

Ultra-high-resolution X-ray structure of proteins

C. Lecomte^{a,*}, B. Guillot^a, N. Muzet^{a,b}, V. Pichon-Pesme^a and C. Jelsch^a

^a LCM3B, UMR CNRS 7036, Faculté des Sciences et Techniques, Université Henri Poincaré, Nancy 1, BP 239, 54506 Vandoeuvre-lès-Nancy (France), e-mail: claudelcomte@lcm3b.uhp-nancy.fr

^b Sanofi-Synthelabo Recherche, Drug Design, 16 rue d'Ankara, 67080 Strasbourg (France)

Abstract. The constant advances in synchrotron radiation sources and crystallography methods and the impulse of structural genomics projects have brought biocrystallography to a context favorable to subatomic resolution protein and nucleic acid structures. Thus, as soon as such precision can be frequently obtained, the amount of information available in the precise electron density should also be easily and naturally exploited, similarly to the field of small molecule charge density stud-

ies. Indeed, the use of a nonspherical model for the atomic electron density in the refinement of subatomic resolution protein structures allows the experimental description of their electrostatic properties. Some methods we have developed and implemented in our multipolar refinement program MoPro for this purpose are presented. Examples of successful applications to several subatomic resolution protein structures, including the 0.66 Å resolution human aldose reductase, are described.

Key words. Subatomic resolution; charge density; aldose reductase.

Introduction

Since 1990, the number of high- or ultra-high-resolution protein X-ray diffraction data has increased almost exponentially, as shown from Protein Data Base records and recent reviews [1–5]. We have recently shown that these data ($d < 1$ Å) deserve a more sophisticated model than the usual spherical free atom model called the Independent Atom Model (IAM), which does not take into account charge transfer and deformation of valence electron density due to chemical bonding and intermolecular interactions [4, 6–8]. Hence at these atomic resolutions, charge transfer and asphericity of electron density can be quantified using a model developed in small-molecule accurate crystallography: the multipole model [9, 10]. This multipole model, described in detail below, gives an analytical representation of the charge density which enables experimental estimation of electrostatic properties: electrostatic potential, electric field, electrostatic interaction energy [11–15] as well as dipole and quadrupole moments [16].

As high-resolution crystallography allows precise location of the atoms in the active site, including hydrogen

atoms, the resulting atomic model may also be used for first principles calculations using density-functional (DFT) theory. Recent DFT developments such as linear scaling, which scales linearly with the number of atoms of the system rather than with the cube [17, 18], allows quantum calculations for biological systems of ~1000 atoms. Such a procedure was first applied to the protein crambin with the software SIESTA [18]. Thus both experimental and theoretical methods are now available for fine estimation of the protein – ligand or protein – protein interactions. The following review describes the experimental X-ray method: It is divided into two sections: the first is devoted to the methods used, whereas the second section gives applications to a scorpion toxin [6] and to an aldose reductase complex [4].

Part I: Methodology

In small-molecule crystallography, experimental electron densities are obtained by analysis of single crystal X-ray diffraction measured to $d = 0.5$ Å resolution. For protein crystals, very high resolution comes from highly ordered crystals with a small proportion of solvent (<35% V/V). It implies that low temperature factors ($B < 8$ Å²) in the

* Corresponding author.

protein part are modelled by an aspherical atom model [19] which can be attained when the data resolution is smaller than 1 Å.

IAM: Independant Atom Model

Single-crystal X-ray diffraction data lead to structure factor amplitudes. The structure factors are the Fourier transform of the electron density $\rho_{\text{dyn}}(\mathbf{r})$ of the unit cell of volume V and parameters $a_i, i = 1, 3$.

$$F(\mathbf{H}) = \int_{\text{unit cell}} \rho_{\text{dyn}}(\mathbf{r}) \exp(2i\pi\mathbf{H}\cdot\mathbf{r}) d^3\mathbf{r}$$

where $|\mathbf{H}| = 2\sin\theta/\lambda$, θ is the Bragg angle, λ is the wavelength. The electron density ρ_{dyn} in the unit cell, which is affected by the atomic thermal motion, is obtained by a summation of convolution products:

$$\rho_{\text{dyn}}(\mathbf{r}) = \sum_{\text{atoms}} \rho_{\text{at, static}} * P_{\text{at}}(\mathbf{r})$$

P_{at} is an atomic probability distribution function, and its Fourier transform is the Debye Waller factor. The crystal is triply periodic; therefore, the Fourier transform has non-zero values only on reciprocal lattice points defined by the reciprocal vectors:

$$\mathbf{a}_i^* = (\mathbf{a}_j \wedge \mathbf{a}_k) V^{-1}$$

$F(\mathbf{H})$ are complex quantities, and both amplitude and phase must be known for directly calculating $\rho_{\text{dyn}}(\mathbf{r})$ by inverse Fourier transform. Methods for ab initio phase determination are still under development, but they are not the aim of this paper. As the electron density is mainly concentrated around atomic positions, the structure factor may be expressed as a sum over pseudoatoms in the unit cell:

$$F(\mathbf{H}) = \sum_j f_j(|\mathbf{H}|) \exp(2i\pi\mathbf{H}\cdot\mathbf{r}_j) \exp(-0.25 B_j |\mathbf{H}|^2)$$

where \mathbf{r}_j is the atomic position of the j^{th} atom, B_j its isotropic Debye Waller factors and f_j its atomic scattering factor; these latter are the Fourier transform of the electron density of the free spherical neutral atom (IAM).

