# The ROK kinase *N*-acetylglucosamine kinase uses a sequential random enzyme mechanism with successive conformational changes upon each substrate binding

Sumita Roy[1],[*] , Mirella Vivoli Vega[1] , Jessica R. Ames[1] , Nicole Britten[1], Amy Kent[1], Kim Evans[1],
Michail N. Isupov[2], and Nicholas J. Harmer [1],[*] , On behalf of the GoVV Consortium
*From the [1]Living Systems Institute, Exeter, UK; [2]Henry Wellcome Building for Biocalalysis, Biosciences, Exeter, UK*

*N*-acetyl-D-glucosamine (GlcNAc) is a major component of bacterial cell walls. Many organisms recycle GlcNAc from the cell wall or metabolize environmental GlcNAc. The first step in GlcNAc metabolism is phosphorylation to GlcNAc-6-phosphate. In bacteria, the ROK family kinase *N*-acetylglucosamine kinase (NagK) performs this activity. Although ROK kinases have been studied extensively, no ternary complex showing the two substrates has yet been observed. Here, we solved the structure of NagK from the human pathogen *Plesiomonas shigelloides* in complex with GlcNAc and the ATP analog AMP-PNP. Surprisingly, *Ps*NagK showed distinct conformational changes associated with the binding of each substrate. Consistent with this, the enzyme showed a sequential random enzyme mechanism. This indicates that the enzyme acts as a coordinated unit responding to each interaction. Our molecular dynamics modeling of catalytic ion binding confirmed the location of the essential catalytic metal. Additionally, site-directed mutagenesis confirmed the catalytic base and that the metal-coordinating residue is essential. Together, this study provides the most comprehensive insight into the activity of a ROK kinase.

*N*-acetylglucosamine (GlcNAc) is a critical monosaccharide for both prokaryotes and eukaryotes. Eukaryotes widely employ GlcNAc in the *N*- and *O*-linked glycans that decorate protein surfaces; in the glycosaminoglycans hyaluronan, heparin sulfate, and keratan sulfate that form a major part of the connective tissues (1, 2); and in chitin (3). GlcNAc is also used as a reversible modification of proteins (4) that is conserved among metazoans and to decorate some growth factors (5). This modification is particularly common on nuclear proteins and generally acts to modulate signaling (often in competition with phosphorylation) and transcription in response to stress and nutrient conditions (6–8).

GlcNAc is essential to most prokaryotes, as the cell wall is formed from a polymer of GlcNAc and *N*-acetylmuramic acid cross-linked with peptides (9). Consequently, the key enzymes required for the biosynthesis of the nucleotide-linked sugar UDP-GlcNAc are essential in all bacteria. Many bacteria also require GlcNAc to form their lipopolysaccharides (with GlcNAc forming the core of lipid A) (10) and capsular polysaccharides (11). The first sugar added to the lipid carrier in many oligosaccharides is GlcNAc, its epimer *N*-acetylgalactosamine, or 6-deoxy versions of these (*N*-acetyl-D-quinovosamine and *N*-acetyl-D-fucosamine, respectively) (10, 12, 13). The wzx flippase that transfers oligosaccharides from the cytoplasmic leaflet of the inner membrane into the periplasm (14, 15) and the wzy O-antigen/capsular polysaccharides polymerase (16) have strong specificity for the membrane proximal sugar. Furthermore, most oligosaccharide transferases (17, 18) are exquisitely specific for the *N*-acetyl group, making the *N*-acetylated sugars intimately linked to the surface biology of bacteria.

GlcNAc is generally synthesized by cells from fructose-6-phosphate from central metabolism (Fig. 1) (19, 20). However, many organisms also have pathways for recycling GlcNAc. This is of particular importance for many bacteria that remodel their cell wall and for intracellular bacteria that have a reduced availability of metabolic precursors in their environmental niches. Loss of the recycling pathway enzymes reduces the capacity of bacteria to remodel their cell walls (21–24). These pathways have been recognized in a wide range of human pathogens (*e.g.*, *Escherichia coli* (22), *Pseudomonas aeruginosa* (25), Enterobacteriaceae, *Staphylococcus aureus* (26, 27), and *Mycobacterium tuberculosis* (28)). Many bacteria utilize chitin as a nutrition resource, using chitinases to recycle it to GlcNAc (29, 30). Chitin is likely to be of particular importance in pathogens of crustaceans and insects (*e.g.*, *Serratia* (31) and *Vibrio* species (32)).

An essential step in GlcNAc metabolism is the phosphorylation of GlcNAc to GlcNAc-6-phosphate (GlcNAc-6P). Eukaryotes isomerize this to GlcNAc-1-phosphate (33, 34) (Fig. 1), as their preferred metabolic route to UDP-GlcNAc. In contrast, bacteria that recycle GlcNAc deacetylate

* For correspondence: Sumita Roy, S.Roy@exeter.ac.uk; Nicholas J. Harmer, N.J.Harmer@exeter.ac.uk.
Present addresses for: Mirella Vivoli Vega, School of Biochemistry, University of Bristol, Bristol BS8 1TD, UK; Jessica R. Ames, H.H. Wills Physics Building, School of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK.
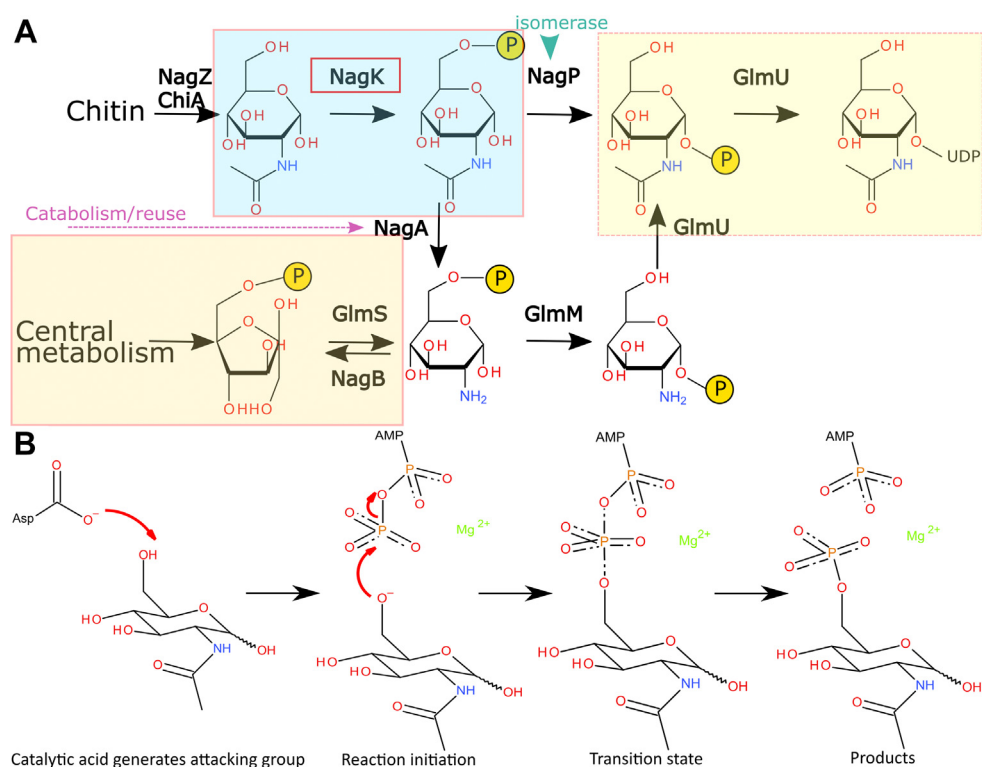
**Figure 1. Biosynthesis and catabolism of *N*-acetylglucosamine.** *A*, cells from bacteria to humans synthesize *N*-acetylglucosamine from fructose-6-phosphate derived from the Embden–Meyerhof–Parnas pathway (*orange shaded box*). Some organisms utilize GlcNAc from the environment (*e.g.*, digested chitin, bacterial cell wall components, or glycosaminoglycans). GlcNAc is converted into GlcNAc-6-phosphate by NagK enzymes (*sky blue shaded box*). GlcNAc-6-P can be deacetylated by NagA (*pink dashed arrow*) for catabolism or reuse or isomerized to GlcNAc-1-phosphate by NagP (*green arrowhead*) for direct transfer to UDP (*yellow shaded box*). *B*, mechanism of ROK kinases. ROK kinases are proposed to catalyze phosphate transfer with an aspartic acid residue acting as a general base to deprotonate a hydroxyl. This hydroxyl attacks the ATP γ-phosphate. The transition state (third image) is stabilized by a catalytic cation (here, $Mg^{2+}$). GlcNAc, *N*-acetylglucosamine; NagK, *N*-acetylglucosamine kinase.

GlcNAc-6P, linking recycled and environmental GlcNAc to their central metabolism (35). Phosphorylation of GlcNAc to GlcNAc-6P is performed by a specific kinase, *N*-acetylglucosamine kinase (NagK). Both mammalian (36) and bacterial NagK enzymes belong to the ROK kinase family of carbohydrate kinases (37). This family phosphorylates a broad range of sugars, with individual kinases showing tight specificity for their substrates (38–41). ROK kinases have a two-domain fold, with the sugar binding between the two domains, causing a structural re-arrangement that forms the active site (42, 43). Other characterized ROK kinases have shown a requirement for either manganese or magnesium for catalysis (40, 44, 45). Existing crystal structures suggest that ROK kinases use a similar mechanism to other classes of carbohydrate kinases (37) (Fig. 1*B*). A conserved aspartic acid side chain deprotonates the 6′-hydroxyl of GlcNAc. This hydroxyl attacks the ATP γ-phosphate, passing through a negatively charged transition state that is stabilized by the catalytic metal. However, current structural information does not include a structure of an ATP analog with an intact γ-phosphate. There is only one structure (from the human *N*-acetylmannosamine kinase [NanK]) that contains a catalytic metal: the metal binding site has not been confirmed by mutations or in bacterial enzymes (36, 46).

Here, we report the activity, structure, and mechanism of NagK from *Plesiomonas shigelloides*. Surprisingly, the enzyme displays a random sequential mechanism, with both GlcNAc and ATP able to bind to the enzyme first. PsNagK showed activity with magnesium and manganese as divalent cofactors. The structure of PsNagK in complex with GlcNAc and the ATP analog AMP-PNP demonstrates how the enzyme catalyzes phosphorylation of GlcNAc. Molecular dynamics simulations allowed us to confirm the location of the catalytic cation binding site. Comparing the ternary complex to the product complex of NagK bound to GlcNAc-6P highlights a possible catalytic mechanism. This provides, for the first time, a comprehensive kinetic and structural characterization of a ROK kinase.

