

A conserved structural role for the Walker-A lysine in P-loop containing kinases

Fatlum Hajredini^{1,2} and Ranajeet Ghose^{1,2,3,4*}

¹Department of Chemistry and Biochemistry, The City College of New York, New York, NY 10031.

PhD Programs in ²Biochemistry, ³Chemistry and ⁴Physics, The Graduate Center of CUNY, New York, NY 10016.

*Correspondence: rgghose@ccny.cuny.edu

Word Count: 3902

Abstract

Bacterial tyrosine kinases (BY-kinases) and shikimate kinases (SKs) comprise two structurally divergent P-loop containing enzyme families that share similar catalytic site geometries, most notably **with respect to** their Walker-A, Walker-B, and DxD motifs. We had previously demonstrated that in BY-kinases, a specific interaction **between** the Walker-A and Walker-B motifs, driven by the conserved “catalytic” lysine housed on the former, leads to a conformation that is unable to efficiently coordinate $\text{Mg}^{2+}\cdot\text{ATP}$ and is therefore incapable of chemistry. Here, using enhanced sampling molecular dynamics simulations, we demonstrate that structurally similar interactions between the Walker-A and Walker-B motifs, also mediated by the catalytic lysine, stabilize a state **in SKs** that deviates significantly from one that is necessary for the optimal coordination of $\text{Mg}^{2+}\cdot\text{ATP}$. This structural role of the Walker-A lysine is a general feature in SKs and is found to be present in members that encode a Walker-B sequence characteristic of the family (*Coxiella burnetii* SK), and in those that do not (*Mycobacterium tuberculosis* SK). Thus, the structural role of the Walker-A lysine in **stabilizing** an inactive state, distinct from its catalytic function, is conserved between two distantly related P-loop containing kinase families, the SKs and the BY-kinases. The universal conservation of this element, and of the key characteristics of its associated interaction partners within the Walker motifs of P-loop containing enzymes, suggests that this structural role of the Walker-A lysine is perhaps a widely deployed regulatory mechanism within this ancient family.

Introduction

P-loop containing proteins are widespread in all three branches of the tree of life suggesting that they emerged prior to the last **universal** common ancestor (LUCA) (Aravind et al., 2002; Edwards et al., 2013; Longo et al., 2020). **Nucleotide triphosphatases (NTPases)** (Leipe et al., 2002) and P-loop kinases constitute two of several divergent families of P-loop containing enzymes that encode sequences known as Walker-A (GxxxGK[S/T], x is any residue) and Walker-B ($\phi\phi\phi\phi$ [D/E]xxG, ϕ is a hydrophobic residue) motifs, or variations thereon (Walker et al., 1982). The BY-kinases (for Bacterial tYrosine kinases), that are widely conserved within the prokaryotic kingdom, but are without archaeal or eukaryotic counterparts (Shi et al., 2014), constitute a unique family of protein tyrosine kinases within the SIMIBI (named for the signal-recognition GTPases, MinD and BioD superfamilies) class of NTPases (Leipe et al., 2002). Instead of the dual-lobed fold characteristic of the eukaryotic protein kinases (Taylor and Radzio-Andzelm, 1994; Scheeff and Bourne, 2005), the catalytic domains (CDs) of BY-kinases (Lee and Jia, 2009; Grangeasse et al., 2012) utilize a fold that closely resembles ATPases of the MinD family. While BY-kinase CDs may be classified as P-loop kinases based on their function, they deviate significantly from the P-loop kinase superfamily (comprising kinases that phosphorylate small molecules e.g., nucleosides, shikimate, 6-phosphofructose etc.) as defined by Leipe *et al.* (Leipe et al., 2003) (we will adhere to that definition here), notably through the absence of a “LID” domain present in the latter (Figure S1A). In addition to modified Walker-A and Walker-B motifs, BY-kinase CDs also encode a sequence, $\phi\phi\phi\phi$ DxDxR, that has been termed the Walker-A’ motif (Grangeasse et al., 2012). The DxD sequence present in the Walker-A’ motif is also found in many members of the SIMIBI class of P-loop enzymes including MinD (Figure S1B). Interestingly, the DxD motif, is characteristic of many P-loop kinases including those that utilize shikimate, gluconate or adenylyl

phosphosulfate (APS) as substrates (Leipe et al., 2003). A key feature of these “DxD-group” P-loop kinases is a Walker-B motif that is degraded at its C-terminal end and marked by the conspicuous absence of the acidic Asp (or Glu) residue (Figure S1C) (Leipe et al., 2003). While the BY-kinase CD and P-loop kinases deviate significantly in their three-dimensional structures, they display similar catalytic site geometries formed by closely related structural motifs (Figures S1C). Given the close similarities in the catalytic site architecture despite being embedded in divergent structural scaffolds, and the fact that both families perform similar phospho-transfer chemistry, we wondered whether these two distantly related families retain common structural/dynamic features that characterize their functional states.

We have previously demonstrated the unique conformational dynamics of the catalytic core of the archetypal BY-kinase, *Escherichia coli* (K12) Wzc (Wzc_{CDAC}); these dynamics appear to have significant functional consequences. Our studies revealed the presence of two major conformations with distinct long- and short-range structural features (Hajredini et al., 2020) in Wzc_{CDAC}. We designated these conformations as open (open state, OS) and closed (closed state, CS) based on their global compactness. The OS has high affinity for ADP but is unable to efficiently engage ATP•Mg²⁺. The CS, that can optimally engage ATP•Mg²⁺, is stabilized upon formation of a closed octameric ring seen both in the structure of the isolated CD (Bechet et al., 2010) as well as in the membrane-anchored full-length protein (Yang et al., 2021). Oligomer formation is necessary to enable intermolecular autophosphorylation. Thus, the differential nucleotide preferences of the OS and CS together with the coupling of their relative stabilities to oligomerization appear to play a critical role in maintaining optimal Wzc function through modulation of the substrate affinity and the prevention of futile ATP hydrolysis (Hajredini et al., 2020). A key structural feature that was found to distinguish the OS from the CS involved an

alteration in the mode of interaction between conserved Walker-A and Walker-B residues (Hajredini et al., 2020; Hajredini and Ghose, 2021). In the OS, the conserved Walker-A Lys (the so-called “catalytic lysine”, K540) establishes a salt bridge with the conserved Walker-B Asp (D642) generating a conformation that cannot effectively coordinate Mg^{2+} (and consequently, ATP• Mg^{2+}). In the CS, the Walker-A Lys disengages from the Walker-B Asp and the latter then contacts the conserved Walker-A Thr (T541) generating a conformation required to optimally coordinate Mg^{2+} . To assess whether similar conformational states stabilized by interactions between key conserved residues also populate the landscape of the P-loop kinases, we performed enhanced sampling molecular dynamics simulations (REST2) (Wang et al., 2011) using *Mycobacterium tuberculosis* SK (SK_{Mtu} from hereon) in its unliganded, and its ATP• Mg^{2+} -bound states. Our results suggest that SK_{Mtu} samples distinct global conformations that, as in the case of the Wz_{CDAC}, are distinguished by local interactions involving the Walker-A and Walker-B motifs, most notably those involving the catalytic Lys (K15). This structural feature of the Walker-A Lys in SKs is also preserved despite sequence variations in the Walker-B motif (as in *Coxiella burnetii* SK, SK_{Cbu}) suggesting that this particular role of this residue in parsing functional conformations may be a general feature in P-loop containing enzymes.

Materials and Methods

Preparation of starting structures

The starting conformation of the SK_{Mtu}•ATP• Mg^{2+} complex was generated using the coordinates of SK_{Mtu} from the structure of the SK_{Mtu}•ADP• Mg^{2+} •shikimate complex (PDB: 1WE2) (Pereira et al., 2004) after removal of the bound ADP and shikimate. The Mg^{2+} ion and all solvent molecules were retained. The coordinates of ATP were derived from the structure of the SK_{Mtu}•AMPPCP•shikimate complex (PDB: 1ZYU) (Gan et al., 2006) by aligning the 1WE2 and

1ZYU structures using the C α atoms of their respective protein components and converting the bound AMPPCP to ATP through the replacement of the relevant carbon atom by an oxygen. Subsequently, the protein and shikimate coordinates of the 1ZYU structure were removed to generate the SK_{Mtu}•ATP•Mg²⁺ complex. For consistency, the corresponding apo simulations utilized the same starting structure after removal of ATP•Mg²⁺. For **unliganded** SK_{Cbu}, the starting structure was obtained directly from the coordinates found in PDB: 3TRF (Franklin et al., 2015) after removal of all ligands except solvent.

Molecular dynamics simulations

The starting structures, generated as discussed above, were equilibrated using previously described protocols (Hajredini et al., 2020). Briefly, the systems were parameterized within the CHARMM36m force-field (Huang et al., 2017), solvated with TIP3P water molecules, energy minimized and then equilibrated, first in an NVT ensemble, and then in an NPT ensemble to generate starting structures for the enhanced sampling REST2 (Wang et al., 2011) simulations. Subsequently, REST2 production runs were carried out for 200 ns using 14 replicas spanning a 300-400 K temperature range. **The final exchange probabilities for the SK_{Mtu}•ATP•Mg²⁺, SK_{Mtu}, and SK_{Cbu} simulations were $23.3 \pm 8.2\%$, $24.2 \pm 8.4\%$ and $29.0 \pm 0.6\%$, respectively. The first 40 ns in each of the simulations were used as “burn-in” periods and discarded, and all subsequent analyses were carried out utilizing the remaining 160 ns. As shown in Figure S2, 40 ns was more than sufficient to ensure equilibration of the key observables (θ and $|h|$, see below) probed in the simulations.**

Definition of the cylindrical coordinate frame

To describe the global conformational landscape of SK_{Mtu}, a cylindrical coordinate frame, comprising of an angle θ and a rise $|h|$ (Figure 1A), was defined as in the case of WzCDAC,

described previously (Figure S3A) (Hajredini et al., 2020). The centers of mass of the C α atoms of two segments on helix α 1 (that contains the Walker-A motif), comprising residues 15-19 (P1, brown) and 21-25 (P2, yellow), were used to define a vector that was aligned to the z-axis of the molecular frame. Each structure within the REST2-generated ensembles was rotated and translated such that point P3 (cyan), defined by the center of mass of the C α atoms of the first turn of helix α 2 (33-37), was placed on the x-axis along the positive direction. θ is then defined as the angle between the xy-projection of a vector extending from point P1 to P4 (blue; defined by the center of mass of the C α atoms of residues 155-159 that form the first turn of helix α 8). The rise $|h|$ is defined as the absolute value of the distance between points P1 and P3 along the z-axis. Thus, the angle θ and the rise $|h|$ provide measures of the outward rotation and the up/down movement, respectively, of the Walker-A motif relative to the Walker-B and DxD motifs. The crystal structures of SK_{Mtu} (listed in Table S1) were projected onto the cylindrical coordinate frame using the same protocol. For the structures of the SK enzymes from other organisms (listed in Table S1), a similar procedure was utilized to define the points P1-P4 (described in Table S2) and subsequently, θ and $|h|$.

