IIJIMA, & ALLPRESS, J. G. (1974b). *Acta Cryst.* A30, 30–36.

IIJIMA, S. & COWLEY, J. M. (1978). *J. Phys. (Paris), 38, Suppl. CF.* 135-144.

IIJIMA, S. & O'KEEFE, M. A. (1978). In preparation.

MAGNÉLI, A. (1953). *Acta Cryst.* 6. 495–500.

O'KEEFE, M. A., BUSECK, P. R. & IIJIMA, S. (1978). *Nature (London).* In the press.

ROTH, R. S. & WADSLEY, A. D. (1965). *Acta Cryst.* 19, 26–32.

SLEIGHT, A. W. (1966). *Acta Chem. Scand.* 20, 1 102–1112.

STONEHAM, A. M. & DURHAM, R. J. (1973). *J. Phys. Chem. Solids, 34, 2* 12 7-2 13 5.

WADSLEY, A. D. & ANDERSSON, S. (1970). *Perspect. Struct. Chem. 3,* 1-58.

Refinement of Large Structures by Simultaneous Minimization of Energy and R Factor

BY A. JACK* AND M. LEVITT

MRC Laboratory of Molecular Biology, University Postgraduate Medical School, Hills Road, Cambridge CB2 2QH, England

An improved method of crystallographic structure refinement, especially suitable for large molecules, is described. It is based on simultaneous minimization of a realistic potential-energy function and a crystallographic residual. The method has already proved its worth in the final stages of refinement of two structures; an application to crude wire-model coordinates of a small protein is described and evaluated.

Introduction

Until recently the refinement of macromolecular structures against X-ray diffraction data was a laborious process, involving many comparisons of a model with a difference map, followed by manual correction of the model, coordinate idealization, and further refinement. The traditional method, Diamond's (197 1) real-space refinement, treats torsion angles (and possibly some bond angles) as variables; it has a wide range of convergence (Diamond, 1976), but suffers from the disadvantage that it is tied to a particular set of 'observed' phases; that is, the function minimized is

$$\int_v (\rho_o - \rho_c)^2 \, dV \equiv \frac{1}{V} \sum (F_o - F_c)^2, \qquad (1)$$

whereas we should ideally seek to minimize

$$A = \sum_h (|F_o| - I |F_c|)^2. \qquad (2)$$

We can approach this ideal with Diamond's method by using a new map for each cycle, the coefficients of which are $|F_o| \exp(i\alpha_c)$ or $(2|F_o| - |F_c|) \exp(i\alpha_c)$. Nonetheless, this is expensive and still biased toward the starting structure, so that convergence eventually becomes slow.

(2) may be minimized by conventional crystallo-

* Died 14 July 1978.

graphic least squares, but only if very high resolution data are available (see, for example, Watenpaugh, Sieker, Herriott & Jensen, 1973). An alternative method is to compute unconstrained shifts, either by least squares or from a difference map, using the gradient/curvature method (Freer, Alden, Levens & Kraut, 1976) to apply these shifts for several cycles, and then to re-idealize the bond lengths and angles. This method also works well with high-resolution (<2 Å) data.

Consider now the problem of refining a rather poorly determined structure, in which atoms may be misplaced by several angstroms. The radius of convergence of least-squares methods is dependent on the data resolution (typically *d/4* where *d* is the spacing of the highest-order reflexion). Thus to have any hope of correcting large errors automatically, we must start with low-resolution data. Unconstrained methods are unlikely to work in this case, since the shifts will be so large as to alter the stereochemistry in a way that may not easily be reversed by idealization. The solution is to use a least-squares technique in which constraints (or restraints) on the stereochemistry are included. Two such methods have recently been reported. Konnert (1976) uses Waser's (1963) method to impose quadratic constraints on bond lengths and angles, and solves the resulting sparse set of equations by the method of conjugate gradients. Sussmann, Holbrook, Church & Kim (1977) have used a mixture of constraints and

restraints, with similar results. Both methods are some-what slow (Schmidt, Girling & Amma, 1977), and neither seems to take any account of close non-bonded contacts. The method proposed here minimizes a complete potential-energy function (Levitt, 1974) which includes terms for bond stretching, bond-angle bending, torsional potentials, and non-bonded and electrostatic forces. The crystallographic residual (2) and its deriva-tives with respect to positional parameters are com-puted externally, and used as input to a modified version of Levitt's energy-minimization program which minimizes $E + k\Delta$ by the conjugate-gradients method. The relative contributions of the X-ray and energy terms to the residual are controlled by the constant k. The best choice of this parameter, so as to produce the most nearly 'correct' structure, is something of a problem; we have chosen values empirically which make E and $k\Delta$ numerically comparable, so that both terms are reduced during minimization. The method is very fast (see below), and has the great advantage of allowing a very flexible specification of restraints.

