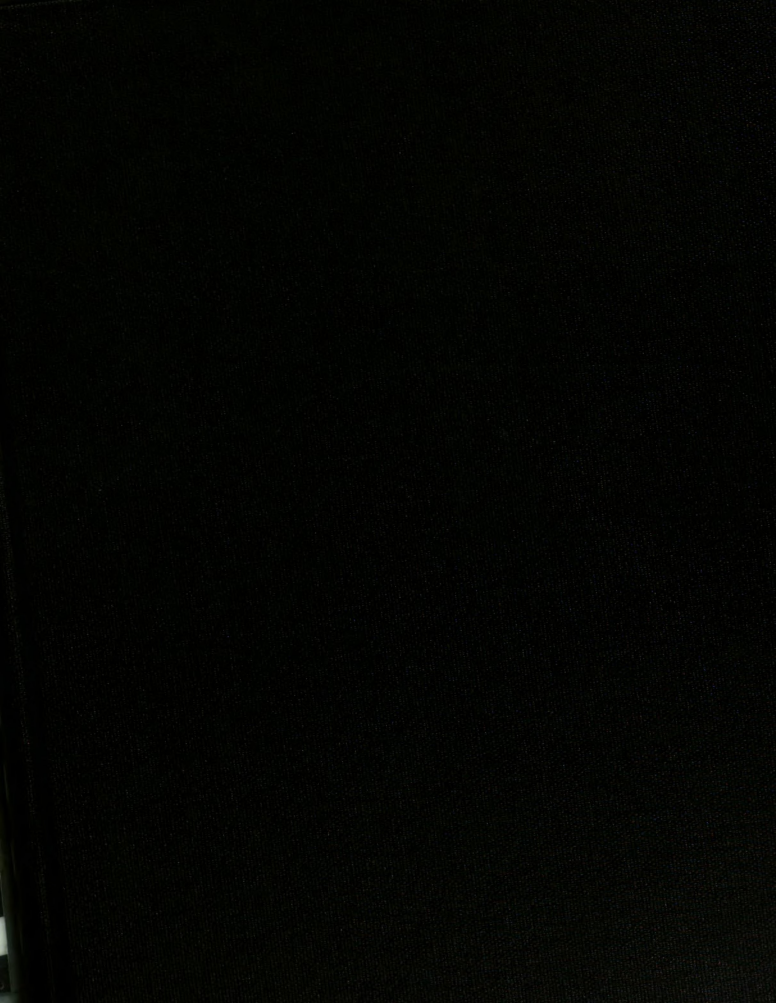






This is to certify that the

dissertation entitled

*Structural Aspects of Thrombin and
KRINGLE DOMAINS: INHIBITION, LIGAND
BINDING MODES AND THEIR BIOLOGICAL
RELEVANCE.*

presented by

IGOR Mochalkin

has been accepted towards fulfillment
of the requirements for

PhD degree in Chemistry

A. Tulinsky

Major professor

Date Aug. 20, 1999.

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

**STRUCTURAL ASPECTS OF THROMBIN AND KRINGLE DOMAINS:
INHIBITION, LIGAND BINDING MODES AND THEIR BIOLOGICAL
RELEVANCE.**

By

Igor Mochalkin

A DISSERTATION

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1999

SEL.

R-va

bind

thron

corre

bindi

Xa. n

conse

in thre

inhibit

and (9

SEL27

fashion

ABSTRACT

STRUCTURAL ASPECTS OF THROMBIN AND KRINGLE DOMAINS: INHIBITION, LIGAND BINDING MODES AND THEIR BIOLOGICAL RELEVANCE.

By

Igor Mochalkin

The X-ray crystal structures of two thrombin active site inhibitors (SEL2711 and SEL2770) have been determined and refined in the (9.0-2.1) Å resolution range to final R-values of 16.5% and 16.7%. Unlike most thrombin inhibitors, SEL2711 and SEL2770 bind in a retro-fashion, forming a parallel β-strand with residues Ser214-Gly216 of thrombin. The thrombin bound crystal structures of SEL2711 and SEL2770 were used to correlate the binding constants and contributions of the different thrombin subsites to the binding mode. Since the Selectide inhibitors are about 10^4 times more specific for Factor Xa, modeling of the inhibitors to the latter suggests that the selectivity can be a consequence of interactions in the S3-S4 binding subsites of Factor Xa that are different in thrombin and Factor VIIa.

The X-ray crystal structures of two β-strand templated thrombin active site inhibitors (MOL376 and MOL592) have been determined and refined in the (9.0-2.3) Å and (9.0-2.0) Å resolution ranges to final R-values of 15.3% and 17.4%. Unlike SEL2711 and SEL2770, MOL376 and MOL592 bind to thrombin in a substrate-like fashion, making an anti-parallel β-strand with residues Ser214-Gly216 of the enzyme.

The n

used

future

throm

of th

patter

of th

resid

plasr

starti

The

of th

K4.

muta

apoli

resol

(1) b

Arg7.

shifts

The network of interactions between the inhibitors and thrombin at the S1 subsite was used to correlated the trend in the binding constants and provide important insight for the future design of therapeutic antithrombotic agents. Probing the operating sites of thrombin further with urea showed that three urea molecules bind to the active site region of the enzyme and two are located next to the heparin binding exosite II. The overall pattern of the interaction at the specificity site is similar to that found in crystal structures of thrombin-PPACK and many other thrombin-inhibitor complexes in which an Arg residue occupies the P1 position.

The X-ray crystal structure of the recombinant kringle (K) 5 domain of human plasminogen has been determined by molecular replacement methods using K1 as a starting model and refined in the (8.0-1.7) Å resolution range to a final R-value of 16.6%. The presence of Leu71 at the ligand binding site (LigBS) leads to a substantial decrease of the binding affinity of K5 for ϵ -aminocaproic acid (EACA), with respect to K1 and K4. Based on the crystal structure, the LigBS was rationally remodeled and a K5[L71R] mutant possesses an affinity for ω -amino acids that is similar to K1 and K4.

Three X-ray crystal structures of the recombinant KIV-10 variants of human apolipoprotein (a) have been determined and refined in the (9.0-1.8) and (9.0-2.1) Å resolution ranges to final R-values of (16.6-18.2)%. The crystal structures indicate that: (1) both KIV-10/Thr66 and KIV-10/Met66 can bind EACA in the embedded way, (2) Arg72 in the KIV-10/W72R mutant occupies the LigBS mimicking ligand and (3) Arg35 shifts upon EACA binding to assist Arg71 at the cationic center.

its suc

Alexan

behind

sugge

assista

Rober

encou

throug

I wis

Padm

SEL2

Raghu

of the

with M

and a

crystal

mentio

crystal

knowle

research

ACKNOWLEDGMENTS

In a project of this magnitude, many individuals make significant contributions to its success. Foremost among those whom I would like to acknowledge for help is Prof. Alexander Tulinsky, my graduate adviser. Prof. Tulinsky not only suggested the ideas behind most of the studies but he also provided a constant flow of recommendations and suggestions and financial support so that most of the work could be done under a research assistantship. I also owe a special debt of gratitude to the members of his team: Dr. Robert St. Charles, Dr. Jorge L. Rios-Steiner and Dr. Raman Krishnan for their guidance, encouragement, generosity and advice in technical, academic and career matters throughout my graduate career.

There are many other people to whom I wish to express my thanks for their help. I wish to express my appreciation to Dr. Irimpan I. Mathews and Dr. Kaillathe Padmanabhan for their help in my earliest project: retro-binding inhibition of thrombin by SEL2711 and SEL2770. I wish to express my thanks to Dr. Raman Krishnan and Dr. Raghuvir K. Arni for their contributions (MOL354, MOL1245 and MOL356) to the study of the non-electrophilic Molecumetics inhibitors as they related to my research project with MOL376 and MOL592. Furthermore, the projects involving plasminogen kringle 5 and apolipoprotein (a) KIV-10 greatly benefited from the expertise in protein crystallization of Dr. Beisong Cheng, who crystallized these kringle modules. I must also mention my appreciation to Dr. Peggy Vanderhoff-Hanaver who introduced me to the crystallization techniques and helped me to build a strong foundation to become knowledgeable in this field. I would like to thank Prof. Francis J. Castellino and his research group from the University of Notre Dame, Prof. Angelo Scanu and Dr. Olga

Klezo

Pharm

kring!

coord

I wou

and 1

Germ

numl

Insti

NMI

my 1

Wal

Med

expa

prog

John

who

funda

neces

repay

Klezovitch from the University of Chicago and Dr. Brian R. Boettcher of Novartis Pharmaceuticals Corp., (Summit, NJ.) for providing protein samples of plasminogen kringle 5 (F.J.C) and apolipoprotein (a) KIV-10 (A.S and O.K.) and unpublished atomic coordinates and observed structure factors of unbounded KIV-10/Thr66/EACA (B.R.B). I would like to thank the Selectide Corp. and Molecumetics Ltd. for providing the SEL and MOL inhibitors and Dr. Herman Schreuder of Hoescht Marion Roussel (Frankfurt, Germany) and Dr. Hiroshi Nakanishi of Molecumetics Ltd. (Bellevue, WA.) for a number of helpful discussions. Thanks are also due to Dr. F. Fraternali of the National Institute for Medical Research (London, England) for providing the coordinates of the NMR structure of Factor Xa complexed with peptide A and B. The writing of most of my published papers was supported by the James L. Dye Endowed Fellowship and the Walter and Margaret Yates Memorial Scholarship.

I gratefully acknowledge the influence of Prof. Gerald Davis (Department of Medical Technology, MSU) and Prof. Houria Hassouna (Clinical Center, MSU) who expanded my knowledge in hematology and hemostasis and maintained an interest in my progress. I also wish to express my thanks to my best friends and my loved ones: Dianna Johnson, René St.Louis and Alexander and Millicent Popov for their support.

And finally, I wish express my most profound thanks to my parents and my sister who since my earliest recollections have always encouraged my curiosity, quest for fundamental knowledge and my attempts to learn. They have backed me financially when necessary, but more importantly, they have backed me with intangibles that I can never repay. It is to my parents that I dedicate this work.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

CHAPTER I

I. HEMODYNAMICS

I.1

I.2

I.3

II. FIBROBLASTS

III. MECHANICAL PROPERTIES

CHAPTER II

I. INTRODUCTION

I.1

I.2

I.3

II. EXPERIMENTAL METHODS

III. RESULTS

IV. DISCUSSION

TABLE OF CONTENTS

ACKONWLEDGEMENT.....	iv
TABLE OF CONTENTS.....	vi
LIST OF FIGURES.....	viii
LIST OF TABLES.....	xiv
ABBREVIATIONS AND SYMBOLS.....	xvii
CHAPTER 1: BLOOD PROTEINS AND COAGULATION	
I. HEMOSTASIS.....	1
I.1. Endothelium.....	1
I.2. Platelet	3
I.3. The Coagulation Cascade.....	8
II. FIBRINOLYSIS.....	12
III. MODULAR NATURE OF BLOOD PROTEINS.....	13
CHAPTER 2: RETRO-BINDING INHIBITION OF THROMBIN	
I. INTRODUCTION.....	22
I.1. Thrombin Binding Sites.....	22
I.2. Thrombin Active Site Inhibitors and Their Binding Modes.....	24
I.3. Selectide Corp. Inhibitors.....	25
II. EXPERIMENTAL.....	28
III. RESULTS.....	33
IV. DISCUSSION.....	39

CHAPTER

I. M

II. E

III.

IV.

CHAPTER

I. D

II. I

III.

CHAPTER

I. D

II. E

III.

IV.

CHAPTER

I. IN

II. E

III. E

**CHAPTER 3: SELECTIVE NON-ELECTROPHILIC INHIBITORS OF
THROMBIN**

I. MOLECUMETICS INHIBITORS.....	52
II. EXPERIMENTAL.....	55
III. RESULTS.....	56
IV. DISCUSSION.....	65

**CHAPTER 4: PROBING THE ACTIVE SITE OF THROMBIN FURTHER WITH
UREA**

I. INTRODUCTION.....	73
II. EXPERIMENTAL.....	76
III. RESULTS AND DISCUSSION.....	84

**CHAPTER 5: STRUCTURE-FUNCION DETERMINANTS OF RECOMBINANT
KRINGLE 5 OF HUMAN PLASMINOGEN**

I. INTRODUCTION.....	94
II. EXPERIMENTAL.....	97
III. RESULTS.....	103
IV. DISCUSSION.....	108

**CHAPTER 6: RECOMBINANT KRINGLE IV-10 MODULES OF HUMAN
APOLIPOPROTEIN (a): STRUCTURES, LIGAND BINDING
MODES AND BIOLOGICAL RELEVANCE**

I. INTRODUCTION.....	115
II. EXPERIMENTAL.....	119
III. RESULTS.....	124

IV.

GLOSSAR

BIBLIOGR

CHAPTER

Figure 1-1.

Figure 1-2.

Figure 1-3.

Figure 1-4.

Figure 1-5.

IV. DISCUSSION.....	133
GLOSSARY.....	140
BIBLIOGRAPHY.....	144

LIST OF FIGURES

CHAPTER 1: BLOOD PROTEINS AND COAGULATION

- Figure 1-1. Functions of endothelium. Anticoagulant factors and their modes of action in the coagulation cascade are represented in bold black. Coagulation factors secreted by the endothelial cells are in thin line text. Double dot denotes binary complex.
- Figure 1-2. Megakaryocyte cell development from the level of the stem cells, through BFU-MK, CFU-MK, and then into the morphologically recognizable platelet precursor cell in stage I-IV.
- Figure 1-3. Progress of platelet plug formation in hemostasis. A – platelet activation and secretory function; B – platelet adhesion; C – platelet aggregation; D – platelet clot formation.
- Figure 1-4. Intrinsic and extrinsic pathways of the blood coagulation cascade.
- Figure 1-5. The prothrombinase complex. Prothrombin is composed of fragment 1 (F1), fragment 2 (F2) and prethrombin 2 (Pre-2). $FVa = Va$; $FXa = Xa$; calcium ions are filled circles. FVa is embedded into the phospholipid layer.

Figure 1-6.

Figure 1-7.

Figure 1-8.

Figure 1-9.

CHAPTER

Figure 2-1.

Figure 2-2.

Figure 2-3.

Figure 2-4.

- Figure 1-6. Structural motifs of various blood proteins.
- Figure 1-7. Molecule of human LDL. Apo(a) consists of 14 to 41 kringle 4-like units, referred to as KIV, distributed into 10 classes (KIV-1...KIV-10). KIV-9 links apo(a) and apo B-100. KIV-10 is followed by one copy of KV and a protease domain.
- Figure 1-8. Kringle domain. Primary structure of human apo(a) KIV-10.
- Figure 1-9. The photograph of normal (A) and 70% occluded (B) coronary arteries. In latter, white and black streaks represent cholesterol depositions and hemorrhage into the atherosclerotic plug, respectively.

CHAPTER 2: RETRO-BINDING INHIBITION OF THROMBIN

- Figure 2-1. Active site retro-binding inhibitors of thrombin.
- Figure 2-2. Stereoview of the final (2Fo-Fc) electron density of the (A) SEL2711 and (B) SEL2770 inhibitors in the thrombin complexes. Contoured at the 1σ level; SEL2711 is shown in bold black. SEL2770 is shown in bold black. Thrombin residues are in thin gray and numbered.
- Figure 2-3. Stereoview of superimposed Selectide inhibitors in the active site of thrombin. SEL2711 and SEL2770 are shown in bold black and gray, respectively. Thrombin residues of the SEL2711 complex are in thin gray and numbered.
- Figure 2-4. Stereoview of superimposed SEL2711 and BMS-183507 bound to thrombin. SEL2711 and BMS-183507 are shown in bold black and gray,

Figure 2-

Figure 2-

CHAPT

Figure 3-

Figure 3-

respectively. Thrombin residues of the SEL2711 complex are in thin gray and numbered.

Figure 2-5. Stereoview of SEL2711 and SEL2770 superimposed on the active site of FXa in a retro-binding mode. SEL2711 and SEL2770 are shown in bold black and gray, respectively. FXa residues are in thin gray and numbered. Since the pyridine of SEL2770 collides with Phe174 of FXa, a conformational change must occur in one or the other, most likely the former.

Figure 2-6. Stereoview of SEL2711 and SEL2770 superimposed on the active site of FVIIa in a retro-binding mode. SEL2711 and SEL2770 are shown in bold black and gray, respectively. FVIIa residues are in thin gray and numbered.

CHAPTER 3: SELECTIVE NON-ELECTROPHILIC INHIBITORS OF THROMBIN

Figure 3-1. Selective non-electrophilic Molecumetics active site inhibitors of thrombin. Electrophilic inhibitor D-Phe-Pro-Arg chloromethyl ketone (PPACK) is shown for comparison. Numbering is defined on MOL376; K_i is shown for both thrombin and trypsin (upper and lower number, respectively); ratio of K_i 's is on the right indicates a degree of selectivity.

Figure 3-2. The Ramachandran plot of the final structure of the thrombin- (A) MOL376 and (B) MOL592 complexes. Glycine residues are shown as triangles; non-glycine residues as squared boxes.

Figure 3-3.

Figure 3-4.

Figure 3-5.

Figure 3-6.

CHAPTER

Figure 4-1.

Figure 4-2.

Figure 3-3. Stereoview of the final (2Fo-Fc) electron density of the (A) MOL376 and (B) MOL592 inhibitors in the thrombin complexes. Contoured at the 1σ level; MOL376 and MOL592 are in bold black. Thrombin residues are in thin gray and numbered.

Figure 3-4. Stereoview of superimposed MOL354/376/1245 bound to thrombin. MOL354 is in bold black; MOL376 is broken black line; MOL1245 is in bold gray. Thrombin residues of the MOL1245 complex are in thin gray and numbered.

Figure 3-5. Stereoview of superposed MOL356 and MOL592 bound to thrombin. MOL356 is in bold black; MOL592 is in bold blue. Thrombin residues of the MOL356 complex are in thin gray and numbered.

Figure 3-6. Schematic diagram of hydrogen bonded salt bridges of Molecumetics inhibitors with Asp189 of the S1 specificity site of thrombin. MOL354/1245 shown together in one schematic, MOL1245 in parentheses.

CHAPTER 4: PROBING THE ACTIVE SITE OF THROMBIN FURTHER WITH UREA

Figure 4-1. Urea binding sites of thrombin. Ure1- Ure3 are in the active site region of the enzyme. Urea4 and Ure5 are located adjacent to heparin binding exosite II. Ser195 of the catalytic triad is numbered.

Figure 4-2. Stereoview of the final (2Fo-Fc) electron density of Ure1-Ure3 and their water environment. Contoured at the 1σ level.

Figure 4-3

CHAPTER

Figure 5-1.

Figure 5-2.

Figure 5-3.

Figure 5-4.

Figure 5-5.

Figure 4-3. Schematic diagram of hydrogen bonds (Å) formed by Ure1 and Ure2 in the active site region of the thrombin-hirugen complex.

CHAPTER 5: STRUCTURE-FUNCTION DETERMINANTS OF RECOMBINANT KRINGLE 5 OF HUMAN PLASMINOGEN

Figure 5-1. Primary structure of plasminogen K5.

Figure 5-2. Space filling and ribbon representations of plasminogen K5 dimer viewed approximately down the local 2-fold axis in the crystal structure.

Figure 5-3. Superposition of the LigBS of K5 molecules 1 and 2 of the dimer. A residue of each segment is labeled. Molecule 1 is shown in gray. Molecule 2 is colored in black.

Figure 5-4. The Ramachandran plot of the final structure of the K5 dimer. Lys48 residues of molecule 1 (Lys148) and molecule 2 (Lys248) are in a generously allowed region. Glycine residues are shown as triangles; non-glycine residues are shown as squared boxes.

Figure 5-5. Stereoview of specific residues of K5 potentially present in the LigBS of molecule 1 of the K5 dimer. The regions of the kringle represented are His33-Phe36, Asp55-Asp57, Trp62-Tyr64, Leu71-Tyr74 shown in gray. A residue of each segment is labeled. Water molecules common to the LigBS of both kringles of the dimer are 1, 4, 5, 7, and 8; those in molecule 1 only are 2, 3, and 6. 1=O_w540, 2=O_w525, 3=O_w419, 4=O_w565, 5=O_w589, 6=O_w524, 7=O_w526, 8=O_w527 of the PDB coordinate file (PDB code: 5HPG).

Figure 5-6.

CHAPTER

Figure 6-1.

Figure 6-2.

Figure 6-3.

Figure 5-6. Superposition of the LigBS of K1 and K5 of plasminogen. K5 is in gray; K1 is shown in black. K1 is inhibited with EACA in the binding pocket. All residues are numbered according to K5.

CHAPTER 6: RECOMBINANT KRINGLE IV-10 MODULES OF HUMAN APOLIPOPROTEIN (a): STRUCTURES, LIGAND BINDING MODES AND BIOLOGICAL RELEVANCE

Figure 6-1. Superposition of EACA of different lysine binding kringles. KIV-10/Thr66/EACA side chains of the LigBS are in gray. Embedded binding modes are labeled: EACA of K1 (1); EACA of K4 (2); EACA of KIV-10/Thr66/EACA (3). EACA of KIV-10/Thr66/EACA (Sandoz) displays the unbound binding mode. Note different conformations of Arg36 in KIV-10/Thr66/EACA and in KIV-10/Met66.

Figure 6-2. Stereoview of the final (2Fo-Fc) electron density of EACA of KIV-10/Thr66/EACA displaying the embedded binding mode. Contoured at the 1σ level. Two conformations of Arg32 are shown.

Figure 6-3. Superposition of KIV-10/W72R (dark gray) on KIV-10/Met66/EACA (light gray) showing Arg72 approximating EACA binding. Note different conformation of Arg36 in the two structures.

CH

Tab

Tab

CH

Tab

Tab

Tabl

Tabl

CHA

Table

Table

LIST OF TABLES

CHAPTER 1: BLOOD PROTEINS AND COAGULATION

Table 1-1. Morphologic criteria for megakaryocytes and platelets.

Table 1-2. Conserved structural modules in different blood proteins.

CHAPTER 2: RETRO-BINDING INHIBITION OF THROMBIN

Table 2-1. Specificity of Selectide inhibitors ($K_i(\mu\text{M})$).

Table 2-2. Crystal data and intensity data collection statistics of SEL2711 and SEL2770.

Table 2-3. Refinement summary of deviations from ideality of thrombin molecules in the SEL2711 and SEL2770 complexes.

Table 2-4. Intermolecular and intramolecular interactions in the active site of the thrombin complexed with SEL2711 and SEL2770 inhibitors.
Numbering and nomenclature as used in text.

CHAPTER 3: SELECTIVE NON-ELECTROPHILIC INHIBITORS OF THROMBIN

Table 3-1. Intensity data collection of Molecumetics inhibitors MOL376 and MOL592 bound to thrombin.

Table 3-2. Refinement summary of deviations from ideality of thrombin molecules in the MOL376 and NOL592 complexes.

Table 3-3. Individual refinement statistics of MOL376 and MOL592 inhibitors.

CHAPTER 4: PROBING THE ACTIVE SITE OF THROMBIN FURTHER WITH UREA

Table 4-1. Crystal data and intensity data collection statistics of the native and urea-treated thrombin-hirugen.

Table 4-2. R_{merge} -value as a function of the (A) structure factors; (B) resolution; (C) $F/\sigma(F)$ of the native and urea-treated crystals of thrombin-hirugen.

Table 4-3. Refinement summary of deviations from ideality and final refinement parameters of the thrombin-hirugen-urea complex.

CHAPTER 5: STRUCTURE-FUNCTION DETERMINANTS OF RECOMBINANT KRINGLE 5 OF HUMAN PLASMINOGEN

Table 5-1. K_d (μM) values of ligand binding to various kringle modules.

Table 5-2. Primary structures of plasminogen K1-K5 and apo(a) KIV-10 aligned in respect of K1 to maximize conservation. The number of residues (1-80) standardized to plasminogen K5. Deletion sites at positions 34 and 59 are indicated with -. Sequence conservation is boxed: identical residues are in the black boxes, homologous residues are in the gray boxes.

Table 5-3. Primary structures of the plasminogen K5 modules from the various species.

Table 5-4. Crystal data and intensity data collection statistics of plasminogen K5.

Table 5-5. Refinement summary of deviations from ideality and individual

Tab

CH

Tab

Table

Table

Table

Table

Table

refinement statistics of plasminogen K5.

Table 5-6. Hydrogen bonds (Å) in the LigBS region of K5.

CHAPTER 6: RECOMBINANT KRINGLE IV-10 MODULES OF HUMAN APOLIPOPROTEIN (a): STRUCTURES, LIGAND BINDING MODES AND BIOLOGICAL RELEVANCE

Table 6-1. Primary structures of various kringles aligned in respect of apo(a) KIV-10 to maximize conservation. The number of residues (1-80) standardized to plasminogen K5. Deletion sites at positions 34, 42A, 47A, 59 and 70A are indicated with –. Sequence conservation is boxed: identical residues are in the black boxes, homologous residues are in the gray boxes.

Table 6-2. Conditions used to crystallized the various KIV-10 kringles.

Table 6-3. Crystal data and intensity data collection statistics of different KIV-10 crystals.

Table 6-4. Summary of final restrained least-square statistics of KIV-10 modules.

Table 6-5. Probable conserved hydrogen bonds (Å) in LigBS of various KIV-10.

Table 6-6. Probable hydrogen bonds (Å) in LigBS of KIV dependent on ligand binding and the Trp72Arg mutation.

ABBREVIATIONS AND SYMBOLS

- A:** 7-AHpA, 7-aminoheptanoic acid; 6-AHx, 6-aminohexane; t-AMCHA, trans-4-aminomethylcyclohexane-1-carboxylic acid; 5-APn, 5-aminopentane; 5-APnA, 5-aminopentanoic acid; apo(a), apoB-100, apolipoprotein B-100; apolipoprotein(a); Argatroban, (2R,4R)-4-methyl-1-[N^α-[(3-methyl-1,2,3,4-tetrahydro-8-quinolinyl)sulfonyl]-L-arginyl]-2-piperidine carboxylic acid; ATIII, antithrombin III;
- B:** B, temperature factor; BFU-MK, burst-forming-unit megakaryocyte;
- C:** CFU-MK, colony-forming-unit megakaryocyte; Chg, L-cyclohexylglycyl; CSF, colony stimulatory factors;
- D:** d, resolution;
- E:** EACA, ε-aminocaproic acid; EGF, epidermal growth factor; EGR, glutamyl-glycyl-arginyl chloromethylketone;
- F:** F, structure factors; FVa, factor Va; FVII, factor VII; FVIIa, factor VIIa; FIX, factor IX; FX, factor X; FXa, factor Xa; FXII, factor XII; FXIa, factor XIa; FXI, factor XI; FXIII, factor XIII; FXIIIa, factor XIIIa; FVIII-FIX-FX, X-ase complex;
- G:** Gla, γ-carboxyglutamic acid;
- H:** hirugen, sulfated Tyr63-N-acetyl-hirudin 53-64;
- I:** IFN, interferon; IL, interleukins;
- K:** k, scale constant; K1 to K5, the kringle 1 to kringle 5 regions (residues Cys84-Cys541) of human plasminogen; K5[L71R] recombinant Leu71Arg kringle 5 mutant; KIV and KV, kringle 4 and kringle 5 of apo(a); KIV-1, KIV-2,...KIV-10, KIV modules, type 1-10; KL, c-kit ligand;

- L:** LDL, low density lipoprotein; LigBS, ligand binding site; Lp(a), lipoprotein(a);
- M:** MØ, macrophages; Meg-POT, megakaryocytes potentiator; MSF, maturation stimulatory factors;
- N:** NAPAP, N^α-(2-naphthyl-sulfonyl-glycyl)-D-*para*-amidinophenylalanyl-piperidine; NCS, non-crystallographic symmetry; NK, natural killer cell; Nlys(2Me), L-(N,N-dimethyl)lysine; NO, nitrogen (II) oxide;
- P:** Pal3(Me), 3-(methyl pyridinium)alanyl; pAph, L-4-amidino-phenylalanyl; PEG, polyethylene glycol; PGI₂, prostacyclin; PGPMC, platelet glycoprotein-bearing mononuclear cell; PT, prothrombin; ; PPACK, D-phenylalanyl-prolyl-arginyl chloromethylketone; PT-F1, prothrombin fragment 1; PT-F2, prothrombin fragment 2;
- R:** rmsΔ, root-mean-square deviation;
- S:** SPB, sodium phosphate buffer;
- T:** T, T-lymphocytes; TF, tissue factor; t-PA, tissue-type plasminogen activator; TPO, thrombopoietin;
- U:** u-PA, urokinase plasminogen activator;
- V:** vWF, van Willebrand factor;
- X:** X-ase complex, FVII-FIX;
- Ω:** Ω, occupancy parameter.

CHAPTER 1

BLOOD PROTEINS AND COAGULATION

I. HEMOSTASIS

Under normal physiological conditions, blood of higher organisms is in a fluid state. Upon tissue damage or vascular injury, a hemostatic response is primed to arrest hemorrhage, seal off the site of damage and begin a process of tissue repair. The delicate and highly dynamic balance between blood clotting and dissolution is regulated by four major processes: endothelial maintenance, platelet functions, the coagulation cascade (hemostasis) and fibrinolysis.

I.1. ENDOTHELIUM

Endothelium is an interfacing single layer of smooth thin cells between the blood vessel lumen and the circulating blood. This unique location demands that endothelial cells participate in both hemostasis and vascular permeability (Colman et al., 1994). On the one hand, endothelial cells possess an adhesive texture on their outer or abluminal surface, which faces the subendothelial tissues. On the other hand, the luminal surface, which is exposed to the bloodstream, is highly non-thrombogenic. Thus, endothelium serves as a shield for circulating platelets against exposure to subendothelial activation factors and a barrier for separation of blood proteins from the tissue thromboplastin or other surface contact activators that, otherwise, would provoke the coagulation cascade.

Endothelium performs both anticoagulant and procoagulant activities due to its synthetic and secretory functions (Figure 1-1). Among the anticoagulant factors

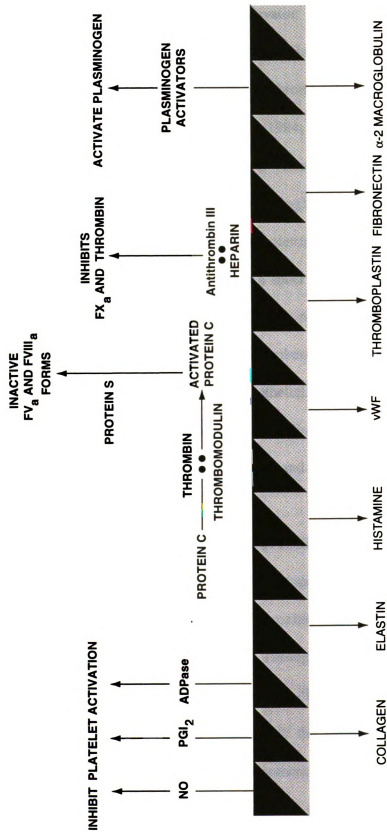


Figure 1-1: Functions of endothelium. Anticoagulant factors and their modes of action in the coagulation cascade are represented in bold black. Coagulation factors secreted by the endothelial cells are in thin line text. Double dot denotes binary complex.

1

a
S
a
h

in
na
lin
for
(C
Pla

synthesized by endothelial cells are prostacyclin (PGI₂), tissue-type plasminogen activator (t-PA) and urokinase plasminogen activator (u-PA), nitrogen (II) oxide (NO), thrombomodulin and heparin. As a procoagulant, endothelium also synthesizes and secretes from its abluminal surface collagen, thromboplastin, fibronectin, von Willebrand factor (vWF), and elastin (Stemerman et al., 1984; Colman et al., 1994).

Endothelium injury could occur in various forms by either direct or indirect induction and range from mechanical (sloughing), physical (freezing, heat or radiation), chemical (carbon monoxide, hypercholesterolemia) to immunological and viral. Under normal physiological conditions, maximum sheer wall stress in the arterial tree is about 200 dynes/cm³. Sheer stress exceeding 400 dynes/cm³ can tear endothelium from the arterial wall exposing the thrombogenic subendothelial matrix to the bloodstream. Sloughing of the non-thrombogenic endothelial layer induced by either mechanical, atherogenic injury, hypertension or thrombotic disorders deprive the vessels of numerous intrinsic anticoagulant protective mechanisms and initiate platelet plug formation.

I.2 PLATELETS

Platelets are derived from their hematopoietic precursor cells, megakaryocytes that have been established to comprise approximately 0.02% - 0.05% of bone marrow nucleated cell population (Harker, 1968). Through a series of poorly understood events, lineage-indifferent stem cells give rise to more differentiated progenitor cells, the burst-forming-unit megakaryocyte (BFU-MK) and the colony-forming-unit megakaryocyte (CFU-MK), that are committed to development within a given lineage (Figure 1-2). Platelet glycoprotein-bearing mononuclear cells (PGPMC) are transitional between the

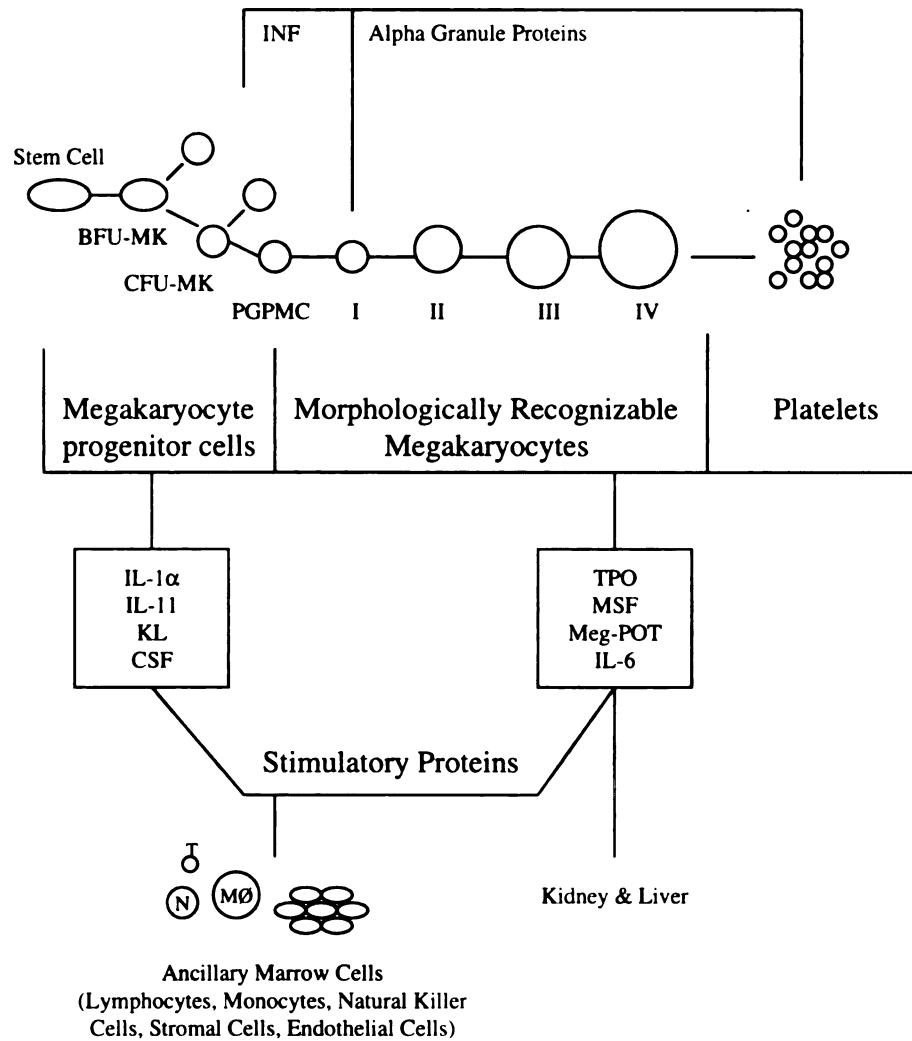


Figure 1-2: Megakaryocyte cell development from the level of the stem cells, through BFU-MK, CFU-MK and then into the morphologically recognizable platelet precursor cell in stages I-IV.

progenitor cells of the first developmental compartment and morphologically recognizable cells of the second developmental compartment where megakaryoblasts, promegakaryocytes and megakaryocytes are undergoing the terminal phases of maturation before releasing platelets into the blood (Table 1-1).

Table 1-1: Morphologic criteria for megakaryocytes and platelets.

Characteristics	Megakaryocytes	Platelets (Thrombocytes)
Cell Size (diameter)	35-160 μm	2-4 μm
Nuclear-Cytoplasmic Ratio	1:1 – 1:3	-
Nuclear Shape	Lobulated (2 or more lobes)	-
Nucleoli	None	-

Some known modulators of human megakaryocytopoiesis are T-lymphocytes (T), macrophages (MØ), natural killer cells (NK), interferon (IFN), c-kit ligand (KL), megakaryocytes potentiator (Meg-POT), maturation stimulatory factors (MSF), colony stimulatory factors (CSF), thrombopoietin (TPO), endothelial cells and various interleukins (IL) (Figure 1-2).

The ability of platelets to form a platelet plug is critical to the maintenance of normal hemostasis (Born & Hardisty, 1972; Colman et al., 1994). Platelet procoagulation activity includes three major steps: activation, adhesion and aggregation (Figure 1-3/A-C). When vascular injury occurs and endothelium is damaged, thrombogenic subendothelial tissues such as collagen and basement membrane are exposed to the

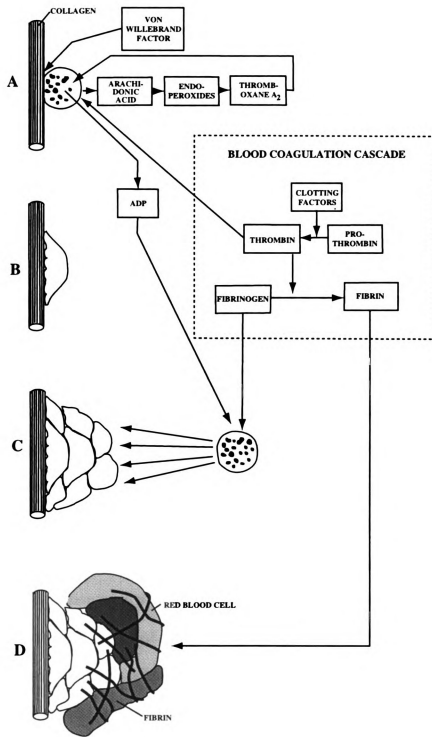


Figure 1-3: Progress of platelet plug formation in hemostasis.
 A-platelet activation and secretory function; B-platelet adhesion;
 C-platelet aggregation; D -platelet clot formation.

bloodstream. Platelets circulating in the blood come into a contact with the exposed subendothelial factors and become activated. Activated platelets: (1) adhere to the surface by the binding of vWF to platelet glycoprotein Ib/IX (2) expose and secrete certain membrane-bound glycoproteins and (3) undergo a change in shape. The new incoming platelets then aggregate at the site of the injury (Figure 1-3/C) through a molecule of fibrinogen linked to glycoproteins IIb/IIIa of two adjacent platelets. Initial platelet plug formation at the site of the endothelial damage is referred to as primary hemostasis.

