Supplementary Methodology

"Orientation and stability of a bi-functional aromatic organic molecular adsorbate on silicon"
K.M. ODonnell, H. Hedgeland, G. Moore, A. Suleman, M. Siegl, L. Thomsen, O. Warschkow, and S.R. Schofield

Computational Methodology for NEXAFS

Figure S2 shows the calculation flowchart for a molecular NEXAFS calculation as employed in the present work. For a molecule with N inequivalent absorbers of a particular species, $2N + 1$ SCF calculations are required:

- 1 ground state calculation to obtain the all-electron ground state energy.
- N calculations with a transition-potential (half core hole) pseudopotential placed at each absorber site in turn, to generate the density inputs for XSpectra and the HOMO/LUMO energy levels for each excited system.
- N calculations with a full core hole pseudopotential placed at each absorber site in turn, to obtain the excited all-electron energy.

Each of the N transition potential calculations is used as an input for 6 XSpectra calculations [one for each of the electric field directions (100), (010), (001), (110), (101) and (011)]. These directional spectra are combined

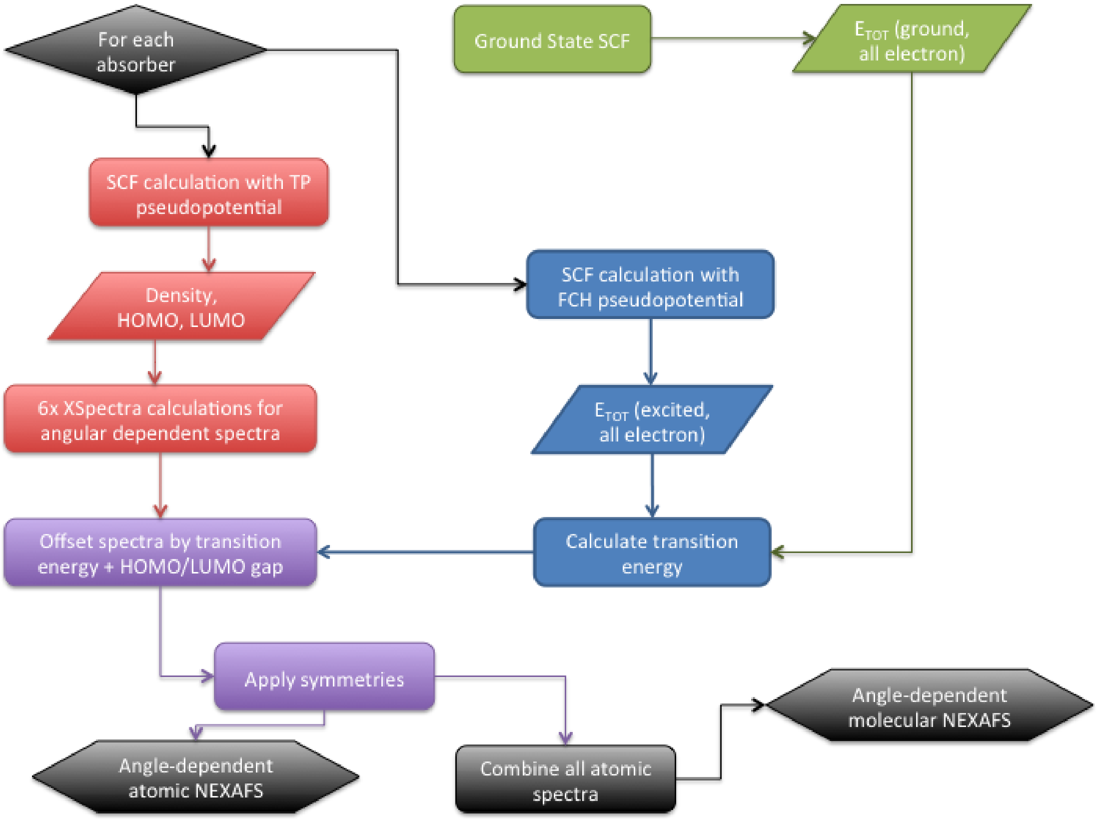

Figure S1: Flowchart for the NEXAFS calculation process used for the present work.

as described in the next section to give the appropriate angle-dependent NEXAFS spectra comparable to experimental measurements.

