GRIMM, H., AXE, J. D. & KRÖHNKE, C. (1982). *Phys. Rev. B*, **25**, 1709-1716.
GUINIER, A. (1958). *Bull. Soc. Fr. Minéral. Cristallogr.* **78**, 680-710.
GUINIER, A. (1963). *X-ray Diffraction in Crystals, Imperfect Crystals and Amorphous Bodies.* New York: Freeman.
HENDRICKS, S. & TELLER, E. (1942). *J. Chem. Phys.* **10**, 147-167.
HOSEMANN, R. & BAGCHI, S. N. (1954). *Phys. Rev.* **94**, 71-74.
HUANG, K. (1963). *Statistical Mechanics.* New York: Wiley.
HUSTER, M. E., HEINEY, P. A., CAJIPE, V. B. & FISCHER, J. E. (1987). *Phys. Rev. B*, **35**, 3311-3326.
JAGODZINSKI, H. (1987). *Prog. Cryst. Growth Charact.* **14**, 47-102.
JAMES, R. W. (1982), *The Optical Principles of the Diffraction of X-rays.* Woodbridge, CT: Ox Bow Press.
JOHNSTON, D. C. & FRYSINGER, S. P. (1984). *Phys. Rev. B*, **30**, 980-984.
KAKINOKI, J. (1967). *Acta Cryst.* **23**, 875-885.
KAKINOKI, J. (1983). *Acta. Cryst.* A**39**, 171-173.
KAKINOKI, J. & KOMURA, Y. (1952). *J. Phys. Soc. Jpn*, **7**, 30-35.
KAKINOKI, J. & KOMURA, Y. (1965). *Acta Cryst.* **19**, 137-147.
KASTELEYN, P. W. (1971). *Trieste Lectures* 1970. Vienna: International Atomic Energy Agency.
KIM, H. J., FISCHER, J. E., MCWHAN, D. B. & AXE, J. D. (1986). *Phys. Rev. B*, **33**, 1329-1339.
KOMURA, Y. & KITANO, Y. (1977). *Acta Cryst.* B**33**, 2496-2501.
LANDAU, L. (1937). *Phys. Z. Sowjetunion*, **12**, 579-585.
LIFSCHITZ, I. M. (1937). *Phys. Z. Sowjetunion*, **12**, 623-643.
LIFSCHITZ, I. M. (1939). *J. Exp. Theor. Phys. USSR*, **9**, 500-509.
MAIRE, J. & MÉRING, J. (1970). In *Chemistry and Physics of Carbon*, Vol. 6, edited by P. L. WALKER. New York: Marcel Dekker.
MÉRING, J. (1949). *Acta Cryst.* **2**, 371-377.
MONCTON, D. E., DISALVO, F. J., AXE, J. D., SHAM, L. J. & PATTON, B. R. (1976). *Phys. Rev. B*, **14**, 3432-3457.
PROST, J. & BAROIS, P. (1983). *J. Chim. Phys.* **80**, 65-81.
REYNOLDS, R. C. (1980). *Crystal Structures of Clay Minerals and Their X-ray Identification*, edited by G. W. BRINDLEY & G. BROWN. Mineral. Soc. Monogr. no. 5.
RIORDAN, J. (1968). *Combinatorial Identities.* New York: Wiley.
SEUL, M. (1988). *Phys. Rev. Lett.* **60**, 1150-1153.
SEUL, M. & EISENBERGER, P. (1989). *Phys. Rev. A.* In the press.
SLATER, L. J. (1960). *Confluent Hypergometric Functions.* Cambridge Univ. Press.
WELBERRY, T. R. (1985). *Rep. Prog. Phys.* **48**, 1543-1593.

Acta Cryst. (1989). A**45**, 396-409

X-ray Refinement of Protein Structures by Simulated Annealing: Test of the Method on Myohemerythrin

BY JOHN KURIYAN

Laboratory of Bioorganic Chemistry and Biochemistry, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

AXEL T. BRÜNGER*

Howard Hughes Medical Institute, Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA

MARTIN KARPLUS

Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

AND WAYNE A. HENDRICKSON

Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168th Street, New York, NY 10032, USA

(*Received* 23 *August* 1988; *accepted* 3 *January* 1989)

Abstract

The recently developed method of structure factor refinement by molecular dynamics with simulated annealing [Brünger, Kuriyan & Karplus (1987). *Science*, **235**, 458-460] is tested on the 118 residue protein myohemerythrin. A highly refined structure for this protein at 1·3/1·7 Å resolution has recently been published [Sheriff, Hendrickson & Smith (1987). *J. Mol. Biol.* **197**, 273-296]. This is compared with the results of simulated annealing refinement (with no manual intervention) starting from an earlier model for the protein from a stage in the refinement when conventional least-squares methods could not improve the structure. Simulated annealing reduces the *R* factor at 2·5 Å from 39 to 31%, with uniform temperature factors and no solvent molecules and with similar stereochemistry; the comparable value for the manually refined structure is 27·9%. Errors in backbone and sidechain positions up to about 3 Å are corrected by the method. The error in backbone positions for roughly 85% of the initial structure is within this range, and in these regions the r.m.s. backbone error is reduced from 1·1 to 0·4 Å. For the

0108-7673/89/060396-14$03.00

rest of the structure, including a region which was incorrectly built due to a sequence error, the procedure does not yield any improvement and manual intervention appears to be required. Nevertheless, the overall improvement in the structure results in electron density maps that are easier to interpret and permit identification of the errors in the structure. The general utility of the simulated annealing methodology in X-ray refinement is discussed.

1. Introduction

The great increase in the number of proteins being studied by X-ray crystallography makes it necessary to develop more rapid methods for refining protein structures. The least-squares refinement methodology that is currently in use (*e.g.* Konnert & Hendrickson, 1980; Jack & Levitt, 1978) is a process that is easily trapped in local minima, resulting in a need for frequent human intervention. This takes the form of manual adjustment of the structure, nowadays using a computer graphics system to display the electron density superimposed on the structural model (Jones, 1985). This part of the process is usually the rate limiting step in the refinement (Deisenhofer, Remington & Steigemann, 1985).

It has recently been shown that the introduction of simulated annealing through molecular dynamics (MD) into the X-ray refinement procedure can reduce the need for manual intervention (Brünger, Kuriyan & Karplus, 1987; Brünger, Petsko & Karplus, 1989; Brünger, 1988*a*, *b*; Gros, Fujinaga, Dijsktra, Kalk & Hol, 1988). The application of molecular dynamics to X-ray refinement problems is a natural extension of the use of high-temperature molecular dynamics and cooling in searching for stable peptide conformations (see, for example, Bruccoleri, 1984; Bruccoleri & Karplus, 1989; Brooks, Pastor & Carson, 1987) and is closely related to the use of molecular dynamics to derive structures from NMR* measurements (Kaptein, Zuiderweg, Scheek, Boelens & van Gunsteren, 1985; Brünger, Clore, Gronenborn & Karplus, 1986). All three are examples of highly non-linear optimization problems. In such cases conventional least-squares optimization will fail unless the starting model is fairly close to optimal. Better convergence is achieved if instead of restricting search directions to be against the energy gradient, one allows the optimization to go uphill as well. This process, implemented as a Monte Carlo algorithm, has been referred to as simulated annealing (Kirkpatrick, Gelatt & Vecchi, 1983); a description of a simple simulated annealing optimization algorithm is given by Press, Flannery, Teukolsky & Vetterling (1986).

Monte Carlo and molecular dynamics simulations are both ways of generating conformations of the system that are consistent with a Boltzmann distribution appropriate to the specified temperature. For large biomolecular structures the molecular dynamics algorithm is generally more efficient at generating equilibrium structures (Northrup & McCammon, 1980). A molecular dynamics rather than a Monte Carlo algorithm was therefore used to implement a 'simulated annealing' procedure for refining protein structures (Brünger *et al.*, 1987, 1989; Brünger, 1988*b*) and we shall refer to this as 'SA refinement'. The SA refinement method has recently been applied to the small protein crambin, starting from a structure derived from simulated NMR data (Brünger *et al.*, 1987, 1989) and to a mutant aspartate amino transferase (MAAT), starting from a model built from MIR phases (Smith, Ringe, Finlayson & Kirsch, 1986; Brünger, 1988*a*). In both cases, the superiority of the SA refinement over conventional least-squares refinement (without manual intervention) was demonstrated. However, in the first example (crambin), the initial structure is generated from simulated NMR data and its relevance to starting structures obtained by fitting electron density maps is not clear. In the second example (MAAT), diffraction data were available only to 2·8 Å resolution and the manual refinement is not yet complete. In this paper we report a test of the method using a protein for which an initial model built using an experimental MIR map is available, and which has been refined to completion at high resolution. Our results provide a realistic evaluation of the SA method.