To form an effective plug, the circulating platelets must be present in adequate numbers ($100-400 \times 10^9/L$) and have normal functions. Decreased platelet counts are referred to as thrombocytopenia and is the most common platelet disorder of insufficient platelet production and excessive removal of platelets from circulation. Failure of sufficient platelet production can result from (1) decreased thrombopoietin, (2) decreased number of megakaryocytes or (3) ineffective thrombopoiesis. The following factors could contribute to removal of platelets from the circulation: (1) platelet destruction, (2) excessive platelet consumption (disseminated intravascular coagulation), (3) loss (hemorrhage) or (4) dilution with platelet-poor blood during massive blood transfusion. Increased platelet counts are classified as thrombocythemia.

While thrombocytopenia is associated with decreased numbers the circulating platelets, thrombocytopathy refers to abnormal platelet functions in adhesion and aggregation. These defects could be classified as inherited or acquired. The inherited defects include Bernard-Soulier Syndrome, pseudo-von Willebrand disease, Glanzmann's thrombasthenia, release disorders and storage pool deficiencies. The acquired defects

occur when the platelet functions are affected by drugs, diet or disease. In thrombocytopathy the platelet count is usually normal or slightly increased, but occasionally both conditions are present simultaneously.

I.3. THE COAGULATION CASCADE

To strengthen the platelet plug formed at the site of vascular injury, a series of enzymatic reactions initially triggered by tissue factor (TF) result in a cross-linked fibrin clot formation (Figure 1-3/D) (MacFarlane et al., 1972). This fibrin network stabilizes the platelet plug against dissolution and is crucial in the coagulation process.

Two similar schemes of the blood coagulation cascade were independently proposed approximately 35 years ago (Davie & Ratnoff, 1964; MacFarlane, 1964). Each proposal suggested two principal mechanisms to stop the loss of blood: these pathways are referred to as extrinsic and intrinsic (Figure 1-4). Present evidence indicates that the extrinsic pathway plays an important role in the initiation of coagulation, while the intrinsic pathway is responsible for growth and maintenance of the fibrin clot.

In general, the coagulation cascade is a series of stepwise reactions that utilize protein molecules circulating in the blood in order to seal an area of endothelial damage and to stop blood loss (Davie et al., 1991). The extrinsic pathway requires TF, a membrane glycoprotein with a molecular weight of approximately 44,000 Da that is located in the tissue adventitia and comes into contact with blood after a vascular injury. The extracellular domain of TF spans from residue Ser1 to Glu219, the trans-membrane region of 23 amino acids extends from Ile220 to Leu242 and the cytoplasmic portion is 21 residues in length (His243 to Ser263) located at the carboxyl end of the protein.

Blood Coagulation Pathways

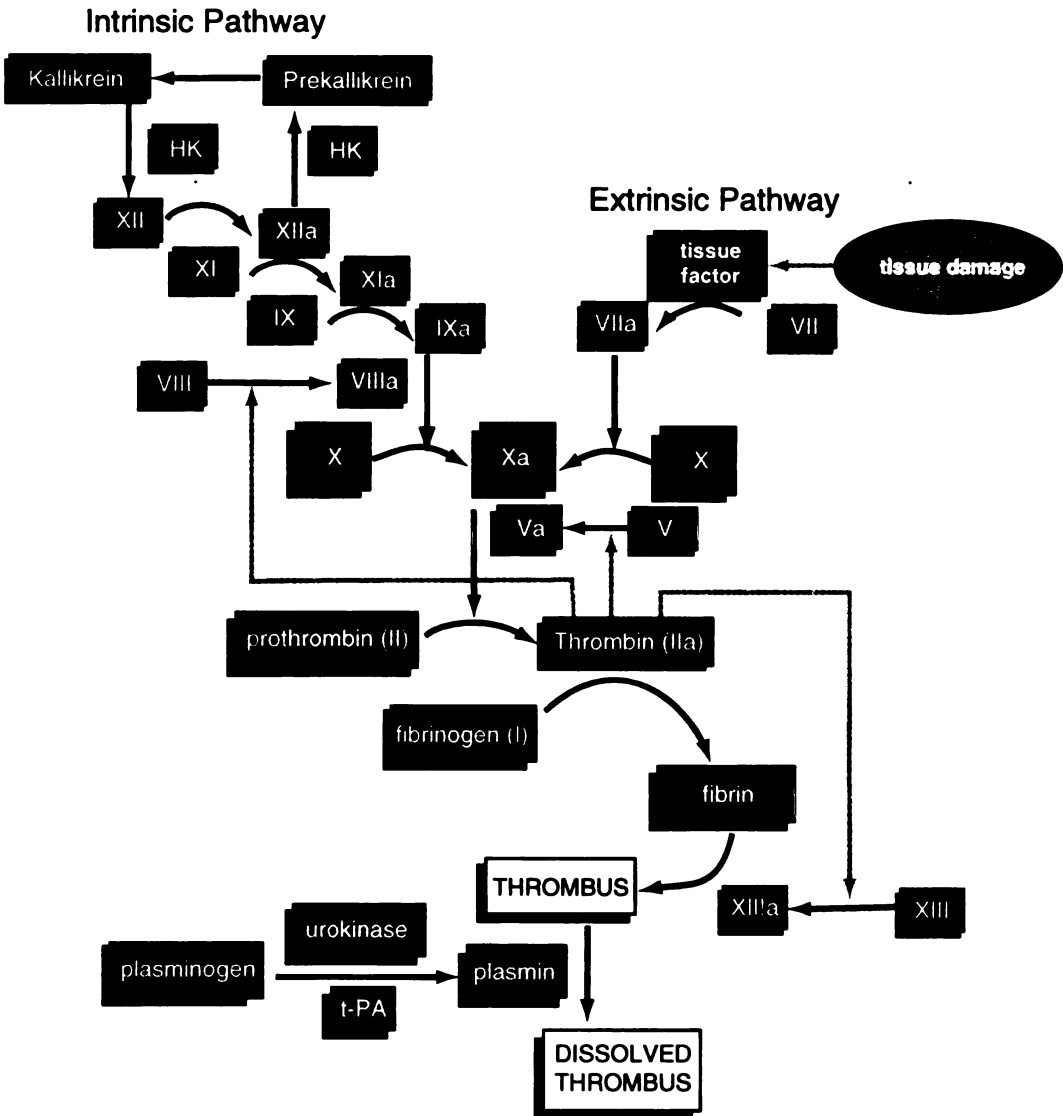


Figure 1-4: Intrinsic and extrinsic pathways of the blood coagulation cascade.

Residue Cys245 is acylated by either palmitic or stearic acid. When a vascular injury occurs and TF is released at the site of the endothelial damage, Factor VII (FVII) forms a one-to-one complex with TF in the presence of calcium ions (Ca^{2+}). This reaction facilitates the conversion of FVII, by minor proteolysis, to an active serine protease, Factor VIIa (FVIIa). Subsequently, the FVIIa-TF complex (Banner et al., 1996; Zhang et al., 1999) triggers the activation of Factor X (FX, Stuart factor) to Factor Xa (FXa) by the cleavage of a single Arg52-Ile53 peptide bond.

In contrast to the extrinsic pathway that is initiated by vascular injury followed by a release of TF, the intrinsic pathway is initiated by the components contained entirely within the vascular system (Figure 1-4). The zymogen Factor XII (FXII, Hageman factor) is the protein that triggers the series of stepwise reactions of the intrinsic pathway. It binds to negatively charged surfaces such as kaolin, dextran sulfate and sulfatides and initiates the sequence of reactions resulting in the activation of Factor IX (FIX) by Factor XIa (FXIa). Thereafter, two pathways are merged at the site of the activation of FX (Figure 1-4).

FXa (Mr. ~ 45,000 Da) is a vitamin K-dependent protein that is generated through limited proteolysis by either the FVIIa-TF complex of the extrinsic pathway (Osterud & Rapaport, 1977; Silverberg et al., 1977) or the X-ase complex (FVIIIa-FIXa) of the intrinsic pathway (Stern et al., 1985). Among all of the obligatory feedback mechanisms of blood coagulation, FXa activates prothrombin to thrombin through a so-called prothrombinase complex (Mann et al., 1990) (Figure 1-5) and can also activate the FVII-TF complex. In the prothrombinase complex, FXa has been shown to be bound to Factor

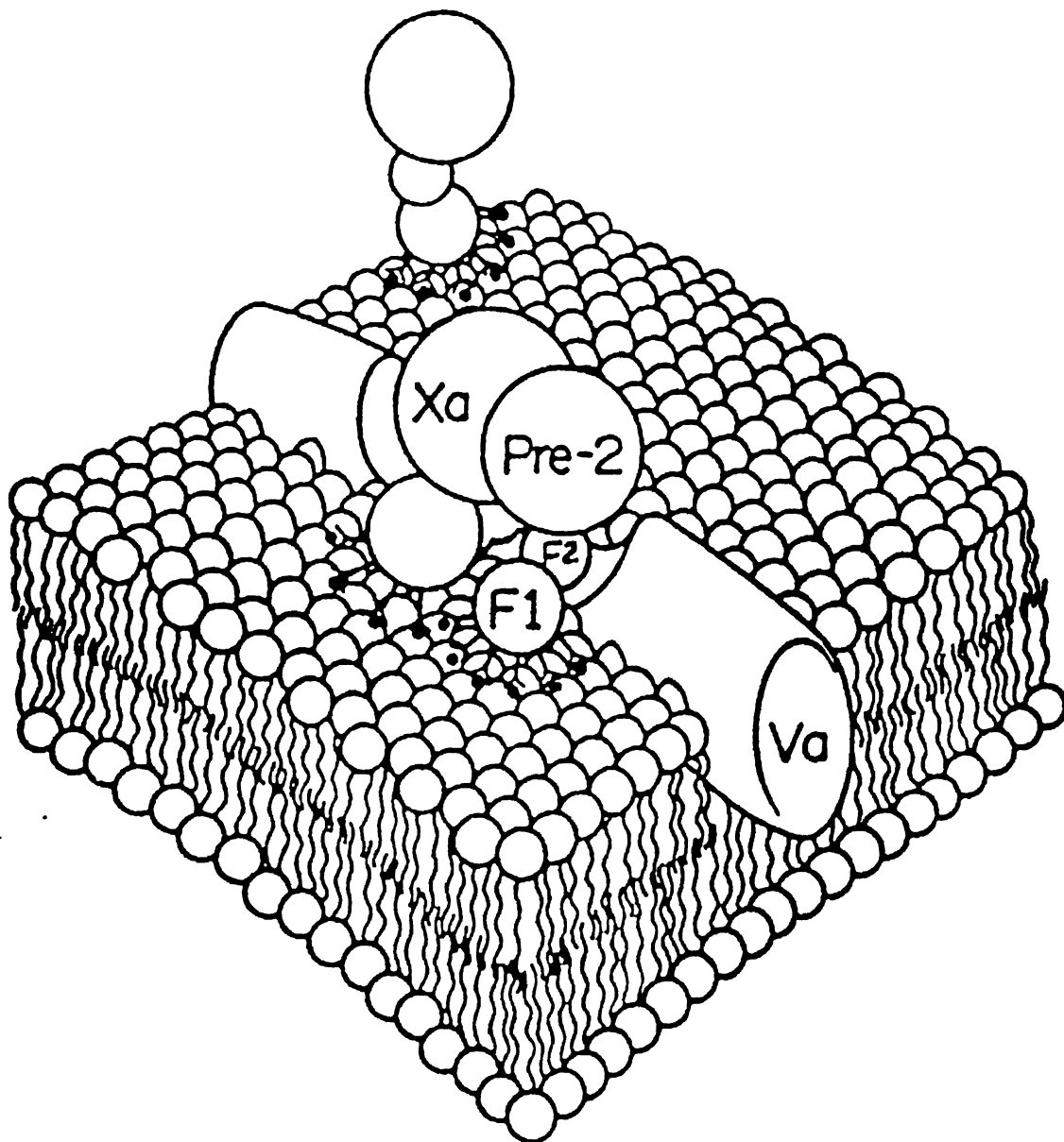


Figure 1-5: The prothrombinase complex. Prothrombin is composed of fragment 1 (F1), fragment 2 (F2) and prethrombin 2 (pre-F2). Factor Va=Va; Factor Xa=Xa; calcium ions are filled circles. FVa is embedded into the phospholipid layer.

Va (FVa) and to the phospholipid surface *via* its γ -carboxyglutamic acid (Gla) containing region. FXa also interacts with a molecule of prothrombin (PT), which is also bound to the phospholipid layer *via* the Gla-domain of fragment 1 (PT F1) and to FVa *via* the fragment 2 region (PT F2 or PT K2).

In the common part of the coagulation cascade, thrombin, generated by through the prothrombinase complex (Figure 1-4) cleaves fibrinogen to form fibrin that leads to the fibrin network of the clot (Figure 1-3/D and 1-4). Factor XIIIa (FXIIIa), generated from its Factor XIII (FXIII) precursor by thrombin in the presence of calcium ions (Lorand & Konishi, 1964; Naski et al., 1991), cross-links fibrin monomers by forming ϵ -(γ -glutamyl)lysine bonds between two adjacent molecules (Folk & Finlayson, 1977). These covalent cross-links result in the formation of the strong fibrin network that stabilizes a pre-formed platelet plug.

II. FIBRINOLYSIS

The phenomenon of clot formation and dissolution is closely balanced and regulated by the hemostatic and fibrinolytic systems. Clot dissolution is necessary for tissue repair and the maintenance of normal blood flow to vital organs. The fibrinolytic system includes: (1) plasma proteins such as the zymogen plasminogen and proteolytically active plasmin, (2) several plasminogen activators and (3) various plasmin inhibitors (Bachmann, 1994).

The essential step in fibrinolysis is the activation of plasminogen to plasmin (Figure 1-4). Plasminogen consists of the N-terminal peptide, an array of five different kringle units (K1-K5) followed by a protease domain that contains a catalytic triad of a

serine protease (Sottrup-Jensen et al., 1978). Two plasminogen activators, t-PA and u-PA, cleave the Arg560-Val561 bond of the zymogen to generate the active serine protease plasmin. Plasmin then degrades the insoluble matrix of the fibrin clot into soluble degradation products.

Physiologically, the fibrinolytic system is designed as a natural response to thrombosis. When thrombotic complications occur, therapeutic use of plasminogen activators becomes important. Due to the fast clearance and low effectiveness of externally administered t-PA, significant quantities of the activator are usually required in practice. However, high dosage of t-PA could lead to hemorrhage, especially in older people with a history of hypertension. From this standpoint, plasminogen activators and inhibitors that can alter delivery mechanisms of plasmin(ogen) to the fibrin clot become important.

III. MODULAR NATURE OF BLOOD PROTEINS

Because of the high sequence and structural homology among various serine proteases participating in the blood coagulation cascade and fibrinolysis, it is likely that proteases originated from a simple trypsin-like ancestor. During evolution, a varying number of modules was inserted between the N-terminal and C-terminal ends giving rise to proteins with highly specialized functions (Figures 1-6 & 1-7, Table 1-2). The structural modules of some blood serine proteases are: (1) Gla domain (2) the finger domain (fibronectin type I and II), (3) the epidermal growth factor (EGF), (4) the apple domain (5) the kringle module and (6) protease or catalytic domain.

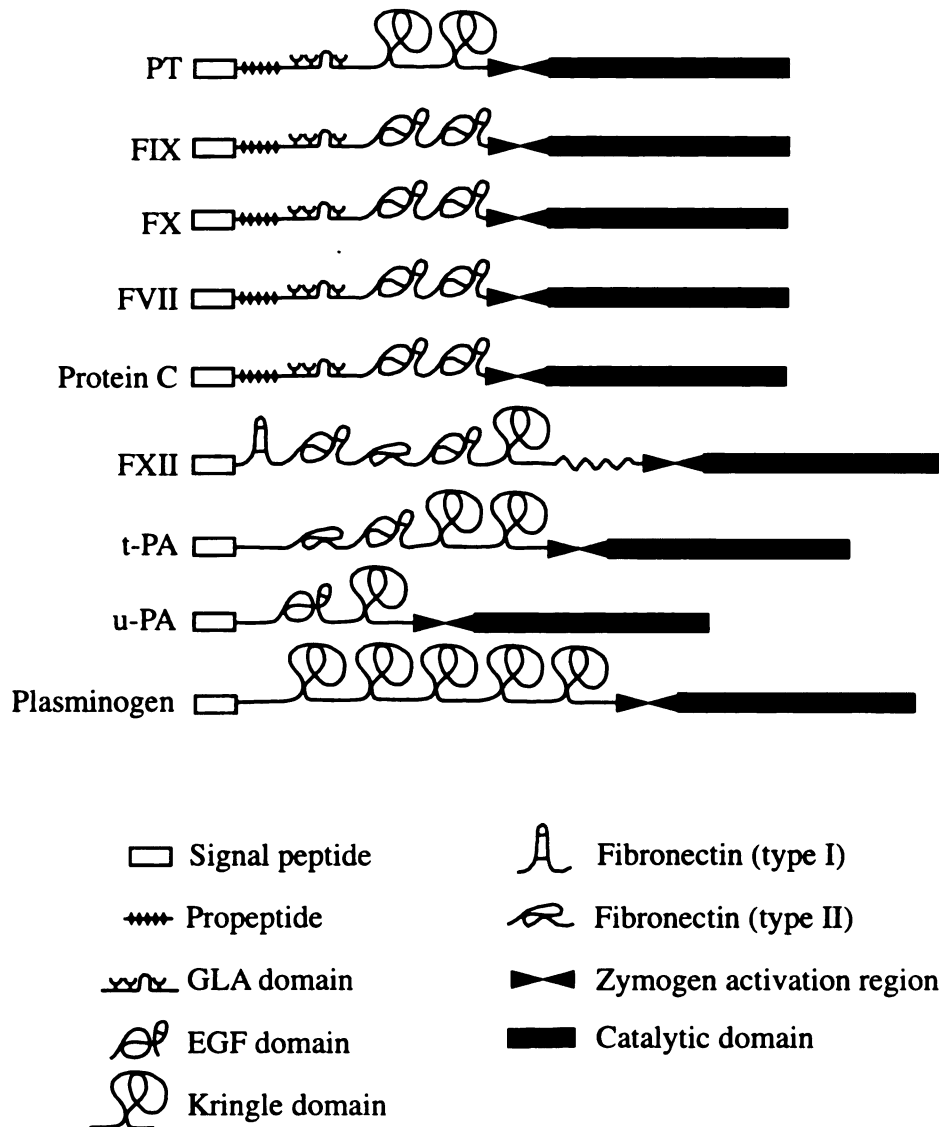


Figure 1-6: Structural motifs of various blood proteins.

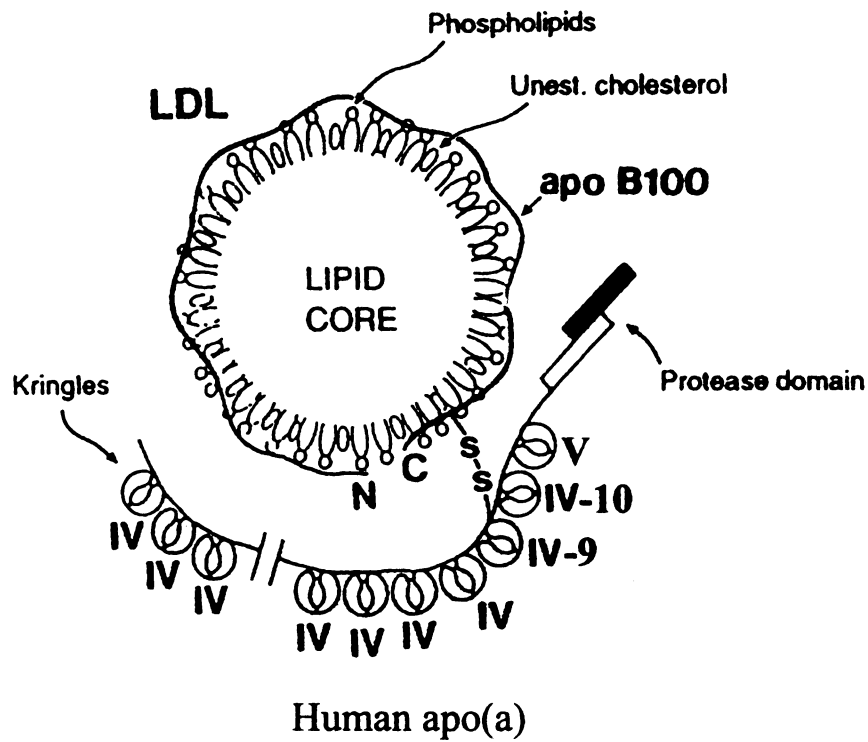


Figure 1-7: Molecule of human LDL. Apo(a) consists of 14 to 41 kringle 4-like units, referred to as KIV, distributed into 10 classes (KIV-1, ... KIV-10). KIV-9 links apo(a) and apo B-100. KIV-10 is followed by one copy of KV and a protease domain.

Table 1-2: Conserved structural modules in different blood proteins.

Structural Module	Blood Proteins
Gla Domain	PT, FVII, FIX, FX, proteins C, S and Z
Finger Domain	Fibronectin, FXII
EGF	FVII, FIX, FX, FXII, proteins C, S and Z, u-PA, t-PA, fibronectin
Apple Domain	Prekallikrein and FXI
Kringle Domain	Apo(a); plasminogen, PT, hepatocyte growth factor, t-PA, u-PA, vampire bat plasminogen activator, angiostatin

(1) The Gla domain is an autonomous structural and functional module of approximately 40 residues found in various blood proteins (Figure 1-6; Table 1-2). The module contains eight to ten Gla residues produced in post-translational modification of glutamic acid by a vitamin K-dependent carboxylase. In the case of the blood proteins, these Gla residues are responsible for binding of proteases to platelets or phospholipid membrane in a calcium (II) dependent manner (Sunnerhagen et al., 1995). It has been shown that in the apo-structure of PT-F1, most of the Gla domain structure is disordered (Park & Tulinsky, 1986; Seshadri et al., 1991). However, when the protein is crystallized in the presence of Ca^{2+} (Soriano-Garcia et al., 1992) or strontium (II) (Sr^{2+}) ions (Seshadri et al., 1994), it exhibits a well folded Gla domain structure. Moreover, based on the fluorescence and circular dichroism (Deerfield et al., 1987), there are two conformational transitions involving two different metal binding sites of the Gla domain: (1) higher

affinity non-cation specific cooperative sites and (2) lower affinity Ca^{2+} and Sr^{2+} ion binding specific sites. However, the mechanism by which these ions promote protein binding to the membrane has yet to be fully elucidated in detail.

(2) The finger domains are homologous to fibronectin type I and type II domains that bind cell surfaces and other molecules including collagen, fibrin, heparin, DNA, and actin (Petersen et al., 1983).

(3) A sequence of about thirty to forty amino-acid residues found in the sequence of epidermal growth factor has been shown (Blomquist et al., 1984; Doolittle et al., 1984; Barker et al., 1986; Appella et al., 1988; Davis, 1990) to be present, in more or less conservative form, in a large number of other proteins including those participating in the coagulation cascade and fibrinolysis (FVII, FIX, FX, FXII, proteins C, S, and Z, u-PA and t-PA) (Table 1-2). The functional significance of EGF domains in these apparently unrelated proteins is not yet clear. A common feature, however, is that these repeats are found in the extracellular domains of membrane-bound proteins or in proteins known to be secreted.

(4) The apple domain, a 90 amino-acid modular unit, has been found only in prekallikrein and Factor XI (FXI) which are two related plasma serine proteases activated by FXIIa (Figure 1-4). The two share the same domain topology: an N-terminal region with a quadruple repeat of the apple domains and a C-terminal catalytic domain. The apple domain contains six conserved cysteines linked together in a pattern 1-6, 2-5 and 3-4 (McMullen et al., 1991a,b).

(5) Kringles are autonomous structural and functional modular units that consist of a characteristic triple loop structure of approximately 80 amino acids cross-linked by

three disulfide bridges in the pattern 1-6, 2-4 and 3-5 (Figure 1-8). Multiple copies of kringle units exist in non-catalytic regions of a variety of blood proteins (Table 1-2). In a striking example there are 14 to 41 kringle copies in apolipoprotein(a) (apo(a)) of lipoprotein(a) (Lp(a)) (Figure 1-7). Plasminogen possesses an array of five highly homologous kringle units (K1-K5) (Figure 1-6) (Sottrup-Jensen et al., 1978). A single copy of a kringle exists in u-PA, vampire bat plasminogen activator and FXII. Kringles are also found as pairs in PT and t-PA and as a quartet in hepatocyte growth factor (Nakamura et al., 1989) and angiostatin (Lijnen et al., 1998).

Kringle domains are also involved in extracellular proteolytic processes such as angiogenesis, atherosclerosis, tumor-growth and metastasis (Cao et al., 1996). In angiogenesis, the formation of new blood vessels supplying nutrients to new cells occurs as a result of the growth of capillaries by a vascular sprouting from pre-existing vessels. Direct experimental studies suggest that tumor growth and metastasis are angiogenesis-dependent processes. An abnormal growth of new blood vessels can result in progression of the tumor, stimulated by angiogenic factors such as fibroblast growth and vascular endothelial growth factors that synergistically promote tumor growth. On the other hand, the process can be highly inhibited by negative angiogenesis regulators such as angiostatin, an endogenous angiogenesis inhibitor of 38,000 Da that consists of the first four kringle units (K1-K4) of plasminogen.

Kringle modules are also found in proteins involved in development of atherosclerosis, another disease threatening the human population in this century. Atherosclerosis, the most common cause of coronary heart disease, myocardial infarction and cerebrovascular disease, is characterized by cholesterol and lipid deposition on the

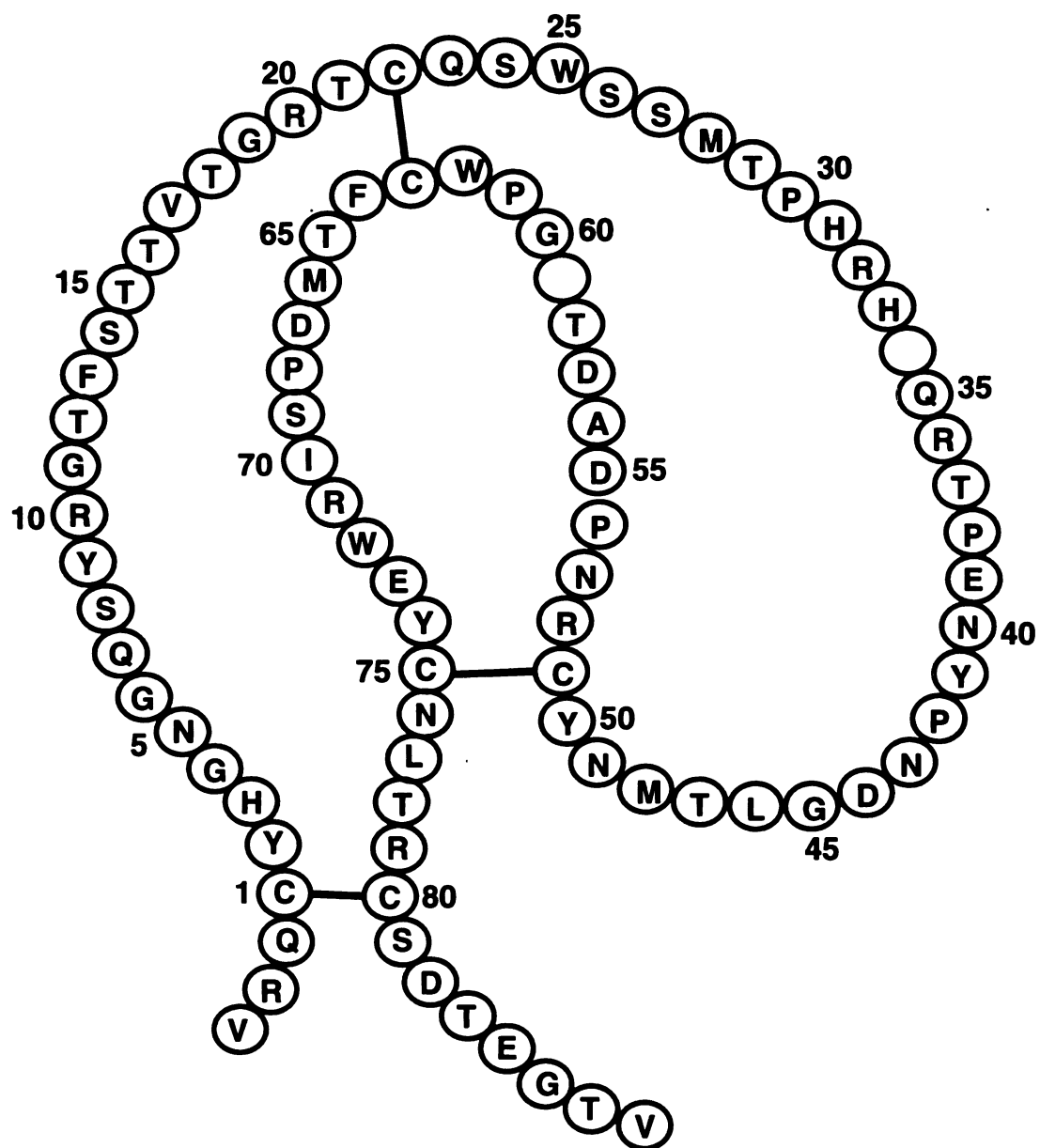
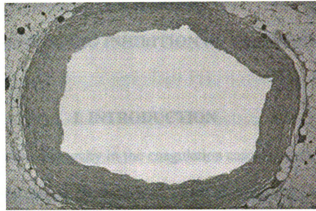
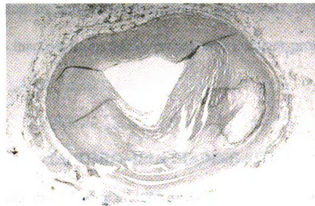


Figure 1-8: Kringle domain. Primary structure of human apo(a) KIV-10

innermost layer of the walls of large and medium sized arteries and results in a vascular obstruction and insufficient blood supply to organs (Figure 1-9). In human subjects, a high incidence of cardiovascular disease has been correlated with high plasma levels of Lp(a) (Dahlen et al., 1978; Kostner et al., 1981; Murai et al., 1986; Dahlen et al., 1986). Lp(a) is a special class of low density lipoproteins (LDL) having as a protein moiety apo(a) covalently linked to apoB100 by a disulfide bond (Figure 1-7) (Brunner et al., 1993; Koschinsky et al., 1993). Kringles are the building blocks of apo(a) the primary sequence of which shares a remarkable degree of homology with those of plasminogen (McLean et al., 1987; Eaton et al., 1987). Apo(a) lacks the first three kringles of plasminogen and contains from 14 to 41 kringle 4-like repeats, referred to as KIV, distributed into 10 classes designated 1 to 10, exhibiting 75% to 94% sequence homology with plasminogen K4. KIV-9 contains an unpaired cysteine that links apo(a) to one of the unpaired cysteines in the C-terminal domain of apoB100. KIV-10 is followed by one copy of KV and a protease domain, which have 91% and 94% sequence homology with K5 and the catalytic domain of plasminogen, respectively (McLean et al., 1987).



(A)



(B)

Figure 1-9: The photograph of normal (A) and 70% occluded (B) coronary arteries. In latter, white and black streaks represent cholesterol depositions and hemorrhage into the atherosclerotic plug, respectively.

CHAPTER 2:

RETRO-BINDING INHIBITION OF THROMBIN

I. INTRODUCTION

Based on the protein hierarchy in the coagulation cascade, thrombin (Mr. ~36,000 Da) is a multifunctional serine protease that plays an important role in both blood coagulation and fibrinolysis (Figure 1-4). As a prominent coagulant, thrombin cleaves fibrinogen, which leads to fibrin, and activates platelets, factors V, VII, VIII and XIII (FV, FVII, FVIII and FXIII) (Colman et al., 1994). Bound to thrombomodulin, thrombin promotes anti-coagulation by activating protein C that converts factor Va (FVa) and factor VIIIa (FVIIIa) to inactive forms in the presence of cofactor protein S (Figure 1-1) (Esmon, 1993).

I.1. THROMBIN BINDING SITES

Thrombin owes its multifunctional role to four distinctive operating sites: (1) the active site, (2) the fibrinogen recognition site (exosite I), (3) the heparin binding site (exosite II) and (4) a sodium (I) (Na^+) ion binding site (Tulinsky, 1996). Like other members of the trypsin-like family, thrombin possesses a catalytic triad of His57-Asp102-Ser195 (chymotrypsinogen numbering, Bode et al., 1992) at the active site surrounded by the Ser214-Cys220 structural segment that forms a loop consisting of two almost orthogonal arrays. The equatorial array of Ser214-Gly216 is critical in its association with the main chain of substrates or peptidic inhibitors by forming an anti-parallel

(substrate binding) (Bode et al., 1989; Mathews et al., 1994) or parallel (retro-binding) β -sheet (Tabernero et al., 1995; Mochalkin & Tulinsky 1999). Both retro- and substrate-like binding active-site inhibitors of thrombin usually possess two or three hydrogen bonds via Ser214-O, Gly216-N and Gly216-O. The axial array of Glu217-Cys220 restricts a specificity pocket that contains Asp189 at its bottom, a key residue of thrombin in the recognition event. An insertion loop at position 60, which partially occludes the active site at the top, is responsible for some additional restricted active site specificity of thrombin compared to that of trypsin or FXa (Locht et al., 1997).

Thrombin fibrinogen binding exosite I is responsible for the association of the natural substrate fibrinogen (Blomback, 1967) and the potent natural inhibitor hirudin (Rydel et al., 1991). The hirudin molecule simultaneously binds to the exosite I and the active site of thrombin. On the other hand, hirugen, a useful fragment of hirudin (sulfated Tyr63-N-acetyl-hirudin 53-64), binds solely to exosite I and thus prevents autocleavage of thrombin with the active site unoccupied. Thrombin heparin binding exosite II binds heparin, a prominent antithrombotic agent that causes a 50,000-fold increase in the rate of inhibition of thrombin by antithrombin III (ATIII) (Olson & Bjork, 1992). An allosteric change from a kinetically slow to a fast form as a result of binding Na^+ at a sodium binding site of thrombin affects the energetics of the active site of the enzyme (Wells & Di Cera, 1992). While fibrinogen binds to the fast form with high affinity (Mathur et al., 1993) and is cleaved with higher specificity, the slow form activates the anticoagulant protein C more specifically (Dang et al., 1995). Two Na^+ binding sites have been identified and described crystallographically with the aid of diffusion-exchange of the

Na⁺ ion by a rubidium (I) (Rb⁺) ion (Zhang & Tulinsky, 1997), one intramolecular, the other intermolecular.

1.2. THROMBIN ACTIVE SITE INHIBITORS AND THEIR BINDING MODES

Since the determination of the crystal structure of thrombin (Bode et al., 1989), an immense array of various thrombin inhibitors has appeared (Balasubramanian, 1995). Selective inhibition of thrombin has been the principal target of antithrombotic drug design ventures aimed at essentially replacing the clinical use of heparin, which is neither highly optimal, convenient nor controllable. There are numerous serine proteases unrelated to the coagulation cascade that play other important physiological roles in life processes. Such enzymes could be accidental targets of less discriminating thrombin inhibitors and could lead to unwanted side effects.

All peptidic thrombin active site inhibitors can be divided into two principal classes according to their binding orientations with respect to residues Ser214-Gly216 of thrombin (antiparallel or parallel β -strand). The antiparallel group, also operative with factors VIIa, IXa and Xa (FVIIa, FIXa and FXa), consists of a variety of inhibitors that mimic natural substrate fibrinogen at the P1-P3 positions, the archetype for thrombin being D-phenylalanyl-prolyl-arginyl chloromethylketone (PPACK) (Bode et al., 1992) and glutamyl-glycyl-arginyl chloromethylketone (EGR) for FXa. There are additionally a large number of other small molecule peptide inhibitors of thrombin (Balasubramanian, 1995). This group includes Argatroban ((2R,4R)-4-methyl-1-[N ^{α} -(3-methyl-1,2,3,4-tetrahydro-8-quinoliny)sulfonyl]-L-arginyl]-2-piperidine carboxylic acid) and NAPAP (N ^{α} -(2-naphthyl-sulfonyl-glycyl)-D-*para*-amidinophenylalanyl-piperidine) (Brandstetter

et al., 1992) that form an antiparallel β -strand even though their binding mode is different at the S1'-S3 subsites (Mathews & Tulinsky, 1995). The second class of thrombin inhibitors consists of hirudin-like active site binding inhibitors (Rydel et al., 1991) that show a retro-binding mode of association by forming a parallel β -strand with Ser214-Gly216. This class inhibitors includes the synthetic peptide BMS-183507 (Tabernero et al., 1995), the natural product Nazumamide A (Nienaber et al., 1996) (Figure 2-1) and a RPPGF breakdown fragment of bradykinin (unpublished results of this laboratory).

