Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 2779–2783, April 1992
Biochemistry

Partial characterization of vertebrate prothrombin cDNAs: Amplification and sequence analysis of the B chain of thrombin from nine different species

(blood coagulation/serine protease/thrombomodulin/polymerase chain reaction/protein evolution)

DAVID K. BANFIELD AND ROSS T. A. MACGILLIVRAY*

Department of Biochemistry, University of British Columbia, Vancouver, V6T 1W5, Canada

Communicated by Earl W. Davie, September 20, 1991

ABSTRACT **The cDNA sequence of the B chain of thrombin (EC 3.4.21.5) has been determined from nine vertebrate species (rat, mouse, rabbit, chicken, gecko, newt, rainbow trout, sturgeon, and hagfish). The amino acid sequence identities vary from 96.5% (rat vs. mouse) to 62.6% (newt vs. hagfish). Of the 240 amino acids spanned in all the species compared, there is identity at 110 (45.8%) positions. When conservative changes are included, the amino acid similarity increases to 75%. The most conserved portions of the B chain are the active-site residues and adjacent amino acids, the B loop, and the primary substrate-binding region. In addition, the Arg-Gly-Asp motif is conserved in 9 of the 11 species compared, and the chemotactic/growth factor domain is well conserved in all of the 11 species compared. The least conserved regions of the B chain correspond to surface loops, including the putative thrombomodulin-binding sites and one of the hirudin-binding regions. The extent of the amino acid sequence similarity and the conservation of many of the functional/structural motifs suggests that, in addition to their role in blood coagulation, vertebrate thrombins may also play an important role in the general mechanisms of wound repair.**

The final reaction of the coagulation pathway is the conversion of fibrinogen to fibrin by the serine protease thrombin (EC 3.4.21.5) (1, 2). In mammals, thrombin is generated from its zymogen, prothrombin, by the limited proteolytic action of factor Xa in the presence of factor Va, calcium ions, and phospholipid (2). In addition, thrombin also activates/inactivates other coagulation factors such as factor XIII, factors V and VIII, and protein C. The anticoagulant activity of thrombin is regulated through the interaction of thrombin with thrombomodulin (2, 3), an endothelial cell membrane protein related in structure to the low density lipoprotein receptor (4). While thrombomodulin likely binds to thrombin by an exposed surface loop, the precise binding site has yet to be resolved (5, 6). The thrombin/thrombomodulin complex activates protein C, which in the presence of protein S then degrades factors Va and VIIIa (2).

The enzymatic activity of thrombin is regulated by the endogenous protease inhibitor antithrombin III (2). Thrombin is also inhibited by protein inhibitors from nonplasma sources such as hirudin, from the European medicinal leech (7). The substrate specificity of thrombin is similar to that of trypsin, cleaving after basic amino acid residues (8). However, thrombin has a more restricted substrate range than trypsin. The molecular basis of this restricted substrate specificity is still unresolved but is thought to be the result of interactions of the substrates with secondary binding sites distant from the active site (9, 10).

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

Thrombin is also a potent stimulator of tissue plasminogen activator release from endothelial cells (11), and various forms of thrombin are chemotactic for monocytes and neutrophils (12–14). This chemotactic activity is blocked by binding with hirudin or antithrombin III (15). For monocytes, the chemotactic activity has been localized to residues 334–399 of thrombin (16). In addition to having chemotactic activity, this polypeptide sequence induces differentiation of certain macrophage lines (17). Thrombin also contains an Arg-Gly-Asp tripeptide analogous to the adhesion site in adhesive proteins such as laminin, fibronectin, and fibrinogen and is a possible site through which thrombin binds to receptors (18). The fact that thrombin possesses these additional bioregulatory and growth-stimulating activities suggests it may also play an important role in the wound healing process as well as fibrinolysis.

