New structural motifs on the chymotrypsin fold and their potential roles in complement factor B

Hua Jing1,2,3, Yuanyuan Xu4, Mike Carson1, Dwight Moore1, Kevin J.Macon1,4, John E.Volanakis4,5 and Sthanam V.L.Narayana1,6

1Center for Macromolecular Crystallography, School of Optometry, 2Graduate Program in Biophysical Sciences and Pharmaceutical Design Program and 3Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA and 4Biomedical Sciences Research Center ‘A.Fleming’, Vari 16672, Greece
3Present address: Center for Blood Research, Harvard Medical School, Boston, MA 02115, USA
6Corresponding author
e-mail: narayana@pearl.cmc.uab.edu

Factor B and C2 are two central enzymes for complement activation. They are multidomain serine proteases and require cofactor binding for full expression of proteolytic activities. We present a 2.1 Å crystal structure of the serine protease domain of factor B. It shows a number of structural motifs novel to the chymotrypsin fold, which by sequence homology are probably present in C2 as well. These motifs distribute characteristically on the protein surface. Six loops surround the active site, four of which shape substrate-binding pockets. Three loops next to the oxyanion hole, which typically mediate zymogen activation, are much shorter or absent. Three insertions including the linker to the preceding domain bulge from the side opposite to the active site. The catalytic triad and non-specific substrate-binding site display active conformations, but the oxyanion hole displays a zymogen-like conformation. The bottom of the S1 pocket has a negative charge at residue 226 instead of the typical 189 position. These unique structural features may play different roles in domain–domain interaction, cofactor binding and substrate binding.

Keywords: complement system/domain structure/ factor B/protein–protein interaction/serine protease

Introduction

Serine proteases with a chymotrypsin-like fold (SP) are ubiquitous and involved in a wide range of biological processes. Structural and functional studies on this family of proteases are directed to answering three major questions: what determines their catalytic efficiency, how do they recognize substrates and how is their activity regulated? Extensive pioneering studies on digestive SPs such as chymotrypsin, trypsin and elastase have provided fundamental answers to these questions. The catalytic efficiency is provided by a His57–Asp102–Ser195 triad, located near the oxyanion hole (reviewed in Kraut, 1977). The substrate-binding specificity and affinity are determined mostly by the geometrical and chemical nature of the substrate-binding pockets, especially the S1 pocket (reviewed in Perona and Craik, 1995; Czapinska and Otlewski, 1999). Blocking these pockets with synthetic peptide analogs or natural protein inhibitors provided insights into the mechanisms of SP inactivation (reviewed in Bode and Huber, 1992). The activation of an SP is attributed in most cases to the zymogen activation where a freely liberated N-terminus induces conformational changes in and around the active site (reviewed in Huber and Bode, 1978; Khan and James, 1998). In some cases, however, SP activation also requires specific protein–protein interactions such as enzyme–cofactor association (Banner et al., 1996; Wang et al., 1998), enzyme–substrate interactions (Volanakis and Narayana, 1996; Jing et al., 1998) and enzyme self-assembly (Pereira et al., 1998). Structural studies of these novel SPs have begun to reveal the diversity of their regulation mechanisms.

The complement system, rich in such regulated SPs, contains eight proteolytic enzymes, all of which belong to the chymotrypsin fold group and cleave specific peptide bonds next to arginine residues (reviewed in Arlaud et al., 1998). Except for factor D, which has a single domain and interacts only with its natural substrate, all others have multiple domains and interact not only with the substrate, but also with certain cofactor proteins. Among these enzymes, factor B (FB) and C2 are homologous, and each supplies a catalytic subunit for the central C3/C5 convertases of complement. They also display structural and functional features distinct from those of the known SPs.

FB and C2 have three domains each: an N-terminal Ba or C2b domain composed of three complement control protein (CCP) modules, followed by a von Willebrand factor type A (vWFA) domain and a C-terminal SP domain. The presence of Mg2+, FB binds cofactor C3b to be activated by factor D, whereas C2 binds cofactor C4b and is activated by C1s (reviewed in Volanakis, 1988). The activation results in the release of their N-terminal domains and the generation of C3b-bound Bb or C4b-bound C2a, respectively, and the consequent protein complexes are named C3 convertases. Hence, the C-terminal SP domains of the catalytic subunits Bb and C2a lack the typical free N-terminus essential for the enzyme activation due to their linkage to the vWFA domains.

