

An ATPase with a twist: a unique mechanism underlies the activity of the bacterial tyrosine kinase, Wzc.

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: rghose@ccny.cuny.edu

Short Title: An ATPase with a twist

Teaser: While BY-kinases share many common features with P-loop enzymes, these are uniquely deployed for tyrosine phosphorylation.

BY-kinases constitute a protein tyrosine kinase family that encode unique catalytic domains that deviate from those of eukaryotic kinases resembling P-loop nucleotide triphosphatases (NTPases) instead. We have utilized computational and supporting biochemical approaches using the catalytic domain of the *Escherichia coli* BY-kinase, Wzc, to illustrate mechanistic divergences between BY-kinases and NTPases despite their deployment of similar catalytic motifs. In NTPases, the “arginine finger” drives the reactive conformation of ATP while also displacing its solvation shell thereby making favorable enthalpic and entropic contributions towards $\beta\gamma$ -bond cleavage. In BY-kinases, the reactive state of ATP is enabled by Mg^{2+} -ATP-induced global conformational transitions coupled to the **conformation** of the Walker-A lysine. While the BY-kinase arginine finger does promote the desolvation of ATP, it does so indirectly by generating an ordered active site in combination with other structural elements. Bacteria, using these mechanistic variations, have thus repurposed an ancient fold to phosphorylate on tyrosine.

INTRODUCTION

The BY-kinases (for bacterial tyrosine kinases) (*1-4*) comprise the largest family of protein tyrosine kinases in the bacterial taxa and are highly conserved in both Gram-positive and Gram-negative species (*5, 6*). BY-kinases have been shown to participate in a variety of cellular processes including the synthesis and export of polysaccharides involved in the formation of biofilms (*7*) and capsules (*8*), in lysogeny (*9*), in regulating the heat shock response (*10*), in DNA metabolism (*11*), just to name a few. BY-kinases have been suggested to be key virulence factors in certain drug-resistant bacterial strains (*12*).

The enzymatic activity of BY-kinases is encoded within a cytoplasmic catalytic domain (CD) with an architecture that is distinct from that of the eukaryotic protein kinases (ePKs) (*13-15*). Instead of the dual-lobed structure **characteristic of** the ePKs, BY-kinases have repurposed (*16*) an ancient fold, that of the P-loop NTPases (*17*), to phosphorylate on tyrosine. BY-kinases, **in lieu** of sequences characteristic of ePKs (*18*), contain those that are **found in** P-loop NTPases (*19, 20*). These include variations on Walker-A (A/GxxxxxGK[S/T], x is any residue) and Walker-B ($\phi\phi\phi\phi DxxP$, ϕ is a hydrophobic residue; the xxP sequence is unique to BY-kinase) motifs, together with an additional Walker-B-like sequence, called the Walker-A' motif ($\phi\phi\phi\phi DxDxR$). BY-kinases also contain a cluster of 5-7 tyrosine residues (the Y-cluster) located on their C-terminal tails. The Y-cluster tyrosines are sites for autophosphorylation (*21*). Additional structural motifs such as the oligomerization motif, characterized by an Ex_2Rx_2R sequence (*13, 14*), and a cluster of basic residues called the RK-cluster (that is partially disordered in the crystal structure) have been shown to contribute to function (Figure S1A).

While the precise mechanism by which BY-kinases achieve functional activation is not well understood, their activity appears to be linked to the phosphorylation state of the Y-cluster

tyrosines (7, 8, 22-26). It has been proposed that BY-kinases function through a phosphorylation-coupled assembly/disassembly process. In this model, oligomer formation enables the interaction between substrate-acting (S-acting) and enzyme-acting (E-acting) protomers to facilitate trans-phosphorylation. In BY-kinases, the functional entity appears to be an octameric ring in which the monomers are organized in a front-to-back arrangement (Figure S1B) (13, 14). This **mode of assembly** allows each monomer (I) to simultaneously function as an S-acting unit, in which its Y-cluster inserts into the active site of the following monomer (I+1), and an E-acting unit, in which its active site receives the Y-cluster of the preceding monomer (I-1). It has been suggested that the octameric ring becomes **increasingly** unstable with each phosphorylation event and ultimately dissociates into its constituent monomers once a certain level of Y-cluster phosphorylation has been achieved (13, 14). The Y-cluster tyrosines are then progressively dephosphorylated by protein tyrosine phosphatases (PTPs) (27-29); the oligomeric state is reconstituted once the Y-cluster is sufficiently dephosphorylated, thus reinitiating the cycle (Figure S1C). A defining feature of this mechanism is the fact that unlike in many ePKs such as c-Src (30) or ERK2 (31), kinase activity is not linked to the phosphorylation state of any specific Y-cluster tyrosine, but rather on the overall level of Y-cluster phosphorylation. Indeed, removal of any individual tyrosine has no appreciable impact on function but removal of several of them results in aberrant physiology (8, 23).

Given their similarities with, and apparent evolutionary origins from P-loop NTPases (4), BY-kinases, not surprisingly, possess appreciable ATPase activity (32) (Figure S2). Therefore, to prevent futile ATP hydrolysis, substrate binding i.e., oligomerization and the placement of an unphosphorylated Y-cluster tyrosine of the S-acting subunit at the active site of the E-acting subunit, must precede ATP•Mg²⁺ binding. Using the catalytic core of the archetypal BY-kinase,

Escherichia coli (K12) Wzc (this construct, Wzc_{CDAC} lacks the C-terminal tail and the Y-cluster therein) and enhanced sampling molecular dynamics (MD) simulations in concert with supporting biophysical studies, we established a mechanistic basis through which BY-kinases temporally regulate binding of its dual substrates, the Y-cluster of a flanking monomer and ATP•Mg²⁺ (33). Using a two-dimensional projection, defined by an angle θ and a rise $|h|$ (see Figure S3 for details), of the global conformation of Wzc_{CDAC}, we established that unliganded monomeric Wzc_{CDAC} exists in an open state (OS, that is more extended than that seen for the individual monomers of the crystallographic octamer). The OS is incapable of stably coordinating ATP•Mg²⁺ or forming an octameric ring. Binding of ATP•Mg²⁺ (or ADP•Mg²⁺) induces partial population transfer to a closed state (CS, similar to those of the crystallographic monomers) that can optimally bind ATP•Mg²⁺ but shows distortions in one or both of the oligomerization interfaces (Figure S1A), either through a bend in the $\alpha 2$ helix (I₁) or through partial unfolding of the $\alpha 7$ and/or $\alpha 9$ helices (I₂). These interfaces are stabilized upon oligomer formation allowing the efficient engagement of ATP•Mg²⁺ and enabling the progression to chemistry.

Our earlier work provided valuable insight into the regulatory mechanisms that couple oligomerization and nucleotide exchange in BY-kinases. During that study, we also developed a set of novel theoretical/computational tools to facilitate the robust analyses of the conformational landscape of Wzc_{CDAC} and its relevant complexes (33). Here, we redeploy several of those tools to explore specific structural features that influence the local and global conformations of Wzc_{CDAC} and those of its ATP substrate within the pre-chemistry Wzc_{CDAC}•ATP•Mg²⁺ complex, and lead to a catalytically competent state. Our computational approaches and supporting biochemical assays utilize Wzc_{CDAC} in wild-type form and functional mutants thereof (14), allowing the assessment of specific factors that lead to activation. Our results suggest that correlated conformational changes

on multiple length scales involving specific structural elements of the enzyme and conserved residues therein, the bound ATP•Mg²⁺, and solvent collectively contribute to form a chemistry-competent state. Our results also reveal specific points of mechanistic divergence between NTPases and BY-kinases despite the deployment of similar structural motifs.

RESULTS

The conformation of the Walker-A lysine is coupled to the global transitions of the catalytic domain

In our previous study (33), we had probed the conformational states of wild-type Wz_{CDAC}. For the various simulations, the procedures used to generate intact wild-type starting structures *in silico* from the available crystal structure of the Walker-A K540M mutant of the catalytic domain of *E. coli* Wzc (PDB: 3LA6) (14) were described at length previously (33). Similar approaches are deployed to generate starting structures of the various mutants described here (see Supplementary Materials for additional details).

As described previously (33), a key structural feature in the OS of unliganded Wz_{CDAC} is the presence of a salt-bridge between the conserved residues, K540 on Walker-A, and D642 on Walker-B. This conformation is unsuitable for optimally engaging Mg²⁺. The presence of nucleotide•Mg²⁺ results in a partial release of this constraint within the structural ensemble allowing the CS to be formed in subset of structures. In the CS, K540 disengages from D642, allowing the latter to form a hydrogen-bond with T541. This conformation enables the optimal engagement of Mg²⁺, and therefore of nucleotide•Mg²⁺. However, the formation of the CS and establishment of the T541-D642 hydrogen-bond results in a destabilization of the two oligomerization interfaces (as mentioned earlier). Thus, for monomeric Wz_{CDAC}, the OS↔CS

transition rates (or the relative populations of the two states in the structural ensemble) are influenced by the opposing effects of an OS stabilized by the presence of the K540-D642 salt bridge (and intact oligomerization interfaces) or a CS stabilized by the Mg^{2+} -induced T541-D642 interaction (and distorted oligomerization interfaces). In this mechanism, it is apparent that the conserved Walker-A K540 represents a key structural component that couples the global conformational changes to a remodeling of the active site. This supplements the presumed catalytic role of K540 (by analogy to P-loop NTPases) in activating the γ -phosphate of ATP for chemistry. To verify this unique secondary function of K540, we performed REST2 (34) simulations on the K540M mutant of WZC_{CDAC} (WZC_{CDAC,K540M}). This mutant is unable to hydrolyze ATP (Figure S2) and does not possess kinase activity (14). Strains expressing the K540M mutant are severely compromised in their ability to produce the polysaccharide colanic acid, a key component of *E. coli* (K12) biofilms (7). The REST2 simulations described in this manuscript utilized the same overall protocols as described previously (33) and are also provided in the Supplementary Materials.

Projection of the conformations sampled in the REST2 simulations of unliganded (apo) WZC_{CDAC,K540M} onto $\theta-|h|$ space (see Figure S3 for details) reveals a distribution that is markedly different from that seen for the WZC_{CDAC}, and characterized by the absence of a well-defined OS, as in the case of the latter (Figure 1A). Apo-WZC_{CDAC,K540M} samples an ill-defined set of states ranging from those that are somewhat more closed (e.g., C1 in Figure 1A) to those that are significantly more open (e.g., C3 in Figure 1A) than the OS in WZC_{CDAC}. This diffuse distribution results from the absence of the constraint imposed by the K540-D642 salt-bridge in the K540M mutant. Instead M540 is buried within the protein core and stabilized by hydrophobic interactions with M531 and V665 (Figure 1B, left panel). A similar mode of interaction is also seen in the

crystal structure of the Walker-A Lys mutant (K15M) of *Erwinia chrysanthemi* shikimate kinase (SK; PDB: 1E6C; Figure 1B, right panel) (35). Evaluation of the local conformations of the structures (Figure 1C) using our previously described clustering approach, EVA-MS (33) (see Figure S4 for details of the analysis) indicates two major structural families. Structures populating the more extended conformation (the C3 region in θ - $|h|$ space) display no remarkable features. Structures corresponding to the less open conformation (the C1 region), display features e.g., bending of helix α_2 , and **resultant** destabilizing of the first oligomerization interface (I_1), that are reminiscent of structures that populate the CS of the $Wz_{CD\Delta C} \bullet \text{nucleotide} \bullet \text{Mg}^{2+}$ complexes (33). This observation reinforces our previous suggestion that the amount of “strain” introduced into the structure is correlated to the degree of closure. In the absence of other stabilizing effects in the monomer, this “strain” must be “released” through other means e.g., destabilization of the oligomerization interfaces (see Figure 8 from Hajredini *et al.* (33)).

As mentioned earlier, the presence of $\text{nucleotide} \bullet \text{Mg}^{2+}$ results in the partial transfer of population from the OS to the CS **for $Wz_{CD\Delta C}$** . The CS then becomes the dominant state, being populated to the extent of 45% and 60% in the $\text{ATP} \bullet \text{Mg}^{2+}$ and $\text{ADP} \bullet \text{Mg}^{2+}$ complexes, respectively; the corresponding populations of the OS, that **still** remains significantly populated, are 28% and 34% (33). In contrast, for the $\text{ATP} \bullet \text{Mg}^{2+}$ complex of the K540M mutant, a CS, that is almost identical to its wild-type counterpart is the only major state seen; no OS or OS-like states are observed (Figure 2). For the $Wz_{CD\Delta C,540M} \bullet \text{ADP} \bullet \text{Mg}^{2+}$ complex, while the distribution deviates slightly from the CS **in the corresponding $Wz_{CD\Delta C}$ complex**, the distribution **remains** unimodal and there is no OS-like state that is significantly populated. Based on these results, it can be argued that in the case of $Wz_{CD\Delta C, K540M}$, the presence of $\text{nucleotide} \bullet \text{Mg}^{2+}$ substantially stabilizes the CS (or CS-like states) relative to the OS-like states (populated in the absence of ligands), and prevents

transitions to the latter. Given that transitions to the OS are more disfavored in the nucleotide•Mg²⁺-bound state of **WzC_{CDΔC},K540M** compared to that of **WzC_{CDΔC}**, one may expect that the affinity of the **K540M** mutant for nucleotide•Mg²⁺ should be higher than that of the wild-type protein. Indeed, the K540M mutant has a measurably higher affinity for ADP•Mg²⁺ over a range of Mg²⁺ concentrations (Figure S5; $K_D=0.5 \pm 0.2 \mu\text{M}$ at 5 mM Mg²⁺ for **WzC_{CDΔC},K540M** compared to $2.6 \pm 1.1 \mu\text{M}$ under the same conditions for **WzC_{CDΔC}**). **It is notable that the K15M mutant of *E. chrysanthemi* SK, mentioned above, has a roughly 6-fold higher affinity for ADP than the wild-type enzyme (35).**

As an additional experimental confirmation of the nature of the states populated by the K540M mutant we relied on gel filtration analysis. We had previously shown that the extended geometry of the OS is not compatible with formation of an octameric ring (33). Given that the simulations described above suggest that the conformational space of unliganded K540M mutant is populated by states by that are even more open than the wild-type OS, one can predict that WzC_{CD},K540M (this construct includes the C-terminal tail that is one of the requirements for stable oligomer formation) would be unable to form an oligomer even when the Y-cluster is fully dephosphorylated, in contrast to wild-type WzC_{CD} (14, 33). Additionally, the fact that only an oligomerization-capable CS is populated in the nucleotide•Mg²⁺ complexes of WzC_{CDΔC},K540M suggests that addition of nucleotide•Mg²⁺ (or Mg²⁺ alone) and induction of the T541-D642 hydrogen-bond, should drive oligomer formation. Indeed, incubation of monomeric WzC_{CD},K540M with ADP•Mg²⁺ or ATP•Mg²⁺ results in the formation of a substantial population of the oligomeric species. Mg²⁺ alone, but not by ADP alone, is able to drive this transition (Figure 3), in line with our predictions.

Global closure drives ATP towards its reactive conformation

It has been suggested that the cleavage of the phosphoanhydride bond between the β - and γ -phosphates of an NTP is mediated by a conformation in which the α -, β -, and γ -phosphates groups **are in** a high-energy eclipsed conformation (36). In P-loop NTPases, the conserved Walker-A lysine and Mg^{2+} serve to lock the β - and γ -phosphates of the NTP in the same plane, while the so-called arginine finger (37, 38) or alternatively, a K^+ (or NH_4^+) ion (39), positioned between the α - and γ -phosphates rotates the locked $\beta\gamma$ -diphosphate moiety with respect to the α -phosphate resulting in **an** eclipsed conformation in which all three non-bridging phosphates lie on the same plane (Figure S6). Given the evolutionary linkage between P-loop NTPases and BY-kinases (4), we wondered whether similar effects also facilitate a chemistry-competent state in the latter.

To compare the conformations of ATP sampled in the REST2 simulations of the $\text{ATP}\cdot\text{Mg}^{2+}$ complexes of $\text{WZ}_{\text{CD}\Delta\text{C}}$ and of its K540M mutant, we defined three dihedral angles, $\varphi_{\beta\gamma}$, $\varphi_{\alpha\gamma}$ and $\varphi_{\alpha\beta}$ between the non-bridging phosphate oxygens of ATP (Figure 4). In our definition, a counterclockwise rotation of each angle towards 0° brings the corresponding phosphates into an eclipsed state, whereas a clockwise rotation towards -60° leads to a staggered conformation. For the $\text{WZ}_{\text{CD}\Delta\text{C}}\cdot\text{ATP}\cdot\text{Mg}^{2+}$ simulations, bimodal distributions of dihedral angles are seen in all cases. The first peak represents structures of the more closed states (this includes the CS, and the so-called hyper-closed state, HS, that is even more closed than the CS and populated to the extent of 11%; see Figure 2A) that show a $\varphi_{\beta\gamma}$ distribution centered around an almost fully eclipsed conformation (6° , Figure 5). The corresponding $\varphi_{\alpha\gamma}$ and $\varphi_{\alpha\beta}$ angles also deviate away from their respective staggered conformations (-60°) and towards the eclipsed conformation (0°) with distributions centered around -33° and -49° , respectively. In contrast, for the second peak that corresponds to the structures of the OS, the $\varphi_{\alpha\gamma}$ angle and the all-important $\varphi_{\beta\gamma}$ angle are both

distributed around staggered conformations with values of -132° and -48° , respectively; the $\phi_{\alpha\beta}$ angle, however, samples a near-eclipsed conformation (-123°). For the $Wz_{CDAC,K540M}\bullet ATP\bullet Mg^{2+}$ complex, that exclusively populates a CS (Figure 2A), the $\phi_{\beta\gamma}$ distribution shows a single peak with a maximum that deviates away from the eclipsed conformation compared to the wild-type, and is centered around 18° . The $\phi_{\alpha\beta}$ distribution is bimodal with maxima centered at -56° and -72° both of which are close to the staggered conformation. The $\phi_{\alpha\gamma}$ distribution is similar to the wild-type case (centered at -34° , Figure 5). Thus, while there are some differences in the ATP conformations in the closed states (CS/HS) corresponding to the $Wz_{CDAC}\bullet ATP\bullet Mg^{2+}$ and $Wz_{CDAC,K540M}\bullet ATP\bullet Mg^{2+}$ (CS) complexes, the most substantial differences are seen between the wild-type CS/HS and the corresponding OS. Based on these observations, one may conclude that ATP is in its most reactive conformation in the wild-type CS/HS and in its least reactive conformation in the OS. This sub-optimal ATP conformation in the OS results from the presence of the K540-D642 salt-bridge and the unproductive engagement of Mg^{2+} in this state. For the K540M mutant, even though only a CS is populated, a deviation the $\phi_{\beta\gamma}$ angle away from the eclipsed, and of the $\phi_{\alpha\beta}$ angle towards the staggered conformation, also suggests an ATP **is in a conformation** that is further away from **one that is suitable for chemistry compared to** the wild-type CS/HS, though less so than the wild-type OS. Of course, a quantitative assessment of the relative reactivities of ATP in these various states will require quantum mechanical approaches.

In the absence of quantitative measures of the reactivity of ATP based on its conformation discussed above, to obtain a qualitative reference, we carried out classical MD simulations on MinD, a close structural homolog of Wz_{CD} and a *bona fide* ATPase of the P-loop family (40). MinD has been shown to dimerize along its ATP binding pocket, to enable the so-called deviant Walker-A lysine (41) to insert into the active site of the neighboring protomer and coordinate the

α - and γ -phosphates of ATP, in much the same fashion as the arginine finger mentioned above (Figure 6A), to activate ATP for hydrolysis. For monomeric MinD, a broad $\phi_{\beta\gamma}$ distribution centered around the eclipsed conformation (-3°) is seen, while the $\phi_{\alpha\gamma}$ (-48°) and $\phi_{\alpha\beta}$ (-53°) distributions are centered closer to the staggered conformation (Figures 6B, C). In contrast, in dimeric MinD, insertion of the deviant Walker-A lysine (K11) into the active site leads to a significant shift of the $\phi_{\alpha\gamma}$ angle towards the eclipsed conformation with the maximum of the distribution shifting to -31° (Figures 6B, C). The center of the $\phi_{\alpha\beta}$ distribution also shifts towards a more eclipsed conformation being centered around -46° . However, the rotation of γ -phosphate with respect to α -phosphate leads to a slight shift (7°) of $\phi_{\beta\gamma}$ away from 0° . The dimeric state of MinD, that represents an active ATPase, features an ATP conformation that is very similar to those seen for the CS of wild-type WzcCDAC (see Figure S7). This suggest that the formation of closed states (CS/HS) in WzcCDAC, like the dimerization of MinD, represents a drive towards the chemistry-competent all-eclipsed state of the bound ATP.

Interestingly, in all our simulations of MinD, a Na^+ ion (added in the simulations for charge balance) is always associated with ATP and occasionally inserts between α - and γ -phosphates in a fashion analogous to the deviant Walker-A lysine (Figure 6B, C). As mentioned earlier, some NTPases deploy monovalent cations *in lieu* of an arginine finger (or a deviant Walker-A lysine) (39). In most cases, K^+ acts as a more significant enhancer of catalytic activity than Na^+ . Indeed, in the presence Na^+ , the distribution of the $\phi_{\alpha\gamma}$ angle, and to some extent, the $\phi_{\beta\gamma}$ angle, show maxima that are intermediate between those of the monomeric and dimeric states of MinD.