Thrombin consists of two polypeptide chains joined by a disulfide bond. The A chain (49 amino acids in bovines and 36 in humans) has no known function. The B chain (259 amino acids in both humans and bovines) is structurally similar to other serine proteases (9, 10). When the amino acid sequence of the thrombin B chain is aligned with that of bovine chymotrypsin, amino acid sequence insertions in the B chain are found at exon junctions (19). The crystal structure of human α-thrombin has placed these amino acid insertions on the surface of the protein (20). Crystal structures have now been described for human α-thrombin (20) as well as the thrombin/hirudin complex (21, 22). The solution of the crystal structure of thrombin should facilitate the identification of regions on the surface of the protein that are potentially involved in protein–protein interactions.

Amino acid sequence data from homologous proteins in divergent species can provide valuable insight into the amino acid residues involved in active-site/substrate specificity, protein–protein interactions, and species-specific differences in biological processes. To examine the structural constraints during the evolution of prothrombin, we have used the polymerase chain reaction (PCR) (23, 24) to amplify and characterize approximately 900 base pairs (bp) of thrombin B-chain sequence from nine vertebrate species.†

MATERIALS AND METHODS

Sample Collection. Liver samples from rat (*Ratus norwegicus*), mouse (*Mus musculus*), rabbit (*Oryctolagus cuniculus*), chicken (*Gallus gallus*), gecko (*Gekko gekko*), Jap-

Abbreviation: UTS, untranslated sequence.

*To whom reprint requests should be addressed.

†The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M81391 for *Gallus gallus*, M81392 for *Gekko gekko*, M81393 for *Eptatretus stouti*, M81394 for *Mus musculus*, M81395 for *Cynops pyrogastor*, M81396 for *Oryctolagus cuniculus*, M81397 for *Rattus norvegicus*, M81398 for *Oncorhynchus mykiss*, and M81399 for *Acipenser transmontanus*).

Table 1. Seven primers for amplification and sequencing of vertebrate thrombin B-chain cDNAs

Primer	Sequence	Amino acids spanned*
Th3	5′-GAGCTGCTGTGTGGGGCCAGCCTCATCAG-3′	347–355
Th4	5′-GGCTTGTAACCAGCACGAACAT-3′	502–508
Th7	5′-AGCGCACCTTGGCAGGTGATG-3′	333–340
Th10	5′-AAGGGCGTGTGACTGGC/ATGGGG-3′	458–465
Ser1†	5′-ACAAAAGCTTGA/AIGGICCICCIC/GT/AA/GTCICC-3′	526–532
His1†	5′-ACAGAATTCTGGGGTIG/CTIACIGCIGCICAC/TTG-3′	360–367
T_{17XSP}	5′-ACACTGCAGGAGCTCTCTAGATTTTTTTTTTTTTTTT-3′	

*Amino acid sequence numbering is based on human prothrombin (26).
†Adapted from ref. 28. The letter I in primers Ser1 and His1 represents deoxyinosine.

anese fire-bellied newt (*Cynops pyrogastor*), rainbow trout (*Oncorhynchus mykiss*), and sturgeon (*Acipenser transmontanus*) were generously provided by colleagues at the University of British Columbia. Hagfish (*Eptatretus stouti*) were purchased from Seacology (Vancouver).

RNA Isolation. Total cellular RNA was isolated from fresh or previously frozen liver samples by using the acid phenol extraction procedure as described by Chomczynski and Sacchi (25).

Oligonucleotide Preparation and Design. Oligonucleotides were synthesized on an Applied Biosystems model 380B or 391 DNA synthesizer. Primer sequences were selected after alignment of cDNA sequences from human (26), bovine (27), and chicken (unpublished data) prothrombin. A total of five oligonucleotides spanning approximately 240 amino acids of the B chain of thrombin were used for the amplification of all vertebrate cDNA fragments with the exception of hagfish prothrombin (see Table 1). Hagfish prothrombin cDNA fragments were amplified by using the degenerate inosine-containing primers Ser1 and His1 (28).

Preparation of Single-Stranded cDNA. Single-stranded cDNA was prepared as described (29) by using the oligonucleotide T_{17XSP} (see Table 1) as the primer. Reaction mixtures were incubated at 37°C for 1 hr, diluted to 200 μl with sterile distilled water, and stored at −20°C.

