Crystal and Molecular Structure of Human Plasminogen Kringle 4 Refined at 1.9-Å Resolution[†,‡]

Anne M. Mulichak,[§] A. Tulinsky,[*] and K. G. Ravichandran[∥]

Department of Chemistry, Michigan State University, East Lansing, Michigan 48824

Received April 15, 1991; Revised Manuscript Received July 23, 1991

ABSTRACT: The crystal structure of human plasminogen kringle 4 (PGK4) has been solved by molecular replacement using the bovine prothrombin kringle 1 (PTK1) structure as a model and refined by restrained least-squares methods to an *R* factor of 14.2% at 1.9-Å resolution. The K4 structure is similar to that of PTK1, and an insertion of one residue at position 59 of the latter has minimal effect on the protein folding. The PGK4 structure is highly stabilized by an internal hydrophobic core and an extensive hydrogen-bonding network. Features new to this kringle include a cis peptide bond at Pro30 and the presence of two alternate, perpendicular, and equally occupied orientations for the Cys75 side chain. The K4 lysine-binding site consists of a hydrophobic trough formed by the Trp62 and Trp72 indole rings, with anionic (Asp55/Asp57) and cationic (Lys35/Arg71) charge pairs at either end. With the adjacent Asp5 and Arg32 residues, these result in triply charged anionic and cationic clusters (pH of crystals at 6.0), which, in addition to the unusually high accessibility of the Trp72 side chain, serve as an obvious marker of the binding site on the K4 surface. A complex intermolecular interaction occurs between the binding sites of symmetry-related molecules involving a highly ordered sulfate anion of solvation in which the Arg32 side chain of a neighboring kringle occupies the binding site.

Kringles are three-disulfide triple-loop folding domains which occur repeatedly among proteins primarily involved in the blood coagulation and fibrinolytic pathways. These domains each contain 80–85 residues and share homologous amino acid sequences. Kringles are known to occur singly in urokinase (Steffens et al., 1982; Gunzler et al., 1982), clotting factor XII (McMullen & Fujikawa, 1985), and vampire bat salivary PG[1] activator (Gardell et al., 1989), and as pairs in PT (Magnusson et al., 1975) and t-PA (Pennica et al., 1983). Hepatocyte growth factor includes a series of four (Nakamura et al., 1989), while PG contains a series of five kringles (Sottrup-Jensen et al., 1978). Most striking, however, is apolipoprotein a, with 38 kringles, 37 of which are highly homologous to K4 of PG (McLean et al., 1987). Moreover, kringles are usually found preceding a catalytic domain of enzymes and are known in many cases to play a key role in protein recognition and enzyme specificity. For example, PTK2 mediates PT attachment to factor Va, K2 of t-PA is involved in fibrin binding, and one or more kringles of PG interact with fibrin and the plasmin(ogen) inhibitor antiplasmin. Of all the kringles, however, only the structure of PTK1 has previously been determined by X-ray crystallographic methods (Park & Tulinsky, 1986; Tulinsky et al., 1988a; Seshadri et al., 1991).

The PGK4 module is among those domains with binding ability and is known to bind fibrin, L-lysine, and a number of analogous zwitterionic ligands such as ACA which inhibit fibrinolysis (Lucas et al., 1983; Markus et al., 1979; Winn et al., 1980). The amino acid sequence and disulfide structure of K4 are shown in Figure 1. Previous spectroscopic studies have indicated that, in accordance with the high degree of sequence homology, K4 also has a high structural homology with PTK1, including a similar internal core of hydrophobic side chains and similar elements of secondary structure (Castellino et al., 1986; DeMarco et al., 1985a; Williams et al., 1986; Atkinson & Williams, 1990). The K4 lysine-binding site in particular has been extensively studied. Biochemical and spectroscopic experiments have identified the K4 residues involved in ligand binding to be the charge centers of Asp57 and Arg71 (Hochschwender & Laursen, 1981; Trexler et al., 1982; Lerch & Rickli, 1980; Ramesh et al., 1987). The three-dimensional structure of the binding site was modeled on the basis of this knowledge and the crystallographic structure of PTK1, where Asp55 was additionally implicated in the binding site (Tulinsky et al., 1988b). We present here the details of a highly refined structure of human PGK4 at 1.9-Å resolution, which reveals the presence of a doubly charged cationic center in the lysine-binding site in addition to the doubly charged anionic one. Furthermore, a tenacious cofacial intermolecular interaction of two neighboring binding sites (Mulichak & Tulinsky, 1990) reinforced by a sulfate ion of solvation suggests that a considerable electrostatic force emanates from the site. The manner in which lysine binds in the site is described in the paper immediately following this one (Wu et al., 1991).

Materials and Methods

Orthorhombic crystals of human PGK4, space group $P2_12_12_1$, were grown from a solution of 30% (w/v) PEG 8000,

[†] This work was supported by NIH Grant HL25942.

[‡] The atomic coordinates have been deposited in The Protein Data Bank, Brookhaven National Laboratories, Upton, NY.

[*] To whom inquiries should be addressed.

[§] Present address: The Upjohn Company, 301 Henrietta St., Kalamazoo, MI 49007.

[∥] Present address: Howard Hughes Medical Institute, 5323 Harry Hines Blvd., Dallas, TX 75235.

[1] Abbreviations: PG, plasminogen; PT, prothrombin; t-PA, tissue-type plasminogen activator; K1, K2, ..., kringle 1, kringle 2, ...; DMF, dimethyl formamide; ACA, ε-aminocaproic acid; F1, fragment 1; Gla domain, N-terminal 48 residues of F1.

0006-2960/91/0430-10576$02.50/0

Table I: 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	0.020/0.030	0.023	0.020	0.018
bond angle	0.040/0.050	0.059/0.035	0.051/0.035	0.044
planar 1–4	0.060/0.070	0.070/0.055	0.069/0.055	0.047
planarity				
dev from plane (Å)	0.020	0.018	0.017	0.015
chirality				
chiral volume (Å^3)	0.15	0.21	0.23	0.23
nonbonded contacts (Å)				
single torsion	0.55	0.24	0.22	0.18
multiple torsion	0.55	0.37	0.36	0.21
X···Y H-bond	0.55	0.33	0.38	0.26
torsion angle (deg)				
planar	3	3	3	3
staggered	15	31	25	17
orthonormal	20	28	23	19
thermal restraints (Å^2)				
main-chain bond	2.0/3.0	2.1	3.1	2.8
main-chain angle	3.0/4.0	3.2	4.2	3.5
side-chain bond	3.0/4.0	2.6	4.0	3.6
side-chain angle	3.0/4.0	3.1	4.6	4.4
$\sigma(\|F_o\|)$ = 17.0–70.0 (sin θ/λ–1/6)				
av angle = 116.7 ± 2.5°				
R factor (%)		19.4	15.1	14.2
$\langle B \rangle$ (Å^2)		12.5	16.1	18.0
no. of reflections in calculations		2312	3286	4919

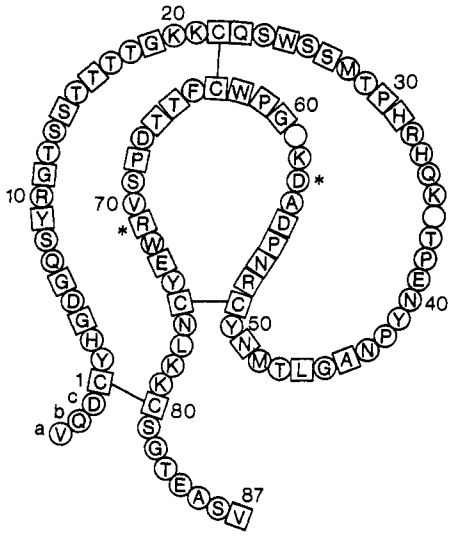

FIGURE 1: Amino acid sequence and disulfide structure of PGK4. The numbering follows the convention of PGK5, with open circles representing deletions with respect to K5. The boxes indicate residues which are conserved with PTK1. Asterisks indicate the charge centers of binding site.

