Solution Structure of Neuronal Bungarotoxin Determined by Two-Dimensional NMR Spectroscopy: Calculation of Tertiary Structure Using Systematic Homologous Model Building, Dynamical Simulated Annealing, and Restrained Molecular Dynamics[†,‡]

Michael J. Sutcliffe,[§] Christopher M. Dobson,[∥] and Robert E. Oswald*,[⊥]

Biological NMR Centre, University of Leicester, P.O. Box 138, Medical Sciences Building, University Road, Leicester LE1 9HN, U.K., Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, U.K., and Department of Pharmacology, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853

Received October 1, 1991; Revised Manuscript Received December 19, 1991

ABSTRACT: Neuronal bungarotoxin has previously been shown, using two-dimensional ^{1}H NMR spectroscopy, to have a triple-stranded antiparallel β-sheet structure which dimerizes in solution [Oswald, R. E., Sutcliffe, M. J., Bamberger, M., Loring, R. H., Braswell, E., & Dobson, C. M. (1991) *Biochemistry 30*, 4901–4909]. In this paper, structural calculations are described which use the 582 experimentally measured NOE restraints in conjunction with 27 ϕ-angle restraints from J-value measurements. The positions of the N-terminal region and C-terminal region were poorly defined in the calculated structures with respect to the remainder of the structure. The region of the structure containing the triple-stranded β-sheet was, however, well defined and similar to that found in the structure of homologous α-bungarotoxin (45% amino acid identity). The experimental restraints did not result in a well-defined dimer interface region because of the small number of NOEs which could be identified in this region. An approach was therefore adopted which produced model structures based to varying degrees on the α-bungarotoxin structure. Fourteen different structures were generated in this manner and subsequently used as starting points for refinement using dynamical simulated annealing followed by restrained molecular dynamics. This approach, which combines NMR data and homologous model building, has enabled a family of structures to be proposed for the dimeric molecule. In particular, Phe 49 has been identified as possibly playing an important role in dimer formation, this residue in one chain interacting with the corresponding residue in the adjacent chain.

The nicotinic acetylcholine receptors (nAChRs)[1] are non-selective cation channels which are activated by the binding of acetylcholine [ACh; for a review, see Galzi et al. (1991)]. In addition to acetylcholine, a variety of other agents interact with nAChRs. Of these agents, snake α-neurotoxins interact with a site which overlaps the ACh-binding sites, thereby blocking the functional activity either reversibly or irreversibly depending upon the type of toxin and nAChR. Over 60 homologous α-neurotoxins have been purified and sequenced from elapid venoms, and several crystal structures are known (Low et al., 1976; Tsernoglou & Petsko, 1976; Walkinshaw et al., 1980; Love & Stroud, 1986). Some aspects of the solution structure of α-bungarotoxin (αBgt; Basus et al., 1988) and a variety of other α-neurotoxins [e.g., Endo et al. (1981), Inagaki et al (1985), and Labhardt et al. (1988)] have been determined by NMR spectroscopy. In addition, the details of the interaction between αBgt and a peptide derived from a nAChR were studied by two-dimensional NMR spectroscopy (Hawrot & Basus, 1991).

In our previous work (Oswald et al., 1991), we assigned the ^{1}H NMR spectrum of neuronal bungarotoxin (nBgt) and determined the NH–αH coupling constants and NH exchange rates. These results have shown that nBgt consists of a three-looped structure with a central triple-stranded antiparallel β-sheet. We have confirmed the fact that nBgt forms a dimer in solution [as shown by Chiappinelli and Lee (1985)] and shown that, at least over the time scale of the NMR experiments, the two individual chains have the same average conformation. In addition, we have found that the interface is along one edge of the β-sheet, thereby forming a six-stranded intermolecular antiparallel β-sheet. A knowledge of the three-dimensional structure of nBgt is important in order to relate the differences in sequence between nBgt and other neurotoxins to differences in function. Determination of the solution structure of a dimeric molecule with identical individual chains is complicated by the difficulty in distinguishing intra- from interchain NOEs. These difficulties can in principle be overcome by using isotopic labeling to distinguish the two components of the dimer (Arrowsmith et al., 1990). Such

[†] This work was supported by grants from the National Institutes of Health (R01 NS18660 and a senior fellowship), the Fulbright Foundation, and the Cornell Biotechnology Institute and Polygen Corporation to REO. M.J.S. is a Royal Society University Research Fellow. C.M.D. is a member of the Oxford Centre for Molecular Sciences, which is supported by the Science and Engineering Research Council and the Medical Research Council.

[‡] The coordinates of STRUCTURES 1, 2, 3, 4, 6, 8, 9, 10, 11, 12, 13, and 14 have been deposited in the Brookhaven Protein Data Bank as entry 1NBT, and the restraints were deposited as entry R1NBTMR.

* Address correspondence to this author.

[§] University of Leicester.

[∥] University of Oxford.

[⊥] Cornell University.

[1] Abbreviations: ACh, acetylcholine; αBgt, α-bungarotoxin; αH, α-proton; C_α, α-carbon; CO, carbonyl oxygen; COSY, correlation spectroscopy; nAChR, nicotinic acetylcholine receptor; nBgt, neuronal bungarotoxin; NH, amide proton; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; RMS, root mean square deviation.

0006-2960/92/0431-2962$03.00/0

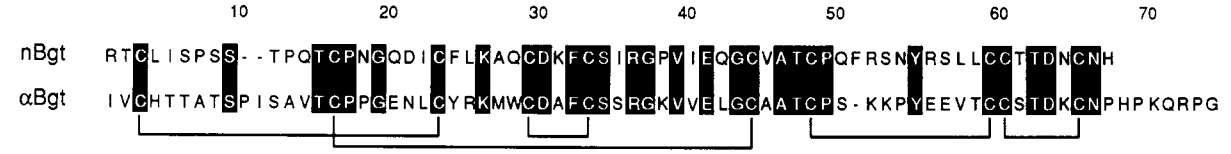

FIGURE 1: Amino acid sequences of nBgt and αBgt. The numbering is that corresponding to αBgt. Identical residues are shown with a black background, and the position of disulfide bonds are indicated by solid lines.

labeling has not, however, yet been possible for nBgt since the protein is not presently available in recombinant form. Over recent years, homologous model building [for a review, see Sutcliffe (1989)] has become established as a technique for predicting the structure of a protein for which the structures of homologous/analogous proteins are known. In this work, we combine the results of ^{1}H NMR spectroscopy with the technique of homologous model building, in conjuction with dynamical simulated annealing and molecular dynamics refinement, in order to propose a structural description of the nBgt dimer.

Materials and Methods

Determination of NMR-Derived Restraints. NMR experiments were performed as described previously (Oswald et al., 1991). A total of 582 NOEs (570 assumed to be intrachain and 12 to be interchain, i.e., 8.7 NOEs per residue) were determined from 2-D NOESY experiments (see Supplementary Material) and classified as strong (<2.5 Å), medium (<3.5 Å), and weak (<5.0 Å) on the basis of their intensities. Distance bounds were based on conservative estimates of the sequential αH–NH separation along a β-strand, the NH–NH interstrand separation across a β-sheet, and the maximum separation expected under the experimental conditions for an observable NOE, respectively (Wüthrich, 1986). NH–αH coupling constants were determined for 27 residues from COSY experiments using procedures described previously (Oswald et al., 1991). Coupling constants greater than 8 Hz were used to define an angular range of −160° to −120°. The pattern of disulfide bonds was assumed to be identical to that in other long α-neurotoxins (see Figure 1). Additionally, eight intrachain and four interchain hydrogen-bonding restraints were used during the initial stages of structural refinement.

