J. Mol. Biol. (1993) **234**, 735–750

Nuclear Magnetic Resonance Solution Structure of Dendrotoxin K from the Venom of *Dendroaspis polylepis polylepis*

Kurt D. Berndt, Peter Güntert and Kurt Wüthrich

Institut für Molekularbiologie und Biophysik
Eidgenössische Technische Hochschule-Hönggerberg
CH-8093 Zürich, Switzerland

(Received 1 March 1993; accepted 9 July 1993)

The solution structure of dendrotoxin K (Toxin K), a protein consisting of one polypeptide chain with 57 residues purified from the venom of the black mamba, *Dendroaspis polylepis polylepis*, was determined by nuclear magnetic resonance (NMR) spectroscopy. On the basis of virtually complete sequence-specific ^{1}H NMR assignments, including individual assignments for 38 pairs of diastereotopic substituents and side-chain amide protons, a total of 818 nuclear Overhauser effect distance constraints and 123 dihedral angle constraints were identified. Using this input, the solution structure of Toxin K was calculated with the program DIANA, and refined by restrained energy-minimization with a modified version of the program AMBER. The average root-mean-square deviation (r.m.s.d.) relative to the mean atomic co-ordinates of the 20 conformers selected to represent the solution structure is 0·31 Å for all backbone atoms N, C^α and C′, and 0·90 Å for all heavy-atoms of residues 2 to 56. The solution structure of Toxin K is very similar to the solution structure of the basic pancreatic trypsin inhibitor (BPTI) and the X-ray crystal structure of the α-dendrotoxin from *Dendroaspis angusticeps* (α-DTX), with r.m.s.d. values of 1·31 Å and 0·92 Å, respectively, for the backbone atoms of residues 2 to 56. Some local structural differences between Toxin K and BPTI are directly related to the fact that intermolelcular interactions with two of the four internal molecules of hydration water in BPTI are replaced by intramolecular hydrogen bonds in Toxin K.

Keywords: dendrotoxin K; nuclear magnetic resonance spectroscopy; protein structure determination; proteinase inhibitors; internal hydration water

1. Introduction

The dendrotoxins are a homologous group of pharmacologically active proteins isolated from the venoms of the black mamba (*Dendroaspis polylepis polylepis*) and the eastern green mamba (*Dendroaspis angusticeps*). Originally identified on the basis of sequence homology as members of the Kunitz proteinase inhibitor family, subsequently they were found to have negligible trypsin and chymotrypsin inhibitor activities (Strydom, 1973). As no physiological role could be attributed to these proteins at that time, they became known simply as "trypsin inhibitor homologues" (Joubert & Strydom, 1978). The mode of action of these toxins was discovered only later when it was shown that contrary to the neurotoxins, which are also present in these venoms and block neuromuscular transmission, dendrotoxins facilitate the release of acetylcholine from presynaptic membranes of cholinergic synapses (Harvey & Karlsson, 1982). Further evidence suggests that this action is mediated by the ability of dendrotoxins to specifically block potassium channels (Harvey & Anderson, 1985). For this reason, dendrotoxins have become important tools for the analysis of voltage-dependent potassium channels (Penner *et al.*, 1986). The presently reported NMR structure of Toxin K† is the first

† Abbreviations used: Toxin K, dendrotoxin K (or trypsin inhibitor homologue K) from the venom of the black mamba, *Dendroaspis polylepis polylepis*; BPTI, bovine pancreatic trypsin inhibitor; α-DTX, α-dendrotoxin from the venom of the eastern green mamba, *Dendroaspis angusticeps*; TFA, trifluoroacetic acid; 2D, two-dimensional; 2QF-COSY, 2D two-quantum-filtered correlation spectroscopy; 2Q-spectroscopy, 2D two-quantum spectroscopy; TOCSY, 2D total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; E.COSY, 2D exclusive correlation spectroscopy; $^3J_{HN\alpha}$, vicinal spin–spin coupling constant between the amide proton and the α-proton; $^3J_{\alpha\beta}$, vicinal spin–spin coupling constant between the α-proton and a β-proton; $^3J_{\beta H\gamma OH}$, vicinal spin–spin coupling constant between the β-proton and the γ-hydroxyl proton in Ser or Thr; REDAC, use of redundant dihedral angle constraints; r.m.s.d., root-mean-square deviation; c.p.u., central processing unit, p.p.m., parts per million.

0022–2836/93/230735–16 $08.00/0

three-dimensional structure determination of a dendrotoxin in solution, an X-ray crystal structure of α-DTX having just appeared (Skarzynski, 1992).

Toxin K was one of the first proteins for which sequence-specific ^1H NMR assignments were obtained (Keller *et al.*, 1983). Furthermore, measurements of backbone–backbone sequential and medium range NOEs, $^3J_{HN\alpha}$ coupling constants and amide proton exchange rates showed that the solution structure of this protein contains the same secondary structures as BPTI (Pardi *et al.*, 1983), which confirmed earlier predictions based on analysis of sequence homologies (Swenson *et al.*, 1978). In view of these results it appeared unwarranted in 1983, to make the then considerable effort of a complete NMR structure determination of Toxin K. In the meantime, several new developments have made it worthwhile to reconsider this decision. First, NMR structures of small proteins in solution can now be determined with high precision. Considering that it is clear from the earlier studies (Pardi *et al.*, 1983) that Toxin K and BPTI adopt the same global polypeptide fold, and that the sequence identity between the two proteins is only 41% (Table 1), it will be of interest to compare at high resolution how the two sequences are accommodated in the same global fold. Second, there are indications that some of the amino acid substitutions between BPTI and Toxin K involve residues next to internal hydration sites, which provides an opportunity for investigating the effects of such mutations on internal hydration (see also Berndt *et al.*, 1993). Third, a domain of the Alzheimer's amyloid β-protein precursor was found to have a Kunitz inhibitor-type structure (Hynes *et al.*, 1990), so that detailed comparisons with this biologically important protein will be of keen interest. These different points will be approached further in the Discussion.

2. Materials and Methods

(a) *Protein preparation*

The Toxin K was isolated from the venoms of either the South African or Tanzanian black mamba, *Dendroaspis polylepis polylepis*, (Latoxan) and batch-purified as described previously (Strydom, 1972). Toxin K obtained from these two venom sources was found to be identical. Residual protein impurities were removed by two subsequent chromatography steps. The protein fraction containing roughly 90% Toxin K was first subjected to ion exchange chromatography using a Mono-S 10/10 column (Pharmacia) equilibrated with 40 mM Hepes buffer (pH 8·0) containing 0·45 M NaCl. The Toxin K was eluted under isocratic conditions. Desalting of the protein was performed by repeated ultrafiltration with water on an Amicron YM-2 membrane after adjustment of the pH to 4·0 to minimize protein–buffer ion pair interactions. The second step involved a conventional reverse phase chromatographic separation using a Nucleopore C_8 column and a linear gradient of 0·1% TFA in water to 0·1% TFA in acetonitrile. Analytical reverse phase HPLC of the purified inhibitor yielded a single, symmetrical peak as monitored by the absorbance of the effluent at 280 nm.

(b) *Nuclear magnetic resonance experiments*

All data used for the structure calculation were obtained from spectra recorded from samples containing 10 mM Toxin K either in a mixture of 95% H_2O/5% ^2H_2O, or in 99·99% ^2H_2O after complete exchange of all labile protons. The pH was adjusted to 4·6 by the addition of minute amounts of NaOH and HCl, or NaO^{2}H and ^{2}HCl, respectively, and the temperature was 36·0(±0·5)°C. The NMR spectra were recorded on either a Bruker AM500 or AM600 spectrometer using the phase-sensitive mode with time-proportional phase incrementation of the initial pulse (Marion & Wüthrich, 1983). Quadrature detection was used in both dimensions, with the carrier placed in the center of the spectrum. To complete the previously reported sequence-specific resonance assignments (Keller *et al.*, 1983) with the identification of the complete spin systems of the long amino acid side-chains, standard homonuclear ^1H–^1H correlation experiments were recorded (Wüthrich, 1986) and analyzed with the interactive program package EASY (Eccles *et al.*, 1991).

NOE upper bounds on ^1H–^1H distances were obtained from 3 NOESY experiments (Anil-Kumar *et al.*, 1980) with mixing times, τ_m, of 40 ms: a NOESY spectrum in H_2O with 920 points in t_1 and 4096 points in t_2 (t_{1max} = 60 ms, t_{2max} = 270 ms) with selective suppression of zero-quantum coherences (Otting, 1990; Rance *et al.*, 1985); a NOESY spectrum in ^2H_2O with 758 points in t_1 and 4096

Table 1

Amino acid sequences of Toxin K and selected homologous proteins

	1	10	20	30	40	50	57
Toxin K	AAKYCKLPLRIGPCKRKIPSFYYKWKAKQC	LPFDYSGCGGNANRFKTIEECRRTCVG					
C13S1C3	AAKYCKLPVRYGPCKKKIPSFYYKWKAKQC	LPFDYSGCGGNANRFKTIEECRRTCVG					
Toxin I	ZPLRKLCILHRNPGRCYQKIPAFYYNQKKKQC	EGFTWSGCGGNSNRFKTIEECRRTCIRK					
α-DTX	ZPRRKLCILHRNPGRCYDKIPAFYYNQKKKQC	ERFDWSGCGGNSNRFKTIEECRRTCIG					
BPTI	RPDFCLEPPYTGPCKARIIRYFYNAKAGLC	QTFVYGGCRAKRNNFKSAEDCMRTCGGA					
APPI	VREVCSEQAETGPCRAMISRWYFDVTEGKC	APFFYGGCGGNRNNFDTEEYCMAVCGSA					

Toxin K is from *Dendroaspis polylepis polylepis* (Strydom, 1973), C13S1C3 from *Dendroaspis angusticeps* (Joubert & Taljaard, 1980), Toxin I from *Dendroaspis polylepis polylepis* (Strydom, 1973), and α-DTX from *Dendroaspis angusticeps* (Joubert & Taljaard, 1980); BPTI is the bovine pancreatic trypsin inhibitor (Kassell *et al.*, 1965) and APPI the human Kunitz-type protease inhibitor domain of Alzheimer's amyloid β-protein precursor (Ponte *et al.*, 1988). In Toxin I and α-DTX, Z stands for pyroglutamate.

points in t_2 (t_{1max}=50 ms, t_{2max}=27 ms) with zero-quantum coherences shifted and suppressed (Otting *et al.*, 1990); a soft-NOESY spectrum (Billeter *et al.*, 1990*a*; Brüschweiler *et al.*, 1988) in ^2H_2O with 470 points in t_1 and 4096 points in t_2 (t_{1max}=t_{2max}=130 ms), with the selective pulse centered at 4·30 p.p.m. Vicinal spin–spin coupling constants $^3J_{HN\alpha}$ were determined from the aforementioned NOESY spectrum in H_2O by inverse Fourier transformation of the in-phase multiplets (Szyperski *et al.*, 1992), and values for $^3J_{\alpha\beta}$ were obtained from an E.COSY spectrum (Griesinger *et al.*, 1985) recorded in ^2H_2O with a digital resolution of 2·2 Hz/point along ω_2. Time-domain data were processed on a Bruker X-32 data station using UXNMR software. Prior to Fourier transformation the NOESY data were weighted with a cosine window and zero-filled 2-fold in both dimensions. Baseline distortions in the NOESY spectra were eliminated using the FLATT procedure (Güntert & Wüthrich, 1992).