PCR. PCR reactions were performed as described (24), using 1–2 units of *Thermus aquaticus* (*Taq*) DNA polymerase (Perkin–Elmer/Cetus), and 2–5 ng of cDNA in 50-μl reaction mixtures.

Sequence Determination of Amplified cDNA Fragments. After amplification, PCR samples were made blunt-ended with the Klenow fragment of *Escherichia coli* DNA polymerase I (30). The amplified blunt-ended DNA was ligated into the *Hin*cII site of pUC19. A minimum of four independent PCR products were sequenced from each species examined.

Nucleic Acid and Amino Acid Sequence Analysis. Sequence data were analyzed with the Delaney DNA sequence program (Delaney Software, Vancouver). Amino acid sequences were aligned by using ESEE (31).

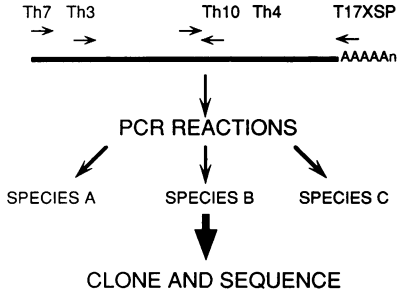

FIG. 1. Amplification strategy. The positions of the oligonucleotide primers used to amplify and sequence portions of the B-chain of vertebrate thrombin (see Table 1).

RESULTS

Amplification of Thrombin B-Chain cDNA Fragments. Amplification of single-stranded cDNA with primers Th7 and T_{17XSP} produced a fragment of ≈900 bp in all species examined except the hagfish (data not shown). Amplification of single-stranded cDNA with primers Th7 and Th4 produced a fragment 500 bp in length, and amplification with primers Th10 and T_{17XSP} yielded a fragment 300 bp in length (see Fig. 1). Amplification using the Ser1 and His1 primers produced fragments of ≈500 bp. This product represents a number of different trypsin-like serine protease cDNA fragments (28), including prothrombin (see Table 1).

Sequence Analysis of Amplified cDNA Fragments. With the exception of the fragments that were amplified by using the degenerate oligonucleotides, all amplified cDNA fragments contained a single species. Prothrombin cDNA fragments were identified by direct sequence analysis (32) and alignment of DNA and amino acid sequences with either human (26) or bovine (27) sequences. To identify the cDNA fragment corresponding to hagfish prothrombin, the Ser1 and His1 oligonucleotide primers were used. Several of the 500-bp fragments amplified by using these primers were cloned and their sequences were determined. Identification of the cDNA corresponding to the B chain of thrombin was based on the presence of amino acid residues Tyr-Pro-Pro-Trp. This sequence represents the B loop of thrombin (19). Among proteases, the B loop is unique to thrombin. The fragment corresponding to the B chain of hagfish prothrombin spanned the region from the active site histidine to the active site serine for a total of 161 amino acid residues.

On completion of sequence analysis, the portion of the B chain corresponding to residues 344–579 of prothrombin had been determined for each of the nine species. While the lengths of the coding sequences are similar in all the species examined (see below for the exceptions), the length of the 3′ untranslated sequence (UTS) is quite heterogeneous. The length of the 3′ UTS varies from 82 nucleotides in the trout and sturgeon to 1145 nucleotides in chicken, with the majority being ≈80–100 nucleotides (see Table 2). With the exception of the mouse and rat B chains, there are no significant regions of nucleotide sequence identity among any of the

Table 2. Length of thrombin B-chain coding sequences and 3′ UTS for the nine species

Species	Length, nucleotides	
	Coding sequence	3′ UTS
Rat	705	115
Mouse	705	114
Rabbit	705	121
Chicken	702	154/1145
Gecko	705	224
Newt	705	108
Rainbow trout	717	82
Sturgeon	702	82
Hagfish	708	407

species compared, excluding the polyadenylylation consensus sequence AATAAA. While this study was in progress the cDNA sequences of rat (33) and mouse (34) prothrombin were published. The sequences determined by PCR in this study are identical to these published sequences.

