Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics Letters 376 (2003) 438–444
CHEMICAL PHYSICS LETTERS
www.elsevier.com/locate/cplett

Stacking modification in DNA-base molecule structures: X-ray diffraction on 5-Chlorouracil on Ag(1 1 1)

H.L. Meyerheim [a,*], H. Maltor [a], I.K. Robinson [b]

[a] *Institut für Kristallographie and Angewandte Mineralogie, Ludwig-Maximilians Universität München, Theresienstraße 41, D-80333, Germany*
[b] *Department of Physics, University of Illinois, Urbana, IL 61801, USA*

Received 2 June 2003; in final form 2 June 2003
Published online: 8 July 2003

Abstract

The geometric structure of the DNA-base derivative 5-Chlorouracil deposited on Ag(1 1 1) was analyzed using surface X-ray diffraction. Although the ad-layer structure bears many similarities with its bulk analog (e.g., lattice metric, hydrogen bonded molecules within each molecular layer and van der Waals interconnected layers), we find that subtle modifications of the interlayer molecular arrangement prevents the growth of films thicker than three layers.

During the last two decades the preparation and characterization of self-assembled organic films deposited on single crystalline substrates has become a major subject in solid-state physics [1]. Organic molecules are seen as highly important materials for potential applications in technology, such as for organic opto-electronic and semiconductor devices [2–6]. The surface interactions of DNA-base molecules as well as their derivatives are of particular interest, since these are the basic components of the genetic information [7,8]. In this context, manipulation of their growth, structure and related properties is a primary issue for understanding fundamentals of genetic mechanisms, which for example can lead to mutations and the development of cancer. Furthermore, the specific synthesis and use of modified molecules is a prerequisite for tailoring more efficient pharmacological substances and procedures.

Since the 1950s the 5-halogenated uracils were used as anti-tumor agents, because they are preferentially incorporated into the DNA instead of thymine and were found to act as anti-metabolites [9,10]. The structure formula of uracil and two of its derivatives are shown in Fig. 1a. In the case of uracil a H-atom is bonded to the C-atom at the position #5 of the pyrimidine ring (R = H). Replacing the H-atom either by $R = CH_3$ or by R = Cl results in the derivatives thymine and 5-Chlorouracil (5-ClU), respectively.

* Corresponding author. Present address: Max-Planck-Institut für Mikrostrukturphysik, Weinberg 2, D-06120 Halle, Germany. Fax: +49-345-5511-223.
E-mail address: hmeyerhm@mpi-halle.mpg.de (H.L. Meyerheim).

0009-2614/03/$ - see front matter
doi:10.1016/S0009-2614(03)00948-5

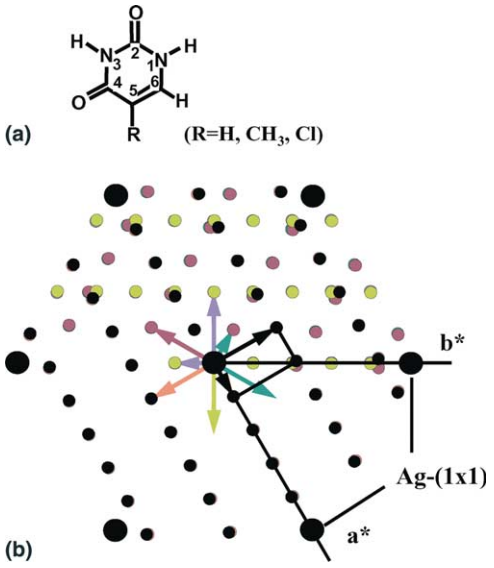

Fig. 1. (a) Structure formula of the pyrimidine ring with the possible residuals (R) bonded to the C-atom at the position 5: R = H (uracil), R = CH_3 (thymine) and R = Cl (5-Chlorouracil). (b) Sketch of the reciprocal space for 5-ClU/Ag(1 1 1) projected along c^*. Large circles represent Ag(1 1 1)–(1 x 1) spots. All other symbols are from the superstructure. Different colors correspond to different domains. Long and short arrows correspond to the b^*- and a^*-basis vectors of the different domains. The lower part shows only the reflections of one domain (•). Its unit cell is indicated by the solid rectangle.