0.12 M ammonium sulfate, and 1.2% DMF at pH 6.0 (Mulichak et al., 1989) using the vapor diffusion sitting drop method. The crystals had unit cell dimensions a = 32.11(1) Å, b = 49.09(2) Å, c = 49.39(3) Å, with four molecules per unit cell, one molecule per asymmetric unit, and an estimated protein content of 62%.

Three-dimensional X-ray diffraction intensities were measured at 1.9-Å resolution from one K4 crystal having dimensions 1.6 × 0.28 × 0.20 mm employing a Nicolet P3/F four-circle diffractometer with a graphite monochrometer and using Cu Kα radiation from a rotating anode X-ray source operating at 3 kW (60 mA). Reflections were measured using a wandering Wyckoff ω-step scan (Wyckoff et al., 1967) of seven 0.03° steps and taking the sum of the five greatest contiguous counts as the integrated intensity. Decay of the crystal was monitored by repeatedly measuring the intensities of three strong reflections after every 100 data measurements. Additional measures of decay included a 6-Å resolution ($0kl$) data set, which was measured before and after data collection, and a short set of strong high-angle reflections, which was measured periodically throughout the data collection. On the basis of these monitors, a 2θ-dependent decay correction was applied, with a maximum decay correction of 14% for 70 h of X-ray exposure. Background measurements were averaged in shells of 2θ and an empirical absorption correction was applied using the procedure of North et al. (1968), with a maximum correction of 1.27. Of 6665 possible reflections at 1.9-Å resolution, 4869 (73%) were taken to be observed based on a cutoff of $2\langle I_{neg} \rangle$, where $\langle I_{neg} \rangle$ is the average negative intensity measurement for a given 2θ range.

The structure of PGK4 was solved by the molecular replacement method using two independent model structures. One was the peptide backbone and conserved side-chain structure of PTK1 (Figure 1) at 2.25-Å resolution (Seshadri et al., 1991) representing 489 atoms (75%) of the K4 structure; the other was that of an unrefined monoclinic PGK4 structure (96%) with two molecules per asymmetric unit (Mulichak et al., 1989), which had just been solved in this laboratory (Ravichandran, 1989). Cross-rotation searches were performed in Patterson space using the program PROTEIN (Steigemann, 1974) with the Patterson maps calculated from 12–3.5-Å resolution (B = 20 Å^2 in the model Patterson). The limits were chosen to omit low-resolution terms including a large solvent contribution and high-resolution terms influenced by the fine structural details.

The rotation search was carried out using only limited sets of the largest model Patterson peaks between different specified vector lengths. Calculations with the PTK1 model were carried out using three vector sets: 1500 vectors between 7 and 23 Å, 1590 between 5 and 20 Å, and 2000 between 5 and 20 Å. One search was made with the monoclinic K4 model corresponding to the third set of conditions. The low limit vector length was used to remove the Patterson origin. The initial searches were made in 5° increments; the orientations with

highest correlation coefficients were then refined in 1° increments. The highest correlation of each search appeared consistently at the same set of rotation angles, with the best solution 7.3σ above background (next highest peak in same solution 6.2σ).

A translation search was performed with the program BRUTE (Fujinaga & Read, 1987) for the three highest rotation solutions, using data from 5.0–4.0, 8.0–4.0, 5.0–2.8, and 8.0–2.8-Å resolution. The best solutions had values ranging between 9σ and 10σ, corresponding to correlation coefficients of 0.35–0.37. Since a single correct solution was not obvious from the translation search results, the possible correctness of a number of highest solutions from each rotation solution was determined by inspection of the resulting C_α-structure packing using an Evans and Sutherland PS390 stereographics system with FRODO (Jones, 1982) software. Eleven translation possibilities from the best and next best rotation solutions and eight of the third best were examined. Due to the tight K4 crystal packing, all but one of the rotation–translation solutions were easily eliminated on the basis of the presence of interpenetrating symmetry mates or unacceptably short contacts. The translation solutions from 8.0 to 4.0 Å of the best rotation solution displayed tight regular packing with no unacceptably close contacts except one interkringle peptide, which could be accounted for by a slight conformational difference of this flexible region from that of the monoclinic crystal form. The combined rotation–translation result was subjected to rigid-body refinement using BRUTE. Most of the omitted side chains were added using an $|F_o|$ electron density map calculated with phases from the final molecular replacement model. The initial K4 model had an R factor of 0.41 ($R = \sum(|F_o| - |F_c|)/\sum|F_o|$).

The K4 structure was refined using the PROFFT (Finzel, 1987) restrained least-squares program. Refinement proceeded in three stages, initially including data at 6.0–2.5-Å, then 2.2-Å, and finally 1.9-Å resolution. The weighting on geometric restraints employed is shown in Table I. The refinement was interrupted regularly, at points of convergence, for interactive FRODO computer graphics sessions, during which the model was manually adjusted to better fit ($2|F_o| - |F_c|$) and ($|F_o| - |F_c|$) density maps. During the first 6.0–2.5-Å stage of refinement, a protocol of alternating sets of refinement cycles having tight and loose geometric restraints was used, with a weighting scheme in which $\sigma(|F|)$ was maintained at approximately half the average $|F_o - F_c|$ discrepancy. In this manner, the crystallographic R factor for the model, with an average thermal parameter of 18 Å^2, was quickly reduced to 0.33. At this point individual isotropic B values were introduced, and the R factor was further reduced to 0.30 ($\langle B \rangle$ = 12 Å^2). Along with remaining omitted side chains, solvent molecules were added to the structure as they became apparent in difference density maps. Additionally, the thermal restraints were relaxed somewhat. 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, but the previous constant structure factor weighting was soon replaced by a θ-dependent variable weighting scheme. Shortly after adding the 2.2-Å data, the alternation of tight and loose geometric restraints was found to be no longer helpful in furthering the refinement. Instead, further improvement was gained primarily through manual adjustment of the model and the continued addition of solvent molecules. Also, thermal restraints were further relaxed. The final 2.2-Å structure included 131 solvent sites, with an R factor of 15.1% and an average of 16 Å^2 for protein atoms. Finally, the 1.9-Å data were added to the calculations. During this stage of refinement, 45 solvent positions which had refined

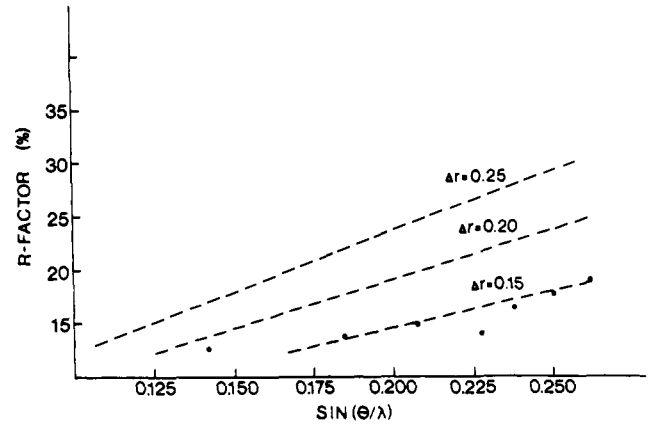

FIGURE 2: Dependence of the R factor on scattering angle for final refined PGK4 structure at 1.9-Å resolution. The dashed lines show theoretical coordinate error curves.

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%. A summary of the least-squares refinement results is given in Table I. The average coordinate error, estimated that from a plot of R factor dependence on resolution (Luzzati, 1952), is 0.15 Å (Figure 2).