Results and Discussion

Unliganded and liganded conformational ensembles of SK_{Mtu} populate distinct structural states

We performed two sets of enhanced sampling simulations utilizing the replica exchange with solute scaling (REST2) approach. REST2 represents a modification of the Hamiltonian scaling protocol of the replica exchange with solute tempering (REST) scheme (Liu et al., 2005) that greatly enhances its efficiency by decoupling the acceptance rate from the solvent. For the REST2 simulations, protocols similar to those described before (Hajredini et al., 2020) were

utilized with either unliganded (apo) or ATP•Mg²⁺-bound SK_{Mtu}. To analyze the global conformational states populated by these two ensembles we utilized the cylindrical coordinate frame defined in Figure 1A. Note that this frame is similar, but not identical to, that previously defined for Wz_{CDAC} (Hajredini et al., 2020) (compare Figure S3A) given the significant structural differences between the two proteins. The orientations of the α 8 helix (carrying P4) and α 1 (carrying P1 and P2) in SK_{Mtu} are similar to the α 2- α 3 pair in Wz_{CDAC} (the latter carries the Walker-A sequence) (Hajredini et al., 2020). *In lieu* of a missing Wz_{CDAC} α 4 equivalent in SK_{Mtu}, the α 2 helix (that hosts the second Asp, D34, of the DxD motif) was utilized to define P3. Thus, while this definition allows for a qualitative comparison of the global conformational states sampled by SK_{Mtu} and Wz_{CDAC} with respect to their key catalytic elements, a quantitative equivalence should not be expected.

Projection of the conformational ensembles of apo-SK_{Mtu} (Figure 1B, top panel) and the SK_{Mtu}•ATP•Mg²⁺ complex (Figure 1C, top panel) onto the 2-dimensional θ - $|h|$ space suggests differences in global conformations between the two cases. The presence of ATP•Mg²⁺ induces a more closed conformation, characterized by smaller θ values, compared to a more open conformation (larger θ values) seen for apo-SK_{Mtu}. We designate these global states as a Walker Closed State (WCS, seen in the SK_{Mtu}•ATP•Mg²⁺ complex), and a Walker Open State (WOS, seen in apo-SK_{Mtu}) to distinguish them from the closed and open classifications of the relative conformation of the SB and the LID domains traditionally used in the SK literature (Hartmann et al., 2006). Contrasts between the θ - $|h|$ projections of SK_{Mtu} and Wz_{CDAC} are evident upon comparing Figures 1B, C with Figures S3B, C. In the case of Wz_{CDAC}, the OS is characterized by high values of both θ and $|h|$, while the CS exhibits low values for both variables (Figures S3B, C), suggesting that these variables are positively correlated (Hajredini et al., 2020). In contrast, the

regions of maximal density for the WOS (Figure 1B, top panel) and the WCS (Figure 1C, top panel) suggest a weak anti-correlation between the θ (WOS: $196 \pm 4^\circ$, WCS: $188 \pm 3^\circ$) and $|h|$ dimensions (WOS: $4.6 \pm 0.7 \text{ \AA}$, WCS: $4.9 \pm 0.4 \text{ \AA}$), with lower θ and higher $|h|$ for the WCS, and the opposite for the WOS. Further, the apparent separation between the WOS and the WCS seems less prominent than that between the OS and CS of Wz_{CCDΔC} (compare Figures 1B, C with Figures S3B, C; also see Figure S4A). As an independent validation of our simulations, we note that the regions in θ - $|h|$ space spanned by the WOS and the WCS in our REST2 ensembles are well represented (Figure S4A) in the several available crystal structures of SKs (and their complexes) from a variety of organisms (listed in Table S1).

Representative structures of apo- (Figure 1B, bottom panel) and ATP•Mg²⁺-bound (Figure 1C, bottom panel) SK_{Mtu} extracted from the regions of maximal density in θ - $|h|$ space of the corresponding ensembles show altered conformations of key catalytic site elements. In the presence of ATP•Mg²⁺ i.e., in the WCS, the Walker-A Lys (K15) assumes a downwards orientation and contacts the β - and γ -phosphates of ATP; the first Asp (D32) of the DxD motif forms a hydrogen bond with the Walker-A S16, thus functionally mimicking the absent Walker-B Asp (Leipe et al., 2003). The D32-S16 interaction is necessary for the optimal coordination of Mg²⁺ (Figure 1C, bottom panel). Indeed, as shown in Figure S4B, the D32,C γ -S16,O γ distance is well defined and shows a narrow distribution ($3.7 \pm 0.2 \text{ \AA}$) within the ensemble. In the absence of ligands, i.e., in the WOS, K15 rotates outwards to establish a hydrogen-bond with S77 of the Walker-B motif (Figure 1B, bottom panel) **in a significant fraction of the structures**. In this configuration, the D32-S16 distance shows a broad distribution ($6.0 \pm 1.2 \text{ \AA}$), **suggestive of increased disorder**, with a maximum reflective of significantly increased separation (Figure S4B) suggesting a conformation that is not conducive for Mg²⁺ coordination. **Conformations displaying**

the K15-S77 hydrogen bond have been observed in several crystal structures of SK_{Mtu}, that are, not surprisingly, free of Mg²⁺. The structures (representative examples shown in Figure S4C) (Hartmann et al., 2006) that display the K15-S77 interaction have the highest values of θ , and correspondingly, the lowest values of $|h|$, and exist at the outer edges of the WOS. Taken together, these results suggest that the global conformations of SK_{Mtu} are coupled to the conformational state of the Walker-A Lys as was observed for Wzc_{CDAC} (Hajredini et al., 2020; Hajredini and Ghose, 2021).

As mentioned above, in a significant fraction of the structures that comprise the WOS of SK_{Mtu}, the Walker-A K15 forms a hydrogen bond with the Walker-B S77 (Figure 1B), generating a conformation that mirrors that seen in the Wzc_{CDAC} OS. In the latter case, the equivalent Walker-A Lys (K540) forms a salt bridge with the Walker-B D642. The β -strand housing the Walker-B motif in SK_{Mtu} is displaced relative to that in Wzc_{CDAC}, this enables S77 from the degraded Walker-B in the former to assume a spatial position that is approximately equivalent to the Walker-B D642 in the latter to facilitate an interaction with K15 (Figure 2A). It would therefore appear from our current results that small local structural rearrangements allow the polar Ser to substitute for the now missing Walker-B Asp in preserving the interaction with the Walker-A K15 to maintain the open conformation (WOS) in SK_{Mtu}. However, despite these similarities, there appear to be differences in stability between the open conformations involving the Walker-A Lys between SKs and BY-kinases. In Wzc_{CDAC}, this interaction mode, as assessed through the K540,N ξ -D642,C γ distance distribution, is well defined over the entire OS ensemble, showing a single peak with a maximum at ~ 3.3 Å (green trace in Figure 2B). The corresponding distance, K15,N ξ -S77,O γ in SK_{Mtu}, shows a broad distribution with the dominant peak centered at ~ 2.8 Å but containing only approximately 50% of the population in the WOS (pink trace in Figure 2B; this disorder is also

evident from the bottom panel of Figure 1B). This suggests that while the orientation of the Walker-A Lys is well-defined in WZC_{CDAC}, it is somewhat more dynamic in SK_{Mtu}. The reasons for this difference are analyzed in detail below.

As mentioned earlier, the conformational dynamics of the LID and SB domains have been suggested to be critical for function in the SKs in serving to stabilize the reaction-compatible state and in facilitating product release (Blanco et al., 2013). The dynamics of these entities have been suggested as suitable targets for drug discovery (Prado et al., 2016). For the sake of completeness, we also analyzed the local conformational variability in the REST2-generated ensembles of SK_{Mtu}. There is, not unexpectedly, a slight decrease in overall flexibility upon binding ATP•Mg²⁺ (Figure S5A), and transition from the WOS to the WCS. The LID domain that is the most dynamic region of the protein in the unliganded state remains so in the presence of ATP•Mg²⁺, though there is some decrease in its overall flexibility. The SB domain is also dynamic in both states of SK_{Mtu}, albeit less so than the LID. To probe the presence of opening/closing motions of the SB domain with respect to the protein core, we defined a distance (d₄₇₋₁₀₉, Figure S5B, top left panel) between the C α atoms of D47 (one of the more dynamic parts of the SB domain and therefore likely to sense these motions, if any) and R109 (near start of the LID domain; one of the more rigid parts of the protein in its ATP•Mg²⁺-bound state). In the absence of ATP•Mg²⁺ a somewhat bimodal distribution of distances is seen (Figure S5B top right panel). Analyses of structures corresponding to the two maxima indicate an open conformation with greater separation between the LID and SB domains (Figure S5B bottom right panel), and a more closed conformation where this distance is reduced (Figure S5B bottom middle panel). Only the latter conformation is seen in the ATP•Mg²⁺-bound state (Figure S4B bottom left panel).