A thorough compilation of the crystal structures of many nucleic acid constituents has proven that the vertical stacking of the pyrimidines and purines in polynucleotides is similar to that observed in crystals of nucleic acid constituents. In general, the layer interaction is dominated by dipole interactions, which involves the polar region of one molecule with the polarizable ring system of the adjacent layer. In particular, the structures of the 5-halogenated uracils are representative for an intimate interaction of the (electronegative) halogen with the C-atom at the position #2 of the adjacent ring, while the intra-layer hydrogen bonding scheme is little affected by halogen substitution [11].

In general the properties of crystals and molecules can be influenced by chemical substitution, but also by the reduction of dimensionality. The transition from three-dimensional to two-dimensional structures is known to significantly influence both, structure and physical properties. In order to study the effect of dimensionality and the influence of a weakly interacting noble metal surface on the intermolecular 5-ClU structure, we have carried out a surface X-ray diffraction (SXRD) study of 5-ClU deposited on Ag(1 1 1). This provides important insights into the mechanisms influencing the structure of this DNA-base molecule.

Surface X-ray diffraction is the method of choice for investigating the geometric structure of large organic molecules. This is because the commonly employed techniques for the analysis of molecular ad-layer structures like scanning tunneling microscopy (STM) and low energy electron diffraction (LEED) either do not provide submolecular resolution or have to surrender in the case of large unit cells. We find that while the incommensurate ad-layer structure in many respects bears close similarities with its bulk analog, only a three-layer film can be grown. This is directly correlated to the slightly modified inter-layer interaction and stacking scheme between the layers.

The experiments were carried out at the beamline X16A of the National Synchrotron Light Source (NSLS) in Brookhaven (USA) using the ultra-high-vacuum (UHV) diffractometer operated in the 5-circle mode [12]. The Ag(1 1 1) surface was cleaned by standard procedures after which 5-ClU was deposited by thermal sublimation from a Knudsen cell on the surface kept at room temperature. Mild annealing at 85 °C for a few minutes significantly improved the surface order as determined from the sharpness of the LEED spots. Fig. 1b schematically shows the reciprocal space viewed along the c^*-axis. The large filled circles correspond to the first-order Ag reflections, while all other symbols represent the superstructure spots of the molecular structure, which is arranged in six domains. The spots of each domain are represented by a different color of the circles. It should be noted that due to the particular lattice metric of the superstructure, the reflections of two domains always nearly overlap and cannot resolved. In the lower part the symbols of only one domain represented by the small filled circles (•) are shown. The solid rectangle indicates the corresponding reciprocal unit cell while the two black arrows represent its a^*- (short arrow) and b^*-axis

(long arrow). The b^* lattice vectors of the other domains are represented by the other long vectors in Fig. 1b. The matrix relating the unit cell to the Ag(1 1 1)–(1 x 1) mesh equals to

$$M = \begin{pmatrix} 5.013 & 2.495 \\ 0.012 & 2.357 \end{pmatrix}.$$

Error bars for the determination of the matrix elements are in the 10^{-4} range. The transformation matrix is related to a nearly rectangular two-dimensional unit cell with parameters $a_0 = 12.54$ Å, $b_0 = 6.79$ Å and $\gamma = 89.94°$, which is very close to the corresponding dimensions of the bulk unit cell in the equivalent setting ($a_0 = 12.42$ Å, $b_0 = 6.85$ Å and $\gamma = 90°$). Thus, a close similarity between ad-layer and bulk can be anticipated. We did not find any LEED spots related to Cl/Ag(1 1 1) superstructures reported so far [13–15], which could indicate the instability of the 5-ClU molecule upon adsorption as in the case of 5-Iodouracil [16].

Integrated SXRD reflection intensities were collected in situ under total reflection conditions of the incident beam by rotating the sample around its normal [17]. Since organic molecules are in general composed of weakly scattering atoms and their structures are characterized by large unit cells (85 Å^2 in the present case) and low plane group symmetries, SXRD reflection intensities are extremely weak [16,18,19]. Good intensity statistics was achieved by long counting times (typically 10 s per data point) resulting in a statistical error below 10%. Furthermore, the high resolution of SXRD allows the separation between reflections belonging to different domains. Note that due to the particular superlattice metric the reflections of two domains always nearly overlap and a k-space resolution in the 10^{-4} Å^{-1} (without the factor 2π) range is a prerequisite for measuring reliable intensities. The measurement of several control reflections indicated that the X-ray beam induces some damage into the molecular film (about 15% intensity reduction during 10 h), making the repeated preparation of the superstructure necessary during data collection.

