Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Comparative experimental electron density and electron localization function study of thymidine based on 20 K X-ray diffraction data

Christian B. Hübschle,[a] Birger Dittrich,[b] Simon Grabowsky,[a] Marc Messerschmidt[c,d] and Peter Luger[a]*

[a]Institut für Chemie und Biochemie/Kristallographie, Freie Universität Berlin, Fabeckstrasse 36a, D-14195 Berlin, Germany, [b]Institut für Anorganische Chemie, Georg-August Universität Göttingen, Tammanstrasse 4, D-37077 Göttingen, Germany, [c]Hasylab at Desy, Notkestrasse 85, 22607 Hamburg, Germany, and [d]SLAC Stanford, 2575 Sand Hill Rd M/S 43A, Menlo Park, CA 94025, USA

Correspondence e-mail: luger@chemie.fu-berlin.de

Received 12 December 07
Accepted 29 February 08

From a high-resolution X-ray data set ($\sin\theta/\lambda = 1.1$ Å^{-1}) measured at 20 K the electron-density distribution of the nucleoside thymidine was derived by a classical multipole refinement and by application of the invariom formalism. Owing to the presence of the heteroaromatic thymine ring system two invariom models were compared which considered the nearest and next-nearest neighbors for the invariom assignments. Differences between the two invariom models were small for the bond topological and atomic properties – about five times smaller than differences with the classical multipole refinement. Even the latter differences are in the uncertainty ranges which are commonly observed in experimental charge-density work and were found in molecular regions involved in intermolecular contacts. The application of the constrained wavefunction-fitting approach allowed the electron localization function (ELF) to be obtained from the experimental X-ray data, which was graphically represented and topologically analyzed. ELF basin populations were derived from experiment for the first time. The electron populations in the disynaptic valence basins were related quantitatively to bond orders.

1. Introduction

Determination of the electron density $\rho(\mathbf{r})$ of a chemical compound from high-resolution single-crystal X-ray diffraction data has seen several favorable developments over the last few years. Experimental advances such as the appearance of CCD area detectors and the use of highly intense synchrotron radiation at short wavelengths around $\lambda = 0.5$ Å have made possible rapid data collections and hence a series of electron density (ED) studies (Luger *et al.*, 2005; Koritsanszky & Coppens, 2001). Bader's quantum theory of atoms in molecules (QTAIM; Bader, 1990) allows a topological analysis of $\rho(\mathbf{r})$ yielding a quantitative description of atoms, bonds, non-bonding interactions, electronic structure and reactivity. As a consequence, the ED research field has recently been declared to have come of age (Coppens, 2005), in that most of the 'teething problems' have been solved. The use of the multipole model for an aspherical description of atomic EDs, introduced by Hansen & Coppens in 1978 (Hansen & Coppens, 1978), can be considered a 'classical' technique today, although some practical difficulties remain (Hübschle *et al.*, 2007). In a 'classical' multipole refinement a large number of multipole parameters are adjusted against the experimentally obtained single-crystal diffraction intensities, so that a considerable number of Bragg reflections is needed, especially in the high-order regions of reciprocal space. Compounds forming only poorly diffracting crystals, which

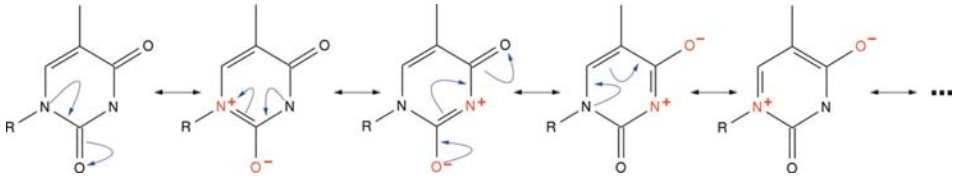

Figure 1
Reasonance formulae of the thymine ring.

holds generally for larger molecules like proteins, are therefore excluded from the application of the method.

To overcome the problem of poorly diffracting crystals, there are considerable research activities in progress to establish databases with transferable non-spherical scattering factors aimed at a significant reduction of parameters in non-spherical atom refinements. All these approaches rely on the transferability of an aspherical atom and the fact that there is additional unmodeled information concerning asphericity in the low-order diffraction data, which is ignored in the independent atom model (IAM). It is expected that once the electron density and the derived properties of an atom or a group in a given chemical neighborhood are known, these properties should be conserved if this atom occurs in another molecule, but in the same immediate chemical environment. This way transferability provides a tool to use these fragments as building blocks for the additive generation of the ED of macromolecules. In various studies the transferability of the electron density has been verified experimentally (Luger, 2007).

The modeling of diffraction data with transferable, experimentally derived, non-spherical scattering factors was first performed in the early nineties (Brock *et al.*, 1991). Transferable non-spherical scattering factors have been introduced by Stewart (Stewart & Bentley, 1975). Koritsanszky *et al.* (2002) have shown that non-spherical scattering factors can be obtained by following a purely theoretical methodology. A database of aspherical scattering factors derived from experimental data has been established since 1995 (Pichon-Pesme *et al.*, 1995; Zarychta *et al.*, 2007). At present, two competing theoretical databases are available (Volkov *et al.*, 2004; Dittrich *et al.*, 2006; Dominiak *et al.*, 2007), one of them represents the invariom[1] approach (Dittrich *et al.*, 2004, 2005) introduced by our group. It uses theoretical multipole population parameters predicted from geometry-optimized model compounds which provide the same (next-) nearest-neighbor environment as the atoms in the molecule of interest. The resulting aspherical pseudoatoms that take into account the effects of chemical bonding are called *invarioms*. If invarioms are entered into the multipole model, the only parameters that are refined against the experimental data are positional and displacement parameters, and the theoretically predicted multipole parameters are kept constant. The invariom model has been tested on several systems (Hübschle *et al.*, 2007; Dittrich *et al.*, 2006).

Our current activities aim to verify entries in a current expanded database, covering a large fraction of organic molecules.[2] Indirectly, such studies also allow empirical rules on the transferability of electron density to be checked. To compare the electron density between experiment and theory we chose to study the second most biologically important class of compounds, nucleosides/nucleotides, the building blocks of DNA, RNA and related macromolecules. As part of several ED studies on nucleosides and bases, thymidine was studied here. The hetero-aromatic thymine ring system in thymidine can be described by the resonance formulae shown in Fig. 1. In this paper we address two related questions. The first chemical question is the extent of delocalization in the ring system of thymidine, where not all double bonds are delocalized to the same extent. Secondly, the following methodological problem of invariom modeling is analyzed: Although the empirical rules in invariom modeling would require next-nearest neighbors for the thymidine ring system (Luger & Dittrich, 2007), we ask whether a nearest-neighbor approximation is sufficient for the reproduction of the ED in thymidine, or whether the consideration of next-nearest neighbors is necessary. In the nearest-neighbors approximation model compounds for the theoretical prediction of multipole parameters include only the first neighbor sphere around the atom of interest. In the next-nearest-neighbor approximation the second neighbor sphere is also taken into account. To answer these two questions for the title compound two models of theoretically predicted multipole populations (ITM and *meso*-ITM, see below) were compared with a 'classical' multipolar refinement (CMR).

To compare the results of the three refinement models, bond topologies, Hirshfeld surfaces, atomic properties and in particular the electrostatic potentials are analyzed. The electrostatic potential was mapped on an electron density isosurface and was quantitatively analyzed following the methodology of Politzer *et al.* (2001). As the electrostatic potential is strongly influenced by the crystal environment, it is a suitable property to investigate differences in the three approaches. To gain deeper chemical insight into the bonding situation in the delocalized ring system, the electron localization function (ELF) was calculated from the X-ray data using the constrained wavefunction approach by Jayatilaka (Jayatilaka & Grimwood, 2001; Grimwood & Jayatilaka, 2001). The X-ray constrained wavefunction fitting (CWF) considers the experimental X-ray data in the self-consistent field calculation. The electron density was integrated in the ELF basins derived from a topological analysis, to show the population of bonds, lone pairs and other bonding effects.

[1] The term invariom was derived from transfer-invariant pseudoatom.

[2] The most current version of the database and the program *InvariomTool* is available online at http://wwwuser.gwdg.de/~chuebsc/dittrich/dev.html.

Table 1
Experimental details.

Crystal data	
Chemical formula	$C_{10}H_{14}N_2O_5$
M_r	242.23
Cell setting, space group	Orthorhombic, $P2_12_12_1$
Temperature (K)	20
a, b, c (Å)	4.846 (2), 13.901 (6), 16.316 (7)
V (Å^3)	1099.2 (8)
Z	4
D_x (Mg m^{-3})	1.464
Radiation type	Mo $K\alpha$
μ (mm^{-1})	0.12
Crystal form, color	Block, colorless
Crystal size (mm)	0.5 × 0.3 × 0.2
Data collection	
Diffractometer	Bruker SMART Apex CCD detector on four-circle diffractometer
Data collection method	φ
Absorption correction	None
No. of measured, independent and observed reflections	61 343, 6295, 5638
Criterion for observed reflections	$I>3\sigma(I)$
R_{int}	0.039
θ_{max}	51.7
Refinement	See Table 3

Computer programs used: *SMART* (Bruker, 2002), *SAINT* (Bruker, 2004).

