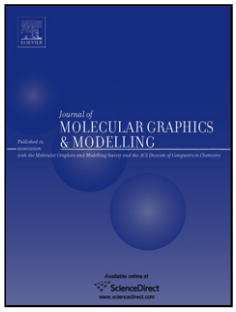
Accepted Manuscript



Title: Structural Modeling and Docking Studies of Ribose 5-Phosphate Isomerase from *Leishmania major* and *Homo sapiens*: A Comparative Analysis for Leishmaniasis Treatment

Author: Priscila V.S.Z. Capriles Luiz Phillippe R. Baptista Isabella A. Guedes Ana Carolina R. Guimarães Fabio L. Custódio Marcelo Alves-Ferreira Laurent E. Dardenne

PII: S1093-3263(14)00179-X
DOI: http://dx.doi.org/doi:10.1016/j.jmgm.2014.11.002
Reference: JMG 6481

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 20-3-2014
Revised date: 22-10-2014
Accepted date: 7-11-2014

Please cite this article as: P.V.S.Z. Capriles, L.P.R. Baptista, I.A. Guedes, A.C.R. Guimarães, F.L. Custódio, M. Alves-Ferreira, L.E. Dardenne, Structural Modeling and Docking Studies of Ribose 5-Phosphate Isomerase from *Leishmania major* and *Homo sapiens*: A Comparative Analysis for Leishmaniasis Treatment, *Journal of Molecular Graphics and Modelling* (2014), http://dx.doi.org/10.1016/j.jmgm.2014.11.002

Structural Modeling and Docking Studies of Ribose 5-Phosphate Isomerase from *Leishmania major* and *Homo sapiens*: A Comparative Analysis for Leishmaniasis Treatment

Priscila V. S. Z. Capriles[1,2*‡], *Luiz Phillippe R. Baptista*[1,3‡], *Isabella A. Guedes*[1‡], *Ana Carolina R. Guimarães*[3,4], *Fabio L. Custódio*[1], *Marcelo Alves-Ferreira*[3], *Laurent E. Dardenne*[1*]

1- Grupo de Modelagem Molecular de Sistemas Biológicos, Laboratório Nacional de Computação Científica, GMMSB/LNCC-MCTI, Petrópolis, Brazil.

2- Programa de Pós-graduação em Modelagem Computacional, Faculdade de Engenharia / Instituto de Ciências Exatas, Universidade Federal de Juiz de Fora, PGMC/UFJF-MEC, Juiz de Fora, Brazil.

3- Laboratório de Genômica Funcional e Bioinformática, Instituto Oswaldo Cruz, IOC/FIOCRUZ-MS, Rio de Janeiro, Brazil.

4- Centro de Desenvolvimento Tecnológico em Saúde, Instituto Oswaldo Cruz, CDTS/FIOCRUZ-MS, Rio de Janeiro, Brazil.

ABSTRACT

Leishmaniases are caused by protozoa of the genus *Leishmania* and are considered the second-highest cause of death worldwide by parasitic infection. The drugs available for treatment in humans are becoming ineffective mainly due to parasite resistance; therefore, it is extremely important to develop a new chemotherapy against these parasites. A crucial aspect of drug design development is the identification and characterization of novel molecular targets. In this work, through an *in silico* comparative analysis between the genomes of *Leishmania major* and *Homo sapiens*, the enzyme ribose 5-phosphate isomerase (R5PI) was indicated as a promising molecular target. R5PI is an important enzyme that acts in the pentose phosphate pathway and catalyzes the interconversion of D-ribose-5-phosphate (R5P) and D-ribulose-5-phosphate (5RP). R5PI activity is found in two analogous groups of enzymes called RpiA (found in *H. sapiens*) and RpiB (found in *L. major*). Here, we present the first report of the three-dimensional (3D) structures and active sites of RpiB from *L. major* (LmRpiB) and RpiA from *H. sapiens* (HsRpiA). Three-dimensional models were constructed by applying a hybrid methodology that combines comparative and *ab initio* modeling techniques, and the active site was characterized based on docking studies of the substrates R5P (furanose and ring-opened forms) and 5RP. Our comparative analyses show that these proteins are structural analogs and that distinct residues participate in the interconversion of R5P and 5RP. We propose two distinct

reaction mechanisms for the reversible isomerization of R5P to 5RP, which is catalyzed by LmRpiB and HsRpiA. We expect that the present results will be important in guiding future molecular modeling studies to develop new drugs that are specially designed to inhibit the parasitic form of the enzyme without significant effects on the human analog.

KEYWORDS

Ribose 5-Phosphate Isomerase; *Leishmania major*; *Homo sapiens*; Leishmaniasis Treatment; Protein Structure Prediction; Protein-Ligand Docking; Reaction Mechanism; Molecular Target; Pentose Phosphate Pathway.

INTRODUCTION

Leishmaniases are infectious diseases that are endemic in 88 countries and that, according to the Center for Disease Control and Prevention (CDC), are found on every continent except Australia and Antarctica.[1,2] These diseases are considered the second-highest cause of death worldwide by parasitic infection and rank third in disease burden in disability-adjusted life years (DALY) caused by neglected tropical diseases.[3] Leishmaniases have a prevalence of 12 million humans infected and an incidence about 14 million humans, and 350 million people are considered at risk of these diseases. The disease burden is calculated at approximately 2.3 million DALY (approximately 0.9 million in women and 1.4 million in men), and approximately 2.0 million new cases are known to occur yearly with many others undetected and underreported.[3-5]

Leishmaniases are caused by protozoa of the genus *Leishmania*, of the family Trypanosomatidae. No vaccine against the parasite is available that can be administered in humans; therefore, disease control relies on therapeutic treatments, such as the administration of drugs, including pentavalent antimonials, pentamidine, amphotericin B, miltefosine and fluconazole.[6,7] Furthermore, the available drugs are becoming ineffective, mainly because the parasite is developing resistance against them. A great deal is known about the resistance mechanisms developed by the parasite, for example, resistances to amphotericin B and pentavalent antimony.[6] Some reports have described toxic effects of these medicines in patients;[8] therefore, it is extremely important to develop new drugs against these parasites.

A crucial aspect of drug design and development is the identification and characterization of novel and promising molecular targets. In most cases, drug targets are specific to the parasite; thus, inhibiting the action of these targets will not cause many side effects in humans. Many parasite-specific targets are currently under study, including thiol-redox enzymes (*e.g.,* trypanothione disulfide reductase)[9] and trans-splicing-specific components.[10] Another way to search for new drug targets is to seek proteins that have analogs in both parasite and host.[11] Analogous enzymes result from independent evolutionary events. Thus, their structures should be significantly different, allowing the possibility to design a ligand that may inhibit one enzyme while having no or a lesser effect on the other.

One of the research groups associated with this study has developed a computational tool called AnEnPi that is used to find analogous enzymes among a group of organisms.[12] This tool was designed to identify, annotate and compare analogous enzymes

by combining several bioinformatics algorithms. AnEnPi is able to perform the following tasks: (i) grouping protein sequences; (ii) analyzing and interpreting all groups generated from all enzymes in the KEGG (Kyoto Encyclopedia of Genes and Genomes) database (http://www.genome.jp/kegg/);[13] (iii) detecting intragenomic analogy (the existence of different forms of the same enzyme within the same organism); (iv) detecting intergenomic analogy between two organisms (different forms of the same enzymatic activity in different organisms); (v) annotating protein sequences (using BLAST or HMMER); (vi) generating metabolic maps. After clustering, sequences representing the same enzymatic activity (Enzyme Comission number - EC) that have been allocated to different groups are considered by AnEnPi as analogous. Analogous enzymes in different clusters have low sequence similarity and will probably have important structural differences.

In this work, similar to the work done in *Trypanosoma cruzi*,[14] we analyzed functional analogous proteins from *Leishmania major* and *Homo sapiens* to identify promising novel molecular targets in the parasite. Ribose 5-phosphate isomerase (R5PI) (EC 5.3.1.6, an important pentose phosphate pathway (PPP) enzyme) emerged from the analysis performed with AnEnPi as having analogous enzymes in *H. sapiens*. R5PI catalyzes the interconversion of D-ribose-5-phosphate (R5P) and D-ribulose-5-phosphate (5RP) (Figure 1). Ribose is a sugar involved in nucleotide and histidine biosynthesis, and in the synthesis of ATP and NAD cofactors. The PPP metabolic pathway produces NADPH, which can be used in biosynthetic pathways and protects the organism from oxidative stress. *Leishmania* parasites, when in the human host, proliferate inside phagolysosomes; therefore, PPP is expected to be important for this protozoa to protect itself against free radicals released by macrophages.[15]

R5PI activity is found in two analogous groups of enzymes called RpiA and RpiB.[16] RpiA is found in all taxonomic groups, whereas RpiB is concentrated in bacteria and lower eukaryotes (including parasites of the family Trypanosomatidae). In most cases, organisms that express both forms of the enzyme use RpiA constitutively whereas RpiB is induced by other factors. In *Escherichia coli*, the *rpiB* gene is part of an operon controlled by the rare sugar allose;[17] it has been shown that this enzyme can catalyze the isomerization of allose 6-phosphate to allulose 6-phosphate, in some cases.[18]

This activity has been described in humans,[19] and genomic data show that the A form catalyzes the reaction. Studies with this enzyme in humans reveal that its absence can cause disease, mainly the slowly progressive leukoencephalopathy.[20] However, due to the absence of experimentally determined structures, no comparative structural analysis has yet been conducted between these analogous enzymes. However, such a structural comparison between the parasite and host enzymes should reveal very interesting insights for drug discovery projects.

In this work, we report the first analysis of *Leishmania major* ribose 5-phosphate isomerase (R5PI). We present a comparative analysis of the three-dimensional (3D) structures of *L. major* (LmRpiB) and *H. sapiens* (HsRpiA) ribose 5-phosphate isomerases as constructed using a hybrid computational methodology that combines comparative and *ab initio* modeling techniques. Docking studies with the substrates R5P (furanose and ring-opened forms) and 5RP are performed to obtain physical-chemical insights into the active site characterization of LmRpiB and HsRpiA. Finally, we propose two distinct reversible catalytic mechanisms (R5P ↔ 5RP) associated with ribose-5-phosphate isomerase type B of *L. major* and ribose-5-phosphate isomerase type A of *H. sapiens.*

EXPERIMENTAL SECTION

1. Identification of R5PI as a molecular target for *Leishmania major*

All predicted proteins from the genome of *Leishmania major* (http://tritrypdb.org/tritrypdb/, release 3.0) were annotated using the "Analogous Enzymes Annotation via Blast" module of AnEnPi[13] with an E-value ≤ e-80. After obtaining the enzyme activity, the sequences were compared against the parasite groups generated by AnEnPi to identify enzymatic activities whose sequences in the parasite were allocated to distinct groups of human sequences. These activities were identified as analogous. Enzymes with intergenomic analogy between the parasite and humans were mapped onto metabolic pathways using the "Metabolic Pathways and Maps" module of AnEnPi, and R5PI was found to be part of a pathway (EC 5.3.1.6) considered critical to the parasite.

2. Three-dimensional protein model construction

Three-dimensional (3D) models of LmRpiB (LmjF28.1970) and HsRpiA (Hsa:22934) were constructed by a hybrid method using comparative modeling and *ab initio* techniques, applying the programs MODELLER 9v11 (www.salilab.org/modeller)[21] and GAPF (www.gmmsb.lncc.br), respectively.[22]

2.1. Choice of template:

Local alignments were performed between the target sequences and proteins from the Protein Data Bank (PDB) (http://www.pdb.org)[23] using the program Blastp.[24] The template structures were selected based on amino acid sequence identity, similarity, number of gaps and on the presence of ligands/heteroatoms in the active site. For LmRpiB model construction, we chose two structures: (i) 3K7S,[25] the R5PI type B structure from *Trypanosoma cruzi* (identity: 51%), and (ii) 3HEE,[26] the R5PI type B structure from *Clostridium thermocellum* (identity: 45%). The templates selected for the 3D modeling of HsRpiA were: (i) 1XTZ,[27] the R5PI type A structure from *Saccharomyces cerevisiae* (identity: 46%); (ii) 1LK7,[28] the R5PI structure from *Pyrococcus horikoshii* (identity: 42%); (iii) 2F8M,[29] the R5PI type A structure from *Plasmodium falciparum* (identity: 35%) and (iv) 3HHE,[30] the R5PI type A structure from *Bartonella henselae* (identity: 39%). The alignment analyses showed that no PDB sequence was able to align against the 20 C-terminal amino acids of LmRpiB and against the 77 N-terminal amino acids of HsRpiA. We investigated the named segments in search of peptide signal or transmembrane regions using the SignalP 4.0 server,[31] but all results were negative, suggesting that these regions are neither signal peptides nor transmembrane regions.