Computational methods

The standard conjugate-gradients minimization algorithm (Fletcher & Reeves, 1964) recomputes the function E and its first derivative many times during a round of refinement, so that it is not limited to strictly quadratic functions. A is very expensive to compute, and so we calculate it and its derivatives at the start of each round, and make the usual least-squares assumption that F_c is a linear function of the positions x_i. Thus

$$\frac{\partial \Delta}{\partial x_i} = -2 \sum_h (|F_o| - |F_c|) \frac{\partial |F_c|}{\partial x_i}$$

$$\frac{\partial^2 \Delta}{\partial x_i \partial x_j} = 2 \sum_h \frac{\partial |F_c|}{\partial x_i} \frac{\partial |F_c|}{\partial x_j} .$$

On each iteration of the minimization, Δ and its first derivative are updated according to

$$\Delta = \Delta_0 + \sum_i \frac{\partial \Delta}{\partial x_i} \delta x_i + \frac{1}{2} \sum_i \sum_j \frac{\partial^2 \Delta}{\partial x_i \partial x_j} \delta x_i \delta x_j \quad (3)$$

and

$$\frac{\partial \Delta}{\partial x_i} = \left(\frac{\partial \Delta}{\partial x_i}\right)_o + \sum_j \frac{\partial^2 \Delta}{\partial x_i \partial x_j} \delta x_j$$

and the updated estimates are used for further iterations. At the end of the process, the actual value of A is recomputed and is found to agree more or less with the value estimated from (3), depending on the validity of the assumption of linearity.

A and its derivatives were initially computed with a conventional least-squares program using the 'all-planes-one-atom' algorithm (Rollett, 1965; Burnett &

Nordman, 1974). The only off-diagonal terms of the second derivative to be calculated were those relating different parameters of the same atom. Despite this, the calculation becomes very time-consuming at high resolution, and was abandoned in favour of a least-squares method based on the fast Fourier transform algorithm. The method used is a slight modification of that developed by Agarwal (1978) and will not be described in detail here. It is based on Cochran's (1948) observation that shifts computed from a difference density map (the gradient/curvature method) are equivalent to those found by diagonal least-squares refinement, where each term in the normal matrix is weighted by the inverse of the atomic form factor. This weighting factor may be removed by convoluting the gradient of the difference map (around the atomic site) with the transform of the form factor, that is, with a modified electron-density profile. In our procedure the gradient of the difference map is computed by numerical difference rather than by computing three differential syntheses (one for each direction); other-wise our procedure seems equivalent to that of Agarwal (1978).

The least-squares algorithm calls for random access to a large difference density map. It can be made extra-ordinarily fast either by holding the whole of the map in computer core or by pre-sorting the atomic coordinates so that only a few map sections are needed at any one time. In the interests of generality, however, we have preferred to use a double-sort technique similar to that used to average electron-density maps (Harrison & Jack, 1975; Bricogne, 1976). Only the diagonal terms of the second derivative are calculated; although in principle the method can handle off-diagonal terms, the amount of computation would become prohibitive. Derivatives of Δ with respect to occupancy and temperature factor may also be computed; these use the difference map itself rather than its gradient.

Results

The method was originally devised for the structure refinement of tRNA (Hingerty, Brown & Jack, 1978), where the conflict between good base-pairing geometry and a good fit to the X-ray data had made real-space refinement of little further use (Jack, Ladner & Klug, 1976). Here it was very successful, reducing R from 0·26 to 0·23 in a single round, after the real-space method had apparently converged. More recently, an application to oxymyoglobin (Phillips, 1978), starting from real-space-refined Met-myoglobin coordinates (Takano, 1977) reduced R from 0·32 to 0·20 at 2 Å resolution.

