*Acta Cryst.* (1998). D**54**, 964–974

# Structure of Dimeric and Monomeric Erabutoxin a Refined at 1.5 Å Resolution

VASSILIOS NASTOPOULOS,[a]* PANAGIOTIS N. KANELLOPOULOS[a,b] AND DEMETRIUS TSERNOGLOU[c]

[a]*Department of Chemistry, University of Patras, 265 00 Patras, Greece,* [b]*European Molecular Biology Laboratory, 69012 Heidelberg, Germany, and* [c]*Dipartimento di Scienze Biochimiche, Università di Roma 'La Sapienza', P. le A. Moro 5, 00185 Roma, Italy. E-mail: nastopoulos@upatras.gr*

(*Received 6 October 1997; accepted 3 April 1998*)

## Abstract

Erabutoxin a has been crystallized in its monomeric and dimeric forms. The structures were refined at 1.50 and 1.49 Å resolution, respectively, using synchrotron radiation data. The crystals belong to space group $P2_12_12_1$, with cell dimensions $a = 49.84$, $b = 46.62$, $c = 21.22$ Å for the monomer and $a = 55.32$, $b = 53.54$, $c = 40.76$ Å for the dimer. Using starting models from earlier structure determinations, the monomeric structure refined to an $R$ value of 16.7% (8004 unique reflections, 17.0–1.50 Å resolution range), while the dimeric structure has been solved by the molecular-replacement method with a final $R$ value of 16.9% (19444 unique reflections, 17.4–1.49 Å resolution range). The high-resolution electron-density maps clearly revealed significant discrete disorder in the proteins and allowed an accurate determination of the solvent structure. For the monomer, the side chains of six residues were modelled with alternate conformers and 106 sites for water molecules and one site for a sulfate ion were included in the final model, whereas for the dimer, 206 sites for water molecules were included and both C-terminal residues together with the side chains of 11 residues adopted alternative conformations. A comparison was made with earlier structure determinations. The features of the solvent structure of the erabutoxin molecules are discussed in detail.

## 1. Introduction

The postsynaptic neurotoxins from the venom of sea and land snakes (*Hydrophiidae* and *Elapidae* families, respectively) are small proteins which exert their lethal action by binding to key elements (usually a receptor, ion channel or enzyme). Snake-venom toxins are currently classified into two distinct classes: the short-chain toxins (60–62 residues, four disulfide bridges) and the long-chain toxins (66–74 residues, usually five disulfide bridges). Both families show large sequence similarities, a common mode of action and have similar folding patterns (Ménez, 1987; Endo & Tamiya, 1987, 1991). They bind with great specificity and high affinity to the nicotinic acetylcholine receptor (AChR) at the postsynaptic membranes and block neuromuscular transmission. Thus, they provoke paralysis of skeletal muscles and induce death as a result of respiratory failure (Chang, 1979; Changeux, 1990; Endo & Tamiya, 1991). These curaremimetic toxins have long been exploited in extensive chemical and biochemical research because of their specific and tight binding to the AChR (dissociation constants, $K_d$, are of the order of $10^{-11}$) and have also been used as probes of receptor structure and function (Weber & Changeux, 1974; Tsetlin *et al.*, 1983; Pillet *et al.*, 1993).

The principal neurotoxic components in the venom of the Japanese sea snake *Laticauda semifasciata* are two small basic proteins, named erabutoxins a and b (Ea and Eb) from the Japanese name of the snake, erabu-umihebi (Tamiya & Arai, 1966). A minor neurotoxic component c (Ec) was later isolated by Tamiya & Abe (1972). Similar work by Tu *et al.* (1971) on the venom from the Philippines *L. semifasciata* indicated the presence of two neurotoxins, a and b. Subsequent crystallographic analyses established the identity of neurotoxins and erabutoxins from the venom of *L. semifasciata* from different regions of the Pacific Ocean. All three erabutoxins are short-chain proteins and comprise a single chain of 62 amino-acid residues. Ea and Ec differ from Eb by a single amino acid at positions 26 and 51, respectively (Ea: Asn26, Lys51; Eb: His26, Lys51; Ec: His26, Asn51). Most of these toxins have been described by X-ray methods using crystals grown under various experimental conditions. Neurotoxin a from the venom of the Philippines sea snake *L. semifasciata* was initially determined at 2.5 Å resolution by Tsernoglou & Petsko (1977) and Ea was determined at 2.0 Å by Corfield *et al.* (1989) (data deposited with the Brookhaven Protein Data Bank, code 5EBX), while the structure of recombinant Ea at 2.0 Å was determined by Arnoux *et al.* (1994). The crystal structure of neurotoxin b was reported at 2.2 Å by Tsernoglou & Petsko (1976) and the structure of Eb has been reported at 2.75 Å (Low *et al.*, 1976; Kimball *et al.*, 1979), followed by refinement at the high resolution of 1.4 Å (Bourne *et al.*, 1985; Low & Corfield, 1986; Smith *et al.*, 1988, PDB code 3EBX). The structure of a dimeric form of Eb, using thiocyanate solution as crystallizing agent, was determined by Saludjian *et al.* (1992) at 1.7 Å (PDB code 6EBX). The polypeptide chain in erabutoxins is folded

Table 1. *Summary of data collection and statistics*

| | Monomer | | Dimer | |
|---|---|---|---|---|
| Resolution range (Å) | 17–1.5 | 1.62–1.5 | 17.4–1.49 | 1.60–1.49 |
| Number of observations | 40456 | 6394 | 103554 | 12335 |
| Unique reflections | 8004 | 1510 | 19444 | 3462 |
| Completeness (%) | 95.1 | 92.2 | 94.8 | 86.0 |
| Mean $I/\sigma(I)$ | 30.9 | 10.2 | 29.7 | 5.25 |
| $I > 3\sigma(I)$ (%) | 89.8 | 77.4 | 81.1 | 56.3 |
| $R_{sym}$† (%) | 4.2 | 11.7 | 6.5 | 20.4 |

† $R_{sym} = \sum_{hkl} \sum_i |F_i^2(hkl) - \langle F^2(hkl)\rangle| / \sum_{hkl} \sum_i F_i^2(hkl)$.

in three adjacent loops rich in $\beta$-pleated sheets. Genetic engineering studies on Ea by site-directed mutagenesis suggest that the 'functional' site of Ea covers a contiguous surface encompassing the three loops and includes both conserved and variant residues (Trémeau *et al.*, 1995).

In this paper we present the refined structures of the monomeric and dimeric Ea from *L. semifasciata* at 1.50 and 1.49 Å resolution, respectively, using synchrotron data. This resolution is higher than any published to date on Ea, whose crystals generally diffract more weakly than those of Eb. Although the crystals used in this work and those reported previously (5EBX) are probably of the same diffracting quality, the better data presented here are presumably due to the larger size of crystals and the stronger X-ray beam. Contrary to the report of Saludjian *et al.* (1992), we found no need for the presence of thiocyanate in the crystallization of the dimer. As both cases reported here were similarly treated with respect to protein crystallization, data collection and processing, it is easier to assess the significance of structural differences observed, including packing effects on structural details and hydration. These results could be useful to current efforts (Pillet *et al.*, 1993; Trémeau *et al.*, 1995) to use protein engineering to probe toxin structure–function relationships.