The esterolytic activities of FB/Bb and C2/C2a towards their best synthetic substrates are ~103-fold lower than that of trypsin (Kam et al., 1987), indicating a possible defect in the catalytic apparatus. No natural inhibitor has been identified for either enzyme. Instead, they are regulated by an association with the respective cofactor proteins. On the other hand, this bimolecular assembly
Table I. Statistics on diffraction data and refined structure

<table>
<thead>
<tr>
<th></th>
<th>Room temperature data</th>
<th>Low temperature data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Native TMLA Sm</td>
<td>Native TMLA</td>
</tr>
<tr>
<td>Diffraction data and phasing statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>51.88, 75.94, 74.04, β = 96.65</td>
<td>51.64, 74.26, 73.26, β = 95.55</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>2.3 2.8 2.8</td>
<td>2.1 2.3</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.7 96.0 91.0</td>
<td>99.2 97.5</td>
</tr>
<tr>
<td>Rsym (%)</td>
<td>7.3 6.0 4.1</td>
<td>6.9 8.1</td>
</tr>
<tr>
<td>Riso (%)</td>
<td>– 12.2 8.9</td>
<td>– 16.4</td>
</tr>
<tr>
<td>Phasing power (acentric/centric)</td>
<td>1.05/0.88 0.84/0.88</td>
<td>1.33/1.00 0.78/0.70</td>
</tr>
<tr>
<td>rRCullis</td>
<td></td>
<td>0.44/0.69</td>
</tr>
<tr>
<td>Figure-of-merit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refinement and structure statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20–2.1</td>
<td></td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>32 125</td>
<td></td>
</tr>
<tr>
<td>Rei</td>
<td>20.9</td>
<td></td>
</tr>
<tr>
<td>Rfree</td>
<td>24.3</td>
<td></td>
</tr>
<tr>
<td>No. of protein atoms</td>
<td>4508</td>
<td></td>
</tr>
<tr>
<td>No. of water atoms</td>
<td>517</td>
<td></td>
</tr>
<tr>
<td>Refined residues</td>
<td>A: A1k-V4, K9-K220a, Q220h-L250</td>
<td>B: S1b-K220a, V220h-L250</td>
</tr>
<tr>
<td>Invisible residues</td>
<td>A: W5-E6-H7-R8, N220b-Q220c-K220d-R220e-Q220f-K220g</td>
<td>B: A1k-D1j-P1i-D1h-E1g-S1f-Q1e-S1d-L1c, N220b-Q220c-K220d-R220e-Q220f-K220g-Q221h</td>
</tr>
<tr>
<td>R.m.s.d. bond length (Å)</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>R.m.s.d. bond angle (°)</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td>Luzzati error (work/test set) (Å)</td>
<td>0.26/0.30</td>
<td></td>
</tr>
<tr>
<td>No. of residues (A/B)</td>
<td>288/232</td>
<td></td>
</tr>
<tr>
<td>Average B-factor for protein (A/B)</td>
<td>31.99/24.04</td>
<td></td>
</tr>
<tr>
<td>Average B-factor for water</td>
<td>37.3</td>
<td></td>
</tr>
</tbody>
</table>

is unstable and exhibits a half-life of <2 min under physiological conditions (Kerr, 1980; Pangburn and Müller-Eberhard, 1986). The dissociated Bb displays residual hemolytic and proteolytic activity (Fishelson and Müller-Eberhard, 1984), albeit with higher esterolytic activity than FB (Kam et al., 1987). The cofactor interacts with the Mg$^2+$-binding motif on the vWFA domain in Bb/C2a (Horiuchi et al., 1991; Hourcade et al., 1995; Tuckwell et al., 1997) and possibly also with the SP domain (Lambris and Müller-Eberhard, 1984; Sanchez-Corral et al., 1990). Very little is known about the mode of cofactor binding and the position of the binding site(s) relative to the active site.

To understand the regulation mechanism and to facilitate the design of complement inhibitors, we determined the crystal structure of the SP domain of FB (FBSP). The structure reveals many motifs novel to the chymotrypsin fold. Sequence homology with C2SP suggests that similar motifs are likely to be present in C2SP as well, although some surface features might be different. The structure is presented below with an emphasis on the novel motifs. Some suggestions on their functional correlates, such as domain–domain interaction and cofactor binding, are presented.