## Results

### NagK activity from divergent species

The enzymatic activity of NagK has previously been described for *E. coli* (47). We determined the activity for a wider range of enzymes to highlight the diversity in activity from different species. We particularly focused on human pathogens with diverse NagK sequences. AlphaFold structures of these orthologs suggest that they have very similar structures (48–50). Recombinant NagK was readily purified for a range of human pathogens (Figs. S1 and S2). The enzymes showed a range of activities (Table 1), with NagK from *Photobacterium damselae* showing the highest activity.

**Table 1**
**Activity of NagK orthologs from diverse bacteria**

| Species | $k_{cat}$ (s$^{-1}$) | $K_{M\ app\ GlcNAc}$ (µM) | $K_{M\ app\ ATP}$ (µM) | Identity to *E. coli* NagK (%) | Enzyme, GlcNAc, and ATP concentrations used (nM, mM, mM) |
|---|---|---|---|---|---|
| *Vibrio vulnificus* | 95 ± 8 | 210 ± 60 | 41 ± 9 | 54.5 | 6.3, 2, 1 |
| *Photobacterium damselae* | 400 ± 20 | 64 ± 10 | 250 ± 10 | 53.8 | 4.23, 2, 1 |
| *Pseudoalteromonas* sp. P1-8 | 25 ± 1 | 1100 ± 200 | 3000 ± 700 | 40 | 13.4, 25, 20 |
| *Plesiomonas shigelloides* | 202 ± 3 | 98 ± 6 | 290 ± 20 | 59.7 | 5.1, 2, 1 |

NagK orthologs from diverse bacteria were purified (Fig. S1) and their activity determined. $k_{cat}$ and $K_{M\ app}$ for both substrates were determined. The table indicates the concentration of NagK used, and the concentration of substrate used for determining $K_{M\ app}$ for the reciprocal substrate. The sequences were aligned using MUSCLE and percentage identities calculated in Geneious v. 8.

## NagK uses a sequential mechanism

We selected NagK from *P. shigelloides* for a more detailed study of the NagK mechanism. The enzyme kinetics showed a sequential mechanism rather than a ping-pong mechanism

(Fig. 2, *A–C*; *p* = 0.0045). The products GlcNAc-6P and ADP showed weak inhibition, with Morrison $K_i$ values one to two orders of magnitude higher than the cognate substrate $K_M$ (Fig. S3). These could not be used to differentiate between an
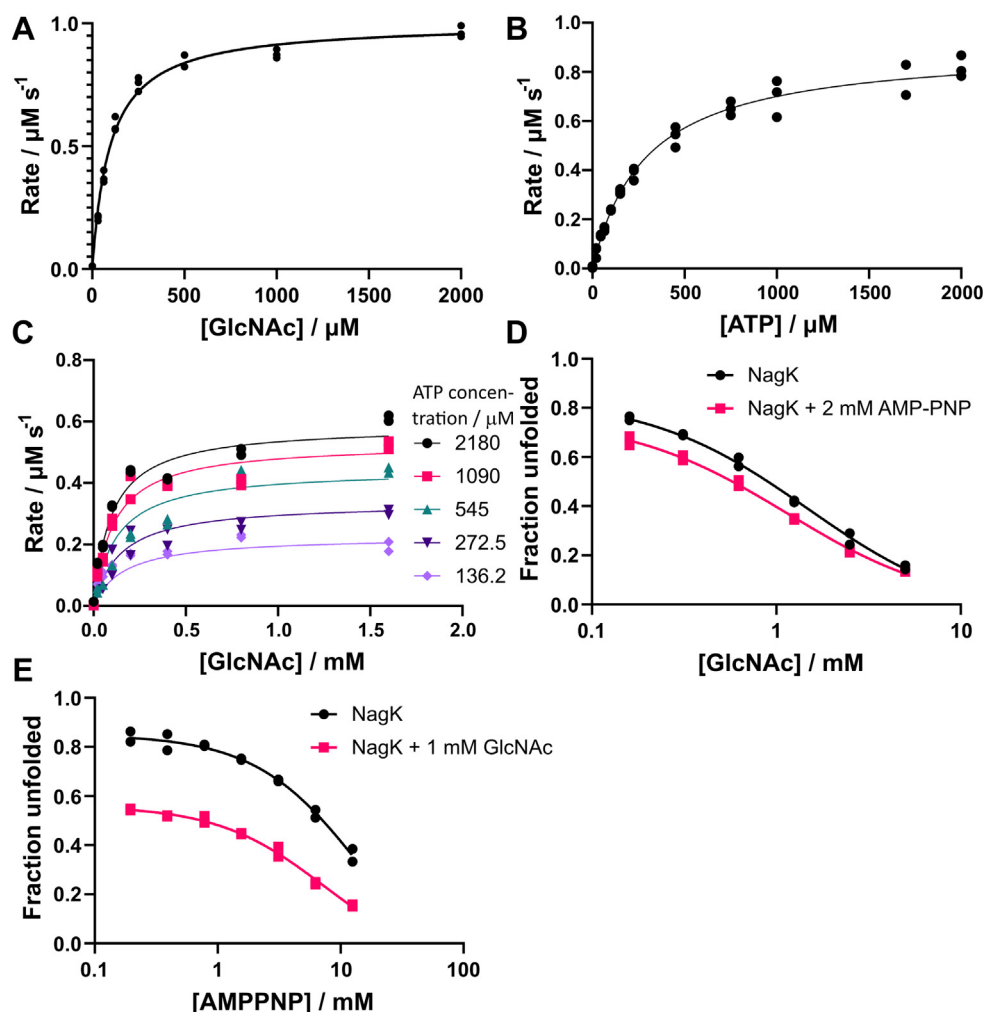


**Figure 2. NagK uses a sequential mechanism of binding.** NagK from *Plesiomonas shigelloides* was selected for a more detailed study of function as it has intermediate activity. *A* and *B*, PsNagK activity fitted to the Michaelis-Menten equation for both GlcNAc (*A*) and ATP (*B*), giving $K_{M\ app}$ values of 98 ± 6 µM and 290 ± 20 µM, respectively. Experiments used 5.1 nM PsNagK and 2 mM GlcNAc or 1 mM ATP to determine $K_{M\ app}$ for the reciprocal substrate. *C*, testing of both substrates together showed a strong preference to the equation for sequential binding rather than a ping-pong mechanism (Akaike's information criteria difference = 10.79; *p* = 0.0045). Neither GlcNAc-6-P nor ADP showed product inhibition at readily testable concentrations (Fig. S3), preventing determination of whether the binding is ordered or random. These experiments used 90 ng/ml PsNagK. *D* and *E*, differential scanning fluorimetry of NagK in the presence of GlcNAc and the ATP analog AMP-PNP. The apparent $K_D$ value of NagK for GlcNAc shows no significant difference in the absence (230 ± 20 µM) or presence (270 ± 20 µM) of 2 mM AMP-PNP. The apparent $K_D$ value of NagK for AMP-PNP increases slightly from 2.2 ± 0.6 mM to 3.2 ± 0.5 mM in the presence of 1 mM GlcNAc. All experiments show results representative of at least two experiments performed on different days, with three experimental replicates per datapoint for panels A and B, and two experimental replicates per datapoint for panels C-E. GlcNAc, *N*-acetylglucosamine; NagK, *N*-acetylglucosamine kinase.

ordered and random sequential mechanism. Examination of the relationship between $K'$ (apparent $K_M$) and the partner concentration can be diagnostic (51, 52). For our enzymes, these showed no decrease in $K'$ with increasing partner concentration (Fig. S4), consistent with a random sequential mechanism. However, counter examples showing this pattern that nevertheless have an ordered mechanism have been described (53, 54). We therefore examined whether the binding of either substrate affects binding of the other. Past studies of enzyme mechanisms have investigated substrate binding using methods such as differential scanning fluorimetry (55) or fluorescence anisotropy (56). We used differential scanning fluorimetry to determine the dissociation constants of GlcNAc and the nonhydrolyzable ATP analog AMP-PNP. Using the isothermal differential scanning fluorimetry approach (57), we determined that the $K_D$ for GlcNAc in the absence and presence of AMP-PNP were $230 \pm 20$ μM and $270 \pm 20$ μM, respectively (Fig. 2D and Fig. S5). We chose a temperature of 68 °C to measure at as this gave the optimal signal to determine $K_D$. The $K_D$ for AMP-PNP in the absence and presence of GlcNAc were $2.2 \pm 0.6$ mM and $3.2 \pm 0.5$ mM, respectively (Fig. 2E). The difference between these $K_D$ values and the apparent $K_M$ values is likely due to the higher temperature at

which these data were collected. There is no significant increase in the affinity for either substrate in the presence of the other. This confirms that NagK likely uses a random sequential mechanism.

### NagK prefers magnesium as the catalytic metal

Most carbohydrate kinases require a metal cofactor. The ROK kinases particularly have previously shown a strong requirement for metals. As our coupled enzyme assay is also dependent on metals, we tested the coupling enzymes and the NagK reaction in the presence of the eight divalent cations observed in the M-CSA database (58) ($Mg^{2+}$, $Ca^{2+}$, $Mn^{2+}$, $Fe^{2+}$, $Co^{2+}$, $Ni^{2+}$, $Cu^{2+}$, and $Zn^{2+}$). NagK showed activity with all the cations that support coupling enzyme activity (Fig. S6). However, the NagK rate approaches zero at low $Co^{2+}$ concentrations, where the coupling enzymes retain activity (Fig. 3A). This suggests that *Ps*NagK has no activity in the absence of divalent cations. We tested all other relevant divalent cations with 10 μM $Co^{2+}$ to support the coupling enzymes (Table 2 and Fig. 3B). No activity was observed with calcium, copper, or zinc. The enzyme showed little preference between manganese ($K_{1/2} = 0.47 \pm 0.06$ mM) and magnesium ($K_{1/2} = 0.7 \pm$
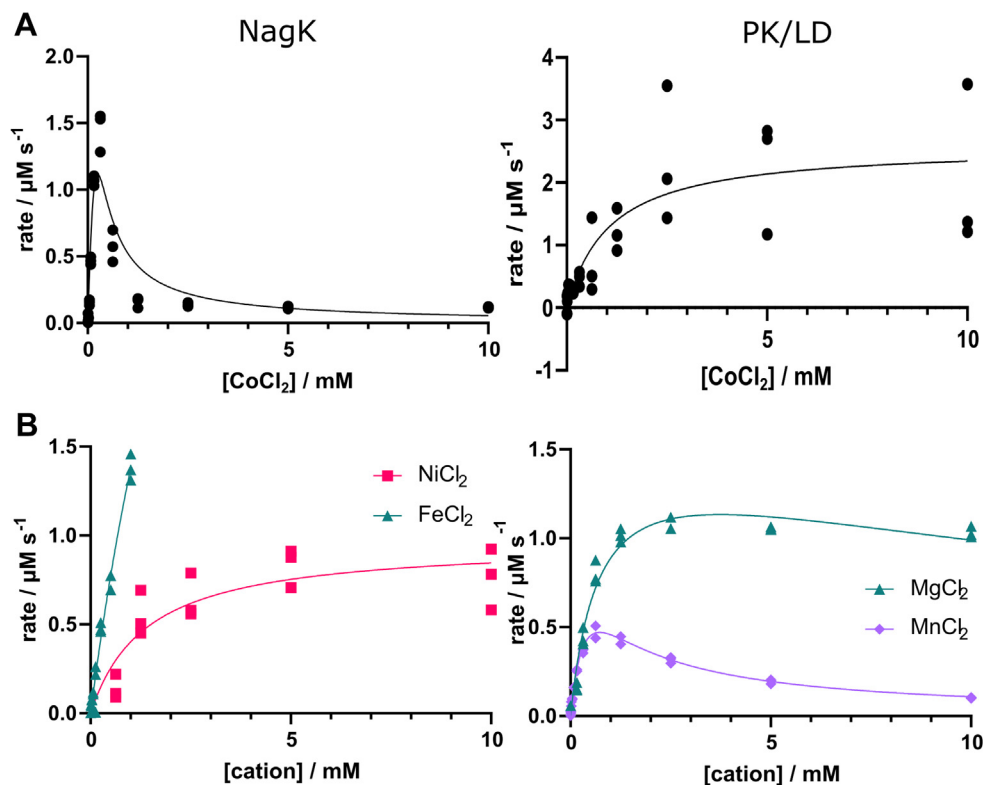