## 2. Experimental

### 2.1. *Purification and crystallization*

*L. semifasciata* sea-snake venom was commercially obtained from Sigma Chemical Company (USA) and highly purified Ea protein was prepared in a similar way to that described previously (Tamiya & Arai, 1966; Tu *et al.*, 1971). Briefly, Ea was isolated and purified by means of Sephadex G-50 superfine gel filtration and CM-sepharose fast-flow column chromatography, after which it was desalted and concentrated to 6 mg $ml^{-1}$. Crystallization experiments were performed using the hanging-drop vapour-diffusion technique following the conditions described by Tsernoglou & Petsko (1977). More specifically, crystals of Ea monomer were grown from 40% saturated ammonium sulfate adjusted to pH 7.5 with 0.05 *M* potassium phosphate buffer. The volume of the drops was 12 µl and that of the reservoir was 1.0 ml. Prism-like crystals formed in a week at room temperature and continued to grow for about a month with maximum dimensions of 0.2 × 0.4 × 1.0 mm. The isomeric form of the protein was checked by electrospray mass spectrometry (ESMS) by means of a sample of dissolved crystals. Unexpectedly, at almost the same conditions (40% saturated ammonium sulfate, pH 7.2), crystals of a dimeric Ea form also appeared, and achieved maximum dimensions of 0.2 × 0.3 × 0.8 mm after three months. The volume of the drops was 10 µl and that of the reservoir 0.8 ml. It is noteworthy that a dimeric form of erabutoxin b has also been crystallized from a thiocyanate solution (Saludjian *et al.*, 1992) in the course of a study of the influence of salts on protein solubility and crystal growth (Riès-Kautt & Ducruix, 1989, 1991; Ménez & Ducruix, 1990).

### 2.2. *Data collection and processing*

X-ray intensities for two single crystals at 295 K were collected to 1.50 and 1.49 Å resolution, respectively, on the EMBL X11 synchrotron-radiation source beamline at the DORIS storage ring, DESY, Hamburg, using a 18 cm MAR image-plate detector at a wavelength of 0.92 Å. Data collection was performed by the rotation method using 2° rotation frames. The X-ray data were processed using the program *DENZO* (Otwinowski & Minor, 1997) and scaled with *SCALEPACK*. The crystals were found to be orthorhombic, belonging to the space group $P2_12_12_1$, with unit-cell dimensions $a$ = 49.84, $b$ = 46.62 and $c$ = 21.22 Å for one crystal and $a$ = 55.32, $b$ = 53.54 and $c$ = 40.76 Å for the second. The asymmetric unit of the first crystal form contained one molecule, while that of the second contained two molecules. The data-collection statistics are given in Table 1.

### 2.3. *Structure determination – molecular replacement*

For the monomeric structure the 2.0 Å structure of erabutoxin a (5EBX) was used as a starting model. This model was used directly in refinement, as described in the following section, after omitting all water molecules and the sulfate group included in the entry. For the 1.49 Å dimeric structure, the orientation of the dimer in the new unit cell was determined by molecular replacement using the program *AMoRe* (Navaza, 1994). The 1.7 Å dimeric form of its isomer erabutoxin b (6EBX) was used as a search model, omitting all water molecules. Applying the fast-rotation function, using reflections in the resolution range 15–3.0 Å and an integration radius of 18 Å, one major peak appeared at 10.7$\sigma$. This solution was used to perform the translation-function calculations, using reflections from 15 to 2.0 Å. A single solution was obtained, and was refined by the rigid-body refinement procedure, implemented in *AMoRe*, to an initial crystallographic $R$ value of 0.41 in the resolution range 15–1.49 Å.

### 2.4. *Refinement*

Both models were initially refined using the *X-PLOR* program package (Brünger, 1992*a*). The starting model was subjected to two cycles of rigid-body refinement, and during the second cycle the number of rigid groups was increased from the number of protein molecules to the number of elements of secondary structure. At this stage, $2F_o - F_c$ and $F_o - F_c$ Fourier syntheses were calculated, and for the dimer the critical residue at position 26 distinguishing between the two isoforms a and b was changed to an asparagine. This was very clearly observed in the electron-density maps. The model building was performed using the program *O* (Jones *et al.*, 1991). The model was then subjected to refinement using simulated-annealing methods. In detail, each refinement step included simulated-annealing refinement with a starting temperature of 3000 K, 20 cycles of restrained *B*-factor refinement, and finally, 160 cycles of conjugate-gradient minimization. For the dimeric form, the two non-crystallographically related molecules *A* and *B* in the asymmetric unit were treated independently during all stages of refinement, as refinement with non-crystallographic symmetry (NCS) restraints resulted in poorer $R$ and $R_{free}$ values. All data in the resolution range 6.0 Å to the maximum resolution were used and no sigma cut-off was applied. The parameter set used was that of Engh & Huber (1991). The weight *wa* in *X-PLOR* was set initially by use of the 'check' procedure and subsequently improved by monitoring $R_{free}$. Water molecules were gradually incorporated at peaks of the $F_o - F_c$ electron-density maps greater than three times the root-mean-square (r.m.s.) electron density. For the monomeric structure a sulfate ion was also added. In the next steps, additional cycles of model building and refinement were performed. To cross-validate in reciprocal space, the free $R$ value was used (Brünger, 1992*b*). During the last round with *X-PLOR*, in addition to the other steps, 15 cycles of occupancy refinement for the water molecules and the sulfate ion were included in the case of the monomer structure, followed by 20 cycles of restrained *B*-factor refinement. The resulting model with *X-PLOR* had a crystallographic $R$ value of 0.206 and an $R_{free}$ value of 0.261 (10% of the data) for the resolution range 6.0–1.5 Å for the monomer structure, while in the case of the dimer the crystallographic $R$ value was 0.198 and the $R_{free}$ value 0.254 (10% of the data) for the resolution range 6.0–1.49 Å. The models were finally refined for a further two cycles using *TNT* (Tronrud *et al.*, 1987) by the conjugate-gradient direction method. Each round of refinement with *TNT* consisted of 20 cycles with constant occupancies, followed by 20 cycles including optimization of the occupancies of all heteroatoms (water molecules and sulfate ion). The standard PROTGEO.DAT parameter set was used, and the WEIGHT RFACTOR parameter was varied to optimize the results. During the last map inspection, alternative conformations were identified and they were included during the last round of refinement with *TNT*. For the monomeric structure, the sites of 106 water molecules and one sulfate ion have been included in the final model, while the side chains of six residues were found to adopt alternative conformations. The crystallographic $R$ value for all data (8004 reflections) in the resolution range 17–1.5 Å was 0.167. In the case of the dimer, the sites of 206 water molecules were included in the final model and both C-terminal residues together with the side chains of 11 residues adopted alternative conformations. The crystallographic $R$ value for all 19444 reflections in the resolution range 17.4–1.49 Å was 0.169.†

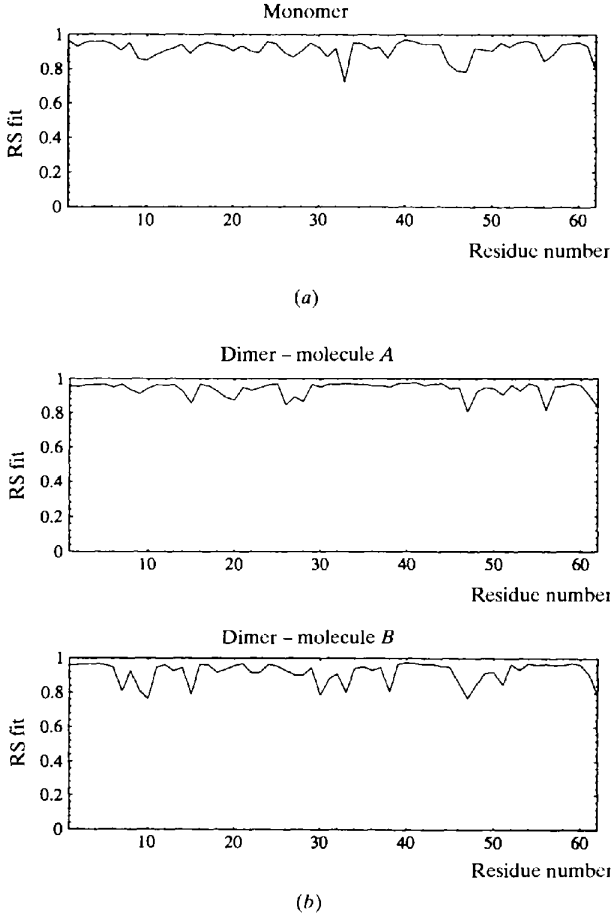



Fig. 1. Real-space correlation coefficients as a function of residue number for (*a*) the monomer and (*b*) the two independent molecules *A* and *B* of the dimer. The program used for the calculations was *O* (Jones *et al.*, 1991).

