BY-kinases: protein tyrosine kinases like no other

Fatlum Hajredini[1,2], Sébastien Alphonse[1] and Ranajeet Ghose[1,2,3,4*]

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

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

*Address correspondence to Ranajeet Ghose – rghose@ccny.cuny.edu

Abstract

BY-kinases constitute a family of protein tyrosine kinases that are highly conserved in the bacterial kingdom and occur most commonly as essential components of multi-component assemblies responsible for the biosynthesis, polymerization, and export of complex polysaccharides involved in biofilm or capsule formation. BY-kinase function has been attributed to a cyclic process involving formation of an oligomeric species, its disassembly into constituent monomers, and subsequent reassembly, depending on the overall phosphorylation level of a C-terminal cluster of tyrosine residues. However, the relationship of this process to the active/inactive states of the enzyme, and the mechanism of its integration into the polysaccharide production machinery remain unclear. Here, we synthesize the substantial body of biochemical, cell-biological, structural, and computational data, acquired over the nearly three decades since the discovery of BY-kinases, to suggest means by which they fulfil their physiological function. We propose a mechanism involving temporal coordination of the assembly/disassembly process with the auto-kinase activity of the enzyme and its ability to be dephosphorylated by its counteracting phosphatase. We speculate that this temporal control enables BY-kinases to function as molecular timers that coordinate the diverse processes involved in the synthesis, polymerization, and export of complex sugar derivatives. We suggest that BY-kinases, that deploy distinctive catalytic domains resembling P-loop nucleoside triphosphatases, have uniquely adapted this ancient fold to drive functional processes through exquisite spatiotemporal control over protein-protein interactions and conformational changes. It is our hope that the hypotheses proposed here will facilitate future experiments targeting these unique protein kinases.

Tyrosine phosphorylation, long recognized as a ubiquitous mode of signal transduction in higher eukaryotes (1,2), is now known to also play a significant regulatory role in bacterial cells (3,4). In the bacterial kingdom, this post-translational modification is largely driven by members of the BY-kinase (for bacterial tyrosine kinase) family (5-7) that are widely conserved in both Gram-positive and Gram-negative species (8) but have no known eukaryotic or archaeal orthologs. While BY-kinases have been shown to contribute to diverse cellular processes (9-11), their most significant, and best studied, role is in the assembly and export of complex polysaccharides that are key components of biofilms (12,13) or virulence-determining capsules (9,14,15) in many bacterial species. In a manner reminiscent of higher eukaryotes (16), BY-kinases appear capable of integrating into complex signaling networks, that also include serine/threonine kinases (17), to drive critical physiological processes.

In the 25 years that have elapsed since the assignment of *Acinetobacter johnsonii* Ptk (18,19) as a *bona fide* protein tyrosine kinase, and a "founding" member of the BY-kinase family, a wealth of information on these enzymes has emerged from structural, biochemical and cell biological studies, reviewed in several excellent treatises (5-7,20,21). Despite the considerable progress, many critical aspects of the functional regulation of BY-kinases remain poorly understood. Recently, *in silico* approaches (22,23) have begun to provide insight into hitherto unknown regulatory features of these unique enzymes. Here, we will draw on these recent results, from both theory and experiment, to interpret the available biochemical, cell-biological, and functional data on BY-kinases to suggest a mechanism for their activation and activity in the context of bacterial physiology, specifically in the synthesis and export of complex polysaccharides. We will also utilize this review as a vehicle to propose hypotheses that are amenable to future experimental validation/testing.

Escherichia coli Wzc, the archetypal (see below) BY-kinase, has been extensively studied using a variety of biophysical and biochemical approaches in the context of the K-12 and K30 strains by the Grangeasse and Whitfield groups among others. Wzc proteins from these two *E. coli* strains are very similar (sequence identity ~52%) but are integrated into slightly different macromolecular assemblies that bridge the inner and outer membranes. The polysaccharide products are also slightly different in the two cases with the exopolysaccharide (EPS) colanic acid (CA) being produced in K-12 (24), and capsular polysaccharides (CPS) utilized in forming Group 1 capsules being generated in K30 (25). However, the overarching mechanisms underlying Wzc regulation appear to be identical in these two cases (we use Wzc interchangeably below unless explicitly stated). Further, broad mechanistic insights obtained from studies on *E. coli* Wzc are expected to be relevant across the bacterial kingdom. Therefore, in our discussion below, we will rely heavily on data for *E. coli* Wzc, invoking available results from other systems as and when necessary.

An ancient fold for a new function

In a manner characteristic of most Gram-negative BY-kinases (Table 1), *E. coli* Wzc comprises an N-terminal "transactivation" domain (TAD) (26) and C-terminal catalytic domain (CD). The TAD consists of a two-pass transmembrane domain (TMD) responsible for inner membrane localization and a periplasmic domain (PD) that is proposed to interact with the outer membrane transporter, Wza (27). The enzymatic function is encoded within the CD, that is fully cytosolic, and terminates in a flexible tail containing a cluster of tyrosine residues (the Y-cluster) (28,29) (Fig. 1A) that are sites for intermolecular autophosphorylation. Both the number and the sequential positions of the Y-cluster tyrosine residues are variable across the BY-kinase family. Contrasting Wzc, in many Gram-positive bacteria (see Table 1), most notably the firmicutes, the

TAD and the CD occur as two separate proteins, encoded by flanking genes, that must associate to form a functional, membrane-anchored complex (30). This latter group has been termed firmicute- or F-type, contrasting the proteobacterial- or P-type BY-kinases (20) like Wzc. In addition, some bacterial genomes (8) contain so-called "orphan" CDs with no obvious TAD partners (Table 1). It has been speculated that these orphan CDs are activated by the TADs of another F-type BY-kinase also present in the cell (31). Interestingly, some of the orphan CDs are predicted to contain transmembrane segments (Table 1) suggesting non-standard arrangements that likely represent species-specific variations on a common theme.

The BY-kinase CD possesses an inherent ability to self-associate, assembling into a ring-shaped octamer both in isolated form (30,32) and in the context of the full-length protein (33) when the Y-cluster is dephosphorylated or sparsely phosphorylated. Within this octameric assembly, individual CD monomers associate in a front-to-back arrangement through a conserved EX_2RX_2R sequence (oligomerization motif, O-motif; where X is any residue) on a helical segment ($\alpha2$), that defines an interaction interface (I_1) on one monomer, and engages helices $\alpha7$ and $\alpha9$, that comprise the complementary interface (I_2) on a flanking monomer (Fig. 1B). This arrangement, with C_8 rotational symmetry, enables each monomer to act as an enzyme, hosting a Y-cluster tyrosine from one flanking monomer at its active site, while simultaneously acting as a substrate by inserting a tyrosine from its own Y-cluster into the active site of the other flanking monomer (Fig. 1A, right panel). In Wzc, the $\alpha2$ helix and the O-motif therein is also utilized to dock its counteracting low molecular weight protein tyrosine phosphatase (LMW-PTP) Wzb (Fig. 1C**)** (34,35).

The BY-kinase CD, illustrated by several structures e.g., of *E. coli* Wzc (32,33) and Etk (36), *Staphylococcus aureus* CapB (30), *Vibrio cholerae* VpsO (13) etc., bears no structural

resemblance to the dual-lobed fold characteristic of eukaryotic protein kinases (37), deploying instead a variation on the P-loop scaffold (38) seen in nucleotide triphosphatases (NTPases). The P-loop fold, a primordial structural module that pre-dates the emergence of the last universal common ancestor (LUCA) (39), is also utilized by many small molecule kinases (40), but is exceedingly rare for a protein kinase; the dual-acting serine kinase/phosphorylase, HprK/P, found in some Gram-positive bacteria, being the only other known example (41). The BY-kinase active site (Fig. 1D) is lined by variants of the nucleoside-binding Walker-A (GX_5GKS/T) and Walker-B (ϕ_4DX_2P; where ϕ is a hydrophobic residue) motifs (42). Also present is a so-called Walker-A' motif (ϕ_4DXDXR) reminiscent of the DXD sequences seen in some P-loop kinases e.g., shikimate and gluconate kinases (40). The second aspartate within the Walker-A' motif has been proposed to play a critical role in recognizing the tyrosine substrate and driving chemistry (32). Mutations of conserved residues within the Walker motifs alter both the intrinsic ATPase activity and the auto-kinase activity in Wzc (23,32,43), and other BY-kinases (44,45) (Table 2). This suggests that BY-kinases utilize an NTPase-like mechanism to transfer the γ-phosphate of ATP to a substrate –OH, that in this case, is housed on a tyrosine rather than on water. It is therefore apparent that BY-kinases have evolved elaborate mechanisms establishing the primacy of their auto-kinase function over their ATPase activity as explored in greater detail below.

Assembly and disassembly, hallmarks of a functional cycle

Wzc function (32,33), and that of BY-kinases (30) in general, has been ascribed to a cyclic assembly/disassembly process coupled to the phosphorylation state of the Y-cluster. As noted above, formation of an octameric ring provides an efficient means of facilitating intramolecular autophosphorylation of the Y-cluster by allowing each monomer to simultaneously serve as both enzyme and substrate. The octameric ring disassembles into its constituent monomers after several

rounds of autophosphorylation on the Y-cluster. Disassembly depends on the overall level of Y-cluster phosphorylation rather than the phosphorylation state of any specific tyrosine or set of tyrosine residues therein (28,46). Reassembly occurs once the Y-cluster is sufficiently dephosphorylated by its counteracting phosphatase, that in the case of Wzc, is the LMW-PTP Wzb (47), and the cycle reinitiates. This oscillatory assembly/disassembly process integrates into an elaborate multi-protein assembly (see for example Fig. 4 of Whitfield (25) for an illustration) responsible for polysaccharide synthesis, polymerization, and export. The mechanism of this integration and the precise role of Wzc therein remains poorly defined.

In its simplest conceptualization, the cyclic process may be conceived as involving transitions between two discrete functional species/states (Fig. 2A), designated Σ_O and Σ_M. For Σ_O, CDs with dephosphorylated Y-clusters, assemble into an octameric ring; Σ_M comprises monomeric CDs with highly phosphorylated Y-clusters. Individual monomers that comprise Σ_O adopt conformations that allow the efficient engagement of ATP•Mg^{2+} and are thus chemistry capable i.e., "active" (discussed in detail below). This assembly allows the Y-cluster tyrosine residues to be successively autophosphorylated at the expense of ATP (generating ADP) thereby enhancing the overall level of phosphorylation. This is accompanied by increasing inter-monomer repulsion due to the accumulation of negative charges within the constituent Y-clusters. Once the level of phosphorylation exceeds a certain critical threshold (C_{phos}), estimated to be an average of about four phosphorylated tyrosine residues per monomer in Wzc (33), the electrostatic repulsion leads to dissociation of the octameric ring into its constituent monomers generating Σ_M. As noted above, Wzc oligomerization and its interaction with Wzb both involve I_1 (Figs. 1B, C) and the O-motif therein. Therefore, these two sets of interactions cannot occur simultaneously. Since Wzc is housed in the bacterial inner membrane, the resultant enhancement in local concentration favors inter-

monomer interactions over the diffusion-controlled association with the wholly cytosolic Wzb. Additional mechanisms that may also serve to exclude Wzb from interacting with CDs assembled within the Σ_O species may be conceived. For example, π-π stacking and hydrogen-bonding between the tyrosine residues have been shown to result in fibril formation (48). Given that the Wzc active site can accommodate only a single tyrosine at any given time, formation of a viscous gel-like phase-separated state (49) resulting from interactions between exposed, still unphosphorylated Y-cluster tyrosines to further spatially exclude Wzb, in a manner somewhat reminiscent of that seen in T-cell receptor signaling (50), is not beyond the realm of possibility. It is worth testing this scenario through future experiments. Irrespective of the detailed mechanisms operative, it may be reasonably argued that the engagement of Wzc by Wzb can occur with any level of efficiency only after the C_{phos} threshold has been exceeded and the octamer has disassembled i.e., in the context of Σ_M. It is also notable that Wzc CD monomers (Σ_M) adopt conformations that are not chemistry-compatible i.e., "inactive" (discussed below). This provides an additional level of control facilitating Y-cluster dephosphorylation catalyzed by Wzb while preventing the competing (and unproductive) re-phosphorylation of the Y-cluster through transient interactions between monomers (or perhaps through intra-molecular phosphoryl transfer). Further, since the Wzc I_1 is protected by the associated Wzb through an interaction that is further enhanced by the engagement of a phosphorylated Y-cluster tyrosine at the Wzb active site and the resultant avidity (~10 μM affinity estimated for interactions outside the Wzb active site, and an additional ~250 μM for the interaction of a phospho-tyrosine with the Wzb catalytic elements) (35), it is possible that dephosphorylation continues to levels below the C_{phos} threshold (hysteresis), perhaps to completion in some cases. Sufficient reduction in Y-cluster phosphorylation levels ultimately regenerates the octameric species (Σ_O) and reinitiates the cyclic process. This proposed bi-stable

behavior and accompanying hysteresis is not unusual in systems regulated through multisite phosphorylation/dephosphorylation cycles (51). It is to be mentioned here, that in some Gram-positive bacteria, the phosphatase partner of the BY-kinase is not an LMW-PTP, but rather belongs to the polymerase and histidinol phosphatase (PHP) family e.g., *Streptococcus pneumoniae* CpsB or *Staphylococcus aureus* CapC (52). Precisely, how these phosphatases engage their cognate BY-kinase CDs is not known. Additional studies, perhaps like those described previously for the Wzc/Wzb pair (34,35), are necessary to establish the universality of the mechanism described above on the precise mechanistic role of phosphatases in the context of the assembly/disassembly process.

