Pyrophosphate release pathway in a sugar nucleotidyltransferase - GlmU.

Neha Vithani1*, Pravin Kumar Ankush Jagtap1*, Sunil Kumar Verma1*, Ravi Tripathi3, Shalini Awasthi3, Nisanth N. Nair3* and Balaji Prakash2*

*These authors contributed equally.

1Department of Biological Sciences and Bioengineering, Indian Institute of Technology, Kanpur 208016, India
2Department of Molecular Nutrition, CSIR-Central Food Technological Research Institute, Mysore, 570020, India.
3Department of Chemistry, Indian Institute of Technology, Kanpur 208016, India
* Authors for correspondence
Email: nnair@iitk.ac.in, balaji.prakash@cftri.res.in
Abstract: Sugar nucleotidyltransferases produce sugar-nucleotides and pyrophosphate using two catalytic Mg$^{2+}$ ions (Mg$^{2+}_A$ and Mg$^{2+}_B$). GlimU (N-acetylglicosamine-1-phoshate uridylyltransferase) is an SNT, important for bacterial cell-wall component synthesis. Here, we study product release pathway in GlimU, using structural and computational studies. In three different crystallographic snapshots of GlimU bound to products, we find that pyrophosphate and Mg$^{2+}_B$ occupy two distinct positions inside a tunnel. Umbrella sampling and temperature accelerated sliced sampling (TASS) simulations elucidate that pyrophosphate and Mg$^{2+}_B$ exit together through the tunnel. We also compute the free energy barrier for pyrophosphate-Mg$^{2+}_B$ release into the solvent. Pyrophosphate-Mg$^{2+}_B$ complex passes through two intermediate states, during its release. One of these intermediates was also captured in our crystallographic snapshots. The residues lining the tunnel stabilize the intermediate states. Conformational change of a tunnel lining residue Arg19 is identified to bring pyrophosphate-Mg$^{2+}_B$ out of the tunnel. Mg$^{2+}_B$ and residues lining the tunnel, in particular Arg19, are highly conserved in SNTs. Hence, the pyrophosphate-Mg$^{2+}_B$ release pathway identified here appears to be common in all SNTs.
GlmU is a member of sugar nucleotidyltransferase (SNT) family. SNTs catalyze nucleotidyl transfer reaction and produce a sugar-nucleotide and pyrophosphate (POP). Sugar-nucleotides participate in diverse biological processes such as cell growth, antigen synthesis, synthesis of cell-wall components in prokaryotes and plants, and sugar metabolism in eukaryotes.\(^1\,^2\,^3\) GlmU synthesizes a sugar-nucleotide called UDP-GlcNAc using the substrates UTP and GlcNAc-1-P (N-acetylglucosamine-1-phosphate).\(^4\,^5\) Like all other SNTs, GlmU requires Mg\(^{2+}\) ions for its nucleotidyltransfer reaction.\(^6\)