Conserved features define the global conformational states of P-loop enzymes

While S77 of SK_{Mtu} appears to substitute for the Asp/Glu residue of a canonical Walker-B motif in maintaining the WOS through its interaction with the Walker-A Lys, this position is poorly conserved in SKs. In contrast to the $\phi\phi$ SLGGG sequence in SK_{Mtu}, the consensus sequence for Walker-B motifs of SKs is $\phi\phi\phi$ TGGG (Leipe et al., 2003); the corresponding consensus in BY-kinases is $\phi\phi\phi\phi$ DT (Leipe et al., 2002). Notably, in Wzc_{CDAC}, the conserved Walker-B Thr (T643) participates in multiple interactions that also serve to stabilize the OS (Figure 3A, top panel). T643 forms a hydrogen bond through its backbone carbonyl with the K540 sidechain (supplementing the interaction of the latter with D642), and an additional hydrogen bond through its sidechain hydroxyl with the backbone carbonyl of M531 on the adjacent β -sheet. Projection of the unliganded Wzc_{CDAC} ensemble (Hajredini et al., 2020) onto the 2-dimensional space spanned by the K540-T643 (D1: K540,N ξ -T643,O) and T643-M531 (D2: T643,O γ 1-M531,O) distances, indicates a single major state (white arrow, Figure 3A, bottom panel). This robust network of interactions (K540-D642, K540-T643 and T643-M531) is the likely reason for the well-defined orientation of K540 in the OS (as is evident from the green trace in Figure 2B). This set of interactions is drastically altered in the CS (T643 now interacts T532, D642 interacts with T541, the latter interaction is necessary to optimally engage Mg²⁺) and this is the likely cause of the well-defined separation between in the OS and the CS upon comparing the apo- and ATP•Mg²⁺ complexes of Wzc_{CDAC} (Figures S3B, C).

To test whether a similar interaction mode involving the Walker-B Thr also stabilizes the WOS in SKs that carry the more conventional $\phi\phi\phi$ TGGG motif, we performed an additional set of REST2 simulations on the unliganded state of *Coxiella burnetii* SK (SK_{Cbu}). Inspection of the structural ensemble of apo-SK_{Cbu} reveals the presence of hydrogen bonds between the backbone carbonyl and sidechain hydroxyl of the Walker-B Thr (T81) with the sidechain of the Walker-A

K18 and the backbone carbonyl of L10 (that lies on the adjacent β -sheet), respectively (Figure 3B, top panel), **mirroring the interaction mode seen** in Wz_{CDAC} . However, unlike in Wz_{CDAC} , projection of the resultant ensemble onto the 2-dimensional space spanned by the K18-T81 (D1: K18,N ξ -T81,O) and T81-**L10** (D2: T81,O γ 1-L10,O) distances reveals a more diffuse set of states (Figure 3B, bottom panel). Nevertheless, a disruption in the T81-K18 interaction leads to a corresponding disruption of the T81-L10 interaction (red arrow in Figure 3B, bottom panel) suggesting that these two sets of interactions are coupled as **for** Wz_{CDAC} . It is of note that the canonical Walker-B Thr in SK_{Cbu} forms two sets of the hydrogen bonds involving both its backbone and its sidechain to stabilize the WOS. This contrasts a single hydrogen bond involving the sidechain of the unique Walker-B Ser (S77) in SK_{Mtu} . One can therefore expect the WOS in SK_{Cbu} to be somewhat more closed than in SK_{Mtu} . Inspection of Figure 3C shows that this is indeed the case with reduced θ values (SK_{Cbu} : $190 \pm 5^\circ$, SK_{Mtu} : $196 \pm 4^\circ$;) and a corresponding increase in the $|h|$ values (SK_{Cbu} : $4.8 \pm 0.6 \text{ \AA}$, SK_{Mtu} : $4.6 \pm 0.7 \text{ \AA}$).

Based on the discussion above, an SK Walker-B motif may be generalized by the following sequence: $\phi\phi\alpha\beta GGG$; the α - and β -positions represent structural equivalents of the **BY-kinase** Walker-B Asp and Thr, respectively (Figure 4A). In unliganded Wz_{CDAC} , both the α - and β -positions are optimal in that they contain polar residues. Thus, the Walker-A Lys (K540) **can form** hydrogen bonds with both the Walker-B Asp (D642 at the α -position) and with the Walker-B Thr (T643); the latter being a polar residue at the β -position is then able to utilize its sidechain to hydrogen bond with M531. This generates a mesh of interactions (shown schematically in Figure 4B) leading to a well-defined orientation of the Walker-A Lys (K540) in a stabilized OS. In contrast, the fact that only one of these two positions contains a polar residue for the two SKs analyzed here, with T81 (**β**) in SK_{Cbu} (Figure 4C) or S77 (**α**) in SK_{Mtu} (Figure 4D). Thus, only a

subset of these interactions is possible in the SKs resulting in less well-defined orientation for the Walker-A Lys and a more diffuse WOS.

Conclusions

Through enhanced sampling MD simulations on unliganded SK_{Mtu} and its ATP•Mg²⁺ complex, we have shown that as in the case of BY-kinases, the conformational landscapes of SKs **contain** open (WOS) and closed (WCS) states. The former is found in the unliganded state, while the latter is induced by the presence of nucleotide and Mg²⁺. In the WOS, the Walker-A K15 of SK_{Mtu} **populates** an extended conformation **allowing the formation of** a hydrogen bond with S77 of the degraded Walker-B motif that mimics the conserved Asp of the intact Walker-B motif in BY-kinases. In SK_{Cbu}, that contains a Walker-B sequence that is more characteristic of P-loop kinases of the DxD group (Leipe et al., 2003), the conserved T81 forms a hydrogen-bond with the Walker-A K18 through its backbone while simultaneously using its polar sidechain to interact with the neighboring L10 backbone to stabilize the WOS. Thus, it appears that specific contacts between the Walker-A and Walker-B motifs, more specifically those involving the so-called catalytic Lys on the former, are necessary in P-loop-containing kinases to maintain the open conformation that cannot efficiently co-ordinate Mg²⁺ (and therefore, ATP•Mg²⁺) and is not conducive for chemistry. The importance of the hydrogen-bond involving the Walker-A Lys is evident from the fact that the absence of a Walker-B Asp as in the SKs, is compensated by the presence of a polar residue at a similar position. While the functional role the OS and CS and the interactions encoded within each state for the catalytic domain of Wzc has been extensively tested experimentally (Hajredini et al., 2020; Hajredini and Ghose, 2021), we expect that the computational analyses presented here will provide guidelines for similar experimental validation of key functional features in SKs in particular, and in P-loop kinases, in general.

Data Availability Statement

All data needed to evaluate the conclusions in the paper are present in the paper and/or in the Supplementary Materials.

Author contributions

RG conceived the project; FH performed and analyzed all simulations. FH prepared a first draft of the paper and figures that were refined by RG with input from FH.

Funding

Support from NSF grants MCB1937937 and PHY1811770 is acknowledged. FH is partially supported by United States Department of Education GAANN award P200A150068.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Aravind, L., Mazumder, R., Vasudevan, S., and Koonin, E.V. (2002). Trends in protein evolution inferred from sequence and structure analysis. *Curr. Opin. Struct. Biol.* 12, 392-399.
- Bechet, E., Gruszczyk, J., Terreux, R., Gueguen-Chaignon, V., Vigouroux, A., Obadia, B., Cozzone, A.J., Nessler, S., and Grangeasse, C. (2010). Identification of structural and molecular determinants of the tyrosine-kinase Wzc and implications in capsular polysaccharide export. *Mol. Microbiol.* 77, 1315-1325.
- Blanco, B., Prado, V., Lence, E., Otero, J.M., Garcia-Doval, C., Van Raaij, M.J., Llamas-Saiz, A.L., Lamb, H., Hawkins, A.R., and Gonzalez-Bello, C. (2013). *Mycobacterium tuberculosis* shikimate kinase inhibitors: design and simulation studies of the catalytic turnover. *J. Am. Chem. Soc.* 135, 12366-12376.

- Edwards, H., Abeln, S., and Deane, C.M. (2013). Exploring fold space preferences of new-born and ancient protein superfamilies. *PLoS Comput. Biol.* 9, e1003325.
- Franklin, M.C., Cheung, J., Rudolph, M.J., Burshteyn, F., Cassidy, M., Gary, E., Hillerich, B., Yao, Z.K., Carlier, P.R., Totrov, M., and Love, J.D. (2015). Structural genomics for drug design against the pathogen *Coxiella burnetii*. *Proteins* 83, 2124-2136.
- Gan, J., Gu, Y., Li, Y., Yan, H., and Ji, X. (2006). Crystal structure of *Mycobacterium tuberculosis* shikimate kinase in complex with shikimic acid and an ATP analogue. *Biochemistry* 45, 8539-8545.
- Grangeasse, C., Nessler, S., and Mijakovic, I. (2012). Bacterial tyrosine kinases: evolution, biological function and structural insights. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2640-2655.
- Hajredini, F., and Ghose, R. (2021). An ATPase with a twist: a unique mechanism underlies the activity of the bacterial tyrosine kinase, Wzc. *Sci. Adv.* 7, eabj5836.
- Hajredini, F., Piserchio, A., and Ghose, R. (2020). Long-range dynamic correlations regulate the catalytic activity of the bacterial tyrosine kinase Wzc. *Sci. Adv.* 6, eabd3718.
- Hartmann, M.D., Bourenkov, G.P., Oberschall, A., Strizhov, N., and Bartunik, H.D. (2006). Mechanism of phosphoryl transfer catalyzed by shikimate kinase from *Mycobacterium tuberculosis*. *J. Mol. Biol.* 364, 411-423.
- Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., De Groot, B.L., Grubmuller, H., and Mackerell, A.D., Jr. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Meth.* 14, 71-73.
- Lee, D.C., and Jia, Z. (2009). Emerging structural insights into bacterial tyrosine kinases. *Trends Biochem. Sci.* 34, 351-357.

