ChemComm

Chemical Communications

rsc.li/chemcomm
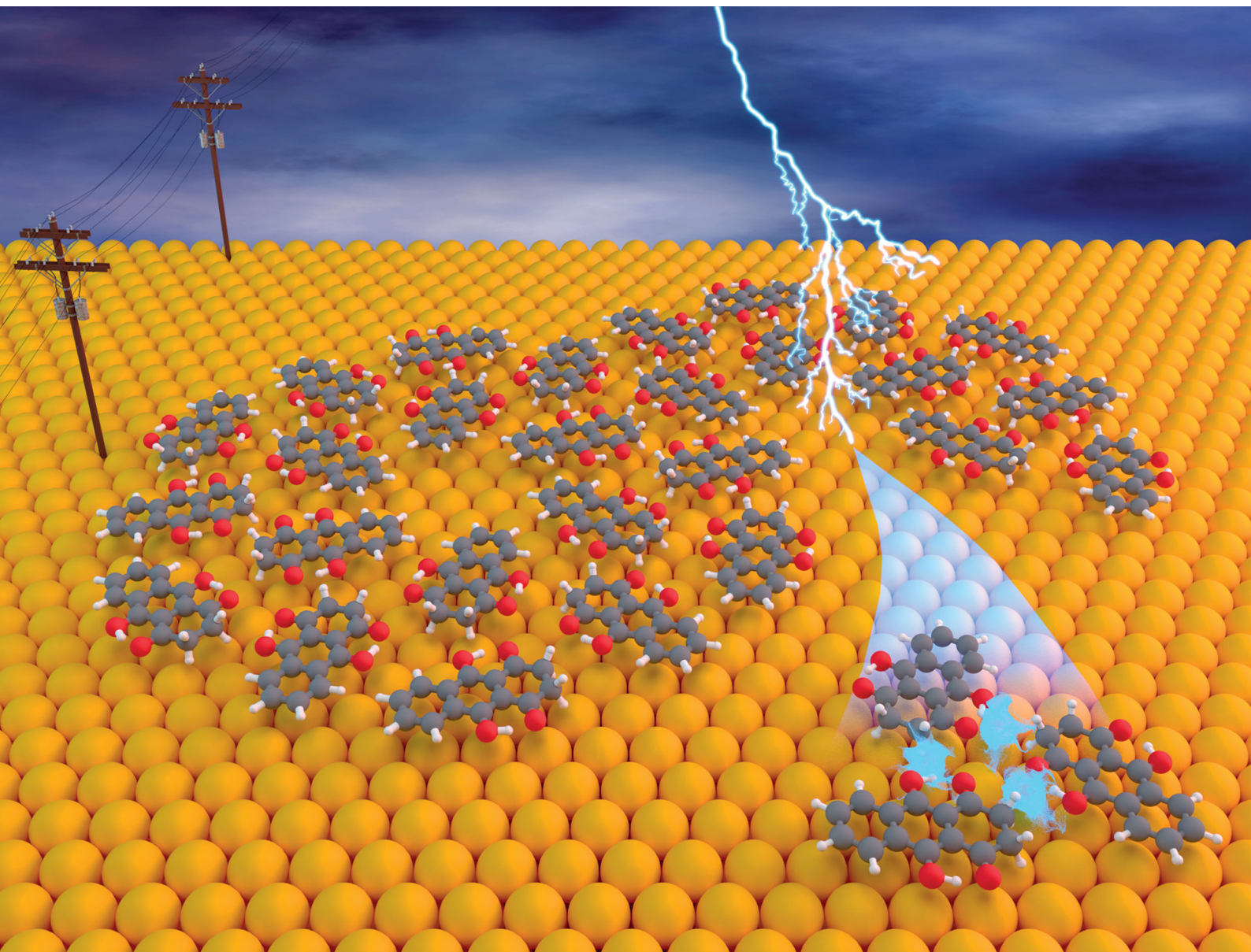
ROYAL SOCIETY
OF CHEMISTRY

COMMUNICATION
Albano Cossaro *et al.*
Keto–enol tautomerization drives the self-assembly of
leucoquinizarin on Au(111)

ROYAL SOCIETY
OF CHEMISTRY

COMMUNICATION

Check for updates

Keto–enol tautomerization drives the self-assembly of leucoquinizarin on Au(111)†

Roberto Costantini, [ID] [ab] Luciano Colazzo, [cd] Laura Batini, [e] Matus Stredansky, [ab] Mohammed S. G. Mohammed, [cd] Simona Achilli, [e] Luca Floreano, [ID] [b] Guido Fratesi, [ID] [e] Dimas G. de Oteyza [cdf] and Albano Cossaro [ID] *[b]

The self-assembly of leucoquinizarin molecules on Au(111) surfaces is shown to be characterized by the molecules mostly being in their keto–enolic tautomeric form, with evidence of their temporary switching to other tautomeric forms. This reveals a metastable chemistry of the assembled molecules, to be considered for their possible employment in the formation of more complex hetero-organic interfaces.

Hydroxyl or oxygen containing derivatives of anthracene are an important class of molecules, widely employed in industrial processes such as dye production,[1,2] drugs synthesis,[3] and others.[4] They are characterized by a fast redox response, which makes them convenient building blocks for the design of cathodes of a new generation of batteries[5–7] and for the synthesis of complex hybrid interfaces for photovoltaic devices.[8] Their color and in general their electronic properties can be finely tailored by varying the functionalization of the aromatic body.[9] The presence of one or more oxygen terminations promotes the occurrence of keto–enol tautomerism in these compounds, which can be of relevant importance for their functionality. For instance, in complexes formed by crown ethers and anthracene derivatives, the ion complexation of the former acts as a switch of the tauto-merism of the anthracene group and promotes its luminescence, making the complexes suitable for the design of ion sensors.[10–13] As a matter of fact, most of the employed anthracene derivatives are based on 9,10-anthraquinone, an anthracene with double bonded

oxygen atoms substituting the 9,10 positions. We report here a combined imaging, spectroscopic and theoretical study of the assembly on Au(111) surfaces of leucoquinizarin (LQZ, $C_{14}H_{10}O_4$), an anthracene with four hydroxyl groups in positions 1, 9, 4, and 10 (Fig. 1), adopted in the chemical industry for the synthesis of dyes. The molecule goes through keto–enol tautomerism and some of the tautomers are represented in Fig. 1. The presence of pairs of hydroxyl groups positioned in a double diol configuration makes LQZ a potential building block for extended covalent networks. In particular, one could exploit its tendency toward aldol condensation with hydroxyl molecules,[14] as well as its covalent coupling with boronic molecules upon a diol-boronic condensation reaction.[15] The latter perspective would be of interest due to the increasing importance that boronic chemistry has gained in recent years as a tool for the synthesis of complex 2D interfaces.[16–18] In the view of exploring these synthesis routes to possibly exploit the electronic and optical properties of anthraquinone derivatives in more extended architectures, the chemistry of LQZ on surfaces needs to be properly investigated first. Both experimental[19] and theoretical[20] findings indicate the T1 configuration as the most stable among the possible tautomers depicted in Fig. 1, both in the gas phase and in solution. Our findings support T1 to be the LQZ form also on the Au(111) surface. The room temperature growth of the LQZ sub-monolayer is characterized by differently ordered assemblies. The imaging of the most ordered phase is reported in Fig. 2 and exhibits a molecular assembly characterized by the presence of chiral pores, spaced by 2.5 nm. The morphology resembles the
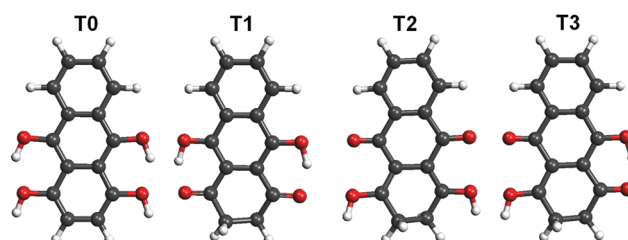
[a] Physics Department of University of Trietse, via A. Valerio 2, 34127 Trieste, Italy
[b] CNR-IOM, Area Science Park, Strada Statale 14, km 163,5, 34149 Trieste, Italy.
E-mail: cossaro@iom.cnr.it
[c] Donostia International Physics Center, Paseo Manuel Lardizabal 4,
E-20018 Donostia-San Sebastián, Spain
[d] Centro de Física de Materiales (CSIC-UPV/EHU) – MPC,
Paseo Manuel de Lardizabal, 5 – E-20018 Donostia-San Sebastián, Spain
[e] Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano,
Milano, Italy
[f] Ikerbasque, Basque Foundation for Science, E-48011 Bilbao, Spain

† Electronic supplementary information (ESI) available: Experimental and theoretical details, NEXAFS spectra, additional STM images, and DFT calculation. See DOI: 10.1039/c9cc09915h