In our previous work, we determined a structure of GlmU bound to two Mg\(^{2+}\) ions (Mg\(^{2+}_A\) and Mg\(^{2+}_B\)) and the products UDP-GlcNAc (termed here as UD1) and POP.\(^7\) This structure for the first time depicted two Mg\(^{2+}\) ions at the active site of GlmU, with simultaneous presence of both the products. In that study, it was inferred that GlmU employs two Mg\(^{2+}\) ions (termed Mg\(^{2+}_A\) and Mg\(^{2+}_B\)) for the catalysis.\(^7\) Mg\(^{2+}_A\) was coordinated to two active site residues, two water molecules and two phosphate oxygens of UD1. Mg\(^{2+}_B\) was coordinated to three water molecules, a phosphate oxygen of UD1 and two phosphate oxygens of POP (Supporting information; Fig. S1B). It was proposed that Mg\(^{2+}_A\) stabilizes the substrates for the reaction and Mg\(^{2+}_B\) would stabilize the transition state. A report on another SNT called Kdo cytidyltransferase (KdsB) also depicted similar roles for the Mg\(^{2+}\) ions.\(^8\) The two-metal-ion mechanism proposed for SNTs is termed ‘Mechanism-B’. DNA/RNA polymerases that perform similar nucleotidyl transfer reaction on the growing nucleic acid chain also employ a two-metal-ion mechanism (termed ‘Mechanism-A’).\(^9\,^10\) Besides, POP is formed as one of the products in SNTs as well as DNA/RNA polymerases. ‘Mechanism-A’ and ‘Mechanism-B’ differ in the way two Mg\(^{2+}\) ions (Mg\(^{2+}_A\) and Mg\(^{2+}_B\)) are stabilized at the active site (Supporting information; Fig. S1). Mg\(^{2+}_B\) in ‘Mechanism-A’ makes one of its coordination interactions with an active site residue, while in ‘Mechanism-B’ it does not make any coordination interaction with the active site residues (Supporting information; Fig. S1). Despite these differences, the catalytic roles assigned to each Mg\(^{2+}\) ion remain the same.\(^10\,^11\) Computational studies on RNA polymerases have shown that POP is released as POP-Mg\(^{2+}_B\) complex.\(^12\,^13\) In the crystal structure of GlmU and a related SNT, it was seen that Mg\(^{2+}_B\) does not make any coordination interaction with the active site residues.\(^7\,^8\) Instead, it makes coordination interactions with the products UD1 and POP, and with water molecules. Based on this observation, it was previously speculated that Mg\(^{2+}_B\) might accompany POP during its exit.\(^8\)
Here, we investigate POP release pathway in GlmU and the participation of Mg\textsuperscript{2+}_B in this process. Product release is as important as product formation, as it could be a rate-determining step in enzymes where products are stabilized at the active site. In three distinct crystal structures of GlmU determined in complex with the products (UD1 and POP) and two Mg\textsuperscript{2+} ions, we find that Mg\textsuperscript{2+}_B and positively charged residues stabilize the products at the active site. These structural snapshots capture POP in complex with Mg\textsuperscript{2+}_B, in two distinct positions inside a positively charged tunnel. This suggested that POP-Mg\textsuperscript{2+}_B complex would exit the active site through the tunnel. To study the process of POP release from the active site into bulk solvent, we employed umbrella sampling (US)\textsuperscript{14,15} and temperature accelerated sliced sampling (TASS)\textsuperscript{16} simulations. These simulations elucidate that POP is released through the tunnel, as POP-Mg\textsuperscript{2+}_B complex. During its exit, POP-Mg\textsuperscript{2+}_B passes through two intermediate states. The residues lining the tunnel were identified to stabilize POP-Mg\textsuperscript{2+}_B in these intermediate states. One of the intermediates was also captured in our crystallographic snapshots. Computational studies in RNA polymerases have also depicted the release of POP-Mg\textsuperscript{2+}_B complex, via a few intermediate states.\textsuperscript{12,13}

A residue lining the tunnel, Arg19, was identified to be important for POP release into bulk solvent. In the intermediate states, Arg19 provides stabilizing interactions to POP inside the tunnel. Following this, Arg19 flips its side chain conformation and brings POP-Mg\textsuperscript{2+}_B out of the tunnel. Arg19 and other residues lining the tunnel are conserved in SNTs. Besides, Mg\textsuperscript{2+}_B is commonly employed by all SNTs.\textsuperscript{7} Therefore, the exit pathway and mechanism for POP release proposed in this study is also likely to be employed in other SNTs.

**Results and Discussion**

*Structural snapshots depicting pyrophosphate release in complex with a Mg\textsuperscript{2+} ion*

We have previously determined a crystal structure of GlmU bound to the products (UD1 and POP) and two Mg\textsuperscript{2+} ions (Mg\textsuperscript{2+}_A and Mg\textsuperscript{2+}_B).\textsuperscript{7} This structure (PDB ID: 4G87) with both the products bound at the active site, following the uridylation reaction, is termed GlmU[S1]. Here, S1 implies ‘snapshot 1’. GlmU[S1] was obtained by soaking GlmU[Apo] crystals with the substrates GlcNAc-1-P and UTP for ~ 45 minutes. In GlmU[S1], Mg\textsuperscript{2+}_B has an octahedral hexa-
coordination geometry. \( \text{Mg}^{2+}_B \) is coordinated to an oxygen of PA phosphate of UD1, two oxygen atoms of POP and three water molecules (Supporting information: Fig. S1B). Unlike \( \text{Mg}^{2+}_A \), \( \text{Mg}^{2+}_B \) does not make any coordination interaction with the active site amino acids. This has also been observed in the crystal structure of another sugar nucleotidyltransferase KdsB, based on which it was speculated that \( \text{Mg}^{2+}_B \) would exit together with POP.\(^8\)

