Role of Structural Features in Oligomerization, Active-Site Integrity and Ligand Binding of Ribose-1,5-Bisphosphate Isomerase

Prerana Gogoi, Shankar Prasad Kanaujia *

Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India

A B S T R A C T

Pentose bisphosphate pathway, exclusively found in archaea, is similar to the pentose phosphate pathway present in bacteria and eukarya. In pentose bisphosphate pathway, the conversion of ribose moieties of nucleosides into 3-phosphoglycerate (3-PGA) involves multiple steps; one of them being the conversion of ribose-1,5-bisphosphate (R15P) to ribulose-1,5-bisphosphate (RuBP) catalyzed by an enzyme ribose-1,5-bisphosphate isomerase (R15Pi). The availability of the three-dimensional structure of R15Pi had facilitated the understanding of various structural and functional aspects of the enzyme. Nevertheless, the structure of R15Pi also left several significant questions unanswered that would aid in understanding the structure-function relationship of the enzyme. Thus, we have taken up a computational approach to further understand the role of various structural features of the enzyme R15Pi. Results obtained from molecular dynamics (MD) simulations aided in understanding the obligation of the enzyme R15Pi to oligomerize and also in deciphering the role of catalytic residue(s) in structural stability. Identification of invariant water molecules of the enzyme R15Pi helped in discerning their significance at the active-site pocket and structurally important regions. Further, molecular docking studies allowed the identification of the amino acid residues essential for holding the substrate (R15P) or product (RuBP) in the vicinity of the active site of the enzyme R15Pi. Interestingly, results of the molecular docking studies assisted in the identification of an "alternate binding site" for the substrate, R15P. Finally, based on these results, we propose a mechanism of "substrate sliding" for the enzyme R15Pi.

© 2019 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The pentose phosphate pathway present in bacteria and eukarya is accountable for the interconversion of pentoses and hexoses required for the nucleic acid biosynthesis and during the glycolysis as well as gluconeogenesis, respectively [1–3]. However, almost all archaea, are known to be devoid of a canonical pentose phosphate pathway [4–6]. In archaea, the production of trioses and hexoses from pentoses occur via a metabolic pathway termed as pentose bisphosphate pathway [7]. In this pathway, nucleosides (the main source of pentoses) are converted into trioses and hexoses via the formation of ribose-1,5-bisphosphate (R15P). The molecule R15P can either be generated through the formation of ribose-1-phosphate (R1P) with the help of

the enzymes, nucleoside phosphorylase and ADP-dependent ribose-1-phosphate (ADP-R1P) kinase or via the nucleoside 5′-monophosphate (NMP) degradation pathway [7]. Apart from the final step of NMP degradation pathway, where ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) produces 3-phosphoglycerate (3-PGA) to be directed towards the glycolysis and gluconeogenesis pathways, another crucial step includes the isomerization of R15P to ribulose-1,5-bisphosphate (RuBP) by an enzyme ribose-1,5-bisphosphate isomerase (R15Pi) [7–11].

Till date, the three-dimensional crystal structure of the enzyme R15Pi have been elucidated from only two archaea *Thermococcus kodakarensis* and *Pyrococcus horikoshii* [11,12]. Structurally, the protomer of the enzyme R15Pi comprises of an N-terminal α-helical domain (NTD) and a C-terminal αβα-sandwich domain (CTD). The active site of the enzyme R15Pi is formed at the interface of the NTD and CTD (Fig. 1A). Functionally, the enzyme R15Pi is made up of six identical protomers with interaction(s) occurring at the CTD while the NTD remains completely isolated (Fig. 1B). The available three-dimensional crystal structures of the enzyme R15Pi can be categorized in three different forms, (i) the unliganded open form (R15Pi-Form I), (ii) the substrate-bound closed form (R15Pi-Form II) and (iii) the product-bound open form (R15Pi-Form III). However, the substrate-bound
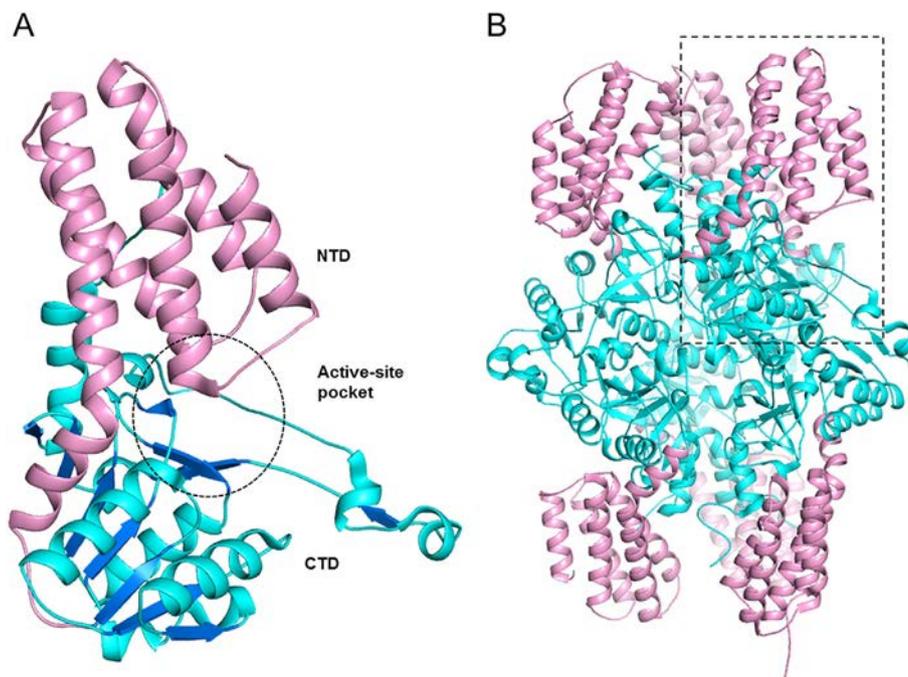
Fig. 1. The three-dimensional structure of the enzyme R15Pi. (A) Protomer of R15Pi (PDB id: 3A11) comprised of NTD (light pink) and CTD (blue and cyan). The active site of the enzyme is encircled by dashed line. (B) The hexameric form of the enzyme, wherein the NTD and CTD are colored in light pink and cyan, respectively. One subunit of the hexamer is marked by dashed line for the clarity of the figure.

structure of the wild type R15Pi (R15Pi-WT) is not available owing to the ability of the enzyme to spontaneously convert R15P to RuBP during the crystallization process. Thus, to obtain substrate-bound enzyme (R15Pi-Form II), mutation of one of the two crucial catalytic residues was carried out [11,12].

In the absence of the substrate (R15P), the enzyme R15Pi retains an open conformation with a completely accessible active site. Upon R15P binding, the NTD move towards the CTD, resulting in the shielding of the active-site cleft from the aqueous environment [11]. Once the conversion of the substrate (R15P) to product (RuBP) is accomplished, the NTD reverts back to its initial position propelling the enzyme into its open conformation, which also allows the product (RuBP) to exit the active site (Fig. S1). This conversion process of R15P to RuBP requires the catalytic activity of two residues Cys133 and Asp202 (numbering according to R15Pi from *T. kodakarensis*, TkR15Pi) via the formation of a *cis*-phosphoenolate intermediate. Although, point mutation of either of these residues abolishes the enzymatic activity completely, it does not hinder the binding of R15P or RuBP to the enzyme [11,12].

Although the three-dimensional structure of the enzyme R15Pi has been elucidated [11,12], the accountability of the enzyme to function as a hexamer, recognition of the essential catalytic residue for structural stability, identification of essential substrate- and/or product-binding residues and structural as well as functional role(s) of invariant water molecules remains undescribed. Thus, in this study, we employed computational approaches such as molecular dynamics (MD) simulation, molecular docking and data mining to address these aspects and consequently to obtain an insight into the role of the structural features of the enzyme R15Pi in rendering adequate stability and optimal ligand binding environment.

2. Materials and Methods

2.1. Molecular Dynamics Simulation

MD simulation was performed using the package GROMACS v.5.1.4 [13]. For all the simulations (Table S1), AMBER03 force field embedded

in GROMACS was utilized [14]. The module *editconf* of GROMACS was used to generate a cubic box keeping the minimum distance between the solute and the edge of the box at least 1.0 nm. The module *genbox* present in the GROMACS suite was used to solvate the protein models with the flexible simple point charge (SPC/E) water model. Chloride and sodium ions were utilized to neutralize the overall charge of the system. To inhibit edge effects, periodic boundary condition (PBC) was applied along the three spatial directions. Energy minimization was performed using the steepest descent method with a maximum force cutoff of 1000 kJ/mol/nm. The solvent and ions around the protein were equilibrated in two phases. Firstly, equilibrations were performed under an NVT (canonical or isothermal-isochoric) ensemble for 100 ps with a reference temperature of 310 K. Subsequently, second phase of equilibration was performed under an NPT (isothermal-isobaric) ensemble for 500 ps with a reference pressure of 1 bar. A velocity rescaling thermostat [15] with coupling constant of 0.1 ps was used to control the reference temperature at 310 K. The reference pressure of 100 kPa (1 bar) with coupling constant 2 ps was controlled using the Parrinello-Rahman barostat [16]. The long-range electrostatic interactions were computed using the particle mesh Ewald (PME) method [17,18] while the short-range van der Waals interactions were computed using Verlet neighbor list calculation with a cutoff of 0.8 nm. P-LINCS algorithm [19] was used to constrain the bond lengths and a time step of 2 fs was used to integrate the equations of motion. Each MD simulation was performed for a time period of 100 ns and 300 ns for R15Pi from *T. kodakarensis* (TkR15Pi) and *P. horikoshii* (PhR15Pi), respectively, resulting in a total of 2.6 μs of simulation. The analyses were performed using the programs available in GROMACS and home-built shell scripts. The program Xmgrace [20] was used to prepare the graphs.

