

Crystallographic Evidence That the F2 Kringle Catalytic Domain Linker of Prothrombin Does Not Cover the Fibrinogen Recognition Exosite*

(Received for publication, September 8, 1995)

Andreas van de Loch†, Milton T. Stubbs, Margit Bauer, and Wolfram Bode

From the Department of Structural Research, Max Planck Institute for Biochemistry, D-82152 Martinsried, Germany

The 2.6-Å x-ray crystal structure of bovine α -thrombin in complex with rhodniin, a protein inhibitor isolated from the bug *Rhodnius prolixus*, has been solved and refined. The structure has enabled us to trace the N-terminal part of the 49-residue A-chain of bovine α -thrombin for the first time, which is fixed in a U-shaped loop on the molecular surface opposite the active site canyon. Model building shows that the 25 amino acid residues that link the A-chain and F2 kringle cannot run through the fibrinogen recognition exosite. This demonstrates that this fibrinogen recognition exosite is available in prothrombin and meizothrombin.

Thrombin plays a central role in hemostasis and thrombosis, performing both coagulatory and regulatory functions (see Ref. 1). Prothrombin circulates in the plasma as a 579-residue (in man; 582 residues in bovine) 4-domain zymogen, consisting of an N-terminal Gla (γ -carboxyglutamic acid) domain, two tandem kringle domains, and the catalytic domain (see Fig. 1). Activation of prothrombin via two cleavages at Arg-PT271 (human; Arg-PT274 in bovine) and Arg-15 yields the two-chain α -thrombin (for thrombin, A-chain residues from Cys-1 to Arg-15 and the B-chain are identified via chymotrypsinogen numbers deduced from topological equivalence (2, 3); residues preceding Cys-1 are numbered according to human prothrombin and designated by the prefix PT (see Table I)). Depending on the reaction conditions, either bond may be cleaved first (4, 5). Under physiological conditions (*i.e.* in the presence of Ca^{2+} and negatively charged phospholipids), activation is effected largely via the prothrombinase complex (factors Xa and Va). In this case, the first cleavage occurs at Arg-15, resulting in the catalytically active membrane-bound intermediate meizothrombin. Subsequent cleavage at Arg-PT271 releases the mature two-chain procoagulant α -thrombin into the blood vessel. In the absence of factor Va, factor Xa performs these cleavages in reverse order at a much slower rate (4). Due to an additional autocatalytic cleavage at Arg-PT285–Thr-PT286, the human α -thrombin A-chain is further truncated to 36 residues.

Crystallographic studies of α -thrombin reveal a compact molecule, with the B-chain and the central portion of the A-chain forming a single globular domain (1–3). The α -thrombin molecule is characterized by a deep narrow canyon-like active site cleft, together with two positively charged surface patches, representing the “fibrinogen recognition exosite” and the “heparin binding site” (3). These two sites are of great functional significance in thrombin’s interactions with macromolecular

substrates and inhibitors (1). Recent crystal structures of α -thrombin in complex with F2 kringle (6) reveal that this kringle domain binds to the heparin binding site, largely via specific electrostatic interactions. The structure of prethrombin 2 has also recently been solved (7), exhibiting an overall organization identical to that of α -thrombin; the C-terminal part of the A-chain shows some displacement due to the intact peptide bond between Arg-15 and Ile-16. As in trypsinogen (8, 9) and chymotrypsinogen (10), the low proteolytic activity of zymogen species can be attributed to the malformation of the oxyanion hole, completed on salt bridge formation between Asp-194 and the new N terminus of Ile-16. The above structures together with that of prothrombin fragment F1 (Gla and F1 kringle domains) (11, 12) have given us a detailed understanding of the structural biology of α -thrombin (1, 3) and a model for the organization of prothrombin (13).