I.3. SELECTIDE CORP. INHIBITORS

A class of tri- to octa-peptide competitive inhibitors of FXa has recently been identified from the screening of a large combinatorial library (Ostrem et al., 1998). The consensus sequence is Tyr-Ile-Arg-X, where X is a hydrophobic amino acid and isoleucine is replaceable with leucine. The minimal sequence is the first three residues and may be replaced by homologous non-natural amino acids. Some of these have been shown to inhibit hydrolysis of chromogenic substrates of FXa but bind to the enzyme differently from substrate such that Ser195 of the catalytic triad does not attack the carbonyl group at the P1 position (Ostrem et al., 1998). Selectide Corp. (a subsidiary of Hoechst Marion Roussel) inhibitors SEL2711 and SEL2770 (Figure 2-1) are members of this class of combinatorially designed and potent active site inhibitors of FXa that are also highly selective against thrombin and other blood proteases (Table 2-1), exhibiting dose-dependent efficacy following both intravenous and oral administration in a rat

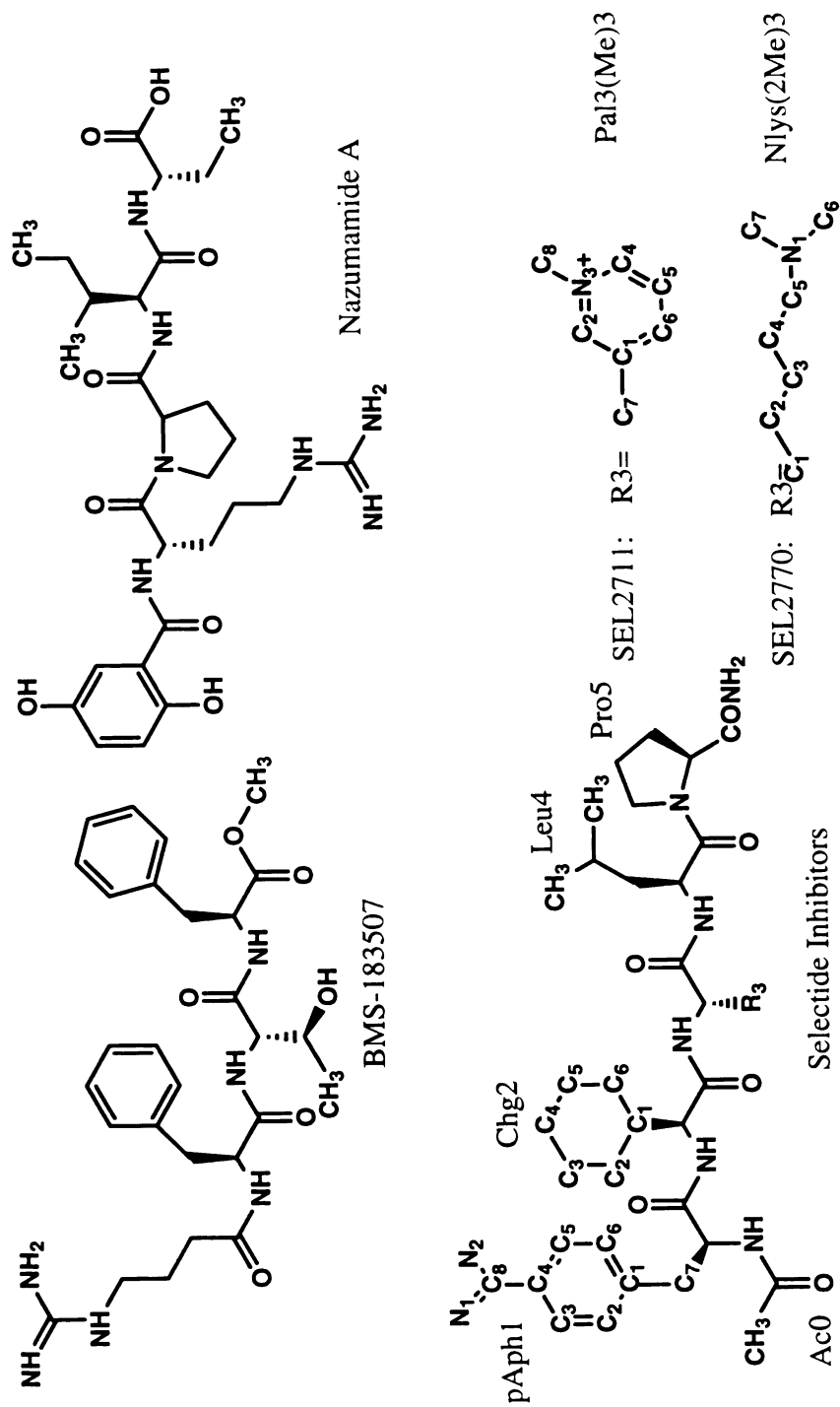


Figure 2-1: Active site retro-binding inhibitors of thrombin.

arteriovenous shunt thrombosis model (Seligmann et al., 1995). A reverse binding mode similar to N-terminal hirudin active site binding to thrombin was suggested to explain the selectivity of this class of inhibitors (Ostrem et al., 1998). However, from an NMR study of two of the inhibitors, Ac-Tyr-Ile-Arg-Ile-Pro-NH₂ and Ac-(4-amino-Phe)-(cyclohexyl-Gly)-Arg-NH₂ (peptides A and B, respectively), bound to the active site of FXa (Fraternali et al., 1998), the authors state that "the inhibitors assume a compact, very well defined conformation, embedded into the substrate binding site not in the same way as a substrate", which was also not a retro-binding mode but with arginine inserted in the S1 specificity site.

Table 2-1: Specificity of Selectide inhibitors (K_i(μM))

Proteins	SEL2711	SEL2770
FXa	0.003	0.002
Thrombin	40	8
Protein C	10	12
Plasmin	130	60
FVIIa/TF	>200	>200
Trypsin	112	NA

Diffusion of the SEL2711 and SEL2770 inhibitors into crystals of FXa proved unsuccessful, probably because, in tetragonal crystals, the active site is blocked by a substrate-like intermolecular interaction (Padmanabhan et al., 1993). On the other hand, co-crystallization attempts with the inhibitors also did not produce crystals. This led us to soak the inhibitors into thrombin-hirugen crystals in an effort to possibly establish and corroborate the different binding mode in thrombin and then model it as it might occur in FXa.

II. EXPERIMENTAL

Crystallization: An approximately 10-fold molar excess of hirugen was added to a frozen sample of human α -thrombin solution (1.0 mg/mL in 0.75 M NaCl) at 4° C to form a 1:1 thrombin-hirugen complex and prevent autocleavage of the enzyme. The solution of the complex was concentrated to about 3.5 mg/mL using a Centricon 10 concentrator with a molecular weight cutoff of 10 kDa in a refrigerated centrifuge at 4° C. The thrombin-hirugen complex was crystallized using the hanging drop crystallization method in a 4 μ L drop, which consisted of 2 μ L of the protein sample and 2 μ L of a well solution. The well solution contained 28% polyethylene glycol (PEG) 8000 in 0.1 M sodium phosphate buffer (SPB), pH 7.3. A repetitive macroseeding technique was applied to enlarge the crystals. X-ray diffraction quality crystals with dimensions of 0.25 x 0.20 x 0.20 mm were transferred into a protein-free storing drop containing 32 % PEG 8000 in 0.1 M SPB, pH 7.3. The SEL2711 and SEL2770 inhibitors were soaked into thrombin-hirugen crystals separately. A 1 μ L aliquot of a 20 mM solution containing the inhibitor was added to the storing drop and 1 μ L of the resulting solution was removed

from the drop in intervals of 10-12 hours for 5 days until the final concentration of inhibitor in the drop was about 15 mM.

Intensity Data Collection: X-ray diffraction data of the ternary SEL2711 and SEL2770 thrombin-hirugen complexes were collected with an R-AXIS II imaging plate detector. The radiation generated from a Rigaku RU200 rotating anode operating at 5 kW power with the fine focus filament (0.3 mm x 3.0 mm) was monochromated ($\text{CuK}\alpha$) and intensified by focusing with Molecular Structure Corp. - Yale mirrors. The crystal to detector distance was 10.0 cm and the detector-swing angle was zero degrees. Both crystals scattered X-rays to 2.1 Å resolution although the diffraction pattern of the SEL2770 complex was considerably weaker (Table 2-2). Autoindexing and processing of the measured intensity data were carried out with the Rigaku R-AXIS II software package (Higashi, 1990). The intensity data collection statistics are summarized in Table 2-2.

Structure Determination: The crystal structure of the SEL2711 complex was determined using thrombin coordinates of the isostructural binary thrombin-hirugen complex (PDB code: 1HAH) (Vijayalakshmi et al., 1994). The initial coordinates were optimized by rigid-body refinement using the X-PLOR program package (Brunger, 1992) to an R-value of 30.4% (7.0-2.5) Å followed by 3 cycles of coordinate and overall temperature factor (B) refinement with a starting temperature factor $B = 30.0 \text{ Å}^2$ using the program PROFFT (R-value = 27.4%) (Hendrickson, 1985; Finzel, 1987). After 6 cycles of individual B refinement, the R-value converged at 21.6%.

Table 2-2: Crystal data and intensity data collection statistics of SEL2711 and SEL2770.

	SEL2711	SEL2770
Space group	C2 ^a	C2 ^a
Cell constants (Å):		
a	71.05	71.66
b	71.82	72.00
c	73.45	73.45
β (deg)	101.2	101.2
Resolution (Å)	2.1	2.1
Observations (I/σ(I)>1.0)	28,464	24,962
I/σ(I) (outermost range) (2.3-2.1) Å	3.2	2.3
Independent reflections (I/σ(I)>1.0)	17,688	13,507
Redundancy	1.6	1.8
Completeness (%)	78	60
outermost range (%)	52	37
R-merge (%)	6.9	4.9
outermost range (%)	14.8	14.3

a - Four ternary complexes per unit cell, one per asymmetric unit.

The crystal structure of the SEL2770 complex was solved with the coordinates of thrombin of the refined thrombin-SEL2711 structure as a starting structure. After 3 cycles of overall and 6 cycles of individual B refinement, the R-value was 20.0% (7.0-2.5) Å. Both complexes at this stage showed electron density corresponding to hirugen and the active site inhibitors (which were omitted from calculations) in (2Fo-Fc) electron and (Fo-Fc) difference density maps. Thereafter, the data were expanded to (9.0 – 2.5) Å and then 2.1 Å resolution and the Na⁺ ions were located and solvent water molecules were progressively found and added by examining difference density maps. The final structure of thrombin complexed with SEL2711 contains 131 molecules of water with occupancies (Ω) > 0.5, R=16.5% (9.0-2.1) Å and $\langle B \rangle = 26.8$ Å². The final SEL2770 structure contains 104 water molecules with Ω > 0.5 refined to R=16.7% (9.0-2.1) Å and $\langle B \rangle = 34.8$ Å². Both complexes also have two Na⁺ ions, one intramolecular and one intermolecular essentially identical to those initially described (Zhang & Tulinsky, 1997). The final refinement statistics of the complexes are presented in Table 2-3. The coordinates of the ternary complexes have been deposited in the Brookhaven Protein Data Bank (SEL2711, access PDB code: 7KME; SEL2770, access PDB code: 8KME).

Table 2-3: Refinement summary of deviations from ideality of thrombin molecules in the SEL2711 and SEL2770 complexes.

Restraints	Target	SEL2711	SEL2770
Distances (Å):			
Bond distance	0.020	0.011	0.014
Angle distance	0.030	0.037	0.036
Planar 1,4 distance	0.050	0.047	0.046
Planes	0.040	0.034	0.030
Chiral volume (Å ³)	0.15	0.17	0.17
Non-bonded contacts (Å)			
Single-torsion	0.60	0.22	0.24
Multiple torsion	0.60	0.25	0.28
Possible H-bond	0.60	0.25	0.28
Isotropic thermal factors (Å ²)			
Main-chain bond	1.0	1.0	0.9
Main-chain angle	1.5	1.7	1.5
Side-chain bond	2.0	2.5	2.1
Side-chain angle	3.0	3.6	3.0

III. RESULTS

Thrombin: The electron density is well defined for most of the residues of thrombin in both complexes except for the termini of the A chain (residues Thr1H-Ala1B, 1Ile4K-Arg15) and the autolysis loop (residues Trp148-Lys149E), which are typically disordered in other isomorphous thrombin structures. The two complexes show no large conformational differences from each other as well as from other thrombin structures. The root-mean-square deviation ($\text{rms}\Delta$) of CA positions between the two structures is 0.25 Å for 253 atom pairs. Comparing the thrombin structures of the ternary complexes with that of the binary thrombin-hirugen complex (PDB code: 1HAH) gives $\text{rms}\Delta$'s of 0.21 Å and 0.30 Å for the SEL2711 and SEL2770 complexes, respectively. The Ramachandran plots (Laskowsky et al., 1993) of the SEL2711 and SEL2770 complexes show that 196 (81.7 %) and 189 (78.8 %) residues of the total 240 non-glycine and non-proline amino acids of thrombin are in most favorable regions while 44 (18.3 %) and 51 (21.2 %) residues occupy additionally allowed areas.

Two Na^+ ion-binding sites, one (intramolecular) responsible for the slow to fast kinetic transition (Wells & DiCera, 1992), and the other intermolecular, have been identified in both structures. The Na^+ ions adopt distorted octahedral arrangements coordinated by solvent water molecules and the carbonyl oxygen atoms of Arg221A, Lys224 (intramolecular ion) and from Lys169, Thr172 of one molecule and Phe204A of a symmetry related molecule for the intermolecular site and are the same as those described previously (Zhang & Tulinsky, 1997). The occupancies and the temperature factors of the Na^+ ions of the SEL2711 and SEL2770 complexes are: $\Omega=1.00$, $B= 23.1 \text{ Å}^2$ and

$\Omega=0.79$, $B=27.6 \text{ \AA}^2$ (intramolecular) and $\Omega=0.84$, $B=22.9 \text{ \AA}^2$ and $\Omega=0.82$, $B=27.2 \text{ \AA}^2$ (intermolecular), respectively.

Selectide Inhibitors: Both SEL2711 and SEL2770 bind to thrombin in a retro-fashion forming a parallel β -strand with residues Ser214-Gly216. The electron density of SEL2711 is well defined except that there is none corresponding to the C-terminal leucine and proline residues (Figure 2-2/A). The (2Fo-Fc) electron density map of SEL2770 has a break at the amide between the L-(N,N-dimethyl)lysine (Nlys(2Me)3) group (nomenclature and numbering used here is the same as that in Ostrem et al., 1998) and the leucine residue and although the density corresponding to C-terminal dipeptide is weaker than the remainder of the inhibitor, it is sufficient to trace the conformations of the leucine and proline (Figure 2-2/B).

As in most crystal structures of thrombin complexed with active site inhibitors containing an arginine-like residue, the specificity site of the enzyme is occupied by the arginine-like L-4-amidino-phenylalanyl (pAph1) residue in both complexes and utilizes doubly hydrogen bonded ionic interactions between the amidino group and side chain of Asp189 (Figure 2-3, Table 2-4). In addition to this hydrogen bonded salt bridge, one of the nitrogen atoms of the amidino group also makes contact with Gly219-O and Ala190-O while the other nitrogen forms a hydrogen bond to a water molecule that bridges to Phe227-O (Table 2-4). The Glu192 residue of thrombin adopts an extended conformation into the solvent region, instead of the more common bent one covering the S1 binding site (Vijayalakshmi et al., 1994), and provides space for the L-cyclohexylglycyl (Chg2) residues of SEL2711 and SEL2770 to occupy the

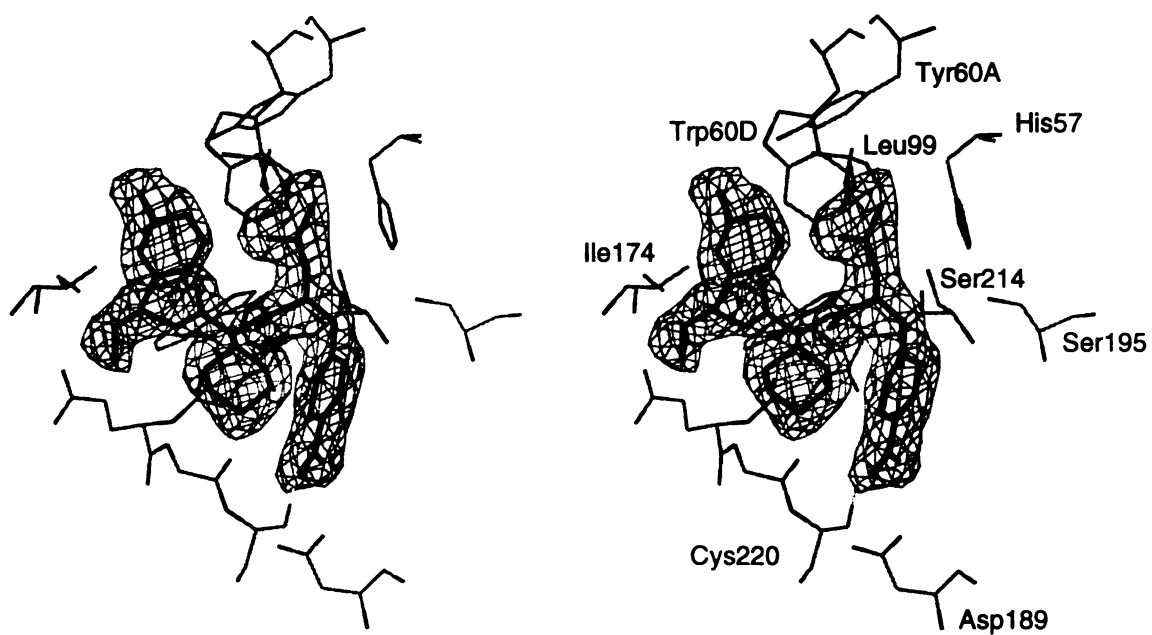


Figure 2-2/A: Stereoview of the final (2Fo-Fc) electron density of the SEL2711 inhibitor in the thrombin complex. Contoured at the 1σ level; SEL2711 is shown in bold black. Thrombin residues are in thin gray and numbered.

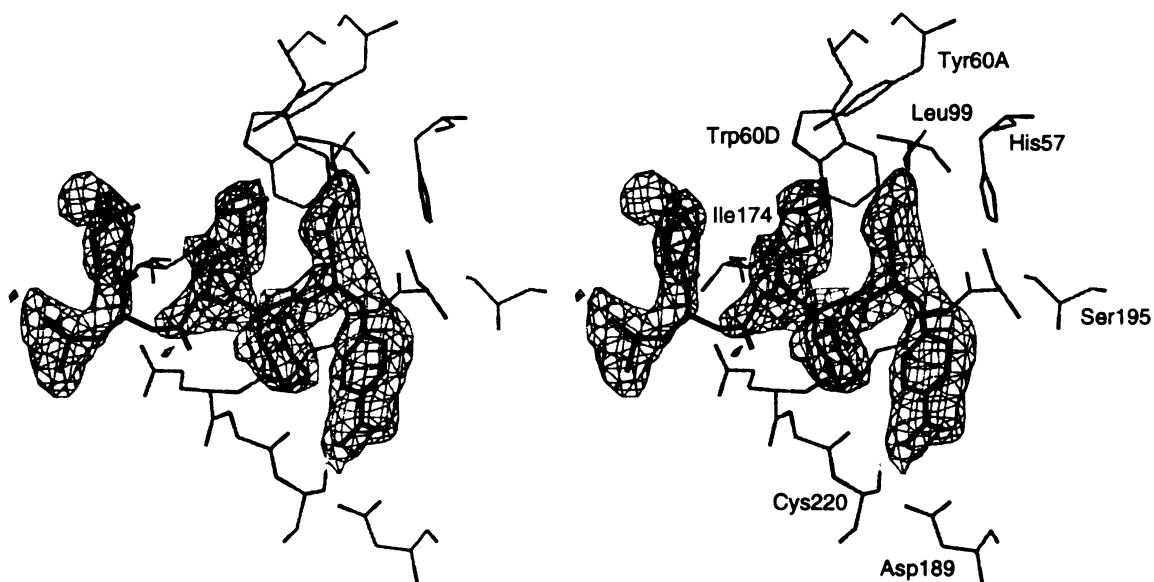


Figure 2-2/B: Stereoview of the final (2Fo-Fc) electron density of the SEL2770 inhibitor in the thrombin complex. Contoured at the 1σ level; SEL2770 is shown in bold black. Thrombin residues are in thin gray and numbered.

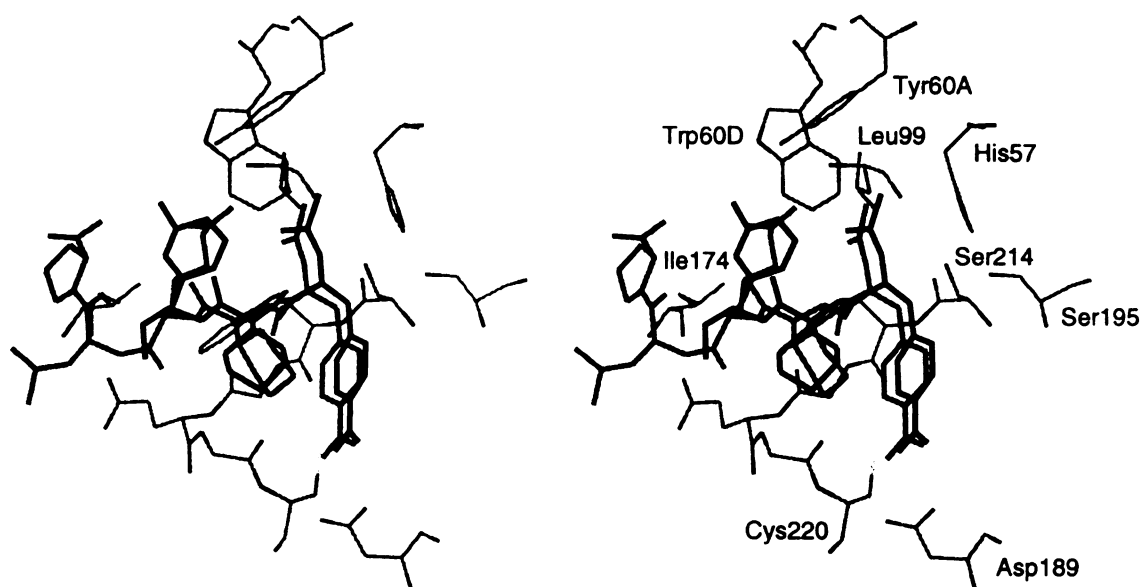


Figure 2-3: Stereoview of the superimposed Selectide inhibitors in the active site of thrombin. SEL2711 and SEL2770 are shown in bold black and gray, respectively. Thrombin residues of the SEL2711 complex are in thin gray and numbered.

Table 2-4: Intermolecular and intramolecular interactions in the active site of the thrombin complexed with SEL2711 and SEL2770 inhibitors. Numbering and nomenclature as used in text.

Intermolecular		SEL2711	SEL2770	Type
pAph1-O	Gly216-N	2.91	2.82	H-bond
-N1	Asp189-OD2	2.89	3.29	H-bond/ion pair
-N2	Asp189-OD1	3.07	2.95	H-bond/ion pair
-N1	Ow438/428	3.07	3.04	H-bond
-N2	Gly219-O	3.28	2.62	H-bond
-N2	Ala190-O	2.87	3.16	H-bond
Phe227-O	Ow438/428	3.57	3.16	H-bond
Pal3(Me)3-N	Gly216-O	2.69	-	H-bond
-N3	Ow455	3.32	-	Ion/dipole
Nlys(2Me)3-N	Gly216-O	-	3.07	H-bond
-N1	Tyr60A-OH	-	3.60	Ion dipole
Leu4-O	Ow679	-	2.85	H-bond
Intramolecular				
Ac0-O	Nlys(2Me)3-N1	-	3.43	Ion/dipole
Chg2-O	Nlys(2Me)3-N1	-	3.46	Ion/dipole

proximity of the L-enantiomeric-S3 site of thrombin similar to the P3 aspartate residue of a thrombin-thrombin platelet receptor peptide complex (Mathews et al., 1994). The Chg2 residues of both peptides are in the lower energy chair conformation; however, they differ in orientation around the CA-CB bond by about 90°. The N-terminal acetyl group of the inhibitors is located in the hydrophobic S2 site of thrombin while the 3-(methyl pyridinium)alanyl residue (Pal3(Me)3) of SEL2711 and Nlys(2Me)3 of SEL2770 are in the aryl D-enantiomeric S3 site; the pyridyl orientation of the former is approximately orthogonal to the plane of D-phenylalanyl of PPACK in the thrombin-PPACK complex (Bode et al., 1992). In the case of SEL2770, the site is occupied by Nlys(2Me)3 in a partially extended conformation (Figure 2-3). The two C-terminal residues, leucine and proline, point into solvent region and have no interaction with the thrombin molecule and appear flexibly disordered in the SEL2711 complex, although they are reasonably fixed in the SEL2770 complex high B-values notwithstanding (Figures 2-2 & 2-3). This is somewhat disturbing since crystals of the SEL2711 complex produced a better diffraction pattern (Table 2-2). Cleavage of the peptide in a substrate binding mode prior to retro-binding is a possibility but highly unlikely with a Pal3(Me)3 group at the P1 specificity position (see also below).

IV. DISCUSSION

Because of the structural similarity of SEL2711 and SEL2770, the binding modes of both are similar (Figure 2-3). While PPACK and many other substrate-like inhibitors

show an anti-parallel β -strand association with Ser214-Gly216 of thrombin, SEL2711 and SEL2770 form a parallel β -strand with these residues making hydrogen bonds with the nitrogen and carbonyl oxygen atoms of Gly216 (Table 2-4). The anti-parallel β -strand association of PPACK in peptidic thrombin inhibition is more abundant because its hydrogen bonding pattern is energetically preferable compared to that in the parallel β -strand alignment of SEL2711 and SEL2770. The parallel β -strand arrangement, however, displays a high degree of fidelity in its positioning in the active site when compared among other known retro-binding thrombin inhibitors.

The largest structural difference in the binding of the two similar inhibitors with thrombin appears at the S3 site of the enzyme where the more flexible Nlys(2Me)₃ group of SEL2770 is able to adopt a favorable conformation to interact internally with itself and thrombin (Figure 2-3). Superposition of the two inhibited structures shows that the quaternary nitrogen atoms of Nlys(2Me)₃ and Pal3(Me)₃ are 3.3 Å apart. However, in the case of the former, free bond rotations place the positively charged nitrogen atom in an electronegative environment consisting of the carbonyl oxygen atoms of the acetyl group and Chg2O of SEL2770, as well as Tyr60A-OH of thrombin (Table 2-4). If the Nlys(2Me)₃ group were fully extended, it could approach the carbonyl oxygen atom of Glu97A within 3.0 Å. Thus, the ion-dipole interactions of SEL2770 appear to be more stabilizing than this possible hydrogen bond. With much less rotational freedom, the positively charged nitrogen atom of Pal3(Me)₃ cannot orient in a direction toward the same internal electronegative environment utilized by Nlys(2Me)₃ and thus achieve a comparable interaction (Pal3(Me)₃-N3 to AcO-O1 = 5.0 Å; to Chg2-O, 5.1 Å; to Tyr60A-OH, 4.7 Å; compare with corresponding Nlys(2Me)₃-N1 distances in Table 2-4).

The Chg2 residues of both inhibitors occupy the L-enantiomeric-S3 region of thrombin similar to aspartate of Leu-Asp-Pro-Arg in a thrombin-thrombin platelet receptor peptide complex (Mathews et al., 1994). The L-S3 site is surface exposed and of considerable size (Figure 2-3): the Chg2 residues do not superpose on the aspartate position of the former complex but rather are located adjacent to it and even appear to be in approximately orthogonal orientations (Figure 2-3) ($\chi_1 = 173^\circ$ and 71° for SEL2711 and SEL2770, respectively). Both the Chg2 groups of SEL2711 and 2770 and the aspartate group of thrombin platelet receptor peptide point into solvent space. The aspartate, however, makes a water mediated hydrogen bond with Arg221A to stabilize its orientation (Ni et al., 1992; Mathews et al., 1994), while the Chg2 groups cover the S1 site and appear to interact with phenyl pAph1 group in the specificity site and the disulfide bridge of Cys191-Cys220. Thus, the Chg2 groups assume the usual role of Glu192 (bent conformation covering S1) and in the process make close contacts with Cys191-S and Cys220-S providing a more hydrophobic environment for the latter by expelling some water molecules from the region.

The different backbone associations with thrombin notwithstanding, PPACK and the Selectide inhibitors display similar binding within the S1 specificity pocket. In both structures, two nitrogen atoms of the guanidino group (PPACK) and the amidino group (SEL2711 and SEL2770) form a doubly hydrogen bonded salt bridge with the carboxylate of Asp189. In addition, one of the nitrogen atoms also interacts with Gly219-O and the other makes a water mediated bridge to Phe227-O. Two more water molecules, one bridging from Asp189-OD2 to Tyr228-OH and the other bound to Asp189-OD1, both

also present in the apo-thrombin structure, remain in the specificity pocket on binding of the inhibitors. In the thrombin-PPACK structure, an additional water molecule in the specificity pocket mediates an interaction between the arginyl NE atom and Glu192-OE1 while in SEL2711 and SEL2770, this interaction is lost because Glu192 is in the thrombin-hirugen conformation (extended) (Vijayalakshmi et al., 1994). The bulkiness of the incoming aromatic ring of pAph1 or that of Chg2 appears to disrupt the usual conformation of Glu192 (bent over the S1 site) that accompanies arginyl binding at the specificity site. Residue Glu192 adopts a fully extended open conformation and does not have any specific contacts, being positioned on the surface of thrombin, which allows the P2 Chg2 group of the inhibitors to occupy the same site as the P3 aspartate of a thrombin-thrombin platelet receptor peptide complex.

Superposition of the Selectide inhibitors and the thrombin retro-binding BMS-183507 inhibitor (PDB code: T5161) reveals a practically identical parallel β -strand association of the backbone of the inhibitors with the Ser214-Gly216 residues of thrombin (Figure 2-4). Like PPACK, BMS-183507 forms three hydrogen bonds with Ser214-O, Gly216-N and Gly216-O of thrombin, while the Selectide inhibitors have only two (Table 2-4). The additional hydrogen bond of BMS-183507 is accomplished by the different positioning of the terminal amide group. The arginyl-like group of the BMS inhibitor has one carbon atom less than an arginyl side chain, so its entrance to the specificity site leads to a bowed conformation that positions the terminal amide within

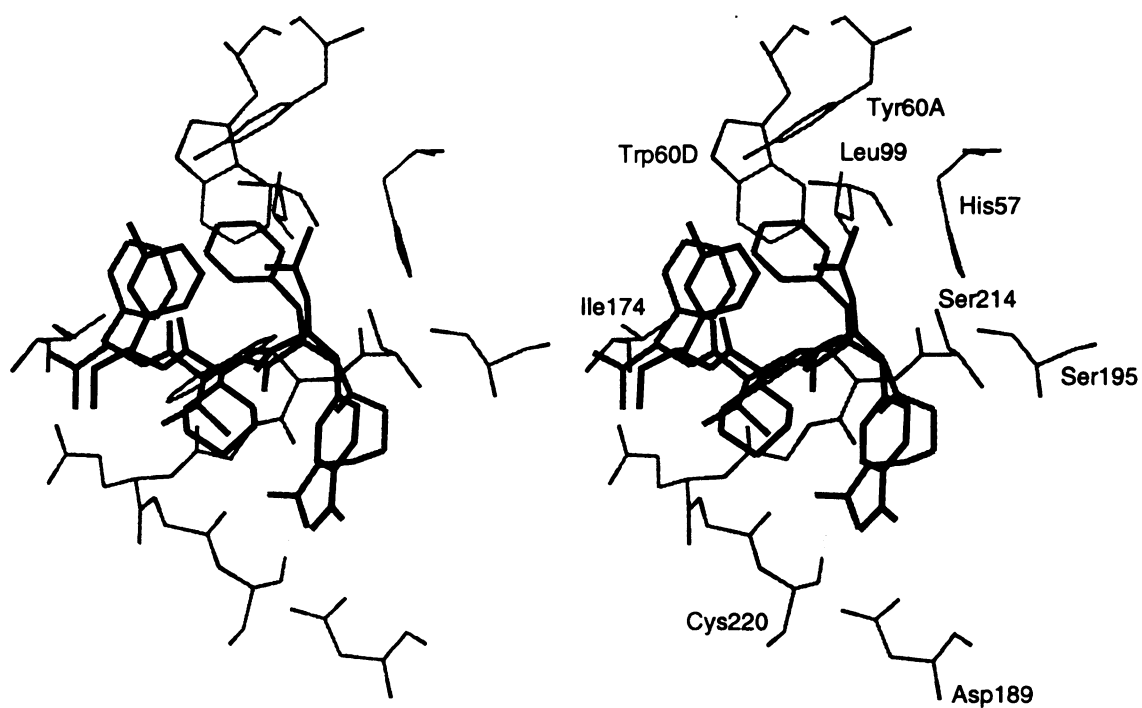


Figure 2-4: Stereoview of superimposed SEL2711 and BMS-183507 bound to thrombin. SEL2711 and BMS-183507 are shown in bold black and gray, respectively. Thrombin residues of the SEL2711 complex are in thin gray and numbered.

hydrogen bonding distance of Ser214-O (Figure 2-4). The N-terminal arginyl-like moiety and phenyl group of BMS-183507 bind in the specificity pocket and S2 site, respectively, whereas the S1 site is occupied by the Aph1 group of the Selectide inhibitors while the N-terminal acetyl group is in the S2 position (Figure 2-4). In addition, the BMS and the Selectide inhibitors display different binding modes within the S1 site. Only one of the nitrogen atoms of the guanidine of BMS-183507 makes contact with a carboxylate oxygen atom of Asp189 while the other hydrogen bonds to Gly219-O like the Selectide inhibitors (Figure 2-4; Table 2-4).

Since the structures of the catalytic domains of FXa (Padmanabhan et al., 1993) and FVIIa (Banner et al., 1996; Zhang et al., 1999) are similar to that of thrombin (rmsΔ = 1.2 Å and 1.1 Å, respectively), which especially applies to the active site regions, it is quite possible that the observed retro-binding mode of the Selectide inhibitors with thrombin may also be operative with FXa and FVIIa. This possibly was explored by transposing the thrombin bound Selectide inhibitors on the active sites of FXa and FVIIa. (Figures 2-5 & 2-6). The principal sequence differences in the residues surrounding the S1-S3 binding subsites of the two enzymes are: Leu99, Ile174 and Glu192 of thrombin are replaced by Tyr99, Phe174 and Gln192 in FXa and by Thr99, Pro170I and Lys192 in FVIIa, respectively. An even larger difference is the absence in FXa and FVIIa of the 60-insertion loop of thrombin. The Tyr60A - Trp60D segment of the loop projects out over the S2 and D-enantiomorphic S3 binding sites in thrombin restricting their access and thus contributing to specificity.

Selectide inhibitors modeled into the active site of FXa: Theoretical considerations have led to the proposal of cation- π electron mediated interactions

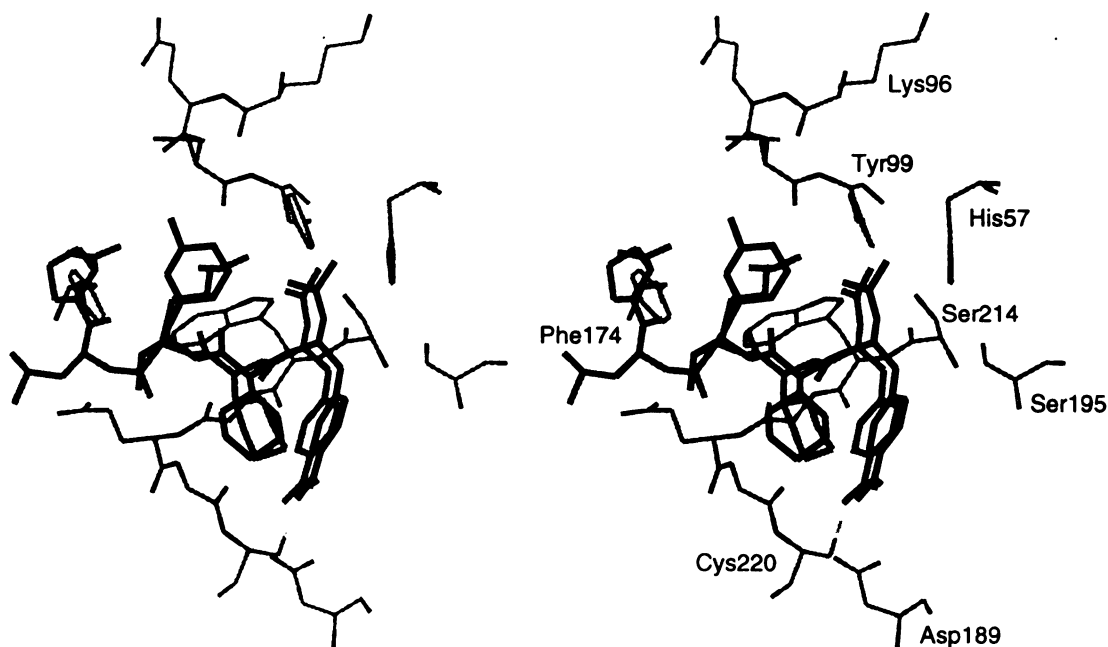


Figure 2-5: Stereoview of SEL2711 and SEL2770 superimposed on the active site of FXa in a retro-binding mode. SEL2711 and SEL2770 are shown in bold black and gray, respectively. FXa residues are in thin gray and numbered. Since the pyridine of SEL2770 collides with Phe174 of FXa, a conformational change must occur in one or the other, most likely in the former.

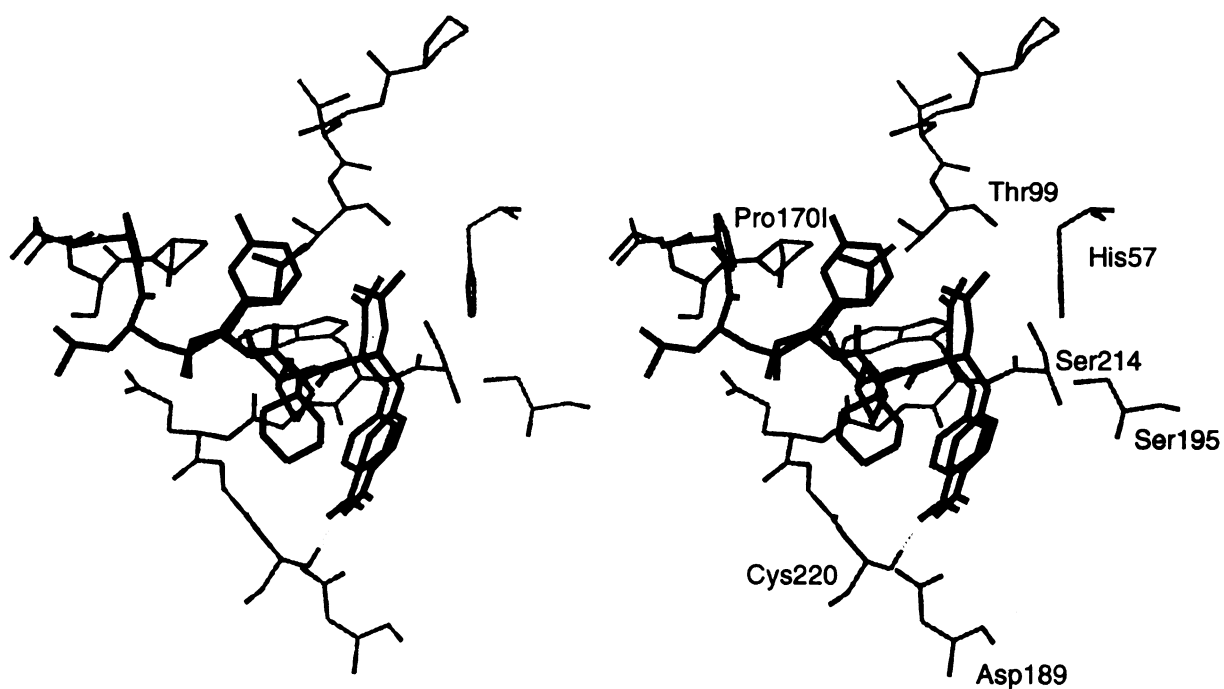


Figure 2-6: Stereoview of SEL2711 and SEL2770 superimposed on the active site of FVIIa in a retro-binding mode. SEL2711 and SEL2770 are shown in bold black and gray, respectively. FVIIa residues are in thin gray and numbered.

