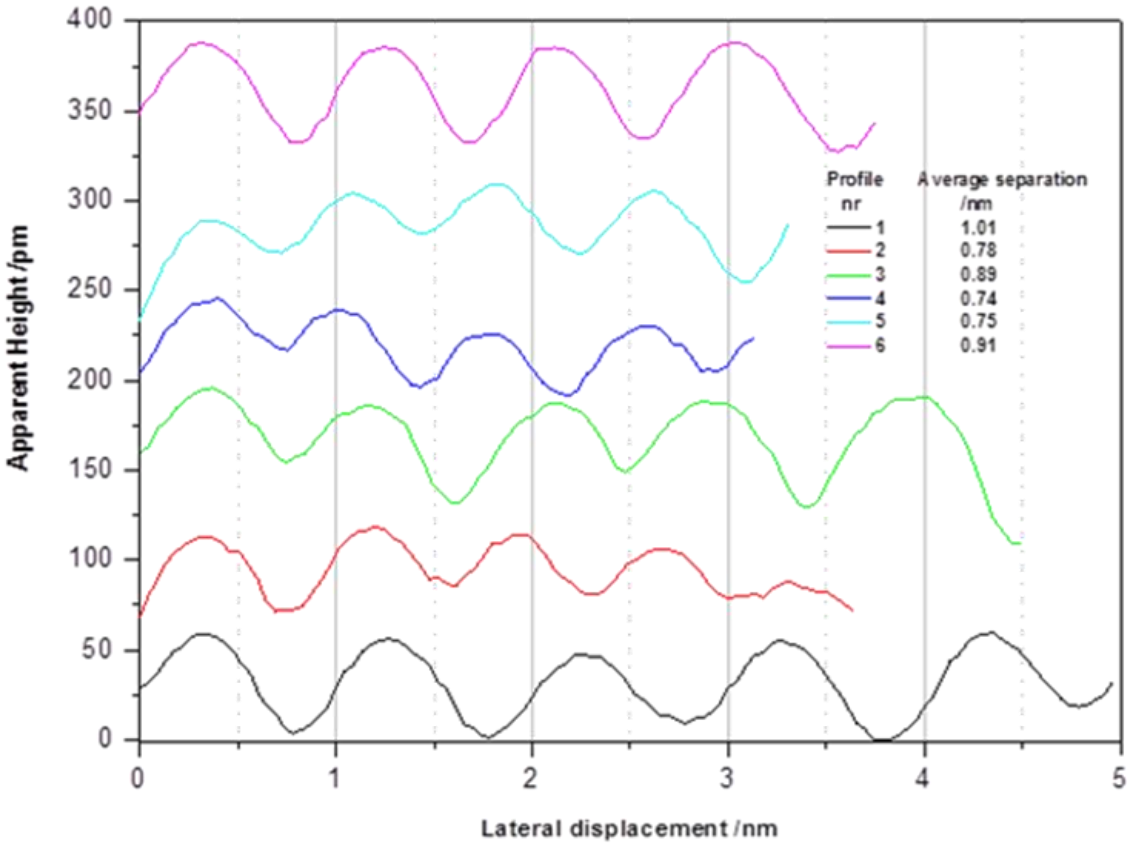


Figure S3.9. STM image obtained at 77 K after annealing a BTAH saturated surface for *ca.* 5 minutes to 308 K; similar area as Fig. 4a in the manuscript; line profiles for different chain topographies. 18.7 × 18.7 nm^2, 1.033 V, 0.077 nA.

ESI4. DFT calculations of BTAH in the gas phase

A variety of functionals implemented in VASP [17] was employed to calculate the difference in energetics between the two BTAH isomers (shown in **Scheme 1a**, manuscript). These include GGA, PBE-D3 [18] (with Becke-Johnson damping [19]), the non-local correlation functionals, vdW-DF[20] and vdW-DF2[21], which account for dispersion inherently, and the hybrid functional HSE06[22], which includes a portion of exact Hartree-Fock exchange, screened by an error function.