In total 94 in-plane ($hk0$), reflection intensities were collected reducing to 48 independent reflections by averaging over symmetry equivalent reflections and over equivalent reflections belonging to different domains. In addition, the intensity distribution along nine independent superlattice rods was measured up to a maximum normal momentum transfer of $\ell = q_z/c^* = 2.4$ reciprocal lattice units (rlu, 1 rlu $= c^* = 0.886$ Å^{-1}). This adds to a total of 202 reflections for the whole data set. The structure factor intensities, $|F(hk\ell)|^2$, were obtained after correcting the intensities for effective sample area, Lorentz- and polarization factor [20]. The standard deviations (σ) of the $|F(hk\ell)|^2$ were derived from the reproducibility of symmetry equivalent reflections and the counting statistics [17]. In general they are in the 10–20% range.

In first step of the analysis the Fourier transformation of the in-plane, $|F(hk0)|^2$ was calculated yielding the Patterson function, $P(u,v)$, of the projected structure. It is shown in Fig. 2a. In general, $P(u,v)$ provides a picture of the interatomic vectors within the unit cell, weighted by the atomic number and their multiplicity [21]. Here, $P(u,v)$ is characterized by a quasi-hexagonal arrangement of broad peaks separated by about 4 Å, all of comparable height. Interestingly, $P(u,v)$ can be interpreted qualitatively by the bulk structure. Figs. 2b, c show the bulk structure and the best fit model for the 5-ClU/Ag(1 1 1) ad-layer, respectively. In both cases the viewing direction is perpendicular to the ring planes. In the bulk structure this corresponds to the (1 0 2) plane. In this setting the lattice parameters of ad-layer and bulk structure differ by less than 1%. Along the normal direction, one bulk unit cell (within the infinite stack of layers) is composed of three molecular sheets, where from the bottom to the top the rings are represented by black (III), gray (II) and open (I) hexagons, respectively. One isolated molecule with its O- and Cl-atoms indicated by the red and green sticks, respectively, is shown in the upper right part of the figure.

The bulk structure is characterized by two basic features: first, within the layers there is a zig-zag pattern of hydrogen bonded molecules. For the top layer this is emphasized by the lines on the left. Second, the layers are interconnected by dipole–dipole interactions between the Cl-ligands and the C-atom at the position #2 of the pyrimidine ring of the adjacent layer (distance 3.35 Å). The Cl- and

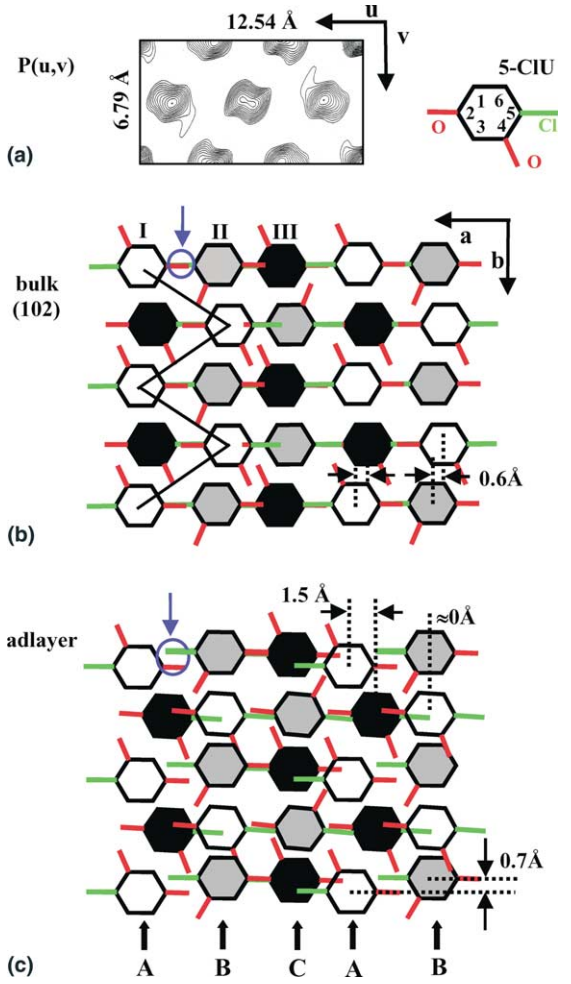

Fig. 2. (a) $P(u,v)$ for 5-ClU/Ag(1 1 1). (b) Structure of bulk 5-ClU in projection normal to the (1 0 2) plane. Molecules are drawn from the bottom to the top by black (III), gray (II) and open (I) hexagons, respectively. Hydrogen bonded molecules within each layer form a zig-zag pattern as indicated by the solid line for the top layer on the left. (c) Structure of 5-ClU deposited on Ag(1 1 1). Molecules are labeled as in (b). Some distances between the centers of the pyrimidine rings are indicated (for details see text).

