The EMBO Journal Vol.18 No.4 pp.804–814, 1999

Structural basis of profactor D activation: from a highly flexible zymogen to a novel self-inhibited serine protease, complement factor D

Hua Jing, Kevin J.Macon[1], Dwight Moore, Lawrence J.DeLucas, John E.Volanakis[1,2] and Sthanam V.L.Narayana[3]

Center for Macromolecular Crystallography and [1]Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, AL 35294, USA

[2]Present address: Biomedical Science Research Center 'A.Fleming', 166 72 Vari, Greece

[3]Corresponding author
e-mail: narayana@pearl.cmc.uab.edu

The crystal structure of profactor D, determined at 2.1 Å resolution with an R_{free} and an *R*-factor of 25.1 and 20.4%, respectively, displays highly flexible or disordered conformation for five regions: N-22, 71–76, 143–152, 187–193 and 215–223. A comparison with the structure of its mature serine protease, complement factor D, revealed major conformational changes in the similar regions. Comparisons with the zymogen–active enzyme pairs of chymotrypsinogen, trypsinogen and prethrombin-2 showed a similar distribution of the flexible regions. However, profactor D is the most flexible of the four, and its mature enzyme displays inactive, self-inhibited active site conformation. Examination of the surface properties of the N-terminus-binding pocket indicates that Ile16 may play the initial positioning role for the N-terminus, and Leu17 probably also helps in inducing the required conformational changes. This process, perhaps shared by most chymotrypsinogen-like zymogens, is followed by a factor D-unique step, the re-orientation of an external Arg218 to an internal position for salt-bridging with Asp189, leading to the generation of the self-inhibited factor D.
Keywords: conformational change/crystal structure/factor D/serine protease/zymogen activation

Introduction

Proteases are often synthesized as larger precursors, also termed pro-enzymes or zymogens, which carry N-terminal extensions that regulate the proteolytic activity (Neurath, 1984). The length of the activation peptides and their mechanisms of action vary largely from one protease to another, even within the same family, such as the serine protease family (Baker *et al.*, 1993; Khan and James, 1998). Serine proteases with a chymotrypsin fold have been studied extensively, and a wealth of information has accumulated on the structural basis of zymogen activation. Structural and functional comparisons between the zymogens and mature enzymes have demonstrated that the pro-enzymes of chymotrypsin-like serine proteases can be activated through several different mechanisms. One common mechanism involves the conformational change of an activation domain, such as in chymotrypsinogen (Freer *et al.*, 1970; Wang *et al.*, 1985), trypsinogen (Bode *et al.*, 1976; Kossiakoff *et al.*, 1977; Stroud *et al.*, 1977; Bode and Huber, 1978; Walter *et al.*, 1982) and prethrombin-2 (Vijayalakshmi *et al.*, 1994; Malkowski *et al.*, 1997). Another mechanism involves removal of a pro-region that blocks accessibility to the active site, such as in α-lytic protease where the C-terminal region of a pro-peptide obstructs the substrate-binding site (Sauter *et al.*, 1998). Additionally, a single-chain tissue-type plasminogen activator was shown to use a unique Lys156 to induce an active conformation in the zymogen (Bode and Renatus, 1997; Renatus *et al.*, 1997). Considering the diversity of the protease structures and functions, additional activation mechanisms can be expected.

The focus of our study is the activation mechanism of human complement factor D (FD). FD is a chymotrypsin-like serine protease essential for the activation of the alternative pathway of the complement system (reviewed in Volanakis and Narayana, 1996). FD circulates in blood devoid of an activation peptide, indicating a unique mechanism regulating its activity. Numerous crystallographic and mutational studies on FD have indeed shown it to be a unique serine protease. On the one hand, it has extremely restricted substrate specificity, cleaving a single bond of its only natural substrate, C3b-bound factor B, and displays low esterolytic activity towards synthetic substrates (Lesavre and Müller-Ebarhard, 1978; Kam *et al.*, 1987). On the other hand, it is a highly efficient protease that apparently does not require regulation by a specific activating enzyme or a natural protease inhibitor. These unique functional characteristics were shown to correlate well with the atypical structural features of the FD active site. In six independent FD structural variants (Narayana *et al.*, 1994a; Cole *et al.*, 1997, 1998; Jing *et al.*, 1998), we observed consistently atypical conformations for the catalytic triad residue His57 and loop 214–218 that lines one side of the primary substrate-binding pocket (S1 pocket). Its overlap with the inhibitor-binding positions in typical serine proteases and a salt bridge formed between Arg218 and Asp189 led to the proposal that loop 214–218 acts as a self-inhibitory loop that dictates the resting state conformation of FD (Jing *et al.*, 1998). This loop has also been proposed to be critical for binding the single natural substrate of FD, together with two other flexible loops near the active site. These structural features support the model proposed by Volanakis and Narayana (1996) that the natural substrate of FD induces a reversible conformational change on FD, from the resting state to an active state where the catalytic triad and substrate-binding site are in catalytic conformation.

Although no activation peptide is present on FD as it circulates in blood, its cDNA encodes, in addition to the

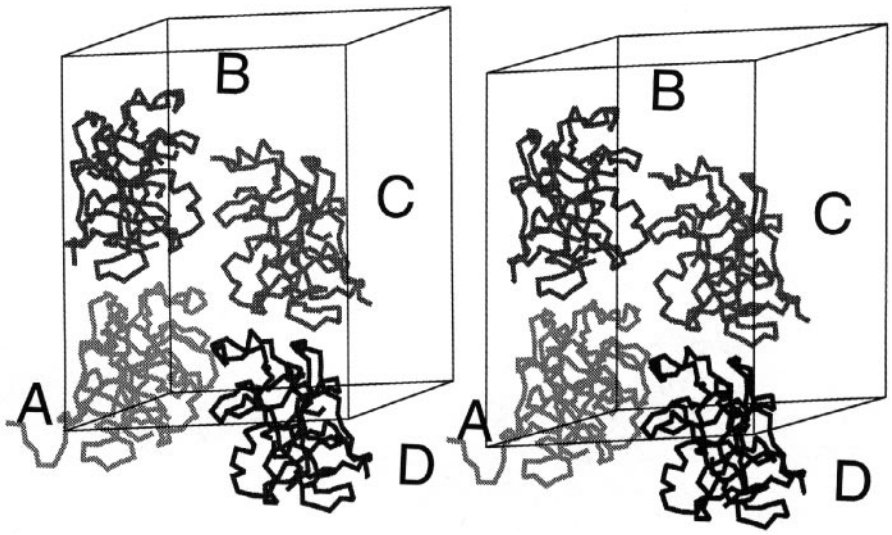

Fig. 1. The packing of four PFD molecules in one asymmetric unit.

mature FD sequence, a putative activation peptide (White *et al.*, 1992). This peptide is apparently cleaved off intracellularly by some unidentified endopeptidase before secretion of FD into blood. Expression of the cDNA using mammalian cells produces mature FD, whereas expression in insect cells using a baculovirus system produces pro-FD (PFD) containing a six or seven residue long activation peptide (Yamauchi *et al.*, 1994). PFD expresses no proteolytic activity, although it can be converted to functional FD by catalytic amounts of trypsin. Thus, it can be considered as a zymogen of FD (Yamauchi *et al.*, 1994). However, unlike zymogen activation of other typical serine proteases, the conversion of PFD to FD results in the atypical, self-inhibited conformation of the active site. To understand the structural basis of the unique PFD activation and its similarities to and differences from the canonical zymogen activation mechanisms, we determined the structure of PFD and compared it with those of FD and zymogen–active enzyme pairs of chymotrypsinogen, trypsinogen and prethrombin-2.

Results

Overall description

The structure of PFD was solved at 2.1 Å resolution in a $P2_1$ crystal form containing four molecules in the asymmetric unit. The four molecules (A, B, C and D) are arranged as two dimers related to each other by 12° of rotation plus approximately a half translational component along the *c* axis (Figure 1). Within each dimer, the two monomers are related to each other by a non-crystallographic 2-fold of 179.1 and 180.0°, respectively. The structure was refined to an R_{free} of 25.1% and an *R*-factor of 20.4 for the reflections ranging from 30 to 2.1 Å resolution (Table I). The Luzzati coordinates error of the structure is 0.25 Å. Of the 821 residues observed, 99.3% are in the core or additionally allowed regions in the Ramachandran plot, five residues (A: Asp153, B: Asp173, C: Ala61A, C: His171, C: His172) are in the generously allowed regions, and no residues are in the disallowed regions.