- Leipe, D.D., Koonin, E.V., and Aravind, L. (2003). Evolution and classification of P-loop kinases and related proteins. *J. Mol. Biol.* 333, 781-815.
- Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol.* 317, 41-72.
- Liu, P., Kim, B., Friesner, R.A., and Berne, B.J. (2005). Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. USA* 102, 13749-13754.
- Longo, L.M., Jablonska, J., Vyas, P., Kanade, M., Kolodny, R., Ben-Tal, N., and Tawfik, D.S. (2020). On the emergence of P-Loop NTPase and Rossmann enzymes from a Beta-Alpha-Beta ancestral fragment. *Elife* 9, e64415.
- Pereira, J.H., De Oliveira, J.S., Canduri, F., Dias, M.V., Palma, M.S., Basso, L.A., Santos, D.S., and De Azevedo, W.F., Jr. (2004). Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2310-2319.
- Prado, V., Lence, E., Maneiro, M., Vazquez-Ucha, J.C., Beceiro, A., Thompson, P., Hawkins, A.R., and Gonzalez-Bello, C. (2016). Targeting the motion of shikimate kinase: development of competitive inhibitors that stabilize an inactive open conformation of the enzyme. *J. Med. Chem.* 59, 5471-5487.
- Scheeff, E.D., and Bourne, P.E. (2005). Structural evolution of the protein kinase-like superfamily. *PLoS Comput. Biol.* 1, e49.
- Shi, L., Ji, B., Kolar-Znika, L., Boskovic, A., Jadeau, F., Combet, C., Grangeasse, C., Franjevic, D., Talla, E., and Mijakovic, I. (2014). Evolution of bacterial protein-tyrosine kinases and their relaxed specificity toward substrates. *Genome Biol. Evol.* 6, 800-817.

- Taylor, S.S., and Radzio-Andzelm, E. (1994). Three protein kinase structures define a common motif. *Structure* 2, 345-355.
- Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequence in the alpha and beta subunits of ATP synthase, myosin, kinases and other ATP requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945-951.
- Wang, L., Friesner, R.A., and Berne, B.J. (2011). Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). *J. Phys. Chem. B* 115, 9431-9438.
- Yang, Y., Liu, J., Clarke, B.R., Seidel, L., Bolla, J.R., Ward, P.N., Zhang, P., Robinson, C.V., Whitfield, C., and Naismith, J.H. (2021). The molecular basis of regulation of bacterial capsule assembly by Wzc. *Nat. Commun.* 12, 4349.

Figure Legends

Figure 1. (A) Points P1 (brown), P2 (yellow), P3 (cyan) and P4 (dark blue) used to define the cylindrical coordinate frame for analyses of the REST2-generated ensembles of SK_{Mtu} are indicated on its structure. The definitions of the angle θ and the rise $|h|$ are shown on the right panel. The REST2-generated structural ensembles of unliganded SK_{Mtu} (B) and the SK_{Mtu}•ATP•Mg²⁺ complex (C) are shown projected onto θ - $|h|$ space and plotted using kernel density estimation. The bottom panels show key interactions that stabilize each state using representative structures drawn from the corresponding regions of highest probability density (194-200° and 4-5 Å in the $|h|$ and θ dimensions, respectively for apo-SK_{Mtu} and 185-190° and 4.5-5.5 Å for the corresponding dimensions of the SK_{Mtu}•ATP•Mg²⁺ complex). A key feature that distinguishes the Walker Open State (WOS) from the Walker Closed State (WCS) is the orientation of the Walker-A Lys (K15 in SK_{Mtu}). For the WOS seen for apo-SK_{Mtu} K15 hydrogen

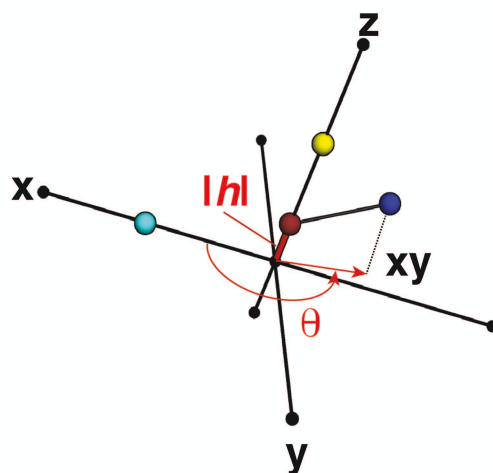
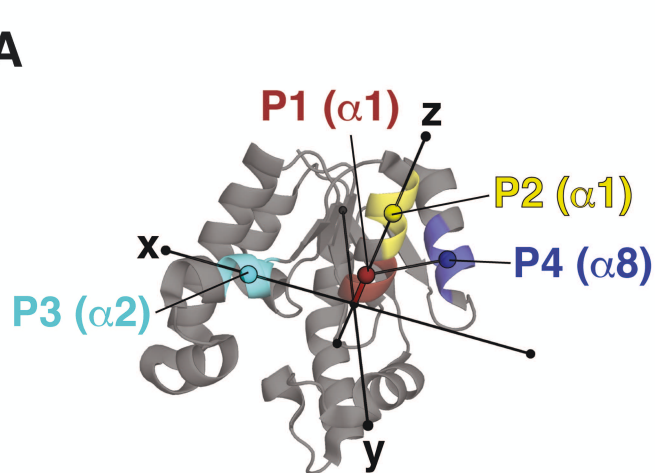
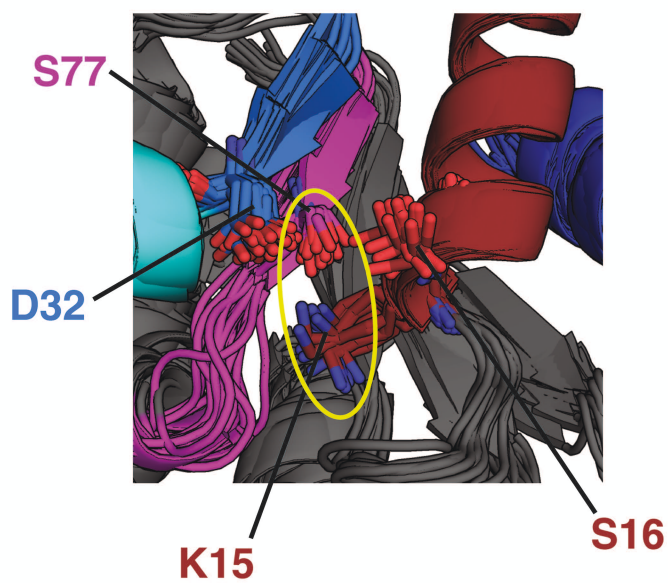
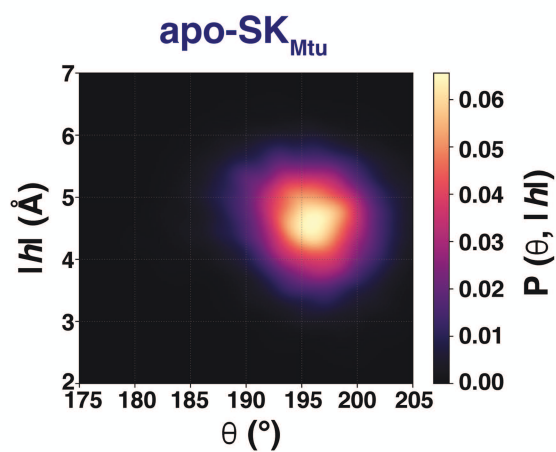
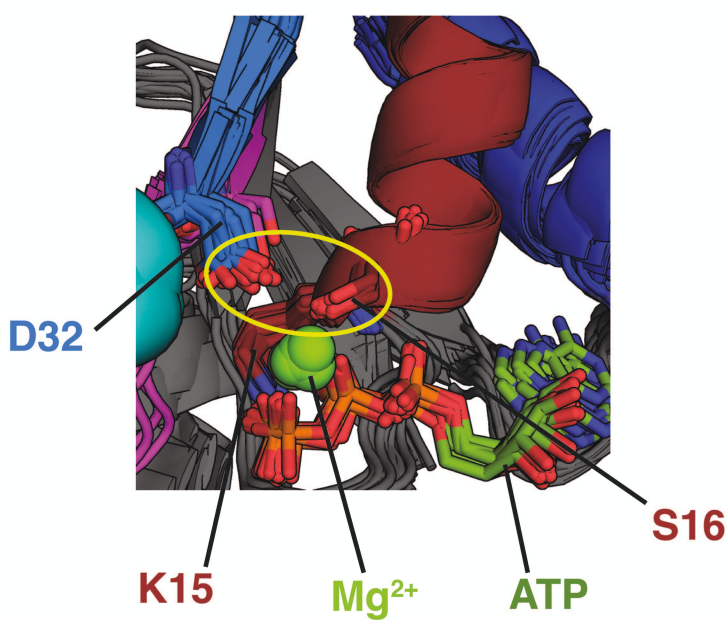
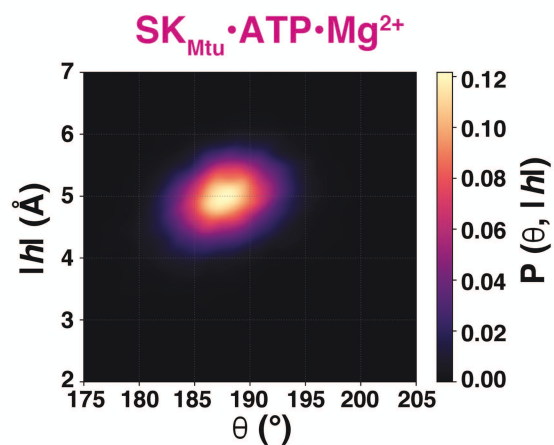
bonds with S77 of the Walker-B motif in the bulk of the structures (K15 and S77 are proximal in all of them). This interaction is broken in the $SK_{Mtu} \cdot ATP \cdot Mg^{2+}$ complex.