Fig. 1 Molecular structure of four tautomers of the LQZ molecule.
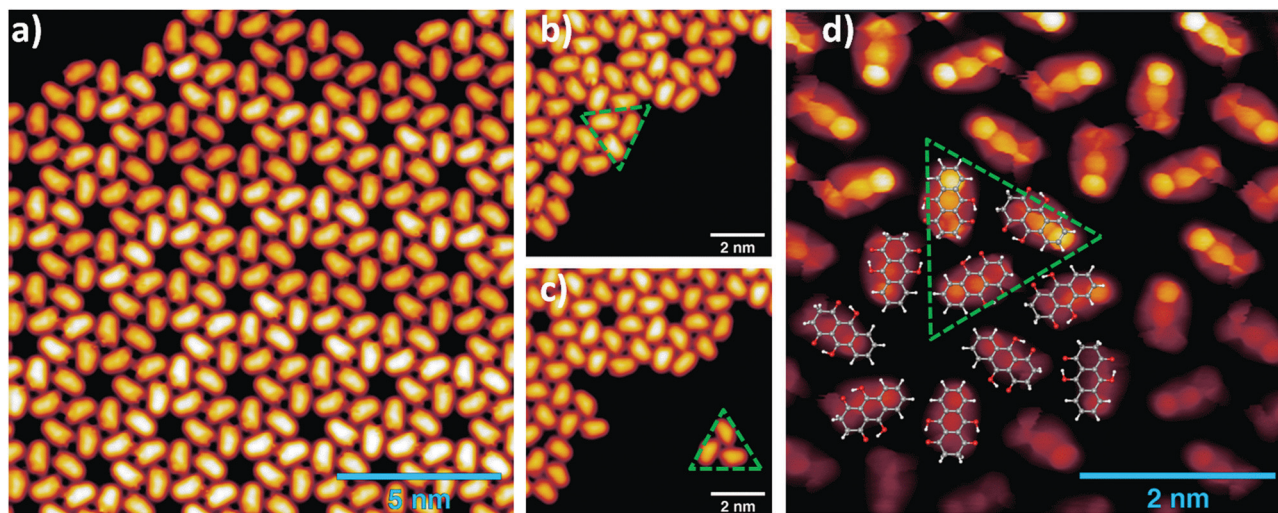
Fig. 2 (a) STM images of a LQZ submonolayer showing the presence of nanometric pores (0.05 V, 100 pA). (b and c) Tip manipulation allows the separation of a trimer from the assembly, which is identified as the building block of the architecture (0.1 V, 50 pA). (d) Bond resolution image reveals a distorted termination of the LQZ molecules, which can be related to the diketo phenol of the T1 tautomer. The superimposed overlayer model results from the DFT calculations of the assembly, obtained as detailed in the ESI.†

giant honeycomb architecture found for the anthraquinone molecule on Cu(111).[21] Here however, the presence of additional oxygen terminations with respect to anthraquinone likely promotes a more compact assembly. Domains with both chiralities are observed on the same surface. The supramolecular structure corresponds to a hierarchical assembly with molecular trimers as building units, which remain together upon controlled molecular manipulation with the scanning probe (Fig. 2b and c). As already mentioned, the ordered assembly of Fig. 2, observed for low coverage (< 0.5 ML), is not the only morphology obtained in the preparations. At higher coverage, or with larger deposition rate, the nanopore domains become less regular due to an increase of the molecular density, finally yielding a disordered phase at completion of the first layer (ESI,† Fig. S1). Interestingly, molecular trimers, even though characterized by a less regular geometry than those reported in Fig. 2, can still be identified as the recurring motif of the architectures. High resolution imaging with a CO functionalized tip is shown in Fig. 2d and allows observation of more details of the LQZ assembly and of the intermolecular interactions. All the molecules display the same aspect: one bright hexagonal carbon ring at one end, a darker and more irregular hexagonal ring in the center, and an even darker and strongly deformed structure at the opposite end. We performed DFT calculations in order to have better insight into both the assembly and the electronic properties of the LQZ molecule. The details of the calculations and of the resulting structures are reported in the ESI.† Our findings confirm the conformer T1 as the most stable form, with the purely hydroxylic T0 having 0.9 eV higher total energy in the gas phase. Moreover, the diketo form of the terminal quinol of T1 introduces a distortion of the anthracene backbone. This is likely the cause of the blurred aspect of one of the molecule terminations as imaged in Fig. 2d and suggests an assembly scheme with the aromatic ring (opposite to the quinol ring) pointing towards the pores. The superimposed molecular model in Fig. 2d is the result of DFT relaxation of a

free-standing LQZ layer, as obtained starting from the structure suggested by STM. The analysis of the intermolecular bonding confirms that the oxygen atoms have a determinant role in the assembly. In particular, the interaction between the keto–enolic terminations of molecules belonging to adjacent trimers drives the assembly of the porous architecture (see the ESI† for details).

A further important aspect of LQZ assemblies, observed by STM measurements, is their unstable imaging. This is shown by way of example in Fig. 3. Panel (b) displays the image of a disordered chain of molecules, measured on a low coverage LQZ layer deposited on Au(111) held at 120 K. Maintaining the probe at the point marked by the blue cross and monitoring the current under open feedback conditions provides a trace as shown in panel (a), characterized by the presence of telegraph noise that is absent on the bare surface. By comparison with similar observations in previous studies, this effect can be related to the switching of the molecule to different tautomeric forms,[22–24] although switching to different adsorption configurations
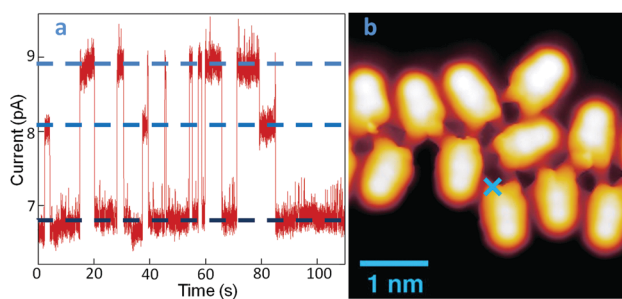
Fig. 3 (a) Tunneling current measured in correspondence of the cross indicated in the image (panel b, −0.15 V, 1 nA) of a very low coverage LQZ layer, as a function of the measuring time. Three different current levels are measured, suggesting a dynamics of the LQZ chemistry in the assembly, possibly due to tautomeric switching.

cannot be excluded either. The curve exhibits two switching levels beside the most stable one, as evidenced by the blue dotted lines in the graph.

However, the switching rate and the different level's relative dominance greatly vary as the tip is positioned differently (Fig. S4, ESI†). While this proves the molecule-related origin of the switching, it also reveals the influence of the location-dependent tip–molecule interactions, in line with other well-known tautomerization switches revealed in porphycenes.[24] Another similarity between the two systems is the promotion of the switching events by vibrational excitations of the molecule, which becomes obvious from the notably increased instabilities at bias values above ±50 meV, and a further increase at around ±350 meV. The symmetric thresholds for both bias polarities are a fingerprint of inelastic tunneling processes that we associate, in the absence of electronic orbitals and transitions at these low energies, to molecular vibrational modes. It should be noted, however, that instabilities in the tunneling gap are observed even at 1 meV (Fig. 3a), whereby tip-induced effects driven by inelastic tunneling excitations are minimized. The DFT energy difference between T1 and the other tautomers, as calculated for a free molecule, is too large for a tautomeric switch to be at the origin of the tunneling current behavior presented in Fig. 3, which has been measured at 4.3 K. However, it has to be considered that experimentally, both the intermolecular hydrogen bonding and the interaction with the substrate can considerably lower this energy barrier. Moreover, transitions to intermediate tautomers, where only one of the enolic terminations of T1 has switched, can be considered at the origin of the telegraph noise, with consequent reduced effort in terms of energy barrier.

To further investigate the chemistry and the morphology of the system we performed X-ray spectroscopy measurements at the ALOISA beamline[25] and at its ANCHOR-SUNDYN endstation[26] at the Elettra Synchrotron. Fig. 4 reports the O and C K-edge NEXAFS spectra taken on an LQZ monolayer. The carbon spectra resemble the anthracene ones,[27] apart from the intensity ratio between the first two resonances, at around 285 eV. Both edges present a strong dichroism, with the lower energy transitions having maximum intensity when the polarization of the electric field is perpendicular to the surface. We calculated the eigenfunctions for the ground state of the four tautomers, as well as the density of states projected onto states of π and σ symmetry. The four conformers exhibit the same symmetry of electronic states around the HOMO. In particular, the HOMO, LUMO and LUMO+1 are π states, whereas σ-symmetry can be found for the HOMO−1 (see the ESI†). The almost perfect dichroism indicates therefore that molecules are adsorbed flat on the surface and that neither the interaction with the substrate, nor the hydrogen bonding scheme the molecules are involved in, introduce relevant distortion of the electronic structure. The lower panels of Fig. 4 report the calculated spectra for the T1 tautomer, which provides a better match to the data than other configurations. Photoemission spectroscopy (XPS) also corroborates the presence of T1 species on the surface. The O1s and C1s peaks are shown in Fig. 5, as measured on an LQZ monolayer, and as calculated for the T1 molecule. No significant differences have been found in the XPS profiles taken on films with lower coverage. The O1s peak presents two broad components, of the same intensity, found at 531.3 eV and 532.7 eV, which can be assigned to carbonyl (=O) and to hydroxyl (–OH) oxygen species respectively.[28] The profile of C1s is more complex and can be fitted by five components. The calculated spectra for T1 nicely reproduce the experimental profiles and confirm the O1s assignment. Regarding the C1s peak, the highest energy component of C1s at ∼289 eV is not reproduced by calculations and can be assigned to energy loss due to inner excitations of the molecule. The C1s components 3 and 4 are due to the carbon atoms linked to the hydroxyl and carbonyl oxygen atoms respectively. We ascribe both the broadness of the O1s components and the minor discrepancies between experimental and calculated C1s profiles to the intermolecular bonds that are not considered by the calculation, as well as to the morphologic disorder of the assembly and to the possible presence of different tautomers on the surface. In this regard, we remark that, the X-ray spectroscopy measurements have been taken at room temperature, whereas STM has been performed at 4 K. We can speculate that these conditions promote a relevant portion of LQZ molecules in a tautomeric form different from T1.
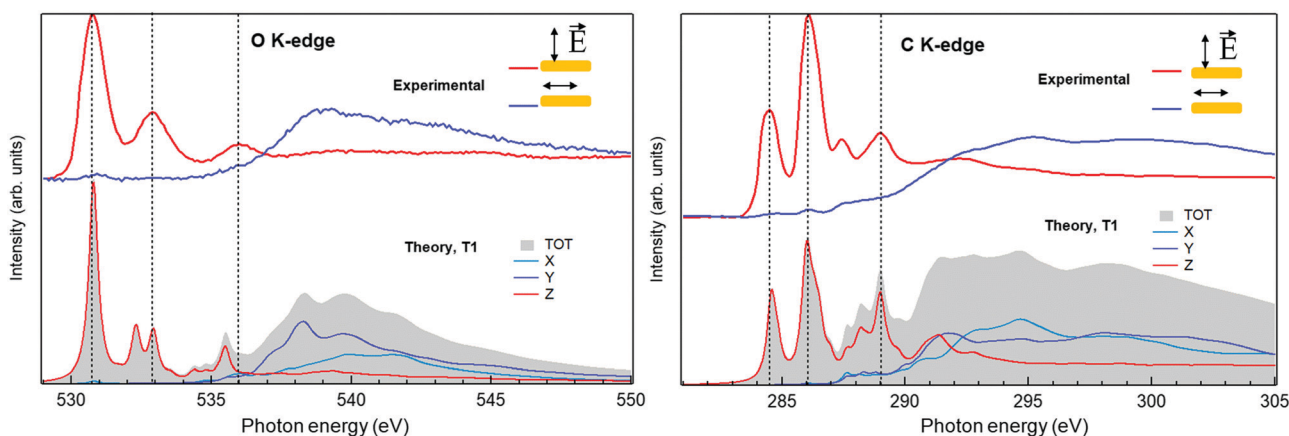

