formalism by allowing for cells of different lengths. This is in contrast to the difference equation method where each member of a group must be treated separately. Another advantage of the matrix method is its ready adaptability to computational methods. With the matrix method the diffuse scattering for a given model is readily calculated throughout reciprocal space. On a high speed computer the model is readily varied. Attempts are underway to refine such scattering models, based on the matrix method and diffuse X-ray data collected by counter methods.

This research was supported under the NSF-MRL program through the Materials Research Center of Northwestern University (Grant DMR76-80847).

References

BEYELER, H. U. (1976). *Phys. Rev. Lett.* **37**, 1557.

COWIE, M., GLEIZES, A., GRYNKEWICH, G. W., KALINA, D. W., MCCLURE, M. S., SCARINGE, R. P., TEITELBAUM, R. C., RUBY, S. L., IBERS, J. A., KANNEWURF, C. R. & MARKS, T. J. (1979). *J. Am. Chem. Soc.* **101**, 2921–2936.

ENDRES, H., KELLER, H. J., MEGNAMISI-BELOMBE, M., MORONI, W., PRITZKOW, H., WEISS, J. & COMÈS, R. (1976). *Acta Cryst.* A**32**, 954–957.

ENDRES, H., KELLER, H. J., MORONI, W. & WEISS, J. (1975). *Acta Cryst.* B**31**, 2357–2358.

GUINIER, A. (1963). *X-ray Diffraction in Crystals, Imperfect Crystals and Amorphous Bodies*. San Francisco: W. H. Freeman and Co.

HENDRICKS, S. & TELLER, E. (1942). *J. Chem. Phys.* **10**, 147–167.

HERBSTEIN, F. H. & KAPON, M. (1972). *Acta Cryst.* A**28**, S74.

HUML, K. (1967). *Acta Cryst.* **22**, 29–32.

JAGODZINSKI, H. (1949*a*). *Acta Cryst.* **2**, 201–207.

JAGODZINSKI, H. (1949*b*). *Acta Cryst.* **2**, 208–214.

JAGODZINSKI, H. (1949*c*). *Acta Cryst.* **2**, 298–304.

KAKINOKI, J. & KOMURA, Y. (1952). *J. Phys. Soc. Jpn*, **7**, 30–35.

KELLER, H. J. (1977). *Chemistry and Physics of One-Dimensional Metals*. New York: Plenum Press.

MARKS, T. J. (1978). *Ann. N.Y. Acad. Sci.* **313**, 594–616.

MILLER, J. S. & EPSTEIN, A. J. (1976). *Prog. Inorg. Chem.* **20**, 1–151.

SMITH, D. L. & LUSS, H. R. (1977). *Acta Cryst.* B**33**, 1744–1749.

TAKAKI, Y. & SAKURAI, K. (1976). *Acta Cryst.* A**32**, 657–663.

WILSON, A. J. C. (1942). *Proc. R. Soc. London Ser. A*, **180**, 277–285.

Acta Cryst. (1979). A**35**, 810–817

Phase Extension and Refinement of Bence–Jones Protein Rhe (1·9 Å)

BY W. FUREY JR, B. C. WANG, C. S. YOO AND M. SAX

Biocrystallography Laboratory, Box 12055, *VA Medical Center, Pittsburgh, PA* 15240, *USA and Department of Crystallography, University of Pittsburgh, Pittsburgh, PA* 15260, *USA*

(*Received* 17 *October* 1978; *accepted* 1 *May* 1979)

Abstract

A procedure is described whereby crude atomic coordinates obtained from a medium-resolution electron density map (~3–4 Å) with the aid of Watson–Kendrew models may be refined to native data of near-atomic resolution using only a limited subset of the data and non-interactive computer graphics. This refinement procedure for the Bence–Jones protein Rhe included phase extension from 3·0 to 1·9 Å and led to an improved crystallographic sequence for the protein. The structure was refined, by a stereochemically restrained least-squares technique in reciprocal space, to a residual value $R_F = 0{\cdot}284$ for a model consisting of 795 non-hydrogen protein atoms with an overall thermal factor using data with d ranging from 5·0 to 1·9 Å. Relative weights for the structure factor and stereochemical-restraint observations were determined empirically and the optimum weights were found to be those which yield values of $W|F_o - F_c|^2$ which are typically 4 to 6 times the value of $W|d_o - d_I|^2$, where d_o and d_I are the current and ideal values for the stereochemically restrained parameters. It was found that reasonable refinement may be obtained with only 32% of the observed data.

0567-7394/79/050810-08$01.00

Introduction

In the application of the method of multiple isomorphous replacement, MIR, to the structure determination of macromolecules, one is often confronted

with the problem of inadequate resolution. The native protein data, however, usually extend to atomic or near-atomic resolution; therefore it is desirable to obtain whatever structural information the MIR data will yield, and proceed to refine and add to this information based on native data alone. In addition, the utilization of native data alone is desirable because this data is free of any perturbing effects introduced by the heavy atoms. In this work we use the high-resolution native data (1·9 Å) to refine the Bence–Jones dimer Rhe for which only 3·0 Å MIR data are available. The technique involves a gradual increase in resolution utilizing subsets of the native data set, simple non-interactive graphics and the Hendrickson–Konnert program system for restrained reciprocal-space refinement of proteins.

