THE UNIVERSITY OF WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): F. Allegretti, M. Polcik and D.P. Woodruff Article Title: Quantitative determination of the local structure of thymine on Cu(1 1 0) using scanned-energy mode photoelectron diffraction

Year of publication: 2007 Link to published version: http://dx.doi.org/ 10.1016/j.susc.2007.07.006 Publisher statement: Allegrett1, F. et al, (2007). Quantitative determination of the local structure of thymine on Cu(1 1 0) using scanned-energy mode photoelectron diffraction. Surface Science, Vol. 601, pp. 3611-3622

Quantitative determination of the local structure of thymine on Cu(110) using scanned-energy mode photoelectron diffraction

F. Allegretti¹⁺, M. Polcik², D.P. Woodruff^{1*}

¹ Physics Department, University of Warwick, Coventry CV4 7AL, UK ² Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D 14195 Berlin, Germany

Abstract

The local adsorption structures of the surface species formed by interaction of thymine with a Cu(110) surface at room temperature, and after heating to ~530 K, have been investigated. Initial characterisation by soft-X-ray photoelectron spectroscopy and O K-edge near-edge X-ray absorption fine structure (NEXAFS) indicates the effect of sequential dehydrogenation of the NH species and provides information on the molecular orientation. O 1s and N 1s scanned-energy mode photoelectron diffraction shows the species at both temperatures bond to the surface through both carbonyl O atoms and the deprotonated N atom between them, each bonding atom adopting near-atop sites on the outermost Cu surface layer. The associated bondlengths are 1.96 ± 0.03 Å for Cu-N and 1.91 ± 0.03 Å and 2.03 ± 0.03 Å for the two inequivalent Cu-O bonds. The molecular plane lies almost exactly in the close-packed [110] azimuth, but with a tilt relative to the surface normal of approximately 20°. Heating to ~530 K, or deposition at this temperature, appears to lead to dehydrogenation of the second N atom in the ring, but no significant change in the adsorption geometry.

Keywords: photoelectron diffraction; chemisorption; surface structure;

⁺ now at Karl-Franzens-Universität Graz, Institut für Physik, Bereich für Experimentalphysik,

Universitätsplatz 5, 8010 – Graz, Austria

^{*} corresponding author. email d.p.woodruff@warwick.ac.uk

1. Introduction

The interaction of biologically-related molecules with surfaces is attracting increasing interest due to its potential relevance to areas such as biocompatibility, biosensors and the fabrication of novel biomaterials. Typical biological molecules such as DNA and proteins are far too complex to be investigated on surfaces at an atomic scale, although remarkably detailed scanning tunnelling microscopy (STM) images have been obtained in the last few years from DNA oligomers on Cu(111) [1, 2]. However, the interaction of such species with surfaces is governed by smaller components, such as simple amino acids and the nucleobases. Some quantitative surface structural methods are capable of providing relatively complete information on the adsorption of these component species which, while simple on the biological scale, are nevertheless rather complex in surface science. Using scanned-energy mode photoelectron diffraction (PhD) [3, 4] we have shown that it is possible to obtain quite detailed quantitative information on the structure of the simplest amino acids, glycine (NH₂CH₂COOH) [5, 6] and alanine (NH₂CH₃CHCOOH) [7], on Cu(110). In this paper we show that the same methodology can be used to determine the structure of the nucleobase molecule, thymine (fig. 1), on this same surface.

Previous investigations of the adsorption of thymine on Cu(110) have been spectroscopic only, and have studied mainly the species formed by room temperature adsorption (or heating to room temperature) [8, 9, 10] or by adsorption at a sample temperature of 558 K (or by heating the room-temperature adsorbed species to approximately this temperature) [9, 10, 11]. In addition, there is some information on low temperature monolayer [12] and multilayer [9] phases. RAIRS (reflection-absorption infrared spectroscopy) data [8, 10, 11] indicate that at both 300 K and at temperatures up to 558 K the adsorbed species has its molecular plane essentially perpendicular to the surface, while NEXAFS (near-edge X-ray absorption fine structure) spectra [8] confirm this result for the room temperature species and indicate that the molecular plane is aligned in the [10] azimuth parallel to the close-packed Cu atomic rows in the surface. Angle-resolved ultra-violet photoemission data also indicate a standing-up geometry of a monolayer

phase adsorbed at low temperature [12]. Temperature-programmed desorption (TPD) spectra show that if the molecule is adsorbed at room temperature and heated, H_2 is desorbed at a temperature around 463 K, indicating that a partial dehydrogenation occurs in this temperature range, while complete dissociation occurs around 638 K [9, 11]. Soft X-ray photoelectron spectroscopy (SXPS) of the species formed by room temperature adsorption [8] indicates that the two O atoms in the molecule are essentially equivalent, but that the two N atoms differ, probably due to deprotonation of one of these NH species. RAIRS data from the species formed at room temperature and at high (558 K) temperature, and also by annealing a surface dosed at room temperature to higher temperatures, clearly show significant differences, but the interpretation of these spectra in terms of models of the surface bonding is controversial. One of the studies of the room temperature species favours a model involving bonding to the substrate through the two O atoms and the (deprotonated) N(3) atom (see fig. 2(a)) [10], a structure favoured in vibrational spectroscopy investigations of thymine adsorption at an Au(111) electrochemical interface [13, 14]. The other study of the room temperature deposition proposes bonding of the intact molecule only through a single O atom [9]. This latter study, on the other hand, suggests that at the higher temperature deprotonation of one or both N-H species has occurred, leading to either O-N-O bonding to the substrate [9] as favoured for the room temperature species in the other RAIRS study [10], or O-N(3) bonding [11], or O-N(1) bonding (fig. 2(b)), or some mixture of these [9]. There is evidence from qualitative LEED observations for the formation of a complex long-range ordered phase (described in the matrix notation as $\begin{pmatrix} 6 & 0 \\ 1 & 2 \end{pmatrix}$) following the high temperature annealing treatment, but no evidence of similar ordering in the species formed at room temperature [9].

