THE PROTEIN SOCIETY

Conserved residues of the Pro103–Arg115 loop are involved in triggering the allosteric response of the *Escherichia coli* ADP-glucose pyrophosphorylase

Benjamin L. Hill, Jennifer Wong, Brian M. May, Fidel B. Huerta, Tara E. Manley, Peter R.F. Sullivan, Kenneth W. Olsen, and Miguel A. Ballicora*

Department of Chemistry and Biochemistry, Loyola University Chicago, 1068 W Sheridan Road, Chicago, Illinois

Received 26 November 2014; Accepted 20 January 2015
DOI: 10.1002/pro.2644
Published online 23 January 2015 proteinscience.org

Abstract: The synthesis of glycogen in bacteria and starch in plants is allosterically controlled by the production of ADP-glucose by ADP-glucose pyrophosphorylase. Using computational studies, site-directed mutagenesis, and kinetic characterization, we found a critical region for transmitting the allosteric signal in the *Escherichia coli* ADP-glucose pyrophosphorylase. Molecular dynamics simulations and structural comparisons with other ADP-glucose pyrophosphorylases provided information to hypothesize that a Pro103–Arg115 loop is part of an activation path. It had strongly correlated movements with regions of the enzyme associated with regulation and ATP binding, and a network analysis showed that the optimal network pathways linking ATP and the activator binding Lys39 mainly involved residues of this loop. This hypothesis was biochemically tested by mutagenesis. We found that several alanine mutants of the Pro103–Arg115 loop had altered activation profiles for fructose-1,6-bisphosphate. Mutants P103A, Q106A, R107A, W113A, Y114A, and R115A had the most altered kinetic profiles, primarily characterized by a lack of response to fructose-1,6-bisphosphate. This loop is a distinct insertional element present only in allosterically regulated sugar nucleotide pyrophosphorylases that could have been acquired to build a triggering mechanism to link proto-allosteric and catalytic sites.

Keywords: alanine scanning mutagenesis; activation signal propagation; glycogen/starch metabolism; correlated movement analysis; network pathway analysis; molecular dynamics

Abbreviations: ADP-Glc PPase, ADP-Glucose pyrophosphorylase; ADP-Glc, Adenosine-5′-diphosphoglucose; AM, ammonium molybdate; AMP, Adenosine-5′-monophosphate; BSA, bovine serum albumin; EDTA, ethylenediaminetetraacetic acid; FBP, fructose-1,6-bisphosphate; Glc1P, glucose-1-phosphate; HEPES, 4-(2-hydroxyethyl)−1-piperazineethanesulfonic acid; HEPPS, 3-[4-(2-Hydroxyethyl)−1-piperazinyl]propanesulfonic acid; LB, Luria broth; MD, molecular dynamics; MG, malachite green; NDP-Glc PPase, NDP-Glucose pyrophosphorylase; NDP-Glc, nucleotide-5′-diphosphoglucose.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: National Science Foundation; Grant number: MCB 1024945 (to M.A.B.); Grant sponsor: National Science Foundation; Grant number: CHE-1004430 (to K.W.O.).

*Correspondence to: Miguel A. Ballicora, Department of Chemistry and Biochemistry, Loyola University Chicago, 1068 W Sheridan Rd, Chicago, Illinois. E-mail: mballic@luc.edu

Bacteria and plants produce glycogen and starch, respectively, as their primary storage polysaccharides. The synthesis of these polymers occurs with the donation of a glucose moiety from ADP-glucose (ADP-Glc) to an existing α-1,4-polyglucan chain. It is the production of ADP-Glc, catalyzed by the enzyme ADP-glucose pyrophosphorylase (ATP:α-D-glucose-1-phosphate adenylyltransferase; EC 2.7.7.27; ADP-Glc PPase), that acts as the rate limiting step in glycogen and starch production.[1] This Mg^{2+} dependent reaction proceeds with ATP binding first to the enzyme, followed by glucose-1-phosphate (Glc1P), with the subsequent production of ADP-Glc and pyrophosphate.[2,3]

All known plant and bacterial ADP-Glc PPases are derived from a common ancestor and are

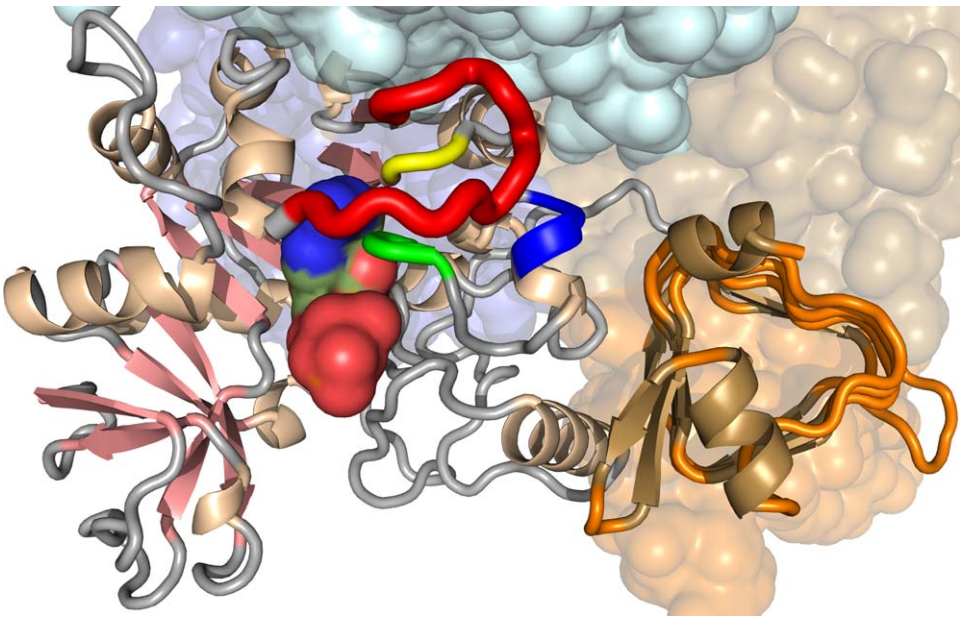

Figure 1. Model of the *E. coli* ADP-Glc PPase subunit, shown in cartoon representation. For reference, the Pro103–Arg115 loop is colored red, the C-terminal domain is in orange, Asn38–Ala40 are colored blue, Thr73–Gln75 are yellow, Leu25–Gly30 are in green, and bound ATP is shown in surface format. Additional subunits of this tetramer can also be seen in surface format. This image was generated using Pymol 1.3 and shows the enzyme 3 ns into the MD simulation.

predicted to share considerable structural similarities.[4,5] These structural features are evident in the crystallographic structures of the potato tuber[6] and *Agrobacterium tumefaciens*[7] enzymes. Notably, the enzymes are tetramers with ~50 kDa subunits. Enteric and cyanobacterial enzymes are homotetrameric (α_4), whereas those of Firmicutes, unicellular algae, and higher plants are heterotetrameric ($\alpha_2\beta_2$).[4,5] In the latter, the two subunit types have similar structure and share a common ancestry, although it is sometimes the case that just one subunit type is catalytic.[8,9] In most organisms, this enzyme is allosterically regulated by intermediates of the main carbon assimilatory pathway that exist within that species.[4,5,10] In the case of *E. coli*, glycolysis is the major pathway and fructose-1,6-bisphosphate (FBP) is the primary activator of the *E. coli* ADP-Glc PPase.[4] Of the ADP-Glc PPases that have been characterized, the *E. coli* enzyme has some of the largest changes in basal activity brought about by the activator. For instance, it increases specific activity and ATP affinity by at least 25- and 15-fold, respectively.[11,12]

Previous mutagenesis studies involving ADP-Glc PPases have provided some information regarding the allosteric activation mechanism of these enzymes. For example, studies on chimeric ADP-Glc PPases have shown that the C-terminal domain is important for allosteric effector specificity.[12] Some alanine mutants of N-terminal arginine residues in the *A. tumefaciens* enzyme had altered allosteric responses, such as R32A which had significantly reduced affinity for the activator fructose-6-P.[13] Several N-terminal residues in the *E. coli* ADP-Glc PPase have also been associated with allosteric activation. Lys39 is part of the FBP binding site,[14] and Gln74, which is universally conserved amongst ADP-Glc PPases, is needed in order for FBP to exhibit regulatory role.[11] The alanine mutant of the nearly universally conserved Trp113 had a similar kinetic profile to Q74A, even though FBP still binds to these enzymes.[11] W116A and Q75A mutants of the potato tuber enzyme also fail to respond to allosteric activator.[15] A model of the *E. coli* ADP-Glc PPase that highlights these regions is shown in Figure 1. Despite all of these data, it is not known how the allosteric signal is transmitted and what specific interactions are involved.

Here, we use molecular dynamics (MD) to get insight into what residues are involved in an allosteric pathway in the *E. coli* ADP-Glc PPase. Allosteric effectors may alter communication pathways that exist within an enzyme.[16] For that reason, correlated movement analysis, which examines cooperative motions within a protein, may be a useful tool for elucidating such interactions.[17] However, because this technique does not consider the physical proximity of correlated regions, further analysis may be needed to identify spatially connected communication pathways. To that end, a network pathway analysis, which accounts for the physical connectivity of correlated regions, can help pinpoint such tracts.[18] These techniques, along with an examination of some hydrogen bond networks and structural comparisons with other ADP-Glc PPases, suggested that the span of residues from Pro103 to Arg115 is

important in an allosteric activation pathway in the *E. coli* enzyme. We tested this hypothesis by performing alanine scanning mutagenesis on the residues of this loop and kinetically characterizing the mutants, several of which were distinguished by a disruption of the allosteric response. In addition to this, sequence alignments of known ADP-Glc PPases that focused on this loop region, as well as structural alignments of different bacterial NDP-glucose pyrophosphorylases (NDP-Glc PPases), provided insight into the evolution of allosteric regulation in ADP-Glc PPases.

