



This is to certify that the

dissertation entitled

Crystal Structure Studies of Kringle
Domains and A Peptidomimetic Inhibitor
Complex of α -Thrombin

presented by

Tswei-Ping Wu

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Chemistry

Major professor

Date 8/2/93

Cr

an

**Crystal Structure Studies of Kringle Domains
and a Peptidomimetic Inhibitor Complex of
 α -Thrombin**

By

Tswei-Ping Wu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

1993

Cr

a

The
plex. PG
a lysine
properties
tic acid re
in PGK
which is c
The lysine
most likely
formation
in the asyn
a lysine res
molecule. T

ABSTRACT

Crystal Structure Studies of Kringle Domains and a Peptidomimetic Inhibitor Complex of α -Thrombin

By

Tswei-Ping Wu

The crystallographic structures of the PGK4- ϵ -aminocaproic acid (ACA) complex, PGK1 and tPAK2 have been determined and show that each of them possesses a lysine binding site composed of three distinct regions according to electrostatic properties: (1) a negatively charged region containing two negatively charged aspartic acid residues, (2) a distinct positively charged region due to one (in tPAK2) or two (in PGK4-ACA and PGK1) side chain amino groups and (3) a hydrophobic region, which is composed of aromatic residues, separates the oppositely charged regions. The lysine binding sites of apo-PGK4 and PGK4-ACA have been compared and most likely the lysine binding site is preformed, lysine binding does not require conformational changes of the host. In the tPAK2 structure, three molecules are found in the asymmetric unit. The crystal structure shows a strong interaction between a lysine residue of one molecule and the lysine binding pocket of the neighboring molecule. This interaction mimics the ligand binding interaction found in the PGK4-

ACA comp

that of apo

intermolec

between tw

α -Thro

tides A + F

Gly bonds

bound str

molecular

tween the

and inhib

plexed w

of FPAM

were mea

chromom

are differe

that of th

critical re

pancreatic

Arg-chlor

ACA complex. The overall conformation of the PGK1 structure is very similar to that of apo-PGK4. The largest deviations are due to insertion of Gln59 in PGK1 and intermolecular interactions occurring around the two-fold axis relating the segment between two neighboring PGK1 molecules.

α -Thrombin displays remarkable specificity, effecting the removal of fibrinopeptides A (FPA) and B (FPB) of fibrinogen through the selective cleavage of two Arg-Gly bonds among 181 Arg/Lys-Xaa bonds in fibrinogen. A model for the thrombin-bound structure of FPA has been proposed based on NMR data and computer-assisted molecular modeling. In order to obtain a better understanding of the interplay between the primary sequence and the conformation required for thrombin substrates and inhibitors, a crystallographic investigation of the FPA mimetic (FPAM) complexed with human α -thrombin was undertaken. Crystals of ternary complexes of FPAM, hirugen and thrombin were grown and three-dimensional intensity data were measured to 2.5 Å resolution. The crystallographic structure of FPA and its chloromethylketone derivative bound to thrombin were determined. Although there are differences between these structures, in the above modeled FPA structure, and that of the crystal structure of FPAM bound to thrombin, the ϕ , ψ angles in the critical region of P1-P2-P3 in all of the structures are similar to those of bovine pancreatic trypsin inhibitor (BPTI) in the BPTI-trypsin complex and D-Phe-Pro-Arg-chloromethylketone (PPACK) in the PPACK-thrombin structure.

To my parents

I wish to
advice, guid
preciation
research, g
Pappas, Pa
for her exp
ble throug
sample, an
Linas, Dr
extended t

My gra
their endi
my husban
in my diff
dissertatio

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Professor Alexander Tulinsky for his advice, guidance and financial support throughout the course of this work. My appreciation is also extended to all present and former members of Professor Tulinsky's research group for their assistance, friendship and good well. I especially thank Dr. Pappan Padmanabhan for the numerous discussions and Dr. Pushpa Padmanabhan for her expert crystallization technique. The completion of this work was made possible through the supplying of plasminogen kringle 4 protein sample, kringle 1 protein sample, and the coordinates of tissue plasminogen kringle 2 structure from Dr. Miguel Llinas, Dr. Francis Castellino and Dr. Bart de Vos, respectively. Thanks are also extended to Dr. Mike Kahn for synthesizing mimetic of fibrinogen peptide A for us.

My gratitude and appreciation for life goes to my parents and parents-in-law for their endless love and support. Last, but not in any way least, I would like to thank my husband Tyan-Shu for his unending understanding and encouragement, especially in my difficult time. His patience for teaching me in using $\text{\LaTeX}$ to word-process this dissertation is greatly appreciated.

LIST OF

LIST OF

1 Fibrin

2 Experi

2.1 C

2.2 D

2.3

2.4

2.5 D

2.6

2.7

2.8 M

2.9

2.10

2.11 S

3 Struct

3.1 I

3.2 M

3.3 R

4 Struct

4.1 I

4.2 M

4.3 R

5 Struct

5.1 I

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	ix
1 Fibrinolysis and Kringles	1
2 Experimental and Computational Methods	9
2.1 Crystallization	9
2.2 Data Collection	12
2.2.1 Diffractometer	12
2.2.2 Area Detector	15
2.3 Data Reduction	18
2.3.1 Diffractometer Data	18
2.3.2 Area Detector Data	19
2.4 Molecular Replacement	22
2.4.1 Rotation Search	24
2.4.2 Translation Search	26
2.5 Structure Refinement	27
3 Structure of the ACA Complex of Human Plasminogen Kringle 4	31
3.1 Introduction	31
3.2 Materials and Methods	35
3.3 Results and Discussion	49
4 Structure of Human Plasminogen Kringle 1	61
4.1 Introduction	61
4.2 Materials and Methods	65
4.3 Results and Discussion	74
5 Structure of Tissue-type Plasminogen Activator Kringle 2 Domain	100
5.1 Introduction	100

5.2 Exp

5.3 Res

6 Compar

7 The Str

α -Thro

7.1 Int

7.2 Exp

7.3 RE

7.4 Dis

A Amino

BIBLIOG

5.2	Experimental Procedure	105
5.3	Results and Discussion	107
6	Comparison of Different Kringles	133
7	The Structure of A Designed Peptidomimetic Inhibitor Complex of α-Thrombin	143
7.1	Introduction	143
7.2	Experimental Procedures	151
7.3	RESULTS	155
7.4	Discussion	170
A	Amino Acid Shorthand Used in the Thesis.	173
	BIBLIOGRAPHY	174

1.1 Co-
ing
pr-
tive
act

2.1 Re

3.1 Su

3.2 A

3.3 D

3.4 W

3.5 S

3.6 N

w

D

3.7 H

3.8 R

3.9 D

4.1 S

4.2 D

s

4.3 W

4.4 S

4.5 S

4.6 H

I

de

4.7 It

LIST OF TABLES

1.1	Comparison of primary structure of homologous kringles. The numbering of the residues is standardized to PG K5. Abbreviations used: PT, prothrombin, PG, plasminogen, uPA, urokinase-type plasminogen activator; HGF, hepatocyte growth factor; tPA, tissue-type plasminogen activator; FXII, factor XII.	6
2.1	Rotation matrix in terms of Eulerian angles θ_1 , θ_2 , θ_3	24
3.1	Summary of crystal data of monolinic K4-ACA	37
3.2	Averaging background in 2θ shells.	44
3.3	Decay rate, S , calculated in each 2θ range.	44
3.4	Weights of Reflections and R Values of the Final Refinement	47
3.5	Summary of Final Least-Squares Parameters and Deviations	48
3.6	New Intramolecular Hydrogen Bonds of K4-ACA. Hydrogen atoms were assigned geometrically idealized positions. Donor atom is denoted (D), acceptor atom (A).	53
3.7	Hydrogen Bonds Between K4 and ACA	56
3.8	RMS difference between the K4-ACA and Apo-K4	56
3.9	Differences in lysine-binding sites between K4-ACA and Apo-K4	58
4.1	Summary of crystal data of tetragonal K1.	68
4.2	Distribution of reflection intensities and R-factors in various resolution shells.	69
4.3	Weights of Reflections and R Values of the Final Refinement	72
4.4	Summary of Final Least-Squares Parameters and Deviations	73
4.5	Secondary structural elements of PGK1 structure.	77
4.6	Hydrogen Bonds found in the lysine binding site of the PGK1 structure. The prime designates a symmetry-related molecule. A donor atom is denoted (D), an acceptor atom (A).	87
4.7	Intermolecular interactions of PGK1 < 3.7Å.	94

4.8 Devel
the E

5.1 Ass
ture

5.2 Wel

5.3 Sum

5.4 Seco

5.5 The

sh

m

5.6 RA

M

pa

5.7 De

3P

6.1 H

PC

7.1 T

re

re

7.2 D

sh

7.3 W

7.4 F

7.5 R

T

4.8	Deviations of the positions of residues in the lysine binding site between the PGK1 and PGK4 structures.	98
5.1	A summary of the final refinement results of the X-PLOR refined structure.	106
5.2	Weights of reflections and R values of the final refinement cycle of tPAK2	108
5.3	Summary of final least-squares parameters and deviations of tPAK2 .	109
5.4	Secondary structural elements of the tPAK2 structure.	115
5.5	The intramolecular and intermolecular interactions present in the lysine binding site of tPAK2 structure. Prime refers to neighboring molecules. Donor atom is denoted (D), acceptor atom (A).	124
5.6	RMS deviations among the three molecules in the asymmetric unit. MA, MB, MC represent molecules A, B and C. Numbers shown in parentheses were calculated without including Met28.	126
5.7	Deviations of the positions of residues in the lysine binding site between tPAK2 and K4-ACA	131
6.1	Hydrogen bonds of the conserved disulfide anti-parallel β -sheets of PGK1, PGK4, tPAK2, PTK1, and PTK2 structures.	138
7.1	The primary sequence of human α -thrombin. The thrombin residue numbers are assigned by homology with chymotrypsin. Insertions are represented by alphabetic characters followed by numbers.	146
7.2	Distribution of reflection intensities and R-factors in various resolution shells.	153
7.3	Weights of reflections and R values of the final refinement	157
7.4	Final least squares parameters and deviations of FPAM-thrombin. . .	158
7.5	RMS Deviations Between the Hirugen-Thrombin and FPAM-Hirugen-Thrombin Complexes	161

1.1 The
min.
min.
1.2 The
ind.
1.3 Lys.
BA.
CH.
1.4 Co.
2.1 Co.
B
2.2 Dia.
2.3 Bio.
2.4 Gra.
2.5 The
to
ro.
3.
2.6 S.
3.1 The
- I
ac
cle
3.2 The
a
3.3 S.
in

LIST OF FIGURES

1.1	The model of the fibrinolytic system. Abbreviations used: PG, plasminogen; PM, plasmin; tPA, tissue plasminogen activator; PAI-1, plasminogen activator inhibitor type 1; LS, lysine sites of fibrin.	2
1.2	The three-disulfide triple-loop kringle structure. Disulfide bridges are indicated by bold lines.	3
1.3	Lysine ligands analogs. Abbreviations used: ACA, aminocarproic acid; BASA, p-benzylamine-sulfonic acid; AcLys, N-acetyl-L-lysine; AM-CHA, trans-4-(aminomethyl)-cyclohexanecarboxylic acid.	5
1.4	Comparison of the folding of PGK4 and PTF1. PTF1 in bold.	7
2.1	Crystallization by vapor diffusion technique. (A) Hanging drop method (B) Sitting drop method.	10
2.2	Diagram of four-circle diffractometer.	13
2.3	Block diagram of four-circle area detector system.	16
2.4	Graph of count rate vs. bias control setting.	17
2.5	The Eulerian angles θ_1 , θ_2 , θ_3 that relate the rotated axes X', Y', Z' to the original axes X, Y, Z. The rotation operation consists of 1.) a rotation by θ_1 around Z-axis. 2.) a rotation by θ_2 around new X-axis. 3.) a rotation by θ_3 around new Z-axis.	25
2.6	Schematic Structure of "PROFFT" and "PROLSQ".	29
3.1	The chain structure of plasminogen. Kringle domains are labeled K1 – K5. Disulfide bridges are shown by bold lines. Asterisks denote active site residues in the serine protease domain. Arrows indicate the cleavage sites.	32
3.2	The primary structure of plasminogen K4. Interkringle residues are a-c and 81-87. (see appendix for abbreviations)	34
3.3	Stereoview of modeled lysine binding site of plasminogen kringle 4.ACA in bold. Hydrogen atoms are put at ideal positions [38].	36

3.4

3.5

3.6

3.7

3.8

3.9

3.10

3.11

3.12

4.1

4.2

4.3

4.4

4.5

4.6

4.7

4.8

4.9

4.10

4.11

4.12

4.13

4.14

4.15

4.16

4.17

4.18

4.19

4.20

3.4	Photograph of K4-ACA crystal. The size of the single crystal is approximately $1.2 \times 0.3 \times 0.1$ mm.	38
3.5	Schematic diagram of K4-ACA crystal morphology with respect to crystallographic axes.	39
3.6	Axial intensity distributions of K4-ACA crystal.	40
3.7	The ω -profile of reflection (5, 5, 5) taken before data collection. . . .	41
3.8	Absorption correction curve of (0, 2, 0) reflection.	43
3.9	Ramachandran plot of final ϕ , ψ angles of K4-ACA. The Gly residues are boxed.	50
3.10	Stereoview of the electron density of the lysine-binding site of K4-ACA. The basket contour is at 1σ ; ACA is between Asp55, Asp57 and Lys35, Arg71.	51
3.11	Stereoview of the lysine-binding site of K4-ACA. ACA is shown in bold; hydrogen bonds are dashed.	55
3.12	Stereoview of the comparison of the lysine-binding site of K4-ACA and apo-K4. K4-ACA is in bold; the sulfate position in apo-K4 is designated by S.	59
4.1	The primary structure of human plasminogen kringle 1. Site numbering deletion is based on homology with PGK5.	62
4.2	Stereoview of the modeled lysine binding site of plasminogen kringle 1; ACA in bold. Hydrogen atoms are shown at ideal positions (taken from [38]).	64
4.3	Photograph of K1 crystal. Crystal length is approximately 1.2 mm. . .	67
4.4	The Ramachandran plot of PGK1. Glycine residues are boxed.	75
4.5	Distribution of ω angles in the PGK1 structure.	76
4.6	Plot of the average thermal factors of main-chain (thin line) and side-chain (thick line) atoms of PGK1 structure.	78
4.7	Distribution of temperature factors for water molecules.	80
4.8	Distribution of occupancies for water molecules.	81
4.9	Stereoview of the internal water molecule in PGK1 structure. Dashed lines indicate hydrogen bonds.	82
4.10	Stereoview of the PGK1 aromatic and proline residues forming hydrophobic core.	83
4.11	Stereoview of the PGK1 lysine binding site.	85
4.12	Electron density in vicinity of Arg34 and Arg71 of PGK1 structure; contour at 1σ	86

4.13 Stere
bind

4.14 RMS
tures

4.15 RMS
Files

4.16 Stere
PGF

4.17 Stere
Res

two

4.18 Stere
site

mode
circle

4.19 Stere
PGF

5.1 SLD
plast

cont
are

5.2 Prim
delet

5.3 Elec
Con

5.4 Stere
Hydr

5.5 Stere
the

5.6 Stere
in bo

5.7 Stere
Retro

axis

5.8 Stere
Retro

is in d

4.13	Stereoview showing intermolecular interactions involved in the lysine binding site. The prime denotes a symmetry-related molecule.	88
4.14	RMS deviations of CA, C, N atoms between PGK1 and PGK4 structures. Filled circles indicate conserved residues.	90
4.15	RMS deviations of side chains between PGK1 and PGK4 structures. Filled circles indicate conserved residues.	91
4.16	Stereoview of the optimally superimposed main chain structures of PGK1 (bold) and PGK4.	92
4.17	Stereoview of crystal packing interactions at a two-fold symmetry axis. Residues involved in intermolecular interactions are shown in bold. The two-fold axis in the a-b diagonal is indicated.	95
4.18	Stereoview of intermolecular interactions found in the lysine binding site of PGK4. The side chains from the second symmetry-related molecule are shown in bold; the sulfate ion is designated with a filled circle [23].	96
4.19	Stereoview of the superposition of the lysine binding sites of the PGK1(bold) and PGK4 structures.	97
5.1	Schematic presentation of the primary structure of human tissue-type plasminogen activator. The arrow indicates the cleavage site for the conversion of single chain tPA to two chain tPA. The active site residues are indicated by asterisks.	101
5.2	Primary structure of tPA kringle 2. Site numbering insertions and deletions are based on homology with PGK5.	103
5.3	Electron density of chloride ion in the lysine binding site of tPAK2. Contour at 1σ	110
5.4	Stereoview showing an internal water molecule in the tPAK2 structure. Hydrogen bonds are represented by dashed lines.	111
5.5	Stereoview showing the water molecule found in an empty pocket of the tPAK2 structure. Hydrogen bonds are represented by dashed lines.	112
5.6	Stereoview of CA, C, and N structure of tPAK2. Disulfides are shown in bold.	114
5.7	Stereoview of the crystal packing of the three molecules in the asymmetric unit of the tPAK2 structure. The non-crystallographic 3_1 screw axis is perpendicular to the plane.	116
5.8	Stereoview of the crystal packing of the three molecules in the asymmetric unit of tPAK2 structure; The non-crystallographic 3_1 screw axis is indicated.	117

5.9

5.10

5.11

5.12

5.13

5.14

5.15

5.16

5.17

6.1

6.2

6.3

6.4

6.5

6.6

6.7

5.9	Stereoview showing the intermolecular interactions between Lys48' and Lys33, Asp55, and Asp57. Prime represents the neighboring molecule.	118
5.10	Stereoview showing the intermolecular interaction between Asn26' and Asp55 bridged by a water molecule. Prime represents the neighboring molecule.	119
5.11	Stereoview of the crystal packing of the three molecules along the a-axis. Residues involved in intermolecular interactions are shown in bold.	121
5.12	Stereoview of the lysine binding site of the tPAK2 structure.	122
5.13	Intramolecular and intermolecular interactions present in the lysine binding site of the tPAK2 structure.	123
5.14	Stereoview of the comparison of the main-chain atoms of the three independent molecules.	127
5.15	Stereoview of the comparison of the side-chain atoms of the three independent molecules.	128
5.16	Stereoview showing the comparison of side chains of His13 in three molecules. His13 in molecule B is indicated.	129
5.17	Stereoview of superimposed lysine binding sites of tPAK2 (bold) and K4-ACA. Lys48' is between Asp55, Asp57 and Lys33 in tPAK2; ACA is between Asp55, Asp57 and Lys35, Arg71 in K4-ACA.	132
6.1	Stereoview of the comparison of the CA, C, N structures of different kringles. PGK1, PGK4, and PTK1 are shown in thin lines; tPAK2 in bold; PTK2 in dashed.	134
6.2	Stereoview of the comparison of the inner disulfide bridges of PGK1, PGK4, tPAK2, PTK1 and PTK2. Disulfide bonds in bold.	136
6.3	Stereoview showing the conserved anti-parallel β -sheets located near disulfide bridges in PGK1, PGK4, tPAK2, PTK1, and PTK2.	137
6.4	Stereoview comparing the lysine binding regions of PTK1 (bold) and PGK4.	140
6.5	Stereoview comparing the lysine binding site regions of PTK2 (bold) and PGK4.	141
6.6	Stereoview of the comparison of internal aromatic residues in PGK1, PGK4, tPAK2, PTK1 and PTK2.	142

7.1

7.2

7.3

7.4 S

F

7.5 F

O

7.6 N

se

N

d

7.7 R

G

7.8 S

FF

su

7.9 S

th

7.10 St

hy

7.11 St

th

7.12 St

ket

7.1	The conversion of prothrombin to thrombin. CHO represents a carbohydrate side chain; a, b, and c represent the cleavage sites of bovine and human prothrombin; b' represents an additional cleavage site in human prothrombin.	144
7.2	The structure of the fibrinogen molecule. The symmetric molecule is composed of a dimeric central domain containing the N-terminal of all six chains (α , β , γ), two connecting coiled coils, two terminal domains, and two A α polar appendages. Four carbohydrate clusters (CHO) occur, and are located on each γ chain near the central domain and on the β chains of each terminal domain. Primary cross-linking sites (XL) can be found near the C-terminal of the γ chain and in the A α polar appendages (taken from [76]).	147
7.3	Scheme for synthesis of FPAM. DEAD, diethyl-azodicarboxylate; EDC, 1-ethyl-3 (3-dimethylaminopropyl) carbodiimide; HOBT, hydroxybentriazole; DMAP, dimethylaminopyridine; DMF, dimethylformamide.	149
7.4	Stereoview of the FPAM structure docked in the thrombin active site. FPAM is in bold.	150
7.5	Photograph of FPAM crystal. Crystal size is approximately $0.5 \times 0.45 \times 0.1$ mm	152
7.6	Numbering of and special restraints used for FPAM in restrained least squares refinement. Since part of FPAM is a non-amino acid group, special bond, angle and planar 1,4 distance restraints were applied during refinement to maintain geometry.	156
7.7	Ramachandran plot of ϕ , ψ angles of FPAM-thrombin structure. Glycines are not displayed.	159
7.8	Stereoview of the comparison of the active sites of thrombin in the FPAM and hirugen complexes. Hirugen-thrombin, broken lines; active site is unoccupied.	160
7.9	Stereoview of the electron density corresponding to FPAM in the thrombin complex. Basket contour at 1σ	162
7.10	Stereoview of FPAM bound in active site of thrombin. FPAM in bold; hydrogen bonds, broken.	163
7.11	Stereoview of the comparison of FPAM (bold) and PPACK in their thrombin complexes.	164
7.12	Stereoview of the comparison of FPA (bold) and FPA-chloromethyl ketone in their thrombin complexes.	167

7.13 Stereoview of the comparison of FPAM (bold) and FPA in their throm-	
bin complexes.	169

CHAPTER 1

Fibrinolysis and Kringles

In the blood coagulation cascade, fibrin is formed to block the flow of blood from a severed vessel. After it has fulfilled its function, it is digested to soluble fragments by the enzyme plasmin; we call this process of dissolution of the fibrin blood clot “fibrinolysis”. The fibrinolytic system, schematically represented in Figure 1.1, consists of three main components: plasma zymogen plasminogen (PG), its activated product, the proteolytic enzyme plasmin (PM) and tissue plasminogen activator (tPA). As proposed by Wiman and Collen [1], when fibrin is formed, a small amount of plasminogen is always specifically bound to it. Plasminogen activators present in the blood or released from the vascular endothelium are adsorbed on the fibrin surface and efficiently activate the plasminogen (Figure 1.1). Plasmin released from the clot is rapidly inactivated by the fast-acting plasmin inhibitor α_2 -antiplasmin ; therefore, other plasma proteins are not attacked by plasmin. Since plasminogen and inhibitor α_2 -antiplasmin are abundantly present in blood at rather stable levels, the blood fibrinolytic activity is determined by the balance between tPA and plasminogen activator inhibitor type 1 (PAI-1).

The interactions between PM, PG and tPA with lysine residues of fibrin are localized at lysine binding sites that reside in the kringle domains of these molecules (Figure 1.1). The interactions are thought to occur with exposed lysine residues of the



Figure 1.1. The
 gen. PM. plasma
 inhibitor type 1

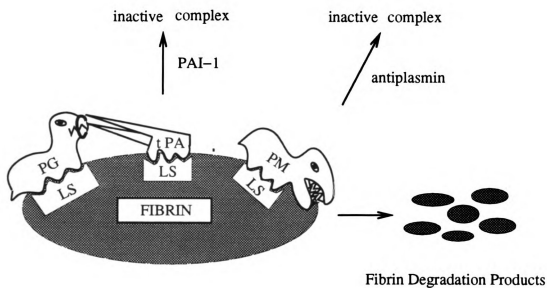


Figure 1.1. The model of the fibrinolytic system. Abbreviations used: PG, plasminogen; PM, plasmin; tPA, tissue plasminogen activator; PAI-1, plasminogen activator inhibitor type 1; LS, lysine sites of fibrin.

horiz

not on

on the

It

affine

within

this st

teracti

sites m

Th

consist

protein

to kring

these fr

they are

Figure 1
indicated

Kring
blood pla
show the

fibrin peptide chain [2]. The affinity between lysine binding sites and lysine residues not only localizes the PG and tPA to the fibrin clot, but also keeps plasmin formed on the fibrin surface from being inactivated by antiplasmin.

There are several lysine binding sites in plasminogen, one binding site with high affinity and four sites with low affinity [3, 4, 5]. The strong binding site is located within the first kringle of plasminogen, which can bind specifically to fibrin through this site. Since the lysine binding sites of plasminogen and plasmin mediate their interactions with fibrin and antiplasmin, the interactions suggest that the lysine binding sites may play a crucial role in the regulation of fibrinolysis.

The kringle domains have three disulfide triple loop patterns (Figure 1.2) and each consists of 80–85 residues of molecular weight approximately 10,000 daltons. Limited proteolysis of prothrombin (PT) and PG is known to yield fragments corresponding to kringles, indicating that kringles are independent structural units [6, 7]. Also, these fragments retain their original lysine binding function after being isolated and they are thought to be independent functional domains [7, 8, 9].

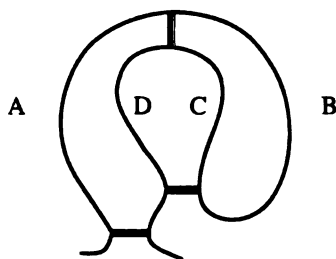


Figure 1.2. The three-disulfide triple-loop kringle structure. Disulfide bridges are indicated by bold lines.

Kringles are present in several non-catalytic regions of trypsin-type proteases of blood plasma that are involved in blood coagulation and fibrinolysis [10]. It has been shown that kringles occur singly in urokinase-type plasminogen activator (uPA) [11],

factor XII [12], as pairs in prothrombin [6], tissue-type plasminogen activator [13], as 4 copies in hepatocyte growth factor (HGF) [14], and as 5 copies in plasminogen [7]. Most interestingly, there are 38 kringles found in apolipoprotein[a] [15], 37 of which display 75–85% conservation with kringle 4 of plasminogen.

Although the different kringles have the same disulfide triple loop pattern, their binding specificities diverge in binding different proteins or low-molecular weight ligands. For instance, the kringle 5 (K5) fragment of PG binds to benzamidine-Sepharose [16], fragment 2 of PT, which corresponds to the second kringle of the PT, has an intact binding site for factor V_a [8]. In addition, K1 and K4 of PG [9, 17] and K2 of TPA [18] not only bind to fibrin and lysine but also to some ω -aminocarboxylic acids, including ϵ -aminocarproic acid (ACA), p-benzylamine-sulfonic acid (BASA), N-acetyl-L-lysine (AcLys) and trans-4-(aminomethyl)-cyclohexanecarboxylic acid (AMCHA) (Figure 1.3). All of the ligands have positive and negative charged groups about 6.8 Å apart and have been found to have an antifibrinolytic effect *in vivo* [19].

A comparison of kringle sequences found in some proteins is given in Table 1.1, from which it can be seen that the kringles of the different proteases show a high degree of sequence homology; in fact, they appear to be more closely related than the protease parts themselves [20]. The comparison also shows that about 25% of the residues are conserved absolutely among kringles from different proteins. These conserved residues are found around Cys22–Cys63 and Cys51–Cys75, which are essential for the kringle-fold. This observation supports the hypothesis that conserved residues in all kringles are essential for the folding autonomy of the domains.

The first kringle structure solved by X-ray crystallography was the K1 of prothrombin fragment 1 (PTF1) [21, 22]. Recently, human plasminogen kringle 4 has also been determined [23] and shows a similar three dimensional kringle structure to that of PTF1. A comparison of the folding of these two kringles is shown in Figure 1.4, from which it can be seen that the structures of the kringles are somewhat

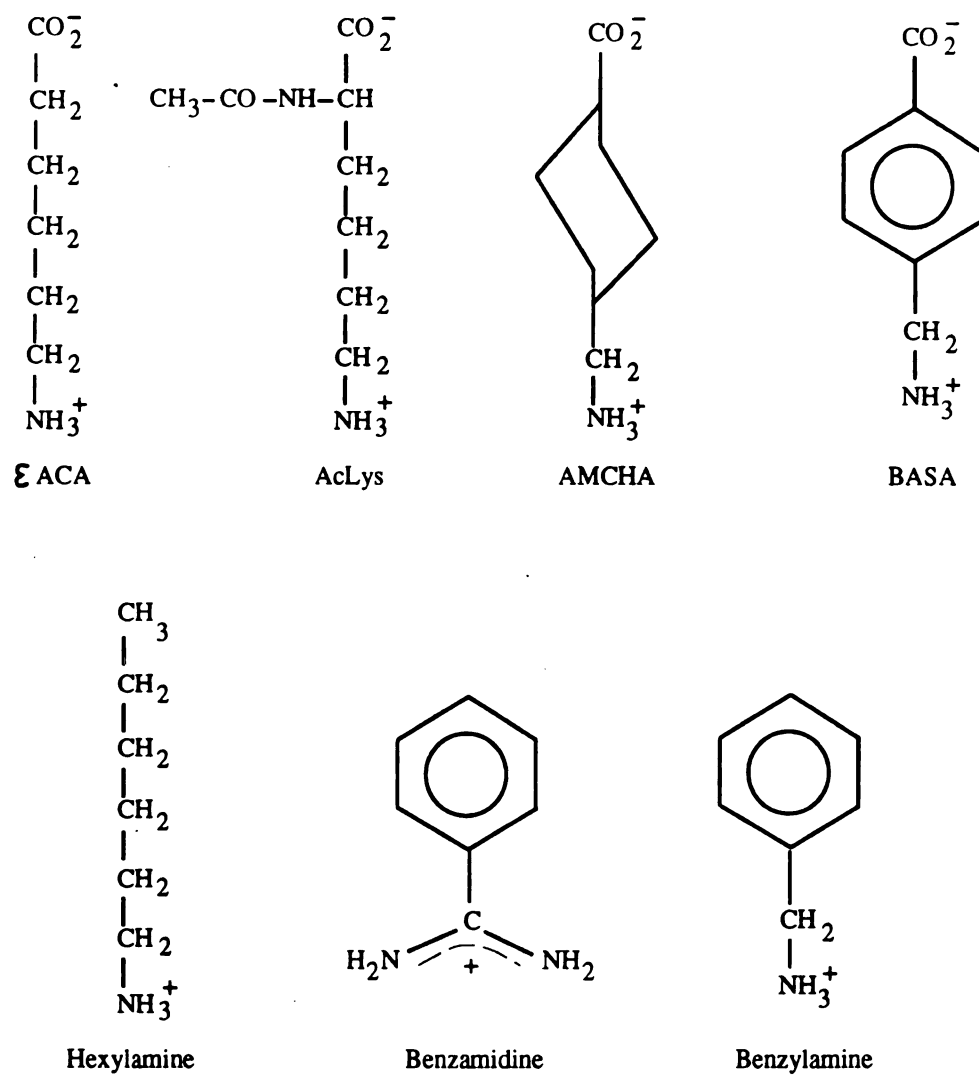


Figure 1.3. Lysine ligands analogs. Abbreviations used: ACA, aminocarproic acid; BASA, p-benzylamine-sulfonic acid; AcLys, N-acetyl-L-lysine; AMCHA, trans-4-(aminomethyl)-cyclohexanecarboxylic acid.

Table 1.1
of the res
plasma
factor 1

CPA K2
CPA K1
CPA K1
PG K4
PG K1
PG K2
PG K3
PG K5
EEF K1
EEF K2
EEF K3
EEF K4
PT K1
PT K2
P XII

CPA K2
CPA K1
CPA K1
PG K4
PG K1
PG K2
PG K3
PG K5
EEF K1
EEF K2
EEF K3
EEF K4
PT K1
PT K2
P XII

Table 1.1. Comparison of primary structure of homologous kringles. The numbering of the residues is standardized to PG K5. Abbreviations used: PT, prothrombin, PG, plasminogen, uPA, urokinase-type plasminogen activator; HGF, hepatocyte growth factor; tPA, tissue-type plasminogen activator; FXII, factor XII.

		1	10	20	30	40
tPA	K2	GNSD	CYFGNGSAYR	GTHSLTESGA	SCLPWNSMIL	IGKVYTAQNP
tPA	K1	DTRAT	CYEDQGISYR	GTWSTAESGA	ECTNWNSSAL	AQKPYSGRRP
uPA	K1	DKSKT	CYEGNGHFYR	GKASTDTMGR	PCLPWNSATV	LQQTYHAHRS
PG	K4	TPVVQD	CYHGDGQSYR	GTSSTTTTGK	KCQSWSSMTP	HRHQKT-PEN
PG	K1	VYLSE	CKTGDGKNYR	GTMSKTKNGI	TCQKWSSTSP	HRPRFS-PAT
PG	K2	LECEEE	CMHCSENYD	GKISKTMISGL	ECQAWDSQSP	HAHGYI-PSK
PG	K3	SGPTYQ	CLKGTGENYR	GNVAVTVSGH	TCQHWSAQTP	HTHNRT-PEN
PG	K5	TPSEED	CMFGNGKGYR	GKRATTVTGT	PCQDWAAQEP	HRHSIFTPET
HGF	K1	NKDYIRN	CIIGKGRSYK	GTVSITKSGI	kCQPWSSMIP	HEHSFL-PSS
HGF	K2	VE	CMTCNESYR	GLMDHTESGK	ICQRWDHQTP	HRHKFL-PER
HGF	K3	DVPLETTE	CIQGQGEGRY	GTVNTIWNIG	PCQRWDSQYP	HEHDMT-PEN
HGF	K4	HGQD	CYRGNGKNYM	GNLSQTRISGL	TCSMWDKNME	DLHRHI-FWE
PT	K1	AACLEGN	CAEGLGTNYR	GHVNITRSGI	ECQLWRSRYP	HKPEIN-STT
PT	K2	SPPLEQ	CVPDRGQQYQ	GRLAVTTHGL	PCLAWASAQA	KALSKH-QDF
F	XII		CYDGRGLSYR	GLARTTISGA	PCQPWASEA-	--TYRNVTAE

		50	60	70	80
tPA	K2	SAQALGLGKHNY	CRNPDGDA-K	PWCHVLKNRR	LTWEYC-DVPSC ST
tPA	K1	DAIRLGLGNHNY	CRNPDRDA-K	PWCYVFKAGK	YSSEFC-STPAC SEG
uPA	K1	DALQLGLGKHNY	CRNPDNRR-R	PWCYVQVGLK	PLVQEC-MVHDC ADG
PG	K4	YP-NAGLT-MNY	CRNPDADK-G	PWCFTTDPST	-RWEYC-NLKKC SGT
PG	K1	HP-SEGLE-ENY	CRNPDNDPQG	PWCYTTPDEK	-RYDYC-DILEC EEE
PG	K2	FP-NKNLK-KNY	CRNPDREL-R	PWCFTTDPNK	-RWELC-DIPRC TTP
PG	K3	FP-CKNLD-ENY	CRNPDGKR-A	PWCHTTNSQV	-RWEYC-KIPSC DSS
PG	K5	NP-RAGLE-KNY	CRNPDGDVGG	PWCYTTPNPK	-LYDYC-DVPQC AAP
HGF	K1	YR-GKDLQ-ENY	CRNPRGEEGG	PWCFTSNPEV	-RYEVC-DIPQC SEV
HGF	K2	YP-DKGFD-DNY	CRNPDGQP-R	PWCYTLDPHT	-RWEYC-AIKTC ADNTMNDT
HGF	K3	FK-CKDLR-ENY	CRNPDGSE-S	PWCFTTDPNI	-RVGYCSQIPNC DMS
HGF	K4	PD-ASKLN-ENY	CRNPDDDAHG	PWCYTGNPLI	-PWDYC-PISRC EGDTPPTIVNL
PT	K1	HP-GADLQ-ENF	CRNPDSSITG	PWCYTTPDPTV	-RRQEC-SIPVC GQD
PT	K2	NS-AVQLV-ENF	CRNPDGDEEG	VWCYVAGKPG	-DFGYC-DLNYC EEA
F	XII	QARNWGLGGHAF	CRNPDNDI-R	PWCFVLNRDR	LSWEYC-DLAQC

A

Figure 1

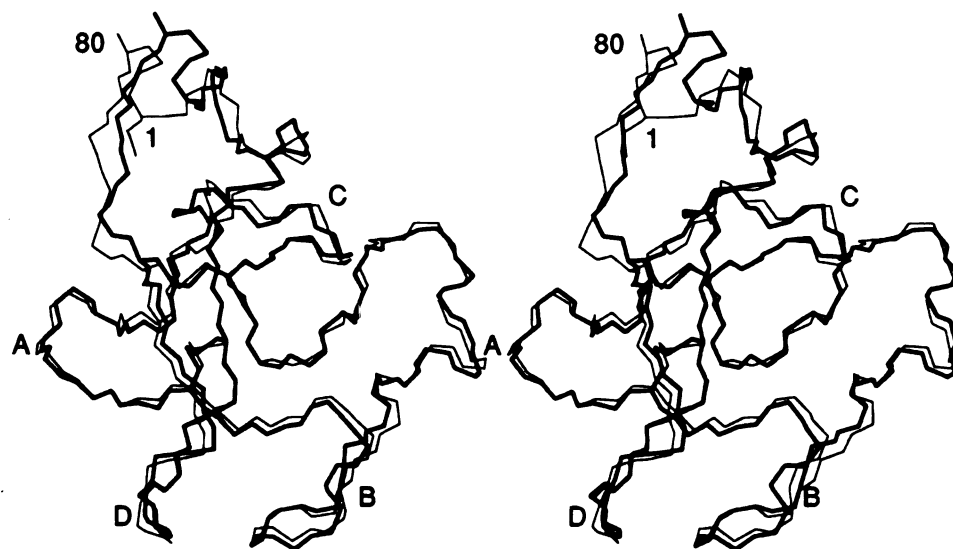


Figure 1.4. Comparison of the folding of PGK4 and PTF1. PTF1 in bold.

symmetric

loop A a

the whole

duplicate

technical

not ident

structure

The di

methods

in chapter

given and

PTF2 will

FPAM can

symmetrical and very similar. The combination of the second half of the first outer loop (A) and the second inner loop (D) is almost the mirror symmetry counterpart of the whole second outer loop (B). The folding of the kringle also shows a complicated duplication by operation of two 90° rotations and a translation. Since apolipoprotein[a] (apo[a]) carries 38 copies of the kringle, 37 of which are highly homologous but not identical to PGK4, it is possible to model kringle structures based on the PGK4 structure and further investigate kringle–kringle interactions of apo[a].

The dissertation is organized as following: the experimental and computational methods used for crystallizing and solving the crystal structures will be introduced in chapter 2. In chapters 3,4 and 5, structures of K4ACA, PGK1 and tPAK2 will be given and discussed. Comparison of the above three kringle structures and PTF1 and PTF2 will be presented in chapter 6. Chapter 7 describes the structure of a designed FPAM complex of α -thrombin.

CHAPTER 2

Experimental and Computational Methods

2.1 Crystallization

The most important factor in solving protein structures is to have satisfactory single crystals. Crystals can be grown from a saturated solution in any of several ways, all of which serve to raise the solute concentration above that which can be supported by the solution: slow evaporation, slow cooling, vapor diffusion. The best crystals are usually produced when the solution is free from mechanical vibration and allowed to evaporate without disturbance.

Among the crystallization methods, vapor diffusion techniques (Figure 2.1) are probably the most widely used to grow protein crystals. As shown in the Figure 2.1, a drop of protein containing precipitant is equilibrated against a reservoir in which the concentration of precipitant is higher than that of the protein solution. Equilibration proceeds by diffusion of the volatile species (water or organic solvent) until vapor pressure in the droplet equals that of the reservoir. Since the drops used in this method can be as small as $2\mu\text{l}$, the vapor diffusion technique is very well suited to screening a large number of conditions when only a small quantity of material is

Figure 21. Cryo-
B String drop

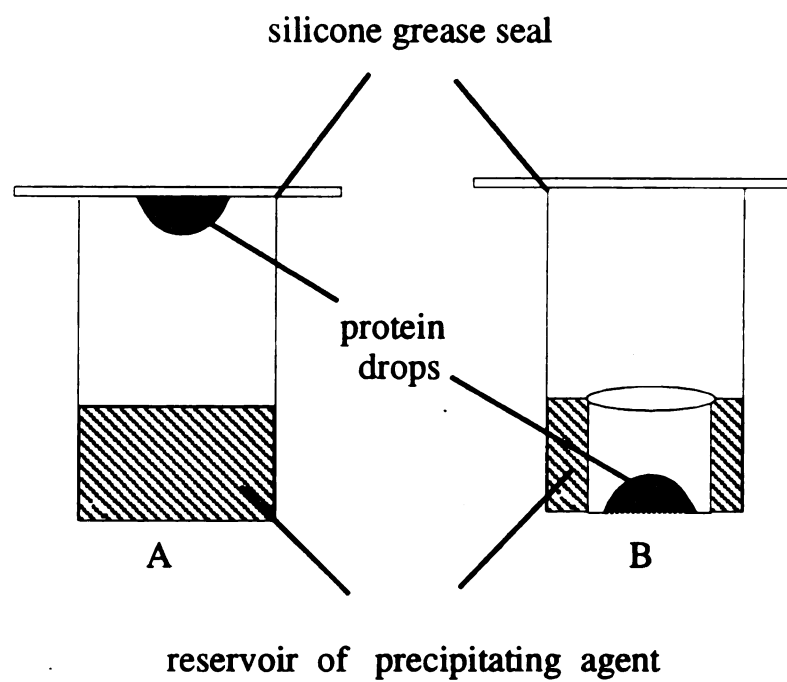


Figure 2.1. Crystallization by vapor diffusion technique. (A) Hanging drop method (B) Sitting drop method.

avail

T

oids

suspen

small

the con

in p f

disposal

tage the

compa

The s

The seal

grease. A

open end

well soluti

solution is

box is seal

When c

crystals res

ters, such as

can be prod

illary in con

work.

available.

The vapor diffusion technique employs both hanging drop and sitting drop methods (Figure 2.1). For the hanging drop method, a microdroplet of mother liquor is suspended from the underside of a microscope cover slip, which is then placed over a small well containing 1 ml of the precipitating solution. An important point is that the cover slips must be thoroughly and carefully coated with silicone to ensure proper drop formation and prevent spreading. The wells are most conveniently supplied by disposable plastic Linbro plates that have 24 wells. These plates provide the advantage that they can be swiftly and easily examined under a microscope and stored compactly.

The sitting drop method shown in Figure 2.1 is usually used to get larger crystals. The seal between the lid and the rim of a clean box is coated with some silicone grease. A clean siliconized glass vial is then placed at the bottom of the box with the open end up. One ml of the well solution is added by pipette. A 20 μ l drop of the well solution is placed in the bottom of the vial and the same amount of the protein solution is then added with slight mixing. The lid is then placed on the box and the box is sealed.

