Structural Basis for the Subcellular Localization Mechanism of a Mammalian Enzyme and the Substrate Recognition Mechanism of a Malaria Parasite Enzyme

January 2018

Ken-ichi AOKI

Structural Basis for the Subcellular Localization Mechanism of a Mammalian Enzyme and the Substrate Recognition Mechanism of a Malaria Parasite Enzyme.

A Dissertation Submitted to

the Graduate School of Life and Environmental Sciences,

the University of Tsukuba

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Biological Science

(Doctoral Program in Biological Sciences)

Ken-ichi AOKI

Table of Contents

目次

General Abstract

The proteins that accelerate chemical and biological processes within cells are called enzymes. The dimensions and shape of a protein molecule are very important to its functionality within the cell as agents that catalyze biochemical reactions. The three-dimensional structure of a protein or enzyme determines how the molecule will react with other molecules, including amino acids, carbohydrates, steroids, hormones, and lipids. However, the three-dimensional structures of many proteins are still unknown.

I have been involved in two major research projects that require protein structural information i) identification of the subcellular localization mechanism of a mammalian carbonyl reductase, and ii) clarification of the substrate recognition mechanisms of phosphoglucose isomerases from a malarial parasite and mouse. Specifically, in this thesis, I summarize in Chapter 1, the structural studies I have performed examining the subcellular localization mechanism of peroxisomal tetrameric carbonyl reductase and in Chapter II I describe the substrate recognition mechanism of phosphoglucose isomerase isolated from a human malarial parasite *Plasmodium falciparum*.

In Chapter I—The peroxisome is an organelle also known as a microbody, which has been identified in virtually all eukaryotic cells. Although almost all peroxisomal enzymes

function within the organelle matrix, the peroxisome, unlike the mitochondrion, has no genetic material of its own and consequently these enzymes must be imported from the cytosol.

Carbonyl reductase (CR) (EC 1.1.1.184) catalyzes the NADPH-linked reduction of a variety of carbonyl compounds to their corresponding secondary alcohols. Recently, cDNA cloning of pig heart CR and rabbit heart CR have revealed that both enzymes belong to the short-chain dehydrogenase/reductase (SDR) family, and that pig heart CR and rabbit heart CR have a peroxisomal target signal located in their C-termini.

With regard to the availability of the C-terminal tripeptide required for peroxisomal import, enzymes containing wild-type (SKR) and mutated (SKL, SHL, SLL, and SL) peroxisomal pig heart CR (PerCR) PTS1 sequences were examined for their ability to be transported into the peroxisome by expressing the respective cDNAs in HeLa cells. Although the mutant enzymes containing the SKL and SHL sequences were imported into the peroxisomes and retained full enzymatic activity, those with the SLL and SL sequences were not targeted to this organelle and were enzymatically inactive. Moreover, interestingly, the wild-type pig heart CR did not show specific targeting when introduced to the cells using a protein transfection reagent.

Thus, I decided to explore the PerCR peroxisomal subcellular localization mechanism

using crystallography.

I purified the PerCR protein from recombinant *Escherichia coli* (*E.coli*) and crystallized it. X-ray diffraction data were collected at the Photon Factory (Tsukuba). Finally, I determined the crystal structure of pig heart PerCR complexed with its coenzyme NADPH at 2.2 Å resolution. The crystal structure shows that the C-terminal Ser-Arg-Leu signal is involved in the R-axis-related inter-subunit interactions of tetrameric PerCR and is buried in the interior of the tetrameric PerCR molecule. The structure therefore reveals why, when injected into cells, the tetrameric PerCR molecule is not imported into the peroxisome, and thus enables me to propose a hypothetical model for the peroxisomal targeting of PerCR. Additionally, analysis of the geometry of the inter-subunit interactions allowed me to understand the role of the C-terminal sequence in maintaining an active tetrameric form of this enzyme.

In Chapter II—Malaria is one of the world's most serious infection diseases caused by parasitic infection of human tissue and blood, and is dependent on blood glucose to sustain the infection. An estimated 300–500 million new cases of malaria and up to 2.7 million deaths due to malaria are reported each year. In addition, recently, resistance to currently used anti-malarial drugs has been described.

Phosphoglucose isomerase (PGI; EC 5.3.1.9), also known as glucose 6-phosphate

isomerase (GPI), or phosphohexose isomerase (PHI), is a key enzyme in both glycolysis and gluconeogenesis. It catalyzes the second step of glycolysis, namely, the interconversion of glucose 6-phosphate (G6P) to fructose 6-phosphate (F6P) and *vice versa*. Since PGI is a housekeeping enzyme, it would be expected that its catalytic residues would be well conserved amongst PGIs from various species. However, significant amino-acid sequence differences (both insertions and deletions) are observed between mammalian PGIs (558 residues) and PfPGI (579 residues). Thus, a detailed comparison of the three-dimensional structures of human and plasmodium PGIs is important to understand whether these differences can be applied to future structure-based drug design and to gain insight into the structural origin of the species specificity of the cytokine activity of PGI.

I purified *Plasmodium falciparum* phosphoglucose isomerase (PfPGI) from recombinant *E. coli* and crystallized it. X-ray diffraction data were collected at the Photon Factory in Tsukuba and I determined the crystal structure in complexes with two kinds of inhibitors (3PGA and 6PGA) as well as its substrate (F6P) at 1.7 to 2.0 Å resolution. The crystal structures show two additional insertion loops (referred to as malaria original loops, MOL1 and MOL2) and these MOLs were found to be only conserved in *Plasmodium* species. Moreover, the MOLs contribute to dimer interactions and the

MOL2 (residues 438-446) disrupts an interaction between the carbohydrate binding helix (CBH) and the α15 helix* by steric hindrance. In the substrate binding site of PfPGI complexed with 6PGA, the conformations of the phosphate binding loop and the 3_{10}-helix containing His411* (i.e. His411 belonging to the other subunit) are quite similar to those observed in the crystal structure of mouse PGI in complex with 6PGA. However, the position of CBH, containing Lys540, in PfPGI was not as similar to the position in mPGI. Although a conformational change in CBH is necessary for it to directly interact with the substrate in order to fix the position of substrate in the protonation step of catalysis, CBH cannot directly interact with the substrate because of the steric hindrance that arises from MOL2 in PfPGI. Though further studies are necessary to make any definitive conclusions, the crystallographic data and structural analyses described in this thesis give new insights into the catalytic mechanism of PfPGI.

Abbreviations

ADH	alcohol dehydrogenase
AKR	aldo-keto reductase
AMF	autocrine motility factor
ASU	asymmetric unit
CBH	carbohydrate binding helix
CR	carbonyl reductase
DNA	deoxyribonucleic acid
E4P	erythrose 4-phosphate
F6P	fructose 6-phosphate
G6P	glucose 6-phosphate
G6PDH	glucose 6-phosphate dehydrogenase
GPI	glucose 6-phosphate isomerase
HEPES	N-2-hydroxyethylpiperazine-N`-2-ethanesulfonic acid
HSD	hydroxysteroid dehydrogenase
MF	maturation factor
MLCR	mouse lung carbonyl reductase
MOL	malaria original loop
mPGI	mouse phosphoglucose isomerase
MR	molecular replacement
NADP+	nicotinamide adenine dinucleotide phosphate
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NLK	neuroleukin
OD	optical density
PCR	polymerase chain reaction
PEG	polyethylene glycol
Pex	peroxine
Pf	*Plasmodium falciparum*
PF	photon factory
PfPGI	*Plasmodium falciparum* phosphoglucose isomerase
PGI	phosphoglucose isomerase
PTCR	peroxisomal tetrameric carbonyl reductase
PTS	peroxisomal targeting signal
RMSD	root-mean-square deviation
RNA	ribonucleic acid
SDR	short-chain dehydrogenase/reductase
TPR	tetratricopeptide repeat

Tris	tris (hydroxymethyl) aminomethane
3PGA	3-phosphoglyceric acid
6PGA	6-phosphogluconic acid

General Introduction

The proteins that accelerate chemical and biological processes within cells are called enzymes. The dimensions and shape of an enzyme molecule are very important to the protein's functionality within the cell as an agent that catalyzes biochemical reactions. Enzymes react with other molecules, including amino acids, carbohydrates, steroids, hormones, and lipids (1).

Enzymes are highly selective catalysts that can be specifically localized in the cells of specific organs. Some enzymes are localized on the surface of specific cells to allow them to be able to degrade large molecules into smaller molecules so that these smaller molecules can be easily taken up by cells. In contrast, some enzymes catalyze the opposite process by combining two substrates to produce a new molecule. In tandem with advanced research technologies such as PCR, DNA sequencing, bioinformatics databases and so on, we can now easily predict each enzyme's subcellular location and function. Currently, according to the UniProt Knowledgebase (UniProtKB: http://www.ebi.ac.uk/uniprot), which is the central access point for extensive protein

information such as function, classification, and cross-reference data, a total of 80 million protein sequences have been registered.

However, in order to obtain more detailed information about enzymes (proteins), such as the phenotypic form of mutant proteins, catalytic mechanisms, and protein-protein interactions at the atomic level, it is necessary to know the structure of the enzyme. To achieve this, various methods such as an X-ray crystallography, electron microscopy, and nuclear magnetic resonance have been developed. As a result, we can now visualize how enzymes work (2, 3).

In microscopy, the resolution is limited by the wavelength of the electromagnetic wave used. Using an optical microscope with a wavelength of approximately 300 nm, we can see individual cells and sub-cellular organelles. For an electron microscope, using a wavelength of under 10 nm, we can see the cellular architecture in more detail and visualize the approximate dimensions of large proteins. However, in both methods, we cannot observe protein behavior at the atomic level. Thus, it is necessary to use an electromagnetic wave with a wavelength of around 0.1 nm (or 1 Å) using X-rays in order to understand protein behavior at the atomic level (4).

X-ray crystallography is a higher resolution microscopy that uses the diffraction data obtained by X-rays and a protein crystal produced by crystallization methodologies (5-7).

Recently, advances in both hardware (brilliant synchrotron radiation sources and highly sensitive detectors) and software (computer programs for protein crystallography) have enabled us to determine new protein structures at high resolution more easily than in the past decade, and a huge number of protein structures are being rapidly solved. According to the Protein Data Bank (PDB; https://www.rcsb.org/pdb/home/home.do), which is an information website related to the 3D conformations of proteins, nucleic acids, and complex assemblies, approximately 800,000 protein structures have now been deposited. As a result, we can understand how proteins interact with other molecules, how they undergo conformational changes, and, in the case of enzymes, how they perform catalysis. Moreover, we can design specific compounds to target proteins, and these resources have contributed to basic research and education in molecular biology, structural biology, computational biology, and so on (8). In addition, X-ray crystallography can structurally classify the protein family type, and can answer research questions based on the 3D structure. I have been involved in two major research projects which require protein structural information: i) the identification of the subcellular localization mechanism of a

mammalian carbonyl reductase and ii) clarification of the substrate recognition mechanism of phosphoglucose isomerases from a malarial parasite and mouse.

Peroxisomal carbonyl reductase from pig heart tissue

Carbonyl reductase (CR; EC 1.1.1.184) catalyzes the NADPH-linked reduction of a variety of carbonyl compounds to their corresponding secondary alcohols, and contains different types of subunit found in several mammalian tissues (9). Some CRs are monomeric enzymes (approximately 280 amino acids) with molecular weights of around 30 kDa. Tetrameric mitochondrial CRs (4 × 250 amino acid acids) are found in the lungs of guinea pigs, mice, and pigs (10-13). The endogenous substrates of the monomeric CRs are thought to be isatin and 20-ketosteroids, whereas those of the mitochondrial tetrameric lung CRs are 3-ketosteroids and carbonyl compounds derived from lipid peroxidation. Both the cytosolic CRs and mitochondrial CRs belong to the short-chain dehydrogenase/reductase (SDR) family (14, 15). In addition to these multiple forms of CR, oligomeric and soluble forms of CR have been isolated from dog liver, rabbit heart, and pig heart (16-18). The properties of the oligomeric liver/heart enzymes differ from those of the monomeric CRs and the tetrameric mitochondrial lung CRs. Partial amino-

acid sequencing and full-length cDNA cloning of both pig heart CR (PHCR) and rabbit heart CR (RHCR) have identified that both enzymes consist of 260 amino-acid residues, and belong to the SDR family (18). cDNA cloning of RHCR has demonstrated that this enzyme is identical to rabbit $NADP^+$ dependent retinol dehydrogenase (NDRD). PHCR has a low sequence identity (<30%) to other mammalian monomeric CRs and mitochondrial tetrameric lung CRs, but has a high (80%) sequence identity to NDRD/RHCR. Characterization of recombinant PHCR and NDRD showed that they are identical enzyme species which reduce alkyl phenyl ketones, dicarbonyl compounds, and retinals, such as all-*trans*-retinal and 9-*cis*-retinal, as endogenous substrates (18). Further studies have shown that PHCR is localized in peroxisomes.

The first CR crystal structure was reported for mouse lung CR (MLCR) (19). The crystal structures of two monomeric CRs, porcine testicular CR and human CR, were subsequently determined (20, 22). The structural features, including the active site and the coenzyme-binding region, of the tetrameric MLCR and the monomeric CRs are similar to those of the other SDR enzymes. However, the substrate-binding region, which is the most variable part of the structure, differs among the various SDR enzymes. The crystal structures of SDR enzymes in the form of a retinal/retinol complex have not been reported (19, 23, 24).

Phosphoglucose isomerase from *Plasmodium falciparum*

Phosphoglucose isomerase (PGI; EC 5.3.1.9), also called as glucose 6-phosphate isomerase (GPI) or phosphohexose isomerase (PHI), is involved in the second step of the glycolytic pathway that interconverts glucose 6-phosphate (G6P) and fructose 6-phosphate (F6P). The pathway is key to the generation of free energy, which is used to form the high-energy molecule ATP (adenosine triphosphate) and NADH (reduced nicotinamide adenine dinucleotide) by converting glucose $C_6H_{12}O_6$, into pyruvate ($CH_3COCOO^- + H^+$) in almost all species (25).

In mammals, PGI also functions as an angiogenic factor (autocrine motility factor, AMF) that is involved in tumor metastasis signaling, and as a neurotrophic factor, neuroleukin (NLK) in spinal and sensory neurons (26-28).

With regard to the structure of PGI, several crystal structures have already been reported (human, pig, mouse, as well as others) (29-35).

PGI is a housekeeping enzyme; therefore, it was expected that its catalytic residues should be well conserved amongst PGIs from various species. However, significant amino-acid sequence differences (including insertions and deletions) are observed between the mammalian PGIs (558 residues) and *Plasmodium falciparum* PGI (579

residues). A structural comparison between mammalian and malaria PGIs should enable us to obtain useful information in order to design a selective inhibitor of PfPGI, which might serve as a lead compound for developing novel anti-malarial compounds.

In this thesis, I describe the subcellular localization mechanism of a mammalian carbonyl reductase, and a difference in the substrate recognition mechanism of mammalian and malaria parasite PGIs. In chapter I, I reveal the structure of tetrameric PerCR and the fact that this tetrameric form cannot be imported into the peroxisome from the cytosol because the peroxisomal targeting signal is buried in the interior of the tetrameric molecule. In Chapter II, I describe the structures of PfPGI in complexes with inhibitors and its substrate F6P, and show that the substrate recognition mechanism of PfPGI is somewhat different from that of mammalian PGIs as a result of steric hindrance caused by the malaria original loop domain.

Chapter I

Crystallization and Preliminary X-ray Crystallographic Studies and Structural Basis for the Peroxisome Localization Mechanism of Pig Heart Carbonyl Reductase

Abstract

Pig heart peroxisomal carbonyl reductase (PerCR), which belongs to the short-chain dehydrogenase/reductase family, and has a C-terminal SRL tripeptide, which is a variant of the type 1 peroxisomal targeting signal (PTS1) Ser-Lys-Leu-C term. PerCR is imported into the peroxisomes of HeLa cells when the cells are transfected with vectors expressing the enzyme. However, PerCR does not show specific targeting when introduced into the cells using a protein transfection reagent. To understand the structural basis for peroxisomal localization of PerCR, I determined the crystal structure of PerCR.

Purified PerCR was crystallized by the hanging drop vapor-diffusion method. Two crystal forms (I and II) were obtained in the presence of NADPH. Form I crystals belong to the tetragonal space group $P4_2$, with unit-cell parameters $a = b = 109.61$, $c = 94.31$ Å, and diffract to 1.5 Å resolution. Form II crystals belong to the tetragonal space group $P4_12_12$, with unit-cell parameters $a = b = 120.10$, $c = 147.00$ Å, and diffract to 2.2 Å resolution. Both crystal forms were suitable for X-ray structure analysis at high resolution.

The data revealed that the C-terminal PTS1 of each subunit of PerCR is involved in inter-subunit interactions and is buried in the interior of the tetrameric molecule. These findings indicate that the PTS1 receptor Pex5p in the cytosol recognizes the monomeric form of PerCR whose C-terminal PTS1 is exposed, and that this PerCR is targeted into the peroxisome, thereby forming a tetramer.

Introduction

The peroxisome is an organelle known as a microbody, that has been identified in virtually all eukaryotic cells (36,37). Peroxisomes are involved in the β-oxidation of fatty acids, the metabolism of amino acids and polyamines, the reduction of reactive oxygen species, specifically hydrogen peroxide, and the biosynthesis of plasmalogens in mammalian brains and lungs (39). While almost of the enzymes function within the organelle matrix, the peroxisome, unlike the mitochondria, has no genome and so these enzymes are all imported from the cytosol. The peroxins, which are referred to as Pex1p, Pex2p, Pex3p, and so forth up to Pex23p, function as part of the mechanism for the import of soluble peroxisomal enzymes from the cytosol into the peroxisomal matrix. Peroxisomal matrix proteins are synthesized by polyribosomes in the cytosol (40). Currently, two type of peroxisomal targeting signals (PTSs): PTS1 and PTS2 have been reported (41-43). Peroxisomal matirix proteins have either of one of the two PTS signals and are recognized in the cytosol by Pexs (44-48). PTS1 is found at the Carboxy- terminus and is encoded by the sequence Ser-Lys-Leu (SKL), or similar sequences (48). In the case of PTS1, proteins are recognized by Pex5p and the protein-Pex5p complexes are transported into the peroxisome via Pex14p. In contrast, PTS2 is an amino-terminal signal, [RK]-[LVI]-x5-[HQ]-[LA], which is formed by an α helix motif and is cleaved in

mammalian cells (42, 43, 49). In the case of PTS2, proteins are recognized by Pex7p and Pex5p and proteins-Pex7p-Pex5p complexes are transported into the peroxisome via Pex19p.

Carbonyl reductase (CR) (EC 1.1.1.184) catalyzes the NADPH-linked reduction of a variety of carbonyl compounds to their corresponding secondary alcohols (9). In mammalian tissues, the enzyme exists in several forms that differ in subcellular localization and/or subunit structure. The cytosolic enzymes purified from several mammalian tissues are monomers (around 30 kDa), and have been shown to catalyze isatin and 20-ketosteroids (11,50). A tetrameric form of CR (around 27 kDa for each subunit) exists in the lungs of guinea pigs, mice, and pigs, and is localized in the mitochondrial matrix (12, 13, 51). Although the two forms of CR belong to the short-chain dehydrogenase/reductase (SDR) family, they have low sequence identity (<30%) and differ in the number of amino acid residues (14, 15).

The first crystal structure of CR was reported for tetrameric mouse lung CR (MLCR) (19). Subsequently, the crystal structures of two monomeric CRs, porcine testicular CR and human CR, were determined (21, 52). In addition to the monomeric and lung CRs, oligomeric and soluble CRs have been isolated from dog liver, rabbit heart, and pig heart (16-18). Protein sequencing and cDNA cloning of pig heart CR and rabbit heart CR have

revealed that both enzymes belong to the SDR family, and that their subunits consist of 260 amino acid residues (18). Interestingly, the properties of these enzymes can be distinguished from those of the monomeric and lung CRs and in particular, the heart CR is localized in peroxisomes. In addition, both pig and rabbit heart CRs are tetrameric and show low sequence identity (<30%) with other mammalian monomeric and lung tetrameric CRs. With regard to the substrate specificity of the recombinant heart CRs, they have been shown to be able to uniquely catalyze the reduction of alkyl phenyl ketones and retinals/ retinols. In particular, pig and rabbit heart CRs can efficiently reduce all-*trans*-retinal. These reactions are not efficiently catalyzed by the monomeric and lung CRs.

The heart CR is localized in peroxisomes, and the enzyme is ubiquitously expressed in pig and rabbit tissues (18). Hence, in this thesis, I will refer to the heart CR as PerCR (peroxisomal CR). As I have discussed previously, pig and rabbit PerCR contain a C-terminal Ser-Arg-Leu (SRL) tripeptide, which is a variant of the PTS1 SKL, while, the C-terminal sequence of dog liver oligomeric CR is Ser-His-Leu (SHL) (48, 53). Notably, with regard to the availability of the C-terminal tripeptide in peroxisomal import, mutation of the PTS1 sequence of pig PerCR from Ser-Arg-Leu (SRL) to Ser-Lys-Leu (SKL), Ser-His-Leu (SHL), Ser-Leu-Leu (SLL), and Ser-Leu (SL) were performed to

assess the effect on the peroxisomal transport of pig PerCR in HeLa cells. Although the mutant enzymes containing the SKL and SHL sequences were imported into the peroxisomes and retained enzymatic activity, those with the SLL and SL sequences were not targeted to this organelle and were enzymatically inactive. Gel filtration data showed that the inactive SLL and SL mutants, as well as the active SKL and SHL mutants formed tetramers. Moreover, interestingly, the wild-type pig heart CR did not show specific targeting when transfected as a protein into cells.

To understand the structural basis for peroxisomal localization of PerCR, I decided to initiate a crystallographic analysis of PerCR.

Material and Methods

Expression and purification

The expression and purification of recombinant pig heart PerCR were performed as described in Usami et al. (18). Briefly, the expression plasmids (pCR-T7/CT TOPO vectors; Invitrogen, Thermo Fisher Scientific K.K., Japan) harboring the cDNA for PHCR were transformed into *Escherichia coli* BL21(DE3)pLysS cells (Invitrogen). Bacterial cultures were grown in LB medium (3 L shaker flask containing 1 L medium) at 310 K to an OD_{600} of 0.4. Expression of PHCR was induced by the addition of 1 mM IPTG for 6 h at 310 K. After this period, cells were harvested by centrifugation at 5000 × *g* for 15 min, suspended in lysis buffer (0.1% Triton X-100, 5 mM 2-mercaptoethanol, and 1 mM EDTA in 10 mM Tris-HCl buffer pH 7.5) and sonicated at 180 W for 10 min using an UH-150 Ultrasonic homogenizer (SMT Co., Tokyo, Japan). The cell extract was obtained by centrifugation at 12 000 × *g* for 15 min. The enzyme was purified from the cell extract by ammonium sulfate fractionation (30–75% saturation) and consecutive column chromatography on Sephadex G-100, Red-Sepharose, and hydroxylapatite. SDS–PAGE of the purified enzyme revealed a single 27 kDa protein band by Coomassie Brilliant Blue staining. The purified enzyme fractions were dialyzed against 200 mM sodium chloride, 5 mM 2-mercaptethanol, and 20%(v/v) glycerol in 10 mM phosphate buffer pH 7.0, and concentrated to 3 mg/mL by ultrafiltration using an Amicon YM-10

membrane and stored at 253 K. The enzyme is a basic protein with a pI value of 9.3 and was easily precipitated when concentrated to a protein concentration greater than 3.3 mg/mL. Avoidance of the precipitation during ultrafiltration resulted in a high yield (27 mg protein from 1 L of culture) of the homogenous enzyme compared with a previous purification (5 mg protein from 1 L culture (18)). The preparation of the *E. coli* cell extract and purification of the enzyme were carried out at 278 K.

Assay for enzyme activity

The reductase activity of the enzyme was determined by recording the change of absorbance of NADPH at 340 nm. The standard reaction mixture for the reductase activity consisted of 0.1 M potassium phosphate (pH 6), 0.1 mM NADPH, 1 mM 4-benzoylpyridine, and enzyme, in a total volume of 2.0 mL. One unit of enzyme activity was defined as the enzyme amount that catalyzes the reduction of 1 mmol of NADPH per minute at 25°C.