Results

MD simulations and correlation analysis

The identification of regions of an enzyme associated with an allosteric pathway can be deduced by coupling an evolutionary analysis of the protein (i.e., the conservancy of involved residues) with MD simulations and a structural analysis that are performed with and without the effector molecule.[17,18] Unfortunately, the activator-bound structure of the *E. coli* ADP-Glc PPase is not known, nor is it for any other ADP-Glc PPase. Nonetheless, studies involving the *E. coli* enzyme have clearly shown that the interaction between ATP and FBP is synergistic, with FBP stimulating the binding of ATP (forward effect) and ATP stimulating the binding of FBP (reverse effect).[2,19] This synergy strongly suggests that a common set of structural components transmit both signals. Based on the Monod-Wyman-Changeux model of allosterism, both the allosteric activator and the substrate selectively bind to the same active conformation of the enzyme (R state). This form will have enhanced interactions with both ligands. In other words, if the allosteric activator enhances the binding of the substrate, the substrate must enhance the binding of the allosteric activator, and both will trigger the same structural pathway.[20] For that reason, information regarding the structural elements associated with the forward effect could lead to information regarding the reverse effect. With this in mind, we ran MD simulations on a homology model of the enzyme with both ATP-bound and unbound subunits.

We produced a correlation map that showed differences in the correlated movements between the ATP-bound versus non-bound subunits (Fig. 2). Residues associated with some of the largest positive differences (indicating that correlations in movement were present or significantly enhanced in the ATP-bound versus unbound subunit) involved a loop structure formed by residues Pro103–Arg115. Specifically, we saw that the greatest differences in movement correlation involved this loop and the FBP binding Lys39 region. Correlation differences were also seen with this loop and a glycine rich

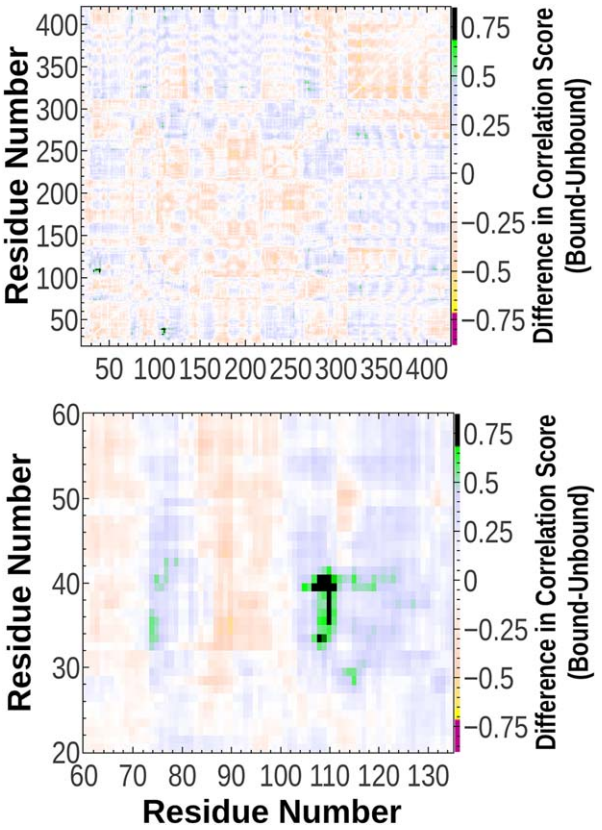

Figure 2. Correlation map showing the differences in correlation scores between ATP-bound and unbound subunits (bound - unbound). Some of the most intense readings involve the Pro103–Arg115 loop with the FBP binding Lys39 region, and this is highlighted in the lower panel of this figure. This image was produced as described in the Materials and Methods.

region (near Gly30) that is part of the ATP binding domain.[4] Furthermore, the network pathway analysis revealed that the dominant path of communication[18] between the FBP binding Lys39 and the adenine ring of ATP mainly involved residues of the Pro103–Arg115 loop, particularly those towards the center and C-terminal half of this structure (Table I, Supporting Information Tables S1, and S2, Fig. S1–I). In addition, a homology model[11] and our simulation indicated that the ATP-bound subunit and the substrate free subunits were structurally very similar, with the exception of the loops bearing regulatory residues Lys39,[14] Gln74,[11] an ATP-binding Leu25-Gly30 loop[4], and the Pro103–Arg115 loop (Fig. 3). Since they are closer together in the ATP-bound subunit, this form is referred to as the closed state whereas the unbound subunit represents the open form of the enzyme.

In the simulations, we observed several hydrogen bond interactions in the ATP-bound subunit involving the Pro103–Arg115 loop, Gln74, Lys39, ATP, and the Leu25–Gly30 loop that were maintained for most of the simulation, and these were generally absent in the open form (Fig. 4). In the

Table I. *Results of the Network Pathway Analysis, Showing the Five Shortest Communication Pathways Linking the FBP Lys39 Residue to the Adenine Ring of ATP in the ATP Bound Subunits (A and C)*

	Path number	Path residues
$Lys39_A \rightarrow ATP_A$	Path 1:	Lys109, Tyr75, Trp113, Arg115
	Path 2:	Lys109, Tyr75, Arg29, Trp113, Arg115
	Path 3:	Lys109, Glu111, Trp113, Arg115
	Path 4:	Lys109, Tyr75, Trp113, Arg115, Thr117
	Path 5:	Lys109, Tyr75, Gly27
$Lys39_C \rightarrow ATP_C$	Path 1:	Lys109, Glu111, Trp113, Arg115
	Path 2:	Lys109, Arg107, Gln105, Gln74, Arg115
	Path 3:	Lys109, Tyr75, Gly27
	Path 4:	Lys109, Arg107, Gln105, Gln74, Arg115
	Path 5:	Lys109, Arg107, Tyr75, Gly27

Most residues in the pathways belong to the Pro103-Arg115 loop. We focused on the adenine ring as target node. Pathways to the ribose ring node also mainly involved the Pro103-Arg115 loop residues. Analysis was performed as described in materials and methods. A detailed list of the 500 shortest pathways is provided in the supplemental information section of this report (Table S1)

ATP-bound subunit, the backbone oxygen of Trp113 hydrogen bonded with the side chain of Gln74, and in turn the backbone of Gln74 formed a hydrogen bond with Ala26 of the ATP binding domain (homologous to the tuber Gln75/Trp116 and Gln75/Gly28 interactions, respectively).[6] We also saw that, simultaneously, the side chain oxygen of Gln74 hydrogen bonded with the amino group of ATP for the entire simulation. The supposition that Gln74 hydrogen bonds with ATP is further supported by similar findings with Gln67 in the *A. tumefaciens* enzyme[7] and, while not present in the crystal structure, a hydrogen bond between Gln75 and ATP in the tuber enzyme formed within the first half nanosecond of an MD simulation (data not shown). For most of the simulation, Arg115 appeared to hydrogen bond with the adenine ring of ATP through both side chain and backbone atoms, homologous to the hydrogen bonds observed between Gln118 and ATP in the crystal structure of the potato tuber enzyme.[6]

The backbone atoms of Gly110 formed hydrogen bonds with backbone atoms of Lys39 or Asn38, which mark the FBP binding region. Further towards the N-terminal end of the loop, we saw that the side chain of Gln106 formed a hydrogen bond with the side chain of Gln76 for almost the entire simulation. Gln106 and Arg107 were also observed interacting with the opposing subunit, although the identity of involved residues was variable but included, amongst others, Asn124, Asp126, and Asp231. Also at this end of the loop, we observed a network of hydrogen bonds between the backbone atoms of Ala104–Arg107 and Gln74–Gln76, though this interaction was generally observed in both ATP bound and unbound subunits. In the potato tuber enzyme, there is a hydrogen bond between Asn113 and Lys40 (homologous to the Gly110 and Asn38/Lys39 interaction), and so are hydrogen bonds between residues homologous to Ala104–Arg107 and Gln74–Gln76.[6]

As a consequence of these computational studies, the Pro103–Arg115 loop seemed like it may serve an important role in enzyme function, particularly regarding allosteric regulation. To test this hypothesis, and to examine the role of the side chains in this loop, we performed alanine scanning mutagenesis and kinetically characterized the mutants.

Kinetic characterization

The results of the kinetic assays showed that the alanine mutants of Pro103–Arg115 loop residues can be divided into three main groups, depending primarily on their response to FBP. Group I mutants

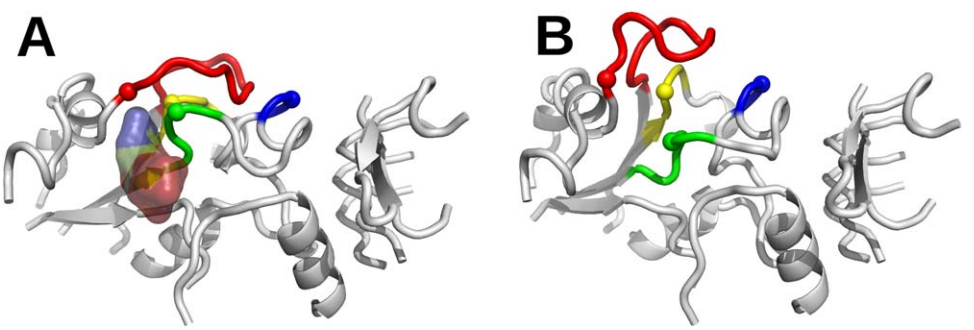

Figure 3. Image showing the differences in position of the Pro103–Arg115 loop, the Leu25-Gly30 region, and loops bearing Gln74 and Lys39 in the ATP bound (A) and unbound (B) subunits. The Pro103–Arg115 loop is colored red, Asn38–Ala40 are colored blue, Thr73–Gln75 are yellow, Leu25–Gly30 are in green. For reference, residues Arg29, Lys39, Gln74, and Arg115 are shown in spheres. ATP is shown in translucent surface format. Asp276 sits near the apex of the loop residing in the image foreground. This image was generated using Pymol 1.3 and shows the enzyme approximately 3 ns into the MD simulation.