A Y569F mutation has little effect on the conformation of bound ATP

It has been shown that even a conservative (Y569F) mutation of Y569 leads to a significant decrease in Y-cluster phosphorylation, and *E. coli* cells expressing the Y569F mutant produce

exopolysaccharides with aberrant sizes (7). Some of these effects have been attributed to a ~6-fold decrease in nucleotide affinity (14) resulting from the loss in the hydrogen-bond between the tyrosyl -OH and the α -phosphate of ATP in the Y569F mutant (14). Given that Y569 contacts ATP, we wondered whether an absence of this contact in the Y569F mutant and its altered activity could also be attributed to perturbed $\phi_{\beta\gamma}$, $\phi_{\alpha\gamma}$ and $\phi_{\alpha\beta}$ distributions. Our earlier REST2 simulations on the $W_{ZC_{DAC}} \cdot \text{ATP} \cdot \text{Mg}^{2+}$ complex had suggested that Y569, in addition to contacting ATP, collaborates with residues R490 and K492 of the RK-cluster in stabilizing the HS mentioned above. In the HS (see Figure S8), Y569 stacks against K492, allowing the latter to form a salt-bridge with E572. R490 inserts into the active site and contacts the α - and γ -phosphates in a fashion analogous to the arginine finger in P-loop NTPases or the deviant Walker-A lysine in the MinD dimer. These interactions result in an “assembled” active site in which the RK-cluster is relatively well-ordered (unlike in the crystal structure and unlike in a majority of conformations sampled in our REST2 simulations) and all key catalytic elements including ATP, are in their appropriate positions to facilitate chemistry (33). Given the apparent structural and functional importance of Y569, and to probe its influence on the conformational landscape of the kinase, we performed REST2 simulations on the $W_{ZC_{DAC}, Y569F} \cdot \text{ATP} \cdot \text{Mg}^{2+}$ complex and compared them with corresponding simulations on the $W_{ZC_{DAC}} \cdot \text{ATP} \cdot \text{Mg}^{2+}$ complex reported earlier (33).

In contrast to the wild-type complex, the Y569F mutant (Figure 7A) samples a relatively diffuse cluster of mostly CS-like states though a substantial proportion of these are more open than the wild-type CS (e.g., C1-C4 in Figure 7A). No conformations that resemble the wild-type OS are seen. In the absence of an OS or similarly open states, ATP remains largely anchored to its binding site in the $W_{ZC_{DAC}, Y569F} \cdot \text{ATP} \cdot \text{Mg}^{2+}$ complex. EVA-MS decomposition of the ensemble into its constituent states (Figures S9 and S10) suggests that for the states that are more closed

(i.e., more like the wild-type CS; e.g., C5 in Figure 7A), the positions of R490 and K492 are reversed with respect to the wild-type HS. Instead of stacking with K492, as in the wild-type HS, Y569 now stacks with R490; K492, rather than R490, now contacts ATP (Figure S8). For states that are more open (e.g., C1 in Figure 7A), F569 is highly dynamic, as is the RK-cluster and R490 and K492 therein, resulting in a more open (“disassembled”) active site (Figure S8). Additionally, all the clusters, given their states of closure (except C2 to some extent), show some degree of distortion at the oligomerization interfaces (I₁, Figure S10). C5 that comprises the cluster with the highest degree of closure is the most distorted, with a bent α 2 (I₁ destabilized), and a partially unfolded α 7 (I₂ destabilized). This further reinforces our suggested correlation between closure and structural strain in the monomeric complexes with nucleotide•Mg²⁺.

To parse out the influence of Y569 on the dihedral angles of ATP, we compared its conformation when engaged to the Y569F mutant with those seen for the wild-type CS. The dihedral angles sampled by Y569F in its complex with ATP•Mg²⁺ are similar to those of the wild-type CS (Figure 7B). This suggests, quite surprisingly, that the presence/absence of the hydrogen bond between Y569 and ATP does not have a major influence on its overall conformation. As mentioned earlier, in the wild-type HS, R490 contacts both the α - and γ -phosphates of ATP and in doing so, mimics the arginine finger of P-loop NTPases. In contrast in the wild-type CS, R490 contacts only the γ -phosphate (Figure S8). Thus, if indeed R490 affects the conformation of ATP in a fashion similar to an arginine finger, one would expect a greater closure of the $\phi_{\alpha\gamma}$ angle in the HS compared to the CS (and the Y569F mutant). Surprisingly, the $\phi_{\alpha\gamma}$ distribution for the HS is shifted more towards a staggered conformation (-42°) compared to either the wild-type CS (-30°) or the Y569F mutant (-32°) (Figure 7B). However, the wild-type HS displays a $\phi_{\beta\gamma}$ distribution that has a maximum corresponding to a more eclipsed state (-1°) compared to either the wild-type

CS (11°) or the Y569F mutant (9°). The $\phi_{\alpha\beta}$ distributions are almost identical in the three cases. Thus, R490 **seems** unable to affect the ATP conformation in a fashion similar to an arginine finger in conventional P-loop NTPases. Further, this observation suggests that despite having the catalytic elements appropriately coordinated to ATP for chemistry, the HS does not appear to represent a true on-pathway intermediate to the transient all-eclipsed ATP **geometry** of the transition state. Therefore, additional conformations, not significantly sampled in our REST2 simulations likely represent a pathway intermediate to the transition state in which the overall **arrangement** of **the** catalytic elements are HS-like but the $\phi_{\beta\gamma}$, $\phi_{\alpha\gamma}$ and $\phi_{\alpha\beta}$ are all closer to the all-eclipsed conformation. Perhaps stable engagement of the substrate Y-cluster is needed to achieve such a state. Taken together our data suggest that despite differences in specific contacts to ATP between the wild-type CS, the wild-type HS and the Y569F mutant, the overall differences in the conformation of ATP appears to be somewhat smaller than the radically different conformation seen in the wild-type OS. One could **therefore** suggest that while differences in local contacts of ATP can lead to subtle changes in its conformation, the largest changes in conformation are driven by transitions from the open to the closed states.

Active site ordering promotes the desolvation of ATP

The importance of Y569 and the RK-cluster, especially residues R490 and K492 therein, in maintaining the assembled state of the active site has been discussed above. The functional importance of the latter residues is also evident from the fact that R490A and K492A mutations each result in substantial reductions in the kinase activity of Wzc (*14*). We wondered whether additional factors could also contribute to the loss of activity in RK-cluster and Y569F mutants. It is evident that in assembled states like the HS in which the active site elements are ordered and the ATP is properly coordinated (Figure S8) would be significantly protected from solvent. Indeed,

exclusion of water molecules in the fully active state and the presence of the substrate Y-cluster would be one of the requirements for the dominance of the kinase over the ATPase function of Wzc. QM/MM calculations by Takahashi et al. (42) have shown that NTP hydrolysis is favored by decreased NTP-solvent interactions due to a stabilizing effect of water on the NTP. In the Ras-GAP complex, insertion of the arginine finger into the active site displaces several water molecules from the vicinity of the triphosphate of GTP and this reflects as a favorable entropic contribution to catalysis (43). Furthermore, differential solvation patterns have been demonstrated to account for activity levels in the ATPase Hsp90 (44) that lacks the P-loop (45), suggesting that desolvation of the bound NTP is perhaps a general requirement for $\beta\gamma$ -bond cleavage.

For the global states sampled in the in the various ATP•Mg²⁺ simulations, one could expect ATP to be better protected from solvent in the more closed states compared to the corresponding open states given its more optimal engagement in the former than in the latter (as discussed above). Indeed, for the wild-type simulations, the distribution of solvent accessible surface area (SASA) of ATP is bimodal. ATP is more buried when engaged to wild-type CS/HS and highly exposed in the OS (Figure 8A). The ATP•Mg²⁺ complex of the K540M mutant populates only a CS. Thus, ATP is highly protected, and the SASA distribution shows a single peak with a maximum that is very similar to the maximum for the wild-type CS/HS (Figure 8A). For the Y569F simulations, that also samples mostly closed states (though some that are more open than the wild-type CS, but not as open as the wild-type OS) the SASA distribution is bimodal. The maximum of the peak at lower SASA values is similar to the wild-type CS/HS (and the K540M mutant). The maximum of the peak at higher SASA values indicate that ATP in these states, while more exposed than the wild-type CS/HS (or the K540M mutant), is significantly more buried than in the wild-type OS.

Next, to test the possible role of active site assembly and RK-cluster ordering in facilitating desolvation of the bound ATP, we applied the grid based discrete implementation of inhomogeneous solvation theory (GIST) (46) to representative structures from two specific structural states obtained in the REST2 simulations. The first, in which the RK-cluster is disordered, and the active site is open, and the second, that has an assembled active site. We utilized structures from the C1 cluster from the $W_{ZCCD\Delta C, Y569F} \bullet ATP \bullet Mg^{2+}$ simulations (Figure S8, bottom right panel) as representative of the former, and structures from the wild-type HS (Figure S8, top left panel) to represent the latter. Both of states have ATP engaged to its appropriate binding site (unlike the wild-type OS) allowing a direct comparison of the influence of active site ordering largely uncontaminated by other factors. Briefly, GIST analyzes water structuring in discrete regions around a protein solute through short MD simulations where the latter is kept “frozen” using position restraints. We focused our attention on water oxygen density, $\rho(O)$, around the triphosphate region of ATP in each case (Figure 8B, C). As expected, the C1 cluster of Y569F with a highly disordered RK-cluster reveals enhanced solvent density around the triphosphate (Figure 8B, white arrows). On the other hand, the Y569-mediated organization and ordering of the RK-cluster in the wild-type HS carves out a significant zone of solvent exclusion around the bound ATP suggesting complete desolvation of the triphosphate moiety (Figure 8C). In the absence of the interactions of ATP with Y569 (resulting from the Y569F mutation) and R490 (resulting from the disorder in the RK-cluster), the first solvation shell of the triphosphates consists of several highly ordered water molecules. Several of these waters are displaced in the HS and the corresponding interactions are replaced by those with enzymic sidechains (Figure S11). In fact, the guanidino and hydroxyl groups of R490 and Y569 in the wild-type HS are exactly in the same position as two specific water molecules (1 and 2 in Figure S11) in the Y569F C1. Obtaining

precise estimates of the entropic gains by the displacement of bound waters into bulk is non-trivial given the difficulties in estimating the entropic contributions resulting from positional and orientational correlations in bulk water (47). We can however provide estimate of 6.7 kcal/mol (referenced to an isotropic distribution of bulk water) as the upper bound of the free energy gain in displacing these bound waters (1 and 2 in Figure S11) comparing the two states at 300 K. Estimates for other displaced waters are shown in Figure S11. Thus, we can state, at least in a qualitative sense, that desolvation of ATP upon formation of an assembled active site (as in the wild-type HS) results in a favorable entropic contribution to the activation energy. A similar mechanism has been suggested for the Ras-GAP complex that, unlike in the present case, features a direct involvement of the arginine finger in desolvating ATP (43).

The CS is stabilized in the octameric ring with characteristic ATP conformations

As discussed above, the monomeric $Wz_{CD\Delta C}\bullet ATP\bullet Mg^{2+}$ complex populates a major CS that distorts the oligomerization interfaces, or a minor OS that sub-optimally engages $ATP\bullet Mg^{2+}$. We had suggested through preliminary simulations that the formation of the octameric ring stabilizes the oligomerization interfaces and the CS enabling the optimal engagement of $ATP\bullet Mg^{2+}$ (33). To confirm this hypothesis and further probe the conformational states of the wild-type species in the context of its oligomeric states we performed unrestrained classical MD simulations on the $ATP\bullet Mg^{2+}$ complexes of $Wz_{CD\Delta C}$ and Wz_{CD} in the context of the octameric ring. As noted in our previous experimental studies, formation of a stable oligomer in solution requires both an intact oligomerization interface as well as an unphosphorylated C-terminal tail (missing in $Wz_{CD\Delta C}$, but present in Wz_{CD}) (33). As expected, the absence of the C-terminal tail in $Wz_{CD\Delta C}$ results in an unstable ring that starts dissociating very early in the simulation (as assessed by an almost linear increase in the radius of gyration, R_g ; Figure S12A) and eventually

ruptures. In contrast, for Wz_{CD} , the complex persists over the entire length of the 200 ns simulation (Figure S12B). A 2-dimensional plot correlating R_g with the root-mean-squared-deviation (RMSD) over all $C\alpha$ atoms of all chains of the $[Wz_{CD}\bullet ATP\bullet Mg^{2+}]_8$ complex suggests that while the ring displays some deviations from C_8 symmetry, its overall integrity is maintained throughout the simulation (Figure S13). Assessment of the conformational landscape of $[Wz_{CD}\bullet ATP\bullet Mg^{2+}]_8$ in $\theta-|h|$ space shows that, unlike in the monomeric simulations, all the individual monomers exist in the CS with no evidence of a tendency to transition to the OS (Figure S14). Our previous classical MD simulations on the monomeric $Wz_{CD\Delta C}\bullet ATP\bullet Mg^{2+}$ complex showed a gradual CS $\rightarrow$ OS transition that is almost complete within the last 10 ns of a 100 ns simulation (33). Our results are thus consistent with the expectation that formation of the octameric ring serves to stabilize the CS.

Finally, we investigated the conformational states of ATP bound to each monomer of the $[Wz_{CD}\bullet ATP\bullet Mg^{2+}]_8$ complex using the set of dihedral angles defined above. For the 8 monomers, the $\phi_{\alpha\beta}$, $\phi_{\beta\gamma}$, and $\phi_{\alpha\gamma}$ angles show distributions centered around 11° to 13° , -28° to -32° , and -47° to -51° , respectively (Figure 9A). These values are very similar to those seen for the corresponding monomeric complex while in the CS (11° , -30° , and -49°). Thus, while formation of the octameric ring stabilizes the CS, the conformation of ATP is largely unchanged from that in the corresponding monomeric complex. Inspection of the active site geometry for individual monomers within the octameric complex reveals that key structural elements are largely in their appropriate conformations and remain so through the course of the simulation (Figure S15). R490 contacts the γ -phosphate of ATP, the K492-E572 distance oscillates about an optimal value allowing Y569 to stably contact the α -phosphate of ATP. The distributions of key distances (R490, C ζ -ATP, P γ ; Y569, O η -ATP, P α and K492, N ζ -E572, C δ) characterizing the active site

geometry are shown for the individual monomers of the $[\text{Wzc}_{\text{CD}}\cdot\text{ATP}\cdot\text{Mg}^{2+}]_8$ complex in Figure 9B. Shorter distances characteristic of an assembled active site are dominant in most cases. However, it is important to note that an optimally assembled active site is only obtained for the low-occupancy HS, that is not sampled in the classical MD simulations.

DISCUSSION

We have used a combination of computational and biochemical studies to investigate key factors that drive the enzymatic activity of the BY-kinase, Wzc. We have previously shown (33) that the Wzc catalytic core ($\text{Wzc}_{\text{CD}\Delta\text{C}}$) samples two major global conformational states, an open state (OS) that deviates significantly from that seen for individual monomers in the crystallographic octamer, and a crystal-like closed state (CS), in the presence of nucleotide and Mg^{2+} . The OS, that is the only major state in unliganded $\text{Wzc}_{\text{CD}\Delta\text{C}}$, is not compatible with the formation of a closed ring and is incapable of stably engaging $\text{ATP}\cdot\text{Mg}^{2+}$. The presence of $\text{ATP}\cdot\text{Mg}^{2+}$ populates the CS (~45%, OS: 28%) together with another state, the HS (~11%) that is more closed than the CS and has all the catalytic elements in an active-like conformation. $\text{ATP}\cdot\text{Mg}^{2+}$ is stably engaged in the CS and the HS. A major difference between the CS/HS and the OS is the conformation of the conserved Walker-A K540 that forms a salt-bridge with Walker-B D642 in the OS, while in the CS/HS, the latter forms a hydrogen bond with T541. On the other hand, formation of the T541-D642 hydrogen-bond in the CS/HS leads to a disruption of the two oligomerization interfaces, highlighting the inverse relationship between intact oligomerization interfaces and the presence/absence of the T541-D642 hydrogen-bond. We had previously shown that introduction of disorder at the C-terminal end of the $\alpha 2$ helix on the first interaction interface shifts the populations entirely to the CS in the $\text{ATP}\cdot\text{Mg}^{2+}$ complex (33). Here we show that the

K540M mutant in its ATP•Mg²⁺ complex also achieves the same effect but through a different mechanism. An unliganded K540M variant samples a diffuse set of global conformational states that display varying degrees of openness including some that are significantly more open than the wild-type OS; experimental data confirm that this mutant cannot oligomerize in the apo state. In the absence of a well-defined OS, the presence of ATP•Mg²⁺ leads to a near complete transition to the CS. The K540M mutant can oligomerize in the presence of Mg²⁺ alone but not in its absence even in the presence of nucleotide. Overall, these findings establish that the Walker-A K540 functions as a key node that couples global transitions with the active site geometry. Thus, by stabilizing a CS, K540 also likely directly influences nucleotide cycling. Our model, proposed earlier (33), suggested that the CS stabilized through oligomerization enables the affinities of ATP•Mg²⁺ and ADP•Mg²⁺ to become comparable and thus the bound ADP can be optimally exchanged out in the context of oligomer given that its cellular concentration is 6-10-fold lower than that of ATP (48). This would represent a characteristic of the Walker-A lysine that is unique to BY-kinases.

It has been suggested that the cleavage of the $\beta\gamma$ -phosphoanhydride bond of ATP is facilitated by a high-energy all-eclipsed conformation of the non-bridging oxygens (36). Consistent with this, the dihedral angles of ATP, most notably the crucial angle between the β - and γ -phosphates, are close to eclipsed in the wild-type CS/HS and in the dimeric ATPase-active form of the closely related P-loop ATPase, MinD. In the case of the K540M mutant of WzCCDAC that also forms a wild-type-like CS, this angle, though a little more open is still close to an eclipsed state. The wild-type OS, on the other hand, generates a staggered $\phi_{\beta\gamma}$ angle for the bound ATP, suggesting a state that is less conducive to chemistry. Surprisingly, altering specific ATP contacts e.g., removal of the hydrogen-bond between Y569 and the α -phosphate of ATP, as in the Y569F

mutant, have small effects on the overall conformational states of ATP. This suggests that the overall conformation of ATP is correlated with the global conformations of the catalytic domain rather than the details of local interactions. Indeed, the conformations of ATP (as defined by the phosphate dihedral angles) bound to individual monomers within the $[Wz_{CD}\bullet ATP\bullet Mg^{2+}]_8$ complex, that are all in a now-stabilized CS, are virtually identical to those seen in the CS of the monomeric $Wz_{CD\Delta C}\bullet ATP\bullet Mg^{2+}$ complex. We also note the sparsely populated HS sampled in the monomeric $Wz_{CD\Delta C}\bullet ATP\bullet Mg^{2+}$ complex, suggested to be a reactive conformation based on the positions of the enzymic sidechains and ATP (33), nevertheless shows a somewhat more staggered $\phi_{\alpha\gamma}$ angle (despite a nearly eclipsed $\phi_{\beta\gamma}$ angle) than the corresponding CS suggesting that it is likely lies off the reactive pathway. It is possible that the presence of a Y-cluster tyrosine is necessary to probe the corresponding on-pathway conformation. One can expect this conformation to resemble the HS sampled in the monomeric $Wz_{CD\Delta C}\bullet ATP\bullet Mg^{2+}$ complex with respect to the orientation of the sidechains and in the positioning of ATP, but with all the relevant dihedral angles of ATP trending towards their eclipsed conformations. Not surprisingly, we did not observe an HS-like conformation in our classical MD on the $[Wz_{CD}\bullet ATP\bullet Mg^{2+}]_8$ complex given the limited conformational sampling in these simulations. Perhaps the far more expensive REST2 simulations on $[Wz_{CD}\bullet ATP\bullet Mg^{2+}]_8$ or sufficiently long classical MD simulations using specialized hardware are needed.

We also investigated the role of active site ordering (e.g., in the wild-type HS) in facilitating chemistry by the desolvation of ATP resulting in a favorable entropic contribution to the free energy of activation (42). Indeed, a comparison of solvent density between the wild-type HS characterized by an ordered RK-cluster facilitated by Y569, E572 and ATP, and the C1 state of the Y569F mutant that has a highly disordered RK-cluster, reveals a substantial difference in ATP

hydration. Disorder in the RK-cluster allows more complete solvation of ATP, a feature that would disfavor chemistry. One could expect that solvent exclusion would be further enhanced by the presence of a Y-cluster tyrosine at the active site as would be the case when oligomerization precedes the binding ATP•Mg²⁺.

Based on the discussion above, it is evident that even though the catalytic domain of Wzc evolved from a P-loop NTPase scaffold (4), there are specific lines of apparent divergence between the two families with respect to how the bound ATP attains its reactive state. For P-loop NTPases (Figure 10A), the Walker-A lysine and the Mg²⁺ ion lock the orientations of the β- and γ-phosphates of ATP into an eclipsed conformation; insertion of the positively charged arginine finger between the α- and γ-phosphates leads to a rotation of the locked βγ-phosphate pair relative to the α-phosphate ultimately resulting in an all-eclipsed high-energy conformation which favors βγ-bond cleavage (37). Additionally, the arginine-finger displaces the solvation shell of NTP further favoring product formation (43). In Wzc_{CD} (Figure 10B) while the Walker-A lysine (K540) and Mg²⁺ ion play the same role as in conventional P-loop NTPases in locking the relative orientation of β- and γ-phosphates (in the CS/HS), the arginine finger equivalent (R490) does not appear to be the principal facilitator of the drive towards the reactive eclipsed conformation of ATP, rather it is the formation closed states (CS/HS) states that seem to do so. Within the CS/HS, the interaction of Y569 with the α-phosphate of ATP and with K492, the salt-bridge of the latter with E572, together with the insertion of R490 into the active site to contact ATP, collectively lead to an ordering of the active site thus significantly desolvating ATP **to facilitate** chemistry. Thus, in P-loop NTPases, the arginine finger plays three distinct roles – (1) it facilitates a transition of ATP to its reactive conformation, (2) it serves to desolvate ATP, and (3) in its “catalytic” role, stabilizes the developing negative charge during chemistry. It is apparent from our results above that the

catalytic domain of Wzc has evolved to separate these functions. The first of these appears largely to be driven by global conformational states (modulated by Walker-A K540 and induced by Mg^{2+}). Desolvation is facilitated by the formation of an ordered active site that is enabled by R490, but only with the support of Y569, E572 and K492. It is likely that R490 plays the assigned third arginine finger-like catalytic role in chemistry, but this needs confirmation through quantum mechanical calculations. Nevertheless, identification of apparently distinct structural origins of these effects in Wzc offers the opportunity to test them rationally by experimental means and additional rounds of carefully designed mutations.