(c) *Determination of the 3-dimensional structure*

The input for the final structure calculation resulted from several rounds of NOESY cross-peak assignments and structure calculations with preliminary sets of input data (Berndt *et al.*, 1992; Güntert *et al.*, 1991*b*). Two different "calibration curves" describing relationships between cross-peak volumes, V, and corresponding upper distance limits, b were used with the program CALIBA (Güntert *et al.*, 1991*b*). A $1/b^6$ dependence was used for the 405 backbone proton–backbone proton or backbone proton–β-proton NOESY cross-peaks , and a $1/b^4$ dependence for the 484 cross-peaks with side-chain protons and the 191 cross-peaks with methyl groups. To obtain resonable upper bounds for both strong and weak NOEs, the proton–proton upper distance bounds were in all cases confined to the range 2·4 to 4·5 Å before the application of appropriate pseudo-atom corrections (Wüthrich *et al.*, 1983). For the final DIANA calculation with Toxin K, refined calibration curves were determined based on plots of the NOESY cross-peak volumes *versus* the average of the corresponding proton–proton distances in a preliminary set of 20 conformers. The structure calculations with the program DIANA were performed using the REDAC strategy (Güntert & Wüthrich, 1991) for improved convergence. Each round of calculations was started with 50 random conformations, and the 20 conformers with the lowest target function values were used to represent the molecular structure. The 20 DIANA conformers with the smallest target function values after the final round of calculations were subjected to restrained energy minimization using a modified version of the program AMBER (Singh *et al.*, 1986), which includes pseudo-energy terms for distance constraints and dihedral angle constraints (Billeter *et al.*, 1990*b*). The resulting 20 energy-minimized DIANA conformers were used to represent the solution structure of Toxin K.

(d) *Structure comparisons*

For visual comparison of different structures, stereo views were produced with the molecular graphics programs MIDAS (Ferrin *et al.*, 1988) or XAM (Xia, 1992). Global superpositions and pairwise r.m.s.d. values for various subsets of atoms were computed as usual (McLachlan, 1979). The mean solution structure was obtained by first superimposing the 20 energy-minimized DIANA conformers so as to minimize the r.m.s.d. for the backbone atoms N, C^α and C' of the residues 2 to 56, and then averaging the Cartesian co-ordinates of corresponding atoms in the superimposed conformers. Displacements, D (Billeter *et al.*, 1989), were used to quantify the local precision of the solution structure and to identify local differences between crystal and solution structures. Displacements are a generalization of the r.m.s.d. values, since the set of atoms used for the superposition of the conformers, M_{sup}, differs from the set of atoms for which the root-mean-square deviation of the positions is actually calculated, $M_{r.m.s.d.}$. For example, for the evaluation of the backbone displacement of a given residue i in Toxin K after global superposition, D^{bb}_{glob}, M_{sup} consisted of the backbone atoms N, C^α and C' of the residues 2 to 56, and $M_{r.m.s.d.}$ of the backbone atoms N, C^α and C' of residue i. To evaluate local backbone displacements for a residue i, D^{bb}_{loc}, M_{sup} consists of the backbone atoms N, C^α and C' of the residues $i-1$, i and $i+1$, and $M_{r.m.s.d.}$ consists of the backbone atoms of residue i.

3. Results

(a) *^1H NMR assignments for Toxin K*

The previously reported sequence-specific resonance assignments at pH 3·4 and 50 °C (Keller *et al.*, 1983) were verified at the new experimental conditions of pH 4·6 and 36 °C used in this study. New, additional assignments were made for the backbone amide protons of residues Arg16, Gly17 and Gly40, and for long amino acid side-chains and labile side-chain protons. As a result, the present list of ^1H chemical shifts of Toxin K includes all non-labile protons with the sole exceptions of δCH_2 of Lys3 and Lys26, all backbone amide protons, and numerous labile side-chain protons (Table 2). The amide proton of Gly37 is shifted to high field at 4·57 p.p.m., which is a similar chemical shift to that of Gly37 in BPTI (Tüchsen & Woodward, 1987).

Hydroxyl protons of serine and threonine are known to have intrinsically different chemical shifts from that of water. Nonetheless, because of rapid exchange with the solvent they are usually not observed as separate signals in standard ^1H NMR experiments (Chazin & Wright, 1988; Otting & Wüthrich, 1989). In Toxin K the hydroxyl protons of Ser20 and Ser36 are observed in ^1H NMR spectra in H_2O (Table 2) even when using pre-irradiation of the solvent resonance. The assignment of these resonances was confirmed by intraresidual $C^\beta H$–$O^\gamma H$ cross-peaks in a 2QF-COSY spectrum, as well as by NH–$O^\gamma H$, $C^\alpha H$–$O^\gamma H$ and $C^\beta H$–$O^\gamma H$ cross-peaks in a TOCSY spectrum. These cross-peaks disappeared after complete exchange of the labile protons with deuterions. The observation of these hydroxyl protons under conditions of solvent pre-irradiation is indicative of low solvent accessibility and suggests their participation as proton donors in intramolecular hydrogen bonds.

(b) *Survey of the final set of conformational constraints and the final structure calculation*

A total of 1080 different NOE cross-peaks were unambiguously assigned and integrated in the three

Table 2

^1H chemical shifts for Toxin K at pH 4·6 and 36 °C

Residue	Chemical shift, δ (p.p.m.)[a]			
	NH	αH	βH	Others
Ala1		4·11	1·44	
Ala2	7·42	4·28	0·87	
Lys3	8·93	3·90	1·81, 1·68	γCH_2, 1·45, 1·50; εCH_2 3·00, 3·00
Tyr4	6·62	4·48	3·22, 2·96	δH 6·91; εH 6·91
Cys5	7·30	4·46	3·14, 3·10	
Lys6	6·93	4·24	1·17, 1·36	γCH_2 0·49, 0·73; δCH_2 0·98, 1·05; εCH_2 2·22, 2·29
Leu7	7·39	4·44	2·00, 1·63	γH 1·91; δCH_3 1·09, 0·99
Pro8		4·57	1·88, 2·36	γCH_2 2·11, 2·11; δCH_2 3·58, 3·80
Leu9	7·80	3·79	0·51, 1·14	γH 0·65; δCH_3 −1·07, 0·22
Arg10	7·78	4·61	1·80, 1·80	γCH_2 1·53, 1·61; δCH_2 3·18, 3·28; εNH 7·20; ηNH_2 6·68
Ile11	8·62	4·12	1·89	γCH_2 1·34, 1·63; γCH_3 1·35; δCH_3 1·01
Gly12	8·20	3·98, 5·26		
Pro13		4·78	2·22, 1·97	γCH_2 2·05, 2·05; δCH_2 3·61, 3·54
Cys14	7·60	4·59	2·64, 3·49	
Lys15	8·37	4·56	1·30, 1·63	γCH_2 2·09, 2·09; δCH_2 1·41, 1·41; εCH_2 2·95, 2·95; ζNH_3^+ 7·40
Arg16	7·94	4·40	1·44, 1·73	γCH_2 1·56, 1·66; δCH_2 3·13, 3·13; εNH 7·12; ηNH_2 6·68
Lys17	8·06	4·15	1·44, 1·62	γCH_2 1·06, 1·23; δCH_2 1·61, 1·61; εCH_2 2·91, 2·91
Ile18	8·98	4·52	1·80	γCH_2 1·01, 1·29; γCH_3 0·85; δCH_3 0·68
Pro19		4·49	1·99, 2·30	γCH_2 1·99, 2·15; δCH_2 4·00, 3·78
Ser20	8·83	4·97	2·46, 3·55	γOH 3·66
Phe21	9·07	5·86	2·63, 2·90	δH 6·75, 6·75; εH 7·39, 7·39; ζH 7·39
Tyr22	10·04	5·35	2·91, 3·00	δH 7·00, 7·00; εH 6·56, 6·56
Tyr23	10·25	4·73	2·96, 3·65	δH 7·40, 7·40; εH 6·46, 6·46
Lys24	8·28	4·38	1·28, 1·75	γCH_2 0·95, 1·27; δCH_2 1·53, 1·53; εCH_2 2·81, 2·99
Trp25	7·88	3·90	3·68, 3·33	δ^1H 7·41; ε^3H 7·52; ε^1NH 10·18; ζ^2H 7·56; ζ^3H 7·09; η^2H 7·25
Lys26	8·48	3·67	1·41, 1·54	γCH_2 0·99, 0·99; εCH_2 2·90, 2·90
Ala27	6·78	4·31	1·22	
Lys28	7·52	3·54	1·87, 1·99	γCH_2 1.23, 1.23; δCH_2 1·56, 1·56; εCH_2 2·84, 2·84
Gln29	6·34	4·61	1·50, 1·81	γCH_2 2·23, 2·23; εNH_2 7·21, 6·72
Cys30	8·60	5·47	3·29, 2·58	
Leu31	9·37	5·12	1·42, 1·69	γH 1·29; δCH_3 0·81, 0·93
Pro32		4·88	1·85, 2·00	γCH_2 2·16, 2·16; δCH_2 3·87, 3·93
Phe33	8·50	4·68	2·85, 3·19	δH 6·99, 6·99; εH 7·25, 7·25; ζH 6·91
Asp34	8·51	4·81	2·58, 1·91	
Tyr35	8·32	4·64	2·76, 2·76	δH 7·79, 6·73; εH 6·56, 6·68; ηOH 9·50
Ser36	7·96	4·40	3·82, 3·48	γOH 5·71
Gly37	4·57	2·95, 4·30		
Cys38	7·76	4·77	3·08, 3·77	
Gly39	8·54	3·88, 3·88		
Gly40	8·12	3·86, 4·21		
Asn41	8·56	4·73	3·24, 2·78	δNH_2 7·98, 8·33
Ala42	7·55	3·90	0·76	
Asn43	7·71	4·75	3·16, 2·99	δNH_2 7·71, 8·29
Arg44	6·78	4·80	1·40, 1·76	γCH_2 1·21, 1·34; δCH_2 2·35, 2·47; εNH 6·73; ηNH_2 6·06, 6·45
Phe45	9·65	5·01	2·59, 3·40	δH 7·20, 7·20; εH 7·80, 7·80; ζH 7·39
Lys46	9·37	4·59	2·09, 2·09	γCH_2 1·63, 1·73; δCH_2 1·80, 1·80; εCH_2 3·08, 3·08
Thr47	7·27	4·83	4·58	γCH_3 1·27
Ile48	8·23	3·05	0·55	γCH_2 0·88, 0·88; γCH_3 0·63; δCH_3 0·75
Glu49	8·35	3·85	1·83, 2·00	γCH_2 2·17, 2·39
Glu50	7·75	3·87	2·15, 2·23	γCH_2 2·37, 2·37
Cys51	7·05	2·00	2·81, 3·13	
Arg52	8·57	3·66	1·78, 1·69	γCH_2 1·46, 1·46; δCH_2 3·05, 3·05; εNH 7·12; ηNH_2 6·76
Arg53	8·21	3·95	1·75, 1·83	γCH_2 1·65, 1·65; δCH_2 3·14, 3·14; εNH 7·37
Thr54	7·36	3·93	3·84	γCH_3 1·36
Cys55	7·79	4·76	1·87, 1·44	
Val56	7·84	3·96	2·23	γCH_3 0·92, 0·92
Gly57	7·90	3·61, 3·80		

[a] For methylene groups, 2 chemical shifts are given only when 2 resolved resonances were observed, or when the presence of 2 degenerate signals had been established unambiguously by observation of a remote peak in a 2-quantum experiment in ^2H_2O. Individual assignments for pairs of diastereotopic substituents or side-chain amide protons of Asn or Gln obtained using the programs HABAS or GLOMSA are indicated by underlined chemical shift values (see text for details). For methylene protons or isopropyl methyl groups, the chemical shift of the proton or methyl group with the lower branch number is then listed first, e.g. the β^2 proton, or the γ^1 methyl in Val. For the side-chain amide protons of Asn and Gln, the first value is the chemical shift of the proton that has the shorter distance to the nearest methylene carbon (C^β for Asn, C^γ for Gln).

