Protein Science (1995), *4*:178–186. Cambridge University Press. Printed in the USA.

NMR structure determination of tick anticoagulant peptide (TAP)

MARGUERITA S.L. LIM-WILBY,[1] KLAAS HALLENGA,[2] MARC DE MAEYER,[2] IGNACE LASTERS,[2] GEORGE P. VLASUK,[1] AND TERENCE K. BRUNCK[1]

[1] Corvas International Inc., 3030 Science Park Road, San Diego, California 92121
[2] Corvas International NV, Plateaustraat 22, B-9000 Gent, Belgium

(RECEIVED August 31, 1994; ACCEPTED November 21, 1994)

Abstract

Tick anticoagulant peptide (TAP) is a potent and selective 60-amino acid inhibitor of the serine protease Factor Xa (fXa), the penultimate enzyme in the blood coagulation cascade. The structural features of TAP responsible for its remarkable specificity for fXa are unknown, but the binding to its target appears to be unique. The elucidation of the TAP structure may facilitate our understanding of this new mode of serine protease inhibition and could provide a basis for the design of novel fXa inhibitors. Analyses of homo- and heteronuclear two-dimensional NMR spectra (total correlation spectroscopy, nuclear Overhauser effect spectroscopy [NOESY], constant time heteronuclear single quantum correlation spectroscopy [CT-HSQC], and HSQC-NOESY; 600 MHz; 1.5 mM TAP; pH 2.5) of unlabeled, ^{13}C-labeled, and ^{15}N-labeled TAP provided nearly complete ^{1}H sequence-specific resonance assignments. Secondary structural elements were identified by characteristic NOE patterns and D_2O amide proton-exchange experiments. A three-dimensional structure of TAP was generated from 412 NOESY-derived distance and 47 dihedral angle constraints. The structural elements of TAP are similar in some respects to those of the Kunitz serine protease inhibitor family, with which TAP shares weak sequence homology. This structure, coupled with previous kinetic and biochemical information, confirms previous suggestions that TAP has a unique mode of binding to fXa.

Keywords: anticoagulant; Kunitz; NMR; serine protease inhibitor; solution structure; tick

Tick anticoagulant peptide is a 60-amino acid polypeptide originally isolated from the soft tick *Ornithodoros moubata* that shows both potent (K_i = 2–200 pM) and exquisitely selective inhibition of blood coagulation Factor Xa, an enzyme of the trypsin-like serine protease family (Waxman et al., 1990). TAP is the smallest protein of the set of specific and potent natural inhibitors of fXa, which includes antistasin (Nutt et al., 1988) and tissue factor pathway inhibitor (Wun et al., 1988). TAP shows weak (~20%) amino acid sequence homology to Kunitz-type serine protease inhibitors, and the connectivity of the disulfide bonds is identical to other members of this family (Sardana et al., 1991). However, TAP shows no ability to inhibit trypsin or other trypsin-like proteases, nor is it cleaved by fXa, trypsin, or any related enzymes. Furthermore, the canonical residues of Kunitz inhibitors required to inhibit trypsin-like proteases are absent from the TAP sequence. Thus, the mechanism of fXa inhibition by TAP is unique compared to that of antistasin and tissue factor pathway inhibitor, both of which appear to bind in a substrate-like manner to fXa.

TAP has been produced recombinantly in *Saccharomyces cerevisiae* and shown to have biological properties identical to native material (Neeper et al., 1990). Additionally, numerous mutants of TAP have been prepared, and a preliminary structure–activity relationship has identified residues crucial for its activity (Dunwiddie et al., 1992). Kinetic analysis of recombinant TAP revealed that it is a competitive inhibitor of fXa with an apparent two-step binding pathway and an overall K_i^* of 0.2 nM (Jordan et al., 1990, 1992). Recombinant TAP is a more potent inhibitor of the prothrombinase complex than of free fXa (Krishnaswamy et al., 1994), and it is effective in vivo as an antithrombotic agent (Dunwiddie et al., 1991; Schaffer et al., 1991; Vlasuk et al., 1991; Sitko et al., 1992).

Similar to the unique specificity that the naturally occurring peptide hirudin exhibits for thrombin, TAP shows absolute selectivity for its target enzyme fXa. It is not known to inhibit any

Reprint requests to: Terence K. Brunck, 3030 Science Park Road, San Diego, California 92121; e-mail: tbrunck@crash.cts.com or corvas.intl@applelink.apple.com (Attn: Terry Brunck).

Abbreviations: TAP, tick anticoagulant peptide; fXa, blood coagulation Factor Xa; 1D, one-dimensional; 2D, two-dimensional; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; CT-HSQC, constant time heteronuclear single quantum correlation spectroscopy; FID, free induction decay; BPTI, bovine pancreatic trypsin inhibitor; RMSD, RMS deviation; ShPI, *Stichodactyla helianthus* protease inhibitor.

other trypsin- or chymotrypsin-like serine proteases. When compared with Kunitz inhibitors, the weak sequence homology and identical pattern of reducing-agent-resistant disulfide bonds originally suggested the possibility of a Kunitz-like structure. The structural elements of TAP, determined by the NMR analysis reported below, show clearly the overall structural similarity of TAP and Kunitz-like protease inhibitors, as was suggested previously based on a preliminary NMR structure (Vlasuk, 1993).[3] However, a key site of interaction on TAP defined by the inhibitory activity of mutants is at or near the N-terminus of the protein (Tyr 1 to Arg 3), in marked contrast to the binding site on Kunitz-like inhibitors (e.g., bovine pancreatic trypsin inhibitor residues 13–20) (Marquart et al., 1983). Although TAP is structurally similar to Kunitz inhibitors, it clearly cannot bind to its cognate protease in the same manner as other members of this family. Thus, the determination of the full structure of TAP provides a key starting point for the understanding of its novel mode of binding and may be of possible use for the design of new selective inhibitors of fXa.

Results

Sequence-specific proton assignments

Initial spectra of TAP acquired at pH 3.5 and 1.6 mM showed concentration-dependent changes in resonant frequencies, which were possibly due to aggregation. No such changes were observed at pH 2.5 for solutions of 0.75–1.5 mM, the conditions for all assignments given in this report. Aggregation was observed visually at pH 7 as a phase-separated gel that formed at 3 mM TAP in D_2O. Thus, to obtain spectra of monomeric TAP, low concentration and pH were essential.