The observed residues generally have high real-space correlation coefficients (>0.9), as illustrated in Figure 2. However, the electron densities are diffused for a few flexible regions. Some of these regions have no traceable density, while others can be traced by lowering the density level to 0.5σ. The traceable flexible regions and the termini of the completely disordered regions usually have high *B*-factors that increase gradually from the ordered part to the flexible part (Figure 2). This pattern and the location of the high *B*-factor regions are consistent among the four molecules, thus indicating a true flexibility for those regions. Molecule A is the most complete structure of the four, missing only the N-terminal extension residues (N-15) and residues 144–151 (chymotrypsinogen numbering is used), while the other three molecules are missing more residues, including N-21, 144–151, 187–192 and 217–223 (Table I). As described in detail below, most of the disordered regions are involved in interactions with the N-terminus in the mature enzyme. Besides these four regions, three other regions also show flexibility in all four molecules (Figure 2): residues 60–61C are in loop L5 which is flexible in almost all FD structural variants (Jing *et al.*, 1998); residues 71–76 are in loop L6 which interacts with residues 144–151 ordered in FD but disordered in PFD; and residues 170–175 are in helix H2, which consistently has several residues with restrained geometry (Narayana *et al.*, 1994a; Jing *et al.*, 1998). After aligning the structures and excluding the regions with C_α–C_α distances >3.8 Å, the overall r.m.s.d. for each pair of the four PFD structures is ~0.5–0.6 Å (Table II).

Similarly to FD, PFD displays a two-domain β-barrel structure with six or seven β-strands in each domain (Figure 3). The flexible loops 60–61C and 71–76 are located within the N-terminal domain but close to the junction of the two domains, whereas all the other flexible regions are located within the C-terminal domain. Interestingly, all of these flexible regions are located on one region of PFD, which in chymotrypsin-like serine proteases carries all the structural elements for efficient catalysis: the catalytic triad (Asp102–His57–Ser195), the oxyanion hole (Gly193–Asp194–Ser195), the S1 pocket and the non-specific substrate-binding site (loop 215–217) (Perona and Craik, 1995). We therefore focus the structural comparisons on these elements to illustrate how their conforma-

Table I. Statistics on diffraction data and refined model

Resolution range (last shell) (Å)	30–2.1 (2.18–2.1)			
Data completeness	98.5% (96.8%)			
R_{merge}	6.6% (23.4%)			
Average I/σ	13.8 (3.9)			
No. of observations	93 825			
No. of unique reflections (⩾2σ)	43 713			
Cell dimensions	65.05, 70.32, 85.30			
	90.00, 101.98, 90.00			
R-factor	20.4%			
R_{free}	25.1%			
Luzzati coordinates error (Å)	0.25			
R.m.s.d. bonds	0.008			
R.m.s.d. angles	1.432			
R.m.s.d. dihedrals	28.415			
R.m.s.d. impropers	0.766			
No. of atoms	6766			
No. of residues in the asymmetric unit	821			
No. of water molecules	636			
Average *B*-factor of protein atoms	23.6			
Average *B*-factor of waters	32.4			
	A	B	C	D
Average *B*-factor	24.71	23.79	24.61	21.31
No. of observed residues	221	202	199	200
Observed residues	16–143	22–142	22–143	22–143
	152–243	152–186	153–186	152–186
		193–216	194–215	194–216
		223A–243	223A–243	224–243

tions can be affected by the N-terminal extension of the zymogen.

Comparison of PFD with FD

The structure of FD in complex with a small inhibitor, isatoic anhydride, was chosen to represent FD because it was determined at high resolution (1.5 Å), structural differences from native FD are minimal and one molecule is present in the asymmetric unit (Jing *et al.*, 1998). Molecule A was chosen to represent PFD because it is more complete and comprises the common structural features of molecules B, C and D.

In FD, the positively charged N-terminus forms a salt bridge with Asp194. The side chains of the first two residues, Ile16 and Leu17, participate in two distinct hydrophobic clusters. The hydrophobic cluster associated with Ile16 is internal, comprising Val138, Val158, Leu160, Leu199, Val213 and Tyr228, whereas that associated with Leu17 is smaller and more external, comprising Ile143, Cys191 and Cys220. Immediately following the N-terminus is a short β-strand (E1) located on the surface of FD and forming main chain hydrogen bonds with β-strand E9. Topologically, the N-terminus and the E1 β-strand seem to make up a 'hook' that extends from the C-terminal domain to the N-terminal domain, thus helping to hold the two domains together (Figure 3). In PFD, because of the presence of additional residues at the N-terminus, the above interactions involving the N-terminal region are disrupted. The extended N-terminal region flips out, forcing the E1 strand to assume a loop conformation (Figure 3). The N-terminal region extends towards a neighboring molecule and differs from the corresponding region of FD starting from residue His25.

Loop 144–151 becomes completely disordered because of the displacement of the N-terminus. In FD, this loop makes several hydrogen bonds with the N-terminal region, and the internal Ile16–Asp194 salt bridge is located directly underneath it. Some of the residues from this loop are close to the Ile16 and Leu17 hydrophobic clusters that become partially destabilized in PFD due to the absence of the interactions with Ile16 and Leu17. Loop 144–151 also interacts extensively with FD loops 71–76 and 187–192, both of which are disordered or flexible in PFD.

Loops 187–193 and 216–223 form two of the walls of the S1 pocket in FD, where the latter is also part of the non-specific substrate-binding site. These two loops are linked to each other through a disulfide bond between Cys191 and Cys220 and the FD-unique salt bridge between Asp189 and Arg218. In PFD, while these loops are disordered in molecules B, C and D, they are traceable in molecule A, and loop 188–193 is re-arranged into a 3_{10} helix. Arg218 present on the left side of the S1 pocket flips out and its guanidino group extends over the pocket to the right side wall, where it makes a pseudo salt bridge with the carboxyl end of the newly formed 3_{10} helix (Figure 3). Thus, the side chain of the Arg218 may act as a latch, blocking the entrance to the S1 pocket. The ordered conformation of these two loops in molecule A preserves the disulfide bond between Cys191 and Cys220, but their C_α atoms are 5 or 7 Å away from the respective positions in FD. It should be noted, however, that some intra- and intermolecular interactions observed in molecule A, but not in molecules B, C and D, may be stabilizing the conformations of these two loops in molecule A.

The disordering of loop 187–193 in molecules B, C and D and its acquisition of 3_{10} helical conformation in molecule A result in the distortion of the geometry of the oxyanion hole geometry in PFD. In the active serine proteases, the oxyanion hole is composed of the amine groups of three highly conserved residues, Gly193, Asp194

