Crystal Structure of Rabbit Phosphoglucose Isomerase, a Glycolytic Enzyme That Moonlights as Neuroleukin, Autocrine Motility Factor, and Differentiation Mediator[†,‡]

Constance J. Jeffery,[§,||] Brian J. Bahnson,[§,⊥] Wade Chien, Dagmar Ringe, and Gregory A. Petsko*

Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02454-9110

Received July 12, 1999; Revised Manuscript Received November 10, 1999

ABSTRACT: The multifunctional protein phosphoglucose isomerase, also known as neuroleukin, autocrine motility factor, and differentiation and maturation mediator, has different roles inside and outside the cell. In the cytoplasm, it catalyzes the second step in glycolysis. Outside the cell, it serves as a nerve growth factor and cytokine. We have determined the three-dimensional structure of rabbit muscle phosphoglucose isomerase complexed with the competitive inhibitor D-gluconate 6-phosphate by X-ray crystallography at 2.5 Å resolution. The structure shows that the enzyme is a dimer with two α/β-sandwich domains in each subunit. The location of the bound D-gluconate 6-phosphate inhibitor leads to the identification of residues involved in substrate specificity (Ser209, Ser159, Thr214, Thr217, and Thr211). The results of previously published kinetic studies suggest that a lysine and a histidine are involved in the catalytic mechanism. The crystal structure suggests active site residues Lys518 and His388 might be these residues. In addition, the positions of amino acid residues that are substituted in the genetic disease nonspherocytic hemolytic anemia suggest how these substitutions can result in altered catalysis or protein stability.

Phosphoglucose isomerase (PGI,[1] EC 5.3.1.9) is a cytosolic enzyme that catalyzes the interconversion of D-glucose 6-phosphate and D-fructose 6-phosphate, an essential reaction of glycolysis and gluconeogenesis. It is found in eukaryotes, bacteria, and archaea, with the exception of some obligate intracellular parasites such as *Rickettsia* (*1*).

PGI is active as a dimer with a subunit molecular mass of approximately 55 kDa. Its enzymatic activity has been characterized extensively (*2*−*4*). The proposed mechanism for sugar isomerization involves several steps and is thought to occur via general acid/base catalysis. Since glucose 6-phosphate and fructose 6-phosphate exist predominantly in their cyclic forms, PGI is believed to catalyze first the opening of the hexose ring to yield the straight chain form of the substrates. Glucose 6-phosphate and fructose 6-phosphate then undergo isomerization via formation of a *cis*-enediol intermediate with the double bond located between C-1 and C-2. This mechanism contains features of the mechanisms for two other sugar isomerases, triose phosphate isomerase, which uses a proton-transfer mechanism via a *cis*-enediol intermediate, and xylose isomerase, which catalyzes the ring opening of its sugar substrate. A crystal structure of PGI could reveal the structural basis for its mechanism.

A crystal structure could also elucidate the structural basis for the enzyme's strict specificity for the substrates glucose 6-phosphate and fructose 6-phosphate. The enzyme can also catalyze epimerization of mannose 6-phosphate, but at a rate that is orders of magnitude slower than the isomerization of glucose 6-phosphate and fructose 6-phosphate (*5*). Many four- to six-carbon phosphosugars, such as gluconate 6-phosphate (*6*), erythrose 4-phosphate (*7*), and related compounds, serve as competitive inhibitors of the reaction. Unphosphorylated sugars are not substrates and do not serve as good competitive inhibitors.

Remarkably, this essential cytosolic enzyme also has roles outside the cell, where it has been observed to function as a cytokine and growth factor. Its sequence appears to be identical to that of neuroleukin (NL) (*8*, *9*), autocrine motility factor (AMF) (*10*), and differentiation and maturation mediator (DMM) (*11*). Neuroleukin is secreted by T cells and promotes the survival of some embryonic spinal neurons and sensory nerves. It also causes B cells to mature into antibody-secreting cells (*12*, *13*). AMF is a product of tumor cells that stimulates cancer cell migration and may be involved in cancer metastasis and invasion (*10*). DMM was isolated from T cell culture media and shown to cause in

† This research was supported by NIH Grant GM32415 to G.A.P. and D.R. and a Cystic Fibrosis Foundation postdoctoral fellowship to C.J.J.

‡ Coordinates for the structure of phosphoglucose isomerase complexed with the competitive inhibitor gluconate 6-phosphate have been deposited in the Protein Data Bank under file name 1DQR. The cDNA sequence for rabbit phosphoglucose isomerase has been deposited in GenBank under file name AF222069.

* To whom correspondence should be addressed. E-mail: petsko@binah.cc.brandeis.edu. Telephone: (781) 736-4903. Fax: (781) 736-2405.

§ The first two authors contributed equally to this work.

|| Current address: Laboratory for Molecular Biology, Department of Biology, University of Illinois at Chicago, Chicago, IL 60607.

⊥ Current address: Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716.

[1] Abbreviations: PGI, phosphoglucose isomerase; MIR, multiple isomorphous replacement; EMP, ethyl mercuric phosphate; PCMB, *p*-chloromercuribenzoate; PCR, polymerase chain reaction; PDB, Protein Data Bank; RACE, rapid amplification of cDNA ends; rms, root-mean-square; AMF, autocrine motility factor; NL, neuroleukin; DMM, differentiation and maturation mediator.

10.1021/bi991604m CCC: $19.00
Published on Web 01/14/2000

vitro differentiation of human myeloid leukemia HL-60 cells to terminal monocytic cells (*11*).

Several experiments have been performed to test if the same polypeptide chain could indeed be carrying out the intracellular and extracellular functions. It was shown that purified rabbit PGI, but not three other glycolytic enzymes (hexokinase, phosphofructokinase, or phosphoglycerate mutase), can cause the increase in cell motility seen with AMF; further, AMF exhibits isomerase enzymatic activity (*10*). In addition, gluconate 6-phosphate and erythrose 4-phosphate block the induced cell motility but not the basal migration rate. DMM was also shown to have isomerase activity, and PGI also causes the dosage-dependent differentiation of human leukemia cells seen with DMM (*11*). PGI has no obvious signal sequence that would direct it to be secreted, and the mechanism by which this cytosolic housekeeping enzyme is brought outside the cell remains to be established. To perform its cytokine and growth factor functions, it is likely that PGI/NL/AMF/DMM binds to a cell surface receptor on target cells. One such receptor has been cloned from fibrosarcoma cells, which have been studied for their response in assays for AMF activity (*14*).

Crystallographic studies of PGI were begun 30 years ago, long before cloning allowed many different species to be surveyed for easily crystallizable proteins. A 3.5 Å resolution tracing of the fold of the polypeptide chain of pig PGI was reported in 1977 (*15*), but the absence of a sequence made it impossible to verify the tracing or to complete the structure determination. In 1981, the resolution was extended to 2.6 Å, but once again the lack of sequence information prohibited refinement of the structure (*16*). As this paper was being prepared for publication, a report appeared of the crystal structure of PGI from *Bacillus stearothermophilus* (*17*). The bacterial PGI structure has no bound ligand. The authors were able to show that the purified bacterial enzyme had activity in an AMF assay.