The protein chosen is myohemerythrin, which is a monomeric oxygen-binding protein from sipunculan worms with 118 amino acid residues. The determination of its structure (Hendrickson, Klippenstein & Ward, 1975) revealed the 4-α helical chain fold, which has since been observed in a number of other proteins. The refinement of a molecular model for this protein has recently been completed (Sheriff, Hendrickson & Smith, 1987); the final model, refined against data to 1·7/1·3 Å resolution with an R value of 15·8%, consists of 979 non-hydrogen protein atoms, with alternative conformations for seven residues. The model for the ordered solvent includes three sulfate ions and 157 water molecules.

The starting model for myohemerythrin was built using an electron density map with phases only to 5·5 Å resolution (Hendrickson & Ward, 1975). This

* Abbreviations: NMR, nuclear magnetic resonance; MIR, multiple isomorphous replacement; r.m.s., root-mean-square; $R_{2\cdot0}$ and $R_{2\cdot5}$, crystallographic R values (expressed as percentages) for data between 2·0 and 7·0 Å and 2·5 and 7·0 Å resolution, respectively; Δb and $\Delta\theta$, r.m.s. deviations of bond lengths and angles from their equilibrium values in the *CHARMM* (Brooks, Bruccoleri, Olafson, States, Swaminathan & Karplus, 1983) force field; ps, fs: picoseconds and femtoseconds; F_c, F_o, calculated and observed structure-factor amplitudes; α_c, calculated phases; δ_1, δ_2, r.m.s. errors in backbone atoms (N, C, C^α), computed over all residues (δ_1) and only over residues 1–10, 15–32 and 45–114 (δ_2); SA refinement, simulated annealing refinement.

polypeptide backbone model was gradually extended to high resolution by restrained refinement (Hendrickson & Konnert, 1981) and model rebuilding. After five iterations of this process the R value stood at 35% at 2·0 Å resolution with quite poor geometry (r.m.s. deviations of bond lengths from ideal values was 0·05 Å). Continued efforts failed to improve this model appreciably. Although some regions, such as the B helix and the dimeric iron center, were essentially correct, other regions clearly contained large errors. The impasse was broken by resorting to experimental phases based on the iron anomalous scattering data to 2·8 Å spacings. Phase ambiguities were resolved by using a model (VIII.K5) that had been refined without water but with individual temperature factors to an R value of 28% at 2·5 Å resolution (Hendrickson, unpublished). The revised model that fitted into this new electron density map refined readily (Smith & Hendrickson, unpublished), and this led eventually to the final model that has been published recently (Sheriff *et al.*, 1987).

The intermediate model VIII.K5 (*i.e.* the model determined before the new experimental phases were obtained) was used to test the SA refinement procedure. This structure represents an early point in the refinement of myohemerythrin, when conventional least-squares optimization was yielding no improvement in the structure and the resulting electron density maps were noisy enough to make rebuilding rather difficult (Hendrickson, unpublished). While the overall chain fold and the placement of the majority of residues in this early model are indeed correct, there are significant (0·5–1·5 Å and greater) errors in backbone and sidechain positions throughout the protein. There are also several regions where even larger errors (3–4 Å in backbone positions) required extensive manual rebuilding before least-squares refinement could continue (Smith & Hendrickson, unpublished). In addition, one of the inter-helix loops was built incorrectly because of an error in the sequence; this resulted in several residues being out of register by one position along the sequence (see below). The improved electron density maps which were obtained after model-resolved anomalous phasing actually led to the identification of this sequence error. In this region, and in a few others, the earlier model has some sidechain positions in error by as much as 8–10 Å. We will refer to this intermediate model as 'INIT' since it was the starting point for the SA refinement.

We have not used the SA procedure to refine the INIT structure to completion. This would require manual intervention in order to place the ordered solvent molecules, to correct the sequence error, and to rebuild regions of especially large positional error which are apparently outside the radius of convergence of the SA method. Knowledge of the final structure is likely to prevent manual intervention from being completely free of bias. We have therefore restricted the present study to running fully automated simulated annealing refinements, starting from the INIT model and varying only parameters such as the resolution of the X-ray data used and the temperature of the molecular dynamics runs. The resulting structures are compared with the final refined structure of myohemerythrin (Sheriff *et al.*, 1987) to evaluate the performance of the SA refinement method.

We first describe the simulated annealing method, the protocols used to carry out the refinements and the model used for the protein. We then discuss the results of the various refinements, first comparing the R values of the annealed structures with those for INIT and the final structure of Sheriff *et al.* (1987). The structural changes brought about by the SA refinement are examined next. An important aspect of refinement which has not yet been automated is the inclusion of ordered solvent atoms in the model. The omission of the solvent molecules causes problems in the refinement of surface sidechain positions and this is discussed in the next section. Finally, the improvement in the electron density maps after SA refinement is demonstrated.

2. Methods

2.1. *Least-squares optimization and simulated annealing*

2.1.1. *The modified energy function.* The standard *CHARMM* (Brooks *et al.*, 1983) empirical energy function was modified as described by Brünger *et al.* (1989) to include the crystallographic residual and the crystal packing interactions:

$$E = E_{\text{charmm}} + E_{\text{packing}} + E_{\text{xray}}. \qquad (1)$$

E_{charmm} is given by the *PARAM*19 version of the *CHARMM* force field (Brooks *et al.*, 1983). No explicit hydrogen-bonding term is included and the hydrogen-bonding energy is accounted for by electrostatic and van der Waals terms (Reiher, 1985). H atoms involved in hydrogen bonding are required to be present for the energy calculation; these were constructed using standard geometry and the initial X-ray structure for the non-hydrogen atoms (Brünger & Karplus, 1988). Only these H atoms were included in the model and none were used in the structure factor computations. The force field was modified slightly as described by Brünger *et al.* (1989) to prevent *cis–trans* isomerization, inversion at chiral centers and buckling of planar groups. In addition, the charges on sidechains of Lys, Arg, Asp and Glu were set to zero. If this is not done, the lack of solvent in the model leads to Coulombic interactions being overemphasized at the surface of the protein, resulting in clustering of oppositely charged sidechains.

E_{packing} is the intermolecular electrostatic and van der Waals energy term computed using the minimum image convention (Brünger *et al.*, 1989). This term is important when running high-temperature dynamics because it prevents surface atoms from moving into density occupied by atoms from a neighboring molecule in the crystal.

E_{xray} is given by

$$E_{\text{xray}} = S \sum_{hkl} w(hkl)(|F_o| - k|F_c|)^2 \qquad (2)$$

where k is a scale factor for F_c, $w(hkl)$ is the weight on a particular reflection and S is a scale factor determining the contribution of the X-ray term relative to the internal and packing energies. Unit weights were used for all reflections. The scale factor, S, is determined by performing a 0·5 ps dynamics run, with the potential $E_{\text{charmm}} + E_{\text{packing}}$ alone, starting from the (minimized) initial structure (Brünger *et al.*, 1989). The resulting structure is used to compute gradients due to $(E_{\text{charmm}} + E_{\text{packing}})$ and E_{xray}. The scale factor S is chosen so as to make the gradients of the same magnitude. The dynamics step is necessary to randomize the structure so as to avoid the unrealistically low gradients that could result if the structure were in a local minimum. The results of the annealing turn out to be not very sensitive to changes in the scale factor that are on the order of 25%. The scale factor is increased slightly during the minimization stages of the refinement. The scale factors used for the refinements described here are 50 000 for the dynamics stages (except for one run, B-1, see below) and 60 000 for the minimization stages.

2.1.2. *X-ray/energy minimization.* In a procedure equivalent to least-squares optimization (*cf. PROLSQ*; Hendrickson, 1985), the pseudo-energy given by (1) is minimized by a conjugate gradient procedure (Jack & Levitt, 1978). Highly vectorized fast Fourier transform routines were used for the computation of the structure factors (Brünger, 1989). As in the Jack & Levitt (1978) refinement, a linear approximation is used for the computation of structure factors (Brünger, 1988*b*). The exact structure factors and structure-factor gradients are computed at the first step of the calculation and are linearly approximated in subsequent steps until at least one atom moves by more than a specified distance (referred to as the "tolerance"); at this point the structure factors and related derivatives are recomputed. Setting tolerance to zero is equivalent to computing structure factors at every step and, although feasible for minimizations, would make molecular dynamics calculations too time consuming. A tolerance of 0·25 Å was found to be satisfactory for the molecular dynamics runs, whereas a tolerance of 0·1 or 0 Å was used in the minimizations.

Pseudo-energy minimizations (including E_{xray}) are performed before and after the molecular dynamics runs. They are necessary at the beginning to remove initial strain in the structure due to bad contacts or stereochemistry. If this is not done, local hot spots are liable to build up in the structure during the high-temperature stage of the calculation and the integration algorithm may break down. After the molecular dynamics run is finished, pseudo-energy minimizations are used to optimize the resulting structure and its stereochemistry, which can be significantly distorted due to the high temperature of the run. The backbone atoms of the structure shift by about 0·3 Å r.m.s. during a typical minimization.