**Results**

**Overall structure**

The FBSP structure was solved by a combination of multiple isomorphous replacement (MIR) and molecular replacement methods (Table I). All residues are in the most favored (85.6%) or additionally allowed (14.2%) regions of the Ramachandran plot except for one (K144) in the generously allowed region. The completeness of the two molecules in the asymmetric unit is slightly different for the N-terminal linker region and a surface loop (Table I). The rest of the two molecules superimpose well with a root-mean-square deviation of 0.34 Å for 277 matched (<3.8 Å) Cα atoms.

FBSP displays a two-domain β-barrel fold in which two β-sheets, composed of six β-strands each, are surrounded by surface helices and loops (Figure 1A and B). Compared with other SPs, the core β-sheets are similar, but many surface loops and helical regions are strikingly different. Long insertions and deletions are quite distinct in the structure-based sequence alignment with other SPs (Figure 1C). Interestingly, the insertions and deletions are clustered characteristically on the back, the right side and around the active site (refer to the orientation in Figure 1A). The active site also displays certain atypical conformations for some critical residues.

**Three insertions on the back**

The N-terminal linker region (residues D1h–T11 in chymotrypsinogen numbering), the 125 insertion (T125a–D133) and the C-terminal insertion (Q243–L250) are clustered on the back of the structure (Figure 1B). The N-terminal linker associates with the main body of the SP domain through a disulfide bond between C1 and C122. The N-terminal end of this linker displays a helical conformation in one molecule but is disordered in the other molecule; the C-terminal end displays a flexible loop conformation. The 125 insertion folds into a helix–loop–helix motif where the N-terminal ends of the two helices are constrained by a unique disulfide
bond (C125–C125p). The second helix is present at a similar position in a few other SPs, but the first helix and the middle loop regions are unique (Figure 1C). The C-terminal insertion is a loop whose conformation is stabilized by specific hydrophobic and electrostatic interactions involving its last four residues.

These three regions are tightly associated with each other and with the main body by a network of disulfide bonds, hydrophobic interactions and hydrogen bonds. A nearby short 203 loop (K203–R206) enhances such an association by complementing with salt bridge partners. The combined motif is obviously rigid and displays three distinct positively charged, negatively charged and hydrophobic patches (Figure 1B).

This motif may display a comparable conformation in C2SP, as indicated by similar residue compositions in the four involved regions (Figure 1C). Some differences, however, can be identified for a few exposed residues in the linker region (residues 3–9) and the C-terminal insertion (residues 245–246). Particularly, N5 and N9 in C2SP are potential N-linked glycosylation sites that are probably occupied (see Xu and Volaranakis, 1999), suggesting that this segment might be surface exposed in C2/C2a. Mutation of a glycine at position 2 to an arginine was described as the cause of C2 secretion deficiency (Westel et al., 1996), probably due to the disruption of the nearby C1–C122 disulfide bond.

**Three deletions on the right side**

Three loops that are present in all other SPs, the 16 loop (16–23), the 71 loop (71–81) and the 142 loop (142–155), are nearly absent in FBSP (Figure 1C). The 16 loop is completely absent, while loops 71 and 142 have only eight residues in total as compared with 24–29 residues observed in other SPs. Notably, seven of the eight residues are charged, but most of them do not form salt bridges. Thus, these three deletions create a unique hydrophilic surface to the right of the active site.

Similarly, C2SP has no 16 loop. Its 71 loop is longer than that in FBSP (seven residues versus three), its 142 loop is shorter (two residues versus five) and only two of the nine residues are charged. The 142 loop contains another potential N-linked glycosylation site at N142, which would make the deletion patch also hydrophilic in C2/C2a. Mutation of a glycine at position 2 to an arginine was described as the cause of C2 secretion deficiency (Westel et al., 1996), probably due to the disruption of the nearby C1–C122 disulfide bond.

