This is to certify that the

dissertation entitled

THE X-RAY CRYSTALLOGRAPHIC STRUCTURE DETERMINATION OF HUMAN PLASMINOGEN KRINGLE 4 AT 1.9 A RESOLUTION

presented by

Anne Marie Mulichak

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Chemistry

A. Tulinsky
Major professor

Date February 12, 1991

THE X-RAY CRYSTALLOGRAPHIC STRUCTURE DETERMINATION OF HUMAN PLASMINOGEN KRINGLE 4 AT 1.9 Å RESOLUTION

By

Anne Marie Mulichak

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Chemistry

1991

ABSTRACT

THE X-RAY CRYSTALLOGRAPHIC STRUCTURE DETERMINATION OF HUMAN PLASMINOGEN KRINGLE 4 AT 1.9 Å RESOLUTION.

by

Anne Marie Mulichak

Human plasminogen kringle 4 was crystallized from a solution of 30% PEG 8000, 0.12 M ammonium sulfate, 1.2 % dimethyl formamide. The crystals were orthorhombic, space group $P2_12_12_1$, with one molecule per asymmetric unit, and unit cell axes of a = 32.11 Å, b=49.09 Å, c= 49.39 Å. Three-dimensional X-ray intensity data were collected at 1.9 Å resolution. The structure was solved by the rotation-translation molecular replacement method, using coordinates of the highly homologous kringle from prothrombin fragment 1 as a model. Structure refinement was performed by restrained least-squares. The final refined K4 structure including 97 solvent molecules has an R-factor of 14.2% and an average B value of 18 $Å^2$. Most of the K4 three-dimensional structure is well-defined, although the terminal regions of the peptide chain preceding Cys1 and beyond Cys80 are disordered. The overall peptide folding, internal side chain structure and intricatehydrogen-bonding network observed for K4 is highly conserved from the prothrombin kringle. Interesting features which were not previously observed for the PT structure include the presence of a *cis* Pro residue and two alternate conformations for the Cys75 side chain. The K4 domain is known to bind lysine and zwitterionic ω-amino-carboxylic acids, and the observed binding-site complements this binding specificity, consisting of two oppositely charged Arg71 and Asp55/Asp57 ionic centers separated by a V-shaped hydrophobic trough formed by the indole side chains of Trp62 and Trp72. The orthorhombic crystal packing involves an intimate intermolecular

interaction in which a highly ordered sulfate ion is coordinated by Arg and Lys side chains from three separate kringle molecules and in which the Arg32 side chain of a symmetry-related molecule occupies the binding site in a ligand-like manner. The diffraction pattern of the orthorhombic K4 crystals displays reproducible relative intensity changes due to X-ray exposure which indicate non-random structural changes occurring within the crystal lattice. A preliminary investigation suggests these changes may involve the disordered terminal regions of the peptide chain and the K4 side chains of the ligand-like intermolecular interaction.

ACKNOWLEDGEMENTS

Sincere thanks are extended to Dr. Alexander Tulinsky for his guidance and support throughout the course of this work. My stay in his laboratory has been a rewarding experience in many ways. I would also like to express my appreciation to Dr. K.G. Ravichandran for his assistance with the molecular replacement methods and to Dr. Miguel Llinas for providing the plasminogen kringle 4 protein sample. Thanks also to all members of the Tulinsky research group, in particular Dr. Ewa Skrzypczak-Jankun and Dr. K. Padmanabhan for their invaluable help and for sharing so much of their time and experience. Finally, special thanks to my family for their constant support and encouragement.

TABLE OF CONTENTS

Chapter		Page
	List of Tables	vi
	List of Figures	viii
I.	Introduction	1
II.	Crystallization	15
III.	ACA Binding Experiments	25
IV.	Data Collection	29
V.	Data Reduction	39
VI.	Molecular Replacement	46
VII.	Structure Refinement	57
VIII.	Main Chain and Disulfide Structure	65
IX.	Side Chain Structure	83
X.	Intermolecular Interactions and Crystal Packing	105
XI.	The Structure of the Lysine-Binding Site	113
XII.	Solvent Structure	120
XIII.	Comparison of Plasminogen K4 and Prothrombin K1	134
XIV.	Comparison of Observed and Modeled Binding Sites	160
XV.	Investigation Into the Cause of Relative Intensity Changes	165
	Appendix A	174
	Appendix B	186
	List of References	189

LIST OF TABLES

Table		Page
1	Comparison of kringle sequences, aligned to show homology. Abbreviations used: PLG, plasminogen; PT, prothrombin; TPA, tissue-plasminogen activator; F XII, factor XII; UK, urokinase; VB, vampire bat salivary plasminogen activator; HGF, hepatocyte growth factor; APO, apolipoprotein(a).	4
2	Initial results of factorial search for K4 crystallization conditions.	16
3	Comparison of two plasminogen kringle 4 crystal forms.	18
4	Intensity measurements of conformation-sensitive reflections taken prior to three-dimensional data collection and between each resolution range of data collection.	36
5	Decay slopes determined based on various intensity measurements. Slope S calculated as (K-1)/(K+t), $K=I_t/I_t$.	44
6	Distribution of structure factor magnitudes (F) in various resolution shells for K4 three-dimensional data collection.	45
7	Highest cross rotation search results	50
8	Translation search results for rotation solution 1: (0,80,40). Highest correlation coefficients (C) are listed.	52
9	Translation search results for rotation solution 2: (85,80,100). Highest correlation coefficients (C) are listed.	53
10	Translation search results for rotation solution 3: (145,125,165). Highest correlation coefficients (C) are listed.	54
11	Results of PROFFT least-squares refinement at successive stages of resolution.	58
12	Comparison of agreement of structure with (+) and without (-) solvent molecules for various resolution ranges (8.0 Å maximum). R-factor is defined by $\Sigma(\|F_o\| - \|F_c\|) / \Sigma \|F_o\|$.	62
13	Observed secondary structural elements of K4 main chain. Reverse turns are classified according to dihedral angles of residues 2,3 [63]. Results of nmr structure are given for comparison [37].	73

14 Hydrogen-bonding interactions of the K4 main chain. Interactions forming β-sheet (β) or reverse turns (T) are indicated. Hydrogen atoms are assigned geometrically idealized positions. 74

15 Geometric parameters of K4 disulfide bonds. 78

16 Accessible surface area of K4 main chain and side chains by residue, calculated using programs ACCESS and ACCFMT [65] with a spherical probe of radius 1.4 Å. Percent accessible surface values are based on calculated accessible surface for the isolated residue. Asterisks indicate residues having disordered side chains. 85

17 Hydrogen bonding interactions involving K4 side chains. Hydrogen atoms were assigned geometrically idealized positions. Donor atom is denoted (D), acceptor atom (A). 87

18 Hydrogen-bonding interactions between sulfate anion and protein side chains. Hydrogen atoms are assigned geometrically idealized positions. Primed residue numbers denote hydrogen donors provided by symmetry-related K4 molecules. 124

19 Protein-solvent hydrogen bonds in which protein atom serves as hydrogen donor (D). 125

20 Protein-solvent hydrogen bonds in which protein atom serves as hydrogen acceptor. Potential interactions were accepted for protein-solvent distances within 3.5 Å. 126

21 Probable solvent-solvent interactions having distances within 3.5 Å. 132

22 Deviations in positions and torsion angles of conserved or highly homologous side chains of PGK4 and PTK1. No torsion angles are given for conserved Ala or Pro residues. 142

23 Final results of PROFFT least-squares refinement at 2.5 Å resolution for second set of orthorhombic data collected after the observation of characteristic relative intensity changes. 168

24 Distribution of magnitude of difference between structure factors calculated from two sets of orthorhombic K4 intensity data measured before and after characteristic relative intensity changes. 169

25 Reflections having largest structure factor discrepancies between first and second K4 data collections ($Diff^2 > 100$, where $Diff^2 = (F_1 - F_2)^2$). 169

LIST OF FIGURES

Figure		Page
1	Triple-loop three-disulfide structure of the kringle domain. Peptide chain cross-links indicate S-S bonds.	1
2	Schematic diagram of plasminogen peptide chain, containing five kringles (K1-K5). Amino (N) and carboxy (C) termini are indicated. Bold chain cross-links indicate disulfide bridges. Arrows designate 1) activation cleavage point and 2) autolytic cleavage point.	2
3	Ligands of kringle Lys-binding site. Abbreviations: ACA, ε-amino-caproic acid; AcLys, N-acetyl-L-lysine; AMCHA, *trans*-4-(amino-methyl)-cyclohexanecarboxylic acid; BASA, *p*-benzylamine sulfonic acid.	6
4	Amino acid sequence and disulfide structure of plasminogen K4. Numbering follows convention of plasminogen K5, with open circles representing deletions with respect to K5. Amino acid types are designated using common one-letter abbreviations.	9
5	Schematic diagram of K4 peptide sequence, showing peptide stretches which define the Lys-binding site. Boxes indicate residues which are conserved with prothrombin kringle 1. Darkened circle indicates position of insertion in prothrombin K1 with respect to K4.	12
6	Stereoview of modeled K4 lysine-binding site [37].	13
7	Orthorhombic crystal of human plasminogen kringle 4. Large crystal is approximately 1.5 mm in length.	19
8	Axial intensity distributions of orthorhombic K4 crystal displaying Type I difraction. X-rays at 2000 Watts power (50 kV,40 mA) for **a** and **c** axes, 4000 Watts (50 kV,80 mA) for **b** axis.	23
9	Axial intensity distributions of orthorhombic K4 crystal displaying Type II diffraction. X-rays at 2000 Watts power (50 kV,40 mA).	24
10	Intensity distributions of **b** and **c** axes for orthorhombic K4 crystal soaked in 10mM ACA for four days. X-rays at 2000 Watts (50kV, 40mA).	26

11 Intensity distributions of **b** axes for a) native and b) ACA-soaked monoclinic K4 crystals. X-rays at 2000 Watts (50kV,40mA). . . . 27

12 Schematic diagram of orthorhombic K4 crystal morphology with respect to crystallographic axes. 30

13 Intensity distributions of **a**, **b**, and **c** axes for orthorhombic K4 crystal prior to intensity data collection. All axes measured using X-rays at 2000 Watts (50 kV,40 mA). 31

14 Schematic diagram of four-circle diffractometer goniostat. The ω, ϕ, and χ circles are used to orient the crystal. The detector is supported by the 2θ circle. 32

15 Intensity peak profile of reflection (1 9 1) versus omega. 34

16 Absorption correction curve showing the dependence of the (2 0 0) reflection intensity on ϕ angle. 35

17 Decay curves for monitor reflections throughout intensity data collection. Resolution range 2-32° began at 0 hrs, 32-41° began at 23 hrs, 41-48° began at 46 hrs, 48-53° began at 71 hrs. . . . 38

18 Intensity distributions of **a**, **b**, and **c** axes for orthorhombic K4 crystal following data collection. X-rays at 2000 Watts (50 kV, 40 mA). 40

19 Two-theta dependence of left (squares) and right (circles) background measurements during intensity data collection. 41

20 Two-theta dependence of average negative intensity measurements during data collection. 42

21 Euler angles which relate rotated axes x', y' and z' to their original positions x, y, and z. 48

22 Dependence of R-factor on scattering angle for final refined K4 structure at 1.9 Å resolution. Dashed lines show theoretical coordinate error curves. 64

23 Ramachandran plot of final refined K4 structure. Open circles denote Gly residues; filled circles denote non-Gly residues. Energetically preferred zones are outlined in dashed lines. 66

24 Distribution of omega angles of refined K4 main chain peptide bonds (single *cis* peptide bond not included). 67

25 Stereoview of $2|F_o|-|F_c|$ electron density for Pro33 which has a *cis* peptide bond configuration. 68

26 Stereoview of K4 CA, C, N backbone and disulfide (bold) structure. 69

27 Stereoview of K4 backbone showing Pro side chains and main chain hydrogen-bonding interactions. Hydrogen bonds are indicated with dashed lines. 70

28 Average thermal factors of K4 main chain (solid lines) and side chain (dashed lines) structure. Breaks due to deletions with respect to K5 numbering convention. 76

29 Stereoview showing two alternate Cys75 side chain conformations. 79

30 Stereoview showing two relative disulfide positions resulting from two Cys75 side chain orientations. 80

31 Stereoview showing a) face-on and b)edge-on views of K4 structure. 84

32 Stereoview showing interactions of adjacent Arg10 and Arg52 side chains. Hydrogen bonds are indicated with dashed lines. . . . 89

33 Stereoview showing hydrogen-bonding interactions (dashed lines) of Gln23 side chain with His31 main chain. 91

34 Stereoview showing interactions of Asn53 side chain. Potential hydrogen bonds to Asp5 O, Asp57 N, and Asp57 O are indicated with dashed lines. 92

35 Stereoview of probable hydrogen bond (dashed line) between Lys35 main chain N and lone pair of His33 imidazole N. Hydrogen atoms are shown at geometric ideal positions. 94

36 Stereoview of K4 backbone and side chains forming internal hydrophobic core. 95

37 Steroview of perpendicular aromatic stacking interactions between Trp25, Trp62, Phe64, Trp72 and Tyr74. 97

38 Stereoviews of perpendicular interactions of a) Pro54 with Tyr41 ring and b) Pro61 with Tyr5 ring. 99

39 Stereoview showing perpendicular interactions of Pro30 and Pro68. 100

40. Stereoview showing parallel stacking of His31 and His33 rings with Pro30 and Phe64 side chains. 101

41 Stereoview showing well-defined electron density observed for Tyr9 side chain. 103

42 Stereoview showing interactions of sulfate anion with side chains of Lys35 and Arg71 of molecule 1, Lys58' of molecule 2, and Arg32" of molecule 3. Additional interactions occur between Arg71/Asp57' and Asp55/Arg32". 106

43 Stereoview illustrating two-fold screw axis in x-direction resulting from trimolecular kringle-kringle interactions at sulfate anion. Sulfate shown in bold. 108

44 Stereoview of intermolecular interaction between Asn43 side chain and Asn76' side chain of symmetry mate. Also shown is a solvent-bridged interaction between Gln7 and Asn76'Hydrogen bonds are indicated with dashed lines. 109

45 K4 molecules related by two-fold screw axis along y-direction. Residues involved in intermolecular side chain hydrogen bonds are shown in bold. 110

46 K4 molecules related by two-fold screw axis along z-direction. Solvent molecules which bridge side chains of adjacent symmetry mates are seen at protein-protein interfaces. 111

47 Solvent-bridged interaction between side chain of Ser14 and Ser69' of neighboring molecule. Hydrogen bonds are indicated with dashed lines. 112

48 Stereoview of K4 lysine-binding site. 114

49 Stereoview showing ion pair interactions which occur at lysine-binding site. Sulfate ion and side chains from second symmetry-related molecule are shown in bold. 118

50 Distribution of occupancies for ordered solvent molecules. Seven water molecules have an occupancy of 1.0. 121

51 Distribution of solvent temperature factors. 122

52 Electron density observed for sulfate anion in orthorhombic K4 crystal structure. 123

53 Stereoview of internal solvent molecule in K4 structrure. Dashed lines indicate hydrogen bonds which bridge Gln23 and Phe64 main chain atoms. 128

54 Stereoview showing two solvent molecules found in an empty cavity off the surface of the K4 structure. Dashed lines indicate potential hydrogen bonds. 131

55 RMS differences between K4 and PTK1 main chain (solid line) and side chain (dashed line) positions. Filled circles indicate conserved residues; open circles indicate highly homologous Tyr/Phe substitutions. 135

56 Comparison of CA backbone and disulfide structures of K4 (bold) and PTK1. 136

57 Comparison of CA,C,N,O main chain positions of K4 (bold) and PTK1. 138

58 Stereoview comparing conformations of Pro30, which appears as *cis* isomer in K4 (bold) but as *trans* isomer in PTK1. 139

59 Stereoview comparing conformations for inner disulfides in K4 (bold) and PTK1. Only the more similar of two observed K4 Cys75 side chain orientations is shown. 141

60 Comparison of hydrophobic core residues in K4 (bold) and PTK1 structures. 145

61 Stereoview showing similar conformations and hydrogen-bonding interactions of conserved residues Glu23 and Asn49 in K4 (bold) and PTK1. 147

62 Stereoview comparing conformations of conserved residues Arg52, Asn53, and Asp55 in K4 (bold) and PTK1. Dashed lines indicate conserved hydrogen-bonding interactions. 148

63 Stereoview comparing conformations of conserved Arg71 and homologous Gln/Glu34 residues in K4 (bold) and PTK1. 149

64 Stereoviews showing approximately 90° torsional differences in the conformations of a)Ser14 and b) Ser27 in K4 (bold) and PTK1. Dashed line indicates a hydrogen-bonding interaction between Ser27 OG and Thr29 N atoms which is observed only in K4. . . 151

65 Stereoview comparing conformations of Arg10 in K4 (bold) and PTK1. Dashed lines indicate K4 hydrogen-bonding interactions not observed in PTK1. 152

66 Stereoview comparing conformations of Glu73 side chains in K4 (bold) and PTK1. Both participate in similar hydrogen-bonding interactions with Thr16 side chain. 153

67 Comparison of K4 lysine-binding site (bold) and corresponding residues of PTK1. 155

68 Stereoview showing similar solvent positions observed in K4 (bold) and PTK1. Solvent molecules, W3 and W18, labeled according to K4 quality factor ranking. 156

69 Stereoview comparing conserved solvent position, W13, in K4 (bold) and PTK1. 158

70 Stereoview comparing positions of K4 internal solvent molecule (W1) and nearest solvent site of PTK1 structure. K4 residues are shown in bold. 159

71 Comparison of observed (bold) and modeled K4 lysine-binding sites. 161

72 Fit of average structure factors from second orthorhombic K4 data collection, applying scale of 1.27 and ΔB correction of +4 (solid line), to those of original data collection (dashed line). 166

73 Distribution of scattering angles for set of 334 reflections having very large discrepancies between original and second orthorhombic data collections ($(|F_1| - |F_2|)^2 > 100$). 170

74 Stereoview showing electron difference density observed near amino terminal of K4 peptide chain when map is calculated using only 334 reflections having very large discrepancies between original and second intensity data sets. 172

75 Stereoview showing electron difference density observed in vicinity of ligand-like intermolecular interaction when difference map is calculated using only 334 reflections having very large discrepancies between original and second intensity data sets. 173

I. INTRODUCTION

The kringle structure is a type of protein folding domain, characterized by a unique pattern of three peptide loops formed by three disulfide bridges (Figure 1), which appears as a structurally and functionally independent unit of protein architecture [1,2]. Kringles are found repeatedly in the structures of certain proteins, particularly those blood plasma proteins involved in the formation and dissolution of blood clots. Usually preceding a catalytic domain of enzymes, kringles are known in many cases to have a role in protein recognition and enzyme specificity.

Among proteins of the blood coagulation cascade, prothrombin includes two kringles [3], one of which mediates the binding of Factor Va [4]; clotting factor XII carries a single kringle [5]. The fibrinolytic proenzyme plasminogen contains five kringles (Figure 2), which are involved in binding to the fibrin matrix of the blood clot and the inhibitor antiplasmin [6,7]. A number of activators of the fibrinolytic pathway also contain kringle domains: urokinase, an activator of plasminogen found in urine, has one kringle [8,9] and tissue-plasminogen activator contains a pair [10], one of which also binds fibrin [11,12]. A kringle domain has also been identified in the structure of a plasminogen activator found in vampire bat saliva and is believed to be involved

Figure 1. Triple-loop three-disulfide structure of the kringle domain. Peptide chain cross-links indicate S-S bonds.

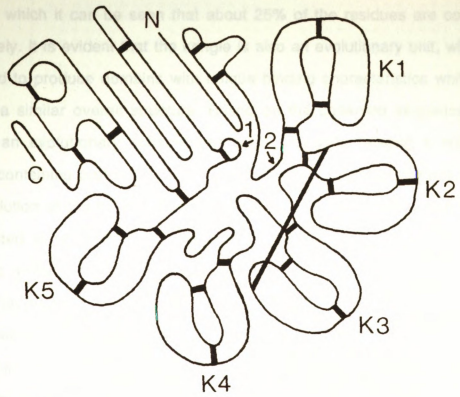


Figure 2. Schematic diagram of plasminogen peptide chain, containing five kringles (K1-K5). Amino (N) and carboxy (C) termini are indicated. Bold chain cross-links indicate disulfide bridges. Arrows designate 1) activation cleavage points and 2) autolytic cleavage point.

in a molecular interaction with fibrin [13]. Recently the sequence of a platelet protein, hepatocyte growth factor, was found to indicate the presence of a series of four kringle domains [14]. However, most impressive is the series of 38 kringles found in the plasma glycoprotein apolipoprotein(a) [15].

The kringles contain 80-85 residues and have highly conserved amino acid sequences. A comparison of selected kringle sequences is shown in Table 1, from which it can be seen that about 25% of the residues are conserved absolutely. It is evident that the kringle is also an evolutionary unit, which has diverged to produce domains with unique binding characteristics while maintaining a similar overall structure. Based on the observed sequence homologies, an evolutionary scheme has been proposed relating a number of kringle-containing proteins [16]. An apparent lack of correspondence between the evolution of the kringle, catalytic, and other domains with which they are associated suggests that the kringle may have been transferred between proteins which had already diverged, and thus, that the kringle may be somewhat genetically autonomous.

Plasminogen (Figure 2) is a 91,500 kD glycoprotein found in all body fluids and most tissues but primarily in the blood plasma. As part of the fibrinolytic pathway, plasminogen is the precursor of the enzyme plasmin, which is directly responsible for digestion of the fibrin matrix of blood clots into soluble fragments. Plasmin itself is not a highly selective enzyme and is capable of hydrolyzing a number of biological proteins; therefore, the regulation of the enzyme lies in its relationships with additional activator and inhibitor proteins. In the bloodstream, plasminogen is converted to plasmin by tissue-plasminogen activator (t-PA) [17], which is released by the endothelial cells of the blood vessel wall. This activation occurs efficiently only at the fibrin surface as it is enhanced by ternary complex formation, in which plasminogen and t-PA are

Table 1. Comparison of kringle sequences, aligned to show homology. Abbreviations used: PLG, plasminogen; PT, prothrombin; TPA, tissue-plasminogen activator; F XII, factor XII; UK, urokinase; VB, vampire bat salivary plasminogen activator; HGF, hepatocyte growth factor; APO, apolipoprotein(a). Only APO K2, 24-fold repeat of APO molecule is listed. Relative sequence numbers given with respect to PLG K5.

```
                1                   10                  20                  30                  40
PLG K1    C K T G N G K N Y R G T M S K T K N G I T C Q K W S S T S P H R P R F   S P A T H P S
   K2     C M H C S G E N Y D G K I S K T M S G L E C Q A W D S Q S P H A H G Y   I P S K F P N
   K3     C L K G T G E N Y R G N V A V T V S G H T C Q H W S A Q T P H T H N R   T P E N F P C
   K4     C Y H G D G Q S Y R G T S S T T T T G K K C Q S W S S M T P H R H Q K   T P E N Y P N
   K5     C M F G D G K G Y R G K R A T T V T G T P C Q D W A A Q E P H R H S I F T P E T N P R
PT K1     C A E G V G M N Y R G N V S V T R S G I E C Q L W R S R Y P H K P E I N S   T T H P G
   K2     C V P D R G R E Y R G R L A V T T S G S R C L A W S S E Q A K A L S K D Q   D F N P
TPA K1    C Y E D Q G I S Y R G T W S T A E S G A E C T N W N S S A L A Q K P Y   S G   R R P D
   K2     C Y F G N G S A Y R G T H S L T E S G A S C L P W N S M I L I G K V Y   T A   Q N P S
F XII     C Y D G R G L S Y R G L A R T T L S G A P C Q P W A S E A     T Y R N V T A E Q A R
UK        C Y E G N G H F Y R G K A S T D T M G R P C L P W N S A T V L Q Q T Y H A H R S D A L
VB        C Y K D Q G V T Y R G T W S T S E S G A Q C I N W N S N L L T R R T Y N G R R S D A I
HGF K1    C I I G K G R S Y K G T V S I T K S G I K C Q P W S S M I P H E H S F   L P S S Y R G
   K2     C I Q G Q G E G Y R G T A N T I W N G I P C Q R W D S Q Y P H K H D M   T P E N F K C
   K3     C Y R G N G K D Y M G N L S Q T R S G L T C S M W N K N M E D L H R H I F W E P D A S
   K4     C Y H G N G Q S Y R G T Y S T T V T G R T C Q A W S S M T P H S H S R   T P E Y Y P N

                                   50                  60                  70                  80
PLG K1    E       G L E E   N Y C R N P D N D P Q G P W C Y T T   D P E K R Y D Y C D   I L E C
   K2     K       N L K K   N Y C R N P D R E L R   P W C F T T   D P N K R W E L C D   I P R C
   K3     K       N L D E   N Y C R N P D G K R A   P W C H T T   N S Q V R W E Y C K   I P S C
   K4     A       G L T M   N Y C R N P D A D K   G P W C F T T   D P S V R W E Y C N   L K K C
   K5     A       G L E K   N Y C R N P D G N V G G P W C Y T T   N P R K L Y D Y C D   V P Q C
PT K1     A       D L R E   N F C R N P D G S I T G P W C Y T T   S P T L R R E E C S   V P V C
   K2     A V P       L A E   N F C R N P D G D E E G A W C Y V A   S Q P G D F E Y C N   L N Y C
TPA K1    A I R L G L G N H N Y C R N P D R D S K   P W C Y V F K A G K Y S S E F C S   T P A C
   K2     A Q A L G L G K H N Y C R N P D G D A K   P W C H V L K N R R L T W E Y C D   V P S C
F XII     N W     G L G G H A F C R N P D N D I R   P W C F V L N R D R L S W E Y C D   L A Q C
UK        Q     L G L G K H N Y C R N P D N R R R   P W C Y V Q V G L K P L V Q E C M   V H D C
VB        T     L G L G N H N Y C R N P D N N S K   P W C Y V I K A S K F I L E F C S   V P V C
HGF K1    K       D L Q E   N Y C R N P R G E E G G P W C F T S   N P E V R Y E V C D   I P Q C
   K2     K       G F D D   N Y C R N P D G Q P R   P W C Y T L   D P H T R W E Y C A   I K T C
   K3     K       D L R E   N Y C R N P D G S E S   P W C F T T   D P N I R V G Y C S Q I P N C
   K4     K           L N E   N Y C R N P D D D A H G P W C Y T G   N P L I P W D Y C P   I S R C
APO K2    A       G L I M   N Y C R N P D A V A A   P Y C Y T R   D P G V R W E Y C N   L T Q C
```

both bound to the fibrin substrate [18]. On the molecular level, activation involves cleavage of an Arg-Val bond of the proenzyme to produce a disulfide-stabilized two-chain structure (Figure 2). This is quickly followed by the autolytic cleavage and release of a 76-residue amino terminal activation peptide [19]. The activated plasmin consists of a light chain (MW ~25,000) containing the catalytic site, and a heavy chain (MW ~60,000) composed largely of the five kringle domains (MW ~10,000 each). Fibrinolysis of the blood clot then proceeds as a series of cleavages at exposed Lys-X peptide bonds of fibrin. Once dissolution of the clot is complete, free plasmin is released into the blood plasma, where it is rapidly inactivated by the circulating inhibitor antiplasmin. Inactivation is a two-step process: rapid formation of a reversible complex between plasmin and antiplasmin, followed by a slower conversion to an irreversible complex [20].

Among these interactions, the plasminogen kringles have been implicated in the binding of the fibrin substrate [21] and antiplasmin [22], and, also, in the intramolecular binding of the activation peptide [7]. Specifically, these interactions occur via a number of lysine-binding sites found among the kringles, one strong site and 3-4 weaker sites [23,24,25]. The strong binding site, found on kringle 1 (K1), appears to be important in the interactions with fibrin and antiplasmin [21,26]. A weaker site occurs on kringle 4 (K4), which also binds fibrin in the isolated state, although its function in the intact protein is uncertain [21]. In addition, K1 and K4 bind a number of ligands which serve as lysine analogs (Figure 3), including ε-aminocaproic acid (ACA), N-acetyl-L-lysine (AcLys), *trans*-4-(aminomethyl)-cyclohexanecarboxylic acid (AMCHA), and *p*-benzyl-amine-sulfonic acid (BASA) [24,27]. Each of these ligands bears a positive and negative ionic group separated at a distance of approximately 6.8 Å by a hydrocarbon chain or ring. The binding of these ligands by

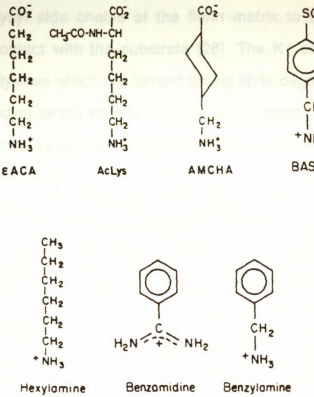


Figure 3. Ligands of kringle Lys-binding site. Abbreviations: ACA, ε-amino-caproic acid; AcLys, N-acetyl-L-lysine; AMCHA, *trans*-4-(amino-methyl)-cyclohexanecarboxylic acid; BASA, *p*-benzylamine sulfonic acid.

plasminogen is known to have an antifibrinolytic effect *in vivo*.

Kringle 5 also carries a weak lysine-binding site with similar binding properties, but which shows no requirement for a negatively charged group, having a high affinity for non-zwitterionic ligands such as benzylamine, benzamidine, and hexylamine (Figure 3), as well as ACA [28]. It has been proposed that the K5 site may be important in the mechanism of fibrinolysis by binding internal lysyl side chains of the fibrin matrix to bring the plasmin catalytic site in contact with the substrate [28]. The K1 site then binds the carboxy terminal lysines which are formed during fibrin degradation. Relatively weak transient binding by the K5 site, coupled with alternation of the K1 site to newly formed terminal lysines, would allow plasmin to move efficiently along the fibrin chain during dissolution. The K5 site probably also corresponds to the "AH-site"[29], which appears to interact with the plasminogen activation peptide and thus to hold the molecule in a closed conformation [30].

Not only is the kringle domain of interest due to its recurrence as a structural unit among this family of proteins, but also the apparent regulatory role of certain kringles in important biological interactions suggests a variety of pharmaceutical opportunities. Thus, a considerable body of research has been concentrated on understanding the structure and behavior of kringles. Prior to this work, however, only the structure of bovine prothrombin kringle 1 (PTK1), as included in PT fragment 1, had been determined crystallographically. This structure was solved in this laboratory at 2.25 Å, with an R-factor of 17%.[31,32] The prothrombin kringle, having the overall form of an oblate ellipsoid, displayed an intricate folding pattern with numerous β-turns, but otherwise little organized secondary structure. The structural "nucleus" of this folding appeared to include three elements: 1) a sulfur cluster near the center of gravity of the kringle, formed by the close approach of two disulfides, 2) antiparallel structure

involving two strands of highly conserved residues, and 3) an aromatic stacked cluster [33].

Various alternate methods have been employed in the investigation of additional kringle structures, particularly those of plasminogen and t-PA. Probably the most thoroughly studied of these is plasminogen K4. The amino acid sequence and disulfide structure of K4 are shown in Figure 4. The K4 folding has been investigated by circular dichroism experiments [34], which indicated the presence of 64% β-structure, 30% β-turns, and 6% random coil. Specific regions of K4 secondary structure were proposed by Castellino *et al.* based on Chou-Fasman calculations [34], including β-sheet in the amino acid stretches Ala44-Asn53, Trp62-Thr66, and Val70-Tyr74 and twelve β-turns, with the greatest probabilities calculated at residues His3-Gly6, Asn53-Ala56, Asp67-Val70, and Cys80-Thr83. Seven of the twelve suggested turns coincide with β-turns reported for the PTK1 folding.

Results from aromatic H-nmr, acid-base titration, nuclear Overhauser, and two-dimensional correlated nmr experiments [35,36] have indicated the presence of a buried hydrophobic core stabilizing the kringle fold. Centered on Leu46, this core is composed largely of an aromatic cluster formed by residues Trp25, Tyr41, Tyr50, and Trp62, most of which are strictly conserved, and may also include His33 and Leu77. The interaction of a Met methyl group with these residues has also been suggested [35]. Models of the K4 backbone folding based on early nmr experiments involved mainly a clustering of the implicated hydrophobic core residues. Based on an observed cross-relaxation between Val(a) and Trp72 resonances, DeMarco *et al.* [35] also suggested that the amino-terminal peptide is folded back toward the kringle.

Recently, a more detailed model of K4 was proposed by Atkinson and Williams [37] from observations of H-nmr spectra and NOE connectivities. In

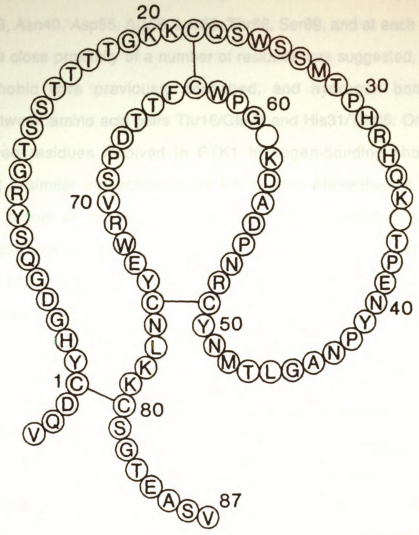


Figure 4. Amino acid sequence and disulfide structure of plasminogen K4. Numbering follows convention of plasminogen K5, with open circles representing deletions with respect to K5. Amino acid types are designated using common one-letter abbreviations.

agreement with the previous work, these authors reported little regular secondary structure other than a short stretch of antiparallel β-sheet between residues Trp62-Phe64 and Tyr74-Trp72 and a possible β-type interaction between Cys22 and Thr65. Turns in the protein backbone were predicted at Gly19, Ser26, Asn40, Asp55, Ala56, Lys58, Thr66, Ser69, and at each of the six prolines. The close proximity of a number of residues was suggested, including the hydrophobic core previously described, and hydrogen bonds were predicted between amino acid pairs Thr16/Glu73 and His31/Thr65. Only one of the conserved residues involved in PTK1 hyrdogen-bonding showed any indication of a similar interaction in the K4 spectra. Nevertheless, a three-dimensional model of the K4 structure, generated by employing distance geometry algorithms with NOE-deduced interproton distances, was found to be highly similar to the crystallographic PTK1 structure.

Considerable effort has been directed at understanding the nature of the K4 lysine-binding site, and numerous binding and spectroscopic studies have identified the key residues involved. Chemical modifications of residues Asp57, Arg71 and Trp72 were found to prevent ACA binding [38,39]. Similarly, photo-oxidation of K4 histidine residues greatly reduced affinity for the ligand [40]. From such studies and a knowledge of the ligand-binding specificity, a model of K4 binding developed in which the ligand amino and carboxyl termini interact electrostatically with the complementary charged centers Asp55/Asp57 and His31/Arg71 [35,40,41]. In aromatic H-nmr experiments, ligand binding was found to significantly perturb the chemical shifts of Trp62, Phe64, and Trp72, suggesting that these residues are located at the surface of the binding site where they interact strongly with the ligand [41,42]. Also affected to a lesser extent were Trp25, His31, and Tyr41. Strong NOE connectivities were observed between these residues and those comprising the hydrophobic core [35],

implying that the binding site is contiguous to the latter and is thus primarily supported by the inner kringle loop. This conclusion is consistent with the observation that K4 binding ability is retained after reduction of the Cys1-Cys80 disulfide bridge [43].

Recently, the K4 lysine-binding site, defined by the amino acid stretches 31-35, 54-58, 61-64, and 71-75, was modeled by Tulinsky *et al.* [44] based on the three-dimensional structure of PTK1 and on information from H-nmr observations. As can be seen in Figure 5, the conservation in amino acid sequence between K4 and PTK1 is concentrated on the inner loop of the kringle, which supports the binding site. However, a potential complication to the modeling is an insertion of a Thr residue in PTK1 relative to K4 at position 59, adjacent to the binding site.

According to the modeling procedure used, the basic conformations of the peptide backbone and conserved side chains of PTK1 were retained. Non-conserved K4 side chains for which the corresponding PTK1 residue was sufficiently similar were modeled to follow guide coordinates of the PTK1 side chain. Examples include the modeling of Arg32 (PT Lys), Gln34 (PT Glu), and Asp55 (PT Ser). In some cases, the modeling was based primarily on nmr observations. In the cases of side chains for which there were no or poor guide coordinates and no additional information, an extended energy minimized conformation was used, as can be seen for Lys35 (PT Ile) and Lys58 (PT Gly) (Figure 6). Finally, the binding site was subjected to energy minimization in both a free and ligand-bound state. The modeled K4 lysine-binding site, shown in Figure 6, is described as a relatively open depression lined by the lipophilic side chains of Trp62, Phe64, Trp72, with anionic Asp55/Asp57 and cationic Arg71 centers at the extremes.

In the present work, the X-ray crystal structure determination of human

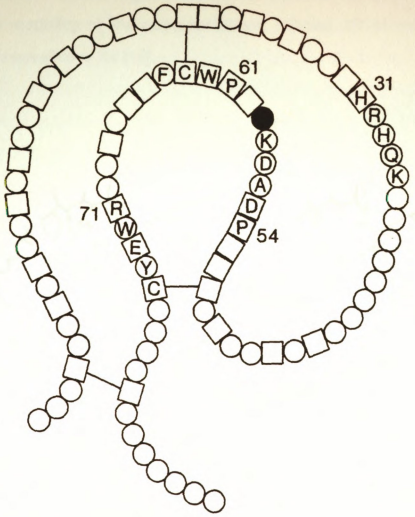


Figure 5. Schematic diagram of K4 peptide sequence, showing peptide stretches which define the Lys-binding site. Boxes indicate residues which are conserved with prothrombin kringle 1. Darkened circle indicates position of insertion in prothrombin K1 with respect to K4.

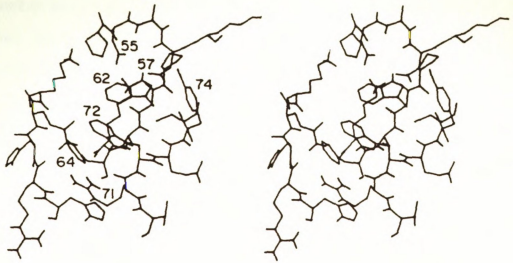


Figure 6. Stereoview of modeled K4 lysine-binding site [37].

plasminogen kringle 4 was undertaken. In addition to obtaining a better understanding of kringle domains and their binding ability, a secondary goal of this research was to assess the validity of the methods used in the previous modeling study. Furthermore, the K4 structure should prove uniquely suited to the further modeling of the apolipoprotein(a) kringles, 37 of which display 75-85% conservation to K4 [15].

II. CRYSTALLIZATION

Human plasminogen kringle 4 is isolated from elastase digestion of the enzyme as a heterogenous mixture, 70% having Ala86 and 30% having Val88 at the carboxy terminus [6]. The K4 sample used was provided in the form of a lyophilized powder by Dr. Miguel Llinas, and was stored at freezer temperatures. Kringle 4 was first crystallized in this laboratory by Dr. Chang Park [45] using the vapor diffusion method, from a drop consisting of one part 2-5 mg/ml aqueous protein solution to one part reservoir solution of 45% saturated ammonium sulfate in tris saline buffer, pH 7.4. The crystals appeared as small needles, tending to form rosettes, and could only be grown to an appreciable size with repeated seeding. They were characterized to be monoclinic, of space group $P2_1$, with four molecules per unit cell, two molecules per asymmetric unit. All crystals observed were found to be twinned along the **a*** axis.

In an effort to find a new set of crystallization conditions which might produce larger, untwinned crystals, a factorial search was made with K4. The factorial search [46] is a method in which a limited set of solutions, representing a random but balanced distribution of common precipitants, salts and buffers, is used to make a rough survey of possible crystallization conditions. Solutions resulting in precipitation of the protein can then be used as a basis for more detailed crystallization experiments. A set of 31 factorial solutions was used, from a suggested list by Dr. Marcos Hatada [47]. Crystallization trials were performed using the vapor diffusion method with 5 µl hanging drops consisting of one part 10 mg/ml K4 solution and one part factorial solution. Results of the initial factorial search are shown in Table 2. To test modified crystallization conditions while conserving protein, hanging drops producing amorphous solid or microcrystals were removed from the well solution, allowed to equilibrate

Table 2. Initial results of factorial search for K4 crystallization conditions.

No.	Buffer	Salt	Precipitant	Results
1	0.1 M Acetate	0.2 M Calcium Chloride	30% MPD	Thin plates, inorganic
2	—	0.8 M Na,K Tartrate	—	1.5 M: brown amorph ppt
3	—	—	0.4 M A.P.	—
4	0.1 M Tris	—	3.0 M A.S.	50 rosettes of microcrystals
5	0.1 M Hepes	0.2 M Na Citrate	30% MPD	50% MPD: phase separation
6	0.1 M Acetate	0.2 M Mg Chloride	27% PEG 4000	—
7	—	1.2 M Na Citrate	—	—
8	0.1 M Acetate	0.2 M Ca Chloride	20% PEG 20K	clusters of thin plates, inorg.
9	0.1 M Citrate	0.2 M Ammon. Acetate	30% PEG 400	—
10	0.1 M Acetate	—	1.5 M A.P.	—
*11	0.1 M Hepes	0.2 M Ammon. Sulfate	2.0 M Na,K Phosphate	rosettes of needles
12	0.1 M Tris	0.2 M Na Citrate	20% PEG 400	—
13	0.1 M Hepes	0.2 M Ca Chloride	25% PEG 4000	dark amorph. ppt.
14	0.1 M Imidazole	0.1 M Ammon. Sulfate	30% PEG 8000	0.5 M A.S.: phase separation
15	0.1 M Tris	0.2 M Li Sulfate	30% PEG 4000	—
16	0.2 M Acetate	1.0 M Li Sulfate	—	—
17	0.1 M Tris	0.2 M Ammon. Phosphate	30% MPD	—
18	0.1 M Tris	0.2 M Ammon. Acetate	2.0 M Na,K Phosphate	amorphous ppt.
19	0.1 M Citrate	0.1 M Ammon. Sulfate	30% PEG 8000	phase separation
20	0.1 M Imidazole	—	30% MPD	—
21	0.1 M Hepes	0.2 M Mg Chloride	30% PEG 400	amorphous ppt.
22	0.1 M Tris	0.2 M Na Acetate	30% PEG 4000	40% PEG: dark amorph ppt.
23	0.1 M Acetate	0.2 M Ca Chloride	30% Isopropanol	thin plates, inorganic
24	0.1 M Tris	0.2 M Na Acetate	—	—
25	0.1 M Tris	0.5 M Ammon. Acetate	30% MPD	—
26	0.1 M Hepes	0.2 M Na Acetate	30% Isopropanol	—
27	0.1 M Imidazole	0.2 M Na,K Tartrate	30% PEG 8000	—
28	0.1 M Hepes	1.0 M Na,K Tartrate	—	amorphous ppt.
29	—	0.2 M Ammon. Sulfate	30% PEG 400	50% PEG: phase separation
*30	—	0.1 M Ammon. Sulfate	20% PEG 4000	40% PEG: column-shaped xtals
31	—	2.0 M Ammon. Sulfate	—	—

Abbreviations: MPD 2-methyl-2,4-pentanediol;A.P. ammonium phosphate;A.S. ammonium sulfate; PEG polyethylene glycol. Buffers: Acetate pH 4.6, Citrate pH 5.6, Imidazole pH 6.6, Hepes pH 7.5, Tris, pH 8.5. Asterisks indicate conditions used for further experimentation.

against a well of the buffer or H_2O to redissolve the protein, and returned to a well of a similar crystallization solution having a lowered precipitant concentration. Similarly, drops producing no precipitate were re-equilibrated against a well solution having an incresed precipitant concentration. Based on such experimentation, subsequent efforts were largely concentrated on variations of two sets of conditions: 1) 0.1 M Hepes buffer, 0.2 M ammonium sulfate, 2.0 M sodium potassium phosphate, and 2) 40% polyethylene glycol (PEG) 4000, 0.1 M ammonium sulfate. The K4 crystals produced with these conditions displayed a strong tendency to grow in rosettes or large clusters rather than individually. A number of organic solvents, including dioxane, n-butanol, dimethyl formamide (DMF), t-butyl alcohol and isoproponal, were added in small amounts (1-2%) to the well solutions, some of which appeared to promote the growth of separate independent crystals. After approximate conditions were determined, crystals were grown from 20-30 µl sitting drops.

From these experiments, several different K4 crystal forms were grown under similar conditions. Monoclinic crystals isomorphous to those of Park were grown using a reservoir solution of 29-30% PEG 8000 (Fisher Carbowax flakes), 0.12 M ammonium sulfate, with 0.8-1.0% n-butanol as an additive, pH 6.2 and an initial protein concentration of 5 mg/ml. These lath-shaped crystals tended to grow to a larger size and were more likely to grow as individual crystals than those monoclinic crystals previously obtained but were crystallographically identical to them. They have unit cell parameters of a=32.78(3) Å, b=49.17(2) Å, c=46.27(3) Å, β=100°, with four molecules per unit cell and two molecules per asymmetric unit. They also have a relatively high protein fraction of 66% [45].

A second crystal form of K4 was grown under similar conditions, but using 1.2% dimethyl formamide as an additive, pH=6.0. These crystals are

orthorhombic, of space group $P2_12_12_1$, with four molecules per unit cell, one molecule per asymmetric unit, and with unit cell parameters a=32.11(1) Å, b=49.09(2) Å, c=49.39(3) Å. They also have a high protein content of 62% [45]. The orthorhombic crystals are similar in morphology to the monoclinic form but are generally thicker in the smallest dimension and often show a characteristic end face development (Figure 7). A comparison of the monoclinic and orthorhombic crystal forms is given in Table 3.

