Crystal Structure of the Rat Liver Fructose-2,6-bisphosphatase Based on Selenomethionine Multiwavelength Anomalous Dispersion Phases[†,‡]

Yong-Hwan Lee,[§] Craig Ogata,[∥] James W. Pflugrath,[⊥] David G. Levitt,[#] Ragupathy Sarma,[○] Leonard J. Banaszak,*[,§] and Simon J. Pilkis[▽]

Departments of Biochemistry and Physiology, University of Minnesota, Minneapolis, Minnesota 55455, Howard Hughes Medical Institute, Brookhaven National Laboratory, Upton, New York 11973-5000, Molecular Structure Corporation, 3200 Research Forest Drive, Woodlands, Texas 77381, and Department of Biochemistry, State University of New York, Stony Brook, New York 11794

Received January 10, 1996; Revised Manuscript Received March 13, 1996[⊗]

ABSTRACT: The crystal structure of the recombinant fructose-2,6-bisphosphatase domain, which covers the residues between 251 and 440 of the rat liver bifunctional enzyme, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase, was determined by multiwavelength anomalous dispersion phasing and refined at 2.5 Å resolution. The selenomethionine-substituted protein was induced in the methionine auxotroph, *Escherichia coli* DL41DE3, purified, and crystallized in a manner similar to that of the native protein. Phase information was calculated using the multiwavelength anomalous dispersion data collected at the X-ray wavelengths near the absorption edge of the K-shell α electrons of selenium. The fructose-2,6-bisphosphatase domain has a core α/β structure which consists of six stacked β-strands, four parallel and two antiparallel. The core β-sheet is surrounded by nine α-helices. The catalytic site, as defined by a bound phosphate ion, is positioned near the C-terminal end of the β-sheet and close to the N-terminal end of an α-helix. The active site pocket is funnel-shaped. The narrow opening of the funnel is wide enough for a water molecule to pass. The key catalytic residues, including His7, His141, and Glu76, are near each other at the active site and probably function as general acids and/or bases during a catalytic cycle. The inorganic phosphate molecule is bound to an anion trap formed by Arg6, His7, Arg56, and His141. The core structure of the Fru-2,6-P_2ase is similar to that of the yeast phosphoglycerate mutase and the rat prostatic acid phosphatase. However, the structure of one of the loops near the active site is completely different from the other family members, perhaps reflecting functional differences and the nanomolar range affinity of Fru-2,6-P_2ase for its substrate. The imidazole rings of the two key catalytic residues, His7 and His141, are not parallel as in the yeast phosphoglycerate mutase. The crystal structure is used to interpret the existing chemical data already available for the bisphosphatase domain. In addition, the crystal structure is compared with two other proteins that belong to the histidine phosphatase family.

The overall flux of glycolytic or gluconeogenic pathways in mammalian cells is determined by the micromolar levels of cellular fructose-2,6-bisphosphate (Fru-2,6-P_2),[1] which is not itself an important metabolic intermediate. Elevated cellular levels of Fru-2,6-P_2 lead to net increases in glycolysis versus gluconeogenesis by allosteric modulation of two major rate-limiting enzymes in sugar metabolism: activation of the 6-phosphofructo-1-kinase (6PF-1-K) and inhibition of the fructose-1,6-bisphosphatase (Fru-1,6-P_2ase). A low level of Fru-2,6-P_2 causes an increased flux in gluconeogenesis by reducing the activation of 6PF-1-K and the inhibition of Fru-1,6-P_2ase (Pilkis et al., 1995). This switching on and off of two opposite carbohydrate metabolic processes by a single molecule, Fru-2,6-P_2, in the liver is an important mechanism for the control of energy storage and blood glucose levels in the rest of the body (Pilkis et al., 1995).

The sugar bisphosphate, Fru-2,6-P_2, is synthesized and hydrolyzed by a bifunctional enzyme consisting of a single polypeptide chain of MW = 55 kDa. The bifunctional enzyme, 6PF-2-K/Fru-2,6-P_2ase, has the kinase domain at residues 1−250 and the bisphosphatase at residues 251−470. The N-terminal kinase domain catalyzes the formation of the β-anomer of Fru-2,6-P_2 using ATP and Fru-6-P as substrates (Lee et al., 1995; Pilkis et al., 1995). The C-terminal bisphosphatase domain catalyzes the hydrolysis

[†] This study was supported by an NIH grant, R37-DK38354 to S.J.P., and by a grant from the Minnesota Supercomputer Institute, MG44901 to S.J.P. and L.J.B.

[‡] The pdb file of the crystal structure of the rat liver fructose-2,6-bisphosphatase domain has been deposited in the Protein Data Bank under the accession code 1FBT.

* To whom correspondence should be addressed: Department of Biochemistry, University of Minnesota, 435 Delaware St. SE, Minneapolis, MN 55455.

[§] Department of Biochemistry, University of Minnesota.

[∥] Brookhaven National Laboratory.

[⊥] Molecular Structure Corp.

[#] Department of Physiology, University of Minnesota.

[○] State University of New York.

[▽] Deceased.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1996.

[1] Abbreviations: Se-Met, selenomethionine; MAD, multiwavelength anomalous dispersion; MIR, multiple isomorphous replacement; Fru-2,6-P_2, fructose 2,6-bisphosphate; Fru-2,6-P_2ase, fructose-2,6-bisphosphatase; 6PF-2-K/Fru-2,6-P_2ase, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase; Fru-6-P, fructose 6-phosphate; E-P, phosphoenzyme; P_i, inorganic phosphate; PEG, poly(ethylene glycol); IPTG, isopropyl thiogalactoside; YPGM, yeast phosphoglycerate mutase; RACP, rat prostatic acid phosphatase; LB, Luria broth; f', real component of the anomalous scattering factor; f'', imaginary component of the anomalous scattering factor; F_o, observed structure factor; F_c, calculated structure factor.

S0006-2960(96)00061-X CCC: $12.00

of the sugar bisphosphate into Fru-6-P and inorganic phosphate (P_i) (Pilkis et al., 1995; Lee et al., 1994a,b).

The bifunctional enzyme is itself subject to modulation by two hormones, glucagon and insulin, both normally associated with the regulation of the blood sugar levels. In addition to the chronic regulation of gene expression that take places in the ranges of minutes to days, glucagon regulation occurs by an acute activation of Fru-2,6-P_2ase activity of the bifunctional enzyme through a cAMP-dependent phosphorylation at Ser32 (Pilkis et al., 1995). The acute control occurs in the ranges of seconds to hours. Ser32 in the bifunctional enzyme somehow interacts with the phosphatase domain. Phosphorylation of the bifunctional enzyme causes increases in the bisphosphatase activity. In the unphosphorylated form of the bifunctional enzyme, the net kinase activity is predominant and higher levels of Fru-2,6-P_2 are present (Pilkis et al., 1995).

To investigate structure/function relationships of rat liver Fru-2,6-P_2ase (residues 251−440) (Lee et al., 1994a), its crystal structure was determined by MAD phasing (Hendrickson, 1991) using a selenomethionine-labeled recombinant form. The overall conformation and the active site region are described in detail below. The single domain enzyme has kinetic characteristics similar to that of the phosphorylated bifunctional enzyme (Lee et al., 1994a,b), and a preliminary hypothesis on the mechanism is proposed. In addition, a comparison is made of Fru-2,6-P_2ase with other members of the histidine phosphatase family such as yeast phosphoglycerate mutase (YPGM) (Campbell et al., 1974) and rat prostatic acid phosphatase (RACP) (Schneider et al., 1993).