**Six loops around the active site**

The active site is surrounded by loops 35, 60, 96, 172, 186 and 218, among which loops 96, 172 and 218 are exceptionally long (Figure 1). The 35 loop (V35–E40) and 60 loop (T60–S63) present on top of the substrate-binding cleft shape the upper wall of the S1’ and S2’ pockets with the side chains of T60, E40 and S41. The long 96 loop (I96–F98) and 172 loop (A172–V175) interdigitate with each other on the left side of the active site. The side chain of Y99, by stacking with that of W215, divides the space between the S2 and S3 pockets (Figure 2B). The S2 pocket is small and the S3 is large, both being hydrophobic in nature, and correlate well with the specificity towards alanine or glycine at the P2 position and leucine at the P3 position of the natural substrates. These two pockets appear to be rigid because of the extensive interactions between loops 96 and 172. This may explain their strict substrate-binding specificity and imply that they are unlikely to undergo any substrate-induced conformational changes. Similar conformations may be expected for the corresponding loops in C2SP, as indicated by their high sequence homology (Figure 1C).

The 186 loop (V186–D187) is located underneath the S1 pocket (Figure 1A). It interacts with a short 163–164c loop through a buried salt bridge between D187 and K163 (Figure 2A) and an aromatic stacking between Y186c and Y161, both of which are unique to FBSP. The 186 loop in C2SP carries somewhat different residues, but the salt bridge between D187 and K163 might be conserved (Figure 1C).

The 218 loop (V218–A224) is in front of the S1 pocket. Three exposed valine residues and a partially buried D218a in the beginning of this loop line the entrance to the S1 pocket (Figure 2B). C220 forms a disulfide bond with C191, which, due to insertions, is shifted by a few angstroms from its typical position (Figure 2A). Following C220, a stretch of approximately six unique charged or hydrophilic residues is invisible in the electron density maps (Table I). The corresponding region in C2SP is longer and has similar residue composition (Figure 1C), and thus may also display a highly flexible conformation. This atypically long and flexible loop, especially in proximity to the unique deletion patch and active site, probably contributes to the cofactor binding as discussed below.

**Within the active site**

The catalytic triad residues (H57, D102 and S195) and the non-specific substrate-binding site (W215–G216) display typical active conformations (Figure 2A). However, the oxyanion hole exhibits an atypical conformation due to an inward orientation of R192 carbonyl oxygen (Figure 2B). It forms a hydrogen bond with the amide group of S195, thus reducing the typical positive charge expected in the oxyanion hole. Residues 191–194 adopt a single turn 310-helix conformation, whose C-terminal end imposes additional negative potential to the oxyanion hole (Figure 2A). Such an inactive oxyanion hole conformation is similar to that observed in some SP zymogens (see Jing et al., 1999) and could possibly explain the low esterolytic activities of FB/Bb (Kam et al., 1987). A conformational...
change to an active oxyanion hole is obviously essential for efficient catalysis.

In the absence of the positively charged N-terminus of the SP domain, D194 forms a hydrogen bond with the hydroxyl group of S140 (Figure 2A). In C2SP, S140 is absent, but a unique R43 residue might be close enough
to form a salt bridge with E194. Other residues within the oxyanion hole, the catalytic triad and the non-specific substrate-binding site in C2SP are nearly identical to those in FBSP.

At the bottom of the S1 pocket, residue D189, typically positioning the positively charged P1 residue of the substrate for the nucleophile attack by S195, is replaced by N189 in FBSP and S189 in C2SP. However, a typical G226, located on the opposite side to D189 at the bottom of the S1 pocket, is replaced by D226 in both enzymes (Figure 1C). Compared with the typical position of the D189 carboxyl group, the D226 carboxyl group in FBSP is more elevated (Figure 2A). Its negative charge might be more dispersed due to hydrogen bonding with the atoms from N189, T190, R225 and a non-crystallographic symmetry (NCS)-related K37. This conformation is similar to that observed in a trypsin D189G/G226D mutant, which is also less active than the wild-type trypsin (Perona et al., 1993). However, relocating the negative charge from 226 to the typical 189 position in FB mutant D226N/N189D did not restore the catalytic efficiency closer to that of trypsin, as shown in our recent mutagenesis experiments (Xu et al., 2000). On the contrary, this mutant exhibits complete loss of hemolytic and C3 cleavage activity and ~50% reduced activity in forming C3 convertase (Figure 3). The esterolytic activity of the mutant is also substantially reduced compared with the wild-type, mainly due to a reduction in the \(k_{cat}\) value. These data indicate that residue D226 is the primary substrate determinant for P1-Arg binding, and a proper registration of the long P1-Arg side chain to D226 on one end and its scissile bond to the nucleophilic S195 on the other end is essential for efficient catalysis (Xu et al., 2000).

