Structure and Function of N-Acetylmannosamine Kinases from Pathogenic Bacteria

Thanuja Gangi Setty, Arunabha Sarkar, David Coombes, Renwick C. J. Dobson, and Ramaswamy Subramanian*

ABSTRACT: Several pathogenic bacteria import and catabolize sialic acids as a source of carbon and nitrogen. Within the sialic acid catabolic pathway, the enzyme N-acetylmannosamine kinase (NanK) catalyzes the phosphorylation of N-acetylmannosamine to N-acetylmannosamine-6-phosphate. This kinase belongs to the ROK superfamily of enzymes, which generally contain a conserved zinc-finger (ZnF) motif that is important for their structure and function. Previous structural studies have shown that the ZnF motif is absent in NanK of Fusobacterium nucleatum (Fn-NanK), a Gram-negative bacterium that causes the gum disease gingivitis. However, the effect in loss of the ZnF motif on the kinase activity is unknown. Using kinetic and thermodynamic studies, we have studied the functional properties of Fn-NanK to its substrates ManNAc and ATP, compared its activity with other ZnF motif-containing NanK enzymes from closely related Gram-negative pathogenic bacteria Haemophilus influenzae (Hi-NanK), Pasteurella multocida (Pm-NanK), and Vibrio cholerae (Vc-NanK). Our studies show a 10-fold decrease in substrate binding affinity between Fn-NanK (apparent $K_a \approx 700 \mu M$) and ZnF motif-containing NanKs (apparent $K_a \approx 60 \mu M$). To understand the structural features that combat the loss of the ZnF motif in Fn-NanK, we solved the crystal structures of functionally homologous ZnF motif-containing NanKs from P. multocida and H. influenzae. Here, we report Pm-NanK:unliganded, Pm-NanK:AMPPNP, Pm-NanK:ManNAc, Hi-NanK:ManNAc, and Hi-NanK:ManNAc-6P:ADP crystal structures. Structural comparisons of Fn-NanK with Hi-NanK, Pm-NanK, and hMNK (human N-acetylmannosamine kinase domain of UDP-N-acetylgalactosamine-2-epimerase/N-acetylmannosamine kinase, GNE) show that even though there is less sequence identity, they have high degree of structural similarity. Furthermore, our structural analyses highlight that the ZnF motif of Fn-NanK is substituted by a set of hydrophobic residues, which forms a hydrophobic cluster that helps the proper orientation of ManNAc in the active site. In summary, ZnF-containing and ZnF-lacking NanK enzymes from different Gram-negative pathogenic bacteria are functionally very similar but differ in their metal requirement. Our structural studies unveil the structural modifications in Fn-NanK that compensate the loss of the ZnF motif in comparison to other NanK enzymes.

1. INTRODUCTION

Zinc (Zn) is a nontoxic metal and is essential for a range of biological processes. Due to its coordination flexibility with different macromolecules, a range of enzymes across different microorganisms, plants, and animals use Zn as a cofactor.1−5 In addition, Zn is required for catalytic, structural, and functional properties of many proteins and enzymes, where it plays an important role in folding and oligomerization of proteins.6−9 In metal-binding enzymes, a Zn-finger (ZnF) motif acts as a structural motif formed by the coordination of Zn with cysteines and histidine residues. Classical ZnF motifs contain Cys$_2$−His$_3$ residues and fold into anti-parallel $\beta$-sheets/$\alpha$-helices to coordinate the Zn. On the other hand, proteins with non-classical Zn-finger (NcZnF) motif are divided into CCCH and CCH[H]C types, and they form a loop kind of structures.5,10

Sialic acids are a family of cyclic nine-carbon amino sugar acids. N-acetylneuraminic acid (Neu5Ac) and N-glycolyneuraminic acid (Neu5Gc) are the most common sialic acids. In eukaryotes, these sugars are present as terminal residues on the cell surface glycans. In the case of prokaryotes, these sialic acids are decorated as the outermost moiety of the lipooligosaccharide/lipopolysaccharide (LOS/LPS) on their cell membranes.11−13 Pathogenic bacteria such as Haemophilus influenzae (Hi), Pasteurella multocida (Pm), Vibrio cholerae...
the biosynthetic pathway of Neu5Ac.11,24 ROK family NanK nosamine kinase, GNE), has been shown to be involved in and V. cholera.

mucous environments of the gut and respiratory tract scavenging sialic acids from the host and transport them into bacteria using the TRAP transport system.14,15 Post scavenging, these bacteria use sialic acids as a carbon and nitrogen source by converting them into fructose 6-phosphate using the enzymes of the nan-nag gene cluster (NanA, NanK, NanE, NagA, and NagB) (Figure 1). NanK is the second enzyme in the sialic acid catabolic pathway and phosphorylates \(N\)-acetylmannosamine (ManNAc) into \(N\)-acetylmannosaminosaccharide (ManNAc-6P) in the presence of ATP, an important step in the sugar metabolism (Figure 1, highlighted in a box).20,21

NanK enzymes belong to the ROK super family of enzymes (bacterial Repressors, uncharacterized Open reading frames, and sugar Kinases), which are functionally diverse and expressed in prokaryotes and eukaryotes.22,23 The human homologue, hMNK (the NanK component of human bifunctional \(N\)-acetylglucosamine-2-epimerase/\(N\)-acetylmannosamine kinase, GNE), has been shown to be involved in the biosynthetic pathway of Neu5Ac.11,24 ROK family NanK enzymes possess conserved signature motifs: (i) DxGxT, located in the ATP binding site at the N-terminus of the protein; (ii) substrate-binding ExGH motif; and (iii) a catalytic aspartate residue. In contrast to traditional CCHH-type ZnF motifs, NanK enzymes have non-classical CCCH-type ZnF motifs, where Zn coordinates with one histidine (from ExGH motif) and three cysteines. Previous structural studies on hMNK (the NanK component of human bifunctional \(N\)-acetylglucosamine-2-epimerase/\(N\)-acetylmannosamine kinase, GNE), have shown that the functional divergence between these NanKs may provide insights into the sequel of presence or absence of the ZnF motif and divalent metal Zn.

Figure 1. Schematic representation of the sialic acid catabolic pathway in the Gram-negative bacteria. The enzymatic reaction of phosphorylation of \(N\)-acetylmannosaminosaccharide (ManNAc) to \(N\)-acetylmannosaminosaccharide-6-phosphate by the enzyme NanK (\(N\)-acetylmannosamine kinase) in the presence of ATP is highlighted in the box. The chemical structures are drawn using ChemDraw. Neu5Ac, \(N\)-acytelyneuraminic acid; NanA, \(N\)-acytelyneuraminic acid lyase; ManNAc, \(N\)-acetylmannosamine, ManNAc-6P, \(N\)-acetylmannosaminosaccharide-6-phosphate; NanE, \(N\)-acytelymannosaminosaccharide-6-phosphate deacetylase; and NagB, glucosamine-6-phosphate deaminase.

NanK is twice as compared to hMNK.27 Thus, understanding the functional difference between these NanKs may provide insights into the sequel of presence or absence of the ZnF motif and divalent metal Zn.

Towards this, we have grouped Hi-NanK, Pm-NanK, and Vc-NanK into ZnF-containing NanK enzymes and Fn-NanK as ZnF-lacking NanK. In this study, we compare the kinetic and thermodynamic properties of these two groups of NanKs. All of the studies on the different NanKs reported here were performed on recombinantly expressed proteins with a polyhistidine tag. In addition, we also report the crystal structures of \(Pm\)-NanK unliganded, \(Pm\)-NanK: adenylylimidodiphosphate (AMPNP), \(Pm\)-NanK: \(N\)-acetylmannosaminosaccharide-6-phosphate and adenosine-\(S\)-diphosphate (\(ManNAc\)-6P:ADP) forms. Further, we compare the structural properties of these two groups of NanK enzymes to understand the structural changes in \(Fn\)-NanK that help in conserving its functional properties. This systematic and comprehensive biophysical characterization of ZnF-containing or ZnF-lacking NanK enzymes elucidate the significance of unique amino acids in the substrate-binding pocket of \(Fn\)-NanK that assist to overcome the requirement of the Zn metal for its enzymatic activity.

2. RESULTS AND DISCUSSION

2.1. NanKs Exhibit Low Sequence Identity but Have Highly Conserved Signature Motifs. Multiple amino acid sequence alignments of NanK enzymes from hMNK, different Gram-negative pathogenic bacteria, and Gram-positive pathogenic bacteria were carried out using ClustalW.28 The analysis reveals that the sequence identity between different NanKs is only 20−25%. Among these sequences, \(Hi\)-NanK and \(Pm\)-NanK show higher sequence identity of approximately 70% (Table S1). Though the sequence identity between the NanK enzymes appears to be moderate, these kinases show high homology at the conserved signature motifs of the ROK family such as the DxGxT motif, catalytic aspartate residue, and ExGH motif. Previous structural studies on hMNK and \(Pm\)-NanK (PDB ID: 2AA4) enzymes also demonstrate the presence of the conserved motifs specific to ROK family enzymes in their structures.26 Further, excluding
**Figure 2.** Kinetic analysis of NanK enzymes. The initial rates for the conversion of ManNac to ManNac-6P by (A) Hi-NanK, (B) Pm-NanK, (C) Vc-NanK, and (D) Fn-NanK were determined by using various concentrations of ManNac (at 0.1 mM ATP), and the luminescence was measured using a plate reader. Data was fitted to the Michaelis–Menten model. The kinetic parameters (apparent $K_M$ and $V_{max}$) were estimated from experimental duplicate and technical triplicate values. The error bars represent the fit to the curve from multiple experiments.