In contrast, the organization of the N-terminal portion of the A-chain has not been determined, which forms the covalent link to the C terminus of the F2 kringle in prothrombin, prethrombin 1, and meizothrombin. Although meizothrombin displays comparable amidolytic activity to α -thrombin against low molecular weight substrates, its activity toward fibrinogen, factor V, and platelets is reduced (14), and prothrombin and prethrombin 1 do not bind to fibrinogen (15). One possible explanation for this was thought to be the partial obstruction of the fibrinogen recognition exosite by the linker peptide. On the other hand, Liu *et al.* (16) have reported that α -thrombin, prethrombin 2, and meizothrombin bind hirugen, suggesting that conformational changes rather than steric hindrance interfere with macromolecular binding at the fibrinogen recognition exosite. Furthermore, Ni *et al.* (17) have shown that the fibrinogen recognition exosite is at least partially accessible in prothrombin. In the present paper, we describe the experimentally determined structure of almost the complete A-chain of bovine α -thrombin and the course of the modeled 25-amino acid residue linker. The N-terminal residues of the A-chain are ordered in such a manner that excludes occupancy of the fibrinogen recognition exosite by the linker.

MATERIALS AND METHODS

Bovine thrombin was cocrystallized with the inhibitor rhodniin isolated from the bug *Rhodnius prolixus* (18). Orthorhombic crystals (space group P2₁2₁2₁, two complexes per asymmetric unit) and monoclinic crystals (space group C2, two complexes per asymmetric unit) were grown from 10% polyethylene glycol 4000, 300 mM phosphate buffer, pH 6.

The structure was solved by Patterson search techniques with bovine thrombin as obtained in complex with N^α-(2-naphthylsulfonyl)glycyl-DL-p-aminophenylalanyl-piperidine (19) as a search model. For the monoclinic data with reflections up to 2.6 Å the model was refined to a *R* value of 0.189 using X-PLOR (20) with the parameters of Engh and

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed.

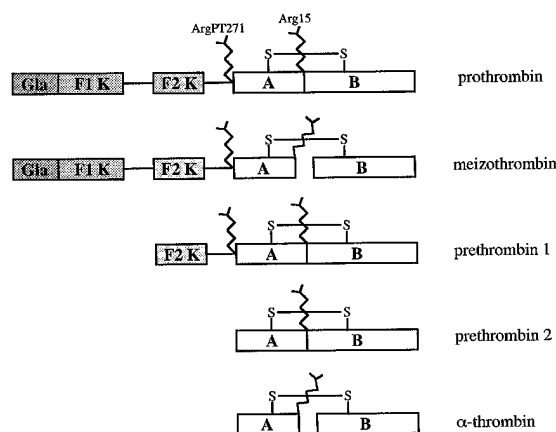


FIG. 1. Schematic representation of prothrombin and some of its activation products. The factor Xa cleavage sites in prothrombin at Arg-PT271 and Arg-15 are indicated by the arginine side chains. In human α -thrombin there is an additional thrombin cleavage site within the A-chain at Arg-PT284. Gla denotes the γ -carboxyglutamic acid domain, which is followed by the F1 and F2 kringle domains.

TABLE I
Alignment of the residues between the last cysteine in F2 kringle and Cys-1 in the thrombin A-chain for the six prothrombin sequences known

The sequences were aligned using the AMPS program package (29, 30) with a gap penalty of 8, a constant of 8, and Dayhoff's mutation matrix as a scoring scheme. The aligned sequences were formatted with the program ALSCRIPT (31). Human prothrombin (PT) and chymotrypsinogen (CHYMO) numbering is given (1, 3). The arginine residues at the factor Xa cleavage site (PT 271) and the site of autocatalytic cleavage in human thrombin (PT 284) are displayed inverse, with the residue numbers of the former one given in parentheses for the individual sequences. Strictly conserved residues in the region PT 280 to PT 287 are marked by boxes, and predominant residues are marked by shading. The sequences were taken from Banfield and MacGillivray (25).