The differences in energy between the two isomers in the gas phase, shown in **Table S1**, are calculated using Eq. 1:

$$\Delta_{1\text{-}2} = E_{1\text{H-BTA}} - E_{2\text{H-BTA}} \qquad (1)$$

The calculations show that for purposes the two isomers in the gas phase can be treated as degenerate, whereby the differences of stability between each other are below the accuracy of the method. Therefore when the molecules are adsorbed on a substrate, the interaction with the surface represents the main contribution to the difference in energetics, lifting the pseudo-degeneracy of the two tautomers in the gas phase.

Table S1. Calculated relative stability of BTAH tautomers. Positive values mean 2H-BTA is more stable

functional	**$\Delta_{1\text{-}2}$ /eV**
PBE-D3(BJ)	+0.009
PBE-D3(BJ)(ZPE)*	-0.017
vdW-DF	-0.001
vdW-DF2	-0.008
HSE06	+0.003

*Zero-point energy (ZPE) correction was considered by calculating the Hessian Matrix using finite differences.

ESI5. DFT calculations of BTAH dimers adsorbed on Au(111)

In order to rationalise the molecular arrangements observed experimentally, DFT calculations for four different configurations were performed using the VASP code [17]. Those were selected as the most likely to occur from the interaction of two BTAH tautomers. The first three types (**Figure S4**, Type 1-3) are discussed in the manuscript, whilst the fourth configuration (**Figure S4**, Type 4) is not considered in the manuscript because this configuration does not allow extension into one-dimensional structures. However, this configuration may occur and represent a discontinuity point in the propagation of a chain.

All geometries were optimised using two different functionals, PBE-D3(BJ)[19] and vdW-DF[20]. Dimers were optimised in (6 × 6) unit cells and a vacuum region > 15 Å was included. The integration of the first Brillouin zone is sampled using 3 × 3 × 1 k-points. All structures were relaxed using the CG method with a convergence criterion for the self-consistent cycle of 10^{-6} eV, with residual forces smaller than 0.015 eV/Å.

