Crystal Structure of Anticoagulant Thrombin Variant E217K Provides Insights into Thrombin Allostery*<sup>§</sup>

Wendy J. Carter†, Timothy Myles§, Craig S. Gibbs§, Lawrence L. Leung§, and James A. Huntington¶

From the †University of Cambridge, Department of Haematology, Division of Structural Medicine, Thrombosis Research Unit, Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 2XY, United Kingdom, the §Department of Medicine, Division of Hematology, Stanford University School of Medicine, Stanford, California 94305, and *Gilead Sciences, Foster City, California 94404

Thrombin is the ultimate protease of the blood clotting cascade and plays a major role in its own regulation. The ability of thrombin to exhibit both pro- and anti-coagulant properties has spawned efforts to turn thrombin into an anticoagulant for therapeutic purposes. This quest culminated in the identification of the E217K variant through scanning and saturation mutagenesis. The antithrombotic properties of E217K thrombin are derived from its inability to convert fibrinogen to a fibrin clot while maintaining its thrombo-modulin-dependent ability to activate the anticoagulant protein C pathway. Here we describe the 2.5-Å crystal structure of human E217K thrombin, which displays a dramatic restructuring of the geometry of the active site. Of particular interest is the repositioning of Glu-192, which hydrogen bonds to the catalytic Ser-195 and which results in the complete occlusion of the active site and the destruction of the oxyanion hole. Substrate binding pockets are further blocked by residues previously implicated in thrombin allostery. We have concluded that the E217K mutation causes the allosteric inactivation of thrombin by destabilizing the Na<sup>+</sup> binding site and that the structure thus may represent the Na<sup>+</sup>-free, catalytically inert “slow” form.

Thrombin activity is central to hemostasis, the balance between thrombosis and bleeding (for review, see Refs. 1, 2), and consequently thrombin is an important target of anticoagulant therapies (for review, see Ref. 3). Thrombin is generated from its zymogen form, prothrombin, at the end of the coagulation cascade and is eventually inhibited by the circulating serpin antithrombin (for review, see Ref. 4). Thrombin has many procoagulant properties, including the cleavage of fibrinogen to fibrin, which then polymerizes to form the fibrin clot. Thrombin is also responsible for activating the transglutaminase factor XIII, which stabilizes the clot by cross-linking the fibrin polymers. Thrombin activates platelets by cleaving protease-activated receptors and stimulates its own generation by activating cofactors V and VIII. Hemostasis is dependent on limiting procoagulant activity to surfaces of the vasculature that have been compromised. Thus, when thrombin leaches away from the site of tissue damage its activity is reversed from pro- to anti-coagulant by binding to the integral membrane protein, thrombomodulin (TM),<sup>1</sup> expressed at the surface of the intact endothelium (5). Once bound to TM, thrombin can no longer cleave fibrinogen and instead cleaves protein C to yield activated protein C. Activated protein C then dampens thrombin generation by cleavage inactivation of cofactors Va and VIIIa (for review, see Ref. 6).

Thrombin interacts with many cofactors capable of inducing conformational change and altered protease activity; however, the physiological significance of thrombin allostery is unclear (7–10). The most relevant alteration of thrombin activity is caused by its binding to TM, but similar changes can also be induced by other cofactors. Of particular interest is the role of Na<sup>+</sup> binding in modulating the activity of thrombin (11). Thrombin binds Na<sup>+</sup> through octahedral coordination to the main chain oxygens of Arg-221a and Lys-224 and four water molecules (12, 13) with a Keq in the range of the physiological Na<sup>+</sup> concentration (11). It has been demonstrated that an equilibrium exists between a Na<sup>+</sup>-free “slow” form and a Na<sup>+</sup>-bound “fast” form, where the slow form cleaves substrates with reduced efficiency because of an increased Keq and a decreased kcat.