Amino Acid Sequence Alignments. The predicted amino acid sequences of the nine vertebrate thrombin B chains are aligned with the previously published sequences of human (26) and bovine (27) thrombin B chains in Fig. 2. The length of this portion of the B chain is relatively invariant; only a single amino acid insertion is present in some species at position 472 (Fig. 2, box D). The amino acid residue at position 472 is absent from chicken, newt, gecko, rainbow trout, sturgeon, and hagfish thrombin. In addition to the variability at position 472, there is variability in the composition and length of the C terminus (see Fig. 2). The amino acid sequences surrounding the active-site residues His-363, Asp-419, and Ser-525 are conserved in all 11 species (see Fig. 2). All of the eight tryptophan and seven cysteine residues seen in the human B-chain sequence are conserved in each of the species examined. Ten of the 12 prolines are also conserved. When the 240 amino acids of the B chain from the 11 different species are aligned, there is amino acid sequence variation at 130 positions (54%).

DISCUSSION

We have amplified, cloned, and determined the DNA sequences of the majority of the B chain of thrombin from nine different vertebrate species. Comparisons of amino acid sequences with those of previously characterized portions of human and bovine prothrombin suggest that thrombin is highly conserved throughout vertebrate evolution (see Table 3). Amino acid sequence identies range from 96.5% (between