To what extent is the method useful for refining crude coordinates such as those measured from a wire model? The structure of bovine pancreatic trypsin

Table *1. Summary of refinement*

Cycle	Resolution (Å) and (number of F's)	Energy (kcal mol⁻¹) Bond	Angle	Torsion	Non-bond	k	A (e² x 10⁸)	R	R.m.s. movement (Å)
0 (WM)	1·5 (7952)						0·562	0·511	
1	5·0 (227)	293	186	164	5384	0·0002	0·177	0·473	
		202	468	186	-1214		0·072	0·345	0·27
2	4·0 (454)						0·199	0·386	
		68	255	164	-1395	0·0001	0·180	0· 349	0·31
3	3·4 (679)						0·248	0·357	
		78	290	171	-1381	0400 1	0·192	0·3 15	0·38
4	3·0 (1078)						0·205	0·384	
		54	236	166	-1436	0·0001	0· 188	0·348	0·41
5	2·8 (1325)						0·186	0·370	
		52	223	172	-1445	o-000 1	0· 183	0·358	0·46
6	2·5 (1848)						0·200	0·400	
		127	336	198	-1419	0·0002	0·156	0·355	0·54
7	2·2 (2677)						0·181	0·382	
		103	316	194	-1481	0·0002	0·192	0·364	0·57
8	2·0 (3539)						0·214	0·387	
		139	340	198	-1457	0·00025	0·202	0·369	0·60
9	1·75 (5164)						0·224	0·394	
		145	332	192	⟶1490	0·00025	0·211	0·381	0·61
10	1·5 (7952)						0·284	0·401	
		157	332	191	-1501	0·00025	0·276	0·39 1	0·63
11	1·5 (7952)						0·149	0·363	*B* only
12	1·5 (7952)						0·141	0·357	*B only*
13	1·5 (7952)	762	603	247	98	0·001	0·129	0·335	0·896
14	1·5 (7952)	133	393	264	⟶1400	0·0005	0· 133	0·337	0·892

The decrease in A between cycles 4 and 5 is a result of re-scaling F_o to F_c. H atoms (at positions calculated assuming exact trigonal or tetrahedral heavy-atom coordination) were included in the energy calculation (but not the structure factor calculation). Later trial calculations suggested that this was an unnecessary sophistication.

inhibitor (PTI) (Deisenhofer & Steigemann, 1975, here-after D&S) is a good test example, since it is small (454 non-hydrogen atoms), and exhaustively refined coor-dinates are available which serve as a standard for comparison with any other sets.

Structure factors computed from the wire-model coordinates of PTI gave $R = 0·5$ 11 for 7952 reflexions with spacings between 7 and 1·5 Å. Ten rounds of $E + k\Delta$ minimization, gradually increasing the data resolution, decreased R to 0·39 1 (Table 1). This compares well with the value of $0·384$ which is the best value obtained by D&S using real-space refinement without study of difference maps. After two cycles of refinement of individual temperature factors (R = 0·357), a difference map clearly showed most of the conformational errors mentioned by D&S (§ 3.7). At this stage the r.m.s. atomic movement for all atoms was 0·63 Å. The largest error which had been corrected was 2·1 Å (NE, of Gln3 1); on the other hand, a 3·6 Å error in the position of Sδ of Met52 (caused by a 180° error in the torsion angle χ^2) was not corrected, although the true position was obvious from the difference map.

Table 1 summarizes the course of refinement. It can be seen that the most dramatic changes occurred in the first round (5 Å resolution), and also when the resolution was increased from 2·8 to 2·5 Å. In both cases a large value of the scale factor k was used; k was
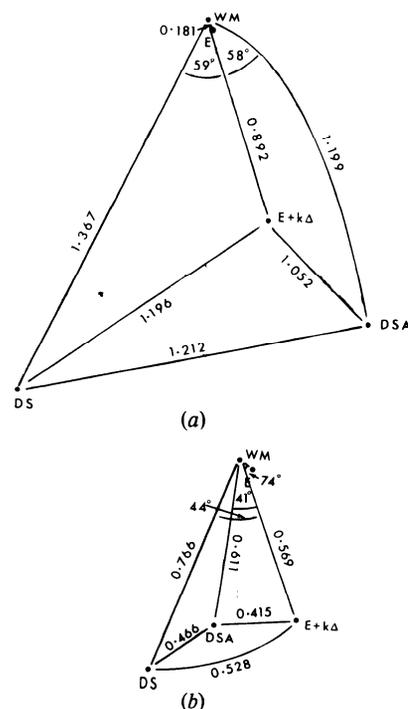

Fig. 1. Shift diagrams illustrating the relationship between coordin-ates generated by different refinement methods for *(a)* all atoms, and *(b)* main-chain atoms only.

halved for the other early rounds to prevent the energy E increasing. Three side-chain errors discovered from the difference map were corrected by hand at cycle 12, when two further cycles reduced R to 0·337. The refinement was not pursued further since the success of the method had become clear.