**Table 1. Comparative Apparent Kinetic Parameters of ZnF-Containing and ZnF-Lacking NanK Enzymes from Gram-Negative Bacteria**

| enzymes | ZnF | apparent $K_M$ ($\mu$M) | $V_{max}$ (µM min$^{-1}$) | $k_{cat}$ (min$^{-1}$) | $k_{cat}/K_M$ (µM$^{-1}$ min$^{-1}$) |
|---------|-----|--------------------------|---------------------------|----------------------|---------------------------------------|
| (a) Hi-NanK | ZnF | 66.3 ± 5 | 1.33 ± 0.03 | 112 | 1.6 |
| (b) Pm-NanK | ZnF | 63.3 ± 45 | 0.97 ± 0.02 | 77 | 1.2 |
| (c) Vc-NanK | ZnF | 71 ± 9 | 0.67 ± 0.02 | 54 | 0.8 |
| (d) Fn-NanK | - | 730 ± 535 | 1.04 ± 0.03 | 82 | 0.1 |
| (e) Sa-NanK | - | 140 ± 8 | 20.50 ± 0.18 | 1680 | 12 |

*Apparent parameters based on Coombes et al. with conditions matching this assay.*

FN-NanK and *Staphylococcus aureus*-NanK (Sa-NanK), all the other bacterial NanK enzymes show the presence of a conserved ZnF motif (Figure S1). Noticeably, NanK from *F. nucleatum* only has His159 (from ExGH motif) and not the three cysteines, resulting in the absence of the ZnF motif and presumably a zinc ion in the structure. Previous functional studies on the ROK family of enzymes such as the transcriptional repressor Mlc from *E. coli* and GlcK from *Bacillus subtilis* reveal that mutating the cysteine residues of the ZnF motif leads to either lack of or reduced enzyme activity. Amino acid insertions, mutations, or deletions may lead to new structural and functional properties of the proteins, and they are significant in protein evolution. Nevertheless, the previous structural studies on Fn-NanK illustrated that in spite of the absence of the ZnF motif, the three-dimensional structure is not affected. However, the alterations in functional properties of Fn-NanK due to the absence of the divalent metal Zn or ZnF motif were not understood.

### 2.2. Comparison of Kinetic Properties between ZnF-Containing and ZnF-Lacking NanK Enzymes

NanK enzymes catalyze the phosphorylation of ManNac. To understand the significance of the ZnF motif in these enzymes, we characterized the steady state kinetic properties of Hi-NanK, Pm-NanK and Vc-NanK (ZnF-containing NanKs), and Fn-NanK (ZnF-lacking NanK) with the substrate ManNac (Figure 2). The apparent $K_M$, $V_{max}$ and $k_{cat}$ for different NanK enzymes with ManNac are presented in Table 1 (due to experimental limitations, the concentration of ATP was limiting in these experiments. Therefore, refer to our kinetic measurements as apparent constants). The kinetic analysis demonstrates that Hi-NanK, Pm-NanK, and Vc-NanK have similar $K_M$ of around 65 µM, which is also close to the $K_M$ of hMNK (95 µM). In contrast, the kinetic studies of Fn-NanK show an increased $K_M$ of 715 µM toward ManNac, which results in Fn-NanK having slightly a 10-fold lower catalytic efficiency than the other ZnF-containing NanK enzymes. However, the difference is not large, and ZnF-lacking Sa-NanK has an apparent $K_M$ (140 µM) that more closely correlates with the ZnF NanKs, and it has a higher $k_{cat}$. Overall, the kinetic studies demonstrate that although the catalytic properties vary between ZnF-containing and ZnF-lacking NanK enzymes, the direct difference is not large.

### 2.3. Reductive Amination of NanK Enzyme Products with AEC and Subsequent Mass Spectrometry Confirms Phosphorylation

In addition to kinetic analysis, we characterized the hydrolyzed product of NanK by mass spectrometry after AEC (3-amino-9-ethylcarbazole) derivatization to confirm phosphorylation activity. In this procedure, the substrate and products of the NanK enzyme ($Pm$-NanK and $Fn$-NanK) were first reduced by AEC in the presence of sodium cyanoborohydrate to ease their quantification by mass spectrometry. Subsequent mass spectrometry analysis revealed the daughter ion signature that can only emanate from phosphorylated NanK.
from the expected product of the NanK enzyme. These studies confirm that both ZnF-containing and ZnF-lacking NanK enzymes catalyze the phosphorylation reaction and generate the product N-acetylmannosamine 6-phosphate (Figure S2A,B).

2.4. Thermodynamic Characterization of ManNAc and AMPPNP Binding to NanK Enzymes. Limited kinetic studies elsewhere have shown that ZnF glucokinases need to bind the sugar substrate first before binding to the nucleotide, whereas ZnF-lacking Su-NanK did not specify substrate-binding order.26,29,35,36 This has led to the hypothesis that ZnF controls substrate entry into the active site pocket, and sugar must be loaded first. We used isothermal titration calorimetry (ITC) to investigate the differences in the thermodynamic properties between these two groups of NanKs binding to the substrates ManNAc and AMPPNP. Figure 3 shows the ITC thermograms, wherein the top panel illustrates the raw heat liberated during protein ligand titrations, and the bottom panel demonstrates the integrated enthalpy changes for these titrations. Single-site binding mode
Table 2. Binding Affinities and Thermodynamic Parameters of NanKs Binding to Their Ligands ManNAc and AMPPNP (ATP Analogue)\(^b\)  

| NanK           | in the presence of ligand | \(K_d\) (\(\mu\)M) | \(\Delta H\) (Kcal/mol) | \(T\Delta S\) (Kcal/mol) |
|----------------|--------------------------|---------------------|-------------------------|--------------------------|
| *P. multocida* | ManNAc                   | 84 ± 9.9            | -1.763 ± 0.158          | 3.78                     |
|                | AMPPNP                   | no measurable binding |                         |                          |
|                | ManNAc                   | 80 ± 1.9            | -13 ± 0.373             | -7.47                    |
| *V. cholera*   | ManNAc                   | 112 ± 14.5          | -4.3 ± 0.35             | 1.1                      |
|                | AMPPNP                   | 39 ± 4.5            | -7.18 ± 0.825           | -1.174                   |
| *H. influenzae*| ManNAc                   | 27 ± 3.1            | -17 ± 0.9               | -10.787                  |
|                | AMPPNP                   | 44 ± 9.2            | -2.1 ± 0.482            | 3.814                    |
| *F. nucleatum* | ManNAc                   | 33 ± 4.29           | -3.15 ± 0.2             | 2.947                    |
|                | AMPPNP                   | 53 ± 6.4            | -7.7 ± 2.44             | -1.85                    |
|                | ManNAc                   | 28 ± 4.76           | -3.49 ± 0.475           | 2.717                    |
|                | AMPPNP                   | 72 ± 12             | -3.7 ± 0.68             | 1.94                     |
|                | MANNAc                   | 76 ± 14             | -0.814 ± 0.08           | 4.8                      |
|                | AMPPNP                   | 61 ± 7.48           | -3.613 ± 0.258          | 2.13                     |
| Hi-NanK        | ManNAc                   | 55 ± 5.5            | -5.25 ± 0.51            | 0.548                    |
|                | AMPPNP                   | 35 ± 1.12           | -8.94 ± 0.18            | -2.87                    |

\(^a\)While most experiments were repeated multiple times, these titrations were only done once due to a paucity of AMPPNP. \(^b\)Isothermal calorimetry (ITC) was used to measure the binding affinities, enthalpy, and entropy for the binding of NanKs with their ligands. The data were analyzed using Origin Analysis software.

analysis shows that NanKs bind to ManNAc and AMPPNP in a 1:1 stoichiometric ratio. Binding affinities and enthalpic and entropic contributions are presented in Table 2. ZnF-lacking *Pm*-NanK binds both ManNAc and AMPPNP independently (Figure S3A,B), which is similar to the ZnF-lacking *Sa*-NanK that was tested previously. This was an expected result as ZnF-lacking NanKs do not have a preference for substrate order of binding.\(^29\) ZnF-containing *Pm*-NanK and *Vc*-NanK bind to ManNAc with micromolar affinity, but they bind to AMPPNP only in the presence of ManNAc (Figure 3C–E and Figure S3A–C). These results suggest that the conformational changes induced by *Pm*-NanK:ManNAc and *Vc*-NanK:ManNAc complex formation are necessary prior to the binding of the nucleotide AMPPNP. This suggests that the ZnF is potentially involved with controlling substrate binding order. However, ZnF-containing *Hi*-NanK binds ManNAc and AMPPNP independently (Figure S3D,E). Therefore, these results suggest that the presence of ZnF alone does not dictate preferential substrate binding, and in the case of *Hi*-NanK, there might be some other structural factors that contribute to the independent binding pattern.