2.1.3. *Simulated annealing refinement protocol.* Molecular dynamics simulations are performed in the normal way (McCammon & Harvey, 1987), except that no heating and equilibration preceded the runs; although the latter are essential for obtaining an equilibrated structure for standard dynamics analysis, this is unnecessary for SA refinement. A trajectory is initiated by assigning a velocity (obtained from a Boltzmann distribution appropriate to the desired temperature) to each atom. Newton's equations of motion are then integrated using a short step size at high temperature (0·25 fs at 3000–9000 K instead of 1·0 fs at 500 K or lower). A tolerance of 0·25 Å is used for the E_{xray} terms during all the dynamics runs. The temperature of the system is checked every 25–50 steps and the velocities are scaled, if necessary, to bring the temperature back to the desired value; this corresponds to coupling the system to a temperature bath with a very short response time. As we show later in this paper, relatively short trajectories (2–3 ps) are sufficient to achieve most of the attainable improvement in the structure. The system is cooled after the trajectory has evolved for the desired length of time by scaling the velocities down to 300 K and letting the trajectory evolve for another 0·5–1 ps at this temperature. Finally, all velocities are set to zero and the energy of the system is minimized, initially with tolerance = 0·1 Å, and finally with tolerance = 0.

Following the standard procedure used in least-squares refinement, we experimented with introducing the X-ray data in stages, starting with low resolution and proceeding to higher resolution with time along the trajectory. There was found to be no apparent advantage to this more complicated procedure.

2.2. *Treatment of temperature factors, solvent and disorder*

The final model for myohemerythrin (Sheriff, Hendrickson & Smith, 1987; from now on referred to as the SHS model) includes individual atomic temperature factors (B factors), 157 water molecules,

three sulfate ions and seven residues with alternative conformations. $R_{2.5}$ and $R_{2.0}$ are 14·1 and 15·9% respectively, using 98·6% of the data to 2·5 Å and 98·1% of the data to 2·0 Å (see below). This structure was obtained by alternating cycles of manual rebuilding and least-squares refinement (Sheriff *et al.*, 1987). Solvent molecules and alternative conformations are not included in the simulated annealing refinements described in this paper. Individual B factors are not used either; instead, all backbone atoms (N, C, C^{α}) are assigned a uniform B factor of 20 $Å^2$ and all other atoms have their B factors set to 24 $Å^2$. These values are close to the average B factors of the SHS model, but the results described here are not sensitive to the chosen values.

Myohemerythrin crystals show a highly anisotropic disorder; the magnitude of the apparent displacement is much larger perpendicular to the helices than along the direction of the helices (Sheriff & Hendrickson, 1987). This effect can be accounted for by an overall anisotropic temperature factor applied to the whole unit cell (or, equivalently, by multiplying the observed structure factors by an anisotropic negative temperature factor). No such correction was applied during two of the simulated annealing runs (A-2 and A-3, see below). For all the other runs an anisotropic correction was applied by multiplying F_o by $\exp(\mathbf{h}^T B \mathbf{h})$ where $\mathbf{h} = (h, k, l)$ and B is the overall temperature factor tensor given by Sheriff & Hendrickson (1987). The structural results are not sensitive to these corrections, although the R values are about 5 to 7 percentage points higher when they are not applied. All R values reported in this paper include the anisotropic correction to F_o of Sheriff & Hendrickson (1987). The R values of the SHS structure calculated without solvent, disorder or individual B factors are $R_{2.5} = 27·9$ and $R_{2.0} = 28·8\%$.

2.3. *Description of different refinements*

The initial model (INIT) has a sequence error in the A-B loop. Residues 34 and 35 had been built as Cys-Asp, in accordance with the chemical sequence determination (Klippenstein, Cote & Ludlan, 1976); later studies (Smith & Hendrickson, unpublished) showed that the correct sequence should be Asp-Cys. This error in the sequence led to an error in the tracing of the backbone chain in this region. The approximate location of the cysteine sulfur was known from a heavy-atom binding site. The A helix was therefore built overwound into a 3_{10} conformation such that residue 34 (wrongly identified as Cys) was placed into the position of residue 35. Subsequent residues in the A-B loop (35-40) are all displaced by one position along the chain, leading to errors on the order of 3-4 Å in backbone atoms.

One of our early aims in performing simulated annealing on myohemerythrin was to see if the refinement would correct the positioning of atoms in the A-B loop. Because of this and also because we did not want to introduce information unavailable to the crystallographer at the time model VIII.K5 was determined, the sequence error was retained in all the calculations described here. In fact we would not expect that use of the correct sequence would greatly affect the results obtained. For example, simulated annealing trials using the neutral sequence of Ala-Ala for these two residues also were trapped in wrong minima, mainly due to the strong electron density of the sulfur atom (see below). Unlike conventional refinement, the high temperature of the SA refinement allows substantial displacements in the positions of atoms, allowing them to locate density peaks unaccounted for in the model. Thus, even with the Ala-Ala sequence, the large peak due to the sulfur causes the backbone to distort at high temperature, moving carbonyl atoms or other groups into the sulfur density (compare with Fig. 8 below).

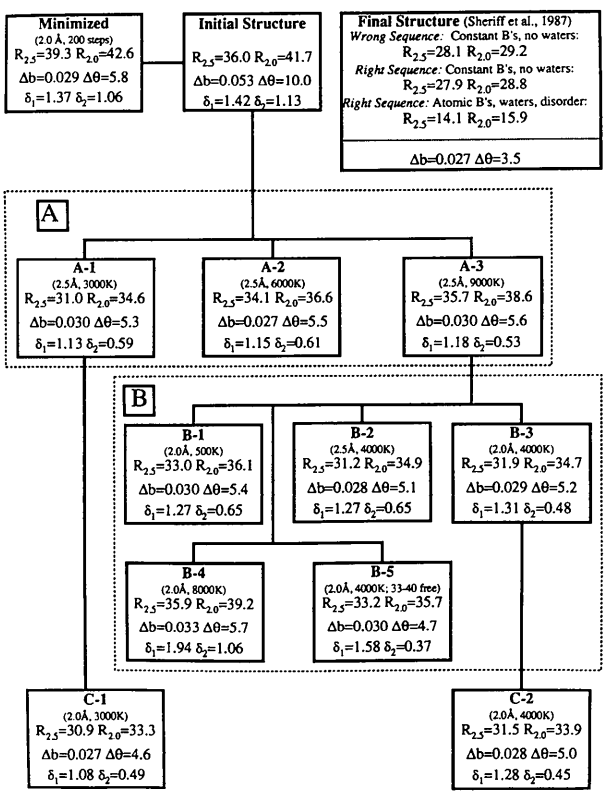
2.3.1. *The X-ray data set used.* Data to only 2·0 Å resolution had been used at the stage of the impasse in the refinement of myohemerythrin. We have also restricted SA refinements to a 2·0 Å data set in order to be consistent with the earlier work. The data set we use has been described by Sheriff *et al.* (1987) and has been deposited by these authors in the protein data bank (Bernstein *et al.*, 1977). In all our calculations we excluded reflections with observed amplitudes less than 10·0; we thus include 4409 of the 4472 unique reflections between 7 and 2·5 Å, and 8606 of the possible 8776 reflections to 2·0 Å resolution. The X-ray data are on an absolute scale.

2.3.2. *Minimization of the initial structure.* The INIT structure has rather poor stereochemistry ($\Delta b =$ 0·053 Å, $\Delta\theta = 10·0°$) compared with the SHS structure ($\Delta b = 0·027$ Å and $\Delta\theta = 3·5°$)* and with the final annealed structures ($\Delta b \simeq 0·030$ Å, $\Delta\theta \simeq 5·0°$). To compare the INIT structure meaningfully with the others, the stereochemistry has to be improved. Furthermore, in order to establish the utility of the molecular dynamics procedure, structures obtained using it must be compared with structures obtained by minimization alone. For these reasons, 200 steps of conjugate gradient minimization were performed, starting with the INIT structure, with a tolerance of 0·1 Å and a weight (S) of 60 000 for the X-ray term. The resulting structure will be referred to as MINIM.

* The deviations in bond lengths and angles are with reference to the *CHARMM* parameter set (Brooks *et al.*, 1983), which differs somewhat from the restraint values used in the refinement program *PROLSQ* used by Sheriff *et al.* (1987). Δb for the SHS structure is 0·017 Å with respect to the *PROLSQ* restraints (Sheriff *et al.*, 1987).

2.3.3. *Simulated annealing.* An outline of all the calculations performed is given in the scheme below. Starting from the initial structure, three annealing runs were performed using X-ray data to 2·5 Å and temperatures of 3000 (A-1), 6000 (A-2) and 9000 K (A-3). Each run was initiated by minimizing the energy, performing high-temperature molecular dynamics for 2–3 ps ($S = 50\,000$), low-temperature molecular dynamics for 1 ps ($S = 50\,000$), and finally minimizing the system ($S = 60\,000$ for the final stages of the minimization).