Fig. 1 compares five different sets of coordinates: WM, the starting wire-model set ($R = 0·511$); DSA, the best coordinates obtained by D&S without calculation of difference maps ($R = 0·384$); DS, the final coordinates of D&S ($R = 0·308$ without individual B's or water molecules); $E + k\Delta$, the coordinates obtained by the method described here ($R = 0·337$ with individual B's, $R = 0·358$ without); and E, the result of energy minimization alone on WM. The method of comparison is that introduced by Diamond (1974): the figures are the principal projections of sets of five points in $3N$-dimensional space (for N atoms). The numbers joining any two vertices give the r.m.s. difference (in A) between the two coordinate sets, and the cosine of any marked angle is the correlation coefficient between two sets of shifts. We note that energy minimization alone (set E) gives very small shifts (r.m.s. movement 0·18 A) which are almost orthogonal to those obtained by any method of refinement against the X-ray data; $E + k\Delta$ shifts, on the other hand, are much larger, and are highly correlated with those of sets DSA and DS, although the final energy is still comparable with that of set E. Table 2 gives the R factors and values of A for the various sets.

Table 2. *R factors of various coordinate sets, with and without individual atomic B's*

	Without $B's$		With $B's$	
	A (x 10^8)	R	A (x 10^8)	R
WM	0·562	0·511		
$E + k\Delta$	0·153	0·358	0·133	09337
DSA	0·194	0·384		
DS	0·116	0·308		

At first sight it is disturbing that $E + k\Delta$ is not much closer to DS (assumed to be the 'correct' answer) than is WM, and that $E + k\Delta$ and DSA are so far apart. However, the *distribution* of errors is reassuring. Fig. 2

shows histograms of the coordinate errors (relative to DS) for sets WM, DSA and $E + k\Delta$. We see that the great majority of the atoms in sets DSA and $E + k\Delta$ are within 0·3 Å of DS, suggesting that both would automatically converge to the same solution on further refinement. The large apparent distance between DSA and $E + k\Delta$ arises mainly from the large (>2 A) errors which have not been corrected by either refinement method. These are the very errors which should be most obvious from a difference map.

Comparing torsion angles of the three sets of coordinates, we find that the largest differences between the starting angles and those of D&S occur in the peptide linkages at residues 2-3, 25-26 and 55-56. Table 3 compares the three sets of angles in these regions, and we see that the largest errors (>110°) have not been corrected, whereas changes of *ca* 60° have occurred automatically. The range of convergence is
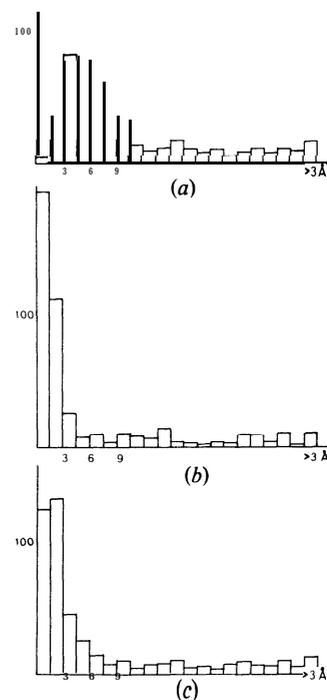

Fig. 2. Histograms of coordinate errors (i.e. shifts relative to set DS) for (a) WM, (b) DSA, and (c) $E + k\Delta$.

Table 3. *Comparison of three peptide conformations*

	Pro-Asp3			Ala25–Lys26			Cys55–Gly56		
	Start	D & S	This work	Start	D & S	This work	Start	D & S	This work
φ	12°	151°	−21°	25"	-28"	−10°	-81"	-4"	-38"
ω	180	177	173	180	170	177	180	190	205
ψ	55	-60	106	-136	-66	-76	-9	-77	-61

therefore comparable to that of real-space refinement, which is able to correct errors of up to twice the atomic radius.

Computational needs

Central-processor times on an IBM 370/165 computer are 1·5 min to compute the X-ray normal matrix (including difference map calculation), 1 min for 30 cycles of conjugate-gradient minimization without H atoms (1500 variables), or 2 min with H atoms (2700 variables), and 2 min for the structure factor calculation. The latter may be reduced to about 0·5 min by constructing a model electron-density map and computing its Fourier transform (Ten Eyck, 1977), although we have not done this in the present work. The programs which compute the normal matrix need 100 kbytes of core (using the IBM sort/merge package for the sort steps); the minimization program can handle up to 1500 atoms (4500 variables) in 260 kbytes.