MATERIALS AND METHODS

Classical and enhanced sampling molecular dynamics simulations

The starting structures for the REST2 simulations were derived in all cases using the procedures described previously (33) (also refer to the Supplementary Materials). For the $Wz_{CDAC,K540M}$ and $Wz_{CDAC,Y569F}$ simulations, the relevant mutants were generated using the mutagenesis tool in PyMOL from the respective starting conformations. For the classical MD simulations of the octameric ring, the starting structure was generated in the following way – a structure from the CS1 cluster of our $Wz_{CDAC} \bullet ATP \bullet Mg^{2+}$ REST2 simulations that most resembled a crystal structure monomer (note that the crystal structure was of the K540M mutant complexed with $ADP \bullet Mg^{2+}$) (14), i.e. in the CS with both interaction interfaces intact was selected and aligned to chains A through H of the crystal structure octamer generate an octamer missing the C-terminal tail i.e. $[Wz_{CDAC} \bullet ATP \bullet Mg^{2+}]_8$. The same starting monomer, as above, was aligned to chain A from the crystal structure and the corresponding C-terminal tail (residues 705-720) was linked to using the Rosetta-KIC (49) loop closure protocol and one round of modeling. This conformation

was then again aligned 8 times to chains A through H, as above, to generate the $[\text{Wz}_{\text{CD}}\cdot\text{ATP}\cdot\text{Mg}^{2+}]_8$ complex.

All REST2 and classical MD simulations were carried out using previously described protocols utilizing identical parameters as before (33) using 200 ns simulation times for individual replicas (see Supplementary Materials for additional details). A simulation time of 200 ns was also utilized for the classical MD on the $[\text{Wz}_{\text{CD}}\cdot\text{ATP}\cdot\text{Mg}^{2+}]_8$ complex while the corresponding simulations on the $[\text{Wz}_{\text{CDAC}}\cdot\text{ATP}\cdot\text{Mg}^{2+}]_8$ complex simulation was stopped after the ring disintegrated.

For the classical MD simulations on MinD, the structure of dimeric MinD (PDB: 3Q9L) (40) together with the bound $\text{ATP}\cdot\text{Mg}^{2+}$ and all crystallographic waters therein were used. For the corresponding monomer simulations, the bound $\text{ATP}\cdot\text{Mg}^{2+}$, and all neighboring solvent molecules were used. All simulation of 100 ns were carried out in triplicate using identical conditions and parameters, as above.

Global structural transitions were determined by projecting the coordinates into the previously defined cylindrical coordinate frame (see Figure S2) (33). Local variations, where applicable, were estimated using the previously described EVA-MS procedure. In all cases, first 40 ns of the simulations were omitted from the analysis. The conformation of ATP was assessed using the GROMACS *angle* utility. The distributions of ATP dihedral angles were assessed by fitting single (or multiple gaussians) through the raw distributions using the program Grace. Radius of gyration (R_g) and root mean squared deviation (RMSD) calculations were carried out using the GROMACS *gyrate* and *rms* utilities, respectively, utilizing the $\text{C}\alpha$ atoms. Solvent accessible surface area for ATP was calculated using the GROMACS *sasa* tool.

Solvation analysis using grid inhomogeneous solvation theory

Representative structures from the HS of $Wz_{CDAC} \cdot ATP \cdot Mg^{2+}$ (33) and C1 cluster of the $Wz_{CDAC,Y569F} \cdot ATP \cdot Mg^{2+}$ complexes were randomly selected and in each case, the first solvation shell of the Mg^{2+} ion was retained. The systems were set-up, energy minimized, and equilibrated in the NVT and NPT ensembles using the same protocols as above. Production runs were carried out with position restraints imposed on all backbone atoms using a force constant of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ acting along each of the x, y, and z directions. Simulation times of 20 ns were used in all cases, and coordinates were saved every 1 ps. Grid inhomogeneous solvation theory (GIST) (46) and hydration site analysis (HSA) calculations were performed on the last 10 ns of each simulation using the SSTMap suite (50). Earlier studies have demonstrated the convergence of water distribution and energy terms (to within $0.04 \text{ kcal mol}^{-1}$) within 10 ns. All simulations were performed in triplicate where the initial velocities were assigned using a random seed. For the GIST calculations the grid box center in each case was defined using two pseudo atoms placed next to the triphosphate of ATP (see Figure 8) with grid dimensions of $40 \times 40 \times 40 \text{ \AA}$. HSA calculations were performed with the same two pseudo atoms described above as a reference for the definition of the hydration site.

Expression and purification of the catalytic domain of Wzc and specific mutants

In all cases protein expression, purification, and preparation of nucleotide-free samples was carried out as previously described (33). In the case of the $Wz_{CDAC,K540M}$ the dialyzed for several days in buffer containing 5 mM $MgCl_2$ to remove bound ADP. In all other cases, the His₆-tagged were bound to a HP HiTrap Ni-NTA column and washed repeatedly as described before (33).

Isothermal titration calorimetry measurements

Purified WZC_{CDAC},K_{540M} was dialyzed against dialysis buffer containing 25 mM piperazine-N, N'-bis(2-ethanesulfonic acid) (PIPES) pH 6.5, 300 mM NaCl, 5 mM β -mercaptoethanol (BME) and 10% glycerol (ITC buffer) either in the presence of 5 or 10 mM MgCl₂. For the titrations, the samples were concentrated to 40 μ M via spin columns. In each case a titrant stock at 100 mM was prepared by dissolving ADP (sodium salt, Sigma-Aldrich) in the dialysis buffer, following which the pH was adjusted to 6.5 using sodium hydroxide (NaOH). Final titrant samples were prepared by further diluting the 100 mM stock with dialysis buffer. All measurements were performed in triplicate. The protein/titrant concentration were 40 μ M/400 μ M in the presence of 5 or 10 mM MgCl₂. ITC measurements were performed using a MicroCal iTC200 isothermal titration calorimeter (Malvern). All titrations were performed with a rotation speed of 500 rpm. A total of 16 injections were made in each case where the first one consisted of 0.4 μ L with a duration 0.8 s, and the rest consisted of 2.4 μ L each with a duration of 4.8 s and filter periods of 5 s. The spacing between injections used was 240 s. In all cases the same experiment was performed by titrating the titrant into buffer alone and used to normalize the data. The normalized data were fitted to the one-site binding model using Origin (OriginLab).

Enzymatic assays

Purified WZC_{CDAC} or WZC_{CDAC},K_{540M} (without the removal of the bound ADP) was exchanged into reaction buffer containing 20 mM Tris (pH 7.5), 15% glycerol, 300 mM NaCl, 5 mM BME, and 50 mM MgCl₂, using spin columns to a final concentration of 3 μ M. All assays were performed in triplicate. The reaction was initiated by the addition of 200 μ M ATP at 37 °C, following which, 80 μ L aliquots were mixed with the malachite green staining solution (Sigma Aldrich) at time intervals of 10, 20, 30, 40, and 50 min. The samples were incubated for 30 min

at room temperature and the absorbance at 620 nm was measured in each case. Absorbance measurements were carried out using a 96-well plate and a SpectraMax M2 plate reader. Additionally, a set of reference experiments were carried out in the absence of protein to assess background ATP hydrolysis. Inorganic phosphate concentrations were determined using the phosphate standard provided in the kit (Sigma Aldrich) diluted into the reaction buffer.

Analysis of the oligomerization states of Wzc_{CD,K540M}

Purified Wzc_{CD,K540M} samples were concentrated to 80 μ M in the ITC buffer (see above) and divided into 5 aliquots of 500 μ L each and subjected to: (1) No treatment, (2) **incubation with** 1.6 mM MgCl₂, (3) **incubation with** 0.8 mM ADP, (4) **incubation with** 1.6 mM MgCl₂ and 0.8 mM ADP or (5) **incubation with** 1.6 mM MgCl₂ and 0.8 mM ATP. All samples were incubated for ~20 h at 4 °C prior to injection into a pre-calibrated Superdex 200 10/300 (GE Healthcare Biosciences) gel filtration column.

Funding

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

Acknowledgements

The authors thank Drs. Tom Kurtzman, Andrea Piserchio and Sébastien Alphonse for critical comments and suggestions.

Author contributions

R. G. conceived the project; F. H. performed and analyzed all simulations and experiments. F. H. prepared a first draft of the paper and figures that were refined by R. G. with input from F. H.

Competing interests

The authors declare that they have no competing interests.

Data and materials availability

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

References

1. A. J. Cozzzone, C. Grangeasse, P. Doublet, B. Duclos, Protein phosphorylation on tyrosine in bacteria. *Arch. Microbiol.* **181**, 171-181 (2004).
2. E. Bechet, S. Guiral, S. Torres, I. Mijakovic, A. J. Cozzzone, C. Grangeasse, Tyrosine-kinases in bacteria: From a matter of controversy to the status of key regulatory enzymes. *Amino Acids* **37**, 499-507 (2009).
3. C. Grangeasse, R. Terreux, S. Nessler, Bacterial tyrosine-kinases: Structure-function analysis and therapeutic potential. *Biochim. Biophys. Acta* **1804**, 628-634 (2010).
4. C. Grangeasse, S. Nessler, I. Mijakovic, Bacterial tyrosine kinases: Evolution, biological function and structural insights. *Philos. T. R. So Lond. B* **367**, 2640-2655 (2012).
5. F. Jadeau, C. Grangeasse, L. Shi, I. Mijakovic, G. Deleage, C. Combet, Bykdb: The bacterial protein tyrosine kinase database. *Nucleic Acids Res.* **40**, D321-324 (2012).
6. J. D. Chao, D. Wong, Y. Av-Gay, Microbial protein-tyrosine kinases. *J. Biol. Chem.* **289**, 9463-9472 (2014).
7. B. Obadia, S. Lacour, P. Doublet, H. Baubichon-Cortay, A. J. Cozzzone, C. Grangeasse, Influence of tyrosine-kinase Wzc activity on colanic acid production in *Escherichia coli* K12 cells. *J. Mol. Biol.* **367**, 42-53 (2007).
8. A. Paiment, J. Hocking, C. Whitfield, Impact of phosphorylation of specific residues in the tyrosine autokinase, Wzc, on its activity in assembly of group 1 capsules in *Escherichia coli*. *J. Bacteriol.* **184**, 6437-6447 (2002).
9. M. Kolot, R. Gorovits, N. Silberstein, B. Fichtman, E. Yagil, Phosphorylation of the integrase protein of coliphage HK022. *Virology* **375**, 383-390 (2008).

10. G. Klein, C. Dartigalongue, S. Raina, Phosphorylation-mediated regulation of heat shock response in *Escherichia coli*. *Mol. Microbiol.* **48**, 269-285 (2003).
11. I. Mijakovic, D. Petranovic, B. Macek, T. Cepo, M. Mann, J. Davies, P. R. Jensen, D. Vujaklija, Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res.* **34**, 1588-1596 (2006).
12. C. M. Ernst, J. R. Braxton, C. A. Rodriguez-Osorio, A. P. Zagieboylo, L. Li, A. Pironti, A. L. Manson, A. V. Nair, M. Benson, K. Cummins, A. E. Clatworthy, A. M. Earl, L. A. Cosimi, D. T. Hung, Adaptive evolution of virulence and persistence in carbapenem-resistant *klebsiella pneumoniae*. *Nat. Med.* **26**, 705-711 (2020).
13. V. Olivares-Illana, P. Meyer, E. Bechet, V. Gueguen-Chaignon, D. Soulat, S. Lazereg-Riquier, I. Mijakovic, J. Deutscher, A. J. Cozzzone, O. Laprevote, S. Morera, C. Grangeasse, S. Nessler, Structural basis for the regulation mechanism of the tyrosine kinase CapB from *Staphylococcus aureus*. *PLoS Biol* **6**, e143 (2008).
14. E. Bechet, J. Gruszczyk, R. Terreux, V. Gueguen-Chaignon, A. Vigouroux, B. Obadia, A. J. Cozzzone, S. Nessler, C. Grangeasse, Identification of structural and molecular determinants of the tyrosine-kinase Wzc and implications in capsular polysaccharide export. *Mol. Microbiol.* **77**, 1315-1325 (2010).
15. D. C. Lee, J. Zheng, Y. M. She, Z. Jia, Structure of *escherichia coli* tyrosine kinase Etk reveals a novel activation mechanism. *EMBO J.* **27**, 1758-1766 (2008).
16. L. Shi, B. Ji, L. Kolar-Znika, A. Boskovic, F. Jadeau, C. Combet, C. Grangeasse, D. Franjevic, E. Talla, I. Mijakovic, Evolution of bacterial protein-tyrosine kinases and their relaxed specificity toward substrates. *Genome Biol. Evol.* **6**, 800-817 (2014).

17. L. M. Longo, J. Jablonska, P. Vyas, M. Kanade, R. Kolodny, N. Ben-Tal, D. S. Tawfik, On the emergence of P-loop ntpase and Rossmann enzymes from a beta-alpha-beta ancestral fragment. *eLife* **9**, e64415 (2020).
18. S. K. Hanks, T. Hunter, Protein kinases 6. The eukaryotic protein kinase superfamily: Kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576-596 (1995).
19. D. D. Leipe, E. V. Koonin, L. Aravind, Evolution and classification of P-loop kinases and related proteins. *J. Mol. Biol.* **333**, 781-815 (2003).
20. D. D. Leipe, Y. I. Wolf, E. V. Koonin, L. Aravind, Classification and evolution of P-loop gtpases and related atpases. *J. Mol. Biol.* **317**, 41-72 (2002).
21. P. Doublet, C. Grangeasse, B. Obadia, E. Vaganay, A. J. Cozzzone, Structural organization of the protein-tyrosine autokinase Wzc within *Escherichia coli* cells. *J. Biol. Chem.* **277**, 37339-37348 (2002).
22. B. Duclos, C. Grangeasse, E. Vaganay, M. Riberty, A. J. Cozzzone, Autophosphorylation of a bacterial protein at tyrosine. *J. Mol. Biol.* **259**, 891-895 (1996).
23. C. Grangeasse, P. Doublet, A. J. Cozzzone, Tyrosine phosphorylation of protein kinase Wzc from *Escherichia coli* K12 occurs through a two-step process. *J. Biol. Chem.* **277**, 7127-7135 (2002).
24. C. Grangeasse, B. Obadia, I. Mijakovic, J. Deutscher, A. J. Cozzzone, P. Doublet, Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J. Biol. Chem.* **278**, 39323-39329 (2003).
25. J. Nourikyan, M. Kjos, C. Mercy, C. Cluzel, C. Morlot, M. F. Noirot-Gros, S. Guiral, J. P. Lavergne, J. W. Veening, C. Grangeasse, Autophosphorylation of the bacterial tyrosine-

- kinase cpsd connects capsule synthesis with the cell cycle in *Streptococcus pneumoniae*. *PLoS Genet* **11**, e1005518 (2015).
26. T. Wugeditsch, A. Paiment, J. Hocking, J. Drummelsmith, C. Forrester, C. Whitfield, Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in escherichia coli. *J. Biol. Chem.* **276**, 2361-2371 (2001).
 27. D. B. Temel, K. Dutta, S. Alphonse, J. Nourikyan, C. Grangeasse, R. Ghose, Regulatory interactions between a bacterial tyrosine kinase and its cognate phosphatase. *J. Biol. Chem.* **288**, 15212-15228 (2013).
 28. C. Grangeasse, A. J. Cozzzone, J. Deutscher, I. Mijakovic, Tyrosine phosphorylation: An emerging regulatory device of bacterial physiology. *Trends Biochem. Sci.* **32**, 86-94 (2007).
 29. A. J. Standish, R. Morona, The role of bacterial protein tyrosine phosphatases in the regulation of the biosynthesis of secreted polysaccharides. *Antioxid. Redox Signal.* **20**, 2274-2289 (2014).
 30. T. J. Boggon, M. J. Eck, Structure and regulation of Src family kinases. *Oncogene* **23**, 7918-7927 (2004).
 31. B. J. Canagarajah, A. Khokhlatchev, M. H. Cobb, E. J. Goldsmith, Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell* **90**, 859-869 (1997).
 32. D. Soulat, J. M. Jault, C. Geourjon, P. Gouet, A. J. Cozzzone, C. Grangeasse, Tyrosine-kinase Wzc from *Escherichia coli* possesses an ATPase activity regulated by autophosphorylation. *FEMS Microbiol. Lett.* **274**, 252-259 (2007).
 33. F. Hajredini, A. Piserchio, R. Ghose, Long-range dynamic correlations regulate the catalytic activity of the bacterial tyrosine kinase Wzc. *Sci. Adv.* **6**, (2020).

34. L. Wang, R. A. Friesner, B. J. Berne, Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2). *J. Phys. Chem. B* **115**, 9431-9438 (2011).
35. T. Krell, J. Maclean, D. J. Boam, A. Cooper, M. Resmini, K. Brocklehurst, S. M. Kelly, N. C. Price, A. J. Laphorn, J. R. Coggins, Biochemical and X-ray crystallographic studies on shikimate kinase: The important structural role of the P-loop lysine. *Protein Sci.* **10**, 1137-1149 (2001).
36. A. Matte, L. W. Tari, L. T. Delbaere, How do kinases transfer phosphoryl groups? *Structure* **6**, 413-419 (1998).
37. D. Mann, C. Teuber, S. A. Tennigkeit, G. Schroter, K. Gerwert, C. Kötting, Mechanism of the intrinsic arginine finger in heterotrimeric G proteins. *Proc. Natl. Acad. Sci. USA* **113**, E8041-E8050 (2016).
38. T. Rudack, F. Xia, J. Schlitter, C. Kötting, K. Gerwert, Ras and GTPase-activating protein (GAP) drive GTP into a precatalytic state as revealed by combining FTIR and biomolecular simulations. *Proc. Natl. Acad. Sci. USA* **109**, 15295-15300 (2012).
39. D. N. Shalaeva, D. A. Cherepanov, M. Y. Galperin, A. V. Golovin, A. Y. Mulkidjanian, Evolution of cation binding in the active sites of P-loop nucleoside triphosphatases in relation to the basic catalytic mechanism. *eLife* **7**, e37373 (2018).
40. W. Wu, K. T. Park, T. Holyoak, J. Lutkenhaus, Determination of the structure of the MinD-ATP complex reveals the orientation of mind on the membrane and the relative location of the binding sites for MinE and MinC. *Mol. Microbiol.* **79**, 1515-1528 (2011).
41. J. Lutkenhaus, M. Sundaramoorthy, MinD and role of the deviant Walker A motif, dimerization and membrane binding in oscillation. *Mol. Microbiol.* **48**, 295-303 (2003).

42. H. Takahashi, S. Umino, Y. Miki, R. Ishizuka, S. Maeda, A. Morita, M. Suzuki, N. Matubayasi, Drastic compensation of electronic and solvation effects on ATP hydrolysis revealed through large-scale QM/MM simulations combined with a theory of solutions. *J. Phys. Chem. B* **121**, 2279-2287 (2017).
43. C. Kotting, A. Kallenbach, Y. Suveyzdis, A. Wittinghofer, K. Gerwert, The GAP arginine finger movement into the catalytic site of Ras increases the activation entropy. *Proc. Natl. Acad. Sci. USA* **105**, 6260-6265 (2008).
44. D. Elnatan, M. Betegon, Y. Liu, T. Ramelot, M. A. Kennedy, D. A. Agard, Symmetry broken and rebroken during the ATP hydrolysis cycle of the mitochondrial Hsp90 TRAP1. *eLife* **6**, e25235 (2017).
45. R. Dutta, M. Inouye, GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* **25**, 24-28 (2000).
46. C. N. Nguyen, T. K. Young, M. K. Gilson, Grid inhomogeneous solvation theory: Hydration structure and thermodynamics of the miniature receptor cucurbit[7]uril. *J. Chem.. Phys.* **137**, 044101 (2012).
47. C. N. Nguyen, T. Kurtzman, M. K. Gilson, Spatial decomposition of translational water-water correlation entropy in binding pockets. *J. Chem. Theory Comput.* **12**, 414-429 (2016).
48. Q. H. Tran, G. Unden, Changes in the proton potential and the cellular energetics of *Escherichia coli* during growth by aerobic and anaerobic respiration or by fermentation. *Eur J Biochem* **251**, 538-543 (1998).
49. A. Stein, T. Kortemme, Improvements to robotics-inspired conformational sampling in rosetta. *PLoS One* **8**, e63090 (2013).

50. K. Haider, A. Cruz, S. Ramsey, M. K. Gilson, T. Kurtzman, Solvation structure and thermodynamic mapping (SSTMap): An open-source, flexible package for the analysis of water in molecular dynamics trajectories. *J. Chem. Theory Comput.* **14**, 418-425 (2018).
51. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. Berendsen, GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701-1718 (2005).
52. C. Kutzner, S. Pall, M. Fechner, A. Esztermann, B. L. de Groot, H. Grubmuller, More bang for your buck: Improved use of GPU nodes for GROMACS 2018. *J. Comput. Chem.* **40**, 2418-2431 (2019).
53. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmuller, A. D. MacKerell, Jr., CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Meth.* **14**, 71-73 (2017).
54. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690 (1984).