PT	248	250		255		260		265		271	
bovine	C	C	P	V	D	G	D	R	L	G	E
human	C	C	P	V	D	G	D	R	L	G	E
rat	C	C	P	V	D	G	D	R	L	G	E
mouse	C	C	P	V	D	G	D	R	L	G	E
chicken	C	C	P	V	D	G	D	R	L	G	E
hagfish	C	C	P	V	D	G	D	R	L	G	E
PT	272	275		280		285		290		293	
CHYMO	IU	IT	IS	IR	IQ	IP	IO	IN			
bovine	T	S	E	D	H	F	Q	P			
human	T	S	E	D	H	F	Q	P			
rat	T	S	E	D	H	F	Q	P			
mouse	T	S	E	D	H	F	Q	P			
chicken	T	S	E	D	H	F	Q	P			
hagfish	T	S	E	D	H	F	Q	P			

Huber (21). The final model showed r.m.s.¹ deviations from ideal values for bond lengths and angles of 0.01 Å and 1.66°, respectively. The orthorhombic crystals diffracted up to 3.1 Å, and the model was refined to a *R* value of 0.175. r.m.s. deviations of the final model are 0.01 Å and 1.62° for bond lengths and angles, respectively. Full details of the refinement and the final model are given in Ref. 22. The coordinates of both crystal forms have been deposited with the Brookhaven Protein Data Bank.

The 25-amino acid residue linker was modeled along the surface of the catalytic module of human thrombin to connect the thrombin A-chain with the F2 kringle using the programs O (23) and MAIN (24) (coordinates of the complex of human thrombin and F2 kringle were kindly provided by Prof. A. Tulinsky). Special attention was paid to burial of hydrophobic residues, charge compensation, and the additional requirements of complementarity to factor Xa. Bond and angle energies of the modeled linker were minimized using X-PLOR.

RESULTS AND DISCUSSION

In three of the four independent bovine thrombin molecules examined here, the complete B-chain and the central and C-

terminal part of the A-chain (Cys-1 to Arg-15) can be traced in continuous electron density; for one molecule no electron density is observed for a short segment (Ala-149–Asp-149) in the autolysis loop. The average r.m.s. deviation for 301 C α -atoms (excluding residues Asp-147–Asp-149) is 0.36 Å, with a conformation virtually identical to those observed previously for thrombin. In addition, we observe for the first time well defined electron density for the 16 residues preceding Cys-1, Phe-PT277 to Asp-PT292 (Fig. 2). These residues exhibit a virtually identical conformation in all four copies, characterized by an average α -carbon r.m.s. deviation of 0.69 Å. Most of these side chains are well defined and exhibit little variability; only those of Lys-PT284 and Gln-PT278 differ considerably between the four thrombin molecules. The five most N-terminal residues of the A-chain can be traced by continuous electron density only for one molecule in the orthorhombic crystals. In the other three thrombin molecules, discontinuous density is found, possibly reflecting slightly different conformations.

The N-terminal fragment (Thr-PT272 to Asp-PT292) of the A-chain forms a large U-shaped loop with multiple turn structure (see Fig. 3). It is situated in an elongated depression on the "back" surface of the catalytic module. These 21 residues cover approximately 700 Å² of a surface patch extending between the C-terminal helix of the B-chain, Arg-206, Arg-50, and Asp-125, enclosing also the basic residues Lys-107, Lys-129B, Lys-235, and Arg-244; this is to be compared with 1700 Å² covered by the whole A-chain.

This 21-residue fragment is stabilized through electrostatic and hydrophobic interactions to the underlying B-chain module. Seven carbonyl groups and the negatively charged side chains of Asp-PT292 and Glu-PT290 are directed toward a positively charged patch of the B-chain surface. Only one carbonyl oxygen (Ala-PT288) is engaged in an inter-main chain hydrogen bond with an amide nitrogen (Asp-49), whereas five of the carbonyl groups are hydrogen bonded to side chain groups. Asp-PT292 is the only residue to be involved in a direct ionic contact with the catalytic module (Lys-9).

