sistent with a greater order, *i.e.* large superstructure domains themselves more perfectly ordered.*

Conclusion

The conclusions of the present study of β-alumina with a composition of 8·3 Al_2O_3–Ag_2O are:

(i) That the low-temperature superstructure arising from the conducting ions never achieves long-range order and saturates around liquid-nitrogen temperature into domains of about 45 Å diameter.

(ii) That even in these short-range superstructures, the order is not complete, decreases with increasing temperature and extrapolates, around room temperature, to values close to the average occupation determined by conventional structure analysis.

We are grateful to D. B. McWhan and S. M. Shapiro for fruitflu discussions and to L. Deschamps for his efficient technical help during the low-temperature experiments.

* Just after this work had been submitted for publication, we received a preprint on the ion–ion correlation and diffusion in several β-alumina-type compounds (McWhan, Allen, Remeika & Dernier, 1975), which yields somewhat different results. Noteworthy is the fact that the Na β-alumina crystals (which are the starting material) used have been grown from a flux and have a slightly different composition ($1{\cdot}6Na_2O.11Al_2O_3$ compared to our $1{\cdot}33Na_2O.11Al_2O$). This could be directly relevant to the present discussion.

References

BEEVERS, C. A. & ROOS, M. A. S. (1937). *Z. Kristallogr.* **97**, 59–66.

BRAGG, W. L., GOTTFRIED, C. & WEST, J. (1931). *Z. Kristallogr.* **77**, 255–274.

COMÈS, R., LAMBERT, M., LAUNOIS, H. & ZELLER, H. R. (1973). *Phys. Rev.* B**8**, 571–575.

GELLER, S. & OWENS, B. B. (1972). *J. Phys. Chem. Solids*, **33**, 1241–1250.

LE CARS, Y., COMÈS, R., DESCHAMPS, L. & THERY, J. (1974). *Acta Cryst.* A**30**, 305-309.

MCWHAN, D. B., ALLEN, S. J. JR, REMEIKA, J. P. & DERNIER, P. D. (1975). *Phys. Rev. Lett.* **35**, 953.

OWENS, B. B. & ARGUE, G. R. (1967). *Science*, **157**, 308–309.

PETERS, C. R., BETTMAN, M., MOORE, J. W. & GLICK, M. D. (1971). *Acta Cryst.* B**27**, 1826–1834.

REIDINGER, F. (1974). Unpublished, referred to by W. L. ROTH, General Electric Rep. No. 74CRDO54, March 1974.

RICE, M. J. & ROTH, W. L. (1972). *J. Solid State Chem.* **4**, 294–310.

ROTH, W. L. (1972). *J. Solid State Chem.* **4**, 60–75.

WHITTINGHAM, M. S. & HUGGINS, R. A. (1971). *J. Chem. Phys.* **54**, 414.

YU YAO, Y. F. & KUMMER, J. T. (1967). *J. Inorg. Nucl. Chem.* **29**, 2455–2475.

Acta Cryst. (1976). A**32**, 255

Test Refinements with Simulated Protein Data Sets

BY R. E. STENKAMP AND L. H. JENSEN

Department of Chemistry and Department of Biological Structure, University of Washington, Seattle, WA. 98195, U.S.A.

(*Received* 27 *June* 1975; *accepted* 26 *September* 1975)

The method of differential difference syntheses combined with idealization for refining protein models has been tested on a dipeptide derivative. An r.m.s. error of 0·5 Å was introduced in the dipeptide model, and the data were modified by applying artificial temperature factors to simulate the case of a protein. In any one of several variations, differential difference syntheses with idealization led to convergence at *R* values substantially lower than for proteins, but refinement is slow, requiring many cycles.

Two methods of refining a protein model based on X-ray diffraction data are by difference syntheses (Watenpaugh, Sieker, Herriott & Jensen, 1973) and by differential difference syntheses combined with application of constraints to maintain acceptable bond lengths, interbond angles and certain torsion angles (Freer, Alden, Carter & Kraut, 1975). To check the differential difference method with the application of constraints we have undertaken a series of test refinements using a small-molecule data set limited and modified to simulate that of a protein. These tests offer two advantages: (1) the parameters from the test refinements can be compared with accurate values from the conventional refinement, and (2) because the test case is small relative to a protein structure, it is economically feasible to compare variations of the basic method of refinement.

