# Thrombin-Bound Structure of an EGF Subdomain from Human Thrombomodulin Determined by Transferred Nuclear Overhauser Effects[†,‡]

Jayashree Srinivasan,[§] Song Hu,[∥] Richard Hrabal,[§] Yi Zhu,[∥] Elizabeth A. Komives,[⊥] and Feng Ni*[,§]

*Biotechnology Research Institute, National Research Council Canada, Montreal, Quebec, Canada H4P 2R2, Department of Biochemistry, McGill University, Montreal, Quebec, Canada H3G 1Y6, and Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, California 92093-0601*

*Received August 12, 1994; Revised Manuscript Received September 14, 1994*[⊗]

ABSTRACT: The EGF-like domains in human thrombomodulin interact with and change the specificity of thrombin from a procoagulant enzyme to an anticoagulant enzyme. Recent experiments identified the minimal thrombin-binding region of thrombomodulin as the most acidic loop of the fifth EGF-like domain with a sequence of $E_{408}$CPEGYILDDGFI$_{420}$CTDIDE. High-resolution NMR spectroscopy was employed to characterize the interaction of a des-Ile420 thrombomodulin peptide, $Cys_{1(409)}Pro_2Glu_3Gly_4Tyr_5Ile_6$-$Leu_7Asp_8Asp_9Gly_{10}Phe_{11}Cys_{12}Thr_{13}Asp_{14}Ile_{15}Asp_{16}Glu_{17(426)}$, with its target coagulation protein, thrombin. The disulfide-bonded peptide was found to be structured only upon binding, while neither the linear nor the cyclized peptide exhibited any structural preference free in solution. The thrombin-bound structure of the cyclic thrombomodulin peptide was determined by transferred nuclear Overhauser effects (transferred NOEs) and by distance geometry and Monte Carlo calculations. The thrombin-bound cyclic peptide assumes an overall conformation similar to those observed in the free but intact EGF molecules. There is a type II $\beta$-turn involving residues Pro2−Tyr5, followed by an optimized antiparallel $\beta$-sheet involving residues Gly4−Asp8 and residues Phe11−Ile15. The thrombomodulin peptide provides a potential thrombin-binding surface between residues Tyr5 and Phe11, which are brought close by a chain reversal within the central $\beta$-sheet. Comparison of the thrombin-bound structure of the EGF-like subdomain with other thrombin−peptide complexes revealed that a common thrombin-binding surface can be organized by different secondary structure elements with entirely different peptide sequences. The thrombin-bound structure of the thrombomodulin peptide may serve as a basis to understand the regulatory functions of thrombomodulin and as a guide for the design of specific inhibitors for thrombin.

The epidermal growth factor- (EGF-)[1] like sequences represent modular domains of ~50 amino acid residues found in many proteins (Campbell & Bork, 1993). The EGF protein itself mediates cell growth by binding to the extracellular domain of the EGF growth factor receptor. The EGF-like domains in plasminogen activators interact with specific cell surface receptors, initiating various cellular events. The EGF repeats are also found in many membrane-bound proteins (or cell surface receptors) believed to be important for the mediation of specific protein−protein interactions (Engel, 1989).

Thrombomodulin is a membrane-bound glycoprotein containing six EGF-like motifs exposed in its extracellular structure (Esmon, 1989). It binds to thrombin via the EGF-like domains and changes the specificity of the latter from a procoagulant enzyme to an anticoagulant enzyme. The binding of thrombin to thrombomodulin results in as much as 20 000-fold faster thrombin-catalyzed protein C activation while the thrombin−thrombomodulin complex no longer interacts with the natural thrombin substrate fibrinogen. The fourth, fifth, and sixth EGF domains are responsible for the thrombin−thrombomodulin interaction as shown by use of natural and recombinant EGF domains derived from thrombomodulin (Kurosawa et al., 1988; Zushi et al., 1989; Hayashi et al., 1990; Ye et al., 1992; Nagashima et al., 1993). These studies were complemented by competitive binding experiments performed with synthetic peptides that span different regions of the EGF-like domains (Hayashi et al., 1990; Tsiang et al., 1990, 1992). Binding experiments with synthetic peptides also identified the minimal thrombin binding region as the most acidic loop of the fifth EGF-like domain with a sequence of ECPEGYILDDGFICTDIDE (Hayashi et al., 1990; Tsiang et al., 1990, 1992). This peptide inhibited the thrombin−thrombomodulin and the thrombin−fibrinogen interactions with an apparent $K_i$ of 85−95 $\mu$M.

[†] NRCC Publication No. 36845. This work was supported in part by research grants from Ciba-Geigy Canada Ltd., from the Natural Science and Engineering Research Council (NSERC), and from the Medical Research Council (MRC) of Canada. Support for E.A.K. from the Searle Scholars Program is gratefully acknowledged.

[‡] The atomic coordinates of the structure have been deposited with the Brookhaven Protein Data Bank, Accession Number 1EGT.

* Author to whom correspondence should be addressed.

[§] National Research Council Canada.

[∥] McGill University.

[⊥] University of California at San Diego.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1994.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; ROESY, two-dimensional rotating-frame nuclear Overhauser spectroscopy; TOCSY, total correlation spectroscopy; DQF-COSY, double-quantum-filtered *J*-correlated spectroscopy; FID, free induction decay; EDTA, ethylenediaminetetraacetic acid; PPACK, D-Phe-Pro-Arg-$CH_2Cl$; EGF, epidermal growth factor; HPLC, high-pressure liquid chromatography.

0006-2960/94/0433-13553$04.50/0 

Despite recent advances in the structure−function relationships of EGF-like domains in proteins (Campbell & Bork, 1993), lack of structural information of the EGF-like domains in complex with their respective targets makes it difficult to predict how changes in amino acids between different EGF domains can be related to their specific functions. We employed high-resolution NMR methods to characterize the interaction of a thrombomodulin-derived EGF subdomain with its target protein, thrombin. Results based on inhibitory activities (Lougheed et al., 1994) and NMR relaxation enhancements showed that only the cyclized synthetic fragment binds to the thrombin exosite specifically. More interestingly, the cyclized peptide was found to be structured upon binding to thrombin while neither the linear nor the cyclized peptide exhibited any structural preference free in solution. In this paper, we present the thrombin-bound structure of the cyclic thrombomodulin peptide as determined by tranferred nuclear Overhauser effects (transferred NOEs). Comparison of the thrombin-bound structure of the EGF-like subdomain with other thrombin−peptide complexes revealed how a common thrombin-binding surface can be organized by different secondary structure elements with entirely different peptide sequences.