Refinement technique and Rhe crystal data

Rhe crystallizes in the orthorhombic space group $P2_12_12$ with the cell dimensions $a = 54{\cdot}63\ (3)$, $b = 52{\cdot}22\ (4)$ and $c = 42{\cdot}62\ (4)$ Å. There are four monomers (V_L domains) per unit cell. The twofold axis utilized by the V_L–V_L dimer upon association is crystallographically exact.

Since chemical-sequence information is presently unavailable, a tentative sequence was deduced from the 3·0 Å MIR map (Wang, Yoo & Sax, 1979). This sequence (113 residues/asymmetric unit), with coordinates of the 795 non-hydrogen protein atoms obtained manually with the aid of an optical comparator from Watson–Kendrew models and the MIR map, provided the starting point for the refinement. The data used in the early steps involved only the native structure factors which were used in the calculation of the MIR map, *i.e.* those reflections which had $I/\sigma(I)$ ratios of 3·0 or greater in each of the native and derivative data sets (2144 reflections). The minimum d spacing for those reflections was 3·0 Å.

Because of the modest size of the protein, the refinement was carried out directly against the observed structure factors with the program system of Hendrickson & Konnert (1977).

The simplicity of their method is easily grasped by examination of the function φ which is minimized, where

$$\varphi = \sum_{i=1}^{NREF} W_i|F_{oi} - F_{ci}|^2 + \sum_{i=1}^{NBOND} W_i|d_{oi} - d_{Ii}|^2 + \sum_{i=1}^{NANG} W_i|d_{oi} - d_{Ii}|^2 + \sum_{i=1}^{NPLN} W_i|d_{oi} - d_{Ii}|^2 + \sum_{i=1}^{NCHIRAL} W_i|V_{oi} - V_{Ii}|^2 + \sum_{i=1}^{NVDW} W_i^2|d_{oi} - d_{Ii}|^4.$$

The first term is the function usually minimized in small-molecule crystallography. The 2nd and 3rd terms minimize the deviations of bond distances and angle-defining distances from their ideal values. The 4th term minimizes the deviations from planarity of groups of atoms expected to be planar. The 5th term minimizes the deviations from the ideal chiral volumes (sign included) required for given chiral centers, thus ensuring the proper handedness. The 6th term prevents unreasonably short non-bonded contacts by minimizing the deviations from ideal contact distances; it is applied only to those contact distances which are deemed too short. All quantities subscripted I are ideal values for the parameters as determined from small-molecule crystal structures. Quantities subscripted with o refer to the current values of the parameters in the trial model (except for term 1 where o denotes F observed).

It is obvious that the minimization of φ with the proper weights should produce a structure that agrees well with the observed data and is consistent with almost everything known about protein structural chemistry. The method is unique in that all structural information is treated as observations. The refinement path should be smooth with respect to the protein geometry as opposed to the saw-tooth refinement paths commonly observed which use Fourier refinement followed by geometry optimization (*cf.* Stenkamp & Jensen, 1976). In addition, by varying the relative weights for the structure factor and stereochemical-restraint contributions one can choose between a free-atom, restrained, or constrained refinement. For further details of the refinement technique the reader is referred to Konnert (1976).

Determination of refinement parameters

Prior to the first cycle of refinement the weighting scheme must be chosen. The weights are defined to be $1{\cdot}0/\sigma^2$, where σ is the estimated standard deviation for the observation. The individual σ's are then defined to produce any desired weight. It is not immediately obvious what weights should be used, particularly in view of the large number of weighting schemes currently in use in small-molecule crystallography, where one is dealing only with term 1 of φ. It was therefore decided that weights would be determined empirically, with the assumption that the best weights are those that produce the greatest decrease in the R factor and still maintain reasonable stereochemistry. In order to determine the weighting scheme several cycles of refinement were performed using the standard deviations listed in Table 1, column A, and an overall thermal factor of 9·0 Å^2. Refinement based on those weights reduced R_F from 0·465 to 0·262 after three cycles ($R_F = \sum|F_o - F_c|/\sum F_o$). However, the r.m.s. deviation from ideal bond distances increased from