Altered cycle, modified product

As mentioned earlier, Wzc is part of a complex multi-protein machine responsible for the synthesis, polymerization, and export of polysaccharides (24,25) utilized in capsule or biofilm formation. It has been shown that mutations that perturb the phosphorylation/dephosphorylation cycle and thereby the associated disassembly/assembly process in Wzc, and other BY-kinases (see below), strongly influence the properties of the polysaccharide product. A mechanism for potentially achieving this effect is through alteration of the relative lifetimes of the Σ_O and Σ_M species (thereby altering the periodicity of the assembly/disassembly process) to perturb the coordination between the synthesis of the sugar subunits and their subsequent polymerization. While the details of how Wzc integrates into this scheme remain obscure, mutations that putatively stabilize it (53), or other BY-kinases (13,15,44,54-57) (see Table 3), in states that mimic Σ_O or Σ_M, provide important clues. Y-cluster Y-to-F mutants, or catalytically dead (or compromised) variants that prevent (or impair) Y-cluster phosphorylation are expected to stabilize Σ_O (increase T_O in Fig. 2A**)**. Mutations that mimic the enhanced negative charge on the Y-cluster due to phosphorylation

(Y-cluster Y-to-E mutants), or variants with modified O-motifs impaired in the ability to oligomerize, would lead to a stabilized Σ_M (increased T_M). Analysis of the available functional data (Table 3) suggests that, in general, putative Σ_O-stabilizers shift the product distribution towards high molecular weight (HMW) polysaccharides or yield bacterial colonies that display enhanced mucoidy (indicative of increased levels of HMW polymers). In contrast, mutations that are expected to stabilize Σ_M result in a higher proportion of low molecular weight (LMW) polysaccharides and smoother bacterial colonies. These observations seem to suggest that increasing T_O results in aberrant extension of the polymer chains, while increasing T_M results in reduced polymerization (Fig. 2B). These correlations are, expectedly, not perfect, given the simplicity of the model shown in Fig. 2A, the diversity of experimental conditions, and system/strain-dependent idiosyncrasies. For example, *E. coli* K-12, cells expressing wild-type Wzc produce EPS with a broad size distribution in 600-2000 kDa range. In contrast, Σ_O-stabilizers (Y-cluster Y-to-F or Y569F mutants; Y569 plays a significant functional role as described below) result in a more uniform distribution centered around ~1350 kDa, while putative Σ_M-stabilizers (Y-cluster Y-to-E mutants) are unable to produce EPS, a similar phenotype as the cells expressing a construct missing the catalytic Walker-A lysine (K540M) that would be expected to stabilize Σ_O (53). Despite these nuances, it would not be too bold to speculate that the cyclic processes attributed to BY-kinase function serve to coordinate the polysaccharide synthesis, oligomerization, and perhaps, the export processes.

Open and closed, inactive and active

As was mentioned earlier, Wzc CD monomers are compatible with chemistry as part of Σ_O, and inactive as Σ_M. This provides additional temporal separation between the autophosphorylation and dephosphorylation events. Enhanced sampling molecular dynamics

(MD) simulations suggest that monomeric Wzc CD accesses distinct conformations in its various liganded states (22). These conformations are distinguished by global changes involving a displacement of helix α4 and rotations of helices α2 and α3 with respect to the protein core (Fig. 3A). These changes are best represented in a two-dimensional coordinate system comprising an angle θ and a rise $|h|$ (Fig. 3B**)** (22). Projection of the structural ensembles onto θ–$|h|$ space reveals two dominant conformations, an open state (OS) with high values of both θ and $|h|$, and a closed state (CS), with low values of both variables. While apo Wzc CD exists almost exclusively in the OS, the presence of ADP•Mg^{2+}or ATP•Mg^{2+}, but not ADP alone, results in a substantial population shift to the CS suggesting that this transition is induced by Mg^{2+}, rather than by nucleotide (Fig. 3C) (22). The OS is not compatible with the formation of the closed octameric ring. This property is clearly illustrated by the catalytically inactive K540M mutant that populates an OS that is significantly more extended than for the wild-type enzyme when free of ligands, and transitions almost completely into a CS when bound to ATP•Mg^{2+} (Fig. 3D, top panels). In line with these predictions, this mutant exists in a monomeric state at low Mg^{2+} levels despite containing a dephosphorylated Y-cluster; the presence of excess Mg^{2+} (also ATP•Mg^{2+}or ADP•Mg^{2+}) results in appearance of the oligomeric species (Fig. 3D, bottom panel) (23).

The global conformational changes that distinguish the CS and the OS are also correlated with local rearrangements of conserved elements at the active site (Fig. 3E). Within the CS, the conserved Walker-B aspartate (D642 in *E. coli* K-12 Wzc) interacts with the Walker-A signature threonine (T541) to yield a configuration necessary for the optimal coordination of the catalytically essential Mg^{2+} cation (23), suggesting a state that encodes features compatible with chemistry. This active site geometry appears to be a universal characteristic of the closed states of P-loop enzymes (58). Perhaps the geometric constraints on such a state are defined by the ability to yield

a Michaelis complex appropriate for NTP hydrolysis. In contrast, in the OS, the Walker-B aspartate (D642) traps the catalytic Walker-A lysine (K540) in a salt-bridge, with the latter also forming a hydrogen-bond with the backbone of a residue (T643) that immediately follows the Walker-B aspartate. This alignment is incompatible with Mg^{2+}, and therefore ATP•Mg^{2+} (ATP is mostly Mg^{2+}-bound, see discussion below), coordination suggesting that the OS represents an inactive state. It is notable that while all P-loop enzymes populate states that are more open than the reactive closed states described above, the former are stabilized by diverse sets of interactions that, unlike the universal conservation of the defining features of the latter, are dependent on the family/class of P-loop enzyme. This suggests that diversification of the open, inactive states has enabled adaption of the basic P-loop fold to different functional contexts by the acquisition of family-specific structural detail (58). This feature, i.e., similar active state geometries with divergent inactive state geometries, has been proposed in eukaryotic protein kinases (59), and we suspect that this may also hold in other large enzyme families.

As noted above, the presence of ATP•Mg^{2+} leads to a shift of population into the CS for Wzc CD monomers. However, in such cases the CS (while in Σ_M) ensemble comprises structures that show significant distortions at the interfaces I_1 and I_2 that are involved in oligomer formation (transition to Σ_O, Fig. 4A). These include a significant bending of $\alpha2$ (I_1) or the partial unfolding of $\alpha7$ and/or $\alpha9$ (I_2) (22) suggesting that formation of the CS upon ATP•Mg^{2+} binding within a monomeric (Σ_M) species results in a loss of several hydrogen bonds at the two interaction interfaces. This instability may be resolved by ejection of ATP•Mg^{2+} and transition back to the OS where $\alpha2$, $\alpha7$, and $\alpha9$, are intact. One may, therefore, attribute the tendency of monomeric Wzc to form the OS to the inherent helical propensity of the two interaction interfaces. Oligomer formation (Σ_O) stabilizes these interfaces through interactions between flanking monomers,

resulting in a stable CS that can then efficiently engage Mg^{2+}•ATP. One can therefore conceptualize the CS formed in the presence of ATP•Mg^{2+} in the CD monomers (Σ_M) as "mechanically strained". This strain is "released" in the oligomer (Σ_O) through inter-monomer interactions and resultant stabilization of the interfaces, I_1 and I_2 (Fig. 4B). For the Wzc CD, the CS may also be stabilized in the context of the monomer (Σ_M) by engineering "strain-releasing" mutations. Such a variant is illustrated by the "G_4-construct" generated by the *in silico* replacement of the 523-526 segment at the C-terminus of α2 by glycine residues thereby introducing disorder but away from the region that participates directly in oligomer formation (Fig. 4C) (22).

An intricate dance in space and time

Based on the observations described above, it is possible to propose a comprehensive model for the regulation of Wzc (as a representative BY-kinase) through Y-cluster phosphorylation (shown schematically in Fig. 5A). Oligomerization (Σ_O) results in a stable CS and formation of the associated critical hydrogen bond between the Walker-B aspartate (D642 in *E. coli* K-12 Wzc) and the Walker-A serine/threonine (T541) allowing optimal engagement of ATP•Mg^{2+}, and the exchange of the bound ADP through mass action. It is worth noting that in the cellular context, ATP, rather than ADP, is significantly more likely to be bound to Mg^{2+}, given the ~20-fold higher Mg^{2+} affinity of the former compared to the latter (60). Formation of the CS also generates additional structural features that are conducive to chemistry (Fig. 5B). ATP (ATP•Mg^{2+}) engages the CS with its triphosphate moiety trending towards a reactive high-energy all-eclipsed conformation (23) known to promote hydrolysis (61,62). Further, formation of the CS (and of the so-called hyper-closed state, HS, that represents a greater degree of closure, only observed in the ATP•Mg^{2+} complex, and may be considered an extreme case of the CS; see Fig. 3C) enables the interaction of a conserved tyrosine (Y569) with the α–phosphate of ATP, and

ordering of a basic element, the so-called RK-cluster (32), to allow the insertion of a positively charged residue (R490) therein into the active site. These events serve to promote the de-solvation of ATP (23), a feature that makes a favorable free-energy contribution to promote hydrolysis (63,64). The basic RK-cluster residue (R490) also appears to play a role equivalent to the so-called "arginine finger" that has been shown to "activate" the γ-phosphate of NTP in P-loop NTPases (62,65). Further, oligomerization and engagement of a substrate Y-cluster tyrosine at the active site prior to the stable binding of ATP•Mg^{2+}, and the consequent displacement of water prevents futile ATP hydrolysis through the intrinsic ATPase activity (43). This sequence of events ensures phospho-transfer to tyrosine rather than to water, thereby establishing the dominance of the kinase activity over the ATPase function of Wzc. Following multiple rounds of phosphorylation at the expense of ATP and increasing negative charge on the Y-clusters, the oligomer becomes progressively unstable, finally dissociating into its constituent monomers (Σ_M) once the C_{phos} threshold (Fig. 2A) is exceeded. While the exchange of the ADP with ATP (ATP•Mg^{2+}) can occur in optimal fashion within Σ_O, it can no longer take place within Σ_M that now adopts an OS. Thus, Wzc CD monomers remain stably bound to the ADP product from the last round of phosphorylation prior to disassembly. Mg^{2+} can no longer be optimally coordinated due to the loss of the D642-T541 hydrogen bond in the OS. Further, the presence of an unobstructed I_1 interface allows Wzb to dock onto Wzc and catalyze Y-cluster dephosphorylation. Once sufficiently dephosphorylated, the monomers can transit back into Σ_O through formation of the octameric ring to reinitiate the cycle. Thus, activity is driven by coordinated conformational transitions on multiple length scales ranging from supramolecular events (assembly and disassembly), at the level of the tertiary fold (CS and OS), and locally at the active site (sidechain reorientation, realignment of secondary elements and exclusion of solvent).