The structure of PGI with a bound competitive inhibitor is important for elucidating the mechanism of its catalytic activity and substrate specificity. The structure of a mammalian PGI is also important for understanding how the enzyme can serve as a growth factor and cytokine and how amino acid substitutions can lead to nonspherocytic hemolytic anemia. We present here the cloning, sequencing, and X-ray crystal structure determination of rabbit PGI to 2.5 Å resolution with bound gluconate 6-phosphate.

MATERIALS AND METHODS

Sequencing of the Rabbit PGI Gene. Plasmid pPGI, which encodes the pig PGI gene, was a gift from V. Claes (Lille, France). A λgt10 cDNA library (Clontech, Palo Alto, CA) from rabbit muscle was screened with a probe made from the DNA sequence of the pig PGI gene (*8*). A partial rabbit PGI cDNA was purified from positive plaques identified during the library screen. Comparison to PGI sequences from other species suggested the partial cDNA lacked the 5′ and 3′ ends of the coding region. The 5′ and 3′ ends of the coding region were isolated using the RACE method (*18*, *19*). The DNA sequences of these overlapping fragments were obtained by standard procedures.

Crystallization of PGI. Rabbit skeletal muscle PGI was purchased from Sigma Chemical Co. and exchanged by gel filtration into a buffer containing 10 mM imidazole (pH 7.5), 50 mM KCl, and 3 mM NaN_3, and then concentrated to 25 mg/mL. The competitive inhibitor D-gluconate 6-phosphate from Sigma Chemical Co. (*6*) was added to a concentration of 10 mM, and the protein stock was filtered through a 0.2 μm cellulose acetate filter. Incomplete factorial analysis (*20*) with the hanging drop vapor diffusion method at room temperature was used to determine the initial crystallization conditions. X-ray diffraction quality crystals of PGI complexed with gluconate 6-phosphate were obtained at 25 °C against a solution of 13% PEG 8000, 250 mM magnesium acetate, and 100 mM sodium cacodylate (pH 6.5). The best diffracting crystals grew as colorless hexagonal rods with the symmetry of orthorhombic space group $C222_1$ ($a = 83.4$ Å, $b = 118.7$ Å, $c = 272.6$ Å). The final size (approximately 0.3 mm × 0.3 mm × 1.0 mm) was obtained through several rounds of macroseeding with freshly grown smaller crystals. Prior to data collection, crystals were removed from the hanging drops and placed in a stabilization soak solution of 20% PEG 8000, 250 mM magnesium acetate, and 100 mM sodium cacodylate (pH 6.5) with 5 mM gluconate 6-phosphate.

Data Collection and Structure Determination. Heavy atom derivatives were prepared by either a 30 min soak in 1 mM ethyl mercuric phosphate (EMP) or a 5 day soak in a saturated solution of *p*-chloromercuribenzoate (PCMB). Diffraction data for multiple isomorphous replacement (MIR) phasing were collected at 3 °C on a Rigaku R-Axis II image plate detector using 0.2 mm collimated, monochromatized CuKα radiation from a Rigaku RU-200 rotating anode generator (Table 1). Diffraction data from crystals of the enzyme−inhibitor complex, flash-frozen at −180 °C, were collected at NSLS beamline X12B at Brookhaven National Laboratories (Brookhaven, NY) using monochromatic X-rays at 1.008 Å. Flash-freezing required the addition of 20% glycerol to the stabilizing soak solution. Native and derivative data sets were processed with the programs DENZO and SCALEPACK (*21*).

Heavy atom data sets were scaled to the native data, and heavy atom positions were located by difference Patterson methods using the program XTALVIEW (*22*). The heavy atom sites were refined for the individual derivative data sets using the program HEAVYREF (*23*) and then refined in concert with the CCP4 program MLPHARE (*24*) to calculate initial MIR phases. The heavy atom positions fell into two groups that define a noncrystallographic 2-fold axis relating the two subunits of the PGI dimer in the asymmetric unit. The initial MIR map was improved using the CCP4 program DM (*24*) to implement solvent flattening, histogram matching, and electron density averaging. An entire protein monomer could then be built into the resulting averaged electron density map using the program O (*25*). Refinement of the model was initially performed using the program X-PLOR (*26*, *27*). Initial refinement was carried out imposing strict noncrystallographic symmetry between the two subunits. Subsequent refinement was carried out using noncrystallographic symmetry with restraints. This model was then refined against the data set collected to higher resolution (2.5 Å) at −180 °C. Several rounds of refinement and model adjusting brought the R factor to 23% with an R_{free} (*28*) of 33% using all data from 8 to 2.5 Å resolution.

Table 1: Statistics for Data Collection and Refinement[a]

Data Collection

	native	EMP	PCMB	native−180
space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$
cell dimensions (Å)				
a	83.40	83.79	82.65	82.70
b	118.66	119.63	119.11	115.27
c	272.56	273.82	271.48	271.84
temperature (°C)	3	3	3	−180
resolution range (Å)	30−3.0	30−3.0	30−3.0	35−2.5
no. of observed reflections				
total	252627	133464	142943	298487
unique	24543	21627	22457	41768
completeness (%)	89 (56)	78 (34)	81 (81)	93 (90)
R_{sym} (% on *I*)	6.6 (17)	6.1 (17)	7.5 (18)	4.1 (13)
no. of heavy atoms bound	0	4	8	0
MFID	−	16.3	22.2	−
empirical *K* anomalous	−	6.4	7.3	−
R_c, isomorphous	−	0.75	0.71	−
R_c, anomalous	−	0.90	0.84	−
phasing power	−	1.28	1.54	−

Refinement

	all data	$I > 2\sigma$ data
resolution range (Å)	35−2.5	35−2.5
$R_{crystal}$ (%)	21.2	21.2
R_{free} (%)	26.5	26.2
no. of reflections	41256	40775
rms deviations from bonds and angles	0.008 Å, 1.4°	0.009 Å, 1.5°

[a] $R_{sym} = |I_i - \langle I\rangle|/I_i$, where I_i is the intensity of an individual reflection and $\langle I\rangle$ is the average intensity over symmetry equivalents. $R_{crystal} = ||F_o| - |F_c||/|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. R_{free} is the equivalent of $R_{crystal}$, but it is calculated for a randomly chosen set of reflections that were omitted from the refinement process. $R_c = \Sigma||F_{PH} \pm F_P| - F_H(calc)|/\Sigma|F_{PH} \pm F_P|$. Phasing power $= \{\Sigma|F_H|^2/\Sigma[|F_{PH}(obs)| - |F_{PH}(calc)|]2\}^{1/2}$.