## MATERIALS AND METHODS

*Peptides, Proteins, and NMR Sample Preparation.* An analog thrombomodulin peptide with a sequence of CPEGY-ILDDGFCTDIDE was synthesized using solid-phase Fmoc chemistry on a Milligen 9050 peptide synthesizer. This peptide corresponds to the C-terminal loop (residues 409−421) of the fifth EGF-like domain of human thrombomodulin plus the connecting sequence TDIDE leading to the sixth EGF-like domain but with a deleted Ile420 residue immediately before the second Cys residue. The analog peptide contained two cysteine residues that are cyclized to form the loop structure which was postulated to be present in the EGF-like domains. Portions of the peptide were reduced and alkylated with iodoacetamide to provide the linear control peptide. The rest of the peptide was oxidized in dilute aqueous solution to produce the intramolecularly disulfide-bonded cyclic peptide. After HPLC purification, the identity of each peptide was confirmed by FAB mass spectrometry, by amino acid analysis, and by N-terminal sequencing. The molecular weights were as predicted on the basis of the amino acid sequence: cyclic peptide, 1903 (expected, 1902.8), and linear peptide, 2019 (expected, 2018.8). The linear or cyclic peptide was also shown to be devoid of any free thiol groups using Ellman's reagent (dithionitrobenzoate). Very importantly, these analog peptides showed the same or better activity compared to the full-length loops; the cyclic peptide had a $K_i$ of 26.5 $\mu$M but had the attribute of being much more soluble than their full-length counterparts (Lougheed et al., 1994).

Bovine α-thrombin was purified by cation-exchange chromatography from the bovine plasma barium citrate eluate (Sigma), containing primarily prothrombin and minor amounts of factors X, IX, and VII, after controlled activation by *Echis carinatus* snake venom (Sigma) (Ni et al., 1990). Stock solutions (10−25 mg/mL) of thrombin were prepared by concentrating the purified enzyme with CentriCon-10 concentration cells (Amicon) and stored at −70 °C before NMR experiments. About 1.2−2 molar equiv of the inhibitor, D-Phe-Pro-Arg-$CH_2Cl$ (PPACK), was used to block the active site of thrombin so that thrombin autolysis was minimized during NMR experiments.

The NMR samples were prepared in an aqueous solution that was 150 mM in sodium chloride, 50 mM in sodium phosphate, and 0.2 mM in EDTA. A sample of the free cyclic peptide was prepared by dissolving 5.4 mg (80% peptide content plus dried salts) in 450 $\mu$L of the aqueous solution with the pH adjusted using dilute HCl or NaOH solutions. A volume of 50 $\mu$L of $D_2O$ was added to this pH-adjusted sample to provide the deuterium lock signal for the NMR spectrometer. Samples of the thrombin−peptide complexes were prepared by adding a stock solution of PPACK−thrombin to a pH-adjusted and dried sample of the peptides. The sample for the cyclic peptide had a peptide-to-thrombin molar ratio of ~10:1 with a thrombin concentration of 0.5 mM while the linear peptide and thrombin concentrations were 4.2 and 0.3 mM, respectively.

*NMR Experiments.* One- and two-dimensional NMR experiments were carried out on a Bruker AM and/or AMX 500-MHz NMR spectrometer, as described previously (Ni et al., 1990, 1992a; Ni, 1992b). Briefly, all two-dimensional NMR data were acquired using time-proportional phase incrementation (TPPI) along the $t_1$ dimension, with a relaxation delay of 1.5−2.0 s, to allow for a more complete magnetization recovery. The spectra were acquired at a controlled temperature of 25 °C without spinning the sample. Sine modulation was applied with an adjusted initial delay to avoid the occurrence of ridges along the $t_1$ dimension (Ni, 1992b). Water suppression was achieved by selectively saturating the solvent signal during relaxation delay and during the mixing time of the NOESY sequence. A jump−return sequence combined with presaturation was used to suppress the water resonance in the ROESY/TOCSY spectra. Each FID of the two-dimensional data required typically 64−96 data accumulations. Spectral processing was carried out using the FELIX software program (Hare Research) and/or an in-house program nmrDSP on a Silicon Graphics 4D/35 workstation. Fourier transformation of the two-dimensional FID data along the $t_2$ dimension involved premultiplication with a cosine square window function and a subtraction procedure to further suppress water signals. Processing along the $t_1$ dimension involved linear prediction to 450 points from the 400 experimental FID slices before premultiplication with a Kaiser window function. A baseline correction procedure was also performed to improve the quality of the 2D NMR spectra (Ni, 1992b).

Sequence-specific assignments of the proton resonances of the cyclic thrombomodulin peptide (Table S1) were determined using standard procedures (Wüthrich, 1986) by the analysis of the cross-peak patterns in the DQF-COSY, TOCSY, and ROESY spectra of the free peptide along with the NOESY spectrum of the same peptide in the presence of α-thrombin. All the chemical shift values were determined from the NOESY spectra of the peptide in the presence of thrombin with a reference chemical shift of 4.80 ppm for the water solvent resonance. These chemical shift values approach those of the free peptide as a result of the low concentration of thrombin used in transferred NOE studies.