2.2. Ab initio protocol:

The three-dimensional (3D) structure of the 20 C-terminal residues of LmRpiB was predicted using the *ab initio* protein structure prediction program GAPF. The parsed sequence contained 25 residues: the 20 final residues (i.e., without template) and five residues whose structure was determined by the comparative modeling protocol. The

selected best structure (called GAPF25) was chosen as it had the lowest energy amongst the structures exhibiting good agreement between the *ab initio* predicted structure and the comparative modeling structure for the five additional residues. The final structure of the 20 C-terminal residues of LmRpiB was used as a template for a second round of the comparative modeling protocol.

The same GAPF protocol was applied to model the 77 N-terminal residues of HsRpiA. However, the *ab initio* protocol used is unreliable for sequences longer than 50 residues. For that reason, additional structures were also modeled using the template-free, fragment-based Rosetta software.[32] The 77 N-terminal residues include a large number of glycines and very few hydrophobic amino acids, suggesting that the region is highly flexible. Nevertheless, in all constructed models, these initial 77 residues appeared to form an additional coiled domain that partially obstructed the active site entrance but never interacted with any active site residues. Due to the low reliability of the generated models, the 3D structure of the 77 N-terminal residues of HsRpiA was not included in the final structure.

2.3. Comparative modeling protocol:

The complete 3D models were constructed using the class automodel from MODELLER 9v11[21] using the structures 3K7S:A, 3K7S:B, 3HEE:A, 3HEE:B and GAPF25 as templates for LmRpiB and the structures of 1XTZ, 1LK7:A, 1LK7:B, 1LK7:C, 1LK7:D, 2F8M:A, 2F8M:B, 3HHE:A and 3HHE:B as templates for HsRpiA. The dimeric protein LmRpiB and the tetrameric protein HsRpiA were constructed by restraining only

alpha-carbon atoms to impose symmetry in the subunits by reporting violations of greater than 1.0 Å. The ligands R5P of 3K7S and 5RP of 3HHE were considered during model construction of LmRpiB and HsRpiA, respectively

The energy quality of the models was assessed using DOPE (Discrete Optimized Protein Energy) and the normalized DOPE methods included in the MODELLER 9v11 program, and the stereochemical quality of the models was assessed using the PROCHECK v.3.5.4[33] and MOLPROBITY v.3.19[34] programs. Additional analysis was performed using the QMEAN server.[35]

3. Molecular docking

3.1. Structural preparation of LmRpiB and HsRpiA:

Previous studies about RpiB of *Mycobacterium tuberculosis* (MtRpiB)[36,] *E. coli* (EcRpiB)[36,50] and *Trypanosoma cruzi* (TcRpiB)[25,46] proposed that H102 is involved in the opening of the furanose ring of R5P. It is important to note that the conformation H102:ND1/NE2 is the most likely form involved in the ring opening of R5P, since it donates the proton from the ND1 to the substrate generating the H102:NE2 form. Additionally, they proposed that C66-EcRpiB/C69-TcRpiB acts as the catalytic base responsible for the isomerization step, accepting a proton from C2 and returning it to C1 in the final step of catalysis. Deprotonated form of C69 was chosen based on the observation that these cysteines are located at the same place of E75-MtRpiB, the catalytic base of *M. tuberculosis*. Based on such informations and supposing that LmRpiB exhibits a similar reaction mechanism as EcRpiB, we considered the following protonation states of active

site residues of LmRpiB: (i) the charged C69 (C69:S^-) and (ii) the two states of H102, one neutral form protonated at NE2 (H102:NE2), and the positively charged form protonated at both ND1 and NE2 (H102:ND1/NE2).

Based on the superposition of 3K7S and 3HEE and redocking studies of 3K7S (results not shown), we defined a set of four water molecules inside the active site as conserved. These four water molecules were added to the final model of LmRpiB by considering the water molecules from 3K7S (HOH177:A, HOH187:B, HOH188:B and HOH194:B) as a reference.

For the HsRpiA model, the protonation states of residues E182 and D160 were defined based on the reaction mechanism proposed by Zhang et al., 2003 (ring opening step)[37] and Hamada et al., 2003 (isomerization step):[38] (i) the charged and neutral forms of D160 at OD2 and (ii) the negatively charged form of E182. A pKa analysis performed using the PROPKA 3.0 program[39] suggested the neutral form of E182, which is protonated at OE2 (E182:OE2) in the HsRpiA model. However, this protonation state of E182 is probably found only within the HsRpiA-intermediate complex. Because our docking studies were not performed with intermediate compounds, we maintained the charged state of E182 in all docking studies.

Analyzing the superposition of 1XTZ, 1LK7, 2F8M and 3HHE and the redocking studies of 3HHE, 3UW1 (R5PI type A from *Burkholderia thailandensis* – BtRpiA),[40] 1UJ5 (R5PI type A from *Thermus thermophilus* – TtRpiA)[38] and 1O8B (R5PI type A from *Escherichia coli* – EcRpiA)[37] (results not shown), we mapped five water molecules from 1O8B into our final model; these water molecules are considered responsible for the

phosphate stabilization and ring opening of the furanose form of R5P (according to reference 37). Water molecules were added to the final HsRpiA model based on the coordinates of the water molecules HOH2020:B, HOH2037:B, HOH2088:B, HOH2089:B and HOH2090:B from 1O8B.

3.2. Docking protocol:

The following substrates were obtained from the PubChem database for use in the docking studies:[41] (i) α-D-ribose-5-phosphate (furanose form - αR5P) (drawn based on CID439167); (ii) β-D-ribose-5-phosphate (furanose form - βR5P) (CID439167); (iii) D-ribose-5-phosphate (R5P) (CID77982) and (iv) D-ribulose-5-phosphate (5RP) (CID439184). These ligands were prepared using the LigPrep application included in the Glide program (version 5.8)[42] using the force field OPLS_2005. The 3D and 2D conformation of the ligands are visualized in Figure 2.

Molecular docking was performed using Glide. The LmRpiB and HsRpiA models were prepared for subsequent grid generation and docking using the Protein Preparation Wizard tool in Glide. In this preparatory step, all hydrogen atoms were added, the protonation states and tautomers of His residues were optimized, and Chi "flip" assignments for Asn, Gln and His residues were performed. All docking studies were performed in the Glide Extra Precision mode (XP) using the post-refinement step with the parameters set to the default values. The center of the grid (with an edge of 15 Å) was defined as: (i) LmRpiB - the centroid of the substrate R5P from chain A of the structure 3K7S aligned with the LmRpiB model (X = 27.47; Y = 5.03 and Z = 41.99) and (ii)

HsRpiA - the centroid of the inhibitor β-D-arabinose-5'-phosphate from chain B of the structure 1O8B aligned with the HsRpiA model (X = 18.36; Y = 11.22 and Z = 89.57). The highest-scoring docking poses were analyzed.

3.3. Electrostatic profile and volume:

To analyze overall and catalytic site differences between LmRpiB and HsRpiA, we compared their electrostatic profiles using the programs PDB2PQR version 1.2[43] and APBS version 1.2.1[44] (available at http://www.poissonboltzmann.org/), programs which implement numerical methods to enable the solution of the Poisson-Boltzmann equation for large systems. The CHARMM force field was applied, and a probe radius of 1.4 Å was used.

The volumes of the active sites were compared using the Pocket-Finder program (available at http://www.modelling.leeds.ac.uk/pocketfinder/). The residues considered for calculations of these properties in the active site were: (i) LmRpiB - L98, H102, N103, R137, H138, R141, D182, H183, A184, Y218, C241, G242, T243, G246, M247 and R285, and (ii) HsRpiA - S104, G105, S106, T107, I108, D160, G161, A162, D163, K173, G174, G175, G177, C178, E182, I195, K200 and K203.

4. Atomic coordinates of structural models

The atomic coordinates of the constructed models (in PDB file format) with the best ligand docking result are presented in the Supporting Information. The structures of LmRpiB complexed with βR5P, R5P and 5RP are presented in Files S1, S2 and S3, respectively. The structures of HsRpiA complexed with βR5P, R5P and 5RP are presented in Files S4, S5 and S6, respectively.

RESULTS AND DISCUSSION

1. Three-dimensional protein model construction

1.1. Overall Structures

The structure of RpiB from *Leishmania major* (LmRpiB) was constructed by comparative modeling using the RpiB molecules from *Trypanosoma cruzi* (TcRpiB) and *Clostridium thermocellum* (CtRpiB) as templates. Considering the high residue conservation between these proteins, the overall structure of the final LmRpiB model is very similar to that observed in the representative structures of RpiB from other organisms (Figure 3A). Ramachandran plot analysis of the final model showed that 98.8% of the amino acids are in favored regions, and 100% are in allowed regions. The global quality of the LmRpiB model calculated from normalized DOPE and QMEAN scores are 1.32 and -1.40, respectively.

Based on previous studies,[36,43] the model was generated as a dimer, with the catalytic cleft lying between the two subunits (Figure 4A). Each subunit of the protein comprises 172 amino acids, assuming a parallel β-sheet with five strands (β1-β5)

sandwiched by five α-helices, three on one side (α1-α3) and two on the other (α4-α5), followed by two other helices (α6 and α7 - the last was predicted using the GAPF program) (Figures 6B and 6C). LmRpiB presents a 3-layer(aba) sandwich architecture with a topology associated with CATH domain 3k7sA00 (CATH superfamily number 3.40.1400.10).

As previously reported, members of RpiA family were found in the following two oligomeric forms: (i) homodimers, commonly found in bacteria (*e.g., Escherichia coli, Haemophilus influenzae, Thermus thermophilus and Plasmodium falciparum*) and (ii) homotetramers, found in archaea and eukaryotes (*e.g., Pyrococcus horikoshii and Saccharomyces cerevisiae*). These differences were attributed to the "tetramerization loop", which is longer in members that form homotetramers.[26,27] The HsRpiA model features a large loop with 15 residues (from R240 to G254), which is very similar to that presented by homotetrameric representatives, suggesting that HsRpiA occurs as a homotetramer (Figure 4C).

Most of the 311 amino acids of the HsRpiA monomer exhibited the conformation previously described for other RpiA proteins,[28] except for the 77 N-terminal residues, for which no templates were available.[21,32] The general structure of each monomer consists of two α/β domains linked by an antiparallel β-sheet (S2) comprising three short strands (β5, β9, β12) (Figure 5).

Domain A is formed by a β-sheet (S1) with six parallel (β1-β4, β7, β13) and one antiparallel (β14) strands protected by four α-helices (α1-α3, α6) on one side and two α-

helices (α4 and α5) on the other. Domain B forms another β-sheet (S3) comprising four antiparallel strands (β9-β11) protected by two further α-helices (α7 and α8).

Between domains A and B, we found helices α4 and α5 from domain A and sheet S2. The region comprising domain A and link region S2 constitute a 3-layer(aba) sandwich architecture with a Rossmann fold topology. Domain B presents a 2-layer sandwich architecture with an alpha-beta plaits topology. These topologies are associated with CATH domains 1lk7A01 (CATH superfamily number 3.40.50.1360) and 1lk7A02 (CATH superfamily number 3.30.70.260), respectively.

The 77 N-terminal residues appear to form an additional domain that partially obstructs the active site; however, no interactions with catalytic residues were observed. Our attempts to model this region suggest that the region is very flexible and lies next to the catalytic cleft in the opposite side of the "tetramerization loop". Due to the low reliability of the generated models, the 77 N-terminal amino acids of HsRpiA were not included in the final structure.

