NMR Sequential Assignments and Solution Structure of Chlorotoxin, a Small Scorpion Toxin That Blocks Chloride Channels[†,‡]

G. Lippens,[*,§,||] J. Najib,[§] S. J. Wodak,[||] and A. Tartar[§]

Service de Chimie des Biomolécules (SCBM), Institut Pasteur de Lille, 1, rue du Professeur Calmette, 59000 Lille, France, and Unité de Conformation des Macromolécules Biologiques (UCMB), CP 160-16, Université Libre de Bruxelles, Avenue P. Héger, B-1050 Brussels, Belgium

Received June 2, 1994; Revised Manuscript Received September 12, 1994[⊗]

ABSTRACT: The solution structure of chlorotoxin, a small toxin purified from the venom of the *Leiurus quinquestriatus* scorpion, has been determined using 2D ^{1}H NMR spectroscopy. Analysis of the NMR data shows that the structure consists of a small three-stranded antiparallel β-sheet packed against an α-helix, thereby adopting the same fold as charybdotoxin and other members of the short scorpion toxin family [Arseniev *et al.* (1984) *FEBS Lett. 165*, 57−62; Martins *et al.* (1990) *FEBS Lett. 260*, 249−253; Bontems *et al.* (1991) *Science 254*, 1521−1523]. Three disulfide bonds of chlorotoxin (Cys5−Cys28, Cys16−Cys33, and Cys20−Cys35), cross-linking the α-helix to the β-sheet, follow the common pattern found in the other short scorpion toxins. The fourth disulfide bridge (Cys2−Cys19) links the small N-terminal β strand to the rest of the molecule, in contrast to charybdotoxin where this disulfide bridge is absent and the first strand interacts with the rest of the molecule by several contacts between hydrophobic residues. Another structural difference between chlorotoxin and charybdotoxin is observed at the level of the $\alpha-\beta$ turn. This difference is accompanied by a change in the electrostatic potential surface, which is largely positive at the level of this turn in chlorotoxin, whereas no such positive potential surface can be found at the same position in charybdotoxin. In the latter protein, the positive surface is formed by different charged residues situated on the solvent-exposed site of the C-terminal β-sheet. This positive region is disrupted in chlorotoxin by the mutation of one of the arginines into a leucine and by the presence of two aspartic acid residues on the α-helix.

Scorpion venoms are known to contain a variety of toxins acting on Na^+ and K^+ channels. Recently, an additional activity on Cl^- channels was discovered in the venom of *Leiurus quinquestriatus* (Debin *et al.*, 1993). Because of their high selectivity, these toxins are employed as pharmacological tools for the study of ionic channels (Moczydolowski *et al.*, 1988; Strong, 1990; Betz, 1990; Cook & Quast, 1990; Dreyer, 1990; Garcia *et al.*, 1991). Despite their large differences in specificities and activities, the toxins of different scorpion species can be considered as one large protein family. They have been classified into two main categories on the basis of their known primary structures: less than 40 amino acids (short-chain toxins) or 60−70 amino acids (long-chain toxins). Both categories of toxins contain three or four intramolecular disulfide bridges, which are at the basis of their high compactness and stability.

Chlorotoxin (ClTx) is a small neurotoxic peptide, belonging to the short-chain toxins isolated from the venom of scorpion *L. quinquestriatus*. It acts as a specific blocker of the small conductance epithelial chloride channels, hence its name of "chlorotoxin" (Debin *et al.*, 1993). Still, the same authors observed an inhibition only when the molecule was applied to the intracellular surface of the channel and hence concluded that it is unlikely that epithelial Cl^- channels are the natural target for chlorotoxin. Sequence analysis showed the molecule to be a 36 amino acid peptide, including eight cysteines, which suggests the presence of four disulfide bonds. The examination of its primary structure revealed a high level of homology with those of the insectotoxins I_5A isolated from *Buthus eupeus* venom (Grishin *et al.*, 1978), peptide I from *Buthus sindicus* (Fazal *et al.*, 1989), and AmmP2 from *Androctonus mauretanicus* (Zlotkin *et al.*, 1978; Rochat *et al.*, 1979) (Figure 1a). Despite a lesser degree of sequence homology, chlorotoxin can also be aligned with an other group of short scorpion toxins which contains only three disulfide bridges: charybdotoxin (Bontems *et al.*, 1991a), iberiotoxin (Galvez *et al.*, 1990; Johnson *et al.*, 1992), scyllatoxin (Martins *et al.*, 1990), PO5 (Meunier *et al.*, 1993), noxiustoxin (Possani *et al.*, 1982), and kaliotoxin (Crest *et al.*, 1992) (Figure 1b). These short toxins inhibit the Ca^{2+}-activated K^+ channels and form a family characterized by a typical helix/sheet organization. In order to determine whether the presence of a fourth disulfide bridge is compatible with this spatial organization, and to gain insight into the channel selectivity of this new member of the family of small scorpion toxins, we determined the complete three-dimensional structure of chlorotoxin by ^{1}H NMR and compared it both from a structural and functional point of view to the other ion channel blockers.

METHODS

Purification of the Natural Chlorotoxin. Semipurified fractions of the scorpion venom containing the chlorotoxin were from Latoxan (Rosans, France). The n-ClTx was

[†] G.L. is a research associate receiving financial support from the Belgian National Science Foundation (NFWO).

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank (file name 1CHL).

* To whom correspondence should be addressed.

[§] Institut Pasteur de Lille.

[||] Université Libre de Bruxelles.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1994.

0006-2960/95/0434-13$09.00/0

```
Chlorotoxin    MCMPCFTTDHQMARKCDDCCGGKGRGKCYGPQCLCR-NH2
I5A            MCMPCFTTDPNMAKKCRDCCGG--NGKCFGPQCLCNR-NH2
Peptide I      RCKPCFTTDPQMSKKCADCCGGKGKGKCYGPQCLC
AmmP2           CGPCFTTDPYTESKCATCCGG--RGKCVGPQCLCNRI

Chlorotoxin    ---MCMPCFTTDHQMARKCDDCCGGKG---RG-KCYGPQCLCR
Charybdotoxin  -EFTNVSC-TT-----SKECWSVCQRLHNTSRG-KCMNKKCRCYS
Iberiotoxin    -EFTDVDCSV------SKECWSVCKDLFGVDRG-KCMGKKCRCYQ
Leiurotoxin    -----AFC--NLRM-----CQLSCRSLGLL--G-KCIGDKCECVKH
P05            -----TVC--NLRR-----CQLSCRSLGLL--G-KCIGVKCECVKH
Noxiustoxin    -TIINVKC-TSPKQ-----CSKPCKELYGSSAGAKCMNGKCKCYN
Kaliotoxin     GVEINVKCSGSPQ------CLKPCKDAAGMRFG-KCMNRKCHCTP
```

FIGURE 1: (a) Sequence alignment of chlorotoxin with those of other small insectotoxins characterized by four disulfide bonds, by matching cysteine residues. (b) Sequence alignment of chlorotoxin with those of small scorpion toxins characterized by three disulfide bonds, by matching cysteine residues. The three conserved cysteine bonds are indicated in bold. Replacement of the positively charged K23 in chlorotoxin by a hydrophobic residue in the other toxins is underlined; the residue between the two cysteines on the C-terminal β-sheet is indicated in italic (see text).

purified by reverse-phase HPLC (Nucleosil, C18, 10.7 × 500 mm, 5 μm, 100 Å) at a flow rate of 2 mL/min, using a gradient from 0% to 18% for 15 min and then from 18% to 36% CH_3CN in aqueous TFA in 60 min.