2. Experimental

Crystals were obtained by slow evaporation of a saturated aqueous solution of thymidine (Sigma-Aldrich). A single-crystal diffraction experiment was carried out on a Huber diffractometer (400 mm Eulerian cradle with offset χ circle) equipped with a double-stage helium cryostat and a Bruker APEX CCD detector. A dataset was collected with Mo $K\alpha$ radiation (graphite monochromator) from a sealed 2 kW X-ray tube at a temperature of 20 K. The reciprocal space was explored by a combination of different φ scans with 2θ (and exposure time) being $-28°$ (5 s) for measurement of low-order reflections, $-58°$ (30 s) for a medium resolution shell and $-80°$ (120 s) for the high-order reflections. For the low- and medium-resolution runs χ positions of 0, 30 and 55° were used, for the high-resolution runs 0 and 30° positions in χ were used. 1150 frames were collected in each of these runs, with a sample-to-detector distance of 5 cm. The scan width was 0.3° in all cases. No significant intensity decay was observed. *SORTAV* (Blessing, 1995) was used for the scaling of runs; absorption was disregarded. Further details of crystal data and measurement conditions are given in Table 1.[3] The molecular structure, which was already described in the literature (Young *et al.*, 1969), thermal motion at 20 K and the atomic numbering scheme are shown in Fig. 2(*a*).

2.1. Aspherical atom models

The data were interpreted with the Hansen and Coppens formalism (Hansen & Coppens, 1978), as implemented in *XD* (Koritsánszky *et al.*, 2003). Three different refinements were performed: a 'classical' multipole refinement (CMR) and two invariom refinements using *InvariomTool* (Hübschle *et al.*, 2007) for the transfer of multipole and κ parameters. As mentioned above, in one model only the nearest neighbors were used for the invariom description of the thymine heterocycle (hereafter called ITM model), while in the *meso*-ITM model the next-nearest neighbors were also considered. Fig. 2(*b*) shows the assignment of invariom names in thymidine and Table 2 lists the model compounds used for theoretically calculated multipole populations. Threshold values of the bond-distinguishing parameter χ were chosen to avoid non-integer bond orders. χ is used to distinguish between bond types based on bond distance (d) and Allred–Rochow electronegativity (EN) using the empirical relation (Schomaker & Stevenson, 1941)

$$\chi = [r_c(\text{atom 1}) + r_c(\text{atom 2}) - 0.08 \times |\Delta(\text{EN})|] - d. \quad (1)$$

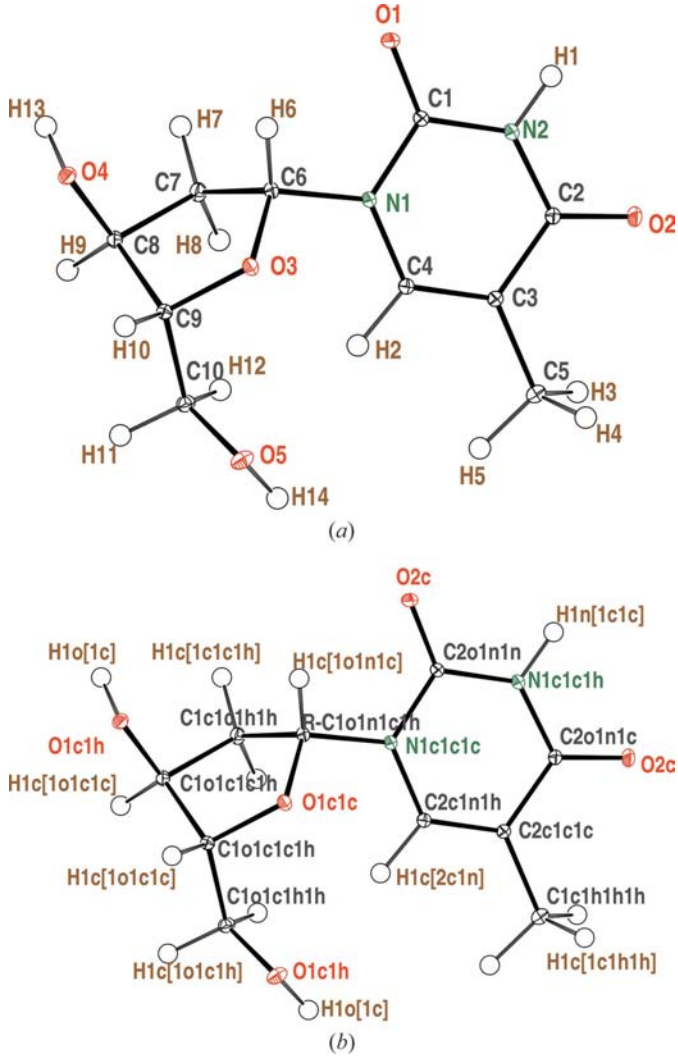

Figure 2
(*a*) *ORTEP* (Burnett & Johnson, 1996) representation of the molecular structure in the crystal at 20 K with atomic numbering scheme; displacement ellipsoids at 50% probability. (*b*) Assignment of invariom names for thymidine.

[3] Supplementary data for this paper are available from the IUCr electronic archives (Reference: SN5064). Services for accessing these data are described at the back of the journal.

Table 2
Invariom names and model compounds.

	Invariom name	Local atomic site symmetry	Model compound for ITM	Model compound for *meso*-ITM
O1	O2c	*m*	Formaldehyde	1-Methylthymine
O2	O2c	*m*	Formaldehyde	1-Methylthymine
O3	O1c1c	*mm*2	Dimethylether	Dimethylether
O4	O1c1h	*m*	Methanol	Methanol
O5	O1c1h	*m*	Methanol	Methanol
N1	N1c1c1c	*m*	Trimethylamine	1-Methylthymine
N2	N1c1c1h	*mm*2	Dimethylamine	1-Methylthymine
C1	C2o1n1n	*mm*2	Urea	1-Methylthymine
C2	C2o1n1c	*m*	Acetamide	1-Methylthymine
C3	C2c1c1c	*m*	Isobutene	1-Methylthymine
C4	C2c1n1h	*m*	Aminoethene	1-Methylthymine
C5	C1c1h1h1h	3	Ethane	1-Methylthymine
C6	R-C1o1n1c1h	1	(*R*)-1-Amino-ethanol	(*R*)-1-Amino-ethanol
C7	C1c1c1h1h	*mm*2	Propane	Propane
C8	C1o1c1c1h	*m*	2-Propanol	2-Propanol
C9	C1o1c1c1h	*m*	2-Propanol	2-Propanol
C10	C1o1c1h1h	*m*	Ethanol	Ethanol
H1	H1n[1c1c]	Linear	Dimethylamine	1-Methylthymine
H2	H1c[2c1n]	Linear	Aminoethene	1-Methylthymine
H3	H1c[1c1h1h]	Linear	Ethane	1-Methylthymine
H4	H1c[1c1h1h]	Linear	Ethane	1-Methylthymine
H5	H1c[1c1h1h]	Linear	Ethane	1-Methylthymine
H6	H1c[1o1n1c]	Linear	(*R*)-1-Amino-ethanol	(*R*)-1-Amino-ethanol
H7	H1c[1c1c1h]	Linear	Propane	Propane
H8	H1c[1c1c1h]	Linear	Propane	Propane
H9	H1c[1o1c1c]	Linear	2-Propanol	2-Propanol
H10	H1c[1o1c1c]	Linear	2-Propanol	2-Propanol
H11	H1c[1o1c1h]	Linear	Ethanol	Ethanol
H12	H1c[1o1c1h]	Linear	Ethanol	Ethanol
H13	H1o[1c]	Linear	Methanol	Methanol
H14	H1o[1c]	Linear	Methanol	Methanol

The values used were 0.11 Å between single and mesomeric bonds, 0.15 Å between mesomeric bonds and double bonds, and 0.27 Å between double and triple bonds. Usually the values employed are 0.0847, 0.184 and 0.27, respectively. This leads to a complete neglect of the second next-neighbor sphere. Hence, for the ITM all model compounds follow the nearest-neighbor approximation. Multipole parameters of the N atoms were obtained from a single-point energy calculation on planar conformations of di- and trimethylamine using the bond lengths of a geometry optimization with the program *GAUSSIAN*98 (Frisch *et al.*, 1998) basis B3LYP/D95++(3df,3pd). This basis set was also used in the calculation of all other model compounds. A theoretical calculation of 1-methylthymine was also carried out to compare the electrostatic potential (see below) with the same basis set (THEO). This calculation was also used to generate multipole parameters for the entire heterocycle which were transferred to the ring system of thymidine (*meso*-ITM). Table 2 lists the local site symmetry used for the 'classical' multipole refinement. Additional to this symmetry, atoms which share the same invariom name were chemically constrained to each other, the only exception being the O atoms O1 and O2, which were refined independently. The only parameters refined in both invariom-transfer models (ITM and *meso*-ITM) were the x, y, z and U^{ij} parameters. κ parameters were refined for the non H-atoms in CMR and kept at 1.2 for H. They were kept constant in the two invariom models according to their values in the database. κ' parameters were kept constant at 1.2 for H atoms and at 1.0 for the other atoms. The residual density shown in Fig. 3 is almost flat and featureless in all three refinement models. In the invariom models some density is left