0·198 to 0·387 Å. It was obvious that the structure factors were overweighted. The standard deviations for structure factors were subsequently set equal to the average value of $|F_o - F_c|/2$ for that particular resolution (a suggestion of Dr W. Hendrickson), hereafter denoted by $0{\cdot}5\langle|\Delta F|\rangle_d$. The subscript d indicates the resolution dependence. This weighting scheme increased the R factor, but improved the geometry so that the r.m.s. deviation from ideal bond distances was 0·046 Å. The R factor and r.m.s. deviations from ideal bond distances are plotted in Fig. 1 *vs* the number of refinement cycles. Based on these results, it appeared that the standard deviations for the structure factors were approximately correct; however, the stereochemistry could still be improved. An improvement in stereochemistry was obtained by arbitrarily increasing the restraint weights. The refinement was then reinitiated using the weights derived from the e.s.d.'s listed in Table 1, column *B*, which place more emphasis on the restraints. After five cycles the R factor was reduced from 0·465 to 0·386 with a corresponding r.m.s. deviation from ideal bond distances of 0·023 Å. The standard deviations for structure factors were then reset to $0{\cdot}33\langle|\Delta F|\rangle_d$ in an attempt to reduce the R factor by placing more emphasis on structure factors. This resulted in a decrease in the R factor to 0·350; however, the r.m.s. deviation from ideal bond distances increased to 0·042 Å. The R factors and r.m.s. deviations from ideality for bond distances are plotted in Fig. 2 *vs* the number of cycles of refinement. In view of the curve in Fig. 2 the optimum weighting scheme appeared to be $\sigma(F) = 0{\cdot}4\langle|\Delta F|\rangle_d$ and e.s.d.'s for the restraints as listed in Table 1, column *B*, since this scheme should produce the minimum R factor consistent with a protein structure with r.m.s. deviations from ideal geometry of about 0.03 Å. All further refinement cycles utilized this weighting scheme with some minor variations which are discussed later.

Table 1. *Estimated standard deviations for weighting schemes A and B*

Observation type	E.s.d. *A*	*B*
Structure factor amplitude (e)	1·	$0{\cdot}5\langle|\Delta F|\rangle_d$
Bond distance (Å)	0·05	0·03
Angle-defining distance (Å)	0·10	0·04
van der Waals contact distance (Å)	0·50	0·30
Chiral volume (Å^3)	0·40	0·40
Planar deviations (Å)	0·05	0·03

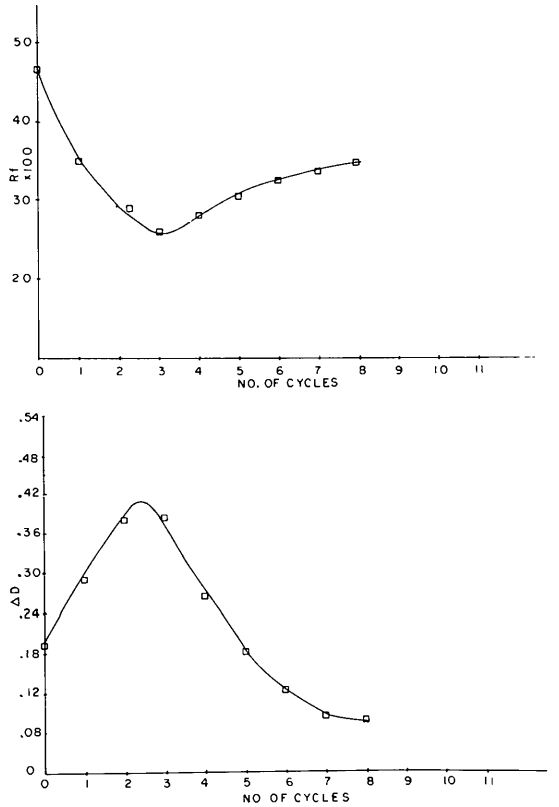

Fig. 1. R_F *versus* number of cycles of refinement. Weights were derived from the e.s.d.'s in Table 1, column *A*, for the first three cycles. Subsequent cycles used the same weights for restraints but the structure factor e.s.d.'s were reset to $0{\cdot}5\langle|\Delta F|\rangle_d$. ΔD is the r.m.s. deviation from ideality for bond distances in ångströms.

Preliminary refinement

With this weighting scheme, the refinement was started with the unrefined coordinates and 3·0 Å data. All low-resolution data were included, despite the fact that the model contained no solvent, in order to take advantage of the power of low-resolution data to move badly misplaced atoms. Some of the pertinent parameters for

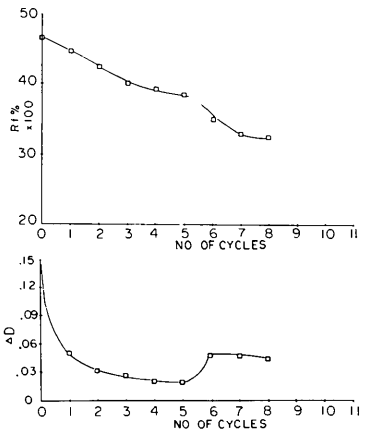

Fig. 2. R_F *versus* number of cycles of refinement. Weights were derived from the e.s.d.'s in Table 1, column *B*, for the first five cycles. Subsequent cycles used the same weights for restraints but the structure factor e.s.d.'s were reset to $0{\cdot}33\langle|\Delta F|\rangle_d$. ΔD is the r.m.s. deviation from ideality for bond distances in ångströms.

the structure and refinement are listed in Table 2. After several cycles of refinement, the R factor began to level off and the r.m.s. shifts dropped from 0·23 to 0·07 Å. The R factors as a function of resolution are given in Table 3. The 96 reflections with $d > 9·0$ Å were then removed from the data set, since they are likely to be heavily dependent on solvent molecules which were omitted from the model. After several more cycles of refinement the R factor had dropped to 0·328 with r.m.s. shifts of 0·03 Å along each cell axis.