MATERIALS AND METHODS

Purification and Crystallization of the Se-Met-Substituted Fructose-2,6-bisphosphatase. The expression plasmid for the rat liver Fru-2,6-P_2ase domain of the bifunctional enzyme (Lee et al., 1994a) was introduced into the *Escherichia coli* methionine auxotroph DL41DE3. An overnight culture was grown in LeMaster media (Hendrickson et al., 1990) enriched by 10% LB media to ensure fast growth of the cells. The Se-Met Fru-2,6-P_2ase was induced in a manner similar to the unlabeled native bisphosphatase domain (Lee et al., 1994a) except LeMaster media was used to guarantee complete substitution of the methionines with selenomethionine (Hendrickson et al., 1990). At an OD_{600nm} of 0.6 and after the ferment was cooled to 20 °C from 37 °C, 0.4 mM IPTG was added to the growth media. Twenty hours after the induction, the cells were harvested.

A cell extract was prepared using a French press followed by centrifugation. Purification of the Se-Met Fru-2,6-P_2ase was done in the same manner as the native Fru-2,6-P_2ase domain (Lee et al., 1994a). The soluble portion of the cell extract was loaded on a Fast Flow Q-Sepharose anion-exchange column and eluted by linear concentration gradient of KCl (50−500 mM). The next step, gel filtration, was carried out on a Sephacryl S-200 column. The final purification step used phenyl-Sepharose chromatography.

The enzymatic activity of the Se-Met Fru-2,6-P_2ase was monitored by determining the extent of formation of a phosphoenzyme intermediate (E-P) (Stewart et al., 1985). The Se-Met Fru-2,6-P_2ase was phosphorylated to the same extent as the native enzyme (data not shown). A total of 68 mg of the pure Se-Met Fru-2,6-P_2ase was obtained from 4 L of media. The Fru-2,6-P_2ase has four Met residues, 54, 85, 128, and 145, in addition to the presence of the N-terminal Met in the recombinant protein. The residue numbers in the text which follows begin with Ser1. Ser1 is equivalent to Ser252 of the bifunctional enzyme.

Although the Se-Met Fru-2,6-P_2ase was less soluble than the native protein, its crystallization occurred using conditions similar to that of the native protein. Both the native and the Se-Met Fru-2,6-P_2ase were crystallized by hanging or sitting drops using a precipitation solution of 30% PEG 4000, 10 mM sodium phosphate, 1 mM dithiothreitol, and 5% glycerol at pH 7.0. Crystallization drops contained a mixture of 5 μL each of the precipitant solution, water, and a protein solution at 15 mg/mL. After 2 weeks, crystals as large as 0.6 mm by 0.4 mm by 0.2 mm were found. The crystals of Fru-2,6-P_2ase were shown to belong to the space group $P2_1$ with unit cell dimensions of $a = 48.4$ Å, $b = 56.6$ Å, $c = 93.7$ Å, and $\beta = 94.4°$. Crystals of space group $P1$ could be achieved under a different condition (Lee et al., 1994a). We estimated that there was 52% of solvent with two protein molecules in the asymmetric unit. The crystals of Fru-2,6-P_2ase produced X-ray diffraction data corresponding to Bragg spacings of 1.8 Å at both the X-12C and X-4A beam lines at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory. However, the usable data were collected to a maximum resolution of 2.5 Å.

X-ray Data Collection and Processing. Multiwavelength diffraction data for MAD phasing of the Se-Met Fru-2,6-P_2ase were collected at the X-4A station at NSLS, Brookhaven National Laboratory. During the data collection, the ring was operated at 2.5 GeV with currents ranging from 250 to 120 mA. A 2 mrad swap of the incident beam was focused in the horizontal direction by the second crystal of a double Si Crystal monochromator and vertically by a Rh-coated spherical mirror. Fuji HR-III phosphorimaging plates were used to record the X-ray data at a distance of 300 mm from the crystal. The data were collected from a single crystal frozen at 95 K by an Oxford Cryostream cooling device. The cryogenic conditioning of both the native and the Se-Met Fru-2,6-P_2ase crystals was done by stepwise soaking of the crystals in 10%, 20%, and 30% glycerol-enriched mother liquor for 20 min each before the crystal was frozen for data collection.

The crystals were aligned such that Bijvoet pairs could be recorded simultaneously on the right and left sides of the image plates (Hendrickson, 1991). Data sets were collected at four different wavelengths, which are the upstream remote (λ_1), the downstream remote (λ_4), the inflection (λ_2), and the maximum of X-ray absorption (λ_3) by K-shell α electrons of Se. Two wavelengths producing the minimum f' (0.9792 Å, λ_2) and the maximum f'' (0.9790 Å, λ_3) anomalous scattering factor components were determined from X-ray fluorescence spectra, which was recorded by a scintillation detector. The two remote wavelengths were at +140 eV (0.9686 Å, λ_4) and −110 eV (0.9873 Å, λ_1) from the λ_3. These wavelengths were chosen to maximize both the Bijvoet and the dispersive differences (Hendrickson, 1991).

A sweep of 200° was collected by rotating around the crystal *b*-axis. The blind cone region was collected by aligning the crystal *a*-axis to the spindle axis and collecting in an inverse beam experiment. In order to minimize the differences among data sets from the effect of varying beam

Table 1: Data Collection and Refinement Statistics of Fru-2,6-P_2ase[a]

(A) Data Collection and Phasing Statistics

data sets (wavelength)	unique reflections	resolution range (Å)	completeness (%)	R_{sym}	phasing power	mean figure of merit
high resolution (1.0400 Å)	15 137	10.0−2.5	93.7	0.068		
λ_1 (0.9873 Å)	7 851	10.0−2.8	75.4	0.046	1.64	
λ_2 (0.9792 Å)	7 829	10.0−2.8	69.7	0.124		
λ_3 (0.9790 Å)	7 875	10.0−2.8	74.4	0.070		0.58
λ_4 (0.9686 Å)	8 132	10.0−2.8	74.2	0.051	1.69	

(B) Refinement Statistics

				rmsd			
resolution range (Å)	total no. of atoms	water molecules	R/R_{free} factor (%)	bonds (Å)	angles (deg)	dihedral (deg)	average B-factor (Å^2)
6.0−2.5	3656	65	20.8/27.3	0.007	1.4	23.1	25.6

[a] $R_{sym} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$, where I = observed intensity and $\langle I\rangle$ = average intensity obtained from multiple observations of symmetry-related reflections. Phasing power = root mean square ($\langle F_h\rangle/E$), where F_h = heavy-atom structure amplitude and E = residual lack of closure error. R-factor = $\sum|F_o - F_c|/\sum F_o$, where F_o and F_c are the observed and calculated structure amplitudes. rmsd is the root mean square deviation from the mean values. A total of 12 891 reflections were used for the refinement, and 1514 reflections randomly selected by X-PLOR were used to evaluate the R_{free} values after each step of the refinement.

intensity, each oscillation frame (typically 1.8−2.5°) was collected at all four wavelengths prior to moving to the next frame.

Reflection data recorded on the image plates were digitized with a Fuji BAS-2000 image scanner. The images were indexed and integrated after profile fitting using the computer program DENZO, and the data from one wavelength were merged and scaled by SCALEPACK (Otwinowski, 1990) with the Bijvoet components processed separately. As can be seen in Table 1, the MAD data were approximately 75% complete at the resolution range of 10.0−2.8 Å with the R_{sym} ranging from 4.6% to 12.4%.