This equation is the basic one for most macromolecular crystallographic refinements which fit the observed $|F|$'s values using a model with four parameters per atom, i. e. atomic coordinates and isotropic Debye Waller factor. At the usual resolution for macromolecular crystallography ($2 \text{ \AA} < d < 3 \text{ \AA}$), the resolution and therefore the number of observations $|F|$ is not enough to determine the total number of parameters, and it has to be completed with relations imposing a standard stereochemistry for the polypeptidic chain. At higher resolution one can observe deviations to the standard geometry. If the data are at atomic resolution ($d < 1.3 \text{ \AA}$), the isotropic temperature factor may be replaced by an anisotropic factor, and the resulting accuracy of atomic positions is good enough to validate shifts from the standard geometry.

Subatomic resolution and charge density modelling

At subatomic resolution ($d < 0.9 \text{ \AA}$), information on valence electron density distribution may be obtained when the anisotropic displacement parameters (ADPs) are small. Hydrogen atoms also clearly show up. Deviations from the spherical atom model appear as electron density peaks in the bonds on deformation electron density maps (calculated by the difference between the observed electron density and the IAM density). For example, in the aldoase reductase structure, 54% of hydrogen atoms were identified as well as most of the bonding density in the bonds of the active site of the protein [20, 21, E. Howard et al., unpublished]. The probability to observe these features is directly related to the thermal displacement parameters or to the equivalent B factor [1–3] (fig. 1).

Charge density refinement

Multipolar model and derived properties

The IAM model is too primitive to take into account all the information existing at subatomic resolution, and a new charge density model derived from small-molecule crystallography has been developed called the multipolar model [9, 22].

In contrast to the IAM model, where all atoms of a molecule or protein are supposed to be neutral with a spherical valence electron distribution (promolecule), the valence charge density is modelled by a sum of multipolar pseudoatoms lying at atomic positions. The valence electron density of such a pseudoatom is projected on the basis of real spherical harmonics functions $y_{lm}(\theta, \varphi)$ centred on each pseudoatom:

$$\rho_{\text{stat}}(\mathbf{r}) = \rho_{\text{core}}(\mathbf{r}) + \kappa^3 P_v \rho_{\text{val}}(\kappa\mathbf{r}) + \sum_{l=0, \text{Imax}} \kappa'^3 R_l(\kappa'\mathbf{r}) \sum_{m=\pm l} P_{lm} Y_{lm}(\theta, \varphi)$$

The radial functions $R_l(r)$ used here are of Slater type: $R_l(r) = r^{\text{nl}} \exp(-\kappa' \xi r)$. First applications of this formalism to mono- or dipeptides were the calibration of ab initio Hartree-Fock (HF) calculations: it was clearly shown that triple zeta basis sets with polarisation functions were necessary to quantitatively reproduce the X-ray diffraction experiment [23]. The charge density parameters P_v , P_{lm} (valence and multipolar populations, respectively), and κ , κ' (dilation and contraction of the spherical and non-spherical valence density, respectively) are directly obtained from least squares refinement against the structure factors amplitudes [8, 9, 22]. This analytical representation of the charge density is used to calculate crystal and molecular properties such as electrostatic potential [24], electric field, net charges, higher moments [16] and topology of the electron density [25]. As an example figure 2 shows the experimental electrostatic potential around NADP^+ , which is the cofactor of numerous enzy-

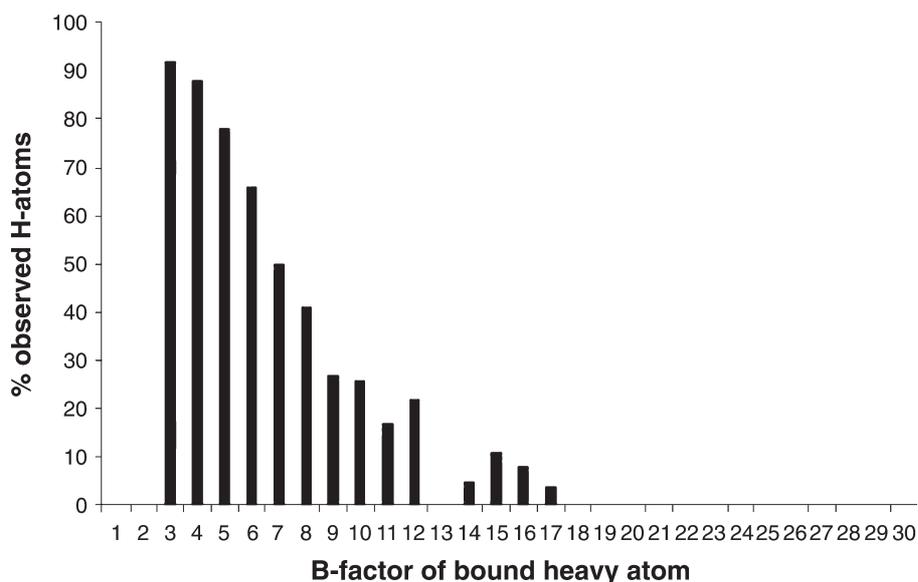


Figure 1. Percentage of the hydrogen atoms observed according to the equivalent B temperature factor of the bonded heavy atom in the structure of aldose reductase, refined at 0.66 Å resolution [20, 21].

matic oxydo-reduction reactions. This potential was modelled from a high-resolution X-ray multipolar analysis of its analogue molecule NAD^+ (see below) [26]. The electropositive and electronegative regions around the molecule are clearly evidenced. The positive electrostatic potential generated by the C18 atom (fig. 2) is coherent with the ability of NAD^+ to accept in this position a negatively charged hydride ion during oxidation reactions.