Ramachandran plot analysis of the final model showed that 96.8% of the amino acids are in favored regions, and 100% are in allowed regions. The global quality measures of the HsRpiA model based on normalized DOPE and QMEAN scores are -0.72 and -0.58, respectively.

1.2. Active Site Characterization and Comparative Analysis

Two active site variants have been proposed for the RpiB family; one was initially described for *Escherichia coli* (EcRpiB),[45] and another was described for *Mycobacterium tuberculosis* (MtRpiB).[18] These variants differ mainly in the catalytic base responsible for extracting a proton from the substrate at C2 and releasing it to C1 (residue C66 for *E. coli* and D75 for *M. tuberculosis*).

The active site of RpiB from *Leishmania major* is similar to that of R5PI type B from *E. coli* (with C69 in LmRpiB playing the role of C66 in EcRpiB). A high structural similarity and sequence conservation is observed when comparing the active sites of EcRpiB and LmRpiB. All amino acids described as important for the active site of *E. coli* have an identical correspondent in *L. major,* suggesting that their catalytic mechanisms are similar.

In the LmRpiB model, the corresponding amino acids (H11:B, R113:B, R137:A and R141:A) that are responsible for stabilizing the phosphate group are well-oriented. The catalytic residues proposed in the literature[42,43] for EcRpiB and TcRpiB are also present in LmRpiB (Figure 5A): (i) H102, which catalyzes the opening of the furanose ring and (ii) T71 and C69, the pair of amino acids responsible for isomerization.[46] Finally, a water molecule proposed to participate in the opening of both α (water W_2) and β (water W_1) isomers of furanose ring[18] also exhibits a well-coordinated position (Figure 4B).

The active site of HsRpiA bears considerable resemblance to the active site of *S. cerevisiae* (ScRpiA). In fact, all key amino acids described in the catalytic cleft of ScRpiA (K22, S48, T49, K147) have an identical match in the sequence alignment with HsRpiA (K82, S106, T107, K200); these residues might interact with the phosphate group of the

substrate[27] (Figure 4D). It is important to note that the structural alignment of ScRpiA and the HsRpiA model suggests that K147 of ScRpiA could alternatively align with K203 instead of K200 of HsRpiA. These lysines reside in disordered loops which can explain the observed difference in the structural alignment relative to the sequence alignment. Another similarity between ScRpiA and the RpiA of *P. horikoshii* is the absence of R100 and T28 counterparts.[28] In HsRpiA, G175 and S104 occupy these positions.[27]

In EcRpiA, Zhang and collaborators[37] proposed that the components responsible for the furanose ring opening are a water molecule (proton donor) and D81 and that the isomerization step is carried out by D81 and E103. Later, Roos and collaborators[36] proposed that glutamic acid (E103 in EcRpiA and E108 in TtRpiA) catalyses all proton transfers required for substrate isomerization. These residues are widely conserved in RpiA proteins, and their relevance was tested in the RpiA of spinach.[37] In HsRpiA, the corresponding catalytic residues are D160 and E182, and the water molecule (W_1) appears to be relevant for the catalytic mechanism (Figures 7D and 8A).

Despite these differences in the active site structures, apparently both HsRpiA and LmRpiB catalyze the isomerization reaction via the intermediate 1,2-*cis*-enediolate. Aiming to analyze further differences between the catalytic sites of LmRpiB and HsRpiA, we compared their electrostatic profiles and volumes of the catalytic sites.

Both enzymes exhibited an electrostatic profile characteristic of a positively charged environment, with HsRpiA (Figure 6D) exhibiting an electrostatic potential more pronounced than that of LmRpiB (Figure 6A). Conversely, at the catalytic site, LmRpiB (Figures 6B and 6C) presents a more positive electrostatic profile than HsRpiA (Figures 6E

and 6F). The positive charge on both catalytic sites could be critical for guiding the negatively charged substrates. The total electrostatic potential values returned by APBS are presented in Table S2. The volumes calculated also presents some differences; whereas the catalytic site of HsRpiA presents 164Å^3 in the ring opening step and 182Å^3 in the isomerization step, the catalytic site of LmRpiB presents larger cavities of 219Å^3 and 222 Å^3, respectively.

2. Docking of substrates: Investigation of binding mechanisms

To assist in the interpretation of the importance of each amino acid in the LmRpiB and HsRpiA active sites, the models were subjected to docking analyses. The substrates used were the α and β isomers of D-ribose-5-phosphate (furanose form, αR5P and βR5P), D-ribose-5-phosphate (R5P) and D-ribulose-5-phosphate (5RP). Docking experiments were performed using Glide XP, and the top ranked conformation was selected according to the XP Rank.[36] The best results for each substrate are presented for LmRpiB in Figure 7 and for HsRpiA in Figure 8. Comparative analyses of XP GScores[42], hydrogen bonds and important contacts are presented in Table 1 and Table S1 (Supporting Information), respectively.

According to our docking results, we suggest that the protonated H102 of LmRpiB opens the furanose ring of βR5P by donating a H^+ to O4 while a water molecule removes the H^+ from O1 (Figures 7A and 7D). Initially, Roos and collaborators[50] proposed that H138 of *M. tuberculosis* was responsible for removing the H^+ from O1, but in 2008,[36] they proposed that a water molecule could play this role for the β-furanose form but that the

participation of this water molecule was not clear for the α-furanose form. In our results, the water molecule W_1 is 3.736Å from O1 of βR5P, and the water molecule W_4 is 3.063Å from O1 of αR5P (Table S1); therefore, different water molecules appear to remove the proton from O1 in each isomer.

In accordance with Zhang et al.[45] and Roos et al.,[36] we propose that the isomerization step in LmRpiB is initiated when the negatively charged C69:S^- acts as a base and accepts a proton from C2 while T71 transfers the H^+ from O2 to O1, forming the enediolate intermediate state. The result of R5P docking (Figures 7B and 7E and Table S1) shows that the negatively charged C69(SG):B atom is 4.271Å and 3.404Å from C1 and C2, respectively, whereas T71(OG1):B is 3.769Å and 3.472Å from O1 and O2, respectively. Additionally, N103(ND2) makes a hydrogen bond with O1, probably assisting with stabilization of the enediolate intermediate. The NH groups of the residues from G70 to G74 appear to help in stabilizing the negative charge of the intermediate (as an oxyanion hole). The final step of isomerization is performed by C69, which returns the proton to C1, thereby forming 5RP (Figures 7C and 7F).

The docking results obtained for HsRpiA suggest that D160(OD2) acts as a base by removing one proton from O1 of βR5P (2.795Å), whereas the water molecule W_1 (in its hydronium form) donates a H^+ to O4 (Figures 8A and 8D). The result of αR5P docking suggests that E182(OE2) catalyzes the opening of the furanose ring removing the H^+ from O1 (2.905Å) (Table S1). According to Zhang and collaborators,[37] a water molecule (probably water 2088 in 1O8B) participates in furanose ring opening in EcRpiA by donating a proton to the equivalent O4 atom of R5P.

Based on results obtained with arabinose-5-phosphate, Zhang et al.[37] suggested for EcRpiA that E103/D81 (E182/D160 of HsRpiA) act in the isomerization step, with E103 acting as the catalytic base and D81 being responsible for proton transfer. However, Hamada et al.[38] proposed another mechanism based on the RpiA structure of *T. thermophilus* (1UJ5) complexed with R5P (in ring-opened form). The authors suggested that E108 (E182 of HsRpiA) was responsible for all steps in the isomerization process. Our docking results for D-ribose (Figures 8B and 8E), in which D160 is neutral because it has accepted a proton in the ring opening step, reproduces the conformation found in *B. thailandensis* (3UW1) and *T. thermophilus* (1UJ5), suggesting that the mechanism proposed by Hamada et al.[38] might also apply in HsRpiA.

In accordance with Zhang et al.[37] and Hamada et al.,[38] we propose that the negatively charged form of E182 of HsRpiA acts as a base by abstracting a proton from C2, thereby forming the enediolate intermediate, which is stabilized by the NH-rich region from G174 to C178 (the oxyanion hole). Later, the neutral form of E182 transfers the H^+ to C1. At this time, the O2 becomes highly polarized due to the formation of a hydrogen bond with K173. Considering the short distances between O1 and O2 of 5RP to E182 (2.817Å and 2.722Å, respectively), this residue probably transfers the H^+ from O2 to O1, finally forming 5RP (Figures 8C and 8F and Table S1).

A comparison of the catalytic mechanisms proposed in this work for RpiB of *Leishmania major* and for RpiA of *Homo sapiens* is presented in Figure 9.

CONCLUSIONS

By analyzing the *Leishmania major* genome using the AnEnPi computational tool, we were able to detect the presence of the RpiB enzyme, but the RpiA appeared to be completely absent. Considering the biological importance of RpiB given its protective role against oxidative stress in parasites,[15] it is natural to consider this enzyme as an important target for the development of new anti-*Leishmania* drugs. In this context, previous studies have shown that double mutants of *E. coli* (*rpiA*$^-$ *rpiB*$^-$) are not viable[51], demonstrating the importance of the RpiB enzyme for bacterial survival.

A structural comparative analysis of LmRpiB and the analogous RpiA form from *Homo sapiens* (HsRpiA) was performed to obtain information that is important for future drug design studies. In the absence of experimental structures, we constructed 3D models of both enzymes by applying a hybrid computational methodology combining comparative and *ab initio* modeling techniques. Unfortunately, we were not able to model the 77 N-terminal residues of HsRpiA appropriately; nevertheless, our tentative models suggest that this region is very flexible and might occlude the active site.

The characterization of structural analogy was performed by analyzing overall and the active sites differences in the 3D structures, electrostatic potential profiles and volumes. Although both enzymes catalyze the same reaction, their overall structure is very different. In fact, even their biological assemblies are different: both enzymes appear to form a homomultimer, but in the enzymes of type B the multimerization is essential to form the active site, since it is formed between the two subunits (Figures 3 and 7). This difference alone could be enough to describe an important regulatory mechanism: if the monomer concentration decays to rates below of the aggregation (multimerization) threshold, the

activity would be compromised. So, any compound that inhibit the formation of the homodimer would also be effective at inactivating the enzyme.

The different composition of residues in LmRpiB and HsRpiA active sites is another key feature that could be exploited to develop a type-specific inhibitor. When Essenberg & Cooper (1975)[52] first described the different types of R5PI in *E. coli*, they used the treatment with iodoacetate to show that the residue composition of the ligand binding sites could be used to specifically inhibit the type B enzyme. This effect was attributed to the inactivation of the catalytic cysteine (equivalent to the C69 in LmRpiB).

In this work, LmRpiB and HsRpiA enzymes were very carefully analyzed with respect to the protonation state of important catalytic residues and the importance of some conserved water molecules within the active site. We observed that distinct residue types participate in the catalytic reaction in the two enzymes. Additionally, some residues confer different physicochemical features that could be exploited to design type-specific inhibitor.

The active site of LmRpiB shows many aromatic residues (Y46, H11, H102 and H138), whereas HsRpiA shows none. These features could favor pi-stacking interactions with a candidate inhibitor. The profile of negatively charged residues are also different, LmRpiB has two residues (C69 and D10) while HsRpiA has three (D163, D160 and E182). Most of the positively charged residues in both enzymes are close to the phosphate group of the substrates in the PDB structures and in our docking results, whereas the residue K173 of HsRpiA is located near catalytic residues D160 and E182, probably contributing to the reaction mechanism. Both enzymes exhibit an electrostatic profile characteristic of a positively charged environment, with HsRpiA exhibiting an electrostatic potential of

overall structure more pronounced than that of LmRpiB, although at the catalytic site the LmRpiB presents electrostatic potential values more positive than those observed for HsRpiA. Moreover, the LmRpiB enzyme exhibits a slightly larger cavity than that of HsRpiA.