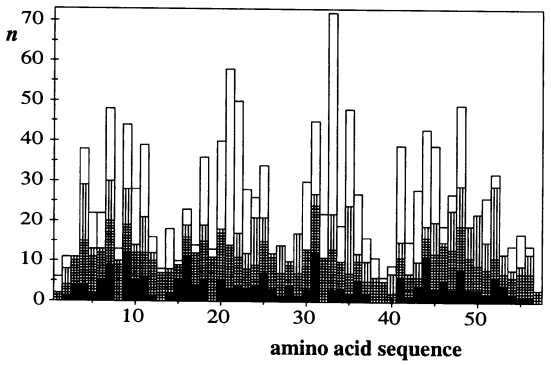

Figure 1. Plot of the number of NOE distance constraints per residue, *n*, *versus* the amino acid sequence of Toxin K (black, intraresidual; cross-hatched, sequential; vertically hatched, medium-range, white, long-range).

NOESY spectra used (see Materials and Methods). Stereospecific assignments for 35 pairs of diastereotopic substituents and individual assignments for the side-chain amide protons of one Gln and two Asn residues (Table 2) had been obtained using the programs HABAS (Güntert *et al.*, 1989) and GLOMSA (Güntert *et al.*, 1991*a*,*b*). After preprocessing of this data set with the program DIANA to eliminate irrelevant distance constraints (i.e. those that are either independent of the conformation or do not constrain the conformation) and to make appropriate corrections for NOEs to diastereotopic substituents without stereospecific assignments (Wüthrich *et al.*, 1983), the number of NOE upper distance constraints was reduced to 809 (Fig. 1). For all 47 non-glycine residues containing backbone amide protons the coupling constants $^3J_{HN\alpha}$ were measured in the same NOESY spectra that were used to collect the NOE distance constraints, using the method of Szyperski *et al.* (1992). The $^3J_{HN\alpha}$ values ranged between 3·5 and 10·2 Hz. Vicinal spin–spin coupling constants $^3J_{\alpha\beta}$ for 35 residues were obtained from an E.COSY spectrum in ^2H_2O solution (Griesinger *et al.*, 1985). Values of $^3J_{\alpha\beta}$ ranged between 1·9 and 12·6 Hz. The six residues Lys17, Lys26, Gln29, Glu49, Glu50 and Arg53 showed $^3J_{\alpha\beta}$ coupling constants indicative of conformational averaging about χ^1 (Nagayama & Wüthrich, 1981) and were not used to generate constraints for the structure calculation. In addition, all NOEs to $H^{\beta 2}$ and $H^{\beta 3}$ of these residues were referred to a pseudo-atom, Q^{β}, with the appropriate correction (Wüthrich *et al.*, 1983). From the treatment of the raw data with HABAS, the numbers of dihedral angle constraints were 44 for ϕ, 44 for ψ and 35 for χ^1. The three disulfide bonds 5–55, 14–38 and 30–51 were explicitly constrained by three upper and three lower distance limits each (Williamson *et al.*, 1985). No supplementary constraints were used to enforce hydrogen bonds during the DIANA structure calculations.

In the final DIANA calculation with the REDAC strategy (Güntert & Wüthrich, 1991), 47 out of 50 starting conformers led to a final target function value of less than 2·0 Å^2. This calculation for Toxin K took a total of 38 minutes of c.p.u. time on a Cray Y-MP computer, using one processor. The best 20 DIANA conformers all have target function values below 0·75 Å^2 and fulfill the NOE distance constraints and dihedral angle constraints almost perfectly (Table 3). The low AMBER energies before energy minimization show that bad steric overlaps were successfully removed already by DIANA. Subsequent restrained energy minimization with AMBER resulted in low-energy conformations with nearly identical molecular geometries, and with only a small increase of the sum of the residual constraint violations. The average of the 20 pairwise r.m.s.d. values between corresponding conformers before and after energy-minimization is

Table 3

Analysis of the 20 best DIANA conformers of Toxin K before and after restrained energy minimization with the program AMBER

Quantity	Average value ± standard deviation (range)[a]			
	Before energy minimization		After energy minimization	
DIANA target function (Å^2)[b]	0·50 ± 0·14	(0·19 . . . 0·74)		
AMBER energy (kcal/mol)	81 ± 37	(0 . . . 162)	−655 ± 57	(−746 . . . −480)
NOE constraint violations				
Number >0·1 Å	3·8 ± 2·1	(0 . . . 7)	0·1 ± 0·2	(0 . . . 1)
Sum (Å)	3·3 ± 0·4	(2·0 . . . 3·8)	5·2 ± 0·4	(4·3 . . . 6·0)
Maximum (Å)	0·15 ± 0·05	(0·07 . . . 0·31)	0·09 ± 0·00	(0·09 . . . 0·11)
Angle constraint violations				
Number >5°	0		0	
Sum (deg.)	6·1 ± 2·4	(2·4 . . . 10·6)	6·2 ± 1·7	(3·2 . . . 10·1)
Maximum (deg.)	1·8 ± 0·9	(0·6 . . . 4·2)	1·9 ± 0·3	(1·2 . . . 2·4)

A total of 50 structures were calculated with DIANA using REDAC, but only the 20 structures with the smallest final target function values were subjected to energy minimization.

[a] Of the values for the 20 individual conformers.

[b] The weighting factors for the NOE upper distance constraints were 1, for the van der Waals lower distance limits 2, and for dihedral angle constraints 5 Å^2. After energy minimization the target function is not defined because the refined conformers do not have the ECEPP standard geometry.

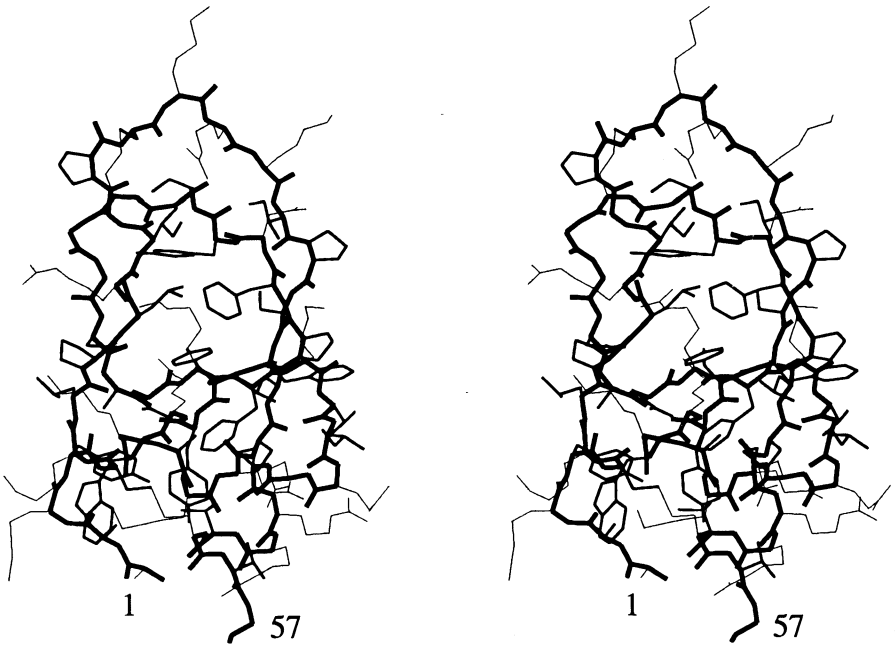

Figure 2. Stereo view of the all-heavy-atom representation of one of the final, energy-minimized DIANA conformers of Toxin K. The backbone heavy-atom positions N, C^{α}, C′ and O′ are connected by thick lines, the best-defined side-chains (see Table 4) are represented by lines of medium thickness, and other side-chains by thin lines. Note that the local conformation of some of the surface side-chains varies widely among the 20 conformers used to represent the solution structure (see Figs 3 and 4). The positions of the first and last residue, 1 and 57, are identified.

only 0·25(±0·08) Å for the backbone atoms N, C^{α} and C′, and 0·37(±0·12) Å for all heavy-atoms in the complete polypeptide chain 1 to 57. The energy-minimized conformers exhibit smaller maximal violations of NOE distance constraints (Table 3), because the corresponding penalty term in the modified AMBER program is steeper than in DIANA (Billeter *et al.*, 1990*b*). The average of the r.m.s. deviations of the bond lengths and bond angles from their equilibrium values is 0·0064 Å and 1·8°, respectively, for the 20 energy-minimized DIANA conformers. The average deviation of the ω dihedral angles from 180° is 5·7°.

(c) *The solution structure of Toxin K*

The complete solution structure of Toxin K is shown in Figure 2, and a visual impression of the quality of the structure determination is afforded by Figure 3. The regular secondary structure elements consist of a 3_{10}-helix of residues 3 to 7, a twisted β-hairpin of residues 18 to 35, a one-residue anti-parallel β-strand consisting of residue Phe45, and an α-helix of residues 47 to 56. This coincides very well with earlier predictions based on sequence homology to the Kunitz proteinase inhibitor family (Swenson *et al.*, 1978; Dufton, 1985) or on NMR pattern recognition (Wüthrich, 1986; Wüthrich *et al.*, 1984) involving NOEs, $^{3}J_{HN\alpha}$ coupling constants, and amide proton exchange data (Pardi *et al.*, 1983). The small global r.m.s.d. value of 0·31 Å obtained for the backbone atoms N, C^{α} and C′ of residues 2 to 56 (Table 4) is representative of a high-quality NMR structure. The same is reflected by the highly restricted ranges for the backbone dihedral angles along the polypeptide chain (Fig. 4), where somewhat larger values occur exclusively for the chain-terminal residues and the glycine-rich segment 37–40. Local r.m.s.d. values (Fig. 5(A)) and displacements after global superposition for the backbone atoms N, C^{α} and C′ (Fig. 5(B)) further document the

Table 4
r.m.s.d. values for different selections of atoms in the solution structure of Toxin K

Atoms used for comparison[a]	r.m.s.d. (Å)[b]
Backbone 2–56	0·31±0·06
Same + best-defined side-chains[c]	0·33±0·06
All heavy-atoms 2–56	0·90±0·06
Backbone 18–35 (β-sheet)	0·17±0·05
All heavy-atoms 18–35	0·56±0·08
Backbone 47–56 (α-helix)	0·12±0·03
All heavy-atoms 47–56	1·01±0·15

[a] Backbone stands for the atoms N, C^{α} and C′ of the residues indicated by the sequence positions.

