Transferability of Multipole Charge-Density Parameters: Application to Very High Resolution Oligopeptide and Protein Structures

Christian Jelsch, Virginie Pichon-Pesme, Claude Lecomte, André Aubry

▶ To cite this version:

HAL Id: hal-01713052

https://hal.science/hal-01713052

Submitted on 20 Feb 2018

Transferability of Multipole Charge-Density Parameters: Application to Very High Resolution Oligopeptide and Protein Structures

Christian Jelsch,* Virginie Pichon-Pesme, Claude Lecomte and André Aubry

Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques (LCM3B), Université Henri Poincaré, Faculté des Sciences, CNRS UPRESA-7036, BP 239, 54506 Vandoeuvre-lès-Nancy CEDEX, France.
E-mail: jelsch@lcm3b.u-nancy.fr

Abstract

Crystallography at sub-atomic resolution permits the observation and measurement of the non-spherical character of the electron density (parameterized as multipoles) and of the atomic charges. This fine description of the electron density can be extended to structures of lower resolution by applying the notion of transferability of the charge and multipole parameters. A database of such parameters has been built from charge-density analysis of several peptide crystals. The aim of this study is to assess for which X-ray structures the application of transferability is physically meaningful. The charge-density multipole parameters have been transferred and the X-ray structure of a 3_{10} helix octapeptide Ac-Aib$_2$-L-Lys(Bz)-Aib$_2$-L-Lys(Bz)-Aib$_2$-NHMe refined subsequently, for which diffraction data have been collected to a resolution of 0.82 Å at a cryogenic temperature of 100 K. The multipoles transfer resulted in a significant improvement of the crystallographic residual factors wR and wR free. The accumulation of electrons in the covalent bonds and oxygen lone pairs is clearly visible in the deformation electron-density maps at its expected value. The refinement of the charges for nine different atom types led to an additional improvement of the R factor and the refined charges are in good agreement with those of the *AMBER* molecular modelling dictionary. The use of scattering factors calculated from average results of charge-density work gives a negligible shift of the atomic coordinates in the octapeptide but induces a significant change in the temperature factors ($\Delta B \simeq 0.4$ Å^2). Under the spherical atom approximation, the temperature factors are biased as they partly model the deformation electron density. The transfer of the multipoles thus improves the physical meaning of the thermal-displacement parameters. The contribution to the diffraction of the different components of the electron density has also been analyzed. This analysis indicates that the electron-density peaks are well defined in the dynamic deformation maps when the thermal motion of the atoms is moderate (B typically lower than 4 Å^2). In this case, a non-truncated Fourier synthesis of the deformation density requires that the diffraction data are available to a resolution better than 0.9 Å.

1. Introduction

In the usual least-squares treatment of X-ray diffraction data the continuous electron density is subdivided into atomic charge densities and the basic assumptions are that atoms are neutral and of spherical shape with a radial dependence equal to that of free atoms in the gas phase. At sub-atomic resolutions, typically 0.6 Å (or $\sin\theta/\lambda$ greater than 0.8 Å^{-1}), the atomic charge and the non-spherical character of the electron density is observable and can be quantified (Lecomte, 1995; Coppens, 1997). Owing to chemical bonding and atom–atom interactions, the atomic electron density is not spherical and the deformation density can be mapped by accurate low-temperature X-ray diffraction. Taking into account the asphericity of the electron density enables the deconvolution of the thermal-motion parameters U^{ij} and the atomic deformation density. The multipole formalism has been developed in the program *MOLLY* (Hansen & Coppens, 1978)

$$\rho_{\text{atom}}(\mathbf{r}) = \rho_{\text{core}}(\mathbf{r}) + P_{\text{val}}\kappa^3\rho_{\text{val}}(\kappa\mathbf{r})$$
$$+ \sum_{l=0, l_{\text{max}}} \kappa'^3 R_l(\kappa'\mathbf{r}). \sum_{m=0,1} P_{\text{lm}}y_{\text{lm}}(\theta, \varphi) \cdot$$

In this formalism, the total density is decomposed into core electrons and valence electrons. The term ρ_{core} represents the spherically symmetric Hartree–Fock core electron density and the term ρ_{val} the spherically averaged free-atom Hartree–Fock valence electron density. The second term gives an estimate of the atomic charge ($N_{\text{val}} - P_{\text{val}}$), where N_{val} is the number of valence electrons. The C, N and O atoms have for instance two core electrons and 2, 3 and 4 valence electrons, respectively; H atoms have one valence electron.

The third term describes the non-spherical part of the valence electron distribution as a multipole density. The y_{lm} are spherical harmonic functions in real form. The R_l are Slater type radial functions. The θ, φ coordinates are expressed in a local axis system centered on the atom.

This facilitates the application of chemical symmetry constraints as well as the comparison atom by atom. The kappa coefficients (κ and κ') describe the expansion/contraction of the perturbed valence density (Coppens *et al.*, 1979).

The object of this paper is to determine to which structures the results of subatomic resolution crystallography can be extended. What are the criteria to be met by a diffraction data set and the derived crystal structure in order to have observable peaks in the covalent bonds in the deformation electron-density maps? How can the multipole description of the electron density be extended to structures of lower resolution for which not enough diffraction data are measured to refine the parameters describing the deformation density? It is possible in that case to apply the notion of transferability, where the *a priori* knowledge of the multipole density for a given chemical group is used.

A database of experimental electron-density parameters has been built in our laboratory from X-ray analyses of peptide crystals (Pichon-Pesme *et al.*, 1995). At present, the database includes results from ultra high resolution ($d < 0.45$ Å) studies of the amino-acid and peptide crystals: Leu-enkephaline trihydrate Tyr-Gly-Gly-Phe-Leu (Wiest *et al.*, 1994; Pichon-Pesme *et al.*, 1992), *N*-acetyl L-tryptophane methylamide (Souhassou *et al.*, 1991), *N*-acetyl-α,β-dehydro-phenylalanine methylamide (Souhassou *et al.*, 1992), L-arginine phosphate (Espinosa *et al.*, 1996), Tyr-Gly-Gly (Lachekar, 1997), Gly-Asp (Lachekar, 1997), Gly-Gly-Gly (Pichon-Pesme & Lecomte, 1998) and L-histidine phosphate (Mata *et al.*, 1997).

The database analyses have shown that, as we would expect from the quasi-constancy of properties of chemical functional groups, there is a high degree of transferability of the experimental electron-density multipole parameters P_{lm} ($l \neq 0$) for atoms of the same type in similar chemical environments. The P_{lm} multipole coefficients of the database have been successfully transferred (Pichon-Pesme *et al.*, 1995) for a thyrotropin-releasing hormone analogue pGlu-Phe-D-Pro-ψ(CN4)-NMe. Two accurate diffraction data sets have been collected for this pseudo tripeptide, one at room temperature and the other at 125 K, up to 0.65 and 0.75 Å resolution, respectively (Howell *et al.*, 1995). After transfer for both temperature data sets, and with a constant number of variables, the crystallographic wR factor decreases nearly by half. The refinement is physically meaningful as it improves significantly the Hirshfeld rigid bond test (Hirshfeld, 1976), which postulates that the difference of the anisotropic thermal motion ($U^{ij} = B^{ij}/8\pi^2$) in the direction of the chemical bond between two neighbouring atoms should be low ($U < 0.001$ Å^2) because of the rigidity of the covalent bond.

In this paper we analyze the benefits of multipole and charge parameters transferability, *versus* the approx-imation of the spherical atom model, for oligopeptide and protein structures when X-ray data are available at atomic resolution. The transferability of the parameters makes the lower data-to-parameter ratios in protein crystallography manageable.

A crystallographic refinement using aspherical scattering factors has been applied to a 3_{10} helix octapeptide Ac-Aib$_2$-L-Lys(Bz)-Aib$_2$-L-Lys(Bz)-Aib$_2$-NHMe (Toniolo *et al.*, 1995). This terminally blocked helical oligopeptide LBZ is structurally restricted as it contains six Aib residues (α-aminoisobutiric acid or α-methyl alanine) which induce a strong 3_{10} helix bias. The LBZ helix was selected for this study as there is no disorder in the molecular structure and the diffraction data had been collected at cryogenic temperature.