In the case of calculations based on a monomeric structure, the dynamical simulated annealing stage was preceded by initial structure determination from distance geometry; details of the specific procedure used here have been described previously (Sutcliffe & Dobson, 1991). In the calculations involving the dimeric structure, the procedure involved generating initial structures by molecular model building using the automated homologous model building program COMPOSER (Blundell et al., 1988; Sutcliffe, 1989). It also involved using, in the simulated annealing and restrained molecular dynamics stages, an additional term in the force field due to the symmetry of the problem to satisfy the experimental observation of a symmetric dimer. In a series of test calculations, the following procedure was found to produce the most satisfactory ensemble of structures when analyzed on the basis of NOE violations, van der Waals energy, free energy of solvation, and a geometric figure of merit (a penalty function which combines bonds, angles, impropers, and van der Waals interactions; Jian-Sheng & Hubbard, 1990). A quadratic function was used to preserve symmetry during the refinement such that the resulting energy was proportional to the square of the distance between equivalent atoms after the two chains had been superimposed. The symmetry force constant (k_{sym}) was initially set to a low value [5 kcal/(mol·Å^2)] and was increased over

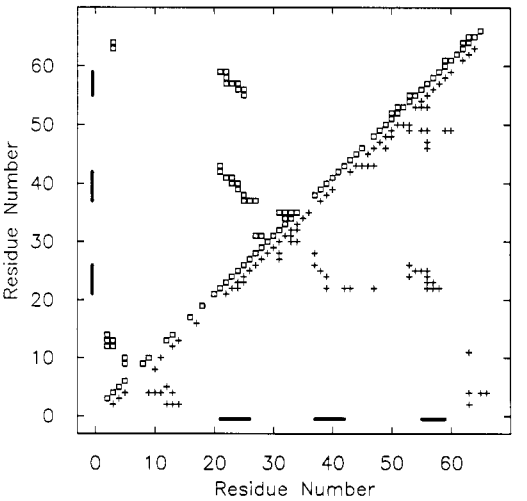

FIGURE 2: Distribution of NOEs for nBgt. NOEs between main-chain protons are shown above the diagonal (□); NOEs between side-chain–main-chain and side-chain–side-chain protons are shown below the diagonal (+).

the first 1.5 ps of the refinement to 50 kcal/(mol·Å^2) in 5 kcal/(mol·Å^2) steps every 150 fs. During this first 1.5 ps, the NOE and dihedral angle force constants were set to a relatively high value, and the repulsive part of the van der Waals interaction was set to a relatively low value, thereby allowing the experimentally derived restaints to be met with little regard for steric interactions. During this period, the disulfide bonds were included as the respective distance restraints rather than as covalent bonds, the latter being the case during the remainder of the refinement. Over the remaining 3 ps of the simulated annealing stage, k_{sym} remained at 50 kcal/(mol·Å^2), whereas the repulsive part of the van der Waals interaction was steadily increased in order to improve the nonbonded interactions. k_{sym} remained at 50 kcal/(mol·Å^2) during the 8.8 ps of restrained molecular dynamics. The explicit hydrogen-bond restraints were removed during the last 3 ps of restrained molecular dynamics.

The distance geometry and dynamics calculations were performed on an Alliant FX-40/3 minisupercomputer and a Convex C210 minisupercomputer using a program kindly supplied by Dr. R. Scheek (Groningen) and XPLOR (Brunger et al., 1987), respectively. Homologous model building and visual inspection of the structures were performed on a Silicon Graphics 4D/25 and a 4D/220 using COMPOSER and QUANTA (Polygen Corporation, Waltham, MA), respectively.

Results and Discussion

Pattern of NOEs. The NOE pattern within a monomer is illustrated in Figure 2 and extends our previous more limited analysis (Oswald et al., 1991). The position of the β-sheet is confirmed by the bands perpendicular to and above the diagonal; the absence of a band parallel to and above the diagonal confirms that no helical structure is present. Apart from the β-sheet, the only other notable long-range contacts are between the N- and C-terminal regions, showing that these two regions are spatially close.

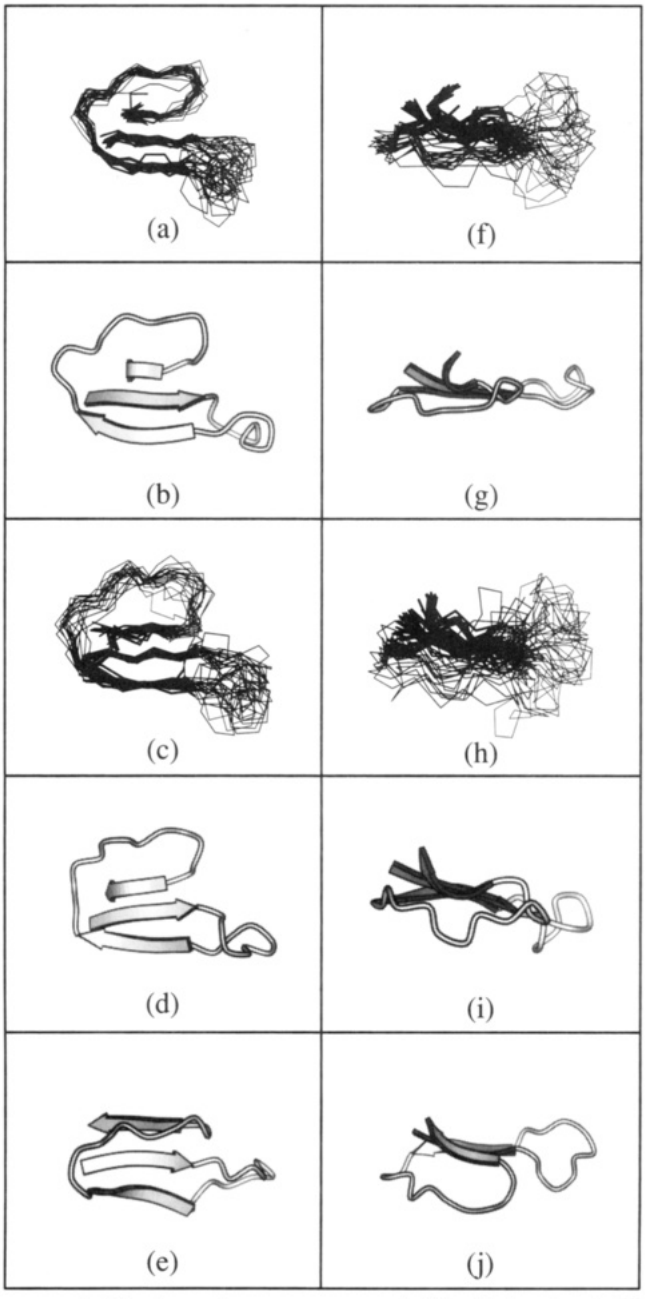

FIGURE 3: Topology of monomers generated for nBgt and a single chain from αBgt. The 20 monomers generated including the 12 NOEs subsequently treated as interchain NOEs (hence the kink in the third β-strand) are shown by a C_α trace in panels a and f (the β-strands are denoted by thicker lines than the remainder of the structure); the topology of one of these is shown schematically in panels b and g. The 20 monomers generated excluding the 12 NOEs treated as interchain are shown by a C_α trace in panels c and h. The topology of one of these is shown schematically in panels d and i, and the topology of a single chain from αBgt is shown in panels e and j. Panels a–e are looking at the β-sheet from above, and panels f–j are looking along the β-sheet from the side. The poorly defined N-terminal (Arg 1–Ile 20) and C-terminal (Thr 60–His 66) regions have been removed for clarity. This figure was generated by the program MolScript (Kraulis, 1991).