When crystallization is finally induced, it is usually so rapid that only microscopic crystals result. Using microseeding methods and by changing crystallization parameters, such as protein concentration, precipitant concentration, large and single crystals can be produced. Once a good crystal is chosen, it is mounted in a sealed glass capillary in contact with its mother liquor and properly aligned for further experimental work.

Two

re and

later

2.2.1

Three

time

from a

be use

and

be

and

this

There

by

and

the

for

the

at

at

the

the

the

the

2.2 Data Collection

Two kinds of detectors are now used for 3-D intensity data collection: the diffractometer and the area detector. The former measures reflection intensities individually; the latter has the ability to measure many reflections at once.

2.2.1 Diffractometer

Three-dimensional intensity data were collected on a Nicolet P3/F four-circle diffractometer (Figure 2.2) using a graphite monochromator with $\text{CuK}\alpha$ radiation (1.54 Å) from a rotating anode X-ray tube. The diffractometer possesses four circles that can be used to adjust the orientation of the crystal and counter, which include the ϕ , χ , ω , and 2θ angles. In order to measure any reflection, the reciprocal lattice point must be brought onto the sphere of reflection in the equatorial plane of the diffractometer and the detector must be positioned at the proper value of 2θ . In order to achieve this, the ω , ϕ , χ angles must be determined and set correctly for each reflection. Therefore, each diffracted X-ray beam is defined by 2θ , ω , ϕ , and χ angles as well as by Miller indices h , k , l . The high sensitivity of the diffractometer to angle settings allows it to achieve an accuracy of alignment and parameter measure that exceeds the accuracy usually obtained photographically by an order of magnitude.

The method used to measure the reflection intensities was the Wyckoff ω -step procedure [24], most widely used in protein crystallography. The crystal is slowly moved through the reflection position, and the profile of the intensity of the reflection and the background are recorded by the stationary detector. In this procedure, every reflection is scanned by a limited number at small intervals in ω , as opposed to the full profile scan from background to background. Background measurements are made on either side of the peak by offsetting by a certain ω value. The scan method therefore allows a slight mis-setting for the crystal. The breadth of a single peak at half the

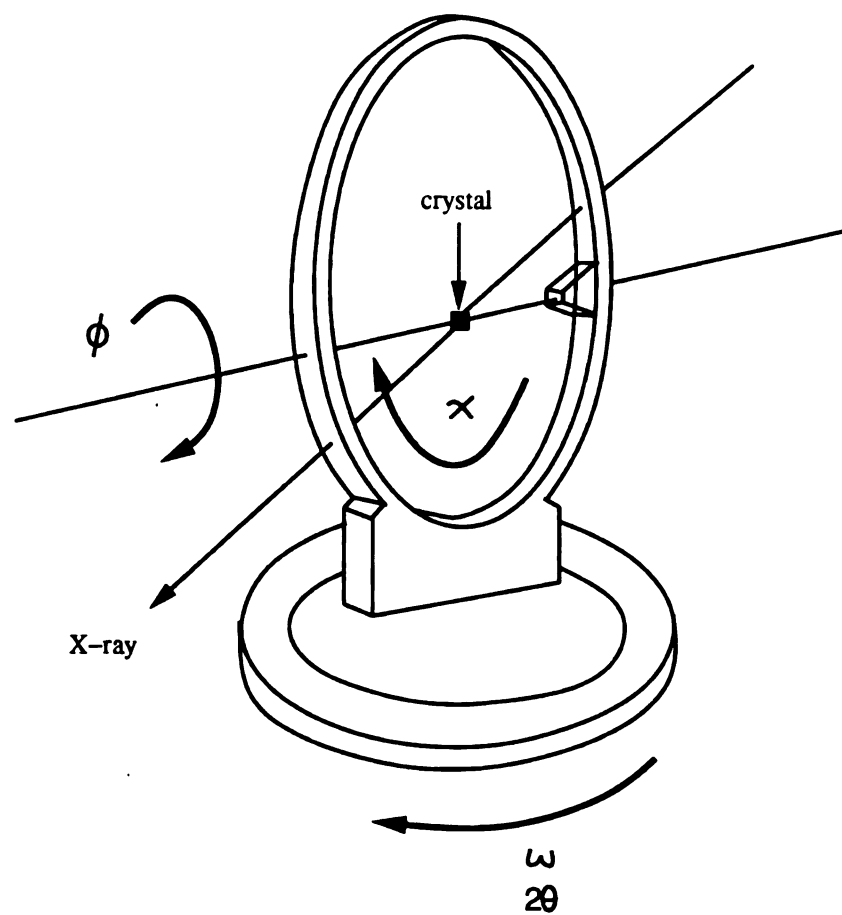


Figure 2.2. Diagram of four-circle diffractometer.

20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

maximal height defines the region of the peak which should be scanned. The ω profile not only can be a basis for measurement but also reports the quality of a crystal. A split profile may indicate that the crystal is cracked and that two closely related but distinct crystal orientations are contributing to the profile.

Since the macroscopic dimensions of a crystal are usually not equal, absorption due to the crystal and the medium around it is unavoidable. As a result, empirical corrections for absorption have become common, especially in protein crystallography, where many reflections are involved. The most quantitative approximation, as suggested by North *et al.* [25], involves setting χ to 90° and measuring the independent intensity of one or more reflections as a function of the ϕ range. This allows us to determine the optimum region of ϕ for carrying out the data collection. Reflections on a principal axis which coincides with the ϕ axis are selected to do the ϕ scan because the Bragg planes are then always in the reflecting position for all values of ϕ . The observed variation can easily be converted to an approximate correction as a function of ϕ and applied to the entire data set.

Many crystals, especially those of protein and other macromolecules, show a more or less steady decrease in diffraction intensity during the process of data collection. This deterioration, which is due to direct X-ray exposure, affects the consistency of the measurements. In order to make the decay correction, a number of measurements of crystal deterioration during data collection are monitored, representing intensities in various resolution ranges. These include a 2-dimensional data set from 2° – 15° in 2θ , collected before and after the data collection, and several monitor reflections measured throughout the data collection after every 100 reflections about every hour. If the intensity of any monitor falls below 80% of the initial measured intensity, the crystal is considered misaligned and this triggers the recentering of the crystal, followed by the calculation of a new orientation matrix and cell parameters by the least squares method.

Therap

sally

under

the

surre

The

sion

to the

He a

person

are p

sion

the c

pion

to r

sion

p

sion

sion

sion

sion

sion

sion

sion

sion

sion

2.2.2 Area Detector

The rapidly increasing interest in the X-ray crystallography of macromolecules has led steadily to the study of larger and larger unit cells. These not only have a greater total number of reflections to be measured, but also cause more to be reflecting at a given time. As a result, area detectors have been devised which combine the sensitivity, accuracy, and convenience of the diffractometer but measure many reflections at once.

The Siemens Xentronics Area Detector with graphite monochromated CuK_α radiation and a Rigaku RU200 rotating anode X-ray generator is used for data collection in this laboratory. A schematic of the area detector system is shown in Figure 2.3. The area detector mounts on a dovetail track on the diffractometer, permitting adjustment of the detector to specimen distance. Area detector operation and control are performed by the frame buffer, in which the FRAMBO (FRAME Buffer Operation) program is used to collect data and display frames. The output signals of the detector are decoded by the Position Decoding Circuit (PDC) to create a digital position of each X-ray event measured by the detector. The resulting position is sent to the frame buffer for processing. The primary purpose of the frame buffer is to receive data from the PDC and convert it into frame information. The frame buffer processor also drives a high resolution color display which shows the output from the area detector in two forms: first, as a real-time color display showing the diffraction pattern building as it is collected; and second, as a data frame showing the results of several seconds (or minutes) of output integrated in the file as a single frame.

Before 3-D X-ray diffraction intensities are measured, some calibrations of the instrument must be performed, first, the bias adjustment. Since the x-ray signal height is varied by adjusting the detector bias, the bias should be set at a proper value. From Figure 2.4 it can be seen that while lowering the bias slowly there will be a point at which the rate meter reading falls. This is the lower window setting.

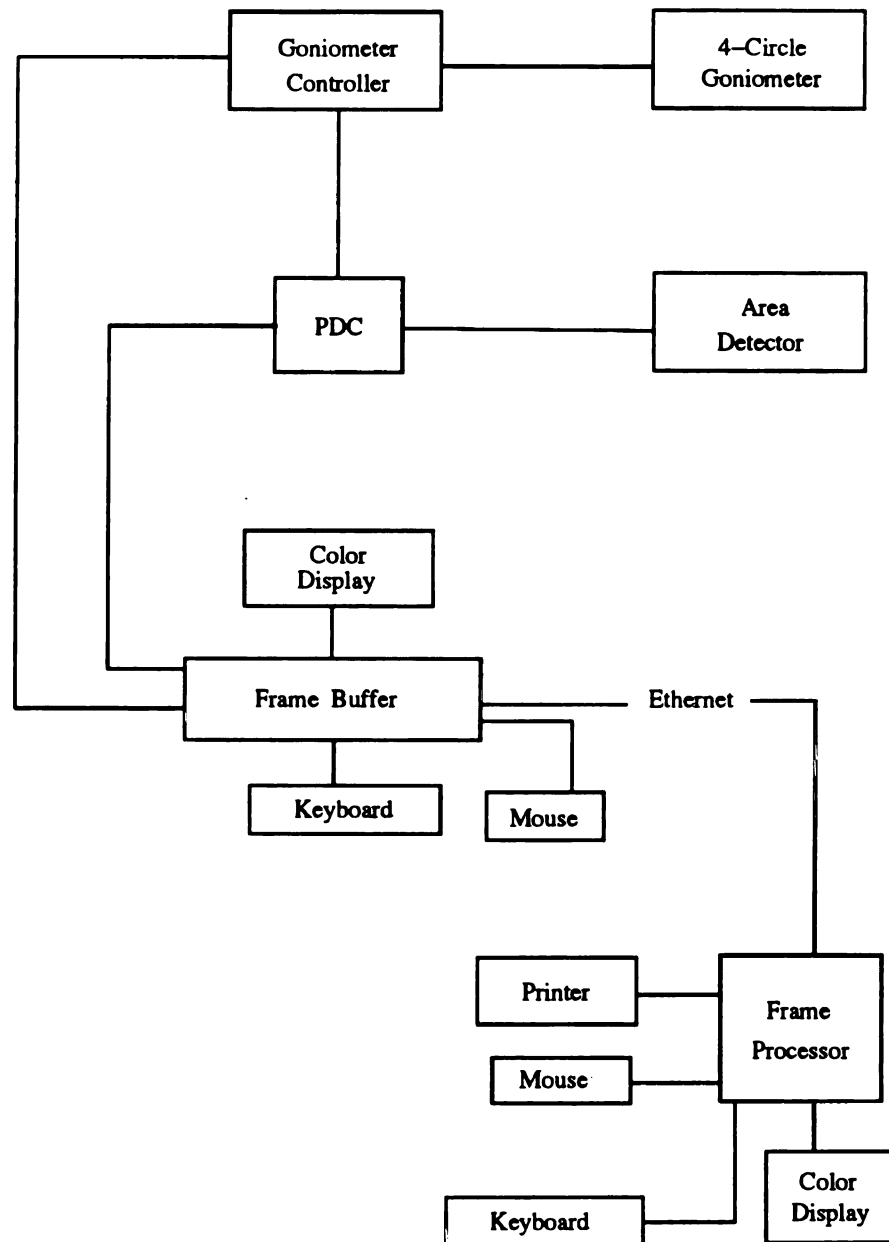


Figure 2.3. Block diagram of four-circle area detector system.

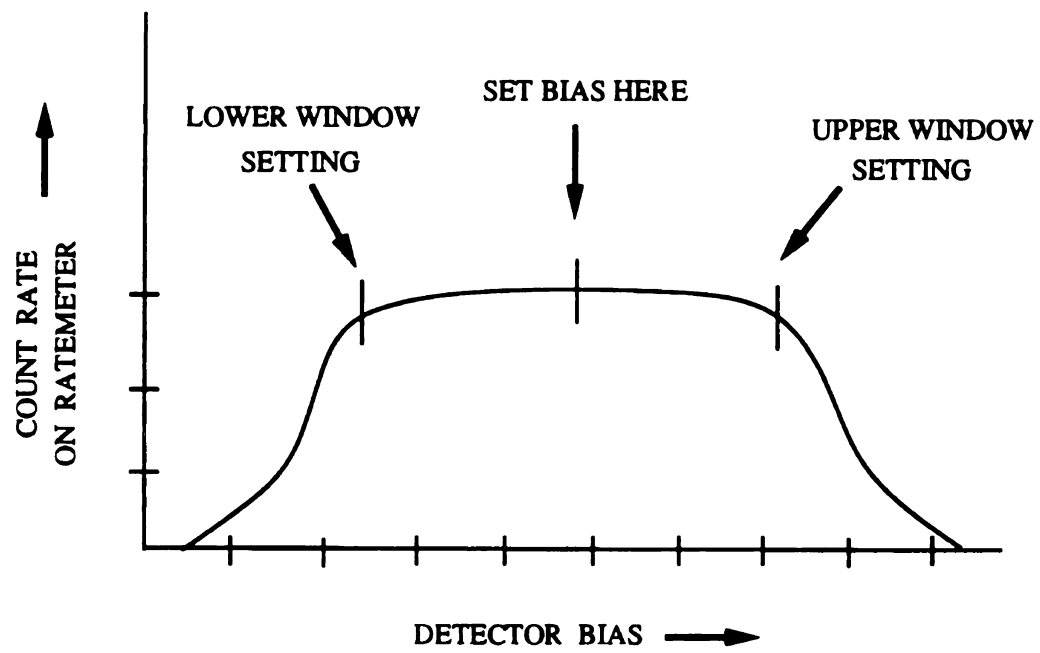


Figure 2.4. Graph of count rate vs. bias control setting:

Increasing the bias setting should cause the rate meter to peak and then fall again at the upper window setting. The bias should be set midway between the upper and lower window. The bias setting can be performed with a ^{55}Fe calibration source. The second calibration collects a flood-field image taken from an ^{55}Fe source to generate a lookup table which corrects the detector response to an uniform field. A third correction collects an image from the iron source with a precisely machined brass plate mounted in front of the active surface of the detector. The data obtained from this calibration frame are used to generate a calibration table for conversion from detector address in pixels to true positions in centimeters.

2.3 Data Reduction

2.3.1 Diffractometer Data

The purpose of data reduction is to convert each reflection intensity $I(hkl)$ to its corresponding structure factor amplitude $|F(hkl)|$. The intensity data are reduced using a program called P-DATA, written by Dr. C. D. Buck in this laboratory. The program calculates the structure factor amplitudes from diffraction intensities based on the equation:

$$|F(hkl)|^2 = CONST \times ABS \times DEC \times LORPOL \times I(hkl) \quad (2.1)$$

where $CONST$ is a constant used to scale the data, ABS is the absorption correction factor, DEC is a correction for intensity decay as a function of X-ray exposure time, $LORPOL$ is the Lorentz-polarization correction factor which is a function of 2θ and monochromator, and $I(hkl)$ is the background-corrected intensity of a reflection.

Before applying the P-DATA program, the background of the intensity data is averaged in shells of 2θ , and ϕ , if there is a ϕ dependence. Background-corrected

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

Inter-

2.3.2

X-ray

the XE

range

as a

intensities are calculated from the equation:

$$I(intensity) = \left[total\ scan\ count - \frac{sum\ of\ background\ counts}{background\ to\ scan\ ratio} \right] \times scan\ rate \quad (2.2)$$

Intensities are considered to be observed according to the minimum intensity (I_{min}) which is taken as 2 times the magnitude of the average of negative intensity ($\langle I_{neg} \rangle$).

The absorption correction applied in the P-DATA program is based on the tables of $I_{max}/I(\phi)$ versus ϕ . The data in these tables are derived from the observed ϕ -dependence of reflections at $\chi = 90^\circ$. The absorption correction may vary with 2θ and up to 6 tables are allowed.

The decay correction factor, DEC, corrects the intensity deterioration of a reflection as a function of exposure time. Using a 2-D set of data measured before and after data collection and 3 monitor reflections, , we can get a decay factor for the crystal based on following equation:

$$DEC = \frac{1}{1 - S \cdot t} \quad (2.3)$$

where $S = -s/I^0$, I^0 is the intensity at zero time, and s is the slope of the intensity versus exposure time. Since the check reflections and 2-D data set have different 2θ values, a 2θ -dependent DEC is derived from the S versus 2θ plot and all reflections collected at different times and 2θ values are corrected by the decay factor.

2.3.2 Area Detector Data

X-ray diffraction intensities collected from area detectors are reduced and scaled with the XGEN packages of programs [26]. Programs are supplied to refine, integrate, merge and scale data, producing a comprehensive listing of integrated intensities and a statistical analysis of the data. There are many steps involved in data reduction.

(A) I

A p

by a

The

ge m

over

where

hence

the p

entire

in the

Any

part

that

(B) C

in the

Program

in the

of the

of the

(A) Defining the active pixels on the detector face

A pixel is inactive if it is outside the overall active area of the detector or is obscured by a beamstop or some other objects in the path between the crystal and detector. The program *BORDER* reads a set of frames which are collected under the same geometric condition. In each frame the program determines the mean count value over the entire detector face according the equation:

$$\langle C(n) \rangle = \frac{\sum_{xy} C(x, y, z)}{(512 \times 512)} \quad (2.4)$$

where n is the frame number. For each pixel it adds a number to the appropriate element in a 512×512 buffer. The number it adds will be the number of counts in that pixel in that frame unless the count value is greater than twice the mean for the entire frame. In this case, twice the mean is added in. Then the value at pixel (x, y) in the buffer after frame n will be:

$$Z(x, y, n) = Z(x, y, n - 1) + \min(C(x, y, n), 2 \times \langle C(n) \rangle) \quad (2.5)$$

Any pixel whose Z value is above a defined fraction of the mean is assumed to be part of the active area of the detector. This effectively defines beamstop shadows and other peculiarly-shaped regions.

(B) Creating the calibration file

In this step the detector pixels are converted to actual distances in centimeters by the program *CALIBRATE*. It reads a fiducial-plate data frame made on the frame buffer by collecting several million counts with an ^{55}Fe source and then finds the centroids of the spots on the fiducial image and indexes them relative to one another. From this information it then constructs a mapping of detector pixels into centimeters from

and

(C) C

The p

of the

of the

U.S.

and

and

D.

On

and

the

has

and

and

(1)

1

;

s

a reference position.

(C) Generating a list of bright spots

The program *SPOTS* reads a series of contiguous data frames and constructs a list of bright spots for the crystal and detector refinement. A spot is regarded as bright if it includes at least one pixel which is N sigma brighter than the local background. Usually N can be adjusted to generate approximately 20 bright spots per data frame, which is a reasonable number. Therefore, the stronger the diffraction the larger is the number N used.

(D) Calculating the orientation matrix and the detector parameters

Once a set of bright spots centroids have been assembled, an initial set of unit cell and crystal orientation parameters can be obtained with the *REFINE* program. It then refines these parameters and the position of the detector and the rocking-curve behavior of the crystal. All reflections with three indices closer to integers than the error limit are used in a linear refinement to minimize the differences between the integerized and unintegerized indices.

(E) Computing integrated data

The program *INTEGRATE* reads a series of contiguous data frames and computes the integrated intensities of the Bragg reflections both by profile-fitting and by simple summation. Both include these corrections : (1) The Lorentz effect; (2) Polarization; (3) Variations in exposure time from frame to frame; (4) Dead-time loss; (5) The contribution to the profile of pixels not included in the summation or the profile-fitting. The program not only integrates intensities but also refines the crystal parameters based on the observed data.

(F) I

to be

of the

of the

about

under

(G) I

The p

data

(H) S

A line

from

state

into

by the

(I) C

The 6

from

from

2.4

to solve

and by

(F) Reformatting the integrated reflections

In order to make integrated reflections compatible with available structure solution software, the program *REDUCE* reads the integrated reflections and then writes a file containing most of the pertinent information and none of the unnecessary information about each observation. Also, the output file created by *REDUCE* is sorted in hkl order and contains symmetry-related observations.

(G) Merging the integrated reflections

The program *MRMERGE* reads the reflections from *REDUCE*, then merges together data from various orientations of one crystal.

(H) Scaling merged data

A linear least-squares scaling of the merged intensity data is performed by the program *SCALEI* based on minimizing the differences among intensities of symmetry-related observations. *SCALEI* breaks the data into 5° ranges and splits each 5° range into two pseudo-films. One, two, or three scaling parameters per film are calculated by the program, and they are used to calculate scaled intensities for the observations.

(I) Creating merged intensities or structure amplitudes

The final step of data reduction is to write an output file of either intensities or structure factors which is performed by the program *MAKEMU*. Once this output is obtained, structure determination can be carried out in further work.

2.4 Molecular Replacement

To solve the structure of any macromolecule, an electron density map must be generated for model-building and structure interpretation. However, calculating such elec-

two of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

the of

tron density maps is not straightforward, since they require the phase angles which can not be measured during data collection. In this investigation, molecular replacement [27] provides the best way for determining initial phases of macromolecular crystal structures. In this technique, a molecule of known structure (search model) is used to probe the diffraction data of an unknown crystal structure; angles and translation vectors are found which will locate the probe structure in the unit cell of the unknown structure to generate an initial phasing model for crystallographic structure fitting and refinement. Molecular replacement is thus concerned with finding the three rotational and three translational parameters that specify the orientation and position of the molecule in the crystal cell with respect to the symmetry elements. If the model is sufficiently similar to the unknown structure, molecular replacement can be rather straightforward to apply.

Molecular replacement is performed by calculating Patterson functions of the known model and the unknown structure based on the equation:

$$P(uvw) = (1/V) \sum |F|^2 \cos(2\pi(hu + kv + lw)) \quad (2.6)$$

where V is the volume of the unit cell, and $|F|^2$ is proportional to the intensity of the diffracted beam and u , v and w are the coordinates of a point within the unit cell. Since no phase information is required for this function, it can be calculated from both unknown structure and search model. The molecular replacement method consists of three stages: (1) Rotation: determination of relative orientation of the independent molecules; (2) Translation: using information from (1) to position the molecule in the unit cell in the correct orientation; (3) After completing (1) and (2), the equivalence between the point X' in the unknown structure and the point X in

the model can be expressed by

$$X' = [C]X + d \quad (2.7)$$

where $[C]$ is the rotation matrix determined in stage (1) and d is the translation vector determined in stage (2).

2.4.1 Rotation Search

The rotation search is carried out using the SEARCH routine in the program “PROTEIN” package [28]. The rotation function (Equation 2.8) measures the degree of coincidence between Patterson functions from unknown structure P_1 and search model P_2 which is rotated by a matrix C with respect to unit cell.

$$R = \int P_1(x) \times C \times P_2(x) dv \quad (2.8)$$

Maximum values of R occur at a certain matrix C which gives close agreement between vector sets of the unknown structure and model. The matrix C (Table 2.1) is usually expressed in terms of three Eulerian angles θ_1 , θ_2 and θ_3 as illustrated in Figure 2.5.

Table 2.1. Rotation matrix in terms of Eulerian angles θ_1 , θ_2 , θ_3 .

$$\begin{pmatrix} -\sin \theta_1 \cos \theta_2 \sin \theta_3 & \cos \theta_1 \cos \theta_2 \sin \theta_3 & \sin \theta_2 \sin \theta_3 \\ +\cos \theta_1 \cos \theta_3 & +\sin \theta_1 \cos \theta_3 & \\ -\sin \theta_1 \cos \theta_2 \cos \theta_3 & \cos \theta_1 \cos \theta_2 \cos \theta_3 & \sin \theta_2 \cos \theta_3 \\ -\cos \theta_1 \sin \theta_3 & -\sin \theta_1 \sin \theta_3 & \\ \sin \theta_1 \sin \theta_2 & -\cos \theta_1 \sin \theta_2 & \cos \theta_2 \end{pmatrix}$$

Page 2
of 2
Date 2/2/2006
Time 12:00

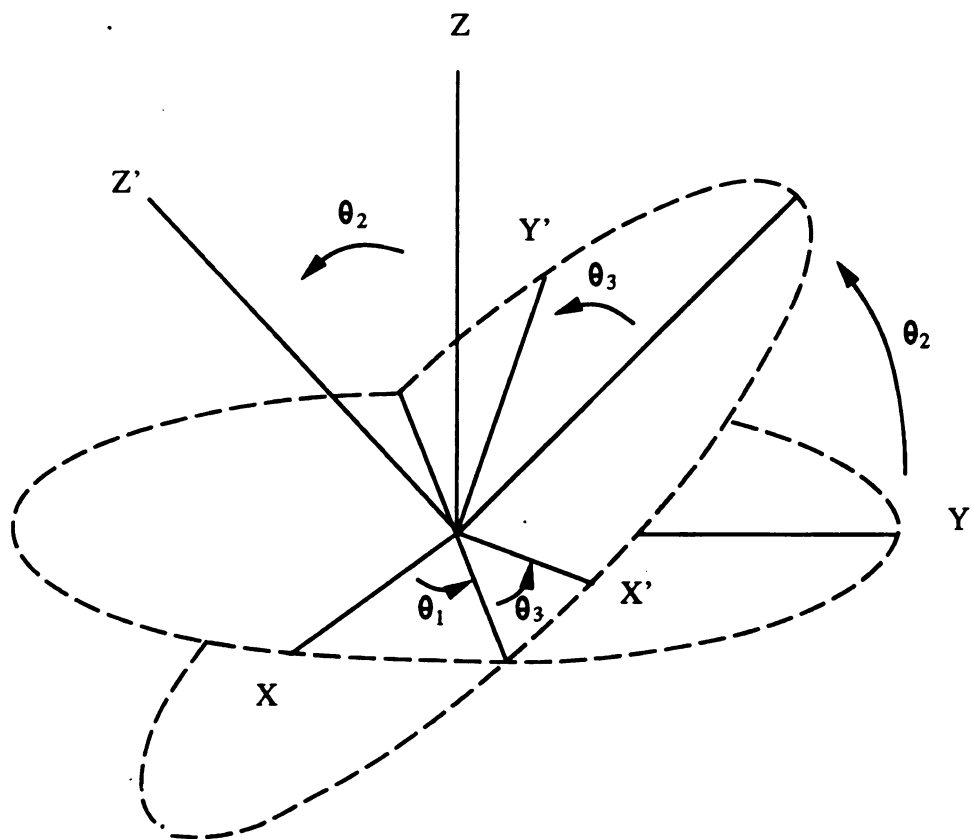


Figure 2.5. The Eulerian angles θ_1 , θ_2 , θ_3 that relate the rotated axes X' , Y' , Z' to the original axes X , Y , Z . The rotation operation consists of 1.) a rotation by θ_1 around Z -axis. 2.) a rotation by θ_2 around new X -axis. 3.) a rotation by θ_3 around new Z -axis.

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

In macromolecular structures, the number of Patterson vectors required for rotation function calculation depends on the number of atoms in the model and in the unknown structure. The Patterson vectors of the search model are rotated and interpolated to the nearest grid in the Patterson of the unknown molecule and a product function calculated. This process is repeated for each set of angles. The angles of the rotation matrix corresponding to the highest value of the product function correspond to the transformation needed to bring the model and the unknown to the same orientation.

2.4.2 Translation Search

Once the relative orientation of the search model and the unknown structure is known, the translational parameters can be determined. A translation function that has been used widely is one that employs a linear correlation coefficient to determine the correct position of an oriented molecule in the crystal cell. The correlation coefficient (C') is represented as:

$$C' = \frac{\sum (|F_o|^2 - |\langle F_o \rangle|^2) \times (|F_c|^2 - |\langle F_c \rangle|^2)}{\sqrt{\sum (|F_o|^2 - |\langle F_o \rangle|^2)^2 \times \sum (|F_c|^2 - |\langle F_c \rangle|^2)^2}} \quad (2.9)$$

Like the R factor, it is basically a measure of the agreement between observed, $|F_o|$, and calculated quantities, $|F_c|$.

The calculation of correlation coefficients has been implemented in the program "BRUTE" [29]. It moves the search model over a grid of points in the crystal cell. At each point the symmetry-related positions are generated and the structure factors are calculated. The amplitudes of these calculated structure factors are then compared with the observed values using the correlation coefficient, C' , as well as the conventional R factor. Therefore the overall calculation time is a function of the number of symmetry operations, the number of reflections and the number of

22

one

246

2.5

The

20

10

22

PR.

9

1

10

30

10

99

10

22

10

22

2

2.4.2

grid points. The most important feature of the program is its ability to adjust the orientation of the model. This allows the errors in the orientation obtained from the rotation function to be corrected.

2.5 Structure Refinement

The crystal structure can be refined employing the restrained least-squares refinement method either implemented in the program “PROLSQ” [30] or “PROFFT” [31] after applying the rotational matrix and translational vector. This method is a very general and powerful technique for optimizing the fit of a model to a set of observations. For PROLSQ, however, application to even moderate-resolution protein structures can result in high computational costs, as structure factors and deviations between observation and calculation are computed directly from contributions by each reflection. Therefore, “PROLSQ” has been modified with the fast-Fourier-transform algorithms added throughout the computation and implemented in the program “PROFFT”. This approach reduces by up to a hundredfold the computation time for a single cycle.

The principle of least squares states that the “best” set of model parameters is that which minimizes the sum of the weighted squares of the difference between observation and calculation. It can be stated by the following mathematical terms:

$$\phi = \sum_h^{\text{reflections}} W_h [F_h^{\text{obs}} - f_h^{\text{calc}}(\{\chi\})]^2 \quad (2.10)$$

where $\{\chi\}$ is the set of parameters, F_h and f_h are observed and calculated values and W_h , the applied weights, are the inverses of the variances of the observations. The parameters included in the function are structure factors, bond distances, angle distances, torsion angles, 1-4 planar distances, planarities, chiralities, nonbonded

could be

the one

and the

the

the

F. and

in B.

and

and

and

and

and

the

and

the

the

the

the

the

the

the

the

the

contacts, and thermal parameters. Sometimes appropriate restraints are placed on the occupancy factor values of atoms in activity effectors, water, and other solvent molecules.

The refinement process used in “PROLSQ” or “PROFFT” are summarized by the schematic shown in Figure 2.6. The program reads diffraction data (h , k , l , F) and scattering factors ($f(s)$) prepared by SCATT, initial atomic coordinates (x , y , z , B) and the restraint specifications (ideal values) prepared by PROTIN, parameter shifts from previous refinement cycles, and control card-images. It then augments the normal-equation elements pertinent to each of the stereochemical restraints and the structure factor observations. Program MERGE then combines the shifts produced by either “PROLSQ” or “PROFFT” with the atomic coordinates to produce an updated coordinate file.

The most commonly used measure of the degree of match between observed and calculated structure amplitudes is the residual R :

$$R = \frac{\sum_{hkl} \left| |F_{hkl}^{obs}| - |F_{hkl}^{calc}| \right|}{\sum_{hkl} |F_{hkl}^{obs}|} \quad (2.11)$$

If the value of the residual after the second Fourier synthesis is smaller than its former value, it indicates that the second electron density map is likely to be a more reliable representation of the actual molecular structure than the original trial model.

The intermittent model building performed on an Evans and Sutherland PS390 interactive stereographics by using program FRODO [32] is carried out after each stage of refinement. During the model building, the structure is manually adjusted to better positions according to $(2|F_o| - |F_c|)$ and $(|F_o| - |F_c|)$ maps and the Ramachandran plot.

At the beginning of the refinement, an average isotropic thermal parameter B is

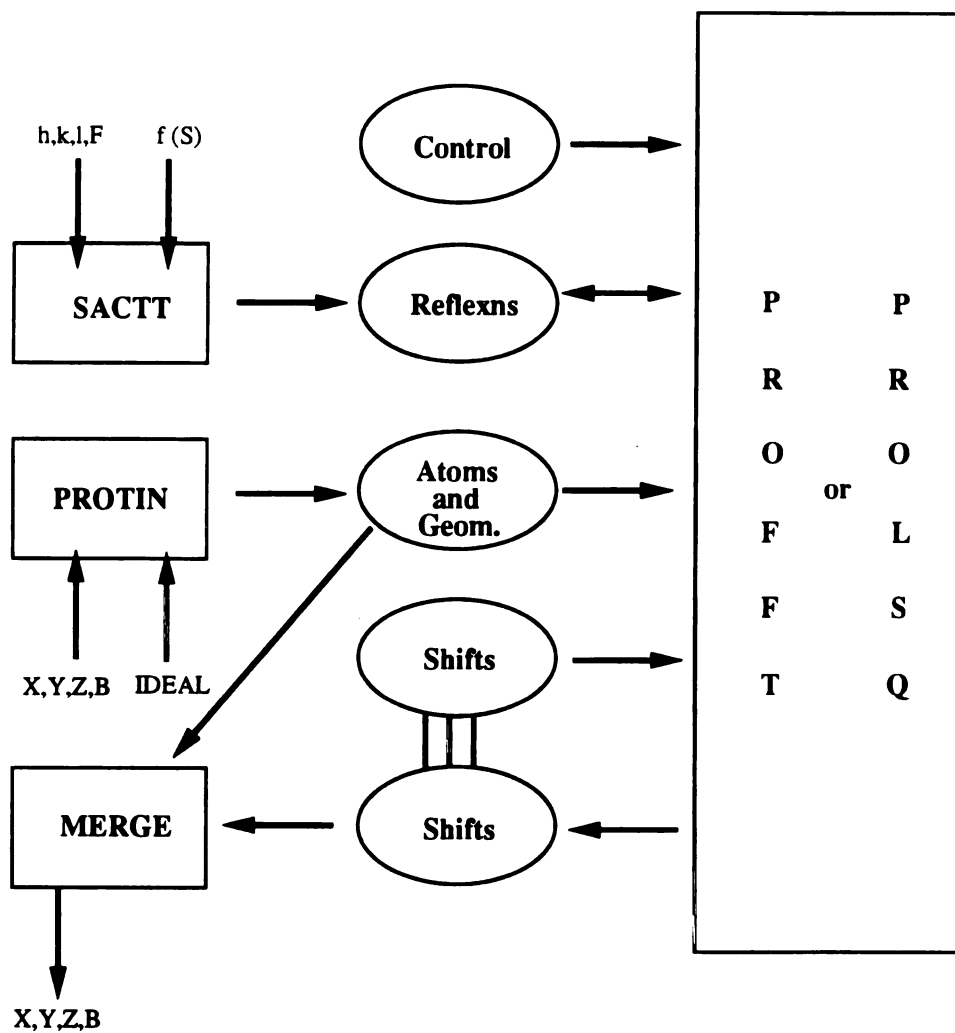


Figure 2.6. Schematic Structure of "PROFFT" and "PROLSQ".

67

67

67

67

67

67

67

67

67

67

67

67

67

67

67

67

applied to the model and structure factors are weighted by using constant σ_F which is approximately half of the discrepancy between F_o and F_c . During the refinement, the variables can be changed alternatively to loose or tight restraints which depend on the agreement of the geometric variables and structure factors. Water molecules are added gradually. Peaks are selected as water for structure factor calculations only if they are within 2.5–4.0 Å of the protein or another water molecule. New water molecules included into the refinement are initially assigned an occupancy of 0.75 and an overall B obtained from the last refinement stage. In the last stages of the refinement, the remainder of the data are assigned in seven shells of $\sin\theta/\lambda$ based on $\langle ||F|_o - |F|_c| \rangle / 2$ of the range:

$$\sigma_f = \sigma_{fa} + \sigma_{fb} * (\sin \frac{\theta}{\lambda} - 1/6) \quad (2.12)$$

In this way, the weights of low and high resolution data can be controlled by adjusting the coefficient, σ_{fa} , and slope, σ_{fb} . Lastly when the cyclic refinement is no longer giving new information and the R value converges to a reasonable value, refinement is terminated and the final model obtained.

Cl

St

H

3.1

Plas

lar

Plas

nerf

Neu

gg

for

and

St

B

Se

by

Co

the

CHAPTER 3

Structure of the ACA Complex of Human Plasminogen Kringle 4

3.1 Introduction

Plasminogen (Figure 3.1) is a glycoprotein produced by the liver with a molecular weight of about 92,000 daltons and carbohydrate content about 2%. A single plasminogen chain has 790 residues with Glu and Asp residues on the amino- and carboxyl- terminals. Sometimes a plasminogen variant that results from autolytic cleavage at the Lys76–Lys77 site, also exists with lysine as the amino- terminal residue [33] This “Lys–plasminogen” is recognized as a degradation form of the native Glu-form and is more readily activated by plasminogen activators. On the molecular level, activation involves cleavage of an Arg560–Val561 bond of the proenzyme to yield the proteolytically active plasmin which has a disulfide-linked two-chain structure. The B or light chain from Val561 to Asn790 consists mainly of the protease domain which is a trypsin-like serine protease. The A or heavy chain from Glu1 to Arg560 is formed by a tandem array of five kringles. These kringle domains are involved mainly in the interactions of PG with substrate and regulatory molecules. There are 48 cysteine residues in the plasminogen molecule, forming 24 disulfide bridges, two of which link

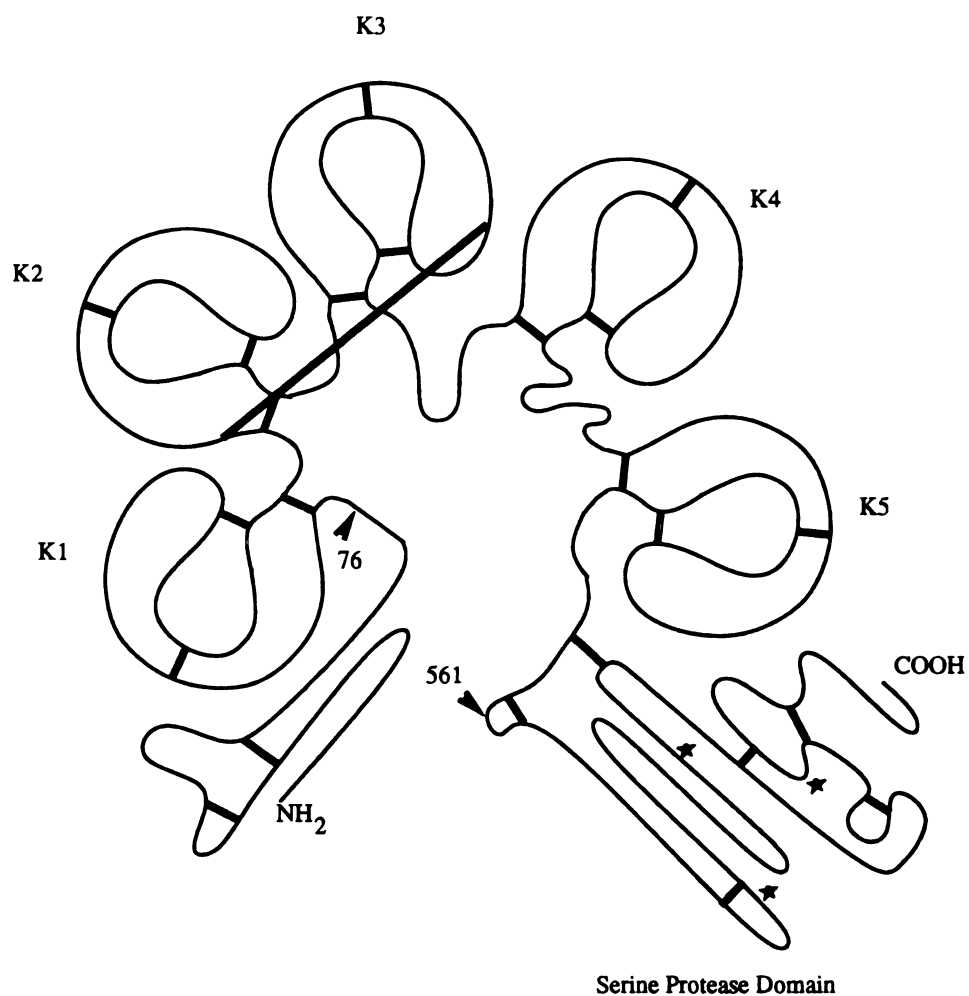


Figure 3.1. The chain structure of plasminogen. Kringle domains are labeled K1 – K5. Disulfide bridges are shown by bold lines. Asterisks denote active site residues in the serine protease domain. Arrows indicate the cleavage sites.

the A- and B- chains (Figure 3.1).

The first and fourth kringles can be obtained from the heavy chain by controlled proteolysis and affinity chromatography on a lysine-conjugated support gel [7, 9]. The experiments of the binding of 6-aminohexanoic acid to K1 and K4 have been carried out [9] and showed that K1 has the higher affinity for lysine binding ($K_a=60\text{mM}^{-1}$) than does K4 ($K_a=28\text{mM}^{-1}$). According to this observation, the first kringle is probably the main contributor to fibrin binding. A weak lysine binding site is also carried by K5 [34] which prefers ligands not having a free carboxylate function and therefore may interact with lysine side chains of proteins such as benzylamine and hexylamine. It has been proposed that K5 may play a key role in binding internal lysyl side chains of intact fibrin to the plasmin catalytic site, which cleaves Lys-X bonds of fibrin generating a number of C-terminal lysine residues.

Chemical modifications of K4 have been carried out in order to identify residues essential for ligand binding. Modification of Asp57 and Arg71 (Figure 3.2) with 1,2-cyclohexanedione and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide [17] showed the loss of lysine-Sepharose affinity. The importance of these two residues for ligand binding is related to the electrostatic interaction of the charged groups with ligands.

In addition to the polar interactions between the kringles and the lysine analogs, hydrophobic interactions between the ligand methylene groups and aromatic amino acids of kringles have also been studied. Reaction of Trp72 with dimethyl(2-hydroxy-5-nitrobenzyl)sulfonium bromide [35] destroyed lysine-Sepharose affinity. From NMR NOE experiments [36, 37], it has been found that the side chains of Trp62, Phe64 and Trp72 are perturbed most by ligand presence and that Trp25, His31, His33, Tyr41 and Tyr74 are also affected, but to a lesser extent. From the above observations, residues Asp57, Trp62, Phe64, Arg71, and Trp72 might be located on the lysine binding surface, where they interact with ligand directly and strongly [17, 35].