Dynamic regions as hotspots for adaptation

BY-kinases show mutation rates that are higher than those of other bacterial genes (31). Overall, there is sparse conservation outside the Walker motifs and for specific residues with known functional roles (e.g., Y569, discussed above). Nevertheless, BY-kinase CD sequences clustered using protein similarity network (PSN) analysis (66) parse into distinct groups each with multiple sub-groups containing signature motifs beyond the broadly conserved "catalytic" residues (Fig. 6; the readers are referred to Shi *et al.* (31) for further detail). A subset of defining features of these motifs is retained within some groups, but not others, suggesting that these regions have differentiated within the sub-groups and across groups of BY-kinase CDs. Interestingly, some of these conserved elements overlap with those that display significant conformational variability between functional states e.g., the CS and OS, discussed above. A prime example of such a region includes helices α7 and α9 that form I_2 (Fig. 6**)** and comprise regions of instability within the CS (CS_M in Fig. 5A) of the monomeric (Σ_M) species.

A possible role of mutations localized within key conformationally labile regions could be to impart fitness in pathogenic species by altering the stabilities of the OS and CS thereby affecting the relative lifetimes of Σ_M and Σ_O and allowing rapid adaption to environmental cues. It has been shown that Wzc mutations enable drug-resistant strains of *Klebsiella pneumoniae* to adapt to the urinary tract by regulating mucoidy (67). G569, that lies on α4 (rigid body motions of α4 distinguish the CS and the OS, see Fig. 3A), was found to be extensively mutated (G569C, G569V and G569D). These modifications were causative for enhanced mucoidy, a characteristic that we attribute to a stabilized Σ_O. Indeed, these mutations also resulted in decreased kinase activity, a reduction that would yield an increased T_O in our simple model (Fig. 2). Further, an insertion at the C-terminal end of α2 (at position 526) was also found to enhance mucoidy without reducing

kinase activity (67). It is notable that this insertion coincides with the region for the generation of the G_4-construct that stabilizes the CS in *E. coli* Wzc (22), as discussed above **(**Fig. 4C). It conceivable that this insertion, by increasing disorder in this region, produces a similar effect in *K. pneumoniae* Wzc of stabilizing the CS and thereby Σ_O species.

A molecular timer for polysaccharide production

As noted above, perturbations in phosphorylation/dephosphorylation-coupled oscillations between the octameric (Σ_O) and monomeric (Σ_M) species modify the physiological output. Specifically, it appears that modulation of BY-kinase activity alters the size distribution of the polysaccharide product suggesting a change in the balance between the synthesis of constituent sugar units and their polymerization within the multi-component machine (24,25). One may therefore speculate that the proposed switch-like behavior driven by Y-cluster phosphorylation could potentially provide a timing mechanism to maintain coherence between these events. We propose that BY-kinase activity serves the role of a "molecular timer" that coordinates the various processes involved the generation of the polysaccharide product. We suggest that this coordination is achieved by modulating the lifetimes of the octameric (Σ_O) and monomeric (Σ_M) forms of the enzyme to enable adaptation to specific conditions in a particular bacterium under altered environmental conditions (short evolutionary timescales) or in different bacterial species (long evolutionary timescales). These lifetimes are related to the degree of instability introduced into the oligomeric form by each successive Y-cluster phosphorylation event, and on the corresponding autophosphorylation and dephosphorylation rates. One could conceive of several ways by which this could potentially be achieved. For example, changes in the rates of phosphorylation and/or dephosphorylation would affect T_O and/or T_M (Fig. 2B). Alternatively, modifications in the relative stabilities of the CS or the OS and therefore the ability to generate Σ_O or Σ_M species with

increased/decreased C_{phos} thresholds could also provide similar outcomes. For example, a CS stabilized through the introduction of disorder at the C-terminus of α2, as in the G_4-construct of Wzc mentioned above (22), likely generates a Σ_O that would require a higher C_{phos} threshold for dissociation compared to the wild-type enzyme.

An intriguing example that suggests functional regulation by modulation of the relative lifetimes of the Σ_O and Σ_M species is provided by the *Xanthomonas campestris* inner-membrane protein, GumC. However, in this case, modulation is achieved through a unique phosphorylation-independent mechanism (68). GumC contains a cytosolic domain that closely resembles a BY-kinase CD but lacks the Y-cluster. Overexpression of GumC results in xanthan molecules with increased chain lengths suggestive of a stabilized Σ_O and an enhanced T_O. While the detailed biochemical characteristics of this system are not yet known, it is conceivable that activity in GumC is regulated through clustering in a manner reminiscent of the scaffolding function of eukaryotic pseudokinases (indeed it would not be too much of a stretch to classify GumC as a "pseudo-BY-kinase") (69). Thus, enhanced local concentration from increased expression of this membrane-anchored protein potentially stabilizes the oligomeric Σ_O and yields the characteristic phenotype. The Σ_O species can be "dissolved" through reduced expression levels and decreased local concentration.

Looking back, looking forward

Here, we have attempted to synthesize the large body of cell-biological, biochemical, structural, and computational data into a coherent framework to interpret the mechanisms underlying the activation and activity of BY-kinases. We suggest that BY-kinases integrate into the polysaccharide synthesis machinery by functioning as molecular timers that serve to synchronize the synthesis and polymerization of complex carbohydrates in a manner unique to a

given bacterial species. Such "timing control" has been suggested for other multisite phosphorylation driven systems (70). Several observations raise the tantalizing possibility of such a role for BY-kinase function. In *S. aureus*, activity of the functional complex comprising the catalytic CapB and its membrane-anchored TAD partner, CapA, appears to synchronize capsule assembly and cell wall biosynthesis (71). In *S. pneumoniae*, the phosphorylation/dephosphorylation cycle of the BY-kinase CpsD is required for the appropriate coupling of cell-division and capsule synthesis (14). However, this idea of a molecular clock remains in the realm of speculation pending verification through precisely designed future experiments. In this regard, a defining set of measurements could involve obtaining quantitative estimates of the cellular lifetimes of the Σ_O and Σ_M species (and perhaps additional ones in more comprehensive models), the influence of alterations (through mutations like those described above) thereon, and resultant changes in the nature of the polysaccharide product.

In the discussions above, we have largely focused on interactions within the cytosol i.e., those involving the BY-kinase CD and the catalytic activity encoded therein. There is mounting evidence that extra-cytoplasmic events also affect the physiological outcomes. For example, despite close similarities, *E. coli* K-12 Wzc is incapable of substituting for the native enzyme in *E. coli* K30 cells apparently decoupling the synthesis of capsular polysaccharides and their display on the cell surface suggesting altered export properties (72). It has been proposed that the outer membrane transporter Wza is gated through interactions with the Wzc PD in the periplasm (33). Deletion of the so-called Motif 3 (33), a helical segment in the Wzc PD that is a likely interaction partner of Wza, results in a phenotype that is similar to that of a Wza deletion strain without affecting the ability to form the octameric species or perturbing kinase activity (33,73). Whether, and how these intra- and extra-cytoplasmic events communicate with each other remains unknown.

It is possible that the extra-cytoplasmic unit, necessarily part of a functional assembly involving a BY-kinase (as a single protein in case of a P-type, or as a separate TAD protein for an F-type BY-kinase), could serve as a sensor to elicit activating/deactivating conformational changes in the CD by means of the TMD, perhaps also influenced by the lipid membrane. Such a mode of communication has been suggested for the eukaryotic receptor tyrosine kinase (eRTK), epidermal growth factor receptor (EGFR) (74,75). Indeed, observation of EPS-mediated regulation of the activity of the BY-kinase EpsB through its membrane-bound TAD EpsA in the Gram-positive *Bacillus subtilis* (57) hints at such possibilities. One can, however, expect a larger diversity in the roles of the extracellular/periplasmic domains and their interactions compared to those within the cytosol across the bacterial kingdom. A comparison of the AlphaFold-predicted structures (76,77) of the TADs of representative P-type and F-type BY-kinases show more variability compared to the corresponding CDs. These differences could, in some cases, be related to the presence/absence of a protein partner (e.g., Wza for Wzc). It has also been suggested that in addition to stimulating kinase function, the TAD may itself encode specific enzymatic activity e.g., for one of the two CapA variants in *S. aureus*, CapA1 (71).

It is evident, that a lot is still unknown about the unique enzymes that are the BY-kinases, in terms of structure, mechanism and the diversity therein. Answering these open questions, testing the many predictions made above, and refining the proposed hypotheses, will involve extensive future experimentation through the application of techniques ranging from structural/biochemical studies of intact functional assemblies to cell-biological studies and single-molecule analyses combined with computational approaches. The design of these approaches can perhaps draw inspiration from those that have been crucial in providing critical insights into the regulation of the eukaryotic *functional* cousins of the BY-kinases, the eRTKs (74,78). We expect that these

studies will provide fertile grounds for BY-kinase research during this, the third decade since their annotation, and beyond, given the importance of these systems both in terms of fundamental biology and the opportunity for the design of novel antibacterial therapies (79).

Conflicts of interest

The authors declare that they have no known conflicts of interest with the contents of this article.

Acknowledgements

This work was supported by NSF grant MCB1937937. Drs. Shi and Mijakovic (Chalmers University of Technology) are thanked for providing a high-resolution version of Fig. 5 from Ref. 60, parts of which have been incorporated into Fig. 6 of the present manuscript. The authors thank Dr. Luciana Giono for input in refining the figures.

Author contributions

F.H., S. A. and R. G. conceptualizing, writing, and editing; F.H., S. A. and R. G. preparing the figures; R. G. supervision and funding.

References

1. Hunter, T. (2009) Tyrosine phosphorylation: thirty years and counting. *Curr. Opin. Cell Biol.* **21**, 140-146
2. Hunter, T. (2014) The genesis of tyrosine phosphorylation. *CSH Persp. Biol.* **6**, a020644
3. Cozzone, A. J., Grangeasse, C., Doublet, P., and Duclos, B. (2004) Protein phosphorylation on tyrosine in bacteria. *Arch. Microbiol.* **181**, 171-181
4. Grangeasse, C., Cozzone, A. J., Deutscher, J., and Mijakovic, I. (2007) Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem. Sci.* **32**, 86-94
5. Bechet, E., Guiral, S., Torres, S., Mijakovic, I., Cozzone, A. J., and Grangeasse, C. (2009) Tyrosine-kinases in bacteria: from a matter of controversy to the status of key regulatory enzymes. *Amino Acids* **37**, 499-507
6. Grangeasse, C., Terreux, R., and Nessler, S. (2010) Bacterial tyrosine-kinases: structure-function analysis and therapeutic potential. *Biochim. Biophys. Acta* **1804**, 628-634
7. Grangeasse, C., Nessler, S., and Mijakovic, I. (2012) Bacterial tyrosine kinases: evolution, biological function and structural insights. *Philos. Trans. R. Soc. Lond. B* **367**, 2640-2655
8. Jadeau, F., Grangeasse, C., Shi, L., Mijakovic, I., Deleage, G., and Combet, C. (2012) BYKdb: the Bacterial protein tYrosine Kinase database. *Nucleic Acids Res.* **40**, D321-324
9. Ernst, C. M., Braxton, J. R., Rodriguez-Osorio, C. A., Zagieboylo, A. P., Li, L., Pironti, A., Manson, A. L., Nair, A. V., Benson, M., Cummins, K., Clatworthy, A. E., Earl, A. M., Cosimi, L. A., and Hung, D. T. (2020) Adaptive evolution of virulence and persistence in carbapenem-resistant *Klebsiella pneumoniae*. *Nat. Med.* **26**, 705-711

10. Kolot, M., Gorovits, R., Silberstein, N., Fichtman, B., and Yagil, E. (2008) Phosphorylation of the integrase protein of coliphage HK022. *Virology* **375**, 383-390

11. Mijakovic, I., Petranovic, D., Macek, B., Cepo, T., Mann, M., Davies, J., Jensen, P. R., and Vujaklija, D. (2006) Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res.* **34**, 1588-1596

12. Danese, P. N., Pratt, L. A., and Kolter, R. (2000) Exopolysaccharide production is required for development of *Escherichia coli* K-12 biofilm architecture. *J. Bacteriol.* **182**, 3593-3596