Generation of Monomeric Structures. The restraints were used in conjunction with the distance geometry and the simulated annealing approach to generate a set of structures on the assumption that all observed NOEs were intramolecular. This procedure resulted in structures with consistent violations of NOE restraints and a severely distorted β-strand (Ser 55–Cys 59; see Figure 3a,b,f,g). This can be attributed to the fact that certain of the NOEs were intermolecular rather than

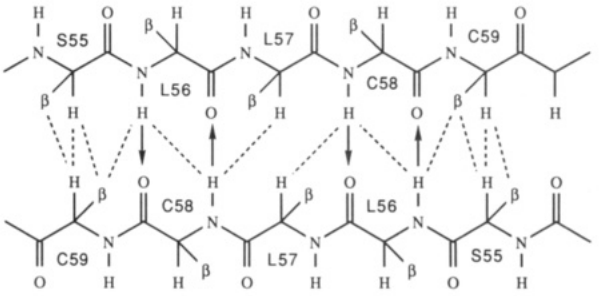

FIGURE 4: Schematic representation of the interface between the two intramolecular β-sheets. Arrows represent hydrogen bonds, dashed lines represent intermolecular NOEs, and "β" represents β-protons.

intramolecular. Examination of these structures indicated 12 NOEs which could be attributed to intermolecular interactions; these were removed from the data set for a subsequent series of calculations. Omission of these NOEs removed the consistent NOE violations and gave a set of 20 structures with a better defined, although still somewhat distorted, triple-stranded β-sheet (see Figure 3c,d,h,i). The N- and C-terminal regions, however, are poorly defined; this can be attributed to the small number of long-range NOEs in these regions. These structures have a C_α RMS from their mean of 3.6 ± 0.5 Å for the entire molecule and 1.1 ± 0.3 Å for the β-sheet region. Their overall topology is closely similar to that of a single chain in the homologous αBgt with the exception of the region Pro 47–Ser 51. In the structures of nBgt calculated at this stage, this region is lateral to the β-sheet; whereas in the crystal structure of αBgt, it is perpendicular to the β-sheet (see Figure 3e,j). In the NOE data set, there are very few NOEs linking this region to the β-sheet; this difference may, therefore, be the result of the tendency for distance geometry to produce extended structures in poorly defined regions of the structure (Metzler et al., 1989).

Generation of Dimeric Structures. Once the structure of the monomer had been determined, attempts were made to generate dimeric structures by docking the predetermined monomeric structures together on the basis of the 12 experimentally observed NOEs that were assumed to be intermolecular (see Figure 4) using a simulated annealing approach. This, however, produced structures with unfavorable steric interactions. Similar steric problems were encountered when the dimeric structures were generated using distance geometry. In order to overcome these problems, an alternative strategy was adopted. The overall chain topology and mode of dimer formation deduced here for nBgt is closely similar to that observed in the crystal structure of αBgt [Love & Stroud, 1986; 2ABX in the Brookhaven Protein Data Bank (Bernstein et al., 1977)]. Furthermore, αBgt has 45% amino acid identity with nBgt (see Figure 1). Therefore, the initial structures were generated using the approach of homologous model building.

Construction of the Initial Model Structures. A prerequisite to the construction of an initial set of model structures is that the tertiary structure of at least one similar protein must be known [for a review, see Sutcliffe (1989)]. Because of the relatively small number of intermolecular NOEs, a similar structure which exists as a dimer is required. Of the known tertiary structures, only the 2.5-Å resolution crystal structure of αBgt meets these requirements. The use of αBgt as a starting model is further justified by the strong correlation of NH and αH chemical shifts of αBgt with the chemical shifts for the corresponding protons of nBgt (Oswald et al., 1991).

The strategy for choosing the starting models was to produce a diverse set of conformations for nBgt based on the αBgt structure for subsequent refinement using the NMR data

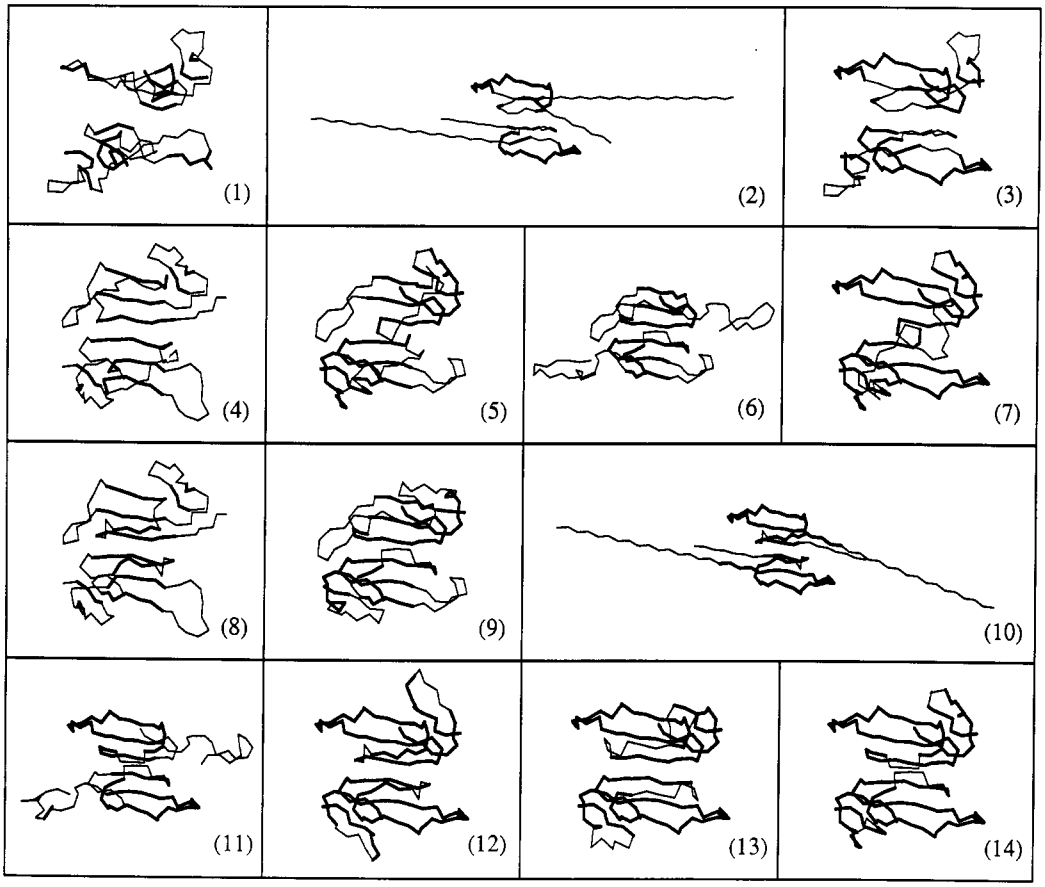

FIGURE 5: Starting models based on αBgt. Shown are structures generated by the homologous model building program COMPOSER. The portions of the structure based on αBgt are shown with thick lines, and the portions modeled from structures in the Brookhaven Protein Data Bank are shown with thin lines. The models were constructed from the αBgt as shown in the legend for Table I. Because of the extended nature of the models 2 and 10, they are shown at lower magnification than the remainder of the structures.

determined experimentally for nBgt. In addition to producing a structure consistent with the experimental data, these particular models were selected in order to answer in particular the following questions: (1) How well defined is the position of the N-terminal region with respect to the remainder of the protein? and (2) Does the Pro 47–Ser 51 loop form a part of the dimer interface?