RESULTS

General Structure. Nearly the entire K4 structure is extremely well defined; however, little or no electron density was observed for the residues of the interkringle regions a–c and 81–87 or for the side-chain atoms of Lys78 and Lys79. These regions are concluded to be randomly disordered. Furthermore, the density observed for the Cys1–Cys80 disulfide was much less well defined than that for the other two disulfides, suggesting that it also suffers from some degree of disorder. Finally, no density was found for the side chains of Thr12 and Glu39 beyond C_β. The observed main-chain structure of K4, illustrated in Figure 3, agrees well with idealized conformational parameters (Table I, Figure 4). A Ramachandran plot (Ramachandran et al., 1963) of the refined K4 structure (Figure 4) shows that the dihedral angles of nearly all the non-glycine amino acids, with the exception of Met48, fall within energetically preferred zones. Lastly, a single cis peptide bond occurs at Pro30 (Figure 5). This bond has an ω angle of 4°, which compares favorably with the ideal value of 0°. The remaining peptide bond angles fall in a narrow range about 180°.

A striking feature of the PGK4 structure is the close proximity of the two inner loop disulfides, Cys22–Cys63 and Cys51–Cys75, noted previously with PTK1 (Park & Tulinsky, 1986; Seshadri et al., 1991), which forms a four-sulfur cluster and serves as the nucleus of the three-dimensional folding. This backbone folding is stabilized by extensive hydrogen-bonding interactions (Table II), with the resulting secondary structure composed primarily of antiparallel β-structure and reverse β-turns (Table III). These hydrogen-bonding interactions are concentrated primarily near the core of the kringle. Here, short stretches of antiparallel β-strands, involving residues Ser14–Thr16/Lys20–Cys22 and Cys63–Thr65/Arg71–Glu73, cross one another at approximately right angles. These stretches of chain are stabilized further with respect to each other by hydrogen bonds between main-chain atoms of Thr15–Asn76, Thr17–Tyr74, and Lys21–Thr66 linking the two β-segments. Most of the remaining interactions involve

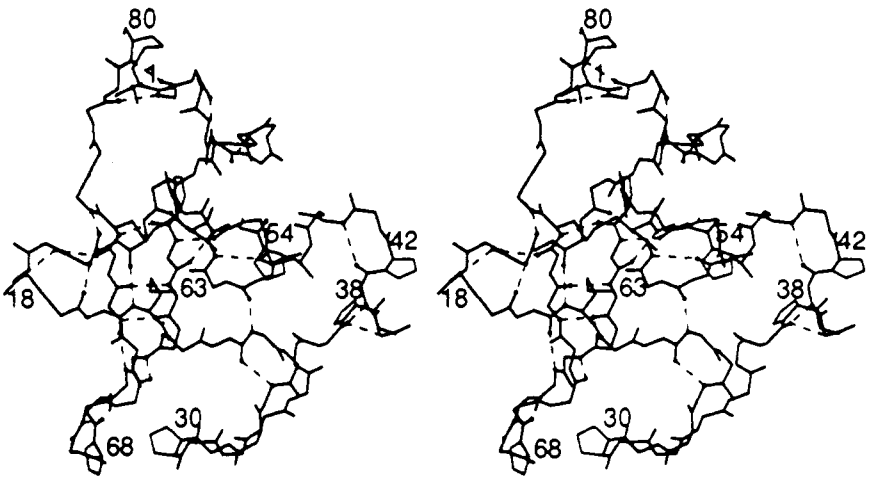

FIGURE 3: Stereoview of backbone and disulfide structure of PGK4, showing positions of proline side chains and main-chain hydrogen bonds (dashed lines).

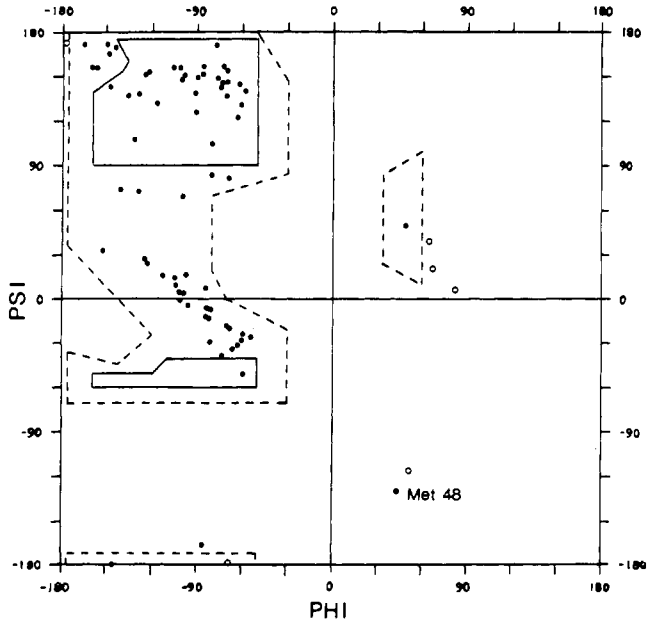

FIGURE 4: Ramachandran plot of final refined PGK4 structure. The open circles denote Gly residues; filled circles denote other residues.

backbone stretches adjacent to disulfide bridges, including a hydrogen bond between Gln23 and 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 (Table II).

The disulfide bridges are worthy of some additional discussion. Interestingly, Cys75 was found to have two alternate side-chain positions which differ by a torsional rotation of approximately 90° (Figure 6). The two positions result in two alternate Cys51–Cys75 disulfide bonds having perpendicular orientations. 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 (Table IV). This orientation positions all four sulfur atoms of the cluster within approximately 4.5 Å of one another in a square 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 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, with a distance between disulfide midpoints of 4.9 Å (Table IV). Though least-squares refinement of the K4 structure was pursued 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. Also, on the basis of 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 if solvent), the occupancies of both alternate Cys75 side-chain orientations refined to roughly 50%.

Of the 76 residues from Cys1 to Cys80, 51 (63%) of the side chains are exposed on the surface of the structure. About 75% of the external side chains are hydrophilic. Since nearly the entire outer surface of the kringle is hydrophilic in nature, it would be expected that, in the absence of specific side-chain interactions with other PG domains, the K4 module 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.

Of the hydrophilic residues, 21 (55%) are involved in intramolecular hydrogen-bonding interactions with adjacent protein side-chain or main-chain atoms. A complete listing of observed PGK4 side-chain interactions is given in Table V. 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 Lys20 side chain is also within 4.6 Å of a carboxy oxygen of the Asp67 side chain. 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, as was also noted in modeling of the PGK4 structure (Tulinsky et al., 1988b). The Asp5 side chain 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. The only significantly exposed hydrophobic residue of the K4 surface is the indole ring of Trp72. This large hydrophobic patch on the otherwise hydrophilic surface, together with the positive and negative charge clusters at either end, serves as a clear marker of the lysine-binding site. The dipolar surface created thusly could be of functional utility in recognizing the charged region of the fibrin ligand site.

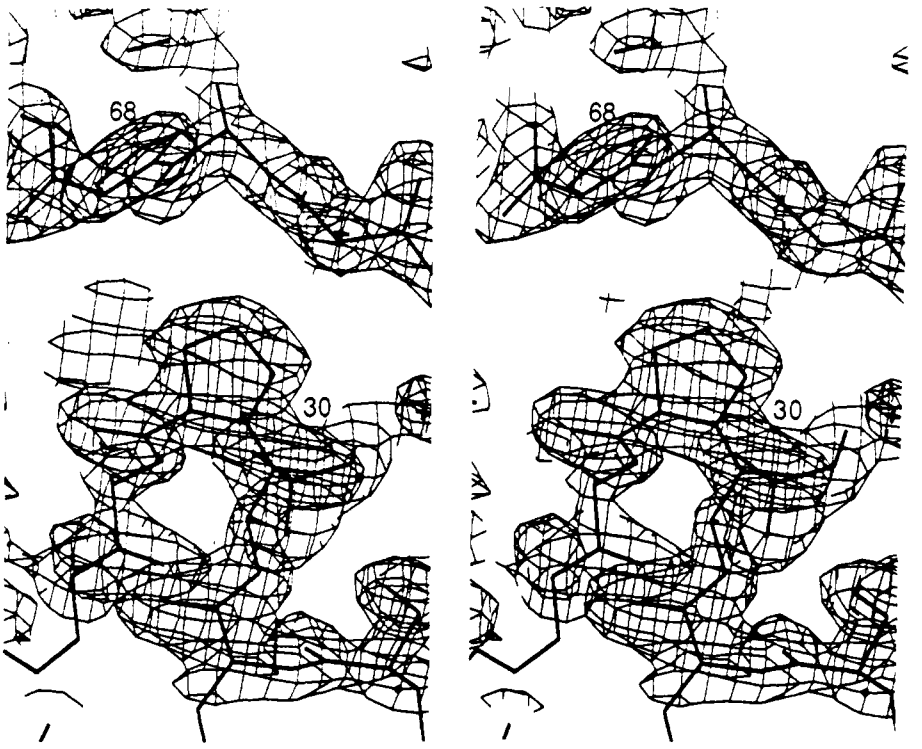

FIGURE 5: Stereoview of $(2|F_o| - |F_c|)$ electron density for Pro30 in cis peptide bond configuration. Contoured at 1σ.