**Figure 1.** Tunnel like structure (front view), enclosed by T1 and T2 regions, in GlmU[S1]. T1 and T2 regions are shown by blue and red colored ribbons, respectively, together with their surface representation. POP (shown in stick representation) is occupied in the tunnel. Interactions of POP (red colored dotted line) and \( \text{Mg}^{2+}_B \) (blue colored dotted line) are depicted. The interacting residues, UD1 and water molecules (\( W_1 \), \( W_2 \) and \( W_3 \)) are shown in ball and stick representation. Color code: gray (carbon), red (oxygen), blue (nitrogen), green (\( \text{Mg}^{2+}_B \)) and white (hydrogen).
Analyzing the structure GlmU[S1], we find that POP is snugly occupied in a tunnel like structure (Fig. 1). A part of this tunnel is lined by positively charged amino acids (Arg19 and Ly26), which are also important for binding and stabilization of the substrate UTP. We term this region of tunnel as T1 region. T1 region is formed by the backbone amide of Gly17, Thr18, Arg19 and the side chains of Thr18, Arg19 and Lys26 (Fig. 1). POP makes H-bond interactions with the backbone amide of Gly17, Thr18, Arg19 and the side chains of Thr18, Arg19 and Lys26. Additionally, coordination interactions of Mg$^{2+}$B with both the products (UD1 and POP), stabilize them together at the active site. Mg$^{2+}$B also makes coordination interactions with three water molecules. Two of these water molecules are stabilized by Glu207 side chain (Fig. 1). Glu207 is present in a loop that forms the upper part of the tunnel. This region is termed as T2. Thus, T1 and T2 regions together form the entire exit tunnel and contribute stabilizing interactions for POP and Mg$^{2+}$B (Fig. 1). The shape and charge complementarity, between POP and the pocket suggested that free diffusion of the POP following product formation, would be prohibited. We presumed this to be a likely reason for the retention of POP at the active site.

Encouraged by this possibility, we collected diffraction datasets for GlmU$^{\text{Mtb}}$[Apo] crystals soaked in substrates GlcNAc-1-P and UTP for varying time, to monitor subsequent events, if possible. Serendipitously, amongst the several dataset collected, we could obtain three different crystal structures showing different positions of POP (termed GlmU[S2], GlmU[S3] and GlmU[S4]) for crystals soaked in substrates for 4.5 hours, 5 hours and 12 hours respectively.

A comparison of the structures GlmU[S1], GlmU[S2], GlmU[S3] and GlmU[S4] suggests that they represent snapshots of POP release process. POP was seen to occupy two distinct positions – position A and position B – in these snapshots (Fig. 1 & 2). In all of these, POP remains coordinated to Mg$^{2+}$B. As detailed above, GlmU[S1] represents a structure where products UD1 and POP are retained at the active site, in a position immediately following their formation (Fig. 1). Interestingly, in the snapshot GlmU[S2], POP and Mg$^{2+}$B were found in two distinct positions. POP occupying the position A is termed APOP, while POP found in the position B is termed BPOP (Fig. 2A). In GlmU[S2], APOP and BPOP were modeled with occupancy of ~0.5 for each, (Supporting information; Fig. S2A). Corresponding to these, Mg$^{2+}$B was also modeled in two distinct positions with occupancies ~0.5 for each (Supporting information; Fig. S2A). APOP superposes well with POP seen in GlmU[S1], while BPOP has its position shifted by ~3 Å with respect to APOP. In the snapshot GlmU[S3], POP and Mg$^{2+}$B
occupy a single position and they were modeled with an occupancy 1.0 (Supporting information; Fig. S2B). Although not identical, POP in GlmU[S3] overlaps well with BPOP seen in GlmU[S2] (Fig. 2B).

Figure 2. UD1, POP, Mg$^{2+}_A$, and Mg$^{2+}_B$ bound at the active site in the snapshots GlmU[S2], GlmU[S3] and GlmU[S4]. (A) APOP and BPOP in GlmU[S2] are shown with their respective Mg$^{2+}_B$ (shown as green and cyan colored spheres, respectively). (B) Structural overlap of BPOP of GlmU[S2], POP of GlmU[S3] and GlmU[S4] (shown in cyan, yellow and magenta, respectively). Mg$^{2+}_B$ are shown as spheres in the same color as that of their respective POP. T1 (shown in pink color) and T2 (shown in blue color) regions are shown in surface representation.

In order to capture a state where POP has moved beyond the position B, structures were determined by soaking the crystals with the substrates for longer times, i.e. 12, 20 and 60 hours. However, all of these snapshots show presence of POP in the position B. This depicts inability of POP to completely exit the tunnel. The snapshot corresponding to a 12 hours soak, GlmU[S4] alone is presented here. In GlmU[S4], POP occupies the position B (Fig. 2B). In the position B, POP is stabilized by T1 region through its interactions with backbone amide of Gly17, Thr18, Arg19 and the side chains of Thr18 and Arg19. In this position, however, the interaction with Lys26 is absent. One of the waters coordinating with Mg$^{2+}_B$ is stabilized by Gln205 and Glu207 present in T2 loop. Thus, T1 and T2 regions continue to provide stabilization to POP as it moves outwards.

