PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

A QM/MM MD Study of the pH-Dependent Ring-Opening Catalysis and Lid Motif Flexibility in Glucosamine 6-Phosphate Deaminase†

Yuan Zhao[‡a], Nanhao Chen[‡b], Ruibo Wu[*b] and Zexing Cao[*a]

[a]*State Key Laboratory of Physical Chemistry of Solid Surfaces and Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, P.R. China.*

E-mail: zxcao@xmu.edu.cn

[b]*School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, P.R. China*

E-mail: wurb3@mail.sysu.edu.cn

†Electronic Supplementary Information (ESI) available: The glucosamine 6-phosphate deaminases (NagBs) from *Escherichia coli*, *Homo sapiens*, *Bacillus subtilis*, and *Streptococcus mutans*. The RMSD values of all the backbone atoms as function of time for the four different models. Relative energy profiles along the defined reaction coordinates for different protonation models. The statistical Mulliken charges of main atoms in the active site of reactant, transition state, and intermediate for different protonation models**.** The relative energies for the ring-opening process in deprotonation model and in the first hydrogen-transfer step of protonation model by the QM/MM energy scanning for original systems and Asn128Ala & Glu135Ala mutant systems. See DOI: 10.1039/ c0xx00000x

‡Yuan Zhao and Nanhao Chen contributed equally to this work

Abstract

The glucosamine 6-phosphate deaminase (NagB), which catalyzes the conversion of D-glucosamine 6-phosphate (GlcN6P) into D-fructose 6-phosphate (F6P) and ammonia, determines the final metabolic fate of the N-acetylglucosamine (GlcNAc). Here using state-of-the-art *ab initio* QM/MM MD simulations, we have explored the plausible mechanisms for the enzymatic ring-opening of GlcN6P in the basic environment. Two different proton-shuttle mechanisms have been proposed. Calculations show that the protonated state of the amino group in the substrate dominates the concerted and stepwise catalytic pathways and a catalytic triad plays an important role in mediating the proton transfer and the resulting ring-opening process. The free energy barrier for the rate-determining step in the low-energy stepwise reaction is 17.9 kcal mol^{-1}. In acidic solution, the lid motif prefers a closed state while it always stays in the open state in basic solution upon the substrate binding, which are basically dominated by the protonated state of the residue His145.

Key words: Glucosamine 6-phosphate deaminase (NagB), QM/MM MD simulations, catalytic mechanism, proton shuttle, lid motif flexibility

1. Introduction

The glucosamine 6-phosphate deaminase (NagB), which belongs to a class of aldose-ketose isomerases, can catalyze the conversion of D-glucosamine 6-phosphate (GlcN6P) into D-fructose 6-phosphate (F6P) and ammonia. This isomerization reaction is the final specific step in the N-acetylglucosamine (GlcNAc) utilization pathway, which leads to the metabolic direction of GlcNAc[1] into glycolysis (see Fig. 1a). The NagB family has attracted considerable attention for a long period of time, and many experimental studies have been conducted to determine its protein structure, kinetic properties, flexibility of lid region, allosteric regulation, genomic organization, function analysis, catalytic mechanism, etc.[2-14]

In recent years, a number of hexameric NagB crystal structures have been characterized from *Escherichia coli*[3-4] and *Homo sapiens*,[5] and several monomeric NagBs have been determined from *Bacillus subtilis*[6] and *Streptococcus mutans*[7] (see Table S1 in Electronic Supporting Information (ESI) material†). However, only one enzyme-reactant complex (PDB ID: 2RI1[7]) has been obtained from *S. mutants* in acid solution among all of the available crystals. Moreover, as Fig. 1(b) shows, the "lid motif", which would rim the mouth of the binding pocket, presents a "open" state in acidic solution regardless of the ligand binding or without in the *Smu*NagB (NagB from *S. mutans*).[7] Nevertheless, they are "closed" in *Eco*NagB and *Bsu*NagB complexes.[4,6]

In regard to the enzymatic catalysis, a "three-step" reaction mechanism was proposed in previous studies,[7] including ring-opening, isomerization and deamination.

