Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

On the possibility of the observation of valence electron density for individual bonds in proteins in conventional difference maps

Pavel V. Afonine,[a,b]‡ Vladimir Y. Lunin,[c] Nicolas Muzet[d] and Alexandre Urzhumtsev[a]*

[a]LCM3B, UMR 7036 CNRS, Université Henri Poincaré, Nancy 1, BP 239, Faculté des Sciences et Techniques, 54506 Vandoeuvre-lès-Nancy, France, [b]Centre Charles Hermite, LORIA, 54602 Villers-lès-Nancy, France, [c]Institute of Mathematical Problems of Biology, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia, and [d]Sanofi–Synthelabo Recherche, 67080 Strasbourg, France

‡ Current address: Lawrence Berkeley National Laboratory, Building 4, Mailstop 4-230, 1 Cyclotron Road, Berkeley, CA 94720, USA.

Correspondence e-mail: alexander.ourjoumtsev@lcm3b.uhp-nancy.fr

Received 31 August 2003
Accepted 12 November 2003

In the last decade, high-resolution data have become available for macromolecular objects. Furthermore, ultrahigh-resolution diffraction data (resolution close to 0.6 Å) have been collected for several protein crystals. This allows the study of fine details of the electron-density distribution such as the deformation density, *i.e.* the deviation of the experimentally determined electron density from the density composed of 'free' non-bonded atoms. This paper discusses the resolution and atomic temperature factors necessary to make the valence electron density visible at individual bonds in conventional difference maps for macromolecules. The study of theoretical maps calculated by quantum-chemistry methods allows estimation of these conditions; these results are confirmed by analysis of experimental maps for Leu-enkephalin and antifreeze protein RD1. A resolution limit close to 0.6 Å was found to be highly important for refinement even when the maps were calculated at lower resolution. The refinement of the same models at near to 0.9 Å resolution results in artificially increased values of the atomic displacement parameters and does not permit bond electron density to be visible in difference maps. To some extent, overestimation of the atomic displacement parameters may be restricted if dummy bond electrons are used in the refinement.

1. Introduction

Recent progress in experimental techniques has made macromolecular crystals that diffract to a resolution higher than 1 Å no longer a dream but a reality. A number of papers have been published on structures determined at a resolution of approximately 0.9 Å. Furthermore, a resolution close to 0.6 Å has been achieved for several proteins [crambin, 0.54 Å (Teeter *et al.*, 1993; Jelsch *et al.*, 2000); antifreeze protein, 0.62 Å (Ko *et al.*, 2003); aldose reductase, 0.66 Å (Howard *et al.*, 2004)]. At this resolution, H atoms become routinely visible in difference Fourier syntheses and the question may be posed as to what other structural details may be recognized when investigating these maps. In the following, we use the term 'high resolution' (HR) to refer to a resolution around 0.9 Å and the term 'ultrahigh resolution' (UHR) for that close to 0.6 Å. Tests show (see §4 below) that there exists a significant difference between the results of refinement carried out at HR and UHR.

In conventional macromolecular study, the distribution of electron density is modelled by a sum of 'spherical' atoms (usually of a multi-Gaussian type) smeared by an anisotropic thermal motion. Obviously, this distribution is not exactly the same as the real electron density in the molecule, where covalent bonding makes the distribution for an atom different from a spherical atom. The 'deformation map' (*i.e.* the

difference between the true electron density and the density corresponding to a 'promolecule' composed of spherical atoms) has been a subject of thorough studies in small-molecule crystallography since the middle of the 20th century (*e.g.* Ewald & Hönl, 1936; Cochran, 1953; for contemporary reviews, see Coppens, 1998; Coppens *et al.*, 1999). A special 'multipolar' model based on the expansion of the density deformation into a series of spherical harmonics has been developed (Stewart, 1969; Hansen & Coppens, 1978) in order to take this deformation into account. Currently, this model is in routine use in studies of inorganic and small organic compounds. The appearance of UHR data for proteins and nucleic acids made it possible to perform the first attempts to transfer these ideas and approaches to macromolecular objects (Wiest *et al.*, 1994; Jelsch *et al.*, 1998, 2000; Housset *et al.*, 2000; Muzet *et al.*, 2003).

A multipolar model of the electron density allows more accurate structure-factor phases to be calculated and used when preparing the difference Fourier synthesis

$$\rho_{\rm res}(\mathbf{r}) = \frac{1}{V}\sum_{|\mathbf{s}|<1/d_{\rm min}} \{F^{\rm obs}(\mathbf{s})\exp[i\varphi^{\rm calc}_{\rm mult}(\mathbf{s})] - F^{\rm calc}_{\rm sph}(\mathbf{s})\exp[i\varphi^{\rm calc}_{\rm sph}(\mathbf{s})]\}\exp[-2\pi i(\mathbf{s},\mathbf{r})]. \quad (1)$$

Here, $F^{\rm obs}$ represents the experimentally observed magnitudes and $F^{\rm calc}_{\rm sph}$ and $\varphi^{\rm calc}_{\rm sph}$ represent structure factors corresponding to a conventional 'spherical-atom' model smeared by anisotropic thermal motion, while $\varphi^{\rm calc}_{\rm mult}$ represents phases calculated from the multipolar model. Such syntheses are routinely used to study precisely the density deformation of inorganic and small organic compounds.

On the other hand, it is well known in crystallography that structural details erroneously included in a model are usually present in the syntheses built with the observed magnitudes and model phases (Main, 1979; Hodel *et al.*, 1992). The 'image' of these false details is reduced in magnitude (Main, 1979) but still exists in such syntheses. Therefore, maps built with multipolar-model phases may be biased to the multipole parameters used and may reveal the multipolar model itself rather than the real density deformation. This explains an interest to the study of the density deformation using conventional crystallographic difference maps

$$\rho_{\rm res}(\mathbf{r}) = \frac{1}{V}\sum_{|\mathbf{s}|<1/d_{\rm min}} [F^{\rm obs}(\mathbf{s}) - F^{\rm calc}(\mathbf{s})]\exp[i\varphi^{\rm calc}(\mathbf{s})] \times \exp[-2\pi i(\mathbf{s},\mathbf{r})], \quad (2)$$

where $F^{\rm calc}$, $\varphi^{\rm calc}$ are calculated from a spherical-atom model and do not contain any prior information on a possible density redistribution. Such a map may present a more objective witness of the existence of a signal, while map (1) may provide a more accurate shape of this signal.

The first attempts to recognize the deformation density in macromolecular Fourier maps [not necessarily in a difference '1–1' map (2) but in some other kinds of maps as well] were made by Jelsch *et al.* (1998) and by Lamzin *et al.* (1999). In particular, Lamzin and coworkers reported that when studying crambin at 0.67 Å resolution they failed to observe deformation density for individual residues, but that averaging over 40 peptide units provided them with a signal. These authors have also obtained similar results for savinase at 0.9 Å resolution. Several structures have been reported (Jelsch *et al.*, 1998, 2000; Schuerman & Van Meervelt, 2000; Ko *et al.*, 2003; Howard *et al.*, 2004) in which peaks in the middle of individual bonds (bond electrons) were seen for well ordered parts of the structure, indicating the deformation density. Our attempts to observe bond electrons for individual bonds in different macromolecular structures resulted in a partial success: these peaks were observed for some structures but not all. This stimulated a more thorough study of the conditions necessary to observe the bond electron density.

Three stages of this study are presented below. We start (§3) from the study of the density-deformation map calculated theoretically (see §2.2 for details) for the molecule $CH_3-NH-CO-CH_3$ containing the peptide unit. The goal of this stage was to estimate to what extent the truncation of Fourier series at the highest presently possible resolution for macromolecules (about 0.6 Å) and smearing by the thermal motion destroy the peaks of the deformation density. It follows from these studies that even in the ideal case when the phase errors are absent, the peaks for individual bonds cannot be seen if the equivalent isotropic temperature-factor values $B_{\rm iso}$ are higher than $\sim$7 Å^2. For lower $B_{\rm iso}$ values, deformation-density peaks are seen at individual bonds in theoretical difference maps but not necessarily in practice, where errors inherent to the calculated structure factors in (2) may result in a further deterioration of the deformation-density image. To check the latter effect, difference maps calculated with observed X-ray data for Leu-enkephalin (Wiest *et al.*, 1994) and eel pout type III antifreeze protein RD1 (Ko *et al.*, 2003) were analysed. It was found that deformation-density peaks are still regularly observed at individual bonds for residues that satisfy the $B_{\rm iso}$ condition when the model is refined at ultrahigh resolution. This agrees with observations made previously for crambin (Jelsch *et al.*, 2000) and for aldose reductase (Guillot, 2002) that the bond electron density is expected to be observed for residues possessing a mean value of the equivalent $B_{\rm iso}$ factor that is lower than 5 Å^2.

The study of images of theoretical deformation maps shows that for well ordered parts of the structure the bond electron density is visible even when the resolution of the difference synthesis is reduced from the UHR to the HR level, for example to 0.9 Å. A further check with experimental diffraction data for enkephalin and antifreeze protein confirmed this conclusion. Nevertheless, all our attempts to detect bond electron density in 0.9 Å syntheses failed for a number of structures when applied to models refined at this resolution (about 0.9 Å). The results of tests for $CH_3-NH-CO-CH_3$, enkephalin and antifreeze protein (discussed in §4) show a drastic change in the difference maps at the same resolution of 0.9 Å when the refinement resolution zone is extended from 0.9 Å to approximately 0.6 Å. When calculated at the same resolution, the maps obtained after refinement at ultrahigh resolution show the bond-electron density, while those obtained after a high-resolution refinement do not.

Comparison of the results of refinements at these resolutions shows that the models refined at about 0.9 Å possess artificially large atomic displacement parameters (ADPs). This possibility that atoms with artificially increased ADPs could mask the deformation density is known (Housset *et al.*, 2000) and has been discussed previously for small molecules (Coppens, 1967). Our tests demonstrate that transition to the ultrahigh-resolution data set restricts such an overestimation of the ADPs. Another possibility for restricting the ADPs when refining atomic models at a high resolution is to use dummy bond electrons (DBEs), as discussed in §4.3. The idea of placing artificial 'electrons' near the middle of the bonds has been discussed for small molecules since the 1960s (Brill, 1960; Brill *et al.*, 1971; Coppens & Lehmann, 1976; Hellner, 1977; Pietsch, 1981) and was recently incorporated in macromolecular studies (Afonine *et al.*, 2002); some details are given in §4.2. This idea is close to the more general approach of the use of dummy atoms in phase improvement (Agarwal & Isaacs, 1977; Lunin & Urzhumtsev, 1984; Lunin *et al.*, 1985; Lamzin & Wilson, 1993; Wilson & Agard, 1993; Urzhumtsev, 1997) realised in the *ARP*/*wARP* program (Lamzin & Wilson, 1993; Lamzin *et al.*, 2001).

Table 1
Main characteristics of the structures used for the tests.

DBEs were included in the models only for some refinements, as indicated in Table 2.