2. Materials and methods

2.1. *Crystal and diffraction data*

The octapeptide LBZ monohydrate has a molecular weight of 1066 g mol^{-1}. It crystallizes in space group $P2_12_12_1$ with unit-cell parameters $a = 11.413$ (2), $b = 13.485$ (2), $c = 38.760$ (2) Å. A well shaped crystal of dimensions $0.3 \times 0.2 \times 0.2$ mm was used for intensity data on a conventional CAD-4F Enraf–Nonius (Delft, The Netherlands) automated diffractometer, using graphite-monochromated Cu $K\alpha$ radiation ($\lambda = 1.54178$ Å). The independent reflections were measured in the θ range 1–70° at cryogenic temperature 115 K, using the Enraf–Nonius model FR558NH liquid nitrogen cryostat. The diffraction data set is complete up
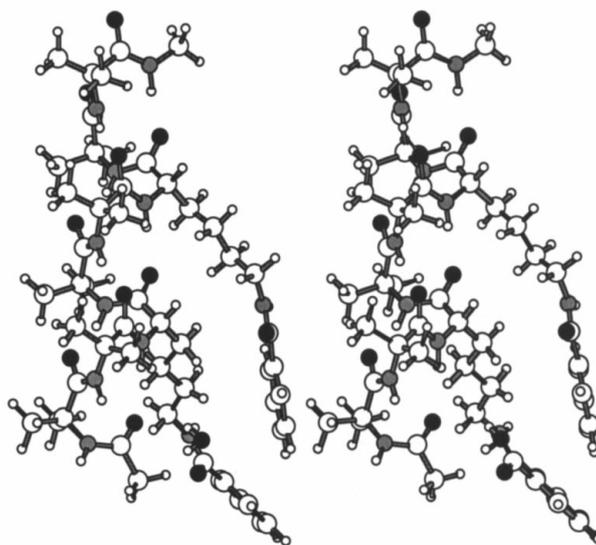
Fig. 1. Stereographic representation fo the LBZ octapeptide helix crystallographic structure. The residues of Ac-Aib$_2$-L-Lys(Bz)-Aib$_2$-L-Lys(Bz)-Aib$_2$-NHMe can be seen along the helix axis from the bottom to the top of the figure. Atom representation: carbon, white; oxygen, grey; nitrogen, black; hydrogen, small circles.

to 0.82 Å resolution. The data had initially been measured for a conventional structure determination and not for charge-density work. Therefore, each reflection was measured only once, the number of independent reflections being 6275. No experimental problem occurred during the data collection.

The structure had initially been refined using the program *SHELXL*93 (Sheldrick, 1993) to an *R* factor of 6.18% (Toniolo *et al.*, 1995). There is one molecule per asymmetric unit. The analysis of the thermal-displacement parameters shows that no atom has an unreasonably elongated thermal ellipsoid, which suggests that there is no disorder in the whole LBZ octapeptide. There is a single water molecule hydrating the LBZ helix in the crystal. This water molecule has a moderate thermal motion ($B_{eq} = 2.3$ Å^2), is ordered and is located in the vicinity of the *N*-acetyl terminus (Fig. 1). A total of 4244 reflections, with a single-to-noise ratio $I/\sigma(I) > 3$, were used for the crystallographic refinement. As the standard deviations of the reflections with high intensities turned out to be strongly underestimated, we applied an intensity-dependent correction of the sigmas $\sigma^2(I) = \sigma^2(I) + (pI)^2$, where $p = 0.01$ can be related to the instrumental instability coefficient (McCandlish *et al.*, 1975).

2.2. *Crystallographic refinement*

For the structure-factor computations, we used the bound-atom form factor for H atoms (Stewart *et al.*, 1965), the form factors calculated from Clementi & Raimondi (1963) wave functions for non-H atoms, and the real and imaginary dispersion corrections to the form factors given by Cromer (1974).

In order to assess the quality of the refinement by computation of the free *wR* factor (Brünger, 1992), 10% of randomly selected reflections were omitted during the refinement. At first, the structure obtained by Toniolo *et al.* (1995) was refined with the *SHELXL*93 program against a complete diffraction data set where a random error of 10% had intentionally been introduced in order to let the structure 'forget' the 10% free reflections. The structure was then refined with *SHELXL*93 using 90% of the reflections with their true measured intensity.

The H atoms were constrained to their ideal position according to the *SHELXL* dictionary of stereochemistry. Three successive refinements I, II and III have been carried out using the *MOLLY* program.

2.2.1. *Refinement I.* The *X*, *Y*, *Z* coordinates and U^{ij} temperature factors were refined with *MOLLY* using the spherical, neutral atom model ($P_{val} = N_{val}$, $P_{lm} = 0$).

2.2.2. *Refinement II.* In this refinement, the multipole parameters P_{lm} from the database (Pichon-Pesme *et al.*, 1995) were transferred to the LBZ structure. Then the H atoms were adjusted by moving the H atoms outward along the C—H and N—H bond directions to bond lengths equal to average values from neutron diffraction

studies (Allen, 1986). The parameters describing the structure were then refined as described below.

(i) *XYZ* U^{ij} refinement (P_{lm} fixed, $P_{val} = N_{val}$).
(ii) Transfer of valence populations P_{val} from the database.
(iii) *XYZ* U^{ij} refinement (P_{lm} and P_{val} fixed).
(iv) P_{val} *XYZ* U^{ij} refinement.
(v) P_{val}, κ, *XYZ* U^{ij} refinement.
(vi) H atoms set to the standard neutron distance.

2.2.3. *Refinement III (kappa refinement).* At the end of the refinement II, in order to evaluate the net atomic charges by crystallographic means, we also performed a (P_{val}, κ) refinement on the final structure, but using the spherical atom model (Coppens *et al.*, 1979). The P_{lm} multipoles were set to zero, while the (P_{val}, κ) variables were refined. During refinement III, the coordinates and temperature factors were kept fixed at their final values of refinement II, as they are the best estimates of these variables.

At the end of refinements II and III, five additional refinement cycles were carried out using a complete diffraction data set including the 10% free reflections in order to obtain a final structure as accurate as possible. During all the refinements, the electroneutrality constraint was applied to the LBZ molecule. Anisotropic temperature factors were assigned to all non-H atoms. Concerning the H atoms, their coordinates were constrained during all the refinements. The temperature factors of the H atoms were kept isotropic and refined but with application of restraints. The target values were 1.2 times (1.5 times for CH$_3$ groups) the equivalent isotropic *B* factors of their neighbouring atom. In the least-squares minimization of the function $E = \sum (F_{obs} - F_{calc})^2/\sigma_F^2 + \sum (B_{restr} - B_{calc})^2/\sigma B_{restr}^2$, the restraints standard deviation was set to $\sigma B_{restr} = 0.1$ Å^2.

2.3. *Atom type definition*

In order to decrease the number of refinable P_{val} and κ parameters, the atoms of similar chemical type were constrained to have the same deformation-density parameters. The nine different atom types used are displayed in Table 1. The non-H atom types are C, O, N (peptide-bond moiety), main-chain Cα, aromatic C, primary C atom CH$_3$ and secondary C atom CH$_2$. The H atoms were differentiated in two types depending on the heavy atom they are bound to: H—N and H—C. The Cα atom type designates simultaneously the tertiary Cα atom found in natural amino acids and the quaternary Cα in residue Aib.

The P_{val} parameters initially transferred, are still preliminary values in our database and were taken from Lachekar (1997). The valence populations P_{val} and the κ contraction/expansion coefficients of the nine atom types were considered refinable parameters, except the κ parameter of the H atoms.

Table 1. *Atomic charges $N_{val} - P_{val}$ and expansion/contraction coefficients κ for chemically similar atoms in the LBZ molecule*

For convenience, the values $N_{\mathrm{val}} - P_{\mathrm{val}}$ rather than the valence populations P_{val} are displayed, as they represent the atomic charges. The standard deviations obtained by full-matrix least-squares refinement with the program *MOLLY* are given in parentheses. The charges from the *AMBER* 4.1 dictionary 'all_amino94.in' (Pearlman *et al.*, 1995; Bayly *et al.*, 1993) are mostly derived from the alanine residue. The aromatic carbon charge is taken from the phenylalanine residue and the secondary carbon CH2 charge from the lysine residue. In the *AMBER* dictionary, the charge of a given atom type takes different values depending on its position on the residue. The charge of a H atom (H—C) also varies slightly depending on the type of C atom to which it is bonded.

Atom	Type	No. atoms	Multipolar model			Spherical model		
			Transfered $N_{\mathrm{val}} - P_{\mathrm{val}}$	Refined $N_{\mathrm{val}} - P_{\mathrm{val}}$	Refined κ	Refined $N_{\mathrm{val}} - P_{\mathrm{val}}$	*AMBER* dictionary	Refined κ
H	(H—N)	11	0.37	0.17 (3)	1.16 fixed	0.20 (2)	0.27	1.16 fixed
	(H—C)	70	0.19	0.09 (2)	1.16 fixed	0.10 (3)	0.06:0.14	1.16 fixed
C	(C═O)	11	−0.01	0.12 (9)	1.04 (2)	0.42 (9)	0.60	1.16 (2)
	(Cα)	11	−0.25	0.02 (8)	1.01 (2)	−0.05 (8)	0.03	1.13 (2)
	Arom	12	−0.15	−0.05 (3)	1.05 (2)	0.01 (3)	−0.17:0.01	1.15 (2)
	(CH₃)	14	−0.15	−0.17 (6)	1.03 (2)	−0.24 (6)	−0.18	1.13 (2)
	(CH₂)	8	−0.15	0.11 (4)	1.02 (1)	0.02 (4)	−0.04:0.02	1.12 (1)
N	(N—H)	11	−0.34	−0.25 (8)	1.00 (2)	−0.44 (8)	−0.42	1.06 (2)
O	(C═O)	11	−0.31	−0.41 (6)	0.97 (1)	−0.65 (6)	−0.57	0.99 (1)

For the transfer of the multipole P_{lm} parameters (Pichon-Pesme *et al.*, 1995), all H atoms were considered to have a same single dipole along the X—H bond. The quaternary Cα atoms from the unnatural amino acid Aib, which were absent in our database, were given the same multipoles as the tertiary Cα from natural amino acids. The list of transferable multipoles is given in Table 2.