Clearly these previous studies provide consistent information regarding the orientation of the adsorbed species due to the interaction of thymine with Cu(110), but the detailed adsorption geometry, and indeed the identification of the adsorbed species at different temperatures, remains far from clear. These are the issues we seek to illuminate, using the PhD technique, combined with further characterisation using SXPS and NEXAFS. The PhD technique [3, 4] exploits the coherent interference of the directly-emitted component of the outgoing photoelectron wavefield from a core level of an adsorbate atom with components of the same wavefield which are elastically backscattered by the nearby substrate atoms. By measuring the photoemission intensity in specific directions as a function of photon energy, the resulting changes in photoelectron energy, and thus photoelectron wavelength, cause specific scattering paths to switch in and out of phase with the directly-emitted component, leading to modulations in the intensity which depend on the relative emitter-scatterer location. Simulations of these PhD modulation spectra, including multiple scattering for the surrounding atoms in 'guessed' model structures, allow one to determine the local adsorption geometry around the emitter atom. One special virtue of the method is that it is both element-specific and chemical-state specific, because it identifies emitters by their core level photoelectron binding energies. In the present case of thymine adsorption this means not only that we can probe separately the local structure around the O and N atoms within the molecule, but we can also distinguish the local structure of the two O and the two N atoms from one another through their 'chemical shifts' if they are rendered inequivalent by the molecule/surface interaction. Indeed, it is also possible to distinguish the emission from the inequivalent C atoms within the molecule, but as these atoms are unlikely to be involved in the surface bonding, they are of somewhat less interest.

2. Experimental details

The experiments were conducted in an ultra-high vacuum surface science end-station equipped with typical facilities for sample cleaning, heating and cooling. This instrument was installed on the UE56/2-PGM-2 beamline of BESSY II which comprises a 56 mm period undulator followed by a plane grating monochromator [15]. Different electron emission directions can be detected by rotating the sample about its surface normal (to change the azimuthal angle) and about a vertical axis (to change the polar angle). Sample characterisation *in situ* was achieved by LEED and by SXPS using the incident synchrotron radiation. Both the wide-scan SXPS spectra for surface characterisation, and the narrow-scan O 1s and N 1s spectra used in the PhD measurements, were obtained

using an Omicron EA-125HR 125 mm mean radius hemispherical electrostatic analyser. This was equipped with seven-channeltron parallel detection, and was mounted at a fixed angle of 60° to the incident X-radiation in the same horizontal plane as that of the polarisation vector of the radiation.

A clean well-ordered Cu(110) surface was prepared from an oriented and polished crystal slice by the usual combination of cycles of Ar ion bombardment and brief annealing to 800 K to give a sharp (1x1) LEED pattern and a SXPS spectrum devoid of impurities. Thymine powder of 99% of purity was obtained from Sigma-Aldrich and loaded into a purpose-built evaporation source. This comprised a glass tube contained within a ceramic cylinder which was heated by passing a current through a copper wire coiled around it. The ceramic tube was used to improve the homogeneity of the heating, and the evaporation rate was controlled by monitoring the temperature of a tantalum end-cap of the source with a thermocouple. The evaporation source was placed in differentially pumped six-way cross separated from the main chamber by a gate valve, a Cu gasket with a 7 mm diameter hole being placed between the thymine source and the gate valve in order to collimate the thymine flux and limit the pressure rise in the experimental chamber during dosing.

O K-edge NEXAFS spectra were recorded in the Auger electron detection mode by measuring the intensity of the electron emission at the energy corresponding to the O KVV Auger transition (513 eV), and scanning the photon energy through the O K-edge. Spectra were recorded at a range of angles of incidence from normal to grazing in each of the two principle azimuths ([001] and $[1\overline{10}]$); these changes in incidence geometry provide information on the dependence of the intensity of the molecular resonance peaks on the direction of the polarisation vector of the linearly-polarised incident radiation. These data provide the basis for the determination of the molecular orientation.

The PhD modulation spectra were obtained by recording a sequence of photoelectron energy distribution curves (EDCs) around the O 1s and N 1s peaks at typically 3 eV or 4 eV steps in photon energy in the photoelectron kinetic energy range of approximately 50-

350 eV for each of a number of different emission directions in the polar emission angle range from 0° (normal emission) to 60° in the [001] and $[1\overline{10}]$ azimuthal planes, mainly in 20° intervals. These data were processed following our general PhD methodology (e.g. [3, 4]) in which the individual EDCs are fitted by the sum of Gaussian peaks, a step and a template background. The integrated areas of each of the individual chemically-shifted component peaks were then plotted as a function of photoelectron energy and each final PhD modulation spectrum was obtained by subtraction of, and normalisation by, a smooth spline function representing the non-diffractive intensity and instrumental factors. These PhD modulation spectra were used in the structure analysis described in the following section.

3. Results

3.1 Spectral characterisation: SXPS and NEXAFS

SXPS data recorded around the O 1s, N 1s and C 1s core level photoemission peaks are shown in fig. 3 for three different preparation conditions, namely immediately following a saturation dose of thymine at room temperature, after the same room temperaturedeposited surface had been heated to approximately 520 K, and following deposition of the same dose at the higher temperature of 530 K. Previous work has shown that multilayers of thymine deposited at low temperature start to desorb to leave a monolayer at ~300 K [9], and a coverage calibration of a saturated layer deposited at room temperature using Auger electron spectroscopy appears to indicate a coverage of ~ 0.3 ML [10]. For the initial room temperature deposited layer Fig. 3 agrees well with the previously-reported O 1s and N 1s SXPS [8] although there are differences in absolute binding energies of ~0.5 eV. This is probably attributable to differences in calibration methods; in our case the absolute energy calibrations was based on the Cu 3s peaks taken to have a photoelectron binding energy of 122.45 eV [16, 17]. For our present purposes, however, the absolute binding energies are not important, and our concern is to spectral fingerprints associated with single or multiple chemically-shifted component peaks. In particular, the N 1s spectrum is characterised by two components with a photoelectron binding energy difference of ~1.7 eV, while the O 1s shows a single peak with no indication of multiple components. The clear implication is that the local bonding environments of the two O atoms are similar, while the two N atoms are clearly in quite different environments. The most obvious interpretation is that the N(3) atom, but not the N(1) atom, is dehydrogenated and that the bonding to the Cu surface is through the two O atoms and the N(3) atom (fig. 2(a)), as suggested in the electrochemical deposition work on Au(111). Based on their similar SXPS data Furukawa et al. [8], proposed that the lower binding energy component of the N 1s emission originates from the deprotonated (surface bonding) N(3) atom and that the peak with the higher binding energy is from the N(1) that is still hydrogenated. As we shall see, our analysis of the PhD modulations of these photoemission components provides clear confirmation of this identification and the associated bonding geometry. The C 1s spectrum from the as-deposited thymine layer at room temperature, shown in fig. 3, clearly contains four distinct components, one of which is approximately twice the intensity of the other three. The most obvious interpretation is that the largest peak corresponds to emission from the two carbonyl C(2)and C(4) atoms, while the other three components are from the methyl C and the C(5)and C(6) atoms in the thymine ring, each being locally inequivalent. On the other hand, this largest peak has the lowest binding energy, whereas carbonyl species typically have higher C 1s binding energies. However, as we make no use of the C 1s SXPS spectra, we do not speculate further on the individual assignments.