The model was then refined using the CNS program (*29*). All data between 30 and 2.5 Å resolution were used in the refinement with a bulk solvent correction applied. The "wat_pick" feature of CNS was used to select water molecules on the basis of omit maps with coefficients $|F_o - F_c|$, together with criteria based on geometry and refined *B* factors. Waters without a reasonable H-bonding environment or with a *B* factor of >60 Å^2 were removed. Several rounds of water generation and removal were repeated with intermittent rounds of positional and individual *B* factor refinement. The inhibitor D-gluconate 6-phosphate was built into each of the two active sites in the PGI molecule. The geometry was assessed and corrected using the CCP4 program PROCHECK (*24*). Geometrically unfavorable regions of the model were adjusted so that the two monomers in the asymmetric unit had similar backbone geometry. The entire structure was checked with a composite simulated annealing omit map (*29*). One final round of refinement was performed with noncrystallographic restraints set at a low force constant. The model was determined to be complete when all further attempts of refinement failed to reduce the R_{free} (*28*). The final *R* factor is 21.2% for all data from 30 to 2.5 Å resolution with an R_{free} of 26.5% (Table 1).

RESULTS AND DISCUSSION

The crystal structure of rabbit skeletal muscle PGI with a bound competitive inhibitor of the isomerase activity, gluconate 6-phosphate, was determined to 2.5 Å resolution by the method of multiple isomorphous replacement. To complete the crystal structure, it was necessary to determine the amino acid sequence of rabbit PGI, so the cDNA for rabbit PGI was isolated and sequenced as described in Materials and Methods. The rabbit PGI amino acid sequence is shown in Figure 1.

The final model includes two PGI subunits, two bound inhibitor molecules, and 370 water molecules per asymmetric unit. The two subunits in the asymmetric unit are arranged as a dimer, with a roughly spherical overall shape (Figure 2A). There is a large cleft at each end and a hook shape that protrudes from each side (Figure 2B). Each subunit of the model contains residues 1−555. The two C-terminal residues of each subunit are disordered. Each subunit is made up of two dissimilar α/β-sandwich domains (Figure 2C). The small domain contains a central five-stranded parallel β-sheet surrounded by loops and α-helices. The large domain has four parallel β-strands, and two additional antiparallel β-strands from near the N-terminus that complete the six-stranded β-sheet core. The subunit interdomain interface and the dimer interface overlap to form the active site of the enzyme. The dimer interface defines a noncrystallographic 2-fold axis. The C-terminus of each subunit is made up of two helices and a loop that extend across the surface of the other subunit in the dimer (Figure 2B).

The topology of rabbit PGI is similar to that originally reported for pig PGI but differs in several respects. Most significantly, the large domain of rabbit PGI contains a mixed β-sheet; the pig enzyme structure was reported to have an all-parallel sheet (*16*). The differences are easily explained by the lack of sequence information for pig PGI at the time that structure was determined, which made it impossible to verify the fold or refine the structure.

As expected, the overall fold of rabbit PGI is very similar to the recent structure of PGI from *B. stearothermophilus* (*17*). The most significant differences between the mammalian and bacterial PGI are located in the large domain. The mammalian PGI contains an additional 36 residues at the N-terminus that form an initial α-helix followed by a loop and an additional β-strand. In addition, a hook feature that may be involved in some of the mammalian protein's extracellular functions has a different location and conformation in the two proteins. The bacterial hook is a loop without any helix or strands, but the mammalian hook feature has defined secondary structure with a helix−turn−helix motif. Panels D and E of Figure 2 compare the subunit structure of mammalian and bacterial PGI, respectively. The relative positions of the mammalian N-terminal extension and the altered hook region are shown on the surface of one subunit. These differences in the structures do not appear to be the result of inhibitor binding because they are not located near the active site. Further comparison of the two structures does not reveal other conformational changes that might be due to inhibitor binding. Superposition of the bacterial structure, which lacks a bound inhibitor, and the rabbit PGI structure, which contains a bound inhibitor in the active site, reveals no significant amino acid shifts in and around the active site and no significant changes in interdomain or intersubunit interfaces.

While the bacterial and rabbit PGI structures share a similar overall fold, this fold differs from that of other

FIGURE 1: PGI sequence alignment. The rabbit PGI sequence from this work is aligned with representative sequences from *Drosophila* (Swissprot accession number P52030), *Escherichia coli* (P11537), yeast (P12709), plants (P49105), and *B. stearothermophilus* (P13376). Residues that interact directly with the inhibitor are in red boxes. Residues that are the same as in the rabbit enzyme are in blue boxes. This figure was prepared using the program ALSCRIPT (*30*).

proteins reported thus far. When the rabbit PGI fold was compared to structures in the Protein Data Bank using the Dali program package (*33*), the program identified many proteins that contain domains with α/β-folds. The highest scoring protein was glucosamine-6-phosphate synthase (1moq) which, like PGI, contains two α/β-sandwich domains. However, neither glucosamine 6-phosphate nor the other α/β-domains identified by the Dali program have the same secondary structure topology as PGI.

Active Site. Electron density corresponding to a bound gluconate 6-phosphate molecule is found in each of the two identical substrate binding pockets (Figure 3A). These pockets are located between the two α/β-sandwich domains of one subunit and at the interface between the two subunits. Many secondary structural elements come together to form the substrate binding pockets, and each pocket consists of residues from both subunits of the dimer and from both domains of a subunit. In the crystal structure, there are several residues that interact with the bound gluconate 6-phosphate inhibitor and are probably involved in binding the substrate and promoting catalysis (Figure 3B). An alignment of representative PGI sequences from mammals, plants, flies, bacteria, and yeast illustrates the absolute conservation of all the active site residues mentioned below (Figure 1).

Side chain hydroxyl groups from Ser159, Ser209, Thr211, and Thr214 form hydrogen bonds with the phosphate group of the inhibitor. An additional hydrogen bond is made between the phosphate group and a water molecule that is bound to the side chain hydroxyl group of Thr217. The phosphate group also interacts with two backbone amide groups, from residues 210 and 211. Together, these residues surround the phosphate group, forming a small pocket with considerable binding energy. Lys210 is another residue that supplies some positive charge to the phosphate binding site. In addition, these residues are clustered near the N-terminus of two α-helices that are within the small domain of the PGI monomer. The partial positive charge at the N-terminus of the α-helices helps to balance the negative charge of the phosphate group. These interactions and partial charges explain the enzyme's specificity for phosphorylated sugars, and why unphosphorylated sugars are not competitive inhibitors. The phosphate binding site of PGI bears a striking resemblance to that of triosephosphate isomerase, in which backbone amides, side chain hydroxyls, and a lysine amino group are all involved (*34*).