Fig. 4 O1s and C1s XPS spectra measured on the LQZ monolayer (top panels) and simulated for the T1 tautomer (bottom).
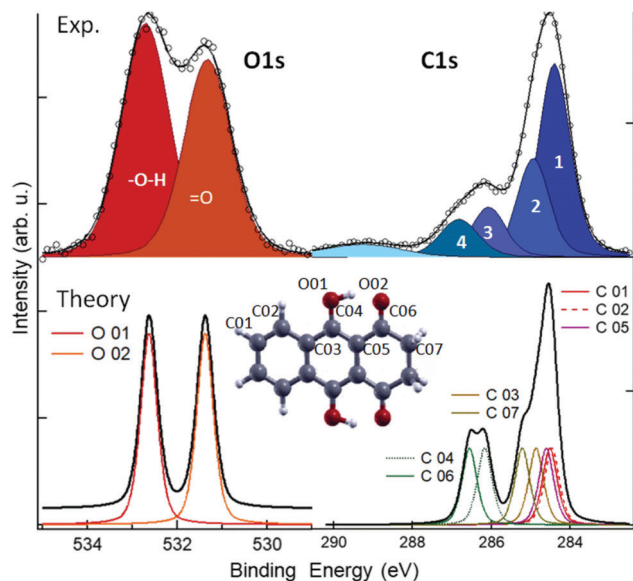
Fig. 5 XPS spectra measured (top) and calculated for the T1 molecule (bottom). See the ESI,† for the comparison with the calculations of the other tautomers.

In conclusion, the assembly of LQZ molecules on Au(111) has been characterized in terms of the morphology and of the electronic properties of the supramolecular architecture. A trimer of hydrogen bonded molecules has been identified as a recurring motif of the assembly and as a building block of differently ordered superstructures. The assembly scheme that can be inferred from the STM images on the highly ordered structures is consistent with the LQZ molecules predominantly in their keto–enolic tautomeric form. The same configuration is compatible with the X-ray spectroscopy results, as simulated by DFT calculations. We ascribe a certain degree of unpredictability of the morphology, as the molecular density increases, to the tautomeric switch the molecules can experience on the surface. This is corroborated by the observation of telegraph noise in the STS measurements. As different tautomers present different chemical properties of their oxygen terminations, our findings have to be taken into account in the perspective of employing LQZ as a building block of more complex interfaces and of the possible on-surface reactions it can be involved in.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1   D. R. Waring and G. Hallas, *The Chemistry and application of dyes*, Plenum Press, 1990.
2   K. Hunger, *Industrial dyes: chemistry, properties, applications*, Wiley-VCH, 2003.
3   J. William Lown, *Pharmacol. Ther.*, 1993, **60**, 185–214.
4   G. Diaz-Muñoz, I. L. Miranda, S. K. Sartori, D. C. de Rezende and M. A. N. Diaz, *Studies in Natural Products Chemistry*, 2018.
5   Z. Song, H. Zhan and Y. Zhou, *Chem. Commun.*, 2009, 448–450.
6   W. Wang, W. Xu, L. Cosimbescu, D. Choi, L. Li and Z. Yang, *Chem. Commun.*, 2012, **48**, 6669–6671.
7   B. Huskinson, M. P. Marshak, C. Suh, S. Er, M. R. Gerhardt, C. J. Galvin, X. Chen, A. Aspuru-Guzik, R. G. Gordon and M. J. Aziz, *Nature*, 2014, **505**, 195–198.
8   M. Castelaın, H. J. Salavagione, R. Gomez and J. Luis Segura, *Chem. Commun.*, 2011, **47**, 7677–7679.
9   D. Jacquemin, J. Preat, M. Charlot, V. Wathelet, J. M. André and E. A. Perpète, *J. Chem. Phys.*, 2004, **121**, 1736–1743.
10 M. Alaparthi, K. Mariappan, E. Dufek, M. Hoffman and A. G. Sykes, *RSC Adv.*, 2016, **6**, 11295–11302.
11 P. N. Basa, A. Bhowmick, L. M. Horn and A. G. Sykes, *Org. Lett.*, 2012, **14**, 2698–2701.
12 M. Kadarkaraisamy and A. G. Sykes, *Inorg. Chem.*, 2006, **45**, 779–786.
13 B. Kampmann, Y. Lian, K. L. Klinkel, P. A. Vecchi, H. L. Quiring, C. C. Soh and A. G. Sykes, *J. Org. Chem.*, 2002, **67**, 3878–3883.
14 C. E. Lewis, *J. Org. Chem.*, 1970, **35**, 2938–2942.
15 Y. Furikado, T. Nagahata, T. Okamoto, T. Sugaya, S. Iwatsuki, M. Inamo, H. D. Takagi, A. Odani and K. Ishihara, *Chem. – Eur. J.*, 2014, **20**, 13194–13202.
16 S. Clair, M. Abel and L. Porte, *Chem. Commun.*, 2014, **50**, 9627–9635.
17 M. Stredansky, A. Sala, T. Fontanot, R. Costantini, C. Africh, G. Comelli, L. Floreano, A. Morgante and A. Cossaro, *Chem. Commun.*, 2018, **54**, 3971–3973.
18 D. Toffoli, M. Stredansky, Z. Feng, G. Balducci, S. Furlan, M. Stener, H. Ustunel, D. Cvetko, G. Kladnik, A. Morgante, A. Verdini, C. Dri, G. Comelli, G. Fronzoni and A. Cossaro, *Chem. Sci.*, 2017, **8**, 3789–3798.
19 M. Kikuchi, T. Yamagishi and M. Hida, *Dyes Pigm.*, 1981, **2**, 143–151.
20 J. O. Morley, *J. Phys. Chem.*, 1995, **99**, 5956–5960.
21 G. Pawin, K. L. Wong, K.-Y. Kwon and L. Bartels, *Science*, 2006, **313**, 961–962.
22 T. Kumagai, F. Hanke, S. Gawinkowski, J. Sharp, K. Kotsis, J. Waluk, M. Persson and L. Grill, *Phys. Rev. Lett.*, 2013, **111**, 246101.
23 T. Kumagai, F. Hanke, S. Gawinkowski, J. Sharp, K. Kotsis, J. Waluk, M. Persson and L. Grill, *Nat. Chem.*, 2014, **6**, 41–46.
24 M. Koch, M. Pagan, M. Persson, S. Gawinkowski, J. Waluk and T. Kumagai, *J. Am. Chem. Soc.*, 2017, **139**, 12681–12687.
25 L. Floreano, A. Cossaro, R. Gotter, A. Verdini, G. Bavdek, F. Evangelista, A. Ruocco, A. Morgante and D. Cvetko, *J. Phys. Chem. C*, 2008, **112**, 10794–10802.
26 R. Costantini, M. Stredansky, D. Cvetko, G. Kladnik, A. Verdini, P. Sigalotti, F. Cilento, F. Salvador, A. De Luisa, D. Benedetti, L. Floreano, A. Morgante, A. Cossaro and M. Dell'Angela, *J. Electron Spectrosc. Relat. Phenom.*, 2018, **229**, 7–12.
27 M. Klues, K. Hermann and G. Witte, *J. Chem. Phys.*, 2014, **140**, 14302.
28 A. Cossaro, M. Puppin, D. Cvetko, G. Kladnik, A. Verdini, M. Coreno, M. De Simone, L. Floreano and A. Morgante, *J. Phys. Chem. Lett.*, 2011, **2**, 3124–3129.

Keto-enol tautomerization drives the self-assembly of Leucoquinizarin on Au(111).

Roberto Costantini, Luciano Colazzo, Laura Batini, Matus Stredansky, Mohammed S. G. Mohammed, Simona Achilli, Luca Floreano, Guido Fratesi, Dimas G. de Oteyza and Albano Cossaro

Supporting Information

1. Materials and methods

Experimental.

The LQZ powder was purchased from TCI (purity >98%) and sublimated from a Knudsen cell at 370K. Deposition was performed at room temperature. The Au(111) sample was cleaned by cycles of Ar+ sputtering and annealing at 800 K.