The S1, S2 and S3 pockets are occupied coincidentally by the side chains of K37, S36c and P36b, respectively, from an NCS-related 35 loop in a substrate-like manner (Figure 2B). The amino group of K37 forms four hydrogen bonds with oxygen atoms from residues D226, T190 and V218. The carbonyl oxygen of P36b forms a hydrogen bond with the backbone amine of G216. This crystal packing artifact results in a shift of the loop 35 tip from its typical position, but it is unlikely to affect the S1 site conformation because each structural element in the active site region is engaged in a network of specific intra-molecular interactions.

The active site of FBSP displays an overall negative potential, which may provide the driving force to attract and position the natural substrates (Figure 1A). The presence of long insertion loops immediately around the active site may limit the accessibility of the FB active site to non-specific macromolecular substrates and natural SP inhibitors in blood.

**Discussion**

The crystal structure of FBSP reveals the presence of a unique bulky and rigid motif on the back, three distinct deletions on the right and six loop insertions around the active site. As described, four of the six loops (loops 35, 60, 96 and 172) dictate the specificity at the S1’, S2’, S2 and S3 pockets. The atypical D226 residue in the S1 pocket determines P1-Arg binding and catalysis (Xu et al., 2000). The remaining structural motifs, which have no corresponding regions in all other SPs with known structures, raise intriguing questions regarding their functional roles. Given that FB/Bb and C2/C2a are multidomain SPs and require cofactor binding to express proteolytic activity fully, it seems possible that these unique regions could be involved in the domain–domain interactions, cofactor binding and substrate binding.

**Domain architecture**

Several biophysical methods have shown that the three domains in FB constitute a compact and globular structure and that the two domains in Bb are in close contact (Smith et al., 1982, 1984; Ueda et al., 1987; Chamberlain et al., 1998; Hinshelwood and Perkins, 1998). The linker between the vWFA domain and the SP domain is inaccessible to limited proteolysis by elastase in the context of Bb, but is susceptible in free FB (Lambris and Müller-Eberhard, 1984), suggesting a possible rearrangement of the two domains from FB to Bb.

To identify the possible domain arrangements, we performed a docking search using the FBSP crystal structure and an FB vWFA model built by Tuckwell et al. (1997), where the C-terminal region of the vWFA domain was restrained together with the N-terminal region of the SP domain. Although there were no clear-cut solutions, most of the best solutions placed the vWFA domain at the right of the back of the SP domain (see an example in Figure 2C). The Mg\(^{2+}\) binding site on top of the vWFA domain and the N-terminus of the vWFA domain are exposed from the domain interface. The residues in the contact regions come from bottom loop regions of the vWFA domain and from the unique insertion motif on the back of the SP domain, suggesting that the latter probably plays a role in the interactions with the vWFA domain.

However, the exact orientation is hard to determine because the linker appears to be flexible in this structure. The N-terminal end of the SP domain could be traced to the end only for one molecule in the asymmetric unit (Table I). The C-terminal helix of the vWFA domain may also be flexible, because in isolated integrin Mac-1 I-domain (also a vWFA domain) this helix shifts in response to ligand binding (Lee et al., 1995; Liddington and Bankston, 1998). Such flexibility of the linker may reflect a structural basis for the domain rearrangement mentioned above. Alternatively, it could be simply a crystal packing artifact in the isolated SP domain. On the other hand, because of the presence of potential carbohydrates at N5, N9 and N142 in C2SP, the deletion patch is unlikely to be buried in the domain interfaces. The location of the Ba
The serine protease domain of complement factor B

The serine protease domain of complement factor B is uncertain, but it should be in contact with both the vWFA and SP domains (Chamberlain et al., 1998; Hinshelwood and Perkins, 1998).

In other multidomain SPs, the three most common domains preceding the SP domain are the EGF domain, the Kringle domain and the CCP module (Table II).
However, only a few multidomain SP structures containing the EGF domains currently are available, e.g. factor Xa, factor IXa, factor VIIa and protein C (reviewed in Bode et al., 1990). Superposition of these structures onto FBSP showed that a similar C1–C12 disulfide bond is utilized to link the SP domain with its preceding domain. Similarly, the regions on the back of these SPs also mediate the domain interface. However, the FBSP- and C2SP-unique helix (E125b–A125g) in the motif on the back overlaps significantly with the core of the proximal EGF domain, suggesting that the vWFA–SP interface would be different from the EGF–SP interface.