† Atomic coordinates and structure factors for the monomer and dimer have been deposited with the Protein Data Bank, Brookhaven National Laboratory (References: 1QKE and 1QKD, respectively).

Table 2. *Refinement statistics*

| | | Monomer | Dimer |
|---|---|---|---|
| Number of | Residues | 62 | 124 |
| | Protein atoms | 489 | 987 |
| | Solvent molecules | 106 | 206 |
| | Sulfate ions | 1 | 0 |
| Average $B$ factors (Å$^2$) | Overall structure | 24.2 | 25.8 |
| | Protein | 21.4 | 23.2 |
| | Solvent | 35.9 | 37.8 |
| R.m.s. deviations † | Bonds (Å) | 0.009 | 0.010 |
| | Angles (°) | 2.07 | 2.17 |

† R.m.s. deviations from ideal geometry.

## 3. Results and discussion

### 3.1. *Quality of the structures*

The refined models exhibit good fit to the $2F_o - F_c$ electron-density maps calculated with reflections in the resolution range 15–1.5 Å for the monomer and 15–1.49 Å for the dimer. The real-space correlation coefficients are plotted against residue number in Fig. 1 and are, in general, very good. The coefficients have been calculated for complete residues as implemented in *O* (Jones *et al.*, 1991). The pattern is similar for the two independently refined molecules *A* and *B* of the dimer. Final refinement statistics for *B* factors and r.m.s. deviations from geometric ideality are summarized in Table 2. It is obvious that both structures possess good stereochemistry. The variation of the *B* parameters along the peptide chain is shown in Fig. 2. The maximum coordinate error was determined from a Luzzati plot (Luzzati, 1952) to be about 0.15 Å for the monomer (Fig. 3) and between 0.15 and 0.20 Å for the dimer. The Ramachandran plot of the $\varphi$ and $\psi$ angles (Ramakrishnan & Ramachandran, 1965) of the monomer shows that only Ser8 lies in the generously allowed regions as calculated with the program *PROCHECK* (Laskowski *et al.*, 1993). For the dimer, out of 114 non-glycine residues, only residues SerA8 and SerB8 lie outside the stereochemically allowed regions. This was also observed in the dimeric Eb (6EBX) and is probably attributable to the fact that these residues are located at the position $i + 2$ of a tight turn and have high *B* factors.

### 3.2. *Description of the monomer*

The overall structure of the 62 amino-acid single chain established here for the monomeric Ea is very similar to those found previously for Ea (5EBX) and Eb (3EBX). The protein has a leaf-shaped structure of about 38 × 28 × 15 Å and from its central core stabilized by four disulfide bridges (3–24, 17–41, 43–54, 55–60) there emerge three long adjacent loops. They consist of a large and flat $\beta$-pleated sheet of five antiparallel $\beta$-strands connected by two hairpin motifs (left to right in Fig. 4*a*) and one additional strand segment (extreme right in Fig. 4*a*). In total, there are five turns where the polypeptide chain reverses direction. The protein displays the characteristic 'three-finger' structure

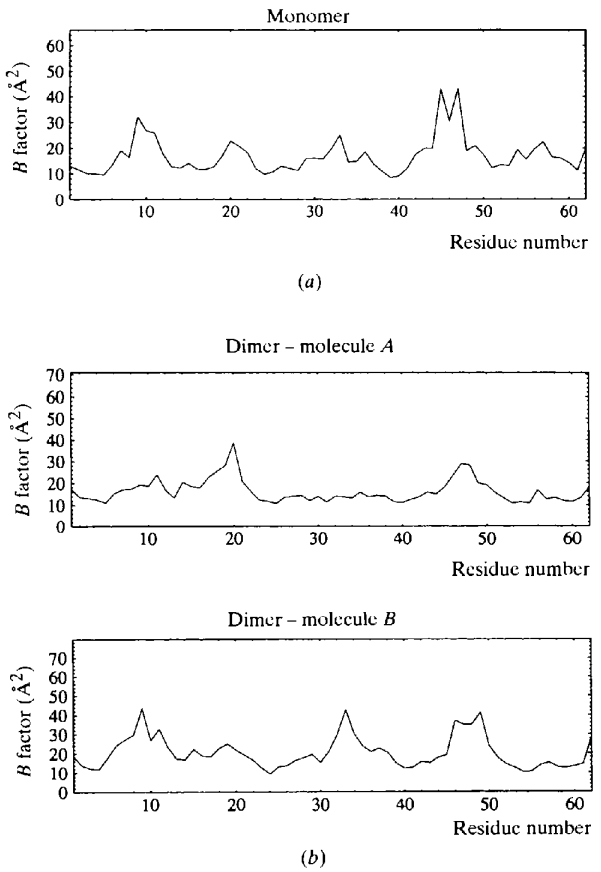



Fig. 2. Average main-chain *B* factors (Å$^2$) for (*a*) the monomer and (*b*) molecules *A* and *B* of the dimer, plotted against the residue number.

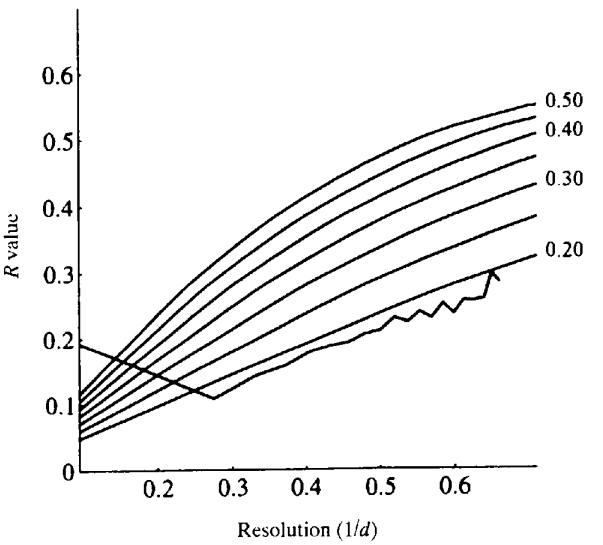



Fig. 3. Luzzati plot of *R* value *versus* resolution (Å$^{-1}$) for the 1.5 Å refined model of the Ea monomer.

(emerging like the fingers of a hand from a palm) commonly encountered in structurally related toxins (Drenth *et al.*, 1980). The extensive $\beta$-sheet hydrogen-bond network of the monomer is similar to that observed in the 3EBX and 5EBX models. The presence of five strong hydrogen bonds, four between residues located close to the N- and C-terminus of the protein (Ile2 N···Glu58 OE1 = 2.74 Å, Ile2 O···Val59 N = 2.97 Å, Phe4 N···Asn61 OD1 = 2.90 Å and Phe4 O···Asn61 ND2 = 2.99 Å) and another between Arg39 and Asn62 (Arg39 NH1···Asn62 OXT = 2.82 Å), contribute to the stabilization of the upper-left part of the structure [note the low $B$ factor (12.1 Å$^2$) of Asn61].