As seen in fig. 3, heating this surface to \sim 520 K, or dosing at this higher temperature, leads to significant changes in all the spectra, although the two different forms of high temperature treatment seem to lead to very similar changes, with the possible exception of the rather complex C 1s spectra. Heating the surface prepared at room temperature was also found to lead to an ordered but somewhat streaked LEED pattern that appears to correspond to that reported previously, although the high temperature dosing seemed to give rise to a more complex pattern that was not studied in detail. The XPS data certainly indicate that the majority species on the surface is probably the same in both forms of high temperature preparation, although there may be differences in ordering or the relative occupation of minority species. Both the higher temperature annealing and deposition lead to shifts in the C 1s and O 1s spectra, but in each case the number of

constituent components remains unchanged; in particular, the O 1s spectra show a single component peak both before and after heating. The N 1s spectra, however, show a more significant change, the higher binding energy peak (assigned above to the N(1) NH species in the room temperature preparation) falling sharply in intensity (and almost vanishing at ~10 K higher temperature) while the lower binding energy peak grows and shifts slightly. As this heating is known to lead to H₂ desorption, it seems reasonable to attribute this change in the N 1s spectrum to the dehydrogenation of the N(1) atom. As is discussed below, the PhD data from the O 1s emission, and from the combined N 1s emission from the two components, is changed little as a result of this heating, so it seems that the adsorption geometry is almost identical before (fig. 2(a)) and after (fig. 2(c)) heating. The changes in the C 1s and O 1s photoelectron binding energies must therefore be a consequence of electron rearrangement in the thymine ring following the N(1)dehydrogenation, rather than any substantial change in the adsorption geometry. Clearly the resultant changes in hybridisation will have some consequences for the SXPS spectra, but there is no simple and reliable way to estimate the magnitude (or sign) of these effects.

Fig. 4 shows the O K-edge NEXAFS spectra recorded at grazing incidence, and at normal incidence in each of the two principle azimuths, from thymine adsorbed on Cu(110) at room temperature, and in a single azimuthal plane from a deposition at ~530 K. The photon energy calibration used in preparing these spectra is based on a very weak O K-edge feature in the incident flux attributable to O contamination in the optics. This feature was assigned an energy of 531 eV; while this calibration method is far from exact, the resulting room temperature preparation data are in excellent agreement with those of Furukawa *et al.*[8]. Notice that the incidence is varied by rotating the sample about a vertical axis, the azimuth of the incidence plane is also the azimuth of the polarisation vector of the incident radiation. Because it is the direction of the polarisation vector, E, that is of significance in determining selection rules in NEXAFS, it is common to quote the incidence angle as the grazing angle (i.e. normal incidence is 90° grazing), as this angle then also corresponds to the angle between the polarisation vector and the surface normal.

The NEXAFS spectra for the two different preparation conditions are extremely similar. In particular, both show a sharp doublet structure close to the photoabsorption threshold, which has low intensity at grazing incidence, and is most intense for normal incidence in the [001] azimuth. These peaks are assigned, as in the earlier study of Furukawa et al.[8], to transitions from the O 1s state to $\pi^*_{C=O}$ antibonding states, for which the cross-section is highest when the polarisation vector lies perpendicular to the molecular plane. The clear implication is therefore that the molecular plane lies approximately perpendicular to the surface aligned in the [110] azimuth. This conclusion was also reached for the asdeposited species at room temperature by Furukawa el al., but fig. 4 indicates this same orientation is adopted by the species on the surface following deposition at a temperature above that at which TPD implies that (further) dehydrogenation occurs. The fact that the $\pi^*_{C=O}$ resonance is a doublet is attributed to the inequivalence of the two C=O carbonyls within the adsorbed molecule, despite the fact that the more local probe of SXPS indicates that the two O atoms are almost equivalent. Essentially the same doublet structure is seen in NEXAFS from evaporated thin films of (intact) thymine molecules, and in this case comparison with similar spectra from cytosine led to the conclusion that the higher energy component is associated with the C(2) atom between the two N atoms [18]. Comparison of the NEXAFS spectra from the adsorbed thymine species, deposited at room temperature and at 530 K, shows a slight change in the relative intensities of the two components of the doublet, presumably a (rather modest) consequence of the rehybridisation following the N(1) dehydrogenation discussed above.

A more quantitative estimate of the exact molecular orientations, based on the full set of NEXAFS spectra recorded at incidence angles in 20° steps between normal and 10° grazing (in both azimuths for the as-deposited thymine, but only in [001] for the surface after heating), can be obtained using the standard analysis approach as described, for example, in another recent application of molecular adsorption on Cu(110) [19]. This exploits the fact that the dependence of the NEXAFS intensity on the angle, δ , between the electric vector of the incident radiation and the final state π -orbital (i.e. the direction perpendicular to the molecular plane) is given by $\cos^2\delta$ [20]. The results of this analysis may be expressed in terms of the angle of tilt of the molecule away from the surface

normal, and the angle of twist out of the [110] azimuth. The best fits to the polarisationdependent π -resonance intensities are also shown in fig. 4, the associated orientational parameter values agreeing within the error estimates for the two different preparation temperatures, with a tilt of $30\pm10^{\circ}$ and a twist of $20\pm15^{\circ}$.