This characterization of the active sites is important, not only because it emphasizes structural differences between these analogous enzymes, but also because it indicates important directions to be explored in specific drug design projects. In this context, several reports[37,45] have shown differences in the kinetics of the RpiA and RpiB enzymes of *E. coli* and different susceptibility to inhibitors, such as the previously described iodoacetate, which only inhibits RpiB.[52]

An important aspect of this work was the docking studies using the substrates R5P (furanose and ring-opened forms) and 5RP; these studies not only validate the characterization of the active site residues, but also allow additional insights into the enzymatic reaction mechanisms. Based on these docking experiments and previous proposals discussed in the literature, we propose two distinct mechanisms for the reversible isomerization of R5P to 5RP that is catalyzed by LmRpiB and HsRpiA.

Knowledge of the correct residues and their protonation forms, and characterization of the catalytic reaction are key factors in performing successful docking and structure-based drug design studies. We expect that the results presented in this paper (particularly the constructed structures, which are available as part of the Supporting Information) might be important aids to the development of new anti-Leishmania drugs that are designed to inhibit the parasite enzyme without significantly affecting the analogous human protein.

REFERENCES

(1) Centers for Disease Control and Prevention (CDC): Parasites – Leishmaniasis. Available: http://www.cdc.gov/parasites/leishmaniasis/index.html. Accessed 26 March **2013**.

(2) World Health Organisation (WHO): Leishmaniasis. Available: http://www.who.int/leishmaniasis/en/. Accessed 26 March **2013**.

(3) den Boer, M.; Argaw, D.; Jannin, J.; Alvar, J. Leishmaniasis impact and treatment access. *Clin Microbiol Infect* **2011**, 17(10): 1471-1477 .

(4) World Health Organisation (WHO). Control of leishmaniasis. Sixtieth World Health Assembly A60/10, **2007**.

(5) Bern, C.; Maguire, J. H.; Alvar, J. Complexities of Assessing the Disease Burden Attributable to Leishmaniasis. *PLoS Negl Trop Dis* **2008**, 2(10): e313. doi:10.1371/journal.pntd.0000313.

(6) Ouellette, M.; Drummelsmith, J.; Papadopoulou, B. Leishmaniasis: Drugs in the clinic, resistance and new developments. *Drug Resist Updat* **2004**, 7: 257-266.

(7) Costa, C. H. N.; Peters, N. C.; Maruyama, S. R.; de Brito, E. C. Jr.; de Miranda Santos, I. K. F. Vaccines for the Leishmaniases: Proposals for a Research Agenda. *PLoS Negl Trop Dis* **2011**, 5(3): e943. doi:10.1371/journal.pntd.0000943.

(8) Griensven, J.; Balasegaram, M.; Meheus, F.; Alvar, J.; Lynen, L.; et al. Combination

therapy for visceral leishmaniasis. *Lancet Infect Dis* **2010**, 10(3): 184-194.

(9) Singh, B.; Sarkar, N.; Jagannadham, M.; Dubey, V. Modeled structure of trypanothione reductase of *Leishmania infantum*. *BMB Rep* **2008**, 41(6): 444-447.

(10) Liang, X.; Haritan, A.; Uliel, S.; Michaeli, S. *Trans* and *cis* splicing in Trypanosomatids: mechanism, factors and regulation. *Eukaryot Cell* **2003**, 2: 830-840.

(11) Capriles, P. V.; Guimaraes, A. C.; Otto, T. D.; Miranda, A. B.; Dardenne, L. E.; et al. Structural modelling and comparative analysis of homologous, analogous and specific proteins from *Trypanosoma cruzi* versus *Homo sapiens*: Putative drug targets for Chagas' disease treatment. *BMC Genomics* **2010**, 11: 610.

(12) Otto, T.; Guimarães, A.; Degrave, W.; Miranda, A. AnEnPi: identification and annotation of analogous enzymes. *BMC Bioinformatics* **2008**, 9: 87-95.

(13) Kanehisa, M.; Goto, S.; Hattori, M.; Aoki-Hinoshita, K.; Itoh, M.; et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* **2006**, 34(D): 354–357.

(14) Alves-Ferreira, M.; Guimarães, A. C. R.; Capriles, P. V. S. Z.; Dardenne, L. E.; Degrave, W. M. A new approach for a potential drug target discovery through *in silico* metabolic pathway analysis using *Trypanosoma cruzi* genome information. *Mem Inst Oswaldo Cruz* **2009**, 104(8): 1100–1110.

(15) Maugeri, D.; Cazzulo, J.; Burchmore, R.; Barret, M.; Ogbunude, P. Pentose phosphate metabolism in *Leishmania mexicana*. *Mol Biochem Parasitol* **2003**, 130: 117-125.

(16) Wiesmeyer, H.; David, J. Regulation of ribose metabolism in *Escherichia coli*: I. The ribose catabolic pathway. *Biochim Biophys Acta* **1970**, 208: 45-55.

(17) Kim, C.; Song, S.; Park, C. The D-allose operon of *Escherichia coli* K-12. *J Bacteriol* **1997**, 179: 7631-7637.

(18) Ross, A.; Mariano, S.; Kowalinski, E.; Salmon, L.; Mowbray, S. D-ribose 5-phosphate isomerase B from *Escherichia coli* is also a functional D-allose 6-phosphate isomerase, while the *Mycobacterium tuberculosis* enzyme is not. *J Mol Biol* **2008**, 382: 667-679.

(19) Stanley, L. S.; Doak, L. S. Studies of the metabolism of human erythrocyte membranes. *J Clin Invest* **1963**, 42(6): 756-766.

(20) Spencer, N.; Hopkinson, D. Biochemical genetics of the pentose phosphate cycle: human ribose 5-phosphate isomerase (RPI) and ribulose 5-phosphate 3-epimerase (RPE). *Ann Hum Genet* **1980**, 43(4): 335-342.

(21) Sali, A.; Blundell, T. Comparative protein modeling by satisfaction of spatial restraints. *J Mol Biol* **1993**, 234: 779-815.

(22) Custódio, F. L.; Barbosa, H. J. C.; Dardenne, L. E. A Multiple Minima Genetic Algorithm for Protein Structure Prediction. *Appl Soft Comput* **2014**, 15: 88-99.

(23) Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; et al. The Protein Data Bank. *Nucleic Acids Res* **2008**, 28: 235-242.

(24) Altschul, S.; Gish, W.; Miller, W.; Myers, E.; Lipman, D. Basic local alignment search tool. *J Mol Biol* **1990**, 215: 403-410.

(25) Stern, A.; Naworyta, A.; Cazzulo, J.; Mowbray, S. Structures of type B ribose 5-phosphate isomerase from *Trypanosoma cruzi* shed light on the determinants of sugar specificity in the structural family. *FEBS J* **2011**, 278: 793-808.

(26) Jung, J.; Kim, J.; Yeom, S.; Ahn, Y.; Oh, D.; et al. Crystal structure of

Clostridium thermocellum ribose-5-phosphate isomerase B reveals properties critical for fast enzyme kinetics. *Appl Microbiol Biotechnol* **2011**, 2: 517-527.

(27) Graille, M.; Meyer, P.; Leulliot, N.; Sorel, I.; Janin, J.; et al. Crystal structure of the S. cerevisiae D-ribose-5-phosphate isomerase: comparison with the archaeal and bacterial enzymes. *Biochimie* **2005**, 87: 763-769.

(28) Ishikawa, K.; Matsui, I.; Payan, F.; Cambillau, C.; Ishida, H.; et al. A Hyperthermostable D-Ribose-5-Phosphate Isomerase from *Pyrococcus horikoshii* Characterization and Three-Dimensional Structure. *Structure* **2002**, 10(6): 877-886.

(29) Holmes, M.; Buckner, F.; Van Voorhis, W.; Verlinde, C.; Mehlin, C.; et al. Structure of ribose 5-phosphate isomerase from Plasmodium falciparum. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2006**, 62: 427-431.

(30) Edwards, T. E.; Abendroth, J. A.; Staker, B. L.; Seattle Structural Genomics Center for Infectious Disease. Crystal structure of ribose-5-phosphate isomerase A from *Bartonella henselae*. "In press".

(31) Petersen, T. N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **2011**, 8: 785-786.

(32) Leaver-Fay, A.; Tyka, M.; Lewis, S.; Lange, O.; Thompson, J.; et al.. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* **2011**, 487: 545-74.

(33) Laskowski, R.; MacArthur, M.; Moss, D.; Thornton, J. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **1993**, 26: 283-291.

(34) Chen, V.; Arendall-III, W.; Headd, J.; Keedy, D.; Immormino, R.; et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010**, D66: 12-21.

(35) Benkert, P.; Künzli, M.; Schwede, T. QMEAN server for protein model quality estimation. *Nucleic Acids Res* **2009**, 37(Web Server issue): W510–W514.

(36) Roos, A. K.; Mariano, S.; Kowalinski, E.; Salmon, L.; Mowbray, S. L. D-ribose-5-phosphate isomerase B from *Escherichia coli* is also a functional D-allose-6-phosphate isomerase, while the *Mycobacterium tuberculosis* enzyme is not. *J Mol Biol* **2008**, 382: 667-679.

(37) Zhang, R-G. Andersson, C. E.; Savchenko, A.; Skarina, T.; Evdokimova, E.; et al.. Structure of *Escherichia coli* Ribose-5-Phosphate Isomerase: A Ubiquitous Enzyme of the Pentose Phosphate Pathway and the Calvin Cycle. *Structure* **2003**, 11(1): 31–42.

(38) Hamada, K.; Ago, H.; Sugahara, M.; Nodake, Y.; Kuramitsu, S.; et al. Oxyanion Hole-stabilized Stereospecific Isomerization in Ribose-5-phosphate Isomerase (Rpi)*. *J Biol Chem* **2003**, 278(49): 49183-49190.

(39) Delphine, C. B.; David, M. R.; Jan, H. J. Very Fast Prediction and Rationalization of pKa Values for Protein-Ligand Complexes. *Proteins* **2008**, 73: 765-783.

(40) Craig, T. K.; Gardberg, A.; Staker, B.; Stewart, L.; Seattle Structural Genomics Center for Infectious Disease. Crystal structure of ribose-5-phosphate isomerase A from *Burkholderia thailandensis* with ribose-5-phosphate. "In press".

(41) Sayers, E.; Barrett, T.; Benson, D.; Bolton, E.; Stephen, H.; et al. Database resources of the National Center for Biotechnology Information Export. *Nucleic Acids Res* **2010**,

36(Database issue): D13–D21.

(42) Friesner, R.; Murphy, R.; Repasky, M.; Frye, L.; Greenwood, J.; et al. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J Med Chem* **2006**, 49 (21): 6177-6196.

(43) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; et al. PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* **2007**, 35: W522-W525.

(44) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci* **2001**, 98: 10037-10041.

(45) Zhang, R.; Andersson, C.; Skarina, T.; Evdokimova, E.; Edwards, A.; et al.. The 2.2 A resolution structure of RpiB/AlsB from Escherichia coli illustrates a new approach to the ribose-5-phosphate isomerase reaction. *J Mol Biol* **2003**, 332: 1083-1094.

(46) Stern, A.; Burgos, E.; Salmon, L.; Cazzulo, J. Ribose 5-phosphate isomerase type B from *Trypanosoma cruzi*: kinetic properties and site-directed mutagenesis reveal information about the reaction mechanism. *Biochem J* **2007**, 401: 279-285.

(47) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, 23: 2947-2948.

(48) Frishman, D.; Argos, P. Knowledge-Based Protein Secondary Structure Assignment. *Proteins* **1995**, 23: 566-579.

(49) Laskowski, R. A. PDBsum new things. *Nucleic Acids Res* **2009**, 37 (Database issue): D355-D359.

(50) Roos, A. K.; Burgos, E.; Ericsson, D. J.; Salmon, L.; Mowbray, S. L. Competitive Inhibitors of *Mycobacterium tuberculosis* Ribose-5-phosphate Isomerase B Reveal New Information about the Reaction Mechanism. *J Biol Chem* **2005**, 280(8): 6416-6422.

(51) Sorensen, K.; Hove-Jensen, B. Ribose catabolism of *Escherichia coli*: characterization of the *rpiB* gene encoding ribose phosphate isomerase B and of the *rpiR* gene, which is involved in regulation of *rpiB* expression. *J Bacteriol* **1996**, 178: 1003-1011.

