munity and the Deutsche Forschungsgemeinschaft for their support of this work.

References

BAGGIO, R., WOOLFSON, M. M., DECLERCQ, J.-P. & GERMAIN, G. (1978). *Acta Cryst.* A34, 883-892.
BIANCHI, R., PILATI, T. & SIMONETTA, M. (1978). *Acta Cryst.* B34, 2157-2162.
BRACKMAN, J. C., DALOZE, D., DUPONT, A., TURSCH, B., DECLERCQ, J.-P., GERMAIN, G. & VAN MEERSSCHE, M. (1981). *Tetrahedron*, 37, 179-186.
COLENS, A., DECLERCQ, J.-P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSSCHE, M. (1974). *Cryst. Struct. Commun.* 3, 119-122.
DEBAERDEMAEKER, T., TATE, C. & WOOLFSON, M. M. (1985). *Acta Cryst.* A41, 286-290.
DEBAERDEMAEKER, T. & WOOLFSON, M. M. (1983). *Acta Cryst.* A39, 193-196.
DECLERCQ, J.-P., GERMAIN, G. & KING, G. S. D. (1977). Fourth Eur. Crystallogr. Meet., Oxford. Abstracts A, pp. 279-280.
DECLERCQ, J.-P., GERMAIN, G. & VAN MEERSSCHE, M. (1972). *Cryst. Struct. Commun.* 1, 13-15.
GILMORE, C. J. (1984). *J. Appl. Cryst.* 17, 42-46.
HULL, S. E. & IRWIN, M. J. (1978). *Acta Cryst.* A34, 863-870.
ROQUES, R., ROSSI, J. C., DECLERCQ, J.-P. & GERMAIN, G. (1980). *Acta Cryst.* B36, 1589-1593.
SCHENK, H. & KIERS, C. T. (1984). In *Methods and Applications in Crystallographic Computing*, edited by S. R. HALL & T. ASHIDA, pp. 96-105. Oxford: Clarendon Press.
SHELDRICK, G. M. (1981). *SHELXTL. A System of Computer Programs for the Solution and Refinement of Crystal Structures.* Nicolet Instruments, Madison, Wisconsin, USA.
SZEIMIES-SEEBACH, U., HARNISCH, J., SZEIMIES, G., VAN MEERSSCHE, M., GERMAIN, G. & DECLERCQ, J.-P. (1978). *Angew. Chem. Int. Ed. Engl.* 17, 848-850.
YAO, J.-X. (1981). *Acta Cryst.* A37, 642-644.
ZHANG, S.-H. & WOOLFSON, M. M. (1982). *Acta Cryst.* A38, 683-685.

Acta Cryst. (1988). A44, 357-368

Refinement at 1·4 Å Resolution of a Model of Erabutoxin b: Treatment of Ordered Solvent and Discrete Disorder

BY JANET L. SMITH*

Department of Biochemistry and Molecular Biophysics and Howard Hughes Medical Institute, Columbia University, New York, NY 10032, *USA, and Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC* 20375, *USA*

P. W. R. CORFIELD

Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, *USA*

WAYNE A. HENDRICKSON

Department of Biochemistry and Molecular Biophysics and Howard Hughes Medical Institute, Columbia University, New York, NY 10032, *USA, and Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC* 20375, *USA*

AND BARBARA W. LOW

Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, *USA*

(*Received* 10 *August* 1987; *accepted* 6 *January* 1988)

Abstract

The latter stages in the refinement of the protein erabutoxin b are described. The crystal structure of the 62-residue protein has been refined to a conventional R factor of 0·144 by stereochemically restrained least-squares methods using diffraction data to a limit of 1·4 Å spacings. Emphasis was placed on determining as accurately as possible the solvent structure and the structures of heterogeneous groups in the protein. The final model includes two conformers for each of seven side chains and for an octapeptide segment. A total of 111 sites for water molecules have been located as well as one sulfate ion with a total of 68 site occupancies. 65 of the solvent sites overlap either with protein atoms belonging to groups in two alternative conformations or with other solvent sites. Dual protein conformers and overlapping solvent sites were both included in the least-squares refinement. Individual thermal and occupancy parameters were refined for solvent molecules. An

* Author to whom correspondence should be addressed at Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA.

0108-7673/88/030357-12$03.00

analysis of these parameters has provided useful structural information.

Introduction

Erabutoxin b is a postsynaptic neurotoxic protein of 62 amino-acid residues found in venom of the Pacific sea snake *Laticauda semifasciata*. The three-dimensional structure of erabutoxin b was reported previously from analyses at lower resolution (Low, Preston, Sato, Rosen, Searl, Rudko & Richardson, 1976; Kimball, Sato, Richardson, Rosen & Low, 1979) as were the results of an earlier stage of the crystallographic refinement at 1·4 Å resolution (Bourne, Sato, Corfield, Rosen, Birken & Low, 1985; Low & Corfield, 1986). The earlier refinement provided starting parameters for the work reported here, corrected errors in the amino-acid sequence of erabutoxin b, and established the identity of erabutoxin b and neurotoxin b, toxins from the venom of *L. semifasciata* from different regions of the Pacific. The crystal structure of neurotoxin b was reported by Tsernoglou & Petsko (1976). We describe here a further refinement of the erabutoxin b model using diffraction data to 1·4 Å spacings.

One of the goals of this project was to determine the structure of the ordered solvent in erabutoxin b crystals. Both thermal and occupancy parameters were refined for solvent molecules in the model, and the physical significance of these highly correlated parameters was considered in interpreting the results. As part of an effort to streamline the time-consuming process of addition of solvent to a protein model, all significant density in $|F_o|-|F_c|$ electron density maps was systematically surveyed. In many cases the difference density was more reasonably interpreted as structural heterogeneity rather than as new solvent molecules. The term structural heterogeneity refers to discrete crystallographic disorder (Smith, Hendrickson, Honzatko & Sheriff, 1986). This is ubiquitous in protein crystals owing to limited lattice contacts and to the large proportion of crystalline volume occupied by solvent. Techniques were developed for accommodating the additional structural information in the stereochemically restrained refinement procedure. Various attempts were also made to treat the highly correlated thermal and occupancy parameters of partially occupied protein and solvent atomic sites.

A summary of the refinement process is presented here. A full description of the structure of erabutoxin b will be presented elsewhere.

Methods

1. *Refinement*

The erabutoxin b model was refined by the stereochemically restrained least-squares methods developed by Hendrickson & Konnert (1980; Konnert, 1976). Each stage of refinement consisted of interpretation of electron density maps, additions or corrections to the protein/solvent model and refinement of the new model. Solvent molecules were added directly to the model at the positions of maxima of appropriate $|F_o|-|F_c|$ electron density peaks, aided by a set of locally developed computer programs. All other significant $|F_o|-|F_c|$ electron density was inspected on an interactive graphics system and interpreted with the aid of the geometric information provided by these programs. Electron density maps calculated from $2|F_o|-|F_c|$ syntheses were also studied. Refinement of the resulting model was then carried out. Five to ten cycles of shifts were applied to the parameters in each phase of refinement. Weights employed for the stereochemical 'observations' in the least-squares refinement were the inverses of the target variances for the stereochemical features and the values used were generally those described by Hendrickson (1981). The weights for the structure amplitudes were adjusted during refinement to be as high as possible without the variances for the stereochemical features greatly exceeding their target values. Positional and isotropic thermal parameters were refined for all atoms in the model.