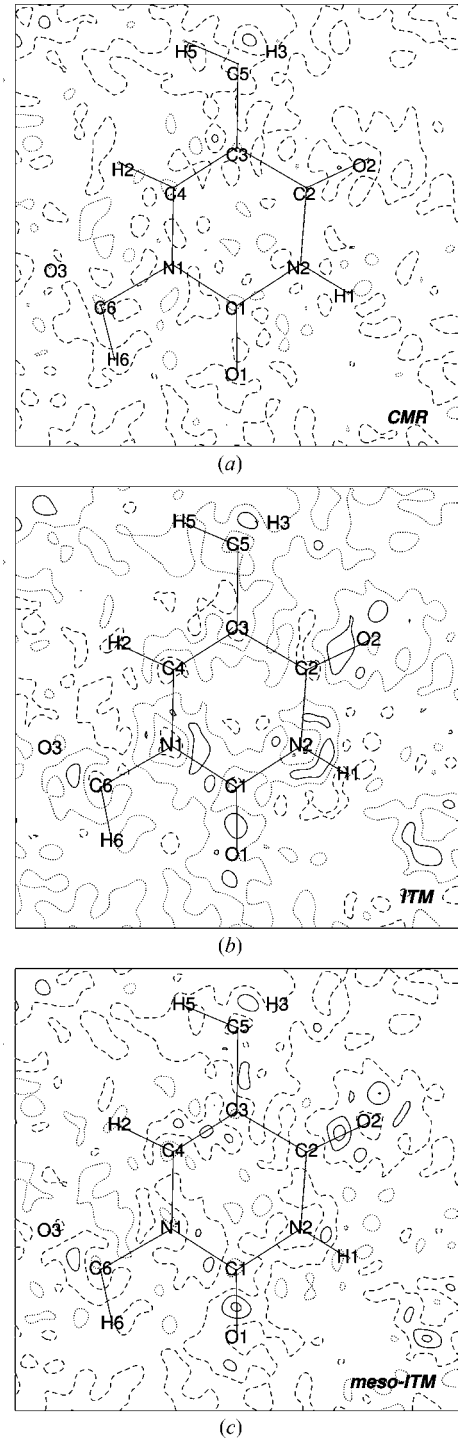

Figure 3
Residual density maps of the thymine ring. (*a*) classical multipolar refinement (CMR); (*b*) invariom-transfer model (ITM); (*c*) 'mesomeric' invariom-transfer model (*meso*-ITM). Contours at 0.1 e Å^{-3}; full lines: positive; dashed lines: zero; dotted lines: negative.

Table 3
Figures-of-merit for the three aspherical models.

	CMR	ITM	*meso*-ITM
Parameters	382	168	168
No. of reflections/No. of parameters	14.8	33.6	33.6
$R(F)$	0.016	0.019	0.018
$R_{all}(F)$	0.019	0.022	0.021
$wR(F)$	0.019	0.023	0.021
GoF	0.94	1.15	1.04
Min r.d. (e Å^{-3})	−0.17	−0.22	−0.18
Max r.d. (e Å^{-3})	0.16	0.29	0.24

unfitted in the double bonds O1═C1 and O2═C2. Figures-of-merit of all three refinements are given in Table 3.

2.2. Constrained wavefunction fitting (CWF)

The constrained wavefunction fitting approach by Jayatilaka (Jayatilaka & Grimwood, 2001; Grimwood & Jayatilaka, 2001) allows the ELF to be obtained from experiment. In this approach a wavefunction is fitted to experimental X-ray data. The Hartree–Fock energy equations $\mathbf{E}(\mathbf{c}, \epsilon)$ are augmented with a term which describes the influence of the experimental data

$$L(\mathbf{c}, \epsilon, \lambda) = \mathbf{E}(\mathbf{c}, \epsilon) - \lambda[\chi^2(\mathbf{c}) - \Delta], \qquad (2)$$

where $\mathbf{c}$ are the molecular-orbital coefficients expressed in a basis set and ϵ are the Lagrange multipliers associated with the orbital orthogonality constraints. Δ is the desired experimental agreement in χ^2 which is defined as

$$\chi^2 = \frac{1}{N_r - N_p} \sum_{\mathbf{h}}^{N_r} \frac{[F_{calc}(\mathbf{h}) - \mathbf{F}_{obs}(\mathbf{h})]^2}{\sigma^2(\mathbf{h})}, \qquad (3)$$

where N_r is the number of reflection and N_p the number of adjustable fitting parameters.

Herein, a constrained wavefunction (Jayatilaka & Grimwood, 2004) was fitted to the experimental data of thymidine. A λ value of 0.306 was used to reach a χ^2 value of 0.523 and a $wR(F)$ value of 0.022.

3. Results

3.1. Electrostatic potential

The electrostatic potential was calculated using the *XDPROP* subprogram of *XD* (Koritsánszky *et al.*, 2003) from the experimental ED using the method of Su & Coppens (1992). Figs. 4 and 5 show the electrostatic potentials mapped on an electron density isosurface of 0.001 a.u., which corresponds to 0.0067 e Å^{-3}. Figures were generated with the program *MOLISO* (Hübschle & Luger, 2006). Fig. 5 represents the sheer plan of the isosurfaces, which is the normally invisible concave inner surface behind the molecule relative to the viewers eye. The convex front face of the surfaces is visualized in Fig. 4.

Although the electrostatic potential at this level in $\rho(\mathbf{r})$ is very sensitive to small changes, the invariom-transfer models (ITM and *meso*-ITM) lead to an almost identical potential around the sugar region (see Figs. 4 and 5; second and third row), which is not quite true in the thymine part. Since the

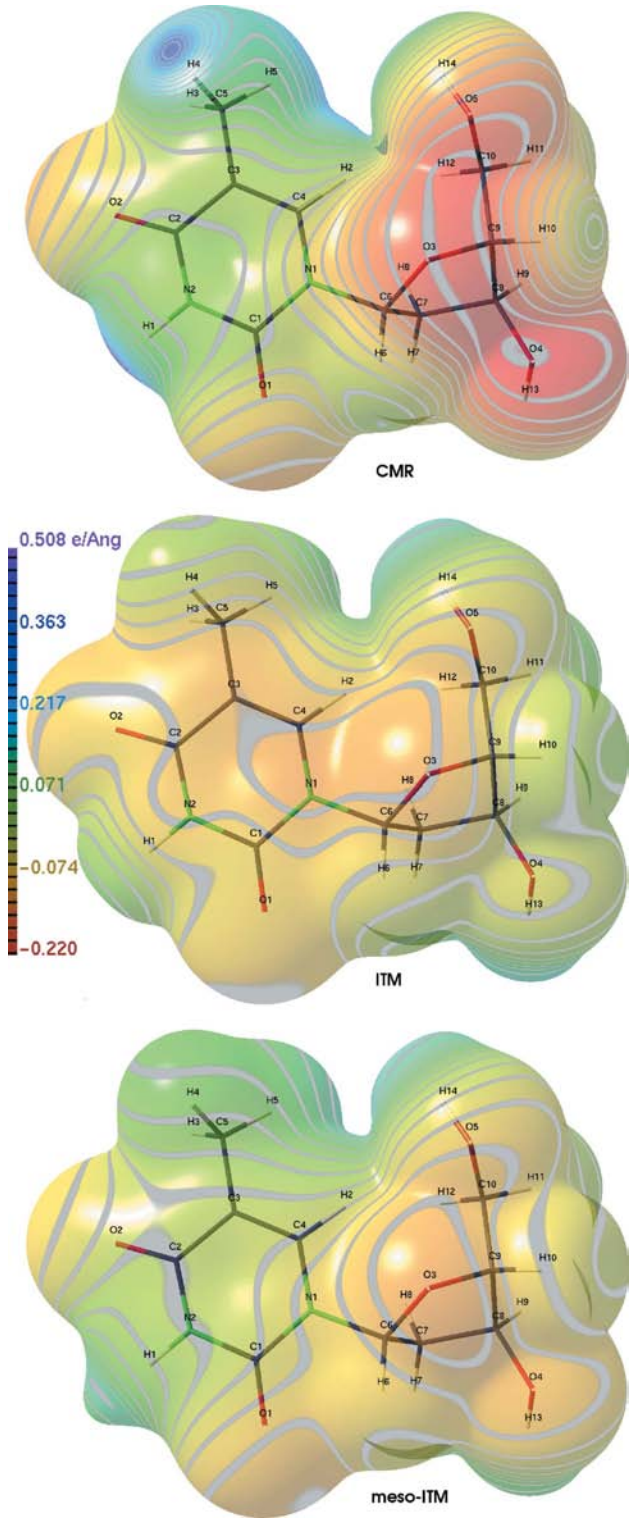

Figure 4
The electrostatic potential of thymidine mapped on an isosurface at $\rho(\mathbf{r})$ = 0.0067 e A^{-3}. Front face: classical multipole refinement (CMR) first row, invariom-transfer model (ITM) second row, and the 'mesomeric' invariom-transfer model (*meso*-ITM) third row.

invariums are fitted to theoretical molecular calculations the most optimal case would be when the invariums reproduce exactly the theoretical calculation of the target molecule.