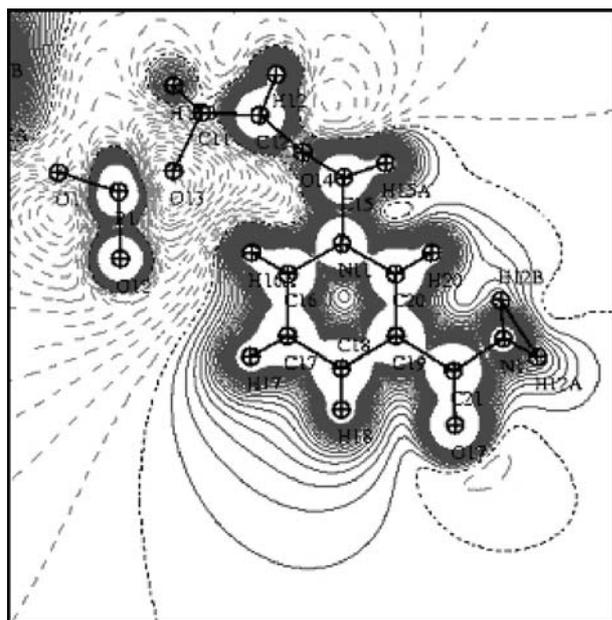


Figure 2. Electrostatic potential Φ generated by NAD^+ (calculated from subatomic resolution diffraction data). Contours 0.05 $\text{e}\text{\AA}^{-1}$, continuous line $\Phi > 0$, broken $\Phi < 0$ and dotted $\Phi = 0$ $\text{e}\text{\AA}^{-1}$.

The multipolar parameters library

High-resolution X-ray diffraction studies have been performed in Nancy on several peptides in order to precise the electron density distribution of all natural amino-acids (see for example 27, 28). These studies allowed building a database of atomic charge density parameters (P_v , P_{lm} , κ , κ') [29]. These parameters were shown to be transferable to amino acids in proteins [30]. Figure 3

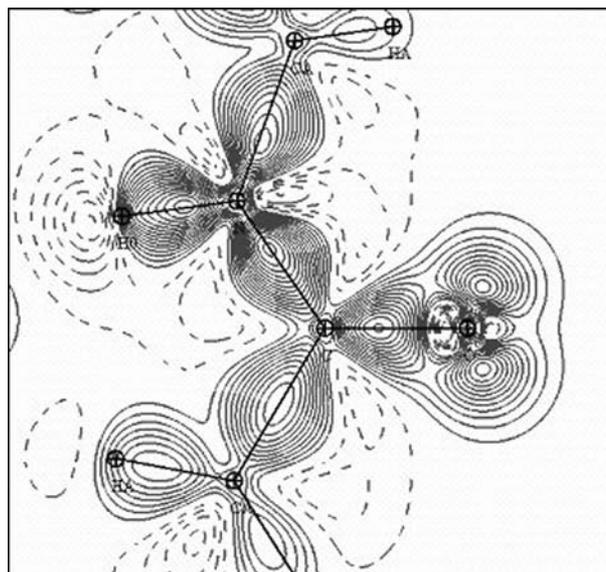


Figure 3. Deformation electron density in the peptide plane calculated from the multipolar database [29]. Contours 0.05 $\text{e}\text{\AA}^{-3}$, continuous lines $\delta\rho > 0$, broken lines $\delta\rho < 0$.

gives the static deformation electron density calculated from this multipolar parameters library for the CONH peptide group, in the following way:

$$\delta\varrho(\mathbf{r}) = \sum_{j=1, \text{Nat}} [\varrho_{\text{base}}(\mathbf{r}-\mathbf{r}_j) - \varrho_{\text{promolecule}}(\mathbf{r}-\mathbf{r}_j)]$$

This deformation density reveals the valence electron redistribution due to covalent interactions. The electrons built up in the bonding region and the oxygen lone pairs clearly show up. Disposing of a reliable charge density library, the next step has been testing this database on a protein. X-ray diffraction data on crambin, which is a small 46-residue protein, have been measured on the BW7A line of the DORIS (Hamburg) synchrotron to a resolution of $d = 0.54 \text{ \AA}$ by Teeter et al., which is still the world record for a protein [31]. This protein possesses all the necessary criteria such as very low Debye Waller factors ($B \sim 3 \text{ \AA}^2$) for ordered parts of the protein. Taking advantage of the repetition of the CONHC α H α chemical motif along the polypeptide main chain, the average dynamic deformation map (fig. 4) over the 34 nondisordered peptide residues was calculated according to:

$$\delta\varrho(\mathbf{r}) = \sum_{\text{H}} (F_o - F_c) \exp(i\varphi_c) \exp(-2i\pi\mathbf{H}\cdot\mathbf{r})$$

where F_c and φ_c are respectively the structure factor amplitude and the phase calculated from the IAM model (neutral, spherical atoms). F_o is the structure factor amplitude derived from the synchrotron experiment. This average deformation density map displays significant residual density in the bonds between nonhydrogen atoms and on oxygen lone pairs. These features clearly demonstrate that the IAM model does not provide a fully adequate fit to the experimental diffraction data.

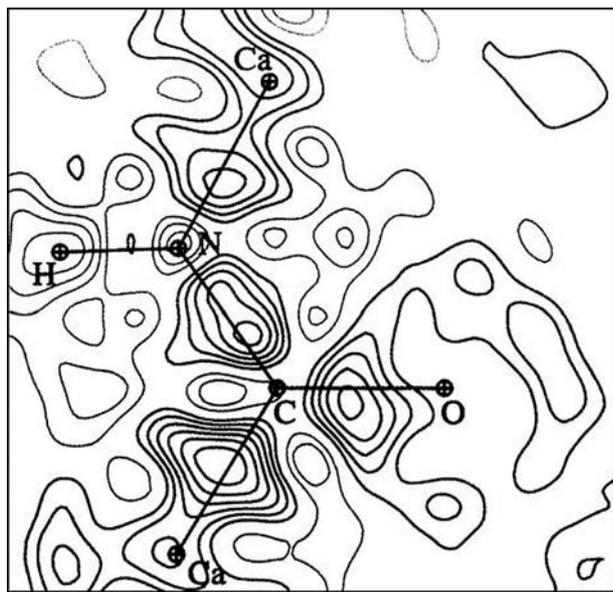


Figure 4. Residual electron density averaged over the 34 nondisordered peptide groups of crambin after IAM modelling. Positive density in black contours and negative in grey.