3.2 PhD data: introduction and preliminary evaluation

In order to determine the adsorption structure from PhD data it is necessary to compare the experimental spectra with the results of simulations based on realistic multiple scattering calculations for a series of model trial structures, adjusting these structures until a good fit is achieved. However, some features of the probable structural models may often be obtained from inspection of the raw experimental data. In particular, if measurements are made in an emission direction that corresponds to that of 180° backscattering from a nearest-neighbour substrate atom, the PhD spectrum is commonly dominated by a single periodicity (in electron momentum) with a relatively large period, corresponding to the (short) scattering pathlength difference from this one neighbour. We also note that PhD spectra are generally very sensitive to the local emitter sites, so similar spectra generally imply closely similar emitter sites.

Fig. 5 shows comparisons of a subset of the experimental PhD spectra which we consider in the light of these general remarks. Fig. 5(a) shows a comparison of the O 1s PhD modulation spectra from the two different surface preparations for polar emission angles of 0° (normal emission), 20° and 40° . Two features of these spectra are notable. Firstly the modulations seen for the two different preparations are essentially identical, most of the differences being attributable to noise and scatter; the implication, remarked upon earlier, is that the local adsorption geometry of the O atoms is the same, and is not affected by the apparent deprotonation that occurs on heating from room temperature to ~530 K. A second notable feature, common to both sets of spectra, is that the normal emission spectrum shows reasonably strong modulations dominated by a single long period, while the modulations in the off-normal emission spectra are significantly weaker. This strongly suggests that normal emission corresponds approximately to a 180° backscattering geometry from a near-neighbour strongly-scattering Cu atom, implying that the O atoms are close to atop surface Cu atoms. Panels (c) and (d) of fig. 5 show the matching set of spectra, in the same emission geometries, from the two N 1s components of the species resulting from depositing thymine at room temperature; (c) is from the lower binding energy component, (d) from the higher binding energy component, these being labelled N(3) and N(1) respectively according to the discussion in section 3.1. The form of these two sets of PhD spectra confirm this labelling. The spectra in (c) are extremely similar to the O 1s PhD spectra of panel (a), with strong long-period modulations at normal emission that are damped out at high emission angles, consistent with an atop emitter at a very similar distance from the Cu scatterer to the O emitters of panel (a). The implication is that the two O atoms and the N(3) atom between are all close to atop surface Cu atoms, as shown schematically in fig. 2(a). By contrast, panel (d) of fig. 5 shows weak modulations, a result expected from the N(1) atom in this adsorption geometry, because there are no strongly-scattering Cu atoms nearby and the PhD is likely to be dominated by intramolecular scattering involving only weakly-scattering nearneighbour C, N and O atoms. Finally, panel (b) of fig. 5 shows a comparison of a similar set of PhD spectra, from the single N 1s peak from the adsorbate species resulting from thymine deposition at a sample temperature of ~530 K, with PhD spectra obtained from the room temperature species by adding the N(3) and N(1) components. Evidently these PhD spectra from the two preparations are extremely similar. The clear implication is that although significant spectral changes have occurred in preparing the higher-temperature species, the sites of the two distinct N atom within the molecule are little changed. This reinforces the view that this higher temperature preparation may produce a species in which the N(1) atom is dehydrogenated, but the overall molecular adsorption geometry is essentially unchanged.

This qualitative evaluation provides strong support for the interpretation of the spectroscopic data of section 3.1, and for the adsorption geometries of figs. 2(a) and 2(c) for the species present after room temperature deposition, and after subsequent heating, respectively. However, a reliable interpretation of the PhD data in terms of a quantitative structural model requires significantly more detailed evaluation of the data.

3.3 PhD data: quantitative modelling and structure determination

Full (quantitative) structure determination in PhD is based on the use of multiple scattering simulations for trial model structures to reproduce the experimental modulation spectra. These calculations were performed with computer codes developed by Fritzsche [21, 22, 23] that are based on the expansion of the final state wave-function into a sum over all scattering pathways that the electron can take from the emitter atom to the detector outside the sample. The quality of agreement between the theoretical and experimental modulation amplitudes is quantified by the use of an objective reliability factor (*R*-factor) defined [3, 4] such that a value of 0 corresponds to perfect agreement and a value of 1 to uncorrelated data. Different model structures were initially tested on a grid-search of structural parameters, but this structural optimisation to locate the minimum *R*-factor was also aided with the use of an adapted Newton-Gauss algorithm to automate the search in a region of multi-parameter space. In order to estimate the errors associated with the individual structural parameters we define a variance in the minimum of the *R*-factor, R_{min} , following an approach developed by Pendry for LEED [24]. All parameter values giving structures with R-factors less than $R_{min} + Var(R_{min})$ are regarded as falling within one standard deviation of the 'best fit' structure [25].

In view of the qualitative discussion of section 3.2, indicating the strong similarity of the local structure of the adsorbed species prepared at the two different temperatures we focus on the data from the room temperature preparation for which the chemical shift in the N 1s spectra between the N(3) and N(1) atoms provides more incisive information. In view of the complexity of the system the structural analysis proceeded through the sequential (rather than combined) analysis of individual PhD data sets for each distinct emitter atom, in effect determining the local position of each of these emitter atom sites independently. As we show below, these provide a highly consistent picture of the local adsorption geometry.

The first step in the quantitative analysis was to provide a more quantitative test of the

adsorption site of the molecule, starting with the location of the N(3) atom that lies between the two C=O carbonyls, that are also thought to be involved in the substrate bonding. In the presence of strong substrate backscattering, intramolecular scattering is usually relatively unimportant, so initially calculations were performed simply for an isolated N atom adsorbed in each of the four relatively high symmetry adsorption sites, namely atop, hollow, long bridge and short bridge. For each of these sites the height of the N emitter above the outermost surface layer was adjusted to obtain the lowest Rfactor, and some optimisation of the vibrational amplitudes was undertaken for each of the best-fit structures, but no further structural optimisation (e.g. including the effects of substrate relaxation or intramolecular scattering) was undertaken at this stage. The results of these calculations were lowest R-factor values for each site of 0.32 (atop), 0.44 (hollow), 0.60 (long bridge) and 0.73 (short bridge). The last two of these can clearly be rejected. For the atop site the optimised N-Cu distance was found to be 1.96 Å, while for the hollow site the N was found to be 1.96 Å above a second layer Cu atom. In effect, therefore, both of these solutions correspond to the N atom occupying an atop site (at the same distance), the difference being whether the N lies atop outermost or second layer Cu atoms. From the structural point of view, of course, this distinction is important, but as the dominant backscattering atoms are in similar positions, the PhD data to be expected from the two sites is similar. In fact further optimisation of the atop and hollow sites, including the effect of correlated vibrations (allowing the near-neighbour scatters to have smaller vibrational amplitudes relative to the emitter) led to a significant improvement in the *R*-factor for the atop site to 0.25, but an even larger reduction for the hollow site to 0.26, attributable to the reduced influence of the outermost near-neighbour scatterers in the outermost layer for the hollow site, relative to that of the nearest-neighbour Cu atom below. We conclude, therefore, that the N atoms clearly does lie atop a Cu surface atom, but cannot, on this basis alone, distinguish the atop and hollow sites, as the R-factor values are too similar.