The mechanism for the isomerization reaction has been proposed to involve general acid/base catalysis with proton transfer via an enediol intermediate (*3*, *35*). pH profiles for the reaction suggest involvement of groups with pK_a values of 6.7 and 9.4, implicating a nonprotonated histidine and a protonated lysine side chain as candidates for general acid/base residues (*3*). As proposed by Rose and co-workers (*2*), in the first step of the reaction, the cyclic form of glucose 6-phosphate binds to the enzyme active site, and a proton is transferred from the enzyme to the ring oxygen of the substrate. Protonation of the ring oxygen is accompanied by

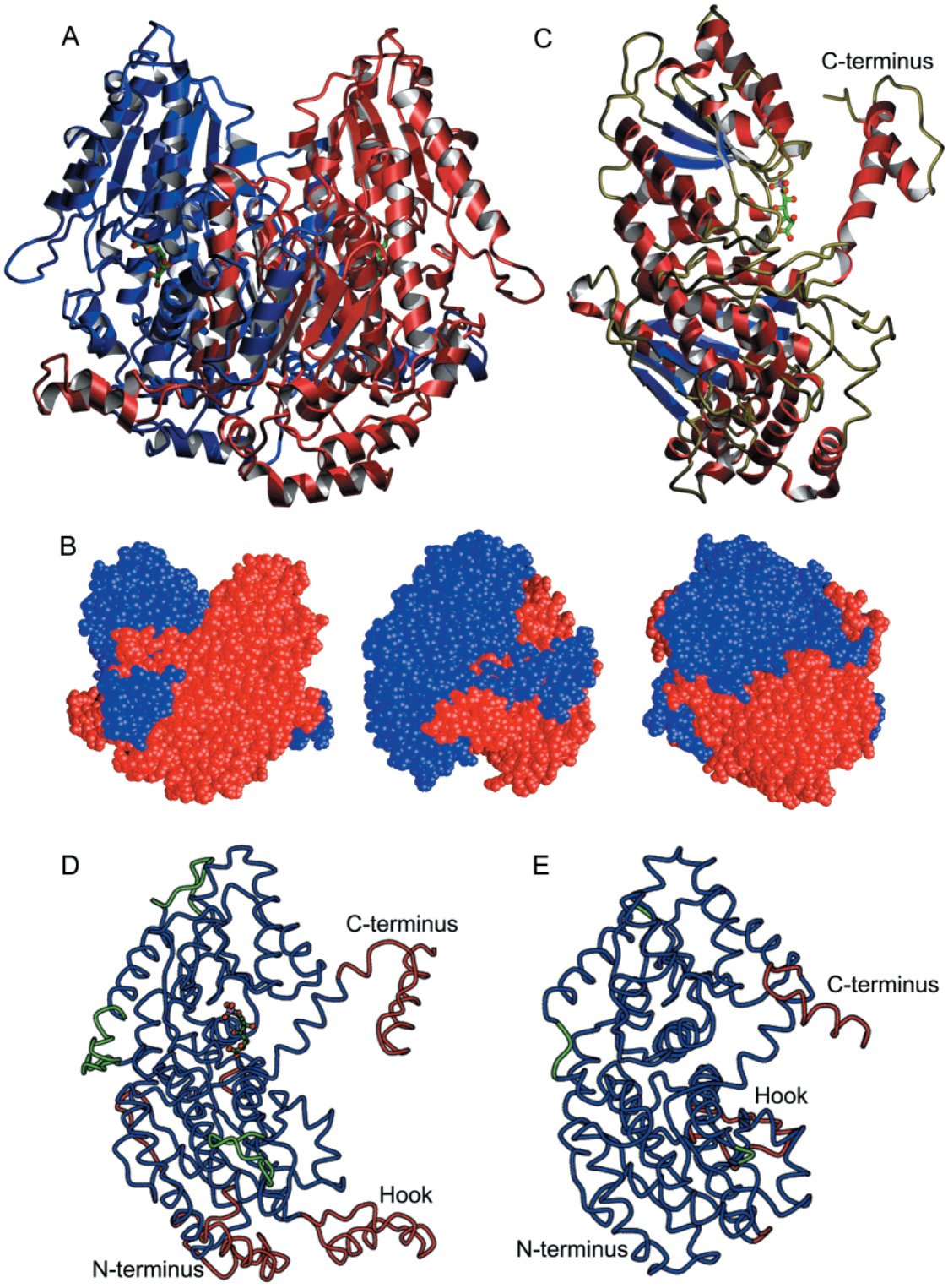

FIGURE 2: (A) A dimer of PGI, the active form of the enzyme, is shown as a ribbon diagram. One subunit is drawn in blue, and the other is in red. The active sites can be located with ball-and-stick models of the gluconate 6-phosphate inhibitor (green). This figure, as well as Figures 2C, 3A,B, 4, and 5A, was made using the programs MOLSCRIPT (*31*), POVSCRIPT (E. Peisach and D. Peisach, Brandeis University, personal communication), and POVRAY (http://www.povray.org). (B) Three views of a space filling model of the dimer. One subunit is shown in red, and the other is in blue. A large cleft is located at the top of the model, and another cleft becomes visible when the protein is turned by 90°. This figure was prepared with the program GRASP (*32*). (C) A ribbon diagram of a single subunit of PGI. Each of the two domains has a central β-sheet (blue) surrounded by α-helices (red). The C-terminal residues form an extended region that wraps around the other subunit in the dimer. The N-terminus is located at the end of one of the helices near the bottom of the figure, in the large domain. The 6-phosphogluconate inhibitor is shown as a green ball-and-stick model. (D and E) Comparison of rabbit (D) and *B. stearothermophilus* (E) enzyme structures. An α-carbon trace of one subunit of each enzyme is shown. The hook region, the N-terminal extension in the rabbit enzyme, and the C-terminal residues for each subunit are shown in red. Three additional loops that are much longer in the rabbit enzyme than in the bacterial enzyme are shown in green. These figures, as well as panels B−D of Figure 5, were made using the program MOLSCRIPT (*31*).