NMR Spectroscopy. All NMR experiments were performed at 10 and 36°C on a Varian Unity 600 MHz spectrometer (Varian Instruments, Palo Alto, CA). We recorded different homonuclear ^1H spectra of a 3 mM sample of the natural chlorotoxin at pH 3.2 (not corrected for isotope effects) dissolved successively in 95% H_2O/5% D_2O, including a DQF-COSY spectrum (Braunschweiler *et al.*, 1983), a TOCSY spectrum with a spin-locking time of 60 ms (Bax & Davis, 1985), and a NOESY spectrum in H_2O (Jeener *et al.*, 1979) with a mixing time of 200 ms. Water suppression was achieved by presaturation, and the SCUBA method for recovery of saturated H_α resonances was used for the homonuclear spectra (Brown *et al.*, 1988). Quadrature detection in F_1 and suppression of axial peaks were achieved according to the TPPI−States method (Marion *et al.*, 1989). The spectra were typically acquired with 2048 points in F_2 (4096 for the DQF-COSY) and 512 complex points in F_1. Zero-filling was applied to obtain a number of points equal to the next power of two. After lyophilization, the sample was dissolved in D_2O and brought immediately in the preshimmed spectrometer at 36 °C. Amide protons exhibiting low rates of exchange with the solvent were identified by a series of short (2 h) TOCSY spectra. Amide protons still giving rise to intraresidual H_α−NH cross-peaks 24 h after dissolution of the sample were considered to be slowly exchangeable. A series of NOESY spectra with mixing times of 50, 100, 150, and 200 ms were then taken at the same temperature, and the coupling constants $J_{\alpha\beta}$ were measured using a E.COSY experiment (Griesinger *et al.*, 1985). These coupling constants, together with the classification of the intraresidual $\alpha-\beta$ and β-NH cross-peaks as strong or weak, allowed us to attribute stereospecifically a majority of the methylene protons (see Results).

Data were processed using the Varian software on a SUN Sparc 2 or the program SNARF (F. H. J. van Hoesel, University of Groningen, The Netherlands) on a Silicon Graphics Indigo R3000 workstation. A routine was incorporated into the latter program, performing automatically (1) the volume integration of the cross-peak, (2) a baseline correction where we subtract from this volume integral the area of a rectangle around the cross-peak considered times the average spectral value over its borders, and (3) calibration of the distance. This last step used the average volume of the cross-peaks between seven methylene proton pairs of the cysteine residues (excluding Cys5; see further) as a yard stick. The same routine also gave a calibration with respect to peak height, and the consistency between the results using volumes and heights was used to monitor the line width and thus the T_2 value of the resonances concerned, in order to evaluate an eventual influence of the mobility on the distances obtained.

The final set of restraints (removing trivial constraints such as the methylene proton distances) consisted of 311 distance constraints, consisting of 156 intraresidue constraints and 155 interresidue constraints; of these latter, 91 were sequential constraints and 64 nonsequential constraints. This last group contained 49 contacts between protons that are three or more residues apart and 35 contacts between protons that are five or more residues apart. The distances were classified into three classes, corresponding to distance ranges of [1.5Å, 2.8Å] (strong), [1.5Å, 3.5Å] (medium), and [1.5Å, 4.5Å] (weak). The accurate distance estimations obtained by comparison of the buildup curves to the averaged one over seven methylene protons were initially not used in the calculations but were used to verify *a posteriori* the quality of the results. The slow exchange of amide protons was introduced as a short distance between the amide proton and the oxygen atom implicated in the hydrogen bond, based on the hypothesis that slow amide exchange is due to a protection of the proton in a hydrogen bond. A total of eight distances (involving 16 residues) was introduced. In addition to the distance constraints, we introduced 14 χ_1 angle restraints for the orientation of the side chains.

Structure Calculations. All structures were generated with the X-PLOR program (Brünger, 1992). We first used the *sub-embed* algorithm, where we extended the subembedded set to the δ and ϵ carbons (because some important constraints concerned the methyl groups of Met12 and Leu34). The covalent constraints concerning the disulfide bridges were at this stage eliminated and were reintroduced as fake NOE distances. NOE constraint terms were calculated using a split harmonic energy term with a force constant of 50 kcal/(mol Å^2). After the subembedding, we performed a simulated annealing step on every structure, consisting of a thousand steps at 2000 K followed by a cooling down during another 1000 steps to room temperature. The final structures were imported into the Brugel software package (Delhaise *et al.*, 1985) and further subjected to a nonconstrained steepest descent refinement of 200 steps using the CHARMM force field version 19. The conformational energies improved considerably during this refinement, even though the structures hardly moved at all (RMS values of less than 0.2 Å). We performed a final check to verify whether the refinement without the NOE constraints had not introduced new NOE constraint violations and found this was not the case. Therefore, we used this final set of structures for evaluation of the different energetic contributions and for the structural comparisons.

Visual analysis and comparison with the structures of charybdotoxin and leiurotoxin I were done with the INSIGHT II module of the BIOSYM package (BIOSYM Technologies, San Diego, CA) running on an Indigo R3000 Silicon Graphics workstation.

Electrostatic Potential Calculations. The electrostatic potential surfaces around chlorotoxin and charybdotoxin were calculated by the DELPHI program (Gilson & Honig, 1988) that solves the linearized Poisson−Boltzmann equation at the nodes of a cubic lattice. The parameters used were 65 nodes by dimension of the cube, successively 30%, 50%, 75%, and 90% for the occupancy of the cube volume by the protein, 2 for the dielectric constant of the protein, 78.6 for the dielectric constant of the solvent, 0.2 M for the ionic strength, 2 Å for the exclusion radius of ions, and 1.4 Å for the exclusion radius of the solvent. The calculated electrostatic potential surfaces were displayed and graphically analysed using the BRUGEL software (Delhaise *et al.*, 1985).