And the initial ring-opening is thought to be a stepwise process catalyzed by a triangle hydrogen network (Asn128-His130-Glu135), and the amino group of GlcN6P was assumed as a protonated state in Liu et al.'s studies,[7] as shown in Fig. 1. The Nε of His130 would abstract one proton from the O1 atom to form a zwitterion configuration, and the subsequent transfer of the proton to O5 finally achieves the ring opening of pyranose. Quite surprisingly, the GlcN6P deamination by *Smu*NagB was found to have remarkable pH dependence,[7] and the enzyme shows the highest catalytic activity in pH values from 8.0 to 8.5, but it is nearly inactive at the pH < 6.

Despite these important contributions, the following two crucial issues remain open: (i) How does the "lid motif" regulate the open-closed state of the active pocket? (ii) Since the pH-dependent catalytic activity is observed, what's the protonation state tailored for the amino group in GlcN6P? And is it related to the ring-opening mechanism?

In the present work, classic molecular dynamics (MD) simulations and the state-of-the-art quantum mechanics/molecular mechanics (QM/MM)[15-23] MD simulations have been employed to analyze the functional roles of the lid motif and explore the ring-opening mechanism, in order to deeply understand the catalytic properties of *Smu*NagB and clarify the essential roles of the conserved residues. Moreover, in consideration of plausible protonation states of the amino group in GlcN6P, distinct catalytic processes have been investigated comparatively here. It can be expected that the detailed interpretations for the catalytic mechanism and the dynamics of lid motif are helpful for the drug design for cure of the dental caries

caused by *S. mutans*.[24-29]

2. Computational Details

2.1 Molecular dynamics simulations

The apo state of SmuNagB and the SmuNagB-GlcN6P complex are constructed based on their X-ray crystal structures (PDB ID: 2RI0 and 2RI1).[7] In order to reduce computational costs, only one chain of the enzyme dimer is considered in calculation. Herein two kinds of computational models have been built for the apo state and the complex state, respectively, and they are named as Models A and B for the acidic solution or Models C and D for the basic environment. The protonation states of residues are determined at pH = 5.5 and pH =8.0 via the program PROPKA 3.1,[30-33] As to the ligand, the charge parameters of substrate are calculated at the HF/6-31G(d) level by the restrained electrostatic potential (RESP) methodology[34] with Gaussian 03 package,[35] and the force field parameters are acquired by the generalized Amber force field (GAFF).[36] The whole systems are solvated into a cubic water box of ~ 68×71×72 $Å^3$ and the necessary counterions are added to neutralize the model systems. The Amber99SB force field[37-39] and TIP3P water model[40] are employed by tleap[41] in Amber12.

For each of the four models, after the energy minimization, the whole systems are heated up gradually from 0 K to 290 K under the NVT ensemble for 50 ps. Then, another 50 ps MD simulations are carried out subsequently to relax the system density to about 1.0 g/cm^3. Finally, 17 ns NPT MD equilibration at 290K are carried out with

the integration time step of 2 fs. Herein, the pressure relaxation time of 5 ps is set to keep trajectories stable. Long-range electrostatic interactions are handled with the particle mesh Ewald (PME) method,[42] and the Langevin method is employed to maintain the temperature at 290K. During the MD simulations, the constant pressure periodic boundary conditions together with a cutoff distance of 12 Å for van der Waals and electrostatic interaction calculations were applied and all of the bonds with hydrogen atoms in the system will be constrained with the SHAKE algorithm.[43-44] The trajectories of last 3 ns have been used for the root mean square fluctuation (RMSF) estimation and the binding free energy analysis. All the molecular dynamic simulations are performed by AMBER12 software.[45]