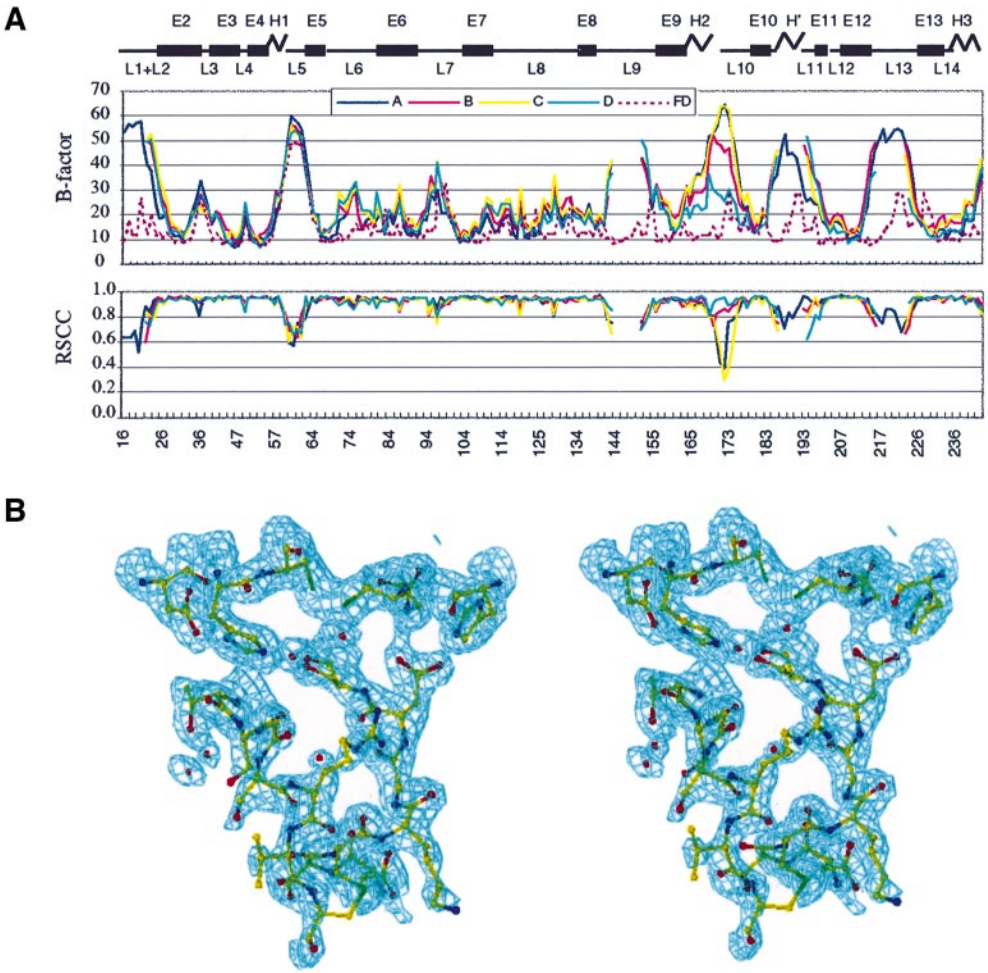

Fig. 2. Structure qualities as indicated by (**A**) *B*-factor and real-space correlation coefficient plots and (**B**) electron density map of the active site. Secondary structure elements shown on the top are numbered sequentially from the N- to the C-terminus, as those in FD. L, loops; E, β-strands; H, helices.

Table II. Comparison of PFD, FD, chymotrypsinogen, trypsinogen and prethrombin-2[a]

	A	B	C	D
A	–	0.492 (193)	0.594 (190)	0.620 (190)
B		–	0.540 (190)	0.604 (196)
C			–	0.581 (190)
D				–
	FD	chymotrypsin	trypsin	thrombin
Zymogen molecule A	0.790 (188)	0.722 (227)	0.620 (205)	0.679 (216)
Zymogen molecule B	0.775 (192)	0.712 (227)	0.498 (212)	0.680 (209)

[a]R.m.s.d.s are calculated for all regions excluding the C_α–C_α distances >3.8 Å. The number of matched C_α atoms is listed in parentheses. The PDB structures selected for comparisons are FD (1BIO, Jing *et al.*, 1998), chymotrypsin (4CHA, Tsukada and Blow, 1985), trypsin (1PPC, Bode *et al.*, 1990), thrombin (1PPB, Bode *et al.*, 1989), chymotrypsinogen (2CGA, two NCS-related molecules), trypsinogen (1TGN and 2TGT) and prethrombin-2 (1HAG and 1MKX).

and Ser195, which together stabilize the negative charge developing on the carbonyl oxygen of the P1 residue during the transition state. In PFD, the carbonyl group of Gly193 points towards the corresponding oxyanion hole position in FD, thus changing the charge composition and geometry of the oxyanion hole, resulting in disfavoring of the binding of the P1 residue. This conformational change is accompanied by a twist in Asp194 which, instead of salt-bridging with the mature N-terminus, forms hydrogen bonds with His40–NE2 and/or the main chain nitrogen or oxygen atoms of residues 43, 139, 141 or 142.

The catalytic triad displays active conformation. This structural difference from FD can be explained by the removal of the steric hindrance between His57 and the self-inhibitory loop 214–218. The flipped conformation of Arg218 described above results in lowering of the backbone of the self-inhibitory loop and subsequently of the side chain of Ser215, thus allowing His57 to take up the

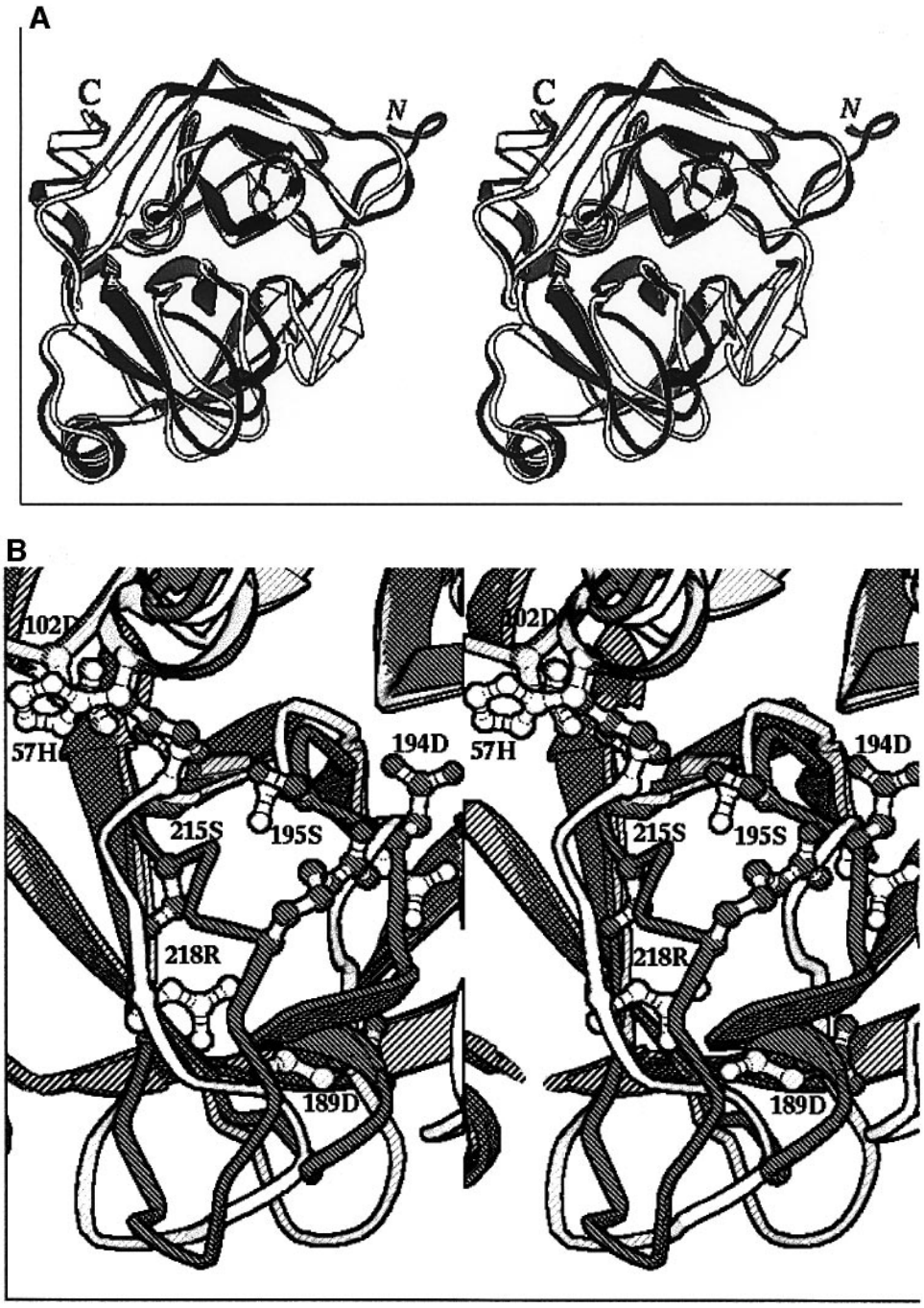

Fig. 3. Stereo view of the overall and active site conformations of PFD in comparison with FD. (**A**) Overall structure and (**B**) active site. Molecule A of PFD is shaded and FD is not shaded. The figure was prepared using RIBBONS (Carson, 1997).

typical serine protease conformation. However, His57 in molecule C displays an atypical conformation similar to that in FD, despite the absence of the self-inhibitory loop. The side chain of His57 seems to have been pushed away by the side chain of Ser195 whose C_α atom is shifted, along with residues 194–197, by 1.5 Å towards His57.