RESULTS

^1H NMR Assignments. Sequence-specific assignment of the resonances of chlorotoxin was accomplished according to the method described by Wüthrich (1986). The six long basic side chains (three Arg and three Lys) could easily be recognized by their TOCSY lines from the $N^{\epsilon}H$ or $C^{\epsilon}H$ to the $C^{\alpha}H$ and by the reverse transfer pathway from the NH proton to the $N^{\epsilon}H$ or $C^{\epsilon}H$. The Thr7, Thr8, Ala13, and Leu34 residues were also easily distinguishable because they are the only residues containing a methyl group. The eight Cys residues are characterized by well-separated resonance frequencies of their amide proton resonances, so no problems of overlap were encountered. The two Pro residues were identified by the TOCSY lines starting from their $C^{\delta}H$ protons around 3.7 ppm to their C^{α} H proton, and a sequence specific assignment could be made by the NOE contact from their α proton to the NH proton from the following residue and from their δ protons to the α proton of the preceding residue. The first Met residue was assigned on the basis of the intense NOE cross-peaks from both its α and β protons toward the NH protons of the Cys2 residue and by a subsequent verification of the spin system in the DQF-COSY spectrum in D_2O. Similarly, the Met3 residue was assigned by the NOE cross-peaks from its α proton toward the Pro4 δ protons, but the NH of the same residue was not detectable at 36 °C, probably because at this low pH (pH = 2.55) and elevated temperature it is exchanging rapidly with the protons of water. A spectrum at 10 °C confirmed this hypothesis, because we found a broadened TOCSY cross-peak from a NH proton resonating at 7.6 ppm toward the earlier found α-proton position at 4.9 ppm. The protons belonging to the aromatic side chains of Phe6 and Tyr29 were assigned by the characteristic NOE cross-peaks between their β protons and their δ protons, where the latter resonate typically around 6.5−7 ppm. For the Phe6, however, all aromatic protons cluster around 7.31 ppm, so we could not distinguish between δ, ϵ, or ζ protons. In Table 1, we give all resonance frequencies at 36 °C.

While performing a comparison with a sample of chlorotoxin obtained by chemical synthesis (J. Najib, to be published) a remarkable pH dependence was found for the amide proton resonances of the glycine residues at positions 24 and 26; whereas most resonances hardly change with pH, the resonances of the amide protons of these two glycines change by as much as 0.7 ppm when the pH is varied from 2.5 to 4.7. This dependence may be explained by the fact that the two NH protons in question point toward the two aspartic acids Asp17 and Asp 18 on the helix. The only other significant variations observed were for the amide protons of His10 and Gln11, but in this case the variation was limited to 0.1 ppm.

Table 1: Chemical Shifts Measured for Chlorotoxin at pH 2.55 and 36 °C[a]

residue	NH	$C^{\alpha}H$	$C^{\beta}H$	others
Met1		4.16	2.18	γ 2.40/2.12
Cys2	9.14	4.44	*3.17*/2.82	
Met3	7.60*	4.89	2.41/2.27	γ 2.55/2.33
Pro4		4.80	2.13/1.19	γ 2.05/1.96, δ 3.82/3.62
Cys5	7.51	5.02	*3.71*/2.85	
Phe6	8.31	4.86	3.21/*2.91*	δ 7.31, ϵ 7.31, ζ 7.31
Thr7	8.50	4.20	4.25	CH_3 1.25
Thr8	7.47	4.17	4.38	CH_3 1.20
Asp9	7.41	4.70	3.24/2.72	
His10	9.18	4.59	3.42/*3.30*	
Gln11	8.76	4.59	2.40/1.92	γ_{12} 2.35
Met12	7.53	4.22	*2.52*/2.22	γ 2.97/2.34
Ala13	8.54	3.95	CH_3 1.40	
Arg14	7.83	4.17	1.91	γ_{12} 1.67, δ_{12} 3.22, $N^{\epsilon}H$ 7.28
Lys15	8.22	4.05	1.90/1.76	γ_{12} 1.43/δ_{12} 1.64/ϵ_{12} 2.93/ $N^{\zeta}H_3^+$ 7.79
Cys16	8.51	4.22	*2.97*/2.79	
Asp17	7.85	4.20	2.93/*2.77*	
Asp18	8.55	4.24	2.92/2.86	
Cys19	8.23	4.11	*3.77*/*3.35*	
Cys20	7.52	4.61	3.14/*2.52*	
Gly21	7.81	4.39/3.81		
Gly22	8.34	4.32/3.71		
Lys23	8.26	4.00	1.76	γ 1.44/1.35, δ_{12} 1.67/ϵ_{12} 2.97/ $N^{\zeta}H_3^+$ 7.50
Gly24	8.61	4.06/3.98		
Arg25	7.76	4.20	1.94/1.70	γ_{12} 1.39, δ_{12} 3.02, $N^{\epsilon}H$ 7.11
Gly26	8.48	5.18/3.55		
Lys27	8.52	4.37	1.64	γ 1.15/0.95, δ 1.52/1.46, ϵ_{12} 2.72/ $N^{\zeta}H_3^+$ 7.18
Cys28	8.68	5.15	2.95/*2.87*	
Tyr29	9.40	4.57	2.98/*2.89*	δ 7.21, ϵ 6.81
Gly30	8.70	4.22/3.50		
Pro31		4.88	2.24/2.21	γ 2.10/1.83, δ 4.00/3.76
Gln32	7.68	4.78	2.18/1.89	γ 2.33
Cys33	8.49	4.96	3.12/*2.93*	
Leu34	9.05	5.04	*1.92*/1.45	γ 1.70, δ_1 CH_3 0.86, δ_2 CH_3 0.67
Cys35	8.87	5.63	3.46/*2.69*	
Arg36	8.67	4.36	1.87/1.78	γ_{12} 1.63, δ_{12} 3.18, $N^{\epsilon}H$ 7.05

[a] All resonance frequencies are given with respect to (trimethylsilyl)propionate. In the case of stereospecific assignments, the frequency of the β_1 proton is indicated in italic. The resonance indicated by an asterisk corresponds to the frequency at 10 °C because it was not visible at 36 °C.

Secondary Structure. The details of the secondary structure were determined from the NMR data summarized in Figure 2. The pattern of d_{NN}, $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ indicates clearly an α-helix from residues 11 to 21 and the presence of four contacts, $\alpha N(i,i+4)$, in the NOE spectrum at 10 °C confirmed this hypothesis.

The pattern of consecutive $\alpha N(i,i+1)$ contacts in the residue 26−29 and 32−36 parts of the molecule together with the long-range NOE's connecting the α protons of residues 26−35 and 28−33 are consistent with a β hairpin having a β turn centered around the residues 30−31. A further indication of the presence of this β hairpin is provided by the slow exchange rate of all amide protons that point to the interior of the hairpin.

The amide proton of Cys33 is also well protected although it points outward from the above mentioned hairpin. This, however, can be explained by the spatial proximity of an additional small β strand formed by the first four residues. The presence of a β strand spanning residues 1−4 is confirmed by the observation of a strong contact between the α proton of Met1 and the amide proton of Cys2 and between the α proton of Met3 and the δ protons of Pro4.

