Transfer of a β-Hairpin from the Functional Site of Snake Curaremimetic Toxins to the α/β Scaffold of Scorpion Toxins: Three-Dimensional Solution Structure of the Chimeric Protein[†,‡]

Sophie Zinn-Justin,* Marc Guenneugues, Eugenia Drakopoulou, Bernard Gilquin, Claudio Vita, and André Ménez

CEA, Département d'Ingénierie et d'Etude des Protéines, CE Saclay, 91191 Gif-sur-Yvette Cedex, France

Received February 26, 1996; Revised Manuscript Received April 23, 1996[⊗]

ABSTRACT: The α/β scorpion fold is shared by scorpion toxins, insect defensins, and plant thionins. This small and functionally versatile template contains an α-helix and a triple β-sheet linked by three disulfide bridges. With the view to introduce novel functional centers within this fold, we replaced the sequence (the cysteines and glycines excepted) of the original β-hairpin of a scorpion toxin by the sequence of a β-hairpin that forms part of the site by which snake neurotoxins bind to nicotinic acetylcholine receptors (AcChOR). The resulting chimeric protein, synthesized by chemical means, binds to AcChOR, though with a lower affinity than the snake toxins [Drakopoulou, E., Zinn-Justin, S., Guenneugues, M., Gilquin, B., Ménez, A., & Vita, C. (1996) *J. Biol. Chem. 271*, 11979−11987]. The work described in this paper is an attempt to clarify the structural consequences associated with the transfer of the β-hairpin. We report the determination of the three-dimensional solution structure of the chimeric protein by proton NMR spectroscopy and molecular dynamics calculations. Comparison of the structure of the chimera with those of the scorpion α/β toxin and of the snake neurotoxin shows that (i) the new protein folds as an α/β motif and (ii) the β-hairpins of the chimera and of the curaremimetic toxin adopt a similar conformation. A closer inspection of the differences between the structures of the original and transferred β-hairpins allows rationalization of the biological properties of the chimera.

Engineering of the functional sites of proteins is now becoming a reality by virtue of the considerable potentialities offered by recombinant technologies and chemical synthesis. However, due to the complexity of the structural and dynamical events that are associated with the expression of biological activities, rationally conducted manipulations of functional sites remain a difficult exercise.

Among the successful attempts of rational handling of protein functional sites are found examples of reconstitutions, or "transfers", of protein sequences in new structural contexts. In this frame, particular attention has been focused on protein loops which form continuous solvent-exposed surfaces and which are frequently involved in molecular recognition and biological functions (Kuntz, 1972; Rose et al., 1985). The conformation of a loop in a protein is determined by both its amino acid sequence (Kabsch & Sander, 1984; Argos, 1987) and the geometry of its flanking residues (Fine et al., 1986; Moult & James, 1986; Chothia et al., 1986). Therefore, transfer of a loop into a new structural context is anticipated not to perturb the structure of the invited sequence, provided the structure of the host protein is adequately chosen. Several studies report transfers of loops and β-turns (Verhoeyen et al., 1988; Harper & Vallee, 1989; Hynes et al., 1989; Allemann et al., 1991; Toma et al., 1991; Wolfson et al., 1991; Braxton & Wells, 1992; Hedstrom et al., 1992; Eijsink et al., 1993; Vuillemer & Fersht, 1994; Raines et al., 1995); however, these operations were all performed at the surface of large proteins, distant from the hydrophobic core of the host protein, and a detailed structural investigation allowing an evaluation of the extent of success of the transfer operation was reported in only a few cases (Hynes et al., 1989; Wolfson et al., 1991).

The goal of the present study was to investigate the structural implications associated with the transfer of a functional β-hairpin from a 61-residue protein where the β-hairpin is sandwiched between two loops rich in β-pleated sheet toward a smaller 31-residue protein containing the α/β scorpion fold where the β-hairpin interacts predominantly with a short α-helix (Figure 1).

The α/β scorpion fold possesses an α-helix and a triple-stranded β-sheet, whose respective positions are stabilized by three disulfide bridges (Bontems et al., 1991a). This scaffold offers a number of advantages which include (i) a small size of about 37 amino acids, making it readily amenable to chemical synthesis and NMR[1] structure determination, (ii) a high stability, due in particular to the presence of three disulfide bridges, and (iii) a high degree of functional permissivity, the fold being adopted by all known scorpion toxins (Bontems et al., 1991b), insect defensins (Hanzawa et al., 1990; Bonmatin et al., 1992), and plant thionins (Bruix et al., 1993, 1995). As a result of these advantages, the α/β scaffold was proposed to constitute an appropriate template to introduce new functional sites (Ménez et al., 1992), and indeed, the metal binding site of human carbonic anhydrase

[†] M. G. was supported by grants from IFSBM.

[‡] The atomic coordinates of the chimeric protein have been deposited in the Brookhaven Protein Data Bank (file name 1CMR).

* To whom correspondence should be addressed. E-mail: szinn@balthazar.saclay.cea.fr. FAX: (33) 1 69089071.

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1996.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D, two dimensional; COSY, correlated spectroscopy; DQF-COSY, double-quantum-filtered correlated spectroscopy; TOCSY, total correlated spectroscopy; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; RMSD, root mean square deviation.

S0006-2960(96)00466-7 CCC: $12.00

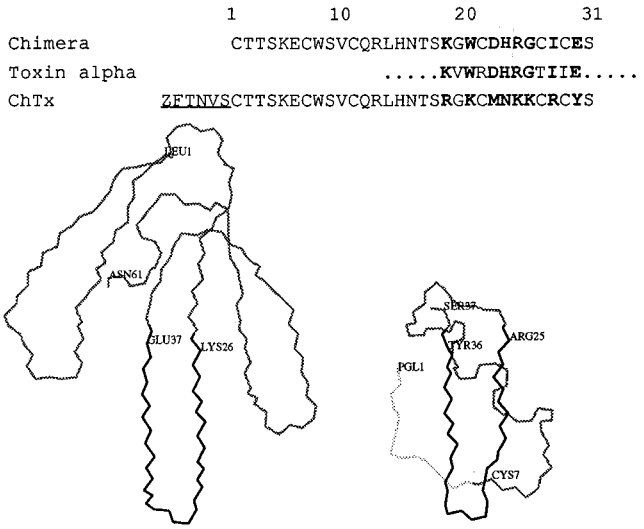

FIGURE 1: Three-dimensional solution structures of toxin α and charybdotoxin. The native and host β-hairpin structures are displayed in black and the part of charybdotoxin deleted in the chimera is displayed in light gray. Alignment of the sequences of toxin α (Toxin alpha), charybdotoxin (ChTx), and the chimeric protein (Chimera) is shown; the mutated positions are displayed in bold type, and the charybdotoxin sequence deleted in the chimera is underlined.

has been successfully reconstituted at the surface of a scorpion toxin (Vita et al., 1995).