Table II: Hydrogen-Bonding Interactions of the PGK4 Main Chain[a]

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

		Main Chain Interactions Bridged by Solvent (W) distances (Å)				angles (deg)		
donor	acceptor	NH···W	CO···W	NH···O	N···O	HWO	NHW	COW
46 Leu N	41 Tyr O	2.0	2.7	4.1	5.1	124	159	150
53 Asn N	44 Ala O	2.1	2.6	4.3	5.1	127	166	127
64 Phe N	23 Gln O	2.0	2.9	4.4	5.0	124	156	139

[a] Interactions forming β-sheet (β) or reverse turns (T) are indicated. Hydrogen atoms are assigned geometrically idealized positions.

Table III: Observed Secondary Structural Elements of PGK4 Main Chain[a]

	β-structure[b]						
	β1			Ser14 O–Thr16 N Cys22 N–Lys20 O			
	β2			Lys21 O-Gln23 N Thr66 N–Phe64 O			
	β3			Pro61 O–Thr65 N Cys75 N–Arg71 O			
turn	C_1–C_4	O–N	$\phi 1$	$\psi 2$	$\phi 3$	$\psi 3$	type
reverse turns							
T1 His3–Gly6	5.32	2.85	53	−115	−99	5	II′
T2 Gly6–Tyr9	4.89	3.07	−69	−18	−96	−3	I
T3 Thr16–Gly19	5.27	3.04	−55	−26	−104	10	I
T4 Ser24–Ser27	5.38	3.14	−63	−31	−84	−12	I
T5 Thr37–Asn40	5.30	2.85	−55	−25	−73	−38	III
T6 Tyr41–Ala44	5.64	3.25	−60	−24	−101	5	I
T7 Thr47–Tyr50	6.50	3.34	44	−130	−125	27	II′
T8 Asn53–Ala56	5.57	3.55	−66	−34	−97	17	I
T9 Asp67–Val70	5.68	3.25	−60	−28	−82	−12	I
other turns[c]							
T10 Pro30							
T11 Lys59							

[a] Reverse turns are classified according to dihedral angles of residues 2 and 3. [b] In addition, the peptide of Arg52 makes two head/tail hydrogen bonds with Trp62. [c] Turn T10 is an open, rather than reverse, turn at the cis Pro30 (C_1–C_4 is 5.44 Å); T11 is an abrupt 90° bend in the main chain which does not meet the criteria for either type of turn (Crawford et al., 1973).

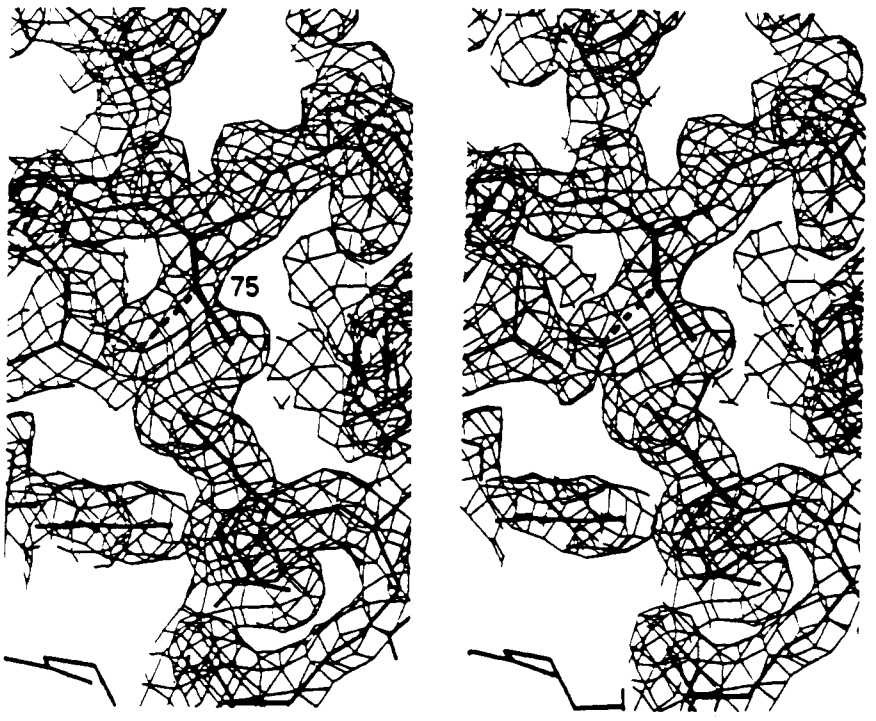

FIGURE 6: Stereoview showing two alternate Cys75 side-chain conformations. Contoured at 1σ.

Among the remaining K4 surface residues, Met28 is particularly interesting as it appears that it may participate in a hydrogen bond in which the side-chain sulfur atom serves as a hydrogen acceptor to the His33 main-chain amide group. The observed N···S and calculated H···S distances of 3.49 and 2.52 Å, respectively, (based on idealized hydrogen positions) agree well with the values of similar hydrogen bonds reported from crystallographic studies of other protein structures (Gregoret et al., 1991).

A dominant feature of the internal structure of PGK4 is a central hydrophobic core, similar to that observed in PTF1 (Seshadri et al., 1991), whicn is formed by a number of aromatic and other bulky hydrophobic side chains (Figure 7). At one end 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, which are in turn adjacent to the side chains of Trp72 and Tyr74. A similar hydrophobic cluster is centered at Leu77, which is surrounded by the side chains of Tyr2, Tyr9, and Pro61. A number of these aromatic side chains participate in highly stabilized perpendicular aromatic stacking interactions. 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 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. The interplanar stacking angles and distances observed in the K4 structure agree well with those reported for other proteins (Burley & Petsko, 1985). Perpendicular stacking interactions

Table IV: Geometrical Parameters of PGK4 Disulfide Bonds

	C–S–S distance (Å)		Cys side-chain angles (deg)		torsion angle (deg)[a]				
disulfide	S1–S2	CA–CA	C1–S–S	C2–S–S	χ1	χ2	χ3	χ2′	χ1′
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

[a] Torsion angles are defined as

CA—CB—SG—SG—CB—CA
χ1 χ2 χ3 χ2′ χ1′

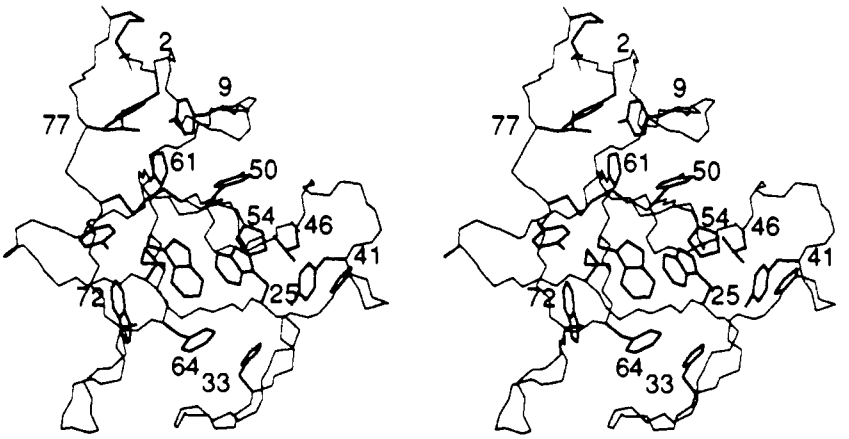

FIGURE 7: Stereoview of PGK4 backbone and the side chains forming the internal hydrophobic core.