C-atoms are arranged almost vertically above or below each other (see blue circle and the arrow on the upper left of Fig. 2b). The molecular arrangement along the *a*-axis is given by a (III, II, I, III, ...) sequence. Along the *b*-axis there is also a sequence of three different types of chains labeled by A, B, C, A, ... (see lower part of the figure) consisting of type (I, III)-, (I, II)- and (II, III)-molecules, respectively. Along these chains the centers of neighboring rings are displaced by about 0.6 Å relative to each other forming a zig-zag chain as well.

To first order, the quasi-hexagonal arrangement of the peaks in $P(u,v)$ derived from the surface structure can be related to the bulk arrangement of the molecules, while individual inter-atomic vectors cannot be resolved. Moreover, the uniform peak height distribution in $P(u,v)$ can be explained by considering each molecule as a single scattering unit. This is because a primitive hexagonal arrangement of identical scatterers leads to a uniform peak height distribution in the corresponding $P(u,v)$. Thus, so far on the basis of the similarity of the lattice parameters and the semi-quantitative analysis of $P(u,v)$ a close resemblance between bulk and surface structure can be anticipated, however, some differences exist in detail as derived on the basis of the quantitative intensity analysis.

Using the bulk structure as a starting model, the analysis was carried out by last-squares refinement using the program SHELX-97 [22]. In order to minimize the number of free parameters, the pyrimidine rings were treated as planar rigid units (inter-atomic distance within the ring: 1.40 Å) keeping the exo-cyclic ligands at fixed distances from the ring (O, 1.20 Å; Cl, 1.7 Å). Thus, for each molecule there are three positional and three rotational free parameters to refine adding up to 36 in total for all six molecules in the unit cell (plane group p1).

The symbols in Figs. 3a–d represent the $|F(hk\ell)|$'s along several rods. The strong intensity modulation along q_z is only explainable by a multilayer structure. A molecular monolayer would only lead to a monotonic intensity decrease along q_z. The best fit (solid lines) quantitatively expressed by the goodness of fit, GOF = 1.5, and the unweighted residuum, $R_u = 0.25$ [17,23], was obtained by a structure model, which is composed of a three-layer stack of intact 5-ClU molecules similar to the bulk. However, several modifications with respect to the bulk are found as discussed in the following using Fig. 2c.

The most important modification is that the top-layer molecules (white hexagons) are shifted

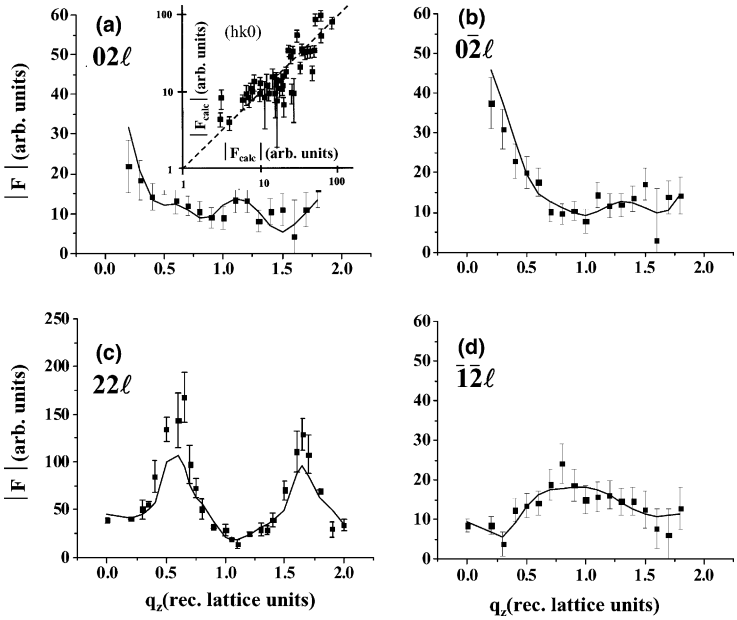

Fig. 3. (a)–(d) Observed (symbols) and calculated (lines) structure factor amplitudes for some superlattice rods. The inset in (a) shows the plot of $|F_{obs}|$ versus $|F_{calc}|$ for the in-plane data.