To investigate the effects of (i) including data to higher resolution and (ii) continuing the annealing procedure further, six more annealing runs were carried out. Five of these started with the minimized structure from run A-3 (9000 K), and one from the minimized structure from A-1. Four of them (B-1, B-3, B-4, B-5) incorporated data to 2·0 Å, while run B-2 used only 2·5 Å data. B-1 was carried out at 500, B-2, B-3 and B-5 at 4000 and B-4 at 8000 K. The lower-temperature run (B-1), was done with a weight of 80 000 on the X-ray term. In run B-5, the A–B loop (residues 33–40) was not included in the calculation of E_{xray}. Finally, the cooled structure from run B-3 (2·0 Å, 4000 K) was used to inititate another annealing run (C-2), at 4000 K, using data to 2·0 Å. The A-1 structure was used to initiate a 3 ps 3000 K dynamics run (2·0 Å data, $S = 50\,000$), which was cooled using a 2 ps 300 K run ($S = 50\,000$) and minimized ($S = 60\,000$) to obtain the C-1 structure.



Several calculations were focused on the A–B loop region, all of them using the incorrect sequence. These included annealing runs where (1) the rest of the protein was frozen and only the A–B loop was allowed to move; (2) the rest of protein was included in the X-ray calculation while the A–B loop was 'free'; and (3) annealing runs with one or two peptide bonds broken in the A–B loop. None of these calculations succeeded in correcting the error in the A–B loop and, except with reference to run B-5, we shall not discuss them further.

All the minimizations reported here used data to 2·0 Å in the final steps, and the same scale factor (60 000), so as to make the resulting structures strictly comparable. The dynamics runs using 2·0 Å data involved approximately 600 R value computations and took about 100 min of Cray-XMP c.p.u. time for 3 ps of dynamics.

3. Results and discussion

3.1. *R values*

$R_{2\cdot5}$ for the INIT structure is 7·9% higher than that of the SHS structure (see scheme). However, this structure has poor stereochemistry, and on refining it further to get the MINIM structure with approximately the same bond deviations as in the SHS structure, the R-value difference increases to 11·2% at 2·5 Å resolution. For the example of the C-1 structure, simulated annealing reduces $R_{2\cdot5}$ to 30·9% (*versus* 28·1% for SHS, with the wrong sequence) and $R_{2\cdot0}$ to 33·3% (*versus* 29·2% for SHS), a considerable improvement over MINIM or INIT. There are, however, some interesting differences between the R values of the various annealed structures that are worth examining further.

For the A series of runs, all starting with the INIT structure, the final R values increase with temperature (3000 K, $R_{2\cdot5} = 31\cdot0\%$; 6000 K, 34·1%; 9000 K, 35·7%). The A-3 structure was continued further at several temperatures to obtain the B series of runs. On looking at the B series, we see that the final R value is lowest for an intermediate temperature. Lower temperature (500 K) and higher temperature (8000 K) yield only modest improvements in R value whereas 4000 K (B-2 and B-3) yields the greatest improvement ($R_{2\cdot5}$ drops by 5·0%). However, on continuing B-3 further at 4000 K (C-2), no significant improvement in R value is seen. Continuation of A-1 (C-1 at 3000 K) yields the lowest R factor of all the structures ($R_{2\cdot5} = 30\cdot9\%$, $R_{2\cdot0} = 33\cdot3\%$).

The B-4 structure (8000 K) has a higher R value than other B structures. As discussed in the next section, the structure actually gets significantly worse during this refinement run. We have generally observed that when the R value during the high-temperature stage of the annealing process consistently stays over 50%, the structure usually

degrades to an unrecoverable state (the R value for a random structure is ~59%). Fig. 1 shows the time evolution of the R value for runs at 4000 K (B-3) and at 8000 K (B-4), both run with 2·0 Å data. The R value for the 4000 K run stays under 47% and decreases continuously with local fluctuations during the high-temperature stage, indicating that the structure is improving. For the B-4 run, however, the R value rises above 50% and does not show a steady decrease with time. Interestingly, the A-3 run was at an even higher temperature (9000 K) and the same weight on the X-ray term, although at lower resolution, and it yielded the best structural results amongst the A series of runs (scheme). High temperature can therefore be useful, and these results suggest that the annealing process may be improved by coupling the temperature control to the R value so that a continued increase in R values above 50% would result in cooling of the system.

In all of the annealing runs the wrong sequence (Cys 34–Asp 35) was used. This means that if the positional errors in the A–B loop are to be corrected, the sulfur of Cys 35 has to be in the wrong place (it would be in the position occupied by Asp 34 in the correct sequence). The effect of moving the sulfur out of its correct position is estimated by computing R values based on a structure obtained by taking SHS and changing the C^{γ} atom of Asp 34 into an S atom and the S^{γ} atom of Cys 35 into a C atom. The two terminal O atoms of the Asp sidechain are deleted. This raises $R_{2\cdot0}$ by 0·5 (see *wrong sequence* entry in scheme), emphasizing that the placement of the strongly scattering sulfur is critical. In fact, as will be shown below, the cysteine sidechain is very tightly held in place by the density, even at high temperature.

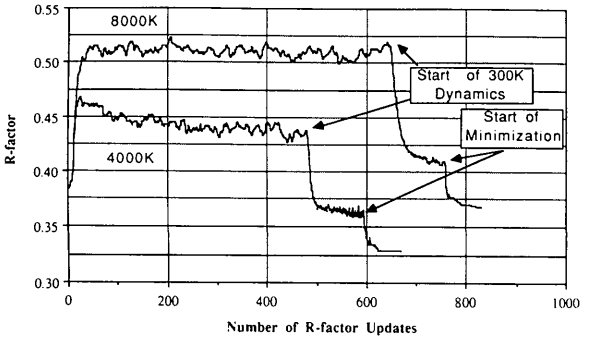


Fig. 1. R-factor history for the B-4 run (8000 K, upper line) and the B-3 run (4000 K, lower line). Both runs started with the A-3 structure and, following an initial minimization, ran at high temperature for 2 ps and at 3000 K for 1 ps, followed by another minimization. The abscissa shows the number of R-factor updates. These updates were done whenever any atom moved more than 0·25 Å away from its position at the last update. There are more updates for the 8000 K run because the greater atomic mobility led to more frequent updates.

3.2. *Structure*

Fig. 2(*a*) superimposes the C^{α} traces of the INIT and SHS structures. Most regions of the molecule show significant deviations between the two structures. Fig. 2(*b*) compares the C-2 structure and

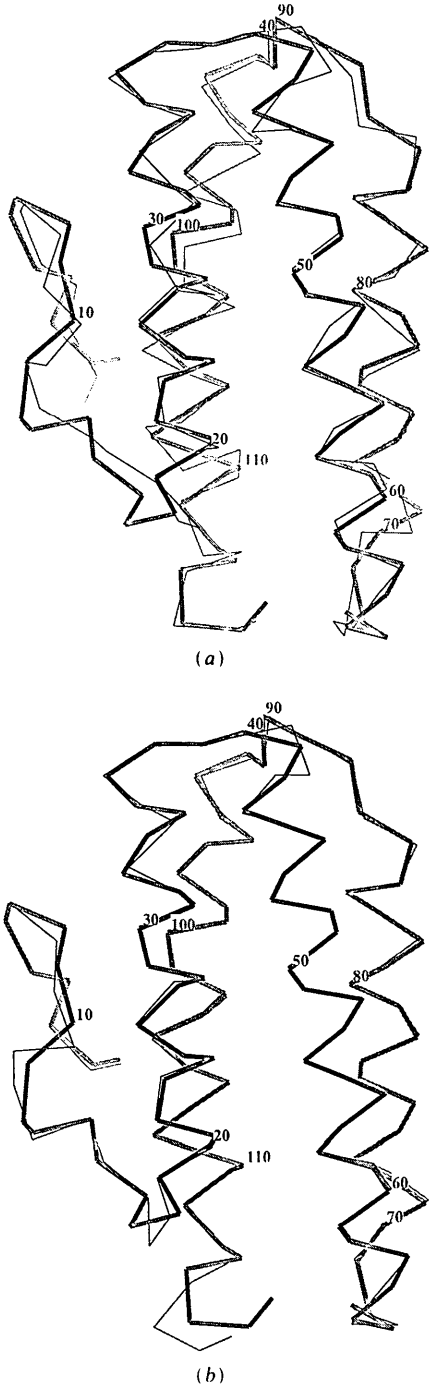


Fig. 2. C^{α} plots of the structure. In (*a*) and (*b*), the thick line represents the SHS structure. The thin line is INIT in (*a*) and C-1 in (*b*).

SHS. The two structures are now very close to each other; large deviations are seen only in three regions: the N-terminal arm, the A–B loop and the C terminus. The result of the annealing is to position correctly the backbone for the four helices and the BC and CD loops. The annealing does not appear to reduce the errors in the three other regions mentioned above.

The differences in backbone position are quantified in Fig. 3(*a*), which compares the errors in the backbone in INIT and A-3, relative to the SHS structure. Fig. 3(*b*) shows the improvement in the backbone structure on continuing A-3 to yield C-2. For most of the molecule, the C-2 structure is somewhat improved over A-3 and substantially improved over INIT. For the regions with large initial error, however, extended annealing increases the deviations from the SHS structure. In computing deviations between structures we therefore find it useful to discuss two types of r.m.s. deviations, δ_1 and δ_2. δ_1 is the r.m.s. deviation between backbone atoms over the whole molecule, whereas δ_2 is the r.m.s. deviation excluding those regions of large uncorrected error (δ_2 is computed over residues 1–10, 15–32, 45–114; *i.e.* 84% of the molecule). δ_1 and δ_2 for the various structures are shown in the scheme.