The two additives appeared to promote the growth of different crystal forms, as evidenced by the fact that orthorhombic crystals were grown using DMF, even though the solution was seeded with monoclinic crystals. However, this tendency is not absolute, and both forms have been found to grow in the presence of either organic solvent. On the other hand, the two forms have never been grown simultaneously.

Less commonly, a third crystal form of K4 has also been grown under seemingly identical conditions, in the presence of either n-butanol or DMF, and in some cases along with developing orthorhombic crystals. The morphology of these crystals is that of flat plates, which may attain very large surface areas

Table 3. Comparison of two plasminogen kringle 4 crystal forms.

Crystal form	Monoclinic	Orthorhombic
Space group	$P2_1$	$P2_12_12_1$
Mol./unit cell	4	4
Mol./asym. unit	2	1
Lattice Parameters		
a (Å)	32.78	32.11
b (Å)	49.17	49.09
c (Å)	46.27	49.39
β (deg)	100.67	90.0
Protein fraction (%)	66	62
Vm (Å^3/dalton)	1.86	1.99

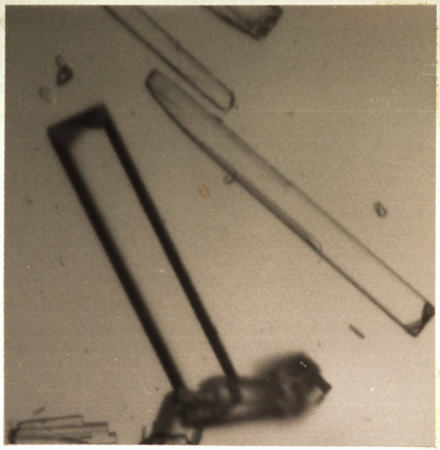


Figure 7. Orthorhombic crystal of human plasminogen kringle 4. Large crystal is approximately 1.5 mm in length.

(over 1 Å^2), but they are extremely thin. As observed with the other two crystal forms, they display a tendency to grow in clusters. Although surprisingly strong in the mother liquor, these crystals splinter when mounted; therefore, this crystal form was not characterized.

Both monoclinic and orthorhombic forms of K4 diffract X-rays well, consistent with the high protein fractions, and are quite stable with respect to deterioration due to radiation exposure. However, the monoclinic crystals invariably exhibit the same twinning as was previously observed, with crystal and twin **c*** axes oriented 20 degrees apart. In addition, they are often further split, with secondary "**c*** axes" found within one to two degrees of both crystal and twin. The principal twinning alone greatly complicates the collection and interpretation of intensity data; the secondary splitting, at such small angular differences, renders the crystals useless. Due to the twinning complications, as well as the fact that the monoclinic crystals contain two rather than one kringle in the asymmetric unit, the orthorhombic form was chosen for the present investigation.

Some additional points regarding the crystallization of orthorhombic K4 deserve mention. Firstly, as mentioned above, the crystallization conditions have a somewhat acidic pH relative to the physiological pH 7.0. Although an effort was made to incorporate buffer into the crystallization solution and thus ensure a neutral pH, this resulted in inferior crystal growth. Once grown, the crystals were stored in a solution having an increased PEG concentration of 38%. During experimentation with both crystallization and storing solutions it was discovered that an increase in ammonium sulfate concentration increases K4 solubility. As the crystallization appears to be highly sensitive to small changes in ionic strength, it was important not to raise the salt concentration in the storing solution. This sensitivity may also account for the difficulty

encountered upon including buffer.

Finally, a novel method of seeding was used during this work which proved to be particularly effective and reproducible. Although the original crystal growth generally succeeded in producing orthorhombic crystals of a size suitable for diffraction experiments, macroscopic seeding was used in many cases to further improve crystal size or to ensure growth of orthorhombic rather than monoclinic crystals. A reverse seeding method was used in which additional dissolved protein was slowly introduced to the crystal, in contrast to the commonly used procedure of depositing a seed crystal into a previously prepared protein drop [48]. Crystals which had reached their maximum size in the original mother liquor were first transferred to a fresh 10 µl sitting drop of reservoir solution containing no dissolved protein. The crystal was then allowed to rest for one day to equilibrate with the new conditions, and in so doing, to allow the outermost layer of the crystal to dissolve, exposing a fresh surface for renewed growth. An approximately 10 mg/ml K4 solution was then added in 2-3 µl aliquots, allowing one day between additions, until crystal growth was observed. Generally, the addition of 5-6 µl protein solution was required to initiate growth, which continued for several days. Once crystal growth had ceased, the entire procedure could be repeated.

The advantage of this method in the present case is that it provides a simple and reliable alternative to the time-consuming repetitive rinsing of seed crystals which is otherwise necessary to expose fresh crystal surface and dissolve any microscopic crystalline material that may serve as a nucleation site for secondary crystal growth. However, it cannot be applied directly for all crystallization conditions but may require some experimentation to find an appropriate drop solution and resting period so that the seed crystal is not immediately dissolved.

During experiments with numerous orthorhombic K4 crystals, two significantly different diffraction patterns have been observed. Although most newly mounted crystals produce a single characteristic pattern (Type I), in some cases, striking changes in relative intensities occur to give a second pattern (Type II), which appears to reflect some conformational change in the kringle structure. Figures 8 and 9 show the axial intensity distributions from a single orthorhombic K4 crystal as it was first observed (Figure 8) and after it had been exposed for approximately 20 hrs to X-rays and, due to problems with the original mount, had been removed from the capillary and remounted (Figure 9). Most notable are the intensity changes observed for reflections (2 0 0), (8 0 0), (0 0 14), (0 0 16), (0 0 18), (0 0 22), (0 10 0) and (0 16 0). In addition to these intensity changes, crystals displaying the Type II diffraction tend to have slightly shorter unit cell dimensions, approximately a=32.15(2) Å, b=49.01(2) Å, c=49.09(3) Å. This phenomenon has also been observed during intensity data collections. In these cases, the changes appear to be gradual and radiation dependent, as mounted crystals do not undergo relative intensity changes with time in the absense of X-ray exposure, even when a partial change has already been induced by previous exposure.

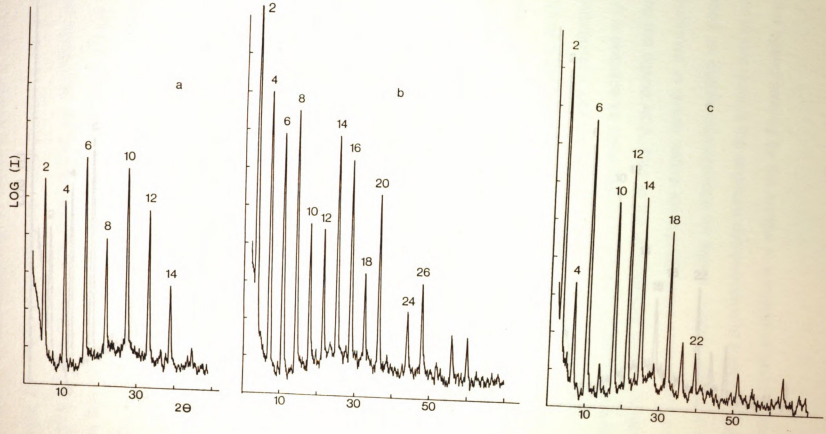


Figure 8. Axial intensity distributions of orthorhombic K4 crystal displaying Type I diffraction. X-rays at 2000 Watts power (50kV,40mA) for **a** and **c** axes, 4000 Watts (50kV,80mA) for **b** axis.

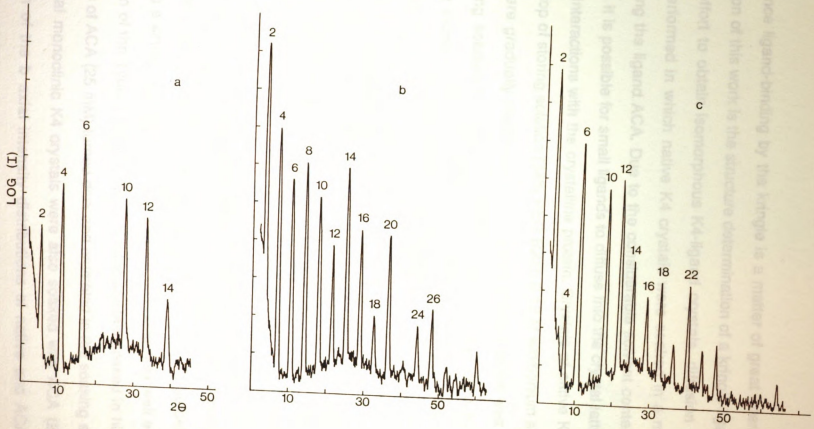


Figure 9. Axial intensity distributions of orthorhombic K4 crystal displaying Type II diffraction. X-rays at 2000 Watts power (50kV,40mA).

III. ACA-BINDING EXPERIMENTS

Since ligand-binding by the kringle is a matter of great interest, a logical extension of this work is the structure determination of a kringle-ligand complex. In an effort to obtain isomorphous K4-ligand crystals, diffusion experiments were performed in which native K4 crystals were soaked in a mother liquor containing the ligand ACA. Due to the considerable solvent content of protein crystals, it is possible for small ligands to diffuse into the crystal lattice and form binding interactions with the crystalline protein. Crystals of native K4 in a 40 µl sitting drop of storing solution (38% PEG 8000, 0.12 M ammonium sulfate, 1.4% DMF) were gradually transferred to a soaking solution which was identical to the storing solution but included 10 mM ACA. This was done by slowly exchanging increasingly large aliquots of the storing solution with the soaking solution (5,10,15,20,25,30,40 µl) at approximately 12 hr intervals. Crystals were soaked for 4 to 7 days.

The intensity distribution for axes **b** and **c** of an orthorhombic K4 crystal soaked in 10 mM ACA are shown in Figure 10. The intensities agree well with those of the native K4 Type II distribution, indicating that no binding of the ligand occurred. The same results were obtained for a number of similarly treated K4 crystals. It should be noted that the ACA-soaked orthorhombic crystals were first mounted and their diffraction examined to confirm the crystal type before being returned to a sitting drop for soaking. This prior mounting may well explain the observation of the Type II pattern. Orthorhombic crystals soaked in higher concentrations of ACA (25 mM) turned dark and crumbled in the soaking solution.

Several monoclinic K4 crystals were also soaked with ACA (25mM). A comparison of the **b** axial intensity distributions of native and ACA-soaked monoclinic crystals is shown in Figure 11. The differences in relative intensities

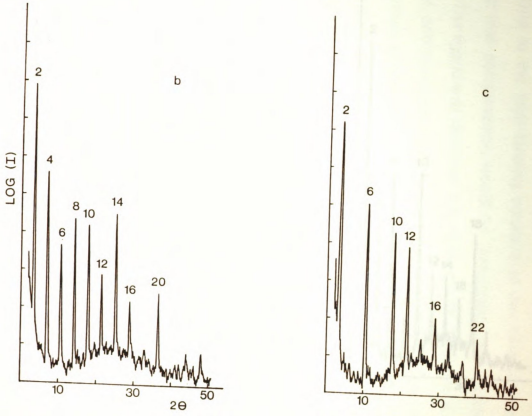


Figure 10. Intensity distributions of **b** and **c** axes for orthorhombic K4 crystal soaked in 10 mM ACA for four days. X-rays at 2000 Watts power (50kV,40mA).

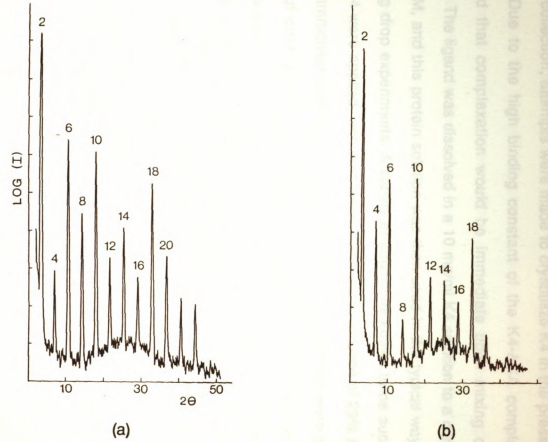


Figure 11. Intensity distributions of **b** axes for a) native and b) ACA-soaked monoclinic K4 crystals. X-rays at 2000 Watts power (50kV,40mA).

of the fourth, eighth, and fourteenth order reflections suggest that in this case ligand-binding was successful. However, all ACA-soaked monoclinic crystals observed were multiply twinned and unacceptable for diffraction experiments.

Since soaking experiments had failed to produce K4-ACA crystals suitable for data collection, attempts were made to crystallize K4 in the presence of the ligand. Due to the high binding constant of the K4-ACA complex, it was assumed that complexation would be immediate upon mixing of the two species. The ligand was dissolved in a 10 mg/ml K4 solution to a concentration of 25 mM, and this protein solution was then used in the typical way in setting up sitting drop experiments. A variety of conditions known to be successful for native K4 were applied, with best results from a well solution of 29% PEG 8000, 0.12 M ammonium sulfate, 1.4% DMF. Most crystals observed were plate-like in habit, but often growing in rosettes. Precession photographs of one of these plates revealed it to be another monoclinic form having lattice parameters a= 40.9(1) Å, b=32.3(1) Å c=25.1(1) Å, β=90°, with two molecules per unit cell.

IV. DATA COLLECTION

An orthorhombic K4 crystal having dimensions 1.60 x 0.28 x 0.20 mm was chosen for intensity data collection. The crystal was mounted in a siliconized glass capillary (1 mm diameter) with the **a** axis parallel to the length of the capillary (Figure 12). A small drop of mother liquor was placed at one end of the capillary, which was then sealed with a plug of silicon grease and a drop of epoxy cement. It has been found that use of a siliconized capillary and the placing of mother liquor at one, rather than at both ends, are preferable techniques for crystals grown with PEG. These measures appear to discourage the tendency of the mother liquor to travel along the inner capillary wall and collect in a droplet around the crystal. The initial axial intensity distributions observed for this crystal are shown in Figure 13. Comparison with Figures 8 and 9 show the intensities to clearly resemble the Type I pattern, as is typical for newly mounted crystals.

Three-dimensional intensity data were collected to $2\theta=52°$ (1.7 Å) on a Nicolet P3/F four-circle diffractometer using a graphite monochromater, with CuK_α radiation (1.5418 Å) from a rotating anode X-ray tube operating at 3000 Watts (50 kV,60 mA). Figure 14 illustrates the goniostat of the four-circle diffractometer, with which the omega, phi and chi circles are used to orient the crystal so that the Bragg condition is satisfied for a particular set of planes. The detector, a scintillation counter, is oriented using the two-theta circle.

Reflections were measured using a wandering Wyckoff omega-step scan [49] in which the detector is held stationary and the crystal is rotated in increments via the omega-circle. Each intensity measurement was a scan of seven steps across the top of the peak (approximately between points of peak half height), with the sum of the five greatest counts taken as the integrated intensity.

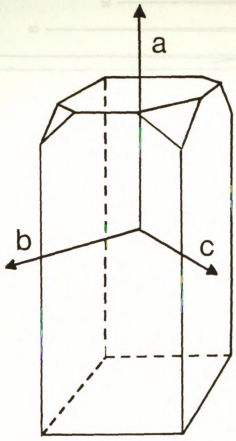


Figure 12. Schematic diagram of orthorhombic K4 crystal morphology with respect to crystallographic axes.

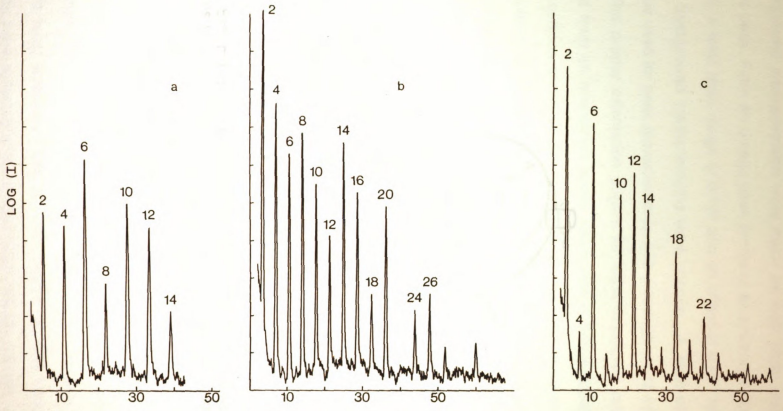


Figure 13. Intensity distributions of **a**, **b**, and **c** axes for orthorhombic K4 crystal prior to intensity data collection. All axes measured using X-rays at 2000 Watts (50kV,40mA).

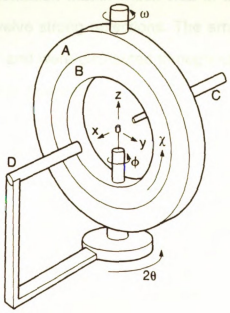


Figure 14. Schematic diagram of four-circle diffractometer goniostat. The ω, φ, and χ circles are used to orient the crystal. The detector is supported by the 2θ circle.

A maximum of three extra steps was allowed, if needed, to cross the top of the peak. Background measurements were made on both high- and low-omega sides of the peak. From an omega peak profile of an arbitrary reflection (1 9 1) (Figure 15), a total scan range of 0.18° and an offset of 0.40° from the peak center for background measurements were chosen. A scan speed of 0.40 deg/min was used for peak measurements, with a background to scan time ratio of 0.20. The proper orientation angles for each measured reflection were determined from an orientation matrix which was in turn calculated from the centered positions of twelve strong reflections. The array of reflections chosen had 2θ values of 39-43° and were distributed throughout the phi and chi ranges defining the data collection quadrant.

Before beginning data collection, the absorption properties resulting from the crystal morphology were determined by measuring the intensity of the (2,0,0) reflection as the crystal was rotated about the phi axis in 10° increments. With the a axis mounted coincident to the phi axis, this set of planes remains in the Bragg condition regardless of the phi angle; however, the intensity of the diffracted beam is affected as the path length of the X-rays through the crystal changes. The resulting absorption correction curve is shown in Figure 16. The purpose of this phi-scan is two-fold; to aid in selecting a quadrant for data collection, and for later use in correcting intensities for crystal absorption.

Due to the symmetry of the crystal system, only one eighth of the total reflection sphere was unique and thus only an octant was measured (reflections having positive indices along each axis). Based on the observed absorption profile (Figure 16), the quadrant between phi angles 132° and 222°, which showed the least absorption effects, was chosen for data collection (chi = 0-90°). Data were collected in the ranges 2-32°, 32-41°, 41-48°, 48-52°, each 2θ range taking approximately one day to complete. Before beginning

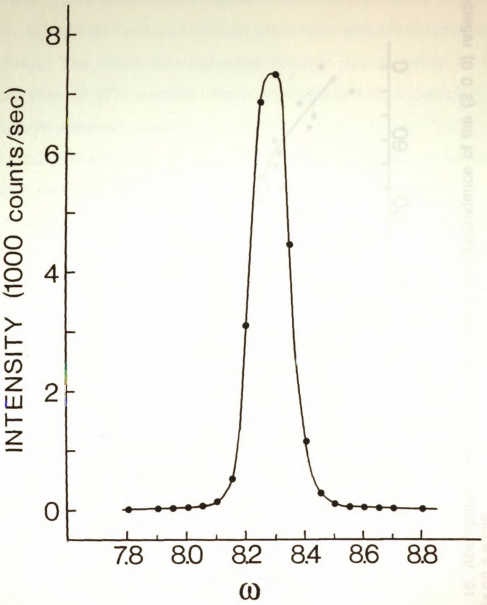


Figure 15. Intensity peak profile of reflection (1 9 1) versus omega.

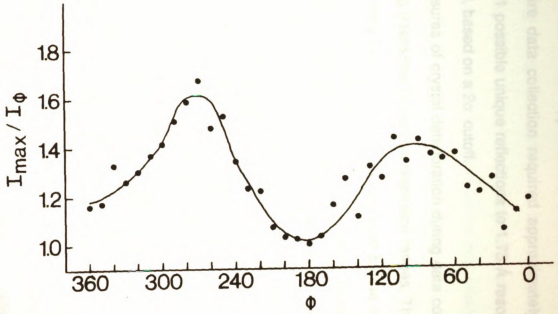


Figure 16. Absorption correction curve showing the dependence of the (2 0 0) reflection intensity on ϕ angle.

collection of each new resolution range, the integrity of the Type I diffraction pattern was checked by measuring the intensities of the reflections (6,0,0), (8,0,0), (0,0,14), and (0,16,0), which were known to differ markedly between the Type I and Type II patterns. The results of these measurements are shown in Table 4. In addition, the **a** axis intensity distribution was remeasured between each range. The entire data collection required approximately 92 hrs of exposure time. Of 8771 possible unique reflections to 1.75 Å resolution, 7647, or 87%, were observed, based on a 2σ cutoff.

A number of measures of crystal deterioration during data collection were monitored, representing intensities in various resolution ranges. These included a two-dimensional 0*kl* data set from 2° to 15°, collected before and following the data collection, and a short set of selected reflections having 2θ values between 45° and 51°, which was measured periodically during data collection. In addition, several monitor reflections were measured throughout the data collection after every 100 reflections. During collection of the 2-32° range, the three reflections (0,14,0), (0,15,9), and (12,5,1), having 2θ values of 25°, 32°,

Table 4. Intensity measurements of conformation-sensitive reflections taken prior to three-dimensional data collection and between each resolution range of data collection.

Time (hrs)	Intensity (counts/sec) 6,0,0	8,0,0	0,0,14	0,16,0
0	3867	246	810	1034
23	4547	220	652	845
46	4390	222	664	817
71	4222	235	640	704
92	5112	169	388	554
Average I, 0-71 hrs	4257	231	692	850

and 35° respectively, were used as monitors. For successive ranges, the first two of these monitors were substituted for reflections having higher Bragg angles: (0,21,8) and (0,15,18) at 41° and 43° respectively. The intensity decay observed for these reflections during the data collection is shown in Figure 17. Another purpose of these monitors was to track any physical movement of the crystal within the capillary during the experiment. In the event that the intensity of any monitor fell below 80% of the initial measured intensity, the centering array reflections were automatically recentered and a new orientation matrix calculated.

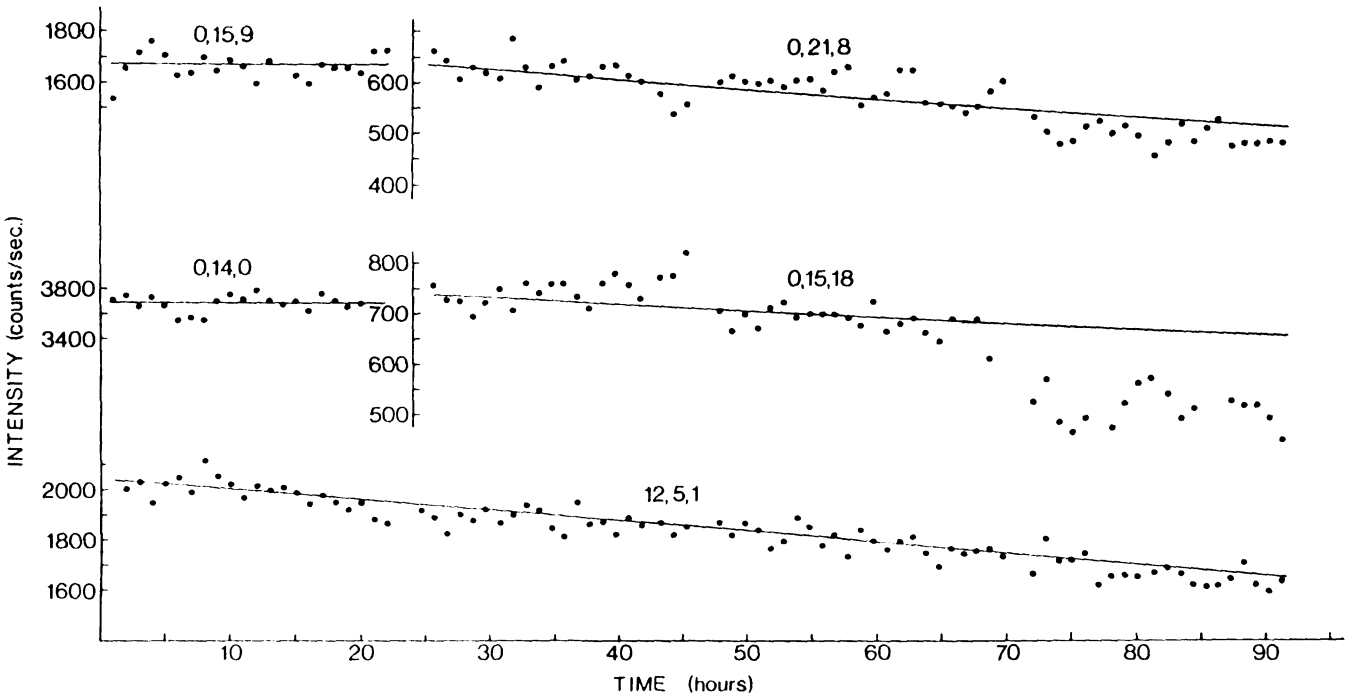


Figure 17. Decay curves for monitor reflections throughout intensity data collection. Resolution range 2-32° began at 0 hrs, 32-41° began at 23 hrs, 41-48° began at 46 hrs, 48-53° began at 71 hrs.

V. DATA REDUCTION

The unusually large intensity changes observed in the final measurement of the conformation test reflections (Table 4), as well as the similar abrupt drop in the intensities of the (0,21,8) and (0,15,18) monitor reflections in the final stages of the data collection (Figure 17), suggested that a conformational change may have occurred during measurement of the 48-53° range of reflections. A comparison of pre- and post-data collection axial intensity distributions supported this conclusion (Figures 13,18). Therefore, it was decided that only data at 1.9 Å (48°) could reliably be assumed to represent the original Type I diffraction, and thus, the higher resolution data were not used in the K4 structure solution.

Before processing the measured data, background measurements were averaged in 2θ shells, with approximately 200-250 reflections per shell. Figure 19 shows the 2θ-dependence of the average background measurements. Similar plots of average background measurements versus phi angle (not shown) indicated no discernible phi dependence. Background-corrected intensities were then calculated as

$$I = (SCAN - (LB + RB)/RATIO) \times SCAN\ RATE$$

where LB and RB designate left and right background measurements and RATIO is the background to scan time ratio. Intensities were considered to be observed according to a cutoff of $2*\langle I_{neg}\rangle$, where $\langle I_{neg}\rangle$ is the magnitude of the average of negative intensity measurements. Based on a plot of $\langle I_{neg}\rangle$ dependence on 2θ (Figure 20), a minimum acceptable intensity of 13 counts/sec was chosen for reflections with 2θ less than 32°, while a value of 4

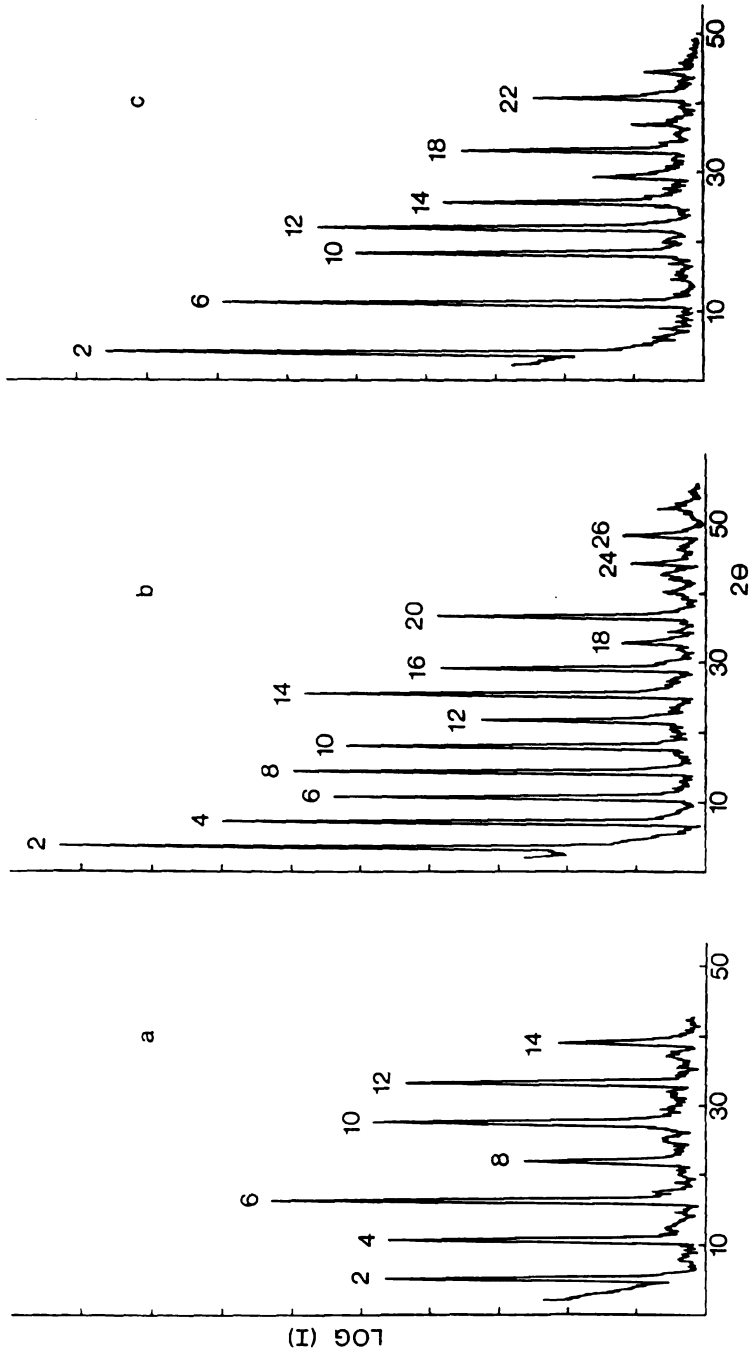


Figure 18. Intensity distributions of **a**, **b**, and **c** axes for orthorhombic K4 crystal following intensity data collection. All axes measured using X-rays at 2000 Watts (50kV,40mA).

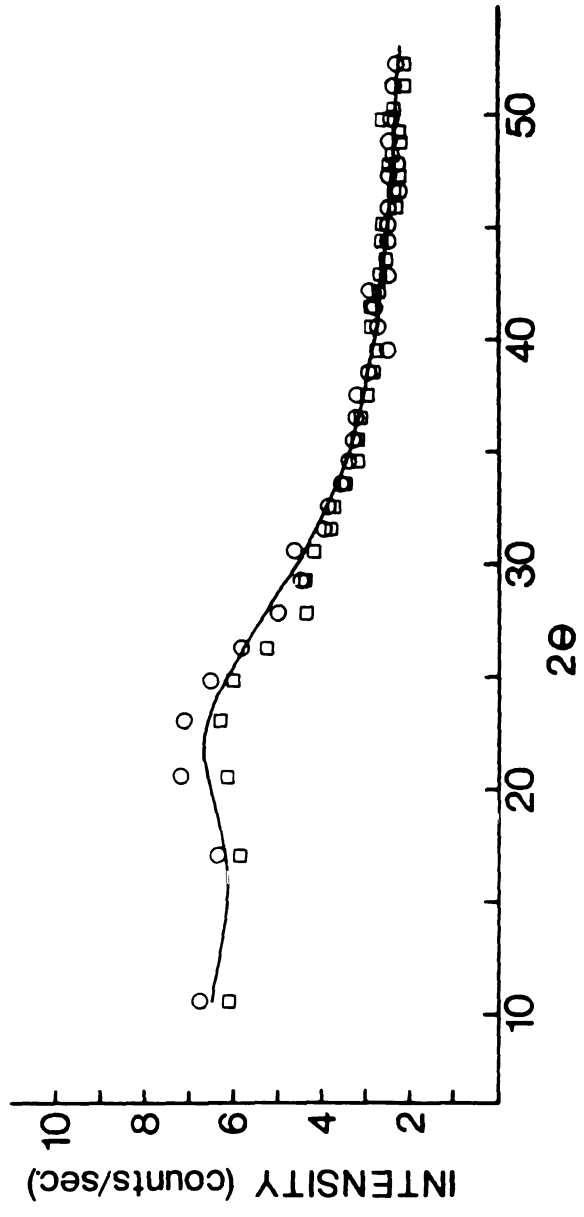


Figure 19. Two-theta dependence of left (squares) and right (circles) background measurements during intensity data collection.

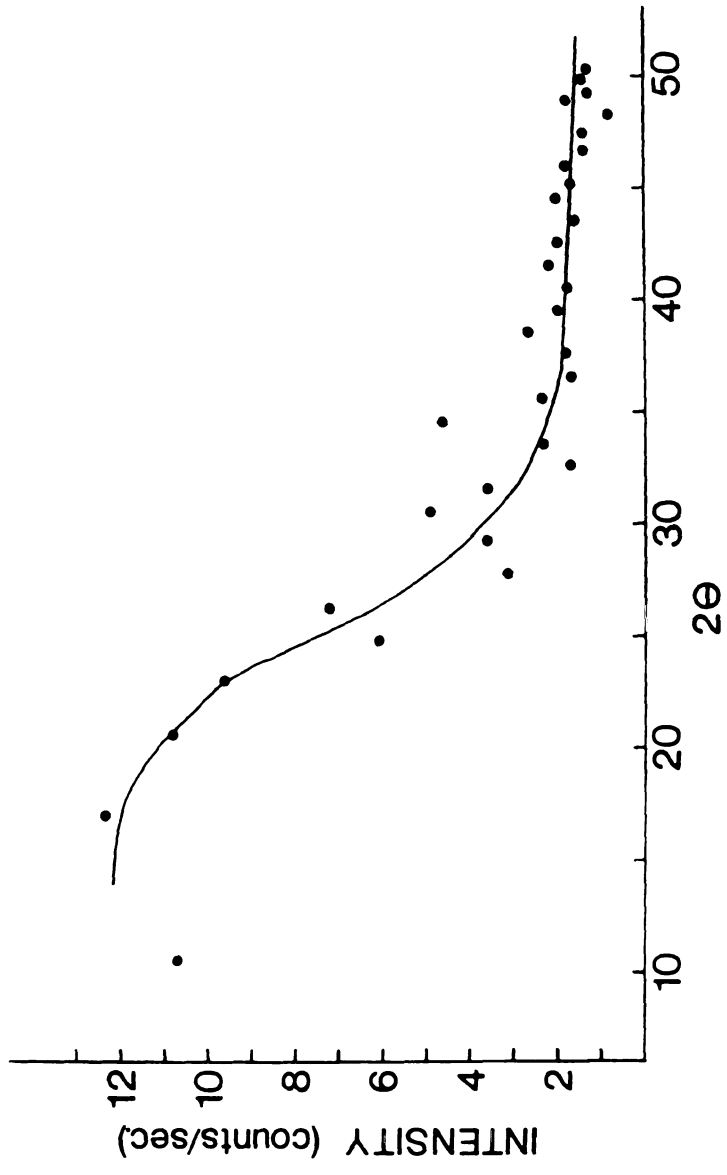


Figure 20. Two-theta dependence of averge negative intensity measurements during data collection.

count/sec was chosen for reflections beyond 32°. A total of 4869, or 75%, of possible reflections to 1.9 Å resolution were accepted as observed.

The structure factor modulus, F(*hkl*), for each reflection was calculated using PDATA [50], a program written in this laboratory, according to the relation:

$$F(hkl)^2 = I(hkl) \times LP \times ABS \times DEC$$

where I(*hkl*) is the background corrected intensity, LP is a Lorentz-polarization factor depending on 2θ and monochromater characteristics, ABS is an absorption factor which corrects for the differing path lengths of diffracted X-rays through the crystal, and DEC is a correction for intensity decay as a function of X-ray exposure time. An empirical absorption correction was used, as suggested by North *et al.* [51], derived from the observed phi-dependence of reflections at chi 90°. The maximum absorption correction applied to the data was 1.27.

The behavior of the monitor reflections, the 0*kl* data set, and the high angle *hkl* data set, as well as the measurements of the array reflections during successive crystal recenterings, were all considered in estimating X-ray deterioration of the crystal. For the latter three, the sums of the individual measured reflection intensities were compared as a function of exposure time. A linear least-squares fit was used to obtain the decay slope for each. The decay factor was then calculated as:

$$DEC = [1/(1 - t * S)],$$

where t is exposure time and S is the rate of decay,

$$S = (K-1)/(K*t)$$
$$K = I(_i)/I(_f).$$

Based on a comparison of the various decay rates (Table 5), a decay slope of 0.0009 hr^{-1} was used for data from 2-32°, while a slope of 0.0017 hr^{-1} was used for data beyond 32°. The maximum decay correction required in 70 hrs of exposure during data collections to 1.9 Å was 1.14. The distribution of structure factor magnitudes observed for various resolution ranges is given in Table 6.

Table 5. Decay slopes determined based on various intensity measurements. Slope S calculated as: $(K-1)/(K*t)$, $K=I_i / I_f$.

Reflections	2θ	S
Monitors:		
0,14,0	25°	0
0,15,9	32°	0
12,5,1	35°	0.00224
0,21,8	41°	0.00224
0,15,18	43°	0.00270
Decay file	45-51°	0.00164
Array reflections	40-49°	0.00162
0*kl* set	2-15°	0.00006

Table 6. Distribution of structure factor magnitudes (F) in various resolution shells for K4 three-dimensional data collection.

Res. (Å)	2θ max	No. of Reflns	Number of F's with magnitude: F<2σ	<5σ	<10σ	<20σ	<30σ	<40σ	<50σ
3.5	25.5	1133	49	4	26	45	31	49	929
3.0	29.8	750	31	15	21	54	30	66	533
2.8	32.0	383	19	16	21	42	13	44	228
2.6	34.5	505	25	11	32	75	62	55	245
2.5	35.9	314	22	12	24	50	62	25	119
2.4	37.5	372	25	18	32	57	74	27	139
2.3	39.2	436	35	30	34	94	92	32	119
2.2	41.0	517	34	44	49	105	127	64	94
2.1	43.1	622	62	73	89	144	120	91	43
2.0	45.3	743	90	105	141	189	131	66	21
1.9	47.9	890	136	169	218	236	91	29	11
1.8	50.7	1139	272	272	304	235	43	11	2
1.75	53.9	967	324	286	250	96	10	1	0
	Total:	8771	1124	1055	1241	1422	886	560	2483

VI. MOLECULAR REPLACEMENT

Since the crystallographic structure of the highly homologous prothrombin K1 had been determined previously in this laboratory, the plasminogen K4 structure was solved using the method of molecular replacement [52], in which a highly similar molecule of known three-dimensional structure is used as a model to calculate initial, approximate phases to be used with the experimental data of the unknown structure.This requires that the model structure first be positioned correctly in the unit cell of the unknown. The correct position can be determined by making use of the Patterson function $P(u,v,w)$,

$$P(u,v,w) = (1/V) \Sigma |F|^2 \cos 2\pi (hu+kv+lw)$$

which is phase independent and therefore can be calculated both from the observed intensities and from the model coordinates. This positioning can be resolved into two parts: a rotation search and a translation search.

The rotation search was carried out using the Search routine of Steigemann's PROTEIN package [53], in which a correlation coefficient of the form

$$C = \Sigma (\text{Patt}(x_o)) (\text{Patt}(R \cdot x_c)),$$

was calculated, multiplying the Patterson functions from the observed data and that calculated from the model as it is rotated by a matrix R with respect to the unit cell. The maximum correlation coefficient observed corresponds to the optimum agreement between unknown and model vector sets, and thus to the appropriate model orientation. Rotations were performed with respect to three

Euler angles θ1, θ2, θ3 (Figure 21), where each complete rotation operation is comprised of the following three components: rotation about the z-axis (θ1), followed by rotation about the new x-axis (θ2), and finally rotation about the new z-axis (θ3).

Two independent model structures were used in solving K4. One model employed was the peptide backbone and conserved side chain structure of the highly refined (R-factor of 0.18 at 2.3 Å resolution) prothrombin kringle 1 (PTK1), representing 489 atoms or 75% of the K4 structure [32]. The second model was the unrefined monoclinic K4 structure, which had also just been solved at the time of this work [54]. Twenty-six atoms which were not well-defined in the monoclinic structure were omitted, the remaining model accounting for 96% of all possible atoms. Prior to calculations involving the monoclinic K4 model, the coordinates were transformed to an orientation corresponding to PTF1, to facilitate comparison of the rotation search results. Based on the greater diffracting power of the K4 crystals relative to those of prothrombin fragment 1 (<B> = 40 $Å^2$), a thermal parameter of 20 $Å^2$ was used in calculating the model Patterson functions. The model Patterson functions were calculated at a 1 Å grid spacing, and an oversized triclinic cell of dimensions 75x75x75 Å was used, so that the calculated vectors would represent only intramolecular distances, which depend solely on the rotational orientation of the model, rather than intermolecular distances, which depend also on the translational placement. Both model and unknown K4 Patterson functions were calculated at 12-3.5 Å resolution (837 reflections), the limits chosen so as to omit very low resolution terms which include a large solvent contribution, and high resolution terms which are influenced by fine structural details.

The rotation search itself was calculated using only a limited set of the largest model Patterson peaks having a specified vector length. Calculations

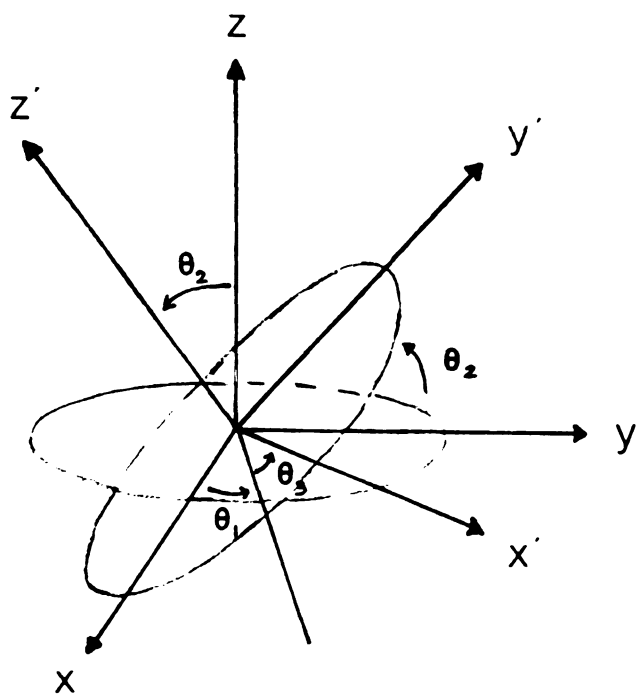


Figure 21. Euler angles which relate rotated axes x', y' and z' to thir original positions x, y, and z.

involving the PTK1 model were carried out using three different vector sets as a verification of the correctness of the rotation solutions. These included 1) the set of 1500 highest vectors having a length of 7-23 Å, 2) the set of 1590 highest vectors having a length of 5-20 Å, and 3) the set of 2000 highest vectors having a length of 5-20 Å. One search was made using the monoclinic K4 model, using the set of 2000 highest peaks with a 5-20 Å vector length. The lower limits on vector length were chosen so as to avoid the Patterson origin, which has an extremely large peak height, but provides no structural information. The upper limits were chosen based on the approximate overall size of the kringle, to exclude vectors having lengths greater than most intramolecular distances. For each of the four model vector sets, an initial search was made in five degree increments over a range of 0-180° through each Euler angle. The orientations resulting in the highest correlation coefficients were then refined using a search in 1 degree increments, ranging 4° on either side of the initial solution. The highest correlations from each rotation search are shown in Table 7. The solution $\theta 1=0$, $\theta 2=75/80$, $\theta 3=40$ appeared consistently among the highest solutions of the PTK1 searches and was the only solution common to both PTK1 and MONOK4 searches.

The second stage of molecular replacement, the translation search, was performed using the program BRUTE [55]. In this procedure, the model structure, which has been correctly orientated in a rotational sense, is translated in increments about the unit cell and at each grid point appropriate symmetry mates are generated and structure factor amplitudes calculated. Agreement with the observed data is determined through calculation of the correlation coefficient:

$$C = \frac{(|F_o|^2 - |F_o|^2) * (|F_c|^2 - |F_c|^2)}{[(|F_o|^2 - |F_o|^2)^2 * (|F_c|^2 - |F_c|^2)^2]^{1/2}}$$

Table 7. Highest cross rotation search results.

Model		θ1	θ2	θ3	Peak Ht.	Ht/σ
PTF1, 7-23 Å	1	0	75	40	113.0	6.1
1500 pks	2	35	135	130	113.0	6.1
	3	25	115	50	103.7	5.6
	4	95	70	40	99.7	5.4
	5	25	40	135	98.6	5.3
	6	155	140	55	98.1	5.3
	7	30	135	60	97.8	5.3
PTF1, 5-20 Å	1	0	75	40	118.9	6.8
1590 pks	2	25	115	50	104.6	6.0
	3	150	140	50	103.7	5.9
	4	35	135	130	102.2	5.8
	5	95	60	135	101.2	5.8
PTF1, 5-20 Å	1	0	75	40	118.0	7.3
2000 pks	2	25	115	50	100.3	6.2
	3	150	140	50	99.0	6.1
	4	30	135	55	96.4	6.0
	5	35	135	130	95.6	5.9
	6	95	60	135	94.4	5.9
MONOK4, 5-20 Å	1	0	80	40	238.6	6.7
2000 pks	2	85	80	100	226.2	6.3
	3	145	125	165	223.5	6.3
	4	65	105	85	218.1	6.1
	5	125	60	55	215.1	6.0
	6	135	95	5	212.5	6.0

from which the F(*000*) term is omitted. BRUTE makes use of a different angular convention than the Euler system which was employed in the rotation search, so the rotation solutions must first be converted to the BRUTE angles α, β, γ by the relations

$$\alpha = \theta 1 - 90°$$
$$\beta = \theta 2$$
$$\gamma = \theta 3 + 90°.$$

To guard against possible oversight of the correct solution, the translation search was performed on the highest three rotation solutions obtained with the monoclinic K4 model, each of which had been refined in 0.3° increments. Since translation search results are often ambiguous, with no single convincing solution, the calculations were repeated at four different resolution ranges to aid in interpreting the results. The ranges used were 5.0-4.0 Å (276 refls), 5.0-2.8 Å (1236 refls), 8.0-4.0 Å (525 refls) and 8.0-2.8 Å (1485 refls). As a result of the $P2_12_12_1$ symmetry, it was only necessary to search an eighth of the unit cell, thus each search was made in 0.5 Å increments over ranges in x, y and z of 16, 25 and 25 Å respectively. The highest solutions observed for each of the three searches are shown in Tables 8-10. Since a single correct translation solution was not obvious in any case, the possible correctness of a number of highest solutions was determined by inspecting the C_α packing resulting from each, using Evans and Sutherland PS390 stereo graphics with FRODO software [56]. Eleven translation solutions from each of rotation solutions I and II, and eight translation solutions for rotation solution III were examined. Due to the tight K4 crystal packing, all but one were easily rejected based on the presence of interpenetrating symmetry mates or unacceptably short contacts (<5-7 Å). Most also showed highly uneven and improbable packing, with very close intermolecular contacts in some regions offset by large solvent cavities. In

Table 8. Translation search results for rotation solution 1: (0,80,40). Highest correlation coefficients (C) are listed.