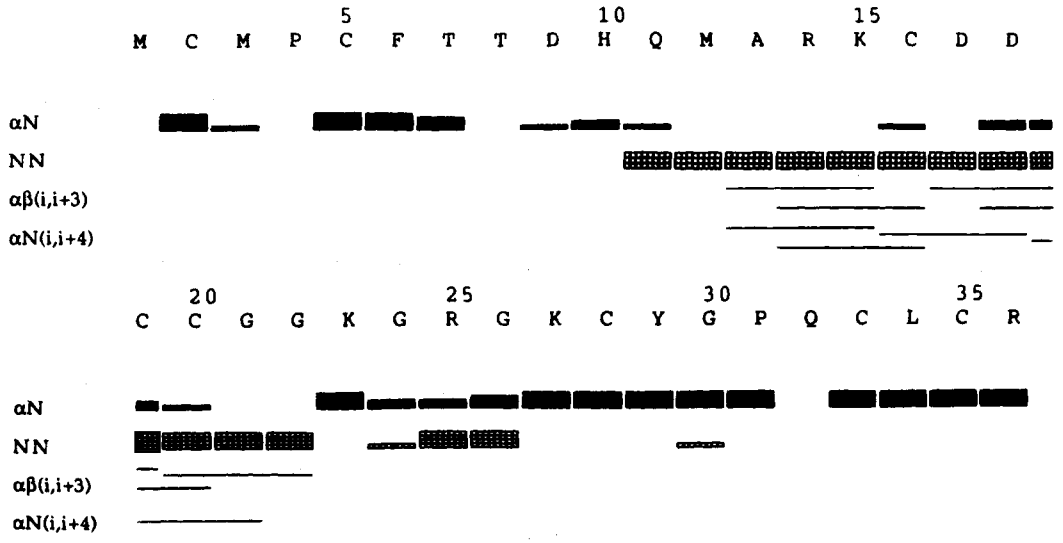

FIGURE 2: Summary of sequential and medium-range NOE's, characteristic for β-sheet and α-helix.

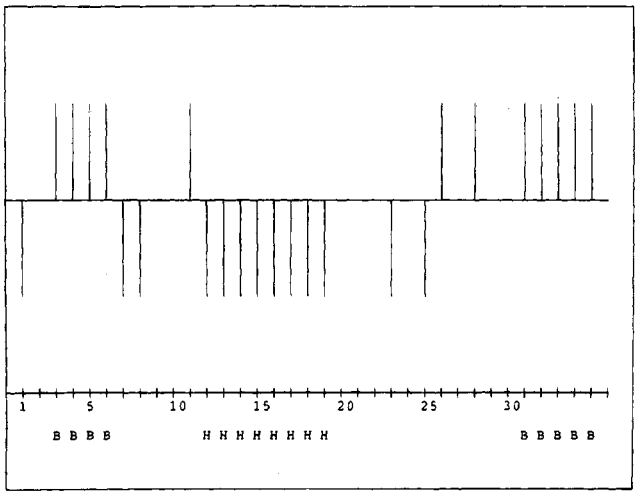

FIGURE 3: Assignment of secondary structure of chlorotoxin, as determined by the chemical shift index method. The observed $C_\alpha H$ chemical shifts are compared with the random coil values given by Wüthrich (1986). A 1, a 0, or a -1 value shows respectively that the backbone α-proton chemical shift of a given residue is above, within, or below the range given by Wishart *et al.* (1992). An α-helix (H) is recognized as a dense group of four or more 1's and a β-sheet (B) as a dense group of 4 or more -1's.

The method described recently by Wishart *et al.* (1992) uses the deviation of the chemical shifts of the α protons from their random coil values as an indication of the secondary structure. In Figure 3, representing graphically these deviations, the first β strand, the α helix, and the C-terminal strand of the β hairpin can be readily recognized. The chemical shift deviations from the random coil values do not attribute clearly the β strand between residues Gly26 and Tyr29.

Disulfide Pattern Determination. First we tried to establish the connectivity pattern among cysteine residues in chlorotoxin by enzymatic methods, which had proved to be successful in the cases of charybdotoxin and PO_5. However, in our case, we were not immediately able to obtain a good proteolytic cleavage and to isolate useful fragments. The failure of this method did not completely surprise us, as it had also been encountered in scyllatoxin, iberiotoxin, and the ragweed allergen Amb t V (unpublished results of our laboratory for scyllatoxin; Johnson & Sug, 1992 ; Metzler *et al.*, 1992).

We therefore attempted to use the NMR data to assign the disulfide pairs, basing our analysis on a recent study (Klaus *et al.*, 1993) which investigated the predictive value of different proton−proton distances for the presence of a disulfide bond. From an analysis of a large number of protein structures deposited in the Brookhaven Protein Data Bank, it was found that in more than 88% of the cases, a distance shorter than 5 Å between the β protons of the two cysteine residues under investigation is indicative of the presence of a disulfide bond between them, and this percentage increases to 95% when the $\beta-\beta$ proton distance is shorter than 4 Å. Moreover, if the $\beta-\beta$ distance shorter than 5 Å is combined with a short $\alpha-\beta$ distance (<5 Å) between protons of the two cysteine residues involved, this percentage increases to 96%.

We show in Figure 4a the aliphatic region corresponding to the β-proton resonances of chlorotoxin. A contact qualified as "medium" can be seen between two β protons from residues Cys20 and Cys35 corresponding most probably to a distance shorter than 4 Å for both protons. This distance limit was subsequently confirmed by the buildup curve of this peak for four different mixing times (50, 100, 150, and 200 ms). Calibration with respect to the average buildup of seven geminal proton pairs (see Methods) yielded a distance of 3.6 Å for this contact. In the region characteristic for the $\alpha-\beta$ contacts (Figure 4b), we observe next to the many intraresidue contacts, the corresponding interresidue contact between Cys^{35} α and Cys^{20} β_1, corresponding to a short distance of 3.4 Å after calibration. These two contacts confirm to a high degree of certainty the presence of a disulfide bridge between cysteines 20 and 35. A larger degree of uncertainty remains for the other disulfide bridges. We observed clear contacts between the α proton of Cys16 and both β protons of Cys19; however, no $\beta-\beta$ contacts were observed between both residues, and we did find an αN contact. These data indicate that the aforementioned contacts are the $\alpha(i)-\beta(i+3)$ contacts characteristic for an α-helical structure rather than an indication of a disulfide bridge between both residues.

In order to resolve the ambiguities over the assignment of the disulfide bridges, we generated structures with all possible disulfide bridges. This brute-force method has been applied earlier (Heitz *et al.*, 1989; Metzler *et al.*, 1992) and