[b] Average and standard deviation of the pairwise r.m.s.d. values between each of the 20 energy-refined DIANA conformers and the mean solution structure.

[c] The best-defined side-chains are those of the following 32 residues: Ala2, Tyr4, Cys5, Leu7, Pro8, Leu9, Ile11, Pro13, Cys14, Ile18, Pro19, Ser20, Phe21, Tyr22, Tyr23, Trp25, Ala27, Cys30, Leu31, Pro32, Phe33, Tyr35, Ser36, Asn41, Ala42, Asn43, Phe45, Thr47, Ile48, Cys51, Thr54, Val56 (see text for details). All heavy-atoms of these residues were used in the calculation of the r.m.s.d. value.

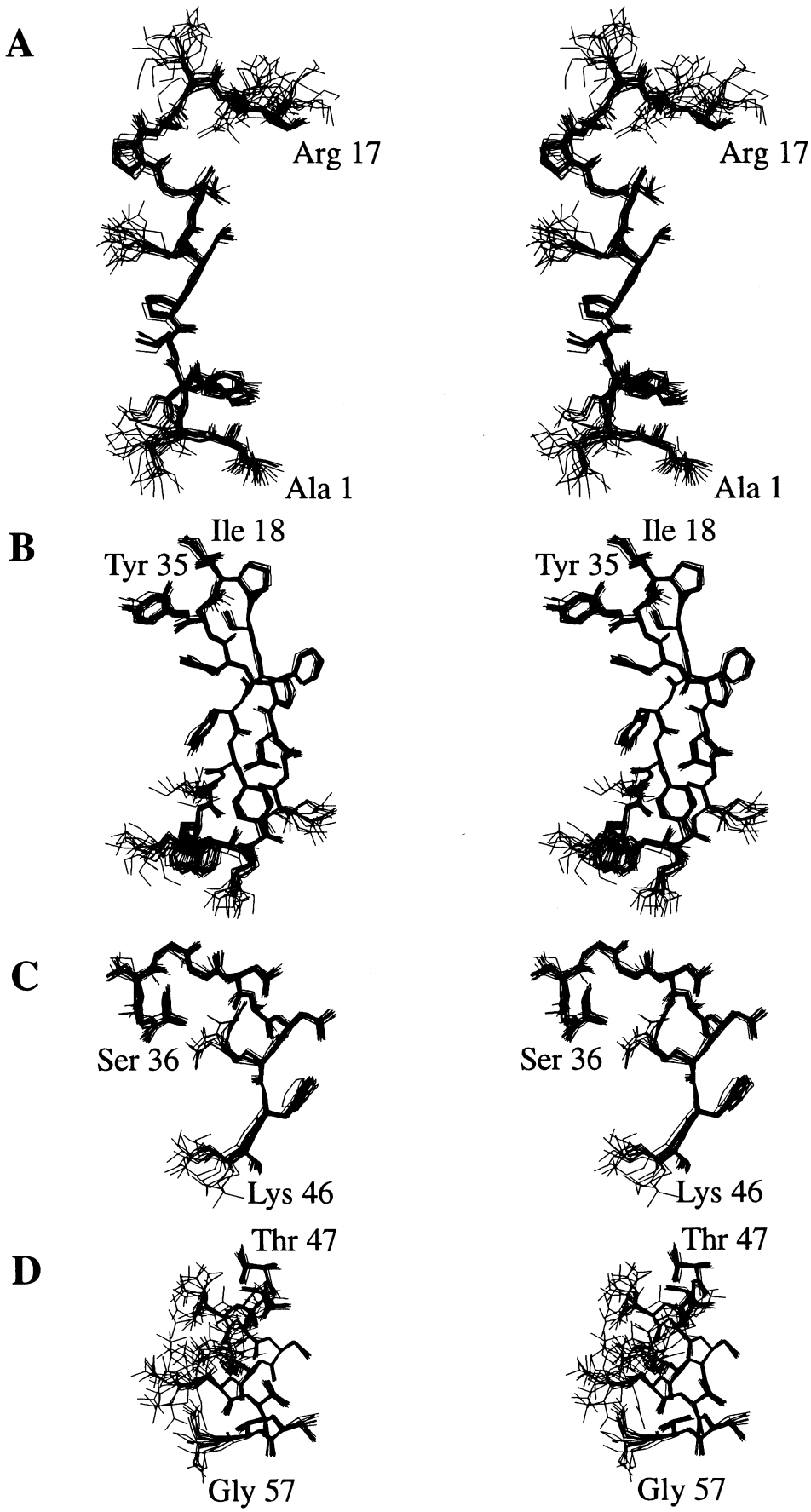

Figure 3. Stereo view of the all-heavy-atom representation of the 20 final, energy-refined DIANA conformers of Toxin K: A, residues 1 to 17 (AAKYCKLPLRIGPCKRK); B, residues 18 to 35 (IPSFYYKWKAKQCLPFDY); C, residues 36 to 46 (SGCGGNANRFK); D, residues 47 to 57 (TIEECRRTCVG). The residues shown are the same as those used for the superposition for minimal r.m.s.d. of the backbone atoms N, C^{α} and C′, except that the first residue in A and the last residue in D were not included in the superposition.

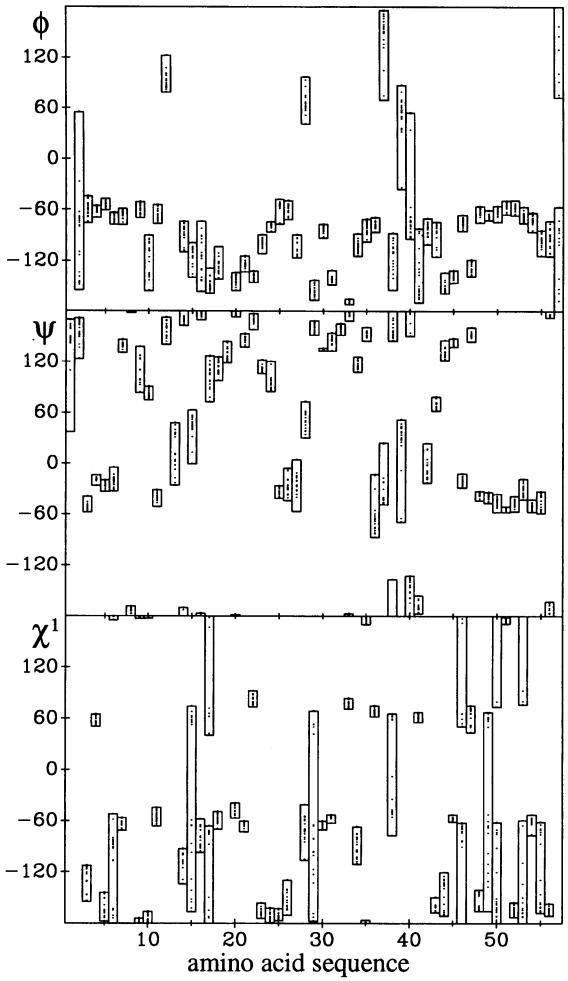

Figure 4. Plots of the dihedral angles ϕ, ψ and χ^1 *versus* the amino acid sequence of Toxin K. The bars indicate the range of values observed in the 20 final, energy-minimized DIANA conformers. Dots within the bars indicate the values for the individual conformers.

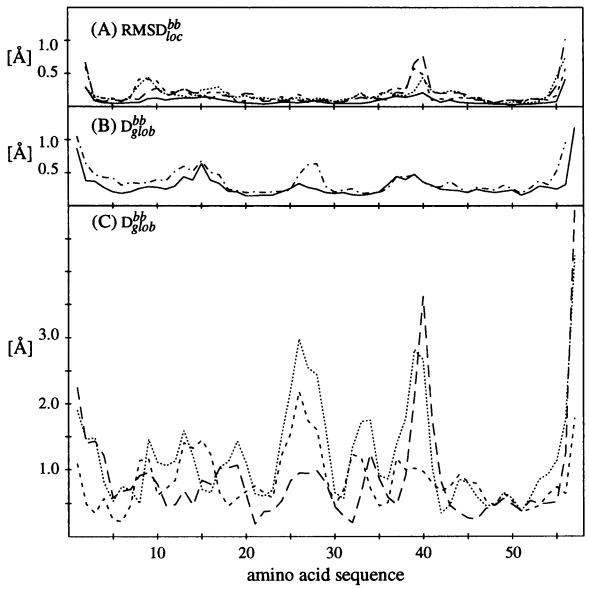

Figure 5. Local precision of the NMR solution structure of Toxin K, and comparisons with BPTI and with the X-ray crystal structure of α-DTX. The following quantities are plotted *versus* the amino acid sequence. (A) The average of the local backbone r.m.s.d. values for tripeptide segments, $RMSD^{bb}_{loc}$, between the 20 energy-minimized DIANA conformers and the mean solution structure of Toxin K (continuous line). Corresponding local $RMSD^{bb}_{loc}$ values between the mean solution structure of Toxin K and the mean solution structure of BPTI (dotted line), between the mean solution structure of Toxin K and the X-ray crystal structure of α-DTX (short-dash line), and between the mean solution structure of BPTI and the X-ray crystal structure of α-DTX (long-dash line) are also shown. (B) The average of the global backbone displacements, D^{bb}_{glob}, of the 20 energy-refined DIANA conformers of Toxin K relative to the mean structure of Toxin K (continuous line), and of the 20 energy-refined DIANA conformers of BPTI relative to the mean structure of BPTI (dot-dash line). For both proteins the superposition was for residues 2 to 56. (C) Global backbone displacements, D^{bb}_{glob}, between the mean solution structures of Toxin K and BPTI (dotted line), the mean solution structure of Toxin K and the X-ray crystal structure of α-DTX (short-dash line), and between the mean solution structure of BPTI and the X-ray crystal structure of α-DTX (long-dash line) after superposition of residues 2 to 56. The data on the NMR solution structure of BPTI are from Berndt *et al.* (1992), those on the crystal structure of α-DTX from Skarzynski (1992).

fact that with the exception of the two terminal residues and a displacement larger than 0·5 Å for Lys15, the polypeptide backbone is very well defined by the NMR data.