The lysine binding subsite structure was first modeled by Tulinsky *et al.* [38]

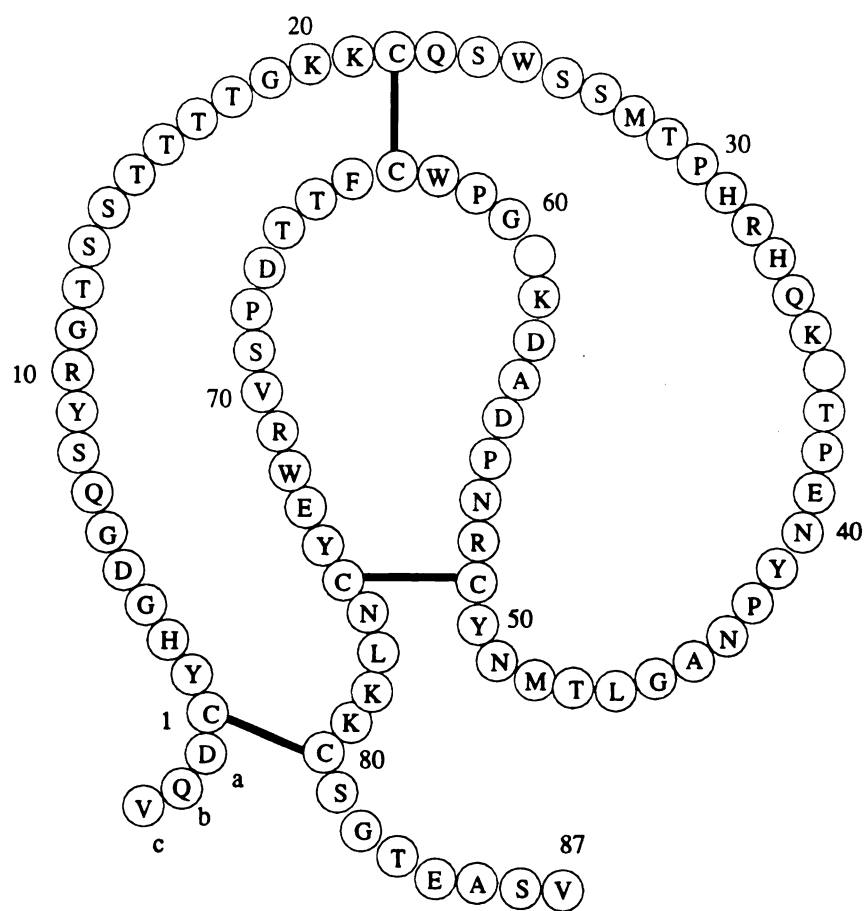


Figure 3.2. The primary structure of plasminogen K4. Interkringle residues are a-c and 81-87. (see appendix for abbreviations)

based on the structure of PTF1, which additionally implied that Asp55 might be important to the lysine binding (Figure 3.3). This was then confirmed with the crystallographic structure of PGK4 [23, 39]. However, a cofacial intermolecular interaction between the binding sites of neighboring molecules excluded access of the site to the small ligands like ACA by diffusion into crystals, thus precluding the formation of the K4-ACA complex. This led us to crystallize PGK4 in the presence of ACA ligand [40]. The crystallographic structure of PGK4-ACA complex not only establishes the nature of ligand binding but provides a good model for fibrin binding.

3.2 Materials and Methods

A. Crystallization

The human plasminogen kringle 4 was generated from plasminogen by elastase digestion and affinity chromatography on lysine-Sepharose [7]. The product is a heterogeneous mixture of approximately 70% C-terminal Ala85 and about 30% C-terminal Ala85-Ser86-Val87. The PGK4 used for crystallization was kindly provided by Dr. M. Llinas in the form of a lyophilized powder, which was stored at freezer temperatures in a dry environment.

Soaking experiments failed to produce K4-ACA crystals [41]. This led us to crystallize the PGK4 complex in the presence of ACA ligand. The K4 was first dissolved in distilled water at a protein concentration of approximately 10 mg/ml, and ACA at 25 mg/ml in water was added. This solution was then used in growing crystals. Crystals of K4-ACA were grown by the sitting drop method using 30% PEG 8000, 0.12 M ammonium sulfate, with 1.4% dimethyl formamide (DMF) as an additive at pH 6.0 in the presence of 25mM ACA. Although the conditions were practically the same as those which produce orthorhombic apo-K4 crystals [42], K4-ACA crystallized in a different crystal form. The unit cell parameters and other

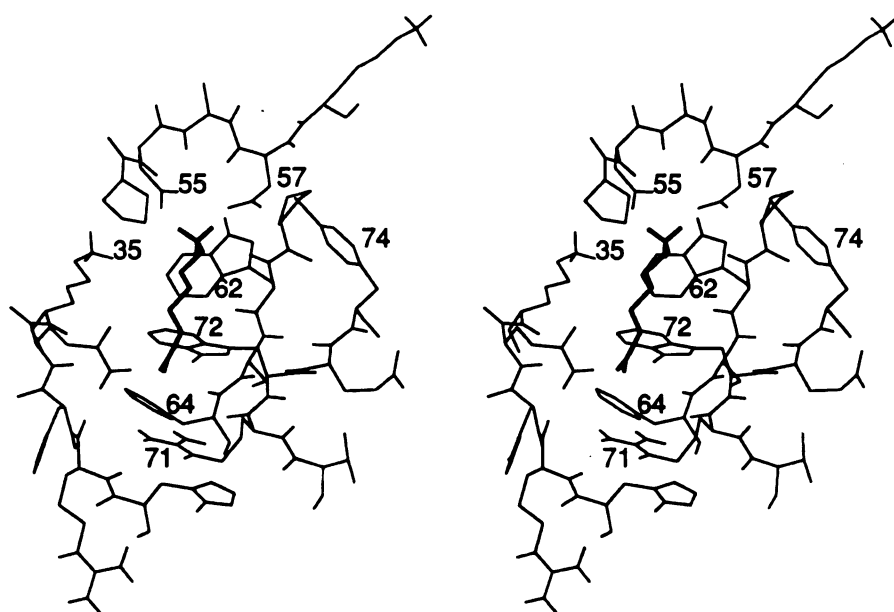


Figure 3.3. Stereoview of modeled lysine binding site of plasminogen kringle 4.ACA in bold. Hydrogen atoms are put at ideal positions [38].

crystal data are shown in Table 3.1. Crystals appeared within a week after seeding with monoclinic apo-K4 crystals. Most of the crystals had a plate-shaped form and tended to grow together. Single and separate crystals could be obtained occasionally as shown in Figure 3.4.

Table 3.1. Summary of crystal data of monoclinic K4-ACA

space group	P2 ₁
crystal system	monoclinic
molecules/asymmetric unit	1
molecular weight (dalton)	10000
a (Å)	42.21
b (Å)	35.46
c (Å)	25.43
β	102.95°
$V_m (\frac{\text{\AA}^3}{\text{dalton}})$	1.89
solvent fraction (%)	34.5

B. Data Collection

The crystal used for intensity data collection had dimensions of $0.8 \times 0.4 \times 0.2$ mm. The crystal was mounted in a siliconized capillary of diameter 1.0 mm with the b axis parallel to the length of the capillary (Figure 3.5). Three-dimensional intensity data were measured to 2.25 Å resolution using a Wyckoff ω -step procedure [24] with a Nicolet P3/F diffractometer at 2.5 kW (50kV, 50mA). Axial intensity distributions of the crystal are displayed in Figure 3.6. From the ω -profile of a reflection (5, 5, 5) (Figure 3.7), a scan range of 0.2° and an offset value of 0.35° from the peak center for background measurement were chosen. A scan speed of 0.4°/min was used for measurements, with a background to scan ratio of 0.1. Each measurement was scanned for seven steps and the highest five counts were summed as the integrated

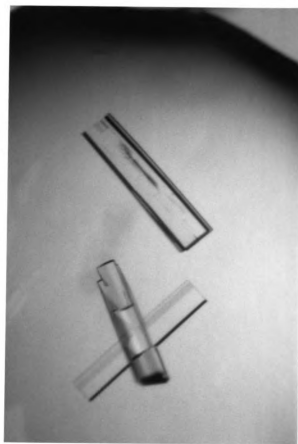


Figure 3.4. Photograph of K4-ACA crystal. The size of the single crystal is approximately $1.2 \times 0.3 \times 0.1$ mm.

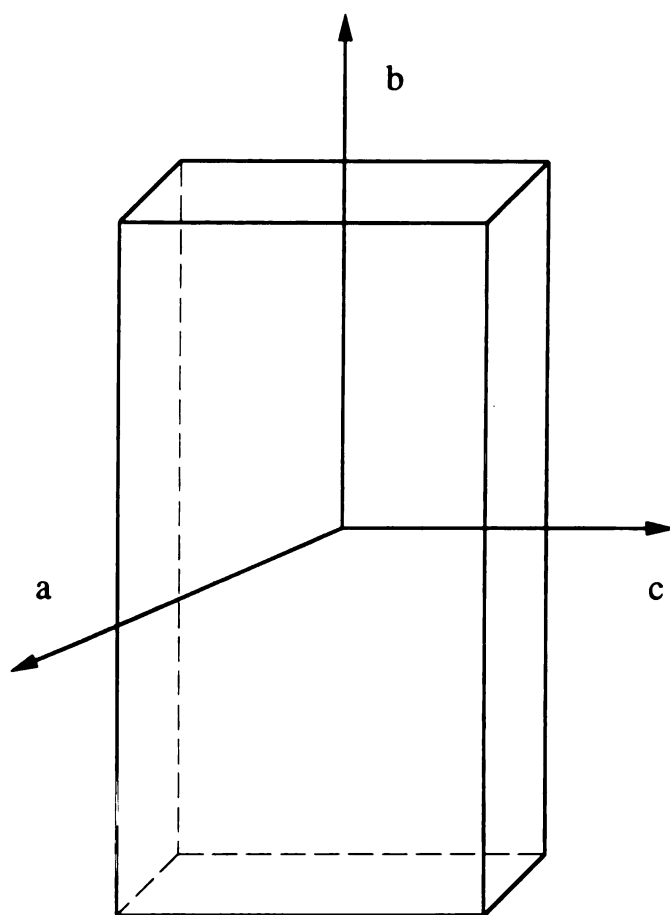


Figure 3.5. Schematic diagram of K4-ACA crystal morphology with respect to crystallographic axes.

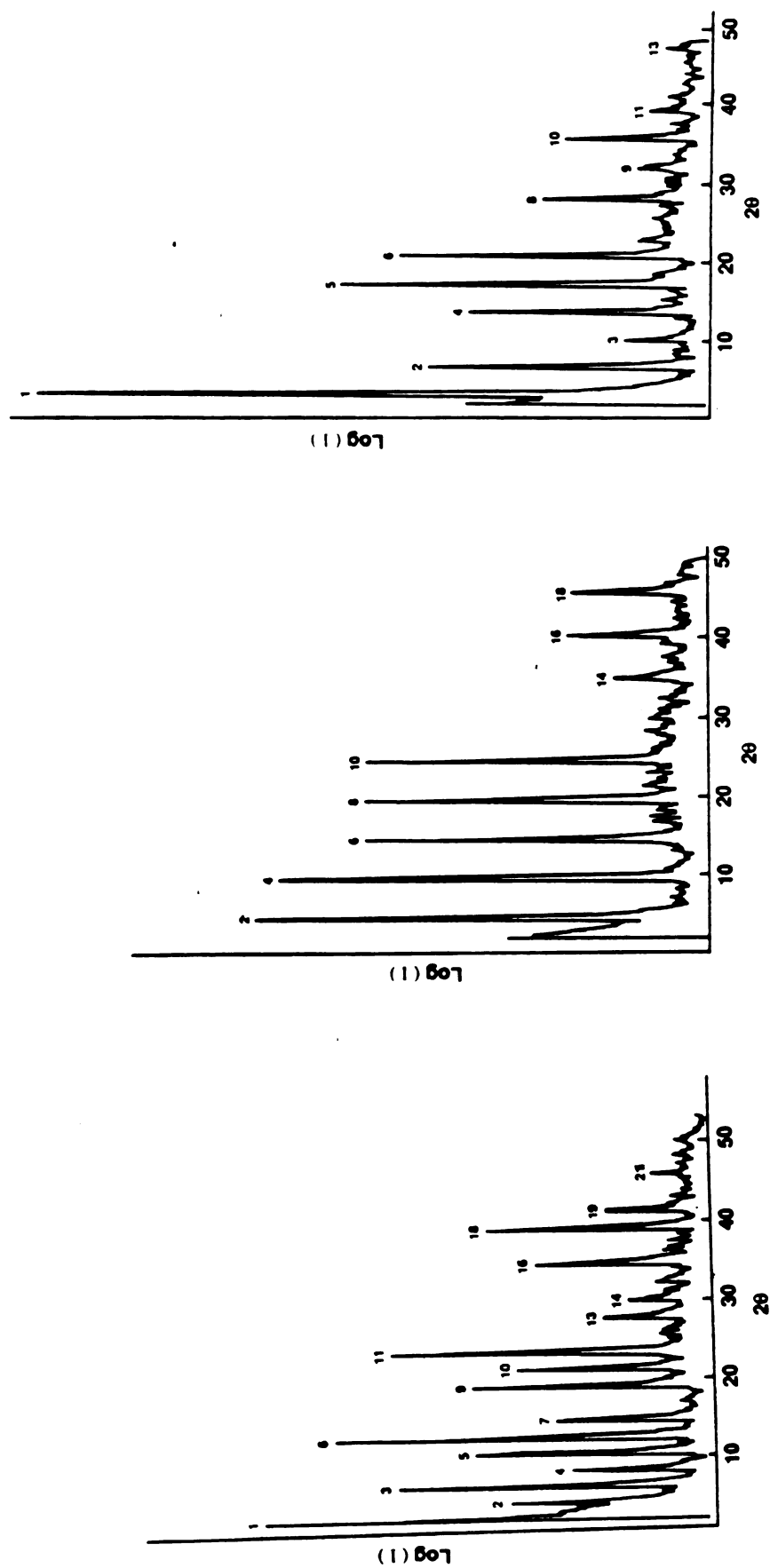


Figure 3.6. Axial intensity distributions of K4-ACA crystal.

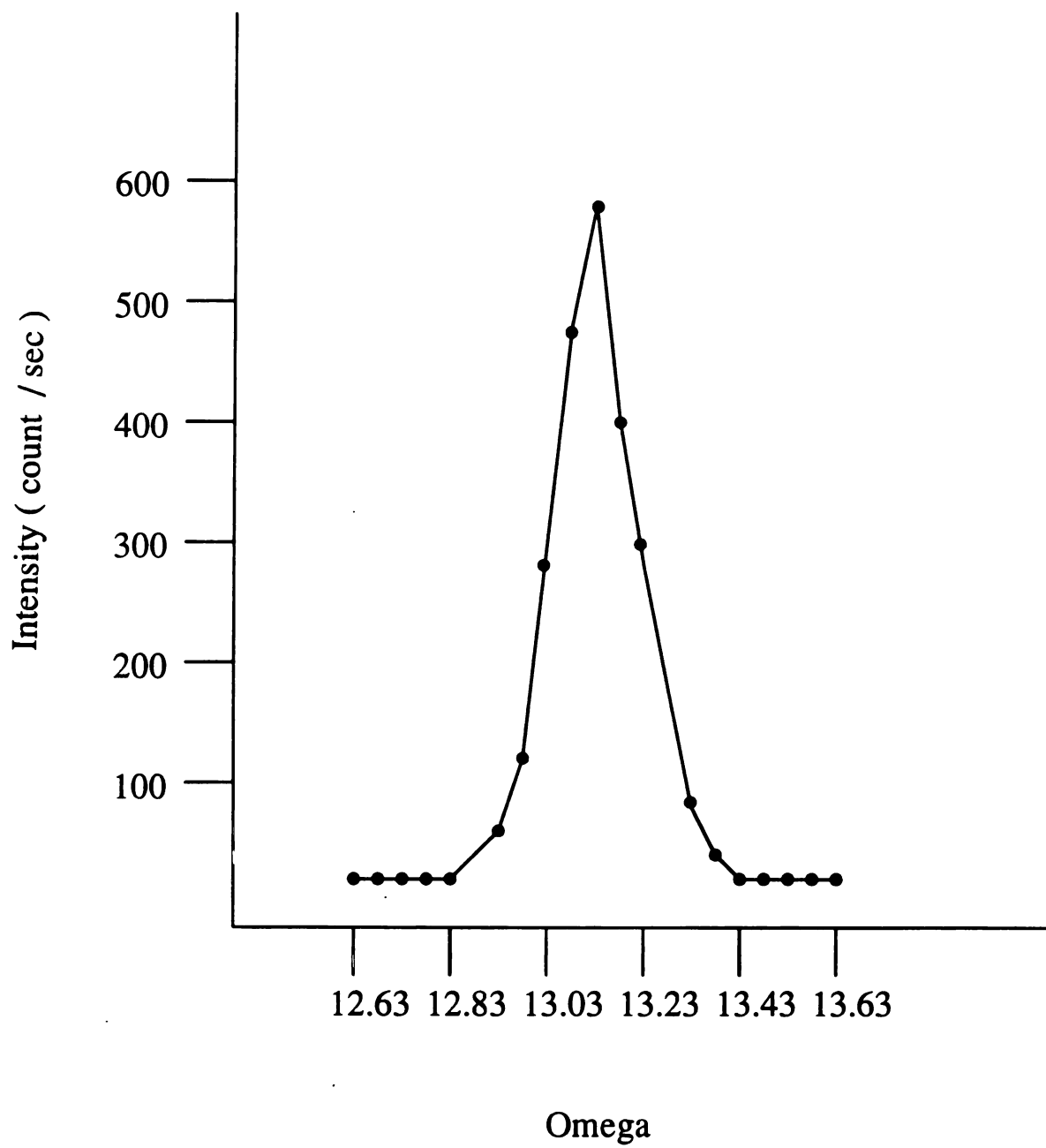


Figure 3.7. The ω -profile of reflection (5, 5, 5) taken before data collection.

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

intensity.

Before 3-D intensity data were collected, an absorption correction was determined for every 10° in ϕ by measuring intensities for reflections (0, 2, 0) and (0, 0, 10). Since both of them showed the same trend, only reflection (0, 2, 0) was used in absorption correction. A plot of $I_{\max}/I(\phi)$ versus ϕ is shown in Figure 3.8; the 180° range between $\phi = 90^\circ$ to 270° was chosen for data collection. The intensity data were collected in three 2θ shells of $(32^\circ-40^\circ)$, $(26^\circ-32.2^\circ)$, $(2^\circ-26.2^\circ)$, in the order of highest to lowest resolution range.

In order to make an intensity decay correction, three check reflections were measured after every 100 reflections, a 2-D $h0l$ data set from 2° to 15° was collected before and after data collection, and 12 reflections having a 2θ range between 18° to 25° were measured periodically during data collection.

C. Data Reduction

The background readings were first averaged in 2θ shells as shown in Table 3.2. Background was also averaged with respect to ϕ angle and showed no angular variation (data are not shown). From Table 3.2 it can be seen that the average corrected negative intensity ($\langle I_{\text{neg}} \rangle$) is approximately -3. As mentioned in chapter 2, $I_{\min}$ was chosen to be 6. Of a total of 3381 reflections, 3198 which were greater than twice $I_{\min}$ were considered observed.

In order to get the decay correction factor, DEC, S values in each 2θ range were calculated, including the 3 check reflections, the 2-D data set, the 12 reflections (in 3 ranges), and the data are shown in Table 3.3. According to this table, a decay slope, s , of $1.0 \times 10^{-4} \text{ hr}^{-1}$ was used for the decay correction.

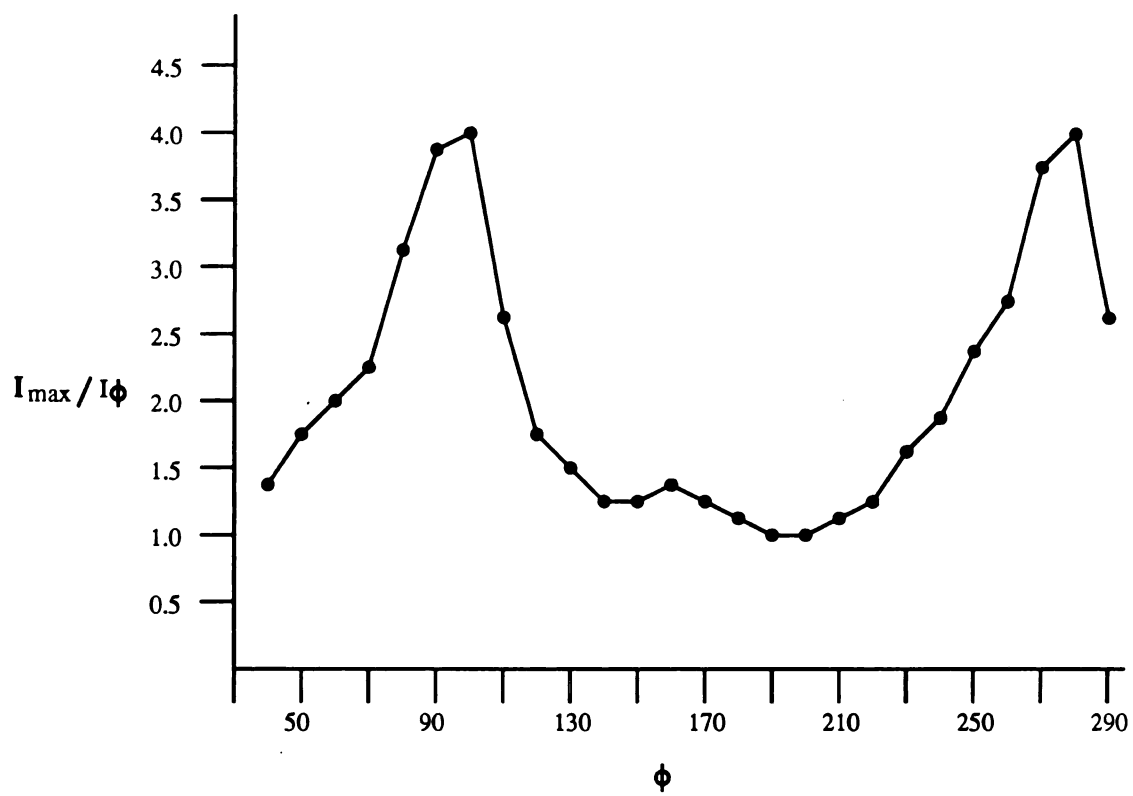


Figure 3.8. Absorption correction curve of (0, 2, 0) reflection.

Table 3.2. Averaging background in 2θ shells.

2θ	$\langle \text{BG} \rangle$	# reflections	$\langle I_{\text{neg}} \rangle$
5.0–18.0	1.4	391	-3.3
18.0–22.0	1.6	307	-2.0
22.0–25.0	1.8	302	-1.5
25.0–28.0	1.8	399	-1.5
28.0–30.5	1.6	381	-2.0
30.5–32.5	1.5	396	-1.7
32.5–34.5	1.4	385	-1.1
34.5–36.0	1.3	322	-1.5
36.0–37.5	1.2	345	-1.7
37.5–38.5	1.1	229	-1.5
38.5–40.0	1.1	392	-1.3

Table 3.3. Decay rate, S , calculated in each 2θ range.

2θ	9.9	17.9	23.0	23.1	23.7	24.7	24.9
$S \times 10^{-3}$	1.5	4.5	2.9	2.6	2.7	1.8	2.2

D. Molecular Replacement

The K4-ACA structure was solved by molecular replacement rotation-translation methods by utilizing the refined PGK4 structure as a search model [23]. The rotation search was performed in Patterson space with the SEARCH routine in the program PROTEIN [28]. A triclinic cell of dimensions $80 \times 80 \times 80$ Å using data from 8.0 to 3.0 Å resolution and an overall thermal parameter of 20 Å² were used in the model structure factor calculation. A large triclinic cell was chosen to separate the intramolecular vectors which depend solely on the rotational orientation. The low resolution range cutoff is to avoid contributions from solvent molecules; the high angle cutoff is to avoid detailed structural information. Both model and unknown K4-ACA Patterson vectors of length greater than 3.0 Å and less than 15.0 Å were selected, since rotation function calculation uses only intramolecular vectors. For kringles having dimensions of approximately $15 \times 30 \times 30$ Å a resolution range from 5.0–20.0 Å is usually used in rotation calculation. The lower limits value is selected to avoid the Patterson function origin. The higher limit was chosen based on the size of the kringle to exclude possible intermolecular vectors. A set of the 1500 highest Patterson peaks was used in the rotation calculation.

In the rotation search calculation, an initial search model was rotated over the range of θ_1 , 0–360°, θ_2 , 0–180°, and θ_3 , 0–180° in 5° increments. The highest peak was found at the Euler angle (80.0°, 35.0°, 125.0°) at 7.34σ above the mean. The refinement of this solution was carried out in 2°, 1° and 0.3° increments, and the final solution had peak height 7.6σ above the mean at (84.5°, 33.4°, 121.3°).

The position of the model in the unit cell after rotation was found using the program BRUTE [29] by a translation search first in 0.5, and then 0.1 Å increments with data from 8.0–3.0 Å resolution. Since the origin along the b-axis is arbitrary in the monoclinic system, the translation search was applied in the XZ plane. To

get a convincing solution, data in the 8.0–2.8 Å resolution range were also used for the translation search, which showed the same result as that in 8.0–3.0 Å resolution range. The final refined peak had a correlation coefficient of 0.43 (about 16σ above the mean).

To verify the rotation matrix and translation vectors, the packing of the molecule was inspected on an Evans and Sutherland PS390 interactive computer graphics system equipped with FRODO [32] software. An electron density map based on this rotation–translation position clearly revealed the K4 structure and, in addition, new electron density extending between the side chains of Asp57 and Arg71.

E. Structure Refinement

The starting PGK4 model of K4–ACA was refitted using interactive computer graphics. The structure was then refined employing the restrained least-squares method implemented in the program PROFFT [31] with intermittent model building performed on an Evans and Sutherland PS390 interactive stereographics system. The refinement proceeded in stages, each of which was followed by model building using $(2|F_o| - |F_c|)$ and $(|F_o| - |F_c|)$ maps and examination of the Ramachandran plot. The initial R value started at 0.43 with an overall thermal parameter of 16 Å^2 ; R decreased to 0.26 after the first stage (2.8 Å resolution). As phases improved, higher resolution data were added to 2.5 Å resolution. The unaccounted for density between Asp57 and Arg71 now appeared convincingly to be due to ACA. Therefore, an ACA was fitted into the density and included in the refinement. Water molecules were also added at 2.5 Å resolution. In the last stages of refinement, the remainder of the data (to 2.25 Å) were included, and reflection weights were assigned in seven shells of $\sin \theta / \lambda$ based on $\langle ||F_o| - |F_c|| \rangle / 2$ of the range. The final reflection weights and R values in each range are given in Table 3.4, and a summary of refinement parameters is listed in Table 3.5. The final K4–ACA structure has a crystallographic R value of 0.148 for

Table 3.4. Weights of Reflections and R Values of the Final Refinement

$D_{\min}(\text{\AA})$	no. reflections	$\sigma(F)^a$	$< F_o - F_c >$	R value	
				shell	sphere
4.00	465	11.2	25.6	0.153	0.153
3.30	465	9.9	19.2	0.124	0.139
2.90	493	9.1	16.0	0.150	0.142
2.60	555	8.4	14.1	0.157	0.145
2.45	375	7.8	13.5	0.171	0.148
2.35	288	7.5	11.5	0.151	0.148
2.25	352	7.2	11.00	0.148	0.148
<i>a.</i> $\sigma(F) = 18.0 - 70.0 [(\sin \theta/\lambda) - (1/6)]$.					

Table 3.5. Summary of Final Least-Squares Parameters and Deviations

	target (σ)	rms (Δ)
Distances ($\text{\AA}$)		
Bond Distance	0.015	0.017
Angle Distance	0.025	0.042
Planar 1,-4 distance	0.035	0.050
Non-bonded distances ($\text{\AA}$)		
Single torsion	0.50	0.20
Multiple torsion	0.50	0.32
Possible H-bond	0.50	0.32
Torsion angles (deg)		
Planar	3	4
Staggered	15	23
Orthonormal	20	16
Plane groups ($\text{\AA}$)	0.02	0.02
Chiral centers ($\text{\AA}^3$)	0.13	0.23
Thermal restraints ($\text{\AA}^2$)		
Main chain bond	1.5	1.8
Main chain angle	2.0	2.6
Side chain bond	2.5	2.3
Side chain angle	2.5	3.6

2993 reflections between 7.0–2.25 Å resolution range with 106 water molecules and an average thermal parameter of 17.9 Å². The average occupancy of the water molecules is about 0.75, and their average B value (19 Å²) is slightly higher than that of the protein (17.5 Å²).

3.3 Results and Discussion

A. General Structure

The refined K4-ACA structure extends from residue Gln-b to Cys80 along with the ACA ligand; there was no electron density for the N-terminal Val-a nor for the C-terminal interkringle pentapeptide. A similarly disordered interkringle region was found in the apo-K4 structure [23] and might be due in part to the interkringle link heterogeneity discussed in section 2. There was also only little electron density for the Thr18 and Arg32 side chains; therefore, they were refined as alanine. The distribution of main-chain torsion angles (ϕ, ψ) is shown in Figure 3.9, from which it can be seen that all the residues conform well with allowed regions except for one outlier (Met48); however, the latter fits in very well-defined electron density and has the same conformation as that observed in apo-K4 [23]. Moreover, a *cis*-proline was found at position 30, similar to the apo-K4 structure. Conversely, in K4-ACA the sulfur atom of Cys75 is localized in one position while it was distributed between two different equally occupied positions in the apo-structure giving rise to different disulfide conformations [23]. The localization in K4-ACA could be a result of ACA ligand binding or possibly due to different packing in the crystal structure. The average thermal parameter of the kringle in K4-ACA is 17.5 Å², but although the electron density of the ACA is quite good (Figure 3.10), it has an average B value which is about twice that of the kringle (34 Å²). Since the ACA is fixed in the lysine-binding site by two doubly charged, hydrogen-bonding, ion pair interactions, the

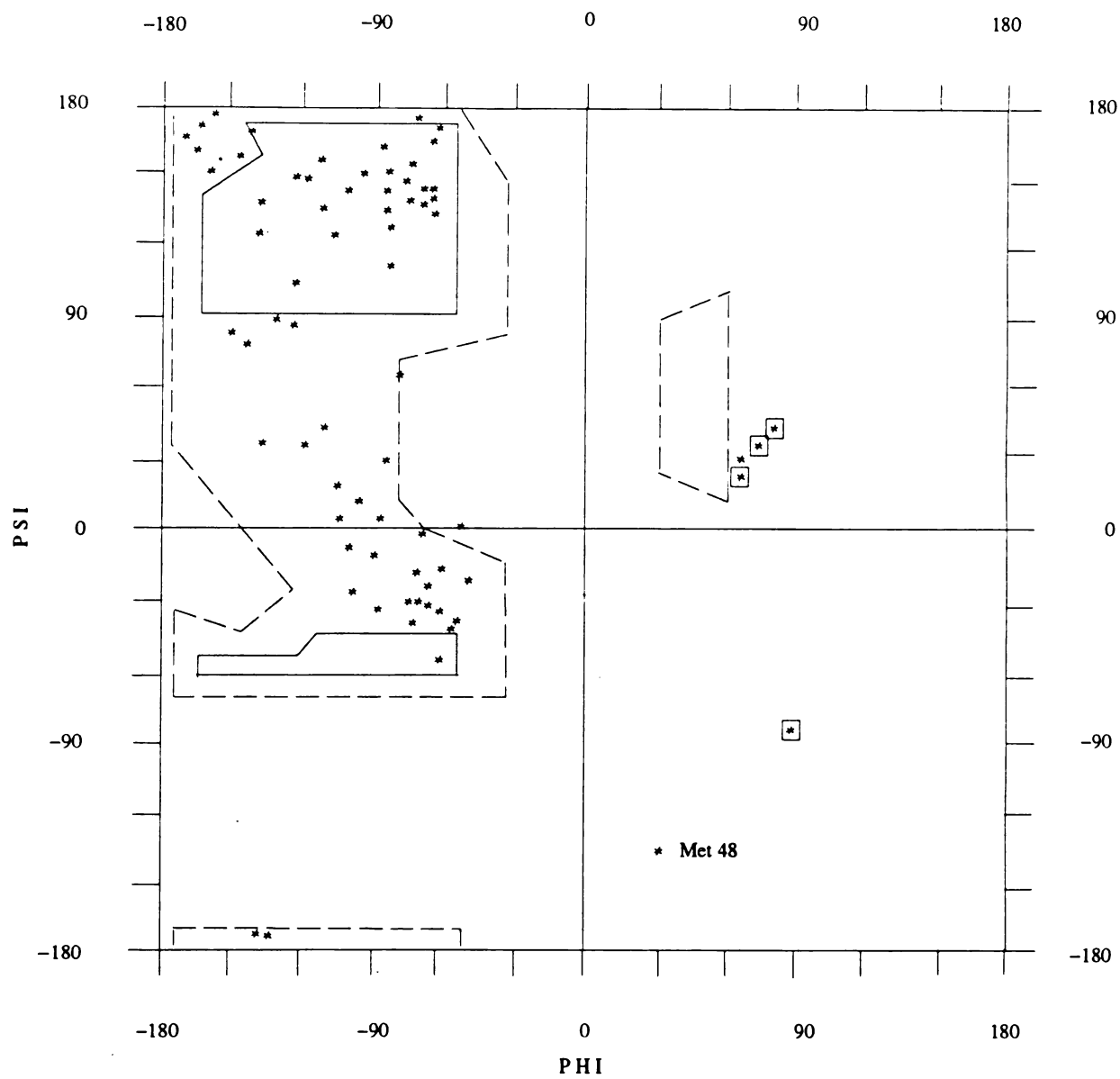


Figure 3.9. Ramachandran plot of final ϕ , ψ angles of K4-ACA. The Gly residues are boxed.

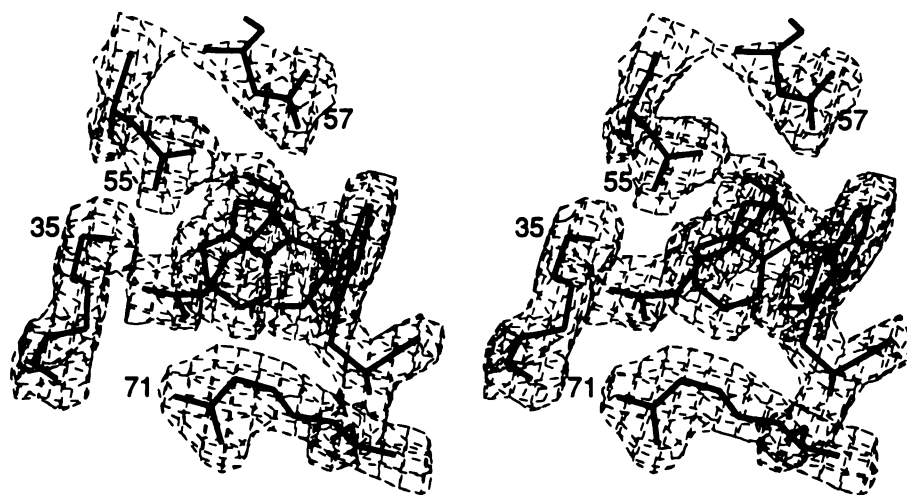


Figure 3.10. Stereoview of the electron density of the lysine-binding site of K4-ACA. The basket contour is at 1σ ; ACA is between Asp55, Asp57 and Lys35, Arg71.

higher B value of ACA must reflect a microheterogeneity of its positioning within the binding site similar to that observed for hirudin in its complex with α -thrombin [43].

Of the hydrogen-bond interactions in the K4-ACA structure, all but eight are also common to the apo-K4 structure [23]; the ones present in K4-ACA and not in apo-K4 are listed in Table 3.6. The criteria used to identify possible hydrogen bonds were (a) a donor-acceptor center distance of less than 3.05 Å and (b) a hydrogen-bond angle of greater than 120°. Of the new hydrogen bonds, that of Gln7N-Asp5OD2 is borderline, that of Tyr9OH-AspcOD1 involves an interkringle residue, which was disordered in the apo-structure, that of Lys20NZ-Glu73OE1 makes the only ion pair of K4 a hydrogen-bonded one, and the hydrogen bond between Arg71NH1 and Arg32O just satisfies the selection criteria. Also listed in Table 3.6 are two apparently important hydrogen bonds involving Asp55/Asp57 of the lysine-binding site of both structures, which will be addressed later.

B. Lysine-Binding Site

The lysine-binding site is a relatively open, elongated, shallow depression located on the kringle surface that is formed by His31-Lys35, Pro54-Lys58, Pro61-Phe64, and Arg71-Cys75. At neutral pH, there are two negatively charged residues located on one end of the binding site (Asp55 and Asp57) with carboxylate oxygens 5.2 Å from one another; there are also two positively charged residues located at the other end of the depression (Lys35 and Arg71) with quaternary amino groups 6.3 Å from one another. Thus, the binding site contains doubly charged anionic and doubly charged cationic centers. The depression, lined by aromatic rings of Trp62, Phe64, and Trp72, also provides a highly nonpolar environment between the charged centers so that the lysine-binding site approximates a dipolar surface as first suggested from modeling [38]. Thus, zwitterionic ligands such as lysine and ACA can first interact at long range in preparation of docking, ultimately being anchored by ion pair interactions.

Table 3.6. New Intramolecular Hydrogen Bonds of K4-ACA. Hydrogen atoms were **assigned** geometrically idealized positions. Donor atom is denoted (D), acceptor atom (**A**).

DONOR	ACCEPTOR	DISTANCES(Å)		ANGLES(deg)	
		D...A	H...A	DHA	CAH
Gln7 N	Asp5 OD2	3.08	2.22	145	156
Tyr9 OH	Aspc OD1	2.63	1.69	156	104
Arg10 NH2	Asn43 O	2.89	2.30	118	157
Lys20 NZ	Glu73 OE1	2.96	2.17	132	152
Ser27 OG	Thr29 O	2.74	1.99	134	123
Thr37 N	Gln29 O	2.89	2.02	142	101
Thr47 N	Gly45 O	2.71	1.86	137	103
Arg71 NH1	Arg32 O	3.01	2.17	143	139
Trp62 NH1	Asp55 OD2	2.99	2.43	116	120
Tyr74 OH	Asp57 OD2	2.59	1.71	152	105

The close van der Waals contacts between the methylene carbons of the ligand and the aromatic residues between the charge centers additionally assist the binding.

C. Kringle 4-ACA Interaction

A zwitterionic ACA molecule in an extended conformation lies between the doubly charged anionic and cationic centers of K4 formed by Asp55/Asp57 and Lys35/Arg71 (pH of crystals is 6.0), which also makes four or five hydrogen bonds with these residues (Table 3.7) and interacts with the lipophilic core formed by Trp62, Phe64, and Trp72 through the methylenes between the zwitterion charges (Figures 3.10, 3.11). The hydrogen bond to Asp55 is questionable since its donor-acceptor angle is so small; however, its ion pair interaction is undeniable and even appears to be stronger than that of Asp57. The oxygen atoms of the carboxylate groups of Asp55 (OD2) and Asp57 (OD2) are hydrogen bonded back to the kringle (Table 3.6), which might aid in aligning and anchoring the side chains for ligand binding. The Asp57 and Arg71 residues were originally implicated in ligand binding through modification of Asp57 with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide and Arg71 with 1,2-cyclohexanedione [17], while the participation of Asp55 was suggested on the basis of computer modeling [38]. However, from Figure 3.10 and Figure 3.11 and Table 3.7, it is clear that Lys35 also participates as an integral member of the principal binding residues as inferred from the apo-K4 structure, in agreement with modification studies which showed that blocking Lys35 decreases the affinity of K4 for lysine-Sepharose [44]. Computer modeling also suggests a doubly charged cationic center in the binding site of PGK1 involves Arg34 and Arg71 [38]. In the case of K4, the guide coordinates for modeling Lys35 had to be based on Ile35 of PTF1, which only extended to CB, so that the lysyl side chain was simply modeled in an extended energy-minimized conformation. This is clearly not the case as Lys35 retreats toward Arg71 in the binding site and gives rise to a doubly charged center in both

the apo- and K4-ACA structures. The new finding thus places the K4 binding site on a comparative level to the modeled PGK1 site with respect to a doubly charged positive center.

Table 3.7. Hydrogen Bonds Between K4 and ACA

DONOR	ACCEPTOR	DISTANCES(Å)		ANGLES(deg)	
		D...A	H...A	DHA	CAH
Lys35 NZ	ACA O1	2.72	1.99	127	148
Arg71 NE	ACA O2	2.71	1.83	145	147
Arg71 NH2	ACA O2	2.89	2.09	134	129
ACA NZ	Asp57 OD1	2.84	2.28	113	107
ACA NZ	Asp55 OD2	2.12	1.97	84	112

D. Comparison of K4-ACA and Apo-K4 Structures

The structures PGK4 and PTK1, which have already been compared elsewhere [39], show that the lysine-binding site is approximated fairly well by PTK1, but with some notable exceptions. These render PTK1 to be a nonbonding kringel. The structure of PGK4 has additionally been compared with the K4-ACA structure by the optimal superpositioning of CA, C, and N atoms. The rms differences between the two structures are listed in Table 3.8. The agreement between the two structures is

Table 3.8. RMS difference between the K4-ACA and Apo-K4

	Δ (Å)	# atom
All protein atoms	0.57	602
Main chain	0.44	237
Carbonyl oxygens	0.60	79
Side chains	0.66	286
Sulfurs (Cys, Met)	0.44	8
α -Carbon	0.43	79

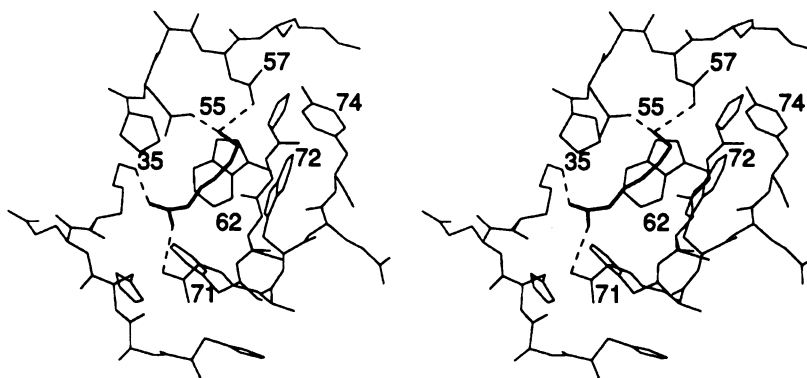


Figure 3.11. Stereoview of the lysine-binding site of K4-ACA. ACA is shown in bold; hydrogen bonds are dashed.

good for the main-chain atoms and superb ($0.27 \text{ \AA}$) after removing about 15% of the atoms having deviations greater than 1σ . The differences between the apo and the complexed structure in the lysine-binding site are listed in Table 3.9 and a stereoview of the superposition of the binding site regions is shown in Figure 3.12. The average rms difference in the main-chain position in the site is only $0.25 \text{ \AA}$ with only one difference as large as $0.50 \text{ \AA}$, while the average of the side groups is $0.69 \text{ \AA}$. The largest deviations in the binding site are due to crystal packing interactions. In the K4-ACA structure, Gln34 forms a hydrogen bond with the carbonyl oxygen of Trp72 of a neighboring molecule, giving rise to a rms deviation of $1.12 \text{ \AA}$. Other significant deviations occur with Asp55 and Asp57 of the anionic charge center (0.75 and $1.23 \text{ \AA}$, respectively). These are the consequence of a cofacial kringle-kringle interaction of the binding site regions of two neighbor molecules in the apo-K4 structure. The Asp55/Asp57 of one molecule form two ion pairs with Arg32/Arg71 of neighboring molecules; thus, Asp55 and Asp57 in apo-K4 point somewhat away from the lipophilic binding core, which leads to relatively large deviations from the K4-ACA structure. It is also noteworthy that Arg32 appears to be disordered in K4-ACA structure where it cannot make a similar intermolecular interaction. In the case of Lys58, it is involved in a complex intermolecular interaction in the apo-structure with a sulfate ion of a neighboring molecule. Interestingly, the position of the sulfate anion in K4 is close to the cationic center in K4-ACA (Figure 3.12). Lastly, the side chain of Lys35 in apo-K4 forms a hydrogen-bonding ion pair with the sulfate thus making the side-chain conformation a little different from that of the cationic center in K4-ACA (Table 3.9). The remainder of the residues in the lysine-binding site are practically identical in the two structures. Thus, from all appearances, the lysine-binding pocket is preformed in the kringle structure, and ACA binding takes place without requiring any conformational changes of the host.