2.2. Identification of Invariant Water Molecules

The three-dimensional atomic coordinates of 13 crystal structures of PhR15Pi are available in closed conformation. These structures include R15Pi-WT, R15Pi-C135S and R15Pi-D204N proteins bound to either R15P or RuBP (Table S2). However, only those structures which harbor greater than or equal to 100 water molecules were considered for the

identification of invariant water molecules. For identification of invariant water molecules in TkR15Pi, all the three available structures were considered even though they contain less than 100 water molecules. All the structures of PhR15Pi belong to the space group $P3_112$ and consist of three subunits in the asymmetric unit (ASU). The three monomeric units in each of the eight structures were considered as an independent molecule. Thus, a total of 18 monomers of PhR15Pi were taken as data set for the analysis. The crystal structure deposited with PDB id 5YFT (chain B) was taken as the reference or fixed molecule as it contains the highest number (236) of water molecules among the considered structures. All the remaining structures were considered as mobile molecules and superimposed on 5YFT (chain B) to identify the invariant water molecule(s). The superimposition of the structures was manually performed in the program Coot [21] as well as using a home-built shell script. The cutoff distance between a pair of superposed water molecules was kept as 1.8 Å. A water molecule was considered to be invariant only when it was present in all the structures considered. An invariant water molecule was considered to be buried when it had a solvent accessible surface area (SASA) of less than equal to 2.5 Å^2 [22]. For calculation of the residence frequency of the invariant water molecules, MD simulation was performed for a time period of 300 ns. The structures generated during MD simulation at every 1 ns were used for comparison and to compute the residence frequency of each invariant water molecule. The interaction(s) between protein atoms and solvent molecules were calculated with a hydrogen-bond distance of 3.5 Å. All the other parameters of MD simulation were kept same as mentioned in the previous section.

2.3. Molecular Docking Studies

To perform molecular docking experiments, the three-dimensional atomic coordinates of the two open forms of R15Pi i.e. R15Pi-Form I (PDB id: 3A11) and R15Pi-Form III (PDB id: 3A9C) from TkR15Pi were extracted from the RCSB Protein Data Bank (PDB) [23]. Since R15Pi-Form II (PDB id: 3VM6) has a closed active site, only the open forms of the protein were used for docking experiments. The ligands (R15P and RuBP) used for the docking experiments were extracted from their bound states of the enzyme TkR15Pi. Molecular docking experiments were carried out using the program Autodock version 4.0 [24]. During each docking experiment, the receptor (i.e. protein) molecule was kept as rigid while the ligand (i.e. R15P and RuBP) molecules were considered both as rigid and flexible along the rotatable bonds. As both the rigid and flexible ligands produced similar outcomes, results of only the rigid-ligand docking were considered for analysis. The size of the grid box was kept to $126 \times 126 \times 126$ with a spacing of 0.375 Å between the grid points and taking the center of mass of the protein as the grid center. To search for the best conformational space of the ligand, the Lamarckian genetic algorithm (LGA) with a total of 2000 runs was performed. The docked conformations of ligands in each molecular docking experiment were clustered with a root mean square deviation (RMSD) cutoff of 2.0 Å. The ligand conformation having the lowest estimated free energy of binding (EFBE) was selected as the final docked ligand to the protein. Identification of the amino acid residues involved in the interaction(s) with the docked ligand was performed using the program Coot [21]. All the figures of the docked ligands were generated using the program PyMOL (PyMOL Molecular Graphics System, Schrodinger, LLC).

3. Results

3.1. Hexameric State is Obligatory for the Structural Stability and Optimal Ligand Binding of R15Pi

To understand the significance of the oligomeric state of R15Pi, MD simulations of R15Pi-monomer and R15Pi-hexamer were performed. The R15Pi-hexamer is found to be more stable than the R15Pi-

monomer during the simulation (Fig. 2A). Nonetheless, binding of either of the ligands i.e. R15P or RuBP to the R15Pi-monomer reasonably increases the stability of the enzyme (Fig. 2A). On the other hand, R15Pi-hexamer remains stable throughout the simulation irrespective of the presence or absence of ligand (Fig. 2A). The stability of the hexameric protein is further supported by the comparison of the potential energies of the R15Pi-monomer and R15Pi-hexamer (Fig. 2B).

An analysis of the root mean square fluctuations (RMSFs) exhibits that the NTD of R15Pi-hexamer displays a slightly higher fluctuation than that of the R15Pi-monomer. In contrast, the RMSFs of the CTD are notably lower for R15Pi-hexamer than that of the R15Pi-monomer. The most significant fluctuation of the CTD in the R15Pi-monomer occurs at a long loop region (residues 245–283) of the protein. Interestingly, this fluctuation is somewhat decreased in the presence of either the substrate (R15P) or the product (RuBP) bound at the active site of the protein (Fig. 2C).

To find whether the oligomeric state of R15Pi had any effect on the binding of the substrate or product, the distance between protein and the ligand (R15P or RuBP) was computed during the 100 ns simulation. Results show a lesser distance between them in the case of R15Pi-hexamer compared to that of the R15Pi-monomer suggesting a better ligand binding to the hexameric protein (Fig. 2D).

3.2. The Catalytic Residue Cysteine is Preferred over Aspartate for Structural Stability

To identify the preferred catalytic residue between Cys135 and Asp204 (Cys133 and Asp202 in TkR15Pi) for the structural stability of enzyme PhR15Pi, MD simulations of the mutant proteins R15Pi-C135S and R15Pi-D204N, both in bound and unbound forms, were performed. The RMSD plot reveals that the R15Pi-C135S mutant protein displays a relatively higher instability in the structure as compared to R15Pi-D204N mutant protein (Fig. 3A). However, in the presence of either R15P or RuBP, the mutant gains structural stability and exhibits a lower RMSD (Fig. S2A). On the other hand, in case of R15Pi-D204N mutant, the presence or absence of ligand (R15P or RuBP) shows no effect on the stability of the protein (Fig. S2A). This is further supported by comparison of the potential energy of two mutants, where the mutant R15Pi-C135S exhibits a higher potential energy than that of the mutant R15Pi-D204N (Fig. 3B and S2B), indicating the importance of the cysteine residue in rendering structural stability.

In the three-dimensional crystal structure of R15Pi-C135S mutant protein, Ser135 interacts with Gln166 (bond distance: 2.8 Å) while in R15Pi-D204N, no such interactions occur between Cys135 and Gln166 (bond distance: 4.2 Å) (Fig. 3C). Hence, in R15Pi-D204N mutant protein, Gln166 is available to interact with Lys164 which in turn coordinates with Glu260 present in the longest loop region. Glu260 further interacts with another amino acid residue Arg254 which is also present in the longest loop region. Thus, Cys135 is involved in indirectly rendering stability to the highly flexible longest loop region by adhering it to the core CTD. However, hydrogen bond formation between Ser135 and Gln166 in R15Pi-C135S mutant protein makes Gln166 unavailable for further interactions (Fig. 3C). The higher fluctuation in the longest loop region in case of R15Pi-C135S mutant protein due to the lack of stabilizing interactions is also evident from the RMSF plot (Fig. 3D). The fluctuation also decreases in the presence of either the substrate or product owing to their interaction with Arg256 which resides in the longest loop region (Fig. S2C and S2D).