Object	No. C, N, O, S atoms	No. water O atoms	No. H atoms	No. DBEs (when used)	Space group	Unit-cell parameters (Å) ($\alpha = \beta = \gamma = 90^\circ$)	Refinement method
$CH_3-HN-CO-CH_3$	5	—	7	11	$P1$	$a = 18.0$, $b = 18.0$, $c = 15.0$	LS
Enkephalin	40	3	43	92	$P2_12_12_1$	$a = 10.9$, $b = 13.1$, $c = 21.2$	CGLS
RD1	490	240	530	321	$P2_12_12_1$	$a = 32.5$, $b = 39.5$, $c = 44.6$	CGLS

Table 2
Main characteristics of test refinements.

Object/ resolution (Å)	DBE	No. of reflections	Refined parameters for: C, N, O, S atoms	Water O atoms	H atoms	DBE	Total No. refinement parameters
$CH_3-HN-CO-CH_3$							
0.90	No	13928	U_{aniso}	—	B_{iso}	—	38
0.60	No	47051	U_{aniso}	—	B_{iso}	—	38
0.90	Yes	13928	Q, U_{aniso}	—	Q, B_{iso}	Q, B_{iso}	72
0.60	Yes	47051	Q, U_{aniso}	—	Q, B_{iso}	Q, B_{iso}	72
Enkephalin							
0.90	No	1887	U_{aniso}	U_{aniso}	B_{iso}	—	302
0.56	No	6894	U_{aniso}	U_{aniso}	B_{iso}	—	302
0.90	Yes	1887	Q, U_{aniso}	Q, U_{aniso}	Q, B_{iso}	Q, B_{iso}	572
0.56	Yes	6894	Q, U_{aniso}	Q, U_{aniso}	Q, B_{iso}	Q, B_{iso}	572
RD1							
0.90	No	42525	U_{aniso}	B_{iso}	B_{iso}	—	3711
0.62	No	106652	U_{aniso}	B_{iso}	B_{iso}	—	3711
0.90	Yes	42525	Q, U_{aniso}	Q, B_{iso}	Q, B_{iso}	Q, B_{iso}	5613
0.62	Yes	106652	Q, U_{aniso}	Q, B_{iso}	Q, B_{iso}	Q, B_{iso}	5613

2. Test schemes and data used

2.1. Test schemes

The three main questions studied were the following.

(i) Are bond-electron peaks visible in difference Fourier synthesis maps of about 0.6 Å resolution (§3)?

(ii) Are bond-electron peaks still visible at a lower resolution, namely close to 0.9 Å (§4)?

(ii) How may artificial growth of ADPs be reduced in the high-resolution refinement (§4)?

The resolution of about 0.6 Å is of special interest for macromolecular crystallography as it marks the highest resolution currently obtainable for macromolecular structures. This resolution is not an object of special study for small molecules, where much higher resolutions may be reached. Another resolution limit that has become important for macromolecular crystallography is that of about 0.9 Å; several dozen macromolecular structures refined to near this resolution are now available in the PDB.

The studies of deformation density discussed in §§3 and 4 below followed the same scheme. Firstly, tests with simulated F^{obs} data were performed. These data were obtained as Fourier transforms of electron-density distributions calculated by quantum-chemistry methods for a peptide-like molecule. The calculated structure-factor magnitudes simulated observed ones. (The calculated structure-factor phases could be used in these tests to estimate results.) Next, conclusions derived from the tests with the simulated data were checked in similar tests with experimentally obtained magnitudes. Two experimental data sets presented below (§2.3) were used for this control. Table 1 summarizes the parameters of the test objects.

2.2. Preparation of theoretical density maps and simulated data sets

We started with the theoretical deformation map calculated by the DFT method for the $CH_3-HN-CO-CH_3$ molecule containing the peptide group. All quantum-chemistry calculations were performed with the program *SIESTA* (Sanchez-Portal *et al.*, 1997). Some of the approximations made are described below; detailed descriptions can be found in Sanchez-Portal *et al.* (1997), Artacho *et al.* (1999) and Soler *et al.* (2002).

Exchange and correlation of the Kohn–Sham theory (Kohn & Sham, 1965) are treated with the Generalized Gradient Approximation (GGA) function proposed by Perdew *et al.* (1996). Core electrons were replaced by scalar-relativistic

norm-conserving pseudopotentials generated using the Troullier–Martins scheme (Troullier & Martins, 1991) in their fully non-local formulation (Kleinman & Bylander, 1982). A uniform mesh with a plane-wave cutoff of 75 Ry was used to represent the Hartree and exchange-correlation potentials and the local part of the pseudopotential. Valence electrons were described using a basis set of finite-range numerical atomic orbitals. The radial parts of these orbitals are based on the scheme of pseudo-atomic orbitals of Sankey & Niklewski (1989), who proposed a scheme to build single-z bases. In *SIESTA*, the bases are generalized to arbitrarily complete bases up to double-z with polarization orbitals (Sanchez-Portal *et al.*, 1997; Artacho *et al.*, 1999). The orbital confinement energies used were defined in earlier work (Guillot *et al.*, 2003), where the final set of cutoff values was selected so that the theoretical electron density would be comparable with the experimental density.

The theoretical deformation-density map was defined as the difference

$$\rho_{\rm def}^{\rm theor}(\mathbf{r}) = \rho_{\rm full}^{\rm bond}(\mathbf{r}) - \rho_{\rm full}^{\rm free}(\mathbf{r}) \qquad (3)$$

between the electron-density distribution in the assembly of bonded atoms, $\rho_{\rm full}^{\rm bond}(\mathbf{r})$, and the sum $\rho_{\rm full}^{\rm free}(\mathbf{r})$ of the electron densities of free atoms. Fig. 1(*a*) presents this theoretical deformation map. The main well known visible features of these deformation areas are the following.

(i) A shift of the electron-density peaks from the H atoms to the heavy atoms (in the following, the term 'heavy' means all atoms other than hydrogen and X—H represents the bond between a heavy atom and an H atom).

(ii) Redistribution of the density along the bonds between heavy atoms; for a particular bond, this redistribution reveals itself as a peak in the deformation density near the middle of the bond.

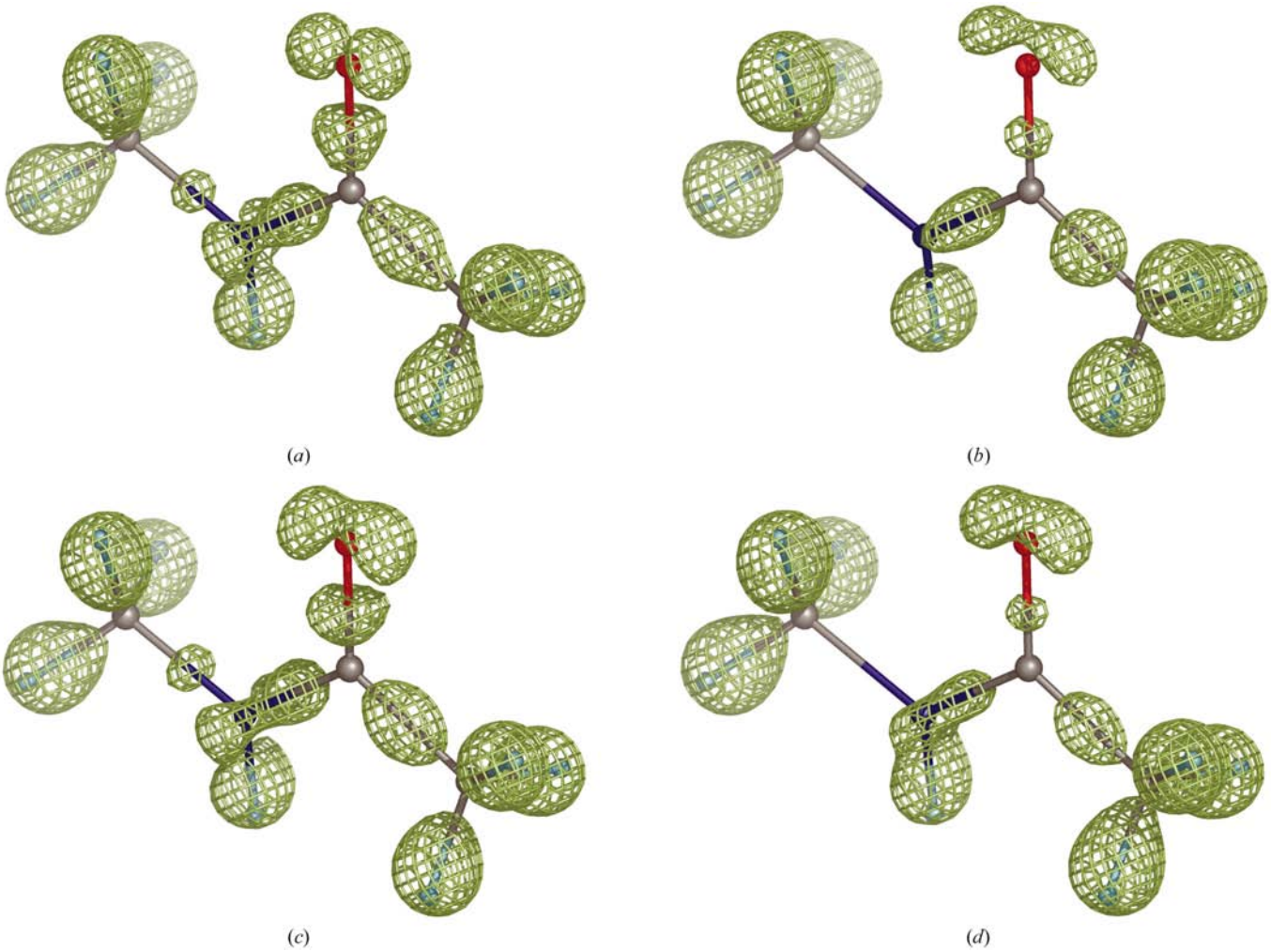

Figure 1
(*a*) The theoretical deformation-density map for the CH_3—HN—CO—CH_3 molecule calculated using the DFT technique (see §2.2 for details). The density is shown at the cutoff level $\rho_{\rm crit} = 0.25$ e Å^{-3}. (*b*) The image of the theoretical deformation-density map obtained as the corresponding Fourier series truncated at $d_{\rm min} = 0.9$ Å; $\rho_{\rm crit} = 0.23$ e Å^{-3}. (*c*) The image of the theoretical deformation-density map obtained as the corresponding Fourier series truncated at $d_{\rm min} = 0.6$ Å; $\rho_{\rm crit} = 0.24$ e Å^{-3}. (*d*) The same image as (*c*), but smeared by isotropic temperature factors with $B_{\rm iso} = 5$ Å^2 ($\rho_{\rm crit} = 0.14$ e Å^{-3}). In the four images the cutoff levels $\rho_{\rm crit}$ are equivalent when converted to the σ scale.

(iii) A deformation density corresponding to lone-pair electrons in the O atom.

To obtain the simulated 'observed' magnitudes, the theoretical density distribution $\rho_{\rm full}^{\rm bond}(\mathbf{r})$ for the molecule $CH_3-HN-CO-CH_3$ was placed in a large enough 'box' ($18 \times 18 \times 15$ Å unit-cell parameters were taken to avoid interactions between the copies linked by unit-cell period translations). The corresponding theoretical structure factors were then calculated as the Fourier transform of the unit-cell content. The theoretical values of the structure factors corresponding to the deformation map were calculated in a similar manner. To simulate isotropic thermal motion with $B_{\rm iso}$ parameters, the calculated magnitudes were multiplied by the factor $\exp(-B_{\rm iso}s^2/4)$.