The aim of each phase of refinement was to account for peaks in the $|F_o|-|F_c|$ difference map by stereochemically sensible additions or corrections to the model. Nearly all peaks higher than 0·3 e Å^{-3} in $|F_o|-|F_c|$ maps could be interpreted as new solvent sites, new positions for incorrectly placed atoms in the model or structural heterogeneity (see below). Most of the peaks lower than 0·3 e Å^{-3} were not interpretable and we therefore did not attempt to analyze $|F_o|-|F_c|$ maps beyond this level.

Most of the calculations were made with a Texas Instruments Advanced Scientific Computer at the Naval Research Laboratory and most of the graphics fitting was done in the program *FRODO* (Jones, 1982) on an MMS-X interactive graphics system (Barry, Molnar & Rosenberger, 1976) at that institution. The final stages of the work were done with a VAX 11/780 computer and an Evans & Sutherland Multi-Picture System, both at Columbia University.

2. *Definition of solvent structure*

Solvent was added to the model as oxygen atoms of water molecules. When a significant peak (height greater than about 0·35 e Å^{-3}) in the $|F_o|-|F_c|$ map had reasonable hydrogen-bonding contacts and no very short contacts with the model or other peaks, a water oxygen was straightaway added to the model at that position. Individual atomic occupancy and thermal parameters were shifted in alternate cycles of refinement. Usually these shifts were not applied

fully, but were damped to 50–80% of their calculated values.

Initial occupancy parameters were set according to the height of the peak in the $|F_o|-|F_c|$ map, with occupancy varying between 1·0 and 0·4 for peak heights of 0·8 to 0·3 e Å^{-3}. Initial thermal parameters were assigned based on the number of potential hydrogen bonds that could form between each solvent molecule and the protein. For three or four hydrogen bonds, a B value of 13 Å^2 was assigned (average B value for the protein when solvent was initially added); 15 Å^2 was assigned for solvent molecules with two hydrogen bonds to the protein, 17 Å^2 for one hydrogen bond, and 20 Å^2 for hydrogen bonds only to solvent atoms. All solvent molecules had potential hydrogen bonds to some other atoms in the model; no isolated peaks were added to the model.

Near the end of the refinement, solvent molecules were removed whose occupancies were less than 0·40 or whose thermal parameters were greater than 30 Å^2 and a few cycles of refinement were then carried out. Of the solvent molecules removed, those with peaks higher than 0·3 e Å^{-3} in the resulting $|F_o|-|F_c|$ map were added back to the model.

3. *Characterization of structural heterogeneity*

In many cases significant difference electron density clearly indicated instances of discrete disorder, or structural heterogeneity, in the erabutoxin b crystals. These instances involved either alternative conformers for groups in the protein or solvent sites partially overlapping other atoms in the model. They were modeled as partially occupied atomic sites, and positions and thermal parameters for all of them were refined along with the nonheterogeneous parts of the model in each refinement cycle. Mindful that the sum of occupancies for dual conformers of a protein group or for a pair of partially overlapping solvent sites could not exceed unity, most of these groups or atoms were initially assigned fractional occupancies of 0·5. Different approaches were taken to refining structural heterogeneity in the protein and in the solvent.

For dual conformers of protein groups, individual atomic thermal parameters, but not occupancies, were refined. The individual atomic occupancies were kept equal for atoms comprising a single conformer of a group and were adjusted occasionally in order that the atomic thermal parameters for alternative conformers refined to similar values. (The preferred approach of refining group occupancy parameters and restraining the total occupancy for all conformers of a group to be unity or less is a planned enhancement to the refinement program.)

For overlapping or heterogeneous solvent sites, both occupancy and thermal parameters were refined initially, as described in the preceding section for nonheterogeneous solvent. This was done to find the best estimates of the occupancies for heterogeneous solvent networks. Once these networks had been determined, their occupancies were altered where necessary and fixed. This was the starting point for subsequent refinement in which thermal parameters for heterogeneous solvent sites, and not occupancies, were varied. Because occupancy (q) and thermal (B) parameters are highly correlated, compensating adjustments were made to the thermal parameters if the occupancies were altered before being fixed. This was an attempt to minimize the perturbation to the refinement caused by occasional large changes in occupancies. Individual B values were adjusted in order to keep constant the quantity $q \exp[-B(\sin^2\theta)/\lambda^2]$ for the diffraction angle (θ) representing the weighted center of data used in the refinement.

Initially, restraints on covalent bonds, bond angles and group planarity were specified individually for 'alternative' conformers of the protein, while the full stereochemical restraints were applied automatically to the 'primary' conformers. Later the program *PROTIN*, which establishes the stereochemical restraints before each phase of refinement, was modified to allow full restraints to be applied to all conformers with the exception that restraints limiting short contacts between nonbonded atoms were not applied to atomic sites determined to be mutually exclusive. It remained necessary only to keep track of which conformers of heterogeneous groups in the protein and which versions of multiple solvent sites could be present simultaneously. This bookkeeping task was important to producing an accurate model for two reasons. First, the sum of occupancies for atoms that appeared to occupy the same volume in the crystal, whether due to static, dynamic or crystal-to-crystal disorder, could not exceed unity. Second, erroneous positions could have resulted from applying nonbonded-contact restraints to pairs of atoms that obviously did not occupy the same volume at the same time.

Diffraction data

Erabutoxin b crystallizes in the space group $P2_12_12_1$ with one protein molecule in the asymmetric unit. The diffraction data employed in the initial stages of this study were those used earlier (Low *et al.*, 1976; Kimball *et al.*, 1979) and in the phases 0 through IV of refinement that are described in the preliminary report (Bourne *et al.*, 1985). This data set did not include estimated standard deviations for structure amplitudes. Details of the data collection were not reported in the earlier publications and are described here in Table 1. Following phase V of refinement, the cell constants were corrected to new values of $a=49{\cdot}94$, $b=46{\cdot}58$ and $c=21{\cdot}59$ Å from the values of $a=49{\cdot}9$, $b=46{\cdot}6$ and $c=21{\cdot}3$ Å that were used in the earlier work. Initially all data from spacings of

Table 1. *Parameters for data collection and reduction*

Crystal	1	2	3	4	5	6
Scan width in ω (°)	0·7	1·0	1·0	1·0	1·0	1·0–1·5
Minimum scan speed (° min^{-1})	1·5	1·5	2·0	2·0	2·0	1·5
Number of points in step scan	13	13	7	7	7	13
Number of check reflections*	3	10	6	7	7	3
Selection of $K\alpha$ peak (M = graphite monochromator, F = nickel filter)	M	F	F	F	F	M
Resolution range (Å)	∞–2·5	2·6–1·9	2·6–1·8	2·0–1·5	1·7–1·4	1·8–1·4†
Number of points measured, includuding Friedel pairs	4006	5174	6836	6608	8479	7833
Number of reflections with $I > 2\sigma(I)$	3664	4398	4097	4708	4271	1940
Range of corrections for fall off of intensities	1·00–1·03	0·97–1·09	0·97–1·09	0·94–1·04	0·91–1·11	0·96–1·06
Range of corrections for absorption	1·01–1·56	1·01–1·16‡	1·00–1·47§	1·01–1·32	1·01–1·14	1·01–1·61
R_{merg}¶	0·032**	0·065	0·100	0·091	0·183	0·182
R_{merg} for $I > 2\sigma(I)$	0·031	0·059	0·080	0·068	0·100	0·110
Scale factor on $\|F_o\|^2$ ††	20·8	5·8	5·9	1·4	1·9	11·8