An overall similarity is observed comparing the distribution of $B$ parameters along the polypeptide chain with those of 3EBX and 5EBX. This is particularly evident for the peripheral-strand region Pro44→Lys51 which, by exhibiting relatively large $B$ factors and fewer

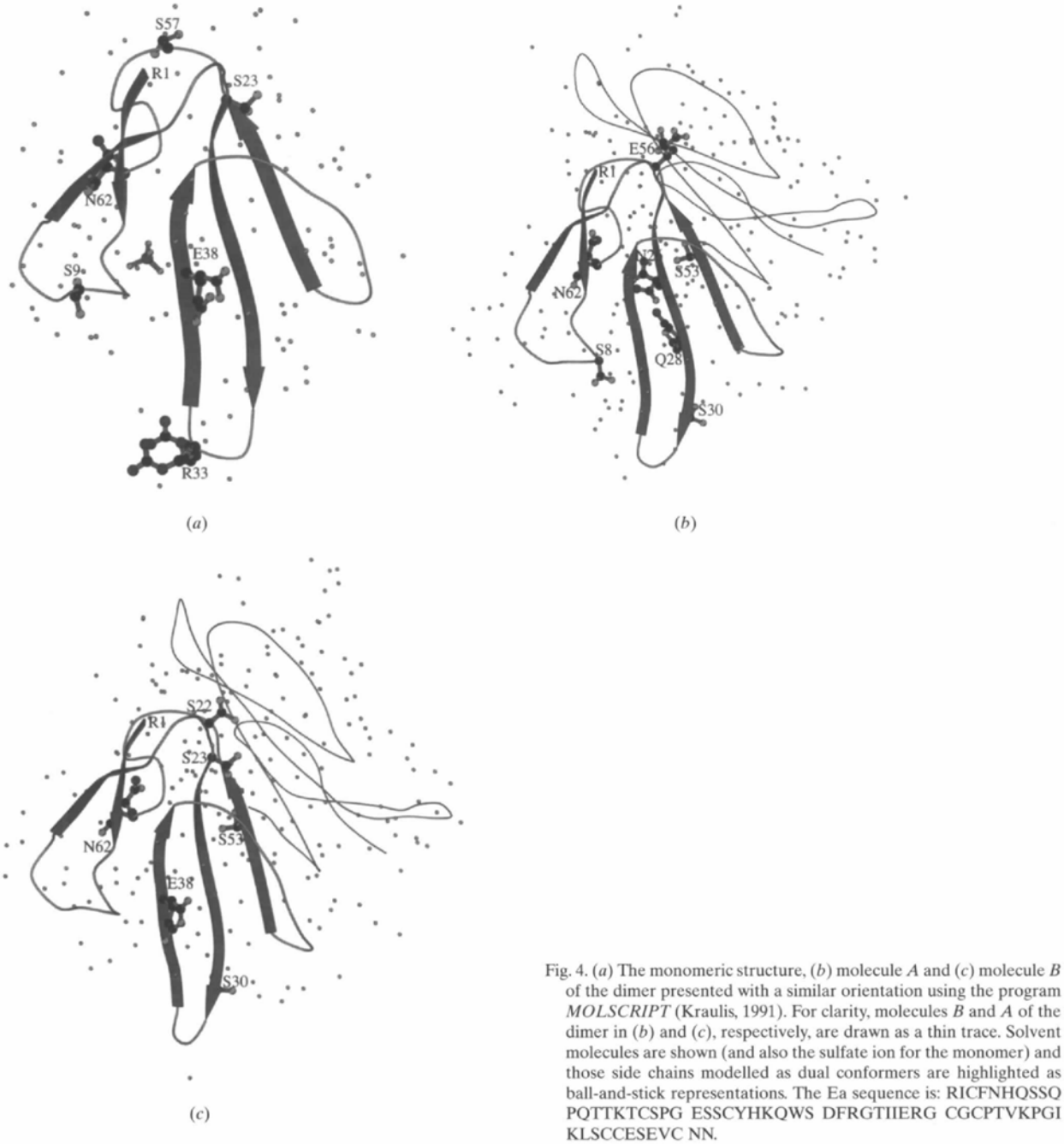



Fig. 4. (*a*) The monomeric structure, (*b*) molecule *A* and (*c*) molecule *B* of the dimer presented with a similar orientation using the program *MOLSCRIPT* (Kraulis, 1991). For clarity, molecules *B* and *A* of the dimer in (*b*) and (*c*), respectively, are drawn as a thin trace. Solvent molecules are shown (and also the sulfate ion for the monomer) and those side chains modelled as dual conformers are highlighted as ball-and-stick representations. The Ea sequence is: RICFNHQSSQ PQTTKTCSPG ESSCYHKQWS DFRGTIIERG CGCPTVKPGI KLSCCESEVC NN.

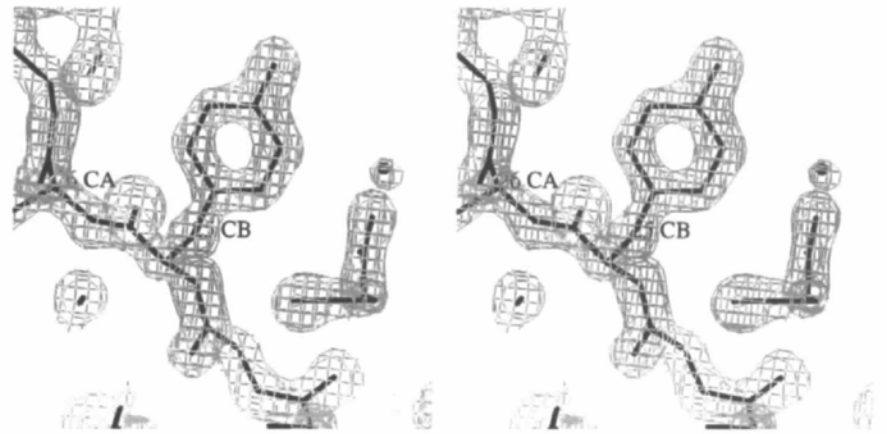
Table 3. *R.m.s. deviations (Å) between all Cα atoms for the superimposed backbones of monomeric and dimeric erabutoxins*

| | Ea† | Ea dimer‡ | Ea dimer molecule *B*‡ |
|---|---|---|---|
| Ea† | — | — | 0.768 |
| Ea dimer molecule *A*‡ | 0.865 | — | 0.680 |
| Ea dimer molecule *B*‡ | 0.768 | — | — |
| Eb (3EBX)§ | 0.137¶, 0.159†† | — | — |
| Ea (5EBX)‡‡ | 0.223 | — | — |
| Eb dimer (6EBX)§§ | — | 0.431 | — |

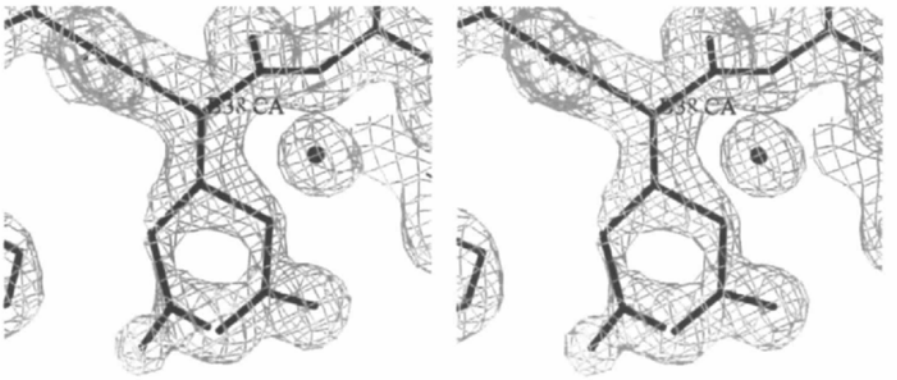
† This study for monomer at 1.50 Å ‡ This study for dimer at 1.49 Å. § Eb at 1.40 Å (PDB code 3EBX; Smith *et al.*, 1988) ¶ Eb at 1.40 Å with residues modelled with the first of the two possible bimodal positions †† Eb at 1.40 Å with residues modelled with the second of the two possible bimodal positions ‡‡ Ea at 2.0 Å (PDB code 5EBX; Corfield *et al.*, 1989). §§ Eb dimer at 1.7 Å (PDB code 6EBX; Saludjian *et al.*, 1992).

stabilizing interactions, makes up a region of high mobility and flexibility. Superposition of the backbones of Ea with 3EBX and 5EBX reveals minimal difference between them. The orientations of the side chains are similar, with certain exceptions: the Arg33 side chain differs from either bimodal position in 3EBX and mostly from 5EBX, while Gln10, Ile50, Leu52 and both conformations of Asn62 (see below) agree with 3EBX but have significantly different orientations to 5EBX. These findings are in agreement with an overall analysis of rotamers of side chains of all deposited erabutoxin structures described in §3.3. For comparison, the average r.m.s. deviations between all Cα atoms of the superimposed backbones of the known refined erabutoxin proteins are given in Table 3.