The spin system assignments were made initially from the NH-αH fingerprint region of a 75-ms TOCSY spectrum for 52 of 60 residues. The spin system assignments for the remaining eight residues were developed from both TOCSY and NOESY spectra concurrent with the generation of the sequence-specific assignments. A ^{13}C CT-HSQC spectrum confirmed the assignments for upfield αH (at 2.20, 2.91, 3.24, and 3.55 ppm), the βH for serine and threonine residues, and the δH for the two proline residues (Fig. 1). The $^{13}C\alpha$ assignments in regions of overlapping αH resonances (~4.34, ~4.40, ~4.74 ppm) were developed from careful analyses of proton chemical shifts (resolution of ~3 Hz or 0.005 ppm before enhancement) and, for Trp 37 and His 43 at ~4.74 ppm, by intensity and shape. With the exception of Phe 26, proton assignments for 10 of the 11 aromatic residues were obtained from the CT-HSQC spectrum in conjunction with the TOCSY and NOESY spectra. A second ^{13}C CT-HSQC at 35 °C was required to establish the assignments for Phe 26 aromatic protons, which were not clearly detected at 25 °C. Confirmation of 52 of the NH and all of the arginine δH assignments and determination of the asparagine βH and δH assignments were obtained from an ^{15}N HSQC-NOESY spectrum. Sequence-specific assignments were developed from the spin system assignments using the short-range NOEs shown in Figure 2. The complete sequential assignments

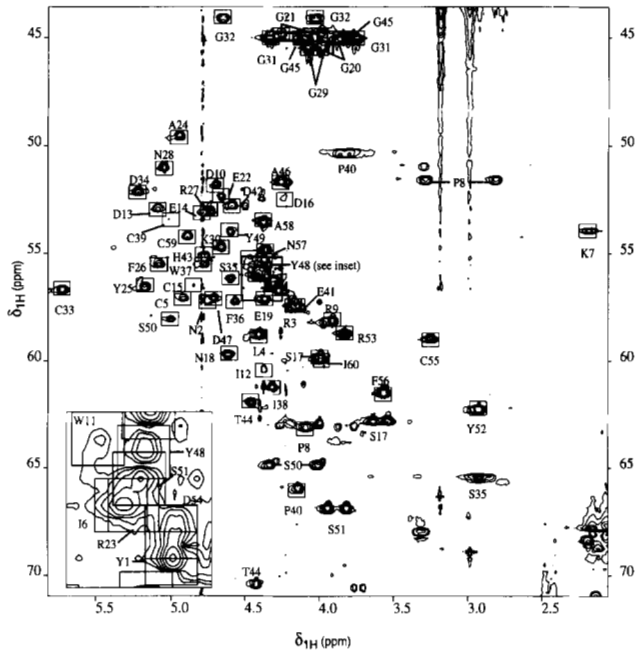


Fig. 1. Cα region of a ^{13}C-constant time HSQC correlation experiment. Peaks are labeled by the single-letter code and sequence number, with the Cα peaks distinguished from Cβ and Cδ peaks by an enclosing box. The constant time period was 17 ms, with a total of 189 increments, each a sum of 64 transients. Processing was performed with TRIAD software (Tripos Associates). The intensity of the first point for each FID was halved, followed by apodization with a squared sine wave, shifted by 80°. Each FID was zero-filled to 2,048 points, Fourier transformed, phased, and then baseline corrected with a polynomial fit. Pairs of interferograms from 512 increments were shuffled to separate real and imaginary parts. The second transform was performed after similar halving of the first point and apodization, followed by zero filling to 1,024 points. The spectrum was referenced to the TSP peak at 0 ppm for both carbon and proton dimensions; two ω1 sweepwidths were added to the former to obtain the scale for this region.

of ^{1}H, $^{13}C\alpha$, ^{13}C aromatic, and ^{15}N chemical shifts are given in the supplementary material (Table S1, Electronic Appendix).

Several regions of the TAP molecule presented special difficulties or showed unusual resonances worthy of further description. Beginning from the N-terminus, Lys 7 was found to have large upfield shifts for αH (2.20 ppm) and γH (−0.11, 0.61 ppm). The side chains of both Lys 7 and Pro 8 showed very broad and weak resonances. TOCSY crosspeaks from NH and ζH were observed for both Lys 7 and Lys 30, although those for the former were relatively weak. Because there were only two proline and two lysine residues in the molecule, sequential assignments of residues 7 and 8 were aided by assigning the more intense resonances and crosspeak patterns of Lys 30 and Pro 40 first. The ^{13}C CT-HSQC spectrum clearly showed multiple resonances for Pro 8 αH and δH resonances at both 25 and 35 °C (Fig. 1). Arg 9 similarly showed multiple αH resonances.

The NH resonances of residues 13–16 are quite broad (peak widths in ω2 up to 0.1 ppm), with Asp 13 and Asp 16 weak, and Glu 14 and Cys 15 barely detectable in H_2O TOCSY spectra. Slices of the fingerprint region of a TOCSY spectrum are shown in Figure 3 for residues 13–16, which illustrate the relative difference in intensity and peak shape for these residues compared to nearby resonances. Also shown in Figure 3 are the corre-

[3] Between the time of our first report on the TAP structure (Vlasuk, 1993) and our revision of this manuscript, a substantially similar NMR structure of TAP at pH 3.6 and 36 °C has been reported (Antuch et al., 1994).

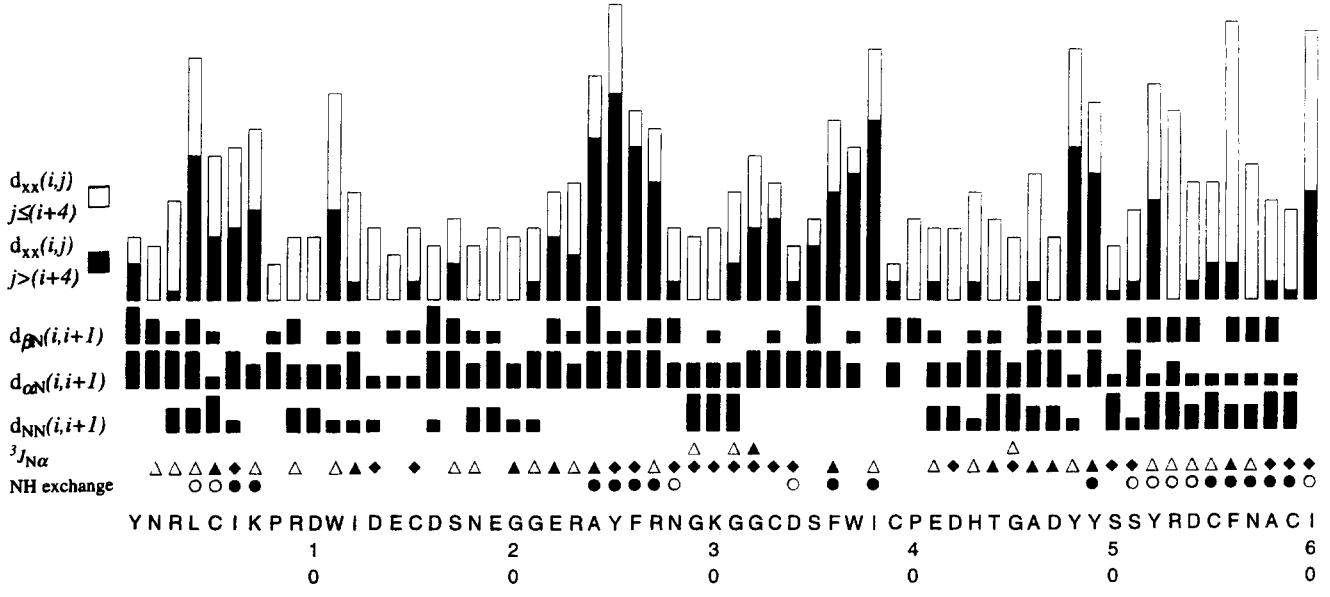


Fig. 2. Summary of NOEs, $^3J_{N\alpha}$ coupling constant data, and slowly exchanging amide protons. The relative intensities of the strong, medium, and weak NOEs are indicated by the height of the bars for $d_{NN}(i, i+1)$, $d_{\alpha N}(i, i+1)$, and $d_{\beta N}(i, i+1)$. The height of the columns for the (i, j) connectivities represents the number of NOEs observed for residue i; the smallest observable unit, found for Arg 3 $j \leq (i+4)$, is equal to one NOE. Symbols for the coupling constants are as follows: △, $^3J_{N\alpha} < 5$ Hz; ▲, 5 Hz $\leq {}^3J_{N\alpha} \leq 9$ Hz; ◆, $^3J_{N\alpha} > 9$ Hz. Slowly exchanging amide protons that are observable in a TOCSY spectrum recorded during the first 16 h after dissolving TAP in D_2O are indicated by circles (●, ○). Filled-in circles (●) are still present in a NOESY spectrum recorded over the next 24 h.

sponding slices of a NOESY spectrum. Peak shifts for the side chains of Glu 14 and Cys 15 can be discerned across their broad NH resonances in the NOESY spectrum and provide evidence for multiple conformations for these residues. Spectra taken over a range of temperatures were necessary to confirm the assignments. Ser 17 also shows multiple Cβ and βH resonances in the ^{13}C CT-HSQC spectrum at 35 °C (Fig. 1).