We identified 32 "best-defined" residues (listed individually in Table 4), for which the side-chain displacements after global backbone superposition are below a threshold of 0·8 Å. The global r.m.s.d. values calculated for the backbone atoms of residues 2 to 56 and the heavy atoms of these 32 best-defined side-chains are virtually identical to those for the backbone atoms alone (Table 4), and for these residues the variation of the dihedral angle χ^1 among the 20 final, energy-minimized DIANA conformers is also very restricted (Fig. 4). A large proportion of the remaining side-chains, most of which are located near the protein surface (Fig. 2), are significantly less well defined (Fig. 3), with outstandingly large variability of χ^1 observed for residues Lys6, Lys15, Lys17, Gln29, Lys46, Glu49, Glu50 and Arg53 (Fig. 4).

Using the criteria described in the footnotes to Table 5, a total of 22 backbone–backbone hydrogen bonds and ten backbone–side-chain hydrogen bonds were identified in Toxin K. Nearly all of the backbone amide protons involved in these hydrogen bonds have been found to be slowly exchanging at pH 4·6 and 36 °C (Antuch *et al.*, 1993), the exceptions being Cys5, Lys6, Ala27, Lys28, Glu50 and Val56 (Table 5). Five of these six residues were previously found to exchange sufficiently slowly at pH 3·5 and 24 °C for their amide proton NMR lines

Table 5

Hydrogen bonds identified in the solution structures of Toxin K and BPTI

Donor[a]		Acceptor[a]		Toxin[b]	BPTI[b]
A. *Backbone amide proton–backbone carbonyl oxygen hydrogen bonds*					
5 Cys	NH	2 Ala (P)	O′	11	7
6 Lys (L)	NH	3 Lys (D)	O′	19	20
7 Leu (E)	**NH**	4 Tyr (F)	O′	20	19
14 Cys	**NH**	12 Gly	O′	6	14
18 Ile	**NH**	35 Tyr	O′	19	20
20 Ser (R)	**NH**	33 Phe	O′	20	20
21 Phe (Y)	**NH**	45 Phe	O′	20	19
22 Phe	**NH**	31 Leu	O′	20	20
24 Lys (N)	**NH**	29 Gln (L)	O′	20	17
27 Ala	NH	24 Lys (N)	O′	18	20
28 Lys (G)	NH	24 Lys (N)	O′	19	20
31 Leu (Q)	**NH**	22 Tyr (F)	O′	20	20
33 Phe	**NH**	20 Ser (R)	O′	20	20
35 Tyr	**NH**	18 Ile	O′	20	20
36 Ser (G)	**NH**	11 Ile (T)	O′	19	15
44 Arg (N)	**NH**	42 Ala (R)	O′	16	20
45 Phe	**NH**	21 Phe (Y)	O′	20	20
51 Cys	**NH**	47 Thr (S)	O′	15	19
52 Arg (M)	**NH**	48 Ile (A)	O′	20	20
53 Arg	**NH**	49 Glu	O′	20	20
54 Thr	**NH**	50 Glu (D)	O′	18	20
55 Cys	**NH**	51 Cys	O′	20	18
56 Val (G)	NH	52 Arg (M)	O′	18	15
B. *Hydrogen bonds involving side-chain atoms*					
20 Ser (R)	γOH	18 Ile	O′	10	—
23 Tyr	**NH**	43 Asn	δO	20	16
24 Lys (N)	(δNH_2)	31 Leu (Q)	(εO)	—	12
26 Lys	NH	24 Lys (N)	(δO)	—	20
35 Tyr	ηOH	37 Gly	O′	10	5
35 Tyr	ηOH	38 Cys	O′	7	12
36 Ser (G)	γOH	11 Ile (T)	O′	20	—
43 Asn	δNH_2	7 Leu (E)	O′	20	15
43 Asn	δNH_2	23 Tyr	O′	20	17
43 Asn	**NH**	41 Asn (K)	δO	19	—
44 Arg (N)	ηNH_2	10 Arg	O′	15	—
50 Glu (D)	NH	47 Thr (S)	γO	17	16
54 Thr	γOH	50 Glu (D)	O′	16	12

In each individual conformer a hydrogen bond is identified if the proton–acceptor distance is less than 2·4 Å, and the angle between the donor–proton bond and the line connecting the donor and acceptor atoms is less than 35°.

[a] The residue type given with the 3-letter code corresponds to the sequence of Toxin K. When different, the one-letter code of the corresponding residue in BPTI is given in parentheses. NH and O′ are the backbone amide proton and the backbone carbonyl oxygen, respectively. Side-chain atoms are identified with greek letters. Amide protons exchanging more slowly than 0·01 min^{-1} at p^2H 4·6 and 36 °C are indicated in bold face. In addition, the residues Arg10, Gln29, Cys38, Asn41 and Thr47, which are not listed here, are also slowly exchanging by this criterion (see text for details).

[b] These columns list the number of conformers among the 20 conformers used to represent the solution structure for which the hydrogen bond was identified. Hydrogen bonds are reported if a hydrogen bond is identified in at least half of the 20 energy-refined conformers of either Toxin K or BPTI. A dash indicates that the corresponding side-chain atoms are not present due to an amino acid substitution between Toxin K and BPTI. The data for BPTI are taken from Berndt *et al.* (1992).

to be observed after dissolving the protein in ^2H_2O (Pardi *et al.*, 1983). Conversely, the four backbone amide protons of Arg10, Gln29, Cys38 and Thr47 were identified as slowly exchanging at pH 4·6 and 36 °C but are not consistently hydrogen bonded in the group of 20 refined DIANA conformers (Table 5). However, if the geometric criteria were slightly relaxed, Lys24 O′ and Ser36 $^{\gamma}O$ would be identified as hydrogen bond acceptors for residues 29 and 38, respectively, in a majority of the conformers. The significance of the slow amide proton exchange of residue Cys38 will be discussed further when comparing Toxin K with BPTI (see below).

Table 6

Pairwise r.m.s.d. values between Toxin K and selected homologous proteins

	r.m.s.d. (Å)[a]			
	⟨BPTI⟩[b]	5PTI[c]	α-DTX[d]	APPI[e]
⟨Toxin K⟩[f]	1·31	1·32	0·92	1·21
⟨BPTI⟩[b]		0·76	1·17	0·91
5PTI[c]			0·95	0·68
α-DTX[d]				0·66

[a] The r.m.s.d. values were calculated after global superposition of the backbone atoms N, C^{α} and C′ of residues 2 to 56.

[b] Mean solution structure of BPTI (Berndt *et al.*, 1992).

[c] X-ray crystal structure of BPTI form II (Wlodawer *et al.*, 1987).

[d] X-ray crystal structure of α-Dendrotoxin (Skarzynski, 1992).

[e] X-ray crystal structure of the Kunitz-type protease inhibitor domain of the Alzheimer's amyloid β-protein precursor (Hynes *et al.*, 1990).

[f] Mean solution structure of Toxin K.

4. Discussion

Structural similarity between Toxin K and homologous proteins has long been the subject of both theoretical (e.g. see Swenson *et al.*, 1978; Dufton, 1985; Harvey & Anderson, 1986) and experimental investigations (e.g. see Arseniev *et al.*, 1982; Keller *et al.*, 1983; Pardi *et al.*, 1983). However, a basis for detailed comparisons of the three-dimensional structures was established only recently with new high-resolution structure determinations for four of the proteins given in Table 1, in addition to the previously determined three crystal structure forms of BPTI (Deisenhofer & Steigemann, 1975; Wlodawer *et al.*, 1984, 1987). These data show that the four proteins (Toxin K, BPTI, α-DTX and APPI) have very similar polypeptide backbone folds (Fig. 6). This is confirmed by the global r.m.s. deviations calculated for the backbone atoms (Table 6); although these are small, they are nonetheless significantly larger than those among the bundle of solution conformers of Toxin K (Table 4). Except for residues 39 and 40 and the chain-terminal dipeptide segments there are no large local backbone r.m.s. deviations between these proteins (Fig. 5(A)). Following a global superposition of residues 2 to 56 (Fig. 5(B) and (C)) significant displacements are confined to residues 24 to 29 in the β-hairpin, and again to residues 39 to

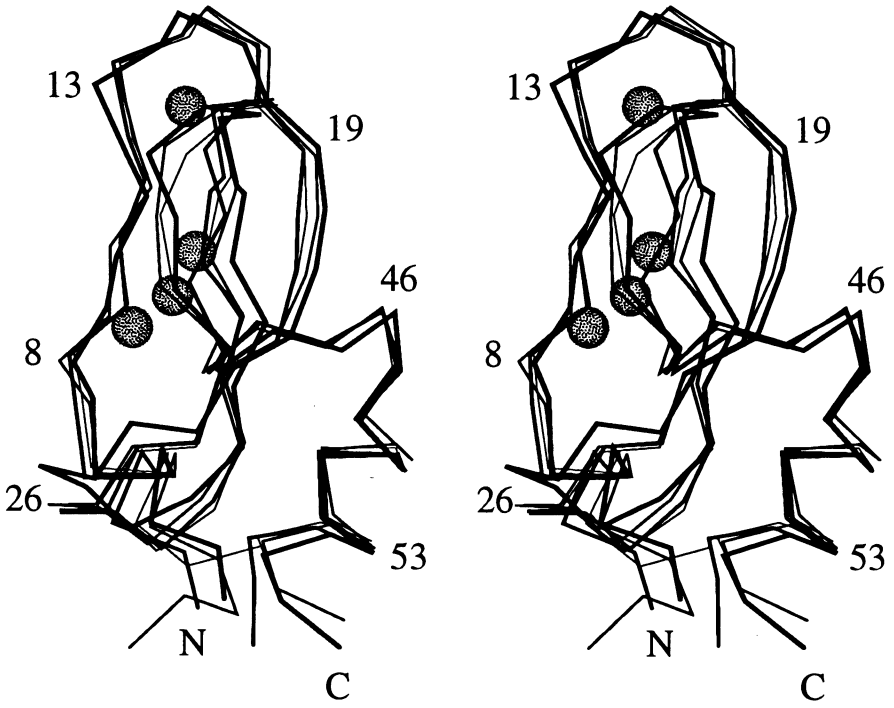

Figure 6. Stereo view of a superposition for best fit of the C^{α} atom positions of residues 2 to 56 of the mean solution structure of Toxin K and the X-ray crystal structures of BPTI form II, α-DTX and APPI. The structures are displayed with decreasing line widths in this order, and only the virtual bonds linking sequentially neighboring C^{α} atoms are shown. The dotted spheres represent the 4 internal water molecules identified in the crystal form II of BPTI (from top to bottom W122, W113, W112 and W111).