FIGURE LEGENDS

Figure 1. Conformational landscape of unliganded $Wz_{CDAC,K540M}$. (A) Probability density of conformational states sampled in the REST2 simulations projected onto $\theta-|h|$ space and plotted using kernel density estimation. Conformations sampled by the K540M mutant are shown as a density plot; results of the Wz_{CDAC} (WT) simulations are represented by the green contours. Specific regions corresponding to states that are more closed (C1) or more open (C3) compared to the WT OS, are indicated. The red dot indicates the crystallographic monomers. (B) Conformations of key active site residues in the WT OS (red) and in a structure representative of the C3 region of the K540M mutant (yellow) are compared on the left panel. The WT OS is stabilized by a Walker-A K540/Walker-B D642 salt-bridge; for the K540M mutant, absent the salt-bridge, M540 is buried within the protein core participating in hydrophobic interactions with V665 and M531. Shown on the right panel, is a similar interaction in K15M mutant of *E. chrysanthemi* shikimate kinase (PDB: 1E6C). (C) Major conformations seen in the K540M simulations identified through EVA-MS clustering (see Figure S4). Structures that populate the C1 region display a bent $\alpha 2$ (distortion of I_1 , the first interaction interface) characteristic of WT nucleotide• Mg^{2+} -bound states; structures from the C3 region show no remarkable features or distortions. Helices $\alpha 2$ (I_1), $\alpha 7/\alpha 9$ (I_2), $\alpha 3$ and $\alpha 4$ are colored blue, red, brown, and yellow, respectively. The disordered region of the RK-cluster is colored cyan.

Figure 2. Conformational landscapes of the nucleotide• Mg^{2+} complexes of $Wz_{CDAC,K540M}$. Conformations sampled in the REST2 simulations of the $Wz_{CDAC,K540M}$ •ATP• Mg^{2+} (left) and $Wz_{CDAC,K540M}$ •ADP• Mg^{2+} (right) complexes are mapped onto $\theta-|h|$ space and represented as density plots using kernel density estimation. The conformations sampled by the corresponding complexes of Wz_{CDAC} (WT) are shown as green contours. The locations of the major

conformations seen in the WT simulations are shown, and the red dot indicates reference values for monomers in the crystal structure.

Figure 3. Oligomerization propensities of Wzc_{CD,K540M}. The effects of incubating Wzc_{CD,K540M} (untreated shown as the pink trace; top panel) with Mg²⁺ alone (20-fold, purple trace; top panel), ADP alone (10-fold, blue trace; middle panel), ADP•Mg²⁺ (10-fold•20-fold, red trace; middle panel) or ATP•Mg²⁺ (10-fold•20-fold, green trace; middle panel) on oligomer formation through gel filtration chromatography (Superdex 200 10/300). The bottom panel shows the calibration traces (blue dashed lines) using molecular weight standards (molecular weights, in KDa, are indicated).

Figure 4. Definition of the dihedral angles representing the conformation of the phosphate backbone of ATP. The α -, β -, and γ -phosphates of ATP are colored blue, yellow, and red, respectively. The top panel shows the sets of four atoms used to define the (A) $\phi_{\beta\gamma}$, (B) $\phi_{\alpha\beta}$, and (C) $\phi_{\alpha\gamma}$ dihedral angles. The lower panels depict the corresponding Newman projections; the black spheres indicate the bridging oxygens. As defined, counterclockwise rotations of each angle towards 0° generates an eclipsed conformation, while clockwise rotations towards -60° brings them into a staggered conformation.

Figure 5. Comparison of ATP dihedral angle distributions in the ATP•Mg²⁺ complexes of Wzc_{CDAC} (WT, cyan) and Wzc_{CDAC,K540M} (purple). The entire ensemble of structures was utilized to generate the distributions in each case. The dashed lines indicate the actual distributions, and the solid lines represent gaussian fits; the corresponding mean values are indicated by the vertical black lines. The red and green vertical lines depict angles for the fully eclipsed (E, 0°) and the staggered (S, -60°) conformations, respectively. Individual distributions (in this and in all other cases) have been normalized within the ensemble selected. The right panels depict the Newman

projections corresponding to the most probable conformation with the numerical values of the dihedral angles indicated. The distributions for the **WT** simulations are bimodal with the left peak indicating values corresponding to the OS and the right peak corresponding to values for the closed states i.e., the CS and the HS. Note that the -123° maximum for the $\phi_{\alpha\beta}$ angle in the wild-type OS also represents a near eclipsed conformation for a pair of non-bridging oxygens.

Figure 6. Conformation of ATP in its complexes with the P-loop ATPase, MinD. **(A)** The crystal structure of dimeric MinD (PDB: 3Q9L) showing dimerization along the active site with the individual monomers colored blue and red (left panel). An expanded view of the dimerization interface showing the so-called deviant Walker-A lysine of one monomer inserting into the active site of a neighboring monomer between the α - and γ -phosphates of ATP (right panel). **(B)** Representative distributions of the dihedral angles of ATP for the ATP•Mg²⁺ complexes of monomeric MinD (pink), dimeric MinD (black) and monomeric MinD complexed with a Na⁺ ion in a position analogous to the deviant Walker-A lysine (cyan). Dashed lines indicate actual distributions and solid lines represent gaussian fits with the corresponding mean values indicated by the vertical black lines. Green and red vertical lines indicate the values for the staggered and eclipsed conformations of ATP, respectively. **(C)** The most probable **conformations** in the three cases shown in **(B)** are illustrated; the top panel showing stick representations of ATP and key coordinating elements, and the lower panels showing the Newman projections corresponding to the most probable $\phi_{\alpha\beta}$ (green), $\phi_{\beta\gamma}$ (orange), and $\phi_{\alpha\gamma}$ (purple) angles for each case.

Figure 7. REST2 simulations on the WzCDAC,Y569F•ATP•Mg²⁺ complex. **(A)** Projection of the conformations sampled in the **WzCDAC,Y569F simulations** onto $\theta-|h|$ space is shown as a density plot; the conformations sampled by the **corresponding WzCDAC (WT)** complex are depicted using green contours. Distinct regions sampled in both cases are labeled (also see Figures S9 and S10).

(B) Comparison of the distribution of the dihedral angles of ATP in 500 representative structures drawn from the WT HS (black), or the WT CS (blue), or the entire Y569F ensemble (purple). Dashed lines indicate raw distributions and solid lines indicate gaussian fits with the corresponding means indicated by black vertical lines.

Figure 8. Comparison of ATP solvation in the $Wz_{CD\Delta C} \cdot ATP \cdot Mg^{2+}$, $Wz_{CD\Delta C, K540M} \cdot ATP \cdot Mg^{2+}$, and $Wz_{CD\Delta C, Y569F} \cdot ATP \cdot Mg^{2+}$ REST2 simulations. (A) Solvent accessible surface areas of ATP for the three cases are shown. For the $Wz_{CD\Delta C}$ (WT) complex (black line), ATP is largely shielded in the closed states (CS/HS) while it is more solvent exposed in the OS. For the K540M mutant complex (red line), which exists exclusively in the CS, ATP is shielded from the solvent. For Y569F mutant complex (purple line), the solvent accessibility of ATP shows a bimodal distribution. (B, C) Solvation densities determined using Grid Inhomogeneous Solvation Theory (GIST) for the C1 cluster from the $Wz_{CD\Delta C, Y569F} \cdot ATP \cdot Mg^{2+}$ simulations as representative of an open active site with a disordered RK-cluster (B) or the HS from corresponding $Wz_{CD\Delta C} \cdot ATP \cdot Mg^{2+}$ simulations as representative of an assembled active site (C). The volumes shown in the right panels in each case, indicate the water oxygen density, $\rho(O)$, within the solvation box defined by two pseudo atoms centered around ATP (gray spheres in the corresponding left panels). The blue coloring indicates bulk water density while red indicates water oxygen density that is higher relative to bulk.

Figure 9. Structural features observed in the classical MD simulation on the $[Wz_{CD} \cdot ATP \cdot Mg^{2+}]_8$ complex. (A) ATP dihedral angle distributions. Dashed lines indicate the actual distributions and solid lines represent the corresponding gaussian fits colored according to individual chains (A-H) of the octamer. Black vertical lines indicate the most probable values; the maximum and minimum values of the means (over the 8 chains) are shown next to the vertical lines. (B) Distributions of

selected distances that define the active site of Wzc_{CD}. Distributions of the R490,C ζ -ATP,P γ (black), Y569,O η -ATP,P α (red) and K492,N ζ -E572,C δ (green) distances over the MD simulation for all 8 chains are shown. The Y569,O η -ATP,P α and K492,N ζ -E572,C δ seen in the crystal structure are 3.9 ± 0.05 Å and 7.6 ± 4.0 Å, respectively. Electron density corresponding to R490 was missing.

Figure 10. Mechanistic deviations of Wzc_{CD} from P-loop NTPases. **(A)** In P-loop NTPases, the Walker-A lysine and the Mg²⁺ ion lock the orientations of the β - and γ -phosphates (represented by the red and yellow lines, respectively) within the same plane. Insertion of the arginine finger (indicated by the blue circle) between the α - (represented by the blue lines) and γ -phosphates leads to a rotation of the α -phosphate relative to the locked $\beta\gamma$ -diphosphate moiety bringing all the non-bridging oxygens of the triphosphate moiety into an eclipsed conformation. Additionally, insertion of the arginine finger into the active site results in the displacement of several ATP-solvating waters. These events make favorable enthalpic and entropic contributions to overcome the activation barrier for product formation. **(B)** In Wzc_{CD}, the Walker-A lysine and Mg²⁺ ion also initially lock the β - and γ -phosphates, however the subsequent rotation of the triphosphate is not dependent on the placement of the arginine finger into the active site, rather it is formation of the CS/HS (that is also coupled to the conformation of the Walker-A lysine) that promotes formation of the eclipsed conformation of ATP. Simultaneous to the formation of the closed states, the appropriate placement of Y569 in contacting the α -phosphate of ATP, its interaction with K492, and of the latter with E572, together with the interaction of the arginine finger equivalent, R490 with ATP, promote desolvation of ATP.

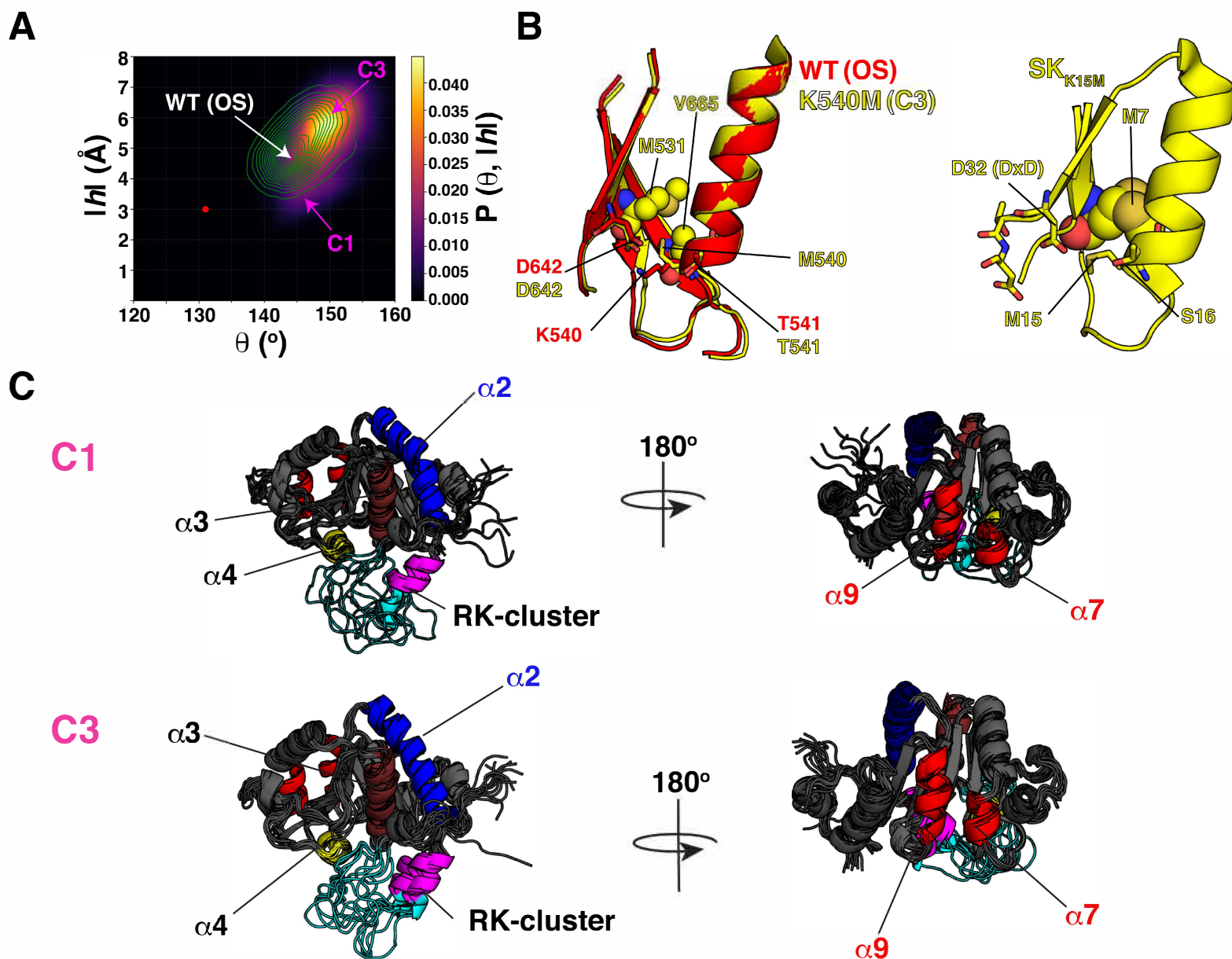


Figure 1

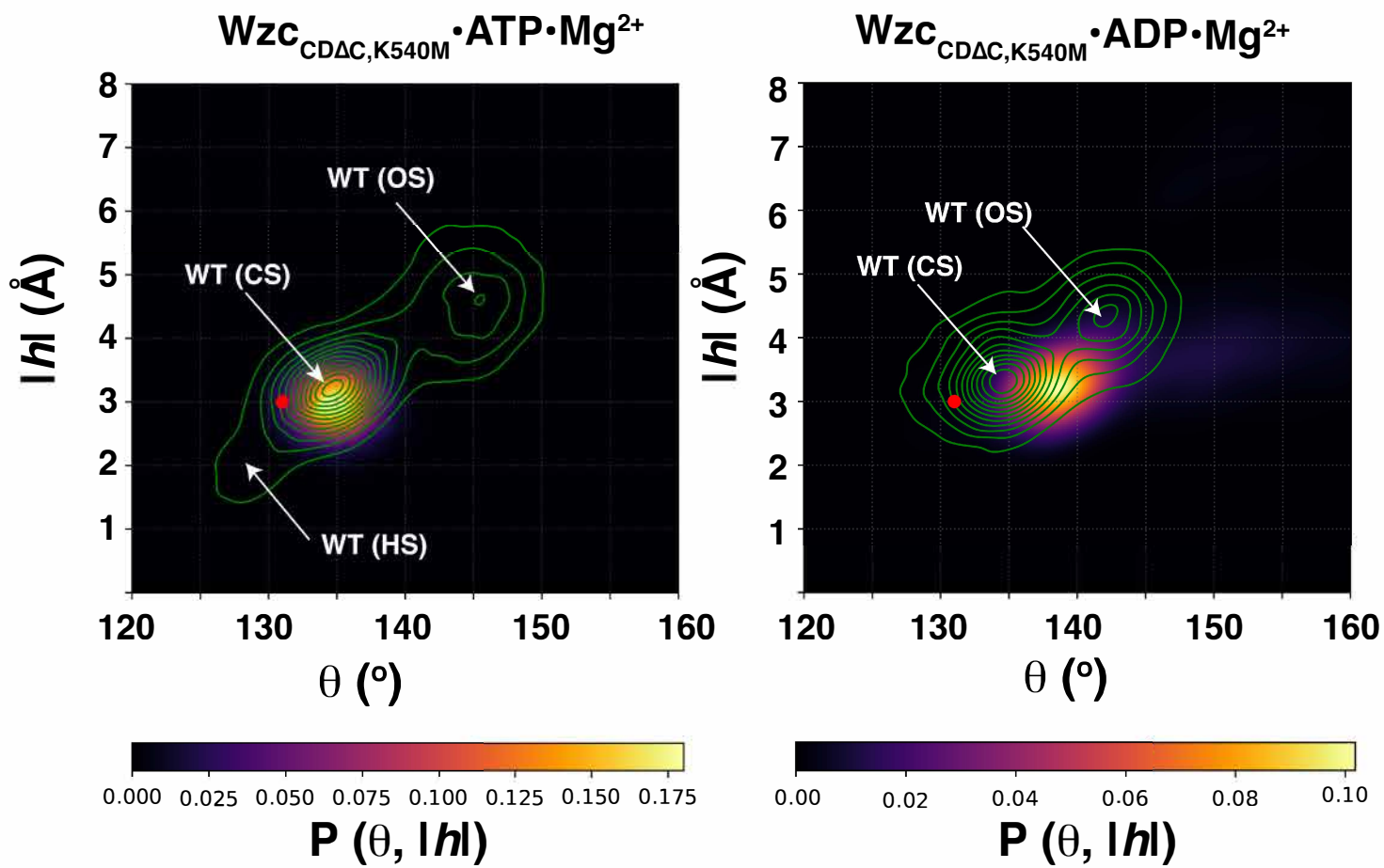


Figure 2

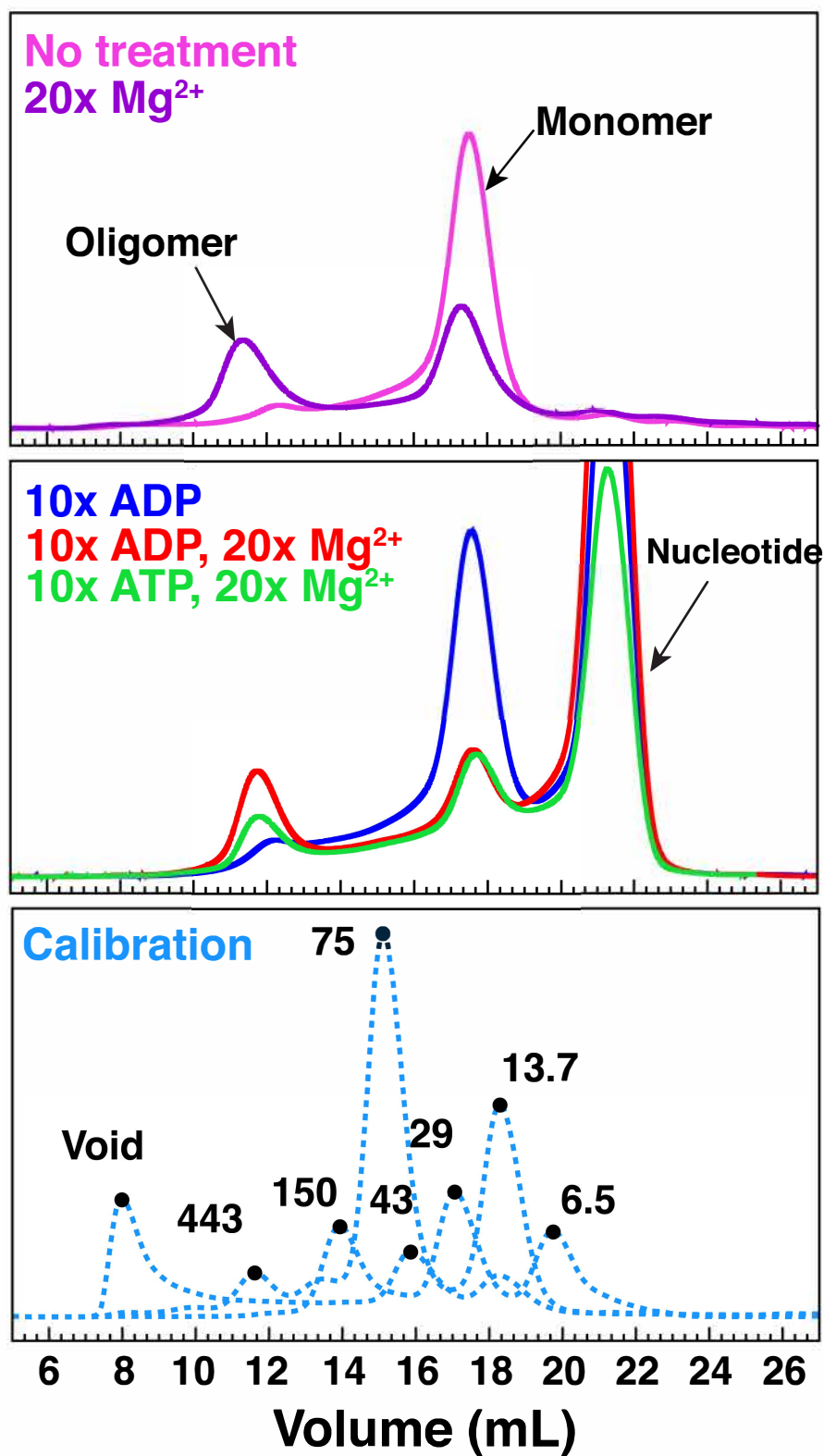


Figure 3

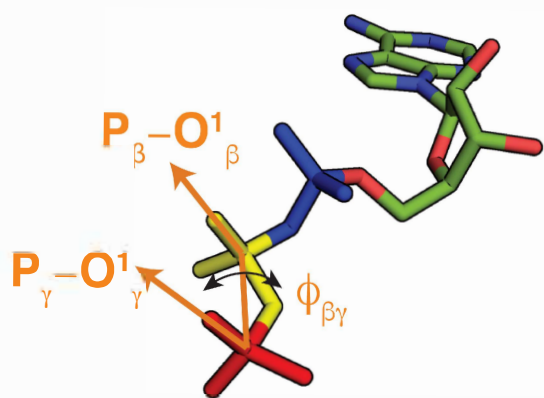
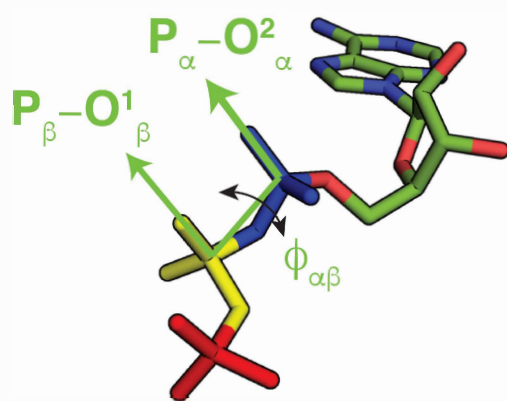
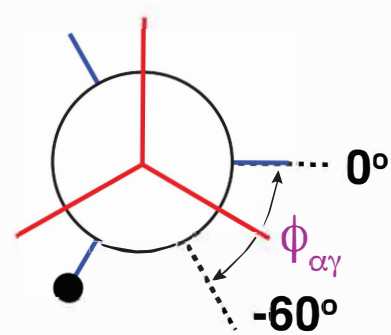
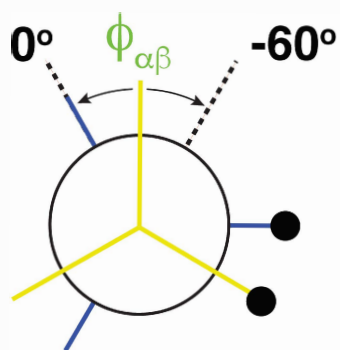
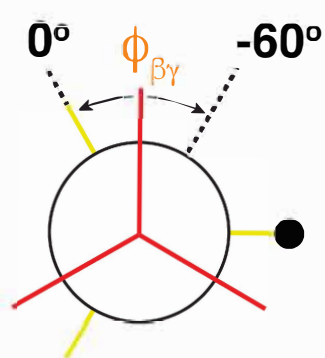
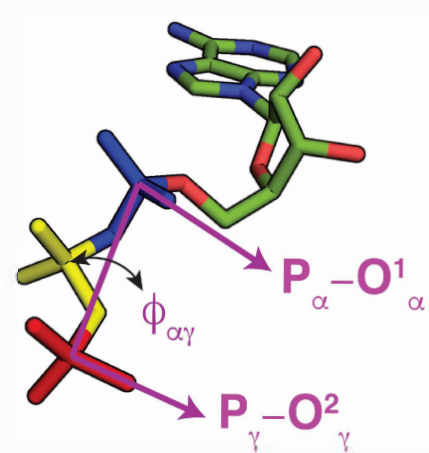
A**B****C**