3.2.1. *Structural heterogeneity*. With the exception of the C-terminal residue Asn62, where the density is diffuse for both monomer and dimer, the excellent quality of the maps (Fig. 5*a*) made possible a clear identification of electron-density branching for some side chains which were modelled as dual conformers. Thus, discrete disorder was observed for the side chains of six residues: Ser9, Ser23, Arg33, Glu38, Ser57 and Asn62 (Fig. 4*a*). Disorder is a common feature in the crystal structures of erabutoxins: in the 2.0 Å Ea model Ser23 and Thr45 were bimodal. Furthermore, the Eb high-resolution structure at 1.4 Å comprises dual conformations for seven single residues (of which five are the same as in Ea) and for the peptide 44→51; many of these disordered residues are considered instances of




(*a*)




(*b*)

Fig. 5. The electron-density map in two regions of the final $2F_o - F_c$ synthesis, calculated using data from 15 to 1.5 Å resolution. (*a*) Residue Tyr25 of the monomer. (*b*) Glu*B*38 side chain of the dimer modelled as double conformer. The first and second maps are contoured at 2.0 and 1.5 times the r.m.s. electron density, respectively. The program used was *O* (Jones *et al.*, 1991).

Table 4. *Hydrogen bonds (<3.30 Å) between side chains with dual conformation and surrounding water molecules for the monomer*

| Side-chain atom | Water molecule(s) | Distance(s) (Å) |
|---|---|---|
| Ser9 OG | — | — |
| Ser9 OG† | W13, W101 | 2.97, 2.78 |
| Ser23 OG | W12, W28 | 2.74, 3.05 |
| Ser23 OG† | W12 | 2.84 |
| Arg33 NE | W87 | 2.45 |
| Arg33 NE† | W68 | 2.53 |
| Arg33 NH1 | W42 | 2.69 |
| Arg33 NH1† | — | — |
| Arg33 NH2 | W42 | 2.32 |
| Arg33 NH2† | W68 | 3.06 |
| Glu38 OE1 | — | — |
| Glu38 OE1† | W60 | 2.58 |
| Glu38 OE2 | W93 | 2.59 |
| Glu38 OE2† | W93 | 3.26 |
| Ser57 OG | W102 | 2.61 |
| Ser57 OG† | W78 | 2.54 |
| Asn62 OD1 | W21 | 2.86 |
| Asn62 OD1† | — | — |
| Asn62 ND2 | W80 | 2.79 |
| Asn62 ND2† | W35 | 2.29 |

† Alternate conformer of disordered side chains.

Table 5. *Contacts in the crystal packing of the monomer (<3.30 Å)*

Intermolecular contacts with symmetry-related molecules

| Residue | Residue | Distance (Å) |
|---|---|---|
| Gln7 NE2 | Cys60$^{i}$ O | 3.15 |
| Gln12 O | Phe32$^{ii}$ N | 2.93 |
| Gln12 OE1 | Leu52$^{iii}$ N | 2.57 |
| Cys17 O | Thr35$^{iv}$ N | 3.00 |
| Asp31 OD1 | Ser53$^{i}$ OG | 2.34 |
| Arg33 NH1 | Asn62$^{i}$ O | 2.29 |
| Glu38 OE1 | Glu56$^{i}$ N | 2.57 |
| Cys43 O | Lys47$^{v}$ NZ | 3.07 |

The bonding network of the sulfate ion

| Sulfate ion | | Atom | Distance (Å) |
|---|---|---|---|
| | O1 | Asn5 ND2 | 3.06 |
| | | W97 (···Arg39 O) | 2.71 (···3.14) |
| | | W102 (···Asn5 O) | 3.27 (···2.35) |
| SO4 | O2 | W17 | 3.23 |
| | O3 | Arg1$^{i}$ NH2 | 2.71 |
| | | W45 (···Glu56$^{i}$ O) | 2.49 (···3.30) |
| | | W17 | 3.21 |
| | O4 | W9 (···Glu58$^{i}$ N) | 2.66 (···3.01) |

Symmetry operators: (i) $x, y, z + 1$; (ii) $-x + 1, y - 1/2, -z + 5/2$; (iii) $-x + 1, y - 1/2, -z + 3/2$; (iv) $x - 1/2, -y + 1/2, -z + 2$; (v) $-x + 1/2, -y + 1, z - 1/2$.

alternative staggered conformations of side chains. In most of these residues, each of the alternate side-chain conformers forms hydrogen bonds to surrounding solvent molecules (Table 4). These bonds presumably induce stabilization of the side chains in discrete conformations, suppressing random orientations. Moreover, these findings are in accordance with Smith *et al.* (1986) who, having studied a sample of four proteins including Eb, suggest that residues for which multiple conformations are observed are likely to be either polar or charged and are exposed to solvent.

3.2.2. *Solvent structure.* The final model includes 106 discrete water sites (numbered from 2 to 107) and one (0.75 occupancy) sulfate ion, with a total of 84 sites with full occupancy (Fig. 4*a*). The solvent content of the crystals is 32% (Matthews, 1968) and the expected solvent content based on the density is 131 water molecules. The 3EBX and 5EBX models contain a total of 68 (spread over 111 sites) and 65 water molecules, respectively. As mentioned, the average $B$ factor of the water molecules is 35.9 Å$^2$ and the shortest water-to-water distance is 2.27 Å. These short interactions are considered to be acceptable because these waters are either found in discrete unconnected strong peaks of electron density or the sum of refined occupancy factors is notably above one; therefore, these molecules do not represent solvent alternates. In addition, these water molecules are neither close to side chains with alternates nor far from the protein. The same applies for the dimeric structure (see §3.3). Adopting 3.3 Å as the upper limit for acceptable hydrogen bonds, 86 water molecules are directly bonded to the protein molecule, constituting the hydration shell of the protein. Specifically, 46 are bonded to one protein molecule only, 16 are bonded to two atoms of the same protein molecule, 15 are involved in bridging two different protein molecules (including two bridges to sulfate) and nine form three hydrogen bonds to the same or two different protein molecules. As in the case of other proteins (Baker & Hubbard, 1984), the protein-bound waters show a preference for interaction with protein O atoms (72%) rather than N atoms (28%). This preferential interaction is explained by the greater geometrical flexibility of hydrogen bonding to O atoms compared with that to N atoms (Frey, 1994). Of the remaining 20 water molecules, 15 are hydrogen bonded only to other water molecules (not protein atoms), three are hydrogen bonded to the sulfate ion, while two which belong to discrete strong peaks of electron density form weak contacts with the protein (W95 at 3.5 Å from Glu58 OE1 and W32 at 3.9 Å from the hydrophobic surface of Pro44). Taking into consideration the leaf-like structure of erabutoxin, favourable solvation is observed on either side of the leaf while the residues at the periphery are mostly involved in crystal contacts (see §3.2.3). In addition, most of the solvent molecules of the protein are connected between them and to others of neighbouring proteins, creating an extensive hydrogen-bonding network throughout the crystal lattice. The network comprises water molecules grouped into short strings linked by hydrogen bonds, while other water molecules form smaller or larger clusters with irregular polygonal arrangements. No immediate correlation of the contacts of the solvent molecules to their $B$ factors was found, yet most of the

solvent atoms not bonded to the protein are characterized by large $B$ values.