Crystallization

The stock solution of 3 mg/mL PerCR described above was dialyzed against 1 mM NADPH and 50 mM NaCl in 20 mM Tris-HCl buffer pH 7.5, and concentrated using a Centricon-30 (Millipore), yielding a working solution of 3 mg/mL PHCR with 1 mM

NADPH and 50 mM NaCl in 20 mM Tris-HCl buffer pH 7.5. Initial sparse-matrix crystal screening was conducted using Crystal Screen I (Hampton Research, USA) (54). Crystallization was carried out by the hanging-drop method, in which 1 μL of working solution was mixed with the same volume of crystallization buffer and incubated at 293 K. The drops were suspended over a 200 μL reservoir solution in 48-well plates.

X-ray data collection

Since the crystallization conditions for PerCR described above included 20%(v/v) glycerol in the reservoir solution, X-ray data collection could be performed under cryogenic conditions, without further addition of cryoprotectant. Thus, crystals from the hanging drop were directly mounted in nylon loops and flash-cooled in a cold nitrogen-gas stream at 100 K just prior to data collection. Data collection for the form I crystal was performed by the rotation method at 100 K using an ADSC Q210 CCD detector with synchrotron radiation (λ = 1.000 Å at beamline NW12 of the Advanced Ring of the Photon Factory, Tsukuba, Japan). The Laue group and unit-cell parameters were determined using the *DPS* program package (55).

Structure determination

The initial phase determination was performed by the molecular replacement (MR) technique with the coordinate set of the ternary complex of mouse lung carbonyl reductase, MLCR (PDB code; 1CYD) as a search model. Crystallographic refinement was performed with the program *REFMAC* (55).

Structural analysis

Least-squares comparisons of the molecular models were performed using the *DALI* server (55). Figures were produced using the *ViewerPro* program (Figures 3, 4, 5, 6, 7A, 7B, and 7D, and Figure S1) and the *Raster3D* and *XtalView* programs (57, 58). Other figures were produced using the program *Chimera*.

Results

Crystallization

Several microcrystals, and two morphologically distinct crystals, were observed within a week. Crystals of form I grew as rod-shaped crystals from condition No. 38 in Crystal Screen I (100 mM HEPES buffer and 1.4 M sodium citrate as a precipitant). Crystals of form II grew as rectangular-shaped crystals from condition No. 4 in Crystal Screen I (100 mM Tris-HCl buffer and 2.0 M ammonium sulfate as a precipitant). Trials to improve the crystallization conditions were performed for the form I and II crystals by varying the pH, buffer system, and precipitant concentration.

Simultaneously, I also examined the addition of glycerol to the reservoir solutions.

In order to obtain form I crystals suitable for X-ray analysis, a droplet was prepared by mixing equal volumes (2.0 + 2.0 μL) of the working solution (3 mg/mL PerCR) described above, and the reservoir solution [1.2 M sodium citrate and 20% (v/v) glycerol in 100 mM HEPES buffer pH 7.5], and this was suspended over a 500 μL reservoir solution in 24-well plates. Rod-shaped crystals with typical dimensions of approximately 0.1× 0.1 × 0.5 mm were grown in one week (Fig. 1A).

In order to obtain form II crystals suitable for X-ray analysis, a droplet was prepared by mixing equal volumes (2.0 + 2.0 μL) of the working solution (3 mg/mL PHCR) described

above, and the reservoir solution [1.4 M ammonium sulfate and 20% (v/v) glycerol in 100 mM MES buffer pH 6.0], and was suspended over a 500 μL reservoir solution in 24-well plates. Rectangular-shaped crystals with typical dimensions of approximately 0.1 × 0.2 × 0.3 mm grew in one week (Fig. 1B).

X-ray data collection

The Laue group was found to be 4/m, the unit-cell parameters were $a = b = 109.61$, $c = 94.31$ Å, and the space group was $P4_2$. For the form II crystal, the Laue group was found to be 4/mmm, and the unit-cell parameters were $a = b = 120.10$, $c = 147.00$ Å and the tetragonal space groups were $P4_12_12$ or $P4_32_12$. The data-collection statistics are summarized in Table 1.

Crystallographic analyses

Initial phase determination for the form I crystal was performed by the molecular-replacement (MR) technique using the coordinate set of the whole tetramer of MLCR (PDB code; 1CYD), which has approximately 27% amino-acid sequence identity to PerCR, as a search model. The bound NADPH, 2-propanol, and water molecules were removed from the search model. Cross-rotation and translation functions were calculated

using the program *AMoRe* from the *CCP*4 suite (59). The results showed a clear solution [correlation coefficient of 0.321 (the first noise solution was 0.286) and *R* factor of 0.539 (0.553) in the resolution range 15.0–2.5 Å] and a reasonable molecular arrangement of PerCR in the ASU. The MR solution was supported by the observation that the directions of the non-crystallographic (NCS) two-fold axes determined by the self-rotation function, showing that the 222 point group symmetry of the tetrameric PerCR molecule, were consistent with the MR solution obtained. The model was improved using a manual model building with the program *XtalView*, and refined to a resolution of 1.5 Å with an R factor of 0.223 (free *R* factor of 0.244), without incorporating the bound NADPH and water molecules, using *REFMAC* (58, 60). A random subset of the data (5%) were not included in the refinement. Refinement of the form I crystal at 1.5 Å is incomplete at this time, and work on it continues. The incompletely refined form I model could nevertheless be used for structure determination of the form II crystal by the MR method, which was performed by procedures similar to those described above. The space-group ambiguity ($P4_12_12$ or $P4_32_12$) for the form II crystal was resolved by calculating the translation function for either case. The results showed a clear solution [correlation coefficient of 0.729 (the first noise solution 0.398) and *R* factor of 0.397 (0.562), in the resolution range 15.0–2.5 Å, and a reasonable molecular arrangement of PerCR in the ASU for the space group $P4_12_12$.

Finally, the model was refined to an *R*-factor of 0.161 (free *R*-factor of 0.188) at 1.5 Å resolution. The final model includes one NADPH molecule, and all of the non-hydrogen atoms except for the nine N-terminal residues (Met1 to Arg9) for each of the four crystallographically independent subunits.

In addition, a total of 772 water molecules are included per ASU. The refinement statistics are summarized in Table 2.

Effects of mutations of the C-terminal tripeptide on peroxisomal localization and enzyme activity

Four mutant pig PerCRs where the C-terminal peroxisomal targeting signal was changed from SRL to SHL, SKL, SLL, and SL were generated and expressed in *Escherichia coli* cells. The enzyme activity of the SHL and SKL mutants could be detected, however the SLL and SL mutants were catalytically inactive, although all the enzymes were expressed appropriately, as assessed by western blotting analysis (Figure 2A). The SHL mutant was purified using a three-column chromatographic fractionation procedure using Sephadex G-100, Red-Sepharose, and hydroxylapatite, and a homogeneous preparation with a 24% yield and a specific activity of 22 U/mg was obtained. However, the SKL mutant gradually lost activity during incubation at 0°C (Figure 2B). The inactivation was prevented by adding 20% (v/v) glycerol containing 0.1 mM NADPH or 0.5 M KCl, but it was not due to the incubation at low temperature, because the enzyme lost 70% of its activity after incubation for 15 min at 25°C. The SKL mutant was further purified by re-chromatography on a Red-Sepharose column using a buffer containing 20% glycerol, and the purified protein was obtained as an electrophoretically homogeneous preparation with a specific activity of 28 U/mg, although the yield was low (5.4%). The molecular weights of the two mutant enzymes (SHL and SKL) were 27 kDa by SDS-PAGE, and

approximately 100 kDa following Sephadex G-100 gel filtration, confirming their tetrameric structures. The substrate specificities of the purified SHL and SKL mutants were essentially identical to that of the wild-type enzyme (18). For example, the K_m values for 4-benzoylpyridine, ethyl benzoylformate, 1-phenylisatin, and NADPH were 380, 19, 6, and 6 mM, respectively, for the purified SHL mutant, whereas the respective values of the purified SKL mutant were 380, 29, 5, and 5 mM. When the inactive recombinant SLL and SL mutants were partially purified, they emerged at the same elution volume as that of the wild-type enzyme on Sephadex G-100 gel filtration, suggesting that these mutant enzymes retain their tetrameric structures.

Tertiary structure of PerCR

The subunit of PerCR is a single-domain protein having an α/β double-wound structure (Figure 3). Seven β-strands are found as a βsheet in the center of the molecule and three helixes sandwich this β-sheet. Two short helixes (αFG1 and αFG2) are located outside of the main body. These helixes are known to be involved in substrate binding among enzymes in the SDR family (19). The C-terminus containing the PTS1 was found to be in the structural core of the subunit.

Quaternary structure

With respect to overall structure, the tetrameric PerCR is similar in conformation to tetrameric SDRs (Figure 4). As I have discussed previously, two short helixes (αFG1 and αFG2) are also found outside the main body of the enzyme (Figures 4B and 4C). In tetrameric SDRs, these four subunits are arranged with a 222-point group symmetry. The three mutually perpendicular two-fold axes in the tetrameric SDRs are designated as the P, Q, and R-axis, respectively (15).

Subunit interactions were found to occur in four main areas; i) αE-αE, ii) around αF-αF, iii) αG-αG, and iv) C-terminal (Figure 4). The most extensive inter-subunit interactions are found between the Q-axis related subunits. The αE and αF helixes of one subunit interact with the αE and αF helixes of the neighboring subunit, respectively. Hydrophilic interactions are found in the αE-αE helixes interaction of PerCR, (Asp110, Glu113, Asp117, His121, Lys125 and Lys132) (Figure 6A, 6B).

Generally, in the interaction between the αE-αE helixes of SDRs, the aromatic side chains face each other. In the case of *E. coli* 7α-hydroxysteroid dehydrogenase (7α-HSD), Phe111, Tyr115, Phe120, and Phe123 are observed, whereas no aromatic side chains, with the exception Trp116 (which corresponds to Phe111 in 7α-HSD), are found in the PerCR αE-αE helixes interaction (19). Tyr115, Phe120, and Phe123 in 7α-HSD are replaced by Leu120, Lys125, and Val128, respectively, in PerCR. Thus, it is remarkable

that the interaction between αE-αE helixes has a hydrophilic nature in PerCR and is more hydrophobic in other SDRs.

With respect to the αF-αF helixes interaction in PerCR, the hydrophobic side chains (Val166, Thr169, and Gly173) interact with each other, as well as in MLCR and 7α-HSD. (Figure 6B.) Interestingly, the hydrophobic pocket, which consists of Leu172, Gly173, and Lys176 in the center (red frame) between αF-αF helixes interfaces to stabilize the side chain of Leu260 which is the C-terminus of neighboring subunits related to the R-axis (Figure 6C).

For the αG-αG helixes interaction, and C-terminal interaction related to the R-axis, the interaction is not as extensive as is found for the αE-αE and αF-αF helixes interactions. The benzyl groups from Phe230 are stacked against each other (Figure 6D).

With respect to the C-terminal interaction, (Figure 7), each C-terminus involving the peroxisomal targeting signal (PTS1: Ser258-Arg259-Leu260) is well ordered and shielded from the solvent in the tetramer (Figure 7A). Interestingly, the P-axis-related PTS1 signals interact with each other by a hydrogen bond/salt bridge network (Figure 7A). The side chain of Ser258 of subunit A [Ser258(A)] is involved in a hydrogen bond network formed by Lys176(D) and Glu248(D) via water molecules (Figure 5B). Moreover, the carbonyl group of the main chain forms a hydrogen bond with His156 of

the same subunit to stabilize the C-terminal chain. The guanidyl group of Arg259(A) mainly interacts with the side chain of Glu248(D) and the C-terminal carboxylate group of Leu260(D). (Figures 7C). Interestingly, the hydroxyl group of Glu248(D) adopts two conformations which are i) a direct interaction between Glu248(D)-Arg259(A) and ii) a water mediated interaction between Glu248(D) and Arg259(A) (Glu248(D)-Wat-Arg259(A)). Occupancy was fixed at 0.7 for a direct interaction and at 0.3 for a water mediated interaction. The side chain of Leu260(A) interacts with hydrophobic pocket, as I have previously described, which consists of Leu172(D), Gly173(D), and Lys176(D) from the P-axis-related subunit, and the side chains Tyr155(C), Thr169(C), and Leu172(C) from the R-axis-related subunit (Figure 7D).

The inter-subunit (intramolecular) PTS1 recognition mechanism in PerCR can be summarized as follows: the side chains of Ser258, Arg259, and Leu260 are recognized by hydrogen bonds (i.e. a steric confinement that may accept only small residues), electrostatic interactions, and hydrophobic interactions, respectively. In the present PerCR structure, only a few hydrophobic interactions are found in the R-axis interface, as described above. In the case of MLCR, the R-axis interface involves one prominent interaction: the C-terminal carboxylate group belonging to one subunit forms a salt bridge with the guanidine group of Arg203 belonging to the other subunit (19).

Thus, the C-terminal carboxylate group of PerCR and that of MLCR are involved in different inter-subunit interactions: the P-axis interfaces in PerCR and the R-axis interfaces in MLCR.

Coenzyme binding site

In the coenzyme binding site of PserCR, the interactions with NAD(P)(H) are well conserved with previously reported SDRs (19). In particular, the coenzyme-binding mode of PerCR is quite similar to that of MLCR (Figure 8).

Active site residues

The catalytic mechanism of the SDRs is also well known. In particular, the Ser-Tyr-Lys catalytic triad plays an important role in the catalytic mechanism. The Ser residue plays a role in the catalytic step as a stabilizer of the reaction species, the Tyr residue is a catalytic residue, and the Lys residue has a dual role in orienting the nicotinamide moiety of the coenzyme and lowering the pK_a value of the Tyr residue (20). As has been observed for many other SDRs, PerCR has the same catalytic triad: Ser151, Tyr164, and Lys168 (Figure 8). A water molecule is found close to the OH groups of Ser151 and Tyr164, and it is suggested that this molecule occupies the binding site for the carbonyl-oxygen atom of the substrate. Filling et al. proposed a new catalytic mechanism that extends the above-

mentioned catalytic triad to form a tetrad of Asn-Ser-Tyr-Lys from a high-resolution crystal structure analysis of 3β/17β-HSD (62). The Asn residue plays a role in proton relay. The side chain atoms of Asn123 contribute to a hydrogen bond with the main chain atoms of Val101 (OD1(Asn123)—N(Val101) and ND2(Asn123)—O(Val101)), which constitute the right-side wall of the catalytic cleft. These residues (Asn-Ser-Tyr-Lys) are also well conserved in PerCR (Asn123, Ser151, Tyr164, and Lys168), suggesting that the catalytic mechanism of PerCR appears to be similar to that proposed for other SDRs (20, 62).

Discussion

I have determined the crystal structure of the binary complex of PerCR. Although, the amino acid identity is low (<30%) compared to other SDRs, the overall structure adopts a similar conformation. In addition, both the coenzyme binding site and the active site are well conserved.

Hypothetical model for peroxisomal localization of PerCR

The structure reveals that the SRL motif located in the C-terminus contributes to inter-subunit interactions in PerCR, and is buried in the internal region of the tetrameric PerCR molecule. The structure provokes the question as to whether tetrameric PerCR molecules

injected into cells would be imported into the peroxisome. With regard to the import of proteins into the peroxisome, Pex5p which is a cytosolic receptor that requires an interaction with PTS1 (43, 44). In other words, the PTS1 recognition sequence cannot be recognized by Pex5p in the tetrameric form and this suggests that the tetrameric form of PerCR is not formed in the cytosol.

Hence, on the basis of the present crystal structure analysis and the biochemical studies of PerCR, I propose a possible model for the peroxisomal targeting mechanism of PerCR, assuming that Pex5p recognizes the monomeric form of PerCR in the cytosol (Figure 9). First, the exposed PTS1 in monomeric PerCR (yellow) is recognized by Pex5p in the cytosol. Second, monomeric PerCR is transported into the peroxisome by the Pex5p/Pex14p complex. Finally, the monomeric PerCR molecules (yellow, blue, red, and green) form a tetramer in the peroxisome.

With respect to the possibility of oligomerization of PerCR in cytosol, an immunofluorescence study of PerCR expressed in HeLa cells showed that the SRL tripeptide in the enzyme sequence also acts as a peroxisomal-targeting signal for this enzyme (18). Moreover, a peroxisomal isoenzyme of $NADP^+$-dependent isocitrate dehydrogenase has been suggested to provide the intra-peroxisomal supply of NADPH (63). Thus, these data suggest that oligomerization of PerCR does not occur in the cytosol.

With respect to the import mechanism from cytosol to the peroxisome, although further detailed research is needed, I assume that molecular chaperones within the cytosol prevent the assembly of PerCR into tetramers and retain it in its monomeric form. The cytosolic chaperone Hsp70, is well-known to unfold proteins and play a role in protein import into mitochondria and the endoplasmic reticulum (64). It has also been reported that Hsp70 is involved in the import of proteins into the peroxisomal matrix (65).

Two previous reports have shown that mitochondrial precursor proteins are partially unfolded to the monomeric state in order to be translocated across the mitochondrial membrane (66). In fact, both catalase and alcohol oxidase are imported into the peroxisome as monomers, and their oligomerization is observed to occur within the peroxisome (36, 67). Although excellent cell-biological studies have been conducted for both catalase and alcohol oxidase, the present study clearly provides the first structural explanation of the PTS1 recognition mechanism and the import model for an oligomeric protein, which is imported into the peroxisome as a monomer.

The PTS recognition of Pex5p

When monomeric PerCR is imported to the peroxisome, I have to consider the C-terminal interactions, including PTS1. In the case of monomeric PerCR, the C-terminal residues (Gly253-leu260) are located outside of the main body of the enzyme structure

and have no electrostatic or hydrophobic interactions (Figure 10A). However, according to the surface model of tetrameric PerCR, the C-terminal residues are buried within the molecule (Figure 10B).

Interestingly, the C-terminal residues are well conserved in other species (Figure 10C). Thus, these residues are more flexible in the PerCR monomeric state.

With regard to the PTS-Pex5p interaction, the crystal structure of Pex5p with the signal peptide has been resolved (68). Based on the crystal structure of Pex5P, for Pex5p-PTS1 recognition, the tetratricopeptide repeat (TPR) domain of Pex5p (which has a wide active site), interact with the PTS within the TPR active site which has a diameter of approximately 8 Å. Based on this report, the Leu (-1) position requires a hydrophilic amino acid and the Lys (-2) position requires a basic amino acid residue. Finally, the Ser (-3) position requires small residues, such as Ser, Ala, and Cys, due to the sterically confined space (Figure 10D).

It is well accepted that PTS1 in metazoans is typically composed of amino acids with small side chain residues (Ser, Ala, Cys) at the -3 position, positively charged residues (Lys, Arg, His) at the -2 position, and large hydrophobic residues (Leu, Met) at the -1 position (69). Thus, the intramolecular (inter-subunit) PTS1 recognition mode observed in PerCR is equivalent to the canonical intermolecular PTS1 recognition mode observed

for the Pex5p/ PTS1 peptide complex.

Significance of the C-terminal S-(R/H)-L tripeptide of PerCR

With regard to the subcellular localization study of the pig PerCR mutants expressed in HeLa cells, the data show that the C-terminal S-(R/H/K)-L tripeptide of the enzyme, which functions as PTS1, is found in peroxisomal proteins and suggests that dog liver oligomeric CR containing the C-terminal SHL tripeptide is also a peroxisomal protein, despite the fact that this enzyme is designated to be a cytosolic oligomeric CR (16, 53). The C-terminal amino acid sequences of the mammalian enzymes are well conserved (Figure 10C), however the SKL sequence which has been reported in many peroxisomal proteins is not well conserved. With respect to subunit-subunit interactions, although the Q-axis-related dimer shows a strong electrostatic interaction between αE-αE and αF-αF helixes and the R-axis-related dimer shows a strong hydrophobic interaction between the hydrophobic pocket in C-terminal and Leu260. In particular, the hydrogen bond between Arg259(A) and Glu248(D) is important. In fact, following mutation of the second residue from SRL to SHL, SKL, SLL, and SL in the C-terminal of pig PerCR, the SLL and SL forms lost enzymatic activity and stability. These data suggest that the C-terminal residues are important for the stability of the architecture of tetramer. In addition, the PerCR enzyme with a Lys (K) mutation had impaired the stability (Figure 2B), whereas

the wild type and SHL mutant had similar stability and activity. This is the reason that the Arg to His mutation can maintain the hydrogen bond network that exists in the wild-type enzyme. The hydrogen bonds in the SHL mutant are predicted to be the hydrogen bond (NH1(Arg259(A))-Wat and NE(Arg259(A))-OXT(Leu260(D)), respectively, in the wild type enzyme. On the other hand, the SKL mutant can retain either of the two hydrogen bonds, because the Lys side chain has only one nitrogen atom. Thus, the SKL sequence is unfavorable as the PTS1 in tetrameric mammalian PerCRs (Figure 10C), rather than their inability to bind Pex5p.

Conclusion

I have determined the crystal structure of pig heart PerCR complexed with its coenzyme NADPH at 2.2 Å resolution. The C-terminal SRL motif is involved in the inter-subunit interactions related to the R-axis of PerCR and is buried in the interior of the tetrameric PerCR molecule. The structure reveals why the tetrameric PerCR molecules injected into the cells are not imported into the peroxisome and it enables me to propose a hypothetical model for the peroxisomal targeting of PerCR. Additionally, the analysis of the geometry of the inter-subunit interactions has allowed me to understand the role of the C-terminal sequence in maintaining an active tetrameric form of this enzyme.

Chapter II

Crystallization and Preliminary X-ray Crystallographic Study and Structural Basis for the Substrate Recognition Mechanism of Phosphoglucose Isomerase from *Plasmodium falciparum*.

Abstract

Malaria is one of the world's most serious infection diseases in which parasites infect human tissues and blood, and rapidly multiply in the presence of blood glucose. An estimated 300–500 million new malarial cases and up to 2.7 million malaria-related deaths are reported each year. In addition, recently, cases of resistance to malarial treatments have been reported. Thus, the development of new anti-malarial drug is urgently needed. Phosphoglucose isomerase (PGI) is a key enzyme involved in glycolysis and gluconeogenesis that catalyzes the interconversion of glucose 6-phosphate (G6P) and fructose 6-phosphate (F6P).

For crystallographic studies, PGI from the human malaria parasite *Plasmodium falciparum* (PfPGI) complexed with 6-phosphogluconic acid (6-PGA) was overproduced in *Escherichia coli*, purified, and crystallized using the hanging-drop vapor-diffusion method. X-ray diffraction data to 1.5 Å resolution were collected from an orthorhombic crystal form belonging to space group $P2_12_12_1$ with unit-cell parameters a = 103.3, b = 104.1, c = 114.6 Å.

A comparison of the PfPGI-6PGA and mPGI-6PGA, revealed there to be differences in the catalytic recognition mechanism, particularly in the interaction between Lys540 of

PfPGI and O3 of 6PGA, a water mediating this interaction, in contract to the direct interact between Lys519-O3 and PGA in mammalian PGI.

In addition, PfPGI has two malaria original loop domains (MOL1; residues 175-200, MOL2; residues 438-442) which are well conserved in *Plasmodium vivax*, *Plasmodium malariae*, and *Plasmodium ovale*, and make a strong contribution to the dimerization between subunit A and subunit B. Moreover, the MOL2 domain is inserted between the carbohydrate moiety-binding helix, including Lys540 (CBH; α37 helix; residue 534-555) and α15 helix* of the small domain (* the neighbor subunit) and maintains the binding site space through a different mechanism from that of mPGI (steric hindrance and numerous hydrogen bonds in CBH*-MOL2-α15 helix* in PfPGI versus one hydrogen bond in CBH-α15 helix* in mPGI). These findings might explain the CBH conformational difference between Lys540 (PfPGI) and Lys519 (mPGI) and thus the difference in catalytic activity observed in mouse and plasmodium.

Introduction

Malaria is one of the world's most serious parasitic diseases. An estimated 300–500 million new cases of malaria and up to 2.7 million malaria-related deaths are reported each year. Human malaria is caused by infection with protozoan parasites of the genus *Plasmodium* that are transmitted by *Anopheles* mosquitoes. *Plasmodium falciparum* is the most lethal of the four species of Plasmodium (*P. falciparum, P. vivax, P. ovale, and P. malariae*) that infect humans. Chemotherapy for malaria is available, but the emergence of strains that are resistant to conventional drug therapy has stimulated the search for anti-malarial agents with novel modes of action. Proteins important for survival of the parasite and that show differences in structure from the host homologue serve as potential drug targets. Such proteins can be classified as surface proteins, proteins involved in invasion, and metabolic enzymes (70).