Although the term slow is commonly used to describe the properties of the ligand-free form, it is more likely a catalytically incompetent state with apparent catalysis by the slow form reflecting a rapid equilibrium between inactive and active forms or an induced-fit substrate binding mechanism (14). Consistent with such a mechanism is the conserved conformational change induced by either Na<sup>+</sup> binding, substrate interaction, or TM binding (15). The conformational change is characterized by an opening of the active site cleft (11), and several residues have been implicated, including Trp-60d and Trp-215. Mutations in the Na<sup>+</sup>-binding region have also been shown to perturb the equilibrium in favor of the slow conformation (16). Defining the structural basis of this equilibrium would help resolve the role of thrombin allostery in hemostasis; however, to date, no structure has been solved that satisfies all of the properties of the slow form (10, 17). The issue of thrombin allostery is also of potential importance for the development of novel therapies, because trapping thrombin in the slow state would allow its administration as an anti-

* This work was supported by National Institutes of Health Grants R01 HL68829 (to J. A. H) and R01 HL57530 (to L. L. L.) and by the Medical Research Council, UK. 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.

The atomic coordinates and structure factors (code 1BD0) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ The on-line version of this article (available at http://www.jbc.org) contains a supplementary video.

† To whom correspondence should be addressed. Tel.: 44-1223-763230; Fax: 44-1223-336827; E-mail: jah52@cam.ac.uk.

¶ The abbreviations used are: TM, thrombomodulin; r.m.s.d., root mean square deviation.
The E217K variant was more effective as an anticoagulant which was protein C system (19). Subsequent saturation mutagenesis ultant in an animal model through the stimulation of the B-chain numbering), which proved effective as an anticoag- in the discovery of the variant E217A (E229A in the thrombin variant). A scanning mutagenesis approach was under- in 50 mM Tris, pH 7.4, and 25% glycerol. Preliminary crystals were taken to identify a recombinant anticoagulant form, resulting 

EXPERIMENTAL PROCEDURES

The E217K variant of human thrombin was produced as before (20). Crystal trials were established after concentrating thrombin to 6 mg/ml in 50 mM Tris, pH 7.4, and 25% glycerol. Preliminary crystals were obtained in Hampton Crystal Screen Cryo 42 (40 mM mono-potassium phosphate, 16% polyethylene glycol 8000, and 20% glycerol). Diffraction quality crystals were obtained from a 4:1 ratio of solution 42 and 20% glycerol. Data were obtained from a single flash-cooled crystal (100 K) at the Daresbury Synchrotron Radiation Source (UK) station 14.2 and were processed using Mosflm, Scala, and Truncate (23). The structure was solved by molecular replacement using the program MolRep (24) with S195A thrombin (1JOU) as the search model, and two molecules were placed in the asymmetric unit. After rigid body refinement, strict non-crystallographic symmetry (NCS) was applied for the first round of refinement, followed by restrained NCS for one round of refinement. NCS restraints were not used in further rounds of refinement. All refinement was conducted using the program CNS (25) (version 1.0), and the program XtaView (25) was employed for rebuilding. Data processing and refinement statistics are given in Table I. Figures were made using Bobscript (26), Raster3D (27), and Spock. Template num- 

### RESULTS

**Overall Structure**—Because thrombin is capable of autolysis (28), most crystal structures of thrombin are obtained from active site-inhibited forms. We set up crystals of uninhibited E217K because the reduced activity of the variant made autolysis unlikely. Thus, the two monomers that constitute the asymmetric unit have a free active site and are more likely to reflect a true solution conformation. The two monomers are nearly identical with a r.m.s.d. of 0.351 for the 258 equivalent carbons (Cα) of the heavy chain. All further structural analysis is based on the structure of the first monomer, AB (A denotes the light chain and B the heavy chain) because it does not differ in any significant way from the second monomer, CD; any comment made concerning the structure of the variant E217K applies equally to both molecules of the asymmetric unit.