* Check reflections were usually measured after every 100 intensity measurements.
† Data collection with this crystal was halted owing to shrinkage of the unit cell before measurements in this shell were complete.
‡ Additional corrections made as a function of χ and φ, ranging from 0·90 to 1·10.
§ Additional corrections made as a function of φ, ranging from 0·85 to 1·00.
¶ $R_{merg} = \sum_{i,h} ||F_h|^2_i - \langle|F_h|^2\rangle| / \sum_{i,h} |F_h|^2_i$.
** There was very little overlap between crystal 1 and the other crystals. Thus, this value reflects mainly the agreement between Friedel pairs for crystal 1.
†† Intensities from individual crystals are multiplied by this factor. Thus, a small scale factor implies a more strongly diffracting crystal.

10 to 1·4 Å were used in the refinement (10 300 reflections). After phase XIV only the strongest 87% of structure amplitudes were used in the refinement, and after phase XVI the set of amplitudes used was reduced to the strongest 79%. After phase XVIII preliminary estimates were made of the standard deviations of the structure amplitudes and subsequently only reflections with $|F|/\sigma(|F|)$ greater than 1·5 (81% of the total data) were used in refinement. The elimination of weak data had a significant effect on the R factor and improved the quality of $|F_o| - |F_c|$ maps.

After phase XXII diffraction data from an additional (sixth) crystal were included and errors were estimated in a more rigorous fashion by the *de novo* reprocessing of data from all six crystals. Details of this data reduction are given in subsequent paragraphs and summarized in Table 1. Cell dimensions of the first five crystals remained constant during data collection, and agreed closely with each other, with individual values lying within ±0·1% of the corresponding means. The cell dimensions of the sixth crystal had decreased during data collection and data measured from this crystal after the cell volume had decreased by 0·7% were not used. With the sixth crystal included, the range from minimum to maximum for each cell axis was 0·15 Å. For each crystal, cell constants were obtained by least-squares refinement of the setting angles of 11 to 15 reflections, measured at frequent points during data collection.

Data for hkl and for the Friedel-related $\bar{h}\bar{k}\bar{l}$ forms were collected by ω-scan methods at room temperature on a Syntex $P2_1$ diffractometer with Cu $K\alpha$ radiation. The first stage in data reprocessing involved examination of the step-scan profiles of the strongest reflections for each crystal in order to determine base widths of the peaks. Net intensities were obtained by subtracting background counts made 0·5° away from each edge of the scans. The total time for background measurement was 70–100% of the total scan time. The scan width was always appreciably greater than the peak width, so that the signal-to-noise ratio in some cases (crystals 1, 4 and 5) could be improved by selection of only the inner points of the step scan.

Empirical absorption corrections were made based on ψ scans at high χ angles (North, Phillips & Mathews, 1968). Averages for several scans were used for each crystal. The effects of absorption by an asymmetric arrangement of mother liquor could be seen in systematic variations of $I(+)/I(-)$ values with angles φ and χ, which were quite substantial for crystals 2 and 3. For these crystals, additional corrections to the data were made based on a least-squares fit of the $I(+)/I(-)$ ratios to a linear function of χ and φ. Corrections for fall-off of standard reflections were made, and then data from the six crystals were merged to obtain weighted averages of the $|F|^2$ values. Weights were taken as $|F|/\sigma(|F|^2)$ rather than the usual $1/\sigma(|F|^2)$, in order to avoid overemphasizing the weakest estimates for the very weak reflections. The data had been put on an approximately absolute scale by comparison with calculated structure factors, and final relative scale factors were found during the merging process by a least-squares method involving only reflections with I greater than $1·5\sigma(I)$. Of a total of 38 936 measurements recorded, 5000 were not included in the averaging: 4250 had I less than zero, 474 represented intensities measured only once, and a further 276 were rejected on the basis of a number of statistical tests. On averaging the remaining 33 936 measurements, the mean deviation of individual $|F|^2$ values from their average, R_{merg}, was 0·069. In a separate calculation where only reflections with I greater than $2\sigma(I)$ were used, R_{merg} was 0·053. The final set of 10 445 structure factors in the resolution range infinity to 1·4 Å had 10 328 reflections with $|F|^2$ greater than zero. In the outermost shell (1·45 to 1·40 Å spacings) 48% of these structure amplitudes had $|F|^2$ greater than $2\sigma(|F|^2)$. The mean deviation

Table 2. *Synopsis of refinement*

Phase	Number of heterogeneous amino acids	Total number of solvent sites	Number of overlapping solvent sites	Resolution (Å)	Number of reflections	Data limits	*R* factor	R.m.s. deviation from ideal bonds (Å)	Summary of action
IV	0	54	0	10–1·4	10 300	$F>0$	0·224	0·019	Starting model
X	0	73	0	10–2·0	3 590	$F>0$	0·181	0·016	Correct cell constants, experiment on building 45–49, adjust some side chains
XI	0	0	0	10–1·7	5 821	$F>0$	0·241	0·019	Remove all solvent, adjust some side chains
XII	0	41	0	10–1·4	10 300	$F>0$	0·209	0·017	Add solvent, adjust some side chains
XIII	1	59	4	10–1·4	10 300	$F>0$	0·199	0·017	Add solvent, begin modeling heterogeneity, adjust some side chains
XIV	3	83	9	10–1·4	10 300	$F>0$	0·197	0·018	Add solvent, model heterogeneity, adjust some side chains
XV	6	101	14	10–1·4	8 976	$F>14$	0·173	0·017	Add solvent, model heterogeneity, adjust some side chains
XIX	9	126	21	10–1·4	8 122	$F>18$	0·151	0·016	Experiments on building 45–49, 33, 62; add solvent; model heterogeneity
XXII	17	124	24	10–1·4	8 364	$F>1·5\sigma$	0·152	0·015	Sort out correlated heterogeneity, model heterogeneity in 44–50
XXIII	17	124	24	10–1·4	7 732	$F>4\sigma$	0·144	0·017	Substitute remerged and scaled data
XXVII	15	112	28	10–1·4	7 732	$F>4\sigma$	0·144	0·016	Remove heterogeneous groups and check for return of density, experiment on building 45–49

between the new $|F|^2$ values and those previously was 7%.

During the data reduction, the variance of every parameter involved was included in the final error estimates, and then the standard deviations of the final set were renormalized with a two-dimensional analysis of variance. These error estimates were used to select the 7732 reflections with $|F|^2$ greater than $2\sigma(|F|^2)$ in the resolution range 10·0 to 1·4 Å used in the refinement of the atomic model. The data processing and merging programs were local versions of those of Blessing (1987).