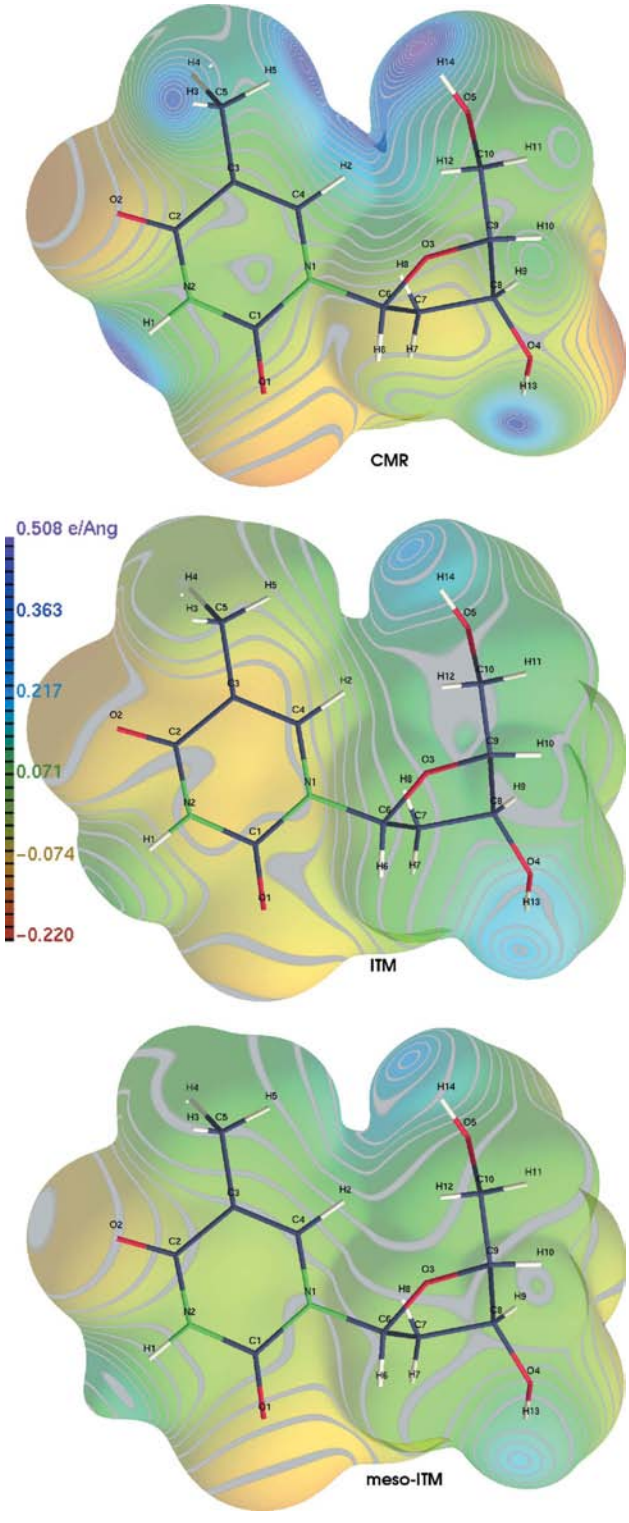

Figure 5
The electrostatic potential of thymidine mapped on an isosurface at $\rho(\mathbf{r})$ = 0.0067 e A^{-3}. Sheer plan: classical multipole refinement (CMR) first row, invariom-transfer model (ITM) second row, and the 'mesomeric' invariom-transfer model (*meso*-ITM) third row.

Differences between the electrostatic potentials therefore reflect:

(i) shortcomings of the multipole model in reproducing a theoretical density,

(ii) the inaccuracy of approximating a molecular electron density by fragments and

(3) the effects of hydrogen bonding and the crystal field.

Agreement of the electrostatic potential around the thymine region between the CMR model (Figs. 4 and 5; first row) and the invariom-transfer models is closer to the 'mesomeric'-invariom-transfer model (Figs. 4 and 5; third row) than to the ITM. The positive potential around H3, H4 and H5, which is typical for hydrogen bonds, is not seen in the ITM and *meso*-ITM. Generally it is found that the potential difference between the maximum and minimum potential on that surface is more pronounced in the classical multipole refinement model compared with the invariom models. We consider this as a result of the crystal field, which is known to stabilize charge separations.

In order to quantify these findings, Table 4 shows the surface quantities of the three refinement models and a model from a theoretical calculation according to Politzer *et al.* (2001). The positive, negative and overall average potential values $\bar{V}_S^+$, $\bar{V}_S^-$ and $\bar{V}_S$ on the surface are calculated as given in (1)–(3), where n and m are the number of grid points with positive and negative potential, respectively

$$\bar{V}_S^+ = \frac{1}{n}\sum_{i=1}^{n} V_S^+(\mathbf{r}_i), \tag{4}$$

$$\bar{V}_S^- = \frac{1}{m}\sum_{i=1}^{m} V_S^-(\mathbf{r}_i), \tag{5}$$

$$\bar{V}_S = \frac{1}{m+n}\left[\sum_{i=1}^{n} V_S^+(\mathbf{r}_i) + \sum_{i=1}^{m} V_S^-(\mathbf{r}_i)\right]. \tag{6}$$

The average deviation from the overall potential on the surface is Π

$$\Pi = \frac{1}{m+n}\sum_{i=1}^{m+n}\left|V_S(\mathbf{r}_i) - \bar{V}_S\right|. \tag{7}$$

The positive, negative and total variances of the surface potential are

$$\sigma_{\rm tot}^2 = \sigma_+^2 + \sigma_-^2 = \frac{1}{n}\sum_{i=1}^{n}[V_S^+(\mathbf{r}_i) - \bar{V}_S^+]^2 + \frac{1}{m}\sum_{i=1}^{m}[V_S^-(\mathbf{r}_i) - \bar{V}_S^-]^2. \tag{8}$$

In nearly all of the calculated quantities (Table 4) a trend can clearly be observed that the classical refinement leads to the more differentiated values than the invariom models, which we ascribe to hydrogen bonding. Even the invariom models have more extreme surface quantities than the model from the theoretical calculation: All three XD models have significantly greater $\bar{V}_S^+$ than $|\bar{V}_S^-|$ values and $V_{S,\rm max}$ is more than two times

Table 4
Computed surface quantities.

Units: $\bar{V}_S^+$, $\bar{V}_S^-$, Π, $\bar{V}_{S,\min}$ and $\bar{V}_{S,\max}$ are in e Å^{-1}; σ_+^2, σ_-^2 and σ_{tot}^2 are in (e Å^{-1})^2.

Model	$\bar{V}_S^+$	$\bar{V}_S^-$	Π	σ_+^2	σ_-^2	σ_{tot}^2	$\bar{V}_{S,\min}$	$\bar{V}_{S,\max}$
CMR	10.72×10^{-2}	-8.18×10^{-2}	9.49×10^{-2}	18.5×10^{-3}	11.7×10^{-3}	30.2×10^{-3}	−0.220	0.509
ITM	7.02×10^{-2}	-5.52×10^{-2}	6.22×10^{-2}	7.03×10^{-3}	4.18×10^{-3}	11.2×10^{-3}	−0.119	0.241
meso-ITM	6.12×10^{-2}	-5.25×10^{-2}	5.70×10^{-2}	4.90×10^{-3}	4.32×10^{-3}	9.22×10^{-3}	−0.120	0.226
THEO†	4.48×10^{-2}	-5.17×10^{-2}	4.68×10^{-2}	2.78×10^{-3}	4.58×10^{-3}	7.37×10^{-3}	−0.123	0.160
CWF	5.84×10^{-2}	-6.49×10^{-2}	5.96×10^{-2}	4.05×10^{-3}	7.16×10^{-3}	11.2×10^{-3}	−0.139	0.191

† B3LYP/6-31G(d,p) geometry optimization (Frisch *et al.*, 1998).

greater than $|V_{S,\min}|$. The theoretical calculation leads to a smaller $\bar{V}_S^+$ than $|\bar{V}_S^-|$ value. This can be ascribed to shortcomings in the multipole model and to the differences in the optimized geometry of the theoretical calculation. For the constrained wavefunction fitting (CWF), which is discussed later, $\bar{V}_S^+$ is smaller than in the *XD* models but higher than for the theoretical optimization. The corresponding $|\bar{V}_S^-|$ value is the closest compared with the CMR model.