More recently, Drakopoulou et al. (1996) proceeded to the transfer of one of the three loops of snake neurotoxins onto the α/β scorpion scaffold. This loop forms the central part of the site by which neurotoxins bind to nicotinic acetylcholine receptors with high affinity and specificity, as deduced from mutational analyses (Hervé et al., 1992; Pillet et al., 1993; Trémeau et al., 1995). It shows six of the ten functionally identified side chains on its concave side. These six side chains are highly conserved in all related snake toxins. The central loop was thus proposed to constitute the "evolutionary preserved functional core" of snake neurotoxins. We aimed to reconstitute part of an identified functional site on a smaller scaffold, which would stabilize the native three-dimensional structure of the functional sequence. Partial reconstitution of a predetermined site on the smallest possible scaffold is anticipated to offer new potentialities for ultimately engineering an appropriate drug with an adequate bioavailability profile, as recently described (Li et al., 1995).

The β-hairpin that was considered in the present transfer operation is originally from toxin α from the snake *Naja nigricollis* (Zinn-Justin et al., 1992). The host protein is charybdotoxin from the scorpion *Leiurus quinquestriatus hebraeus* (Bontems et al., 1991a, 1992). The design of the sequence to be introduced in charybdotoxin had to comply with the presence of the three disulfide bridges of charybdotoxin that cannot be removed without perturbing the stability of the scaffold (Vita et al., 1994). However, the cysteines associated with the β-sheet of the scaffold all point toward the buried face of the β-sheet, whereas the functional site of snake neurotoxins should be located on the opposite and hence accessible face. Therefore, the transfer operation has mostly consisted in introducing the functional residues of the toxin α β-hairpin into appropriate positions in the scorpion toxin β-hairpin and in preserving the residues of the buried side of the charybdotoxin β-hairpin, including three cysteines and one glycine (Figure 1). With the view to further decrease the size of the host scaffold and to increase the solvent accessibility for the grafted side chains, the first six N-terminal residues of charybdotoxin were deleted. This last operation was shown to have no major influence on the three-dimensional structure of charybdotoxin (Vita et al., 1993). The resulting chimeric protein displayed binding properties to the nicotinic acetylcholine receptor with a K_d value in the micromolar range, while neither the host structure (charybdotoxin) nor the denatured chimera show any affinity for the same receptor. The affinity of the chimeric protein is 5 orders of magnitude lower compared to that of the original snake toxin. In order to clarify these results in structural terms, we have now determined the three-dimensional structure of the chimeric protein by NMR and molecular modeling. The three following questions have been addressed. First, does the chimeric sequence fold into a globular protein? Second, does the chimeric protein adopt the anticipated α/β scaffold? Third, how do the structures of the β-hairpin compare in the visited scaffold and in the snake toxin?

MATERIALS AND METHODS

Sample Preparation. Synthesis of the chimeric protein is described elsewhere (Drakopoulou et al., 1996). The stability of the protein at pH 3.5 was studied by circular dichroism and HPLC in different temperature conditions. Incubating the protein at 45 °C induced no more than 5% degradation after 30 days.

For the NMR experiments, 9 mg of the chimeric protein was dissolved in 0.4 mL of solvent, so that the final concentration was 7 mM. The solvents used were either a mixture of 95% (v/v) H_2O and 5% (v/v) D_2O at pH 3.5 or 100% D_2O at pD 3.5. The pH (pD) was adjusted by addition of microliter amounts of dilute HCl (DCl). For the preparation of samples in D_2O, the amide protons were previously fully exchanged with deuterium by keeping the solution at 20 °C for 5 days. Chemical shifts were measured referenced to internal 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate.

NMR Experiments. Proton 2D COSY (Aue et al., 1976), DQF-COSY (Rance et al., 1983), TOCSY (Braunschweiler & Ernst, 1983), and NOESY (Kumar et al., 1980) spectra were recorded at 600 MHz on a Bruker AMX600 spectrometer. Assignment of the resonance frequencies was achieved on the basis of experiments carried out at two temperatures (20 and 30 °C). A WALTZ 16 composite pulse sequence (Shaka et al., 1983) was used for the isotropic mixing (80 ms) in the TOCSY experiments. NOESY spectra were recorded at 20 °C with mixing times of 50, 100, and 200 ms in H_2O and 80, 160, and 250 ms in D_2O; a preparation period of 2 s was used during the NOESY experiments.

The water signal was suppressed by low-power irradiation of the solvent resonance at all times except during t_1 and t_2 and by using the "jump-and-return" sequence (Plateau & Guéron, 1982). All the experiments were performed in hypercomplex mode. The spectra were recorded with 512 $t_1 \times 1024$ t_2 points [1024 $t_1 \times 4096$ t_2 points for the DQF-COSY] and with a spectral width of 7812.5 Hz at 600 MHz. Data processing was carried out using the FELIX program (Biosym Technologies, San Diego) on a SUN SPARC station 1.

Experimental Restraints. The volumes of the cross-peaks from NOESY spectra recorded at 20 °C were integrated after

zero completion of the matrices, multiplication by a cosine bell function in both dimensions, Fourier transformation, and base-plane correction using fifth-order polynomials. For each correlation peak, a build-up curve was constructed going through the origin, and the points corresponding to the volumes were measured at the three mixing times. This curve was then fitted by a second-order polynomial. The coefficient of the first-order term of the polynomial was taken as the build-up rate of the corresponding NOE.

Calibration of these dipolar correlation rates was achieved using a rigid body model of the protein: we assumed that $\sigma(\text{NOE}) = k/r^6$. The k constant was calculated so that the distance values derived from the (NH,CαH) dipolar correlation rates were in the 2.2−3.5 Å allowed range. We checked that the distances between the protons of aromatic rings calculated with such a calibration were consistent with the known distance values. Calibration of the dipolar correlation rates observed in D_2O was obtained by scaling the rates of 30 cross-peaks measured in both H_2O and D_2O. When a distance was calculated in both H_2O and D_2O, only the value coming from the D_2O spectra was retained, because of the higher signal/noise ratio of experiments recorded in this solvent.

In order to evaluate the error from measurement of NOE build-up rates, a set of 40 cross-peaks was integrated after two or three distinct spectrum processings. The root mean square deviation of the volume measurements was 100%, 50%, and 40% for cross-peaks observed in the 50, 100, and 200 ms mixing time NOESY spectra, respectively. This led to an error of about 50% on the rate of cross-relaxation, which corresponds to an error of 8% on the distances. This error and others, due to NMR acquisition, fitting of the build-up curve by a second-order polynomial, and use of a rigid body model, were taken into account by allowing a range of ±25% to the distance restraints.