This attachment on the back of other types of domains could be a general feature in the multidomain SPs. Such a domain arrangement would allow the other domains to facilitate the regulation of the SP domain through binding various cofactors without hindering the accessibility of the active site in the front.

**Cofactor binding and possible activation mechanisms**

Various biochemical experiments have demonstrated the presence of cofactor-binding sites in each of the three domains (reviewed in Volanakis, 1988). The binding site on the vWFA domain has been located by mutagenesis experiments to be around the Mg2+ -binding motif of FB and C2 (Horiuchi et al., 1991; Hourcade et al., 1995; Tuckwell et al., 1997). The binding site on the SP domain of FB was also suggested for a 33 kDa fragment corresponding to FBSP based on its C3b-binding activity (Lambris and Müller-Eberhard, 1984) and limited C3 cleavage activity that is independent of Mg2+ but dependent on C3b (Sanchez-Corral et al., 1990). This fragment behaves like Bb in that they have similar proteolytic activity and their cofactor-assisted proteolysis can be inhibited similarly by factor H, whereas, in the presence of Mg2+, Bb has enhanced proteolytic activity, suggesting that the vWFA domain assists C3b binding (Sanchez-Corral et al., 1990). Additionally, some mutations at residues D226, D187 and N189 in the SP domain were shown to reduce the binding of FB and Bb to a C3b analog, cobra venom factor (Xu et al., 2000; Figure 3). However, the presence of the cofactor-binding sites on both domains is not supported by an early electron microscopy visualization of the bimolecular complexes, suggesting that in most populations only one blob attached the cofactor protein (Smith et al., 1984). Given that the half-life of the bimolecular complexes is ~2 min in solution (Kerr, 1980; Pangburn and Müller-Eberhard, 1986), the interaction between the SP domain and the cofactor protein can not be ruled out.

From the structure of FBSP, the cofactor-binding site on the SP domain can be postulated based on the unique regions where FBSP and C2SP differ, because FB/Bb binds C3b but C2/C2a binds C4b. Two regions stand out from such comparisons. One region is the deletion patch, which in FBSP displays mostly charged residues, but in C2 displays neutral and hydrophilic residues. This patch is bordered by the C-terminal end of the linker region on the right and loops 35, 186 and 218 on the left (Figure 2C), all of which are unique to but different between FBSP and C2SP (Figure 1C). The deletions or shortening of loops 16, 71 and 142, which typically mediate zymogen activation, appear to leave a space for something to access and activate the oxyanion hole. This flat and hydrophilic surface also fits the predominant features suitable for weak hetero protein–protein interactions (Jones and Thornton, 1996) and seems capable of accommodating a large molecule such as C3b or C4b (~200 kDa). However, the potential presence of a glycan on the 142 loop of C2SP would alter its surface features, which may or may not exclude C4b binding.

Another region where FBSP and C2SP differ is underneath or in front of the S1 pocket, near the 164 insertion, 186 loop and 218 loop (Figure 2C). This region features
Thus also playing a substantial role in activation, which is of the SP domain in both cases is free and tucked inside, the element responsible for SP regulation can be as small and understanding of SP regulation mechanisms. The structural complex (Banner 164–170b) as in the case of the factor VIIa–tissue factor sites.

These possibilities and to pinpoint the cofactor-binding assists in the binding of cofactor and the substrates while the other site below or in front of the S1 pocket generated by either C3b-bound Bb or C4b-bound C2a, is could be activated by the natural substrate, C3. C3b, from either the right side (loops 16, 71 and 142), as in the case of the factor VIIa–tissue factor (Ikari et al., 1998), or from the bottom (residues 125–134 and 164–170b) as in the case of the factor VIIa–tissue factor complex (Banner et al., 1996). However, the N-terminus of the SP domain in both cases is free and tucked inside, thus also playing a substantial role in activation, which is not possible in Bb/C2a.

Taken together, the activation mechanism of complement FB and C2 further broadens our current understanding of SP regulation mechanisms. The structural element responsible for SP regulation can be as small and efficient as a single residue insertion, such as Arg218 in complement factor D (Volanakis and Narayana, 1996; Jing et al., 1998), or as large and elaborate as the requirements of other domains and cofactors, such as the EGF domains and tissue factor for factor VIIa (Banner et al., 1996) or the vWFA domain and C3b/C4b cofactor for Bb/C2a. More diversity may be expected for other regulated SPs, depending on the specific control needed.