**Figure 3. Magnesium is the preferred metal cofactor of NagK.** The activity of *P. shigelloides* NagK was tested in the presence of diverse cofactors. *A*, cobalt supports the coupling enzymes pyruvate kinase and lactate dehydrogenase at concentrations that give minimal NagK activity. NagK (*left*) shows activity less than 10 times background at concentrations below 20 μM, while supporting much higher rates of the coupling enzyme (*right*) at these concentrations. Conditions used 2 mM GlcNAc, 1 mM ATP, 5.1 nM NagK; for the coupled reaction, 2 U/ml was used; for determination of coupling enzyme efficiency, 0.2 U/ml was used. No activity was observed without divalent cations (Fig. S6). *B*, testing of other cations showed that magnesium, manganese, iron, and nickel all support NagK activity. The saturating concentration for iron is not reached within the constraints of the assay (*left*). Both magnesium and manganese show a substrate inhibition effect (*right*). Nickel shows a maximum rate of $1.0 \pm 0.2$ μM s$^{-1}$, with a $K_{1/2}$ of $1.4 \pm 0.4$ mM. The maximum rates for magnesium and manganese are $1.1 \pm 0.1$ μM s$^{-1}$ and $0.47 \pm 0.04$ μM s$^{-1}$ at 3.7 mM and 0.73 mM, respectively. The inhibition constant $K_I$ is $20 \pm 6$ mM for magnesium and $1.1 \pm 0.1$ mM for manganese. Three experimental replicates were taken for each datapoint, and all experiments are representative of at least two experiments performed on different days. GlcNAc, *N*-acetylglucosamine; LD, lactate dehydrogenase; NagK, *N*-acetylglucosamine kinase; PK, pyruvate kinase.

**Table 2**
Effect of divalent cations on enzyme activity

| Metal | Maximum rate ($\mu M\ s^{-1}$) | $K_{1/2}$ (mM) | $K_I$ (mM) |
|---|---|---|---|
| $Mg^{2+}$ | 1.1 ± 0.1 | 0.7 ± 0.1 | 20 ± 6 |
| $Ca^{2+}$ | 0 | N/A[a] | N/A[a] |
| $Mn^{2+}$ | 0.47 ± 0.04 | 0.07 ± 0.01 | 1.1 ± 0.1 |
| $Fe^{2+}$ | Not determined | >>1 mM | N/A[a] |
| $Co^{2+}$ | 1.2 ± 0.2 | >10 | <0.05 |
| $Ni^{2+}$ | 1.0 ± 0.2 | 1.4 ± 0.4 | N/A[b] |
| $Cu^{2+}$ | 0 | N/A[a] | N/A[a] |
| $Zn^{2+}$ | 0 | N/A[a] | N/A[a] |

All experiments were performed using 180 ng/ml *P. shigelloides* NagK, 2 mM GlcNAc, and 1 mM ATP. Experiments for metals other than cobalt contained 10 μM CoCl₂ to support coupling enzyme activity. Data for iron and nickel were fitted to the Michaelis–Menten equation. Data for magnesium, manganese, and cobalt were fitted to the substrate inhibition equation. Data were fitted in GraphPad v. 9.4.1.
[a] This cation did not support NagK activity and so no constants were determined.
[b] This cation did not show evidence of an inhibitory effect and so no $K_I$ was determined.

0.1 mM) at low concentrations. Both showed substrate inhibition, with manganese inhibiting at lower concentrations ($K_i$ = 1.1 ± 0.1 mM; maximum rate 0.47 ± 0.04 $\mu M\ s^{-1}$ at 0.73 mM; Fig. 3B). Magnesium shows a higher maximum rate (1.1 ± 0.1 $\mu M\ s^{-1}$ at 3.7 mM) and would be strongly preferred at physiological concentrations (Fig. 3B). Both ferrous iron and nickel also supported NagK activity. Nickel supported activity moderately (maximum 1.0 ± 0.2 $\mu M\ s^{-1}$, $K_{1/2}$ = 1.4 ± 0.4 mM). The rate supported by ferrous iron reached 1.4 $\mu M\ s^{-1}$ at 1 mM and was clearly not saturating. In both cases, 1 mM represents a far higher concentration than would be found in a bacterial cell, again supporting magnesium as the physiological cation.

### The NagK active site is formed by enzyme closure around the GlcNAc and ATP substrates

Although a structure of *Vibrio vulnificus* NagK has been solved (59), there is no structure of a ligand-bound NagK. We therefore determined the structure of *P. shigelloides* NagK, as this crystallized readily with and without its substrates (Table 3). As expected, *Ps*NagK forms a two-domain fold with a large domain (including the structural zinc characteristic of ROK kinases (37)) and a small domain (Fig. 4A). The enzyme closes around the GlcNAc substrate, with the small domain rotating by 23° (moving up to 15 Å) relative to the large domain (Fig. 4B). The GlcNAc is bound specifically by the side chains of residues S78, N104, D105, E154, H157, and D187 (Fig. 4C).

We then soaked the ATP analog AMP-PNP into the *Ps*NagK structure. A structure with both GlcNAc and AMP-PNP shows the location of the γ-phosphate in a position poised for catalysis (Fig. 5A): the best previous ROK kinase ligand structures showed density only to the β-phosphate (42, 46). The small domain rotates a further 16° to engage the AMP-PNP (Fig. S7C). This suggests that the full catalytic complex is formed only when both substrates are bound and that there is a "two-step" closing of the gap between two domains on binding of each substrate. Consistent with this, binding of AMP-PNP alone caused only a small closing of the gap between the domains (Fig. S7D). In contrast, structures

solved with products showed a conformation similar to the ternary complex. Both the abortive complex of NagK bound to GlcNAc and ADP (Fig. S7E), and the complex of NagK with GlcNAc-6P (Fig. S7F) resemble the GlcNAc-AMP-PNP complex. AMP-PNP is held in place by the side chains of residues T10, D105, T132 and E196, with the phosphates being coordinated by the main chain of G9, T10, and G255 (Fig. 5B). Most of these side chains are well conserved amongst NagKs, consistent with a role in substrate binding (Fig. S8).

We were unable to obtain a structure that contained the catalytic cation. However, our ternary complex with GlcNAc and AMP-PNP is structurally very similar to the previously solved NanK structure that included a catalytic magnesium ((46); Fig. S9). The cation binding site is adjacent to a water molecule in our structure coordinated by D6, the main chain carbonyl of I7, and the γ-phosphate (Fig. 5C). To test the hypothesis that this is the metal binding site, we performed molecular dynamics simulations of the active site with divalent cations added in this location and AMP-PNP replaced by ATP. Molecular dynamics of the solved structure over 5 ns showed no significant changes in the structure, aside from a minor re-arrangement of the ATP phosphates (Fig. S10A). When magnesium, manganese, or calcium was added to the protein structure, the cation and ATP phosphates re-arrange to form a binding site for the divalent cation. Counterintuitively, in the cases of magnesium and manganese, the re-arrangement brings the cation close to the side chain of D105 and the GlcNAc 6'-O as well as the D6 side chain, I7 main chain carbonyl, and the γ-phosphate (Fig. S10, B, D). These cations are coordinated to five ligands as one face is partially blocked by the side chain of I127. In contrast, the calcium ion forms a classical octahedral coordination with the side chains of D105 and D6 (both oxygens), I7 main chain, and two oxygens from the ATP γ-phosphate. In this case, GlcNAc 6'-O is excluded from the coordination. This may reduce the acidity of the GlcNAc 6'-O, consistent with calcium not supporting catalysis. The rapid, reproducible re-arrangement of the active site under molecule dynamics strongly supports the hypothesis that this is the cation binding site. However, it is likely that a further re-arrangement of the enzyme active site is necessary for catalysis, as the ATP γ-phosphate remains too far away from GlcNAc 6'-O (4.3 Å) to support a reaction.