In summary, the comparison between FD and PFD shows that the major conformational changes are located in the regions that are disordered or flexible in PFD. Except for the catalytic triad, the other three catalytic elements, oxyanion hole, S1 pocket and non-specific binding site, have either disordered (molecule B, C and D) or inactive (molecule A) conformations in PFD. This deformed catalytic apparatus could explain the resistance of PFD to diisopropyl fluorophosphate (DFP) inhibition and its lack of proteolytic activity (Yamauchi *et al.*, 1994).

Comparison of PFD with other zymogens

The flexible regions and active catalytic triad conformation described above for PFD had also been observed previously in a number of other zymogens, such as chymotrypsinogen, trypsinogen and prethrombin-2 (Table III). In a review by Huber and Bode (1978), the disordered regions in trypsinogen were first said to comprise an 'activation domain' that is inter-digitized in the active enzyme. Glycine residues are found frequently at the

Table III. Known zymogen structures in the chymotrypsin-fold family

Zymogen	Conformation state	PDB code (resolution)	Reference
Chymotrypsinogen	native	1CHG (2.5)	Freer *et al.* (1970)
		2CGA (1.8)	Wang *et al.* (1985)
	+ natural inhibitor	1CGI (2.3), 1CGJ (2.3)	Hecht *et al.* (1991)
	+ procarboxypeptidase and proproteinase E	1PYT (2.35)	Gomis-Ruth *et al.* (1995, 1997)
Trypsinogen	native	1TGN (1.65)	Kossiakoff *et al.* (1977)
		1TGB (1.8)	Fehlhammer *et al.* (1977)
		1TGC (1.8), 1TGT (1.7)	Walter *et al.* (1982)
		2TGA (1.8), 2TGT (1.7)	Walter *et al.* (1982)
	+ natural inhibitor	4TPI (2.2)	Bode *et al.* (1978)
		2TPI (2.1)	Walter *et al.* (1982)
		1TGS (1.8)	Bolognesi *et al.* (1982)
		3TPI (1.9), 2TGP (1.9)	Marquart *et al.* (1983)
	+ DFP	2TGD (2.1)	Jones and Stroud (1986)[a]
Prethrombin-2	+ hirugen	1HAG (2.0)	Vijayalakshmi *et al.* (1994)
	+ α-thrombin	1MKW (2.3), 1MKX (2.2)	Malkowski *et al.* (1997)
Proproteinase E	truncated form	1FON (1.7)	Pignol *et al.* (1994)
	+ chymotrypsinogen and procarboxypeptidase	1PYT (2.35)	Gomis-Ruth *et al.* (1995, 1997)
Single-chain tissue-type plasminogen activator	+ synthetic inhibitor	1BDA (3.25)	Renatus *et al.* (1997)
α-Lytic protease	complex with pro-region	4PRO (2.4)	Sauter *et al.* (1998)
Profactor D	native	1FDP (2.1)	this study

[a]Protein Data Bank, accession code 2tgd.

termini of the disordered regions, and some of them have aromatic residues next to the glycines. These residues have been proposed to act as hinges and anchors, respectively, facilitating the process of zymogen activation. We find that the overall distribution of flexible regions in PFD is similar to those in chymotrypsinogen, trypsinogen and prethrombin-2 and correlates with the distribution of the high *B*-factor regions (Figure 4). However, the magnitude and width of the conformational changes following activation peptide cleavage are clearly different (Figure 5). In an attempt to understand further the structural basis of zymogen activation, especially the uniqueness of PFD to FD conversion, we compared these four zymogen–active enzyme pairs and correlated the differences in conformational changes with the differences in primary and secondary structures of the regions involved.

The activation peptides of PFD and trypsinogen are six or seven residues long, whereas those of chymotrypsinogen and prethrombin-2 are much longer and linked to the main body of the structures by identical disulfide bonds between Cys1 and Cys122. Consistently, the activation peptides of chymotrypsinogen and prethrombin-2, despite having high *B*-factors, patch on the protein surface with some loop and some helical conformations, whereas those of PFD and trypsinogen are completely missing (Figure 4). Therefore, the length of an activation peptide and its disulfide linkage with the rest of the structure probably determine its degree of ordering. When the activation peptide gets even longer, such as that in α-lytic protease, the pro-region folds up into an individual domain and facilitates protein folding (Sauter *et al.*, 1998). Interestingly, the fold displayed by this pro-region is similar to those in procarboxypeptidase (Coll *et al.*, 1991; Guasch *et al.*, 1992) and pro-subtilisin (Gallagher *et al.*, 1995) that nevertheless display a serine protease fold different from the chymotrypsinogen-fold.

The N-terminal region of PFD is distinct among the four zymogens because it deviates significantly from that in FD up to residue His25, as compared with Gly19 in the other three zymogens. Also, the C_α atom of Ile16 undergoes a 34.7 Å shift during the PFD to FD conversion, as compared with shifts of 10.7 Å in chymotrypsinogen, 13.0 Å in trypsinogen and 16.5 Å in prethrombin-2. This implies that the N-terminal region of PFD would be more exposed, as perhaps required for efficient cleavage and secretion. Gly19 is conserved among the four zymogens and probably plays two roles in zymogen activation: in the mature enzymes, it allows the main chain of the N-terminal region to bend and thus helps to position the N-terminus near the Asp194; in the zymogens, its flexibility helps to make the N-terminal regions accessible to activating enzymes. However, in FD–PFD it appears that the conserved Gly19 plays only the former role, while the FD-unique His25 contributes to the latter function.

As in PFD, loop 144–151 also changes dramatically during the activation of chymotrypsinogen, trypsinogen and prethombin-2 (Figure 5). While it is completely disordered in PFD, a shift of up to 9.8 Å (residue 145) is associated with the activation of chymotrypsinogen, 14.4 Å (residue 149) with that of trypsinogen, and 8.6 Å (residue 144) with that of prethrombin-2. These changes are apparently due to the absence of the interactions with the N-termini, as described above for PFD. A sequence motif Gly140–Trp141–Gly142 present at the N-terminal end of this loop is conserved among the four enzymes, and Trp141 constantly marks the beginning point of the conformational change (Figure 5). The C-terminal end of this loop, however, does not contain any glycine residue, and residue 153 (Asp or Ser), or Ser154 in FD, constantly marks the ending point of this conformational change. The next residue, Leu155, is conserved among the four zymogens, and thus may act as the anchor defining the extent of the conformational change. Loop 144–151 is