After transfer of the statistically significant multipole from the database and after multipolar refinement (using the software MOLLY [9]), the residual density does not exceed 0.06 e\AA^{-3} , which is about the estimated error: the progressive flattening of residual density features through the refinement stages was convincing physical evidence of real improvement in the modelling [7]. At the end of the refinement, the static deformation electron density of the average peptide residue (fig. 5) is in almost quantitative agreement with that derived from a triple zeta Hartree-Fock calculation on a single mono-peptide [23].

The MoPro refinement

As these results were very encouraging, a new software refinement program (MoPro) was written [22]; it combines the advantages of small molecule multipolar refinement with anisotropic atom refinement of proteins and runs on scalar or parallel computers. In order to extract the aspherical features of the electron density, the following strategy was designed. First, a full anisotropic spherical atom refinement using starting parameters from a SHELXL [32] refinement is performed on all nondisordered atoms with all diffraction data. Bond lengths, angle, planarity and rigid bond [33] restraints are applied or not, depending on the data quality and on the atomic B factors. Then, a subset of the structure based on equivalent B factors less than a given threshold ($8-12 \text{ \AA}^2$) is selected to perform the electron density analysis. At first, a high-order spherical atom refinement (HO) on the non-H atoms is performed to get least-biased positional and thermal parameters: hence, due to the properties of Fourier transform, this HO procedure refines the positional and anisotropic thermal motion parameters only on core electrons (fig. 6). Because valence electron scattering factors diffuse only at low resolution, HO refinement gives precise positions and displacements of the core electrons only.

In the second stage of the refinement, the starting deformation valence density parameters of the atoms of the se-

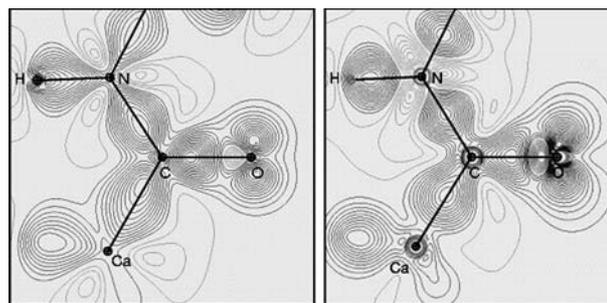


Figure 5. Static deformation density of a peptide plane in crambin (a), compared with a HF self-consistent fields (SCFs) calculation on a single peptide (b). Contours positive in black and negative in grey.

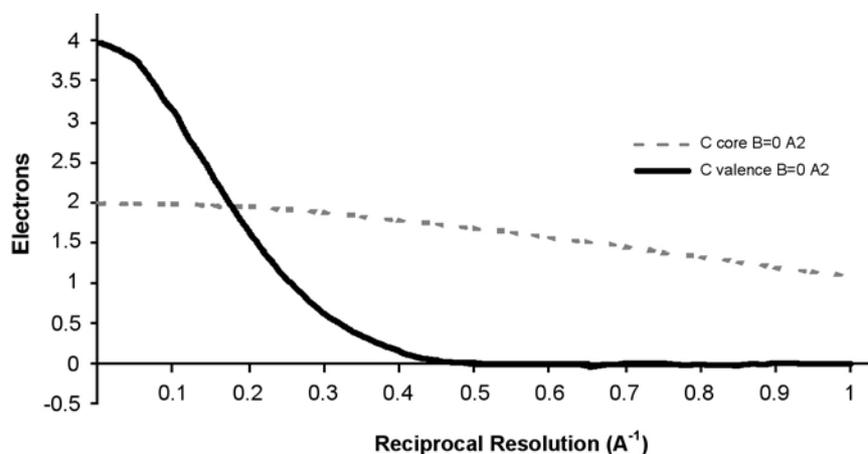


Figure 6. Atomic scattering factor for a hypothetical static carbon atom ($B = 0 \text{ \AA}^2$) as a function of reciprocal resolution s . The black line stands for valence electrons, and dotted grey line for core electrons.

lected fragments are automatically transferred from the charge density parameter database using a well-defined MoPro procedure. This means that all atoms of the selected fragment are charged atoms and are assigned nonspherical atom-scattering factors. All H atoms are displaced along the X-H bonds to standard bond distances obtained from neutron studies [4]. Usually, after a few refinement cycles of X , Y , Z and U_{ij} , the transfer procedure leads to a large improvement of the refinement. Because a part of the nonspherical atom-scattering density is taken into account, it also leads to more accurate positions and thermal motion parameters (see for example the charge density analysis of NAD^+ [26]). Then the charge density parameters, including H atoms, can be refined using or not, depending on the data quality, electron density restraints or constraints (chemical equivalence, site symmetry) that are part of the MoPro algorithms.

At the end of the refinement, one gets the best experimental description of the electron density, which can be used to compute the electrostatic potential [14]. This procedure, as described below, was successfully applied to a human aldose reductase complex, for which ultra-high-resolution data (0.66 \AA) were collected by Podjarny et al. [4, 20] on the ID19 beamline at the Advanced Photon Source synchrotron (APS-ANL, Argonne, IL, in collaboration with A. Joachimiak).

Part II: some examples

Protein electrostatics using the charge density database

A first application of these methods is the calculation of electrostatic properties for proteins with X-ray data at slightly lower resolution (about $0.9 < d < 1.5 \text{ \AA}$). At these resolutions, the atomic positions, including H atoms, of the active site are clearly defined. Therefore a direct use

of the charge density database [29] permits, using the MoPro software, a quick calculation of the electrostatic potential, which was shown to compare very well with more time-consuming theoretical methods such as Density Functional Theory (DFT) calculations [4]. This can now be almost routinely performed, at very low cost, after any atomic spherical refinement. A successful application has been obtained with the 0.66 \AA aldose reductase complex, although this resolution clearly allows an effective refinement of the transferred charge density parameters (see below).