For crystallographic refinement at high resolution, another data set of the Se-Met Fru-2,6-P_2ase was collected at the X-12C beamline at a wavelength of 1.04 Å with beam conditioning similar to that used for collecting the MAD data. The crystal was treated so that the cryogenic data collection procedure as described above could be used. The data were recorded on a MAR image disk, processed, and reduced using DENZO and SCALEPACK (Otwinowski, 1990). The native data covered 10−2.5 Å resolution with an overall merging R_{sym} of 6.8%. It was 93.7% complete with the I/σ cutoff of 3. The results for this data set are also summarized in Table 1 under the heading "high resolution".

RESULTS AND DISCUSSION

MAD Phasing and Structure Refinement. We decided to follow the MIR-style phasing (Blundell & Johnson, 1976) procedures with the MAD data since overall the MAD data were merely 75% complete in the resolution range of 10−2.8 Å. It has been reported that, with a less complete data set, the MAD maps calculated by MIR procedures implemented with a maximum-likelihood algorithm were somewhat more interpretable than the maps calculated by the algebraic formalism (Ramakrishnan et al., 1993). Each of the MAD data sets was then treated as a conventional heavy-atom derivative to exploit the dispersive differences caused by different wavelengths as a source of phase information (Ramakrishnan et al., 1993; Terwilliger, 1994). The data set from λ_2 (0.9792 Å) was not used for phase calculation because of its unusually high R_{sym} value compared with other data sets (see Table 1).

We decided to consider the data set from λ_3 as native data and the data from λ_1 and λ_4 as isomorphous derivatives. This selection was made to keep the signs of the occupancies of the Se sites positive and is a result of the fact that the f'' scattering factor at λ_3 is lowest among the three data sets used for the phase calculation. To carry out MIR-style phase calculation, the data sets from λ_1 and λ_4 were merged and scaled against the pseudonative λ_3 data, using the CCP4 package (Collaborative Computational Project, Number 4, 1994).

The positions of Se atoms were obtained by Patterson analyses using the Bijvoet differences. The Se sites were located from the Bijvoet difference-Patterson map calculated using the data measured at λ_3. The Se positions were confirmed using the $|\lambda_4 - \lambda_3|$ dispersive differences. In principle, the Bijvoet difference at λ_3 and the $|\lambda_4 - \lambda_3|$ dispersive differences should have produced the highest Patterson peaks (Hendrickson, 1991) for the data used. A total of four of the eight expected Se sites were clearly located from the difference-Patterson map. Preliminary phases were then calculated by the maximum-likelihood algorithm (Otwinowski, 1991) implemented in MLPHARE (Collaborative Computational Project, Number 4, 1994) using the four initial Se sites. By cross-Fourier synthesis using the preliminary phases, four additional sites were found. After refinement, a mean figure of merit was 0.58 using data to 2.8 Å resolution (Table 1).

Density modification (Wang, 1985) methods which included the use of noncrystallographic symmetry were applied to improve the initial phases using DM (Collaborative Computational Project, Number 4, 1994). Masks for symmetry averaging were drawn from the initial electron density map using MAID (Levitt & Banaszak, 1993). The correlation coefficient for the dimeric structure was improved to 0.91 from 0.45. The resulting electron density map from the pseudonative data at λ_3 was readily interpretable with major secondary structure and well-defined side chains all visible. A small portion of the MAD map is shown in Figure 1A. It is readily apparent when comparing Figure 1A to the same electron density calculated with refined model phases shown in Figure 1B that the initial map was of high quality. A starting model was built into this electron density map using the program MAID (Levitt & Banaszak, 1993).

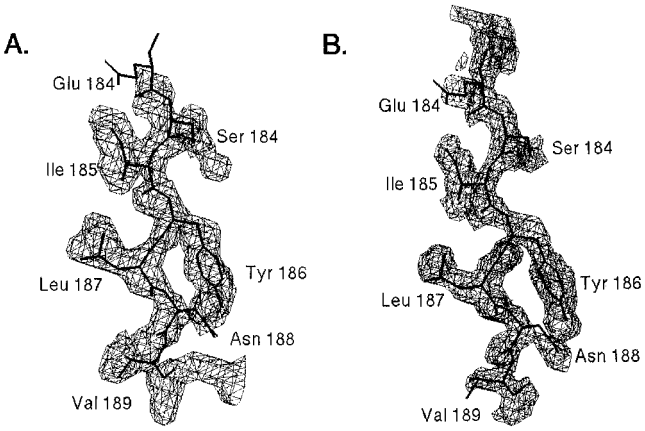

FIGURE 1: Electron density maps drawn in the region of the C-terminal β-strand of Fru-2,6-P_2ase. (A) The electron density map was calculated using phases obtained with the MAD data as described in the text followed by density modification with symmetry averaging. The amplitudes were obtained at λ_3. (B) The $2|F_o| - |F_c|$ map was calculated for the same region using refined model phases. Electron density in map A is contoured at a threshold level of 1.0σ and in map B at 1.8σ. The contours were drawn using the MAID program.

The initial model was improved by using the high-resolution data set and the simulated annealing method implemented in X-PLOR (Brunger, 1989) using a total of 12 891 reflections. The free *R*-factors were used to justify each refinement step with 1430 reflections set aside. After the refinement, the resulting *R*-factor dropped from 0.47 to 0.25. After some manual coordinate adjustments using the program O (Jones et al., 1991), positional refinement by X-PLOR reduced the *R*-factor to 0.240. The symmetry restraints of the two noncrystallographically related protein molecules were released after this stage to provide more degrees of freedom in further refining. At two sites in the refined map where a $|F_o| - |F_c|$ difference peak higher than 15σ was found, inorganic phosphate (P_i) molecules were incorporated into the coordinate list. Next, 65 water molecules also were placed in the electron density maps. The usual criteria for solvent sites were adhered to and included a distance of less than 4.0 Å from a polar protein atom with the σ for electron density higher than 3.5. The incorporation of P_i and water molecules improved both the geometry of the protein structure and the *R*-factor to 0.208 and the free *R*-factor from 0.288 to 0.273. An electron density map based on phases from the refined coordinates is shown in Figure 1B.

The current crystallographic model for the two subunits of Fru-2,6-P_2ase has 3656 protein atoms, 2 phosphates, and 65 solvent sites. As examined by PROCHECK (Laskowski et al., 1993), more than 86% of the residues in the crystallographic model have the main chain dihedral angles in "most favored" regions with the rest in "additional allowed" regions (Laskowski et al., 1993). Exceptions occur at five residues that are found in "generously allowed" regions. The five residues are Leu64, Glu34, and Cys140 of subunit A and Arg107 and Cys140 of the B subunit. Except for Cys140, they are all involved in sharp turns. Cys140 is at the C-terminus of β-strand β4 and is adjacent to the catalytic residue His141. His141 will be discussed in more detail in the section describing the active site. The dihedral angles of Cys140 were not improved by the refinement in spite of the well-defined electron density. The refined ϕ, ψ values of Cys140 in molecule A were 127° and 139°, respectively. It is unclear whether or not the unusual configuration of Cys140 is crucial for the proper conformation of the catalytic site.