Coupling constants were measured with a resolution of 1.9 Hz/point, by evaluating the antiphase cross-peak separation on COSY and DQF-COSY spectra. ϕ angle restraints were derived from the $^3J_{\text{HN-H}\alpha}$ constants on the basis of the empirical Karplus relation (Pardi et al., 1984). A restraint of $-120° \pm 45°$ was used when $^3J_{\text{HN-H}\alpha}$ values were higher than 8 Hz, and a restraint of $70° \pm 45°$ was used for $^3J_{\text{HN-H}\alpha}$ values lower than 6 Hz in the α-helical region of the protein. χ_1 angle values were determined using the method of Hyberts et al. (1987), and a range of ±45° was taken into account to define the corresponding restraints.

Protocol for Structure Calculations. A first set of 50 structures was generated using the DIANA program (Güntert et al., 1991). The 18 structures corresponding to the lower target functions were further submitted to a refinement protocol using the X-PLOR program (Brünger et al., 1992).

The protocol was derived from that described by Gippert et al. (1990) and was similar to that already used in our laboratory (Bontems et al., 1992; Zinn-Justin et al., 1992; Segalas et al., 1995). A force field adapted to NMR structure calculation (file parallhdg.pro and topallhdg.pro in X-PLOR 3.1) was used in the simulated annealing part of the refinement protocol. The final minimization was carried out in the force field derived from CHARMM22 (file parallh22x.pro and topall22x.pro in X-PLOR 3.1).

Comparison of the Chimeric Protein Structure with the Structures of Charybdotoxin and Toxin α. The structures of charybdotoxin and toxin α were loaded from the Protein Data Bank (Bernstein et al., 1977), reference 2CRD and 1NEA, respectively. All these structures were displayed, analyzed, and compared on a Silicon Graphics 4D/25 station using the SYBYL package (Tripos Associates, Inc.).

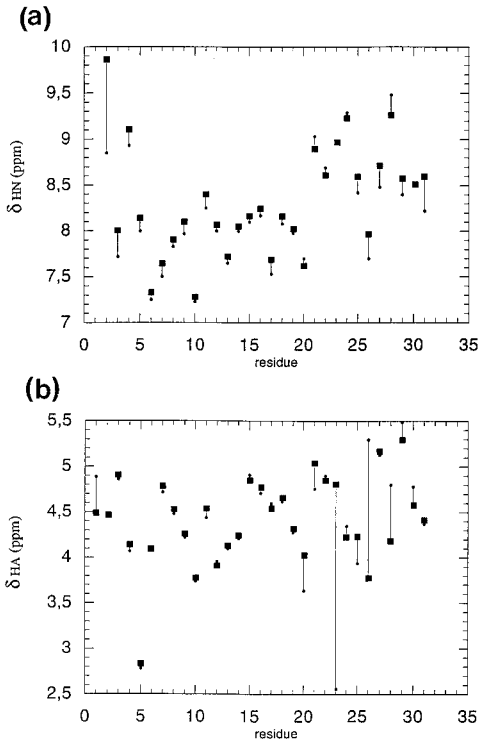

FIGURE 2: Comparison of (a) HN and (b) $C_\alpha H$ chemical shifts of the chimeric protein (■) with the corresponding chemical shifts in charybdotoxin (●). Note that the chemical shifts have been measured at 20 °C in the chimera and at 45 °C in charybdotoxin. Thus, differences of less than 0.25 ppm between the two proteins should not be taken into account. The chemical shifts of charybdotoxin are from Bontems et al. (1991).

RESULTS

Resonance Assignment. The complete sequence-specific assignment of the chimeric protein was achieved according to the standard method developed by Wüthrich and coworkers (Wüthrich, 1986). Comparison of the backbone proton chemical shifts of the new protein with those of charybdotoxin (Figure 2) shows that differences greater than 0.25 ppm are observed for a few residues located at the N-terminus (C1, T2, T3) or in the mutated part of the sequence (G20, W21, D23, R25, G26, I28, S31). These differences are probably due to both the deletion of the six N-terminal residues as compared to charybdotoxin and the mutations, in particular, the replacement of a tryptophan for a lysine at position 21. Besides these differences, the chemical shifts in the two proteins remain close enough to suggest that the overall folds of the chimeric and native proteins are similar.

Experimental Data. The secondary structure of the chimeric protein was established by analyzing the backbone proton NOEs and the $^3J_{\text{HN-H}\alpha}$ coupling constants. The pattern of strong sequential and long-range d_{NN} connectivities for residues 5−14 indicates the presence of a helical conformation. This is confirmed by coupling constants lower than 7 Hz for residues 5−6 and 8−14. Furthermore, the presence of strong sequential $d_{\alpha\text{N}}$ NOEs, the typical pattern of long-range $d_{\alpha\alpha}$ (Figure 3), d_{NN}, and $d_{\alpha\text{N}}$ connectivities,

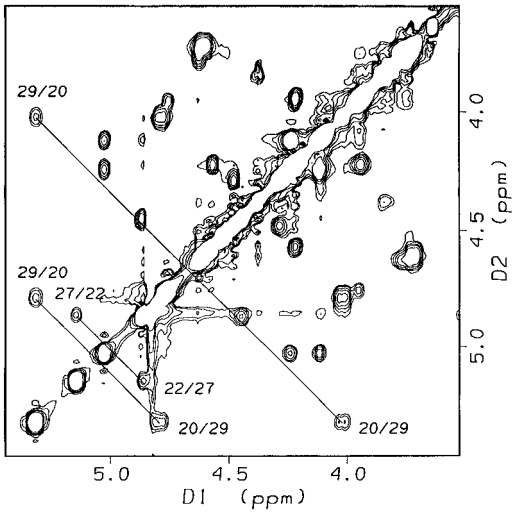

FIGURE 3: Part of a NOESY spectrum recorded in D_2O at 20 °C and pH 3.5 showing the long-range $d_{\alpha\alpha}$ connectivities characteristic of the β-sheet.

and the large values of the coupling constants indicate that residues 19−22 and 27−30 are involved in a two-stranded antiparallel β-sheet. Thus, the secondary structure elements of the chimera are found at the same positions as in charybdotoxin (Bontems et al., 1991a).

A total of 450 cross-relaxation rates were measured (290 in H_2O and 190 in D_2O; 30 common effects were used for the calibration). Approximately 30 of these rates correspond to fixed distances, and an additional 80 rates correspond to distances greater than 5 Å. They have not been used for structure calculations. On the basis of the remaining rates and after treatment of the pseudoatoms, 273 upper distance restraints and 145 lower distance restraints were defined. The upper distance restraint set contained 125 intraresidual, 62 sequential, and 86 long-range restraints. Furthermore, 28 rates were treated as ambiguous and assigned to several pairs of protons in the X-PLOR calculations. From these rates, 28 upper and 17 lower distance restraints were obtained. Finally, 9 upper and 9 lower distance restraints were included in the DIANA calculations to constrain the three disulfide bridges. The number of lower distance restraints is always less than the number of upper distance restraints because no lower distance restraint was calculated for the distances involving a pseudoatom. The average number of restraints per residue is 10.

Twenty-three angle restraints were defined, comprising 15 ϕ angle restraints and 8 χ_1 angle restraints.