### Confirmation of proposed ligand-interacting residues by site-directed mutagenesis

Site-directed mutagenesis of proposed ligand-binding and catalytic residues support the role of these amino acids in *Ps*NagK activity. Mutation to either D105N (catalytic base) or D6N/A (metal coordinating negatively charged group) results in a loss of activity below the limit of detection (at least 1000-fold; Table 4). Mutation of the phosphate coordinating T10V and T132V results in a loss of activity, without substantially affecting the $K_M$ for either substrate. Mutation of some side chains that coordinate GlcNAc (N104D, E154Q or double mutant, and D187N) results in substantial increases in $K_M$ for

≋ASBMB

*J. Biol. Chem.* (2023) 299(4) 103033 **5**

**Table 3**
Crystal information, data collection, and refinement

| Project | NagK | NagK-GlcNAc | NagK-GlcNAc-ADP | NagK-GlcNAc AMP-PNP | NagK-GlcNAc-P | NagK-AMP-PNP |
|---|---|---|---|---|---|---|
| Data collection statistics | | | | | | |
| Beamline | I04 Diamond | I03 Diamond | I04−1 Diamond | I04 Diamond | I03 Diamond | I03 Diamond |
| Wavelength (Å) | 0.9795 | 0.9763 | 0.9159 | 0.9795 | 0.9763 | 0.9763 |
| Space group | $P3_121$ | $P3_221$ | $P3_221$ | $P3_221$ | $P3_221$ | $P6_5$ |
| Unit cell parameters a, b, c (Å) | 95.2, 95.2, 180.6 | 115.5, 115.5, 119.7.3 | 115.1, 115.1, 120.3 | 115.2, 115.2, 120.4 | 114.5, 114.5, 119.3 | 121.3, 121.3, 91.7 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90.0, 90.0, 120.0 | 90.0, 90.0120.0 | 90.0, 90.0, 120.0 | 90.0, 90.0, 120.0 | 90.0, 90.0, 120.0 | 90.0, 90.0, 120.0 |
| Resolution range (Å)[a] | 75.01−1.70 (1.73−1.70) | 57.74−1.94 (1.99−1.94) | 99.68−1.57 (1.60−1.57) | 57.60−2.11 (2.17−2.11) | 99.17−1.75 (1.78−1.75) | 60.63−2.20 (2.27−2.20) |
| Total reflections[a] | 521,556 (25,890) | 457,524 (29,499) | 1,273,708 (54,820) | 535,990 (43,641) | 346,458 (17,705) | 127,050 (10,922) |
| Unique reflections[a] | 104,319 (5073) | 68,568 (4548) | 128,365 (6335) | 53,533 (4354) | 90,389 (4505) | 38,637 (3361) |
| Completeness (%)[a] | 99.6 (99.7) | 100.0 (100.0) | 100.0 (100.0) | 100.0 (100.0) | 99.1, (100.0) | 99.2 (99.8) |
| Multiplicity[a] | 5.0 (5.1) | 6.7 (6.5) | 9.9 (8.7) | 10.0 (10.0) | 3.8 (3.9) | 3.3 (3.2) |
| $R_{merge}$ (%)[a,b] | 4.6 (262.5) | 8.6 (231.3) | 10.1 (321.7) | 17.3 (336.0) | 9.1 (169.1) | 17.8 (272.5) |
| $<I>/<\sigma(I)>$[a] | 13.8 (0.4) | 11.0 (0.8) | 12.7 (0.6) | 9.5 (0.7) | 8.4 (0.8) | 5.2 (0.4) |
| $CC_{1/2}$[a,c] | 0.999 (0.270) | 0.999 (0.282) | 0.999 (0.291) | 0.998 (0.336) | 0.998 (0.194) | 0.964 (0.396) |
| Wilson B-factor[d] (Å$^2$) | 41.3 | 48.5 | 33.4 | 49.9 | 35.1 | 48.9 |
| Refinement statistics | | | | | | |
| $R_{work}$ | 0.183 | 0.208 | 0.185 | 0.191 | 0.194 | 20.9 |
| $R_{free}$ | 0.210 | 0.250 | 0.212 | 0.233 | 0.223 | 25.2 |
| No. of protein monomers in a.u. | 2 | 2 | 2 | 2 | 2 | 2 |
| Number of atoms | | | | | | |
| Macromolecules | 4891 | 4679 | 4976 | 4741 | 4899 | 4666 |
| Ligands and metal ions | 2 | 34 | 86 | 96 | 42 | 64 |
| Solvent | 597 | 382 | 749 | 360 | 511 | 242 |
| Number of protein residues | 609 | 608 | 610 | 610 | 609 | 608 |
| RMS bond lengths (Å) | 0.009 | 0.009 | 0.007 | 0.007 | 0.012 | 0.006 |
| RMS bond angles (°) | 1.62 | 1.67 | 1.46 | 1.57 | 1.77 | 1.42 |
| Ramachandran favored (%)[e] | 97.9 | 94.4 | 98.0 | 96.2 | 97.9 | 96.2 |
| Ramachandran outliers (%)[e] | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.5 |
| Clashscore[e] | 4.05 | 12.6 | 8.8 | 6.9 | 7.0 | 6.3 |
| Average B-factor protein (Å$^2$) | 42.2 | 55.3 | 27.8 | 49.3 | 35.1 | 46.5 |
| Average B-factor ligands (Å$^2$) | 39.2 | 69.1 | 30.3 | 56.8 | 40.3 | 72.7 |
| Average B-factor solvent (Å$^2$) | 55.4 | 56.9 | 43.5 | 57.3 | 49.6 | 50.7 |
| RCBS PDB code | 7P7I | 7P9Y | 7P7W | 7P9P | 7P9L | 7PA1 |

[a] Values for the highest resolution shell are given in parentheses.
[b] $R_{merge} = \Sigma_h \Sigma_i |I_{h,i} - <I_h>|/\Sigma_h \Sigma_i I_{h,i}$.
[c] $CC_{1/2}$ is defined in (99).
[d] Wilson B-factor was estimated by SFCHECK (100).
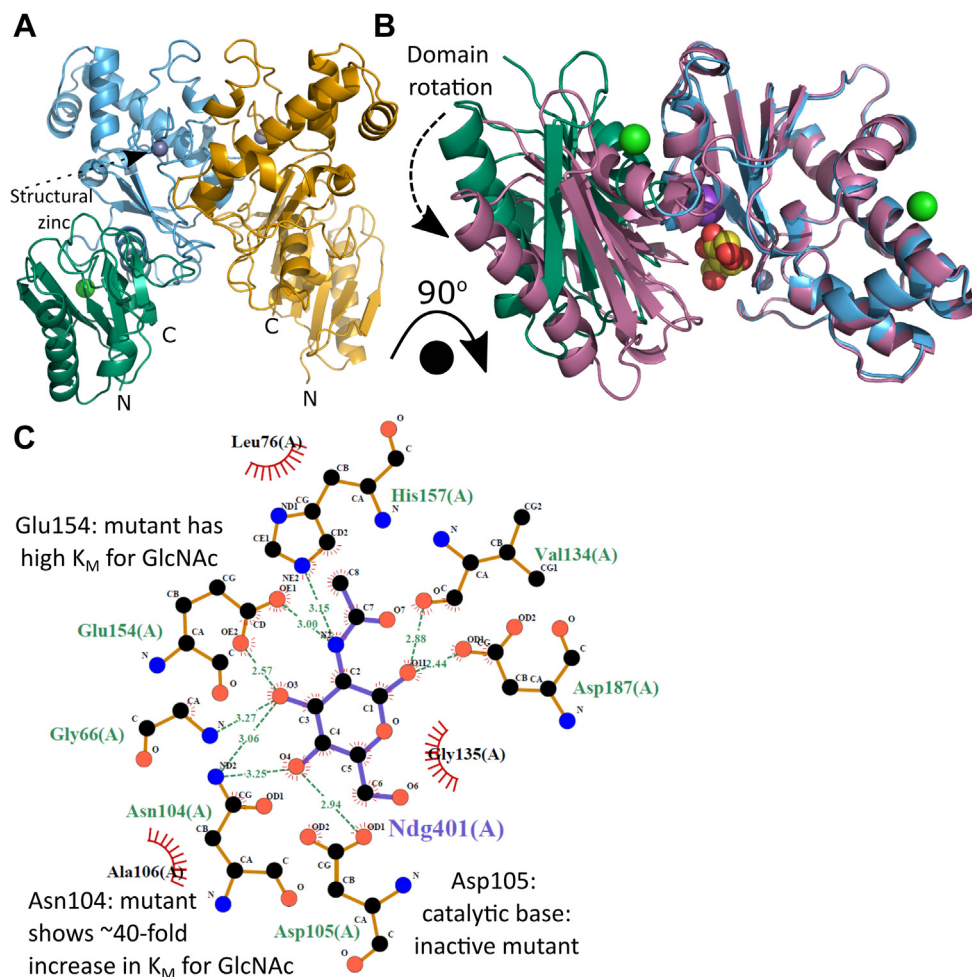[e] The Ramachandran statistics and clashscore were calculated using MOLPROBITY (89).

**Figure 4. NagK closes around the *N*-acetylglucosamine substrate.** *A*, overall structure of the NagK dimer (PDB: 7P7I). NagK has two domains: an N-terminal small domain (*green*) that includes the C-terminal helix and a large domain (*sky blue*). NagK forms a dimer (second molecule in *yellow*), with the interface between two large domains. A conserved structural zinc ion (*gray sphere*) is seen in the large domain (*dashed arrow*). The N and C termini are indicated. *B*, upon binding of the ligand *N*-acetylglucosamine (*spheres*, carbon atoms in *yellow*), the small domain rotates approximately 15° relative to the large domain to close around the sugar. Unbound structure (7P7I) colored as in A; bound structure (7P9Y) shown in *magenta*. Structures were superimposed over the large domain. *C*, GlcNAc is held in place by a network of amino acids from both small (residues 1–104, 291–303) and large (residues 105–290) domains (7P7Y). Structure images shown as cartoon with ligand atoms shown as *spheres*. Atom colors where not indicated: nitrogen, *blue*; oxygen, *yellow*; chloride, *green*, potassium, *purple*. Panels A and B generated using PyMOL v. 2.4.1 (27); panel C generated using LigPlot+ v2.2 (97, 98). NagK, *N*-acetylglucosamine kinase; GlcNAc, *N*-acetylglucosamine.

both substrates; for the E154Q mutants, the rate is substantially reduced. Mutation of other conserved GlcNAc binding residues S78A and E196Q resulted in clear increases in rate without affecting $K_M$. These two residues are not well conserved (Fig. S8), and the residues mutated to are found in other orthologs. We did not mutate H157 as this residue also coordinates to the structural zinc atom, and mutation would likely significantly affect the protein structure; all the tested mutants showed good stability in differential scanning fluorimetry (Fig. S12).

## Discussion

GlcNAc recycling from the cell wall is important for the biology of many human pathogens. These include some of the ESKAPE pathogens (60) of greatest concern for antimicrobial resistance (22, 25–27). To efficiently recycle cell wall GlcNAc, bacteria phosphorylate and then de-acetylate GlcNAc to form glucosamine-6-phosphate (35), an intermediate in the essential UDP-GlcNAc biosynthesis pathway (Fig. 1 and (19, 20)). Here, we have thoroughly characterized the first enzyme that performs the first of these steps, NagK. This enzyme belongs to the ROK kinase family of carbohydrate kinases (37). Key questions arising from previous studies of ROK kinases were the order of binding of substrates, confirming the location of the catalytic metal ion and the location of the γ-phosphate.