2.4. *Electron-density maps*

After crystallographic refinement I with the spherical neutral atom approximation, we computed residual electron-density maps using the Fourier coefficients $(F_{\mathrm{obs}} - F_{\mathrm{sph}}, \varphi_{\mathrm{sph}})$ where the F_{sph} and φ_{sph} are the structure factor and phase calculated from the structure with the spherical neutral atom density model

$$\Delta\rho_{\mathrm{res}}(\mathbf{r}) = V^{-1} \sum_{\mathbf{h}} [k^{-1}|F_{\mathrm{obs}}(\mathbf{h})| - |F_{\mathrm{sph}}(\mathbf{h})|] \\ \times \exp i(\varphi_{\mathrm{sph}} - 2\pi\mathbf{h}.\mathbf{r}) \quad (1)$$

After the multipolar refinement II, the same corresponding residual maps were computed with the coefficients $(F_{\mathrm{obs}} - F_{\mathrm{mult}}, \varphi_{\mathrm{mult}})$.

The dynamic deformation electron-density maps were computed using the following equation,

$$\Delta\rho_{\mathrm{def}}(\mathbf{r}) = V^{-1} \sum_{\mathbf{h}} [k^{-1}|F_{\mathrm{obs}}(\mathbf{h})| \exp(i\varphi_{\mathrm{mult}}) \\ - |F_{\mathrm{sph}}(\mathbf{h})| \exp(i\varphi_{\mathrm{sph}})] \exp(-2i\pi\mathbf{h}.\mathbf{r}) \quad (2)$$

In this case, the φ_{mult} phases were calculated from the structure at the end of refinement II. Spherical atomic scattering factors were used in the calculation of F_{sph} and φ_{sph}. For the calculation of all the terms in equation (2), the coordinates and temperature factors used are derived from the multipolar refinement II, as they are the best estimate of the true parameters.

Static deformation maps, corresponding to the deformation electron density that would be observed for an immobile molecule ($B = 0$ Å^2), can also be computed directly using the XYZ, P_{lm}, P_{val} and κ parameters of the molecular structure.

2.5. *Electrostatic potential*

The electrostatic potential of the isolated molecule has been computed using the program *ELECTROS* (Bouhmaida *et al.*, 1997; Ghermani *et al.*, 1993). The electrostatic potential Φ corresponding to the molecular charge-density distribution is given by the integrated form of the Poisson equation

$$\Phi(\mathbf{r}) = \sum_{\mathrm{atoms}} Z_a/|\mathbf{r} - \mathbf{r}_a| - \int \rho(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'| \mathrm{d}^3\mathbf{r}', \quad (3)$$

where the Z_a and $\mathbf{r}_a$ are the nuclear charges and positions.

3. Results and discussion

3.1. *Structure refinement of the LBZ octapeptide*

A representation of the LBZ helical structure is shown in Fig. 1. No static disorder is found in the crystal structure. After refinement II, the isotropic temperature factor B is in the range 1.0–3.8 Å^2 for the non-H atoms, the average B factor being 1.9 Å^2. The total number of variables refined (XYZ, U^{ij}, κ, P_{val}) is of 785 compared to 4244 reflections used, which results in a ratio of data per variable of 5.4.

The signal-to-noise ratio for the diffraction intensities is shown in Fig. 2. The average $I/\sigma(I)$ is 15 and at the highest resolution its value is 5.2. The percentage of reflections with $I/\sigma(I)$ lower than 3, which were excluded from the refinement, is also shown. The average goodness of fit (g.o.f.), after correction of the intensity errors,

Table 2. *Transfered multipole populations for the different atom types*

Local axes and notations are the same as in Pichon-Pesme *et al.* (1995).

Atom	Type	Transferred multipole parameters		
H	(H—N) (H—C)	$P_{11+} = 0.15$		
C	(C=O) (arom)	$P_{11+} = 0.12$ $P_{20} = -0.23$	$P_{20} = -0.32$ $P_{33-} = -0.32$	$P_{22+} = 0.13$
	(Cα) (CH$_3$)	$P_{31+} = -0.16$	$P_{31-} = -0.21$	$P_{33+} = 0.25$
	(CH$_2$)	$P_{31+} = -0.19$	$P_{31-} = -0.24$	$P_{33+} = 0.21$
N	(N—H)	$P_{33+} = 0.27$		
O	(C=O)	$P_{11-} = -0.10$	$P_{20} = -0.06$	$P_{22+} = -0.06$

is 2.7 at the end of refinement II, indicating that there is still an underestimation of the intensity errors in the reflection data set, especially at low resolution.

The different steps of the multipole refinement of the helix octapeptide and the corresponding wR free and wR factors are displayed in Fig. 3. The first refinement step is conventional and the initial drop of the crystallographic residual is due to the refinement procedure which was performed against the reflection intensities $|F^2_{hkl}|$ in the program *SHELXL*93 and against the structure factors $|F_{hkl}|$ in *MOLLY*. Also the weighting scheme used is different in *SHELXL*93 ($\sigma^2 = k\sigma^2_I + pI^2$, with k and p refined during the least-squares process) and *MOLLY* (σ^2 fixed at its value described in §2.1).

There is an increase of the crystallographic residual factor at each parameter transfer and a subsequent improvement upon positional and temperature-factor refinement. The improvement of both R factors is noticeable after transfer of the multipole parameters, the free R factor decreases from 7.13 to 6.51%. The transfer of the valence populations from the database, on the other hand, did not lead to an improvement of the wR free factor. This might be due to the fact that the P_{val} values were transferred alone without the corre-

Table 3. *Some bond lengths in the LBZ structure after refinement I and II and comparison with the Engh & Huber (1991) dictionary ($\mathring{A} \times 10^3$)*

The standard deviations in the samples are given in parentheses. Cα, tertiary alpha C atom; Cα(Aib), quaternary alpha C atom.

Bond type	No. bonds	Refinement I	Refinement II	Engh & Huber dictionary
C=O	11	1244 (010)	1242 (013)	1231 (020)
C—N	11	1334 (010)	1334 (011)	1341 (016)
C$_{arom}$—C$_{arom}$	12	1386 (012)	1384 (016)	1382 (030)
Cα—N	4	1462 (001)	1455 (001)	1466 (015)
CH$_2$—N	2	1479 (011)	1474 (009)	
Cα(Aib)—N	6	1482 (015)	1481 (018)	
Cα—C	6	1515 (014)	1508 (004)	1525 (021)
Cα(Aib)—CH$_3$	12	1526 (012)	1524 (016)	
CH$_2$—CH$_2$	6	1529 (005)	1529 (013)	1520 (030)
Cα—CH$_2$	2	1531 (006)	1530 (004)	1530 (020)
Cα(Aib)—C	6	1537 (018)	1542 (020)	

sponding κ parameters. The refinement of the valence populations P_{val} alone, and in conjunction with the κ parameters resulted in a modest gain (6.53 → 6.36) in the wR free factor.

3.2. Stereochemistry

The octapeptide is folded in a regular right-handed helix with average absolute values for the φ, ψ backbone torsion angles equal to 56 and 28°, very close to those expected for a 3_{10} helix ($\varphi = 57$, $\psi = 30°$).

The bond lengths of the LBZ structure are listed in Table 3. After spherical and multipole structure refinements I and II, which have been carried out without application of distance restraints for the non-H atoms, the bond lengths are similar within the standard deviation. Also the bond lengths are within the expected range of the Engh & Huber (1991) dictionary. The root-mean-square deviation between the structures obtained from the two refinements is 0.0017 Å, which implies that

Fig. 2. Average $I/\sigma(I)$ ratio, percentage of reflections with $I/\sigma(I) < 3$ and goodness of fit (g.o.f.) of the diffraction data as a function of the resolution. g.o.f. $= [(N_{obs} - N_{var})^{-1} \sum (F_{obs} - kF_{calc})^2/\sigma^2(F_{obs})]^{1/2}$.
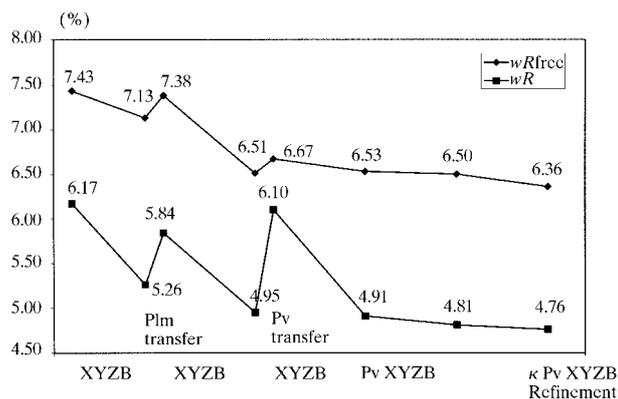
Fig. 3. Evolution of the crystallographic residual factors wR and wR free during the refinement. The variables refined are shown under the x axis. $wR = [\sum w(F_{obs} - kF_{calc})^2 / \sum wF^2_{obs}]^{1/2}$; $w = 1/\sigma^2(F_{obs})$.

the bond lengths are very similar. The main effect of the multipole transfer and charge refinement in the LBZ structure shows up in the thermal-motion parameters rather than in the coordinates, as will be seen later.