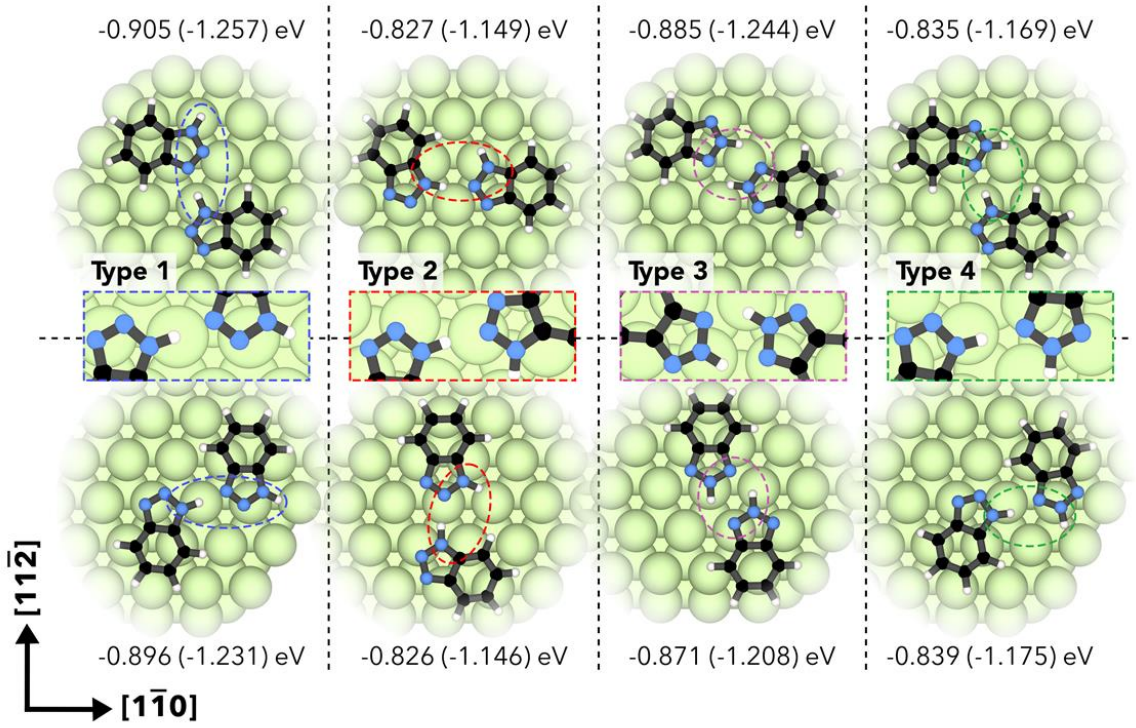


Figure S4. Optimised structures and vdW-DF (PBE-D3(BJ)) adsorption energies for the four types of configurations considered.

In terms of stability, type 1, 2, and 3 configurations are found to adsorb on Au(111) in a way that favours chain growth along $[11\bar{2}]$-type directions, independently of the functional employed. For a type 2 configuration, the difference between the two functionals is minimal. While type 4 configurations are more stable than type 2, type 4 do not generate a dimer that can propagate to form a chain, so they are unlikely to form a major part on the observed structures. Nevertheless, such configuration can occur within a chain and alter the alternation sequence of the different tautomers.

The computed energetics enable an insight into the preferential adsorption sites of the different dimers and growth directions of the different chains, for those dimers allowing chain propagation. This is particularly the case for type 1. On the surface, the type 3 configuration is energetically much more favourable than type 2, even though 2H-BTA is considered a minority species in the condensed phase [23]. Therefore, adsorption on the surface allows for the possibility of having the 2H-BTA isomer stabilised by the substrate, taking also into account that type 3 would have more favourable adsorption energy per unit area than type 1 when forming chains, as shown in Table 3 of the manuscript.

ESI6. Vibrational spectra calculations

Vibrational spectra using the Born Effective Charges (BEC) matrix are calculated using density functional perturbation theory (DFPT) intrinsically implemented in VASP [17]. The intensities are calculated similarly to the dipole approximation proposed by David Karhánek [24]:

$$I(\omega) = \sum_{\alpha=1}^{3} \left| \sum_{l=1}^{M} \sum_{\beta=1}^{3} {Z_{\alpha\beta}}^{*}(l)\, e_{\beta}(l) \right|^{2} \qquad (2)$$

where $e_{\beta}(l)$ is the normalised eigenvector of the ω-th mode. α and β range over the three Cartesian directions for the (l = 1-M) different atoms composing the system. $Z_{\alpha\beta}$ *(l) are the ionic effective charges and can be understood as a coefficient which is proportionally related to the change of the polarisation when displacing an atom in the different directions in the absence of an external electric field.

Typically, the modes producing a polarisation in the z direction, perpendicular with respect to the substrate which lays on the (x, y)-plane, are detected with more intensity than the remaining ones, despite the fact that the other two components can play an important role in the resulting simulated spectra. In the mathematical formalism proposed here, α can be projected in its components x, y and z independently. That means (Z_x,β), (Z_y,β) and (Z_z,β) can be understood as a polarisation change in x, y and z produced by the displacement in one of the three directions (β). Therefore, the total intensity can be separated in terms of the three directions; thus, the different components can be weighted independently:

$$I(\omega) = a \left| \sum_{l=1}^{M} \sum_{\beta=1}^{3} {Z_{x\beta}}^{*}(l)\, e_{\beta}(l) \right|^{2} + b \left| \sum_{l=1}^{M} \sum_{\beta=1}^{3} {Z_{y\beta}}^{*}(l)\, e_{\beta}(l) \right|^{2} + c \left| \sum_{l=1}^{M} \sum_{\beta=1}^{3} {Z_{z\beta}}^{*}(l)\, e_{\beta}(l) \right|^{2} \qquad (3)$$