The solvent structure of Ea was compared to that of 3EBX (the solvent structure of 5EBX is less complete and was not compared). Water molecules were classified as equivalent if, after least-squares fitting of Cα backbones, they lay within 1.0 Å of one another and formed hydrogen bonds to the same protein, water or $SO_4^{2-}$ atom(s). The superposition (r.m.s. deviation 0.137 Å) gave 62 such equivalent solvent sites (38 were within 0.5 Å). Of these, 57 (66% of the protein hydration shell) are directly bonded to the protein molecule, whereas three are bonded to other water molecules and two are bonded to the sulfate ion. These results indicate that a substantial proportion of the bound solvent molecules occupy discrete sites and that the crystallization conditions have little effect on their location, as the crystals of Ea and 3EBX were grown under different conditions.

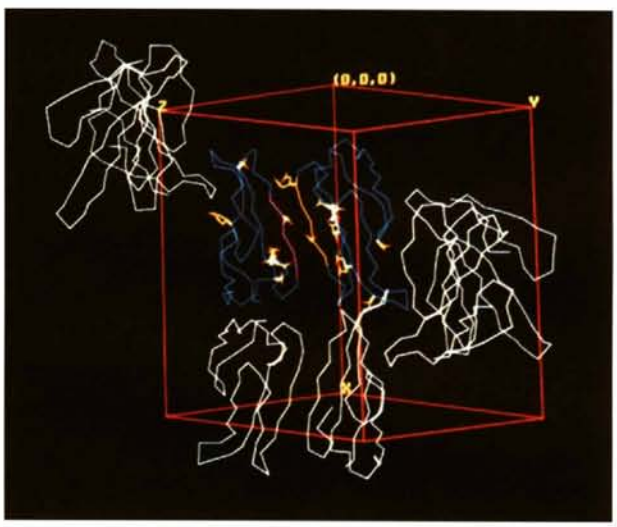



Fig. 6. View of the packing of the dimer with molecules $A$ and $B$ (blue). Residues of the antiparallel β-sheet are coloured orange in molecule $A$ and magenta in molecule $B$. Side chains modelled as dual conformers are coloured yellow. The tail-to-tail interaction can be seen between molecule $A$ and a symmetry-related molecule $B$ below (white) along the $x$ axis ($O$ program, Jones *et al.*, 1991).

3.2.3. *Crystal packing*. Various contacts occur within the erabutoxin crystal structure: (i) each reference molecule is directly linked to eight other symmetry-related molecules by 16 hydrogen bonds (made up of two equivalent sets of eight different contacts). They range from 2.3 to 3.2 Å, two of the eight hydrogen bonds involve main-chain–main-chain interactions and five connect molecules related by translation along the $c$ axis (Table 5). (ii) The hydrated sulfate ion (47.8 $Å^2$ mean $B$ factor) forms a salt bridge to the side chain of Asn5 in one protein molecule, and another to the side chain of Arg1 in a symmetry-related molecule translated along the $c$ axis. The bonding network of the sulfate ion is shown in Table 5. (iii) The aforementioned intermolecular bridges mediated by the hydration shell water molecules create a network that contributes to the stabilization of the crystal structure. The high number of intermolecular contacts in the erabutoxin crystal structure may partially account for the high diffraction quality of its crystals. The same pattern of crystal contacts has been found in 5EBX; likewise in the 3EBX structure, with the exception that only six of the eight intermolecular hydrogen bonds are observed.

### 3.3. *Description of the dimer*

The association of two independent Ea molecules to form a dimer leads to a structure with characteristic features as far as the interaction between the two molecules as well as the entire packing arrangement is concerned (Fig. 6). The two molecules $A$ and $B$ are in close contact through an intermolecular antiparallel two-stranded β-sheet between the same residues (52→56) of both molecules, which are related by a twofold non-crystallographic axis perpendicular to the β-sheet (Fig. 7). The transformation matrix giving the

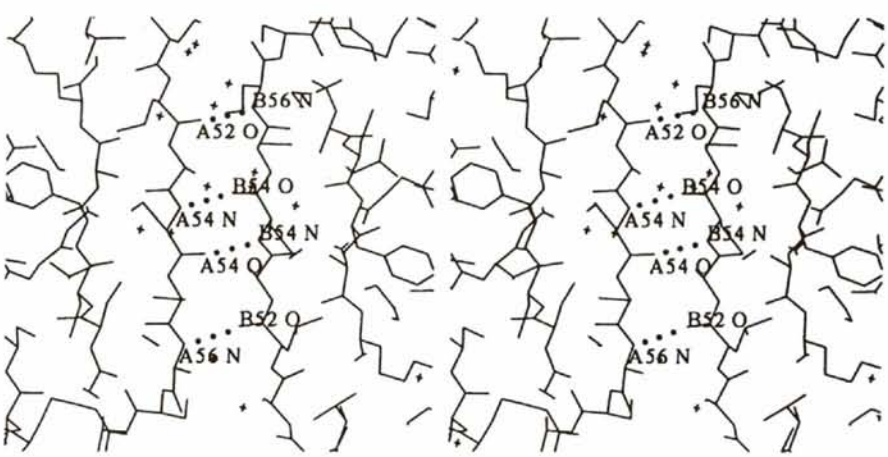



Fig. 7. The interface between the two dimer molecules consists of an antiparallel β-sheet formed by the same amino acids (52→56) in both molecules. They are related by a twofold non-crystallographic axis running perpendicular in the middle of the β-sheet ($O$ program, Jones *et al.*, 1991).

Table 6. *Transformation operations relating molecules A and B of the dimer, $X_A = MX_B + V$*

| Rotations† $M$ | | | Translations $V$ |
|---|---|---|---|
| −0.9980 | −0.0264 | −0.0579 | 48.3697 |
| −0.0635 | 0.3626 | 0.9298 | −8.3110 |
| −0.0035 | 0.9316 | −0.3635 | 14.5737 |

† The rotations for the $B \rightarrow A$ transformation around the crystallographic axes $a$, $b$ and $c$ are 111.3°, 0.2° and −176.3°, respectively.

relation between the two independent molecules is given in Table 6. The rotational relation of molecule $B$ to molecule $A$ around the crystallographic axes $a$, $b$ and $c$ is 111.3, 0.2 and −176.3°, respectively. This kind of dimeric association is also encountered in several snake-venom toxins, *e.g.* cardiotoxin (Rees *et al.*, 1990) and $\alpha$-bungarotoxin (Love & Stroud, 1986). Inspection of the superimposed backbones of molecules $A$ and $B$ with Ea, as well as of the entire dimer with the 6EBX dimer shows the similarity of the main chains (Table 3). The most striking differences appear in the long loop regions (Fig. 8). In general terms, the dimeric association does not significantly influence the overall backbone conformation of the two erabutoxin molecules. Side chains diverge more, probably to accommodate crystal packing. The essential differences in the side chains of molecules $A$ and $B$ are those of Phe32, Arg33 and Glu56 which adopt significantly different orientations observed on visual inspection of the superimposed molecules. As in the case of the monomeric structure, these differences are also obvious after a residue-by-residue analysis of side-chain rotamers of all deposited erabutoxin structures [seven models: 1QKE (present monomer), 3EBX, 5EBX, 1QKD molecule $A$, 1QKD molecule $B$ (present dimer), 6EBX molecule $A$ and 6EBX molecule $B$] performed using the program *PROCHECK-COMP* (Laskowski *et al.*, 1993). This analysis shows that the side chains of all erabutoxins are in general similar with few exceptions, the most significant of which are mentioned above. More specifically, out of 53 residues only 19 differ significantly in their $\chi_1$ angle and out of 38 residues only 12 exhibit significant differences in their $\chi_2$ angle. It is noteworthy that only six residues differ in both their $\chi_1$ and $\chi_2$ angles (Gln10, Lys15, Phe32, Ile50, Glu56 and Asn62).