2.2 QM/MM MD simulations

Since high catalytic activity (k_{cat} > 30 s^{-1}) has been achieved within the pH range from 7.5 to 9.5 experimentally,[7] the enzyme-substrate structures in alkaline solution after MD simulations were thus employed as the initial configuration in the following research. As depicted in Fig. 2, there are two QM/MM systems (deprontonation model and protonation model) constructed by removing all water molecules beyond 24 Å radiuses from the C3 atom of substrate according to the pKa measure of the protonation glucosamine in experiment.[46] The only difference between the two models is the protonation state of amino ($-NH_2$ and $-NH_3^+$) linked to the C2 atom in substrate. As Fig. 2 shows, the QM region includes Asn128, Glu135, His130 and GlcN6P, and they are treated by the B3LYP functional[47-50] with the 6-31G(d) basis set,

which have been successfully used for enzymatic systems.[16-23] The MM subsystem is treated by Amber99SB force field,[37-39] which is the same with previous MD simulations. The QM/MM boundary is described by the improved pseudobond approach.[51-53] The spherical boundary condition is adopted, and the atoms within 20 Å from the spherical center are free to move. The cutoff values of 18 and 12 Å are used for electrostatic and van der Waals interactions among MM atoms, respectively. For the QM/MM system, the optimization is performed with the iterative minimization procedure,[54] and then the reaction coordinate driving (RCD) method[54] is used to describe the minimum reaction energy path (MEP) of the ring-opening. For all of the electronic structures along the reaction path, the MM regions are equilibrated for 500 ps MD simulations with the QM subsystem frozen. Then, the snapshots are used as starting structures for *ab initio* QM/MM MD calculations with the umbrella sampling. For each window, a 20 ps simulation is carried out with 1 fs time step which employed by Beeman algorithm.[55] Meanwhile, Berendsen themostat[56] is used to control the system temperature around 290 K. The probability distribution along the defined reaction coordinate is obtained after 5ps for each window combined with the weighted histogram analysis method (WHAM)[57-60] to determine the free energy profiles of the two models. All the *ab initio* QM/MM calculations are performed with the modified Q-Chem 4.0[61] and Tinker 4.2[62] programs.

3. Results and Discussions

3.1 The flexibility of lid motif

For the each model mentioned above, 17 ns MD simulations were performed to equilibrate the whole system. According to the root-mean-square deviation (RMSD) results of all the backbone atoms (see Fig. S1 in ESI†), all trajectories get to equilibrium after about 12 ns MD simulations and the structures of last 3 ns were abstracted for the subsequent analysis.

As shown in Fig. 3, the secondary structures of two models in acidic solution (Models A and B) and basic solution (Models C and D) are aligned respectively. Obviously, the lid motif is at the "close state" when the enzyme binds with the substrate in acidic solution (Model B) while it is still at the "open state" in other three models. According to the root-mean-square fluctuation (RMSF) of four models, the lid motif (residues 152~168) becomes less flexible in Model B, and it is almost unchanged in comparison with Models C & D. We also estimated the volume around the active site by twelve random structures from last 3 ns MD simulations. In acidic solution, the volumes of the pocket for the apo state and the enzyme-substrate complex are 543.2 ± 56.2 and 416.2 ± 18.5 $Å^3$, respectively, suggesting that the pocket would become smaller when the substrate entrances the active site. In basic solution, the volumes of the pocket are 1125.3 ± 60.0 and 1258.0 ± 73.5 $Å^3$. All these results reveal that the lid motif has a tendency to move from "open" to "close" in acidic solution upon the substrate binding at the active site. However, we should note that the lid region of crystalline complex determined by Liu et al.[7] maintains an open state like the apo state of enzyme. Accordingly, we examined the structures along the MD simulations and RMSF values in different time periods (see Fig. 4). We found

that the lid motif approaches to the substrate gradually and shows a tendency to close. The flexibility of the lid motif also decreases as the simulation proceeds. Clearly, here the predicted dynamical behaviors of the lid motif in *Smu*NagB are comparable with *Eco*NagB and *Bs*uNagB,[4,6] and they undergo similar conformation changes upon the substrate binding.

In alkaline environment, the flexibility of the lid motif always stays at the open state after the substrate bound to the active site, which is quite different with the situation obtained in acidic solution. This phenomenon stimulates us to carry out statistical analysis on the hydrogen bonds which are most likely to be formed between the active site and the lid motif by the last 1 ns in MD simulations.