3.3. Invariant Water Molecules are Strategically Arranged Around the Active Site and Structurally Important Regions

A total of 29 water molecules were identified to be invariant in a protomer of the closed conformation of PhR15Pi. Out of these, 12 invariant water molecules are present either in the active-site pocket or in its vicinity (Table S3), five (IW1, IW2, IW3, IW4 and IW5) of which make
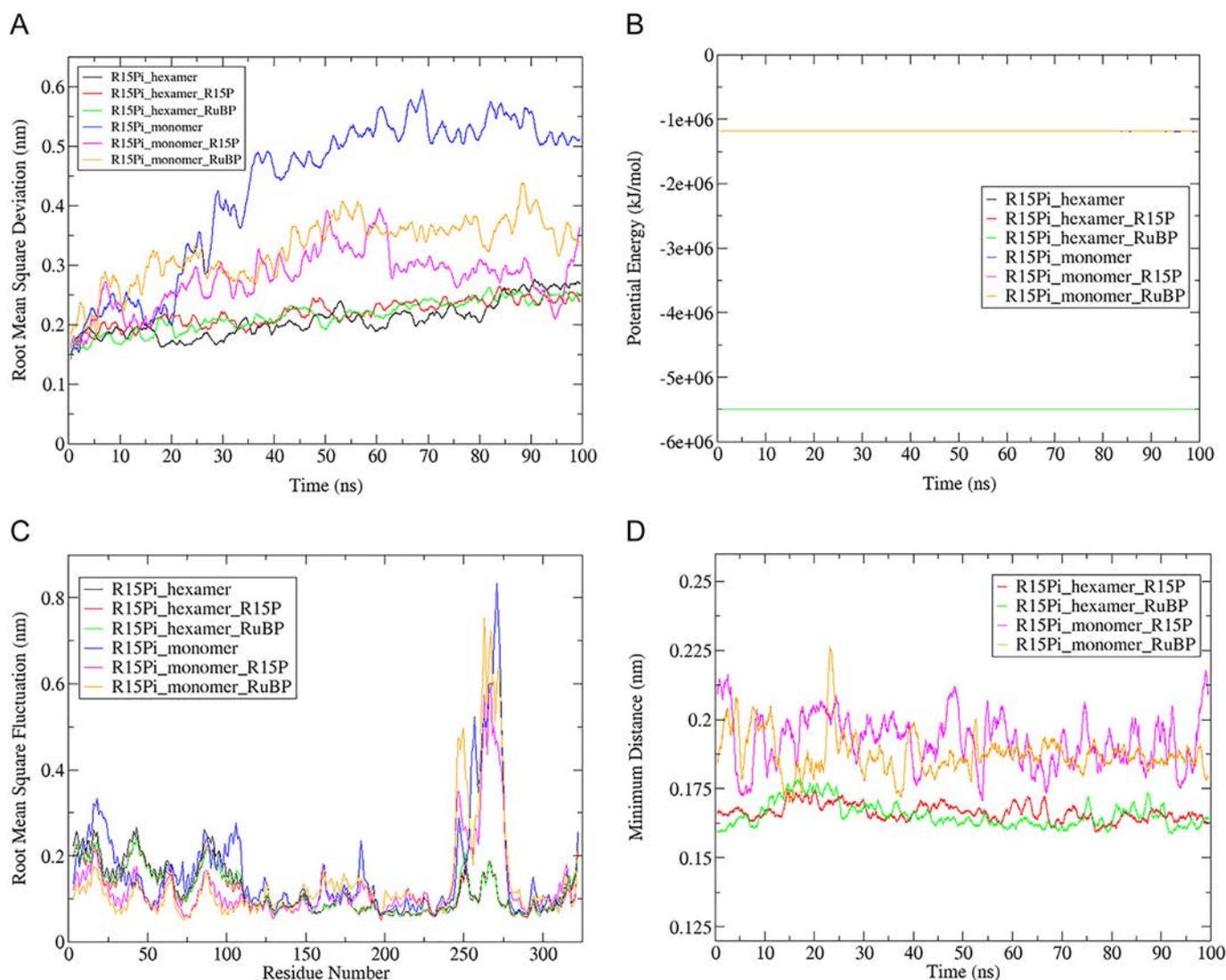
Fig. 2. Dynamics of R15Pi-monomer and R15Pi-hexamer. The plot for the (A) root mean square deviation (RMSD), (B) potential energy, (C) root mean square fluctuation (RMSF) and (D) the minimum distance between the protein and ligand. The different forms of the enzyme R15Pi for which simulations were performed include R15Pi-monomer (blue), R15P-bound R15Pi-monomer (magenta), RuBP-bound R15Pi-monomer (orange), R15Pi-hexamer (black), R15P-bound R15Pi-hexamer (red) and RuBP-bound R15Pi-hexamer (green).

direct interaction with both R15P and RuBP in a similar manner. For instance, IW2 is hydrogen bonded to the 1-phosphate group of the ligand while the 5-phosphate group is held by three other invariant water molecules IW1, IW3 and IW5. The invariant water molecule IW4 forms hydrogen bond with the O3 oxygen atom of the ribose sugar (Fig. 4A and Table S3). All these water molecules are highly stable as indicated by their low B factors (Table S4). Moreover, these invariant water molecules stabilize residues in the vicinity of the active site of the enzyme. One of the invariant water molecules IW2 forms hydrogen bond with $O^{\delta 1}$ atom of the catalytic residue Asp204 of R15Pi-WT. Similarly, invariant water molecules IW1, IW3, IW4, IW5 and IW6 forms hydrogen bonds with the residues Arg26, His134, Gly202, Ala203, Asn214, Lys215, Ala235, Lys240 and Glu287 (Fig. 4B and Table S3).

A total of another five invariant water molecules (IW7, IW8, IW9, IW10 and IW11) do not directly interact with the ligand, however, make contact with amino acid residues which form the "active-site roof" of the enzyme. The water molecule IW8 is hydrogen bonded to backbone O atoms of Ile16 and Ile21 as well as to $N^{\eta 2}$ atom of Arg65. The invariant water molecule IW9 hydrogen bonds with the backbone O atom of Pro66 and networks three other water molecules (W585, W666 and W701 of PhR15Pi). Two invariant water molecules IW7 and IW11 are hydrogen bonded to the backbone O and $N^{\eta 2}$ atoms of

Arg22 while IW10 holds the residue Arg29 through $N^{\eta 1}$ and N^{n2} atoms (Fig. 4C and Table S3). Interestingly, only one invariant water molecule (IW12) interacts with amino acid residues forming the "active-site floor" of the enzyme. More precisely, IW12 interacts with the N atom of Pro63 & Lys164, O atom of Thr161 & Arg162 and $O^{\epsilon 1}$ atom of Gln166. These five amino acid residue are part of a loop region which shield the active-site pocket from underneath (Fig. 4D and Table S3). Notably, three invariant water molecules IW4, IW5 and IW12 are present in both open and closed conformations of the enzyme TkR15Pi as well, signifying their importance.

Apart from the invariant water molecules at the active-site pocket and its vicinity, 17 invariant water molecules are present at various structurally relevant regions (Fig. 5A). Four invariant water molecules (IW15, IW16, IW17 and IW18) are involved in holding the long loop region. Three invariant water molecules (IW19, IW20 and IW21) interact with the longest helix, α_5 while two other invariant water molecules (IW27 and IW28) interact with helices α_7 and α_4, respectively. On the other hand, IW29 forms hydrogen bond with three water molecules (W603, W559 and W549 of PhR15Pi) which in turn interacts with the loop region connecting helices α_3 and α_4. Five invariant water molecules (IW22, IW23, IW24, IW25 and IW26) interact with amino acid residues at the inter-subunit region. Although these invariant water
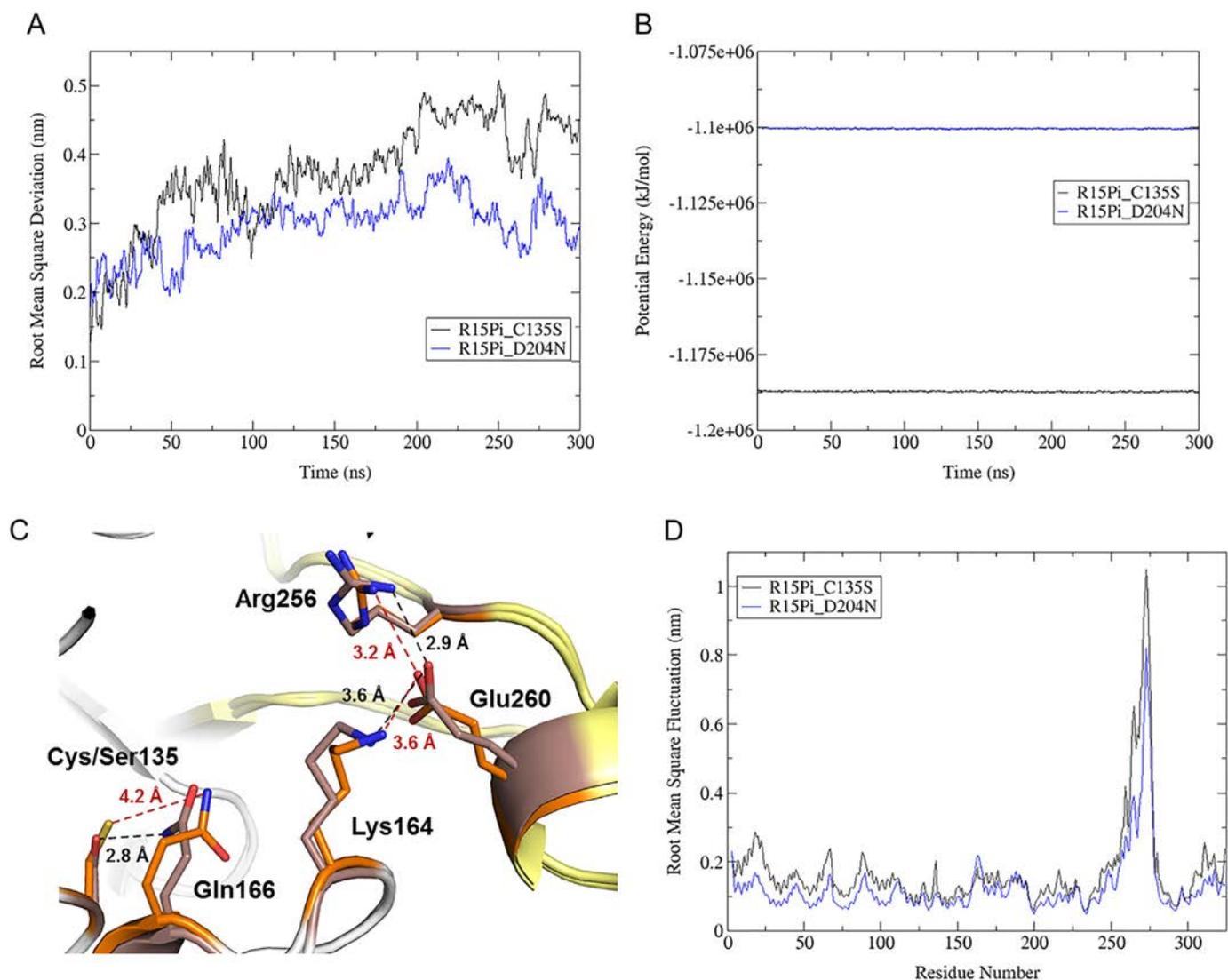
Fig. 3. Dynamics of the mutants R15Pi-C135S and R15Pi-D204N. The plot of (A) root mean square deviation (RMSD) and (B) potential energy of the mutant proteins R15Pi-C135S (black) and R15Pi-D204N (blue) during the MD simulation. (C) Absence (presence) of interaction between Gln166 and Cys135 (Ser135) in R15Pi-C135S (R15Pi-D204N) mutant protein(s). The amino acid residues involved in conferring structural stability in R15Pi-C135S (orange) and R15Pi-D204N (light brown) mutant proteins are shown as lines and their interactions are depicted as red and black dashed lines. The distances between the atoms are indicated adjacent to their corresponding dashed lines. The flexible longest loop region is highlighted in yellow. The amino acid residues are numbered according to the enzyme PhR15Pi. (D) The plot of root mean square fluctuation (RMSF) of the mutant proteins R15Pi-C135S (black) and R15Pi-D204N (blue) during the MD simulation.