Table 2: Secondary Structure in Fru-2,6-P_2ase[a]

	residue numbers in			residue numbers in	
α-helix	Fru-2,6-P_2ase	bif enzyme	β-sheet and loop	Fru-2,6-P_2ase	bif enzyme
α1	11−14	262−265	β1	2−6	253−257
α2	26−42	277−293	β2	48−50	299−301
α3	55−63	306−314	β3	69−71	320−322
α4	72−74	323−325	β4	136−140	387−391
α5	87−93	338−344	β5	169−176	420−427
α6	95−102	346−353	β6	179−186	430−437
α7	116−131	367−382			
α8	142−152	393−403	loop 1	8−21	259−272
α9	157−162	408−413	loop 2	80−110	331−361

[a] Residue Ser1 of Fru-2,6-P_2ase is the same as residue Ser252 of the bifunctional enzyme.

The average thermal factor (*B*-factor) for all of the protein atoms is 25.6 Å^2, and high *B*-factors are found at the residues involved in turns or at the protein surface. For help with the discussion that follows, the position of secondary structure in the crystallographic model is defined in terms of residue numbers in Table 2. The highest *B*-factors are found at the end of helix α3 and the beginning of β-strand β3 (residues 64−70). β3 is the strand at one edge of the central β-sheet. High temperature factors are also found at the end of helix α6 including residues 103−106.

Details regarding discrepancies from canonical model values are summarized in Table 1B. All are within acceptable ranges. A more critical and experimental evaluation of the coordinate errors can be obtained by comparing the two crystallographically independent molecules. The final model has sufficient degrees of freedom to evaluate the errors, since the symmetry restraints for the refinement of the two protein molecules were removed after the *R*-factor reached 0.24 as mentioned in the previous section. After rotating one set of coordinates by the 2-fold rotational symmetry found in the crystal, the agreement between the two models has been checked. After superposition of all the main chain atoms, the two independent crystallographic subunits agreed with a positional rmsd of 0.68 Å and 4.23 Å^2 for the *B*-factors.

Overall Structure of the Fructose-2,6-bisphosphatase. The polypeptide conformation of the rat liver Fru-2,6-P_2ase is shown as a ribbon diagram in Figure 2, and as noted above, the secondary structural elements are listed in Table 2. The combination of β-structures and α-helices is typical of a nucleotide-binding fold. However, major structural excursions occur in the formation of the nucleotide-binding domain so that the components are spread throughout much of the linear sequence. Furthermore, although the core β-sheet includes six strands, one of them is antiparallel β5. In terms of viewing, the β-strands appear to be somewhat longer than as listed in Table 2 where, for example, β3 only includes amino acids 69−71. This is because of the requirement for a hydrogen-bonding ladder in defining a β-strand (Kabsch & Sander, 1983). In nearly all of the strands, ϕ, ψ values alone would include extra residues.

Since the Fru-2,6-P_2ase domain is derived from the C-terminal portion of the bifunctional enzyme, the location

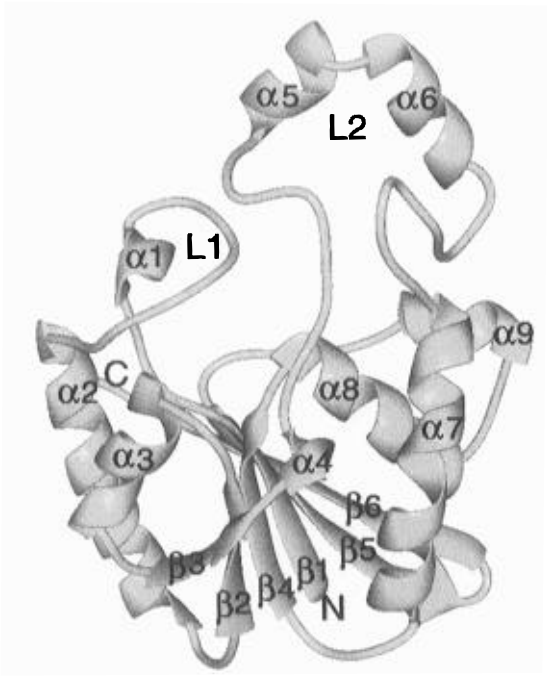

FIGURE 2: Crystal structure of the rat liver Fru-2,6-P_2ase. The secondary structures of Fru-2,6-P_2ase are drawn in cartoon fashion with arrows for β-strands and coils for α-helices. The nine α-helices are labeled from α1 to α9. The central β-sheet contains six strands which are labeled from β1 to β6. The two loops constituting the active pocket are labeled L1 and L2 and are discussed in more detail in the text.

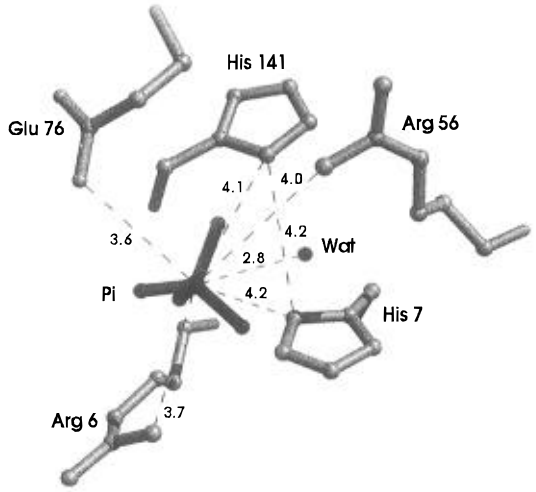

FIGURE 3: Inorganic phosphate site in crystalline Fru-2,6-P_2ase. The ball and stick cartoon depicts the protein atoms interacting with the bound P_i, which is a competitive inhibitor of Fru-2,6-P_2ase. This is believed to be the active site of the enzyme. The five side chains were selected by listing the atoms within 5 Å of the bound P_i and include Arg6, His7, Arg56, Glu76, and His141. A water molecule with a *B*-factor of 30.25 Å^2 is shown close to both the bound P_i and His7. Distances between certain atom pairs are indicated in angstroms.

of the N-terminus is of special interest and could crudely place the location of the kinase domain. On the basis of the crystallographic model, the kinase domain would be located near the N-terminus which is at the bottom and right of Figure 2. This places the kinase region well away from the active site of the phosphatase domain. The phosphatase active site is situated near the C-termini of the three parallel β-strands (β1, β4, and β2) and the N-termini of the four helices (α2, α3, α7, and α8). As can be seen in Figure 2, the two antiparallel strands nearest the C-terminus, β5 and β6, are important to the core β-sheet structure. Earlier studies have demonstrated that C-terminal truncation mutagenesis beyond Val189 abolished the bisphosphatase activity (Y.-H. Lee, unpublished data).

Two large loop structures, loop 1 (residues 9−22) and loop 2 (residues 77−106), constitute a wide arch above the active site by forming strong main chain interactions with each other (Figures 2 and 4). Thus, the top of the core β-sheets and the two loops together lead to a funnel-shaped active pocket. The catalytic site is near the bound inorganic phosphate as shown in Figure 3. It will be described in more detail below. The opening of the funnel at the surface of the protein has a diameter of about 15 Å and appears to be the point where substrates would enter the catalytic site. The opening on the narrow side of the funnel, more toward the center of the molecule, has a diameter of about 6−7 Å and appears to be too small for substrate passage. It would be wide enough for one water molecule to pass through.

Loop 1 and loop 2 position a number of residues close to the active site, and these residues presumably are involved in binding of the substrate and the product inhibitors. Although loop 1 contains only a short helix, α1, loop 2 ends in two consecutive helices, α5 and α6. Because of the importance of these two loops to the juxtapositioning of side chains at the active site, we have tried to examine the interactions that bring them together. The interactions between the two loop segments are summarized in Figure 4.