2.5. Structural Analysis of ZnF-Containing and ZnF-Lacking NanK Enzymes. 2.5.1. Crystallographic Details of *Pm*-NanK and *Hi*-NanK. To define the structural basis for the differences in kinetic efficiency and binding properties between the two groups of NanK enzymes, here we report the crystal structures of *Pm*-NanK and *Hi*-NanK. Further, we compared this structural information with previously reported NanK structures.

*Pm*-NanK-unliganded, AMPPNP-bound, and ManNAc-bound forms were crystallized. The *Pm*-NanK:ManNAc crystal structure was refined to a resolution of 2.7 Å, and the positive electron density in the binding pocket demonstrates the presence of ManNAc (Figure S4A). Crystal structures *Pm*-NanK:apo and *Pm*-NanK:AMPPNP (AMPPNP modeled at 0.76 occupancy) were refined to a resolution of 1.9 Å.

Similarly, *Hi*-NanK:ManNAc and *Hi*-NanK:ADP:ManNAc-6P were crystallized. The *Hi*-NanK:ADP:ManNAc-6P crystal structure was refined to a resolution of 2.65 Å. These crystals were obtained by co-crystallization, wherein *Hi*-NanK was incubated with ATP, MgCl\(_2\), and ManNAc at room temperature before setting up of crystallization trays. Interestingly, the crystal structure shows the presence of positive electron density for both products ManNAc-6P and ADP in the binding pocket. This structure illustrates the transfer of γ-phosphate from the ATP to ManNAc due to the enzyme catalytic activity (Figure 4). Likewise, *Hi*-NanK: ManNAc crystals were also obtained by co-crystallization. *Hi*-NanK incubated with sodium orthovanadate, MgCl\(_2\), and ManNAc was crystallized, and the structure was refined with data extending to 2.27 Å resolution. However, the crystal structure shows the presence of only ManNAc and not sodium orthovanadate and/or ADP in the binding pocket (Figure S4B).

The crystallographic details and the refinement statistics of *Pm*-NanK and *Hi*-NanK are presented in Table 3.
Table 3. Crystallographic Data Collection and Refinement Statistics of NanKs<sup>a</sup>

| parameter                                      | Pm-NanK:apo | Pm-NanK:AMPPNP | Pm-NanK:ManNAc | Hi-NanK:ManNAc | Hi-NanK:ManNAc-6P<sup>a</sup> ADP |
|------------------------------------------------|-------------|----------------|--------------|---------------|---------------------------------|
| PDB ID                                         | 6JDH        | 6JDO           | 6JDA         | 6JDC          | 6JDB                            |
| wavelength (Å)                                 | 0.97856     | 0.97857        | 0.97857      | 0.86500       | 0.97857                         |
| resolution range                               | 45.6−1.9 (1.92−1.90) | 45.7−2.0 (2.05−2.00) | 48.1−2.9 (3.08−2.90) | 46.0−2.2 (2.29−2.22) | 46.3−2.6 (2.77−2.64) |
| space group                                    | P 3 2 1     | P 3 2 1        | C 2 2 1      | 4 2 2         | 4 2 2                           |
| Data Collection Statistics                     |             |                |              |               |                                 |
| unit cell a, b, and c (Å)                       | 126.3, 126.3, and 82.4 | 126.4, 126.4, and 82.9 | 103.8, 173.7, and 48.1 | 91.1, 91.1, and 184.0 | 92.6, 92.6, and 183.3 |
| a, b, and c (°)                                 |             |                |              |               |                                 |
| total reflections                               | 579,287 (29614) | 514,396 (36298) | 62,338 (9925) | 69,837 (4746) | 71,480 (9401)                   |
| unique reflections                              | 59,936 (3836) | 51,733 (3739)  | 9983 (1573)  | 19,197 (1661) | 12,021 (1528)                   |
| multiplicity                                    | 9.7 (7.7)   | 9.9 (9.7)      | 6.2 (6.3)    | 3.6 (2.9)     | 5.9 (2.6)                       |
| completeness (%)                                | 99.9 (99.5) | 99.9 (98.5)    | 99.5 (99.8)  | 98.1 (93.8)   | 99.6 (98.0)                     |
| mean I/σ(I)                                    | 17.2 (1.8)  | 15.0 (2.3)     | 13.4 (3.7)   | 11.3 (2.1)    | 11.9 (1.9)                      |
| Wilson B factor                                 | 27.14       | 35.27          | 56.20        | 55.67         | 57.23                           |
| R<sub>merge</sub>                               | 0.08 (1.2)  | 0.07 (1.2)     | 0.09 (0.45)  | 0.07 (0.41)   | 0.09 (0.95)                     |
| R<sub>free</sub>                                | 0.09 (1.3)  | 0.08 (1.2)     | 0.09 (0.49)  | 0.08 (0.52)   | 0.10 (1.04)                     |
| R<sub>free</sub> for ADP                       | 0.03 (0.56) | 0.03 (0.39)    | 0.04 (0.20)  | 0.04 (0.28)   | 0.04 (0.41)                     |
| CCl1/2                                         | 0.9 (0.8)   | 0.9 (0.9)      | 0.9 (0.9)    | 0.9 (0.3)     | 0.9 (0.8)                       |
| Refinement Statistics                          |             |                |              |               |                                 |
| reflections used in refinement                 | 59,897 (5917) | 51,682 (5110)  | 9975 (977)   | 17,929 (1706) | 11,928 (1175)                   |
| reflections used for R free                   | 2997 (321)  | 2624 (230)     | 471 (31)     | 909 (96)      | 572 (57)                        |
| R<sub>work</sub>                                | 0.196 (0.347) | 0.192 (0.298)  | 0.194 (0.345) | 0.217 (0.332) | 0.214 (0.323)                   |
| R<sub>free</sub>                                | 0.230 (0.394) | 0.229 (0.342)  | 0.243 (0.528) | 0.266 (0.362) | 0.276 (0.435)                   |
| RMS (bonds)                                    | 0.006       | 0.007          | 0.009        | 0.008         | 0.008                           |
| RMS (angles)                                   | 0.74        | 0.83           | 1.06         | 0.93          | 0.96                            |
| Ramachandran favored (%)                       | 98.30       | 98.10          | 93.77        | 95.80         | 93.70                           |
| Ramachandran allowed (%)                       | 1.7         | 1.9            | 5.88         | 4.2           | 6.3                             |
| Ramachandran outliers (%)                      | 0.00        | 0.00           | 0.35         | 0.00          | 0.00                            |
| rotamer outliers (%)                           | 0.00        | 0.68           | 0.00         | 0.00          | 0.00                            |
| clashscore                                     | 2.7         | 3.3            | 8.7          | 5.9           | 7.0                             |
| average B factor (Å<sup>2</sup>)               | 38.6        | 43.7           | 63.4         | 44.6          | 64.8                            |
| macromolecules                                 | 38.2        | 43.2           | 63.5         | 44.7          | 64.7                            |
| ligands (occupancy)                            |             |                |              |               |                                 |
| AMPNPNP/ADP                                    |             |                |              |               | 80 (1)                          |
| ManNAc/ManNAc-6P                               |             |                |              |               | 53.3 (1)                        |
| Zn                                             | 30.2 (1)    | 49.7 (1)       | 49.5         | 42.4          | 57.8                            |
| % solvent                                      | 42.6        | 47.7           | 42.4         |               |                                 |

Values in parentheses correspond to the highest resolution shell. Except for the column where the number in the parentheses is the occupancy in the penultimate row.

2.5.2. Structural Analysis of Pm-NanK and Hi-NanK Shows that ManNAc Binding Induces Conformational Change in These Enzymes. Similar to the previously reported bacterial and hMNK structures, the structures reported here from H. influenzae and P. multocida contain N- and C-terminal domains composed of two α/β domains and connected by hinge regions. The monomeric structure is “V” shaped. The C-terminal domain of NanK possesses the dimerization domain, and the resulting dimeric structure is shaped like a butterfly. The N-terminal domain of Hi-NanK consists of residues 1–117 and residues 270–291 of the C-terminus. These are arranged into five long and three short β-strands sandwiched between four α-helices. The C-terminal domain contains residues from 118 to 269 and arranged in four α-helices and four β-strands, sandwiched between N- and C-terminal domains, respectively. Pm-NanK and Hi-NanK display similar structural folds, and the α-helices are numbered from α1 to α9, and β-sheets are numbered from β1 to β13 (Figure S5). The crystal structures of Pm-NanK:apo and Pm-NanK:AMPPNP forms show open conformations, whereas Pm-NanK:ManNAc and Hi-NanK:ManNAc structures exist in a closed conformation. Similar to the previously reported SgGlk, hMNK, and Sa-NanK (ZnF-containing and ZnF-lacking ROK kinases) structures, our structural studies demonstrate that upon ManNAc binding, the N-terminal domain closes over ManNAc to trap the substrate, resulting in the closed conformation of NanK (Figure S5).<sup>26,29,35</sup> The RMS deviation after superposition of Pm-NanK:AMPPNP and Pm-NanK:ManNAc structures is 2.6 Å for 2020 atoms. The RMS deviation after superposition of Pm-NanK:AMPPNP and HiNanK:ManNAc-6P structures is 2.9 Å for 1892 atoms.