Structure Calculations. Table 1 displays the statistics concerning the 18 final X-PLOR structures. The calculated structures are consistent with the experimental distance data; they present less than four distance violations larger than 0.2 Å and none larger than 0.5 Å; moreover, no dihedral angle violation exceeds 5°. The averaged *R*-factor, which measures the agreement of the structures with the experimental NOE intensities, was calculated according to James and co-workers (Thomas et al., 1991) and is equal to 20.1%. The covalent geometry is respected, as indicated by the low average root mean square deviation (⟨RMSD⟩) values for the bond lengths (0.012 Å), the 1−3 angles (2.41°), the torsion angles (57.8°), and the impropers (1.24°). The van der Waals energy is negative (average value, −41.8 kcal/mol), indicating that there is no important nonbonded contact.

Table 1: Statistical Parameters Characterizing the 18 Final Structures of the Chimeric Protein[a]

	average value	RMS	mean structure
E (bonds)	19.40	0.40	19.20
E (1−3 angles)	72.10	2.60	73.90
E (1−4 angles)	124.00	5.00	121.00
E (impropers)	1.32	0.21	1.54
E (VdW)	−41.80	4.70	−47.7
E (electrostatic)	−1.71	0.30	−2.11
E (exptl angles)	0.00	0.00	0.00
E (exptl NOE)	8.63	1.88	6.87
E (total)	182.00	9.00	172.00
RMS (bonds)	0.012	0.000	0.012
RMS (1−3 angles)	2.41	0.04	2.44
RMS (1−4 angles)	57.80	0.30	57.90
RMS (impropers)	1.24	0.15	1.47
RMS (exptl angles)	0.002	0.007	0.020
RMS (exptl NOE)	0.036	0.003	0.033

[a] The energies are expressed in kcal/mol, the bonds in Å, and the angles in deg. The values of the experimental NOE (exptl NOE) and dihedral angle (exptl angles) energies are calculated with force constants of 20 kcal/(mol·Å^2) and 50 kcal/(mol·rad^2), respectively. The VdW energy is derived from the Lennard-Jones expression, and the electrostatic energy is obtained using a dielectric constant of 80. All of these energies are calculated using the CHARMM22 (Brooks et al., 1983) empirical function.

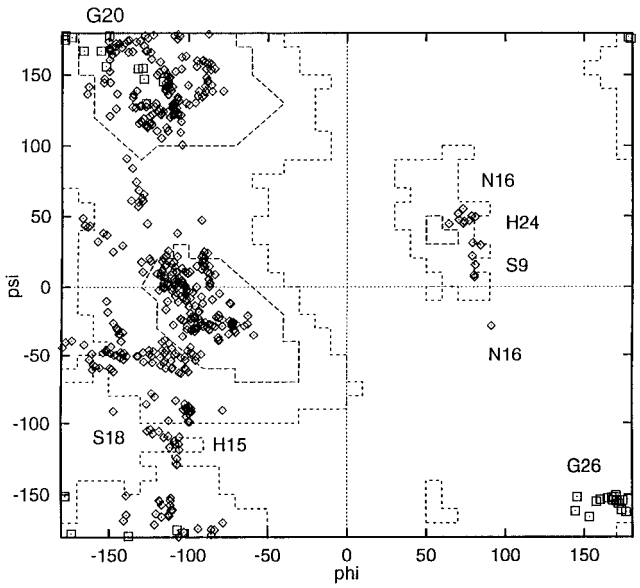

FIGURE 4: Ramachandran plot of the 18 structures of the chimera. The regions displayed in dotted lines are extracted from MacArthur and Thornton (1993) and correspond to the parts of the diagram where 50% and 70% of the residues of the well-defined three-dimensional structures of the Protein Data Bank are found (Bernstein et al., 1977). Squares correspond to glycines while diamonds correspond to the other amino acids.

Analysis of the Ramachandran plot (Figure 4) shows that the (ϕ,ψ) angle values of the calculated structures are in the allowed regions. Eight residues are occasionally found in the $\phi > 0$ region of the graph. These are the two glycines, G20 and G26, and S9, N16, and H24. Residues H15, N16, and S18 are found at the limits of the regions allowed for non-glycine amino acids.

Description of the Backbone Structure. Figure 5 shows the backbone superposition of the 18 structures of the chimeric protein. All these structures have similar folds. The RMSD value between each structure and the mean structure averaged over the 18 structures is 0.65 Å for the backbone atoms. The structure of the chimeric protein presents a 3_{10}

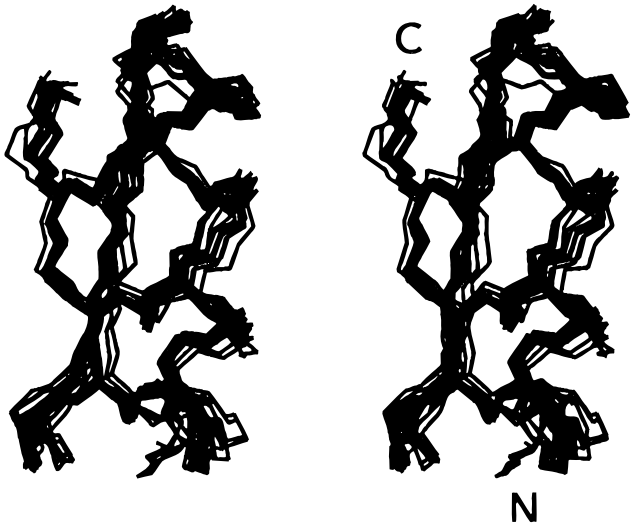

FIGURE 5: Stereoview of the 18 backbone structures of the chimeric protein. Disulfide bonds are displayed.

helix formed by residues 5−7 and an α-helix formed by residues 10−14 ($\langle\phi\rangle = -106°$; $\langle\psi\rangle = 2°$). These helices are characterized by the following hydrogen bonds: 3O−6HN and 4O−7HN for the 3_{10} helix and 11O−15HN for the α-helix. Furthermore, it shows a two-stranded antiparallel β-sheet formed by residues 19−22 and 27−30 ($\langle\phi\rangle = -113°$; $\langle\psi\rangle = 145°$). This β-sheet is stabilized by the hydrogen bonds 26O−23HN, 21O−28HN, and 19O−30HN. The average RMSd values for the coordinates of the backbone atoms belonging to the 3_{10} helix, the α-helix, and the β-sheet are 0.65, 0.53, and 0.49 Å, respectively. The α-helix and the β-sheet are better defined than the rest of the molecule.

The two residues, W8 and S9, located between the 3_{10} and the α-helices show high average RMSD values of 1.25 and 0.62 Å, respectively. This is essentially due to the presence of two different conformations for these residues, characterized by two sets of values for $\psi(8)$ and $\phi(9)$. The proton NMR data were insufficient to overcome this structural ambiguity.