Test refinement

The structure chosen for the test was *N*-acetyl-L-phenylalanyl-L-tyrosine (NAPT) (Stenkamp & Jensen, 1973). Table 1 contains relevant crystal, refinement, and data set information. The data set was truncated at *d* spacings typical of high-resolution protein data that

have been used in actual refinement, and it was modified by applying temperature factors of the form $\exp(-\Delta B \sin^2\theta/\lambda^2)$ to raise the effective thermal parameter, B, to 20 Å^2 for one of the tests and to 12 Å^2 for the rest of them. The test refinements are described in Table 2.

Table 1. *Crystal, refinement, and data set information*

$a=$ 11·530 (5) Å	22 Nonhydrogen and
$b=$ 8·589 (3)	27 hydrogen atoms
$c=$ 10·635 (3)	Refinement used 2725
$\beta=114{\cdot}52$ (2)°	reflections $>\sigma(I)$
$Z=$ 2	$\sin\theta_{max}/\lambda=0{\cdot}715$
	$R=0{\cdot}047$

For the modified data set:
148 reflections to 2·0 Å resolution ($\sin\theta/\lambda=0{\cdot}250$)
334 reflections to 1·5 Å resolution ($\sin\theta/\lambda=0{\cdot}333$)
$\exp(-8{\cdot}83\sin^2\theta/\lambda^2)$ applied to F_o to give data set with overall $B=12$ Å^2.
$\exp(-16{\cdot}83\sin^2\theta/\lambda^2)$ applied to F_o to give data set with overall $B=20$ Å^2.

Table 2. *Description of tests*

Test No.	Resolution	Overall B	Description
I	1·5 Å	12·0 Å	Idealization after every third refinement cycle.
II	1·5	12·0	Free refinement (no constraints).
III	1·5	20·0	Idealization after every third refinement cycle.
IV	1·5	12·0	Idealization after each refinement cycle.
V	2·0	12·0	Idealization after each refinement cycle, curvatures for 2·0 Å resolution.
VI	2·0	12·0	Idealization after each refinement cycle, curvatures for 1·5 Å resolution.

Random rotations about the free torsion angles were applied to the model from the conventional structure refinement to give an approximate one with an r.m.s. deviation of the atoms from their true positions of 0·5 Å. This served as the starting model for tests in which refinement was by differential difference synthesis and application of constraints was by a program originally written by J. Hermans (Hermans & McQueen, 1974).

Test I

Fig. 1 is a plot of $R(=\sum||F_o|-|F_c||/\sum|F_o|)$ and the r.m.s. distance of the atoms from their true positions for the coordinate sets generated in Test I. The behavior of R with refinement is similar to that found for proteins; successive cycles of differential difference syntheses decrease it while constraining bond lengths and angles (idealization) raises it. However, in this test, R converges to 0·1 in contrast to values in the region of 0·2 for protein models with constraints. A noteworthy feature of these tests which cannot be observed in protein refinements is the decrease in the r.m.s. distance of the atoms in the model from their true positions for the first three idealizations.

Table 3 shows the results of refinement in Test I on three parameters of the model: the distance of the acetyl methyl carbon atom, C(1), from its true position, its B value, and the peptide dihedral angle, ω. The refinement greatly improved the position of C(1), the atom with the greatest error, reducing the distance from its true position from 1·16 Å to 0·06 Å, as close as the other atoms to their true positions. The peptide dihedral angle, ω, has a value of 162·3 (4)° in the precise model. Idealization tends to drive it toward 180°, but the refinement returns it to the vicinity of its true value.

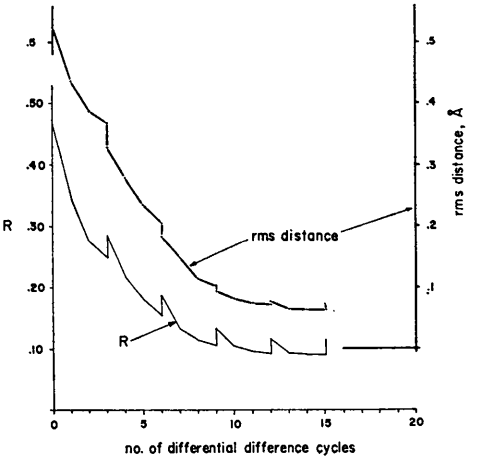

Fig. 1 Test I: R and r.m.s. distance from true atomic positions *vs* refinement cycle. 1·5 Å data set, idealization after every third cycle.