Fourteen different models of the nBgt dimer were constructed using the automated homologous model building program COMPOSER (Blundell et al., 1988; Sutcliffe, 1989). These are shown in Figure 5 and summarized in Table I. The models incorporated the structure of αBgt to varying degrees. All models were obtained by specifying the region of the models to be based on the αBgt structure. (Note that COMPOSER requires each structural fragment from αBgt to be a minimum of three residues in length.) The remainder of the structure was then generated by searching a database of protein fragments to find the fragment with both an end-point geometry consistent with the parts of the model based on αBgt (to which it is to be joined) and the conformation predicted to correspond most closely to that predicted by the sequence of nBgt in that region (Sutcliffe, 1989). The conformations of side chains not identical to those in nBgt were defined by a set of rules based on observed amino acid substitutions (Sutcliffe et al., 1987b). The models[2] varied from one that only had the disulfide architecture conserved (MODEL 1), through a series containing the triple-stranded β-sheet, to several in which a minimum number of insertions and deletions were used (MODELS 11, 12, 13, and 14).

Calculation of the Dimer Structures. Refinement of the 14 models was performed using the program XPLOR (Brunger et al., 1987) to apply the dynamical simulated annealing–restrained molecular dynamics approach. Some of the properties of the final structures are listed in Table II. In all of these structures, the six β-strands have a relative orientation similar to that found in the αBgt crystal structure. On the basis of van der Waals energy, geometric figure of merit, and NOE violations, two of the structures (STRUCTURES 5 and 7), however, are relatively poor when compared with the other 12 structures. In the case of one of these, the β-sheet was not formed completely because residues Phe 49 and Arg 50 lie between residues Ala 25 and Ser 55. In the other, the region Phe 49–Tyr 53 from one chain was intertwined with the other chain in the starting model, and the long-range NOEs were not sufficiently restraining to resolve this problem during subsequent refinement. C_α representations of 12 structures (i.e., the ensemble excluding these two high-energy structures) are shown in Figure 6a,b (this does not show the poorly defined N- and C-terminal regions) and illustrate the overall fold of nBgt. The relationship between the conformations of the starting models and the final structures is illustrated in Figure 7A. Apart from the two high-energy structures, the final structures have refined away from the region of conformational space occupied by the starting models to a somewhat more limited region of conformational space. The range of conformational space represented by the ensemble of final structures could be due either to the conformations of the starting models or to the random number seeds used to assign

[2] For clarity, the term "MODEL" will refer to the starting conformation and the term "STRUCTURE" will refer to the refined conformation following simulated annealing and restrained molecular dynamics.

Table I: Composition of the Starting Models[a]

MODEL	residue range						
1[a]	1–4[b]	15–17	22–24	28–34	43–49	58–61	64–67
	1–4[c]	13–15	20–22	26–32	41–47	57–60	63–66
2	(1–4)[d]	(15–19)	23–26	29–49	53–55	59–(67)	
	(1–4)	(13–17)	21–24	27–47	52–54	58–(66)	
3	1–4	15–19	23–26	29–49	53–55	59–67	
	1–4	13–17	21–24	27–47	52–54	58–66	
4	2–5	15–17	23–28	39–45	55–61		
	2–5	13–15	21–26	37–43	54–60		
5	1–7	12–27	40–45	56–67			
	1–7	10–25	38–43	55–66			
6	[1–3][e]	[15]–27	40–45	47–49	52–[67]		
	[1–3]	[13]–25	38–43	45–47	51–[66]		
7	1–7	12–45	56–67				
	1–7	10–43	55–66				
8	2–5	15–17	23–28	39–45	48–52	55–61	
	2–5	13–15	21–26	37–43	46–50	54–60	
9	1–3	15–27	40–45	47–49	52–67		
	1–3	13–25	38–43	45–47	51–66		
10	(1–10)	(17)–52	55–(67)				
	(1–10)	(15)–50	54–(66)				
11	[1–7]	[12]–49	52–[67]				
	[1–7]	[10]–47	51–[66]				
12	1–10	17–52	55–67				
	1–10	15–50	54–66				
13	1–6	15–48	55–67				
	1–6	13–46	54–66				
14	1–7	12–49	52–67				
	1–7	10–47	51–66				

[a] The models were constructed from the αBgt dimer as follows: (1) Based on the positions of the disulfide bonds. (2) Based on MODEL 3 with the N-terminal (Arg 1–Ile 20) and C-terminal (Thr 60–His 66) regions fully extended. (3) Based on amino acid identities. (4) Based on the three β-strands in the monomer together with residues in the N-terminal region. (5) Based on the three β-strands in the monomer together with residues in the N-terminal region. (6) Based on MODEL 9 with part of the N-terminal region (Arg 1–Gly 17) rotated as a rigid body to lie as far away from the rest of the structure as possible. (7) Based on MODEL 14 but with the loop Pro 47–Ser 51 not based on αBgt. (8) Based on MODEL 5 but including the region Pro 47–Ser 51. (9) Based on the three β-strands, the Pro 47–Ser 51 loop, and part of the N-terminal region. (10) Based on MODEL 12 with the N-terminal (Arg 1–Ile 20) and C-terminal (Thr 60–His 66) regions fully extended. (11) Based on MODEL 14 with part of the N-terminal region (Arg 1–Gly 17) rotated as a rigid body to lie as far away from the rest of the structure as possible. (12) Based on the maximum number of residues from αBgt using one possible sequence alignment. (13) Similar to MODEL 12 but with fewer residues based on αBgt in both the N-terminal region and between the second and third β-strands. (14) Based on the maximum number of residues from αBgt using another possible sequence alignment. [b] Denotes the residues from αBgt incorporated into the model; the remaining residues are constructed as described in the text. [c] Denotes the corresponding residue numbering in the model of nBgt. [d] Parentheses denote those regions which are fully extended. [e] Square brackets denote those regions which have been rotated away from the rest of the structure as a rigid body.

the initial velocities for the simulated annealing and restrained molecular dynamics, or a combination of these two factors. In order to investigate this, two models (8 and 14) were used to generate another nine structures each, using different random number seeds for the assignment of initial velocities in each case. Figure 7B illustrates that the conformations of the final structures are determined at least as much by the random number seed as by the conformation of the starting model. In the following sections, we analyze these final structures in order to assess the persistence of specific global and local features. Such persistence within the set of calculated structures should provide some indication as to the probability that individual features contribute to the average solution structure.

The Dimer Interface. The nature of the interfaces in the 12 final structures formed by the corresponding β-strands involving Ser 55–Cys 59 in each of the dimer molecules was examined first. In particular, the existence of hydrogen bonds between Leu 56 and Cys 58 in one chain and Cys 58 and Leu 56, respectively, in the other chain was examined (see Figure 4). The respective donors and acceptors for these four hydrogen bonds adopt a relative orientation consistent with a hydrogen bond (as defined by the program QUANTA) in some (five), but not all, of the 12 final structures. In some cases, the lack of intermolecular restraints in the other part of the dimer interface (region Pro 47–Ser 51, see below) may influence the formation of hydrogen bonds in the region Ser 55–Cys 59. This, coupled with the similar topology of all 12 structures, leads us to conclude that these four interchain hydrogen bonds make some contribution to the stability of the dimer. The dimer interface may be further stabilized by hydrophobic interactions between the side chains of Leu 57 in the respective chains which are in close proximity in all 12 structures (Figure 8). Another possible contribution to the dimer interface is the region around Pro 47–Ser 51. In all but one structure, Phe 49 from one chain is in close proximity to Phe 49 from the other chain in the dimer. This suggests that a hydrophobic interaction between the two aromatic rings could contribute to the stability of the dimer; a schematic representation of the dimer interface involving such an interaction is shown in Figure 8. Interestingly, Phe 49 is conserved in all four known neuronal α-neurotoxins (Grant & Chiappinelli, 1985; Loring et al., 1986; Grant et al., 1988; Danse & Garnier, 1990), all of which form dimers in solution (Chiappinelli & Wolf, 1989), and is absent in all other α-neurotoxins [for a review, see Karlsson (1979)], most of which do not form dimers in solution. The dimer interface is illustrated in Figure 8.