Table 3.9. Differences in lysine-binding sites between K4-ACA and Apo-K4

residue	rms Δ (Å) (main chain)	rms Δ (Å) (side chain)
His31	0.21	0.28
Arg32	0.17	0.50
His33	0.16	0.19
Gln34	0.23	1.12
Lys35	0.26	0.70
Pro54	0.36	0.23
Asp55	0.19	0.75
Ala56	0.50	0.35
Asp57	0.25	1.23
Lys58	0.25	1.27
Pro61	0.10	0.54
Trp62	0.26	0.20
Cys63	0.19	0.35
Phe64	0.15	0.12
Arg71	0.10	0.21
Trp72	0.23	0.44
Glu73	0.19	0.65
Tyr74	0.10	0.53
Cys75	0.20	0.29
av rms	0.25	0.69

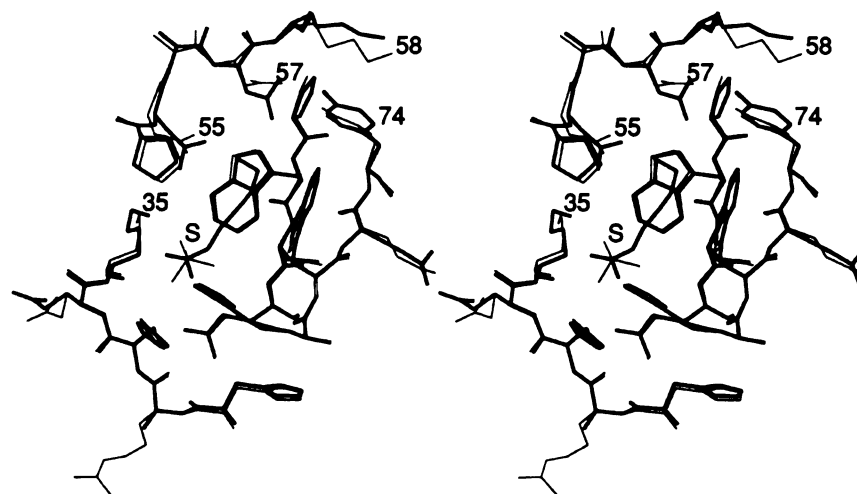


Figure 3.12. Stereoview of the comparison of the lysine-binding site of K4-ACA and apo-K4. K4-ACA is in bold; the sulfate position in apo-K4 is designated by S.

E. Ligand Interaction and NMR Results

The indole side chains of Trp62 and Trp72 form end-to-face contacts with the rings of Phe64 and Tyr74, respectively (Figure 3.11); such aromatic clustering is common in proteins and provides enhanced stability [45] to the hydrophobic depression. NMR NOE experiments [36, 37] found that the side chains of Trp62, Phe64 and Trp72 are perturbed most by ligand presence and that Trp25, His31, His33, Tyr41 and Tyr74 are also affected, but to a lesser extent, in good agreement with the crystal structure. Most of these aromatic residues are near the surface in the binding site region within which the ligand lies in an extended conformation (Figure 3.10, 3.11). The ring of Phe64 is as close as 3.2 Å from CA of ACA and approximately 3.9 Å from the end of the Trp62 indole ring; thus, it appears that the phenylalanyl ring can be affected through the aromatic stacking interaction with the Trp62 side chain or possibly by a substitution at the CA position of the ligand. This conforms with Phe64 having large ligand-induced chemical shifts with such ligands [37]. Although somewhat removed from the binding center, the tyrosyl ring of Tyr74 is positioned 3.7 Å from the end of the indole of Trp72 while its hydroxyl group appears to be making an important hydrogen bond orienting the carboxylate group of Asp57. The Tyr74 may sense ligand-binding effects indirectly from the aromatic interaction with Trp72 and/or the hydrogen-bonding interaction with Asp57. The His33 residue, which is at the perimeter of the binding site, has its imidazole ring stacked parallel to the phenyl ring of Phe64 and within 4.5 Å of it. The binding of bulkier ligands such as AMCHA (Figure 1.3) may have a greater influence on this imidazole ring. In contrast to His33, His31 is located away from the binding core and in agreement with NMR observations which show that it only has a small chemical shift upon ligand binding [36, 37]. The indole ring of Trp25 is in the next layer below the surface, but its end is oriented about 4.0 Å from the face of the indole of Trp62. It also displays only minor ligand effects

which most likely transmit through the aromatic-aromatic interaction with Trp62; however, it too may experience greater effects from bulkier ligands. Lastly, the side chain of Tyr41 is located distantly from the center of the binding site. Therefore, it can only sense the presence of the ligand through secondary effects transmitted through residues in the immediate vicinity of the ligand.

CHAPTER 4

Structure of Human Plasminogen Kringle 1

4.1 Introduction

Human plasminogen kringle 1 (PGK1) and kringle 4 (PGK4) both carry lysine binding sites with the former having a higher affinity for lysine binding. It has been shown by NMR that PGK1 and PGK4 have a similar three-dimensional structure [46, 47], which is compact and globular and built around a core of hydrophobic, aromatic amino acid side chains [46].

The secondary structure of PGK1 [48] has been predicted based on the statistical methodology of Chou and Fasman. Of the 80 residues (Cys1 – Cys80) (Figure 4.1) of the PGK1 region and the PGK1-PGK2 interconnecting peptide (Glu81-Glu83), approximately 4% has been predicted as α -helix, 30% of β -sheet, 53% of β -turn and 13% of random coil. According to the results of Overhauser experiments, acid-base titration and two-dimensional chemical shift correlated spectroscopy, it has been shown that Leu46 is surrounded by a cluster of interactive aromatic side chains, which include Trp25, Phe35, His41, Tyr50, Trp62, and Tyr64 [47]. The same results were also found in the case of PGK4, in which His41 and Tyr64 are replaced by Tyr41

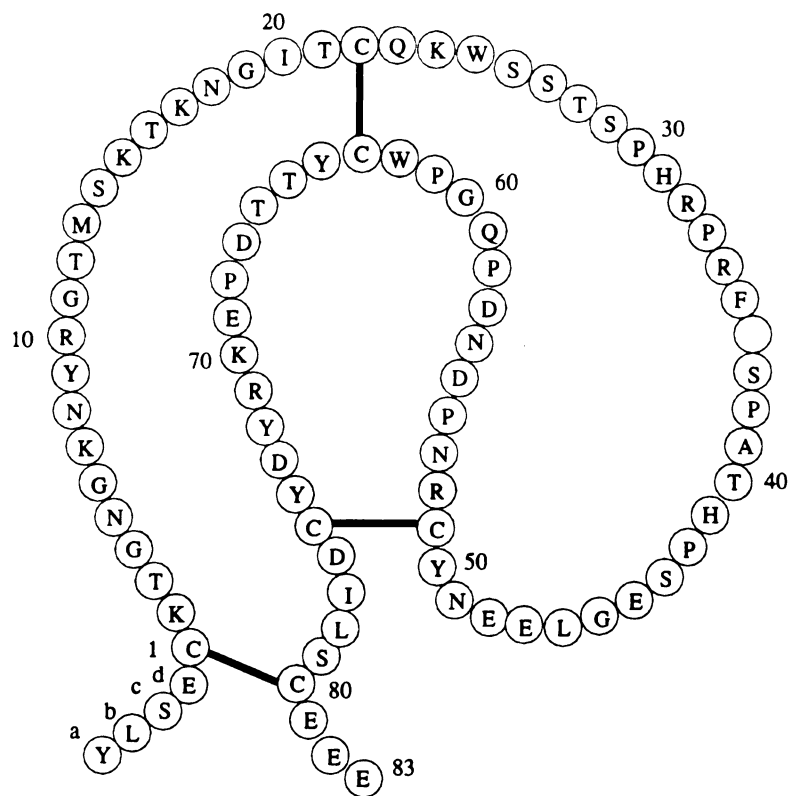


Figure 4.1. The primary structure of human plasminogen kringle 1. Site numbering deletion is based on homology with PGK5.

and Phe64 in PGK4. The observation supported the hypothesis that the buried hydrophobic core is fundamental for the overall kringle folding [49].

Since PGK1 can bind to fibrin and some ω -aminocarboxylic acids, considerable effort has been made to understand the nature of the lysine binding site. From the proton Overhauser experiment [50], it has been found that the ligand interacts directly with Phe36, Trp62 and Tyr72. A close interaction which has also been found between Tyr64 and Tyr72 indicates that residue 64 is positioned close to the binding site. Chemical modifications made with 1,2-cyclohexanedione [51] showed that both Arg32 and Arg34 are involved in fibrin binding; additionally, Arg34 is also involved in the binding of ω -aminocarboxylic acids. In the PGK4 lysine binding site, Asp55/Asp57 and Arg71/Lys35 are located on a dipolar surface to form the anionic and cationic centers. The same arrangement, except for Lys35, was also observed in the PGK1 structure. Trexler *et al.* [17] have pointed out that only PGK1 and PGK4 have both Asp at 57 and Arg at 71; in addition, the lysine binding sites present in PGK1 and PGK4 have been shown to have the same specificity [19]. Therefore, Asp57 and Arg71 in the PGK1 binding site were suggested to act as anionic and cationic centers, respectively.

Recently, based on the three-dimensional structure of PTF1 and NMR observations [50], the PGK1 lysine binding site was modeled by Tulinsky *et al.* [38] and a stereoview of the modeled structure is shown in Figure 4.2. In the modeling, conserved side chain atoms between PTF1 and PGK1 were kept constant and like atoms of unconserved residues served as guide coordinates for replaced side chains. As shown in Figure 4.2, Asp55 and Asp57 serve as anionic center, and Arg34 participates with Arg71 in making up the cationic center of the lysine binding site. The groove of the binding site is lined by the aromatic rings of Phe35, Trp62, and Tyr72, with Tyr72 interacting with both Trp62 and, in a second aromatic plane, with Tyr64. One boundary of the site is determined by the Pro54-Pro58 pentapeptide stretch,

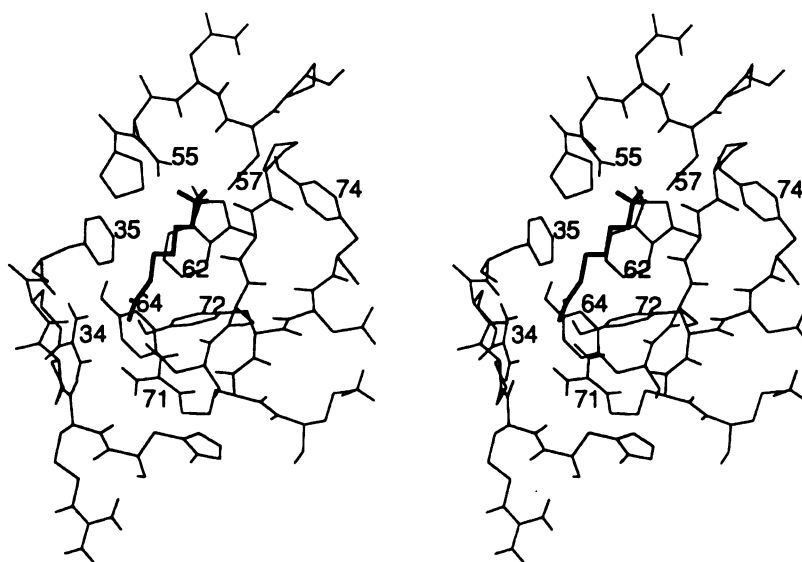


Figure 4.2. Stereoview of the modeled lysine binding site of plasminogen kringle 1;
ACA in bold. Hydrogen atoms are shown at ideal positions (taken from [38]).

which contains both Asp55 and Asp57. The other consists of two short, disconnected peptide segments containing Arg34 and Arg71.

The X-ray crystal structure determination of PGK1 was undertaken to provide a better understanding of the kringle structure and the nature of lysine binding. The structures of PGK1 and PGK4 give suitable information for modeling kringles of plasminogen, which has five kringles.

4.2 Materials and Methods

A. Crystallization

The PGK1 was isolated intact via proteolytic fragmentation of plasminogen. The PGK1 used for crystallization was kindly provided by Dr. F. J. Castellino in the form of a lyophilized powder, which was stored at freezer temperatures in a dry environment.

Because of the homology between PGK1 and PGK4, the conditions used for crystallizing PGK4 crystals were first applied to PGK1. A drop of 10 mg/ml of protein solution was equilibrated against a 1 ml well solution of 0.12 M ammonium sulfate, with 30% PEG 8000 and 1.2% DMF as an additive, at pH 6.0. However, no crystals formed. In an effort to find a new crystallization condition for PGK1, a factorial search [52] was carried out. The search consisted of a set of 31 factorial solutions and consisted of a random but balanced distribution of common precipitants, salts, and buffers. Crystallization was carried out using the vapor diffusion method at room temperature: 2 μ l hanging drops of the protein to which an equal volume of precipitant was added were used for each individual trial. The factorial solution with 0.1 M magnesium chloride, 0.1 M (N-[2-Acetamido]-2-iminodiacetic acid, pH 6.5)(ADA), 30% PEG 8000 produced clusters of long, fine needle-like crystals after a few days. Trials to refine the crystallization condition by reducing the precipitant concentra-

tion and also introducing 1–2% of the organic additives in the well, either resulted in clusters of needles or clear drops. In order to improve the crystal size, a macroseeding technique was adopted. However, using needles as seeds in macroseeding is more complicated since needles have a tendency to bend while being transferred from the original growth solution to the new solution. The stress created in the crystals during this process may result in defects at each stress point, and each of them may act as a nucleation site for growth of new needles. In order to avoid such phenomenon, the needles were broken into smaller segments by a sharp glass fiber and the resulting fragments were used as seeds. The seeds were transferred to fresh protein drops every two weeks, since they started dissolving upon standing for more than two weeks. It was thought that there might have been a serine protease contaminant in the protein sample which made the crystals dissolve. Crystallization in the presence of serine protease inhibitors such as benzamidine (50mM), which can bind to PGK1 also, was tried. However, crystals still dissolved after 2–3 weeks. Crystallization at the cold room temperature (4° C) was then carried out. The crystallization trays were transferred to cold room immediately after the hanging drops were set up. The crystals again dissolved but only after 5–6 weeks. A crystal of size $1.2 \times 0.2 \times 0.1$ mm was grown (Figure 4.3) by a macroseeding procedure after reducing the PEG 8000 concentration to 23% in the well at 4° C. This crystal which was the only one grown to X-ray diffractable size was mounted and data were collected. The unit cell parameters and some crystal data are listed in Table 4.1. From the above observations, it appears that the solubility of PGK1 is temperature dependent; i.e., the solubility of the PGK1 protein increases with increasing temperature.

B. Data Collection and Data Reduction

X-ray diffraction intensity data were measured at 2.48 Å resolution from the crystal specimen having dimensions $1.2 \times 0.2 \times 0.1$ mm employing a Siemens Xentronics

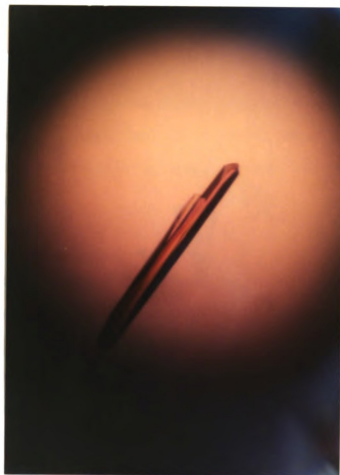


Figure 4.3. Photograph of K1 crystal. Crystal length is approximately 1.2 mm.

Table 4.1. Summary of crystal data of tetragonal K1.

space group	P4 ₃ 2 ₁ 2
crystal system	tetragonal
a (Å)	58.93
b (Å)	58.93
c (Å)	54.64
molecules/unit cell	8
molecules/asymmetric unit	1
molecular weight (dalton)	10000
$V_m (\frac{\text{\AA}^3}{\text{dalton}})$	2.37
solvent fraction (%)	48.0

Area Detector with graphite monochromated CuK α radiation generated by a Rigaku **RU200** X-ray generator operating at 7.5 KW (50KV, 150MA). The crystal-detector **distance** was set to allow neighboring Bragg reflections to be distinguished and is **usually** determined from

$$D = a_{\text{MAX}}/8, \quad (4.1)$$

where a_{MAX} is the largest unit cell length in Å. The distance was 11.65 cm in the data **collection** because: (1) the dimensions of the unit cell of the kringle are smaller than **93 Å** (Table 4.1); and (2) the closest distance possible with this instrument is 11.65 **cm**. The detector swing angle was set at 12.5°, the scan range was 0.2° per frame of **measurement** and each frame was collected for 120 seconds.

The raw intensity data reduction was carried out using the XENGEN **programs** [26]. A total of 17,701 observations were measured, of which 3,444 were **independent**. The distribution of intensities observed for various resolution ranges **is given** in Table 4.2. After reflections with $I/\sigma(I) < 2$ were removed, a set containing **2,987** independent reflections remained (87% observed, $R_{\text{MERGE}} = 0.059$).

Table 4.2. Distribution of reflection intensities and R-factors in various resolution shells.

Res. (Å)	# refs	0	<2 σ	<5 σ	<10 σ	<20 σ	<40 σ	<60 σ	>60 σ
4.5	673	3	12	13	16	39	80	94	416
3.57	631	9	23	15	40	47	119	117	216
3.12	606	22	36	40	60	117	165	81	82
2.84	609	44	65	79	112	144	124	26	13
2.63	596	55	107	121	99	134	71	9	0
2.48	329	40	82	92	68	34	12	0	0
Totals:	3444	173	325	360	395	515	571	327	772

C. Molecular Replacement

Since PGK1 has a different space group from PGK4, Patterson rotation/translation molecular replacement methods were used to obtain initial phases for structure determination and refinement. Molecular replacement was carried out with the program X-PLOR [53]. In the rotation search, the selected vectors were rotated by application of a rotation matrix $C(\theta_1, \theta_2, \theta_3)$ where $\theta_1, \theta_2, \theta_3$ are the Eulerian angles. The orientation of the search model was sampled by using the Lattman angles [54]:

$$\theta_+ = \theta_1 + \theta_3 \quad \theta_- = \theta_1 - \theta_3 \quad \theta_2 = \theta_2 \quad (4.2)$$

The interval Δ for θ_2 is a constant, the interval for θ_+ is given by $\Delta/\cos(\theta_2/2)$ and the interval for θ_- is given by $\Delta/\sin(\theta_2/2)$.

In a conventional rotation search, the highest peak is usually assumed to be the correct orientation, but sometimes this is not true. In the program X-PLOR, the

highest peaks of the rotation function are refined prior to a translation search. This refinement is performed by minimizing the target function of individual atomic coordinates (r) :

$$E_{rot}(r) = E_{PC}(r) + E_i(r) \quad (4.3)$$

where E_i is an empirical energy function that describes the geometry and non-bonded interactions of the molecule and E_{PC} is an effective energy term that is proportional to the standard linear correlation coefficient between the squared amplitudes of the observed and the calculated normalized structure factors. The normalized structure factors are computed with the search model placed in a triclinic unit cell identical in geometry to that of the crystal. After this refinement, the refined search model with the highest correlation coefficient is used for a translation search.

The model employed in the molecular replacement was the entire peptide backbone and conserved side chains of the refined PGK4 structure, representing 504 atoms or 82% of the PGK4 structure. A triclinic cell of dimensions $90 \times 90 \times 90$ Å using data from 8.0 to 4.0 Å resolution was used in the model structure factor calculation. Both the model and unknown PGK1 Patterson vectors of length greater than 3.0 Å and less than 15.0 Å were selected. The set of the 1500 highest Patterson peaks were used in the rotation search. The highest peak was found with Euler angles (105° , 155° , 80°). The refinement of this solution was carried out and the final solution had peak height 3.5σ above the mean at (112.2° , 153.7° , 88.8°).

Translation search of the rotated search model was also carried out with the program X-PLOR by computing the standard linear correlation coefficient

$$TF(\mathbf{xyz}, C) = \frac{\langle |E_{obs}|^2 |E_{calc}(\mathbf{xyz}, C)|^2 \rangle - \langle |E_{obs}|^2 \rangle \langle |E_{calc}(\mathbf{xyz}, C)|^2 \rangle}{\sqrt{[\langle |E_{obs}|^4 \rangle - \langle |E_{obs}|^2 \rangle^2] \times [\langle |E_{calc}(\mathbf{xyz}, C)|^4 \rangle - \langle |E_{calc}(\mathbf{xyz}, C)|^2 \rangle^2]}} \quad (4.4)$$

where E_{obs} are the normalized observed structure factors and E_{calc} are the normalized

calculated structure factors. The search model was placed in the unit cell of the crystal with the position given by the coordinates x , y , z of the center of gravity and with the orientation given by the rotation matrix C . Data from 8.0 to 4.0 Å resolution was used in the translation calculation through which both enantiomorphic space groups $P4_12_12$ and $P4_32_12$ were examined. The latter was taken as the correct space group because its solution was 5.7σ above the mean, which was significantly larger than that of the former space group (3.9σ). The correlation coefficient was 0.34, with the next highest being 0.28. After the orientation and position of the molecule had been determined, a rigid body refinement was carried out giving an R-factor of 0.44 at 8.0 – 4.0 Å resolution.

D. Structure Refinement

The K1 structure was refined using the program PROLSQ [30]. The refinement was carried out in two stages, initially with data at 7.0–2.8 Å, then extending the resolution to 2.48 Å. At the beginning of the refinement, an overall thermal parameter was assigned as 20 Å^2 . After 10 cycles, the refinement converged at a R-factor of 0.34 and the $(2|F_o| - |F_c|)$ and $(|F_o| - |F_c|)$ electron density maps were calculated. The maps clearly indicated the positions for side chains which were not included in the calculation. Model building was then performed on an Evans and Sutherland PS390 interactive stereographics system. The R-factor was reduced to 0.22 at the end of the first stage of refinement. The remainder of the data (to 2.48 Å) were included and water molecules were also gradually added. The reflection weights were assigned in seven shells of $\sin \theta / \lambda$ based on $(||F_o| - |F_c||) / 2$ of the range. The final PGK1 structure has a crystallographic R value of 0.16 for 2695 reflections between 7.0–2.48 Å resolution range with 73 water molecules and two chloride ions. The final reflection weights and R values in each range are given in Table 4.3, and a summary of refinement parameters is listed in Table 4.4. The average thermal parameter is

Table 4.3. Weights of Reflections and R Values of the Final Refinement

$D_{\min}(\text{\AA})$	no. reflections	$\sigma(F)$	$< F_o - F_c >$	R value	
				shell	sphere
4.70	357	24	57	0.189	0.189
3.90	393	21	41	0.140	0.164
3.45	391	19	37	0.142	0.157
3.15	381	17	30	0.155	0.157
2.92	379	15	26	0.161	0.157
2.72	406	14	23	0.165	0.158
2.40	388	12	21	0.173	0.159
$\sigma(F) = 15 - 120 (\sin \theta / \lambda - 1/6); < F_o - F_c > = 34.$					

Table 4.4. Summary of Final Least-Squares Parameters and Deviations

	target (σ)	rms (Δ)
Distances ($\text{\AA}$)		
Bond Distance	0.020	0.015
Angle Distance	0.035	0.043
Planar 1,-4 distance	0.050	0.042
Non-bonded distances ($\text{\AA}$)		
Single torsion	0.55	0.22
Multiple torsion	0.55	0.30
Possible H-bond	0.55	0.27
Torsion angles (deg)		
Planar	3	2
Staggered	15	23
Orthonormal	20	25
Plane groups ($\text{\AA}$)	0.02	0.02
Chiral centers ($\text{\AA}^3$)	0.15	0.17
Thermal restraints ($\text{\AA}^2$)		
Main chain bond	1.5	0.8
Main chain angle	2.0	1.3
Side chain bond	2.0	1.2
Side chain angle	2.5	2.0

27.6 Å², which is higher than that of PGK4 (18 Å²). This is consistent with the high solvent content of PGK1 (48% versus 38%).

4.3 Results and Discussion

A. General Structure

Nearly the entire structure of PGK1 is well defined in the electron density; however, no density was found for interkringle regions of a-c and 81-83. Thus, the refined PGK1 structure extends from residue Glu-d to Cys80. The distribution of the main-chain torsion angles (ϕ , ψ) of the refined PGK1 structure is shown in Figure 4.4, from which it can be seen that all the non-glycine amino acids, except Glu48, conform satisfactorily to the conformationally allowed regions. Although the residue Glu48 is outside the allowed region, it fits in electron density very well. The same conformation is also observed in the PGK4 structure, where position 48 is methionine. The distribution of the omega angles of the refined PGK1 structure is shown in Figure 4.5. All residues, except one, have ω angles within the $180^\circ \pm 5^\circ$ region. A *cis*-proline with an ω angle of 2° at position 30 is also found in this structure similar to that in the apo-K4 structure. The secondary structural features of PGK1 are listed in Table 4.5. The closed interkringle C, D loops (Cys51-Cys75) form two distinct antiparallel β -sheets due to close contact of two disulfide bridges in the folded state. There are 10 reverse turns in the PGK1 structure. It appears that about half of the kringle residues are involved in these turns, of which one third are conserved among different kringle sequences. Thus, the similar folding of kringles could be due to the large number of conserved turns.

The final r.m.s. B-values of main-chain and side-chain atoms along the sequence of PGK1 are shown in Figure 4.6. The magnitudes and variations of B values are likely reflective of the positions of the residues with respect to the three-dimensional

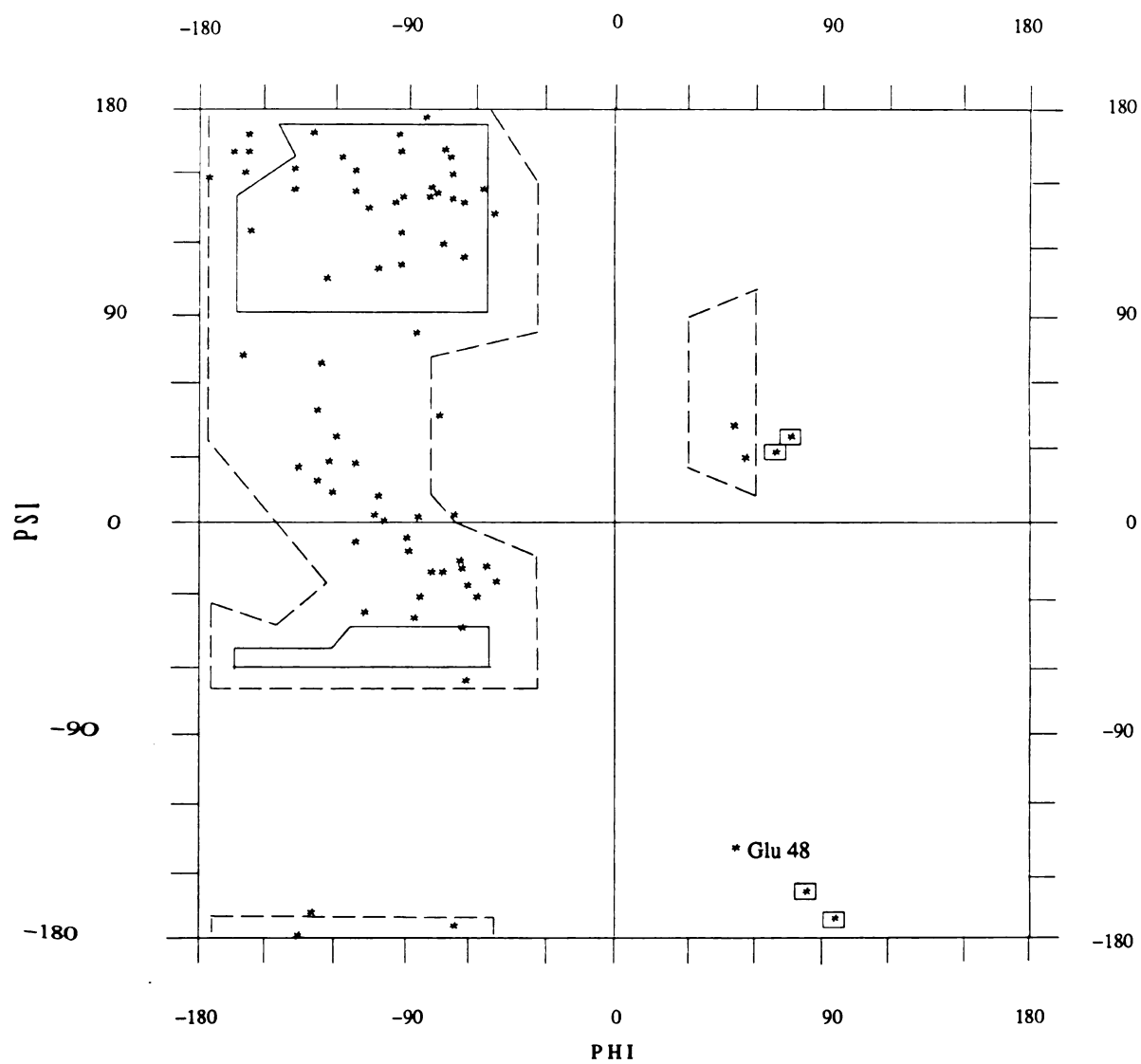


Figure 4.4. The Ramachandran plot of PGK1. Glycine residues are boxed.

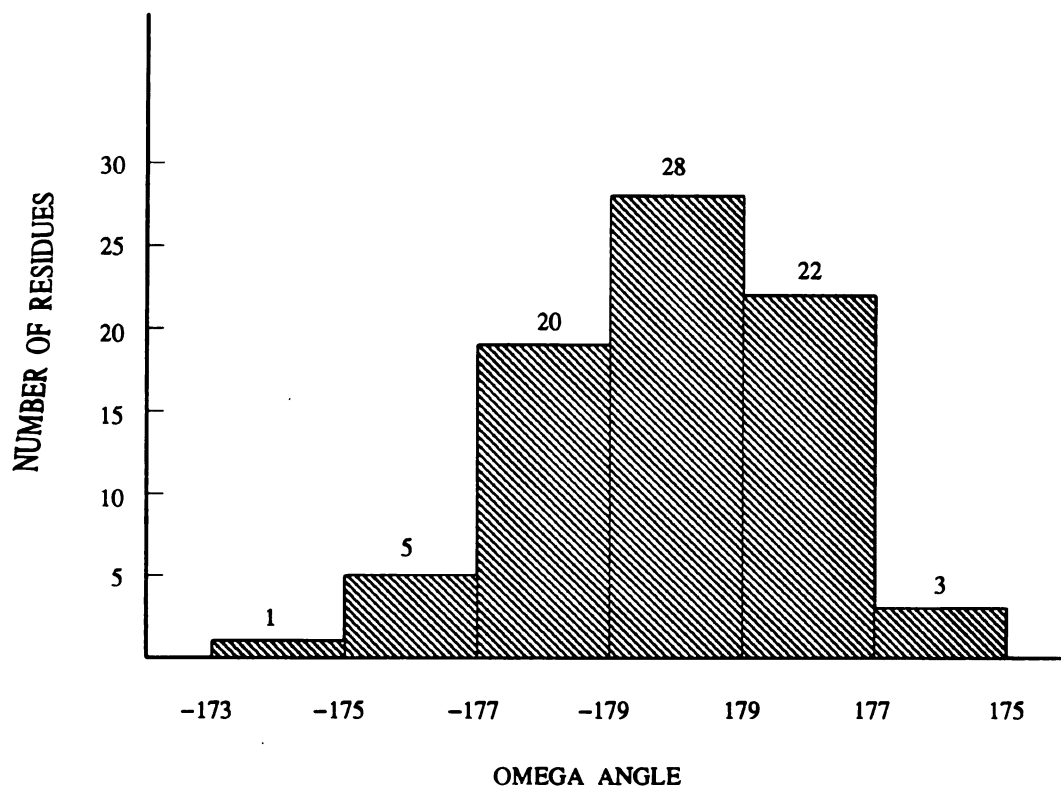


Figure 4.5. Distribution of ω angles in the PGK1 structure.

Table 4.5. Secondary structural elements of PGK1 structure.

β Structures	Type	Residues
$\beta 1$	Antiparallel	Ser14 – Cys22 Lys15 – Thr21 Thr16 – Ile20
$\beta 2$	Antiparallel	Gln23 – Tyr50 Lys24 – Asn49 Trp25 – Glu48
$\beta 3$	Antiparallel	Arg52–Trp62 Asn53 – Pro61
$\beta 4$	Antiparallel	Cys63 – Cys75 Tyr64 – Tyr74 Thr65 – Asp73 Thr66 – Tyr72
Reverse Turns	Type	Residues
T1	Type II'	Thr3–Gly4–Asn5–Gly6
T2	Type I	Gly6–Lys7–Asn8–Tyr9
T3	Type I	Thr16–Lys17–Asn18–Gly19
T4	Type III	Lys24–Trp25–Ser26–Ser27
T5	Type III	Ser37–Pro38–Ala39–Thr40
T6	Type I	His41–Pro42–Ser43–Glu44
T7	Type II'	Glu47–Glu48–Asn49–Tyr50
T8	Type I	Asn53–Pro54–Asp55–Asn56
T10	Type I	Asp67–Pro68–Glu69–Lys70

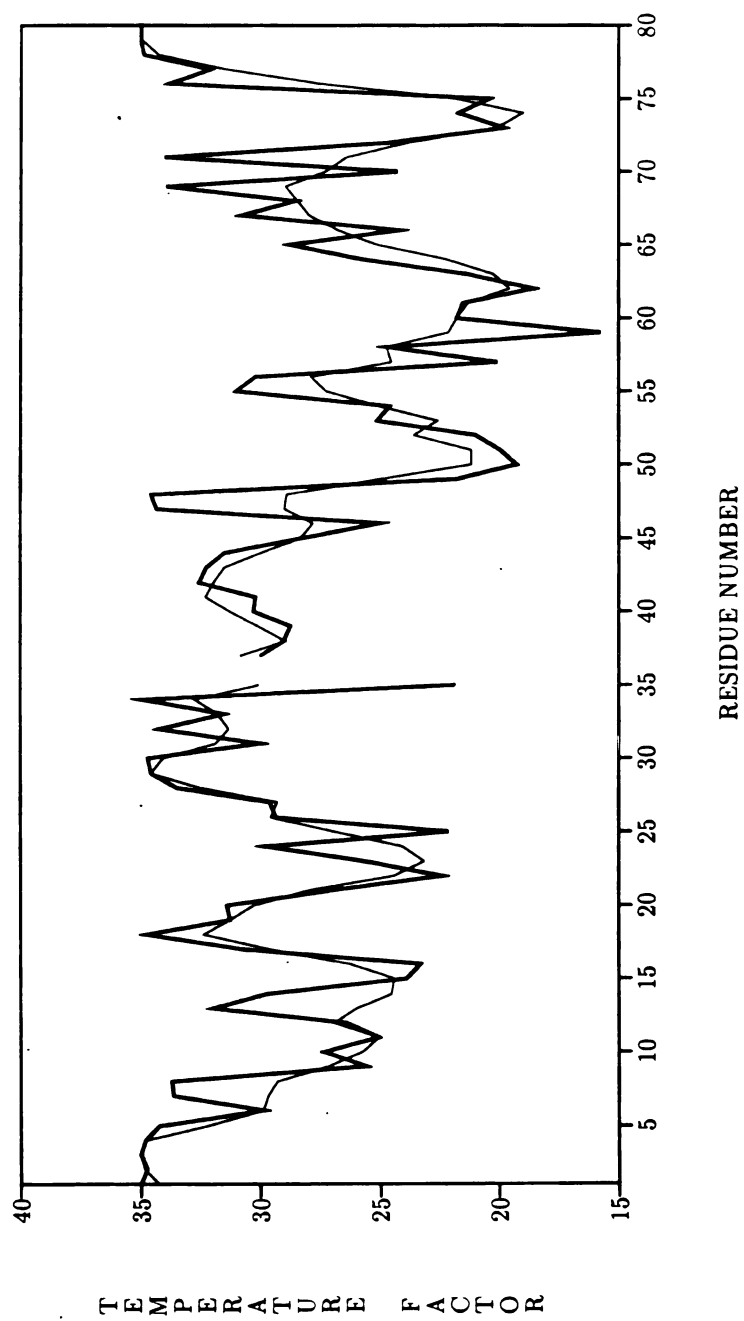


Figure 4.6. Plot of the average thermal factors of main-chain (thin line) and side-chain (thick line) atoms of PGK1 structure.

folding. The lowest values shown are associated with residues involved in the disulfide groups, especially the central disulfide cluster, which is buried in the interior of the kringle domain (Cys22–Cys63, Cys51–Cys75). The stabilizing effect of the disulfide bridges on kringle folding is very visible. The high B values in Figure 4.6 correspond to residues that are poorly defined and are somewhat disordered, especially residues near the terminals of the peptide chain. The B values of the side chain groups are generally higher than that of main chain showing that side chain atoms have more flexibility than main chain atoms. From Figure 4.6, it can also be seen that low values are associated with the C, D inner loops (Cys51–Cys75) of the kringle while the larger B loop (Cys22–Cys51) which is six residues longer than A loop, displays more flexibility than most other parts of the kringle.

The refined PGK1 structure contains 71 water molecules with average occupancy of 0.68 and average thermal parameter of $24 \text{ \AA}^2$. The distributions of thermal parameters and occupancies are shown in Figures 4.7 and 4.8. Of the 71 water molecules, 47 have occupancies between 0.5 and 1.0, and 10 have full occupancy. There is only one water molecule buried within the PGK1 structure, where Gln23 O and Tyr64 N make hydrogen bonds with the water molecule of $2.80 \text{ \AA}$ and $2.71 \text{ \AA}$, respectively (Figure 4.9). The internal water molecule has the highest quality factor (Occ^2/B) ranking among all of the water molecules, with an occupancy, 1.0, and the lowest B value of $12 \text{ \AA}^2$.

The hydrophobic core formed by a number of stacked aromatic and proline residues, centered around Leu46, is shown in Figure 4.10. Of these side chain groups, all make close van der Waals contacts with one another and a number of the aromatic rings participate in stacking interactions in which one aromatic ring is pointed in a perpendicular manner toward another. The edge-to-face interactions observed in the PGK1 structure also include some proline residues. The edge of Pro61 is positioned $4.5 \text{ \AA}$ from the tyrosyl ring of Tyr9 at an angle of approximately 100° . In addition,

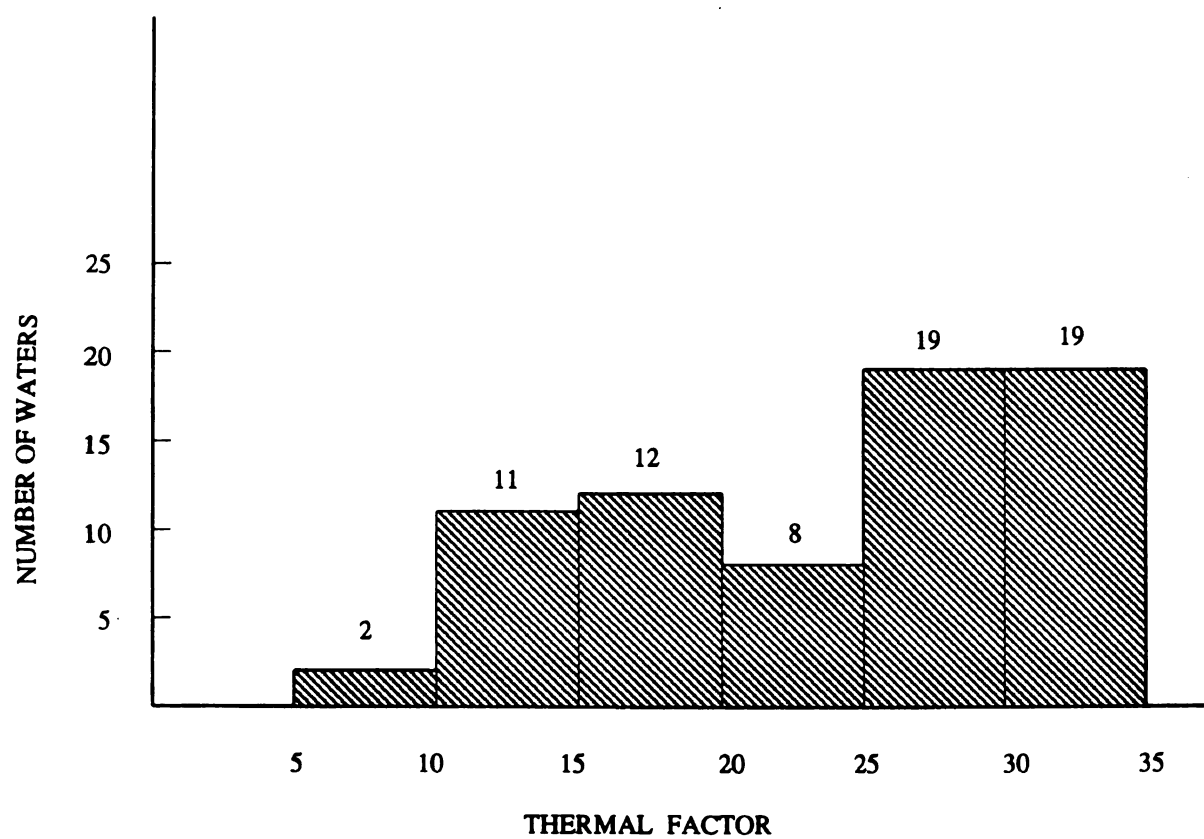


Figure 4.7. Distribution of temperature factors for water molecules.

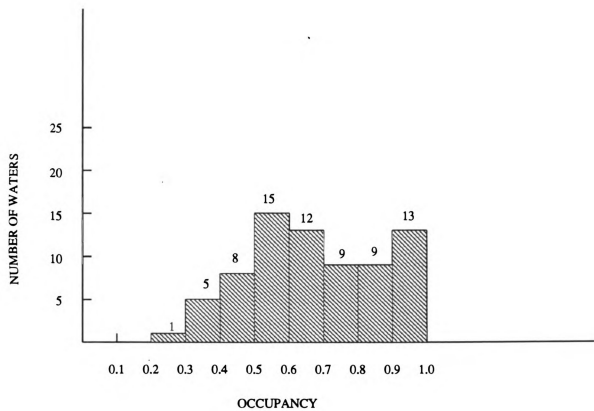


Figure 4.8. Distribution of occupancies for water molecules.