Table V: Hydrogen-Bonding Interactions Involving PGK4 Side Chains[a]

		distances (Å)		angles (deg)		
donor	acceptor	H···A	D···A	DHA	DHA	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	33 His ND	1.94	2.92	147		
40 Asn ND2	34 Gln O	2.14	2.96	136	135	125
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
69 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	1.86	2.88	169	133	136
71 Arg NH1	57 Asp OD2	2.28	3.14	140	93	94

[a] Hydrogen atoms were assigned geometrically idealized positions. Donor atom is denoted (D), acceptor atom (A).

involving proline residues are also observed in the PGK4 structure, similar to those present in the structure of α-chymotrypsin (Tulinsky et al., 1973). 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 ~3 Å. In addition, the edge of the Pro30 side chain is oriented similarly approximately 3.5 Å from the Pro68 ring face.

An NMR study of PGK4 suggested the presence of two alternate orientations for Tyr9, one in which the side chain was free to rotate and one in which rotation was severely hindered (DeMarco et al., 1985b). In the crystal structure of PGK4, Tyr9 has a single well-defined conformation in which the tyrosyl ring is highly localized (Figure 8). Stabilizing influences in the surrounding structure include van der Waals interactions of the ring faces to the Leu77 methyl and Gly11 amide groups and the hydrogen bonding of the tyrosyl hydroxyl oxygen to the Ser13 amide group. In the crystal structure, torsional rotation about the Tyr9 C_α–C_β bond is precluded by the path of 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. A similarly localized tyrosine at this position was also observed with PTF1 (Seshadri et al., 1991).

The internal residues of PGK4 include three hydrophilic amino acids: Gln23, Arg52, and Asn53. The hydrocarbon chain of Arg52, although buried, extends toward the kringle surface where the terminal guanidinium group is slightly accessible to the solvent. Two of the Arg52 side-chain nitrogen atoms serve as hydrogen donors to the carbonyl oxygen atoms of Gly11 and Gly45 (Table V). Similarly, Gln23 and Asn53 are also fully compensated with hydrogen-bonding interactions to the protein backbone. The side-chain amino groups of both display bifurcated hydrogen-bonding interactions with the

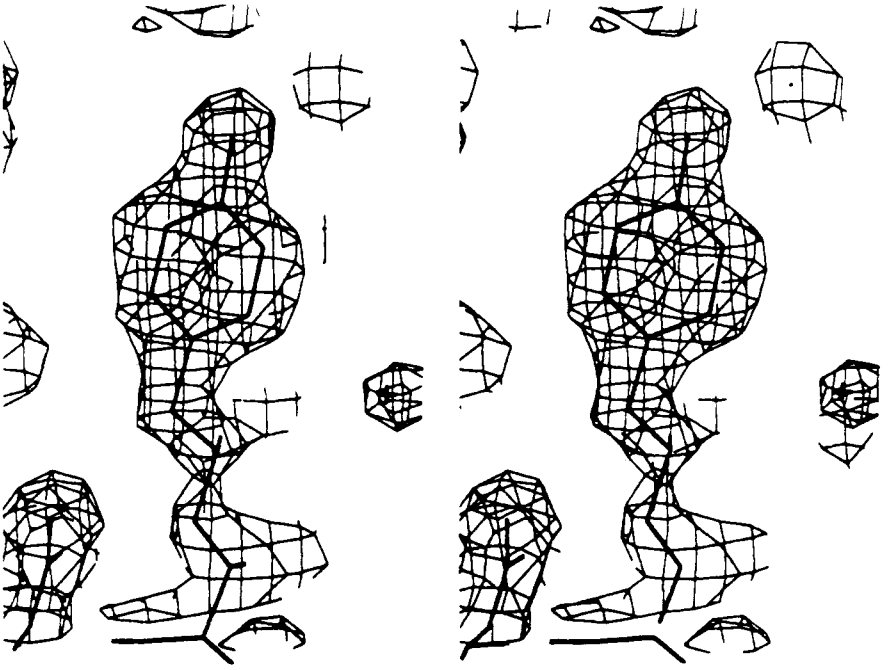

FIGURE 8: Stereoview showing well-defined localized electron density observed for Tyr9 side chain. Contoured at 1σ.

carbonyl oxygen atoms of Ser27/His31 and Asp5/Asp57, respectively, while the side-chain carboxy oxygens hydrogen bond to the amide nitrogens of His31 and Asp57, respectively. Thus, these internal hydrophilic side chains add further stability to the three-dimensional folding.

Nearly all the polar groups provided by other internal side chains are also involved in similar interactions. For example, the hydroxyl group of the Thr16 side chain acts as a hydrogen donor to the carboxy group of Glu73. The side chain of Tyr9 also provides a buried hydroxyl group which, in this case, appears to act as a hydrogen acceptor to the Ser13 main-chain amide group. Also, the indole side chains of Trp25 and Trp62 display hydrogen-bonding interactions to the Tyr50 carbonyl and Asp55 carboxy oxygens, respectively. Of particular note, His33 appears to be unprotonated and to participate in a hydrogen-bonding interaction in which the ND atom of the imidazole ring serves as a hydrogen acceptor to the Lys35 amide group. In many cases these side-chain interactions occur between residues of separate peptide loops and therefore appear to contribute toward stabilization of the overall protein structure.

The variation of the PGK4 thermal parameters with sequence number is given in Figure 9, from which it can be seen that the fluctuations in *B* faithfully reflect the positions of residues with respect to the three-dimensional folding. The thermal parameters are largest 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 main-chain *B* values occur near the ends of three-dimensional loops. Likewise, the average thermal factors of the side chains also correspond closely to their positions in the three-dimensional structure. 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, Gln34, Asn40, Thr47, and Met48. All of these are also highly exposed on the surface of the kringle and do not form stabilizing interactions with neighboring side chains (cf. Table V). The side chains of Thr12 and Glu39, for which no density was observed, have similar exposed environments.

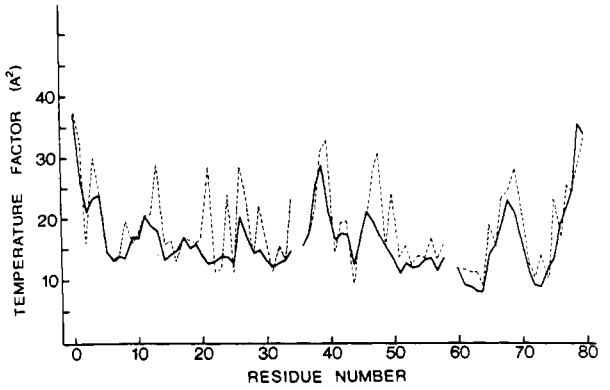

FIGURE 9: Average thermal factors of PGK4 main-chain (solid lines) and side-chain (dashed lines) structure. The breaks are due to deletions with respect to the K5 numbering convention.

Lysine-Binding Site. The K4 lysine-binding site is defined by the peptide segments His31–Lys35, Pro54–Lys58, Pro61–Phe64, and Arg71–Cys75 and is thus supported primarily by the inner kringle loop. 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 10a). These two side chains are oriented in an antiparallel manner with an interplanar angle of approximately 80°, forming a hydrophobic V-shaped trough. Although the Trp62 side chain 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%).[2] The Trp62 and Trp72 side chains display perpendicular aromatic 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 histidine residues, which lend further structural stability to this region.

[2] Accessibility calculations were performed using the ACCESS and ACCFMT programs (Lee & Richards, 1971). Percentages are the actual accessible area divided by the accessible area of a completely exposed side chain.