13. Schwechheimer, C., Hebert, K., Tripathi, S., Singh, P. K., Floyd, K. A., Brown, E. R., Porcella, M. E., Osorio, J., Kiblen, J. T. M., Pagliai, F. A., Drescher, K., Rubin, S. M., and Yildiz, F. H. (2020) A tyrosine phosphoregulatory system controls exopolysaccharide biosynthesis and biofilm formation in *Vibrio cholerae*. *PLoS Pathog.* **16**, e1008745

14. Nourikyan, J., Kjos, M., Mercy, C., Cluzel, C., Morlot, C., Noirot-Gros, M. F., Guiral, S., Lavergne, J. P., Veening, J. W., and Grangeasse, C. (2015) Autophosphorylation of the bacterial tyrosine-kinase CpsD connects capsule synthesis with the cell cycle in *Streptococcus pneumoniae*. *PLoS Genet.* **11**, e1005518

15. Nakamoto, R., Kwan, J. M. C., Chin, J. F. L., Ong, H. T., Flores-Kim, J., Midonet, C., VanNieuwenhze, M. S., Guan, X. L., and Sham, L. T. (2021) The bacterial tyrosine kinase system CpsBCD governs the length of capsule polymers. *Proc. Natl. Acad. Sci. USA* **118**

16. Brognard, J., and Hunter, T. (2011) Protein kinase signaling networks in cancer. *Curr. Opin. Genet. Dev.* **21**, 4-11

17. Grangeasse, C. (2016) Rewiring the pneumococcal cell cycle with serine/threonine- and tyrosine-kinases. *Trends Microbiol.* **24**, 713-724

18. Duclos, B., Grangeasse, C., Vaganay, E., Riberty, M., and Cozzone, A. J. (1996) Autophosphorylation of a bacterial protein at tyrosine. *J. Mol. Biol.* **259**, 891-895

19. Grangeasse, C., Doublet, P., Vaganay, E., Vincent, C., Deleage, G., Duclos, B., and Cozzone, A. J. (1997) Characterization of a bacterial gene encoding an autophosphorylating protein tyrosine kinase. *Gene* **204**, 259-265

20. Chao, J. D., Wong, D., and Av-Gay, Y. (2014) Microbial protein-tyrosine kinases. *J. Biol. Chem.* **289**, 9463-9472

21. Mijakovic, I., Grangeasse, C., and Turgay, K. (2016) Exploring the diversity of protein modifications: special bacterial phosphorylation systems. *FEMS Microbiol. Rev.* **40**, 398-417

22. Hajredini, F., Piserchio, A., and Ghose, R. (2020) Long-range dynamic correlations regulate the catalytic activity of the bacterial tyrosine kinase Wzc. *Sci. Adv.* **6**, eabd3718.

23. Hajredini, F., and Ghose, R. (2021) An ATPase with a twist: A unique mechanism underlies the activity of the bacterial tyrosine kinase, Wzc. *Sci. Adv.* **7**, eabj5836

24. Schmid, J., Sieber, V., and Rehm, B. (2015) Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Front. Microbiol.* **6**, 496

25. Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* **75**, 39-68

26. Jadeau, F., Bechet, E., Cozzone, A. J., Deleage, G., Grangeasse, C., and Combet, C. (2008) Identification of the idiosyncratic bacterial protein tyrosine kinase (BY-kinase) family signature. *Bioinformatics* **24**, 2427-2430

27. Collins, R. F., Beis, K., Dong, C., Botting, C. H., McDonnell, C., Ford, R. C., Clarke, B. R., Whitfield, C., and Naismith, J. H. (2007) The 3D structure of a periplasm-spanning

platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **104**, 2390-2395

28. Grangeasse, C., Doublet, P., and Cozzone, A. J. (2002) Tyrosine phosphorylation of protein kinase Wzc from *Escherichia coli* K12 occurs through a two-step process. *J. Biol. Chem.* **277**, 7127-7135

29. Wugeditsch, T., Paiment, A., Hocking, J., Drummelsmith, J., Forrester, C., and Whitfield, C. (2001) Phosphorylation of Wzc, a tyrosine autokinase, is essential for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *J. Biol. Chem.* **276**, 2361-2371

30. Olivares-Illana, V., Meyer, P., Bechet, E., Gueguen-Chaignon, V., Soulat, D., Lazereg-Riquier, S., Mijakovic, I., Deutscher, J., Cozzone, A. J., Laprevote, O., Morera, S., Grangeasse, C., and Nessler, S. (2008) Structural basis for the regulation mechanism of the tyrosine kinase CapB from *Staphylococcus aureus*. *PLoS Biol.* **6**, e143

31. Shi, L., Ji, B., Kolar-Znika, L., Boskovic, A., Jadeau, F., Combet, C., Grangeasse, C., Franjevic, D., Talla, E., and Mijakovic, I. (2014) Evolution of bacterial protein-tyrosine kinases and their relaxed specificity toward substrates. *Genome Biol. Evol.* **6**, 800-817

32. Bechet, E., Gruszczyk, J., Terreux, R., Gueguen-Chaignon, V., Vigouroux, A., Obadia, B., Cozzone, A. J., Nessler, S., and Grangeasse, C. (2010) Identification of structural and molecular determinants of the tyrosine-kinase Wzc and implications in capsular polysaccharide export. *Mol. Microbiol.* **77**, 1315-1325

33. Yang, Y., Liu, J., Clarke, B. R., Seidel, L., Bolla, J. R., Ward, P. N., Zhang, P., Robinson, C. V., Whitfield, C., and Naismith, J. H. (2021) The molecular basis of regulation of bacterial capsule assembly by Wzc. *Nat. Commun.* **12**, 4349

34. Alphonse, S., Djemil, I., Piserchio, A., and Ghose, R. (2022) Structural basis for the recognition of the bacterial tyrosine kinase Wzc by its cognate tyrosine phosphatase Wzb. *Proc. Natl. Acad. Sci. USA* **119**, e2201800119

35. Temel, D. B., Dutta, K., Alphonse, S., Nourikyan, J., Grangeasse, C., and Ghose, R. (2013) Regulatory interactions between a bacterial tyrosine kinase and its cognate phosphatase. *J. Biol. Chem.* **288**, 15212-15228

36. Lee, D. C., Zheng, J., She, Y. M., and Jia, Z. (2008) Structure of *Escherichia coli* tyrosine kinase Etk reveals a novel activation mechanism. *EMBO J.* **27**, 1758-1766

37. Taylor, S. S., and Radzio-Andzelm, E. (1994) Three protein kinase structures define a common motif. *Structure* **2**, 345-355

38. Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990) The P-loop - a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**, 430-434

39. Longo, L. M., Jablonska, J., Vyas, P., Kanade, M., Kolodny, R., Ben-Tal, N., and Tawfik, D. S. (2020) On the emergence of P-Loop NTPase and Rossmann enzymes from a beta-alpha-beta ancestral fragment. *eLife* **9**, e64415

40. Leipe, D. D., Koonin, E. V., and Aravind, L. (2003) Evolution and classification of P-loop kinases and related proteins. *J. Mol. Biol.* **333**, 781-815

41. Fieulaine, S., Morera, S., Poncet, S., Mijakovic, I., Galinier, A., Janin, J., Deutscher, J., and Nessler, S. (2002) X-ray structure of a bifunctional protein kinase in complex with its protein substrate HPr. *Proc. Natl. Acad. Sci. USA* **99**, 13437-13441

42. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) Distantly related sequence in the alpha and beta subunits of ATP synthase, myosin, kinases and other ATP requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951

43. Soulat, D., Jault, J. M., Geourjon, C., Gouet, P., Cozzone, A. J., and Grangeasse, C. (2007) Tyrosine-kinase Wzc from *Escherichia* coli possesses an ATPase activity regulated by autophosphorylation. *FEMS Microbiol Lett.* **274**, 252-259

44. Niemeyer, D., and Becker, A. (2001) The molecular weight distribution of succinoglycan produced by *Sinorhizobium meliloti* is influenced by specific tyrosine phosphorylation and ATPase activity of the cytoplasmic domain of the ExoP protein. *J. Bacteriol.* **183**, 5163-5170

45. Mijakovic, I., Poncet, S., Boel, G., Maze, A., Gillet, S., Jamet, E., Decottignies, P., Grangeasse, C., Doublet, P., Le Marechal, P., and Deutscher, J. (2003) Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J.* **22**, 4709-4718

46. Paiment, A., Hocking, J., and Whitfield, C. (2002) 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

47. Vincent, C., Doublet, P., Grangeasse, C., Vaganay, E., Cozzone, A. J., and Duclos, B. (1999) Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb. *J. Bacteriol.* **181**, 3472-3477

48. Anand, B. G., Prajapati, K. P., Shekhawat, D. S., and Kar, K. (2018) Tyrosine-Generated nanostructures initiate amyloid cross-seeding in proteins leading to a lethal aggregation trap. *Biochemistry* **57**, 5202-5209

49. Mittag, T., and Pappu, R. V. (2022) A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **82**, 2201-2214

50. Su, X., Ditlev, J. A., Hui, E., Xing, W., Banjade, S., Okrut, J., King, D. S., Taunton, J., Rosen, M. K., and Vale, R. D. (2016) Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**, 595-599

51. Markevich, N. I., Hoek, J. B., and Kholodenko, B. N. (2004) Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J. Cell Biol.* **164**, 353-359

52. Standish, A. J., and Morona, R. (2014) The role of bacterial protein tyrosine phosphatases in the regulation of the biosynthesis of secreted polysaccharides. *Antioxid. Redox. Signal.* **20**, 2274-2289

53. Obadia, B., Lacour, S., Doublet, P., Baubichon-Cortay, H., Cozzone, A. J., and Grangeasse, C. (2007) Influence of tyrosine-kinase Wzc activity on colanic acid production in *Escherichia coli* K12 cells. *J. Mol. Biol.* **367**, 42-53

54. Geisinger, E., and Isberg, R. R. (2015) Antibiotic modulation of capsular exopolysaccharide and virulence in *Acinetobacter baumannii*. *PLoS Pathog.* **11**, e1004691

55. Ferreira, A. S., Leitao, J. H., Sousa, S. A., Cosme, A. M., Sa-Correia, I., and Moreira, L. M. (2007) Functional analysis of *Burkholderia cepacia* genes *bceD* and *bceF*, encoding a phosphotyrosine phosphatase and a tyrosine autokinase, respectively: role in exopolysaccharide biosynthesis and biofilm formation. *Appl. Environ. Microbiol.* **73**, 524-534

56. Nakar, D., and Gutnick, D. L. (2003) Involvement of a protein tyrosine kinase in production of the polymeric bioemulsifier emulsan from the oil-degrading strain *Acinetobacter lwoffii* RAG-1. *J. Bacteriol.* **185**, 1001-1009

57. Elsholz, A. K., Wacker, S. A., and Losick, R. (2014) Self-regulation of exopolysaccharide production in *Bacillus subtilis* by a tyrosine kinase. *Genes Dev.* **28**, 1710-1720

58. Hajredini, F., and Ghose, R. (2022) conserved link between catalytic site interactions and global conformation in P-loop enzymes. *bioRxiv*, DOI: 2022.2007.2013.499785

59. Huse, M., and Kuriyan, J. (2002) The conformational plasticity of protein kinases. *Cell* **109**, 275-282

60. Gout, E., Rebeille, F., Douce, R., and Bligny, R. (2014) Interplay of Mg^{2+}, ADP, and ATP in the cytosol and mitochondria: unravelling the role of Mg^{2+} in cell respiration. *Proc. Natl. Acad. Sci. USA* **111**, E4560-4567

61. Matte, A., Tari, L. W., and Delbaere, L. T. (1998) How do kinases transfer phosphoryl groups? *Structure* **6**, 413-419

62. Mann, D., Teuber, C., Tennigkeit, S. A., Schroter, G., Gerwert, K., and Kotting, C. (2016) Mechanism of the intrinsic arginine finger in heterotrimeric G proteins. *Proc. Natl. Acad. Sci. USA* **113**, E8041-E8050

63. Takahashi, H., Umino, S., Miki, Y., Ishizuka, R., Maeda, S., Morita, A., Suzuki, M., and Matubayasi, N. (2017) 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

64. Kotting, C., Kallenbach, A., Suveyzdis, Y., Wittinghofer, A., and Gerwert, K. (2008) The GAP arginine finger movement into the catalytic site of Ras increases the activation entropy. *Proc. Natl. Acad. Sci. USA* **105**, 6260-6265