As we can see in Fig. 5, in acidic solution, for the complex state, Asn157 and the substrate can form a hydrogen bond network through water molecules to make the lid motif close to the active site. Meanwhile, Thr150 & Glu135, and Asn134 & Hip145 trend to form the hydrogen bond with the possibility of 32% and 99.7%, respectively, for the enzyme-substrate complex during the sampling procedure, wherein the four residues may play an inductive role for the close of lid motif. However, the phenomenon cannot be seen for the apo-state enzyme, in which the lid motif stays in the open state. Thereby, the conformation change from the open state to the close state for the lid motif arises from the substrate binding in acidic solution.

In basic solution, the lid motif always stays at the open state regardless of the complex or the apo-state enzyme. We checked the region between the lid motif and the active site, and noted that this is no meaningful hydrogen bond responsible for the

close of lid region. This quite differs from the situation in acidic solution, and the protonation state of His145 may account for this phenomenon to some extent. As Fig. 5 shows, the doubly protonated His145 induces formation of a hydrogen bond chain linking to the lid motif through one water molecule, leading to a closed state in acidic solution. However, the hydrogen bond interactions disappear in basic solution, and the lid motif keeps an open state. Such distinct pH dependence of the lid motif could be ascribed to the polarization effect of different protonation states of His145 as well as the resulting hydrogen-bond network rearrangement in the active domain.

3.2 Catalytic mechanisms modulated by the protonated state of substrate

To figure out a rough picture of the enzymatic reaction, possible concerted and stepwise ring-opening processes have been investigated by QM/MM calculations, and the QM/MM-predicted relative energies along predefined reaction coordinates are shown in Fig. S2 (see ESI†). The minimal energy paths (MEP) of the catalytic ring-opening processes are obtained by QM/MM calculations for the deprotonated model (named as Model E) and the protonated model (named as Model F), where the reaction coordinates are defined by $d_{H1\text{-}O5} + d_{N\varepsilon\text{-}H1}$ (RC1) and $d_{H1\text{-}O1} - d_{N\varepsilon\text{-}H1}$ (RC2), respectively. The concerted ring-opening product state can be obtained along the reaction coordinate of RC1 for Model E (the deprotonated amino group of substrate). We note that one intermediate can be located along the reaction coordinate of RC2 for Model F (the protonated amino group of substrate) and this model is responsible for a stepwise process. Therefore, RC1 is more appropriate as the reaction coordinate for

Model E, and RC2 can be served for the first step of Model F. Besides that, for the second step of reaction for Model F, d_{C1-O5} - d_{H1-O5} (RC3) is chosen as the reaction coordinate. To determine more reliable catalytic mechanism, *ab initio* QM/MM MD simulations along with umbrella sampling are performed to acquire the relative free energy profiles for the enzymatic reaction.

The potential of mean forces (PMFs) along the defined reaction coordinates of two models are displayed in Fig. 6, and all of the corresponding important states involved in reaction are collected in Figs. 7 and 8. We find that the difference among different time periods is very small, showing that the convergence for the PMF profiles is reliable. As to the model E, the PMF curve clearly shows that this ring-opening process should be concerted and only one transition state has been located. The calculated free energy barrier is around 21.0 kcal mol^{-1}, and the energy of ring-opening state (**IM**) is about 20.0 kcal mol^{-1} above the initial reactant state, just about 1.0 kcal mol^{-1} lower than the corresponding **TS**, suggesting that the intermediate configuration is less stable.

As Fig. 7 shows, the optimized structures of the reactant state (**R**), the transition state (**TS**), and the intermediate (**IM**) reveal the presence of a catalytic triad through a hydrogen bond network among the residues Asn128, Glu135, and His130. The catalytic triad makes the deprotonated Nε of His130 nicely positioning toward the substrate and forms a hydrogen bond. Presumably, the polarization and hydrogen-bond cooperation effects of the catalytic triad can facilitate the proton transfer in the enzymatic reaction.