For isotropic systems, because the matrix is diagonal, only displacements in the same direction as the polarisation change are considered to be non-zero ($\alpha \neq \beta \rightarrow {Z_{\alpha\beta}}^{*}(l) = 0$), and hence the intensity result is:

$$I(\omega) = a \left| \sum_{l=1}^{M} {Z_{xx}}^{*}(l)\, e_{x}(l) \right|^{2} + b \left| \sum_{l=1}^{M} {Z_{yy}}^{*}(l)\, e_{y}(l) \right|^{2} + c \left| \sum_{l=1}^{M} {Z_{zz}}^{*}(l)\, e_{z}(l) \right|^{2} \qquad (4)$$

x, y, z components and total spectra calculated for the three types of interactions (see scheme 1), are reported in **Figure S5.1,** with a = b = c = 1.

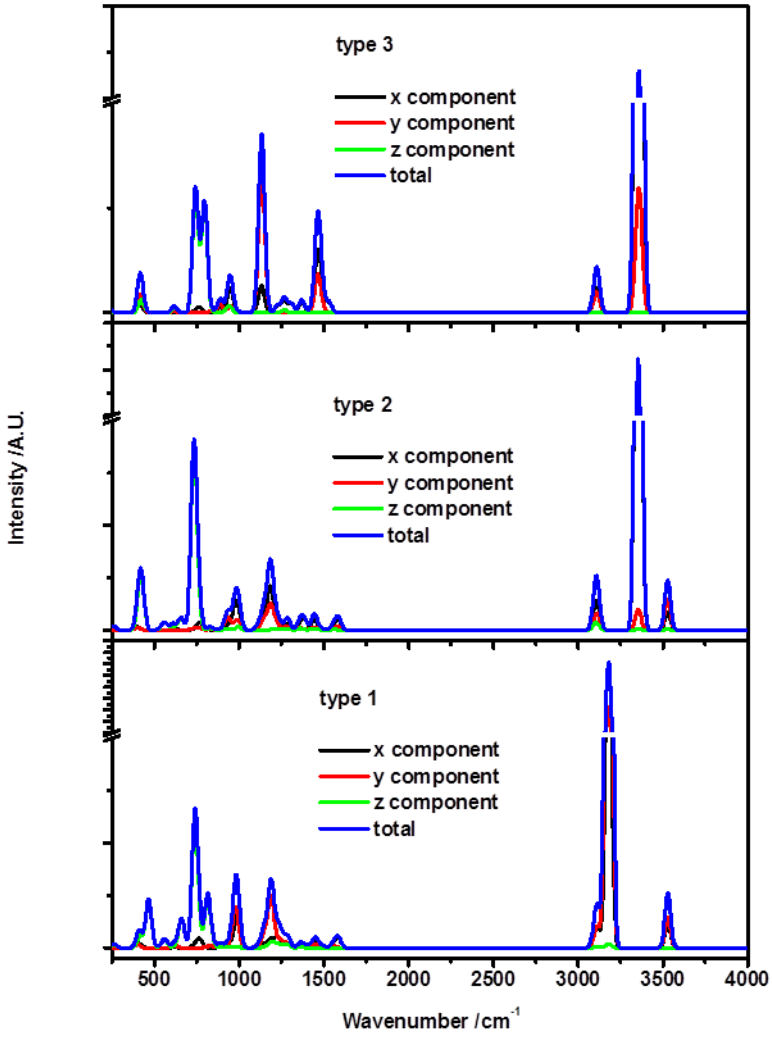


Figure S5.1. x, y, z components and total trace of the vibrational spectra calculated for type 1, 2, and 3 interactions as in Figure S4.

The weights of the coefficients *a*, *b,* and *c* are subsequently optimised to reproduce the dipole contribution separately for each component involved in HREELS. In this case, *a* and *b* are considered to be 0.01, hence *c* is set as 0.98, which means that a calculated spectrum will have 98% contribution derived from the component normal to the surface. Gaussian broadening of 40 cm^{-1} FWHM is applied to the calculated spectra, as this is the resolution typically attained for enhanced experimental spectra.