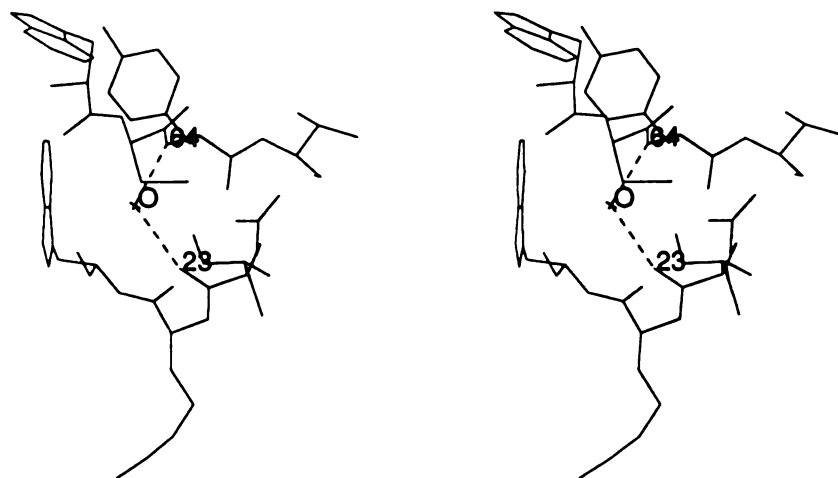


Figure 4.9. Stereoview of the internal water molecule in PGK1 structure. Dashed lines indicate hydrogen bonds.

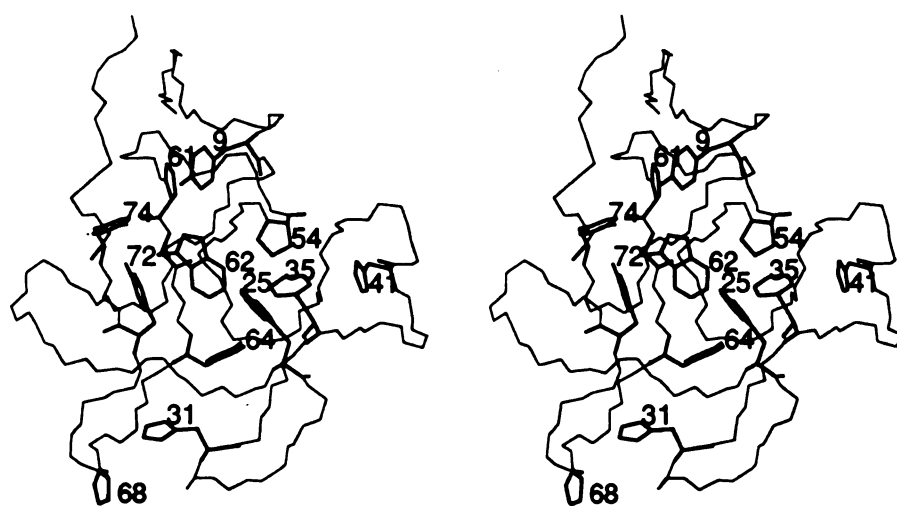


Figure 4.10. Stereoview of the PGK1 aromatic and proline residues forming hydrophobic core.

the edge of Pro54 is directed at an angle of 95° toward the phenyl ring of Phe35, at a separation of about 3.8 Å. Such aromatic clustering is common in proteins and reinforces structural stability [45].

B. Structure of the Lysine Binding Site

The lysine binding site observed in the PGK1 structure is similar to that of the PGK4 structure. It is a relatively open, elongated, shallow depression and is located on the surface of one of oblate faces of the kringle, bounded by the segments His31–Phe35, Pro54–Pro58, Pro61–Tyr64, Arg71–Cys75 (Figure 4.11). Thus, this region corresponds mostly to the inner loops of the kringle sequence and is formed by the B, C and D segments (Figure 1.2). The two negatively charged Asp55 and Asp57 residues are located on one end of the lysine binding site having carboxylate oxygens 4.9 Å from each other. The two positively charged Arg34 and Arg71 residues are located on the other end of the depression and act as a cationic center. Although Arg34 participates with Arg71 in making up the cationic center, it has a somewhat disordered guanidinium group. As can be seen in Figure 4.12, a long channel of electron density is observed between Arg34 and Arg71 but the amino group of the former is outside the density map. In spite of the lack of density on the complete guanidino group, from the electron density up to Arg34 NZ, it is clear that the amino group of Arg34 must be close to the guanidino group of Arg71 so that the two can act as a cationic center. The aromatic cluster in the binding pocket, formed by the Phe35, Trp62, Tyr64, Tyr72, and Tyr74 side chains, provides a highly nonpolar environment between the charged centers. The indole ring of Trp62 and the tyrosyl ring of Tyr72 conform a V-shaped groove with an interplanar angle of approximately 80° , make aromatic stacking interactions with Tyr64 and Tyr74, respectively, and produce a stabilized structural framework for ligand binding. Thus, in PGK1, the binding site is a well-defined area in which zwitterionic ligands such as lysine and ACA can be

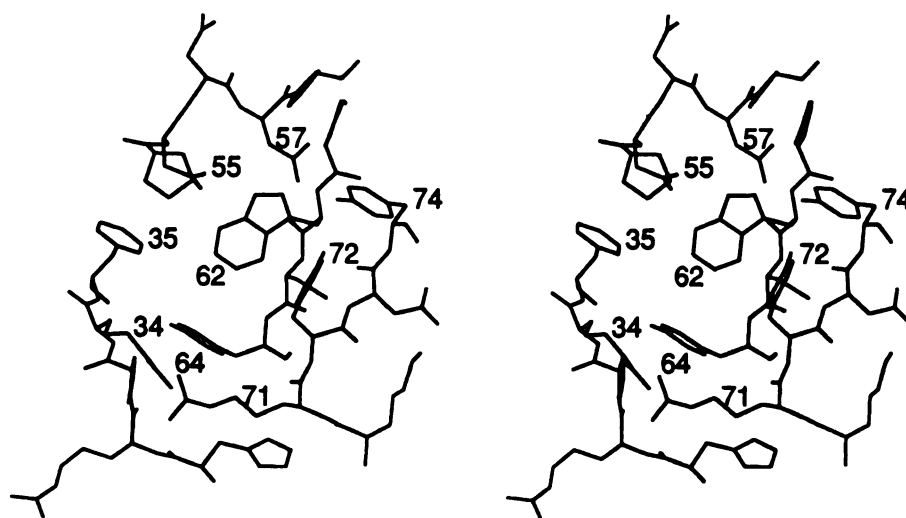


Figure 4.11. Stereoview of the PGK1 lysine binding site.

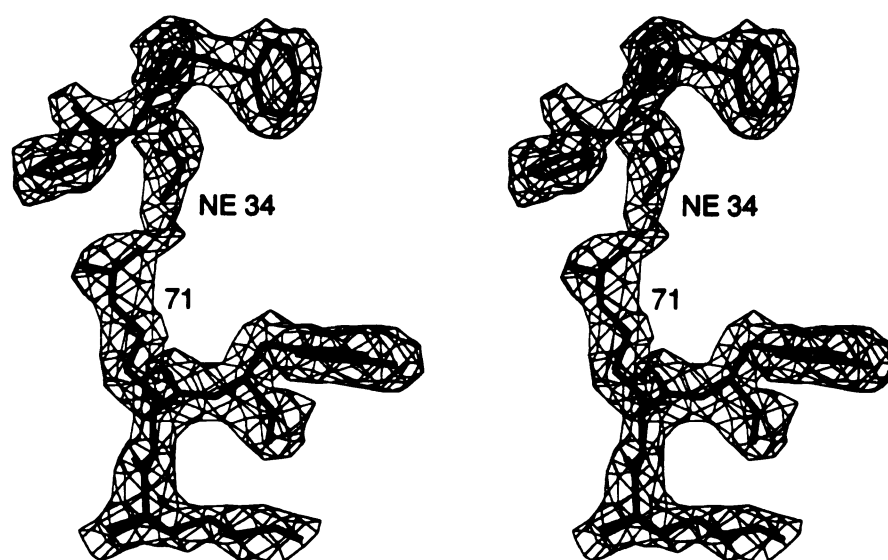


Figure 4.12. Electron density in vicinity of Arg34 and Arg71 of PGK1 structure; contour at 1σ .

docked by making ion pair interactions with anionic and cationic centers defined by the side chain carboxylate groups of Asp55 and Asp57 on one side and conformed by the guanidinium groups of Arg34 and Arg71 on the other side. The exposed lipophilic groove on the kringle surface then interacts with the nonpolar portion of the ligand by making close van der Waals contacts.

The most significant intermolecular interaction found in the crystal structure involves the lysine binding site of one molecule and the Lys15 side chain of a neighboring molecule: the oxygen atoms of the carboxylate groups of Asp55 and Asp57 interact with the ammonium group of the lysine (Figure 4.13). The positively charged Lys15' side chain of a symmetry mate makes hydrogen bonds with Asp55 OD1 and Asp57 OD1 at distances of 3.1 Å and 3.0 Å, respectively. In addition, the latter two side chains are oriented by hydrogen bonds between Asp55 OD2 and Trp62 NE1 and between Asp57 OD1 and Tyr74 OH. The hydrogen bonds found in the lysine binding site are summarized in Table 4.6.

Table 4.6. Hydrogen Bonds found in the lysine binding site of the PGK1 structure. The prime designates a symmetry-related molecule. A donor atom is denoted (D), an acceptor atom (A).

Donor	Acceptor	D...A (Å)
Tyr74 OH	Asp57 OD1	2.3
Trp62 NE1	Asp55 OD2	3.0
Lys15' NZ	Asp55 OD1	3.1
Lys15' NZ	Asp57 OD1	3.0

A second interaction observed in the binding site is mediated via chloride ions. The discovery of the two chloride ions was based on : (1) two large blocks of electron density near the guanidino groups of Arg34 and Arg71 were observed in the difference map and (2) charges are offset in the binding site by their presence. Since the Lys15

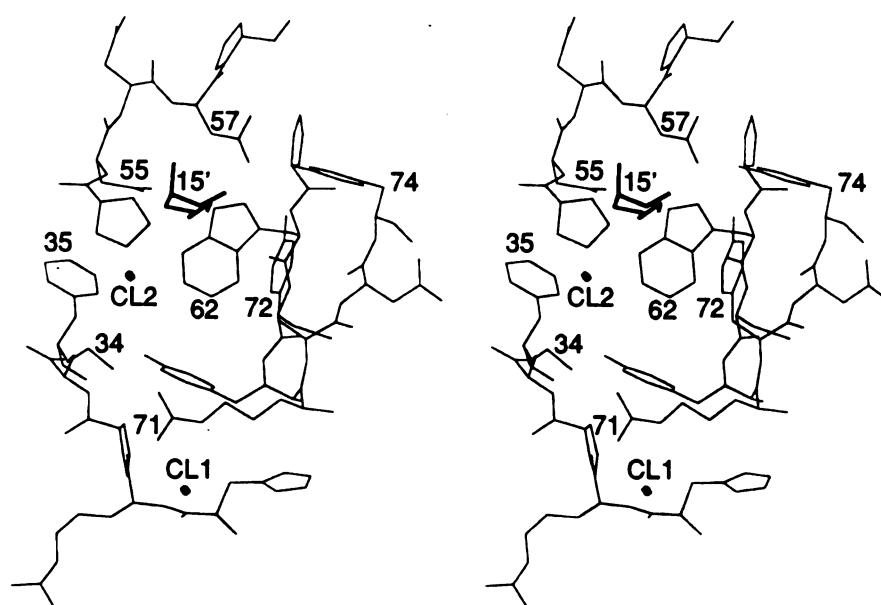


Figure 4.13. Stereoview showing intermolecular interactions involved in the lysine binding site. The prime denotes a symmetry-related molecule.

side chain from the symmetry-related molecule makes hydrogen bonds with Asp55 and Asp57, the inclusion of chloride ions in the cationic center produces interactions with positively charged residues Arg34 and Arg71 and all the charges in the binding site are essentially offset. The observed Cl-N distances are 3.4 Å (71NH2-Cl1) and 5.6 Å (34NH2-Cl2), respectively. Although the electron density of the two chloride ions is quite good, they have partial occupancies and average B values (Cl1 = 0.5, 29 Å²; Cl2 = 0.6, 35 Å²). The unexpected low occupancies and high B values might be due to the fact that two chloride ions, which are located on the kringle surface, are shared by two molecules in two different asymmetric unit and only partially occupy the lysine binding site. The slight disorder on the amino group of Arg34 might reflect the longer distance of Arg34-Cl2 (5.6 Å) than that of Arg71-Cl1 (3.4 Å).

C. Comparison of Plasminogen K1 and Plasminogen K4

The structures of PGK1 and PGK4 have been compared by calculating the rotation-translation components which minimize the difference in positions of the CA, C, and N atoms of the kringle. The rms deviation for the main chain atoms between the two structures is 0.88 Å. After removing 59 atoms (25%) having deviations greater than 1σ , the rms deviation was reduced to 0.49 Å, which is close to the expected error between the independent determinations. The average deviations of individual residues are shown in Figures 4.14 and 4.15, and the stereoview of the superposition of both backbone structures is shown in Figure 4.16. As can be seen in Figure 4.16 the overall conformation of the polypeptide chain of PGK1 is very similar to that of PGK4 except in two regions, around residue 58 and the region from residue 32 to 45. In the former region, the large deviations are due to insertion of Gln59 in PGK1, which gives rise to an expansion of the C loop. In the latter region, the rms deviation of the main chain is a surprisingly large 1.2 Å; the large deviations here are most likely due to different intermolecular interactions occurring around the two-fold axis relating

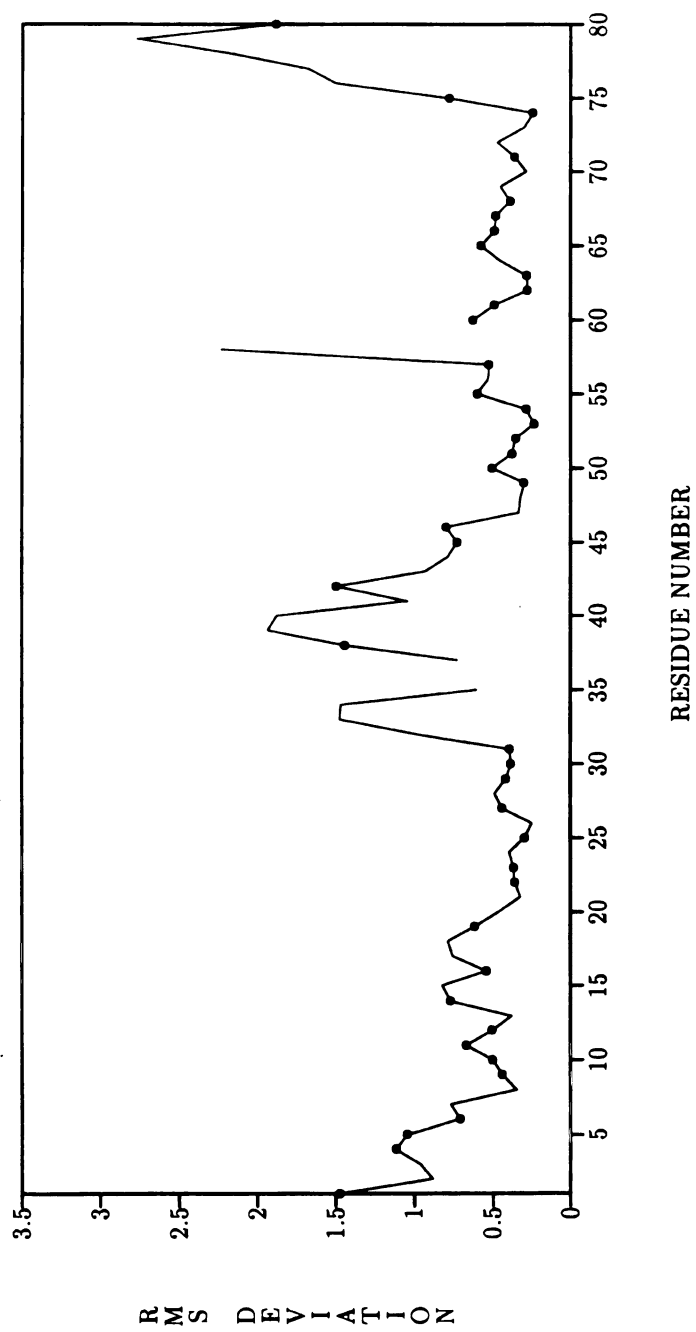


Figure 4.14. RMS deviations of CA, C, N atoms between PGK1 and PGK4 structures. Filled circles indicate conserved residues.

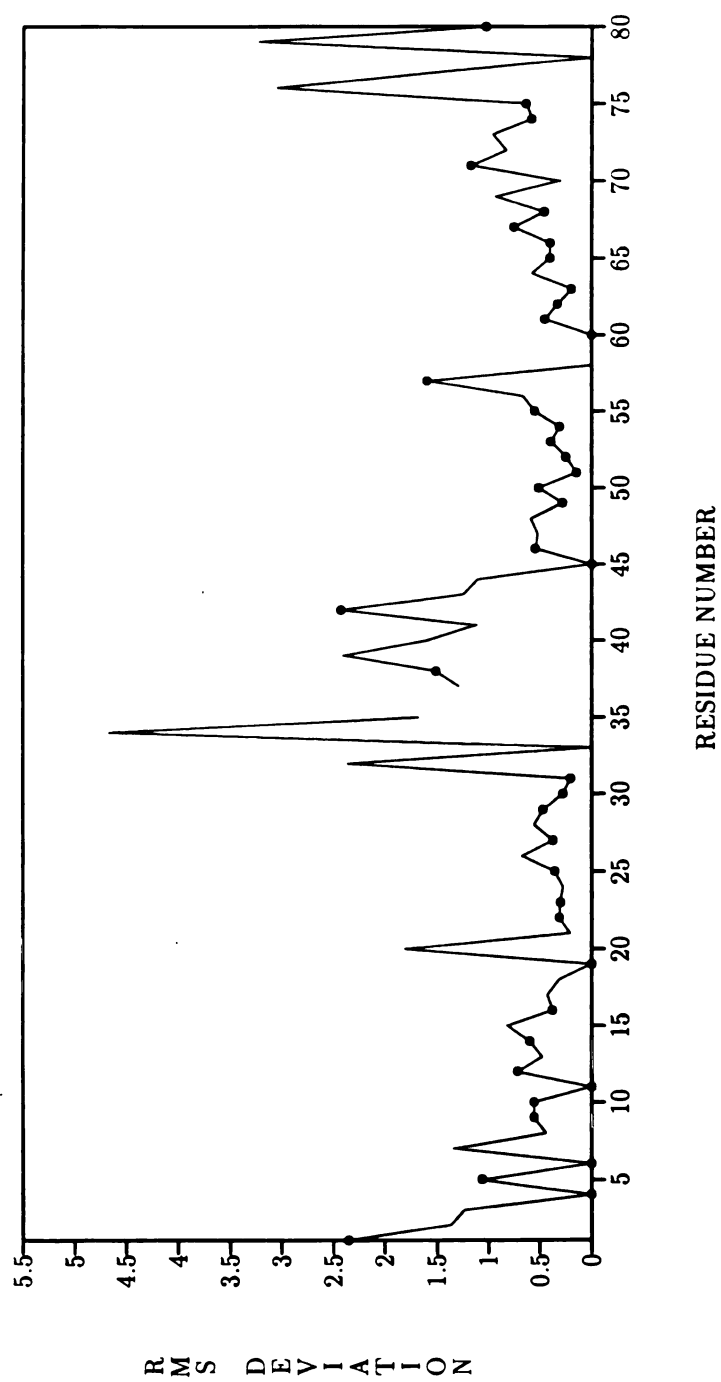


Figure 4.15. RMS deviations of side chains between PGK1 and PGK4 structures. Filled circles indicate conserved residues.

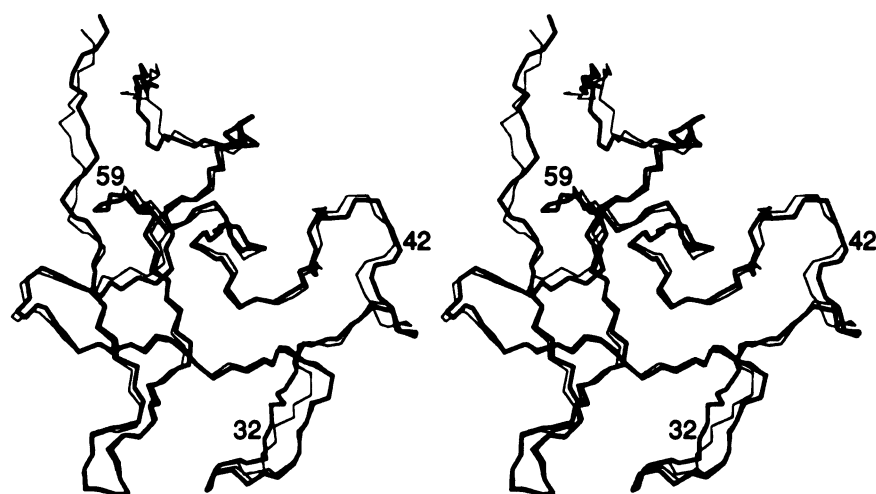


Figure 4.16. Stereoview of the optimally superimposed main chain structures of PGK1 (**bold**) and PGK4.

the segment between two neighboring PGK1 molecules (Table 4.7, Figure 4.17). Such interactions not only affect the conformations of the side chain groups but also make the position of the main chain backbone different from that of the PGK4 structure. In addition, in the PGK4 structure, Arg32 forms ligand-like interactions with Asp55 and the sulfate ion of a neighboring molecule (Figure 4.18) [23] that also may affect the main chain conformation of the B loop (Gln23-Tyr50). Both of these observations are in agreement with the second outer loop of kringles being relatively flexible with different conformations in tissue plasminogen kringle 2 [55], in the second kringle of prothrombin [56] and in the kringle of urokinase [57].

Since PGK1 and PGK4 have the ability to bind lysine, a comparison of the lysine binding site of the two structures has been carried out and a stereoview of the superposition of the binding site regions is shown in Figure 4.19. The rms differences between the two structures in the lysine binding sites are listed in Table 4.8. The average rms difference in main chain positions in this site is 0.42 Å except for residues 31-35 and 58. In the PGK1 structure, Arg34 projects into the binding pocket to participate with Arg71 in making up the cationic ionic center of the binding site, while in the PGK4 structure, Gln34 points to the bulk solvent without making any interaction with other residues. Arg32, which is conserved in both structures, has a huge deviation of 2.3 Å in the side chain position. One reason for this is the different free bond rotations of the side chain groups. Another reason is an intermolecular interaction in the PGK4 structure (Figure 4.18). Besides Arg32, other residues found in ligand-like binding in the PGK4 structure were Asp57 and Arg71. Similar to the interaction between Arg32 and Asp55, the positively charged Arg71 is positioned near the oppositely charged Asp57 residue of a neighboring kringle and participates in an ion pair interaction with it. Such interactions not only make the side chain conformations of Asp57 and Arg71 different in both structures but also account for the side chain group of Arg71 in the PGK4 structure extending outward away from

Table 4.7. Intermolecular interactions of PGK1 < 3.7A.

Molecule A	Molecule B	Distance(Å)
Arg32 NH1	Pro37 O	3.67
Arg32 CG	Ala38 O	3.06
Arg32 CD	Ala38 O	2.96
Arg32 NE	Ala38 O	3.63
Arg32 CD	Pro42 CG	3.49
Pro33 O	Thr40 CA	3.58
Pro33 O	Thr40 OG1	3.33
Arg34 CA	Thr40 CG2	3.52
Arg34 C	Thr40 CG2	3.20
Arg34 O	Thr40 CG2	2.70
Pro38 O	Arg32 NH1	3.67
Ala39 O	Arg32 CG	3.06
Ala39 O	Arg32 CD	2.96
Ala39 O	Arg32 NE	3.63
Thr40 CA	Pro33 O	3.58
Thr40 OG1	Pro33 O	3.33
Thr40 CG2	Arg34 CA	3.52
Thr40 CG2	Arg34 C	3.20
Thr40 CG2	Arg34 O	2.70
Pro42 CG	Arg32 CD	3.49

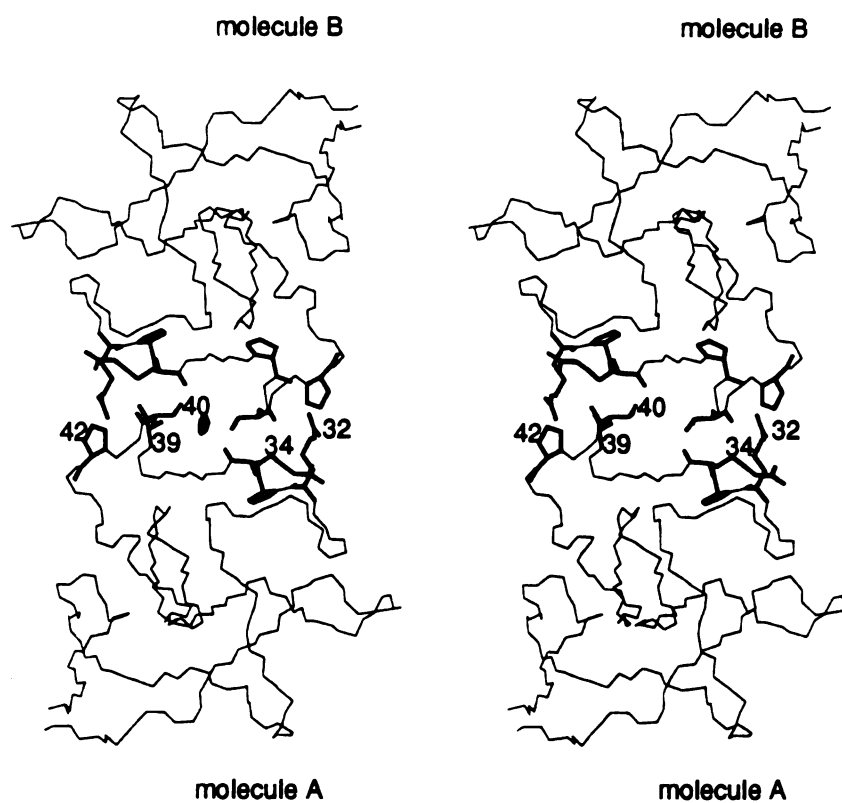


Figure 4.17. Stereoview of crystal packing interactions at a two-fold symmetry axis. Residues involved in intermolecular interactions are shown in bold. The two-fold axis in the a-b diagonal is indicated.

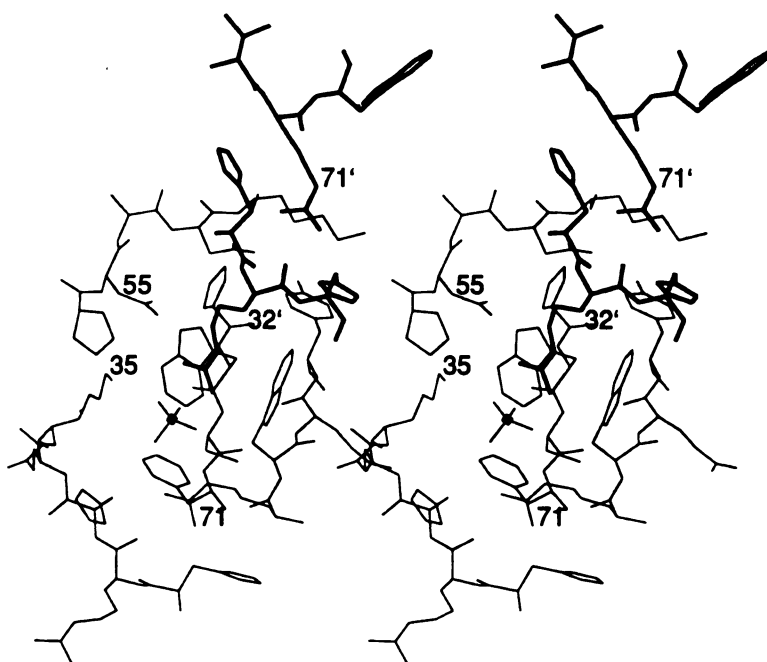


Figure 4.18. Stereoview of intermolecular interactions found in the lysine binding site of PGK4. The side chains from the second symmetry-related molecule are shown in bold; the sulfate ion is designated with a filled circle [23].

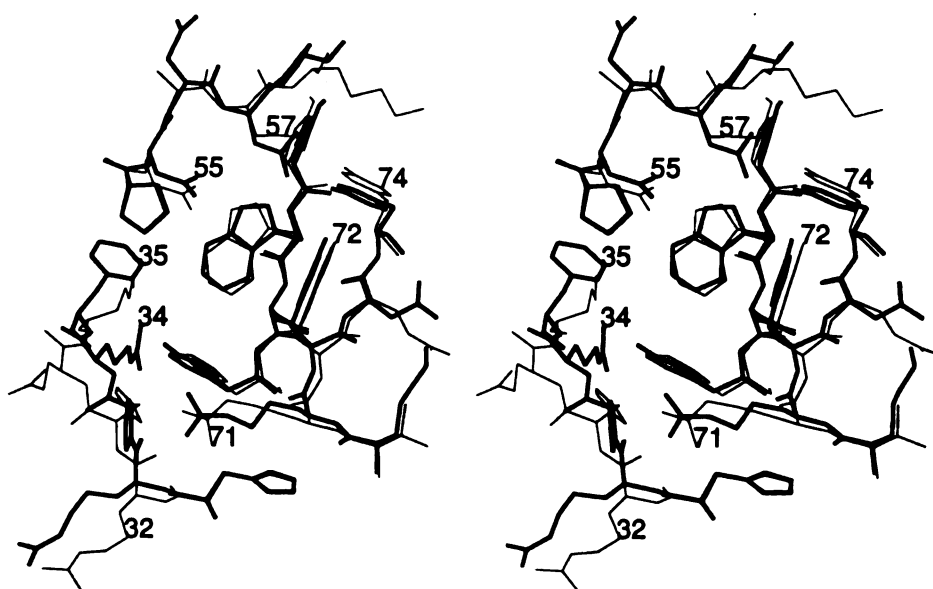


Figure 4.19. Stereoview of the superposition of the lysine binding sites of the PGK1(bold) and PGK4 structures.

Table 4.8. Deviations of the positions of residues in the lysine binding site between the PGK1 and PGK4 structures.

residue number	PGK1	PGK4	rms $\Delta(\text{\AA})$ (main chain)	rms $\Delta(\text{\AA})$ (side chain)
31	H	H	0.39	0.20
32	R	R	0.97	2.35
33	P	H	1.47	—
34	R	Q	1.46	4.66
35	F	K	0.60	1.67
54	P	P	0.28	0.31
55	D	D	0.59	0.55
56	N	A	0.52	0.66
57	D	D	0.52	1.59
58	P	K	2.22	—
61	P	P	0.48	0.45
62	W	W	0.27	0.32
63	C	C	0.28	0.19
64	Y	F	0.45	0.57
71	R	R	0.35	1.17
72	Y	W	0.46	0.82
73	D	E	0.30	0.95
74	Y	Y	0.24	0.58
75	C	C	0.77	0.63

the binding site. Lastly, the inclusion of Lys35 within the PGK4 binding site as a cationic center introduces a large rms difference of 1.6 Å between this residue and the Phe35 residue in the PGK1 structure. The two aromatic rings in the lysine binding site of PGK4 which are replaced by analogous residues in PGK1, are expected to behave similarly. For instance, PGK4 Phe64 and Trp72 are replaced by tyrosines at both positions in PGK1. Figure 4.19 shows that the agreement of the positions of the aromatic rings between the two structures is good for both main and side chain atoms. The remainder of the residues in the lysine binding site are practically identical in the two structures. A comparison of the two structures shows that even though the binding site is confined by a rather fixed set of residues in both PGK1 and PGK4, they show some flexibility. The absence of the aromatic ring at position 35 in PGK4 makes the lysine binding site region more open and might make it kinetically faster for ligand binding, especially for bulkier ligands such as AMCHA. On the other hand, the association constant of AMCHA of PGK1 is about twice as large as that of PGK4, probably because Phe35 is in the PGK1 lysine binding site.

CHAPTER 5

Structure of Tissue-type

Plasminogen Activator Kringle 2

Domain

5.1 Introduction

Tissue type plasminogen activator (tPA) is a serine protease that converts proenzyme plasminogen into plasmin, which, in turn, degrades the fibrin network into soluble products. The native tPA is a single polypeptide chain (sctPA) with 527 amino acids and a molecular weight of about 70,000 daltons [13, 58]. Upon limited plasmin hydrolysis the molecule is cleaved at the Arg278-Ile279 peptide bond and converted to a two chain activator (tctPA) linked by one disulfide bond (Figure 5.1) [59]. The sctPA is less active than the tctPA form toward small substrates [60], but both forms possess similar plasminogen activator activities [61]. It has been known that tPA consists of five domains which are homologous to parts of other proteins. Starting from the N-terminal, it consists of a finger domain, homologous with the finger domain responsible for the fibrin-affinity of fibronectin, a growth factor domain, similar to those of factor X, factor IX, bovine protein C, and the epidermal growth factor

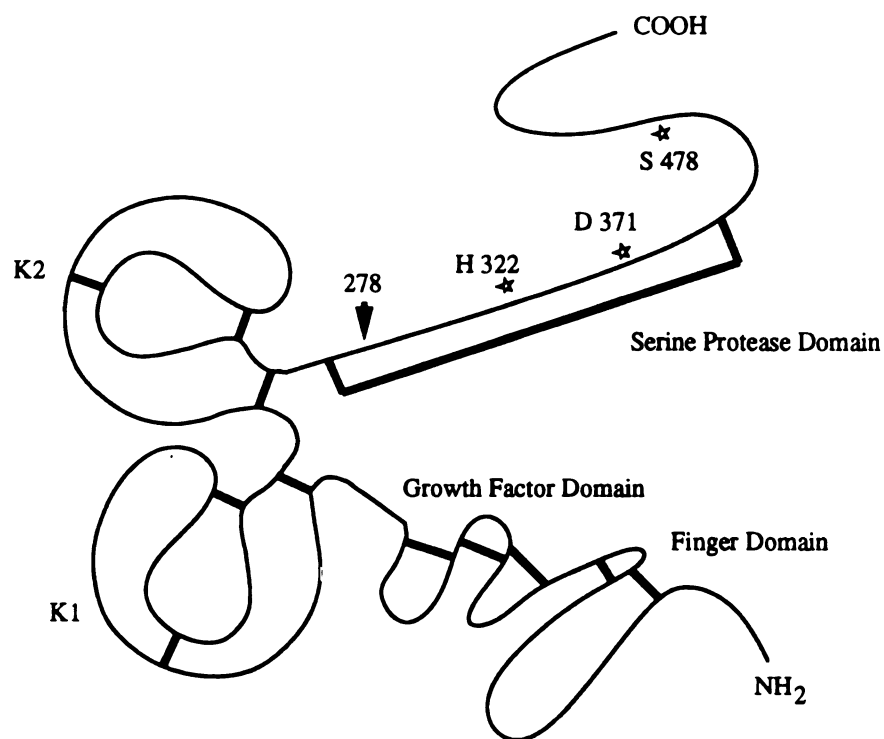


Figure 5.1. Schematic presentation of the primary structure of human tissue-type plasminogen activator. The arrow indicates the cleavage site for the conversion of single chain tPA to two chain tPA. The active site residues are indicated by asterisks.

domain, two kringle domains (K1 and K2), homologous to kringles of plasminogen, urokinase, factor XII, and prothrombin, and lastly, a protease domain, similar to those in trypsin or chymotrypsin [13].

The tPA can bind to fibrin like plasminogen and plasmin where it activates plasminogen to plasmin. Several studies have been done on the nature of the interaction between tPA and fibrin. Deletion mutant experiments have been performed and studied on the function of the individual domains of tPA, and it was concluded that stimulation of the plasminogen activator activity by fibrin was mediated both by the finger domain and the K2 domain [62]. A lysine binding site was found in K2 of tPA and its involvement in fibrin binding was strongly suggested [62]. The lysine binding site in K2 of tPA was shown to have equal affinity for lysine analogs with and without a free carboxylate group, suggestive that tPA, unlike plasminogen, does not discriminate between C-terminal and interchain lysine residues [63]. From chemical modification experiments, it was shown that one or more glutamic acid or aspartic acid residues in K2 of tPA are involved in the interaction with fibrin and lysine analogs [64]. Mutation experiments showed that Asp55 and Asp57 (Figure 5.2) were both essential for lysine binding, while Glu73 might be involved but was not essential [64].

Modeling of tPAK2 based on the structure of PTF1 was carried out by Tulinsky *et al.* [38], where the binding pocket consists of residues Asp55, Asp57, Trp62, and Trp72. The His64 residue, substituted for Tyr64 and Phe64 in PGK1 and PGK4, respectively, is located just below the surface of the binding site near the expected cationic site. Therefore, it was thought to act as a cationic center and provide a positive charge to the negative carboxylate group of the ligand.

Recently, the crystal structure of tPAK2 was determined by multiple isomorphous replacement (MIR) and refined by the program X-PLOR at 2.4 Å resolution [55]. Before applying the MIR method, molecular replacement was attempted to determine

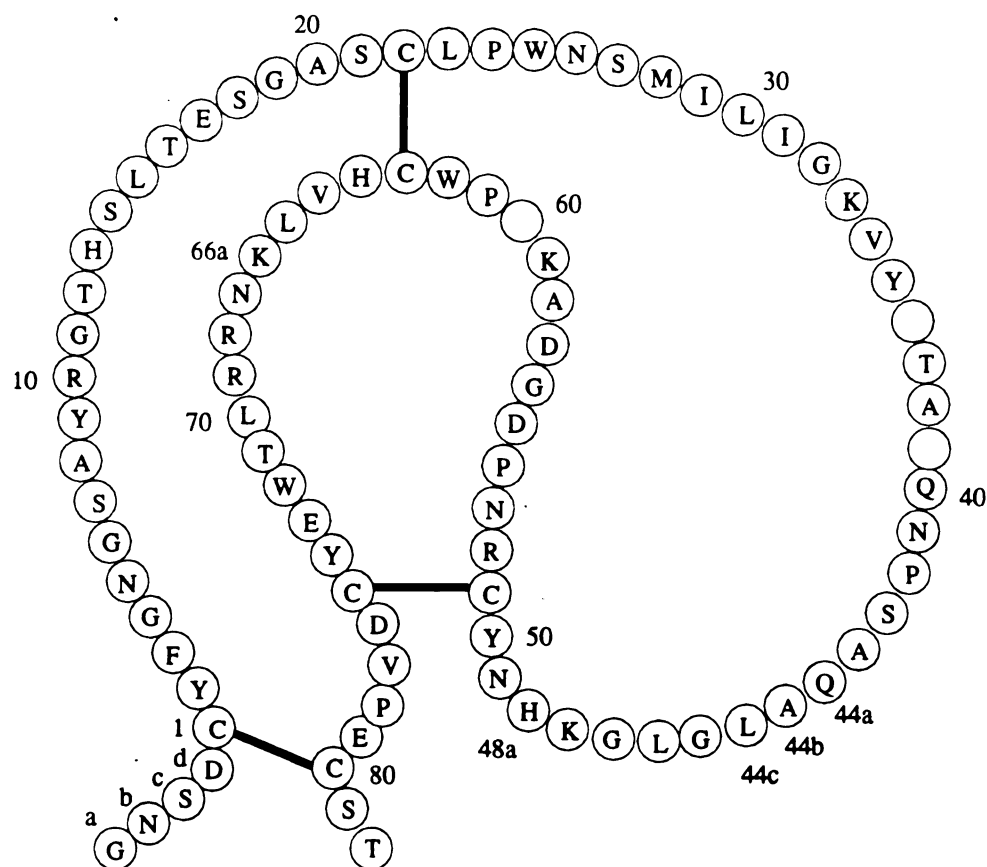


Figure 5.2. Primary structure of tPA kringle 2. Site numbering insertions and deletions are based on homology with PGK5.

initial phases for tPAK2 by using PGK4 as a model. However, due to three molecules in the asymmetric unit, the translation search proved to be unsuccessful. The structure of tPAK2 showed a similar overall folding to that of PGK1 and PGK4 except for an α -helix in the tPAK2 structure. The crystal structure also showed a strong interaction between the lysine residue of one molecule and the lysine binding site of a non-crystallographically related neighboring molecule. The lysine pocket observed in this structure is similar to that of PGK1 and PGK4. The negatively charged side chains of Asp55 and Asp57 were proposed to be involved in the interaction with the positively charged amino group of the ligand, and the hydrophobic surfaces of Trp62 and Trp72 were thought to form van der Waals contacts with the methylene groups of the ligand. However, the important positively charged residue in this structure was found to be Lys33, whose amino group is virtually positionally identical to that of the guanidinium group of Arg71 in the PGK4 structure.

In addition to the X-PLOR refined structure, a solution structure of tPAK2 complexed with 6-aminohexanoic acid was determined by NMR spectroscopy and dynamical simulated annealing calculations [65]. This work showed that the structure of tPAK2 has a compact globular conformation characterized by a number of turns as well as by one right-handed α -helix and five antiparallel β -sheets. The α -helix is formed by the Ser43-Gly45 segment, which contains three insertions in the tPAK2 (Figure 5.2). The binding site was also defined by two negatively charged residues, Asp55 and Asp57, and an aromatic pocket lined by Tyr36, Trp62, His64 and Trp72 side chains. The positively charged side chain of Lys33 and Arg69 may favor interactions with the carboxylate group of the ligand in the structure.

This chapter describes the refinement of the tPAK2 structure by the restrained least-squares method implemented by the program "PROSLQ" with the diffraction data (2.43 Å resolution) used for the X-PLOR refinement. The final 2.43 Å resolution structure from the X-PLOR refinement was used as the initial model. A comparison

of the final "PROSLQ" and "X-PLOR" structures revealed that both showed similar lysine binding sites, and the rms deviation on the main-chain atoms is as small as 0.18 Å.