Figure 4

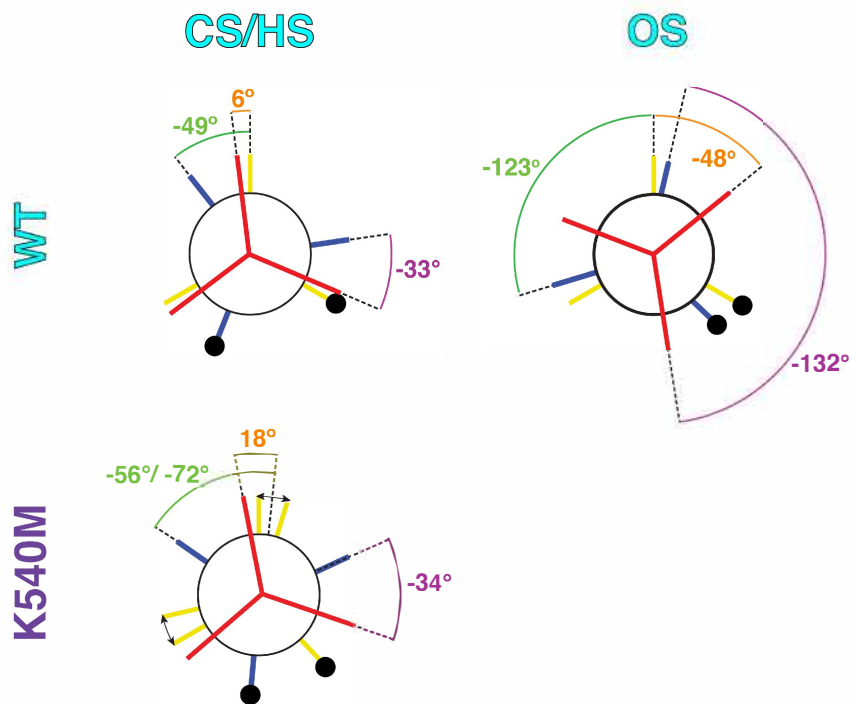
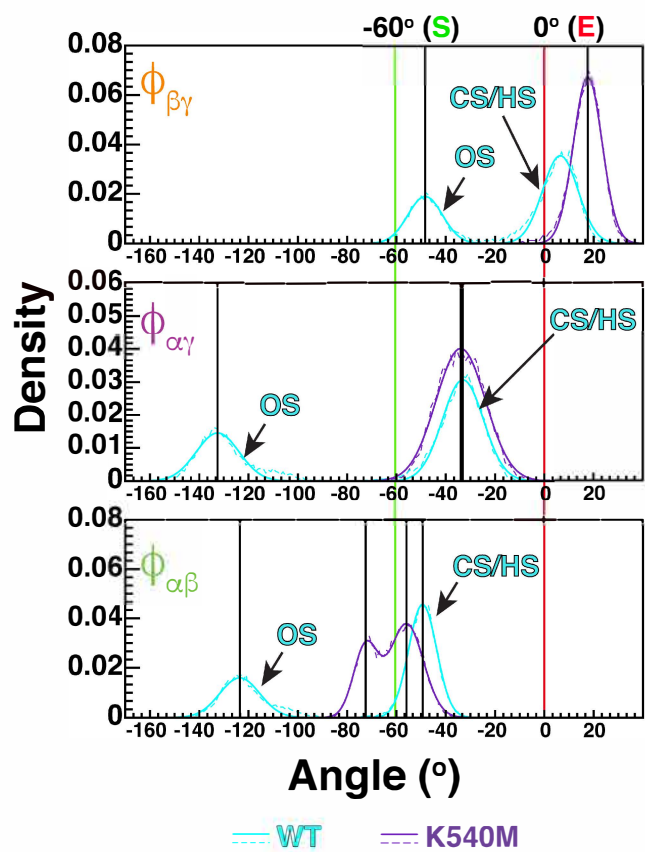


Figure 5

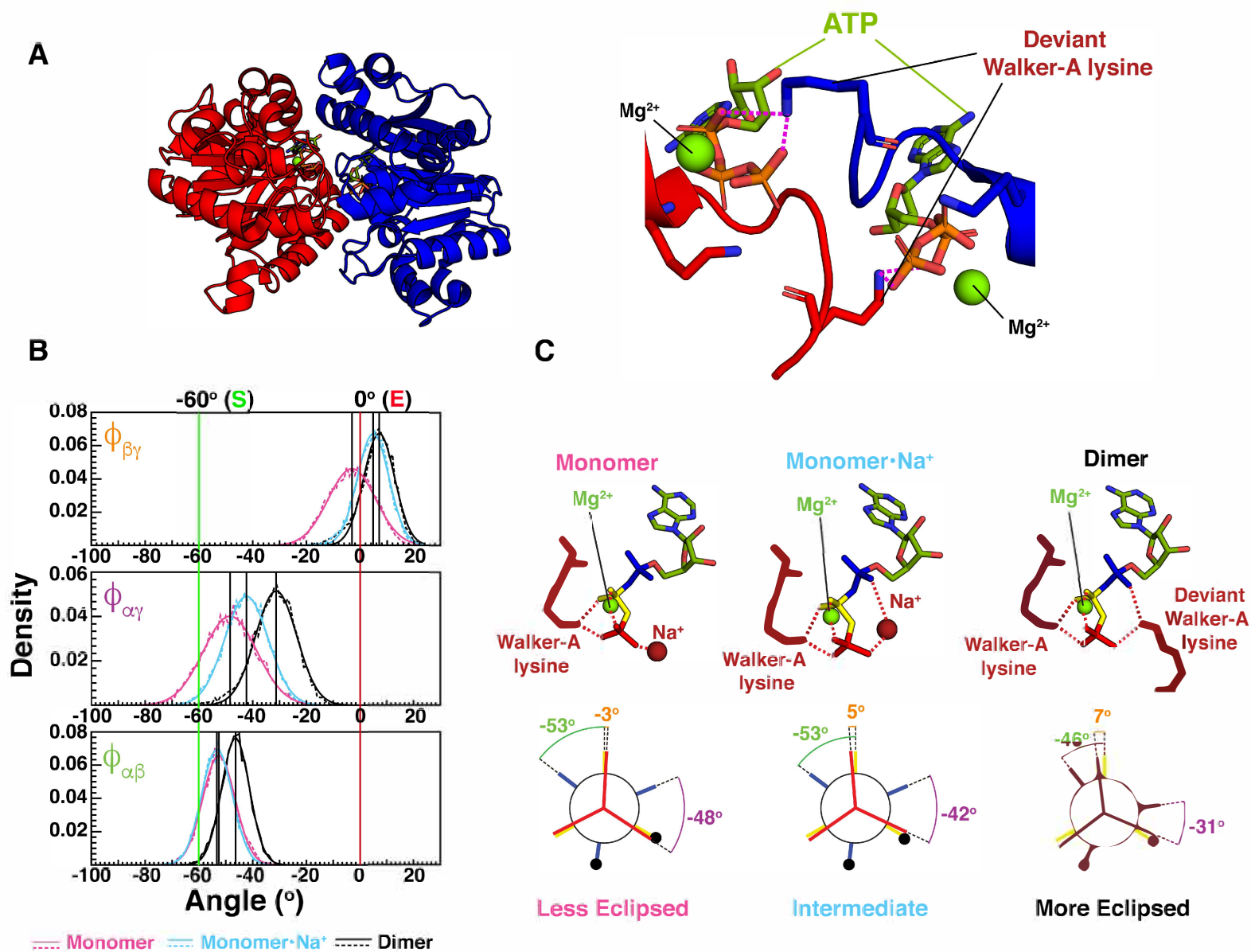


Figure 6

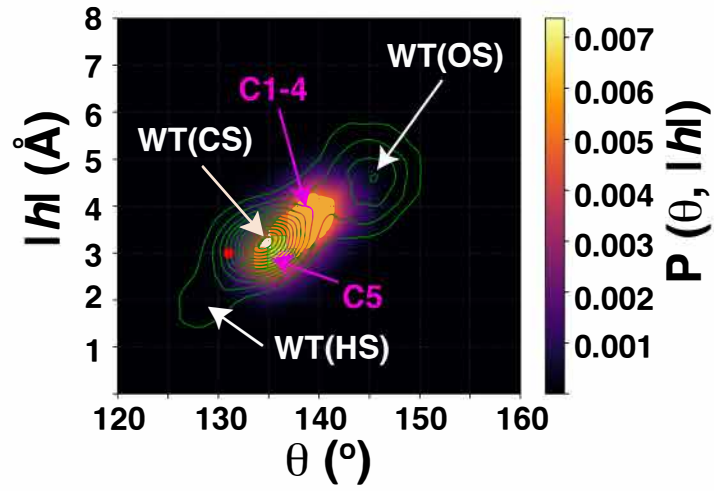
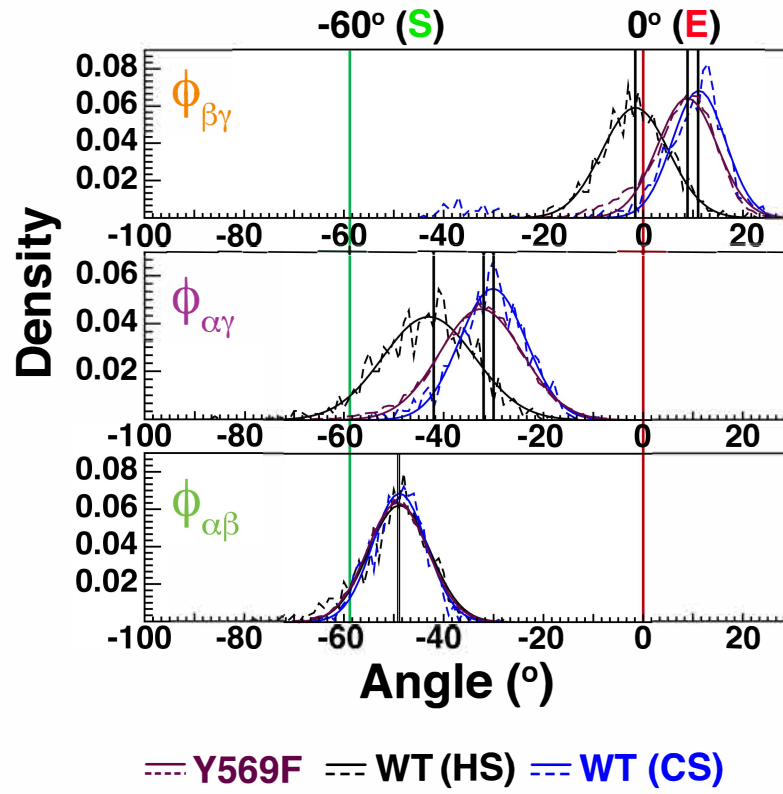
A**B**

Figure 7

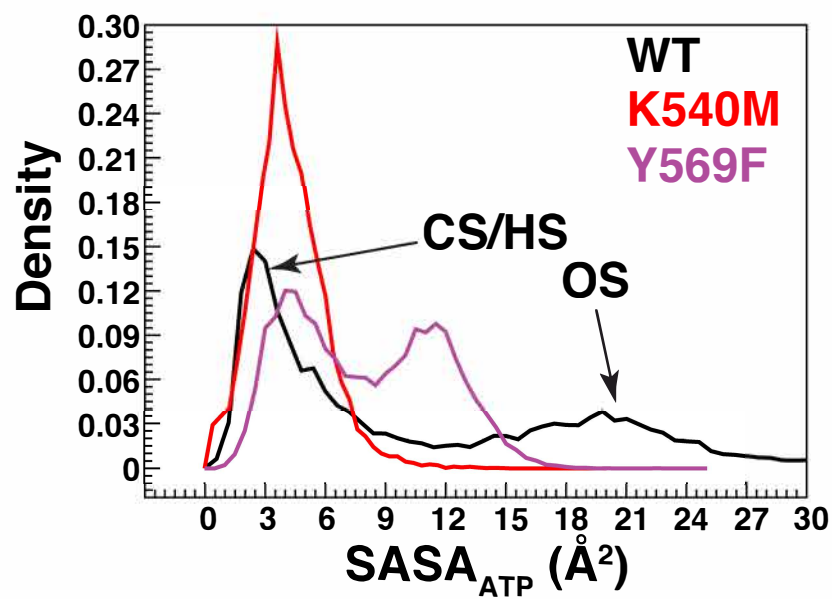
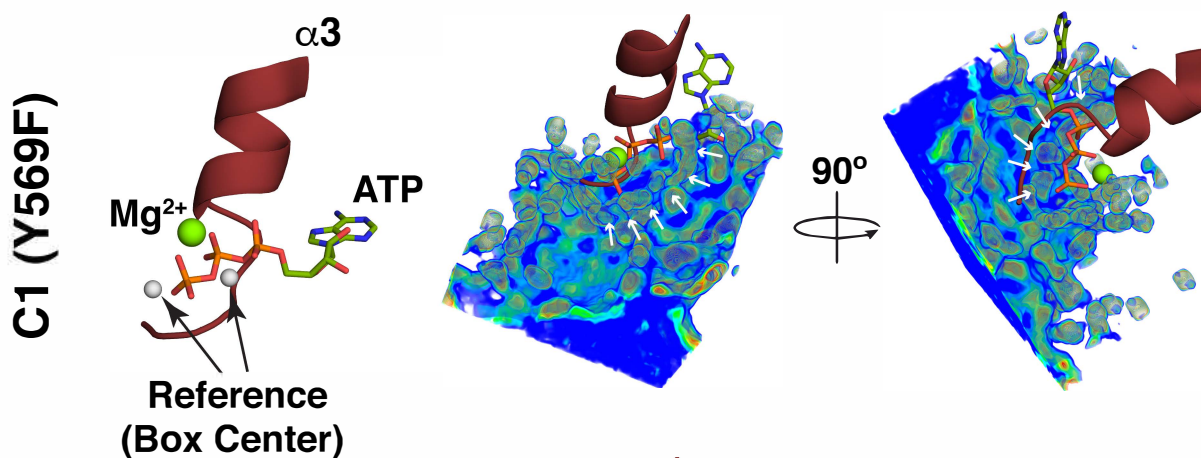
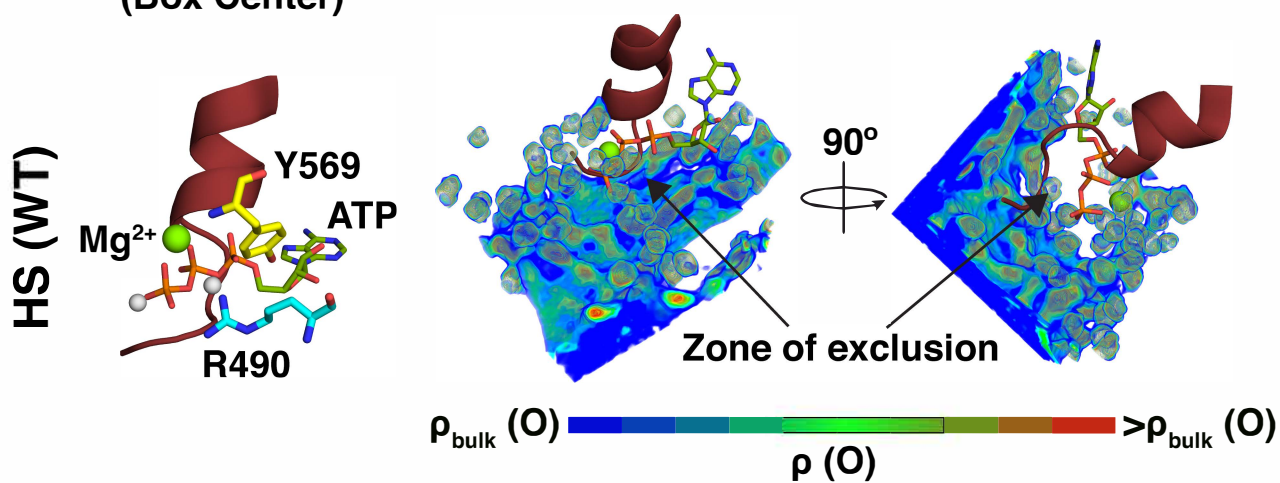
A**B****C**