As shown on figure 7, the complex is made from the aldose reductase protein, the NADP^+ cofactor and an inhibitor. In order to analyse the interaction between NADP^+ and the protein, first, a charge density study was made on a NADP^+ analogue: the NAD^+ cofactor. From this experimental multipolar analysis, the deformation density parameters of NADP^+ have been modelled, then added to the database. The electrostatic potential calculation was therefore performed on a 64-amino acids (711-atoms) substructure (included in the region highlighted in fig. 7b) surrounding the active site, with and without the NADP^+ molecule, using the charge and multipolar parameters of the database. As the inhibitor definition was not available in the data base, the system used for the electrostatic potential computations do not include its contribution.

The electrostatic potential of the free NADP^+ cofactor has already been discussed above and is shown in the active site orientation on figure 2. Figure 8a gives the electrostatic potential generated by the holoenzyme structure in the active site, plotted in the plane of the NADP^+ nicotinamide ring. Figure 8b shows the electrostatic potential, in the same orientation as figures 8a and 2, obtained for the apoenzyme, i.e. without the NADP^+ contribution. The apoenzyme electrostatic potential of the binding pocket shows two electropositive regions located (top left and

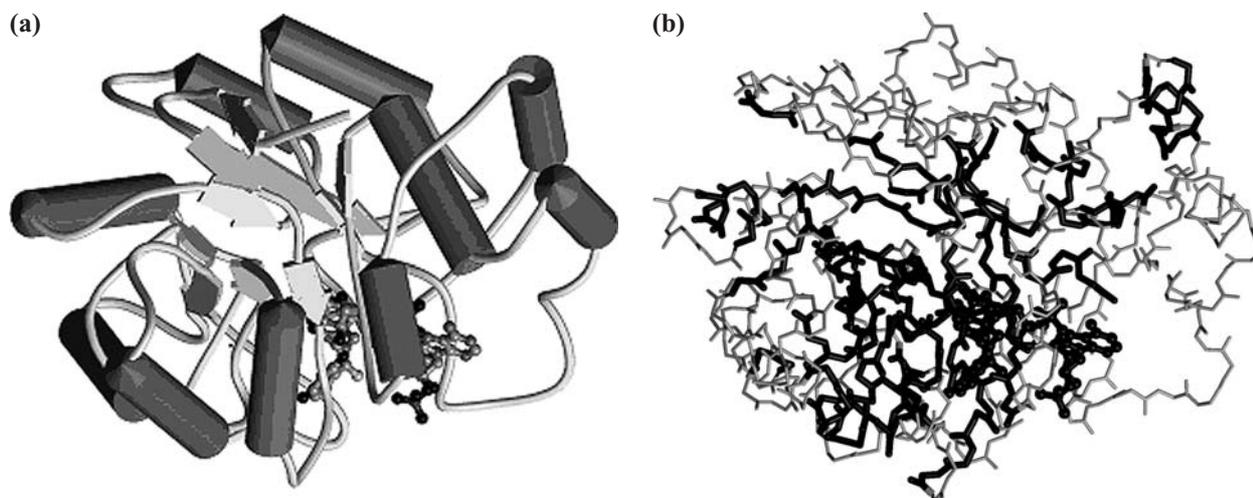


Figure 7. (a) Schematic view of the human aldose reductase structure, with alpha helices represented as tubes, beta strands as arrows and loops as coils. The NAD⁺ cofactor and the inhibitor are shown in ball-and-stick-mode, indicating the position of the active site. (b) View of the protein C-alpha trace, showed in the same orientation as (a), with highlighted in bold black – the 119 residues of the structure for which the equivalent *B* factors (averaged over non-H atoms) are lower than 4 Å².

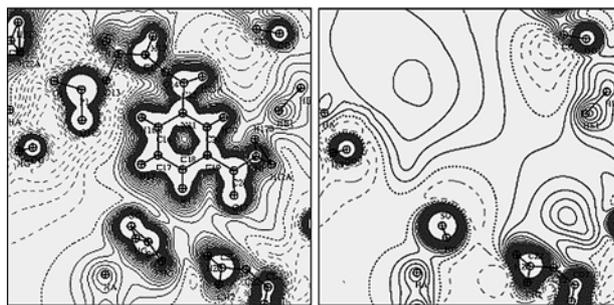


Figure 8. (a) Electrostatic potential computed with transferred multipolar parameters as generated by the holoenzyme in the active site, represented in the plane of the NAD⁺ nicotinamide ring. (b) Electrostatic potential computed without the NAD⁺ contribution, represented in the same orientation as (a). Contours are 0.05 eÅ⁻¹, positive, negative and zero potentials, respectively, in full, broken and dotted lines.

bottom right in fig. 8b) and two electronegative regions (top right and bottom left). Comparison of figures 2 and 8 shows that the electrostatic potential generated by the cofactor alone (in the holoenzyme conformation, fig. 2) and the one obtained in the apoenzyme active site (fig. 8b) are clearly complementary. For instance, the electronegative potentials generated by the pyrophosphate and the amide oxygen atom of NAD⁺ are qualitatively complementary to the electropositive potential in the active site, and the same observation can be made for active site electronegative potentials and NAD⁺ positive ones. This is all the more remarkable as the NAD⁺ electrostatic potential has been obtained in a totally independent way: without any assumption related to the active site geometry or electrostatics. The results of this study were

the first experimental charge density demonstration of an electrostatic complementarity between a protein environment and its ligand.

Another application has been performed with the allosteric insulin hexamer structure, solved at atomic resolution [34].