Phosphoglucose isomerase (PGI; EC 5.3.1.9), also known as glucose 6-phosphate isomerase (GPI) or phosphohexose isomerase (PHI), is a key enzyme in glycolysis and gluconeogenesis. It catalyzes the second step of glycolysis, namely the interconversion between glucose 6-phosphate (G6P) and fructose 6-phosphate (F6P). In recent years, several other protein factors have been confirmed to be identical to PGI: autocrine motility factor (AMF), neuroleukin (NLK), maturation factor (MF), antigens involved in

rheumatoid arthritis and sperm agglutination, and a novel serine-protease inhibitor (MBSPI) (26-28, 71-74). PGI therefore not only acts as a housekeeping enzyme for sugar metabolism inside the cell, but also has various cytokine properties (AMF, NLK, MF), as well as several other functions outside the cell. The cytokine activity of PGI is specific to mammalian PGIs, and the enzymatic activity of PGI is not essential for either the receptor-binding or the cytokine activity of mammalian PGIs (75, 76). The enzymatic activity of the PGI from *P. falciparum* (PfPGI) was found to be comparable to that of mammalian PGIs, but the cytokine activity of PfPGI against mammalian cells was not detectable (Haga et al., unpublished data). Since PGI is a housekeeping enzyme, its catalytic residues should be well conserved amongst PGIs from various species. However, significant amino-acid sequence differences (insertions and deletions) are observed between the mammalian PGIs (558 residues) and PfPGI (579 residues). Thus, a detailed comparison of the three-dimensional structures of human and plasmodium PGIs is important to understand whether these differences can be applied to future structure-based drug design, and to gain insight into the structural origin of the species specificity of the cytokine activity of PGI. To date, crystal structures of PGI from human, pig, mouse, rabbit, *Bacillus stearothermophilus*, *Pyrobaculum aerophilum*, *Leishmania mexicana*, *Thermus thermophiles*, and *Trypanosoma brucei* have been reported (29-36, 77-82). With respect

to the crystal structure of *Plasmodium falciparum*, although Gileadi et al. resolved the substrate-free form of PfPGI and PfPGI complexed with fructose 6-phsophate (F6P) and deposited the structures in the PDB (PDB code; 3QKI and 3PR3), these data have not been published. Here, I determined the crystal structure of PfPGI complexed with two carbohydrate inhibitors of different lengths (6-phosphogluconic acid (6PGA) and 3-phosphoglyconate (3PGA)) and its substrate (fructose 6-phosphate (F6P)) at high resolution (1.5–2.0 Å). In addition, I also compared the structural differences between these PfPGIs and the crystal structures of mPGI-6PGA (PDB code; 2CXR), mPGI-3PGA (in-house data), and mPGI-F6P (PDB code; 2CXS) (35). These findings therefore provide key insights into the design of new anti-malarial compounds.

Material and Methods

Overproduction and purification

The cDNA encoding the full-length *P. falciparum* PGI (residues 1–579) was obtained by reverse-transcription PCR. Reverse transcription was carried out using SuperScript II reverse transcriptase, as described in the user's manual (Invitrogen, Thermo Fisher Scientific K.K., Japan), with total RNA derived from *P. falciparum* (FCR-3) as the template. The target DNA was PCR amplified from the reverse-transcription products using AccuPrime Pfx DNA polymerase (Invitrogen) with 5'-CGC**GGATCC**ATGAATATGGAGATTACAAAT-3' and 5'-CGC**AAGCTT**ATTTGGACAAGTAATAATTTA-3' as the forward and reverse primers, respectively. The PCR product was cloned into the pQE30 expression plasmid (Qiagen, QIAGEN K.K., Japan.) with the BamHI and HindIII cloning sites (bold). The construct was verified by sequencing.

Escherichia coli BL21 (DE3) cells (Novagen; Merck Millipore, USA) harboring the expression plasmid were grown in LB medium (3 L shaker flask containing 1 L medium) at 310 K to an OD_{600} of 0.6. Overproduction of PfPGI was induced by the addition of 0.5 mM IPTG for 4 h at 310 K. After this period, the cells were harvested by centrifugation at 8000 × *g* for 15 min, suspended in buffer A (20 mM Bis-HCl pH 6.5, 50 mM NaCl), and disrupted by ultrasonication on ice for 4 × 30 s. The cell extract was obtained by

centrifugation at 15000 × *g* for 15 min and was applied onto a 5 mL HiTrap SP HP column (GE Healthcare, UK) equilibrated with buffer A. The column was washed with 20 column volumes of wash buffer (0.3 M NaCl in buffer A). After washing, the PfPGI was eluted with a linear gradient of 0.3–0.7 M NaCl in buffer A. The PfPGI was further purified by gel chromatography using a Superdex 200 PG column (GE Healthcare, UK) equilibrated with buffer A. Fractions containing PfPGI were pooled and concentrated to 5 mg/mL using an Amicon Ultra-15 (Merck Millipore, USA).

Crystallization

The protein solution (5 mg/mL PfPGI) was mixed with 10 mM 6-phosphogluconic acid (6PGA) dissolved in buffer A at a volume ratio of 1:1. Initial sparse-matrix crystal screening was conducted using Crystal Screen I (Hampton Research) and Cryo I and II (Emerald BioSystems) (54). Crystallization was carried out by the hanging-drop vapor-diffusion method. Briefly, 1 mL of protein solution (5 mg/mL protein and 5 mM 6PGA) was mixed with the same volume of reservoir solution and incubated at 293 K. Drops of the solution were then suspended over a 200 μL reservoir solution in 48-well plates.

X-ray data collection

Since the crystallization conditions for PfPGI included 38% (v/v) PEG 400 in the

reservoir solution, X-ray data collection could be performed under cryogenic conditions without further addition of cryoprotectant. Crystals from the hanging drop were directly mounted in nylon loops and flash-cooled in a cold nitrogen-gas stream at 100 K just prior to data collection. Data collection was performed by the rotation method at 100 K using an ADSC Q315 CCD detector with synchrotron radiation [λ = 1.000 Å on beamline 5A at the Photon Factory (PF), Tsukuba, Japan]. The Laue group and unit-cell parameters were determined using the *DPS* program package (55).

Structure determination

The initial phase determination was performed using the molecular replacement technique with the coordinate set of mouse Autocrine Motility Factor, mPGI (PDB code; 2CXR) as a search model (34). Crystallographic refinement was performed with the *REFMAC* program (59).

Least-squares comparisons of the molecular models were performed using the *DALI* server. Figures were generated using the *CHIMERA* program package.

Results

Purification and crystallization

PfPGI was successfully cloned, overproduced, and purified to homogeneity. SDS–PAGE of the purified enzyme revealed a single 65 kDa protein band by Coomassie Brilliant Blue staining (Figure 11A).

Initial crystal screening produced several microcrystals within one week. Micro-plate crystals grew from condition Nos. 6, 13, 24, and 38 for Cryo I [No. 6, 40% (v/v) PEG 600 and 0.2 M calcium acetate in 0.1 M sodium cacodylate–NaOH pH 6.5; No. 13, 30% (v/v) PEG 200 and 5% (w/v) PEG 3000 in 0.1 M MES–NaOH pH 6.0; No. 24, 40% (v/v) PEG 400 and 0.2 M sodium chloride in 0.1 M sodium/potassium phosphate pH 6.2; No. 38, 40% (v/v) PEG 400 and 0.2 M lithium sulfate in 0.1 M Tris-HCl pH 8.5] and No. 4 for Cryo II [40% (v/v) PEG 400 and 0.2 M calcium acetate in 0.1 M HEPES–NaOH pH 7.5]. Trials to improve the crystallization conditions were performed by varying the pH, the buffer system, and the concentration of the crystallizing agent. To obtain crystals suitable for X-ray analysis, a droplet was prepared by mixing equal volumes (2 μL + 2 μL) of the working solution described above, and a reservoir solution [38%(v/v) PEG 400 and 0.2 M calcium acetate in 0.1 M sodium cacodylate–HCl pH 6.5] and was suspended over a 500 μL reservoir solution in 24-well plates. Plate-shaped crystals with typical dimensions of approximately 0.3 × 0.1 × 0.03 mm grew within one week (Figure 11B).

The crystallization of the PfPGI with the inhibitor, 3-phosphoglycerate (3PGA) a phosphate containing carbohydrate, and the substrate F6P were also prepared using the same regimen described above at a final inhibitor concentration of 10 mM.

Data collection

The Laue group of the PfPGI crystals was found to be mmm, with unit-cell parameters a = 103.3, b = 104.1, c = 114.6 Å. Only reflections with h = 2n, k= 2n and l = 2n were observed along the [h00], [0k0], and [00l] axes, respectively, indicating the orthorhombic space group $P2_12_12_1$. Assuming the presence of two subunits (one dimer) per asymmetric unit led to an empirically acceptable V_M value of 2.29 $Å^3$ Da^{-1}, corresponding to a solvent content of 46.3% (83). The current best diffraction data from the PfPGI crystal were collected to 1.5 Å resolution. The data-collection statistics are summarized in Table 3.

Initial phase determination

Initial phase determination was performed by molecular replacement (MR) with the coordinates of the dimeric molecule of the mouse PGI–6PGA complex (PDB code; 2CXR), which shares approximately 36% amino-acid sequence identity with PfPGI, as a search model. The bound inhibitor and water molecules were removed from the search model. Cross-rotation and translation functions were calculated using the *MOLREP*

program from the *CCP4* suite (84). The results showed a clear solution [correlation coefficient of 0.313 (the first noise solution was 0.215) and an *R* factor of 0.519 the first noise solution was 0.554) in the resolution range 40.0–3.0 Å] and a reasonable molecular arrangement of PfPGI in the asymmetric unit. The MR solution was supported by the observation that the directions of the non-crystallographic two-fold axes determined by the self-rotation function (data not shown) were consistent with the MR solution obtained. Automatic model building and refinement using the *ARP*/*wARP* and *REFMAC5* programs, as well as manual model building and refinement with the *XtalView* (McRee, 1999) and *REFMAC5* programs (58, 60, 85).

Structure determination

I determined the crystal structure of PfPGI-inhibitor (6PGA) complex at 1.7 Å resolution and refined it to an *R*-factor of 0.166 (R_{free} of 0.193). The final model comprised four identical protein subunits (two homodimers) with 2 × 579 amino acid residues and 1016 water molecules per asymmetric unit (ASU). The amino acid residues between Met1 – Met3 at the N-terminus of each subunit, and between Lys557 – Thr562 in the C-terminus, were disordered in the B subunit. The crystal structure of the PfPGI-the 3PGA and the PfPGI-F6P complexes were are also determined by the difference Fourier method, using the refined model of the PfPGI-6PGA complex model (excluding water molecules)

as a starting model, and further refined to an *R*-factor of 0.190 (R_{free} of 0.231) at 2.0 Å resolution, and an *R*-factor of 0.183 (R_{free} of 0.209) at 1.7 Å resolution, respectively. In each final model, two identical protein subunits (two homodimers) were also comprised of 2 × 579 amino acid residues, 2 × 3PGA molecules (one molecule per subunit)/777 water molecules per asymmetric unit, and 2 × F6P molecules (one F6P molecule per subunit) /925 water molecules per asymmetric unit, respectively. In the 3PGA complex model, electron density map 3PGA, with the exception of the phosphate region, was not observed. The data collection and refinement statistics are summarized in Table 4.

Overall structure of the PfPGI subunit

The overall structure of PfPGI (Figure 12A, 12B) is similar to that of mPGI, although, the amino acid sequence homology is low (approximately 36%) between *Plasmodium falciparum* and mouse (30, 35). One dimer of PfPGI was contained in the asymmetric unit in the PfPGI crystal. The dimer is spherical in shape. The subunits interact in an arm-to-arm hugging manner, with intimate contacts at the catalytic site.

Based on previous PGI reports, a PfPGI subunit can be characterized as having three domains; the large (residues 1–101 and 310–533), small (residues 102–309), and C-terminal (residues 533–579) domains (Figure 12A, 12B). Interestingly, in PfPGI, an additional loop (Maralia Original Loop; MOL) domain was found, which is not found in

mammalian PGIs.

The large domain has an α/β structure composed of a six-stranded β-sheet. The β-sheet core is surrounded by the α helices, loops, and a "hook" region (residues 446–489)(Figure 12A).

The small domain has an α/β structure composed of a five stranded parallel β-sheet core (β4–β7) surrounded by α helixes, and loops. (Figure 12A).

The C-terminal domain is composed of two α helixes (α37 and α38) and a loop (Figure 12B).

The MOL domain was found in two insertion loops (MOL1; residues 175-200, and MOL2; residues 438-446) (Figure 12C). It is interesting to note that MOL1 is rich in asparagine (Asn) residue and forms an Asn-Asn interaction network (Asn176, Asn182, and Asn192) to maintain the folding of the loop. MOL2, which is inserted between the small domain and the C-terminal domain, contribute to dimer interactions.

Dimeric interaction between the two subunits

In PfPGI, a large number of hydrogen bonds are found compared to mammalian PGI (Figure 13A, 13B). In mammalian PGIs, only four hydrogen bonds have been reported. However, two out of these four hydrogen bonds are lost in PfPGI. Hydrogen bonds are found between the carboxylate group of Glu238 and the hydroxyl group of Thr407

(OG1(Thr407)-OE2(Glu238*); 2.74 Å, *: neighbor subunit) and the hydroxyl group of Gln365 and the amide group of Asn207* (OE1(Gln365)-ND2(Asn207); 2.87 Å). These hydrogen bond are also in the same positions in mammalian PGIs. The Glu238 residue (corresponding to Glu217 in mPGI and hPGI, and Glu216 in rabbit PHI) is assumed to be part of the catalytic diad (77).

In addition in PfPGI, additional dimeric interactions were found in MOL1, MOL2/α33 helix, and the C-terminal domain (Figure 13B, 13C, 13D).

In MOL1, hydrogen bonds between the amide group of Asn176 and the hydroxyl group of Asn214* (ND2(Asn176)-OD1(Asn214*); 2.92 Å), and between the hydroxyl group of Tyr194 and the amide group of Asn214* (OH(Tyr194)-ND2(Asn214*); 2.87 Å) are observed (Figure 13D).

In MOL2 and α33 helix, a dimeric interaction is found involving four hydrogen bonds and a CH/π interaction, namely between the carboxylate group of Glu442 and the guanidyl group of Arg244* (NE(Glu442)-OE2(Arg 244*); 2.7 Å and OE1(Glu442)-NH2(Arg244*); 2.96 Å) , between the amide group of Asn446 and the hydroxyl group of Ser245* (ND2(Asn446)-OG(Ser245*): 3.10 Å) and the CH/π interaction (CE(Phe439)-CD1(Lys248*): 4.05 Å, CE1(Phe439)-CD(Lys248*); 3.98 Å and CZ(Phe439)-CG(Lys248*); 3.81 Å) . Additionally, residues Asp448 and Glu449 in α33 helix help

MOL2 to form hydrogen bonds with Arg554* and Tyr576* as follows: Asp448* (OD2(Asp448)-NE(Arg554*);2.84 Å and OD1(Asp448)-OH(Tyr576*); 2.7 Å) and Arg547* (OE1(Glu449)-NE(Arg547*); 2.75 Å and OE2(Glu449)-NH2(Arg547*); 2.91 Å) respectively (Figure 13d, 13E). The dimer interface which involves MOL1, MOL2 and α33 helix is not observed in mammalian PGIs.

In the C-terminal domain, a hydrophobic pocket is formed between α33 helix and the C-terminal domain (α37 helix and α38 helix). Moreover, although, Tyr576 in the C-terminal tail forms a hydrogen bond with Asp448*, Tyr576 interacts with Arg547 of the same subunit and contributes to stabilizing the two helixes (α37 and α38) in the C-terminal domain (Figure 13F).

Inhibitor molecule and inhibitor binding site

I solved three PfPGI crystal structures containing either inhibitors or substrates (6-phosphogluconic acid; 6PGA and 3-phosphoglyceric acid; 3PGA as inhibitors and fructose 6-phosphate; F6P as substrate. See Material and methods)

The substrate-binding sites are located between the large and small domains of one subunit and at the interface between the two subunits (see Figure 12A, 12B).

The electron density map for each of the 6PGA and F6P are very well-defined (data are not shown).

In the binding site, hydrogen bonds and electrostatic interactions between the negatively charged phosphate group and positively charged Arg or His side-chains were found (Figure 14).

With regard to the phosphate binding site, the phosphate group of 6PGA strongly interacts with the hydroxyl groups of Ser159 (OG(Ser159)-O2P):2.57 Å, Ser231 (OG(Ser231)-O1P): 2.52 Å, Thr233 (OG1(Thr233)-O3P): 2.75 Å, and Thr236 (OG1(Thr236)-O1P): 2.62 Å and the nitrogen atoms of Lys232 (N(Lys232)-O2P) and Thr233 (N(Thr233)-O3P): 2.92 Å) (Figure 14A). These interactions and the positions of the residues and the phosphate group, overlap with the reported mammalian PGI structures (Figure 14D). In addition to the PfPGI complex with 3PGA and F6P, these phosphate interactions and the position of side chain residues are also similar for the PfPGI-6PGA structure (Figure 14B, 14C).

With regard to the sugar binding site, numerous hydrogen bonds were observed.

In the case of the PfPGI-6PGA complex, the following hydrogen-bonds were found; (N(Arg294)-N(Gly293)-O1A(6PGA)), OE1(Glu380)-O1B(6PGA), (NE2(Gln376) - O(His411*)-O3(6PGA)), N(Gly158)-O4(6PGA), and (ND1(His411*)-O5(6PGA)) (Figure 14A).

Superimposing the structures of both PfPGI-6PGA and mPGI-6PGA revealed that the

individual position of residues and these hydrogen bonds overlapped with the mPGI-6PGA complex (Figure 14D). Importantly, it should be noted that the interaction between the amino group of Lys540 and O5 of 6PGA in PfPGI is different from that seen in mPGI (Figure 14D). In mPGI-6PGA, a direct hydrogen bond between the NZ of Lys519 and the O5 of 6PGA was reported in 6PGA. However, in the PfPGI-6PGA complex, a water molecules bridges the interaction between the NZ of Lys540 and O5 of 6PGA (Figure 14D).

In the case of the PfPGI-F6P structure, previous reports have suggested that the omit map of F6P is difficult to define due to PGI catalyzing the forward and reverse catalytic reaction between F6P and G6P, however my omit map was well defined. The high-resolution omit map at the substrate-binding site of the PfPGI-F6P complex clearly demonstrated the presence of a five membered sugar ring in both subunits in the ASU (Figure 14B). However, unfortunately, the open chain form of the sugar ring for F6P was not observed in this crystal.

The O1, O3, O4, and Oxygen atoms of ring form F6P formed the following hydrogen-bonds; (OE2 (Glu380)-O1(F6P); 2.48 Å, OE1(Glu380)-O3(F6P);2.93 Å, N(G159)-O3(F6P); 2.73 Å, and (ND1(H411*) and a water molecule). A water molecule formed a bridge interaction between the side chain of Lys540 and O4 of F6P (Figure 14B) (86).

In the case of 3PGA, although the electron density of the phosphate group was found, the electron density of the carbon portion of 3PGA could not be defined. The main reason is that carbon portion of 3PGA is too short to fill the binding site space and interact with the substrate binding residues. It follows that carbon portion of 3PGA might adopt a flexible conformation due its lower interaction with the binding pocket (Figure 14C), although, the 3PGA was well defined in mPGI.

A comparison of the PfPGI-F6P and mPGI-F6P structures revealed that the hydrogen bond between NZ(K519)-O1(F6P) in mPGI-F6P is not observed in PfPGI. This is likely because although O1 (F6P) in mPGI can adopt a double conformation, and only one conformation can interact with the NZ of Lys519 in mPGI, however, in PfPGI, a double conformation for O1 (F6P) was not found and the O1-NZ Lys540 hydrogen bond was not found (Figure 14E).

Conformational change in the inhibitor complex

In the case of the complex with the inhibitor, a conformational change in the mPGI binding site has been reported which involves three domains i) the phosphate binding loop (residues 210-216), ii) 3_{10}-helix (residues 386*-389* containing the His411* neighboring subunit), and iii) the carbohydrate moiety-binding helix (CBH; α37 helix;

residues 513-529) including Lys519 of the C-terminal domain (34). The conformation change of each domain has been classified as “Open/Closed” based on the sugar form bound. For an open-chain sugar such as 6PGA, all three domains are closed to the inhibitor molecule (“Closed”). In contrast, in the case of combination with ring-form sugars, such as F6P and G6P, although the phosphate binding loop and 3_{10}-helix come close to the inhibitor molecule (Closed), the CBH does not undergo a conformational change (Open).

Comparing PfPGI-6PGA and mPGI-6PGA, as I shown previously, both the phosphate binding loop and the 3_{10}-helix are close to the inhibitor molecule. These individual residues overlap closely with the mPGI-6PGI structure. However, a change in the conformation of the CBH in PfPGI does not occur (Open) and a water molecule and a water-mediated interaction between NZ of Lys540 and O5 of 6PGA is observed instead (Figure 15).

In the case of the PfPGI-ring sugar form (F6P) complex, the conformational change is also quite similar to the mPGI-F6P complex. This conformation is also the same seen in the PfPGI-6PGA complex.

These data importantly suggest that a conformational change in CBH does not occur, regardless of whether the sugar is in the open or ring form in PfPGI.

Discussion

I have determined the crystal structure of *Plasmodium falciparum* (*Pf*) PGI in complex with 6PGA at 1.7 Å, with 3PGA at 2.0 Å, and with F6P at 1.7 Å resolution. While the structures of the substrate-free PfPGI and PfPGI complexed with F6P have been solved and deposited in PDB (PDB code; 3QKI and 3PR3), this is the first report of PfPGI bound to the open sugar form. These crystal structures show that the key catalytic amino acids, which are reported to be Arg272 (Arg294 in PfPGI), Glu358 (Glu380 in PfPGI), His388 (His411 in PfPGI), Lys518 (Lys540 in PfPGI) for mammalian PGIs are well conserved in PfPGI, and they are superimposable (see Figure 14). However, an MOL domain and the unique dimeric interactions (CBH-α33 helix-MOL2-α15 helix* interaction) suggest that there may be differences in the catalytic mechanism between mouse and plasmodium PGIs.

The role of the MOL domain

I have detected two MOL domains (MOL1; residues 175-200, MOL2; residues 438-446) in the overall structure. As I have discussed above, MOL1 and MOL2 play important roles in maintaining dimerization in PfPGI. Although mammalian PGIs have only one hydrogen bond, a total of eleven hydrogen bonds are found in PfPGI (see Figure 13). This

presumably results in higher stability of the dimer conformation compared to mammalian PGIs.

MOL2 is found between the CBH including C-terminal domain and the Small domain of the neighboring subunit and cause a steric hindrance to prevent the moving of α33 helix. The MOL2 is located outside part of the structure as loop structure. In general, a loop structure is flexible, but the MOL2 is more rigid due to interactions between CBH and Small domain and MOL2 fixes the domains. Thus, MOL2 is assumed to play an important role as a "Frame" to retain space in the binding pocket. In addition, the α33 helix was found to form tight hydrogen bonds with the C-terminal domain (α37 (CBH) helix and α38 helix). It is likely that these interactions strongly affect the C-terminal domain.

Interestingly, the CBH-α33 helix-MOL2-Small domain interaction is only found in plasmodium PGI. In particular, residues (Phe435, Glu442, Asn446, Asp448, Arg547, Arg554, and Tyr576) which are important in the CBH-α33 helix-MOL2-Small domain interaction, are conserved completely in plasmodium PGIs, while in mammalian PGIs, the interaction motif is very simple, consisting of only a hydrogen bond between Glu526 in the α37 helix and Gln216 in the α15 helix and a "Frame" is not required (Figure 17). Notably, residues Gln216 and Glu526 in mPGI correspond to Ala237 and Arg547,

respectively, suggesting that the binding site retention mechanisms are very different, even though the catalytic residues are the same between mammalian and plasmodium PGIs (Figure 16A, 16B).