Next, we analyzed the domain motions caused by substrate binding apparent in the structures of the NanK enzymes using Dyndom.<sup>37</sup> Dyndom conformational analysis considers the N-terminal domain containing residues from 1 to 63 and the 277-end as the moving domain and the C-terminal domain containing amino acids 64–276 as the fixed domain. The Dyndom analysis of open and closed forms of Pm-NanK:AMPPNP and Pm-NanK:ManNAc shows that upon ManNAc binding, the N-terminal domain rotates by 22.2°
toward the C-terminal domain. The analysis also shows residues 64−66, 70−83, 104−108, and 270−278 as hinge regions, while residues 104−108 and 270−278 are located between the two domains. Further, Dyndom conformational analysis of Hi-NanK:ManNAc-6P:ADP and Hi-NanK:ManNAc structures suggests similar hinge regions.

2.5.3. ManNAc Binding Pocket. The binding pocket for ManNAc is buried deep inside the cleft region between N- and C-terminal domains. The electron density difference map, F_o−F_c and 2F_o−F_c shows the presence of ManNAc in the binding pocket, and it appears in α-D chair conformation in the complex structures. Previous multiple sequence alignment studies of functionally distinct ROK family members clearly show the presence of a very-well-conserved ExGH motif that interacts with ManNAc. However, in the case of NanK enzymes from Gram-negative bacteria, Glu in the ExGH motif is replaced by a His residue. Whereas in the case of ZnF-lacking Fn- and Sr-NanK enzymes, the Glu residue in the ExGH motif is retained. Further, in Sa-NanK, the His residue in the ExGH motif is also replaced by tyrosine29 (Figure S7).

The ManNAc binding sites in both Pm-NanK and Hi-NanK are very well conserved, and the majority of the residues contributing to ManNAc binding are from the N-terminal domain. Only His153, His156 from the ExGH motif, and Glu175 from the C-terminal domain play a role in ManNAc binding. The crystal structures of Pm- and Hi-NanK enzymes clearly show that the hydroxyl group at the C1 position forms hydrogen bonds with His156 (part of the ExGH and ZnF motif) and Glu175. His153 and Gly64 form hydrogen bonds with the hydroxyl group at the C3 position. Similarly, Asn103 and Asp104 form hydrogen bonds with the hydroxyl group at the C4 position, while main chains of Leu74 and Asn75 form hydrogen bonds with the hydroxyl group at the C7 position. Phosphorylation occurs at the C6 hydroxyl group position, and it interacts with the catalytic Asp104.
2.5.4. Detailed Structural Characterization of the ATP Binding Pocket and Its Interacting Residues. The binding pocket for the nucleotide is located toward the C-terminal domain. In the Pm-NanK:AMPPNP structure, the γ-phosphate of AMPPNP shows a hydrogen bonding interaction with Ser130. Additionally, adenine and ribose moieties of AMPPNP are sandwiched between residues Gly180, Pro196, Phe200, Ala246, Ser242, and Val243 from the C-terminal domain. In the closed conformation of the Hi-NanK:ManNAc-6P:ADP structure, pyrophosphate makes extensive interactions with residues from both the N- and C-terminal domains. Since the C-terminal domain is the dimerization domain, we fixed the C-terminal domain and superposed Pm-NanK:AMPPNP (open conformation) with the Hi-NanK-ManNac-6p (closed conformation) structure. This superposition analysis highlights that upon ManNac binding, the loop between the β1 and β2 strands, which holds the ATP binding motif DIGG, makes a large movement (approximately moves about 5.7 Å) and moves closer to the AMPPNP, which further facilitates the phosphorylation event. A representative example is Gly10 (from the DIGG motif of the ATP binding loop), where the distance between β-phosphate of AMPPNP and Gly10 in open and closed conformations of NanK enzymes is 9.2 and 3.5 Å, respectively (Figure 6). This structural information suggests that there are no conformational changes in the nucleotide-bound states, but upon ManNac binding, there are dramatic conformational changes leading to the hydrolysis of ATP and product formation. These observations support the hypothesis that conformational change in NanK only allows for ATP hydrolysis when N-acetylmannosamine is present.

Superimposing the structures of Hi/Pm-NanK on hMNK shows a very-well-conserved nucleotide binding pocket with only minor differences. Previous structural studies of hMNK have shown that a Mg2+ ion coordinates axially with both β-phosphate of ADP and Asp413 in the active site and is crucial for ADP binding.46 Unexpectedly, density for Mg2+ is not observed in any of the structures reported here. Notably, the overall distance between Asp7 and β-phosphate-oxygen of ADP from Hi-NanK:ManNac-6P:ADP is similar to the distance between Asp413 and β-phosphate-oxygen of ADP from hMNK. This structural information along with our thermodynamic data studies demonstrates that in these structures without the Mg2+, the productive orientation of ATP and ManNac is stabilized by their neighboring amino acids.

2.5.5. Thr131 is a Catalytically Important Residue. The open conformation evident in the Pm-NanK:AMPPNP structure suggests that the residues Ser130 and Thr131 are at a hydrogen bond distance from the triphosphate group of AMPPNP. Whereas, the Hi-NanK:ManNac-6P:ADP structure suggests that upon ManNac binding, the movement of the loop containing the ATP-binding motif results in the β-phosphate being positioned in between Thr11 (from the ATP-binding motif loop) and Thr131. This is similar to hMNK and Sa-NanK structures, where the β-phosphate is positioned in between residues Thr417 and Thr544 in hMNK and the residues Thr11 and Thr134 in Sa-NanK. Previous reports by Martinez et al. suggest that Thr544 is important for stabilization of the β-phosphate, which is supported by Coomes et al., who suggest that Asp7, Thr11, and Thr134 are important to support electron withdrawal during the catalytic cycle.26,29 Based on these similar structural features, we believe that bacterial NanK enzymes studied here follow a similar phosphorylation mechanism that was previously proposed.26,29 Consistent with this, sequence and structural alignment of NanK enzymes show that Thr131 is broadly conserved among the different NanK enzymes (Figure S1). Thus, we hypothesize that Thr131 plays a key catalytic role during the phosphorylation event.

To test this hypothesis, we generated and purified the Hi-NanK T131V mutant enzyme and then compared its thermodynamic and kinetic properties to the Hi-NanK wild-type enzyme. Consistent with our hypothesis, kinetic studies demonstrate that the Hi-NanK T131V mutant enzyme is enzymatically inactive—it is a kinase-dead mutant. Surprisingly, ITC studies demonstrate that the Hi-NanK T131V mutant is still able to bind the substrates with a similar affinity to the wild-type Hi-NanK enzyme (Table S2 and Figure S6), suggesting that Thr131 mutation does not affect the overall fold of the enzyme or its ability to bind the substrates. As previously hypothesized by two other groups, our results support the role of Thr131 in the electron withdrawal from γ-phosphate of ATP during substrate phosphorylation.

2.5.6. ZnF-Binding Pocket in NanK Enzymes. Apart from the Pm-NanK:AMPPNP, the other refined structures of Pm-NanK and Hi-NanK show clear density positioned near the ManNac-binding pocket that we modeled as a Zn ion. We attempted modeling other ions, such as Mg2+, but they did not satisfy either the coordination or the electron density during refinement. Similar to the previously reported ZnF-containing NanK structures25,26 (putative Ec-NanK PDB ID: 2AA4), tetrahedral coordination of Zn is satisfied by three cysteines (Cys166, Cys168, and Cys173) and His156 from the ExGH motif (numbering corresponds to Hi-NanK), which are positioned approximately at a distance of 2.3 Å. In all these structures, the B factors for Zn2+ are similar to B factors of the coordinating residues, which suggests a full occupancy for Zn2+ in all the structures. Apart from its association with Zn2+, His156 also forms a hydrogen bond with the hydroxyl group at the C1 position of ManNac. These structural details suggest that the ZnF motif plays an important role in proper positioning of His156, which further helps in the positioning of ManNac in the active site (Figure 7A). Further, structural and sequence comparison of NanK enzymes shows that Glu175, which is present on the short α4 helix following the ZnF motif, interacts with the C1 hydroxyl group and is highly conserved across the species ranging from prokaryotes to eukaryotes. Structural details also indicate that the ZnF motif also plays an important role in positioning of the Glu175, which in turn helps in the positioning of the short α4 helix and the 310-helix between the β11 and α5 helix. These structural details illustrate that the ZnF motif along with highly conserved Glu175 plays an important structural role and orients ManNac in the active site.