40, which are part of a loop that is not well determined by the NMR data for Toxin K. These structure deviations correlate with differences in the amino acid sequence of BPTI when compared with all the other proteins of Table 1 (i.e. BPTI contains the tripeptide Arg-Ala-Lys in positions 39 to 41, whereas the tripeptide Gly-Gly-Asn is found in these positions in the other proteins). Furthermore, BPTI does not contain any prolyl residues in the regular β-structure, which also contrasts with the other proteins (Table 1). In the following we concentrate on a comparison between the NMR solution structures of BPTI and Toxin K, where the latter is selected to broadly represrent the group of proteins that also includes APPI (Ponte *et al.*, 1988) and α-DTX (Strydom, 1973; Joubert & Taljaard, 1980). The BPTI structure (Berndt *et al.* 1992) was obtained with identical procedures to those used here for Toxin K. The main emphasis will be focussed on analyzing how the many amino acid substitutions between Toxin K and BPTI are accommodated in the same global polypeptide fold, and how the numbers and locations of the interior hydration water molecules compare in the two proteins. The question of the structural basis for the variable inhibitory activity of BPTI and the dendrotoxins *versus* proteases has recently been addressed by Skarzynski (1992). His treatment applies to toxin K in a similar way as for the other dendrotoxins.

(a) *The primary and tertiary structures of BPTI and Toxin K*

Although the sequence identity between Toxin K and BPTI is only 41% (Table 1), the total number of non-hydrogen atoms in Toxin K (458) is almost the same as in BPTI (453), which provides a rationale for the fact that the side-chains of both proteins can be accommodated within a nearly identical backbone fold. The local structural similarity is even higher than what one might have expected from the global r.m.s.d. values (Table 6). Thus, differences between corresponding backbone dihedral angles ϕ and ψ do in no instance consistently exceed 20° between all 20 energy-minimized DIANA conformers Toxin K and BPTI, and differences between corresponding side-chain dihedral angles χ^1 exceeding 20° were observed only for the residues in positions 3, 5, 7, 34, 41 and 52. Further inspection of the two structures then provided direct evidence that the differences in the backbone conformations of the two proteins manifested in the global r.m.s.d. value are due primarily to displacements of larger segments of the polypeptide chain with respect to the rest of the protein molecule, rather than to local structural variations on the level of individual backbone dihedral angles. In the standard orientation of Figure 1, the top parts of the two molecules containing the anti-protease loop with residues 11 to 18 and 34 to 39 can be super-

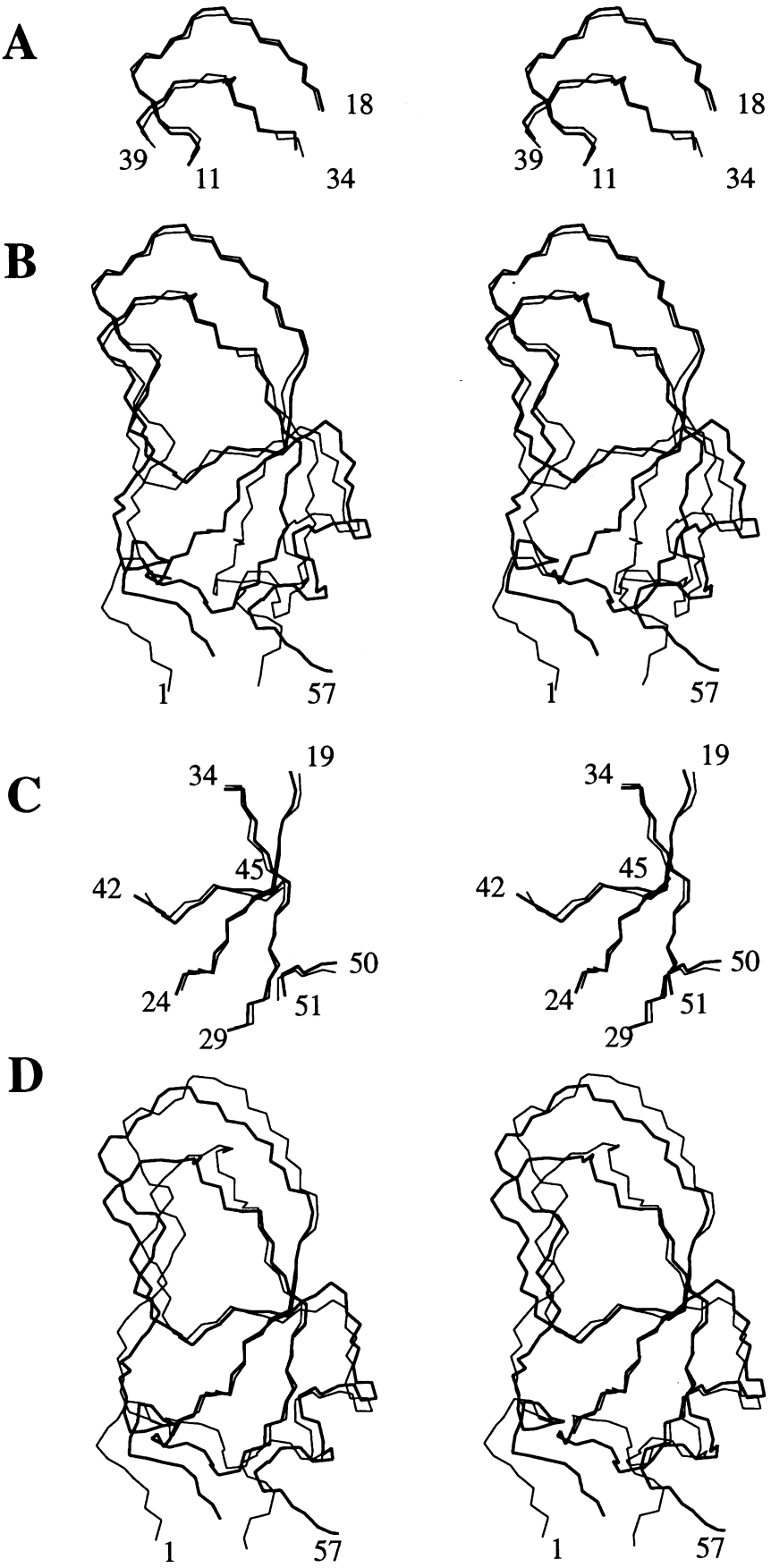

Figure 7. Stereo views of superpositions of the mean solution structures of Toxin K (thick line) and BPTI. In A and B the superposition was made for the backbone atoms N, C^{α} and C′ of the polypeptide segments 11–18 and 34–39 as identified in A, and in C and D for those of segments 19–24, 29–34, 42–45 and 50–51 as identified in C.

Table 7

Amino acid substitutions between Toxin K and BPTI, accessible surface areas of the individual residues, and comments on the conformational consequences of the amino acid substitutions

Position[a]	Toxin K[b]		BPTI[b]		Comments[c]
1	A	(46)	R	(45)	Chain end; largely exposed
2	**A**	(15)	**P**	(20)	Chain end
3	K	(62)	D	(48)	Largely exposed
4	**Y**	(26)	**F**	(16)	Similar ring conformations
6	K	(39)	**L**	(32)	Lys in Toxin K is largely exposed and disordered
7	**L**	(25)	E	(17)	Involved in H_2O replacement (see the text)
9	**L**	(27)	**P**	(25)	
10	R	(35)	Y	(31)	Largely exposed and disordered
11	**I**	(28)	**T**	(24)	Similar rotamer states of $-60°$ about χ^1
16	R	(48)	**A**	(24)	Arg in Toxin K is disordered
17	K	(52)	R	(55)	Largely exposed
19	**P**	(47)	**I**	(42)	Different backbone conformations (see the text)
20	**S**	(6)	R	(19)	See position 44
21	**F**	(18)	**Y**	(19)	Similar ring conformations
22	**Y**	(3)	**F**	(7)	Similar ring conformations
24	K	(24)	**N**	(11)	Different backbone conformations (see the text)
25	**W**	(27)	**A**	(24)	Trp in Toxin K is packed against the protein surface
28	K	(43)	G	(21)	Largely exposed
29	Q	(36)	L	(35)	Largely exposed
31	**L**	(20)	Q	(24)	
32	**P**	(39)	**T**	(34)	Different backbone conformations (see the text)
34	D	(22)	**V**	(29)	Different χ^1 rotamer states; Asp is disordered beyond βC
36	**S**	(1)	G	(0)	Involved in H_2O replacement (see the text)
39	G	(27)	R	(55)	Involved in H_2O replacement (see the text)
40	G	(13)	**A**	(30)	Involved in H_2O replacement (see the text)
41	**N**	(15)	K	(32)	Involved in H_2O replacement (see the text)
42	**A**	(31)	R	(52)	Largely exposed
44	R	(22)	N	(9)	AsnδO in BPTI forms H-bond to Arg20 ηNH_2
47	**T**	(24)	**S**	(19)	γOH forms N-cap of the α-helix (see the text)
48	**I**	(21)	A	(21)	
50	E	(29)	D	(24)	Largely exposed
52	R	(35)	M	(30)	Largely exposed
56	**V**	(31)	G	(21)	Chain end

[a] The numbering is identical for the 2 proteins (Table 1).

[b] These columns list the amino acid types found in the 2 proteins and, in parentheses, the average of the solvent-accessible surface of the residue (Richmond, 1984) in the 20 energy-refined DIANA conformers in % of the total surface. Residues listed with bold letters are among those with best-defined side-chains, i.e. the side-chain displacements after global backbone superposition are smaller than 0·8 Å (see also Table 4).

[c] Brief characterization of the conformational states of the 2 different residues; "exposed" stands for solvent-exposed, "disordered" indicates that the conformation is not uniquely defined by the NMR data and, by implication, probably flexible and disordered.

imposed nearly exactly (Fig. 7A and B), with concomitantly increased deviations for the other molecular regions, and similarly a close fit can be obtained for the central parts of the two molecules with residues 19 to 24, 29 to 34, 42 to 45 and 50 to 51 (Fig. 7C and D).

The Table 7 lists the 33 differences between the amino acid sequences of Toxin K and BPTI, together with the accessible surface areas of the individual residues and with brief comments on the nature of the conformational effects of the amino acid substitutions. Three substitutions at the chain termini (residues 1, 2 and 56) can readily be accommodated, although all of these residues have only limited solvent accessibility. Near the protein surface the positions 3, 6, 10, 16, 17, 28, 42, 50 and 52 all involve a charged residue in at least one of the two proteins, and the side-chains are largely disordered in the NMR structures and exposed to the solvent, with an average surface accessibility of 38%. In the positions 4, 21 and 22 we have conservative substitutions of Tyr with Phe, and the positioning of the aromatic rings is preserved. In positions 11, 20, 24, 25, 31, 44, 47 and 48, which have surface accessibilities of 6 to 29%, the dihedral angles ϕ, ψ and χ^1 are well preserved between the two proteins. In several cases, such as Ile/Thr11 and Thr/Ser47, the similarities extend all the way to the periphery of the side-chains, so that these occupy approximately the same locations in space, or provide similar structural functionality. For example, the γ-oxygens of Thr/Ser47 are both acceptors for the "N-cap" hydrogen bond of the α-helix (Presta & Rose, 1988; Richardson & Richardson, 1988). The presence of the long, flexible disordered side-chain of Arg16 in Toxin K, which replaces Ala in the anti-protease loop of BPTI, is presumably one of the factors responsible for the fact that

Toxin K does not inhibit the proteases trypsin or chymotrypsin (Strydom, 1973). In contrast to BPTI, Toxin K contains two Pro residues within the β-sheet at positions 19 and 32; the slightly different backbone conformations of the β-hairpin 18–35 in the two proteins might be partially due to these substitutions, as Pro residues do not fit into a standard β-sheet structure (Schulz & Schirmer, 1979). There are nine additional amino acid differences in the β-hairpin; in particular, although this side-chain is partially solvent-exposed near the protein surface, the packing requirements of Trp25 in Toxin K, which replaces Ala25 in BPTI, are likely to contribute to the observed difference in the backbone conformation of residues 24 to 29. For Asp/Val34, the backbone dihedral angles have similar values in Toxin K and BPTI, but the χ^1 dihedral angles are in different rotamer positions, and the rest of the Asp side-chain is disordered in Toxin K. The residues 7, 36 and 39 to 41 are involved in the replacement of an internal water molecule, which will be discussed in detail below. Overall, from inspection of the individual amino acid substitutions we find for most positions that both side-chains can be readily accommodated in the given three-dimensional molecular architecture. A small number of amino acid substitutions can be quite clearly related to the previously discussed local differences in the polypeptide backbone conformations.