5.2 Experimental Procedure

A. X-PLOR Structure

Crystals of tPAK2 were grown by the vapor diffusion method by using 4mg/ml protein, 7% saturated NH_4Cl , and 50 mM NH_4HCO_3 , at pH 8.0. Crystals of tPAK2 are monoclinic, space group $\text{P}2_1$; the unit cell with dimensions $a = 54.80$ Å, $b = 63.58$ Å, $c = 46.58$ Å, $\beta = 106.73^\circ$ has six molecules per unit cell with three molecules per asymmetric unit. Three dimensional intensity data were originally measured with a Rigaku AFC6 diffractometer and $\text{CuK}\alpha$ radiation at 9kW power (50kV, 180mA). Three crystals were used to collect 2.8 Å resolution data with $R_{\text{merge}}(\text{I})=0.087$. The initial structure of tPAK2 was determined by the technique of multiple isomorphous replacement. The heavy atom derivatives used for crystal soaking were 6 mM, $(\text{NH}_4)_2\text{PtCl}_4$, 10mM of K_2AuCl_4 , or a 200-fold dilution of a saturated KI solution. A 3.0 Å intensity data set was collected for each of the derivatives. The resulting electron density maps clearly showed the boundaries for three molecules in the asymmetric unit and good connectivity for most parts of the main chain. The starting model used in the refinement was obtained by fitting the structure of PGK4 to the map, with corrections to the sequence of PGK4 to correspond to that of tPAK2. Refinement of the 10.0–2.8 Å data set was performed with the program X-PLOR [53]. The R-factor was reduced to 20% after three iterations of refinement and map fitting. The final R-factor was 17% between 10.0–2.8 Å resolution with 20 water molecules and three chloride ions (one in each molecule).

A much better set of 2.36 Å intensity data was collected later with the Siemens

area detector. The raw data were reduced and scaled with the XENGEN program [26]. A total of 30,526 observations were measured, of which 11,623 were independent ($R_{\text{sym}} = 0.068$). The previous structure at 2.8 Å resolution was used as a starting model for the refinement of the new set of data. Refinement proceeded in stages, initially including data at 10.0–2.8 Å, then at 2.43 Å resolution. A summary of the final refinement results is given in Table 5.1.

Table 5.1. A summary of the final refinement results of the X-PLOR refined structure.

Resolution (Å)	10.0–2.43
No. of reflections	9687
Crystallographic R	0.184
No. of protein atoms	2025
No. of water molecules	92
No. of chloride ions	3
RMS deviation in bond length (Å)	0.016
RMS deviation in bond angle (deg)	3.3
RMS deviation in torsion angles (deg)	26

B. PROLSQ Structure

The X-PLOR refined structure devoid of chloride ions and water molecules was taken as the initial model for the PROLSQ refinement. The R value started at 0.26 with an overall thermal parameter of 16 Å², and R decreased to 0.20 after the first stage (8.0–2.8 Å). Chloride ions were added after 32 cycles in accordance with the $|F_o| - |F_c|$ map. As phases improved, the remainder of the data (to 2.38 Å) were included, and reflection weights were assigned in seven shells of $\sin \theta / \lambda$ based on $(||F_o| - |F_c||) / 2$ of the range. Since three molecules were in an asymmetric unit, non-crystallographic symmetry restraints were applied between main-chain and side-chain atoms of the

three independent molecules. The final reflection weights and the R values in each range are shown in Table 5.2, and a summary of the refinement parameters is listed in Table 5.3. The final tPAK2 structure has a crystallographic R value of 0.145 for 8827 reflections between 8.0–2.38 Å resolution with 203 water molecules and three chloride ions and an average thermal parameter of 21.7 Å².

5.3 Results and Discussion

A. General Structure

The refined tPAK2 structure consists of 3 molecules in the asymmetric unit; nearly all three molecules are well defined and extend from Ser-c to Ser81. Three chloride ions were confirmed in the refined tPAK2 structure, one in each molecule. The presence of chloride ions agrees with the observation that even relatively low concentrations of chloride can precipitate the protein, and in fact, crystals can not be obtained in the absence of the chloride ion [55]. The three chloride ions, which have well-defined electron densities (Figure 5.3) are at full occupancies and have low thermal parameters (~ 16 Å²). The average occupancy of the 203 water molecules is about 0.76, and their average thermal B value is 25.6 Å², slightly higher than that of the protein (21.3 Å²). Of the 203 water molecules, two are internal and conserved in the three independent molecules. One of the the two water molecules is buried within the tPAK2 structure and bridges the main chain Leu23 O and His64 N atoms at distances of 2.7 Å and 3.0 Å, respectively (Figure 5.4). Interestingly, this water molecule was also observed in the PGK1 (Chapter 4) and PGK4 [41] structures. The other internal water molecule was found in an empty pocket of the protein structure, that is removed from, but accessible to the bulk solvent. It bridges three main-chain atoms (Gly6 O, Tyr9 O, and Asn53 N) and serves as a hydrogen donor and acceptor (Figure 5.5).

Although tPAK2 has four more residues than PGK4, it has a globular conforma-

Table 5.2. Weights of reflections and R values of the final refinement cycle of tPAK2

$d_{\min}(\text{\AA})$	no. reflections	$\sigma(F)^a$	$< F_o - F_c >$	R value	
				shell	sphere
4.56	1142	23	47	0.172	0.172
3.75	1301	19	35	0.126	0.147
3.32	1299	17	31	0.129	0.142
3.01	1332	15	26	0.148	0.143
2.78	1309	13	22	0.149	0.144
2.60	1195	12	20	0.155	0.145
2.20	1249	10	17	0.151	0.145
a. $\sigma(F) = 26 - 258 [(\sin \theta/\lambda) - (1/6)]$.					

Table 5.3. Summary of final least-squares parameters and deviations of tPAK2

	target (σ)	rms (Δ)
Distances ($\text{\AA}$)		
Bond Distance	0.020	0.015
Angle Distance	0.040	0.046
Planar 1,-4 distance	0.060	0.055
Non-bonded distances ($\text{\AA}$)		
Single torsion	0.55	0.20
Multiple torsion	0.55	0.24
Possible H-bond	0.55	0.23
Torsion angles (deg)		
Planar	3	2
Staggered	15	20
Orthonormal	20	23
Plane groups ($\text{\AA}$)	0.02	0.01
Chiral centers ($\text{\AA}^3$)	0.15	0.15
Thermal restraints ($\text{\AA}^2$)		
Main chain bond	1.5	0.9
Main chain angle	2.0	1.5
Side chain bond	2.0	1.4
Side chain angle	2.5	2.0
Non-Crystallographic Symmetry ($\text{\AA}$)		
Main chain atoms	0.40	0.26
Side chain atoms	0.80	0.45

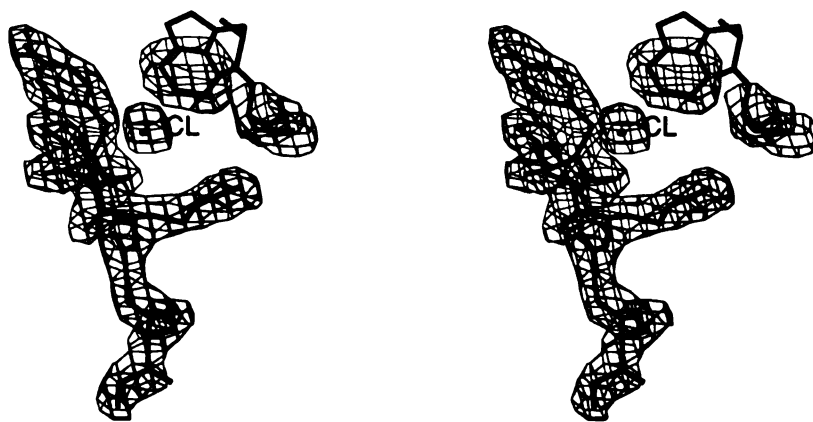


Figure 5.3. Electron density of chloride ion in the lysine binding site of tPAK2. Contour at 1 σ .

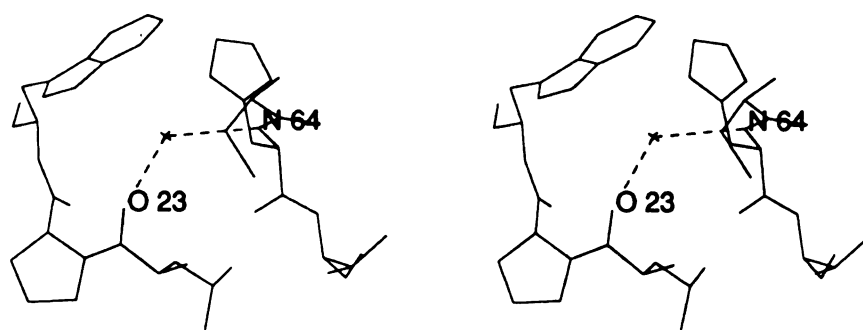


Figure 5.4. Stereoview showing an internal water molecule in the tPAK2 structure. Hydrogen bonds are represented by dashed lines.

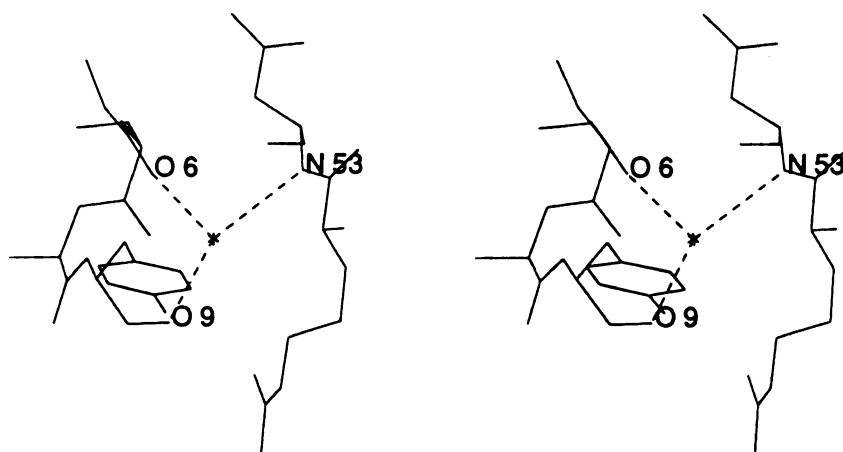


Figure 5.5. Stereoview showing the water molecule found in an empty pocket of the tPAK2 structure. Hydrogen bonds are represented by dashed lines.

tion similar to that of PGK4. The three-dimensional main chain folding is shown in Figure 5.6, from which it can be seen that the overall structure of the kringle is related to the close contact between the inner loop disulfide groups of Cys22–Cys63 and Cys51–Cys75, which orient nearly perpendicular to each other and serve as a core of the overall folding. The secondary structural elements present in the tPAK2 molecule are listed in Table 5.4. The four antiparallel β -sheets and seven β -turns observed in the crystal structure are consistent with the NMR results [66], with the exception of the Gly6–Tyr9 turn. In addition, one α -helix turn was found in the Thr37–Ala38–Gln40–Asn41–Pro42–Ser43–Ala44–Gln44a–Ala44b–Leu44c segment, which contains a Pro residue and three insertions. The distortion of the α -helix that results from insertion of a Pro residue in this helix region was also observed in this structure.

B. Crystal Packing

The tPAK2 molecule crystallized in the monoclinic system, space group $P2_1$. As such, a crystallographic two-fold screw axis along b resides in the crystalline structure. However, since there are three molecules in the asymmetric unit, a non-crystallographic 3-fold screw axis nearly parallel to the crystal a -axis is also present. As is shown in Figures 5.7 and 5.8, the three molecules in the asymmetric unit are related to each other by an approximate 3_1 screw axis. The most significant interaction among the three molecules is the ligand-like binding interaction which occurs in the lysine binding sites of the molecules (Figure 5.9). The NZ of Lys48' of a neighboring molecule hydrogen bonds to the negatively charged Asp55 OD2 and Asp57 OD2 atoms at a distance of 2.8 Å. The carbonyl oxygen of Lys48' is 2.6 Å from the amino group of the positively charged Lys33 side chain. The other intermolecular interaction found between asymmetric units is a solvent-bridge between the Asp55 OD2 and Asn26' ND2 of the neighboring molecule (Figure 5.10). This water molecule makes hydrogen bonds to the Asp55 OD2 and Asn26' ND2 at distances of 2.6 Å and 2.9 Å,

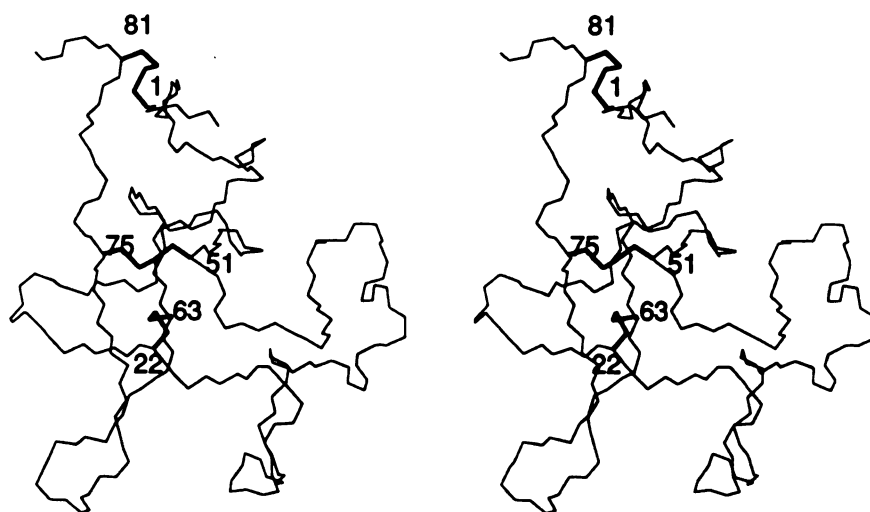


Figure 5.6. Stereoview of CA, C, and N structure of tPAK2. Disulfides are shown in bold.

Table 5.4. Secondary structural elements of the tPAK2 structure.

β Structures	Type	Residues
$\beta 1$	Antiparallel	Ser14 – Cys22 Leu15 – Ser21 Thr16 – Ala20
$\beta 2$	Antiparallel	Pro24–Asn49 Trp25–His48a
$\beta 3$	Antiparallel	Arg52–Trp62 Asn53 – Pro61
$\beta 4$	Antiparallel	Cys63 – Cys75 His64 – Tyr74 Val65 – Glu73 Leu66 – Thr72
Reverse Turns	Type	Residues
T1	Type II'	Phe3–Gly4–Asn5–Gly6
T2	Type I	Gly6–Ser7–Ala8–Tyr9
T3	Type I	Thr16–Glu17–Ser18–Gly19
T4	Type I	Pro24–Trp25–Ser26–Ser27
T5	Type III	Ser27–Met28–Ile29–Leu30
T6	Type II	Leu30–Ile31–Gly32–Lys33
T7	Type I'	Lys66a–Asn67–Arg68–Arg69
α -Helix		Thr37–Leu44c



Figure 5.7. Stereoview of the crystal packing of the three molecules in the asymmetric unit of the tPAK2 structure. The non-crystallographic 3_1 screw axis is perpendicular to the plane.



Figure 5.8. Stereoview of the crystal packing of the three molecules in the asymmetric unit of tPAK2 structure; The non-crystallographic 3₁ screw axis is indicated.

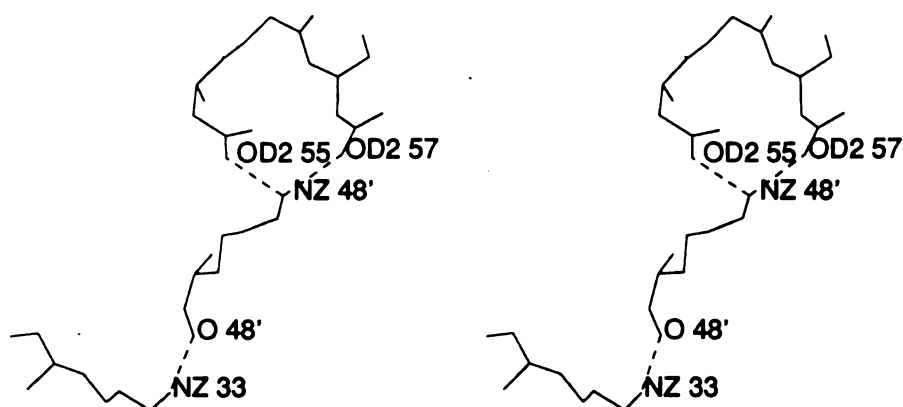


Figure 5.9. Stereoview showing the intermolecular interactions between Lys48' and Lys33, Asp55, and Asp57. Prime represents the neighboring molecule.

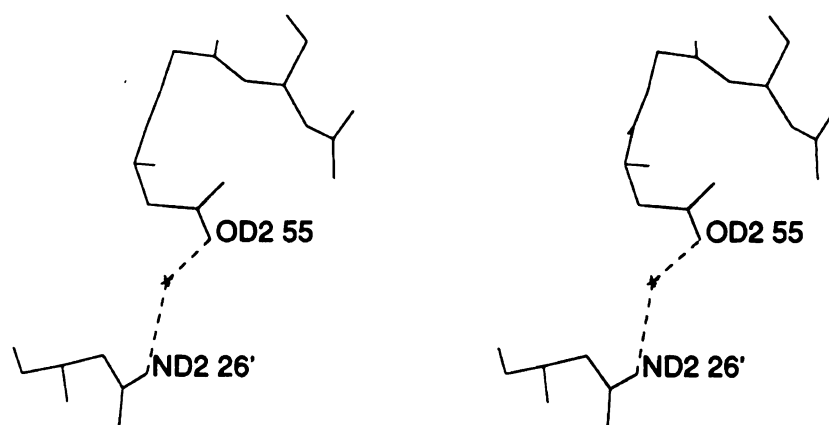


Figure 5.10. Stereoview showing the intermolecular interaction between Asn26' and Asp55 bridged by a water molecule. Prime represents the neighboring molecule.

respectively. Not surprisingly, the strong interaction along the non-crystallographic 3-fold screw axis (nearly parallel to the crystal a-axis) enhances crystal growth along the a-axis [55] (Figure 5.11).

C. Structure of the Lysine Binding Site

The lysine binding site in the tPAK2 structure is located on the kringle surface and is supported mainly by the inner kringle loop, which consists of central section of the loop B and the central parts of loops C and D (Figures 1.2 and 5.12). The binding site residues can be divided into three groups. The negatively charged residues Asp55 and Asp57, with carboxylate oxygens 4.3 Å from one another, serve as an anionic center at one end of the binding site. An elongated depression, lined by the indole rings of Trp62 and Trp72, provides a highly nonpolar environment for the methylene groups of the ligand. The two indole rings are oriented in an antiparallel manner with an interplanar angle of approximately 80° and form aromatic stacking interactions with the rings of His64 and Tyr74, respectively. Finally, a cationic center is formed by the side chain of the positively charged residue Lys33. The lysine binding site not only is stabilized by the symmetric structural framework but also by a number of inter- and intra-molecular interactions (Figure 5.13, Table 5.5). The OD2 atoms of Asp55 and Asp57 are 2.8 Å from the Lys48 NZ atom of the neighboring molecule. In addition, OD1 atoms of Asp55 and Asp57 hydrogen bond back to the protein with Trp62 NE1 and Tyr74 OH at distances of 3.0 Å and 2.4 Å, respectively. The carbonyl oxygen of Lys48 of the adjacent molecule forms a hydrogen bond with the positively charged Lys33 side chain at a distance of 2.6 Å. The chloride ion, which is found in the interface between two molecules, interacts with the main chain amide atoms of Val34, Tyr35 and His64 with Cl to N distances of 3.2 Å, 3.2 Å, and 4.0 Å, respectively. The chloride ion not only makes these three intramolecular interactions, but also forms an intermolecular interaction with ND2 of Asn26' of the next kringle. Two solvent-

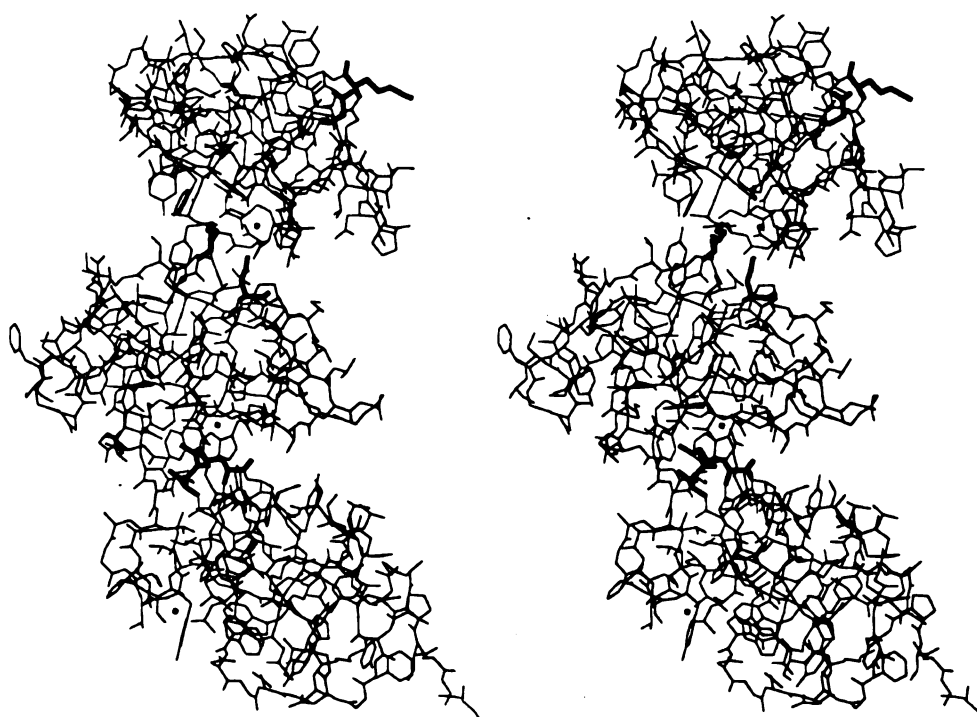


Figure 5.11. Stereoview of the crystal packing of the three molecules along the *a*-axis. Residues involved in intermolecular interactions are shown in bold.

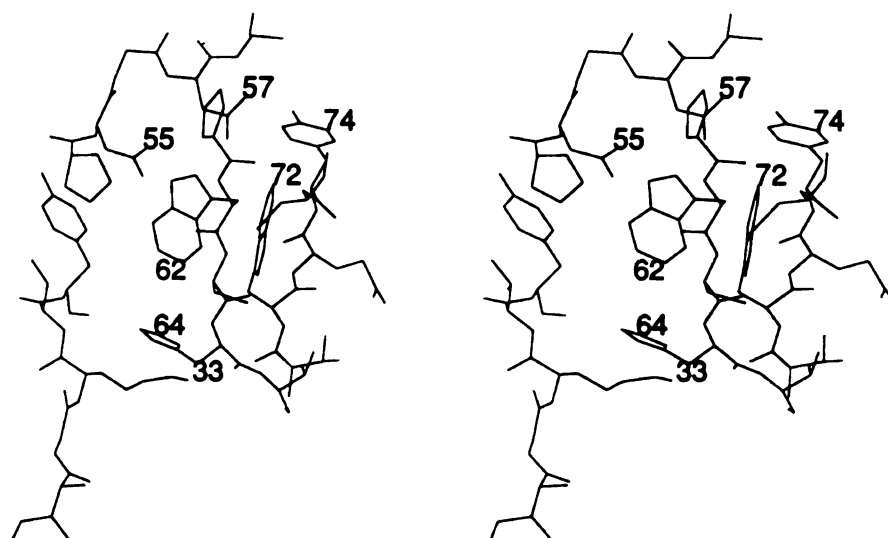


Figure 5.12. Stereoview of the lysine binding site of the tPAK2 structure.

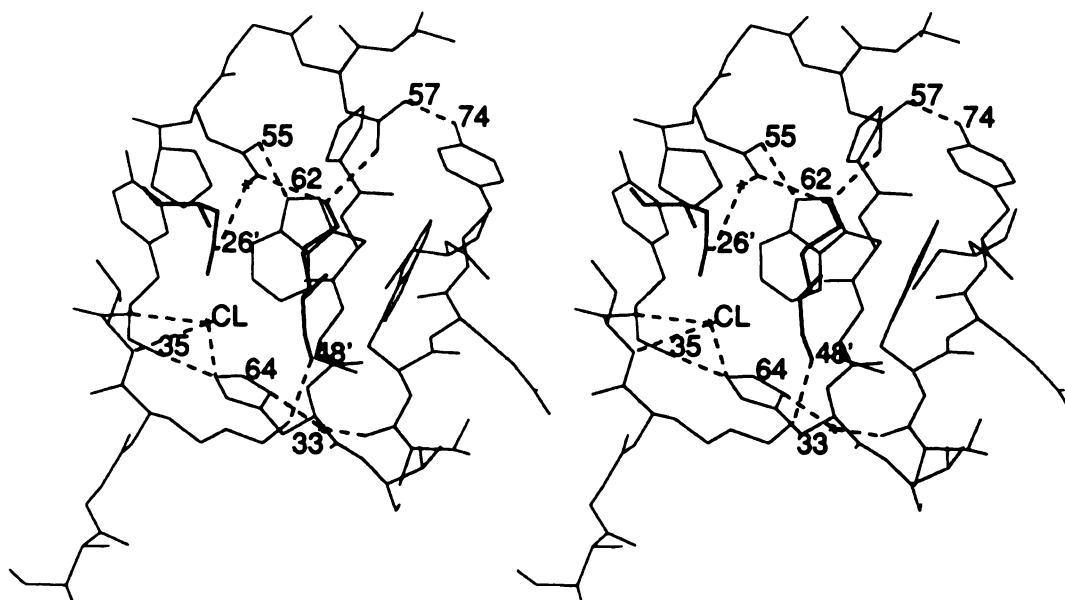


Figure 5.13. Intramolecular and intermolecular interactions present in the lysine binding site of the tPAK2 structure.

bridged interactions were found between His64 ND1 and Thr71 O, Asp55 OD2 and Asn26' ND2 of the neighboring molecule. Interestingly, the two water molecules are conserved in all three molecules of the asymmetric unit with unit occupancies and average thermal parameters of only 16.0 Å².

Table 5.5. The intramolecular and intermolecular interactions present in the lysine binding site of tPAK2 structure. Prime refers to neighboring molecules. Donor atom is denoted (D), acceptor atom (A).

Donor	Acceptor	D...A (Å)
Tyr74 OH	Asp57 OD1	2.40
Trp62 NE1	Asp55 OD1	3.04
Lys48' NZ	Asp55 OD2	2.82
Lys48' NZ	Asp57 OD2	2.80
W386 O	Asp55 OD2	2.61
Asn26' ND2	W386 O	2.92
Val34 N	CL	3.18
Tyr35 N	CL	3.20
His64 NE2	CL	4.00
His64 ND1	W339 O	3.01
W339 O	Thr71 O	2.32
His64 NE2	Tyr35 O	2.55
Lys33 NZ	Lys48' O	2.55

NMR NOE experiments [65, 67] found that the side chains of Tyr36, Trp62, His64, Trp62 and Trp72 were perturbed most by ligand presence. The Trp25 and Tyr74 aromatic rings were also shifted, but not to undergo direct contacts with the ligand. In the crystal structure, the side chain of His64 is approximately 3.9 Å from the end of the Trp62 indole ring; thus, it appears that the imidazole ring is affected through the aromatic stacking interaction with the side chain of Trp62. The tyrosyl ring of Tyr74, which is 3.7 Å from the edge of the indole ring of Trp72 and has a hydrogen bond between its hydroxyl atom and the carboxylate group of Asp57, may

sense ligand-binding effects indirectly from the aromatic interaction with Trp72 or the hydrogen-bonding interaction with Asp57. The indole ring of Trp25 is in the layer below the surface with its end about 3.8 Å from the face of the indole ring of Trp62; thus, it may sense the presence of the ligand through an aromatic-aromatic interaction with Trp62. Although the side chain of Tyr36 appears to be involved in ligand-binding in the NMR NOE experiments, there is no obvious direct interaction between Tyr36 and Lys48' in the crystal structure (Figure 5.13). Therefore, the side chain of Tyr36 may experience the presence of the ligand from secondary effects or in the case of the bulkier ligands. Although the lysine binding site of the tPAK2 is blocked by Lys48' from the neighboring molecule, the lysine binding site structure agrees well with the NMR observation results. Thus, the intermolecular binding interaction mimics a ligand-like binding site interaction and might be useful as a model for the tPAK2-fibrin interaction.

D. Comparison of the Three Molecules in the Asymmetric Unit

The three molecules in the asymmetric unit have been compared, and the r.m.s. deviations are listed in Table 5.6. From a superposition of the three molecules (Figures 5.14, 5.15), it can be seen that the agreement of the main-chain atoms among the three molecules is quite good and almost all the side-chain groups have a similar conformation, except His13, Met28, and Arg68. In molecules A and C, the His13 side chain has the same conformation with $\chi_1 = -63^\circ$ and $\chi_2 = -81^\circ$, while in molecule B, $\chi_1 = -148^\circ$ and $\chi_2 = -86^\circ$. Figure 5.16 shows that the His13 ND2 makes a hydrogen bond with the Tyr9 OH at a distance of 2.8 Å in molecules A and C, but not in B. This observation is consistent with the NOE results and acid/base titration experiments [67], which found two side chain positions for His13, one located on the

kringle surface (as in molecule B) and the other making intramolecular interactions with other protein groups (as in molecules A and C). The residue of Met28, which is located on the kringle surface and exposed to the bulk solvent, shows different conformations in all three molecules (Figure 5.15). This difference results in the large r.m.s. deviations of sulfur atoms (Table 5.6). Residue Arg68 shows different side-chain conformations in molecules A and B (Figure 5.15). Inspection of this region in the crystal reveals that the side chains are involved in intermolecular interactions in molecules A and B, but not in C. In molecule C, no electron density was observed for side chains of the Asn67-Arg68-Arg69 segments which were refined as glycine residues.

Table 5.6. RMS deviations among the three molecules in the asymmetric unit. MA, MB, MC represent molecules A, B and C. Numbers shown in parentheses were calculated without including Met28.

	MA & MB (Å)	MA & MC (Å)	MB & MC (Å)
All protein atoms	0.83	0.49	0.68
Main chain	0.34	0.30	0.40
Carbonyl oxygens	0.45	0.42	0.53
Side chains	1.16	0.63	0.88
Sulfurs (Cys, Met)	0.82 (0.54)	0.70 (0.32)	1.43 (0.70)
Carbon alphas	0.35	0.31	0.41

E. Comparison of the Structures of Lysine Binding Sites of tPAK2 and K4-ACA

In order to compare the lysine binding sites of tPAK2 and K4-ACA, an optimal superposition of the two structures based on the main chain atoms CA, C, N was obtained. The r.m.s. deviations of residues in the lysine binding sites of the two

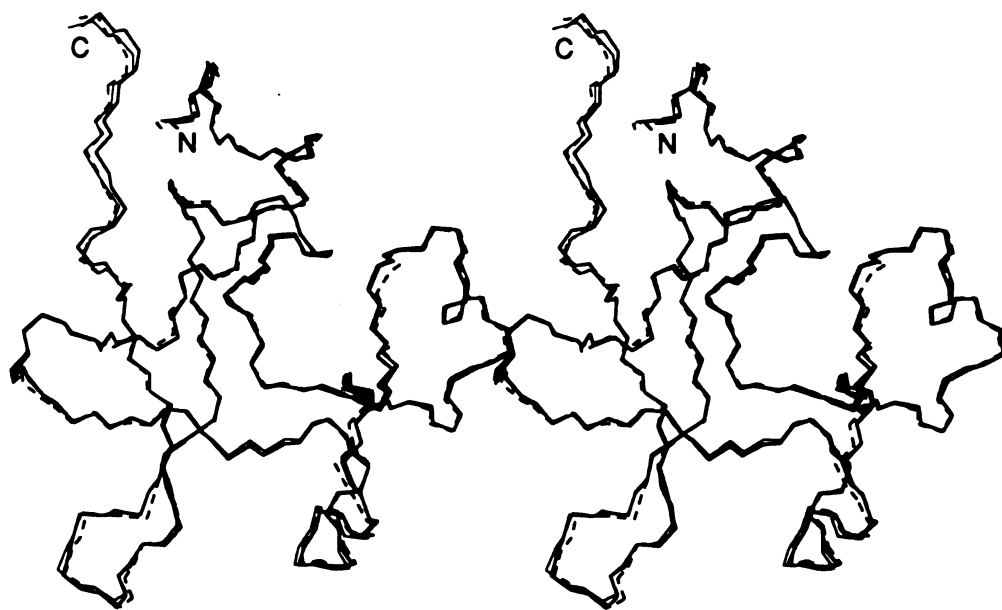


Figure 5.14. Stereoview of the comparison of the main-chain atoms of the three independent molecules.

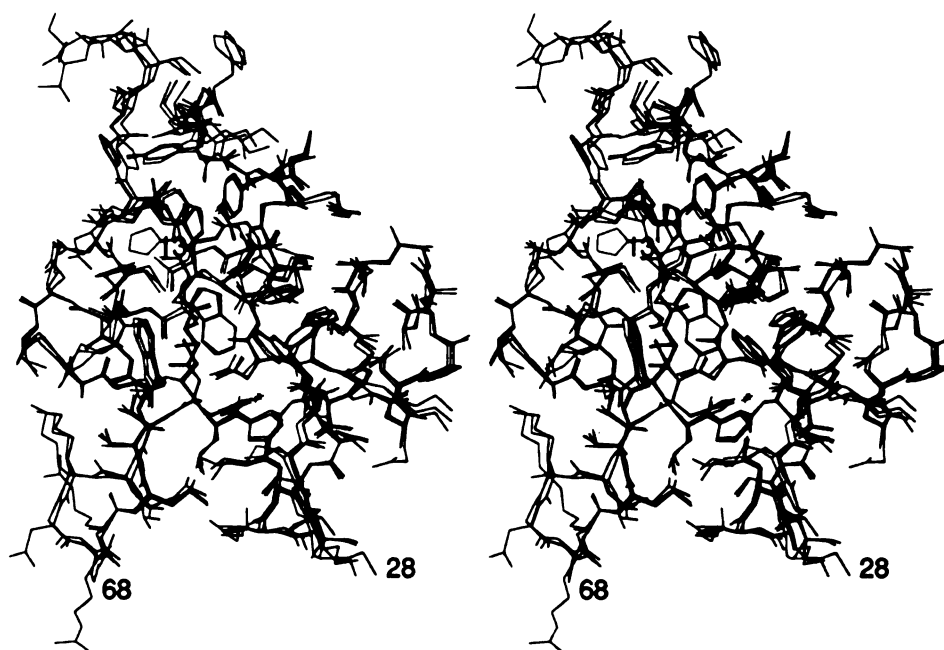


Figure 5.15. Stereoview of the comparison of the side-chain atoms of the three independent molecules.

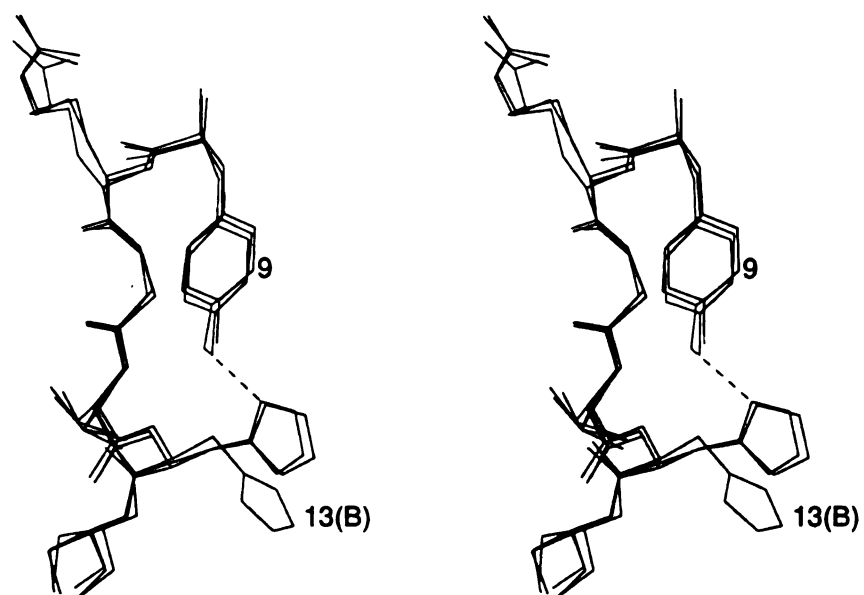


Figure 5.16. Stereoview showing the comparison of side chains of His13 in three molecules. His13 in molecule B is indicated.

structures are listed in Table 5.7. A superposition of lysine binding sites of the tPAK2 and K4-ACA structures is shown in Figure 5.17. An examination of Figure 5.17 reveals that although the Ile31-Tyr35 segment of tPAK2 is different from that of K4-ACA, the lysine binding site structures of the two kringles are in general similar, with an anionic center formed by Asp55 and Asp57, an aromatic groove conformed by the indole rings of Trp62 and Trp72, and a cationic center provided by the amino group of Lys33 in tPAK2 or the guanidino group of Arg71 in K4-ACA. Comparison of the main-chain structures in the lysine binding site gives a r.m.s. deviation of 1.5 Å; and the difference reduces to 0.6 Å when atoms with deviations greater than 1σ are removed. Not surprisingly, the r.m.s. deviation in the B loop is extremely large. This significant structural difference results from the different number of residues of the B loop in the two structures; in which tPAK2 has three more than PGK4. Except for the 31-35 peptide segment, good agreement was observed between the two structures for the rest of the lysine binding sites, especially the residues which are conserved. A comparison of the position of Lys48' and that of ACA shows that both are ducked in the aromatic groove with their methylene groups making van der Waals contacts with indole rings of Trp62 and Trp72. One point regarding the cationic centers of tPAK2 and PGK4 is noteworthy: on PGK4, Arg71 is the residue that interacts with the carboxylate group of the ligand, whereas, in tPAK2, the arginine residue at position 71 is lacking but has the Lys33 side chain at the cationic center in the lysine binding site. From Figure 5.17, it can be seen that the position of the amino group of Lys33 in the tPAK2 structure is virtually identical to that of the guanidinium group of the Arg71 in the PGK4 structure. *Thus, it provides an interesting example of structural but not sequential homology.*

Table 5.7. Deviations of the positions of residues in the lysine binding site between tPAK2 and K4-ACA

residue number	tPAK2	K4-ACA	rms $\Delta(\text{\AA})$ (main chain)	rms $\Delta(\text{\AA})$ (side chain)
31	I	H	5.60	9.60
32	G	R	4.30	—
33	K	H	1.14	1.06
34	V	Q	0.80	1.82
35	Y	K	1.35	0.91
54	P	P	0.57	0.59
55	D	D	0.63	1.19
56	G	A	0.64	—
57	D	D	0.50	1.40
58	A	K	0.55	1.71
61	P	P	0.19	0.76
62	W	W	0.16	0.44
63	C	C	0.22	0.46
64	H	F	0.49	0.69
71	T	R	0.70	3.70
72	W	W	0.46	0.96
73	E	E	0.33	0.98
74	Y	Y	0.45	0.54
75	C	C	0.64	0.34

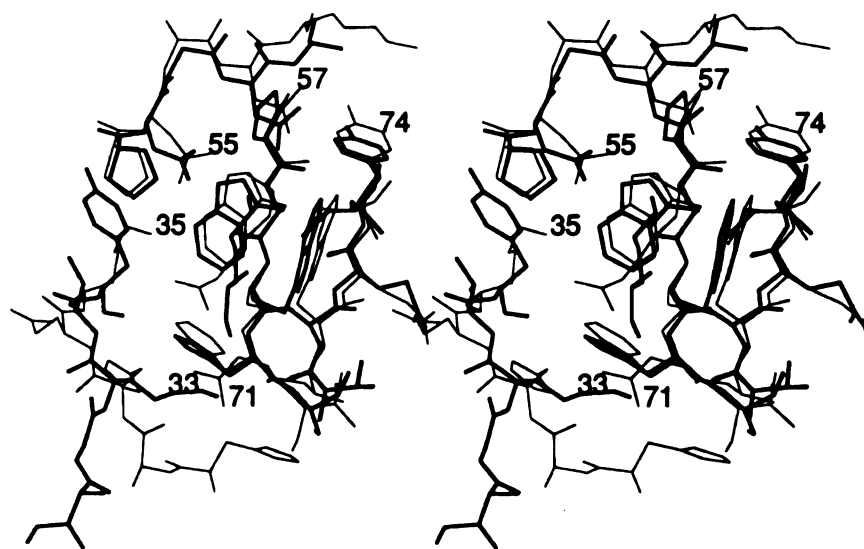


Figure 5.17. Stereoview of superimposed lysine binding sites of tPAK2 (bold) and K4-ACA. Lys48' is between Asp55, Asp57 and Lys33 in tPAK2; ACA is between Asp55, Asp57 and Lys35, Arg71 in K4-ACA.

CHAPTER 6

Comparison of Different Kringles

In order to ascertain the structural changes of kringles, all the kringle structures that have been solved are compared and shown in Figure 6.1, which includes prothrombin kringle 1 (PTK1, 2.25 Å resolution) [22], prothrombin kringle 2 (PTK2, in PPACK-thrombin complex at 3.3 Å resolution), [56], plasminogen kringle 1 (PGK1, 2.38 Å resolution), plasminogen kringle 4 (PGK4, 1.9 Å resolution) [23], tissue plasminogen activator kringle 2 (tPAK2, 2.43 Å resolution). From Figure 6.1, three different kinds of kringle folding (thin, bold, and dashed lines) are observed and the largest deviation among them is in the B and D loops. Except for these two regions the five kringles show nearly identical conformations in their A and C loops having a r.m.s. deviation of 0.5 Å. This is consistent with the high degree of homology in the A and C loops (40%) of kringles from different proteins (Table 1.1). As mentioned in Chapter 4, the different folding between tPAK2 and PGK4 is due to four insertions in the former structure, in which three are in the B loop and one is in the D loop. These add a helical turn to the conformation of the B loop of the tPAK2 structure and change the conformation of the loop dramatically. It is somewhat surprising that the PTK2 fold differs markedly from those of the four kringles. It has a distorted two-turn helix at Ala28-Lys35 and the hairpin β -turn of the D loop pivots as a unit about 60° at Val65 and Asp71. The two-turn helix in the B loop might be induced by a close approach

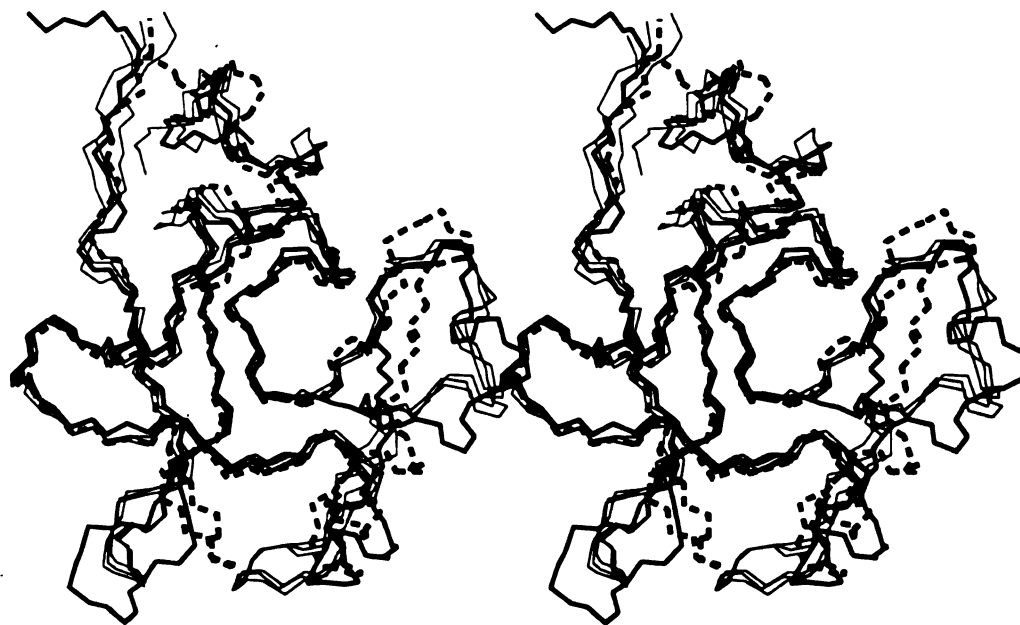


Figure 6.1. Stereoview of the comparison of the CA, C, N structures of different kringles. PGK1, PGK4, and PTK1 are shown in thin lines; tPAK2 in bold; PTK2 in dashed.

to the C-terminal helix of the B-chain of thrombin. The residues in the D loop of PTK2 make 18 van der Waals contacts less than 4.0 Å with thrombin. Since the PTK2 structure was solved as the PTK2-PPACK-thrombin complex, the features of the B and D loops of PTK2 could be either due to the complexation or inherent to the kringle fold. If the former, the conformational change may be necessary to maintain the structure of the complex. The structure of urokinase type plasminogen activator kringle 1 (uPAK1), which was determined by NMR [57], reveals two helix turns in the kringle. It is noteworthy that one of the two helices corresponds to the one observed in tPAK2 (Ser40-Gly45); the other one is in the region of Asn26-Gln33, which corresponds to a helical turn in PTK2. Although the three kinds of kringles have different conformations on their B and D loops, they all share a similar overall size with approximate dimensions of $15 \times 30 \times 30$ Å.