Vector Set	X	Y	Z	C	C/σ
5-4 Å	2.0	21.0	11.0	0.3574	5.3
mean C=	2.0	11.5	11.5	0.3352	4.9
0.0759	2.0	18.5	12.0	0.3343	4.9
	2.0	16.0	17.5	0.3310	4.9
5-2.8 Å	2.0	18.5	12.0	0.3654	9.9
mean C=	2.0	11.5	19.0	0.3580	9.7
0.2212	2.5	16.5	12.5	0.3551	9.6
	6.0	4.0	19.0	0.3535	9.6
	2.5	16.0	15.0	0.3516	9.6
	2.0	3.5	19.0	0.3508	9.5
8-4 Å	3.5	9.0	10.0	0.2645	5.4
mean C=	14.5	13.5	10.5	0.2363	4.9
0.0449	14.5	9.0	10.0	0.2300	4.7
	10.0	11.5	6.0	0.2237	4.6
	9.0	19.5	3.0	0.2235	4.6
	9.5	13.5	10.5	0.2200	4.5
8-2.8 Å	3.5	9.5	10.0	0.3025	9.8
mean C=	10.0	16.0	18.0	0.2815	9.1
0.1599	10.0	11.5	6.0	0.2789	9.0
	10.5	18.0	15.0	0.2734	8.8

Table 9. Translation search results for rotation solution 2: (85,80,100). Highest correlation coefficients (C) are listed.

Vector Set	X	Y	Z	C	C/σ
5-4 Å	11.5	2.0	12.0	0.2552	4.8
mean C=	6.0	6.0	24.5	0.2530	4.8
0.0297	6.5	6.5	18.5	0.2258	4.3
	6.0	21.0	24.5	0.2235	4.2
5-2.8 Å	7.0	23.5	18.5	0.2969	10.9
mean C=	6.0	6.0	24.5	0.2942	10.8
0.1850	7.0	6.5	18.5	0.2937	10.8
	11.0	2.0	12.0	0.2924	10.8
	13.0	3.0	21.5	0.2809	10.3
	11.0	11.0	7.0	0.2807	10.3
8-4 Å	6.5	21.0	19.5	0.1308	3.6
mean C=	4.0	25.0	21.0	0.1271	3.5
0.0074	7.0	1.5	16.0	0.1242	3.4
	1.5	14.0	22.0	0.1229	3.4
8-2.8 Å	7.0	14.0	19.0	0.2090	8.9
mean C=	7.0	1.5	16.0	0.2068	8.8
0.1193	14.5	21.5	19.5	0.2050	8.8
	6.5	21.0	19.5	0.2039	8.7
	11.0	2.0	12.5	0.2035	8.7
	15.0	19.0	22.0	0.2023	8.6

Table 10. Translation search results for rotation solution 3: (145,125,165). Highest correlation coefficients (C) are listed.

Vector Set	X	Y	Z	C	C/σ
5-4 Å	10.5	4.0	4.5	0.3354	4.8
mean C=	3.5	7.0	9.5	0.3307	4.8
0.0913	15.0	22.5	8.5	0.3288	4.7
	3.0	19.5	10.0	0.3261	4.7
	7.5	15.0	17.0	0.3253	4.7
5-2.8 Å	3.5	11.5	13.0	0.3367	9.6
mean C=	10.5	4.0	4.5	0.3363	9.6
0.2149	3.5	7.0	9.5	0.3348	9.6
	3.5	19.5	13.0	0.3337	9.5
	15.0	4.5	8.0	0.3317	9.5
	15.0	22.5	8.5	0.3304	9.4
8-4 Å	3.5	7.0	9.0	0.1977	4.8
mean C=	15.0	9.5	8.5	0.1920	4.6
0.0081	15.0	24.5	8.5	0.1864	4.5
8-2.8 Å	3.5	7.0	9.0	0.2485	9.1
mean C=	15.0	24.5	8.5	0.2429	8.9
0.1260	15.0	9.5	8.5	0.2394	8.8
	3.5	19.5	9.5	0.2329	8.5

contrast, translation solution (3.5, 9.5, 10.0) of rotation solution I displayed tight, regular packing with no close contacts except involving one peptide tail, which could be easily accounted for by a slight difference in conformation of this unconstrained, flexible region from that of the monoclinic form.

The combined rotation-translation solution was refined using BRUTE. The final solution, with a correlation coefficient of 0.3469, had a rotational orientation of α=267.77, β=76.68, γ=128.67, defined by the matrix:

0.7858	-0.6174	-0.0379
0.1134	0.2040	-0.9724
0.6080	0.7598	0.2303

and a translation vector of (3.5, 9.5, 10.1 Å).

Initial electron density maps were calculated using phases derived from the monoclinic K4 model, transformed according to the above molecular replacement solution. The model structure included the K4 backbone from Cys1 to Ser82 and those side chains conserved from PTF1, accounting for 506 atoms, or 75% of the complete K4 structure. Two F_o density maps were calculated at resolution ranges 8-3.5 Å and 8-2.8 Å, using an average thermalparameter of 20 Å^2. During these and subsequent electron density calculations, extensive use was made of the crystallographic program package PROTEIN [53].

The maps were examined using FRODO graphics, during which 18 of the 35 missing side chains were added, to include 567, or 84% of all possible atoms. Based on a Wilson plot from this new expanded model, the average B was corrected to 18.5 Å^2. This model, having an R-factor of 0.45, was used for the further calculation of a number of density maps: $(2|F_o| - |F_c|)$ maps at resolution ranges of 8-2.8 Å, 8-3.5 Å, 5-2.8 Å, and a 5-2.8 Å $(|F_o| - |F_c|)$ map.

Inspection of these maps yielded five more side chains (25 atoms) to include 87.6% of the structure. Refinement of the structure began at this point.

VII. STRUCTURE REFINEMENT

The K4 structure was refined by the method of restrainedleast squares [57], using the fast Fourier transform program PROFFT [58]. This method makes use of idealized geometric parameters which relate atomic positions, such as bond lengths and bond angles, to reduce the number of unknown variables to be found. The deviations of these geometric parameters from their ideal values and the discrepancy between observed and model-calculated structure factors are reduced simultaneously, both properly weighted. The agreement of the model structure and observed intensities is determined by calculation of an R-factor,

$$R = \Sigma (|F_o| - |F_c|)/ \Sigma |F_o|,$$

which compares observed and calculated structure factors.

The K4 structure was refined in three stages, initially including data at 6.0-2.5 Å, then extending the resolution to 2.2 Å, and finally to 1.9 Å. Periodically, the refinement was interrupted for interactive computer graphics sessions using FRODO software, during which the model was manually adjusted to better fit calculated ($2|F_o|-|F_c|$) and ($|F_o|-|F_c|$) density maps.

During the first stage of refinement, at 6.0-2.5 Å resolution, the model structure was assigned an average isotropic thermal parameter (B) of 18 $Å^2$. The values applied in weighting the geometrical restraints are given in Table 11. In the early stages of the refinement, a protocol of alternating sets of refinement cycles having tight and loose geometric restraints was used, alternately emphasizing the agreement of geometric variables and structure factors. The structure factors were weighted according to a scheme in which σF was maintained at approximately half the average discrepancy between F_{obs} and F_{cal}.

Table 11. Results of PROFFT least-squares refinement at successive stages of resolution.

RMS Deviations:	Target(T/L)	-- Resolution -- 2.5 Å	2.2 Å	1.9 Å
Distances (Å):				
Bond length	.020/.030	.023	.020	.018
Bond angle	.040/.050	.059/.035	.051/.035	.044
Planar 1-4	.060/.070	.070/.055	.069/.055	.047
Planarity				
Dev from plane (Å)	.020	.018	.017	.015
Chirality				
Chiral volume ($Å^3$)	.150	.208	.233	.234
Nonbonded Contacts (Å):				
Single torsion	.550	.241	.217	.178
Multiple torsion	.550	.373	.359	.211
X..Y H-bond	.550	.334	.382	.260
Torsion Angle (deg):				
Planar	3.0	2.5	3.0	2.6
Staggered	15.0	31.0	24.7	16.9
Orthonormal	20.0	28.1	23.4	18.9
Thermal Restraints ($Å^2$):				
Main chain bond	2.0/3.0	2.11/2	3.14/3	2.76/3
angle	3.0/4.0	3.20/3	4.23/4	3.48/4
Side chain bond	3.0/4.0	2.64/3	3.98/4	3.61/4
angle	3.0/4.0	3.09/3	4.57/4	4.39/4
R-Factor (%)		19.4	15.1	14.2
<B> ($Å^2$)		12.5	16.1	18.0
No. of reflections in calculations		2312	3286	4919

The unrefined K4 structure had an R-factor of 40.7%, but with successive refinement cycles, alternating sets having tight and loose restraints, the R-factor was quickly reduced to 33.4%. At this point, individual B's were introduced and the R-factor was further reduced to 30.3%, the average B falling to 12 $Å^2$. Approximately 300 weak reflections having unusually large individual R-factors (>85%) were removed from the calculations, bringing the overall R-factor down to 27.8%. Additionally, the thermal restraints were relaxed somewhat to 2.0, 3.0, 3.0, 3.0 $Å^2$ (Table 11), whereas restraints on bond angle and 1-4 distances were tightened to 0.035 A and 0.055 Å respectively, to better control the geometry of the structure.

During FRODO graphics interventions, omitted side chains of non-conserved residues were included as suitable density appeared in the electron density maps. In addition, water molecules were added to the structure in the form of oxygen atoms as they became apparent in difference electron density maps. Possible water positions were chosen in the following manner. Difference density ($|F_o|-|F_c|$) maps were calculated at both 8-2.5 Å and 6-2.5 Å resolution ranges. Difference peaks above a 2.5σ threshhold which were common to both maps (within 1 Å of each other) were considered. These peaks were evaluated based on computer graphics examination of the difference maps, as well as a 6-2.5 Å ($2|F_o|-|F_c|$) map. The criteria used for accepting possible solvent positions included that the peak appear in all three maps and be approximately spherical in shape, with the 8-2.5 Å peaks being preferably as large or larger than the corresponding 6-2.5 Å peaks. In addition, it was required that potential water positions be within hydrogen-bonding distance (2.5-3.5 Å) of an appropriate protein atom or another solvent site. New water molecules added to the refinement calculations were assigned an initial B of 18 $Å^2$, a slightly higher value than the average protein B, and an occupancy of 0.75

to sp

relate

sets

refine

updat

solver

cule b

as its

specie

and wa

Tl

ficant e

that tig

contrib

refinem

21.6%.

had an

Da

initially

tightly

weight

θ-depe

In this

relative

to speed convergence. As the thermal parameter and occupancy are closely related variables, they were not refined simultaneously. Rather, several initial sets of cycles refining solvent coordinates and B values were followed by refinement of solvent coordinates and occupancies, returning occasionally to update refinement of the thermal parameters. During the 2.5 Å refinement, one solvent molecule was suspected to be a sulfate ion rather than a water molecule based on its refinement to a very low B value and unit occupancy, as well as its position between a number of positively charged K4 sidechains. This species was accordingly reassigned as the central sulfur atom of a sulfate ion and was refined using a constant occupancy of 1.00.

The inclusion of solvent structure in the model was found to have a significant effect on the R-factor. This observation is consistent with the expectation that tightly bound waters of a relatively small molecule should make a sizable contribution to the diffraction pattern. The addition of solvent, along with further refinement cycles alternating tight and loose restraints, reduced the R-factor to 21.6%. The final structure at 2.5 Å resolution, containing 45 water molecules, had an average B of 12.5 $Å^2$ and an R-factor of 19.4%.

Data at 2.2 Å were then added and similarly refined. The R-factor, which initially increased to nearly 27%, was reduced to 23% using alternating sets of tightly and loosely restrained refinement cycles. At this point the previous weighting of structure factors, using a constant σ, was replaced by a θ-dependent variable weighting scheme in which

$$\sigma = \sigma_a + \sigma_b(\sin \theta/\lambda - 1/6).$$

In this way, not only the overall weighting on structure factors, but also the relative weights of higher and lower resolution terms could be controlled by

adjusting the coefficient, σ_a, and slope, σ_b. Although the variable weighting scheme did not significantly affect the overall R-factor, agreement was improved for the higher resolution terms. Agreement of low resolution terms, which was sacrificed, was subsequently improved with the continued addition of solvent structure.

Approximately 260 low intensity reflections ($F_{obs}<7.5$) having unusually large discrepancies were omitted from the calculations. In addition, the thermal restraints were relaxed by increasing the target values to 3.0, 4.0, 4.0 and 4.0 Å^2 (Table 11). Soon after beginning refinement at 2.2 Å, alternation of tight and loose restraints was found to be no longer helpful in the progress of the refinement. Instead, further improvement was gained primarily through manual adjustment of the structure based on electron density maps, where necessary corrections were beyond the capability of the least-squares program, and also through the continued addition of solvent. At this stage the criteria used in accepting solvent positions were eased somewhat. Peaks occuring in two of the three maps were considered, and greater flexibility was given to the apparent H-bonding distances, allowing for slight readjustments of the protein structure during subsequent refinement of the altered model. Particularly, distances between solvent positions were evaluated bearing in mind the possibility of non-simultaneous partially occupied sites. The final 2.2 Å structure, with 131 solvent molecules, had an R-factor of 15.1% and an average B value of 16 Å^2 for protein atoms.

Addition of 1.9 Å resolution data to the refinement increased the R-factor to 22.6%, but the K4 structure was quickly refined to 19.5% using tight restraints and a variable weighting scheme. Difference density maps distinctly showed the positions of oxygens for the suspected sulfate ion and these atoms were added to the model at this time. With further refinement, the R-factor was

reduced to 17.9%. Although most of the K4 structure corresponded to the electron density maps exceptionally well, Pro30 was changed from the *trans* to *cis* isomer at this point, based on the relatively poor fit and geometric distortions in this region. This substitution further reduced the R-factor and dramatically improved rms deviations of the geometric parameters, allowing the distance restraints to be relaxed to their initial values. During this stage of the refinement, 45 solvent positions, which had refined to very low occupancies and high B values and for which there was poor or no density, were removed from the structure. The final K4 structure at 1.9 Å resolution included 97 solvent molecules and had an R-factor of 14.2%. The R-factor after removal of the solvent molecules was 23% (Table 12), indicating a considerable contribution of the solvent structure to the observed diffraction pattern.

A summary of the least-squares refinement results at each stage of refinement is given in Table 11. During the course of refinement, little or no density was observed for the residues of the tail regions, a-c and 81-87, nor for the side chain atoms of Lys78 and Lys79. These regions of the K4 structure were therefore concluded to be disordered. Furthermore, the density observed for the

Table 12. Comparison of agreement of structure with (+) and without (-) solvent molecules for various resolution ranges (8.0 Å maximum). R-factor is defined by $\Sigma |F_o - F_c| / \Sigma |F_o|$.

Resolution	Shell			Sphere		
(Å)	R+	R-	Diff	R+	R-	Diff
3.5	12.6	13.9	1.3	12.6	13.9	1.3
2.7	13.7	15.1	1.4	13.1	14.5	1.4
2.4	14.7	16.1	1.4	13.4	14.8	1.4
2.2	13.7	15.0	1.3	13.5	14.8	1.3
2.1	16.0	17.7	1.7	13.7	15.0	1.3
2.0	16.6	18.4	1.8	13.9	15.3	1.4
1.9	18.8	20.0	1.2	14.2	15.6	1.4

Cys

disu

was

mea

beha

coorc

that t

result

0.15 /

termir

final a

‡ Ab

side

symb

oxyge

E, Z,

side

Cys1-Cys80 disulfide was much less well defined than that for the other two disulfides, suggesting that it also suffers from some disorder. Finally, no density was found for the side chains of Thr12 and Glu39 beyond CB‡.

The variation in R-factor with scattering angle may be used to estimate the mean coordinate error [59]. As can be seen from Figure 22, the R-factor behavior for K4 agrees well with the theoretical curve corresponding to a coordinate error of 0.15 Å. This method of determining error, however, assumes that the discrepancies between observed and calculated structure factors are a result of positional errors only. The actual coordinate error probably exceeds 0.15 Å for portions of the structure having high thermal parameters, such as the terminal regions of the peptide chain at and near the Cys1-Cys80 disulfide. The final atomic parameters of the K4 structure are listed in Appendix A.

‡ Abbreviation CA denotes main chain carbon atom from which amino acid side chain branches. Side chain atoms designated by common biochemical symbols in which first character indicates atom type (carbon=C, nitrogen=N, oxygen=O) and additional character(s) indicates position on side chain (B,G,D, E, Z,H proceeding from main chain). Numerals distinguish between equivalent side chain positions.

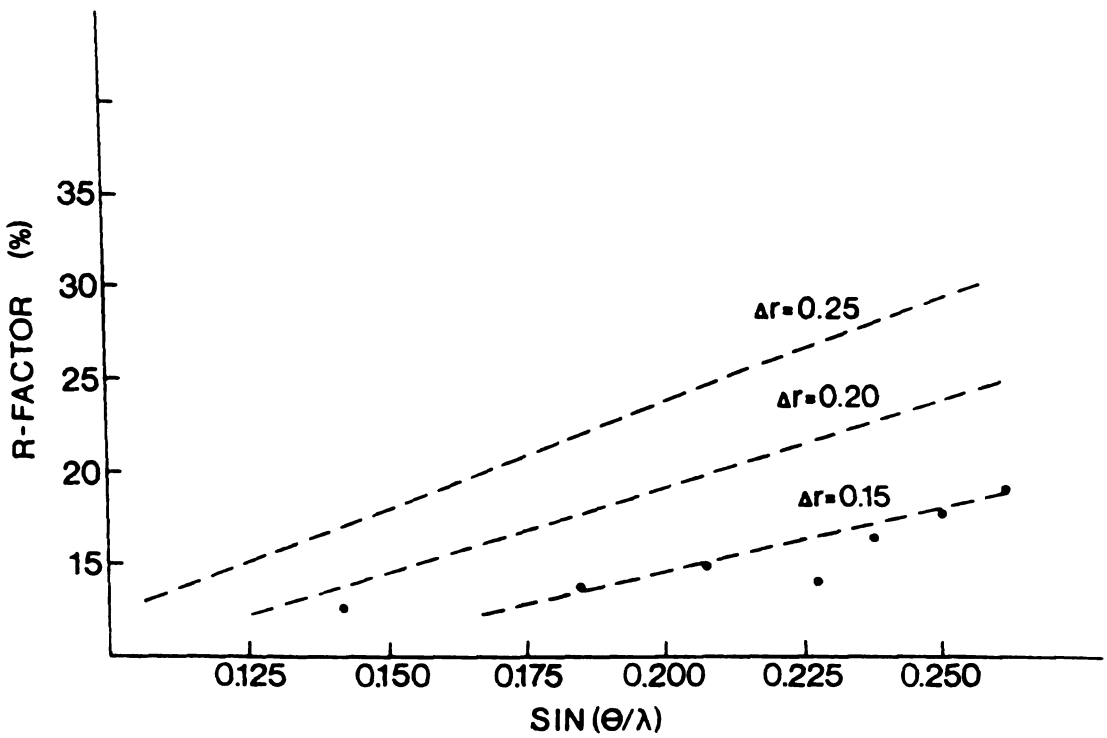


Figure 22. Dependence of R-factor on scattering angle for final refined K4 structure at 1.9 Å resolution. Dashed lines show theoretical coordinate error curves.

VIII. MAIN CHAIN AND DISULFIDE STRUCTURE

The observed main chain structure of K4 agrees well with idealized conformational parameters. A Ramachandran plot of the refined K4 structure is presented in Figure 23, from which it can be seen that the dihedral angles of nearly all the non-Gly amino acids, with the exception of Met48, fall within the energetically preferred zones. Similarly, the distribution of main chain omega angles (Figure 24) shows the observed values for the *trans* residues to be concentrated in a narrow range about 180°. A single *cis* peptide bond is found in the K4 structure, occurring at Pro30 (Figure 25). This bond has an omega angle of 4°, which compares favorably with the ideal value of 0°.

The three-dimensional main chain folding and disulfide structure of K4 are shown in Figure 26. The backbone structure can be subdivided into four segments which result from the characteristic kringle disulfide bridging pattern and which are defined by amino acid stretches Tyr2-Lys21, Gln23-Tyr50, Arg52-Trp62, and Phe64-Lys79. These loops shall be referred to as loops A, B, C, and D, respectively, according to the convention of the prothrombin fragment 1 literature [33]. A striking feature of the K4 structure is the close proximity of the two disulfides Cys22-Cys63 and Cys51-Cys75, which form a four-sulfur cluster and serve as the nucleus of the three-dimensional folding. Due to this arrangement, segments B and C, which are bounded by these two disulfides, are each nearly closed three-dimensional loops. The third disulfide, Cys1-Cys80, is approximately 12 Å away at the edge of the kringle structure, causing loops A and D to have a more extended conformation.

The Pro residues, hydrogen-bonding interactions, and turns (Tables 13,14) which influence the finer details of the K4 folding within each of the four loops are identified in Figure 27. Loop A is a generally S-shaped stretch, with two

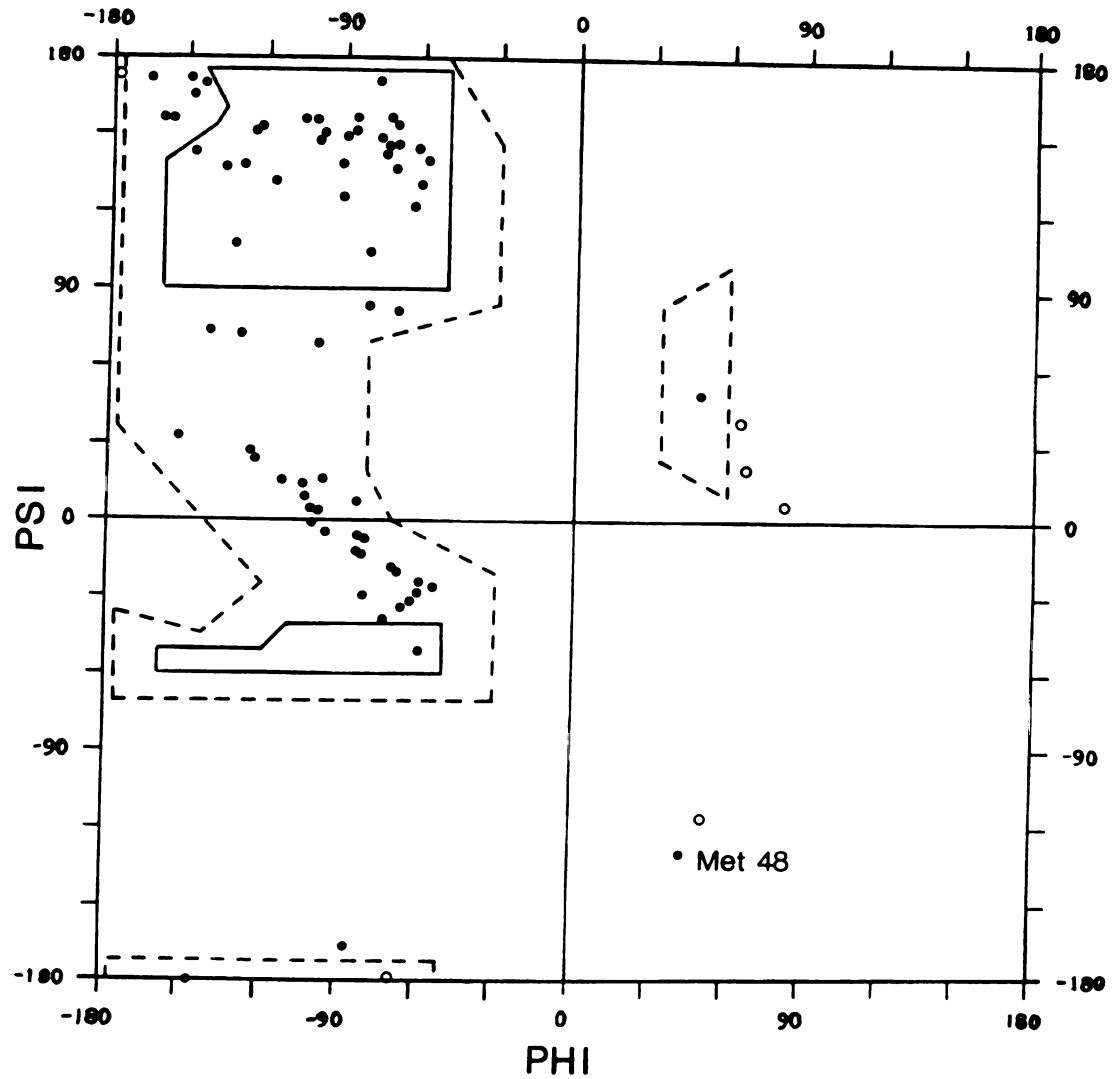


Figure 23. Ramachandran plot of final refined K4 structure. Open circles denote Gly residues; filled circles denote non-Gly residues. Energetically preferred zones are outlined in dashed lines.

NUMBER OF RESIDUES

Fig
(Sin

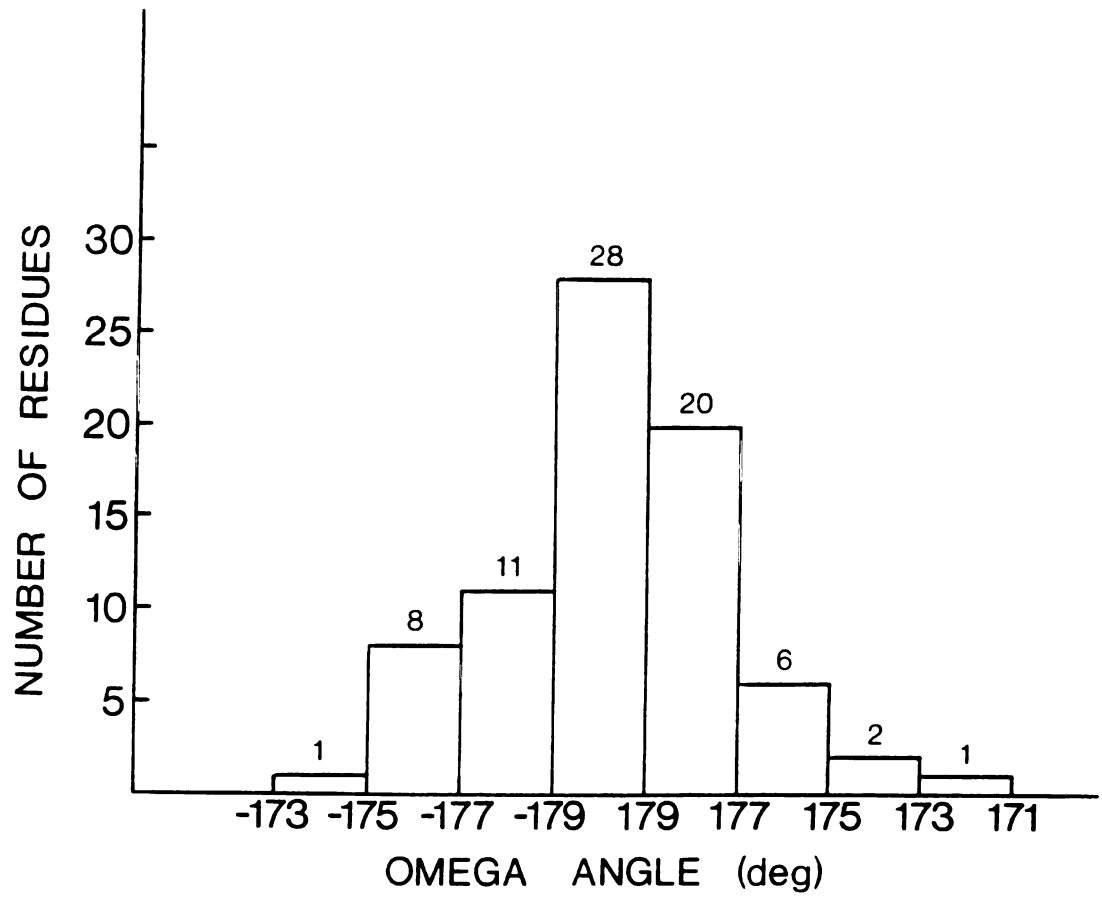


Figure 24. Distribution of omega angles of refined K4 main chain peptide bonds (single *cis* peptide bond not included).

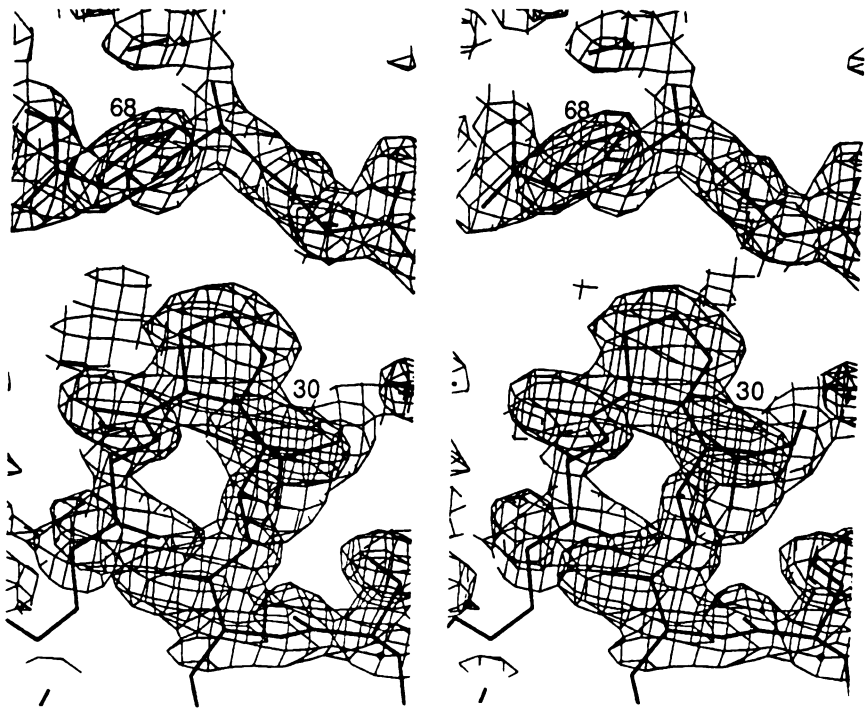


Figure 25. Stereoview of $2|F_o|-|F_c|$ electron density for Pro33 which has a *cis* peptide bond configuration.

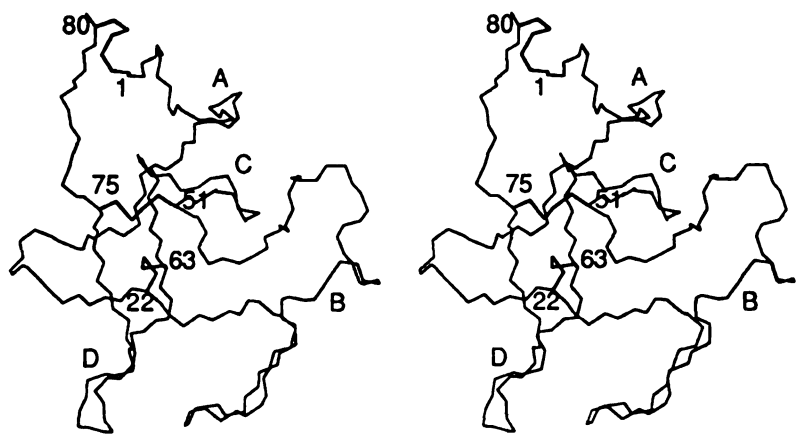


Figure 26. Stereoview of K4 CA,C,N backbone and disulfide (bold) structure.

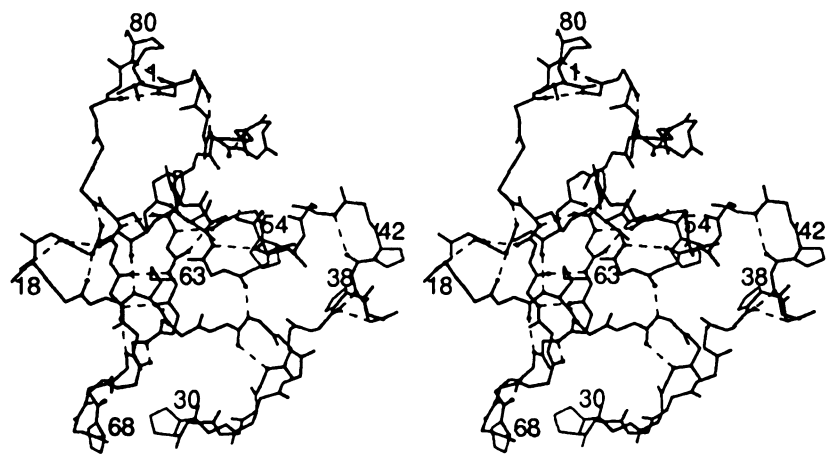


Figure 27. Stereoview of K4 backbone showing Pro side chains and main chain hydrogen-bonding interactions. Hydrogen bonds are indicated with dashed lines.

β-turns at His3-Gly6 and Gly6-Tyr9 combining to form the large upper curve. A sharper hairpin turn at Thr16-Gly19 forms the second curve, which terminates at the central disulfide cluster. Loop B is the longest of the amino acid stretches and its residues comprise nearly half the volume of the kringle. Shortly beyond the disulfide core, residues Ser24-Ser27 form a reverse β-turn, after which the B loop doubles back again in a tight reverse open turn at Pro30, as is commonly observed for *cis* Pro residues [60]. A short extended stretch of main chain separates these turns from a second bending region, in which a number of consecutive bends in the chain are caused by Pro38, Pro42, and a β-type turn involving residues Thr37-Tyr41. The combined effect of these turns results in minimal change in the direction of the main chain; however, an additional β-turn at Tyr41-Ala44 directs the chain back toward the disulfide cluster. One last β-turn of loop B occurs at Thr47-Tyr50. Loop C is the shortest amino acid stretch and is more tightly constrained by the central disulfide pair. This loop has a single wide turn resulting from a β-turn at Asn53-Ala56 and a 90° bend at Gly60 which does not appear to be stabilized by any hydrogen-bonding interactions of the main chain. Two Pro residues occuring in this region have little effect on the path of the main chain: Pro54 results in only a small bend, and Pro61 produces no deviation at all. Finally, loop D has only a single reverse hairpin turn at Pro68, stabilized by a β-type interaction at Asp67-Val70. The remainder of the chain follows a nearly fully extended path, traversing the entire 31 Å length of the kringle to the Cys1-Cys80 disulfide.

In addition to the β-turn interactions, a number of additional main chain hydrogen-bonding interactions contribute to stabilization of the K4 structure. Extensive hydrogen-bonding is found in regions of loops A and D which cross one another near the disulfide cluster. Short stretches of antiparallel β-sheet occur on both loops involving residues Ser14-Thr16/Lys20-Cys22 and

Cys63-Thr65/Arg71-Glu73, respectively. These perpendicular stretches are then stabilized further with respect to each other by hydrogen bonds between Thr15-Asn76, Thr17-Tyr74, and Lys21-Thr66 linking the two loops. A third short stretch of antiparallel β-sheet at Arg52-Cys63 stabilizes the ends of loop C. A single main chain hydrogen bond anchors the ends of loop B near the disulfides at Trp25-Met48. The remaining interactions stabilize backbone stretches adjacent to disulfide bridges, including a hydrogen bond at Gln23-Phe64 near the Cys22-Cys63 disulfide and two involving Tyr2-Lys78 and Tyr2-Cys80 near the outer Cys1-Cys80 disulfide. Three solvent-mediated hydrogen-bonding interactions also occur in the K4 crystal structure between residue pairs Gln23-Phe64, Tyr41-Leu46, and Ala44-Asn53, the last of which stabilizes the interface between loops B and C.

It is evident that the disulfide cluster, adjacent stretches of β-sheet, and inter-loop hydrogen bond links act together to form a highly stabilized zone comprised of the C loop and lower portions of the A and D loops. This is significant in that the K4 Lys-binding site is found in the same region, extending from Asp55/Asp57 on the far edge of the C loop to Arg71 near the lower turn of the D loop. Thus these structural elements appear to serve not only to support the overall kringle folding, but also to rigidly maintain the conformation of the crucial binding region.

A summary of the K4 secondary structural elements is given in Table 13, and a complete listing of observed main chain hydrogen-bonding interactions is presented in Table 14. The K4 β-interactions have an average acceptor-to-hydrogen distance of 2.08 Å and average hydrogen-bonding angle of 159°, in good agreement with values observed for other protein structures [61]. Two potential hydrogen bonds having good inter-atomic distances but unusually small bond angles are included in Table 14. The Cys80 N-Tyr2 O interaction,

Table 13. Observed secondary structural elements of K4 main chain. Reverse turns are classified according to dihedral angles of residues 2,3 [63]. Results of nmr structure are given for comparison [37].

	X-RAY		NMR
β Structure:			
β1	Ser14 O - Cys22 N Thr16 N - Lys20 O		
β2	Lys21 O - Thr66 N Gln23 N - Phe64 O		
β3	Arg52 N - Trp62 O Arg52 O - Trp62 N		
β4	Pro61 O - Cys75 N Cys63 N - Glu73 O Cys63 O - Glu73 N Thr65 N - Arg71 O		Trp62 - Tyr74 Phe64 - Trp72
Reverse Turns:			
T1	His3-Gly4-Asp5-Gly6	(Type II')	
T2	Gly6-Gln7-Ser8-Tyr9	(Type I)	
T3	Thr16-Thr17-Thr18-Gly19	(Type I)	Gly19
T4	Ser24-Trp25-Ser26-Ser27	(Type I)	Ser26
T5	Thr37-Pro38-Glu39-Asn40	(Type III)	Pro,Asn40
T6	Tyr41-Pro42-Asn43-Ala44	(Type I)	Pro
T7	Thr47-Met48-Asn49-Tyr50	(Type II')	
T8	Asn53-Pro54-Asp55-Ala56	(Type I)	Pro,Asp55,Ala56
T9	Asp67-Pro68-Ser69-Val70	(Type I)	Thr66,Pro,Ser69
Other Turns*:			
T10	Pro30		Pro
T11	Lys59		Lys60

* Turn T10 is an open, rather than reverse, turn. T11 is an abrupt 90° bend in the main chain which does not meet the criteria for either type of turn [64].

Table 14. Hydrogen-bonding interactions of the K4 main chain. Interactions forming β-sheet (β) or reverse turns (T) are indicated. Hydrogen atoms are assigned geometrically idealized positions.

DONOR	ACCEPTOR		Distances (Å) H..O	N..O	Angles (deg) NHO	COH	CON
Tyr 2 N	Lys 78 O		2.15	2.97	144	140	151
Gly 6 N	His 3 O	T1	1.91	2.85	149	147	149
Tyr 9 N	Gly 6 O	T2	2.13	3.07	162	100	105
Thr 16 N	Lys 20 O	β1	1.99	2.93	160	149	155
Thr 17 N	Tyr 74 O		2.04	3.02	154	155	150
Gly 19 N	Thr 16 O	T3	2.04	3.04	163	118	120
Cys 22 N	Ser 14 O	β1	2.01	2.98	167	149	145
Gln 23 N	Phe 64 O	β2	1.76	2.75	164	130	135
Trp 25 N	Met 48 O		2.30	3.18	150	164	161
Ser 27 N	Ser 24 O	T4	2.26	3.14	148	103	112
Asn 40 N	Thr 37 O	T5	2.14	2.85	124	114	128
Tyr 41 N	Thr 37 O		1.92	2.80	151	156	165
Ala 44 N	Tyr 41 O	T6	2.24	3.25	178	122	123
Tyr 50 N	Thr 47 O	T7	2.39	3.34	161	150	153
Cys 51 N	Ser 13 O		1.90	2.87	163	146	151
Arg 52 N	Trp 62 O	β3	2.14	3.07	160	165	159
Ala 56 N	Asn 53 O	T8	2.60	3.55	160	106	106
Trp 62 N	Arg 52 O	β3	1.86	2.86	161	144	145
Cys 63 N	Glu 73 O	β4	2.26	3.19	156	161	164
Thr 65 N	Arg 71 O	β4	2.11	3.02	153	128	134
Thr 66 N	Lys 21 O	β2	1.99	2.86	151	179	170
Val 70 N	Asp 67 O	T9	2.27	3.25	168	119	122
Glu 73 N	Cys 63 O	β4	2.25	3.17	151	170	172
Cys 75 N	Pro 61 O	β4	2.09	3.06	167	132	130
Asn 76 N	Thr 15 O		1.85	2.86	179	145	145
Cys 80 N	Tyr 2 O		1.99	2.74	128	122	137

MAIN CHAIN INTERACTIONS BRIDGED BY SOLVENT (W)

DONOR	ACCEPTOR	Distances (Å) NH..W	CO..W	NH..O	N..O	Angles (deg) HWO	NHW	COW
Leu 46 N	Tyr 41 O	1.97	2.66	4.09	5.09	124	159	150
Asn 53 N	Ala 44 O	2.12	2.64	4.28	5.14	127	166	127
Phe 64 N	Gln 23 O	2.00	2.90	4.35	4.97	124	156	139

has a hydrogen bond angle of 128°, which may simply reflect the general disorder accorded to the structure in the vicinity of the Cys1-Cys80 disulfide. A similar hydrogen bond angle of 124° is measured for the possible Asn40 N-Thr37 O interaction. In this case, the residues occur at the point of a reverse turn in the main chain (T5), and, although Thr37 carbonyl group actually appears to make a more favorable hydrogen bond with the amide nitrogen of Tyr41 (bond angle 151°), Asn40 is at the proper sequence position for the expected 1-4 β-type interaction. Furthermore, a survey of NH..O hydrogen bonds reported from neutron diffraction studies [62] found there to be a wider distribution of hydrogen bonding angles, with a lower mean value of 132°, when intramolecular interactions are considered separately from intermolecular interactions. Therefore, both potential hydrogen bonds are listed in Table 14 for consideration.

Of the 78 residues of the Cys1-Cys80 kringle peptide chain, 47 of the main chain nitrogen atoms and 64 of the carbonyl oxygen atoms participate in some form of hydrogen-bonding. In each case, 25 are accounted for by direct main chain-main chain interactions. Of the remaining hydrogen-bonds, 7 of the amine groups and 13 of the carbonyl groups interact with protein side chains. The balance consists of solvent interactions. A comparison of the observed K4 secondary structure with that proposed by Atkinson and Williams [37] in a recent nmr study shows that many of the β-turns agree or are immediately adjacent to one another (Table 13). One stretch of antiparallel β-sheet, Trp62-Tyr74 and Phe64-Trp72, was also correctly identified in the nmr structure, although the additional two observed stretches on loops A and C were not reported.

The distribution of average thermal parameters for the main chain as a function of residue number is shown in Figure 28, from which it can be seen that the K4 backbone has a rather low overall average B of approximately 16 $Å^2$.