Conformational change in the PGI substrate binding site

In the structure of the PfPGI-6PGA complex, both the phosphate binding loop and the 3_{10} helix are in similar "Closed" form as in mPGI; in contrast, the CBH is not in a"Closed" form. In order to clarify the conformational change, I performed an overlap structure analysis of the substrate binding site between PfPGI-6PGA complex and the substrate-free form of PfPGI, which had already been resolved at 1.9 Å and deposited in the PDB. (PDB code; 3QKI) (Figure 18). It showed that the phosphate binding loop shifts to the inhibitor around 1.5 Å and a change in the 3_{10}-helix conformational occurred with the imidazole group of His411* rotating and shifting the ND1 of the imidazole group close to O5 of 6PGA (3.98 Å to 2.69 Å) to become more suitable for substrate binding in the active site. These conformational changes are considered to cause steric hindrance between Thr236 and His411* .This is because the Cα of Thr236 is shifted 1.5 Å to tightly contact the phosphate group of 6PGA. As a result, His411* seems to move toward the hydroxyl group of 6PGA to avoid steric hindrance.

These conformational changes are similar to conformational changes previously

reported in mammalian PGIs (28-34). However, the CBH in the PfPGI complex with 6PGA does not shift toward the 6PGA inhibitor (i.e. there is no conformational change).

With respect to the conformational change for the CBH, there is no uniform change, and several reports of conformational change in PGI with substrates in the open chain form (F6P and G6P) and open chain inhibitors (6PGA, E4P, sorbitol 6-phospate (S6P), 5-phospho-darabinonate (5PAA), and 5-phospho-arabinonohydroxamate (5PAH)) have been published (28-30, 32-34, 74, 83-86).

Table 5 shows a summary of the conformational changes in the three domains in previously reported PGI-substrate/inhibitor complexes. Both the phosphate binding loop (residue 230-236 in PfPGI) and the 3_{10} helix are "Closed" in complexes with all inhibitors and the open chain form of substrates. Interestingly, in contrast, the changes in the conformation of CBH have remained unclear. In fact, in PGIs complexed with almost all inhibitors it shifted from "Open" to "Closed", however, in PGIs complexed with open chain form of substrates it did not shift from "Open" to "Closed". As a specific example, in the complex with S6P, Lee and Jeffery et al., have reported it as being "Open" in the rPGI-S6P complex, whereas Tanaka et al. have reported it as being "Closed" in the mPGI-S6P complex (35).

For this question regarding substrate-structure (open or closed forms) dependent

conformational change of CBH, based on X-ray crystal structures, several proposals have been made about the catalytic mechanism in substrate binding site

From the structure of the rabbit PGI F6P complex, Glu216, His388, Lys518, and a water molecule were identified as having roles in the ring opening step, and the structure of rabbit PGI complexed with 5PAA suggested that Glu357 and a water molecule act in the proton transfer mechanism, whereby a proton is moved from C1 to C2 in the *cis*-enediol(ate) intermediate step (34, 83-86). Arg272 has also been proposed to have a role in helping to stabilize the *cis*-enediol(ate) intermediate (84).

In addition, detailed catalytic mechanisms have also been suggested from high resolution structures containing the open chain/ring forms of G6P. The phosphoglucose isomerase catalytic reaction is thought to proceed in seven steps; i) ring opening by His388 and Lys518, ii) rotation about C3-C4, iii) substrate stabilization by Lys518, iv) proton abstraction to the form *cis*-enediol(ate) intermediate by Glu358 and stabilization of the intermediate by Arg272, v) proton re-donation to Glu358, vi) rotation about C3-4, and vii) ring closure by His388 and Lys518 (28, 29, 34, 83-86).

With regard to the Lys518 (corresponding to Lys540 in PfPGI) conformational change, Lee and Jeffery et al. have concluded that Lys518 has an important role in the ring opening/closure step, and in the substrate stabilization (intermediate) step before

protonation by Glu358 and formation of the *cis*-enediol(ate) intermediate (33, 83). The Lys residue is involved in the water-mediated interaction with the ring oxygen atom of the ring form of F6P to deprotonate the hydroxyl group on C2 of F6P in the ring opening step. Thus, in this state, the CBH conformation is "Open". In contrast, in the substrate stabilization step, after C3-C4 rotation of the open chain F6P, the water molecule which contributes the water-mediated interaction between Lys and substrate, is lost, and the Lys residue shifts and comes into direct contact with the C5 hydroxyl group and the oxygen atom of the phosphate group. In this state, the CBH conformation is "Closed".

Because open chain inhibitors such as 5PAH, A5P, S6P, and 6PGA (especially 5PAH and A5P), are similar to the *cis*-enediol(ate) intermediate this suggest that the status of shifted Lys residue appears to be an intermediate step in this isomerization reaction that is involved in both glycolysis and gluconeogenesis. Thus, based on the reported crystal structure, the conformational change of CBH involving Lys518 might occur in intermediate step in which Glu358 protonates the substrate (28, 34, 83, 86-88).

From my study, the data showing there is no CBH shift in PfPGI-6PGA may be a snapshot of the intermediate step based on the proposal of Lee and Jeffery et al. however, the further concern is needed to be conclusion (33, 83). There is a steric hindrance caused by the additional MOL domain. As described previously, MOL2 is found in-between

CBH and α15 helix* and creates a steric hindrance to inhibit the movement of the CBH. In addition, two hydrogen bonds between CBH and α33 helix, which interact for subunit dimerization, strongly restrain the movement of the CBH, and also inhibit the shift to substrate. In mammalian PGIs, this steric hindrance does not occur and the Gln216-Gln526 hydrogen bond maintain the substrate binding space. Interestingly, while two interactions occur between Gln216-Gln526 (NE2 (Gln216)-OE1(Gln526); 2.93 Å and OE1(Gln216)-OE2(Gln526); 3.37 Å) in the case of substrate/inhibitor free mPGI, only one interaction occurs (NE2(Gln216)-OE2(Glu526); 2.84 Å) in the case of mPGI complexed with 6PGA. This is because the position of Gln216 is close to the phosphate binding loop and shifts to become close to the substrate binding pocket (approximately 1.0 Å, Cα of Gln216) after substrate/inhibitor binding. Therefore, the space retention motif breaks down and Lys519 can easily contact 6PGA directly. These differences suggest that the binding site in PfPGI is more rigid than those of mammalian PGIs, and that Lys540 may not contribute to stabilization at the substrate intermediate step due to the fact that it cannot form a direct contact with the substrate. With regard to substrate stabilization, the C1-C2 interaction pocket, which consists of residues Gln533, Gly293, and Arg294 is sufficient to stabilize the open chain form of the substrate and to protonate C1-C2 by Glu380, even though Lys540 does not shift and stabilize the substrate by

interaction with the oxygen atoms of C5 and C6.

This structural differences would affects the enzyme catalytic activity and indicate that the catalytic activities of mammalian PGIs might be greater than those of plasmodium PGIs.

In fact, Table 6 shows the enzyme catalytic activity of PGIs where the K_m values for mammalian PGIs are lower than for PfPGI (K_m = 0.037–0.119 mM for mammalian PGI, K_m = 0.26 mM for PfPGI) for the forward reaction (i.e. F6P converting to G6P).

Moreover, in PGI from *Trypanosoma brucei* (tPGI), a K_m value of 0.122 mM has been reported, and the *Ki* value of 5PAH revealed that the preference for tPGI was a four-fold that for rPGI (*Ki* = 0.05 μM for tPGI, *Ki* = 0.20 μM for rPGI) (89-95). In the tPGI crystal structure, the CBH does not form hydrogen bonds, but hydrophobic interactions are used instead. Hence, CBH is able to stabilize the substrate to allow for efficient catalysis of the substrate.

In the PGI from *Pyrobaculum aerophilum* (PaPGI), interestingly, the K_m value is different at different temperatures. (K_m = 0.06 μM at 80°C, and K_m = 0.3 μM at 50°C). In crystal structure of PaPGI, the CBH domain is stabilized by one hydrogen bond between Glu250*-Arg300. These data show that CBH become more flexible and loses hydrogen bonds at high temperatures (94).

Thus, these results support the notion that CBH flexibility contributes to substrate stabilization for protonation and can affect the catalytic activity.

Conclusion

I have determined the crystal structure of *Plasmodium falciparum* PGI complexed with two inhibitors (3PGA and 6PGA) at 2.0 Å and 1.7 Å and a substrate (F6P) at 1.7 Å, respectively. The overall of crystal structure shows two additional insertion loops (MOLs) and these MOLs are only conserved in the plasmodium species. Moreover, these MOLs contribute significantly to the dimeric interaction and MOL2 (residues 438-446) creates a steric hindrance by inserting itself between CBH and α15 helix*. In the substrate binding site, the position of the phosphate binding loop domain and the 3_{10}-helix containing His411* in PfPGI complexed with 6PGA are quite similar as in the crystal structure of mouse PGI with 6PGA, however, the position of the CBH containing Lys540 is more extended than in mPGI. Although, the CBH conformational change should allow the CBH to closely and directly interact with the substrate to help position the substrate for the protonation step, CBH cannot get close to the substrate due to steric hindrance. Although further studies are required to form a firm conclusion, my crystallographic data and review of the data may allow me to propose a new catalytic mechanism which is somewhat different from those proposed in previous reports.

General Discussion

This thesis reports on the mechanism of peroxisomal localization of tetrameric carbonyl reductase and on the substrate binding recognition mechanism in phosphoglucose isomerase from *Plasmodium falciparum* using X-ray crystallography. From the results described in Chapter 1, I found that the peroxisomal targeting signal is not only an import signal for the peroxisome, but also has a role in inter-subunit interactions, and plays a role in enzyme activity.

Although, PerCR forms a homo-tetrameric molecule consisting of four monomeric PerCRs, a recent crystal structure showing a hetero-tetrameric SDR has been reported (96). This enzyme, which is called 3-ketoacyl-acyl carrier protein reductase (KAR), catalyzes the second step in the mitochondrial fatty acid synthesis (mtFAS) pathway. Although all of the other mtFAS enzymes identified thus far are encoded by a single gene, human KAR (HsKAR) is a hetero-tetrameric α2β2 enzyme formed by two subunits, namely, 17β-hydroxysteroid dehydrogenase type 8 (HSD17β8) and carbonyl reductase type 4 (CBR4) (97). This finding suggest a new question that can only be addressed currently by X-ray crystallography. The question is whether PerCR can form hetero-tetramers with other enzyme SDRs.

From the crystal structure of HsKAR, both dimers of HSD17β8 and CBR4, each of which is related by the Q-axis (please see Figure 5) comprise the hetero-tetramer. When the possible dimeric form of PerCR is considered, the Q-axis related dimer is a candidate form. This is because the Q-axis related inter-subunit interaction is stronger than the interaction between the R-axis related subunits. In this state, the dimeric form of PerCR related to the Q-axis, requires complementary amino acid residues of the P-axis related subunit for peroxisomal targeting signal (PTS) of PerCR (Ser-Arg-Leu). In my research, Ser258 (involved in a hydrogen-bond network between Lys176 and Glu248), Arg259 (forming a hydrogen bond with Glu248) and Leu260 (forming a hydrogen bond between Lys172 and hydrophilic interaction with R-axis related hydrophobic residues) were identified and have been used to seek for candidate enzymes with complementary amino acids. Although I carried out an amino acid sequence alignment using the UniProt database (URL; http://www.uniprot.org/) using the following key words “Peroxisome”, “SDRs or reductase or oxidase” and “pig or human”, unfortunately, I have not yet been able to identify any candidate partner enzyme, so further research is need to resolve this question.

In Chapter II, with regard to the substrate binding mechanism of PfPGI, I found a hint of the reason why the CBH conformational change occurs in this important reaction that is involved in both glycolysis and glucogenogenesis. This finding could be of relevance to the design of selective anti-malarial drugs. With respect to *Trypanosoma brucei* PGI, several selective anti-malarial PGI candidate compounds have been reported (suramin, and agaricic acid) (Figure 19A, 19B) (82, 98, 99). However, suramin cannot fit into the substrate binding site due to fact that it is a large molecule and it is therefore unclear how it inhibits tPGI. As for agaricic acid inhibition, its mechanism may be explained based on the CBH interaction motif (Figure 19E). The CBH interaction motif is very different among species, as I discussed in Chapter II. In particular, the tPGI CBH – α helix which maintain the binding site space, forms only hydrophobic interactions, whereas the mammalian PGI CBH – α helix forms hydrogen bond between Glu526-Gln216* (Figure 19D). This difference suggests two important points, first that tPGI CBH is more flexible and second that the space between CBH and its α helix is more extensive. This may explain why the long carbon chain of agaricic acid can fit into this space.

In my research, the MOL2 is identified as being a cause of steric hindrance and creates an extensive space between CBH and α15 helix* with the same volume as in tPGI (Figure 19F).

Currently, I am focusing on this space and I am studying the compound salicin which is an alcoholic β-glucoside, comprised of glucose and salicyl alcohol, as a new selective anti-malarial drug. Salicin is the historical origin of aspirin, and is chemically related to it, with a similar mechanism of action (100) (Figure 19C).

If salicin, or its derivatives, are capable of inhibiting PfPGI, salicin could become a candidate for a selective anti-malarial drug. In particular, this compound is very well known, is safe in humans, and is inexpensive to produce. Since malarial disease most often occurs in developing country, an inexpensive and new selective anti-malaria drug could make a significant contribution to the treatment of malarial disease.

In summary, I have solved the crystal structures of PerCR and PfPGI, and as a result I identified the mechanism for peroxisomal localization and the mechanism of substrate recognition for these two enzymes, respectively.

These findings could only be made and explained through the use of X-ray crystallography and I am happy to be able to make a contribution to biological science.

Tables and Figures

(a)

(b)

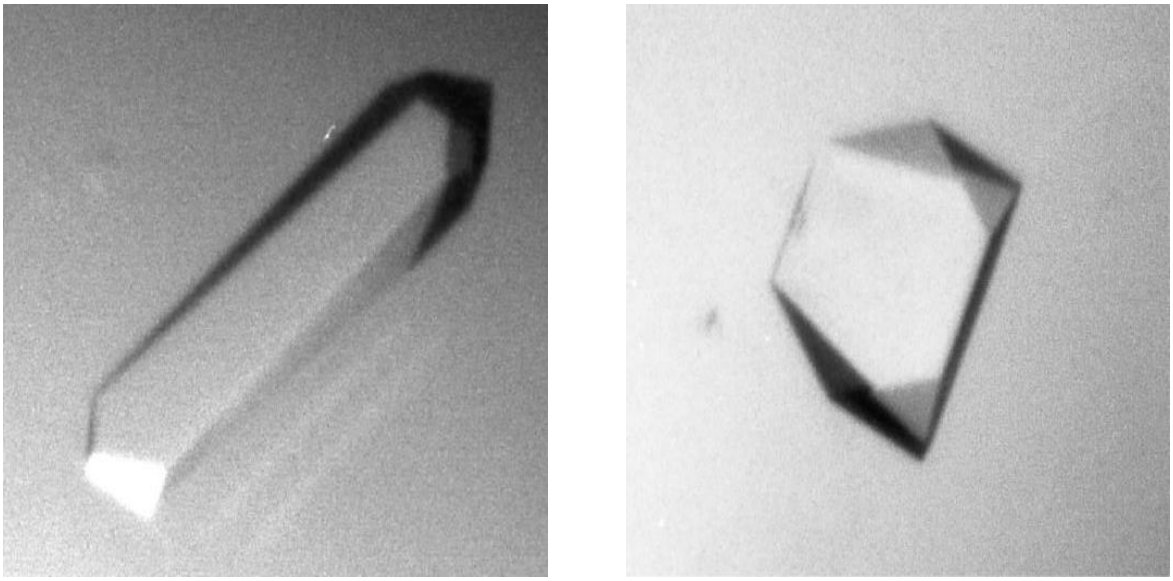


Figure 1
Tetragonal crystals of PerCR. (a) Form I. (b) Form II.

Table 1.
Data-collection statistics for the forms I and II crystals of PerCR.

	Form I	Form II
Space group	$P4_2$	$P4_12_12$
Unit-cell parameters (Å)	a = 109.6, b = 109.6, c = 94.3	a = 120.1, b = 120.1, c = 147.0
No. of subunits per asymmetric unit	4 [one tetramer]	4 [one tetramer]
Solvent content (%)	48	44
X-ray source	PF- AR NW12	PF- AR NW12
Detector	ADSC Q210	ADSC Q210
Wavelength (Å)	1.000	1.000
Resolution (Å)	1.5 (1.58-1.5)*	2.2 (2.32-2.2)
No. of unique reflections	174385	54751
Multiplicity	4.9 (4.9)	8.3 (6.8)
Mean I/σ	6.5 (2.3)	5.4 (3.3)
B factor (Wilson plot) ($Å^2$)	13.4	34.8
R_{merge}† (%)	8.2 (31.0)	8.1 (19.4)
Completeness (%)	98.0 (98.4)	99.3 (96.4)

*values in parentheses are for the outer shell.

Table 2.

Refinement statistics for the form II crystal of PerCR.

Resolution range (Å)	40-1.5
No. of reflections	
Working set	167,340
Test set	8,816
R-factor	0.161
Free *R*-factor	0.188
No. of protein atoms[a]	7,516 (1,879 × 4)
No. of NADPH atoms	192 (48 × 4)
No. of water molecules	772
RMS deviations	
Bond distances (Å)	0.014
Bond angles (°)	1.541
Ramachandran plot	
Most favored (%)	90.4
Additional allowed (%)	8.8
Generously allowed (%)	0.8

[a] N-terminal residues (Met1 to Arg9) are not included in each of the four subunits in ASU.

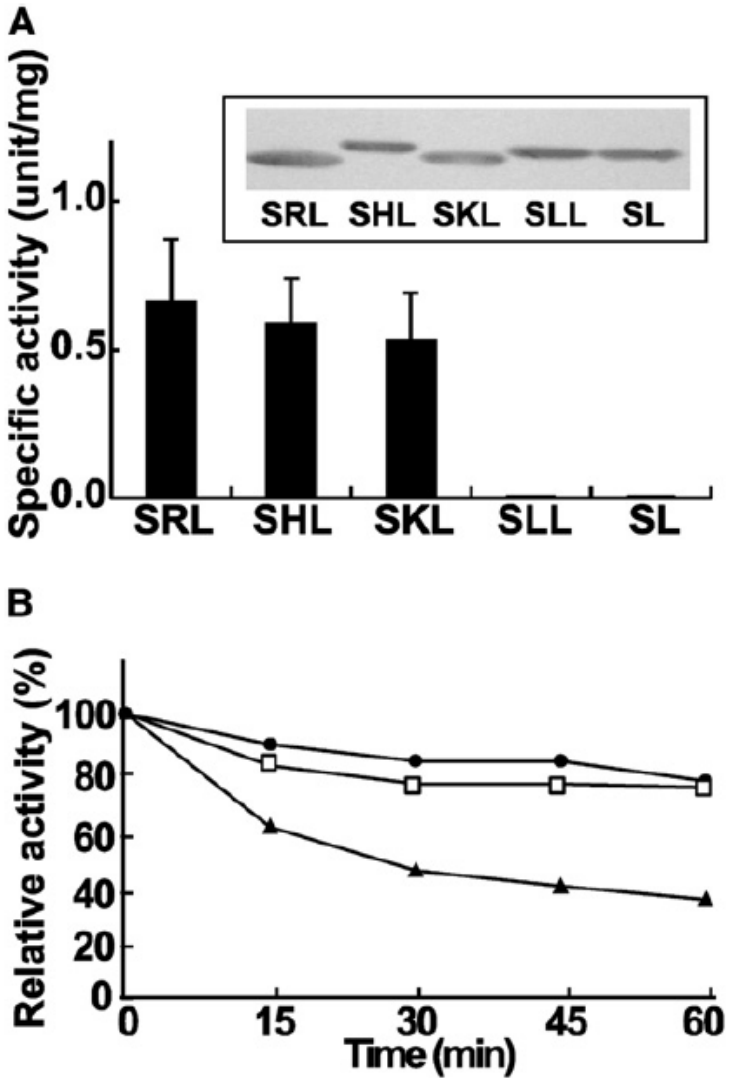


Figure 2. Effects of mutation of the C-terminus on enzyme activity and stability*.
(A) Activity analysis of *E. coli* extracts expressing wild-type pig PerCR (SRL) and the C-terminal mutants (SHL, SKL, SLL, and SL). Standard deviation obtained from three independent experiments. No enzymatic activity was detected in extracts expressing the C-terminal SLL and SL mutants. The inset depicts a representative western blot analysis using the anti-pig PerCR antibody. The western blot was used to detect the expression levels of the different PerCR enzymes and shows only the monomeric enzyme size.
(B) Stability of the purified enzymes. The wild-type pig PerCR (●), and the C-terminal mutants SHL (□) and SKL (▲) were diluted to 2 mg/mL with 10 mM potassium phosphate (pH 7) and incubated in an ice bath for the indicated times. Reductase activity was then measured and is expressed as the activity relative
to that at 0 min.
*Dr. Makoto Nagano & Prof. Akira Hara (Gifu Pharmaceutical University)

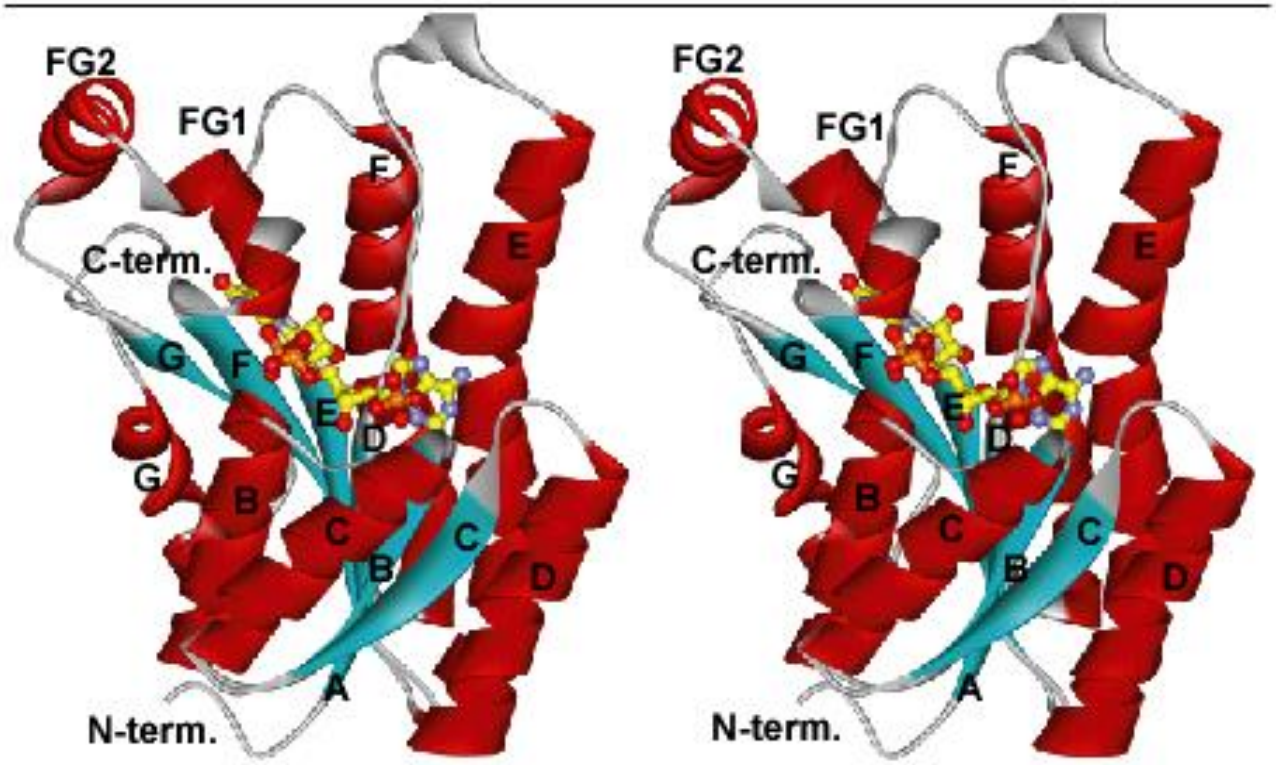


Figure 3. Stereo view of the PerCR Subunit.
The ball-and-stick model shows the bound NADPH molecule.
The α helices (red), β strands (cyan), and the N and C termini are marked.