Starting model

The starting point for this work was the model described by Bourne *et al.* (1985). This model is the result of phases 0 through IV of refinement and is composed of the 475 atoms of erabutoxin b, 51 water molecules, one sulfate S atom and one sulfate O atom. All atoms have individual isotropic thermal parameters. The conventional *R* factor $(\sum||F_o|-|F_c||/\sum|F_o|)$ for the model is 22·4% for all data from 10·0 to 1·4 Å spacings and the root-mean-square (r.m.s.) deviation of covalent bond lengths from ideal values is 0·019 Å. In order to develop reasonable estimates for their occupancies and thermal parameters, the solvent molecules were removed from the model after phase X of refinement and new positions from $|F_o|-|F_c|$ maps were established. Thus the solvent structure determined here is not exactly that described in the earlier report, and the enumeration given to solvent molecules does not correspond to that in the preliminary report (Bourne *et al.*, 1985).

Refinement progress

A synopsis of the refinement process is contained in Table 2. Phases V through X involved primarily corrections to the starting protein model: several side chains were adjusted and various models were built for the peptide segment, residues 45 through 49. The electron density representing this peptide, while continuous along the backbone, was weaker and more poorly formed than that for other parts of the structure. In addition, discrete isolated peaks of electron density did not appear where solvent molecules were expected, as was the case for other areas of ordered solvent. Instead, rather elongated poorly formed peaks were found. Residues 45 through 48 were modeled in several conformations with the peptide going through the putative solvent density, but all of these models resulted in unacceptably large thermal parameters after refinement. We thus concluded that, while multiple conformers exist in the crystal for the 45–49 peptide, there is not a significant population of molecules whose backbone position differs radically from that of our original model. The elongated peaks in the solvent regions then represent overlapping solvent sites most probably correlated with the multiple protein conformers.

The cell constants were corrected in phase VI. The largest change was a 1·5% increase in the *c* cell edge. This had a significant detrimental effect on the *R* factor, which was not overcome for several cycles of refinement.

Phases XI through XV principally involved the identification of solvent molecules and alternative side-chain conformers.

In phase XI of refinement all solvent was removed from the model in order to develop a rational scheme for assigning and refining occupancy and thermal parameters. Solvent sites were added back to the model as described under *Methods* during phases XII through XV. No attempt was made at this point to keep at unity the sum of occupancies for heterogeneous solvent sites, although when an overlapping site was added to the model the occupancies of both sites were generally set to 0·5. Conformational heterogeneity in several side chains was observed and modeled during this segment of the refinement.

The weakest 21% of the data were eliminated from the refinement process by phase XVII. This had a significant effect on the R factor. Beginning with phase XVIII an overall anisotropic correction to the thermal parameters was included, using a method developed by Sheriff & Hendrickson (1987). An overall anisotropic thermal parameter was calculated outside the structure-factor least-squares refinement of the atomic model by minimizing the fit of calculated to observed structure amplitudes. The components of this overall thermal parameter were adjusted so that the net change to the individual isotropic thermal parameters was zero and the overall anisotropic correction was then included in structure-factor calculations during refinement of the atomic model. The magnitude of this correction for erabutoxin b data was about 2·4 Å^2 (see Table 3), increasing B values in the b and c directions and decreasing B values in the a direction. This correction was redetermined between phases of refinement. It did not have a large impact on the R factor but did significantly reduce the noise in $|F_o|-|F_c|$ electron density maps. The 45–49 pentapeptide segment, three side chains with weak electron density and several associated water sites were deleted from the model in phase XVI and gradually rebuilt into new electron density in phases XVII through XIX.

Correlations among heterogeneous protein groups and solvent atoms were established in phases XX to XXII.

In phase XX the results of interatomic distance calculations were studied to determine which discretely disordered atoms could not coexist with either partially occupied solvent sites or heterogeneous parts of the protein. If interatomic contacts were shorter than 3·0 Å, or 2·4 Å for pairs of atoms that could form hydrogen bonds, the pair of atoms was classified as mutually exclusive. These cutoffs were relaxed by 0·1 Å for contacts between atoms in different asymmetric units since there were no restraints applied during refinement to keep such atoms apart. Estimates of occupancies were made based on the results of the prior refinements, and occupancy and thermal parameters of affected atoms were adjusted as described under *Methods*. The occupancies of affected solvent sites were not refined further. At this point errors were estimated for the structure amplitudes and reflections eliminated whose magnitudes were less than 1·5 times their standard deviation (19% of the data). Minor adjustments to the solvent structure and modeling of heterogeneity in the 45–50 peptide segment took place in phases XXI and XXII. The entire peptide segment encompassing residues 44 through 51 was modeled in two conformations, even though there are only subtle differences between the primary and alternative backbone conformations of the terminating Pro44 and Lys51 residues. The side chain of Lys51 had previously been observed in two conformations.

The new processing, scaling and merging of data, including measurements from an additional crystal, were completed and these data were used for refinement beginning with phase XXIII. Only data with $|F_o|$ greater than four times its estimated standard deviation were included (74% of the total).

Minor adjustments were made to the model in phases XXIV through XXVII. In phases XXIV and XXV more rebuilding experiments were done on the 45–49 peptide. In five subphases of phase XXVI all heterogeneous groups in the protein, most overlapping solvent sites and solvent molecules with occupancy less than 0·40 or B value greater than 30 Å^2 were removed from the model. Several cycles of refinement were carried out and those groups for which strong electron density reappeared were returned to the model. Correlations among heterogeneous solvent sites were redetermined at this stage. Dual conformers for side chains were confirmed further by refining models with first one and then the other alternative for these side chains removed, and verifying that electron density appeared for each excluded conformer. Examples are shown in Fig. 1. Phase XXVII is the latest refinement with the full model in place.

Refined model

A stereo drawing of the erabutoxin b model is given in Fig. 2. Table 3 summarizes the stereochemical features of the model and the agreement with diffraction data. Coordinates and structure amplitudes have been deposited in the Brookhaven Protein Data Bank (Bernstein, Koetzle, Williams, Meyer, Brice, Rodgers, Kennard, Shimanouchi & Tasumi, 1977).* The final

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 3EBX, R3EBXSF), and are available in machine readable form from the Protein Data Bank at Brookhaven or one of the affiliated centres at Melbourne or Osaka. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP37023 (as microfiche). Free copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England. At the request of the authors, the list of structure factors will remain privileged until 1 January 1992.

model comprises the full erabutoxin b molecule with dual conformations for seven single residues and the peptide encompassing residues 44 through 51, 111 water sites together with one sulfate ion having a total site occupancy of 68. Twenty solvent sites overlap partially occupied sites of flexible protein groups. There are four isolated pairs of overlapping solvent sites and seven 'networks' of disordered solvent molecules. The largest of these involves seven sites in two groups - four molecules in one group and three in the other - and the smallest involves three sites.