2.3. Experimental data sets

2.3.1. Enkephalin. The anisotropic model of a trihydrate form of Leu-enkephalin (Tyr-Gly-Gly-Phe-Leu) was initially determined by Aubry *et al.* (1989). This model was used as a starting point to build and refine a multipolar model (Wiest *et al.*, 1994) against very high-resolution data (the formal resolution limit was about 0.44 Å; the data set was practically complete at resolution lower than 0.56 Å). The corresponding crystals in space group $P2_12_12_1$ have unit-cell parameters a = 10.851 (4), b = 13.095 (3), c = 21.192 (4) Å and contain one molecule per asymmetric unit. The multipolar model composed of 43 heavy (non-H) and 43 H atoms includimg three water molecules refined against 5444 reflections has a crystallographic R factor equal to 0.038. In this model, all heavy atoms were considered to be multipolar with anisotropic ADPs, while the displacement factor for the H atoms was isotropic. The experimental data and the model were kindly provided by V. Pichon-Pesme.

This multipolar model was not used directly in our study, but was used to derive a conventional model. To remove the 'memory' about the multipoles (Brünger, 1993), a spherical-atom model was reconstructed from this multipolar model and refined. A zero value was then assigned to all ADP values of this 'relaxed' model and the model was refined at the resolution of 0.56 Å using the *SHELX* program as an anisotropic model. Most of the atoms in the refined structure have an equivalent isotropic B value close to 1 Å^2.

2.3.2. Antifreeze protein RD1. The second experimental data set used was that for the eel pout type III antifreeze protein RD1, the structure of which was solved by Ko *et al.* (2003). This relatively small protein (7 kDa; 64 residues) crystallizes in space group $P2_12_12_1$ with unit-cell parameters a = 32.50, b = 39.50, c = 44.64 Å; the crystals contain one molecule per asymmetric unit. These crystals diffract to 0.62 Å, allowing 118 101 independent reflections to be measured; the data-set completeness in the highest resolution shell, 0.62–0.64 Å, is close to 92%. The model deposited in the PDB was refined to R = 0.133 and $R_{\rm free}$ = 0.155 (5% reflections in the free data set). We were extremely happy to discover that this excellent set of experimental structure-factor magnitudes has been immediately released in the PDB (PDB code 1ucs; Bernstein *et al.*, 1977; Berman *et al.*, 2000).

The antifreeze protein model deposited in the PDB does not contain ADP values. These values were recovered after several cycles of refinement of anisotropic ADP values with *SHELX* at a resolution of 0.62 Å starting from the isotropic values.

2.4. Test refinements

All test refinements were performed using the *SHELX* program (Sheldrick & Schneider, 1997). The coordinates of all atoms were fixed during refinement, *i.e.* only the anisotropic atomic displacement parameters of non-H atoms and isotropic ADPs of the H atoms were normally refined. When using dummy bond electrons (DBEs; see §4.2), the occupancy factors of these atoms were refined as well as their parameters. When DBEs were used, their starting occupancy factor was defined as 0.5 and the starting $B_{\rm iso}$ value for each DBE was taken as the average of the equivalent B values of the atoms forming the corresponding bond. The starting values of the displacement parameters of all atoms were set to zero. The overall scale factor was also considered as a parameter to be refined, although it changed insignificantly during the refinements.

At the low-resolution end, all reflections were always included in the refinement. To refine the antifreeze protein RD1 models, standard *SHELX* bulk-solvent correction was used with the same parameters g = 0.629 and U = 6.09 as reported by Ko *et al.* (2003).

Full-matrix least-squares optimization (corresponding *SHELX* keyword is LS 100) was used in the tests with the $CH_3-NH-CO-CH_3$ model and Konnert–Hendrickson conjugate-gradient optimization (corresponding *SHELX* keyword is CGLS 100) was used in the tests with enkephalin and antifreeze protein RD1. Table 2 gives some refinement statistics. In all refinements 100 cycles of minimization were performed, although in fact the convergence (as indicated by the behaviour of the R factors) required no more than 30–70 cycles. Depending on the refinement protocol, the $X-H$ bond distances were either shortened as suggested by *SHELX* or fixed at standard values (Allen, 1986).

3. Ultrahigh-resolution study of the deformation density in macromolecules

3.1. Finite-resolution images of the theoretical deformation-density map

When studying macromolecular electron density, one usually deals with a map of a Fourier synthesis calculated at some finite resolution $d_{\rm min}$. This resolution cutoff smears density details and produces series-truncation ripples. Such finite-resolution Fourier syntheses are referred to below as images of the corresponding density distribution in order to emphasize this difference. To study series-truncation effects, the complex structure factors corresponding to the theoretical deformation map were calculated as the Fourier transform of $\rho_{\rm def}^{\rm theor}(\mathbf{r})$ density and used further to calculate Fourier syntheses of different resolutions. Figs. 1(*b*) and 1(*c*) present the

images of the theoretical deformation map calculated as Fourier syntheses at resolutions of 0.9 and 0.6 Å, respectively. The features of these finite-resolution maps reflect the general tendency towards image deterioration when decreasing the resolution and agree with the results of previous studies of deformation maps (O'Connell *et al.*, 1966).

An essential feature of macromolecular structures in comparison with that of small molecules is that the former are composed of atoms with larger atomic displacement parameters. This results in an additional smearing of the Fourier images and in a further loss of detail of the density deformation (Fig. 1*d*). Fig. 2 shows the variation of the deformation map image with an increase in the isotropic atomic displacement parameter, which was supposed to be the same for all atoms in these tests. It follows from this analysis that the lowest deformation-density peak (C—N bond) disappears in 0.6 Å resolution maps if the ADP exceeds 4 Å^2. The limiting values for the presence of peaks at N—C(O), C—O and C(O)—C bonds are a little higher, but do not exceed 7–8 Å^2. These values are in agreement with remarks made previously (for example, by Jelsch *et al.*, 1998), where the value 5 Å^2 was pointed out as the limiting value.

3.2. Residual maps with simulated data

In crystallographic study, in contrast to theoretical density analysis, only the structure-factor magnitudes $F^{\rm obs}(\mathbf{s})$ are usually known from experiment, not their phases. The last statement must be made with caution, since in recent years MAD-based phases have been obtained in some advanced macromolecular studies at resolutions of 1.0 Å (Schmidt *et al.*, 2002) and 0.9 Å (Podjarny *et al.*, 2004). However, the accuracy of these phases is as yet far from the accuracy of phases calculated from a refined model. As a consequence, phases $\varphi_1^{\rm calc}(\mathbf{s})$ calculated from an available atomic model are necessary to calculate the image of the distribution $\rho_{\rm full}^{\rm bond}(\mathbf{r})$ as a Fourier synthesis appropriate for our goals. On the other hand, a finite-resolution image of $\rho_{\rm full}^{\rm free}(\mathbf{r})$ in (3) may be presented by a Fourier synthesis calculated with structure-factor magnitudes $F^{\rm calc}(\mathbf{s})$ and phases $\varphi^{\rm calc}(\mathbf{s})$ obtained from

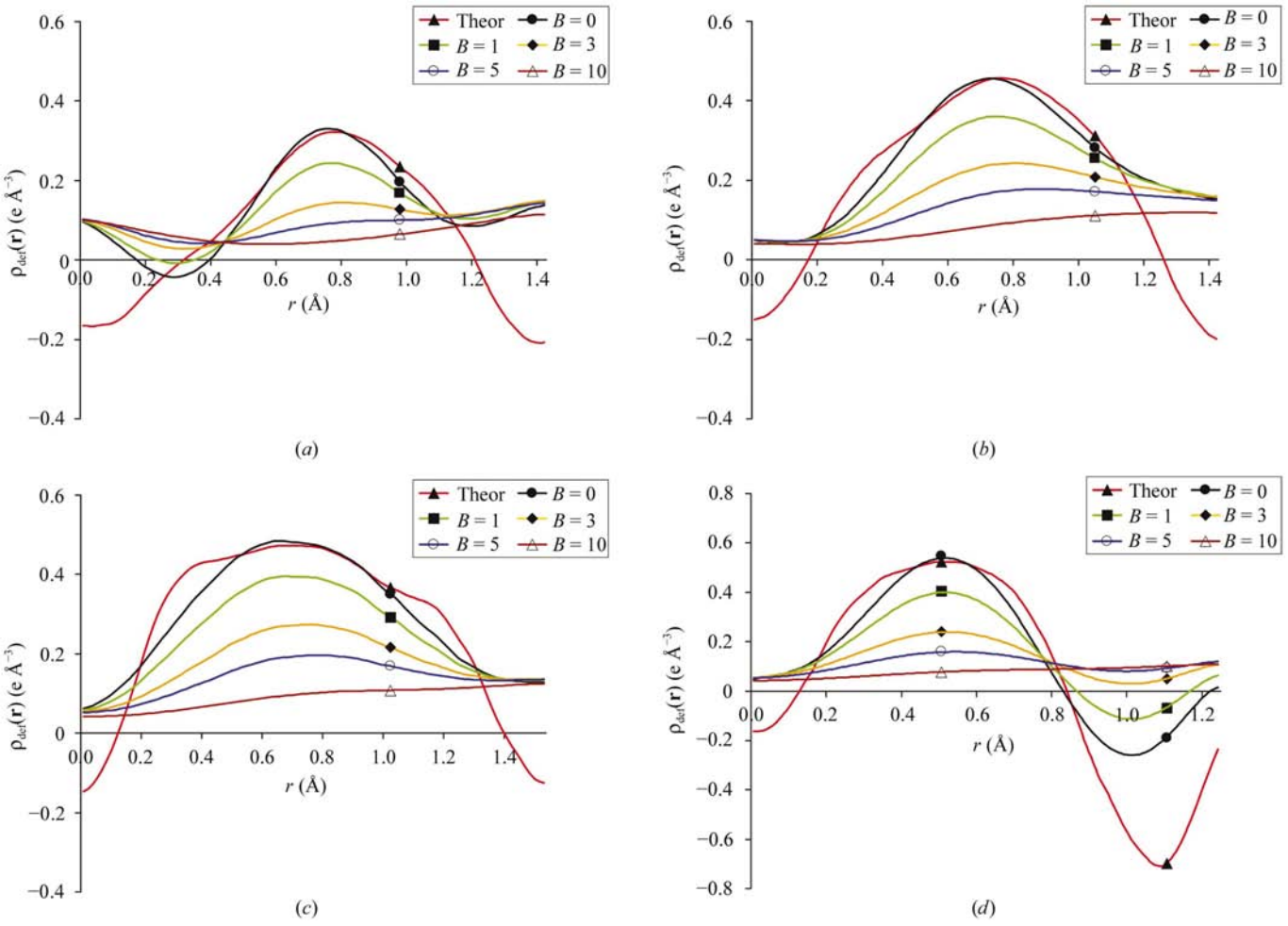

Figure 2
Variation of the electron-density values along the bonds in the theoretical deformation map ('Theor') and in its 0.6 Å resolution image smeared with different $B_{\rm iso}$ values. (*a*) H3C—N, (*b*) N—CO, (*c*) C—CH3, (*d*) C═O.