An analysis of the crystal contacts in these snapshots reveals that the path by which POP would exit the pocket is occluded by a surface patch of negative potential from a neighboring
symmetry related molecule in the crystal (Supporting information; Fig. S3A, S3B). This is likely to prevent negatively charged POP from leaving the pocket. Thus, in all the snapshots, POP moves only up to the position B and does not completely exit the active site. Consecutively, the soaking time of crystals with the substrates does not reflect or correlate to the real time durations for POP release. In the absence of the occlusion by a symmetry related molecule, as would be the case in solution, POP would exit through the tunnel into the solvent.

In order to study the mechanism of complete release of POP from the active site of GlmU into the solvent, we employed MD simulations combined with US and TASS\textsuperscript{16} simulations. MD simulations were carried out using GlmU[S1] as the initial coordinates (PDB ID: 4G87) which represents the state where POP has just formed. In this structure, POP was retained at the active site in complex with Mg\textsuperscript{2+} ion. We carried out equilibrium MD simulations for the solvated enzyme prior to performing US and TASS simulations (See ‘Methods and Models’ for further details). Root mean square deviation (RMSD) of the backbone atoms measured with respect to the original crystal structure (GlmU[S1]) is less than 1.6 Å during 5 ns NVT simulation (Supporting information; Fig. S5A). RMSD for active site residues Gly17, Thr18, Arg19, Lys26, Ala238 and Asn239 (which are important for POP and Mg\textsuperscript{2+} ion stabilization) is less than 1.6 Å. Similarly, RMSD measured for the products (UD1 and POP) and two Mg\textsuperscript{2+} ions is less than 1.6 Å during 5 ns NVT simulation (Supporting information; Fig. S5). During 5 ns NVT simulation, Mg\textsuperscript{2+}\textsubscript{A} and Mg\textsuperscript{2+}\textsubscript{B} retained a stable hexa-coordination geometry, which is also seen in the crystal structure. Mg\textsuperscript{2+}\textsubscript{A} has one coordination interaction with Asp114, one with Asn239, two with water molecules and two with UD1. Similarly, the hexa-coordinated geometry of Mg\textsuperscript{2+}\textsubscript{B} formed by its interactions with three water molecules, UD1 and POP remains stable. Stabilizing interactions of POP with the residues Gly17, Thr18, Arg19 and Lys26 provided by T1 region remain intact throughout the simulation. The interactions of POP with T1 region of the tunnel include: (i) H-bond interactions made by P1 phosphate of POP with the backbone amide of Gly17 and the side chain of Lys26 and, (ii) H-bond interactions made by P2 phosphate with the side chains of Thr18 and Arg19.

**The complete release of pyrophosphate into the solvent**

In order to examine entire process of POP release from the active site into the solvent, we carried out US simulations, starting from the structure obtained from the last frame of NVT
US simulations were carried out using $d[PA-POP_{\text{COM}}]$ as the collective variable (CV). $d[PA-POP_{\text{COM}}]$ is the distance between PA phosphorous of UD1 and the center of mass of P1, O$_b$ and P2 of POP (Fig. S6). P1 and P2 are two phosphorus atoms of POP and O$_b$ is the bridging oxygen between P1 and P2. Umbrella windows were placed for the distance $d[PA-POP_{\text{COM}}]$, in the range of 5.4 Å to 21 Å. The restraint potential applied for each umbrella is listed in Table S3. In each umbrella window, simulations were run for 30 ns.

Free energy profile obtained from the US simulations shows that before POP is completely released in bulk solvent, it passes through two intermediate states - S2 and S3 (Fig. 3A). The initial state S1 is the most stable and represents the state where UD1, POP and two Mg$^{2+}$ ions are bound at the active site, following uridyltransfer reaction. Structures obtained from the umbrella corresponding to S1, highly resemble with the snapshot GlmU[S1] in terms of the position occupied by POP and its interactions with T1 region (Fig. 4A, Table S4). Mg$^{2+}$ remains coordinated to POP in S1 state. Glu207 from T2 stabilizes the water molecules coordinating to Mg$^{2+}$. In S1, POP makes interactions with the side chains of Thr18, Arg19 and Lys21, and the backbone of Gly16 and Arg19. These interactions with T1 and T2 are also present in GlmU[S1]. Comparison of the interactions observed in S1 and GlmU[S1] is provided in Table S4 (Supporting information; section 3.2).

