



# CCP4 NEWSLETTER ON PROTEIN CRYSTALLOGRAPHY

An informal Newsletter associated with the BBSRC Collaborative  
Computational Project No. 4 on Protein Crystallography.

Number 41

Fall 2002

---

## Contents

### CCP4 News

1. **News from CCP4.** [html](#)  
Alun Ashton, Charles Ballard, Peter Briggs, Maeri Howard-Eales, Pryank Patel and Martyn Winn.
2. **Developments with CCP4i: October 2002.** [html](#)  
Peter Briggs.
3. **Developments with the CCP4 library II.** [html](#)  
Martyn Winn, Charles Ballard and Eugene Krissinel.

### General News

4. **BioMed College at Daresbury.** [html](#)  
Gareth Jones.

### Software

5. **Handling reflection data using the Clipper libraries.** [html](#)  
Kevin Cowtan, Department of Chemistry, University of York, UK.

### Theory and Techniques

6. **Atomic displacement in incomplete models caused by optimisation of crystallographic criteria.** [html](#) [pdf](#) [doc](#)  
P.V Afonine, Université Henri Poincaré, Nancy, France, Centre Charles Hermite, LORIA, Villers-lès-Nancy, France.
7. **Modelling of bond electron density by Gaussian scatters at subatomic resolution.** [html](#) [pdf](#) [doc](#)  
P. Afonine<sup>(1,2)</sup>, V. Pichon-Pesme<sup>(1)</sup>, N. Muzet<sup>(1)</sup>, C. Jelsh<sup>(1)</sup>, C Lecomte<sup>(1)</sup> and A. Urzhumtsev<sup>(1)</sup>, <sup>(1)</sup> Université Henri Poincaré, Nancy, France, <sup>(2)</sup> Centre Charles Hermite, LORIA, Villers-lès-Nancy, France.
8. **Bulk-solvent correction for use with the CCP4 version of AMoRe.** [html](#) [pdf](#)  
Guido Capitani<sup>(1)</sup> and Andrei Fokine<sup>(2)</sup>, <sup>(1)</sup> University of Zürich, Switzerland, <sup>(2)</sup> Université Henry Poincaré, Nancy, France.
9. **Variation of solvent density and low-resolution *ab initio* phasing.** [html](#) [pdf](#) [doc](#)  
Andrei Fokine, Université Henry Poincaré, Nancy, France.
10. **Retrieval of lost reflections in high resolution Fourier syntheses by 'soft' solvent flattening.** [html](#) [pdf](#) [doc](#)

Natalia L. Lunina<sup>(1)</sup>, Vladimir Y. Lunin<sup>(1)</sup> and Alberto D. Podjarny<sup>(2)</sup>, <sup>(1)</sup> Russian Academy of Sciences, Russia, <sup>(2)</sup> CU Strasbourg, France.

## Bulletin Board

### 11. **Summaries.** [html](#)

Maria Turkenburg, University of York, York, UK.

---

**Editors:** Charles Ballard and Maeri Howard-Eales

Daresbury Laboratory, Daresbury,  
Warrington, WA4 4AD, UK

---

**NOTE:** The CCP4 Newsletter is not a formal publication and permission to refer to or quote from the articles reproduced here must be referred to the authors.

**Contributions** are invited for the next issue of the newsletter, and should be sent to Charles Ballard by e-mail at [c.c.ballard@ccp4.ac.uk](mailto:c.c.ballard@ccp4.ac.uk) by 30th April 2003. HTML is preferred but other formats are also acceptable.

---

[CCP4 Main Page](#)



# Modelling of bond electron density by Gaussian scatters at subatomic resolution

P. Afonine<sup>1,2</sup>, V. Pichon-Pesme<sup>1</sup>, N. Muzet<sup>1</sup>, C. Jelsch<sup>1</sup>, C. Lecomte<sup>1</sup> & A. Urzhumtsev<sup>1</sup>

<sup>1</sup>LCM3B, UPRESA 7036 CNRS, Faculté des Sciences, Université Henri Poincaré, Nancy I, 54506, Vandoeuvre-lès-Nancy, France

<sup>2</sup>Centre Charles Hermite, LORIA, Villers-lès-Nancy, 54602 France

## Abstract

At resolutions of 1-2 Å and lower, traditional for macromolecular crystallography, the electron density of the crystal unit cell is modelled by a sum of contributions from individual atoms (plus the bulk solvent). The contribution of each atom is calculated as a spherical or an elliptic function centred at the atomic position. Such models are not adequate at a subatomic resolution, higher than 1 Å, as proven by residual Fourier maps. In particular, these maps show extra density peaks between the bonded atoms. A modelling of this density by placing extra 'gaussian' electrons at the corresponding positions allows the improvement of crystallographic criteria. This new model allows also the determination of some important characteristics of the crystal such as critical points of the electron density distribution.