A second region with broad, weak resonances is located in the sequence near Cys 39, which is disulfide-bonded to Cys 15. Slices of the TOCSY and NOESY spectra for residues 35–39 are shown in Figure 3. In this region, Ser 35, Trp 37, Cys 39, and Pro 40 are especially affected, with broad, weak crosspeaks for all protons of Ser 35 and barely detectable crosspeaks for NH and αH of Trp 37 and Cys 39 in the TOCSY spectrum. Interestingly, the neighboring residues (Asp 34, Phe 36, Ile 38, and Glu 41) have reasonably sharp and intense resonances and the NH-αH connectivities are reasonably strong.

Residues near the C-terminus of TAP (Tyr 52, Cys 55, and Phe 56) show significantly (>1 ppm) upfield-shifted αH resonances (at 2.92, 3.24, and 3.53 ppm, respectively).

Amide proton exchange in D_2O

A 16-h TOCSY data set was acquired at 25 °C immediately after dissolution of TAP in D_2O and pH adjustment to 2.5 (Fig. 4). Residues with slowly exchanging amide protons over this time period are shown in Figure 2. The residues with amide protons still detectable in a 24-h NOESY spectrum acquired immediately after the TOCSY experiment were: Ile 6, Lys 7, Ala 24, Tyr 25, Phe 26, Arg 27, Phe 36, Ile 38, Tyr 49, Cys 55, Phe 56, Asn 57, Ala 58, and Cys 59, as shown in Figure 2.

Secondary structure elements of TAP

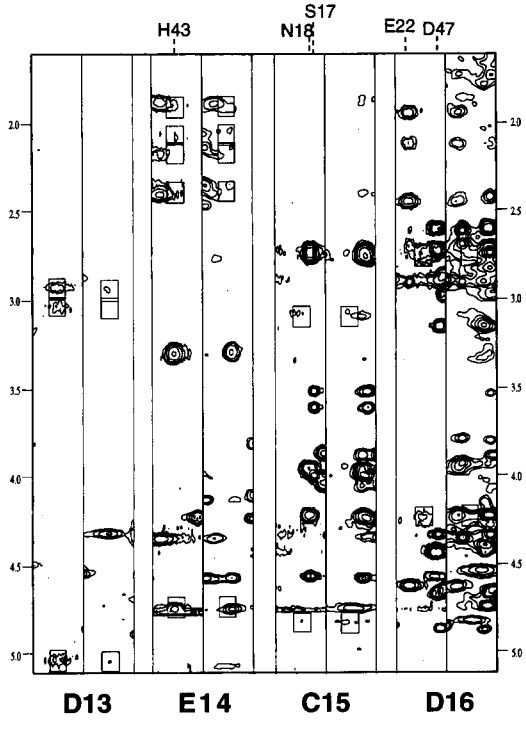
Based on the slow exchange of amide protons and characteristic patterns of NOEs (Fig. 2), several regions of secondary structure were identified. Slow amide exchange for residues 4–7, together with NN$(i, i+1)$ and βN$(i, i+1)$ sequential NOEs for Arg 3, Leu 4, Cys 5, Ile 6, and Lys 7, suggest a helical structure in this region. φ Angles determined for residues 2, 3, 4, and 7 were also consistent with a helical structure.

The NN$(i, i+1)$, αN$(i, i+1)$, βN$(i, i+1)$, βN$(i, i+3)$, βN$(i, i+4)$ NOEs were also observed for residues in the segment Ser 51–Ile 60. All of these interactions were quite strong except those for Ser 51–Tyr 52. The NN sequential NOEs were confirmed in the ^{15}N-HSQC-NOESY spectrum. Additionally, the amide protons of residues Ser 51–Ile 60 were observed to exchange slowly (Figs. 2, 4), and φ angles determined for residues 52–55 and 57 were also consistent with a helical structure.

Finally, a region of double-stranded, antiparallel β-sheet was identified in the central portion of the sequence. Figure 5 shows the characteristic NN, αN, and αα NOEs that were observed for the sheet-turn-sheet fragment of the TAP structure. Additional evidence for the β-sheet was obtained from the TOCSY and NOESY amide-exchange experiments (with slowly exchanging protons shown as wavy N-H bonds in Fig. 5).

Three-dimensional structure of TAP

Fifty distance geometry solutions were generated with the final set of NMR constraints as described in the Materials and methods. The 10 solutions with the smallest distance violations were refined using energy minimization and dynamic simulations to



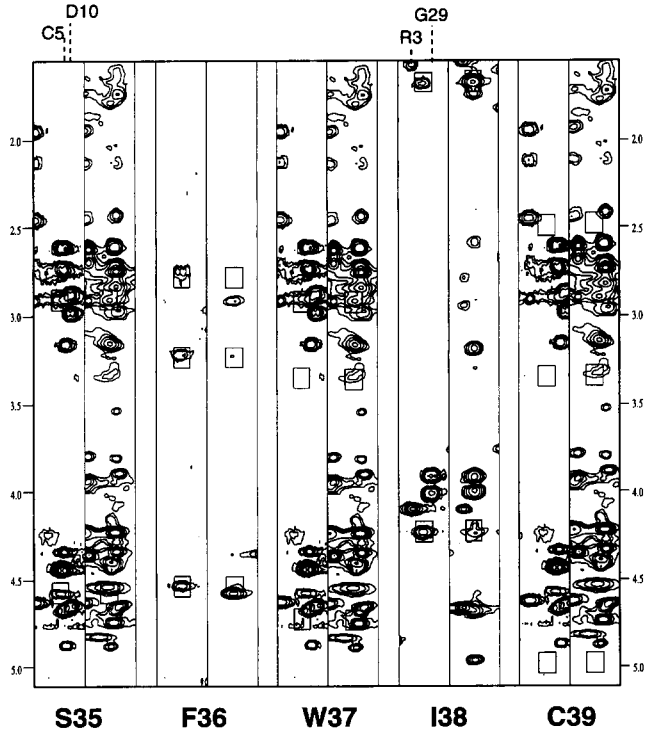


Fig. 3. Corresponding slices from a 75-ms TOCSY and a 180-ms NOESY in the fingerprint region for residues involved in internal motion. For each residue in a region of slow internal motion, a fingerprint slice of the TOCSY (left) and corresponding NOESY (right) is shown. Residues 13–16 (above) are linked to residues 35–38 (below) by the 15–39 disulfide bond. Boxes (dimension: 0.04 ppm horizontal, 0.12 ppm vertical) show the locations of intraresidue crosspeaks. For comparison of peak shape and intensity, other residues with similar NH chemical shifts are included in some of the slices; assignments for some of these are indicated above the spectral slices.