65. Komoriya, Y., Ariga, T., Iino, R., Imamura, H., Okuno, D., and Noji, H. (2012) Principal role of the arginine finger in rotary catalysis of F_1-ATPase. *J. Biol. Chem.* **287**, 15134-15142

66. Zhang, Y., Zagnitko, O., Rodionova, I., Osterman, A., and Godzik, A. (2011) The FGGY carbohydrate kinase family: insights into the evolution of functional specificities. *PLoS Comput. Biol.* **7**, e1002318

67. Khadka, S., Ring, B., Krzeminski, L. R., Hathaway, M., Walker, R. S., Mobley, H. L. T., and Mike, L. A. (2022) Regulation of *Klebsiella pneumoniae* mucoidy by the bacterial tyrosine kinase Wzc. *bioRxiv*, DOI: 2022.2006.2005.494587

68. Galvan, E. M., Ielmini, M. V., Patel, Y. N., Bianco, M. I., Franceschini, E. A., Schneider, J. C., and Ielpi, L. (2013) Xanthan chain length is modulated by increasing the availability of the polysaccharide copolymerase protein GumC and the outer membrane polysaccharide export protein GumB. *Glycobiol.* **23**, 259-272

69. Mace, P. D., and Murphy, J. M. (2021) There's more to death than life: Noncatalytic functions in kinase and pseudokinase signaling. *J. Biol. Chem.* **296**, 100705

70. Salazar, C., Brummer, A., Alberghina, L., and Hofer, T. (2010) Timing control in regulatory networks by multisite protein modifications. *Trends Cell Biol* **20**, 634-641

71. Rausch, M., Deisinger, J. P., Ulm, H., Muller, A., Li, W., Hardt, P., Wang, X., Li, X., Sylvester, M., Engeser, M., Vollmer, W., Muller, C. E., Sahl, H. G., Lee, J. C., and Schneider, T. (2019) Coordination of capsule assembly and cell wall biosynthesis in Staphylococcus aureus. *Nat. Commun.* **10**, 1404

72. Reid, A. N., and Whitfield, C. (2005) functional analysis of conserved gene products involved in assembly of *Escherichia coli* capsules and exopolysaccharides: evidence for

molecular recognition between Wza and Wzc for colanic acid biosynthesis. *J. Bacteriol.* **187**, 5470-5481

73. Drummelsmith, J., and Whitfield, C. (2000) Translocation of group 1 capsular polysaccharide to the surface of *Escherichia coli* requires a multimeric complex in the outer membrane. *EMBO J.* **19**, 57-66

74. Endres, N. F., Das, R., Smith, A. W., Arkhipov, A., Kovacs, E., Huang, Y., Pelton, J. G., Shan, Y., Shaw, D. E., Wemmer, D. E., Groves, J. T., and Kuriyan, J. (2013) Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543-556

75. Srinivasan, S., Regmi, R., Lin, X., Dreyer, C. A., Chen, X., Quinn, S. D., He, W., Coleman, M. A., Carraway, K. L., 3rd, Zhang, B., and Schlau-Cohen, G. S. (2022) Ligand-induced transmembrane conformational coupling in monomeric EGFR. *Nat. Commun.* **13**, 3709

76. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589

77. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Zidek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D., and Velankar, S. (2022) AlphaFold

Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439-D444

78. Arkhipov, A., Shan, Y., Das, R., Endres, N. F., Eastwood, M. P., Wemmer, D. E., Kuriyan, J., and Shaw, D. E. (2013) Architecture and membrane interactions of the EGF receptor. *Cell* **152**, 557-569

79. Cozzone, A. J. (2009) Bacterial tyrosine kinases: novel targets for antibacterial therapy? *Trends Microbiol.* **17**, 536-543

Table 1. Distribution of BY-kinases in the bacterial kingdom[a,b]

	P-type	F-type	Orphan CD with TMDs[c]	Orphan CD	Total
Gram-positive					
Actinomycetota (Actinobacteria)	1588	2	265	46	1901
Bacillota (Firmicutes)	**281**	**441**	**133**	**5025**	**5880**
Mycoplasmatota (Tenericutes)	-	-	1	39	40
Deinococcota (Deinococcus-thermus)	2	-	3	-	5
Gram-negative					
Abditibacteriota	2	-	1	-	3
Acidobacteriota (Acidobacteria)	424	-	74	238	736
Aquificota (Aquificae)	2	-	2	-	4
Armatimonadota (Armatimonadetes)	17	-	9	2	28
Atribacterota	1	-	-	2	3
Bacteroidota (Bacteroidetes)	792	3	237	20	1052
Balneolota (Balneolaeota)	13	-	6		19
Calditrichota (Calditrichaeota)	30	-	10	7	47
Chlorobiota (Chlorobi)	34	-	4	-	38
Chloroflexota (Chloroflexi)	126	2	29	69	226
Chrysiogenota (Chrysiogenetes)		-	-	1	1
Cyanobacteria	537	-	75	4	616
Elusimicrobiota (Elusimicrobia)	19	-	4	2	25
Euryarchaeota (Euryarchaeota)	2	-	-	4	6
Fibrobacterota (Fibrobacteres)	82	-	9	1	92
Fusobacteriota (Fusobacteria)	-	3	-	14	17
Gemmatimonadota (Gemmatimonadetes)	115	-	60	14	189
Ignavibacteriota (Ignavibacteriae)	46	-	20	-	66
Kiritimatiellota (Kiritimatiellaeota)	12	-	3	-	15
Lentisphaerota (Lentisphaerae)	19	-	12	1	32
Nitrospirota (Nitrospirae)	34	-	9	24	67
Planctomycetota (Planctomycetes)	256	-	215	21	492
Pseudomonadota (Proteobacteria)	**10741**	28	422	592	11783
Rhodothermota (Rhodothermaeota)	7	-	3	-	10
Spirochaetota (Spirochaetes)	22	-	-	2	24
Thermodesulfobacteriota (Thermodesulfobacteria)	-	-	-	3	3
Thermotogota (Thermotogae)	41	-	1	-	42
Verrucomicrobiota (Verrucomicrobia)	308		236	11	555
Unclassified/Unknown	286	1	79	173	539
Total	15839	480	1922	6315	24556

[a]Based on sequence dataset in BYKdb release 67.0 (8).
[b]Phyla are named according to the International Code of Nomenclature for Prokaryotes (ICNP), historical synonyms are shown in brackets.
[c]Transmembrane segments predicted by DeepTMHMM (biolib.com/DTU/DeepTMHMM/).

Table 2. Influence of key mutations on the kinase and ATPase activities of select BY-kinases

BY-kinase	Mutations	ATPase activity	Kinase activity	Reference
B. subtilis YwqD	D81A, D83A (Walker-A')	Greatly reduced	Abolished	Mijakovic (45)
	Y225F/Y227F/Y228F (Y-cluster)	Enhanced	N/A	Mijakovic (45)
S. meliloti ExoP	G588V, G588E, K589I (Walker-A)	Abolished	Abolished	Niemeyer (44)
E. coli K-12 Wzc	D562A, D564A (Walker-A')	Reduced	Reduced	Soulat (43)
	Y569N, Y569F	Reduced	Reduced	Soulat (43)
	K540M (Walker-A)	Abolished[a]	Reduced[b]	[a]Hajredini (23) [b]Bechet (32)

Table 3. Effects of BY-kinase mutations on observed phenotype

BY-kinase	Mutation	Kinase activity	Phenotype	Proposed to stabilize	Reference
V. cholerae VpsO	K551A (Walker-A)	Abolished	Rougher colonies, higher EPS levels	Σ_O	Schwechheimer (13)
	VpsU C12S (Inactive cognate LMW-PTP)	-	Smooth colonies, lower EPS levels	Σ_M	
	E519A/R522A/R525A (O-motif)	Reduced	Smooth colonies, lower EPS levels	Σ_M	
	Y720F/Y721F/Y726F/Y727F (Y-cluster)	-	Smooth colonies, lower EPS levels	Σ_M	
E. coli K-12 Wzc	Y708E/Y710E/Y11E/Y713E/Y715E (Y-cluster)	-	No EPS	Σ_M	Obadia (53)
	ΔY-cluster	-	No EPS	Σ_M	
	K540M (Walker-A)	Abolished	No EPS	Σ_O	
	Y708F/Y710F/Y11F/Y713F/Y715F (Y-cluster)	-	Uniform EPS size distribution ~1350 kDa (600-2000 kDa in WT)	Σ_O	
	Y569F	Reduced	Uniform EPS size distribution ~1350 kDa (600-2000 kDa in WT	Σ_O	
S. pneumoniae CpsD	Y215F/Y218F/Y221F (Y-cluster)	-	Decreased levels of LMW EPS, no effect on total abundance	Σ_O	Nakamoto (15)
	Y215E/Y218E/Y221E (Y-cluster)	-	Increased levels of LMW EPS no effect on total abundance	Σ_M	
A. baumannii Wzc	Y712F/Y714F/Y716F/Y718F/Y720F/Y722F (Y-cluster)	-	Increased levels of HMW CPS	Σ_O	Geisinger (54)
	D569N (Walker-A')	Reduced	Increased levels of HMW CPS	Σ_O	
	K547Q (Walker-A)	Not affected	Increased levels of HMW CPS	Σ_O	
	D649N (Walker-B)	Not affected	Increased levels of HMW CPS	Σ_O	
B. cepacia BceF	ΔBceD (Deletion of cognate LMW-PTP)	-	Reduced viscosity suggestive of lower MW EPS	Σ_M	Ferreira (55)

A. lwoffii Wzc	Y712F/Y714F/Y716F/Y718F/Y720F (Y-cluster)	-	Increased mucoidy, Enhanced levels of higher MW EPS	Σ_O	Nakar (56)
B. subtilis EpsB	Y225F/Y227F (Y-cluster)	-	Hyper-wrinkled colonies	Σ_O	Esholz (57)
	Y225E/Y227E (Y-cluster)	-	Flat colonies	Σ_M	

Figure Legends

Fig. 1: Structural organization of a functional BY-kinase. (A) Structure of full-length *E. coli* K30 Wzc (K540M mutant; PDB: 7NHR; a P-type BY-kinase) showing the periplasmic domain (PD), the two-pass transmembrane domain (TMD) that is embedded in the bacterial inner membrane, and the cytoplasmic catalytic domain (CD). The TMD and the PD together form the transactivation domain (TAD). The TAD and CD are separate proteins in F-type BY-kinases. The middle inset shows a representative monomer from the octameric assembly indicating the N-terminal transmembrane helix (TM1) that leads to the PD followed by a second transmembrane helix (TM2) that terminates in the CD. The top right inset shows a view of the CD octamer looking up from the cytoplasm into the periplasmic space; the TAD is hidden for ease in visualization. The C-terminal Y-cluster is colored red and the constituent tyrosine residues are shown in stick representation. Shown on the bottom right panel is a representative trimer extracted from the octameric assembly that illustrates the enzyme-substrate dual role of each CD. A tyrosine from the Y-cluster of monomer *a* is inserted into the active site of monomer *b*, which in turn inserts a tyrosine from its own Y-cluster into the active site of monomer *c*. Thus, *b* can simultaneously act both as an enzyme (with respect to *a*) and as a substrate (with respect to *c*). **(B)** Interactions at the inter-monomer interface of the CD of *E. coli* K-12 Wzc (K540M mutant; PDB: 3LA6) involving conserved O-motif residues E508, R511 and R514 on α2 (yellow) at the the first interface (I_1) with residues on α7 (light blue) and α9 (dark blue) that form the second interface (I_2). The two interacting monomers (designated *a* and *b*) are colored pink and purple. Key interface residues are shown as spheres. **(C)** The interaction of the *E. coli* K-12 Wzc CD with its cognate phosphatase, the LMW-PTP Wzb (teal); the latter interaction also involves α2 and the O-motif therein. Selected interactions at the interface are illustrated with the residues involved (Wzb residues are indicated

using 3-letter codes) shown as spheres. **(D)** The characteristic P-loop fold of Wzc CD is shown indicating the conserved motifs (key sidechains are shown in stick representation; the motifs they belong to are indicated by the corresponding subscripts): Walker-A (^{533}**G**VSPSI**GKT**^{541}, teal; conserved residues are indicated in bold font), Walker-A' (^{558}VLLI**D**C**D**M**R**^{566}, orange) and Walker-B (^{639}VLI**D**TP**P**^{645}, purple). Also indicated are α2, α7 and α9, colored as in (B) above. The Y-cluster tyrosine residues (Y705, Y708, Y710, Y711, Y713 and Y715) are also indicated (in red as sticks). Y705 (indicated by the '*') is not autophosphorylated, nor can it optimally access the Wzb active site. A substrate tyrosine residue ($Y715_S$) from the Y-cluster of a flanking CD docked at the active site is also shown.