(Dougherty & Stauffer, 1990) in the S4 binding site of FXa (Lin & Johnson, 1995) where the π -faces of the three aromatic residues Tyr99, Phe174 and Trp215 are sufficiently rich in π electrons that they not only form a hydrophobic pocket but also act as a cation recognition site. This region also corresponds to the D-enantiomorphous S3 site, which is occupied by D-Phe in the thrombin-PPACK complex. The P4, L-leucine residue of the thrombin-LDPR complex of thrombin platelet receptor peptide also occupies this site in thrombin (Mathews et al., 1994).

Thus, the charge interaction is basically an attraction of a positive charge to the π -electrons of the aromatic rings, alternatively called an ion-quadrupole attraction (Stauffer et al., 1990; Schwabacher et al., 1993). The cation- π electron binding hypothesis was modified somewhat with the structure determination of the DX-9065a (Daiichi compound) inhibited structure of FXa (Brandstetter et al., 1996), which showed that the above aryl-binding site was further assisted by the carbonyl oxygen atoms of Lys96-Glu97 and the side chain of the latter to form a so-called cation hole that expands the S4 binding site region. It is noteworthy that the main chain of FXa between 94 - 98 is positioned differently in thrombin due to an insertion in the vicinity of Glu97A of the latter. In the absence of the insertion, Lys96 - Glu97 of FXa make closer contacts with groups binding at the D-S3 site. It is also note that the Arg93 - Arg101 stretch has been implicated in thrombin heparin binding by electrostatic potential energy calculations (Karshikov et al., 1992) and by mutagenesis (Gan et al., 1994; Ye et al., 1994; Sheehan et al., 1994). Thus, under appropriate circumstances, the two sites (S4 and the heparin binding exosite II) could interact with each other.

If the Selectide inhibitors bind to FXa in a similar *retro-manner* as with thrombin, the positively charged Pal3(Me)₃ and Nlys(2Me)₃ groups occupy the aryl site formed by Tyr99, Phe174 and Trp215 (Figure 2-5). In the case of SEL2711, the Pal3(Me)₃ group, modeled as it binds to thrombin, makes edge-to-face interactions with Tyr99 and Phe174. The latter aromatic ring systems are both parallel to the pyrrolidine ring of the inhibitor in the DX-9065a-FXa complex (Brandstetter et al., 1996). The different orientations could reflect the difference between an aromatic and aliphatic group interacting with the aryl site of FXa or it could reflect a more accurate determination of orientation of the Pal3(Me)₃ group (2.1 Å resolution) compared to the pyrrolidine (3.0 Å resolution). If no conformational changes occur to FXa accompanying binding, the positively charged nitrogen atoms of the two Selectide inhibitors are too far away (> 4.5 Å) to hydrogen bond with the carbonyl oxygen atoms of Lys96 - Glu97 of the S4 cation hole. The three residues of the aryl site, however, can interact with the delocalized positive charge of the Pal3(Me)₃ ring similar to that described for the Daiichi compound (Brandstetter et al., 1996) to approximate a cation- π electron mediated ion-quadrupole attraction (Lin & Johnson, 1995). Since the binding constants of the two Selectide inhibitors are practically the same for FXa (Table 2-1), the Nlys(2Me)₃ group most likely achieves a similar cation - π electron mediated interaction as the Pal3(Me)₃ group of SEL2711.

Conversely, if the Selectide inhibitors bind to FXa in a *substrate-like binding mode*, the amidino nitrogen atoms of the pAph1 groups of both inhibitors could form hydrogen bonds with Lys96-O and Glu97-O of the S4 cation hole; however, the binding of the Pal3(Me)₃ and Nlys(2Me)₃ groups in the S1 specificity site is clearly inferior to that of the doubly hydrogen bonded salt bridge of pAph1 in retro-binding. Thus, a retro-

binding mode in FXa would appear to be more preferable. If the inhibitors bind in a retro-manner in both thrombin and FXa, the difference of a factor of 1.3×10^4 (SEL2711) and 4×10^3 (SEL2770) in selectivity (Table 2-1) would be due to the contribution of the Pal3(Me)₃ and Nlys(2Me)₃ groups binding at the S4 - S3 cation-aryl site of FXa. This only corresponds to 5 - 6 kcal mol⁻¹ and from all appearances, the site can either operate through cation- π electron mediated interactions (Lin & Johnson, 1995) or with a longer inhibitor group to reach Lys96 - Glu97, hydrogen bonding could also occur in the cation hole (Brandstetter et al., 1996).

Selectide inhibitors modeled into the active site of FVIIa: If the Selectide inhibitors bind to FVIIa in a retro-manner, as they do in thrombin, and if there are no protein conformational changes accompanying binding, the inhibitors retain their intermolecular interactions in the specificity pocket as well as with Gly216-O (Figure 2-6). However, the longer ion-dipole interaction between positively charged Pal3(Me)₃-N3 (SEL2711) or Nlys(2Me)₃-N1 (SEL2770) with Tyr60A-OH is lost due to the absence of the 60-insertion loop of thrombin. While in FXa this loss is compensated through cation- π electron mediated interactions at the aryl site formed by Tyr99, Phe174 and Trp215 (Figure 2-5) which result in a significant increase of selectivity of the inhibitors for FXa (Table 2-1), in FVIIa this loss remains uncompensated. The residues Thr99 and Pro170I of FVIIa, which are structurally equivalent to Tyr99 and Phe174 of the aryl site of FXa, form neither a hydrophobic pocket nor a cation recognition site. These loss of interactions results in weaker binding of the Selectide inhibitors to FVIIa: the differences are a factor >5 (SEL2711) and >25 (SEL2770) with respect to thrombin (Table 2-1). The difference of five between these two difference arises from the closer contact observed in

SEL2770, between Nlys(2Me)₃-N1 and Tyr60A-OH (3.6Å), while Pal3(Me)₃-N3 and Tyr60A-OH are separated by 4.7Å in the SEL2711 complex.

Binding of Peptides A and B to the active site of FXa: Two other inhibitors developed by Selectide Corp. (peptides A and B mentioned earlier), have been shown by using transferred nuclear Overhauser effect NMR spectroscopy (Fraternali et al., 1998) to have a very compact conformation displaying a severe hydrophobic collapse of the P2 – P3 groups in the substrate binding site of FXa and not resembling the usual substrate binding mode. The arginine of the peptides enters the specificity S1 site differently from substrate, in an extended but curved conformation, surprisingly forming a doubly hydrogen bonded salt bridge with Asp189, while the P2 and P3 residues (Ile-Tyr for peptide A, Chg-4-amino-Phe for peptide B), which must be the source of the selectivity for FXa, interact together above the S1 site with the S2 site and Tyr99. These bound structures are very different from the retro-binding mode modeled here for SEL2711 and SEL2770 where selectivity for FXa appears to come from a cationic interaction with the S3 – S4 sites. The only significant factor that would seem to underlie the binding difference thus appears to center about the S1 specificity site. It has already been indicated that the Pal3(Me)₃ and Nlys(2Me)₃ groups of SEL2711 and SEL2770 are inferior P1 residues compared to the pAph1 group and is most likely the reason why these inhibitors do not bind like peptides A and B. This is also generally in agreement with the weaker binding constants of peptides A and B for FXa, being only 1.6 and 0.3 μM, respectively, which essentially result from side chain-side chain hydrogen bond interactions between inhibitor and enzyme. Thus, three or more different active site binding modes (including anti-parallel β-strand substrate binding) may be operative in

FXa depending on the exact structure of the inhibitor, the selectivity of which is derived from interactions with the S4 region in the retro-binding case.

CHAPTER 3

SELECTIVE NON-ELECTROPHILIC INHIBITORS OF THROMBIN

I. MOLECUMETICS INHIBITORS

Peptidic thrombin active site inhibitor can be broadly classified under two types, electrophilic and non-electrophilic, according to their ability to directly interact with Ser195-OG of the catalytic triad of thrombin. Electrophilic inhibitors, such as the widely studied model ligand PPACK, form a covalent tetrahedral intermediate between an activated carbonyl group of the P1 residue of the inhibitor and Ser195-OG, while all interactions between thrombin and non-electrophilic inhibitors are strictly non-covalent. Since structural similarity of the catalytic centers of thrombin and other serine proteases that also play important physiological roles in life processes, electrophilic inhibitors can operate with a very large array of serine proteases without much selectivity, causing unwanted side effects. On the other hand, non-electrophilic inhibition of thrombin is highly sensitive to the geometry and charge distribution of the active site of the enzyme as well as structural elements of the inhibitors that all lead to greater discriminative inhibition between thrombin and other serine proteases.

With this notion, Molecumetics Ltd. (Bellevue, WA) has used a combinational chemistry approach to generate a diverse variety of non-electrophilic conformationally-constrained mimetics of the principal structural elements of proteins (α -helices, β -

strands, reverse turns) (Kahn et al., 1988; Nakanishi et al., 1992; Eguchi et al., 1996; Kim & Kahn, 1997). The syntheses (Boatman et al., 1999) and the crystal structures (St. Charles et al., 1999) of four Molecumetics (MOL) thrombin active site inhibitors bound to the enzyme were recently reported and classified as first generation of (P3–P1') β -strand templated thrombin inhibitors with the hetero-bicyclic P2 group. The binding mode of these inhibitors is similar to that of PPACK: the bicyclic and phenylalanyl groups of each inhibitor approximate the extended conformation of PPACK in PPACK-bound thrombin (St. Charles et al., 1999).

The second generation of MOL inhibitors (Figure 3-1) has an amino, amidino or amino-imidazole group attached to a 1,4-substituted cyclohexyl ring at the P1 position that can simulate arginine- or lysine-like binding at the specificity pocket of thrombin. A modified, fused three-nitrogen, hetero-bicyclic ring (Bic) constitutes the central core of the inhibitors expected to associate with the Ser214-Gly216 stretch of thrombin in the antiparallel β -strand manner. The P3 benzyl (MOL354, MOL356, MOL376 and MOL592) or bromobenzyl (MOL1245) group is designed to occupy the D-enantiomorphic-S3 binding site of thrombin. From analysis of the binding constants, the second generation of MOL inhibitors possesses more optimal binding features at the S1-S3 sites that result in highly potent and selective inhibition of thrombin. The three best thrombin inhibitors (MOL354, MOL1245 and MOL376) are also the most selective (2-3 orders of magnitude) against trypsin (Figure 3-1).

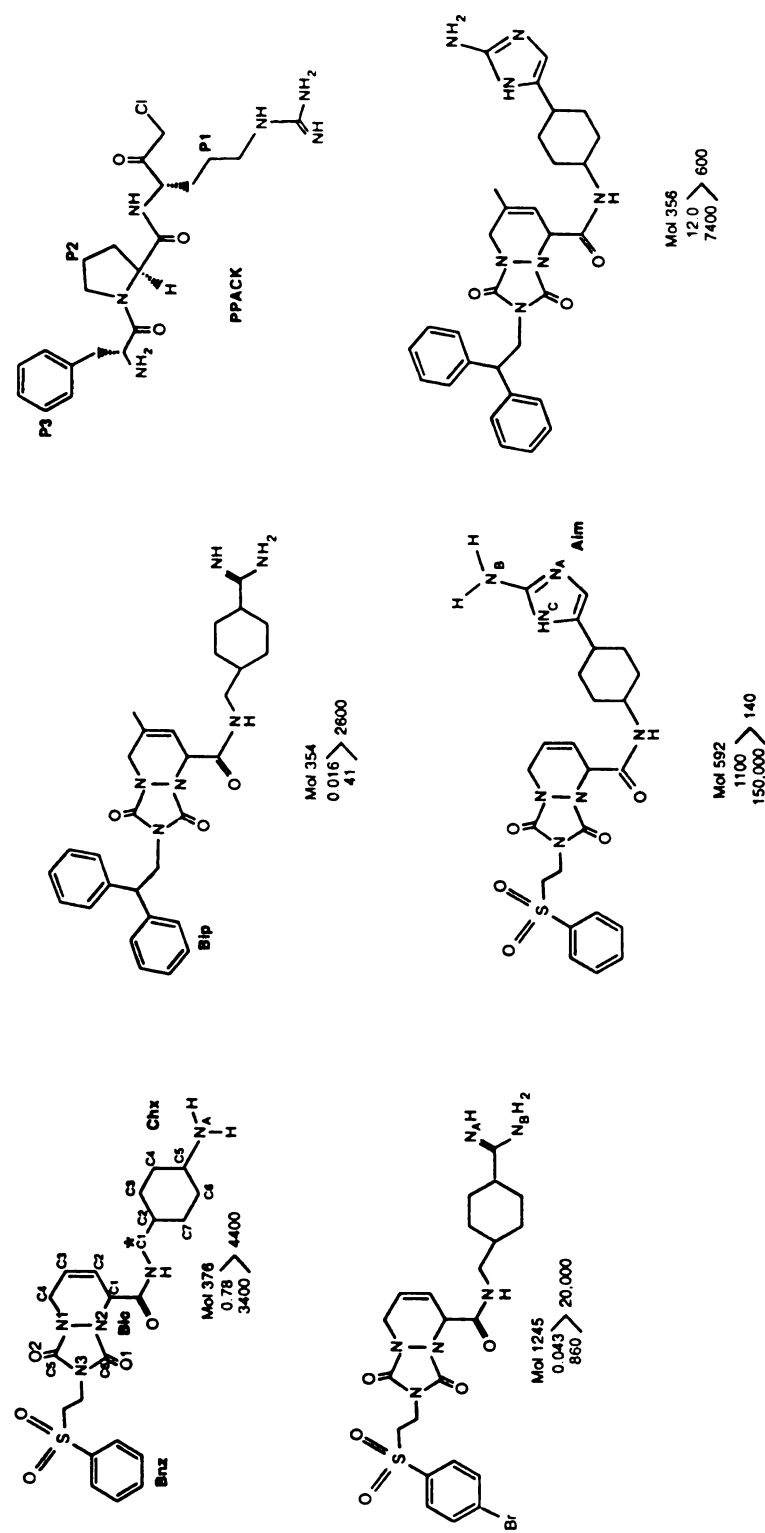


Figure 3-1: Selective non-electrophilic Molecumetics active site inhibitors of thrombin. Electrophilic inhibitor D-Phe-Pro-Arg chloromethyl ketone (PPACK) is shown for comparison. Numbering is defined on MOL 376; K_i is shown for both thrombin and trypsin (upper and lower number, respectively); ratio of K_i 's is on the right indicates a degree of selectivity.

II. EXPERIMENTAL

*Crystallization of the MOL376 and MOL592 complexes*¹: A 7-fold molar excess of hirugen was added at 4°C to a frozen human α -thrombin solution (1.0 mg/mL in 0.75 M NaCl) to form a 1:1 complex and prevent autocleavage of the enzyme. In approximately 24 hours, the solution of the formed binary thrombin-hirugen complex was concentrated in a refrigerated centrifuge (4°C) to about 7.0 mg/mL using a Centricon 10 concentrator with a molecular weight cutoff of 30 kDa. This concentrated thrombin-hirugen solution were left overnight before MOL376 and MOL592 were added to the thrombin-hirugen complex in a 50 to 60-fold molar excess.

The ternary complexes were crystallized using the hanging drop crystallization technique in a 2 μ L drop, which consisted of 1 μ L of the protein sample and 1 μ L of a well solution. The well solution contained 24% PEG 8000 in 0.1 M SPB, pH 7.3. The repetitive macroseeding technique was applied to enlarge the crystals to the size of 0.2 x 0.2 x 0.1 mm (MOL376) and 0.3 x 0.2 x 0.2 mm (MOL592).

Intensity Data Collection: The X-ray diffraction data of the thrombin-hirugen complexes inhibited with MOL376 and MOL592 were collected at -150 °C with a R-Axis II imaging plate detector using 30% PEG 4000, 15% glycerol in 50 mM SPB, pH 7.3 as cryo-solvent. The radiation generated from a Rigaku RU200 rotating anode

¹ Crystal structure and the binding modes of MOL376 and MOL592 to thrombin presented in this chapter are the part of larger study that includes structure determination and comparative analysis of the binding modes of the MOL354, MOL356 and MOL1245 complexes.

operating at 5 kW power with the fine focus filament (0.3 x 3.0 mm) was monochromated (CuK α) and intensified by focusing with Molecular Structure Corporation - Yale mirrors. The crystal to detector distance was 10.0 cm with a swing angle of zero degrees. Autoindexing and processing of the measured intensity data were carried out with the Rigaku R-Axis software program package (Higashi, 1990). Characteristics of the MOL376 and MOL592 crystals and the intensity data collection statistics are given in Table 3-1.

Structure Determination: The MOL376- and MOL592-thrombin-hirugen crystals belong to the monoclinic space group C2 with four ternary complexes per unit cell. Their crystal structures were determined directly using the isomorphous thrombin coordinates of the binary thrombin-hirugen complex (Vijayalakshmi et al., 1994; PDB code: 1HAH). The initial coordinates were optimized by rigid-body rotation-translation refinement in the (9.0 - 3.0) Å resolution range using the X-PLOR program package (Brunger, 1992). The structures were refined with the programs PROLSQ (Hendrickson, 1985) and PROFFT (Finzel, 1987) generally following procedures described in Chapter 2. The final refinement statistics of the complexes are presented in Table 3-2 and 3-3.

III. RESULTS

Thrombin: Electron density is well defined for most of the residues of thrombin in both MOL the complexes, except for the terminal regions of the A chain (Thr1H - Ala1B, Ile14K - Arg15) and the autolysis loop in the vicinity of Trp148 - Lys149E, both of which are also typically disordered in most other isomorphous thrombin structures. The

Table 3-1: Intensity data collection of Molecumetics inhibitors MOL376 and MOL592 bound to thrombin.

	MOL376	MOL592
Space group	C2	C2
Cell constants		
a (Å)	71.46	71.43
b (Å)	72.19	71.67
c (Å)	73.12	72.45
β (deg)	100.7	101.0
Complexes/asymmetric unit	1	1
Resolution (Å)	2.3	2.0
Observations	14,987	33,066
Outermost range (Å)	2.5-2.3	2.2-2.0
R-merge (%)	5.8	5.7
Outermost range	13.1	15.8
Independent reflections	9,294	17,371
Redundancy	1.6	1.9
Completeness (%)	56	72
Outermost range	38	42
$I/\sigma(I)$ (outermost range)	2.6	2.8

Table 3-2: Refinement summary of deviations from ideality of thrombin molecules in the MOL376 and MOL592 complexes.

Restraints	Target	MOL376	MOL592
Distances (Å):			
Bond distance	0.020	0.011	0.014
Angle distance	0.030	0.037	0.036
Planar 1,4 distance	0.050	0.047	0.046
Planes	0.040	0.034	0.030
Chiral volume (Å ³)	0.15	0.17	0.17
Non-bonded contacts (Å)			
Single-torsion	0.60	0.22	0.24
Multiple torsion	0.60	0.25	0.28
Possible H-bond	0.60	0.25	0.28
Isotropic thermal factors (Å ²)			
Main-chain bond	1.0	1.0	0.9
Main-chain angle	1.5	1.7	1.5
Side-chain bond	2.0	2.5	2.1
Side-chain angle	3.0	3.6	3.0

Table 3-3: Individual refinement statistics of MOL376 and MOL592 inhibitors.

	MOL376	MOL592
Resolution range (Å)	9.0-2.3	9.0-2.0
No. of reflections	9092	17152
$ F_o /\sigma$	1.5	1.5
R-value (%)	15.3	17.4
No. of water molecules	98	134
$\langle B \rangle$ (Å ²)	28.4	28.3

final thrombin structures of the ternary complexes are similar and show no significant conformational differences from each other as well as other thrombin structures ($\text{rms}\Delta < 0.4\text{\AA}$ between 240-250 CA positions). The Ramachandran plots (Laskowsky et al., 1993) of the MOL376 and MOL592 complexes show that 191 (78.9 %) and 200 (82.3 %) residues of thrombin are in most favorable regions while 50 (20.7 %) and 43 (17.7 %) residues occupy additionally allowed areas (Figure 3-2). There is no residue in the disallowed regions and only one residue (MOL356) is in the generously allowed region. Two Na^+ ion sites have been found in the monoclinic crystal structures like those initially described (DiCera et al., 1995; Zhang & Tulinsky, 1997). Both sodium ions adopt slightly distorted octahedral arrangements coordinated by solvent water molecules and carbonyl oxygen atoms: those of Arg221A and Lys224 for the intramolecular

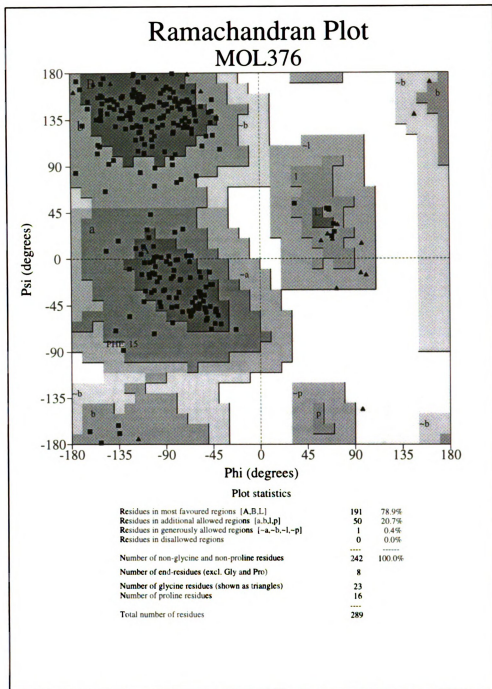


Figure 3-2/A: The Ramachandran plot of the final structure of the thrombin-MOL376 complex. Glycine residues are shown as triangles. Non-glycine residues as squared boxes.

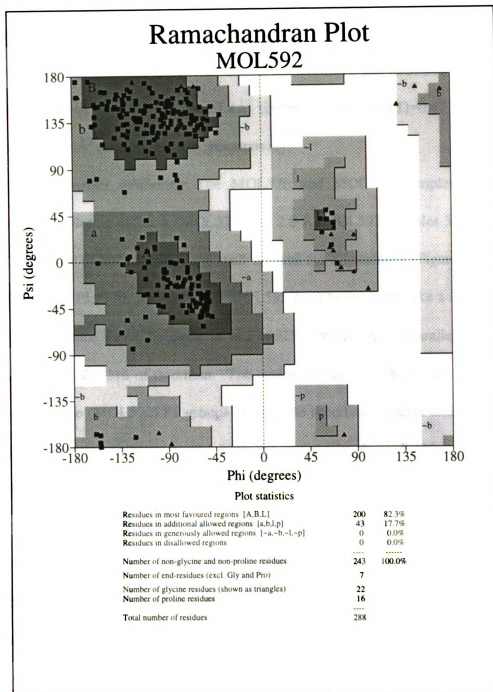


Figure 3-2/B: The Ramachandran plot of the final structure of the thrombin-MOL592 complex. Glycine residues are shown as triangles. Non-glycine residues as squared boxes.

site while those of Lys169 and Thr173, as well as Phe204A from a symmetry related molecule, coordinate the intermolecular ion (Zhang & Tulinsky, 1997). The occupancies and the temperature factors of the Na⁺ ions of the MOL376 and MOL592 complexes are: $\Omega=0.97$, $B=26.9\text{\AA}^2$ and $\Omega=0.98$, $B=27.8\text{\AA}^2$ (intramolecular) and $\Omega=0.67$, $B=28.6\text{\AA}^2$ and $\Omega=0.95$, $B=31.0\text{\AA}^2$ (intermolecular), respectively.

Molecumetics inhibitors: The MOL376 and MOL592 complexes have been refined at a resolution of 2.3 Å (MOL376) and 2.0 Å (MOL592) (Tables 3-1 & 3-2). The electron density of both inhibitors is well defined for all residues (Figure 3-3). Unlike SEL2711 and SEL2770, MOL376 and MOL592 bind to thrombin like a substrate or the archetype inhibitor PPACK (Bode et al., 1992) forming an antiparallel β -strand with residues Ser214-Gly216 of thrombin. The cyclohexyl amino (MOL376) and cyclohexyl amino-imidazole (MOL592) P1 groups occupy the specificity pocket and make one and two direct interactions, respectively, with the negatively charged carboxylate group of Asp189. The P2 hetero-bicyclic Bic groups are practically planar with some small deviations from planarity in the six-membered ring. The Bic group of each inhibitor occupies the apolar S2 and part of the S3 region of the active site making hydrophobic contacts with the side chains of His57, Tyr60A, Trp60D and Trp215 like the first generation of bicyclic groups (St. Charles et al., 1999). The six-membered ring occupies approximately the same position as the proline of PPACK in thrombin-PPACK (Bode et al., 1992). Bridged by the sulfonyl, the P3 benzyl groups of MOL376 and MOL592 curl up toward the 60-insertion loop of thrombin and make numerous hydrophobic

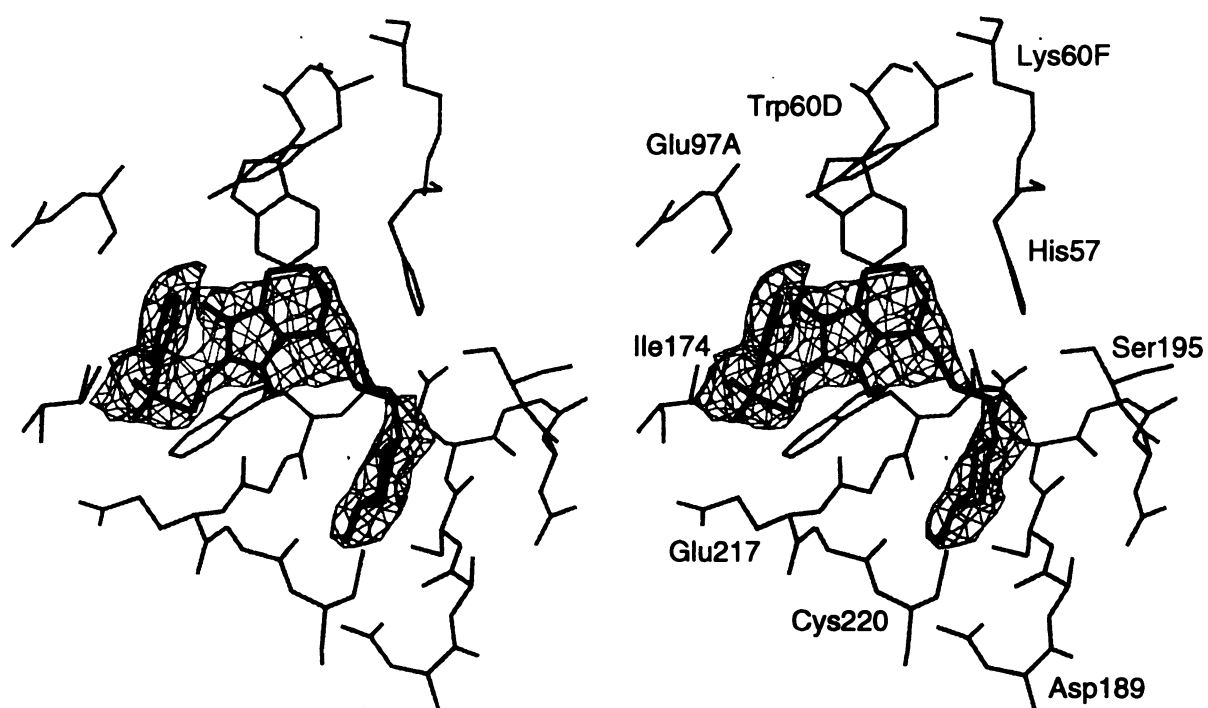


Figure 3-3/A: Stereoview of the final (2Fo-Fc) electron density of the MOL376 inhibitor in the thrombin complex. Contoured at the 1 σ level. MOL376 is in bold black. Thrombin residues are in thin gray and numbered.

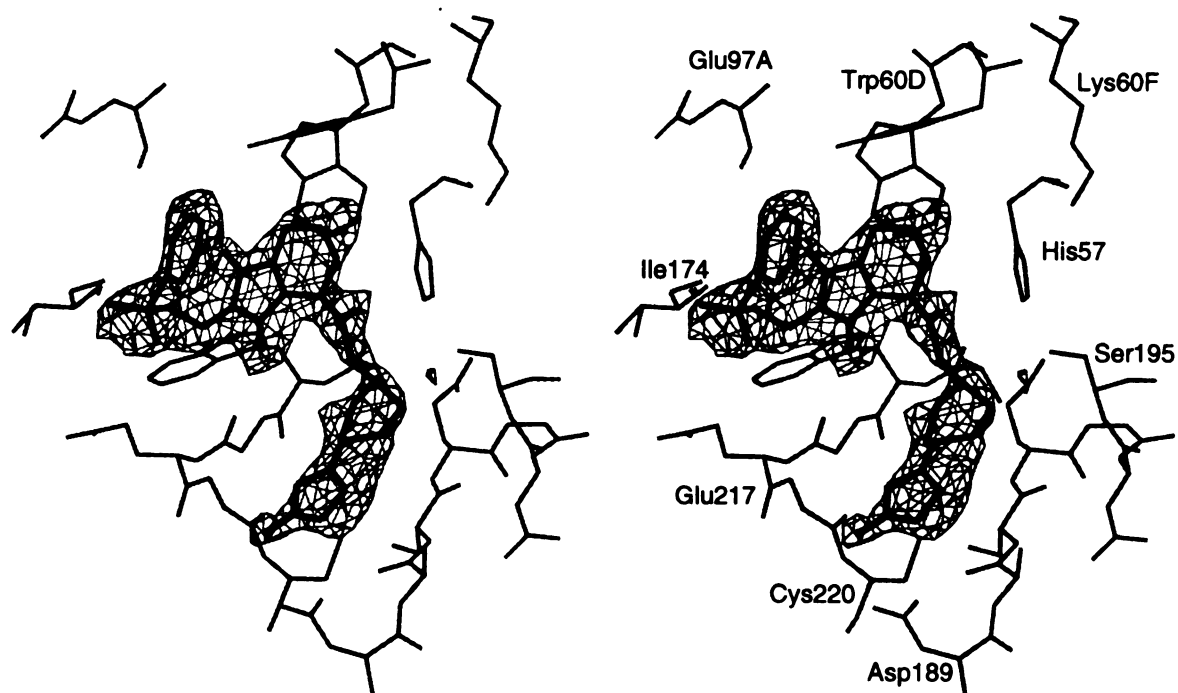


Figure 3-3/B: Stereoview of the final (2Fo-Fc) electron density of the MOL592 inhibitor in the thrombin complex. Contoured at the 1σ level. MOL592 is in bold black. Thrombin residues are in thin gray and numbered.

interactions with thrombin residues Asn98-Leu99, Ile174 and Trp215 of the D-enantiomeric-S3 binding site. The orientation of the benzyl ring of both inhibitors is approximately orthogonal to the plane of the indole ring of Trp215, in a fashion typical of aromatic-aromatic interactions in proteins (Figure 3-3).

IV. DISCUSSION

As mentioned earlier, MOL376 and MOL592 together with MOL354, MOL356 and MOL1245 belong to a second generation of conformationally-constrained Molecumetics inhibitors (Figure 3-1). Because of the close structural similarity of these inhibitors, the overall binding modes are similar and relate to that observed in the PPACK complex (Bode et al., 1992). The amino, amidino or amino-imidazole charged group attached to a cyclohexyl ring at the P1 position form hydrogen bonded salt bridges with the negatively charged carboxylate of the Asp189 side chain (Figures 3-4 & 3-5). Differences in the chemical nature of the charged basic groups, however, and the overall size and exact conformation of the P1 groups produce variations in binding interactions and the role of solvent in the stabilization of the thrombin-MOL inhibitor complexes (Figure 3-6). The cyclohexyl group of MOL376, like those in MOL354 and MOL1245, appears to be in the energetically favored chair conformation (Figure 3-4); however, in MOL592 and MOL356 the cyclohexyl ring is in a boat conformation (Figure 3-5).

The cyclohexylamino group of MOL376 interacts with Asp189 of thrombin more like the lysine residue of the first generation MOL106 inhibitor (St. Charles et al., 1999) rather than the lysine analogue of DUP714 (Weber et al., 1995), where a water molecule

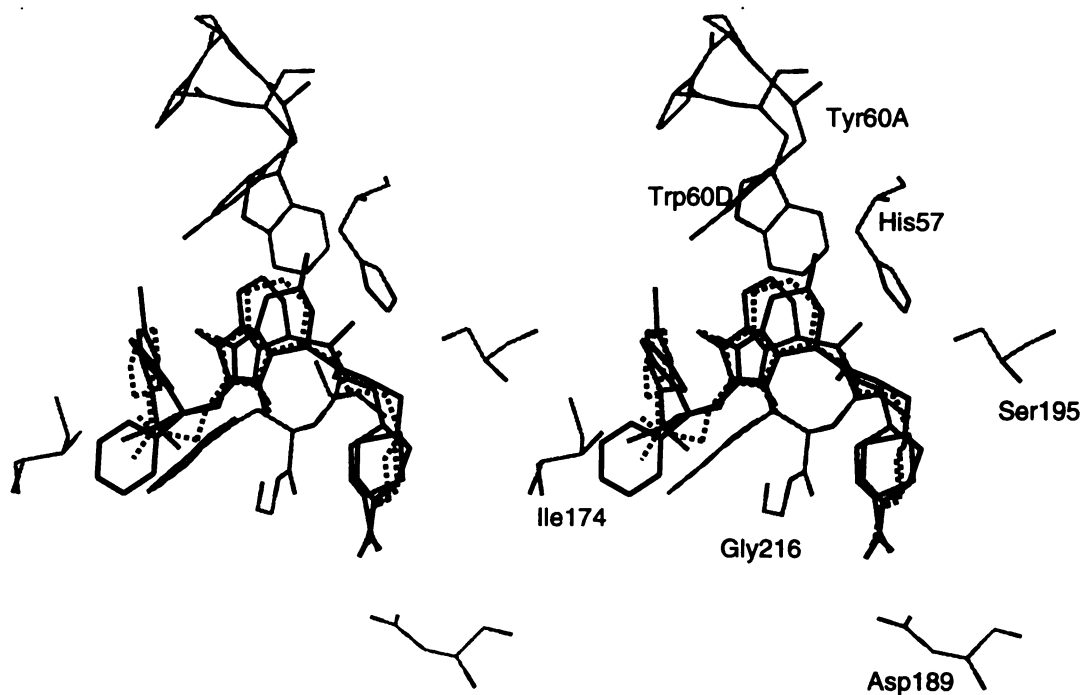


Figure 3-4: Stereoview of superimposed MOL354/376/1245 bound to thrombin. MOL354 is in bold black; MOL376 is broken black line; MOL1245 is in bold gray. Thrombin residues of the MOL1245 complex are in thin gray and numbered.

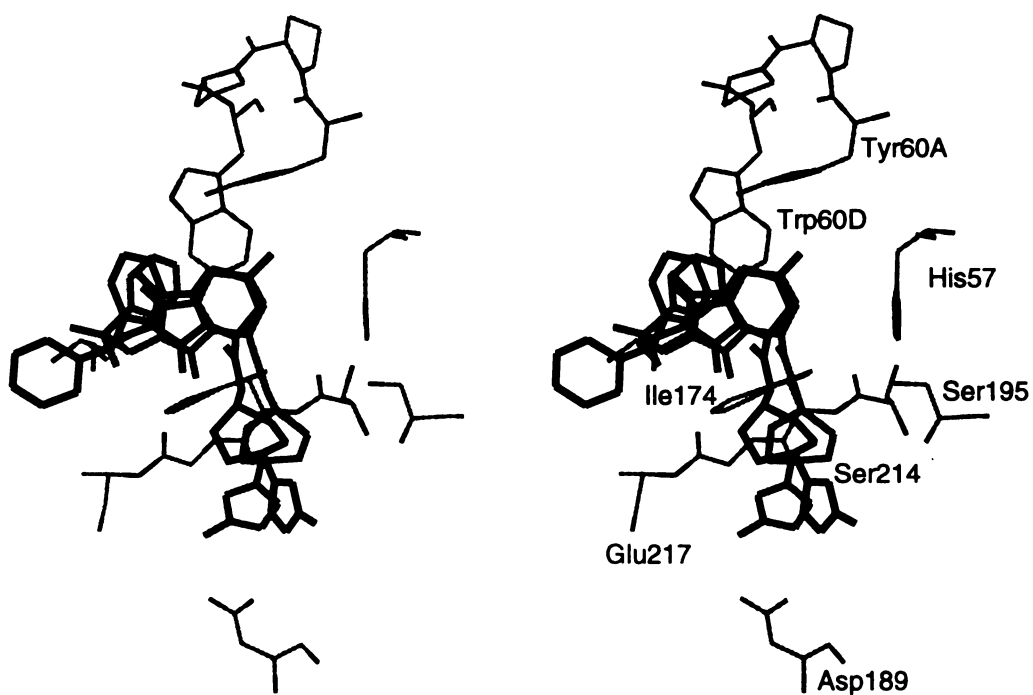


Figure 3-5: Stereoview of superimposed MOL356 and MOL592 bound to thrombin. MOL356 is in bold black; MOL592 is in bold gray. Thrombin residues of the MOL356 complex are in thin gray and numbered.