3.3. Atomic charges

3.3.1. Atomic charges in refinement II. The concept of atomic charge is fundamental to chemistry as it gives powerful insights towards the understanding of chemical reactivity and physical properties. There is however no clear definition of the true atomic charge, given the distributed nature of the electrons around the nuclei in a molecule. The $N_{val} - P_{val}$ values can be considered as atomic charges obtained by crystallographic fitting in the reciprocal space. The values of the valence populations transferred from the database and after refinement II with *MOLLY* are given in Table 1. The $(N_{val} - P_{val})$ values in the case of refinement II cannot be considered as the total atomic charges as the multipoles contribute also to the electron transfers from one atom to the other.

The transferred and refined P_{val} parameters agree moderately as the correlation coefficient reaches 73%. The O and N atoms have negative $N_{val} - P_{val}$ charges as expected. The two H-atom types have positive charges, the values obtained in our refinement are, however, lower than those transferred from the database. For the C atoms, the transferred and refined $(N_{val} - P_{val})$ in Table 1 are not always in good agreement as they are sometimes of opposite signs. In the case of the C atom, as there are four electrons involved in the covalent bonds, the impact of the multipoles on the total charge is the highest.

We have attempted to define more atom types, for instance differentiating the tertiary and quaternary Cα atoms. But as there were only two tertiary Cα atoms, the refined charge turned out to be physically unrealistic and with a high estimated error. Similarly, distinguishing the H atoms, according to the type of C atom they are bound to, lead to unrealistic values. The refinement of individual atomic charges, without application of several equality constraints over equivalent atoms, would not have been meaningful for this 0.82 Å resolution structure. We have thus considered only nine different atoms types for a total number of 156 atoms in the LBZ molecule (Table 1). The averaging of the parameters through application of constaints on chemically equivalent atoms, rendered the P_{val} and κ variables refinement relevant.

3.3.2. Atomic charges in refinement III. In order to obtain a better estimation of the atomic charges, unbiased by the multipoles, the valence populations P_{val} have also been refined, under the spherical atom approximation (Coppens *et al.*, 1979). The refined charges with the spherical model, depend however partly on the transferred multipoles, as the coordinates and temperature factors used are those after refinement

II. The transferable multipoles describing the peptide main chain are the best determined in our database as they are the most represented in our peptide sample. The charges obtained in refinement III are generally meaningful and agree well with the charges used in the *AMBER* molecular modelling software (Pearlman *et al.*, 1995). The point charges in the *AMBER* dictionary, have been obtained by a restrained electrostatic potential fit, where the potential is calculated from the quantum mechanical wavefunction using a medium-sized 6-31G* basis set (Bayly *et al.*, 1993). The agreement is especially

(a)

(b)
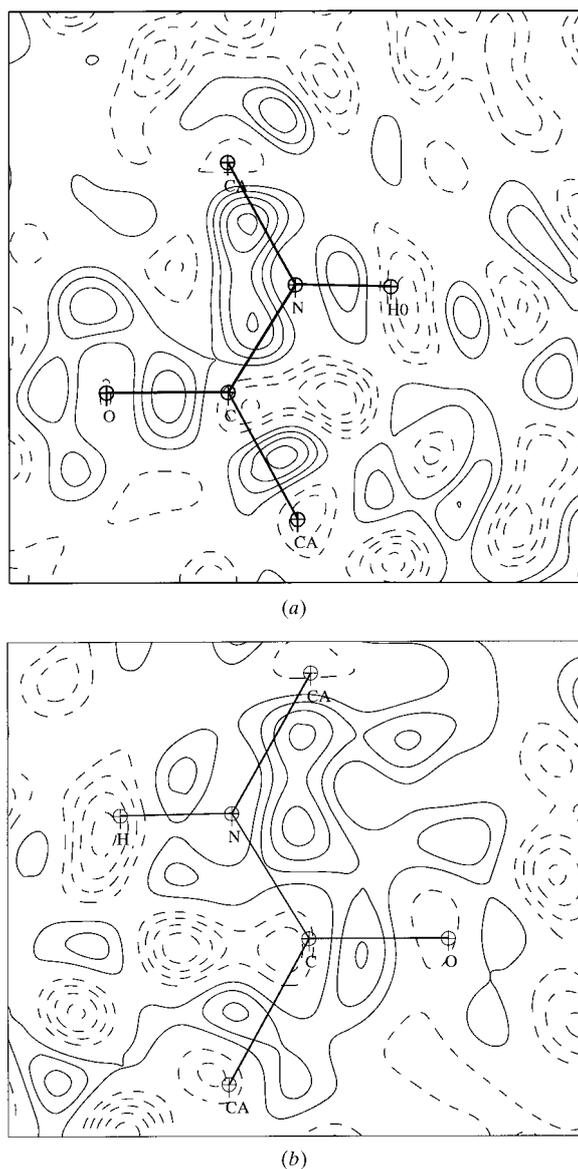
Fig. 4. Residual electron-density maps in the plane of the peptide bond Aib²-Lys(Bz)³. (*a*) after refinement I using the spherical atom model, (*b*) after refinement II using the multipole–P_{val} atom model. Contours: ±0.05 e Å⁻³, positive contours: solid line, negative contours: dashed line, zero contour omitted.

good for the atoms of the peptide-bond moiety O=C—NH and the C atom CH_3 which have the highest charges in absolute value.

It has to be reminded that a consensus charge cannot be found for a given atom type in all amino acids. The atomic charges of a molecule are interdependent as they have to match the electroneutrality constraint and can vary slightly with the conformation of the molecule (Bayly *et al.*, 1993). Also the charge for a given atom type can vary with the location of the atom within the amino-acid residue, as it is the case in the *AMBER* dictionary (Table 1). In the construction of our database, we have indeed observed a lower degree of transferability for the valence populations than for the P_{lm} multipole parameters (Lachekar, 1997).

The refined κ values are also shown in Table 1. After refinement II, the negatively charged O atom has an expanded valence density ($\kappa < 1$) whereas the positively charged carbonyl C atom has a contracted density ($\kappa > 1$) as expected. For the other atom types, the refined κ coefficients have less meaningful values. After refinement III, all the κ values are higher and overestimated, specially for the N and CH_3 atom types. Hence, these atoms are negatively charged and one would expect an extension of the radial electron density. Two comments can be made about Table 1. At first the set of (P_{val}, κ) parameters derived from the multipolar refinement II is more consistent than that obtained from the kappa refinement (refinement III). This may be due to the better valence electron density modelling and improved phase estimate (see §3.6) in the multipolar case. Secondly, the ($P_{val} - N_{val}$) values obtained from refinement III agree better with the *AMBER* dictionary point charges as a spherical atom model is considered.

3.4. Residual electron-density maps

According to equation (1), the cleaner these residual maps are, the better the scattering factor model. After refinement I, the electron-density residual maps show some non-random density, located at the site of the covalent bonds. In the plane of peptide bond Aib[2]-Lys(Bz)[3] these peaks reach 0.3 e Å^{-3} (Fig. 4a). However, these peaks are flatter than what is expected for a covalent bond density. This is because the least-squares fit of the XYZ and U^{ij} parameters takes into account the bonding density by overestimating the U^{ij} parameters (Coppens, 1967). The bias introduced by the spherical atom approximation is in a way compensated by the too powerful least-squares fit of the six U^{ij} parameters per atom. These peaks appear reduced and more randomly distributed in the residual density maps when the multipole–P_{val} atom model was used (Fig. 4b).

3.5. Deformation electron-density maps

3.5.1. *Quality of the maps*. An accumulation of valence electrons in the covalent bonds is clearly visible in the

dynamic electron-density deformation maps (Figs. 5 and 6a). For both the peptide bond Aib[2]-Lys(Bz)[3] and the phenyl ring of residue Lys(Bz)[6] the bonding densities are well defined and all the bond peaks appear at the right place on the bonds. Table 4 gives the peak heights averaged over the 11 amide bonds, and over the two phenyl rings, as well as the corresponding averaged values observed in three charge-density studies of peptides at cryogenic temperature used for building the parameter database (Pichon-Pesme *et al.*, 1995). On average, the bond peaks in the dynamic deformation density observed for the LBZ molecule, reach 85% of the value observed for these peptides. The density peaks on the carbonyl bonds C=O are the highest reaching 0.5 e Å^{-3}. The peaks on the N—Cα bonds are higher in the LBZ molecule than the value averaged over the three peptides, this might be due to an overestimation of the corresponding multipole parameters in the database.