Overall, these structural data highlight that although the Zn-binding and ATP-binding pockets are located on the opposite sides of the enzyme, upon substrate binding, there are series of inter-residue interactions between different regions of the enzyme that extend from the ZnF motif to Glu175, to ManNac, and then to ATP. These interactions are responsible for inactive (open) and active (closed) conformations of the enzyme. This analysis also shows the significance of a metal-binding center in the structural integrity of the ZnF-containing NanK enzymes (Figure S8).

2.5.7. Differences between ZnF-Containing and ZnF-Lacking NanK Enzymes. In this paper, we explore the
structural differences in the binding pockets of ZnF-containing and ZnF-lacking NanK enzymes by comparing the Gram-negative bacterial NanK structures of Pm/Hi-NanK with the Fn-NanK structure. Structural superposition of Hi-NanK and Pm-NanK with the previously reported hMNK and putative E. coli NanK enzymes shows conserved ExGH and ZnF motifs (Figure 7A). In contrast, structural superposition of ZnF-containing and ZnF-lacking NanKs shows that there are clear differences near the putative Zn-binding region. Closer examination of the structural details of the Fn-NanK structure, which lacks the ZnF motif, shows that ManNAc is held in the active site through its interactions with the residues Glu156, His159 (from ExGH motif), and Glu168 (equivalent to Glu175 of Hi-NanK). The cysteine-rich loop that holds Zn is absent in Fn-NanK, and it is replaced by a short loop of eight amino acids. Surprisingly, the distance between His159 and C1 of ManNAc is somewhat further compared to its equivalent residues in Hi/Pm-NanK. Though there are no observable interactions between His159 and ManNAc, we tested its role in ManNAc binding by mutational analysis. In contrast to the Fn-NanK wild type, the thermodynamic properties of H159L show that it binds to ManNAc only in the presence of AMPPNP (Table S3). The enzymatic assay shows that the catalytic efficiency of the enzyme is reduced (apparent $K_M = 994 \mu M$, $V_{max} = 1.147 \mu M/min$, $K_{cat} = 37 \text{ min}^{-1}$, and $K_{cat}/K_M = 0.04 \mu M^{-1} \text{ min}^{-1}$). Active-site residues, Glu156, His159, and Glu168 residues in wild-type Fn-NanK are held in place by a combination of hydrophobic and hydrogen bonding interactions formed by Phe167 present in the helix following the eight-amino acid-short loop. The side chains of Phe167 in Fn-NanK are held tightly in the hydrophobic pocket formed by Ile136, Ile160, Ala171, Leu215, and Gly218. Among these five residues, Leu215 and Gly218 are located on the large a7 helix, which runs under the putative Zn-binding motif. Further, the main chain amino group of Phe167 makes hydrogen bonds with main chain carbonyl residues of Gly158 and Ile160 (Figure 7B).

Based on this structural information, we mutated F167 to F167A and F167Q in Fn-NanK to check the effect of these substitutions on the hydrophobic pocket. We observed very low protein expression of the Fn-NanK F167A mutant in addition to its high instability. Similarly, there was no protein expression for the Fn-NanK F167Q mutant. These studies clearly show that a hydrophobic environment is essential for the structural organization of the Fn-NanK active site.

Previous reports by Coombes et al. have shown that the ZnF motif is absent in Gram-positive bacterial Sa-NanK, and ExGH is replaced by the ExGY motif. Comparison of the structure demonstrates that Sa-NanK has overall structural resemblance to ZnF-lacking Fn-NanK. However, the region equivalent to the ZnF motif is structurally maintained by the arginine-stacking interaction, and tyrosine in the ExGY motif is longer than histidine, which enables it to associate with the substrate (Figure 7C). Potentially, this is beneficial for Sa-NanK as it has an apparent $K_M$ similar to the ZnF-containing NanKs and a higher catalytic efficiency. While, the hydrophobic pocket is involved in the proper structural organization of the active site in Fn-NanK, His159 from the ExGH motif is approximately 5 Å away from the C1 hydroxyl group of ManNAc. Therefore, we hypothesize that this difference might be the probable reason for 10-fold increase in apparent $K_M$ of the Fn-NanK enzyme in comparison to other NanK enzymes.

2.5.8. Mutational Analysis of Residues in the Hinge Region. To understand the differential binding properties of Hi-NanK to its substrate, we compared Hi- and Pm-NanK structures. Based on the structural information, we hypothesized that though their binding pocket residues are very well conserved, the residues in the hinge regions might play a role in the differential binding properties of these two enzymes. We purified the mutants Pm-NanK D115L (mutation in the a1-helix between N- and C-terminal domains), Pm-NanK H273F (mutation in the hinge region, located in the loop connecting the last $\beta$-12 strand to the last $\alpha$-11 helix), and the double
mutant *Pm*-NanK D115L, H273F, subjected to ITC studies to test whether the random binding pattern of *Hi*-NanK will be observed in *Pm*-NanK mutants. Surprisingly, no changes were observed in the substrate-binding properties of *Pm*-NanK mutants compared to the wild type. The structural reasons for the differential binding properties of *Hi*-NanK in comparison to other ZnF-containing NanK enzymes will have to be addressed in the future.

3. CONCLUSIONS

*N*-Acetylmannosamine kinases are involved in the anabolic and catabolic pathways of sialic acids in human and bacteria, respectively. These kinases belong to the ROK superfamily of enzymes, and most of the kinases from this family contain a conserved ZnF motif, where the Zn ion is coordinated by one histidine and three cysteines. Previous structural studies have shown that the ZnF motif is absent in *Fn*-NanK, but its structural properties are well maintained. To understand the importance of the ZnF motif in NanKs, we have carried out functional studies on *Hi*-NanK, *Pm*-NanK, and *Vc*-NanK (contains the ZnF motif) in comparison to *Fn*-NanK (lacks the ZnF motif) to its substrates ManNAc and ATP, plus with its nucleotide analogueAMPPNP. These functional studies show few minor changes between these two groups of kinases.

Similar to earlier studies on hMNK and putative *E. coli*-NanK, our structural analysis on NanKs from other Gram-negative bacteria like *H. influenzae* and *P. multocida* shows that the ZnF motif helps in proper conformation of His156 and Glu167 (numbering corresponds to *Hi*-NanK). These residues in turn are responsible for the proper conformation of ManNAc in the binding pocket. Interestingly, previous studies on Ros homologues have shown that during the course of evolution, prokaryotic ZnFs are successfully replaced by either polar or hydrophobic residues, which surrogate the function of ZnF domains. Likewise, recent studies by Coombes et al. have shown that the loss of the ZnF motif in the Gram-positive bacterial *S. aureus*-NanK is replaced by an arginine stack that plays a similar role. Our current study on *Fn*-NanK shows yet another solution, whereby the ZnF motif is replaced by a hydrophobic pocket formed by Phe167 and its interacting hydrophobic residues.

The exact evolutionary reason for the loss of the ZnF motif in *Fn*-NanK is not clear. However, previous phylogenetic analysis shows that the core genome of *F. nucleatum* is very different when compared to other bacterial lineages. Due to horizontal gene transfer, it has parts of the genome from Firmicutes, Bacteroidetes, Proteobacteria, and Spirochaetes. Early studies have also hypothesized that as *F. nucleatum* is closely associated with other bacteria in dental plaque, it has acquired niche-specific genes due to strong environmental selection pressure, but it retained its Gram-negative cell wall genes to protect itself from the host immune system.

As the ZnF motif is present in species within the genus, such as *F. mortiferum* NanK (60% sequence identity), we can speculate that a *Fusobacterium* ancestral species lost the ZnF motif only recently.

Altogether, to preserve functional properties, the loss of the ZnF motif is compensated by alternative changes in protein sequences at the Zn metal-binding center. These changes in *Fn*-NanK guaranteed the correct, stable, and functional fold of the protein for its kinase activity.

In summary, our work shows that in spite of low sequence identity among NanK enzymes, different pathogenic bacteria have evolved to carry out the function of phosphorylation by conserving most of their signature motifs. Further, our structural and functional studies highlight that though the ZnF motif is absent in certain bacteria, they evolved with few compensatory amino acid changes near the Zn-binding pocket. This systematic characterization of amino acid network using biophysical studies provides the deeper understanding of molecular-level details in enzyme activity and function. In the future, these distinct structural features can be leveraged for designing specific inhibitors that can target the NanK enzymes of pathogenic bacteria.