Fig. 2: A simple model for BY-kinase function. **(A)** Schematic representation of the effects of autophosphorylation and dephosphorylation (red line) on the temporal distribution of the octameric (Σ_O) and monomeric (Σ_M) species (dark blue line). The left ordinate indicates the fraction of phosphorylated Y-cluster tyrosine residues; the right ordinate indicates the oligomer order. Autophosphorylation occurs at a given rate within the octameric state until the Y-cluster phosphorylation levels exceed a critical threshold (C_{phos}, dashed orange line), following which the octameric ring dissociates into its constituent monomers. The phosphatase (light green) can then engage the kinase, and perform multiple rounds of Y-cluster dephosphorylation, following which the oligomer reassembles and the cycle repeats (two cycles are shown). Assuming fixed rates of autophosphorylation and dephosphorylation (in reality, these rates are likely to be non-linear functions of the level of Y-cluster phosphorylation), the system cycles between Σ_O and Σ_M; the corresponding lifetimes of T_O and T_M, respectively define the periodicity of the process. **(B)** Increasing the lifetime of Σ_O (relative to the lifetime of Σ_M) i.e., enhancement of T_O (relative to T_M) shifts the distribution towards high molecular weight polysaccharides (HMW) and increased

mucoidy (blurred circles in the inset); a larger T_M (relative to T_O) results in increased production of low molecular weight polysaccharides (LMW) and smoother colonies (sharp circles in the inset).

Fig. 3: Coupling between global and local conformations in the Wzc CD. (A) Global conformational changes that drive the transition from the closed state (CS, cyan) to the open state (OS, lime) are represented by the red arrows. **(B)** Definition of the 2-dimensional frame of reference that parses the OS and the CS. P1: center of mass (COM, indicated by colored spheres in all cases) of the Cα atoms of the first two turns of the N-terminus of α3 (541-547; *E. coli* K-12 Wzc numbering, brown); P2: COM of Cα atoms of the penultimate turn (548-552, red) from the C-terminus of α3; P3: COM of Cα atoms of the 569-573 segment (pink) on α4; P4: COM of the Cα atoms of the first turn of the N-terminus of α2 (505-509, green). The structures are rotated and translated to place P1 and P2 along the Cartesian z–axis, then rotated about the z-axis to place P3 on the x–axis in the positive direction. The $|h|$ coordinate is then defined as the absolute value of the z-projection of P1; θ is defined as the angle between the x–axis and the xy–projection of a vector directed from P1 to P4. **(C)** Probability density plots of the conformational ensembles obtained through enhanced sampling MD simulations of apo Wzc CD (top left), and the corresponding ATP•Mg^{2+} (top right), ADP•Mg^{2+} (bottom left), and ADP (bottom right) complexes in θ–$|h|$ space. The green circle represents the position occupied by the crystallographic monomers. Regions corresponding to the OS and CS are indicated. An additional "hyper-closed state" (HS), that contains structural features characteristic of a fully active state, sampled only in the ATP•Mg^{2+} complex, is also shown. **(D)** Probability density plots of the structural ensembles of the K540M mutant of Wzc CD in its unliganded state (top left), and its ATP•Mg^{2+} (top right) complex, are represented in θ–$|h|$ space; the plots for the corresponding wild-type (WT) cases are shown as

green contours with the positions of key states indicated. The lower plot shows gel filtration traces of the untreated mutant CD and that incubated with excess Mg^{2+}. **(E)** OS/CS transitions are correlated with local rearrangements at the active site. In the CS, the Walker-B D642 forms a hydrogen bond (lime dashed lines) with the Walker-A T541 enabling coordination of the catalytically essential Mg^{2+} ion (not shown for visual ease). This interaction is broken in the OS with the Walker-A K540 establishing interactions with the D642 sidechain and the T643 backbone in a manner that neither allows optimal coordination of Mg^{2+} nor the availability of K540 for chemistry. Parts of this figure were adapted from Hajredini *et al.* (22,23);

Fig. 4: Structural destabilization in the CS of monomeric Wzc CD. (A) Major conformations selected from within the CS ensemble obtained through enhanced sampling MD simulations of monomeric Wzc CD bound to ATP•Mg^{2+}. These structures are characterized by distortions of the two oligomerization interfaces including a significant bending (red arrows) of α2 (yellow; I_1; left panel), or a partial unfolding (red circles) of α7 (light blue; I_2) and/or α9 (dark blue; I_2; right panel). ATP•Mg^{2+} is not shown for ease in visualization. **(B)** Classical MD simulations of the ATP•Mg^{2+} complex of Wzc CD (red) as part of a trimeric "pseudo-ring" (used *in lieu* of the octamer for computational convenience) where positional restraints are applied to the flanking monomers (cyan) to mimic the octameric ring (top left) or in its monomeric state (top right). The corresponding projections of structures from representative 100 ns classical MD simulations onto θ–|*h*| space are shown on the lower panels; the points are color coded according to the simulation timestamps. The black circle indicates the reference value in the crystal structure of the octamer. The CS is stabilized in the pseudo-ring preventing transition to the OS. **(C)** The CS for the Wzc

CD in the monomeric state may be stabilized through the introduction of a 4-glycine sequence (G_4) at the C-terminus of the α2 away from the O-motif (G_4-construct). The structural ensemble of the ATP•Mg^{2+} complex (top) of the G_4-construct obtained through enhanced sampling MD simulations populates a stable CS in which the constituent structures display intact oligomerization interfaces free from distortions in α2, α7 or α9. The disordered G_4 region at the C-terminus of α2 is indicated by the orange circle; the positions of the conserved O-motif residues are also indicated in green for reference. The bottom panel shows the probability density plot of the G_4-construct•ATP•Mg^{2+} complex projected onto in θ-$|h|$ space. The structure seen *in crystallo* (representative of the CS) is shown as a green circle. Parts of this figure were adapted with permission from Hajredini *et al.* (22,23); © the authors, some rights reserved; exclusive licensee AAAS. Distributed under a Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC).

Fig. 5: Global and local features in the regulation of Wzc activity. (A) In clockwise fashion starting from the point indicated by the red '*'. In the oligomer (Σ_O), individual Wzc CD monomers (indicated by the green outlines) with dephosphorylated Y-clusters (cyan sticks) exist in the CS. As shown in the inset, the two oligomerization interfaces, I_1 (α2 in yellow) and I_2 (α7 and α9 in light and dark blue, respectively), are stabilized through inter-monomer interactions though some disorder persists (red arcs) at the C-terminus of α2 away from the O-motif. The CS allows formation of the characteristic $D642_B$-$T541_A$ hydrogen bond (see Fig. 3E) and thereby the stable engagement of ATP•Mg^{2+} (orange) that displaces the bound ADP (purple) from the active site. Binding of ATP•Mg^{2+}, multiple rounds of Y-cluster phosphorylation (indicated by progressively darker shades of orange), and nucleotide exchange within Σ_O ultimately results in the C_{phos} threshold being exceeded and disassembly of the oligomer into its constituent monomers (Σ_M).

These monomers exist in the OS (monomers shown with dark blue outlines) bound to ADP but unable to engage Mg^{2+} (and therefore ATP•Mg^{2+}). While a CS (CS_M, shown in the inset) may be forced within Σ_M by the presence of ATP•Mg^{2+}, such a state shows instability in I_1 (through bending of α2; green arrows) or I_2 (through partial unfolding of the α7 and/or α9). The RK-cluster (in brown) is disordered in Σ_M; its ordering for a stabilized CS facilitates the formation of a reactive state. In Σ_M, stable engagement with Wzb (shown in green) through I_1 initiates Y-cluster dephosphorylation. Σ_O formation can occur after several rounds of Wzb-mediated dephosphorylation, and the functional cycle can be reinitiated. It is likely that the assembly/disassembly process also involves partially assembled/disassembled states (dashed boxes) representing a deviation from the simple 2-state scheme presented in Fig. 2. Some individual monomers in such cases also likely contain features that deviate (pink outlines) from those that are characteristic of the OS and CS, perhaps containing features of both. **(B)** Engagement within the CS facilitates the formation of a reactive conformation of ATP. ATP is shown as a Newman projection with its α–, β–, and γ–phosphate moieties indicated by blue, yellow, and red lines, respectively. The Walker-A lysine ($K540_A$, scarlet sphere) and the catalytic Mg^{2+} ion (green circle), lock the β– and γ–phosphates within the same plane. Rotation of the α–phosphate is enabled through its engagement with Y569 thereby facilitating formation of a high-energy all-eclipsed conformation. The ordering of the RK-cluster and insertion of R490 (cyan circle) therein, together with Y569, and a substrate tyrosine (from the Y-cluster of a flanking monomer; red), into the active site facilitate de-solvation (dark blue dashes) providing favorable free energy contributions to overcome the activation barrier for chemistry. The precise sequence of the conformational events mentioned that lead to the reactive state remains unclear. Parts of this figure were adapted with permission from Hajredini *et al.* (22);

Fig. 6: Dynamic regions as hotspots for adaptation in BY-kinases. Clustering of 796 BY-kinase CD sequences results in three large groups – A (labeled in magenta), B (brown), and C (grey). Group A comprises 3 sub-groups – A1 (mostly Actinobacteria, dominant constituent phyla indicated in all cases using historical phyla names), A2 (Cyanobacteria) and A3 (Firmicutes). Group B comprises 2 sub-groups – B1 (Gammaproteobacteria) and B2 (Betaproteobacteria). Group C comprises 9 sub-groups – C1 (Deltaproteobacteria), C2 (Alphaproteobacteria), C3 (Alphaproteobacteria), C4 (Gammaproteobacteria), C5 (Alphaproteobacteria), C6 (Firmicutes), C7 (Bacteroidetes), C8 (Actinobacteria) and C9 (Bacteroidetes). Each of these groupings display specific signature motifs (strongly conserved positions are indicated by black and grey bars; conserved motifs are indicated by the red, dark blue and purple boxes) beyond the universally conserved Walker motifs (indicated by the colored bars; Walker-A: pink, Walker-A': orange, Walker-B: blue). Key features of some of these motifs are conserved across some groups but not others. Some of these motifs overlap with dynamic regions e.g., α7 (purple; similar in A1, C4 and C6) and α9 (dark blue; similar in A1 and C4). The corresponding sequence logos are indicated for each case, and that from the *E. coli* (K-12) Wzc sequence is shown below; residues with high level of conservation are underlined. Parts of this figure have been modified with permission from Shi *et al.* (31).