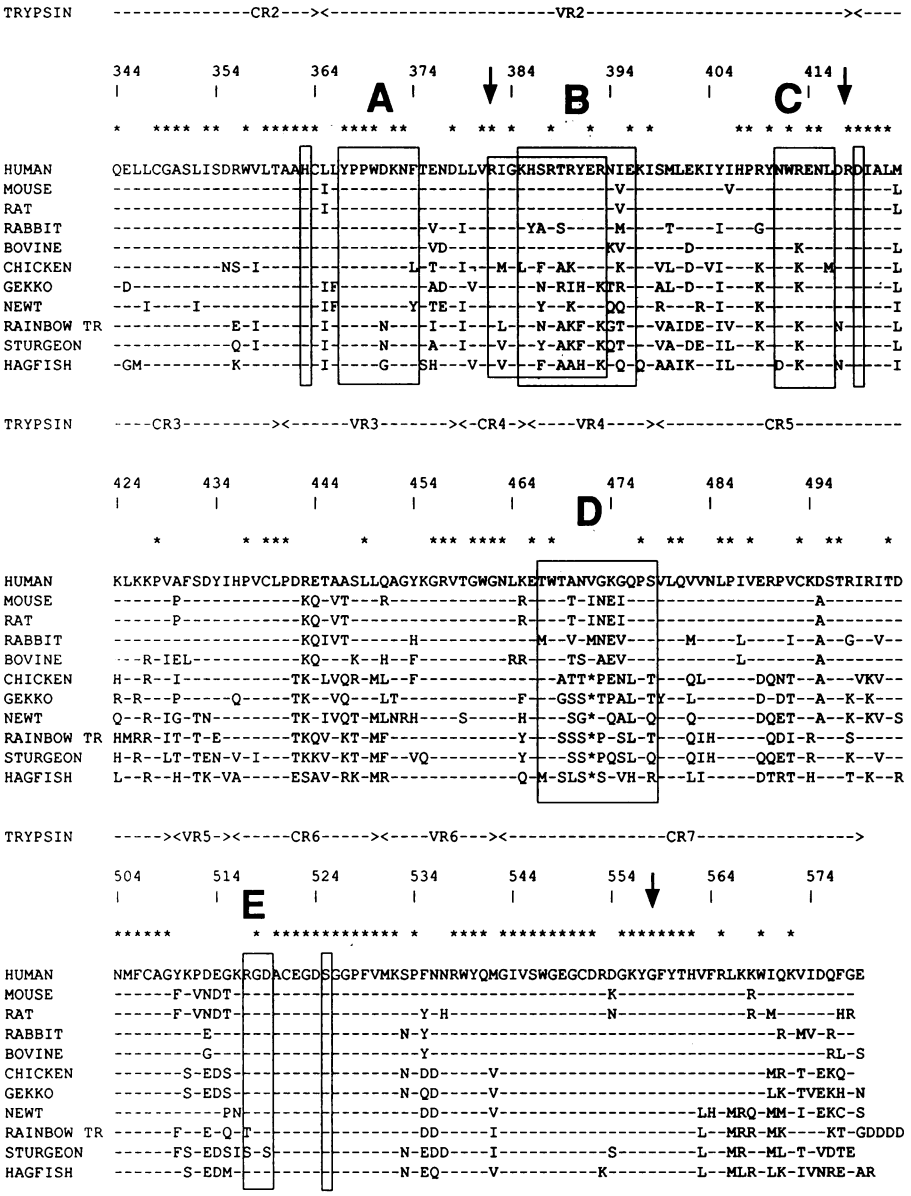

FIG. 2. Amino acid sequence alignment of vertebrate thrombin B-chain sequences. All of the aligned amino acid sequences were predicted from the translation products of the cDNA sequences. Amino acid sequence numbering is based on human prothrombin (26). Residues identical to the human sequence are indicated by a dash. The * above the human sequence identifies amino acid residues identical in all 11 species. Amino acid substitutions are as indicated, and the deletion at position 472 in some species is marked by an *. The active-site His, Asp, and Ser residues are marked by the narrow boxes. Arrows above the human sequence indicate the positions of Arg-382, Arg-418, and Gly-558 (see text for details). For descriptions of residues bounded by boxes A, B, C, D, and E, see text. The broken lines refer to positions representing either conserved regions (CR) or variable regions (VR) as described by Furie *et al.* (19).

Table 3. Percent amino acid sequence identity among vertebrate thrombin B-chains

	R	M	Rbt	B	C	G	N	RTr	S	Hgf
Human	88.5	89.8	85.1	87.3	72.2	72.8	70.2	68.6	65.0	65.3
Rat		96.5	82.6	86.0	71.4	73.2	68.5	68.5	65.5	63.8
Mouse			83.0	86.4	71.4	73.2	68.5	69.4	65.0	64.7
Rabbit				82.6	73.1	71.9	68.9	66.4	64.1	64.1
Bovine					72.2	72.3	68.5	68.8	65.0	64.0
Chicken						77.8	73.5	72.2	72.2	65.8
Gecko							71.1	70.2	70.1	69.4
Newt								69.4	69.2	62.6
Rainbow trout									82.5	68.6
Sturgeon										66.2

Percent identity data was generated with PALIGN (IntelliGenetics) using the structure-genetic matrix, open gap = 5; unit gap = 50. R, rat; M, mouse; Rbt, rabbit; B, bovine; C, chicken; G, gecko; N, newt; RTr, rainbow trout; S, sturgeon; Hgf, hagfish.

mouse and rat) to 62.6% (between newt and hagfish). Thrombins from mammalian species share greater than 82% amino acid sequence identity. The percent identity is much lower among the nonmammalian species (66% average). Rainbow trout and sturgeon share 82.5% amino acid sequence identity; this is 10% more identity than either shares with any of the other nine species. This high amino acid sequence identity may reflect a slower rate of nucleotide sequence change in fish or constraints on the B chain through interactions with other proteins involved in hemostasis. The overall amino acid sequence identity in this region of prothrombin is 43.3% between the 11 vertebrates; when conservative changes are included, the overall amino acid sequence similarity increases to 74.6%.

The specificity of thrombin for its substrates is likely determined by its insertion loops (10, 19, 20). The locations of these insertion loops are indicated in Fig. 2 and correspond to the VR regions defined by Furie *et al.* (19). A number of the loop segments are similar to those of chymotrypsin but contain additional residues, which alter the overall three-dimensional structure of the region (20). Several of the loop structures in the B chain of thrombin have been implicated in interactions with substrate (boxes A and C and the first four residues of box D in Fig. 2) and as sites of interaction with hirudin (box B in Fig. 2) (21, 22) and thrombomodulin (residues 382–392; see box B and box D in Fig. 2) (5, 6). With the exception of the B loop, corresponding to residues 365–369 (box A in Fig. 2), and the loop highlighted by box C (Fig. 2), all loop structures have highly variable amino acid sequences. This variability in surface loop amino acid sequence may contribute to some of the species-specific differences observed between thrombin and fibrinogen (35–37).