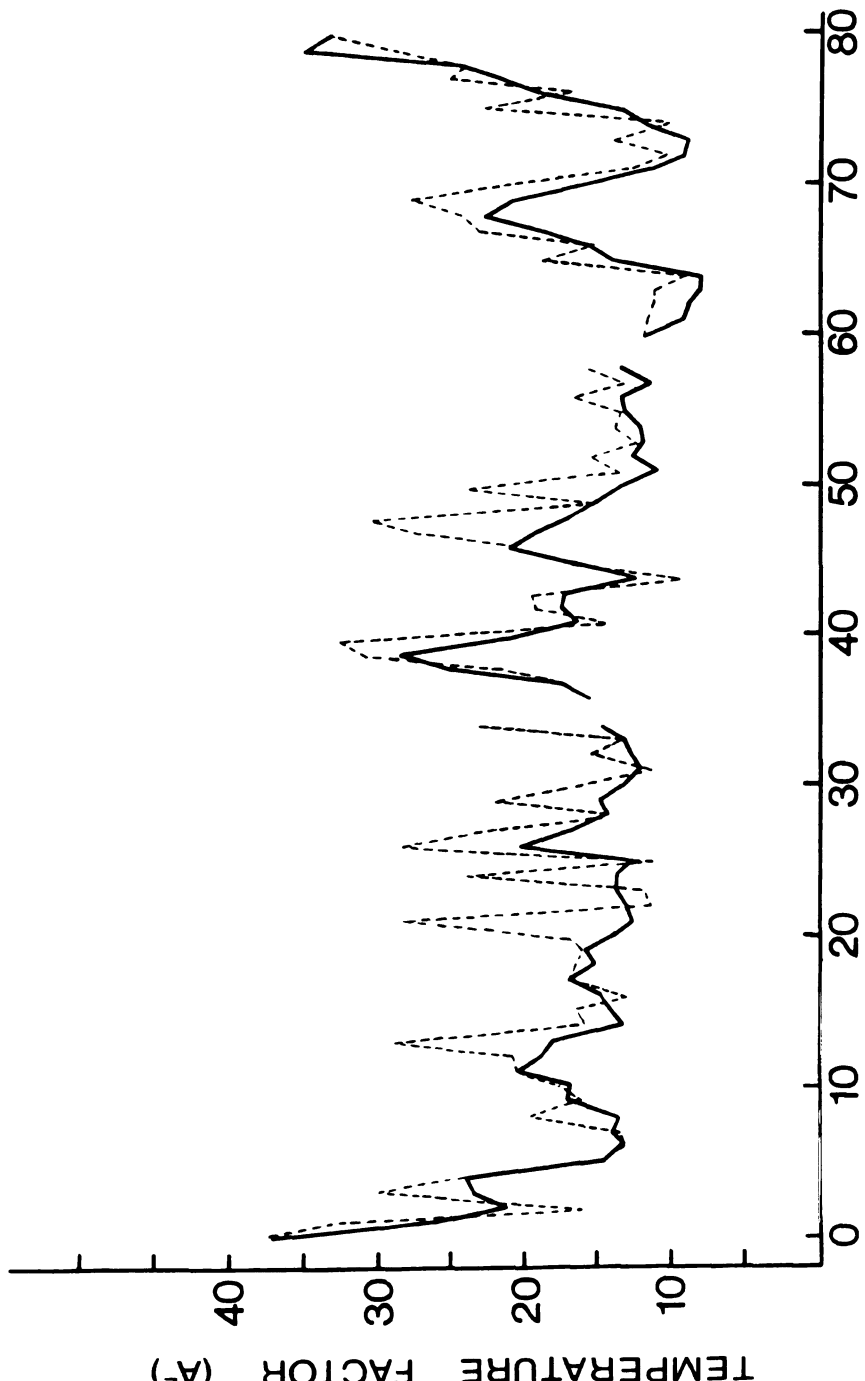


Figure 28. Average thermal factors of K4 main chain (solid line) and side chain (dashed line) structure. Breaks due to deletions with respect to K5 numbering.

The fluctuations in B faithfully reflect the positions of residues with respect to the three-dimensional folding. The thermal parameters are highest in the terminal regions, c,1-5 and 76-80, which have electron density of marginal quality and are considered to be somewhat disordered. Other notable peaks in B value occur near the ends of three-dimensional loops or bends. These are observed at Ser26, Glu39, and Leu46 and all occur in the generally less stabilized B loop. The peak centered at Pro68 corresponds to the long hairpin turn of loop D. Conversely, residues adjacent to the central disulfide cluster and residues participating in β-sheet interactions all have B values which are among the lowest observed.

The three disulfide bonds, which not only define the characteristic kringle primary loop structure, but also have a fundamental role in the three-dimensional folding, are worthy of some additional discussion. The observed bond distances and bond angles for the three disulfides are presented in Table 15. Interestingly, Cys75 was found to have two alternate side chain positions which differ by a torsional rotation of approximately 90° (Figure 29). Through least-squares refinement of the K4 structure with both Cys75 side chain positions, the difference in orientation was found to have no noticeable effect on the conformation of the adjacent backbone structure or of the Cys51 side chain. The two Cys75 side chain positions result in two alternate Cys51-Cys75 disulfide bonds having perpendicular orientations (Figure 30). In one case, the Cys51 and Cys75 side chains form a left-handed disulfide bond (S-S dihedral angle ~ -90°) which is parallel to and nearly aligned with the neighboring Cys22-Cys63 bridge. This orientation positions all four sulfur atoms of the cluster within approximately 4.5 Å of one another in a box-like arrangement, with a distance of only 3.8 Å between the midpoints of the two disulfide bonds. In the second case, the Cys75 sulfur atom is rotated away from the

Table 15. Geometric parameters of K4 disulfide bonds.

Torsion angles are defined as:

CA---CB---SG---SG---CB---CA
X1 X2 X3 X2' X1'

	C-S-S Distance (Å)		Cys Side Chain Angles (deg)		Tors Angle (deg)				
Disulfide	S1-S2	CA-CA	C1-S-S	C2-S-S	X1	X2	X3	X2'	X1'
Cys1-Cys80	2.07	4.81	104	106	56	92	82	-40	-66
Cys22-Cys63	1.98	5.55	105	100	-70	-64	-73	-167	64
Cys51-Cys75a	1.97	6.14	99	100	-67	-94	-82	-136	-86
Cys51-Cys75b	2.10	6.09	101	94	-66	-175	86	103	-172

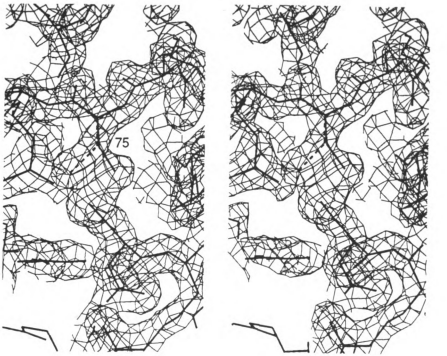


Figure 29. Stereoview showing two alternate Cys75 side chain conformations.

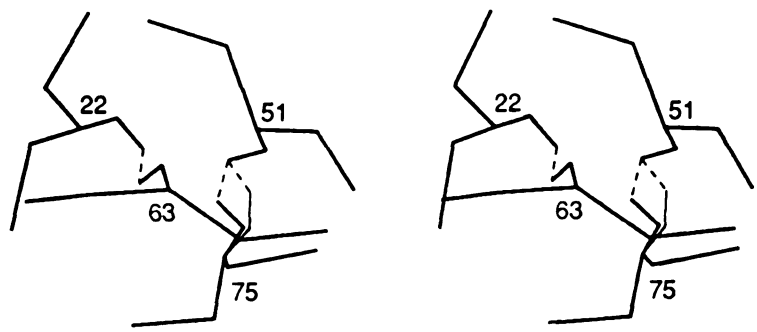


Figure 30. Stereoview showing two relative disulfide positions resulting from two Cys75 side chain orientations.

Cys22-Cys63 disulfide, to form a right-handed Cys51-Cys75 bridge (S-S dihedral angle ~ 90°) which is perpendicular to and directed away from its neighbor. The four sulfur atoms now have a T-shaped arrangement, with a distance between disulfide midpoints of 4.9 Å, and with the Cys75 sulfur atom approximately 6 Å from either sulfur of the Cys22-Cys63 bridge.

Based on separate least squares refinements in which one, then the second, side chain position was used, with the other position filled by a free sulfur atom (as solvent), it was estimated that the two alternate Cys75 side chain orientations are each roughly 50% occupied. With the Cys51-Cys75 disulfide in the left-handed orientation, the free sulfur refined to an occupancy of 0.53; with the disulfide in the right-handed orientation, the free sulfur assumed an occupancy of 0.68. In both cases the B factor of the free sulfur refined to the same value of 21 $Å^2$. At the same time, the sulfur assigned as SG of the Cys75 side chain had a fixed occupancy of 1.00, along with all protein atoms, but refined to a lower B value in the left-handed disulfide orientation (28 vs 31 $Å^2$). These observations may suggest some slight preference for an arrangement in which the two disulfides are parallel. Alternatively, since the K4 structure was refined using only the left-handed disulfide arrangement until the final stages, this difference may merely reflect a bias of the refinement procedure itself. The roughly even distribution in occupancy of the two orientations implies that there is no significant energetic advantage to either arrangement of the four sulfur atoms. Nonetheless, it is surprising that such conformational disorder is permissible at the structural core of the tightly constrained kringle fold.

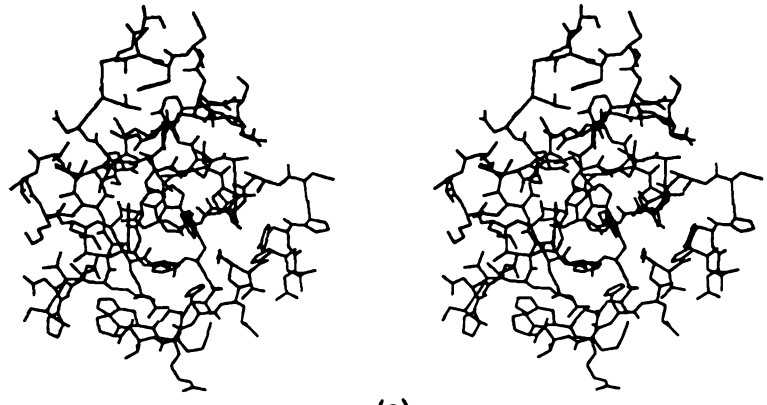
It is probable that the observed Cys75 disorder is static, rather than dynamic. Simple torsional rotation of the Cys75 side chain between the two observed orientations results in the unacceptably close approach of approximately 1.5 Å between the Cys51 and Cys75 sulfur atoms. Thus, a

conformational change would be extremely hindered without considerable adjustment in position by the neighboring atoms. Moreover, the observed thermal parameters of the Cys51 side chain atoms, as well as of the main chain atoms adjacent to the Cys51-Cys75 disulfide, are quite low, suggesting no such unusual movement. Since the residues following Cys75 in the K4 peptide chain become steadily less ordered, ending in a completely disordered carboxy tail, the presence of two Cys75 conformations may be a result of general flexibility of the terminal peptide regions. The greater conformational freedom of the K4 terminal region in solution may allow flexibility of the Cys51-Cys76 disulfide bridge itself, which is then trapped in one of the two observed orientations upon crystallization.

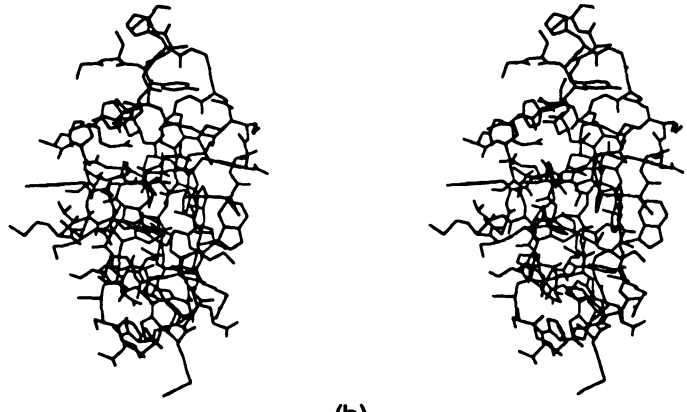
IX. SIDE CHAIN STRUCTURE

With the addition of the amino acid side chains, the overall three-dimensional structure of K4 has the appearance of a disc with one convex surface and one flat surface (Figure 31). Of the 76 residues from Cys1 to Cys80 on the K4 peptide chain, 51 (63%) of the side chains are exposed on the surface of the structure while the remaining 27 (37%) are primarily internal. The solvent-accessible surface areas calculated for each residue [65] are listed in Table 16, in which an accessible area of 10% effectively distinguishes buried from external side chains. Several residues, such as Asn49 and Glu73, which appear to be only marginally exposed based on surface accessibility fractions, are observed in the three-dimensional structure to be indeed external but shielded by hydrogen-bonding interactions with neighboring residues.

About 75% of the external side chains are hydrophilic. Of these, 21 (55%) are involved in hyrdogen-bonding interactions with adjacent protein side chain or main chain atoms. A complete listing of probable K4 side chain interactions is given in Table 17. Nearly all of the charged side chains are among those found on the surface of the kringle. Although there appears to be no overall pattern to the distribution of charge on the surface, a clustering of like charges is found at either end of the lysine-binding region. The side chain of Asp5 is adjacent to the Asp55/Asp57 pair of the binding site, forming a triply anionic cluster at one end, while Arg71 of the binding site, Arg32, and Lys35 form a similar cationic group at the opposite end. A repeated pairing of similar charged residues is also observed elsewhere on the K4 surface. In addition to the sequential pairs of Lys20/Lys21 and Lys78/Lys79, residues Arg10 and Arg52 are spatially adjacent in the three-dimensional structure. These side chains are oriented in an antiparallel manner, with the guanidinium groups of each involved in



(a)



(b)

Figure 31. Stereoview showing a) face-on and b) edge-on views of K4 structure.

Table 16. Accessible surface area of K4 main chain and side chains by residue, calculated using programs ACCESS and ACCFMT [65] with a spherical probe of radius 1.4 Å. Percent accessible surface values are based on calculated accessible surface for the isolated residue. Asterisks indicate residues having disordered side chains.

Residue Number	Type	Main Chain Accessible Surface Å²	%	Side Chain Accessible Surface Å²	%
1	Cys	11.3	8	42.3	38
2	Tyr	1.4	1	34.4	16
3	His	6.5	5	122.6	71
4	Gly	74.6	38	-	-
5	Asp	3.7	3	51.0	41
6	Gly	0.2	0	-	-
7	Gln	6.7	5	71.0	43
8	Ser	14.7	11	66.2	70
9	Tyr	6.4	5	11.4	5
10	Arg	23.9	18	92.6	41
11	Gly	24.8	13	-	-
12	Thr	*	*	*	*
13	Ser	9.9	7	21.8	23
14	Ser	1.7	1	25.0	26
15	Thr	0.0	0	61.2	49
16	Thr	0.5	0	0.0	0
17	Thr	29.8	23	60.7	48
18	Thr	35.3	27	48.8	39
19	Gly	58.6	30	-	-
20	Lys	8.2	6	30.7	24
21	Lys	3.8	3	124.8	63
22	Cys	0.0	0	0.0	0
23	Gln	1.4	1	2.0	1
24	Ser	2.3	2	36.2	38
25	Trp	0.3	0	0.8	0
26	Ser	20.8	15	65.5	69
27	Ser	8.4	6	27.5	29
28	Met	9.3	39	52.6	37
29	Thr	3.8	3	98.0	78
30	Pro	26.9	24	26.9	19
31	His	0.9	1	12.5	7
32	Arg	3.6	3	183.1	80
33	His	12.4	10	0.5	0
34	Gln	2.3	2	131.8	81
35	Lys	0.0	0	57.2	29
37	Thr	0.0	0	24.3	19
38	Pro	4.7	4	50.9	37
39	Glu	*	*	*	*
40	Asn	25.8	20	63.2	45
41	Tyr	4.5	3	59.6	27
42	Pro	28.2	25	75.3	55
43	Asn	16.4	13	117.0	83

Table 16 (cont'd)

44	Ala	9.9	7	3.6	5
45	Gly	20.3	10	-	-
46	Leu	6.3	5	0.0	0
47	Thr	6.1	5	65.9	52
48	Met	16.3	87	116.7	78
49	Asn	0.1	0	15.7	11
50	Tyr	0.0	0	58.5	27
51	Cys	0.0	0	0.0	0
52	Arg	0.1	0	7.1	3
53	Asn	4.4	3	0.0	0
54	Pro	14.1	12	0.0	0
55	Asp	25.1	19	40.1	32
56	Ala	22.1	15	38.1	48
57	Asp	7.4	6	51.0	41
58	Lys	18.2	13	140.9	72
60	Gly	3.4	2	-	-
61	Pro	1.7	2	1.5	1
62	Trp	0.0	0	11.1	5
63	Cys	0.0	0	0.0	0
64	Phe	0.1	0	5.8	3
65	Thr	0.0	0	0.0	0
66	Thr	12.5	10	39.2	31
67	Asp	3.6	3	42.2	34
68	Pro	17.9	82	92.7	67
69	Ser	36.9	27	67.3	71
70	Val	4.3	3	39.7	28
71	Arg	5.9	4	114.0	50
72	Trp	9.4	8	93.0	38
73	Glu	3.9	3	23.1	15
74	Tyr	0.7	1	62.0	29
75	Cys	7.4	5	0.2	0
76	Asn	8.5	7	85.7	61
77	Leu	24.5	19	2.8	2
78	Lys	*	*	*	*
79	Lys	*	*	*	*
80	Cys	64.2	47	13.1	12

Table 17. Hydrogen bonding interactions involving K4 side chains. Hydrogen atoms were assigned geometrically idealized positions. Donor atom is denoted (D), acceptor atom (A).

DONOR			ACCEPTOR			Distances (Å) H..A	D..A	Angles (deg) DHA	CAH	CAD
INTRAMOLECULAR SIDE CHAIN INTERACTIONS										
7	Gln	NE2	5	Asp	OD1	2.07	3.02	157	114	119
10	Arg	NH1	7	Gln	O	2.02	3.02	161	138	136
10	Arg	NH1	9	Tyr	O	2.06	3.01	158	117	116
13	Ser	N	9	Tyr	OH	2.04	2.98	161	149	148
16	Thr	OG	73	Glu	OE1	1.87	2.55	129	172	162
23	Gln	NE2	27	Ser	O	2.19	2.83	119	116	120
23	Gln	NE2	31	His	O	2.04	3.02	161	153	148
25	Trp	NE	50	Tyr	O	1.77	2.71	154	155	158
29	Thr	N	27	Ser	OG	2.15	3.07	165	126	131
31	His	N	23	Gln	OE1	1.98	2.82	149	120	129
31	His	NE	65	Thr	O	2.29	2.92	117	129	145
31	His	NE	67	Asp	O	2.24	2.97	124	130	135
33	His	N	28	Met	SD	2.52	3.49	166	146	142
33	His	NE2	25	Trp	O	2.09	2.83	125	130	134
35	Lys	N	31	His	ND	1.94	2.92	147	-	-
40	Asn	ND2	34	Gln	O	2.14	2.96	136	135	125
43	Asn	ND2	76	Asn	OD1	1.86	2.88	169	133	136
49	Asn	ND2	22	Cys	O	1.79	2.72	161	135	134
52	Arg	NE	45	Gly	O	2.01	3.00	159	130	130
52	Arg	NH2	11	Gly	O	1.89	2.86	160	121	127
52	Arg	NH2	51	Cys	O	2.12	2.92	137	157	147
53	Asn	ND2	5	Asp	O	1.86	2.77	143	145	148
53	Asn	ND2	57	Asp	O	1.89	2.87	167	111	111
57	Asp	N	53	Asn	OD1	1.78	2.68	154	126	117
58	Lys	N	57	Asp	OD1	1.87	2.77	153	89	85
62	Trp	NE1	55	Asp	OD2	1.99	2.79	137	148	136
29	Ser	N	67	Asp	OD2	1.75	2.67	151	152	146
71	Arg	NH2	32	Arg	O	1.99	2.82	137	145	151
74	Tyr	OH	57	Asp	OD1	1.81	2.74	174	142	143
INTERMOLECULAR INTERACTIONS										
32	Arg	NH1	55	Asp	OD1	1.98	2.96	160	156	149
32	Arg	N	57	Asp	OD2	1.74	2.60	153	135	136
43	Asn	ND2	76	Asn	OD1	1.79	2.82	174	140	141
71	Arg	NH1	57	Asp	OD2	2.28	3.14	140	93	94

hydrogen-bonding interactions with main chain carbonyl oxygens, Arg10 with Gln7 and Tyr9, Arg52 with Gly11 and Cys51, thus distancing the similar charges from one another (Figure 32). Additionally, the NE atom of the Arg52 side chain forms a hydrogen bond with the Gly45 carbonyl oxygen.

One ion pair is formed between charged side chains on the K4 surface: the positively charged Lys20 residue extends toward the negatively charged Glu73 side chain, bringing the charge centers within 4.0 Å of one another. The Lys side chain is also within 4.6 Å of a carboxy oxygen of the Asp67 side chain. Among the remaining hydrophilic K4 surface residues, Met28 is particularly interesting as it appears to participate in a hydrogen bond in which the side chain sulfur atom serves as a hydrogen acceptor to the His33 main chain amide group (Table 17). The observed N..S and H..S distances of 3.49 Å and 2.52 Å, respectively, (based on idealized hydrogen positions) agree well with the values of similar NH..S hydrogen bonds which have been reported from crystallographic studies of small organic molecules [66,67].

The hydrophobic amino acids which are exposed to the surface include four Pro residues and a number of residues having relatively small side chains, including Gly, Ala, and Val. The bulky His3 side chain is also highly exposed in the K4 crystal structure, but as this residue occurs in the terminal region, it may have a quite different environment in the intact plasminogen molecule. The aromatic rings of residues Tyr41, Tyr50, and Tyr74 are also somewhat exposed, having accessible fractions near 25% (Table 16). In each case, it is the hydrophilic tyrosyl hydroxyl group which is directed outward toward the solvent. The only remaining significantly exposed hydrophobic surface is the indole ring of Trp72. This large hydrophobic patch on the otherwise hydrophilic surface serves as an obvious marker of the lysine-binding site. Since nearly the entire external surface of the kringle is hydrophilic in nature, it would be expected that,

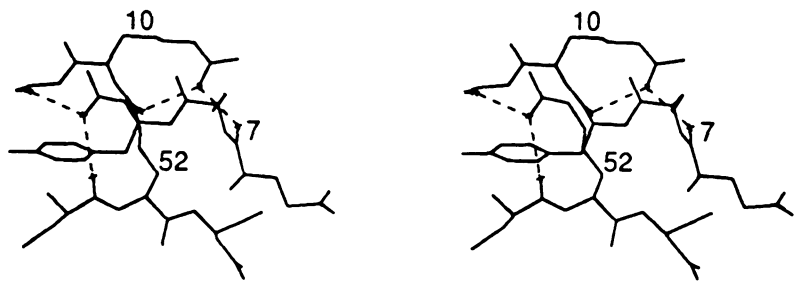


Figure 32. Stereoview showing interactions of adjacent Arg10 and Arg52 side chains. Hydrogen bonds are indicated with dashed lines.

in the absense of specific side chain interactions with other plasminogen domains, the K4 domain would tend to assume a solvent-exposed conformation in the intact protein. This observation agrees well with the previous interpretations of the kringle domain as a structurally and functionally autonomous unit.

The internal residues of K4 include three hydrophilic amino acids: Gln23, Arg 52, and Asn53. The hydrocarbon chain of the Arg52 side chain, although buried, extends toward the kringle surface where the terminal guanidinium group is slightly accessible to the solvent. Two of the internal side chain nitrogen atoms serve as hydrogen-donors to the carbonyl oxygen atoms of Gly11 and Gly45 (Table 17). In the cases of Gln23 and Asn53, the side chains also appear to be fully involved in hydrogen-bonding interactions with the protein backbone, which may add further stability to the three-dimensional folding. The Gln23 side chain carboxy and amino groups bond with main chain His31 N and O atoms, respectively (Figure 33). The Asn53 carboxy group bonds to Asp57 N and the side chain amino group is within hydrogen-bonding distance of both Asp5 O and Asp57 O (Figure 34).

Nearly all the polar groups provided by less hydrophilic internal side chains are also involved in similar interactions. Thr16 and Thr65, which have side chains bearing both hydrophilic and hydrophobic groups, are internal to the K4 structure. The hydroxyl group of the Thr16 side chain acts as a hydrogen donor to the carboxy group of Glu73; however, the side chain of Thr65 appears to have no similar binding partner. The side chain of Tyr9 also bears a buried hydroxyl group which, in this case, appears to act as a hydrogen acceptor to the Ser13 main chain amide group. Lastly, the indole side chains of Trp25 and Trp62 display hydrogen-bonding interactions to the Tyr50 carbonyl and Asp55 carboxy oxygens, respectively. In many cases these side chain interactions occur between residues of separate peptide loops and thus appear to play a

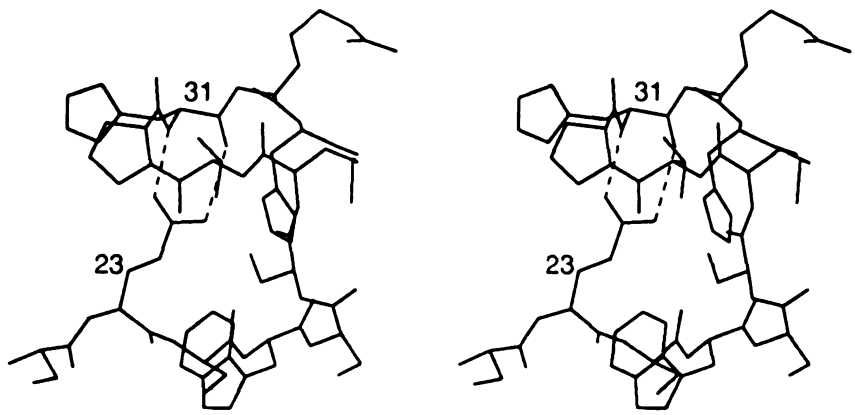


Figure 33. Stereoview showing hydrogen-bonding interactions (dashed lines) of Gln23 side chain with His31 main chain.

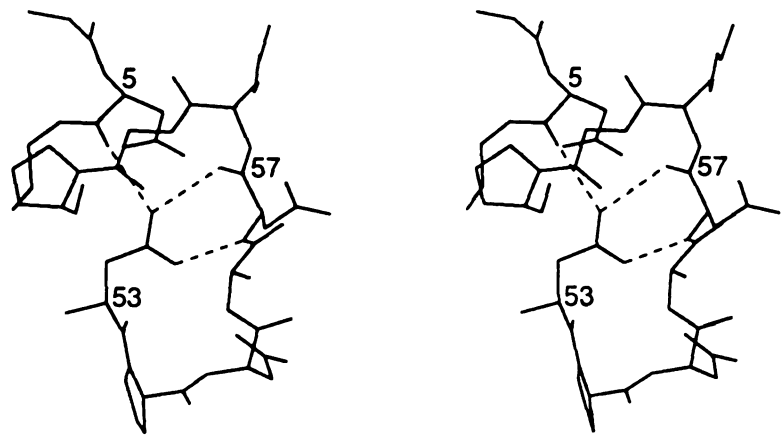


Figure 34. Stereoview showing interactions of Asn53 side chain. Potential hydrogen-bonds to Asp5 O, Asp57 N, and Asp57 O are indicated with dashed lines.

role in stabilizing the protein secondary structure.

A final observation on the hydrogen-bonding interactions of the internal K4 side chains regards the imidazole group of His33. Although there is ample room in the K4 structure to accommodate this side chain, it is positioned so as to make an unusually close contact of 2.92 Å between the ND atom and the main chain amide nitrogen of Lys35 (Figure 35). This suggests that the imidazole ring, which has a pKa of 6.0, is not protonated in this case, allowing for a hydrogen bond between the Lys35 amide hydrogen and the lone pair of the His33 side chain nitrogen. Based on the calculated geometric ideal position of the amide hydrogen, this interaction has a hydrogen-to-acceptor distance of 2.09 Å and a hydrogen-bonding angle of 147°, both of which agree well with typical values. Since the orthorhombic crystals were obtained from a mother liquor at approximately pH 6.0, the His residues might easily exist in either the protonated or de-protonated form; however, the His33 side chain is not accessible to the solvent (Table 16) and probably does not recognize this influence. On the other hand, the His31 side chain is slightly accessible to the solvent. This side chain has no available neighboring hydrogen donors, and although several potential oxygen hydrogen acceptors are found in close proximity to His31, the calculated hydrogen-bonding distances and angles are poor. Thus, in addition to its possible protonation, even the rotational orientation of the imidazole ring about the CB-CG bond (± 180°) is ambiguous in this case.

A dominant feature of the internal structure of K4 is a central hydrophobic core formed by a number of aromatic and other bulky hydrophobic side chains (Figure 36). Near the center of this core is Leu46, surrounded by the side chains of Trp25, Pro38, Tyr41, and Pro54. Through Trp25, the core extends to include Trp62, Phe64, and nearby His33. The aromatic rings of Trp62 and Phe64 are also adjacent to the side chains of Trp72 and Tyr74. Tyr50 is near the

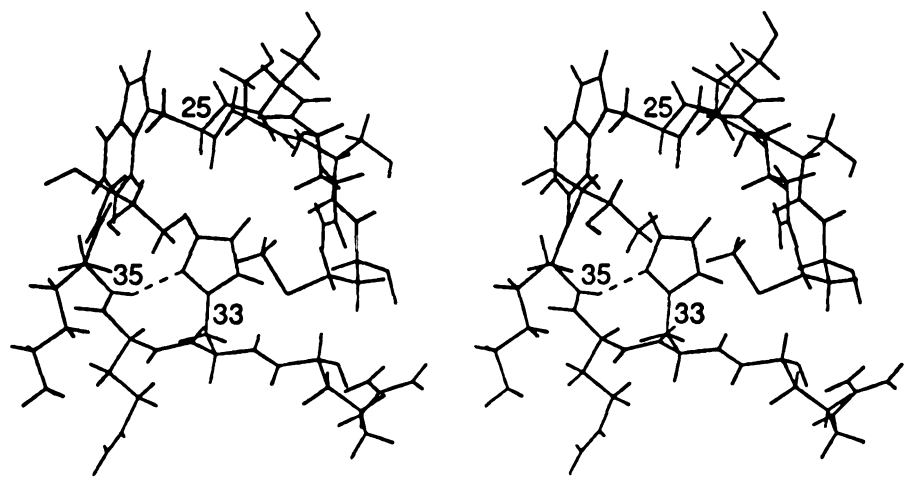


Figure 35. Stereoview of probable hydrogen bond (dashed line) between Lys35 main chain N and lone pair of His33 imidazole N. Hydrogen atoms are shown at geometric ideal positions.

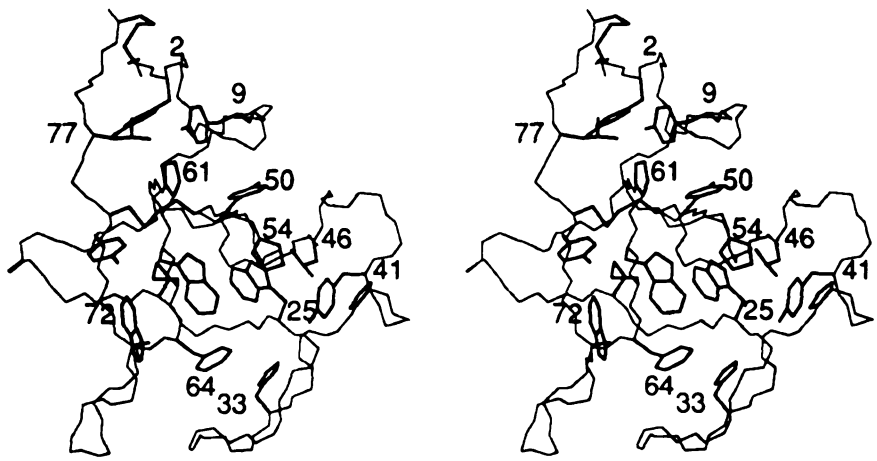


Figure 36. Stereoview of K4 backbone and side chains forming internal hydrophobic core.

hydrophobic cluster, and, although the side chain is directed outward toward the solvent, it may communicate with the core via a hydrogen-bonding interaction between the Tyr50 carbonyl group and the nitrogen atom of the Trp25 indole ring. A similar but much smaller hydrophobic cluster occurs at Leu77, which is surrounded by the side chains of Tyr2, Tyr9 and Pro61.

The specific interactions observed between the residues of the hydrophobic core are interesting in themselves. A number of the aromatic side chains participate in stacking interactions in which one aromatic ring is oriented in a perpendicular manner toward a second (Figure 37). For example, the edge of the Trp25 indole ring is directed toward the face of the six-membered ring of the Trp62 indole at an angle of approximately 94° and at a distance of 3.2 Å from the the plane of the Trp62 ring. Similarly, the Trp62 indole is oriented at an angle of approximately 92° to the face of the Phe64 ring, and the Trp72 indole lies at an angle of 85° to the Tyr74 ring. In both cases, the calculated geometric ideal position of the nearest radial indole hydrogen is 2.8 Å from the midpoint of the corresponding benzene ring. In each case it is the six-membered ring of the Trp indole, rather than the nitrogen-containing five-membered ring, which is directly involved in the interaction.

This type of perpendicular aromatic stacking interaction is commonly found in protein structures and is a result of the planar structure of the aromatic side chains [68]. Since the hydrogen atoms, which are positively charged relative to the ring carbon atoms, are oriented radially in the plane of the ring, the edge of the side chain has a positive charge with respect to the face. Thus the perpendicular stacking arrangement is the most energetically favorable and in fact has a stabilizing effect on the protein structure. The interplanar stacking angles and distances observed in the K4 structure agree well with those reported for other proteins [68]. In several cases, the planar surfaces of these aromatic K4 side

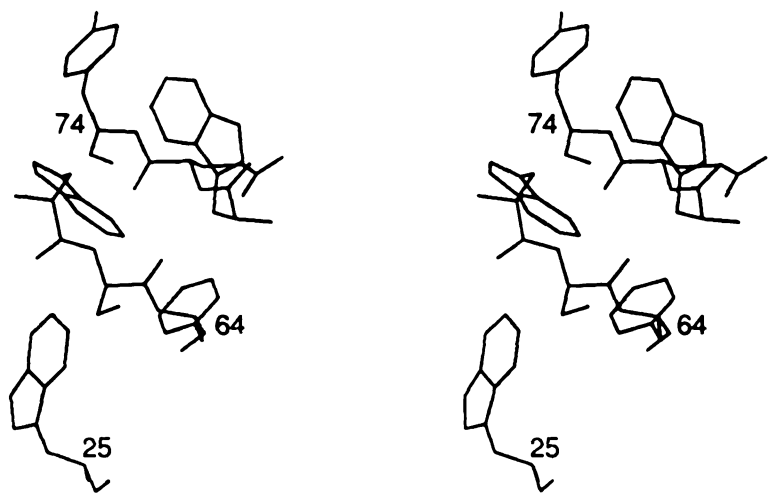
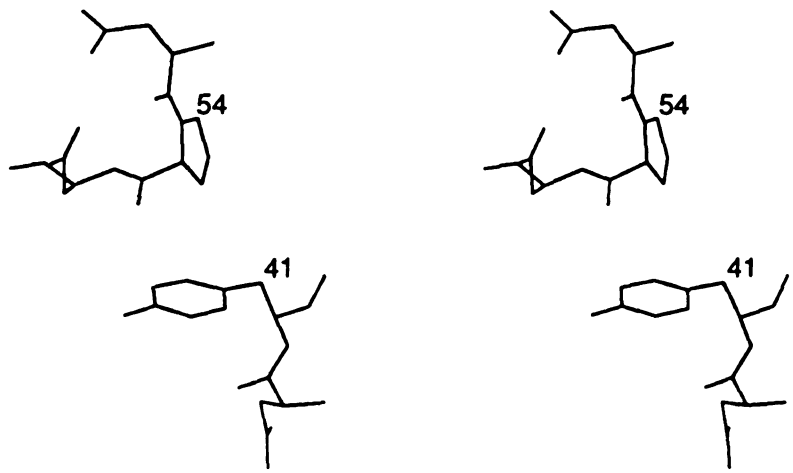


Figure 37. Stereoview of perpendicular aromatic stacking interactions between Trp25, Trp62, Phe64, Trp72 and Tyr74.

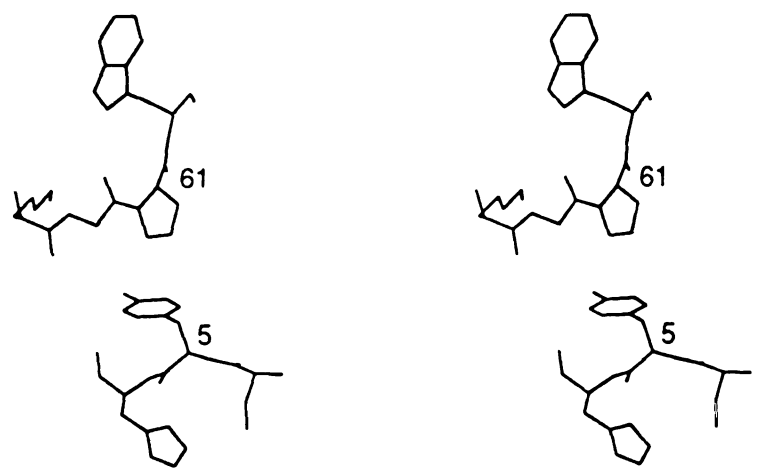
chains appear to interact in a similar manner with other groups having a partial positive charge. A Leu46 methyl group is positioned 2.8 Å from the plane of the Trp25 indole. Likewise, a Leu77 methyl group is 2.7 Å from one face of the Tyr9 ring, while the Gly11 main chain amine hydrogen is directed at a distance of 2.5 Å toward the opposite face.

Perpendicular stacking interactions very similar to those of the aromatic rings but involving Pro residues are also observed in the K4 structure. The edges of the Pro54 and Pro61 side chains are directed at angles of approximately 80° toward the faces of the tyrosyl rings of Tyr41 and Tyr2 respectively, at separations of about 3 Å (Figure 38). In addition, the edge of the Pro30 side chain is oriented similarly about 3.5 Å from the Pro68 ring face (Figure 39). This type of clustering between Pro and aromatic side chains has been observed previously for other protein structures [69]. This behavior of the Pro residues may be rationalized using the same principles which influence the contacts of the aromatic residues since the Pro side chains also have a nearly planar conformation. Although the hydrogen atoms of Pro side chains do not lie in the plane of the ring, they are directed away from the center, giving the side chain a charge distribution which is less pronounced but analogous to that of the aromatic residues.

In contrast, although the two His residues of K4 both have neighboring planar ring side chains, neither displays an edgewise interaction. The imidazole groups of His31 and His33 are adjacent to the Pro30 and Phe64 side chains respectively, and in both cases, the neighboring rings are approximately parallel to one another (Figure 40). In the case of His33, which appears to be de-protonated, the radial lone electron pair of the imidazole nitrogen atom would disrupt the edge/face dipolar charge distribution of the ring which encourages the perpendicular stacking arrangements. Similarly, although the



(a)



(b)

Figure 38. Stereoviews of perpendicular interactions of a) Pro54 with Tyr41 ring and b) Pro61 with Tyr5 ring.

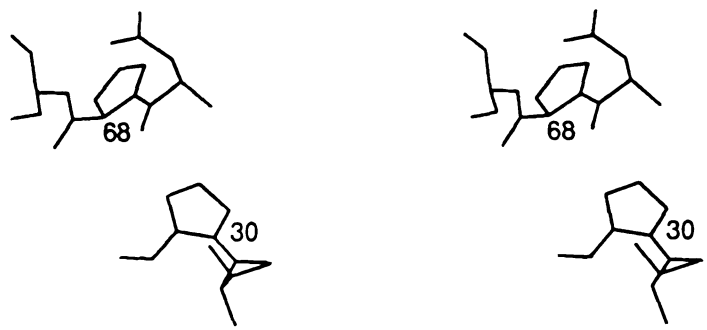


Figure 39. Stereoview showing perpendicular interactions of Pro30 and Pro68.

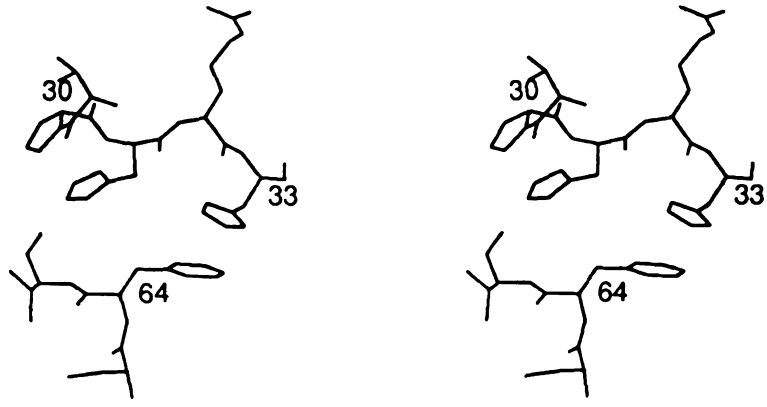


Figure 40. Stereoview showing parallel stacking of His31 and His33 rings with Pro30 and Phe64 side chains.

form of the His31 side chain is uncertain, the presence of an equilibrium between protonated and de-protonated states may also alter the nature of its interactions.

The observed K4 side chain structure is generally in accordance with previous spectroscopic results. The central hyrdophobic core and aromatic side chain interactions agree well with nmr and NOE experiments [35,37,42], which indicated connectivity between these same residues. Some more detailed observations of Atkinson and Williams [37] are also confirmed, including the specific hydrogen-bonding interactions between Thr16 OGH and a Glu73 carboxy group and between Thr65 and His31 (Table 17). Although NOE connectivity was also reported between His33 and Leu46, there is no direct contact observed between the two residues. There may, however, be some indirect communication between the two via the Trp25 side chain.

A particular point of interest regards the side chain conformation of Tyr9. An nmr investigation of K4 suggested the presence of two alternate orientations for this residue, one in which the side chain was free to rotate and one in which rotation was severely hindered [70]. However, in the crystal structure of K4, Tyr9 has a single well-defined conformation in which the tyrosyl ring is highly immobilized (Figure 41). Stabilizing influences in the surrounding structure include the previously mentioned van der Waals interactions of the ring faces to the Leu77 methyl and Gly11 amide groups and hydrogen-bonding of the side chain hydroxyl oxygen. In the present structure, a torsional rotation of 60° about the CA-CB bond would cause the Tyr9 side chain to collide with the amino terminal region of the peptide chain; however, given adequate flexibility of the terminal regions in solution, a rotation of 60-100° would expose most of the tyrosyl side chain to the solvent where it could rotate freely.

As was observed for the main chain atoms, the average thermal factors of

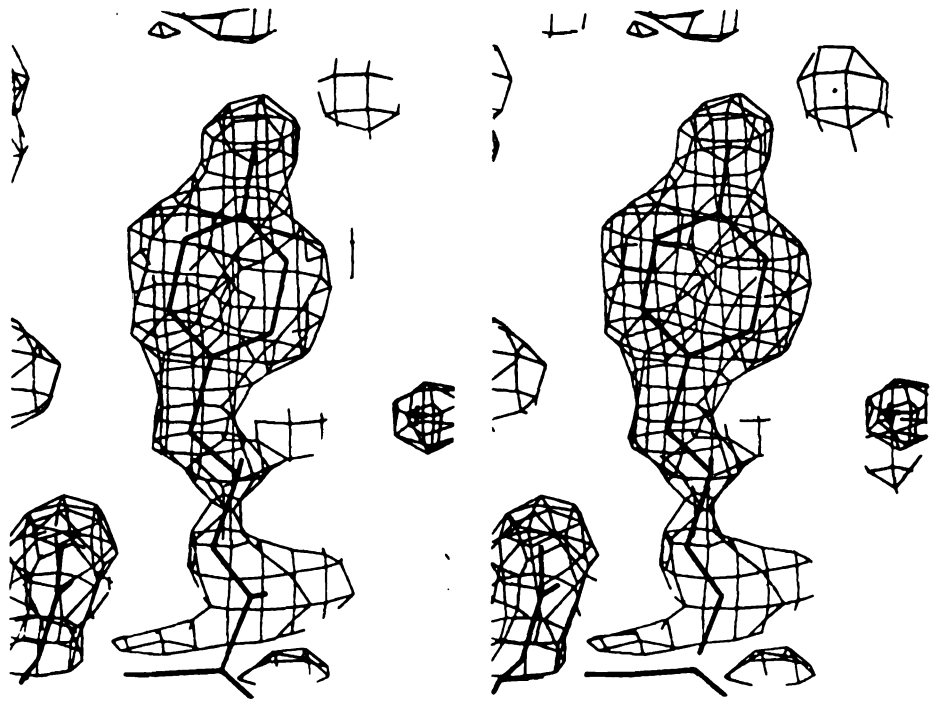


Figure 41. Stereoview showing well-defined electron density observed for Tyr9 side chain.

the side chains correspond closely to their positions in the three dimensional structure (Figure 28). Large thermal factors are observed for the terminal residues of the peptide chain as expected due to the general disorder in this region. The B value of His3 is especially large because the side chain is completely exposed and free to rotate in the solvent. Other residues throughout the peptide chain having particularly large B values include Ser residues 13, 24, 26, and 27, Lys21, Thr29, Gln 34, Asn40, Thr47 and Met48. All are also highly exposed on the surface of the kringle and have no neighboring side chains with which to form stabilizing interactions (cf. Table 17). The side chains of Thr12 and Glu39, for which no density was observed, have similar exposed environments. In contrast, internal side chains, such as those of the Cys, Pro, His, and aromatic residues, have consistently low thermal factors. In some cases, including Tyr2, Trp25, Tyr41 and Tyr74, the side chain B values are actually lower than the values of the corresponding main chain atoms, emphasizing the structural importance of these side chains. Among the aromatic residues, the thermal factor of Tyr50 is relatively high. This can be accounted for by the less restricted environment of this side chain. The Tyr50 ring has a comparatively high accessible surface area, with no interference to movement in the edgewise direction. Furthermore, the Thr12 and Met48 residues adjacent to the faces of the Tyr50 ring are both, to some extent, disordered themselves.

X. INTERMOLECULAR INTERACTIONS AND CRYSTAL PACKING

In the orthorhombic crystal structure, a number of side chain interactions occur between neighboring kringle molecules of the lattice. The most significant of these is a very complex binding interaction which involves three symmetry-related molecules and a sulfate ion (Figure 42). The sulfate anion can be considered to be primarily associated with one of the kringle domains, which binds the ion via the positively charged side chains of Lys35 and Arg71. In addition, the Arg71 guanidinium group also particiates in an ion pair interaction with the Asp57' side chain carboxy group of a second kringle. Further linking the two molecles, the adjacent positively charged Lys58' side chain of the symmetry mate also binds the same sulfate ion. Finally, the Arg32" side chain of yet a third kringle interacts with both the sulfate ion and Asp55 of the first molecule. Thus, all charges in this region are offset, with the negative charges contributed by the sulfate ion (2-), Asp55 (1-), and Asp57 (1-) balanced by the four positively charged side chains Lys35, Arg71, Lys58', and Arg32". It is obvious that, due to the presence of numerous interactions within a small region on the kringle surface, there is a very strong association between these symmetry mates. Interestingly, the interaction involves three residues which are crucial to the kringle lysine-binding site: Asp55 and Asp57, which comprise the negative ionic center, and Arg71, the positive ionic center. This is highly significant, and the interaction will be mentioned again in a later detailed discussion of the binding site.