The uncorrected backbone errors all lie in regions where the initial structure has backbone dihedral angles (φ and ψ) deviating by more than 100° from the final values (see Fig. 3*c*). These large deviations in φ and ψ are due to localized errors in the tracing of the backbone, often resulting in sidechains being placed in directions diametrically opposite to the correct position. For example, the accumulation of backbone dihedral angle errors at residue 12 (see Fig. 3*c*) results in the C^{δ} atom of the Glu sidechain (in MINIM) being placed 11 Å away from the correct position. Molecular dynamics is unable to correct these kinds of errors because the atoms of the sidechain would have to pass through other atoms in order to get to the final structure.

δ_1 is dominated by the large localized errors; since some of these regions get worse on annealing (see below), δ_1 shows only a modest decrease (from 1·4 to 1·1 Å) on annealing. The behavior of δ_2 is more dramatic. It decreases from 1·13 Å in INIT to 0·45 and 0·37 Å in the best structures (C-2 and B-5). The high temperature of the A-3 structure appears to have led to the best structure in the A series ($\delta_2 = 0{\cdot}53$). Continued annealing of A-3 leads to a further improvement in structure and a significant drop in R value (scheme). The bulk of the rearrangements appears to be complete by this stage because further annealing (C-2) does not lead to a significant drop in δ_2. The effect of increasing the resolution of the data from 2·5 to 2·0 Å is not a very marked one, though there is a slight improvement of the structure at higher resolution (compare B-2 with B-3, scheme).

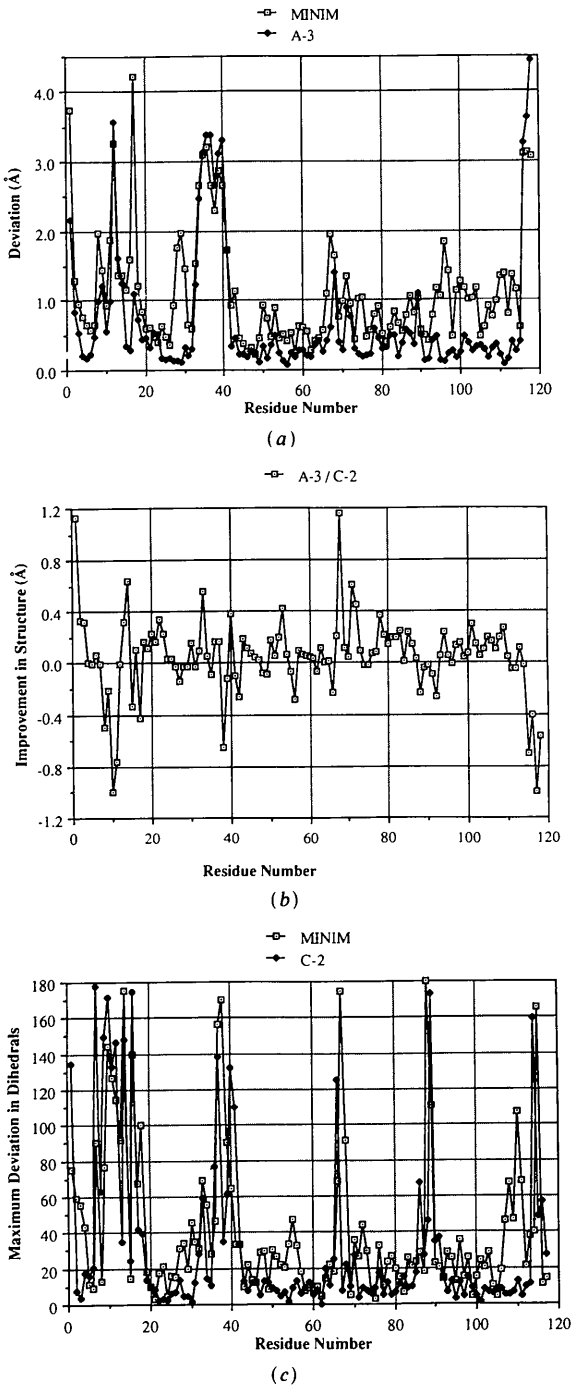


Fig. 3. (*a*) Residue average backbone deviations (C, N, C^{α}). Open squares, MINIM *vs* SHS. Closed diamonds, A-3 *vs* SHS. (*b*) Improvement in backbone structure in C-2 over A-3. The residue-averaged error in backbone position for the C-2 structure is subtracted from that for the A-3 structure. A positive value corresponds to an improvement in the structure, whereas a negative value shows that the errors increase on going from A-3 to C-2. (*c*) Deviation, in degrees, of φ and ψ. The patterns of errors in φ are very similar to those for ψ. The figure shows the deviation in either φ or ψ, whichever is greater, as a function of residue number for the MINIM structure (open squares) and the C-2 structure (closed diamonds).

In this test case we can easily pinpoint regions where the structure is in error because we know the final refined structure. However, the crucial question is whether one can determine errors in the absence of the known structure. There are internal indicators of the regions of the structure that are in error. One such indicator is the deviation of stereochemistry from ideal values. The most useful parameters to examine are the softer terms such as the bond angles or the torsion angles; the bond lengths are more tightly restrained by the energy function, E_{charmm}, and tend to remain close to ideal in all parts of the structure. Fig. 4 shows the deviation of the bond angles from ideality as a function of residue number for the C-2 structure. The regions of large deviation clearly correspond to the regions of large error (Fig. 3) and the pattern of deviations is quite distinct from that seen for the well refined SHS structure (Fig. 4). If multiple annealing runs are performed then another useful parameter is the deviation between different annealed structures; this result is analogous to that obtained in NMR structure determinations (Brünger *et al.*, 1986). Fig. 4 shows, as a function of residue number, the deviations of the A-1, A-2 and A-3 structures from the average structure determined by combining A-1, A-2 and A-3. This plot is very similar to that showing the actual errors in the structure (Fig. 3) and yet was generated with no reference to the SHS structure. Plots of the type shown in Fig. 4 allow one to identify regions of the molecule that need special attention; *i.e.* when SA refinement is not sufficient and manual rebuilding may be required. An interesting feature of the deviation curve (Fig. 4) is the dip near residue 35 (Cys, in the wrong sequence). As mentioned before, the strong density due to the sulfur holds this region in place, even when high temperatures are used.

Fig. 5 shows the improvement in sidechain structure in C-2, over INIT; in some sidechains the improvements are 2 to 3 Å. The situation here is similar to that for the backbone. For the bulk of the protein, the annealing process results in substantial improvement in sidechain positions. In regions of large backbone error, however, the annealing tends to introduce larger deviations from the SHS structure. For some surface sidechains another problem is that the lack of ordered water molecules in the model results in sidechains moving into solvent density (see below).

The structural changes brought about by annealing are illustrated by two examples. Fig. 6(*a*) compares the structures of residues 27-30 in SHS and INIT. In the latter structure the backbone and sidechain positions are significantly in error. Fig. 6(*b*) compares the SHS structure with the annealed C-2 structure. The ring of Phe 29 has rotated by 90° and moved by 3 Å, almost into register with the SHS structure. The backbone atoms in the two structures are almost superimposable, although the two lysine sidechains still show some deviations. Fig. 7 shows a similar comparison for residues 95-97. The C-2 structure is again almost superimposable with the SHS structure. The C^{γ} atom of Val 96 moves by 3 Å in this case.

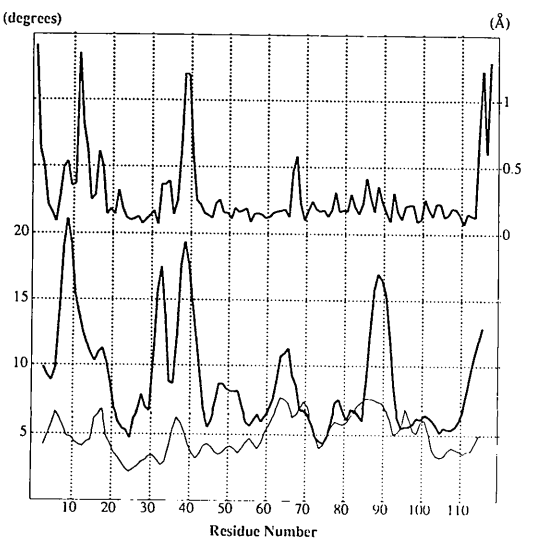


Fig. 4. Upper thick line: average deviation (in Å) of A-1, A-2 and A-3 backbone atoms (N, C, C^{α}), from the average of A-1, A-2 and A-3. Middle thick line: deviations of backbone bond angles from ideal values (*CHARMM*; Brooks *et al.*, 1983) in the C-1 structure. The maximum value per residue is computed and averaged over the neighboring five residues to obtain a smoothed plot. Bottom thin line: deviations of bond angles from ideality in the SHS structure, as before.