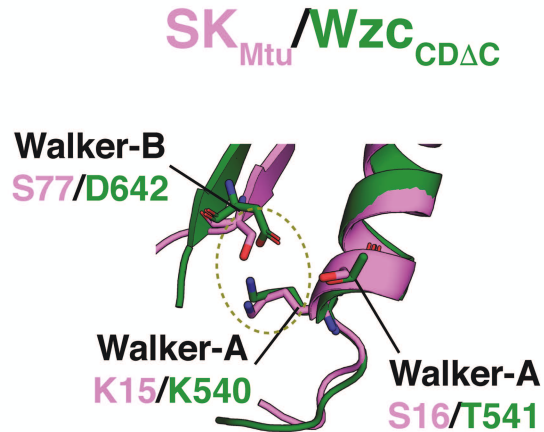
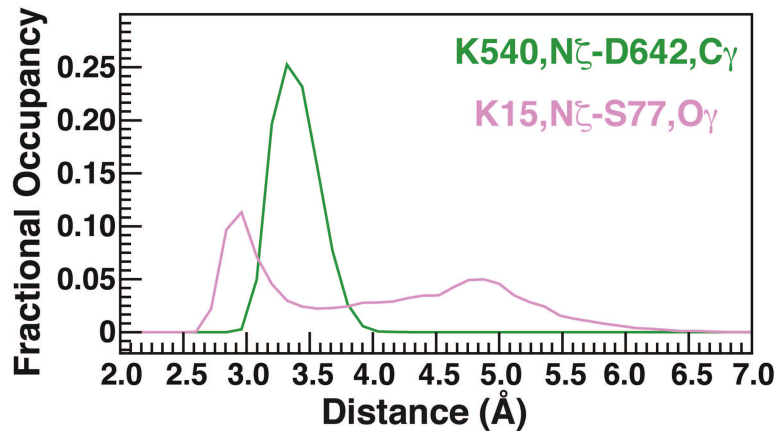
Figure 2. (A) Interactions that are characteristic of the open states of BY-kinases (OS) and SKs (WOS). S77 from the degraded Walker-B motif from SK_{Mtu} (pink) occupies approximately the same spatial position as the Walker-B D642 from Wz_{CDAC} (green) allowing the formation of a S77-K15 hydrogen bond reminiscent of the D642-K540 interaction in Wz_{CDAC} . **(B)** Distribution of the K540, N ξ -D642, C γ (green) and K15, N ξ -S77, O γ (pink) distances from the REST2-generated ensembles of unliganded Wz_{CDAC} and SK_{Mtu} , respectively.

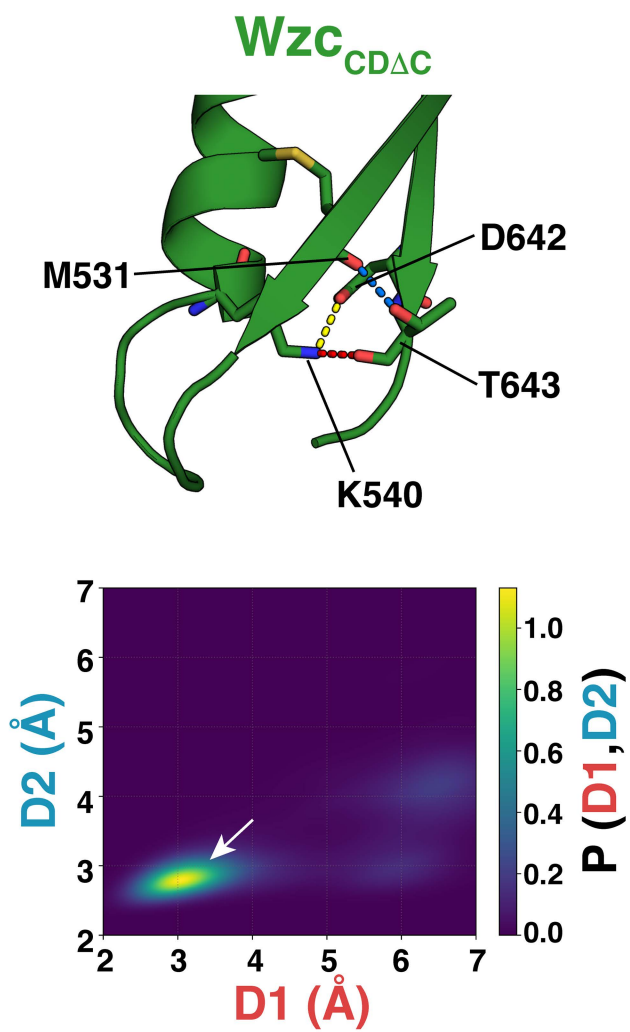
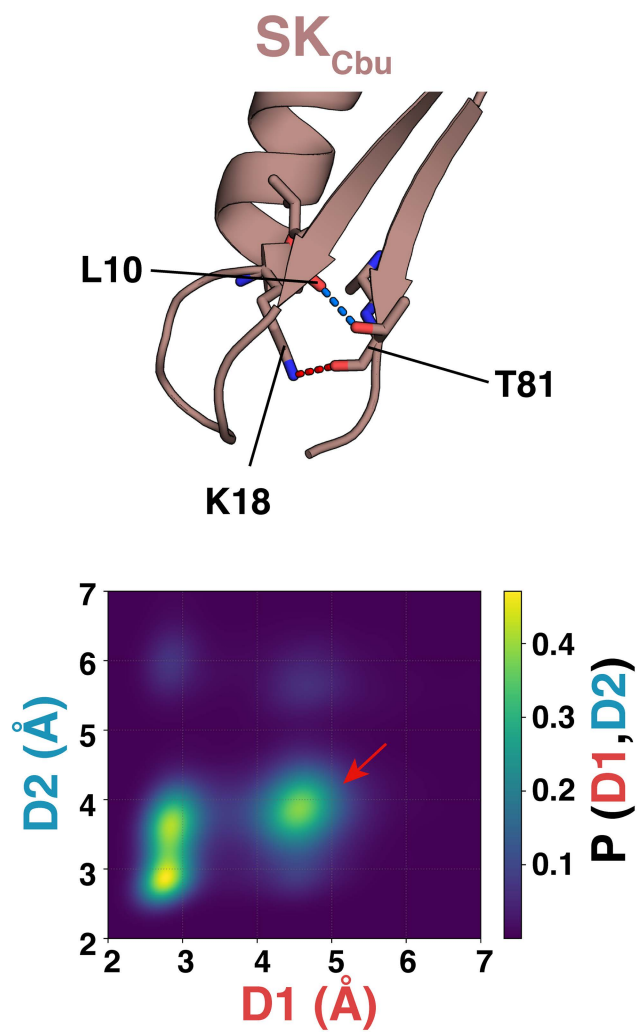
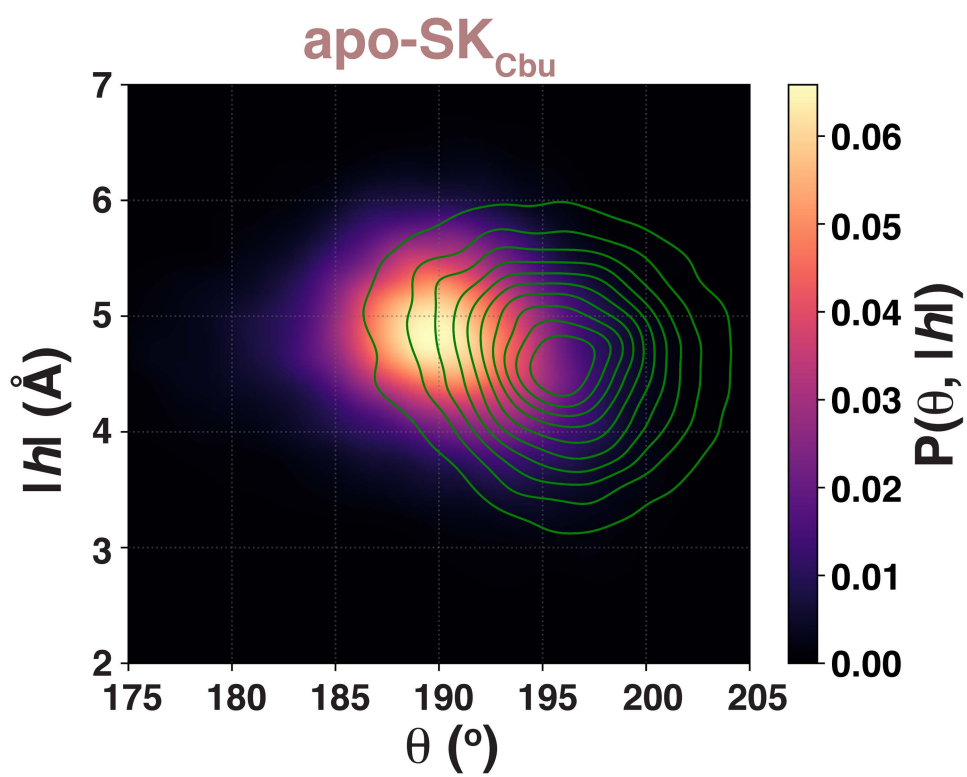
Figure 3. (A) Interactions of the Walker-A K540 with the Walker-B D642 (yellow dashed line), of K540 with T643 (red dashed line), and of T643 with M531 (blue dashed line) for Wz_{CDAC} in the OS are shown. **(B)** Interactions of the Walker-A K18 with Walker-B T81 (red dashed line) and of T81 with L10 (blue dashed line) for SK_{Cbu} in the WOS are shown. The lower panels, in each case, show the structural ensembles of unliganded Wz_{CDAC} and SK_{Cbu} projected onto the 2-dimensional space spanned by the {K540-T643 (D1: K540, N ξ -T643, O), T643-M531 (D2: T643, O γ 1-M531, O)} and {K18-T81 (D1: K18, N ξ -T81, O), T81-L10 (D2: T81, O γ 1-L10, O)} distances, respectively. **(C)** The structural ensemble of unliganded SK_{Cbu} is projected onto θ - $|h|$ space and plotted using kernel density estimation. Also shown as green contours in the corresponding projection of the unliganded SK_{Mtu} ensemble.

Figure 4. (A) Comparison of the Walker-B sequences in Wzc , SK_{Mtu} and SK_{Cbu} . Wzc contains polar residues at both the α - and β -positions while only one of these positions is polar in SK_{Cbu} (β) and SK_{Mtu} (α). **(B)** The presence of polar residues at both the α - (D642) and β -positions (T643) in the Walker-B motif of Wz_{CDAC} allows the intricate network of interactions to appropriately align the Walker-A Lys (K540) leading to its well-defined orientation in the OS.

These interactions are depicted schematically; hydrogen-bonds are indicated using purple and blue lines. Since only one of the α - or β -positions contain a polar residue in **(C)** SK_{Cbu} (β :T81) and **(D)** SK_{Mtu} (α :S77), only one of these sets of interactions (blue or purple) is preserved allowing a greater degree of orientational freedom in the WOS for the corresponding Walker-A Lys (K18 in SK_{Cbu}, K15 in SK_{Mtu}).