The majority of hydrophobic contacts is mediated through the 4 phenylalanine residues Phe-PT286, Phe-PT281, Phe-PT280, and Phe-PT277 (see Fig. 3). Their bulky aromatic side chains are almost completely buried in the interface between the A- and the B-chain, with residues Phe-PT286 and Phe-PT281 in the center of a large loop pointing toward a hydrophobic surface patch formed by Ile-47, Trp-51, Leu-53, Leu-105, Ile-242, Ile-238, and Leu-123. Three of these phenylalanine residues (Phe-PT286, Phe-PT281, and Phe-PT280) are clustered in an aromatic triad, with Phe-PT281 making face-to-face contacts with Phe-PT286 and face-to-edge contacts with Phe-PT280. Furthermore, Lys-235 points toward the electron-rich face of Phe-PT286 to form a 3.5-Å hydrogen bond (see Fig. 2). The benzene ring of the fourth phenylalanine, Phe-PT277, is oriented parallel to the guanidino group of Arg-206. These four phenylalanyl residues certainly represent a central element for the stabilization of this N-terminal A-chain fragment. This is further underscored by the fact that amino acids PT280 to PT287 are highly conserved in all six prothrombin sequences known (see Table I); residues Gly-PT287, Phe-PT286, Thr-PT285, Phe-PT281, and Phe-PT280 (*i.e.* including three of the phenylalanines) are identical, and the three intervening amino acids, PT282–PT284, exhibit only two different types. The supporting amino acids are also highly conserved; five of the seven hydrophobic residues are identical in all 11 thrombin sequences known (25), and in the other positions only hydrophobic residues are found.

In addition to these A-B interchain contacts, the loop is also stabilized through intrachain interactions. An important con-

¹ The abbreviation used is: r.m.s., root mean square.

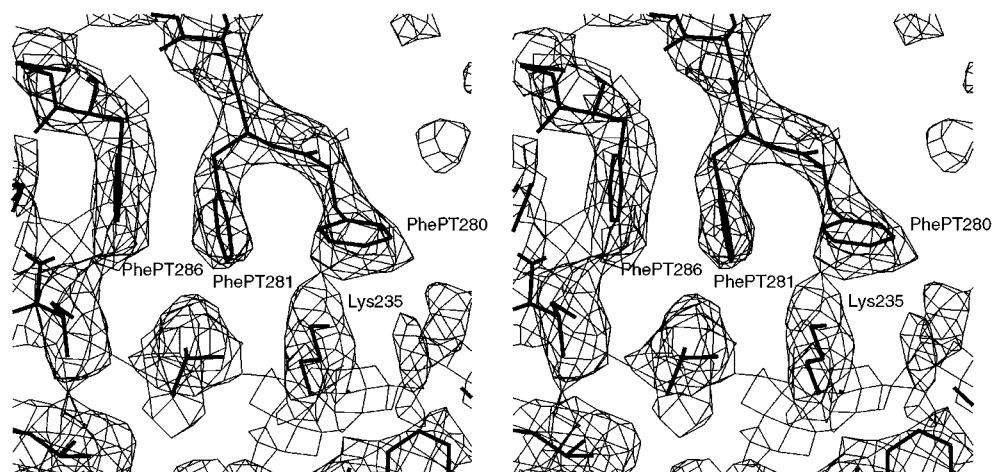


FIG. 2. **Stereo view of the electron density of the extended N-terminal fragment of the A-chain.** The phenylalanines PT281 and PT286 make a face-to-face contact, and Phe-PT280 makes a face-to-edge contact with Phe-PT281. Furthermore, the side chain of Lys-235 makes a hydrogen bond to the face of the phenyl group of Phe-PT280. Contour surface is shown at 0.8σ . This figure was produced using the program MAIN.