From S1, POP proceeds to an intermediate state S2, crossing free energy barrier of 5 kcal mol$^{-1}$. Interestingly, the position occupied by POP in S2 is similar to the position B occupied by POP in GlmU[S2], GlmU[S3] and GlmU[S4] snapshots (Fig. 4B). During the transition from S1 to S2, POP retains its coordination interactions with Mg$^{2+}$B. This is in accordance with the crystal structures GlmU[S2], GlmU[S3] and GlmU[S4], where POP is always observed to make coordination interactions with Mg$^{2+}$B (Fig. 2). However, the coordination interactions between POP and Mg$^{2+}$B in the S2 state are not identical to those seen in the structural snapshots (Supporting information; section 3.3). One of the water molecules coordinating Mg$^{2+}$B is stabilized by Glu205 present in the T2 region of the exit tunnel. Stabilizing interactions of Glu205 and Gln207 with the water molecules coordinating Mg$^{2+}$B are also observed in GlmU[S2], GlmU[S3] and GlmU[S4]. In S2, POP is stabilized by H-bond interactions with the backbone amides and side chains of Thr18 and Arg19 of T1 region. The same residues are observed to make stabilizing interactions with POP (in position B) in snapshots GlmU[S2], GlmU[S3] and GlmU[S4] (Supporting information; section 3.2). Unlike in GlmU[S2],
GlmU[S3] and GlmU[S4], POP in S2 retains its H-bond interaction with Lys26. Comparison of the interactions made by POP in S2 and the snapshots GlmU[S2], GlmU[S3] and GlmU[S4] is provided in Table S5 (Supporting information; section 3.2).

Figure 3. Free energy barrier and pathway for POP-Mg$_{2+}$ release obtained from the US simulations. (A) Free energy profile for release of POP-Mg$_{2+}$ from S1 (the initial state) into the bulk solvent is shown. The intermediate states (S2 and S3), structures on the plateau (P1, P2, P3) and the state PE where POP is released into the solvent are marked on the free energy profile. (B) POP, Mg$_{2+}$ and the interacting tunnel residues are shown for the states S2 (blue), S3 (red), P1 (cyan), P2 (magenta), P3 (green) and PE (yellow). For each state, only the residues interacting with POP and Mg$_{2+}$ are shown in the transparent sticks. The water molecules coordinating to Mg$_{2+}$ are not shown for the purpose of clarity. T1 and T2 region are shown in ribbon form.
Figure 4. (A) Comparison of the S1 state with GlmU[S1]. UD1, POP, Mg\(^{2+}\)A and Mg\(^{2+}\)B bound at the active site of GlmU[S1] are shown in thicker ball and stick representation, while those obtained from umbrella window corresponding to the S1 state are shown as transparent ball and stick. (B) Comparison of the S2 state with GlmU[S3]. Products and two Mg\(^{2+}\) ions from GlmU[S3] are shown in thicker ball and stick form, while those from US simulation are shown in transparent ball and stick form. POP and Mg\(^{2+}\)B are shown for the frames obtained from US simulations performed for 30 ns.

For the reasons stated in section ‘Structural snapshots depicting pyrophosphate release in complex with a Mg\(^{2+}\) ion’, structural studies could not capture the movement of POP beyond the position B i.e. beyond the state S2. In our US simulations, we find that POP release beyond S2 is the rate determining step. Transition of POP from the state S2 to S3 requires crossing a free energy barrier of ~9 kal mol\(^{-1}\) (Fig. 3A). This could be due to breaking and rearrangement of several H-bonds between POP and tunnel residues during this transition (Supporting information; section 4). H-bond interactions of POP with Thr18 backbone amide and Lys26 side chain of T1 region break when it passes from S2 to S3 state. During this transition, Arg19 side chain flips from an inward facing conformation to a conformation facing towards the bulk solvent (Fig. 3 & Fig. 5). Rotation of Arg19 side chain appears to drive POP from S2 to S3 state. As a result of Arg19 side chain rotation and the movement of POP, H-bond interactions of POP with Arg19 are rearranged. Specifically, Arg19 side chain breaks it H-bonds with P1 phosphate and reforms H-bonds with P2 phosphate of POP (Supporting information; section 4).
Figure 5. Rotation of Arg19 side chain during POP release from the S2 to S3 state. (A) Arg19 side chain and POP are shown for the S2 (green color), S3 (magenta color) and for the transition state (gray color) between S2 and S3 states. UD1 is shown in stick representation. (B) Probability distribution of $d[O_{UD1} - C_{Arg19}]$, the distance between UD1 oxygen and side chain Cz carbon of Arg19 for the umbrella windows for the S2 (green), S3 (magenta) states and the transition state for S2 to S3 transition i.e. TS$_{S2-S3}$ (black). This distance is labeled in the figure 5A. For the TS$_{S2-S3}$ umbrella window, two peaks are seen for $d[O_{UD1} - C_{Arg19}]$, representing Conf-A and Conf-B for Arg19 side chain. (C) Reconstructed free energy surface from TASS simulation for S2 to S3 transition. The free energy surface is projected as a function of two CVs, $d[PA-POMB]$ and $\chi_2^{Arg19}$ (i.e. TORSION[C$_B$:C$_6$:C$_D$:Ne]). $\chi_2^{Arg19}$ is defined by rotation around CG-CD bond of Arg19 side chain (shown in figure 5A with black curved arrow). Contour lines are drawn at every 1 kcal mol$^{-1}$. The minimum energy path for S2 to S3 transition is shown by dotted lines.