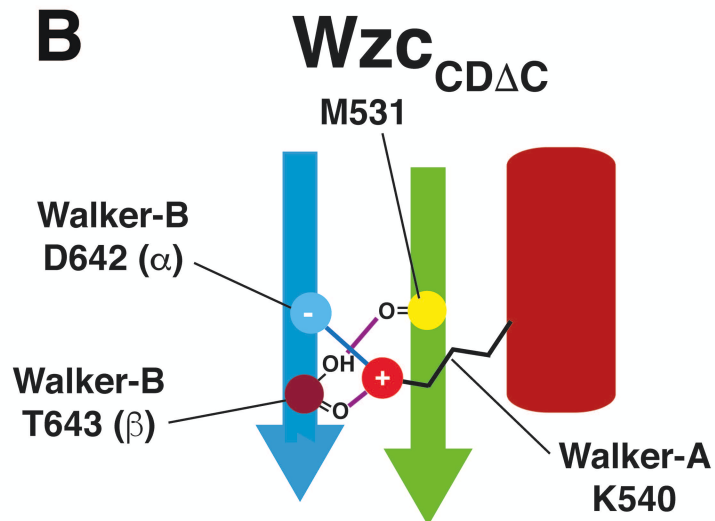
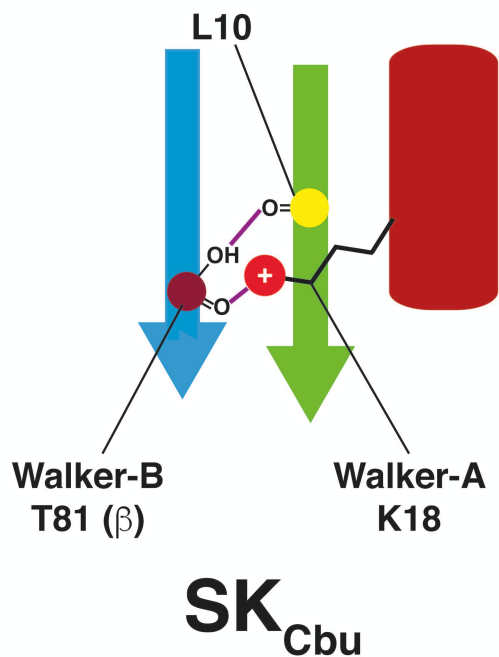
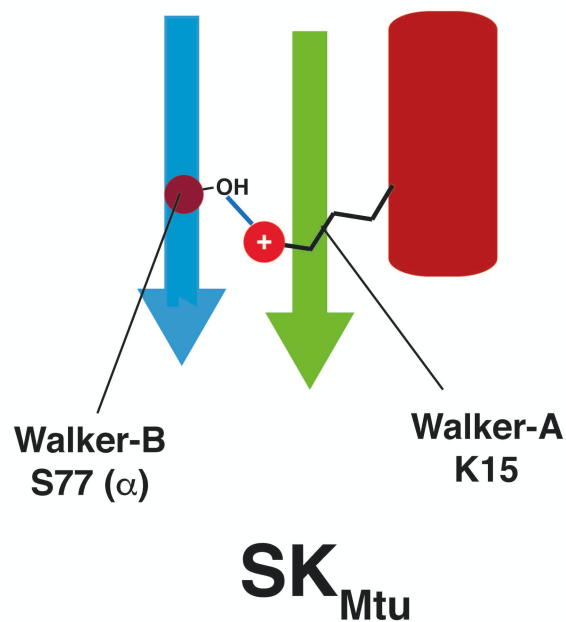
A**B****C**

A**B**

A**B****C**

A**Walker-B**

Wzc	637	LVLII	<u>DT</u>	PP-	645
SK_{Mtu}	75	--VL	<u>SL</u>	G	81
SK_{Cbu}	78	--IL	<u>AT</u>	G	84

B**C****D**

Supplementary Material

for

A conserved structural role for the Walker-A lysine in P-loop containing kinases

Fatum Hajredini and Ranajeet Ghose

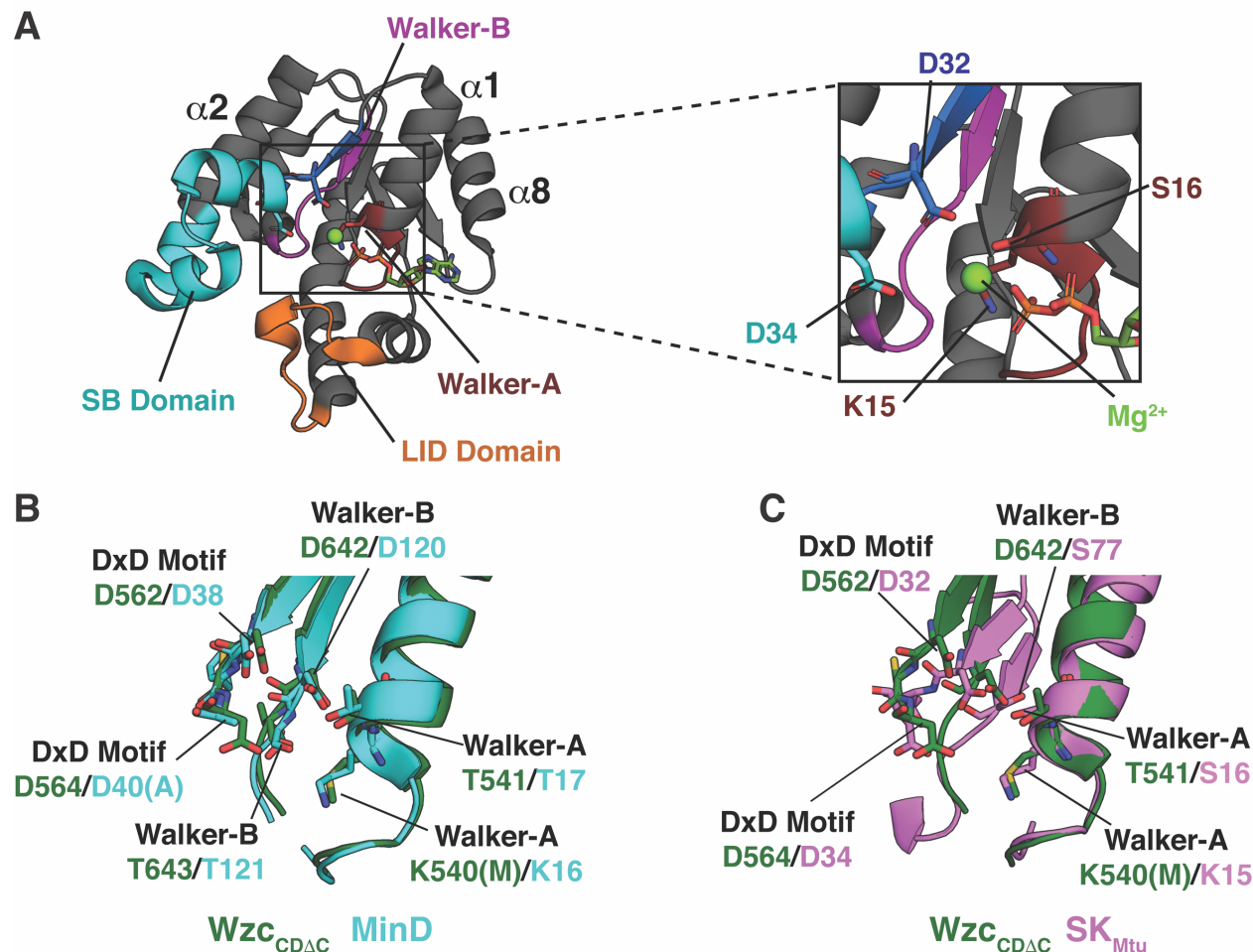


Figure S1. (A) Structure of *Mycobacterium tuberculosis* shikimate kinase (SK_{Mtu}). The shikimate binding (SB) and LID domains are colored cyan and orange, respectively. The Walker-A, Walker-B and the DxD motifs are colored brown, purple and dark blue, respectively. An expanded view of the catalytic site is shown on the right panel; key residues of the Walker-A (K15, S16) and DxD (D32, D34) motifs are indicated. Comparison of the catalytic site geometry of Wzc_{CDΔC} (K540M mutant, green) with (B) MinD (D40A mutant, cyan) and (C) SK_{Mtu} (purple); corresponding residues are indicated in each case. Note that like Wzc_{CDΔC}, MinD also contains both a Walker-B Asp and a DxD motif.

