

## Local Improvement of Electron-Density Maps

A. G. URZHUMTSEV

IGBMC, Parc d'Innovation, BP 163, 67404 Illkirch, c.u. de Strasbourg, France, and IMPB, RAS, Puschino, Moscow Region, 142292, Russia. E-mail: sacha@igbmc.u-strasbg.fr

(Received 23 January 1997; accepted 27 March 1997)

### Abstract

A method is proposed for improvement of local portions of electron-density maps. The main difference between this method and several other dummy-atoms techniques is that the dummy atoms are placed independently of the initial weak density and are not biased by it. An example of an application of the method is given.

### 1. Introduction

Currently, the method of improving electron-density maps through refinement of an atomic model in reciprocal space and recalculation of phases based on the refined model, first applied by Watenpaugh, Sieker, Herriott & Jensen (1973), is a routine procedure. In the case when the stereochemically correct atomic model is not yet available, Agarwal & Isaacs (1977) suggested that the density map could be improved by placing dummy atoms with certain stereochemical constraints. This method was shown to be highly effective; however its authors noted some artificial breaks in the electron density where their procedure did not place enough dummy atoms. This can be explained by the observation that statistically the density corresponding to the missing atoms is weaker than that for the atoms already included in the phase calculation (Raman, 1959; Main, 1979). Since dummy atoms are usually missed in the regions of poorer electron density, the corresponding density contrast for well and poorly defined regions becomes higher and decreases the signal for the latter to the level of noise.

This idea of dummy atoms was further developed by Lunin, Urzhumtsev and co-workers (Lunin & Urzhumtsev, 1984; Lunin *et al.*, 1985) by suggesting a new scheme for model construction and modification and by developing a maximum-likelihood approach to phase-error estimation. Their mixed atomic models contained the stereochemically correct model for the interpreted portion of the map and dummy atoms for an uninterpreted one. The method was successfully applied in several cases at a resolution ranging from 2.3 to 3.5 Å. An extension of the use of dummy atoms for the high-resolution case (Lamzin & Wilson, 1993) is implemented in the package *ARP* (Collaborative Computational Project, Number 4, 1994).

However, even with a number of improvements, the method still had three major problems. First, phase errors for the refined models were essentially underestimated (Lunin & Urzhumtsev, 1984; Read, 1986). Second, it was impossible to retrieve weak density. Third, it was impossible to extend the phase information starting from a resolution lower than approximately 3.5 Å. The first problem has recently been solved by Lunin & Skovoroda (1995). To solve the second problem, Agard and co-authors (Wilson & Agard, 1993; Baker, Bystroff, Fletterick & Agard, 1993; Bystroff, Baker, Fletterick & Agard, 1993) used the properties of atomicity and connectivity simultaneously. Their approach, realised through the program *PRISM* (Wilson & Agard, 1993), allows one to build 'bridges' between regions with reasonably strong electron density and to recover missing density between them. However, the presence of some initial peaks of density near the poorly defined regions is needed.

### 2. Density growing: method description

A technique of 'density growing' is proposed to improve poor density in the regions where, in general, the idea of 'connectivity bridges' is not always applicable, *e.g.* in terminal regions or for side chains. The basic idea of the current approach is to fill a non-interpreted region of an electron-density map with some scatterers placed independently of the current density values in order to increase the contrast for a weak density. The number of such scatterers should be small enough so that their parameters can be optimized by fitting to the available structure-factor amplitudes, *i.e.* that the ratio  $N_{\text{data}}/N_{\text{parameters}}$  is large enough (at least greater than 1). Hence, the total volume occupied by these dummy scatterers should be small relative to the unit cell and the contribution of the rest of the scattering model (stereochemical atomic model) should be reasonably correct.

The dummy atoms are placed on a fine grid whose spacing should be 1 Å or less, depending on resolution, and every node inside a chosen region is occupied by a dummy atom. The initial atomic parameters correspond to weak electron density, *i.e.* the atom type is an H atom and occupancies are assigned equal small values (usually between 0.1 and 0.5). The atoms of the stereochemical

Table 1. Mean density values (in  $\sigma$ ) corresponding to the main-chain ( $N$ ,  $C\alpha$ ,  $C$ ) atomic position of the refined model

The phases for the initial map were calculated from an available atomic model without residues 2–6.