A strong ion pair is formed between Arg 17 and Glu89. Separated by a slightly longer distance is the ion pair formed between Arg17 and Glu84. In addition, the main chains of the two loops come close enough at one point that main chain hydrogen bonds are formed between the carbonyl oxygen of Ile18 and the peptide nitrogen of Glu84, between the carbonyl oxygen of Met85 and the peptide nitrogen of Ile18, the finally between the peptide nitrogen of Gly20 and the side chain of Glu83. Together loop 1 and loop 2 form a side of the funnel-shaped active pocket.

The Active Site. Previous modeling studies based on the structure of YPGM have suggested that the active site resides are on the C-terminal end of the four parallel β-strands and near the imidazole rings of two key histidine residues, His-7 and His-141 (Bazan et al., 1989; Campell et al., 1974). This hypothesis was confirmed by the present crystallographic studies based on the position of P_i, as shown in Figure 3. Inorganic phosphate is a competitive inhibitor (K_i of 74 μM) in the phosphatase reaction. It is located 4.2 Å distant from ND1 of His7 and 4.1 Å from ND1 of His141. Although these imidazole rings in Fru-2,6-P_2ase are not parallel as in YPGM (Figures 3 and 7), they are near to each other with a distance of 4.2 Å between ND1 of His7 and ND1 of His141. Their orientation is shown in Figure 3. Glu76 is also located near the bound P_i, and the distance between OE1 of Glu76 and the bound P_i is 3.6 Å. Glu76 is located about 5 Å from ND1 of His141. This is farther than expected to form a typical histidine carboxylate proton shuttle.

The results from previous site-directed mutagenesis studies also suggest that, besides His7 and His141, Glu76 is also a key residue for the dephosphorylation reaction (Tauler et al., 1992; Lin et al., 1992a). In fact, these earlier chemical studies have shown that His7 is transiently phosphorylated to form a covalent phosphoenzyme intermediate (E-P) during

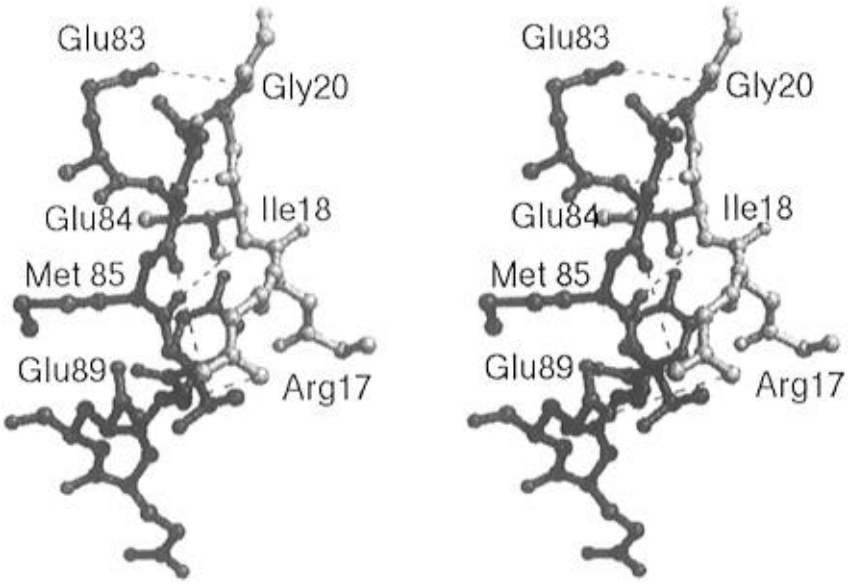

FIGURE 4: Stereoview of the interactions between loop 1 and loop 2. The cross-eyed stereo ball and stick drawing depicts the juxtapositioning of loops 1 and 2 in the crystal structure of Fru-2,6-P_2ase. Loop 1 includes residues 17−20 and is drawn with light shading. Loop 2 includes residues 83−89, drawn with heavy shading. Forming a significant part of the active pocket, the two loops interact with each other through the hydrogen bonds represented by dotted lines. These interactions include hydrogen bonds between O of Ile18 and N of Glu84 and N of Ile18 and O of Met85, between side chains of Arg17 and Glu89, between main chain atoms of N of Gly20 and Glu84 and side chains of Arg17 and Glu83.

a catalytic cycle and that His141 activates the E-P formation (Tauler et al., 1992). It is also known that mutation of Glu76 diminishes the catalytic activity with no significant effect on E-P formation (Lin et al., 1992a). The steric relationship among these three residues in the crystal structure confirms their overall proximity to each other and the bound phosphate ion.

The crystal structure also points to Arg6 and Arg56 as being important for substrate binding. As is apparent in Figure 3, Arg6 and Arg56 are located within hydrogen-bonding distances to oxygens belonging to the bound P_i. The structural data suggest that the same two side chains would serve as ligands to the 2-phosphate of the Fru-2,6-P_2. In trial substrate-docking models where the 2-phosphate was positioned at the P_i site, Arg6 and Arg56 were not close enough to interact with the phosphate at the 6-position. The functional role of these arginines as contributing to substrate binding is strengthened by the fact that they are conserved in other members of the histidine phosphatase family (Bazan et al., 1989; Fothergill-Gillmore & Watson, 1989). Furthermore, their functional role in Fru-2,6-P_2ase was suggested by earlier site-directed mutagenesis studies (Lin et al., 1992b).

Fru-2,6-P_2 has a K_m of 4 nM while Fru-6-P is a noncompetitive inhibitor with a K_i of 157 nM (Lee et al., 1994a). This suggests simultaneous recognition by the enzyme of both phosphate moieties. Earlier site-directed mutagenesis studies on the basic amino acid residues at loop 2 have demonstrated that Arg101 and Lys105 were involved in binding of the phosphate at the 6-position of both Fru-2,6-P_2 and Fru-6-P (Lin et al., 1992a,b). The crystal structure with the docked substrate confirms these assignments and adds the nearby presence of the side chain of Arg107. Thus, the 6-phosphate moiety of Fru-2,6-P_2 would be located midway between Arg101, Arg107, and Lys105. Arg109 reported earlier to be near this portion of the binding site is in fact a surface residue in the crystal structure (Li et al., 1992b). Crystal soaking experiments using Fru-6-P are underway to experimentally confirm the location of the residues interacting with the 6-phosphate moiety.

The Catalytic Site and a Reaction Mechanism. The crystal structure of Fru-2,6-P_2ase with substrate docked by modeling so that the 2-phosphate is positioned at the inorganic phosphate site generally agrees with functional groups defined in an earlier mechanism derived from site-directed mutagenesis and kinetic studies (Pilkis et al., 1995). A schematic illustration of the active site as it might be involved in catalysis is shown in Figure 5. The 2-P of Fru-2,6-P_2 would interact with Arg6 and Arg56 at the active site. Both are between 3.0 and 3.5 Å from oxygens covalently linked to the 2-phosphate of the substrate. After binding of the substrate, the next stage of the catalytic mechanism involves several different steps. The order in which they occur is unclear, but the net result is the formation of a phosphohistidine intermediate.