Figure 1

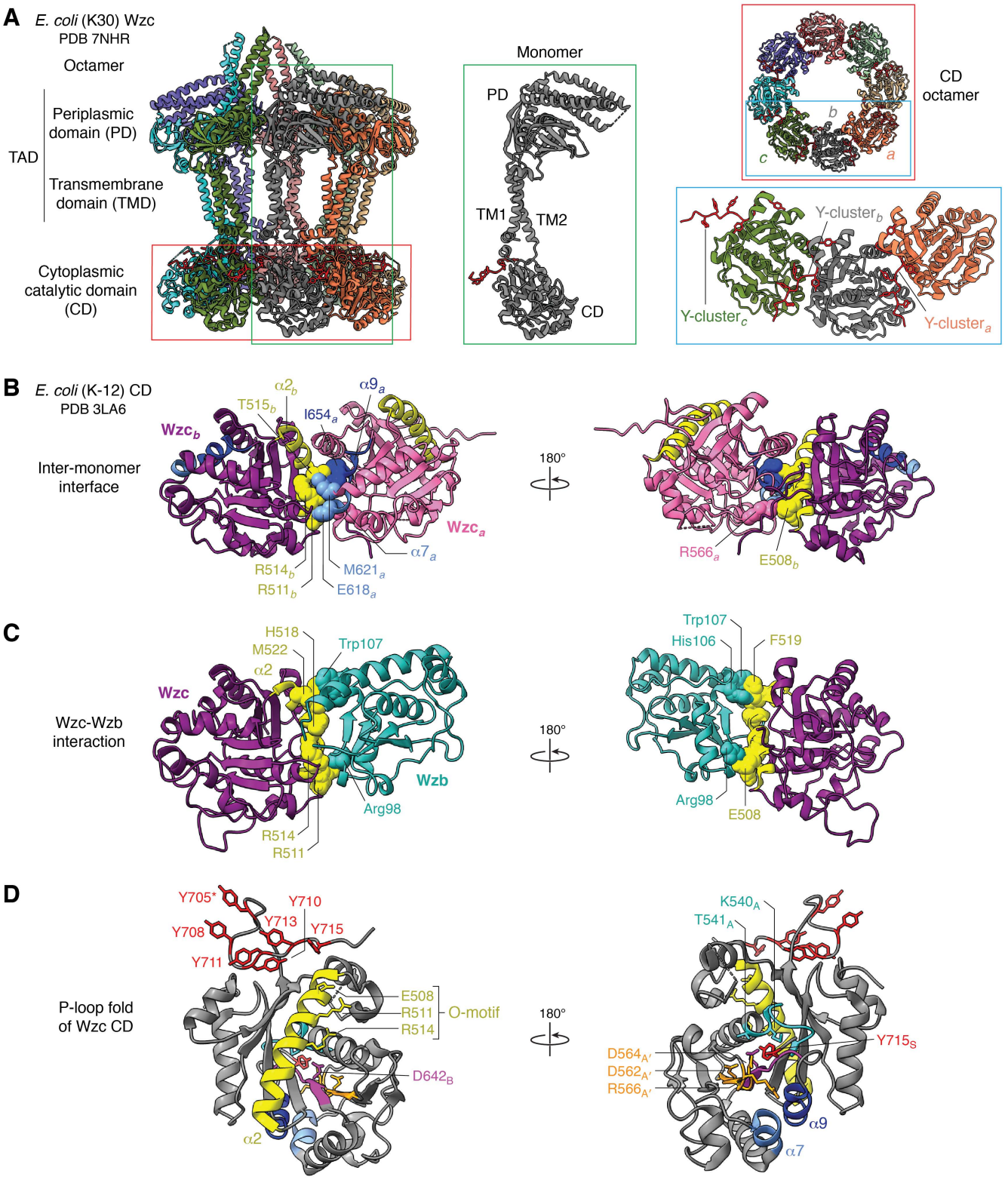

A
E. coli (K30) Wzc
PDB 7NHR
Octamer
TAD
Periplasmic domain (PD)
Transmembrane domain (TMD)
Cytoplasmic catalytic domain (CD)
Monomer
PD
TM1
TM2
CD
CD octamer
b
c
a
Y-cluster$_b$
Y-cluster$_c$
Y-cluster$_a$
B
E. coli (K-12) CD
PDB 3LA6
Inter-monomer interface
$\alpha2_b$
$\alpha9_a$
T515$_b$
I654$_a$
Wzc$_b$
Wzc$_a$
$\alpha7_a$
R514$_b$
R511$_b$
M621$_a$
E618$_a$
180°
R566$_a$
E508$_b$
C
Wzc-Wzb interaction
H518
M522
$\alpha2$
Trp107
Wzc
Wzb
Arg98
R514
R511
180°
Trp107
His106
F519
Arg98
E508
D
P-loop fold of Wzc CD
Y705*
Y710
Y708
Y713
Y715
Y711
E508
R511
R514
O-motif
D642$_B$
$\alpha2$
180°
K540$_A$
T541$_A$
Y715$_S$
D564$_{A'}$
D562$_{A'}$
R566$_{A'}$
$\alpha9$
$\alpha7$

Figure 2

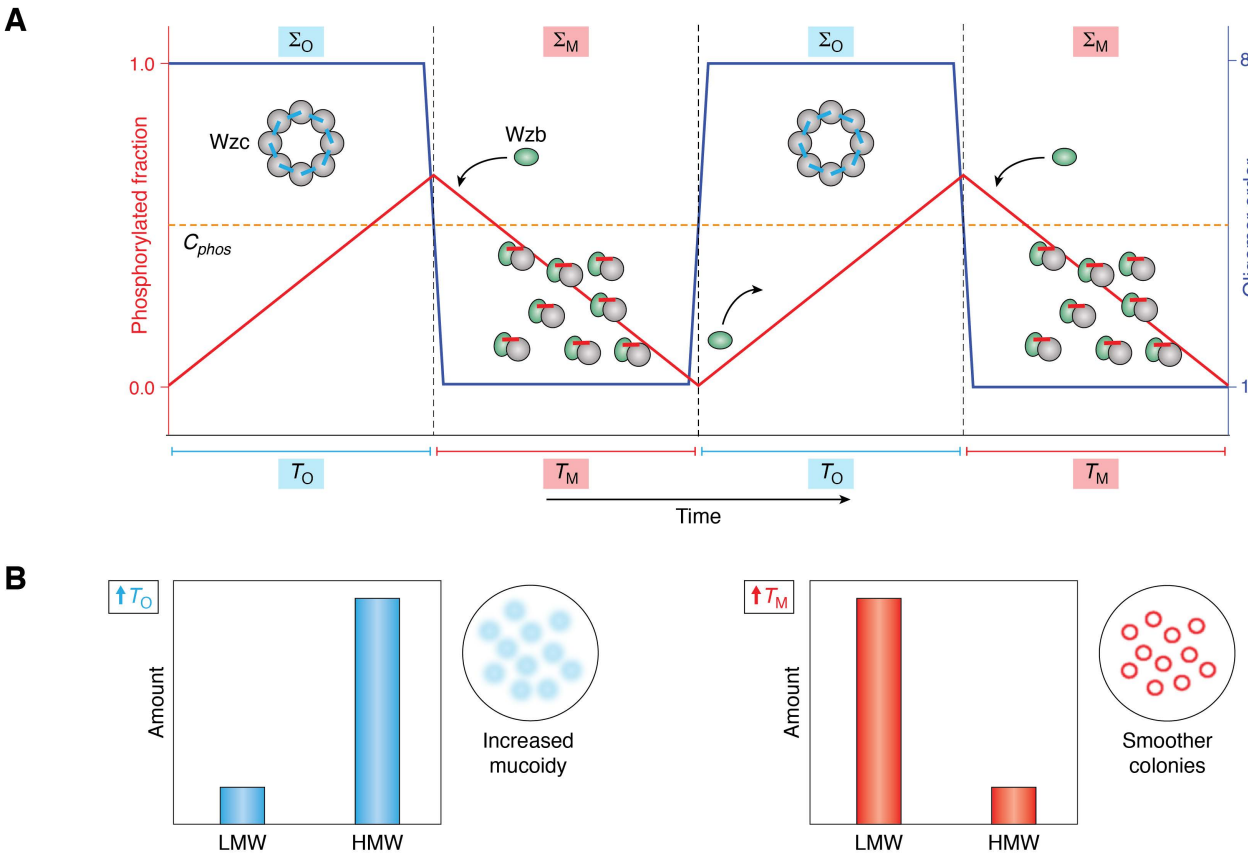

A
Σ_O
Σ_M
Σ_O
Σ_M
1.0
0.0
8
1
Phosphorylated fraction
Oligomer order
Wzc
Wzb
C_{phos}
T_O
T_M
T_O
T_M
Time
B
$\uparrow T_O$
Amount
LMW
HMW
Increased mucoidy
$\uparrow T_M$
Amount
LMW
HMW
Smoother colonies

Figure 3

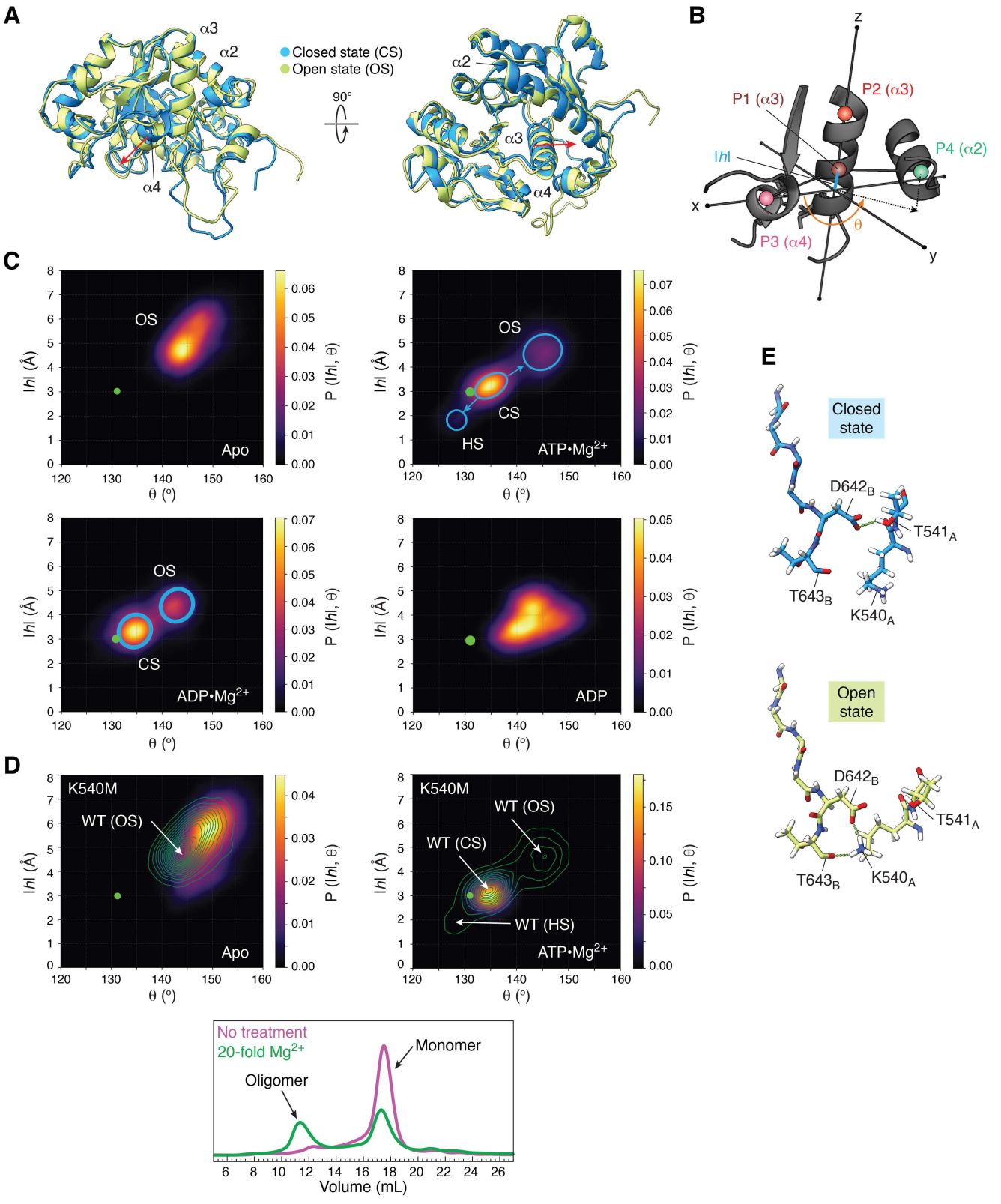

A
α3
α2
α4
Closed state (CS)
Open state (OS)
90°
α2
α3
α4
B
z
P1 (α3)
P2 (α3)
P4 (α2)
|h|
x
θ
P3 (α4)
y
C
OS
Apo
OS
CS
HS
ATP•Mg2+
OS
CS
ADP•Mg2+
ADP
|h| (Å)
θ (°)
P (|h|, θ)
D
K540M
WT (OS)
Apo
K540M
WT (OS)
WT (CS)
WT (HS)
ATP•Mg2+
E
Closed state
D642B
T541A
T643B
K540A
Open state
D642B
T541A
T643B
K540A
No treatment
20-fold Mg2+
Monomer
Oligomer
Volume (mL)