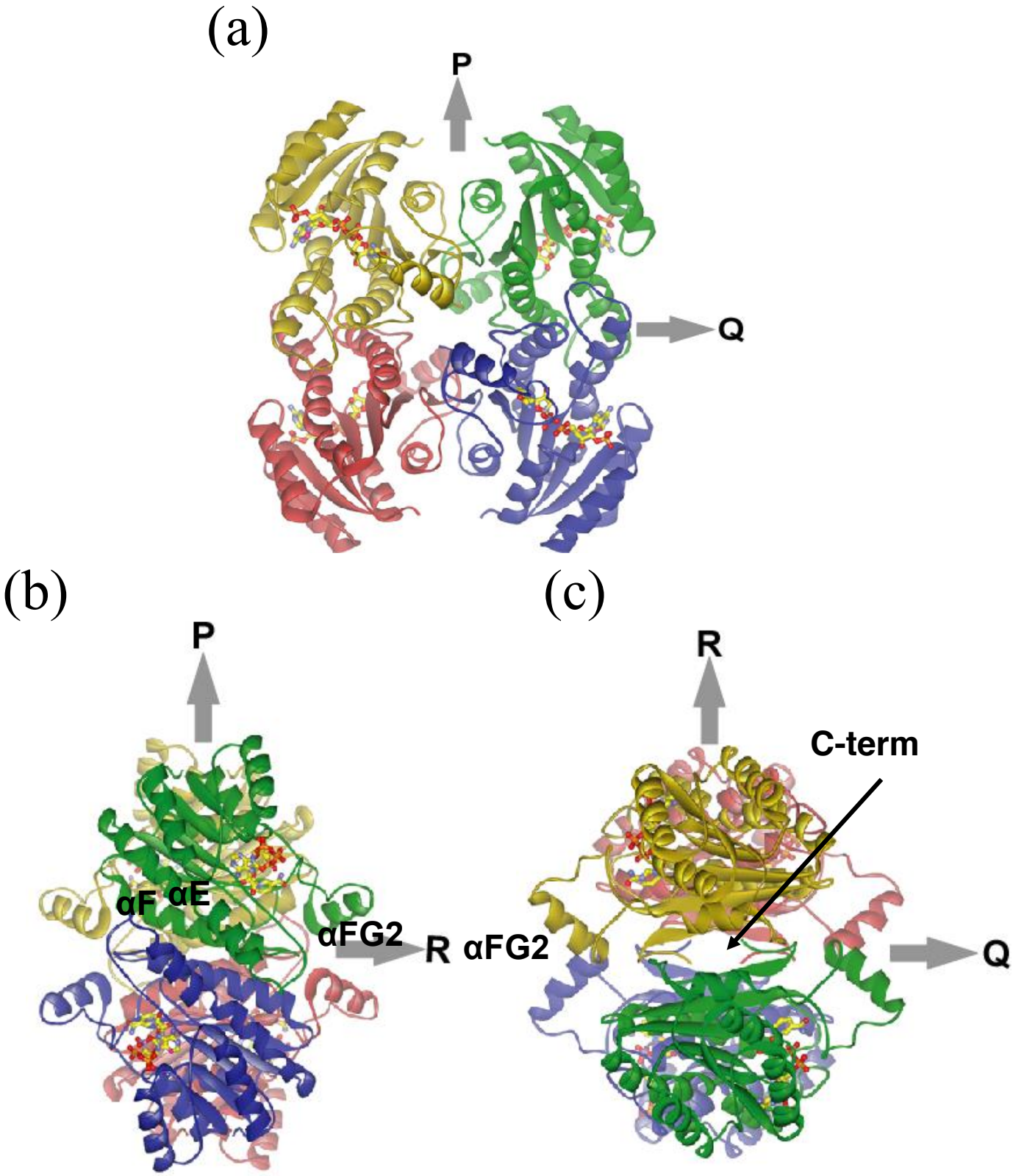


Figure 4. Ribbon representations of the PerCR tetramer viewed.
along each of the three non-crystallographic two-fold axes
The subunits A, B, C, and D are shown in yellow, blue, red, and green, respectively.
Bound NADPH molecules are shown as space-filling models (cyan).
(A) View along the R-axis, (B) View along the Q-axis, (C) View along the P-axis.
Please define what αF, αE, and αFG2 refer to.

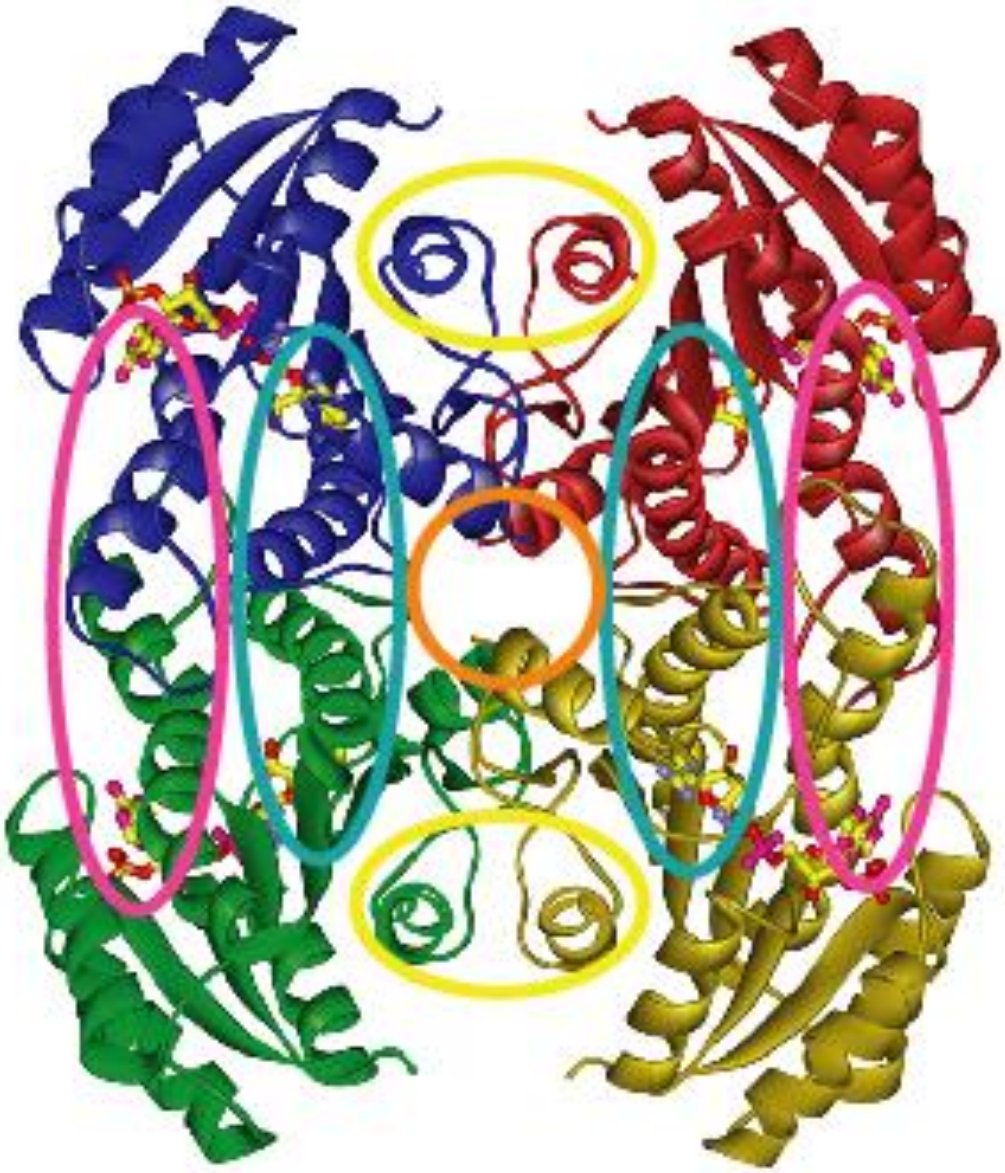


Figure 5. Overview of the inter-subunit interactions.
The inter-subunit interactions are αE-αE (pink), αF-αF(cyan), αG-αG (yellow), and C-terminal (orange).

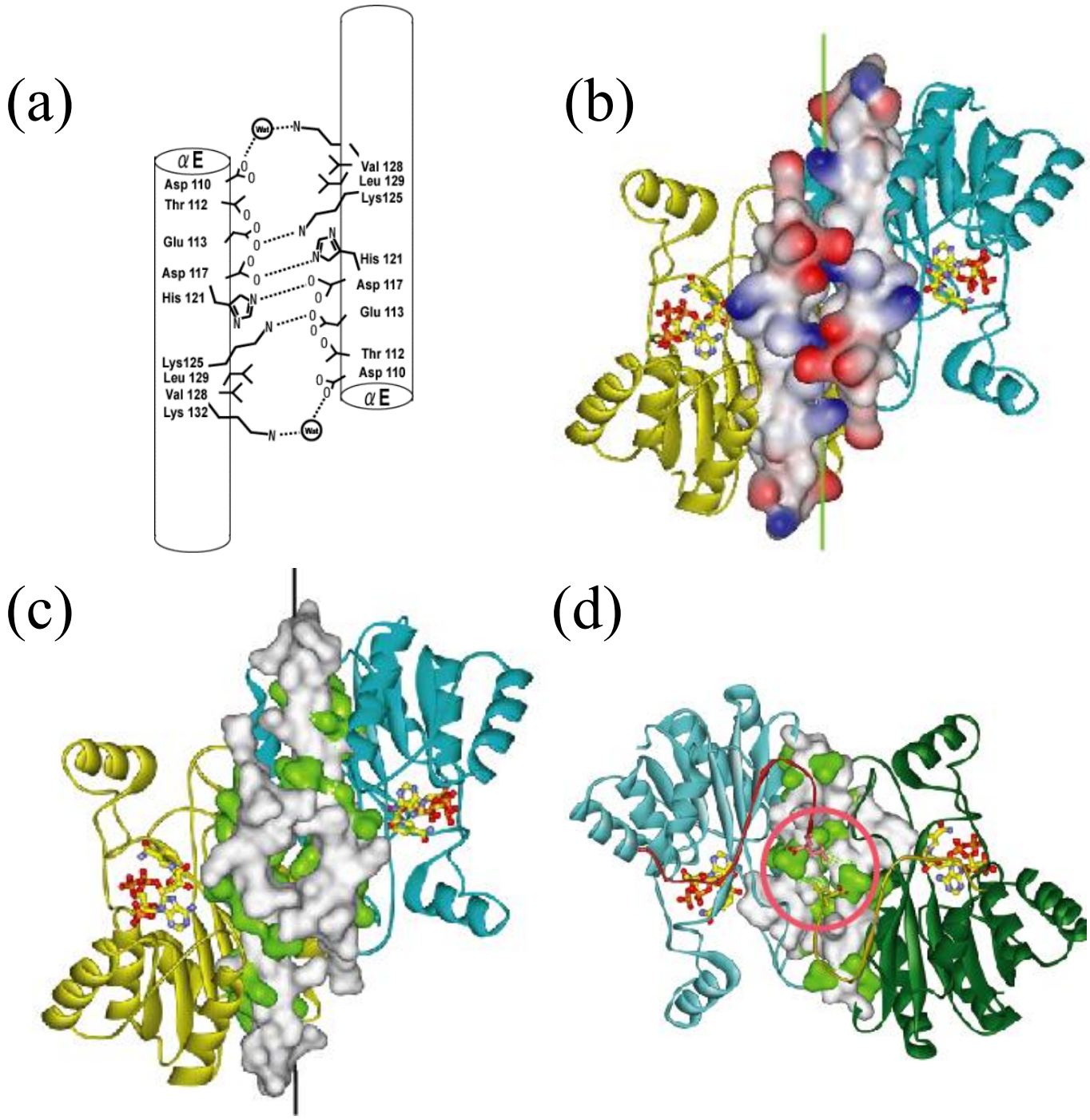


Figure 6. Inter-subunit interaction between the Q-axis related subunits and the R-axis related subunits.
(a) Pattern diagram of the αE-αE interaction: the electrostatic interactions are shown as a dashed line. (b) Surface model of the αE-αE interaction: anionic and cationic charged residues are shown in blue and red, respectively. (c) Surface model of the αF-αF interaction; hydrophilic residues are shown in light green. (d) Surface model of the αG-αG interaction: hydrophilic residues are shown in light green. The C-terminal tail interaction and the hydrophilic interactions are shown as a pink circle.

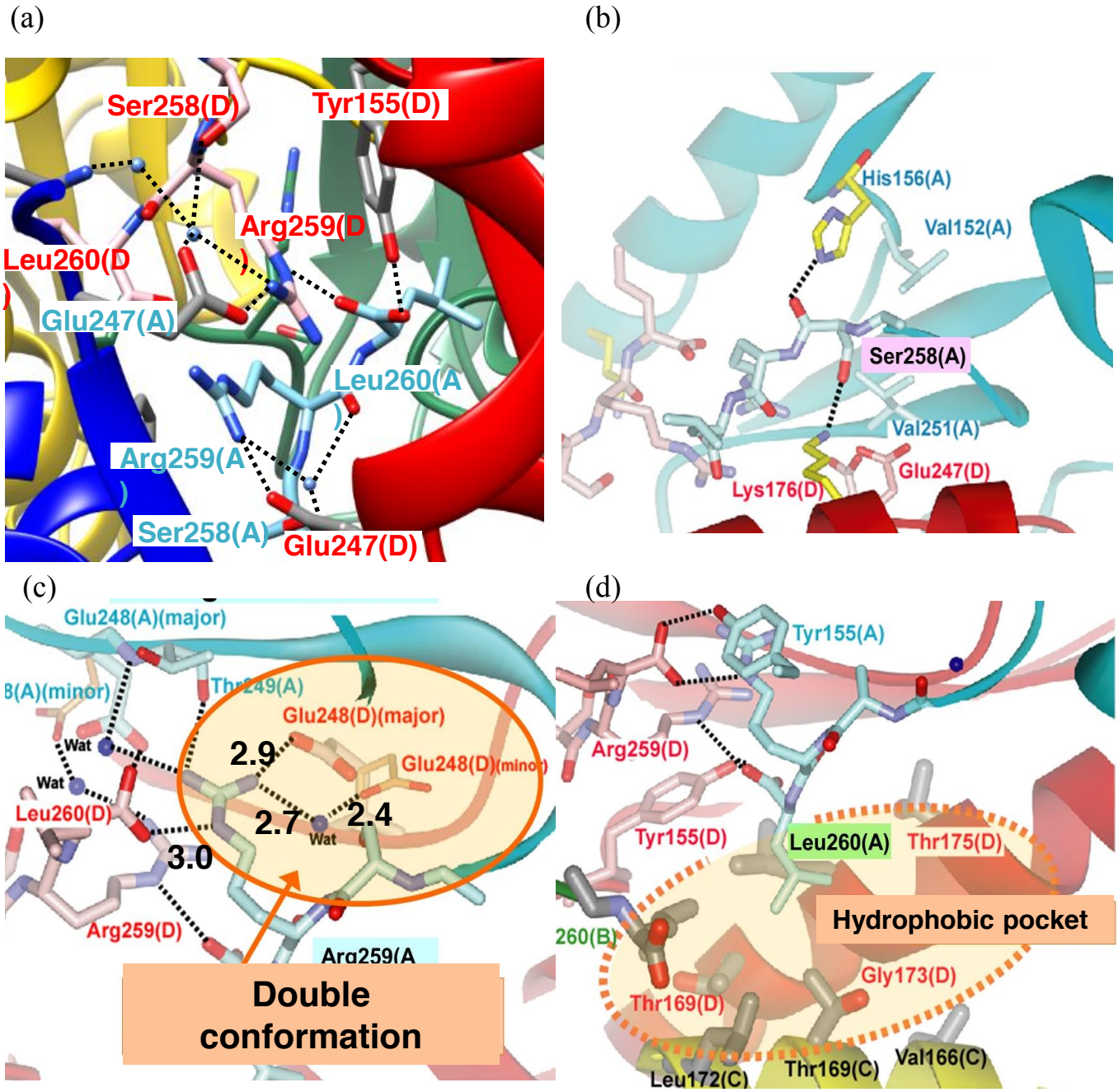


Figure 7. Intramolecular and inter-subunit PTS1 recognition mode observed in PerCR. Stick models of the PTS1 (Ser258-Arg-259-Leu260) belonging to subunit A and the surrounding amino acids. Water molecules are shown as spheres (blue). The dashed lines indicate possible hydrogen bonds (or salt bridges); all distances are given in Å.

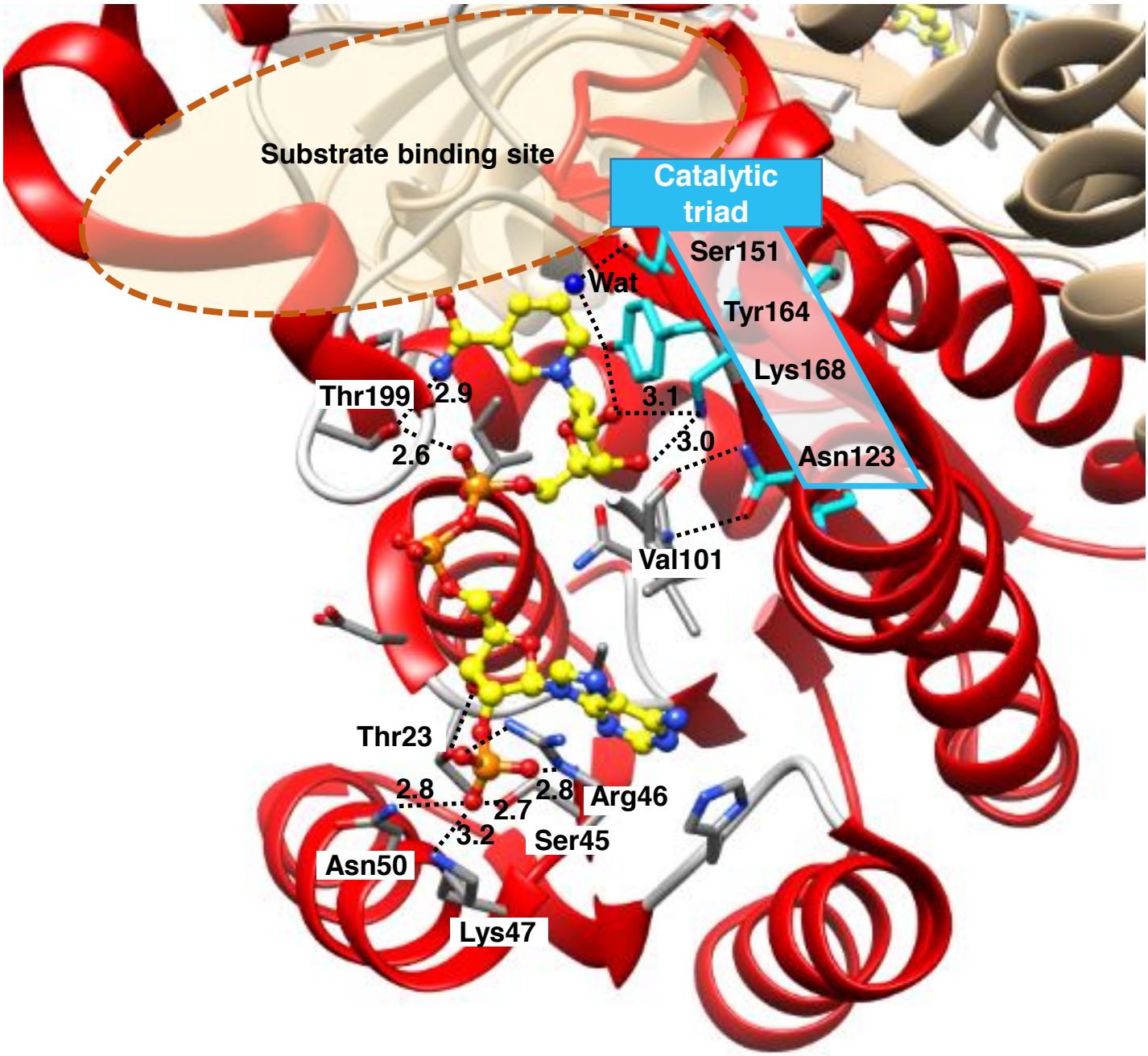


Figure 8. Active site of PerCR complexed with NADPH.
Possible hydrogen bonds (or salt bridges) are shown as dashed lines. The bound NADPH molecule and the surrounding amino acids are shown as ball-and-stick, and stick models, respectively. A water molecule bound to the side chains of Ser151 and Tyr164 is shown as a sphere (blue). The catalytic triad are shown as diamonds (Cyan) and the substrate binding site is highlighted as a circle (light yellow).

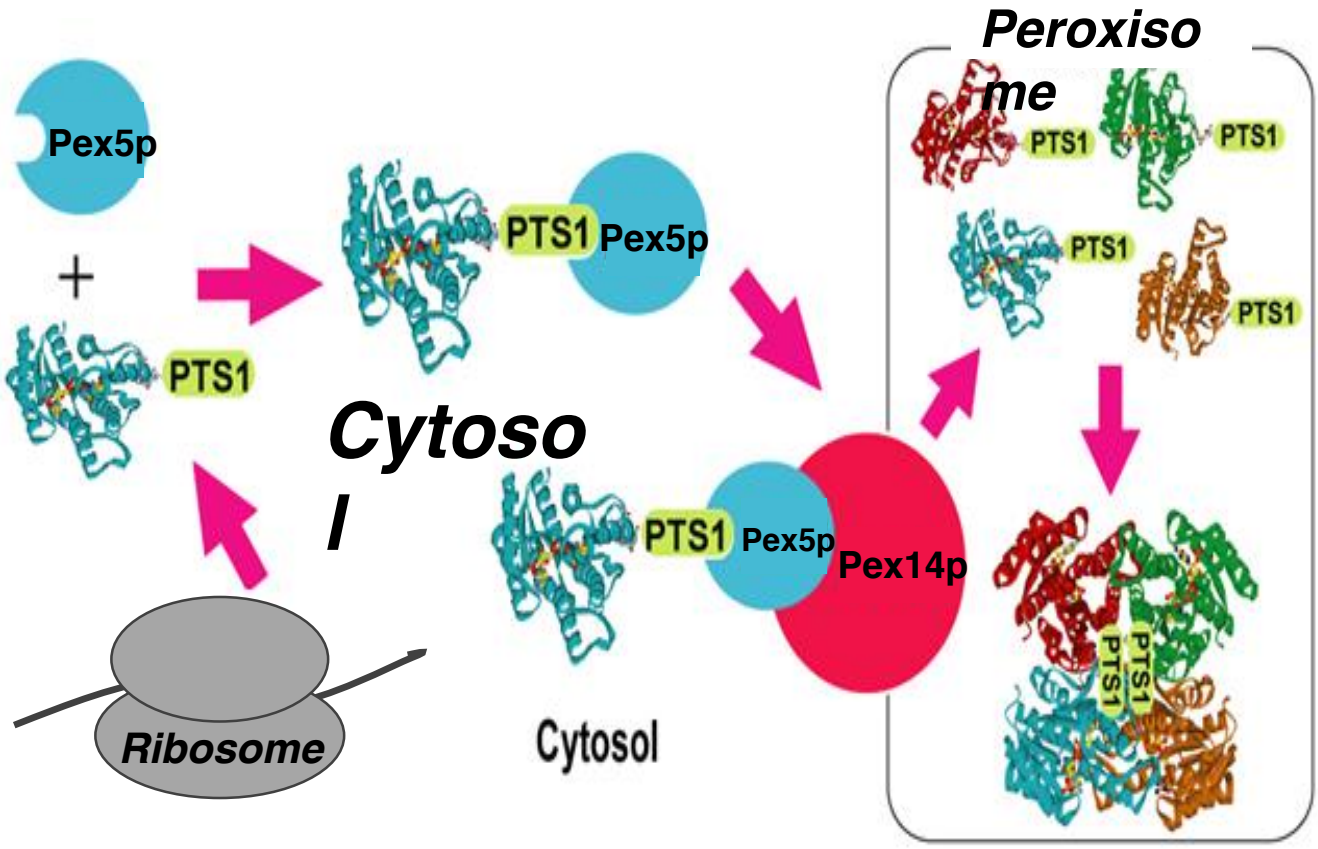


Figure 9. The most likely scenario for the transport of PerCR into the peroxisome.
1) Translated monomeric PerCR is recognized by Pex5p in the cytosol through PTS1.
2) The monomeric PerCT is the imported into the peroxisome by the Pex5p-Pex14p complex.
3) The monomeric PerCR form adopts a tetrameric structure inside the peroxisome.