A static deformation electron-density map has also been computed using the multipole parameters P_{lm} transferred from the database and the refined valence populations (Fig. 7). This map represents mainly the transferred charge-density model used. The dynamic deformation density peaks reach, on average, 52% of the values found in the corresponding static deformation maps (Table 4). This comparison of the dynamic and static maps indicates the level of attenuation of the deformation density due to the thermal smearing. The two oxygen electron lone pairs are clearly distinguishable in the static deformation map, but appear as two distinct peaks in the dynamic deformation maps of the LBZ molecule only in four cases out of 11. For the peptide bond 2–3, the oxygen lone pairs appear as an elongated density connecting the two electron pairs (Fig.
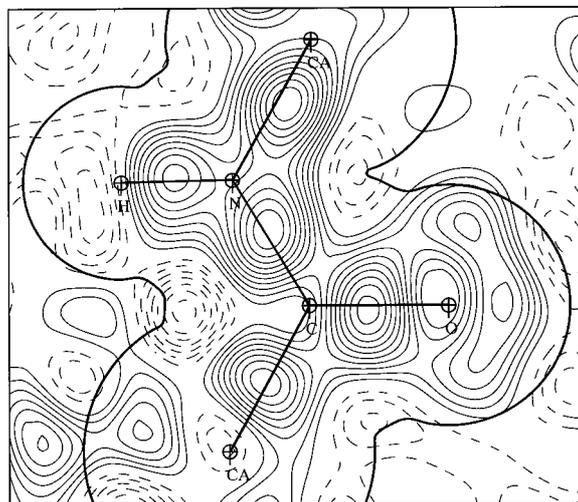
Fig. 5. Dynamic deformation electron-density map in the plane of the peptide bond Aib[2]-Lys(Bz)[3]. Contours: same as in Fig. 4. The thick line represents a contracted van der Waals surface of the molecule (atomic radii multiplied by 0.7).

5). The oxygen lone pairs are the finest feature in the non-spherical electron density of peptides; they require a good description by Fourier synthesis data up to 0.5 Å resolution.

3.5.2. *Effect of a high- and low-resolution cutoff.* The effect of high-resolution diffraction data truncation on the dynamic deformation maps has been analysed on the aromatic ring of residue Lys(Bz)[6]. This phenyl ring has moderate temperature factors ($1.25 < B < 2.5$ Å^2) and the deformation density has a regular shape (Fig. 6a). This dynamic deformation density computed by Fourier synthesis is nearly the same when using all the reflections or when truncating the data at 0.9 Å resolution. However, when a 1 Å resolution cutoff is applied, a lot

of information has been lost (Figs. 6b and 6c). The covalent-bond density peaks decrease by 0.1 e Å^{-3} on the benzene ring of residue LysBz[6] when the 1 Å resolution cutoff is applied.

Low-resolution cutoffs were also applied in order to determine the resolution range which is important for deformation-density parameters. The corresponding dynamic deformation density maps were unaltered for the 2.5 and 2 Å cutoffs and largely attenuated at 1.5 Å (data not shown). The contribution of the very low resolution diffraction data to the deformation density is limited as the number of reflections involved is small. These results show that the essential information on the deformation electron density is contained in the reflec-

(a)

(b)

(c)
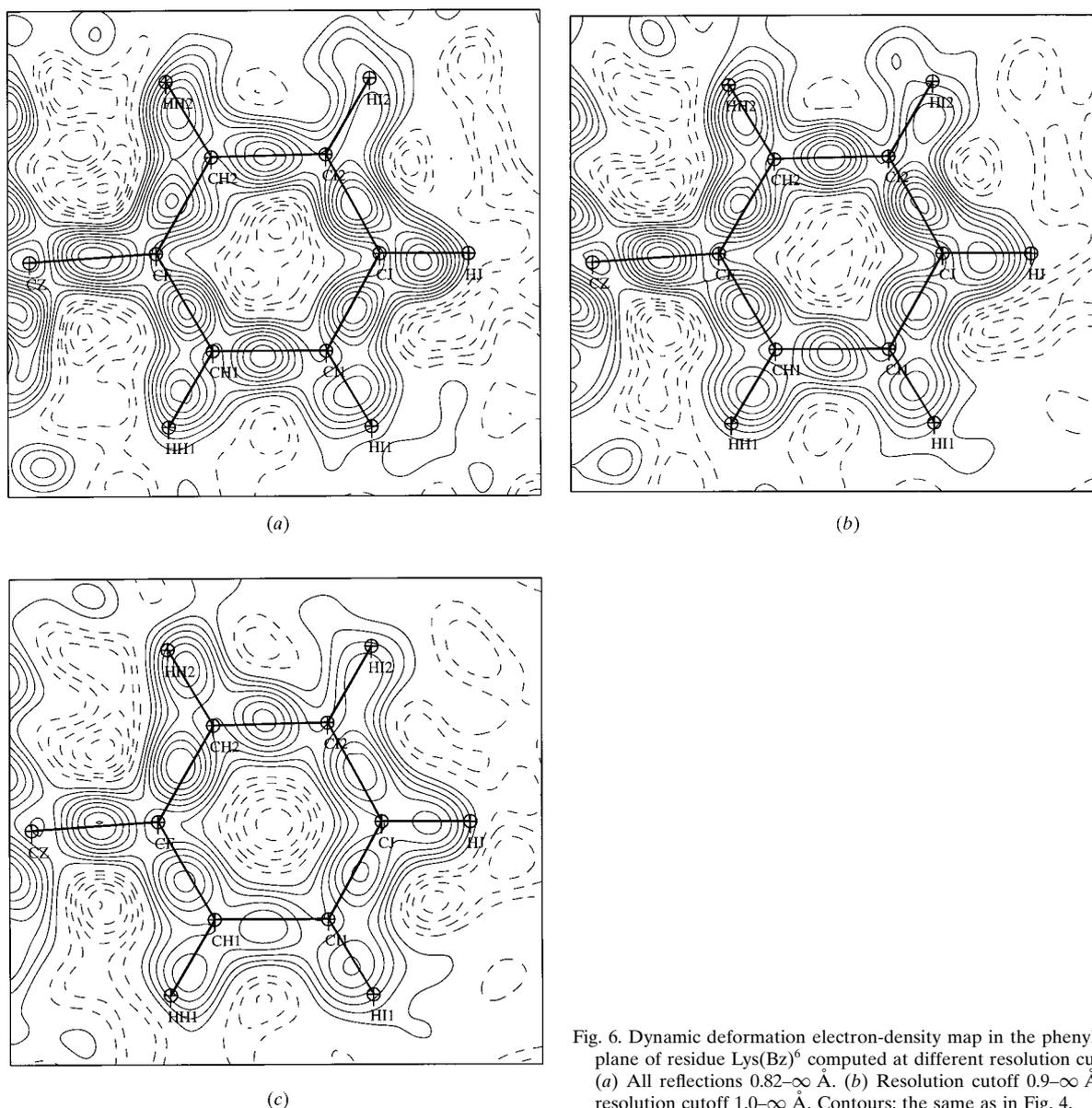
Fig. 6. Dynamic deformation electron-density map in the phenyl ring plane of residue Lys(Bz)[6] computed at different resolution cutoffs. (a) All reflections 0.82–∞ Å. (b) Resolution cutoff 0.9–∞ Å. (c) resolution cutoff 1.0–∞ Å. Contours: the same as in Fig. 4.

Table 4. *Selected bond peaks (e Å^{-3}) in the dynamic and static deformation electron-density maps of the LBZ helix and average values found in dynamic deformation maps for three peptide structures*

Peptide structures: *N*-acetyl L-tryptophane methylamide (Souhassou *et al.*, 1991), *N*-acetyl-($\alpha\beta$)-dehydro-phenylalanine methylamide (Souhassou *et al.*, 1992) and leuenkephaline trihydrate (Wiest *et al.*, 1994).

LBZ	C—N	N—H	N—Cα	Cα—C	C=O	O lone pairs	C$_{arom}$—H	C$_{arom}$—C$_{arom}$
Average	0.37	0.38	0.42	0.39	0.49	0.26	0.33	0.35
R.m.s. deviation	0.12	0.06	0.07	0.09	0.08	0.06	0.07	0.06
Static	0.85	0.75	0.75	0.6	1.05	0.45	0.65	0.75
Average in Enk, AcTrp, AcPhe	0.51	0.46	0.36	0.4	0.51	0.38	0.39	0.56

tions between 0.9 and 2 Å resolution, a region of the reciprocal space where the contribution of the solvent molecules to the X-ray diffraction is small in the case of protein crystals.