*Structural Calculations.* A variable target function (VTF) implementation (Vásquez & Scheraga, 1988; Ni et al., 1989) of distance geometry was used to generate peptide conforma-

tions that are free of steric overlaps and that are consistent with the NMR information. The covalent geometries were fixed at their equilibrium values as documented in the ECEPP/2 (Emprical Conformational Energy Parameters for Peptides) library (Vásquez & Scheraga, 1988). The complete set of backbone and side-chain dihedral angles were considered as independent variables during minimization of a distance target function or the ECEPP/2 potential energy (Vásquez & Scheraga, 1988; Ni et al., 1989). The distance penalty function explicitly took into account groups of protons with overlapping resonances by requiring a proper average of individual distances to satisfy both the upper and lower bounds (Ni et al., 1989). Resolution of some of the overlapped resonances was achieved in an iterative manner (Ni et al., 1990). This involved computing a set of initial structures by using distance constraints derived from well-resolved NOEs. These initial structures provided information for the identification of some overlapped NOEs to specific proton pairs. Further structure refinement was carried out by restrained Monte Carlo conformational sampling and energy minimization (Ripoll & Ni, 1992). The side chains of aspartic and glutamic acid residues were treated as charged groups for energy computations, but with a dielectric constant of 4 for electrostatic interactions at short distances (<5 Å) and a dielectric constant of 78 for longer distances (>13 Å), implemented through the use of a sigmoidal dielectric function (Daggett et al., 1991).

*Simulation of Transferred NOE Spectra.* Theoretical transferred NOEs were calculated for all the model structures of the cyclic thrombomodulin peptide using a program PDB2NOE for the complete relaxation matrix analysis of multi-spin-exchanging systems including transferred NOEs (Ni, 1992a, 1994; Ni & Zhu, 1994). The tumbling correlation time for the thrombin–peptide complex was taken to be 20 ns, expected for a globular protein of molecular weight 40 000 (Ni et al., 1992a). A short correlation time of 0.356 ns was assumed for the peptide in accordance with the close-to-null NOEs observed for the free peptide. The ligand dissociation rate constant $k_{off}$ was assumed to be 200 $s^{-1}$ since, in this exchange regime, transferred NOEs from the thrombin–peptide complexes are no longer very sensitive to the exchange kinetics (Ni & Zhu, 1994). The dissociation constant $K_d$ was set to a value of 0.03 mM, estimated on the basis of the $K_i$ value of 26.5 $\mu$M for the interaction of the peptide with thrombin (see the previous section). The exact values for $K_d$ do not affect the transferred NOE calculations since the binding sites are saturated with the peptides at the high ratios of peptide to thrombin concentrations adopted in the current study. The computed NOE intensities were converted to two-dimensional FIDs by use of an in-house program, GFIDSJ. Calculated FIDS were Fourier transformed with identical processing procedures as used for the experimental NOE spectra.

## RESULTS AND DISCUSSION

*Transferred Nuclear Overhauser Effects.* The 17-residue cyclic thrombomodulin peptide $Cys_1Pro_2Glu_3Gly_4Tyr_5Ile_6$-$Leu_7Asp_8Asp_9Gly_{10}Phe_{11}Cys_{12}Thr_{13}Asp_{14}Ile_{15}Asp_{16}Glu_{17}$ itself exhibited very few NOE interactions in aqueous solution (spectra not shown). On the other hand, a large number of NOE cross peaks were observed for the same peptide in the presence of less than stoichiometric amounts of thrombin (Figure 1A). Most of these NOE cross peaks were not

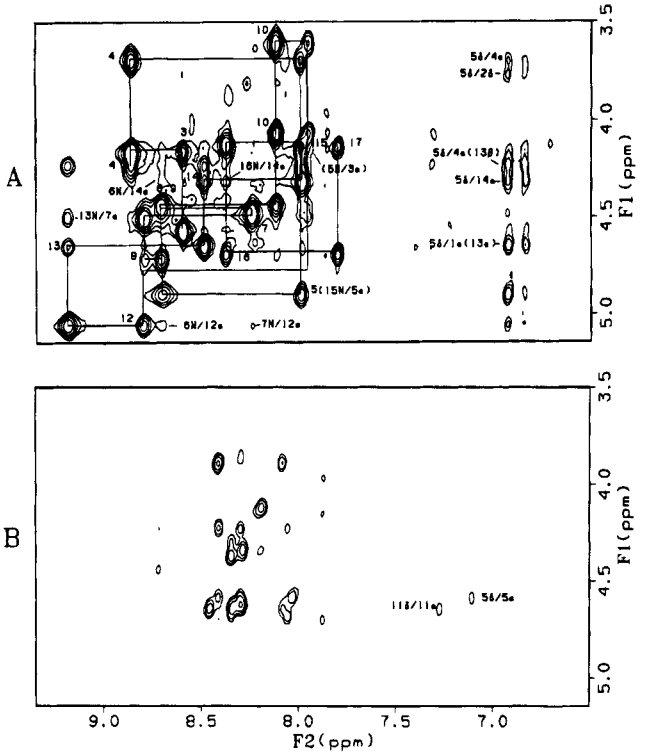


FIGURE 1: Fingerprint regions of the NOESY spectra of the cyclic thrombomodulin peptide (A) and the linear counterpart (B) in complex with thrombin. Both spectra were acquired with a mixing time of 200 ms and a sample pH of 5.5 at 25 °C. The peptide concentrations were 5.5 mM (cyclic peptide) with a peptide–thrombin molar ratio of 11:1 and 4.2 mM (linear peptide) with a peptide–thombin molar ratio of 15:1. (A) Sequential αCH–NH connectivities for the cyclic peptide with each intraresidue NH–αCH NOE labeled by the residue number (see Figure 2A). The long-range NOEs are indicated by the corresponding assignment. Assignments in parentheses refer to NOEs which were unresolved at pH 5.5 due to partial peak overlap but were resolved in a sample at pH 7.2. The NOE with the label 5δ/1α(13α) indicates that the αCH proton of Cys1 overlaps with that of Thr13 and similarly for 5δ/4α(13β). The αCH proton of Phe11 overlaps with that of the water solvent peak (Table S1) and was thus attenuated by solvent presaturation (see also Figure 3). (B) Corresponding NOEs for the linear peptide. Long-range NOEs involving Tyr5 are absent in this spectrum. NOESY spectra of the cyclic peptide in the absence of thrombin, conducted under identical conditions, showed no cross peaks in the spectral regions represented in (A) and (B).