Because no chemical sequence was available and several regions of the structure were poorly defined in the MIR map, the refined structure was checked for sequence and peptide plane orientation errors, using a program written to superimpose regions of the protein (5 to 10 residues) on the MIR map and plot the results in stereoscopic projection. The plots indicated 16 sequence changes (mostly single-atom additions or deletions) which were implemented by optically estimating new coordinates from the map. In several places gross coordinate shifts were indicated for groups of atoms. These changes were implemented by placing clear plastic sheets over the appropriate projections of the map and structure plot, tracing in the structure, and then rotating or translating the tracing to get a better fit. This method is quite inexpensive, and is capable of correcting groups of atoms (anywhere from 1 to 25) in terms of only one or two parameters. The maintenance of proper stereochemistry is not essential since the restrained refinement procedure will correct these errors. Primary concern is with moving the group of atoms into regions of electron density. After resetting the overall thermal factor to 14 Å^2, several cycles of refinement reduced the R factor to 0·304 and the r.m.s. shifts to 0·02 Å per axial direction for the last cycle.

Examination of the R factors now indicated poor agreement for reflections with $d > 6$ Å and these reflections were removed from the data on the basis of solvent contributions. This was also taken as an indication of a well ordered solvent structure.

The structure was now re-examined three residues at a time, applying corrections by rotating about individual bonds. The procedure involved calculating and contouring the Fourier map, superimposing the current structure, storing the plot orientation and scale information on a disk file, plotting the results in stereoscopic projection on a Tektronix 'scope and producing the hard-copy image. One would next examine the image, decide which atoms are to be moved, and execute a small program to generate and plot the revised structure. The program (Furey, 1978) uses the disk file previously prepared to plot the new atom positions on the same scale and with the same viewing orientation as the original plot. When a hard copy of the new plot is produced, it can be superimposed on the original copy of the density map and viewed in a light box in stereoscopic projection. The results of the corrections are seen quickly (about 3 min* on our system) using a standard Tektronix 'scope with hard-copy unit. A complete pass through the 113 residues of Rhe required less than 16 h.

After adjusting the residues as described above, more cycles of refinement reduced the R factor to 0·258 with a r.m.s. deviation from ideal bond distances of 0·028 Å. The phases when compared with the MIR phases gave the phase shifts listed in Table 4. It is interesting to note that in every figure of merit range, even the highest, a phase shift was observed that is very near the maximum possible shift of 180·00°. The overall r.m.s. phase shift of 42·83° is in excellent agreement with the phase error (and therefore the expected phase shift, if the refined structure is the correct structure) implied by the average figure of merit

* This time delay is due mainly to the 300 baud line we are using and could be reduced considerably with a faster line.

Table 2. *Some refinement parameters*

795 atoms
2197 distances
144 planes
112 chiral centers
5768 possible contacts (473 actually requiring restraints)
2387 parameters varied
2144 structure factors (3 Å–50 Å data)

Table 3. R_F *values as a function of resolution*

d_{min}	Number of reflections	R (shell)	R (sphere)
9·	96	0·591	0·591
5·	420	0·400	0·443
4·5	188	0·289	0·394
4·0	286	0·295	0·365
3·5	444	0·325	0·354
3·0	710	0·342	0·350

Table 4. *Comparison of refined and MIR phases as a function of figure of merit*

100 × figure of merit*	Number of reflections	Maximum shift	$\langle \Delta\varphi^2 \rangle^{1/2}$†
90–100	513	172·63°	26·61°
80–90	492	179·89	37·26
70–80	281	179·85	44·48
60–70	171	179·68	52·77
50–60	135	179·77	55·30
40–50	103	178·96	64·62
30–40	68	179·63	76·37
20–30	46	170·28	68·86
10–20	27	179·35	91·87
0–10	5	132·58	69·18

* $\langle$Figure of merit$\rangle$ = 0·75.
† $\langle \Delta\varphi^2 \rangle^{1/2}$ = 42·83° (1841 reflections).

of 0·75 which is 41·41°. Fig. 3 is a plot of the R factor *versus* the number of cycles of refinement.

High-resolution data collection and reduction

All previous refinement cycles involved reflection data with a minimum d spacing of 3·0 Å. The next step was to extend the phasing to higher resolution by way of structure factor calculations based on the refined atomic coordinates. High-quality diffraction data to 1·5 Å were collected on a single crystal which showed a maximum decrease in standard reflection intensities of 15% for 1·6 Å data and 25% for 1·5 Å data. Intensity data were collected on a four-circle diffractometer by the θ–2θ scan technique at room temperature with graphite-monochromatized Cu $K\alpha$ radiation. No background counts were made which permitted data collection on one crystal by reducing the data-acquisition time. The background counts were determined by the method of Krieger, Chambers, Christoph, Stroud & Trus (1974) and applied after all data were collected.