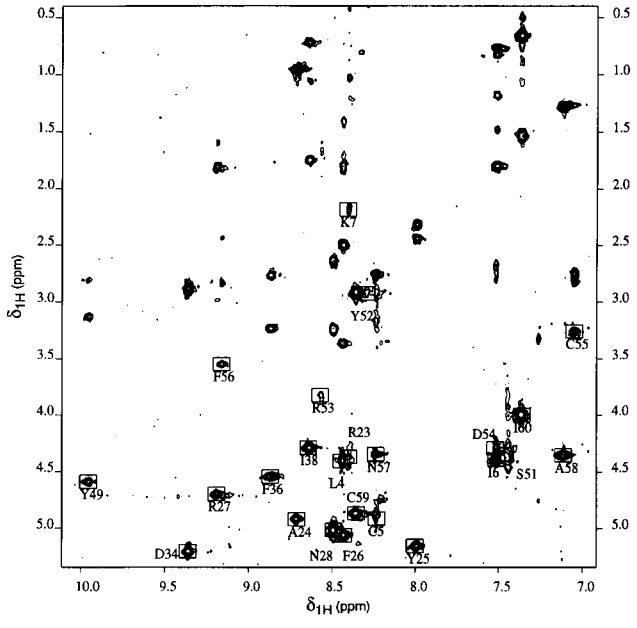


Fig. 4. The fingerprint region of a TOCSY spectrum acquired immediately after dissolving TAP in D_2O displaying peaks for the residues with slowly exchanging amide protons. The NH-αH crosspeaks are boxed and labeled with the letter code and sequence number of each contributing residue.

produce a set of final TAP structures. The inclusion of constraints based on the identified secondary structure elements was found to improve distance geometry computation speed without creating additional constraint violations. Energy refinement included constraint terms only for the observed NOE interactions and dihedral angles. The final refined structures were aligned with respect to the secondary structure elements (residues 2–7, 22–28, 32–38, 51–60) yielding average pairwise RMSDs of 0.87 ± 0.15 Å for backbone alignments. On average, the structures had 3.5 distance violations over 0.5 Å based on the relatively tight constraints (see the Materials and methods); these violations were mostly located in the regions with the fewest number of constraints. The backbone atoms of the aligned structures are shown in Figure 6. It is clear that the quality of the structure is not uniform, with loops 8–22 and 39–47 not well defined. This is due, in part, to the absence of observable NOEs for the broad resonances in these regions.

The refined structures, one of which is depicted in Figure 7, show several interesting features. Lys 7 is seen to lie in an aromatic groove composed of residues Trp 11, Phe 26, and Phe 36, which accounts for the large upfield shift in its side-chain resonances. There is motion in this region that is seen by the broadness of side-chain resonances for Lys 7, Pro 8, and Phe 26. Although not broadened, Phe 36 shows NOEs to both Asp 34 and Ile 38 that can only occur if its side chain moves between conformations pointing up and down the β-strand. In some structures, Asp 13 is able to form a salt bridge with Lys 7, a more likely possibility at pH 7 than under the acidic conditions of this work. The large upfield shifts for the αH resonances of residues Tyr 52 and Cys 55 can be explained by the ring current effects of Tyr 25 and Tyr 49, respectively. In the region of Cys 15 and Cys 39, where resonances are again broad, solutions are ob-

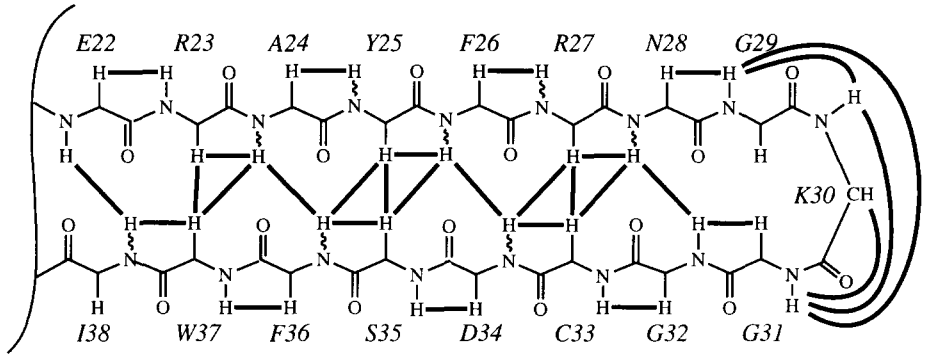


Fig. 5. Schematic representation of the backbone of the Glu 22 to Ile 38 segment of TAP, showing the sequential and long-range NOE connectivities indicative of an antiparallel β-sheet. The NOEs observed between atom pairs are depicted as bold lines. Only one αH is presented for each Gly residue, although NOEs were assigned to both. Slowly exchanging amide protons as observed in the TOCSY exchange experiment (see the Materials and methods) are indicated with wavy N-H bonds.

tained with both possible disulfide conformations and with relatively larger variations between structures.

Discussion

Sequence-specific proton assignments

Approximately two-thirds of the sequential proton assignments were developed with minimal difficulty based on reasonably sharp and intense crosspeaks in the TOCSY and NOESY spectra of the unlabeled protein. Three regions were substantially more troublesome due to broad, weak crosspeaks or the occasional absence of critical crosspeaks. Spin systems for residues 7-9, 13-16, and 35-40 were particularly difficult to identify and place sequentially. For residues 8, 9, 14, and 15, evidence was found for the presence of two sets of resonances. The broadness of residues near the Cys 15-Cys 39 disulfide bond suggests that two or more conformations are present in this region. Protons for which no or very weak resonances were found must be interconverting between environments at rates comparable to the chemical shift differences between their various conformations. As described below, TAP is structurally homologous to the Kunitz inhibitor class. It has been shown that for the Kunitz inhibitor BPTI, using rotating frame ^{15}N spin relaxation times, there is an interconversion of Cys 14-Cys 38 disulfide bond chirality with a correlation time of 1.3 ms for BPTI residues 38 and 39 coupled to an even faster segmental motion for BPTI residues 14-16 (Szyperski et al., 1993). From the crystallographic structure of the Alzheimer's amyloid β-protein precursor Kunitz domain, it is known that residue changes (in the absence of insertions or deletions) in the 39-41 region (BPTI numbering) can lead to significant changes in the backbone orientation in this region (Hynes et al., 1990). In the homologous region of TAP, the NMR evidence suggests that there is a slow and rather widespread motion that broadens the resonances of residues 13-16 and 35-40.