Protein charge density refinement

As previously noticed, the subatomic resolution in conjunction with low-to-moderate atomic thermal motion allows the refinement of the charge density parameters. The starting values for the refinement of the multipolar parameters can be either taken as IAM (neutral valence populations and null multipolar parameters), or as transferred from the multipolar database. We have shown that the latter option leads to more precise charge density description [35].

The first test was performed on the toxin II of the *Androctonus Australis* Hector scorpion, for which diffraction data at room temperature were collected to $d = 0.96$ Å resolution [6]. Although the thermal smearing of electron density was higher (average 8 Å²), the procedure described above allowed enhancement of the nonspherical electron density, as shown on figure 9, where strong electron density peaks are visible on each covalent bonds of the peptide plane. However, the effects of both limited resolution and quite high atomic thermal motion are noticeable around the carbonyl oxygen atom, where no clearly defined accumulation of electron density accounting for the electron lone pairs are visible.

The next application of the method was the 0.54 Å crambin data [7] collected at DORIS (Hamburg) by M. Teeter

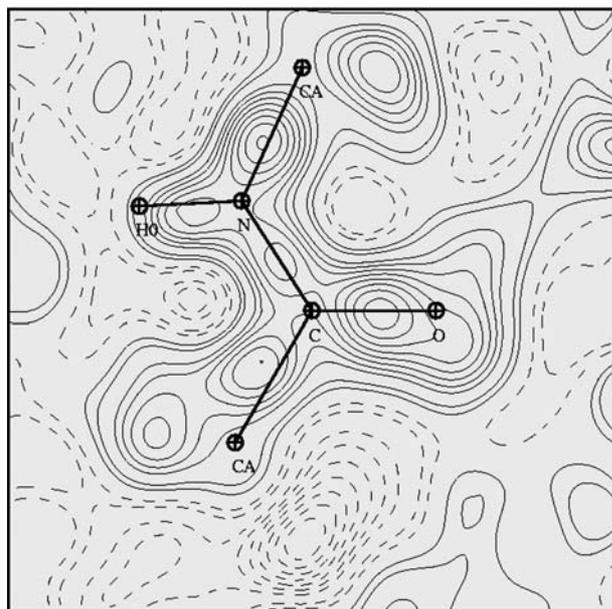


Figure 9. Dynamic deformation density on the Tyr47-Cys48 peptide group in a scorpion toxin II, computed with transferred multipolar parameters. Contours are $0.02 \text{ e}\text{\AA}^{-3}$, positive in solid lines, negative in dashed lines.

and V. Lamzin. This study has demonstrated the feasibility of the method and was discussed in part I.

The third application, which is at this time still under way, is the 0.66 \AA resolution structure of the ternary complex human aldose reductase – NADP⁺ cofactor – IDD594 inhibitor. A multipolar refinement of the transferred charge density parameters against the experimental data has been performed. Again, the system effectively used in the refinement was not the full atomic model but only regions of the structure with moderate thermal motion (fig. 7b). In other words, regions of multiple conformations were avoided, and both C and N terminus as well as water molecules were not included in the structure subset. The working subset is made of 119 of the 316 amino acids of the enzyme, including the active site residues and the two bonded molecules NADP⁺ and inhibitor.

All the methods depicted above have been applied. First, an High order (Ho) refinement of structural parameters has been performed with MoPro [22] for atoms of the structure subset, against successive narrowing high-resolution ranges: the first one between 1.0 and 0.66 \AA and the last one between 0.75 and 0.66 \AA . This method avoids instabilities in the refinement due to large parameter shifts when the working resolution range goes directly from the data to very high resolution only. The presence of electron density on the covalent bonds in the residual maps ($F_{\text{obs}} - F_{\text{calc}}$) in the signal that significant charge density needs to be modelled for an accurate refinement. This HO refinement leads to significant sharpening of these features in the residual electron density maps. Strong bonding density peaks are visible in the middle of most of

the covalent bonds of the structure subset, as well as around many oxygen atoms, corresponding to the lone pairs. Even in regions usually less ordered in protein structures, such as long amino acid side chains, the bonding densities are clearly visible, as illustrated in figure 10 in the case of the Lys262 residue.

The next stage of the procedure consists in the transfer of database multipolar parameters to the atomic model obtained at the end of the high-order refinement. This transfer leads to an immediate and drastic improvement of the crystallographic agreement factors, with $R(F)$ dropping from 9.28 to 8.79% , and $R_{\text{free}}(F)$ from 9.45 to 9.16% (using all the $491,000$ experimental data). It means that this procedure accounts for most of the residual densities, which are the signature of aspherical features due to covalent bonding, nonmodelled by the standard spherical refinement using an AIM model. The transferred charge density parameters are then refined against the full-resolution range with the application of symmetry and chemical equivalence constraints on all moieties. Resulting static deformation densities are represented in figure 11 for the protein peptide plane (fig. 11 a) and in the plane of a tyrosine residue side chain (fig. 11 b). A comparison between figure 11 a and the peptide group deformation electron density as described in the database (fig. 3) reveals that no significant deviation occurs for the refined parameters (root mean square deviation = $0.06 \text{ e}\text{\AA}^{-3}$) when the starting values are taken from the database. Actually, this is expected from such a procedure, as this constrained refinement leads to the average protein deformation density, which is close to that described in the database. Figure 11 b shows the static deformation density in the side-chain plane of a tyrosine residue. The quality of the average charge density in this protein case is comparable to the results one could expect for the electron density on an individual moiety in small-molecule refinement.