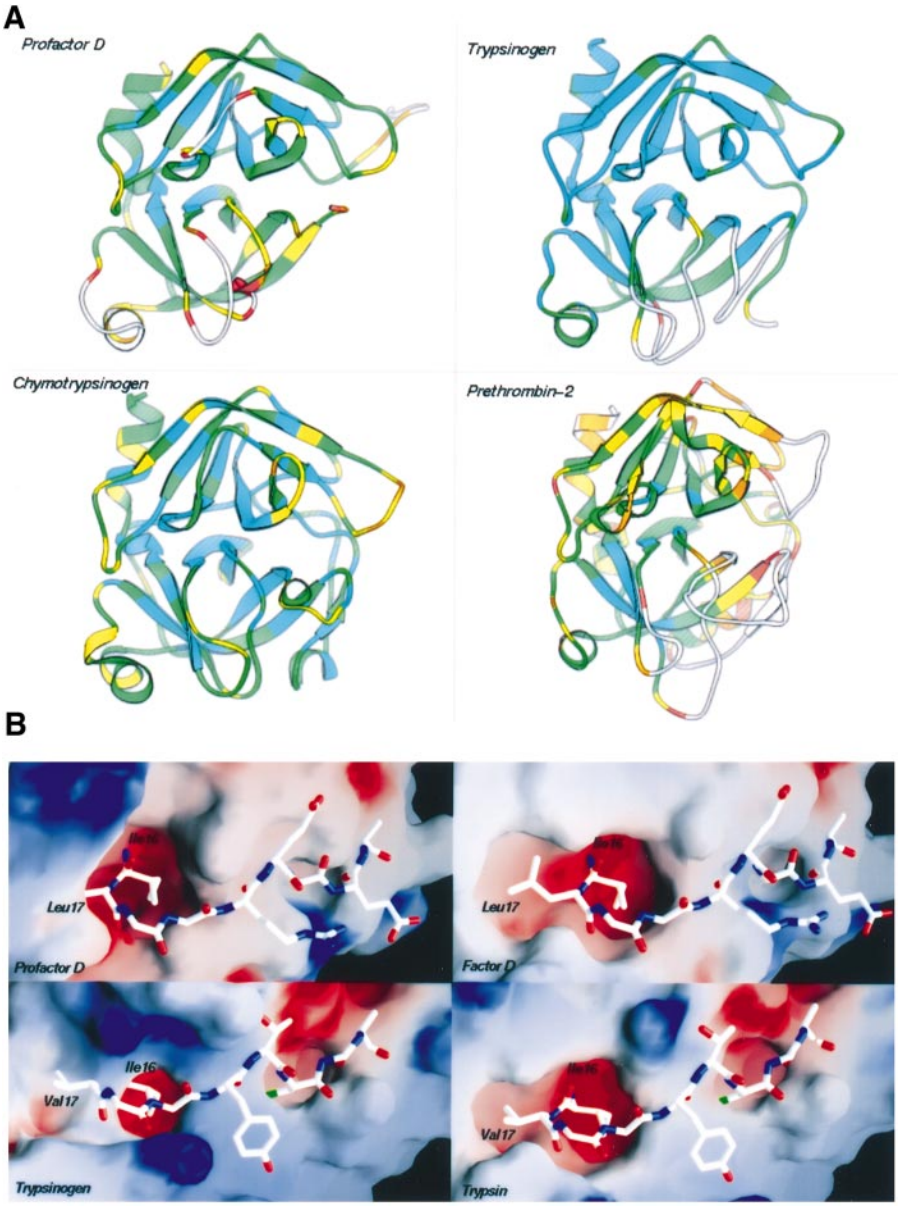

Fig. 4. Comparisons of four zymogen–enzyme pairs. (**A**) High *B*-factor regions in PFD, chymotrypsinogen (2CGA-A), trypsinogen (2TGT) and prethrombin-2 (1HAG). From low to high *B*-factors are shown in cyan, green, yellow, orange, red and white. This panel was prepared using RIBBONS (Carson, 1997). (**B**) Surface charge of the N-terminus-binding pockets in PFD–FD and trypsinogen–trypsin. Residues 16–25 of the mature enzymes are shown in ball-and-stick model and are excluded in the surface calculations. They are superimposed onto the surface of the zymogens based on structural alignments. This panel was prepared using GRASP (Nicolls *et al.*, 1991).

connected by hydrogen bonding to another loop, 70–79, which also undergoes hinge conformational change during the activation of PFD and chymotrypsinogen. Thus, it probably should also be considered as a component of the activation domain.

The 3_{10} helical conformation of residues 185–193 of PFD was also identified in two independently determined prethrombin-2 structures (Vijayalakshmi *et al.*, 1994; Malkowski *et al.*, 1997), although the helices are a little shorter than the one in PFD. The longer helix in PFD could result from the stabilization of the helix dipole by the guanidino group of Arg218. The maximum shift of this loop during zymogen activation could range from 5.5 to 8.6 Å (Figure 5). Despite the dissimilar conformations for this loop, a common feature shared by the four zymogens is the distortion of the oxyanion hole geometry. This is perhaps the common mechanism used by the zymogens to disable the premature substrate binding and cleavage. The N-terminal end of this loop in PFD lacks the conserved Gly184, and the C-terminal end has a highly conserved GDSGGP sequence patch for residues 193–198. However, no hydrophobic residues are conserved next to these glycines.

Loop 216–223 is strikingly different in PFD compared with the other zymogens. Its conformational change for PFD is the largest among the four zymogens, in terms of both magnitude and width. A shift up to 10 Å was observed during PFD conversion to FD, as compared with 2.6 Å during chymotrypsinogen, 4.3 Å during trypsinogen, and 6.4 Å during prethrombin-2 activation. Twelve residues differ by >1.0 Å from the corresponding positions in FD, as compared with 6–8 residues in the activation of

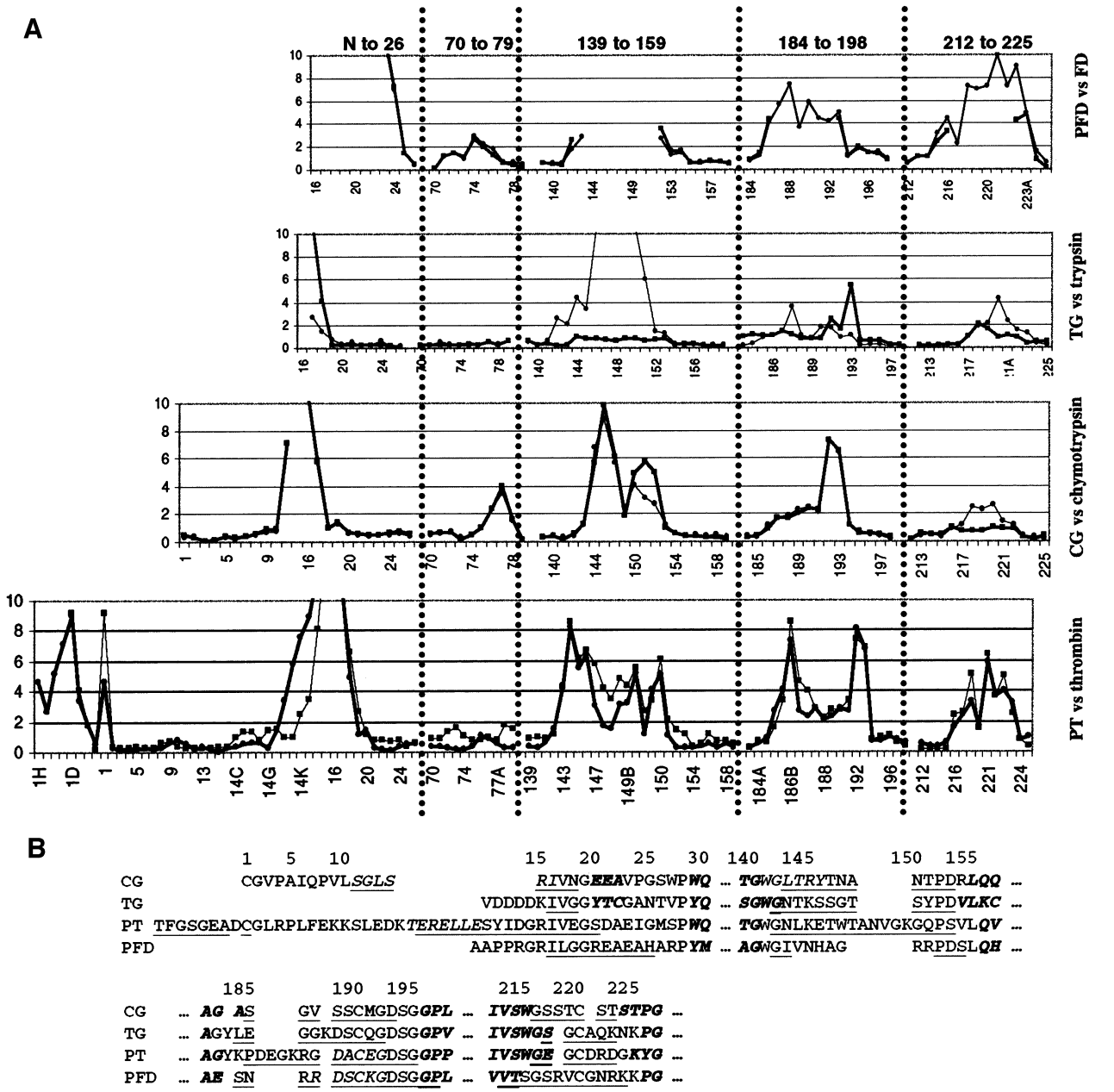

Fig. 5. The regions with substantial conformational changes during zymogen activation. (**A**) The magnitudes and widths of conformational change, as indicated by the C_α–C_α distances between a mature enzyme and two of its zymogen structures (indicated by thick and thin lines, respectively). The PDB codes of the compared structures are shown in Table II. Shifts of >10 Å are not shown. (**B**) Amino acid sequences of the activation peptides and the flexible regions. The residues in β-strands are bold and italicized, the residues in helices are italicized, and the residues with >1.0 Å shifts are underlined.