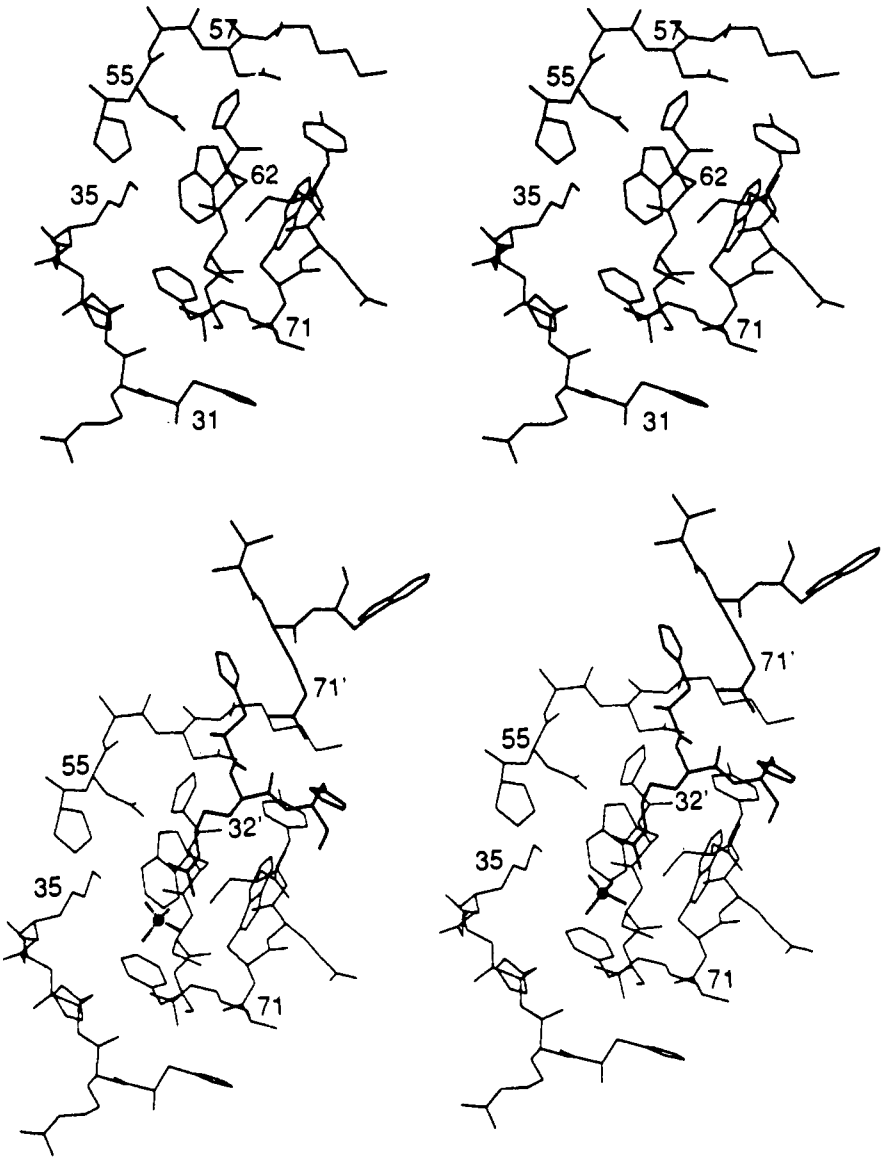

FIGURE 10: Stereoviews of the PGK4 lysine-binding site. (a, top) Binding site in a given molecule. (b, bottom) Intermolecular ion pair interactions which occur at the lysine-binding site. The side chains from the second symmetry-related molecule are shown in bold; sulfate ion is also shown in bold and designated with a filled circle.

The charged residues which have been implicated in ligand binding form a semicircle along the outer edge of the Trp62 indole (Figure 10a), 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. Both carboxylates are hydrogen-bonded to other residues within the site (Trp62 and Tyr74). 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 previously identified as necessary to ligand binding, the Lys35 side chain also retreats into the site, with the amino group approximately 7.6 Å from the Arg71 side chain. That Lys35 is integral to the site has unequivocally been established with the structure determination of the K4–ACA complex (Wu et al., 1991).

The general three-dimensional structure of the binding site corresponds well to the observed K4 binding specificity. The terminal positive and negative charged groups of zwitterionic ligands are likely 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. Although the general characteristics of the lysine-binding site fit the predicted model of ligand binding (Tulinsky et al., 1988b), 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 the 12-Å separation between the charged centers poorly matching the observed optimal ligand length of 6.8 Å (Winn et al., 1980). This unexpected result arises from intermolecular interactions between Arg71 and Asp57 side chains of neighboring molecules, where these residues are both directed outward to make contact with a symmetry mate, rather than inward in a ligand-binding fashion (Figure 10b). However, free bond torsional rotations of the Arg71 side chain easily brings the guanidinium group within 8.5 Å of the anionic center, in good position for ligand binding.

Intermolecular Interactions. The most significant intermolecular contact between neighboring kringle molecules in the orthorhombic PGK4 crystal structure is a very complex binding interaction which involves three symmetry-related molecules and a sulfate ion (Figures 10b and 11). Inter-

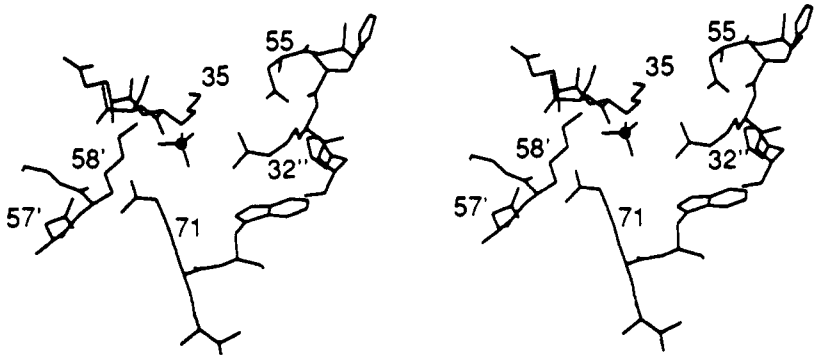

FIGURE 11: Stereoview showing intermolecular interactions of sulfate anion with three symmetry-related PGK4 molecules. The sulfate ion is designated with a filled circle; primes and double prime designate neighboring molecules.

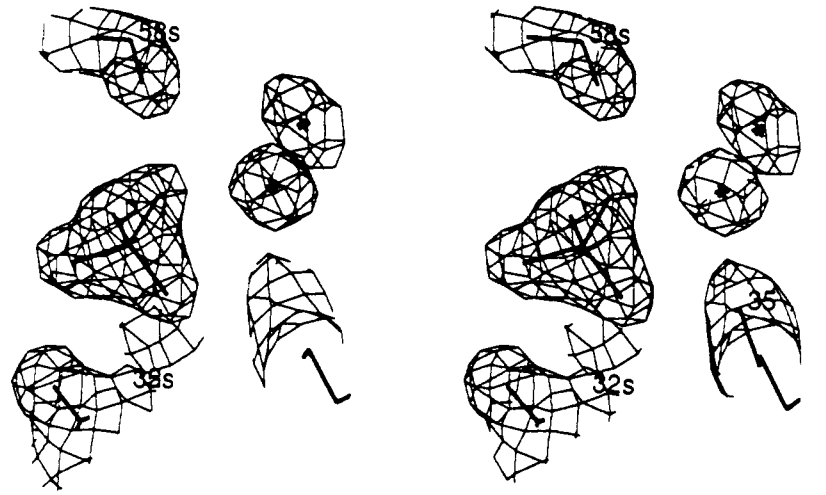

FIGURE 12: Stereoview of $(2|F_o| - |F_c|)$ electron density observed for sulfate anion in orthorhombic PGK4 crystal structure. Contoured at 1σ; two nearby water molecules are also shown.

estingly, this interaction involves four residues which are crucial to the kringle lysine-binding site: Asp55 and Asp57, which comprise the negative ionic center, Arg71 of the positive ionic center, and the newly implicated Lys35 residue. 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. The same Arg71 guanidinium group also participates in an ion pair interaction with the Asp57′ (prime refers to symmetry-related molecules) side-chain carboxy group (Figures 10b and 11) of a second kringle (Mulichak & Tulinsky, 1990). Further linking the two molecules is the adjacent positively charged Lys58′ side chain of a symmetry mate, which also interacts with the 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, Asp55, and Asp57 balanced by the four positively charged side chains Lys35, Arg71, Lys58′, and Arg32″. These hydrogen-bonded salt-bridged interactions are summarized in Table VI, which shows particularly excellent donor–acceptor hydrogen bond angles. All of the foregoing suggests that an electrostatic force of considerable proportions emanates from the binding site.