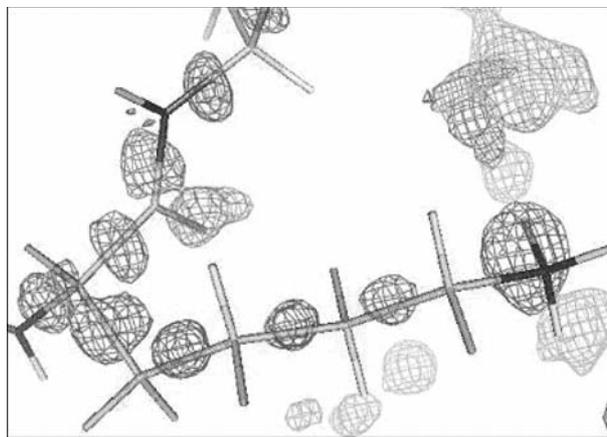


Figure 10. Residual electron density in the Lys262 region of the 0.66-\AA resolution aldose reductase complex, showing bonding densities along the lysine main and side chains. Contours are 2.4 sigma units.

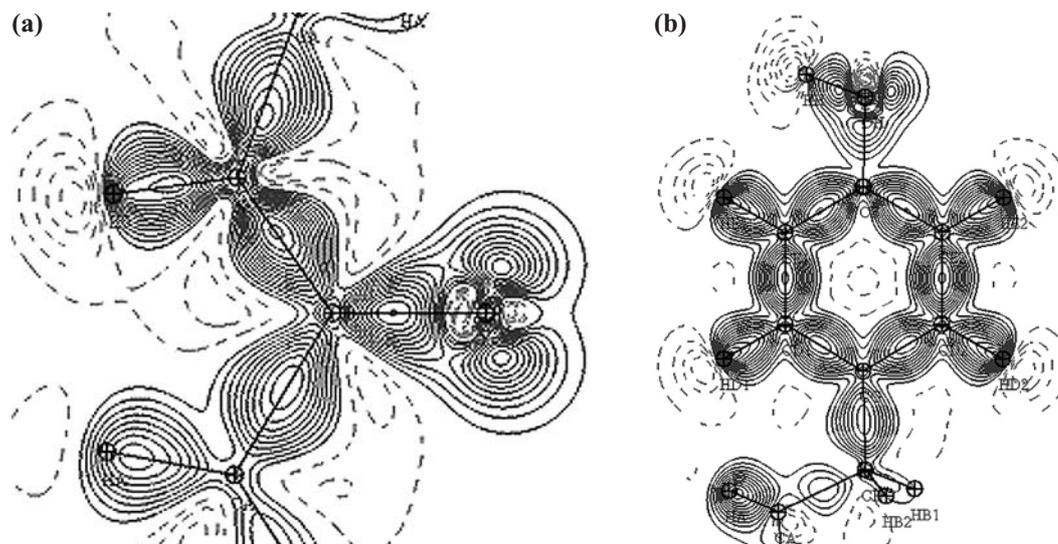


Figure 11. Static average deformation densities after a constrained multipolar refinement of the 0.66 Å Human aldose reductase structure. (a) in the protein peptide plane, (b) in the C_6 plane of a tyrosine residue. Contours are $0.05 \text{ e} \cdot \text{Å}^{-3}$, positive in full and negative in broken lines.

Conclusion

We have shown that high-resolution and ultra-high-resolution refinement of proteins is now possible using the multipole model and the new software MoPro. This permits calculation of electrostatic properties, which allows understanding of interactions among proteins, ligands and cofactors. Recent developments of diffraction techniques at third-generation synchrotrons, progress in cryocrystallography and crystallisation promise more high-resolution data sets which will require aspherical models yielding more accurate structure and protein electrostatics. Electrostatic energy calculations are under way, and the resulting energies will be used, for example, in enzymology to calibrate inhibitors in relation to biochemical activity. Further development of the charge density library will also allow a more quantitative description of the interactions between proteins and nucleic acids.

Acknowledgements. This work was financed by the Centre National de la Recherche Scientifique (CNRS) and by the University Henri Poincaré-Nancy-1 (UHP). N. M. was supported by a grant from the Association pour la Recherche sur le Cancer and B. G. by a PhD fellowship from the French Ministry of Research and Technology.

- 1 Dauter Z., Lamzin V. S. and Wilson K. S. (1997) The benefit of atomic resolution. *Curr. Opin. Struct. Biol.* **7**: 681–688
- 2 Longhi S., Czjzek M. and Cambillau C. (1998) Messages from ultrahigh resolution crystal structures. *Curr. Opin. Struct. Biol.* **8**: 730–737
- 3 Dauter Z., Lamzin V. S. and Wilson K. S. (1995) Proteins at atomic resolution. *Curr. Opin. Struct. Biol.* **5**: 784–790
- 4 Muzet N., Guillot B., Jelsch C., Howard E. and Lecomte C. (2003) Electrostatic complementarity in an aldose reductase

complex from ultra-high-resolution crystallography and first-principles calculations. *Proc. Natl. Acad. Sci. USA* **100**: 8742–8747