Progress in the last stages of refinement was monitored by the appearance of $|F_o|-|F_c|$ difference Fourier maps. We ended the refinement when the strongest features in the difference maps could not be accounted for by rebuilding the model or adding more solvent or heterogeneity. The final $|F_o|-|F_c|$ difference map has nine peaks higher than 0·30 e Å^{-3} for which we have no satisfactory explanation. The largest peak (0·38 e Å^{-3}) is adjacent to the Ser8–Ser9 peptide bond, just above and inside the β turn formed by residues 7 through 10. The peak center is 1·45 Å from Ser9 N. Peaks 2 (0·33), 4 (0·31) and 5 (0·30 e Å^{-3}) are too near side chains to be solvent, yet are inaccessible as alternative conformers for the side chains. The third (0·32), seventh (0·30) and eighth (0·30 e Å^{-3}) peaks are all less than 2 Å from carbon atoms in the polypeptide backbone. The sixth peak (0·30 e Å^{-3}) is in the solvent continuum, 3·20 Å

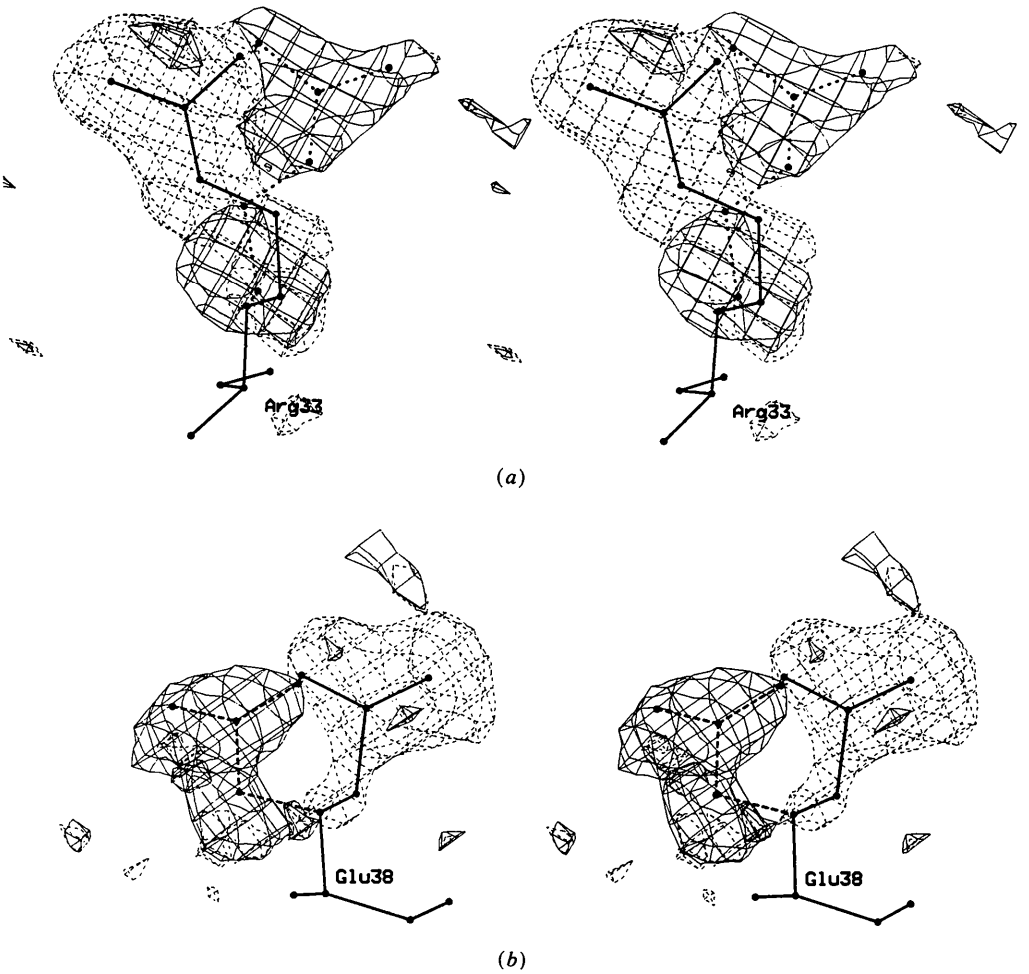

Fig. 1. Difference electron densities in two of the disordered side groups: (*a*) Arg33; (*b*) Glu38. Two sets of refinements were carried out, involving side chains for the groups Arg33, Glu38, Glu56 and Asn62. For each of these groups, there are two sets of positions for atoms beyond the β carbons. In the stereo diagrams, these alternative positions are shown with solid and broken bonds. In the first refinement, side chains shown in solid lines were assigned half weights, and the other side chains were left out. The $|F_o|-|F_c|$ synthesis calculated after several cycles of refinement is shown as solid contours. In the second refinement, the roles of the alternatives were reversed, and the resulting difference density is shown in broken contours. The diagrams clearly indicate bimodal positions for these side groups. Contour levels are at 0·21 e Å^{-3}; the estimated standard deviation for both maps is 0·1 Å^{-3}.

Table 3. *Statistics for the refined model*

	R.m.s. deviation from ideality	Target variance
Stereochemical features		
Covalent bonds (Å)	0·016	0·020
1-3 distances (Å) [bond angles (°)]	0·033, [1·5]	0·030
Planar 1-4 distances (Å)	0·049	0·060
Least-squares planes (Å)	0·016	0·020
Volume at chiral centers (Å^3)	0·175	0·150
Single-torsion contacts	190 at 0·174 Å	0·500
Multiple-torsion contacts	133 at 0·174 Å	0·500
Possible H-bond contacts	62 at 0·203 Å	0·500
Planar torsions (°)	2·9	3·0
Staggered torsions (°)	16·8	15·0
Orthonormal torsions (°)	22·8	20·0
ΔB for main-chain bonds (Å^2)	1·23	1·50
ΔB for main-chain angles (Å^2)	1·84	2·00
ΔB for side-chain bonds (Å^2)	2·49	2·00
ΔB for side-chain angles (Å^2)	3·67	3·00
Thermal parameters		
Mean B for protein atoms (Å^2)	14·19	
Mean B for solvent atoms (Å^2)	25·74‡	
Overall anisotropic ΔB(Å^2)		
a component	−1·42	
b component	0·39	
c component	1·03	
Agreement with diffraction data		
Standard R factor*	0·144	
Error R factor†	0·055	

Weights on structure amplitudes in last cycle $= 1/\sigma^2$

$\sigma = 12{\cdot}7 - 38{\cdot}2[(\sin\theta)/\lambda - 1/6]$

Average $||F_o| - |F_c|| = 12{\cdot}65$

* Standard R factor $= \sum ||F_o| - |F_c|| / \sum |F_o|$.
† Error R factor $= \sum \sigma(|F_o|) / \sum |F_o|$.
‡ 112 sites, total occupancy = 68.

from the side chain of Ile2, and the ninth peak (0·30 e Å^{-3}) is too near both conformers of Glu38 to coexist with either.

Discussion

Examples of two alternative conformations for parts of the protein molecule are numerous in the crystal structure of erabutoxin b. Many of these are instances of alternative staggered conformations for side chains (Smith, Hendrickson, Honzatko & Sheriff, 1986). The two models for the octapeptide segment of residues 44 through 51 differ primarily in their backbone conformations. This backbone segment, the most difficult to fit in the protein model, was built into strong and continuous electron density that clearly represented more than a single conformer. It is likely that there are more than two conformations for this segment given the number of conformational degrees of freedom and the lack of hydrogen bonds to other parts of the protein.