In common with previous ROK kinases, we determined that NagK has an absolute requirement for divalent cations (40, 45, 61). Magnesium, manganese, iron, cobalt, and nickel support NagK function, while calcium, copper, and zinc do not. Physiologically, magnesium would likely be preferred as bacterial intracellular free magnesium concentrations (~2 mM) exceed $K_{1/2}$ (0.7 mM), while free iron and manganese concentrations (1–15 μM) are well below concentrations where these support a high NagK rate, while only low micromolar
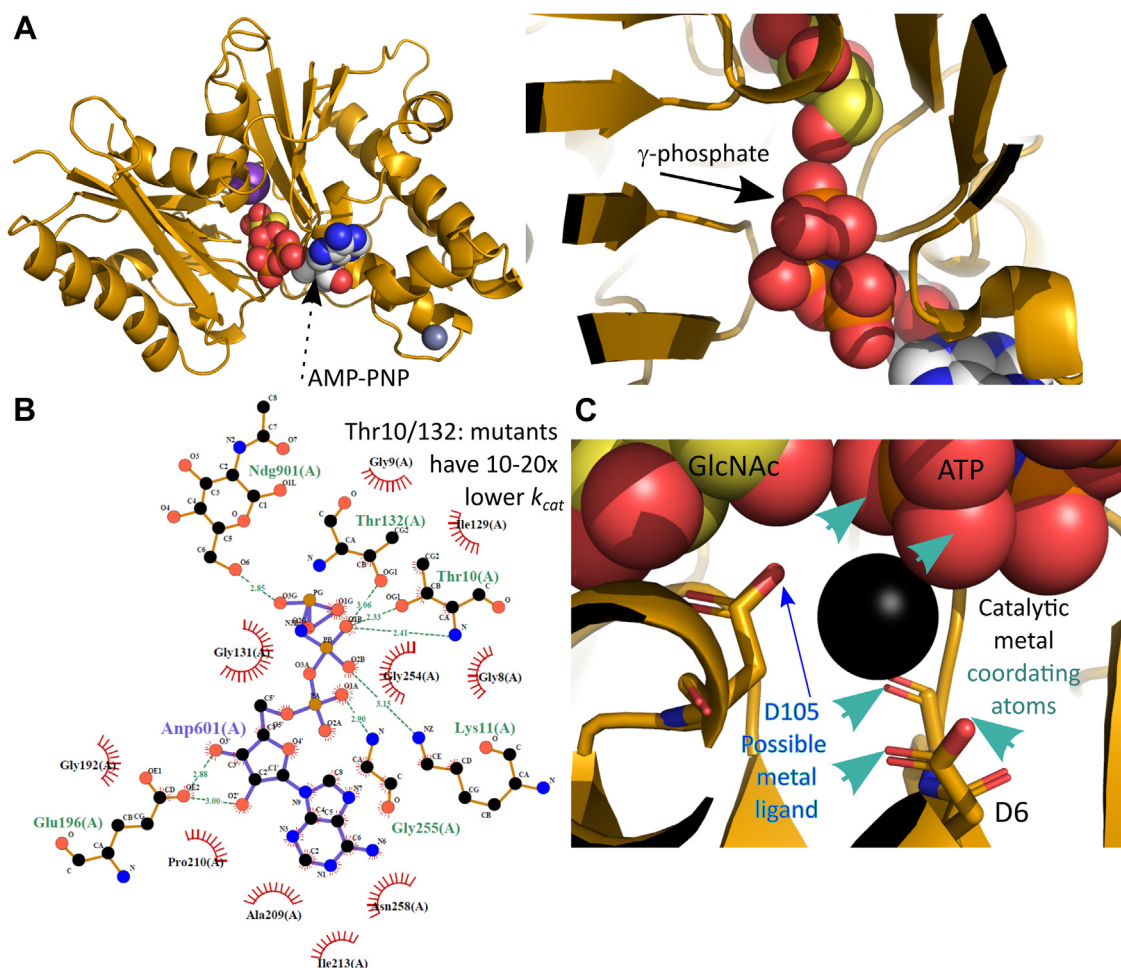
**Figure 5. The NagK ternary complex with GlcNAc and AMP-PNP highlights the likely catalytic mechanism.** *A*, structure of the NagK-GlcNAc-AMP-PNP ternary complex (PDB: 7P9P). *Left*: overview of the structure in the same conformation as Fig. 4*B*. The AMP-PNP (*black dashed arrow*) is shown as *spheres* with carbon atoms colored *white*. *Right*: close-up view of the interaction between the two ligands. The terminal phosphate is indicated with the *black arrow*. *B*, AMP-PNP is held in place by hydrophobic contacts to the adenine ring, hydrogen bonds from the ribose ring to E196, and interactions of the phosphate groups with T10, K11, T132, and the protein main chain. *C*, the ternary complex creates a metal binding site that is occupied by water in the structure. Likely water molecule shown as a *black sphere*. Likely metal coordinating atoms (acid oxygens of D6, main chain carbonyl of I7, γ-phosphate atoms) are indicated with *green arrowheads*; the side chain of D105 (*blue arrow*) may act as a sixth ligand. Atom colors where not indicated: nitrogen, *blue*; oxygen, *yellow*; chloride, *green*, potassium, *purple*; zinc, *gray*; phosphorus, *orange*. Panels A and C generated using PyMOL v. 2.4.1 (27); panel B generated using LigPlot+ v2.2 (97, 98). GlcNAc, *N*-acetylglucosamine; NagK, *N*-acetylglucosamine kinase.

concentrations of nickel and cobalt are tolerated by bacteria (62–65). Comparison of the crystal structure of NagK bound to GlcNAc and AMP-PNP to the human NanK structure (46) suggested that the metal ion should bind into a pocket adjacent to the γ-phosphate. This pocket would be coordinated by two oxygens from the γ-phosphate, the main chain carbonyl of I7, and the side chain of D6. An alignment of ROK kinases shows that D6 is strongly conserved as an acidic residue (Fig. S11). This has previously been proposed (albeit with limited evidence) as a metal ion binding residue (36). To support this proposal, we added a magnesium ion to this site in our structure and performed a molecular dynamics simulation. The maintenance of the ion in this location is strongly supported in the simulation, with both magnesium and manganese predicted to coordinate to both substrates. Furthermore, mutation of D6 to either asparagine or alanine completely abolishes the activity of the enzyme. Given that D6 is not close to either substrate in the crystal structure, this very strong

phenotype strongly supports a role in binding to the catalytic metal ion. These observations strongly support this pocket as the metal binding site for a wide range of ROK kinases.

The effect of mutations in GlcNAc binding residues is in accordance with previous studies. A detailed phylogenetic study proposed that the 3'-OH is coordinated by asparagine (N104) and glutamic acid (E154) residues (39). Mutations in either of these residues significantly reduced the activity of NagK. In contrast, two side chains that contact GlcNAc in the crystal structures (S87 and E196) are not evolutionarily conserved (Fig. S8). Mutation of these side chains increases the catalytic efficiency of NagK *in vitro*. The 1'-OH is engaged by an aspartic acid (D187), mutation of which reduces catalytic efficiency: this is not conserved in *Pseudoalteromonas* NagK and may explain the reduced activity of this ortholog. The highly conserved aspartic acid (here D105) coordinates the 6'-OH, in common with previous NagK structures (39, 41–43, 46). As was previously observed for human *N*-mannosamine

**Table 4**
**Effect of mutants of ATP and GlcNAc binding residues**

| *Ps*NagK variant | $k_{cat}/s^{-1}$ | $K_{M\ app\ GlcNAc}$ (µM) | $K_{M\ app\ ATP}$ (µM) | $k_{cat}/K_{M\ app\ GlcNAc}$ (µM$^{-1}$ s$^{-1}$) | $k_{cat}/K_{M\ app\ ATP}$ (µM$^{-1}$ s$^{-1}$) | Enzyme, GlcNAc, and ATP concentrations used (nM, mM, mM) |
|---|---|---|---|---|---|---|
| Wildtype | 202 ± 3 | 98 ± 6 | 290 ± 20 | 2.1 ± 0.1 | 0.70 ± 0.05 | 5.1, 2, 1 |
| D6A | N/A[a] | | | | | Attempted up to 82 nM with 2 mM GlcNAc, 1 mM ATP |
| D6N | N/A[a] | | | | | As for D6A |
| T10V | 12.5 ± 0.2 | 157 ± 10 | 420 ± 50 | 0.080 ± 0.005 | 0.030 ± 0.004 | 82, 0.8, 1 |
| S78A | 416 ± 9 | 126 ± 7 | 130 ± 10 | 3.3 ± 0.2 | 3.2 ± 0.3 | 2.5, 0.8, 1 |
| N104D | 364 ± 36 | 4400 ± 300 | 3000 ± 700 | 0.08 ± 0.01 | 0.12 ± 0.03 | 28, 30, 5 |
| D105N | N/A[a] | | | | | As for D6A |
| T132V | 18.4 ± 0.3 | 200 ± 20 | 220 ± 10 | 0.092 ± 0.009 | 0.084 ± 0.004 | 170, 2, 1.1 |
| E154Q | 12.3 ± 1.5[b] | $K_M$ increases | 2400 ± 600 | N/A[b] | 0.005 ± 0.001 | 82, 30, 2; 4 U/ml PK/LD used |
| D187N | 110 ± 2.2 | 430 ± 20 | 580 ± 90 | 0.26 ± 0.01 | 0.19 ± 0.03 | 14, 2, 5 |
| E196Q | 665 ± 15 | 230 ± 20 | 200 ± 20 | 2.9 ± 0.3 | 3.3 ± 0.3 | 5.1, 1.2, 1 |
| N104D E154Q | 6.4 ± 0.9[b] | $K_M$ increases | 1200 ± 500 | N/A[b] | 0.005 ± 0.002 | 280, 30, 10 |

Site-directed mutants were prepared for *P. shigelloides* NagK at key side chains that coordinate ATP, GlcNAc, or magnesium. $k_{cat}$ and $K_{M\ app}$ for both substrates were determined as for the wildtype enzyme. The table indicates the concentration of NagK used, and the concentration of substrate used for determining $K_{M\ app}$ for the reciprocal substrate.
[a] D6A, D6N and D105N mutants caused a loss of activity below the limit of detection of the assay ($k_{cat} < 0.01$ s$^{-1}$); no kinetic constants could be determined.
[b] The E154Q and N104D/E154Q mutants caused the apparent $K_M$ for GlcNAc to increase to above 50 mM (*i.e.*, the plot of rate against substrate concentration was a straight line). An apparent $k_{cat}$ at 50 mM GlcNAc was determined; $K_{M\ app\ GlcNAc}$ and $k_{cat}/K_{M\ app\ GlcNAc}$ cannot be determined.

kinase (46), mutation of the aspartic acid coordinating the 6′-OH (here D105) abolishes activity. This strongly suggests that this is the catalytic base (42, 46).