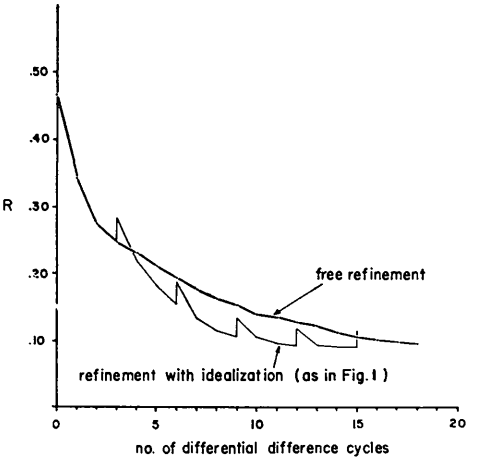

Fig. 2. Test II: R *vs* refinement cycle. 1·5 Å data set.

Test II

To check the effects of idealization, the refinement tests were repeated without constraints, Fig. 2 showing the R values for this test. The free refinement requires more cycles to converge, and some of the bond lengths at the end of the refinement are in error by as much as 0·1 Å. Although the constrained refinement results in more acceptable bond lengths and angles because they have been idealized, the differences in the atomic positions at the end of the two refinements are, nevertheless, insignificant, the r.m.s. distance between corresponding atoms being 0·1 Å.

It is surprising that atom C(1) with an error of 1·16 Å refined to the correct position. In general atoms with errors of this magnitude cannot be expected to converge to their true positions without the application of constraints.

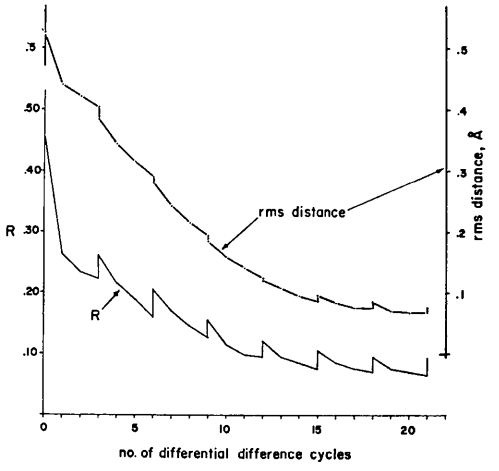

Fig. 3. Test III: R and r.m.s. distance from true atomic positions *vs* refinement cycle. 1·5 Å data set, overall $B=20$ Å^2, idealization after each cycle.

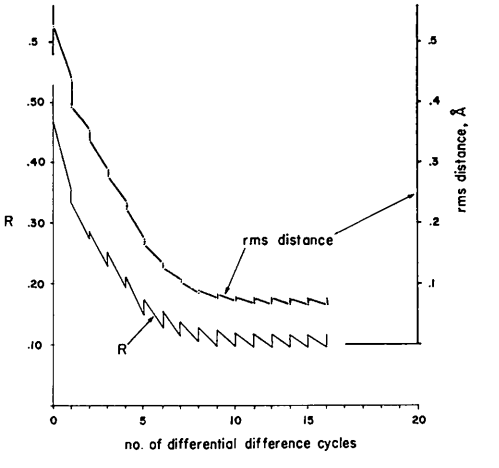

Fig. 4. Test IV: R and r.m.s. distance from the true atomic positions *vs* refinement cycle. 1·5 Å data set, idealization after each cycle.

Table 3. *Some structural information from* 1·5 Å *refinement, idealization after every third refinement cycle*

	atom C(1)		
Refinement cycle	Distance from true value	B	Peptide dihedral angle, ω [true value = 162·3 (4)°]
Initial values	1·16 Å	12·0 Å	180·0°
1	1·21	12·2	157·9
2	1·10	12·9	159·3
3	1·06	13·6	158·3
Idealized	0·98	—	175·9
4	0·83	14·7	164·2
5	0·63	15·5	161·7
6	0·49	16·2	160·6
Idealized	0·50	—	174·6
7	0·29	16·1	167·6
8	0·16	16·0	168·0
9	0·07	15·9	166·7
Idealized	0·13	—	174·1
10	0·05	15·9	168·3
11	0·04	15·9	170·2
12	0·04	15·9	166·7
Idealized	0·06	—	173·8
13	0·03	15·9	167·7
14	0·04	15·9	168·2
15	0·04	15·9	166·7
Idealized	0·06	—	173·6

Test III

To test the effect of refining data with large B values, the NAPT data set was modified to an overall $B=20$ Å^2 and the refinement scheme of Test I was repeated. The results are shown in Fig. 3. The decrease in R is much greater in the first cycle for the data set with an overall B of 20 Å^2 and it is reduced to lower values in the later stages of the refinement. The lower R stems from the fact that the higher B value reduces the intensities of the high-angle reflections so that the low-angle, more intense ones which are less sensitive to errors in the model are dominant.