Experimental HREEL spectra can be best fitted using contributions of these three spectra which, in the first approximation, correspond to the amount of species coordinated through type 1, 2, and 3 interactions. In principle, the ratio used in the fitting should reflect the relative amount of different chains seen on STM. However, the area imaged by STM and the one scanned by the HREELS are

different, as HREELS gives an average result on a wider area. Therefore some discrepancy with the ratio that may be determined by analysing STM images may be expected. This does not contradict the general idea that one needs contributions from the three types of chains to fit the experimental spectrum. **Figure S5.2** shows the fitting of the as-prepared experimental HREEL spectrum as an example.

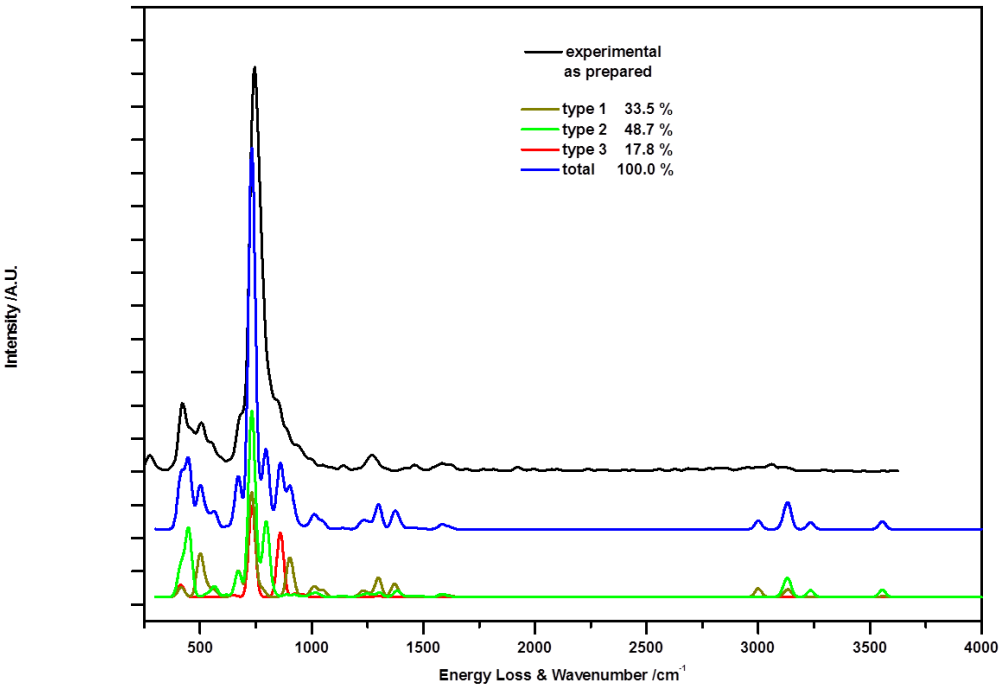


Figure S5.2. Fitting of the HREEL spectrum obtained by dosing BTAH on Au(111) to saturation coverage at room temperature as in Fig 1. The best fit is obtained by weighting differently the spectra calculated for the three types of interactions.