The hydrogen transfer from O1 to Nε of His130 initiates the reaction, and the consecutive hydrogen transfer from Nε to O5 induces the ring-opening of pyranose, leading formation of the intermediate (**IM**) as shown in Fig. 7. Fig. 6a indicates that the H1 transfer from O1 to Nε to form the protonated His130 is quite slow, and the late-transition state (**TS**) bears a remarkable character of doubly protonated His130 (Hip). At the transition state, the hydrogen bond network among the catalytic triad strengthens and stabilizes this active Hip-like transition state. The consecutive H1 transfer to O5 is quite facile, which induces the ring-opening of pyranose. Simultaneously, the C1–O1 bond is shortened from 1.39 ± 0.03 to 1.29 ± 0.03 Å, showing a tendency to form a carbonyl. The statistical analysis of Mulliken charges in reaction (see Table 1a for key atoms) suggests that the overall hydrogen transfer-mediated ring-opening process follows a concerted proton-shuttle mechanism.

QM/MM MD simulations reveal that during the H1 transfer from Nε to O5 of substrate the Glu135 residue moves far away from His130 and then the ring is entirely opened. At the intermediate state (**IM**), the normal carbonyl of C1=O1 and hydroxyl of O5–H1 are formed with the bond lengths of 1.22 ± 0.01 and 1.00 ± 0.02 Å, respectively. We note that the configuration twisting of the ring-opened D-glucosamine 6-phosphate through the dihedral angle ((C2-C3-C4-C5)) should be facile at the **IM** state, which may be the reason for the appearance of a flat-plateau intermediate state.

Alternatively, the ring-opening process may follow a stepwise mechanism as

shown by Model **F** in Fig. 6b with **TS1** and **TS2**. The first proton transfer from O1 of the substrate to Nε of His130, yielding a zwitterionic intermediate complex (**IM1**), is quite facile with a free energy barrier of 2.1 kcal mol^{-1}. As Fig. 8 shows, the presence of $–NH_3^+$ in the substrate can stabilize the deprotonated $(O1)^-$. The intermediate configuration (**IM1**) proceeds to the ring-opened intermediate (**IM2**) through **TS2** with a free energy barrier of 18.0 kcal mol^{-1} relative the initial reactant state, which is the rate-determining step. The intermediate configuration **IM2** is higher in energy than the initial configuration by about 12.5 kcal mol^{-1}. The statistical analysis of the Mulliken charges in Table 1b and Fig. S3 (see ESI†) supports that the proton serves as the chemical entity for the hydrogen transfers in the stepwise process.

In **IM2**, the distances of O5–H1 and Nε···H1 are 1.01 ± 0.03 and 1.76 ± 0.20 Å, respectively, suggesting strong hydrogen bond interactions between His130 and the ring-opened substrate. We note that **IM2** in Model F is lower in energy than its corresponding **TS2** by 5.5 kcal mol^{-1}, showing relatively higher stability than **IM** in Model E, and furthermore no the flat plateau region for the intermediate configuration appears on the relative free energy profiles. Such differences may be ascribed into distinct twisted configurations of the dihedral angle (C2-C3-C4-C5) and the different polarization and hydrogen-bond interactions in both models E and F. As Fig. 9 shows, the dihedral angles in Model E fluctuate around 160° during the sampling at the **IM** state, while in Model F, the dihedral angle oscillates around 60°, suggesting that the ring-opened substrate **IM2** in Model F is less flexible and it is stabilized by relatively strong interactions.

As Fig. 6 shows, both ring-opening processes in Models E and F are endothermic by QM/MM MD simulations. Here we approximately estimated the relative energies between the ring-opened intermediate and the final product (fructose 6-phosphate and ammonium ion), and the primary QM calculations indicate that the final step is notable exothermic by 13.8 kcal mol^{-1}, which can drive the initial ring-opening step.