(52) Essenberg, M.; Cooper R. Two ribose 5-phosphate isomerases from *Escherichia coli* K12: partial characterization of the enzymes and consideration of their possible physiological roles. *Eur J Biochem* **1975**, 55: 323-332.

(53) Maugeri, D.; Cazzulo, J. The pentose phosphate pathway in *Trypanosoma cruzi*. *FEMS Microbiol Lett* **2004**, 234: 117-123.

AUTHOR INFORMATION

Corresponding authors

* E-mail: Priscila Capriles (capriles@lncc.br) and Laurent Dardenne (dardenne@lncc.br)

Author Contributions

PVSZC, LED and ACRG conceived and designed the experiments. PVSZC, LPRB, IAG, FLC and ACRG performed the experiments. PVSZC, IAG, ACRG and MAF analyzed the data. PVSZC, LPRB, IAG, LED and MAF wrote the paper. All authors read and approved

the final manuscript. ‡ These authors contributed equally to this work and should be considered co-first authors.

Funding Sources

The authors thank the Brazilian National Council of Research (CNPq), the FAPERJ Foundation and the Coordination for Enhancement of Higher Education Personnel (Capes) for supporting this work. Contract grants: N. 307062/2010-4, CNPq/MS- SCTIE-DECIT 41.0544/2006-0, CNPq/MS-SCTIE-DECIT 409078/2006-9, CNPq/ MCT 15/2007-Universal.

Notes

The authors declare no competing financial interests.

FIGURES

Figure 1. Reaction mechanisms for the reversible isomerization of R5P and 5RP, assuming the R5P as a starting point and the formation of the 1,2-cis-enediol(ate) intermediate. The isomerization reaction can be catalyzed in both directions depending on the substrate concentration.

Figure 2. The 3D and 2D conformation of the ligands used in docking analysis. (A) α-D-ribose-5-phosphate (furanose form - αR5P) (drawn based on CID439167); (B) β-D-ribose-5-phosphate (furanose form - βR5P) (CID439167); (C) D-ribose-5-phosphate (R5P) (CID77982) and (D) D-ribulose-5-phosphate (5RP) (CID439184).

Figure 3. Analysis of the 3D structure of LmRpiB model. (A) Primary and secondary structural alignment between LmRpiB enzyme (LmjF28.1970) and the selected templates: (i) 3K7S[25] (TcRpiB - identity: 51%); (ii) 3HEE[26] (CtRpiB - identity: 45%) and (iii) GAPF25 (fragment generated by GAPF program). The catalytic residues are highlighted in gray, according to references 37 and 44. Strands (templates are in yellow and *L. major* are in blue) are labeled from β1 to β4 and helices (templates are in pink and *L. major* are in red) are labeled from α1 to α7. The sequence alignment was performed using ClustalX (v2.1)[47] and secondary structure description were obtained via Stride Web Interface.[48] (B) Cartoon representation of LmRpiB (monomer A). The label of each secondary structure corresponds to that presented in Figure 3A. The image was generated using PyMOL, version 1.4.1, Schrödinger, LLC. (C) Topology diagram of the monomer presented in

Figure 3B. The diagram was constructed using the program PDBSum.[49]

Figure 4. The 3D structural comparison between the LmRpiB and HsRpiA models and their templates. (A) Structural alignment between LmRpiB model (yellow) and the templates 3K7S[25] (TcRpiB - orange) and 3HEE[26] (CtRpiB - green). 3D structures are presented in Cartoon, ligands (R5P) and water molecules are presented in Licorice, and the active site are marked with a black circle. (B) The active site of LmRpiB, TcRpiB and CtRpiB are presented in detail. Residue numbering according to LmRpiB index. The conserved water molecules (W_{1-4}) from TcRpiB are presented in orange (HOH177:A, HOH187:B, HOH188:B and HOH194:B), and from CtRpiB in green (HOH201:A, HOH155:B, HOH183:B and HOH184:B). (C) Structural alignment between HsRpiA model (yellow) and the templates 1XTZ[27] (ScRpiA- orange), 1LK7[28] (PhRpiA - green), 2F8M[29] (PfRpiA - blue) and 3HHE[30] (BhRpiA - cian). 3D structures are presented in Cartoon, ligands (5RP) and water molecules are presented in Licorice. The active site are marked with a black circle and the tetramerization loop with a black arrow. (D) The active site of HsRpiA, ScRpiA, PhRpiA, PfRpiA and BhRpiA are presented in detail. Residue numbering according to HsRpiA index, except the residue D99* (green) from PhRpiA. Water molecules (W_{1-5}) considered responsible for the phosphate stabilization and ring opening of the furanose form of R5P (according to reference 37) are displayed in red for 1O8B[37] (EcRpiA - HOH2088:B, HOH2090:B, HOH2020:B, HOH2089:B and HOH2037:B), in purple for 1UJ5[38] (TtRpiA - HOH403:A, HOH434:A, HOH448:A and HOH405:A) and in black for 3HHE[40] (BtRpiA - HOH257:A, HOH283:A, HOH287:A and HOH248:A). Structural alignments and images were generated using PyMOL, version

1.4.1, Schrödinger - LLC.

Figure 5. Analysis of the 3D structure of HsRpiA model. (A) Primary and secondary structural alignment between HsRpiA enzyme (Hsa:22934) and the selected templates: (i) 1XTZ[27] (ScRpiA - identity: 46%); (ii) 1LK7[28] (PhRpiA - identity: 42%); (iii) 2F8M[29] (PfRpiA - identity: 35%) and (iv) 3HHE[30] (BhRpiA - identity: 39%). The catalytic residues are highlighted in gray, according to reference 37. Strands from templates are colored in yellow and strands S1 (β1-β4, β7, β13, β14), S2 (β5, β6, β12) and S3 (β8-β11) from *H. sapiens* are respectively colored in blue, orange and green. The helices from templates are in pink and in *H. sapiens* the helices from domain A (α1 to α6) are in red and from domain B (α7 and α8) are in purple. The sequence alignment was performed using ClustalX (v2.1)[47] and the secondary structure description was obtained via Stride Web Interface.[48] (B) Cartoon representation of HsRpiA (monomer A). The label of each secondary structure corresponds to that presented in Figure 5A. The image was generated using PyMOL, version 1.4.1, Schrödinger, LLC. (C) Topology diagram of the monomer presented in Figure 5B. The diagram was constructed using the program PDBSum.[49]

Figure 6: Electrostatic potential profile of LmRpiB and HsRpiA models. (A) The electrostatic potential profile of LmRpiB. The active site is marked with a black circle. The electrostatic potential profile of the catalytic site of LmRpiB in the ring opening and isomerization steps are presented in detail in Figures (B) and (C), respectively. (D) The

electrostatic potential profile of HsRpiA. The active site is marked with a black circle. The electrostatic potential profile of the catalytic site of LmRpiB in the ring opening and isomerization steps are presented in detail in Figures (E) and (F), respectively. The substrates βR5P (9B and 9E) and R5P (9C and 9F) are represented as green carbon sticks and 5RP (9C and 9F) as gray carbon sticks. The water molecules are presented as red spheres. The electrostatic potential vary from -5 (red) to 5 (blue) kT/e. The images were generated using PyMOL, version 1.4.1, Schrödinger - LLC.

Figure 7. Docking results of the substrates βR5P, R5P and 5RP into the catalytic site of LmRpiB. 3D visualization of docking analysis with substrates: (A) β-D-ribose-5-phosphate (βR5P) (CID439167); (B) D-ribose-5-phosphate (R5P) (CID77982) and (C) D-ribulose-5-phosphate (5RP) (CID439184). Protein structure is presented in Cartoon (monomer A in yellow and monomer B in blue). Substrates are represented as green carbon sticks and R5P from 3K7S[25] is in gray carbon sticks. The water molecules are presented as red spheres. Residue numbering according to LmRpiB index. All referenced distances are in angstrom (Å) units. (D-F) 2D visualization of results presented in Figures 7A-7C. Filled and dashed arrows represent hydrogen bond with backbone and side chain, respectively. Residues are colored in accordance with the following properties: negatively charged (pink), positively charged (purple), polar (blue), hydrophobic (green), glycine (yellow), C69:S^- (CYT69:B - dark gray), H102:ND1/NE2 (HIP102:A - purple) and water (light gray). The 3D and 2D images were generated using PyMOL v.1.4.1 and Maestro v.9.2, both from Schrödinger - LLC, respectively.

Figure 8. Docking results of the substrates βR5P, R5P and 5RP into the catalytic site of HsRpiA. 3D visualization of docking analysis with substrates: (A) β-D-ribose-5-phosphate (βR5P) (CID439167); (B) D-ribose-5-phosphate (R5P) (CID77982); (C) D-ribulose-5-phosphate (5RP) (CID439184). Protein structure of monomer A is presented in Cartoon and residues (G174-C178) that participate of the oxyanion role are represented by stars enumerated from 1 to 5 (blue). Substrates are represented as green carbon sticks and βABF from 1O8B[37], R5P from 3UW1[40] and 5RP from 1UJ5[38] are in gray carbon sticks. The water molecules are presented as red spheres. Residue numbering according to HsRpiA index. All referenced distances are in angstrom (Å) units. (D-F) 2D visualization of results presented in Figures 8A-8C. Filled and dashed arrows represent hydrogen bond with backbone and side chain, respectively. Residues are colored in accordance with the following properties: negatively charged (pink), positively charged (purple), polar (blue), hydrophobic (green), glycine (yellow), D160:OH (ASH160 - dark gray) and water (light gray). The 3D and 2D images were generated using PyMOL v.1.4.1 and Maestro v.9.2, both from Schrödinger – LLC, respectively.

Figure 9. Reaction mechanisms proposed in this work for the reversible isomerization of R5P to 5RP catalyzed by LmRpiB and HsRpiA. The amino acid residues and water molecules suggested in this work to be directly involved in the catalytic mechanism are colored in blue (LmRpiB) and red (HsRpiA). Initially, one proton is donated for the O4 atom of the furanose form and a basic molecule (W_1 of LmRpiB and D160 of HsRpiA)

removes the H^+ from O1, opening the ring and forming the linear form of D-ribose-5-phosphate (R5P). The isomerization step involves the formation of 1,2-cis-enediol(ate) intermediate and consists of two proton transfers. First, the negatively charged residue (C69 of LmRpiB and E182 of HsRpiA) acts as a base abstracting a proton from C2 followed by a transfer of H+ to C1 (T71 of LmRpiB and E182 of HsRpiA), forming the enediolate intermediate state, which is stabilized by the NH-rich region (from G70 to G74 of LmRpiB and from G174 to C178 of HsRpiA). The final step of isomerization is performed by C69 of LmRpiB and E182 of HsRpiA, transfering the H^+ from O2 to O1, finally forming D-ribulose-5-phosphate (5RP).

TABLES

Table 1: XP GScore of substrates docked into LmRpiB and HsRpiA.