molecules do not directly interact with the amino acid residues of the associating monomer, they interact with water molecules present at the void created at the hexameric interface (Fig. 5A and Table S3). Two invariant water molecules (IW13 and IW14) are involved in holding together the three regions NTD, CTD and the long loop region of R15Pi by anchoring the residues Thr67, Pro163, Glu254 and Glu260 (Fig. 5B,C and Table S3). All the amino acid residues involved in interacting with the invariant water molecules are well conserved in the enzyme R15Pi (Fig. S3).

3.4. The Substrate R15P Seems to Bind at an "Initial Binding Site" Before Sliding Into the Active Site

Among the available three-dimensional structures of R15Pi, only two (R15Pi-Form I, PDB id: 3A11 and R15Pi-Form III, PDB id: 3A9C) are present in an open conformation. To gain insights into the early binding events of the substrate to the protein R15Pi, docking of the R15P with the R15Pi-Form I was performed. Interestingly, the result reveals that the substrate R15P is docked at a site (hereafter referred to as "initial binding site"), which overlaps the active site (Fig. 6A), with a

binding energy of −10.18 kcal mol^{-1} (Table S5). At this site, the 1-phosphate group of the substrate (R15P) forms hydrogen bond with the main chain N atoms of Lys136 and Ala137 while the 5-phosphate group is held by the N$^{\zeta}$ atoms of Lys24 and Lys136. Although, Lys24 of TkR15Pi is substituted by arginine in PhR15Pi, the availability of a positively charged amino acid residue will favor similar interaction with the negatively charged 5-phosphate group of the ligand (Fig. S3). The main chain N atom of Gly23, O$^{\varepsilon2}$ of Glu235 and N$^{\zeta}$ of Lys238 stabilize the ribose sugar oxygen atoms (Fig. 6B and Table S5). Notably, the residues Lys136, Ala137 and Lys238 bridge the "initial binding site" and active site of the protein.

However, the docking of R15P to R15Pi-Form III reveals its binding directly to the active site instead of the "initial binding site", although with an identical binding energy (−10.18 kcal mol^{-1}), to that of the latter site. Notably, the docked conformation of the substrate is identical to that of the crystal structure. At the active site, the 1-phosphate group of the substrate (R15P) interacts with N$^{\zeta}$ of Lys213 and N$^{\eta1}$ of Arg254 while the 5-phosphate group is coordinated by O$^{\gamma}$ of Ser135, main chain N atoms of Lys136 & Ala137 and N$^{\zeta}$ atoms of Lys213 & Lys238 (Table S5). The oxygen atoms (O2 and O3) of the ribose sugar are
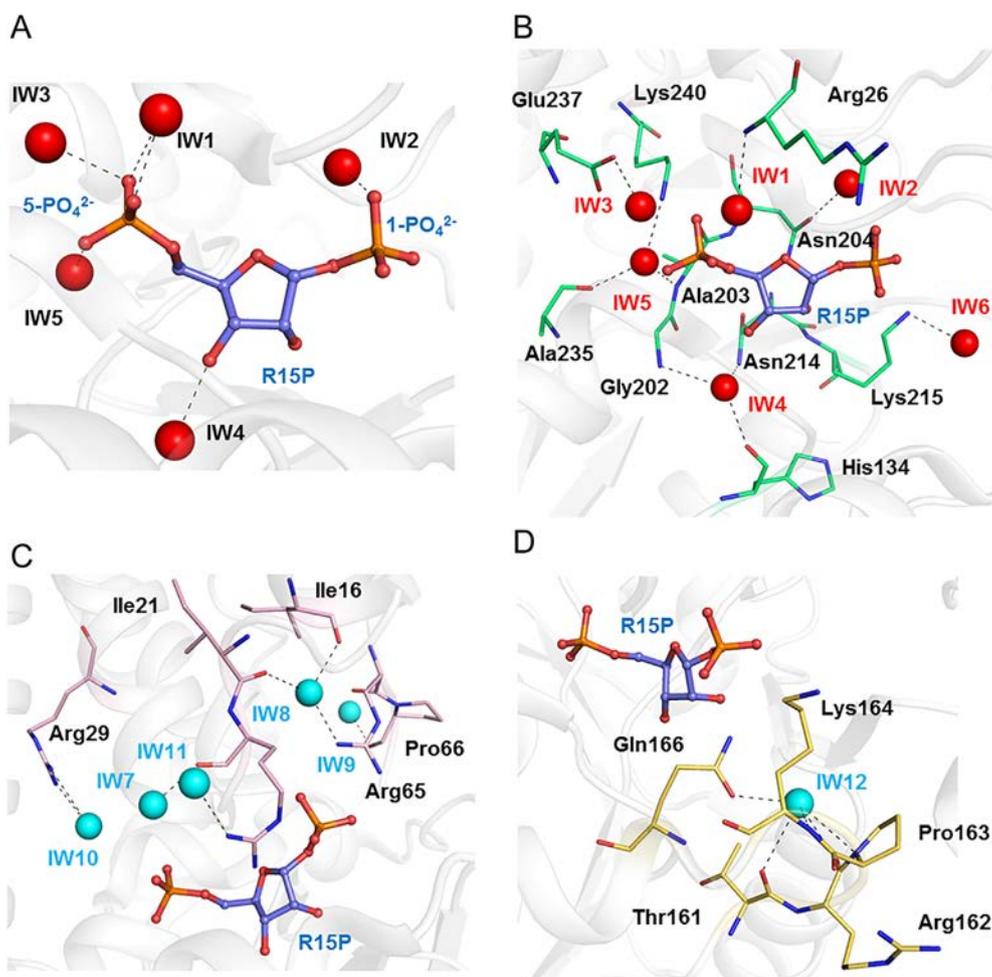
Fig. 4. Invariant water molecules at the active site and in its vicinity. (A) Invariant water molecules (red sphere) interacting directly with R15P (blue ball-and-stick model) and (B) amino acid residues (green lines) at the active-site pocket. (C,D) Invariant water molecules (cyan sphere) interacting with the amino acid residues forming the "active-site roof" (pink lines) and "active-site floor" (yellow lines), respectively, of the protein. The interaction(s) between invariant water molecule and amino acid residues are shown in black dashed lines. All the amino acid residues are numbered according to the enzyme PhR15Pi.

anchored by the $O^{\delta 1}$ & $N^{\delta 2}$ of Asn212 and N of Lys213 (Fig. 6C and Table S5).

However, docking of the product (RuBP) to both the forms i.e. R15Pi-Form I as well as R15Pi-Form III establishes binding directly at the active site with an estimated free binding energies of −9.46 and −11.88 kcal mol^{-1}, respectively. Expectedly, in the R15Pi-form III, the coordinating residues of the protein to the docked RuBP at the active-site pocket are almost identical to that of the substrate, except for the interaction between N^{ζ} of Lys136 and 5-phosphate (Fig. 6D,E and Table S5).

The amino acid residues interacting with the substrate or product at the "initial binding site" and active-site pocket are well conserved in the members of the order thermococcales while less conserved in the members from other archaeal orders (Fig. S3).

3.5. The Residues Lys24 and Lys136 Dictate the Binding of the Substrate at the "Initial Binding Site"

The 1- and 5-phosphate groups of the docked substrate (R15P) at the "initial binding site" interact with the residues Lys136 & Ala137 and Lys24, respectively. Among these, the residue Lys24 is exclusive to the "initial binding site" while Lys136 and Ala137 lie at the overlapping region. Thus, to estimate the significance of Lys24 in the substrate binding at the "initial binding site", it was mutated to alanine. Interestingly, the substrate docked at the active site of R15Pi-K24A mutant, instead of

the "initial binding site", with an estimated free binding energy of −7.89 kcal mol^{-1} (Fig. 7A and Table S5).

Docking of the R15P to R15Pi-Form III resulted in its binding to the active site instead of the "initial binding site". Thus, to identify the factors responsible for the R15P binding at the "initial binding site", the structures of both the forms, R15Pi-Form I and R15Pi-Form III were compared. Notably, among the residues interacting with the R15P at the "initial binding site", two residues Lys24 and Lys136 adapt different conformations in the two forms i.e. R15Pi-Form I and R15Pi-Form III (Fig. 7B). To verify this, the orientation of the residues Lys24 and Lys136 in R15Pi-Form I were altered according to that of R15Pi-Form III. The docking result reveals the binding of the substrate R15P at the active site rather than the "initial binding site" with the side chain N^{ζ} atom of Lys136 interacting with 5-phosphate group of R15P (Fig. 7C and Table S5).