The turn 23−26, which connects the two strands of the β-sheet, is close to a type I β-turn (Wilmot & Thornton, 1988) in most of the structures: $\langle\phi(24)\rangle = -128°$; $\langle\psi(24)\rangle = -26°$; $\langle\phi(25)\rangle = -94°$; $\langle\psi(25)\rangle = -29°$. As the variations around these mean angles are quite important, the average RMSD value for the coordinates of the backbone atoms corresponding to the turn yields 0.69 Å. The β-turn is less well resolved than the flanking strands of the β-sheet.

The loop 15−18, which connects the α-helix to the β-sheet, is also poorly resolved (average RMSD = 1.06 Å). Moreover, three of its residues have (ϕ,ψ) angle values at the limits of the Ramachandran plot. The structure of this loop could not be correctly characterized, and the bad fit within the Ramachandran plot suggests some conformational exchange in the loop, making it difficult to satisfy all the NMR data and the covalent geometry constraints in a single structure.

Description of the Side Chains. A number of hydrophobic contacts between the side chains of the chimeric protein contribute to the stability of the structure (Figure 6). In particular, the interactions between the cystines, V10 and L14, are probably involved in the positioning of the β-sheet (C22, C27, C29) relative to the α-helix (C7, V10, C11, L14). In the same way, the side chain of W8 is clustered with the aliphatic parts of K5 and K19. This may stabilize the α-helical structure (K5, W8), as the positioning of the β-sheet (K19) relative to the α-helix (W8). Finally, on the solvent-exposed face of the β-hairpin, the packing of the large side chain of W21 with those of D23, I28, and E30, characterized by the hydrophobic interactions between W21, I28 and the aliphatic part of E30, as well as by the hydrogen bond formed between the indole nitrogen of W21 and the δ oxygen of D23, probably contributes to the stabilization of the β-sheet structure; it also explains why the structures of the three central residues of this face, W21, D23, and I28, are particularly well defined (averaged RMSD <1 Å).

DISCUSSION

Analysis of the Refinement Protocol. In order to take into account spin diffusion and relaxation leakage in our calculations, the initial slopes of build-up curves constructed on the basis of volumes measured at three mixing times were calculated. The resulting structures are consistent with the three sets of intensities, as shown by the value of the mean *R* factor (20.1%). Such a value is comparable to those found for example by White et al. (1992) or Bonvin et al. (1994). Direct refinement of NOE intensities could improve the *R* factor, but this would lead to structures very similar to the present ones. Thus, we chose to use the intensities to

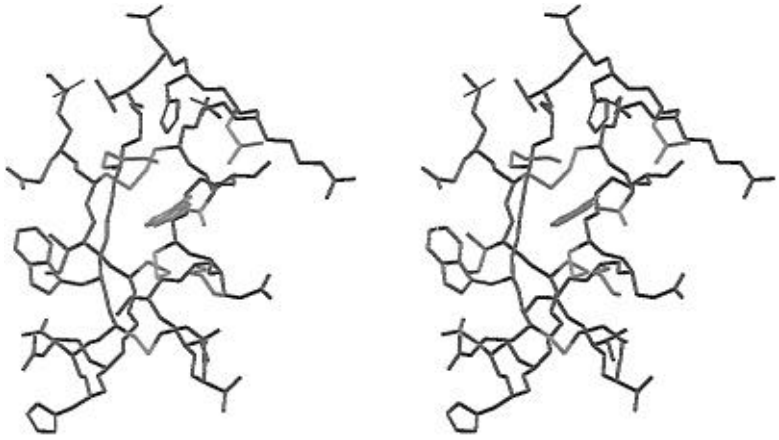

FIGURE 6: Stereoview of the minimized averaged structure of the chimeric protein. The three groups of side chains which play a role in the stabilization of the structure are displayed in red (the cystines, V10 and L14), green (K5, W8, and K19), and blue (W21, D23, I28, and E30), respectively.

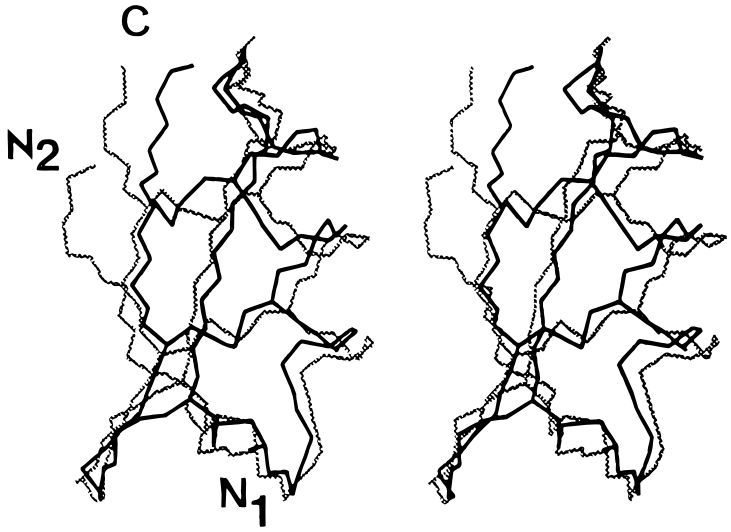

FIGURE 7: Superposition of the backbone structure of the chimeric protein (in black; from N_1 to C) with that of charybdotoxin (in gray; from N_2 to C). The disulfide bridges are displayed for the two proteins.

measure only the *R* factor, which seems to be the best way to quantitate the fit between the experimental NOE data and the three-dimensional structures.

Two methods could lead to a more realistic description of the solution structure of the chimeric protein. First, an "ensemble" refinement approach permits calculation of a set of structures where a unique structure does not obligatorily satisfy all the restraints, but where the structures all together satisfy the restraints (Bonvin et al., 1993). Second, the inclusion of internal mobility via simulated order parameters (Olejniczak et al., 1984; Koning et al., 1991) in the calculations of distance restraints from NOE intensities allows us to take into account part of the dynamics of the protein in solution when extracting distance data from the NOE intensities. These approaches should yield structures with a lower *R* factor, indicating better agreement with the experimental NOE data (Bonvin et al., 1993, 1994).

Comparison with Charybdotoxin. Superposition of the three-dimensional structures of the chimeric protein and charybdotoxin (Figure 7) shows that the two proteins have a similar fold. The RMSD between the two backbones is 1.53 Å. Their secondary structures are closely related to each other.