O1s and C1s XPS were measured at 650 eV and 515 eV with overall resolution of 0.25 eV and 0.2 eV respectively. Binding energy calibration was done by measuring the $Au4f_{7/2}$ peak and by aligning its bulk component to 84.0 eV[1]. NEXAFS spectra were acquired in Auger yield measuring the Auger signal at 252 eV and 507 eV respectively.

STM measurements have been performed at the *Donostia* International Physics Center on a commercial Scienta-Omicron LT-SPM system at 4.3 K. A mechanically clipped PtIr wire was used as tip, sharpened by voltage pulses and indentation into clean Au(111) patches. The images were analysed with the WSxM software.[2]

Calculations.
We performed the theoretical analysis based on density functional theory (DFT) simulations within the Perdew-Burke-Ernzerhof (PBE) approximation to the exchange-correlation functional[3], using the Quantum-ESPRESSO simulation package[4] and the same norm-conserving pseudopotentials and numerical setup as in our previous studies[5]. Free molecules have been optimized in periodically repeated orthorhombic cells with a vacuum separation between replicas of 11 Å. We evaluate core-level spectra as described elsewhere.[6] Self-consistent calculations with a C pseudopotential generated with a 1s full core hole (FCH)[7] at a given atom site provide the XPS core level shifts between inequivalent carbon (oxygen) atoms: those range from C01 to C07 (O01 to O02) for tautomers T0, T1, and T2, and up to C14 (O04) for T3. Next, we evaluate NEXAFS within the half-core-hole approach (HCH) [8,9] by using the xspectra code[10]. Within a pseudopotential approach, the XPS and NEXAFS spectra of the molecule are both defined up to an energy constant, which is adjusted to match the experimental features.

2. STM imaging
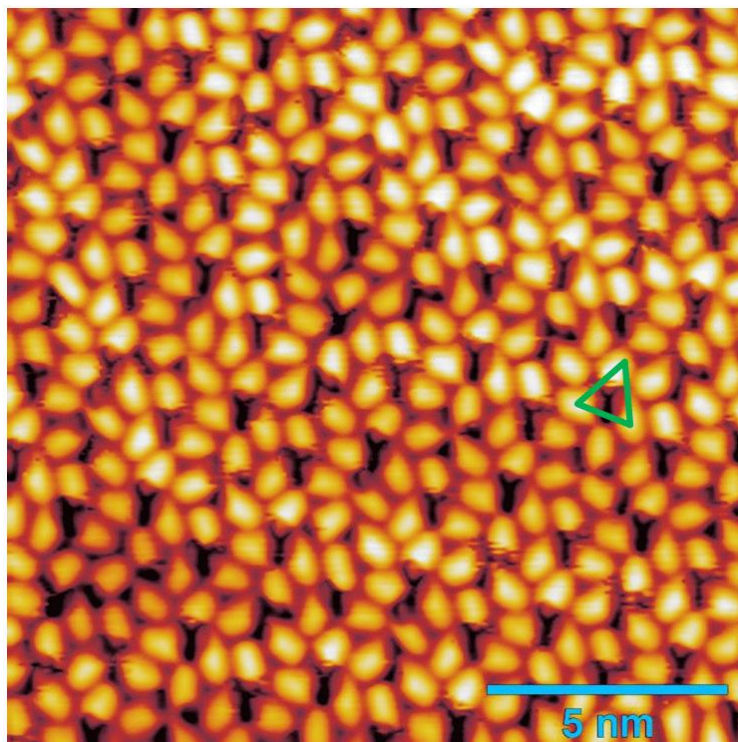
Figure S1 The RT growth of the monolayer of LQZ leads to less ordered phases with respect to the honeycomb structure presented in the manuscript. As shown in the image, a quite compact assembly of molecules is formed, where molecular trimers similar to the ones discussed in the manuscript, exhibiting a more distorted geometry, can be identified (green triangle). STM parameters: 0.03 V, 200 pA.
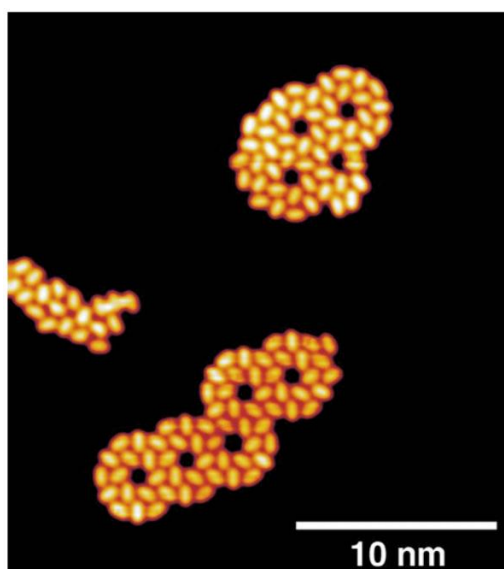
Figure S2. Thermal treatment of the LQZ films reveald that the desorption of the molecules is obtained at sample temperature Ts~500K. Image shows a very low coverage film, obtained upon annealing of a submonolayer at 420K, where both chiralities of the pores are present.(0.05 V, 100 pA).
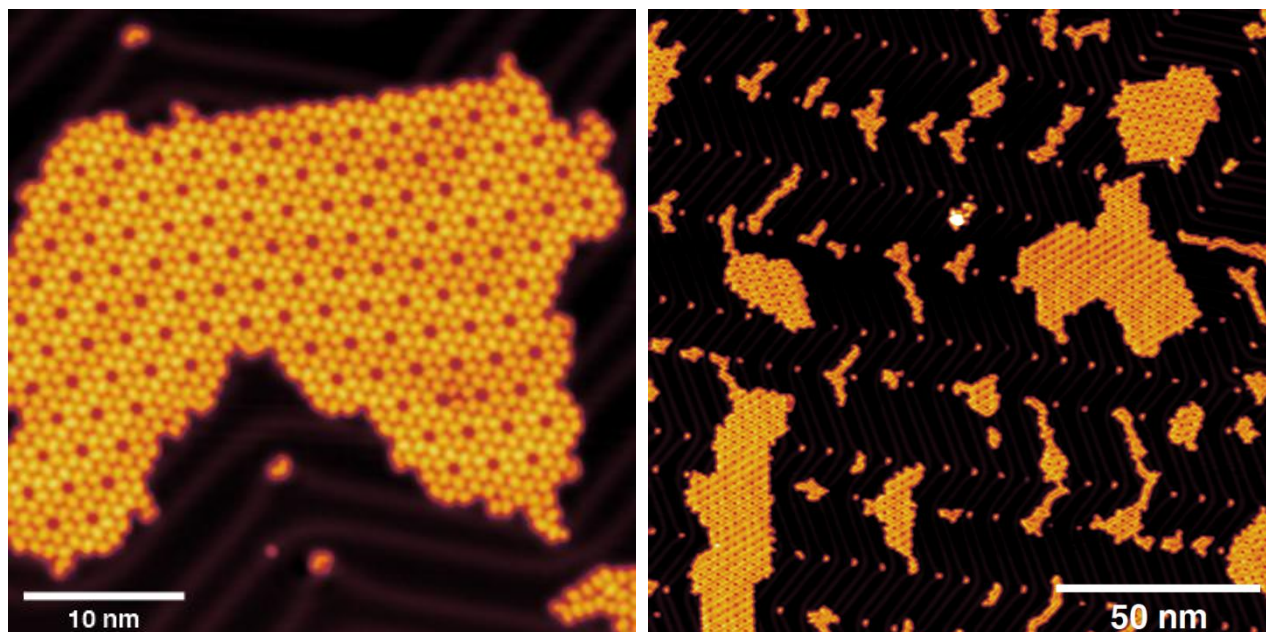
Fig S3 Two large scale images of the porous phase presented in the manuscript (0.4V, 20pA).

3. Telegraph noise
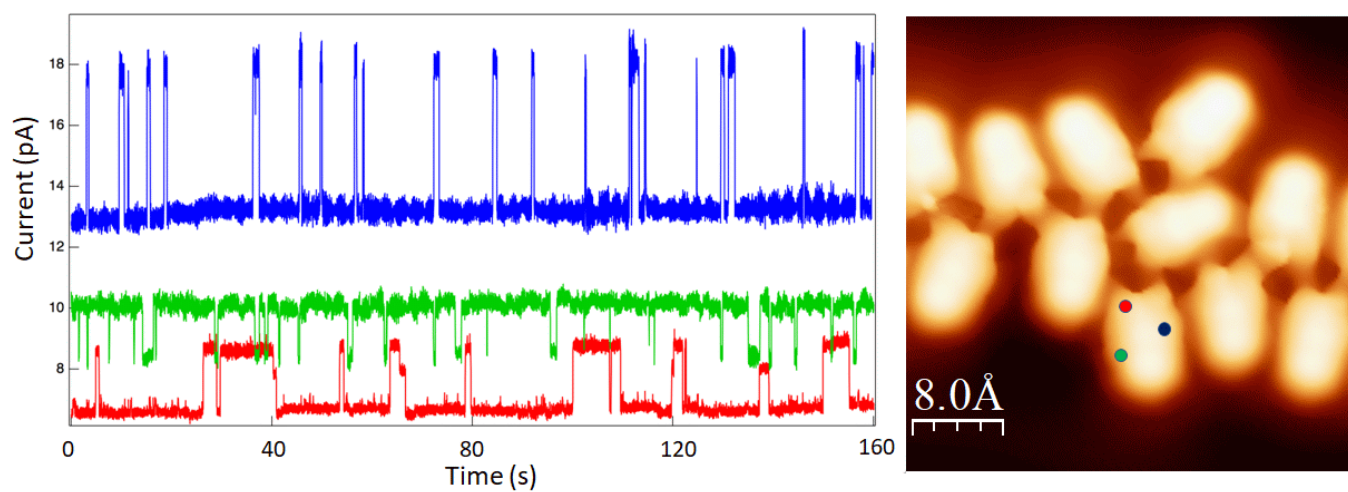
Figure S4. Three telegraph noise traces taken with different tip positioning, as indicated by the coloured circles in the image. The red curve is taken from the same measurement presented in figure 3 of the manuscript (last part of this curve corresponds to the first part of the curve of the manuscript, upon shifting the time scale). As stated in the manuscript, both rate and different level's dominance vary as the tip positioning is changed.