3.6. The Substrate (R15P) Slides From the "Initial Binding Site" to the Active Site

The docking results suggest that the substrate (R15P) binds firstly at the "initial binding site" rather than the active site of the enzyme R15Pi. Thus, we speculated that the substrate may slide into the active site after binding briefly at the "initial binding site". To test this hypothesis, we calculated the distance between R15P and the residues Lys24 (interacts with the 5-phosphate group at the "initial binding site") and Lys213 & Arg254 (interacts with the 1-phosphate group at the active site)
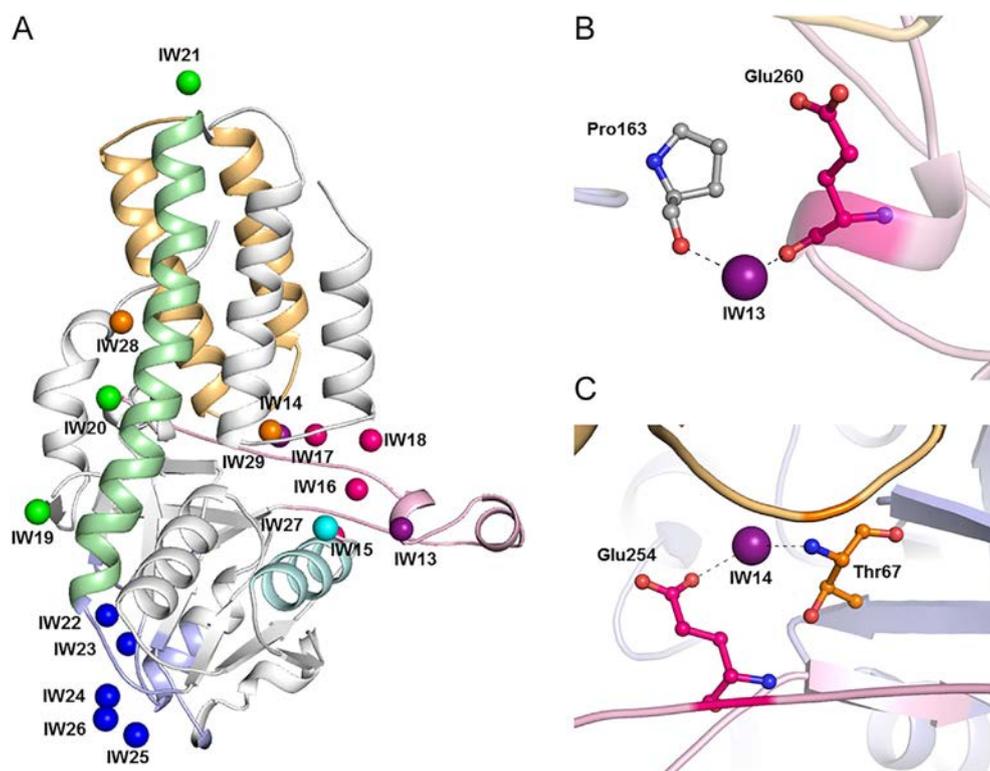
Fig. 5. Invariant water molecules at structurally important regions. (A) Invariant water molecules interacting with amino acid residues at the helices α_4 (light orange), α_5 (pale green) & α_7 (pale cyan), the long loop region (light pink), the loop connecting α_3 and α_4 (light orange) and at the inter-subunit region (light blue). The invariant water molecules (sphere) interacting with the corresponding structural elements are shown in a similar colour. (B,C) Invariant water molecules (purple sphere) involved in holding together the CTD, NTD and the long loop region. The residues interacting with the invariant water molecules are shown as ball-and-stick model and the interaction(s) are shown as black-dashed lines. The amino acid residues are numbered according to the enzyme PhR15Pi.

during the MD simulation. Quite interestingly, the distance between R15P and Lys24 increases from 4.4 Å to 6.8 Å while that between R15P and Lys213 & Arg254 decreases from 6.8 Å to 2.1 Å indicating the movement of the substrate (R15P) from the "initial binding site" to the active site of the protein (Fig. 8A). This is further supported by the decrease (from −9.3 to 0 kJ/mol) and increase (from 0 to −47 kJ/mol) in interaction energies between the R15P and Lys24 and Lys213 & Arg254, respectively (Fig. 8B).

3.7. The Residues Arg20, Arg63, Lys213, Arg238 and Arg254 Are Essential for the Product (RuBP) Binding at the Active Site

The molecular docking results demonstrate that, unlike the substrate, the product (RuBP) binds at the active site instead of the "initial binding site" irrespective of the protein forms i.e. R15Pi-Form I and R15Pi-Form III. Thus, to identify the residues aiding the binding of the product (RuBP) at the active site, we mutated the conserved residues Ser135, Lys213, Lys238 and Arg254 (Fig. S3), which holds the 1- and 5-phosphate groups of the RuBP, individually to alanine. The docking results reveal that except S135A, the other point mutations K213A, K238A and R254A affect the binding of the product at the active site (Fig. 9A,B,C,D and Table S5). Similarly, point mutation of conserved residues Arg20 and Arg63 (Fig. S3), which interact with the 5-phosphate group of the substrate (R15P) in the R15Pi-Form II (i.e. closed conformation), to alanine causes the binding of the product (RuBP) to the "initial binding site" instead of the active site of the protein (Fig. 9E,F and Table S5).

4. Discussion

The enzyme R15Pi has been well studied in terms of its three-dimensional structure, domain motions, catalytic residues and enzyme

reaction mechanism [10–12]. However, some of the questions that still need to be addressed include an understanding of the obligation of R15Pi to be present in a hexameric state, role of the catalytic residue in structural stability, role of invariant water molecules and identification of amino acid residues essential for ligand binding at the active site of the protein. Furthermore, acquisition of structural information related to the initial stages of substrate binding might provide new insight into the mechanism of enzyme substrate interaction. Overall, understanding these aspects of the enzyme R15Pi would aid in further comprehending the structure-function relationship of the enzyme. From MD simulation results, it is evident that R15Pi-hexamer is more stable than R15Pi-monomer. In the R15Pi-monomer, higher fluctuations could be perceived at the CTD including the long loop region. However, these fluctuations are stabilized to a large extent in the R15Pi-hexamer owing to the strong interaction present between the CTD of each monomer. Since the CTD remains rigid, only the NTD of R15Pi accomplishes the domain motion essential for ligand binding. R15Pi-hexamer also forms more stable interaction with the ligand which might also be due to the fact that most of the amino acid residues of the active-site pocket which interacts with the ligand are present at the CTD. A flexible CTD, as in the case of R15Pi-monomer, might hinder a stable binding with the ligand moiety, thus necessitating the requirement of a hexameric enzyme.

The two catalytic residues (cysteine and aspartate) required for the enzymatic activity are well identified and studied. Initially, the residue aspartate (202/204 in TkR15Pi/PhR15Pi) was considered to be more vital than the residue cysteine (133/135 in TkR15Pi/PhR15Pi) owing to the inability of the ligand to bind to the D202N mutant of TkR15Pi [11]. However, elucidation of R15P-bound structure of R15Pi-D204N mutant protein from *P. horikoshii* later negated this observation [12]. Thus to recognize the structurally essential catalytic residue, MD simulation was performed for the two mutant proteins. A higher instability
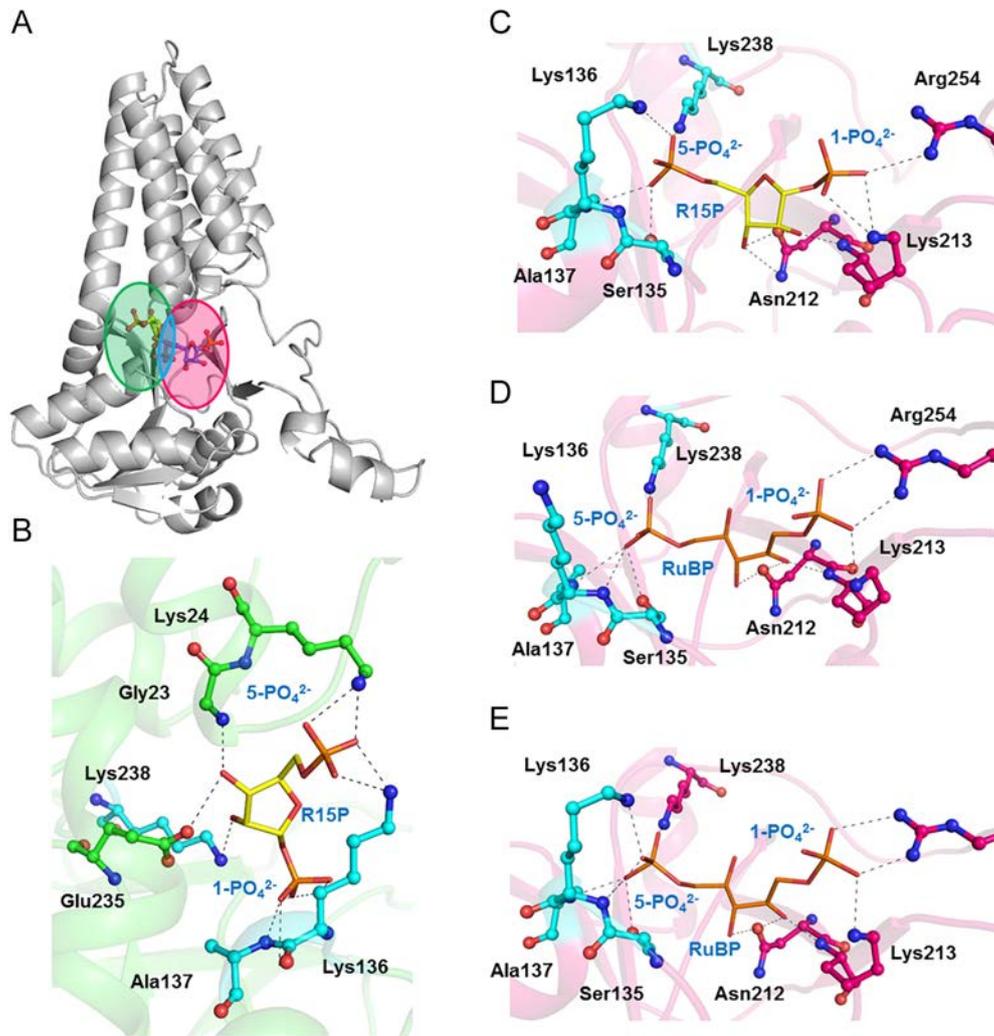
Fig. 6. Docking of R15P and RuBP to R15Pi-Form I and R15Pi-Form III. (A) The two proposed substrate-binding sites viz. "initial binding site" (green) and active site (magenta) in the enzyme R15Pi. The overlapping region between these two sites is highlighted by cyan oval. (B,C) The substrate (R15P, yellow lines) bound at the "initial binding site" in R15Pi-Form I and at the active site in R15Pi-Form III, respectively. (D,E) The product (RuBP, orange lines) bound at the active sites of R15Pi-Form I and R15Pi-Form III. The amino acid residues involved in the interaction at the "initial binding site", the overlapping region and the active site are shown in green, cyan and magenta, respectively. The interaction between ligand and amino acid residues are shown in black dashed lines. All the amino acid residues are numbered according to the enzyme TkR15Pi.