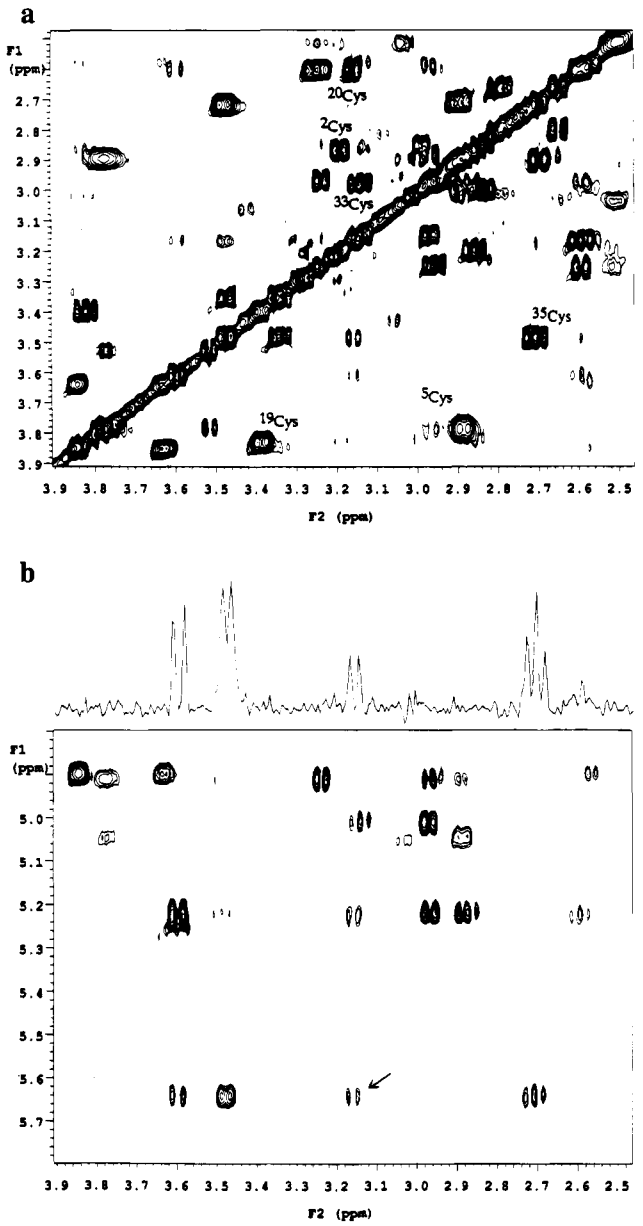

FIGURE 4: (a) Aliphatic region showing the β-proton resonances of a 200 ms NOESY spectrum of chlorotoxin (temperature = 36 °C, pH = 2.55). Indicated are the contacts between the geminal protons of the different cysteine residues (with an increased line width for the Cys5 $\beta_1-\beta_2$ contact) and the NOE contact between the β protons of Cys20 and Cys35 (arrow). (b) Region of the same spectrum as in (a) connecting the α and β protons of chlorotoxin. The trace is taken at the chemical shift of Cys35 αH, and the cross peak toward the β_2 proton of Cys20 at 3.14 ppm is indicated by an arrow.

allows one to look for those disulfide pairing that are compatible with *all* distance constraints, rather than only those concerning protons on the cysteine residues. Even though the method is conceptually straightforward, the most important drawback is the computational cost (see Methods). For a protein with eight disulfides, the number of possible combinations is 105, and the generation of 10 structures for each of them did require a full week of CPU time on a Silicon Grapics R3000 machine. The outcome of these calculations yielded a surprisingly large number of structures with no NOE violations larger than 1 Å (40 out of 105), but the majority of them had unacceptably large values of van der Waals and NOE constraint energies. When placing a cutoff of 0.3 Å on the NOE violations and 70 kcal/mol for the total energy after the X-PLOR refinement step, only the above mentioned disulfide pairing lead to two accepted structures (out of ten). Although this again confirms the proposed disulfide bridge pattern, the low overall yield of accepted structures suggests that one should generate more structures per permutation when trying to obtain the correct pattern solely by this brute-force method.

Table 2: Energetic Statistics of the Seven Final Structures after a Nonconstrained Steepest Descent Minimization with the CHARMM Force Field Using the BRUGEL Package[a]

av energies	kcal/mol
total	12.8 ± 8
van der Waals	−163 ± 6
NOE constraints	10.1 ± 1.5
electrostatic	33.6 ± 4.4
bonds	6.2 ± 0.4
bond angles	25.5 ± 1.9
torsions	116.8 ± 11.8
planar + tetragonal	3.5 ± 0.6

[a] Delhaise et al., 1985. NOE constraint terms were calculated using a split harmonic energy term with a force constant of 50 kcal/(mol Å^2).

In conclusion, the presence of the Cys20−Cys35 disulfide bridge as verified by the NMR data, the identity of the secondary structure elements and their relative ordering between chlorotoxin and charybdotoxin, and more generally, the evident alignment of the primary structures of both molecules indicate that the cysteine pairing in chlorotoxin (and probably also in homologuous toxin I_5A) does follow the common motif, with covalent links between Cys16−Cys33, Cys5−Cys28, and Cys20−Cys35. The two additional cysteines Cys2 and Cys19 in chlorotoxin are linked together by a disulfide bridge, which cross-links the N-terminal strand of the β-sheet to the α-helix.

Structure Calculation by Distance Geometry and Simulated Annealing. A total of 100 structures with the above described disulfide pairing was generated with the program X-PLOR. Of the 100 structures, 45 had no violations larger than 1 Å, 28 passed the test with no violations larger than 0.5 Å, and 16 with an upper limit of 0.3 Å for the violations. Considering the relative success of the structure generation, we decided to use the 0.3 Å as cutoff. A last selection, however, was based on an energetic criterium: of the 16 structures, 7 had a total potential energy of less than 70 kcal/mol, whereas the other ones had energies ranging from 75 up to 118 kcal/mol. Because the major contribution of the high total energies came predominantly from the van der Waals and NOE terms (despite the absence of any distance violation larger than 0.3 Å, the latter structures accumulated a large number of distance violations between 0.15 and 0.3 Å), and because visual inspection of those structures indicated that they were not substantially different from the 7 lower-energy structures, we decided to limit our structural analysis to the 7 structures that displayed no NOE violation larger than 0.3 Å and had a total potential energy below 70 kcal/mol. In order to evaluate their conformational energies using a full force field, we introduced them into the BRUGEL package (Delhaise *et al.*, 1985) and subjected them to 200 steps of steepest descent reminimization in absence of NOE constraints. The structures hardly changed, but their energies did improve considerably. In Table 2, we summarize the conformation energies as averaged over the 7 structures. The