by about 0.7 Å along the positive *a*-axis relative to the second and bottom layer. The intra-layer zig-zag arrangement of the molecules is basically unchanged. The relative arrangement between the deeper layers (II) and (III) are much less modified, although in detail differences exist with respect to the bulk structure as well (vide infra). The top layer shift is most evident by inspecting the molecular structure along the *b*-axis in comparison with the bulk. In the surface structure the molecular (zig-zag) arrangement within chain C, which contains only the layers (III) and (II), i.e., the layers next to the substrate and the second one, is similar to the bulk. In contrast, due to the shift of the top layer molecules, (I), the corrugation in chain A is increased to approximately 1.5 Å, while it is almost zero along chain B. Furthermore, while in the bulk all molecules are perfectly aligned along the *a*-axis, an about 0.7 Å relative shift and a 5–10° rotation of molecules (with the axis normal to the ring plane) is observed. This affects all molecules and directly indicates that differences to the bulk structure exist in all layers.

As a consequence of the translations and rotations of the molecules relative to each other, the *interlayer* Cl–C distances (3.35 Å in the bulk) are considerably elongated. We find 3.6 Å on average and 4.2 Å for one position between (II) and (III) marked by the blue circle and the arrow in the upper left of Fig. 2c (the error bars for the distance determination are estimated to be ±0.25 Å). Here, the Cl-atoms are no longer closely above (or below) the C-atom of the adjacent layer as in the bulk but are laterally shifted. Consequently, the dipole–dipole interlayer bond is considerably weakened. This is in correspondence with the observation that even mild annealing to 100 °C destroys the superstructure. Finally, the increasing disorder and progressive bond-weakening in the layers is also reflected in the intra-layer (N···O) distances between hydrogen bonded molecules. For the layer (III) next to the substrate N–O distances in the range between 2.7 and 3.1 Å are determined, but in the top layer we find larger distances of 3.2–3.8 Å. At least the latter cannot be seen as an intact hydrogen bond when compared

with the bulk value (2.82 Å). Note, however, that even in the first layer the hydrogen bond length is increased by up to 0.2 Å, reflecting the perturbation of the substrate (see below).

The structure model outlined so far leads to a significant improvement of the fit relative to that obtained for the three-layer bulk starting model (GOF = 1.5 versus 3.0 and $R_u = 0.25$ versus 0.60). Some small deviations along the $(2\,2\,\ell)$ rod are attributed to some extra scattering contribution as a result of a small (≈1%) normal substrate relaxation. This conclusion is supported by good fit ($R_u \approx 0.15$) of the in-plane intensities, $|F(hk0)|^2$, which are not affected by the substrate relaxation. In the inset of Fig. 3a, the $|F_{obs}|$ are closely aligned along the dashed diagonal representing the (ideal) condition $|F_{obs}| = |F_{calc}|$.

We have evidence that from the bottom to the top of the molecular film there is a progressive modification of the mutual arrangement of the molecular sheets relative to the bulk structure leading to a progressive weakening of the inter-layer bonding strength. This mechanism is accounted for inhibiting the growth of a film thicker than three layers. We argue that this can be traced back to the (weak) chemisorptive interaction of the first layer Cl-atoms and to the interaction of the whole pyrimidine π-system with the metal surface. In this model, the dipole interaction to the second layer is weakened by the depolarization of the Cl-atoms as well as by the depolarization of the pyrimidine rings, which in turn interact with the Cl-atoms of the second layer. Our structure analysis shows that the weakening of the inter-layer dipole–dipole interaction is already more pronounced at the second interface (i.e., between layers II and I). This is indirectly evident by the significant mis-arrangement of the top layer relative to the second one. Thus, we conclude, that once the intra-layer structure and bonding is disturbed by the substrate it rapidly shakes up and no ordered films thick than three layers can be grown at room temperature.