Symmetries are accounted for prior to combining the spectra from inequivalent sites. Under experimental conditions, symmetries represent the set of equivalent molecular/substrate orientations relative to the (fixed) X-ray polarization. The theoretical equivalent is the set of equivalent electric field orientations relative to a fixed molecular/substrate orientation. This set is of course system dependent. For the present work, the underlying Si(001)-(2x1) surface has dihedral symmetry due to the existence of two perpendicular dimer row directions.

Finally, to combine inequivalent sites into one spectrum, a transition energy is calculated as $E_{\text{excited}} - E_{\text{ground}}$ using the ground state and full-core-hole total energies. It is possible that the HOMO-LUMO gap is different for each excited absorber site so we also add the HOMO-LUMO gap to the resulting transition energy to represent an excitation into the first unoccupied state. In practice the variation of gap with absorber site is small (~ 0.1 eV) compared to the different in transition energies between inequivalent absorber sites (up to ~ 2 eV).

Calculation of angle-dependent spectra

The common practice for experimental NEXAFS on molecular adsorbates is to measure five angles of X-ray polarization relative to the surface. The reason for this is to use the angular variation of peak intensities to fit one of the Stohr equations and so determine the tilt angle of an unoccupied state with respect to the substrate normal. One could replicate this in XSpectra directly with five calculations, one for each electric field polarization. However, this makes symmetrization difficult. Instead, we construct a basis set of spectra consisting of calculations for the electric field directions (100), (010), (001), (110), (101), (011), and use linear combinations of these to generate the appropriate angle-dependent spectra. The reason six spectra are required and not just the three orthogonal directions is because the matrix element term of Fermi's golden rule is squared:

$$\sigma_v(\omega) \propto \sum_f |\langle f|\vec{v} \cdot \nabla|i\rangle|^2 \delta(E_f - E_i - \hbar\omega)$$

If the electric field vector is expressed in component form relative to some basis ($\vec{v} = \epsilon_x \vec{x} + \epsilon_y \vec{y} + \epsilon_z \vec{z}$), the matrix element $M_{\vec{v}}^2 = |\langle f|\vec{v} \cdot \nabla|i\rangle|^2$ expands as:

$$\begin{aligned} M_{\vec{v}^2} = \epsilon_x^2 M_x^2 + \epsilon_y^2 M_y^2 + \epsilon_z^2 M_z^2 + \epsilon_x \epsilon_y (M_x \overline{M_y} + M_y \overline{M_x}) \\ + \epsilon_x \epsilon_z (M_x \overline{M_z} + M_z \overline{M_x}) + \epsilon_y \epsilon_z (M_y \overline{M_z} + M_z \overline{M_y}) \end{aligned}$$

where $M_x = \langle f|\partial_x|i\rangle$, for example. It is then clear that judicious choice of the components $\epsilon_{x,y,z}$ in the calculation of a spectrum allows the spectral components arising from cross terms such as $M_x \overline{M_y} + M_y \overline{M_x}$ to be deduced from linear combinations of the basis spectra set listed earlier. For example,

$$\sigma_{xy} = \frac{1}{\epsilon_x \epsilon_y} \left(\sigma_{110} - \epsilon_x^2 \sigma_{100} - \epsilon_y^2 \sigma_{010} \right)$$

These cross-terms are typically small but not negligible and cannot be ignored since they too vary with polarization angle and hence affect the theoretical tilt angle determined via an application of Stohr's equations to the theoretical NEXAFS intensities.

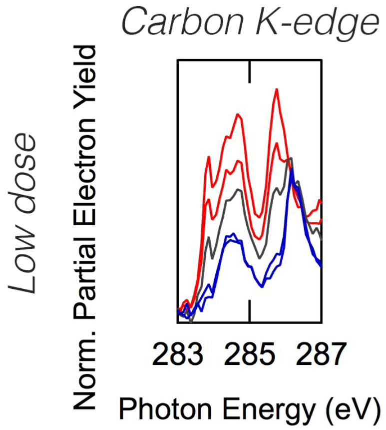

Figure S2: Enlargement of the inset from the main text, panel Fig. 2a