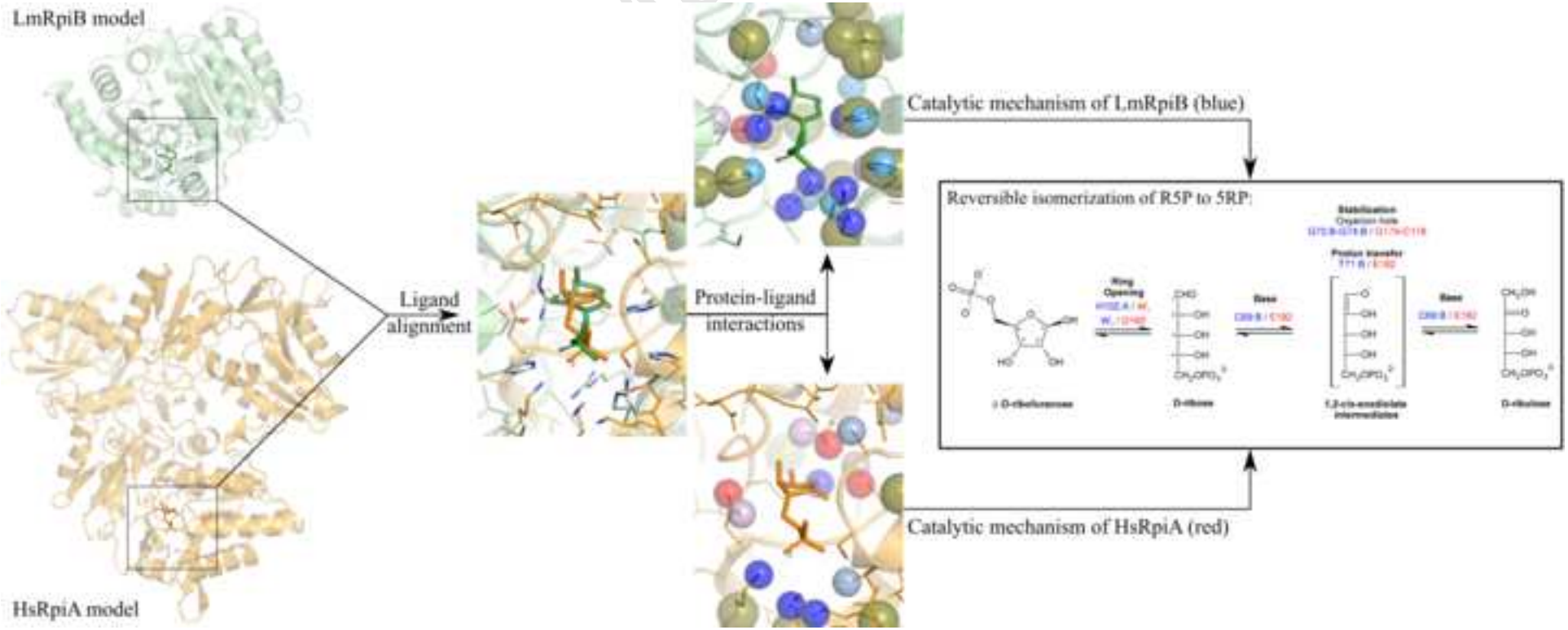
Molecular docking was performed using Glide XP and the top ranked conformation was selected according to XP Rank.[42]

Table 1: XP GScore of substrates docked into LmRpiB and HsRpiA.

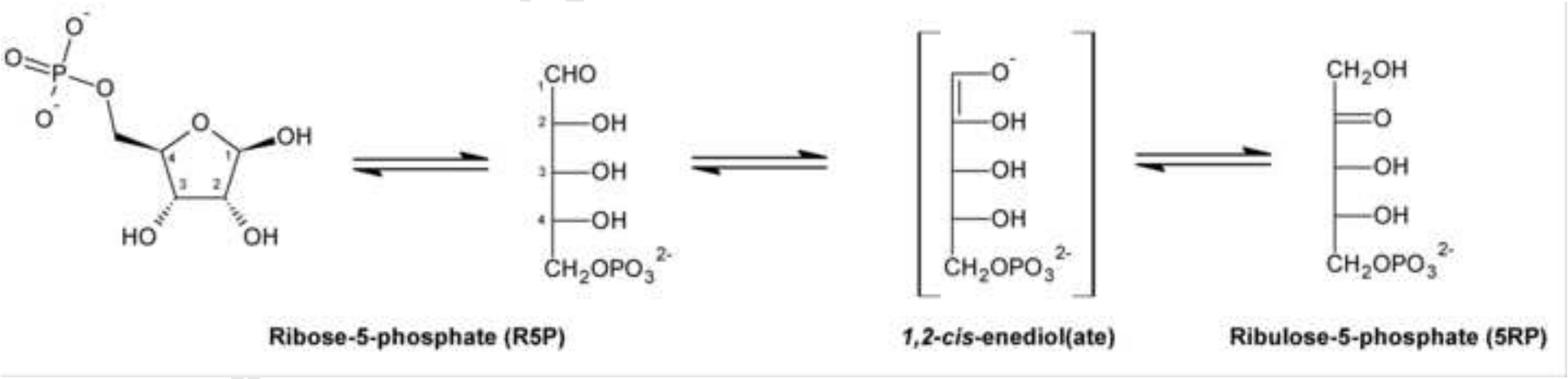
Substrate	XP GScore	
	Leishmania major	***Homo sapiens***
α-D-ribose-5-phosphate (αR5P)	-7.898	-9.177
β-D-ribose-5-phosphate (βR5P)	-7.954	-8.443
D-ribose-5-phosphate (R5P)	-10.460	-9.073
D-ribulose-5-phosphate (5RP)	-8.258	-9.572

Molecular docking was performed using Glide XP and the top ranked conformation was selected according to XP Rank.[42]

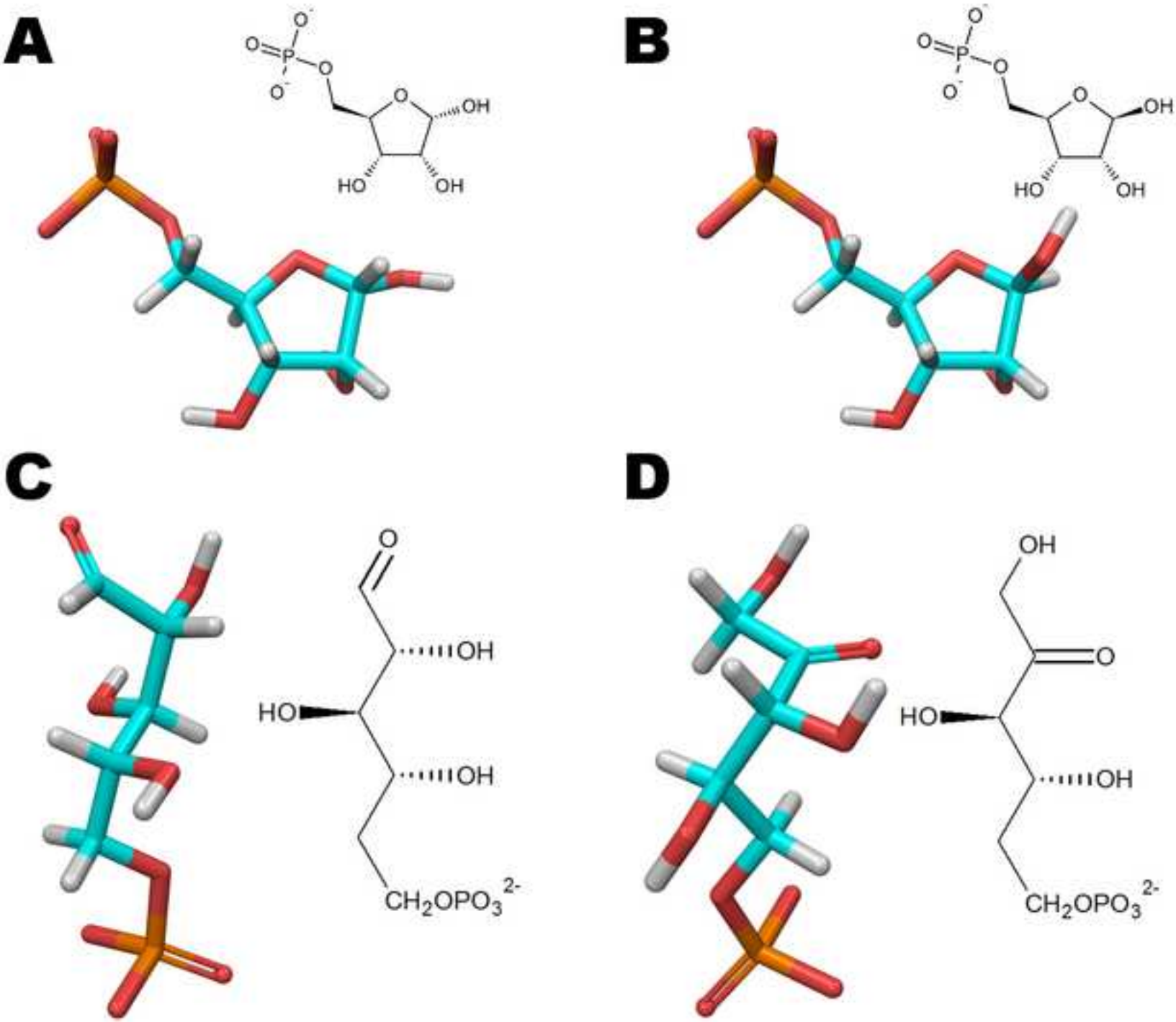
Graphical Abstract (for review)

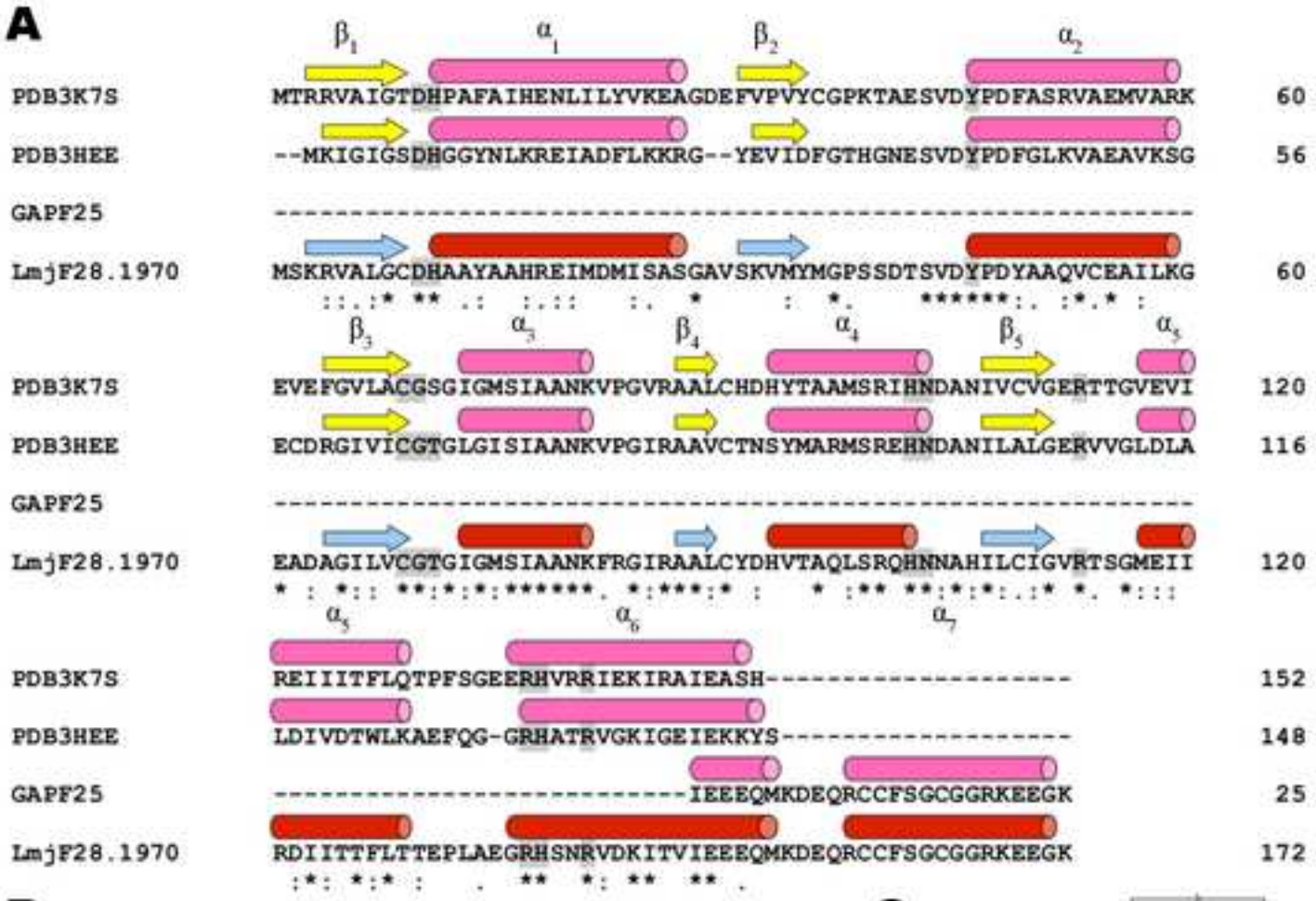


Figure(s)