The initial two steps in the reaction pathway are outlined in Figure 5A. (1) A nucleophilic attack by ND1 of the unprotonated imidazole ring of His7 occurs at the 2-P position of Fru-2,6-P_2. This would result in the formation of a phosphoamide bond (E-P formation) at His7. The E-P formation at His7 was identified by labeling the bifunctional enzyme with the radioactive substrate and amino acid sequencing of the proteolytic fragments (Pilkis et al., 1987). (2) A proton provided by the protonated form of His141 would add to the anionic O-2 of the leaving Fru-6-P. As noted above, the order in which these two intermediate reactions occur is unknown. Protonation of oxygen (O-2) could occur before, after, or simultaneously with phosphoamide formation. Considering that Fru-2,6-P_2 is labile to acidic conditions and that the His7Ala mutant still has about 5% the activity of the wild type, (2) could precede (1). Furthermore, it is important to note that, in the crystal structure, the inorganic phosphate is located equidistant from His7 and His141. Without evidence to the contrary, either

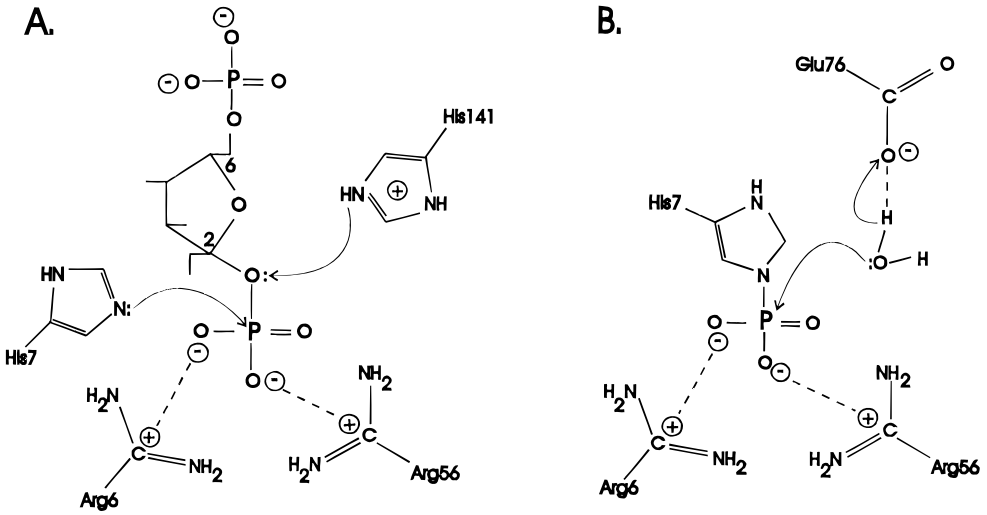

FIGURE 5: Hypothetical mechanism of the catalytic reaction for Fru-2,6-P_2ase. Two stages of the mechanism are shown. (A) The 2-P of Fru-2,6-P_2 is positioned close to the two catalytic His residues, His7 and His141, by the charge attraction provided by Arg6 and Arg56 and represented by dotted lines. (1) Nucleophilic attack of His7 at the 2-P leads to the formation of the E-P covalent phosphoenzyme intermediate and produces a water molecule (not shown). (2) His141 donates a proton to the 2-O of the leaving Fru-6-P. (B) An OH^- ion is formed from a water molecule by the unprotonated form of Glu76. The E−P is broken down by nucleophilic attack of the OH^- ion, producing an inorganic phosphate molecule. During this process, a proton harvested by Glu76 from the water molecule could be recycled by transfer to His141.

histidine could be the one that is phosphorylated initially during a catalytic cycle.

In a later stage of the catalytic pathway, it has been proposed that a water molecule is activated by Glu76 (Lin et al., 1992a). This is illustrated in Figure 5B. The hydroxyl ion formed by the interaction of water and the carboxylate of Glu76 attacks the phosphohistidine and is responsible for the release of inorganic phosphate (Lin et al., 1992b). In the hypothetical model generated with the crystal structure and a model of Fru-2,6-P_2, Glu76 is located at the active site but is closer to His141 than to His7. This fits into the reaction scheme because the proton generated by water activation could be used to reprotonate His141 (see Figure 5A). A water molecule with a *B*-factor of 30.3 $Å^2$ is found in the active site of the crystal structure and was shown in Figure 3. It could be the water indicated in Figure 5B.

Dimeric Interface in the Fru-2,6-P_2ase Domain. Although the intact bifunctional enzyme is active only as a dimer, the Fru-2,6-P_2ase domain is active as a monomer (Lee et al., 1994a; Pilkis et al., 1995). However, it also has been observed that the bisphosphatase dimerizes at high protein concentrations. We believe that the two subunits found in the crystallographic asymmetric unit may represent the dimeric form of the bisphosphatase domain and perhaps the interface contained in the bifunctional enzyme. The crystallographic dimer is shown in Figure 6. As is typical in homodimers, the structural relationship between the two subunits is defined by the point symmetry group−*C*2. The local dyad, also depicted as a dotted line between the two small balls in Figure 6, is tilted 17° from the crystallographic *b*-axis toward the *a*-axis. In this dimer, the active sites of the two protein molecules face each other, and the distance between the two bound inorganic phosphates, which are depicted as large balls in Figure 6, is 24.1 Å.

Although only the C_α atoms are shown in Figure 6, most of the intersubunit interactions appear to involve side chains. Some of these polar interactions occur between Glu88 and Arg6, Gln142 and Gln91, Glu92 and Arg27, His93 and Tyr186, and finally His98 and Glu92. There is a main chain−side chain interaction involving Lys164 of one subunit and Pro95 of the other. Overall, the main chains of each subunit are closest together around Glu92 of one subunit and Pro166 of another.

It is unproven whether the dimeric interaction observed in the crystalline state of the Fru-2,6-P_2ase domain would also be present in the dimeric form of the bifunctional enzyme. However, in steric terms, the quaternary structure shown in Figure 6 has space above and below for the kinase domain. This being the case, we would predict little or no intersubunit interactions that would involve the kinase domain in the bifunctional enzyme.

Comparison with YPGM and RACP. The structure of the Fru-2,6-P_2ase domain was compared with other proteins to understand the structure/function relationships. Among them were YPGM, RACP, and Fru-1,6-P_2ase. Although both the Fru-2,6-P_2ase domain and Fru-1,6-P_2ase are sugar bisphosphatases, no significant structural similarity was found. That was not surprising since Fru-1,6-P_2ase (335 residues) not only is longer than the Fru-2,6-P_2ase domain (220 residues) but also has no significant amino acid sequence homology.

Knowing of the functional similarity and the amino acid sequence homology, the crystal structure of Fru-2,6-P_2ase was compared with that determined for YPGM and RACP. To illustrate the conformational similarity between Fru-2,6-P_2ase and YPGM, which shows an amino acid sequence homology of 36% with the bisphosphatase domain, a stereo representation of the two C_α models is shown in Figure 7. The amino acid sequence homology following the conformational alignment is shown in Figure 8. Two short helical segments, labeled "a" in Figure 8, had similar but not helical conformations in YPGM. For ease of viewing, the crystal structure of RACP is not shown in Figure 7. When the four equivalent parallel β-strands in the core structures of YPGM and RACP were superposed on those of the Fru-2,6-P_2ase, the C_α's in this region agree very well. The rmsd was 0.45 and 0.85 Å for YPGM and RACP, respectively. Many of the remaining β-strands and α-helices were also superimposable. However, as can be seen in Figure 7, conforma-