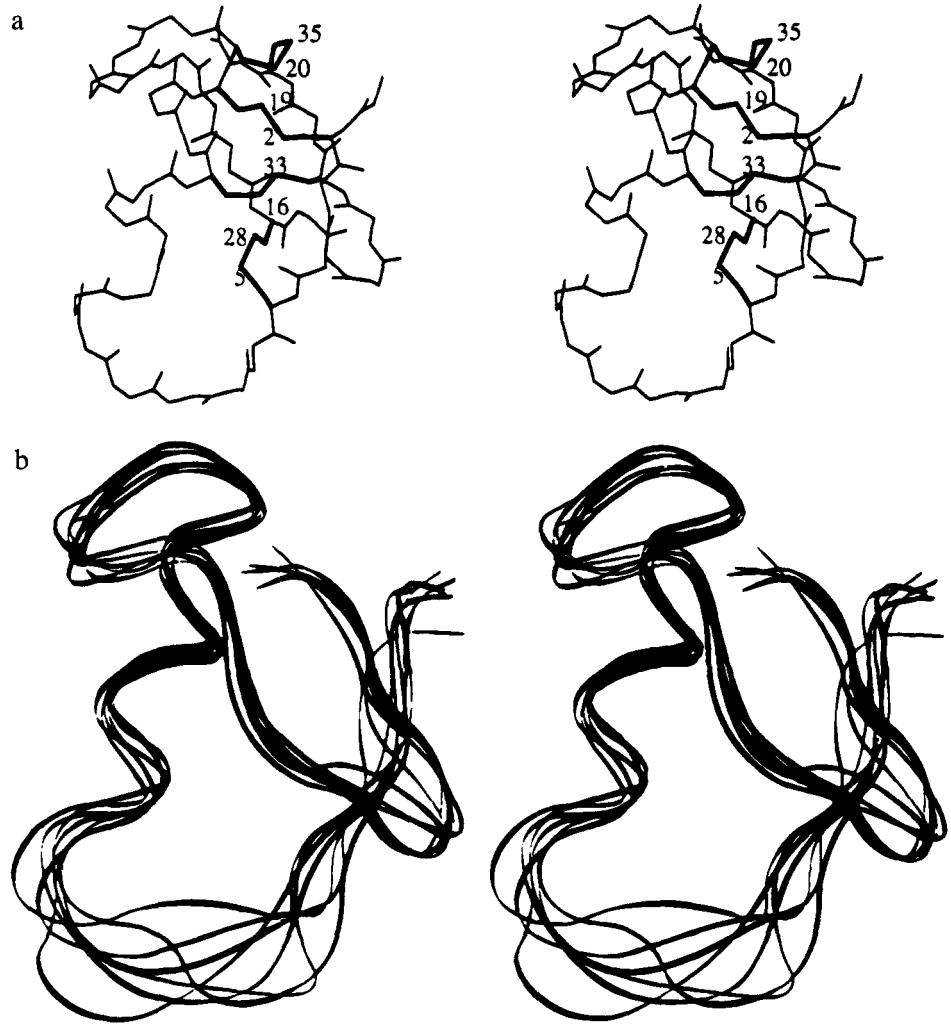

FIGURE 5: (a) Stereoview of the structure of chlorotoxin. The eight cysteine residues are shown in bold. (b) Superposition of the best seven chlorotoxin molecular structures. A ribbon has been drawn through every structure of the final set (see text).

27 accurate distance limits determined by integration of the cross-peaks and calibration with respect to the averaged integrated values over 7 geminal β protons (see Methods) were used to verify the final structures; when admitting deviations of 0.3 Å of those distances, no violations were observed for the seven final structures.

Analysis of the Final Structures. In Figure 5a, we show one individual structure of chlorotoxin, indicating the four disulfide pairs in bold. Figure 5b shows a superposition of the backbones of the 7 best structures after distance geometry and simulated annealing and using all backbone atoms for the superposition. The RMSD for all backbone atoms is 1.07 Å. It increases to 1.8 Å when all non-hydrogen atoms are included in the RMSD calculation. The different secondary structure elements are particularly well-defined: we measured backbone RMSD values of 0.7 Å for the α-helix, 0.86 Å for the C-terminal β-sheet (from residues 26 to 36), and 1.0 Å for the short N-terminal β-strand (residues 1 to 4). The turn linking this first β-strand to the α-helix is however less well-defined, with a backbone RMSD value of 2.0 Å. This last feature was not completely unexpected, as the number of NOE contraints in this region was smaller than in the regular secondary structure elements. Moreover, the line width of the β resonances of the Cys5 residue was a factor of 2.5 larger than that of the β resonances of the other cysteine residues (Figure 4a). This is probably due to exchange broadening, resulting from an increased mobility of the loop defined by residues 4 to 10.

DISCUSSION

(a) Structure. The secondary structure of chlorotoxin agrees well with those observed in the known 3D structures of the other short toxins: we find the characteristic fold where an α-helix is packed against a small β-sheet and is cross-linked to it by three disulfide bonds (Figure 5a,b). However, when overlaying the structures of chlorotoxin and the other members of the short toxin family, some interesting differences were observed. In Figure 6, we show the structures of chlorotoxin and charybdotoxin (Bontems *et al.*, 1991), where the three common disulfide bridges and the C-terminal β-sheet are superimposed. A first remarkable difference is found at the level of the α-helix: whereas the central part of the helix with the characteristic C-X-X-X-C motif is highly conserved between both molecules, the helix in chlorotoxin is shifted by one turn toward the N-terminus. This shift is already clear from the primary structure comparison (Figure 1b): the optimal alignment between chlorotoxin and charybdotoxin or iberiotoxin contains four deletions in the latter sequences, positioned before the central cysteines in the helix.

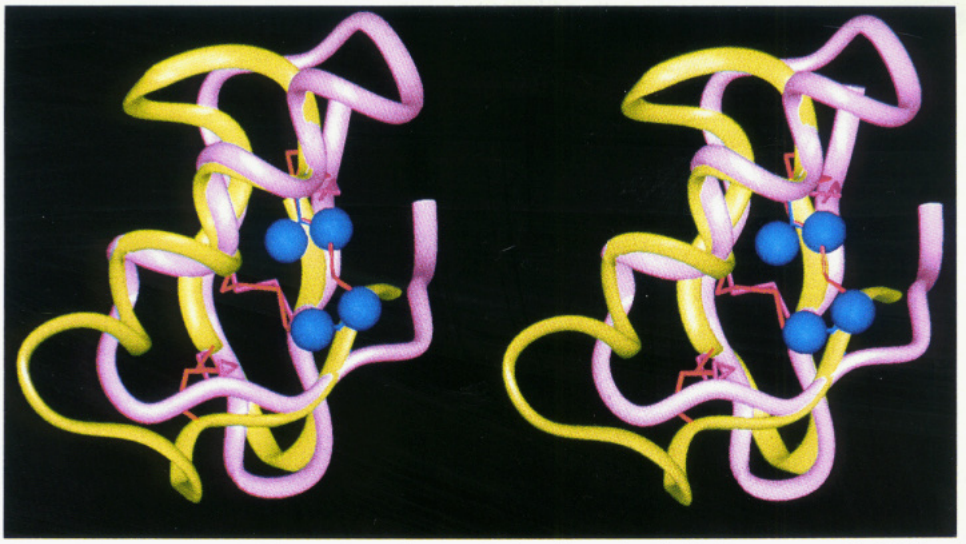

FIGURE 6: Superposition of the structures of chlorotoxin (yellow) and charybdotoxin (pink) (Bontems *et al.*, 1991), where the C-terminal β-sheet and the three disulfide bridges in common are superimposed. The disulfide bridges of chlorotoxin are shown in red and those of charybdotoxin in magenta. The two valines of charybdotoxin replacing the Cys2−Cys19 bridge in chlorotoxin are shown as full blue spheres (CPK representation).