From the perspective of crystal packing, this trimolecular interaction may be viewed as the result of crystallographic translation of a bimolecular ligand-like interaction, in which Asp55 and Asp57 from the lysine-binding site of one kringle interact with Arg71 and Arg32, respectively, of a second kringle.

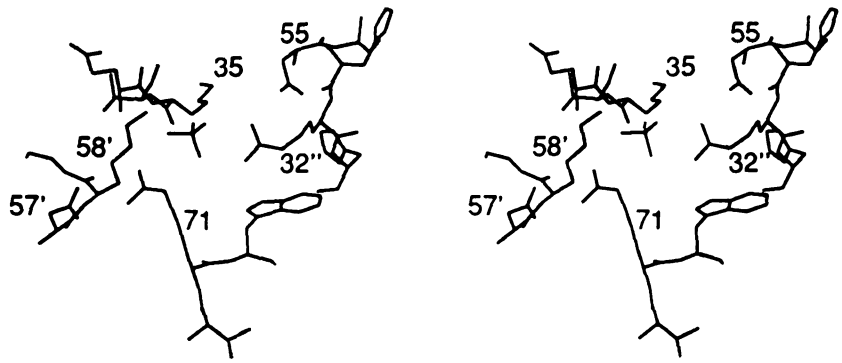


Figure 42. Stereoview showing interactions of sulfate anion with side chains of Lys35 and Arg71 of molecule 1, Lys58' of molecule 2, and Arg32" of molecule 3. Additional interactions occur between Arg71/Asp57' and Asp55/Arg32'.

Repeated throughout the crystal lattice, this coupling of negative and positive centers forms a kringle chain which corresponds to a crystallographic two-fold screw axis in the x-direction (Figure 43).

On a second face of the kringle, an intermolecular hydrogen-bonding interaction occurs between Asn43 ND and Asn76' OD of a neighboring molecule. There is also a solvent-bridged interaction between the Gln7 carboxy group and the side chain amide group of the same Asn76' residue (Figure 44). The symmetry mates linked by these interactions are related crystallographically by a two-fold screw along the y-axis (Figure 45). Finally, the only contact between K4 molecules along the third two-fold screw axis in the z-direction (Figure 46) appears to be a solvent-bridged interaction between the Ser14 OG atom of one molecule and the Ser69' OG atom of a second (Figure 47).

A comparison of the contacts between kringle molecules related by each of the three $P2_12_12_1$ symmetry elements easily explains the observed crystal face development. The intermolecular interactions in the direction of the x-axis are by far the strongest, and, not surprisingly, the orthorhombic K4 crystals grow most rapidly along the **a** axis. Lastly, it is also important to recognize that the interactions observed between K4 domains in the crystal lattice may perturb the participating residuesfrom the conformations which would be found for the isolated molecule in solution. This is certainly true and especially pertinent for the residues of the lysine-binding site, at which the interactions are unusually intimate.

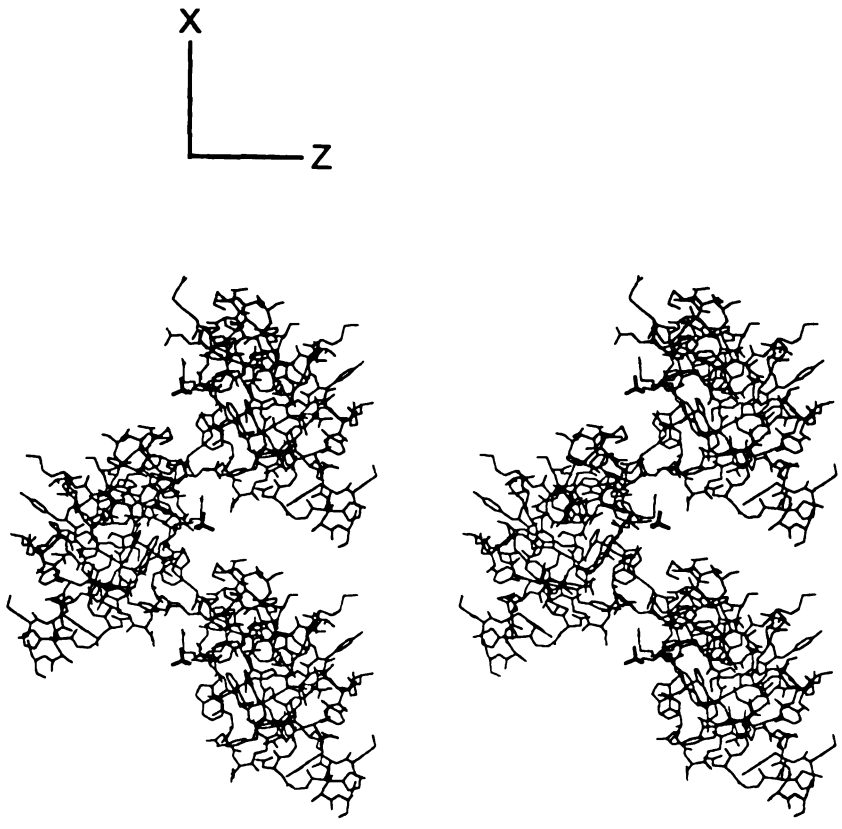


Figure 43. Stereoview illustrating two-fold screw axis in x-direction resulting from trimolecular kringle-kringle interactions at sulfate anion. Sulfate shown in bold.

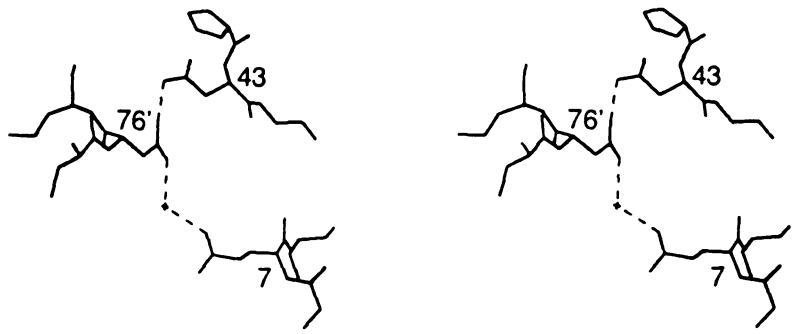


Figure 44. Stereoview of intermolecular interaction between Asn43 side chain and Asn76' side chain of symmetry mate. Also shown is a solvent-bridged interaction between Gln7 and Asn76'. Hydrogen bonds are indicated with dashed lines.

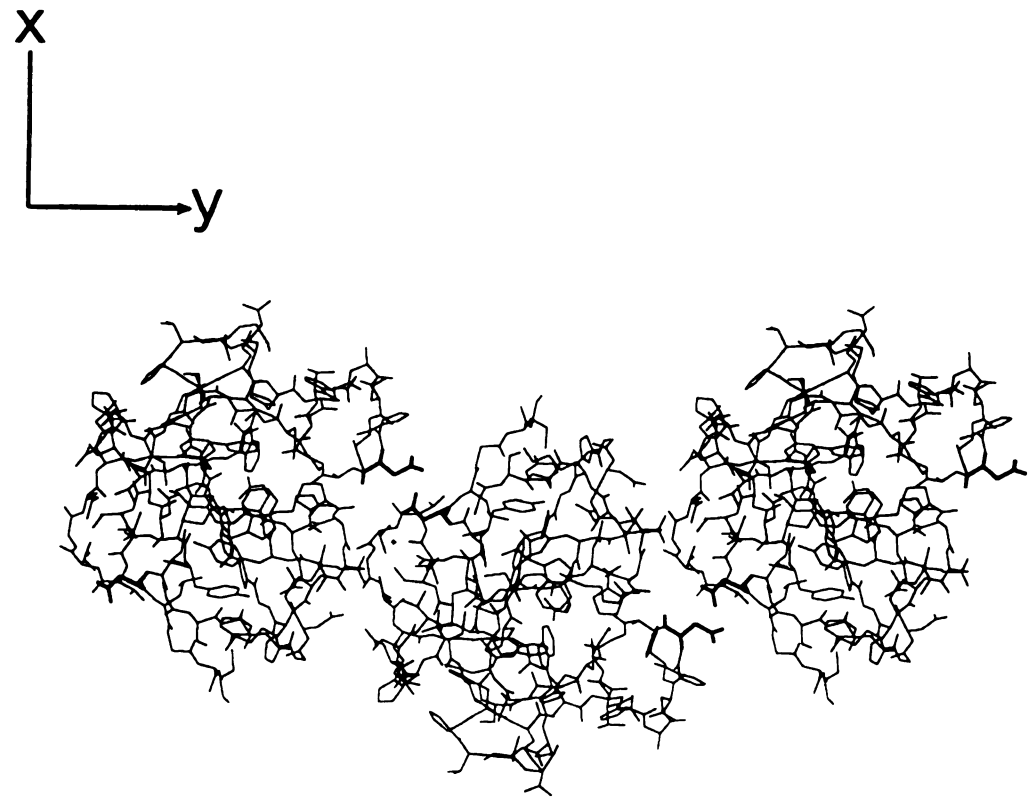


Figure 45. K4 molecules related by two-fold screw axis along y-direction. Residues involved in intermolecular side chain hydrogen bonds are shown in bold.

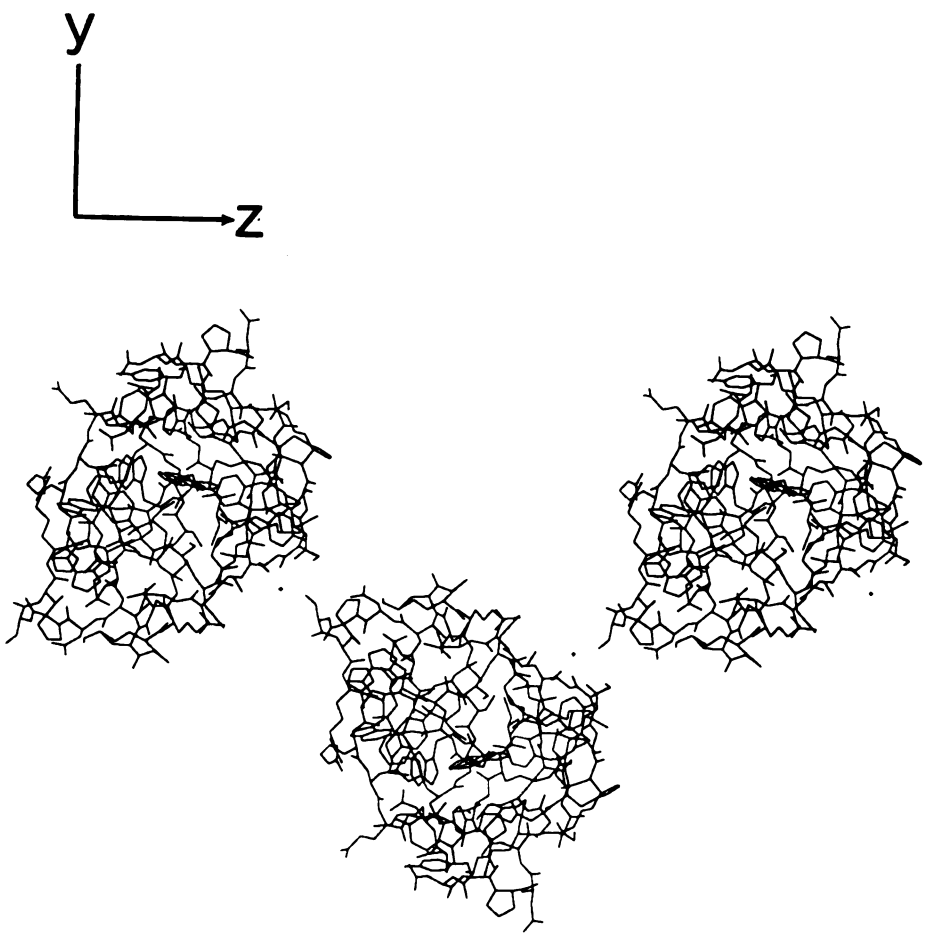


Figure 46. K4 molecules related by two-fold screw axis along z-direction. Solvent molecules which bridge side chains of adjacent symmetry mates are seen at protein-protein interfaces.

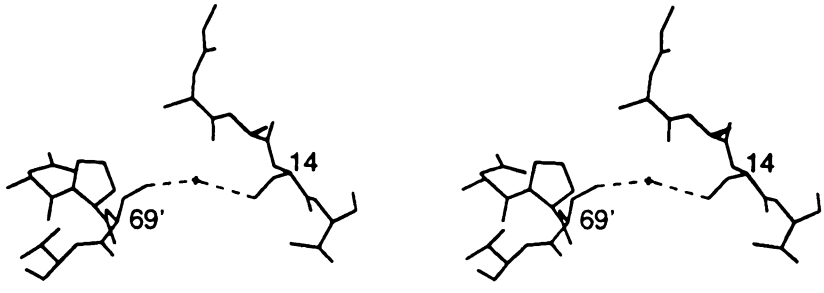


Figure 47. Solvent-bridged interaction between side chain of Ser14 and Ser69' of neighboring molecule. Hydrogen bonds are indicated with dashed lines.

XI. THE STRUCTURE OF THE K4 LYSINE-BINDING SITE

The K4 lysine-binding site is found on the flat surface of the overall plano-convex disc, where it is bordered by the peptide segments 31-35, 54-58, 61-64 and 71-75. Thus, the site is supported primarily by the inner kringle loop, extending from the outer rim of loop C to the far end of loop D. The binding site itself has the form of an elongated depression on the kringle surface which is lined by the indole rings of Trp62 and Trp72 (Figure 48). These two side chains are oriented in an anti-parallel manner with an interplanar angle of approximately 80°. This arrangement forms a hydrophobic trough which lines the bottom and one wall of the binding site depression. Although the Trp62 side chain, which comprises the bottom surface, has minimal solvent accessibility (5%), the long edge of the Trp72 side chain projects out into the solvent and is exposed to an unusually high extent for this residue type (38%). The Trp62 and Trp72 side chains display perpendicular stacking interactions with the rings of Phe64 and Tyr74 respectively, resulting in a symmetric, stabilized structural framework for the binding site. In addition, the 31-35 peptide segment, which bounds the site across one end of the aromatic trough, includes two His residues, which lend further structural stability to this region.

The charged residues which have been implicated in ligand binding are arranged in a semi-circle along the outer edge of the Trp62 indole, completing the remaining walls of the binding site depression. The two negatively charged Asp55 and Asp57 residues are located at one end of the hydrophobic trough, with the carboxy centers within 5.5 Å of one another. The positively charged Arg71 residue is positioned at the opposite end of the trough, with the side chain extending outward away from the site, placing the guanidinium group about 12 Å from the Asp55/Asp57 pair. In addition, although not identified as

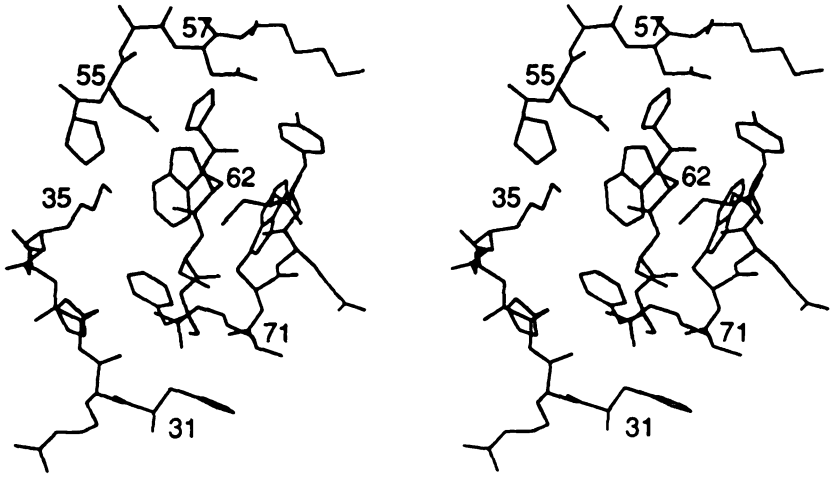


Figure 48. Stereoview of K4 lysine-binding site.

necessary to the binding site, Lys35 is also found along the aromatic trough between the Asp and Arg centers. The Lys35 side chain extends into the site where the amino group is approximately 7.6 Å from the Arg71 side chain. Thus, there appears to be a charge pair at the positive as well as negative center.

It can be seen that the general three-dimensional structure of the binding site corresponds nicely to the observed K4 binding specificity. The terminal positively and negatively charged groups of zwitterionic ligands such as ACA are attracted electrostatically to the complementary opposed anionic and cationic centers of the binding site, while the aromatic trough provides favorable van der Waals interactions for the hydrocarbon body of the ligand. It is believed that the physiological ligand of this binding site is a carboxy terminal Lys residue of fibrin [20], which also has positively and negatively charged ends and therefore may bind in a similar fashion. Though the basic binding mechanism may be the same, it is complicated by the fact that, in this case, the Lys ligand carboxy group is simultaneously the terminus of a fragment of fibrin polypeptide. It appears that certain characteristics of the binding site may help to accommodate the fibrin Lys ligand. Of the two ionic centers, the positively charged end of the binding site is considerably more open. The 54-58 peptide stretch, which serves as the upper boundary of the site, crosses closely to one end of the Trp indole surface and the Asp side chains have limited freedom of movement. However, the lower 31-35 boundary recedes slightly from the aromatic trough and the cationic center is supported by the much more flexible Arg (and possibly Lys) side chain. Thus, although the amino group at the end of the fibrin lysine side chain may fit easily into the tightly held anionic region of the binding site, with a less relaxed topology at the opposite end, steric interference by the fibrin chain might prevent an efficient interaction of the ligand with the cationic center.

The observed binding site also corresponds well to previous experimental results. Aromatic nmr experiments [41,42] found ligand-binding to perturb most strongly the chemical shifts of the side chains of Trp62, Phe64, and Trp72, in accordance with their observed positions at the center of the site. Signals of the residues Trp25, His31, Tyr41 and Tyr74 were also shifted, but to a lesser extent. Of these, Trp25 and Tyr74 communicate with the binding site through aromatic stacking interactions with Trp62 and Trp72 respectively. The His31 side chain is adjacent to the base of the Arg71 side chain and may sense the effect of the ligand on this ionic center. The Tyr41 ring, however, is spatially quite remote from the binding site and may only be affected indirectly through the residues of the hydrophobic core. The side chain of His33 was reported to be sensitive to the binding of bulky, but not linear, ligands [42]. It appears that this ring, which lies parallel to to Phe64, is also indirectly affected by perturbations of the phenylanyl side chain caused by the greater contact with bulkier ligands. The projection of the Lys35 side chain toward the binding site positive center agrees well with the results of chemical modification studies [71] which found the blocking of this residue to weaken affinity of K4 for Lys-sepharose. On the other hand, although a role in ligand-binding has been proposed for His31 [41], this side chain appears to be too far removed to participate directly. Ligand-binding has also been reported to change the flip rate of Tyr9 [37]; however, there is no obvious reason for this observation based on the crystal structure, in which Tyr9 appears only to communicate with the binding site very indirectly. The amino terminal of the K4 peptide chain, which partially restricts the Tyr9 side chain conformation and which is not fully determined in the present X-ray structure, may be responsible for this effect.

Although the general characteristics of the lysine-binding site fit the predicted model of ligand-binding, the detailed side chain conformations at the

ionic centers do not. Both Asp57 and Arg71 side chains are directed away from the binding pocket, with a separation between the charged centers of approximately 12 Å, poorly matching the observed optimal ligand length of 6.8 Å. This unexpected result stems from the intermolecular interactions mentioned previously between Arg71 and Asp57 side chains of neighboring molecules (Figure 49). Thus, these residues are both directed outward to make contact with a symmetry mate, rather than inward in a ligand-binding fashion. However, torsional rotation of the Arg71 side chain easily brings the guanidinium group within 8.5 Å of the anionic center, in good position for ligand binding.

The intermolecular interactions occurring at the binding site appear to affect the conformations of other residues as well. For example, the Lys35 side chain is directed toward the cationic end of the binding pocket to form an ion pair with the sulfate anion which is docked there. This observed behavior suggests the possibility that this residue may similarly reinforce the cationic center during ligand-binding. In the crystal structure Lys58 binds to the sulfate associated with the binding site of a neighboring molecule, drawing the side chain away from the binding site of the host molecule in an extended conformation. In solution, the conformation of this side chain may be very different, especially considering the availability of the adjacent oppositely charged Asp57 side chain of the binding site. Perhaps the most interesting intermolecular contact is the ligand-like binding of the Arg32 residue of a second symmetry mate to both the sulfate ion and Asp55 of the binding site. The Arg32 side chain projects into the hydrophobic trough, with the aliphatic chain abutted within 2.5 Å of the face of the Trp72 indole, mimicking the expected behavior of more typical binding site ligands. Thus, although the kringle-kringle interactions occurring in the crystal structure distort an accurate picture of the free binding site, at the same time they present an interesting view of the kringle ligand

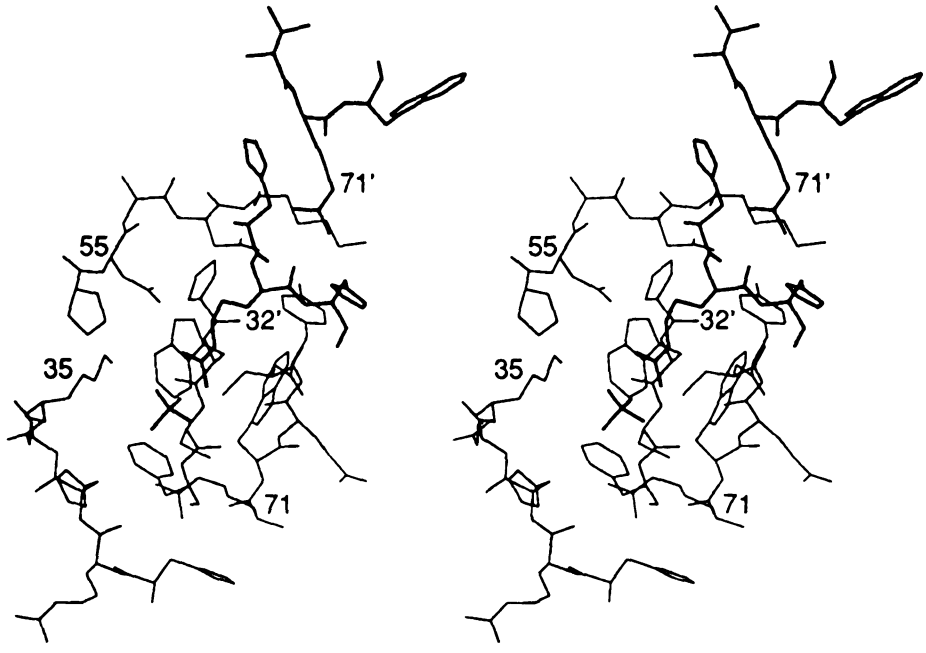


Figure 49. Stereoview showing ion pair interactions which occur at lysine-binding site. Sulfate ion and side chains from second symmetry-related molecule are shown in bold.

binding ability.

The binding site interactions explain the observed inability to bind ACA in the orthorhombic K4 crystals during soaking experiments. The close interactions, which occupy both the ionic centers and the hydrophobic pocket, preclude the access of ACA to the binding site. Should this small ligand eventually gain access to the site and displace the protein-protein interactions, the disruption of these important intermolecular contacts would naturally lead to destruction of the crystal structure itself. This possibility may account for the cracking and crumbling of orthorhombic K4 crystals soaked with high concentrations of the ligand.

XII. SOLVENT STRUCTURE

The final K4 crystal structure includes 97 ordered solvent molecules. These are listed, along with their crystallographic parameters, in Appendix B. The solvent occupancy and thermal factor distributions are shown in Figures 50 and 51. There is a normal distribution of thermal factors about the mean value of 25 $Å^2$, with nearly 75% of the values falling between 20 and 30 $Å^2$. This type of distribution does not occur for solvent occupancies; in this case there is a roughly even spread between values of 0.5 and 1.0.

Among the solvent molecules, one sulfate ion was found to be associated with each kringle molecule (Figure 52). As was previously described, the sulfate anion is located at the cationic center of the K4 lysine-binding site, where it is coordinated by four Arg and Lys side chains from three separate symmetry related molecules. The geometric parameters of these interactions are given in Table 18. As might be expected from the number of stabilizing hydrogen-bonding interactions, the sulfate has a relatively low average thermal factor of 19 $Å^2$, a value which is comparable to the average value for protein atoms. A comparison of the individual atomic B values also agrees well with the observed structure. The central sulfur atom has a B value of 16 $Å^2$, which is significantly lower than the average value of nearly 20 $Å^2$ related to the surrounding oxygen atoms. This difference can easily be explained by a certain degree of rotational freedom about the center of mass. Furthermore, the fact that atoms O2 and O3 each have B values of 18 $Å^2$ whereas O1 and O4 have slightly higher values of 21 and 22 $Å^2$ corresponds nicely with the observed binding interactions. The former two oxygens each have contacts with two protein donors whereas the latter two oxygens each have only one.

The remaining solvent sites are all assigned as water molecules and are

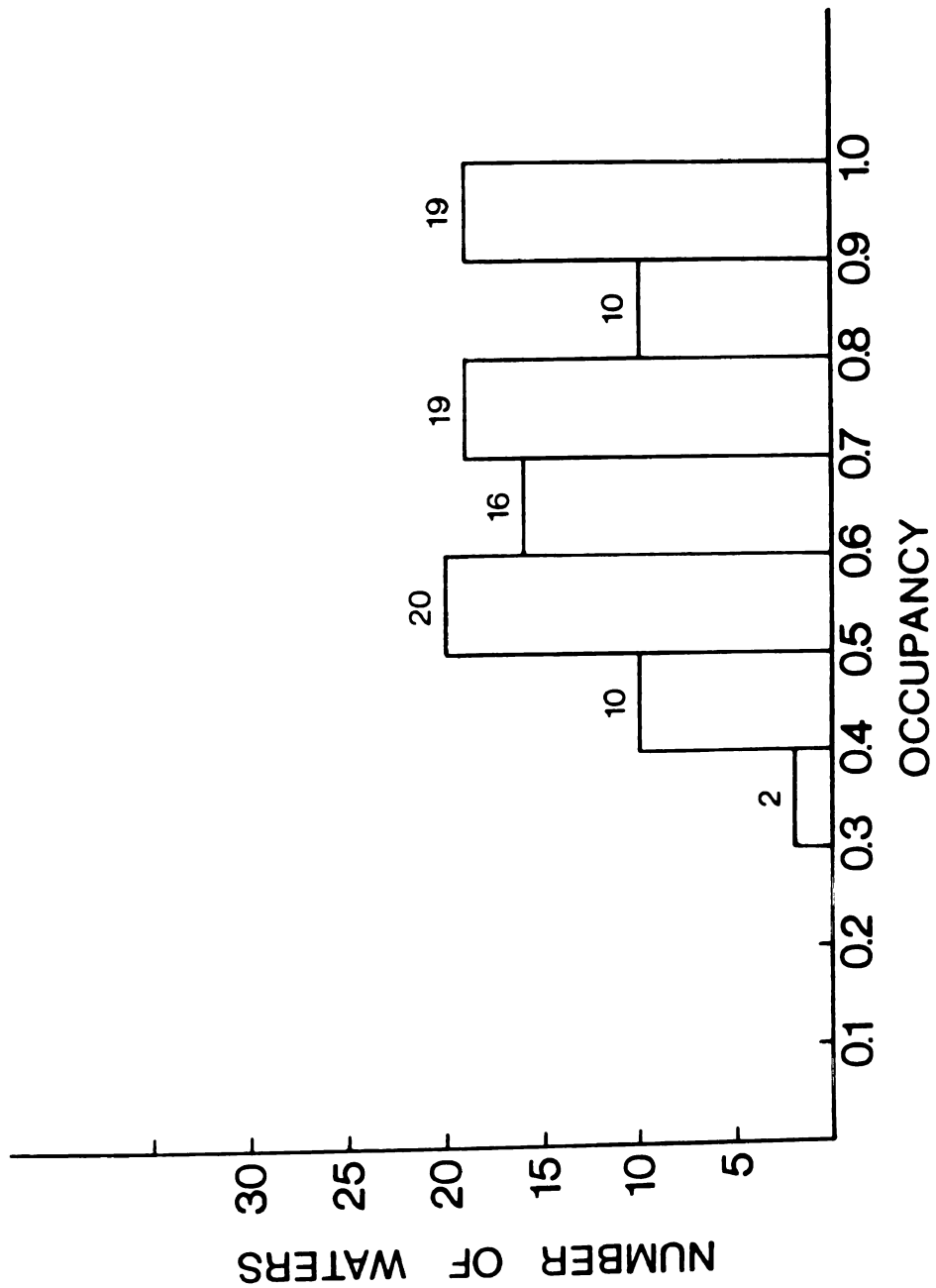


Figure 50. Distribution of occupancies for ordered solvent molecules. Seven water molecules have an occupancy of 1.0.

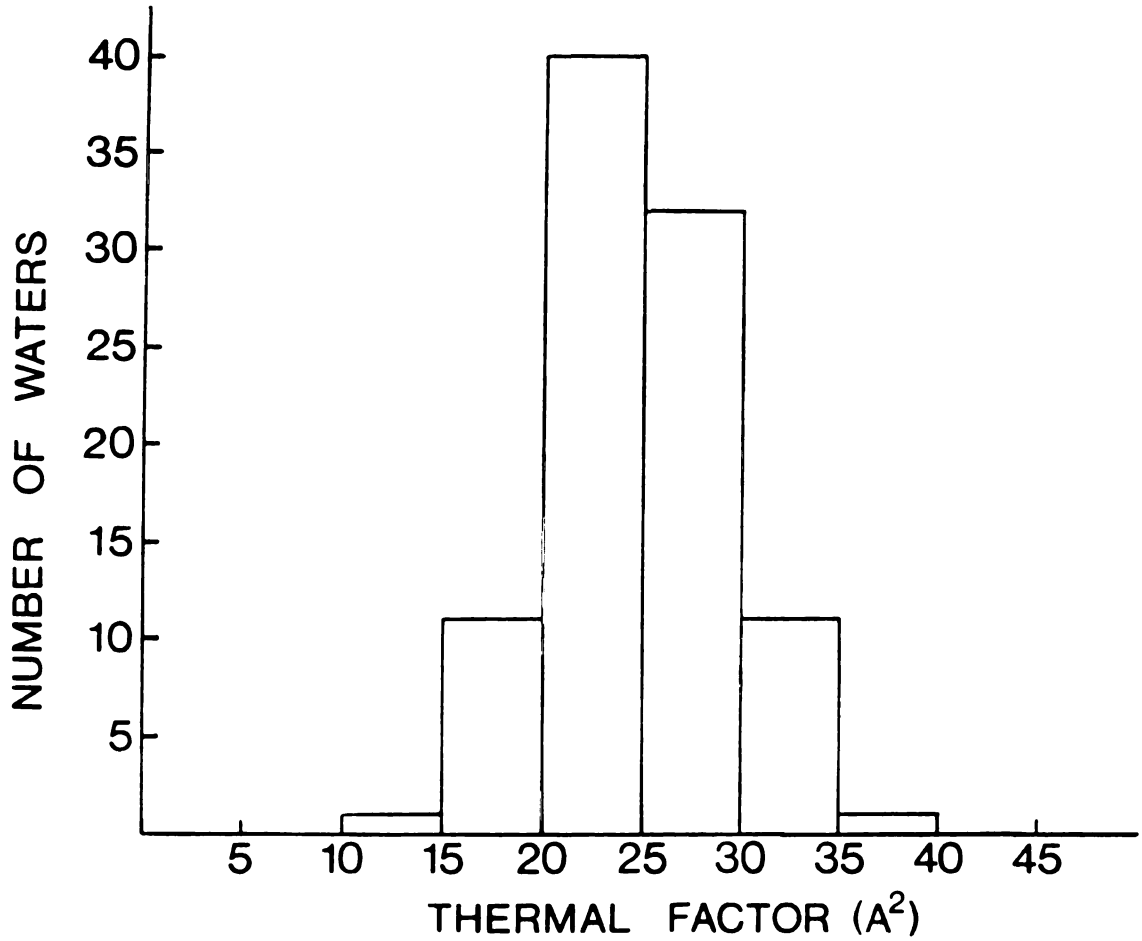


Figure 51. Distribution of solvent temperature factors.

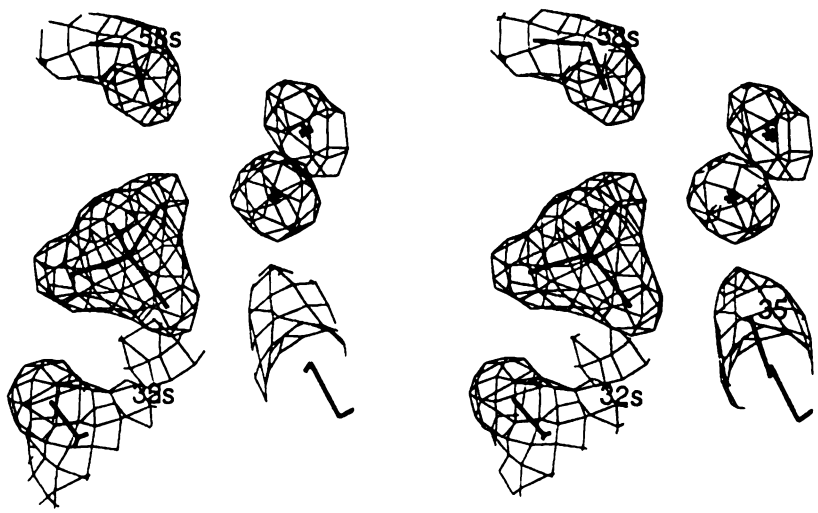


Figure 52. Electron density observed for sulfate anion in orthorhombic K4 crystal structure.

Table 18. Hydrogen-bonding interactions between sulfate anion and protein side chains. Hydrogen atoms are assigned geometrically idealized positions. Primed residue numbers denote hydrogen donors provided by symmetry-related K4 molecules.

ATOM	DONOR	Distances (Å) H..O	N..O	Angles (deg) NHO	SOH	SON
O1	72 Arg NH2	1.85	2.82	162	126	127
O2	35' Arg NH2	1.93	2.85	153	102	108
O2	72 Arg NE	1.94	2.94	162	112	117
O3	35' Arg NH1	1.77	2.75	162	122	123
O3	38 Lys NZ	1.70	2.71	169	128	131
O4	60" Lys NZ	2.08	2.99	148	102	112

ranked in decreasing order according to the quality factor occ^2/B of James and Sielecki [72]. The quality factor gives an approximate measure of relative solvent reliability and ranges in value from 0.087 to 0.005 for the 96 water sites of the K4 structure. Of these solvent molecules, most are closely associated with the K4 molecules, with 73 (76%) forming one or more hydrogen bonds to the protein. Probable solvent-protein hydrogen bonds are given in Tables 19 and 20.

One solvent site is buried within the K4 protein structure, completely isolated from the surrounding solvent (Figure 53). This water molecule bridges the main chain Gln23 O and Phe64 N atoms, but lies in an otherwise hydrophobic pocket, approximately 3.2 Å above the six-membered ring of the Trp25 indole group and adjacent to the additional hydrophobic Trp62 and Phe64 side chains. Consistent with this stabilized and restricted environment, this water molecule has the highest quality factor ranking, with unit occupancy and an exceptionally low B value of 11.5 $Å^2$, which is only slightly greater than the average B values for the neighboring aromatic rings. Two additional solvent sites occur in an empty pocket of the protein structure which is removed from,

Table 19. Protein-solvent hydrogen bonds in which protein atom serves as hydrogen donor (D).

DONOR			SOLVENT		Occ^2/B	Distances (Å) H..O	D..O	Angle (deg) NHO
5	Tyr	OH	Wat	75	0.011	2.09	3.04	155
7	Gly	N	Wat	87	0.009	2.52	3.49	155
8	Asp	N	Wat	20	0.035	2.28	3.07	138
11	Ser	N	Wat	22	0.033	2.17	3.11	153
		OG	Wat	22	0.033	1.71	2.55	145
12	Tyr	OH	Wat	16	0.037	1.72	2.68	154
13	Arg	N	Wat	66	0.013	1.67	2.65	176
		NE	Wat	91	0.007	2.19	3.04	138
		NH2	Wat	61	0.015	1.61	2.60	161
17	Ser	N	Wat	7	0.052	1.93	2.87	166
		OG	Wat	57	0.016	2.29	3.21	163
23	Lys	NZ	Wat	54	0.016	2.08	2.98	147
27	Ser	N	Wat	10	0.045	1.99	2.99	168
		OG	Wat	37	0.022	2.18	3.07	152
31	Met	N	Wat	44	0.019	1.95	2.95	171
37	Gln	N	Wat	5	0.057	1.97	2.90	161
		NE2	Wat	33	0.024	1.87	2.79	153
38	Lys	NZ	Wat	26	0.030	2.14	2.92	133
		NZ	Wat	39	0.021	1.94	2.84	140
43	Tyr	OH	Wat	25	0.031	1.84	2.83	160
		OH	Wat	26	0.030	2.21	3.12	144
45	Asn	N	Wat	4	0.059	2.00	2.89	149
		ND2	Wat	46	0.018	1.97	2.92	155
47	Gly	N	Wat	18	0.037	2.21	3.14	147
48	Leu	N	Wat	3	0.062	1.97	2.95	160
54	Arg	NH1	Wat	31	0.025	1.77	2.71	156
55	Asn	N	Wat	2	0.065	2.12	3.05	166
65	Phe	N	Wat	1	0.087	2.00	2.95	156
70	Ser	OG	Wat	32	0.024	2.00	2.89	153
72	Arg	NH1	Wat	70	0.013	1.85	2.70	141
73	Trp	N	Wat	36	0.023	1.94	2.93	167
		NE	Wat	47	0.018	2.14	2.93	137
75	Tyr	N	Wat	14	0.039	1.97	2.85	150
77	Asn	ND2	Wat	6	0.055	1.92	2.76	139
		ND2	Wat	9	0.048	2.23	3.00	132

Table 20. Protein-solvent hydrogen bonds in which protein atom serves as hydrogen acceptor. Potential interactions were accepted for protein-solvent distances within 3.5 Å.

ACCEPTOR			SOLVENT		Occ^2/B	O..O Distance (Å)	COO Angle (deg)
4	Cys	O	Wat	38	0.022	2.53	151
6	His	O	Wat	22	0.033	2.92	134
8	Asp	OD2	Wat	15	0.038	2.65	116
		OD2	Wat	20	0.035	3.29	93
		OD1	Wat	45	0.019	2.54	124
9	Gly	O	Wat	8	0.051	2.84	154
10	Gln	O	Wat	61	0.015	2.48	144
		OE1	Wat	6	0.055	2.72	138
		OE1	Wat	11	0.043	3.14	122
11	Ser	O	Wat	17	0.037	3.00	118
		O	Wat	74	0.012	3.36	171
12	Tyr	O	Wat	8	0.051	2.90	132
13	Arg	O	Wat	89	0.009	3.06	157
15	Thr	O	Wat	57	0.016	2.51	158
19	Thr	O	Wat	63	0.014	3.08	148
20	Thr	O	Wat	82	0.010	2.80	134
		O	Wat	63	0.014	3.30	94
21	Thr	O	Wat	54	0.016	3.11	131
		O	Wat	17	0.037	2.84	135
22	Gly	O	Wat	56	0.016	2.50	135
		O	Wat	15	0.038	2.71	162
25	Cys	O	Wat	12	0.040	2.80	141
26	Gln	O	Wat	1	0.087	2.89	139
31	Met	O	Wat	30	0.027	2.94	146
29	Ser	O	Wat	60	0.015	2.87	113
32	Thr	O	Wat	67	0.013	2.82	141
		OG	Wat	29	0.028	3.44	131
33	Pro	O	Wat	40	0.020	2.88	144
		O	Wat	59	0.015	3.18	133
36	His	O	Wat	23	0.031	2.73	142
37	Gln	O	Wat	71	0.013	3.25	138
		OE1	Wat	69	0.013	2.47	108
		OE1	Wat	14	0.039	3.19	98
42	Asn	O	Wat	88	0.009	3.16	131
		O	Wat	78	0.011	2.29	139
43	Tyr	O	Wat	3	0.062	2.67	151
44	Pro	O	Wat	18	0.037	2.68	127
45	Asn	O	Wat	63	0.014	3.38	113
		O	Wat	61	0.015	3.38	131
46	Ala	O	Wat	2	0.065	2.64	127
47	Gly	O	Wat	31	0.025	3.14	119
48	Leu	O	Wat	86	0.010	2.73	138
50	Met	O	Wat	27	0.029	2.95	125
51	Asn	OD1	Wat	41	0.020	2.88	129
55	Asn	O	Wat	11	0.043	2.92	132

Table 20 (cont'd)

56	Pro	O	Wat	24	0.031	2.55	142
		O	Wat	9	0.048	2.95	128
57	Asp	O	Wat	40	0.020	2.76	135
		OD1	Wat	30	0.027	2.84	116
		OD1	Wat	39	0.021	3.27	122
58	Ala	O	Wat	45	0.019	3.21	110
		O	Wat	59	0.015	3.30	123
59	Asp	O	Wat	45	0.019	3.26	121
		OD2	Wat	70	0.013	3.16	104
		OD2	Wat	40	0.020	3.06	108
		OD2	Wat	59	0.015	3.31	119
60	Lys	O	Wat	35	0.023	2.93	124
62	Pro	O	Wat	13	0.040	2.92	138
67	Thr	O	Wat	52	0.017	2.80	146
68	Asp	OD2	Wat	32	0.024	2.71	127
		OD1	Wat	80	0.011	3.11	137
69	Pro	O	Wat	76	0.011	3.24	138
70	Ser	O	Wat	76	0.011	3.40	83
		OG	Wat	57	0.016	2.68	160
73	Trp	O	Wat	69	0.013	2.78	143
		O	Wat	36	0.023	2.60	131
74	Glu	OE2	Wat	93	0.007	2.80	176
76	Cys	O	Wat	34	0.023	2.72	155
77	Asn	O	Wat	53	0.017	2.30	146
78	Leu	O	Wat	46	0.018	2.71	132
		O	Wat	75	0.011	3.20	156

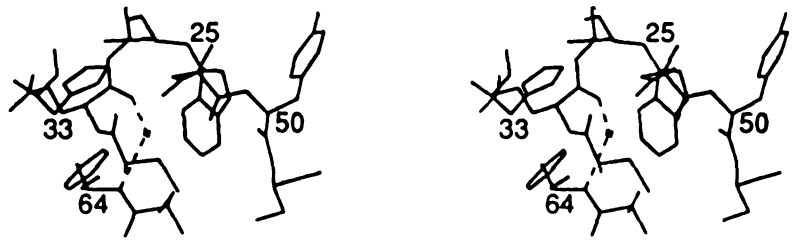


Figure 53. Stereoview of internal solvent molecule in K4 structure. Dashed lines indicate hydrogen bonds which bridge Gln23 and Phe64 main chain atoms.

but accessible to, the bulk solvent (Figure 54). One of these water molecules bridges the main chain carbonyl and amide groups of Ala44 and Asn53 respectively, while the second is within hydrogen-bonding distance of both Gly6 and Tyr9 carbonyl oxygens. Both sites are among the highest ranked solvent positions (Wat2, Wat8), having unit occupancies and relatively low thermal factors.

The remaining protein-bound solvent molecules are found at the K4 surface, particularly in regions near the protein-protein interfaces. Eleven water molecules have positions within hydrogen-bonding distance of two separate symmetry related molecules. With the exception of the three described above, the highest ranked water molecules are all found in contact regions between symmetry-related protein molecules, and most form hydrogen bonds with protein main chain atoms. The Wat6 molecule was mentioned previously as the solvent bridge between the Asn43 side chain and the Asn76' side chain of a symmetry mate related by the y-screw axis. In the same vicinity, Wat3 bridges the main chain Tyr41 O and Leu46 N atoms, Wat4 binds to the Asn43 main chain N, and Wat9 interacts with both the Pro54 carbonyl oxygen and the Asn76' side chain amino group. At the lysine-binding site region, Wat5 is within binding distance of both the sulfate anion and the amide nitrogen of Gln34. Finally, Wat7 is found within hydrogen-bonding distance of the main chain amide nitrogen of Ser14, a residue which also forms a solvent-bridged interaction with a K4 neighbor related along the z-axis.

The outer shell solvent molecules which do not have direct hydrogen-bonding interactions with the protein are also concentrated mainly in the more stabilized regions of close approach between neighboring K4 domains. Nineteen of these sites are within hydrogen bonding distance of other solvent molecules (Table 21). The five remaining sites are all within 5 Å of polar protein

groups or other solvent molecules.

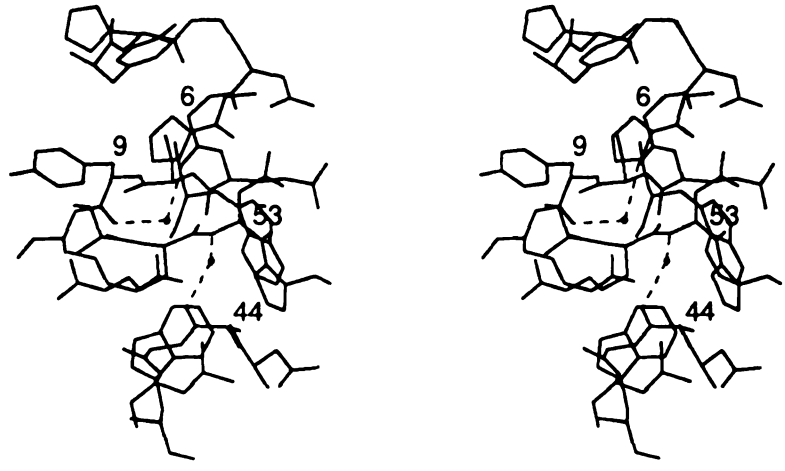


Figure 54. Stereoview showing two solvent molecules found in an empty cavity off the surface of the K4 structure. Dashed lines indicate potential hydrogen bonds.