## 1. Introduction

Crystal models play several roles. First, they give a possibility to describe the content of a crystal by a finite number of parameters, usually with a clear physical meaning. Secondly, these models allow the determination of properties which are not available directly from the experiment. For example, model parameters can be used to calculate the electron density distribution while the diffraction experiment gives only a limited set of its Fourier coefficients and allows thus the calculation of Fourier syntheses only at a limited resolution. Naturally, such a model should be quite precise and detailed.

In macromolecular crystallography, for many years the usual resolution of the experimental data was about 2-3 Å. At such a resolution, the models, composed from isotropic atoms and refined using crystallographic restraints or constraints, give a good description of experimental diffraction data. During last years, many macromolecular structures were reported at atomic resolution, about 1 Å (for a review, see Dauter *et al.*, 1997; Longhi *et al.*, 1998). Such data require more detailed models. In particular, anisotropic thermal displacement parameters become necessary. An introduction of extra parameters is justified by increased number of structure factor magnitudes as confirmed by R-free (Brünger, 1992) calculations. Recently, several cases were reported where macromolecular crystals diffracted to the resolution higher than 0.9 Å. An example is aldose reductase (Lamour *et al.*, 1999) crystals of which diffract to 0.64 Å even when the size of this protein is quite large, 315 amino acid residues. At such a subatomic resolution, the spherical electron density model even with anisotropic thermal parameters is already inadequate. In particular, difference maps show significant density peaks around the atoms, essentially at the interatomic bonds (Housset *et al.*, 1994; Lamzin *et al.*, 1999). On the other hand, a large amount of data could allow the detailed study of physical properties of electron density of the crystal.

In small molecules crystallography such subatomic resolution is usual and the problem was solved by introduction of multipolar models (Stewart, 1969; Hansen & Coppens, 1978). Here the density of each atom is modelled by a function:

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

where the total atomic electron density ( $\rho_{atom}$ ) is decomposed into three terms corresponding to the core electrons, valence electrons and non-spherical part of the valence electron distribution as a multipole density (see Hansen & Coppens, 1978, for details and notation). The number of parameters for each atom grows with the number of included spherical harmonics. For traditional models, a multipolar atom, different from a hydrogen, can have up to 28 parameters (18 multipolar, 3 positional, 6 temperature parameters, 1 occupancy). Recently, Pichon-Pesme *et al.* (1995) and Jelsch *et al.* (1998) demonstrated the transferability of such models to macromolecular crystals when a large enough number of structure factor magnitudes allows this (in other words, when the resolution of the diffraction data set is high enough). Such models have been successfully refined for crystals of crambin (Jelsch *et al.*, 1999), for a scorpion toxin (Housset *et al.*, 2000) and for an aldose reductase complex (Guillot *et al.*, 2000), respective resolution of these crystals is 0.54, 0.9, and 0.65 Å.

Some other protein crystals, for example those of other complexes of aldose reductase, diffract to slightly lower resolution, about 0.9 Å. Use of multipolar atoms and refinement of hydrogen atoms reduces the ratio  $N_{data}/N_{parameters}$  to the limits when such type of a model can be hardly used. The major risk is, as in any modelling, to overfit the data introducing too many parameters in the model.

In order to answer this problem, a project was started to develop a simplified molecular model which is more detailed than a usual model of anisotropic atoms but contains less parameters than a multipolar model. This model can be used when the amount of experimental data is not sufficient enough to use multipolar models. Additionally, such a model would accelerate the calculations of crystallographic values (structure factors, electron density, and electrostatic potential) which are quite time consuming when multipolar models are employed for macromolecules.

## 2. Model of bond electrons

As traditionally for macromolecular crystals, incorrect parts of an available model are highlighted in difference density maps. In particular, so called ‘deformation density maps’ can be calculated as the Fourier series

$$\Delta\rho_{def} = V^{-1} \sum_{\mathbf{h}} \left( k^{-1} F_{obs}(\mathbf{h}) e^{i\varphi_{best}} - F_{sph}(\mathbf{h}) e^{i\varphi_{sph}} \right) e^{-2\pi i \mathbf{h} \cdot \mathbf{r}}$$

Here  $V$  – unit cell volume,  $k$  – scale factor,  $F_{obs}(\mathbf{h})$  – observed structure factor amplitudes,  $\varphi_{best}(\mathbf{h})$  – best phases available,  $F_{sph}(\mathbf{h})$  and  $\varphi_{sph}(\mathbf{h})$  – structure factors amplitudes and phases obtained from spherical atomic model. These maps show an excess and a missed part in the available model by negative and positive peaks whose size does not allow seeing them at lower resolution. Such maps show the redistribution of the electron density due to formation of interatomic bonds and other interactions. Among various ‘deformations of density’, large peaks are seen most clearly on the bonds (Fig. 1A). This extra density, which represents reorganisation of electrons to bonding, seems to be the major need for a necessary model modification.