observed when the linear peptide was analyzed in the same manner (Figure 1B), indicating that they are thrombin-induced transferred NOEs (Albrand et al., 1979; Ni & Scheraga, 1994) unique to the cyclic peptide. The appearance of transferred NOEs indicates that there is a dynamic exchange for the cyclic thrombomodulin peptide between the free and the thombin-bound states, in agreement with results based on differential line broadening of peptide resonances (Figure S1B). In contrast, the same cyclic peptide did not bind to γ-thrombin (Figure S1C), a proteolyzed form of thrombin with an impaired fibrinogen recognition exosite (Ni et al., 1993). In a separate experiment, the cyclic peptide protected α-thrombin from being autolyzed at the exosite while the linear peptide failed to prevent thrombin autolysis (SDS gels not shown). All these results indicated that the cyclic thrombomodulin peptide binds specifically to the fibrinogen recognition exosite of thrombin, in agreement with results based on competitive inhibition studies (Tsiang et al., 1990; Lougheed et al., 1994).

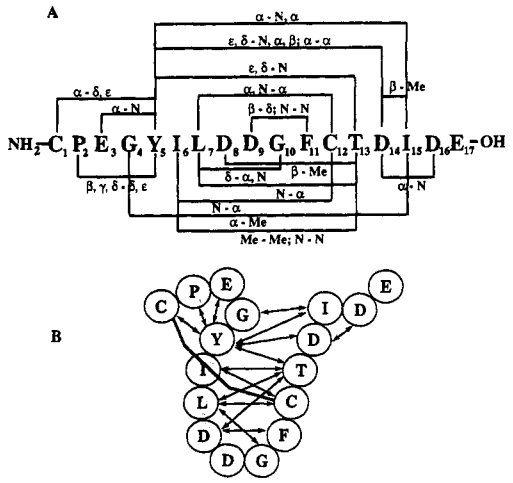


FIGURE 2: (A) A summary of the long-range transferred NOEs observed for the cyclic thrombomodulin peptide. The numbering starts with the first Cys residue of the synthetic peptide which includes the following thrombomodulin residues (Zushi et al., 1989; Nagashima et al., 1993): $C_1$(Cys409), $P_2$(Pro410), $E_3$(Glu411), $G_4$(Gly412), $Y_5$(Tyr413), $I_6$(Ile414), $L_7$(Leu415), $D_8$(Asp416), $D_9$(Asp417), $G_{10}$(Gly418), $F_{11}$(Phe419), $C_{12}$(Cys421), $T_{13}$(Thr422), $D_{14}$(Asp423), $I_{15}$(Ile424), $D_{16}$(Asp425), and $E_{17}$(Glu426). NOE connectivities were identified from the transferred NOE spectra of the cyclic peptide at both pHs of 5.5 and 7.2 with NOE mixing times of 100, 150, 200, and 300 ms. (B) A schematic folding model of the cyclic peptide incorporating the transferred NOEs (shown by arrows) from panel A. A disulfide bond is placed between Cys1 and Cys12 since the peptide is cyclic in the free state, and in the bound state, the disulfide bond is confirmed by transferred NOEs between the αCH proton of Cys1 and the $\beta CH_2$ protons of Cys12 (Figure 3) along with many other long-range NOEs indicated in panel A.

Apart from an indication of thrombin binding, the observation of transferred NOEs also suggests that the range of conformations accessible to the bound cyclic peptide is restricted by the geometry of the binding site on thrombin. The thrombin−peptide complex exhibited unique long-range transferred NOE interactions between the $\delta CH_2$ ring protons of Tyr5 and the αCH protons of Cys1 and/or Thr13, between the αCH proton of Leu7 and the NH proton of Thr13, and between the $\delta CH_2$ ring protons of Tyr5 and the αCH proton of Asp14 (Figure 1A). These NOEs are accompanied by other weaker cross peaks connecting the NH proton of Ile6 (and/or Asp9) and the αCH protons of Cys12 and Asp14, connecting the NH proton of Leu7 and the αCH proton of Cys12, and connecting the αCH proton of Asp14 and the NH proton of Asp16. This is in sharp contrast with the linear peptide where all the observed NOEs in the presence of thrombin are intraresidue and sequential in nature (Figure 1B). The lack of long-range NOEs for the linear peptide clearly indicates that disulfide bond formation is essential for maintaining a correct conformation for the bound peptide. The residual NOEs observed for the linear peptide may be a result of nonspecific interactions of the linear peptide with thrombin.

Figure 2A shows a summary of most of the significant transferred NOE conectivities observed for the cyclic peptide. The NOEs between the αCH proton of Glu3 and the NH proton of Tyr5 and between the αCH proton of Tyr5 and the NH proton of Ile15 were observed only at pH 7.2 due the slight shift of the NH resonance of Tyr5. However, the Glu3−Tyr5 connectivity plus a strong sequential αCH−NH NOE between Glu3 and Gly4 (Figure 1A) and a strong NH−NH NOE between residues Gly4 and Tyr5 (Figure S2A) clearly identifies a type II β-turn involving residues Pro2, Glu3, Gly4, and Tyr5. The side-chain and NH−NH NOEs between residues Asp9 and Phe11 suggest a second chain reversal involving residues Asp8, Asp9, Gly10, and Phe11. A third turn-like structure may be present at residues Ile15 and Asp16 due to the presence of an NOE between the αCH proton of Asp14 and the NH proton of Asp16 (Figure 1A).