Table 21. Probable solvent-solvent interactions having distances within 3.5 Å.

SOLVENT 1		SOLVENT 2		Distance (Å) 1-2
SO4	O1	Wat	5	2.65
SO4	O4	Wat	19	2.68
Wat	2	Wat	8	2.89
Wat	2	Wat	11	3.32
Wat	3	Wat	62	2.61
Wat	4	Wat	9	2.84
Wat	4	Wat	88	3.08
Wat	5	Wat	33	2.54
Wat	6	Wat	24	3.00
Wat	7	Wat	72	3.16
Wat	9	Wat	11	2.98
Wat	10	Wat	67	2.42
Wat	10	Wat	12	3.02
Wat	11	Wat	61	3.08
Wat	13	Wat	21	3.04
Wat	13	Wat	34	2.62
Wat	14	Wat	49	3.08
Wat	15	Wat	56	2.17
Wat	15	Wat	20	2.90
Wat	15	Wat	28	3.07
Wat	17	Wat	54	3.34
Wat	17	Wat	73	2.64
Wat	19	Wat	44	3.11
Wat	19	Wat	39	3.50
Wat	20	Wat	56	3.30
Wat	20	Wat	84	3.39
Wat	20	Wat	90	2.81
Wat	21	Wat	77	3.13
Wat	22	Wat	68	2.76
Wat	23	Wat	69	3.17
Wat	23	Wat	47	3.34
Wat	26	Wat	71	3.15
Wat	27	Wat	86	2.37
Wat	29	Wat	30	2.72
Wat	30	Wat	40	3.21
Wat	30	Wat	39	2.59
Wat	34	Wat	46	3.34
Wat	34	Wat	77	2.64
Wat	34	Wat	75	3.35
Wat	40	Wat	59	2.91
Wat	41	Wat	57	3.03
Wat	42	Wat	48	2.60
Wat	45	Wat	92	3.14
Wat	46	Wat	75	3.49
Wat	49	Wat	77	3.10
Wat	50	Wat	72	3.44

Table 21 (cont'd)

Wat	51	Wat	65	2.21
Wat	56	Wat	68	2.41
Wat	59	Wat	70	2.24
Wat	60	Wat	64	2.23
Wat	60	Wat	83	3.46
Wat	61	Wat	63	2.77
Wat	72	Wat	95	2.54
Wat	89	Wat	96	2.49

XIII. COMPARISON OF PLASMINOGEN K4 AND PROTHROMBIN K1

To compare the refined K4 structure with that of the previously solved prothrombin K1 [32], the two structures were first optimally aligned so as to minimize the deviations of the CA, C, N main chain atoms. Only residues 5-57 and 61-75 were used for this procedure, omitting residues in the less ordered terminal regions and in the immediate vicinity of the one residue insertion in PTK1. At this point the rms deviation between the two backbone structures was 0.75 Å. After removing 62 atoms having deviations greater than this value and re-optimizing the alignment, the rms deviation was reduced to 0.47 Å. Those atoms removed came mostly from the segments of loop B on either side of Pro30. Average deviations of individual residues are shown in Figure 55.

From a superposition of both K4 and PTK1 CA backbone structures (Figure 56) it can be seen that there are few large deviations between the two structures. One difference which was anticipated occurs at the insertion of one residue on the C loop of PTK1. The actual observed perturbation in the conformation of the backbone caused by the insertion is minimal. The loop is merely expanded in the immediate vicinity of the insertion, and only the positions of two residues residues on either side of the insertion are significantly affected. The conformations of Asp55 and Trp62 are nearly identical. The most noticeable difference in the two structures occurs in the terminal regions of the peptide chain, preceding residue 9 and following residue 75. The Cys1-Cys80 disulfides of K4 and PTK1 are parallel but displaced by approximately 3.3 Å. The average deviation between the two structures for the terminal regions alone is 2.09 Å. Several factors may influence the large general difference in this region. The conformation may be indirectly affected by the insertion at position 59, which relaxes the C loop of PTK1 slightly, thus allowing the terminal regions of

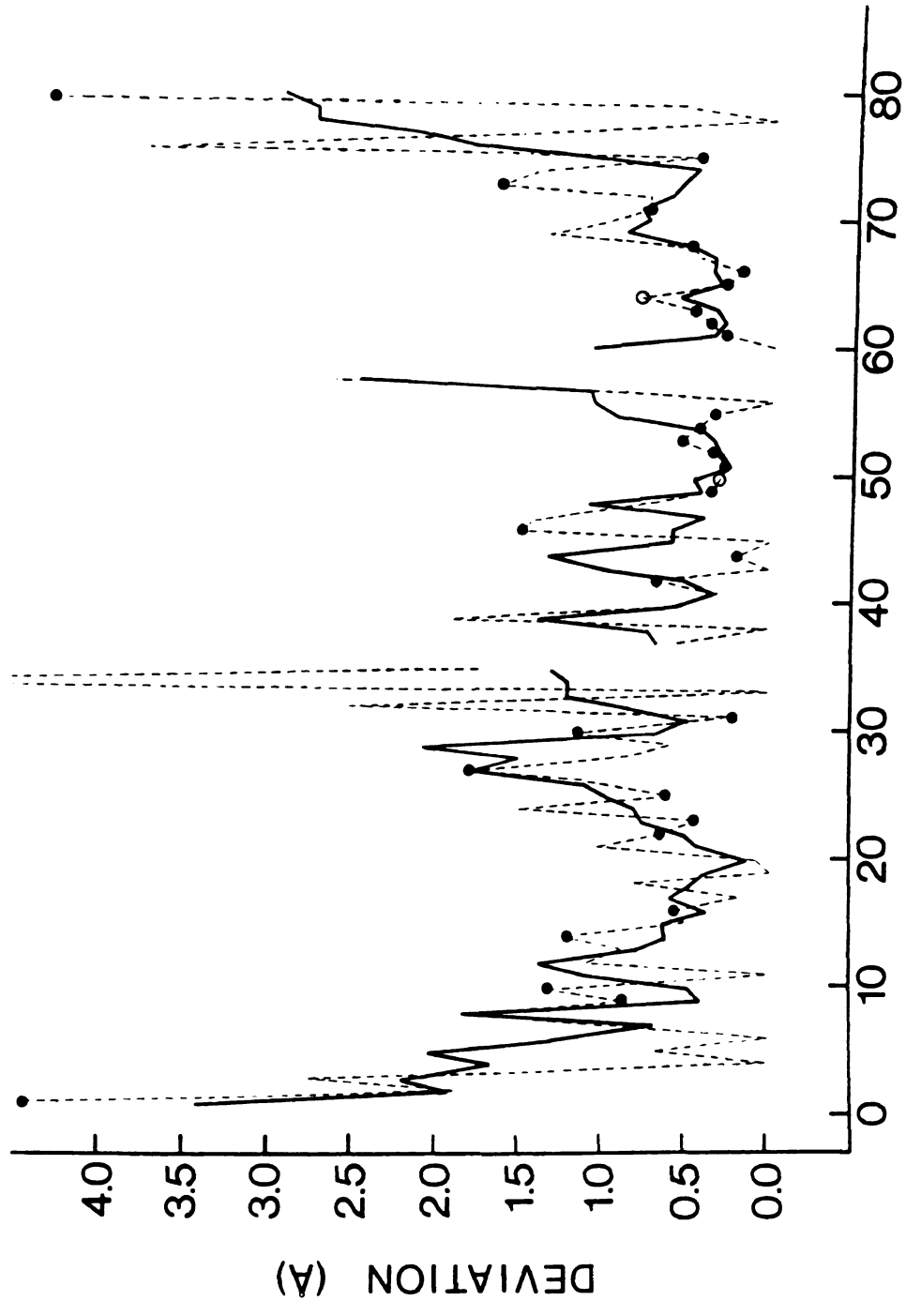


Figure 55. RMS differences between K4 and PTK1 main chain (solid line) and side chain (dashed line) positions. Filled circles indicate conserved residues; open circles indicate highly homologous Tyr/Phe substitutions.

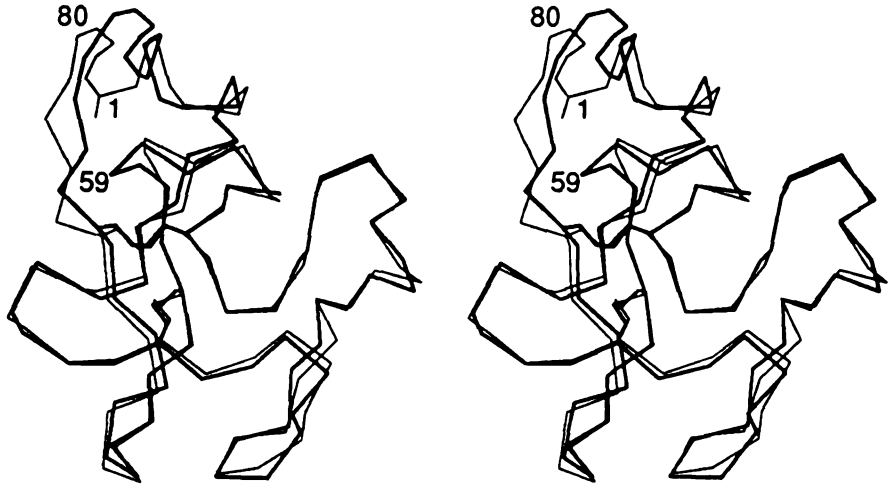


Figure 56. Comparison of CA backbone and disulfide structures of K4 (bold) and PTK1.

the backbone to shift toward the amino terminal side. In addition, PTK1 residues 77 and 78 form an antiparallel β-sheet with the carboxy tail in the PT fragment 1 crystals whereas the plasminogen K4 structure has no similar stabilizing interactions. Finally, as is suggested by the general disorder observed in this region for K4, the terminal peptide segments of kringles in general may have some inherent flexibility.

A comparison of the complete backbones, including positions of the carbonyl oxygen atoms, highlights some additional differences between the K4 and PTK1 main chain structures (Figure 57). Most notable is the conformation of the conserved Pro30 residue, which has a *trans* configuration in PTK1 but was changed to the *cis* isomer in the K4 structure (Figure 58). This difference in configuration between the two kringles is surprising and may not be real. Although there is well-defined electron density for the carbonyl oxygen atoms throughout the K4 structure, the electron density maps for PTK1 were not found to be as clear and thus this apparent difference may be due to a misinterpretation of the latter.

The orientations of the carbonyl oxygen atoms indicate large differences between the dihedral angles of the two kringles in several other regions, such as the relative twist between the two backbones from residue 33 to 35. Other particularly large differences can be seen between the carbonyls of residues 39 and 55, which are 90° apart, and those of residues 44 and 48, which are 180° apart. However, the majority of the main chain atoms have very similar positions. The agreement is particularly striking (< 0.40 Å) for segments 50-55 and 61-65, which form part of the binding site in K4. Again, due to the difference in the resolution and quality of the electron density maps for the two structures, it is difficult to ascertain whether the differences observed are meaningful. However, most of the residues for which the main chain positions are identical are

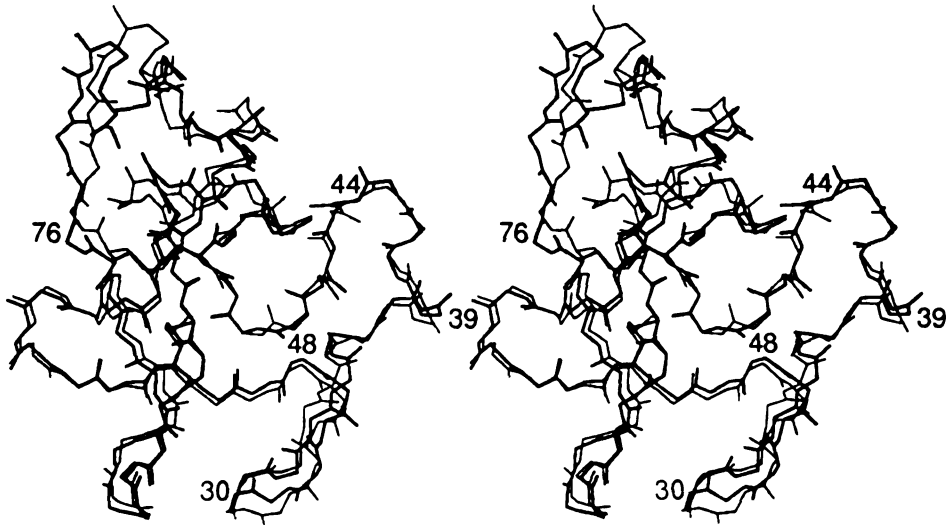


Figure 57. Comparison of CA, C, N, O main chain positions of K4 (bold) and PTK1.

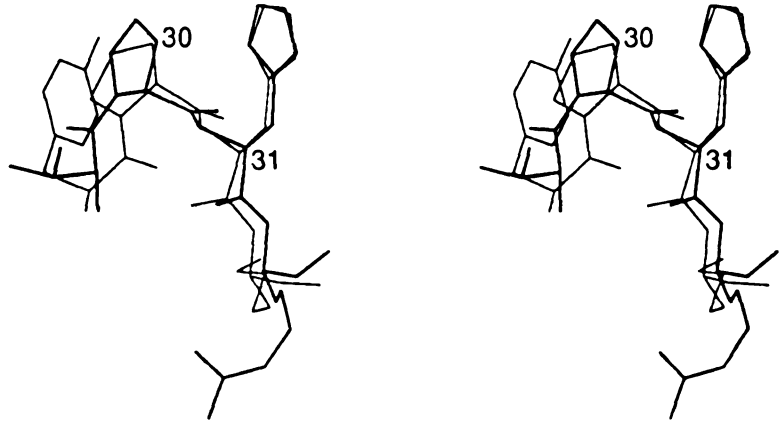


Figure 58. Stereoview comparing conformations of Pro30, which appears as *cis* isomer in K4 (bold) but as *trans* isomer in PTK1.

conserved between the two kringles and many participate in hydrogen-bonding interactions.

The hydrogen-bonding scheme and secondary structure observed in K4 are nearly identical to that reported for PTK1. Turns and β-sheet are observed in corresponding positions. Only one main chain-main chain hydrogen bond link observed for PTK1 is not seen in K4: 47 NH-45 CO [32]. On the other hand, the main chain interactions of PTK1 account for only 10 of the 25 such hydrogen bonds of K4. Most of the non-conserved hydrogen bonds are prevented in the PTK1 structure by a relative twist between the two backbones or by displacements of the backbone in the terminal or insertion regions.

The conformations of the central disulfides of K4 and PTK1 are also very similar, with most of the deviations between the positions of the Cys side chains appearing to be caused by deviations of the backbone (Figure 59). There is an approximately 30° difference between the torsional angles of the Cys63 and Cys75 side chains of K4 and PTK1; however, the deviations between the midpoints of the Cys22-Cys63 and Cys51-Cys75 disulfides in the K4 and PTK1 structures are 0.49 Å and 0.34 Å respectively, close to the combined coordinate error of the two structures. Unlike K4, the PTK1 structure includes only one side chain conformation for Cys75 - that which results in the more parallel relative orientation of the two interior disulfides.

The conformations of side chains conserved between K4 and PTK1 are also generally very similar, with an average deviation of 0.66 Å. A comparison of the observed side chain torsional angles is given in Table 22. There is particulary good agreement of side chains which appear to be important in maintaining the basic kringle structure. Of the 35 conserved side chains, 28 (80%) are structurally important. These include Cys, Pro, and Gly residues, as well as internal side chains. There are only two non-conserved Pro residues:

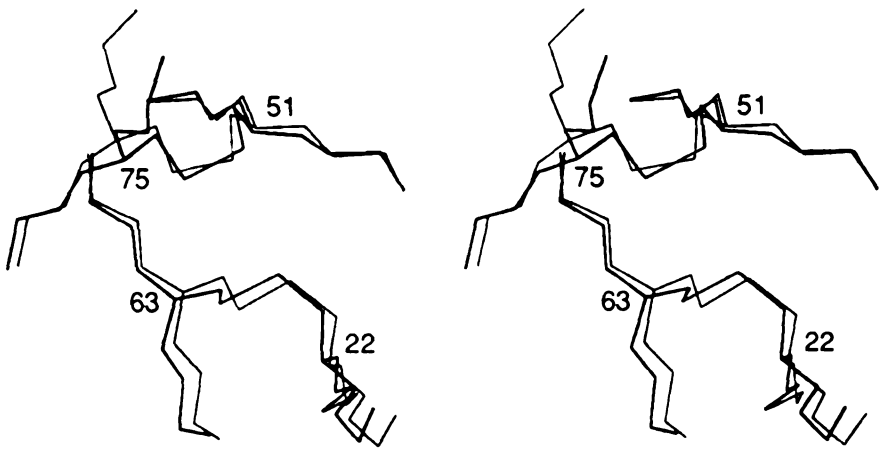


Figure 59. Stereoview comparing conformations for inner disulfides in K4 (bold) and PTK1. Only the more similar of two observed K4 Cys75 side chain orientations is shown.

Table 22. Deviations in positions and torsion angles of conserved or highly homologous side chains of PGK4 and PTK1. No torsion angles are given for conserved Ala or Pro residues.

Residue	Coord. (Å)	Torsion angles (deg) K4	PTK1	DIFF
Cys 1	4.45	56	47	9
Tyr 9	0.87	178	-171	11
		70	74	4
Arg 10	1.31	-68	-52	16
		-17	-56	39
		-177	174	9
		-67	154	139
		166	-163	31
Ser 14	1.23	58	-38	96
Thr 16	0.56	-43	-26	17
Cys 22	0.65	-70	-74	4
Gln 23	0.45	178	163	15
		-173	-164	9
		138	137	1
Trp 25	0.62	-61	-72	11
		107	106	1
Ser 27	1.82	172	-108	80
Pro 30	1.18	-	-	-
His 31	1.18	-56	-41	15
		-71	-70	1
Gln/Glu 34	5.83	-59	-35	24
		-178	67	115
		114	102	12
Tyr/His 41	0.30	-69	-62	7
		-97	-110	13
Pro 42	0.68	-	-	-
Ala 44	0.19	-	-	-
Leu 46	1.49	-88	-70	18
		22	-149	171
Asn 49	0.35	65	102	37
		-27	-34	7
Tyr 50	0.30	-56	-60	4
		-34	-55	21
Cys 51	0.27	-69	-76	7
Arg 52	0.33	-59	-81	22
		-75	-29	46
		-176	168	16
		-95	-129	34
		-6	4	10
Asn 53	0.53	-173	-153	20
		-173	-161	12
Pro 54	0.42	-	-	-
Asp 55	0.33	61	29	32
		170	-149	41
Pro 61	0.28	-	-	-

Table 22 (cont'd)

Trp 62	0.36	61	67	6
		-85	-98	13
Cys 63	0.46	64	96	32
Phe/Tyr 64	0.78	-72	-63	9
		-72	-78	6
Thr 65	0.27	-87	-95	8
Thr 66	0.17	56	68	12
Pro 68	0.48	-	-	-
Arg 71	0.74	177	-151	32
		178	-180	2
		-174	145	41
		83	75	8
		2	6	4
Glu 73	1.63	-66	178	116
		175	155	20
		146	159	13
Cys 75	0.43	-79	-49	30
Cys 80	4.22	-67	-79	12

Pro40 which appears in K4, and Pro78 in PTK1. The lack of conservation of these Pro residues may account in part for the main chain deviations between the two kringles which are also observed in the same regions.

Nearly all of the internal residues are conserved. Those which are not conserved are substituted with residues having analogous properties. The Phe64 side chain of K4 is replaced with the nearly identical Tyr side chain in PTK1. Similarly, His33 and Leu77 are replaced by analogous bulky hydrophobic Pro and Val residues, respectively, in PTK1. This high degree of internal conservation maintains the hydrophobic core, which, as previously mentioned, appears to be an important structural foundation of the kringle fold (Figure 60). Nearly all the core residues have deviations less than 0.50 Å. The seemingly large deviation of 1.5 Å between Leu46 residues is the result of a 180° torsional rotation of the side chain. Although the torsional orientation of this residue is certain in the case of K4, the resolution at which the PTK1 structure was refined (2.25 Å) is not sufficient to distinquish reliably between these two conformations, and thus, this difference is probably not real. As was observed for buried side chains in general, residues of the K4 hydrophobic core which are not conserved, both buried and somewhat exposed, are replaced by similar bulky hydrophobic side chains in the PTK1 structure (Figure 60). For instance, two substitutions involving the nearly identical Tyr and Phe residues occur at positions 50 and 64. In addition, the K4 Tyr41 residue is replaced by His in PTK1, with the positions of the side chain rings deviating by only 0.47 Å. Similarly, although K4 His33 is replaced by a less homologous Pro residue in PTK1, the CA carbon is displaced so as to bring the midpoints of the His and Pro rings within 1.6 Å of one another.

The hydrophilic internal side chains also have very similar conformations in the two structures and, in some cases, share similar hydrogen-bonding

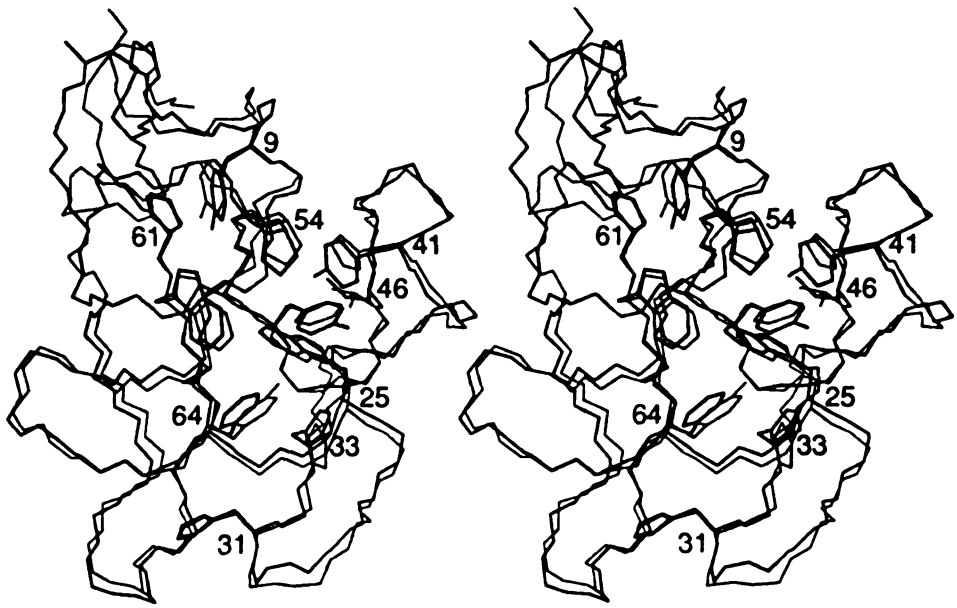


Figure 60. Comparison of hydrophobic core residues in K4 (bold) and PTK1 structures.

interactions. Residues Arg52 and Thr16 have deviations of 0.33 and 0.56 Å, respectively, and both have conserved intramolecular interactions. The Glu23 residue shows a similar deviation of 0.45 Å (Figure 61). In the K4 structure, this side chain is fully involved in hydrogen-bonding with the His31 main chain, and although analogous interactions were not identified for the PTK1 structure, it is likely that they exist since the His31 atomic positions are also very similar between the two structures. The conservation of these internal interactions supports the role of these residues in further stabilizing the kringle fold. The positions of the Asn53 side chains differ by only 0.53 Å, although the intramolecular interactions with the carbonyl oxygens of residues 5 and 57 which are observed in the K4 structure are not possible in PTK1 due to differences in the main chain conformation at the amino terminus and at the site of the residue 59 insertion (Figure 62). Finally, Thr65 has a small average deviation of 0.27 Å although no stabilizing interactions are observed in either structure.

Several of the conserved residues found at the kringle surface also have similar side chain conformations in K4 and PTK1. Residues Asn49 and Asp55 both have deviations of approximately 0.35 Å and participate in the same hydrogen-bonding interactions in the two kringles (Figures 61,62). The side chain of Thr66 also shows little discrepancy (0.17 Å), but does not have common hydrogen bonds in both kringles. Most interesting is the similarity in the conformations of the K4 and PTK1 Arg71 side chains despite the fact that they have entirely different binding interactions (Figure 63). The Arg side chain in K4 interacts with a symmetry mate and sulfate ion in a way which is particular to the crystal form, whereas the corresponding side chain in PTK1 appears to form an ion pair with the Glu34 residue, which is not conserved between the two kringles. Nevertheless, the guanidinium groups of both are oriented similarly and are within 0.75 Å of one another.

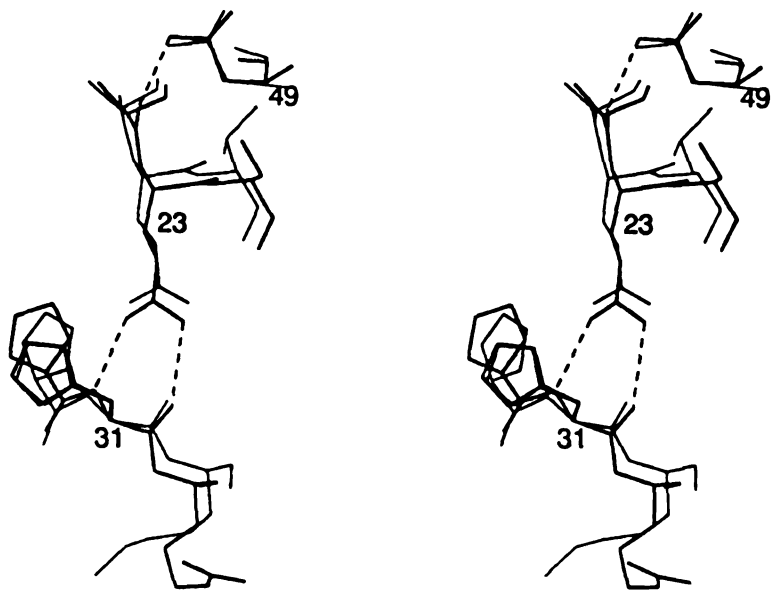


Figure 61. Stereoview showing similar conformations and hydrogen-bonding interactions of conserved residues Glu23 and Asn49 in K4 (bold) and PTK1.

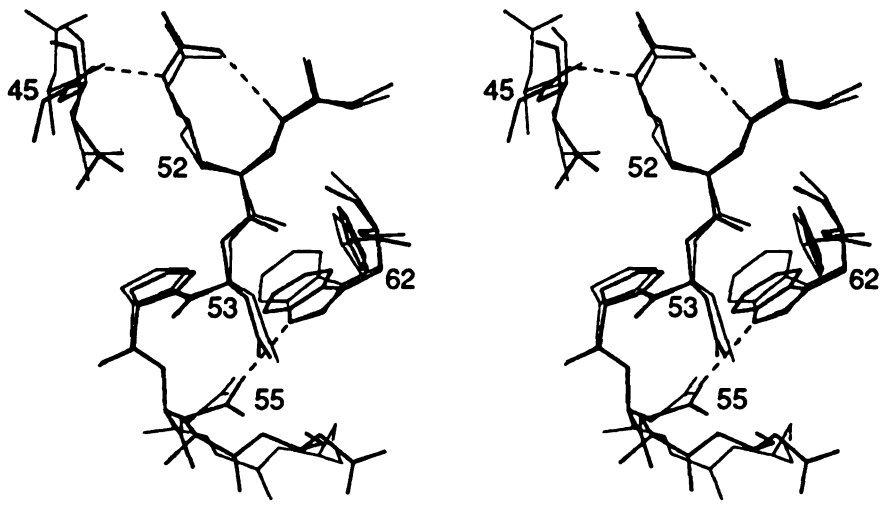


Figure 62. Stereoview comparing conformations of conserved residues Arg52, Asn53, and Asp55 in K4 (bold) and PTK1. Dashed lines indicate conserved hydrogen-bonding interactions.

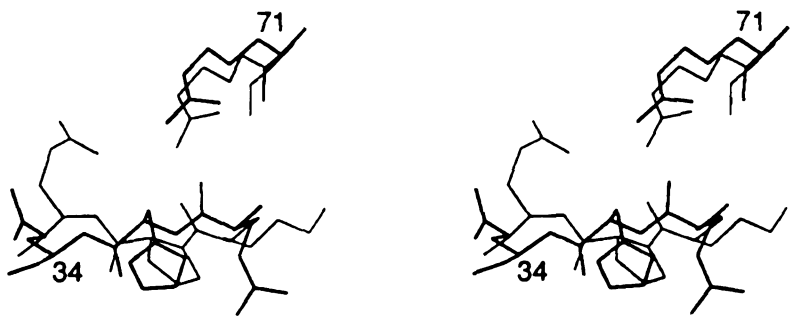
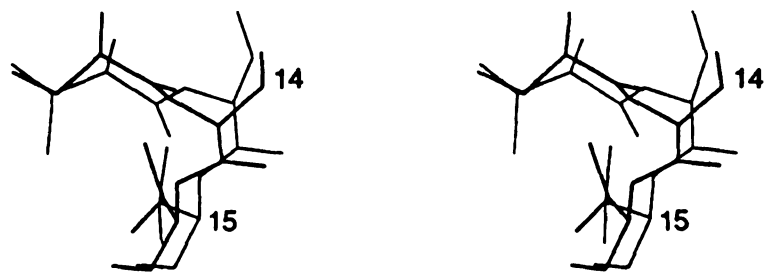


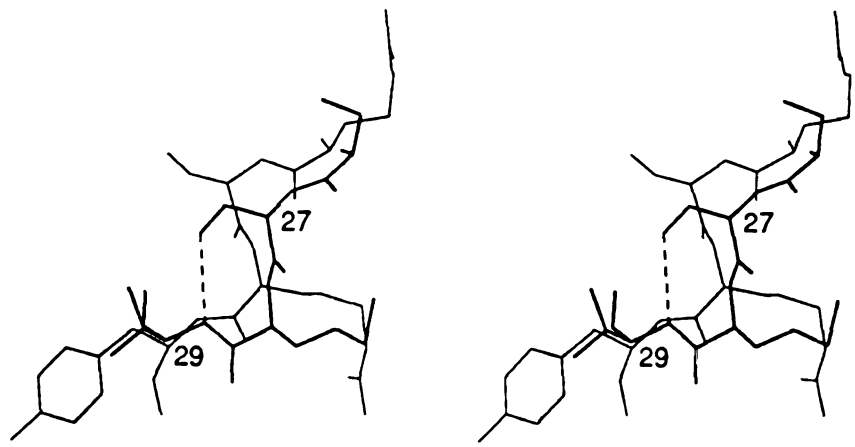
Figure 63. Stereoview comparing conformations of conserved Arg71 and homologous Gln/Glu34 residues in K4 (bold) and PTK1.

In some cases, larger differences exist between conserved residues. For example, the side chains of Ser14 and Ser27 have an approximately 90° difference in torsion angle between the two structures (Figure 64). In the first case no stabilizing interactions are observed in either structure; for the latter, the Ser side chain forms a hydrogen bond with the Thr29 main chain N atom in K4 but forms no similar interactions in PTK1. The 1.3 Å deviation between Arg10 conformations is due mainly to differences in the midsection of the side chain; the CZ atoms of both guanidinium groups are within 0.40 Å of one another (Figure 65). However, due to a difference in the rotational orientation of the guanidinium groups, the PTK1 side chain does not display the hydrogen bonding interactions observed in K4. Another large difference occurs at Glu73, for which there is an approximately 120° difference in torsion angle about the CA-CB bond (Figure 66). This single large difference in torsional rotation causes the K4 and PTK1 side chains to appear nearly as mirror images of one another with a 2.3 Å distance between respective CG atoms. Nevertheless, both side chains interact similarly with the hydroxyl oxygen of conserved Thr16. The observed difference in side chain conformation may be due to the close proximity of a bulky neighboring Leu70 side chain which may have a steric influence in the PTK1 structure, but which is replaced by the smaller Val residue in K4.

Some K4 surface side chains are replaced by analogous residues in PTK1 which might be expected to behave similarly. For instance, K4 Arg32 is replaced by a Lys residue, which is also positively charged. The 2.6 Å deviation between corresponding atoms (CB,CG,CD,NE/CE,CZ/NZ) reflects the fact that whereas the PTK1 Lys side chain adopts an extended conformation, the K4 Arg side chain is perturbed due to its involvement in the intermolecular interaction at the lysine-binding site. A higher degree of homology exists between the Gln34 side chain near the binding site of K4 and the Glu34 residue of PTK1; however,



(a)



(b)

Figure 64. Stereoviews showing approximately 90° torsional differences in the conformations of a) Ser14 and b) Ser27 in K4 (bold) and PTK1. Dashed line indicates a hydrogen-bonding interaction between Ser27 OG and Thr29 N atoms which is observed only in K4.

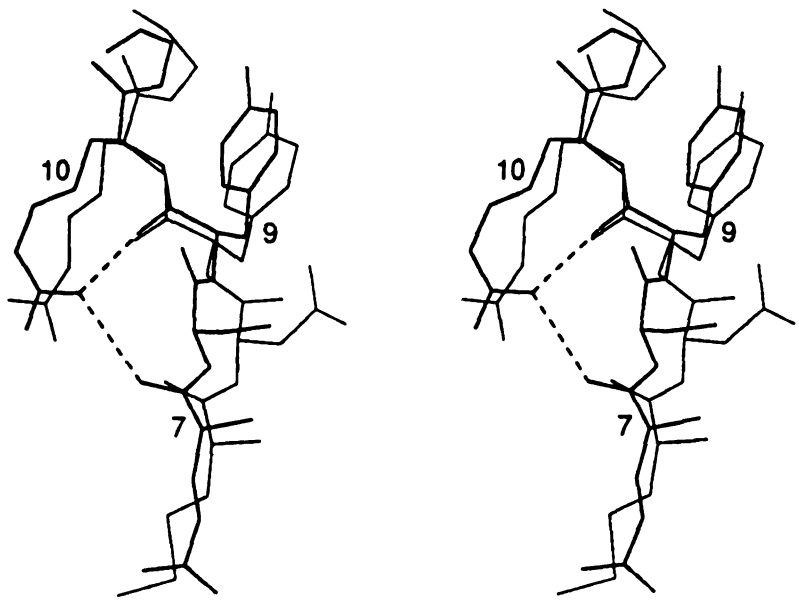


Figure 65. Stereoview comparing conformations of Arg10 in K4 (bold) and PTK1. Dashed lines indicate K4 hydrogen-bonding interactions not observed in PTK1.

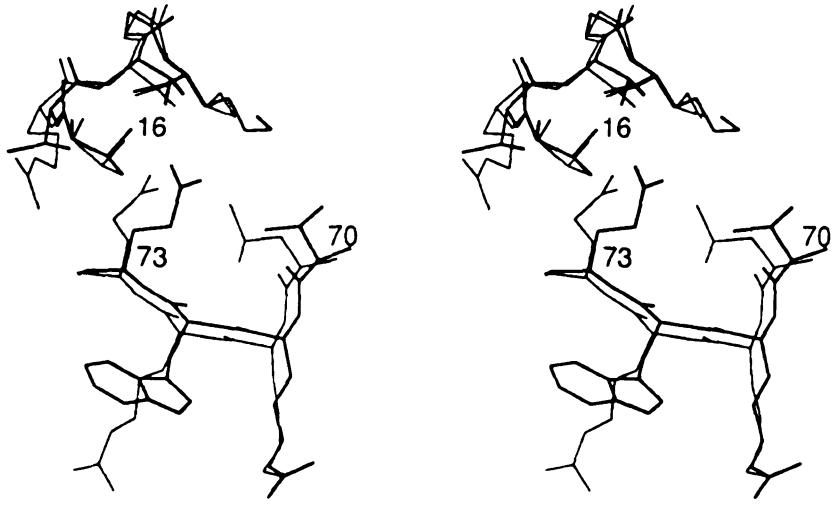


Figure 66. Stereoview comparing conformations of Glu73 side chains in K4 (bold) and PTK1. Both participate in similar hydrogen-bonding interactions with Thr16 side chain.

there is a very large average deviation of 5.8 Å between the two (Figure 63). The K4 Gln side chain has an extended conformation which directs it away from the lysine-binding site pocket, whereas the PTK1 Glu side chain extends in the opposite direction to form an ion pair with Arg71.

Although many residues of the K4 lysine-binding site are conserved in PTK1, the latter kringle does not share its binding properties. A comparison of the K4 lysine-binding site and the corresponding region of PTK1 is shown in Figure 67. It is obvious that the Trp/Arg substitution at position 72 and the Gln/Glu substitution at position 34 entirely change the nature of this site and destroy its ability to bind ligands. Though Asp55 and Arg71, essential residues of the negative and positive centers of the K4 binding site, are conserved in PTK1, both are occupied by intramolecular ion pair interactions. As previously mentioned, Arg71 forms an ion pair with Glu34. Similarly, non-conserved Arg72 extends across the potential binding site to form an ion pair with Asp55. Steric effects of the PTK1 Arg72/Asp55 ion pair formation in PTK1 may be responsible for the slight torsional differences observed between the Trp62 and Phe/Tyr64 side chains of the two kringles. Furthermore, the substitution of aromatic Trp72 and Tyr74 side chains with charged Arg and Glu residues in PTK1 eliminates the large exposed hydrophobic surface which provides favorable ligand interactions at the K4 binding site.

Surprisingly, there are only a few solvent sites which are common to both the K4 and PTK1 structures. All of these common solvents are bound to main chain protein atoms at the kringle surface and have high occupancies in the K4 structure. Two of these sites are adjacent, one being within hydrogen-bonding distance of residue 41 O and 45 N atoms, and the other within bonding distance of 42 O and 45 N atoms (Figure 68). These solvent molecules correspond to Wat3 and Wat18 of K4, which have occupancies of 1.00 and 0.78, respectively.

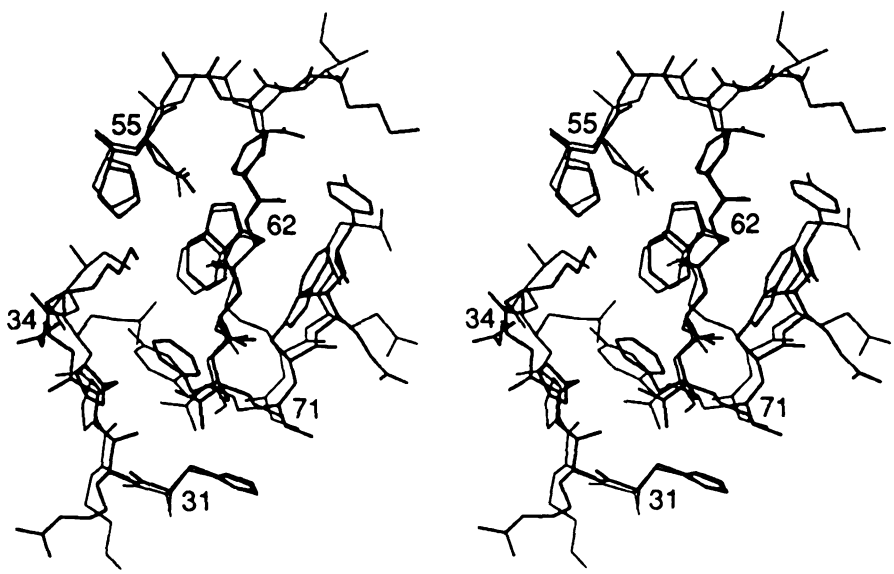


Figure 67. Comparison of K4 lysine-binding site (bold) and corresponding residues of PTK1.

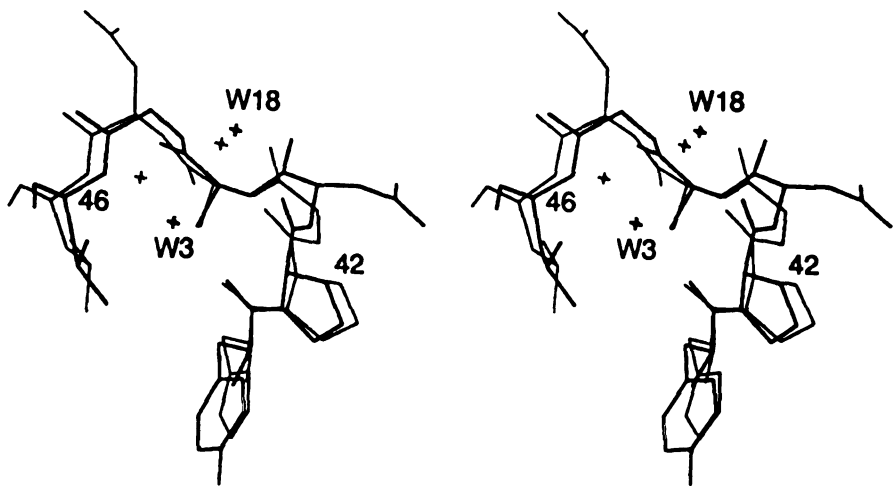


Figure 68. Stereoview showing similar solvent positions observed in K4 (bold) and PTK1. Solvent molecules, W3 and W18, labeled according to K4 quality factor ranking.

The third common solvent site occurs near the carbonyl oxygen of residue 62 and corresponds to K4 Wat13, which has an occupancy of 0.97 (Figure 69). The low agreement between ordered surface solvent sites of the two kringles is probably due to the presence of different surface side chains and the additional non-kringle protein structure of prothrombin fragment 1. Furthermore, the crystal packing and intermolecular contacts, which appear to significantly influence the ordered solvent of K4, also differ between the two structures.

More unexpected is the lack of agreement between the internal solvent positions. The single internal solvent molecule found in the hydrophobic core of K4 does not occur in PTK1, although an analogous cavity exists (Figure 70). The nearest solvent molecule found in PTK1 is 4.2 Å away and is in closer contact with Pro30 than Trp25. Conversely, a number of internal solvent sites of the PTK1 structure are not observed in K4, including two which occur adjacent to the inner disulfide bonds.

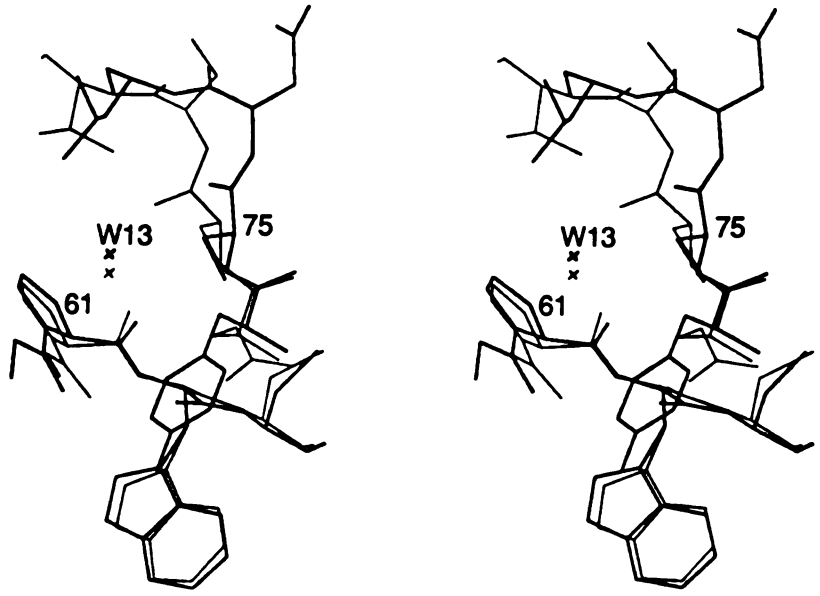


Figure 69. Stereoview comparing conserved solvent position, W13, in K4 (bold) and PTK1.

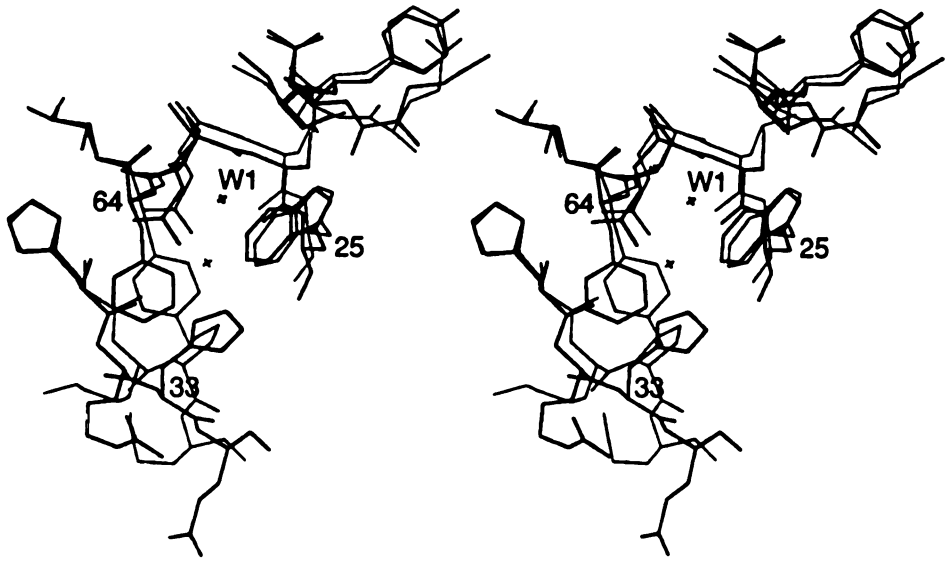


Figure 70. Stereoview comparing positions of K4 internal solvent molecule (W1) and nearest solvent site of PTK1 structure. K4 residues are shown in bold.