3.2. Hirshfeld surface

The Hirshfeld surface (Spackman & Byrom, 1997; McKinnon *et al.*, 1998) is a computationally simple and elegant way of partitioning molecules in a crystal. It is defined by the weighting function

$$W = \frac{\rho(\mathbf{r})_{\text{iam, molecule}}}{\rho(\mathbf{r})_{\text{iam, crystal}}} = 0.5. \quad (9)$$

We have calculated Hirshfeld surfaces according to this formula using the IAM. However, we used them to map the non-spherical multipole model electron density on them, which allows the amount of density in the region of the hydrogen bonds to be visualized (Chęcińska *et al.*, 2006). It should be emphasized that bond-critical points (b.c.p.s) of a hydrogen bond do not necessarily lie on the Hirshfeld surfaces, since different electron densities were used. Since the b.c.p. is a minimum of the ED along the bond path and the ED which is mapped here on the Hirshfeld surface (HS) has its maximum on the bond path, the ED mapped on the HS is always higher than on the b.c.p. Fig. 6 shows the multipole ED of the crystal mapped on the Hirshfeld surface making sites and strengths of intermolecular interactions visible. The color gradient is transparent at regions with very low ED. This enables the viewer to recognize regions with higher electron density even if they are occluded by the surface itself. For the majority of the intermolecular contacts the ED on the Hirshfeld surface in the CMR (first row) is somewhat smaller than on surfaces of the invariom-transfer models. This can be considered as a result of the stronger electrostatic component of these contacts. Except for this small difference all three visualizations share qualitatively the same features.

3.3. Bond-topological analysis

The bond-topological analysis of the experimental ED was carried out with the *XD* subprogram *XDPROP* (Koritsánszky *et al.*, 2003). A critical point in $\rho(\mathbf{r})$ is found when the gradient $\nabla\rho(\mathbf{r})$ vanishes. All the expected b.c.p.s could be located. Fig. 7 shows $\rho(\mathbf{r}_{bcp})$ and $\nabla^2\rho(\mathbf{r}_{bcp})$ values of all the covalent bonds in thymidine. An almost perfect agreement between the two invariom models can be seen. The average difference between the two invariom models is 0.023 e Å^{-3} for $\rho(\mathbf{r})$ and 0.7 e Å^{-5} for $\nabla^2\rho(\mathbf{r})$. Compared with the average differences found in an experimental ED with two independent molecules (Chęcińska *et al.*, 2006), where the differences were 0.07 e Å^{-3} for $\rho(\mathbf{r})$ and 4.9 e Å^{-5} for $\nabla^2\rho(\mathbf{r})$, and compared with analyses of two datasets of the same compound with a similar range (0.1 e Å^{-3}/4 e Å^{-5}; Dittrich *et al.*, 2002) the differences found here are quite small. Between the classical refinement and the ITM the average differences are 0.09 e Å^{-3} for $\rho(\mathbf{r})$ and 3.7 e Å^{-5} for $\nabla^2\rho(\mathbf{r})$. The corresponding quantities between CMR and *meso*-ITM are 0.10 e Å^{-3} for $\rho(\mathbf{r})$ and 3.8 e Å^{-5}. These differences are still in the range of the accuracy of an experimental ED approach. One can clearly see that the C═O double bonds O1═C1 and O2═C2 have a higher electron density and a more negative Laplacian at the b.c.p. than the other C—O bonds. The ED fits perfectly in the list of observations compiled for carbonyl bonds for other experiments. In the C—N bonds the single bond C6—N1 has a lower density than the C—N bonds in the hetero-aromatic ring. The latter ones have a slightly lower ED at the b.c.p. than C—N peptide bonds (Chęcińska *et al.*, 2006). The C3—C4 bond is stronger than all other C—C bonds, as indicated by its $\rho(\mathbf{r})_{bcp}$ value of 2.30 e Å^{-3} compared with the average $\rho(\mathbf{r})_{bcp}$ of all other C—C bonds being 1.82 e Å^{-3}. From Bader's exponential formula for the bond order n of a C—C bond $\{n = \exp[A(\rho(\mathbf{r})_{bcp}) - B)]\}$,[4] the bond order for C3—C4 is calculated to $n = 2.06$, while for the other C—C bonds $n = 1.20$ is obtained.

The bond strengths found here can be well explained with the resonance formulae of the thymine ring (Fig. 1). There is only one resonance formula whith single charge separations where the C3—C4 bond is not a double bond. Between the other C atoms and the N atoms the double bonds are more

[4] *A* and *B* from a theoretical calculation on model compounds (Luger *et al.*, 2004).

delocalized. These findings will be discussed in more detail with the help of ELF (see below).

3.4. Hydrogen bonds

Thymidine takes part in three classical hydrogen bonds and some weaker C—H···*X* contacts in the crystal structure. Table 5 shows the results of the topological analysis of the hydrogen bonds and other contacts. B.c.p.s could be found for all of them except C10—H12···O3, whereas no critical point was found in the CMR. The average differences for $\rho(\mathbf{r})$ are 0.003 e Å^{-3} between the invariom models, 0.07 e Å^{-3} between CMR and the invariom models. The average differences for $\nabla^2\rho(\mathbf{r})$ are 0.06 e Å^{-5} between ITM and *meso*-ITM, 0.29 e Å^{-5} between CMR and ITM, and 0.32 e Å^{-5} between CMR and *meso*-ITM. It is evident that for most of the contacts the CMR has the lowest density values. Exceptions are the C—H···O contacts donated by the deoxyribose where the density in all three models is more or less the same. The contacts seem to have a more electrostatic component in the CMR, which can be seen at the atomic charges discussed in the next section. H atoms involved in hydrogen contacts to O, N or C atoms of the thymine subsystem have a more pronounced charge depletion in the CMR while donor atoms have increased charges.

3.5. Atomic properties

A well defined partitioning procedure is provided by Bader's QTAIM theory, making use of the zero-flux surfaces in the gradient vector field $\nabla\rho(\mathbf{r})$ which separate atomic basins from their neighbors. Integration of these basins, which was carried out with the program *TOPXD*, gave atomic volumes and electron populations of the defined atoms. These quantities are displayed in Figs. 8 and 9 (the volumes V_{001} given in Fig. 8 are defined by a cutoff at ρ = 0.001 a.u.).

As for bonding properties and likewise for the atomic properties no important differences between the two considered invariom models exist, being on average 0.03 e and 0.09 Å^3 for population and volume. The average differences between CMR and the invariom models are somewhat higher, being 0.13 e for the populations, while the volume differences are 0.63 Å^3. In all cases these differences are smaller or in the range commonly found in other studies (Chęcińska *et al.*, 2006).

Nevertheless, if there are differences seen in atomic properties between CMR and the invariom models they occur mainly at atoms involved in intermolecular interactions. The volumes of the three H atoms involved in classical hydrogen bonds (H1, H13 and H14; see Table 5) are quite small in the CMR (V_{001} = 1.46–2.43 Å^3) compared with the invariom models (V_{001} = 2.18–3.07 Å^3). Consequently, the CMR volumes of the donor O and N atoms are larger than in the invariom cases. Also an increase of charge around these donor atoms and a decrease around the corresponding H atoms can be observed. These findings for the CMR properties are in agreement with the eight criteria developed by Koch & Popelier (1995) on the basis of hydrogen-bond topological and atomic properties.

A similar trend as discussed above but less pronounced is observed for the atoms involved in weaker C—H···*X* contacts listed in Table 5. Moreover, exceptions exist. For example, H3, H4 and H5 have a smaller CMR volume (V_{001} = 5.07–5.36 Å^3) compared with the invariom models (V_{001} = 6.45–6.84 Å^3),

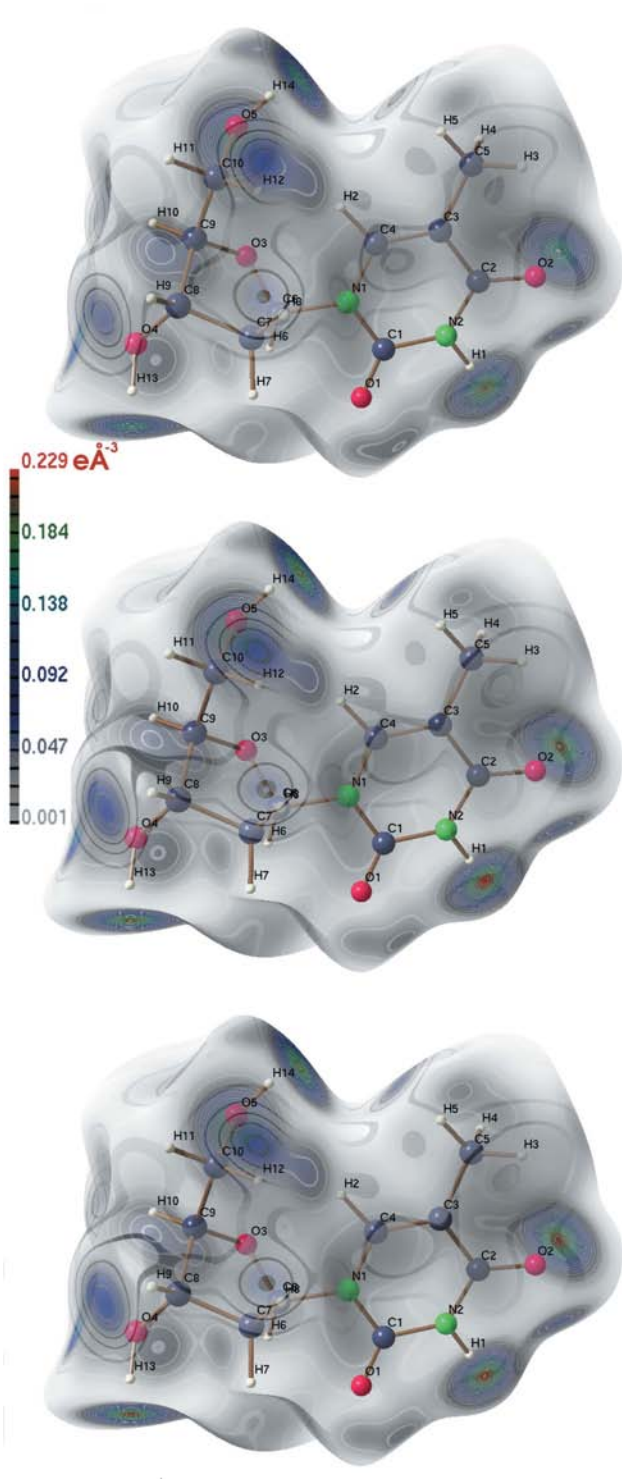

Figure 6
The ED mapped the Hirshfeld surface: classical multipole refinement (CMR) first row, invariom-transfer model (ITM) second row, and the 'mesomeric' invariom-transfer model (*meso*-ITM) third row.