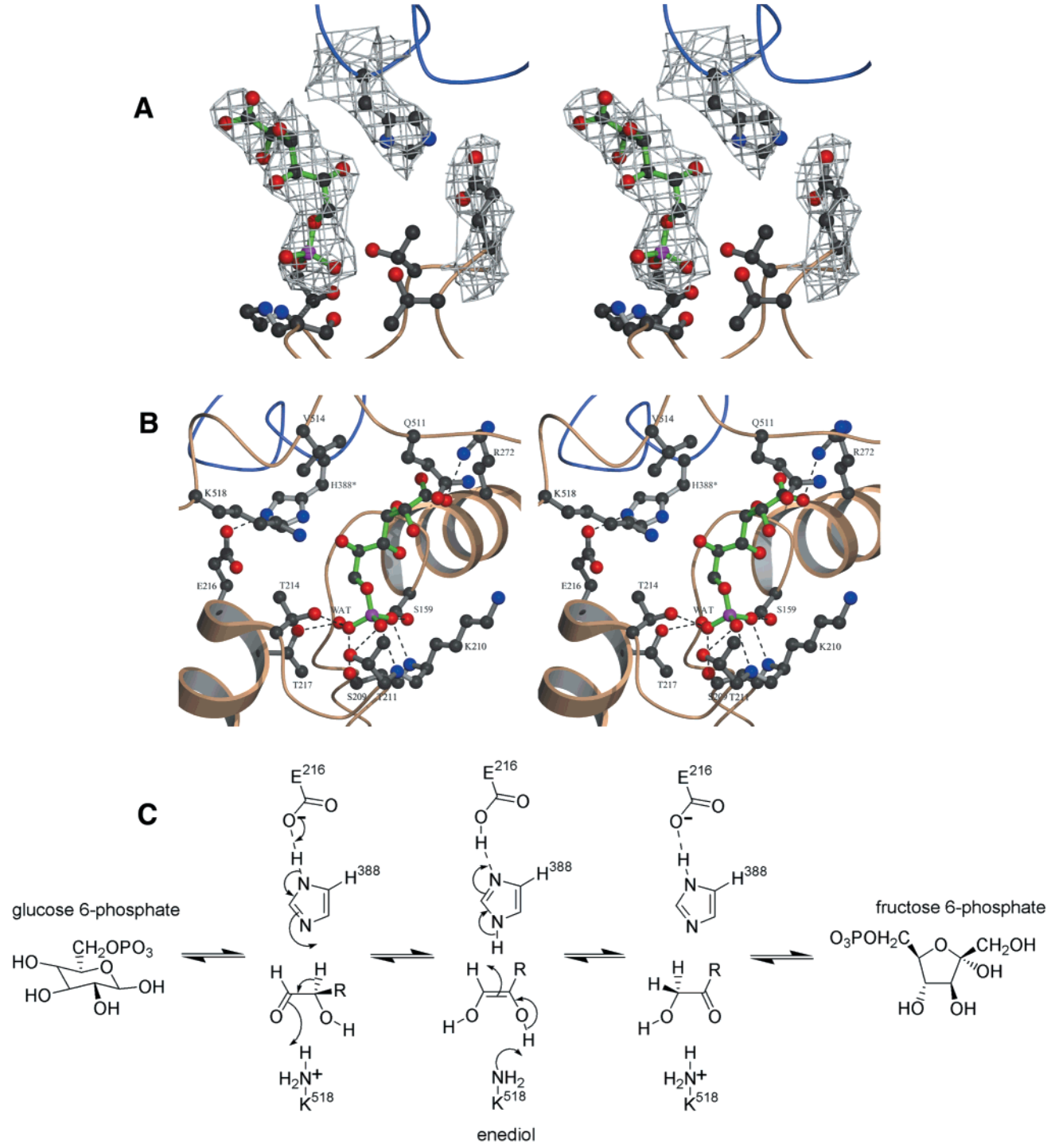

FIGURE 3: Isomerase active site. (A) Part of an electron density map surrounding the gluconate 6-phosphate inhibitor and residues H388 and E216 is shown. The composite annealing electron density map was calculated with coefficients of $|2F_o - F_c|$ and shown with a 1σ contour level (*29*). (B) Ball-and-stick model of the active site. The gluconate 6-phosphate inhibitor is shown (green) relative to essential residues of the active site (gray). Most of the residues are from one subunit. His388 is a residue that is from the other subunit of the dimer. Hydrogen bond interactions are represented by dashed lines (this view is rotated 180° relative to panel A). (C) Catalytic mechanism of PGI. His388 and Glu216 make up a catalytic diad to abstract the proton off C-2 of the ring-opened form of glucose 6-phosphate. Lys518 acts as the general acid catalyst to complete the formation of the enediol intermediate. The enediol is reprotonated by the catalytic diad, and Lys518 acts as a general base catalyst to form the ring-opened form of fructose 6-phosphate.

an electron shift, opening the ring. The C-2 proton of the substrate is now relatively acidic because it is α to a carbonyl group. An enzymic base abstracts the C-2 proton, and a pair of electrons shift to form a double bond between C-1 and C-2, yielding the *cis*-enediol intermediate (Figure 3C). The results of solvent exchange experiments indicated that the same proton that was removed from C-2 is now added back to the intermediate on C-1. This second proton transfer is accompanied by electron shifts that result in the loss of a proton from the oxygen at C-2 and the gain of a proton by the oxygen at C-1 of the substrate, resulting in the formation of the open chain form of fructose 6-phosphate. The final step involves transfer of a pair of electrons from the incipient ring oxygen to C-2. Upon formation of this bond, the ketose

oxygen accepts a proton from the solvent, to complete the ring-closed form of fructose 6-phosphate.

Affinity labeling and chemical modification studies have also predicted the involvement of a histidine (*36*) and a lysine (*37*) residue in the reaction, as well as suggested the importance of a glutamate (*38*) and an arginine (*39*) residue. This structure of PGI with a bound inhibitor is consistent with these predictions. A catalytic diad defined by His388 and Glu216 could provide the predicted histidine side chain for abstracting the C-2 proton of glucose 6-phosphate, while the amino group of Lys518 could act as a general base catalyst to protonate the carbonyl and complete the formation of the putative enediol intermediate (Figure 3B).

The glucose 6-phosphate substrate is likely to bind so that C-1 and C-2 are positioned closer to His388 and Lys518 than is seen for the inhibitor in the crystal structure. The 6-phosphogluconate inhibitor differs from the glucose 6-phosphate substrate because the inhibitor has a charged carboxylate group at C-1. The inhibitor interacts with Arg272 in the crystal structure, but the glucose 6-phosphate substrate could bind in a different conformation so that C-1 and C-2 are on the other side of Val514 (Figure 3B). Also, since the crystals were grown at pH 6.5, the NE2 atom of His388 is likely protonated and found to be hydrogen bonding with a carboxylate oxygen of Glu216. However, in the productive mode at pH 7, the NE2 atom of His388 would be unprotonated and is predicted to rotate toward C-2 of the substrate, as shown in the mechanism in Figure 3C.

The details of this mechanism show both similarities with and important differences from the mechanisms of triosephosphate isomerase and xylose isomerase (*40*). Both PGI and xylose isomerase catalyze sugar ring opening, but in the case of xylose isomerase, a possible candidate for the ring-opening acid is a water molecule bound to a magnesium ion in the active site (*41*). No metal ion is found in our structure despite the presence of 250 mM magnesium ion. The triosephosphate isomerase active site contains a histidine and a lysine, but the histidine is used to protonate the oxygens of the substrates (*42*), a role here proposed for Lys518 in PGI. The catalytic base in triosephosphate isomerase is a glutamate (Glu165) with an elevated pK_a; in PGI, use of a histidine in a catalytic diad would provide a group with a similar basicity. Arg272 in PGI may serve a function similar to that of Lys12 in triosephosphate isomerase (*42*); it sets the overall electrostatic potential at the substrate to be positive and provides stabilization for the negative charges that develop on the enediolate-like transition states.

A thorough analysis of the precise positioning of residues and bound substrate in the PGI active site is not warranted at present mainly due to chemical differences between the inhibitor gluconate 6-phosphate and the substrate glucose 6-phosphate, the enediol intermediate, or the fructose 6-phosphate product. It is expected that with a true substrate bound, the C-2 position will be closer to NE2 of His388 than is seen in the structure with bound inhibitor.