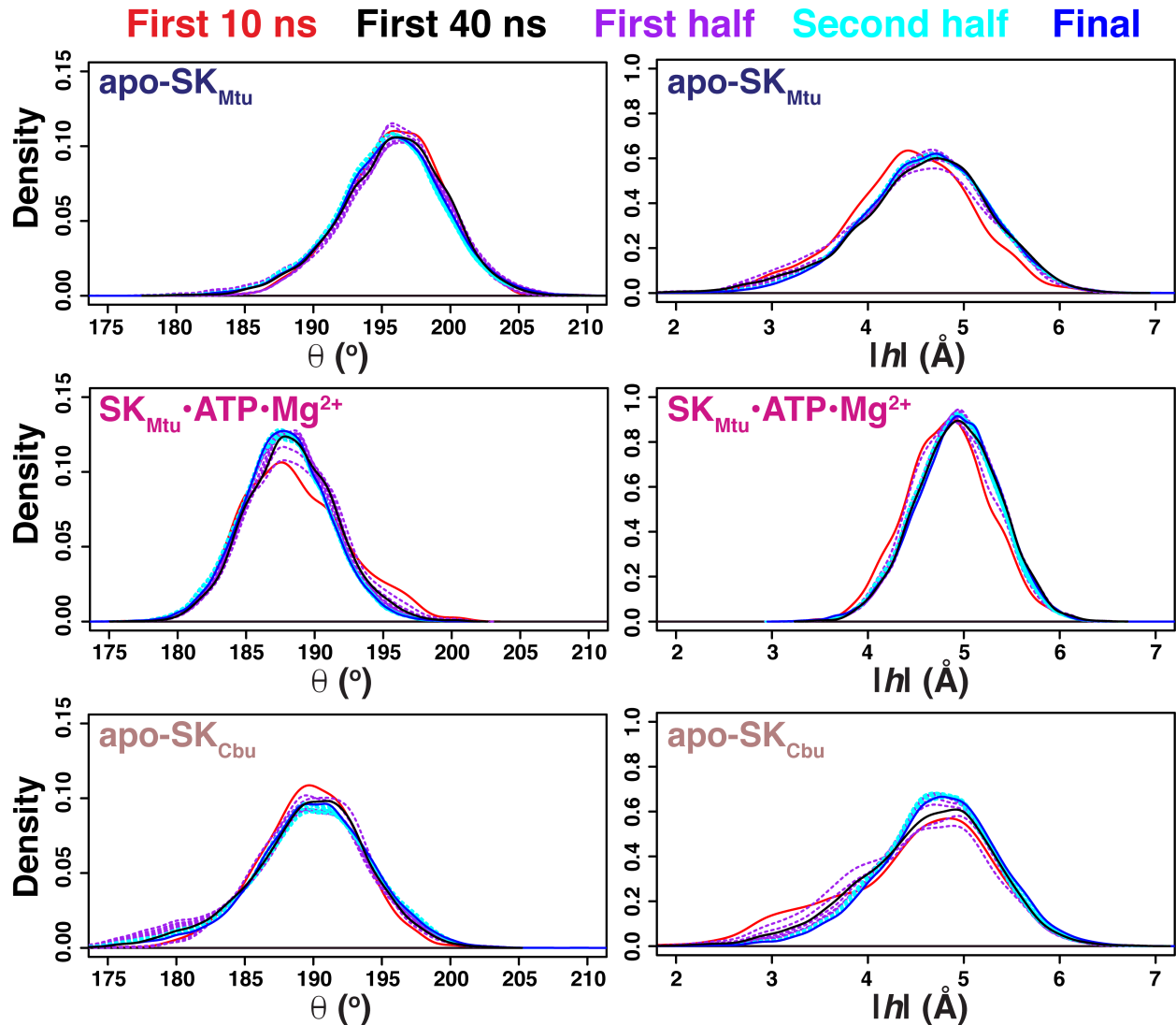


Figure S2. Convergence of the θ and $|h|$ distributions in the REST2 simulations on apo-SK_{Mtu} (top panels), SK_{Mtu}·ATP·Mg²⁺ (middle panels) and apo-SK_{Cbu} (bottom panels). Cumulative densities of states along the θ (left panels) and $|h|$ (right panels) dimensions calculated in 10 ns increments using kernel density estimation are shown as colored lines. Data for various durations are shown together with the final distributions (blue solid line) obtained over the last 160 ns. As is evident, the distributions obtained after 40 ns (black solid line) are essentially the same as the final distributions in contrast to the large deviations of the distributions calculated from the first 10 ns (red solid line) of the 200 ns trajectories. Also shown for reference are distributions calculated over 10 ns intervals from the first half (purple dashed lines) and second half (cyan dashed lines) of the simulations.

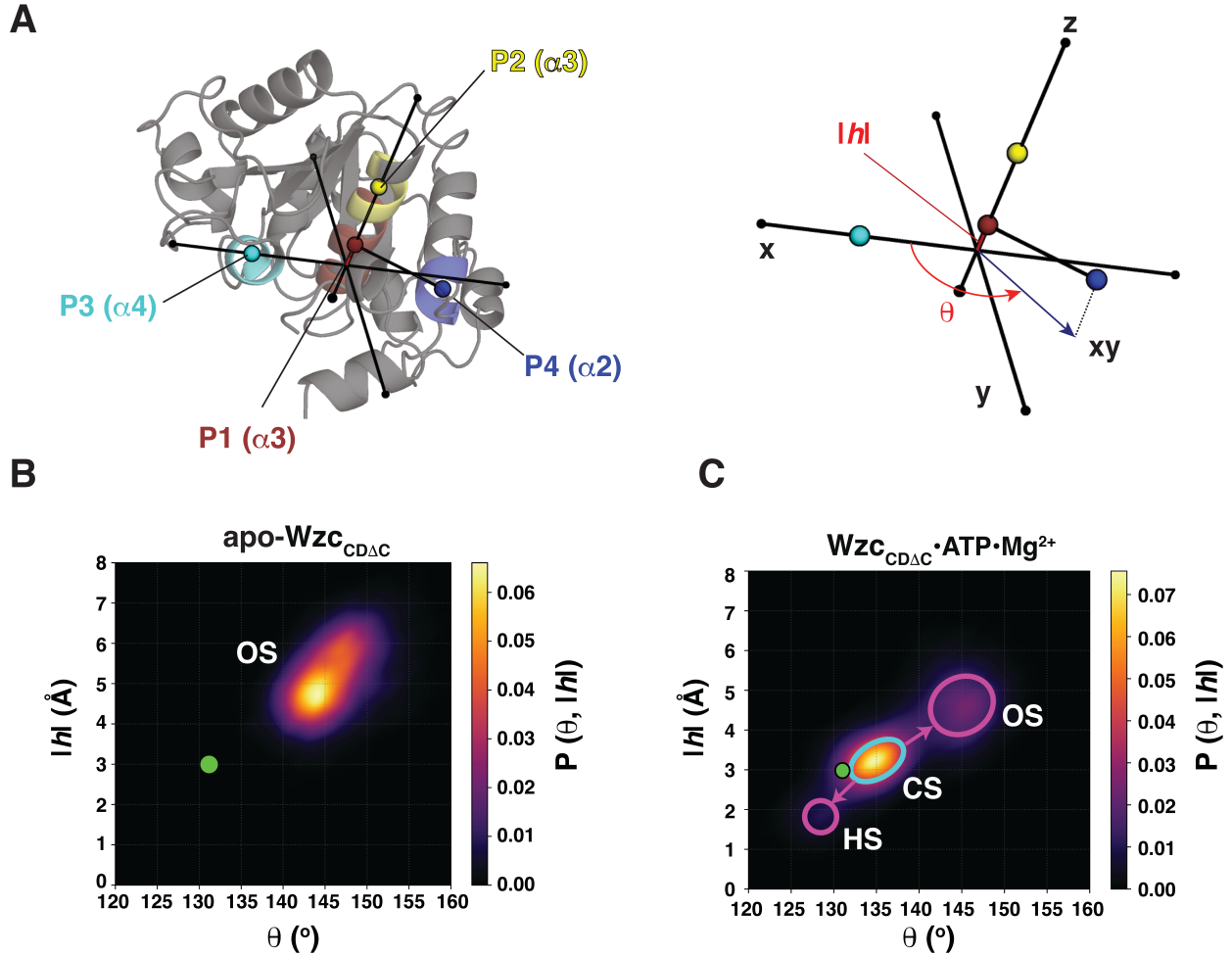


Figure S3. (A) Points P1 (brown), P2 (yellow), P3 (cyan) and P4 (blue) that define the cylindrical coordinate frame used to analyze the REST2-generated structural ensembles of Wzc_{CDAC} are indicated on its structure. The definitions of the angle θ and the rise $|h|$ are shown on the right panel. The REST2-generated ensembles for apo- Wzc_{CDAC} (B) and the $Wzc_{CDAC} \cdot ATP \cdot Mg^{2+}$ complex (C) are shown projected onto θ - $|h|$ space. Only an open state (OS) is populated in unliganded Wzc_{CDAC} . The presence of $ATP \cdot Mg^{2+}$ leads to the partial occupancy of a dominant closed state (CS) and a sparsely populated hyper-closed state (HS). The green filled circle represents the conformation seen in the crystallographic monomers (PDB: 3LA6).