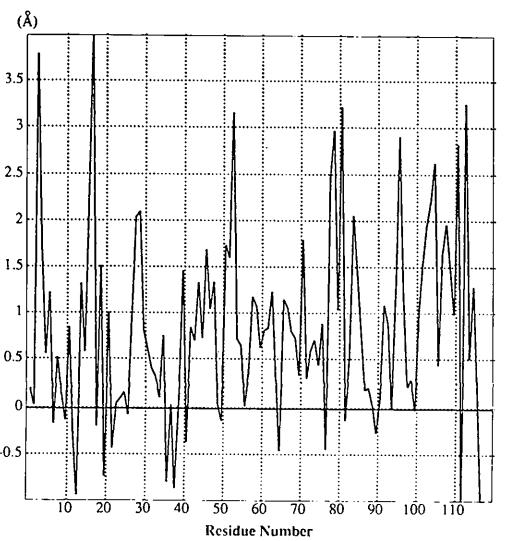


Fig. 5. Average improvement in sidechain structure for C-1 over INIT. $\langle d_1 - d_2 \rangle$ is plotted, where d_1 is the deviation between an atom in INIT and an atom in SHS and d_2 is the deviation between C-2 and SHS. A positive value implies that the structure improved in C-2. For this comparison the sidechains of Asp, Glu, Asn, Gln, Arg, His, Thr, Tyr, Phe, Leu and Val were considered to be symmetrical beyond the branch point.

The region of sequence error is examined next, in Fig. 8, which shows how the error results in the residues being misaligned. Annealing does not improve the structure and merely brings the sulfur atom of Cys 34 (wrong sequence) closer to the position of S^{γ} of Cys 35 in SHS (Fig. 8*b*). Finally, the structure resulting from run B-5 is shown in Fig. 8(*c*). In this run, residues 33–40 were allowed to move without reference to the X-ray term in the high-temperature stage of the dynamics; consequently, the loop structure changes considerably. After the high-temperature stage, the X-ray term was turned back on for these residues, but it only succeeded in locking them into a wrong conformation. The sidechain of Phe 33 leaves its correct position and moves into the unoccupied density of Asp 34 (correct sequence). In a domino effect, the sidechain of Tyr 8 (not shown in figure) moves in to fill the now vacant density of Phe 33. The other residues of the A–B loop are also in wrong positions. This is a case where the power of high-temperature molecular dynamics to move sidechains results in damage to parts of the structure that were initially correct. Such possibilities provide a note of caution; the more powerful a method, the more care must be used in its application.

3.3. *Ordered solvent molecules*

In the early stages of the refinement of a protein structure the ordered solvent molecules are not usually placed. This poses a problem for the SA refinement method. We have experimented with simulated annealing runs starting from a well refined structure of ribonuclease-A where solvent molecules were included. We find that unless the solvent molecules are constrained to remain near their original positions, the elevated kinetic energies of these simulations cause them to wander irreversibly out of the electron density (Kuriyan, unpublished). The application of constraints on water positions is problematic; if the initial electron density maps are noisy (as is the case here), the water positions are often poorly identified. However, if the more tightly bound solvent molecules are not included in the model, the SA

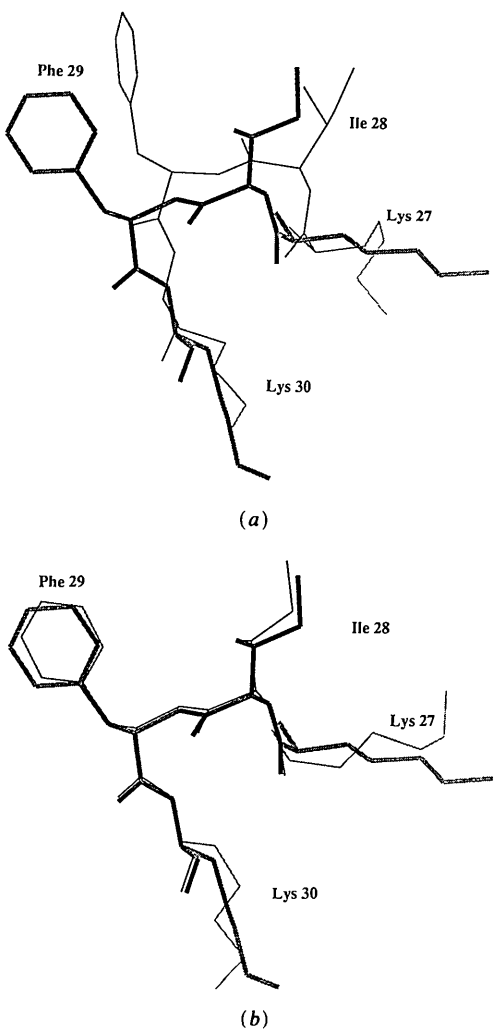


Fig. 6. Residues 27–30. Thick line in both figures: SHS. Thin line: INIT in (*a*) and C-2 in (*b*).

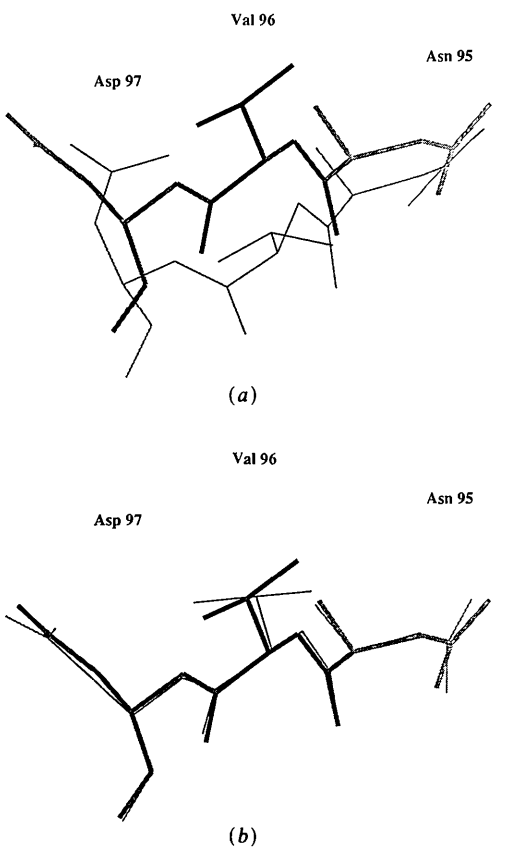


Fig. 7. Residues 95–97. Thick line in both figures: SHS. Thin line: INIT in (*a*) and C-2 in (*b*).

refinement moves many poorly built surface sidechain atoms into density associated with solvent.

In the test refinements reported here we did not include any solvent atoms. The INIT structure has ten sidechains that are in (or very close to) density that was later identified by Sheriff *et al.* (1987) as due to bound solvent (Table 1). Some of the errors introduced are localized, such as the terminal atoms of lysine sidechains falling into solvent density, but for others the misinterpretation of the solvent structure results in large backbone errors as well. For example, in INIT, Tyr 18 was wrongly built into density later associated with a sulfate ion. This error in the INIT structure is not corrected in any of the annealed structures; the strong electron density due to the sulfur and the four oxygen atoms provides too deep a local minimum even for the high-temperature dynamics. Some of the errors are corrected after the SA refinement, but on the whole more sidechains are trapped in water positions, with 17 sidechains now having some atoms in positions identified as solvent sites in the SHS structure (Table 1). The errors in the positions of most sidechains are actually decreased; an interesting example is that of Phe 17. The INIT structure has this sidechain placed about 14 Å from its position in the SHS structure (Table 1). The annealing moves the sidechain towards the correct position, reducing the error in the phenyl ring to about 7 Å, but it gets trapped in density belonging to a four-water-molecule cluster, from which it does not escape (Table 1).