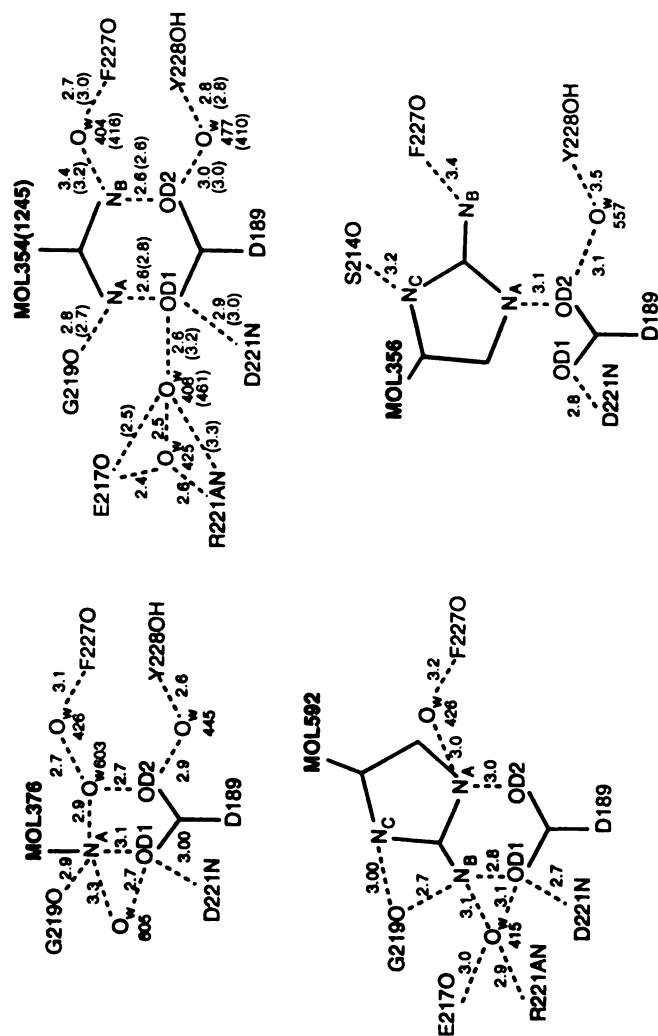


Figure 3-6: Schematic of hydrogen bonded salt bridges of Molecumetics inhibitors with Asp189 of the S1 specificity site of thrombin. MOL354/1245 shown together in one schematic, MOL1245 in parentheses.

bridges the lysine amine and the Asp189 carboxylate group. This is most likely because the similarity of the large hydrophobic bicyclic P2 groups of MOL106 and MOL376 which align the inhibitors in the active site of thrombin in a similar orientation. In both MOL complexes, the amino group makes a hydrogen bonded salt bridge with Asp189-OD1 (Figures 3-4 & 3-6). Moreover, in the MOL376 complex, the bond to the cyclohexyl ring is equatorial and an equatorial cyclohexylamino group occupies a position equivalent to one of the terminal nitrogen atoms of the guanidinium group of arginine- or amidino-based inhibitors (Figures 3-4 & 3-6). Other conserved interactions of such inhibitors are also observed in the S1 site (Zhang & Tulinsky, 1997), which are a hydrogen bond to Gly219-O and a doubly hydrogen bonded water molecule mediating Asp189-OD2 and Tyr228-OH (Figure 3-6). Another water molecule, O_w603 positionally mimics a NH1 or NH2 atom of thrombin-bound arginine and completes the interactions of the amino group bridging between it and Asp189-OD2 (Figure 3-6). The second solvent shell of MOL376 includes the conserved water site of O_w426 hydrogen bonded to Phe227-O and O_w603. Although Asp189-OD1 of thrombin and the cyclohexylamine of MOL376 hydrogen bond to a water molecule (O_w605) as in thrombin complexes with arginine-based inhibitors, unlike the latter, no further interactions ensue from this water because it is positionally a little different. Some of the foregoing interactions have also been observed with other thrombin bound inhibitors containing a cyclohexylamine or an aminopyridine group at the P1 position (Tucker et al., 1997; Feng et al., 1997).

Both MOL356 and MOL592 have cyclohexyl-aminoimidazole groups occupying the S1 site but since the amide nitrogen lacks a methylene carbon, it is linked directly to the cyclohexyl ring so the latter does not penetrate as deep into the S1 site and only

overlaps about half of the cyclohexyl ring of MOL354, MOL376 or MOL1245. The cyclohexyl rings of MOL356 and MOL592 are in the boat conformation and the stereochemistry of the bonds to the cyclohexyl and imidazole groups of the two (both apical positions) are the same being axial and equatorial, respectively; however, the interactions of the amino-imidazole groups are very different in the specificity pocket (Figures 3-5 & 3-6). The imidazole ring is parallel to the plane of the boat in MOL592 whereas it is rotated about 75° in MOL356. Moreover, the MOL592 inhibitor forms a doubly hydrogen bonded salt bridge with Asp189, through a nitrogen atom of the imidazole ring and its amino group. This is unlike MOL356, which interacts with Asp189 uniquely (Figures 3-5 & 3-6). Thus, the terminal imidazole of the P1 group of MOL592 simulates more a guanidinium of arginine utilizing all three nitrogen atoms (Figure 3-6). This amino-imidazole mimics arginine further by having other conserved hydrogen bonded interactions found in thrombin-arginine complexes (Krishnan et al., 1998). The imidazole ring of MOL356 is flipped 180° with respect to MOL592 about its bond to the cyclohexyl ring and only one imidazole nitrogen atom forms a hydrogen bond with Asp189-OD2 *while the amino nitrogen makes a direct hydrogen bonding contact with Phe227-O* occupying the position of a conserved water molecule of other complexes (Figure 3-6). This water molecule is linked to the charged terminal guanidinium or amidine group by a hydrogen bond in P1 arginyl or amidino inhibitors.

The O1 oxygen atoms of the Bic groups of the inhibitors form a hydrogen bond with Gly216-N like the bicyclic P2 group of the first generation Molecumetics inhibitors (St. Charles et al., 1999). This interaction is in all the thrombin-MOL structures (< 3.0 Å) and corresponds to one of the three usually found in the antiparallel β -strand between

P2-P3 residues of inhibitor and Ser214-N, Gly216-N and Gly216-O. The usually accompanying interaction from P3 nitrogen to Gly216-O, which is observed in the SEL2711 and SEL2770 complexes (Table 2-3), is missing with the present Molecumetics inhibitors because they lack the nitrogen atom. A slightly longer hydrogen bond contact, which varies between (3.1 – 3.3) Å, in the antiparallel β -stretch is found between the amide nitrogen of the MOL354/376/1245 inhibitors and Ser214-O. This interaction is observed in thrombin complexed with PPACK (substrate binding mode) and with BMS-183507 (retro-binding mode).

The sulfonyl groups of MOL376 and MOL592 are located on the surface and have no interactions with thrombin. This is in agreement with other P3 benzylsulfonyl inhibitors bound to thrombin (Krishnan et al., 1998). In comparable complexes with trypsin, the benzyl group is disordered and the sulfonyl is positionally displaced and interacts with Gly218-N through a hydrogen bond (Krishnan et al., 1998). A similar hydrogen bond has been reported with thrombin and two benzylsulfonamide inhibitors (Tucker et al., 1997; Feng et al., 1997). However, these particular inhibitors have a D-Phe at P3 that occupies P3, with the benzylsulfonated N-terminus of the inhibitor. The benzyl group thus occupies a uniquely different position close to a P1 cyclohexyl or pyridylamine.

The crystallographic determination of the binding modes of the Molecumetics inhibitors is used here to correlate the trends in the binding constants of the inhibitors for thrombin. The two best inhibitors for thrombin are MOL354 and MOL1245, which have a cyclohexylamidino P1 group and sub-nanomolar binding constants (Figure 3-1). Based on the binding constants, the amidino group of MOL354 or MOL1245 mimicks the

guanidinium of arginine and substantially contributes to the binding constant, while the contribution of the P3 groups should be comparable. The majority of the loss of a factor of about 20-30 in thrombin binding of MOL376 is clearly the result of the replacement of the amidino with an amine group at P1. Although the amine of MOL376 copies amidino binding to some extent (Figure 3-6), the overall energetics of binding appears to be weakened. The cyclohexyl-aminoimidazole inhibitors are the poorest inhibitors of thrombin (Figure 3-1), which must be due to the bulkiness of their P1 group rather than delocalization of the positive charge within the imidazole ring. It is difficult, however, to account for the markedly poorer binding constants of MOL356 and MOL592 from their specificity site interactions (Figure 3-6) that suggest more or less comparable binding to the other Molecumetics inhibitors. MOL356 is a better inhibitor of thrombin than MOL592 by two orders of magnitude. It is most likely due to a larger number of direct interactions of MOL356 with thrombin, rather than a larger number of water mediated ones observed in MOL592 (Figure 3-6).

CHAPTER 4:

PROBING THE ACTIVE SITE OF THROMBIN FURTHER WITH UREA

I. INTRODUCTION

A way to determine relationships between tertiary structure and biological activity of enzymes is to alter the structure and study the effect of the alternation on the biological functions of the enzyme. Structure alternations in proteins caused by heat, extreme pH, or by exposure to organic solvents, detergents or certain solutes are referred to as denaturation. The structural changes caused by denaturation are generally associated with loss of catalytic activity and are a consequence of catalytic functions depending on its specific three-dimensional structure. The process of denaturation is usually reversible: most of the denaturated proteins will regain their native structure and, therefore, biological activity, if they are returned to a condition in which the native conformation is stable. This phenomenon of reversibility is one of the proofs that the tertiary structure of proteins is determined by amino acid sequences and not by the path by which a protein folds into its native conformation. Previous studies of the effect of denaturing agents on the biological properties of various proteins may be broadly classified into two categories: 1) those that examine dependence of the protein activity as a function of concentration of denaturing agents and 2) those that examine denaturant binding sites in the protein at the molecular level.

The effect of denaturing agents on biological activity of protein is a multi-facetted subject. Many enzymes with covalently intact structures exhibit sigmoidal, almost fully reversible unfolding upon treatment with denaturants in non-reducing

conditions. These enzymes include widely used blood proteins such as trypsin (Harris, 1956), α -chymotrypsin (Martin & Frazier, 1963) and α -thrombin (Chang et al., 1980). The degree of the recovery of activity is sensitive to a molar ratio of enzyme to denaturant, time and pH of the incubation. Multidomains non-covalently associated in proteins also exhibit disassociation and reassociation as entire units. While covalently intact α -thrombin refolds from urea in less than two minutes with complete return of both clotting and esterase activity, three (noncovalent)-domain γ -thrombin requires up to 90 minutes to regain full esterase activity under the same conditions (Chang et al, 1980). It was suggested that γ -thrombin renaturation represents a two step process: (1) the rapid refolding of the covalently intact units and (2) subsequent reassociation of the three non-covalent domains of the protein to yield a lower energy, fully active conformation. A possible side effect of the treatment of proteins with denaturing agents is self-association or an oligomerization of protein molecules trapped in conformationally altered forms (Zhuang et al., 1996^{a,b}).

The main incentive to study the pattern of interactions between denaturants and proteins at the molecular level is to gain insight into the nature of protein stability, susceptibility to different levels and types of denaturants and structural rearrangements associated with denaturation. It is expected that denaturing agents are not only randomly distributed on the surface of a protein but also that they bind to specific operating sites in characteristic patterns. A total of nine urea molecules have been crystallographically identified in the structure of hen egg-white lysozyme from crystals soaked in 9.0 M urea (Pike & Acharya, 1994). While eight of these nine molecules were

also found in lysozyme soaked in 4.0 and 5.0 M urea, only three urea binding sites were occupied in the lysozyme structures of crystals soaked in 0.7, 2.0 and 3.0 M urea. Three-dimensional X-ray intensities of crystals of α -chymotrypsin dialyzed against either 2.0 M guanidine hydrochloride or 3.0 M urea solutions at pH 3.6 were measured at 2.8 Å resolution and examined to identify denaturant binding sites (Hibbard & Tulinsky, 1978). Guanidinium ions appear to have bound to the protein at two distinct locations on the surface of the protein: (1) in the intermolecular uranyl heavy atom binding site (Tulinsky et al., 1973; Vandlen & Tulinsky, 1973) and (2) in a solvent cavity in the dimer interface region. Structural analysis of the urea treated crystal of chymotrypsin displayed changes occurring not only on the surface of the protein but also in the hydrophobic interior close to nonpolar side chains of Ile6, Val9 and Val23 of chymotrypsin. Surprisingly, neither guanidine hydrochloride nor urea was found in the active site region of the protein (Hibbard & Tulinsky, 1978).

Crystallographic identification and characterization of hydrogen bonding and van der Waals interactions of denaturing agents bound to the active site region of an enzyme could make significant contributions to the process of drug design. However, to present, neither the sites by which proteins interact with denaturing agents nor the consequent structural alternations upon binding of the denaturants have been classified or precisely determined at the molecular level for most of the proteins. Factors, such as (1) poor diffraction of protein crystals treated with denaturing agents and (2) conformational heterogeneity of protein molecules upon treatment with denaturing agents have been major obstacles to fully addressing the process of denaturation.

The multifunctional biological activity of thrombin resulting from the presence of four operating sites, the active site, the fibrinogen and heparin binding exosite I and exosite II and the sodium binding site (Chapter 2), stimulated our interest to study the relationships between biochemical and structural alternations caused by denaturation to this multifunctional enzyme. To fully address the problem: (1) all transition states from the intact to fully denatured thrombin structure should be characterized; (2) the changes that occur at the catalytic and the allosteric sites during denaturation have to be analyzed; (3) the catalytic activity of thrombin must be measured as a function of the concentration of the denaturant; and (4) observed and simulated structural changes have to be correlated with the measured catalytic activity of the enzyme.

Due to fundamental limitations of crystallography, the technique can not address all of the foregoing. Crystallography, nevertheless, may shed light on the onset of the process of denaturation. At low concentrations of urea: (1) the thrombin-hirugen complex should still be crystallizable and (2) the denaturing agent should not yet be randomly localized on the surface but rather bound to specific sites on the protein. Identification of urea binding sites in the active site region of thrombin could make a significant contribution to the process of drug design and to an understanding of the initial phase of protein denaturation.

II. EXPERIMENTAL

Crystallization: An approximately 7-fold molar excess of hirugen was added to a frozen sample of human α -thrombin solution (1.0 mg/mL in 0.75 M NaCl) at 4°C to form a 1:1 complex thrombin-hirugen complex. Bound to the fibrinogen exosite I of thrombin,

hirugen prevents autocleavage of the enzyme. In approximately 24 hours, the solution of the binary complex was concentrated to about 7.0 mg/mL using a Centricon 10 concentrator with a molecular weight cutoff of 30 KDa in a refrigerated centrifuge at 4° C. The hanging drop technique and repetitive macroseeding were applied to obtain and enlarge the crystals, respectively, as well as to increase the concentration of urea in the crystals to the point where the crystals stop growing and start to dissolve. All crystallization procedures were carried out in a 4 µL drop consisting of 2 µL of the binary thrombin-hirugen complex and 2 µL of a well solution at 4° C. First, thrombin-hirugen-urea crystals were obtained in the drop over the well solution containing 30 % PEG 8000, 2.0 M urea in 0.1 M SPB, pH 7.3. Then, a crystal selected for X-ray diffraction was enlarged to dimensions of 0.70 x 0.40 x 0.20 mm in a fresh 4 µL drop which was pre-incubated for three days over a well solution containing 32 % PEG 8000 and 3.0 M urea in 0.1 M SPB, pH 7.3. Finally, the crystal was transferred into a fresh drop pre-incubated for four days over 36 % PEG 8000 and 4.0 M urea in 0.1 M SPB, pH 7.3 and was left to equilibrate for 12 days. The final concentration of urea in the thrombin-hirugen crystal before X-ray data collection was between the initial concentration of the denaturant in the drop (2.0 M) and the initial concentration of the denaturant in the well (4.0 M). Thus, the final concentration of urea in the thrombin-hirugen crystal will be hereinafter defined as (2.0+X) M and is expected to be between the (2.0-3.0) M concentration range.

Intensity Data Collection. The X-ray diffraction data from the thrombin-hirugen complex co-crystallized in the presence of (2.0+X) M urea were collected with an R-AXIS II imaging plate detector. The radiation generated from a Rigaku RU200 rotating

anode operating at 5 kW power with the fine focus filament (0.3 mm x 3.0 mm) was monochromated ($\text{CuK}\alpha$) and intensified by focusing with Molecular Structure Corporation-Yale mirrors. The crystal to detector distance was 10.0 cm and the swing angle was 10 degrees. The crystal scattered X-rays to 1.8 Å resolution. Autoindexing and processing of the measured intensity data were carried out with the Rigaku R-Axis software package (Higashi, 1990). Characteristics of the crystal and the intensity data collection statistics are given in Table 4-1.

Preliminary structure analysis: Scaling the X-ray diffraction data between native thrombin-hirugen and that crystallized in the presence of urea was employed to estimate the degree of possible structural re-arrangements induced by the denaturant. Two independent reflection data sets from two different thrombin-hirugen crystals (R-merge 3.1% and 3.5%) were merged to produce the data set that was chosen as native (Vijayalakshmi et al., 1994). This native data set contains total 15,157 reflections with R-merge of 6.8% (Table 4-1). The collected intensity data set of thrombin-hirugen treated with urea contains 23,588 reflections with R-merge of 5.9 % (Table 4-1). Comparison of the cell parameters between the native and derivative crystals indicates shrinkage of the thrombin-hirugen-urea unit cells in all three directions; however, the changes are marginal, all within two percent. The $I/\sigma(I) = 4$ cut-off criteria was applied to the native and derivative data sets to illuminate weak reflections. Then, the reflections from two data sets were sorted, matched according to the Miller indices and scaled with PROTEIN (Steigemann, 1974). The scale constant (k) and the average B-value between

Table 4-1: Crystal data and intensity data collection statistics of the native^a and urea-treated thrombin-hirugen complex.

	Native ^b	Urea-treated
Space group	C2 ^c	C2 ^c
Cell constants (Å):		
a	71.75	70.59
b	72.39	71.23
c	73.52	72.42
β (deg)	101.3	100.5
Resolution (Å)	2.3	1.8
Observations (I/σ(I)>1.0)	36,665	36,134
outermost range (Å)	2.5-2.3	2.5-1.8
I/σ(I) (outermost range)	6.3	7.0
Independent reflections (I/σ(I)>1.0)	15,157	23,588
Redundancy	2.4	1.5
Completeness (%)	83	71
outermost range (%)	70	56
R-merge (%)	6.8	5.9
outermost range (%)	9.7	8.5

a - Vijayalakshmi et al., 1994.

b – native data set consists of two merged independent data sets from two different crystals.

c - four complexes per unit cell, one per asymmetric unit.

native (N) and derivative (D) data sets ($|F_D|^2 = k |F_N|^2 \exp[-2(B_D - B_N)(\sin\theta/\lambda)^2]$) were 2.2 and $-3.1 \text{ \AA}^2$, respectively. The preliminary crystallographic analysis was based on a total 2,880 pairs of equivalent reflections distributed into ten shells according to their mean amplitudes ($\langle F \rangle$) mean resolution ranges ($\langle d \rangle$) and as a function of mean $\langle F/\sigma(F) \rangle$. The merging R-values between the native and derivative data sets calculated for individual shells ranged from 6.7% to 19.3% as a function of $\langle F \rangle$ (Table 4-2A), from 9.4% to 13.8% as a function of $\langle d \rangle$ (Table 4-2B) and from 9.6% to 15.8% as a function of $\langle F/\sigma(F) \rangle$ (Table 4-2C). The overall merging R-value between the native and derivative data is 10.8%.

Structure Determination. The thrombin-hirugen-urea crystal belongs to space group C2 with four ternary complexes per unit cell (one molecule per asymmetric unit). The crystal structure was determined using thrombin coordinates of the isostructural binary thrombin-hirugen complex (PDB code: 1HAH) (Vijayalakshmi et al, 1994). The initial coordinates were optimized to an R-value of 34.0% by a rigid-body refinement in the (9.0-3.0) $\text{\AA}$ resolution range using the X-PLOR program package (Brunger, 1992). Additional positional refinement decreased the R-value to 26.7%. The structure of thrombin optimized by X-PLOR was then refined with the program PROLSQ (Hendrickson, 1985) and PROFFT (Finzel, 1987). After 3 cycles of overall and 6 cycles of individual B-refinement, the R-value converged at 24.6% (7.0-2.5) $\text{\AA}$. At this stage, hirugen and both intramolecular and intermolecular sodium ions, identical to those initially described (Zhang & Tulinsky, 1997), were included based on the examination of

Table 4-2A: R_{merge} -values as a function of the structure factor magnitudes of the native and urea-treated crystals of thrombin-hirugen.

No. of Shell	< F >	Merging R-value
1	116.8	19.3
2	166.2	19.1
3	208.4	16.9
4	250.2	15.1
5	296.5	12.9
6	347.8	12.4
7	413.8	11.3
8	497.9	9.1
9	625.2	8.6
10	933.9	6.7

Table 4-2B: R_{merge} -values as a function of the resolution of the native and urea-treated crystals of thrombin-hirugen.

No. of Shell	$\langle d \rangle$ (Å)	Merging R-values
1	2.3	13.8
2	2.5	12.5
3	2.6	11.9
4	2.8	10.8
5	2.9	11.9
6	3.2	11.3
7	3.4	12.1
8	3.9	10.0
9	4.6	8.1
10	7.3	9.4

Table 4-2C: R_{merge} -values as a function of $\langle F/\sigma(F) \rangle$ of the native and urea-treated crystals of thrombin-hirugen.

No. of Shell	$\langle F/\sigma(F) \rangle$	Merging R-values
1	7.3	15.8
2	12.2	13.5
3	16.8	12.5
4	22.2	11.6
5	29.6	11.1
6	40.1	10.6
7	53.7	9.6
8	75.8	9.9
9	116.7	10.1
10	238.8	9.6

(2Fo-Fc) and (Fo-Fc) difference electron density maps. Extending the data further to 2.1 Å, to 1.9 Å and finally to 1.8 Å and conducting further refinements resulted in converging of the R-value to 23.1% (the model included thrombin, hirugen and two sodium ions; no water or urea molecules). Thereafter, solvent and urea molecules were progressively found and added into the refinement. The final structure contains five molecules of urea, two sodium ions and 190 molecules of water with occupancies greater than 0.5. The final R-value is 19.1 % (8.0-1.8) Å. The final refinement and R-value statistics are summarized in Table 4-3.

IV. RESULTS AND DISCUSSION

The presence of urea during the crystallization of the thrombin-hirugen complex had little effect on the dimensions of the monoclinic unit cell (Table 4-1). The denaturant, however, did have a substantial effect on crystal growth. To grow thrombin-hirugen crystals suitable for X-ray diffraction in the presence of the denaturant, it is necessary to increase the concentrations of both the protein complex and the precipitant PEG and to lower the crystallization temperature to 4° C. It was noticed that the growth rate of thrombin-hirugen-urea in the hanging drop over the well solution containing 30 % PEG 8000, 2.0 M urea in 0.1 M SPB, pH 7.3 was significantly faster compared to that of the native thrombin-hirugen. However, when the crystals were transferred into the fresh pre-incubated drop over the well solution containing 32 % PEG 8000, 3.0 M urea and then 36% PEG 8000 and 4.0 M urea, both in 0.1 M SPB, pH 7.3, the apparent crystal growth rate became comparable to that of the native thrombin-hirugen in the case of former and

Table 4-3: Refinement summary of deviations from ideality and final refinement parameters of the thrombin-hirugen-urea complex.

Restraints	Target	RMSΔ
Distances (Å):		
Bond distance	0.015	0.011
Angle distance	0.020	0.021
Planar 1,4 distance	0.030	0.034
Planes	0.035	0.041
Chiral volume (Å³)	0.15	0.18
Non-bonded contacts (Å)		
Single-torsion	0.60	0.20
Multiple torsion	0.60	0.25
Possible H-bond	0.60	0.24
Isotropic thermal factors (Å²)		
Main-chain bond	1.0	1.2
Main-chain angle	1.5	1.8
Side-chain bond	2.5	3.0
Side-chain angle	3.5	4.3
Resolution range (Å)	8.0 – 1.8	
No. of reflections	21,529	
 (Å²)	24.1	
R-value (%)	19.1	

then ceased. Attempts to further increase the concentration of urea in crystals resulted in their dissolution.

Electron density was well defined for most of the residues of thrombin except the termini of the A chain (residues Thr1H-Ala1B, Ile14K-Arg15) and the autolysis loop (residues Trp148-Lys149E), which are typically disordered in other isomorphous thrombin structures. A comparison of CA positions of thrombin in native thrombin-hirugen (PDB code: 1HAH) (Vijayalakshmi et al., 1994) and refined thrombin-hirugen co-crystallized in the presence of (2.0+X) M urea indicates no significant conformational changes: the rms Δ of CA positions is only 0.30 Å. The refined thrombin-hirugen structure is also similar to that of the PPACK- (PDB code: 1PPB) (Bode et al., 1992) and FPA- (PDB code: 1FPH) (Stubbs et al., 1992) thrombin bound complexes: rms Δ are 0.33 Å and 0.45 Å, respectively. The conformational resistance of the main chain of thrombin to urea denaturation at (2.0+X) M concentration of the denaturant is not surprising and is consistent with the sigmoidal curve representing protein denaturation. The conformational changes in thrombin become noticeable only at the inflection point of 3.5 M or higher concentrations of urea when protein unfolding occurs spontaneously within a very narrow concentration range of the denaturant. All residues of the active site of thrombin are essentially positioned optimally and remain so on urea binding in the thrombin-hirugen-urea complex. The sidechain of Glu192 residue, a part of the active site region of thrombin, adopts an extended conformation into solvent region similar to those observed in the structures of Selectide and Molecumetics inhibitors as well as in the native thrombin-hirugen (Vijayalakshmi et al., 1994). The “up” conformation of the

Glu192 side chain is stabilized by a water molecule. Furthermore, the Ramachandran plot of the thrombin-hirugen-urea complex shows that 204 (83.3 %) residues of the total 245 non-glycine and non-proline amino acids of thrombin are in the most favorable regions while 41 (16.7 %) residues occupy additionally allowed areas. There is no residue in disallowed regions.

Five urea molecules have been clearly identified in the thrombin-hirugen complex (Figure 4-1) based on the (2Fo-Fc) and (Fo-Fc) difference electron density maps. Three molecules of the denaturant (Ure1-Ure3) occupy the active site region of thrombin and can directly affect the catalytic activity of the enzyme. Two urea molecules (Ure4 and Ure5) are located next to the heparin binding exosite II and may indirectly affect the catalytic functions of thrombin by causing conformational changes at the active site of the enzyme similar to those observed in heparin-ATIII inactivation of thrombin. It is also known that changes at the fibrinogen exosite I or the sodium binding sites can have allosteric consequences on both kinetic and thermodynamic characteristics of the enzyme (Wells & DiCera, 1992). However, in the thrombin-hirugen complex, the fibrinogen exosite I is occupied with hirugen that not only prevents autocleavage of the enzyme but also occludes the site from being accessible to the urea molecules. Finally, based on the examination of (2Fo-Fc) and (Fo-Fc) difference electron density maps, the two Na⁺ binding sites of thrombin-hirugen remain intact. To summarize, the primary binding sites of urea in thrombin are in the active site region and in the heparin binding exosite of the enzyme.

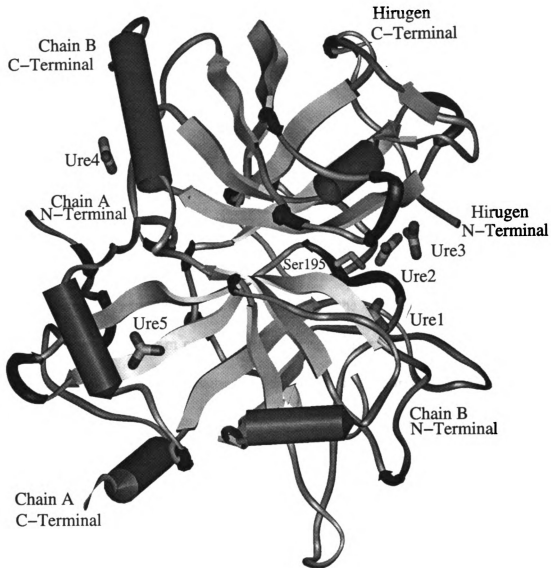


Figure 4-1: Urea binding sites of thrombin. Ure1-Ure3 are in the active site region of the enzyme. Ure4 and Ure5 are located adjacent to heparin binding exosite II. Ser195 of the catalytic triad is numbered.

Three urea molecules and two water molecules (O_w698 & O_w700) adopt a semicircular arrangement at the catalytic site stretching from the S1 to the S2 site (Figure 4-2). The first urea molecule (Ure1) lies in the upper part of the specificity pocket, a subsite that accommodates arginine-like residues. It is located approximately 3.3 Å below Ser195-OG and 5.5 Å above Asp189-OD, at the position corresponding to CG-CD-NE of Arg in thrombin-bound PPACK. This urea molecule is incorporated into an extensive solvent network interacting with the enzyme *via* multiple hydrogen bonds and is stabilized through two hydrogen bond bridges (O_w668 and O_w695) that link Ure1-N1 and Gly219-O (Figure 4-3). In addition, two water molecules (O_w661 and O_w668) mediate hydrogen bonding between the urea molecule and the carboxylate oxygen atoms of Asp189. The overall pattern of the interactions at the specificity site is essentially the same as that found in crystal structures of thrombin-PPACK and many other thrombin-inhibitor complexes in which an Arg residue occupies the P1 position.

The second urea molecule (Ure2) interacts directly with Ser195-OG of the catalytic triad and with Gly193-N (2.71 Å) and Ser195-N (3.14 Å) of the oxyanion hole of thrombin (Figure 4-3). It is positioned to mimic inhibitors with C-terminal carboxylate or terminal peptide-like groups. Water molecule O_w698 adopts a distorted tetrahedral arrangement coordinated by Ser195-OG, O_w683 and nitrogen atoms of Ure1 and Ure2 positioned almost edge to plane to each other (Figure 4-2). A third urea molecule (Ure3) located in the active site region of thrombin utilizes dipole- π electron mediated interactions with Tyr60A (2.70 Å) and Trp60D (3.70 Å) at the P2 subsite similar to

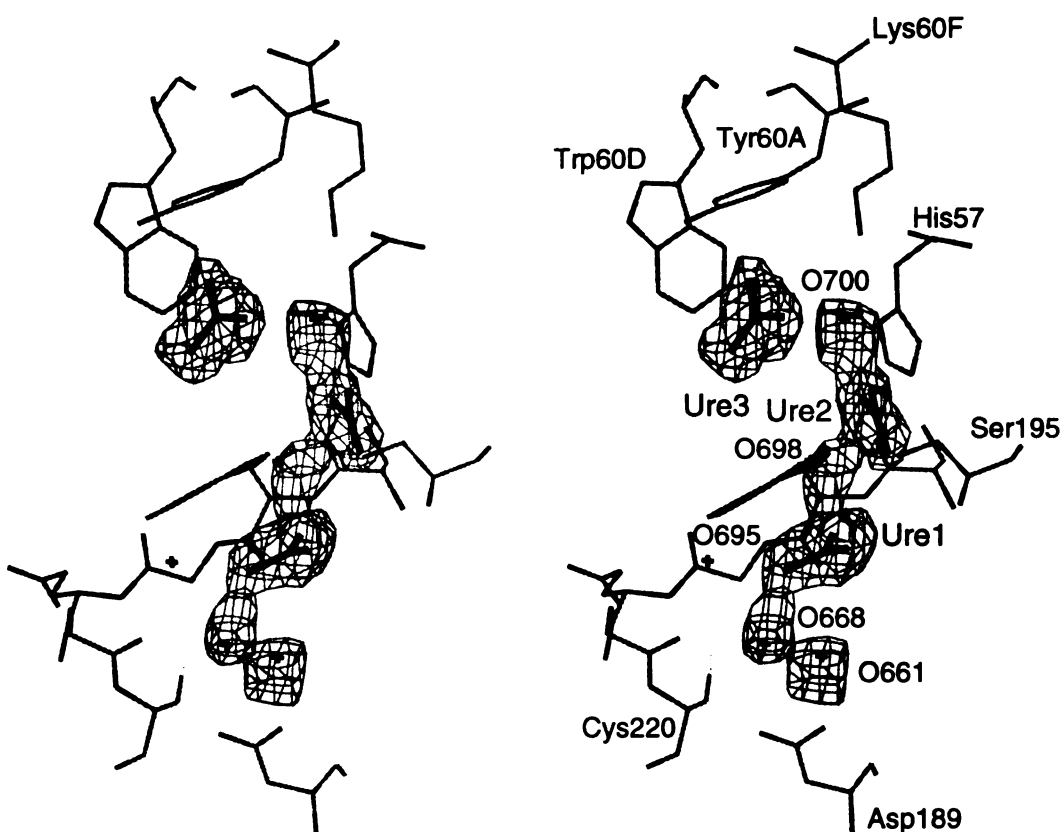


Figure 4-2: Stereoview of the final (2Fo-Fc) electron density of Ure1-Ure3 and their water environment. Contoured at the 1σ level.

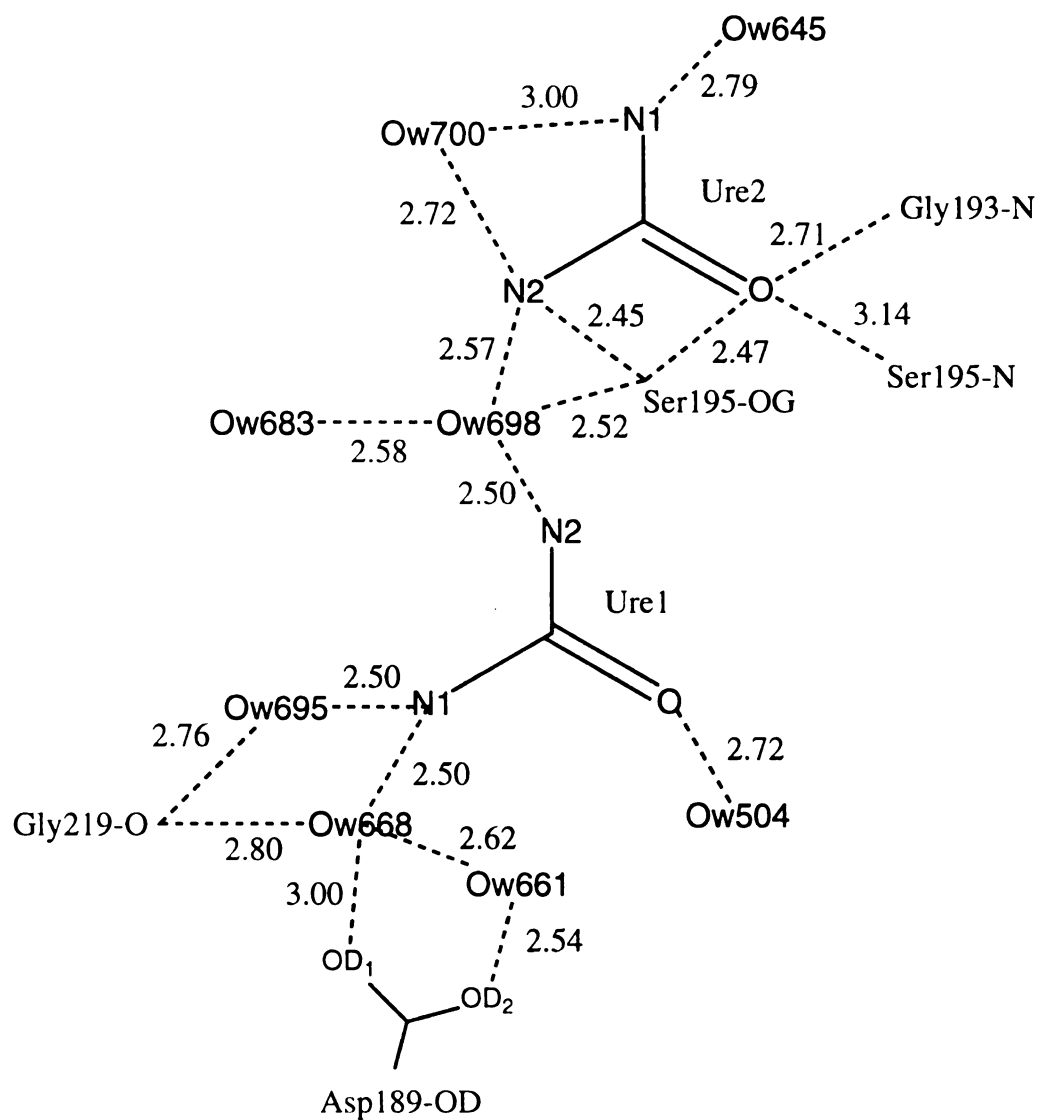


Figure 4-3: Schematic diagram of hydrogen bonds (Å) formed by Ure1 and Ure2 in the active site region of the thrombin-hirugen complex.

cation- π electron mediated interaction (Lin & Johnson, 1995). The π -faces of the aromatic residues, Tyr60A and Trp60D are sufficiently rich in π -electrons to form not only a hydrophobic pocket but also a cation-like recognition site. In addition, Ure3-O forms hydrogen bonds with a water molecule O_w664 (3.00 Å) and His57-ND1 (3.14 Å) of the catalytic triad of thrombin. The fourth and fifth urea molecules are located on two different sides of the heparin binding exosite II, approximately 5 to 7 Å away from residues Arg126, Arg165, Lys169 and Arg233 and about 20 Å apart from each other (Figure 4-1). These two urea molecules form hydrogen bonds with the Leu130-O (2.98 Å) and Ile147-O (3.00 Å) of thrombin, respectively.

The two Na⁺ binding sites remain intact in the thrombin crystal structure upon treatment with urea and are practically identical to those described previously (Zhang & Tulinsky, 1997). Resistance to the effect of urea is most likely due to the extensive water network around the sodium ions and octahedral coordination of Na⁺ by water molecules. This results in the low mobility of the ions reflected in their occupancies, and temperature factors ($\Omega = 1.00$, $B = 20.6 \text{ Å}^2$, intramolecular Na⁺; $\Omega = 1.00$, $B = 18.7 \text{ Å}^2$, intermolecular Na⁺).

Even though thrombin shows no significant signs of denaturation in (2.0+X) M urea, we have been able to crystallographically detect the primary urea binding sites in thrombin. It comes as no surprise that the binding of the denaturant is selective: all primary urea binding sites are localized at the operating functional sites of thrombin, such as the active site and the heparin binding exosite II. Identification and analysis of the effect of urea on the crystal structure of thrombin-hirugen should be valuable to other

biochemical and crystallographic studies of blood proteins that are less resistant to denaturation effects. It should also provide better insights into the molecular mechanism of the denaturation process.

CHAPTER 5:

STRUCTURE-FUNCTION DETERMINANTS OF RECOMBINANT KRINGLE 5 OF HUMAN PLASMINOGEN

I. INTRODUCTION

Kringle domains found in non-catalytic regions in a variety of blood proteins (Figure 1-7 & 1-8; Table 1-2) are autonomous structural and functional modular units of approximately 80 amino acids cross-linked by three disulfide bridges in the pattern 1-6, 2-4 and 3-5 (Figure 1-9). Biological evidence suggests the ligand binding site (LigBS) of most kringles binds lysine residues of several different proteins, such as fibrin, α 2-antiplasmin, tetranectin and thrombospondin, and is important in molecular recognition. These kringle-dependent interactions can be competitively inhibited by lysine zwitterionic analogues such as ϵ -aminocaproic acid (EACA), 5-aminopentanoic acid (5-APnA), 7-aminoheptanoic acid (7-AHpA), benzamidine, trans-4-aminomethylcyclohexane-1-carboxylic acid (t-AMCHA) and p-amino-methylbenzoic acid (Table 5-1). Three-dimensional structures of uncomplexed and liganded kringle domains of most of the foregoing kringles have been determined by X-ray crystallography (Mathews et al., 1996 and references therein) and by NMR methods (Rejante & Llinas, 1994 ^{a,b}).