the available model of the structure under study. As a result, the '1–1' difference map (the residual map) may approximate a finite resolution image of the deformation density

$$\rho_{\rm res}(\mathbf{r}) = \frac{1}{V}\sum_{|\mathbf{s}|<1/d_{\rm min}} \{F^{\rm obs}(\mathbf{s})\exp[i\varphi_1^{\rm calc}(\mathbf{s})] - F^{\rm calc}(\mathbf{s})\exp[i\varphi^{\rm calc}(\mathbf{s})]\}\exp[-2\pi i(\mathbf{s},\mathbf{r})]. \quad (4)$$

In the phase $\varphi_1^{\rm calc}(\mathbf{s})$ we use the index '1' in order to emphasize that, generally speaking, different sets of phases may correspond to the distributions $\rho_{\rm full}^{\rm bond}(\mathbf{r})$ and $\rho_{\rm full}^{\rm free}(\mathbf{r})$ in difference Fourier series. The magnitudes and phases $F^{\rm calc}(\mathbf{s})$, $\varphi^{\rm calc}(\mathbf{s})$ are usually calculated from a model composed of spherically symmetric atoms smeared by ADPs, eventually anisotropic. The phases $\varphi_1^{\rm calc}(\mathbf{s})$ may be calculated from more sophisticated models (*e.g.* multipolar models) and may be used to obtain more accurate estimates of the phases of $\rho_{\rm full}^{\rm bond}(\mathbf{r})$. At the same time, it is well known that the structural details included in the calculation of phases $\varphi_1^{\rm calc}(\mathbf{s})$ reveal themselves in the combined $[F^{\rm obs}(\mathbf{s}), \varphi_1^{\rm calc}(\mathbf{s})]$ syntheses independently of their real presence in the crystal (Main, 1979; Hodel *et al.*, 1992). Therefore, to avoid the appearance of predetermined deformation features in residual maps, we used the same phases $\varphi^{\rm calc}(\mathbf{s})$ in both terms in (4). In other words, in our tests we calculated the residual map as the Fourier series (2).

Calculation of the image of the deformation density $\rho_{\rm def}^{\rm theor}(\mathbf{r})$ in the form (2) instead of (4) causes additional distortions of its image (beside the series termination and thermal motion smearing). These extra distortions may arise from the following.

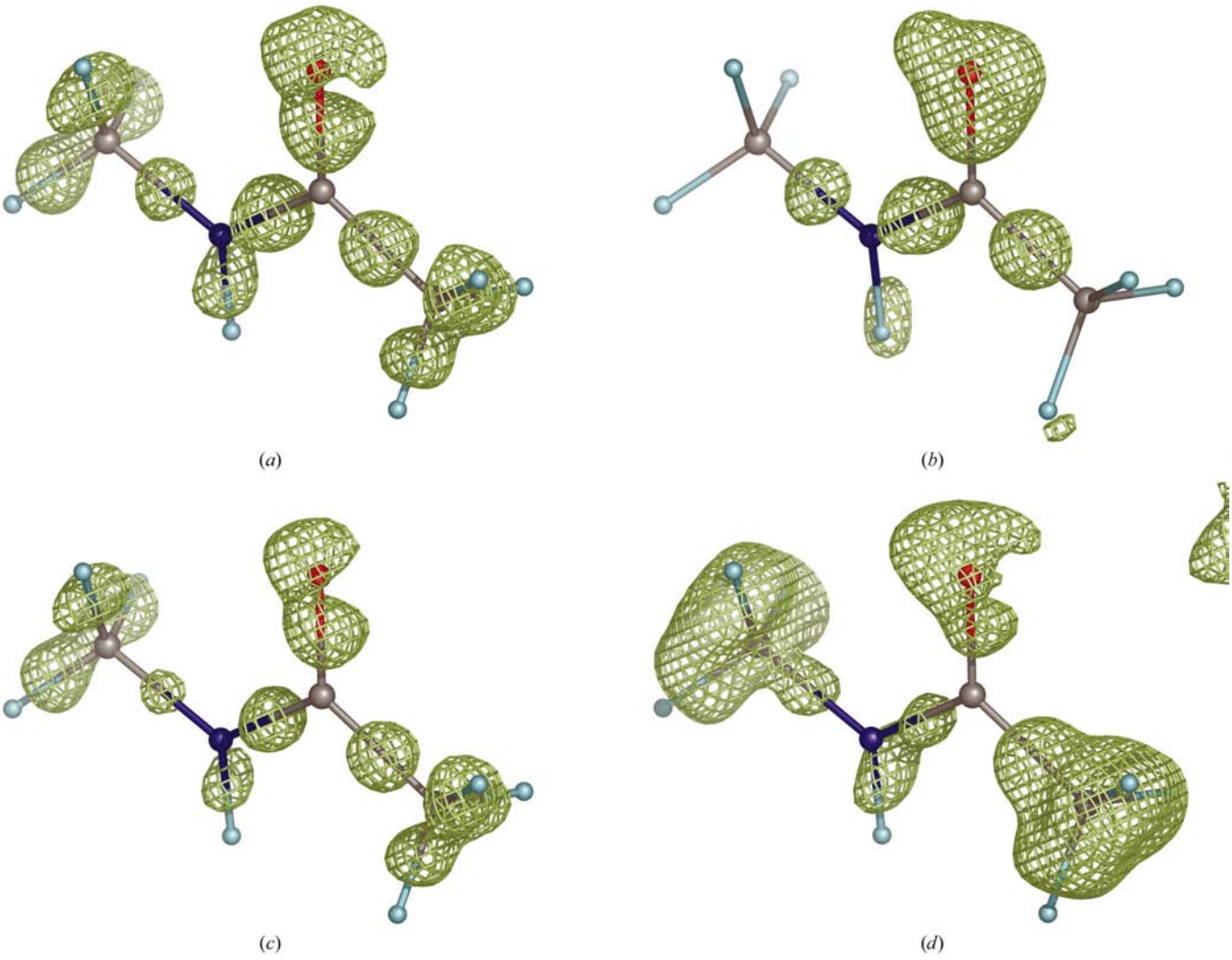

Figure 3
Residual ($F^{\rm obs} - F^{\rm calc}$, $\varphi^{\rm calc}$) maps calculated at a resolution of 0.6 Å with the simulated data. $F^{\rm obs}$ values were calculated from the theoretical density distribution for the $CH_3-HN-CO-CH_3$ molecule obtained by the DFT technique and smeared with isotropic temperature factor $B_{\rm iso}$. $F^{\rm calc}$ and $\varphi^{\rm calc}$ were calculated from the atomic model refined by the *SHELX* program using anisotropic ADPs. Depending on the protocol, the refinement used either the 'standard' $X-$H bond lengths (obtained by neutron diffraction studies) or shortened ones (the default *SHELX* values). (*a*) $B_{\rm iso} = 2$ Å^2, standard $X-$H bonds, $\rho_{\rm crit} = 0.11$ e Å^{-3}; (*b*) $B_{\rm iso} = 2$ Å^2, shortened $X-$H bonds, $\rho_{\rm crit} = 0.086$ e Å^{-3} (the peaks at $X-$H bonds have mostly disappeared); (*c*) $B_{\rm iso} = 5$ Å^2, standard $X-$H bonds, $\rho_{\rm crit} = 0.076$ e Å^{-3}; (*d*) $B_{\rm iso} = 8$ Å^2, standard $X-$H bonds, $\rho_{\rm crit} = 0.014$ e Å^{-3}. The cutoff levels $\rho_{\rm crit}$ in the different parts of the figure correspond to the same value in the σ scale with the exception of Fig. 3(*d*), where it is four times lower (at the equivalent σ level the map does not reveal any density).

(i) Approximation of the free-atom density distribution used to calculate $F^{\rm calc}(\mathbf{s})$, $\varphi^{\rm calc}(\mathbf{s})$; this approximation includes both the substitution of the proper free-atom density by a spherically averaged one and the use of approximate atomic scattering factors (*e.g.* five-gaussians or others; Brown *et al.*, 1999) when calculating the structure factors.

(ii) Errors in the parameters of the model used to calculate the structure factors; in particular, anisotropic ADPs may be overestimated in the refinement (see §4 for discussion).

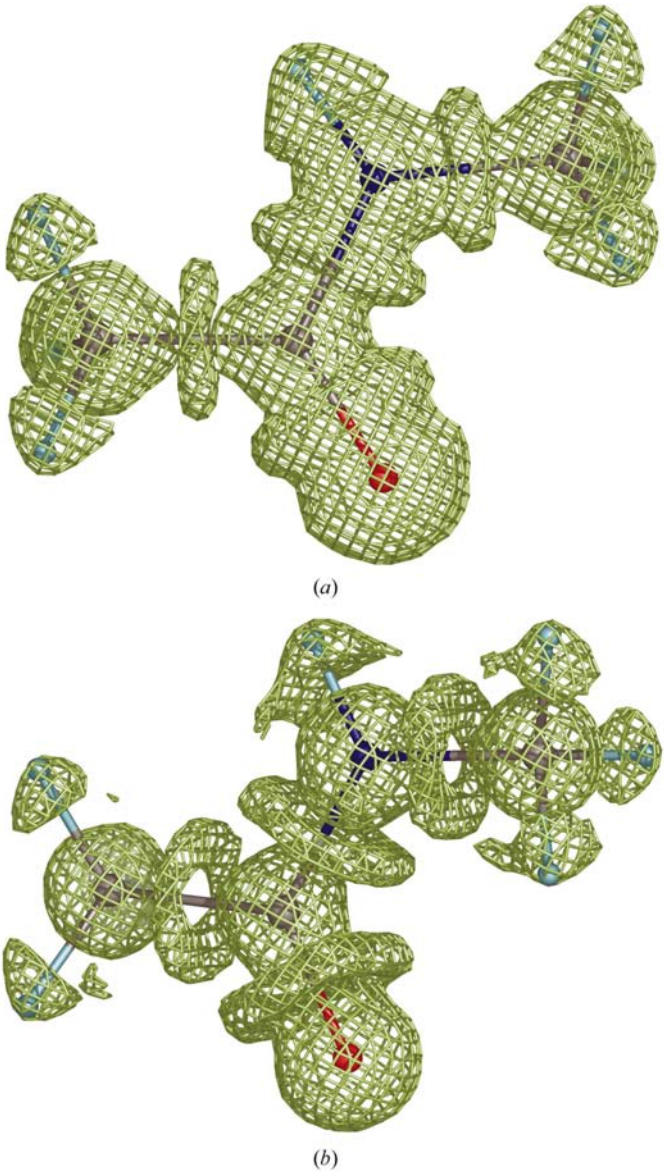

Figure 4
Fourier syntheses ($F^{\rm calc}$, $\varphi^{\rm calc}$) calculated at the resolution of 0.6 Å with the structure factors corresponding to the ideal spherical-atom model. The peaks between the atoms along the bonds appear purely because of Fourier series truncation and have nothing to do with the deformation electron density. The chosen cutoff level $\rho_{\rm crit}$ is significantly higher than the usual level of the deformation density (see Figs. 1 and 2 and the corresponding figure legends for comparison). The two views are slightly different in order to show different aspects of the peaks shape. (*a*) All atoms have isotropic temperature factor $B_{\rm iso}$ = 2 Å^2, $\rho_{\rm crit}$ = 1.0 e Å^{-3}; (*b*) all atoms have isotropic temperature factor $B_{\rm iso}$ = 1 Å^2, $\rho_{\rm crit}$ = 1.25 e Å^{-3}.