(b) *Hydrogen bonds*

A comparison of the hydrogen bonds in the solution structures of Toxin K and BPTI (Berndt *et al.*, 1992) shows that in addition to those involved in regular secondary structures, all backbone–backbone hydrogen bonds are conserved between the two structures (Table 5). Five side-chain–backbone hydrogen bonds are also conserved between the two structures. The N and C-caps to the α-helix in BPTI (Berndt *et al.*, 1992), which are thought to stabilize the helical conformation (Presta & Rose, 1988; Richardson & Richardson, 1988), are a conserved feature despite the substitution of Thr for Ser at position 47. The fact that the side-chains of these "capping" residues are among the best-defined ones in the solution structures of both proteins, although they are located on the surface of the molecule, is evidence for the high population of the observed spatial arrangements and their importance in the overall solution structure. The other three conserved side-chain–backbone hydrogen bonds all involve the side-chain of Asn43, which is strictly conserved among all proteins of the Kunitz-type proteinase inhibitor family (Creighton & Charles, 1987), bridge the C-terminal end of the 3_{10}-helix with the β-hairpin, and probably form a cornerstone for this type of polypeptide fold.

As might be expected from the residue identity of only 41% between Toxin K and BPTI (Table 1), there are a number of differences in hydrogen bonds involving side-chain atoms. The two hydrogen bonds Asn24 δNH_2–Gln31 εO and Asn24 δO–Lys26 HN, present in the solution and crystal structures of BPTI, cannot be formed in Toxin K due to substitutions of Lys and Leu at positions 24 and 31, respectively. This may also contribute to the significant displacement observed for the loop connecting the two strands of the β-hairpin (Figs 5(C) and 6). Four hydrogen bonds with side-chain atoms of Toxin K involve non-conserved residues and cannot occur in BPTI (Table 5). The NMR observation of the hydroxyl protons of Ser20 and Ser36 in Toxin K under conditions where these hydroxyl protons normally exchange too rapidly to be seen as separate lines, is a direct result of their involvement in hydrogen bonds in the protein interior.

(c) *Internal hydration water molecules*

BPTI contains four "internal" water molecules, which participate in hydrogen bonds to protein atoms and constitute an integral part of the native protein structure. They were identified by both X-ray diffraction (Diesenhofer & Steigemann, 1975; Wlodawer *et al.*, 1984, 1987) and NMR (Otting & Wüthrich, 1989). A single water molecule (W122) forms hydrogen bonds to residues Thr11, Cys14 and Cys38, and a group of three water molecules (W111, W112 and W113) is located in a cleft formed by the residues 7 to 10 and 40 to 44, with hydrogen bonds to residues Glu7, Pro8, Tyr10, Lys41, Asn43 and Asn44 (Fig. 6). Amino acid substitutions in the homologous proteins APPI and α-DTX have resulted in the displacement of the water molecule W111 (Hynes *et al.*, 1990) and the waters W111 and W122 (Skarzynski, 1992), respectively, as determined by X-ray diffraction. The present determination of the solution structure of Toxin K confirms earlier suggestions (Pardi *et al.*, 1983; Hollecker & Creighton, 1983) that residue Ser 36 in Toxin K replaces Gly36 as well as the internal water molecule W122 in BPTI. This represents a unique structural feature of the family of dendrotoxins, because residue 36 is Ser in all known dendrotoxins (Table 1) and Gly in all protease inhibitors (Dufton, 1985). In the following, the local environments of the two molecular regions which contain internal water molecules in BPTI are examined in the solution structure of Toxin K.

The side-chain conformation of Ser36 is well-defined by four intraresidual and six interresidual NOEs involving the hydroxyl proton (see also Berndt *et al.*, 1993). The resulting orientation of the hydroxyl proton could further be confirmed by coupling constants between the hydroxyl proton and the β-protons, $^3J_{\beta H\gamma OH}$, since the βH-γOH 2QF-COSY cross-peak patterns in H_2O are indicative of a small and a large coupling constant. The well-defined orientation of the hydroxyl proton determined only by NOEs and $^3J_{\alpha\beta}$ is fully consistent with this observation. In all 20 energy-refined NMR structures, the hydroxyl proton of Ser36 is hydrogen-bonded to the carbonyl oxygen of Ile11 (Table 5). If one uses a hydrogen bond length

of 2·7 Å (rather than 2·4 Å as in Table 5), the hydrogen bonds HN Cys14–O$^{\gamma}$ Ser36 and HN Cys38–O$^{\gamma}$ Ser36 are also present in at least half of the energy-minimized NMR structures. The slow exchange of the amide proton of Cys38 (Table 5) further supports its involvement in a hydrogen bond. Compared to the hydrogen bonding network involving the internal water W122 in BPTI, this scheme leaves only the carbonyl of Cys38 as an unsatisfied potential hydrogen bond acceptor, since the water proton interacting with this carbonyl oxygen in BPTI has no counterpart in Ser36 of Toxin K. In the X-ray crystal structure of α-DTX (Skarzynski, 1992), where the hydroxyl moiety of Ser36 also occupies the position of the internal water molecule W122 in BPTI, it forms four hydrogen bonds, with the additional hydrogen bond in O$^{\gamma}$H Ser36–O′ Cys38 being part of a bifurcated hydrogen bond involving the hydroxyl proton. In the solution structure of Toxin K such an interaction would have an unfavorable hydrogen bond angle (>67°). This apparent discrepancy between the two dendrotoxin structures probably results from the uncertainty in the hydroxyl proton position in the X-ray crystal structure of α-DTX.

In the crystal structure form II of BPTI (Wlodawer *et al.*, 1984) the internal water molecule W111 forms three hydrogen bonds with εO Glu7, O′ Pro8 and HN Asn43, and one hydrogen bond to H W112, the neighboring internal water. In Toxin K the substitution of Leu for Glu at position 7 removes one hydrogen bond acceptor, and the side-chain carbonyl oxygen of Asn41 hydrogen bonds with the amide proton of Asn43. A hydrogen bond between the side-chain amide proton of Asn41 and the carbonyl oxygen of Pro8 is also found in most of the NMR conformers, with a hydrogen bond length of 2·7 Å. In this way, all hydrogen donors and acceptors that interact with the internal water W111 in BPTI are involved in intramolecular hydrogen bonds. The aforementioned structural differences between BPTI and Toxin K for the tripeptide segment 39–41 are probably required for proper positioning of the side-chain of Asn41 in Toxin K, since the amide group can otherwise not match the geometry of the replaced water molecule, which is also replaced by identical amino acid substitutions (Table 1) in both α-DTX and APPI.

In a different, complementary approach to the analysis of the internal hydration of Toxin K, NMR experiments using a selective spin-lock instead of selective irradiation to suppress the water signal in NOESY and ROESY spectra (Otting, *et al.*, 1991) were used to identify protein-bound water molecules in Toxin K. A number of intermolecular protein–water NOEs were observed at the ω_1 ^{1}H chemical shift of the solvent resonance, which could be unambiguously assigned to specific protons in Toxin K. Cross-peaks of the water resonance with the residues 11 to 14 and 36 to 38 that are characteristic of the internal water molecule W122 in BPTI (Otting & Wüthrich, 1989) are absent in corresponding spectra of Toxin K, and nearly all these cross-peaks in BPTI are replaced by corresponding cross-peaks with the hydroxyl proton of Ser36 in Toxin K. These data were also compared with identical experiments performed with a mutant BPTI, BPTI(G36S), which differs from the wild-type protein only by a substitution of Gly36 by Ser (Berndt *et al.*, 1993). This comparison confirmed again that the hydroxyl group of Ser36 in Toxin K occupies the position of the single internal water molecule W122 in BPTI. Furthermore, the significantly reduced trypsin and chymotrypsin inhibitory activity of the mutant BPTI(G36S) will compel us to re-examine the cause of the abolition of such activities in the family of dendrotoxins.

Six protons from the residues Glu7 H$^{\beta 2,3}$, H$^{\gamma 2,3}$) and Asn43 (H$^{\beta 2}$, NH$^{\delta}$) have short distances of less than 4·0 Å to the internal water molecule W111 in the crystal structure of BPTI form II (Otting & Wüthrich, 1989). None of the corresponding protons of Toxin K show cross-peaks with the water resonance. In contrast, the protein–water NOEs to the two other internal water molecules in BPTI, W112 and W113, are also present in corresponding spectra of Toxin K, showing that these two internal water molecules are also present in Toxin K. Overall, these comparative studies provide a nice illustration that structural integrity of globular proteins can be based either on the presence of some internal water molecules or on the formation of corresponding intramolecular hydrogen bonds in mutant polypeptide chains. This underscores both the importance of fulfilling the hydrogen bonding requirement of buried polar atoms in globular proteins, and the ability of a stable polypeptide fold to accommodate local structural changes with minimal loss of overall conformational stability.

Financial support by the Schweizerischer Nationalfonds (project 31.32033.91) and the use of the Cray Y-MP computer of the ETH Zürich are gratefully acknowledged. We thank Mr R. Marani for the careful processing of the manuscript. The atomic co-ordinates have been deposited in the Brookhaven Protein Data Bank, accession number 1DTK.

References

Anil-Kumar, Ernst, R. R. & Wüthrich, K. (1980). A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **95**, 1–6.

Antuch, W., Berndt, K. D., Chávez, M. A., Delfín, J. & Wüthrich, K. (1993). The NMR solution structure of a Kunitz-type proteinase inhibitor from the sea anemone *Stichodactyla helianthus*. *Eur. J. Biochem.* **212**, 675–684.

Arseniev, A. S., Wider, G., Joubert, F. J. & Wüthrich, K. (1982). Assignment of the ^{1}H nuclear magnetic resonance spectrum of the trypsin inhibitor E from *Dendroaspis polylepis polylepis*. Two-dimensional nuclear magnetic resonance at 500 MHz *J. Mol. Biol.* **159**, 323–351.

Berndt, K. D., Güntert, P., Orbons, L. P. M. & Wüthrich, K. (1992). Determination of a high-quality NMR solution structure of the bovine pancreatic trypsin inhibnitor (BPTI) and comparison with three crystal structures. *J. Mol. Biol.* **227**, 757–775.