Previously, several attempts have been done (Ewald & Höln, 1936a, 1936b); Brill, 1960; Hellner, 1977; Dietrich & Scheringer, 1978; Scheringer, 1980; Pietsch, 1981; Pietsch & Unger, 1981; Pietsch, 1985; Pietsch *et al.*, 1986) to model this density by placing there an additional scattering matter, a kind of pseudo atom. All these attempts were done with small-molecular crystals with a small number of diffraction intensities. On contrary, macromolecular crystals have a very large number of reflections at such a subatomic resolution. This allows an

easy application of the R-free methodology (Brünger, 1992) which has been shown as a powerful tool to indicate the data overfitting. First checks (Cetin *et al.*, 2000) showed a feasibility of this modelling.

### 3. Preparation of the test model

More detailed tests were done using the leu-enkephalin peptide (Wiest *et al.*, 1994). This pentapeptide (Tyr<sup>1</sup>-Gly<sup>2</sup>-Gly<sup>3</sup>-Phe<sup>4</sup>-Leu<sup>5</sup>) crystallises in space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> with unit cell parameters  $a = 10.851 \text{ \AA}$ ,  $b = 13.095 \text{ \AA}$ ,  $c = 21.192 \text{ \AA}$  and  $Z = 4$ . The diffraction data were collected to  $0.43 \text{ \AA}$ . However, the tests described below were done at a lower resolution,  $d > 0.56 \text{ \AA}$  ( $\sin\theta/\lambda < 0.89 \text{ \AA}^{-1}$ ), beyond which the completeness of data is insufficient (Sheldrick, 1990), see Fig. 2. At the resolution of  $0.56 \text{ \AA}$  the number of independent reflections is 8707 and completeness for the last high-resolution shell is higher than 60 %. The enkephalin model contains 43 non-hydrogen and 43 hydrogen atoms. A refined multipolar model (M1) was obtained (Wiest *et al.*, 1994) using the program MOLLY (Hansen & Coppens, 1978). This refinement was done without separation of the data set into work and test (*R*-free) subsets.

Because the *R*-free values were important for our tests (one of the indicators of modelling progress) and in order to estimate the quality of previously refined model (M1) in terms of ‘*R*-free’, a usual procedure was applied (Brünger, 1993). First, a test set was chosen, composing 20% of the total amount of reflections belonging to  $0.56\text{-}11.0 \text{ \AA}$  resolution. This selection was done randomly and uniformly in several resolution shells. Random and independent errors were introduced into atomic coordinates of M1 model by performing molecular dynamics simulation at 2000K with subsequent energy minimisation using CNS (Brünger *et al.*, 1998). Such procedure removed some of the ‘memory’ of previous refinement towards test set and gave the model differing from M1 by  $\sim 0.06 \text{ \AA}$  shift in coordinates. Then, this model was refined using only the 80% rest of reflections (work set) by the program MOPRO (Guillot *et al.*, 2001) at the exactly same conditions as previously, reported by Wiest *et al.* (1994). The desired  $R / R_{\text{free}}$ -factor statistics for the model obtained (M2) is summarised in Table 1.

### 4. Test of bond electron models

All the subsequent tests were done with the program suite SHELX (Sheldrick & Schneider, 1997). No stereochemical constraints were used in these tests.

First of all, the conventional anisotropic refinement of enkephalin was done at  $0.56\text{-}11.0 \text{ \AA}$  resolution in order to have reference values for such type of models at this resolution (model A1 in Table 2). It is not surprising that the results of standard anisotropic refinement ( $R = 9.08 \%$ ,  $R_{\text{free}} = 9.74 \%$ ) are worse than those after multipolar refinement ( $R = 7.90 \%$ ,  $R_{\text{free}} = 8.63 \%$ ; M2 model in Table 1). This tendency is true in every resolution shell (Fig. 3).

For the first test with bond electron (BE in what follows) models, standard hydrogen atoms were used. These atoms were placed in the middle of interatomic bonds (as in Fig. 1B). The optimal starting values of occupancies for BEs were found as 0.5, agreeing with previous tests (Cetin *et al.*, 2000) and the starting values of isotropic temperature factors were taken equal to the average temperature parameters of the neighbouring atoms. Different set of parameters were refined and various refinement strategies were tested. In many cases, including the case of anisotropic BEs, the *R*-free criterion increased or refinement was unstable (details of these numerous tests will be discussed elsewhere). The best improvement in both *R* and *R*-free values, to 8.43 and 9.16 % (model BE1 in Table 2), was obtained when refining the following parameters:

- for each non-hydrogen atom of the model: *coordinates, anisotropic temperature factor and occupancy*;

- for each hydrogen atom of the model placed according to ideal stereochemistry: *isotropic temperature factor (while occupancy is fixed equal to 1)*;
- for each bond electron (BE): *isotropic temperature factor and occupancy*.