Discrete disorder is also abundant in the solvent regions of the crystals, where many pairs of incompatible sites were located. The overall pattern of relationships between solvent sites is shown graphically in Fig. 3. Extensive interlacing of incompatible solvent structures is not limited to solvent molecules associated with protein groups found in two alternative conformations; it occurs throughout the regions of ordered solvent in the crystal. These observations are not overinterpretations of random noise in difference electron density maps. The large number of instances of discrete disorder in the solvent were interpreted from significant difference density that in all cases appeared too near partially occupied atomic sites. Except for the few peaks discussed in the previous section, there were not significant difference peaks too near fully occupied atomic sites in the protein model. We also did not observe, with the single exception noted above, difference peaks too far from all atoms in the model to form any hydrogen bonds. In several cases, pairs of water sites refined to positions that were incompatible after initially being greater than 2·4 Å apart. In most cases these pairs of sites were in the same asymmetric unit and nonbonded-contact restraints in the refinement program were acting against their close approach. Our assignment

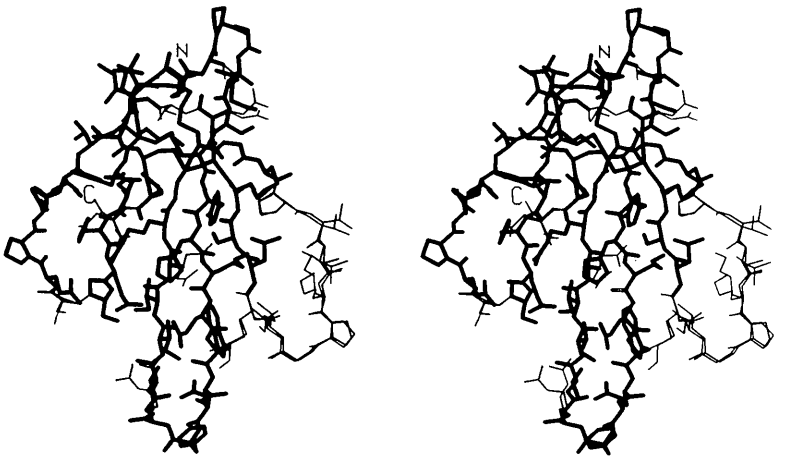

Fig. 2. Stereodrawing of the molecular structure of erabutoxin b as seen in the crystal. Thin lines are used to indicate positions of alternative conformers.

of them as mutually exclusive was thus not an artifact of the lack of restraints in the refinement program between atoms in different asymmetric units.

Discrete disorder was defined in the crystal structure by interatomic contacts that prohibited the affected atoms or chemical groups from occupying the same unit cell at the same time. The extensive discrete disorder in erabutoxin b crystals means that the crystal structure is actually the composite of the many unique structures that arise from combinations of local structures at all the sites of discrete disorder or heterogeneity in the crystals. This single crystallographic experiment cannot determine whether the observed multiple structures are due to static or dynamic disorder or even to differences among the crystals contributing to the data set. In any case, both complete description of the crystal structure and accurate refinement of the model demand that the sets of incompatible local structures be defined. These are summarized in Table 4. For each of the 16 sets of atoms or chemical groups that are related by prohibitive interatomic contacts, only two mutually exclusive groups need be invoked to account for all the close contacts.

Two aspects of the refinement procedure require knowledge of which local structures can coexist. Restraints against short nonbonded contacts are an important stereochemical feature of the refinement program, but should not be applied to pairs of atoms that are not present simultaneously in the same unit cell. In addition, fractional occupancies for atoms that occupy the same volume should not sum to greater than unity. The relationships among the sites of heterogeneity in erabutoxin b crystals were determined twice in the course of refinement and this information was used in subsequent refinement.

The limiting 1·4 Å spacing of the diffraction data limits the observable separation of dual conformers for protein groups and of overlapping solvent sites. In the erabutoxin b model heterogeneous protein groups are separated from the most distant atoms of their alternative conformers by 1·2 to 2·8 Å (excluding heterogeneous prolines). Comparable separations involving solvent atoms, with one exception, occur over a range of 1·2 to 3·0 Å. The exception involves H_2O64, which is only 0·65 Å from Glu38′ $O_{\varepsilon 1}$. This water molecule forms a hydrogen bond to Glu38 $O_{\varepsilon 2}$ and was placed in the model before the alternative conformer Glu38′ was located. Both Glu side chains and the water molecule have thermal parameters below the average for protein atoms, and the electron density for these atoms is strong. In general, however, discrete positions this close cannot be resolved in this analysis.

Values for occupancy (q) and thermal (B) parameters are highly correlated for those components of the model with variable occupancies: solvent sites and heterogeneous protein groups. The limits of the diffraction data obtainable from erabutoxin b crystals enhance this correlation and make difficult the simultaneous determination of q and B values. In this work q and B values for solvent atoms were shifted in

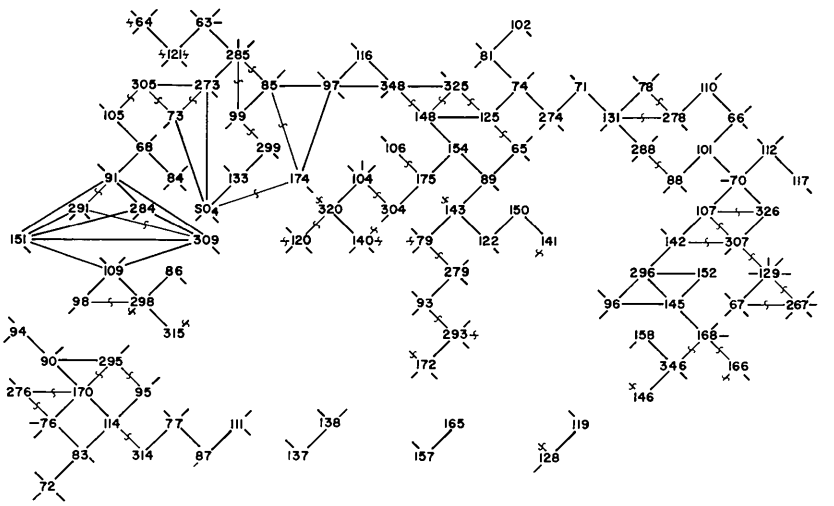

Fig. 3. Compatibility of solvent sites in the erabutoxin b crystals. Numbers refer to water sites. Sites that are separated by a reasonable hydrogen-bonding distance are linked by unbroken lines, except when they belong to the mutually incompatible groups listed in Table 4. Sites that are too close together to be occupied simultaneously are linked by lines with a tilde. Short non-terminating lines indicate hydrogen-bonding contacts to protein residues, while such lines with a tilde indicate contacts to a discretely disordered amino-acid residue that are too short for simultaneous occupancy. Some water molecules form multiple hydrogen bonds that cannot occur simultaneously. For example, H_2O97 can be hydrogen bonded to both H_2O85 and H_2O174. Since the latter two water molecules are mutually incompatible, H_2O97 can only be hydrogen bonded to one of them at a time. This network does not include nine water molecules that can form hydrogen bonds only to protein atoms and not to other water molecules.