The RMSD between the backbone atoms of the helical regions is 0.94 Å. The differences between the two helical strands is essentially associated with the presence in the chimeric protein of a 3_{10} helix which is interrupted for two residues before becoming an α-helix, whereas charybdotoxin has a corresponding continuous and regular α-helix (Bontems et al., 1991a, 1992). A similar change from one helix type to another has already been observed in two scorpion toxins (Fernandez et al., 1994; Dauplais et al., 1995), and a high value of the $^3J_{NH-H\alpha}$ coupling constant at the first cysteine of the helix has also been found for two scorpion toxins (Johnson et al., 1994; Dauplais et al., 1995). All these data show that the helix of some scorpion toxins can be disrupted near the second cysteine of the α/β motif, emphasizing the high adaptability of the helix structure in the α/β scorpion scaffold. However, the clear interruption of the helix at residues 8 and 9 in the chimeric protein is particular to this protein and may be related to the absence of the six N-terminal residues or to the additional mutations. It is

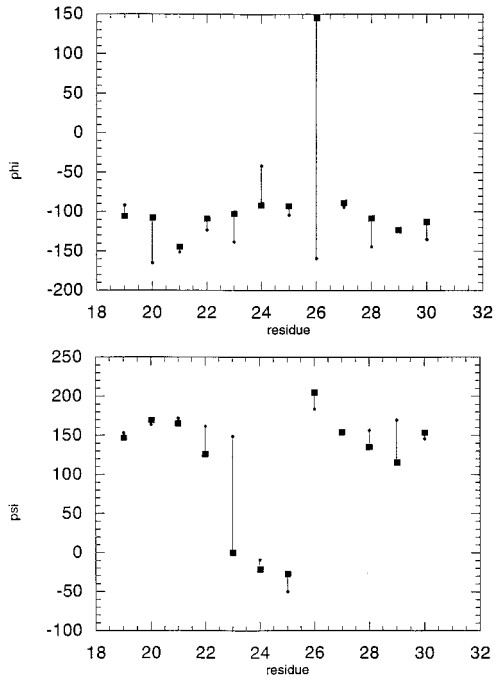

FIGURE 8: Φ,Ψ angle values of the residues from the β-hairpin of the chimeric protein (■) and charybdotoxin (●).

associated with a relative orientation of the helix with regard to the β-sheet that differs by nearly 10° between the chimera and charybdotoxin.

The β-hairpin is relatively well conserved in the two proteins, the RMSD between the backbone atoms being 1.16 Å. Comparison of the ϕ and ψ angles (Figure 8) shows that differences are essentially localized at the β-turn residues (D23, H24, G26). Both β-turns are type I though they differ in the ψ value of residue i ($\Delta\psi$ is about 150°), the ϕ value of residue $i + 1$ ($\Delta\phi$ is about 50°), and the ϕ value of residue $i + 3$ ($\Delta\phi$ is about 50°). This last residue is a lysine in charybdotoxin. In the chimera, it is a glycine and has (ϕ,ψ) angles forbidden to non-glycine residues (Figure 4), which may explain the ϕ value difference.

Among the nine side chains which are well defined in the two proteins [i.e., ⟨RMSD (χ_1)⟩ lower than 25°], eight have the same χ_1 rotamer. Moreover, the positioning of the side chains is essentially the same in charybdotoxin and the

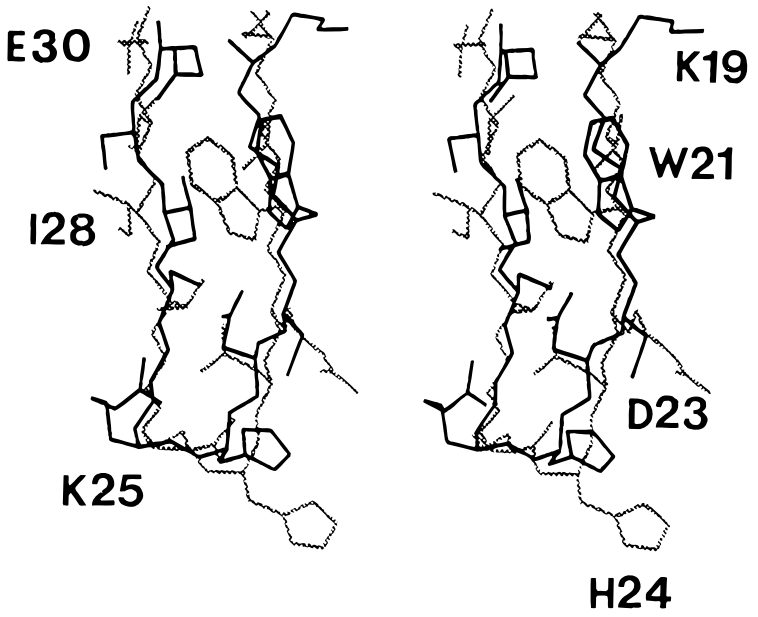

FIGURE 9: Superposition of the β-hairpin of the chimeric protein (in black) with that of toxin α (in gray). The names of the mutated side chains are displayed.

chimera (Figure 7). Differences essentially come from the fact that a few side chains which were directly interacting with the six N-terminal residues in charybdotoxin have unresolved orientations in the engineered protein. In particular, R25, which is a lysine in charybdotoxin and which interacts with the backbone portion of residues V5 and S6, is now completely solvent-exposed and poorly resolved.

Thus, in spite of eight substitutions (R25/K19, K27/W21, M29/D23, N30/H24, K31/R25, K32/G26, R34/I28, Y36/E30) and the deletion of the six first residues of charybdotoxin, the tertiary structure of the scorpion toxin is largely conserved. This result confirms our previous proposal (Vita et al., 1995) that the use of the charybdotoxin motif (7−37) as a structural template is a reasonable strategy.

Comparison with Toxin α. Figure 9 shows that the backbones of the β-sheet in the chimeric protein (residues 19−22 and 27−30) and in the snake toxin (residues 26−29 and 34−37) nicely superimpose. The RMSD value between the backbones of the two β-sheet strands is 0.92 Å. The RMSD value for the β-hairpin backbones yields 1.31 Å. This higher value results from structural differences at D23 in position *i* of the β-turn: $\Delta\phi(23)$ is about 50° and $\Delta\psi(23)$ is about 180° (Figure 10). Nevertheless, the turns in both proteins are close to type I β-turns.

The conformations and positions of the well-defined side chains of the solvent-exposed face of the chimeric β-sheet are close to their corresponding positions in toxin α. Thus, the structures of W21 and I28 are correctly reproduced. In the two proteins, W21 has a χ_1 rotamer of +60°, and I28 has a χ_1 rotamer of −60°. The hydrophobic side chains of W21 and I28 are only slightly displaced relative to the β-hairpin backbone structures. They move by approximately 1.5 Å from toxin α to the chimera when fitting the two backbones. Such a global shift may be associated with the absence of the two adjacent loops as they exist in toxin α (Figure 1). In particular, in the native toxin α, a third β-strand interacts with the (19−22) strand and forms a triple-stranded β-sheet with the two strands of the central loop. In the chimera, the absence of that strand and especially of one of its hydrophobic residues, which interacts with W21 (Zinn-Justin et al., 1992), could induce a positioning of the hydrophobic residues W21 and I28 which is not completely native-like. A similar explanation has been proposed by Cuniasse et al. (1995) to account for the difference between the conformation of a cyclic peptide corresponding to the central loop of toxin α and the structure of the native loop. No β-sheet was detected in the cyclic peptide. The authors suggested that, as in toxin α, several hydrophobic residues of the central β-sheet interact with residues present in the two adjacent loops; the absence of these loops in the cyclic peptide could perturb the formation of the β-sheet (Cuniasse et al., 1995). In relation to the slightly different positioning of W21 and I28, the side chain of E30 has also moved by about 1 Å and its χ_1 rotamer is −60° in the chimera but +60° in toxin α. On the opposite strand of the β-sheet, K19