It is important to note that this refinement decreased the occupancies of the enkephalin atoms showing the redistribution of the density to newly placed BEs.

The number of BEs is approximately equal to the number of bonds (some extra BEs can be also introduced for lone pair electrons) or, in other words, is approximately equal to the total number of atoms in the molecule. Therefore, the number of additional parameters in BE-model can be estimated as 2 parameters per atom, resulting in 12 parameters per atom in total.

The model was improved when the BEs, taken previously as hydrogens, were replaced by Gaussian scatterers defined from DFT calculations by program SIESTA (Sanchez-Portal *et al.*, 1997) as follows. First of all, the theoretical deformation density maps were obtained for all types of residues as the difference between the exact electron density distribution calculated by quantum-chemical methods and that calculated from the model using standard scattering factors for neutral atoms. For these calculations, each idealised single residue was taken as an isolated molecule in gas phase. The peaks at the bonds were approximated by a single-gaussian function, and the corresponding distance between the peak centre and the neighbouring atoms was calculated. Then, such gaussian peaks were placed in the covalent bond positions of enkephalin. Similarly to the previous case, the starting values of isotropic temperature factors were taken equal to the average temperature parameters of the neighbouring atoms. The refinement of this model decreased the *R* and *R*-free factors to 8.12 and 8.76 % (model BE2 in Table 2). The behaviour of *R*-factors as a function of resolution is shown in Fig. 3.

Summarising, it should be noted that the proposed BE-model:

- is better than the best anisotropic model;
- requires about 12 parameters per atom, in comparison with 10 parameters for the anisotropic model and 28 for a multipolar model;
- gives values of overall *R*- and *R*<sub>free</sub>-factors close to that for the refined multipolar model.

## 5. Further validation of bond electrons models

A lower value of the *R*-free factor is a good proof of a higher quality of BE-models in comparison with classical anisotropic models. However, more tests were done to confirm the quality and physical meaning of the BE-model.

First, the rigid-bond test (Hirshfeld, 1976) showed that BEs do not cause unphysical perturbations in the thermal parameters of atoms. A refinement even without rigid-bond constraints gave the rigid-bond criterion at the same level as for the multipolar model (Fig. 4). An inclusion of the rigid-bond criterion into refinement (not shown) improved the similarity of the projections of thermal ellipsoids on the bond without increasing the *R* and *R*-free values.

The second check consisted in verification of a predicting power of such BE-models. Multipolar models allow the calculation of accurate electron density. In small-molecule crystallography, these electron density maps are called ‘experimental’ because the models used for their generation are obtained from the experimental data. One of the important features of this distribution is critical points, where the gradient of the density is equal to zero (Bader, 1990). In particular, these points allow the characterisation of atomic interactions. Each such point is characterised by three numbers which are the eigenvalues of the normal matrix of the electron density calculated at this critical point (the matrix of second derivatives with respect to three coordinates). These values allow one to characterise the type of covalent bonds. A special role is played by density Laplacian which is the sum of the three eigenvalues.

In order to obtain the parameters of critical points for BE-models, a special program was written allowing a very fast calculation of the electron density map, exact maps of the

gradient of electron density and that of the density Laplacian directly from the BE model. Electron density maps calculated from the best available isotropic and anisotropic models did not reproduce the critical points derived from the multipolar model. On the contrary, the BE2 model gave both the correct position of the critical points and, for most of bonds, the eigenvalues were very close to those obtained from the multipolar model (Table 3). A few cases with incorrect eigenvalues, for example that for the C=O bonds, prove that for some types of bonds the parameters of BEs, including their position, should be improved in further studies.

## 6. Conclusions and discussions

Several independent criteria confirmed that modelling of the density on the interatomic bonds plays the major role in the improvement of anisotropic models on the way to multipolar models. The simplest way to improve an anisotropic model is its completion by appropriate gaussian scatters for bond electrons. The available software does not allow the complete refinement of parameters of these electrons as it seems to be necessary. Nevertheless, by refining only the shape and not the position of these BEs,  $R$  and  $R$ -free factors can be significantly decreased in comparison with anisotropic models. Moreover, our tests demonstrated that the refinement of BE-models gives the values of  $R$ - and  $R_{\text{free}}$ -factors similar to those for multipolar models.

These models reproduce the experimental structure factor values better thus allowing the calculation of improved Fourier syntheses using the experimental structure factors. In addition, such models provide electron density maps that cannot be obtained using isotropic or anisotropic models. The maps reproduce important features of the electron density distribution, for example, its critical points.

A slight decreasing in occupancies of C, N and O atoms in the refined BE-model indicates the tendency to conserve the total charge of the model without electroneutrality constraint; this confirms physical meaning of such a modelling.