the other three zymogens (Figure 5). In the other three zymogens, the conformations of the loop are the same as those in the active enzymes up to Trp215 because the interactions mediated by Trp215 are the same in the zymogens as in the active enzymes. These interactions, including a hydrophobic clustering with neighboring residues, a β-sheet interaction between the backbone of residues 212–215 and 226–229, and several other hydrogen bonds such as that between Ser214 and Asp102, make this loop very stable in both the zymogens and active enzymes. In contrast, FD has a threonine at position 214, serine at position 215 and arginine at position 218. As a result, this region takes a loop conformation in FD, losing not only the hydrophobic interactions mediated in other serine proteases by the conserved Trp215, but also the β-sheet-forming potential. A slight compensation for this loss is an FD-unique salt bridge between Arg218 and Asp189. This interaction is probably enhanced by the nearby disulfide bond between Cys220 and Cys191, which, however, is common to every trypsin-like serine protease. In PFD, this salt bridge is absent, although Arg218 in molecule A forms a less stable pseudo-salt bridge with the carboxyl end of the 3_{10} helix.

Discussion

In the present study, we determined the crystal structure of PFD and compared it with that of FD and three other

zymogen–active serine protease pairs. The comparisons were focused on the effects of the activation peptide on the geometry of the active site, including that of the substrate-binding pocket, the non-specific substrate-binding site, the catalytic triad and the oxyanion hole. The similarities and differences derived from the comparisons revealed, in addition to our previous understanding of the zymogen activation process, other structural aspects of zymogen activation, such as the degree of the conformational changes and its relationship with the primary structure. Here, we summarize the common features shared by the zymogen activation, explore the possible origin of these conformational changes and address in particular how the self-inhibited conformation of FD could be shaped during the zymogen activation.

Zymogen activation in general

The comparison of the four zymogen–active enzyme pairs showed that some aspects of zymogen activation are common. These include the similar distribution of the flexible regions in the zymogens, the maintenance of a typical catalytic triad conformation during the activation (except for FD, as described below) and the induction of the active conformation of the primary substrate-binding site and oxyanion hole in the zymogens. Therefore, zymogen activation can be considered as a process in which functional catalytic elements, namely the substrate-binding site and the oxyanion hole, mature by direct or indirect interactions with the newly formed N-terminus. This process is accompanied by large conformational changes, from a highly flexible zymogen to an ordered mature enzyme. This ordered conformation, shown to be similar in many mature serine proteases, is perhaps critical for specific substrate binding and cleavage, whereas the flexible and proteolytically non-functional conformation of the zymogens is essential for preventing premature proteolysis. It should be mentioned that in some of the zymogen structures compared, intermolecular contacts contribute to a certain extent to the stabilization of regions of the activation domain. Such an effect is described here for PFD and has been described previously for chymotrypsinogen (Wang *et al.*, 1985) and prethrombin-2 (Vijayalakshmi *et al.*, 1994; Malkowski *et al.*, 1997). However, in solution, these regions could be expected to be more flexible.

The origin of the conformational change during zymogen activation had been considered commonly to be driven by the electrostatic interaction between Asp194 and the N-terminus. Recently, mutagenesis and thermodynamics studies on trypsinogen showed that the hydrophobic moiety of Ile16 contributes more to activation than its N-terminal positive charge (Hedstrom *et al.*, 1996). An examination of the surface charge distribution around the N-terminus-binding pocket suggested, in addition to these interactions, a possible hydrophobic interaction mediated by residue 17. As shown in Figure 4, the binding pocket for the N-terminus is pre-formed in both PFD and trypsinogen and is slightly negatively charged. Thus, it may have a potential to attract the positively charged N-terminus, Ile16, for initial positioning. The active enzymes also have binding pockets for the next residue, Leu17 in FD and Val17 in trypsin. In the zymogens, while the pocket for Ile16 is partially shaped and spacious, the pocket for residue 17 requires substantial conformational change to accommodate its side chain. As described above in FD, Leu17, Cys191 and Cys220 form a small hydrophobic cluster located near the surface of the mature enzyme. The hydrophobic cluster is also conserved in the other three mature enzymes. Therefore, the association of residue 17 with the two cysteine resides, upon cleavage of the activation peptide, would probably result in the correct positioning of the disulfide bond and its associated loops, 185–193 and 216–223. This process may be facilitated by the formation of the Ile16–Asp194 salt bridge and the association of Ile16 with its hydrophobic cluster, and it may occur in concert with the correct positioning of Ile16. Conformational change of the activation domain can also be induced by strong ligand binding in the presence or absence of the activation peptide (Bode *et al.*, 1978; Huber and Bode, 1978; Bode, 1979). Therefore, ordering of the activation domain is a cooperative phenomenon.

Unique conversion of PFD to FD

Unlike the other zymogen activation processes, the conversion of PFD to FD is unique in that the product is not an active enzyme, but an enzyme displaying a unique resting state conformation that is not regulated by any ligands (Volanakis and Narayana, 1996). As the resting state conformation is dictated by the self-inhibitory loop 214–218 (Jing *et al.*, 1998), the formation of this loop is critical for the PFD to FD conversion. The structural comparisons have shown that, unlike the other zymogen–active enzyme pairs, the self-inhibitory loop does not have strong interactions with the rest of the structure, except for forming a salt bridge between Arg218 and Asp189. In PFD, Arg218 seems to be associated temporally with the carboxyl end of a helix that is highly flexible by itself. In addition, the distinct plasticity of PFD, especially around the substrate-binding site, seems to provide a suitable environment allowing Arg218 to flip easily. Therefore, once the Cys191–Cys220 disulfide bond interacts with Leu17, the nearby Arg218 probably assumes a stable conformation by salt-bridging with Asp189, which seems to be the only and the best choice. Consequently, Ser215 from this loop will flip upwards and push His57 away to the atypical position, which is nevertheless a common conformation among the normal histidines in protein structures (Sprang *et al.*, 1987).

Compared with the other mature serine proteases, the fewer interactions of the self-inhibitory loop with the rest of the structure imply that this loop has the potential to undergo further substrate-induced conformational change (Volanakis and Narayana, 1996). Thus, the flexibility of the loop is intrinsic, but hidden in the resting state of FD.

Taken together, the structural comparisons suggest that activation of FD may proceed through two steps. The first step would be the irreversible zymogen activation step where the active site changes from the one displaying molecule A-like conformation, with typical catalytic triad, deformed S1 pocket and non-functional oxyanion hole, to that displaying the self-inhibited conformation, with atypical catalytic triad, obstructed S1 pocket and functional oxyanion hole. The second step would be the *in situ*, reversible conformational change induced only by the single natural substrate, where the active site changes to an active conformation, with typical catalytic triad, available S1 pocket and functional oxyanion hole. The three

different conformational states are highly correlated with their distinct functions: the first one (PFD) is proteolytically inactive and DFP resistant; the second (resting-state FD) is also proteolytically inactive, but is susceptible to DFP inhibition and has low reactivity towards peptide thioester substrates; and the third is perhaps similar to trypsin, expressing proteolytic activity as well as normal esterolytic activity, although the latter cannot be demonstrated.

Two complement serine proteases, factor B and C2, have no functional N-termini in the proteolytically active forms (reviewed in Volanakis and Arlaud, 1998). Together with the unique PFD to FD conversion and substrate-induced FD activation, these examples certainly add more diversity to the zymogen activation mechanisms and to the regulation of serine protease activities.

Materials and methods

Protein expression and purification

The recombinant PFD was expressed and purified as described previously (Yamauchi *et al.*, 1994). Briefly, a plasmid containing hg31 factor D cDNA was used to co-transfect Sf9 insect cells with wild-type baculovirus AcNPV DNA. Recombinant virus stocks were used to transfect insect cells in serum-free media. The recombinant PFD was isolated from the cell growth media using Bio-Rex 70 and MonoS cation exchange chromatography.