3.5.3. *Deformation density and thermal motion.* The influence of the atomic thermal motion on the height of the deformation density has also been analyzed. This effect can be readily seen in the dynamic deformation density of the aromatic ring for the residue Lys(Bz)^3 (Fig. 8*a*). The atomic *B* factors for the residue Lys(Bz)^3 increase gradually along this elongated side chain from 1.5 Å^2 to higher values close to 4 Å^2 (Fig. 8*b*). This aromatic ring has the highest thermal-displacement parameters in the helix octapeptide. Despite their chemical equivalence, the dynamic C—C bonding densities decrease from 0.45 to 0.25 e Å^{-3} as the thermal-displacement parameters increase. The deformation density is well defined on the three bonds formed by the atom CF and its peak value reaches 0.4–0.5 e Å^{-3}. At the end of the side chain, the peaks of the defor-

mation density are very attenuated and their shape is deformed by thermal displacements. This effect has already been observed by Pichon-Pesme *et al.* (1995). From these observations, we can draw the conclusion that the multiple description of the electron density is meaningful for X-ray structures, when the thermal displacement or the non-resolved disorder is moderate ($B < 4$ Å^2). For greater displacement parameters, the dynamic deformation density of the atoms is largely diluted.

3.6. *Phase and deformation density*

The phase error has been evaluated using the R_{free} likelihood estimates method, implemented in the

(*a*)

(*b*)
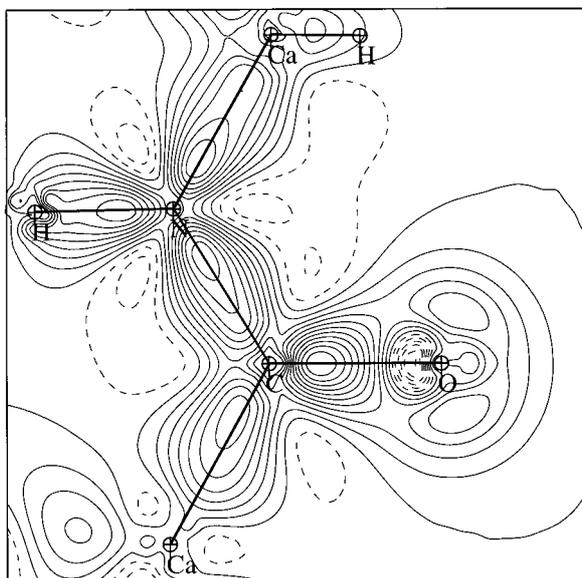
Fig. 7. Static electron-density deformation map in the plane of the peptide bond Aib^2-Lys(Bz)^3. Contours: ±0.1 e Å^{-3}, positive contours: solid line, negative contours: dashed line, zero contour omitted.
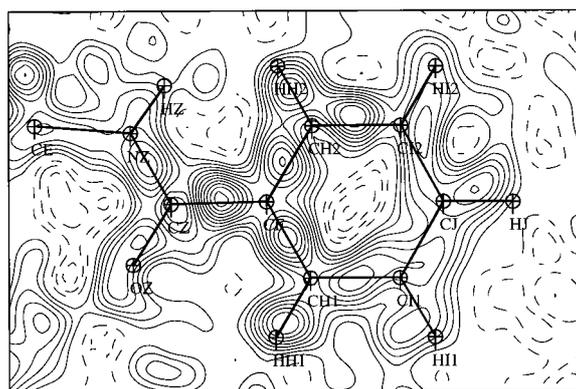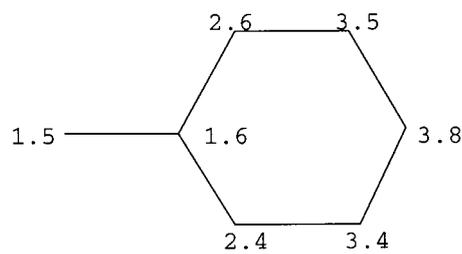
Fig. 8. (*a*) Dynamic deformation electron-density map in the plane of the benzene ring [residue Lys(Bz)^3]. Contours: the same as in Fig. 4. (*b*) The equivalent *B* factors (Å^2) of the (non-H) atoms are indicated.

program *RFLEXPL* (Urzhumtsev *et al.*, 1996). The average phase error of the acentric reflections after refinements I and II are 3.3 and 2.6°, respectively. This 0.7° gain in the phase error denotes the benefit of the multipole parameters transfer to the LBZ structure.

The phase shift due to the atom model has also been analyzed. The unweighted root-mean-square (r.m.s.) phase difference between refinements I and II is 2.6° on the acentric reflections used in the refinement. The r.m.s. phase differences of 2.6 and 2.5° between refinements II and III, and refinements I and III, respectively, are quite similar. These figures illustrate that in the case of the LBZ crystal structure, the phase error and phase shift due to the deformation density are of the same magnitude. Thus, a refinement instead of a transfer of the multipole charge density would not have been physically meaningful.

The dynamic deformation density can be decomposed as the sum of an amplitude-difference density and a phase-difference density (Souhassou *et al.*, 1991),

$$\Delta\rho_{def} = \rho(F_{obs} - F_{sph}, \varphi_{mult}) + \rho(F_{sph}, \varphi_{mult} - \varphi_{sph})$$
$$= \rho(\Delta F) + \rho(\Delta\varphi)$$

The total deformation density has an r.m.s. value of 0.103 e Å^{-3} over the asymmetric unit; the density terms due to ΔF and $\Delta\varphi$ have r.m.s. values of 0.099 and 0.029 e Å^{-3}, respectively. The LBZ molecule crystallizes in an orthorhombic space group and up to 20% of the reflections are centric, which explains the relatively low contribution of the phase difference $\Delta\varphi$ to the deformation density.

3.7. *Scattering factors of the electron-density components*

The scattering factors of the different components of the electron density have been analyzed as a function of the resolution. In Fig. 9(*a*), the scattering factors of the core and valence electrons are plotted for a carbonyl sp^2 C atom. The atom has been assigned a temperature factor $B = 2$ Å^2, which is typical for peptide crystals at cryogenic temperature. The core electrons, which are located close to the nucleus, have a scattering factor which decreases slowly with the resolution. For the valence population and the multipole components, the diffraction decreases rapidly and at resolutions better than 0.9 Å, their contribution becomes negligible. These high-order reflections are thus essentially representative of the core electrons.

As a consequence, in charge-density analysis, the structure refinement is at first performed against the reflections at resolutions better than typically 0.9 Å. The resulting atomic positions and temperature factors are nearly devoid of errors due to the spherical neutral atom approximation. In a second step, the valence populations and multipole parameters are refined using all the reflections. This refinement strategy permits the decon-

volution of the thermal-motion parameters and of the deformation part of the electron density. In the case of the LBZ octapeptide, this refinement strategy could not be applied due to the limited number of measured reflections at very high resolution; the transfer of the multipole parameters was thus carried out.

In the case of an atom with a higher thermal motion $B = 6$ Å^2 (Fig. 9*b*), there is a sharp decline of all the three components of the diffraction at 0.9 Å resolution and the refinement of the multipole parameters becomes delicate. This observation in the reciprocal space has to be related to the smearing of the multipole
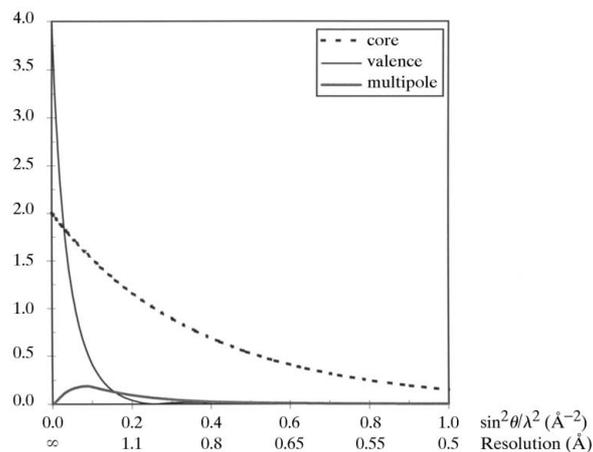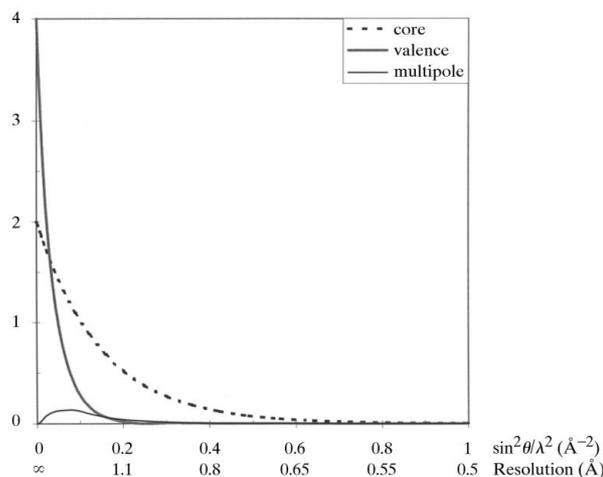
Fig. 9. Scattering factors of the different components of the electron density as a function of resolution or $\sin^2\theta/\lambda^2$ for a carbonyl atom. A C atom with $P_{val} = N_{val} = 4$ is considered here and the multipole parameters have been transferred from the database. For the multipole component, the calculated structure factors have been averaged for reflections in the same resolution shells. (*a*) Isotropic temperature factor of the atom, $B = 2$ Å^2. (*b*) Isotropic temperature factor of the atom, $B = 6$ Å^2.

electron density in the dynamic maps when the B factor increases (Fig. 8).