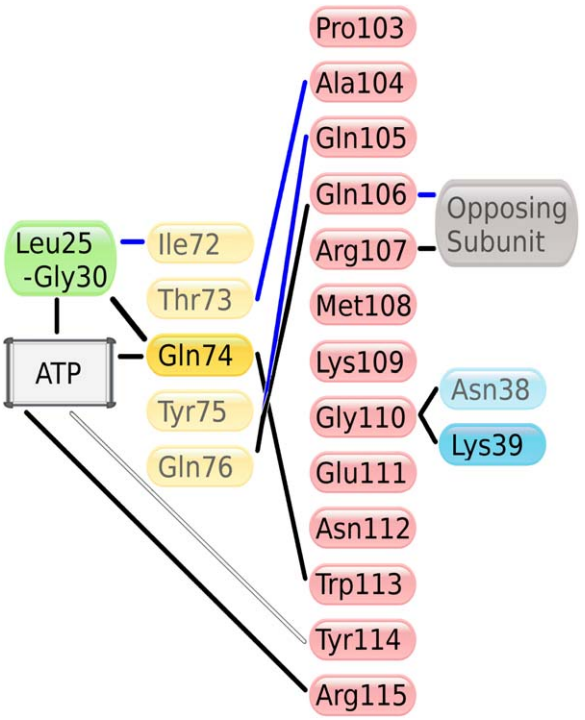

Figure 4. Schematic of some of the key interactions involving the Pro103–Arg115 loop observed in the MD simulations. The following hydrogen bonds were present in ATP bound subunits for at least 70% of the simulation: (Oxygen and nitrogen atoms that are labeled O and N represent backbone atoms, while OE1, NE1, NE2, and NH2 belong to side chains. Atoms N1 and N6 reside on the adenine ring of ATP): (*ATP*·**N6**–*Arg115*·**O**), (*ATP*·**N6**–*Gln74*·**OE1**), (*Gln74*·**N**–*Ala26*·**O**), (*Gln76*·**NE2**–*Gln106*·**OE1**), (*Ile72*·**O**–*Ala26*·**N**), (*Thr73*·**O**–*Ala104*·**N**), (*Asn38*·**O**–*Gly110*·**N**), and (*Gln74*·**NE2**–*Trp113*·**O**). Additional (*ATP*·**N6**–*Arg115*·**NE**) and (*Gln105*·**O**–*Gln76*·**N**) interactions were maintained for about half of the simulation. When not hydrogen bonding with the backbone of Asn38, Gly110 would hydrogen bond with the backbone of Lys39. None of these interactions, except (*Thr73*·**O**–*Ala104*·**N**), (*Ile72*·**O**–*Ala26*·**N**), and (*Gln105*·***O***–*Gln76*·***N***) were observed in the non-ATP bound subunit (blue lines). Previous experimental data suggests that Tyr^{114} comes within close contact of ATP during some stage of the binding process, as mentioned in the Discussion section of this report.

include Q105A, M108A, K109A, E111A, and N112A, which had kinetic profiles that resembled the wild-type enzyme. Upon the addition of FBP, there was a marked increase in the apparent affinities for all substrates in this group (Tables II and III, Figs. 5 and 6). The ATP $S_{0.5}$ values shifted from 1.7–3.3 m*M* to 0.3–0.7 m*M* when activator was added, which nears the 2.5 m*M* to 0.2 m*M* shift observed with the wild-type. Similar effects were observed for the apparent Glc1P and Mg^{2+} affinities. An exception here would be the 1.5 fold change in Mg^{2+} apparent affinities for M108A. FBP also induced a significant rise in V_{max} with Group I mutants under standard conditions, defined here as 1.5 m*M* ATP, 1.0 m*M* Glc1P, and 10 m*M* Mg^{2+} at pH 8.0 (Table IV, Fig. 6). The specific activities increased 25-fold for the wild-type enzyme, and 10- to 18-fold for these mutants. They also had similar FBP $A_{0.5}$ values as the wild-type enzyme. All of them were near 0.08 m*M*, with only Q105A requiring slightly higher FBP concentrations (0.17 m*M*). Overall, these enzymes generally had specific activities that were at least 50–80% those of the wild-type.

Group II mutants comprise Q106A, R107A, W113A, and Y114A. They had kinetic profiles which differed significantly from the wild-type, particularly with regards to their response to FBP (Tables II–IV, Figs. 5 and 6). In the absence of activator, their behaviors were similar to the wild-type enzyme. ATP, Glc1P, and Mg^{2+} affinities were within 1–2 fold of those seen in the wild-type (Tables II and III). Specific activities under standard conditions without FBP ranged from 2.1 to 3.2 U/mg, which is close to the value of 3.6 U/mg seen for wild type. Large differences between Group II mutants and the wild-type were observed in the presence of FBP. Whereas apparent affinities for ATP, Glc1P, and Mg^{2+} for the wild-type enzyme were found to increase approximately 15-, 16-, and 3.5-fold, respectively, all such values were found to increase by no more than 2-fold for Group II mutants. Specific activities increased by no more than 4-fold for these mutants

Table II. *Effect of Ala mutations on the kinetic parameters of the Hill equation for ATP saturation curves, as determined with and without 1.5 m*M *FBP.*[a]

		ATP		
	FBP (m*M*)	V_{max} (U/mg)	$S_{0.5}$ (m*M*)	*n*
WT	1.5	88.0 ± 0.6	0.17 ± 0.01	2.0 ± 0.1
	0	21.8 ± 0.8	2.5 ± 0.1	2.4 ± 0.2
P103A	1.5	23.7 ± 0.8	0.81 ± 0.06	1.5 ± 0.1
	0	18.3 ± 0.2	0.89 ± 0.02	1.9 ± 0.1
Q105A	1.5	52.0 ± 1.9	0.74 ± 0.06	1.4 ± 0.1
	0	13.0 ± 0.3	1.91 ± 0.05	2.3 ± 0.1
Q106A	1.5	32.2 ± 4.5	4.4 ± 1.1	1.2 ± 0.1
	0	11.1 ± 0.6	2.5 ± 0.2	1.7 ± 1.6
R107A	1.5	27.2 ± 0.6	1.63 ± 0.05	2.3 ± 0.1
	0	16.8 ± 0.5	2.5 ± 0.1	2.1 ± 0.1
M108A	1.5	42.6 ± 2.0	0.31 ± 0.03	2.2 ± 0.5
	0	21.5 ± 0.3	2.54 ± 0.06	1.8 ± 0.1
K109A	1.5	55.7 ± 0.5	0.38 ± 0.01	2.2 ± 0.1
	0	16.0 ± 0.3	2.11 ± 0.06	1.5 ± 0.1
E111A	1.5	66.7 ± 1.9	0.29 ± 0.02	1.3 ± 0.1
	0	15.5 ± 1.5	3.3 ± 0.5	1.7 ± 0.2
N112A	1.5	41.9 ± 2.0	0.43 ± 0.05	1.7 ± 0.3
	0	13.3 ± 1.7	1.7 ± 0.3	1.7 ± 0.4
W113A	1.5	15.0 ± 2.3	3.1 ± 1.2	1.1 ± 0.1
	0	11.3 ± 1.0	3.5 ± 0.7	1.2 ± 0.1
Y114A	1.5	14.1 ± 1.5	2.3 ± 0.3	1.9 ± 0.3
	0	10.1 ± 0.3	2.5 ± 0.1	2.0 ± 0.1
R115A	1.5	13.2 ± 0.7	1.3 ± 0.1	1.7 ± 0.2
	0	10.1 ± 1.3	3.0 ± 0.5	2.1 ± 0.5

[a] Enzyme assays were performed as described in Materials and Methods.

Table III. *Effect of Ala mutations on the kinetic parameters of the Hill equation for Glc-1P and Mg^{2+} saturation curves, as determined with and without 1.5 m*M *FBP.*[a]

	Glc1P			Mg^{2+}	
	FBP (m*M*)	$S_{0.5}$ (m*M*)	*n*	$S_{0.5}$ (m*M*)	*n*
WT	1.5	0.023 ± 0.002	1.4 ± 0.2	1.3 ± 0.1	1.8 ± 0.1
	0	0.37 ± 0.04	1.5 ± 0.1	6.0 ± 0.4	1.1 ± 0.4
P103A	1.5	0.17 ± 0.01	1.4 ± 0.1	1.5 ± 0.2	1.0 ± 0.2
	0	0.22 ± 0.02	1.1 ± 0.1	1.5 ± 0.2	1.3 ± 0.2
Q105A	1.5	0.018 ± 0.002	1.1 ± 0.2	2.0 ± 0.1	2.8 ± 0.5
	0	0.33 ± 0.01	1.8 ± 0.1	7.4 ± 1.9	1.1 ± 0.1
Q106A	1.5	0.120 ± 0.009	1.3 ± 0.1	4.8 ± 1.3	1.1 ± 0.2
	0	0.25 ± 0.03	1.6 ± 0.2	4.6 ± 0.4	1.9 ± 0.4
R107A	1.5	0.12 ± 0.02	1.0 ± 0.1	6.7 ± 1.9	1.0 ± 0.2
	0	0.26 ± 0.12	0.9 ± 0.1	6.4 ± 1.2	1.3 ± 0.3
M108A	1.5	0.019 ± 0.002	1.1 ± 0.1	3.2 ± 0.2	2.2 ± 0.3
	0	0.24 ± 0.01	1.8 ± 0.1	4.6 ± 1.6	1.3 ± 0.3
K109A	1.5	0.025 ± 0.001	1.6 ± 0.1	0.8 ± 0.1	1.3 ± 0.2
	0	0.27 ± 0.02	2.5 ± 0.5	4.0 ± 0.3	2.8 ± 0.8
E111A	1.5	0.029 ± 0.002	1.9 ± 0.2	1.0 ± 0.1	1.7 ± 0.3
	0	0.39 ± 0.05	1.5 ± 2.6	4.7 ± 0.9	1.4 ± 0.2
N112A	1.5	0.040 ± 0.009	1.5 ± 0.8	0.99 ± 0.04	4.1 ± 0.7
	0	0.363 ± 0.005	1.5 ± 0.1	4.5 ± 0.5	1.7 ± 0.3
W113A	1.5	0.34 ± 0.07	1.3 ± 0.2	5.8 ± 1.4	1.6 ± 0.4
	0	0.46 ± 0.02	1.7 ± 0.1	5.0 ± 2.0	1.6 ± 0.7
Y114A	1.5	0.27 ± 0.16	0.2 ± 0.2	6.2 ± 2.1	1.2 ± 0.2
	0	0.43 ± 0.14	0.3 ± 0.2	6.9 ± 2.1	1.2 ± 0.2
R115A	1.5	0.143 ± 0.002	1.7 ± 0.1	5.4 ± 0.6	1.3 ± 0.1
	0	0.28 ± 0.01	1.7 ± 0.1	5.4 ± 0.7	1.7 ± 0.4

[a] Enzyme assays were performed as described in Materials and Methods.

upon the addition of FBP under standard conditions, which is significantly less than the 25-fold change seen for wild type.