The crystal (sealed in a capillary along with mother liquor) was located 230 mm from the focal spot and 228 mm from the scintillation counter. The scan range was 1·4° per reflection with a scan rate of 1° min^{-1} in 2θ. Intensity data were corrected for crystal deterioration by four-point interpolations of three standard reflection intensities (Shiono, 1971), which were measured at intervals of 50 reflections throughout the data collection. Absorption corrections were applied using the method of North, Phillips & Mathews (1968), except that the relationship $\zeta = \delta = \tan^{-1}(\tan\theta\cos\chi)$ was used to calculate the values of ζ and δ directly from the diffractometer setting angles.

A Wilson plot, calculated from the 19 906 unique structure amplitude observations to 1·5 Å, was based on scattering content deduced from the tentative sequence and solvent estimates reported earlier (Wang & Sax, 1974). The Wilson plot, shown in Fig. 4, indicated an overall thermal factor of 11·54 Å^2 for the new data set. There are no indications of intensities leveling off until at least 1·6 Å.

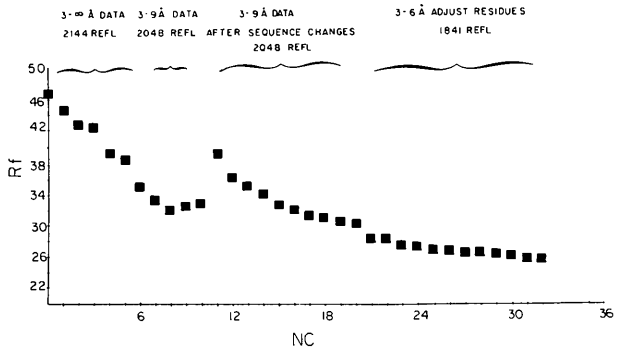

Fig. 3. R_F *versus* number of cycles of refinement. R_F values are ×100. Weights for restraints were derived from the e.s.d.'s in Table 1, column B. Structure factor e.s.d.'s were set to $0{\cdot}4\,\langle|\Delta F|\rangle_d$.

Phase extension and high-resolution refinement

Prior to introducing the high-resolution data, it is instructive to consider the ratios of observations to parameters varied which were observed in the previous refinement cycles. The refinement cycles involving the 3·0–6·0 Å data utilized only 1841 structure factor observations for the 2387 parameters varied yet the refinement showed little tendency to diverge. We believe that the reason for the success of the method relates to the treatment of restraints as if they were observations. The first effect introduced by this treatment is to increase the number of observations by the number of restraints, and therefore increase the overdeterminancy ratio simply by adding more 'observations'. A second and more powerful effect is that because the restraints force the shifts to be highly correlated the trial model behaves somewhat like a rigid body, despite the fact that each parameter is still permitted to vary. It is, therefore, difficult to assess the 'effective' overdeterminancy ratio; however, it is clear that the ratio of structure factor observations to parameters varied is an underestimation.

In view of the smooth refinement path achieved with such a small number of structure factor observations, it was decided to introduce subsets of the reflection data rather than the complete data set when extending the phasing and therefore reduce computer time. The

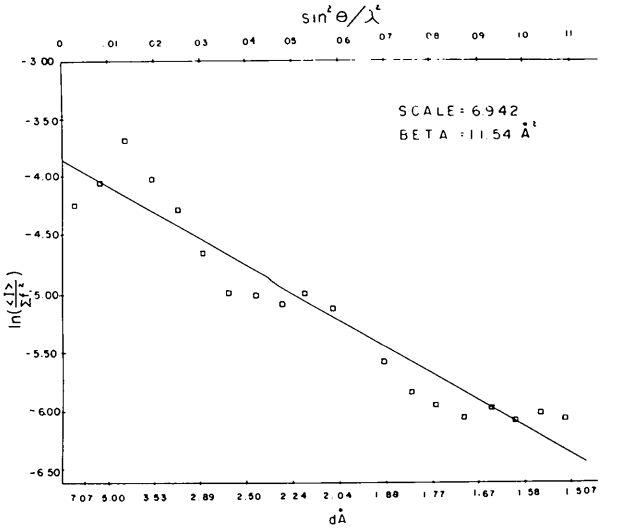

Fig. 4. Wilson plot for Rhe. All data were collected on the same crystal.

phasing was initially extended to a subset of reflections with $d_{min} = 2{\cdot}25$ Å. The subset was selected by:

(*A*) preparing a file with reflections with *d* spacings ranging from 5·0 to 2·25 Å;

(*B*) reordering the file according to increasing $\sin\theta/\lambda$ and then dividing the file into ten equally populated segments;

(*C*) reordering each segment individually according to increasing $|F|$;

(*D*) selecting 30% of the reflections in each region with the largest $|F|$'s and 5% of the reflections in each region with the lowest $|F|$'s [provided $I/\sigma(I) \geq 3$].

Using 35% of the new data (2212 reflections) with *d* ranging from 2·25 to 5 Å, an overall thermal factor of 12 Å^2, and the previously refined coordinates as initial parameters, the refinement was continued for several cycles. Following an initial increase in the residual R_F, due mainly to scaling of the new data set, the residual was reduced to 0·247.