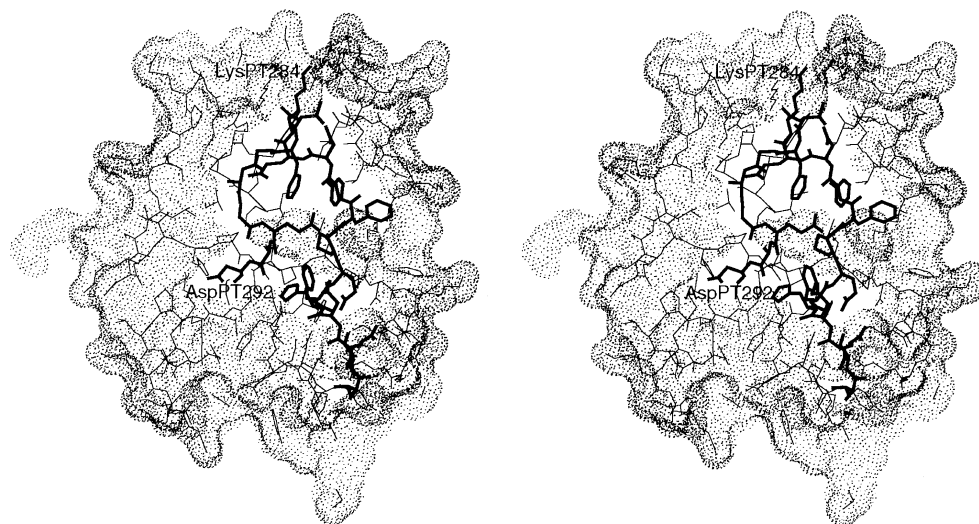


FIG. 3. **Connolly surface (calculated and displayed with MAIN) of the observed structure of bovine thrombin.** The N-terminal residues PT272 to PT292 are shown in *thick lines* and the central and C-terminal part of the A-chain and the B-chain in *thin lines*. The surface of the N-terminal residues is not shown. The view is onto the reverse side of thrombin, rotated 180° about the y axis compared with the standard orientation. The active site therefore runs from *right to left* along the back of the figure, with the fibrinogen recognition exosite at the *left*.

tribution is made by the side chain of Glu-PT290, which extends toward the loop center and connects its two peripheral ends by hydrogen bonds to main chain amide nitrogens. This loop is further stabilized by a few inter-main chain hydrogen bonds. Intermolecular crystal contacts, in contrast, do not seem to affect the loop structure to a large extent. In only two of the four structures examined a weak ionic contact between the partially disordered side chain of Lys-PT284 and a symmetry-related rhodniin molecule is observed; the side chain is disordered in the other cases.

Inspection of the electron densities for some other bovine thrombin structures in different space groups determined in our laboratory² suggest a similarly folded A-chain segment, PT277–PT292. It might be interesting to note that the homologous PT285–PT292 segment of human thrombin is found with a similar conformation in the D-Phe-Pro-Arg-chloromethylketone-thrombin complex, albeit in much weaker electron density (3). The congruence in the conformation of segment PT277–PT292 observed in the bovine thrombin structures leads us to assume that this conformation and location are

inherently preferred and also maintained in the uncleaved proforms.

In human thrombin, Lys-PT284 is replaced by an arginine residue, and its peptide bond becomes autocatalytically cleaved leading to the N-terminally truncated A-chain that starts with Thr-PT285. In the bovine thrombin crystal structures presented here the six amino acids flanking this residue on both sides form one turn of an approximate α -helix (see Fig. 3). Despite this P1-arginine and the preceding P2-proline residue (replacing the bovine thrombin glutamic acid) this segment presumably adopts a similar conformation in the intermediate elongated human (pro)thrombin forms. Residues P4–P4' would presumably require considerable rearrangements to facilitate binding into the active site of an attacking human thrombin molecule and to adopt a canonical conformation for cleavage (26). Only human and rat prothrombin exhibit an arginine at position P1, and only human prothrombin possesses a proline at position P2 that accurately matches thrombin cleavage site specificity.

In the intact form of bovine prothrombin, a 25-residue linker peptide connects the last cysteine of F2 kringle (Cys-PT249) with Thr-PT272 of the A-chain (see Table I). The corresponding

² R. Engh, personal communication.