R115A is also considered a Group II enzyme because it had a kinetic profile markedly different from wild-type enzyme. This includes decreased substrate affinities along with specific activities that were markedly lower than the wild type (Tables II–IV, and Figs. 5 and 6). Under standard conditions in the absence of FBP, R115A had, in fact, the lowest specific activity of any of the mutants studied in this report, at 1.1 U/mg (Table IV, Fig. 6). The FBP saturation curves, however, revealed that under standard conditions there was a nearly 10 fold increase in V_{max} for this enzyme upon the addition of FBP. However, it had the highest FBP $A_{0.5}$ of any of the enzymes, at 0.38 m*M*, and it also had significantly higher ATP $S_{0.5}$ than wild-type (Fig. 5).

At last, a third group could be defined consisting only of P103A. Like most of the Group II enzymes, it did not respond significantly to activator FBP. There was less than 1.5 fold shift in either specific activities or substrate affinities with the activator (Tables II–IV, Figs. 5 and 6). What is unique about this mutant, however, is that apparent affinities for ATP and Mg^{2+} in the absence of activator were

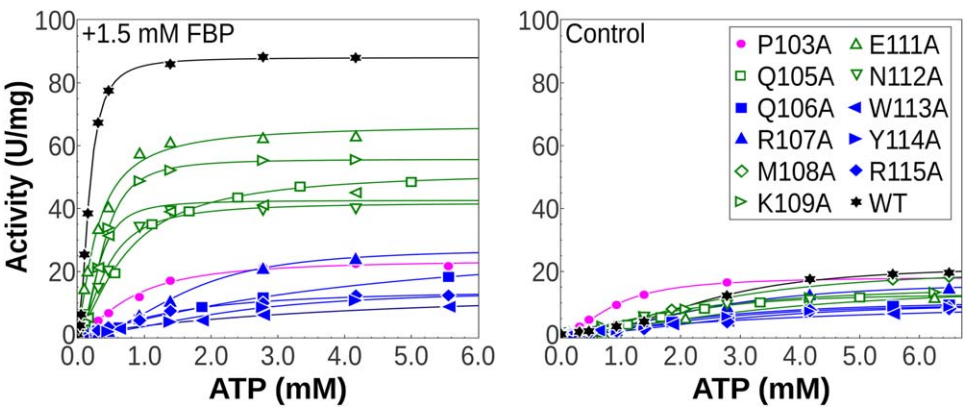

Figure 5. ATP saturation curves of wild-type and alanine mutants of *E. coli* ADP-Glc PPases conducted with 1.5 m*M* FBP or without FBP (control). Assays were performed as described in the Materials and Methods section.

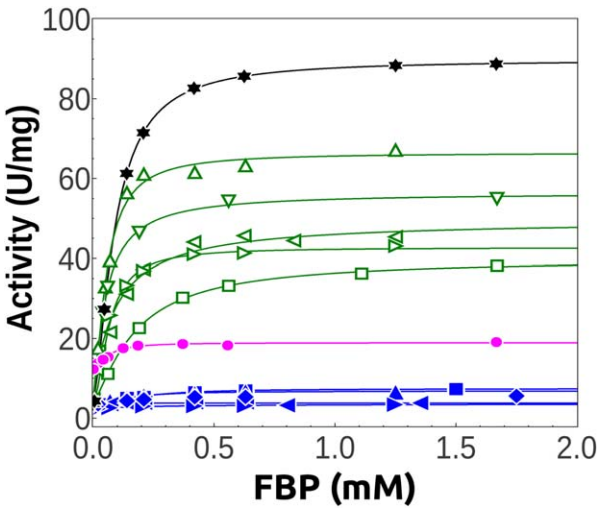

Figure 6. FBP-activation kinetics of wild-type and alanine mutants of *E. coli* ADP-Glc PPases. P103A (magenta ●), Q105A (green □), Q106A (blue ■), R107A (blue ▲), M108A (green ◇), K109A (green ▷), E111A (green △), N112A (green ▽), W113A (◀), Y114A (▶), R115A (◆), and wild type (✱). Assays were performed as described in the materials and methods section.

substantially higher than those of the wild-type enzyme. The ATP $S_{0.5}$ was just 0.9 m*M* (2.5 m*M* for wild type) and the Mg^{2+} $S_{0.5}$ was 1.5 m*M* (5.9 m*M* for wild type). Glc1P $S_{0.5}$ values were also lower, at 0.2 m*M*, compared with 0.4 m*M* for wild type. Furthermore, under standard conditions in the absence of activator, P103A has specific activity that is substantially higher than the wild-type enzyme, at approximately 14.8 U/mg, compared to 3.6 U/mg for wild type (Table IV, Fig. 6). However, upon the addition of FBP, P103A loses any kinetic advantages over the wild type. Whereas the wild-type enzyme achieves a specific activity of about 90 U/mg under standard conditions with FBP, this value is just over 20 U/mg for P103A. The ATP and Glc1P apparent affinities of the activated P103A enzyme are lower than wild type, at 0.8 m*M* and 0.17 m*M*, respectively, compared with 0.2 m*M* and 0.02 m*M* for wild type.

All mutant and wild-type forms of ADP-Glc PPase studied in this work exhibited some inhibition by AMP (Table V). However, enzymes which had the least sensitivity to FBP were also the least sensitive to AMP inhibition. For instance, whereas the wild-type enzyme retains just over 3% of its activity with saturating AMP (Table V), this figure is 50–70% for enzymes like P103A, W113A, and Y114A.

Sequence alignment

The results of the sequence alignment (Fig. 7) reveal that several of the Pro103–Arg115 loop residues are highly conserved. This was particularly true for Group II and III residues, with the exception of Arg115. Trp113 is universally conserved in all known ADP-Glc PPases, with the exception of Lactobacilli. Pro103 and Tyr114 are similarly conserved in non-photosynthetic bacteria. In photosynthetic species, Tyr114 is replaced by phenylalanine. In the homology model, Pro103 terminates a preceding beta sheet (formed by residues 98–102), and it marks the beginning of the Pro103–Arg115 loop. In ADP-Glc PPases from photosynthetic species, this residue is an alanine. Specifically, two consecutive alanine residues break the preceding beta sheet and mark the beginning of the homologous loop structure in these enzymes. A glutamine at positions akin to residue 106 of the *E. coli* enzyme is present nearly 70% of the time in plant and bacterial ADP-Glc PPases. However, the identity of its replacement in other ADP-Glc PPases is more variable, with both polar and hydrophobic residues observed in the remaining 30% of sequences. While Gln105 is comparatively less conserved, it is interesting that nearly 80% of all ADP-Glc PPase sequences have a glutamine at either positions 105 or 106, and about 45% have glutamines at both sites, including the

Table IV. *Effect of Ala mutations on the FBP saturation curves.*[a]

	FBP				
	$A_{0.5}$ (m*M*)	*n*	V_0[b] (U/mg)	V_{max} (U/mg)	Activation-Fold
WT	0.077±0.002	1.8 ± 0.1	3.6 ± 0.6	88.9 ± 0.7	25.0
P103A	0.062±0.009	1.8 ± 0.4	14.8 ± 0.3	21.6 ± 0.4	1.5
Q105A	0.174±0.006	1.3 ± 0.1	3.7 ± 0.4	39.8 ± 0.4	10.7
Q106A	0.18±0.03	1.7 ± 0.5	2.7 ± 0.7	7.4 ± 0.3	2.7
R107A	0.08±0.01	1.2 ± 0.3	3.2 ± 0.3	11.0 ± 0.4	3.5
M108A	0.09±0.02	1.0 ± 0.2	5.0 ± 2.0	49.6 ± 2.8	9.9
K109A	0.064±0.005	1.6 ± 0.2	2.7 ± 1.4	42.7 ± 1.1	15.7
E111A	0.053±0.005	1.5 ± 0.2	3.6 ± 2.2	66.5 ± 1.7	18.3
N112A	0.052±0.002	1.2 ± 0.1	4.5 ± 0.7	56.3 ± 0.6	12.5
W113A	0.090±0.009	1.6 ± 0.3	3.1 ± 0.1	4.1 ± 0.1	1.3
Y114A	0.08±0.01	1.4 ± 0.3	2.1 ± 0.1	3.2 ± 0.1	1.5
R115A	0.38±0.06	1.4 ± 0.2	1.1 ± 0.2	11.2 ± 0.8	10.0

[a] Enzyme assays were performed as described in Materials and Methods.
[b] V_0 represents the specific activity with 0 m*M* FBP.

Table V. *Effect of Pro[103]–Arg[115] Ala mutations on AMP induced inhibition.*[a]

	AMP				
	$I_{0.5}$ (m*M*)	V_0[b] (U/mg)	V_∞[c] (U/mg)	*n*	V_∞/V_0 (%)
WT	0.064 ± 0.009	85.5 ± 3.4	3.2 ± 4.0	1.1 ± 0.2	3.7
P103A	0.079 ± 0.005	14.6 ± 0.1	7.5 ± 0.1	1.3 ± 0.1	51.1
Q105A	0.184 ± 0.012	28.4 ± 0.9	2.3 ± 0.7	2.1 ± 0.3	8.1
Q106A	0.064 ± 0.008	7.1 ± 0.2	1.9 ± 0.2	1.5 ± 0.2	26.0
R107A	0.083 ± 0.012	11.7 ± 0.3	2.5 ± 0.4	1.4 ± 0.2	20.9
M108A	0.049 ± 0.002	44.2 ± 0.8	4.6 ± 0.6	3.2 ± 0.3	10.3
K109A	0.056 ± 0.003	42.1 ± 1.0	4.4 ± 0.9	2.9 ± 1.1	10.6
E111A	0.103 ± 0.013	63.3 ± 2.8	4.5 ± 2.3	1.9 ± 0.4	7.1
N112A	0.089 ± 0.005	41.6 ± 1.2	1.7 ± 1.1	2.6 ± 0.4	4.2
W113A	0.065 ± 0.012	2.9 ± 0.1	1.5 ± 0.1	0.9 ± 0.1	51.2
Y114A	0.034 ± 0.004	2.6 ± 0.1	1.7 ± 0.4	1.4 ± 0.1	66.7
R115A	0.073 ± 0.003	4.1 ± 0.2	0.5 ± 0.2	0.9 ± 0.2	12.7

[a] Enzyme assays were performed as described in Materials and Methods, using 1.5 m*M* ATP, 1.0 m*M* Glc1P, 10.0 m*M* $MgCl_2$, 0.18 m*M* FBP, at pH 8.0.
[b] V_0 represents the specific activity with 0 m*M* AMP.
[c] V_∞ represents the velocity at saturating inhibitor concentrations.