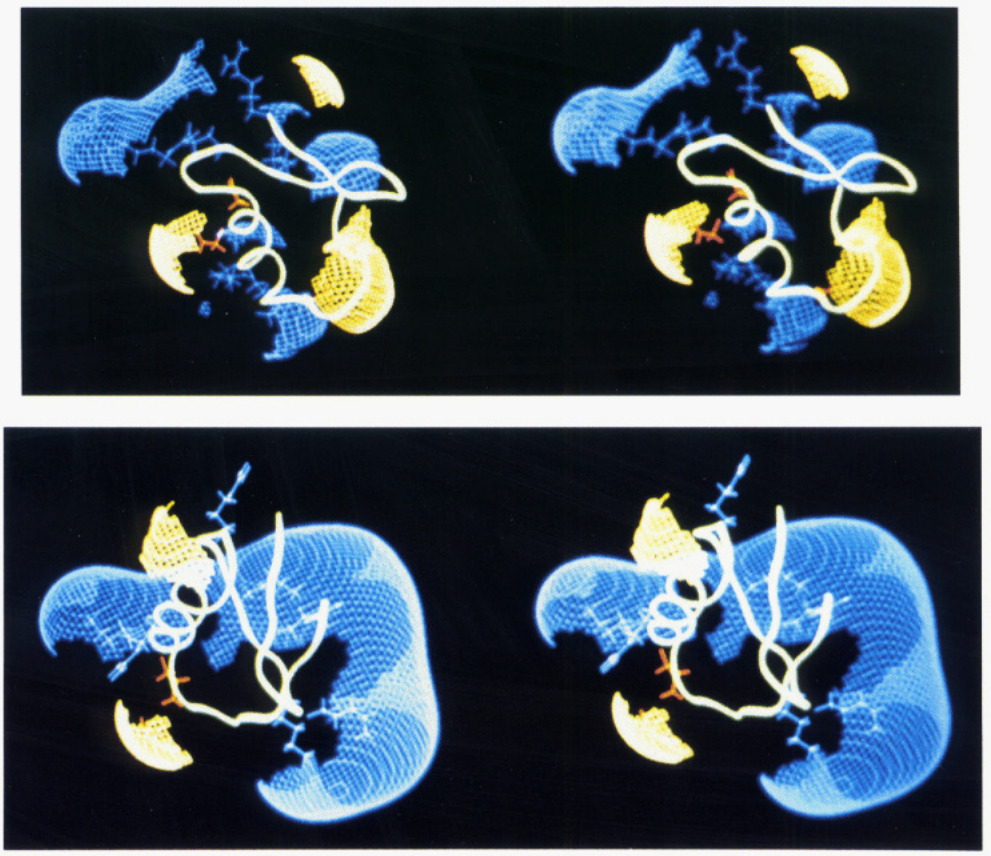

FIGURE 7: Electrostatic surface of chlorotoxin (a, top) and charybdotoxin (b, bottom) as calculated by the program DELPHI. Different color codes are -2 *kT*/e (red), -1 *kT*/e (yellow), and 1.5 *kT*/e (blue) (where *k* is the Boltzmann constant, *T* is the temperature in kelvin, and e is the charge of the electron. With *T* set to 298 K, *kT*/e amounts to 25.6 mV or 0.6 kcal/mol of electrons). Positively charged residues are shown in blue and negatively charged in red.

The different extensions of the central part of the helix have their influence on the two turns connecting the helix with the N- and C-terminal β-strands. The N-terminal $\beta\alpha$ turn of chlorotoxin is wider than that of charybdotoxin and is simultaneously characterized by a larger mobility (*vide supra*).

The $\alpha\beta$ turn connecting the α-helix with the C-terminal β-strand adopts a completely different orientation in both molecules. In charybdotoxin, the turn is oriented toward the N-terminal β-strand. This orientation is not fortuitious but serves to bring the hydrophobic residues of the turn (Leu20, His21, and Thr23) in close contact with the first N-terminal residue of charybdotoxin. In this manner, those residues form a hydrophobic cluster linking the first β-strand to the rest of the molecule (Bontems, 1992). This Leu together with the aromatic nature of the following residue is strictly conserved between charybdotoxin, iberiotoxin, and noxiustoxin, and the orientation of the turn is identical in all of

them. The primary structures of leiurotoxin and PO5 deviate slightly from this pattern, because of the replacement of two residues in the turn by one Gly. The $\alpha\beta$ turn is shortened due to this modification but is still oriented in the same direction as in charybdotoxin. The hydrophobic cluster in these molecules is not formed by the residues of the (lacking) first β-strand but by three Leu residues in the turn itself (Martins, 1993). Chlorotoxin and its direct analogues (Figure 1a), however, deviate substantially from this common pattern: the turn is formed by a triple (or double) repeat of the dipeptide Gly-Arg(Lys), representing thereby a positively charged rather than a hydrophobic cluster.

The different orientation of the $\alpha\beta$ turn leads to a loss of hydrophobic contacts compared to charybdotoxin, but the first β-strand in chlorotoxin is linked to the rest of the molecule by the fourth disulfide bridge between Cys2 and Cys19. This fact is illustrated even more clearly when we compare the cysteines with their corresponding side chains in charybdotoxin , i.e., Val5 and Val16. In Figure 6, we show both valines as full spheres (CPK representation), and one can see that they occupy exactly the same space as the fourth disulfide bridge in chlorotoxin. This also fits well with the lack of the first two residues in chlorotoxin: in charybdotoxin, they are necessary from a structural point of view as they allow the hydrophobic contact with the $\alpha\beta$ turn. In chlorotoxin, however, these structural requirements seem to be fulfilled by the extra cysteine bridge.

(b) Structure–Function Relationships. It has been shown in several cases that electrostatic effects are of uttermost importance for the interaction between the channels and the small basic peptides that are the scorpion toxins (Meunier *et al.*, 1993). The calculated electrostatic potential surface around the PO5 molecule was found to be highly asymmetric. In this molecule, three arginine residues, all located on the solvent-exposed side of the helix which is believed to interact with the receptor, form a positively charged potential surface centered around the neutral polar group of Gln9. The influence of the electrostatic contribution to the charybdotoxin specificity was investigated by a site-directed mutagenesis study, where each of the positively charged groups was replaced by a glutamine, and the resulting changes in the toxin's association–dissociation kinetics with single K^+ channels were examined (Park & Miller, 1992). As a first step in investigating the possible role of electrostatics for the chlorotoxin function, we calculated its electrostatic potential and compared it with the results of identical calculations on the charybdotoxin structure.

The electrostatic surfaces for both molecules as calculated by the program DELPHI are shown in Figure 7. Their inspection allows to rationalize the results from the site-directed mutagenesis experiments on charybdotoxin: around the three critical residues Arg25, Lys27, and Arg34, we see a positively charged surface that forms half a torus around the central part of the C-terminal β-sheet. In the case of chlorotoxin, this positive surface has almost completely disappeared, even though the first two positive residues (Arg25 and Lys 27) are strictly conserved. However, their effect seems to be attenuated by the presence of two negatively charged aspartic acid residues on the α-helix. The third positively charged residue (Arg34) of charybdotoxin is replaced by a hydrophobic leucine in chlorotoxin, and this leucine is conserved over all its direct analogues (Figure 1a). It is also at the position of this residue that we find a glutamic acid in the case of leiurotoxin and PO5, with consequently a shift of the positively charged surface from the C-terminal β-sheet side exposed to the solvent to the solvent-exposed side of the helix (Meunier *et al.*, 1993).