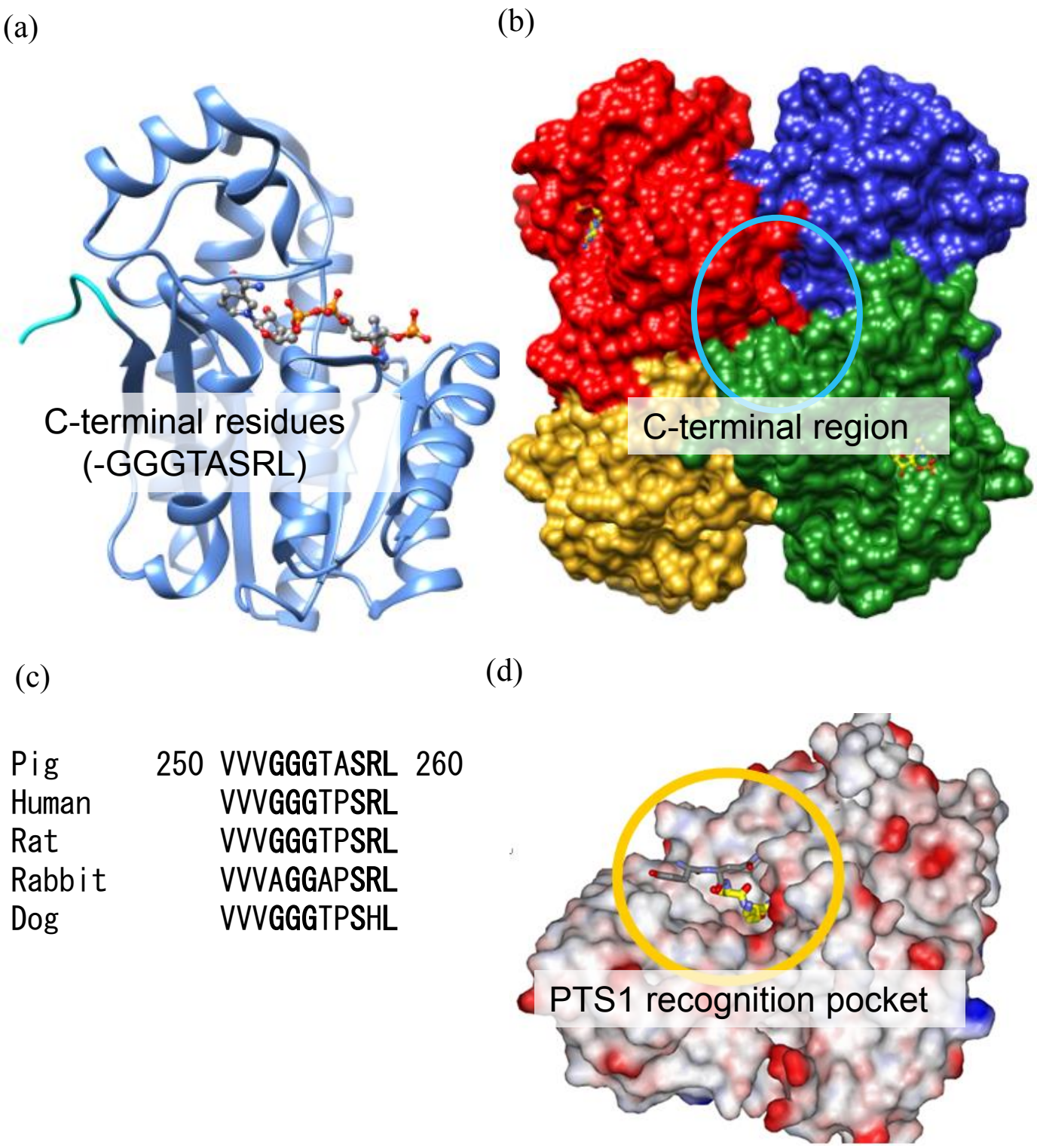
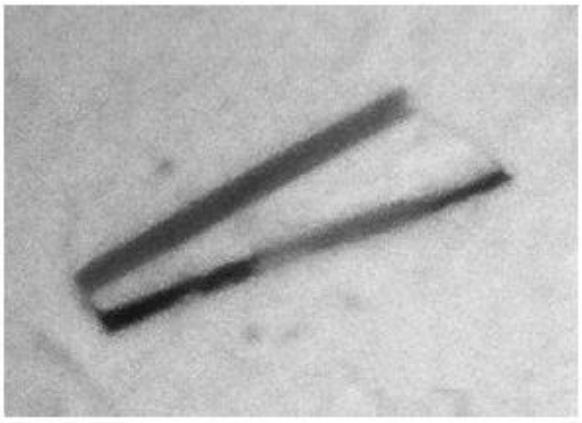


Figure 10. C-terminal residues of PerCR and insights into recognition.
(a) Monomeric PerCR; the C-terminal residues are shown in cyan. (b) Surface model of tetrameric PerCR. The blue circle shows the C-terminal region. (c) C-terminal amino acid alignments for different mammalian PerCRs. (d) Surface model of the Pex5p-PTS1 peptide complex model (PDBID; 1FCH, Gatto et al., 2000). The orange circle shows the recognition pocket which interact with PTS1

(a)



(b)

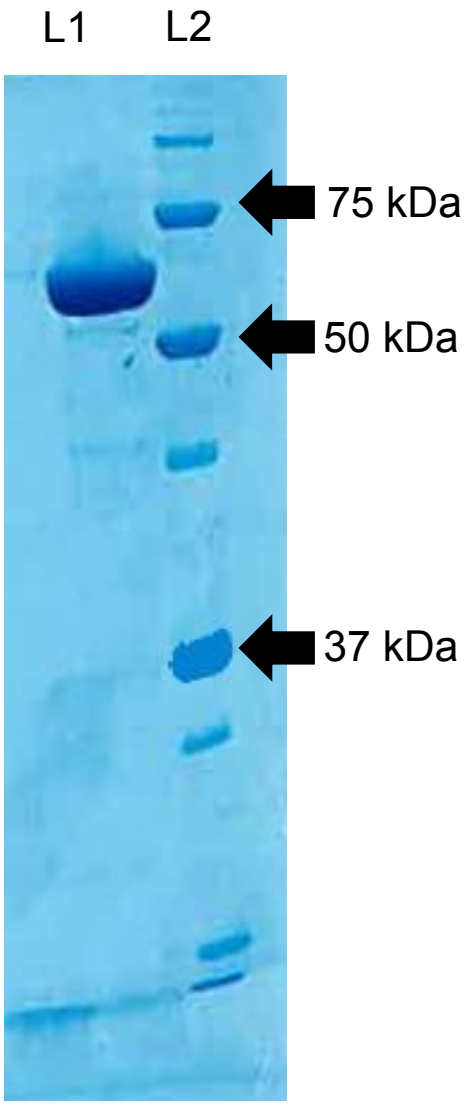


Figure 11. Result of purification and crystallization.
(a) An orthorhombic crystal of PfPGI (b) 10% SDS-PAGE of purified PfPGI. Purified PfPGI is shown in L1 after concentration. L2 contains the molecular mass markers.

Table 3
Data-collection statistics for *Plasmodium falciparum* Phosphoglucose isomerase (PfPGI) values in parentheses are for the outer shell.

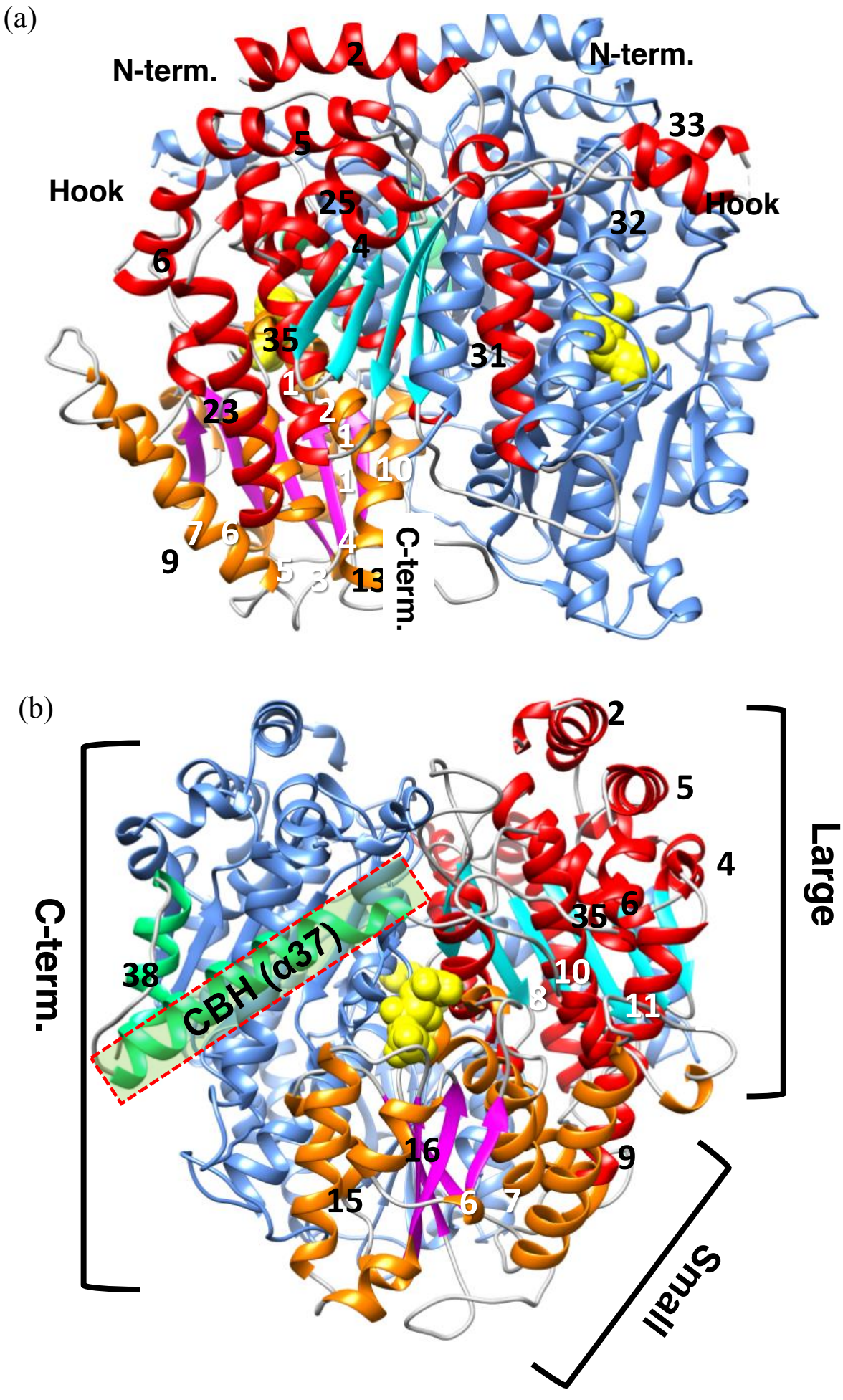
	PfPGI-6PGA
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 103.3, b = 104.1, c = 114.6
No. of subunits per asymmetric unit	2 [one dimer]
Solvent content (%)	46.3
X-ray source	PF BL5A
Detector	ADSC Q315
Wavelength (Å)	1.000
Resolution range (Å)	50–1.5 (1.58–1.50)
No. of observed reflections	933398
No. of unique reflections	195801
Multiplicity	4.8 (4.7)
Mean I/σ	7.6 (2.9)
B factor (Wilson plot) (Å^2)	16.4
R_{merge}† (%)	5.7 (26.2)
Completeness (%)	99.3 (96.5)

† $R_{merge} = \sum_{hkl} \sum_i | I_i(hkl) - \langle I(hkl) \rangle | / \sum_{hkl} \sum_i I(hkl)$, where $I_i(hkl)$ is the ith measurement and $\langle I(hkl) \rangle$ is the weighted mean of all measurements of $I(hkl)$.

Table 4.
Refinement statistics for PfPGI.
Values in parentheses are for the outer shell.

	PfPGI-6PGA
Resolution range (Å)	40-1.5
No. of reflections	
Working set	167,340
Test set	8,816
R-factor	0.161
Free R-factor	0.188
No. of protein atoms[a]	7,516 (1,879 × 4)
No. of NADPH atoms	192 (48 × 4)
No. of water molecules	772
RMS deviations	
Bond distances (Å)	0.014
Bond angles (°)	1.541
Ramachandran plot	
Most favored (%)	90.4
Additional allowed (%)	8.8
Generously allowed (%)	0.8

[a] N-terminal residues (Met1 to Arg9) are not included in each of the four subunits in ASU.


(a)
N-term.
N-term.
2
5
33
Hook
Hook
25
32
4
6
35
31
1
23
2
1
1
10
7
6
4
9
5
3
13
C-term.
(b)
2
5
4
Large
C-term.
35
6
38
CBH (α37)
10
8
11
16
9
15
6
7
Small

(c)

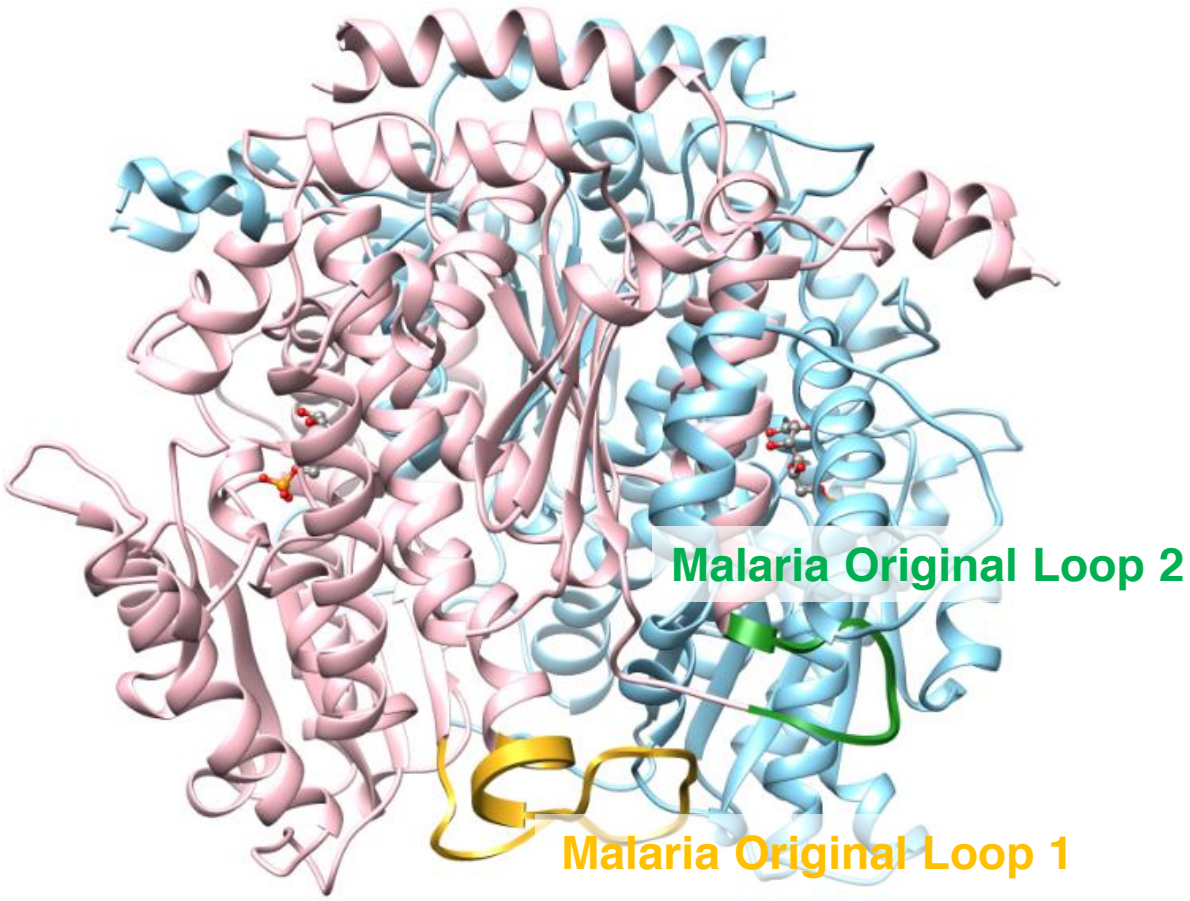


Figure 12. Dimeric structure of the Plasmodium falciparum PGI/6-phosphogluconoic acid (6PGA) complex.
One subunit is colored light blue, and the other is colored red (α-helixes) and cyan (β-strands) for the large domain, orange (α-helixes) and magenta (β-strands) for the small domain, and green (α-helixes) for the C-terminal domain. The bound 6PGA molecules (yellow) are shown as space-filling models for each of the two subunits. Some short helices, and the 310-helix, are not shown. The non-crystallographic two-fold axis is parallel with the plane. (a) Front view of the molecule. The N and C termini, and the hooks are marked. (b) A side-view of the molecule. The large, small, and C-terminal domains are marked. The view in (b) is obtained by rotating the view in (a) by 90º. (c) malaria original loop domains 1 and 2 (MOL1 and Mol2) are shown in yellow and green respectively.

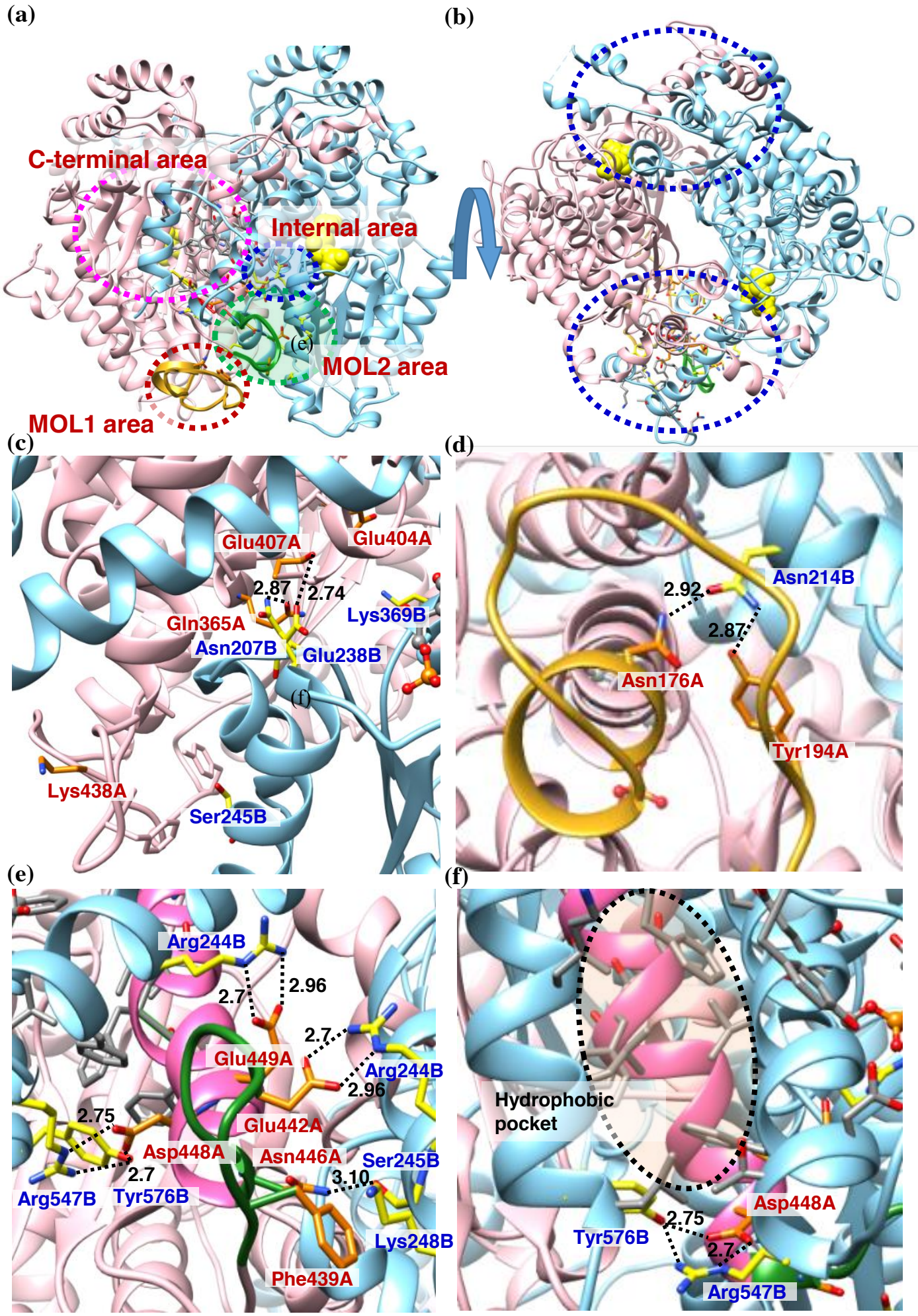

(a)
(b)
C-terminal area
Internal area
(e)
MOL2 area
MOL1 area
(c)
(d)
Glu407A
Glu404A
2.87
2.74
Gln365A
Lys369B
Asn207B
Glu238B
(f)
Lys438A
Ser245B
2.92
Asn214B
2.87
Asn176A
Tyr194A
(e)
(f)
Arg244B
2.7
2.96
2.7
Glu449A
Arg244B
2.96
2.75
Glu442A
Asp448A
Asn446A
Ser245B
2.7
3.10
Arg547B
Tyr576B
Lys248B
Phe439A
Hydrophobic
pocket
Asp448A
2.75
Tyr576B
2.7
Arg547B

Figure 13. Dimeric interaction between subunits.
Overview of dimeric interactions. The four interaction areas (internal, MOL1, MOL2, and C-terminal) are shown as dashed circles colored in blue, red, green and pink, respectively. (b) Location overview of the dimeric interaction indicated by the dashed circle through rotating the view in (a) by 90°. (c) The dimeric interaction is shown in the internal area. Dashed lines indicate possible hydrogen bonds. Distances are given in Å. (d) Dimeric interaction in MOL1. (e) Dimeric interaction in MOL2. (f) Dimeric interaction in the C-terminal domain.