Berndt, K. D., Beunink, J., Schröder, W. & Wüthrich, K. (1993). Designed replacement of an internal hydration water molecule in BPTI: structural and functional implications of a Gly-to-Ser mutation. *Biochemistry,* **32**, 3564–3570.

Billeter, M., Kline, A. D., Braun, W., Huber, R. & Wüthrich, K. (1989). Comparison of the high-resolution structures of the α-amylase inhibitor Tendamistat determined by nuclear magnetic resonance in solution and by X-ray diffraction in single crystals. *J. Mol. Biol.* **206**, 677–687.

Billeter, M., Quian, Y. Q., Otting, G., Müller, M., Gehring, W. J. & Wüthrich, K. (1990*a*). Determination of the three-dimensional structure of the *Antennapedia* homeodomain from *Drosophila* in solution by ^{1}H nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **214**, 183–197.

Billeter, M., Schaumann, Th., Braun, W. & Wüthrich, K. (1990*b*). Restrained energy refinement with two different algorithms and force fields of the structure of the α-amylase inhibitor Tendamistat determined by NMR in solution. *Biopolymers,* **29**, 695–706.

Brüschweiler, R., Griesinger, C., Sørensen, O. W. & Ernst, R. R. (1988). Combined use of hard and soft pulses for ω_1-decoupling in two-dimensional NMR spectroscopy. *J. Magn. Reson.* **78**, 178–185.

Chazin, W. J. & Wright, P. E. (1988). Complete assignment of the ^{1}H nuclear magnetic resonance spectrum of French bean plastocyanin: sequential resonance assignments, secondary structure and global fold. *J. Mol. Biol.* **202**, 623–636.

Creighton, T. E. & Charles, I. G. (1987). Biosynthesis, processing and evolution of bovine pancreatic trypsin inhibitor. *Cold Spring Harbor Symp. Quant. Biol.* **52**, 511–519.

Deisenhofer, J. & Steigmann, W. (1975). Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at 1·5 Å resolution. *Acta Crystallogr. sect. B,* **31**, 238–250.

Dufton, M. J. (1985). Proteinase inhibitors and dendrotoxins: sequence classification, structural prediction and structure/activity. *Eur. J. Biochem.* **153**, 647–654.

Eccles, C., Güntert, P., Billeter, M. & Wüthrich, K. (1991). Efficient analysis of protein 2D NMR spectra using the software package EASY. *J. Biomol. NMR,* **1**, 111–130.

Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988). The MIDAS display system. *J. Mol. Graph.* **6**, 13–27.

Griesinger, C., Sørensen, O. W. & Ernst, R. R. (1985). Two-dimensional correlation of connected NMR transitions. *J. Amer. Chem. Soc.* **107**, 6394–6396.

Güntert, P. & Wüthrich, K. (1991). Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J. Biomol. NMR,* **1**, 447–456.

Güntert, P. & Wüthrich, K. (1992). FLATT—a new procedure for high-quality baseline correction of multidimensional NMR spectra. *J. Magn. Reson.* **96**, 403–407.

Güntert, P., Braun, W., Billeter, M. & Wüthrich, K. (1989). Automated stereospecific ^{1}H NMR assignments and their impact on the precision of protein structure determinations in solution. *J. Amer. Chem. Soc.* **111**, 3997–4004.

Güntert, P., Braun, W. & Wüthrich, K. (1991*a*). Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517–530.

Güntert, P., Qian, Y. Q., Otting, G., Müller, M., Gehring, W. & Wüthrich, K. (1991*b*). Structure determination of the *Antp(C39→S)* homeodomain from nuclear magnetic resonance data in solution using a novel strategy for the structure calculation with the programs DIANA, CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 531–540.

Harvey, A. L. & Anderson, A. J. (1985). Dendrotoxins: snake toxins that block potassium channels and facilitate neurotransmitter release. *Pharmac. Ther.* **31**, 33–55.

Harvey, A. L. & Karlsson, E. (1982). Protease inhibitor homologues from mamba venoms: facilitation of acetylcholine release and interactions with prejunctional blocking toxins. *Brit. J. Pharmac.* **77**, 153–161.

Hollecker, M. & Creighton, T. E. (1983). Evolutionary conservation and variation of protein folding pathways. *J. Mol. Biol.* **168**, 409–437.

Hynes, T. R., Randal, M., Kennedy, L. A., Eigenbrot, C. & Kossiakoff, A. (1990). X-ray crystal structure of the protease inhibitor domain of Alzheimer's amyloid β-protein precurser. *Biochemistry,* **29**, 10018–10022.

Joubert, F. J. & Strydom, D. J. (1978). Snake venoms: the amino acid sequence of trypsin inhibitor E of *Dendroaspis polylepis polylepis* (black mamba) venom. *Eur. J. Biochem.* **87**, 191–198.

Joubert, F. J. & Taljaard, N. (1980). Snake venoms: the amino acid sequences of two proteinase inhibitor homologues from *Dendroaspis angusticeps* venom. *Hoppe-Seylers Z. Physiol. Chem.* **361**, 661–674.

Kassell, B., Radicevic, M., Ansfield, M. J. & Laskowski, M. (1965). The basic trypsin inhibitor of bovine pacreas. IV. The linear sequence of the 58 amino acids. *Biochem. Biophys. Res. Commun.* **18**, 255–258.

Keller, R. M., Baumann, R., Hunziker-Kwik, E. H., Joubert, F. J. & Wüthrich, K. (1983). Assignment of the ^{1}H nuclear magnetic spectrum of the trypsin inhibitor homologue K from *Dendroaspis polylepis polylepis*: two-dimensional nuclear magnetic resonance at 360 and 500 MHz. *J. Mol. Biol.* **163**, 623–646.

Marion, D. & Wüthrich, K. (1983). Application of phase-sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H–^{1}H spin–spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967–974.

McLachlan, A. D. (1979). Gene duplication in the structural evolution of chymotrypsin. *J. Mol. Biol.* **128**, 49–79.

Nagayama, K. & Wüthrich, K. (1981). Structural interpretation of vicinal proton–proton coupling constants $^3J_{\alpha\beta}$ in the basic pancreatic trypsin inhibitor measured by two-dimensional *J*-resolved NMR spectroscopy. *Eur. J. Biochem.* **115**, 653–657.

Otting, G. (1990). Zero-quantum suppression in NOESY and experiments with a *z*-filter. *J. Magn. Reson.* **86**, 496–508.

Otting, G. & Wüthrich, K. (1989). Studies of protein hydration in aqueous solution by direct NMR

observation of individual protein-bound water molecules. *J. Amer. Chem. Soc.* **111**, 1871–1875.

Otting, G., Orbons, L. P. M. & Wüthrich, K. (1990). Suppression of zero-quantum coherence in NOESY and soft-NOESY. *J. Magn. Reson.* **89**, 423–430.

Otting, G., Liepinsh, E., Farmer, B. T. & Wüthrich, K. (1991). Protein hydration studied with homonuclear 3D ^{1}H NMR experiments. *J. Biomol. NMR*, **1**, 209–215.

Pardi, A., Wagner, G. & Wüthrich, K. (1983). Protein conformation and proton nuclear magnetic resonance chemical shifts. *Eur. J. Biochem.* **137**, 445–454.

Penner, R., Peterson, M., Pierau, F. K. & Dreyer, F. (1986). Dendrotoxin, a selective blocker of non-activating potassium currents in guinea pig dorsal root ganglion neurons. *Pflügers Arch.* **407**, 365–369.

Ponte, P., Gonzalez-DeWhitt, P., Schilling, J., Miller, J., Hsu, D., Greenberg, B., Davis, K. Wallace, W., Lieberberg, I., Fuller, F. & Cordell, B. (1988). A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors. *Nature (London)*, **331**, 525–527.

Presta, L. G. & Rose, G. D. (1988). Helix signals in proteins. *Science*, **240**, 1632–1641.

Rance, M., Bodenhausen, G., Wagner, G., Wüthrich, K. & Ernst, R. R. (1985). A systematic approach to the suppression of J cross peaks in 2D exchange and 2D NOE spectroscopy. *J. Magn. Reson.* **62**, 497–510.

Richardson, J. S. & Richardson, D. C. (1988). Amino acid preferences for specific locations at the ends of α-helices. *Science*, **240**, 1648–1652.

Richmond, T. J. (1984). Solvent-accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J. Mol. Biol.* **178**, 63–89.

Schulz, G. E. & Schirmer, R. H. (1979). *Principles of Protein Structure*, Springer-Verlag, New York.

Singh, U. C., Weiner, P. K., Caldwell, J. W. & Kollman, P. A. (1986). Amber 3.0, University of California at San Francisco.

Skarzynski, T. (1992). Crystal structure of α-dendrotoxin from the green mamber venom and its comparison with the structure of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **224**, 671–683.

Strydom, D. J. (1972). Snake venom toxins: the amino acid sequences of two toxins from *Dendroaspis polylepis polylepis* (black mamba) venom. *J. Biol. Chem.* **247**, 4029–4042.

Strydom, D. J. (1973). Protease inhibitors as snake venom toxins. *Nature New Biol.* **243**, 88–89.

Swenson, M. K., Burgess, A. W. & Scheraga, H. A. (1978). In *Frontiers in Physicochemical Biology* (Pullman, B., ed.), pp. 115–142, Academic Press, New York.

Szyperski, T., Güntert, P., Otting, G. & Wüthrich, K. (1992). Determination of scalar coupling constants by inverse Fourier transformation of in-phase multiplets. *J. Magn. Reson.* **99**, 552–560.

Tüchsen, E. & Woodward, C. (1987). Assignment of asparagine-44 side-chain primary ^{1}H NMR resonances and the peptide amide ^{1}H resonance of glycine-37 in basic pancreatic trypsin inhibitor. *Biochemistry*, **26**, 1918–1925.

Williamson, M. P., Havel, T. F. & Wüthrich, K. (1985). Solution conformation of proteinase inhibitor IIa from bull seminal plasma by ^{1}H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **182**, 295–315.

Wlodawer, A., Walter, J., Huber, R. & Sjölin, L. (1984). Structure of bovine pancreatic trypsin inhibitor. Results of joint neutron and X-ray refinement of crystal form II. *J. Mol. Biol.* **180**, 301–329.

Wlodawer, A., Nachman, J., Gilliland, G. L., Gallagher, W. & Woodward, C. (1987). Structure of form III crystals of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **198**, 469–480.

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

Wüthrich, K., Billeter, M. & Braun, W. (1983). Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton–proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949–961.

Wüthrich, K., Billeter, M. & Braun, W. (1984). Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton–proton distances. *J. Mol. Biol.* **180**, 715–740.

Xia, T. (1992). *Software for Determination and Visual Display of NMR Structures of Proteins: the Distance Geometry Program DGPLAY and the Computer Graphics Programs CONFOR and XAM*. Ph.D. thesis Nr. 9831, ETH Zürich, Switzerland.

Edited by P. E. Wright