The BE-models are described by a smaller number of parameters than multipolar models (approximately 2 times less) and therefore can be used at lower resolution when the number of experimental structure factor magnitudes is smaller. In addition, gaussian scattering factors of all atoms of the model (a 5-gaussian approximation of atomic scattering factor works well at least up to the resolution of 0.25 Å, see International Crystallographic Tables, 1998) should allow a very fast and direct calculation of the exact maps of the gradient of electron density and the maps of the density Laplacian.

The refinement of BE-models does not require any special refinement programs. Nevertheless, more sophisticated software would allow the use of special criteria and refinement strategies.

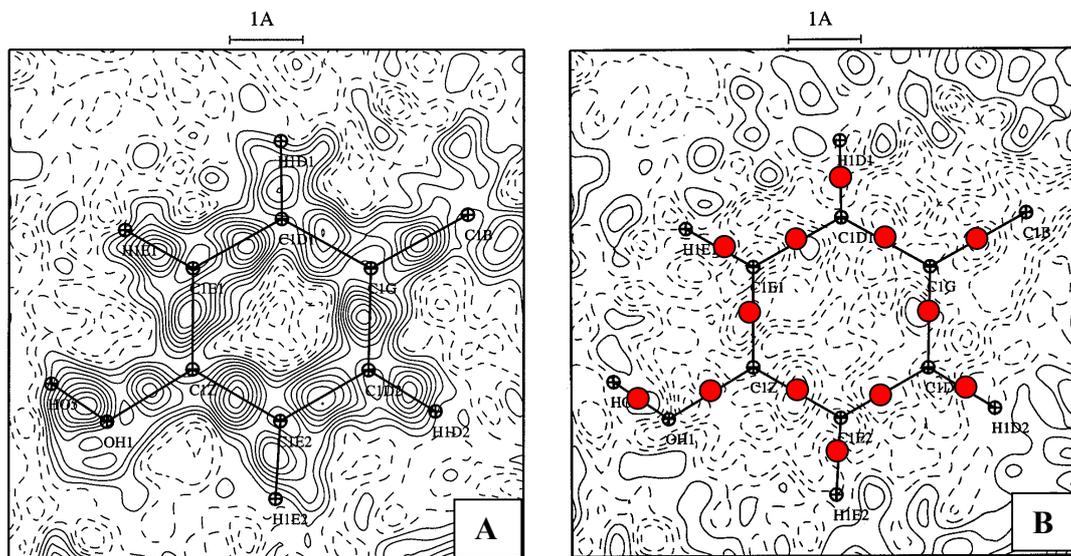
The work is in progress.

## Acknowledgment

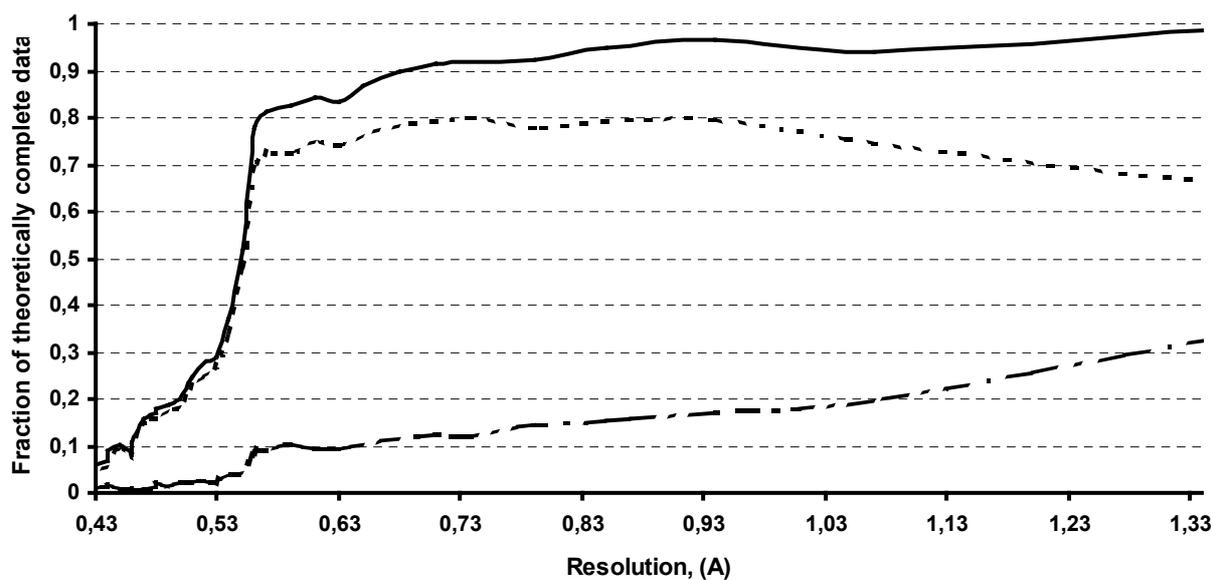
The authors thank C. Katan, M. Souhassou, N.-E. Ghermani for useful discussions of high-resolution features of crystals, I. Uson and G. Sheldrick for their help with the use of SHELX program suite and V. Lunin for fruitful discussions. The authors are participants of the GdR 2417 CNRS.