of R15Pi-C135S mutant protein owing to the absence of amino acid interactions essential for stabilizing the longest loop region indicated that the cysteine residue is more crucial for the structural stability of enzyme R15Pi.

Although the enzyme R15Pi is known to undergo an open-to-closed transition during ligand binding, a set of invariant water molecules still populate the active-site pocket. These water molecules are not displaced even after ligand binding indicating their importance in the
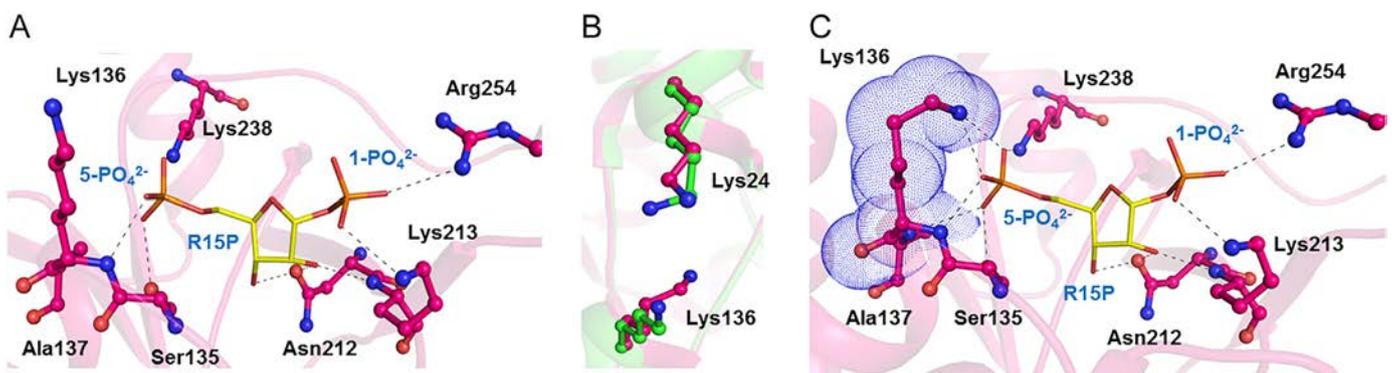

Fig. 7. The substrate (R15P) binding to R15Pi mutant protein. (A) R15P (yellow lines) bound at the active site of R15Pi-K24A mutant; the amino acid residues involved in interaction are shown as ball-and-stick model in magenta. (B) Superimposition of R15Pi-Form I (green) and R15Pi-Form III (magenta) proteins. The residues Lys24 and Lys136, which show difference in their orientation, are highlighted. (C) R15P (yellow lines) bound at the active site after the alteration in the orientation of the residues Lys24 and Lys136. The altered orientation of Lys136 which interact with R15P via the 5-phosphate group is covered with blue dots. The amino acid residues interacting with the R15P at the active site are shown in magenta. All the amino acid residues are numbered according to the enzyme TkR15Pi.
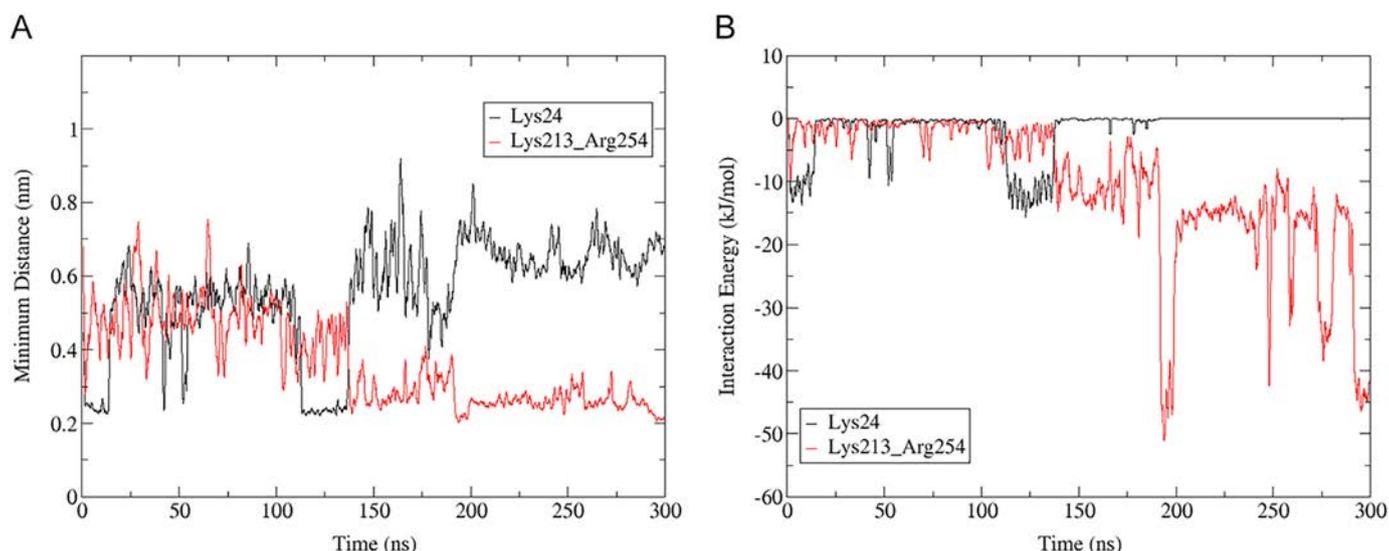
Fig. 8. The movement of the substrate (R15P) during MD simulation. (A) Distance and (B) interaction energy between the substrate (R15P) and the residues Lys24 (black) and Lys213 & Arg254 (red) during the MD simulation.

functioning of the enzyme. Few of the water molecules assist in the binding of the ligand at the active-site pocket by directly interacting with R15P and RuBP in a similar manner while the others interact with the conserved amino acid residues in the vicinity of the active-site pocket. The water molecules concentrated in the vicinity of the active-site pocket mostly interacts with those amino acid residues which are part of the flexible loop regions around the active-site pocket.