4. DFT calculations: properties of gas phase tautomers

The energies of tautomers in the gas phase and their frontier orbitals are reported in Table S1 and Figure S5, respectively. For simplicity, calculations of orbitals and spectra have been performed for molecules with reflection symmetry with respect to the *xy* plane. These tautomers are referred as TN$_{sym}$. In the case of T1, the tautomer with lowest energy, it has been verified that this symmetry assumption does not significantly affect the results when calculating the orbitals allowing for the distortion of the anthracene backbone. In the following paragraphs it will be shown that the spectroscopy simulations are also weakly affected by this relaxation. The T2$_{sym}$ geometry we considered, with the hydroxyl hydrogen atoms pointing away from the central ketones, represents the stable configuration for this tautomer. In fact, the configuration in which the hydrogen atoms point towards the ketones spontaneously switched to T1$_{sym}$ while performing structural optimization.

Tautomer	T0$_{(sym)}$	T1$_{sym}$	T1	T2$_{sym}$	T2	T3$_{sym}$	T3
Energy relative to T1 (eV)	0.90	0.09	0	2.10	1.98	1.09	1.01

Table S1. Calculated Total Energy for the four tautomers of Figure 1, in their symmetric form (TN$_{sym}$) and in their relaxed form (TN). T0 preserves the symmetry after relaxation so T0sym=T0.
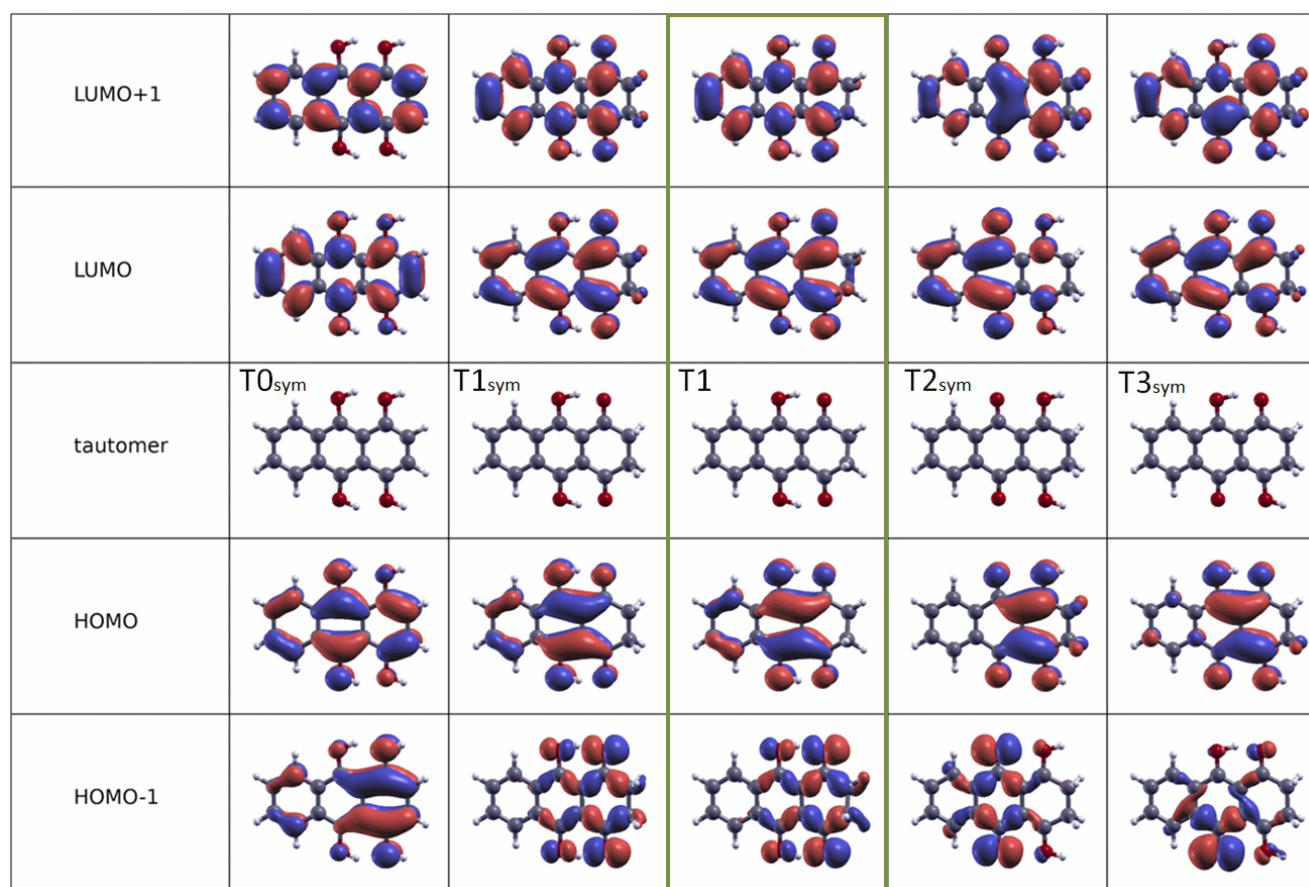

Figure S5. Graphical representation of the first occupied and unoccupied orbitals calculated for the different LQZ symmetric tautomers and for T1.
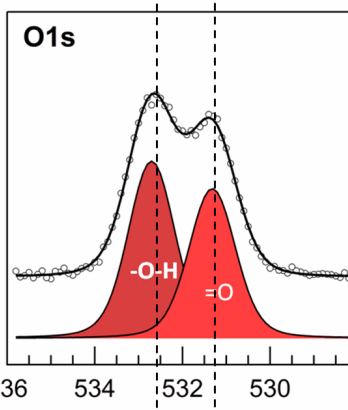
5. DFT calculations: Spectroscopic assignments
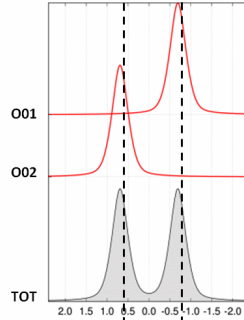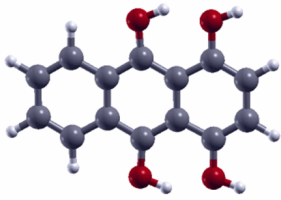
5.1: T1sym vs T1

The presence of the diketo-quinol upon tautomerization promotes the geometrical distortion of the corresponding termination of the anthracene backbone. Figure S6 reports the C1s NEXAFS and XPS calculations for the T1 tautomer in its distorted form (T1) and for the case in which the symmetry of the backbone was forced to be preserved (T1sym). Our findings show that the distortion introduced by the tautomerization poorly affects the spectroscopic properties. The same result was found for O1s (data not shown). For this reason, for simplicity, the calculations of the spectra for the comparison between the different tautomers have been done adopting the symmetric form of the molecules. We remark that the T1 spectra presented in the manuscript are calculated for the most stable distorted case.
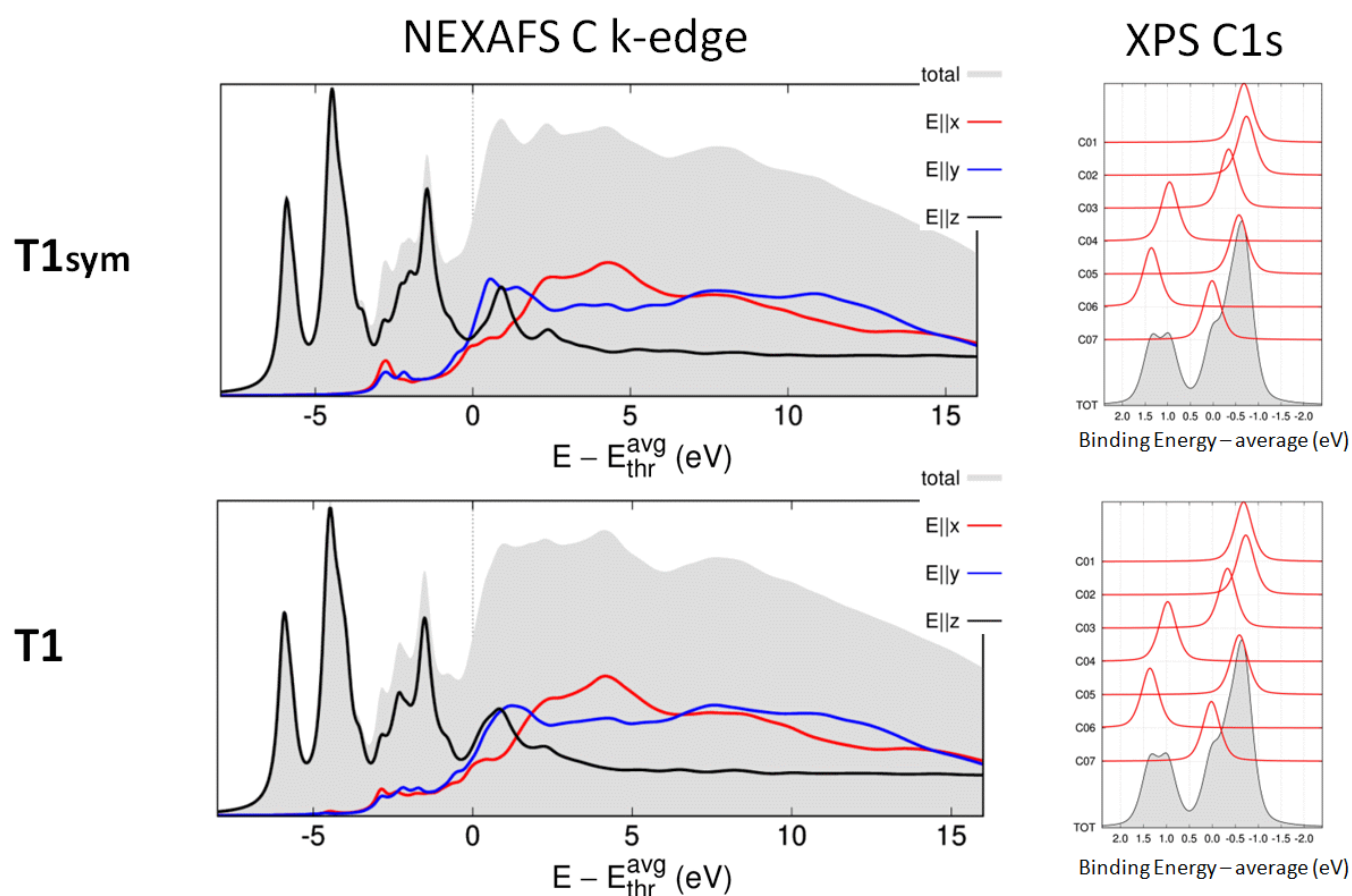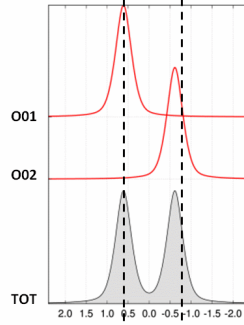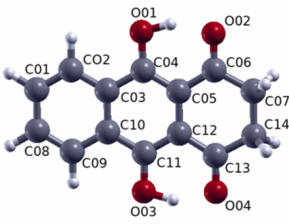