E. coli, A. tumefaciens, and potato tuber enzymes. Arginine is observed at position 107 in about 60% of all sequences, and is typically replaced by other polar amino acids when not present as arginine. Group I residues 108, 109, 111, and 112 (along with the previously mentioned 105) are comparatively less conserved, and so is residue 115 (Group II). Despite having decreased substrate affinities and lower specific activity, the alanine mutant of Arg115 did (like the Group I mutants) show a partial response to FBP with regards to V_{max} (Table IV, Fig. 6). Overall, we saw that the alanine mutants of the

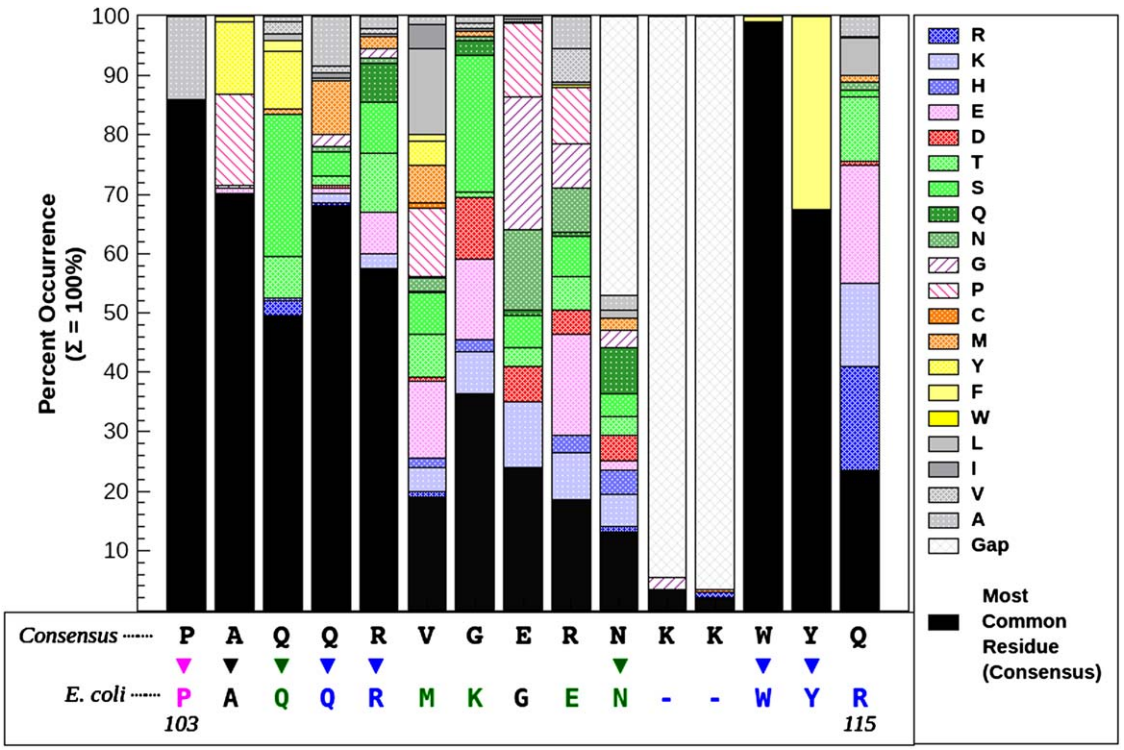

Figure 7. Sequence conservation of the Pro103–Arg115 loop. Here, 242 non-redundant ADP-Glc PPase sequences are represented. The y-axis represents the percent abundance of a particular amino acid type (summing to 100%). The most commonly observed amino acid at each site are represented by black bars, whereas the color representations for less commonly observed amino acids are described in the figure legend. The consensus and *E. coli* sequences are shown at the bottom of the figure, and alignments are indicated by downward arrows. Several residues of the Pro103–Arg115 loop, which reads **PAQQRMKGENYWR**, are highly conserved, and these tend to be Group-II or III residues (colored magenta and blue, respectively). Group I residues (green) are generally less conserved.

more conserved Pro103–Arg115 loop residues were likely to have diminished FBP response compared to the alanine mutants of less conserved residues.

Discussion

MD simulations provided important information on building a hypothesis about where a FBP triggered allosteric pathway is located in the *E. coli* ADP-Glc PPase. As shown in Figure 2, strong differences in correlated movement were observed with regions of the enzyme associated with ATP binding (near Leu25–Gly30),[21] regulation (near Gln74 and Lys39),[11,14] and a span which included a Pro103–Arg115 loop (Fig. 2). Overall, however, the most prominent differences in correlation involved the Pro103–Arg115 loop with residues including Lys39. In addition, the optimal network pathway[18] linking the FBP binding Lys39 to ATP mainly involved residues of the Pro103–Arg115 loop (Table I, Supporting Information Tables S1, and S2, Fig. S1). Subsequent analysis of the simulation revealed that these regions were within close contact of each other in the ATP bound subunit. In the subunit without the ligand, however, these correlated movements are absent or greatly diminished and the distance between them significantly increased (Figs. 2 and 3). As was mentioned, residues along these regions, including the Pro103–Arg115 loop (Fig. 7), are well conserved and their interactions (Fig. 4) coincide with features observed in the *A. tumefaciens* and potato tuber ADP-Glc PPase enzymes. These findings led to the hypothesis that the Pro103–Arg115 loop may be important in allosteric regulation, and we tested it by performing alanine scanning mutagenesis on the loop residues and kinetically characterizing the mutants.

We found that several of the alanine mutants had kinetic profiles that were very different from the wild-type enzyme, particularly with regards to FBP activation. Overall, the mutant enzymes which were less responsive to activation (Table IV) tended be less responsive to AMP inhibition, supporting the idea that one of the main mechanisms of AMP inhibition involves competition with FBP (preventing it from activating the enzyme).[4] Based on our results, we classified the alanine mutants into three main categories. Group I enzymes, which were Q105A, M108A, K109A, E111A, and N112A, had kinetic profiles most closely matching the wild-type ADP-Glc PPase. These residues are comparatively less conserved than the other loop residues (Fig. 7). Group II included Q106A, R107A, W113A, Y114A, and R115A, which differed most from the wild-type enzyme (particularly with regards to FBP activation), and the Group III mutant, P103A, appeared to exist in a partially pre-activated state. Group II and III mutants highlight the most critical residues, which are structurally located at the ends of the loop.

There is evidence supporting a relationship between ATP with Gln74, Trp113, Tyr114, and Arg115 (Fig. 4). In addition to the previously proposed interaction between Gln74 and Trp113,[11] the side chain oxygen of Gln74 maintained a hydrogen bond with the amino group of ATP for the entire simulation. Furthermore, the backbone atoms of Gln74 also formed hydrogen bonds with the backbone of the ATP binding Leu25–Gly30 loop (Fig. 4). As was mentioned above, homologous and universally conserved glutamines in other ADP-Glc PPases probably also interact with ATP. This observation, together with kinetics data on Q74A suggests that Gln74 exerts its influence by interacting with ATP, and that this interaction is stabilized by the presence of FBP.[11]

In our simulation, Tyr114 stayed at least ~4 Å away from the ATP substrate. However, previously research strongly suggests that this residue is involved in ATP contact. First, it was shown that 8-azido ATP would crosslink with the hydroxyl of Tyr114 upon UV radiation.[22] The Y114F mutant was also found to have decreased ATP affinities and diminished FBP activation.[23] Our findings on Y114A were more dramatic, which indicates that the presence of an aromatic side chain plays an important role, as has been shown with Trp113.[11] In addition, steered MD simulations in which the ATP was pulled out of the binding site showed that the adenine ring would come into close contact (< 3.5 Å) with the side chain of Tyr114 during this process (data not shown). Therefore we cannot discard that the side chain of Tyr114 interacts with ATP at some intermediate stage of the binding process.

Arg115 also interacted with the adenine ring of ATP through both backbone oxygen and side chain nitrogens (Fig. 4). Although Arg115 is not highly conserved (Fig. 7), homologous residues tend to be polar and capable of hydrogen bonding, and the most common residue for this site is Gln. Gln118 of the potato tuber enzyme interacts with ATP in a similar manner.[6] Furthermore, the R115Q mutant of the *E. coli* enzyme not only had improved FBP response compared to the alanine mutant without FBP, it also had higher apparent ATP affinity than the wild-type enzyme (Table VI, Fig. 8). It is interesting that while R115A had lowered apparent substrate affinities and specific activities (like the other Group II mutants), it had a partial response to FBP under standard conditions (Table IV, Fig. 6). However, even here, the relative FBP-induced increase in V_{max} was less than half that of WT (Fig. 6), and ATP $S_{0.5}$ values remained comparatively high (Table II). As such, Arg115 could be considered a residue that is important for both ATP binding and FBP activation.