The β-Sheet. The secondary structure of the final structures was analyzed using the molecular graphics program QUANTA. All 12 structures have an antiparallel β-sheet comprising residues Leu 23–Ala 25 and Leu 57–Ser 55. Additionally, this extends to include Cys 21–Phe 22 and Cys 59–Cys 58 in five of the structures and the two β-strands Gly 41–Val 37 and Phe 22–Gln 26 in 10 of the structures (in the two exceptions, the β-strand Gly 41–Val 37 is distorted).

The six-stranded antiparallel β-sheet is shown in Figure 6 and schematically in Figure 9. Figure 10 shows that those residues toward the center of the β-strands are better defined than those toward the edges. It is interesting that the structural

Table II: Properties of the Final Dimeric Structures[a]

	STRUCTURE													
	1[b]	2	3	4	5	6	7	8	9	10	11	12	13	14
NOE_{viol}[c]	8	10	4	6	26	0	38	2	8	6	2	2	6	6
E_{NOE}[d]	69	98	66	67	263	72	416	42	92	96	41	27	65	78
MAX_{VIOL}[e]	0.8	0.6	0.7	0.7	1.4	0.5	2.1	0.6	1.1	0.9	0.5	0.6	0.7	0.9
E_{VDW}[f]	−561	−575	−571	−576	−442	−572	−245	−587	−573	−527	−598	−562	−579	−585
GFOM[g]	0.117	0.127	0.120	0.116	0.137	0.120	0.162	0.112	0.119	0.122	0.117	0.116	0.110	0.115

[a] Although the force field used permitted a transition between the trans and cis conformations in the case of proline residues, the five proline residues adopted the trans conformation in all 14 structures. [b] The number of each structure refers to the 14 starting models following the refinement protocol described in the text. [c] NOE_{viol} is the number of NOE violations greater than 0.5 Å. [d] E_{NOE} is the energy associated with the NOE violations (kcal/mol). [e] MAX_{VIOL} is the maximum NOE violation (Å). [f] E_{VDW} is the Lennard-Jones energy of the structure (kcal/mol). Note that a van der Waals energy of −506 kcal/mol was observed for the dimeric crystal structure of αBgt after 1000 cycles of energy minimization. [g] GFOM is the geometric figure of merit.

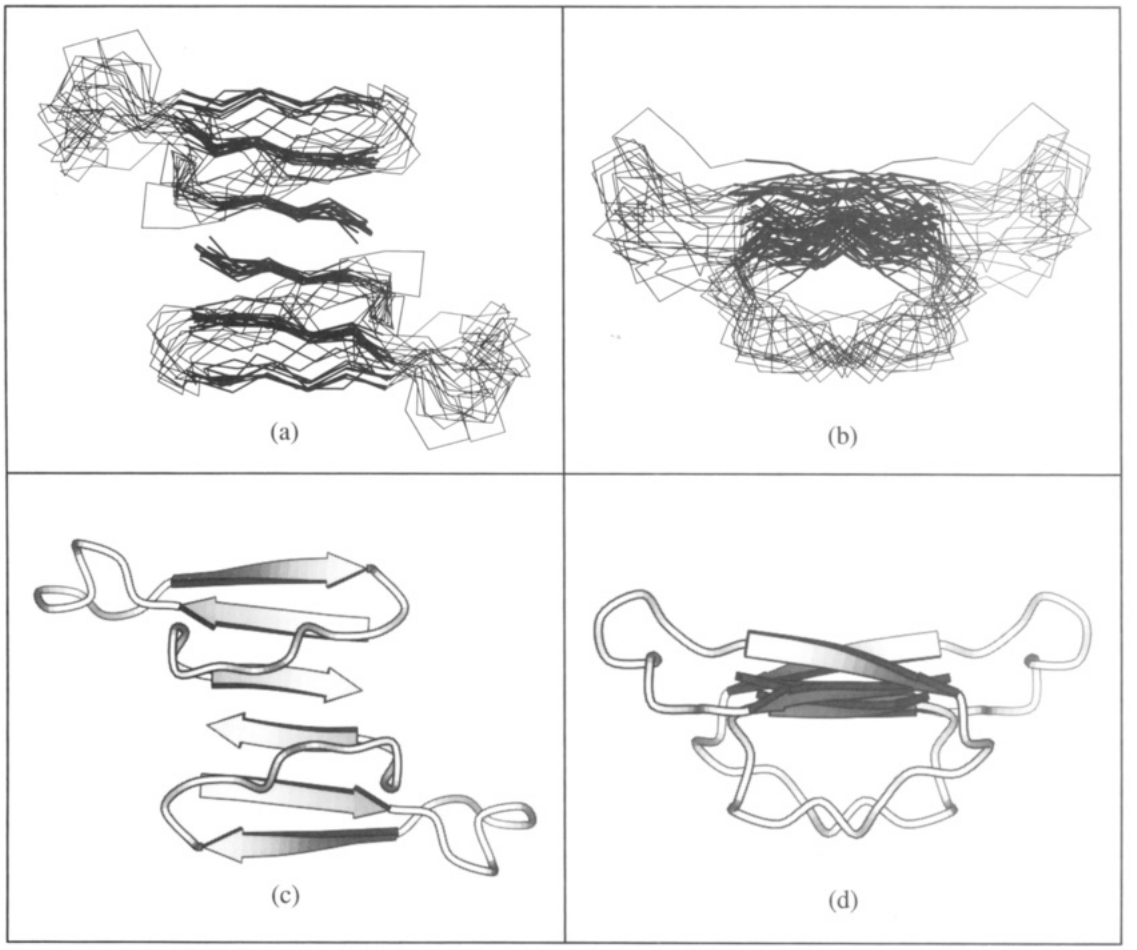

FIGURE 6: Topology of nBgt. Panels a and b show STRUCTURES 1–14, excluding 5 and 7 (see text), represented by their C_α atoms after being aligned on the β-sheet using the program MNYFIT (Sutcliffe et al., 1987a; the β-strands are denoted by thicker lines than the remainder of the structure). Panels c and d illustrate the overall fold of nBgt using a schematic representation. Panels a and c are views from above the β-sheet, and panels b and d are side views along the β-sheet. The poorly defined N-terminal (Arg 1–Ile 20) and C-terminal (Thr 60–His 66) regions have been removed for clarity.

deviations observed over all six β-strands in the dimers are very similar to those for the corresponding triple β-strands of a single chain (Figure 10); the C_α RMS from the mean of the β-sheets is 1.3 ± 0.5 Å for the intermolecular β-sheet and 0.8 ± 0.2 Å for the intramolecular β-sheet. Although the position of corresponding β-strands is on the whole similar across the 12 structures, a few exceptions were observed (Figure 6). For example, in one structure all three β-strands deviate from those in the other structures, and, in three other structures, the position of β-strand Gly 41–Val 37 differs from that observed in the other eight structures. These deviations appear to arise for a variety of reasons. In one case, the β-strand Cys 59–Ser 55 is distorted, and this deviation propogates through the β-sheet. In another two cases, the β-strand Gly 41–Val 37 is distorted; and, in another, the angle between the two intramolecular β-sheets is greater than in the other 11 structures.