The human thrombin B chain contains a region with growth factor and chemotactic activity *in vitro* (16, 17). This chemotactic/growth factor domain corresponds to residues Leu-335 to Met-400 of prothrombin (see Fig. 2). In addition to the chemotactic/growth factor activity attributed to this region, residues 335–400 also contains the B loop (box A in Fig. 2), active-site histidine (His-363), a hirudin-binding loop (box B in Fig. 2), a potential thrombomodulin-binding site (Arg-382 to Arg-393; see Fig. 2), and a fibrinogen-binding site (6). With the exception of the hirudin binding site and the putative thrombomodulin-and fibrinogen-binding regions, this portion of thrombin contains very few amino acid sequence differences among the species examined. Analysis of the thrombin crystal structure has identified the B loop (box A in Fig. 2) as forming an extended loop structure that restricts access to the active-site cleft, which may be responsible for the limited substrate specificity of thrombin (20). Due to the rigid kinked structure of the B loop, it has been suggested that this loop may also represent part of the active substructure of the chemotactic/growth factor domain (10, 20). The amino acid sequence similarity in this region (28/56 identical residues, 8/56 conservative changes) suggests that the chemotactic/growth factor activity attributed to this region in human thrombin may have a similar activity in other vertebrate thrombins. In addition to the putative role of the B loop in substrate specificity and chemotactic/growth factor activity, this portion of the B chain contains a site for N-linked glycosylation (Asn-373). The asparagine residue at position 373 is conserved in all of the 11 species.

Residues Lys-385 to Glu-396 (box B in Fig. 2) represent one of the surface contact loops of hirudin with human thrombin (21, 22). Interestingly, only 4 of the 12 residues in this region are identical or have conservative changes: Arg-388, Arg/Lys-390 (except hagfish and newt), Arg/Lys-393, and Lys-397 (except hagfish) (see box B in Fig. 2). Analysis of the crystal structure of the thrombin–hirudin complex reveals a number of possible electrostatic interactions between these residues and the acidic C-terminal residues of hirudin (21).

To date, two putative thrombomodulin-binding sites have been identified in human thrombin (5, 6). One binding site corresponds to residues Arg-382 to Arg-393 (see Fig. 2), a region of thrombin that has also been identified as binding hirudin (21, 22) and fibrinogen (6). The other thrombomodulin-binding site corresponds to residues Thr-468 to Ser-478 (box D in Fig. 2), where the surface loop formed by Glu-467 to Tyr-470 is located at the entrance of the binding cleft. The putative thrombomodulin-binding site at Arg-382 to Arg-393 (6) overlaps a hirudin-binding site (see box B in Fig. 2). The amino acid sequence of this region is conserved among the mammals compared, with the exception of the rabbit (see Fig. 2); all five of the basic residues in this region are conserved. The amino acid sequences of the nonmammalian species share an average of 50% amino acid sequence identity with the mammalian species. With the exception of the chicken sequence, 4/5 basic residues are conserved in the nonmammalian species compared. The putative thrombomodulin-binding site corresponding to residues Thr-468 to Ser-478 (box D in Fig. 2) (5) is highly variable, with only 2 of the 12 amino acid residues being invariant (Trp-468 and Pro-477). Unlike the putative thrombomodulin-binding site at Arg-382 to Arg-393, the amino acid sequence in this region varies significantly among the mammalian species compared. The majority of the amino acid sequence changes in this region are nonconservative. In addition, this is the only region of the B chain to contain an amino acid deletion. The precise location of the thrombomodulin-binding site(s) in human α-thrombin has yet to be resolved (5, 6, 38, 39).