Table VI. *Effect of Asn, Asn, and Gln mutations of Gln^{106}, Arg^{107}, and Arg^{115} (respectively) on the kinetic parameters of the Hill equation for ATP saturation curves, as determined with and without 1.5 mM FBP.*[a]

		ATP		
	FBP (m*M*)	V_{max} (U/mg)	$S_{0.5}$ (m*M*)	*n*
WT	1.5	88.0 ± 0.6	0.17 ± 0.01	2.0 ± 0.1
	0	21.8 ± 0.8	2.5 ± 0.1	2.4 ± 0.2
Q106N	1.5	11.6 ± 0.9	1.8 ± 0.5	1.1 ± 0.1
	0	5.1 ± 0.4	1.3 ± 0.3	1.5 ± 0.3
R107N	1.5	68.5 ± 1.3	0.56 ± 0.02	3.0 ± 0.2
	0	23.0 ± 1.7	2.6 ± 0.2	2.2 ± 0.2
R115Q	1.5	64.4 ± 0.5	0.43 ± 0.01	3.9 ± 0.2
	0	5.2 ± 0.2	0.69 ± 0.04	2.6 ± 0.3

[a] Enzyme assays were performed as described in Materials and Methods.

Further towards the N-terminal end of the loop, the side chain of Gln106 formed a hydrogen bond with the side chain of Gln76 for almost all of the simulation, and Arg107 was observed interacting with the opposing subunit, although the identity of involved residues was variable. This may be noteworthy because inter-subunit interactions have previously been suggested as being important for the regulation of ADP-Glc PPases.[24,25] Nonetheless, while Arg107 had some electrostatic interactions with opposing subunit residues like Asp126 or Asp231, the R107N mutant had significantly improved response to FBP compared to R107A (Table VI). This means that electrostatic interactions are not crucial for this residue site. On the other hand, the Ala mutant of Gln106 had a kinetic profile similar to the Q106N, as both failed to appreciably respond to FBP (Table VI). This would suggest the existence of specific interactions involving Gln106, although further studies would be needed to determine involved residues. Interestingly, nearly 80% of the sequences referenced in Figure 7 have a Gln at either positions 105 or 106, and about half have Gln at both sites, as is the case with the *E. coli* enzyme. In fact, even though Q105A is classified here as a Group II mutant, it had a substantially elevated FBP $A_{0.5}$, as did Q106A (Table IV).

P103A had a unique kinetic profile. In this case, FBP did not cause a significant change in substrate affinities or specific activity. This is similar to a previously studied mutant bearing a pentapeptide insertion at Leu102[26] (as well as the Group II mutants studied here). However, unlike these enzymes, our kinetic assays indicate that P103A exists in a partially pre-activated state. This is demonstrated by substrate affinities determined in the absence of FBP that rival or exceed those of the activated wild-type enzyme (Tables II–IV, Figs. 5 and 6). Despite this fact, P103A with FBP had an impaired kinetic profile compared to the activated WT enzyme (Tables II–IV). We suggest that proline's unique structural characteristics help to ensure the proper positioning of downstream loop residues, some of which may lie near FBP and ATP binding sites (Fig. 1). Other mutants of the *E. coli* ADP-Glc PPase that are pre-activated or even fully activated include G336D, P295D, P295E,[27,28] as well as a mutant missing the first 11–15 residues at the N-terminus.[29] Residues 336 and 295 lie in the proximity of the interface between the C- and N-terminal domains, as do parts of the Pro103–Arg115 loop (Fig. 1), and this could support the idea that interactions between C- and N-terminal domains are important in the overall allosteric mechanism.[12] It is possible all of these mutants, including P103A, favor a common active form of the enzyme.

The side chains of Group I residues, which include Gln105 and Met108–Asn112 and are comparatively less conserved (Fig. 7), appear to play no critical role in the regulation or functioning of this enzyme. However, we cannot discard that the backbone of these residues participates in important interactions. These may include hydrogen bonds between the backbone atoms of Gly110 and residues at or near Lys39, as well as hydrogen bonds between the backbone atoms of Ala104–Arg107 and residues

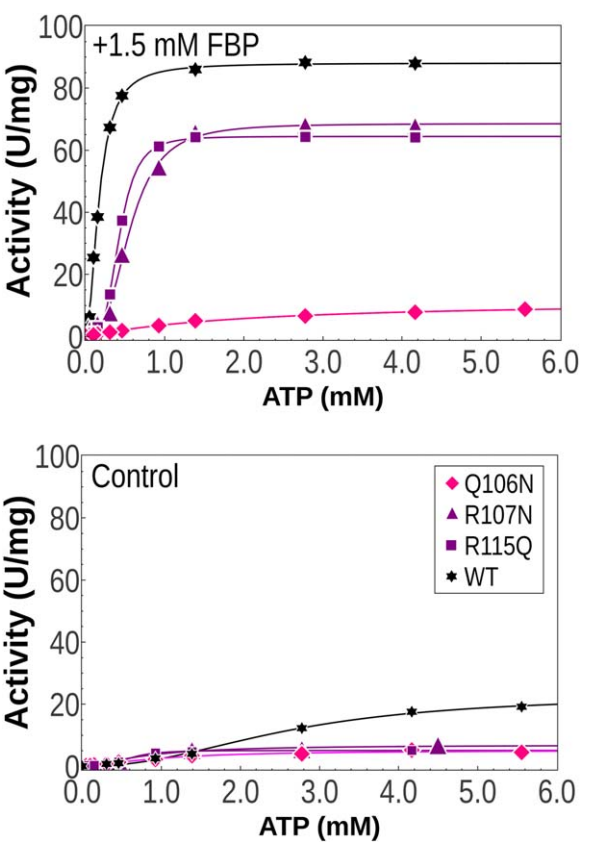

Figure 8. ATP saturation curves of wild-type and Q106N, R107N, and R115Q mutants of *E. coli* ADP-Glc PPases conducted with 1.5 m*M* FBP and without FBP (control). Assays were performed as described in the materials and methods section.

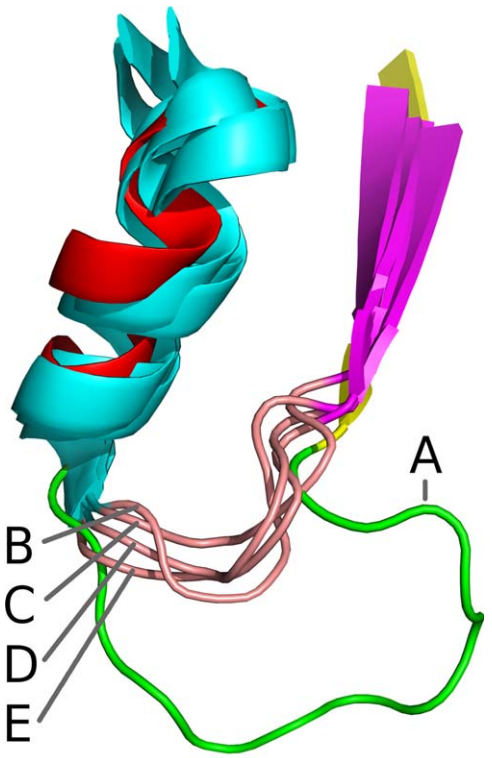

Figure 9. *E. coli* Pro103–Arg115 loop with flanking α-helix and β-sheet overlayed on top of the homologous structures found in other NDP-Glc PPases, which are not allosterically regulated. The *E. coli* ADP-Glc PPase (A) is shown with red α-helix, green loop, and yellow β-sheet. The other structures that are shown are the CDP-Glc PPase from *Salmonella enterica* (B), the *E. coli* dTDP-Glc PPase (C), the *E. coli* UDP-Glc PPase (D), and the UDP-Glc PPase from *Helicobacter pylori* (E) (PDB accession codes 1WVC, 1H5T, 2E3D, and 3JUK, respectively).

near Gln74 (Fig. 4). If this were the case, then an examination of Figures 1 and 3 would suggest that the interaction between Gly110 and Asn38/Lys39 functions as a latch (with hinges at the ends of the loop), helping to modulate the "closure" of the Pro103–Arg115 loop, thereby facilitating the aforementioned interactions that involve these regions.

Comparisons with analogous loops of other NDP-Glc PPases

Aside from analyzing the sequences of the *E. coli* ADP-Glc PPase with other ADP-Glc PPases, it is worth examining other NDP-Glc PPases. All the ADP-Glc PPases as well as other sugar nucleotide pyrophosphorylases appear to have evolved from a common ancestor.[4] Despite low sequence similarity, many of the structural features observed in the ADP-Glc PPases are observed in the other bacterial NDP-glucose pyrophosphorylases.[4,26] For instance, the catalytic domains in all of these enzymes have virtually identical folds. These features are represented in the crystal structure of the *E. coli* UDP-Glc PPase and dTDP-Glc PPase, the CDP-Glc PPase from *Salmonella enterica*, and the UDP-Glc PPase from *Helicobacter pylori* (PDB accession codes 2E3D, 1H5T, 1WVC, and 3JUK, respectively).[30–33] Other similarities between these enzymes included regions homologous to the Pro103–Arg115 loop. An overlay of the *E. coli* ADP-Glc PPase homology model with these NDP-Glc PPases that focuses on this loop with flanking α helix and β sheets is shown in Figure 9. There is remarkable overlap between the structures, with one distinct difference. The other NDP-Glc PPases have a short loop structure (5–7 residues) linking the α helix and β sheet. This compares to the 13-residue long loop in the *E. coli* ADP-Glc PPase (with a range of 12–15 and average of 12.6 residues amongst the 242 ADP-Glc PPases sequences described in Fig. 7). What is perhaps more interesting is that all of the other bacterial NDP-Glc PPases are unregulated, whereas the ADP-Glc PPases are generally highly regulated allosterically.[4] As we have shown in this article, disruption of this loop (i.e., single alanine mutation of conserved residues) causes dramatic loss of allosteric activation. Since ADP-Glc PPase is the only NDP-Glc PPase with this loop, we speculate that it originally appeared as an insertion that evolved over time to serve a regulatory role. Another example of such an "insertion" present in ADP-Glc PPases is the C-terminal domain (Fig. 1). When the C-terminal domain of the *E. coli* ADP-Glc PPase was added to the *Streptococcus mutans* UDP-Glc PPase, the resultant enzyme had a partial response to activators like FBP. However, the effect was small compared to the change in activity seen in WT ADP-Glc PPase.[34] As such, the Pro103–Arg115 loop may be one of the necessary, but not solely sufficient, elements for the evolution of allosteric control in this enzyme.