Most, but not all, sidechains that become trapped in solvent density during the simulated annealing have large errors in the INIT structure (see Table 1). Most have errors that are larger than 3 Å and more than half of them have at least one atom that is

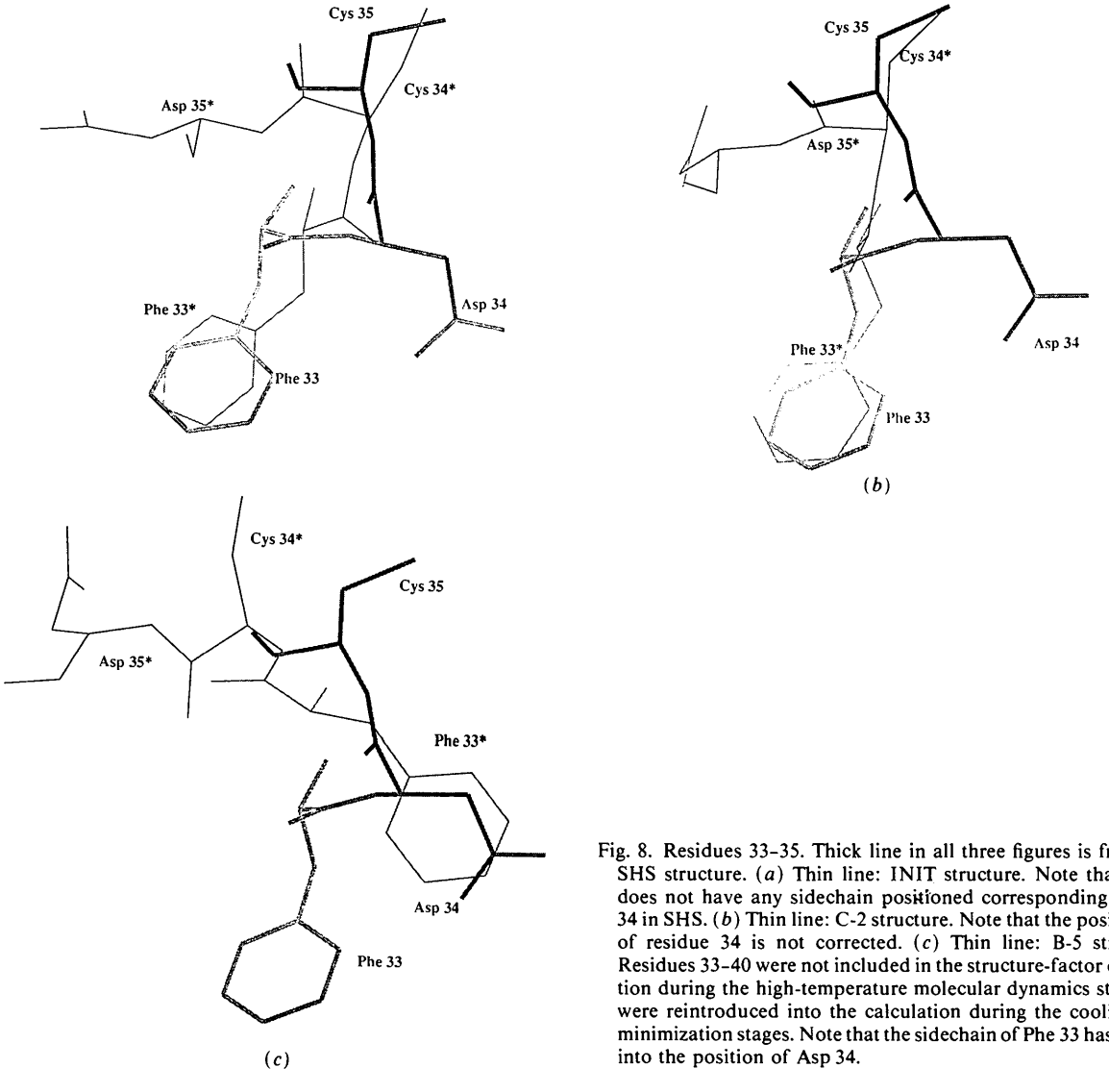


Fig. 8. Residues 33–35. Thick line in all three figures is from the SHS structure. (*a*) Thin line: INIT structure. Note that INIT does not have any sidechain positioned corresponding to Asp 34 in SHS. (*b*) Thin line: C-2 structure. Note that the positioning of residue 34 is not corrected. (*c*) Thin line: B-5 structure. Residues 33–40 were not included in the structure-factor calculation during the high-temperature molecular dynamics stage but were reintroduced into the calculation during the cooling and minimization stages. Note that the sidechain of Phe 33 has moved into the position of Asp 34.

Table 1. *Details of residues in the INIT and C-1 structures that have one or more atoms overlapping a solvent molecule in the SHS structure*

The numbers in the second and third columns are the residue identifiers for water molecules (and one sulfate ion) in the structure deposited in the Protein Data Bank by Sheriff *et al.* (1987) that are within 1·5 Å of an atom in the sidechain of the specified residue. The last two columns gives the deviation (in Å) between a reference atom in the sidechain (fourth column) in the INIT and C-1 structures from the SHS structure. The reference atom is the outermost atom for unbranched and aromatic sidechains and is the outermost atom before the branch for the others.

Residue	INIT	C-1	Reference atom	Error in INIT	Error in C-1
Asp 3	155, 284	—	CG	6·4	2·1
Tyr 8	126, 143	143, 179	OH	10·0	10·9
Val 9	—	135	CB	4·1	3·8
Trp 10	122	122	CZ3	3·6	5·4
Glu 12	—	186	CD	11·0	11·2
Arg 15	—	218	CZ	0·6	0·4
Phe 17	—	127, 136, 156, 254	CZ	14·4	7·5
Tyr 18	Sulfate-138, 312	Sulfate-138, 312	OH	10·0	9·8
Gln 20	297	147, 297	CD	3·0	3·3
Glu 23	—	210	CD	1·1	0·3
Arg 37	—	148, 177	CZ	6·3	6·6
Ser 40	—	160	OG	6·9	4·7
Asn 43	121	121	CG	3·3	2·8
Val 71	137	—	CB	2·7	0·2
Lys 75	—	294	NZ	4·4	3·3
Asp 79	217, 196	—	CG	4·2	0·2
Lys 78	188	—	NZ	6·0	2·0
Leu 81	300	—	CG	3·8	0·5
Lys 94	—	256	NZ	5·9	5·3
Lys 100	—	241	NZ	3·2	4·1
Phe 112	—	272	CZ	2·1	8·2
Lys 117	—	182	NZ	4·8	4·8

wrongly placed by more than 4 Å in the INIT structure. In all cases but one the SA refinement results in comparable or lower errors than the INIT structure for these residues. The exception is Phe 112, which has been moved from an essentially correct conformation in INIT into solvent density (C-2), resulting in errors of about 8 Å in the sidechain.

3.4. *Electron density maps*

One very promising result of the SA refinement is that electron density maps computed using the refined structures are easier to interpret, even when there remain significant local errors, as in the present case. This is not surprising because the substantial improvement in the model is expected to reduce the noise in these maps. Fig. 9 shows electron density for residues 34–35, calculated using INIT (Fig. 9*a*) and C-2 (Fig. 9*b*) as the models for the phase information. As shown in Fig. 8, one problem in this region for both models is that no sidechain has been placed in the position of Asp 34 (correct sequence). The INIT map in fact shows no density for this sidechain. However, in the C-2 map, even though this sidechain is positioned elsewhere in the model, the density clearly shows a feature corresponding to the correct placement of the Asp sidechain. This feature shows up in the C-2 map even though the wrong structure in this region was included in the calculation of F_c. The prospects for rebuilding this region are therefore significantly improved.

4. Concluding remarks

In this paper we provide a realistic test of X-ray refinement by simulated annealing, in that we begin with a crude initial model obtained from experimental X-ray data and compare the annealed structures with a well refined model. The results of these test refinements are very encouraging since they demonstrate that SA refinement, without any manual intervention, is able to correct errors of 2–4 Å in backbone positions. This clearly saves a great deal of human effort, albeit at a non-negligible cost of supercomputer time. In regions of the protein where mistakes in the chain tracing have led to larger errors (3–5 Å in backbone positions and substantially greater in sidechain positions), the procedure is not able to correct the structure. Even in these regions, SA refinement improves the results because the resulting electron density maps are less noisy. This implies that any required manual rebuilding will be easier, as was suggested for MAAT by Brünger (1988*a*).

Most of the improvement in the structure occurs in a relatively short time (3–4 ps, corresponding to less than 30 min on the Cray-XMP) when high-temperature dynamics is used and continuing the simulations further does not lead to significant improvement in the structure. The heating and cooling protocol could no doubt be improved, but it appears that the SA refined model is at a stage where it needs to be manually adjusted and water molecules added before continuing the refinement. This suggests

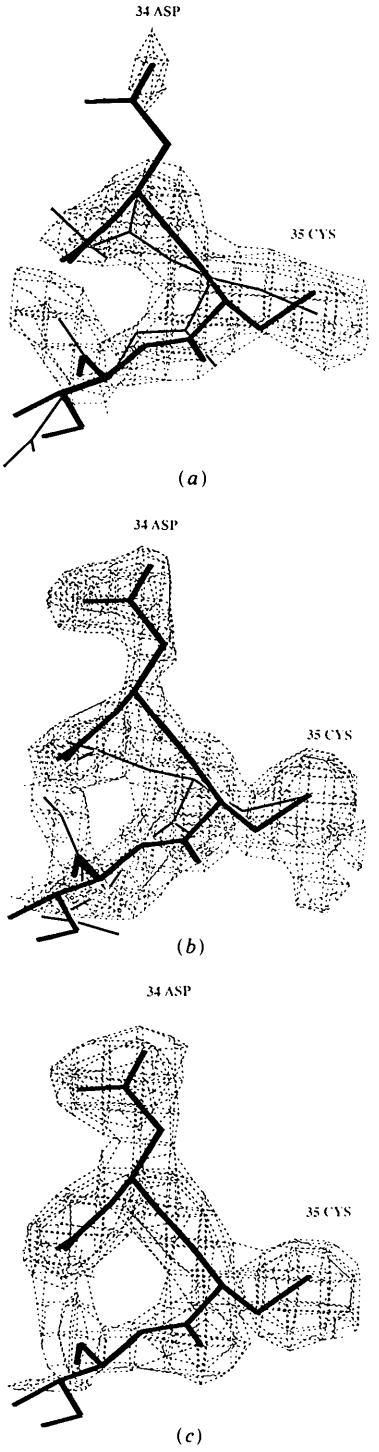


Fig. 9. Electron density from $(2|F|_o - |F|_c)\exp(i\alpha_c)$ maps. The model used to calculate F_c is INIT in (*a*), C-2 in (*b*) and SHS in (*c*). In all three figures the thick lines represent the SHS structure while the thin lines represent INIT in (*a*) and A-3 in (*b*). Electron density features corresponding to the Asp 34 sidechain clearly show up in the map using C-2 for the phasing model (*b*), even though this structure does not have this residue in the correct position. Maps calculated using the MINIM structure (*a*) do not show these features.