Secondary structure elements of TAP

Four regions of secondary structure are present in the TAP structures: a 3_{10} helix (residues 2-7), two β-strands (residues 22-28 and 32-38), and an α-helix (residues 51-60). The 3_{10} helix is not required to accommodate NOESY-derived constraints, but it is consistent with them, with the results of the amide proton-exchange experiments, and with the ϕ angles determined. It is also consistent with the structure of BPTI in this region (Marquart et al., 1983). The deduced structural elements and conventional amino acid homologies were used to align the sequence

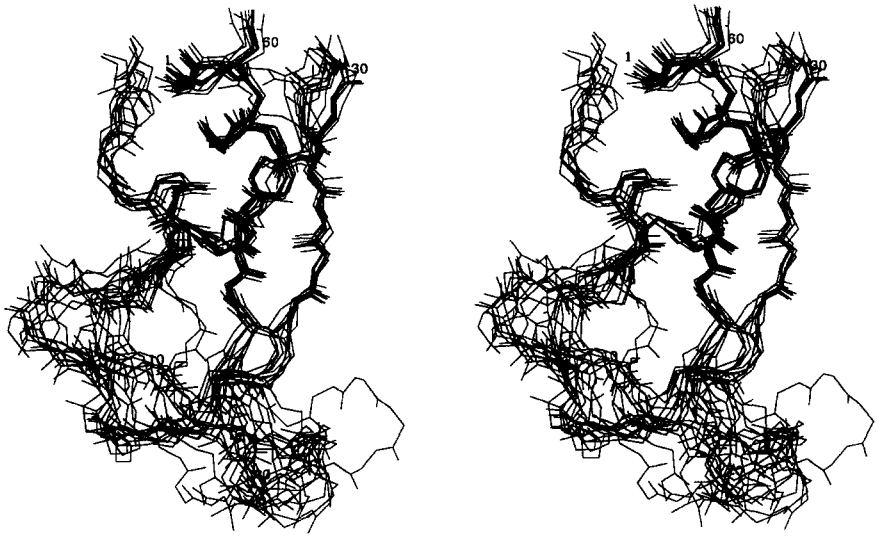


Fig. 6. Superposition of the backbone atoms of 10 refined TAP structures. The structures were aligned using backbone atoms (N, Cα, C, O) of residues 2-7, 22-28, 32-38, and 51-60 of each of the structures 2-10 on structure 1. Selected residue numbers are shown to provide orientation.

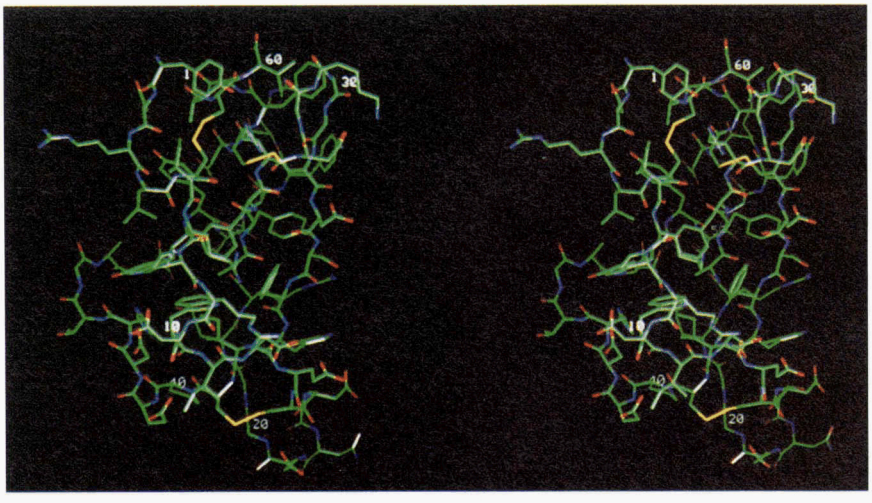


Fig. 7. Solution structure of TAP is shown with all side chains. Residues 1, 10, 20, 30, 40, 50, and 60 are numbered to provide an orientation for the text. Atoms are colored by atom type (C, green; O, red; N, blue; S, yellow). The orientation is identical to Figure 6.

of TAP with two Kunitz-domain proteins, BPTI (Wagner & Wüthrich, 1982) and *Stichodactyla helianthus* proteinase inhibitor (Antuch et al., 1993), for which NMR structures have been determined (Fig. 8). Four of the cysteine residues (TAP residues 5, 33, 55, and 59) are in conserved regions of secondary structure as determined by NMR. Cys 15 and Cys 39 are located in or near regions where sequence gaps occur and where multiple conformations are suggested by the spectra.

The slowly exchanging amide protons are generally located within the secondary structure elements (residues 5–7, 24, 26, 28, 34, 36, 38, and 55–60). Of those not in secondary structures, the following donor–acceptor pairs were observed in the refined structures: Leu 4–Asn 2Oδ or Tyr 49OH, Tyr 25–Tyr 48O, Arg 27–Cys 5O, Tyr 49–Asp 47Oδ, Ser 51–Asp 54Oδ, Arg 53 and Asp 54 share Ser 51Oγ. Only one slowly exchanging amide proton (Tyr 52) is unexplained by the structures.

Three-dimensional structure of TAP

The refined structures of TAP (Fig. 6) show good conservation of the well-defined secondary structure elements and considerable variation in the loop regions where resonances were broad at 25 °C. The refined structures (residues 22–28, 32–38, 51–60) were aligned against BPTI (residues 18–24, 29–35, and 47–56) in the regions of conserved secondary structure yielding average pairwise RMSDs 1.13 ± 0.16 Å for backbone atoms. One such structural alignment is shown in Figure 9, based on the coordinate set 4PTI (Marquart et al., 1983) from the Protein Data Bank (Bernstein et al., 1977; Abola et al., 1987). Note the significant differences in structure around the TAP 15–39 and BPTI 14–38 disulfide bonds due to the insertion of residues in TAP relative to BPTI (Figs. 6, 7). The N-terminal regions of TAP and BPTI also differ with respect to the orientation of the 3_{10} helices.

In contrast to the basic nature of BPTI, TAP at pH 7 is predicted to be highly dipolar with many of the anionic glutamic and aspartic acid residues located around the Cys 15–Cys 39 disulfide bond. Most of the cationic arginine and lysine residues are located on the side of the molecule opposite to the Cys 15–Cys 39 bridge. This may explain the tendency of TAP to aggregate at concentrations and pH higher than that of this work.

Having the refined structures, it is possible to address some of the questions about the unique manner in which TAP and fXa interact. A stereo view of one complete structure is shown in Figure 7. The Cys 15 and Cys 39 region of TAP, which is more variable in the refined structures, is homologous in sequence to the region of BPTI that represents the binding site for trypsin, shown as a molecular surface in Figure 9. However, single point

```
                    *         *         *         *         *         *
TAP          YNRLCIKPRDWIDECDSNEGGERAYFRNGKG GCDSFWI  CPEDHTGADYYSSYRDCFNACI

SecStr        HHHHHH              SSSSSSS    SSSSSSS             HHHHHHHHHH

                   *           *         *          *         *
BPTI         RPDFCLEPP YTGPCK AR  IIRYFYNAKAGLCQTFVYGGC   RAKRNNFKSAEDCMRTCGGA

                     *           *         *          *         *
ShPI           SICSEPK KVGRCK GY  FPRFYFDSETGKCTPFIYGGC   GGNGNNFETLHQCRAICRA
```

Fig. 8. Sequence alignment of TAP with two Kunitz-domain proteins for which NMR structures have been reported, shown with the single-letter amino acid codes. The alignment is based on both sequence and secondary structure homologies. The secondary structure elements of TAP (H, helix; S, β-strand) were determined from characteristic NOE patterns and amide proton-exchange experiments (see text). Above each sequence the asterisks denote residue numbers that are multiples of 10.