Table 4. *Mutually exclusive groups of amino-acid residues and solvent molecules*

	Group 1	Group 2
1.	Ser9	Ser9′
	H_2O79	H_2O279
	H_2O143	
2.	Ser18	Ser18′
3.	Arg33	Arg33′
	Asn62	Asn62′
	H_2O166	H_2O69
	H_2O146	H_2O168
	H_2O346	H_2O172
	H_2O93	H_2O293
4.	Glu38	Glu38′
	Ser57	Ser57′
	H_2O64	
	H_2O121	
5.	Pro44	Pro44′
	Thr45	Thr45′
	Val46	Val46′
	Lys47	Lys47′
	Pro48	Pro48′
	Gly49	Gly49′
	Ile50	Ile50′
	Lys51	Lys51′
	Glu56	Glu56′
	H_2O113	H_2O120
	H_2O124	H_2O128
	H_2O140	H_2O298
	H_2O320	H_2O315
	H_2O141	
	H_2O98	
	H_2O176	
6.	H_2O67	H_2O267
	H_2O107	H_2O307
	H_2O129	H_2O326
	H_2O142	
7.	SO_475	H_2O174
	H_2O85	H_2O285
	H_2O99	H_2O299
8.	H_2O76	H_2O276
	H_2O95	H_2O295
	H_2O170	
9.	H_2O125	H_2O325
	H_2O148	H_2O348
		H_2O65
10.	H_2O73	H_2O273
	H_2O105	H_2O305
11.	H_2O78	H_2O278
	H_2O131	
12.	H_2O91	H_2O291
	H_2O309	
13.	H_2O88	H_2O288
14.	H_2O104	H_2O304
15.	H_2O106	H_2O175
16.	H_2O114	H_2O314

In this table, the pairs of groups are distinguished from one another by exclusion. For the sets 1 through 16, each residue or solvent molecule listed in group 1 is incompatible with at least one member of group 2 for the same set. Thus all members of each group for each set may be present simultaneously, although they are not necessarily all involved in interactions with one another. Potentially hydrogen-bonded pairs of atoms closer than 2·4 Å were classified as incompatible and a cutoff of 3·0 Å was used for other contacts. These cutoffs were relaxed by 0·1 Å for contacts between atoms in different asymmetric units, since there are no restraints applied during refinement to keep such atoms apart.

alternate refinement cycles, and the applied shifts were damped to about half of their calculated values. Occasionally q and B values were readjusted, as described under *Methods*, to what seemed reasonable values in response to a change in some feature of the model. The adjusted parameters tended to refine towards the values held before adjustment in nearly all cases. This indicates that the refined occupancy and thermal parameters have some validity.

Further evidence for the validity of refined occupancy and thermal parameters is provided by the results on poorly refined solvent sites. Two classes of apparently unreasonable combinations of q and B values for solvent molecules arose during refinement. These were low q/low B and high q/high B combinations. Solvent molecules that are only occasionally bound to the crystalline protein, and thus have low q values, would not be expected to have exceptionally low B values as well. Conversely, solvent molecules that have very high thermal parameters are not expected to be bound a high proportion of the time. A physical explanation for the high q/high B combination is the modeling of multiple unresolved sites with one solvent atom. The occupancy would tend towards the sum of individual site occupancies and the B value would become large enough to encompass all sites. No simple physical explanation is apparent for the low q/low B combination. When these two classes of 'poor' solvent molecules were removed from the model near the end of refinement, they included 34 water molecules with B values greater than 30 and 17 water molecules with q values less than 0·40. These two groups had no water molecules in common. Of the group with low q, only five waters (29%) reappeared with significant density in an $|F_o|-|F_c|$ difference map following a few cycles of refinement, while 29 of the waters with high B values (85%) returned. This lends further support to the hypothesis that refined occupancy and thermal parameters have meaning.

The assumptions made in the initial assignment of occupancy and thermal parameters for solvent molecules can be tested by comparing the final refined values with initial estimates. These data are shown graphically in Fig. 4. The initial occupancy parameters were estimated according to the interpolated height of the corresponding peaks in $|F_o|-|F_c|$ difference maps. While there is an appreciable scatter of points in Fig. 4(a), occupancy estimates track well with their final values and peak heights seem to be a reasonable indicator of solvent occupancy. The correlation coefficient between initial and final occupancy parameters for waters not discretely disordered is 0·64. On the other hand, the correlation coefficient is 0·41 between initial and final thermal parameters for the same waters (see Fig. 4b). The initial estimates of thermal parameters came from the number of hydrogen bonds each solvent could form with protein atoms. The inadequacy of the underlying assumption is shown in Fig. 5 where solvent B values are plotted as a function of number of hydrogen bonds in the final model. While there is a trend towards lower B values for solvents with more hydrogen bonds (especially when only hydrogen bonds to protein atoms are considered), there is such a wide range of B values for each number of hydrogen bonds that number of hydrogen bonds must be considered a very unreliable predictor of B value. Refined B values for water oxygens also are not predicted by the refined B values for the protein atoms to which they hydrogen bond. The correlation coefficient between these sets of B values is 0·12.

Several observations during the course of this refinement indicated to us that many of the solvent sites in the erabutoxin crystal actually represent multiple unresolved discrete sites. The high B/high q water molecules that were removed near the end of refinement and returned with very strong electron

density were those with the highest B values and with q values of 1·0. Two of these returned with double-lobed electron-density peaks and are now both modeled as two water sites. This behavior was also observed for the sulfate ion, which was initially modeled as a water. Its B value increased from 20 to 29 Å^2 and its q value from 0·5 to 1·0 in ten cycles of refinement (five sets of shifts each on B and on q) before it was changed to a sulfate ion. The strength of the difference density peaks in combination with the sulfate results suggest discrete unresolved sites for these water molecules over single highly mobile sites. The largest thermal and occupancy parameters for solvent molecules in the final model occur in water molecules associated with the heterogeneous peptide segment of residues 45 through 50. Unresolved discrete heterogeneity is not surprising in this region, since the protein conformers differ only slightly.

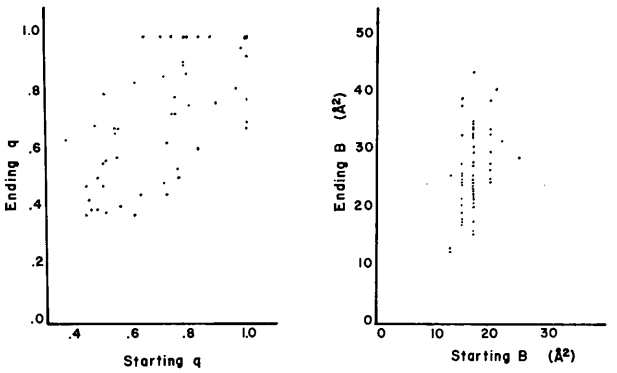

Fig. 4. (*a*) Plot of occupancy parameter (q) at the end of refinement *vs* its initial value for solvent molecules. Solvents for which occupancy was fixed due to correlated heterogeneity have not been included. (*b*) Plot of thermal parameter (B) at the end of refinement *vs* its initial value for solvent molecules.

As discussed above, the range of separations of solvent sites that are observed to be overlapping is limited by the diffraction data. The lack of stereochemical restraints on positional and thermal parameters in the refinement of water molecules makes this effect more pronounced in the solvent than in the protein, where restraints help to prevent discrete conformers from collapsing onto one another. At one point in the refinement an attempt was made to model as double sites some of the water molecules that had very large q and B values and were associated with heterogeneous protein groups. Hydrogen-bonding geometry was used to determine in what direction to separate each pair of sites. However, without extensive stereochemical restraints all double sites collapsed back onto each other when refined. Thus, although we have been able to model a considerable amount of discrete heterogeneity in the solvent regions of the erabutoxin crystal, the evidence suggests that much more heterogeneity is present. This effect undoubtedly accounts for some of the spread in B values and their lack of correlation to number of hydrogen bonds.