4. MATERIALS AND METHODS

4.1. Cloning and Expression of *Hi*-NanK, *Pm*-NanK, *Fm*-NanK, and *Vc*-NanK

The genes corresponding to NanK from *Fusobacterium nucleatum* (*Fn*-NanK; NCBI reference sequence no.: WP_011017180.1), *Haemophilus influenzae* (*Hi*-NanK; NCBI reference sequence no.: WP_011271901.1), *Pasteurella multocida* (*Pm*-NanK; NCBI reference sequence no: WP_005752223.1), and *Vibrio cholerae* (*Vc*-NanK; NCBI reference sequence no.: WP_001259414.1) were synthesized from GeneArt Gene Synthesis (Thermo Fisher Scientific). These genes were cloned into a pET300/NT-DEST vector with an N-terminal 6X His tag using an Invitrogen Gateway cloning technology (Table S4), and the details of these have been previously described by Bairy et al. Further, these plasmids were transformed into BL21(DE3) cells for protein expression. *Hi*-NanK T131V, *Pm*-NanK D115L, *Pm*-NanK H273F, *Pm*-NanK D115L H273F, *Fn*-NanK H159L, *Fn*-NanK F167A, and *Fn*-NanK F167A, respectively. These kinases belong to the ROK superfamily of catabolic pathways of sialic acids in human and bacteria, suggesting that the core genome of *F. nucleatum* is protected by an arginine stack that plays a similar role.

**Table 4.** Comparison of *Hi*-NanK, *Pm*-NanK, *Fm*-NanK, and *Vc*-NanK with other bacterial lineages. Due to horizontal gene transfer, it has parts of the genome from Firmicutes, Bacteroidetes, and Spirochaetes. Early studies have also hypothesized that as *F. nucleatum* is closely associated with other bacteria in dental plaque, it has acquired niche-specific genes due to strong environmental selection pressure, but it retained its Gram-negative cell wall genes to protect itself from the host immune system.38,39 As the ZnF motif is present in species within the genus, such as *F. mortiferum* NanK (60% sequence identity), we can speculate that a *Fusobacterium* ancestral species lost the ZnF motif only recently.

Altogether, to preserve functional properties, the loss of the ZnF motif is compensated by alternative changes in protein sequences at the Zn metal-binding center. These changes in *Fn*-NanK guaranteed the correct, stable, and functional fold of the protein for its kinase activity.

In summary, our work shows that in spite of low sequence identity among NanK enzymes, different pathogenic bacteria have evolved to carry out the function of phosphorylation by conserving most of their signature motifs. Further, our structural and functional studies highlight that though the ZnF motif is absent in certain bacteria, they evolved with few compensatory amino acid changes near the Zn-binding pocket. This systematic characterization of amino acid network using biophysical studies provides the deeper understanding of molecular-level details in enzyme activity and function. In the future, these distinct structural features can be leveraged for designing specific inhibitors that can target the NanK enzymes of pathogenic bacteria.

**Table 4.** Comparison of *Hi*-NanK, *Pm*-NanK, *Fm*-NanK, and *Vc*-NanK with other bacterial lineages.

- **Table 4.** Comparison of *Hi*-NanK, *Pm*-NanK, *Fm*-NanK, and *Vc*-NanK with other bacterial lineages. Due to horizontal gene transfer, it has parts of the genome from Firmicutes, Bacteroidetes, and Spirochaetes. Early studies have also hypothesized that as *F. nucleatum* is closely associated with other bacteria in dental plaque, it has acquired niche-specific genes due to strong environmental selection pressure, but it retained its Gram-negative cell wall genes to protect itself from the host immune system.38,39 As the ZnF motif is present in species within the genus, such as *F. mortiferum* NanK (60% sequence identity), we can speculate that a *Fusobacterium* ancestral species lost the ZnF motif only recently.

Altogether, to preserve functional properties, the loss of the ZnF motif is compensated by alternative changes in protein sequences at the Zn metal-binding center. These changes in *Fn*-NanK guaranteed the correct, stable, and functional fold of the protein for its kinase activity.

In summary, our work shows that in spite of low sequence identity among NanK enzymes, different pathogenic bacteria have evolved to carry out the function of phosphorylation by conserving most of their signature motifs. Further, our structural and functional studies highlight that though the ZnF motif is absent in certain bacteria, they evolved with few compensatory amino acid changes near the Zn-binding pocket. This systematic characterization of amino acid network using biophysical studies provides the deeper understanding of molecular-level details in enzyme activity and function. In the future, these distinct structural features can be leveraged for designing specific inhibitors that can target the NanK enzymes of pathogenic bacteria.

**Table 4.** Comparison of *Hi*-NanK, *Pm*-NanK, *Fm*-NanK, and *Vc*-NanK with other bacterial lineages. Due to horizontal gene transfer, it has parts of the genome from Firmicutes, Bacteroidetes, and Spirochaetes. Early studies have also hypothesized that as *F. nucleatum* is closely associated with other bacteria in dental plaque, it has acquired niche-specific genes due to strong environmental selection pressure, but it retained its Gram-negative cell wall genes to protect itself from the host immune system.38,39 As the ZnF motif is present in species within the genus, such as *F. mortiferum* NanK (60% sequence identity), we can speculate that a *Fusobacterium* ancestral species lost the ZnF motif only recently.

Altogether, to preserve functional properties, the loss of the ZnF motif is compensated by alternative changes in protein sequences at the Zn metal-binding center. These changes in *Fn*-NanK guaranteed the correct, stable, and functional fold of the protein for its kinase activity.

In summary, our work shows that in spite of low sequence identity among NanK enzymes, different pathogenic bacteria have evolved to carry out the function of phosphorylation by conserving most of their signature motifs. Further, our structural and functional studies highlight that though the ZnF motif is absent in certain bacteria, they evolved with few compensatory amino acid changes near the Zn-binding pocket. This systematic characterization of amino acid network using biophysical studies provides the deeper understanding of molecular-level details in enzyme activity and function. In the future, these distinct structural features can be leveraged for designing specific inhibitors that can target the NanK enzymes of pathogenic bacteria.
theoretical molar extinction coefficients of different NanKs. Molar extinction coefficients were obtained by using ProtParam on the ExPASy web server. The purified protein was further utilized for structural and functional studies.

4.3. Kinetic Enzymatic Assays. Kinase activities for different NanKs were carried out using an ADP-Glo kinase assay kit (Promega). The reaction was performed in a buffer containing 40 mM Tris–HCl at pH 7.5, 20 mM MgCl₂, and 0.1 mg/mL BSA. The kinase assay with ManNAc was performed using 10 ng of the enzyme, 100 μM ATP, and varying concentrations of ManNAc. The reactions were incubated at room temperature for 15 min. Reagent 1 was added to 5 μL of the reaction mixture to deplete the unutilized ATP, and the reaction was further incubated for 40 min. Next, reagent 2 was added to the reaction mixture that converts ADP (liberated during phosphorylation reaction) to ATP, which was further converted into light by luciferase. The luminescence was measured post 1 h of incubation using a plate reader (Tecan). The data was analyzed, and the apparent Km values for ManNAc were calculated by nonlinear fitting of the data into the Michaelis–Menten equation. Each point represents the average of three independent samples, and the curves were plotted using GraphPad Prism (version 7.0b). These kinetic enzymatic assays were performed in experimental duplicates with technical triplicates.

4.4. Analysis of ManNAc and ManNAc-6P by Mass Spectrometry. To measure the amount of the substrate and product formed after the NanK enzymatic activity, three different quantities (1, 100, and 1000 ng) of Pm-NanK and Fn-NanK enzymes were used for each reaction and followed the protocol described previously. Briefly, Fn-NanK or Pm-NanK at a final concentration of 1, 100, and 1000 ng were added to 50 μL of the reaction mixture containing 50 mM Tris–HCl at pH 8.0, 5 mM ManNAc, 10 mM ATP, and 10 mM MgCl₂. These reactions were carried out at 37 °C for 20 min and stopped by adding TCA (trichloroacetic acid) to a 10% final concentration followed by placing the reaction mixture on ice for 10 min. These reactions were centrifuged at 13,000 rpm for 10 min, then the supernatants were collected and subjected to mass spectrometry analysis. The 10 μL supernatant was diluted with 90 μL of MS-grade water in a low-protein binding microcentrifuge tube. Subsequently, freshly made 100 μL of 25 mM 3-aminoo-9-ethylcarbazole and 50 μL of 50 mM NaCNBH₃ were added to the supernatant. The reaction mixtures were further incubated at 70 °C for 60 min and then kept on ice for 1 min. Next, 300 μL of a dichloromethane/hexane (2:1) mixture was added to the above 300 μL of MS-grade water. These reaction mixtures were vortexed and centrifuged at 10,000 rpm for 5 min. Approximately 300 μL of the upper aqueous phase was transferred to a fresh low-protein binding microcentrifuge tube without disturbing the lower phase. Finally, the sample was further injected into a TSQ Vantage-Agilent 1290 UHPLC (LC–MS) machine for MRMs (multiple reaction monitoring) of parent and daughter ions of both the reactant and product at a constant injection rate using a Hamilton syringe with a flow rate of 5 μL/min in positive ion mode. The daughter ion of the 3-aminoo-9-ethylcarbazole characteristic (AEC) derivatization moiety of 210 Da from the parent ion was used for quantification.