3.8. *Thermal-motion analysis*

The difference of the refined B factors when using the multipole and spherical atom models has been analyzed in Fig. 10. The equivalent isotropic B factors (*i.e.* B-factor norms) in both models are highly correlated (corr. = 97%). The points (B_{sph}, B_{mult}) are generally located under the diagonal in Fig. 10. The B-factor norms after refinement II, due to the non-spherical atom model, are thus on average, slighly lower than after refinement I ($B_{mult} = 1.07 \times B_{sph} - 0.30$ Å^2). The r.m.s. difference $||B_{sph} - B_{mult}||$ between both atom models is about 0.38 Å^2. The r.m.s. difference of the anisotropic B-factor norms is $\Delta ||B|| = 0.18$ Å^2 whereas the orientation change of the thermal ellipsoids causes a 0.29 Å^2 r.m.s. change of the B factors. The use of scattering factors that include the multipolar part of the atomic electron-density induces thus a significant change, greater than 10% in relative value, on the thermal-motion parameters, when the B factors are moderate, lower than 4 Å^2.

The Rigid-Bond Test (Hirshfeld, 1976) is a way to determine if the anisotropic displacement parameters U^{ij} obtained from a crystallographic refinement are physically realistic. Fig. 11 gives the differences ΔZ of the thermal factors B_{aniso} between bonded pairs of atoms

in the direction of the covalent bond. For the LBZ structure, the transfer of the non-spherical scattering factors decreases slightly the ΔZ discrepancies (ΔZ_{mult} = 89% ΔZ_{sph}); the significance of the thermal-motion parameters is thus marginally improved.

The ΔZ values of the LBZ structure are on average larger ($\langle |\Delta Z| \rangle = 0.46$ Å^2) than the upper threshold used for the Hirshfeld Rigid-Bond Test in sub-atomic resolution crystallography ($|\Delta Z| < 0.001 \times 8\pi^2 = 0.1$ Å^2). The poor quality of the LBZ structure with respect to the Rigid-Bond Test might be due to the diffraction data which have been measured with a precision required for a simple structure determination but not for charge-density work. In addition, the LBZ octapeptide is larger and has slightly higher thermal-displacement parameters than the usual small molecules studied in charge-density work.

3.9. *Electrostatic potential*

The electrostatic potential provides more information than the electron-density mapping about chemical reactivity or molecule–molecule interactions (Lecomte *et al.*, 1992). The potential map in the plane of the peptide bond, when using the spherical and multipole atom models is shown in Figs. 12(*a*) and 12(*b*). The potential on the van der Waals surface is negative in the vicinity of the carbonyl O atoms and generally positive near the C and H atoms. What is the effect of the
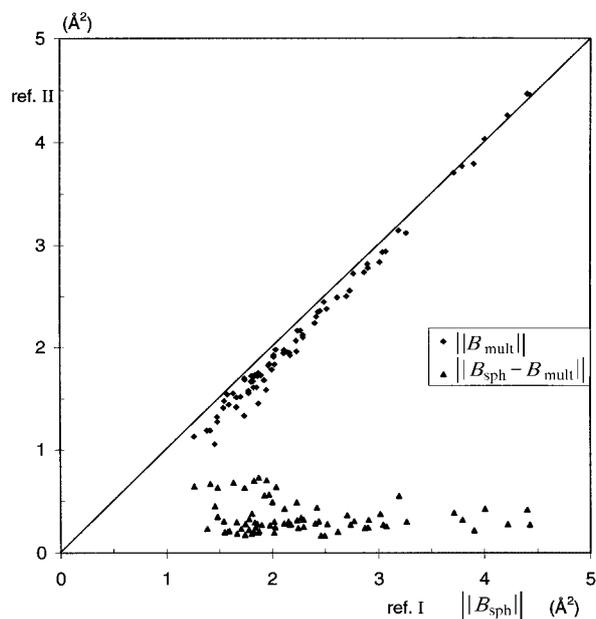
Fig. 10. Effect of the multipole refinement on the temperature factors of the atoms. For all the non-H atoms, the norms of the B factors after refinements I and II are plotted on the x and y axes, respectively. The norm of the anisotropic B factor 3×3 matrix has been defined as: $||B|| = [(\sum_i B_{ii}^2 + 2. \sum_{i < j} B_{ij}^2)/3]^{1/2}$. The norm of the B-factor difference ΔB is also shown: $\Delta B = ||B_{sph} - B_{mult}||$.
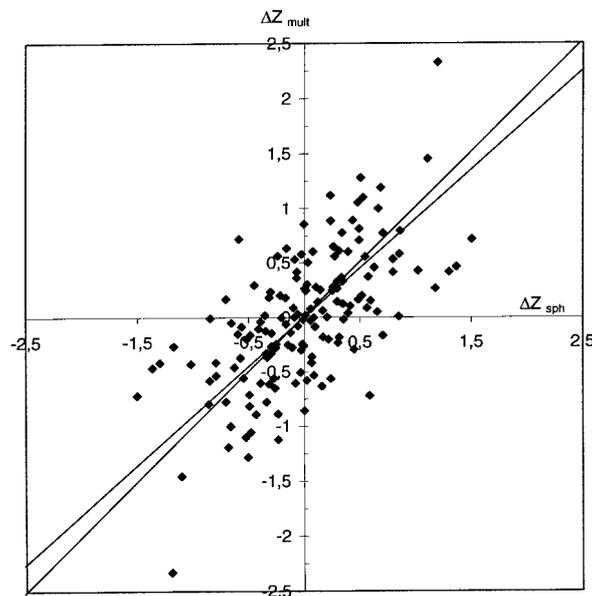
Fig. 11. Differences ΔZ (Å^2) of the projection of the thermal displacements B between bonded pairs of atoms as defined by Hirshfeld (1976). Each point on the graph represents a covalent bond in the LBZ structure after refinements I and II ($x = \Delta Z_{sph}, y = \Delta Z_{mult}$). The diagonal $y = x$ is plotted on the graph. The continuous line $y = 0.89x$ represents the linear regression for $\Delta Z_{mult} = f(\Delta Z_{sph})$.

multipole or spherical description of the atomic density on the electrostatic potential evaluation? The two potentials are qualitatively similar but display locally some differences. The impact of the atomic density model can be seen in Fig. 12(c) which displays the potential difference $\Delta\Phi$ between refinements II and III. The effect of the multipoles is the highest on the covalent bonds where the accumulation of deformation electron density causes the potential to be less positive.