## References

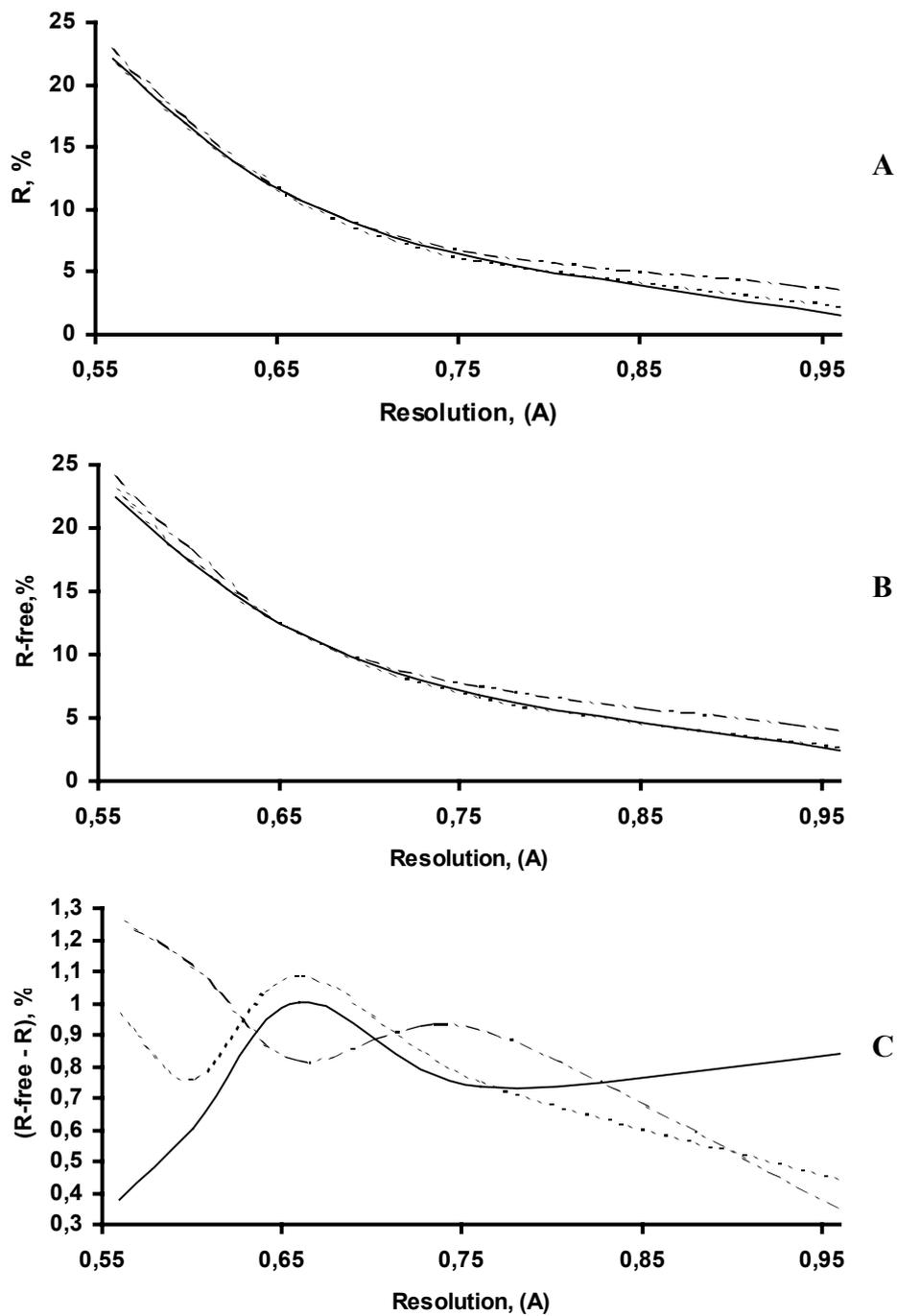
- Bader, R. W. F. (1990). *Atoms in Molecules. A Quantum Theory*. Oxford University Press.
- Brill, R. (1960). *Acta Cryst.*, **13**, 275-276
- Brünger, A. T. (1992). *Nature*, **355**, 472-475
- Brünger, A. T. (1993). *Acta Cryst.*, **D49**, 24-36
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.*, **D54**, 905-921.
- Cetin, A., Pichon-Pesme, V., & Urzhumtsev, A. (2000). *University of H. Poincaré, Nancy 1, LCM3B; internal report*.
- Dauter, Z., Lamzin, V. S. & Wilson, K.S. (1997). *Curr. Opin. Struct. Biol.*, **7**, 681-688
- Dietrich, H. & Scheringer, C. (1978). *Acta Cryst.*, **B34**, 54-63
- Ewald, P. P. & Höln, H. (1936a). *Ann. Phys., Lpz.*[5], **25**, 281
- Ewald, P. P. & Höln, H. (1936b). *Ann. Phys., Lpz.*[5], **26**, 173
- Guillot, B., Jelsch, C., Muzet, N., Lecomte, C., Howard, E., Chevrier, B., Mitschler, A., Podjarny, A., Cousson, A., Sanishvili, R. & Joachimiak, A. (2000). *Acta Cryst.*, **A56** (Supplement), s199
- Guillot, B., Viry, L., Guillot, R., Lecomte, C. & Jelsch, C. (2001). *J. Appl. Cryst.*, **34**, 214-223
- Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.*, **A34**, 909-921
- Hellner, E. (1977). *Acta Cryst.*, **B33**, 3813-3816
- Hirshfeld, F. L. (1976). *Acta Cryst.*, **A32**, 239-244
- Housset, D., Habersetzer-Rochat, C., Astier, J.-P. & Fontecilla-Camps, J. C. (1994). *J.Mol.Biol.*, **238**, 88-103
- Housset, D., Benabicha, F., Pichon-Pesme, V., Jelsch, C., Maierhofer, A., David, S., Fontecilla-Camps, J. C. & Lecomte, C. (2000). *Acta Cryst.*, **D56**, 151-160
- Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* **D54**, 1306-1318
- Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. H. & Lecomte, C. (2000). *PNAS*, **97**, no. 7, 3171-3176
- Lamour, V., Barth, P., Rogniaux, H., Poterszman, A., Howard, E., Mitschler, A., Van Dorsselaer, A., Podjarny, A. & Moras, D. (1999). *Acta Cryst.* **D55**, 721-723
- Lamzin, V., Morris, R.J., Dauter, Z., Wilson, K.S. & Teeter, M. M. (1999). *J.Biol.Chem.*, **274**, 20753-20755.
- Longhi, S., Czjzek, M. & Cambillau, C. (1998). *Curr. Opin. Struct. Biol.*, **8**, 730-737
- Pichon-Pesme, V., Lecomte, C. & Lachekar, H. (1995). *J. Phys. Chem.*, **99**, 6242-6250.
- Pietsch, U. (1981). *Phys. Stat. Sol., (b)* **103**, 93
- Pietsch, U. & Unger, K. (1981). *Phys. Stat. Sol., (b)* **104**, 253
- Pietsch, U. (1985). *Phys. Stat. Sol., (a)* **87**, 151
- Pietsch, U., Tsirelson, V. G. & Ozerov, R. P. (1986). *Phys. Stat. Sol., (b)* **138**, 47-52
- Sanchez-Portal, D., Ordejon, P., Artacho, E. & Soler, J.M. (1997). *Int. J. Quant. Chem.*, **65**, 453.
- Scheringer, C. (1980). *Acta Cryst.*, **A36**, 205-210
- Sheldrick, G. M. (1990). *Acta Cryst.*, **A46**, 467-473
- Sheldrick, G. M. & Schneider, T. R. (1997). *Methods in Enzymology*, **277**, 319-343
- Stewart, R.F. (1969). *J. Chem. Phys.*, **51**, 4569
- Wiest, R., Pichon-Pesme, V., Bénard, M. & Lecomte, C. (1994). *J. Phys. Chem.*, **98**, 1351-1362