Figure 4

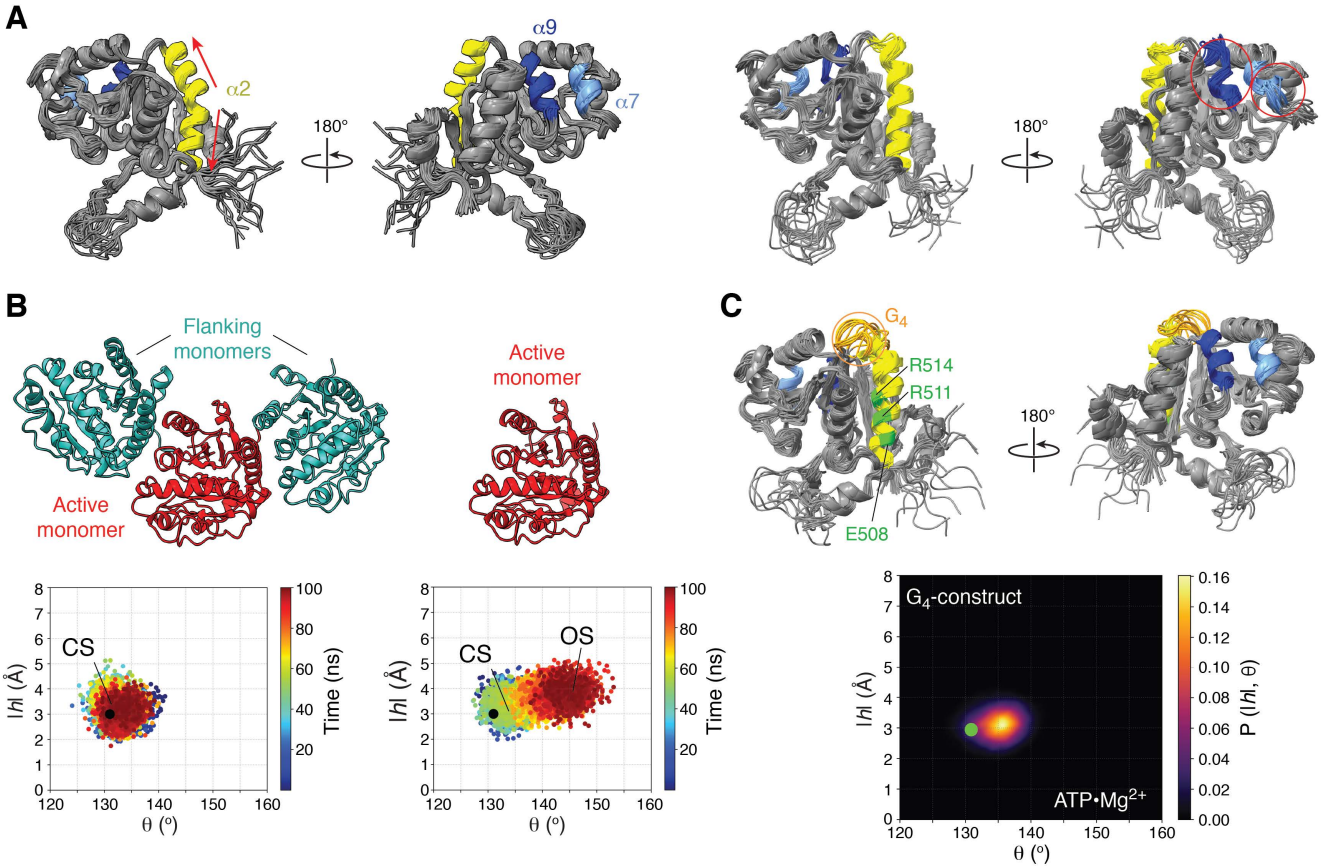

A
α2
α9
α7
180°
180°
B
Flanking monomers
Active monomer
Active monomer
C
G4
R514
R511
E508
180°
CS
OS
|h| (Å)
θ (°)
Time (ns)
G4-construct
ATP•Mg2+
P (|h|, θ)

Figure 5

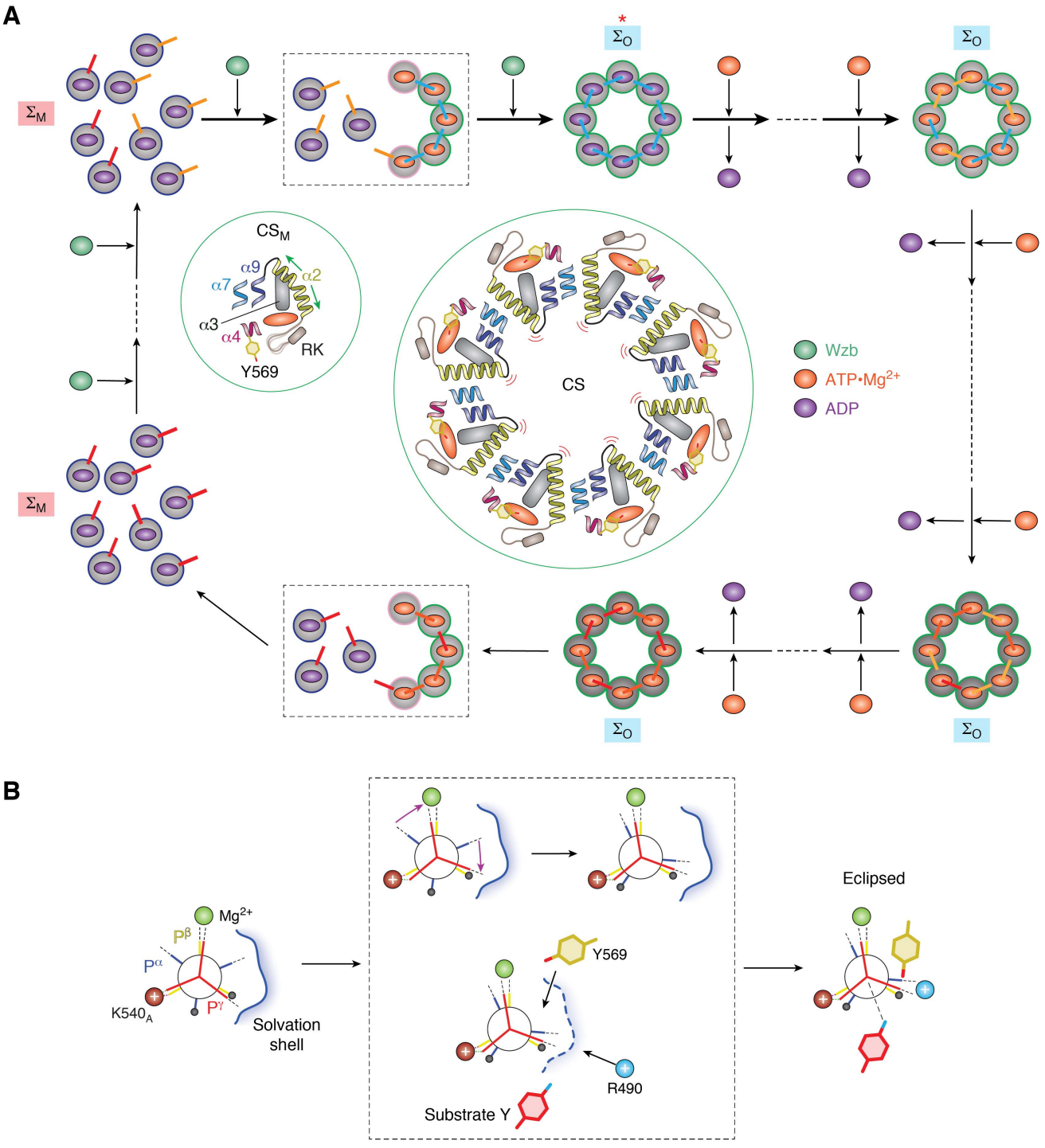
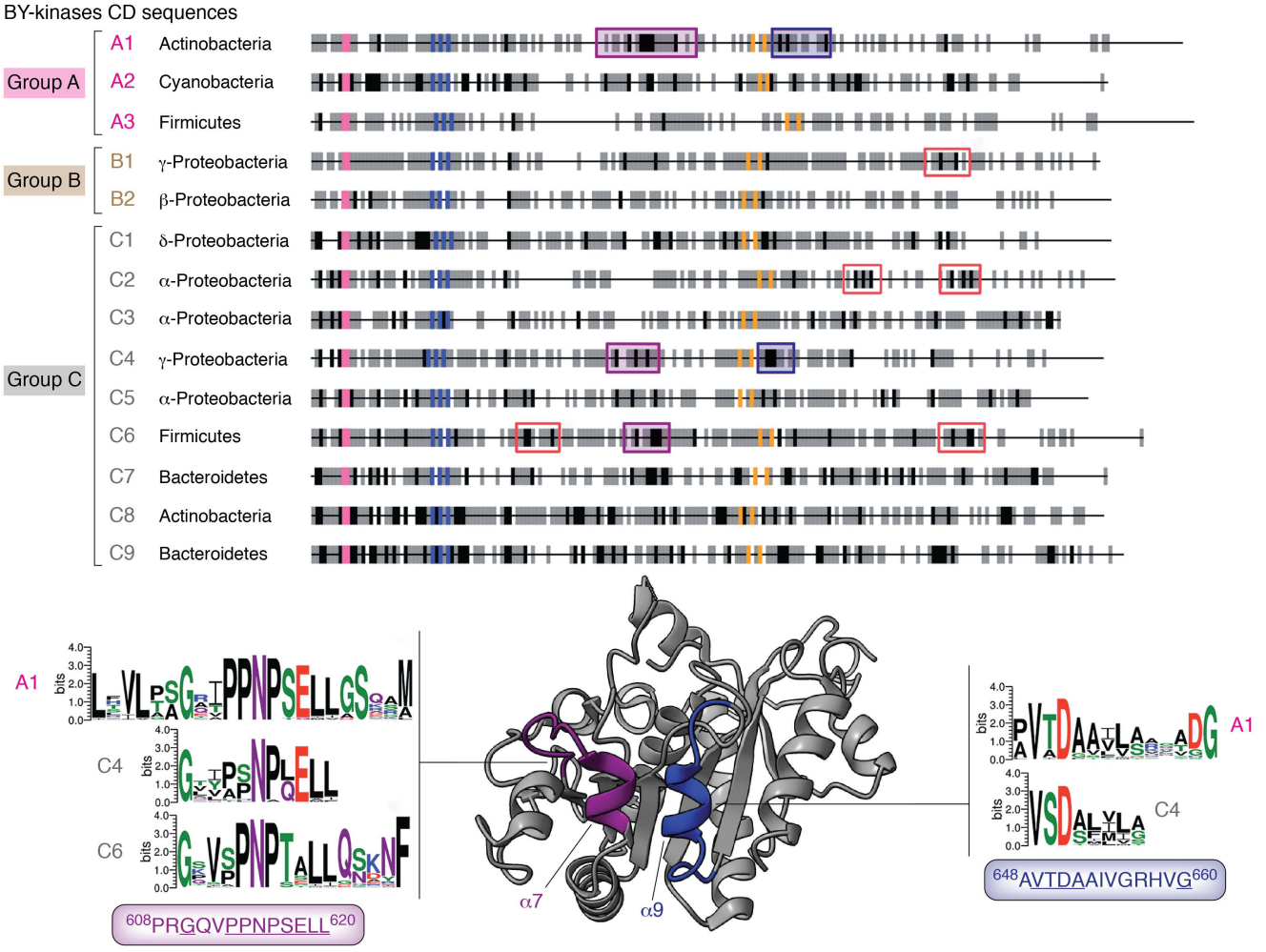

A
Σ_M
Σ_O
*
CS_M
α9
α2
α7
α3
α4
RK
Y569
CS
Wzb
ATP•Mg^{2+}
ADP
B
Mg^{2+}
Pβ
Pα
Pγ
$K540_A$
Solvation shell
Y569
R490
Substrate Y
Eclipsed

Figure 6


BY-kinases CD sequences
Group A
A1 Actinobacteria
A2 Cyanobacteria
A3 Firmicutes
Group B
B1 γ-Proteobacteria
B2 β-Proteobacteria
Group C
C1 δ-Proteobacteria
C2 α-Proteobacteria
C3 α-Proteobacteria
C4 γ-Proteobacteria
C5 α-Proteobacteria
C6 Firmicutes
C7 Bacteroidetes
C8 Actinobacteria
C9 Bacteroidetes
A1
C4
C6
bits
608PRGQVPPNPSELL620
α7
α9
A1
C4
bits
648AVTDAAIVGRHVG660