Table 5
Hydrogen-acceptor b.c.p.s.

Model	D—H···A	ρ	$\nabla^2\rho$	H···A (Å)	A—CP (Å)	CP—H (Å)	D—H···A (°)
CMR	O4—H13···O2^{i}	0.13	3.9	1.787	1.183	0.604	176.8
ITM	O4—H13···O2^{i}	0.19	3.3	1.789	1.161	0.629	174.2
meso-ITM	O4—H13···O2^{i}	0.20	3.3	1.789	1.158	0.630	174.9
CMR	N2—H1···O4^{ii}	0.14	4.4	1.835	1.206	0.629	178.1
ITM	N2—H1···O4^{ii}	0.21	3.2	1.834	1.177	0.656	178.3
meso-ITM	N2—H1···O4^{ii}	0.20	2.8	1.834	1.193	0.641	179.0
CMR	O5—H14···O5^{iii}	0.13	3.7	1.840	1.223	0.617	170.3
ITM	O5—H14···O5^{iii}	0.18	3.1	1.838	1.193	0.645	169.3
meso-ITM	O5—H14···O5^{iii}	0.18	3.1	1.840	1.191	0.650	168.7
CMR	C5—H5···O5^{iii}	0.02	0.6	2.548	1.544	1.005	159.7
ITM	C5—H5···O5^{iii}	0.05	0.7	2.521	1.489	1.033	161.4
meso-ITM	C5—H5···O5^{iii}	0.05	0.7	2.526	1.491	1.035	160.3
CMR	–	–	–	–	–	–	–
ITM	C10—H12···O3^{iv}	0.05	0.8	2.529	1.041	1.488	156.2
meso-ITM	C10—H12···O3^{iv}	0.05	0.8	2.535	1.052	1.483	155.2
CMR	C8—H9···O2^{v}	0.05	0.9	2.463	1.457	1.006	148.3
ITM	C8—H9···O2^{v}	0.07	0.8	2.453	1.420	1.038	147.4
meso-ITM	C8—H9···O2^{v}	0.07	0.8	2.453	1.417	1.036	147.6
CMR	C7—H8···O4^{iv}	0.04	0.7	2.630	1.135	1.496	131.6
ITM	C7—H8···O4^{iv}	0.05	0.7	2.580	1.080	1.494	135.4
meso-ITM	C7—H8···O4^{iv}	0.05	0.7	2.591	1.093	1.498	134.3
CMR	C5—H4···O1^{vi}	0.02	0.5	2.634	1.543	1.091	143.8
ITM	C5—H4···O1^{vi}	0.04	0.6	2.636	1.522	1.118	141.0
meso-ITM	C5—H4···O1^{vi}	0.04	0.6	2.626	1.518	1.107	142.2
CMR	C9—H10···O1^{vii}	0.04	0.7	2.719	1.483	1.236	113.6
ITM	C9—H10···O1^{vii}	0.04	0.6	2.675	1.514	1.161	112.9
meso-ITM	C9—H10···O1^{vii}	0.04	0.6	2.672	1.511	1.161	113.1
CMR	C4—H2···C10	0.04	0.6	2.750	1.743	1.007	148.5
ITM	C4—H2···C10	0.06	0.8	2.693	1.654	1.040	148.7
meso-ITM	C4—H2···C10	0.06	0.8	2.716	1.690	1.027	147.5

Symmetry codes: (i) $-x,\frac{1}{2}+y,\frac{3}{2}-z$; (ii) $-\frac{1}{2}-x,1-y,\frac{1}{2}+z$; (iii) $\frac{1}{2}+x,\frac{1}{2}-y,1-z$; (iv) $1+x,y,z$; (v) $\frac{1}{2}-x,1-y,-\frac{1}{2}+z$; (vi) $-x,-\frac{1}{2}+y,\frac{3}{2}-z$; (vii) $-\frac{1}{2}-x,1-y,-\frac{1}{2}+z$. Units: distances in Å, angles in °, ρ in e Å^{-3}, $\nabla^2\rho$ in e Å^{-5}.

while the parent C atoms have an increased CMR volume. Two of them (H4 and H5) are involved in C—H···X contacts, while H3 is involved in a C—H···π contact to the thymine ring. Also a C—H···π contact is present for C6—H6···thymine, which explains the decreased CMR volume of (V_{001} = 4.78 Å^3) compared with the volume of (V_{001} = 6.40–6.46 Å^3) in the invariom models and an increase of the CMR volume of C6 by 1.4 Å^3. On the other hand, no volume differences could be found for C8 and C9 and the H atoms bonded to these atoms.

3.6. ELF from constrained wavefunction

The electron localization function (ELF; Becke & Edgecombe, 1990) provides useful information about the nature of chemical bonds. The ELF separates the electron density into spatial regions which correspond to the chemical image of electron pairs. It is related to the first term of a Taylor expansion D_σ of the spherically averaged conditional-pair probability distribution for electrons of the same spin around a reference point. If the reference electron is highly localized then the probability of finding a like-spin electron nearby is small. Necessarily, the same holds for electron pairs. Due to its definition the ELF reaches its maximum value of one where electron pairs are perfectly localized. Unfortunately, ELF cannot be calculated directly from the experimental electron density satisfactorily. With the CWF approach ELF can be determined from the experiment.

Fig. 10 shows the experimental ELF of thymidine. The isosurfaces are at a value of 0.856 in ELF and are color coded by the number of the 70 basins sorted by their volume. Additionally, a cut-plane is visualized in the thymine ring, color mapped by the ELF. The disynaptic valence basin C3—C4 differs in volume (represented by the color of the isosurface) and shape from all the other disynaptic valence basins and thus behaves like a real double bond. As expressed before by Grimwood & Jayatilaka (2001), zero-synaptic valence basins can be observed in the spatial regions where hydrogen bonding in the crystal takes place (see Fig. 10, bottom-left corner). Zero-synaptic valence basins are discussed as artefacts in Jayatilaka & Grimwood (2004), but we think that they might be useful as indicators for hydrogen bonds in the crystal.

This seems to be a special feature of the experimental ELF and shows that crystal effects are accounted for. Therefore, analysis of the populations of the ELF basins and the comparison with the topological analysis of the electron density should be of interest.

The experimental ELF was topologically analyzed and the electron density integrated with the program *BASIN*4.2 from Kohout (2007). The results are listed in Table 6. The electron population in the disynaptic valence basins can be connected with the bond order. The basins O1—C1 and O2—C2 are populated with 2.2 e, while other O—C basins are populated with only 1.3 e. These two bonds are partial double bonds in a 'mesomeric' system, while the others are C—O single bonds. The populations of the C—N basins are in agreement with the expected order of strengths from the resonance formulae (Fig. 1). The C—N basins in the ring are populated with 2.0–2.2 e, which are more than single bonds according to the resonance formulae, while C6—N1 is populated with 1.8 e, which refers to a single bond. The disynaptic valence basin C3—C4 with 3.5 e is the highest populated basin in the molecule which

characterizes it as a double bond, while the C2—C3 basin counts 2.3 e, which is a little more than for the other single C—C bonds in the molecule, where the basins are occupied by only 1.9–2.0 e. This is in agreement with Fig. 10, the resonance formulae (Fig. 1) and the CMR-topological analysis of $\rho(\mathbf{r})_{\rm bcp}$ (C3—C4 2.30 e Å^{-3} and C2—C3 1.90 e Å^{-3}).

Monosynaptic valence basins referring to O-atom lone pairs of the thymine ring are more occupied (2.7 e) than those of the sugar (2.2–2.4 e). The lone pairs of the N atoms are split into two basins below and above the ring. Summed over the two basins, the lone pair of N1 counts 1.7 e and N2 1.5 e.

The electron count of protonated monosynaptic valence basins results in 1.8 e for H13 and H14, which are involved in O—H···O hydrogen bonds. However, H1 involved in a N—H···O hydrogen bond is occupied by 2.1 e like most of the other H atoms (2.0–2.1 e). This behavior is not yet fully understood, but it must be connected with the higher electronegativity of the O atoms with respect to the N atoms.