Residue	2	3	4	5	6	7	8
Initial map	0.46	0.70	0.56	1.03	1.66	2.83	2.90
Improved map	0.83	0.71	1.33	2.63	2.33	2.43	2.60

model which are inside this region are removed from further calculations.

Optimization of dummy models (let the term refinement be reserved for stereochemical models in order to avoid possible confusion) is an important step in the phase improvement (Lunin & Urzhumtsev, 1984). In the current procedure, the stereochemical atomic model is not refined at all. The most important parameters to be fitted are the occupancies of dummy atoms (one parameter per atom). Most of the occupancies do not increase; however some of them can reach quite significant values (2–4). After the initial fitting, the dummy atoms with low occupancies (usually about 60–90% of atoms) can be removed from the model. This reduction of the number of dummy atoms is important to have a reasonably high ratio  $N_{\text{data}}/N_{\text{parameters}}$  for the subsequent positional fitting, where every atom has three parameters and no geometrical restraints are used. A higher ratio should lead to a correct minimum of the criterion in a faster way. Note that dummy-atom parameters are used only as a tool to recalculate the density and have no relation to the stereochemical model and that the starting dummy-atom component has relatively weak contribution to structure factors. Therefore, a low value of the ratio (about 2 or even less) should not result in an introduction of incorrect map details, but it could slow down the convergence of the procedure.

This occupancy and positional fitting procedure is iterated several times while the low-resolution limit of the data is slowly increased. Usually, five to six iterations suffice, with 20–30 refinement cycles per iteration. It is important to include all low-resolution data in the initial iteration in order to form the correct shape of the dummy-atom field. However, it is advisable, as one approaches final iterations, to restrict the low-resolution limit to 6–7 Å to avoid the problems associated with bulk-solvent contribution. An alternative could be to use all data and apply a structure-factor correction for bulk solvent in all cycles.

Once the model optimization is completed, structure-factor amplitudes are calculated from the final model and used to estimate errors in the corresponding calculated phases (Lunin & Skovoroda, 1995; Urzhumtsev, Skovoroda & Lunin, 1996). Subsequently, a weighted electron-density map in the form suggested by Read (1986) can be calculated using all available structure factors.

The procedure is performed with the program *X-PLOR* (Brünger, 1992) and requires an atomic model and a set of structure factors (with FOBS, SIG, TEST values) as starting data. The unit-cell region is defined as a set of spheres of a given radius. Two additional FORTRAN programs are used to generate dummy atoms