Test IV

The R values for Test IV are given in Fig. 4. Comparison with Fig. 1 shows that idealization after each refinement cycle results in more rapid convergence but to essentially the same R as Test I. The main difference between the two schemes is that the first application of constraints in Test IV decreases R, presumably because the model was still so poor at that point.

Test V

Since protein data sets have not usually extended to d spacings less than 2 Å, the NAPT data set was limited to 2 Å resolution ($\sin\theta/\lambda=0{\cdot}25$) and the refinement was repeated as in Test IV. The results are shown as A in Fig. 5. Comparison with Fig. 4 shows that refinement with 2 Å data converges to essentially the same R as that with the 1·5 Å data set, but initially the decrease in R is more rapid. This suggests that a judicious choice of resolution limit early in a refinement could lead to more rapid convergence. Moreover, atoms greatly in error can be moved farther with low-resolution data than with high, but it is essential to

minimize the number of parameters by use of constraints when refining with low-resolution data.

Comparison with Fig. 4 indicates that near convergence idealization causes larger increases in R for the refinement with 2 Å data than with the 1·5 Å data. With the 2 Å data the model adjusts to give lower R values, but idealization raises R to essentially the same value for both data sets.

Test VI

In refining by differential difference syntheses, corrections to the positional parameters for structures referred to orthogonal axes are given by the expression $\delta(\text{coord.}) = -\text{slope/curvature}$, and the appropriate curvature for an atom will depend on the resolution of the data set used in the refinement. In Tests I–IV the curvatures used were 2, 2·5, and 3 e $Å^{-5}$ for the 1·5 Å resolution data and half these values for Test V with the 2 Å resolution data. They are likely to be underestimates of the true curvatures, possibly by as much as two times, so that their use avoids the need for the double-shift factor of two for non-centrosymmetric structures (Cruickshank, 1950).

The results of refinement with inappropriately large curvatures is shown as B in Fig. 5. After the initial decrease in R, refinement is much slower with the larger curvatures because the atoms are undershifted. These results suggest that choosing a smaller curvature or using a convergence factor greater than unity may be useful in accelerating convergence.

Additional tests

To determine the effect of idealization on the precise model (Stenkamp & Jensen, 1973), we applied constraints to it and found an r.m.s. deviation for the coordinates of 0·05 Å from their true values. This is in the same range as the r.m.s. differences among the coordinate sets from the various refinements, all being less than 0·1 Å.

Since the differential difference tests used a model with isotropic thermal factors, we subjected the *precise model* with *isotropic thermal* factors to full-matrix, least-squares refinement using the 1·5 Å resolution data set as in Test I. R converged to 0·063 and the r.m.s. deviation of the atoms from their true positions was 0·08 Å.

Conclusions

The results reported here show that differential difference syntheses with constraints are effective in refining a dipeptide model having errors approximating those expected for protein models derived from electron density maps based on experimental phases. Although the refinements converged slowly, the schemes tested all led to essentially the true model. In contrast to proteins which converge at R values in the region of 0·2, the dipeptide model refined to an R of approximately 0·1 by differential difference syntheses. The higher R's for protein refinement can be attributed in part to the greater errors inherent in collecting protein data. It appears likely, however, that a significant source of error is the model itself and that differential difference syntheses have insufficient power to lead to the best model for the limited data sets available.

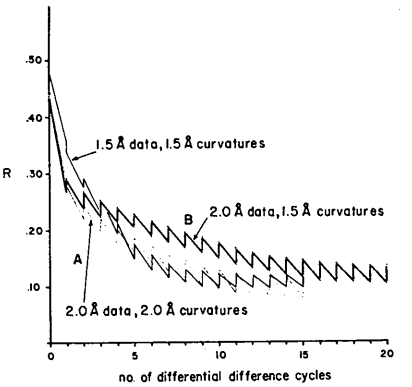

Fig. 5. Tests V and VI: R *vs* refinement cycle. 2·0 Å data set with two different sets of curvatures, idealization after each cycle.

This work was supported by Grant GM-10828 from the National Institutes of Health, U.S. Public Health Service.

References

CRUICKSHANK, D. W. J. (1950). *Acta Cryst.* **3**, 10–13.
FREER, S. T., ALDEN, R. A., CARTER, C. W. & KRAUT, J. (1975). *J. Biol. Chem.* **250**, 46–54.
HERMANS, J. & MCQUEEN, J. (1974). *Acta Cryst.* A**30**, 730–739.
STENKAMP, R. E. & JENSEN, L. H. (1973). *Acta Cryst.* B**29**, 2872–2877.
WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1973). *Acta Cryst.* B**29**, 943–956.