Figure 2B is a schematic representation of the chain folding incorporating the chain reversals identified above based on the transferred NOE connectivity patterns. Other transferred NOEs suggest that the turn structures may be folded by a long strand of antiparallel β-sheet. The telltale NOEs for the β-sheet structure include the strong cross peaks between the αCH protons of residues Leu7 and Cys12 and residues Tyr5 and Asp14 (Figure 3) and the NH−NH NOEs between residues Ile6 and Thr13 (Figure S2A) along with the long-range αCH−NH NOEs indicated in Figures 1A and 2A. The presence of the β-sheet is further supported by the cross-strand side-chain interactions, in particular those between the $\beta CH_2$ protons of Asp8 and the methyl protons of Thr13 (Figure 2A). The aromatic protons of Tyr5 also interact with the main-chain and the side-chain protons of several residues, Pro2, Thr13, Asp14, and Ile15, distant in sequence. In addition, there are NOEs between the ring protons of Tyr5 and the overlapped methyl resonances of Ile6, Leu7, and/or Ile15. These NOEs suggest that the Tyr5 side chain is surrounded by both polar and nonpolar residues in the bound peptide. Interestingly, the $\delta CH_2$ proton resonance of Tyr5 is also very sensitive to the cyclization state of the peptide in that this resonance is shifted downfield upon the reduction and alkylation of the disulfide bond between Cys1 and Cys12 (Figures 1 and S2). Furthermore, this same resonance showed a dramatic broadening in the thrombin−peptide complex (Figure S1B), indicating that the aromatic ring of Tyr5 may be involved in binding to thrombin as well.

In the thrombin-bound state, the Cys1−Pro2 peptide bond assumes a trans conformation as is evident from the very intense transferred NOE cross peaks between the αCH proton of Cys1 and the $\delta CH_2$ protons of the Pro2 (Figure 3). There is also an NOE between the αCH proton of Cys1 and the downfield of the two $\beta CH_2$ protons of Cys12, which is slightly separated from the closeby $\beta CH_2$ protons of Cys1, $\delta CH_2$ protons of Pro2, $\beta CH_2$ protons of Tyr5, and $\beta CH_2$ protons of Phe11 (Figure 3 and Table S1). This Cys1−Cys12 NOE along with many other NOEs involving residues Cys1, Pro2, Tyr5, Thr13, Asp14, and Ile15 (Figure 2) suggests that the disulfide pairing is intact for the thrombomodulin peptide in complex with thrombin. This is in agreement with the observation that the disulfide-reduced form of the same peptide failed to bind thrombin and to produce transferred NOEs (Figures 1B and S2B).

*Thrombin-Bound Structure of the Cyclic Thrombomodulin Peptide.* Transferred NOEs have been used to determine the bound structures of peptide inhibitors targeted for both the active site and the fibrinogen recognition exosite of thrombin (Ni et al., 1989, 1990, 1992, 1994; Ni & Scheraga, 1994). The quantitation of transferred NOEs required for structure determination can be complicated if the free peptide shows some kind of NOE cross peaks, especially for sequential and intraresidue interactions (Ni, 1994). In the current study, the analysis of transferred NOEs was simpli-

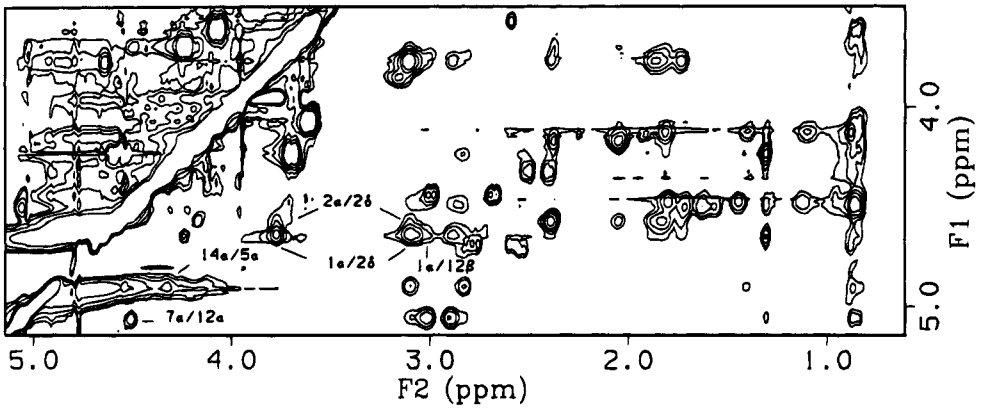


FIGURE 3: The αCH−aliphatic region of the transferred NOE spectra of the cyclic peptide. The experimental conditions are as indicated in Figure 1. A strong NOE between the αCH protons of Tyr5 and Asp14 (14α/5α) sits on a baseline ridge, but it is visible in other NOESY spectra with different NOE mixing times (not shown) but with slightly different baseline properties. The αCH proton of Phe11 overlaps with that of the water solvent peak (Table S1) and was thus attenuated by solvent presaturation (see also Figure 1A).

fied considerably since there are many long-range and side-chain NOE interactions (Figure 2A) which were not seen for the free peptide. These unique transferred NOEs may be interpreted by the isolated spin-pair/initial-slope approximation to yield distances between bound ligand protons (Ni, 1994).

An initial set of distances was derived from only the long-range or side-chain NOE interactions (Figure 2A) roughly classified as strong, medium, and weak intensities with the following assigned distance bounds: strong NOEs, 1.8−2.6 Å; medium NOEs, 1.8−3.3 Å; weak NOEs, 1.8−4.2 Å. These distance assignments are consistent with the fact that there are only weak NOEs between the αCH proton and the $\delta CH_2$ protons of Pro2 (expected distances 3.85−4.06 Å) even though these NOEs may include some spin-diffusion contributions through other short distances within the proline ring. The NOE distances were subsequently refined by a comparison of experimental and theoretical transferred NOEs expected from a model structure of the bound peptide, as used in other studies (Ni et al., 1992a, 1994; Ni, 1994; Ni & Scheraga, 1994). The calculated spectra were inspected to identify cross peaks that were absent in the experimental data or cross peaks that had grossly different transferred NOE intensities. Protons with weaker (or missing) experimental NOEs were pushed apart by incrementing the distance constraints during the subsequent refinement cycle. Otherwise, they were pulled together by decrementing the distances. Lower bounds between some protons were also identified for proton pairs which participate in many NOE interactions and between protons which tended to be forced too close by NOE constraints after energy minimization. This strategy helped to surmount the issue of local mobility, which can also result in reduced NOE intensities or even the absence of transferred NOEs.