Although the overall kringle folds are not absolutely the same for the five kringles, the conformations of the central disulfides, which are perpendicular to each other, are very similar, with r.m.s. deviation of 0.25 Å (Figure 6.2). The disulfide cluster has close intramolecular contacts between Cys22-Cys63 and Cys51-Cys75 and gives rise to two approximately perpendicular anti-parallel stretches of β -sheet (β_1 and β_2) which are conserved in these five kringles (Figure 6.3, Table 6.1). The conservation of the two anti-parallel β -sheets most likely results in a highly stabilized zone and serves to support the overall folding.

As is described in Chapters 3, 4 and 5, each of the kringle domains with affinities for ω -aminocarboxylic acids (includes PGK1, PGK4, and tPAK2) possesses a binding site composed of three distinct regions according to electrostatic properties. The first is a negatively charged region containing two negatively charged aspartic acid residues. The second is a distinct positively charged region due to one (in tPAK2) or two (in PGK1 and PGK4) side chain amino groups. The third region separates the other two and appears as a cleft between the oppositely charged regions. This hydrophobic

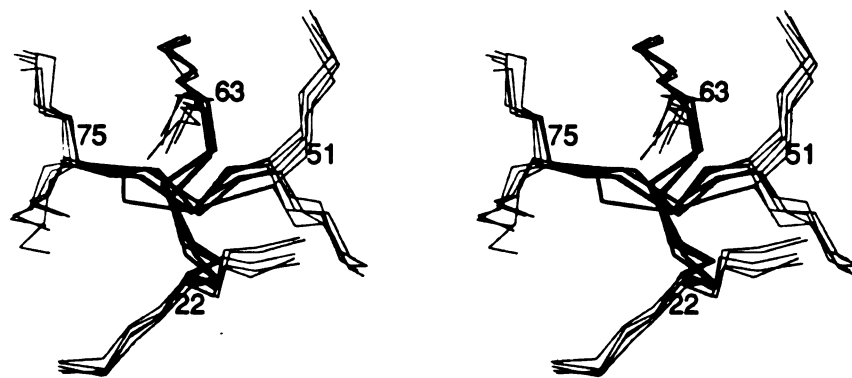


Figure 6.2. Stereoview of the comparison of the inner disulfide bridges of PGK1, PGK4, tPAK2, PTK1 and PTK2. Disulfide bonds in bold.

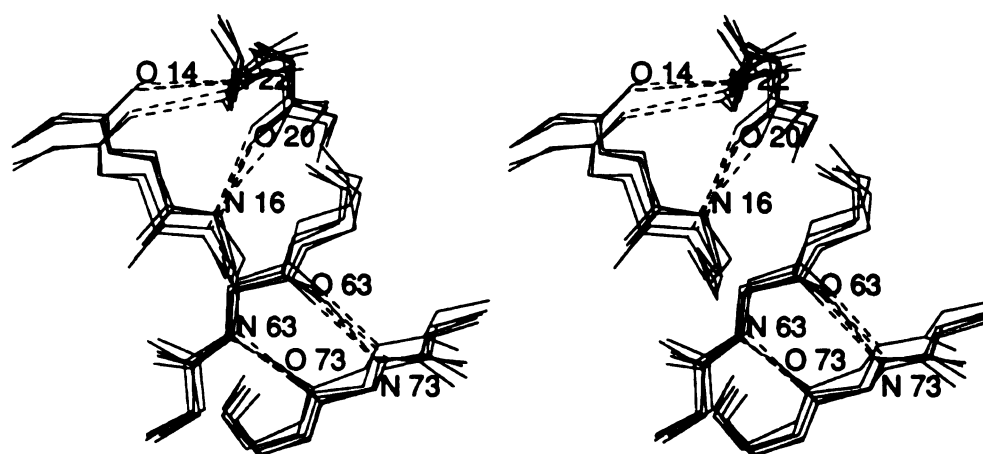


Figure 6.3. Stereoview showing the conserved anti-parallel β -sheets located near disulfide bridges in PGK1, PGK4, tPAK2, PTK1, and PTK2.

Table 6.1. Hydrogen bonds of the conserved disulfide anti-parallel β -sheets of PGK1, PGK4, tPAK2, PTK1, and PTK2 structures.

	Donor	Acceptor	N...O (Å)	N...O (Å)	N...O (Å)	N...O (Å)	N...O (Å)
			PGK1	PGK4	tPAK2	PTK1	PTK2
$\beta 1$	16 N	20 O	2.65	2.93	2.90	2.57	2.76
$\beta 1$	22 N	14 O	2.83	2.98	2.93	3.01	2.88
$\beta 2$	63 N	73 O	3.17	3.18	3.02	2.89	2.98
$\beta 2$	73 N	63 O	3.44	3.17	2.94	2.93	3.05

region is composed of aromatic residues, one of which is a Trp that is conserved even in non-binding kringle domains, and is the residue immediately preceding the fourth Cys of the kringle. Although many residues of the PGK4 lysine binding site are conserved in PTK1, it has no affinity to fibrin, lysine and analogous ligands. As seen in Figure 6.4, Arg71, which serves as a cationic center in PGK4, is conserved in PTK1, but is involved in an intramolecular ion pair interaction with Glu34 that blocks one end of the lysine binding site. Though Asp55, an essential residue of the anionic center of the PGK4 binding site, is also conserved in the PTK1 structure, it makes an ion pair interaction with Arg72 and precludes access to the site by lysine or other ω -aminocarboxylic acids. Furthermore, the substitution of the aromatic side chains of Trp72 and Tyr74 with charged Arg and Glu residues in PTK1 eliminates the large hydrophobic surface that binds methylene groups of ligands. In the PTK2 (Figure 6.5) structure, the two negatively charged Asp55 and Asp57 residues and one of the positive residues, Lys35 of PGK4, important for lysine binding, are also conserved. However, Lys31–Lys35 of PTK2 is in a helical conformation that causes Lys35 to project out into solvent region and disrupts this cationic center feature. Since

the structure of PTK2 was determined as the PPACK-thrombin complex, it is not clear whether the different conformation of this loop is inherent to the kringle, as in tPAK2 and the uPAK1, or whether it might be due to a conformational change upon complexation. Since PTK2 is generally thought not to bind lysine, it is consistent that it has a different native folding conformation. Further support for this inference is the observation of the same two-turn helix in the uPAK1 [57].

Nearly all the internal aromatic residues are conserved or are highly homologous in the five different kringles (Table 1.1). For instance, Tyr50 of tPAK2, PGK1 and PGK4 is replaced with Phe in PTK1 and PTK2. Similarly, His64 of tPAK2 is replaced by Phe in PGK4 and Tyr in PGK1, PTK1 and PTK2. Not surprisingly then, these aromatic residues have very similar conformations in the five different kringles (Figure 6.6), with an r.m.s. deviation of 0.45 Å. As mentioned in Chapter 1, the high degree of internal conservation around Cys22–Cys63 and Cys51–Cys75 maintains the hydrophobic core, which appears to be important for the three-dimensional kringle folding.

Thus, four types of kringle folding are known: (1) PGK1, PGK4, PTK1; (2) tPAK2; (3) PTK2 ; (4) uPAK1. Interestingly, the lengths of the loops of PTK1 and tPAK2 are identical to those of PTK2 and uPAK1, respectively, however, the kringle foldings of the former are different from those of the latter. It appears that kringles in general may have inherent flexibility in their three-dimensional folding.

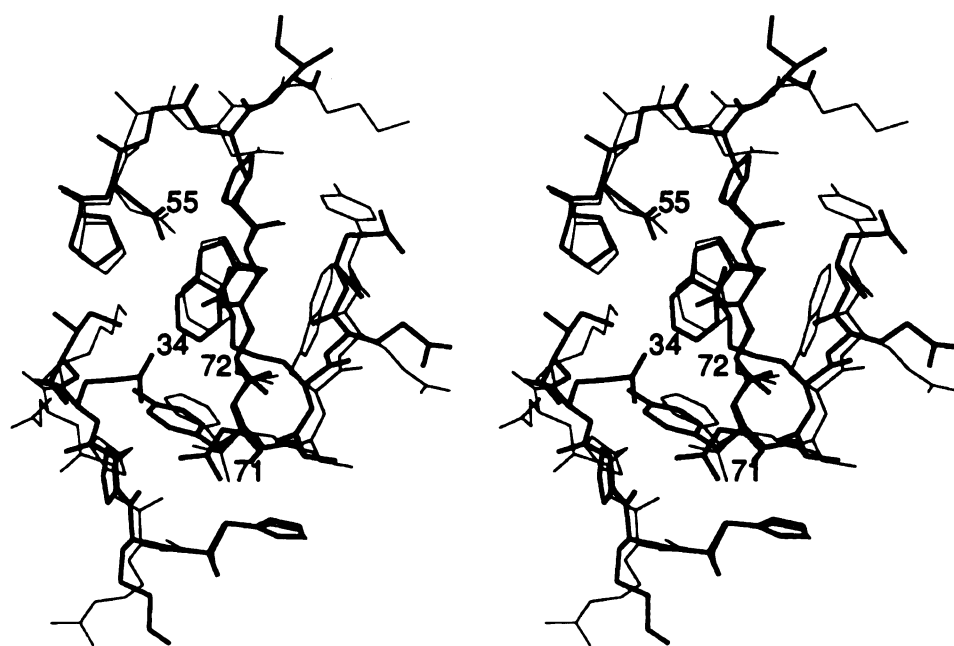


Figure 6.4. Stereoview comparing the lysine binding regions of PTK1 (bold) and PGK4.

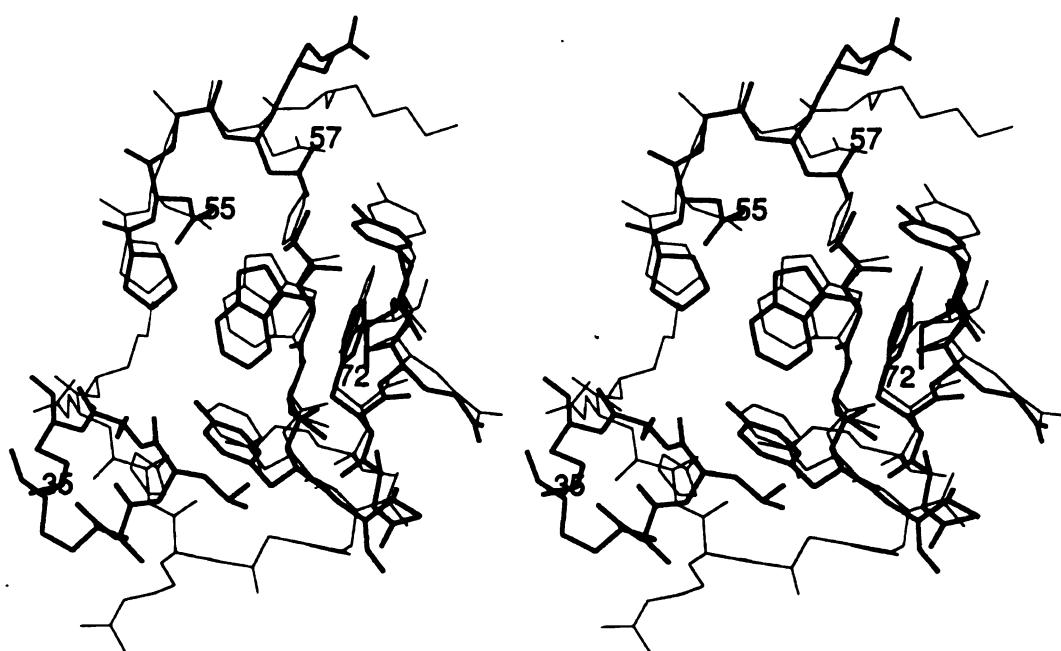


Figure 6.5. Stereoview comparing the lysine binding site regions of PTK2 (bold) and PGK4.

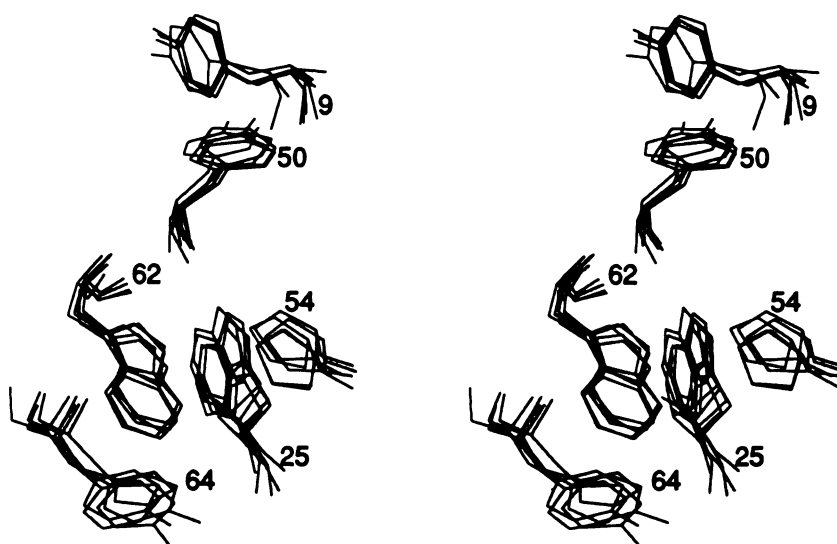


Figure 6.6. Stereoview of the comparison of internal aromatic residues in PGK1, PGK4, tPAK2, PTK1 and PTK2.

CHAPTER 7

The Structure of A Designed Peptidomimetic Inhibitor Complex of α -Thrombin

7.1 Introduction

The blood coagulation mechanism consists of a series of linked proteolytic reactions in which zymogens are converted into trypsin-like enzymes. The activation events take place on the surfaces of cells such as platelets, white blood cells, and endothelial cells. These transformations are accelerated by non-enzymatic protein cofactors that act either by altering the conformation of the zymogens or by binding converting enzymes and zymogens in close proximity on the surface.

α -Thrombin is a trypsin-like serine proteinase that plays a key role both in blood coagulation and in other physiological processes that involve catalytic functions and non-enzymatic intermolecular associations [68, 69, 70]. In the penultimate step of the blood coagulation cascade, thrombin is generated from prothrombin by limited proteolysis [71] (Figure 7.1); thrombin then transforms fibrinogen to fibrin. The human α -thrombin used in this study is a two chain molecule which possesses a molecular

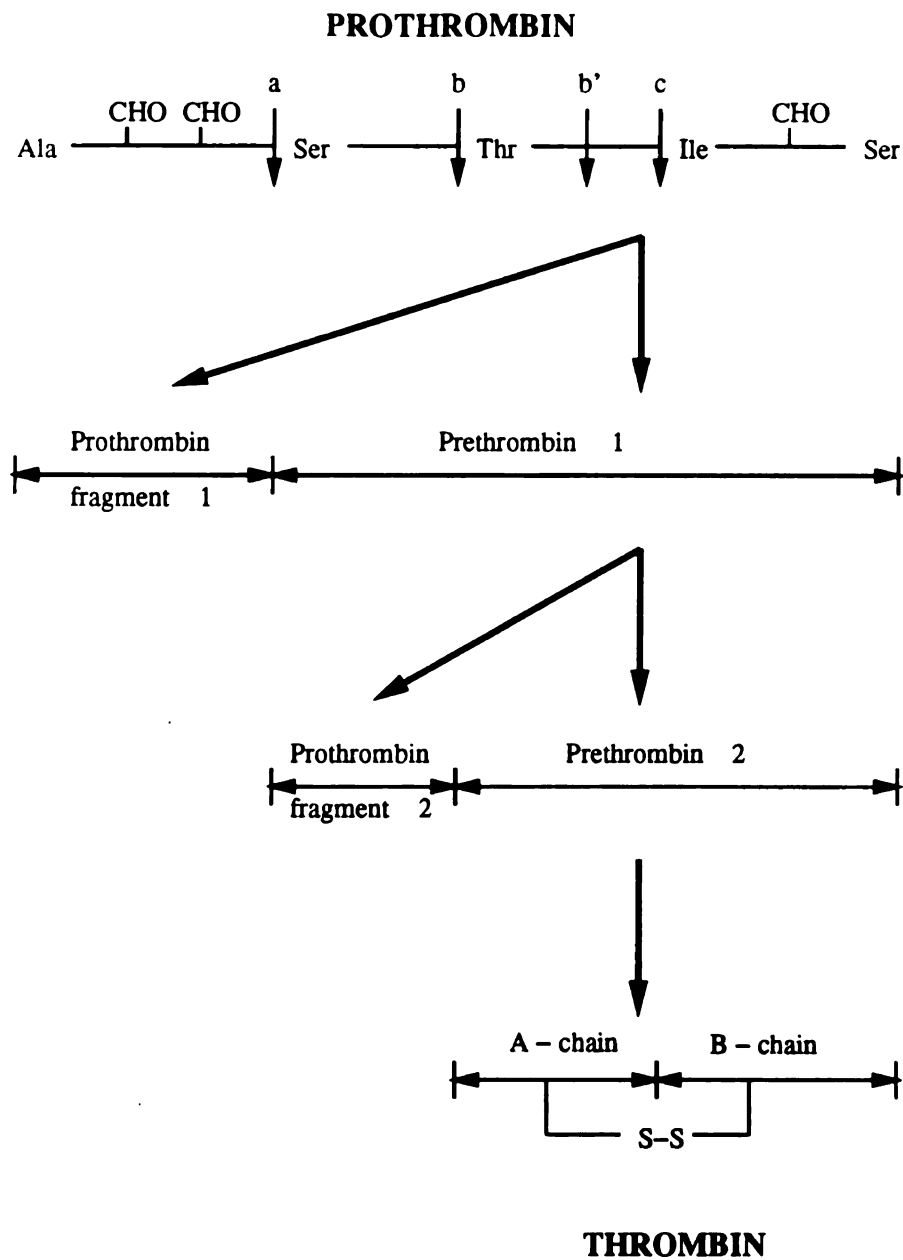


Figure 7.1. The conversion of prothrombin to thrombin. CHO represents a carbohydrate side chain; a, b, and c represent the cleavage sites of bovine and human prothrombin; b' represents an additional cleavage site in human prothrombin.

weight of approximately 36,600 [72, 73]. The A-chain consists of 36 residues linked via a disulfide bond to a B-chain of 259 residues at residues 1 and 122, respectively. The B-chain is highly homologous to other coagulation/fibrinolytic serine proteases (Factor IXa, Factor Xa, protein C, urokinase, tissue plasminogen activator, and plasmin) and contains the active site residues His57, Asp102, and Ser195. The sequence of human α -thrombin is listed in Table 7.1 [74].

The primary function of thrombin is to convert fibrinogen to fibrin, which is the major protein component of a blood clot and is responsible for mechanically binding together platelets and several plasma proteins into a network of fibrin polymers that block the flow of blood from a severed vessel. Fibrinogen is a disulfide-linked dimer of three peptide chains with stoichiometry ($A\alpha, B\beta, \gamma$) and that contain 625, 461, 411 amino acids, respectively (Figure 7.2) [75, 76]. Fibrin assembly begins with the cleavage of two $A\alpha$ chains of fibrinogen at Arg16–Gly17 with the subsequent release of two molecules of fibrinopeptide A (FPA) [77]. The FPA deficient fibrinogen molecule is known as the fibrin monomer. Fibrin monomers proceed to polymerize to fibrin protofibrils, in which the monomers are laid end-to-end with an overlap equal to one-half of the monomer's length of 450 Å [78], i.e. each protofibril contains one monomer 225 Å long. The second stage of fibrin assembly involves the lateral association of protofibrils to form fibers. This stage is usually accompanied by thrombin cleavage of $B\beta$ chains of fibrinogen at Arg14–Gly15 and release of fibrinopeptide B (FPB). The protofibrils are not able to form a blood clot until at least three additional types of bonds are formed. First, the protofibrils must be linked to one another through a lateral attachment site in the α chain [79, 80, 81]. At this point, the clot consists of a loosely linked, unstable network. Second, the entire clot is stabilized by the formation of covalent, end-to-end, antiparallel bonds between the C-terminal of adjacent γ chains in the protofibrils [82]. Finally, the fibrin structure is covalently stabilized as the result of the action of Factor XIIIa (fibrin stabilizing factor), an

Thr 1H

B Chain :

[illegible]

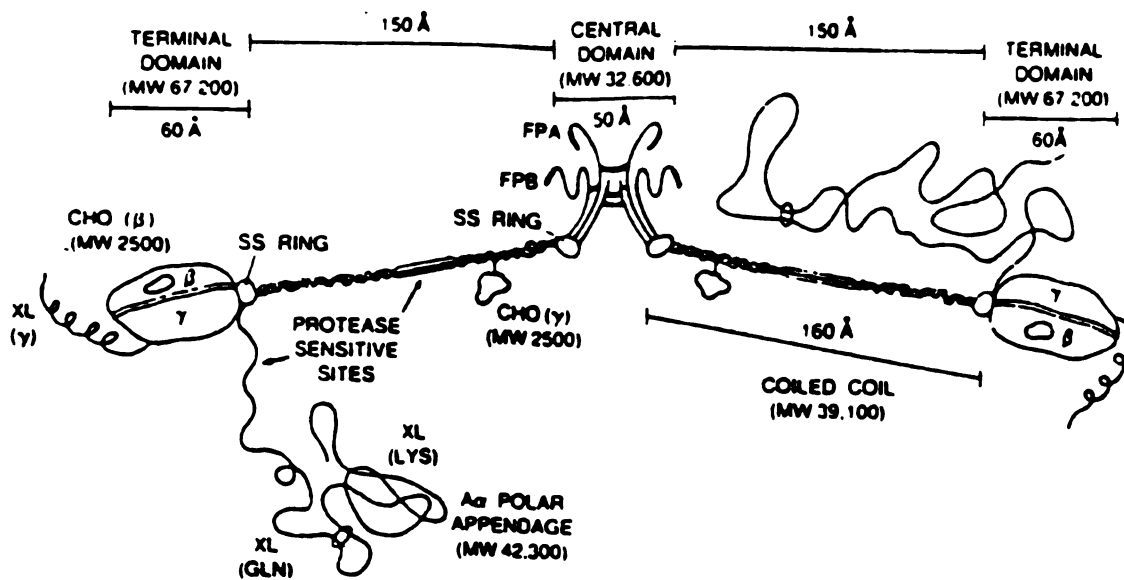
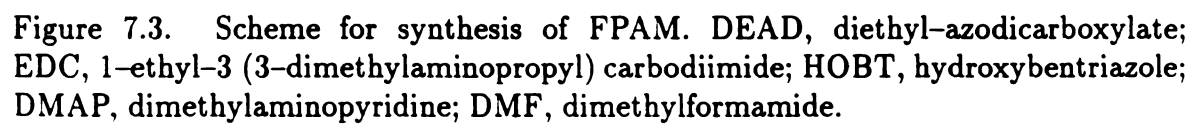
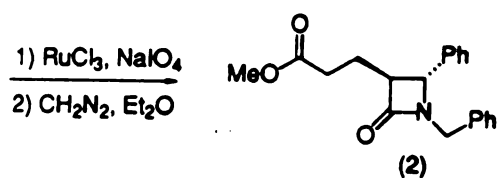


Figure 7.2. The structure of the fibrinogen molecule. The symmetric molecule is composed of a dimeric central domain containing the N-terminal of all six chains (α , β , γ), two connecting coiled coils, two terminal domains, and two A α polar appendages. Four carbohydrate clusters (CHO) occur, and are located on each γ chain near the central domain and on the β chains of each terminal domain. Primary cross-linking sites (XL) can be found near the C-terminal of the γ chain and in the A α polar appendages (taken from [76]).

enzyme which introduces bonds between ϵ amino group of lysine and the γ carboxy group of glutamine.

Various alternate methods [83, 84, 85, 86, 87] have been employed in the investigation of the mechanism of the interaction of thrombin with fibrinogen, especially about the cleavage of the Arg-Gly peptide bond in the A α chain of human fibrinogen. These studies have shown that the first six residues of FPA do not interact with thrombin, whereas Asp7 and Phe8, which are located 10 and 9 residues away, respectively, from the thrombin cleavage site, influence the effectiveness of the binding of synthesis peptide substrates to thrombin. Since Asp7 and Phe8 are relatively far in sequence from the scissile bond at Arg16, several investigations have proposed that the N-terminal of FPA binds to thrombin in a bent configuration with Phe8 close to Arg16 [88, 89]. Through NMR results [90], it has been shown that residues Asp7, Phe8, Leu9, Val15, and Arg16 are involved in the interaction of the peptide with thrombin. Furthermore, transfer NOE measurements [90] indicated that a β -bend might exist within the segment from Gly12 to Val15 in the FPA-thrombin complex.

In order to examine the role that this reverse turn plays in the thrombin active site, a chloromethyl ketone inhibitor mimetic of FPA (FPAM) has been designed and synthesized as shown in Figure 7.3 [91]. A model for the thrombin-bound structure of FPAM has been proposed (Figure 7.4) [92] based on the crystallographic thrombin structure [93, 43], NMR data [94], computer assisted molecular modeling and peptidomimetic substrates and inhibitors [91]. Subsequently, the X-ray crystal structures of human FPA bound to bovine thrombin [95] and a chloromethyl ketone FPA derived thrombin complex [96] were determined; some variance between these and the predicted FPAM structures were observed. In order to obtain a better understanding of the interplay between the primary sequence and the conformation required for thrombin substrates and inhibitors, a crystallographic investigation of the previously modeled mimetic FPAM complexed with human α -thrombin was undertaken [97].



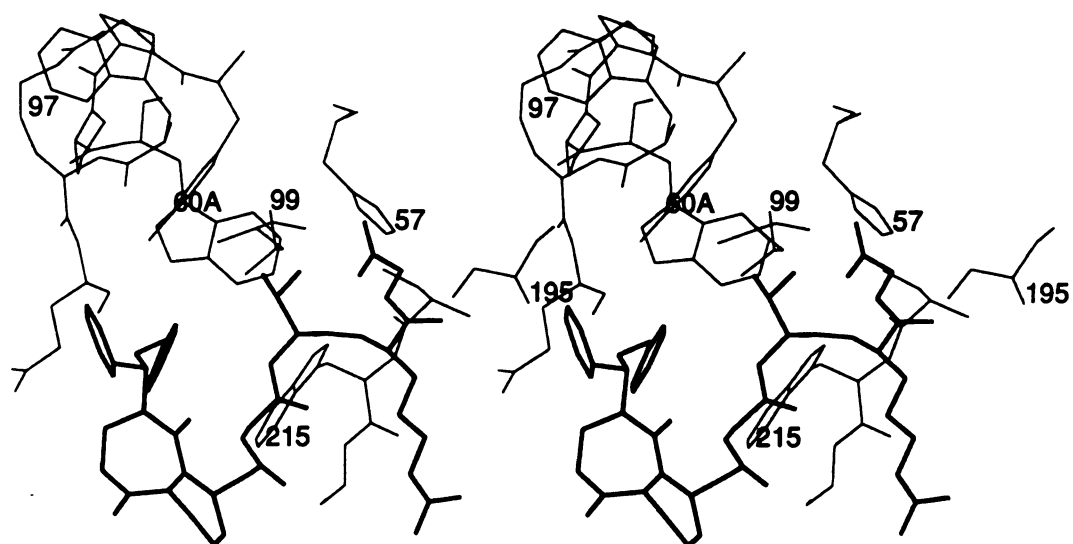


Figure 7.4. Stereoview of the FPAM structure docked in the thrombin active site. FPAM is in bold.

7.2 Experimental Procedures

A. Crystallization

The FPAM–thrombin–hirugen (N-acetylhirudin 53'–64' with sulfato–Tyr63') complex was crystallized in a similar way to that of the hirugen–thrombin complex [98]. An approximately 10 fold molar excess of hirugen was added to a frozen 1 ml sample of thrombin solution (2.6mg/ml) at 4°C. The solution was then diluted to 2 ml with 0.1 M phosphate buffer at pH 7.3 and 225 μ l of FPAM chloromethyl ketone solution at a concentration of 2.5 mg/ml in methanol (10 molar excess) was added. The solution of the ternary complex was concentrated to about 5 mg/ml using a Centricon 10 miniconcentrator (MW cutoff 10K) in a refrigerated centrifuge. Crystallization was carried out in 10 μ l hanging drops against 1 ml of well solution containing 0.1 M sodium phosphate buffer (pH 7.3), 28% PEG 8000. Autolysis was prevented by the hirugen in crystals grown in this manner (Figure 7.5). The crystals were found to be isomorphous to those of the hirugen thrombin complex : monoclinic, space group C2, four molecules per unit cell, $a=71.13$ Å, $b=72.43$ Å, $c=73.00$ Å, $\beta = 101.09^\circ$ with an estimated protein content of 50% and $V_m=2.5$ Å³/ dalton.

B. Data Collection

X-ray diffraction intensities were measured at 2.5 Å resolution from one crystal having dimensions $0.40 \times 0.27 \times 0.10$ mm employing a Siemens multiwire area detector with graphite monochromated CuK $_{\alpha}$ radiation from a Rigaku RU200 rotating anode tube operating at 50KV and 150mA. The crystal–detector distance was 11.65 cm, the detector swing angle was set at 12°, the scan angle was 0.2° per frame of measurement and each frame was measured for 90 seconds. The raw data which were reduced and scaled with the XENGEN programs [26], yielded 37,533 reflections of which 11,675

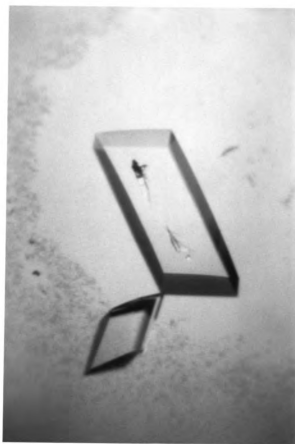


Figure 7.5. Photograph of FPAM crystal. Crystal size is approximately $0.5 \times 0.45 \times 0.1 \text{ mm}$.

were unique. The distribution of intensities observed for various resolution ranges is given in Table 7.2. After removing reflections with $I/\sigma(I) < 2$, a set containing 10,918 independent reflections remained (93% observed. $R_{\text{merge}} = 0.039$).

Table 7.2. Distribution of reflection intensities and R-factors in various resolution shells.

Res. (Å)	# refs	0	$<2\sigma$	$<5\sigma$	$<10\sigma$	$<20\sigma$	$<40\sigma$	$<60\sigma$	$>60\sigma$
4.54	2155	0	5	32	38	84	259	329	1408
3.60	2122	9	25	55	69	172	425	381	986
3.15	2109	24	85	133	203	374	624	378	288
2.86	2107	48	193	297	390	517	510	130	22
2.66	2112	38	245	496	487	517	293	35	1
2.50	1070	28	160	374	299	179	28	2	0
Totals:	11675	147	713	1387	1486	1843	2139	1255	2705

C. Refinement

The FPAM–thrombin crystal structure was solved using isomorphous thrombin coordinates of the hirugen–thrombin complex [98]. Since no electron density was found for the autolysis insertion loop from Thr147–Lys149E in other isomorphous hirugen complexes, the initial model only included the A chain from Thr1H to Arg15, and Ile16–Glu146 and Gly150–Glu247 of the B chain of thrombin. The structure was refined employing the restrained least-squares method implemented in the program PROLSQ [30] with intermittent model building performed on an Evans and Suther-

land PS390 interactive stereographics system with the program FRODO [32]. The refinement proceeded in two stages, data from 7.0–2.8 Å resolution were included initially, then data from 7.0–2.5 Å resolution. Each major round of refinement was followed by model building; $(2|F_o| - |F_c|)$ and $(|F_o| - |F_c|)$ electron density maps were used in conjunction with the Ramachandran plot. The R value started at 28% with hirugen not considered in the calculation and with an overall thermal parameter of 25.0 Å². The first $(2|F_o| - |F_c|)$ electron density map at 2.8 Å resolution showed good density for most of the thrombin–hirugen residues and Val5–Arg16 of the FPAM (Figure 7.6). The hirugen and FPAM were fitted into the density and gradually included throughout further calculations. The special aromatic groups in FPAM, not being regular amino acids, did not have normal peptide bonds connecting to the bicyclic ring. In order for the PROLSQ program to recognize and refine the aromatic rings, the dictionary had to be modified and some additional restraints were applied in the control file (Figure 7.6). The R value decreased to 19.2% after the first refinement stage (2.8 Å resolution) and water molecules were located at 2.5 Å resolution. Peaks considered to be possible water molecules were identified by comparing (7.0–2.5) Å and (8.0–2.5) Å resolution difference maps. In addition, the hirugen and the FPAM positions were updated according to the electron density maps. The final FPAM–thrombin–hirugen structure has a crystallographic R value of 13.8% for 10,139 reflections between 7.0 and 2.5 Å resolution with 234 water molecules and an average thermal parameter of 29 Å². The average occupancy of the water molecules is about 0.67 and their average thermal parameter is 29 Å². The final reflection weights and R values in each range are given in Table 7.3, and a summary of refinement parameters is listed in Table 7.4. The distribution of main-chain torsion angles (ϕ, ψ) is shown in Figure 7.7; nearly all the non-glycine amino acids fall within or close to conformationally allowed regions (except Glu1C and Ser1E of the A-chain). A *cis* peptide bond occurs at Pro37 with an ω angle of 0.77°. The ω angles of the remaining peptide

bonds are in a narrow range close to planarity.

7.3 RESULTS

A. Structure of Thrombin

Nearly the entire structure of thrombin is well defined; however, as in other hirugen-thrombin complexes, little or no electron density was found for N-terminal residues Thr1H-Glu1C of the A chain and C-terminal residues Phe245-Glu247 of the B chain. In the present structure, no electron density was observed for the sidechain atoms of Asn62, Lys110, Gln151 and the sometimes important Glu192 residue. The structure of thrombin in the ternary complex was compared with that of the hirugen-thrombin complex by the optimal superposition of CA, C, and N atoms; agreement between the two is excellent (Table 7.5). The residues in the active sites of the two complexes have been compared in detail and a stereoview of the superposition of the active site regions is shown in Figure 7.8. The only significant change in the active site induced by the binding of FPAM to thrombin is associated with Trp60D. Another large deviation in the region that occurs at the sidechain of Ile174 results in the sidechain pointing toward the face of the phenyl ring of Phe1 of FPAM overall, the residues in the catalytic site have practically the same conformations as those in the hirugen-thrombin complex [98], where the active site is unoccupied. Thus, the conformation of the active site of thrombin is conserved and binding of FPAM or D-Phe-Pro-Arg-chloromethylketone (PPACK) does not induce much change in the region.

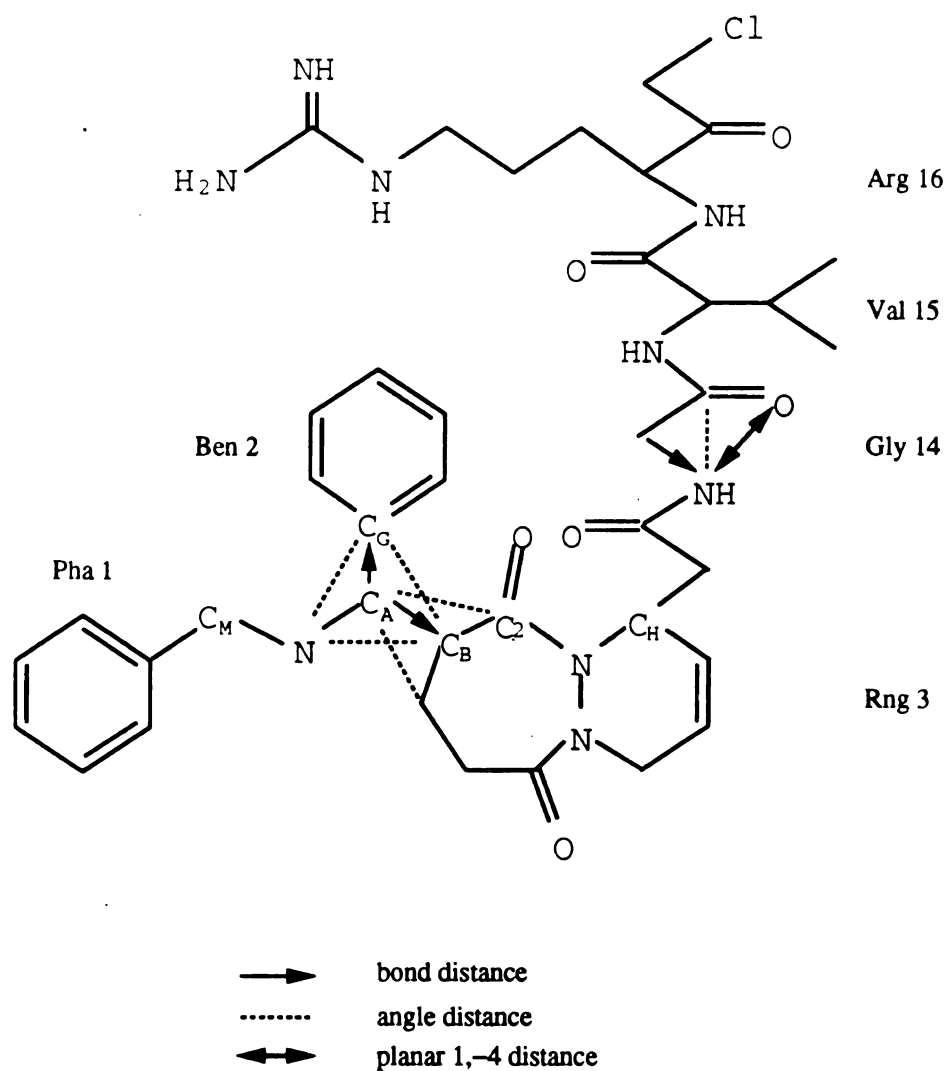


Figure 7.6. Numbering of and special restraints used for FPAM in restrained least squares refinement. Since part of FPAM is a non-amino acid group, special bond, angle and planar 1,-4 distance restraints were applied during refinement to maintain geometry.

Table 7.3. Weights of reflections and R values of the final refinement

$d_{\min}(\text{\AA})$	no. reflections	$\sigma(F)$	$\langle F_o - F_c \rangle$	R value	
				shell	sphere
4.60	1388	39	80	0.146	0.146
3.85	1406	34	64	0.111	0.128
3.42	1423	31	56	0.118	0.125
3.13	1387	28	51	0.141	0.128
2.90	1482	26	44	0.157	0.132
2.72	1521	23	41	0.165	0.135
2.40	1532	21	37	0.163	0.138
$\sigma(F) = 26 - 180 [(\sin \theta/\lambda) - (1/6)]; \langle F_o - F_c \rangle = 53$					

Table 7.4. Final least squares parameters and deviations of FPAM–thrombin.

	target σ	rms δ
Distances ($\text{\AA}$)		
Bond distance	0.020	0.016
Angle distance	0.030	0.044
Planar 1,-4 distance	0.050	0.049
Non-bonded distances ($\text{\AA}$)		
Single torsion	0.50	0.23
Multiple torsion	0.50	0.32
Possible H-bond	0.50	0.31
Torsion angles (deg)		
Planar	3	2
Staggered	15	24
Orthonormal	20	31
Plane groups ($\text{\AA}$)		
Plane groups ($\text{\AA}$)	0.02	0.01
Chiral centers ($\text{\AA}^3$)	0.15	0.18
Thermal restraints ($\text{\AA}^2$)		
Main chain bond	1.5	1.1
Main chain angle	2.0	1.9
Side chain bond	2.5	1.7
Side chain angle	2.5	2.6

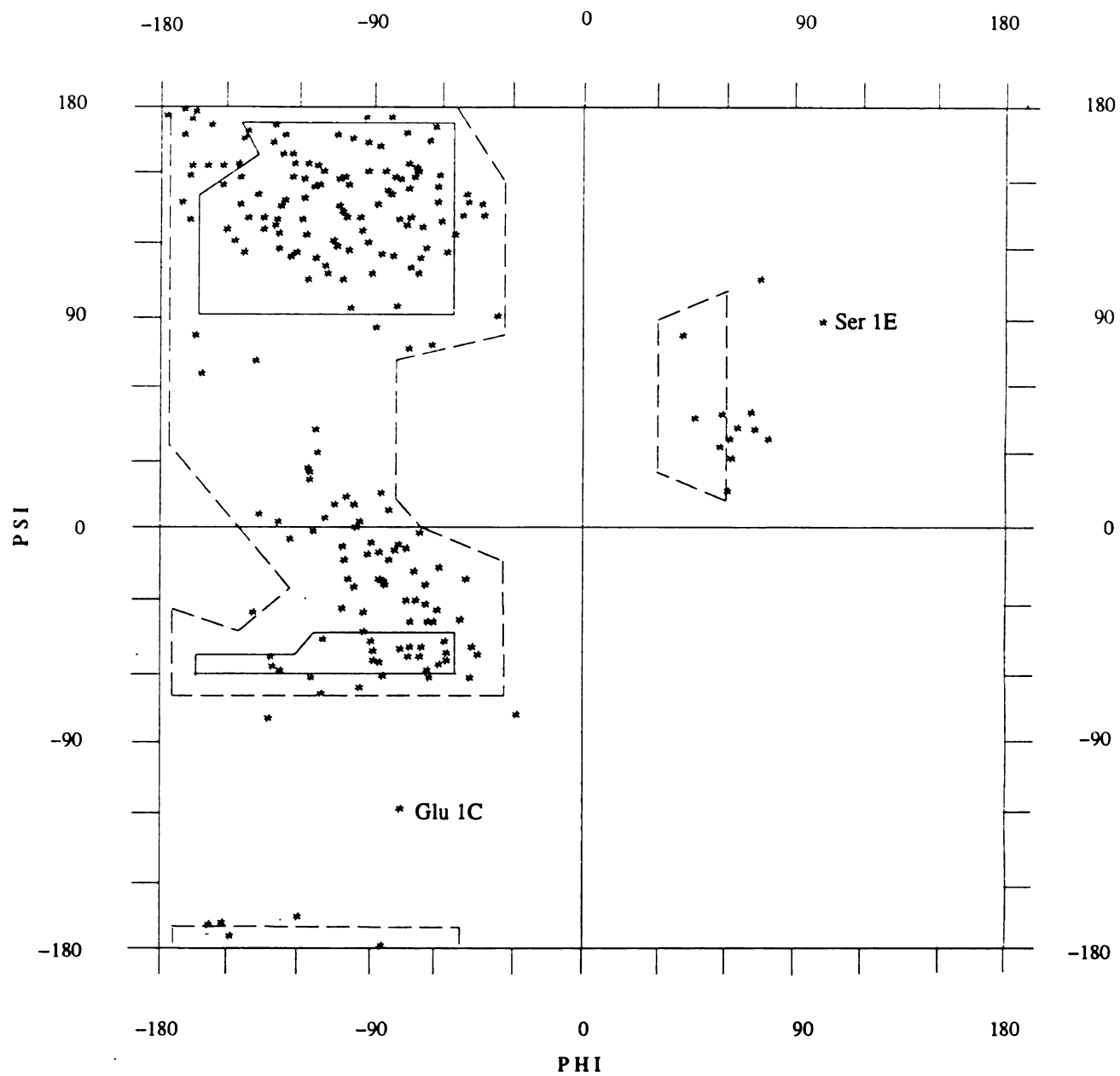


Figure 7.7. Ramachandran plot of ϕ , ψ angles of FPAM-thrombin structure. Glycines are not displayed.