- 5 Schmidt A., Jelsch C., Rypniewski W. and Lamzin V. S. (2003) Trypsin revisited: crystallography at (sub)atomic resolution and quantum chemistry revealing details of catalysis. *J. Biol. Chem.* **278**: 43357–43362
- 6 Housset D., Benabicha F., Pichon-Pesme V., Jelsch C., Maierhofer A., David S. et al. (2000) Towards the charge density of proteins: a scorpion toxin at 0.96 Å resolution as a first test case. *Acta Cryst. D* **56**: 151–160
- 7 Jelsch C., Teeter M. M., Lamzin V., Pichon-Pesme V., Blessing R. H. and Lecomte C. (2000) Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin. *Proc. Natl. Acad. Sci. USA* **97**: 3171–3176
- 8 Stewart R. F. (1969) Generalized X-ray scattering factors. *J. Chem. Phys.* **51** **10**: 4569–4577
- 9 Hansen N. K. and Coppens P. (1978) Testing aspherical refinement on small molecules data sets. *Acta Cryst. A* **34**: 909–921
- 10 Spackman M. A. (1998) Charge densities from X-ray diffraction data. *Ann. Rep. Prog. Chem. C: Phys. Chem.* **94**: 177–207
- 11 Bouhaida N., Ghermani N. E., Lecomte C. and Thalal M. (1999) Molecular fragment electric moments derived from the fit of the experimental electrostatic potential. Application to the water molecule. *Acta Cryst. A* **55**: 729–738
- 12 Stewart R. F. and Craven B. M. (1993) Molecular electrostatic potentials from crystal diffraction: the neurotransmitter gamma-amino butyric acid. *Biophys. J.* **65**: 998–1005
- 13 Spackman M. A. (1986) Atom-atom potentials via electron gas theory. *J. Chem. Phys.* **85**: 6579–6586
- 14 Bouhaida N., Ghermani N. E., Lecomte C. and Thalal M. (1997) Modelling electrostatic potential from experimentally determined charge densities. II. Total potential. *Acta Cryst. A* **53**: 556–563
- 15 Abramov Y. A., Volkov A., Wu G. and Coppens P. (2000) The experimental charge-density approach in the evaluation of intermolecular interactions. Application of a new module of the XD programming package to several solids including a pentapeptide. *Acta Cryst. A* **56**: 585–591
- 16 Spackman M. A. (1992). Molecular electric moments from x-ray diffractions data. *Chem. Rev.* **92**: 1769–1797

- 17 Soler J. M., Artacho E., Gale J. D., Garcia A., Junquera J., Ordejon P. and Sanchez-Portal D. (2002) The SIESTA method for ab initio order-N materials simulation. *J. Phys. Condens. Matter* **14**: 2745–2779
- 18 Fernandez-Serra M. V., Junquera J., Jelsch C., Lecomte C. and Artacho E. (2000) Electron density in the peptide bonds of crambin. *Solid State Comm.* **116**: 395–400
- 19 Podjarny A., Howard E., Mitschler A., Chevrier B., Lecomte C., Guillot B. et al. (2002) X-ray crystallography at subatomic resolution. *Europhysics News* **33** **4**: 113–117
- 20 Howard E., Cachau R. E., Mitschler A., Chevrier B., Barth P., Lamour V. et al. (2002) Reference removed in proof
- 21 Cachau R., Howard E., Barth P., Mitschler A., Chevrier B., Lamour V. et al. (2000) Model of the catalytic mechanism of human aldose reductase based on quantum chemical calculations. *Journal de Physique* **10**: 3–13
- 22 Guillot B., Viry L., Guillot R., Lecomte C. and Jelsch C. (2001) Refinement of proteins at subatomic resolution with MoPro. *J. Appl. Cryst.* **34**: 214–223
- 23 Souhassou M., Lecomte C., Ghermani N. E., Rohmer M. M., Wiest R., Bénard M. et al. (1992) Electron distributions in peptides and related molecules. 2. An experimental and theoretical study of (Z)-N-acetyl- α , β -dehydro-phenylalanine methylamide. *J. Am. Chem. Soc.* **114**: 2371–2382
- 24 Ghermani N. E., Lecomte C. and Bouhaida N. (1993) Electrostatic potential from high resolution X-ray diffraction data: application to a pseudopeptide molecule. *Z. Naturforsch. Teil A*. **48**: 91–98
- 25 Souhassou M. and Blessing R. H. (1999) Topological analysis of experimental electron densities. *J. Appl. Cryst.* **32**: 210–217
- 26 Guillot B., Muzet N., Artacho E., Lecomte C. and Jelsch C. (2003) Experimental and theoretical electron density studies in large molecules: NAD⁺, beta-nicotinamide adenine dinucleotide. *J. Phys. Chem. B*. **107**: 9109–9121
- 27 Souhassou M., Lecomte C., Blessing R. H., Aubry A., Rohmer M. M., Wiest R. et al. (1991) Electron distributions in peptides and related molecules. 1. An experimental and theoretical study of N-acetyl-L-tryptophan methylamide. *Acta Cryst. B* **47**: 253–266
- 28 Pichon-Pesme V., Lachekar H., Souhassou M. and Lecomte C. (2000) Electron density and electrostatic properties of two peptide molecules: tyrosyl-glycyl-glycine monohydrate and glycyl-aspartic acid dihydrate. *Acta Cryst. B* **56**: 728–737
- 29 Pichon-Pesme V., Lecomte C. and Lachekar H. (1995) On building a databank of transferable experimental electron density parameters: application to polypeptides. *J. Phys. Chem.* **99**: 6242–6250
- 30 Jelsch C., Pichon-Pesme V., Lecomte C. and Aubry A. (1998) Transferability of multipole charge-density parameters: application to very high resolution oligopeptide and protein structures. *Acta Cryst. D* **54**: 1306–1318
- 31 Lamzin V. S., Morris R. J., Dauter Z., Wilson K. S. and Teeter M. M. (1999) Experimental observation of bonding electrons in proteins. *J. Biol. Chem.* **274**: 20753–20755
- 32 Sheldrick G. M. and Schneider T. S. (1997) SHELXL: high resolution refinement. *Methods Enzymol.* **276**: 319–343
- 33 Hirshfeld F. L. (1976). Can X-ray data distinguish bonding effects from vibrational smearing? *Acta Cryst. A* **32**: 239–244
- 34 Smith G. D., Pangborn W. A. and Blessing R. H. (2003) The structure of T6 human insulin at 1.0 Å resolution. *Acta Cryst. D* **59**: 474–482
- 35 Guillot B. (2002) Extension des méthodes d'affinement multipolaire des petites molécules aux macromolécules. Thèse de l'Université Henri Poincaré – Nancy 1



To access this journal online:
<http://www.birkhauser.ch>