After a few cycles of distance refinement, the underlying model structures of the thrombomodulin peptide were inspected for candidates of hydrogen bond formation. These hydrogen bonds were already apparent during initial stages of structure calculations incorporating all the long-range NOEs (Figure 2A) but were added explicitly only during later stages of structure calculations. The identified hydrogen bonds included the one between Pro2 CO and Tyr5 HN, expected within the type II β-turn involving residues Pro2−Tyr5, and the ones within the long antiparallel β-sheet for residuesGly4−Asp8 and residues Phe11−Ile15 (Figure 2B). For every potential hydrogen bond, a loose constraint of 1.7−2.2 Å was imposed for the distance between the amide proton and the carbonyl oxygen atoms. Incorporation of these distance constraints resulted in better agreement between the experimental and the computed transferred NOEs, confirming the validity of the hydrogen bond assignments.

Figure 4 shows a superposition of the converged structures of the cyclic thrombomodulin peptide obtained by distance geometry and followed by restrained Monte Carlo calculations incorporating the refined NOE and the hydrogen bond constraints listed in Table S2. Extensive distance geometry and Monte Carlo sampling showed that transferred NOE distances constrained residues Cys1−Ile15 of the cyclic peptide to almost unique conformations with RMS deviations of less than 0.6 Å for all the backbone atoms among the various converged structures. Residues Asp16 and Glu17, on the other hand, are not as well defined due to the lack of NOE interactions with the rest of the molecule. There are also variations among some of the side-chain orientations due to the lack of stereospecifically assigned distances (Table S2). The converged structures reproduced the transferred NOE data as shown by a comparison of experimental and theoretical transferred NOESY spectra (Figure S3). As expected, there is a long strand of antiparallel β-sheet, but very interestingly, the β-strands showed a right-handed twist in all the structures (Figure 4). One face of the structure is primarily hydrophobic with a distinct nonpolar cluster comprised of the side chains of Cys1, Pro2, Tyr5, Leu7, and Cys12. With the exception of Ile6, the other face is largely hydrophilic with residues Glu3, Asp8, Asp9, Thr13, Asp16, and Glu17. The charged side chains of residues Glu3, Asp8, Asp9, and Asp16 and Glu17 decorate both ends of the three-stranded sheet structure.

NMR, X-ray crystallographic, and binding studies have firmly established that the thrombin−ligand complexes are stabilized by hydrophobic interactions between the thrombin exosite and the nonpolar residues provided by the binding ligand (Ni et al., 1990, 1992b; Skrzypczak-Jankun et al., 1991; Stone & Hofsteenge, 1991; Yue et al., 1992; Qiu et al., 1993; Mathews et al., 1994). On the basis of the bound structure of the thrombomodulin peptide, the side chains of residues Cys1, Pro2, Tyr5, Phe11, and Cys12 may all make hydrophobic contacts with the thrombin exosite. Very

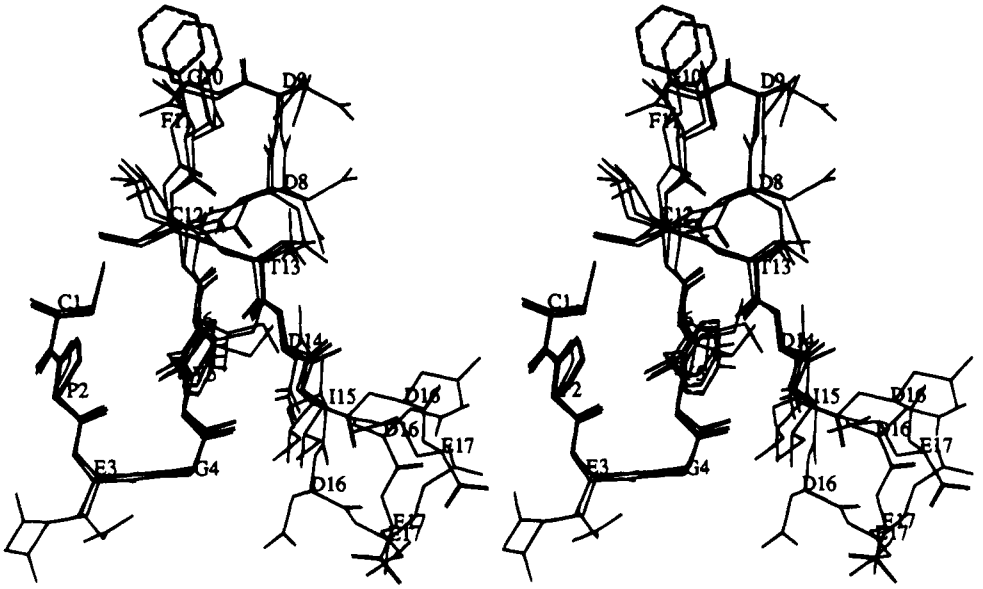


FIGURE 4: A cluster of three low-energy structures of the cyclic peptide based upon restrained Monte Carlo calculations. RMS deviations among the backbone atoms are less than 0.6 Å for residues Cys1−Ile15. Structures of residues Asp16 and Glu17 are not well defined due to the absence of NOEs to the rest of the molecule (Figure 2A). All the calculated structures reproduce the experimental transferred NOE spectra as illustrated in Figure S3. The atomic coordinates of the structures shown have been deposited with the Brookhaven Protein Data Bank.

interestingly, the Cα−Cα distance of ~9.5 between residues Tyr5 and Phe11 in the thrombomodulin peptide (Figure 4) closely approaches the distance of ~9 Å between the Cα atoms of residues Phe56 and Ile59 in the complex of a hirudin peptide, Asp-Phe56-Glu-Glu-Ile59-Pro-Glu-Glu-Tyr-Leu-Gln, with thrombin (Skrzypczak-Jankun et al., 1991). On the other hand, residues Phe56 and Tyr63 in the hirudin peptide are around 15 Å away as measured by their Cα distances. Recent studies with other thrombin−peptide complexes all point to the importance of the sequence motif Phe/Tyr-Glu-Glu/Pro-Ile/Phe in binding to the thrombin exosite (Qiu et al., 1993; Mathews et al., 1994). In these thrombin−peptide complexes, residues Phe/Tyr-Glu-Glu/Pro-Ile/Phe must assume extended backbone conformations in order to make hydrophobic contacts with the two distinctive nonpolar patches within the thrombin exosite (Yue et al., 1992). The thrombomodulin peptide, on the other hand, provides a potential thrombin-binding surface between residues Tyr5 and Phe11 which are brought close by a chain reversal within the two-stranded and antiparallel β-sheet. Therefore, a common surface topology may be organized by different secondary structure elements with entirely different peptide sequences. The thrombin-bound structure of the thrombomodulin peptide may serve as a basis to understand the regulatory functions of thrombomodulin and as a guide for the design of specific inhibitors for thrombin. Work is in progress to further define the binding site on thrombin and to refine the structure of the peptide−thrombin complex in solution.