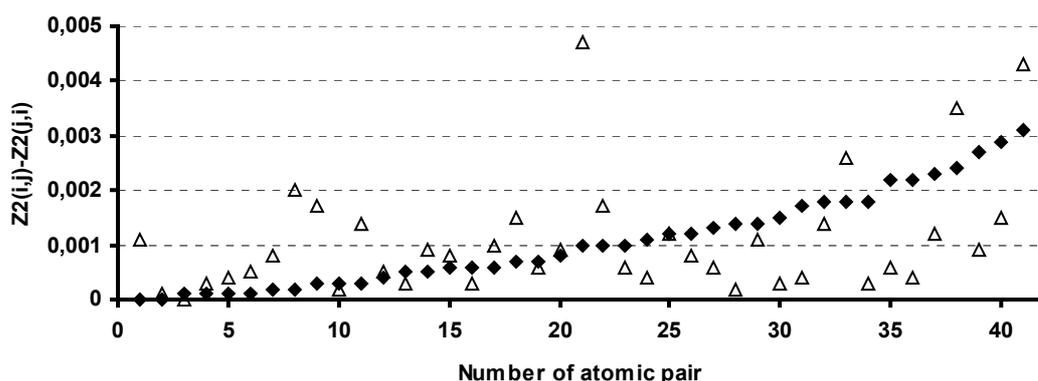
**Fig. 1.** (A) The model of spherical atoms and corresponding deformation density map showing the extra electron density peaks at interatomic bond distances. Such peaks are subject to modelling at resolution higher than 0.9 Å. (B) The same model completed by ‘bond electrons’ (red circles) placed between chemically bonded atoms and difference Fourier map after refinement of this BE-model. The maps calculated at 0.45 Å resolution show the Tyr-side chain of the YGG polypeptide (Cetin *et al.*, 2000).