4.5. Isothermal Titrination Calorimetry (ITC). The binding affinities between NanK and its substrates, ManNAc and AMPPPNP (ATP analogue), were measured using a MicroCal ITC system (Malvern). The reactions were carried out in a buffer containing 20 mM Tris–HCl at pH 8.0, 300 mM NaCl, 5% glycerol, and 1 mM DTT. The protein concentration in the reaction cell was varied from 100 to 200 μM. Correspondingly, the syringe concentration used in the experiments was 10–20 times higher than the cell concentration. Two different series of thermodynamic reactions were carried out to measure the binding affinities of NanKs to their substrates.

The first set of titrations was carried out between different NanKs and ManNAc in the presence and absence of AMPPPNP. The second set of titrations was carried out between different NanKs and AMPPPNP in the presence and absence of ManNAc. Concentrations of NanK, AMPPPNP, and ManNAc were varied in the titrations, and the experiments were performed in duplicates with technical duplicates or triplicates. The nonspecific heat released by the dilution of the protein, nucleotide, and substrate was calculated by averaging the heat liberated during the last three to five injections post saturation. The values were further subtracted from the raw heat released during each injection, which eliminates the heat of dilution. The calorimetric data sets were further analyzed using Origin ITC analysis software (MicroCal, USA) for the single-site binding model. Nonlinear least-square analysis was used to calculate the stoichiometry, enthalpy (ΔH), and binding affinity.

4.6. Protein Crystallization and Data Collection. 4.6.1. Pm-NanK. Post purification, the protein was concentrated to 20 mg/mL using 10 kDa cutoff Amicon tubes (Millipore). Then, 96-well crystallization trays were used for setting up crystal trays, and 100 μL of different crystallization solutions were manually pipetted into all the wells. Further crystallization trays were set up by a Mosquito (nanodrop liquid handling machine, TTP Labtech) using the hanging drop vapor diffusion method. In these crystallization drops, 200 μL of protein was mixed with 200 μL of different crystallization solutions.

4.6.1.1. Pm-NanK Unliganded (apo). The crystals were obtained in an optimized condition from a Qiagen classic suite crystallization screen, and the buffer contained 0.1 M Tacsimate at pH 5.0, 20% v/v isopropanol, and 20% w/v PEG 4000 with an additive containing 2 M NaCl.

4.6.1.2. Pm-NanK:AMPPNP. Initially, the crystals were obtained in a buffer containing 0.1 M Tacsimate at pH 5.0, 20% v/v isopropanol, and 20% w/v PEG 4000 with an additive containing 100% v/v ethylene glycol. These crystals were reproduced in similar crystallization conditions in 24-well crystallization trays, where 0.5 μL of protein was mixed with 0.5 μL of crystallization solution. These crystals were further soaked for 7 days in the same buffer containing AMPPPNP and ManNAc that were 20 times higher in concentration than the protein concentration (however, the obtained crystal structure contains only AMPPPNP).

4.6.1.3. Pm-NanK:ManNAc. The protein and ManNAc were mixed at a 1:10 molar ratio and incubated on ice for 1 h. The crystals were obtained in a buffer containing 0.2 M lithium citrate tribasic tetrahydrate and 20% w/v PEG 3350 (PEG/ion screen from Hampton Research). Pm-NanK unliganded, Pm-NanK:AMPNP, and Pm-NanK:ManNAc crystals were mounted in loops and flash cooled in the mother liquor containing 10% v/v glycerol. The X-ray diffraction data was collected at 100 K at the Proxima 1 beamline, SOLEIL Synchrotron, France.
4.6.2. Hi-NanK. Post purification, the protein was concentrated to 20 mg/mL using 10 kDa cutoff Amicon tubes (Millipore). For Hi-NanK also, 96-well crystallization trays were used for setting up crystal trays, and 100 µL of different crystallization solutions was taken in these trays. The crystallization trays were set up at a protein concentration of 21 mg/mL. The protein was further incubated with ATP, MgCl₂, and ManNAc at a 1:5:5:5 molar ratios at RT for 30 min. The crystals were obtained in a buffer containing 0.2 M sodium malonate at pH 7.0 and 20% w/v PEG 3350 (PEG/ion 2 screen from Hampton Research).

4.6.2.1. Hi-NanK:ManNAc. The crystal trays were set up at a protein concentration of 20 mg/mL. Further, the protein was further incubated with ATP, MgCl₂ and ManNAc at a pH 7.0 and 20% w/v PEG 3350 (PEG/ion 2 screen from Hampton Research). The protein was mixed with 200 nL of different crystallization solutions.

4.6.2.2. Hi-NanK:ManNAc-6P:ADP. The crystals were set up at a protein concentration of 21 mg/mL. The protein was further incubated with ATP, MgCl₂ and ManNAc at a pH 8.5 and 50% v/v MPD (classic suite from Hagena, Germany).

Hi-NanK:ADP:ManNAc-6P crystals were mounted in loops, and the X-ray diffraction data was collected at 100 K using the Proxima 1 beamline, SOLEIL Synchrotron, France. Hi-NanK:ManNAc complex crystals were mounted in loops, and X-ray diffraction data was collected at the ID29 beamline, ESRF, Grenoble, France.

4.7. Refinement of NanK Protein Structures. All diffraction images were processed by XDS, and the data were scaled using Aimless in the CCP4 program suite. The structure of Pm-NanK:ManNAc was determined by molecular replacement using the putative NanK from E. coli (PDB entry 2AA4) as the search model. Further, the Pm-NanK:ManNAc structure was used as the search model for Pm-NanK apo and AMPNNP-bound structures. The structure of the Hi-NanK:ADP:ManNAc-6P complex was also determined by molecular replacement using the structure of the Pm-NanK:ManNAc complex as the search model. Further, the Hi-NanK:ADP:ManNAc-6P complex structure was used as a search model for the molecular replacement of the Hi-NanK:ManNAc structure. Finally, manual model building was carried out using COOT, and further refinements and processing were performed using Phenix. Atomic coordinates and structure factors for the five structures were deposited in Protein Data Bank (PDB), and their IDs are as follows: Pm-NanK.apo - 6JDH, Pm-NanK:AMPNNP - 6JDO, Pm-NanK:ManNAc - 6JDA, Hi-NanK:ManNAc-6P:ADP - 6JDB, and Hi-NanK:ManNAc - 6JDC. Crystallographic data for data collection and refinement are presented in Table 3. Molecular graphics figures were prepared by PyMol (The PyMOL Molecular Graphics System, Schrödinger). DynDom analysis of the resulting protein structures were carried out in the DynDom server (http://dyndom.cmp.uea.ac.uk/dyndom/) by providing the coordinates of the pair of structures with and without the ligand.

**ASSOCIATED CONTENT**

1. Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.0c03699.
contribution from co-authors. D.C. and R.C.J.D contributed to analyzing kinetic data, critical reading of the manuscript, and discussion of the results.

Notes
The authors declare no competing financial interest.

## ACKNOWLEDGMENTS
We thank the ESRF Access Program of RCB (supported by Department of Biotechnology grant BT/INF/22/SP22660/2017). We thank SOLEIL Synchrotron (Proxima 1 beamline) and ESRF (ID29 beamline) for providing the beamtime for data collection. This research was supported by DDT-Indo Swedish grant (BT/IN/SWEDEN/41/01/2013), DRT-B-life grant (BT/PR5801/INF/156/2012), NCBS X-ray facility grant (BT/PR12422/MED/31/287/214), and a grant from the DST for SR. We thank Nitish Sathyanarayana, Dr. Swagatha Ghosh, and Dr. Rhawnie Caing-Carlsson for their contributions toward the initial stages of the work; Dr. Chetan Arya for the help with the ChemDraw figures of the sialic acid pathway; and Dr. Sanchari Banerjee for critical reading of the manuscript. T.G.S. would like to acknowledge senior research fellowship support from the Council of Scientific and Industrial Research (CSIR).

## ABBREVIATIONS

| Abbreviation | Description |
|--------------|-------------|
| NanK | N-acetylmannosamine kinase |
| h | human |
| GNE | UDP-GlcNAc 2-epimerase/ManNAc kinase |
| ITC | isothermal titration calorimetry |
| AEC | 3-amino-9-ethylcarbazole |
| ZnF | zinc Finger |
| NanA | N-acetyleneuraminic acid lyase |
| ManNAc | N-acetylmannosamine |
| ManNAc-6P | N-acetylmannosamine-6-phosphate |
| NanE | N-acetylenamnosamine-6-P epimerase |
| NagA | N-acetylglucosamine-6-phosphate deacetylase |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| TCA | trichloroacetic acid |
| MRMs | multiple reaction monitors |
| Neu5Aca | N-acetylneuraminic acid |
| Neu5Gc | N-glycolylneuraminic acid |
| LOS | lipooligosaccharide |
| LPS | lipopolysaccharide |
| LOS | lipooligosaccharide |

## REFERENCES

1. Vallee, B. L.; Auld, D. S. Zinc Coordination, Function, and Structure of Zinc Enzymes and Other Proteins. Biochemistry 1990, 29, 5647–5659.