The six zero-synaptic valence basins found were unoccupied and are nearest to the H atoms.

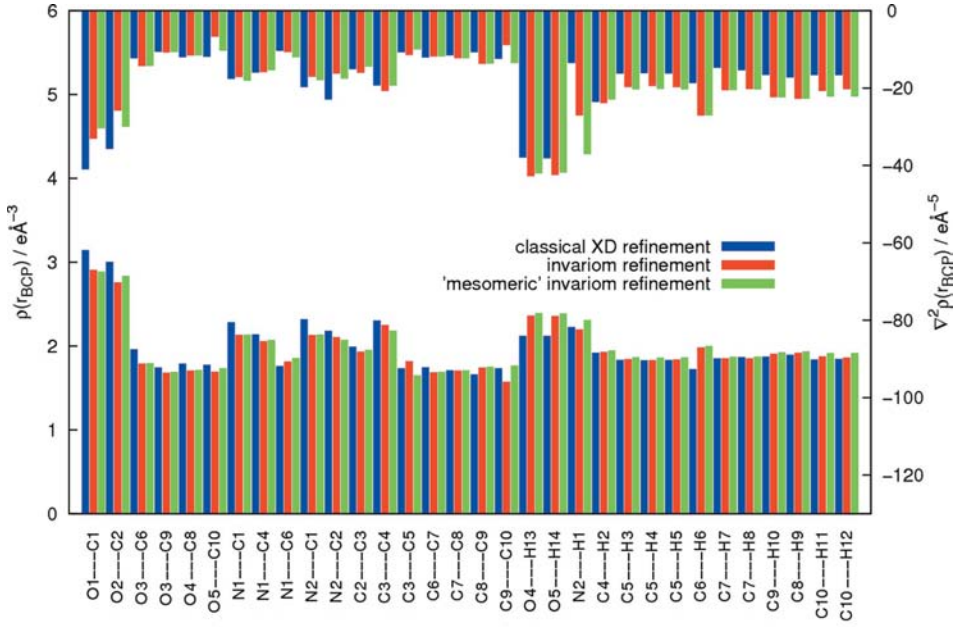

Figure 7
Electron density $\rho(\mathbf{r})$ e Å^{-3} and $\nabla^2\rho(\mathbf{r})$ e Å^{-5} at the b.c.p.s of thymidine.

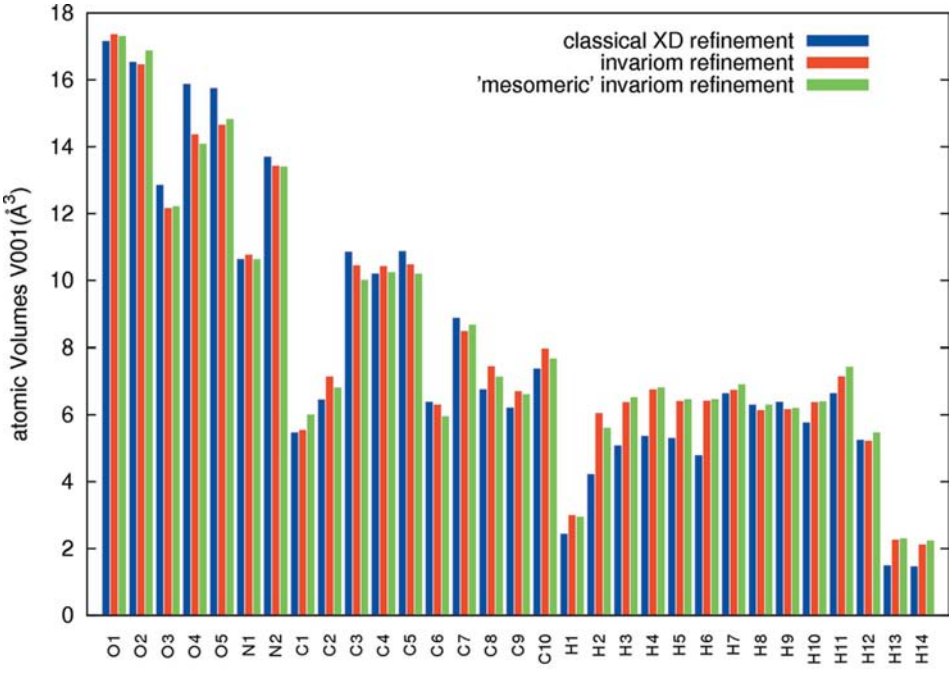

Figure 8
Volumes of the atomic basins in thymidine in Å^3.

4. Conclusions

For a high-resolution 20 K X-ray dataset of thymidine a classical multipole refinement was compared with two invariom-transfer models. Based on topological properties a good agreement between the classical and the invariom approach was found. Differences can be attributed to the crystal field, as they are small and occur mainly for atoms involved in hydrogen bonding. The invariom models, one with the neglect of the second next-neighbor sphere and the other one with the complete thymine ring system as the model compound, differ only slightly in terms of b.c.p.s. However, the electrostatic potentials from the classical multipole refinement differ, again mostly in the thymine ring in regions of hydrogen bonding. The electrostatic potential mapped on the molecular surface at the $\rho(\mathbf{r})$ = 0.0067 e Å^{-3} level has been shown to be a good indicator for effects of the crystal environment, such as hydrogen bonds. Classical hydrogen bonds in the molecule fulfill the criteria introduced by Koch & Popelier (1995).

Although in the case of the hetero-aromatic thymine ring system the next-neighbor approximation seems already sufficient to model the electron density of the C atoms, N atoms with three single

Table 6
Integrated ELF basins.

$N_{electrons}$	ELF_{max}	Atom	$N_{electrons}$	ELF_{max}	Bond
Core basins			Disynaptic valence basins		
2.123	1.0000	O1	2.256	0.8729	C1—O1
2.133	1.0000	O2	2.232	0.8750	C2—O2
2.150	1.0000	O3	1.388	0.8951	C6—O3
2.132	1.0000	O4	1.296	0.8905	O3—C9
2.128	1.0000	O5	1.267	0.8938	C8—O4
2.162	1.0000	N1	1.275	0.8953	C10—O5
2.112	1.0000	N2	2.156	0.9358	C1—N1
2.092	1.0000	C1	1.999	0.9347	N1—C4
2.083	1.0000	C2	1.806	0.9436	N1—C6
2.091	1.0000	C3	2.142	0.9386	C1—N2
2.088	1.0000	C4	2.122	0.9353	N2—C2
2.087	1.0000	C5	2.327	0.9692	C2—C3
2.092	1.0000	C6	3.490	0.9384	C4—C3
2.090	1.0000	C7	1.981	0.9664	C3—C5
2.101	1.0000	C8	1.951	0.9705	C6—C7
2.088	1.0000	C9	1.902	0.9669	C8—C7
2.087	1.0000	C10	1.988	0.9725	C9—C8
			1.976	0.9734	C10—C9

$N_{electrons}$	ELF_{max}	Atom	$N_{electrons}$	ELF_{max}	Atom
Monosynaptic valence basins			Protonated monosynaptic valence basins		
2.767	0.9164	O1	2.106	0.9994	H1
2.758	0.9186	O1	2.163	0.9999	H2
2.769	0.9215	O2	1.981	0.9999	H3
2.716	0.9224	O2	1.993	0.9999	H4
2.394	0.9233	O3	2.000	0.9999	H5
2.475	0.9267	O3	2.100	0.9999	H6
2.362	0.9264	O4	2.046	0.9999	H7
2.282	0.9248	O4	2.028	0.9999	H8
2.214	0.9221	O5	2.114	0.9999	H9
2.430	0.9253	O5	2.072	0.9999	H10
0.867	0.8636	N1	2.062	0.9999	H11
0.881	0.8680	N1	2.068	0.9999	H12
0.713	0.8580	N2	1.781	0.9977	H13
0.777	0.8652	N2	1.750	0.9979	H14

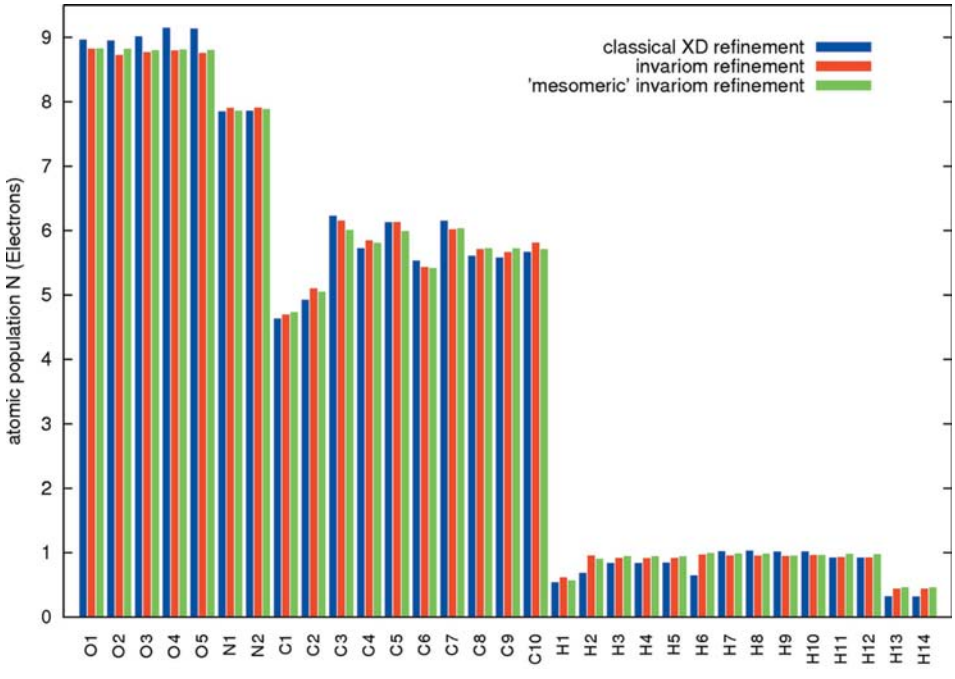

Figure 9
Populations of the atomic basins in thymidine in e.