Our structures provide for the first time a complex of a ROK kinase poised for activity. The structure shows the AMP-PNP γ-phosphate positioned above the 6′-OH group of GlcNAc. The catalytic base, D105, is in position to de-protonate the 6′-O and turn this into a strong nucleophile. The location of the phosphate group allows coordination of two oxygens with the catalytic metal ion. Other carbohydrate kinases generally follow a mechanism of a nucleophilic substitution with a negatively charged intermediate stabilized by a metal ion (37, 66–68). Based on ours and others' structures, it seems highly likely that ROK kinases follow a similar mechanism.

In conclusion, our study provides further detail explaining the catalytic power of ROK kinases. Our structures demonstrate the choreography of the two enzyme domains as they bind partners to form the ternary complex and release these partners and the location of the critical γ-phosphate in the ternary complex. We demonstrate that a metal ion is required for NagK enzymes and that the conserved ROK kinase metal coordinating acid is essential for enzyme activity. Our data confirm the critical side chains that support NagK binding to its substrate GlcNAc. The availability of a detailed structure of the catalytic state of ROK kinases will enable the engineering of these enzymes to phosphorylate alternative substrates to support synthetic biology. This enzyme would also be an attractive target for the development of small molecule inhibitors to target bacteria that rely on cell wall remodeling as part of their pathogenic processes.

## Experimental procedures
### Construction of expression vectors

A codon-optimized *nagK* gene from *P. shigelloides* was cloned into the pOPINS3C expression plasmid (N-terminal polyhistidine and SUMO tags; Addgene #41115 (69)). The gene sequence was obtained as a gBlock (IDT). DNA was amplified by PCR using the following primers: forward:

5′- cagcggtctggaagttctgtttcagggtacc-3′; reverse: 5′- aagctttcta-gaccagtttgtgattaacctc-3′. Each PCR reaction contained 1 ng/µl of gBlock DNA, 0.25 µM each primer, 2.5 mM dNTPs, 1× Phusion buffer and 2 U Phusion Polymerase (NEB). The PCR protocol used was an initial denaturation of 30 s at 98 °C, followed by 35 cycles of 10 s at 98 °C, 10 s at 55 °C and 1 min at 72 °C, with a final elongation step of 5 min at 72 °C. The PCR fragment and the plasmid (pOPINS3C) were assembled using the NEBuilder HiFi DNA Assembly kit (NEB), following the manufacturer's recommendations. The assembled product was transformed into 5-alpha competent cells (NEB), and the insert sequence was confirmed by Sanger sequencing (Source Bioscience). Codon optimized *nagK* genes from *P. shigelloides, V. vulnificus, Pseudoalteromonas* sp. *P1-8,* and *P. damselae* were cloned into pNIC28-Bsa4 (N-terminal polyhistidine tag; Addgene #26103 (70)), pOPINS3C, and pGAT2 (N-terminal polyhistidine and GST tags; Addgene #112588; last two genes (71)) by Twist Bioscience. Genbank files for all plasmids are available in Supplementary Data. Plasmids were transformed into the expression strain *E. coli* BL21 (DE3) (Novagen) using ampicillin (100 µg/ml; pGAT2 and pOPINS3C clones) or kanamycin (50 µg/ml; pNIC28 clones) for selection.

### Expression and purification of NagK

NagK was expressed in 1 L of high salt LB broth supplemented with 100 µg/ml ampicillin or 50 µg/ml kanamycin as appropriate. Each flask was inoculated with 10 ml of an overnight culture and grown at 37 °C with shaking at 200 rpm until $A_{600}$ reached 0.6. NagK expression was induced with 200 µM isopropyl thio-β-D-galactoside, and cultures were grown at 20 °C for 18 h. Cells were harvested by centrifugation at 4500 × *g* for 30 min at 4 °C. The pellet was resuspended in binding buffer (20 mM Tris-HCl, 500 mM NaCl, and 10 mM imidazole, pH 8.0) and lysed by sonication (SONIC Vibra cell VCX130). The lysed sample was clarified by centrifugation (24 000 × *g* for 30 min at 4 °C). The soluble fraction was purified using an ÄKTAxpress chromatography system (GE Healthcare). The sample was purified firstly using a 1 ml HisTrap

crude column (GE Scientific). After loading sample, the column was washed with binding buffer, and the protein eluted into binding buffer with imidazole at 250 mM. The product was purified over a Superdex 200 16/60 size-exclusion column (GE Healthcare) and eluted isocratically into 10 mM Hepes, 500 mM NaCl, pH 7.5. The eluted protein was concentrated using a Vivaspin centrifugal concentrator (Generon) to 1 mg/ml and stored at −20 °C with 20% (v/v) glycerol for enzymatic assays or concentrated to 11.5 mg/ml and stored at −80 °C in small aliquots without any glycerol for crystallization. Protein concentration was determined using a Nanodrop 2000 nano-spectrophotometer (Thermo). The extinction coefficient for NagK was determined using the Protparam tool (https://web.expasy.org/protparam/) (72).

### Kinetic analysis

NagK activity was assayed using the previously described coupling reaction with pyruvate kinase (PK) and lactate dehydrogenase (LD (73)). For *P. shigelloides*, the His-tagged protein was used. Reactions contained 90 to 6000 ng/ml NagK, 40 mM Hepes, pH 7.5, 100 mM KCl, 8 mM MgCl₂, 5 mM DTT, 100 μg/ml BSA, 200 μM NADH, 500 μM phosphoenolpyruvate, 2 U/ml PK-LD (Merck #P0294), 2 mM GlcNAc, and 1 mM ATP. Reactions were performed in 96 well flat-bottomed plates (Greiner #655001) in a total reaction volume of 200 μl. Reactions were monitored by measurement of the absorbance at 340 nm over 40 min in an Infinite M200PRO plate reader (Tecan) with incubation at 37 °C. Datasets were examined individually to determine the region of each experiment that corresponds to the initial rate. Three experimental replicates were performed for all reactions.

Kinetic parameters ($K_M$ and $k_{cat}$) for ATP and GlcNAc were determined by varying either ATP or GlcNAc concentrations between 2 to 0.02 mM and 2 to 0.03 mM, respectively, at constant concentrations of enzyme and partner substrate (values detailed in figure/table legends). The data were fitted to the Michaelis–Menten equation in Prism 9.0.1 (GraphPad). To determine the substrate mechanism, the initial reaction rates were measured with a two-fold dilution of GlcNAc from 2 mM in eight steps and with a two-fold dilution of ATP from 2180 μM in five steps. Two experimental replicates were taken for each data point. Data were fitted to the sequential bi-bi and ping-pong Equations (1 and 2) in Prism 9.0.1 (GraphPad) (73–75). To determine the effect of divalent cations, the coupling enzymes were first tested in a mixture of 40 mM Hepes, pH 7.5, 100 mM KCl, 5 mM DTT, 100 μg/ml BSA, 200 μM NADH, 500 μM phosphoenolpyruvate, 1 mM ADP, and 0.2 U/ml PK-LD. MgCl₂, CaCl₂, MnCl₂, FeCl₂, CoCl₂, NiCl₂, CuCl₂, and ZnCl₂ were tested at 1 to 8 mM. 180 ng/ml NagK was initially tested in the same conditions with ADP substituted with 2 mM GlcNAc, 1 mM ATP, and 10 μl CoCl₂, and the PK-LD increased to 1 U/ml. $K_M$ and $V_{max}$ were determined for MgCl₂, MnCl₂, FeCl₂, CoCl₂, and NiCl₂ by varying the concentration between 0 and 10 mM (0–1 mM for FeCl₂), in the same conditions. Three experimental replicates were performed for all reactions. The data were fitted to the Michaelis–

Menten equation or substrate inhibition Equation (3) as appropriate in Prism v. 9.0.1.

$$v = \frac{V_{max}[A][B]}{K_{iA}K_B + K_A[B] + K_B[A] + [A][B]} \tag{1}$$

$$v = \frac{V_{max}[A][B]}{K_A[B] + K_B[A] + [A][B]} \tag{2}$$

$$v = \frac{v_{max}[S]}{K_M + [S]\left(1 + \frac{[S]}{K_I}\right)} \tag{3}$$

### Differential scanning fluorimetry

The dissociation constant ($K_D$) for NagK with its substrates was determined using differential scanning fluorimetry (57). Each sample contained 0.1 mg/ml NagK, 8× SYPRO Orange dye (Fisher Scientific #10338542), 10 mM Hepes pH 7.5, 100 mM KCl, and varying concentrations of either GlcNAc, AMP-PNP, or the combination of these in a total volume of 10 μl. Data were collected on a Rotorgene Q (Qiagen) using the ROX channel to collect data. The melt curves showed a monotonic melt. Raw data were converted to a percentage unfolded using the fluorescence readings at the start and end of the melt to define 0 and 100% unfolded. 68 °C was selected as the temperature giving an optimal range of unfolding percentages. Data were fitted to Equation 4 using GraphPad v. 9.0.1.

$$f_u = Bottom + (Top - Bottom)\left(1 + \frac{[S]}{EC_{50}}\right) \tag{4}$$

Where $f_u$ is the fraction unfolded, Top and Bottom are the maximum and minimum unfolded fractions, [S] is the varied substrate concentration, and $EC_{50}$ is the substrate concentration that reduces the unfolded fraction by half. Equations fixing Bottom as zero and including a Hill slope were rejected as inferior to three for these data based on Akaike's information criterion.

The fitted $EC_{50}$ values were converted to $K_D$ using Equation 5 (57).

$$K_D = (1 - f_{u0})\left(EC_{50} - \frac{[p]_T}{2}\right) \tag{5}$$

where $f_{u0}$ is the fraction unfolded at zero substrate concentration, and $[P]_T$ is the total protein concentration.

For verifying mutant stability, the melting temperature was determined using the Boltzmann method using Protein Thermal Shift Software v. 1.4 (Applied Biosystems).

### Crystallization

For crystallization, the His-SUMO tagged *P. shigelloides* NagK was used. Crystals were grown using the microbatch

method using an Oryx8 crystallization robot (Douglas Instruments). Initial crystals grew in well E6 of the Morpheus I screen (Molecular Dimensions), mixed 1:1 with 5 mg/ml NagK. Seed stocks were prepared from these crystals in 0.1 M Mops pH 7.5, 30% (v/v) ethylene glycol, and 10% (w/v) PEG 8000. Final crystals were grown by matrix microseeding (76) these crystals into the Morpheus I screen using a mix of 3:2:1 5 mg/ml NagK: mother liquor: seeds. The successful crystallization conditions, soaking conditions, and cryoprotectants used are detailed in Table S1.