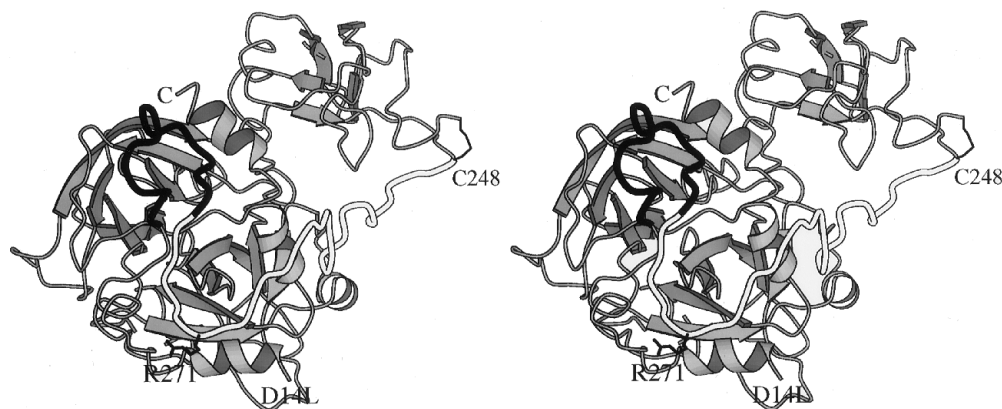


FIG. 4. Stereo view of the prethrombin 1 model in the same orientation as Fig. 3. A ribbon representation shows secondary structure elements for human thrombin and F2 kringle (6). The C α trace for residues PT249–PT292 with the observed N-terminal fragment of the A-chain (residues PT277–PT292) is shown in black, and the 28 residues of the modeled linker peptide are shown in white. This figure was drawn using the program MOLSCRIPT (28).

linkers in human, rat, mouse, and chicken species are shorter by two, seven, and eight residues, respectively. Assuming the conformation observed in our four bovine thrombin molecules from Phe-PT277 onward, we have tried to model an appropriate peptide along the surface of human thrombin linking Phe-PT277 with the C terminus of the F2 kringle (see Fig. 4). This linking peptide was positioned to cover a hydrophobic patch centered around Ile-PT268 formed by Leu-14G, Tyr-14J, Tyr-134, Pro-204, and Phe-204A. Ile-PT268 is conserved in the prothrombin sequences of human, bovine, rat, mouse, and chicken species, and even in hagfish a leucine is found at that position. Tyr-14J and Pro-204 of the supporting surface patch are also strictly conserved, and the residues found at the other three positions are mainly hydrophobic (25).

In a factor Xa-prothrombin enzyme-substrate encounter complex the position of Ile-PT268 and the overall apolar nature of this region would nicely match the extended hydrophobic patch west of the S1 pocket in factor Xa (27). Furthermore, the charged patches on the module surface around the cleavage site between linker and A-chain are complementary to an attacking factor Xa molecule. The modeled linker peptide also covers part of the heparin binding site not occupied by F2 kringle, with the five glutamate residues in the N-terminal part of the linker contacting the basic residues Arg-165, Lys-169, and Arg-175. The highly negatively charged linking peptide together with at least two uncompensated negatively charged side chains of F2 kringle (*i.e.* Asp-PT173 and Asp-PT207) might mediate binding to factor Xa, whose large positively charged heparin binding site would otherwise be repulsed by that of thrombin. The proposed conformation would also allow this linking part to follow the same course in rat and mouse prothrombin, where the linker is seven residues shorter. In our model for prethrombin 1, the two factor Xa cleavage sites are only approximately 15 Å apart. Thus both sites would be accessible to the same attacking factor Xa molecule in the prothrombinase complex without large rearrangements, in agreement with experimental results.

Although the linker chain could take several other adjacent courses, a course through the active site cleft and/or through the fibrinogen recognition exosite would require a linker peptide of at least 115-Å length. Such a trace would not be possible for the chains of rat, mouse, chicken, or hagfish, and even for human and bovine prothrombin the linker chain would have to be extremely extended. Therefore, a direct influence of the linking peptide on this important fibrinogen recognition exosite can probably be excluded.

Acknowledgments—We thank Drs. H. Brandstetter, R. A. Engh, and D. Lamba for stimulating and critical discussions, D. Grosse for technical assistance, and Prof. A. Tulinsky for providing the coordinates of the complex of human thrombin and F2 kringle.