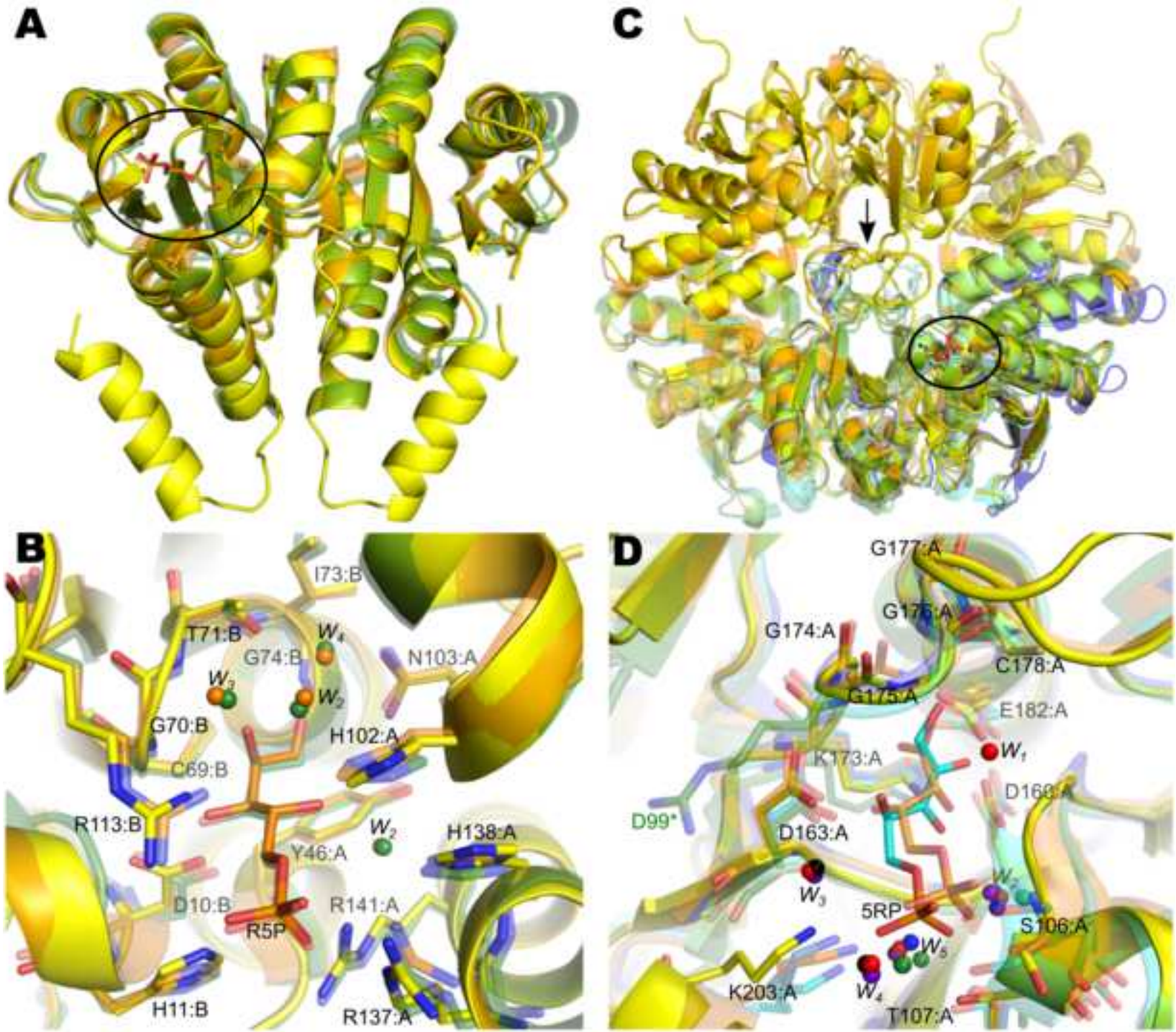

Ribose-5-phosphate (R5P)
1,2-cis-enediol(ate)
Ribulose-5-phosphate (5RP)

Figure(s)