Gp120 Homology. A portion of PGI was reported to be homologous to a portion of the gp120 env protein of HIV-1 (*43*). The sequence alignment is suggestive of homology between neuroleukin (residues 403−447) and HIV-1 (residues 238−282), with a stretch of 45 amino acids that are 31% identical and 73% similar. Secondary structure predictions of this sequence predicted a β-sheet for the N-terminal portion and an α-helix for the C-terminal section (*43*). The crystal structure of rabbit PGI has a β-strand of residues 403−409 and α-helices of residues 419−436 and 440−450, confirming the earlier prediction. However, the crystal structure of the HIV gp120 glycoprotein (*44*) does not support the conclusion of structural homology with PGI or neuroleukin within this region. The aligned stretch of amino acids in the gp120 protein (residues 238−282) has entirely β-strand structure, with no α-helix present.

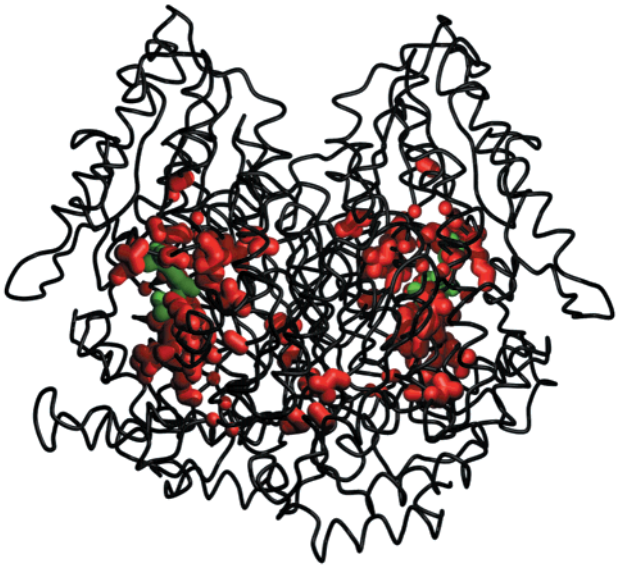

FIGURE 4: Conserved residues. An α-carbon trace of the phosphoglucose isomerase structure is shown in black. Residues that are conserved in all known PGI sequences are shown as red ball-and-stick models. The bound gluconate 6-phosphate inhibitors are shown in green, to indicate the location of the two active sites.

Conservation of the Active Site. Phosphoglucose isomerase is nearly ubiquitous in evolution, and sequences are available from more than 60 sources. An alignment of several PGI sequences showing representatives from mammals, plants, bacteria, flies, and yeast is shown in Figure 1. Almost 10% of the sequence, 47 residues, is conserved. The conserved residues cluster around each active site (Figure 4). As is true for many enzymes, they include the residues that interact directly with the inhibitor, and others that form several layers that help to shape the active site pocket. A few are involved in holding the catalytic residues in place. During the billion or so years since these different organisms diverged, this region was preserved, so from the structure, we have an idea of many of the residues that are important in the isomerase function of this protein.

At this time, it is not clear which regions of PGI are important for binding to receptors that may be involved in its "moonlighting" functions as AMF, DMM, and NL. However, Figure 4 can be evaluated in another way; during a billion years of evolution, a cluster of about 40 residues remained the same, but at the same time, most other residues changed. As is the case with most proteins, many of the residues that diverged are located on the protein's surface. In this large dimeric protein of more than 1000 amino acids, there are whole helices made up of nonconserved residues. It is quite possible that one of these regions, perhaps a hook or a cleft, could have gained an additional binding function. As long as that new function did not interfere with the