Materials and methods

Protein expression, purification and crystallization

The cDNA sequence coding for FBSP was amplified by PCR using the FB cDNA clone BHL4-1 as template. The PCR product was restricted by BamHI and EcoRI and subcloned into a baculovirus transfer vector, pAcGP67-A (PharMingen, San Diego, CA). Transfection and expression of FBSP in baculovirus-infected S9 cells were carried out according to the manufacturer’s protocols. The recombinant FBSP consists of a vector-derived tripeptide Ala–Asp–Pro at the N-terminus and 295 C-terminal amino acid residues of FB. The protein was purified by ion exchange chromatography using a CM–Sepharose column followed by a Mono S column (Amersham Pharmac ia Biotech, Piscataway, NJ). The starting buffer for the CM–Sepharose column is 0.1 M sodium acetate, 0.02 M EDTA (pH 6.3), and for the Mono S column is 0.05 M phosphate (pH 6.8). The elution buffer contains a linear gradient of the starting buffer and 1.0 M NaCl. Purified protein was dialyzed against 20 mM Tris, 100 mM NaCl (pH 7.5) and concentrated to 20 mg/ml. Crystals were obtained at room temperature by the hanging drop vapor diffusion method. The reservoir solution contained 35% PEG-1000 and 30 mM MES (pH 6.0), and the drops contained an equal volume of reservoir solution and protein solution.

Data collection and phasing

Two native data sets were collected on in-house X-ray generators (40 kV, 100 mA), one at room temperature on a RAXIS-II detector, and the other one at 95 K on a DIP detector using the mother liquor as cryoprotectant, both processed using the DENZO and SCALEPACK programs (Otwinowski and Minor, 1997). The two data sets were non-isomorphous as the b-axis differed by 1.7 Å (Table I). The crystals belonged to space group P2₁ with two molecules in the asymmetric unit. While searching for molecular replacement solutions, we also collected heavy atom derivative data at both room and low temperatures (Table I). The positions of the heavy atoms were identified and refined using the XTAListView (McRee, 1992) and MLPHARE (in CCP4, 1994) programs. Later, molecular replacement solutions were found in the AmoRe program (Navaza, 1994) using an SP model containing only the core β-sheets. The phases calculated using these solutions confirmed the heavy atom sites on the difference Fourier maps and brought the sites from different derivatives to one relative origin. Next, 4-fold multiple crystal averaging was performed using the DMMulti program (CCP4, 1994), which not only combined the phases from independent sources but also extended the phases to the highest resolution in the native data (reviewed in Kleywegt and Read, 1997). This averaged map helped to interpret the positions of ~90% of the residues.

Structure refinement

The structure was refined using the CNS program with its maximum likelihood-targeted torsion angle dynamics, positional refinement and B-factor refinement protocols (Brünger et al., 1998). Except for 10% reflections excluded for Rfree calculation (Brünger, 1992; Kleywegt and Brünger, 1996), all reflections from 20.0 to 2.1 Å were used in the refinement. Sigma-weighted electron density maps (Read, 1986) were calculated using CNS and guided manual rebuilding on O (Jones et al., 1991). Structure quality (Table I) was checked in each cycle using OOPS (Kleywegt and Jones, 1996, 1997), and later with CNS, PROCHECK (Laskowski et al., 1993) and WHAT_CHECK (Hooft et al., 1998).
structure and diffraction data are deposited in the PDB with the accession code 1DLE.

Acknowledgements

The authors are grateful to the staff at the Center for Macromolecular Crystallography (CMC) for their enormous support in using the state of the art X-ray crystallography and computational facilities. H.J. deeply appreciates Yu Luo, Todd Bice and Champion Deivayamong (CMC) for discussions on the MIR method; Kevin Cowtan (York University, UK) for suggestions on multiple crystal averaging; Danny Tuckwell (University of Manchester, UK) for the model of the FB vWFα domain; Xuejun Zhang (Oklahoma Medical Research Foundation) for Cα coordinates of the plasmin–streptokinase complex before release; and Lei Jin and Timothy Springer (Harvard University) for generous support, help in computer usage and enlightening discussions. The project is funded by NIH grants AI 39818 to S.V.L.N. and AR 44506 to J.E.V.

References


Received August 9, 1999; revised and accepted November 3, 1999