From the perspective of crystal packing, this trimolecular interaction may be viewed as the result of a crystallographic rotation/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 (Mulichak & Tulinsky, 1990). Repeated throughout the crystal lattice, this coupling of negative and positive centers forms a kringle chain which corresponds to a crystallographic 2-fold screw axis in the x-direction. A comparison of the contacts between kringle molecules related by each of the three $P2_12_12_1$ symmetry elements also easily accounts for the observed crystal face development: the intermolecular interactions in the direction of the a-axis are by far the strongest, and, not surprisingly, the orthorhombic PGK4 crystals grow most rapidly along the a^*-direction.

Table VI: Hydrogen-Bonding Interactions between Sulfate Anion and Protein Side Chains[a]

		distances (Å)		angles (deg)		
atom	donor	H···O	N···O	NHO	SOH	SON
O1	71 Arg NH2	1.85	2.82	162	126	127
O2	32″ Arg NH2	1.93	2.85	153	102	108
O2	71 Arg NE	1.94	2.94	162	112	117
O3	32″ Arg NH1	1.77	2.75	162	122	123
O3	35 Lys NZ	1.70	2.71	169	128	131
O4	58′ Lys NZ	2.08	2.99	148	102	112

[a] Hydrogen atoms are assigned geometrically idealized positions. Primed residue numbers denote hydrogen donors provided by symmetry-related PGK4 molecules.

The binding site interactions explain an observed inability to bind ACA in the orthorhombic K4 crystals in diffusion soaking experiments. Crystals soaked up to 7 days in a 10 mM ACA solution showed unchanged diffraction patterns. Crystals soaked in higher ligand concentrations turned dark and crumbled. From the crystal structure it is clear that the close intermolecular interactions, which occupy both the ionic centers and the hydrophobic pocket, preclude the access of ACA to the binding site. Should the ligand eventually gain

Table VII: Protein–Solvent Hydrogen Bonds in Which a Protein Atom Serves as a Hydrogen Donor (D)

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

access to the site and displace the protein–protein interactions, the disruption of these important intermolecular contacts naturally leads to destruction of the crystal structure itself.

Solvent Structure. The $(2|F_o| - |F_c|)$ electron density for the sulfate anion is well-defined and clearly shows the tetrahedral positions of oxygen atoms (Figure 12), which were also confirmed in a difference density map calculated with only a sulfur atom representing the sulfate group. The geometrical parameters of the sulfate–protein interactions have been given in Table VI. As might be expected from the number of stabilizing hydrogen-bonding interactions, the sulfate ion has a relatively low average thermal factor of 19 Å^2, a value which is comparable to the average value for protein atoms.

The remaining 96 solvent sites are all assigned as water molecules and have been ranked in decreasing order according to the quality factor Q^2/B (James & Sielecki, 1983), where Q is the fractional occupancy. 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 PGK4 structure (a fully occupied water with an average B value corresponds to 0.056). There is a normal distribution of solvent thermal factors about a mean value of 25 Å^2, with nearly 75% of the values falling between 20 and 30 Å^2. In the case of occupancies there is roughly even spread between values of 0.5 and 1.0. 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 VII and VIII.

One solvent site is buried within the K4 protein structure, completely isolated from the surrounding solvent. This water

Table VIII: Protein–Solvent Hydrogen Bonds in Which a Protein Atom Serves as a Hydrogen Acceptor[a]

acceptor	solvent	Q^2/B	O···O distance (Å)	COO angle (deg)
1 Cys O	wat 38	0.022	2.53	151
3 His O	wat 22	0.033	2.92	134
5 Asp OD2	wat 15	0.038	2.65	116
5 Asp OD2	wat 20	0.035	3.29	93
5 Asp OD1	wat 45	0.019	2.54	124
6 Gly O	wat 8	0.051	2.84	154
7 Gln O	wat 61	0.015	2.48	144
7 Gln OE1	wat 6	0.055	2.72	138
7 Gln OE1	wat 11	0.043	3.14	122
8 Ser O	wat 17	0.037	3.00	118
8 Ser O	wat 74	0.012	3.36	171
9 Tyr O	wat 8	0.051	2.90	132
10 Arg O	wat 89	0.009	3.06	157
12 Thr O	wat 57	0.016	2.51	158
16 Thr O	wat 63	0.014	3.08	148
17 Thr O	wat 82	0.010	2.80	134
17 Thr O	wat 63	0.014	3.30	94
18 Thr O	wat 54	0.016	3.11	131
18 Thr O	wat 17	0.037	2.84	135
19 Gly O	wat 56	0.016	2.50	135
19 Gly O	wat 15	0.038	2.71	162
22 Cys O	wat 12	0.040	2.80	141
23 Gln O	wat 1	0.087	2.89	139
26 Ser O	wat 60	0.015	2.87	113
28 Met O	wat 30	0.027	2.94	146
29 Thr O	wat 67	0.013	2.82	141
29 Thr OG	wat 29	0.028	3.44	131
30 Pro O	wat 40	0.020	2.88	144
30 Pro O	wat 59	0.015	3.18	133
33 His O	wat 23	0.031	2.73	142
34 Gln O	wat 71	0.013	3.25	138
34 Gln OE1	wat 69	0.013	2.47	108
34 Gln OE1	wat 14	0.039	3.19	98
40 Asn O	wat 88	0.009	3.16	131
40 Asn O	wat 78	0.011	2.29	139
41 Tyr O	wat 3	0.062	2.67	151
42 Pro O	wat 18	0.037	2.68	127
43 Asn O	wat 63	0.014	3.38	113
43 Asn O	wat 61	0.015	3.38	131
44 Ala O	wat 2	0.065	2.64	127
45 Gly O	wat 31	0.025	3.14	119
46 Leu O	wat 86	0.010	2.73	138
48 Met O	wat 27	0.029	2.95	125
49 Asn OD1	wat 41	0.020	2.88	129
53 Asn O	wat 11	0.043	2.92	132
54 Pro O	wat 24	0.031	2.55	142
54 Pro O	wat 9	0.048	2.95	128
55 Asp O	wat 40	0.020	2.76	135
55 Asp OD1	wat 30	0.027	2.84	116
55 Asp OD1	wat 39	0.021	3.27	122
56 Ala O	wat 45	0.019	3.21	110
56 Ala O	wat 59	0.015	3.30	123
57 Asp O	wat 45	0.019	3.26	121
57 Asp OD2	wat 70	0.013	3.16	104
57 Asp OD2	wat 40	0.020	3.06	108
57 Asp OD2	wat 59	0.015	3.31	119
58 Lys O	wat 35	0.023	2.93	124
61 Pro O	wat 13	0.040	2.92	138
66 Thr O	wat 52	0.017	2.80	146
67 Asp OD2	wat 32	0.024	2.71	127
67 Asp OD1	wat 80	0.011	3.11	137
68 Pro O	wat 76	0.011	3.24	138
69 Ser O	wat 76	0.011	3.40	83
69 Ser OG	wat 57	0.016	2.68	160
72 Trp O	wat 69	0.013	2.78	143
72 Trp O	wat 36	0.023	2.60	131
73 Glu OE2	wat 93	0.007	2.80	176
75 Cys O	wat 34	0.023	2.72	155
76 Asn O	wat 53	0.017	2.30	146
77 Leu O	wat 46	0.018	2.71	132
77 Leu O	wat 75	0.011	3.20	156

[a] Potential interactions were accepted for protein–solvent distances within 3.5 Å.