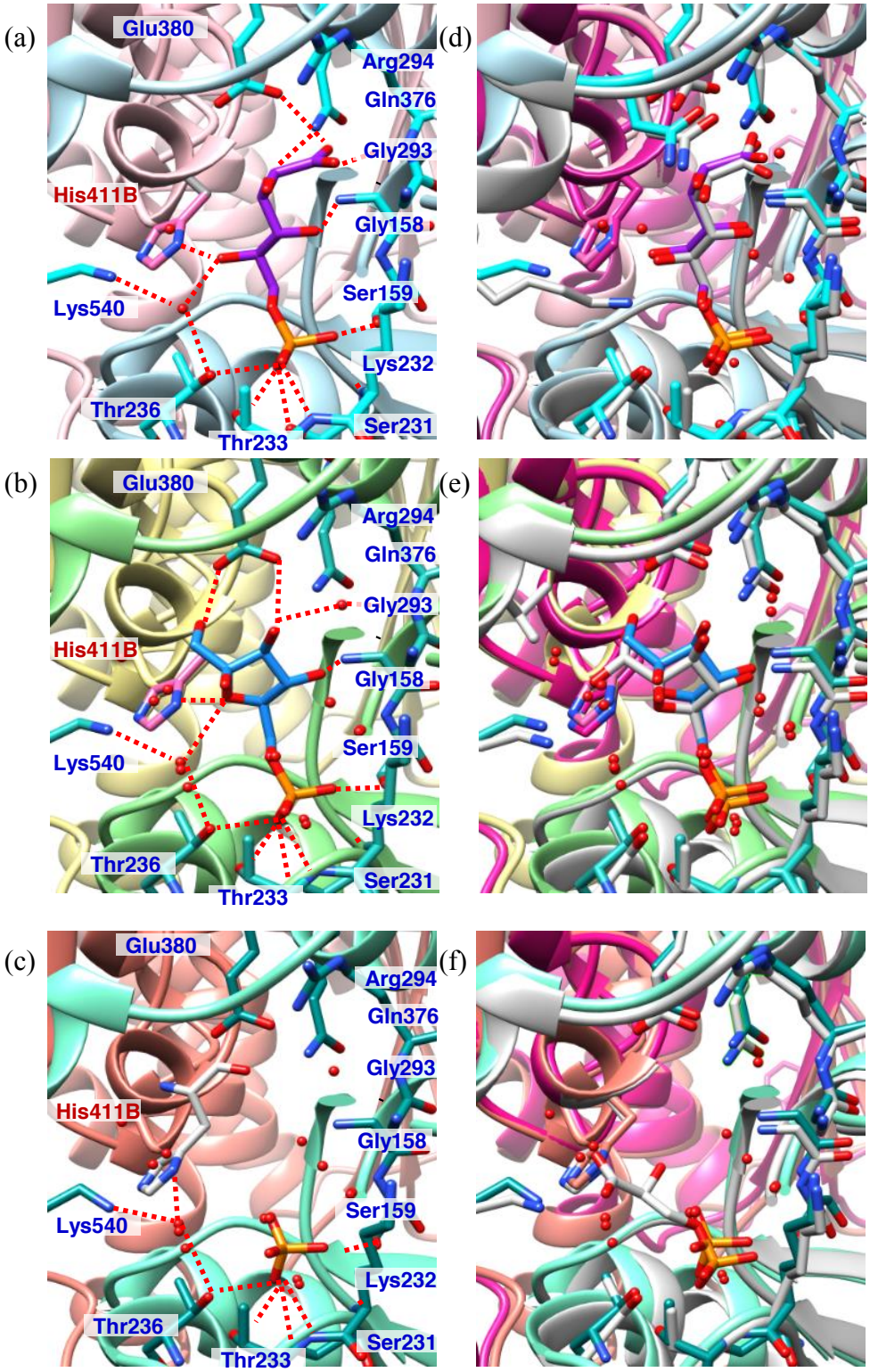


Figure 14. The substrate-binding site in subunit A of PfPGI.
The carbon and phosphorus atoms of bound inhibitor/substrate molecules are shown as follows; purple and orange for 6PGA, and sky blue and orange for fructose 6-phosphate. Possible hydrogen bonds are indicated by the dashed lines (red). The bound water molecules are shown as ball models (red). (a) 6PGA, (b) F6P, and (c) 3PGA. The structural overlap between the PGI substrate binding in PfPGI and mouse PGI. The ribbon diagram and each inhibitor/substrate for mouse PGI is colored in gray; (d) 6PGA, (e) F6P, (f) 3PGA; (g) F6P.

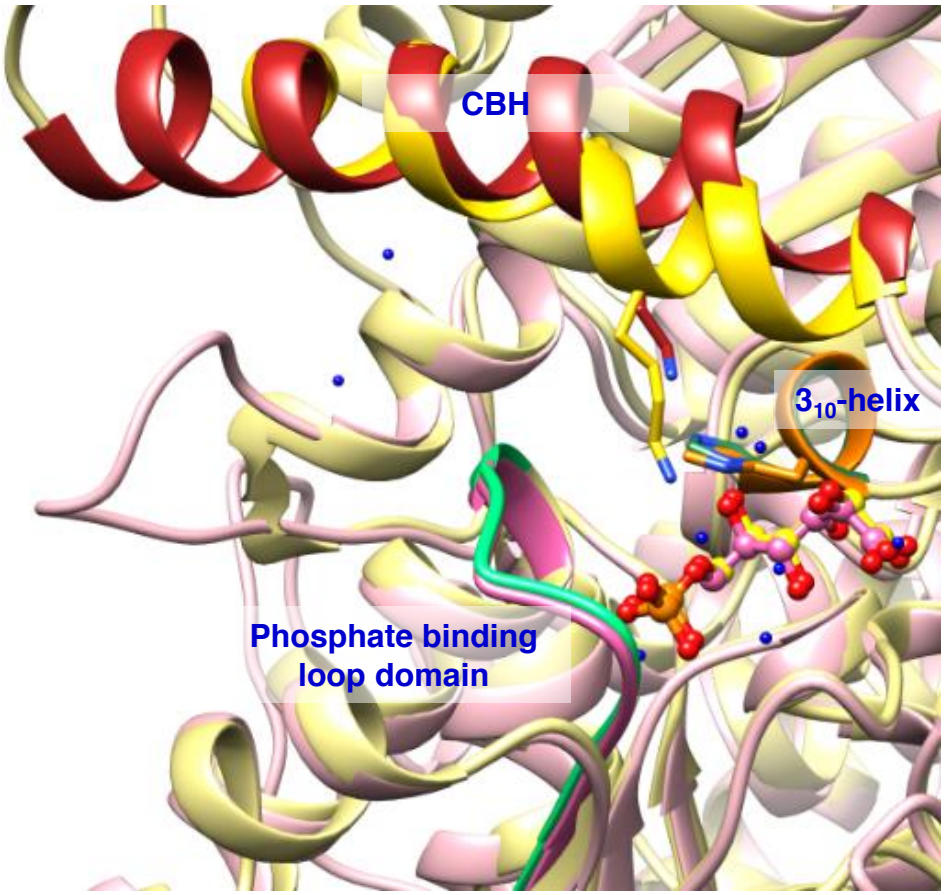
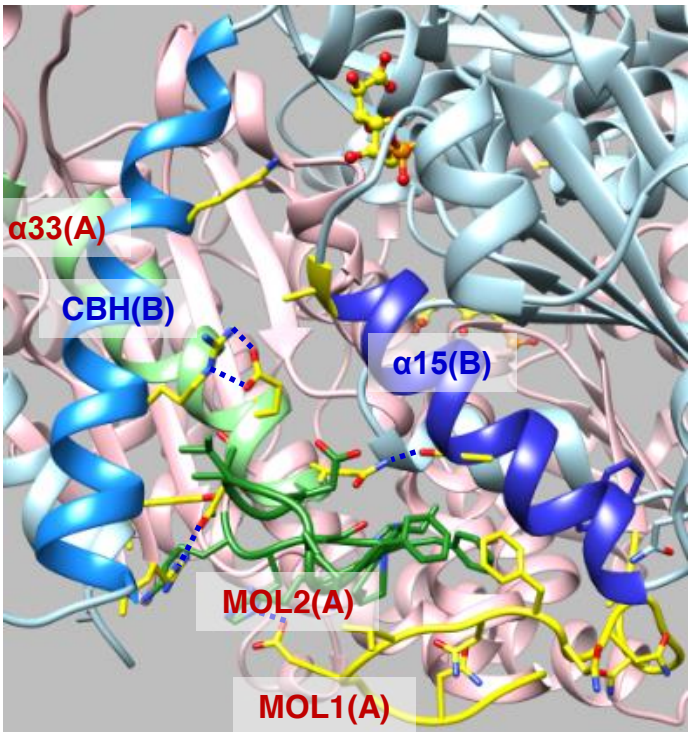


Figure 15. The structure overlap of conformational change between PfPGI and mPGI complex with 6PGA.
This ribbon diagram shows a superimposition of PfPGI in complex with 6PGA (pink, red and orange) and mPGI in complex with 6PGA .
Each phosphate binding loop, 3_{10}-helix and CBH are colored by pink, orange and red in PfPGI, and by light green, green and yellow, in mPGI

(a)



(b)

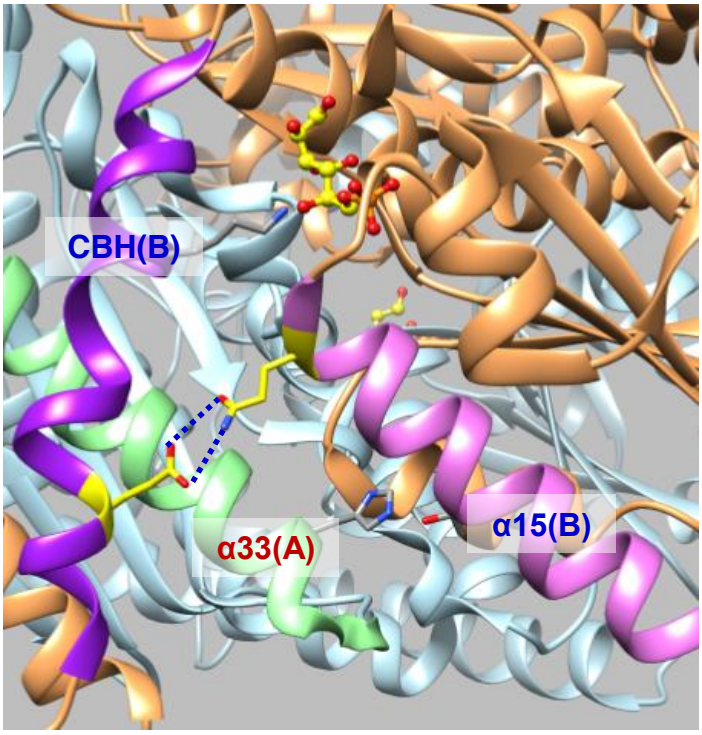


Figure 16. The space retention mechanism of the binding site.
(a) CBH, α33-helix, MOL1, MOL2, and the α15-helix are shown in light blue, light green , yellow, green, and blue in PfPGI, respectively. (b) CBH, α33, and the α15-helix are shown in purple, light green, and pink in mPGI, respectively. Possible hydrogen bonds are indicated by the dashed line (blue). Helixes and loops are colored for each subunit; subunit A and B are in red and blue respectively.

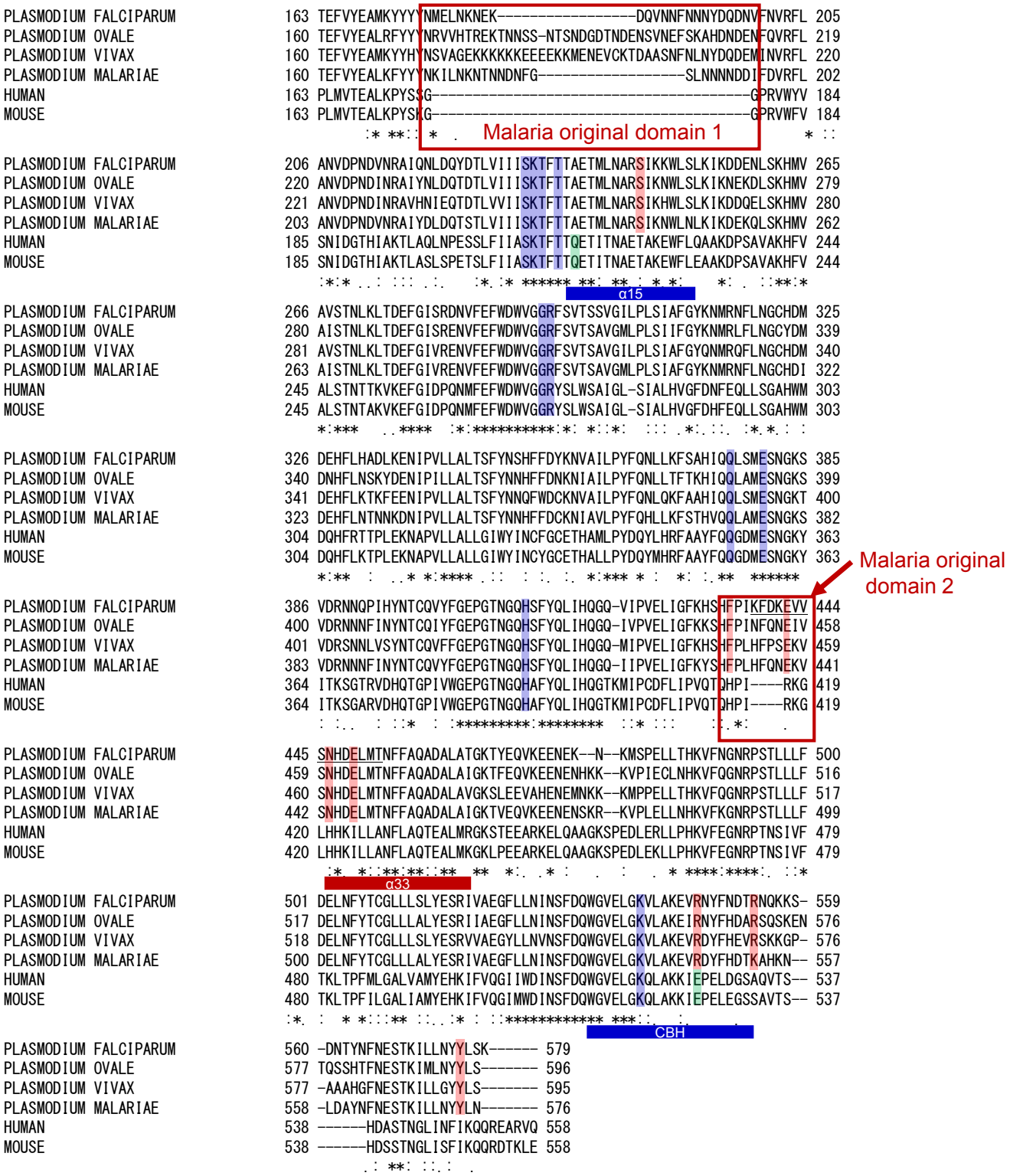


Figure 17. A sequence alignment of PfPGI with PGIs from other species.
The conserved substrate binding site residues are highlighted in blue. The MOL domains are indicated by red boxes. Key dimeric interaction residues for malaria and mouse PGI are highlighted in red and green, respectively.

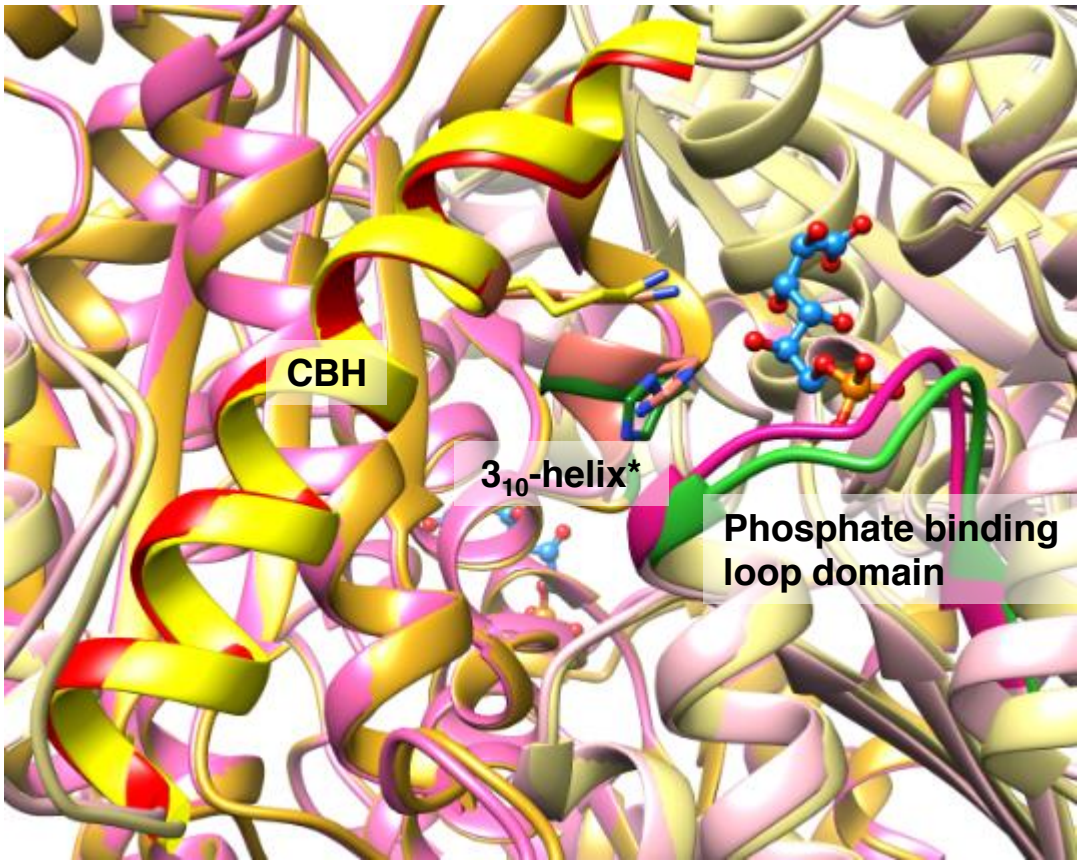


Figure 18. Conformational differences between the substrate-free PfPGI and the PfPGI–6PGA complex.
Each phosphate binding loop, 3_{10}-helix, and CBH are colored pink, orange and red in the substrate free form, and light green, green, and yellow in the 6PGA complexed form.

Table 5. Previous reported PGI complex with inihibitors / open chain form of substrate.

Specie	human [32]	pig [33]	Rabbit [88]	Rabbit [88]	Rabbit [89]	Mouse [34]	Mouse [35]	Mouse [35]	Mouse [35]	Mouse [35]	Mouse [35]	Mouse [35]	Trypanosoma brucei [82]	Francisella tularensis [e]	Pyrobaculum aerophilum [79]
PDB ID	1nuh	1gzv	1g98	1koj	1xtb	1u0f	2cxo	2cxp	2cxq	2cxr	2cxt	2cxu	2o2c	3q7i	1tzc
Ligand	5PHA	A5P	A5P	5PHA	S6P	G6P (open)	E4P	A5P	S6P	6PGA	F6P	M6P	G6P (open)	6PGA	A5P
Phospate binding loop	+ [d]	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3_{10}-Helix [a]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
CBH [b]	+	+	+	+	- [c]	-	+/-	+/-	+/-	+/-	+/-	+/-	-	+	+

[a] Belonging to the other subunit containing His*.
[b] Carbohydrate moiety-binding helix containing Lys in the C-terminal domain.
[c] Open conformation
[d] Closed conformation
[e] Data was only published in PDB

Table 6. K_m value of previous reported crystal structure of PGIs.

Km value (mM)	Spies	pH	Temp. (℃)	Structure cinsideration of CBH
0.037	Human	pH7.5	30	One hydrogen bond between CBH and α15(Gln215-Glu525) [89] Hydrophilic interaction
0.071	Human	pH8.3	30	One hydrogen bond between CBH and α15 (Gln215-Glu525) [90] Hydrophilic interaction
0.058	Mouse	pH7.5	25	One hydrogen bond (Gln216-Glu526) [91] Hydrophilic interaction
0.116	Mouse	pH8.5	30	One hydrogen bond between CBH and α15 (Gln216-Glu526) [92] Hydrophilic interaction
0.119	Rabbit	pH7.5	25	One hydrogen bond (Gln215-Glu525) [93] Hydrophilic interaction
0.06	Pyrobaculum aerophilum	pH7.4	80	One hydorgen bond in CBH (Glu250*-Arg300) [94] Hydrophilic interaction
0.3	Pyrobaculum aerophilum	pH7.4	50	One hydorgen bond in CBH (Glu250*-Arg300) [94] Hydrophilic interaction
0.122	Trypanosoma brusei	pH7.5	25	Hydrophilic interaction [93]
0.26	Plasmodium farciparum	pH8.3	37	Two hydrogenbond (Asp448*-Arg554, Glu449*-Arg547), MOL2 Steric hindrance [95]

* neighbor subunit Km values were collect form BRENDA (The Comprehensive Enzyme Information System) URL: https://www.brenda-enzymes.org/

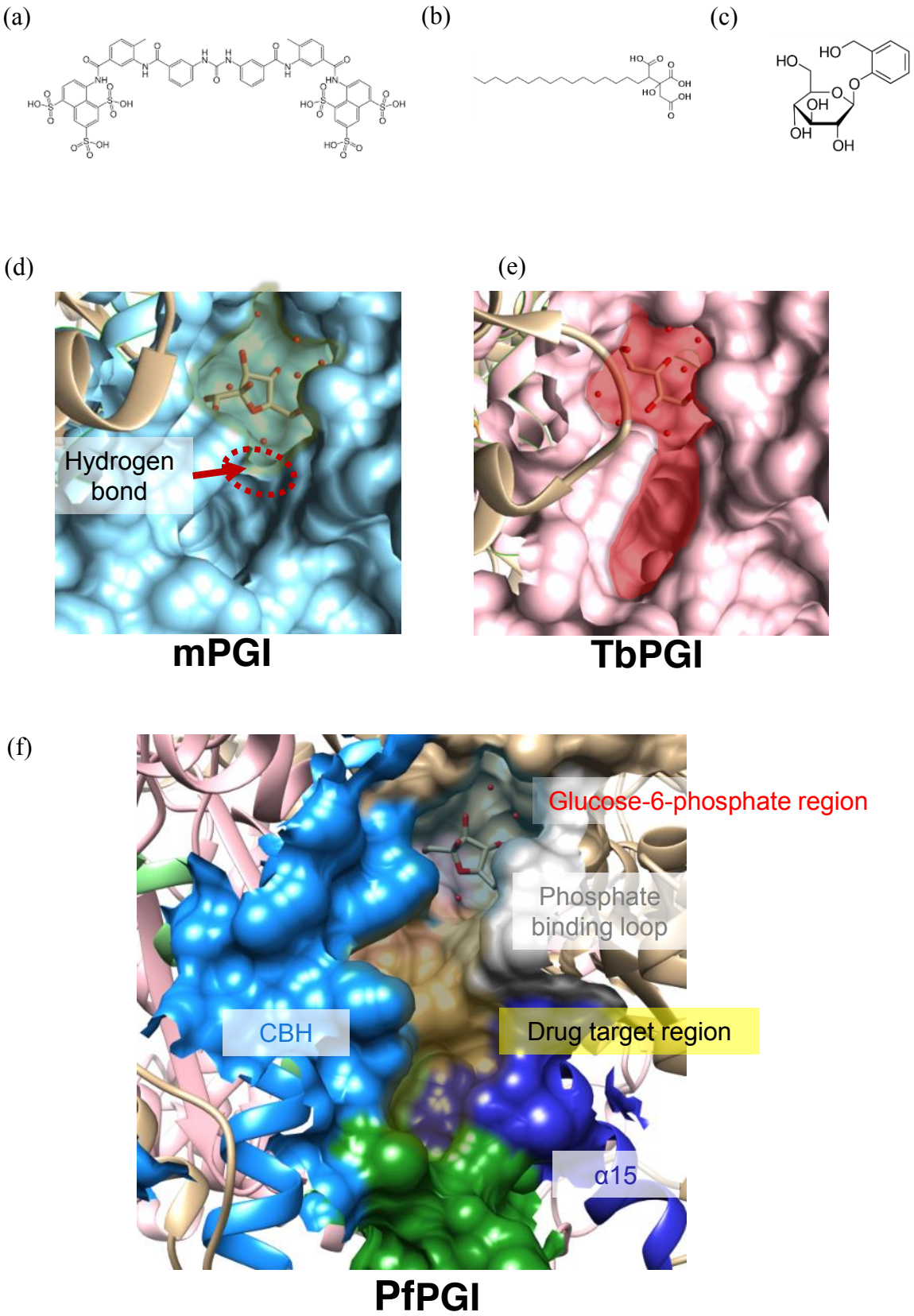


Figure 19. (a) suramin structure, (b) agaricic acid structure, (c) salicin structure , (d) Surface model of the CBH-α15 region in mPGI, (e) Surface model of the CBH-α15 arregion in PGI from *Trypanosoma brucei*, (f) Surface model of the CBH-α15 region in mPGI. (d), (e) and (f) Each pocket between CBH and α15 are shaded by yellow, red and light blue. Drug target region of PfPGI is colored by yellow

Acknowledgements

I am deeply grateful to Associate Professor Ryusuke Niwa, University of Tsukuba, and Associate Professor Nobutada Tanaka, Showa University for guiding my work and valuable discussions through my doctoral program.

I am very thankful to Dr. Shuhei Ishikura; Gifu Pharmaceutical University, Professor Yorishige Imamura; Graduate School of Pharmaceutical Sciences, Kumamoto University, Professor Akira Hara; Gifu Pharmaceutical University, Professor Kazuo T. Nakamura; School of Pharmaceutical Sciences, Showa University, Senior Lecturer Yoshio Kusakabe; Teikyo University, Ms. Chiharu Fukumi; School of Pharmaceutical Sciences, Showa University, Associate Professor Arayo Haga; Gifu University of Medical Science, Associate Professor Masayuki Nakanishi; Matsuyama University, Professor Hiroaki Gouda; Showa University, and Professor Yukio Kitade; Aichi Institute of Technology, for guiding my work and for their valuable discussions throughout my work.