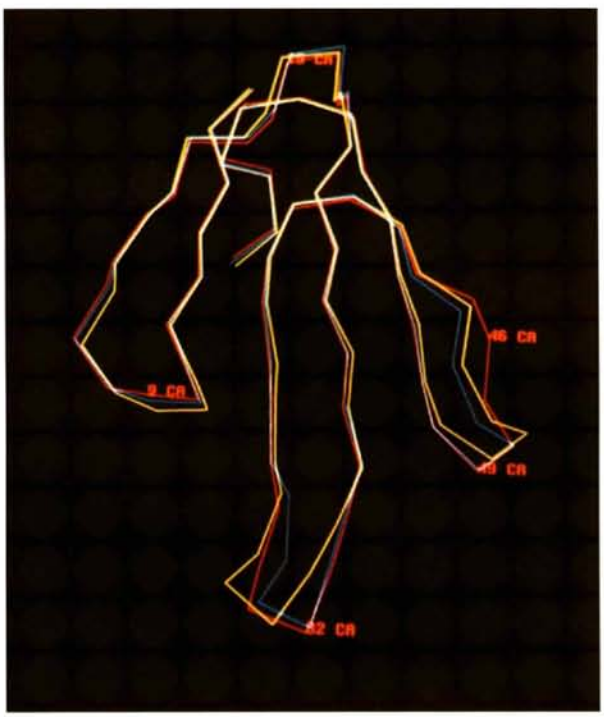



Fig. 8. Superposition of the Cα trace of the monomer (red) with molecule $A$ (yellow) and molecule $B$ (blue) of the dimer (*O* program, Jones *et al.*, 1991).

In general, the variation of $B$ factors along the chains of molecules $A$ and $B$ resembles the monomer (Fig. 2). However, there is a higher similarity between the $B$-factor profiles of the monomer and molecule $B$ of the dimer which cannot be explained on the basis of the crystal contacts described in Tables 5 and 7. Molecule $B$ exhibits a slightly higher degree of motion with respect to molecule $A$. This is particularly apparent in the agitated region 31→34 which is located at the tight turn of the middle loop; the same feature is also present in the 6EBX dimeric model. As expected, slightly lower $B$ factors are observed for the residues engaged in the intermolecular $\beta$-sheet.

3.3.1. *Structural heterogeneity*. Distinct alternative orientations were observed for both C-terminal residues Asn$A$62 and Asn$B$62 as well as for the side chains of 11 residues: Ser$A$8, Asn$A$26, Gln$A$28, Ser$A$30, Ser$A$53, Glu$A$56, Ser$B$22, Ser$B$23, Ser$B$30, Glu$B$38 and Ser$B$53, which were modelled as dual conformers (Figs. 4*b*, 4*c* and 5*b*). We observe that only two residues with alternate conformations (Ser23 and Glu38) are common to the monomer and molecule $B$ of the dimer. These residues have similar rotamers for their alternate side chains. The C-terminal Asn62 has alternate conformers in all three molecules but their rotamers are different in each case because of the flexibility in this region. We also note that three peaks were detected for Ser$B$30 OG and this side chain was modelled with three different orientations (35:35:30). It is noteworthy that Ser30 is involved in an intermolecular hydrogen-bonding interaction between molecule $A$ and a symmetry-related molecule $B$ (see §3.3.3). The first and second conformer of Ser30 in both molecules $A$ and $B$ are exactly the same ($\chi_1$ = 61° and −35°, respectively) and the third conformer for Ser$B$30 has a value of $\chi_1$ = 165°, all of which are different from that found in the monomeric structure ($\chi_1$ = −71°). Ser53 is the third residue with an alternative conformation common to molecule $A$ and molecule $B$ of the dimer and it has very similar rotamers in both molecules. In the 6EBX model only the Phe$B$32

Table 7. *Contacts in the crystal packing of the dimer (<3.30 Å)*

Contacts at the interface of the two independent molecules

| Molecule *A* | Molecule *B* | Distance (Å) |
|---|---|---|
| Leu52 O | Glu56 N | 2.97 |
| Cys54 N | Cys54 O | 2.81 |
| Cys54 O | Cys54 N | 2.81 |
| Glu56 N | Leu52 O | 3.08 |
| Ser53 OG | Ser53 OG | 2.91 |
| Glu56 OE2 | Leu52 N | 2.42 |
| Ser57 OG | Lys51 N | 2.52 |
| Glu58 O | Lys51 NZ | 2.66 |

Contacts of molecule *A* with symmetry-related molecules *B*

| Molecule *A* | Symmetry-related molecule *B* | Distance (Å) |
|---|---|---|
| Gln7 OE1 | Ser22$^{i}$ OG† | 2.61 |
| Gln7 OE1 | Ser22$^{i}$ N | 2.86 |
| Gln7 NE2 | Glu21$^{i}$ OE1 | 3.07 |
| Gln10 NE2 | Glu56$^{i}$ O1 | 2.95 |
| Gln12 O | Thr16$^{ii}$ OG1 | 2.78 |
| Thr14 O | Thr14$^{ii}$ OG1 | 2.29 |
| Thr16 OG1 | Gln12$^{ii}$ O | 2.90 |
| Glu21 OE1 | Lys47$^{iii}$ NZ | 2.52 |
| Ser30 N | Ser30$^{iv}$ O | 2.89 |
| Pro48 O | Thr35$^{iv}$ N | 3.34 (>3.30) |
| Gly49 O | Ser30$^{iv}$ OG | 2.64 |
| Gly49 O | Ser30$^{iv}$ OG† | 3.04 |

Symmetry operators: (i) $-x+1, y+1/2, -z+1/2$; (ii) $x-1/2, -y+1/2, -z+1$; (iii) $-x+1/2, -y, z-1/2$; (iv) $x+1/2, -y+1/2, -z+1$. † Alternate conformer of disordered side chains.

side chain showed some degree of disorder. As in the monomer, most of these alternate side-chain conformers are hydrogen bonded to surrounding solvent molecules.

3.3.2. *Solvent structure*. A total of 206 sites for water molecules (numbered from 1 to 206) were included in the final model (Figs. 4*b* and 4*c*) and the sum of all water occupancy factors equals 152. The calculated solvent content is 44% and, from density considerations, the expected solvent content is 447 water molecules. In the 6EBX dimer, a total of 99 water molecules were located. We therefore observe that a higher amount of solvent has been located for both monomer and dimer in the present study. The average *B* factor of the water molecules is 37.8 Å$^2$ and the shortest water-to-water distance is 2.28 Å. Each of the molecules *A* and *B* is surrounded by 76 water molecules which are directly linked to the molecule by one or more hydrogen bonds (hydration shell). Of these water molecules, 24 are also bonded to symmetry-related protein atoms and 57 are bonded to other water molecules of the crystal lattice. Moreover, five other waters are located between molecules *A* and *B* and form hydrogen bonds to both of them. Of the remaining water molecules (most of which are marked by high *B* factors), 42 are only bonded to other waters and five form weak contacts (greater than the 3.3 Å cut-off limit) with the protein or other water molecules with distances ranging from 3.4 to 3.9 Å. Finally, two water molecules were found at distances above 4 Å but were accepted because they occupy strong discrete peaks of electron density in the neighbourhood of which weak peaks were found that were not modelled. The distribution of the solvent sites is similar to that of the monomer with a preference for the faces of the protein. Again, short string and cluster arrangements exist in the solvent network. Applying the same criteria for equivalence, superposition of the monomer and molecule *A* (r.m.s. deviation 0.865 Å) revealed 23 equivalent water sites (26% of the protein hydration shell) and superposition of the monomer and molecule *B* (r.m.s. deviation 0.768 Å) revealed 32 equivalent sites (41% of the protein hydration shell). Finally, superposition of the two molecules *A* and *B* (r.m.s. deviation 0.680 Å) gave 36 equivalent solvent sites. This represents 43% of the protein hydration shell of each molecule, which is a considerable percentage taking into account the different packing arrangements of the two independent molecules (see below).