Human thrombin contains an Arg-Gly-Asp tripeptide sequence at positions 517–519 (box E in Fig. 2) analogous to the adhesion site in adhesive proteins such as laminin, fibronectin, and fibrinogen (18). This region of thrombin has been shown to promote endothelial cell adhesion, spreading, and cytoskeletal reorganization, potentially contributing to repair mechanisms and maintenance of the internal blood vessel

lining (18). Although the cell adhesion activity was greatest with a chemically modified form of thrombin (NO_2-α-thrombin), native thrombin was also found to promote endothelial cell adhesion. While a portion of the loop segment Tyr-509 to Gly-518 is exposed to the solvent, the tripeptide Arg-Gly-Asp is not (20). The Arg-Gly-Asp sequence is conserved in 9 of the 11 species, supporting a possible role for this sequence in vertebrate thrombins *in vivo*.

The C-terminal region of thrombin is variable in both composition and length. Within the C-terminal 10 residues of human thrombin, only Trp-569 and Lys-572 are conserved in all of the species compared. According to the crystal structure, the C terminus of human thrombin is exposed to the solvent (20); however, no interactions or functions have been assigned to this region.

A number of abnormal human prothrombins have been characterized and their molecular defects have been identified. Three of these are found in the B chain. Prothrombin Quick I (40) is characterized by an Arg → Cys change at residue 382 (between boxes A and B in Fig. 2). The Arg residue is conserved in all 11 species. This amino acid has been identified from the three-dimensional structure as one of the residues lining the long groove extending from the active site and may form part of the putative fibrinogen secondary binding site (20). Substitution of Cys for Arg at this position probably disrupts thrombin/fibrinogen interactions. In prothrombin Tokushima, the Arg at position 418 is replaced by a Trp (41). This Arg residue is also conserved in all 11 species. Arg-418 is adjacent to box C (Fig. 2), which forms one of the surface loops projecting out from the active site cleft (20). It may be the nature of the amino acid change at position 418 that leads to the decreased enzyme efficacy observed in prothrombin Tokushima. A Gly → Val substitution at position 558 is found in prothrombin Quick II (42). Gly-558 is adjacent to the Cys-551 to Tyr-557 loop segment (20) and is conserved in all 11 species. Substitution of Val for Gly appears to alter the primary substrate-binding pocket (42).

We are indebted to W. Kelley Thomas, Roger Graham, and Billy Chow for helpful discussions and to Tina Umelas for excellent technical assistance. This work was supported by Grant MT 7716 from the Medical Research Council of Canada to R.T.A.M.