The N- and C-Terminal Regions. The N-terminal region (Arg 1–Ile 20) has a fairly well-defined conformation (Figure 11 and insert in Figure 10) when considered in isolation but is not well defined with respect to the rest of the structure (Figure 10). This is the origin of the large deviation of C_α positions in this region (Figure 10) and can be attributed to the relatively small number of long-range NOEs linking this region to the main β-sheet region of the molecule (Figure 2). The final ensemble of structures shows, however, that this loop is always situated on the opposite side of the β-sheet from that region of the structure thought to bind nAChR (the concave surface; Tsetlin et al., 1982; Johnson & Taylor, 1983; Hawrot & Basus, 1991). This is in contrast to the position defined for the N-terminal region in the structure of αBgt, where it is situated lateral to the β-sheet (Love & Stroud, 1986; V. J. Basus, private communication). Whether this difference is real or whether it is a consequence of the limited NMR data presently available for this region of nBgt remains to be established.

Like the N-terminal region, the C-terminal region (Thr 60–His 66) is not well defined (Figure 10). In this case, the lack of definition persists whether these residues are considered in isolation or as part of the whole molecule. It is interesting in this respect that the only long-range NOEs involving this region are with the N-terminal region.

Comparison of the Structures with the Properties of Other Neurotoxins. In the light of the structural analysis presented above, we are now in a position to discuss the significance of some of the individual residues in related α-neurotoxins. Ser 8 corresponds to a highly conserved threonine in long toxins. However, the significance of this amino acid change on the tertiary structure is impossible to determine due to the poor definition of the N-terminal region with respect to the remainder of the molecule. Phe 22 corresponds to a highly conserved tyrosine in other toxins. This phenylalanine is oriented similarily in all 12 structures of nBgt and indeed is similar to the orientation of the corresponding tyrosine in the known crystal structures of related α-neurotoxins (αBgt, 2ABX in the Brookhaven Protein Data Bank; α-cobratoxin, 1CTX; and erabutoxin, 3EBX). In αBgt, the hydroxyl group of this

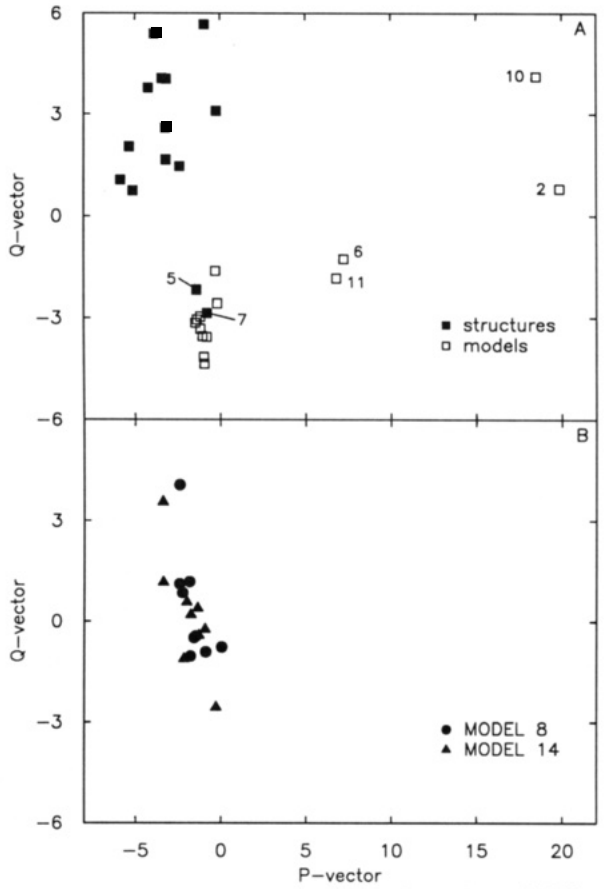

FIGURE 7: Relationship between the conformations of different structures. In this figure, the smaller the distance between two structures, the more similar their overall conformation. (a) Relationship between the starting models (□) and the final structures (■) (MODELS 2, 6, 10, and 11 are labeled, as these deviate significantly in overall conformation from the remaining models as illustrated by Figure 5). (b) Relationship between the 10 structures generated from MODEL 8 (●; STRUCTURE 8 plus nine others), MODEL 14 (▲; STRUCTURE 14 plus nine others), the starting models (included in the calculation and occupying closely similar positions to those in panel a but not shown), and the final structures (included in the calculation and occupying closely similar positions to those in panel a but not shown). This figure is a two-dimensional projection of the multidimensional space defined by the eigenvalues and eigenvectors derived from the matrix of pairwise RMS values [see, for example, Levitt (1983)].

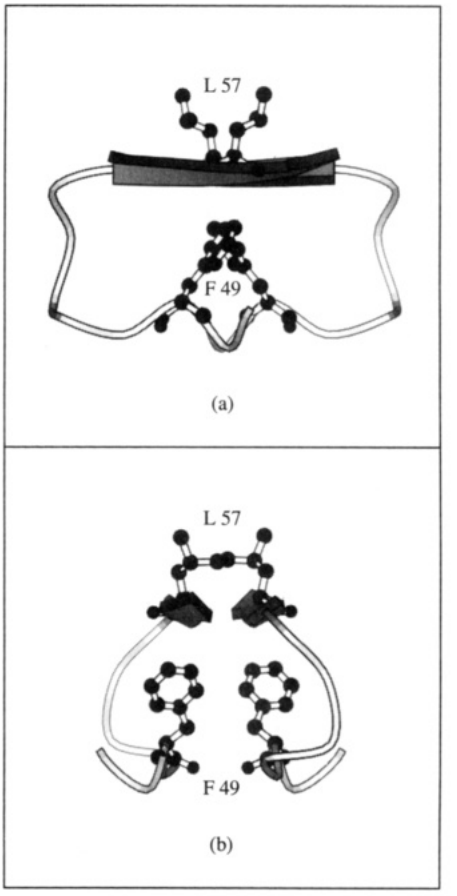

FIGURE 8: Dimer interface illustrated schematically. Phe 49 and Leu 57 are shown, together with a schematic representation of the backbone for the region Pro 47–Cys 59. (a) View from the side of the dimer interface; (b) view along the dimer interface.