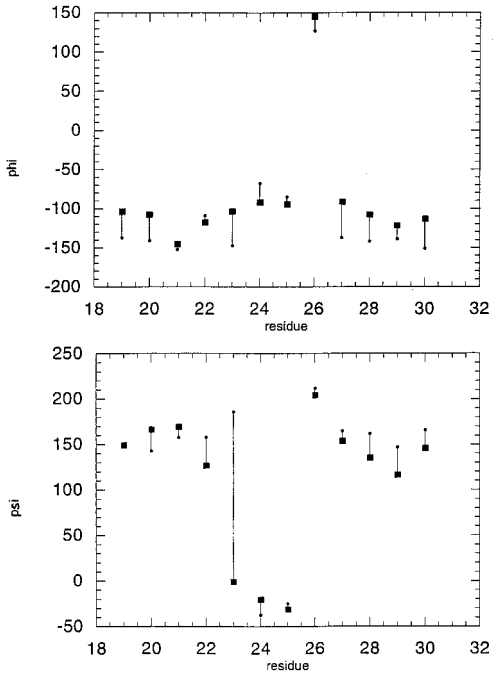

FIGURE 10: Φ,Ψ angle values of the residues from the β-hairpin of the chimeric protein (■) and toxin α (●).

shows completely different positions in the two proteins (Figure 9). It interacts with W8 in the helix of the chimera and with W21 in the β-sheet of toxin α. Thus, the presence of an alternative hydrophobic residue in the helix region of the chimera perturbs the positioning of K19 which has fewer van der Waals contacts with W21 in the chimera than in toxin α. In the β-turn, D23 has similar positions but different conformations in the two proteins (Figure 9). Its χ_1 angle value is $-83° \pm 4°$ in the chimera and $199° \pm 9°$ in toxin α (Zinn-Justin et al., 1992). Moreover, the hydrogen bond involving the oxygen of the aspartic acid and the indole nitrogen of the tryptophan observed in several structures of the chimera is not found in the solution structure of toxin α. Such different conformations of the D23 side chain in the two proteins is probably related to the local differences in the backbone conformations, but no clear evidence permits an explanation of these differences. Finally, the side chains of H24 and R25 are largely solvent-exposed and thus badly resolved in the two proteins.

Conclusion. The present study has further confirmed our earlier claims that the α/β scorpion fold is an appropriate template to successfully achieve the molecular grafting of novel functions. Previously, it was shown that a metal binding site, predominantly composed of three coordinating histidines, could be adequately engineered on the surface of the β-sheet hairpin of the fold (Vita et al., 1995). More recently, the transfer of residues associated with a receptor binding function was achieved (Drakopoulou et al., 1996). The analysis reported in the present paper shows that, in the last chimeric protein, introduction of eight mutations and deletion of the first six N-terminal residues as compared to the amino acid sequence of charybdotoxin did not prevent the original α/β motif to be folded. The structural study also indicates that, when present in a snake toxin or transferred into a scorpion toxin, the functional β-hairpin adopted a similar backbone structure. Even the overall positions of side chains located on the solvent-exposed side of the β-sheet, including W21/W28, I28/I35, and E30/E37, are conserved in the original and transferred β-hairpins. Finally, the side chains located at positions $i + 1$ and $i + 2$ of the β-turn, which are largely solvent-exposed, seem to be relatively mobile in both the chimera and toxin α.

However, despite this successful structural transfer, the binding affinity of the chimera is characterized by a dissociation constant of about 1 μM (Drakopoulou et al., 1996), a value which is nearly 5 orders of magnitude higher than that of the original snake toxin. The related question is whether or not this result is in accordance with our conclusion that the snake toxin β-hairpin has been adequately reproduced in the chimera. On the basis of approximate theoretical calculations, Drakopoulou et al. (1996) estimated that, as only six of the ten residues of the active site of snake neurotoxins have been transferred to the chimera, the new protein should recognize the toxin receptor with a much lower affinity, consistent with the experimental value of 1 μM. This explanation suggests that the structural transfer has been correctly achieved. Nevertheless, the low binding affinity of the new protein for the nicotinic acetylcholine receptor might also partially result from an approximate structural transfer of the snake β-hairpin. How then could this transfer be improved? Our study noted some structural differences between the original and grafted β-hairpins. In particular, the two side chains of K19 and D23 adopted different conformations. One way to erase these differences may consist of mutating W8 into a polar residue, to allow K19 to retrieve a native-like positioning in the chimeric β-hairpin. This proposal may be important as this lysine is one of the most critical residues for the function of snake neurotoxins: the single mutation of the lysine into a glutamic acid induces a 175-fold affinity decrease for the nicotinic acetylcholine receptor (Hervé et al., 1992; Trémeau et al., 1995). On the other hand, it is difficult to evaluate the impact of the observed structural differences for the binding of the chimera to the nicotinic acetylcholine receptor, especially because the three-dimensional structure of the receptor is unknown. The dynamics of the β-hairpin in the snake toxin and in the chimera constitute another parameter which has not been considered upon transfer of the β-hairpin. Since it is well established that intramolecular motions of proteins on several distinct time scales may be crucial for their biological functions, a study of the compared dynamics of toxin α and the chimeric protein is suitable.

Thus, structural and dynamic analyses may be needed for increasing the binding affinity of an engineered molecule. These studies could be combined to random mutagenesis at some chosen positions in the engineered protein (Li et al., 1995). The rational approach described in this paper associated with a combinatorial methodology is anticipated to lead to ligands with a higher affinity for the nicotinic acetylcholine receptor.

REFERENCES

Allemann, R. K., Presnell, S. R., & Benner S. A. (1991) *Protein Eng. 4*, 831−835.

Argos, P. (1987) *J. Mol. Biol. 197*, 331−348.

Aue, W. P., Bartholi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229−2246.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Bonmatin, J.-M., Bonnat, J.-L., Gallet, X., Vovelle, F., Ptak, M., Reichhart, J.-M., Hoffmann, J. A., Keppi, E., Legrain, M., & Achstetter, T. (1992) *J. Biol. NMR 2*, 235−256.

Bontems, F., Roumestand, C., Boyot, P., Gilquin, B., Doljansky, Y., Ménez, A., & Toma, F. (1991a) *Eur. J. Biochem. 196*, 19−28.

Bontems, F., Roumestand, C., Gilquin, B., Ménez, A., & Toma, F. (1991b) *Science 254*, 1521−1523.

Bontems, F., Gilquin, B., Roumestand, C., Ménez, A., & Toma, F. (1992) *Biochemistry 31*, 7756−7764.