(iii) The substitution of $\varphi^{\rm calc}(\mathbf{s})$ values for the true $\varphi^{\rm bond}_{\rm full}(\mathbf{s})$ phases.

To study the influence of these additional errors on the image of the deformation density, we performed the following tests.

(i) The theoretical distribution $\rho^{\rm bond}_{\rm full}(\mathbf{r})$ was calculated by the DFT method (*SIESTA* program; Sanchez-Portal *et al.*, 1997); the corresponding structure-factor magnitudes smeared with an isotropic temperature factor $B_{\rm iso}$ [*i.e.* multiplied by $\exp(-B_{\rm iso}s^2/4)$] were considered in tests as the 'observed' structure factors, namely $F^{\rm obs}(\mathbf{s})$.

(ii) The ADPs of the $CH_3-NH-CO-CH_3$ model were refined against these magnitudes using the *SHELX* program (Sheldrick & Schneider, 1997); some details of refinement protocols are discussed in §2.4 and are given in Table 2.

(iii) The magnitudes and phases $F^{\rm calc}(\mathbf{s})$, $\varphi^{\rm calc}(\mathbf{s})$ calculated from the refined model were used to calculate the residual map (2).

Fig. 3 shows the residual maps corresponding to different values of the overall isotropic temperature factor $B_{\rm iso}$. The quality of these maps is obviously lower than that for the theoretical deformation maps with the same $B_{\rm iso}$ value. Nevertheless, the main features of the density redistribution are still present, being more or less pronounced depending on the resolution and the $B_{\rm iso}$ factor.

The models used to produce Figs. 3(*a*), 3(*c*) and 3(*d*) were obtained by applying the refinement protocol (§2.4) in which the distances between heavy and H atoms are kept equal to the distance determined from neutron diffraction studies (Allen, 1986). The residual maps reflect the features of the theoretical maps shown at Fig. 1. These maps show a strong shift of the hydrogen electron density to the corresponding heavy atom. To take this effect into account, the *SHELX* program includes an option allowing shortening of the $X-$H bonds. Fig. 3(*b*) presents the residual maps obtained under the same conditions as for Fig. 3(*a*), but with the use of the shortened bond lengths as refinement restraints. As could be expected, the shift of the hydrogen centres in the model to the $X-$H bond peaks of the deformation density removes these peaks from the residual map.

3.3. Fourier series-truncation effects in high-resolution density studies

In the two previous sections we studied the influence of the Fourier series truncation on the deformation map and its approximations presented by residual (or '1–1') difference syntheses. Obviously, this type of Fourier synthesis is not the only one used in crystallography and some attempts have been described in the literature to observe deformation density in 'standard' ($F^{\rm obs}$, $\varphi^{\rm calc}$) syntheses or in other 'unbalanced' ($mF^{\rm obs} - nF^{\rm calc}$, $\varphi^{\rm calc}$), $m \neq n$, syntheses. Such a study must be performed with a great caution, as the series-truncation effects in the unbalanced syntheses may be much stronger than the deformation density. It is worthy of note that the peak value of the theoretical deformation density (for $B_{\rm iso}$ = 0) is about 0.5 e Å^{-3}. At the same time, the full density value near the centre of a typical macromolecular atom (C, O, N) may exceed

250 e Å^{-3} and the full density at the middle of the bond is about 2–3 e Å^{-3}. The ripple corresponding to this high density can bury the fine deformation-density image.

The main features of the ripples corresponding to an isolated atom may be described roughly as a 'negative density shell' at the distance about $0.9d_{\min}$ from the atomic centre and a 'positive density shell' at the distance near to $1.4d_{\min}$. The shells caused by different atoms interfere, producing the noise, which can become significant. In protein crystallography necessary attention is not always paid to series-truncation effects. This can be explained by several reasons. Firstly, the thermal motion of atoms attenuates these effects and macromolecular structures generally have higher ADPs than do small molecules. Secondly, for a relatively low resolution the ripple shells have large radii and are mutually cancelled, thus reducing the amplitude of the noise. Thirdly, the main goal of macromolecular studies is usually the search for atomic positions; in other words, the search for a strong signal with a background of a relatively small noise. In small-molecule crystallography, a number of papers have been devoted to the problem of Fourier series truncation; however, most of them study the variation of the atomic shape (for example, Scheringer, 1977) rather than the ripple.

The extension of macromolecular studies into the ultrahigh-resolution range changes the situation. Firstly, small B values in the interval 2–5 Å^2 are not rare for macromolecular crystals diffracting to such a resolution. Secondly, when the resolution increases the ripple shells move closer to the atomic centres and the peaks resulting from ripple-shell superposition appear near the middle of bonds. Finally, details of the density distribution come into the focus of the study rather than the atomic positions themselves. These details, *e.g.* bond electron density, are now comparable in magnitude with the series-truncation ripple and may be buried in the noise if unbalanced difference syntheses such as '2–1' '3–2', '1–2', *etc.* are used. Fig. 4 shows a Fourier synthesis built at a resolution of 0.6 Å with the structure factors calculated from an ideal free spherical-atom model. The series-truncation ripple appearing in this synthesis may be confused with the bond electron density, which obviously does not exist in this case. The '1–1' difference Fourier syntheses ($F^{\rm obs} - F^{\rm calc}$, $\varphi^{\rm calc}$) allows a more reliable study of the density deformation, as now the series-truncation effects are linked to the deformation density itself and are small in comparison with this density.

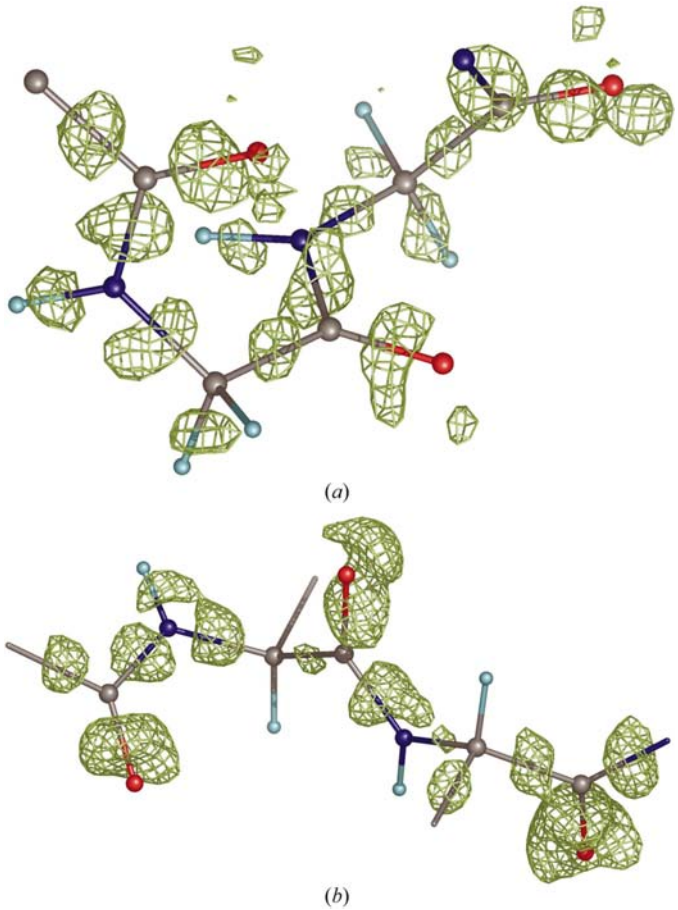

Figure 5
Fragments of residual maps ($F^{\rm obs} - F^{\rm calc}$, $\varphi^{\rm calc}$). Peaks for the bond electron density are clearly seen. The density is shown in the vicinity of the presented atoms only. (*a*) The map for enkephalin calculated at 0.56 Å resolution. Anisotropic ADPs were refined by *SHELX* at the same resolution. The values of equivalent isotropic B parameters are close to 1 Å^2. The standard X—H bond lengths were used in the refinement, therefore peaks for these bonds are mostly present in the map. The cutoff level $\rho_{\rm crit}$ = 0.15 e Å^{-3} (2.5σ). (*b*) The map for antifreeze protein RD1 calculated at 0.62 Å resolution (the main-chain atoms for the residues Val5–Val6 are shown). Anisotropic ADPs were refined by *SHELX* at the same resolution. The values of equivalent isotropic B parameters vary between 3 and 4 Å^2. The shortened X—H bond lengths were used in the refinement, therefore peaks for these bonds are mostly absent. The cutoff level $\rho_{\rm crit}$ = 0.39 e Å^{-3} (3.0σ).

3.4. Residual maps with observed data

The tests with simulated data demonstrated the features of the density deformation that could be seen in residual maps. Two crystal structures (see §2.3) were studied as examples in order to check to what extent these features can be perturbed by experimental errors.

Fig. 5(*a*) presents a fragment of the residual map (2) for

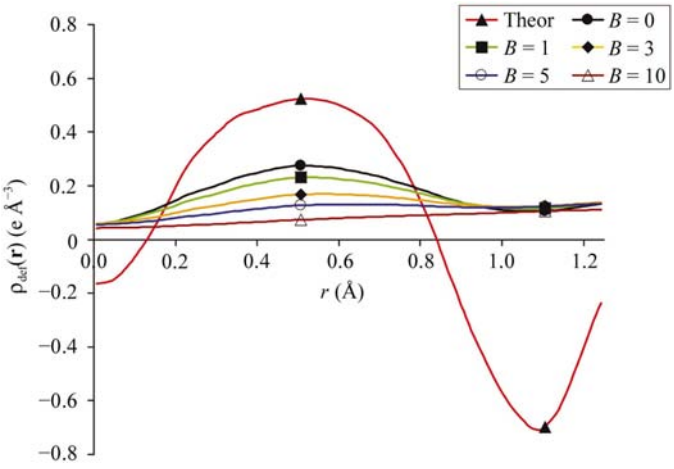

Figure 6
Variation of the electron-density values along the bond C═O in the theoretical deformation map ('Theor') and in its 0.9 Å resolution image smeared with different $B_{\rm iso}$ values.

enkephalin. The peaks corresponding to the bond electron density are obviously seen in this map. Standard X—H bond lengths were used in this refinement and the map clearly demonstrates peaks for these bonds as well as peaks for bonds between heavy atoms. Bond electron density for other parts of the molecule is also observed.

Fig. 5(*b*) shows a fragment of the residual map for the antifreeze protein RD1. Since in our test the refinement of the ADPs was carried out with shortened X—H bonds (default *SHELX* values), the peaks corresponding to X—H bonds are absent in the residual map, similar to the map shown at Fig. 3(*b*).

The two tests performed with the observed data confirm that ultrahigh-resolution difference Fourier syntheses do allow the bond-deformation density for individual bonds to be seen if the conditions found in the tests with the theoretical deformation density are satisfied.