REFERENCES

- Stubbs, M. T., and Bode, W. (1993) *Thromb. Res.* **69**, 1–58
- Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. R., and Hofsteenge, J. (1989) *EMBO J.* **8**, 3467–3475
- Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci.* **1**, 426–471
- Krishnaswamy, S., Church, W. R., Nesheim, M. E., and Mann, K. G. (1987) *J. Biol. Chem.* **262**, 3291–3299
- Walker, R. K., and Krishnaswamy, S. (1994) *J. Biol. Chem.* **269**, 27441–27450
- Arni, R. K., Padmanabhan, K., Padmanabhan, K. P., Wu, T. P., and Tulinsky, A. (1993) *Biochemistry* **32**, 4727–4737
- Vijayalakshmi, J., Padmanabhan, K. P., Mann, K. G., and Tulinsky, A. (1995) *Protein Sci.* **3**, 2254–2271
- Huber, R., and Bode, W. (1978) *Acc. Chem. Res.* **11**, 114–122
- Bode, W. (1979) *J. Mol. Biol.* **127**, 357–374
- Wang, D., Bode, W., and Huber, R. (1985) *J. Mol. Biol.* **185**, 595–624
- Seshadri, T. P., Tulinsky, A., Skrzypczak-Jankun, E., and Park, C. H. (1991) *J. Mol. Biol.* **220**, 481–494
- Soriano-Garcia, M., Padmanabhan, K., de Vos, A. M., and Tulinsky, A. (1992) *Biochemistry* **31**, 2554–2566
- Arni, R. K., Padmanabhan, K., Padmanabhan, K. P., Wu, T. P., and Tulinsky, A. (1994) *Chem. Phys. Lipids* **67**, 59–66
- Doyle, M. F., and Mann, K. G. (1990) *J. Biol. Chem.* **265**, 10693–10701
- Kaczmarek, E., Kaminski, M., and McDonagh, J. (1987) *Biochim. Biophys. Acta* **914**, 275–282
- Liu, L.-W., Ye, J., Johnson, A. E., and Esmon, C. T. (1991) *J. Biol. Chem.* **266**, 23632–23636
- Ni, F., Ning, Q., Jackson, C. M., and Fenton, J. W., II (1993) *J. Biol. Chem.* **268**, 16899–16902
- Friedrich, T., Kröger, B., Bialojan, S., Lemaire, H. G., Höffken, H. W., Reuschenbach, P., Otte, M., and Dödt, J. (1993) *J. Biol. Chem.* **268**, 16216–16222
- Brandstetter, H., Turk, D., Höffken, H. W., Grosse, D., Stürzebecher, J., Martin, P. D., Edwards, B. F. P., and Bode, W. (1992) *J. Mol. Biol.* **226**, 1085–1099
- Brünger, A. T. (1992) *XPLOR Manual*, Version 3.1, Yale University, New Haven, CT
- Engh, R. A., and Huber, R. (1991) *Acta Crystallogr. Sec. A* **47**, 392–400
- van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kröger, B., Höffken, W., and Bode, W. (1995) *EMBO J.* **14**, 5149–5157
- Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sec. A* **47**, 110–119
- Turk, D. (1992) *Weiterentwicklung eines Programms für Molekülgrafik und Elektronendichte-Manipulation und seine Anwendung auf verschiedene Protein-Strukturaufklärungen*. Ph.D. thesis, Technische Universität München
- Banfield, D. K., and MacGillivray, R. T. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2779–2783
- Bode, W., and Huber, R. (1992) *Eur. J. Biochem.* **204**, 433–451
- Padmanabhan, K., Padmanabhan, K. P., Tulinsky, A., Park, C. H., Bode, W., Huber, R., Blankenship, D. T., Candin, A. D., and Kisier, W. (1993) *J. Mol. Biol.* **232**, 947–966
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950
- Barton, G. J., and Sternberg, M. J. E. (1987) *J. Mol. Biol.* **198**, 327–337
- Barton, G. J. (1990) *Methods Enzymol.* **183**, 403–428
- Barton, G. J. (1993) *Protein Eng.* **1**, 37–40