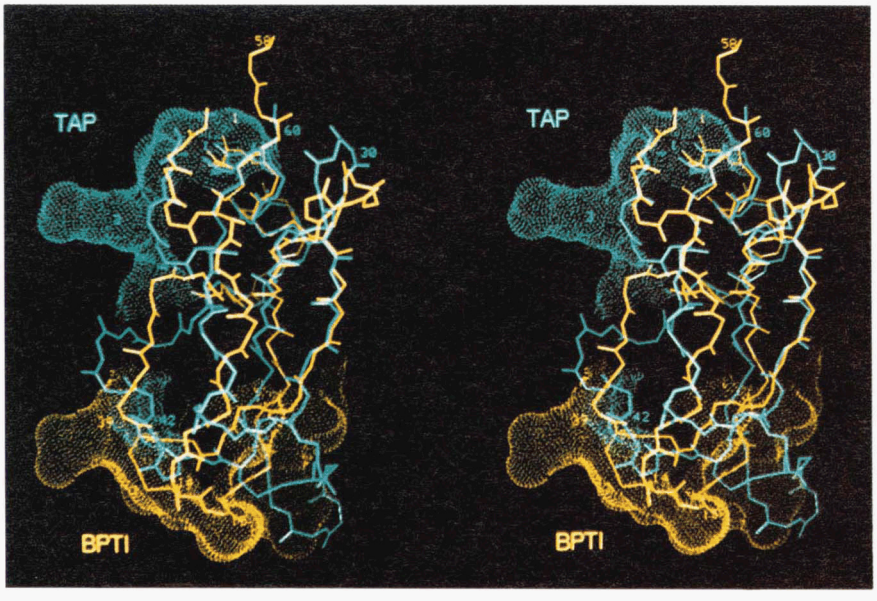


Fig. 9. Superposition of the backbone atoms of a refined TAP structure (cyan) and BPTI (yellow) (Marquart et al., 1983) aligned using backbone atoms (N, Cα, C, O) of residues 22–28, 32–38, and 51–60 of TAP with backbone atoms of residues 18–24, 29–35, and 47–56 of BPTI, respectively. Selected residue numbers of TAP (1–4, 30, 42, 60) and BPTI (1, 14, 15, 38, 39, 58) are shown to provide an orientation. Cystine side chains are depicted for each molecule. Solvent-accessible surfaces (probe radius 1.4 Å) are shown as dot surfaces for the critical binding residues of TAP (1–4, 42) and BPTI (11, 13–19, 34, 36–39). The orientation is identical to Figures 6 and 7.

mutations of TAP in this region (e.g., Asp 13 → Asn, Glu 14 → Gln, Asp 16 → Asn, Glu 19 → Gln) had less than twofold effects on fXa activity. Furthermore, the mutations Asp 16 → Arg and Asp 16 → Lys, at the position sequentially homologous to the primary specificity residue of BPTI, did not lead to trypsin activity (Dunwiddie et al., 1992). From the structural alignment of TAP and BPTI (Fig. 9) and from the observed internal motion of TAP in this region, it is perhaps not surprising that trypsin inhibition is not observed even in the Asp 16 → Arg mutation.

Point mutations of TAP replacing every charged residue with either Gln or Asn, and deletion mutants removing from one to four N-terminal residues have been made (Dunwiddie et al., 1992). The mutations of TAP that most affect the inhibition of fXa are near the N-terminus of TAP (i.e., Arg 3 → Asn: 40,000-fold potency reduction; removal of one to four N-terminal residues: 10,000–800,000-fold potency reduction), not near the known binding site on BPTI (residues 13–20). The solvent-accessible surface of residues 1–4 and 42 of TAP (whose mutation or deletion gave at least a twofold effect on fXa inhibition) are shown in Figure 9. All other point mutations had less than a twofold effect on fXa activity. Of all reported point mutations of TAP, the most effective for reducing activity was Arg 3 → Asn (Dunwiddie et al., 1992). The Cα–Cα distance from Arg 3 to Asp 16, the residue homologous to the primary specificity residue of BPTI, is about 22 Å, almost the length of the TAP molecule.

It is clear from Figure 9 that there is almost no three-dimensional overlap between TAP and BPTI in the binding sites for their cognate enzymes. Considering the structural similarities in the catalytic domain of fXa with other trypsin-like proteases (Padmanabhan et al., 1993) and the fact that TAP is not cleaved by fXa as BPTI is by trypsin, the binding modes of these two inhibitors to their cognate enzymes must be quite different. As would be expected from the foregoing discussion, the superposition of TAP on BPTI and fXa on trypsin as in the BPTI:trypsin complex shows that this binding mode cannot involve residues at the N-terminus of TAP.

The TAP mutagenesis work also revealed a weak secondary site of interaction from residues 40 to 54 (Dunwiddie et al., 1992). As can be seen from the solution structure (Figs. 7, 9), residues 44–47 are in a highly exposed loop on the same face of the molecule and in proximity to Arg 3, the site of most significant mutagenesis effects.

In a very general way, the solution structures are compatible with the reported structure–activity relationships. Recent work involving the modeling of docked TAP/fXa structures and newly designed mutants will be reported separately.

Thus, the current NMR-derived structural information supports the notion of significant structural similarities between TAP and other Kunitz proteins, with the region neighboring the Cys 15–Cys 39 disulfide bond of TAP a notable exception. Although TAP is structurally similar to other Kunitz proteins, the mutagenesis results (Dunwiddie et al., 1992) and the analysis presented here show that TAP must bind to fXa in a manner quite different from the way Kunitz proteins normally bind their cognate enzymes. In spite of its structural similarities to Kunitz proteins, TAP is clearly a novel inhibitor of fXa from which new information about serine protease inhibition can be developed.

Materials and methods

Materials

HPLC-purified recombinant TAP, ^{13}C-labeled TAP, and ^{15}N-labeled TAP, obtained from a *Pichia pastoris* expression system, was used in all experiments. The materials were characterized by mass spectrometry, N-terminal sequencing, and inhibitory activity against free fXa (Laroche et al., 1994). Quantitative determination of the extent of isotope incorporation was accomplished by ^{13}C-nondecoupled 1D proton NMR (>98% via integration of the αH resonance at 5.70 ppm) and by ^{15}N-nondecoupled 1D proton NMR (>98% via integration of the NH resonance at 9.32 ppm). Solutions were prepared by dissolving weighed amounts of TAP directly into H_2O/D_2O 90%/10% (v/v) or D_2O, as required by the experiment. The pH was lowered to 2.5 by addition of a solution of DCl in D_2O monitored by a pH meter equipped with a microelectrode. One microliter

of a solution containing 100 mM sodium 3-trimethylsilyl-(2,2,3,3, $^2H_4$)-propionate was added as an internal proton/carbon standard. In all cases the total sample volume was 500 μL.

NMR experiments

All data were acquired at 25 °C, unless otherwise noted, on a two-channel Varian 600-MHz Unity spectrometer equipped with waveform generators and a standard variable temperature unit. Proton experiments in D_2O were performed either immediately after preparation of the sample, to study amide proton exchange, or after two cycles of lyophilization from D_2O.