By interacting with these amino acid residues, the water molecules permit sufficient movement of the loops around the active-site pocket for accommodating the ligand, without rendering the whole area too flaccid. Furthermore, the invariant water molecules at the active-site pocket and in its vicinity are also observed to be completely buried. This implies that these invariant water molecules might have formed interaction with the amino acid residues at the active-site pocket or its vicinity
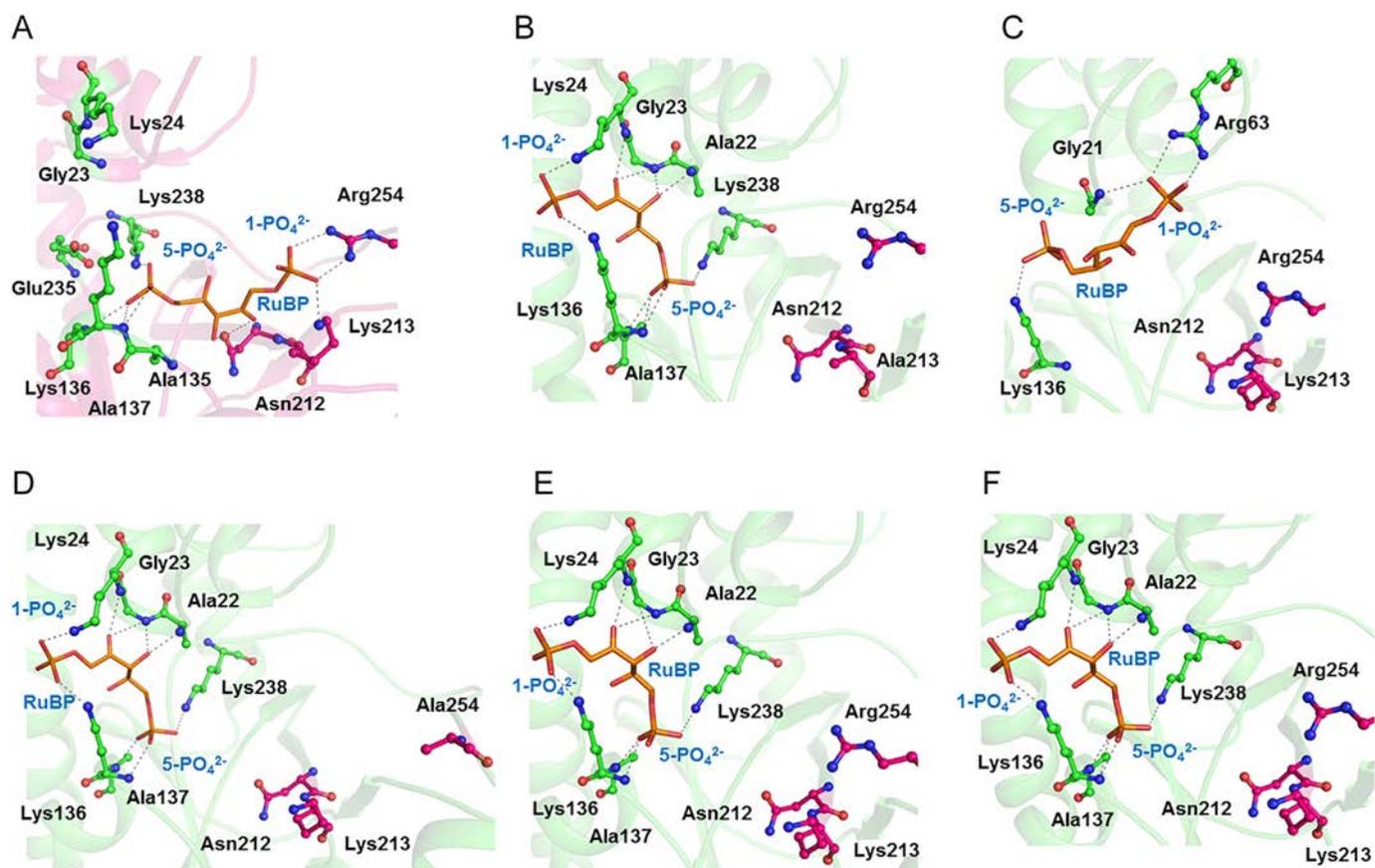
Fig. 9. The product (RuBP) binding to the R15Pi mutants. The binding of RuBP (orange lines) at the active site of (A) R15Pi-S135A mutant and "initial binding site" of (B) R15Pi-K213A, (C) R15Pi-K238A, (D) R15Pi-R254A, (E) R15Pi-R20A and (F) R15Pi-R63A mutant proteins. The amino acid residues interacting with RuBP at the active site and the "initial binding site" are shown in magenta and green, respectively. The numbering of the residues shown in the figure is according to the enzyme TkR15Pi.
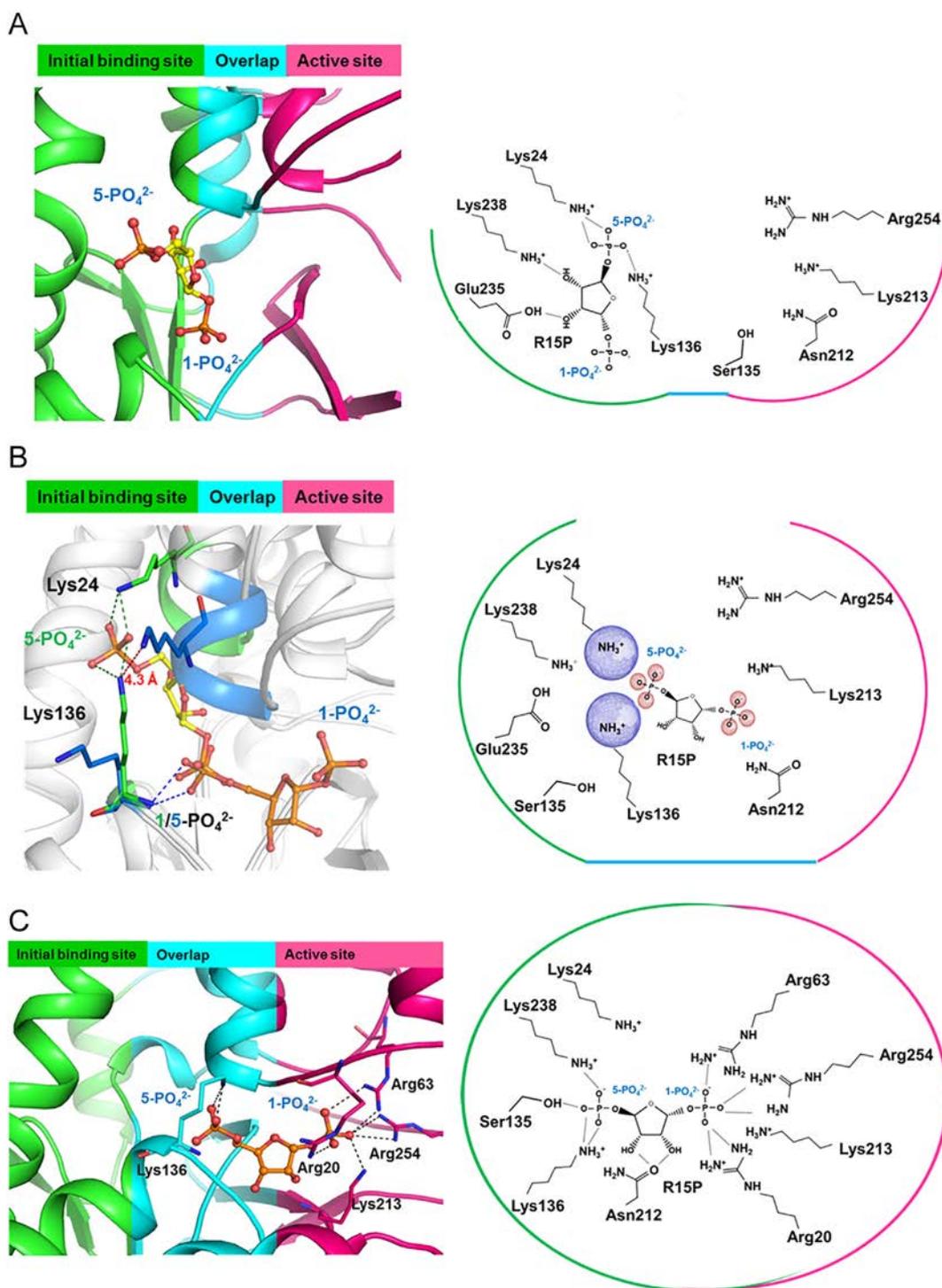
Fig. 10. Proposed mechanism of the "substrate sliding" in the enzyme R15Pi depicted in a three-dimensional model (left) and chemical scheme (right). (A) *Initial step*, R15P (yellow) binds at the "initial binding site" of R15Pi present in an open conformation. (B) *Intermediate step*, transition from open to closed conformation causes Lys24 to move forward resulting in the disruption of the bond between Lys24 and the 5-phosphate group of R15P at the "initial binding site". The forward movement of Lys24 also causes Lys136 to change its orientation. This changes in interaction and orientation of amino acid residues as well as spatial constraints in the closed state of R15Pi obliges R15P to slide into the active-site pocket. In the three-dimensional model, the residues Lys24 and Lys136 in open and closed conformation of R15Pi are highlighted in green and blue lines, respectively. The substrate R15P at the "initial binding site" and the active-site pocket are shown in yellow and orange, respectively. The interactions of R15P with Lys24 and Lys136 at the "initial binding site" and the active-site pocket are shown in green and blue dashes, respectively. The repulsion of Lys24 and Lys136 in the closed and open conformation of R15Pi, respectively, is shown in red dash with the distance indicated in red. The phosphate groups of R15P at the "initial binding site" and the active-site pocket are labeled in green and blue, respectively. In the chemical scheme, the repulsion between Lys24 and Lys136 as well as the spatial constraint between R15Pi and amino acid residues at the "initial binding site" is highlighted with blue and red spheres, respectively. (C) *Final step*, the closed conformation of the enzyme helps R15P to interact with the amino acid residues at the active-site pocket. The "initial binding site", overlapping region and the active-site pocket are shown in green, cyan and magenta, respectively. The open and closed conformation of the active-site pocket are represented in open and closed circles, respectively. All the amino acid residues are numbered according to the enzyme TkR15Pi.

prior to the folding of the protein and thus contributed towards the formation of the active-site pocket. Even though, the crystal structures of TkR15Pi contain low water molecules, three of the invariant water molecules are found to be present in 87% of the structures. Out of these, one of the invariant water molecules (IW4) interacts with the O3 oxygen atom while the other (IW5) interacts with the 5-phosphate group of the ligand. This indicates that among the five invariant water molecules that interact with the ligand, two (IW4 and IW5) are most essential for ligand binding at the active-site pocket. Another invariant water molecule IW12 is involved in holding a loop region that forms the "active-site floor". The presence of this water molecule in TkR15Pi protein hints towards the necessity of the water molecule for the stability of the loop.