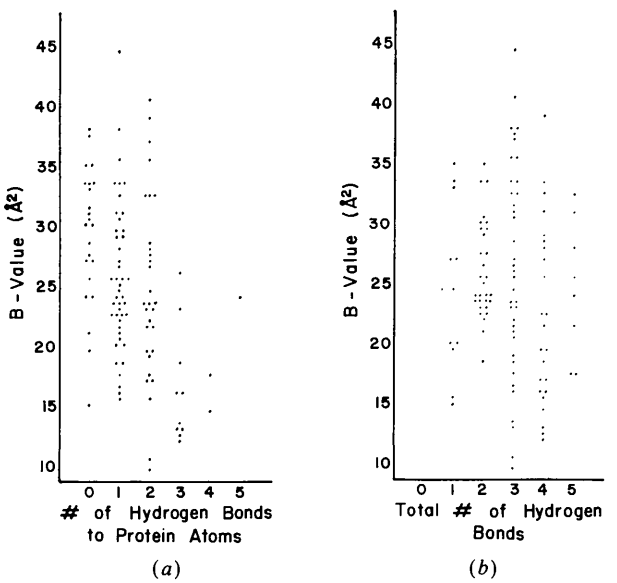

Fig. 5. (*a*) Plot of thermal parameter (B) for solvent molecules in the final model *vs* number of potential hydrogen bonds each could form to protein atoms. (*b*) Plot of B for solvent molecules *vs* total number of potential hydrogen bonds to both protein and other solvent atoms.

This work was supported by grants NS22719 and NS07747 (BWL) from the National Institute of Neurological and Communicative Diseases and Stroke, by federally-funded Biomedical Sciences Research Grants (BWL) from the College of Physicians and Surgeons of Columbia University, by grant DMB 84-09658 from the National Science Foundation (WAH), and by general research funds of the Office of Naval Research at the Naval Research Laboratory. We thank Dr Philip E. Bourne and Ms Linda Austrian who collected the data from the sixth crystal of erabutoxin b.

References

BARRY, C. D., MOLNAR, C. E. & ROSENBERGER, F. U. (1976). *Tech. Memo.* no. 229. St Louis: Computer Systems Lab., Washington Univ.

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.

BLESSING, R. B. (1987). *Crystallogr. Rev.* **1**, 3-58.

BOURNE, P. E., SATO, A., CORFIELD, P. W. R., ROSEN, L. S., BIRKEN, S. & LOW, B. W. (1985). *Eur. J. Biochem.* **153**, 521-527.

HENDRICKSON, W. A. (1981). In *Refinement of Protein Structures*, edited by P. A. MACHIN, J. W. CAMPBELL & M. ELDER, pp. 1-7. Warrington: SERC Daresbury Laboratory.

HENDRICKSON, W. A. & KONNERT, J. H. (1980). In *Computing in Crystallography*, edited by R. DIAMOND, S. RAMASESHAN & K. VENKATESAN, pp. 13.01-13.23. Bangalore: Indian Academy of Sciences.
JONES, T. A. (1982). In *Computational Crystallography*, edited by D. SAYRE, pp. 303-317. Oxford: Clarendon Press.
KIMBALL, M. R., SATO, A., RICHARDSON, J. S., ROSEN, L. S. & LOW, B. W. (1979). *Biochem. Biophys. Res. Commun.* **88**, 950-959.
KONNERT, J. H. (1976). *Acta Cryst.* A**32**, 614-617.
LOW, B. W. & CORFIELD, P. W. R. (1986). *Eur. J. Biochem.* **161**, 579-587.
LOW, B. W., PRESTON, H. S., SATO, A., ROSEN, L. S., SEARL, J. E., RUDKO, A. D. & RICHARDSON, J. S. (1976). *Proc. Natl Acad. Sci. USA*, **73**, 2991-2994.
NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351-359.
SHERIFF, S. & HENDRICKSON, W. A. (1987). *Acta Cryst.* A**43**, 118-121.
SMITH, J. L., HENDRICKSON, W. A., HONZATKO, R. B. & SHERIFF, S. (1986). *Biochemistry*, **25**, 5018-5027.
TSERNOGLOU, D. & PETSKO, G. A. (1976). *FEBS Lett.* **68**, 1-4.

Acta Cryst. (1988). A**44**, 368-373

A Reconciliation of Extinction Theories

BY T. M. SABINE

NSW Institute of Technology, Sydney, NSW 2007, *Australia*

(*Received* 20 *December* 1986; *accepted* 15 *January* 1988)

Abstract

The differences between previous theoretical treatments of extinction based on the Darwin intensity equations arise because of the different functional form chosen for the coupling constant σ. When the same function is used these theories make closely similar predictions. It is shown that a limiting condition on integrated intensity as the crystal size increases puts restrictions on the functions which may be used. A Lorentzian or Fresnellian function can be used for primary extinction while secondary extinction requires a Gaussian, rectangular or triangular function. An analytical expression is given for the variation in the value of the extinction factor with scattering angle.

1. Introduction

The kinematic theory of the diffraction of X-rays or neutrons predicts that the intensity of the diffracted beam is proportional to the volume of the crystal. If this were the case the intensity of the diffracted beam would exceed the intensity of the incident beam for sufficiently large crystals. The drawback of the kinematic theory is that it ignores the possibility of rescattering of the diffracted beam as it passes through the crystal. When the incident beam satisfies the Bragg condition in a perfect crystal so must the diffracted beam, and an interchange of energy between the diffracted beam and the incident beam will occur as both beams flow through the crystal. The rescattering probability increases as the size of the crystal increases.

For the imperfect crystal (Darwin, 1922), which is composed of blocks of perfect crystal (called mosaic blocks) tilted at small angles to each other, the beam diffracted by one block has a probability of being scattered again by a block of identical orientation during the passage of the diffracted beam through the crystal. The rescattering probability is proportional to the distribution of mosaic block orientations and the size of the crystal. For the perfect crystal spatial coherence between scattering centres is preserved across the entire specimen. In the imperfect crystal each block is a perfect crystal but there is no coherence between scattering centres located in different blocks.

In crystal structure analysis the phenomenon of increasing reduction of the intensity of the diffracted beam from the prediction of the kinematic theory as the crystal volume increases is termed extinction. The extinction factor, y, is defined by $I^{obs} = yI^{kin}$. I^{obs} is the integrated intensity measured in an experiment. I^{kin} is the integrated intensity a Bragg reflection would have if the kinematic theory applied exactly to the system being examined.

Extinction within a perfect crystal is termed primary extinction. In an ideally imperfect crystal, which is one in which extinction within mosaic blocks can be ignored, it is termed secondary extinction. Both types can occur in the same specimen.

The randomly oriented powder is a special case of the imperfect crystal. For this specimen the mosaic block distribution is known explicitly.

The object of the theories of extinction is to obtain an expression for y in terms of the dimensions and microstructure of the crystal so that I^{obs} can be correc-

0108-7673/88/030368-06$03.00