Bonvin, A. M. J. J., Rullman, J. A. C., Boelens, R., & Kaptein, R. (1993) *Proteins: Struct., Funct., Genet. 15*, 385−400.

Bonvin, A. M. J. J., Vis, H., Breg, J. N., Burgering, M. J. M., Boelens, R., & Kaptein, R. (1994) *J. Mol. Biol. 236*, 328−341.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.

Braxton, S., & Wells, J. A. (1992) *Biochemistry 31*, 7796−7801.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Bruix, M., Jiménez, M. A., Santoro, J., Gonzalez, C., Colilla, F. J., Mendez, E., & Rico, M. (1993) *Biochemistry 32*, 715−724.

Bruix, M., Gonzalez, C., Santoro, J., Soriano, F., Rocher, A., & Mendez, E. (1995) *Biopolymers 36*, 751−763.

Brünger, A. T. (1992) *X-PLOR version 3.1: a System for X-ray Crystallography and NMR*, Yale University, New Haven and London.

Chothia, C., Lesk, A. M., Levitt, M., Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., & Poljak, R. J. (1986) *Science 233*, 755−758.

Cuniasse, P., Thomas, A., Smith, J. C., Lam Thanh, H., Léonetti, M., & Ménez, A. (1995) *Biochemistry 34*, 12782−12789.

Dauplais, M., Gilquin, B., Possani, L. D., Gurrola-Briones, G., Roumestand, C., & Ménez, A. (1995) *Biochemistry 34*, 16563−16573.
Drakopoulou, E., Zinn-Justin, S., Guenneugues, M., Gilquin, B., Ménez, A., & Vita, C. (1996) *J. Biol. Chem. 271*, 11979−11987.
Eijsink, V. G. H., Vriend, G., van den Burg, B., van der See, J. R., Veltman, O. R., Stulp, B. K., & Venema, G. (1992) *Protein Eng. 5*, 157−163.
Fernandez, I., Romi, R., Szendeffy, S., Martin-Eauclaire, M. F., Rochat, H., Van Rietschoten, J., Pons, M., & Giralt, E. (1994) *Biochemistry 33*, 14256−14263.
Fine, R. M., Wang, H., Shenkin, P. S., Yarmush, D. L., & Levinthal, C. (1986) *Proteins 1*, 342−362.
Gippert, G. P., Yip, P. F., Wright, P. E., & Case, D. A. (1990) *Biochem. Pharmacol. 40*, 15−22.
Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
Hanzawa, H., Shimada, I., Kuzuhara, T., Komano, H., Kohda, D., Inagaki, F., Natori, S., & Arata, Y. (1990) *FEBS Lett. 269*, 413−420.
Harper, J. W., & Vallee, B. L. (1989) *Biochemistry 28*, 1875−1884.
Hedstrom, L., Szilagyi, L., & Rutter, W. J. (1992) *Science 255*, 1249−1253.
Hervé, M., Pillet, L., Humbert, P., Trémeau, O., Ducancel, F., Hirth, C., & Ménez, A. (1992) *Eur. J. Biochem. 208*, 125−131.
Hyberts, S. G., Märki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625−635.
Hynes, T. R., Kautz, R. A., Goodman, M. A., Gill, J. F., & Fox, R. O. (1989) *Nature 339*, 73−76.
Johnson, B. A., Stevens, S. P., & Williamson, J. M. (1994) *Biochemistry 33*, 15061−15070.
Kabsch, W., & Sander, C. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1075−1078.
Koning, T. M. G., Boelens, R., van der Marel, G. A., van Boom, J. H., & Kaptein, R. (1991) *Biochemistry 30*, 3787−3797.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 64*, 2229−2246.
Kuntz, I. D. (1972) *J. Am. Chem. Soc. 94*, 4009−4012.
Li, B., Tom, J. Y. K., Oare, D., Yen, R., Fairbrother, W. J., Wells, J. A., & Cunningham, B. C. (1995) *Science 270*, 1657−1660.
MacArthur, M. W., & Thornton, J. M. (1993) *Proteins: Struct., Func., Genet. 17*, 232−251.
Ménez, A., Bontems, F., Roumestand, C., Gilquin, B., & Toma, F. (1992) *Proc. R. Soc. Edinburgh 99B*, 83−103.
Moult, J., & James, M. N. G. (1986) *Proteins 4*, 146−163.
Olejniczak, E. T., Dobson, C. M., Karplus, M., & Levy, R. M. (1984) *J. Am. Chem. Soc. 106*, 1923−1930.
Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741−751.
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.
Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.
Raines, R. T., Toscano, M. P., Nierengarten, D. M., Hoi ha, J., & Auerbach, R. (1995) *J. Biol. Chem. 270*, 17180−17184.
Rance, M., Sorensen, O., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
Rose, G. D., Gierasch, L. H., & Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1−109.
Segalas, I., Roumestand, C., Zinn-Justin, S., Gilquin, B., Ménez, R., Ménez, A., & Toma, F. (1995) *Biochemistry 34*, 1248−1260.
Shaka, A. J., Keeler, J., & Freeman, R. (1983) *J. Magn. Reson. 53*, 313−334.
Thomas, P. D., Basus, V. J., & James, T. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1237−1241.
Toma, S., Campagnoli, S., Margarit, I., Gianna, R., Grandi, G., Bolognesi, M., Defillipis, V., & Fontana, A. (1991) *Biochemistry 30*, 97−106.
Trémeau, O., Lemaire, C., Drevet, P., Pinkasfeld, S., Ducancel, F., Boulain, J.-C., & Ménez, A. (1995) *J. Biol. Chem. 270*, 9362−9369.
Verhoeyen, M., Milstein, C., & Winter, G. (1988) *Science 239*, 1534−1536.
Vita, C., Bontems, F., Bouet, F., Tauc, M., Poujeol, P., Vatanpour, H., Harvey, A. L., Ménez, A., & Toma, F. (1993) *Eur. J. Biochem. 217*, 157−169.
Vita, C., Aniort, V., Ménez, A., Toma, F., & Neyton, J. (1994) in *Innovation and Perspectives in Solid Phase Synthesis* (Epton, R., Ed.) pp 201−206, Mayflower Worldwide, Birmingham, U.K.
Vita, C., Roumestand, C., Toma, F., & Ménez, A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6404−6408.
Vuillemer, S., & Fersht, A. (1994) *Eur. J. Biochem. 221*, 1003−1012.
White, S. A., Nilges, M., Huang, A., Brünger, A. T., & Moore, P. B. (1992) *Biochemistry 31*, 1610−1621.
Wilmot, C. M., & Thornton, J. M. (1988) *J. Mol. Biol. 203*, 221−232.
Wolfson, A. J., Kanaoka, M., Lau, F. T. K., & Ringe, D. (1991) *Protein Eng. 4*, 313−317.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
Zinn-Justin, S., Roumestand, C., Gilquin, B., Bontems, F., Ménez, A., & Toma, F. (1992) *Biochemistry 31*, 11335−11347.

BI960466N