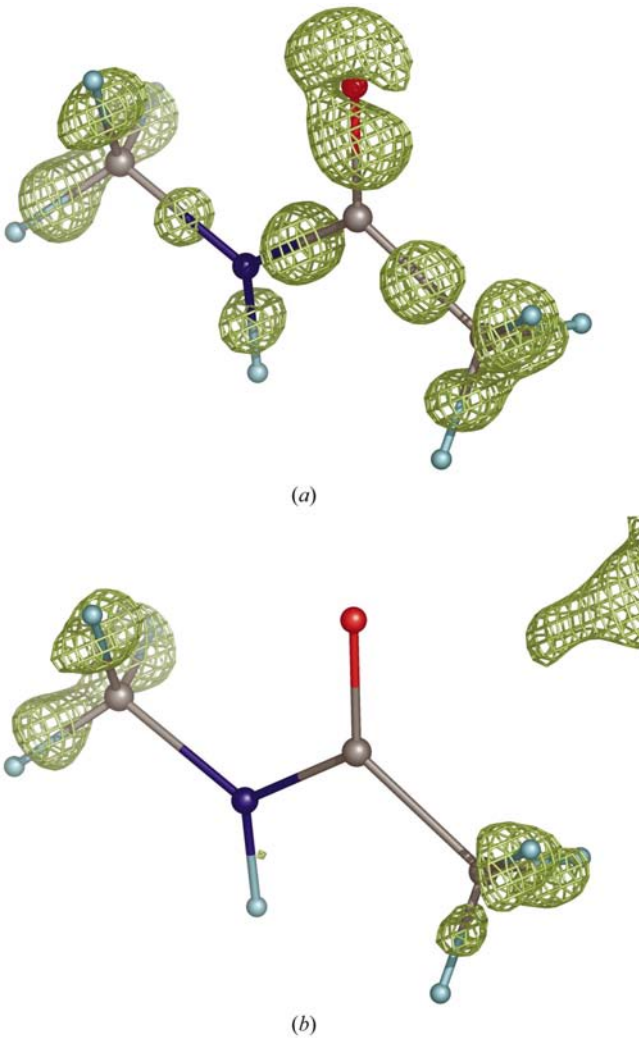

Figure 7
A fragment of the residual ($F^{obs} - F^{calc}$, φ^{calc}) maps calculated at a resolution of 0.9 Å with the simulated CH_3—HN—CO—CH_3 data (B_{iso} = 2 Å^2). Standard X—H bond lengths were used in the refinement. The density is shown in the vicinity of the presented atoms only. (*a*) The anisotropic ADPs were refined by *SHELX* at a resolution of 0.6 Å; ρ_{crit} = 0.11 e Å^{-3}; peaks for the bond electron density are clearly seen. (*b*) The structure was refined with the same protocol but at a resolution of 0.9 Å; ρ_{crit} = 0.02 e Å^{-3} (this is the level of the noise in this map); no peaks for the peptide-bond electron density can be identified in this map.

4. Diversity in the results of high- and ultrahigh-resolution refinements

4.1. High-resolution residual maps

The examples adduced to in the previous section demonstrate that the bond electron density can be observed for well ordered individual residues in conventional crystallographic residual maps (calculated without the use of phases from multipolar models) if the phases used were calculated from the model refined at ultrahigh resolution. A similar observation was made previously for the residual maps for crambin refined at 0.54 Å resolution (Jelsch *et al.*, 2000) and for aldose reductase refined at 0.66 Å (Howard *et al.*, 2004). Nevertheless, we failed to reproduce this effect in residual maps

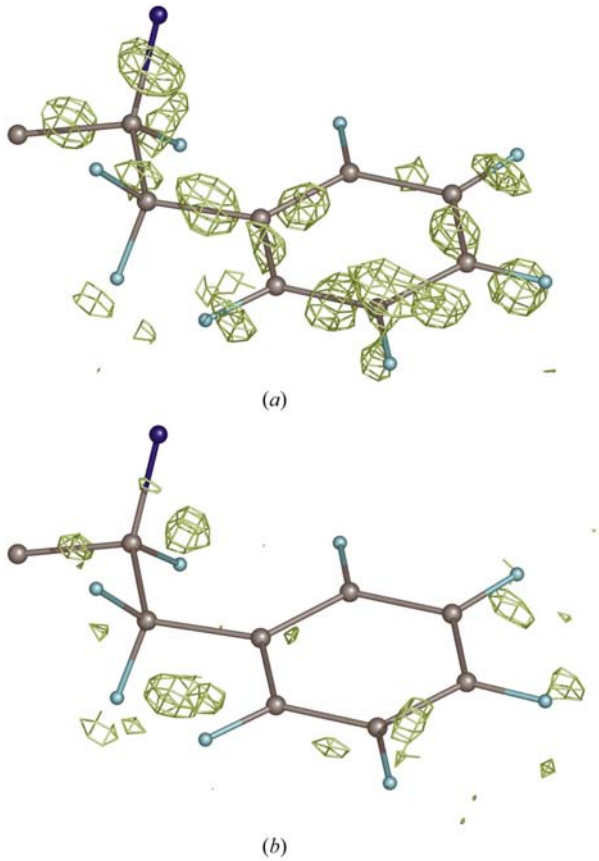

Figure 8
A fragment of the residual ($F^{obs} - F^{calc}$, φ^{calc}) map for enkephalin calculated at a resolution of 0.9 Å. Standard X—H bond lengths were used in the refinement. The density is shown in the vicinity of the presented atoms only. In both cases, the resulting values of the equivalent isotropic B parameter for the most of atoms vary between the limits 1–2 Å^2. (*a*) The anisotropic ADPs were refined by *SHELX* at a resolution of 0.56 Å; ρ_{crit} = 0.06 e Å^{-3} (2.0 σ); peaks for the bond electron density are clearly seen. (*b*) The structure was refined with the same protocol but at the resolution of 0.9 Å; ρ_{crit} = 0.07 e Å^{-3} (2.0σ); peaks for the bond electron density can hardly be identified in this latter map.

Table 3
The refined ADPs and anisotropy parameters for non-H atoms in the $CH_3-NH-CO-CH_3$ molecule.

The refinement was performed at a resolution of 0.6 Å and independently at 0.9 Å with the *SHELX* program against F^{obs} values obtained as the Fourier transform of the theoretical map and smeared by $B_{iso} = 2$ Å^2. Two independent refinements were carried out at each resolution, namely with dummy bond electrons included and not included. The anisotropy parameter (Trueblood *et al.*, 1996; Schneider, 1996; Merritt, 1999) is defined as $A = \lambda_{min}/\lambda_{max}$, where λ_{min} and λ_{max} are the smallest and the largest eigenvalues of the ADP matrix ($A = 1$ corresponds to an isotropic atom as is the case in this model).

Atom	B_{iso} (Å^2)				A			
	0.6 Å	0.6 Å + BE	0.9 Å	0.9 Å + BE	0.6 Å	0.6 Å + BE	0.9 Å	0.9 Å + BE
$C(H_3-N)$	2.18	2.15	2.52	2.17	0.94	0.95	0.89	0.94
N	2.27	2.21	2.63	2.30	0.93	0.94	0.90	0.78
C(O)	2.10	2.07	2.34	1.97	0.96	0.97	0.76	0.93
O	2.09	2.11	2.33	2.20	0.90	0.91	0.82	0.85
$(C-)C(H_3)$	2.19	2.19	2.53	2.33	0.99	0.97	0.96	0.91

calculated for various published structures refined at resolutions close to 0.9 Å. It may be believed that the resolution 0.9 Å is too low to reveal the deformation density. However, residual maps calculated at this resolution for the simulated data (Figs. 6 and 7*a*) as well as for the experimental data of enkephalin (Fig. 8*a*) and of antifreeze protein (Fig. 9*a*) show

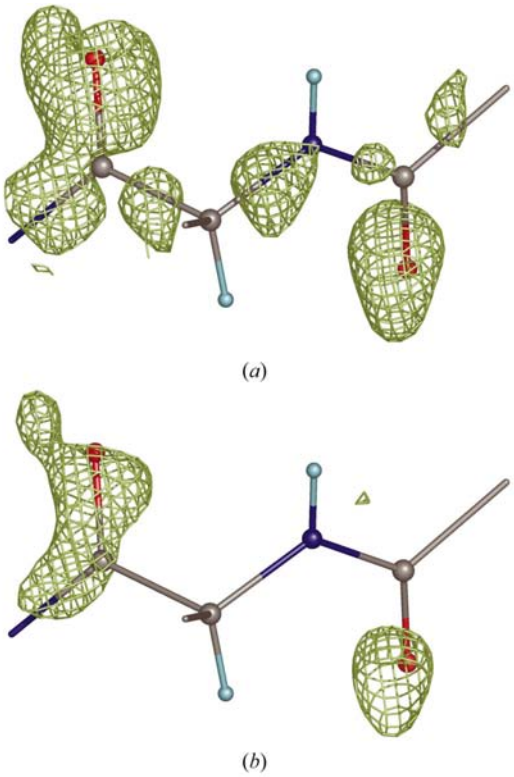

Figure 9
A fragment of the residual map ($F^{obs} - F^{calc}$, φ^{calc}) calculated at a resolution of 0.9 Å for antifreeze protein RD1 around the peptide group of Met56. The refinement was performed with *SHELX*. Shortened $X-H$ bonds were used in the refinement. The density is shown in the vicinity of the presented atoms only. In both cases, the resulting values of equivalent isotropic B parameters vary between 3 and 4 Å^2. (*a*) Refinement of the anisotropic ADPs was performed at a resolution of 0.62 Å; $\rho_{crit} = 0.33$ e Å^{-3} (2.5σ); a number of peaks for the bond electron density are clearly seen. (*b*) The ADPs were refined with the same protocol but at a resolution of 0.9 Å; $\rho_{crit} = 0.27$ e Å^{-3} (2.5σ); peaks for the bond electron density can hardly be identified in this latter map.

such a bond density well enough. Therefore, the resolution cutoff 0.9 Å is not in itself the reason that the deformation density is masked. At the same time, it may be expected that modification of such tiny density details may be caused by small phase errors, as has been discussed previously when studying deformation maps for small molecules (see, for example, Coppens, 1974; Mullen & Scheringer, 1978; Souhassou *et al.*, 1991 and references therein). These phase errors may be caused by inappropriate ADP values. As has been previously pointed out (Coppens, 1967; Housset *et al.*, 2000), artificially increased anisotropic ADPs may extend the density of free atoms to the bond-density region and mask the latter. We performed further tests to check the hypothesis that too large ADPs may be the reason why the deformation density is not always seen in 0.9 Å resolution maps.

As before, we started with tests with the simulated data. Table 3 presents the ADP values for the test model $CH_3-NH-CO-CH_3$ after two independent refinements starting from zero values: one at a resolution of 0.6 Å and the second at a resolution of 0.9 Å. The ADP values after the refinement at 0.9 Å are indeed higher than those obtained after the refinement at 0.6 Å resolution. The same is valid for the degree of anisotropy. (It is worth reiterating that the test object was composed of isotropic atoms.) This means that the inclusion of additional experimental data in the resolution zone $0.6 < d < 0.9$ Å into refinement imposes strong restrictions on the artificial growth of ADPs. In order to check the importance of the 0.6–0.9 Å resolution zone for the appearance of bond peaks, similar tests were performed with the simulated data as well as with the enkephalin and antifreeze protein experimental data.