To compare the structure of E217K with other thrombin structures, it is necessary to choose the most relevant of the nearly 150 deposited coordinates. Because the E217K variant has an unoccupied active site, we compared its structure with other active site free forms (1JOU, Na⁺-bound monomer EF (10); 1MH0, Na⁺-free (17); and 1HAB, bound to C-terminal hirudin peptide (29)). For reference, the active site-occupied forms 1JOU (monomer AB) and 1PPB (the original thrombin
structure covalently bound to PPACK (30) were also used. The Ca r.m.s.d. for all were ~1.3 Å when the flexible γ-loop is excluded from the comparison, indicating no radical restructuring of the backbone in response to the point mutation at residue 217. There were, however, significant differences in the main chain conformation at certain positions when comparing the reference structures with that of the E217K variant. Fig. 1a is a Ca trace of E217K colored according to r.m.s.d. with thrombin in its normal, fast conformation (1JOU monomer AB) reveals significant conformational changes. The color ramp is from 0–3 Å, from gray to red. Thrombin is in the standard orientation, with disulfide bonds in yellow, significant regions labeled, and the active site Ser-195 represented as a green ball. b, stereo representation of the Ca trace of E217K thrombin (in the same orientation as above), colored from N to C terminus (blue to red), is superimposed on the structure of active thrombin (1PPB, gray). The position of the mutation (217) is shown by a ball in each structure.

Fig. 1. The structure of E217K thrombin differs significantly from that of wild-type thrombin. a, stereo representation of the Ca trace of E217K thrombin colored according to r.m.s.d. with thrombin in its normal, fast conformation (1JOU monomer AB) reveals significant conformational changes. The color ramp is from 0–3 Å, from gray to red. Thrombin is in the standard orientation, with disulfide bonds in yellow, significant regions labeled, and the active site Ser-195 represented as a green ball. b, stereo representation of the Ca trace of E217K thrombin (in the same orientation as above), colored from N to C terminus (blue to red), is superimposed on the structure of active thrombin (1PPB, gray). The position of the mutation (217) is shown by a ball in each structure.

structure covalently bound to PPACK (30) were also used. The Ca r.m.s.d. for all were ~1.3 Å when the flexible γ-loop is excluded from the comparison, indicating no radical restructuring of the backbone in response to the point mutation at residue 217. There were, however, significant differences in the main chain conformation at certain positions when comparing the reference structures with that of the E217K variant. Fig. 1a is a Ca trace of E217K colored according to r.m.s.d. with thrombin in its normal, fast conformation (1JOU monomer AB) reveals significant conformational changes. The color ramp is from 0–3 Å, from gray to red. Thrombin is in the standard orientation, with disulfide bonds in yellow, significant regions labeled, and the active site Ser-195 represented as a green ball. b, stereo representation of the Ca trace of E217K thrombin (in the same orientation as above), colored from N to C terminus (blue to red), is superimposed on the structure of active thrombin (1PPB, gray). The position of the mutation (217) is shown by a ball in each structure.

The Active Site—The catalytic efficiency of E217K thrombin is significantly reduced toward fibrinogen, chromogenic peptide S2238, and antithrombin, with second-order rate constants 270, 250, and 22 times lower than with wild-type (20). This is due both to an increase in $K_m$ and a decrease in $k_{cat}$. It is clear from the structure of the E217K variant what factors lead to its reduced catalytic activity. Electron density for the active site is given in Fig. 2a and unequivocally demonstrates a high degree of certainty in both side and main chain position. The hydrogen bonding in the active site of the E217K variant is highly un-
usual and results in the steric blocking of substrate binding sites and a catalytically incompetent active site architecture (Fig. 2b). Precise orientation of the catalytic triad, Asp-102, His-57, and Ser-195, is required for the nucleophilic attack of the Ser-195 O\textsubscript{y} on the carbonyl carbon of the scissile bond. The resultant tetrahedral transition state is stabilized by the oxyanion hole formed by the amide hydrogens of Gly-193 and Ser-195. In the structure of E217K thrombin, Glu-192 has shifted position to make a side chain and a main chain hydrogen bond with the H\textgamma of Ser-195 and a main chain hydrogen bond with the amide hydrogen of Ser-195. The oxyanion hole is further perturbed by the flipping of Gly-193 in response to the movement of the preceding Glu-192. The dramatic movement of Glu-192 in the structure also results in the complete burying of Ser-195 (see Fig. 3). Thus, the altered hydrogen bonding pattern in the active site of the E217K thrombin variant renders it effectively non-catalytic by blocking substrate access to the active site and by perturbing the precise geometry required for catalysis.