A second stage of refinement of the optimum structure to account for the N(3) N 1s PhD spectra involved the inclusion of all the intramolecular scattering. To achieve this it was assumed that the relative positions of the C, N and O atoms in the adsorbate were the

same as in the free thymine molecule [26]. In these calculations the orientation of the molecule was defined by a set of three Euler angles, (Φ, Θ, Ψ) , as used in earlier PhD studies of planar molecular adsorbates [27, 28]. The Euler angles relate the Cartesian coordinate systems fixed relative to the surface $(x_L y_L z_l)$, and relative to the molecule (x_M, y_M, z_M) . Two somewhat different conventions are used to define these angle; in the 'main convention' the three rotation angles are performed in sequence about z_I , x_M and z_M , whereas in the 'x-convention' the axes of these sequential rotations are z_I , x_I and z_I . These two conventions can be shown to be equivalent if the angles Φ and Ψ are switched around. Here we quote the angles according to the former convention. The case of $\Phi=\Theta=\Psi=0^{\circ}$ describes the situation in which the molecule is fully upright with its molecular plane in the [110] azimuth, the nominal ideal geometry initially discussed in the Introduction, for thymine on Cu(110). Φ defines an azimuthal rotation of the molecular coordinate system relative to the surface about the common z axes. Θ defines a rotation of the molecular axes about x_M . Ψ defines the subsequent rotation of the molecular frame about z_M . Notice that the order of application of these rotations is uniquely defined. A few specific examples of some simple orientations and the associated combinations of Euler angles were given in the earlier paper concerned with adsorbed pyridine that used this same convention [27]. One important special case that is relevant to the present study is when $\Psi=0^{\circ}$; in this case Θ and Φ are the tilt and twist of the molecular plane (as discussed in section 3.1 in the context of the NEXAFS data) relative to the idealised case of $\Phi = \Theta = \Psi = 0^{\circ}$.

This second stage of structural optimisation based on the N(3) N 1s PhD data also allowed for local relaxation of the outermost Cu atoms perpendicular to the surface, specifically considering relaxation of the Cu atom directly below the N atom, and the Cu atoms almost directly below the O atoms, for the atop adsorption geometry. The optimised N-atop geometry led to an *R*-factor value of 0.149; in this solution the optimum values of the Euler angles were $\Phi=2^\circ$, $\Theta=22^\circ$ and $\Psi=9^\circ$, showing twist and tilt angles in good agreement with the NEXAFS results (the full set of final optimised parameter values are given in Table 1). For adsorption in the hollow site a local minimum in the *R*-factor was found with a value of 0.184, and associated Euler angles similar to those for the atop site, so the *R*-factor value exceeds that of the atop geometry by exactly the variance in R_{\min} (=0.035), and is formally just within the acceptable range of possible solutions. In fact, additional local minima in the *R*-factors were found for both adsorption sites with much larger tilt angles of 65°; for the atop site the associated value of *R* was 0.166, and for the hollow site *R*=0.175, both within the variance of R_{\min} . However, this large tilt angle is clearly inconsistent with the NEXAFS results.

For the $N(1) \times 1$ PhD spectra, calculations were only performed for the basic adsorption sites clearly favoured by the N(3) N 1s data, namely those with the N(3) atom in atop or hollow sites. Whereas the N(3) PhD data are clearly dominated by backscattering from the near-neighbour Cu substrate atoms, the weaker N(1) PhD modulations are expected to arise mainly from intramolecular scattering, because although the C, N and O atoms are weak scatterers, the more strongly scattering Cu atoms are much further away from the N(1) atom. Necessarily, therefore, these N(1) PhD simulations included the intramolecular scattering from the outset, but assuming, always, a rigid molecular thymine structure to define the relative locations of the C, O and N atoms. For the geometry with the N(3) atom in an atop site a lowest R factor value of 0.35 was obtained, the optimised Euler angles being $\Phi=2^\circ$, $\Theta=18^\circ$ and $\Psi=3^\circ$, in rather good agreement with the values obtained from the NEXAFS and the N(3) PhD spectra. The fact that the lowest *R*-factor value for the N 1s PhD data from the N(1) atom is very significantly larger than that from the N(3) atom is not surprising; the PhD modulations are much weaker for the N(1) emitter due to the absence of near-neighbour strongly-scattering Cu atoms, and under these circumstances the experimental scatter is larger and the fit will become more sensitive to any weaknesses in the theory. Interestingly, the fit did nevertheless show some sensitivity to the location of the molecule relative to the substrate, yielding almost the same N(3)-Cu optimum distance as for the N(3) PhD data (1.98 Å and 1.96 Å, respectively), albeit with lower precision. Further evidence of the role of substrate scattering was the fact that the lowest *R*-factor obtained for hollow site adsorption of the molecule through the N(3) atom was notably higher at 0.40, although because R_{\min} is large, this value lies within the (larger) variance of R_{\min} . Of course, we might expect that

the N(1) PhD spectra would also be sensitive to subtle changes in the intramolecular bondlengths and bond angles, but the data range relative to the potential number of variables do not warrant detailed investigation of this dependence, although the effect of a small change in the molecular conformation was investigated in a final optimisation stage described below.