Figure 8

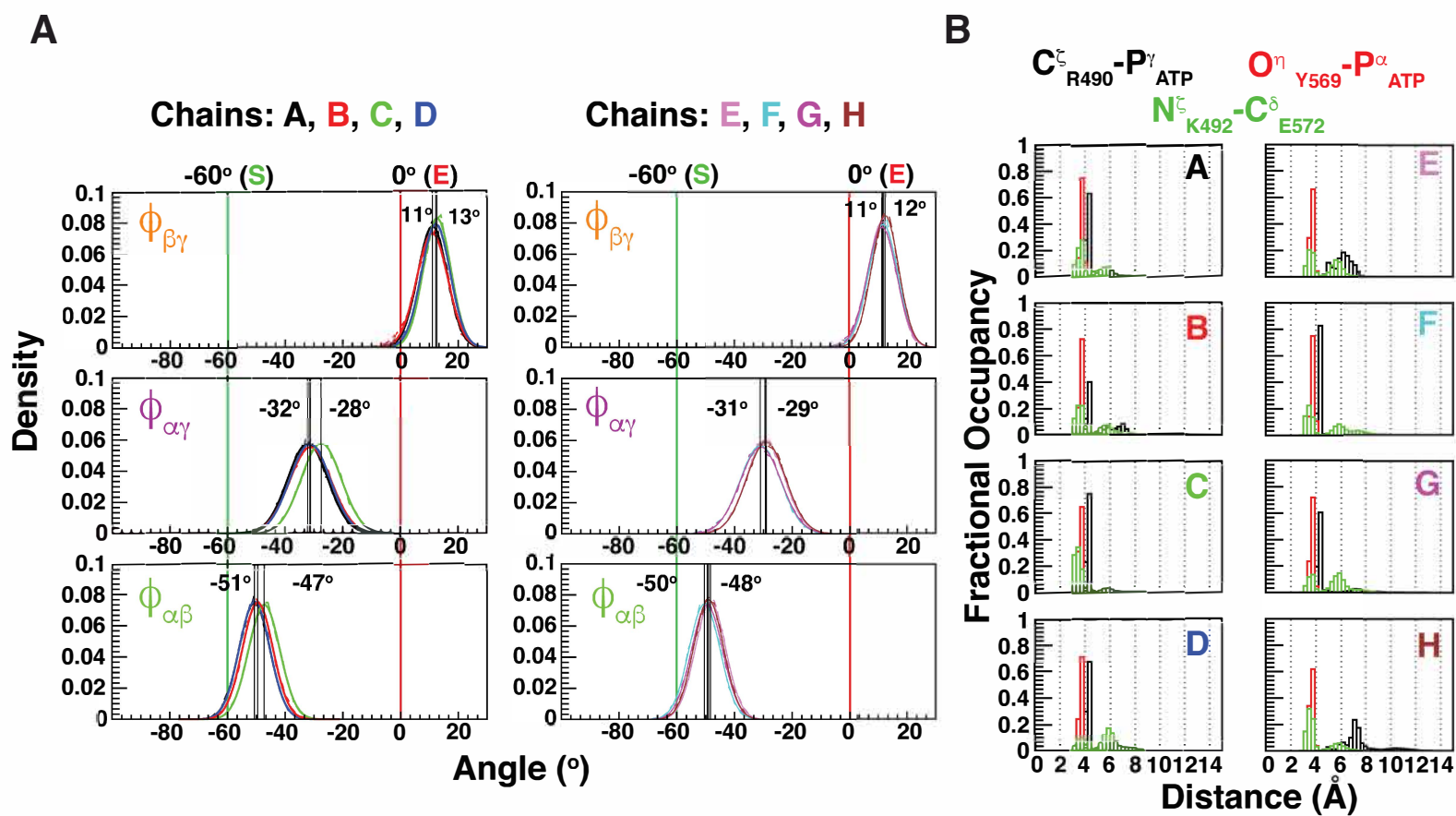
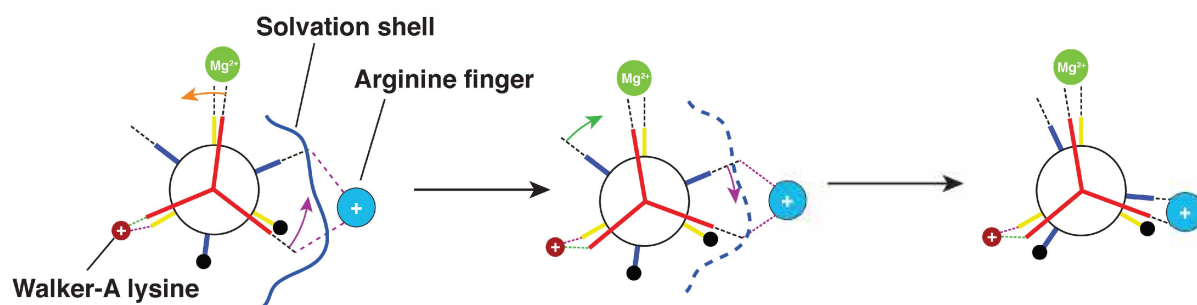
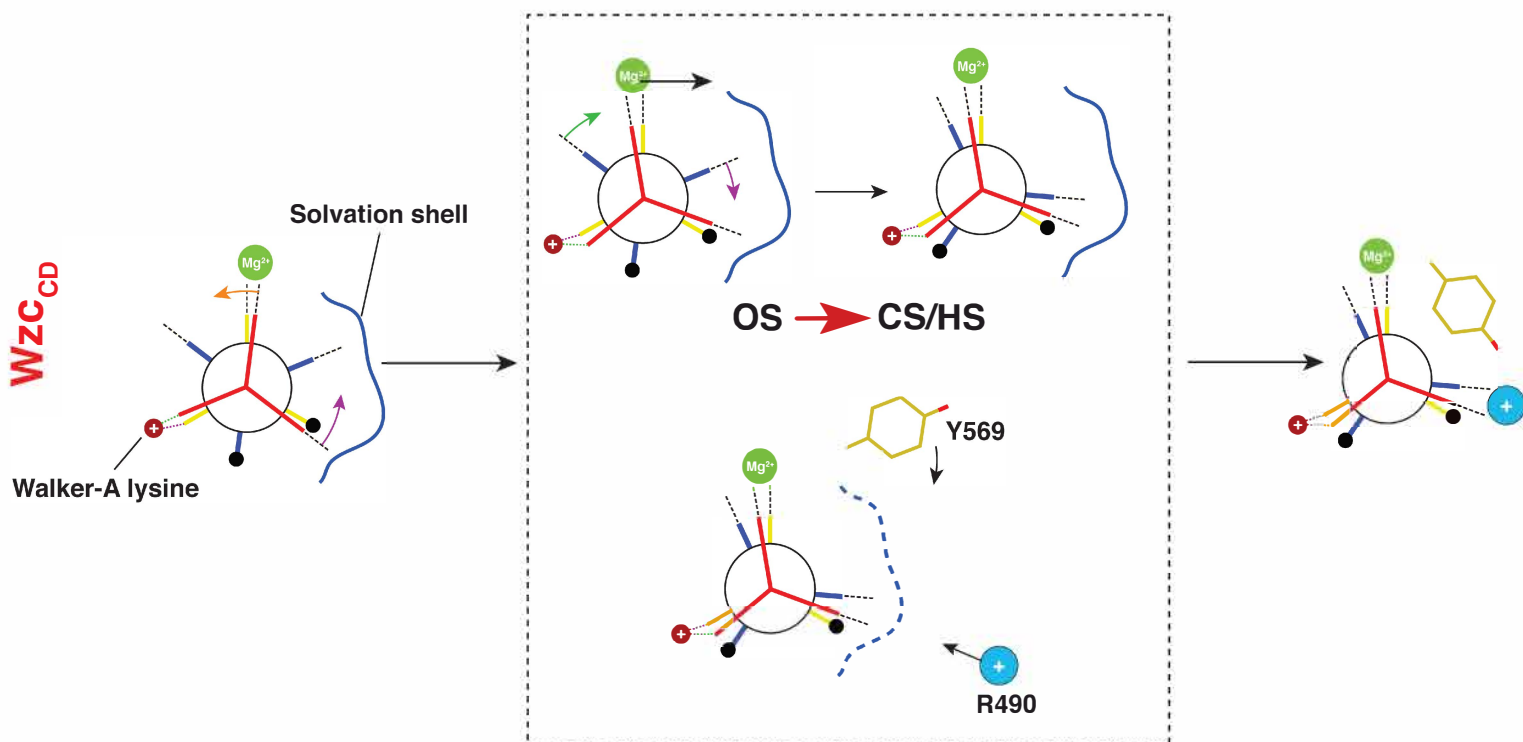


Figure 9

A**P-loop NTPase****B****Wzc_{CD}****Figure 10**

Supplementary Materials

for

An ATPase with a twist: a unique mechanism underlies the activity of the bacterial tyrosine kinase, Wzc.

SUPPLEMENTARY METHODS

Structure Preparation

The procedures followed to prepare the structures for the simulations are the same as those described in Materials and Methods section of Hajredini *et al.* (33). We reproduce this here, with some additions relevant to the present study, for the benefit of the reader.

A chain (Chain A, average RMSD over all chains based on C α atoms: 0.12 $\pm$ 0.04 Å) was selected at random from the crystal structure (PDB: 3LA6) of the catalytic domain of the K540M mutant of Wzc (Wzc_{CD}) (14). The missing segment of the RK-cluster was constructed using the Rosetta KIC protocol (49) and a single round of modeling without further energy minimization. In all cases, the bound Ca²⁺ ion present in the crystal structure (that contains bound ADP) was replaced by a Mg²⁺. M540 on the Walker-A motif was mutated back to the native lysine using the mutagenesis tool in PyMOL. Additionally, all residues spanning the dynamic Y-cluster at the C-terminal tail (705-720) were removed to generate the Wzc_{CDAC} kinase core construct that was in the simulations. A series of classical molecular dynamics (MD) simulations (see below) were performed using the structure generated above containing an intact RK-cluster with ADP•Mg²⁺ bound to the active-site to generate starting structures for all subsequent simulations. The structure was selected based on the presence of all ADP contacts that are observed in the crystal structure after equilibration of the conformation of the RK-cluster in the MD runs. In case of the Wzc_{CDAC}•ADP•Mg²⁺ simulations, this structure was used as the starting structure, without additional processing. For the Wzc_{CDAC}•ATP•Mg²⁺ simulations, the ATP•Mg²⁺ conformation was obtained from the crystal structure of the Wzc homolog MinD (PDB: 3Q9L) (40) and used to replace ADP•Mg²⁺ moiety. A short classical MD simulation was performed (see below for the setup) using Wzc_{CDAC}•ATP•Mg²⁺ at 400 K to achieve proper solvation of the Mg²⁺ ion. The appropriately solvated Wzc_{CDAC}•ATP•Mg²⁺ structure was then used as the starting structure for further simulations. For REST2 simulations involving the K540M or Y569F mutants, the starting Wzc_{CDAC}•ATP•Mg²⁺ structure was used to obtain the mutants that were generated using the PyMOL mutagenesis tool.

All simulations were performed using the GROMACS 2019 software suite (51, 52); the CHARMM36m (53) force-field was used to parameterize the system. The system was solvated using TIP3P waters in a cubic box with 1.5 nm edge distance from the solute. In all cases, charge

neutralizing ions were added to the system. Energy minimization with gradient descent was performed with a 0.1 fs time step until energy convergence was obtained. Prior to initiating the production runs the system was equilibrated, first in the NVT, and then in the NPT ensemble, for 100 ps each. NVT and NPT equilibration were conducted with position restraints on all heavy atoms using a 2 fs integration step. The production run was performed in the NPT ensemble using the Berendsen (54) and the Parrinello-Rahman (55) thermostat and barostat, respectively. Classical MD simulations were carried out at 300 K with a 2 fs integration step. The particle-mesh Ewald (PME) summation procedure with a cutoff of 1.0 nm was used to compute long-range electrostatic interactions and periodic boundary conditions were used to handle edge effects. In all cases, coordinates were stored every 10 ps.

REST2 Simulations

The REST2 simulations utilized the same protocols as in Hajredini *et al.* (33). Structures were prepared as described above and used after NVT and NPT equilibration to initiate the REST2 simulations (34). Simulations were performed using the PLUMED 2 (56) plugin patched GROMACS 2019 software suite. The “hot” region in our simulation encompassed all solute atoms including the nucleotide and Mg^{2+} (where present). A total of 14 replicas were used, where the lowest temperature (T_{min}) was set to 300 K. A series of short 2 ns REST2 simulations were performed by varying the temperature of the highest temperature replica (T_{max}) until the average replica exchange probability fell in the acceptable range. The desired exchange probability was achieved with a T_{max} of 400 K using 14 replicas with exchange probabilities of $23.0 \pm 1.0\%$ for $\text{WzC}_{\text{CDAC}} \cdot \text{ATP} \cdot \text{Mg}^{2+}$, $23.2 \pm 0.9\%$ for $\text{WzC}_{\text{CDAC}} \cdot \text{ADP} \cdot \text{Mg}^{2+}$, $21.0 \pm 5.0\%$ for $\text{WzC}_{\text{CDAC}, \text{K540M}} \cdot \text{ATP} \cdot \text{Mg}^{2+}$, $22.1 \pm 5.0\%$ for $\text{WzC}_{\text{CDAC}, \text{K540M}} \cdot \text{ADP} \cdot \text{Mg}^{2+}$, $23.5 \pm 1.0\%$ for $\text{WzC}_{\text{CDAC}, \text{K540M}}$, and $20.1 \pm 6.2\%$ for $\text{WzC}_{\text{CDAC}, \text{Y569F}} \cdot \text{ATP} \cdot \text{Mg}^{2+}$. In all cases the temperature difference between replicas was obtained using the following expression:

$$T_n = T_{min} * e^{n \cdot \ln \left(\frac{T_{max}}{T_{min}} \right) / (n-1)} \quad (\text{S1})$$

Where n is the n^{th} replica, T_n is its corresponding temperature; T_{min} and T_{max} are the minimum and maximum temperatures, 300 K and 400 K, respectively. The factor λ , that describes that scaling of the solute intramolecular potentials, was defined by the T_{min}/T_n ratio which then was used to re-parameterize the system for each replica n . Production runs in all cases were carried out for 200 ns per replica, for an effective 28000 ns sampling time; exchange attempts were made every 20

ps. Only the final 160 ns was used in the analyses in each case. The final temperatures and scaling factors (λ) for each replica were 1: 300.0, 1.000; 2: 306.7, 0.978; 3: 313.6, 0.957; 4: 320.6, 0.936; 5: 327.8, 0.915; 6: 335.1, 0.895; 7: 342.6, 0.875; 8: 350.3, 0.856; 9: 358.1, 0.838; 10: 366.1, 0.819; 11: 374.3, 0.801; 12: 382.7, 0.784; 13: 391.3, 0.767 and 14: 400.0, 0.750.

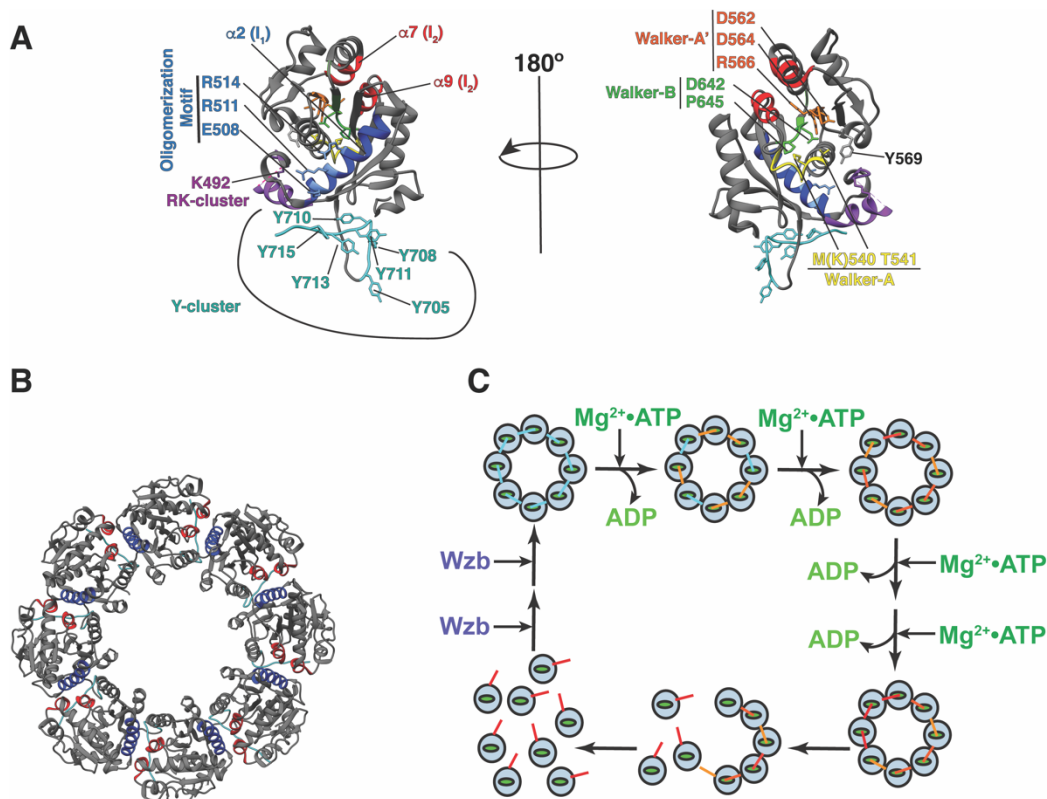


Figure S1. Structure and organization of the catalytic domain (CD) of the *Escherichia coli* BY-kinase, Wzc. **(A)** Key motifs are indicated on the structure of Wzc_{CD} (PDB: 3LA6, left panel). The Walker-A (yellow, ⁵³³GVSPSIGKT⁵⁴¹, K540 has been mutated to methionine in the crystallographic construct), Walker-B (green, ⁶³⁹VLIDTPP⁶⁴⁵) and Walker-A' (orange, ⁵⁵⁸VLLIDCDMR⁵⁶⁶) motifs. Also shown is the C-terminal tail (cyan) containing a cluster of 6 tyrosine residues (Y-cluster; Y705, Y708, Y710, Y711, Y713, and Y715). The oligomerization motif harbored on helix α2 (blue) that encodes a conserved ⁵⁰⁸EX₂RX₂R⁵¹⁴ sequence and forms the first inter-monomer interaction interface (I₁); the second interface (I₂) comprises of helices α7 and α9 (both in red). The RK-cluster comprises of helix αB at its N-terminus and a largely disordered C-terminus, is shown in purple; residue K492, discussed in the text is indicated. Y569 that contacts the α-phosphate of ATP and discussed in the text is also shown. **(B)** Individual monomers associated in a front-to-back arrangement to form a closed octameric ring resulting from the association of I₁ (blue) with I₂ (red) of neighboring monomers. **(C)** Functional cycle of phosphorylation and dephosphorylation of Wzc_{CD}. The octameric ring is stable when the Y-cluster is dephosphorylated (or contains extremely low levels of phosphorylation). In this conformation, Wzc_{CD} can stably engage ATP•Mg²⁺ and trans autophosphorylate on the Y-cluster tyrosines (orange). After several rounds of phosphorylation, the Y-cluster achieves a critical phosphorylation state (red) leading to instability in the ring that begins to rupture, finally dissociating into its constituent monomers. The Y-cluster tyrosines can then be dephosphorylated by the low molecular weight protein tyrosine phosphatase, Wzb. After a sufficiently low levels of phosphorylation of the Y-cluster is achieved through successive rounds of Wzb-catalyzed dephosphorylation, the Wzc_{CD} monomers are able to reassociate culminating in the formation of the octameric ring and the cycle reinitiates.

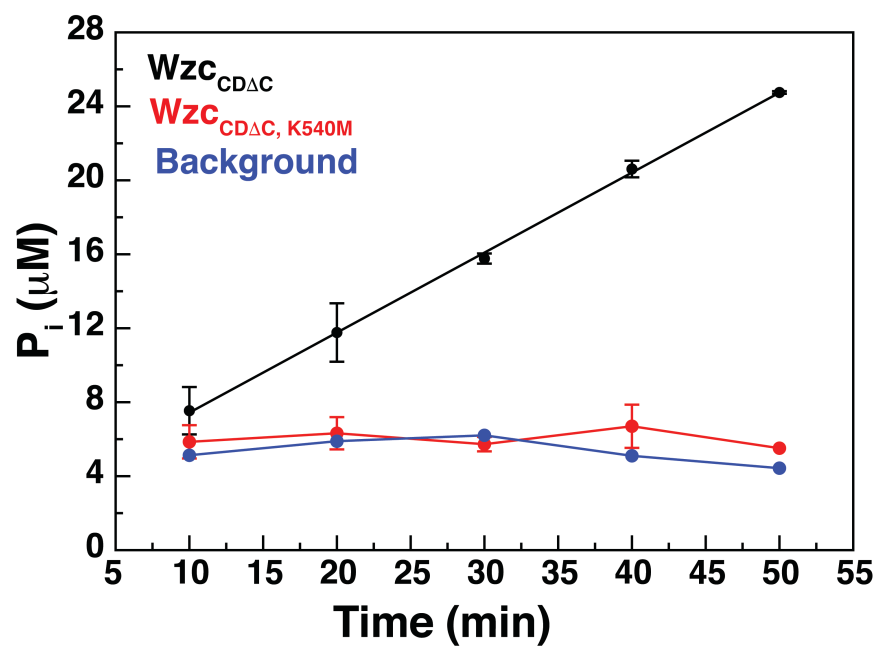


Figure S2. Comparison of the ATPase activity of $Wzc_{CD\Delta C}$ (black circles) and $Wzc_{CD\Delta C, K540M}$ (red circles). Error bars represent standard deviations across three independent measurements and the blue line shows background ATP hydrolysis in the absence of enzyme. A k_{cat} of 0.144 min^{-1} was estimated for $Wzc_{CD\Delta C}$ using a linear initial rates analysis (black line); ATP hydrolysis by $Wzc_{CD\Delta C, K540M}$ is indistinguishable from the background suggesting that it is catalytically inactive.