TOCSY experiments used mixing times ranging from 30 ms to 90 ms with the clean-TOCSY variation (Bearden et al., 1988), using a spinlock field of 15 kHz in the MLEV17 mode including 1.5-ms trim pulses. A coherent presaturation field (Zuiderweg et al., 1986) of 5 Hz in D_2O and 30 Hz in H_2O was applied during the relaxation delay of 0.7 s. Partial recovery of α-proton intensity was achieved by introducing a $\tau_1 - 180 - \tau_2$ sequence after the presaturation, where τ_1 and τ_2 were 0.08 s (Brown et al., 1987).

NOESY experiments used a series of mixing times ranging from 60 to 240 ms and a relaxation time of 0.7 s. Solvent suppression for experiments in H_2O was accomplished with the NOESY-11 sequence (Sklenar & Bax, 1987) and a 10-Hz presaturation during the relaxation delay and mixing time, with α-proton intensity recovery as described previously. Careful adjustment of the initial delays in the evolution as well as in the detection time was used to obtain either a zero or a $-180°$ first-order phase correction and a flatter baseline (Zhou et al., 1993). The NOESY experiments in D_2O were performed with 5-Hz presaturation during a 0.7-s relaxation delay and the mixing time, again with recovery of a proton intensity. In addition, a spin-echo sequence ($\tau_1 - 180 - \tau_1$, where τ_1 was 10 μs) was introduced before detection to flatten the baseline (Lippens & Hallenga, 1990).

The heteronuclear 2D correlation experiments were performed as HSQC and HSQC-NOESY in the case of ^{15}N (Bodenhausen & Ruben, 1980; Norwood et al., 1990) and as CT-HSQC in the case of ^{13}C (Vuister & Bax, 1992; Hallenga & Lippens, 1994) with a constant evolution time of 17 ms. The constant evolution time ^{13}C experiments, in which both aliphatic and aromatic regions can be observed with full intensity and without artifacts, produces correlation peaks of opposite sign for primary and tertiary, versus secondary, carbons.

Amide proton exchange

As a crude measure of proton exchange, a TOCSY spectrum was acquired during the first 16 h immediately after sample preparation in D_2O. A NOESY spectrum was acquired subsequently, between 21 and 45 h following sample dissolution, to identify more slowly exchanging protons.

Determination of dihedral angles and stereospecific assignments

Constraints for the ϕ dihedral angles were obtained by estimating $^3J_{N\alpha}$ from the splitting of the H_N-H_α crosspeak doublets in homonuclear TOCSY and NOESY spectra of TAP. The ϕ angles were not derived for the Gly residues because the α protons were not stereospecifically assigned, except for Gly 21, where both $^3J_{N\alpha}$ were small. Residues with $^3J_{N\alpha}$ values less than 5 Hz were assigned as having ϕ angles of $-60°$, whereas those with $^3J_{N\alpha}$ greater than 9 Hz were assigned with $\phi = -120°$.

Knowing all four $^3J_{N\beta}$ and $^3J_{\alpha\beta}$ for the AMX spin systems and both $^3J_{N\beta}$ and $^3J_{\alpha\beta}$ for the Ile residues, it was possible to simultaneously obtain χ^1 dihedral angles and the stereospecific assignments for these β protons (Wagner, 1990). $^3J_{N\beta}$ was obtained from the $\omega 2$ separation between the two multiplet components of the H_N-H_β crosspeaks for each H_β, in a homonuclear TOCSY spectrum of ^{15}N-labeled TAP (Wagner, 1990). $^3J_{N\beta}$ of magnitudes less than 2 Hz and greater than 4 Hz were considered as small and large coupling constants, respectively. The multiplicity of each H_N-H_β peak in the same spectrum was used to estimate $^3J_{\alpha\beta}$; H_β *trans* to H_α had triplet character in $\omega 2$ in the crosspeak from the additional $^2J_{\beta\beta}$ coupling, whereas H_β *cis* to H_α had the appearance of a doublet. Only residues with clearly triplet or doublet splittings were assigned large (~12 Hz) or small (~3 Hz) coupling constants, respectively. The $^3J_{\alpha\beta}$ values were also obtained and confirmed from the H_α-H_β crosspeak splitting in $\omega 2$ in D_2O NOESY and TOCSY spectra (Wagner, 1990).

Structure determination

NOESY crosspeaks were assigned visually to three classes (strong, medium, and weak) from a 180-ms H_2O and from 60-ms and 180-ms D_2O spectra. Upper-bound distance constraints were assigned as follows for the three categories: strong, 3.0 Å; medium, 4.0 Å; and weak, 5.0 Å. For distance geometry and refinement calculations, these upper bounds were adjusted (+1.4 Å) for the aromatic pseudo-atoms, where necessary. Only 24 very strong intraresidue NOEs (e.g., NH-$\beta,\gamma,\delta,\epsilon$ or αH-δ,ϵ) were included as indications of side-chain orientation.

For distance geometry calculations, the NMR-derived constraints were augmented by distance and chiral volume constraints that (1) represented proper disulfide bonds, (2) forced amide bonds to be planar and *trans*, (3) constrained ϕ and $\chi 1$ dihedral angles to those observed, and (4) constrained elements of secondary structure (helices and sheets) to their approximate standard geometries. The secondary structure constraints were added iteratively based on the observed NOESY crosspeaks and the amide exchange results. At each stage, it was determined that the secondary structure constraints did not cause additional violations of the pure experimental constraints. DGEOM (Quantum Chemistry Program Exchange, Indiana University) was used to perform the computations, and the initial structure was a linear polypeptide generated using InsightII/Discover (Biosym Technologies, San Diego, California). Only heavy atoms, and the protons NH and αH, were included explicitly in the distance geometry calculations.

The assignment of NOESY crosspeaks and their inclusion in the data set were performed iteratively with distance geometry calculations to resolve ambiguous assignments and to eliminate obvious inconsistencies due to spin diffusion and side-chain motion. A total of 412 distance and 47 dihedral constraints were determined experimentally for use in the final structure refinement.

Ten distance geometry solutions, selected from a set of 50 to have minimal constraint violations, were refined using Discover. For these refinements, the direct NMR distance and dihedral

constraints were augmented only with constraints to force amide bonds to be *trans* because these are not guaranteed to be *trans* by DGEOM. The proteins were protonated as they would be at pH 2 and a distance-dependent dielectric constant was used. The refinement consisted sequentially of (1) 500 steps of conjugate gradient minimization, (2) 1,000 steps of dynamic simulation at 450 °C, and (3) 2,000 steps of conjugate gradient minimization, all with weak quadratic energy terms (15 kcal/mol/Å^2 force constant) for the distance constraints and weak (20 kcal/mol/rad^2) force constants for the dihedral constraints. The distance violation force constant has a value in the Discover force field that would be about 5% of a C–C bond or equivalent to a weak hydrogen bond. Qualitatively, the bounds are relatively narrow and the forces relatively weak to ensure that misassignments or structural problems would be detected.

Graphics display, violations statistics, and RMSD comparisons were performed with the program InsightII/NMR_Refine (Biosym Technologies, San Diego, California).

Supplementary material in Electronic Appendix

Two files are included in the Electronic Appendix for this report (SUPLEMNT directory, LimWilby.SUP subdirectory). File LimWilby.TS1 lists all ^{1}H and ^{15}N assignments and all ^{13}Cα and aromatic assignments. File LimWilby.NMR is a complete listing of the Biosym formatted restraint file used in the structure refinement.