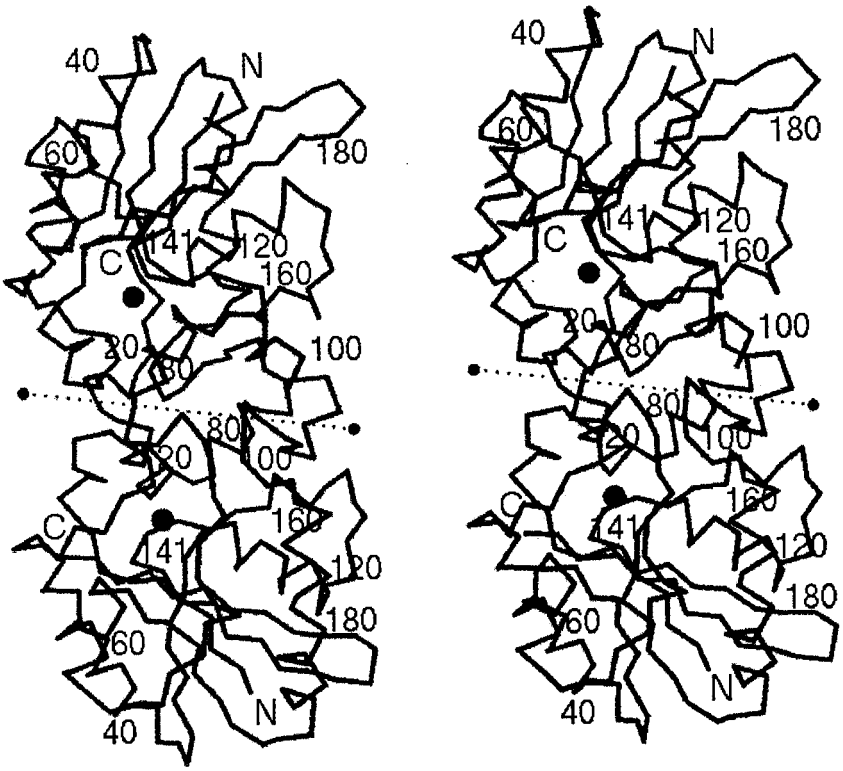

FIGURE 6: Rat liver Fru-2,6-P_2ase dimer. The stereo drawing in cross-eyed mode shows how the two Fru-2,6-P_2ase molecules are related in the crystalline state. Only C_α atoms are shown and every 20th residue is numbered. Although the Fru-2,6-P_2ase domain exists as a monomer in solution, it forms a dimer in the crystalline state. The crystal dimer has *C*2 symmetry, and the local dyad is illustrated by the dotted line. Two large spheres represent the P_i molecules bound to the active sites. Refer to the text for additional details.

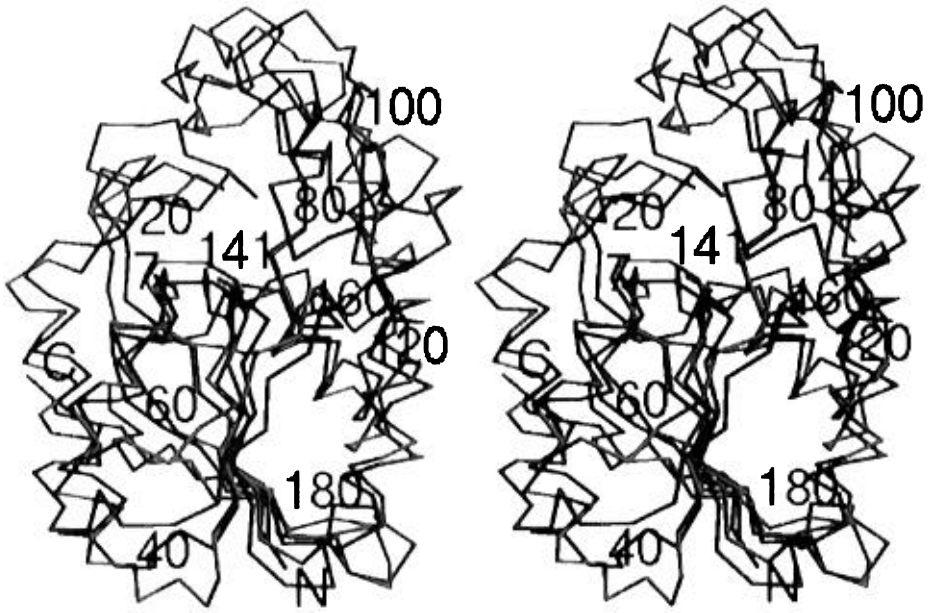

FIGURE 7: Cross-eyed stereoview of the superimposed structures of Fru-2,6-P_2ase and YPGM. The C_α backbones of Fru-2,6-P_2ase (red) and YPGM (blue) are shown after the crystal coordinates are superimposed using least squares methods for the core four parallel β-strands. The YPGM coordinates had a PDB accession code 3PGM. The stereo diagram shows well-matched secondary structure for the core region and many of the helices. Note the conformational differences for the two catalytic histidines of both proteins labeled as 7 and 141. The structure of YPGM has a major insertion which includes residues 122 through 147. The loop−loop interaction near residue 20 in Fru-2,6-P_2ase is also found in YPGM despite the differences in loop 2.

tional differences are apparent. In the course of the comparisons, one unusual observation was made. The helices in RACP that are equivalent to α3,4 or Fru-2,6-P_2ase seem to be pushed toward the core β-sheet in RACP.

The conformational homology between Fru-2,6-P_2ase and YPGM breaks down at nearly all of the surface turns. In addition, between Lys112 and Gly113 in Fru-2,6-P_2ase, there is a major insertion in YPGM. In YPGM, the loop equivalent to loop 2 of Fru-2,6-P_2ase includes residues Pro122 to Thr148. This 25 residue insertion in YPGM should be visible toward the front in Figure 7.

At the catalytic sites of each enzyme, two histidines are placed similarly. The catalytic histidine residues of Fru-2,6-P_2ase (His7, 141) and RACP (His12, 278) have similar

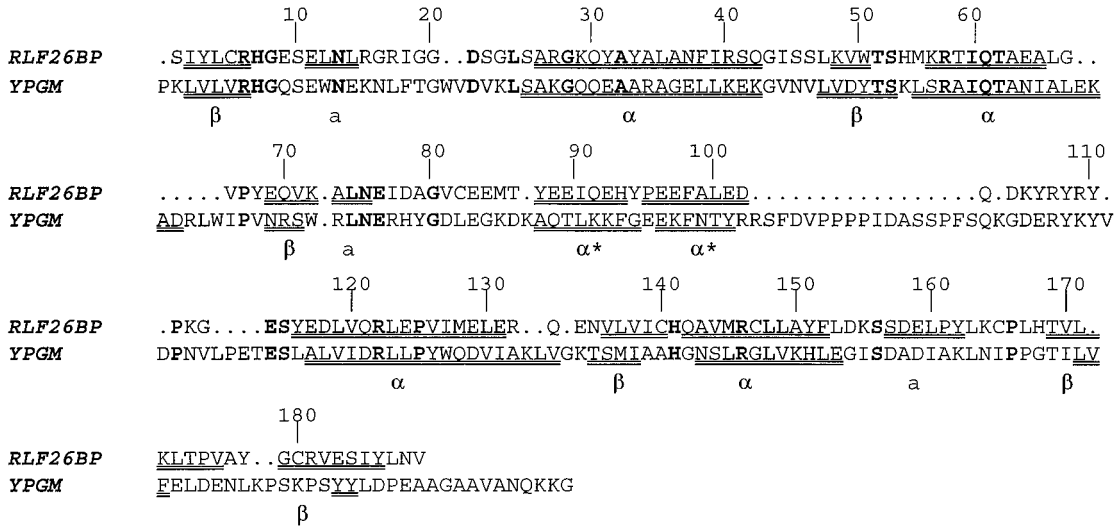

FIGURE 8: Comparisons of Fru-2,6-P_2ase and YPGM. The amino acid sequences were aligned such that the secondary structures of both proteins overlapped as based on the conformational comparison. The structural overlaps are emphasized by double underlines. The amino acids known to be conserved throughout the histidine phosphatase family were matched and are shown in bold type. The rest are aligned on the basis of the sequence similarity, and the gaps were filled with dots. The secondary structures conserved in both of the proteins are labeled with Greek symbols, and the short helices are shown only in the bisphosphatase with the letter "a". The residue numbers are on the top of the sequences.

rotamers. However, the imidazole rings of Fru-2,6-P_2ase are not facing each other as in the YPGM where the two histidines (His8, 179) are in the so-called "clapping hands" conformation (Bazan et al., 1989). The side chains for these two important residues in Fru-2,6-P_2ase and YPGM are included in Figure 7.