### X-ray data collection and structure determination

Data were collected at Diamond Light Source (Didcot) at 100 K using Pilatus 6M-F detectors and wavelengths of 0.92 to 0.98 Å. All data were processed using XDS (77). Further data processing and structural studies was carried out using CCP4 program package (78, 79). The apo structure of NagK was solved by the molecular replacement (MR) using the MR pipeline MORDA (80) with the best solution found for the model (PDB ID: 4DB3). The model was refined using REFMAC5 (81) and *PHENIX* (82) and rebuilt using COOT (83). The refined apo NagK model was used as a MR search model in MOLREP (84) for the NagK-GlcNAc-ADP data, which crystallized in a different space group. The MR solution was refined using *Buccaneer* (85), following which further refinement was performed as above. The crystals of NagK-GlcNAc-AMP, NagK-GlcNAc, NagK-GlcNAc-6'-phosphate, and NagK-GlcNAc-AMP-PNP were in the same space group as the NagK-GlcNAc-ADP complex; however, phased MR (86) was used to reposition the small domain in the NagK-GlcNAc structure. All structures were subjected to phased refinement in REFMAC5 (87) with input density modification phases (88) from non-crystallographic symmetry averaging. The models were validated using MOLPROBITY (89) implemented in the CCP4i2 interface (90).

### Molecular dynamics

Molecular dynamics was performed in YASARA v.20.12.24 (91). The structure of NagK complexed with GlcNAc and AMP-PNP was cleaned to remove water and PEG molecules. Molecular dynamics was run using the md_runfast macro for 5 ns using the AMBER15FB force field (92). Simulations including divalent cations were performed by replacing water molecule 97 with the relevant cation.

### Data availability

All data underpinning this work are publicly available. Structure coordinates and structure factor files are deposited with the Protein Data Bank (accession numbers: 7PA1, 7P7I, 7P7W, 7P9L, 7P9P and 7P9Y). Enzymatic and biophysical data are available as Supplementary Files or from Open Research Exeter (doi: to be confirmed on acceptance).

*Supporting information*—This article contains supporting information (34, 42, 51, 68, 73, 93–96, https://pymol.org/2/).

## References

1. Hascall, V., and Esko, J. D. (2015) Hyaluronan. In: Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Etzler, M. E., eds. *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, New York: 197–206
2. Lindahl, U., Couchman, J., Kimata, K., and Esko, J. D. (2015) Proteoglycans and sulfated glycosaminoglycans. In: Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Etzler, M. E., eds. *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, New York: 207–221
3. Tiemeyer, M., Nakato, H., and Esko, J. D. (2015) Arthropoda. In: Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Etzler, M. E., eds. *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, New York: 335–349
4. Zachara, N., Akimoto, Y., and Hart, G. W. (2015) The O-GlcNAc modification. In: Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Etzler, M. E., eds. *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, New York: 239–251
5. Ogawa, M., and Okajima, T. (2019) Structure and function of extracellular O-GlcNAc. *Curr. Opin. Struct. Biol.* **56**, 72–77
6. Groves, J. A., Lee, A., Yildirir, G., and Zachara, N. E. (2013) Dynamic O-GlcNAcylation and its roles in the cellular stress response and homeostasis. *Cell Stress Chaperones* **18**, 535–558
7. Hardiville, S., and Hart, G. W. (2014) Nutrient regulation of signaling, transcription, and cell physiology by O-GlcNAcylation. *Cell Metab.* **20**, 208–213
8. Hart, G. W. (2019) Nutrient regulation of signaling and transcription. *J. Biol. Chem.* **294**, 2211–2231
9. Seltmann, G., and Holst, O. (2001) *The bacterial cell wall* **ix**. Springer, Berlin; New York: 280
10. Whitfield, C., and Trent, M. S. (2014) Biosynthesis and export of bacterial lipopolysaccharides. *Annu. Rev. Biochem.* **83**, 99–128
11. Reid, A. N., and Szymanski, C. M. (2010) Biosynthesis and assembly of capsular polysaccharides. In: Holst, O., Brennan, P. J., von Itzstein, M., eds. *Microbial Glycobiology*, Academic Press, London: 351–373
12. Forsberg, L. S., Noel, K. D., Box, J., and Carlson, R. W. (2003) Genetic locus and structural characterization of the biochemical defect in the O-antigenic polysaccharide of the symbiotically deficient Rhizobium etli mutant, CE166. Replacement of N-acetylquinovosamine with its hexosyl-4-ulose precursor. *J. Biol. Chem.* **278**, 51347–51359

13. Visansirikul, S., Kolodziej, S. A., and Demchenko, A. V. (2020) Staphylococcus aureus capsular polysaccharides: a structural and synthetic perspective. *Org. Biomol. Chem.* **18**, 783–798
14. Hong, Y., and Reeves, P. R. (2014) Diversity of o-antigen repeat unit structures can account for the substantial sequence variation of wzx translocases. *J. Bacteriol.* **196**, 1713–1722
15. Islam, S. T., and Lam, J. S. (2014) Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Can. J. Microbiol.* **60**, 697–716
16. Hong, Y., Morcilla, V. A., Liu, M. A., Russell, E. L., and Reeves, P. R. (2015) Three Wzy polymerases are specific for particular forms of an internal linkage in otherwise identical O units. *Microbiology* **161**, 1639–1647
17. Lizak, C., Gerber, S., Numao, S., Aebi, M., and Locher, K. P. (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474**, 350–355
18. Terra, V. S., Mills, D. C., Yates, L. E., Abouelhadid, S., Cuccui, J., and Wren, B. W. (2012) Recent developments in bacterial protein glycan coupling technology and glycoconjugate vaccine design. *J. Med. Microbiol.* **61**, 919–926
19. Barreteau, H., Kovac, A., Boniface, A., Sova, M., Gobec, S., and Blanot, D. (2008) Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiol. Rev.* **32**, 168–207
20. Milewski, S., Gabriel, I., and Olchowy, J. (2006) Enzymes of UDP-GlcNAc biosynthesis in yeast. *Yeast* **23**, 1–14
21. Park, J. T. (2001) Identification of a dedicated recycling pathway for anhydro-N-acetylmuramic acid and N-acetylglucosamine derived from Escherichia coli cell wall murein. *J. Bacteriol.* **183**, 3842–3847
22. Plumbridge, J. (2009) An alternative route for recycling of N-acetyl-glucosamine from peptidoglycan involves the N-acetylglucosamine phosphotransferase system in Escherichia coli. *J. Bacteriol.* **191**, 5641–5647
23. Popowska, M., Osinska, M., and Rzeczkowska, M. (2012) N-acetylglu-cosamine-6-phosphate deacetylase (NagA) of Listeria monocytogenes EGD, an essential enzyme for the metabolism and recycling of amino sugars. *Arch. Microbiol.* **194**, 255–268
24. Yadav, V., Panilaitis, B., Shi, H., Numuta, K., Lee, K., and Kaplan, D. L. (2011) N-acetylglucosamine 6-phosphate deacetylase (nagA) is required for N-acetyl glucosamine assimilation in Gluconacetobacter xylinus. *PLoS One* **6**, e18099
25. Fisher, J. F., and Mobashery, S. (2014) The sentinel role of peptidoglycan recycling in the beta-lactam resistance of the Gram-negative Enterobacteriaceae and Pseudomonas aeruginosa. *Bioorg. Chem.* **56**, 41–48
26. Borisova, M., Gaupp, R., Duckworth, A., Schneider, A., Dalugge, D., Muhleck, M., *et al.* (2016) Peptidoglycan recycling in gram-positive bacteria is crucial for survival in stationary phase. *mBio* **7**, e00923-16
27. Komatsuzawa, H., Fujiwara, T., Nishi, H., Yamada, S., Ohara, M., McCallum, N., *et al.* (2004) The gate controlling cell wall synthesis in Staphylococcus aureus. *Mol. Microbiol.* **53**, 1221–1231
28. Zhang, W., Jones, V. C., Scherman, M. S., Mahapatra, S., Crick, D., Bhamidi, S., *et al.* (2008) Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase. *Int. J. Biochem. Cell Biol.* **40**, 2560–2571
29. Oyeleye, A., and Normi Yahaya, M. (2018) Chitinase: Diversity, limitations, and trends in engineering for suitable applications. *Biosci. Rep.* **38**. BSR2018032300
30. Rathore, A. S., and Gupta, R. D. (2015) Chitinases from bacteria to human: properties, applications, and future perspectives. *Enzyme Res.* **2015**, 791907
31. Aggarwal, C., Paul, S., Tripathi, V., Paul, B., and Khan, M. A. (2015) Chitinolytic activity in Serratia marcescens (strain SEN) and potency against different larval instars of Spodoptera litura with effect of sublethal doses on insect development. *BioControl* **60**, 631–640
32. Meibom, K. L., Li, X. B., Nielsen, A. T., Wu, C. Y., Roseman, S., and Schoolnik, G. K. (2004) The Vibrio cholerae chitin utilization program. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 2524–2529
33. Fang, W. X., Du, T., Raimi, O. G., Hurtado-Guerrero, R., Marino, K., Ibrahim, A. F. M., *et al.* (2013) Genetic and structural validation of Aspergillus fumigatus N-acetylphosphoglucosamine mutase as an antifungal target. *Biosci. Rep.* **33**, 689–699
34. Nishitani, Y., Maruyama, D., Nonaka, T., Kita, A., Fukami, T. A., Mio, T., *et al.* (2006) Crystal structures of N-acetylglucosamine-phosphate mutase, a member of the alpha-D-phosphohexomutase superfamily, and its substrate and product complexes. *J. Biol. Chem.* **281**, 19740–19747
35. Dik, D. A., Fisher, J. F., and Mobashery, S. (2018) Cell-wall recycling of the gram-negative bacteria and the nexus to antibiotic resistance. *Chem. Rev.* **118**, 5952–5984
36. Weihofen, W. A., Berger, M., Chen, H., Saenger, W., and Hinderlich, S. (2006) Structures of human N-acetylglucosamine kinase in two complexes with N-acetylglucosamine and with ADP/glucose: insights into substrate specificity and regulation. *J. Mol. Biol.* **364**, 388–399
37. Roy, S., Vega, M. V., and Harmer, N. J. (2019) Carbohydrate kinases: a conserved mechanism across differing folds. *Catalysts* **9**, 29
38. Brigham, C. J., and Malamy, M. H. (2005) Characterization of the RokA and HexA broad-substrate-specificity hexokinases from *Bacteroides fragilis* and their role in hexose and N-acetylglucosamine utilization. *J. Bacteriol.* **187**, 890–901
39. Conejo, M. S., Thompson, S. M., and Miller, B. G. (2010) Evolutionary bases of carbohydrate recognition and substrate discrimination in the ROK protein family. *J. Mol. Evol.* **70**, 545–556
40. Hansen, T., and Schonheit, P. (2003) ATP-dependent glucokinase from the hyperthermophilic bacterium *Thermotoga maritima* represents an extremely thermophilic ROK glucokinase with high substrate specificity. *FEMS Microbiol. Lett.* **226**, 405–411
41. Coombes, D., Davies, J. S., Newton-Vesty, M. C., Horne, C. R., Setty, T. G., Subramanian, R., *et al.* (2020) The basis for non-canonical ROK family function in the N-acetylmannosamine kinase from the pathogen Staphylococcus aureus. *J. Biol. Chem.* **295**, 3301–3315
42. Miyazono, K., Tabei, N., Morita, S., Ohnishi, Y., Horinouchi, S., and Tanokura, M. (2012) Substrate recognition mechanism and substrate-dependent conformational changes of an ROK family glucokinase from *Streptomyces griseus*. *J. Bacteriol.* **194**, 607–616
43. Nakamura, T., Kashima, Y., Mine, S., Oku, T., and Uegaki, K. (2012) Characterization and crystal structure of the thermophilic ROK hexokinase from *Thermus thermophilus*. *J. Biosci. Bioeng.* **114**, 150–154
44. Dorr, C., Zaparty, M., Tjaden, B., Brinkmann, H., and Siebers, B. (2003) The hexokinase of the hyperthermophile *Thermoproteus tenax*. ATP-dependent hexokinases and ADP-dependent glucokinases, two alternatives for glucose phosphorylation in Archaea. *J. Biol. Chem.* **278**, 18744–18753
45. Hansen, T., Reichstein, B., Schmid, R., and Schonheit, P. (2002) The first archaeal ATP-dependent glucokinase, from the hyperthermophilic crenarchaeon Aeropyrum pernix, represents a monomeric, extremely thermophilic ROK glucokinase with broad hexose specificity. *J. Bacteriol.* **184**, 5955–5965
46. Martinez, J., Nguyen, L. D., Hinderlich, S., Zimmer, R., Tauberger, E., Reutter, W., *et al.* (2012) Crystal structures of N-acetylmannosamine kinase provide insights into enzyme activity and inhibition. *J. Biol. Chem.* **287**, 13656–13665
47. Uehara, T., and Park, J. T. (2004) The N-acetyl-D-glucosamine kinase of Escherichia coli and its role in murein recycling. *J. Bacteriol.* **186**, 7273–7279
48. Hekkelman, M. L., de Vries, I., Joosten, R. P., and Perrakis, A. (2022) AlphaFill: Enriching AlphaFold models with ligands and cofactors. *Nat. Methods* **20**, 205–213
49. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589
50. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., *et al.* (2022) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444