Association of plasmin(ogen) with fibrin via the LigBS located in the K1 and K4 module essential for the degradation of the fibrin clot. The LigBS of plasminogen K1 (Wu et al., 1994; Mathews et al., 1996) and K4 (Mulichak et al., 1991; Wu et al., 1991) is characterized by short arrays of residues His33-Arg35 (His33-Lys36 in K4), Asp55-

Asp57, Trp62-Phe64 (Tyr62-Tyr64 in K1) and Arg71-Tyr74, which together are arranged in a relatively open elongated surface depression approximately 9 Å wide and 12 Å long. Aromatic residues Trp62, Phe64, Trp72 and Tyr74 form a hydrophobic pocket bounded by a negatively charged anionic subsite containing Asp55 and Asp57 at the top and a positively charged cationic one containing Arg35 or Lys36 and Arg71 at the bottom. In a display of polar efficiency, the anionic and cationic centers of the LigBS simultaneously stabilize the amino and carboxylic group of ligands (Wu et al., 1991; Mathews et al., 1996). The hydrophobic pocket interacts with the methylene chain of ω -amino acids so that all three binding centers of the LigBS are simultaneously operative in an unrestricted way.

Table 5-1: K_d (μM) values of ligand binding to various kringle modules.

Kringles	5-APnA	EACA	7-AHpA	t-AMCHA	PnA	HxA
K1	24	12	250	1	N/A	N/A
K4	29	26	280	5	3000	N/A
K5	580	140	367	22	168	470
K5[L71R]	47	26	68	3	N/C	N/C
KIV-10	66	20	230	N/A	N/A	N/A
PnA – 5 aminopentane; HxA – 6-aminohexane.						
N/A – Dissociation constant is not available.						
N/C - No fluorescent changes						

Our interest in the structure of plasminogen K5 (Figure 5-1) lies in the fact that this kringle module possesses a unique LigBS. Compared to the LigBS of K1 and K4, the unique feature of the binding region of plasminogen K5 lacks a cationic center

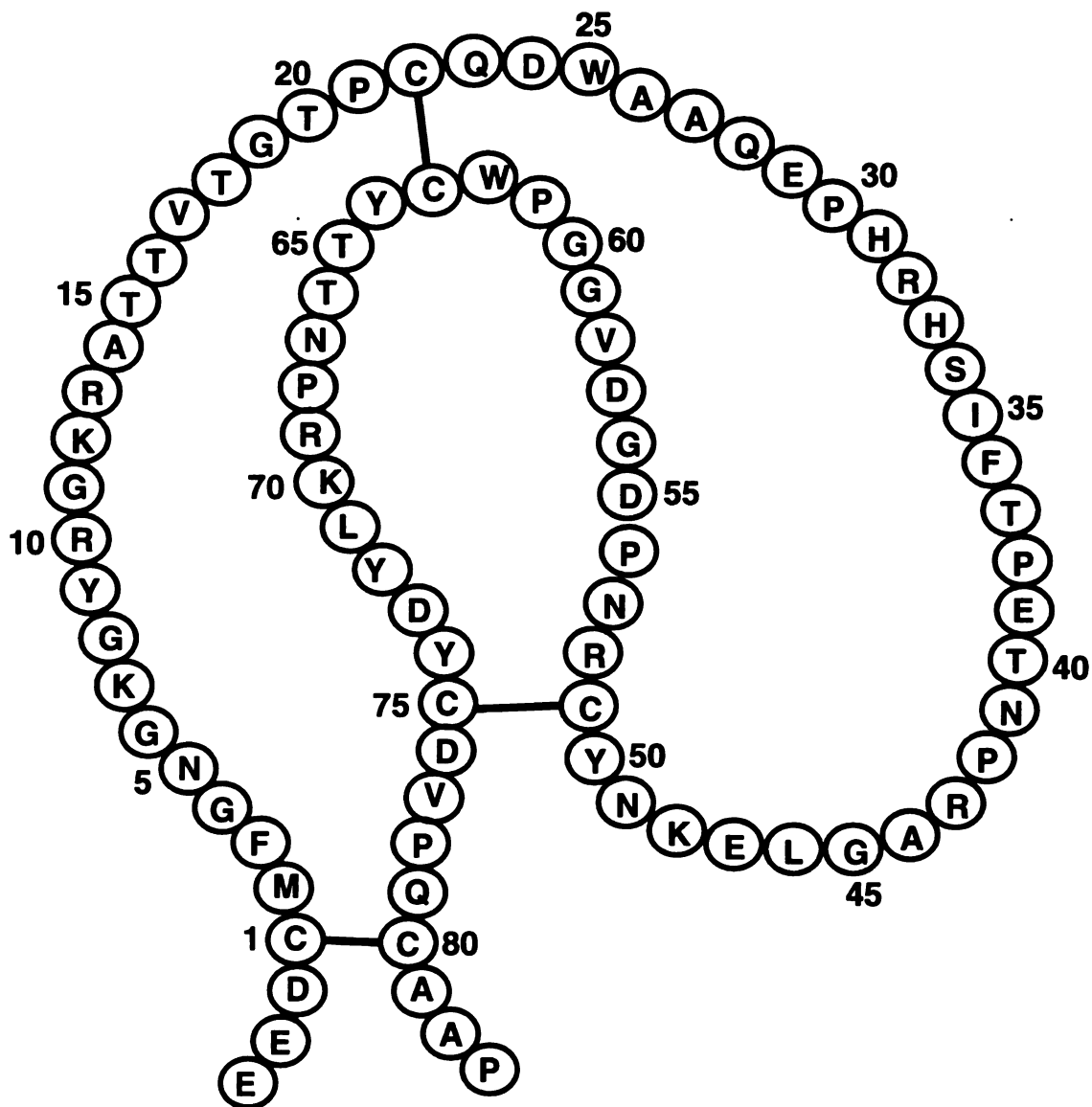


Figure 5-1: Primary structure of human plasminogen K5.

(Tables 5-2 & 5-3) so that a ligand can only occupy the anionic binding site. In K5 the residues positioned near the cationic center are His33-Ser34-Ile35 and Leu71. As a result, the affinities of this domain for ω -amino acids are weaker, but better for alkylamines as compared to K1 and K4 (Table 5-1). In order to test the role of Leu71 in the LigBS, the recombinant Leu71Arg mutant (K5[L71R]) was constructed, expressed and purified (Chang et al., 1998). Using intrinsic fluorescence titration methods, the dissociation constants of the wild type K5 domain as well as the K5[L71R] mutant have been determined for various inhibitors. Comparison of the K_d values of K5 and K5[L71R] for ligand binding clearly demonstrates the enhancement of the ω -amino acids for the mutant kringle (Table 5-1). The binding extent of K5[L71R] for EACA is essentially the same as that of K4 and only 2-fold weaker than the best affinity of K1.

II. EXPERIMENTAL

Crystallization: Previous experience with crystallization of kringles (Mulichak et al., 1991; Wu et al., 1994; Mathews et al., 1996) suggested that PEG 400 to 8000 in Na^+ -Hepes buffer, pH 7, with or without salt, would be more suitable for crystallizing K5. An in-house set of factorial solutions utilizing Li_2SO_4 as salt (protein concentration of 60 mg/mL) produced small crystals using the vapor diffusion hanging drop method, from a solution of 24% PEG 8000/0.15 M Li_2SO_4 in 0.1 M Na^+ -Hepes, pH 7.0. The reservoir contained 1 mL of the solution, and the hanging drop consisted of 1 μL of the protein solution added to 1 μL of the reservoir solution. X-ray diffraction quality crystals with

	1	5	10	15	20	25	30	35	40																															
K1	C	K	T	G	N	G	K	N	Y	R	G	T	M	S	K	T	K	N	G	I	T	C	Q	K	W	S	S	T	S	P	H	R	P	-	R	F	S	P	A	T
K2	C	M	H	C	S	G	E	N	Y	D	G	K	I	S	K	T	M	S	G	L	E	C	Q	A	W	D	S	Q	S	P	H	A	H	-	G	Y	I	P	S	K
K3	C	L	K	G	T	G	E	N	Y	R	G	N	V	A	V	T	V	S	G	H	T	C	Q	H	W	S	A	Q	T	P	H	T	H	-	N	R	T	P	E	N
K4	C	Y	H	G	D	G	Q	S	Y	R	G	T	S	S	T	T	T	T	G	K	K	C	Q	S	W	S	S	M	T	P	H	R	H	-	Q	K	T	P	E	N
K5	C	M	F	G	N	G	K	G	Y	R	G	K	R	A	T	T	V	T	G	T	P	C	D	W	A	A	Q	E	P	H	R	H	S	I	F	T	P	E	T	
KIV	C	Y	H	G	N	G	Q	S	Y	R	G	T	F	S	T	T	V	T	G	R	T	C	Q	S	W	S	S	M	T	P	H	R	H	-	Q	R	T	P	E	N

	45	50	55	60	65	70	75	80																																
K1	H	P	S	E	G	L	E	E	N	Y	C	R	N	P	D	N	D	P	Q	G	P	W	C	Y	T	T	D	P	E	K	R	Y	D	Y	C	D	I	L	E	C
K2	F	P	N	K	N	L	K	K	N	Y	C	R	N	P	D	R	E	L	-	R	P	W	C	F	T	T	D	P	N	K	R	W	E	L	C	D	I	P	R	C
K3	F	P	C	K	N	L	D	E	N	Y	C	R	N	P	D	G	K	R	-	A	P	W	C	H	T	T	N	S	Q	V	R	W	E	Y	C	K	I	P	S	C
K4	Y	P	N	A	G	L	T	M	N	Y	C	R	N	P	D	A	D	K	-	G	P	W	C	F	T	T	D	P	S	V	R	W	E	Y	C	N	L	K	K	C
K5	N	P	R	A	G	L	E	K	N	Y	C	R	N	P	D	G	D	V	G	G	P	W	C	Y	T	T	N	P	R	K	L	Y	D	Y	C	D	V	P	Q	C
KIV	Y	P	N	D	G	L	T	M	N	Y	C	R	N	P	D	A	D	T	-	G	P	W	C	F	T	M	D	P	S	I	R	W	E	Y	C	N	L	T	R	C

Table 5-2: Primary structures of plasminogen K1-K5 and apo(a) KIV-10 aligned in respect of K1 to maximize conservation. The number of residues (1-80) standardized to plasminogen K5. Deletion sites at positions 34 and 59 are indicated with -. Sequence conservation is boxed: identical residues are in black boxes, homologous residues are in the gray boxes.

	1	5	10	15	20	25																					
human-K5	C	M	F	G	N	G	K	G	Y	R	G	K	R	A	T	T	V	T	G	T	P	C	Q	D	W	A	A
bovine-K5	C	M	I	G	T	G	K	S	Y	R	G	K	K	A	T	T	V	A	G	V	P	C	Q	E	W	A	A
porcine-K5	C	M	F	G	N	G	K	R	Y	R	G	K	R	A	T	T	V	A	G	V	P	C	Q	E	W	A	A
murine-K5	C	M	Y	G	N	G	K	D	Y	R	G	K	T	A	V	T	A	A	G	T	P	C	Q	G	W	A	A

	30	35	40	45	50																						
human-K5	Q	E	P	H	R	H	S	I	F	T	P	E	T	N	P	R	A	G	L	E	K	N	Y	C	R	N	P
bovine-K5	Q	E	P	H	Q	H	S	I	F	T	P	E	T	N	P	Q	S	G	L	E	R	N	Y	C	R	N	P
porcine-K5	Q	E	P	H	R	H	S	I	F	T	P	E	T	N	P	R	A	G	L	E	K	N	Y	C	R	N	P
murine-K5	Q	E	P	H	R	H	S	I	F	T	P	Q	T	N	P	R	A	D	L	E	K	N	Y	C	R	N	P

	55	60	65	70	75	80																				
human-K5	D	G	D	V	G	G	P	W	C	Y	T	T	N	P	R	K	L	Y	D	Y	C	D	V	P	Q	C
bovine-K5	D	G	D	V	N	G	P	W	C	Y	T	M	N	P	R	K	P	F	D	Y	C	D	V	P	Q	C
porcine-K5	D	G	D	D	N	G	P	W	C	Y	T	T	N	P	Q	K	L	F	D	Y	C	D	V	P	Q	C
murine-K5	D	G	D	V	N	G	P	W	C	Y	T	T	N	P	R	K	L	Y	D	Y	C	D	I	P	L	C

Table 5-3: Primary structures of the plasminogen K5 modules from the various species.

Identical residues are boxed.

dimensions of 0.3 x 0.15 x 0.05 mm were obtained by multiple macroseeding while the PEG 8000 concentration was stepwise decreased from 24% to 22% to 18% and to 16%.

Intensity Data Collection: The X-ray diffraction data of the K5 crystals were collected with an R-Axis II imaging plate detector using Molecular Structure Corporation-Yale focusing mirrors. The CuK α radiation was generated with a Rigaku RU200 rotating anode operating at 5kW power with a fine focus filament (0.3 mm x 3.0 mm). Intensity data were measured at -150° C with crystals flash frozen in a cryosolvent consisting of 20% PEG 4000, 20% glycerol, and 0.1M Na⁺-Hepes buffer, pH 7.0. The crystal-detector distance was 10.0 cm and the detector-swing angle was 12 degrees. Autoindexing and processing of the diffraction data were carried out with the Rigaku R-Axis software package (Higashi, 1990). The intensity data collection statistics are summarized in Table 5-4.

Structure Determination: The crystal structure of K5 was determined by the molecular replacement method using the coordinates of K1 (Wu et al., 1994) as an initial model. All residues of K1 that differed in sequence from K5 were replaced with alanine in the model and both N- and C-terminal interkringular peptides were ignored. Rotation/translation searches were carried out with the program AMoRe (Navaza et al., 1994) in the range 10.0-3.0 Å resolution. The rotation search provided two separate solutions with correlation coefficients of 0.28 and 0.19, corresponding to each of the two molecules in the asymmetric unit related by a non-crystallographic 2-fold rotation axis approximately parallel to the z-direction. The two solutions were also in agreement with the self-rotation search. The position of one molecule of the rotated structure was determined with a translation function that provided a R-value of 46.8% (correlation

Table 5-4: Crystal data and intensity data collection statistics of plasminogen K5.

Space group	P2 ₁ 2 ₁ 2
Cell constant (Å)	
a	77.43
b	79.20
c	30.78
Molecules/asymmetric unit	2
Solvent fraction (%)	42
Resolution (Å)	1.66
Observations	42,366
Independent reflections	16,720
Redundancy	2.5
Outermost range (Å)	1.80-1.66
R-merge (%)	4.2
outermost range (%)	8.7
Completeness (%)	72
outermost range (%)	37
I/σ outermost range	3.6

coefficient, 0.31); R was 50.1% (correlation coefficient 0.18) for the other molecule. The coordinates a model of both rotated and translated molecules were optimized by rigid body refinement to an R-value of 38.1% (correlation was 0.59). Additional positional refinement was conducted with the X-PLOR program package (Brunger et al., 1992) by energy minimization to relieve close contacts of the model. This decreased the R-value to 33.5%.

Structure Refinement: The dimeric structure of K5 was refined with the program PROLSQ (Hendrickson, 1985) and PROFFT (Finzel, 1987). The first stage of refinement, (7.0 - 2.8) Å, employed tight positional non-crystallographic symmetry (NCS) and an overall B-factor. After three cycles, starting with an average $B = 15 \text{ Å}^2$, the R-value only changed from 35.7% to 34.5%. The next three cycles of the individual B-refinement with tight NCS geometry and temperature factors restraints for atoms of both the main and side chains (0.5 Å, 3.0 Å²), and six cycles of looser NCS geometry and temperature factors restraints (1.0 Å, 10.0 Å²) converged at R = 31.9%. Continued refinement using alternate tight-loose NCS restraints decreased R-value to 26.5%. At this stage, the (2Fo-Fc) electron and (Fo-Fc) difference density maps had most of the side chains of K5 that were arbitrary alanine residues in the initial model. The insertion residue (Ser34) with respect to K1 (Table 5-2) was fixed based on an analysis of the electron and difference density maps after deletion and then reconstruction of the His33-Thr37 region. This clearly showed that the insertion was Ser34. Further refinement was conducted including most of the K5 residues in the model. Solvent water molecules was found and added periodically, by examining difference density maps, during the refinement at 2.5 Å and higher resolution. In the final stage, beyond 2.1 Å resolution,

the NCS restraints were abandoned and the two molecules of K5 were refined independently. Summary of final refinement parameters and their deviations are listed in Table 5-5.

III. RESULTS

The X-ray crystal structure of K5 has been employed to examine critical aspects of the LigBS of this kringle as it relates to those of K1 and K4. Crystals of K5 were grown in the orthorhombic space group $P2_12_12$ with dimensions 77.43 Å x 79.20 Å x 30.78 Å (Table 5-4). The asymmetric unit contains two molecules (Figure 5-2) related by a non-crystallographic 2-fold rotation axis approximately parallel to the z-direction. The final model converged at an R-value of 16.6% at 1.66 Å resolution and consisted of 1294 protein atoms and 193 water molecules with occupancy > 0.6. Two K5 molecules in the dimer are almost identical, which especially applies to the LigBS regions (Figure 5-3). The rmsΔ between the superimposed CA atoms of the two K5 units is 0.31 Å. The Ramachandran plot of the two kringles, generated by PROCHECK (Laskowsky et al., 1993) shows that 106 (81.5%) residues of a total of 130 non-glycine and non-proline amino acids occupy the most favorable regions, while 22 (16.9%) residues are in additionally allowed areas (Figure 5-4). Similar to other kringle structures, such as K1 (Mathews et al., 1996), K4 (Mulichak et al., 1991), KIV-10 (Mikol et al., 1996; Mochalkin et al., 1999), the residue Lys48, which is located in a hairpin turn, is in a generously allowed area of the Ramachandran plots. Notwithstanding, it has well-defined main and side chain electron density. Moreover, the Ramachandran plot of K5 indicates a close coupling of the ϕ , ψ angles of the two molecules in a K5 dimer (Figure 5-4).

Table 5-5: Refinement summary of deviations from ideality and individual refinement statistics of plasminogen K5.

	Target	RMSΔ
Distances (Å):		
Bond lengths	0.020	0.016
Bond angles	0.040	0.034
Planar 1-4	0.050	0.050
Planes (Å):		
Peptides	0.030	0.032
Aromatic groups	0.030	0.030
Chiral volumes (Å ³)	0.130	0.160
Non-bonded contacts (Å)		
Single torsion	0.60	0.18
Multiple torsion	0.60	0.24
Possible H-bond	0.60	0.23
Thermal parameters (Å ²)		
Main-chain bond	1.5	1.4
Main-chain angle	2.0	1.7
Side-chain bond	2.5	3.5
Side-chain angle	3.0	4.1
Refinement range (Å)		9.0-1.66
Final R-value (%)		16.6
 (Å ²)		15.4
PDB code		5HPG

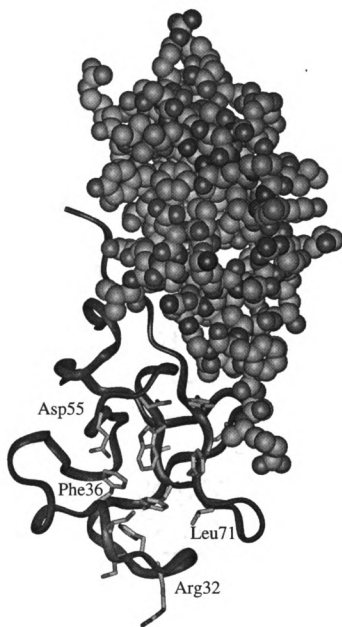


Figure 5-2: Space filling and ribbon representations of plasminogen K5 dimer viewed approximately down the local 2-fold axis in the crystal structure.

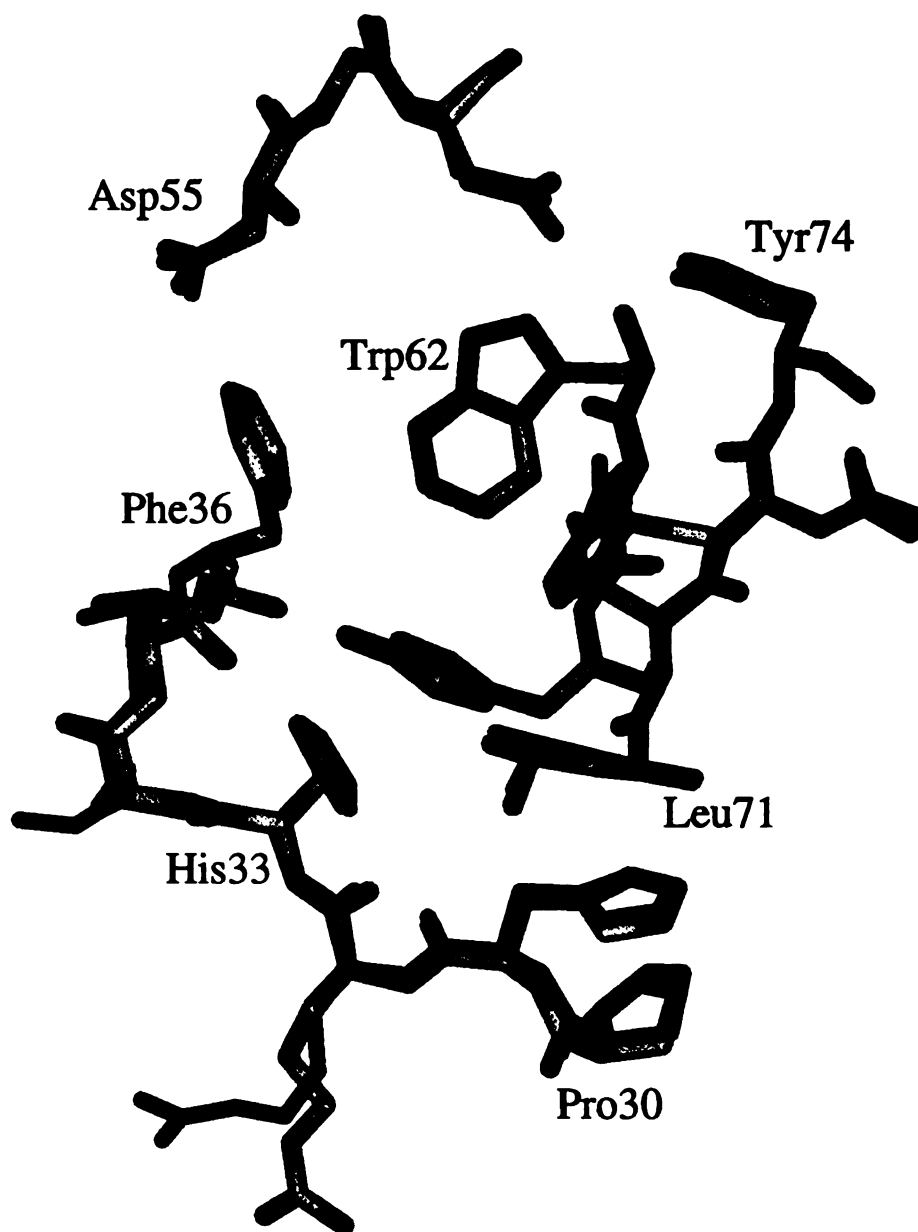


Figure 5-3: Superposition of the LigBS of K5 molecules 1 and 2 of the dimer. A residue of each segment is labeled. Molecule 1 is shown in gray. Molecule 2 is colored black.

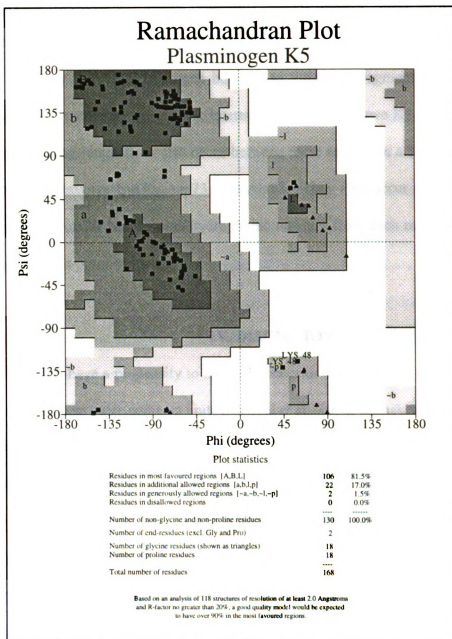


Figure 5-4: The Ramachandran plot of the final structure of the K5 dimer. Lys48 residues of molecule 1 (Lys148) and molecule 2 (Lys248) are in a generously allowed region. Glycine residues are shown as triangles. Non-glycine residues are shown as squared boxes.

The interface between the 2-fold rotation-related kringles consists of an antiparallel arrangement of the Leu71-Pro83 segments of the C-terminus of each kring molecule. Of the 34 water molecules hydrogen bonding to these two segments, seven hydrogen bond with both, mediating the dimeric interaction between the two K5 molecules. Direct hydrogen-bonding interactions also occur between these regions in a 2-fold manner, near their ends, and involve Lys70, Asp73, and Ala82. They also occur close to the 2-fold axis involving Asp76, while seven hydrophobic residues from each molecule participate in stabilizing non-polar contacts.

IV. DISCUSSION

Kringles have a propensity to naturally occur in tandem arrays and associate with one another to form more compact units (Mangel et al., 1990; Ramakrishnan et al., 1991). The latter consideration also applies to crystal structures when more than one kring molecule constitutes the asymmetric unit (Padmanabhan et al., 1994; Mathews et al., 1996). Although a number of different arrangements utilizing local rather than crystallographic symmetry elements have been identified in crystal structures, such as 2-fold rotation axis and pseudo 2_1 and 3_1 screw axis-like symmetries, none corresponds to the unique antiparallel C-terminal interface arrangement of K5 in which 389 Å² or 13% of kring surface area is involved in the formation of the dimer (Figure 5-2). However, the most distinguishing characteristic of the K5 dimer, as compared to those of other kringles, is the large number of water molecules mediating the dimer interface. Whether this or any of the other known kring associations in crystals are physiologically relevant remains to be shown by structure determination of a tandem kring array.

Similar to other kringle modules, the LigBS of K5 is defined by His33-Ile35, Asp55-Asp57, Trp62-Tyr64 and Leu71-Tyr74 as an elongated depression on the kringle surface that is lined with solvent molecules (Figure 5-5). The hydrogen bonds that occur in the LigBS of K5 are provided in Table 5-6. Those listed from Asn53 to Thr65, especially when involving water molecules, reflect the excellent 2-fold symmetry between two independent K5 molecules. The 16 hydrogen bonds in each LigBS strongly support the notion that the LigBS of kringles is essentially preformed prior to, and independent of, ligand binding. In the case of molecule 1, residues Gln28, Arg32 and Ser34 from molecule 2 of the neighboring dimer make several hydrogen bonds that do not possess 2-fold symmetry because of local rather than crystallographic symmetry relationships between these contacts. The Arg32 of the neighboring molecule limits access to the LigBS of molecule 1 in the crystal (Figure 5-5). However, in molecule 2, the LigBS is essentially unobstructed. It is also noteworthy that the LigBS of molecule 1 is systematically occupied by eight water molecules (Figure 5-5), while molecule 2, which lacks intermolecular contacts in this location, contains only five water molecules. In fact, the (2Fo-Fc) electron density of seven of the water molecules of molecule 1 is connected at the 1σ level, resembling that of ligand binding. This apparent relationship to ligand binding is probably the result of the intermolecular interaction of the LigBS with Gln28, Arg32 and Ser34 from the adjacent molecule(s), which decreases the mobility of the water molecules at the LigBS of K5.

In plasminogen K1, K4 and apo(a) KIV-10, residue Asp55 of the LigBS is in the same orientation in both inhibited and unliganded kringles ($\chi_1 \approx 60^\circ$). In plasminogen

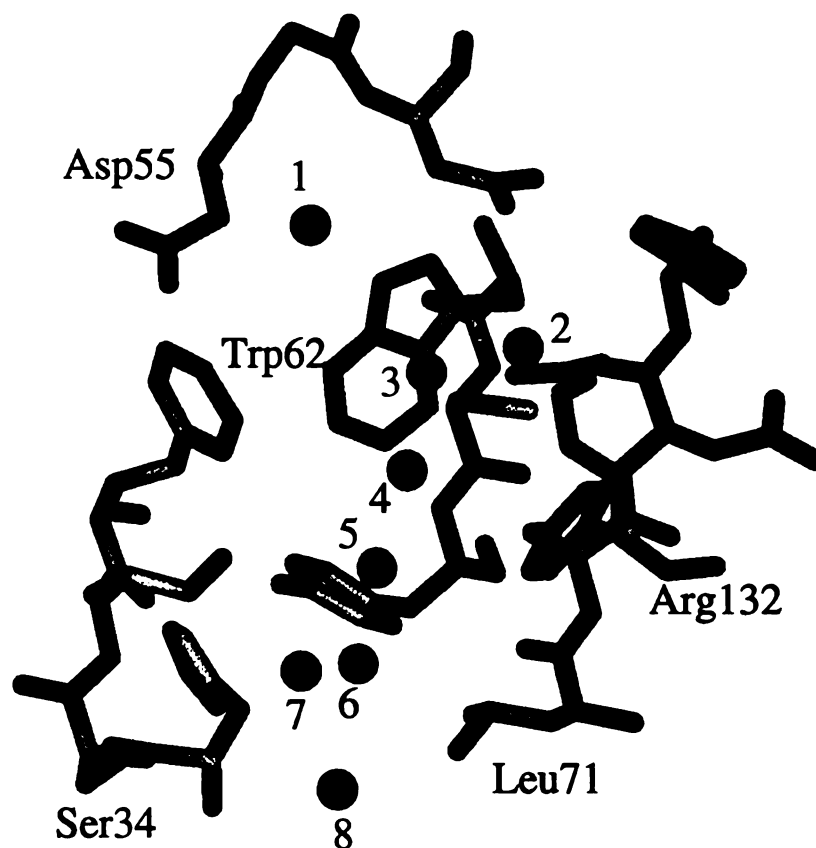


Figure 5-5: Specific residues of K5 potentially present in the LigBS of molecule 1 of the K5 dimer. The regions of the kringle represented are His33-Phe36, Asp55-Asp57, Trp62-Tyr64, Leu71-Tyr74 are shown in gray. A residue of each segment is labeled. Water molecules common to the LigBS of both kringles of the dimer are 1, 4, 5, 7 and 8; those in molecule 1 only are 2, 3 and 6. 1=O_w524, 2=O_w525, 3=O_w419, 4=O_w565, 5=O_w589, 6=O_w524, 7=O_w526, 8=O_w527 of the PDB coordinate file (PDB code: 5HPG).

Table 5-6: Hydrogen bonds (Å) in the LigBS region of K5

		Molecule 1	Molecule 2
Asn53-OD1	Asp57-N	2.84	2.83
-OD1	Ow540/470	2.93	2.97
-ND2	Asn57-O	2.70	2.70
-ND2	Asp57-O	2.97	2.94
Asp57-OD1	Gly59-N	2.81	2.93
-OD2	Gly60-O	3.12	2.95
Pro61-O	Cys75-N	3.02	2.99
Trp62-N	Arg52-O	2.91	2.87
-O	-N	2.94	2.98
-NE1	Ow540/470	3.10	2.92
Cys63-N	Asp73-O	2.87	2.96
-O	-N	2.93	2.98
Tyr64-N	Ow502/418	2.97	2.84
-O	Gln23-N	2.61	2.81
-OH	Ow526/425	2.71	2.84
Thr65-N	Lys71-O	2.97	2.92
Tyr72-N	Ser34-OG ^a	2.90	
-O	Ser34-OG ^a	2.55	
-O	Ow582 ^a		2.76
-OH	Ow491 ^a	2.64	
-OH	Gln28-OE1 ^a	2.81	
Tyr74-OH	Arg32-NH2 ^a	3.00	

^a Neighboring molecules in crystal; hydrogen bond has no local 2-fold rotation equivalent.

K5, however, Asp55 extends into the solvent region in a different orientation ($\chi_1 \approx 100^\circ$) apparently nonconductive to ligand binding (Figures 5-3, 5-5 & 5-6). It is most likely that the Asp55 residue of K5 reorients on ligand binding to form a doubly charged anionic center along with Asp57, which then interacts with the amino group of ligand as it is observed in the K1 and K4 (Figure 5-6).

When structurally superposed on K1, the segment His33-Thr37 of K5 contains residue Ser34, which is deleted in K1. This part of the kringle molecule appears to possess the most flexibility (de Vos et al., 1991; Arni et al., 1993) and represents the only region of the backbones of a number of kringle domains that differ from each other. Although the difference between CA atoms of Ser34 of the two K5 molecules is 1.1 Å, the main chain and side chains of His33 and Phe36 compensate for this displacement of Ser34 and adopt the same orientation in both molecules (Figure 5-3). This is probably due to the hydrophobic edge-face interaction of Phe36 with Tyr64 and Trp62, respectively.

On the basis of similarities with K1 (Menhart et al., 1991; Mathews et al., 1996), and K4 (Wu et al., 1991; McCance et al., 1994), the cationic residues of K5, Arg32, Arg69 and Lys70, are candidates for interactions with the carboxylate of the ligand. However, as highlighted in Figures (5-5) these residues are either clearly removed from the binding pocket of K5 and/or have unfavorable orientations for interaction with the ligand. This is also revealed by the distances of the functional groups from side chains of Asp55, Asp57, Trp62 and Tyr72 residues that are uniformly critically positioned in the LigBS in the other ligand binding kringles. Examination of all other cationic side chains of K5 shows that none can contribute to the putative ligand pocket,

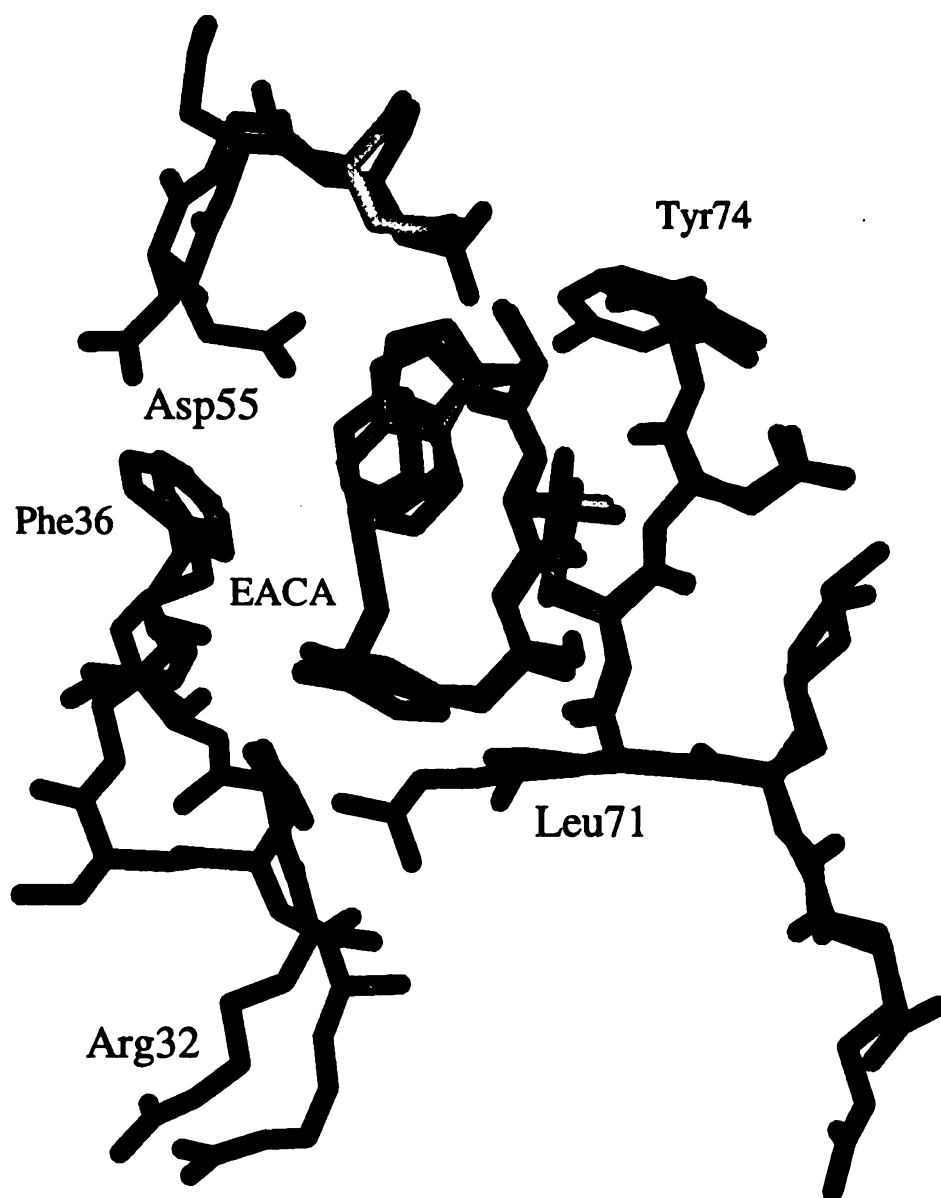


Figure 5-6: Superposition of the LigBS of K1 and K5 of plasminogen. K5 is in gray; K1 is shown in black. K1 is inhibited with EACA in the EACA binding pocket. All residues are numbered according to K5.

except

O. In o

moiety

180°.

was co

Using

ApnA

first fo

AMC

obser

the sa

the h

natur

hydro

wild

poly

for

mon

liga

except possibly His33 (Figure 5-6). However, His33-NE is hydrogen bonded to Trp25-O. In order for this group to play a role as a cationic donor to the ligand carboxylate moiety, the orientation of the imidazole would have to be repositioned by approximately 180°.