When the r.m.s. shifts became negligible (~0.017 Å per axial direction) the phasing was extended further to 1·9 Å using a subset of the data with *d* ranging from 1·9 to 5·0 Å selected by the process previously described, except that the cut-offs were modified to accept 29% of the largest $|F|$'s and 3% of the lowest $|F|$'s per region. The resulting 3021 reflections were introduced and the refinement was resumed. After several cycles the residual had been reduced to 0·238 with r.m.s. shifts of 0·02 Å per axial direction indicated on the last cycle.

More data were introduced next by reselecting the 1·9 Å subset utilizing 64% of the data (59% largest $|F|$'s and 5% lowest $|F|$'s). The resulting 6042 reflections were used in several cycles of refinement. After two cycles, the r.m.s. deviation from ideal bond distances was found to have increased to 0·044 Å and the estimated standard deviations for structure factors were therefore increased from 0·4 $\langle|\Delta F|\rangle_d$ to 0·5 $\langle|\Delta F|\rangle_d$ to put more emphasis on the restraints. In addition, some of the weaker reflections were removed by raising the minimum $I/\sigma(I)$ cut-off ratio to 5.

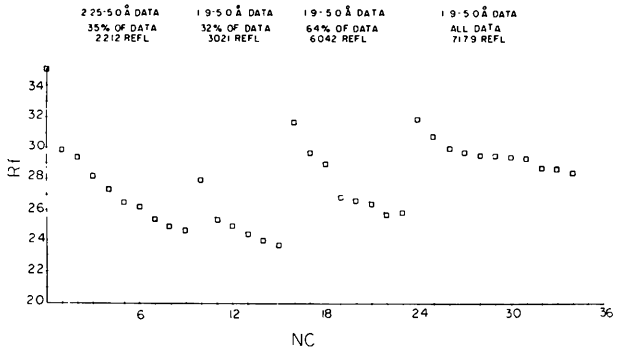

Fig. 5. R_F *versus* the number of cycles of refinement during the phase extension. R_F values are × 100.

Several more cycles of refinement were carried out resulting in a leveling off of the residual at $R_F = 0{\cdot}258$ and r.m.s. shifts of 0·01 Å indicated on the last cycle.

The next step was to include all reflection data with *d* ranging from 1·9 to 5·0 Å and $I/\sigma(I) \geq 5$. Several cycles of refinement resulted in an increase in the residual to $R_F = 0{\cdot}284$. At this point the r.m.s. shifts were ~0·01 Å per axial direction. The r.m.s. deviations from ideal bond distances and planarity were 0·038 and 0·028 Å respectively. The r.m.s. deviations from ideality for all other restraint types were approximately equal to their respective estimated standard deviations. The overall r.m.s. shift from the initial atomic positions was 0·599 Å and the 1·9 Å refinement was complete. Fig. 5 is a plot of the residual as a function of the number of refinement cycles during the phase extension.

Assessment of refinement and phase extension

A program was written to assess the quality of the refined structure by calculating the electron density and difference electron density at each atom site. The results are given in Table 5. The maximum electron density was observed at the position of the sulfur in Cys 89 and was 3·9 e Å^{-3}, a site which had an electron density of 2·0 e Å^{-3} in the MIR map.

Table 5. *Electron density [including F(000) term, F(000)/V ~ 0·39 e Å^{-3}; coefficients were $F_o \exp(i\varphi_c)$]*

Main chain atom types	Minimum	Maximum	Average
N	1·12 e Å^{-3}	2·41 e Å^{-3}	1·80 e Å^{-3}
C*A*	0·97	2·05	1·47
C	1·17	2·46	1·79
O	1·39	2·84	2·22
All side-chain atoms	0·59	3·90	1·66

Minimum difference electron density [calculated with coefficients $(F_o - F_c)\exp(i\varphi_c)$, omitting the $F(000)$ term] = −0·44 e Å^{-3}.

Fig. 6. Stereoscopic projection of residues Gln 6, Pro 7, Pro 8, and Ser 9 superimposed on a residue-deleted 1·9 Å $2F_o - F_c$ map. The illustrated atoms were omitted from the structure factor calculation which produced the coefficients $(2F_o - F_c)\exp(i\varphi_c)$ used in the Fourier synthesis. Contours start at 0·30 e Å^{-3} and increase in steps of 0·75 e Å^{-3}.

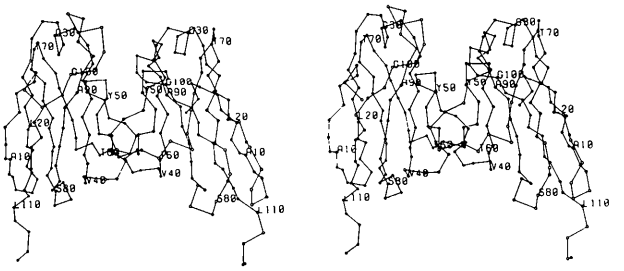

Fig. 7. Stereoscopic drawing of the α carbon structure of the Rhe dimer.