1. Jackson, C. M. & Nemerson, Y. (1980) *Annu. Rev. Biochem.* **49**, 765–811.
2. Rosenberg, R. D. (1987) in *The Molecular Basis of Blood Diseases*, eds. Stamatoyannopoulos, G., Nienhuis, A. W., Leder, P. & Majerus, P. W. (Saunders, Philadelphia) pp. 534–574.
3. Jackman, R. W., Beeler, D. L., VanDeWater, L. & Rosenberg, R. D. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8834–8838.
4. Jackman, R. W., Beeler, D. L., Fritze, L., Soff, G. & Rosenberg, R. D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6425–6429.
5. Suzuki, K., Nishioka, J. & Hayashi, T. (1990) *J. Biol. Chem.* **265**, 13263–13267.
6. Noe, G., Hofsteenge, J., Ravelli, G. & Stone, S. R. (1988) *J. Biol. Chem.* **263**, 11729–11735.
7. Markwardt, F. (1970) *Methods Enzymol.* **19**, 924–932.
8. Fenton, J. W., II, (1981) *Ann. N.Y. Acad. Sci.* **370**, 468–495.
9. Fenton, J. W., II, & Bing, D. H. (1986) *Semin. Thromb. Hemostasis* **12**, 200–206.
10. Fenton, J. W. (1988) *Semin. Thromb. Hemostasis* **14**, 234–240.
11. Levin, E. G., Stern, D. M., Nawroth, R. A., Marlar, R. A., Fair, D. S., Fenton, J. W., II, & Harker, L. A. (1986) *Thromb. Haemostasis* **56**, 115–119.
12. Bar-Shavit, R., Kahn, A. J., Wilner, G. D. & Fenton, J. W., II (1983) *Science* **220**, 728–731.
13. Bizios, R., Lau, L., Fenton, J. W., II, & Malik, A. B. (1986) *J. Cell Physiol.* **128**, 485–490.
14. Morin, A., Arvier, M., Doutremepuich, C. & Vigneron, C. (1990) *Thromb. Res.* **60**, 33–42.
15. Bar-Shavit, R., Bing, D. H., Kahn, A. J. & Wilner, G. D. (1985) in *Membrane Receptors and Cellular Regulation*, UCLA Symposium in Molecular and Cellular Biology, New Series, eds. Czech, M. P. & Kahn, C. P. (Liss, New York), Vol. 23, pp. 329–338.
16. Bar-Shavit, R., Kahn, A., Mudd, S., Wilner, G. D., Mann, K. G. & Fenton, J. W. (1984) *Biochemistry* **23**, 400–403.
17. Bar-Shavit, R., Kahn, A., Mann, K. G. & Wilner, G. D. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 976–980.
18. Bar-Shavit, R., Sabbah, V., Lampugnani, M. G., Marchisio, P. C., Fenton, J. W., II, Vlodavsky, I. & Dejana, E. (1991) *J. Cell Biol.* **112**, 335–344.
19. Furie, B., Bing, D. H., Feldman, R. J., Robison, D. J., Burnier, J. P. & Furie, B. C. (1982) *J. Biol. Chem.* **257**, 3875–3882.
20. Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. T. & Hofsteenge, J. (1989) *EMBO J.* **8**, 3467–3475.
21. Grutter, M. G., Priestle, J. P., Rahuel, J., Grossenbacher, H., Bode, W., Hofsteenge, J. & Stone, S. R. (1990) *EMBO J.* **9**, 2361–2365.
22. Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C. & Fenton, J. W., II (1990) *Science* **249**, 277–280.
23. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230**, 1350–1354.
24. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. (1988) *Science* **239**, 487–491.
25. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159.
26. Degen, S. J. F., MacGillivray, R. T. A. & Davie, E. W. (1983) *Biochemistry* **22**, 2087–2097.
27. MacGillivray, R. T. A. & Davie, E. W. (1984) *Biochemistry* **23**, 1626–1634.
28. Sakanari, J. A., Staunton, C. E., Eakin, A. E., Craik, C. S. & McKerrow, J. H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4863–4867.
29. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
30. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 394–395.
31. Cabot, E. L. & Beckenbach, A. T. (1989) *Comput. Appl. Biosci.* **5**, 233–234.
32. Gyllensten, U. B. & Erlich, H. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7652–7656.
33. Dihanich, M. & Monard, D. (1990) *Nucleic Acids Res.* **18**, 4251.
34. Friezner-Degen, S. J., Schaefer, L. A., Jamison, C. S., Grant, S. G., Fitzgibbon, J. J., Pai, J.-A., Chapman, V. M. & Elliot, R. W. (1990) *DNA Cell Biol.* **9**, 487–498.
35. Doolittle, R. F. (1965) *Biochem. J.* **94**, 735–741.
36. Doolittle, R. F., Oncley, J. L. & Surgenor, D. M. (1962) *J. Biol. Chem.* **237**, 3123–3127.
37. Ratnoff, O. D. (1987) *Perspect. Biol. Med.* **31**, 1–33.
38. Tsiang, T., Lentz, S. R., Dittman, W. A., Wen, D., Scarpati, E. & Sadler, J. E. (1990) *Biochemistry* **29**, 10602–10612.
39. Hofsteenge, J., Braun, P. J. & Stone, S. R. (1988) *Biochemistry* **27**, 2144–2151.
40. Henriksen, R. A. & Mann, K. G. (1988) *Biochemistry* **27**, 9160–9165.
41. Miyata, T., Morita, T., Inomoto, T., Kawauchi, S., Shikakami, A. & Iwanaga, S. (1987) *Biochemistry* **26**, 1117–1122.
42. Henriksen, R. A. & Mann, K. G. (1989) *Biochemistry* **28**, 2078–2082.