To test the importance of Leu71 in the LigBS, the recombinant mutant K5[L71R] was constructed, expressed in the same yeast system, and purified (Chang et al., 1998). Using intrinsic fluorescence titrations, the binding (dissociation) constants of K5 for 5-ApnA, EACA, 7-AhpA, t-AMCHA, PnA and HxA have been determined along with the first four for K5[L71R] (Table 5-1). The magnitudes of the latter, and in the case of t-AMCHA, the sign, of the fluorescence changes possess some differences to those observed for K5. This indicates that the environment of the residues, likely Trp62, is not the same in two ligand-bound kringles. Such dissimilarities could include the nature of the hydration of the LigBS, the orientation of the ligand in the binding pocket, and/or the nature of side-chain-reporter group contacts. Since the ligand must bind in the more hydrophilic pocket of the mutant quite differently than in the more hydrophobic site of its wild-type counterpart, the differences in the intrinsic fluorescence of the ligand-saturated polypeptide are not surprising. In any case, this intrinsic fluorescence change is saturable for all of the ligands that exhibit binding, and, thus can be effectively employed to monitor ligand-kringle interactions (Chang et al., 1998). The K_d values obtained for the ligand-K5[L71R] interactions are also compared to those for other kringle domains.

Al

and K

via th

has b

(Scar

al., 1

expr

mon

the

a un

role

con

Ar

Lig

no

hu

CHAPTER 6:

RECOMBINANT KRINGLE KIV-10 MODULES OF HUMAN APOLIPOPROTEIN (a): STRUCTURE, LIGAND BINDING MODES AND BIOLOGICAL RELEVANCE

I. INTRODUCTION

Plasminogen is known to bind lysine-Sepharose via the LigBS located in the K1 and K4 modules (Miles & Plow, 1990) while Lp(a) (Figure 1-8) binds lysine-Sepharose via the LigBS located in apo(a) KIV-10 (Figure 1-9; Table 6-1) (Eaton et al., 1987). This has been supported by studies of defective lysine binding of intact rhesus monkey Lp(a) (Scanu et al., 1993) and human Lp(a), both with Arg72Trp mutation in KIV-10 (Scanu et al., 1994), as well as wild type KIV-10 (Trp72) and mutant KIV-10 (Arg72) individually expressed in *E.Coli* (LoGrasso et al., 1994; Klezovitch & Scanu, 1996). Apo(a) of rhesus monkey differed from the human counterpart by having: 1) an unpaired catalytic triad in the protease region, 2) Arg in position 72 in the LigBS of KIV-10 and 3) no copy of KV, a unit located between KIV-10 and the catalytic domain of human apo(a). The critical role of Trp72 in lysine binding, which has emerged from the above observations, is consistent with the results of the structure of PT-K1 (Tulinsky et al., 1988), which has an Arg in position 72. This residue extends diagonally across the region equivalent to the LigBS, occupying it and sterically precluding binding of lysine-like ligands. The same notion also emerges from the modeling of the LigBS of rhesus (Scanu et al., 1993) and human mutant KIV-10/W72R (Scanu & Edelstein, 1995).

KIV-10

K1

K2

K3

K4

K5

PT-K1

PT-K2

I-PA K1

I-PA K2

U-PA K

KIV-10

K1

K2

K3

K4

K5

PT-K1

PT-K2

I-PA K

I-PA K

U-PA

Tab

ma

De

co

the

	1	5	10	15	20	25	30	35	40																																	
KIV-10	C	Y	H	G	N	G	Q	S	Y	R	G	T	F	S	T	T	V	T	G	R	T	C	Q	S	W	S	S	M	T	P	H	R	H	-	Q	R	T	P	E	N	Y	P
K1	C	K	T	G	N	G	K	N	Y	R	G	T	M	S	K	T	K	N	G	I	T	C	Q	K	W	S	S	T	S	P	H	R	P	-	R	F	S	P	A	T	H	P
K2	C	M	H	C	S	G	E	N	Y	D	G	K	I	S	K	T	M	S	G	L	E	C	Q	A	W	D	S	Q	S	P	H	A	H	-	G	Y	I	P	S	K	F	P
K3	C	L	K	G	T	G	E	N	Y	R	G	N	V	A	V	T	V	S	G	H	T	C	Q	H	W	S	A	Q	T	P	H	T	H	-	N	R	T	P	E	N	F	P
K4	C	Y	H	G	D	G	Q	S	Y	R	G	T	S	S	T	T	T	T	G	K	K	C	Q	S	W	S	S	M	T	P	H	R	H	-	Q	K	T	P	E	N	Y	P
K5	C	M	F	G	N	G	K	G	Y	R	G	K	R	A	T	T	V	T	G	T	P	C	Q	D	W	A	A	Q	E	P	H	R	H	S	I	F	T	P	E	T	N	P
PT-K1	C	A	E	G	V	G	M	N	Y	R	G	N	V	S	V	T	R	S	G	I	E	C	L	W	R	S	R	Y	P	H	K	P	-	E	I	N	S	T	T	H	P	
PT-K2	C	V	P	D	R	G	R	E	Y	R	G	R	L	A	V	T	T	S	G	S	R	C	L	A	W	S	S	E	Q	A	K	A	L	-	S	K	D	O	D	F	N	P
t-PA K1	C	Y	E	D	O	G	I	S	Y	R	G	T	W	S	T	A	E	S	G	A	E	C	T	N	W	N	S	S	A	L	A	Q	K	P	Y	S	G	R	R	P	D	A
t-PA K2	C	Y	F	G	N	G	S	A	Y	R	G	T	H	S	L	T	E	S	G	A	S	C	L	P	W	N	S	M	I	L	I	G	K	V	Y	T	A	Q	N	P	S	A
u-PA K1	C	Y	E	G	N	G	H	F	Y	R	G	K	A	S	T	D	T	M	G	R	P	C	L	P	W	N	S	A	T	V	L	Q	Q	T	Y	H	A	H	R	S	D	A

	42A	45	47A	50	55	60	65	70A	75	80																															
KIV-10	-	N	D	G	L	T	-	M	N	Y	C	R	N	P	D	A	D	T	-	G	P	W	C	F	T	M	D	P	S	I	-	R	W	E	Y	C	N	L	T	R	C
K1	-	S	E	G	L	E	-	E	N	Y	C	R	N	P	D	N	D	P	Q	G	P	W	C	Y	T	T	D	P	E	K	-	R	Y	D	Y	C	D	I	L	E	C
K2	-	N	K	N	L	K	-	K	N	Y	C	R	N	P	D	R	E	L	-	R	P	W	C	F	T	T	D	P	N	K	-	R	W	E	L	C	D	I	P	R	C
K3	-	C	K	N	L	D	-	E	N	Y	C	R	N	P	D	G	K	R	-	A	P	W	C	H	T	T	N	S	Q	V	-	R	W	E	Y	C	K	I	P	S	C
K4	-	N	A	G	L	T	-	M	N	Y	C	R	N	P	D	A	D	K	-	G	P	W	C	F	T	T	D	P	S	V	-	R	W	E	Y	C	N	L	K	K	C
K5	-	R	A	G	L	E	-	K	N	Y	C	R	N	P	D	G	D	V	G	G	P	W	C	Y	T	T	N	P	R	K	-	L	Y	D	Y	C	D	V	P	Q	G
PT-K1	-	G	A	D	L	R	-	E	N	F	C	R	N	P	D	G	S	I	T	G	P	W	C	Y	T	T	S	P	T	L	-	R	R	E	E	C	S	V	P	V	C
PT-K2	-	A	V	P	L	A	-	E	N	F	C	R	N	P	D	G	D	E	E	G	A	W	C	Y	V	A	D	Q	P	G	-	D	F	E	Y	C	D	L	N	Y	C
t-PA K1	I	R	L	G	L	G	N	H	N	Y	C	R	N	P	D	R	D	S	-	K	P	W	C	Y	V	F	K	A	G	K	Y	S	S	E	F	C	S	T	P	A	G
t-PA K2	Q	A	L	G	L	G	K	H	N	Y	C	R	N	P	D	G	D	A	-	K	P	W	C	H	V	L	K	N	R	R	L	T	W	E	Y	C	D	V	P	S	C
u-PA K1	L	Q	L	G	L	G	K	H	N	Y	C	R	N	P	D	N	R	R	-	R	P	W	C	Y	V	Q	V	G	L	K	P	L	V	Q	E	C	M	V	H	D	C

Table 6-1: Primary structures of various kringle aligned in respect of apo(a) KIV-10 to maximize conservation. The number of residues (1-80) standardized to plasminogen K5. Deletion sites at positions 34, 42A, 47A, 59 and 70A are indicated with -. Sequence conservation is boxed: identical residues are in black boxes, homologous residues are in the gray boxes.

region

with

from

2) on

Howe

functi

KIV-1

KIV-1

Scanu

deficie

et al.,

Trp62

elonga

earlier

hydro

and A

at the

LigBS

1991;

of ω -

opera

Recently, Chenivasse et al., (Chenivasse et al., 1998) reported that the C-terminal region of apo(a) of chimpanzee (*Pan troglodytes*) exhibits a high degree of homology with the corresponding region of human apo(a). In turn, chimpanzee apo(a) differed from the rhesus counterpart by having: 1) an intact catalytic triad in the protease region, 2) one copy of the KV motif and 3) Trp in position 72 in the LigBS of KIV-10. However, like the rhesus product, chimpanzee Lp(a) was deficient in lysine binding, a functional impairment attributed to the presence of Asn instead of Asp at position 57 in KIV-10 (Chenivasse et al., 1998). Another sequence polymorphism, Met66Thr, in apo(a) KIV-10 has also been reported in 42-50% of the human population (Kraft et al., 1995; Scanu et al., 1995). This mutation, however, is silent and causes no lysine binding deficiency in Lp(a).

Like K1 and K4 of human plasminogen, the LigBS of apo(a) KIV-10 (Mochalkin et al., 1999) is characterized by short arrays of residues His33-Arg36, Asp55-Asp57, Trp62-Phe64 and Arg71-Tyr74, which together are arranged in a relatively open elongated surface depression of approximately 9 Å wide and 12 Å long. As mentioned earlier (Chapter 5), aromatic residues Trp62, Phe64, Trp72 and Tyr74 form a hydrophobic pocket bounded by a negatively charged anionic sub-site containing Asp55 and Asp57 at the top and the positively charged cationic one containing Arg36 and Arg71 at the bottom. In a display of polar efficiency, the anionic and cationic centers of the LigBS simultaneously stabilize the amino and carboxylic group of ligands (Wu et al., 1991; Mathews et al., 1996). The hydrophobic pocket interacts with the methylene chain of ω-amino acids so that all three binding centers of the LigBS are simultaneously operative in an unrestricted way. This binding mode will be hereinafter called embedded.

observ

et al..

to inter

extend

leaving

be her

might

Arg711

thus fo

larger

(Mikol

sugges

All th

(Kd=2

(Table

differ

intera

with

LoGr

anom

LigB

struc

An unexpected binding mode in KIV-10/Thr66 complexed with EACA was observed by the crystallography group at Sandoz Pharma AG, Basel Switzerland (Mikol et al., 1996). In this case, only the ϵ -amino group of the zwitterionic ligand was reported to interact with the anionic binding center of the kringle while the remainder of the ligand extended into the solvent region, almost perpendicular to the surface of the kringle, leaving the cationic site and the hydrophobic pocket unoccupied. This binding mode will be hereinafter called unbounded, and such a structure (Sandoz KIV-10/Thr66/EACA) might be expected with EACA bound to plasminogen K5 where the Arg35Ile and Arg71Leu substitutions in the LigBS remove the cationic site of the kringle and could thus force a ligand to solely interact with the anionic binding site with a correspondingly larger K_d (140 μ M) (Table 5-1). The unbounded binding mode of KIV-10/Thr66/EACA (Mikol et al., 1996), however, is not consistent with biochemical studies, which clearly suggest the embedded binding mode displayed by K1 and K4 (Mochalkin et al., 1999). All three binding centers are present in KIV-10/Thr66 and the K_d value of EACA ($K_d=20$ μ M) is similar to those of plasminogen K1 ($K_d=12$ μ M) and K4 ($K_d=26$ μ M) (Table 5-1). Furthermore, the K_d values of ligands in the unbounded mode should not differ significantly with the length of the ligand because only the terminal ϵ -amino group interacts with the anionic center. The measured dissociation constants of KIV-10/Thr66 with 5-APa ($K_d=66$ μ M), EACA ($K_d=20$ μ M) and 7-AHa ($K_d=230$ μ M) (Table 5-1; LoGrasso et al., 1994) suggest otherwise and the embedded binding mode. These anomalies and inconsistencies led us to investigate the binding mode of EACA in the LigBS of human KIV-10/Thr66 during the course of a more extensive study of the structures of the KIV-10/M66 variant in its free and ligand-bound states.

EACA

diffrac

succes

crysta

hangin

single

grow-

two l

sugg

migh

prep

10/V

singl

10 k

meas

Corp

RU20

mm).

consis

the c

II. EXPERIMENTAL

Crystallization. Relatively large crystals of the KIV-10 variants, with and without EACA, could be grown from ammonium sulfate-PEG 8000 solutions. Their X-ray diffraction patterns, however, showed a complex splitting pattern that could not be successfully interpreted and indexed. All indications suggested that the apparently single crystals were actually several small, similarly oriented fused crystals. Consequently, hanging drops of Hampton factorial solutions were examined. Excellent X-ray quality single crystals of KIV-10/Met66/EACA and KIV-10/Thr66/EACA could be readily grown from Hampton condition #40 (Table 6-2). However, crystallization of the other two kringles was not successful with this condition. Previous experience with kringles suggested PEG 8000 as precipitant in a buffer with pH around 7.0, with or without salt, might be suitable. An in-house set of factorial solutions based on the foregoing was prepared and the hanging drop method applied again to the KIV-10/Met66 and KIV-10/W72R samples. The conditions were eventually refined to produce X-ray quality single crystals. The full crystallization conditions and crystal dimensions of all the KIV-10 kringles are given in Table 6-2.

Intensity Data Collection. The X-ray diffraction data of the KIV-10 crystals were measured with an R-Axis II imaging plate detector using Molecular Structure Corporation-Yale focusing mirrors. The CuK α radiation was generated with a Rigaku RU200 rotating anode operating at 5kW power with a fine focus filament (0.3 mm x 3.0 mm). Intensity data were measured at -150° C with crystals flash frozen in a cryosolvent consisting of 20% PEG 4000/8000, 20% glycerol, and 0.1M buffer that was also used in the crystallization (Table 6-2). The crystal-detector distance was 10.0 cm

Table 6-2: Conditions used to crystallize the various KIV-10 kringles^a.

	KIV-10/Met66	KIV-10/Met66/EACA	KIV-10/Thr66/EACA	KIV-10/W72R
Protein concentration	20 mg/mL	20 mg/mL 10-fold molar excess of EACA	24 mg/mL 10-fold molar excess of EACA	20 mg/mL
Precipitant	10% Acetone 15% PEG 8000 0.1 M HEPES, 0.1 M Li ₂ SO ₄ pH 7.0	20% 2-propanol 20% PEG 4000 0.1 M Na Citrate, pH 5.6	20% 2-propanol 20% PEG 4000 0.1 M Na Citrate, pH 5.6	20% PEG 8000 0.1 M Na Cacodylate, 0.1 M Li ₂ SO ₄ pH 6.5
Crystal dimensions	0.18 × 0.15 × 0.02 mm	0.23 × 0.13 × 0.02 mm	0.50 × 0.10 × 0.02 mm	0.45 × 0.15 × 0.04 mm
Cryo condition	20% PEG 8000 20% Glycerol 0.1 M HEPES, pH 7.0	20% PEG 4000 20% Glycerol 0.1 M Na Citrate, pH 5.5	20% PEG 4000 20% Glycerol 0.1 M Na Citrate, pH 5.5	20% PEG 8000 20% Glycerol 0.1 M Na Cacodylate, pH 6.5

a - 2.0 µL hanging drop consists of 50 % protein - 50 % precipitant solution suspended over 0.75 mL of the latter.

and

eigh

Auto

AXI

Thre

shran

deter

plas

of R

Rota

in th

with

tran

bod

Ado

reli

(Br

the

cy

at

and the detector-swing angles were zero degrees for KIV-10/Met66 and KIV-10/W72R, eight degrees for KIV-10/Met66/EACA and ten degrees for KIV-10/Thr66/EACA. Autoindexing and processing of the diffraction data were carried out with the Rigaku R-Axis software package (Higashi, 1990) and the results are summarized in Table 6-3. Three of the crystals are isomorphous; the c-axis of the KIV-10/W72R mutant, however, shrank by about 7.0 Å.

Structure Determination. The crystal structure of KIV-10/Thr66/EACA was determined by molecular replacement using the coordinates of K4 of human plasminogen (Mulichak et al., 1991) (PDB code: 1PK4) as an initial model. All residues of K4 that differed in sequence from KIV-10/Thr66 were replaced with alanine. Rotation/translation searches were conducted with the program AMoRe (Navaza, 1994) in the resolution range (10.0 - 3.5) Å. The rotation search provided one distinct solution with a correlation coefficient of 0.30. The position of the molecule was fixed by a translation function that increased the correlation coefficient to 0.39 (R=48.5%). Rigid body optimization further increased the correlation coefficient to 0.64 (R=37.6%). Additional positional refinement by energy minimization in the (10.0 - 3.0) Å range to relieve close contacts of the model was carried out with the X-PLOR program package (Brunger, 1992) decreasing the R-value to 27.6 %.

The crystal structure of KIV-10/Thr66 complexed with EACA was refined with the programs PROLSQ (Hendrickson, 1985) and PROFFT (Finzel, 1987). First, three cycles of overall B-factor refinement starting with $B = 18.0 \text{ Å}^2$ converged to an R-value at 32.9% (7.0 - 2.8) Å. Then three cycles of the individual temperature factor refinement

Table 6-3: Crystal data and intensity data collection statistics of different KIV-10 crystals.

Table 6-3: Crystal data and intensity data collection statistics of different KIV-10 crystals.

	KIV-10/Met66	KIV-10/Met66/EACA	KIV-10/Thr66/EACA	KIV-10/W72R
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell constant (Å)				
a	24.61	24.37	24.30	24.30
b	45.68	45.60	45.67	45.82
c	63.70	62.90	63.36	56.23
Resolution (Å)	2.07	1.78	1.80	2.20
Independent reflections	3822	5717	5656	2241
Redundancy	2.5	2.6	2.5	2.9
Observations	9644	14679 ^a	12758	6494
R-merge (%)	6.2	5.7	6.0	8.9
outermost range (%)	15.3 (2.25-2.07) Å	9.3 (2.00-1.78) Å	10.1 (2.00-1.80) Å	8.9 (2.50-2.20) Å
Completeness (%)	81	79	81	63
outermost range (%)	64	53	53	45
I/σ outermost range	3.0	5.7	4.7	2.0

a - I/σ(I)>1.5; all others, I/σ(I)>1.0.

brought

geomet

differe

were a

calcula

kringl

exami

with t

using

varia

optim

the F

34.5

posi

thre

usin

10/7

sho

stru

me

3)

brought the R-value to 29.4%. Further refinement applying alternate tight-loose geometry restraints decreased R to 28.3%. At this stage, (2Fo-Fc) electron and (Fo-Fc) difference density maps indicated most of the correct side-chains of KIV-10/Thr66 that were alanine residues in the initial model and revealed EACA, which was not included in calculations. Further refinement was conducted with most of the actual residues of the kringle, as well as EACA. Solvent water molecules were added periodically based on examination of difference density maps at 2.5 Å and higher resolution.

Since the crystals of KIV-10/Met66 and KIV-10/Met66/EACA are isomorphous with that of KIV-10/Thr66, the structure determinations of the first two were conducted using isomorphous replacement methods. The refined coordinates of the KIV-10/Thr66 variant, with Thr66 replaced by alanine, were used as an initial model. Rigid-body optimization with the X-PLOR program package in the (10.0 - 3.0) Å range decreased the R-value from 38.6 % to 33.2 % for unliganded KIV-10/Met66 and from 43.7 % to 34.5 % for KIV-10/Met66/EACA (without the ligand in the calculations). Additional positional refinement improved the R-values to 25.1% and 28.1%, respectively. After three cycles of overall B-factor refinement and three cycles of individual B refinement using program PROFFT, R converged to 24.6% for KIV-10/Met66 and 26.4% for KIV-10/Met66/EACA. The (2Fo-Fc) electron and (Fo-Fc) difference density maps clearly showed Met66 in KIV/Met66 and the EACA in addition, in the KIV-10/Met66/EACA structure.

The crystal structure of KIV-10/W72R had to be determined *de novo* by the molecular replacement method (because of the dramatic shrinkage of the c-axis, Table 6-3) using the coordinates of KIV-10 from the KIV-10/Thr66/EACA complex. The

residu

model.

one dis

further

square

Table

varian

structu

indica

Rama

residu

plasm

1998)

48 ha

a hair

resid

resid

T

Arg⁷

three

residues Thr66 and Arg72 were replaced with alanine and EACA was omitted from the model. The rotation/translation search in the resolution range (10.0 - 3.0) Å provided one distinct solution with a correlation coefficient of 0.71 and a R-value of 32.6%, while further rigid-body refinement decreased the R-factor to 30.5%. Final restrained least-squares parameters and R-factor statistics for the KIV-10 complexes are presented in Table 6-4.

III. RESULTS

The electron density is well defined for most of the residues of the kringle variants except the interkringle peptides that are disordered here and in other kringle structures. The rmsΔ of the position of CA atoms between the different KIV-10 units indicates that the molecules exhibit very similar structural folding (Table 6-4). The Ramachandran plots of the KIV-10 variants reveal that all non-glycine and non-proline residues, except Met48, lie in the allowed regions. Similar to other kringle structures, plasminogen K1 (Mathews et al., 1996), K4 (Mulichak et al., 1991), K5 (Chang et al., 1998) and unbounded KIV-10/Thr66/EACA (Mikol et al., 1996;), the residue at position 48 has well-defined main and side chain electron density in all of them corresponding to a hairpin turn but is in a generously allowed area of the Ramachandran plot. No kringle residue was found in a disallowed region. As in other kringle structures, the Pro30 residue is in a *cis* conformation.

The LigBS of KIV-10 is defined by His33-Arg36, Asp55-Asp57, Trp62-Phe64, and Arg71-Tyr74 (Figure 6-1). Hydrogen bonds in the LigBS region can be divided into three principal groups according to several factors, which can depend on the ligand

Table 6-4: Summary of final restrained least-squares statistics of KIV-10 modules.

		KIV-10/ Met66	KIV-10/ Met66/ EACA	KIV-10/ Thr66/ EACA	KIV-10/ W72R
	Target	RMSΔ	RMSΔ	RMSΔ	RMSΔ
Distances (Å):					
Bond lengths	0.020	0.012	0.017	0.016	0.011
Bond angles	0.040	0.032	0.039	0.039	0.032
Planar 1-4	0.050	0.043	0.047	0.047	0.041
Planes (Å):					
Peptides	0.030	0.024	0.029	0.027	0.026
Aromatic groups	0.030	0.029	0.037	0.034	0.029
Chiral volumes (Å ³)	0.130	0.125	0.165	0.169	0.118
Non-bonded contacts (Å):					
Single torsion	0.60	0.18	0.18	0.20	0.22
Multiple torsion	0.60	0.26	0.21	0.28	0.29
Possible H-bond	0.60	0.32	0.24	0.32	0.28
Thermal parameters (Å ²):					
Main-chain bond	1.5	1.1	1.5	1.6	1.1
Main-chain angle	2.0	1.8	1.8	2.2	1.7
Side-chain bond	2.5	2.1	2.6	3.0	1.8
Side-chain angle	3.0	3.0	3.3	3.5	2.5
Range (Å)		9.0-2.1	9.0-1.8	9.0-1.8	9.0-2.2
R-value (%)		17.8	17.6	18.2	16.6
No. of water molecules ^a		67	77	109	41
 (Å ²)		21.8	15.6	15.9	21.5
CA RMSΔ (Å) ^b		0	0.30	0.30	0.43
PDB code		1KIV	2KIV	3KIV	4KIV
a - Occupancy > 0.5					
b- Calculated versus KIV-10/Met66					

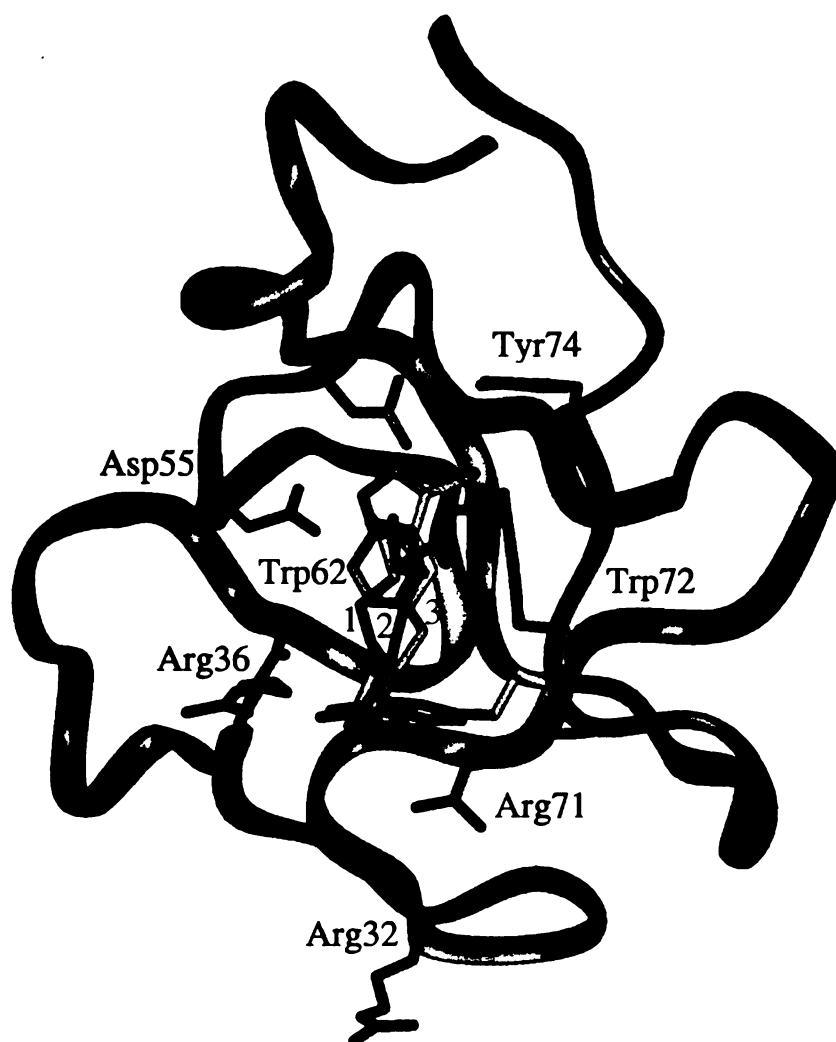


Figure 6-1: Superposition of EACA of different lysine binding kringles. KIV-10/Thr66/EACA side chains of LigBS are in gray. Embedded binding modes are labeled: EACA of K1 (1); EACA of K4 (2), EACA of KIV-10/Thr66/EACA (3). EACA of KIV-10/Thr66/EACA (Sandoz) displays the unbound binding mode. Note different conformations of Arg36 in KIV-10/Thr66/EACA and in KIV-10/Met66.

(EACA) or on intermolecular interactions of neighboring molecules in crystals. The first group includes hydrogen bonds that are common to all the structures (Table 6-5). These interactions in the LigBS mainly involve atoms of the backbone and are generally not completely exposed on surface. The second group consists of the hydrogen bonds that result from the binding of EACA ligand (Table 6-6). The residues of the anionic and cationic binding centers and several water molecules in the LigBS participate in the hydrogen bonding interactions of this group. Residue Arg32, which is located on the surface of the LigBS, gives rise to the third group of hydrogen bonds in the LigBS region that depend on crystal packing.

In the liganded structures of the Met66 and Thr66 variants of KIV-10, the electron density of the EACA in the LigBS is similar and well defined (Figure 6-2) and EACA interacts in the *embedded* way with all three binding centers in both kringles (Figures 6-1 & 6-2). The anionic (Asp55/Asp57) and cationic (Arg36/Arg71) centers stabilize the amino and carboxylic groups of EACA, respectively, while the hydrophobic pocket (Trp62, Phe64, Trp72, Tyr74) interacts with the methylene chain of the ligand. Thus, the KIV-10/Thr66/EACA structure determined here differs very significantly from the unbounded KIV-10/Thr66/EACA structure reported previously (Mikol et al., 1996), which only showed ligand binding at the anionic center (Figure 6-1). Except for Arg36, the binding of EACA to both the Met66 and Thr66 variants occurs with minimal structural reorganization so the LigBS appears to be essentially preformed anticipating ligand.

Because of the decrease of the lysine binding affinity of KIV-10/W72R compared to that of the Thr66 and Met66 variants (Klezovitch & Scanu, 1996), its LigBS is of

Table 6-5: Probable conserved hydrogen bonds (Å) in LigBS of various KIV-10.

		KIV-10/ Met66	KIV-10/ Met66/ EACA	KIV-10/ Thr66/ EACA	KIV-10/ W72R
Arg36-NH2	Pro54-O	3.00	a	a	3.20
Asn53-OD1	Asp55-OD1	2.59	2.85	3.02	3.76
-OD1	-N	2.75	2.82	2.44	2.51
-ND2	Asp57-O	2.77	2.90	3.02	3.17
-ND2	Asn5-O	2.50	2.75	2.63	2.54
-OD1	Trp62-NE1	2.51	2.74	2.70	3.01
Asp57-OD1	Tyr74-OH	2.88	2.41	2.87	2.44
-OD1	Thr58-N	2.64	2.68	2.59	2.46
-OD2	O _w 242/212/255 ^b	3.60	3.06	2.96	c
Pro61-O	Cys75-N	3.00	2.85	2.94	3.43 ^c
-O	O _w 228/264/258 ^b	2.72	2.74	3.23	c
Trp62-N	Arg52-O	2.77	2.81	2.95	2.66
-O	-N	3.50	3.11	3.06	3.45
Cys63-N	Glu73-O	3.04	2.98	3.01	3.49 ^c
-O	-N	3.03	2.99	3.15	3.37 ^c
Phe64-N	O _w 201/202/208/219 ^b	2.73	2.88	2.74	2.58
-O	Gln23-N	2.61	2.64	2.69	3.38 ^c
Arg71-O	Thr65-N	3.10	3.10	2.96	2.58
-NH1	Arg32-O	2.65	2.75	2.86	3.37 ^c
Trp72-O	O _w 222/249/239 ^b	3.48	3.05	2.96	c
Tyr74-N	O _w 219/237/306 ^b	2.72	2.96	2.88	c
-O	Val17-N	2.88	3.00	3.06	3.68 ^c

a - Arg36 undergoes a conformational change.

b - Numbering of water molecules in respective structures.

c - Due to the shrinkage of the c-axis of the mutant, the solvent content is reduced to 22% (total 41 water molecules, Table 6-3) resulting in the loss of hydration interactions and possibly some hydrogen bonds.

Table 6-6: Probable hydrogen bonds (Å) in LigBS of KIV dependent on ligand binding and the Trp72Arg mutation.

Kringle	Ligand	KIV-10/ Met66/EACA	KIV-10/ Thr66/EACA	KIV-10/ W72R
Arg36-NE	EACA-OZ	2.98	2.72	-
-NH1	O _w 242/291	3.01	3.12	-
-NH2	EACA-OZ	3.05	2.87	Pro54-O (3.20)
-NH2	Arg32-NH1 ^a	-	2.67	3.22
-NH2	Arg32-NH2 ^a	-	-	3.25
Pro54-O	O _w 261/244	2.97	2.98	Arg36-NH2 (3.20)
Asp55-OD2	EACA-NZ	2.62	2.75	Arg32-NE (3.16) ^a
Asp57-OD2	EACA-NZ	2.53	2.87	Arg72-NH2 (3.01)
-NE	EACA-O	2.65	3.13	-
Arg71-NH1	EACA-O	3.00	2.75	-
-NH2		-	-	Thr29-OG1 (3.20) ^a

a - Symmetry related molecule.

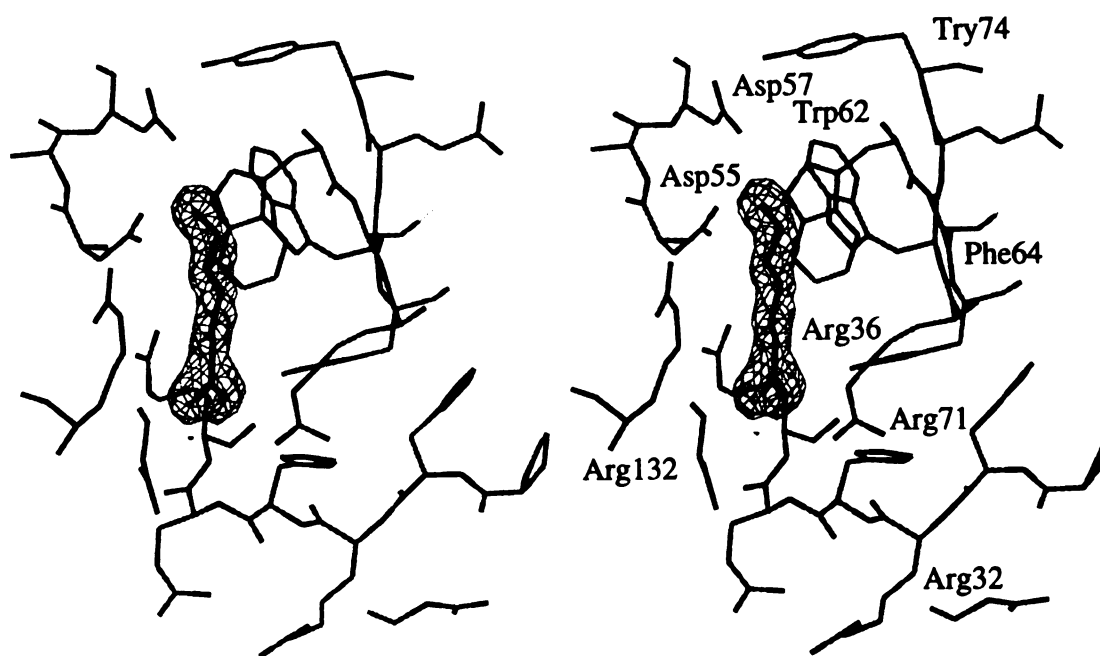


Figure 6-2: Stereoview of final (2Fo-Fc) electron density of EACA of KIV-10/Thr66/EACA displaying the embedded binding mode. Contoured at the 1σ level. Two conformations of Arg32 are shown. Arg32 of symmetry related molecule (Arg132) is in red.

special interest. Substituting a non-aromatic amino acid for Trp72, an obviously important residue of the hydrophobic pocket along with Trp62 (Figure 6-3), would in itself be expected to lead to some impairment of binding affinity. The structure of KIV-10/W72R additionally shows that Arg72 extends across the LigBS toward the negatively charged Asp55/Asp57 pair making a hydrogen bonded salt bridge with Asp57 (Table 6-6) and thusly obstructs access to the LigBS. Since the cationic center consisting of Arg36 and Arg71 does not undergo any structural changes and is positionally similar to that of the unliganded KIV-10 molecule (Figure 6-1 & 6-3), the arginyl of Arg72 occupying the binding groove of the LigBS appears to be the structural deterrent responsible for the lack of binding affinity of the KIV-10 mutant.

All four KIV-10 variants crystallized in the orthorhombic space group $P2_12_12_1$. Three of the crystals are isomorphous, while the c-axis of the fourth, KIV-10/W72R, is about 7.0 Å shorter (Table 6-3). This 10% shrinkage is the result of different molecular packing of the KIV-10/W72R molecules along the c-axis of the crystal. Superposition of the mutant with any of the three isomorphous variants showed very little rotation of the molecules with respect of each other, with only a translational difference along the z-direction. In all four crystal arrangements, kringle molecules form infinite chain columns parallel to the y-direction with a 15 Å wide channel between chains. In KIV-10/W72R, two neighboring symmetry related molecules are aligned approximately 3.5 Å closer along the crystallographic c-axis, while the packing along the two other axes remains about the same. The net result is to shrink the channel by about 7 Å at the expense of solvent water molecules. The solvent fraction of the three isomorphous KIV-10 structures is approximately 30.6 % ($V_M = 1.76 \text{ Å}^3/\text{Da}$). In the case of the W72R mutant,

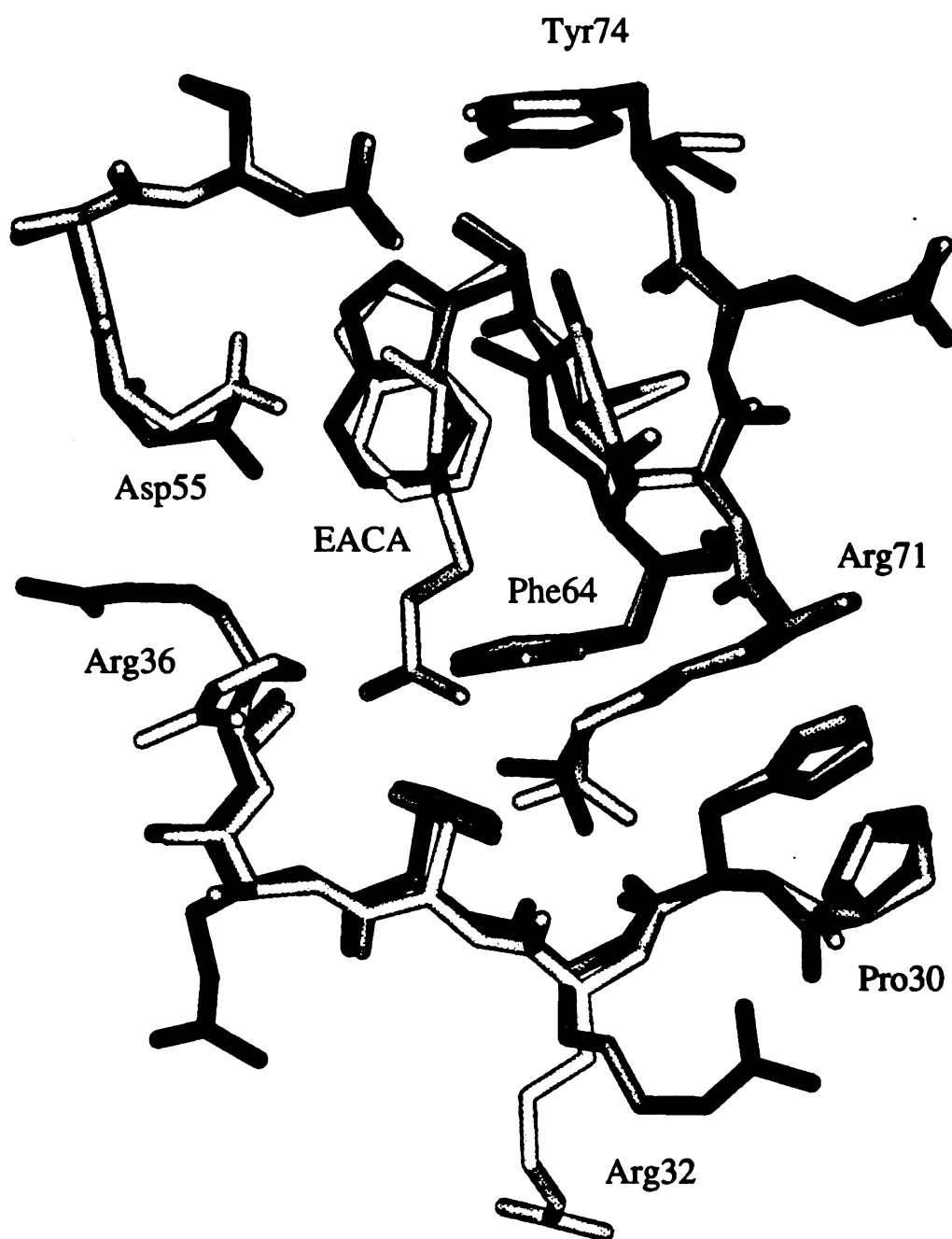


Figure 6-3: Superposition of KIV-10/W72R (dark gray) and KIV-10/Met66/EACA (light gray) showing Arg72 approximating EACA binding. Note different conformation of Arg36 in the two structures.

the shrinkage along the c-axis decreases the solvent content to 22.0 % ($V_M = 1.56 \text{ \AA}^3/\text{Da}$). The number of water molecules found in the four different kringle structures varies from 41 to 109 and is generally related to the resolution of the diffraction data but in the case of KIV-10/W72R, it depends primarily on the solvent fraction of the crystals.