For the O 1s PhD data, the initial calculations followed the same approach as those for the N(1) emitter. The data analysis described so far clearly identifies the N(3) atoms as atop a surface Cu atom (in the outermost or second layer), and shows that the molecular plane is almost aligned within the [110] azimuth (the Euler angles Φ and Ψ are both small). On this basis, the two O atoms must be both close to atop sites on the Cu(110) surface, so their PhD spectra can be expected to be dominated by scattering from the near-neighbour substrate atoms. However, the relative positions of the O atoms to the substrate must also be defined by the molecular conformation, so we again first considered a rigid thymine molecular structure on the Cu(110) substrate, although the outermost Cu atoms were allowed to move perpendicular to the surface to produce a locally relaxed (rumpled) surface layer. Notice that because of the intrinsic asymmetry of the thymine molecule, we cannot assume that the two O atoms are in symmetrically equivalent sites, so the calculated PhD spectra from each of the two inequivalent emitter sites must be summed (incoherently). This further increases the computational time which, for all the different emitters considered, must also sum over all symmetricallyequivalent domain structures for each low-symmetry local geometry. For the O 1s PhD data this led to a lowest *R*-factor value of 0.19 for adsorption with the N(3) atom atop an outermost layer Cu atom, and a value of 0.22 with the N(3) atom in the hollow site (atop a second layer Cu atom). Thus, the N(3) atop geometry is clearly favoured over the hollow site geometry for all three emitter species, but here too, for the O 1s PhD spectra, the difference in the *R*-factors for the two sites is not formally statistically significant, lying within the variance of 0.035. The optimised values of the Euler angles obtained from the O 1s PhD spectra also fit the consistent picture of the molecule tilted by $\sim 20^{\circ}$ and almost exactly aligned along the [110] azimuth.

Of course, the O 1s PhD spectra are dominated by backscattering from the underlying Cu atoms, and particularly from the nearest-neighbour Cu atoms almost directly below these emitters. Thus, the structural parameters to which these spectra are most sensitive are the O-Cu nearest-neighbour distances, while the next most important parameters are the distances to the underlying Cu layers, parameters that define the local outermost-layer Cu atom relaxations. Clearly, the fact that both the N(3) atom and the two O atoms are bonding to the substrate must mean that the constraint that the relative positions of these three bonding atoms is defined exactly by the geometry of the isolated thymine molecule is unreasonable, and some slight distortion of the molecule may be permitted. To address this point a final set of calculations was performed for all three types of emitter atom in which the two oxygen atoms were allowed to move within the molecular plane, both parallel and perpendicular to the N(3)-C(6) direction, and thus effectively (largely) perpendicular to the surface and parallel to the surface along (110). This allowed the O and N(3) heights above their nearest neighbour Cu atoms, and also above the underlying substrate, to vary independently, as well as allowing the off-atop distance of the O atoms to vary, both movements introducing a small distortion into the adsorbed thymine molecule. In this final set of calculations the N(3)-Cu bondlength (equal to the difference in height of these two atoms above the surface, z_{N-Cu}) was constrained for the N(1) PhD calculations to the (most precise) value of 1.96 Å given by the fit to the N(3) PhD data. Moreover, in this final N(1) PhD optimisation, just one intramolecular distance, that between the N(1) atom and the adjacent C(2) atom, was allowed to change; this C atom is almost directly below the N(1) emitter and so ideally located to contribute significantly to the PhD modulations in near-normal emission. The optimal change in this parameter, defined as $\Delta d(N(1)-C(2))$, is given in Table 1 together with values of all other structural parameter obtained from this final optimisation (although, for completeness, the actual value and precision of the z_{N-Cu} value obtained from the N(1) PhD data when this parameter was a free variable is included). Fig. 6 shows the generally good agreement of the experimental and theoretically simulated PhD spectra for this best-fit structure, while fig. 7 provides a schematic diagram of the structure. The resulting *R*-factor values, with the molecule atop the outermost layer, were 0.126, 0.364 and 0.162 for the N(3), N(1)and O emitters, respectively. Similar optimisations for the molecule atop the second Cu

layer (i.e. with the N(3) atom in a hollow site) yielded *R*-factors for the N(3) and O emitters of 0.181 and 0.203, both outside the variance of the best-fit atop site. We may therefore formally conclude that the molecule does lie atop the outermost layer Cu atoms.

4. General discussion and conclusions

This detailed structural analysis of the surface species produced by deposition of thymine onto a Cu(110) surface at room temperature provides a clear and consistent picture of the surface geometry. The strong and almost identical PhD modulation energies at normal emission from the O emitter and one of the chemically-shifted N emitters clearly shows that these three atoms all lie close to atop surface Cu atoms, a situation only possible if this N 1s component is assigned to a deprotonated N(3) atom. The significant chemical shift between the two N emitters then strongly suggests that the N(1) atom retains its H atom. The fact that this N(1) N 1s component is transformed to the same energy as that of the N(3) atom on heating to a temperature known to lead to desorption of molecular hydrogen clearly indicates that this heating leads to deprotonation of the N(1) atom. In view of this, one might have expected some change in the substrate bonding, but the PhD data show clearly that is cannot be the case: the O 1s PhD data from the adsorbate before and after this heating are essentially identical, while the PhD spectra from the single N 1s component after the heating are also essentially identical to the PhD spectra obtained from the sum of the N(1) and N(3) emitters before heating. Of course, the modulations of the combined N 1s spectra are dominated by those of the N(3) emitter which shows strong substrate backscattering, so some displacement of the N(1) due to the electronic rearrangement in the molecule following the deprotonation, and a consequential change in the exact molecular conformation, cannot be excluded. However, the bonding geometry to the surface is unchanged, as is the molecular orientation, the latter conclusion being supported by the NEXAFS measurements.

Both the NEXAFS and the PhD analysis show that the molecular plane stands up on the surface, a conclusion consistent with the prior studies using vibrational spectroscopy, although our results show a consistent trend for a preferred tilt of the molecule plane

away from the surface normal of approximately 20°. We should remark that the agreement regarding this tilt angle between the two methods is rather significant. In the case of NEXAFS, one might suggest that the average tilt angle actually results from the coexistence of two species, one untilted and the other 'lying down' (i.e. heavily tilted). The origin of the tilt angle determination in PhD, however, is completely different, and it is unlikely that in such a two-state scenario the two methods would reach the same conclusion as to the tilt angle. Nevertheless, the origin of this tilt is unclear, but could, perhaps, be related to intermolecular interactions. In the same way, the reason for a possible twist of the molecule out of the high-symmetry azimuth has no obvious origin. The numerical values for the optimal twist angles for NEXAFS (20°) and PhD (4°) have estimated precision values that could be consistent with zero twist, yet the non-zero amplitude of the NEXAFS resonance in the geometry for which this would be forbidden in the absence of twist does suggest that the molecules are not perfectly azimuthally aligned. This effect could, however, be due to the presence of a second minority species on the surface.