ESI7. References

1. A. Baraldi, M. Barnaba, B. Brena, D. Cocco, G. Comelli, S. Lizzit, G. Paolucci and R. Rosei, *J. Electron. Spectrosc. Relat. Phenom.*, 1995, **76**, 145.
2. A. Baraldi, G. Comelli, S. Lizzit, M. Kiskinova and G. Paolucci, *Surf. Sci. Reports*, 2003, **49**, 169.
3. G. Cautero, R. Sergo, L. Stebel, P. Lacovig, P. Pittana, M. Predonzani and S. Carrato, *Nucl. Instrum. Meth. A*, 2008, **595**, 447.
4. J. F. Moulder, W. F. Stickle, P. E. Sobol and K. D. Bomben, Handbook of X-ray Photoelectron Spectroscopy, Physical Electronics Inc., Eden Prairie (Minnesota), USA 1995
5. T. Kosec, D. K. Merl and I. Milošev, *Corr. Sci.*, 2008, **50**, 1987.
6. D. Chadwick and T. Hashemi, *Corros. Sci*., 1978, **18**, 39.
7. H. G. Tompkins and S. P. Sharma, *Surf. Interface Anal.*, 1982, **6**, 261.
8. G. Xue, J. Ding, P. Lu and J. Dong, *J. Phys. Chem*., 1991, **95**, 7380.
9. I. A. Arkhipushkin, L.I. Yesina, Yu. Ya. Andreev, L. P. Kazansky and Yu. I. Kuznetsov, *Int. J. Corros. Scale Inhib*., 2012, **1**, 107.
10. A. Mezzi, E. Angelini, T. De Caro, S. Grassini, F. Faraldi, C. Riccucci and G. M. Ingo, *Surf. Interface Anal.: European Applications of Surface and Interface Analysis*, 2012, **44**, 968.
11. M. Finšgar, J. Kovač and I. Milošev, *J. Electrochem. Soc*., 2010, **157**, C52.
12. T. Hashemi and C. A. Hogarth, *Electrochim. Acta*, 1988, **33**, 1123.
13. a) Y. Park, S. H. Choi, H. Noh and Y. Kuk, *Proc. of SPIE*, 2006, **6172**, 617214-1. b) Y. Park, H. Noh, Y. Kuk, K. Cho and T. Sakurai, *J. Korean Phys. Soc.*, 1996, **29**, 745.
14. a) F. Grillo, D. W. Tee, S. M. Francis, H. Früchtl and N. V. Richardson, *Nanoscale*, 2013, **5**, 5269; b) F. Grillo, D. W. Tee, S. M. Francis, H. A. Früchtl and N. V. Richardson, J. Phys. Chem. C, 2014, **118**, 8667.
15. C. Gattinoni and A. Michaelides, *Faraday Discuss*. 2015, **180**, 439.
16. a) M. F. Luo, G. R. Hu and M. H. Lee, *Surf. Sci*., 2007, **601**, 1461; b) G. Lee and E. W Plummer, *Surf. Sci.*, 2002, **498**, 229; c) M. F. Luo, D. A. MacLaren, I.G. Shuttleworth and W. Allison, *Chem. Phys. Lett.*, 2003, **381**, 654; d) K. Mudiyanselage, Y. Yang, F. M. Hoffmann, O. J. Furlong, J. Hrbek, M. G. White, P. Liu and D. J. Stacchiola, *J. Chem. Phys.*, 2013, **139**, 044712.
17. a) G. Kresse and J. Hafner, *Phys. Rev. B*, 1993, **47**, 558; b) G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169.
18. S. Grimme, J. Antony, S. Ehrlich and S. Krieg, *J. Chem. Phys*., 2010, **132**, 154104.
19. S. Grimme, S. Ehrlich and L. Goerigk, *J. Comp. Chem.*, 2011, **32**, 1456.
20. M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, *Phys. Rev. Lett.*, 2004, **92**, 246401.
21. K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist and D. C. Langreth, *Phys. Rev. B*, 2010, **82**, 081101.
22. a) J. Heyd, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.,* 2003, **118**, 8207; b) J. Heyd and G. E. Scuseria, *J. Chem. Phys.,* 2004, **121**, 1187; c) J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.*, 2006, **124**, 219906.
23. a) C. Borin, L. Serrano-Andrés, V. Ludwig and S. Canuto, *Phys. Chem. Chem. Phys.*, 2003, **5**, 5001; b) F. Tomás, J. Catalán, P. Pérez and J. Elguero, *J. Org. Chem.*, 1994, **59**, 2799.
24. D. Karhánek, T. Bučko and J. Hafner, *J. Phys.: Condens. Matter*, 2010, **22**, 265006.