3.3.3. *Crystal packing*. Two important types of intermolecular contacts are present in the crystal packing of the dimer: (i) four contacts at the interface of the two independent molecules *A* and *B*, forming an antiparallel $\beta$-sheet between the same residues (52→56) of both molecules (Fig. 7). This association is further enhanced by four short hydrogen bonds between residues of *A* and *B* located in the $\beta$-sheet region as well as by five cross-linking solvent molecules. These *A*–*B* interactions are shown in Table 7. (ii) Each molecule *A* and *B* is directly hydrogen bonded to four different symmetry-related molecules *B* and *A*, respectively, resulting in a total of 12 hydrogen bonds for each molecule *A* and *B* (Table 7). Among them, of particular interest is the interaction involving residues Ser30, Pro48 and Gly49 of molecule *A* with Ser30 and Thr35 of a symmetry-related molecule *B* translated along the $\alpha$ axis. These residues are located at the end of the long loops and this kind of interaction could best be described as tail-to-tail (Fig. 6). A similar interaction (reported as head-to-head), but with longer distances, was observed in the 6EBX model. Crystal contacts of the type *A*–*A*$^{i}$ and *B*–*B*$^{i}$ were not observed (where *A*$^{i}$ and *B*$^{i}$ are symmetry-related molecules). As in the monomer, the close packing of the protein is stabilized by the hydrogen-bond network mediated by the solvent molecules of the crystal structure.

## 4. Concluding remarks

The structures of the monomeric and dimeric erabutoxin a were compared with each other and with the PDB-deposited structures 3EBX, 5EBX and dimeric 6EBX. The molecules are, in general, similar and only certain differences appear in the orientation of the side chains, which is usually dependent on the specific crystal used for data collection. Dimerization has little influence on the overall conformation of the protein molecule. The improved diffraction data of the current report,

compared with those of the structure 5EBX, are presumably due to the larger crystal size and the use of synchrotron radiation. The high-resolution density maps provide detailed structural information regarding multiple side-chain conformers and made possible an accurate determination of the solvent structure. This study, together with earlier determinations, confirms that ordered water structure is an important feature of the erabutoxin protein. In particular, this protein is associated with a characteristic hydration shell, a significant part of which remains invariant irrespective of crystal form (monomer or dimer) or conditions of crystallization.

We wish to thank Dr P. A. Tucker and Dr M. Vlassi for their advice during data processing and refinement, and Dr Z. Dauter and Dr K. Wilson of the EMBL Hamburg Outstation for their help during the synchrotron experiments. This work has been supported financially by the Greek Secretariat of Research and Technology.

## References

Arnoux, B., Ménez, R., Drevet, P., Boulain, J.-C., Ducruix, A. & Ménez, A. (1994). *FEBS Lett.* **342**, 12–14.

Baker, E. N. & Hubbard, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.

Bourne, P. E., Sato, A., Corfield, P. W. R., Rosen, L. S., Birken, S. & Low, B. W. (1985). *Eur. J. Biochem.* **153**, 521–527.

Brünger, A. T. (1992*a*). *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR.* Yale University, Connecticut, USA.

Brünger, A. T. (1992*b*). *Nature (London)*, **355**, 472–475.

Chang, C. C. (1979). *Handbook of Experimental Pharmacology: Snake Venoms*, edited by C. Y. Lee, Vol. 52, pp. 309–376. Berlin: Springer-Verlag.

Changeux, J.-P. (1990). *Trends Pharmacol. Sci.* **11**, 485–492.

Corfield, P. W. R., Lee, T.-J. & Low, B. W. (1989). *J. Biol. Chem.* **264**, 9239–9242.

Drenth, J., Low, B. W., Richardson, J. & Wright, C. (1980). *J. Biol. Chem.* **255**, 2652–2655.

Endo, T. & Tamiya, N. (1987). *Pharmacol. Ther.* **34**, 403–451.

Endo, T. & Tamiya, N. (1991). *Snake Toxins*, edited by A. L. Harvey, pp. 165–222. New York: Pergamon Press.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Frey, M. (1994). *Acta Cryst.* D**50**, 663–666.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kimball, M. R., Sato, A., Richardson, J. S., Rosen, L. S. & Low, B. W. (1979). *Biochem. Biophys. Res. Commun.* **88**, 950–959.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Love, R. A. & Stroud, M. (1986). *Protein Eng.* **1**, 37–46.

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.

Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Ménez, A. (1987). *Recherche*, **18**, 886–893.

Ménez, R. & Ducruix, A. (1990). *J. Mol. Biol.* **216**, 233–234.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pillet, L., Trémeau, O., Ducancel, F., Drevet, P., Zinn-Justin, S., Pinkasfeld, S., Boulain, J.-C. & Ménez, A. (1993). *J. Biol. Chem.* **268**, 909–916.

Ramakrishnan, C. & Ramachandran, G. N. (1965). *Biophys. J.* **5**, 909–933.

Rees, B., Bilwes, A., Samama, J. P. & Moras, D. (1990). *J. Mol. Biol.* **214**, 281–297.

Riès-Kautt, M. & Ducruix, A. (1989). *J. Biol. Chem.* **264**, 745–748.

Riès-Kautt, M. & Ducruix, A. (1991). *J. Cryst. Growth*, **101**, 20–25.

Saludjian, P., Prangé, T., Navaza, J., Ménez, R., Guilloteau, J. P., Riès-Kautt, M. & Ducruix, A. (1992). *Acta Cryst.* B**48**, 520–531.

Smith, J. L., Corfield, P. W. R., Hendrickson, W. A. & Low, B. W. (1988). *Acta Cryst.* A**44**, 357–368.

Smith, J. L., Hendrickson, W. A., Honzatko, R. B. & Sheriff, S. (1986). *Biochemistry*, **25**, 5018–5027.

Tamiya, N. & Abe, H. (1972). *Biochem. J.* **130**, 547–555.

Tamiya, N. & Arai, H. (1966). *Biochem. J.* **99**, 624–630.

Trémeau, O., Lemaire, C., Drevet, P., Pinkasfeld, S., Ducancel, F., Boulain, J.-C. & Ménez, A. (1995). *J. Biol. Chem.* **270**, 9362–9369.

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.

Tsernoglou, D. & Petsko, G. (1976). *FEBS Lett.* **68**, 1–4.

Tsernoglou, D. & Petsko, G. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 971–974.

Tsetlin, V., Pluzhnikov, K., Karelin, A. & Ivanov, V. (1983). *Toxins as Tools in Neurochemistry*, edited by F. Hucho & Y. A. Ovchinnikov, pp. 159–169. Berlin: Walter de Gruyter.

Tu, A. T., Hong, B.-S. & Solie, T. N. (1971). *Biochemistry*, **10**, 1295–1304.

Weber, M. & Changeux, J.-P. (1974). *Mol. Pharmacol.* **10**, 1–14.