Figure S6. Calculated C1s spectra calculated for the distorted (top) and for the symmetric (bottom) T1 tautomer. The distortion only poorly affects the spectroscopic features on both NEXAFS and XPS spectra.
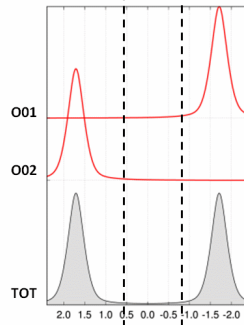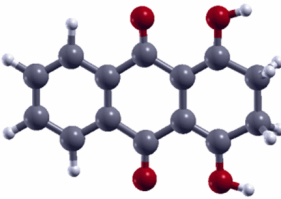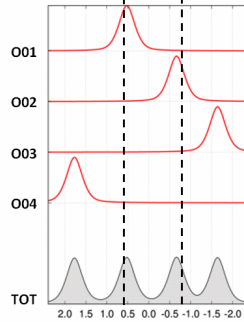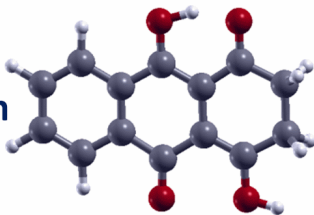
O1s

-O-H

=O

536   534   532   530

T0sym

O01
O02
TOT

2.0 1.5 1.0 0.5 0.0 -0.5 -1.0 -1.5 -2.0

T1sym

O01
C02   C04   O02
C01   C06
C03   C05   C07
C10   C12   C14
C08   C13
C09   C11
O03   O04

O01
O02
TOT

2.0 1.5 1.0 0.5 0.0 -0.5 -1.0 -1.5 -2.0

T2sym

O01
O02
TOT

2.0 1.5 1.0 0.5 0.0 -0.5 -1.0 -1.5 -2.0

T3sym

O01
O02
O03
O04
TOT

2.0 1.5 1.0 0.5 0.0 -0.5 -1.0 -1.5 -2.0
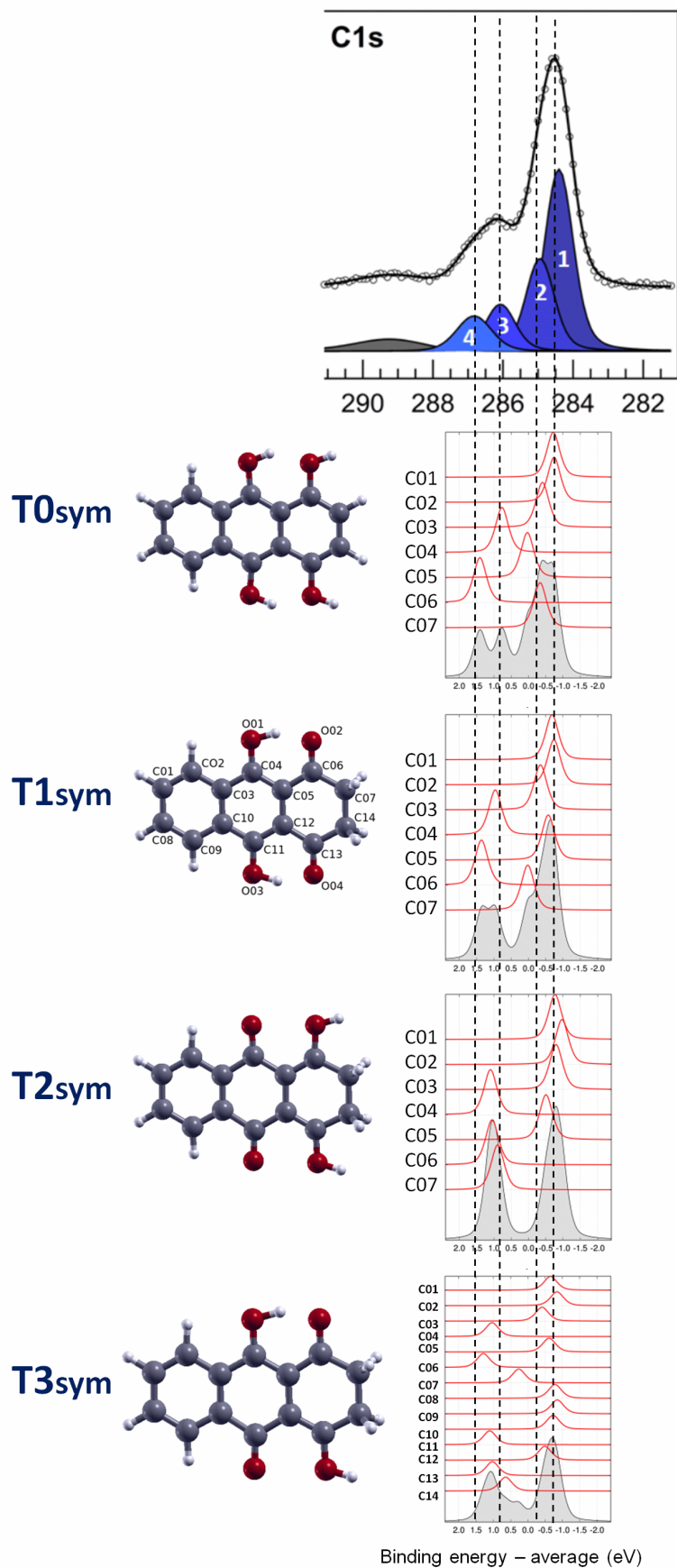
Binding energy − average (eV)

Figure S7. O1s XPS DFT simulated spectra for the T0sym, T1sym, T2sym and T3sym tautomers, compared with the experimental result. The simulated T1sym spectrum matches the experimental one, similarly to the T0sym spectrum. Conversely, features computed for T2sym and T3sym differ significantly from observations.