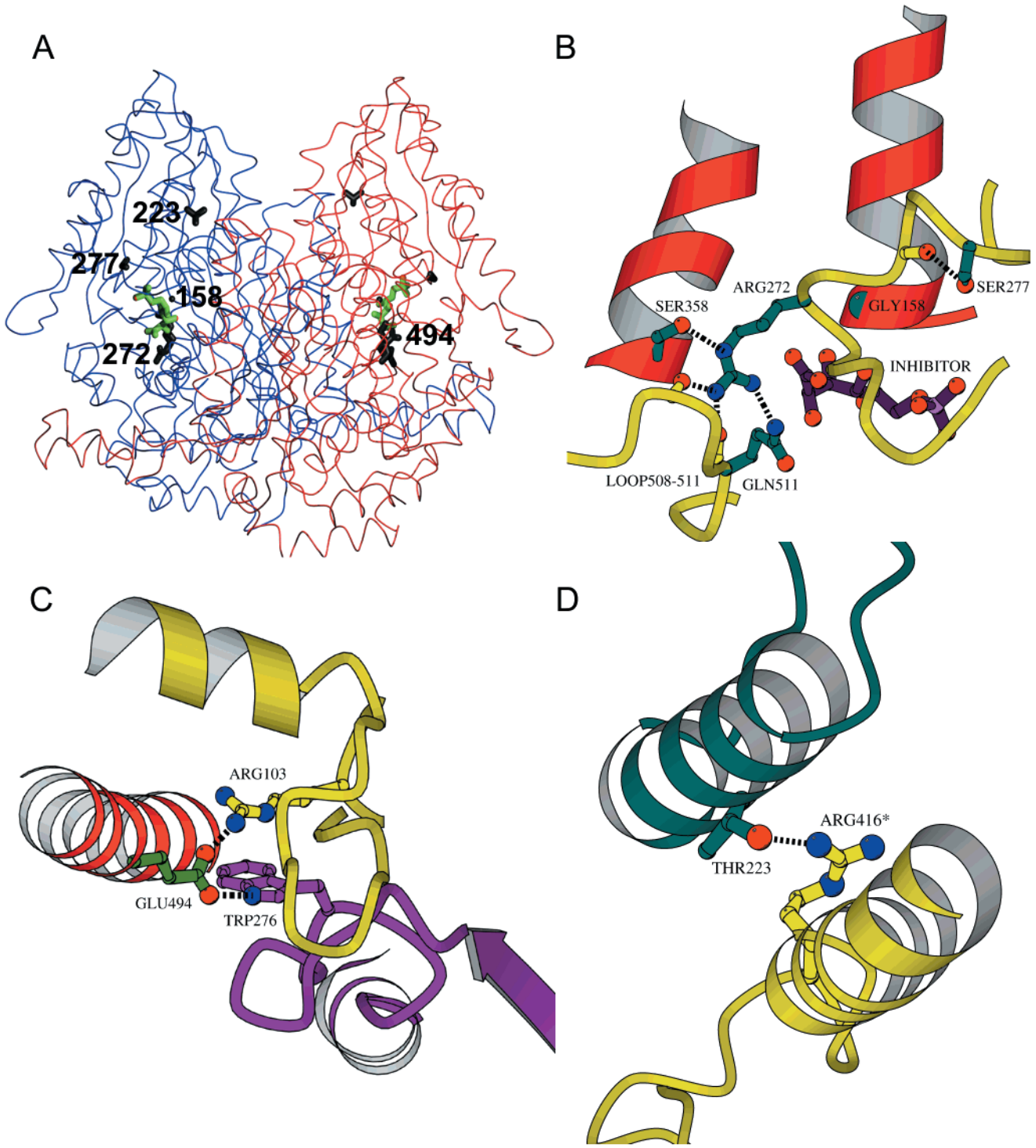

FIGURE 5: (A) Locations of amino acid substitutions that can cause nonspherocytic hemolytic anemia are indicated by ball-and-stick models of the wild type residue (black) superposed on an α-carbon trace of the PGI dimer (red and blue). (B) The substitution of Ser277 with leucine, Gly158 with serine, or Arg272 with histidine can also cause nonspherocytic hemolytic anemia. Ser277 normally makes a hydrogen bond to the backbone carbonyl of residue 274, which helps form a tight loop near the bound inhibitor. Arg272 forms hydrogen bonds to the side chains of Ser358 and Gln511 and to the backbone carbonyls of residues 508 and 510. These interactions between the loop containing residue 272, the loop containing residues 508−511, and the helix containing residue 358 are probably important in forming the shape of the active site. In addition, Gly158 is located at the end of helix 6. The gluconate 6-phosphate inhibitor (green) binds close to the end of this helix. Substitution of Gly158 with a larger residue would result in steric hindrance of substrate binding. (C) The substitution of Glu494 with lysine results in the loss of hydrogen bonds to the side chains of Arg103 and Trp276. These hydrogen bonds are probably important in stabilizing the interactions between these three regions of one subunit. (D) The substitution of residue Thr223 with methionine results in the loss of a hydrogen bond between Thr223 and Arg416, which is from the other subunit in the dimer.

original function of the protein, it might have benefited the organism and its offspring and been retained during evolution. This is one possible way to explain how a protein can moonlight, or have two apparently unrelated functions, and specifically how this key glycolytic enzyme may have gained a second function as a growth factor.

Other examples of moonlighting proteins have been described in the literature (reviewed in ref *45*). For example,

thymidylate phosphorylase is another intracellular enzyme that is also an extracellular cytokine (*46*). For some such proteins, there is not a clear connection between their different functions, but it may be a matter of a ubiquitous protein being recruited for a new function as, for instance, a new organ evolves. For example, this seems to have been the origin of some of the crystallins that form the lens of the eye (reviewed in ref *47*).

In the case of PGI, Watanabe and co-workers have identified a glycosylated transmembrane protein as a receptor for at least some of the cytokine functions (*14*). The N-termini, hook shapes, and other surface loops that vary between the rabbit and bacterial PGI structures make these features attractive candidates for regions of the protein that might bind a receptor. However, in light of the results of Sun and co-workers (*17*) who found that the bacterial protein retains at least some of the cytokine function of the mammalian protein, at least part of the receptor binding is probably due to structurally homologous regions. In addition, since binding of gluconate 6-phosphate to autocrine motility factor inhibits cell motility (*10*), it appears the sugar binding site is an essential component of binding to at least one receptor; however, it is not clear if this interaction completely determines the specificity or binding affinity. One likely scenario is that the conserved PGI active site binds to a carbohydrate moiety of the receptor, and additional regions of PGI form protein−protein interactions with the receptor. Without structural information about how this protein binds to its receptor, there is no direct evidence for this hypothesis.

Nonspherocytic Hemolytic Anemia. Nonspherocytic hemolytic anemia (NHA) is a rare, autosomal recessive disease that occurs in a variety of ethnic groups. It is caused by a decrease in the activity of triosephosphate isomerase, pyruvate kinase, or PGI. One of the characteristic features of the disease is fragile red blood cell membranes, resulting in jaundice and anemia (*48*−*51*). The less severe forms of the disease can be treated by blood transfusions and splenectomy. More severe forms can also result in neurological, cardiac, and additional problems that can lead to death by age 5.

Gene sequencing has identified 26 amino acid substitutions in human PGI that can result in NHA (*52*−*56*). The sequences of the human and rabbit proteins are 93% identical, so the rabbit structure can be used to elucidate the possible roles individual residues have in the wild type protein and how substitutions can affect activity, folding, and/or stability. The positions of five amino acid residues that are mutated in the disease are shown superposed on an α-carbon trace of the rabbit PGI structure in Figure 5A. Three of these five residues are completely conserved (Gly158, Arg272, and Gly494), and a fourth, Ser277, is conserved as a serine in eukarya and eubacteria, and is replaced by a similar amino acid, threonine, in archea.

Three amino acid residues that are substituted in NHA (Gly158, Arg272, and Ser277) are located in the active site (Figure 5B). (1) The side chain hydroxyl group of Ser277 forms a hydrogen bond to the carbonyl oxygen of Leu274 and helps stabilize a tight loop. In some patients, Ser277 is changed to a leucine, which would result in the loss of this hydrogen bond and probably affect folding in this region. (2) Arg272 is in a position where its side chain could form hydrogen bonds to the side chains of Ser358 on helix 14 and Gln511 on a loop between helix 19 and helix 20 and the carboxyl groups of residues 508 and 510. These interactions connect two loops and a helix and help form the shape of the active site. As mentioned above, Arg272 might also be important for establishing the proper electrostatic environment of the active site. In some nonspherocytic hemolytic anemia patients, Arg272 is changed to a histidine. (3) Gly158 is located at the end of helix 6. The substrate binds closely against the end of this helix. Substitutions of Gly158 would result in the addition of a bulkier residue at the end of helix 6, which would be likely to interfere with substrate binding.

NHA can also be caused by amino acid substitutions at positions outside the active site, including substitutions of Glu494 and Thr223. Glu494 forms two hydrogen bonds (Figure 5C). The first is to the side chain of Trp276, which is located in a loop connecting helix 11 to strand 7. The second is to Arg103, which is located in a loop between helix 5 and strand 3. Together, these interactions help stabilize the subunit fold. The substitution of Glu494 with a lysine, found in anemia patients, results in the loss of these two hydrogen bonds and could lead to problems in folding and/or a decrease in stability.

The side chain of Thr223 forms a hydrogen bond with the side chain of Arg416 from the other subunit in the dimer (Figure 5D). In some nonspherocytic hemolytic anemia patients, Thr223 is replaced by a methionine, which would result in the loss of the hydrogen bond as well as possible steric hindrance, preventing correct packing between the subunits. Thr223 is partially conserved: a threonine or a serine in mammals, yeast, and bacteria.

Summary. The crystal structure of rabbit PGI with bound gluconate 6-phosphate inhibitor reported here is the first refined structure of a mammalian phosphoglucose isomerase. It also provides the first view of the active site with a specifically bound substrate analogue. This structure leads to the identification of several residues that appear to be important for substrate specificity and catalysis. The structure is consistent with the results of previously reported studies of the activity of the enzyme, with the locations of conserved active site residues, and also with the phenotypes of several naturally occurring amino acid substitutions. While it is not yet clear how this protein might bind to a cell surface receptor, the PGI fold has several unusual features, including large clefts and hooks in addition to the active site, that could be sites for receptor binding.