The bonding geometry, with the two O atoms and the N(3) atom all having nearneighbour Cu atoms at distances consistent with chemisorption bonds, is that favoured by some, but not all, of the previous spectroscopic studies; interpretation of the vibrational spectroscopic data has certainly proved controversial, but the quantitative determination of the local geometry fully resolves this issue. The actual Cu-O and Cu-N bonding distances are very consistent with previous determinations of other molecular adsorbates on Cu surfaces, and indeed on Cu(110). For example, the Cu-O distances for thymine on Cu(110) of 1.91 ± 0.03 Å and 2.03 ± 0.03 Å may be compared with values of 1.91 ± 0.04 Å for the acetate species, CH₃COO, [29] and 1.91 ± 0.02 Å for the benzoate species, C₆H₅COO [30], both on the same surface and both also with the O atom in atop sites. The two distinct Cu-O distances in the case of thymine presumably reflect the constraints imposed by the matching of all three bonding atoms (the two O atoms and the N atom) to three Cu atoms in the Cu(110) surface. By contrast, the (more symmetric) acetate and benzoate species bond only through the two O atoms which are expected to be in symmetrically equivalent sites. Similarly, the Cu-N distance for the thymine bonding of 1.96 ± 0.02 Å may be compared with values of 2.00 ± 0.02 Å for pyridine, C_5H_5N , [31], 2.04 ± 0.02 Å for 2-methyl-pyridine [32], both in near atop sites and 2.00 ± 0.04 Å for ammonia, NH₃ [25] in an off-atop site. Strictly, these comparisons indicate that in the thymine bonding the C-N bond is slightly shorter, and the Cu-O bonds slight longer, than in these other examples in which the molecular bonding is only through either O or N atoms. This, too, may be attributed to the constraints of matching the thymine molecular conformation to that of the underlying Cu(110) surface. In this case achieving optimum local bondlengths may require a distortion of either the Cu(110) surface, or the molecule, or both. Consistent with this idea is the fact that the PhD analysis shows a trend for the Cu atoms bonded to the O atoms to relax outwards towards the molecule and for the Cu atoms bonded to the N atom to relax inwards away from the molecule, distortions consistent with an attempt to shorten the Cu-O bonds and lengthen the Cu-N bond. We note, however, that none of these distortions in Table 1 are formally significant, so it is not clear that firm conclusions of this kind are warranted. A similar caveat applies to the intramolecular distortions listed in Table 1.

Nevertheless, the results presented here provide the first definitive quantitative structure determination of any nucleobase on a well-characterised surface, and resolve qualitative controversies over the local bonding of thymine to Cu(110) arising from ambiguities in the interpretation of prior spectroscopic studies.

Acknowledgements

The authors acknowledge the financial support of the Physical Sciences and Engineering Research Council (UK) and the award of beamtime at the BESSY II facility.

Table 1

Optimised values of the structural parameters obtained from the different components of the NEXAFS and PhD analyses for the surface species resulting from room temperature exposure of the Cu(110) surface to molecular thymine. All Δz values are defined as positive for displacements outwards from the surface; $\Delta x(O)$ is defined as positive outwards away for the N(3) atom. The O(2) and O(4) labelling refers to the O atoms bonded to the C(2) and C(4) atoms, respectively.

parameter	N(3) PhD	N(1) PhD	O PhD	NEXAFS
z _{N-Cu} (Å)	1.96±0.02	1.98±0.09	1.94±0.10	-
$z_{\text{O-Cu}}(\text{\AA})$	1.78±0.17 O(2)	1.93±0.18 O(2)	1.91±0.03 O(2)	-
	2.02±0.17 O(4)	2.02±0.18 O(4)	2.03±0.03 O(4)	
Φ (°)	6±9	4±9	2±5	20±15
Θ (°)	24±10	16±6	22±5	30±10
Ψ (°)	8±8	5±9	5±3	0 assumed
$\Delta z(Cu_N)$ (Å)	-0.08±0.10	-0.05±0.25	-0.05±0.10	-
$\frac{\Delta z(Cu_N) (\text{\AA})}{\Delta z(Cu_O) (\text{\AA})}$	-0.08±0.10 0.08±0.15	-0.05±0.25 0.02±0.12	-0.05±0.10 0.05±0.10	-
$\begin{array}{c} \Delta z(Cu_{N}) (\text{\AA}) \\ \\ \hline \Delta z(Cu_{O}) (\text{\AA}) \\ \\ \hline \Delta x(O) (\text{\AA}) \end{array}$	-0.08±0.10 0.08±0.15 0.04±0.08	-0.05±0.25 0.02±0.12 0.07±0.22	-0.05±0.10 0.05±0.10 0.03±0.07	-
$\begin{array}{c} \Delta z(Cu_N) (\text{\AA}) \\ \\ \Delta z(Cu_O) (\text{\AA}) \\ \\ \Delta x(O) (\text{\AA}) \\ \\ \Delta z(O) (\text{\AA}) \end{array}$	-0.08±0.10 0.08±0.15 0.04±0.08 0.04±0.12	-0.05±0.25 0.02±0.12 0.07±0.22 0.09±0.16	-0.05±0.10 0.05±0.10 0.03±0.07 0.14±0.12	-

Figure Captions



Fig.1 Molecular structure of thymine showing the labelling convention for the different constituent atoms



Fig. 2 Some possible adsorption bonding configurations of different thymine-derived species on Cu(110)



Fig. 3 Soft X-ray photoelectron spectra in the energy range of the O 1s, N 1s and C 1s emission peaks from thymine deposited on Cu(110) and room temperature, at 530 K, and deposited at room temperature and subsequently heated to ~520 K. All absolute binding energies are uncalibrated.



Fig. 4 O K-edge NEXAFS data from thymine deposited onto Cu(110) at either room temperature or at 530 K. On the left are shown the NEXAFS spectra, recorded at grazing incidence and at normal incidence, in each of the two principle azimuths from the room temperature deposition. On the right are shown the variation in the intensity of the near-edge π -resonance as a function of the angle of the polarisation E-vector of the incident radiation relative to the surface normal. The dashed lines passing through the experimental data points are theoretical fits yielding information of the molecular orientation, as described in the text.