(i) The ADPs were assigned equal values of 0 and then refined against data of about 0.9 Å resolution; the phases calculated from the refined model were used to construct the residual map at a resolution of 0.9 Å.

(ii) The same starting ADPs were refined independently against data of about 0.6 Å resolution (0.56 Å for enkephalin and 0.62 Å for antifreeze protein); the phases calculated from the refined model were used to the construct a residual map again at a resolution of 0.9 Å.

(iii) For each data set, the two calculated residual maps were compared.

The atomic coordinates and occupancies were fixed in these refinements (§2.4). For each structure, the two refinement protocols were exactly the same, with the only exception being the set of reflections used. Figs. 7, 8 and 9 unambiguously show the difference between the two residual maps obtained using the models refined at 0.9 Å and at approximately 0.6 Å resolution, arguing for the importance of ultrahigh-resolution refinement. Comparison of the refined ADPs shows that refinement at 0.9 Å gives ADPs that are larger (Fig. 10) and have a higher anisotropy than those obtained after refinement at 0.6 Å.

4.2. Dummy bond electrons

The study of images of the deformation density for proteins at the highest resolution available nowadays for macromolecules (about 0.6 Å) shows that at this level of accuracy the deformation map may be reasonably approximated by a set of additional diffractors (sometimes called dummy electrons, bond charges, bond electrons *etc.*) placed at the bonds or at lone-pair density peaks. This idea has been discussed previously for small molecules (some initial references are Brill, 1960; Brill *et al.*, 1971; Coppens & Lehmann, 1976; Hellner, 1977; Pietsch, 1981). Recently, Afonine *et al.* (2002) extended this dummy bond electron (DBE in the following) modelling to macromolecular studies, first illustrating its efficiency for the enkephalin model. The introduction of DBEs into the anisotropically refined model of enkephalin resulted in a decrease in both the R and $R_{\rm free}$ factors and an improvement in the results of the rigid-bond test (Hirshfeld, 1976). Additionally, more sophisticated checks were performed, analysing topological features of the electron density, namely its behaviour at critical points (Bader, 1990). These points can be derived from the multipolar model and used to study molecular properties (see, for example, Souhassou & Blessing, 1999 and references therein). It is known that the electron-density maps calculated from the best available anisotropic models do not reproduce these critical points. At the same time, tests with enkephalin showed that the model with DBEs provides the correct positions of the critical points for most of bonds as well as their correct parameters.

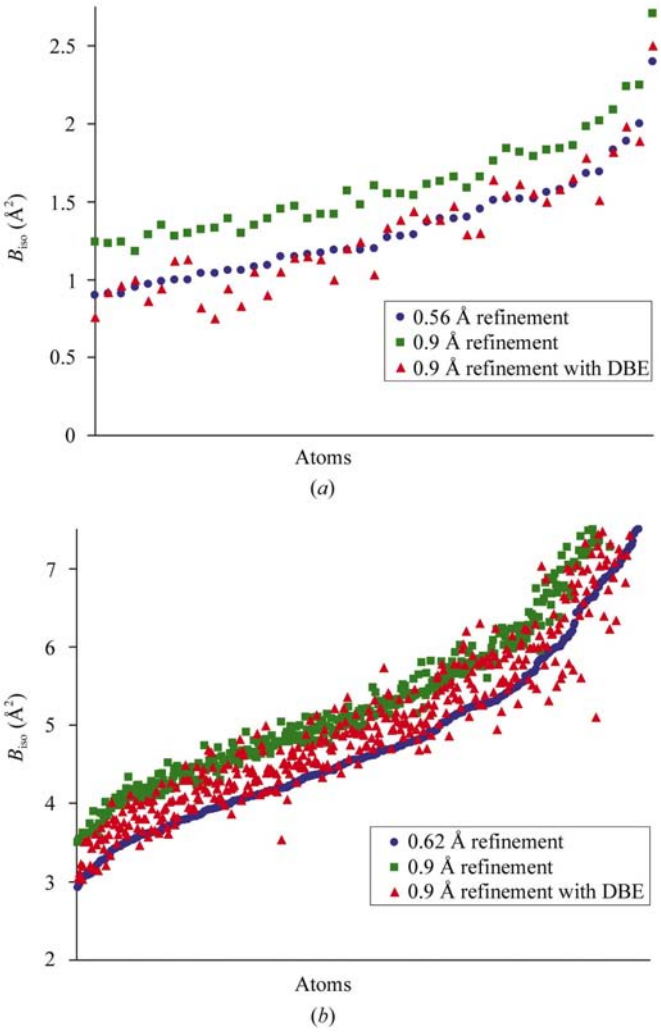

Figure 10
Equivalent isotropic ADPs of the atoms of (*a*) enkephalin and (*b*) antifreeze protein RD1 after conventional refinement at a resolution of 0.56 Å (enkephalin, blue circles), 0.62 Å (RD1, blue circles) and 0.9 Å (both, green squares) and after refinement at a resolution of 0.9 Å with the use of DBEs (both, red triangles). The use of DBEs essentially restricts artificial growth of the ADPs. The atoms are ordered in the diagram with the increase in their $B_{\rm iso}$ values obtained after refinement at the highest resolution. For RD1 only atoms with $B_{\rm iso} < 7.5$ Å^2 are shown.

In the current realisation, each DBE has three parameters: one positional parameter defining its position at the bond with respect to the model atoms (or, for the lone-pair density, following a given direction from a heavy atom) and two parameters of its single-Gaussian scattering curve. In order to estimate these parameters, the theoretical deformation-density maps for the typical bonds in proteins and nucleotides were obtained from DFT calculations using the program *SIESTA* (Sanchez-Portal *et al.*, 1997). Every deformation-density peak at the bond was approximated by a single-Gaussian function

$$\rho(r) = A\exp(-Br^2) \qquad (5)$$

and the corresponding distances between the peak centre and the neighbouring atoms were calculated in order to define the position of the DBE. In reciprocal space, the scattering curve for DBE can be obtained as a Fourier transformation of the corresponding $\rho(r)$,

$$f(s) = a\exp(-bs^2) = A(B/\pi)^{-3/2}\exp\left(-\frac{\pi^2 s^2}{B}\right). \qquad (6)$$

The parameters A, B and a, b thus obtained form the library of DBE parameters for the most characteristic bond types in amino acids and nucleotide fragments. These parameters are available on request from the first author. The parameters a and b can be used directly in the SFAC instruction of *SHELX* in order to define new types of scatterers in the model.

To obtain the mixed model, *i.e.* the model composed from conventional atoms and DBEs, these latter should be placed in the appropriate positions in accordance with the parameter library (for example, at interatomic bonds, at lone-pair elec-

tron peaks *etc.*). It was found that the optimal starting value of the DBE occupancy is 0.5 and that the starting value of its isotropic temperature factors is the mean value of the isotropic equivalents of the neighbouring atoms. A number of different refinement strategies were tried, resulting in a rough conclusion that the best progress may be obtained when refining the following parameters.

(i) For each non-H atom of the model: coordinates, anisotropic ADPs and occupancy.

(ii) For each H atom of the model placed accordingly to ideal stereochemistry (and not to the shortened distances as by *SHELX*): isotropic temperature factor and occupancy.

(iii) For each dummy bond electron (DBE): isotropic temperature factor and occupancy.

Therefore, the number of additional refined parameters added by introduction of DBEs can be estimated as two parameters per atom of the structure, resulting in total in 12 parameters per atom. This can be compared with 9–10 parameters per atom for an anisotropic model (the occupancy may be included in this list for some atoms) and with at least 28 parameters per atom for a multipolar model.

4.3. Use of dummy bond electrons in the refinement of ADPs

Beside its main goal, *i.e.* to be a tool for phase improvement, DBEs may serve as a tool to restrict to some extent abnormal growth of ADPs in refinement performed at a resolution of about 0.9 Å. Indeed, when reproducing a part of the bond density, DBEs may prevent real atoms from attempts to 'grab' this density by increasing their size in the corresponding direction. To check this hypothesis, the refinement of ADPs at 0.9 Å resolution, first for the simulated and then for the observed data, was repeated but now with the use of a mixed density model composed of both real atoms and DBEs (§2.4). Table 3 illustrates the decrease in model ADPs after this introduction of DBEs into refinement. Fig. 10 shows ADPs obtained in similar tests with the observed data for enkephalin and antifreeze protein. As a result of the use of DBEs, the ADP values for enkephalin became practically the same as those obtained after refinement at a resolution at about 0.6 Å. For antifreeze protein, these values also decreased significantly toward the accurate values. The tests with simulated data (where the exact phases are known) allow the calculation of the error for the phases obtained after the refinement either with DBEs included or without them (Fig. 11). It should be stressed that the DBEs were excluded from the model after the refinement and the phases for the comparison were calculated from the 'real' atoms only. The phase difference resulting in the disappearance of bond peaks is quite small, with an average of 1–2°. It is close to that reported previously for multipolar models (Coppens, 1974; Spackman & Byrom, 1997).

In the current work, besides enkephalin, the mixed-model refinement was performed for antifreeze protein RD1 at the

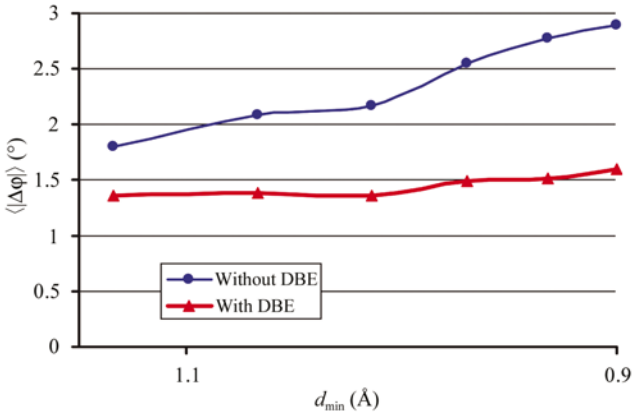

Figure 11
Mean phase error as a function of resolution obtained after the refinement of ADPs for the $CH_3-NH-CO-CH_3$ molecule with the conventional model of spherical atoms (blue) and with dummy bond electrons (red) included. The refinement was performed at a resolution of 0.9 Å. The control structure factors (F^{obs}, φ^{exact}) were calculated from the map obtained by the DFT technique. In both cases (F^{calc}, φ^{calc}) were calculated without using the DBEs.