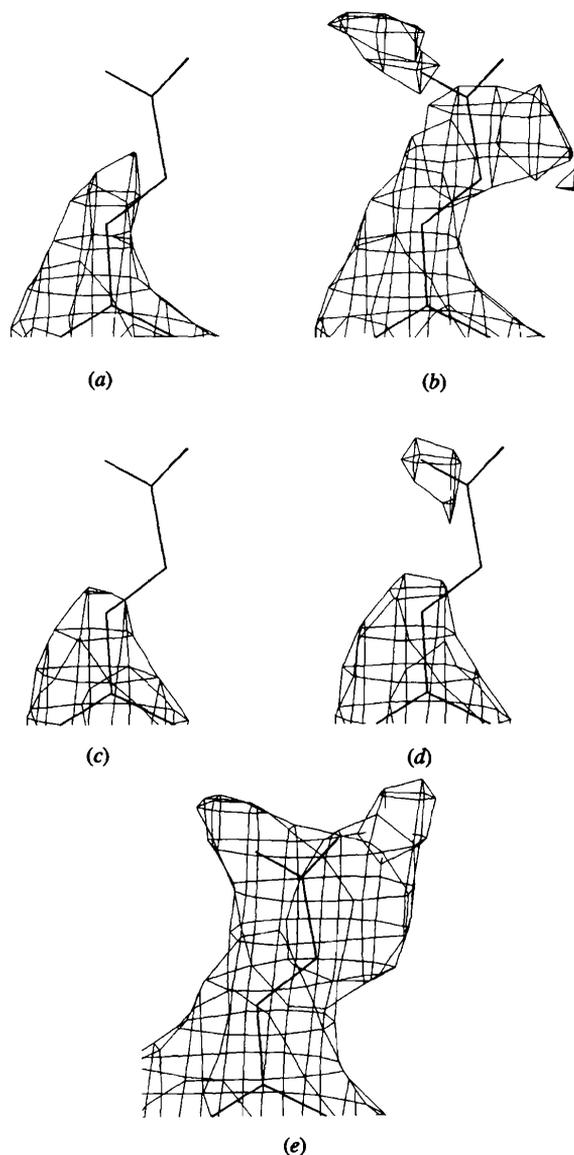


Fig. 1. Electron-density maps corresponding to the Glu174 of the Y5S mutant of the N-terminal 43 K domain of gyrase B. The side chain of Glu174 was excluded from all calculations. The main chain of Glu174 was excluded from the calculations of the difference density map (c and d). The final model is superposed on all maps. The program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991) was used for graphic analysis and figure preparation. (a)  $2F_o - F_c$ -weighted map (2.3–20.0 Å) contoured at  $1.0\sigma$ ; (b)  $2F_o - F_c$ -weighted map (2.3–20.0 Å) contoured at  $0.7\sigma$ ; (c)  $F_o - F_c$  difference map (2.3–6.0 Å) contoured at  $3.0\sigma$ ; (d)  $F_o - F_c$  difference map (2.3–6.0 Å) contoured at  $2.5\sigma$ ; (e) improved  $2F_o - F_c$  weighted map (2.3–20.0 Å) contoured at  $1.0\sigma$ .

and to estimate phase errors; these are available from the author by e-mail.

### 3. Applications of the procedure

The density growing procedure has been successfully applied to the structure of the N-terminal (43 K) fragment of the mutant Y5S gyrase B (Brino *et al.*, 1997). The mutant protein crystallized in the space group  $P2_12_12_1$ , with unit-cell dimensions  $a = 84.7$ ,  $b = 137.4$ ,  $c = 78.9$  Å, and with two monomers per asymmetric unit. 33 856 diffraction amplitudes were measured in the resolution range from 20 to 2.3 Å. The procedure was applied when residues 6–392 of the model were placed in density and refined an  $R$  factor of 0.21. The N-terminal residues and some of the side chains in the regions of close intermolecular contacts needed to be rebuilt.

One of such regions included residues 174–179 where, in particular, the position of the side chain Glu174 in both monomers was not clear either in the  $2F_o - F_c$  or in the difference  $F_o - F_c$  map (Figs. 1*a*–1*d*). A region for improvement was defined by a set of 6 Å radius spheres around the C $\alpha$  atoms of residues 174–179 of every monomer (plus a sphere around residue 74 whose side chain was also analysed). 6621 dummy H atoms, all with occupancy equal to 1 and temperature factor equal to 20 Å<sup>2</sup>, were placed on a 1 Å grid within the region, whereas the 318 protein atoms within the same region were removed. Five iterations of the occupancy/positional fitting were carried out (protein atoms outside the region were fixed), starting with a resolution range of 2.3 Å to infinity and finishing with the resolution range of 2.3–6.0 Å (30 656 reflections). The map calculated using the mixed model (protein atoms outside the region plus 975 dummy atoms with optimized parameters) showed clear density for the side chain at the level of  $1\sigma$  (Fig. 1*e*).

A similar procedure was applied for the N-terminal regions, where residues 2–6 of the native model were excluded. It improved the contrast of the density (Table 1) and made the density continuous, which allowed us to rebuild the residues. It is interesting to note that the improved density also showed four strong peaks corresponding to water molecules, for which the initial map had a very weak or practically no signal. Since the procedure is local, general characteristics are not useful. Instead, some local ones can be applied to illustrate the result (Table 1).