that the SA methodology in its present form is most effective for structures of intermediate accuracy with errors in the atomic positions on the order of 2–3 Å. For starting structures of significantly better quality than the INIT structure, the benefits of using the SA refinement over conventional refinement is less clear. The critical factors in these cases will probably be the modelling of the solvent and relatively fine adjustments to the structure. For these purposes an extended version of the SA algorithm including explicit treatment of solvent is required.

In spite of the success of the SA methodology, the importance of manual examination of the electron density maps at various stages of the refinement cannot be overemphasized. This is essential for the placement of surface sidechains and solvent molecules and for checking regions of the protein where the high-temperature dynamics leads to large excursions. Such large changes can take the structure from one wrong local minimum to another wrong, and perhaps deeper, local minimum.

The present results when combined with those already published (Brünger *et al.*, 1987, 1989; Brünger, 1988*a*) make clear that SA refinement is a powerful addition to the arsenal of macromolecular crystallographers. It is hoped (expected even) that, with its widespread use, improvements will be introduced which will extend its range of applicability and lead to further reduction in the time required for the refinement process.

We are grateful for grants of Cray supercomputer time from the Office of Naval Research (MK) and the Pittsburgh Supercomputer Center (JK, grant No. DMB-88001P). JK is supported by the Andre Meyer University Fellowship. This work was supported in part by grants from the National Science Foundation (grant Nos. DMB-8409658 to WAH and CHE-85-14245 to MK).

References

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F., BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BROOKS, B. R., BRUCCOLERI, R. E., OLAFSON, B. D., STATES, D. J., SWAMINATHAN, S. & KARPLUS, M. (1983). *J. Comput. Chem.* **4**, 187–217.

BROOKS, B. R., PASTOR, R. W. & CARSON, F. W. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 4470–4474.

BRUCCOLERI, R. E. (1984). PhD thesis. Harvard Univ., Cambridge, MA 02138, USA.

BRUCCOLERI, R. E. & KARPLUS, M. (1989). In preparation.

BRÜNGER, A. T. (1988*a*). *J. Mol. Biol.* **203**, 803–816.

BRÜNGER, A. T. (1988*b*). In *Crystallographic Computing* 4: *Techniques and New Technologies*, edited by N. W. ISAACS & M. R. TAYLOR. Oxford Univ. Press.

BRÜNGER, A. T. (1989). *Acta Cryst.* A**45**, 42–50.

BRÜNGER, A. T., CLORE, G. M., GRONENBORN, A. M. & KARPLUS, M. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 3801.

BRÜNGER, A. T. & KARPLUS, M. (1988). *Proteins: Struct. Function Genetics.* In the press.

BRÜNGER, A. T., KURIYAN, J. & KARPLUS, M. (1987). *Science*, **235**, 458-460.
BRÜNGER, A. T., PETSKO, G. A. & KARPLUS, M. (1989). *Acta Cryst.* A**45**, 50-61.
DEISENHOFER, J., REMINGTON, S. J. & STEIGEMANN, W. (1985). *Methods Enzymol.* **115**, 303-323.
GROS, P., FUJINAGA, M., DIJKSTRA, B. W., KALK, K. H. & HOL, W. G. J. (1988). Abstr. Int. School of Crystallography, 14th Course, 'Crystallography of Molecular Biology', Erice, Italy.
HENDRICKSON, W. A. (1985). *Methods Enzymol.* **115**, 252-270.
HENDRICKSON, W. A., KLIPPENSTEIN, G. L. & WARD, K. B. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 2160-2164.
HENDRICKSON, W. A. & KONNERT, J. H. (1981). In *Biomolecular Structure, Function, Conformation and Evolution*, edited by R. SRINIVASAN, Vol. 1, p. 43. Oxford: Pergamon.
HENDRICKSON, W. A. & WARD, K. B. (1975). *Biochem. Biophys. Res. Commun.* **66**, 1349-1356.
JACK, A. & LEVITT, M. (1978). *Acta Cryst.* A**34**, 931-935.
JONES, T. A. (1985). *Methods Enzymol.* **115**, 157-171.
KAPTEIN, R., ZUIDERWEG, E. R. P., SCHEEK, R. M., BOELENS, R. & VAN GUNSTEREN, W. F. (1985). *J. Mol. Biol.* **182**, 179.
KIRKPATRICK, S., GELATT, C. D. JR & VECCHI, M. P. (1983). *Science*, **220**, 671-680.
KLIPPENSTEIN, G. L., COTE, J. L. & LUDLAN, S. E. (1976). *Biochemistry*, **15**, 1128-1136.
KONNERT, J. H. & HENDRICKSON, W. A. (1980). *Acta Cryst.* A**36**, 344-349.
MCCAMMON, J. A. & HARVEY, S. C. (1987). *Dynamics of Proteins and Nucleic Acids.* Cambridge Univ. Press.
NORTHRUP, S. H. & MCCAMMON, J. A. (1980). *Biopolymers*, **19**, 1001.
PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. & VETTERLING, W. T. (1986). *Numerical Recipes.* Cambridge Univ. Press.
REIHER, W. E. (1985). PhD thesis. Harvard Univ., Cambridge, MA 02138, USA.
SHERIFF, S. & HENDRICKSON, W. A. (1987). *Acta Cryst.* A**43**, 118-121.
SHERIFF, S., HENDRICKSON, W. A. & SMITH, J. L. (1987). *J. Mol. Biol.* **197**, 273-296.
SMITH, D. L., RINGE, D., FINLAYSON, W. L. & KIRSCH, J. F. (1986). *J. Mol. Biol.* **191**, 301-302.

Acta Cryst. (1989). A**45**, 409-415

The Influence of Coulombic Interactions on Thermal Parameters for Naphthalene and Anthracene: a Lattice-Dynamical Approach

BY A. CRIADO

Departamento de Fisica de la Materia Condensada, Instituto de Ciencia de los Materiales, Universidad de Sevilla, Aptdo 1065, 41080 *Sevilla, Spain*

(*Received* 8 *March* 1988; *accepted* 16 *January* 1989)

Abstract

A lattice-dynamical calculation of thermal parameters is performed for naphthalene and anthracene in the external Born-von Karman formalism using 6-exp potential functions and electrostatic interactions in the form of an atomic point-charge model. Phonon dispersion curves and density-of-states functions are calculated with and without charges. In general, the effect of the Coulombic forces is small and sensibly affects optical branches linked to the B_u mode in both compounds, improving the agreement with the experimental data. The effect on thermal parameters is roughly opposite to that of molecular non-rigidity.

Introduction

The calculation of crystallographic thermal parameters based on a knowledge of the crystal structure and the interatomic forces has been considered an interesting subject because it constitutes an independent test of the thermal parameters obtained from crystal structure analysis as well as of the proposed interatomic force model. In many cases, the hypothesis of a rigid molecule and 6-exp semiempirical potentials has been found to be successful. Most cases where such calculations have been performed concern molecules where Coulombic effects are not important (Filippini, Gramaccioli, Simonetta & Suffritti, 1973, 1976, 1978).

Recently, an extension of the method has been made by taking into account the effects of molecular non-rigidity. For a 'rigid' molecule, the contribution of intramolecular modes (though little populated) added to the external ones gives rise to a small increase of the calculated parameters (Gramaccioli, Filippini & Simonetta, 1982) especially for peripheral atoms. On the other hand, the coupling of external modes with intramolecular degrees of freedom produces a relaxation of lattice frequencies (Pawley & Cyvin, 1970); consequently, a rise of thermal parameters is observed which is comparable to the separate contribution of internal modes (Gramaccioli & Filippini, 1983). For all these cases and for others where non-rigidity is beyond question, the normal coordinates of the isolated molecule have been taken as a basis for the crystal calculations (Bonadeo & Burgos, 1982; Filippini, Simonetta & Gramaccioli, 1984; Gramaccioli & Filippini, 1985; Filippini 1985; Filippini & Gramaccioli, 1986).

0108-7673/89/060409-07$03.00