**Substrate Binding Pockets**—In addition to the predicted effect of the altered active site hydrogen bonding on the \( k_{\text{cat}} \) of E217K thrombin, the structure in the regions surrounding the catalytic residues will also predictably lead to a significant increase in \( K_m \). Physiological thrombin specificity is often determined by exosite interactions, but for isolated peptide sub-
strates and inhibitors specificity is determined primarily by the P1, P2, and P4 positions (31) (substrate nomenclature of Schechter and Berger (32) where cleavage occurs between P1 and P1’). This corresponds to binding pockets S1, S2, and the aryl binding site (or S4) on thrombin. The global effect of the structural rearrangement for the E217K variant is the constriction of the active site cleft of thrombin. Fig. 3 is a surface representation of thrombin in the classic “Bode” orientation with the active site facing and substrate running from left to right for N to C terminus. The surface is colored according to hydrophobicity with the Oy of Ser-195 colored red. Wild-type thrombin (Fig. 3a) has an open active site cleft with an exposed catalytic serine Oy, whereas the E217K mutant has a closed active site and a completely buried Ser-195 Oy. The structure of the Michaelis complex between thrombin and the serpin heparin cofactor II (HCII) provides the best model for the thrombin-substrate interaction (33). When the structure of allosterically activated thrombin is superimposed on that of thrombin in complex with HCII, the reactive center loop is easily accommodated from P4 to P4’, as shown in Fig. 3a. However, when the same is done with thrombin variant E217K, the substrate loop of HCII cannot be accommodated (Fig. 3b). Major clashes are observed between the P4-P1 region of the substrate with main and side chain atoms of residues in the active site cleft: P1 Leu is blocked from the S1 pocket by Cys-191, Glu-192, and Gly-216, and the effect would be more severe with the normal Arg in the P1 position; consensus P2 Pro clashes with Trp-60d and Glu-192; P3 Met clashes with Trp-215 and Asp-221; and consensus P4 Phe overlaps with the side chain of Trp-215.

Fig. 3. Surface representations of thrombin reveal the occlusion of the active site cleft caused by the E217K mutation. a, the surface of active thrombin (1HAH, colored green for hydrophobicity) in the standard orientation demonstrates the open active site cleft of thrombin and the accessibility of the catalytic Oy of Ser-195 (red). This conformation represents an allosterically activated thrombin with nothing bound in the active site cleft. However, the structure can easily accommodate a natural thrombin substrate derived from the reactive center loop of heparin cofactor II (P4- P4’, shown as rods). b, in contrast, the active site cleft of E217K is in a closed conformation, and although the catalytic Oy is colored as above, it is fully blocked by the conformational changes in the active site. In addition, it is clear that overlaps and steric clashes from P4 to P1 would prevent substrate binding to E217K. A video depiction of the structural transition from a closed (E217K) to an open (1JOU, monomer AB) active site is given as supplementary material.
FIG. 4. Conformational changes in the Na\(^+\) binding site and the allosteric switch. a, the sodium binding sites (residues 217–225) for E217K (yellow), active Na\(^+\)-bound state (1JOU, monomer AB, magenta), and Na\(^+\)-free 1MH0 (transparent) with residues labeled in the color of the corresponding structure. Na\(^+\) (purple ball) coordinates to wild-type thrombin via main chain oxygens of Arg-221 and Lys-224 and four conserved water molecules (red balls). This coordination is dependent on the salt bridge between Lys-224 and Glu-217 (broken rod). The conformation of this loop is dramatically altered by the substitution of Glu-217 for Lys (E217K variant in yellow) so that monovalent cations can no longer be.

b.
The discovery of the E217K mutation was the result of an initial alanine scanning mutagenesis of surface-exposed thrombin residues, subsequently followed by saturation mutagenesis of the hits at Trp-60d, Lys-60f, Glu-217, and Arg-221a (19, 20). The substitution of lysine for glutamate at position 217 resulted in the most dramatic shift in specificity with a 270-fold reduction in the efficiency of cleavage of the fibrinogen Aα chain and only a 2-fold reduction in efficiency of TM-catalyzed protein C activation. These results were initially explained sterically because Glu-217 of wild-type thrombin and the P5 Gly of fibrinopeptide A are in van der Waals contact in the structure of the complex (34). The substitution for lysine would predictably block the P5 binding site for fibrinogen but not for protein C, which is thought to bind in an extended manner. This hypothesis, however, does not explain why the alanine variant has properties similar to the E217K variant nor does it agree with the structural results described here.