*Comparison with the Solution Structures of EGF Molecules.* Most EGF or EGF-like domains have a 10-residue third disulfide loop, except for the fifth EGF-like repeat in thrombomodulin and the carboxy-terminal EGF-like domain of factor X, both of which have a longer loop of 13 and 14 residues, respectively (Campbell & Bork, 1993; Padmanabhan et al., 1993). On the other hand, the analog cyclic peptide has an intermediate loop length of 12 residues. Figure 5 shows a comparison of the thrombin-bound structure of the cyclic thrombomodulin peptide with the refined structures of the corresponding residues in the murine EGF molecules (Kohda & Inagaki, 1992; Montelione et al., 1992). The first seven residues in the three disulfide loops have almost identical backbone and side-chain conformations with the first four residues involved in a type II β-turn. The invariant aromatic (or Tyr) residue at the fourth position starts the first strand of the antiparallel β-sheet while its aromatic side chain packs against the disulfide bond. Interestingly, the variation in chain length hardly changes the spatial positions of the disulfide-bonded cysteine residues. In contrast, the shorter 10-residue disulfide loops have a significantly distorted complementary β-strand (Campbell & Bork, 1993) to accommodate the proper formation of the disulfide bond. The 12-residue cyclic thrombomodulin peptide assumes an optimized antiparallel β-sheet in the thrombin-bound state, allowing residues Tyr5−Asp8 and Phe11−Ile15 to participate in structure formation and in intermolecular interactions with thrombin. A similar sheet structure involving eight residues was also found for the larger 14-residue third loop in the carboxyl-terminal EGF-like domain of factor X as part of an intimate molecular complex with the factor X protease domain (Padmanabhan et al., 1993). The shortening of two residues in the synthetic thrombomodulin peptide apparently only affected the size of the turn between the two β-strands. These results provide further support to the view that the core structure of the disulfide loop may be very much conserved across all EGF-like domains in proteins (Campbell & Bork, 1993).

At this point, it is interesting to point out that the cyclic thrombomodulin peptide has a structure similar to the corresponding subdomains in EGF molecules only when it is in complex with thrombin. In the free state, the same cyclic peptide has no predominant conformational preferences as there are no NOEs of significant intensity for the free peptide (spectra not shown) and the chemical shift values of the thrombomodulin peptide approach those of a random structure (Table S1). The lack of structure for the free

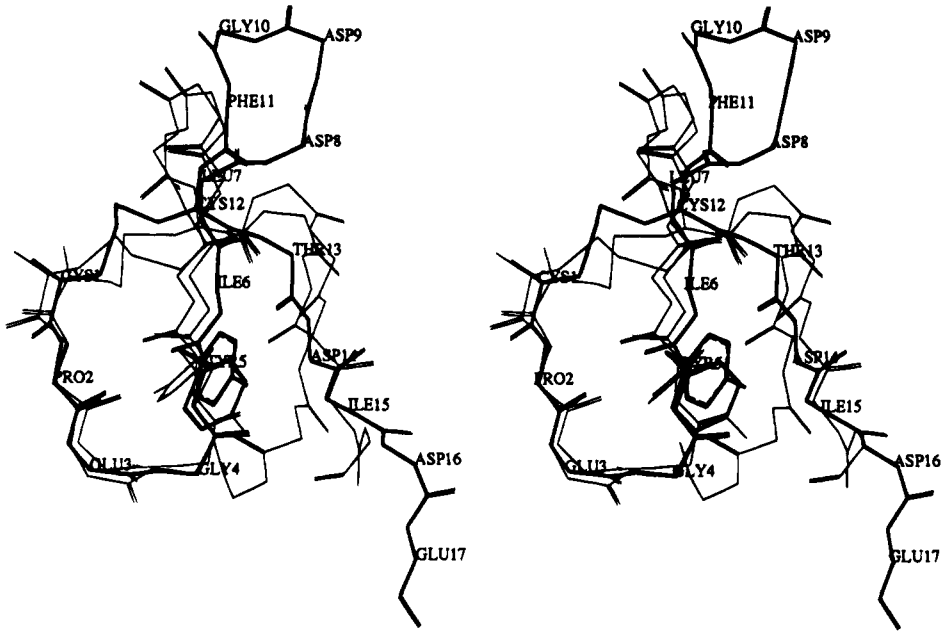


FIGURE 5: A superposition of a refined thrombin-bound structure of the cyclic thrombomodulin peptide (thick line) with the corresoponding loop structures (thin lines) of murine EGF molecules (Kohda & Inagaki, 1992; Montelione et al., 1992) available in the Protein Data Bank as 3EGF and 1EPI, respectively. Superposition was carried out with the Cα atoms of residues Cys1–Leu7, Cys12, and Asp14. The structurally equivalent residues in murine EGF are Cys33–Gly39, Cys42, and Arg45, respectively.

peptide could be due to the lack of long-range interactions when it is part of the EGF domain. In fact, there is close contact between the third disulfide loop and another loop immediately after the second cysteine residue in intact EGF domains (Campbell & Bork, 1993). It is possible that the thrombomodulin peptide will assume an folded structure when attached to the rest of the fifth EGF-like domain. However, a recent study showed that the thrombomodulin peptide has a nearly identical thrombin-binding activity whether it is isolated or part of the entire fifth EGF-like domain (Hunter et al., 1994). Further insight into this intriguing observation may come from the determination of the structures of the full-length EGF domains in solution and in complex with thrombin.