Despite the fact that the R factor increased by several per cent when all reflection data were introduced, the use of data subsets still appears to be valid in terms of the atomic coordinates. The r.m.s. difference in atomic positions between results obtained with all the data and 32% of the data was only 0·057 Å.

The refinement process has resulted in several sequence changes which will be discussed elsewhere. There are two features of the refined structure which are noteworthy. Pro 8 is in the *trans* as opposed to the *cis* conformation (see Fig. 6) observed for Pro 8 in Rei (Huber & Steigemann, 1974). There are no indications

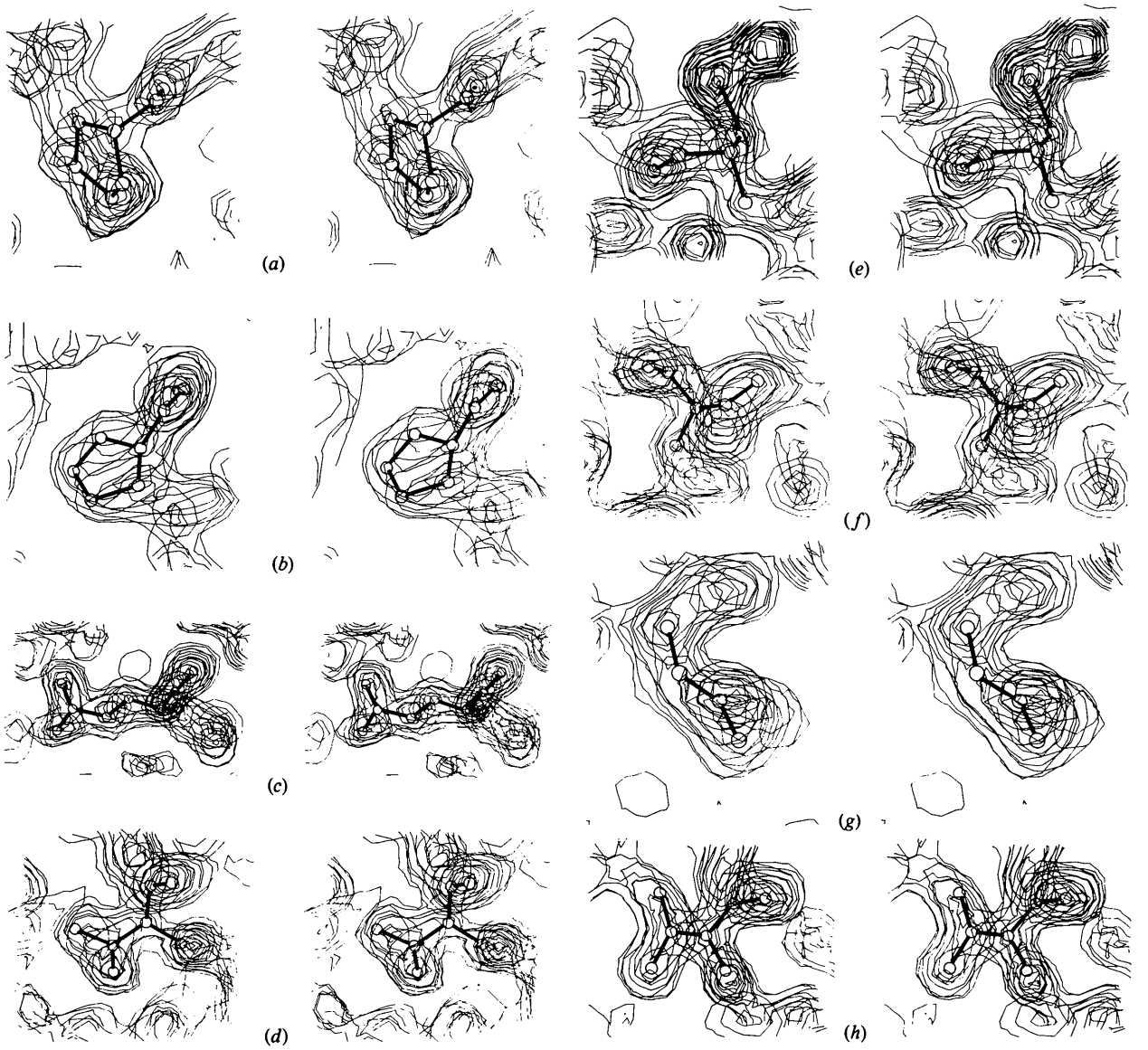

Fig. 8. Some typical residues superimposed on the 1·9 Å residue-deleted electron density map. The illustrated atoms were omitted (approximately 7% of total protein scattering matter) from the structure factor calculation which produced the coefficients $(2F_o - F_c) \exp(i\varphi_c)$ used in the Fourier synthesis. Contours start at 0·30 e Å^{-3} and increase in steps of 0·50 e Å^{-3}. The residues are: (*a*) Pro 7, (*b*) Pro 8, (*c*) Gln 16, (*d*) Thr 19, (*e*) Cys 22, (*f*) Ser 27, (*g*) Gly 30, (*h*) Val 34.

of disorder among the disulfides, as was observed in Rei. Fig. 7 is a stereoscopic drawing of the α carbon structure of the Rhe dimer.