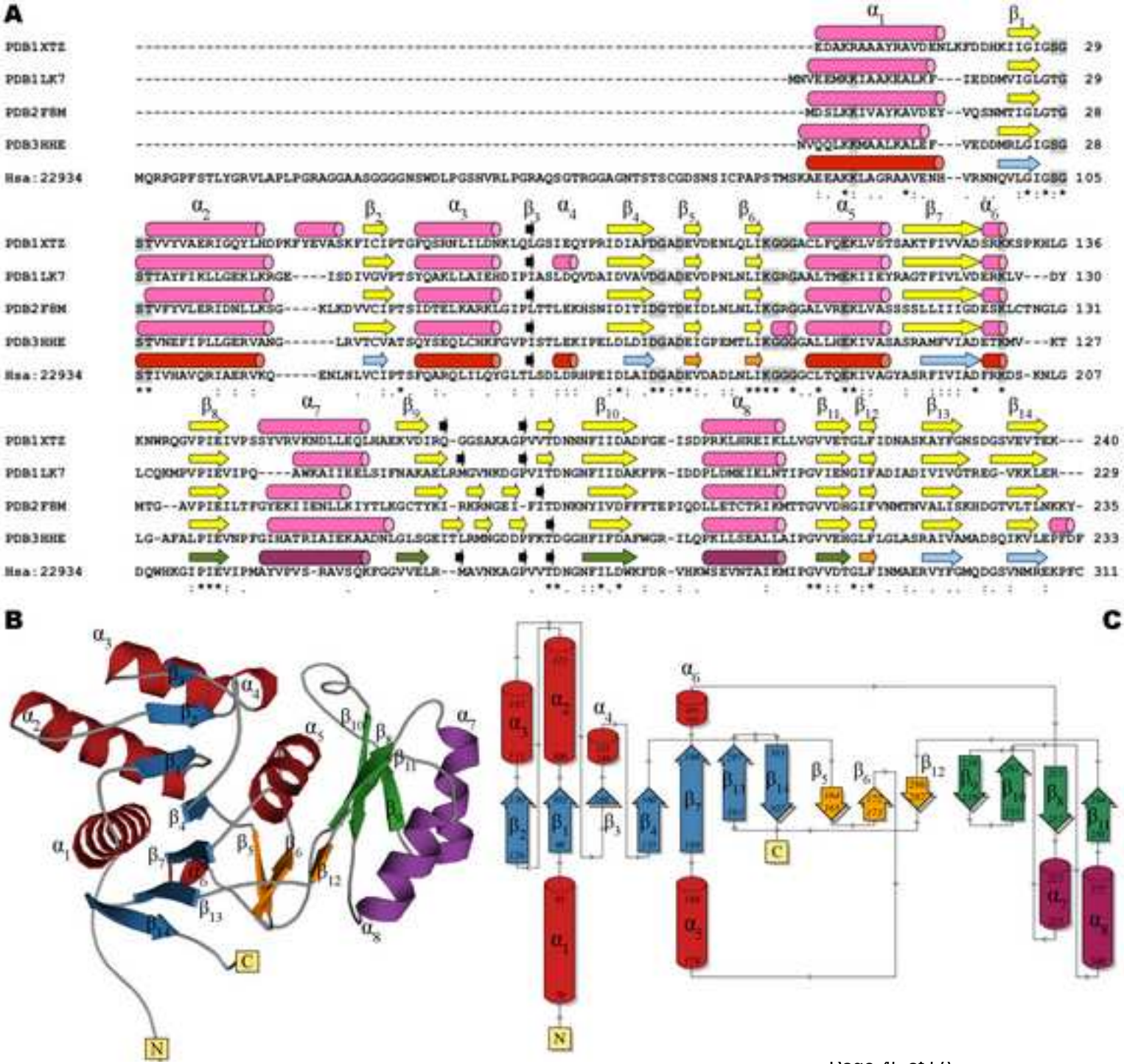


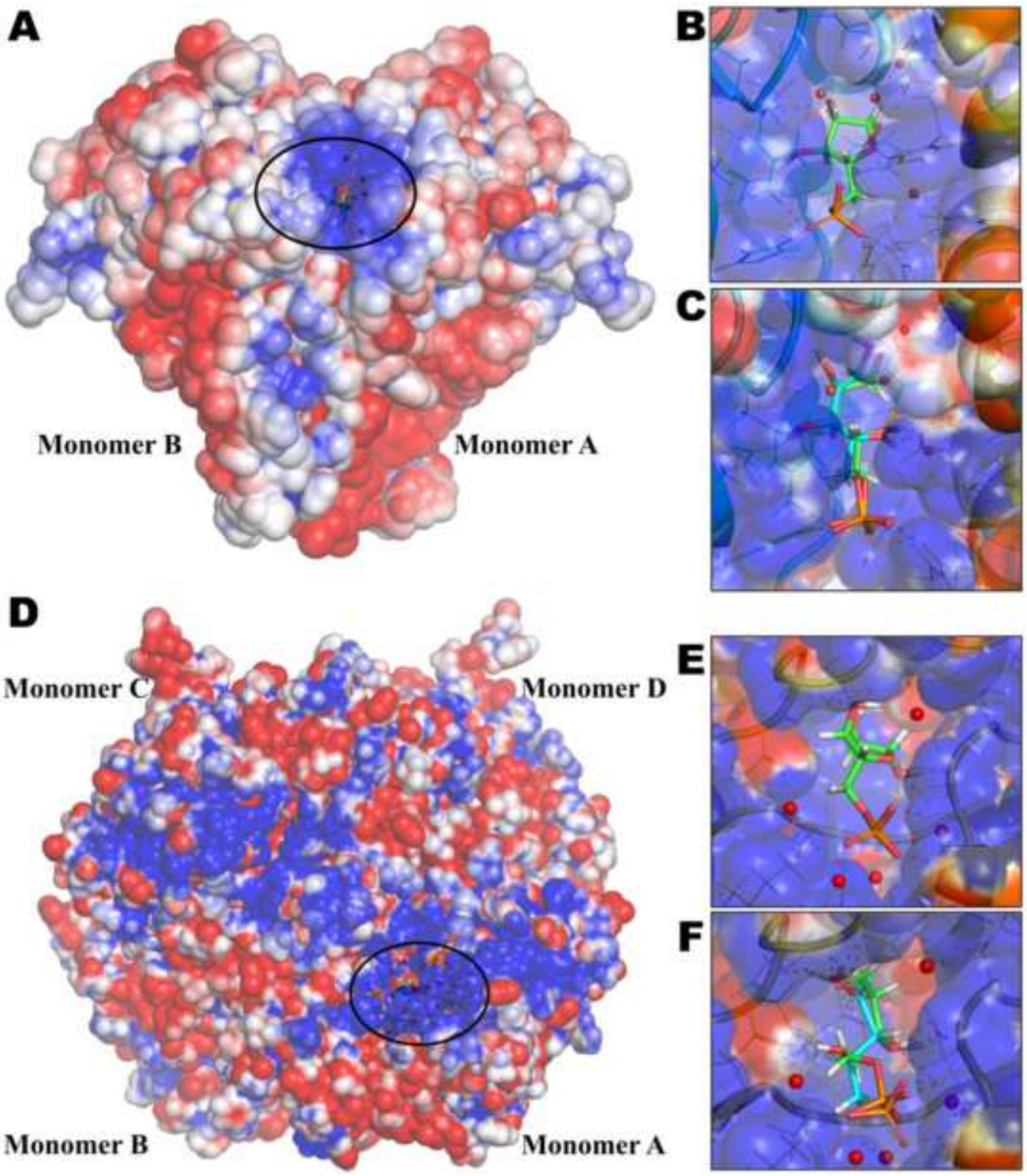


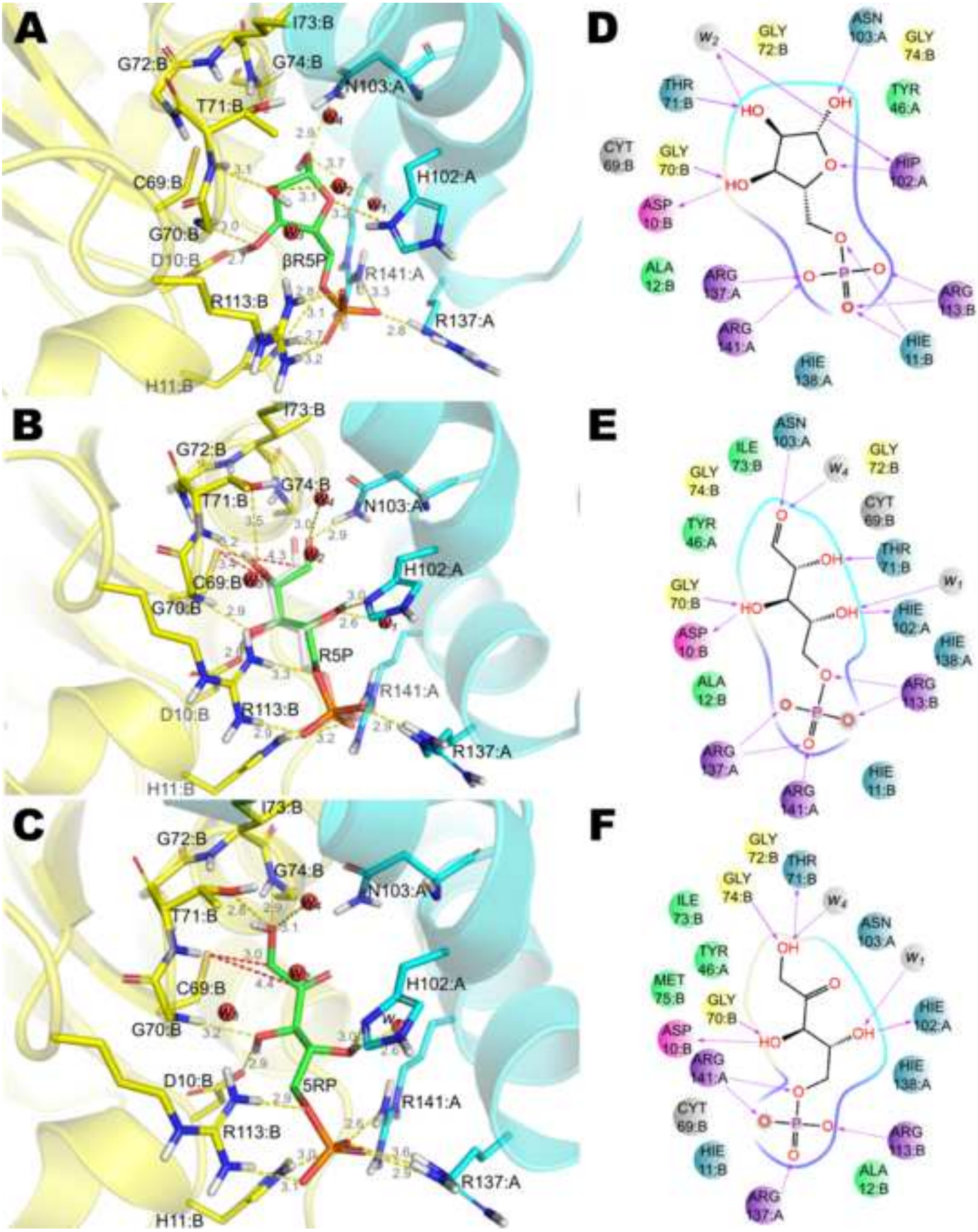
Figure(s)

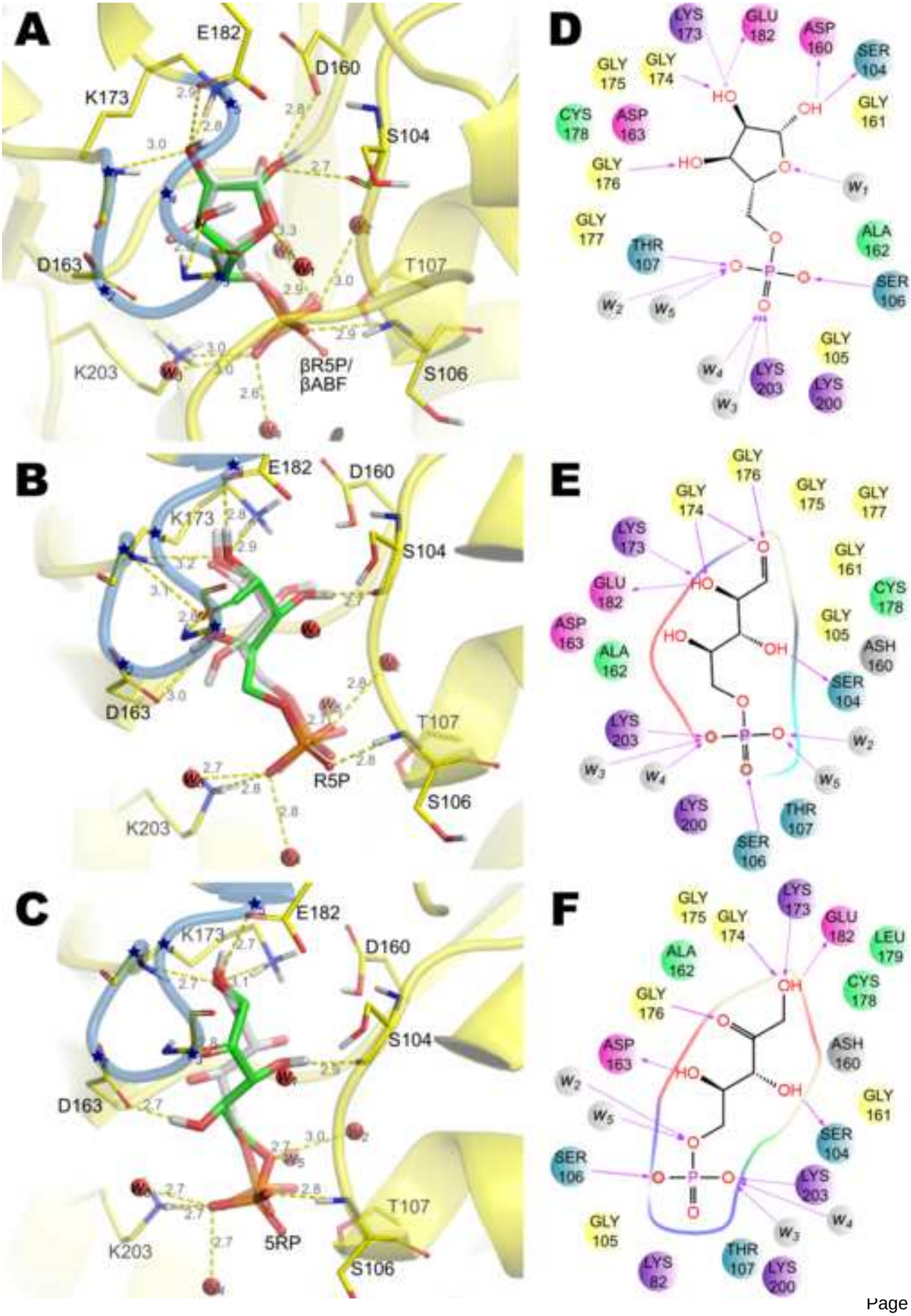

A
C
B
I73:B
T71:B
G74:B
N103:A
G70:B
H102:A
C69:B
R113:B
H138:A
Y46:A
D10:B
R141:A
R5P
H11:B
R137:A
D
G177:A
G174:A
C178:A
E182:A
K173:A
D99*
D163:A
5RP
S106:A
K203:A
T107:A

Figure(s)

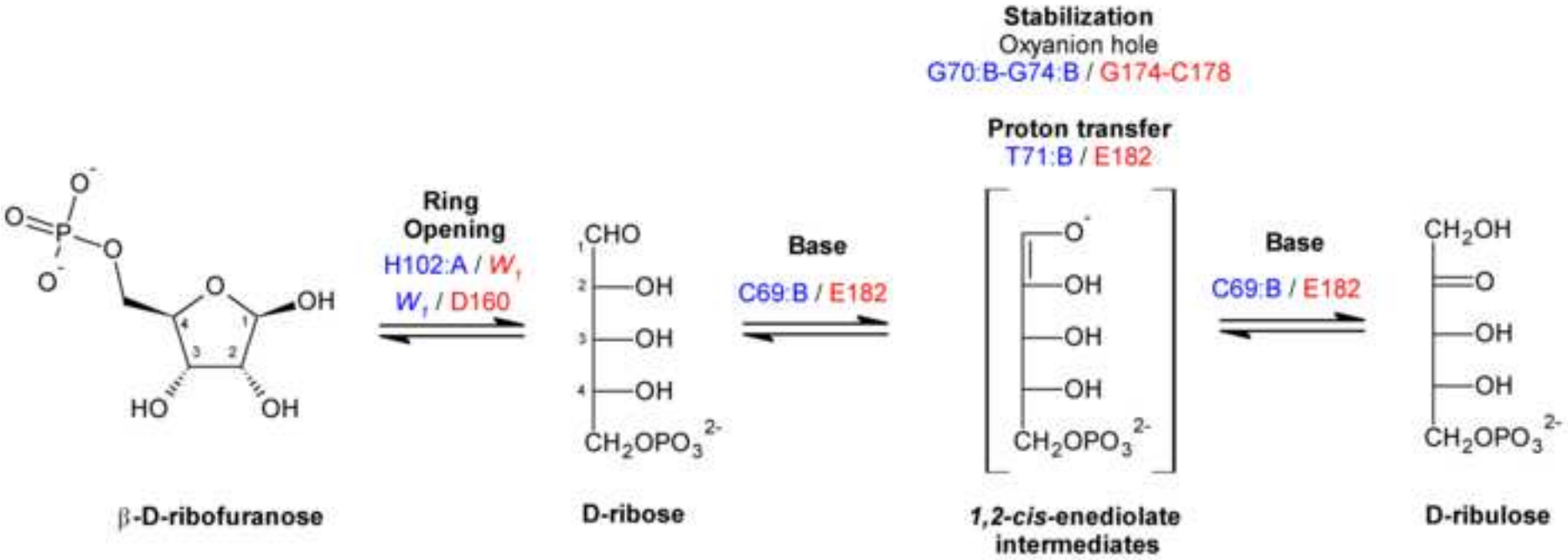








Figure(s)



Structural Modeling and Docking Studies of Ribose 5-Phosphate Isomerase from *Leishmania major* and *Homo sapiens*: A Comparative Analysis for Leishmaniasis Treatment

Priscila V. S. Z. Capriles[1,2‡], Luiz Phillippe R. Baptista[1,3‡], Isabella A. Guedes[1‡], Ana Carolina R. Guimarães[3,4], Fabio L. Custódio[1], Marcelo Alves-Ferreira[3], Laurent E. Dardenne[1*]*

1- Grupo de Modelagem Molecular de Sistemas Biológicos, Laboratório Nacional de Computação Científica, GMMSB/LNCC-MCTI, Petrópolis, Brazil.
2- Programa de Pós-graduação em Modelagem Computacional, Faculdade de Engenharia / Instituto de Ciências Exatas, Universidade Federal de Juiz de Fora, PGMC/UFJF-MEC, Juiz de Fora, Brazil.
3- Laboratório de Genômica Funcional e Bioinformática, Instituto Oswaldo Cruz, IOC/FIOCRUZ-MS, Rio de Janeiro, Brazil.
4- Centro de Desenvolvimento Tecnológico em Saúde, Instituto Oswaldo Cruz, CDTS/FIOCRUZ-MS, Rio de Janeiro, Brazil.

HIGLIGHTS

- *In silico* comparative analysis of ribose 5-phosphate isomerase (R5PI) type A and B.
- 3D model construction of analogous *Leishmania major* and *Homo sapiens* R5PI.
- 3D models were constructed combining comparative and *ab initio* modeling techniques.
- Active site characterization based on docking studies of the substrates R5P and 5RP.
- We propose two distinct reaction mechanisms (R5P↔5RP) catalyzed by LmRpiB and HsRpiA.