## ACKNOWLEDGMENT

Thanks are due to Bruce Fulton for help with spectrometer hardware and to Hervé Hogue for support in computer graphics. We are grateful to Dr. John Fenton, II, for supplying the human γ-thrombin used in control binding studies.

## SUPPLEMENTARY MATERIAL AVAILABLE

Table S1 listing the proton resonance assignments for the cyclic thrombomodulin peptide, Table S2 containing the refined transferred NOE distance constraints used in structure calculations, Figure S1 displaying the aromatic region of the one-dimensional NMR spectra of the free cyclic peptide (S1A) and those in the presence of α-thrombin (S1B) and γ-thrombin (S1C), Figure S2 showing the transferred NOEs involving the NH and the aromatic protons of the cyclic peptide and the linear peptide in complex with thrombin, and Figure S3 showing the computed transferred NOE spectra corresponding to the experimental spectra shown in Figures 1A, S2A, and 3 (11 pages). Ordering information is given on any current masthead page.

## REFERENCES

Albrand, J. P., Bridsall, B., Feeney, J., Roberts, G. C. K., & Burgen, A. S. V. (1979) *Int. J. Biol. Macromol. 1*, 37–41.

Campbell, I. D., & Bork, P. (1993) *Curr. Opin. Struct. Biol. 3*, 385–392.

Daggett, V., Kollman, P. A., & Kuntz, I. D (1991) *Biopolymers 31*, 285–304.

Engel, J (1989) *FEBS Lett. 251*, 1–7.

Esmon, C. T. (1989) *J. Biol. Chem. 264*, 4743–4746.

Hayashi, T., Zushi, M., Yamamoto, S., & Suzuki, K. (1990) *J. Biol. Chem. 265*, 20156.

Hunter, M. J., Plesniak, L. A., TenEyck, L. F., & Komives, E. A. (1994) *Protein Sci. 3* (Suppl. 1), 122.

Kohda, D., & Inagaki, F. (1992) *Biochemistry 31*, 11928–11939.

Kurosawa, S., Stearns, D. J., Jackson, K. W., & Esmon, C. T. (1988) *J. Biol. Chem. 263*, 5993–5996.

Lougheed, J. L., Bowman, C. A., Meininger, D. P., & Komives, E. A. (1994) *Biochemistry* (submitted for publication).

Mathews, I. I., Padmanabahn, K. P., Ganesh, V., Tulinsky, A., Ishii, M., Chen, J., Turck, C. W., Coughlin, S. R., & Fenton, J. W., II (1994) *Biochemistry 33*, 3266–3279.

Montelione, G. T., Wüthrich, K., Burgess, A. W., Nice, E. C., Wagner, G., Gibson, K. D., & Scheraga, H. A. (1992) *Biochemistry 31*, 236–249.

Nagashima, M., Lundh, E., Leonard, J. C., Morser, J., & Parkinson, J. F. (1993) *J. Biol. Chem. 268*, 2888–2892.

Ni, F. (1992a) *J. Magn. Reson. 96*, 651–656.

Ni, F. (1922b) *J. Magn. Reson. 99*, 391–397.

Ni, F. (1994) *Prog. NMR Spectrosc. 26*, 517–606.

Ni, F., & Scheraga, H. A. (1994) *Acc. Chem. Res. 27*, 257–264.

Ni, F., & Zhu, Y. (1994) *J. Magn. Reson. B102*, 180–184.

Ni, F., Meinwald, Y. C., Vásquez, M., & Scheraga, H. A. (1989) *Biochemistry 28*, 3094–3105.

Ni, F., Konishi, Y., & Scheraga, H. A. (1990) *Biochemistry 29*, 4479–4489.

Ni, F., Ripoll, R. D., Martin, P. D., & Edwards, B. F. P. (1992a) *Biochemistry 31*, 11551.

Ni, F., Ripoll, D. R., & Purisima, E. O. (1992b) *Biochemistry 31*, 2545−2554.

Ni, F., Ning, Q., Jackson, C. M., & Fenton, J. W., II (1993) *J. Biol. Chem. 268*, 16899.

Ning, Q., Ripoll, D. R., Szewczuk, Z., Konishi, Y., & Ni, F. (1994) *Biopolymers 34*, 1125−1137.

Padmanabhan, K., Padmanabhan, K. P., Tulinsky, A., Park, C. H., Bode, W., Huber, R., Blankenship, D. T., Cardin, A. D., & Kisiel, W. (1993) *J. Mol. Biol. 232*, 947−966.

Qiu, X., Yin, M., Padmanabahn, K. P., Krstenansky, J. L., & Tulinsky, A. (1993) *J. Biol. Chem. 268*, 20318−20326.

Ripoll, D. R., & Ni, F. (1992) *Biopolymers 32*, 359−356.

Skrzypczak-Jankun, E., Carperos, V. E., Ravichandran, K. G., Tulinsky, A., Westbrook, M., & Maraganore, J. M. (1991) *J. Mol. Biol. 221*, 1379−1393.

Stone, S. R., & Hofsteenge, J. (1991) *Biochemistry 30*, 3950−3955.

Tsiang, M., Lentz, S. R., Dittman, W. A., Wen, D. Z., Scarpati, E. M., & Sadler, J. E. (1990) *Biochemistry 29*, 10602−10612.

Tsiang, M., Lentz, S. R., & Sadler, J. E. (1992) *J. Biol. Chem. 267*, 6164−6170.

Vásquez, M., & Scheraga, H. A. (1988) *J. Biomol. Struct. Dyn. 5*, 757−784.

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

Ye, J., Liu, L. W., Esmon, C. T., & Johnson, A. E. (1992) *J. Biol. Chem. 267*, 11023−11028.

Yue, S. Y., DiMaio, J., Szewczuk, Z., Purisima, E. O., Ni, F., & Konishi, Y. (1992) *Protein Eng. 5*, 77−85.

Zushi, M., Gomi, K., Yamamoto, S., Maruyama, I., Hayashi, T., & Suzuki, K. (1989) *J. Biol. Chem. 264*, 10351−10353.