Loop 1 in the three structures appears to have similar conformations, which may indicate that this is a crucial structural segment for the histidine phosphatases. Reflecting variations of lengths and sequence, the largest structural variation among the three proteins is in loop 2, which probably plays a major role in determining the specificity for substrates. Interestingly and in spite of differences, all three proteins conserve the same sort of loop−loop interactions which were described in the previous section. Recall that these interactions form a portion of the funnel-shaped active pocket. Hence, the variations in loop 2, away from the site of interactions with loop 1, may be the key to substrate specificity.

CONCLUSIONS

Although it is a relatively slow enzyme with a turnover number of 10.7/min, the Fru-2,6-P_2ase domain of the bifunctional kinase/phosphatase enzyme is of great importance for the control of glucose metabolism. Amino acid sequence homology and now structural data have shown that the phosphatase enzyme is homologous with phosphoglycerate mutases and acid phosphatases. The similarity among this group of enzymes is based on the fact that they are phosphoryl-transfer proteins. The mutases move the phosphate moiety from one position to another while the phosphatase transfers the organic phosphate to an OH^- ion to form inorganic phosphate. In the case of the mutases and Fru-2,6-P_2ase, it has been shown that the phosphoryl transfer occurs through a phosphohistidine intermediate (Fothergill-Gillmore & Watson, 1989).

In the crystal structure of Fru-2,6-P_2ase described in this report, the overall conformation of the protein and the steric relationship of two histidine side chains confirm the overall structural and mechanistic similarity in this class of phosphoryl-transfer proteins. The crystals of Fru-2,6-P_2ase were grown in the presence of 10 mM phosphate, and the electron density map contained a region that is best explained by the presence of a bound phosphate. The phosphate position was assumed to mark the location of the 2-phosphate of the substrate and was used to model build Fru-2,6-P_2 into the active site. A catalytic mechanism already suggested by chemical and mutagenic studies was reinforced. It involves the two clapping hands histidines—one serving as a proton donor to the ester oxygen and the other as a nucleophile which involves the formation of a phosphohistidine as part of the enzyme. On the basis of the crystal structure, it should now be possible to begin to design compounds that might serve as Fru-2,6-P_2ase inhibitors and suppress the effects of phosphorylation on the bifunctional enzyme. Such compounds should mimic the effects of insulin.

ACKNOWLEDGMENT

This work is dedicated to the memory of Dr. Simon J. Pilkis, who was the central figure in these studies. He devoted his research career to studying diabetes and the factors important in glucose metabolism. Dr. Pilkis died suddenly and unexpectedly in August of 1995. Y.-H.L. is grateful to Dr. Wayne A. Hendrickson at Columbia University for advice in the use of selenomethionine and the application of MAD phasing. The authors thank Ed Hoeffner for maintaining the X-ray and computing facilities at the University of Minnesota. We are grateful to Todd Weaver and Jim Thompson for advice in the application of CCP4, X-PLOR, and O software. We are also grateful to Dr. Bob Sweet and his staff at X-12C and Mr. X. Yang at X-4A at the National Synchrotron Light Source at Brookhaven National Laboratory for helping with data collection. X-4A station is supported by the Howard Hughes Medical Institute, and X-12C is supported partly by the National Science Foundation.

REFERENCES

Bazan, J. F., Fletterick, R. J., & Pilkis, S. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9642−9646.

Blundell, T. L., & Johnson, L. N. (1976) *Protein Crystallography*, Academic Press Inc., Olando, FL.

Brunger, A. (1989) *Acta Crystallogr. A45*, 50−61.

Campbell, J. W., Watson, H. C., & Hodgson, G. I. (1974) *Nature 250*, 301−303.

Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.

Fothergill-Gillmore, L. A., & Watson, H. C. (1989) *Adv. Enzymol. 62*, 227.

Hendrickson, W. A. (1991) *Science 254*, 51−58.

Hendrickson, W. A., Horton, J. R., & LeMaster, D. M. (1990) *EMBO J. 9*, 1665−1672.

Jones, T. A., Zou, J.-Y., Cowan, S., & Kijeldgarrd, M. (1991) *Acta Crystallogr. A47*, 110−119.

Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.

Lee, Y.-H., Lin, K., Okar, D., Alfano, N. L., Sarma, R., Pflugrath, J. W., & Pilkis, S. J. (1994a) *J. Mol. Biol. 235*, 1147−1151.

Lee, Y.-H., Okar, D., Lin, K., & Pilkis, S. J. (1994b) *J. Biol. Chem. 269*, 11002−11010.

Lee, Y. H., Picardt, F., & Pilkis, S. J. (1995) *Arch. Biochem. Biophys. 332*, 357−360.

Levitt, D. G., & Banaszak, L. J. (1993) *J. Appl. Crystallogr. 26*, 736−745.

Li, L., Lin, K., Correia, J. J., & Pilkis, S. J. (1992a) *J. Biol. Chem. 267*, 16669−16675.

Li, L., Lin, K., Pilkis, J., Correia, J. J., & Pilkis, S. J. (1992b) *J. Biol. Chem. 267*, 21588−21594.

Lin, K., Li, L., Correia, J. J., & Pilkis, S. J. (1992a) *J. Biol. Chem. 267*, 6556−6562.

Lin, K., Li, L., Correia, J. J., & Pilkis, S. J. (1992b) *J. Biol. Chem. 267*, 19163−19171.

Otwinowski, Z. (1990) *The DENZO Manual*, Yale University, New Haven, CT.

Otwinowski, Z. (1991) *Proceedings of the CCP4 Study Weekend Jan 25−26, 1991* (Wolf, W., et. al., Eds.) pp 80−86, Daresury Laboratory, Warrington WA 4AD., U.K.

Pilkis, S. J., Lively, M. O., & El-Maghrabi, M. R. (1987) *J. Biol. Chem. 262*, 12672−12675.

Pilkis, S. J., Claus, T. H., Kurland, I. J., & Lange, A. J. (1995) *Annu. Rev. Biochem. 64*, 799−835.

Ramakrishnan, V., Finch, J. T., Granziano, V., Lee, P. L., & Sweet, R. M. (1993) *Nature 362*, 219−223.

Schneider, G., Lindquist, Y., & Vihko, P. (1993) *EMBO J. 13*, 2609−2615.

Stewart, H., El-Maghrabi, M. R., & Pilkis, S. J. (1985) *J. Biol. Chem. 260*, 12935.

Tauler, A., Lin, K., & Pilkis, S. J. (1990) *J. Biol. Chem. 265*, 15617−15622.

Terwilliger, T. C. (1994) *Acta Crystallogr. D50*, 17−23.

Wang, C. (1985) *Methods Enzymol. 115*, 90−112.

BI9600613