The restrained reciprocal-space refinement scheme appears to work well even when subsets of the master data set are used, provided the relative emphasis on structure factor observations is restricted to the range of 4 to 6 times the emphasis placed on restraint observations. These figures are arrived at by evaluating typical entries from terms 1 and 2 of φ. If $\sigma(F) = 0{\cdot}5\langle|\Delta F|\rangle_d$, then a typical entry to the sum in term 1 is $W|F_o - F_c|^2 = (4|F_o - F_c|^2)/\langle|\Delta F|\rangle_d^2 \sim 4$ provided the particular ΔF is not far from the average value of ΔF for that particular resolution. A similar evaluation of a typical entry to term 2 of φ yields $W|d_o - d_I|^2 = |d_o - d_I|^2/(0{\cdot}03)^2 \sim 1$ provided the particular Δd is not far from the average Δd which is taken to be $\sim 0{\cdot}03$ Å as indicated by the overall r.m.s. deviation from ideality. Use of $0{\cdot}4\langle|\Delta F|\rangle_d$ as $\sigma(F)$ would then yield $W|F_o - F_c|^2 \sim 6$.

As expected, the refinement path followed was shown to be smooth with the exception of places where the model atoms or reflection data were altered by user intervention. The lack of a high overdeterminancy ratio did not appear to hinder the refinement. In general the refined atoms fit well into the calculated electron density for most residues, particularly those which are not on the surface of the protein. This fit is illustrated in Fig. 8 where some typical residues are seen superimposed on the 1·9 Å electron density maps.

A disappointing aspect of the refinement is illustrated in Fig. 9 where it can be seen that the side chain of Thr is incorrectly oriented by a rotation about the $CA-CB$ bond. However, once OG1 refined into a site with high electron density, the use of the restraints prevented CG2 from moving into its proper location. This suggests that it may be useful to occasionally remove the restraints in order to reposition badly misplaced atoms.

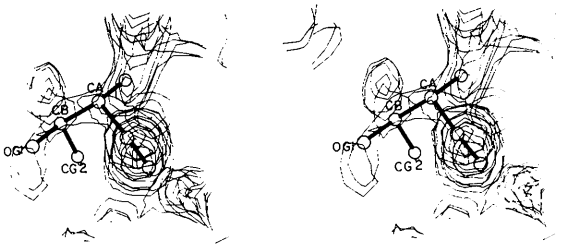

Fig. 9. 1·9 Å residue-deleted electron density map with the illustrated atoms omitted from the structure factor calculation which produced the coefficients $(2F_o - F_c)\exp(i\varphi_c)$ used in the Fourier synthesis. Contours start at 0·30 e Å^{-3} and increase in steps of 0·50 e Å^{-3}. Thr is apparently incorrectly placed. The correction implied is a rotation about the $CA-CB$ bond.

All calculations were performed on a PDP 10 computer with a KL10 processor. The refinement cycles with ~2400 variables and numbers of structure factors of 1841, 3021 and 7179 required 8·8, 13·9 and 27·0 min/cycle respectively. The latter time compares well with the time of 26 min/cycle for a similar size problem, cytochrome C2 (Freer, Alden, Levens & Kraut, 1976), which was refined by the conventional technique of Fourier synthesis followed by geometry optimization.

The authors are grateful to Dr W. Hendrickson for supplying a copy of the refinement program and for some helpful suggestions regarding weighting. We are also grateful to the University of Pittsburgh Computer Center for the use of their computer and to Dr W. Shin and Mr A. Turano for their assistance in the programming and calculations.

This work was supported by the Medical Research Service of the VA Medical Center and by NIH Grant AM-CA18827.

References

FREER, S. T., ALDEN, R. A., LEVENS, S. A. & KRAUT, J. (1976). *Crystallographic Computing Techniques*, edited by F. R. AHMED, p. 319. Copenhagen: Munksgaard.

FUREY, W. (1978). *ROPLOT*. Unpublished.

HENDRICKSON, W. & KONNERT, J. H. (1977). Private communication.

HUBER, R. & STEIGEMANN, W. (1974). *FEBS Lett.* **48**, 235–237.

KONNERT, J. H. (1976). *Acta Cryst.* A**32**, 614–617.

KRIEGER, M., CHAMBERS, J. L., CRISTOPH, G. G., STROUD, R. M. & TRUS, B. L. (1974). *Acta Cryst.* A**30**, 740–748.

NORTH, A. C., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

SHIONO, R. (1971). Tech. Rep. 49. Department of Crystallography, Univ. of Pittsburgh, Pittsburgh, PA, USA.

STENKAMP, R. E. & JENSEN, L. H. (1976). *Acta Cryst.* A**32**, 255–258.

WANG, B. C. & SAX, M. (1974). *J. Mol. Biol.* **87**, 505–508.

WANG, B. C., YOO, C. S. & SAX, M. (1979). *J. Mol. Biol.* **129**, 657–674.