Apart from the active-site pocket, invariant water molecules are also involved in providing structural integrity to the protein. They are present along the longest helix, α_5 which exhibits the most pronounced structural transition during the domain movement. The availability of water at this region might contribute towards the stability of the helix α_5 during the domain movement. The long loop region, which harbors one of the conserved active-site residues Arg254, is also held by invariant water molecules. Presence of invariant water molecules provides rigidity to the otherwise flexible loop and in turn helps in maintaining the structural integrity of the active site of the enzyme. The structural integrity of the active-site pocket is further maintained by two invariant water molecules which couples the long loop region with the NTD and CTD. This interaction also results in the formation of a more compact protein molecule. The invariant water molecule at the intersubunit region forms hydrogen bond(s) with other water molecule (s) concentrated in the void area created at the hexameric interface of R15Pi. These water molecules, in turn, interact with amino acid residues of the neighboring protomer. In this manner, a water mediated network is established that is involved in maintaining an intact hexamer.

An attempt to obtain a substrate-bound structure of R15Pi-WT resulted in the attainment of a product-bound structure [11,12]. This inhibited in achieving a clear picture of the initial stages of substrate (R15P) binding to the enzyme R15Pi in its native form. Molecular docking of substrate to R15Pi-Form I interestingly led to the identification of an "initial binding site" which overlaps with the active-site pocket. However, docking of the substrate to R15Pi-Form III did not result in binding of the R15P to the "initial binding site". This occurs due to the difference in the orientation of the two amino acid residues Lys24 and Lys136 in R15Pi-Form III, even though both forms are present in an open conformation. This difference in orientation of amino acid residues in R15Pi-Form III might be to ensure the binding of RuBP at the active-site pocket and inhibit its access to the "initial binding site".

The binding of the substrate R15P at the "initial binding site" hints towards a probable early binding event which could not be captured during the process of crystallization. Binding of R15P to the "initial binding site" might resemble the "Two-step model" proposed by Stank et al. [25] where a ligand first binds to an adjacent site transiently which, in turn, increases the possibility of its binding to its desired site. Similarly, in R15Pi, it might be possible that R15P initially binds at the "initial binding site" with one of the phosphate groups interacting with residues, which are part of the active-site pocket, thus, guiding R15P to the actual catalytic site. Another possibility might be to render binding specificity to the substrate over the product. As observed from the docking studies, RuBP shows the tendency of binding to the active-site pocket rather than the "initial binding site". Although, the conversion of R15P to RuBP is a reversible process, the subsequent step i.e. the conversion of RuBP to 3-PGA by the enzyme RuBisCO is irreversible and thus would constantly require the availability of RuBP. Henceforth, the binding of R15P at the "initial binding site" might help in the displacement of RuBP from the active-site pocket. However, it is very preliminary to arrive at such a conclusion based only on the docking studies. It would require a more detailed investigation supported by experimental evidences to prove the hypothesis.

To identify the amino acid residues which are crucial for ligand binding at the "initial binding site" and the active-site pocket, point mutations were performed. Mutation of Lys24 led to the displacement of R15P from the "initial binding site" to the active-site pocket indicating it to be the most crucial amino acid residue at the "initial binding site". At the active-site pocket, all the positively-charged conserved amino acid residues e.g. Arg20, Arg63, Lys213, Lys238 and Arg254 interacting with the 1-phosphate group are more crucial than the conserved amino acid residues Ser135, Lys136 and Ala137 interacting with the 5-phosphate group. The amino acid residues interacting with the 5-phosphate group act as an "anchor" to bring the substrate R15P into a close proximity of the active-site pocket while the actual binding site is determined by the amino acid residues which interacts with the 1-phosphate group of the ligand.

4.1. A Plausible Mechanism of "Substrate Sliding" From the "Initial Binding Site" to the Active Site of the Enzyme R15Pi

Binding of the substrate R15P to the "initial binding site" instead of directly to the active-site pocket have led us to propose a probable "substrate sliding" mechanism for the enzyme R15Pi. Initially, the substrate R15P binds to the "initial binding site" with the 1-phosphate group interacting with amino acid residues at the overlapping region (Fig. 10A). The binding of R15P to the "initial binding site" might compel the NTD to move towards the CTD. Owing to this conformational change, R15P can no longer maintain its interaction with the amino acid residues at the "initial binding site" and, in turn, is forced to slide into the active-site pocket. The most significant bond breakage due to the forward movement of the NTD occurs in between Lys24 and the 5-phosphate group of R15P. This also causes Lys24 to come into close proximity of Lys136 and as a result, the positive-positive charge repulsion compels Lys136 to change its orientation. The change in the orientation of Lys136 changes the interacting atoms of 5-phosphate group from side chain N^ζ atom to the main chain N atom (Fig. 10B). Simultaneously, forward movement of the NTD impels the other essential amino acid residues to come into close proximity to R15P. Once R15P forms a stable interaction with amino acid residues at the active-site pocket, Lys136 again changes its orientation in order to interact with the 5-phosphate group via its side chain atom N^ζ (Fig. 10C). This movement of R15P from the "initial binding site" to the active-site pocket can also not be ignored based on the change in orientation of Lys136 in the three different available forms of R15Pi.

In summary, our study provides an understanding of the different structural aspects of the enzyme R15Pi that are indispensable for its effective functioning. Furthermore, identification of an "initial binding site" in R15Pi and the proposition of a "substrate sliding" movement towards the active-site pocket provide a new perspective to the mechanism of action of the enzyme R15Pi.

Declarations of Interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2019.02.009.

References

[1] Lunt SY, Vander Heiden MG. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. Annu Rev Cell Dev Biol 2011;27:441–64.

[2] Stanton RC. Glucose-6-phosphate dehydrogenase, NADPH, and cell survival. IUBMB Life 2012;64:362–9.

[3] Stincone A, Prigione A, Cramer T, Wamelink Campbell K, Cheung E, Olin-Sandoval V, et al. The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. Biol Rev 2015;90:927–63.

[4] Grochowski LL, Xu H, White RH. Ribose-5-phosphate biosynthesis in *Methanocaldococcus jannaschii* occurs in the absence of a pentose-phosphate pathway. J Bacteriol 2005;187:7382–9.

[5] Soderberg T. Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes. Archaea 2005;1:347–52.

[6] Orita I, Sato T, Yurimoto H, Kato N, Atomi H, Imanaka T, et al. The ribulose monophosphate pathway substitutes for the missing pentose phosphate pathway in the archaeon *Thermococcus kodakaraensis*. J Bacteriol 2006;188:4698–704.

[7] Aono R, Sato T, Imanaka T, Atomi H. A pentose bisphosphate pathway for nucleoside degradation in Archaea. Nat Chem Biol 2015;11:355–60.

[8] Finn MW, Tabita FR. Modified pathway to synthesize ribulose 1, 5-bisphosphate in methanogenic archaea. J Bacteriol 2004;186:6360–6.

[9] Sato TH, Atomi H, Imanaka T. Archaeal type III RuBisCOs function in a pathway for AMP metabolism. Science 2007;315:1003–6.

[10] Aono R, Sato T, Yano A, Yoshida S, Nishitani Y, Miki K, et al. Enzymatic characterization of AMP phosphorylase and ribose-1,5-bisphosphate isomerase functioning in an archaeal AMP metabolic pathway. J Bact 2012;194:6847–68.

[11] Nakamura A, Fujihashi M, Aono R, Sato T, Nishiba Y, Yoshida S, et al. Dynamic, ligand-dependent conformational change triggers the reaction of ribose-1, 5-bisphosphate isomerase from *Thermococcus kodakarensis* KOD1. J Biol Chem 2012;287:20784–96.

[12] Gogoi P, Kanaujia SP. A presumed homologue of the regulatory subunits of eIF2B functions as ribose-1, 5-bisphosphate isomerase in *Pyrococcus horikoshii* OT3. Sci Rep 2018;8:1891.

[13] Abraham MJ, Murtola T, Schulz R, Pall S, Smith JC, Hess B, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1:19–25.

[14] Ponder JW, Case DA. Force fields for protein simulations. Adv Prot Chem 2003;66:27–85.

[15] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126:014101.

[16] Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 1981;52:7182–90.

[17] Darden T, York D, Pedersen L. Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. J Chem Phys 1993;98:10089–92.

[18] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103:8577–93.

[19] Hess B. P-LINCS: a parallel linear constraint solver for molecular simulation. J Chem Theory Comput 2008;4:116–22.

[20] Turner PJ. XMGRACE, Version 5.1.19. Beaverton, OR: Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology; 2005.

[21] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–32.

[22] Kanaujia SP, Sekar K. Structural and functional role of water molecules in bovine pancreatic phospholipase A2: a data-mining approach. Acta Crystallogr D Biol Crystallogr 2008;65:74–84.

[23] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res 2000;28:235–42.

[24] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 2009;30:2785–91.

[25] Stank A, Kokh DB, Fuller JC, Wade RC. Protein binding pocket dynamics. Acc Chem Res 2016;49:809–15.