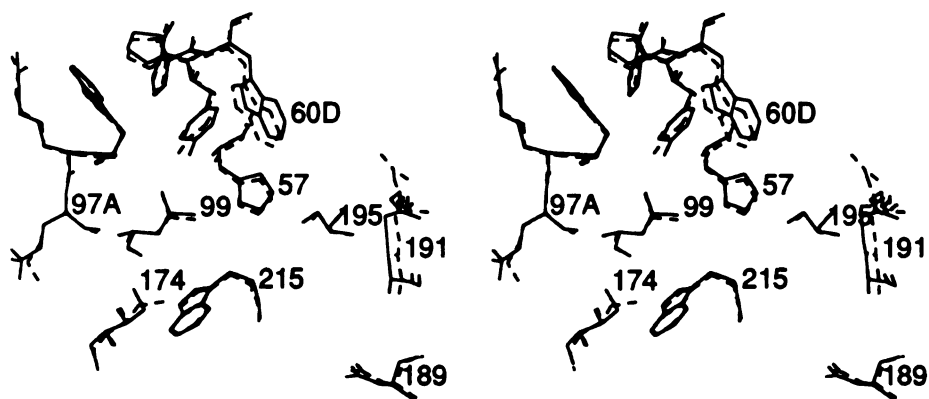


Figure 7.8. Stereoview of the comparison of the active sites of thrombin in the FPAM and hirugen complexes. Hirugen-thrombin, broken lines; active site is unoccupied.

Table 7.5. RMS Deviations Between the Hirugen–Thrombin and FPAM–Hirugen–Thrombin Complexes

	Δ (Å)	atom no.
All protein atoms	0.54	1972
Main chain	0.28	738
Carbonyl oxygens	0.35	246
Side chains	0.70	988
Sulfurs (Cys, Met)	0.36	14
Carbon alphas	0.29	246

B. Structure of FPAM

The FPAM structure is nearly completely defined by electron density in the active site of the thrombin (Figure 7.9). However, the two phenyl rings lack continuity to the aromatic groups. The two phenyl rings are about 7.8 Å from each other in the apolar binding site region of thrombin (Figure 7.10), and each of them interacts with thrombin through a number of hydrophobic contacts. The conformation of Arg16 to Rng3 of FPAM (Figure 7.6) is very similar to that of PPACK in the PPACK–thrombin complex (Figure 7.11) [99], with the former making an approximate helical turn between Gly14–Arg16, which is followed by the mimetic β -bend of the bicyclic ring. There are some minor differences in the main chain positions between the two structures in the helical-like turn. The main chain nitrogen atoms of Gly14 and Arg16 make a two strand antiparallel β -sheet with the thrombin Ser214–Gly216 segment (Figure 7.10) so this region, like that of PPACK, also possesses very favorable interactions. Although the electron density of FPAM is generally quite good (Figure 7.9), the mimetic has an average thermal parameter of 49 Å², which is almost twice that of thrombin (28 Å²). A similar value was observed for hirudin in the hirudin–thrombin complex [43], where the difference was attributed to imprecision in positioning of the inhibitor.

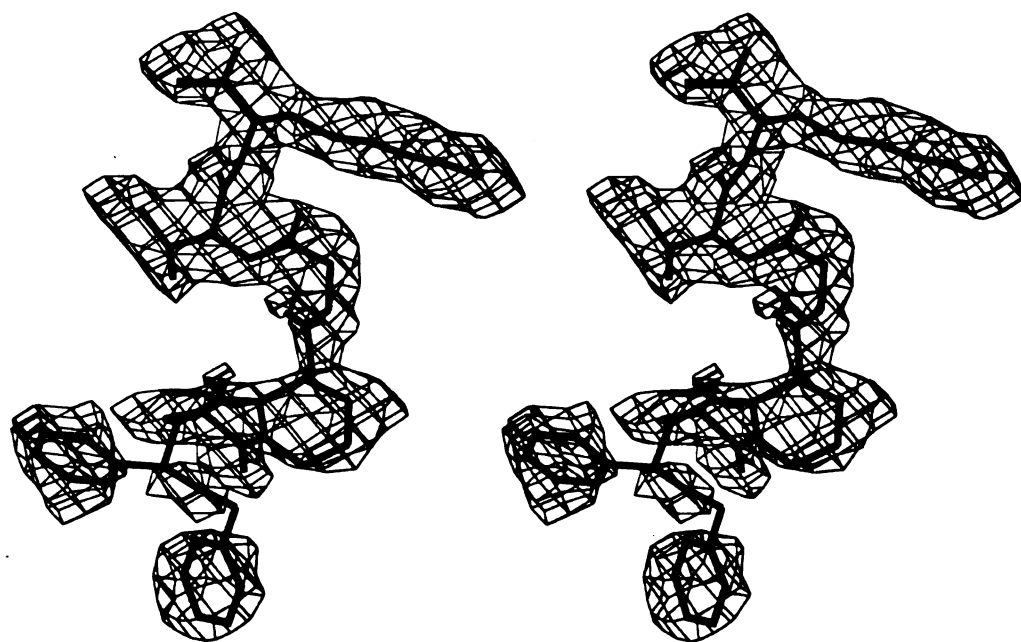


Figure 7.9. Stereoview of the electron density corresponding to FPAM in the thrombin complex. Basket contour at 1σ .

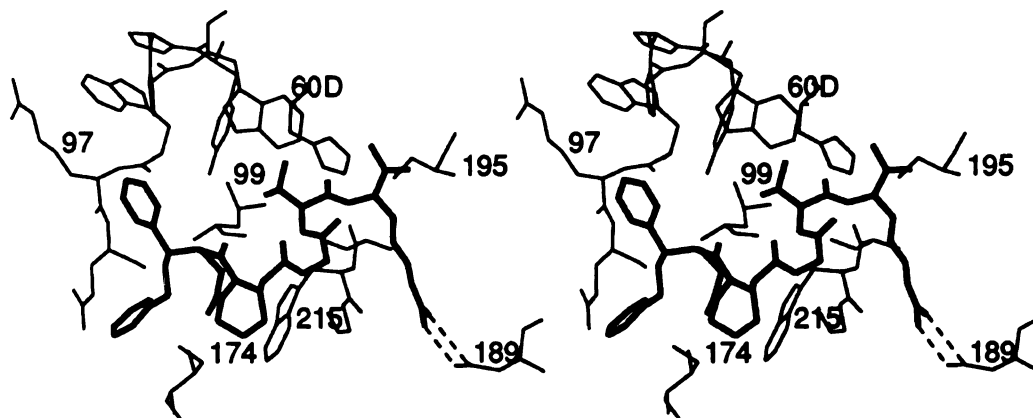


Figure 7.10. Stereoview of FPAM bound in active site of thrombin. FPAM in bold; hydrogen bonds, broken.

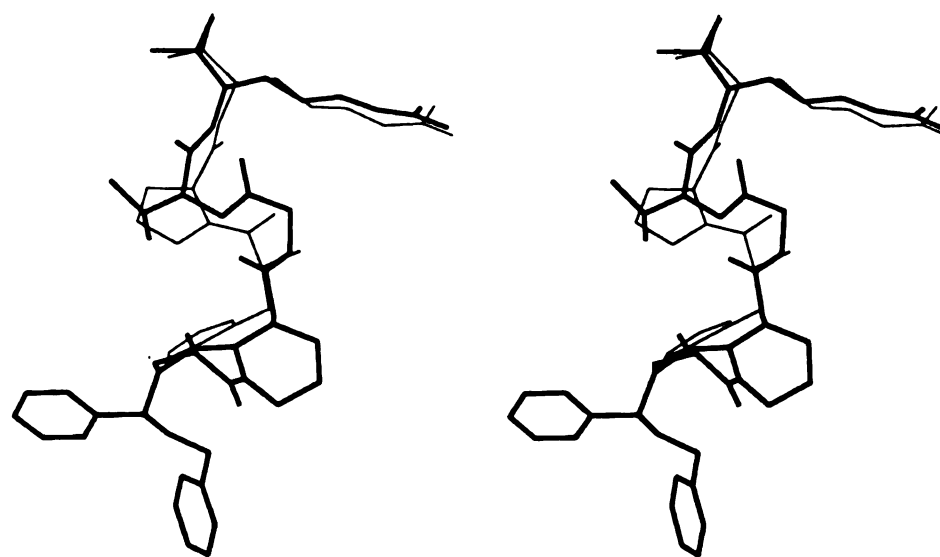


Figure 7.11. Stereoview of the comparison of FPAM (bold) and PPACK in their thrombin complexes.

C. FPAM–Thrombin Interaction

The active site of thrombin, which displays a preference for arginyl and lysyl sidechains, has the form of an elongated channel and is defined by peptide segments Tyr60A–Trp60D, Arg97–Leu99, Thr172–Arg175, Cys191–Gly196 and Ser214–Glu217 (Figure 7.10). The density observed for the catalytic triad indicates an intermediate hemiketal is formed between the carbonyl group of Arg16 and Ser195 OG (2.2 Å) (Figure 7.10). Both FPAM and PPACK being chloromethylketone derivatives have the same contacts at this region. The S1 specificity pocket of the FPAM complex is occupied with an arginyl group with geometry similar to that of arginine in PPACK– [99], hirulog 1– [98], and β -homoarginine in hirulog 3–thrombin [100]: the guanidinium group of the Arg16 forms a doubly hydrogen bonded salt-bridge with the carboxyl oxygens of Asp189 (2.5 Å and 3.1 Å) (Figure 7.10). Moreover, the guanidinium group makes a close contact with the carbonyl oxygen of Gly219 (3.1 Å) and the main chain nitrogen atom of Arg16 may be involved in a hydrogen bond with the carbonyl oxygen of Ser214 (3.1 Å). The sidechain of Val5 in the S2 subsite is buried within a hydrophobic cage that is the apolar binding site of the thrombin (Figure 7.10). The valyl group makes hydrophobic contacts with sidechains of Tyr60A, Trp60D and Leu99 (3.7 Å, 3.8 Å, 4.3 Å respectively) and occupies a spatial region similar to Pro of PPACK in PPACK–thrombin complex [99]. The P3 interaction observed in this structure is due primarily to an anti-parallel β -sheet hydrogen bond between the amide nitrogen of Gly14 and the carbonyl group of Gly216 (2.4 Å) that appears to be important in positioning the bicyclic ring and is different from that in PPACK–thrombin. The bicyclic ring corresponding to a β -bend, which was presumed to be at the 11–12 position of FPA [94], has a (S,S) conformation according to the chirality at carbon atoms C_H and C_B (Figure 7.6) and is located in the region bordered by residues of Tyr60A, Trp60D, Leu99, Trp215 and Glu217 (Figure 7.10). The ring, al-

though not aromatic, forms an end-to-face contact with the indole ring of Trp215 and produces an aromatic-like interaction that is common in proteins [45]. In addition to the stacking interaction, the bicyclic ring is also stabilized by other hydrophobic contacts formed with sidechains of Tyr60A, Trp60D and Leu99. The N-terminal of the FPAM concludes with two phenyl groups that are also located near the non-polar region of the S2, S3 subsites. The Phal ring makes a good van der Waals contact with the sidechain of Ile174 (2.6 Å) and shows density for its phenyl ring. However, Ben2 on the other side has no significant interaction with thrombin except for loose hydrophobic contacts with residues of Tyr60A and Pro60C, and is most likely the reason why the density is not as well-defined as that of Phal (Figures 7.9 and 7.10).

D. Comparison of FPAM Related Structures

There are five FPA derived structures that are relevant in any comprehensive comparison of thrombin in its bound state: (1) the FPA bovine thrombin complex [95], (2) the chloromethyl ketone of FPA alkylating His57 of the active site [96], (3) the solution NMR structure of FPA bound to thrombin [94], (4) the present FPAM-thrombin structure and (5) the modeled FPAM-thrombin complex structure [92]. The structures of FPA in (1) and (2) are practically identical with $\text{rms}\Delta=0.6$ Å and only 0.3 Å if side chains of Leu9 and Glu11 are omitted (Figure 7.12). The ϕ , ψ , conformational angles of the P1-P2-P3 residues of (1), (2), (3), and (5) are all in close agreement with the conformation of PPACK-thrombin, except for ψ_3 of PPACK which is negative due to the D-Phe enantiomer [99]. In addition, the conformational angles of the P1-P2-P3 residues of (1), (2), (3), and (5) are also similar to that of Pro13-Cys14-Lys15 of bovine pancreatic trypsin inhibitor (BPTI) bound in the BPTI-trypsin complex [101]. The P1-P2-P3 residues of FPA of the NMR structure bound to thrombin [94] are not in agreement with the foregoing, especially with respect to the ψ_2 and ϕ_3 , ψ_3 angles. This lack of agreement was first noticed

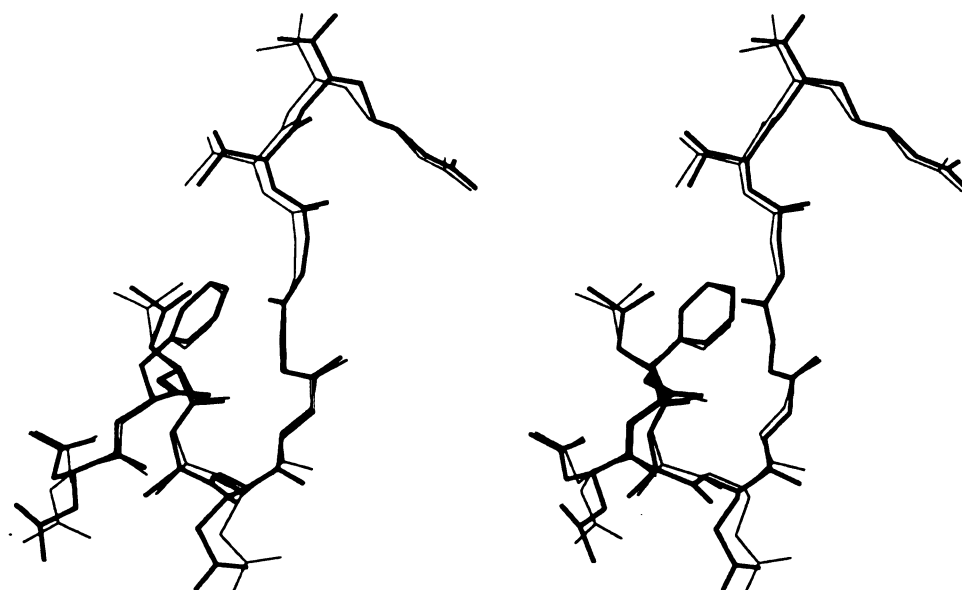


Figure 7.12. Stereoview of the comparison of FPA (bold) and FPA-chloromethyl ketone in their thrombin complexes.

when attempts to dock and model the NMR structure in the active site of thrombin by placing Arg16 in the S1 subsite, failed because of massive collisions of FPA with the enzyme [92]. When the BPTI trypsin bound structure was used as a template to reorient the P1-P2-P3 residues of the NMR structure, an excellent fit of these in the active site of thrombin was achieved. In the same work, similar rationalizations were employed to model FPAM in the active site.

The structure of FPAM bound to thrombin is compared with that of FPA in the thrombin complex [95] in Figure 7.13, from which it will be seen that the first generation FPA mimetic does not correspond to FPA in two important aspects. The first aspect is that a peptide insertion that follows the bicyclic system could place the β -turn of the mimetic more optimally with respect to the turn of FPA. The initial positioning of the turn in FPAM was based on the NMR position (between Gly11 and Val15). The second aspect is in the conformation of the Phal moiety: rather than reversing to interact with Val15 as in FPA, the FPAM molecule assumes a more or less extended conformation in the thrombin complex (Figures 7.10 and 7.13). Both of these shortcomings can be easily rectified with some additional design features and synthesis (insertion of a peptide, reconfirming the bicyclic system). Another notable difference from the FPA complex is that Gly14 N of FPAM makes a hydrogen bond with Gly216 O thereby altering the conformation somewhat between the P2-P3 positions. The most important aspect that emerges, however, is that the FPAM-thrombin complex displays yet another binding mode in the active site of thrombin [102] bringing the total to four. The other three are: (1) FPA/substrate-like, (2) N-terminal hirudin-like [93, 43] and (3) argatroban-like [103].

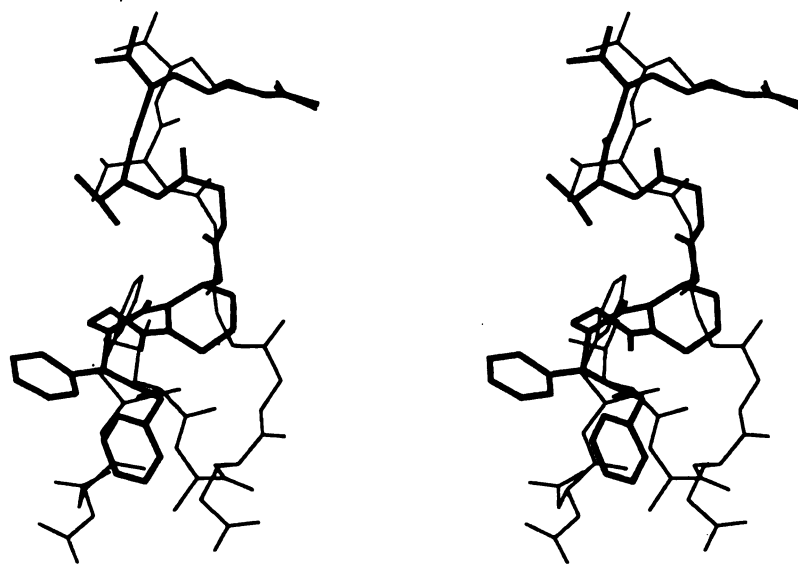


Figure 7.13. Stereoview of the comparison of FPAM (bold) and FPA in their thrombin complexes.

7.4 Discussion

A hallmark of the enzyme thrombin is its remarkable specificity. This is due to the inherently deep and constricted binding site of thrombin and a requirement of its substrates to adopt a specific conformation to productively bind in the active site. Limited proteolysis by trypsin-like serine proteases plays an important role in coagulation/hemostasis [104] and complements activation [105]. The potential to specifically intervene in these processes to ameliorate a number of disease states is significant and well recognized. However, controlled and selective interference is critical to the success of the strategy that is made considerably more difficult by the high degree of sequence homology within this family of proteases, which generally contain a trypsin-like core with insertions which modify specificity and are responsible for interaction with additional macromolecular components [106].

Of the 181 Arg/Lys-Xaa sequences in fibrinogen [107], only two bonds are cleaved by thrombin. Experimental rationalization of this was provided through NMR investigations of the complex between FPA and bovine thrombin [94]. A striking feature that emerged in this study is the cluster of nonpolar residues (Phe8, Leu9, and Val15) that were apparently brought into close proximity by a reverse turn. Reverse turns have been implicated in enhancing the specificity of proteolytic processing of prohormones, zymogens and viral proteins [108, 109, 110]. The role of secondary structural elements in proteolysis has been investigated through the incorporation of peptidomimetic prosthetic units [91, 111, 112] and recently a model was proposed for the bound structure of FPA [92]. The most critical feature of the model involved the reorientation of the *P1 to P3* residues of the NMR derived FPA structure to coincide in alignment with the active site conformation of BPTI, which is believed to represent a canonical loop proteolysis substrate mimic [101, 113, 114]. This proved to be very effective in that the reoriented FPA model then fit well into the thrombin active site and satisfied all

of the previously reported NOE data [94]. Hybrid mimetic substrates based on the foregoing were found to have similar kinetic parameters to those of FPA₁₋₅₂, and thus believed to effectively mimic the bound conformation of natural substrates.

We anticipated that because the FPAM inhibitor was designed around a specific natural substrate for thrombin, it would exhibit a high degree of selectivity. This is indeed the case, in that FPAM exhibits a degree of specificity similar to that of PPACK, a well known selective inhibitor of thrombin. Utilization of the canonical loop motif [113] of a natural proteinaceous inhibitor as a lead for designing reverse turn peptidomimetic inhibitors may provide a general strategy for introducing specificity into an inhibitor.

Although the model for the bound structure of FPAM is consistent with the observed crystallographic structure [95], particularly in the orientation of the P1 to P3 sites, not surprisingly there are some significant differences. Most striking is the fact that the hydrophobic pocket, formed by the 60 insertion loop of thrombin and residues Leu99, Ile174 and Trp215, is not fully occupied by the N-benzyl group back-tracking to Val15 as anticipated, but rather by the bicyclic β -turn template in a similar manner to the dansyl group of DAPA-thrombin [102] and related molecules [103]. This is due to the relatively extended structure of FPAM and is yet another example of the intriguing dichotomy that exists between the specificity and promiscuity of thrombin [102]. A peptide insertion between Gly14 and Rng3 (Figure 7.6) and an alternate conformation or stereochemistry in the C-7 ring of the bicyclic β -turn prosthetic unit could conceivably match the bound FPA conformation with considerable fidelity (Figure 7.13).

Information garnered from these investigations is being utilized in the design and synthesis of novel nonpeptidic thrombin inhibitors by Kahn and his collaborators. This stepwise process, starting with natural protease substrates or inhibitors and culminating in truly nonpeptide inhibitors will generate new structures that should

maintain the specificity that nature has so elegantly and carefully crafted.

APPENDICES

APPENDIX A

Amino Acid Shorthand Used in the Thesis.

Amino Acid	Three-letter abbreviation	one-letter symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] B. Wiman and D. Collen *Nature*, vol. 272, pp. 549–550, 1978.
- [2] B. Wiman and P. Wallen *Thromb. Res.*, vol. 1, pp. 213–222, 1977.
- [3] G. Markus, J. L. DePasquale, and F. C. Wissler *J. Biol. Chem.*, vol. 253, pp. 727–732, 1978.
- [4] G. Markus, R. L. Priore, and F. C. Wissler *J. Biol. Chem.*, vol. 254, pp. 1211–1216, 1979.
- [5] G. Markus, J. L. Evers, and G. H. Hobika *J. Biol. Chem.*, vol. 253, pp. 733–739, 1978.
- [6] S. Magnusson, T. E. Petersen, L. Sottrup-Jensen, and H. Claeys in *Proteases and Biological Control* (E. Reich, D. B. Rifkin, and E. Shaw, eds.), pp. 123–149, Cold Spring Harbor Laboratory, Cold Spring, NY, 1975.
- [7] L. Sottrup-Jensen, H. Claeys, M. Zajdal, T. E. Peetersen, and S. Magnusson *Prog. Chem. Fibrinolysis Thrombolysis*, vol. 3, pp. 191–209, 1978.
- [8] C. T. Esmon and C. M. Jackson *J. Biol. Chem.*, vol. 249, pp. 7791–7797, 1974.
- [9] P. G. Lerch, E. E. Rickli, W. Lergier, and D. Gillessen *Eur. J. Biochem.*, vol. 107, pp. 7–13, 1980.
- [10] L. Patthy *Cell*, vol. 41, pp. 756–663, 1985.
- [11] G. J. Steffens, W. A. Gunzler, F. Otting, E. Frankus, and L. Flohe *Hoppe-Seyler's Z. Physiol. Chem.*, vol. 363, pp. 1043–1058, 1982.
- [12] B. A. McMullen and K. Fujikawa *J. Biol. Chem.*, vol. 260, pp. 5328–5340, 1985.
- [13] D. Pennica, W. E. Holmes, W. J. Kogr, R. N. Harkins, G. A. Vehar, C. A. Ward, W. F. Bennett, E. Yelverton, P. H. Seeberg, H. L. Hrynobar, D. V. Goeddel, and D. Collen *Nature*, vol. 301, pp. 214–221, 1983.

- [14] T. Nakaamura, T. Nishizawa, M. Hagiya, T. Seki, M. Shimonishi, A. Sugimura, K. Tashiro, and S. Shimizu *Nature*, vol. 342, pp. 440–443, 1989.
- [15] J. W. Mclean, J. E. Tomlison, W.-J. K., D. L. Eaton, E. Y. Chen, G. M. Fless, A. M. Scanu, and R. M. Lawn *Nature*, vol. 330, pp. 132–137, 1987.
- [16] A. Varadi and L. Patthy *Biochem. Biophys. Res. Commun.*, vol. 103, pp. 97–102, 1981.
- [17] M. Trexler, Z. Váli, and L. Patthy *J. Biol. Chem.*, vol. 257, pp. 7401–7406, 1982.
- [18] A. J. van Zonneveld, H. Veerman, and H. Pannekoek *Proc. Nat. Acad. Sci. U.S.A.*, vol. 86, pp. 4670–4674, 1986.
- [19] E. S. Winn, S. P. Hu, S. M. Hochschwender, and R. A. Laursen *Eur. J. Biochem.*, vol. 104, pp. 579–586, 1980.
- [20] C. L. Young, W. C. Barker, C. M. Tomaselli, and M. O. Dayhoff in *Atlas of Protein Sequence and Structure* (M. O. Dayhoff, ed.), pp. 73–93, National Biomedical Research Foundation, Silver Spring, Maryland, 1978.
- [21] A. Tulinsky, C. H. Park, and E. Jankun *J. Mol. Biol.*, vol. 202, pp. 885–901, 1988.
- [22] T. P. Seshadri, A. Tulinsky, E. Skrzypczak-Jankun, and C. H. Park *J. Mol. Biol.*, vol. 220, pp. 481–494, 1991.
- [23] A. M. Mulichak, A. Tulinsky, and K. G. Ravichandran *Biochemistry*, vol. 30, pp. 10576–10588, 1991.
- [24] H. W. Wyckoff, M. Doscher, D. Tsernoglou, T. Inagami, L. N. Johnson, K. D. Hardman, N. M. Allewell, D. M. Kelly, and F. M. Richards *J. Mol. Biol.*, vol. 27, pp. 563–578, 1967.
- [25] A. C. T. North, D. C. Philips, and F. S. Mathews *Acta Cryst.*, vol. 24A, pp. 351–359, 1968.
- [26] A. J. Howard, G. L. Gilliland, B. C. Finzel, T. L. Poulos, D. H. Olhendorf, and F. R. Salemme *J. Applied. Cryst.*, vol. 20, pp. 383–387, 1987.
- [27] M. G. Rossmann and D. M. Blow *Acta Cryst.*, vol. 15, pp. 24–31, 1962.

- [28] W. Steigemann, *Dissertation*. PhD thesis, Technische Unicersitat, Munchen, 1974.
- [29] M. Fujinaga and R. J. Read *J.Appl.Cryst.*, vol. 20, pp. 517–521, 1987.
- [30] W. A. Hendrickson and J. H. Konnert in *Biomolecular Structure, Function, Conformation and Evolution* (R. Srinivasan, ed.), pp. 43 – 57, Oxford, Pergamon, 1980.
- [31] B. C. Finzel *J.Appl.Cryst.*, vol. 20, pp. 53–55, 1987.
- [32] T. A. Jones in *Computational Crystallography* (D. Sayre, ed.), pp. 303–317, Oxford, Clarendon Press, 1982.
- [33] B. N. Violand and F. J. Castellino *J. Biol. Chem.*, vol. 251, pp. 3906–3912, 1976.
- [34] L. Sottrup-Jensen, H. Claeys, M. Zajdel, T. E. Petersen, and S. Magnusson *Prog. Chem. Fibrinolysis Thrombolysis*, vol. 3, pp. 139–209, 1977.
- [35] S. M. Hochschwender and R. A. Laursen *J. Biol. Chem.*, vol. 256, pp. 11166–11171, 1981.
- [36] V. Ramesh, A. M. Petros, M. Llinas, A. Tulinsky, and C. H. Park *J. Mol. Biol.*, vol. 198, pp. 481–498, 1987.
- [37] M. Andrew, V. Ramesh, and M. Llinas *Biochemistry*, vol. 28, pp. 1368–1376, 1989.
- [38] A. Tulinsky, C. H. Park, B. Mao, and M. Llinas *Protein*, vol. 3, pp. 85–96, 1988.
- [39] A. M. Mulichak and A. Tulinsky *Blood Coagulation and Fibrinolysis*, vol. 1, pp. 673–679, 1990.
- [40] T. P. Wu, K. Padmanabhan, A. Tulinsky, and A. M. Mulichak *Biochemistry*, vol. 30, pp. 10589–10594, 1991.
- [41] A. M. Mulichak, *Dissertation*. PhD thesis, Michigan State Univeraity, East Lansing, Michigan, 1991.
- [42] A. M. Mulichak, C. H. Park, A. Tulinsky, and M. Llinas *J.Biol.Chem.*, vol. 264, pp. 1922–1923, 1989.

- [43] T. J. Rydel, A. Tulinsky, W. Bode, and R. Huber *J. Mol. Biol.*, vol. 221, pp. 583–601, 1991.
- [44] M. Trexler, L. Banyai, L. Patthy, N. D. Pluck, and R. J. P. Williams *Eur.J.Biochem.*, vol. 152, pp. 439–446, 1985.
- [45] S. K. Burley and G. A. Petsko *Science*, vol. 229, pp. 23–28, 1985.
- [46] M. Llinas, A. D. Marco, S. M. Hochschwender, and R. A. Laursen *Eur. J. Biochem.*, vol. 135, pp. 379–391, 1983.
- [47] A. Motta, R. A. Laursen, N. Rajan, and M. Llinas *J. Biol. Chem.*, vol. 261, pp. 13684–13692, 1986.
- [48] J. R. Powell, J. M. Beals, and F. J. Castellino *Arch. Biochem. Biophys.*, vol. 248, pp. 390–400, 1986.
- [49] M. Trexler and L. Patthy *Proc. Natl. Acad. Sci. U. S. A.*, vol. 80, pp. 2457–2461, 1983.
- [50] A. Motta, R. A. Laursen, M. Llinas, A. Tulinsky, and C. H. Park *Biochemistry*, vol. 26, pp. 3827–3836, 1987.
- [51] Z. Vali and L. Patthy *J.Biol.Chem.*, vol. 259, pp. 13690–13694, 1984.
- [52] J. C. W. Carter and C. W. Carter *J. Biol. Chem.*, vol. 254, pp. 12219–12223, 1979.
- [53] A. T. Brunger, “X-plor manual (version 2.1),” 1990.
- [54] E. E. Lattman *Acta Cryst.*, vol. B28, pp. 1065–1068, 1972.
- [55] A. M. de Vos, M. H. Ultsch, R. F. Keeley, K. Padmanabhan, and A. Tulinsky *Biochemistry*, vol. 31, pp. 270–279, 1992.
- [56] R. K. Arni, K. Padmanabhan, K. P. Padmanabhan, T. P. Wu, and A. Tulinsky *Biochemistry*, vol. 32, 1993.
- [57] X. Li, R. A. G. Smith, and C. M. Dobson *Biochemistry*, vol. 31, pp. 9562–9571, 1992.
- [58] T. Ny, F. Elgh, and B. Lund *Proc. Natl. Acad. Sci. U.S.A.*, vol. 81, pp. 5355–5359, 1984.

- [59] G. Pohl, M. Kallstrom, N. Bergsdorf, P. Wallen, and H. Jornvall *Biochemistry*, vol. 23, pp. 3701–3707, 1984.
- [60] P. Wallen, N. Bergsdorf, and M. Ranby *Biochim. Biophys. Acta*, vol. 719, pp. 318–328, 1982.
- [61] D. C. Rijken, M. Hoylaerts, and D. Collen *J. Biol. Chem.*, vol. 257, pp. 2920–2925, 1982.
- [62] A. J. V. Zonneveld, H. Veerman, and H. Pannekoek *Proc. Natl. Acad. Sci. U.S.A.*, vol. 83, pp. 4670–4674, 1986.
- [63] G. A. W. de Munk, M. P. M. Caspers, G. T. C. Chang, P. H. Pouwels, B. E. Engervalk, and J. H. Verheijen *Biochemistry*, vol. 28, pp. 7318–7325, 1989.
- [64] E. J. D. Weening-Verhoeff, P. H. A. Quax, R. T. van Leeuwen, E. F. Rehberg, K. R. Marotti, and J. H. Verheijen *Protein Engineering*, vol. 4, pp. 191–198, 1990.
- [65] I. L. Byeon and M. Llinas *J. Mol. Biol.*, vol. 222, pp. 1035–1051, 1991.
- [66] I. L. Byeon, R. F. Kelley, and M. Llinas *Eur. J. Biochem.*, vol. 197, pp. 155–165, 1991.
- [67] I. L. Byeon, R. F. Kelley, and M. Llinas *Biochemistry*, vol. 28, pp. 9350–9360, 1989.
- [68] J. W. I. Fenton *Ann. N. Y. Acad. Sci.*, vol. 370, pp. 468–495, 1981.
- [69] L. J. Berliner *Mol. Cell. Biochem.*, vol. 61, pp. 159–172, 1984.
- [70] K. G. Mann and R. L. Lundblad in *Hemostasis and Thrombosis: Basic Principles and Clinical Practice* (R. W. Coleman, J. Hirsh, V. J. Marder, and E. W. Salzman, eds.), pp. 148–161, J. B. Lippincott Co., Philadelphia, 1987.
- [71] M. E. Nesheim, L. S. Hibbard, P. B. Tracy, J. W. Bloom, K. H. Myrnes, and K. G. Mann in *The Regulation of Coagulation* (K. G. Mann and F. B. Taylor, eds.), pp. 145–149, Elsevier–North Holland, New York, 1980.
- [72] R. J. Butkowski, J. Elion, M. R. Downing, and K. G. Mann *J. Biol. Chem.*, vol. 252, pp. 4942–4957, 1977.
- [73] A. R. Thompson, D. L. Enfield, L. H. Ericsson, M. E. Legaz, and J. W. F. II *Arch. Biochem. Biophys.*, vol. 178, pp. 356–367, 1977.

- [74] W. Bode, I. Mayr, U. Baumann, R. Huber, S. R. Stone, and J. Hofsteenge *EMBO J.*, vol. 8, pp. 3467–3475, 1989.
- [75] A. Henschen, F. Lottspeich, M. Kehl, and C. Southan *Ann. N. Y. Acad. Sci.*, vol. 408, pp. 28–43, 1983.
- [76] R. F. Doolittle *Annu. Rev. Biochem.*, vol. 53, pp. 195–229, 1984.
- [77] B. Blomback, B. Hessel, D. Hogg, and L. Therkildsen *Nature*, vol. 275, pp. 501–505, 1978.
- [78] E. F. Casassa *J. Chem. Phys.*, vol. 23, pp. 596–597, 1955.
- [79] R. R. Hantgan and J. Hermans *J. Biol. Chem.*, vol. 254, pp. 11272–11281, 1979.
- [80] R. R. Hantgan, W. E. Fowler, H. P. Erickson, and J. Hermans *Thromb. Haemost.*, vol. 44, pp. 119–125, 1980.
- [81] P. H. Ehrlich, J. H. Sobel, Z. A. Moustafa, and R. E. Canfield *Biochemistry*, vol. 22, pp. 4184–4192, 1983.
- [82] R. Chen and R. F. Doolittle *Proc. Natl. Acad. Sci. USA*, vol. 66, pp. 472–479, 1970.
- [83] B. Blomback *Ann. N. Y. Acad. Sci.*, vol. 484, pp. 120–123, 1986.
- [84] H. A. Scheraga *Ann. N. Y. Acad. Sci.*, vol. 485, pp. 124–133, 1986.
- [85] Y. C. Meinwald, R. A. Martinelli, J. W. Nispen, and H. A. Scheraga *Biochemistry*, vol. 19, pp. 3820–3825, 1980.
- [86] H. C. J. Marsh, Y. C. Meinwald, S. Lee, and H. A. Scheraga *Biochemistry*, vol. 21, pp. 6167–6171, 1982.
- [87] H. C. J. Marsh, Y. C. Meinwald, T. W. Thannhauser, and H. A. Scheraga *Biochemistry*, vol. 22, pp. 4170–4174, 1983.
- [88] B. Blomback, M. Blomback, P. Olsson, L. Svendsen, and G. Aberg *Scand. J. Clin. Lab. Invest.*, vol. 24, pp. 59–64, 1969.
- [89] I. D. Rae and H. A. Scheraga *Int. J. Rept. Protein Res.*, vol. 13, pp. 304–314, 1979.
- [90] F. Ni, Y. Konishi, R. B. Frazier, and H. A. Scheraga *Biochemistry*, vol. 28, pp. 3082–3094, 1989.

- [91] M. Kahn, S. Wilke, B. Chen, and K. Fujita *J. Am. Chem. Soc.*, vol. 110, pp. 1638–1639, 1988.
- [92] H. Nakanishi, R. A. Chrusciel, R. Shen, S. Bertenshaw, M. E. Johnson, R. T. J., A. Tulinsky, and M. Kahn *Proc. Natl. Acad. Sci. USA*, vol. 89, pp. 1705–1709, 1992.
- [93] T. J. Rydel, K. G. Ravichandran, A. Tulinsky, W. Bode, R. Huber, C. Roitsch, and I. Fenton, J. W. *Science*, vol. 249, pp. 277–280, 1990.
- [94] F. Ni, Y. C. Meinwald, M. Vasquez, and H. A. Scheraga *Biochemistry*, vol. 28, pp. 3094–3105, 1989.
- [95] P. D. Martin, W. Robertson, D. Turk, R. Huber, W. Bode, and B. F. P. Edwards *J. Biol. Chem.*, vol. 267, pp. 7911–7930, 1992.
- [96] M. T. Stubbs, H. Oschkinat, I. Mayr, R. Huber, H. Angliker, S. R. Stone, and W. Bode *Eur. J. Biochem.*, vol. 206, pp. 187–195, 1992.
- [97] T. P. Wu, V. Yee, A. Tulinsky, R. A. Chrusciel, H. Nakanishi, R. Shen, C. Priebe, and M. Kahn *Protein Engineering*. in press.
- [98] E. Skrzypczak-Jankun, V. Carperos, K. G. Ravichandran, A. Tulinsky, M. Westbrook, and J. M. Maraganore *J. Mol. Biol.*, vol. 221, pp. 1379–1393, 1991.
- [99] W. Bode, D. Turk, and A. Karshikov *Prot. Sci.*, vol. 1, pp. 426–471, 1992.
- [100] X. Qiu, K. Padmanabhan, V. E. Carperos, A. Tulinsky, T. Kline, J. M. Maraganore, and J. W. Fenton II *Biochemistry*, vol. 31, pp. 11689–11697, 1992.
- [101] R. Huber and W. Bode *Acc. Chem. Res.*, vol. 11, pp. 114–122, 1978.
- [102] A. Tulinsky and X. Qiu *Blood Coagulation and Fibrinolysis*, vol. 4, pp. 305–312, 1993.
- [103] D. W. Banner and P. Hadvary *J. Biol. Chem.*, vol. 266, pp. 20085–20093, 1992.
- [104] E. W. Davie and K. Fujikawa *Annu. Rev. Biochem.*, vol. 44, pp. 799–829, 1975.
- [105] H. J. Muller-Eberhard *Annu. Rev. Biochem.*, vol. 44, pp. 697–724, 1975.
- [106] S. Magnusson, L. Sottrup-Jensen, T. E. Petersen, G. D. Wojciechowska, and H. Claey in *Miami Winter Symposia* (D. W. Ribbons and K. Brew, eds.), vol. 11, pp. 203–239, Academic Press, New York, N.Y., 1976.

- [107] H. Blomback in *Blood Clotting Enzymology* (W. H. Seegers, ed.), pp. 143–215, Academic New York, 1967.
- [108] M. Rholam, P. Nicolas, and P. Cohen *FEBS Lett.*, vol. 207, pp. 1–6, 1986.
- [109] M. Rhloam, P. Cohen, P. Brakch, N. Paolillo, A. Scarrurin, and C. DiBello *Biochem. Biophys. Res. Commun.*, vol. 168, pp. 1066–1073, 1990.
- [110] E. Bek and R. Berry *Chemistry*, vol. 29, pp. 178–183, 1990.
- [111] M. Kahn, H. Nakanishi, R. A. Chrusciel, D. Fitzpatrick, and M. E. Johnson *J. Med. Chem.*, vol. 34, pp. 3395–3399, 1991.
- [112] R. Shen, C. Proebe, C. Patel, L. Rubo, T. Su, M. Kahn, and R. Sugawara *Tetrahedron Lett.*, vol. 33, pp. 3417–3420, 1992.
- [113] M. Laskowski and I. Kato *Annu. Rev. Biochem.*, vol. 49, pp. 593–626, 1980.
- [114] R. Huber and R. W. Carrell *Biochemistry*, vol. 28, pp. 8951–8966, 1989.

MICHIGAN STATE UNIV. LIBRARIES



31293008853289