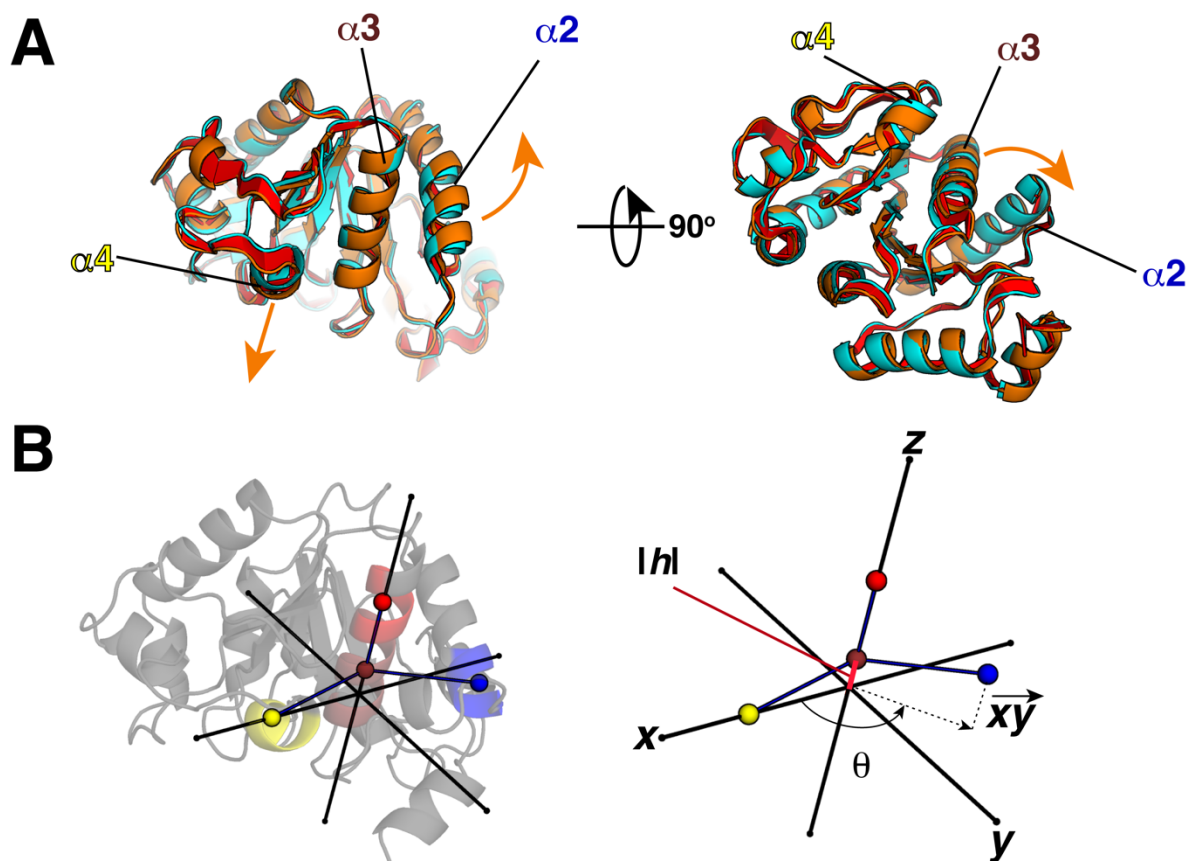


Figure S3. Definition of the two-dimensional cylindrical coordinate frame to represent the global conformational states of WzCDAC. (A) REST2 simulations on WzCDAC (unliganded or in a variety of nucleotide complexes) result in conformational distributions about two major global states - a compact closed state (CS), similar to that seen for the crystallographic monomers, and an extended open state (OS). These states are related by transformations that involve a downward displacement of helix $\alpha4$ coupled to outward rotations of helices $\alpha2$ and $\alpha3$ (orange arrows, and red interpolated path) with respect to the protein core. (B) The CS and OS, and other related states, are best represented in a two-dimensional coordinate frame defined by an angle θ , and a translation (rise), $|h|$. To define the conformations in θ - $|h|$ space, individual structures from the ensemble are aligned as follows: (1) Helix $\alpha3$ is aligned to the z-axis using the centers of mass of the C α atoms of residues 541-547 (ruby sphere) and 548-552 (red sphere). (2) Helix $\alpha4$ (yellow sphere, defined by the C α atoms of the 569-573 segment) is rotated and translated about the z-axis to place it along the positive x-axis. The rise $|h|$ is then defined as the distance between the yellow and ruby spheres along the z-axis. The angle θ is defined as the angle between the x-axis and the xy-projection of a vector extending from the ruby sphere to the center of mass of the first turn of helix $\alpha2$ (blue sphere, defined by the center of mass of the C α atoms of the 505-509 segment).

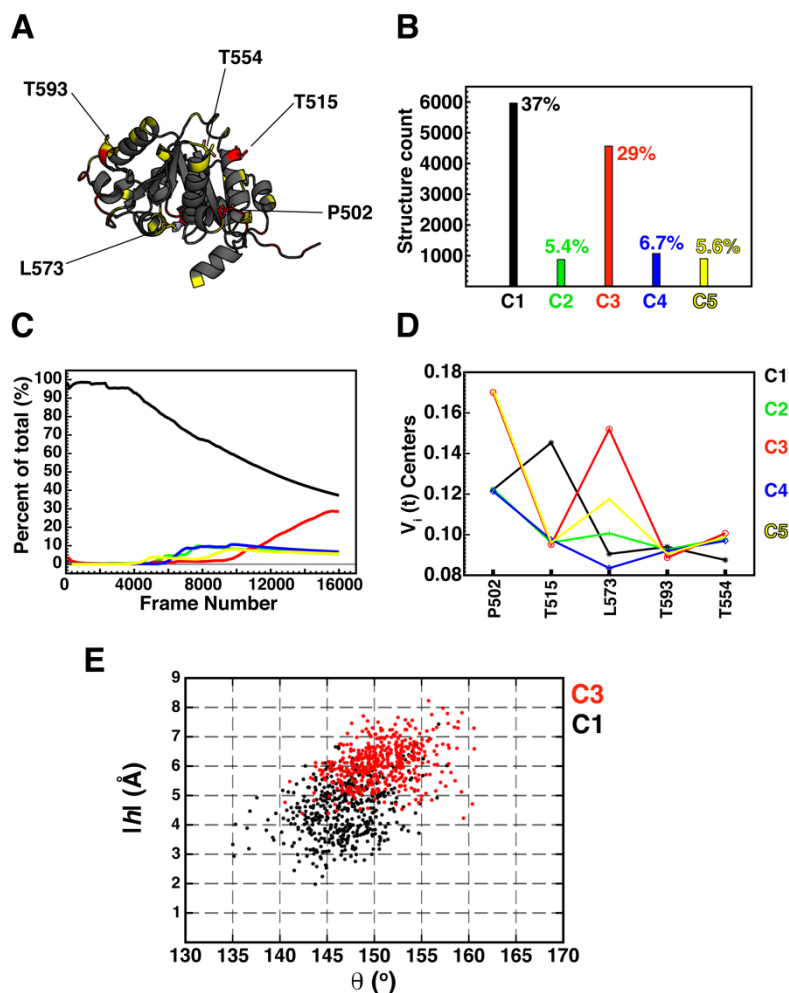


Figure S4. EVA-MS clustering of conformations from the REST2-generated ensemble of unliganded WZC_{CDAC,K540M}. **(A)** Residues that display the highest degree of variability, disregarding loops and termini, determined through the EVA procedure are colored according to the corresponding iteration: 4 (red), 5 (yellow), and 6 (light blue). Residues used for clustering are shown in stick representation and labeled. Overall, the procedure produces 408 distinct clusters, of which only 5 have significant occurrences. The five clusters (C1-C5), shown in **(B)**, account for 83.7% of the structures. The relative occurrences of each of these clusters is indicated next to the bar graph as a percent of total structures (16000). **(C)** Percent occurrence of each of the major clusters as a function of the frame count over the course of the REST2 simulations and calculated every 100 frames (1 ns) is shown. The clusters have been colored as in **(B)**. Clusters emerge early in the simulation and their relative proportions largely stabilize (or they interconvert e.g., C1 and C3) suggesting convergence. **(D)** 2-dimensional projection of the centers of each cluster where the x-axis denotes residues whose $V_i(t)$ values were used to define the 5-dimensional probability distribution function (PDF) used in the Mean Shift (MS) clustering procedure. The y-axis denotes the $V_i(t)$ value corresponding to the center of each cluster along each dimension (residue). **(E)** Projection of 500 structures randomly drawn from each of the two major clusters C1 (black) and C3 (red) onto θ - $|h|$ space. The theory behind the EVA-MS approach is described at length in the Supplementary Materials accompanying Hajredini *et al.* (33).

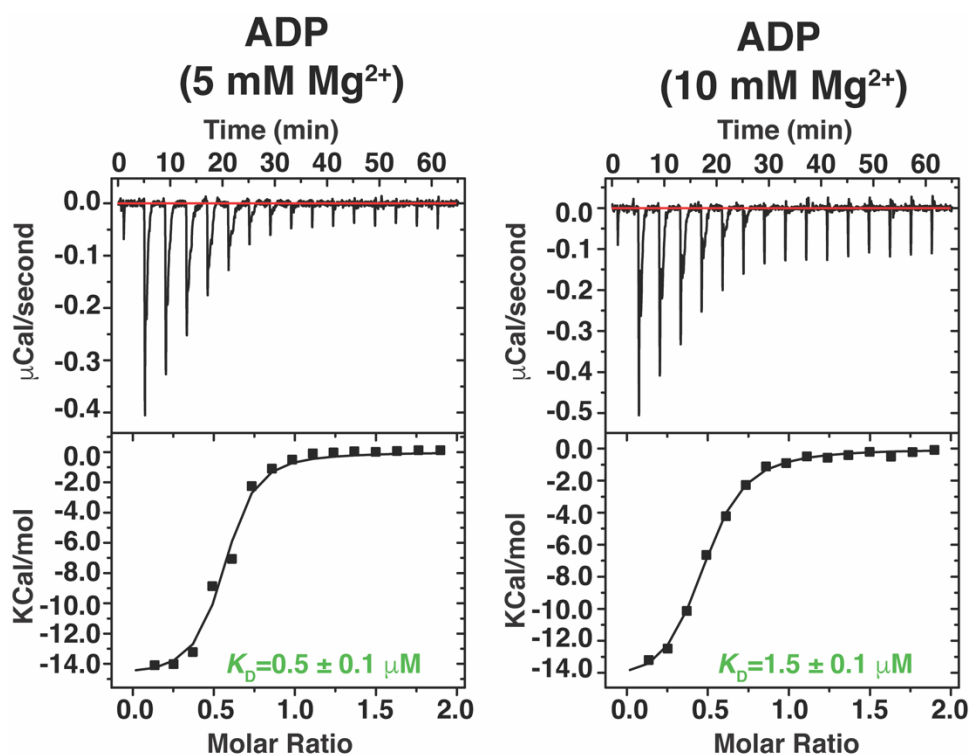


Figure S5. Representative thermograms for the interaction of WzCDD Δ C,K540M with ADP in the presence of 5 or 10 mM MgCl₂. The K_D values indicated on the figures represent those obtained from fits to a 1-site model for the thermograms shown. The corresponding mean (and standard deviation) values from three independent measurements are $K_D = 0.5 \pm 0.2 \mu\text{M}$ and $1.7 \pm 0.9 \mu\text{M}$, respectively.

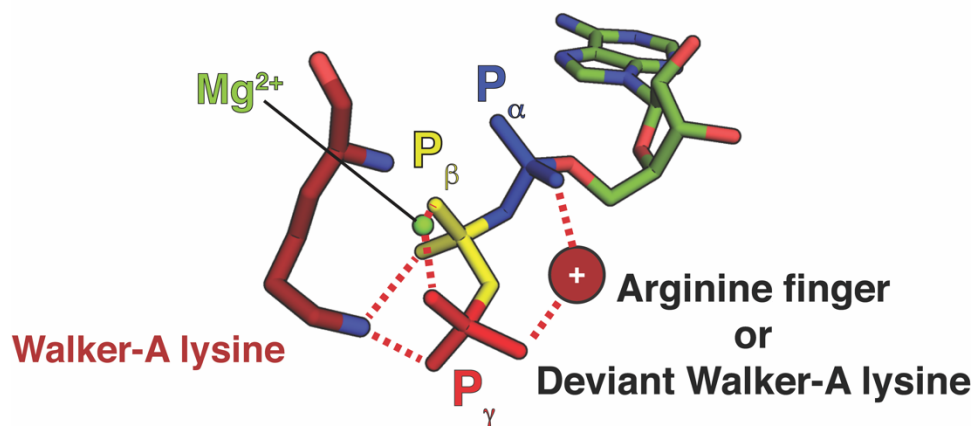


Figure S6. ATP engaged to a canonical P-loop ATPase showing the relative orientation of key catalytic elements with respect to the α - (blue), β - (yellow), and γ -phosphates (red). Key contacts of enzymic sidechains and the Mg^{2+} ion with ATP, are indicated by the dashed lines. The arginine finger (or deviant Walker-A lysine as in case of MinD discussed in the main text) inserts between the α - and γ -phosphates of ATP. This interaction leads to a rotation of the α -phosphate with respect to the β - and γ -phosphates, that are already locked into an eclipsed (or near-eclipsed) conformation by the Walker-A lysine and Mg^{2+} into an all-eclipsed conformation. This high-energy conformation facilitates cleavage of the $\beta\gamma$ -phosphoanhydride bond.

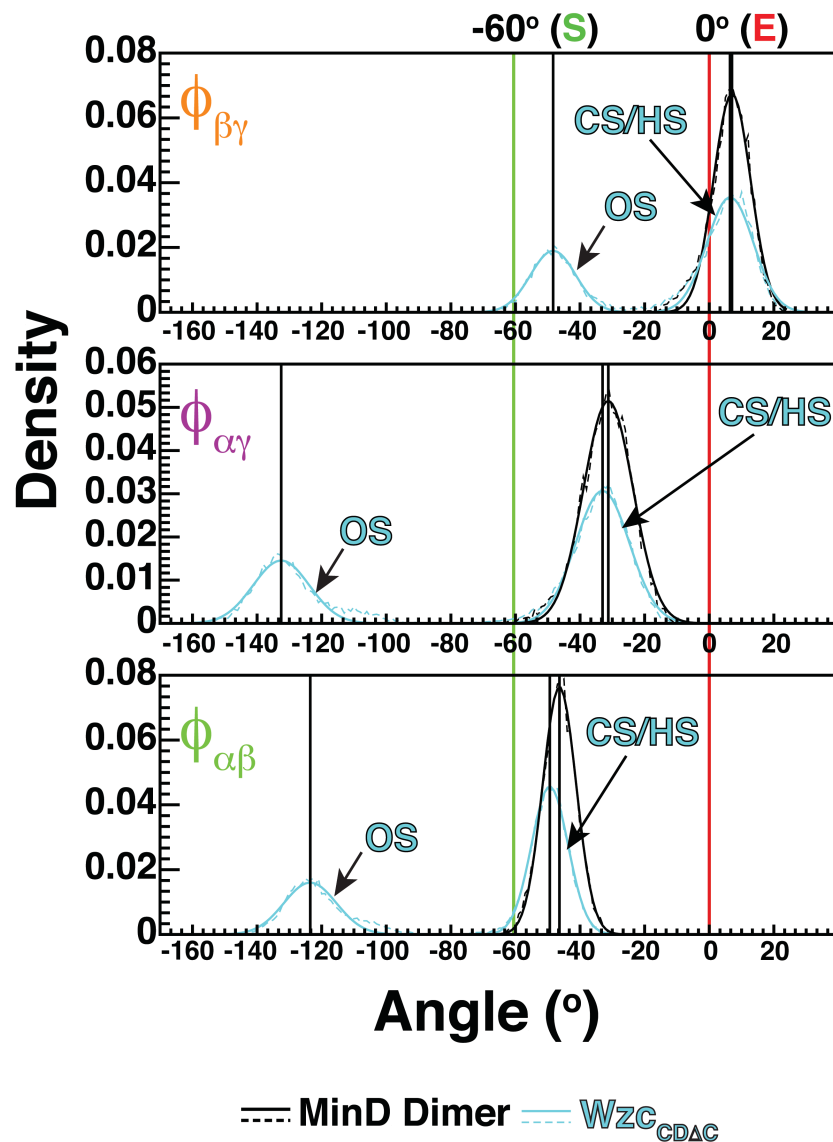


Figure S7. Comparison of ATP dihedral angle distributions in the ATP•Mg²⁺ complexes of Wzc_{CDAC} (cyan) and dimeric MinD (black). The maxima corresponding to the MinD distributions are very similar to those of the WT CS/HS.

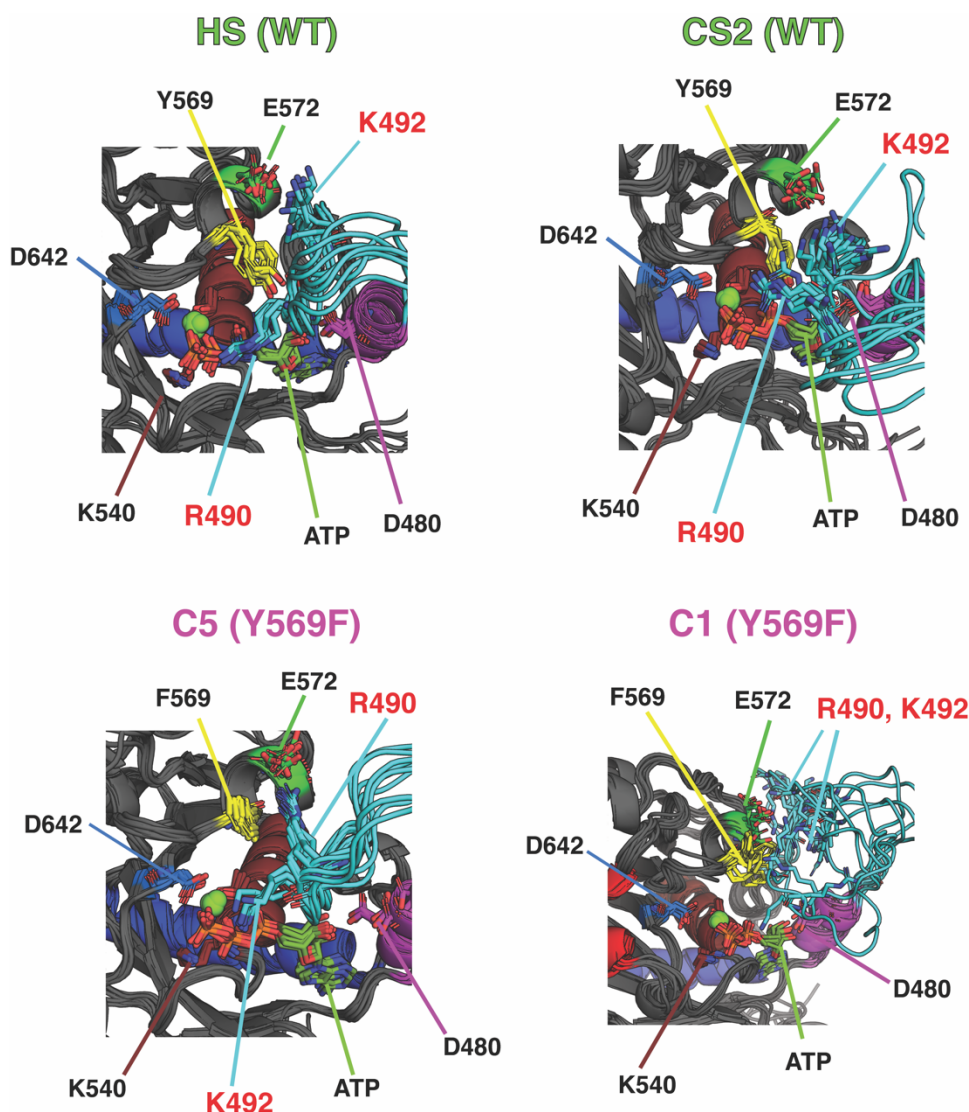


Figure S8. Comparison of the active site geometries for the HS (11% populated) and CS2 (the most dominant cluster in the CS comprising 67% of structures therein, and 30% of all structures) for the $WZ_{CDAC} \cdot ATP \cdot Mg^{2+}$ complex. Also shown are the C1 and C5 clusters of the corresponding complex of the Y569F mutant. These structures illustrate the close structural correlation between Y569 and the RK-cluster. For the wild-type complex, Y569 interacts with the α -phosphate of ATP enabling R490 to interact with the γ -phosphate (in CS2) or both the α - and γ -phosphates (in the HS) of ATP. The Y569F mutation alters the conformational characteristics of the RK-cluster. In the C5 cluster of the $WZ_{CDAC, Y569F} \cdot ATP \cdot Mg^{2+}$ complex, the RK-cluster is relatively ordered and its degree of order is comparable to the wild-type HS, but R490 and K492 have switched positions. The C5 cluster in the Y569F complex is similar to the wild-type CS in its global characteristics (see Figure 7A). It comprises 16% of all structures in the ensemble (see Figure S9 below). In the C1 cluster, that is more distant from the wild-type CS (see Figure 7A) and comprises 24% of all structures (see Figure S9 below), the side chain of F569 is highly dynamic and the RK-cluster is highly disordered.

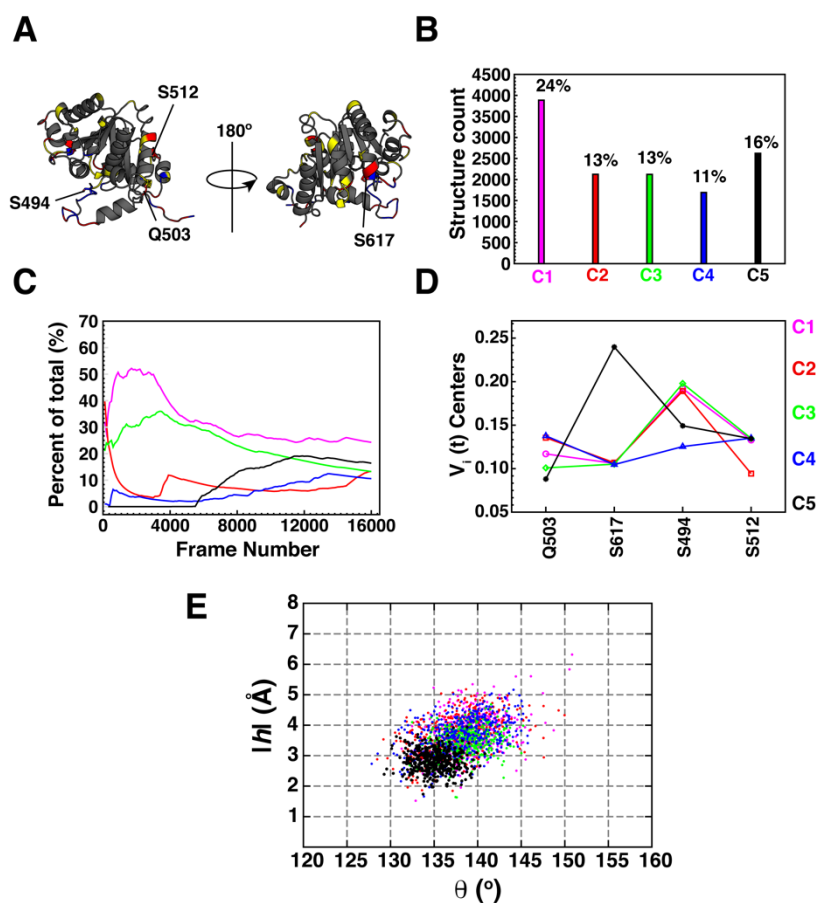


Figure S9. EVA-MS clustering of the REST2-derived ensemble of the $Wz_{CDAC,Y569F} \cdot ATP \cdot Mg^{2+}$ complex. **(A)** The most varying residues determined by the EVA procedure through iterations 2 (blue), 3 (red), 4 (yellow), and 5 (light blue). Residues used for the 4-dimensional MS clustering are shown in stick representation and labeled on the structure. **(B)** The MS clustering procedure produces a total of 153 distinct clusters of which only 5 have significant occupancies (these account for 77% of all structures); the total number of structures in each of these 5 major clusters and their relative proportions (percent of total structures) are indicated. **(C)** Fractional occurrence of each cluster in **(B)** as a function of frame number calculated every 100 frames (1 ns) is shown. The lines represent the evolution of each cluster and are colored as in **(B)**. All clusters emerge early in the simulation and their relative proportions stabilize suggesting convergence. **(D)** 2-dimensional projection of the centers of each cluster as in Figure S4D. **(E)** Projection of 500 representative structures randomly drawn from each cluster onto $\theta-|h|$ space; the points are colored as in **(B)** based on the cluster they represent.