But regarding the impact on the intermolecular interactions, the effect of the electrostatic potential is of essential importance in the outer part of the molecule. Outside of the contracted van der Waals surface in the plane section shown in Figs. 12(a) and 12(b) the potential Φ has an r.m.s. value of 0.14 eÅ^{-1}, whereas for the potential difference $\Delta\Phi$ (Fig. 12c) the r.m.s. mean value is 0.045 eÅ^{-1}. In relative value, the difference $\Delta\Phi$ reaches thus 28% of the electrostatic potential, indi-
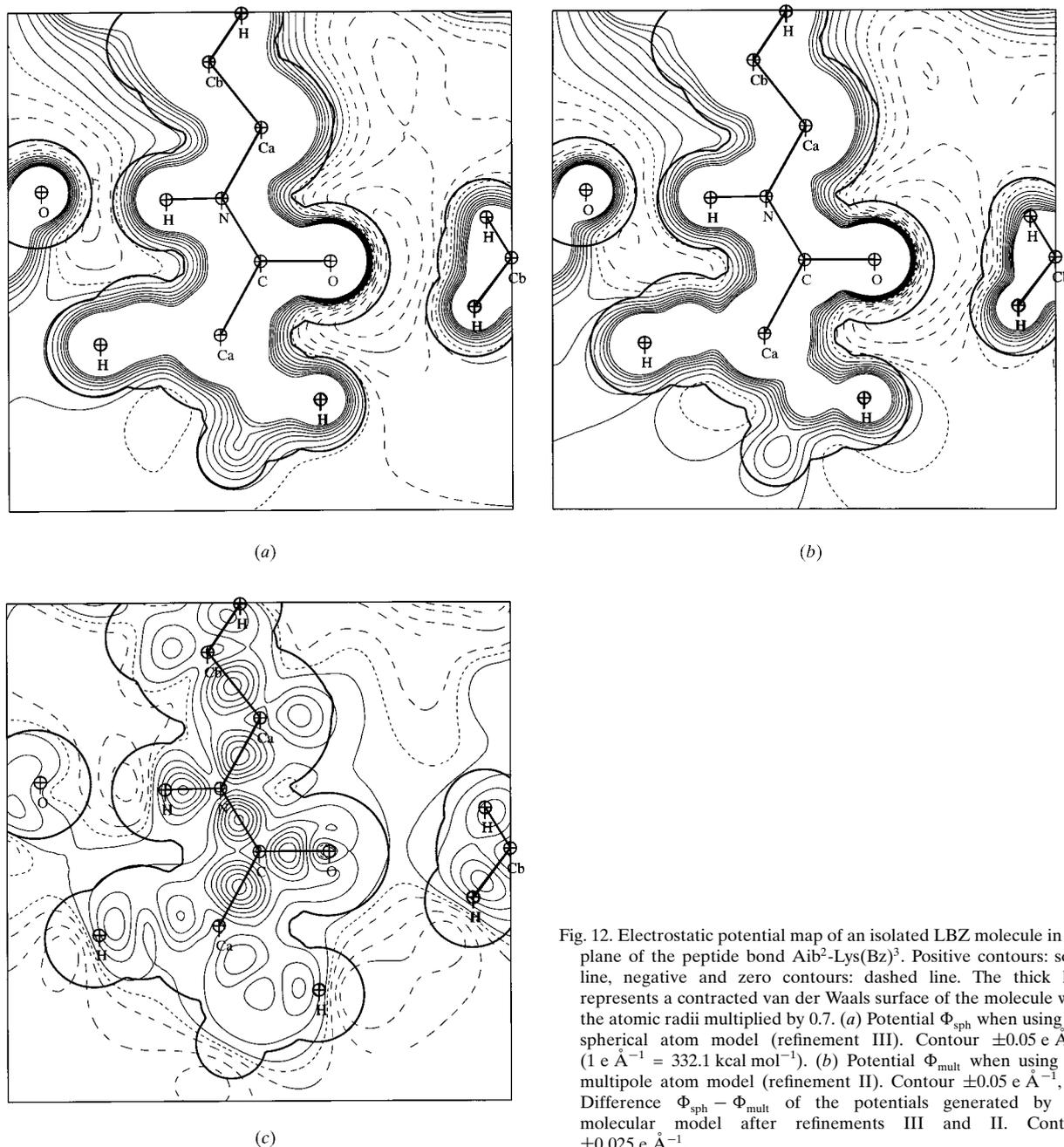
Fig. 12. Electrostatic potential map of an isolated LBZ molecule in the plane of the peptide bond Aib^2-Lys(Bz)^3. Positive contours: solid line, negative and zero contours: dashed line. The thick line represents a contracted van der Waals surface of the molecule with the atomic radii multiplied by 0.7. (a) Potential Φ_{sph} when using the spherical atom model (refinement III). Contour ± 0.05 e Å^{-1}. (1 e $\text{Å}^{-1} = 332.1$ kcal mol^{-1}). (b) Potential Φ_{mult} when using the multipole atom model (refinement II). Contour ± 0.05 e Å^{-1}, (c) Difference $\Phi_{sph} - \Phi_{mult}$ of the potentials generated by the molecular model after refinements III and II. Contour ± 0.025 e Å^{-1}.

cating that a multipole description of the electron density has a non-negligible impact on the electrostatic potential calculation, even in the intermolecular region.

4. Concluding remarks

We are building an aspherical scattering-factor table for the chemical atom types found in all natural amino acids. Such a table includes the valence population P_{val}, the κ contraction/expansion coefficients and the multipole P_{lm} parameters. In the future this will allow an experimental determination of the charges for all the atom types found in peptides and proteins. Also the multipole description of the electron density permits the evaluation of a more accurate electrostatic potential.

This study on the LBZ helix octapeptide has shown that the reflection intensities should be collected to at least 0.9 Å resolution, if one wants all the deformation-density information to be reflected in the diffraction data. Moreover the temperature factors should be low: when the atomic B factors are higher than 4 Å^2, the dynamic deformation density starts to be dispersed. In the LBZ molecule, most of the atoms have moderate temperature factors and have well defined dynamic deformation densities.

What are the potential proteins for which this multipole formalism can be applied? With the availability of synchrotron X-ray sources to biocrystallography and the use of cryogenics, an increasing number of protein structures are known to atomic resolution (Dauter *et al.*, 1995). The protein crambin (Teeter *et al.*, 1993) is an exceptional candidate for an extension of the multipole formalism to proteins. The diffraction data for this 46 amino acids protein have been collected up to 0.54 Å resolution at the EMBL synchrotron in Hamburg. The atomic temperature factors of crambin are remarkably low for a protein, they are typically between 2 and 5 Å^2 for most of the atoms. For the crambin structure, the transfer of the deformation-density parameters should have a meaningful impact on the crystallographic refinement and the valence populations and multipoles might even be refinable. Further work is under way in that direction.

References

Allen, F. H. (1986). *Acta Cryst.* B**42**, 515–522.

Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. (1993). *J. Phys. Chem.* **97**, 10269–10280.

Bouhmaida, N., Ghermani, N. E., Lecomte, C. & Thalal, K. (1997). *Acta Cryst.* A**53**, 556–563.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

Clementi, E. & Raimondi, D. L. (1963). *J. Phys. Chem.* **41**, 2686–2689.

Coppens, P. (1967). *Science*, **158**, 1577.

Coppens, P. (1997). *X-ray Charge Densities and Chemical Bonding.* Oxford: IUCr/Oxford University Press.

Coppens, P., Guru Row, T. N., Leung, P., Stevens, E. D., Becker, P. J. & Yang, Y. W. (1979). *Acta Cryst.* A**35**, 63–72.

Cromer, D. T. (1974). *International Tables for X-ray Crystallography*, edited by J. A. Ibers & W. E. Hamilton, Vol. IV, pp. 148–151. Birmingham: Kynoch Press.

Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1995). *Curr. Opin. Struct. Biol.* **5**, 784–790.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**. 392–400.

Espinosa, E., Lecomte, C., Molins, E., Veintemillas, S., Cousson, A. & Paulus, W. (1996). *Acta Cryst.* B**52**, 519–534.

Ghermani, N. E., Bouhmaida, N. & Lecomte, C. (1993). *Acta Cryst.* A**49**, 781–789.

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

Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239–244.

Howell, P. L., Pangborn, W. A., Marshall, G. R., Zabrocki, J. & Smith, G. D. (1995). *Acta Cryst.* C**51**, 2575–2579.

Lachekar, H. (1997). PhD thesis, Université Henri Poincaré Nancy I, France.

Lecomte, C. (1995). *Advances in Molecular Structure Research*, Vol. 1, edited by I. Hargittai & M. Hargittai, pp. 261–302. Greenwich, CT, USA: JAI Press Inc.

Lecomte, C., Ghermani, N. E., Pichon-Pesme, V. & Souhassou, M. (1992). *J. Mol. Struct. (Theochem.)*, **255**, 241–260.

McCandlish, L. E., Stout, G. H. & Andrews, L. C. (1975). *Acta Cryst.* A**31**, 245–249.

Mata, I., Maniukiewicz, W., Molins, E., Espinosa, E. & Lecomte, C. (1997). *17th Meeting of European Crystallographic Association, Abstracts*, p. 131. Lisbon University, Lisbon, Portugal.

Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P. K. & Kollman, P. A. (1995). *AMBER* 4.1, University of California, San Francisco, USA.

Pichon-Pesme, V. & Lecomte, C. (1998). *Acta Cryst.* B**54**, 485–493.

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

Pichon-Pesme, V., Lecomte, C., Wiest, R. & Bénard, M. (1992). *J. Am. Chem. Soc.* **114**, 2713–2715.

Sheldrick, G. M., (1993). *SHELXL*93, *Program for crystal structure refinement.* University of Göttingen, Germany.

Souhassou, M., Lecomte, C., Blessing, R. H., Aubry, A., Rohmer, M. M., Wiest, R., Bénard, M. & Marraud, M. (1991). *Acta Cryst.* B**47**, 253–266.

Souhassou, M., Lecomte, C., Ghermani, N. E., Rohmer, M. M., Wiest, R., Bénard, M. & Blessing, R. H. (1992). *J. Am. Chem. Soc.* **114**, 2371–2382.

Stewart, R. F., Davidson, E. R. & Simpson, W. T. (1965). *J. Chem. Phys.* **43**, 175–187.

Teeter, M. M., Lamzin, V. S., Dauter, Z. & Wilson, K. (1997). Am. Crystallogr. Assoc. Invited lecture. Ann. Meet. St Louis. Abstracts p. 67. Buffalo: ACA.

Teeter, M. M., Roe, S. M. & Heo, N. H. (1993). *J. Mol. Biol.* **230**, 292–311.

Toniolo, C., Bianco, A., Formaggio, F., Crisma, M., Bonora, G. M., Benedetti, E., Del Duca, V., Savanio, M., Di Blasio, B., Pedone, C. & Aubry, A. (1995). *Bioorg. Med. Chem.* **3**, 1211–1221.

Urzhumtsev, A. G., Skovoroda, T. P. & Lunin, V. Y. (1996). *J. Appl. Cryst.* **29**, 741–744.

Wiest, R., Pichon-Pesme, V., Bénard, M. & Lecomte, C. (1994). *J. Phys. Chem.* **98**, 1351–1362.