Following the state S3, POP passes through the states on a free energy plateau, before its complete release into the bulk solvent (Fig. 3). In these states, POP retains its interactions with Arg19 side chain, while rest of its interactions with T1 and T2 regions breaks (Fig. 3). Additionally, POP now makes H-bond interactions with Arg21 which is also a part of T1 region. Beyond the plateau, interactions of POP with Arg19 break. However, its interactions with the
side chain of Arg21 are retained. Thus, POP passes from one positively charged residue to another on the surface of T1 region (Fig. 3), before it is completely released into bulk solvent. We anticipate that the same Arg residues, Arg19 and Arg21, would participate in substrate (UTP) binding. These positively charged residues - Arg19 (in its outside flipped conformation) and Arg21 can initially interact with triphosphate group of the substrate UTP at the gate of the tunnel. Later, Arg19 can pull UTP into the tunnel, by flipping its side chain conformation.

In US simulations, convergence of free energy profile can be slow if the auxiliary CVs such as H-bond interactions of POP with the tunnel residues are not sampled properly. This could result in erroneous free energy estimates in US simulations. Besides, there is also a possibility of POP-Mg$^{2+}$ complex dissociation during the process of POP release. In order to take into account all of these interactions of POP with tunnel residues and Mg$^{2+}$ ion, one needs to sample a large number of CVs along with $d[PA$-$POP_{COM}]$. Such extensive sampling would not be possible with US simulations. Thus, we employ TASS technique that allows sampling of a large number of CVs without requiring long simulations.\textsuperscript{16} As we found from our US simulations that there is major energy barrier (9 kcal mol$^{-1}$) during S$_2$ to S$_3$ transition which requires rotations of Arg19 side chain and rearrangement of several H-bonds between POP and tunnel residues, we carried out TASS simulations for S$_2$ to S$_3$ transition. For this purpose, $d[PA$-$POP_{COM}]$ was sampled by umbrella bias, where the sampling window is from 6.1 to 9.0 Å (Supporting information; Table S3). Additionally, torsional angles of Arg19 side chain, H-bond interactions of POP with Arg19 and Lys26, and coordination interactions between POP and Mg$^{2+}$ were sampled at a higher temperature (600 K) in each umbrella window (Supporting information; section 5). Free energy profile obtained from TASS simulations, performed for 30 ns per umbrella window, provides the free energy barrier of 7 kcal mol$^{-1}$ for S$_2$ to S$_3$ transition (Fig. 5C, Supporting information; Fig. S9). The free energy surface, reconstructed as a function of $d[PA$-$POP_{COM}]$ and Arg19 side chain dihedral angle, suggests that Arg19 rotation would precede the movement of POP-Mg$^{2+}$ from S$_2$ to S$_3$. The minimum energy path, denoted by dotted line on the free energy profile (Fig. 5C), shows that Arg19 side chain, in the S$_2$ state, rotates and leads to S$_2'$ state which can facilitate the transition to S$_3$ state would precede the movement of POP-Mg$^{2+}$ from S$_2$ to S$_3$. Sampling of coordination interactions between POP and Mg$^{2+}$ in TASS simulations confirms that POP-Mg$^{2+}$ complex always remains intact. The free energy barrier obtained from US and TASS simulations, for complete release of POP from
the active site pocket into bulk solvent, is 18 kcal mol\(^{-1}\) (Supporting information; Fig. S9). With such high energy barrier, product release appears to be a rate limiting factor in the catalytic cycle of the nucleotidyltransfer reaction of GlmU enzyme.

Previously, the movement of Mg\(^{2+}\) ion during product release was speculated in an SNT.\(^8\) Metal ion mediated product release has also been suggested for glycosyltransferases, which share structural homology to SNTs.\(^{17,18}\) A study on phosphoribosyltransferase enzymes has also elucidated that POP release occurs in complex with two Mg\(^{2+}\) ions.\(^{19}\) MD simulation studies carried out on RNA polymerases demonstrated that POP exits a tunnel in complex with a Mg\(^{2+}\) ion (the one that is analogous to Mg\(^{2+}\)\(_B\) of SNT).\(^{12,13}\) In this study, we provide structural evidence and molecular mechanism for the POP release in complex with Mg\(^{2+}\)\(_B\).