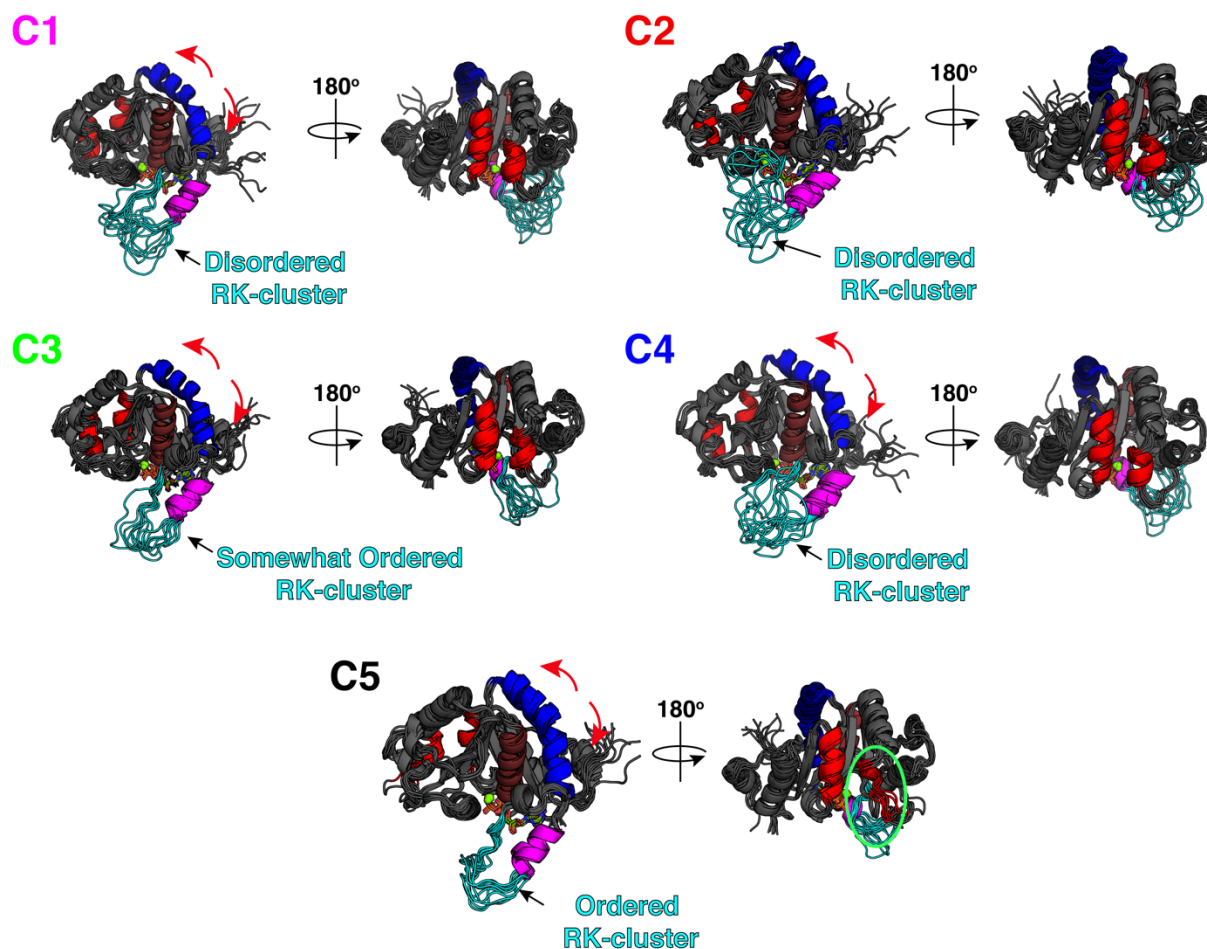
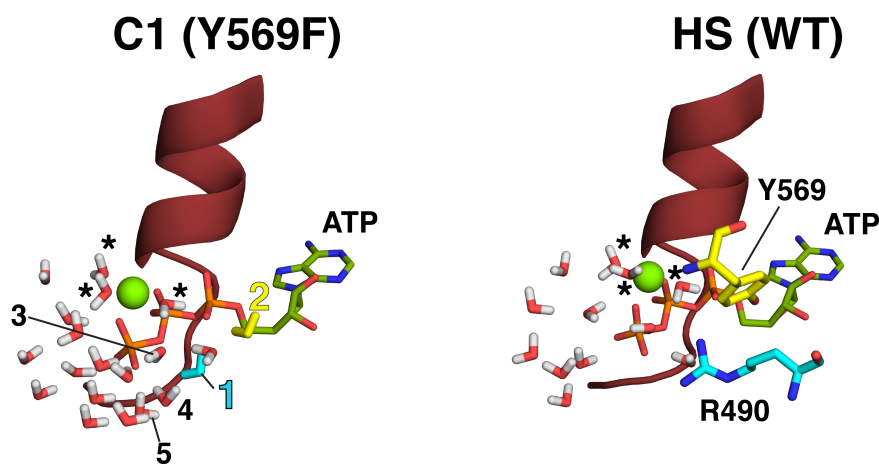


Figure S10. Structural characteristics of the 5 major clusters (C1-5) identified by application of the EVA-MS procedure on the REST2 derived ensemble of the $Wz_{CCDAC,Y569F} \cdot ATP \cdot Mg^{2+}$ complex (see Figure S9 above). Key structural features of each cluster are indicated. These features involve order/disorder in the RK-cluster, bending of $\alpha 2$ (red arrows; disruption of I_1) or unfolding of $\alpha 7$ (green ellipse; disruption of I_2).



Hydration sites displaced from C1

Site	$T\Delta S_{\text{trans}}$	$T\Delta S_{\text{orient}}$	$T\Delta S_{\text{total}}$
1	-1.85	-1.90	-3.75
2	-1.44	-1.51	-2.95
3	-1.93	-2.28	-4.21
4	-1.23	-1.17	-2.40
5	-1.81	-2.10	-3.91

Figure S11. Hydration Site Analysis (HSA) of a representative 20 ns simulation of the REST2-derived C1 cluster of the $\text{WzCCD}_{\Delta\text{C},\text{Y569F}}\cdot\text{ATP}\cdot\text{Mg}^{2+}$ complex (left; also see Figure S8) or the REST2-derived HS of $\text{WzCCD}_{\Delta\text{C}}\cdot\text{ATP}\cdot\text{Mg}^{2+}$ complex (right; also see Figure S8). Simulations and subsequent analyses were carried out in triplicate. A small segment of the $\alpha 3$ helix together with a part of the P-loop, ATP and the Mg^{2+} ion, are shown. Also shown for the HS are the sidechains of Y569 and the RK-cluster residue, R490, that coordinate ATP. The most probable distributions of water molecules in all HSA-identified hydration sites are shown in stick representation for each case. Water molecules in the first solvation shell of Mg^{2+} are indicated by the “*”. Hydration sites in the vicinity of the ATP triphosphate that are seen for the C1 structure but not for the HS structure are indicated. Particularly notable are two sites, 1 (yellow) and 2 (cyan), since these occupy positions in the C1 structure that are virtually identical to those occupied by the guanidino and hydroxyl groups of R490 and Y569, respectively, in the $\text{WzCCD}_{\Delta\text{C}}\cdot\text{ATP}\cdot\text{Mg}^{2+}$ HS. Entropic gains (translational/orientational/total) for water molecules displaced from the active site on going from an open (C1 of the Y569F complex) to a closed active site (HS), are shown in the table. Note that these values represent upper bounds given that they are referenced to an isotropic distribution in bulk solvent that contains significant transitional and orientational correlations that are notoriously difficult to estimate. The $T\Delta S_{\text{total}}$ values averaged over all three simulations are -3.70 ± 0.05 (1), -2.96 ± 0.05 (2), -4.20 ± 0.06 (3), -2.47 ± 0.06 (4), and -4.00 ± 0.25 (5).

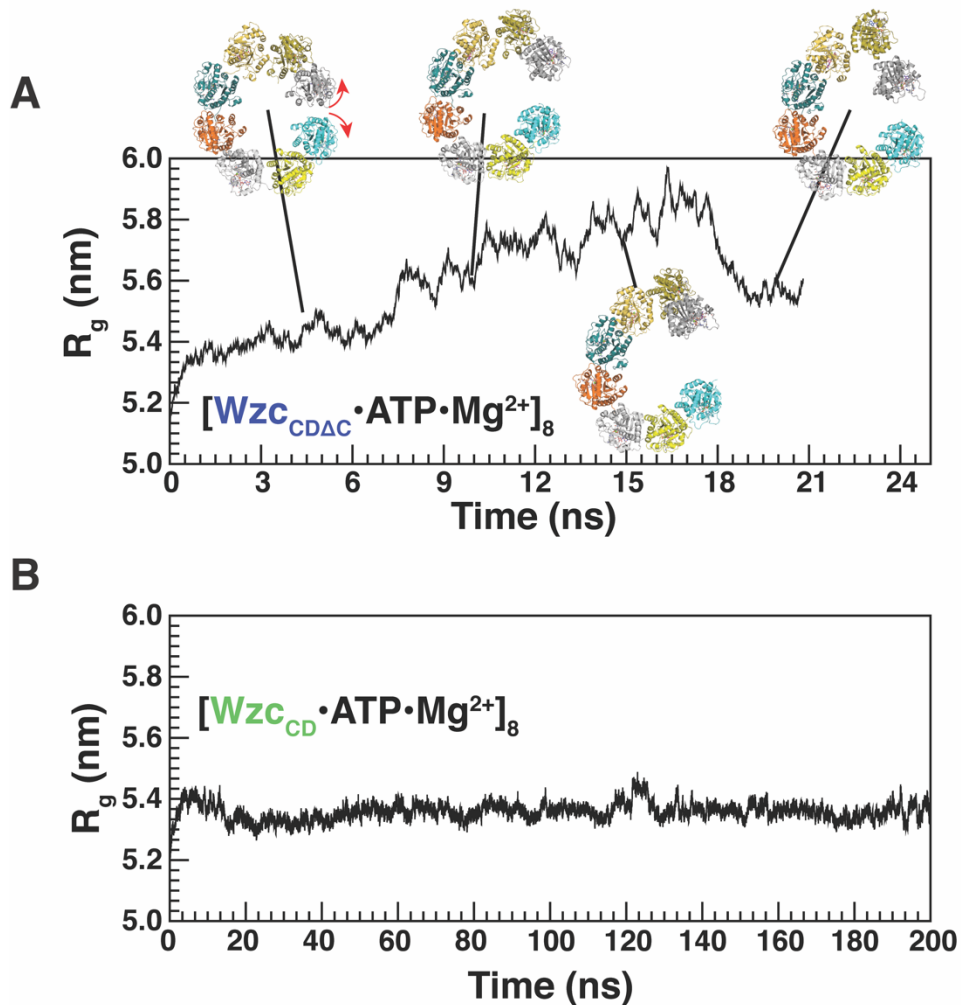


Figure S12. Integrity of the octameric ring formed by the $ATP \cdot Mg^{2+}$ complexes of Wzc_{CDAC} (that lacks the C-terminal tail represented by the 705-720 segment, and the Y-cluster therein) **(A)** or from Wzc_{CD} (that contains the C-terminal tail) **(B)** estimated using the corresponding radii of gyration (R_g). In the absence of the C-terminal tail, the octameric ring is highly unstable and begins to rupture within the first few ns of a classical MD simulation. Structures drawn from at specific timepoints in the trajectory confirm ring opening. In contrast, presence of the C-terminal tail ensures the persistence of the octameric ring over the length of a 200 ns simulation.

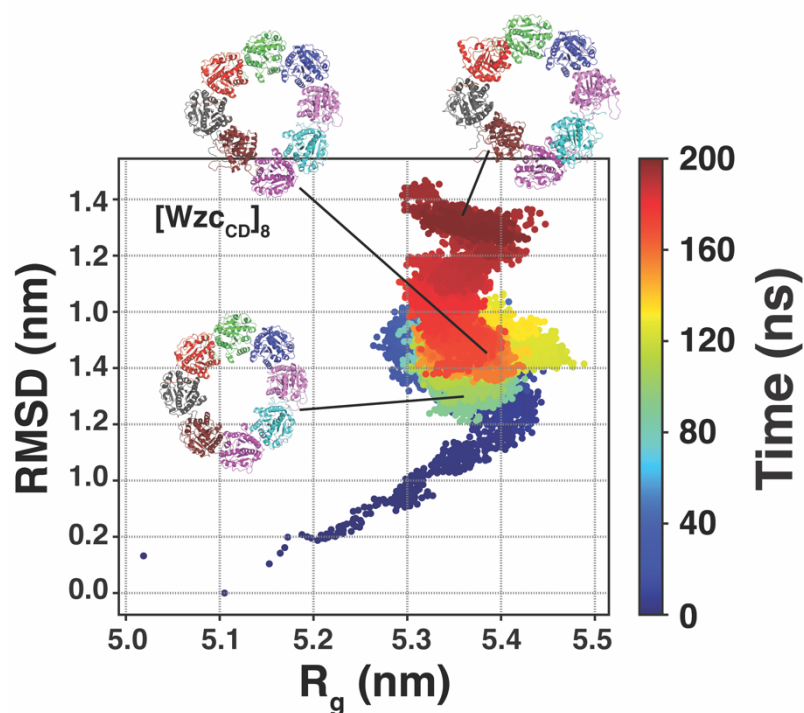


Figure S13. Representations of the ring conformations using a 2-dimensional plot correlating the radius of gyration (x-axis) and root mean squared deviation (RMSD) across all C α atoms of each chain (y-axis) from the 200 ns classical MD simulation on the $[\text{Wzc}_{\text{CD}} \cdot \text{ATP} \cdot \text{Mg}^{2+}]_8$ complex. The points are color coded based on the time stamps of the simulation. Overall, the complex shows various degrees of distortions (see representative structures at various points in R_g -RMSD space) however the integrity of the closed ring is maintained. Structures drawn from various parts of the map reveal some deviations from C_8 symmetry but no disruption in the integrity of the ring.

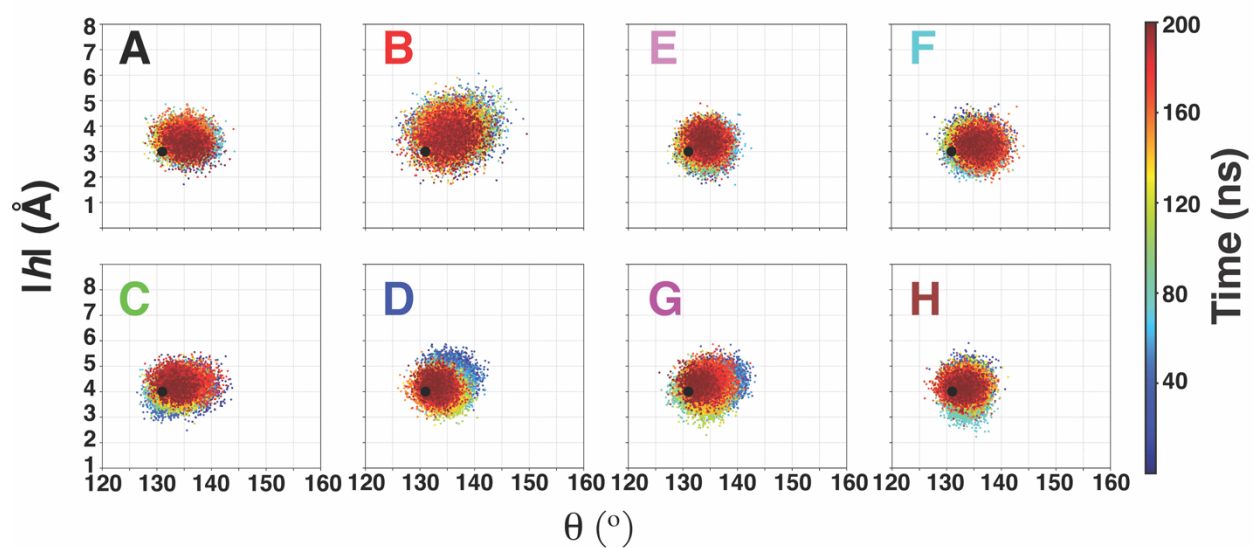


Figure S14. Conformations of individual monomers (chains A through H) from the 200 ns classical MD simulation on the $[\text{Wzc}_{\text{CD}}\cdot\text{ATP}\cdot\text{Mg}^{2+}]_8$ complex projected onto θ - $|h|$ space. Points are color coded using the simulation time stamp; the black dots indicate reference values for the corresponding monomers in the crystal structure. No OS is seen in any of the cases.

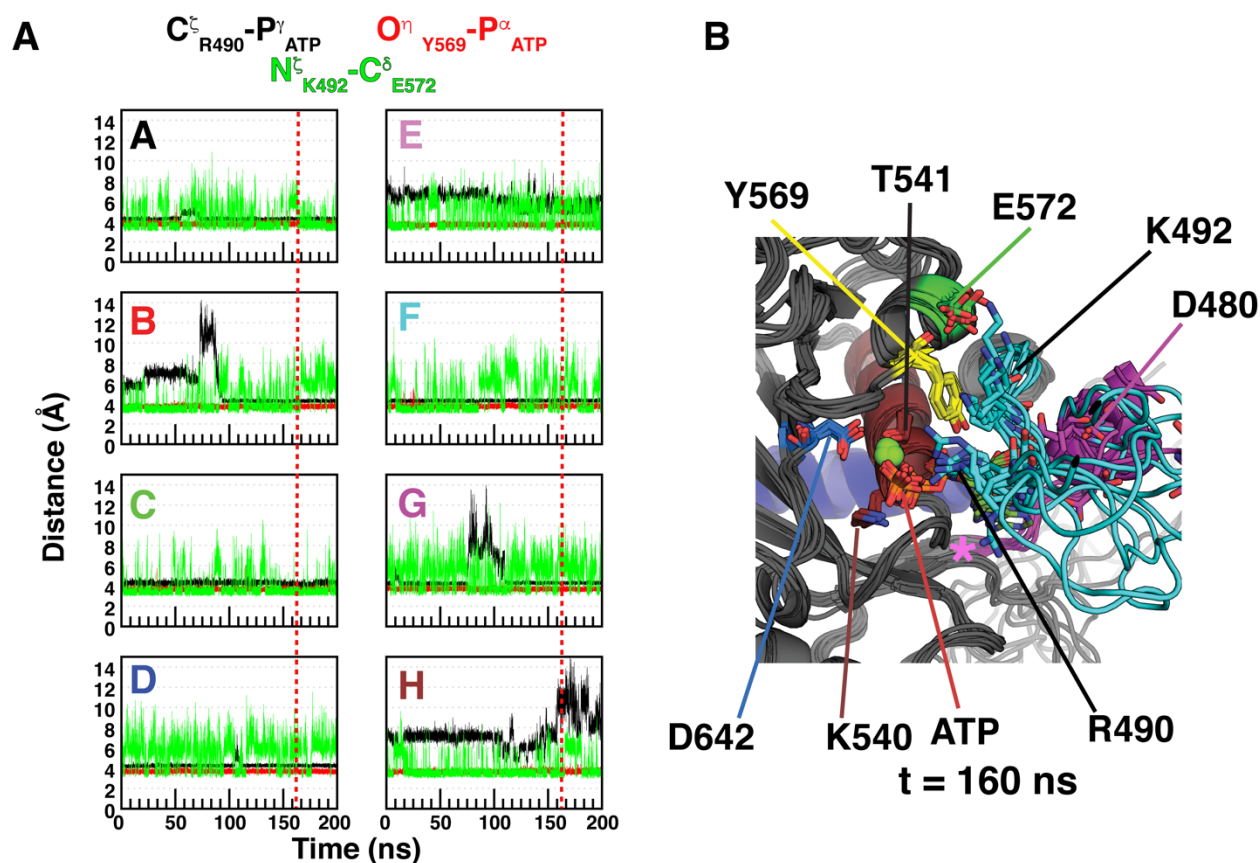


Figure S15. Active site geometry in the classical MD simulation of the $[Wz_{CD}\cdot ATP\cdot Mg^{2+}]_8$ complex. **(A)** Time-evolution of the R490-ATP contact (black), the Y569-ATP contact (red) and K492-E572 salt-bridge (green) during the simulation. The arginine finger equivalent, R490, makes stable contact with γ -phosphate of ATP in most cases but does get displaced away occasionally (chains B, G and H) though contact is restored in most cases (chains B and G). **(B)** Superposition of all 8 chains selected from the t=160 ns frame (indicated by the red dashed line); all the relevant catalytic and nucleotide coordinating residues are shown in stick representation. Relevant contacts are maintained in all cases except for chain H for which R490 (indicated by the pink ‘*’) is displaced away.