It should be emphasized that on the basis of the SXRD data alone it is not directly possible to unambiguously decide whether layers (I) and (III) are at the top and at the bottom of the three-layer stack, respectively, or vice versa. However, the model based on a reversed layer sequence [i.e., with (III) top and (I) bottom next to the Ag surface] is much less plausible, since it would not explain why the film-growth stops just after three layers. One could expect a film composed of more than three layers. This is because in this model the structural mis-arrangement between the layers *decreases* with increasing thickness, which could favor the adsorption of additional layers.

Finally, it should be noted that the basic similarity between bulk and ad-layer structure is a clear indication for the tendency of the 5-halogenated uracils to form its characteristic stacking scheme even on an 'arbitrary' an organic substrate. It also proves the importance of self-assembly processes in DNA-base molecule structures even in complicated cases. The main result of our study – the substrate induced modifications might have wide implications for developing suitable procedures for the specific manipulation of the structural arrangement in this important class of molecules.

In summary, our SXRD study on has provided a detailed picture of the structure of 5-ClU deposited on Ag(1 1 1). Although the basic features of the molecular ad-layer structure resemble those found in the bulk structure, we determine subtle changes in the inter-layer bonding configuration, which are made responsible for the fact, that the ordered growth of the molecular stack is limited to three layers only.

Acknowledgements

The authors thank W. Moritz and S. Werner for helpful discussions. The authors also acknowledge the help of M. De Santis during the experiments. Support by the Sonderforschungsbereich 338 of the German Science Foundation (DFG) is acknowledged. The X16A facility at NSLS is supported by the US Department of Energy under contracts DEAC02-98CH10886 and DEFG02-ER45439.

References

[1] S.R. Forrest, Chem. Rev. 97 (1997) 1793.
[2] C. Väterlein et al., J. Appl. Phys. 82 (1997) 3003.

[3] M. Wohlgenannt, K. Tandon, S. Mazumdar, S. Ramasesha, Z.V. Vardeny, Nature 409 (2001) 494.
[4] M. Baldo, M. Thomson, S.R. Forrest, Nature 403 (2000) 750.
[5] M.O. Lorenzo, C.J. Baddeley, C. Muryn, R. Raval, Nature 404 (2000) 376.
[6] S. Alavi et al., Phys. Rev. Lett. 85 (2000) 5372.
[7] T. Boland, B. Ratner, Langmuir 10 (1994) 3845.
[8] J.E. Freund, M. Edelwirth, P. Kröbel, W.M. Heckl, Phys. Rev. B 55 (1997) 5394.
[9] R.J. Rutman, A. Cantarow, Cancer Res. 14 (1954) 19.
[10] P. Duschinsky, E. Pleven, C. Heidleberger, J. Am. Chem. Soc. 79 (1957) 4559.
[11] H. Sternglanz, C.E. Bugg, Biochim. Biophys. Acta 378 (1975) 1.
[12] P.H. Fuoss, I.K. Robinson, Nucl. Instrum. Methods 222 (1984) 171.
[13] G. Rovida, F. Pratesi, Surf. Sci. 51 (1975) 270.
[14] P.J. Goddard, R.M. Lambert, Surf. Sci. 67 (1977) 180.
[15] M. Bowker, K.C. Waugh, Surf. Sci. 134 (1983) 639.
[16] H.L. Meyerheim, Th. Gloege, Chem. Phys. Lett. 326 (2000) 45.
[17] I.K. Robinson, in: G.S. Brown, D.E. Moncton (Eds.), Handbook of Synchrontron Radiation, vol. 3, Elsevier, Amsterdam, 1991.
[18] H.L. Meyerheim, Th. Gloege, H. Maltor, Surf. Sci. Lett. 442 (1999) L1029.
[19] H.L. Meyerheim, T. Gloege, H. Maltor, M. Sokolowski, E. Umbach, P. Bäuerle, Surf. Rev. Lett. 6 (1999) 883.
[20] E. Vlieg, J. Appl. Cryst. 30 (1997) 532.
[21] M.J. Buerger, Contemporary Crystallography, McGraw-Hill, New York, 1970.
[22] G.M. Sheldrick, University of Göttingen.
[23] R_u is defined as $R_u = \sum \left| |F_{(calc)}|^2 - |F_{(obs)}|^2 \right| \Big/ \sum |F_{(obs)}|^2$; $F_{(obs)}$ and $F_{(calc)}$ are the observed and calculated structure factor amplitudes, respectively. The summation runs over all reflections.