I also wish to thank Dr. Ryuichi Kikuchi, Dr. Kaori Ouchi, Mr. Yoshikazu Yoshida, Dr. Junko Ohata, Dr. Osamu Kuromaru, Mr. Daiji Obara, Mr. Hideo Morikawa and Mr. Motoichiro Fukuchi; Chugai Pharmaceutical Company Limited, for supporting materials, valuable suggestions and comments.

Finally, I wish to thank my family for supporting my life at the University of Tsukuba.

References

1. Stryer L, Berg JM, Tymoczko JL. *Biochemistry (5th ed.)*. 2002; San Francisco: W.H. Freeman. ISBN 0-7167-4955-4956.

2. Anfinsen CB. Principles that govern the folding of protein chains. *Science*. 1973 Jul 20;181(4096):223-230.

3. Debra DM. Enzyme Function Discovery. *Structure*. 2008; 16(11) 1599-1600.

4. Moras D, Billas IM, Rochel N, Klaholz BP. Structure-function relationships in nuclear receptors: the facts. *Trends Biochem Sci*. 2015 Jun;40(6):287-290.

5. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*. 1958 Mar 8;181(4610):662-666.

6. Ducruix A. Giegè R. (1992) Crystallization of Nucleic Acids and Proteins, *Oxford University Press*, NY.ver

7. Hahn T. (ed.) (2002) International Table of Crystallography, *Kluwer Academic Publishers*, Dordrecht.

8. Keith JM. Data, sequence analysis and evolution. Preface. *Methods Mol Biol*. 2008;452:v-vi.

9. Forrest G. L, Gonzalez B. Carbonyl reductase. *Chem Biol Interact*. 2000 Dec 1;129 (1-2):21-40.

10. Wermuth B, Bohren KM, Heinemann G, von Wartburg JP, Gabbay KH. Human carbonyl reductase. Nucleotide sequence analysis of a cDNA and amino acid sequence of the encoded protein. *J Biol Chem*. 1988 Nov 5;263(31):16185-16188.

11. Tanaka M, Ohno S, Adachi S, Nakajin S, Shinoda M, Nagahama Y. Pig testicular 20 beta-hydroxysteroid dehydrogenase exhibits carbonyl reductase-like structure and activity. cDNA cloning of pig testicular 20 beta-hydroxysteroid dehydrogenase. *J Biol Chem*. 1992 Jul 5;267(19):13451-13455.

12. Nakayama T, Yashiro K, Inoue Y, Matsuura K, Ichikawa H, Hara A, Sawada H. Characterization of pulmonary carbonyl reductase of mouse and guinea pig. *Biochim Biophys Acta*. 1986 Jun 19;882(2):220-227.

13. Oritani H, Deyashiki Y, Nakayama T, Hara A, Sawada H, Matsuura K, Bunai Y, Ohya I. Purification and characterization of pig lung carbonyl reductase. *Arch Biochem Biophys*. 1992 Feb 1;292(2):539-547.

14. Jörnvall H, Persson B, Krook M, Atrian S, Gonzàlez-Duarte R, Jeffery J, Ghosh D. Short-chain dehydrogenases/reductases (SDR). *Biochemistry*. 1995 May 9;34(18):6003-6013.

15. Oppermann U, Filling C, Hult M, Shafqat N, Wu X, Lindh M, Shafqat J, Nordling E, Kallberg Y, Persson B, Jörnvall H. Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem Biol Interact*. 2003 Feb 1;143-144:247-253.

16. Hara A, Nakayama T, Deyashiki Y, Kariya K, Sawada H. Carbonyl reductase of dog liver: purification, properties, and kinetic mechanism. *Arch Biochem Biophys*. 1986 Jan;244(1):238-247.

17. Imamura Y, Migita T, Otagiri M, Choshi T, Hibino S. Purification and catalytic properties of a tetrameric carbonyl reductase from rabbit heart. *J Biochem*. 1999 Jan;125(1):41-47.

18. Usami N, Ishikura S, Abe H, Nagano M, Uebuchi M, Kuniyasu A, Otagiri M, Nakayama H, Imamura Y, Hara A. Cloning, expression and tissue distribution of a tetrameric form of pig carbonyl reductase. *Chem Biol Interact*. 2003 Feb 1;143-144:353-361.

19. Tanaka N, Nonaka T, Nakanishi M, Deyashiki Y, Hara A, Mitsui Y. Crystal structure of the ternary complex of mouse lung carbonyl reductase at 1.8 A resolution: the structural origin of coenzyme specificity in the short-chain dehydrogenase/reductase family. *Structure*. 1996 Jan 15;4(1):33-45.

20. Tanaka N, Nonaka T, Tanabe T, Yoshimoto T, Tsuru D, Mitsui Y. Crystal structures of the binary and ternary complexes of 7 alpha-hydroxysteroid dehydrogenase from Escherichia coli. Biochemistry. 1996 Jun 18;35(24):7715-7730.

21. Ghosh D, Sawicki M, Pletnev V, Erman M, Ohno S, Nakajin S, Duax WL. Porcine carbonyl reductase. structural basis for a functional monomer in short chain dehydrogenases/reductases. *J Biol Chem*. 2001 May 25;276(21):18457-18463.

22. Tanaka M, Bateman R, Rauh D, Vaisberg E, Ramachandani S, Zhang C, Hansen KC, Burlingame AL, Trautman JK, Shokat KM, Adams CL. An unbiased cell morphology-based screen for new, biologically active small molecules. *PLoS Biol*. 2005 May;3(5):e128.

23. Duax WL, Ghosh D, Pletnev V. Steroid dehydrogenase structures, mechanism of action, and disease. *Vitam Horm*. 2000;58:121-148.

24. Tanaka N, Nonaka T, Nakamura K T, Hara A. SDR Structure, Mechanism of Action, and Substrate Recognition. *Curr*. *Org*. *Chem*.2001;5:89–111.

25. Webster's New International Dictionary of the English Language, 2nd ed. (1937) Merriam Company, Springfield, Mass.

26. Niinaka Y, Paku S, Haga A, Watanabe H, Raz A. Expression and secretion of neuroleukin/phosphohexose isomerase/maturation factor as autocrine motility factor by tumor cells. *Cancer Res*. 1998 Jun 15;58(12):2667-2674.

27. Gurney ME, Heinrich SP, Lee MR, Yin HS. Molecular cloning and expression of neuroleukin, a neurotrophic factor for spinal and sensory neurons. *Science*. 1986 Oct 31;234(4776):566-574.

28. Chaput M, Claes V, Portetelle D, Cludts I, Cravador A, Burny A, Gras H, Tartar A. The neurotrophic factor neuroleukin is 90% homologous with phosphohexose isomerase. *Nature*. 1988 Mar 31;332(6163):454-455.

29. Read J, Pearce J, Li X, Muirhead H, Chirgwin J, Davies C. The crystal structure of human phosphoglucose isomerase at 1.6 A resolution: implications for catalytic mechanism, cytokine activity and haemolytic anaemia. *J Mol Biol*. 2001 Jun 1;309(2):447-463.

30. Tanaka N, Haga A, Uemura H, Akiyama H, Funasaka T, Nagase H, Raz A, Nakamura KT. Inhibition mechanism of cytokine activity of human autocrine motility factor examined by crystal structure analyses and site-directed mutagenesis studies. *J Mol Biol*. 2002 May 10;318(4):985-997.

31. Cordeiro AT, Godoi PH, Silva CH, Garratt RC, Oliva G, Thiemann OH. Crystal structure of human phosphoglucose isomerase and analysis of the initial catalytic steps. *Biochim Biophys Acta*. 2003 Feb 21;1645(2):117-122.

32. Davies C, Muirhead H, Chirgwin J. The structure of human phosphoglucose isomerase complexed with a transition-state analogue. *Acta Crystallogr D Biol Crystallogr*. 2003 Jun;59(Pt 6):1111-1113

33. Davies C, Muirhead H. Crystal structure of phosphoglucose isomerase from pig muscle and its complex with 5-phosphoarabinonate. *Proteins*. 2002 Dec 1;49(4):577-579.

34. Solomons JT, Zimmerly EM, Burns S, Krishnamurthy N, Swan MK, Krings S, Muirhead H, Chirgwin J, Davies C. The crystal structure of mouse phosphoglucose isomerase at 1.6A resolution and its complex with glucose 6-phosphate reveals the catalytic mechanism of sugar ring opening. *J Mol Biol*. 2004 Sep 17;342(3):847-860.

35. Tanaka N, Haga A, Naba N, Shiraiwa K, Kusakabe Y, Hashimoto K, Funasaka T, Nagase H, Raz A, Nakamura KT. Crystal structures of mouse autocrine motility factor in complex with carbohydrate phosphate inhibitors provide insight into

structure-activity relationship of the inhibitors. *J Mol Biol*. 2006 Feb 17;356(2):312-324.

36. Lazarow PB, Fujiki Y. Biogenesis of peroxisomes. *Annu*. *Rev*. *Cell Biol*. 1985 1, 489–530.

37. Purdue PE, Lazarow PB. Peroxisome biogenesis. *Annu Rev Cell Dev Biol*. 2001;17:701-752.

38. Bonekamp NA, Völkl A, Fahimi HD, Schrader M. Reactive oxygen species and peroxisomes: struggling for balance. *Biofactors*. 2009 Jul-Aug;35(4):346-355.

39. Wanders RJ, Waterham HR. Biochemistry of mammalian peroxisomes revisited. *Annu Rev Biochem*. 2006;75:295-332.

40. Distel B, Erdmann R, Gould SJ, Blobel G, Crane DI, Cregg JM, Dodt G, Fujiki Y, Goodman JM, Just WW, Kiel JA, Kunau WH, Lazarow PB, Mannaerts GP, Moser HW, Osumi T, Rachubinski RA, Roscher A, Subramani S, Tabak HF, Tsukamoto T, Valle D, van der Klei I, van Veldhoven PP, Veenhuis M. A unified nomenclature for peroxisome biogenesis factors. *J Cell Biol*. 1996 Oct;135(1):1-3.

41. Gould SJ, Keller GA, Subramani S. Identification of peroxisomal targeting signals located at the carboxy terminus of four peroxisomal proteins. *J Cell Biol*. 1988 Sep;107(3):897-905.

42. Osumi T, Tsukamoto T, Hata S, Yokota S, Miura S, Fujiki Y, Hijikata M, Miyazawa S, Hashimoto T. Amino-terminal presequence of the precursor of peroxisomal 3-ketoacyl-CoA thiolase is a cleavable signal peptide for peroxisomal targeting. *Biochem Biophys Res Commun*. 1991 Dec 31;181(3):947-954

43. Swinkels BW, Gould SJ, Bodnar AG, Rachubinski RA, Subramani S. A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. *EMBO J*. 1991 Nov;10(11):3255-3262.

44. Brocard C, Kragler F, Simon MM, Schuster T, Hartig A. The tetratricopeptide repeat-domain of the PAS10 protein of Saccharomyces cerevisiae is essential for binding the peroxisomal targeting signal-SKL. *Biochem Biophys Res Commun*. 1994 Nov 15;204(3):1016-1022.

45. Fransen M, Brees C, Baumgart E, Vanhooren JC, Baes M, Mannaerts GP, Van Veldhoven PP. Identification and characterization of the putative human peroxisomal C-terminal targeting signal import receptor. *J Biol Chem*. 1995 Mar 31;270(13):7731-7736.

46. Rehling P, Marzioch M, Niesen F, Wittke E, Veenhuis M, Kunau WH. The import receptor for the peroxisomal targeting signal 2 (PTS2) in Saccharomyces cerevisiae is encoded by the PAS7 gene. *EMBO J*. 1996 Jun 17;15(12):2901-2913.

47. Purdue PE, Zhang JW, Skoneczny M, Lazarow PB. Rhizomelic chondrodysplasia punctata is caused by deficiency of human PEX7, a homologue of the yeast PTS2 receptor. *Nat Genet*. 1997 Apr;15(4):381-384.

48. Braverman N, Steel G, Obie C, Moser A, Moser H, Gould SJ, Valle D. Human PEX7 encodes the peroxisomal PTS2 receptor and is responsible for rhizomelic chondrodysplasia punctata. *Nat Genet*. 1997 Apr;15(4):369-376.

49. Gould SJ, Keller GA, Hosken N, Wilkinson J, Subramani S. A conserved tripeptide sorts proteins to peroxisomes. *J Cell Biol*. 1989 May;108(5):1657-1664.

50. Legakis JE, Terlecky SR. PTS2 protein import into mammalian peroxisomes. *Traffic*. 2001 Apr;2(4):252-260.

51. Usami N, Kitahara K, Ishikura S, Nagano M, Sakai S, Hara A. Characterization of a major form of human isatin reductase and the reduced metabolite. *Eur J Biochem*. 2001 Nov;268(22):5755-5763.

52. Matsuura K, Bunai Y, Ohya I, Hara A, Nakanishi M, Sawada H. Ultrastructural localization of carbonyl reductase in mouse lung. *Histochem J*. 1994 Apr;26(4):311-316.

53. Endo S, Matsunaga T, Nagano M, Abe H, Ishikura S, Imamura Y, Hara A. Characterization of an oligomeric carbonyl reductase of dog liver: its identity with

peroxisomal tetrameric carbonyl reductase. *Biol Pharm Bull*. 2007 Sep;30(9):1787-1791.

54. Jancarik J, Scott WG, Milligan DL, Koshland DE Jr, Kim SH. Crystallization and preliminary X-ray diffraction study of the ligand-binding domain of the bacterial chemotaxis-mediating aspartate receptor of Salmonella typhimurium. *J Mol Biol*. 1991 Sep 5;221(1):31-34.
55. Rossmann MG, van Beek CG. Data processing. *Acta Crystallogr D Biol Crystallogr*. 1999 Oct;55(Pt 10):1631-1640.
56. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. *J Mol Biol*. 1993 Sep 5;233(1):123-138.
57. Merritt EA, Murphy ME. Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 1994 Nov 1;50(Pt 6):869-873.
58. McRee DE. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol*. 1999 Apr-May;125(2-3):156-165.
59. Golinelli-Pimpaneau B, Gigant B, Bizebard T, Navaza J, Saludjian P, Zemel R, Tawfik DS, Eshhar Z, Green BS, Knossow M. Crystal structure of a catalytic antibody Fab with esterase-like activity. *Structure*. 1994 Mar 15;2(3):175-183.

60. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*. 1997 May 1;53(Pt 3):240-255.

61. Ghosh D, Weeks CM, Grochulski P, Duax WL, Erman M, Rimsay RL, Orr JC. Three-dimensional structure of holo 3 alpha,20 beta-hydroxysteroid dehydrogenase: a member of a short-chain dehydrogenase family. *Proc Natl Acad Sci U S A*. 1991 Nov 15;88(22):10064-10068

62. Filling C, Berndt KD, Benach J, Knapp S, Prozorovski T, Nordling E, Ladenstein R, Jörnvall H, Oppermann U. Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J Biol Chem*. 2002 Jul 12;277(28):25677-25684.

63. Yoshihara T, Hamamoto T, Munakata R, Tajiri R, Ohsumi M, Yokota S. Localization of cytosolic NADP-dependent isocitrate dehydrogenase in the peroxisomes of rat liver cells: biochemical and immunocytochemical studies. J *Histochem Cytochem*. 2001 Sep;49(9):1123-1131.

64. Deshaies RJ, Koch BD, Werner-Washburne M, Craig EA, Schekman R. A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. *Nature*. 1988 Apr 28;332(6167):800-805.

65. Walton PA, Wendland M, Subramani S, Rachubinski RA, Welch WJ. Involvement of 70-kD heat-shock proteins in peroxisomal import. *J Cell Biol*. 1994 Jun;125(5):1037-1046.

66. Eilers M, Schatz G. Protein unfolding and the energetics of protein translocation across biological membranes. *Cell*. 1988 Feb 26;52(4):481-483.

67. Faber KN, van Dijk R, Keizer-Gunnink I, Koek A, van der Klei IJ, Veenhuis M. Import of assembled PTS1 proteins into peroxisomes of the yeast Hansenula polymorpha: yes and no! *Biochim Biophys Acta*. 2002 Aug 19;1591(1-3):157-162.

68. Gatto GJ Jr, Geisbrecht BV, Gould SJ, Berg JM. Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat Struct Biol*. 2000 Dec;7(12):1091-1095.

69. Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F. Motif refinement of the peroxisomal targeting signal 1 and evaluation of taxon-specific differences. *J Mol Biol*. 2003 May 2;328(3):567-579.

70. Gayathri P, Balaram H, Murthy MR. Structural biology of plasmodial proteins. *Curr Opin Struct Biol*. 2007 Dec;17(6):744-754.

71. Xu W, Seiter K, Feldman E, Ahmed T, Chiao JW. The differentiation and maturation mediator for human myeloid leukemia cells shares homology with neuroleukin or phosphoglucose isomerase. *Blood*. 1996 Jun 1;87(11):4502-4506.

72. Yakirevich E, Naot Y. Cloning of a glucose phosphate isomerase/neuroleukin-like sperm antigen involved in sperm agglutination. *Biol Reprod*. 2000 Apr;62(4):1016-1023.

73. Matsumoto I, Staub A, Benoist C, Mathis D. Arthritis provoked by linked T and B cell recognition of a glycolytic enzyme. *Science*. 1999 Nov 26;286(5445):1732-1735.

74. Cao MJ, Osatomi K, Matsuda R, Ohkubo M, Hara K, Ishihara T. Purification of a novel serine proteinase inhibitor from the skeletal muscle of white croaker (Argyrosomus argentatus). *Biochem Biophys Res Commun*. 2000 Jun 7;272(2):485-489.

75. Amraei M, Nabi IR. Species specificity of the cytokine function of phosphoglucose isomerase. *FEBS Lett*. 2002 Aug 14;525(1-3):151-155.

76. Tsutsumi S, Gupta SK, Hogan V, Tanaka N, Nakamura KT, Nabi IR, Raz A. The enzymatic activity of phosphoglucose isomerase is not required for its cytokine function. *FEBS Lett*. 2003 Jan 16;534(1-3):49-53.

77. Jeffery CJ, Bahnson BJ, Chien W, Ringe D, Petsko GA. Crystal structure of rabbit phosphoglucose isomerase, a glycolytic enzyme that moonlights as neuroleukin, autocrine motility factor, and differentiation mediator. Biochemistry. 2000 Feb 8;39(5):955-964.

78. Sun YJ, Chou CC, Chen WS, Wu RT, Meng M, Hsiao CD. The crystal structure of a multifunctional protein: phosphoglucose isomerase/autocrine motility factor/neuroleukin. *Proc Natl Acad Sci U S A*. 1999 May 11;96(10):5412-5417.

79. Swan MK, Hansen T, Schönheit P, Davies C. A novel phosphoglucose isomerase (PGI)/phosphomannose isomerase from the crenarchaeon Pyrobaculum aerophilum is a member of the PGI superfamily: structural evidence at 1.16-A resolution. *J Biol Chem*. 2004 Sep 17;279(38):39838-39845.

80. Cordeiro AT, Michels PA, Delboni LF, Thiemann OH. The crystal structure of glucose-6-phosphate isomerase from Leishmania mexicana reveals novel active site features. *Eur J Biochem*. 2004 Jul;271(13):2765-2772.

81. Yamamoto H, Miwa H, Kunishima N. Crystal structure of glucose-6-phosphate isomerase from Thermus thermophilus HB8 showing a snapshot of active dimeric state. *J Mol Biol*. 2008 Oct 10;382(3):747-762.

82. Arsenieva D, Appavu BL, Mazock GH, Jeffery CJ. Crystal structure of phosphoglucose isomerase from Trypanosoma brucei complexed with glucose-6-phosphate at 1.6 A resolution. *Proteins*. 2009 Jan;74(1):72-80.

83. Matthews BW. Solvent content of protein crystals. *J Mol Biol*. 1968 Apr 28;33(2):491-497.

84. Vagin A, Teplyakov A, MOLREP: an Automated Program for Molecular Replacement. *J*. *Appl*. *Cryst*. 1997 30, 1022-1025

85. Pignol D, Gaboriaud C, Fontecilla-Camps JC, Lamzin VS, Wilson KS. How to escape from model bias with a high-resolution native data set - structure determination of the PcpA-S6 subunit III. Acta Crystallogr D Biol Crystallogr. 1996 Mar 1;52(Pt 2):345-355.

86. Lee JH, Chang KZ, Patel V, Jeffery CJ. Crystal structure of rabbit phosphoglucose isomerase complexed with its substrate D-fructose 6-phosphate. *Biochemistry*. 2001 Jul 3;40(26):7799-7805.

87. Jeffery CJ, Hardré R, Salmon L. Crystal structure of rabbit phosphoglucose isomerase complexed with 5-phospho-D-arabinonate identifies the role of Glu357 in catalysis. *Biochemistry*. 2001 Feb 13;40(6):1560-1566.

88. Lee JH, Jeffery CJ. The crystal structure of rabbit phosphoglucose isomerase complexed with D-sorbitol-6-phosphate, an analog of the open chain form of D-glucose-6-phosphate. *Protein Sci*. 2005 Mar;14(3):727-734.

89. Lin HY, Kao YH, Chen ST, Meng M. Effects of inherited mutations on catalytic activity and structural stability of human glucose-6-phosphate isomerase expressed in Escherichia coli. *Biochim Biophys Acta*. 2009 Feb;1794(2):315-323.

90. Gracy RW, Tilley BE. Phosphoglucose isomerase of human erythrocytes and cardiac tissue. *Methods Enzymol*. 1975;41:392-400.

91. Pretsch W, Merkle S. Glucose phosphate isomerase enzyme-activity mutants in Mus musculus: genetical and biochemical characterization. *Biochem Genet*. 1990 Feb;28(1-2):97-110.

92. Röhm KH, Schneider F. Glucose-6-phosphate isomerase from Ehrlich ascites tumor cells: purification and some properties. *FEBS Lett*. 1973 Jun 15;33(1):89-92.

93. Marchand M, Kooystra U, Wierenga RK, Lambeir AM, Van Beeumen J, Opperdoes FR, Michels PA. Glucosephosphate isomerase from Trypanosoma brucei. Cloning and characterization of the gene and analysis of the enzyme. *Eur J Biochem*. 1989 Sep 15;184(2):455-464.

94. Hansen T, Urbanke C, Schönheit P. Bifunctional phosphoglucose/phosphomannose isomerase from the hyperthermophilic archaeon Pyrobaculum aerophilum. *Extremophiles*. 2004 Dec;8(6):507-512.

95. Srivastava IK, Schmidt M, Grall M, Certa U, Garcia AM, Perrin LH. Identification and purification of glucose phosphate isomerase of Plasmodium falciparum. *Mol Biochem Parasitol*. 1992 Sep;54(2):153-164.

96. Venkatesan R, Sah-Teli SK, Awoniyi LO, Jiang G, Prus P, Kastaniotis AJ, Hiltunen JK, Wierenga RK, Chen Z. Insights into mitochondrial fatty acid synthesis from the structure of heterotetrameric 3-ketoacyl-ACP reductase/3R-hydroxyacyl-CoA dehydrogenase. *Nat Commun*. 2014 Sep 9;5:4805

97. Chen, Z. et al. 17beta-hydroxysteroid dehydrogenase type 8 and carbonylreductase type 4 assemble as a ketoacyl reductase of human mitochondrial FAS.FASEB J. 23, 3682–3691 (2009)

98. Hardré R, Bonnette C, Salmon L, Gaudemer A. Synthesis and evaluation of a new inhibitor of phosphoglucose isomerases: the enediolate analogue 5-phospho-D-arabinohydroxamate. *Bioorg Med Chem Lett*. 1998 Dec 1;8(23):3435-3438.

99. Hardré R, Salmon L, Opperdoes FR. Competitive inhibition of Trypanosoma brucei phosphoglucose isomerase by D-arabinose-5-phosphate derivatives. *J Enzyme Inhib*. 2000;15(5):509-515.

100. Wick JY. Aspirin: a history, a love story. *Consult Pharm*. 2012 May;27(5):322-329.