**Fig. 2.** Fraction of theoretically complete data: dashed line for acentric, dashed-dotted for centric reflections accordingly, bold line for total completeness.



**Fig. 3.** Graphics A, B and C show respectively  $R$ -,  $R_{\text{free}}$ - and  $(R_{\text{free}}-R)$ -factors as a function of resolution. Solid line: multipole model; dotted line: BE-model; dashed-dotted line: standard anisotropic model.



**Fig. 4.** Results of Hirshfeld's rigid bond test. For every covalently bonded pair of atoms  $i$  and  $j$ , the discrepancy of the mean square vibration amplitudes  $Z^2(i,j)-Z^2(j,i)$  along their mutual bond is presented as a function of index number for each bond. The bonds are given in increasing order of the discrepancy for the multipole model. Filled rhombs are for BE- and triangles are for previously refined multipolar model. Mean value is equal to  $0.001 \text{ \AA}^2$  for both cases, corresponding to the commonly acceptable limit.

**Table 1.** Retrieved  $R / R_{\text{free}}$ -factor statistics for previously refined multipolar model of enkephalin. M1 is the initial multipolar model (Wiest *et al.*, 1994), and M2 is the same model with recovered  $R_{\text{free}}$  statistics (see text, section 3, for details). Statistics is given for all reflections with  $I > 0$ .

Models (resolution 0.56 -11.0 Å)	$R$ (F), %	$R_{\text{free}}$ (F), %	$(R_{\text{free}} - R)$ , %	Number of reflections		
				work set	test set	total
M1	8.09	-	-	8707	-	8707
M2	7.90	8.63	0.73	6894	1813	8707

**Table 2.** Refinement statistics for different models

Model composition		Parameters <sup>*)</sup>				$R$ ( $I > 0$ ), % and data / parameters	$R_{\text{free}}$ ( $I > 0$ ), %	$(R_{\text{free}} - R)$ , %
		$U_{\text{aniso}}$	$B_{\text{iso}}$	$Q$	$XYZ$			
Model A1	→ non-H atoms of the model	+	-	+	+	9.08 6894 / 474	9.74	0.66
	→ H-atoms of the model	-	+	fixed =1.0	-			
Model BE1	→ non-H atoms of the model	+	-	+	+	8.43 6894 / 658	9.16	0.73
	→ H-atoms of the model	-	+	fixed =1.0	-			
	→ BE (bond electrons) taken as H-atoms	-	+	+	-			
Model BE2	→ non-H atoms of the model	+	-	+	+	8.12 6894 / 658	8.76	0.64
	→ H-atoms of the model	-	+	fixed =1.0	-			
	→ BE (bond electrons) taken from <i>ab initio</i> calculations	-	+	+	-			

<sup>\*)</sup> 1)  $U_{\text{aniso}}$  and  $B_{\text{iso}}$  – anisotropic and isotropic displacement parameters, respectively;  $Q$  – occupancy coefficient;  $XYZ$  – atomic coordinates

2) The sign '+' or '-' means that the corresponding parameter was refined or fixed, respectively

**Table 3.** Selected examples of parameters for bond critical points obtained from BE2-model (Table 2)

Atomic pair, A-B	R <sub>A</sub> (Å)	R <sub>B</sub> (Å)	ρ (eÅ <sup>-3</sup> )	λ <sub>1</sub> (eÅ <sup>-5</sup> )	λ <sub>2</sub> (eÅ <sup>-5</sup> )	λ <sub>3</sub> (eÅ <sup>-5</sup> )	∇ <sup>2</sup> (eÅ <sup>-5</sup> )	η
C1-C2	0.774	0.762	1.60	-8.8	-8.6	5.7	-11.7	1.54
C1-N1	0.621	0.881	1.59	-8.9	-8.8	12.3	-5.5	0.73
C25-C26	0.775	0.747	1.56	-8.9	-8.8	5.2	-12.5	1.72
C23-C24	0.781	0.756	1.53	-9.7	-9.5	4.5	-14.2	2.14
C24-C25	0.768	0.757	1.53	-8.4	-7.9	5.8	-10.5	1.44

R – distance from the critical point to the neighbouring atom

ρ – value of electron density at the critical point

λ<sub>1</sub>, λ<sub>2</sub>, λ<sub>3</sub> – eigenvalues of the normal matrix of the electron density calculated in this critical point (the matrix of second derivatives with respect to three coordinates)

∇<sup>2</sup> – Laplacian

η = |λ<sub>1</sub>| / λ<sub>3</sub>