As discussed above, chlorotoxin and charybdotoxin differ significantly in the region of the $\alpha\beta$ turn. This has also its consequences on the electrostatic properties in this zone: whereas it forms a neutral to slightly negative zone in charybdotoxin, the presence of the three posive residues in chlorotoxin gives rise to a positive "cap" of the molecule that extends over a considerable distance into the solvent. Although we have no direct data of site-directed mutagenesis affecting those residues, this clustering of positively charged residues and the strong positive field they produce might well have important implications on the functional properties of this molecule.

ACKNOWLEDGMENT

We thank LATOXAN, France, for their generous gift of the scorpion venom and Dr. Bontems and Dr. Martins for providing us with a copy of their Ph.D. thesis. Prof. F. Borremans provided us generously with the coordinates of leiurotoxin. We thank D. Horvath for generating the permutations of the cysteine patterns, C. Cerf for help with the DELPHI program, Dr. M. Froeyen for providing his routine of visualization of electrostatic potential surfaces, and A. Waeghe for the photography.

REFERENCES

Arseniev, A. S., Kondakov, V. I., Maiorov, V. N., Volkova, T. M., Grishin, E. V., Bystrov, V. F., & Ovchinnikov, Y. A. (1983) *Bioorg. Khem. (USSR) 9*, 768.

Arseniev, A. S., Kondakov, K. I., Maiorov, V. N., & Bystrov, V. F. (1984) *FEBS Lett. 165*, 57–62.

Auguste, P., Hugues, M., Gravé, B., Gesquière, J. C., Maes, P., Tartar, A., Romey, G., Schweitz, H., & Lazdunski, M. (1990) *J. Biol. Chem. 265*, 4753–4759.

Bax, A., & Davis, D. (1985) *J. Magn. Reson. 65*, 355–360.

Betz, H. (1990) *Biochemistry 29*, 3591–3599.

Bontems, F. (1992) Ph.D. Thesis, University of Paris Sud, Centre d'Orsay, France.

Bontems, F., Roumestand, C., Boyot, P., Gilquin, B., Doljanski, 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.

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535–560.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

Brünger, A. T. (1992) *X-PLOR Version 3.1 Manual*, Yale University, New Haven, CT.

Cook, N. S., & Quast, U. (1990) in *Potassium channels*, Chapter 8, pp 181–255, Ellis Horwood Limited, Chichester U.K.

Crest, M., Jacquet, G., Gola, M., Zerrouck, H., Benslimane, A., Rochat, H., Mansuelle, P., & Martin-Eauclaire, M. F. (1992) *J. Biol. Chem. 267*, 1640.

Debin, J. A., Maggio, J. E., & Strichartz, G. R. (1993) *Am. J. Physiol. Soc. 264*, C369.

Delhaise, P., Van Belle, D., Bardiaux, M., Alard, P., Hamers, P., Van Cutsem, E., & Wodak, S. J. (1985) *J. Mol. Graphics 3*, 116–119.

Dreyer, F. (1990) *Rev. Physiol., Biochem., Pharmacol. 115*, 93–136.

Fazal, A.,. Beg, O. V, Shafgat, J., Zaidi, Z. H., & Jornvall, H. (1989) *FEBS Lett. 257*, 260−262.

Fontecilla-Camps, J. C., Almassy, R. J., Suddath, F. L., & Bugg, C. E. (1982) *Toxicon 20*, 1−7.

Galvez, A., Gimenez, G., Reuben, J. P., Ray-Contanain, L., Feigenbaum, P., Kaczorowski, G. J., & Garcia, M. L. (1990) *J. Biol. Chem. 265*, 11083.

Garcia, M. L., Galvez, A., Garcia-Calva, M., King, F., Vasquez, J., & Kaczorowski, G. J. (1991) *J. Bioenerg. Biomembr. 2*, 615−645.

Gilson & Honig, B. (1988) *Proteins 4*, 7−18.

Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1985) *J. Am. Chem. Soc. 107*, 6394−6396.

Grishin, E. V., Soldatov, N. M.,Tashmuchamedov, B. A., & Atakuziev, B. U. (1978) *Bioorg. Khem. (USSR) 4*, 450−461.

Grishin, E. V., Volkova, T. M., & Soldatova, L. N. (1982) *Bioorg. Khem. (USSR) 8*, 155−164.

Jeener, J., Meyer, P. H., Bachmann, P., & Ernst, R. R., (1979) *J. Chem. Phys. 71*, 4546−4553.

Johnson, B. A., & Sugg, E. E. (1992) *Biochemistry 31*, 8151−8159.

Klaus W., Broger, C., Gerber, P. & Senn, H. (1993) *J. Mol. Biol. 232*, 897.

Lambert, P., Kuroda, H., Chino, N., Watanabe, T. X., Kimura, T., & Sakakibara, S. (1990) *Biochem. Biophys. Res. Commun. 170*, 684−690.

Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.

Martins, J. C. (1993) Ph.D. Thesis, University of Ghent, Ghant, Belgium.

Martins, J. C., Zhang, W., Tartar, A., Lazdunski, M., & Borremans F. A. M. (1990) *FEBS Lett. 260*, 249−253.

Ménez, A., Bontems, F., Roumestand, C., Gilquin, B., & Toma, F. (1992) *Proc. R. Soc. Edinburgh 99B* (1/2), 83−103.

Meunier, S., Bernassau, J.-M., Martin-Eauclaire, M. F., Van Rietschoten, J., Cambillau, C., & Darbon, H. (1993) *Biochemistry 32*, 11969.

Miller, C., Moczydlowski, E., Lattore, R., & Philips, M. (1985) *Nature 313*, 316−318.

Moczydlowski, E., Lucchesi, K., & Ravindran, A. (1988) *J. Membr. Biol. 105*, 95−111.

Nutt, R. F., Arison, B. H. & Smith, J. S. (1992) *Peptides*, 101−102.

Park, C.-S., & Miller, C. (1992) *Biochemistry 31*, 7749−7755.

Possani, L. D., Martin, B. M., & Svendsen, I. (1982) *Carlsberg Res. Commun. 47*, 285.

Rochat, H., Bernard, P., & Courand, F. (1979) in *Neurotoxins: Tools in Neurobiology* (Ceccarelli, B., Ed.) pp 325−334, Raven, New York.

Sabatier, J.-M., Zerrouk, H., Darbon, H., Mabrouk, K., Benslimane, A., Rochat, H., Martin-Eau-claire, M.-F., & Van Rietschoten, J. (1993) *Biochemistry 32*, 2763−2770.

Strong, P. N. (1990) *Pharmacol. Ther.* 46, 137−162.

Veber, D. F., Mikowski, J. D., Varga, S., Denkewalter, R. G., & Hirschmann, R. (1972) *J. Am. Chem. Soc. 94*, 5456−5461.

Wüthrich, K., (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.

Zlotkin, E., Miranda, F., & Rochat, H. (1978) in *Arthropod Venoms* (Bettini, S., Ed.) pp 317−369, Springer-Verlag, Berlin.

BI941215M