**Conservation of Mg\(^{2+}\)\(_B\) and the tunnel in SNTs.**

POP-Mg\(^{2+}\)\(_B\) release along a tunnel occurs via two intermediate states which are stabilized by the residues present in T1 and T2 regions lining the exit tunnel. T1 region, formed by the side chains of Arg19, Lys26 and the backbone amides of Gly17, Thr18 and Arg19, provide the positively charged surface for stabilization of POP in GlmU. These residues are conserved (Supporting information; Fig. S10) and form a signature motif L-X\(_2\)-G-X-G-T-R-M-X\(_4\)-P-K common to all SNTs.\(^{20}\) We identify that Arg19 also plays a crucial role in POP release by driving POP out of the tunnel.

We have previously shown that the two-metal ion mechanism identified in GlmU would be operative in all SNTs.\(^7\) We divided SNTs into two major groups Group-I and Group-II, based on the presence or absence of Mg\(^{2+}\)\(_A\) in SNTs.\(^7\) Group-II SNTs employs a single metal ion for catalysis, where the role of Mg\(^{2+}\)\(_A\) is replaced by a lysine residue. However, Mg\(^{2+}\)\(_B\) is invariably recruited for nucleotidyltransfer by both Group-I and Group-II SNTs.\(^7\) Mg\(^{2+}\)\(_B\) and the tunnel residues, more importantly Arg19, are conserved across all SNTs (Fig. S10). Therefore, we propose that POP release in complex with Mg\(^{2+}\)\(_B\) would follow the same pathway and mechanism in all SNTs.

In DNA and RNA polymerases too, two Mg\(^{2+}\) ions (Mg\(^{2+}\)\(_A\) and Mg\(^{2+}\)\(_B\)) are utilized for the catalysis. Computational studies on POP release mechanism in RNA polymerase II and prokaryotic RNA polymerase suggest that POP exits as POP-Mg\(^{2+}\)\(_B\) complex.\(^{12,13}\) In phosphoribosyltransferases also, POP release occurs in complex with two Mg\(^{2+}\) ions.\(^{19}\) This
suggests that the POP release accompanied by Mg\(^{2+}\) ion(s) could be a common theme in a diverse class of enzymes that perform phosphoryl transfer reactions in the presence of Mg\(^{2+}\) ions.

**Conclusions**

We elucidate the process of POP release from the active site of GlmU – a nucleotidytransferase enzyme, using structural and computational study. In the product bound structure of GlmU, POP was observed to make coordination interactions with one of the catalytic Mg\(^{2+}\) ions (Mg\(^{2+}\)\(_B\)). The simulations presented here depict that POP exits the active site in complex with Mg\(^{2+}\)\(_B\) ion through a tunnel. Release of POP-Mg\(^{2+}\) has been observed through computational studies in diverse enzymes catalyzing phosphoryltransfer reactions.\(^{12,13,19}\) This study provides a structural evidence for this phenomenon. The complete release of POP-Mg\(^{2+}\)\(_B\) complex in bulk water occurs via two intermediate states. One of these intermediates is also captured in our crystallographic snapshots. Based on structural snapshots and the simulations, we identify the tunnel residues stabilizing the intermediates. A conserved arginine (Arg19 in GlmU), which also lines the exit tunnel, is identified to drive POP release from the tunnel into the bulk water. POP release in SNTs appears to be the rate determining step in the catalytic cycle of the nucleotidyl transfer reaction; further investigation of the rate of the chemical transformation per se would be required for a conclusive remark.

**Methods and Models**

**Crystallization and structure determination**

*a) Crystallization*

The wild type GlmU from *Mycobacterium tuberculosis* was purified and crystals of GlmU in apo (termed GlmU[Apo]) were obtained as described previously.\(^{21}\) Three different snapshots of GlmU bound to UD1, POP and two Mg\(^{2+}\) ion - GlmU\(^{\text{Mtb}[S2]}\), GlmU\(^{\text{Mtb}[S3]}\) and GlmU\(^{\text{Mtb}[S4]}\) were obtained by soaking GlmU[Apo] crystals in a soak solution at 4°C for 4.5 hours, 5 hours and 12 hours, respectively. The soak solution consists of 10% PEG8000, 100 mM HEPES (pH 7.5), 20 mM MgCl\(_2\), 4 mM CoCl\(_2\), 10 mM UTP and 10 mM GlcNAc-1-P. The ligand bound crystals were cryo-protected in a cryo solution consisting of 20% ethylene glycol, 15% PEG8000, 100 mM HEPES (pH 7.5), 20 mM MgCl\(_2\) and 4 mM CoCl\(_2\) prior to X-ray diffraction
experiments and data collection.

b) Data collection and processing
X-ray diffraction data were collected using an in-house Rigaku MicoMax007HF X-ray source equipped with a copper rotating anode, Varimax optics, Mar345dtb image plate detector and Oxford cryosystem 700 series cryostream. XDS program package was used to index, integrate and scale the data. The crystals belong to H3 (146) space group. The solvent content of the crystals was ~58% with one protein molecule per asymmetric unit. The details of structure determination and refinement are provided in the supporting information (section 1). The structure factors and atomic coordinates for GlmU[S2], GlmU[S3] and GlmU[S4] have been deposited in the Protein Data Bank (PDB) with accession codes 4G3S, 4G3P and 4G3Q, respectively.