Figure S8. XPS DFT simulated spectra for the T0sym, T1sym, T2sym and T3sym tautomers, compared with the experimental result.
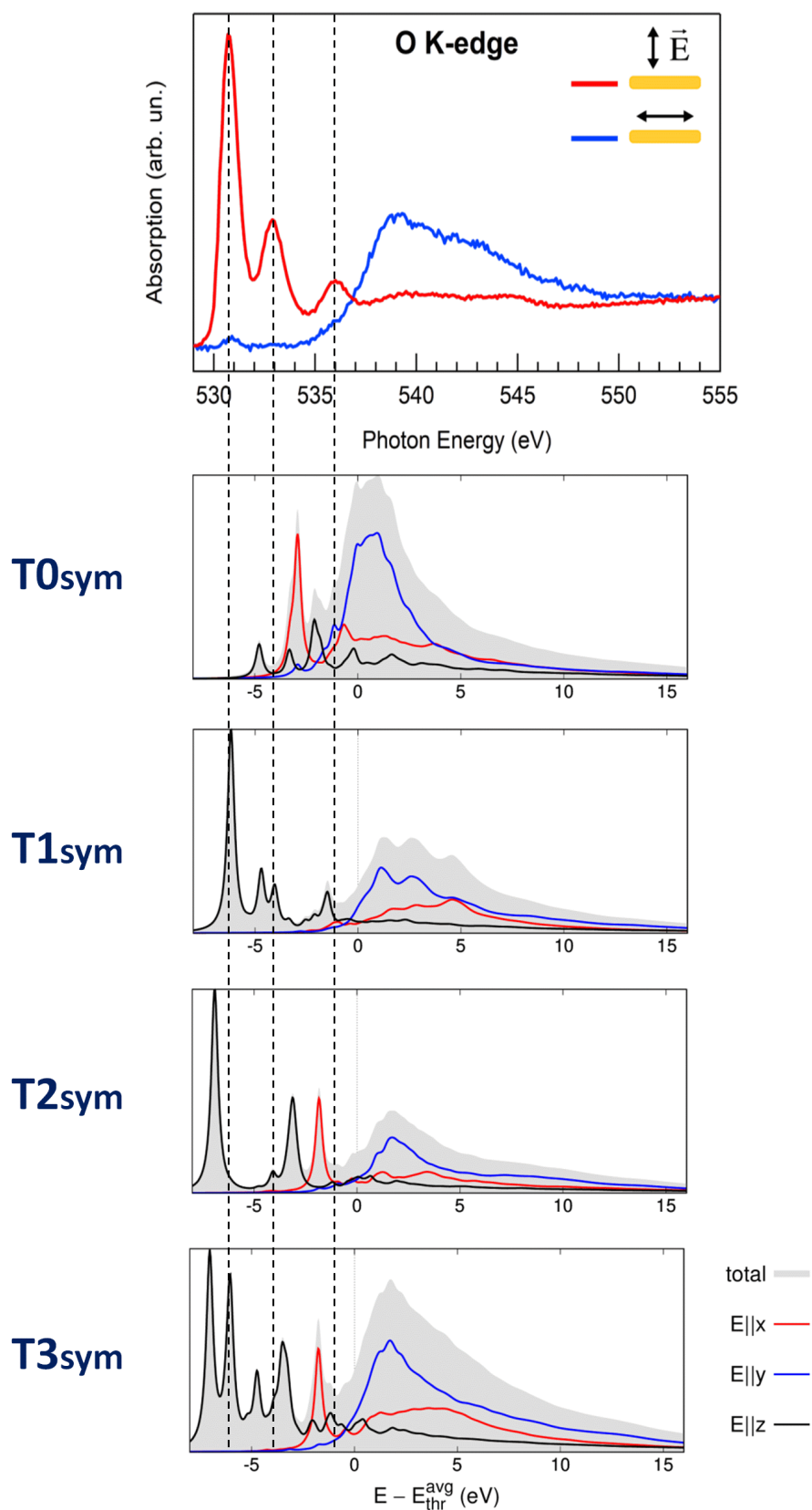
Figure S9 O1s NEXAFS spectra measured (top panel) and simulated for the T0sym T1sym, T2sym and T3sym tautomers (lower panels). T1sym spectrum better agrees with experimental curve with respect to other tautomers, as it properly reproduces the three π* transitions indicated by the dotted black lines, in terms of both relative intensities and position. Energy alignment between theoretical and experimental curve is made by optimizing the overall superposition of the spectral features.
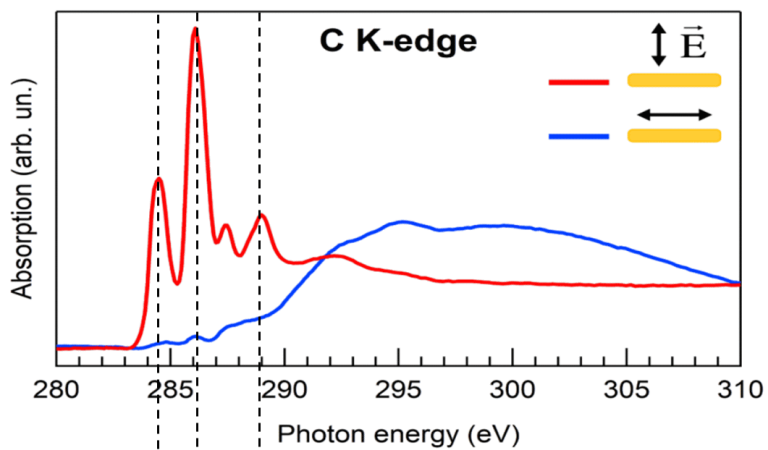
Figure S10 C1s NEXAFS spectra measured (top panel) and simulated for the T0sym, T1sym, T2sym and T3sym tautomers (lower panels). As for O1s, T1sym spectrum better agrees with experimental curve with respect to other tautomers. Energy alignment between theoretical and experimental curve is made by optimizing the overall superposition of the spectral features.
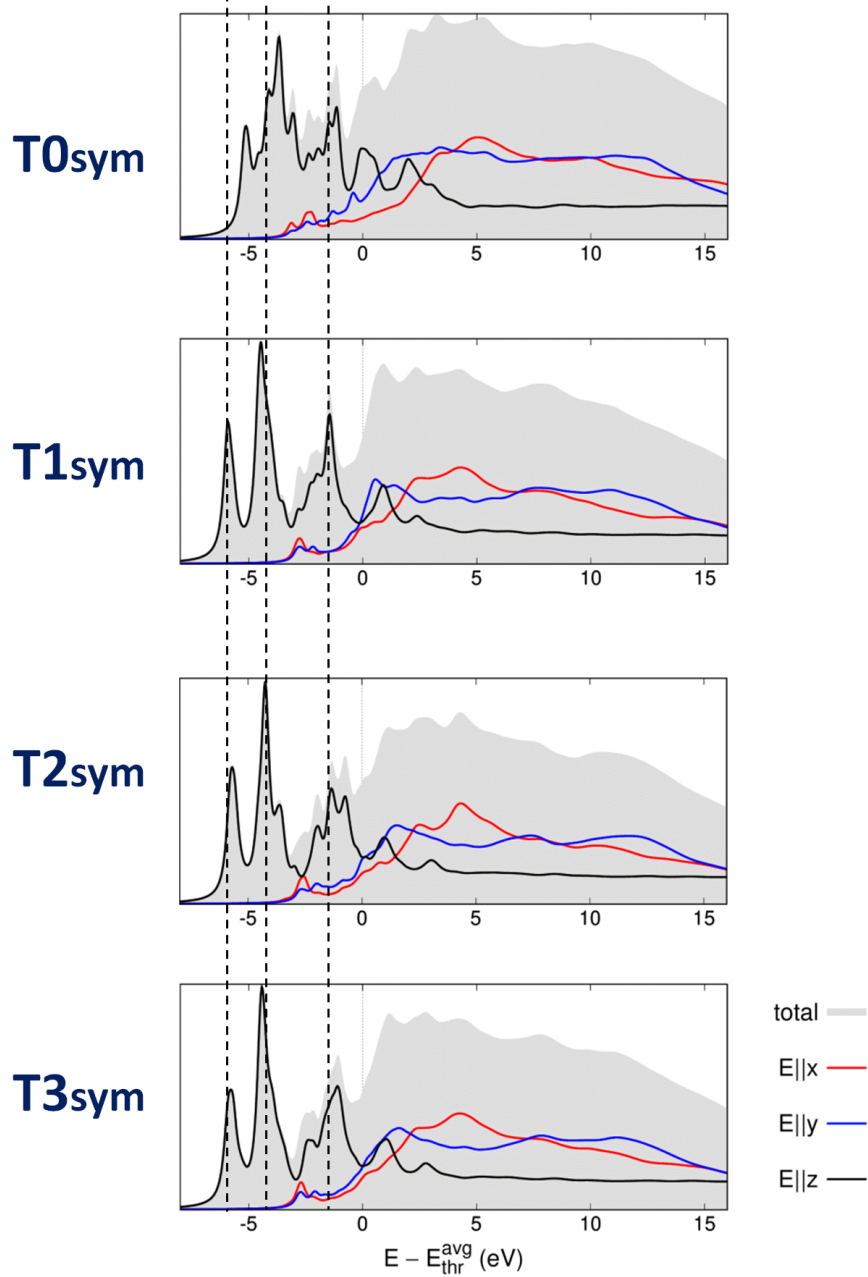
# T1 tautomer: C1s Nexafs assignments
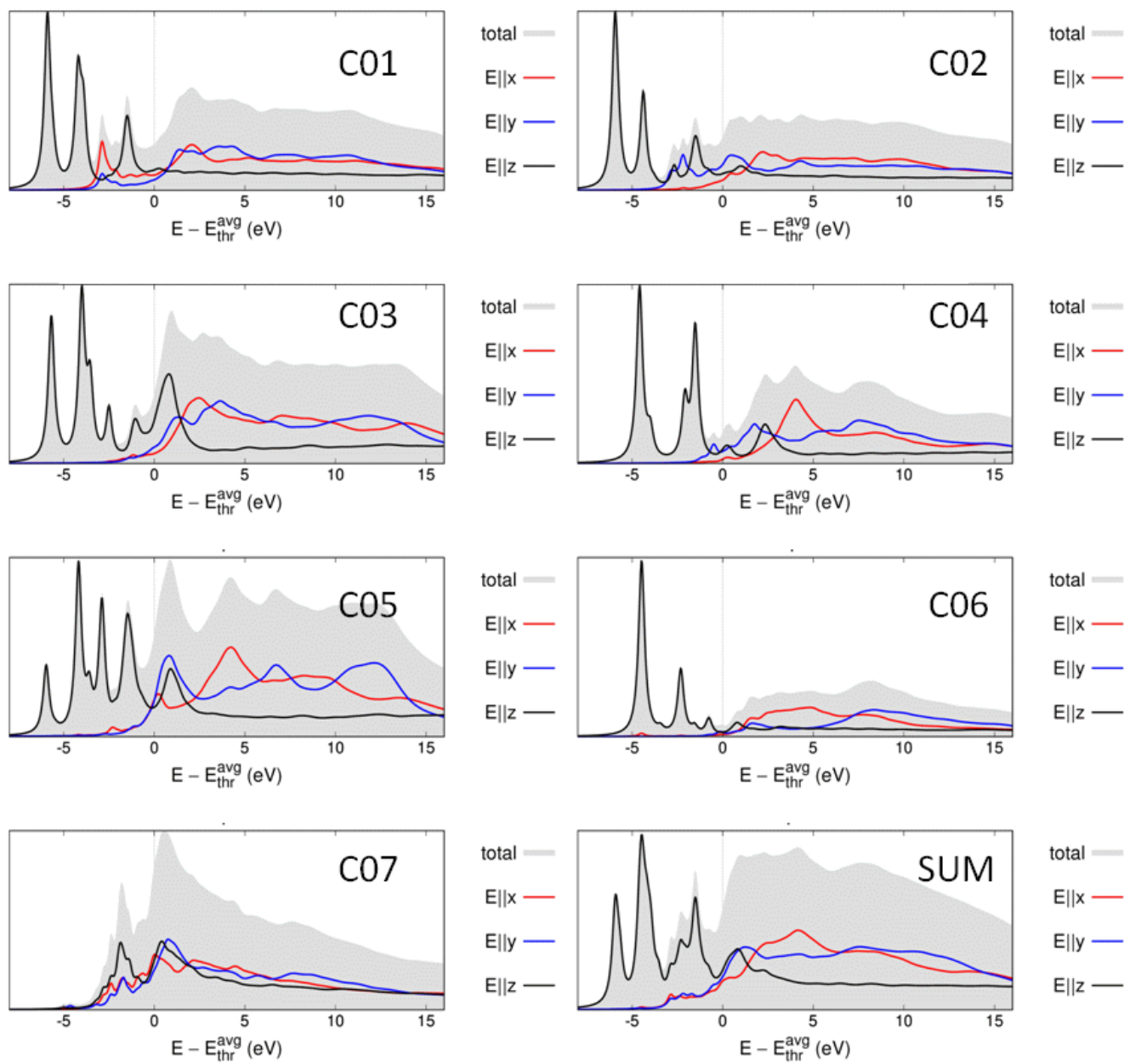
Figure S11 Calculated T1 NEXAFS C1s spectral contributions of the non equivalent carbon atoms. See model in figure S6 for carbon atoms numbering.