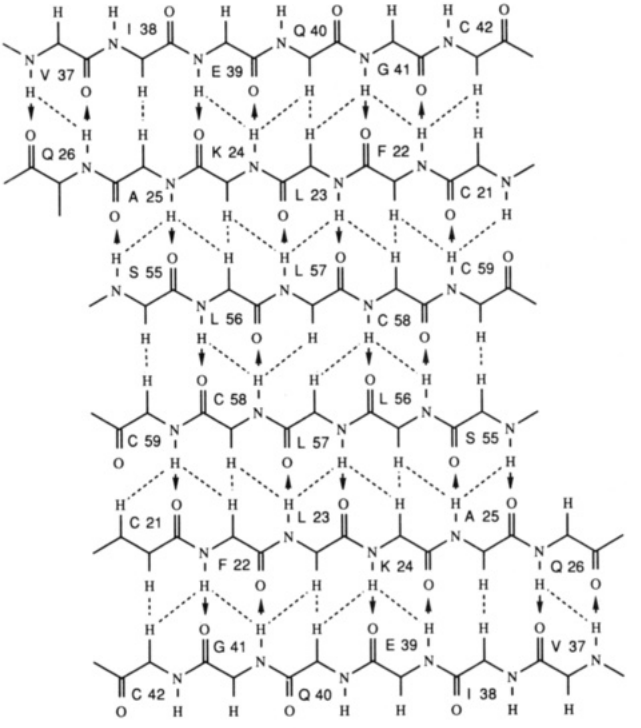

FIGURE 9: Schematic representation of the intermolecular β-sheet. Arrows represent hydrogen bonds, and dashed lines represent NOEs.

tyrosine (Tyr 24) forms a hydrogen bond with Glu 41. Interestingly, no equivalent hydrogen bond is formed in either α-cobratoxin or erabutoxin despite the fact that the tyrosine residue and an appropriate hydrogen-bonding acceptor are present. Although a hydrogen bond cannot be formed between Phe 22 and Glu 39 in nBgt, in all but one structure Glu 39 is oriented similarly to Glu 41 in αBgt, Glu 38 in erabutoxin, and Asp 38 in α-cobratoxin. It is thought that this conserved tyrosine (Tyr 24 in αBgt) interacts with the nAChR (Hawrot & Basus, 1991). Although Phe 22 is unable to interact within the nAChR via a hydrogen bond, it may well interact via its aromatic ring.

In all 12 nBgt structures, Lys 24 is oriented similarly to Lys 26 in αBgt, Lys 27 in erabutoxin, and Lys 23 in α-cobratoxin. It has been proposed that this residue interacts with negatively charged groups on the nAChR (Tsetlin et al., 1982; Johnson & Taylor, 1983), although this has been questioned (Johnson et al., 1990). Gln 26 is an important residue as it corresponds to Trp 28 in αBgt. Trp 28 is in a region of the αBgt structure found to be in a different conformation in the solution and crystal structures (Basus et al., 1988). Like the corresponding tryptophan residue in α-cobratoxin and erabutoxin, Gln 26 is positioned on the concave side of the molecule in all 12 nBgt structures, in accord with the solution structure of αBgt. In

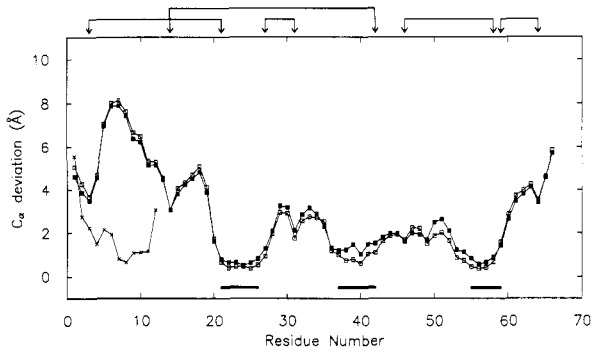

FIGURE 10: Average C_α distance between the 12 final structures and the average structure. This was computed after first aligning the 12 structures using the six β-strands and the program MNYFIT (■) and repeated using the three intramolecular β-strands (□). Also shown is the average C_α distance in the N-terminal region obtained by aligning residues 2–11 (×). The thick solid bars represent the position in the structure of the β-strands, and the arrows at the top of the figure indicate the positions of the disulfide bonds.

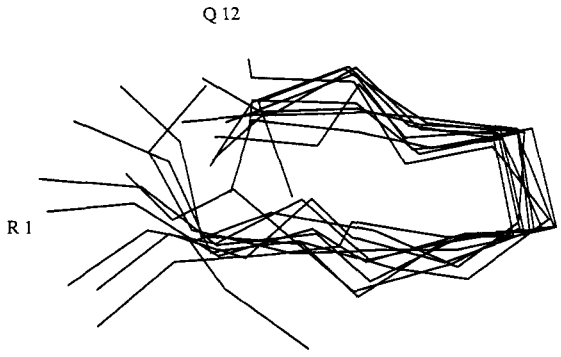

FIGURE 11: Topology of the N-terminal region of nBgt. Residues 2–11 were aligned by MNYFIT and residues 1–12 are shown by their C_α atoms.

an attempt to explain the difference in specificity for the nAChR between nBgt and αBgt, erabutoxin and α-cobratoxin, an investigation was made of both the distribution of surface charge and the nature of the electrostatic surface in the region of the concave surface. However, no significant differences could be found between the properties of nBgt and the other structures.

In conclusion, nBgt consists of a triple-stranded antiparallel β-sheet with a closely similar topology to that found for other α-neurotoxins. In contrast to other α-neurotoxins for which the structure is known, however, it forms a dimer in solution. The dimer interface is stabilized by hydrogen bonds between Leu 56 and Cys 58 in one chain and Cys 58 and Leu 56, respectively, in the second chain, thereby forming a six-stranded antiparallel intermolecular β-sheet similar to that seen in the crystal structure of αBgt. The possibility that in nBgt this interface is stabilized by aromatic–aromatic interactions between Phe 49 in the two chains may be the origin of additional dimer stability. These results illustrate the close similarity of nBgt to other α-neurotoxins and suggest that the difference in specificity arises either from the dimerization itself or from detailed differences in the surface properties, rather than large differences in the fold of individual chains.

ADDED IN PROOF

In the monomeric structures described above, the region Pro 47–Ser 51 was lateral to the β-sheet. We suggested that this could be a shortcoming of the metric matrix distance geometry algorithm used and have substantiated this by subsequent work (M. J. Sutcliffe, unpublished results). A series of monomeric structures was generated using DIANA [Guntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517–530]. The region Pro 47–Ser 51 was oriented with respect to the β-sheet as in the dimeric structure in all of the structures with acceptable target function values.

ACKNOWLEDGMENTS

We thank Michelle Bamberger, Ralph Loring, and Emory Braswell for their invaluable contributions to this project. We are grateful to Per Kraulis for providing MolScript. M.J.S. is a member of the Leicester Centre for Molecular Recognition.

SUPPLEMENTARY MATERIAL AVAILABLE

One table listing the NOE restraints used and a second listing the ϕ angle restraints used (22 pages). Ordering information is given on any current masthead page.