The data presented here clearly demonstrates that the Pro103–Arg115 loop plays an important role in the allosteric regulation of the *E. coli* ADP-Glc PPase. We propose that the Pro103–Arg115 loop and loops bearing Lys39, Gln74, and Leu25-Gly30 form a network of interactions that may help propagate the allosteric signal from the FBP binding site to the ATP binding domain. Efforts are currently underway to crystallize ADP-Glc PPases in the presence of activator, which we hope will further the understanding of allosteric activation in this enzyme.

Materials and Methods

Chemicals and supplies

Adenosine-5'-triphosphate disodium salt (ATP), Adenosine-5'-monophosphate sodium salt (AMP), α-D-glucose-1-phosphate disodium salt (Glc1P), D-fructose-1,6-bisphosphate tetra[cyclohexylammonium] salt (FBP), sucrose, magnesium chloride ($MgCl_2$),

sodium chloride (NaCl), bovine serum albumin (BSA), 3-[4-(2-Hydroxyethyl)−1-piperazinyl]propanesulfonic acid (HEPPS), 4-(2-hydroxyethyl)−1-piperazineethanesulfonic acid (HEPES), ethylenediaminetetraacetic acid (EDTA), ammonium molybdate (AM), malachite green oxalate salt (MG), tween-20 (T20), nalidixic acid, trisodium citrate, and inorganic pyrophosphatase from yeast (PPtase) were obtained from Sigma-Aldrich (St. Louis, MO). Luria broth (LB) media was purchased from USA Scientific (Ocala. FL). StrataClone Blunt PCR Cloning Kits were acquired from Agilent Technologies (Santa Clara, CA). Mutant primers for the production of *E. coli* ADP-Glc PPase mutants were acquired from Integrated DNA Technologies (Coralville, IA) and their sequences are listed in Supporting Information. T7 DNA ligase, NEB turbo competent cells, Nde1 and Sac1 restriction enzymes and Phusion DNA polymerase were purchased from New England Biolabs (Ipswich, MA). Wizard-Plus SV Miniprep and Wizard SV Gel and PCR Clean-Up kits were obtained from Promega (Madison, WI). The Malachite Green-ammonium molybdate-tween20 (MG-AM-T20) solution used in the kinetics assays was prepared as previously described.[35]

Site-directed mutagenesis and protein expression

Site-directed mutagenesis of a pMAB6 plasmid bearing the *E. coli* ADP-Glc PPase gene was performed as previously described[8] with few modifications. The mutant primers used to create the mutants are shown in the Supplemental section of this report (Supporting Information Table S3). All genetic sequencing was performed by the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility in Chicago, Illinois. Protein expression and subsequent purification of enzymes was performed as previously described.[8] Final enzyme purity was estimated at 95% or greater for all enzymes as determined by SDS page.[36]

Enzyme assays

Activity of ADP-Glc PPase was assayed in the direction of ADP-glucose synthesis using the highly sensitive phosphate detecting colorimetric assay developed by Fusari *et al*.[35] The unit of enzyme activity (U) is defined as 1.0 micromole of ADP-glucose formed per minute. Unless otherwise stated, assay conditions were as follows. In addition to the 10 μl of dilute ADP-Glc PPase enzyme solution, the reaction mixtures (with or without 1.5 m*M* FBP) contained 50 m*M* Hepes pH 8.0, 10.0 m*M* $MgCl_2$, 1.0 m*M* Glc-1P, 1.5 m*M* ATP, 1.5 U/ml inorganic pyrophosphatase, 0.2 mg/ml BSA, in a total volume of 60 μl. These concentrations define the "standard conditions" mentioned in this report. Reactions were initiated by the addition of the diluted enzyme solutions and proceeded for 10 minutes at 37°C in a sealed 1.7 ml Eppendorf tube. After this time, 400 μl of the MG-AM-T20 solution was added to halt the reaction and to induce a color change, followed one minute later by the addition of 50 μl of color stabilizing 34% (w/v) sodium citrate. Photometry was performed in polystyrene flat-bottom microplates, measuring absorbance at 595 nm. For all enzyme assays, a phosphate calibration curves were produced for the conversion of absorbance readings to units of specific activity. Exceptions to these assay conditions are as follows. For assays varying inhibitor AMP concentrations, 0.180 m*M* FBP was present in assay mixtures. For assays varying $MgCl_2$ concentration, the protocol was slightly modified since the inorganic pyrophosphatase requires Mg^{2+}. These assays were performed in thin walled glass test tubes instead of Eppendorf tubes. After allowing the reaction to proceed for 10 minutes at 37°C, the reactions were halted by transferring the glass tubes to a boiling water bath for 1 minute, and they were subsequently placed on ice for 1 minute. Tubes were then transferred to a 37°C water bath at which point 25 μl of a solution containing 50 m*M* Hepes pH 8.0, 15.0 m*M* $MgCl_2$, 0.2 mg/ml BSA and 10.0 U/ml pyrophosphatase was added. After 30 seconds, the solution was mixed by pipetting, and after another 30 seconds, 400 μl of the MG-AM-T20 solution was added, and the previously described protocol was resumed.

Data was plotted as specific enzyme activity (U/mg) versus substrate or effector concentration using the program QTI-plot. Kinetic constants of a modified Hill equation[11,37] $V = V_0 + [(V_m - V_0) \times C^n/(B^n + C^n)]$, where V_0 is the velocity in absence of the substrate or effector being analyzed, V_m is the maximal velocity with saturating effector, C is the concentration of substrate or effector under study, B is the concentration of substrate or effector producing half of the maximal velocity ($S_{0.5}$), activation ($A_{0.5}$) or inhibition ($I_{0.5}$), and n is the Hill coefficient. Standard deviations were calculated with the same software. The "fold" change in activity that is reported for the assays is defined as V_m/V_0. Values reported are the mean of at least two to three independent sets of data which agree within 10% of each other.

Sequence alignment of ADP-Glc PPases and structural alignment of NDP-Glc PPases

The sequences of known plant and bacterial ADP-Glc PPases were downloaded from the Uniprot website (www.Uniprot.org). Redundant sequences were removed, yielding 242 unique sequences. The UniRef100 accession ID numbers[38] are provided in the Supporting Information Table S4. The software used for the alignments was Unipro Ugene and the

MUSCLE algorithm was used for the fitting.[39] The sequence alignment figure was produced from the alignment data using QTIplot v.0.9.8.9. The structural image showing the overlap of NDP-Glc PPases was made using Pymol 1.3, and includes the structures of the *E. coli* ADP-Glc PPase and the crystal structure of the *E. coli* UDP-Glc PPase, *E. coli* dTDP-Glc PPase, *Salmonella enterica* CDP-Glc PPase, and *Helicobacter pylori* UDP-Glc PPase (PDB accession codes 2E3D, 1H5T, 1WVC, and 3JUK, respectively.)[30–33]

Computational methods

The ATP bound *E. coli* ADP-Glc PPase tetramer exists as a dimer of dimers. Each dimer has one ATP-bound and one unbound subunit.[4,6] A homology model of a dimer, based on the crystal structures of *Agrobacterium tumefaciens*[7] and the potato tuber[6] ADP-Glc PPases was obtained as previously reported.[11] The ATP molecule was inherited by modeling from the potato tuber template (1YP3) using Modeler 9.11.[11,40] The adenine ring and ribose ring coordinates were not modified, but two of the phosphates were remodeled as follows. Since the β- and γ-phosphate of ATP in the potato tuber enzyme are not in a catalytic position (they partially overlap with the Glc1P site),[6] they were reconciled to mimic the conformation observed in other sugar nucleotide pyrophosphorylases. This was performed using the structure of TDP-Glc pyrophosphorylase[41] (1MC3) as template, as was previously described when modeling ATP into ADP-Glc PPases.[7,9] The dimer was duplicated here (by rotation about the two-fold axis) to create the tetramer used for MD studies. Using the program NAMD 2.9[42] with the CHARMM 27 forcefield[43] and the NPT ensemble, molecular dynamic simulations were performed as previously described with few modifications.[44] Following equilibration, a 1.5 fs timestep and a Langevin dampening coefficient of 2.0 ps^{-1} were used, and the SHAKE algorithm was applied to all covalent hydrogen bonds. Data from the last 15 ns of the 18 ns simulation was used in all analysis. Unless otherwise noted, images in this report show the enzyme 3 ns into the simulation, and all statements made on the enzyme structure pertain to both dimers. Analysis of the simulation data was performed using VMD.[45]

Correlated movement analysis of the protein backbone was performed as previously described[17] with few modifications. In conjunction with VMD,[45,46] the program Carma[47] was used to produce a cross correlation matrix, with correlation scores produced for each pairwise interaction. Correlation scores (C_{ij}) were defined by $C_{ij} = \langle \Delta r_i(t) \cdot \Delta r_j(t) \rangle / (\langle \Delta r_i(t)^2 \rangle \langle \Delta r_j(t)^2 \rangle)^{1/2}$, with $\Delta r_i(t) = \Delta r_i(t) - \langle \Delta r_i(t - \Delta t) \rangle$ and $r_i(t)$ denoting positional vector of node i at time t, and $\langle ... \rangle$ indicating average values over the course of the simulation. A map showing the differences in correlated movement between ATP-bound and unbound subunits was made by averaging the correlation data from the two ligand free subunits and subtracting it from the averaged correlation data of the two ATP bound subunits. The correlation map image shown in this report was prepared using QTIplot v.0.9.8.9. For the pathway analysis,[18] the programs *subopt* and VMD[45,46] were used for the determination of network pathways linking ATP and the FBP binding Lys39. A network is defined as a set of nodes. Each amino acid of the enzyme was represented by a single node. The ATP ligand was represented by three nodes; the adenine ring, ribose ring, and phosphate tail. Nodes are said to be connected if they were within 4.5 Å of each other for at least 75% of the simulation.[46] Any contacts between sequential (i.e., covalently bonded) nodes were not considered in the pathway analysis so as to ignore trivial network connections. The length (w) connecting nodes i and j were weighted based on the correlated movement analysis (described above) with $w = -\log(|C_{ij}|)$. The path length between two distant nodes is defined as the sum of the weighted lengths traversed when connecting them. The shortest (optimal) pathway is reported as the dominant mode of communication between the nodes.[18,46] The significance of signaling pathways was evaluated by plotting the distribution of path lengths (Supporting Information Figs. S1-II, S1-III) as previously described.[48]

Acknowledgments

B.L.H. was a recipient of the Advanced Doctoral Fellowship (Loyola University Chicago). J.W., F.H., B.M., and T.M. received the Mulcahy Scholarship (Loyola University Chicago). P.S. was supported by a Research Experiences for Undergraduates grant.