Crystallization and data collection

PFD was dialyzed against buffer 10 mM Tris, 0.1 M NaCl, pH 7.0, and concentrated to 10 mg/ml. Crystals were obtained at room temperature by the hanging drop vapor diffusion method. The reservoir solutions contained 10–12% PEG-6000 and 30 mM MES (pH 5.2) (Narayana *et al.*, 1994b). The drops contained an equal volume of reservoir solution and PFD. Diffraction data were collected at 95 K using mother liquor plus 8% ethylene glycol as cryo-protectant, on an in-house X-ray generator (40 kV, 100 mA) and RAXIS-IV detector. The solvent content of the crystal is 36% and the V_m value is 1.9 $Å^3$/Da. Diffraction data were processed using DENZO and SCALEPACK programs (Otwinowski and Minor, 1997).

Structure determination and analysis

The structure was solved by the molecular replacement method using the FD structure as the search model. The orientation and location of the four molecules were determined using AMoRe (Navaza, 1994). Only three cross Patterson peaks were identified initially, and the fourth peak was deduced from the relationship of these three and a sense for favorable crystal packing. After rigid-body refinement, the correlation coefficient and *R*-factor are 55.7 and 43.2%, respectively.

Real-space 4-fold non-crystallographic symmetry (NCS) averaging was performed using the program DM (Cowtan and Main, 1993), and a monomer mask built using MAMA (Kleywegt and Jones, 1994) and O (Jones *et al.*, 1991). For the regions with no obvious densities in the averaged map, the corresponding residues are removed from the molecular replacement model, and they were added later only when the density showed up. With the phase restraints from averaging, the partial structure was subject to torsion angle dynamics and positional refinements in X-PLOR (v3.851) (Brünger, 1996) without NCS restraints. Except for 10% of the reflections excluded for R_{free} calculation (Brünger, 1992; Kleywegt and Brünger, 1996), all reflections from 8.0 to 2.1 Å were used in the refinement, and at later stages solvent corrections allowed the inclusion of low-resolution data to 20 Å. Sigma-weighted electron density maps (Read, 1986) were calculated with reflections from 20 to 2.1 Å and they guided manual rebuilding on O (Jones *et al.*, 1991). Special care was given to the flexible regions. The OOPS program (Kleywegt and Jones, 1996, 1997) was used throughout cycles of rebuilding for quality checks.

The final structure was checked using PROCHECK (Morris *et al.*, 1992; Leskowski *et al.*, 1993), WHAT_CHECK (Hooft *et al.*, 1996) and X-PLOR (Brünger, 1996). The PFD structures were aligned with each other and with FD and three other zymogens structures using O (Jones *et al.*, 1991) and LSQMAN (Kleywegt, 1996). The structure and diffraction data have been deposited in the Protein Data Bank with accession codes of 1fdp and r1fdpsf, respectively.

Acknowledgements

The authors are grateful to staff at the Center for Macromolecular Crystallography for excellent technical assistance on X-ray facilities and computers and many helpful discussions. H.J. appreciates the discussions on CCP4, X-PLOR and O newsgroups, especially the inspirations from Dr Gerard J.Kleywegt (Uppsala), Dr Kevin D.Cowtan (York) and Dr Eleanor J.Dodson (York) on averaging and other refinement practices. The projects is funded by NIH grant (AI 39818) to S.V.L.N.

References

Baker,D., Shiau,A.K. and Agard,D.A. (1993) The role of pro regions in protein folding. *Curr. Opin. Cell Biol*., **5**, 966–970.

Bode,W. (1979) The transition of bovine trypsinogen to a trypsin-like state upon strong ligand-binding. II. The binding of the pancreatic trypsin inhibitor and of isoleucine–valine and of sequentially related peptides to trypsinogen and *p*-guanidinobenzoate-trypsinogen, *J. Mol. Biol.*, **127**, 357–374.

Bode,W. and Huber,R. (1978) Crystal structure analysis and refinement of two variants of trigonal trypsinogen. *FEBS Lett*., **90**, 265–269.

Bode,W. and Renatus,M. (1997) Tissue-type plasminogen activator: variants and crystal/solution structures demarcate structural determinants of function. *Curr. Opin. Struct. Biol.*, **7**, 865–872.

Bode,W., Fehlhammer,H. and Huber,R. (1976) Crystal structure of bovine trypsinogen at 1.8 Å resolution. I. Data collection, application of Patterson search techniques and preliminary structural interpretation. *J. Mol. Biol.*, **106**, 325–335.

Bode,W., Schwager,P. and Huber,R. (1978) The transition of bovine trypsinogen to a trypsin-like state upon strong ligand-binding. The refined crystal structures of the bovine trypsinogen–pancreatic trypsin inhibitor complex and of its ternary complex with Ile–Val at 1.9 Å resolution. *J. Mol. Biol.*, **118**, 99–112.

Bode,W., Mayr,I., Baumann,U., Huber,R., Stone,S.R. and Hofsteenge,J. (1989) The 1.9 Å refined structure of human α-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J*., **8**, 3467–3475.

Bode,W., Turk,D. and Stuerzebecher,J. (1990) Geometry of binding of the benzamidine- and arginine-based inhibitors NAPAP and MQPA to human α-thrombin: X-ray crystallographic determination of the NAPAP–trypsin complex and modeling of NAPAP–thrombin and MQPA–thrombin. *Eur. J. Biochem.*, **193**, 175–182.

Bolognesi,M., Gatti,G., Menegatti,E., Guarneri,M., Marquart,M., Papamokos,E. and Huber,R. (1982) Three-dimensional structure of the complex between pancreatic secretory trypsin inhibitor (Kazal-type) and trypsinogen at 1.8 Å resolution. Structure solution, crystallographic refinement and preliminary structural interpretation. *J. Mol. Biol*., **162**, 839–868.

Brünger,A.T. (1992) Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.

Brünger,A.T. (1996) *X-PLOR (Version 3.851): A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.

Carson,M. (1997) Ribbons. *Methods Enzymol.*, **277**, 493–505.

Cole,L.B., Chu,N., Kilpatrick,J.M., Volanakis,J.E., Narayana,S.V.L. and Babu,Y.S. (1997) Structure of diisopropyl fluorophosphate-inhibited factor D. *Acta Crystallogr.*, **D53**, 143–150.

Cole,L.B., Kilpatrick,J.M., Chu,N. and Babu,Y.S. (1998) Crystal structure of 3,4-dichloroisocoumarin inhibited factor D. *Acta Crystallogr.*, **D54**, 711–718.

Coll,M., Guasch,A., Aviles,F.X. and Huber,R. (1991) Three-dimensional structure of porcine procarboxypeptidase B: a structural basis of its inactivity. *EMBO J.*, **10**, 1–9.

Cowtan,K.D. and Main,P. (1993) Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallogr*., **D49**, 148–157.

Fehlhammer,H., Bode,W. and Huber,R. (1977) Crystal structure of bovine trypsinogen at 1.8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J. Mol. Biol*., **111**, 415–438.

Freer,S.T., Kraut J., Robertus,J.D., Wright,H.T. and Xuong N.H. (1970) Chymotrypsinogen: 2.5 Å crystal structure, comparison with α-chymotrypsin and implications for zymogen activation. *Biochemistry*, **9**, 1997–2009.

Gallagher,T., Gilliland,G., Wang,L. and Bryan,P. (1995) The prosegment–subtilisin BPN′ complex: crystal structure of a specific foldase. *Structure*, **3**, 907–914.

Gomis-Ruth,F.X., Gomez-Ortiz,M., Bode,W., Huber,R. and Aviles,F.X. (1995) The three-dimensional structure of the native ternary complex of bovine pancreatic procarboxypeptidase A with proproteinase E and chymotrypsinogen C. *EMBO J.*, **14**, 4387–4394.

Gomis-Ruth,F.X., Gomez-Ortiz,M., Vendrell,J., Ventura,S., Bode,W., Huber R. and Aviles,F.X. (1997) Crystal structure of an oligomer of proteolytic zymogens: detailed conformational analysis of the bovine ternary complex and implications for their activation. *J. Mol. Biol.*, **269**, 861–880.