2. Vallee, B. L.; Falchuk, K. H. The Biochemical Basis of Zinc Physiology. Physiol. Rev. 1993, 73, 79–118.

3. Amann, B. T.; Worthington, M. T.; Berg, J. M. A Cys3His Zinc-Binding Domain from Nup475/Tristetraprolin: A Novel Fold with a Switches? Antioxid. Redox Signaling 2003, 5, 1261–1268.

4. Michalek, J. L.; Besold, A. N.; Michel, S. L. J. Cysteine and Histidine Shuffling: Mixing and Matching Cysteine and Histidine Residues in Zinc Finger Proteins to Afford Different Folds and Function. Dalton Trans. 2011, 40, 12619–12632.

5. Malgieri, G.; Palmieri, M.; Russo, L.; Fattorusso, R.; Pedone, P. V.; Isernia, C. The Prokaryotic Zinc-Finger I: Structure, Function and Comparison with the Eukaryotic Counterpart. FEBS J. 2015, 282, 4480–4496.

6. Baglivo, I.; Russo, L.; Esposito, S.; Malgieri, G.; Gendra, M.; Salluzzo, A.; Di Blasio, B.; Isernia, C.; Fattorusso, R.; Pedone, P. V. The Structural Role of the Zinc Ion Can Be Dispensable in Prokaryotic Zinc-Finger Domains. Proc. Natl. Acad. Sci. U. S. A. 2009, 106, 6933–6938.

7. Coleman, J. E. Zinc Proteins: Enzymes, Storage proteins, transcription factors, and Replication Proteins. Annu. Rev. Biochem. 1992, 61, 897–946.

8. McCall, K. A.; Huang, C.-C.; Fierke, C. A. Zinc and Health: Current Status and Future Directions. J. Nutr. 2000, 122, 1437–1446.

9. Kluska, K.; Adamczyk, J.; Kręžel, A. Metal Binding Properties, Stability and Reactivity of Zinc Fingers. Coord. Chem. Rev. 2018, 367, 18–64.

10. Tanner, M. E. The Enzymes of Sialic Acid Biosynthesis. Bioorg. Chem. 2003, 33, 216–223.

11. Varki, A.; Gagneux, P. Multifarious Roles of Sialic Acids in Immunity. Ann. N. Y. Acad. Sci. 2012, 1253, 16–36.

12. Kwon, J.; Shao, Z.; Shim, M.; Moore, L. B.; Thompson, S. M.; Miller, B. G. Divergent Evolution of Function in the ROK Sugar Kinase Superfamily: Role of Enzyme Loops in Substrate Specificity. Biochemistry 2007, 46, 13564–13572.

13. Conejo, M. S.; Thompson, S. M.; Miller, B. G. Evolutionary Bases of Carbohydrate Recognition and Substrate Discrimination in the ROK Protein Family. J. Mol. Evol. 2010, 70, 545–556.

14. Nocek, B.; Stein, A. J.; Jezdzejczak, R.; Cuff, M. E.; Li, H.; Volkart, L.; Joachimiak, A. Structural Studies of ROK Fructokinase YdB from Bacillus subtilis: Insights into Substrate Binding and Fructose Specificity. J. Mol. Biol. 2011, 40, 325–342.

15. Tillgenmeyer, F.; Reizer, J.; Reizer, A.; Saier, M. H., Jr. Evolutionary Relationships between Sugar Kinases and Transcriptional Repressors in Bacteria. Microbiology 1994, 140, 2349–2354.

16. Reinker, S. O.; Lehner, G.; Hinderlich, S.; Reutter, W. Regulation and Pathophysiologica! Implications of UDP-GlcNAc 2-Epimerase / ManNAc Kinase (GNE) as the Key Enzyme of Sialic Acid Biosynthesis. Biol. Chem. 2009, 390, 591–599.

17. Tong, Y.; Tempel, W.; Nedryalkova, L.; MacKenzie, F.; Park, H.-W. Crystal Structure of the N-Acetylmannosaminic Kinase Domain of GNE. PLoS One 2009, 4, e7105.

18. Martinez, J.; Nguyen, L. D.; Hinderlich, S.; Zimmer, R.; Tauberger, E.; Reutter, W.; Saenger, W.; Fan, H.; Moniot, S. Crystal Structures of N-Acetylmannosaminic Kinase Provide Insights into Enzyme Activity and Inhibition. J. Biol. Chem. 2012, 287, 13656–13665.

19. Caing-carlsson, R.; Goyal, P.; Sharma, A.; Ghosh, S.; Setty, T. G.; North, R. A.; Friemann, R.; Ramaswamy, S. Crystal Structure of N-Acetylmannosaminic Kinase from Fusobacterium nucleatum. Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 2017, 73, 356–362.

20. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Weight Sequencing, Position-Specific Gap Penalties and Weight Matrix Choice. Nucleic Acids Res. 1994, 22, 4673–4680.
Non-Canonical ROK Family Function in the N-Acetylmuramoyl-L-Araylamine Kinase from the Pathogen Staphylococcus aureus. *J. Biol. Chem.* 2020, 295, 3301–3315.

(30) Schiefer, A.; Gerber, K.; Seitz, S.; Welte, W.; Diederichs, K.; Boos, W. The Crystal Structure of Mlc, a Global Regulator of Sugar Metabolism in *Escherichia coli*. *J. Biol. Chem.* 2005, 280, 29073–29079.

(31) Mesak, L. R.; Mesak, F. M.; Dah1, M. K. *Bacillus subtilis* GlcK Activity Requires Cysteine within a Motif That Discriminates Microbial Glucokinases into Two Lineages. *BMC Microbiol.* 2004, 4, 6.

(32) Wolf, Y.; Madej, T.; Babenko, V.; Shoemaker, B.; Panchenko, A. R. Long-Term Trends in Evolution of Indels in Protein Sequences. *BMC Evol. Biol.* 2007, 7, 19.

(33) Zhang, Z.; Huang, J.; Wang, Z.; Wang, L.; Gao, P. Impact of Indels on the Flanking Regions in Structural Domains. *Mol. Biol. Evol.* 2011, 28, 291–301.

(34) Han, J.; Tscherneuttet, V.; Yang, J.; Eckle, T.; Borchers, C. H. Analysis of Selected Sugars and Sugar Phosphates in Mouse Heart Tissue by Reductive Amination and Liquid Chromatography-Electrospray Ionization Mass Spectrometry. *Anal. Chem.* 2013, 85, 5965–5973.

(35) Miyazono, K.-i.; Tabei, N.; Morita, S.; Ohnishi, Y.; Horinouchi, S.; Tanokura, M. Substrate Recognition Mechanism and Substrate-Dependent Conformational Changes of an ROK Family Glucokinase from *Streptomyces griseus*. *J. Bacteriol.* 2012, 194, 607–616.

(36) Imrikov, I.; Arreguin-Espinosa, R.; Guzmán, S.; Rodríguez-Sanoja, R.; Langley, E.; Sanchez, S. Biochemical Characterization of the Glucose Kinase from *Streptomyces coelicolor* Compared to *Streptomyces peucetius* Var. Caesius. *Res. Microbiol.* 2005, 156, 361–366.

(37) Hayward, S.; Berendsen, H. J. C. Systematic Analysis of Domain Motions in Proteins From Conformational Change: New Results on Citrate Synthase and T4 Lysozyme. *Proteins: Struct., Funct., Bioinf.* 1998, 30, 144–154.

(38) Alamgro-Moreno, S.; Boyd, E. F. Insights into the Evolution of Sialic Acid Catabolism among Bacteria. *BMC Evol. Biol.* 2009, 9, 118.

(39) Mira, A.; Pushker, R.; Legault, B. A.; Moreira, D.; Rodriguez-Valera, F. Evolutionary Relationships of Fusobacterium nucleatum Based on Phylogenetic Analysis and Comparative Genomics. *BMC Evol. Biol.* 2004, 4, 50.

(40) Bairy, S.; Gopalan, L. N.; Setty, T. G.; Srinivasachari, S.; Manjunath, L.; Kumar, J. P.; Gunjupalli, S. R.; Bose, S.; Nayak, V.; Ghosh, S.; Sathyaranayan, N.; Caing-Carlsson, R.; Wahlgren, W. Y.; Friemann, R.; Ramaswamy, S.; Neerathilingam, M. Automated Optimization of Cloning, Expression and Purification of Enzymes of *Fusobacterium nucleatum* Valera, F. Evolutionary Relationships of *Fusobacterium nucleatum* To Generate Mutants for Protein Structure-Function Studies. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2002, 58, 1948–1954.

R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX : A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr., Sect. D: Struct. Biol.* 2010, 66, 213–221.

(47) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Iserger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Saschettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX : Building New Software for Automated Crystallographic Structure Determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2002, 58, 1948–1954.