It is possible to re-evaluate the initial hits from alanine scanning mutagenesis in light of the mounting body of evidence that thrombin is an allosteric enzyme, existing in an equilibrium between a catalytically competent fast form and an inactive, so-called slow form. Na+ binding alters the equilibrium position in favor of the fast form and is coordinated to the main chain oxygens of Arg-221a and Lys-224. Thus, two of the hits are related to the Na+ binding site: Arg-221a directly binds Na+, and its substitution to alanine results in properties similar to the wild-type slow form (35), and Glu-217 forms a salt bridge with Lys-224. The involvement of the 60-loop (in particular, Trp-60d) in the equilibrium has also been established through mutagenesis and structural studies (36). It is thus likely that the E217K and the E217A mutations alter the equilibrium between the slow and fast forms by a mechanism that is conserved for the two mutations, with the effect of the lysine substitution being more dramatic.

How then does the substitution E217K result in the conformation observed in the crystal structure, and how can an alanine produce a similar result? The mutations must either destabilize the fast conformation, stabilize the slow conformation, or do both. It is possible to discount stabilization of the slow conformation because an alanine side chain cannot participate in any interactions that can be shared by a lysine. In addition, Lys-217 in the structure of the E217K variant is not likely to be sufficient to alter the equilibrium position, as seen with the E217A variant, a mutation that engendered even more electrostatic repulsion would have an even greater effect, as seen for E217K.

Therefore it is clear from the structure of the active site why E217K is effectively a dead enzyme and also how substitutions at Glu-217 can perturb the pre-existing equilibrium between the slow and fast forms. What then does the structure reveal about the equilibrium in wild-type thrombin? Because the lysine side chain of E217K does not make significant stabilizing contacts in the structure and the effect of the mutation appears to be limited to destabilizing the Na+-bound, fast conformation, we conclude that the E217K thrombin structure presented here may represent the actual Na+-free, slow form. How then does the structure of E217K thrombin compare with the two other structures claiming to represent the slow form? The structure by Di Cera and co-workers (17), although grown in conditions devoid of Na+, possesses a fully formed Na+ binding site and no significant conformational changes (Fig. 4a). Because the structural reorganization of the Na+ binding site seen in E217K thrombin is required for the disruption of the catalytic architecture and for some of the observed blocking of substrate binding pockets, the structure of Di Cera is unlikely to represent the slow form. It must be concluded that crystal contacts are capable of trapping out the fast conformation even in the absence of Na+. The converse of this has been observed when no conformational change took place after Na+ was soaked out of crystals formed in its presence (12). The other structure of slow thrombin (10) also has a major drawback in that Na+ was included in the crystallization conditions. Because of this, the active site free form has Na+ coordinated in the normal fashion and has thus not released the loop containing Glu-217. There were, however, significant changes in the positions of several side chains associated with the Glu-217 loop, in particular the disulfide bond between 168 and 182, Trp-215 and Phe-227. These side chains adopted a previously unobserved conformation that resulted in the blocking of the substrate binding pockets. These residues were proposed to constitute an allosteric switch (including Trp-60d from the 60-insertion loop) and the concerted movement of the switch would result in the opening of the active site cleft upon Na+ or other cofactor binding. The side chains of the allosteric switch are shown in Fig. 4b for the E217K thrombin variant with electron density and for equivalent residues from wild-type fast thrombin (1PPB) and the proposed slow form from Huntington and Esmon (10). The position of the side chains involved is unequivocal from the electron density and closely corresponds to that of the structure of Huntington and Esmon. The only exception is the position of Trp-60d, which is in an intermediate position, probably because of crystal contacts. In conclusion, the struc-
ture of Di Cera and co-workers (17) shares none of the structural characteristics of E217K and that of Huntington and Esmon (10) shares the otherwise unique feature of an allosteric switch in a position that would preclude substrate binding in the S2 and aryl binding pockets.