XIV. COMPARISON OF OBSERVED AND MODELED BINDING SITES

The K4 lysine-binding site model, which was previously proposed based primarily on the three-dimensional structure of PTK1 [45], shows general agreement with the actual observed binding site. A superposition of the modeled and observed binding sites is shown in Figure 71. It can be seen that residues which are conserved between K4 and PTK1, such as His31, Pro54, Pro61, and Trp62, as well as Phe64, which replaces a highly homologous Tyr residue, show no dramatic differences from the modeled conformations. The conserved Asp55, which comprises the crucial anionic center, agrees very well, with a difference of only 0.68 Å between modeled and observed CG centers. The Arg71 residue of the cationic center also agrees reasonably well, although the guanidinium group is directed across the binding site toward the anionic center in the modeled structure but extends in the opposite direction, away from the binding site, in the observed structure. This discrepancy is due in part to the intermolecular interactions involving Arg71 in the orthorhombic crystal structure, and in part to conformational adjustments and energy minimizations used in the modeling procedure to optimize an ACA ligand-binding interaction at the binding site. Finally, the side chain of conserved Glu73, although not directed toward the binding site in either case, also shows a somewhat larger deviation. This residue was previously shown to differ between K4 and the PTK1 structure from which it was modeled due to differences in steric influences of residues further removed from the binding site region.

As might be expected, there are generally larger differences between the observed and modeled conformations of residues not conserved between K4 and PTK1. Such is the case for Lys35 and Lys58, which are not conserved and

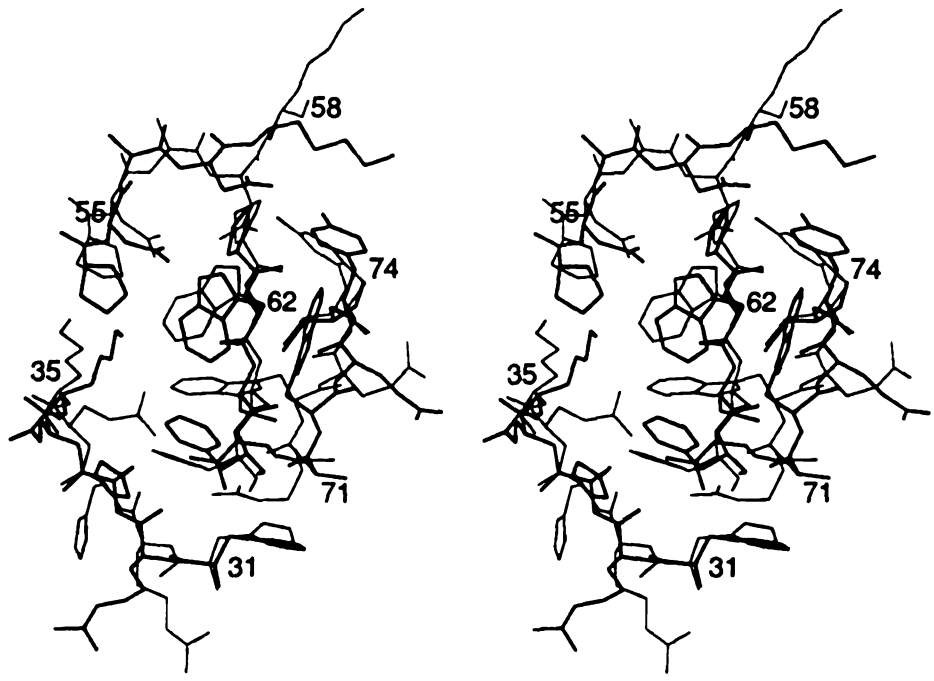


Figure 71. Comparison of observed (bold) and modeled K4 lysine-binding sites.

for which the corresponding PTK1 residues (Ile,Thr) provide guide coordinates to CB only. Both side chains were modeled in extended, energy minimized conformations; however, the observed conformations differ considerably. In the orthorhombic crystals both Lys side chains are perturbed by their participation in a sulfate-binding interaction occurring at the lysine-binding site: Lys35 approaches the Arg72 cationic center and Lys58 extends toward the analogous center of a neighboring protein molecule. Similarly, although the PTK1 Lys32 residue provides good guide coordinates for Arg32 of K4, this side chain also deviates considerably from the modeled position. Again, the K4 side chain is perturbed from the PTK1 Lys extended conformation due to its participation in an intermolecular ion pair. Thus, in all three cases, though the modeled side chains disagree considerably with the observed X-ray crystal structure, they may better approximate the free solution structure. An opposite situation arises for Gln34. In this case the observed side chain has an extended conformation which agrees poorly with the modeled counterpart, even though the PTK1 Glu34 residue provides excellent guide coordinates. Despite the high structural homology between Gln and Glu side chains, the negative charge carried by the latter is an important difference, as the PTK1 Glu side chain conformation is directed by an ion pair formation not common to both structures. For some residues, the differences between modeled and observed conformations appear to be caused predominantly by deviations of the K4 and PTK1 main chains. For instance, adjustments in the backbone near the PTK1 insertion at position 59 surely affect the ability to model the adjacent Asp57 side chain. Likewise, the difference in His33 conformations appears to result largely from the relative twist between the two main chain structures.

The most striking difference between the modeled and observed binding sites regards the position of Trp72. The corresponding PTK1 Arg72 residue,

which forms an ion pair across the binding site with Asp55, provides poor guide coordinates. In this case, the modeling was based on nmr observations and the known tendencies of hydrophobic indole rings. The strong interaction between Trp72 and bound ligands which was observed in nmr experiments suggested two possible modeling conformations for this side chain: one in which the indole ring is oriented parallel to the length of the binding site, and one in which it is directed across the binding site. Typically, the large hydrophobic indole ring of Trp side chains prefers a buried conformation, and, whereas the crosswise alternative protects one end of the Trp side chain in a cluster of aromatic residues, the parallel orientation exposes much more of the indole surface to the solvent. Thus, the former conformation was chosen during the modeling procedure. However, as can be seen in Figure 71, it is the unexpected surface-accessible conformation which is actually observed for this residue. Furthermore, it appears that energy minimization of the modeled site using this incorrect Trp72 conformation also resulted in unnecessary adjustments in the conformations of Trp62 and Phe64, which actually have only minor deviations from the corresponding PTK1 residues.

A more successful aspect of the modeling involves the conformation of the non-conserved Tyr74 side chain. Although the PTK1 Glu74 guide coordinates would have directed the Tyr side chain away from the binding site, NOE experiments indicated an interaction between the tyrosyl ring and a methine proton of Trp62. Therefore, torsional rotations of the side chain were used to position the modeled Tyr ring perpendicular to and within 4.6 Å of the Trp62 indole. This modeled conformation is seen to agree extremely well with the observed structure, differing only by a torsional rotation of approximately 40° about the CA-CB bond.

Thus, although it does not have similar binding properties, the PTK1

structure is generally found to be a very good model for the K4 lysine-binding site. Many of the essential residues which support the binding site are conserved between the two kringles, and their conformations, particularly those of side chains having important structural functions (Trp,Phe,His), are remarkably invariant. Furthermore, the strikingly accurate modeling achieved for the Tyr74 side chain demonstrates the power of combining both homologous crystallographic structural data and other types of experimental observations for the purposes of molecular modeling.

XV. INVESTIGATION INTO THE CAUSE OF RELATIVE INTENSITY CHANGES

As a final stage of this work on the determination of the orthorhombic K4 crystal structure, an investigation was made into the cause of the large relative intensity changes which have been observed in the orthorhombic crystals. These reproducible changes are puzzling in that they must reflect some non-random structural change occurring in the crystal lattice. Furthermore, they appear to be provoked not only by exposure to X-radiation, but also merely by resubmersing the crystal in the mother liquor once it has been removed therefrom.

To determine the structural changes underlying these observed intensity changes, a second set of intensity measurements was collected from the same crystal used in the original structure determination. The crystal, which had begun to show relative intensity changes during the original data collection, was first left idle for approximately one month. During this time, the axial intensity distributions were re-measured periodically to monitor the progress of the intensity changes. When it appeared that the diffraction pattern had stabilized, the second data collection was begun. At this point the crystal had already accumulated about 130 hours of prior X-ray exposure. The second data set was collected in the same manner as the first, using identical parameters (X-ray power, scan speed, scan width, etc.) and measuring data from the same octant of the limiting reflection sphere. However, due to the decreased diffracting power of the crystal, data were only collected to 36° (2.5 Å).

A scale factor of 1.27 and ΔB correction of 4 $Å^2$ had to be applied to the processed data for an optimal fit with the original measurements (Figure 72). Agreement between the two data sets was calculated as:

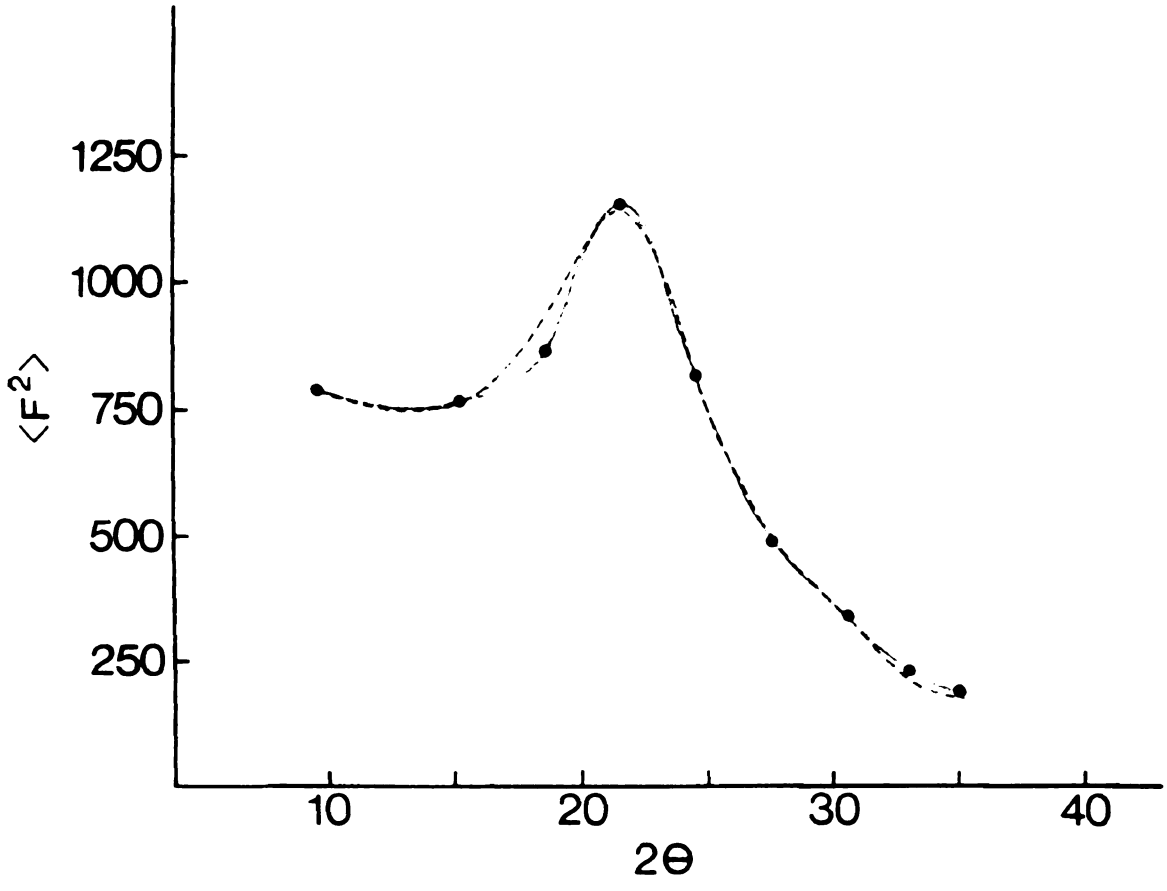


Figure 72. Fit of average structure factors from second orthorhombic K4 data collection, applying scale of 1.27 and ΔB correction of +4 (solid line), to those of original data collection (dashed line).

$$R = \Sigma(|F_n| - |F_o|)/\Sigma F_n .$$

The value of R varies from approximately 6% at low resolution to 11% at higher resolution, with an overall agreement of 9.5% for the complete 2.5 Å set.

A structure refinement was performed at 6-2.5 Å resolution with the second set of observed intensities, starting with the previously refined orthorhombic K4 coordinates (protein only). New solvent positions were determined independently. The refinement proceeded smoothly to an R-factor of 13.5 % with good geometry (Table 23); however, the resulting structure was virtually identical to that previously determined, including the positions of solvent molecules.

Since refinement using the complete second data second failed to indicate any structural differences to account for the observed changes in diffraction, the individual measurements having large discrepancies were considered separately. From the distribution of structure factor differences between the two data sets (Table 24), the set of 334 reflections having squared differences greater than 15 were chosen (11% of total reflections to 2.5 Å) for further consideration. The distribution of these 334 reflections according to scattering angle is shown in Figure 73, and reflections showing particularly large discrepancies are listed in Table 25. In an attempt to identify the regions of the orthorhombic unit cell giving rise to the observed intensity changes, a ($|F_o|$-$|F_c|$) difference electron density map was calculated using only these 334 observed reflections having large discrepancies. Phases and calculated structure factors were derived from the original refined structure. The difference map was calculated at a resolution range of 99 to 2.5 Å so as to include all the reflections having square differences greater than the cutoff value of 15; however, only 21 of the reflections have Bragg angles corresponding to resolutions lower than 6 Å (Figure 73).

Table 23. Final results of PROFFT least-squares refinement at 2.5 Å resolution for second set of orthorhombic data collected after the observation of characteristic relative intensity changes.

	Target	RMS Deviation
Distances (Å):		
Bond length	0.020	0.013
Bond angle	0.040	0.039
Planar 1-4	0.060	0.047
Planarity		
Dev from plane (Å)	0.020	0.011
Chirality		
Chiral volume ($Å^3$)	0.200	0.201
Nonbonded Contacts:		
Single torsion	0.550	0.184
Multiple torsion	0.550	0.238
X..Y H-bond	0.550	0.302
Torsion Angle (deg)		
Planar	3.0	2.0
Staggered	15.0	17.8
Orthonormal	20.0	16.3
Thermal Restraints ($Å^2$):		
Main chain bond	3.00	2.30
angle	4.00	3.34
Side chain bond	4.00	3.55
angle	4.00	4.36
Average $\|F_o-F_c\|$		20.6
Structure Factor Weight		
σ_a		16.0
σ_b		-70.0
R (%)		13.4
<B> ($Å^2$)		20.4

Table 24. Distribution of magnitude of difference between structure factors calculated from two sets of orthorhombic K4 intensity data measured before and after characteristic relative intensity changes.

Range $(Diff)^2$	Number of Measurements
0 - 1	1130
1 - 5	901
5 - 10	430
10 - 15	161
15 - 20	105
20 - 25	71
25 - 30	39
30 - 35	39
35 - 40	22
40 - 45	16
45 - 50	8
50 - 100	23
100+	11

$Diff^2 = (|F_1| - |F_2|)^2$

Table 25. Reflections having largest structure factor discrepancies between first and second K4 data collections ($Diff^2 > 100$, where $Diff^2 = (|F_1| - |F_2|)^2$).

	h,k,l	2θ (deg)	Resolution (Å)	$Diff^2$
1	7 0 14	32.0	2.8	636
2	5 15 5	32.0	2.8	506
3	7 0 15	33.5	2.7	474
4	6 12 11	34.1	2.6	463
5	1 11 0	20.1	4.4	207
6	7 5 0	21.4	4.2	148
7	2 4 8	17.0	5.2	130
8	0 10 0	18.1	4.9	117
9	8 2 0	22.4	4.0	111
10	11 8 5	35.3	2.5	109
11	2 3 9	17.9	5.0	106

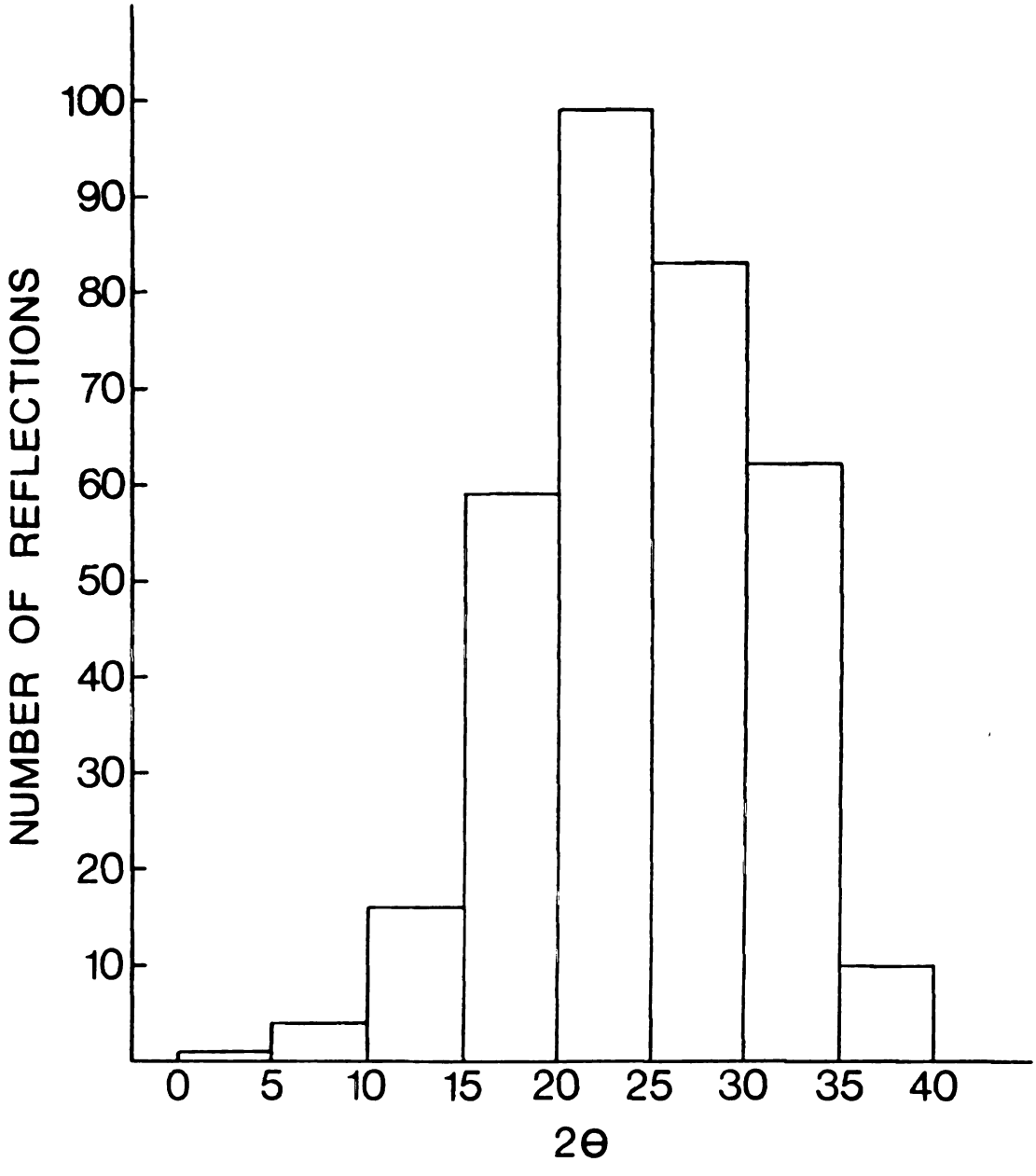


Figure 73. Distribution of scattering angles for set of 334 reflections having very large discrepancies between original and second orthorhombic K4 data collections ($(|F_1|-|F_2|)^2 > 100$).

The observed map is generally featureless, showing primarily fragments of positive difference density in the empty solvent regions between K4 molecules. However, two larger regions of difference density are observed. Some unusually shaped density is found near the last ordered residue of the amino terminal (Figure 74). A second region occurs at the lysine-binding site between the Trp72 indole ring and the side chains of symmetry mate Arg32 and Met28 residues (Figure 75). The latter region is the only difference density observed in the immediate vicinity of well-ordered protein structure. A similar difference density map calculated using the remaining 2622 reflections at the same resolution range displays density only in the intermolecular solvent cavities, none of which coincides with that observed using the set of 334 highly discrepant reflections. Thus, these observations suggest that the intensity changes of orthorhombic diffraction pattern may be related to structural changes occurring at the poorly ordered amino terminal tail region, which was not observed at all in the K4 structure determination, and at the site of the ligand-like kringle-kringle interaction. A fuller understanding of these changes will require further investigation of these preliminary results.

Figure 74. Stereoview showing electron difference density observed near amino terminal of K4 peptide chain when map is calculated using only 334 reflections having very large discrepancies between original and second intensity data sets.

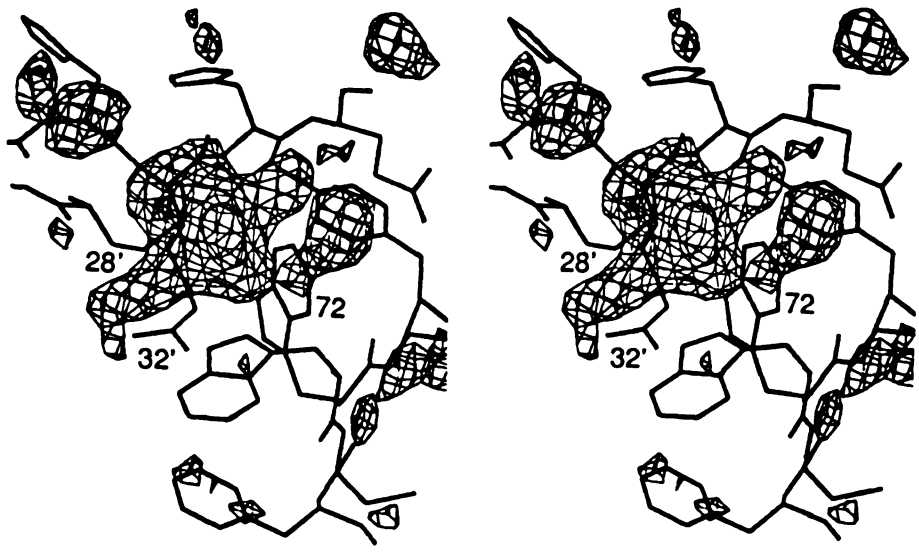


Figure 75. Stereoview showing electron difference density observed in vicinity of ligand-like intermolecular interaction when difference map is calculated using only 334 reflections having very large discrepancies between original and second intensity data sets.

APPENDICES

APPENDIX A

Atomic coordinates and crystallographic parameters of orthorhombic plasminogen K4 crystal structure.

Seq. No.	Res. Type	Atom	Coordinates (Å) X	Y	Z	B (Å^2)	Occupancy
c	ASP	N	-2.60960	5.00590	20.70820	39.44000	1.00000
c	ASP	CA	-3.96030	5.36460	21.21420	37.01870	1.00000
c	ASP	C	-4.82330	5.87650	20.06620	34.01270	1.00000
c	ASP	O	-5.25320	5.13260	19.16120	36.64960	1.00000
1	CYS	N	-5.08620	7.17910	20.13850	31.06120	1.00000
1	CYS	CA	-5.91320	7.80140	19.08820	28.42870	1.00000
1	CYS	C	-5.43360	9.23300	18.82540	21.04750	1.00000
1	CYS	O	-4.64700	9.81020	19.58930	23.79590	1.00000
1	CYS	CB	-7.40010	7.84170	19.46210	29.37410	1.00000
1	CYS	SG	-7.67380	8.69730	21.04680	35.36780	1.00000
2	TYR	N	-5.99120	9.78500	17.78000	23.51860	1.00000
2	TYR	CA	-5.61910	11.21650	17.53920	20.95490	1.00000
2	TYR	C	-6.90010	12.05200	17.49130	20.02810	1.00000
2	TYR	O	-8.00040	11.56600	17.10570	19.37630	1.00000
2	TYR	CB	-4.84870	11.37270	16.22390	17.85780	1.00000
2	TYR	CG	-5.60630	11.01830	14.95920	17.59420	1.00000
2	TYR	CD1	-5.62590	9.68930	14.46600	15.64240	1.00000
2	TYR	CE1	-6.27680	9.41780	13.26630	16.37740	1.00000
2	TYR	CZ	-6.94110	10.38900	12.56130	13.97010	1.00000
2	TYR	CE2	-6.94680	11.69640	13.05260	13.93040	1.00000
2	TYR	CD2	-6.25320	11.97870	14.22090	12.43000	1.00000
2	TYR	OH	-7.59830	10.20400	11.38000	20.51060	1.00000
3	HIS	N	-6.74660	13.32200	17.84770	22.59660	1.00000
3	HIS	CA	-7.82330	14.32210	17.75970	22.58150	1.00000
3	HIS	C	-7.66680	15.18520	16.49710	22.74330	1.00000
3	HIS	O	-6.61830	15.46040	15.91720	25.77080	1.00000
3	HIS	CB	-7.92600	15.27850	18.94160	26.44550	1.00000
3	HIS	CG	-7.89930	14.52780	20.23270	26.07700	1.00000
3	HIS	ND1	-9.01980	14.03100	20.83400	30.09910	1.00000
3	HIS	CE1	-8.65640	13.45440	21.96050	31.88160	1.00000
3	HIS	NE2	-7.34770	13.56810	22.11580	34.49220	1.00000
3	HIS	CD2	-6.85250	14.24940	21.02990	28.36130	1.00000
4	GLY	N	-8.77480	15.64590	15.97010	23.57470	1.00000
4	GLY	CA	-8.91950	16.44790	14.77520	25.73750	1.00000
4	GLY	C	-8.22760	15.86660	13.55340	22.86140	1.00000
4	GLY	O	-8.61200	14.79150	13.05650	25.79960	1.00000
5	ASP	N	-7.22990	16.58520	13.06130	20.29500	1.00000
5	ASP	CA	-6.48320	16.08770	11.87160	16.47700	1.00000
5	ASP	C	-5.18020	15.41530	12.32650	11.05580	1.00000
5	ASP	O	-4.36220	15.05680	11.49310	11.62980	1.00000
5	ASP	CB	-6.34690	17.19250	10.81850	15.78720	1.00000
5	ASP	CG	-5.53110	18.34360	11.37960	14.09670	1.00000
5	ASP	OD1	-5.35060	19.29220	10.62430	17.20790	1.00000
5	ASP	OD2	-5.06590	18.26610	12.51850	11.68460	1.00000

6	GLY	N	-4.97920	15.28400	13.61490	9.63040	1.00000
6	GLY	CA	-3.85030	14.68320	14.24510	13.04590	1.00000
6	GLY	C	-2.52720	15.42980	14.09460	17.67000	1.00000
6	GLY	O	-1.49240	14.77870	14.28100	13.50620	1.00000
7	GLN	N	-2.56260	16.71620	13.77930	16.18800	1.00000
7	GLN	CA	-1.31060	17.49870	13.67680	15.03470	1.00000
7	GLN	C	-0.77790	17.60790	15.09670	10.33340	1.00000
7	GLN	O	0.40880	17.87790	15.33020	14.28420	1.00000
7	GLN	CB	-1.47280	18.86710	13.02560	17.30170	1.00000
7	GLN	CG	-1.66130	18.90940	11.51060	13.37330	1.00000
7	GLN	CD	-1.68430	20.39950	11.07570	11.04240	1.00000
7	GLN	OE1	-0.58130	20.90570	10.97350	13.02340	1.00000
7	GLN	NE2	-2.83030	20.94970	10.87470	12.24860	1.00000
8	SER	N	-1.52910	17.39710	16.13290	10.24800	1.00000
8	SER	CA	-1.08980	17.48110	17.50620	12.23250	1.00000
8	SER	C	-0.71250	16.12800	18.07200	13.75170	1.00000
8	SER	O	-0.27530	16.06000	19.21700	17.66110	1.00000
8	SER	CB	-2.15240	18.14180	18.42020	17.88670	1.00000
8	SER	OG	-3.35650	17.37720	18.39800	20.54300	1.00000
9	TYR	N	-0.90010	15.09080	17.27270	17.86730	1.00000
9	TYR	CA	-0.58040	13.69120	17.74840	16.46040	1.00000
9	TYR	C	0.89080	13.51600	18.06720	15.88040	1.00000
9	TYR	O	1.68400	13.74570	17.15860	18.05230	1.00000
9	TYR	CB	-1.08640	12.64900	16.73930	17.85800	1.00000
9	TYR	CG	-0.74540	11.22510	17.16510	14.92670	1.00000
9	TYR	CD1	-1.33780	10.56100	18.24660	13.33350	1.00000
9	TYR	CD2	0.26290	10.59190	16.44050	16.02040	1.00000
9	TYR	CE1	-0.94410	9.24750	18.54310	15.79380	1.00000
9	TYR	CE2	0.69770	9.30270	16.75630	13.17220	1.00000
9	TYR	CZ	0.06980	8.64010	17.80070	18.22460	1.00000
9	TYR	OH	0.56160	7.35330	18.05710	19.25690	1.00000
10	ARG	N	1.21900	13.13370	19.28340	13.51460	1.00000
10	ARG	CA	2.61620	12.93940	19.63000	15.54950	1.00000
10	ARG	C	2.92500	11.53460	20.13260	18.63550	1.00000
10	ARG	O	3.92950	11.32640	20.81450	20.13970	1.00000
10	ARG	CB	3.07740	14.07800	20.55170	21.05560	1.00000
10	ARG	CG	3.13910	15.43530	19.84390	16.58900	1.00000
10	ARG	CD	4.31360	15.63070	18.96750	18.37720	1.00000
10	ARG	NE	4.40640	16.95350	18.33200	16.96580	1.00000
10	ARG	CZ	3.56790	17.39030	17.42680	16.02260	1.00000
10	ARG	NH1	2.64610	16.58320	16.92400	14.93750	1.00000
10	ARG	NH2	3.56350	18.69200	17.12930	17.93590	1.00000
11	GLY	N	2.12010	10.57350	19.76190	19.02930	1.00000
11	GLY	CA	2.24010	9.17960	20.13000	22.09100	1.00000
11	GLY	C	3.42050	8.43660	19.55790	21.22440	1.00000
11	GLY	O	4.29120	8.90500	18.82570	20.69700	1.00000
12	ALA	N	3.41360	7.15670	19.92040	20.00350	1.00000
12	ALA	CA	4.44380	6.18340	19.52080	19.02900	1.00000
12	ALA	C	4.01560	5.28560	18.36590	20.22260	1.00000
12	ALA	O	4.80650	4.39030	17.95910	17.73360	1.00000
12	ALA	CB	4.73270	5.43530	20.85370	21.35670	1.00000
13	SER	N	2.80330	5.41530	17.81110	19.92990	1.00000
13	SER	CA	2.38350	4.54810	16.69510	17.49400	1.00000
13	SER	C	3.33380	4.68110	15.50750	20.28460	1.00000

13	SER	O	3.82310	5.79620	15.26560	15.59020	1.00000
13	SER	CB	0.94220	4.76740	16.22760	24.84850	1.00000
13	SER	OG	0.10620	3.99570	17.08900	33.31560	1.00000
14	SER	N	3.62160	3.53230	14.84380	14.57160	1.00000
14	SER	CA	4.54620	3.62090	13.73170	11.76980	1.00000
14	SER	C	4.28750	2.52960	12.69110	14.15020	1.00000
14	SER	O	5.30250	2.12180	12.09770	13.59220	1.00000
14	SER	CB	5.98470	3.52030	14.18880	14.67690	1.00000
14	SER	OG	6.29630	2.34600	14.89080	18.10470	1.00000
15	THR	N	3.00050	2.19440	12.56890	13.00840	1.00000
15	THR	CA	2.64990	1.18640	11.53870	13.11610	1.00000
15	THR	C	1.53290	1.77760	10.70820	13.42020	1.00000
15	THR	O	0.66540	2.45050	11.26580	16.43490	1.00000
15	THR	CB	2.52920	-0.29740	12.05350	17.55060	1.00000
15	THR	OG1	1.47700	-1.12060	11.44590	17.93550	1.00000
15	THR	CG2	2.61050	-0.42360	13.54420	13.77540	1.00000
16	THR	N	1.64480	1.53750	9.41810	12.36260	1.00000
16	THR	CA	0.65150	2.05230	8.46060	12.16110	1.00000
16	THR	C	-0.59950	1.14370	8.52270	18.41360	1.00000
16	THR	O	-0.68790	0.07210	9.13770	15.97260	1.00000
16	THR	CB	1.24450	2.08830	7.00470	13.53530	1.00000
16	THR	OG1	1.24030	0.69170	6.49950	14.65210	1.00000
16	THR	CG2	2.67740	2.61020	6.91460	11.48240	1.00000
17	THR	N	-1.64120	1.59650	7.84050	14.69090	1.00000
17	THR	CA	-2.96040	0.98540	7.65100	16.14960	1.00000
17	THR	C	-2.88430	-0.43120	7.07720	18.34280	1.00000
17	THR	O	-3.82650	-1.23400	7.34200	19.12300	1.00000
17	THR	CB	-3.58490	2.15200	6.79140	17.88530	1.00000
17	THR	OG1	-4.40220	3.00640	7.71960	18.58360	1.00000
17	THR	CG2	-3.95420	1.80340	5.41970	12.51520	1.00000
18	THR	N	-1.85640	-0.79910	6.37130	15.80590	1.00000
18	THR	CA	-1.61580	-2.14080	5.77060	16.37400	1.00000
18	THR	C	-0.57990	-2.98530	6.51320	13.25580	1.00000
18	THR	O	-0.11900	-4.02620	6.01580	15.35810	1.00000
18	THR	CB	-1.21270	-2.06220	4.26900	15.14690	1.00000
18	THR	OG1	0.17490	-1.56230	4.27810	16.72810	1.00000
18	THR	CG2	-2.13480	-1.18840	3.40980	17.49040	1.00000
19	GLY	N	-0.17870	-2.56890	7.67620	14.92410	1.00000
19	GLY	CA	0.79280	-3.21740	8.57680	16.00880	1.00000
19	GLY	C	2.26050	-2.96830	8.27770	17.26360	1.00000
19	GLY	O	3.10340	-3.55780	8.99110	16.00320	1.00000
20	LYS	N	2.60210	-2.08980	7.32390	14.54220	1.00000
20	LYS	CA	4.00130	-1.75550	7.00600	13.73360	1.00000
20	LYS	C	4.64240	-0.95210	8.11720	13.72090	1.00000
20	LYS	O	3.91860	-0.21900	8.83880	13.86710	1.00000
20	LYS	CB	4.13430	-1.03070	5.67200	15.13420	1.00000
20	LYS	CG	3.60580	-1.83390	4.51080	15.16760	1.00000
20	LYS	CD	3.73870	-1.01390	3.24470	16.99810	1.00000
20	LYS	CE	3.39400	-1.59770	1.92170	19.57120	1.00000
20	LYS	NZ	1.95320	-1.96090	1.91520	20.21440	1.00000
21	LYS	N	5.96710	-1.00720	8.25240	10.35960	1.00000
21	LYS	CA	6.61300	-0.25240	9.33430	15.02520	1.00000
21	LYS	C	7.00120	1.13610	8.78460	11.76510	1.00000
21	LYS	O	7.47320	1.14560	7.63250	14.34990	1.00000

21	LYS	CB	7.86730	-1.03400	9.73530	20.14700	1.00000
21	LYS	CG	8.95030	-0.19150	10.42160	23.52850	1.00000
21	LYS	CD	9.96620	-1.03120	11.16430	28.64450	1.00000
21	LYS	CE	10.70210	-0.17810	12.18860	36.25110	1.00000
21	LYS	NZ	11.69700	0.62550	11.40960	37.76410	1.00000
22	CYS	N	6.83520	2.16520	9.54010	12.95220	1.00000
22	CYS	CA	7.22440	3.51680	9.07830	13.80370	1.00000
22	CYS	C	8.75090	3.71250	9.07030	13.60340	1.00000
22	CYS	O	9.48850	3.25230	9.96490	13.12200	1.00000
22	CYS	CB	6.72680	4.56460	10.14050	12.20350	1.00000
22	CYS	SG	4.91820	4.75180	10.16230	11.07200	1.00000
23	GLN	N	9.16400	4.43170	8.04720	13.42120	1.00000
23	GLN	CA	10.55560	4.88550	7.82790	13.38300	1.00000
23	GLN	C	10.81360	6.05320	8.77220	15.69940	1.00000
23	GLN	O	9.89620	6.85760	9.09670	13.20980	1.00000
23	GLN	CB	10.69580	5.31360	6.36750	12.14200	1.00000
23	GLN	CG	12.05890	5.85240	6.02830	13.34080	1.00000
23	GLN	CD	12.23500	6.43150	4.66510	8.51330	1.00000
23	GLN	OE1	11.85790	5.96910	3.60690	11.76330	1.00000
23	GLN	NE2	12.93300	7.54940	4.67160	12.51700	1.00000
24	SER	N	12.04130	6.23460	9.25680	14.09030	1.00000
24	SER	CA	12.46070	7.35360	10.12850	18.13800	1.00000
24	SER	C	12.47300	8.66280	9.30230	13.07500	1.00000
24	SER	O	12.90360	8.66890	8.15540	10.65080	1.00000
24	SER	CB	13.87340	7.27300	10.73270	18.09750	1.00000
24	SER	OG	13.85240	6.05900	11.49610	29.57710	1.00000
25	TRP	N	11.95840	9.69830	9.93990	12.75000	1.00000
25	TRP	CA	11.86040	11.03730	9.25020	12.33310	1.00000
25	TRP	C	13.23740	11.57990	8.90300	12.37460	1.00000
25	TRP	O	13.37880	12.25520	7.85130	14.44400	1.00000
25	TRP	CB	11.06470	11.94310	10.17620	11.12440	1.00000
25	TRP	CG	9.67080	11.50390	10.47260	12.34430	1.00000
25	TRP	CD1	9.21180	10.97900	11.65800	9.60350	1.00000
25	TRP	CD2	8.55270	11.62180	9.60170	8.79140	1.00000
25	TRP	NE1	7.87100	10.71130	11.50980	11.63120	1.00000
25	TRP	CE2	7.42680	11.14850	10.29010	10.94620	1.00000
25	TRP	CE3	8.39030	12.11390	8.29970	10.59190	1.00000
25	TRP	CZ2	6.16150	11.09590	9.71080	9.78210	1.00000
25	TRP	CZ3	7.15010	12.09360	7.72460	11.31010	1.00000
25	TRP	CH2	6.04540	11.56950	8.39770	12.31050	1.00000
26	SER	N	14.22060	11.24800	9.71190	15.85120	1.00000
26	SER	CA	15.58880	11.74640	9.39310	22.03010	1.00000
26	SER	C	16.32270	10.87060	8.36820	20.30440	1.00000
26	SER	O	17.32820	11.32380	7.83050	24.02910	1.00000
26	SER	CB	16.41450	11.98840	10.64260	23.99700	1.00000
26	SER	OG	16.64830	10.63720	11.08030	34.19180	1.00000
27	SER	N	15.86770	9.68580	8.00290	17.94130	1.00000
27	SER	CA	16.54770	8.89560	6.98400	16.09620	1.00000
27	SER	C	16.03480	9.18880	5.58230	17.76490	1.00000
27	SER	O	14.85120	9.52130	5.33660	16.65910	1.00000
27	SER	CB	16.34270	7.40960	7.38800	21.22670	1.00000
27	SER	OG	16.81890	6.61620	6.33290	23.97470	1.00000
28	MET	N	16.90700	9.06570	4.57340	16.96970	1.00000
28	MET	CA	16.58050	9.26340	3.16960	15.58530	1.00000

28	MET	C	16.46530	7.91310	2.45910	12.23670	1.00000
28	MET	O	16.48030	7.95350	1.21550	13.42210	1.00000
28	MET	CB	17.50590	10.17820	2.38410	15.35900	1.00000
28	MET	CG	17.70620	11.52980	2.97840	13.80920	1.00000
28	MET	SD	16.18080	12.49610	2.67970	14.62400	1.00000
28	MET	CE	15.59030	12.70440	4.36130	10.10080	1.00000
29	THR	N	16.35800	6.89720	3.28480	17.13850	1.00000
29	THR	CA	16.24660	5.54680	2.64880	16.99270	1.00000
29	THR	C	15.15990	4.69830	3.30380	13.10750	1.00000
29	THR	O	15.03240	4.70390	4.53570	12.41770	1.00000
29	THR	CB	17.66110	4.86000	2.83300	17.74420	1.00000
29	THR	OG1	17.70280	3.77530	1.85720	26.19160	1.00000
29	THR	CG2	17.88580	4.28790	4.22370	22.63000	1.00000
30	PRO	N	14.35030	3.97910	2.56010	11.63490	1.00000
30	PRO	CA	14.40020	3.82350	1.14220	12.74340	1.00000
30	PRO	C	13.76760	4.98940	0.39250	14.83420	1.00000
30	PRO	O	13.81140	4.95940	-0.83830	15.07750	1.00000
30	PRO	CB	13.44110	2.67090	0.82670	16.87020	1.00000
30	PRO	CG	12.79690	2.25780	2.08640	17.10190	1.00000
30	PRO	CD	13.35240	3.10160	3.19250	15.60960	1.00000
31	HIS	N	13.15790	5.91510	1.11310	14.98630	1.00000
31	HIS	CA	12.53770	7.02780	0.29350	13.64850	1.00000
31	HIS	C	13.22850	8.36360	0.66070	9.76950	1.00000
31	HIS	O	13.28180	8.58740	1.85380	10.01640	1.00000
31	HIS	CB	11.01510	7.13970	0.55160	12.60080	1.00000
31	HIS	CG	10.36130	5.78190	0.27120	13.13200	1.00000
31	HIS	ND1	9.93790	4.95560	1.27930	12.05850	1.00000
31	HIS	CD2	10.15260	5.15170	-0.90460	10.97320	1.00000
31	HIS	CE1	9.42400	3.87810	0.71810	8.37090	1.00000
31	HIS	NE2	9.54870	3.94240	-0.54300	11.45010	1.00000
32	ARG	N	13.59120	9.09780	-0.39750	11.44620	1.00000
32	ARG	CA	14.22280	10.42260	-0.00890	12.12040	1.00000
32	ARG	C	13.01560	11.36320	0.15900	13.65320	1.00000
32	ARG	O	11.98450	11.09330	-0.49170	14.02450	1.00000
32	ARG	CB	15.14650	10.98090	-1.05930	14.47490	1.00000
32	ARG	CG	15.70240	10.23310	-2.21470	16.49670	1.00000
32	ARG	CD	16.66840	11.01610	-3.05280	14.49550	1.00000
32	ARG	NE	17.60980	11.78630	-2.32330	18.61810	1.00000
32	ARG	CZ	18.74250	11.48450	-1.67630	19.81650	1.00000
32	ARG	NH1	19.21080	10.23590	-1.61200	13.03100	1.00000
32	ARG	NH2	19.51620	12.46070	-1.14730	12.29410	1.00000
33	HIS	N	13.13550	12.38630	0.96780	10.61420	1.00000
33	HIS	CA	12.04430	13.30920	1.18280	14.28330	1.00000
33	HIS	C	12.50490	14.60600	1.88800	15.32020	1.00000
33	HIS	O	13.66310	14.71670	2.33020	11.77690	1.00000
33	HIS	CB	10.89370	12.60310	1.95790	13.40320	1.00000
33	HIS	CG	11.38390	12.30230	3.34500	13.72860	1.00000
33	HIS	ND1	11.31170	13.22340	4.36260	11.83110	1.00000
33	HIS	CD2	11.94620	11.20690	3.84200	10.44740	1.00000
33	HIS	CE1	11.79170	12.69580	5.47290	14.47800	1.00000
33	HIS	NE2	12.17920	11.46840	5.17140	16.16700	1.00000
34	GLN	N	11.55160	15.53750	1.91010	10.53720	1.00000
34	GLN	CA	11.84100	16.82840	2.58580	17.10560	1.00000
34	GLN	C	11.05050	17.06690	3.85070	17.48650	1.00000