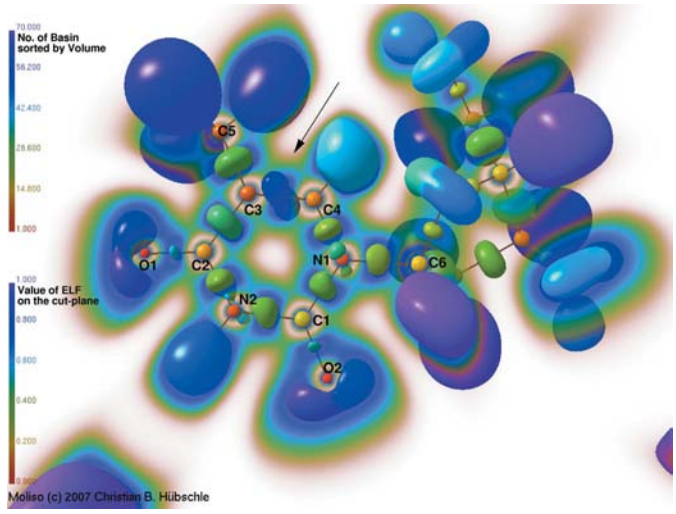

Figure 10
The ELF from constrained wavefunction fitting. Isosurfaces are at a value of 0.856 in ELF and are color coded by the number of the 70 ELF basins sorted by their volumes. The cut plane lies in the thymine ring and is mapped by the ELF. The arrow indicates the disynaptic valence basin C3—C4.

bonds currently cause problems in automatic invariom assignment. We are working on a solution to this problem.

Concerning the question of the extent of the delocalization in thymidine, most of the possible resonance formulae have a positive charge at a hetero atom (Fig. 1). The resonance formula with a double bond at C3—C4, C1—O1 and C2—O2 is the only neutral form, which makes it the predominant resonant state. This chemical insight was supported by a topological analysis of the experimental electron density and an ELF analysis. The values obtained for the $\rho(\mathbf{r})_{bcp}$ and the population of the ELF basins agree very well in their description of the bonding situation. The bonds C3—C4 [$\rho(\mathbf{r})_{bcp}$ = 2.30 e Å^{-3}, N(ELF) = 3.5 e], C1—O1 [$\rho(\mathbf{r})_{bcp}$ = 3.14 e Å^{-3}, N(ELF) = 2.3 e] and C2—O2 [$\rho(\mathbf{r})_{bcp}$ = 3.00 e Å^{-3}, N(ELF) = 2.2 e] were unambiguously identified as double bonds. Therefore, the thymine ring should be considered as only a partially delocalized system. The analysis of the experimentally derived ELF proved to be a very helpful tool for revealing the bonding situation.

This investigation was supported by grants of the Deutsche Forschungsgemeinschaft (DFG), Grants Lu222/

29-2 and DI 921/3-1, within or associated to the priority program *SPP*1178, and by the Graduate School 788 (hydrogen bonding and hydrogen transfer).

References

Bader, R. F. W. (1990). *Atoms in Molecules: A Quantum Theory*, 1st ed. Oxford: Clarendon Press.

Becke, A. D. & Edgecombe, K. E. (1990). *J. Chem. Phys.* **92**, 5397–5403.

Blessing, R. H. (1995). *Acta Cryst.* A**51**, 33–38.

Brock, C. P., Dunitz, J. D. & Hirshfeld, F. L. (1991). *Acta Cryst.* B**47**, 789–797.

Bruker (2002). *SMART*. Bruker AXS Inc., Madision, USA.

Bruker (2004). *SAINT*. Bruker AXS Inc., Madision, USA.

Burnett, M. N. & Johnson, C. K. (1996). *ORTEP*III. Report ORNL-6895. Oak Ridge National Laboratory, Tennessee, USA.

Chęcińska, L., Mebs, S., Hübschle, C. B., Förster, D., Morgenroth, W. & Luger, P. (2006). *Org. Biomol. Chem.* **4**, 3242–3251.

Coppens, P. (2005). *Angew. Chem. Int. Ed.* **44**, 6810–6811.

Dittrich, B., Hübschle, C. B., Luger, P. & Spackman, M. A. (2006). *Acta Cryst.* D**62**, 1325–1335.

Dittrich, B., Hübschle, C. B., Messerschmidt, M., Kalinowski, R., Girnt, D. & Luger, P. (2005). *Acta Cryst.* A**61**, 314–320.

Dittrich, B., Koritsánszky, T., Grosche, M., Scherer, W., Flaig, R., Wagner, A., Krane, H. G., Kessler, H., Riemer, C., Schreurs, A. M. M. & Luger, P. (2002). *Acta Cryst.* B**58**, 721–727.

Dittrich, B., Koritsánszky, T. & Luger, P. (2004). *Angew. Chem.* **43**, 2773–2776.

Dominiak, P. M., Volkov, A., Li, X., Messerschmidt, M. & Coppens, P. (2007). *J. Chem. Theory Comput.* **3**, 232–247.

Frisch, M. J. *et al.* (1998). *GAUSSIAN*98, Revision A.7. Technical Report. Gaussian, Inc., Pittsburgh PA, USA.

Grimwood, D. J. & Jayatilaka, D. (2001). *Acta Cryst.* A**57**, 87–100.

Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.

Hübschle, C. B. & Luger, P. (2006). *J. Appl. Cryst.* **39**, 901–904.

Hübschle, C. B., Luger, P. & Dittrich, B. (2007). *J. Appl. Cryst.* **40**, 623–627.

Jayatilaka, D. & Grimwood, D. J. (2001). *Acta Cryst.* A**57**, 76–86.

Jayatilaka, D. & Grimwood, D. (2004). *Acta Cryst.* A**60**, 111–119.

Koch, U. & Popelier, P. L. A. (1995). *J. Phys. Chem.* **99**, 9747–9754.

Kohout, M. (2007). *DGrid and Basin*, Version 4.2. Technical Report. Max-Planck Institut für Chemische Physik fester Stoffe, Dresden.

Koritsánszky, T. *et al.* (2003). *XD*. Technical Report. Freie Universität Berlin, Germany.

Koritsanszky, T., Volkov, A. & Coppens, P. (2002). *Acta Cryst.* A**58**, 464–472.

Koritsanszky, T. S. & Coppens, P. (2001). *Chem. Rev.* **101**, 1583–1628.

Luger, P. (2007). *Org. Biomol. Chem.* **5**, 2529–2540.

Luger, P. & Dittrich, B. (2007). *The Quantum Theory of Atoms in Molecules*, edited by C. F. Matta & R. J. Boyd, pp. 317–342. Weinheim: Wiley VCH.

Luger, P., Messerschmidt, M., Scheins, S. & Wagner, A. (2004). *Acta Cryst.* A**60**, 390–396.

Luger, P., Wagner, A., Hübschle, C. B. & Troyanov, S. I. (2005). *J. Phys. Chem. A*, **109**, 10177–10179.

McKinnon, J. J., Mitchell, A. S. & Spackman, M. A. (1998). *Chem. Eur. J.* **4**, 2136–2141.

Pichon-Pesme, V., Lecomte, C. & Lachekar, H. (1995). *J. Phys. Chem.* **99**, 6242–6250.

Politzer, P., Murray, J. S. & Preralta-Inga, Z. (2001). *Int. J. Quantum Chem.* **85**, 676–684.

Schomaker, V. & Stevenson, D. P. (1941). *J. Am. Chem. Soc.* **63**, 37–40.

Spackman, M. A. & Byrom, P. G. (1997). *Chem. Phys. Lett.* **267**, 215–220.

Stewart, R. F. & Bentley, J. (1975). *J. Phys. Chem.* **63**, 3786–3793.

Su, Z. & Coppens, P. (1992). *Acta Cryst.* A**48**, 188–197.

Volkov, A., Koritsanszky, T., Li, X. & Coppens, P. (2004). *Acta Cryst.* A**60**, 638–639.

Young, D. W., Tollin, P. & Wilson, H. R. (1969). *Acta Cryst.* B**25**, 1423–1432.

Zarychta, B., Pichon-Pesme, V., Guillot, B., Lecomte, C. & Jelsch, C. (2007). *Acta Cryst.* A**63**, 108–125.