6. DFT calculations: Intermolecular hydrogen bonding

In order to describe the intermolecular bonding scheme of the LQZ porous assembly, we built a model for a 2D crystal with 6 molecules / unit cell as inspired by the experimental STM images. The initial molecular coordinates are obtained from the gas phase structure by subsequent 60° rotations around a common axis located at the nanopore, about 10 Å from the center of the molecule. Variable-cell structural relaxations are then used to optimize the in-plane lattice constant and coordinates.

The unit cell of the overlayer contains two trimers (two of them indicated by the green dotted triangles in Figure S11). As a result of the calculations, the lattice constant of the molecular overlayer is 25.8 Å and the bonding energy is of 0.25 eV / cell. Each molecule (all are equivalent by symmetry) has 5 nearest neighbours, pairing in three different ways: (i) with two molecules within the same trimer; (ii) with one molecule lying nearly parallel; (iii) with two molecules facing the same nanopore. The resulting interaction scheme is depicted for one trimer in Fig S11. The coloured arrows identify, by proximity, the molecular terminations involved in the different interactions. In order to estimate the different interaction energies, we calculate the total energy of molecular overlayers where only pairs or trimers of the molecules of the unit cell involved in the considered interaction are included.
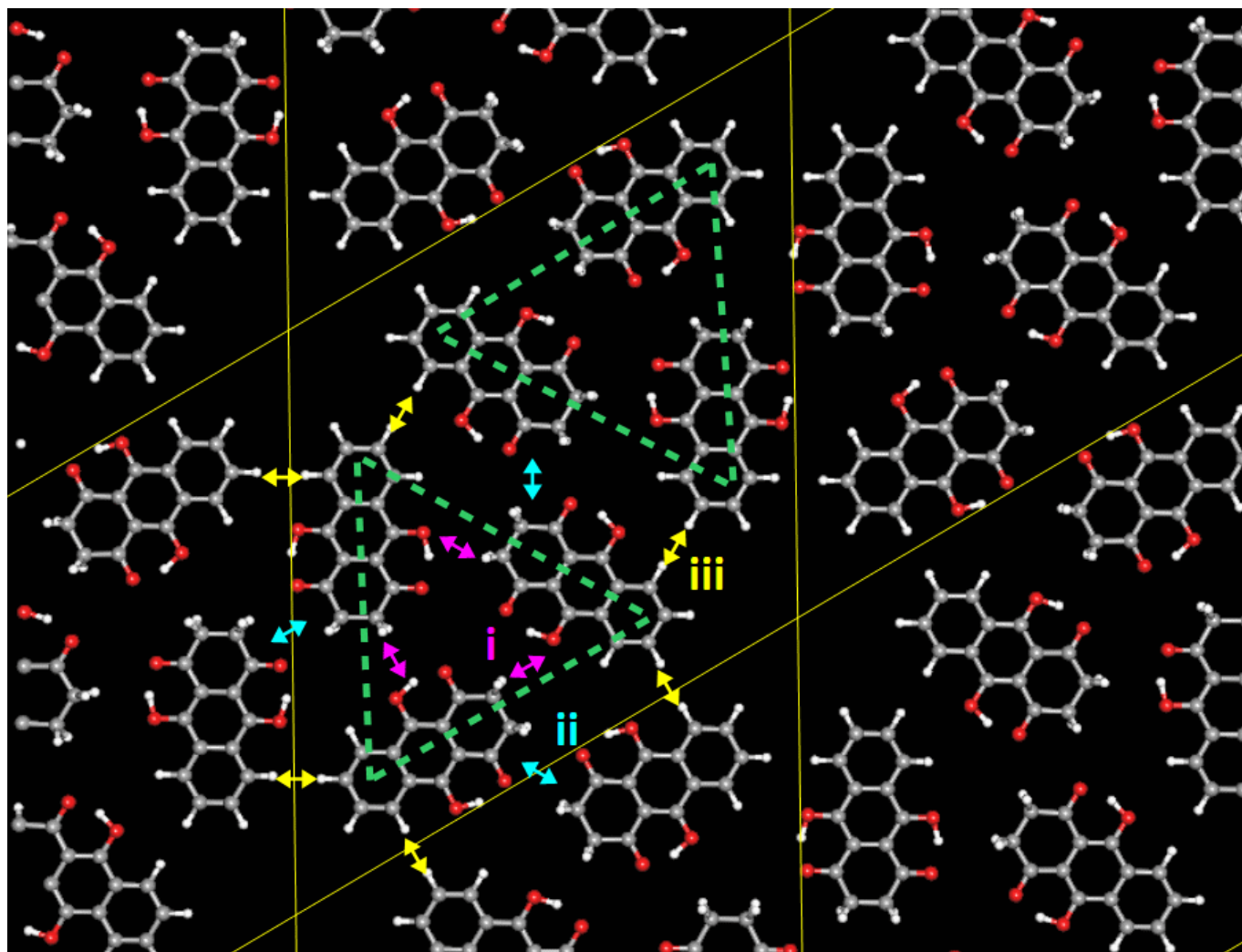
Figure S11. Structure of the molecular overlayer obtained by DFT calculations, by relaxation of the structure suggested by STM images

It results that : the trimers, which when isolated are stabilized by three (i) bonds, have a binding energy of 19 meV; the (iii) interaction in fact does not give any contribution to the bonding energy of the overlayer; the (ii)

bonding between the trimers is 71 meV.

Both (i) and (ii) interactions are related to the oxygen atoms which, as expected play a fundamental role in the assembly process. The strongest intermolecular bond, (ii), is due to the facing of the keto-enolic termination of adjacent molecules. We can speculate that the formation of the trimers is promoted by the symmetry of the substrate and that the high (ii) affinity stabilizes clusters of trimers. On the other hand the (ii) interaction may be seen as a competitor of the trimer formation and lead to less ordered assembly of the LQZ molecules. The low coverage assembly reported in Figure S2 shows both situations, with most of the molecules forming regular chains of pores and part of them assembled in a disordered stripe. We recall however that we are neglecting the interaction of the molecules with the substrate, which could in part modify the relative importance of the different intermolecular interactions.

Bibliography

1   A. Cossaro, L. Floreano, A. Verdini, L. Casalis and A. Morgante, *Phys. Rev. Lett.*, 2009, **103**, 119601.

2   I. Horcas, R. Fernández, J. M. Gómez-Rodríguez, J. Colchero, J. Gómez-Herrero and A. M. Baro, *Rev. Sci. Instrum.*, 2007, **78**.

3   J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, 77, 3865–3868.

4   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. de Gironcoli, S. Fabris, G. Fratesi, 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 and R. M. Wentzcovitch, *J. Phys. Condens. Matter*, 2009, **21**, 395502–395520.

5   G. Fratesi, V. Lanzilotto, S. Stranges, M. Alagia, G. P. Brivio and L. Floreano, *Phys. Chem. Chem. Phys.*, 2014, **16**, 14834–14844.

6   G. Fratesi, V. Lanzilotto, L. Floreano and G. P. Brivio, *J. Phys. Chem. C*, 2013, **117**, 6632–6638.

7   E. Pehlke and M. Scheffler, *Phys. Rev. Lett.*, 1993, **71**, 2338–2341.

8   M. Leetmaa, M. P. Ljungberg, A. Lyubartsev, A. Nilsson and L. G. M. Pettersson, *J. Electron Spectros. Relat. Phenomena*, 2010, **177**, 135–157.

9   L. Triguero, L. G. M. Pettersson and H. Ågren, *Phys. Rev. B*, 1998, **58**, 8097–8110.

10   C. Gougoussis, M. Calandra, A. P. Seitsonen and F. Mauri, *Phys. Rev. B - Condens. Matter Mater. Phys.*, 2009, **80**.