A successful application of the procedure to the high-resolution data (1.4 Å) for the rac1 protein (Hirshberg, Stockley, Dodson & Webb, 1997), where it allowed the building of 14 residues, will be discussed elsewhere.

### 4. Conclusions

A map calculated using the described procedure showed improved density in several cases of missing side

chain(s), flexible loops, poor density at the N terminus or poor density for an inhibitor in protein-inhibitor complexes. Unfortunately, it was not possible to use the positions of dummy atoms as a guide for the construction of the missing part of the molecular model.

The method can be used together with other methods for density improvement especially with those which could retrieve systematically missing structure factors (naturally, with special care for correction of the bulk solvent contribution at low resolution).

This procedure can also supplement the classical procedure of calculating omit maps (Bhat & Cohen, 1984) by filling with dummy atoms the omit zone, which is normally empty. It is important, however, that the known parts of the model undergo simulated-annealing refinement (Hodel, Kim & Brünger, 1992) before fitting the parameters of dummy atoms.

The procedure currently uses the knowledge of the interpreted part of electron density in the form of structure factors calculated from the model. It is conceivable that the procedure can also be used in some special cases of local density improvement when the model is not yet available and the structure factors from the rest of the unit cell are calculated directly from the electron-density map.

The author thanks L. Moulinier for his help in testing the method, V. Lunin and A. Podjarny for many discussions on the dummy-atoms ideas, and D. Moras for his support of the work. The author also thanks J. Arnez who carefully read the manuscript and corrected the language and the text. The structure of the gyrase mutant was solved in collaboration with L. Brino, P. Oudet, A. Mitschler and D. Moras. The work was supported by the CNRS through the UPR 9004, by the Institut National de la Santé et de la Recherche Médicale and the Centre Hospitalier Universitaire Régional. The referees criticism was very useful for the improvement of the manuscript.

### References

- Agarwal, R. C. & Isaacs, N. W. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 2835–2839.
- Baker, D., Bystrhoff, C., Fletterick, R. J. & Agard, D. A. (1993). *Acta Cryst.* **D49**, 429–439.
- Bhat, T. N. & Cohen, G. H. (1984). *J. Appl. Cryst.* **17**, 244–248.
- Brino, L., Urzhumtsev, A. G., Dupuis-Hamelin, C., Mitschler, A., Bronner, C., Mauvais, P., Klich, M., Moras, D. & Oudet, P. (1997). In preparation.
- Brünger, A. T. (1992). *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR*. Yale University, CT, USA.
- Bystrhoff, C., Baker, D., Fletterick, R. J. & Agard, D. A. (1993). *Acta Cryst.* **D49**, 440–448.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.

- Hirshberg, M., Stockley, R. W., Dodson, G. & Webb, M. R. (1997). *Nature Struct. Biol.* **4**, 147–151.
- Hodel, A., Kim, S.-H. & Brünger, A. T. (1992). *Acta Cryst.* **A48**, 851–858.
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* **D49**, 129–147.
- Lunin, V. Yu. & Skovoroda, T. P. (1995). *Acta Cryst.* **A51**, 880–887.
- Lunin, V. Yu. & Urzhumtsev, A. G. (1984). *Acta Cryst.* **A40**, 269–277.
- Lunin, V. Yu., Urzhumtsev, A. G., Vernoslova, E. A., Chirgadze, Yu. N., Nevskaya, N. A. & Fomenkova, N. P. (1985). *Acta Cryst.* **A41**, 166–171.
- Main, P. (1979). *Acta Cryst.* **A35**, 779–785.
- Raman, S. (1959). *Acta Cryst.* **12**, 964–975.
- Read, R. (1986). *Acta Cryst.* **A42**, 140–149.
- Urzhumtsev, A. G., Skovoroda, T. P. & Lunin, V. Yu. (1996). *J. Appl. Cryst.* **29**, 741–744.
- Watenpaugh, K. D., Sieker, L. C., Herriott, J. R. & Jensen, L. H. (1973). *Acta Cryst.* **B29**, 943–956.
- Wilson, C. & Agard, D. A. (1993). *Acta Cryst.* **A49**, 97–104.