34	GLN	O	11.01250	18.26040	4.22230	15.17810	1.00000
34	GLN	CB	11.64990	18.01100	1.60770	10.92750	1.00000
34	GLN	CG	12.62870	17.75710	0.44070	19.47760	1.00000
34	GLN	CD	12.58990	18.84930	-0.57360	25.53290	1.00000
34	GLN	OE1	13.55880	19.54700	-0.73650	28.34660	1.00000
34	GLN	NE2	11.43560	18.95570	-1.23270	30.03060	1.00000
35	LYS	N	10.46410	16.01290	4.40470	16.52970	1.00000
35	LYS	CA	9.67930	16.26490	5.65790	12.54610	1.00000
35	LYS	C	10.57350	15.89470	6.82310	14.83030	1.00000
35	LYS	O	10.43970	14.92520	7.56550	17.39640	1.00000
35	LYS	CB	8.47380	15.31590	5.64940	14.68480	1.00000
35	LYS	CG	7.23480	15.94700	4.99850	16.33100	1.00000
35	LYS	CD	7.25660	15.98780	3.48250	14.49650	1.00000
35	LYS	CE	6.05730	16.81460	2.98220	13.60380	1.00000
35	LYS	NZ	6.28540	16.91630	1.51120	17.71470	1.00000
37	THR	N	11.60160	16.71910	7.01450	17.90770	1.00000
37	THR	CA	12.59050	16.51160	8.08280	16.60080	1.00000
37	THR	C	12.34050	17.53200	9.18340	16.50850	1.00000
37	THR	O	11.78410	18.60570	8.93470	18.19550	1.00000
37	THR	CB	14.00970	16.66120	7.42160	20.73270	1.00000
37	THR	OG1	14.03700	18.08270	6.98710	18.58920	1.00000
37	THR	CG2	14.25370	15.82540	6.14630	13.75040	1.00000
38	PRO	N	12.82780	17.16530	10.35420	22.33080	1.00000
38	PRO	CA	12.70120	18.03740	11.52640	23.37130	1.00000
38	PRO	C	13.26560	19.41200	11.22940	26.90280	1.00000
38	PRO	O	12.78200	20.42350	11.78970	28.01710	1.00000
38	PRO	CB	13.31890	17.26310	12.65520	24.05430	1.00000
38	PRO	CG	13.73650	15.93520	12.13580	22.95290	1.00000
38	PRO	CD	13.45560	15.86940	10.67430	20.31520	1.00000
39	ALA	N	14.21640	19.49930	10.29400	28.14320	1.00000
39	ALA	CA	14.76730	20.83090	9.95720	26.31530	1.00000
39	ALA	C	13.76740	21.67130	9.18970	28.44780	1.00000
39	ALA	O	13.66050	22.89730	9.48010	31.70050	1.00000
39	ALA	CB	16.13870	20.67480	9.30090	31.25860	1.00000
40	ASN	N	13.00660	21.11480	8.26820	24.63710	1.00000
40	ASN	ND2	13.31890	20.06120	4.50070	33.42000	1.00000
40	ASN	OD1	13.60200	22.22550	5.05110	35.03130	1.00000
40	ASN	CG	12.97580	21.15040	5.17640	33.01360	1.00000
40	ASN	CB	11.76610	21.03070	6.09920	24.70070	1.00000
40	ASN	CA	12.00660	21.74430	7.43610	22.77740	1.00000
40	ASN	C	10.66020	21.92220	8.15880	19.54140	1.00000
40	ASN	O	9.93460	22.88790	7.82170	18.81350	1.00000
41	TYR	N	10.39730	21.03620	9.10020	17.49650	1.00000
41	TYR	CA	9.12750	21.15460	9.87080	14.75820	1.00000
41	TYR	C	9.43020	21.11220	11.34400	17.54330	1.00000
41	TYR	O	9.15930	20.12150	12.04080	17.41090	1.00000
41	TYR	CB	8.23570	19.96550	9.48980	13.59450	1.00000
41	TYR	CG	7.73820	20.03370	8.07600	11.07120	1.00000
41	TYR	CD1	6.47000	20.53240	7.75710	14.09380	1.00000
41	TYR	CD2	8.57710	19.67540	7.06360	13.29680	1.00000
41	TYR	CE1	6.00240	20.61440	6.47060	12.32900	1.00000
41	TYR	CE2	8.15180	19.75970	5.74440	16.32410	1.00000
41	TYR	CZ	6.87420	20.22090	5.45000	17.47020	1.00000
41	TYR	OH	6.57970	20.25470	4.11480	18.25490	1.00000

42	PRO	N	9.99550	22.20180	11.89290	17.50050	1.00000
42	PRO	CA	10.35920	22.24930	13.26790	16.41940	1.00000
42	PRO	C	9.30070	22.05930	14.32040	15.72270	1.00000
42	PRO	O	9.72600	21.67350	15.44100	21.95940	1.00000
42	PRO	CB	10.94780	23.67550	13.47810	18.25610	1.00000
42	PRO	CG	10.83020	24.35700	12.17620	22.09740	1.00000
42	PRO	CD	10.27660	23.42640	11.14880	19.15520	1.00000
43	ASN	N	8.05620	22.33690	14.01770	15.43760	1.00000
43	ASN	CA	7.01170	22.18080	15.02930	16.17630	1.00000
43	ASN	C	6.22120	20.88290	14.83500	17.39920	1.00000
43	ASN	O	5.27560	20.81360	15.63240	19.25880	1.00000
43	ASN	CB	6.01080	23.34680	14.88540	17.51300	1.00000
43	ASN	CG	6.73430	24.70360	15.04400	18.46060	1.00000
43	ASN	OD1	7.56440	24.84290	15.91440	22.15520	1.00000
43	ASN	ND2	6.38530	25.68170	14.23930	19.97530	1.00000
44	ALA	N	6.51760	19.96560	13.94140	14.94220	1.00000
44	ALA	CA	5.73320	18.77780	13.69890	10.87260	1.00000
44	ALA	C	5.94750	17.57850	14.59410	12.50860	1.00000
44	ALA	O	5.22910	16.58300	14.42700	11.66240	1.00000
44	ALA	CB	5.98870	18.28820	12.24850	9.42520	1.00000
45	GLY	N	6.87960	17.61120	15.46740	14.82520	1.00000
45	GLY	CA	7.35240	16.56410	16.36330	16.33960	1.00000
45	GLY	C	7.96550	15.43220	15.49950	17.35930	1.00000
45	GLY	O	7.72220	14.26410	15.86620	20.09010	1.00000
46	LEU	N	8.69810	15.63400	14.44430	13.47900	1.00000
46	LEU	CA	9.29970	14.58160	13.61780	20.47230	1.00000
46	LEU	C	10.44650	13.80450	14.28280	22.66040	1.00000
46	LEU	O	11.60840	14.13850	13.97510	27.84500	1.00000
46	LEU	CB	9.65820	15.16400	12.24920	20.41880	1.00000
46	LEU	CG	8.62810	15.16490	11.12730	26.24330	1.00000
46	LEU	CD1	7.14040	15.06640	11.42500	19.05700	1.00000
46	LEU	CD2	8.84290	16.32560	10.16460	17.72640	1.00000
47	THR	N	10.14840	12.81140	15.09870	17.76700	1.00000
47	THR	CA	11.18460	11.99760	15.78450	21.31720	1.00000
47	THR	C	11.03630	10.51580	15.44300	19.60880	1.00000
47	THR	O	9.89240	10.06210	15.16950	18.74460	1.00000
47	THR	CB	11.19170	12.28590	17.32040	21.37420	1.00000
47	THR	OG1	10.33460	11.35620	18.06100	32.20110	1.00000
47	THR	CG2	10.65500	13.71330	17.56880	28.70860	1.00000
48	MET	N	12.15320	9.79030	15.50650	19.05280	1.00000
48	MET	CA	12.17570	8.34870	15.18780	18.09020	1.00000
48	MET	C	11.36810	8.07340	13.91850	14.83100	1.00000
48	MET	O	11.58870	8.76990	12.95090	17.25400	1.00000
48	MET	CB	11.60730	7.55830	16.34090	20.45360	1.00000
48	MET	CG	12.20780	8.01900	17.63980	28.96210	1.00000
48	MET	SD	13.88970	7.31380	17.75740	36.42290	1.00000
48	MET	CE	13.97890	7.30560	19.58690	36.59890	1.00000
49	ASN	N	10.46410	7.09720	13.97560	15.09220	1.00000
49	ASN	CA	9.63040	6.74450	12.81920	17.23290	1.00000
49	ASN	C	8.13300	6.80430	13.18740	14.00790	1.00000
49	ASN	O	7.33640	6.02760	12.63930	15.07070	1.00000
49	ASN	CB	9.94960	5.33030	12.27570	15.89860	1.00000
49	ASN	CG	9.60270	4.22370	13.27470	17.05890	1.00000
49	ASN	OD1	9.61100	4.47200	14.49290	14.89480	1.00000

49	ASN	ND2	9.31190	3.04950	12.66360	12.85170	1.00000
50	TYR	N	7.80310	7.63850	14.15870	10.88760	1.00000
50	TYR	CA	6.41880	7.74740	14.58040	13.21310	1.00000
50	TYR	C	5.50480	8.42420	13.55800	13.88100	1.00000
50	TYR	O	5.97660	9.32440	12.86520	17.20920	1.00000
50	TYR	CB	6.40750	8.56690	15.86620	17.57610	1.00000
50	TYR	CG	7.27020	8.00850	16.95830	20.75600	1.00000
50	TYR	CD1	7.92970	8.92110	17.77930	20.21410	1.00000
50	TYR	CD2	7.44680	6.63260	17.14120	22.34790	1.00000
50	TYR	CE1	8.72190	8.44760	18.82920	25.89360	1.00000
50	TYR	CE2	8.23080	6.14880	18.17370	24.76730	1.00000
50	TYR	CZ	8.84990	7.07230	19.01650	29.00820	1.00000
50	TYR	OH	9.64940	6.65900	20.04090	34.34030	1.00000
51	CYS	N	4.24460	8.00130	13.50320	11.06220	1.00000
51	CYS	CA	3.24830	8.61340	12.62740	13.24870	1.00000
51	CYS	C	3.07440	10.12910	13.01180	9.97620	1.00000
51	CYS	O	2.85070	10.40680	14.20250	9.20030	1.00000
51	CYS	CB	1.91310	7.89790	12.91910	11.43400	1.00000
51	CYS	SG	2.13590	6.17270	12.27710	16.28430	1.00000
52	ARG	N	3.12760	10.96780	12.00240	10.30190	1.00000
52	ARG	CA	2.94330	12.41620	12.24200	14.25930	1.00000
52	ARG	C	2.18680	13.02910	11.06990	14.11390	1.00000
52	ARG	O	1.84090	12.39410	10.05410	13.05940	1.00000
52	ARG	CB	4.26410	13.18490	12.53270	14.27760	1.00000
52	ARG	CG	5.07210	12.70890	13.74850	16.32140	1.00000
52	ARG	CD	4.42270	13.15490	15.01500	12.76630	1.00000
52	ARG	NE	5.14270	12.83770	16.22870	14.77490	1.00000
52	ARG	CZ	4.99410	11.79510	17.03110	16.01850	1.00000
52	ARG	NH1	5.84200	11.66020	18.06680	18.28030	1.00000
52	ARG	NH2	3.99840	10.92360	16.82790	14.50100	1.00000
53	ASN	N	1.94480	14.36250	11.16370	11.31110	1.00000
53	ASN	CA	1.20130	15.01120	10.05840	11.52120	1.00000
53	ASN	C	1.84500	16.40010	9.83820	13.58050	1.00000
53	ASN	O	1.22000	17.39870	10.27180	11.78300	1.00000
53	ASN	CB	-0.27380	15.10480	10.49160	11.69980	1.00000
53	ASN	CG	-1.16340	15.59220	9.37530	13.01740	1.00000
53	ASN	OD1	-0.73390	15.73050	8.23020	13.06600	1.00000
53	ASN	ND2	-2.43330	15.85210	9.63660	14.07390	1.00000
54	PRO	N	3.03220	16.39120	9.25140	11.22050	1.00000
54	PRO	CA	3.77260	17.62760	9.06450	11.41650	1.00000
54	PRO	C	3.12400	18.63230	8.11810	13.38730	1.00000
54	PRO	O	3.27420	19.87570	8.33390	12.82520	1.00000
54	PRO	CB	5.11970	17.21840	8.42900	14.23850	1.00000
54	PRO	CG	5.06050	15.73600	8.11620	15.33780	1.00000
54	PRO	CD	3.76180	15.23840	8.72940	12.84220	1.00000
55	ASP	N	2.47790	18.10710	7.10450	15.18710	1.00000
55	ASP	CA	1.85190	18.90760	6.01840	13.12600	1.00000
55	ASP	C	0.36670	19.19950	6.18400	11.38800	1.00000
55	ASP	O	-0.27120	19.49250	5.14230	12.83300	1.00000
55	ASP	CB	2.25170	18.26080	4.69520	10.51120	1.00000
55	ASP	CG	1.78290	16.79670	4.51000	13.72180	1.00000
55	ASP	OD1	1.93660	16.34080	3.38570	16.01240	1.00000
55	ASP	OD2	1.33500	16.15850	5.48790	13.42990	1.00000
56	ALA	N	-0.09510	19.03170	7.40710	11.68160	1.00000

56	ALA	CA	-1.52680	19.26970	7.73360	13.28000	1.00000
56	ALA	C	-2.53680	18.59680	6.82150	17.27430	1.00000
56	ALA	O	-3.48530	19.23950	6.25720	12.94670	1.00000
56	ALA	CB	-1.72540	20.79710	7.70580	16.64670	1.00000
57	ASP	N	-2.29230	17.26190	6.63120	10.90280	1.00000
57	ASP	CA	-3.20920	16.47740	5.81770	11.40680	1.00000
57	ASP	C	-4.37000	15.99790	6.69290	11.25650	1.00000
57	ASP	O	-4.55990	16.46640	7.80220	12.53130	1.00000
57	ASP	CB	-2.46380	15.38650	5.05450	12.56140	1.00000
57	ASP	CG	-3.29880	14.95530	3.86160	9.31180	1.00000
57	ASP	OD1	-3.88510	13.85770	3.96810	14.59780	1.00000
57	ASP	OD2	-3.37700	15.64080	2.83390	16.40770	1.00000
58	LYS	N	-5.14900	15.03270	6.17140	14.63630	1.00000
58	LYS	CA	-6.33440	14.49300	6.91380	13.11570	1.00000
58	LYS	C	-6.02510	13.69210	8.15870	12.95630	1.00000
58	LYS	O	-6.89670	13.58110	9.03700	13.77270	1.00000
58	LYS	CB	-7.24700	13.80230	5.91130	12.41050	1.00000
58	LYS	CG	-6.60270	12.54490	5.34480	12.21690	1.00000
58	LYS	CD	-7.78010	11.59490	5.05460	17.28150	1.00000
58	LYS	CE	-7.27410	10.33060	4.38680	17.21310	1.00000
58	LYS	NZ	-8.43300	9.53360	3.94800	20.36770	1.00000
60	GLY	N	-4.83080	13.15130	8.36370	10.73520	1.00000
60	GLY	CA	-4.50110	12.39190	9.58350	15.39730	1.00000
60	GLY	C	-3.02510	11.98520	9.50600	8.97370	1.00000
60	GLY	O	-2.44650	12.17970	8.41140	13.38480	1.00000
61	PRO	N	-2.54470	11.42840	10.59100	8.42290	1.00000
61	PRO	CA	-1.11450	11.04430	10.67540	9.25970	1.00000
61	PRO	C	-0.76200	10.00240	9.61480	11.05180	1.00000
61	PRO	O	-1.63590	9.19070	9.24300	8.59910	1.00000
61	PRO	CB	-0.89510	10.77530	12.14780	14.73470	1.00000
61	PRO	CG	-2.17070	11.05730	12.87210	10.04330	1.00000
61	PRO	CD	-3.24540	11.35590	11.86540	11.26080	1.00000
62	TRP	N	0.44710	10.08850	9.08750	8.44130	1.00000
62	TRP	CA	0.99790	9.29140	8.00230	9.56880	1.00000
62	TRP	C	2.49780	9.14510	8.20740	8.63000	1.00000
62	TRP	O	2.97300	9.76130	9.16940	9.67660	1.00000
62	TRP	CB	0.62400	9.87900	6.63800	8.02550	1.00000
62	TRP	CG	1.13910	11.27550	6.41910	11.91790	1.00000
62	TRP	CD1	0.58400	12.49440	6.80070	10.45270	1.00000
62	TRP	CD2	2.40340	11.56470	5.77620	12.97060	1.00000
62	TRP	NE1	1.42970	13.50730	6.35920	11.94570	1.00000
62	TRP	CE2	2.53580	12.98390	5.73660	10.56000	1.00000
62	TRP	CE3	3.36580	10.71590	5.18210	9.65610	1.00000
62	TRP	CZ2	3.63970	13.57860	5.18380	9.06270	1.00000
62	TRP	CZ3	4.46480	11.33860	4.59310	14.13680	1.00000
62	TRP	CH2	4.60740	12.72320	4.59410	13.64050	1.00000
63	CYS	N	3.11270	8.30620	7.39480	9.74700	1.00000
63	CYS	CA	4.59030	8.11700	7.45290	8.41670	1.00000
63	CYS	C	5.01420	7.54610	6.11410	7.00000	1.00000
63	CYS	O	4.20360	6.96420	5.37120	8.16420	1.00000
63	CYS	CB	5.09360	7.22430	8.58240	10.76770	1.00000
63	CYS	SG	4.47160	5.48430	8.37390	11.80500	1.00000
64	PHE	N	6.28700	7.77370	5.81350	7.40510	1.00000
64	PHE	CA	6.97230	7.19470	4.64040	9.44350	1.00000

64	PHE	C	7.12070	5.72940	5.17240	7.00000	1.00000
64	PHE	O	7.23670	5.52850	6.41420	8.32230	1.00000
64	PHE	CB	8.27570	7.93940	4.32560	9.59600	1.00000
64	PHE	CG	8.08250	9.29730	3.76070	8.65530	1.00000
64	PHE	CD1	7.62660	9.41570	2.44820	7.87570	1.00000
64	PHE	CD2	8.34280	10.40710	4.57920	10.63800	1.00000
64	PHE	CE1	7.37470	10.67020	1.87700	11.47980	1.00000
64	PHE	CE2	8.11870	11.69610	3.98590	9.15630	1.00000
64	PHE	CZ	7.60770	11.78620	2.67770	7.00000	1.00000
65	THR	N	7.03430	4.78620	4.26490	10.78030	1.00000
65	THR	CA	7.08000	3.33800	4.64660	15.78060	1.00000
65	THR	C	8.43370	2.69370	4.41440	16.06150	1.00000
65	THR	O	9.15630	3.23460	3.55410	14.08520	1.00000
65	THR	CB	6.03500	2.74360	3.60450	20.23390	1.00000
65	THR	OG1	4.73350	2.85920	4.22630	27.28010	1.00000
65	THR	CG2	6.45630	1.52950	2.83140	10.67040	1.00000
66	THR	N	8.74740	1.62370	5.13180	12.61690	1.00000
66	THR	CA	10.07620	1.00810	4.85220	15.44580	1.00000
66	THR	C	10.02540	0.06370	3.67160	16.20060	1.00000
66	THR	O	10.99890	-0.65820	3.44310	18.35680	1.00000
66	THR	CB	10.59280	0.25710	6.13760	14.56670	1.00000
66	THR	OG1	9.56080	-0.69130	6.50010	14.56790	1.00000
66	THR	CG2	10.85340	1.29380	7.23430	16.90410	1.00000
67	ASP	N	8.96770	-0.03210	2.91930	17.10900	1.00000
67	ASP	CA	8.85750	-0.90970	1.74640	19.04450	1.00000
67	ASP	C	9.31210	-0.04200	0.58380	18.46930	1.00000
67	ASP	O	8.69810	1.01290	0.41400	19.72200	1.00000
67	ASP	CB	7.47050	-1.49460	1.56480	20.19260	1.00000
67	ASP	CG	7.09350	-2.21430	0.30620	21.93500	1.00000
67	ASP	OD1	6.14490	-3.00850	0.39750	24.41730	1.00000
67	ASP	OD2	7.63530	-2.02750	-0.80170	27.35660	1.00000
68	PRO	N	10.27130	-0.51610	-0.18560	20.58290	1.00000
68	PRO	CA	10.76920	0.25180	-1.33380	23.19560	1.00000
68	PRO	C	9.70650	0.55670	-2.34960	21.70190	1.00000
68	PRO	O	9.80390	1.62950	-2.97510	25.71440	1.00000
68	PRO	CB	11.90870	-0.54390	-1.93010	24.56820	1.00000
68	PRO	CG	12.02780	-1.76280	-1.09030	25.68550	1.00000
68	PRO	CD	10.99660	-1.76800	-0.00500	23.26330	1.00000
69	SER	N	8.67240	-0.25110	-2.48050	20.89680	1.00000
69	SER	CA	7.58750	-0.03980	-3.43650	21.80540	1.00000
69	SER	C	6.46200	0.91820	-3.08420	18.75410	1.00000
69	SER	O	5.70290	1.25750	-4.01130	24.43830	1.00000
69	SER	CB	6.95190	-1.42750	-3.68010	23.49460	1.00000
69	SER	OG	7.80710	-2.14700	-4.56320	32.42190	1.00000
70	VAL	N	6.38310	1.31410	-1.84310	18.44730	1.00000
70	VAL	CA	5.25910	2.17980	-1.37890	18.10820	1.00000
70	VAL	C	5.87470	3.32670	-0.60750	13.95070	1.00000
70	VAL	O	6.33700	3.14550	0.52500	14.74790	1.00000
70	VAL	CB	4.33830	1.26140	-0.53250	22.17210	1.00000
70	VAL	CG1	3.12090	1.87950	0.09810	18.44520	1.00000
70	VAL	CG2	3.85930	0.09580	-1.42140	23.54400	1.00000
71	ARG	N	5.77270	4.45820	-1.30290	13.76180	1.00000
71	ARG	CA	6.40020	5.64920	-0.68620	10.19370	1.00000
71	ARG	C	5.84120	5.99440	0.67040	10.33660	1.00000

71	ARG	O	6.61180	6.18660	1.62790	10.85500	1.00000
71	ARG	CB	6.33370	6.74920	-1.76470	10.08970	1.00000
71	ARG	CG	6.88560	8.08830	-1.21200	9.32700	1.00000
71	ARG	CD	6.86390	9.14310	-2.28540	13.69190	1.00000
71	ARG	NE	7.52330	10.36890	-1.80240	11.57790	1.00000
71	ARG	CZ	8.83340	10.58880	-1.80110	13.51210	1.00000
71	ARG	NH1	9.67120	9.68180	-2.28660	13.95590	1.00000
71	ARG	NH2	9.32250	11.68810	-1.19530	15.57520	1.00000
72	TRP	N	4.52490	6.10120	0.86280	9.51010	1.00000
72	TRP	CA	3.90460	6.51090	2.11280	10.20180	1.00000
72	TRP	C	2.44620	6.01290	2.23470	7.93470	1.00000
72	TRP	O	1.83830	5.72170	1.19410	10.86100	1.00000
72	TRP	CD2	1.53060	9.03200	1.59020	10.04970	1.00000
72	TRP	CE3	0.80360	8.92520	2.77060	11.51360	1.00000
72	TRP	CZ3	-0.53240	9.29850	2.79880	11.45020	1.00000
72	TRP	CH2	-1.09900	9.75440	1.63190	7.00000	1.00000
72	TRP	CZ2	-0.43580	9.86260	0.43660	12.68570	1.00000
72	TRP	CE2	0.91670	9.47580	0.41870	8.77600	1.00000
72	TRP	NE1	1.82130	9.48250	-0.58940	11.48460	1.00000
72	TRP	CD1	3.01960	9.03190	-0.07990	12.68320	1.00000
72	TRP	CG	2.88660	8.69570	1.26070	10.72930	1.00000
72	TRP	CB	3.92400	8.09120	2.17690	7.81340	1.00000
73	GLU	N	1.94910	5.88650	3.42670	7.00000	1.00000
73	GLU	CA	0.55960	5.51670	3.67190	10.23500	1.00000
73	GLU	C	0.11550	6.17900	5.00220	8.76970	1.00000
73	GLU	O	0.97520	6.57430	5.79310	9.47020	1.00000
73	GLU	CB	0.37790	3.98230	3.92990	7.82000	1.00000
73	GLU	CG	0.71060	3.18020	2.65240	11.53200	1.00000
73	GLU	CD	0.66500	1.67350	2.93600	16.68700	1.00000
73	GLU	OE1	0.99940	1.16420	3.99680	15.81660	1.00000
73	GLU	OE2	0.25780	1.00200	1.97290	20.19500	1.00000
74	TYR	N	-1.20820	6.15980	5.18270	12.64910	1.00000
74	TYR	CA	-1.75310	6.68970	6.45810	11.21280	1.00000
74	TYR	C	-1.43570	5.69240	7.53750	10.08070	1.00000
74	TYR	O	-1.20360	4.52430	7.27740	13.67630	1.00000
74	TYR	CB	-3.24850	7.07830	6.43390	10.96720	1.00000
74	TYR	CG	-3.49820	8.18760	5.46160	7.07190	1.00000
74	TYR	CD1	-3.18660	9.49000	5.84850	11.47500	1.00000
74	TYR	CD2	-4.01060	7.95830	4.17630	8.20970	1.00000
74	TYR	CE1	-3.40570	10.55650	4.95620	8.64230	1.00000
74	TYR	CE2	-4.22500	9.00070	3.27850	12.65590	1.00000
74	TYR	CZ	-3.93030	10.32870	3.73210	11.85320	1.00000
74	TYR	OH	-4.11120	11.37410	2.84480	11.45420	1.00000
75	CYS	N	-1.42380	6.16590	8.74130	11.50630	1.00000
75	CYS	CA	-1.09010	5.28500	9.87030	14.01500	1.00000
75	CYS	C	-2.31050	4.54180	10.39690	13.97630	1.00000
75	CYS	O	-3.44650	4.95880	10.24900	13.82810	1.00000
75	CYS	CB	-0.40500	6.09960	10.94850	17.74710	1.00000
*75	CYS	SG	1.40190	6.36590	10.46170	27.96850	1.00000
76	ASN	N	-2.02770	3.41650	11.04620	15.48530	1.00000
76	ASN	CA	-3.03820	2.56740	11.71270	15.13620	1.00000
76	ASN	C	-3.36460	3.20250	13.06840	18.82780	1.00000
76	ASN	O	-2.75570	2.87270	14.12360	23.23940	1.00000
76	ASN	CB	-2.37040	1.15440	11.86570	16.11740	1.00000

76	ASN	CG	-3.32660	0.27210	12.64950	15.71670	1.00000
76	ASN	OD1	-4.53100	0.51170	12.49430	18.70470	1.00000
76	ASN	ND2	-2.83000	-0.63690	13.47310	18.31940	1.00000
77	LEU	N	-4.25840	4.16370	13.08770	15.68120	1.00000
77	LEU	CA	-4.66140	4.90640	14.26450	22.32530	1.00000
77	LEU	C	-6.13440	5.29930	14.10220	22.46190	1.00000
77	LEU	O	-6.58280	5.55610	12.96100	27.33470	1.00000
77	LEU	CB	-3.84030	6.21140	14.35470	23.34450	1.00000
77	LEU	CG	-2.68130	6.48220	15.24970	24.91510	1.00000
77	LEU	CD1	-2.33700	7.99840	15.10700	27.05950	1.00000
77	LEU	CD2	-2.98300	6.18180	16.70860	23.24050	1.00000
78	GLY	N	-6.81320	5.27390	15.22600	23.63340	1.00000
78	GLY	CA	-8.25020	5.63690	15.21020	26.18590	1.00000
78	GLY	C	-8.27620	7.11340	15.69360	22.96030	1.00000
78	GLY	O	-7.39270	7.52670	16.46000	23.76350	1.00000
79	ALA	N	-9.28850	7.75750	15.18440	31.45840	1.00000
79	ALA	CA	-9.62190	9.15490	15.48200	33.29060	1.00000
79	ALA	C	-10.32850	9.18210	16.83870	34.60390	1.00000
79	ALA	O	-11.49240	8.76000	16.80220	42.25590	1.00000
79	ALA	CB	-10.55680	9.73740	14.41930	29.76180	1.00000
80	CYS	N	-9.73000	9.61180	17.91530	32.58660	1.00000
80	CYS	CA	-10.32220	9.69670	19.23990	32.84410	1.00000
80	CYS	C	-11.83050	9.99200	19.20470	35.98390	1.00000
80	CYS	O	-12.34510	10.45460	18.15770	35.42010	1.00000
80	CYS	CB	-9.67570	10.81710	20.04750	33.88050	1.00000
80	CYS	SG	-7.97390	10.67690	20.53690	33.55440	1.00000

* Cys75 SG position for left-handed Cys51-Cys75 disulfide is listed. Alternate position:

75	CYS	SG	0.05650	5.74360	12.58280

APPENDIX B

Crystallographic parameters of solvent molecules included in orthorhombic plasminogen K4 structure.

SULFATE ION

Atom		Coordinates (Å) X	Y	Z	B (Å^2)	Occupancy
SUL	S	6.04600	13.79300	-0.71820	16.11150	1.00000
SUL	O1	7.54750	13.85560	-0.94280	21.15120	1.00000
SUL	O2	5.63250	12.32430	-0.66800	18.73570	1.00000
SUL	O3	5.68230	14.45160	0.53690	18.19410	1.00000
SUL	O4	5.49010	14.35350	-1.96820	22.26630	1.00000

WATER MOLECULES

Molecule	Coordinates (Å) X	Y	Z	B (Å^2)	Occupancy	Occ^2/B
WAT 1	8.14400	8.80200	7.85500	11.52090	1.00000	0.0868
WAT 2	2.93120	16.49340	13.10580	15.33130	1.00000	0.0652
WAT 3	9.87130	18.26770	13.82180	16.01820	1.00000	0.0624
WAT 4	6.55930	22.40100	11.55240	16.09920	0.97110	0.0586
WAT 5	9.17060	15.57890	0.25740	17.58870	1.00000	0.0569
WAT 6	-0.51150	-1.37260	14.76610	18.06600	1.00000	0.0553
WAT 7	2.91490	0.91610	15.77660	18.50280	0.98150	0.0521
WAT 8	1.31900	14.45220	14.36750	19.61940	1.00000	0.0510
WAT 9	4.08680	21.20420	10.83830	20.77920	1.00000	0.0481
WAT10	14.03830	4.04510	8.83350	21.24800	0.98060	0.0452
WAT11	1.98940	19.52340	12.12240	20.95460	0.94420	0.0426
WAT12	12.09470	2.31630	10.36810	20.52660	0.90380	0.0398
WAT13	-4.42460	8.60070	9.87010	23.66090	0.96730	0.0395
WAT14	13.09140	20.17470	-3.84530	21.95910	0.92770	0.0392
WAT15	-5.09810	20.50150	13.93880	20.42830	0.87610	0.0376
WAT16	-0.95950	5.96410	19.76940	23.48920	0.93240	0.0370
WAT17	8.76720	19.66080	16.92080	16.54330	0.78040	0.0368
WAT18	1.17260	18.44640	20.30720	21.93220	0.89850	0.0368
WAT19	-11.34030	7.62520	2.05880	23.58300	0.92060	0.0359
WAT20	-7.81180	19.48780	13.86410	24.03110	0.92290	0.0354
WAT21	-6.32700	9.44520	7.65280	23.54150	0.88310	0.0331
WAT22	-4.64140	17.59170	16.19520	24.76660	0.90440	0.0330
WAT23	15.63340	16.55410	2.76790	18.52510	0.76320	0.0314
WAT24	2.30020	22.15740	7.74780	20.68870	0.80380	0.0312
WAT25	4.10970	21.52670	3.58500	27.77960	0.92640	0.0309
WAT26	8.09450	19.18900	1.60580	24.71720	0.86380	0.0302
WAT27	13.87010	10.61930	12.70080	24.42410	0.84400	0.0292
WAT28	5.63950	-1.50940	12.39980	23.12480	0.80820	0.0282
WAT29	17.32760	3.82280	-1.56320	21.48180	0.77200	0.0277

WAT30	1.07520	18.02580	1.26310	23.82720	0.80780	0.0274
WAT31	7.37460	13.71020	18.94010	22.53120	0.75150	0.0251
WAT32	8.59300	-3.94450	-2.47020	22.25090	0.73510	0.0243
WAT33	-7.01080	7.05680	1.41020	24.99070	0.77610	0.0241
WAT34	-5.41130	6.33840	8.97950	28.15760	0.81300	0.0235
WAT35	-9.14960	15.44410	9.31830	22.77290	0.72780	0.0233
WAT36	2.54900	5.92680	-1.30420	22.50030	0.72240	0.0232
WAT37	14.22730	6.06660	14.53030	24.27070	0.72720	0.0218
WAT38	-3.98990	11.82920	20.96460	30.63600	0.81660	0.0218
WAT39	19.71730	6.59780	-1.11220	25.24530	0.73400	0.0213
WAT40	-1.51440	18.03670	3.15820	25.34530	0.71900	0.0204
WAT41	9.28310	2.67000	16.71740	27.92460	0.75350	0.0203
WAT42	9.21980	29.13880	3.97850	25.07610	0.70990	0.0201
WAT43	16.55120	0.92310	3.03780	27.15210	0.73770	0.0200
WAT44	19.79080	8.52060	4.86650	28.67870	0.74510	0.0194
WAT45	-6.05990	19.33930	8.18540	21.00730	0.62630	0.0187
WAT46	-6.95920	3.98320	10.77360	27.98210	0.70870	0.0180
WAT47	18.58560	15.13360	3.42730	28.22940	0.70870	0.0178
WAT48	-10.37590	6.44610	19.31160	26.98050	0.69210	0.0178
WAT49	-5.94720	4.67950	4.52860	28.11790	0.70400	0.0176
WAT50	3.01100	1.77990	19.72750	27.46170	0.68990	0.0173
WAT51	0.20930	1.20000	22.63510	28.72050	0.70440	0.0173
WAT52	13.74920	-1.03770	3.79360	26.72870	0.67520	0.0171
WAT53	-2.97660	2.19530	16.31060	28.84340	0.69510	0.0168
WAT54	1.60560	-4.48170	3.47160	25.77120	0.65100	0.0164
WAT55	16.66230	16.78990	17.42910	21.81740	0.59470	0.0162
WAT56	5.43820	-4.37160	8.63590	27.60390	0.66770	0.0161
WAT57	6.57880	2.61340	18.08440	28.15420	0.67280	0.0161
WAT58	-14.87130	18.55570	11.54520	25.07690	0.63110	0.0159
WAT59	11.73540	5.74080	-3.10090	21.18020	0.56450	0.0150
WAT60	17.67050	14.16970	7.84700	26.86530	0.63430	0.0150
WAT61	2.11940	19.67240	15.19930	22.47940	0.57090	0.0145
WAT62	11.83690	18.57560	15.50610	28.86080	0.63830	0.0141
WAT63	-2.47280	-2.21180	10.18750	28.34880	0.63210	0.0141
WAT64	17.07320	15.00490	9.82540	30.86680	0.64200	0.0133
WAT65	15.71770	0.40070	-0.53480	23.42530	0.55530	0.0132
WAT66	-0.47660	12.57540	21.23530	19.05330	0.50070	0.0132
WAT67	14.51460	2.98170	6.70790	22.87120	0.54810	0.0131
WAT68	-5.51850	19.19510	18.26410	26.94980	0.59460	0.0131
WAT69	-0.53460	6.14900	-0.19010	33.75030	0.66120	0.0130
WAT70	-6.00600	17.34550	3.28200	23.84430	0.55060	0.0127
WAT71	10.14880	21.14330	2.98850	24.37450	0.55200	0.0125
WAT72	11.21850	0.46650	-6.82690	24.91860	0.55550	0.0124
WAT73	3.48610	18.85100	21.51430	28.46260	0.59000	0.0122
WAT74	0.40760	15.87420	22.49770	23.96800	0.52680	0.0116
WAT75	-8.10570	7.27740	10.73090	25.00740	0.53350	0.0114
WAT76	7.84460	3.89870	-4.20870	27.20930	0.55310	0.0112
WAT77	-6.95080	6.48910	6.84190	30.01110	0.57540	0.0110
WAT78	9.11830	23.86170	5.90840	35.84140	0.62820	0.0110
WAT79	18.89580	25.60080	4.21400	34.74710	0.61240	0.0108
WAT80	11.89480	4.40620	23.15630	27.38410	0.53920	0.0106
WAT81	-11.50910	11.68690	23.04300	29.76500	0.55910	0.0105
WAT82	5.76950	22.00980	18.90260	21.67910	0.47580	0.0104
WAT83	-11.21510	13.23320	7.02330	23.78250	0.49630	0.0104

WAT84	-9.06020	17.72120	22.52770	31.14770	0.55710	0.0100
WAT85	-10.65770	19.04250	12.08140	30.07720	0.54720	0.0100
WAT86	13.93810	12.92780	13.21270	18.96070	0.42710	0.0096
WAT87	-11.50570	13.92780	17.29600	33.59600	0.56470	0.0095
WAT88	-7.15570	-0.63800	15.75910	23.78800	0.47050	0.0093
WAT89	6.89400	11.41370	21.57550	27.39240	0.49660	0.0090
WAT90	-9.87550	19.82060	15.74930	29.60170	0.49620	0.0083
WAT91	6.91670	17.80590	19.81600	31.93040	0.48380	0.0073
WAT92	7.98050	2.85690	-7.63720	24.61100	0.41880	0.0071
WAT93	-0.58860	-0.69120	-0.09150	33.15100	0.48350	0.0071
WAT94	13.57400	26.87990	7.03230	33.90540	0.47950	0.0068
WAT95	3.45270	-1.69900	19.60070	26.79520	0.39450	0.0058
WAT96	9.36070	11.63420	21.80340	22.14050	0.34580	0.0054

LIST OF REFERENCES

LIST OF REFERENCES

1 Castellino,F.M., Ploplis,V.A., Powell,J.R., Strickland,D.K. (1981) *J.Biol. Chem.* **256**, 4778-4782.

2 Trexler,M., Patthy,L. (1983) *Proc.Natl.Acad.Sci. USA* **80**, 2457-2461.

3 Magnusson,S., Petersen,T.E., Sottrup-Jensen,L., Claeys,H. (1975) in *Proteases and Biological Control* (Reich,E., Rifkin,D.B., Shaw,E., eds) pp 123-149, Cold Spring Harbor Laboratories, Cold Spring Harbor,NY.

4 Esmon,C.T., Jackson,C.M. (1974) *J.Biol.Chem.* **249**, 7791-7797.

5 McMullen,B.A., Fujikawa,K. (1985) *J.Biol.Chem.* **260**, 5328-5340.

6 Sottrup-Jensen,L., Claeys,H., Zajdal,M., Petersen,T.E.,Magnusson,S. (1978) *Prog.Chem. Fibrinolysis Thrombolysis* **3**,191-209.

7 Lerch,P.G., Rickli,E.E., Lergier,W., Gillessen,D. (1980) *Eur.J.Biochem.* **107**, 7-13.

8 Steffens,G.J., Gunzler,W.A., Otting,F., Frankus,E., Flohe,L. (1982) *Hoppe-Seyler's Z. Physiol.Chem.* **363**, 1043-1058.

9 Gunzler,W.A., Steffens,G.J., Otting,F., Kim,S-M.A., Frankus,E., Flohe,L. (1982) *Hoppe-Seyler's Z. Physiol.Chem.* **363**, 1155-1165.

10 Pennica,D., Holmes,W.E., Kohr,W.J., Harkins,R.N., Vehar,G.A., Ward,C.A., Bennett,W.F., Yelverton,E., Seeberg,P.H., Heynober,H.L., Goeddel,D.V., Collen,D. (1983) *Nature* **201**, 214-221.

11 Verheijen,J.H., Caspers, M.P.M., Chang,G.T.G., de Munk,G.A.W., Pouwels, P.H.,Enger-Valk,B.E. (1986) *EMBO J.* **5**, 3525-3530.

12 van Zonneveld,A.-J., Veerman,H., Pannekoek,H. (1986) *J.Biol.Chem.* **261**, 14214-14218.

13 Gardell,S.J., Duong,L.T., Diehl,R.E., York,J.D., Hare,T.R.,Register,R.B., Jacobs,J.W., Dixon,R.A.F., Friedman,P.A. (1989) *J.Biol.Chem.* **264**, 17947-17952.

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

15 McLean,J.W., Tomlinson,J.E., Wun-Jing,K., Eaton,D.L., Chen,E.Y., Fless, G.M., Scanu,A.M., Lawn,R.M. (1987) *Nature*, **330**, 132-137.

16 Patthy,L. (1985) *Cell* **41**, 657-663.

17 Collen,D. (1980) *Thromb.Haemostasis* **43**, 77-89.

18 Wiman,B., Collen,D. (1978) *Nature* **272**, 549-550.

19 Violand,B.N., Castellino,F.J. (1976) *J.Biol.Chem.* **251**, 3906-3912.

20 Christensen,U., Clemmensen,I. (1977) *Biochem J.* **163**, 389-391.

21 Lucas,M., Fretto,L., McKee,P. (1983) *J.Biol.Chem.* **258**, 4249-4256.

22 Hochschwender,S., Laursen,R. (1981) *J.Biol.Chem.* **256**, 11166-11171.

23 Markus,G., DePasquale,J.L., Wissler,F.C. (1978) *J.Biol.Chem.* **253**, 727-732.

24 Markus,G., Priore,R.L., Wissler,F.C. (1979) *J.Biol.Chem.* **254**, 1211-1216.

25 Markus,G., Evers,J.L., Hobika,G.H. (1978) *J.Biol.Chem.* **253**, 733-739.

26 Thorsen,S., Clemmensen,I., Sottrup-Jensen,L., Magnusson,S.(1981) *Biochim.Biophys.Acta* **668**, 377-387.

27 Winn,E., Hu,S., Hochschwender,S., Laursen,R. (1980) *Eur.J.Biochem.* **104**, 579-586.

28 Thewes,T., Constantine,K., Byeon,I-J.L., Llinas,M. (1990) *J.Biol.Chem.* **265**, 3906-3915.

29 Christensen,U. (1984) *Biochem J.* **223**, 413-421.

30 Mangel,W.F., Lin,B., Ramakrishnan,V. (1990) *Science* **248**,69-73.

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

32 Seshadri,T.P. unpublished results of this laboratory.

33 Park,C.H., Tulinsky,A. (1986) *Biochemistry* **25**, 3977-3982.

34 Castellino,F.J., De Serrano,V.S., Powell,J.R., Johnson,W.R.,Beals,J.M. (1986) *Arch.Biochem.Biophys.* **247**, 312-320.

35 De Marco,A., Laursen,R.A., Llinas,M. (1985) *Biochim.Biophys.Acta* **827**, 369-380.

36 Williams,R.J.P., Esnouf,P., Lawrence,M., Cederholm-Williams,S.A. (1986) *FEBS Lett.* **209**, 111-116.

37 Atkinson,R.A., Williams,R.J.P. (1990) *J.Mol.Biol.* **212**, 541-552.

38 Hochschwender,S.M., Laursen,R.A. (1981) *J.Biol.Chem.* **256**, 11172-11176.

39 Trexler,M., Vali,Z., Patthy,L. (1982) *J.Biol.Chem.* **257**, 7401-7406.

40 Lerch,P.G., Rickli,E.E. (1980) *Biochem.Biophys.Acta* **625**, 374-378.

41 Ramesh,V., Petros,A.M., Llinas,M., Tulinsky,A., Park,C.H. (1987) *J.Mol.Biol.* **198**, 481-498.

42 DeMarco,A., Laursen,R.A., Llinas,M. (1986) *Arch.Biochem.Biophys.* **244**, 727-741.

43 Trexler,M., Patthy,L. (1983) *Proc.Natl.Acad.Sci. USA* **80**, 2457-2461.

44 Tulinsky,A., Park,C.H., Mao,B., Llinas,M. (1988) *Proteins* **3**, 85-96.

45 Mulichak,A.M., Park,C.H., Tulinsky,A., Petros,A.M., Llinas,M. (1989) *J.Biol. Chem.* **264**, 1922-1923.

46 Carter,Jr.,C.W., Carter,C.W. (1979) *J.Biol.Chem.* **254**, 12219-12223.

47 Hatada,M. Hoffmann-LaRoche Inc., (1987) personal comm.

48 Thaller,C., Eichele,G., Weaver,L.H., Wilson,E., Karlsson,R.,Jansonius,J.N. (1985) *Methods Enzymol.* **114**, 132-135.

49 Wyckoff,H.W., Doscher,M., Tsernoglou,D., Inagami,T., Johnson,L.N., Hardman,K.D., Allewell,N.M., Kelly,D.M., Richards,F.M. (1967) *J.Mol.Biol.* **27**, 563-578.

50 Buck,D., Tulinsky,A., Gaier,J. (1982) unpublished results.

51 North,A.C.T., Philips,D.C.,Mathews,F.A. (1968) *Acta Cryst.* **24A**, 351-359.

52 Rossman,M.G., Blow,D.M. (1962) *Acta Cryst.* **15**, 24-31.

53 Steigemann,W. (1974) Dissertation, Technische Universitat, Munchen.

54 Ravichandran,K.G. (1989) Dissertation, Michigan State Univ., East Lansing, Michigan.

55 Fujinaga,M., Read,R.J. (1987) *J.Appl.Cryst.* **20**, 517-521.

56 Jones,T.A. (1982) In: Computational Crystallography, Sayre D. ed. Oxford: Clarendon Press, pp 303-317.

57 Hendrickson,W.A., Konnert,J.H. (1980) In *Biomolecular Structure, Function, Conformation and Evolution*, R. Srinivasan ed. **1**, 43-57, Oxford: Pergamon.

58 Finzel,B.C. (1987) *J.Appl.Cryst.* **20**, 53-55.

59 Luzzati,P.V. (1952) *Acta Cryst.* **5**, 802-810.

60 Ramachandran,G.N., Mitra,A.K. (1976) *J.Mol.Biol.* **107**, 85-92.

61 Gorbitz,C.H. (1989) *Acta Cryst.* **B45**, 390-395.

62 Taylor,R., Kennard,O., Versichel,W. (1984) *Acta Cryst.* **B40**, 280-288.

63 Venkatachalam,C.M. (1968) *Biopolymers* **6**, 1425-1436.

64 Crawford,J.L., Lipscomb,W.N., Schellman,C.G. (1973) *Proc.Nat.Acad.Sci. USA* **70**, 538-542.

65 Lee,B., Richards,F.M. (1971) *J.Mol.Biol.* **55**, 379-400.

66 Downie,T.C., Harrison,W., Raper,E.S., Hepworth, M.A. (1972) *Acta Cryst.* **B28**, 283-290.

67 Form,G.R., Raper,E.S., Downie, T.C. (1973) *Acta Cryst.* **B29**, 776-782.

68 Burley,S.K., Petsko,G.A. (1985) *Science* **229**, 23-28.

69 Tulinsky,A., Vandlen,R.L., Morimoto,C.N., Mani,N.V., Wright,L.H. (1973) *Biochemistry* **12**, 4185-4192.

70 DeMarco,A., Pluck,N.D., Banyai,L., Trexler,M., Laursen,R.A., Patthy,L., Llinas,M., Williams,R.J.P. (1985) *Biochemistry* **24**, 748-753.

71 Trexler,M., Banyai,L., Patthy,L., Pluck,N.D., Williams,R.J.P. (1985) *Eur.J. Biochem.* **152**, 439-446.

72 James,M.N.G., Sielecki,A.R. (1983) *J.Mol.Biol.* **163**, 299-361.