Fig. 5 Summary of the experimental O 1s and N 1s PhD modulations spectra obtained in the [001] azimuth at different polar emission angles from both room temperature and 530 K deposition of thymine on Cu(110). In the case of the N 1s data (c) and (d) show the separate spectra obtained from the two distinct chemically-resolved N 1s photoemission peaks seen after room temperature deposition. The panel (b) compares the PhD spectrum from the sum of these components with that obtained from the single unresolved N 1s component from the 530 K deposition.



Fig. 6 Comparison of experimental (full lines) N 1s and O 1s PhD modulation spectra from thymine deposited at room temperature onto Cu(110) with theoretical simulations (dashed lines) for the best-fit structure, the parameter values of which are given in Table 1.



Fig. 7 Schematic diagram of the optimised structure for adsorbed thymine on Cu(110) looking along the $[1\overline{10}]$ and [001] azimuthal directions. Note that the H atoms are omitted from this figure as the results presented here provide no direct information on the location of these atoms.

References

- 1 C. Hamai, H. Tanaka, T. Kawai, J. Phys. Chem. B 104 (2000) 9894
- 2 H. Tanaka, T. Kawai, Surf. Sci. 539 (2003) L531
- 3 D.P.Woodruff and A.M.Bradshaw Rep.Prog.Phys. 57 (1994) 1029
- 4 D. P. Woodruff, Surf. Sci. Rep. 62 (2007) 1
- 5 N.A.Booth, D.P.Woodruff, O.Schaff, T.Gieβel, R.Lindsay, P.Baumgartel and A.M.Bradshaw Surf.Sci. 397 (1998) 258
- 6 J.-H. Kang, R. L. Toomes, M. Polcik, M. Kittel, J.-T Hoeft, V. Efstathiou, D. P. Woodruff and A. M. Bradshaw, J.Chem.Phys. 118 (2003) 6059
- 7 D.I. Sayago, M. Polcik, G. Nisbet, C.L.A. Lamont and D.P. Woodruff, Surf. Sci. 590 (2005) 76
- 8 M Furukawa, H. Fujisawa, S. Katano, H. Ogasawara, Y. Kim, T. Komeda, A. Nilsson,
 M. Kawai, Surf. Sci. 532-535 (2003) 261
- 9 A. McNutt, PhD thesis, University of Liverpool, 2002
- 10 T. Yamada, K. Shirasaka, A. Takano, M. Kawai, Surf. Sci. 561 (2004) 233
- 11 A. McNutt, S. Haq, R. Raval, Surf. Sci. 502-503 (2002) 185
- 12 M. L. M. Rocco, R. Dudde, K. –H. Frank, E. –E. Koch, Chem. Phys. Lett. 160 (1989) 366
- 13 W. Haiss, B. Roelfs, S. N. Port, E. Bunge, H. Baumgärtel, R. J. Nichols, J. Electroanal. Chem. 454 (1998) 107
- 14 W. -H. Li, W. Haiss, S. Floate, R. J. Nichols, Langmuir 15 (1999) 4875
- 15 K. J. S. Sawhney, F. Senf, M. Scheer, F. Schäfers, J. Bahrdt, A. Gaupp, W. Gudat, Nucl. Instrum. Methods A 390 (1997) 395
- 16 J. C. Fuggle, N. Mårtensson, J. Electr. Spectrosc. Rel. Phenom. 21 (1980) 275
- 17 B.V. Christ, Handbook of Monochromatic XPS Spectra (XPS International, Inc., 1999)
- 18 K. Fujii, K. Akamatsu, A. Yokota, J. Phys. Chem. B 108 (2004) 8031
- 19 S. Söhnchen, S. Lukas, G. Witte, J. Chem. Phys. 121 (2004) 525
- 20 J. Stöhr and R. Jaeger, Phys. Rev. B 26 (1982) 4111
- 21 V. Fritzsche, J. Phys.: Condens. Matter 2 (1990) 1413

- 22 V. Fritzsche, Surf. Sci. 265 (1992) 187
- 23 V. Fritzsche, Surf. Sci. 213 (1989) 648
- 24 J. B. Pendry, J. Phys.C: Solid State Phys. 13 (1980) 937
- 25 N. A. Booth, R. Davis, R. Toomes, D. P. Woodruff, C. Hirschmugl, K.-M. Schindler,
 O. Schaff, V. Fernandez, A. Theobald, Ph. Hofmann, R. Lindsay, T. Giessel, P.
 Baumgärtel, A. M. Bradshaw, Surf. Sci. 387 (1997) 152
- 26 http://ndbserver.rutgers.edu/archives/proj/valence/bases4.html
- 27 T.Gieβel, O.Schaff, R.Lindsay, R.Terborg, P.Baumgärtel, J.T.Hoeft, M.Polcik,
 A.M.Bradshaw, A.Koebbel, D.R.Lloyd and D.P.Woodruff, J.Chem.Phys. 110 (1999)
 9666
- 28 R.Terborg, M.Polcik, J-T.Hoeft, M.Kittel, M.Pascal, J.H.Kang, C.Lamont, A.M.Bradshaw and D.P.Woodruff, Surf.Sci. 457 (2000) 1
- 29 K-U.Weiss, R.Dippel, K-M.Schindler, P.Gardner, V.Fritzsche, A.M.Bradshaw, A.L.D.Kilcoyne and D.P.Woodruff Phys.Rev.Lett. 69 (1992) 3196
- 30 M. Pascal, C.L.A. Lamont, M. Kittel, J.T. Hoeft, R. Terborg, M. Polcik, J.H. Kang, R.Toomes and D.P. Woodruff, Surf. Sci. 492 (2001) 285
- 31 T.Gieβel, O.Schaff, R.Lindsay, R.Terborg, P.Baumgartel, J.T.Hoeft, M.Polcik, A.M.Bradshaw, A.Koebbel, D.R.Lloyd and D.P.Woodruff, J. Chem. Phys. 110 (1999) 9666
- 32 R.Terborg, M.Polcik, J-T.Hoeft, M.Kittel, M.Pascal, J.H.Kang, C.Lamont, A.M.Bradshaw and D.P.Woodruff, Surf. Sci. 457 (2000) 1