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 solvent sites occur in an empty pocket of the K4 surface which is removed from, but accessible to the bulk solvent. 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. 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 internal and partly internal sites, 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 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. The five remaining sites are all within 5 Å of polar protein groups or other solvent molecules.

Discussion

As expected, the folding of the PGK4 structure corresponds closely to that of PTK1, with an rms deviation of 0.47 Å between main-chain atoms of the two structures (210 atom pairs). The hydrogen-bonding scheme and secondary structure observed in PGK4 are nearly identical to that of PTK1 (Mulichak & Tulinsky, 1990). One possible difference which had been anticipated was the region adjacent to an insertion at position 59 in the PTK1 sequence; however, the actual perturbation of the backbone conformation is minimal. This loop of the peptide chain is merely expanded in the immediate vicinity of the insertion in PTK1, and only the positions of two residues on either side of the insertion are significantly affected. The most pronounced difference in the structures occurs in the terminal regions of the peptide chain, preceding residue Tyr9 and following residue Cys75. The average deviation between the two structures for the terminal regions alone is 2.1 Å. Furthermore, neither the discrete Cys75 side-chain disorder nor the general disorder of the interkringle region of PGK4 was observed in the PTK1 crystal structure. These discrepancies are all likely to be consequences of the fact that residues 77 and 78 form an antiparallel β-strand with the interkringle peptide in the PTF1 crystals (Tulinsky et al., 1988a), whereas the PGK4 structure has no similar stabilizing interactions. Another significant difference is that the conserved Pro30 residue originally had a trans configuration in PTK1 but proved to be the cis isomer in the PGK4 structure. This apparent difference has since been removed with a reinterpretation and refinement of the latter (Seshadri et al., 1991). Although there is well-defined electron density for the carbonyl oxygen atoms throughout the K4 structure, the electron density maps for PTK1 were not as clear because of Gla domain and carbohydrate disorder (Park & Tulinsky, 1986).

The conformations of side chains conserved between PGK4 and PTK1 are also generally very similar, with an average deviation of 0.66 Å for strictly conserved residues. The high degree of internal conservation maintains the hydrophobic core, which, as previously mentioned, appears to be an important structural foundation of the kringle fold. Nearly all the core residues have deviations less than 0.50 Å. The conserved hydrophilic internal side chains also have very similar conformations in the two structures and, in some cases, share similar hydrogen-bonding interactions. For example, 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 only 0.45 Å. In the K4 structure, this side chain is fully involved in hydrogen bonding with the His31 main chain, and although an analogous interaction was not identified for the PTK1 structure, it may in fact 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.

Although many residues of the PGK4 lysine-binding site are conserved in PTK1, the latter kringle does not share its lysine-binding properties. An Arg/Trp substitution at position 72 and a Glu/Gln and an Ile/Lys substitution at positions 34–35 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: Arg71 forms an ion pair with Glu34, and Arg72 extends across the potential binding site to form an ion pair with Asp55 (Mulichak & Tulinsky, 1990). Furthermore, the substitution of aromatic Trp72 and Tyr74 side chains for charged Arg and Glu residues in PTK1 eliminates the large exposed hydrophobic surface found in PGK4, which provides favorable nonpolar ligand interactions in the lysine-binding site. A more detailed comparison of the K4 and PTK1 structures has been given elsewhere (Mulichak & Tulinsky, 1990).

The crystallographically observed PGK4 structure also generally agrees well with previous spectroscopic results. The central hydrophobic core and aromatic side-chain interactions correlate with NMR and NOE experiments (DeMarco et al., 1985a; Atkinson & Williams, 1990), which indicated interactions between these same residues. The Tyr50 side chain had also been implicated to be part of the hydrophobic core, but although this residue lies near the hydrophobic cluster, the tyrosyl side chain is directed out toward the solvent. However, there may be communication between Tyr50 and the core residues via a hydrogen-bonding interaction between the Tyr50 carbonyl group and the nitrogen atom of the Trp25 indole ring. Some detailed observations are also confirmed, including the specific hydrogen-bonding interactions between Thr16 OG and the Glu73 carboxyl group and between Thr65 and His31 (Atkinson & Williams, 1990). Although NOE connectivity was also reported between His33 and Leu46, there is no direct contact observed between the two residues. Again, however, there may be some indirect communication between the two via the Trp25 side chain.

Finally, it is important to recognize that the intimate intermolecular contacts observed in the PGK4 crystal structure undoubtedly influence the conformations of the crucial Asp55/Asp57 and Arg71 side chains of the binding site. The natural conformations of other residues near the binding site may be altered 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 behavior suggests the possibility that this residue may similarly reinforce the cationic center during ligand binding. This is confirmed in the subsequent accompanying paper (Wu et al., 1991). In addition, in the crystal structure, Lys58 is drawn away from the binding site of the host molecule in an extended conformation to bind to the sulfate associated with the binding site of a neighboring molecule. 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 component of this intermolecular contact is the ligand-like binding of the Arg32″ side chain to Asp55 of the binding site (Figure 10b). The Arg32″ side chain projects into the hydrophobic trough, with the aliphatic chain abutted within 3 Å of the face of the Trp72 indole. Although the guanidinium group mimicks the behavior of more typical binding site ligands, the direction of the arginyl group is opposite that expected of a ligand. 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 driving force for ligand binding. A remarkably similar interaction also occurs in the binding sites of noncrystallographic trimers of K2 of t-PA (de Vos et al., 1991) where a chloride ion is found instead of a sulfate. Such a force may be important for kringle–kringle interactions where lysine binding is not a factor.

ACKNOWLEDGMENTS

We thank Dr. Miguel Llinas for providing us with samples of PGK4.

Registry No. PG, 9001-91-6.

REFERENCES

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

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

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

Crawford, J. L., Lipscomb, W. N., & Schellman, C. G. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 538–542.

DeMarco, A., Laursen, R. A., & Llinas, M. (1985a) *Biochim. Biophys. Acta 827*, 369–380.

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

de Vos, A. M., Ultsch, M. H., Kelley, R. F., Padmanabhan, K., Tulinsky, A., Westbrook, M., & Kossiakoff, A. (1991) *Biochemistry* (in press).

Finzel, B. C. (1987) *J. Appl. Crystallogr. 20*, 53–55.

Fujinaga, M., & Read, R. J. (1987) *J. Appl. Crystallogr. 20*, 517–521.

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.

Gregoret, L. M., Rader, S. D., Fletterick, R. J., & Cohen, F. E. (1991) *Proteins 9*, 99–107.

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.

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

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

Jones, T. A. (1982) in *Computational Crystallography* (Sayre, D., Ed.) pp 303–317, Clarendon Press, Oxford.

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

Lerch, P. G., & Rickli, E. E. (1980) *Biochim. Biophys. Acta 625*, 374–378.

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

Luzzati, P. V. (1952) *Acta Crystallogr. 5*, 802–810.

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.

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

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.

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

Mulichak, A. M., & Tulinsky, A. (1990) *Blood Coagulation Fibrinolysis 1*, 673–679.

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

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

North, A. C. T., Philips, D. C., & Mathews, F. S. (1968) *Acta Crystallogr. 24A*, 351–359.

Park, C. H., & Tulinsky, A. (1986) *Biochemistry 25*, 2977–2982.

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.

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963) *J. Mol. Biol. 7*, 95–99.

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

Ravichandran, K. G. (1989) Ph.D. Dissertation, Michigan State University, East Lansing, Michigan.

Seshadri, T. P., Tulinsky, A., Skrzypczak-Jankun, E., & Park, C. H. (1991) *J. Mol. Biol. 220*, 481–494.

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

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

Steigemann, W. (1974) Ph.D. Dissertation, Technische Universitat, Munchen.

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

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

Tulinsky, A., Park, C. H., Mao, B., & Llinas, M. (1988b) *Proteins 3*, 85–96.

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

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

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

Wu, T.-P., Padmanabhan, K., Tulinsky, A., & Mulichak, A. M. (1991) *Biochemistry* (following paper in this issue).

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.