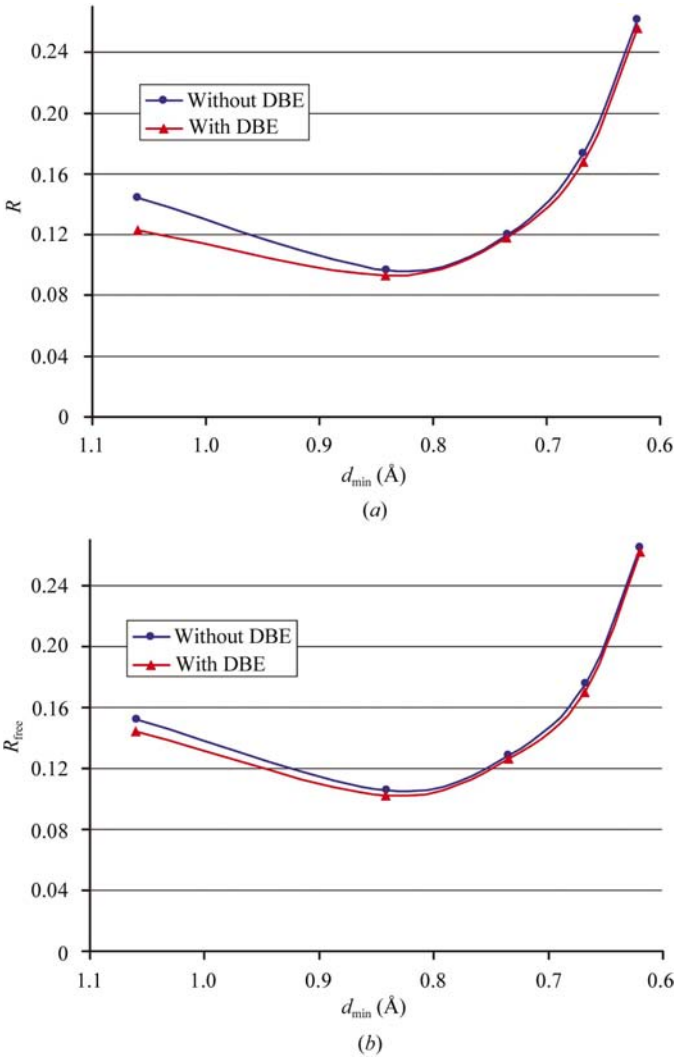

Figure 12
Crystallographic R factors for two models of antifreeze protein RD1 shown as a function of resolution. The models were obtained after anisotropic refinement without DBEs (blue) and after a similar refinement with the DBEs (red). (*a*) R factors; (*b*) R_{free} factors.

resolution of 0.62 Å. The refinement with DBEs lowered both the R and $R_{\rm free}$ factors from 0.137 and 0.155 as reported by the authors (Ko *et al.*, 2003) to 0.125 and 0.141, respectively. Fig. 12 shows the $R_{\rm free}$ factor as a function of resolution. The rigid-bond test for the mixed model did not reveal visible differences in comparison with the results obtained for the conventional refinement.

5. Discussion

Deformation-density studies have been in the focus of small-molecular crystallography for many years. The development of experimental techniques has made it possible to start experimental studies of the deformation density for macromolecular objects. The examples presented above show that in favourable cases (small ADPs and availability of ultrahigh-resolution data), deformation-density peaks become visible in conventional difference syntheses even for individual bonds.

At the same time, a drastic difference must be noted in the possibility of recognizing bond electron density in residual maps obtained as a result of model refinement at a resolution of approximately 0.6 Å or at 0.9 Å. Refinement at 0.9 Å overestimates the atomic displacement parameters. This makes the bond electron density 'invisible' in residual maps. The extension of the resolution zone in the refinement from 0.9 to 0.6 Å essentially restricts this artificial growth of ADPs and results in the appearance of deformation-density peaks both in 0.6 and 0.9 Å resolution residual maps. It is worthy of note that such a drastic difference in maps obtained after high- and ultrahigh-resolution refinement is caused by a quite small difference (mean 1–2°) in phase values.

The use of dummy bond electrons in the refinement restricts the overestimation of ADPs and reduces the $R_{\rm free}$ factor (this is not always the case for refinement when introducing multipolar models; a possible reason may be too large a number of additional parameters in such models) even when the refinement is performed at a resolution close to 0.9 Å. At the same time, some caution is necessary when using phases calculated after refinement performed with the use of DBEs. Some bias in residual maps becomes theoretically possible. This question is the subject of further studies.

The authors thank C. Lecomte for bringing their interest to the problems of crystallographic studies at ultrahigh resolution and for useful discussions. They also thank V. Pichon-Pesme for providing them with the data for enkephalin and G. Sheldrick for useful discussion conserning DBEs. The program *PyMOL* (DeLano, 2002) was used to dysplay the maps. This work was partially supported by the CNRS–RAS collaboration. VYL was supported by RFBR grant 03-04-48155. NM was financially supported by ARC. PA and AU are members of the GdR 2417CNRS. The authors thank the referees and the coeditor for valuable advice that improved the presentation of the results.

References

Afonine, P. V., Pichon-Pesme, V., Muzet, N., Jelsch, C., Lecomte, C. & Urzhumtsev, A. (2002). *CCP4 Newsl. Protein Crystallogr.* **41**. http://www.ccp4.ac.uk/newsletter41/00_contents.html.
Agarwal, R. C. & Isaacs, N. W. (1977). *Proc. Natl Acad. Sci. USA*. **74**, 2835–2829.
Allen, F. (1986). *Acta Cryst.* B**42**, 515–522.
Artacho, E., Sanchez-Portal, D., Ordejon, P., Garcia, E. & Soler, J. M. (1999). *Phys. Status Solidi. B*, **215**, 809–817.
Aubry, A., Birlirakis, N., Sakarellos-Daitsiotis, M., Sakarellos, C. & Marraud, M. (1989). *Biopolymers*, **28**, 27–40.
Bader, R. W. F. (1990). *Atoms in Molecules. A Quantum Theory*. Oxford University Press.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Brill, R. (1960). *Acta Cryst.* **13**, 275–276.
Brill, R., Dietrich, H. & Dierks, H. (1971). *Acta Cryst.* B**27**, 2003–2018.
Brown, P. J., Fox, A. G., Maslen, E. N., O'Keefe, M. A. & Willis, B. T. M. (1999). *International Tables for Crystallography*, Vol. *C*, edited by A. J. C. Wilson & E. Prince, pp. 548–589. Dordrecht: Kluwer Academic Publishers.
Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.
Cochran, W. (1953). *Acta Cryst.* **6**, 260–268.
Coppens, P. (1967). *Science*, **158**, 1577–1579.
Coppens, P. (1974). *Acta Cryst.* B**30**, 255–261.
Coppens, P. (1998). *Acta Cryst.* A**54**, 779–788.
Coppens, P. & Lehmann, M. S. (1976). *Acta Cryst.* B**32**, 1777–1784.
Coppens, P., Su, Z. & Becker, P. J. (1999). *International Tables for Crystallography*, Vol. *C*, edited by A. J. C. Wilson & E. Prince, pp. 706–727. Dordrecht: Kluwer Academic Publishers.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*, http://www.pymol.org.
Ewald, P. P. & Hönl, H. (1936). *Ann. Phys. Leipzig (Germany)*, **25**, 281.
Guillot, B. (2002). PhD Thesis. University of Nancy, France.
Guillot, B., Muzet, N., Artacho, E., Lecomte, C. & Jelsch, C. (2003). *J. Phys. Chem. B*, **107**, 9109–9121.
Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.
Hellner, E. (1977). *Acta Cryst.* B**33**, 3813–3816.
Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239–244.
Hodel, A., Kim, S.-H. & Brünger, A. (1992). *Acta Cryst.* A**48**, 851–858.
Housset, D., Benabicha, F., Pichon-Pesme, V., Jelsch, C., Maierhofer, A., David, S., Fontecilla-Camps, J. C. & Lecomte, C. (2000). *Acta Cryst.* D**56**, 151–160.
Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A. & Podjarny, A. (2004). In the press.
Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* D**54**, 1306–1318.
Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. & Lecomte, C. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3171–3176.
Ko, T.-P., Robinson, H., Gao, Y.-G., Cheng, C.-H. C., DeVries, A. L. & Wang, A. H.-J. (2003). *Biophys. J.* **84**, 1228–1237.
Kleinman, L. & Bylander, D. M. (1982). *Phys. Rev. Lett.* **48**, 1425–1428.
Kohn, W. & Sham, L. J. (1965). *Phys. Rev.* **140**, 1133–1136.
Lamzin, V. S., Morris, R. J., Dauter, Z., Wilson, K. S. & Teeter, M. M. (1999). *J. Biol. Chem.* **274**, 20753–29755.
Lamzin, V. S., Perrakis, A. & Wilson, K. S. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E.

Arnold, pp. 720–722. Dordrecht: Kluwer Academic Publishers.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Lunin, V. Y. & Urzhumtsev, A. (1984). *Acta Cryst.* A**40**, 269–277.
Lunin, V. Yu., Urzhumtsev, A. G., Vernoslova, E. A., Chirgadze, Yu. N., Nevskaya, N. A. & Fomenkova, N. P. (1985). *Acta Cryst.* A**41**, 166–171.
Main, P. (1979). *Acta Cryst.* A**35**, 779–785.
Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1109–1117.
Mullen, D. & Scheringer, C. (1978). *Acta Cryst.* A**34**, 476–477.
Muzet, N., Guillot, B., Jelsch., C., Howard, E. & Lecomte, C. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 8742–8747.
O'Connell, A. M., Rae, A. I. M. & Maslen, E. N. (1966). *Acta Cryst.* **21**, 208–219.
Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). *Phys. Rev. Lett.* **77**, 3865–3868.
Pietsch, U. (1981). *Phys. Status Solidi B*, **103**, 93–100.
Podjarny, A., Schneider, T. R., Cachau, R. E. & Joachimiak, A. (2004). In the press.
Sanchez-Portal, D., Ordejon, P., Artacho, E. & Soler, J. M. (1997). *Int. J. Quant. Chem.* **65**, 453–461.
Sankey, O. F. & Niklewski, D. J. (1989). *Phys. Rev. B*, **40**, 3979–3995.
Scheringer, C. (1977). *Acta Cryst* A**33**, 588–592.
Schmidt, A., Gonzales, A., Morris, R. J., Costabel, M., Alzari, P. M. & Lamzin, V. S. (2002). *Acta Cryst.* D**58**, 1433–1441.
Schneider, T. R. (1996). *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 133–144. Warrington: Daresbury Laboratory.
Schuerman, G. S. & Van Meervelt, L. (2000). *J. Am. Chem. Soc.* **122**, 232–240.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Soler, J. M., Artacho, E., Gale, J. D., Garcia, A., Junquera, J., Ordejon, P. & Sanchez-Portal, D. (2002). *J. Phys. Condens. Matt.* **14**, 2745–2779.
Souhassou, M., Lecomte, C., Blessing, R., Aubry, A., Rohmer, M.-M., Wiest, R., Bénard, M. & Marraud, M. (1991). *Acta Cryst.* B**47**, 253–266.
Souhassou, M. & Blessing, R. (1999). *J. Appl. Cryst* **32**, 210–217.
Spackman, M. A. & Byrom, P. G. (1997). *Acta Cryst.* B**53**, 553–564.
Stewart, R. F. (1969). *J. Chem. Phys.* **51**, 4569–4577.
Teeter, M. M., Roe, S. M. & Heo, N. H. (1993). *J. Mol. Biol.* **230**, 292–311.
Troullier, N. & Martins, J.-L. (1991). *Phys. Rev. B*, **43**, 1993–2006.
Trueblood, K. N., Bürgi, H.-B., Burzlaff, H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., Shmueli, U. & Abrahams, S. C. (1996). *Acta Cryst.* A**52**, 770–781.
Urzhumtsev, A. (1997). *Acta Cryst.* D**53**, 540–543.
Wiest, R., Pichon-Pesme, V., Benard, M. & Lecomte, C. (1994). *J. Phys. Chem.* **98**, 1351–1362.
Wilson, C. & Agard, D. A. (1993). *Acta Cryst.* A**49**, 97–104.