51. Marangoni, A. G. (2002) Two-substrate reactions *Enzyme Kinetics*. Wiley, Hoboken, NJ: 90–101

52. Morpeth, F. F., and Massey, V. (1982) Steady-state kinetic studies on D-lactate dehydrogenase from Megasphera elsdenii. *Biochemistry* **21**, 1307–1312

53. Katz, M., and Westley, J. (1979) Enzymic memory. Steady state kinetic and physical studies with ascorbate oxidase and aspartate aminotransferase. *J. Biol. Chem.* **254**, 9142–9147

54. Katz, M., and Westley, J. (1980) Enzymic memory studies with nucleoside-5′-diphosphate kinase. *Arch. Biochem. Biophys.* **204**, 464–470

55. Brown, J. I., Page, B. D. G., and Frankel, A. (2019) The application of differential scanning fluorimetry in exploring bisubstrate binding to protein arginine N-methyltransferase 1. *Methods* **175**, 10–23

56. Llano-Sotelo, B., Azucena, E. F., Jr., Kotra, L. P., Mobashery, S., and Chow, C. S. (2002) Aminoglycosides modified by resistance enzymes display diminished binding to the bacterial ribosomal aminoacyl-tRNA site. *Chem. Biol.* **9**, 455–463

57. Bai, N., Roder, H., Dickson, A., and Karanicolas, J. (2019) Isothermal analysis of ThermoFluor data can readily provide quantitative binding affinities. *Sci. Rep.* **9**, 2650

58. Ribeiro António, J. M., Holliday, G. L., Furnham, N., Tyzack, J. D., Ferris, K., and Thornton, J. M. (2017) Mechanism and catalytic site atlas (M-CSA): A database of enzyme reaction mechanisms and active sites. *Nucleic Acids Res.* **46**, D618–D623

59. Minasov, G., Wawrzak, Z., Onopriyenko, O., Skarina, T., Papazisi, L., Savchenko, A., *et al.* (2012) *1.95 Angstrom Resolution Crystal Structure of N-Acetyl-D-Glucosamine Kinase from Vibrio Vulnificus*. Protein Data Bank PDB ID (4DB3)

60. Pendleton, J. N., Gorman, S. P., and Gilmore, B. F. (2013) Clinical relevance of the ESKAPE pathogens. *Expert Rev. Anti Infect. Ther.* **11**, 297–308

61. Shakir, N. A., Aslam, M., Bibi, T., and Rashid, N. (2021) ADP-dependent glucose/glucosamine kinase from Thermococcus kodakarensis: cloning and characterization. *Int. J. Biol. Macromol.* **173**, 168–179

62. Chandrangsu, P., Rensing, C., and Helmann, J. D. (2017) Metal homeostasis and resistance in bacteria. *Nat. Rev. Microbiol.* **15**, 338–350

63. Szatmari, D., Sarkany, P., Kocsis, B., Nagy, T., Miseta, A., Barko, S., *et al.* (2020) Author Correction: intracellular ion concentrations and cation-dependent remodelling of bacterial MreB assemblies. *Sci. Rep.* **10**, 18185

64. Tottey, S., Harvie, D. R., and Robinson, N. J. (2005) Understanding how cells allocate metals using metal sensors and metallochaperones. *Acc. Chem. Res.* **38**, 775–783

65. Zeinert, R., Martinez, E., Schmitz, J., Senn, K., Usman, B., Ananthara-man, V., *et al.* (2018) Structure-function analysis of manganese exporter proteins across bacteria. *J. Biol. Chem.* **293**, 5715–5730

66. Li, J., Wang, C., Wu, Y., Wu, M., Wang, L., Wang, Y., *et al.* (2012) Crystal structure of Sa239 reveals the structural basis for the activation of ribokinase by monovalent cations. *J. Struct. Biol.* **177**, 578–582

67. Nishimasu, H., Fushinobu, S., Shoun, H., and Wakagi, T. (2007) Crystal structures of an ATP-dependent hexokinase with broad substrate specificity from the hyperthermophilic archaeon Sulfolobus tokodaii. *J. Biol. Chem.* **282**, 9923–9931

68. Nocek, B., Stein, A. J., Jedrzejczak, R., Cuff, M. E., Li, H., Volkart, L., *et al.* (2011) Structural studies of ROK fructokinase YdhR from Bacillus subtilis: insights into substrate binding and fructose specificity. *J. Mol. Biol.* **406**, 325–342

69. Bird, L. E. (2011) High throughput construction and small scale expression screening of multi-tag vectors in Escherichia coli. *Methods* **55**, 29–37

70. Savitsky, P., Bray, J., Cooper, C. D., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A., *et al.* (2010) High-throughput production of human proteins for crystallization: the SGC experience. *J. Struct. Biol.* **172**, 3–13

71. Peranen, J., Rikkonen, M., Hyvonen, M., and Kaariainen, L. (1996) T7 vectors with modified T7lac promoter for expression of proteins in Escherichia coli. *Anal. Biochem.* **236**, 371–373

72. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., *et al.* (2005) Protein identification and analysis tools on the ExPASy server. In: Walker, J. M., ed. *The Proteomics Protocols Handbook*, Humana Press, Totowa, NJ: 571–607

73. Cook, P. F., and Cleland, W. W. (2007) *Enzyme kinetics and mechanism* **xxii**. Garland Science, London; New York: 404

74. Cornish-Bowden, A. (2012) *Fundamentals of Enzyme Kinetics. 4th, Completely Revised and Greatly Enlarged Edition* **xviii**. Wiley-Blackwell, Weinheim, Germany: 498

75. Harmer, N. J., and Vivoli Vega, M. (2019) Reaction Chemical Kinetics in Biology. In: Ramesh, V., ed. *Biomolecular and Bioanalytical Techniques*, Wiley, Hoboken, NJ: 179–217

76. D'Arcy, A., Villard, F., and Marsh, M. (2007) An automated microseed matrix-screening method for protein crystallization. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 550–554

77. Kabsch, W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132

78. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 1131–1137

79. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242

80. Vagin, A., and Lebedev, A. (2015) *MoRDa*, an automatic molecular replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr.* **A71**, s19

81. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367

82. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367

83. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501

84. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25

85. Cowtan, K. (2012) Completion of autobuilt protein models using a database of protein fragments. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 328–335

86. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 1451–1456

87. Pannu, N. S., Murshudov, G. N., Dodson, E. J., and Read, R. J. (1998) Incorporation of prior phase information strengthens maximum-likelihood structure refinement. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 1285–1294

88. Cowtan, K. (2010) Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 470–478

89. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., *et al.* (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315

90. Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P. R., *et al.* (2018) CCP4i2: The new graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **74**, 68–84

91. Krieger, E., and Vriend, G. (2015) New ways to boost molecular dynamics simulations. *J. Comput. Chem.* **36**, 996–1007

92. Wang, L. P., Martinez, T. J., and Pande, V. S. (2014) Building force fields: an automatic, systematic, and reproducible approach. *J. Phys. Chem. Lett.* **5**, 1885–1891

93. Mukai, T., Kawai, S., Mori, S., Mikami, B., and Murata, K. (2004) Crystal structure of bacterial inorganic polyphosphate/ATP-glucomannokinase. Insights into kinase evolution. *J. Biol. Chem.* **279**, 50591–50600

94. Olsen, L. R., Vetting, M. W., and Roderick, S. L. (2007) Structure of the E. coli bifunctional GlmU acetyltransferase active site with substrates and products. *Protein Sci.* **16**, 1230–1235

95. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324

96. Zhao, X., Creuzenet, C., Belanger, M., Egbosimba, E., Li, J., and Lam, J. S. (2000) WbpO, a UDP-N-acetyl-D-galactosamine dehydrogenase from Pseudomonas aeruginosa serotype O6. *J. Biol. Chem.* **275**, 33252–33259

97. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model* **51**, 2778–2786

98. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134

99. Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality. *Science* **336**, 1030–1033

100. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 191–205

ASBMB