Acknowledgments

We thank Yves Laroche, Jorris Messens, and Marc Lauwereys for providing samples of TAP, ^{13}C-, and ^{15}N-labeled TAP. We also thank Bill Ripka for continuous support and encouragement during this project.

References

Abola EE, Bernstein FC, Bryant SH, Koetzle TF, Weng J. 1987. Protein Data Bank. In: Allen FH, Bergerhoff G, Sievers R, eds. *Crystallographic databases – Information content, software systems, scientific applications*. Bonn/Cambridge/Chester: Data Commission of the International Union of Crystallography. pp 107–132.

Antuch W, Berndt KD, Chavez MA, Delfin 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.

Antuch W, Güntert P, Billeter M, Hawthorne T, Grossenbacher H, Wüthrich K. 1994. NMR solution structure of the recombinant tick anticoagulant protein (rTAP), a factor Xa inhibitor from the tick *Ornithodoros moubata*. *FEBS Lett 352*:251–257.

Bearden DW, Macura S, Brown LR. 1988. Suppression of cross relaxation in TOCSY experiments on macromolecules. *J Magn Reson 80*:534–538.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett 69*:185–189.

Brown SC, Weber PL, Mueller L. 1987. Towards complete 1H NMR spectra in proteins. *J Magn Reson 77*:166–169.

Dunwiddie CT, Neeper MP, Nutt EM, Waxman L, Smith DE, Hofmann KJ, Lumma PK, Garsky VM, Vlasuk GP. 1992. Site-directed analysis of the functional domains in the factor Xa inhibitor tick anticoagulant peptide: Identification of two distinct regions that constitute the enzyme recognition sites. *Biochemistry 31*:12126–12131.

Dunwiddie CT, Smith DE, Nutt EM, Vlasuk GP. 1991. Anticoagulant effects of the selective factor Xa inhibitors tick anticoagulant peptide and antistasin in the APTT assay are determined by the relative rate of prothrombinase inhibition. *Thromb Res 64*:787–794.

Hallenga K, Lippens GM. 1994. A constant time ^{13}C-^{1}H HSQC with uniform sensitivity over the complete ^{13}C chemical shift range. *J Biomol NMR*. Forthcoming.

Hynes RH, Randal M, Kennedy LA, Eigenbrot C, Kossiakoff AA. 1990. X-ray crystal structure of the protease inhibitor domain of Alzheimer's amyloid β-protein precursor. *Biochemistry 29*:10018–10022.

Jordan SP, Mao SS, Lewis SD, Shafer JA. 1992. Reaction pathway for inhibition of blood coagulation factor Xa by tick anticoagulant peptide. *Biochemistry 31*:5374–5380.

Jordan SP, Waxman L, Smith DE, Vlasuk GP. 1990. Tick anticoagulant peptide: Kinetic analysis of the recombinant inhibitor with blood coagulation factor Xa. *Biochemistry 29*:11095–11100.

Krishnaswamy S, Vlasuk GP, Bergum PW. 1994. Assembly of the prothrombinase complex enhances the inhibition of bovine factor Xa by tick anticoagulant peptide. *Biochemistry 33*:7897–7908.

Laroche Y, Storme V, De Meutter J, Messens J, Lauwereys M. 1994. High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotropic yeast, *Pichia pastoris*. *Biotechnology 12*:1119–1124.

Lippens G, Hallenga K. 1990. Perfectly flat baselines in 1D and 2D spectra with optimized spin-echo detection. *J Magn Reson 88*:619–625.

Marquart M, Walter J, Deisenhofer J, Bode W, Huber R. 1983. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen, and its complexes with inhibitors. *Acta Crystallogr B 39*:480–490.

Neeper MP, Waxman L, Smith DE, Schulman CA, Sardana M, Ellis RW, Schaffer LW, Siegl PKS, Vlasuk GP. 1990. Characterization of recombinant tick anticoagulant peptide. *J Biol Chem 265*:17746–17752.

Norwood TJ, Boyd J, Heritage JE, Soffe N, Campbell ID. 1990. Comparison of techniques for 1H detected heteronuclear 1H-15N spectroscopy. *J Magn Reson 87*:488–501.

Nutt E, Gasic T, Rodkey J, Gasic G, Jacobs J, Friedman P, Simpson E. 1988. The amino acid sequence of antistasin. *J Biol Chem 263*:10162–10167.

Padmanabhan K, Padmanabhan KP, Tulinsky A, Park CH, Bode W, Huber R, Kisiel W. 1993. Structure of human des(1-45) factor Xa at 2.2 Å resolution. *J Mol Biol 232*:947–966.

Sardana M, Sardana V, Rodkey J, Wood T, Ng A, Vlasuk GP, Waxman J. 1991. Determination of disulfide bond pairs and stability in recombinant tick anticoagulant peptide. *J Biol Chem 266*:13560–13563.

Schaffer LW, Davidson JT, Vlasuk GP, Siegl PKS. 1991. Antithrombotic efficacy of recombinant tick anticoagulant peptide. *Circulation 84*:1741–1748.

Sitko GR, Ramjit DR, Stabilito II, Lehman D, Lynch JJ, Vlasuk GP. 1992. Conjunctive enhancement of enzymatic thrombolysis and prevention of thrombotic reocclusion with the selective factor Xa inhibitor, tick anticoagulant peptide. *Circulation 85*:805–814.

Sklenar V, Bax A. 1987. Spin-echo water suppression for the generation of pure phase two-dimensional NMR spectra. *J Magn Reson 74*:469–483.

Szyperski T, Luginbuhl P, Otting G, Guntert P, Wüthrich K. 1993. Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *J Biomol NMR 3*:151–164.

Vlasuk GP. 1993. Structural and functional characterization of tick anticoagulant peptide (TAP): A potent and selective inhibitor of blood coagulation factor Xa. *Thromb Haemostasis 70*:212–216.

Vlasuk GP, Ramjit D, Fujita T, Dunwiddie CT, Nutt EM, Smith DE, Shebuski RJ. 1991. Comparison of the in vivo anticoagulant properties of standard heparin and the highly selective factor Xa inhibitors antistasin and tick anticoagulant peptide (TAP) in a rabbit model of venous thrombosis. *Thromb Haemostasis 65*:257–262.

Vuister G, Bax A. 1992. Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J Magn Reson 98*:428–435.

Wagner G. 1990. NMR investigations of protein structure. *Prog NMR Spectrosc 22*:101–139.

Wagner G, Wüthrich K. 1982. Sequential resonance assignments in protein ^{1}H nuclear magnetic resonance spectra. *J Mol Biol 155*:347–366.

Waxman L, Smith DE, Arcuri KE, Vlasuk GP. 1990. Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation factor Xa. *Science 248*:593–596.

Wun TC, Kretzmer K, Girard TJ, Miletich JP, Broze GJ. 1988. Cloning and characterization of a cDNA coding for the lipoprotein-associated coagulation inhibitor shows that it consists of three tandem Kunitz-type inhibitory domains. *J Biol Chem 263*:6001–6004.

Zhou G, Torchia DA, Bax A. 1993. Discrete Fourier transformation of NMR signals. The relationship between sampling delay time and spectral baseline. *J Magn Reson 105*:219–222.

Zuiderweg ERP, Hallenga K, Olejniczak ET. 1986. Improvement of 2D NOE spectra of biomolecules in H_2O solution by coherent suppression of the solvent resonance. *J Magn Reson 70*:336–343.