REFERENCES

Arrowsmith, C. H., Pachter, R., Altman, R. B., Iyer, S. B., & Jardetzky, O. (1990) *Biochemistry 29*, 6332–6341.
Basus, V. J., Billeter, M., Love, R. A., Stroud, R. M., & Kuntz, I. (1988) *Biochemistry 27*, 2763–2771.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanovichi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
Blundell, T. L., Carney, D., Gardner, S., Hayes, F., Howlin, B., Hubbard, T., Overington, J., Singh, D. A., Sibanda, B. L., & Sutcliffe, M. (1988) *Eur. J. Biochem. 172*, 513–520.
Brunger, A. T., Kurian, J., & Karplus, M. (1987) *Science 235*, 458–460.
Chiappinelli, V. A., & Lee, J. C. (1985) *J. Biol. Chem. 260*, 6182–6186.
Chiappinelli, V. A., & Wolf, K. M. (1989) *Biochemistry 28*, 8543–8547.
Danse, J.-M., & Garnier, J.-M. (1990) *Nucleic Acids Res. 18*, 1050.
Endo, T., Inagaki, F., Hayashi, K., & Miyazawa, T. (1981) *Eur. J. Biochem. 120*, 117–124.
Galzi, J.-L., Revah, F., Bessis, A., & Changeux, J.-P. (1991) *Annu. Rev. Pharmacol. 31*, 37–72.
Grant, G. A., & Chiappinelli, V. A. (1985) *Biochemistry 24*, 1532–1537.
Grant, G. A., Frazier, M. W., & Chiappinelli, V. A. (1988) *Biochemistry 27*, 3794–3798.
Hawrot, E., & Basus, V. J. (1991) *Biophys. J. 59*, 189a.
Inagaki, F., Hider, R. C., Hodges, S. J., & Drake, A. F. (1985) *J. Mol. Biol. 183*, 575–590.
Jian-Sheng, J., & Hubbard, R. E. (1990) *Proceedings of the 2nd York Meeting on Generation of 3-D Structures from Distance Information*, York, England.
Johnson, D. A., & Taylor, P. (1983) *J. Biol. Chem. 257*, 5632–5636.
Johnson, D. A., Cushman, R., & Malekzadeh, R. (1990) *J. Biol. Chem. 265*, 7360–7368.
Karlsson, E. (1979) in *Handbook of Experimental Pharmacology* (Lee, C. Y., Ed.) pp 159–212, Springer, Berlin.
Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Labhardt, A. M., Hunziker-Kwik, E. H., & Wüthrich, K. (1988) *Eur. J. Biochem. 177*, 295–305.
Levitt, M. (1983) *J. Mol. Biol. 168*, 621–657.
Loring, R. H., Andrews, D., Lane, W., & Zigmond, R. E. (1986) *Brain Res. 385*, 30–37.
Love, R. A., & Stroud, R. M. (1986) *Protein Eng. 1*, 37–46.

Low, B. W., Preston, H. S., Sato, H., Rosen, L., Searl, J. E., Rudko, A. D., & Richardson, J. S. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2991-2994.
Metzler, W., Hare, D. B., & Pardi, A. (1989) *Biochemistry 28*, 7045-7052.
Oswald, R. E., Sutcliffe, M. J., Bamberger, M., Loring, R. H., Braswell, E., & Dobson, C. M. (1991) *Biochemistry 30*, 4901-4909.
Sutcliffe, M. J. (1989) in *Protein Engineering, Crystallography and Computer Graphics* (Bolognesi, M., Ed.) pp 19-27, Clas International, Brescia, Italy.
Sutcliffe, M. J., & Dobson, C. M. (1991) *Proteins 10*, 117-129.
Sutcliffe, M. J., Haneef, I., Carney, D., & Blundell, T. L. (1987a) *Protein Eng. 1*, 377-384.
Sutcliffe, M. J., Hayes, F. R. F., & Blundell, T. L. (1987b) *Protein Eng. 1*, 385-392.
Tsernoglou, D., & Petsko, G. A. (1976) *FEBS Lett. 68*, 1-4.
Tsetlin, V. I., Karlsson, E., Utkin, Y. N., Pluzhnikov, K. A., Arseniev, A. S., Surin, A. M., Kandakov, V. V., Bystrov, V. F., Ivanov, V. T., & Ovhinnikov, Y. A. (1982) *Toxicon 20*, 83-93.
Walkinshaw, M. D., Saenger, W., & Maelicke, A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2400-2404.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.

Nucleotide Binding and GTP Hydrolysis by Elongation Factor Tu from *Thermus thermophilus* As Monitored by Proton NMR†

Stefan Limmer,‡ Christian O. A. Reiser, Norbert K. Schirmer, Norbert W. Grillenbeck, and Mathias Sprinzl*
Laboratorium für Biochemie, Universität Bayreuth, Postfach 10 12 51, W-8580 Bayreuth, FRG
Received September 17, 1991; Revised Manuscript Received November 27, 1991

ABSTRACT: Proton NMR experiments of the GTP/GDP-binding protein EF-Tu from the extremely thermophilic bacterium *Thermus thermophilus* HB8 in H_2O have been performed paying special attention to the resonances in the downfield region (below 10 ppm). Most of these downfield signals are due to hydrogen bonds formed between the protein and the bound nucleotide. However, three downfield resonances appear even in the nucleotide-free EF-Tu. The middle and C-terminal domain (domain II/III) of EF-Tu lacking the GTP/GDP-binding domain gives rise to an NMR spectrum that hints at a well-structured protein. In contrast to native EF-Tu, the domain II/III spectrum contains no resonances in the downfield region. Several downfield resonances can be used as a fingerprint to trace hydrolysis of protein-bound GTP and temperature effects on the EF-Tu·GDP spectra. NMR studies of the binding of guanosine nucleotide analogues (GMPPNP, GMPPCP) to nucleotide-free EF-Tu have been carried out. The downfield resonances of these complexes differ from the spectrum of EF-Tu·GTP. Protected and photolabile caged GTP was bound to EF-Tu, and NMR spectra before and after photolysis were recorded. The progress of the GTP hydrolysis could be monitored using this method. The downfield resonances have been tentatively assigned taking into account the known structural and biochemical aspects of EF-Tu nucleotide-binding site.

The elongation factor Tu (EF-Tu)[1] is a member of the group of GTP/GDP-binding proteins. This group includes translation, initiation, and elongation factors (Kaziro, 1978), signal-transducing G-proteins (Gilman, 1987), proteins involved in protein transport and secretion (Balch, 1990; Rapoport, 1990), and the *ras* gene family (Barbacid, 1987). The hydrolysis of the bound GTP to GDP and the associated conformational change lead to the deactivation of the protein (signal "off"), whereas the binding of GTP leads to an active conformation (signal "on") which enables the protein to interact with the respective effector.

The elongation factor Tu in its GTP conformation promotes the binding of aminoacyl-tRNA (aa-tRNA) to the bacterial ribosome during polypeptide chain elongation. The nucleotide exchange factor, EF-Ts, facilitates the dissociation of GDP from EF-Tu·GDP and subsequent binding of GTP. EF-Tu·GTP can then interact with aa-tRNA, and the thus formed EF-Tu·GTP·aa-tRNA complex binds to the mRNA-programed ribosome. Following GTP hydrolysis, triggered by association with the ribosome, EF-Tu·GDP dissociates from aa-tRNA and ribosome. This functional cycle of EF-Tu resembles that of other GTP/GDP-binding proteins (Bourne et al., 1990) and is important for kinetic control of the fidelity of translation (Thompson & Karim, 1982; Thompson et al., 1986).

The tertiary structure of bacterial EF-Tu was determined by X-ray analysis of the protein crystals originating from *Escherichia coli* (Jurnak, 1985; la Cour et al., 1985). EF-Tu is composed of three structural domains. Domain I, also referred to as the G domain, is the GTP/GDP-binding domain which shares sequence (Bourne et al., 1991) and structural homology with other GTP/GDP-binding proteins, including the *ras* gene product p21 (Jurnak, 1990a,b). Little is known about the structure of the domain II and the C-terminal do-

† This work was supported by the Deutsche Forschungsgemeinschaft (SFB 213/D5) and Fonds der Chemischen Industrie.
* To whom correspondence should be addressed.
‡ On leave of absence from Sektion Physik, Universität Leipzig, FRG.

[1] Abbreviations: DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; EF-Tu, elongation factor Tu; $EF\text{-}Tu_f$, nucleotide-free EF-Tu; GTP and GDP, guanosine 5′-triphosphate and guanosine 5′-diphosphate; GMPPNP, guanosine 5′-(β,γ-imido)triphosphate, GMPPCP, guanosine 5′-(β,γ-methylene)triphosphate; caged GTP, P^3-1-(2-nitrophenyl)ethylguanosine 5′-triphosphate; FID, free induction decay; HMQC, heteronuclear multiple-quantum coherence.

0006-2960/92/0431-2970$03.00/0