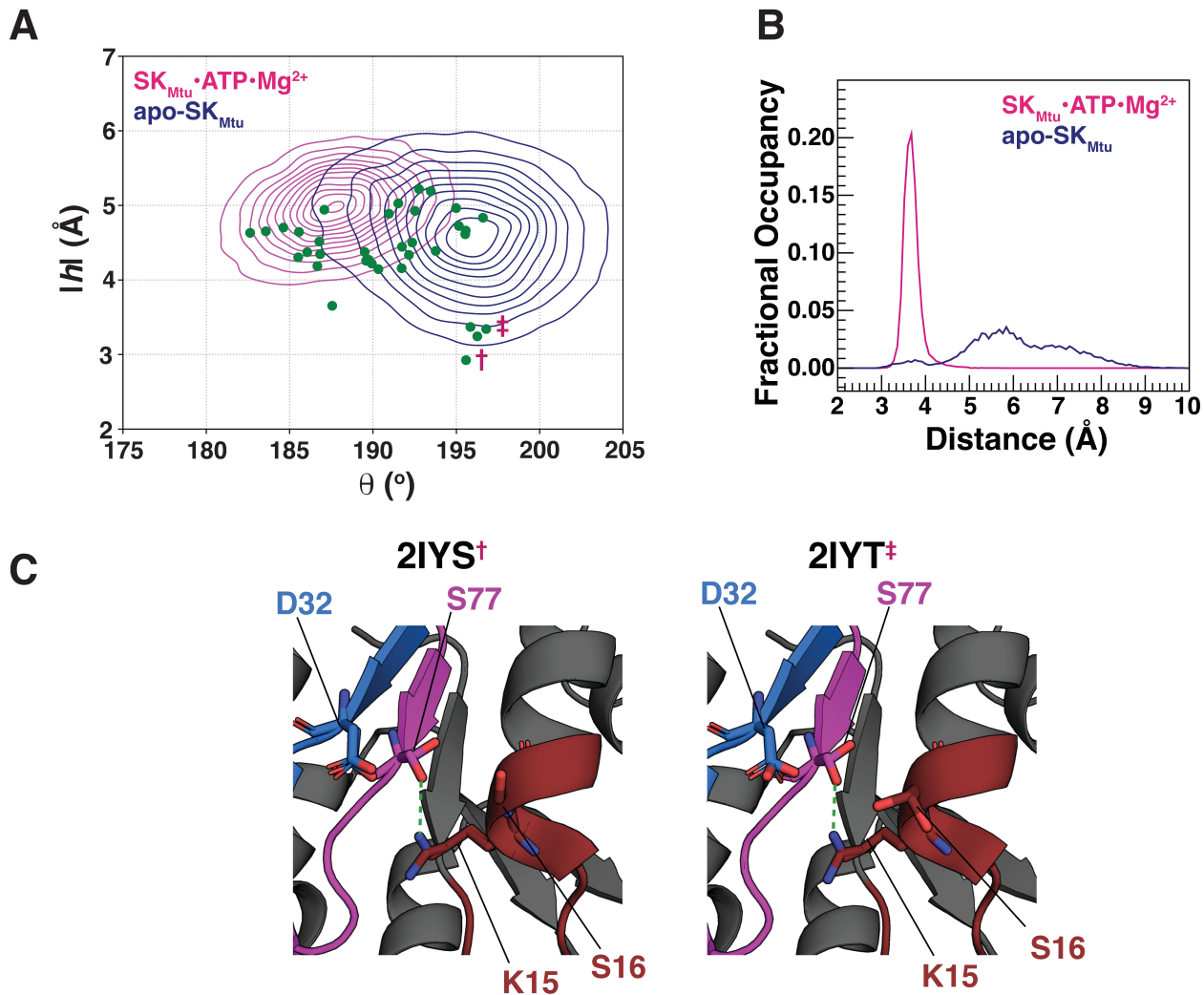


Figure S4. (A) The conformations sampled in the REST2-generated ensembles of apo-SK_{Mtu} (dark blue) and the SK_{Mtu}•ATP•Mg²⁺ complex (pink) are plotted using kernel density estimation and represented as contours. Projections of the conformations seen in the crystal structures of several SK enzymes (and their complexes) are indicated by the green circles. The '†' and '‡' indicate two sets of Mg²⁺-free structures that are the most open with the highest θ and the lowest $|h|$ values. (B) Distributions of the distance between the D32, C γ (DxD motif) and the S16, O γ (Walker-B motif) atoms for apo-SK_{Mtu} (dark blue) and the SK_{Mtu}•ATP•Mg²⁺ complex (pink) in their respective REST2-generated ensembles are shown. (C) Crystal structures of unliganded SK_{Mtu}, indicated by the '†' and '‡' symbols in (A), illustrating proximity (green dashed line) between the catalytic Walker-A Lys (K15) and the Walker-B S77.

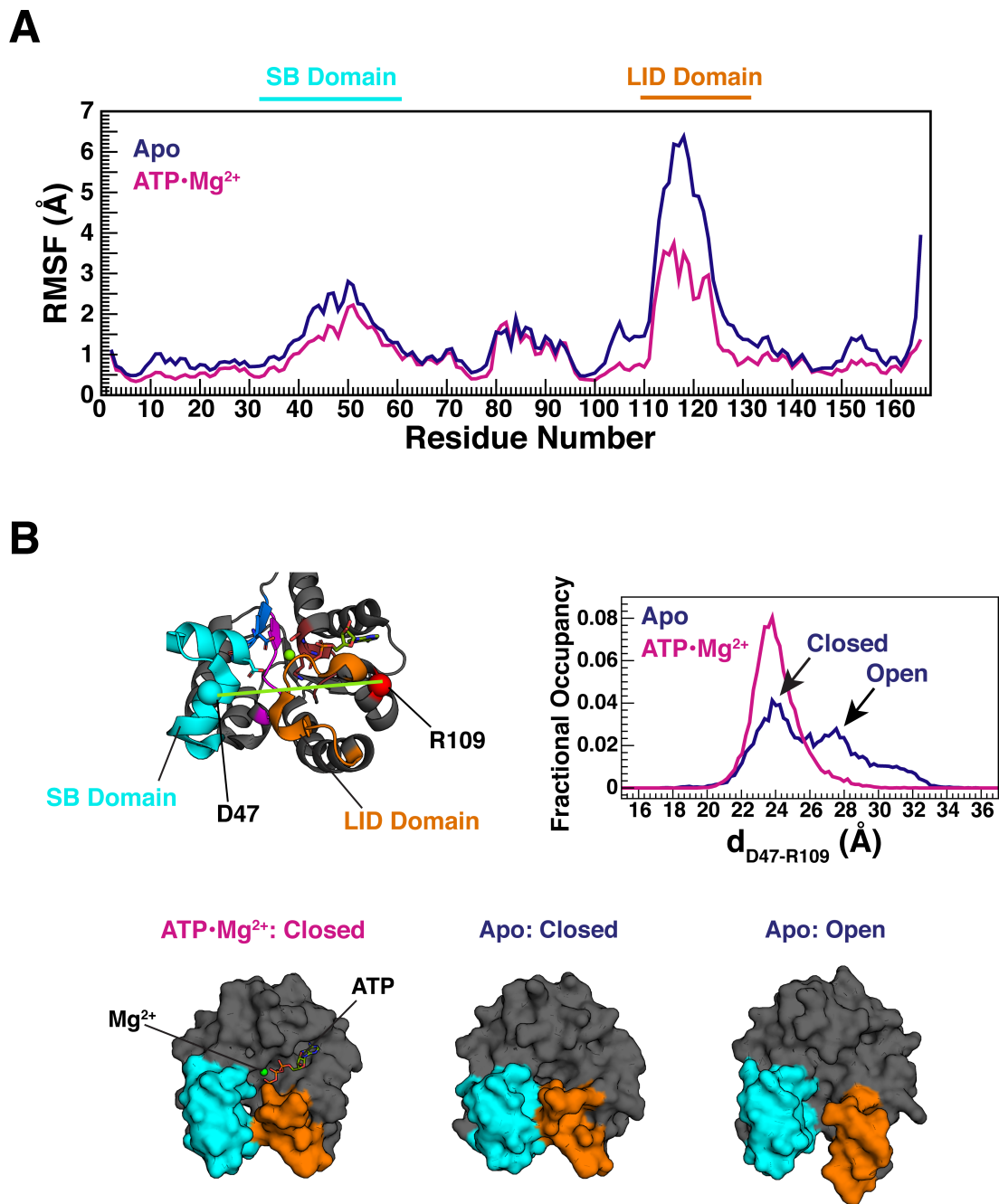


Figure S5. (A) Root mean squared fluctuations (RMSF) for apo-SK_{Mtu} (dark blue) and the SK_{Mtu}•ATP•Mg²⁺ complex (pink) in their respective REST2-generated ensembles are plotted against residue number. The positions of the SB (cyan) and LID (orange) domains along the sequence are indicated. (B) Distribution of the distance (d_{47-109} , shown as a light green line on the left panel) between the C α atoms of D47 (SB domain) and R109 (start of the LID domain) in the REST2-generated ensembles of apo-SK_{Mtu} (dark blue) and the SK_{Mtu}•ATP•Mg²⁺ complex (pink), are shown. The bottom panel shows representative structures drawn from the maxima of the distributions in each case.

Table S1. Shikimate kinase crystal structures used in the analysis

PDB ID	Organism	Ligand
1L4U	<i>M. tuberculosis</i>	ADP•Mg ²⁺
1L4Y	<i>M. tuberculosis</i>	ADP•Mg ²⁺
1U8A	<i>M. tuberculosis</i>	ADP•SA
1WE2	<i>M. tuberculosis</i>	ADP•Mg ²⁺ •SA
1ZYU	<i>M. tuberculosis</i>	AMPPCP•SA
2DFN	<i>M. tuberculosis</i>	ADP•SA
2DFT	<i>M. tuberculosis</i>	ADP•Mg ²⁺
2G1J	<i>M. tuberculosis</i>	Apo
2IYQ	<i>M. tuberculosis</i>	ADP•SA
2IYR	<i>M. tuberculosis</i>	SA
2IYS	<i>M. tuberculosis</i>	SA
2IYT	<i>M. tuberculosis</i>	Apo
2IYU	<i>M. tuberculosis</i>	ADP
2IYV	<i>M. tuberculosis</i>	ADP
2IYW	<i>M. tuberculosis</i>	ATP•Mg ²⁺
2IYX	<i>M. tuberculosis</i>	Apo
2IYY	<i>M. tuberculosis</i>	pSA
2IYZ	<i>M. tuberculosis</i>	ADP•pSA
3BAF	<i>M. tuberculosis</i>	AMPPNP•SA
4BQS	<i>M. tuberculosis</i>	ADP•SA
2G1K	<i>M. tuberculosis</i>	SA
1ZUH	<i>H. Pylori</i>	Apo
1ZUI	<i>H. pylori</i>	SA•PO ₄ ²⁻
3HR7	<i>H. pylori</i>	Apo
3MRS	<i>H. pylori</i>	Apo
3MUF	<i>H. pylori</i>	ADP•pSA
1E6C	<i>E. chrysanthemi</i>	Apo (Walker-A (K15M))
1SHK	<i>E. chrysanthemi</i>	Apo
2SHK	<i>E. chrysanthemi</i>	ADP•Mg ²⁺
3TRF	<i>C. burnetii</i>	Apo
3VAA	<i>B. thetaiotaomicron</i>	Apo
4Y0A	<i>A. baumannii</i>	SA
2PT5	<i>A. aeolicus</i>	Apo
1VIA	<i>C. jejuni</i>	Apo
1KAG	<i>E. coli</i>	Apo

The liganded states and mutations (if any are indicated; SA shikimate). Structures are labeled as “apo” if they do not contain nucleotides (or variants thereof) and/or shikimate. The structure of *H. pylori* SK (3N2E) complexed to an inhibitor has not been included in the analysis.

Table S2. Residues used for points P1-P4 to define the cylindrical coordinate frame in SKs from a variety of organisms

Organism	P1	P2	P3	P4
<i>M. tuberculosis</i>	15-19	21-25	33-37	155-159
<i>H. pylori</i>	14-18	20-24	32-36	150-154
<i>D. chrysanthemi</i>	15-19	21-25	33-37	157-161
<i>C. burnetii</i>	18-22	24-28	36-40	161-165
<i>B. thetaiotaomicron</i>	14-18	20-24	32-36	162-166
<i>A. baumannii</i>	31-35	37-41	49-53	172-176
<i>A. aeolicus</i>	13-17	19-23	31-35	150-154
<i>C. jejuni</i>	17-21	23-27	35-39	153-157
<i>E. coli</i>	17-21	23-27	35-39	159-163