REFERENCES

1. Andersson, S. G. E., Zomorodipour, A., Andersson, J. O., Sicheritz-Pontén, T., Alsmark, U. C. M., Podowski, R. M., Näslund, A. K., Eriksson, A.-S., Winkler, H. H., and Kurland, C. G. (1998) *Nature 396*, 133−140.
2. Rose, I. A. (1962) *Brookhaven Symp. Biol. 15*, 293−309.
3. Dyson, J. E. D., and Noltmann, E. A. (1968) *J. Biol. Chem. 243*, 1401−1414.
4. Rose, I. A., and O'Connell, E. L. (1961) *J. Biol. Chem. 236*, 3086−3092.
5. Seeholzer, S. H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1237−1241.
6. Parr, C. W. (1956) *Nature 178*, 1401.
7. Chirgwin, J. M., Parsons, T. F., and Noltmann, E. A. (1975) *J. Biol. Chem. 250*, 7277−7279.
8. Chaput, M., Claes, V., Portetelle, D., Cludts, I., Cravador, A., Burny, A., Gras, H., and Tartar, A. (1988) *Nature 332*, 454−455.

9. Faik, P., Walker, J. I. H., Redmill, A. A. M., and Morgan, M. J. (1988) *Nature 332*, 455−457.
10. Watanabe, H., Takehana, K., Date, M., Shinozaki, T., and Raz, A. (1996) *Cancer Res. 56*, 2960−2963.
11. Xu, W., Seiter, K., Feldman, E., Ahmed, T., and Chiao, J. W. (1996) *Blood 87*, 4502−4506.
12. Gurney, M. E., Heinrich, S. P., Lee, M. R., and Yin, H.-S. (1986) *Science 234*, 566−574.
13. Gurney, M. E., Apatoff, B. R., Spear, G. T., Baumel, M. J., Antel, J. P., Bania, M. B., and Reder, A. T. (1986) *Science 234*, 574−581.
14. Watanabe, H., Carmi, P., Hogan, V., Raz, T., Silletti, S., Nabi, I. R., and Raz, A. (1991) *J. Biol. Chem. 266*, 13442−13448.
15. Shaw, P. J., and Muirhead, H. (1977) *J. Mol. Biol. 109*, 475−485.
16. Achari, A., Marshall, S. E., Muirhead, H., Palmieri, R. H., and Noltmann, E. A. (1981) *Philos. Trans. R. Soc. London, Ser. B 293*, 145−157.
17. Sun, Y.-J., Chou, C.-C., Chen, W.-S., Wu, R.-T., Meng, M., and Hsiao, C.-D. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 5412−5417.
18. Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L., and Davis, M. M. (1989) *Science 243*, 217−220.
19. Ohara, O., Dorit, R. L., and Gilbert, W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 15*, 5673−5677.
20. Jancarik, J., and Kim, S.-H. (1991) *J. Appl. Crystallogr. 24*, 409−411.
21. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
22. McRee, D. E. (1999) *J. Struct. Biol. 125*, 156−165.
23. Terwilliger, T. C., and Eisenberg, D. (1983) *Acta Crystallogr. A39*, 813−817.
24. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
25. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
26. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science 235*, 458−460.
27. Brünger, A. T. (1992) *X-PLOR Version 3.1. A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
28. Kleywegt, G. J., and Brünger, A. T. (1996) *Structure 4*, 897−904.
29. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
30. Barton, G. J. (1993) *Protein Eng. 6*, 37−40.
31. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
32. Sharp, K., Fine, R., and Honig, B. (1987) *Science 236*, 1460−1463.
33. Holm, L., and Sander, C. (1994) *Nucleic Acids Res. 22*, 3600−3609.
34. Lolis, E., and Petsko, G. A. (1990) *Biochemistry 29*, 6619−6625.
35. Schray, K. J., Benkovic, S. J., Benkovic, P. A., and Rose, I. A. (1973) *J. Biol. Chem. 248*, 2219−2224.
36. Gibson, D. R., Gracy, R. W., and Hartman, F. C. (1980) *J. Biol. Chem. 255*, 9369−9374.
37. Schnackerz, K. D., and Noltmann, E. A. (1971) *Biochemistry 10*, 4837−4843.
38. O'Connell, E. L., and Rose, I. A. (1973) *J. Biol. Chem. 248*, 2225−2231.
39. Riordan, J. F., McElvany, K. D., and Borders, C. L., Jr. (1977) *Science 195*, 884−886.
40. Allen, K. N., Lavie, A., Farber, G. K., Glasfeld, A., Petsko, G. A., and Ringe, D. (1994) *Biochemistry 33*, 1481−1487.
41. Allen, K. N., Lavie, A., Glasfeld, A., Tanada, T. N., Gerrity, D. P., Calrson, S. C., Farber, G. K., Petsko, G. A., and Ringe, D. (1994) *Biochemistry 33*, 1488−1494.
42. Joseph-McCarthy, D., Lolis, E., Komives, E. A., and Petsko, G. A. (1994) *Biochemistry 33*, 2815−2823.
43. Lee, M. R., Ho, D. D., and Gurney, M. E. (1987) *Science 237*, 1047−1051.
44. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) *Nature 393*, 648−659.
45. Jeffery, C. J. (1999) *Trends Biochem. Sci. 24*, 8−11.
46. Furukawa, T., Yoshimura, A., Sumizawa, T., Haraguchi, M., Akiyama, S.-I., Fukui, K., Ishizawa, M., and Yamada, Y. (1992) *Nature 356*, 668.
47. Piatigorsky, J. (1998) *Ann. N.Y. Acad. Sci. 842*, 7−15.
48. Eber, S. W., Gahr, M., Lakomek, M., Prindull, G., and Schröter, W. (1986) *Blut 53*, 21−28.
49. Schröter, W., Eber, S. W., Bardosi, A., Gahr, M., Gabriel, M., and Sitzmann, F. C. (1985) *Eur. J. Pediatr. 144*, 301−305.
50. Shalev, O., Shalev, R. S., Forman, L., and Beutler, E. (1993) *Ann. Hematol. 67*, 197−200.
51. Whitelaw, A. G. L., Rogers, P. A., Hopkinson, D. A., Gordon, H., Emerson, P. M., Darley, J. H., Reid, C., and Crawfurd, M. A. (1979) *J. Med. Genet. 16*, 189−196.
52. Walker, J. I. H., Layton, D. M., Bellingham, A. J., Morgan, M. J., and Faik, P. (1993) *Hum. Mol. Genet. 2*, 327−329.
53. Fujii, H., Kanno, H., and Miwa, S. (1994) *Blood 84* (Suppl. 1), 544A.
54. Xu, W., and Beutler, E. (1994) *J. Clin. Invest. 94*, 2326−2329.
55. Beutler, E., West, C., Britton, H. A., Harris, J., and Forman, L. (1997) *Blood Cells Mol. Dis. 23*, 402−409.
56. Huppke, P., Wünsch, D., Pekrun, A., Kind, R., Winkler, H., Schroter, W., and Lakomek, M. (1997) *Eur. J. Pediatr. 156*, 605−609.

BI991604M