The minimization routine used is a modified version of the Harwell program *VA08A*. A routine with automatic restarts (*VA* 14A; Powell, 1975) usually finds a lower minimum in fewer cycles, but requires too much storage to make its use generally worth while.

Conclusions

The refinement method described here combines advantages of three other methods: real-space refinement, gradient/curvature refinement, and conventional least squares. It has the speed of gradient/curvature methods, whilst maintaining proper stereochemistry at all stages. Because a new map is computed on each round, the function minimized is equivalent to a difference of *moduli*; the phases are not treated as observations. The convergence radius may be made comparable to that of real-space refinement if the process is started with low-resolution data. The convolution of the difference-map gradient with an electron-density profile makes the process equivalent to one of volume fitting rather than peak-maximum fitting, and is thus capable of increasing the theoretical radius of convergence (Diamond, 197 1). Since the convolution is calculated as a sum over grid points, the correction discussed by Diamond (197 1, p. 440) should be applied; we have not done this, but if the difference map is computed on a sufficiently fine grid the errors are no more than a few per cent.

The work on PTI has shown that positional errors of up to 2 or 3 Å may be corrected automatically. Regions of the molecule where errors were greater are still incorrect, but the overall improvement of the model means that difference maps are much more easily interpretable. The method has also proved valuable in the final stages of refinement when convergence of the real-space method has become slow. Unfortunately there is still no sign of any short cut through the intermediate stages: the only way to be certain that no part of the structure has refined into a false minimum is to make a systematic study of difference maps, possibly omitting doubtful regions of the molecule from the structure factor calculation.

References

AGARWAL, R. (1978). *Acta Cryst.* A**34**, 791–809.
BRICOGNE, G. (1976). *Acta Cryst.* A**32**, 832–847.
BURNETT, R. M. & NORDMAN, C. J. (1974). *J. Appl. Cryst.* **7**, 625–627.
COCHRAN, W. (1948). *Acta Cryst.* **1**, 138-142.
DEISENHOFER, J. & STEIGEMANN, W. (1975). *Acta Cryst.* B**31**, 238–250.
DIAMOND, R. *(197* 1). *Acta Cryst.* A**27**, 436–452.
DIAMOND, R. (1974). *J. Mol. Biol.* **82**, 37 1-39 1.
DIAMOND, R. (1976). *Crystallographic Computing Techniques,* edited by F. R. AHMED, pp. 29 1-301. Copenhagen: Munksgaard.
FLETCHER, R. & REEVES, C. M. (1964). *Comput. J. 7,* 149–154.
FREER, S. T., ALDEN, R. A., LEVENS, S. A. & KRAUT, J. *(1976). Crystallographic Computing Techniques,* edited by F. R. AHMED, pp. 3 17-32 1. Copenhagen: Munksgaard.
HARRISON, S. C. & JACK, A. (1975). *J. Mol. Biol. 97,* 173–191.
HINGERTY, B. E., BROWN, R. S. & JACK, A. (1978). *J. Mol. Biol.* Submitted.
JACK, A., LADNER, J. E. & KLUG, A. (1976). *J. Mol. Biol.* **108**, 6 19-649.
KONNERT, J. H. (1976). *Acta Cryst.* A**32**, 614–617.
LEVITT, M. (1974). *J. Mol. Biol.* **82**, 393–420.
PHILLIPS, S. E. V. (1978). In preparation.
POWELL, M. J. D. (1975). Report CSS24, AERE, Harwell.
ROLLETT, J. S. (1965). *Crystallographic Computing,* pp. 38–46. Oxford: Pergamon Press.
SCHMIDT, W. C., GIRLING, R. L. & AMMA, E. L. (1977). *Acta Cryst.* B33, 36 18-3620.
SUSSMANN, J. L., HOLBROOK, S. R., CHURCH, J. M. & KIM, S.-H. (1977). *Acta Cryst.* A**33**, 800–804.
TAKANO, T. (1977). *J. Mol. Biol.* **110**, 537–568.
TEN EYCK, L. F. (1977). *Acta Cryst.* A**33**, 486–492.
WASER, J. (1963). *Acta Cryst.* 16, 1091-1094.
WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1973). *Acta Cryst.* B**29**, 943–956.