Molecular Dynamics (MD) Simulations
a) System setup and methods
The initial coordinates for the classical simulation were obtained from the crystal structure of GlmU determined in complex with both the products (UD1 and POP) and two Mg$^{2+}$ ions (PDB ID: 4G87). In this structure, the products are retained at the active site of GlmU.\(^7\) See supporting information (section 2.1), for additional details on system setup. The enzyme was solvated with 14453 TIP3P water molecules within a periodic box of dimension 70×88×96 Å$^3$. The system was neutralized by adding 15 Na$^+$ counter ions. The whole protein was treated with parm99 version of the AMBER force field.\(^22\) Force field parameters for the ligands (UD1, POP and two Mg$^{2+}$ ions) were derived from the GAFF force field\(^23\) which is compatible with standard AMBER force field, as available in the AMBER suite of programs.\(^24\) Missing force field parameters for the ligands were not present in the GAFF force field, and were derived using Antechamber\(^25\) tool available in AMBER package (parameters provided in Supporting information; section 2). RESP charges were calculated for UD1, POP and two Mg$^{2+}$ ions by the R.E.D.\(^26-28\) package. See supporting information (section 2) for details of the RESP charge calculation. Classical MD simulations were carried out using the AMBER suite of programs.\(^24\) A 12Å cutoff distance was used while computing the non-bonded interactions. The long range electrostatic interactions were computed using the particle mesh Ewald (PME) method.\(^29\) After the preliminary minimization of
the water molecules and the subsequent minimization of the whole system, NPT simulations were carried out for 1 ns at 300 K using Langevin thermostat\textsuperscript{30} and 1 atm using Berendsen barostat. This was followed by 5 ns NVT simulation at the equilibrium volume. The time step of 1 fs was used for integrating the equations of motion.

\textit{b) Umbrella sampling (US) simulations}

US simulations were carried out using GROMACS program\textsuperscript{31} with PLUMED plugin\textsuperscript{32}. The CV used for US is $d[PA-\text{POP}_{\text{COM}}]$ – the distance between PA phosphorous of UD1 and the center of mass of P1, O\textsubscript{b} and P2 of POP (Supporting information; section 3.1). The umbrella windows were placed for $d[PA-\text{POP}_{\text{COM}}]$ CV in the range of 5.4 Å to 21.0 Å. The details of restraint parameters used for each umbrella are provided in the supporting information, section 3.1 (Table S3). The starting structure for the first window set at 5.4 Å was obtained from the NVT simulation run for 5 ns (described above). For rest of the umbrellas, the starting structure was obtained from adjacent umbrella equilibrated for ~ 1 ns. In each umbrella, NVT simulation was performed for 30 ns.

\textit{c) Temperature accelerated sliced sampling (TASS) simulations}

TASS has been recently developed in our group which combines the techniques of US, metadynamics and temperature accelerated sampling for enhanced sampling of a large number of CVs.\textsuperscript{16} Details of this method can be found in Ref.\textsuperscript{16} In the present study, we have used TASS method without adding any external bias potential i.e. only US and temperature accelerated sampling were combined. The parameters and starting structures used for each umbrella window remain the same as described above in US simulations. Additionally, orthogonal CVs (listed in section 5, supporting information) were sampled at 600 K temperature in each umbrella. For this, each CV was linked to an auxiliary variable using a mass $\mu_a$ of 100 a.m.u and $k_a$ of $2.5 \times 10^3$ kcal mol$^{-1}$. Auxiliary variables associated with each CV were sampled at 600 K using Langevin thermostat. Simulations were performed for 30 ns, in each umbrella.

\textbf{Acknowledgement}

US and TASS simulations were performed using the High Performance Computing (HPC) facility at IIT Kanpur.
References

(18) Hosfield, D. J., Zhang, Y., Dougan, D. R., Broun, A., Tari, L. W., Swanson, R. V., and
Finn, J. (2004) Structural Basis for Bisphosphonate-mediated Inhibition of Isoprenoid Biosynthesis. J. Biol. Chem. 279, 8526–8529.