Guasch,A., Coll,M., Aviles,F.X. and Huber,R. (1992) Three-dimensional structure of porcine pancreatic procarboxypeptidase A. A comparison of the A and B zymogens and their determinants for inhibition and activation. *J. Mol. Biol.*, **224**, 141–157.

Hecht,H.J., Szardenings,M., Collins,J. and Schomburg,D. (1991) Three-dimensional structure of the complexes between bovine chymotrypsinogen A and two recombinant variants of human pancreatic secretory trypsin inhibitor (Kazal-type). *J. Mol. Biol.*, **220**, 711–722.

Hedstrom,L., Lin,T.-Y. and Fast,W. (1996) Hydrophobic interactions control zymogen activation in the trypsin family of serine proteases. *Biochemistry*, **35**, 4515–4523.

Hooft,R.W., Vriend,G., Sander,C. and Abola,E.E. (1996) Errors in protein structures. *Nature*, **381**, 272.

Huber,R. and Bode,W. (1978) Structural basis of the activation and action of trypsin. *J. Am. Chem. Soc.*, **11**, 114–122.

Jing,H., Babu,Y.S., Moore,D., Kilpatrick,J.M., Liu,X.-Y., Volanakis,J.E. and Narayana,S.V.L. (1998) Structures of native and complexed complement factor D: implications of the atypical His57 conformation and self-inhibitory loop in the regulation of specific serine protease activity. *J. Mol. Biol.*, **282**, 1061–1081.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.

Kam,C.-M., McRae,B.J., Harper,J.W., Niemann,M.A., Volanakis,J.E. and Powers,J.C. (1987) Human complement proteins D, C2 and B: active site mapping with peptide thioester substrates. *J. Biol. Chem.*, **262**, 3444–3451.

Khan,A.R. and James,M.N.G. (1998) Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci.*, **7**, 815–836.

Kleywegt,G.J. (1996) Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr.*, **D52**, 842–857.

Kleywegt,G.J. and Brünger,A.T. (1996) Checking your imagination: application of the free *R*-value. *Structure*, **4**, 897–904.

Kleywegt,G.J. and Jones,T.A. (1994) Halloween... masks and bones. In Bailey,S., Hubbard,R. and Waller.,D. (eds), *From First Map to Final Model.* SERC Daresbury Laboratory, Warrington, pp. 59–66.

Kleywegt,G.J. and Jones,T.A. (1996) Efficient rebuilding of protein structures. *Acta Crystallogr.*, **D52**, 829–832.

Kleywegt,G.J. and Jones,T.A. (1997) Model rebuilding and refinement practice. *Methods Enzymol.*, **276**, 208–230.

Kossiakoff,A.A., Chambers,J.L., Kay,L.M. and Stroud,R.M. (1977) Structure of bovine trypsinogen at 1.9 Å resolution. *Biochemistry*, **16**, 654–664.

Lesavre,P.H. and Müller-Eberhard,H.J. (1978) Mechanism of action of factor D of the alternative complement pathway. *J. Exp. Med.*, **148**, 1498–1509.

Leskowski,R.A., MacArthur,M.W., Moss,D.S. and Thronton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.

Malkowski,M.G., Martin,P.D., Guzik,J. C and Edwards,B.F.P. (1997) The co-crystal structure of unliganded bovine α-thrombin and prethrombin-2: movement of the Tyr-Pro-Pro-Trp segment and active site residues upon ligand binding. *Protein Sci.*, **6**, 1438–1448.

Marquart,M., Walter,J., Deisenhoffer,J., Bode,W. and Huber,R. (1983) The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallogr.*, **B39**, 480–490.

Morris,A.C., Macarthur,M.W., Hutchinson,E.G. and Thornton,J. (1992) Stereochemical quality of protein structure coordinates. *Proteins: Struct. Funct. Genet.*, **12**, 345–364.

Narayana,S.V.L., Carson,M., El-Kabbani,O., Kilpatrick,J.M., Moore,D., Chen,X., Bugg,C.E., Volanakis,J.E. and DeLucas,L.J. (1994a) Structure of human factor D. A complement system protein at 2.0 Å resolution. *J. Mol. Biol.*, **235**, 695–708.

Narayana,S.V.L., Yamauchi,Y., Macon,K.J., Moore,D., DeLucas,L.J. and Volanakis,J.E. (1994b) Preliminary crystallographic studies on human complement profactor D. *J. Mol. Biol.*, **235**, 1144–1146.

Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.

Neurath,H. (1984) Evolution of proteolytic enzymes. *Science*, **224**, 350–357.

Nicolls,A., Sharp,K. and Honig,B. (1991) Protein folding and association: insights from the interfacial and thermodynamics properties of hydrocarbons. *Proteins: Struct. Funct. Genet.*, **11**, 281–296.

Otwinowski,Z. and Minor,W. (1997) Processing X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.

Perona,J.J. and Craik,C.S. (1995) Structural basis of substrate specificity in serine proteases. *Protein Sci.*, **4**, 337–360.

Pignol,D., Gaborioud,C., Michon,T., Kefelec,B., Chapus,C. and Fontecilla-Camps,J.C. (1994) Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a highly structured truncated zymogen E. *EMBO J.*, **13**, 1763–1771.

Read,R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr.*, **A42**, 140–149.

Renatus,M., Engh,R.A., Stubbs,M.T., Huber,R., Fischer,S., Kohnert,U. and Bode,W. (1997) Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray crystal structure of single-chain human tPA. *EMBO J.*, **16**, 4797–4805.

Sauter,N.K., Mau,T., Rader,S.D. and Agard,D.A. (1998) Structure of α-lytic protease complexed with its pro region. *Nature Struct. Biol.*, **5**, 945–950.

Sprang,S., Standing,T., Fletterick,R.J., Stroud,R.M., Finer-Moore,J., Xuong,W.H., Hamlin,R., Rutter,W.J. and Craik,C.S. (1987) The three-dimensional structure of Asn102 mutant of trypsin: role of Asp102 in serine protease catalysis. *Science*, **237**, 905–909.

Stroud,R.M., Kossiakoff,A.A. and Chambers,J.L. (1977) Mechanisms of zymogen activation. *Annu. Rev. Biophys. Bioeng.*, **6**, 177–193.

Tsukada,H. and Blow,D.M. (1985) Structure of α-chymotrypsin refined at 1.68 Å resolution. *J. Mol. Biol.*, **184**, 703–711.

Volanakis,J.E. and Narayana,S.V.L. (1996) Complement factor D, a novel serine protease. *Protein Sci.*, **5**, 553–564.

Volanakis,J.E. and Arlaud G.J. (1998) Complement enzymes. In Volanakis,J.E. and Frank,M.M. (eds.) *The Human Complement System in Health and Disease*. Marcel Dekker, Inc., New York, NY, pp. 49–81.

Vijayalakshmi,J., Padmanabhan,K.P., Mann,K.G. and Tulinsky,A. (1994) The isomorphous structures of prethrombin2, hirugen- and PPACK-thrombin: changes accompanying activation and exosite binding to thrombin. *Protein Sci.*, **3**, 2254–2271.

Walter,J., Steigemann,W., Singh,T.P., Bartunik,H., Bode,W. and Huber,R. (1982) On the disordered activation domain in trypsinogen: chemical labeling and low-temperature crystallography. *Acta Crystallogr.*, **B38**, 1462–1472.

Wang,D., Bode,W. and Huber,R. (1985) Bovine chymtrypsinogen A. X-ray structure analysis and refinement of a new crystal form at 1.8 Å resolution. *J. Mol. Biol.*, **185**, 595–624.

White,R.T., Damm,D., Hancock,N., Rosen,B.S., Lowell,B.B., Usher,P., Flier,J.S. and Spiegelman,B.M. (1992) Human adipsin is identical to complement factor D and is expressed at high levels in adipose tissue. *J. Biol. Chem.*, **267**, 9210–9213.

Yamauchi,Y., Stevens,J.W., Macon,K.J. and Volanakis,J.E. (1994) Recombinant and native zymogen forms of human complement factor D. *J. Immunol.*, **152**, 3645–3653.

Received October 12, 1998; revised December 7, 1998; accepted December 9, 1998