One of the surprising observations concerning the E217K variant is that it is a dead enzyme toward fibrinogen but its reactivity toward protein C is almost unaltered in the presence of TM. In light of the structural explanation for its catalytic inactivity given here, the effect of TM is even more perplexing. Presumably, the conformational changes seen in the active site cleft of E217K must somehow be reversed by TM binding; however, it is not evident from the structure of TM fragment (EGF domains 4–6) how this could happen (37). Some of the exosite I residues that interact with TM have a different conformation in the structure of the E217K variant (residues 74–76), but this region is remote from the conformational changes that block the active site of the variant and changes in this region may be because of crystal packing. Thus, the mechanism by which Na\(^+\) binding activates thrombin may now be resolved, but the mechanism by which TM binding results in recovery of the catalytic activity of E217K thrombin remains unclear.

REFERENCES
1. Stubbs, M. T. & Bode, W. (1995) Thromb. Res. 69, 1–58
2. Stubbs, M. T. & Bode, W. (1995) Trends Biochem. Sci. 20, 23–28
3. Huntington, J. A. & Baglin, T. P. (2003) Trends Pharmacol. Sci. 24, 589–595
4. Huntington, J. A. (2003) J. Thromb. Haemostasis 1, 1535–1549
5. Esmon, C. T. (1987) Science 235, 1348–1352
6. Esmon, C. T. (2003) Chest 124, 268–268S
7. Guillen, M. C., Benezaud, A., Bontin, M. C. & Jandrot-Perrus, M. (1995) Thromb. Haemostasis 74, 129–133
8. Di Cera, E., Dang, Q. D. & Ayala, Y. M. (1997) Cell Mol. Life Sci. 53, 701–730
9. Verhamme, F. M., Olson, S. T., Tsilefse, D. M. & Bock, P. E. (2002) J. Biol. Chem. 277, 6788–6798
10. Huntington, J. A. & Esmon, C. T. (2003) Structure 11, 469–479
11. Wells, C. M. & Di Cera, E. (1992) Biochemistry 31, 11721–11730
12. Zhang, E. & Tulinsky, A. (1997) Biophys. Chem. 63, 185–200
13. Di Cera, E., Guinto, E. R., Vindigni, A., Dang, Q. D., Ayala, Y. M., Wuyi, M. & Tulinsky, A. (1995) J. Biol. Chem. 270, 23089–23092
14. Dang, O. D., Vindigni, A. & Di Cera, E. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 5977–5981
15. Parry, M. A., Stone, S. R., Hofsteenge, J. & Jackman, M. P. (1993) Biochem. J. 290, 665–670
16. Lai, M. T., Di Cera, E. & Shafer, J. A. (1997) J. Biol. Chem. 272, 30275–30282
17. Pineda, A. O., Savvides, S., Wakeham, G. & Di Cera, E. (2002) J. Biol. Chem. 277, 40177–40180
18. Hall, S. W., Gibbs, C. S. & Leung, L. L. (1997) Cell Mol. Life Sci. 53, 731–736
19. Guinto, E. R. & Di Cera, E. (1997) Biophys. Chem. 63, 185–200
20. Matsuura, S. Y. & Mejia, S. J. (1965) Nature 198, 413–416
21. Tsiang, M., Paborsky, L. R., Li, W. X., Jain, A. K., Dunn, K. E., Lee, D. W., Matsuura, S. Y., Matteucci, M. D., Coutre, S. E., Leung, L. L. & Gibbs, C. S. (1996) Biochemistry 35, 16449–16457
22. Cantwell, A. M. & Di Cera, E. (2000) J. Biol. Chem. 275, 39827–39830
23. Gruber, A., Cantwell, A. M., Di Cera, E. & Hanson, S. R. (2002) J. Biol. Chem. 277, 27561–27584
24. Collaborative Computing Project Number 4 (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 780
25. Trumper, A. T., Bode, W. (1992) Trends Biochem. Sci. 17, 569–569