References

1. Espada J (1962) Enzymic synthesis of adenosine diphosphate glucose from glucose 1-phosphate and adenosine triphosphate. J Biol Chem 237:3577–3581.
2. Haugen TH, Preiss J (1979) Biosynthesis of bacterial glycogen. The nature of the binding of substrates and effectors to ADP-glucose synthase. J Biol Chem 254: 127–136.
3. Kleczkowski LA, Villand P, Preiss J, Olsen OA (1993) Kinetic mechanism and regulation of ADP-glucose pyrophosphorylase from barley (*Hordeum vulgare*) leaves. J Biol Chem 268:6228–6233.
4. Ballicora MA, Iglesias AA, Preiss J (2003) ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. Microbiol Mol Biol Rev 67:213–225.
5. Ballicora MA, Iglesias AA, Preiss J (2004) ADP-glucose pyrophosphorylase: a regulatory enzyme for plant starch synthesis. Photosynth Res 79:1–24.
6. Jin X, Ballicora MA, Preiss J, Geiger JH (2005) Crystal structure of potato tuber ADP-glucose pyrophosphorylase. EMBO J 24:694–704.

7. Cupp-Vickery JR, Igarashi RY, Perez M, Poland M, Meyer CR (2008) Structural analysis of ADP-glucose pyrophosphorylase from the bacterium *Agrobacterium tumefaciens*. Biochemistry 47:4439–4451.
8. Kuhn ML, Falaschetti CA, Ballicora MA (2009) Ostreococcus tauri ADP-glucose pyrophosphorylase reveals alternative paths for the evolution of subunit roles. J Biol Chem 284:34092–34102.
9. Ballicora MA, Dubay JR, Devillers CH, Preiss J (2005) Resurrecting the ancestral enzymatic role of a modulatory subunit. J Biol Chem 280:10189–10195.
10. Kuhn ML, Figueroa CM, Iglesias AA, Ballicora MA (2013) The ancestral activation promiscuity of ADP-glucose pyrophosphorylases from oxygenic photosynthetic organisms. BMC Evol Biol 13:51.
11. Figueroa CM, Esper MC, Bertolo A, Demonte AM, Aleanzi M, Iglesias AA, Ballicora MA (2011) Understanding the allosteric trigger for the fructose-1,6-bisphosphate regulation of the ADP-glucose pyrophosphorylase from *Escherichia coli*. Biochimie 93: 1816–1823.
12. Ballicora MA, Sesma JI, Iglesias AA, Preiss J (2002) Characterization of chimeric ADPglucose pyrophosphorylases of *Escherichia coli* and Agrobacterium tumefaciens. Importance of the C-terminus on the selectivity for allosteric regulators. Biochemistry 41: 9431–9437.
13. Gomez-Casati DF, Igarashi RY, Berger CN, Brandt ME, Iglesias AA, Meyer CR (2001) Identification of functionally important amino-terminal arginines of *Agrobacterium tumefaciens* ADP-glucose pyrophosphorylase by alanine scanning mutagenesis. Biochemistry 40:10169–10178.
14. Gardiol A, Preiss J (1990) *Escherichia coli* E-39 ADPglucose synthetase has different activation kinetics from the wild-type allosteric enzyme. Arch Biochem Biophys 280:175–180.
15. Figueroa CM, Kuhn ML, Falaschetti CA, Solamen L, Olsen KW, Ballicora MA, Iglesias AA (2013) Unraveling the activation mechanism of the potato tuber ADP-glucose pyrophosphorylase. PLoS One 8: e66824.
16. Tsai CJ, Del Sol A, Nussinov R (2009) Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. Mol Biosyst 5:207–216.
17. Kormos BL, Baranger AM, Beveridge DL (2006) Do collective atomic fluctuations account for cooperative effects? Molecular dynamics studies of the U1A-RNA complex. J Am Chem Soc 128:8992–8993.
18. Ghosh A, Vishveshwara S (2007) A study of communication pathways in methionyl- tRNA synthetase by molecular dynamics simulations and structure network analysis. Proc Natl Acad Sci USA 104:15711–15716.
19. Gentner N, Preiss J (1968) Biosynthesis of bacterial glycogen. VI. Differences in the kinetic properties of the *Escherichia coli* B adenosine diphosphate glucose pyrophosphorylase depending on whether Mg++ or Mn++ serves as divalent cation. J Biol Chem 243: 5882–5891.
20. Changeux JP, Edelstein S (2011) Conformational selection or induced fit? 50 years of debate resolved. F1000 Biol Rep 3:19.
21. Preiss J (1996) ADPglucose pyrophosphorylase: basic science and applications in biotechnology. Biotechnol Annu Rev 2:259–279.
22. Lee YM, Preiss J (1986) Covalent modification of substrate-binding sites of *Escherichia coli* ADP-glucose synthetase. Isolation and structural characterization of 8-azido-ADP-glucose-incorporated peptides. J Biol Chem 261:1058–1064.
23. Kumar A, Tanaka T, Lee YM, Preiss J (1988) Biosynthesis of bacterial glycogen. Use of site-directed mutagenesis to probe the role of tyrosine 114 in the catalytic mechanism of ADP-glucose synthetase from *Escherichia coli*. J Biol Chem 263: 14634–14639.
24. Tuncel A, Kavakli IH, Keskin O (2008) Insights into subunit interactions in the heterotetrameric structure of potato ADP-glucose pyrophosphorylase. Biophys J 95:3628–3639.
25. Cross JM, Clancy M, Shaw JR, Greene TW, Schmidt RR, Okita TW, Hannah LC (2004) Both subunits of ADP-glucose pyrophosphorylase are regulatory. Plant Physiol 135:137–144.
26. Ballicora MA, Erben ED, Yazaki T, Bertolo AL, Demonte AM, Schmidt JR, Aleanzi M, Bejar CM, Figueroa CM, Fusari CM, Iglesias AA, Preiss J (2007) Identification of regions critically affecting kinetics and allosteric regulation of the *Escherichia coli* ADP-glucose pyrophosphorylase by modeling and pentapeptide-scanning mutagenesis. J Bacteriol 189: 5325–5333.
27. Meyer CR, Yirsa J, Gott B, Preiss J (1998) A kinetic study of site-directed mutants of *Escherichia coli* ADP-glucose pyrophosphorylase: the role of residue 295 in allosteric regulation. Arch Biochem Biophys 352:247–254.
28. Meyer CR, Bork JA, Nadler S, Yirsa J, Preiss J (1998) Site-directed mutagenesis of a regulatory site of *Escherichia coli* ADP-glucose pyrophosphorylase: the role of residue 336 in allosteric behavior. Arch Biochem Biophys 353:152–159.
29. Bejar CM, Ballicora MA, Iglesias AA, Preiss J (2006) ADPglucose pyrophosphorylase's N-terminus: structural role in allosteric regulation. Biochem Biophys Res Commun 343:216–221.
30. Thoden JB, Holden HM (2007) The molecular architecture of glucose-1-phosphate uridylyltransferase. Protein Sci 16:432–440.
31. Zuccotti S, Zanardi D, Rosano C, Sturla L, Tonetti M, Bolognesi M (2001) Kinetic and crystallographic analyses support a sequential-ordered bi bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidylyltransferase. J Mol Biol 313:831–843.
32. Koropatkin NM, Cleland WW, Holden HM (2005) Kinetic and structural analysis of alpha-D-Glucose-1-phosphate cytidylyltransferase from Salmonella typhi. J Biol Chem 280:10774–10780.
33. Kim H, Choi J, Kim T, Lokanath NK, Ha SC, Suh SW, Hwang HY, Kim KK (2010) Structural basis for the reaction mechanism of UDP-glucose pyrophosphorylase. Mol Cells 29:397–405.
34. Diez MD, Ebrecht AC, Martinez LI, Aleanzi MC, Guerrero SA, Ballicora MA, Iglesias AA (2013) A chimeric UDP-glucose pyrophosphorylase produced by protein engineering exhibits sensitivity to allosteric regulators. Int J Mol Sci 14:9703–9721.
35. Fusari C, Demonte AM, Figueroa CM, Aleanzi M, Iglesias AA (2006) A colorimetric method for the assay of ADP-glucose pyrophosphorylase. Anal Biochem 352: 145–147.
36. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685.
37. Hill AV (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J Physiol (Lond) 40:4–7.

38. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics 23:1282–1288.
39. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.
40. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815.
41. Sivaraman J, Sauve V, Matte A, Cygler M (2002) Crystal structure of *Escherichia coli* glucose-1-phosphate thymidylyltransferase (RffH) complexed with dTTP and Mg2+. J Biol Chem 277:44214–44219.
42. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802.
43. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, MacKerell AD (2010) CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J Comput Chem 31:671–690.
44. Kim Y, Cunningham MA, Mire J, Tesar C, Sacchettini J, Joachimiak A (2013) NDM-1, the ultimate promiscuous enzyme: substrate recognition and catalytic mechanism. FASEB J 27:1917–1927.
45. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph Model 14:33–38.
46. Sethi A, Eargle J, Black AA, Luthey-Schulten Z (2009) Dynamical networks in tRNA: protein complexes. Proc Natl Acad Sci USA 106:6620–6625.
47. Glykos NM (2006) Software news and updates—Carma: a molecular dynamics analysis program. J Comput Chem 27:1765–1768.
48. Van Wart AT, Durrant J, Votapka L, Amaro RE (2014) Weighted implementation of suboptimal paths (WISP): an optimized algorithm and tool for dynamical network analysis. J Chem Theory Comput 10:511–517.