IV. DISCUSSION

The overall topology of human KIV-10 is an oblate ellipsoid of approximately 28 Å in diameter and 18 Å thick (Figure 6-1). The LigBS of KIV-10 is similar to those of plasminogen K1 and K4 (Mulichak et al., 1991; Wu et al., 1991; Mathews et al., 1996). All three key segments, the anionic, cationic and hydrophobic centers, are essentially positioned optimally for ligand binding in KIV-10/Met66 and except for Arg36 (Figure 6-1), remain so on ligand binding in KIV-10/Met66/EACA and KIV-10/Thr66/EACA. In KIV-10/Met66, residue Asp55 is in the same orientation ($\chi_1 \approx 60^\circ$) as in plasminogen K1 and K4.

All three binding centers are operative in KIV-10/Met66/EACA and KIV-10/Thr66/EACA examined in the present work and EACA binds to the LigBS in the embedded way (Figures 6-1 & 6-2). This binding mode, also observed in plasminogen K1 and K4 (Figure 6-1), is consistent with binding constants of different ligands discussed earlier. Thus, the binding of embedded KIV-10/Thr66/EACA is very different from the unbounded mode previously reported for the KIV-10/Thr66/EACA structure (Mikol et al., 1996). It is noteworthy, therefore, that lysine binding kringles can bind ligand in two different ways.

The source of the binding mode variance between our work and that of Mikol et al., (Mikol et al., 1996) most likely resides in the different conditions used for crystallization. The major differences are: unbound KIV-10/Thr66/EACA, protein concentration 5 mg/mL, pH 7.5, in the presence of 40 mM NaCl; present work KIV-10/Thr66/EACA, protein concentration 12 mg/mL, pH 5.6, no NaCl but with 10% 2-propanol and a 10-fold molar excess of EACA (latter unreported in the Sandoz crystallization). Although the different pH conditions might seem important, liganded K1 and K4 crystals in the embedded mode were grown at higher pH of 6.5 and 6.0, respectively. It is of some note that in the LigBS of unliganded plasminogen K1 crystals, there are two chloride ions in the vicinity of the cationic center, which appear to interact with and neutralize Arg35 and Arg71 electrostatically (Wu et al., 1994). In the case of unliganded plasminogen K4, the anion is a sulfate and the interaction is more complicated by the additional presence of a symmetry related molecule (Mulichak et al., 1991). Thus, it is conceivable that the unbound ligand binding mode of KIV-10/Thr66/EACA (Mikol et al., 1996) could result because of an anion compensated cationic center. However, examination of the B-values of the water molecules in the LigBS of the latter reveals only evidence for water molecules of solvation (no inordinately small B-values suggesting anions).

The unique features of the LigBS of the apo(a) KIV-10 of rhesus and chimpanzee are the point mutations that lead to an impairment EACA binding and binding to lysine-sepharose (Scanu et al., 1993; Klezovitch & Scanu, 1996; Chenivasse et al., 1998). The structure of the recombinant mutant KIV-10/W72R shows that Arg72 extends along the ligand binding groove parallel to the expected position of EACA, toward the negatively charged Asp55/Asp57 anionic center and makes a hydrogen bonded salt bridge with

Asp57 (Figure 6-3, Table 6-6). Thus, the Arg72 side chain mimics ligand binding and loss of binding ability is the result of a steric blockage of the LigBS by Arg72 physically occupying part of the site. The LigBS structure of KIV-10/W72R is similar to the LigBS region of PT-K1, which has an Arg72 equivalent but lacks the Asp57 to make a salt bridge. In this case, Arg72 appears to form a salt bridge with Asp55. This structure was used previously to closely model the LigBS of rhesus KIV-10/W72R (Scanu et al., 1993). In turn, the Asp57Asn point mutation in the LigBS of KIV-10/Asn57 of chimpanzee simply lowers the negative electrostatic charge of the anionic center that appears to be so critical for the binding of the quaternary, positively charged ϵ -amino group of lysine and other zwitterionic ligands. This only leads to an electrostatic impairment of an otherwise capable LigBS. Since the Asn57 side chain is geometrically nearly isostructural with the Asp57, which it replaces and can be a hydrogen bond donor or acceptor, it would seem that the LigBS function of KIV-10 of chimpanzee should be less impaired compared to that of the rhesus monkey, primarily because the steric impairment of the latter excluded ligand physically. This suggested difference in LigBS function between the KIV-10 of rhesus and chimpanzee could not be assessed further because the lysine binding conditions were not provided for the chimpanzee product.

Superposition of the KIV-10/Met66 and the ligand bound KIV-10/Met66/EACA and KIV-10/Thr66/EACA structures shows that the binding of EACA in KIV-10 is accompanied by: (1) a displacement of water molecules from the EACA binding groove and (2) a movement of the arginyl residue of Arg36 (Figure 6-1). Three water molecules (O_w258 , O_w259 , O_w260) extending from the anionic to the cationic center in KIV-10/Met66 are displaced by EACA with ligand binding. Two other hydrogen bonding

water sites, located in the LigBS region but not in the EACA binding groove, are conserved in both structural types (Table 6-6). Similar to binding of EACA, the Trp72Arg mutation causes the displacement of water molecules from the binding groove. Additional examination of the residues within 5.0 Å of the LigBS indicates only a conformational change in residue Arg36 on binding EACA. In the ligand-free structures of KIV-10/Met66 (Figure 6-1) and KIV-10/W72R (Figure 6-3), residue Arg36 adopts an extended conformation, almost parallel to the LigBS groove, and utilizes a hydrogen bond with Pro54-O (Table 6-5). Upon ligand binding, Arg36 of the Met66 and Thr66 KIV-10 variants swings around the CA-CB bond by about -120° toward the cationic center, but remains extended, bringing the guanidinium group close to the carboxylate of EACA to assist Arg71 in stabilizing the anionic end of the ligand as in plasminogen K4 (Figures 6-1, 6-2 & 6-3, Table 6-6). The movement is accompanied by an approximately 1.3 Å shift of the tyrosyl ring of the sidechain of Tyr41 into the vacancy created by the Arg36 conformational change. When Arg36 is in the ligand binding posture, two water molecules (O_w261, O_w244, Table 6-6) also occupy the vacancy created by the shift. The KIV-10/Thr66/EACA also shows the arginyl reorientation but the guanidinium amino groups correspond to weaker electron density so their orientation is not as certain as that of the Met66 variant.

Superposition of the KIV-10 variants also revealed different conformations for Arg32. This residue is located on the surface of the kringle and is adjacent to the LigBS of a symmetry-related molecule and appears to play a role in molecular packing. The conformation of Arg32 is similar in the KIV-10/Met66 and KIV-10/Met66/EACA structures. In addition to a favorable hydrogen bonding interaction between its carbonyl

oxygen atom and Arg71-NH1 in the free and ligand bound structures (Table 6-5), the terminal guanidino nitrogen atoms of Arg32 in KIV-10/Met66/EACA participate in a complicated electrostatic cluster comprised of the carboxyl group of EACA of the symmetry related molecule and Arg36/Asp55 of the corresponding cationic and anionic binding centers. Apparent charge neutralization is achieved by the cluster through an pseudo cyclic arrangement of positive and negative centers (Arg32 of a neighboring kringle, Asp55, quaternary amine of EACA, the carboxylate of EACA and Arg36 (Figure 6-2). In the KIV-10/Thr66/EACA structure Arg32 has two alternate conformations, which have been refined to an occupancy of 0.6 (extended) and 0.4 (bent). The Arg32 in the extended conformation interacts with a symmetry related kringle unit (Figure 6-2), whereas in the bent conformation it forms a hydrogen bond with *cis* Pro30-O (3.00 Å) of the same molecule. An additional water molecule participates in the stabilization of both the extended and the bent conformation. In the KIV-10/W72R mutant, Arg32 is in an intermediate bent conformation displaying both intermolecular and intramolecular interactions while bridging two kringle molecules. The NH1 and NH2 atoms interact with *cis* Pro30-O as in KIV-10/Met66 and KIV-10/Met66/EACA but in addition, there are intermolecular contacts between Arg32-NE and Asp55-OD2 (3.16 Å) as well as Arg32-NH2 and Arg36-NH2 (3.25 Å) of the LigBS of a symmetry related molecule.

The current crystallographic studies of a wild type (Trp72) and mutant (Arg72) of human apo(a) KIV-10 provide a detailed account of the way apo(a) KIV-10 binds to lysine and its analogs. This information is of biological relevance in that lysine binding and related activities of Lp(a) depend on the function of LigBS of KIV-10 (Scanu et al., 1993; Scanu et al., 1994). The kringle, according to current dogma, is considered to be

exposed because it has no apparent formal linkages with the lipoprotein component of Lp(a). In keeping with this notion is the lysine binding deficiency of rhesus monkey and the human mutant Lp(a) having the critical Trp72 replaced by Arg in the LigBS of KIV-10. From the standpoint of the athero-thrombogenic potential of Lp(a), this type of mutation has been suggested to be benign on the premise that it would be unable to interfere with the plasminogen to plasmin conversion (Scanu et al., 1993; Scanu et al., 1994). This view has recently received support from studies in transgenic mice (Hughes et al., 1997). Because of the pathobiological significance of the LigBS and its potential for mutability, it would be of further interest to study additional natural and/or artificial mutations in order to define their impact on the lysine binding activity of apo(a) KIV-10. In the current study we have shown that the replacement of Met by Thr in position 66 had no significant structural consequences in agreement with the finding that this common human mutation is functionally silent. The example provided by chimpanzee Lp(a) of a substitution at Asp55 by Asn is of interest in that we predict that such a substitution would lead to only a relatively modest degree of a functional impairment of lysine binding function compared to that of the rhesus. A functional hierarchy in the LigBS might account for the variability in lysine binding activity among human subjects based on measurements in the whole plasma (Scanu et al., 1994; Hoover-Plow et al., 1996). In this context, however, we cannot rule out the contribution to the binding by a second LigBS contained in the region between apo(a) KIV-5 and KIV-8 (Edelstein et al., 1995). This region is usually open in free apo(a) whereas it is essentially masked when apo(a) is a constituent of the Lp(a) particle (Edelstein et al., 1995; Ernst et al., 1995). A structural understanding of the region may be of some consequence in view of its suggested

inv

199

involvement in the lysine-mediated binding of apo(a) to fibrinogen (Klezovitch et al., 1996).

GLOSSARY

Angiogenesis	formation of new blood vessels that occurs as a result of the growth of capillaries by vascular sprouting from pre-existing vessels
Arteries	large blood vessels that lead blood away from the heart.
Atherosclerosis	deposition of fatty compounds on the inner lining of the coronary arteries or any other arteries. The originally smooth lining of the artery becomes roughened as the atherosclerotic plaque collects in the artery.
Blood Coagulation	series of stepwise reactions that utilize the protein
Cascade	molecules circulating in the blood in order to seal an area of endothelial damage and to stop blood loss.
Blood Clotting	see Blood Coagulation Cascade.
Cerebrovascular Disease	disruption of the normal blood supply to the brain
Coagulation Cascade	see Blood Coagulation Cascade
Denaturation	Structure alternations in proteins caused by heat, extreme pH or by exposure to organic solvents, detergents or certain solutes.
Differentiation	change in structure and function of a cell as it matures
Dyne	force necessary to give acceleration of one centimeter per second per second to one gram of mass.

En

Ep

Fi

He

He

Hy

In

K

M

M

Endothelium	type of epithelium composed of a single layer of smooth thin cells that lines the heart, blood vessels, lymphatics and serous cavities.
Epithelium	any tissue which covers a surface or lines a cavity and which performs protective, secreting or other functions.
Fibrinolysis	phenomenon of clot dissolution.
Hemorrhage	discharge of blood from a ruptured blood vessel.
Hemostasis	stoppage of bleeding or the stoppage of the circulation of blood in a part of the body.
Hypertension	high blood pressure.
Infarction	area of dead myocardial tissue. The severity of a myocardial infarction (also known as a heart attack) depends on the size of the artery that is blocked and the extend of the blockage.
Kringle domain	structural and functional modular unit of approximately 80 amino acids cross-linked by three disulfide bridges in the characteristic pattern. Kringle motives are found in various blood proteins.
Megakaryocyte	platelet precursor formed in the bone marrow.
Metastasis	ability of tumor cells to leave their site of origin and migrate to other locations in the body, where a new colony is established.
Occlusion	closure of a blood vessel.

Platelet count	number of platelets per cubic millimeter. Platelets normally average between 140,000 and 400,000/mm ³ .
Primary structure	the order of amino acid residues in polypeptide chain.
Quaternary structure	the arrangement and nature of the binding of the subunits and domains together.
Secondary structure	the conformation of a polypeptide chain and the intra-chain hydrogen binding scheme.
Stem cell	cell in bone marrow that gives rise to different types of blood cells.
Tertiary structure	the folding of the polypeptide which results in a three-dimensional structure.
Thrombocythemia	increased platelet counts.
Thrombocytopathy	abnormal platelet functions in adhesion and aggregation.
Thrombocytopenia	decreased platelet counts is the most common platelet disorder of insufficient platelet production and/or excessive platelet removal from circulation.
Thromboplastin	membrane glycoprotein and a clotting factor that is located in the tissue adventitia and comes into contact with blood after a vascular injury. In combination with FVII/FVIIa and calcium, it initiates the blood coagulation cascade.
Thrombopoiesis	production of platelets
Thrombosis	condition of clot formation

Names, three- and one-letter codes and molecular weights of the common amino acids

Amino Acid	Three- and one-letter code		Mr.	Amino Acid	Three- and one-letter code		Mr.
Glycine	Gly	G	75	Alanine	Ala	A	89
Valine	Val	V	117	Leucine	Leu	L	131
Isoleucine	Ile	I	131	Phenylalanine	Phe	F	165
Tyrosine	Tyr	Y	181	Tryptophan	Trp	W	204
Serine	Ser	S	105	Threonine	Thr	T	119
Cystein	Cys	C	121	Methionine	Met	M	149
Asparagine	Asn	N	132	Glutamine	Gln	Q	146
Aspartic acid	Asp	D	133	Glutamic acid	Glu	E	147
Lysine	Lys	K	146	Arginine	Arg	R	174
Histodine	His	H	155	Proline	Pro	P	115

BIBLIOGRAPHY

- Appella, E., Weber, I. T. & Blasi, F. (1988) *FEBS Lett.* 231, 1-4.
- Arni, R. K., Padmanabhan, K., Padmanabhan, K. P., Wu, T. P. & Tulinsky, A. (1993) *Biochemistry* 32, 4727-4737.
- Bachmann, F. (1994). Hemostasis and Thrombosis, 3rd Ed., edited by Colman, R. W., Hirsch, J., Marder, V. J., Salzman, E. W., pp. 1592-1622, J.B. Lippincott Co., Philadelphia.
- Balasubramanian, B. N. Ed. (1995). *Biorg. & Med. Chem.* 3, 499-1156.
- Banner, D. W., D'Arcy, A., Chène, C., Winkler, F. K., Guna, A., Konigsberg, W. H., Nemerson, Y. & Kirchofer, D. (1996) *Nature* 380, 41-46.
- Barker, W. C., Johnson, G. C., Hunt, L. T. & George, D. G. (1986) *Protein Nucleic Acid Enz.* 29, 54-68.
- Blomback, B. (1967). Blood Clotting Enzymology, edited by Seegers, W. H., pp143-215, New York: Academic Press.
- Blomquist, M. C., Hunt, L. T. & Barker, W. C., (1984) *Proc. Natl. Acad. Sci U.S.A.* 81, 7363-7367.
- Boatman, P. D., Ogbu, C. O., Eguchi, M., Kim, H-O., Nakanishi, H., Cao, B. & Kahn, M. (1999) *J. Med. Chem.* 42,1367-1375.
- Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. & Hofsteenge, J. (1989) *EMBO J.* 8, 3467-3475.
- Bode, W., Turk, D. & Karshikov, A. (1992) *Protein. Sci.* 1, 426-471.
- Born G. V. R. & Hardisty, R. M. (1972) Human Blood Coagulation, Haemostasis and Thrombosis, edited by Biggs, R., pp. 159-187, Alden & Mowbray, Ltd., Oxford.

- Brandstetter, H., Turk, D., Hoeffken, H. W., Grosse, D., Sturzebecher, J., Martin, P.D.,
Edwards, B. F. P. & Bode, W. (1992) *J. Mol. Biol.* 226, 1085-1099.
- Brandstetter, H., Kuhne, A., Bode, W., Huber, R., van der Saal, W., Wirthensohn, K. &
Engh, R. A. (1996) *J. Biol. Chem.* 271, 29988-29992.
- Brunger, A. T. (1992) X-PLOR Version 3.1: A System for Crystallography and NMR.
Yale University Press, New Haven.
- Brunner, C., Kraft H-G., Utermann, G. & Müller H-J. (1993) *Proc. Natl. Acad. Sci. USA*,
90, 11643-11647
- Cao, Y., Ji, R. W., Davidson, D., Schaller, J., Marti, D., Söndel, S., McCance, S. G.,
O'Reilly, M. S., Llinas, M., Folkman, J. (1996) *J. Biol. Chem.* 271, 29461-29467.
- Chang, T-L., Bauer, R. S. & Berlinger, L. J. (1980) *J. Biol. Chem.* 255, 5904-5906.
- Chang, Y., Mochalkin, I., McCance, S. G., Cheng, B., Tulinsky, A. & Castellino, F. J.
(1998) *Biochemistry* 37, 3258-3271.
- Chenivasse, X., Huby, T., Wickins, J., Chapman, J. & Thillet, J. (1998) *Biochemistry*
37, 7213-7223.
- Colman, R. W., Marder, V. J., Salzman, E. W. & Hirsch, J. (1994) Hemostasis and
Thrombosis, 3rd ed., edited by Colman, R.W., Hirsch, J., Marder, V. J. & Salzman,
E. W. pp. 3-18, J. B. Lippincott Co., Philadelphia.
- Dahlen, G. H., Ericson, C. & Berg, K. (1978) *Clin. Genet.* 14. 36-42
- Dahlen, G.H., Guyton, J. R., Attar, M., Farmer, J. A., Kautz, J. A. & Gotto, A. M. (1986)
Circulation 74, 758-765.
- Dang, Q. D., Vindigni, A. & Di Cera, E. (1995). *Proc. Natl. Acad. Sci USA* 92, 5977-
5981.

- Davie, E. W. & Ratnoff, O. D. (1964) *Science* 145, 1310-1312.
- Davie, E. W., Fujikawa, K. & W. Kisiel (1991) *Biochemistry* 30, 10363-10370.
- Davis, C., G. (1990) *New Biol.* 2, 410-419.
- de Vos, A. M., Ultsch, M. H., Kelley, R. F., Padmanbhan, K., Tulinsky, A., Westbrook, M. L. & Kossiakoff, A. A. (1991) *Biochemistry* 31, 270-279.
- Deerfield D., Olson, D., Berkowitz, P., Byrd, P., Koehler, K. A., Pedersen, L. & Hiskey, R. (1987) *J. Biol. Chem.* 262, 4017-4023.
- DiCera, E., Guinto, E. R., Vindigni, A., Dang, Q. O. Alaya, Y. M., Wuyi, M. & Tulinsky, A. (1995) *J. Biol. Chem.* 270, 22089-22092.
- Doolittle, R. F., Feng, D. F. & Johnson, M. S. (1984) *Nature* 307, 558-560.
- Dougherty, D. A. & Stauffer, D. A. (1990) *Science* 250, 1558-1560.
- Eaton, D. L., Fless, G. M., Kohr, W. J., McLean, J. W., Xu, Q., Miller, C. G., Lawn, R. M. & Scanu A. M. (1987) *Proc. Natl. Acad. Sci. USA* 84, 3224-3228.
- Edelstein, C., Mandala, M., Pfaffinger, D. & Scanu, A. M. (1995) *Biochemistry* 34, 16483-16492.
- Eguchi, M., Boatman, P. D., Ogbu, C. O., Nakanishi, H., Bolong, C., Lee, M. S. & Kahn, M. (1996) 212th American Chemical Society National Meeting, Orlando, FL; MEDI125.
- Ernst, A., Helmhold, M., Brunner, C., Pethoschramm, A., Armstrong, V. W. & Muller, H. J. (1995) *J. Biol. Chem.* 270, 6227-6234.
- Esmon, C. T. (1993) *Thrombo. & Haemost.* 70, 29-35.
- Feng, D. M., Gardell, S. J., Lewis, S. D., Bock, M. G., Chen, Z., Freidinger, R. M.,

- Naylor-Olsen, A. M., Ramjit, H. G., Woltmann, R., Baskin, E. P., Lynch, J. J., Lucas, R., Schafer, J. A., Dancheck, K. B., Chen, I-Wu, Mao, S. S., Krueger, J. A., Hare, T. R., Mulichak, A. M. & Vacca, J. P. (1997) *J. Med. Chem.* 40, 3726-3733.
- Finzel, B. C. (1987) *J. Appl. Crystallogr.* 20, 53-55.
- Folk, J. E. & Finlayson, J. S. (1977) *Adv. Protein Chem.* 31, 1-133.
- Fraternali, F., Do, Q.-T., Doan, B.-T., Atkinson, R. A., Palmas, P., Skelnar, V., Safar, P., Wildgoose, P., Strop, P. & Saudek, V. (1998) *Proteins* 30, 264-274.
- Gan, Z. -R., Li, Y., Chen, Z., Lewis, S. D. & Shafer, J. A. (1994) *J. Biol. Chem.* 269, 1301-1305.
- Harker, L. A. (1968) *J. Clin. Invest.* 47, 452-458.
- Harris, J. I. (1956) *Nature* 177, 471-472.
- Hendrickson, W. A. (1985) *Methods Enzymol.* 115, 252-270.
- Hibbard, L. S. & Tulinsky, A. (1978) *Biochemistry* 17, 5460-5468.
- Higashi, T. (1990) *J. Appl. Crystallogr.* 23, 252-257.
- Hoover-Plow, J. L., Boonmark, N., Skocir, P., Lawn, R. M. & Plow, E. F. (1996) *Arterioscler. Thromb. Vasc. Biol.* 16, 656-664.
- Hughes, S. D., Lou, X. J., Ighani, S., Verstuyft, J., Grainger, D. J., Lawn, R. M. & Rubin, E. M. (1997) *J. Clin. Invest.* 100, 1493-1500.
- Kahn, M., Wilke, S., Chen, B. & Fujita, K. (1988) *J. Am. Chem. Soc.* 110, 1638-1639.
- Karshikov, A., Bode, W., Tulinsky, A. & Stone, S. R. (1992) *Prot. Sci.* 1, 727-735.
- Kim, H.-O. & Kahn, M. (1997) *Tetrahedron Lett.* 38, 6483-6484.
- Klezovitch, O. & Scanu, A. M. (1996) *Arter. Thromb. Vasc. Biol.* 16, 392-398.

- Klezovitch, O., Edelstein, C. & Scanu, A. M. (1996) *J. Clin. Invest.* 98, 185-191.
- Kornfeld, S. (1992) *Annu. Rev. Biochem.* 61, 307-330.
- Koschinsky, M. L., Côté, G. P., Gabel, B. & van der Hoek, Y. Y. (1993) *J. Biol. Chem.* 268, 19819-19825.
- Kostner, G. M., Avogaro, P., Gazzolato G., Marth, E., Bittolo-Bon, G. & Qunici, G. B. (1981) *Atherosclerosis* 38, 51-61.
- Kraft, H. G., Haibach, C., Lingenhel, A., Brunner, C., Trommsdorff, M., Kronenberg, F., Muller, H. J. & Utermann, G. (1995) *Hum. Genet.* 95, 275-282.
- Krishnan, R., Zhang, E., Hakansson, K., Arni, R. K., Tulinsky, A., Lim-Wilby, M. S. L., Levy, O. E., Semple, J. E. & Brunck, T. K. (1998) *Biochemistry* 37, 12094-120103.
- Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thorton, J. (1993) *J. Appl. Crystallogr.* 26, 283-286.
- Lijnen, H. R., Ugwu, F., Bini, A. & Collen, D. (1998), *Biochemistry* 37, 4699-4702.
- Lin, Z. & Johnson, M. (1995) *FEBS Letters* 370, 1-5
- Locht, A., Bode, W., Huber, R., Bonniec, B. F., Stone, S. R., Esmon, C. & Stubbs, M. T. (1997) *EMBO J.* 16, 2977-2984.
- LoGrasso, P. V., Cornell-Kennon, S. & Boettcher, B. R. (1994) *J. Biol. Chem.* 269, 34, 21820-21827.
- Lorand, L. & Konishi, K. (1964) *Arch. Biochem. Biophys.* 105, 58-67.
- MacFarlane, R. G. (1964) *Nature* 202, 498-499.
- MacFarlane R. G. (1972) *Human Blood Coagulation, Haemostasis and Thrombosis*, edited by Biggs, R., pp. 159-187, Alden & Mowbray, Ltd., Oxford.

- Mangel, W. F., Lin, B., & Ramakrishnan, V. (1990) *Science* 248, 69-73.
- Mann, K. G., Nesheim, M.E., Church, W. R., Haley, P. & Krishnaswamy, S. (1990) *Blood* 76, 1-16.
- Martin, C. J. & Frazier, A. L. (1963) *J. Biol. Chem.* 238, 3268-3273.
- Mathews, I. I., Padmanabhan, K. P., Ganesh, V., Tulinsky, A., Ischii, M., Chen, J., Turck, C. W., Coughlin, S. R. & Fenton, J. W. (1994) *Biochemistry* 33, 3266-3276.
- Mathews, I. I. & Tulinsky, A. (1995) *Acta Crystallogr.* D51, 550-559.
- Mathews, I. I., Vanderhoff-Hanaver, P., Castellino, F. J. & Tulinsky, A. (1996) *Biochemistry* 35, 2567-2576.
- Mathur, A., Schlapkohl, W. & Di Cera, E. (1993) *Biochemistry* 32, 7568-7573.
- McCance, S. G., Menhart, N. & Castellino, F. J. (1994) *J. Biol. Chem.* 269, 32405-32410.
- McLean, J. W., Tomlinson, J. E., Kuang, W-J., Eaton, D. L., Chen, E. Y., Fless, G. M., Scanu, A. M. & Lawn, R. M. (1987) *Nature* 330, 132-137.
- McMullen, B. A., Fujikawa, K. & Davie, E. W. (1991^a) *Biochemistry* 30, 2050-2056.
- McMullen, B. A., Fujikawa, K. & Davie, E. W. (1991^b) *Biochemistry* 30, 2056-2060.
- Menhart, N., Sehl, L. C., Kelley, R. F. & Castellino, F. J. (1991) *Biochemistry* 30, 1948-1957.
- Mikol, V., LoGrasso, P. V. & Boettcher, B. R. (1996) *J. Mol. Biol.* 256, 751-761.
- Miles, L. A. & Plow, E. F. (1990) *Thromb. & Haemost.* 63, 331-335.
- Miyazawa, K., Shimomura, T., Kitamura, A., Kondo, J., Morimoto, Y. & Kitamura, N. (1993) *J. Biol. Chem.* 268, 10024-10028.
- Mochalkin, I. & Tulinsky, A. (1999) *Acta Crystallogr.* D55, 785-793.
- Mochalkin, I., Cheng, B., Klezovitch, O., Scanu, A. & Tulinsky, A. (1999) *Biochemistry*

38, 1990-1999.

Mulichak, A. M., Tulinsky, A. & Ravichandran, K. G. (1991) *Biochemistry* 30, 10576-10588.

Murai, A. T., Miahara, T., Fujimoto, N., Matsuda, M. & Kameyama, M. (1986) *Atherosclerosis* 59, 199-204.

Nakamura, T., Nishizawa, T., Hagiya, M., Seki, T., Shimonishi, M., Sugimura, A., Tashiro, K. & Shimzu, S. (1989) *Nature* 342, 440-443.

Nakanishi, H., Chrusciel, R. A., Shen, R., Bertenshaw, S., Johnson, M. E., Rydel, T. J., Tulinsky, A. & Kahn, M. (1992) *Proc. Natl. Acad. Sci. USA* 89, 1705-1709.

Naski, M. C., Lorand, L. & Shafer, J. A. (1991) *Biochemistry* 30, 934-941.

Navaza, J. (1994) *Acta Crystallogr.* A50, 157-163.

Ni, F., Ripoll, D. R., Martin, P. D. & Edwards, B. F. P. (1992) *Biochemistry* 31, 11551-11557.

Nienaber, V. L. & Amparo, E. C. (1996) *J. Am. Chem. Soc.* 118, 6807-6810.

O'Connor, B. H. (1984) A Color Atlas and Instruction Manual of Peripheral Blood Cell Morphology, pp. 107-108, Williams&Wilkins, Baltimore.

Olson, S. T. & Bjork, I. (1992) Thrombin. Structure and Function, edited by Berliner, L.J., pp.159-217. Plenum Press, New York.

Osterud, B. & Rapaport, S. I. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5260-5264.

Ostrem, J. A., Al-Obeidi, F., Safar, P., Safarova, A., Stringer, S. K., Patek, M., Cross, M. T., Spoonamore, J., LoCascio, J. C., Kasireddy, P., Thorpe, D., Sepetov, N., Lebl, M., Wildgoose, P. & Strop, P. (1998) *Biochemistry* 37, 1053-1059.

- Padmanabhan, K., Padmanabhan, K. P., Tulinsky, A., Park, C. H., Bode, W., Huber, R., Blankenship, D. T., Cardin, A. D. & Kisiel, W. (1993) *J Mol. Biol.* 232, 947-966.
- Padmanabhan, K., Wu, T. P., Ravichandran, K. G. & Tulinsky, A. (1994) *Protein Sci.* 3, 898-910.
- Park, C. H. & Tulinsky, A. (1986) *Biochemistry* 25, 3977-3982.
- Petersen, T. E., Thogersen, H. C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen L. & Magnusson, S. (1983) *Proc. Natl. Acad. Sci. USA* 80, 137-141.
- Pike, A. C. & Acharya, K. R. (1994) *Protein Science* 3, 706-710.
- Ramakrishnan, V., Patthy, L. & Mangel, W. F. (1991) *Biochemistry* 30, 3963-3969.
- Rejante M. R. & Llinas. M. (1994^a) *Eur. J. Biochem.* 221, 927-937.
- Rejante M. R. & Llinas. M. (1994^b) *Eur. J. Biochem.* 221, 939-949.
- Rydel, T. J., Tulinsky, A., Bode, W. & Huber, R. (1991) *J. Mol. Biol.* 221, 583-601.
- Scanu, A. M., Miles, L. A., Fless, G. M., Pfaffinger, D., Eisenbart, J., Jackson, E., Hoover-Plow, Brunck, T. & Plow, E.F. *J. Clin. Invest.* (1993) 91, 283-291.
- Scanu, A. M., Pfaffinger, D., Lee, J. C. & Hinman J. (1994) *Biochem. Biophys. Acta* 1227, 41-45.
- Scanu, A. M. & Edelstein, C. *Biochem., Biophys. Acta* (1995) 1256, 1-12.
- Scanu, A. M., Pfaffinger, D., Klezovitch, O. & Edelstein, C. (1995) *Cardiovascular Disease*, edited. by Gallo, L. L., pp. 45-47, Plenum Press, New York.
- Schwabacher, A. W., Zhang, S. & Davy, W. (1993) *J. Am. Chem. Soc.* 115, 6995-6996.

Seligm

Sesh

Ses

Sh

Sh

S

S

- Seligmann, B., Stringer, S. K., Ostrem, J. A., Al-Obeidi, F., Wildgoose, P., Walser, A., Sarfar, P., Safarova, A., Lo Cascio, J. C., Sponamore, J., Thrope, D. S., Kasireddy, P., Ashmore, B. & Strop, P. (1995) The Sixth IBC International Symposium, Washington, D.C., Oct. 23-25.
- Seshadri, T., Tulinsky, A., Skrzypczak-Jankun, E. & Park, C. H. (1991) *J. Mol. Biol.* 220, 481-494.
- Seshadri, T., Skrzypczak-Jankun, E., Yin, M. & Tulinsky, A. (1994) *Biochemistry* 33, 1087-1093.
- Sheehan, J. P., Tollefsen, D. M. & Sadler, J. E. (1994) *J. Biol. Chem.* 269, 32747-32751.
- Shing, Y. W., Christofori, G., Hanahan, D., Ono, Y. Sasada, R., Igarashi, K. & Folkman J. (1993) *Science* 259, 1604-1607.
- Silverberg, S. A., Nemerson, Y. & Zur, M. (1977) *J. Biol. Chem.* 252, 8481-8488.
- Soriano-Garcia, Pabmanabhan, K., deVos, A. M. & Tulinsky, A. (1992) *Biochemistry* 31, 2554-2566.
- Sottrup-Jensen, L., Clayes, H., Zajdal, M., Petersen, T. E. & Magnusson, S. (1978) *Prog. Chem. Fibrinolysis Thrombolysis* 3, 191-209.
- St. Charles, R., Matthews, J. H., Zhang, E., Tulinsky, A. & Kahn, M. (1999) *J. Med. Chem.* 42, 1376-1383.
- St. Laurent, D. R., Balasubramanian, N., Han, W. T., Trehan, A., Federici, M. E., Meanwell, N. A., Wright, J. J. & Seiler, S. M. (1995) *Bioorg. & Med. Chem.* 3, 1145-1156.
- Stauffer, D. A., Barrans, R. E. & Dougherty, D. A. (1990) *J. Org. Chem.* 55, 2762-2767.
- Steigemann, W (1974) Die Entwicklung und Anwendung von Rechenverfahren und

St

St

St

Su

Ta

Tu

Tu

Tu

Tu

Va

Vi

Rechenprogrammen zur Strukturanalyse von Proteinen am Beispiel des Trypsin-Trypsininhibitor Komplexes, des Freien Inhibitors und der L-Asparagine [dissertation]. München; Technische Universität.

Stemerman, M. B., Colton, C. & Morell, E. (1984) *Progress in Hemostasis and Thrombosis*, V.7, edited by Spaet, T. H., pp. 289-324, Grune & Stratton, Inc., Orlando.

Stern, D. M., Naworth, P. P., Kisiel, W., Verhar, G. & Esmon, C. T. (1985) *J. Biol. Chem.* 260, 6717-6722.

Stubbs, M. T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S. R. & Bode, W. (1992) *Eur. J. Biochem*, 206, 187-195.

Sunnerhagen, M., Forsen, S., Hoffren, A. M., Drakenberg, T., Teleman, O. & Stenflo, J. (1995) *Nature Struct. Biol.* 2, 504-509.

Taberner, L., Chang, C. Y., Ohringer, S. L., Lau, W. F., Iwanowicz, E. J., Han, W. C., Wang, T. C., Seiler, S. M., Roberts, D. G. M. & Sack, J. S. (1995) *J. Mol. Biol.* 246, 14-20.

Tucker, T. J., Lumma, W. C., Mulichak, A. M., Chen, Z., Naylor-Olsen, A. M., Lewis, S. D., Lucas, R., Freidingel, R. M. & Kuo, L. C. (1997) *J. Med. Chem.* 40, 830-832.

Tulinsky, A., Mani, N. V., Marimoto, C. N. & Vandlen, R. L. (1973) *Acta Crystallogr.*, B29, 1309-1319.

Tulinsky, A., Park, C. H. & Skrzypczak-Jankun, E. (1988) *J. Mol. Biol.* 202, 885-901.

Tulinsky, A. (1996). *Seminars in Thrombosis and Hemostasis*. 22, 117-124.

Vandlen, R. L. & Tulinsky, A (1973) *Biochemistry* 12, 4193-4203.

Vijayalakshmi, J., Padmanabhan, K. P., Mann, K. G. & Tulinsky, A. (1994) *Protein Sci.*

Webe

Well

Wu.

Wu.

Wu

Ye.

Zh.

Zh

Zh

Zh

3, 2254-2271.

Weber, P. C., Lee, S. L., Lewandowski, F. A., Schadt, M. C., Chang, C. H. & Kettner, C.

A. (1995) *Biochemistry* 34, 3750-3757.

Wells, C. M. & Di Cera, E. (1992). *Biochemistry* 31, 11721-11730.

Wu, T-P., Padmanabhan, K., Tulinsky, A. & Mulichak, A. (1991) *Biochemistry* 30,
10589-10594.

Wu, T-P., Yee, V., Tulinsky, A., Chrusciel, R. A., Nakanishi, H., Shen, R., Priebe, C. &
Kahn, M. (1993) *Protein Eng.* 6, 471-478.

Wu, T-P. Padmanabhan, K. P. & Tulinsky, A. (1994) *Blood Coag. Fibrinolys.* 5, 157-
166.

Ye, J., Rezaie, A. R. & Esmon, C. T. (1994) *J. Mol. Biol.* 269, 17965-17970.

Zhang, E. & Tulinsky, A. (1997) *Biophys. Chem.* 63, 185-200.

Zhang, E., St. Charles, R. & Tulinsky, A. (1999) *J. Mol. Biol.* 285, 2089-2104.

Zhuang, P., Blackburns, M. N. & Peterson, C. B. (1996^a) *J. Biol. Chem.* 271, 14323-
14332.

Zhuang, P., Li, H., Williams, J. G., Wagner, N. V., Seiffer, D. & Peterson, C. B. (1996^b)
J. Biol. Chem. 271, 14333-14343.