From QM/MM MD simulations, we have identified different ring-opening mechanisms for Models E and F. The only difference between both models is the protonated state of amino group in the substrate. For Model E with a neutral amino group, the ring-opening process of substrate follows a concerted mechanism, whereas for Model F with the protonated amino group, the reaction proceeds through a stepwise mechanism. The stepwise process in Model F is more favorable than the concerted process in Model E, showing a remarkable effect of the protonated state of substrate on the enzymatic activity.

The statistical analysis of the distance between the amino group and the O1 atom in the substrate indicate that the $–NH_2$ group in Model E is gradually far away from the O1 atom during the ring-opening process, suggesting that the neutral amino group has no obvious effect on the catalytic process, which could not stabilize the deprotonated O1 atom and thus no the zwitterionic intermediate state survives, leading to a concerted mechanism. However, for Model F, upon the deprotonation of O1 by Nε, the negatively charged O1 can be stabilized by the proximal $–NH_3^+$ group of substrate to achieve the zwitterionic intermediate state, which accounts for the stepwise mechanism. Such remarkable effects of the protonation state of the key His

residue on the catalytic mechanism and the reaction energetics may be applied to the other enzymatic systems.

3.3 Roles of the catalytic triad

It has been assumed that the residue network through hydrogen bond interactions plays a significant role in catalytic process.[63-70] For the catalytic ring-opening by *Smu*NagB, Liu et al. suggested that Asn128, His130 and Glu135 form a catalytic triad to take part in the catalytic reaction, where the neutral His130 mediates the proton transfer directly during the ring-opening process. Asn128 localizes the His130 positioning toward to the substrate through the hydrogen bond interactions between the carbonyl and Nδ. Glu135 can stabilize the Hip-like state whether Model E or F, which may facilitate the proton abstraction of the substrate by His130. The distance between the COO^- group of Glu135 and Nδ of His130 is significantly shortened upon the proton transfer from O1 to Nε.

To evaluate the effect of residues Asn128 and Glu135 on the reaction, QM(B3LYP)/MM calculations has been performed on the double mutation and single mutation for Asn128Ala and Glu135Ala, respectively. By means of the QM/MM optimization, the geometries and charge populations of the reactants for the mutant models E and F along with the intermediate for the mutant model F have been determined. Based on the QM/MM energy scanning, the approximately relative reaction energies for one step of the model E mutant system and two steps of the model F mutant system have been estimated. In previous experimental studies, the

mutation of either of both residues would completely abolish the enzyme activity. Herein we try to figure out which model is consistent with the experimental results better. Moreover, the single mutation of either of both residues on Models E and F has been considered to have an insight into their respective contribution to the catalytic activity.

As Fig. S4 shows in ESI†, the barrier for the double mutant system is higher in energy than that for the corresponding wild system by 10.2 kcal mol^{-1} based on Model E. For the model F mutant, the barriers for the proton transfer and C5–O1 bond cleavage increase by 4.3 and 39.9 kcal mol^{-1}, respectively. Such loss of activity might be ascribed to changes of the hydrogen bonding and polarization interactions in the active domain arising from the mutation. For example, the distances of Nε···H1 in the mutant systems are generally larger than those in the original system, and the relative basicity of O1 increases notably, which makes the initial hydrogen abstraction from O1 become difficult. We note that the mutation of any residues of Asn128 and Glu135 will shut down the enzymatic activity (see Fig. S4(c) in ESI) for Model F, as observed in experiment, which means that Model F should be responsible for the ready state of enzyme in the catalytic process.

Among the single mutations, the barrier of Asn128Ala mutant system is higher than the wild system by 7.4 kcal mol^{-1} for Model E, while the barrier for the Glu135Ala mutant system is close to that for the WT system. With regard to the model F, the barriers for the proton transfer in the Asn128Ala and Glu135Ala systems increase by 6.1 and 3.0 kcal mol^{-1} relative to the original system, respectively. The

barriers for the C5–O1 bond cleavage increase by 50.2 and 46.3 kcal mol^{-1}. Both results indicate that Asn128 plays more important role in the catalytic process, compared to Glu135. Population analyses show that the charge differences between C1 and O5 atoms are 0.36 a.u. and 0.59 a.u. for Asn128Ala and Glu135Ala mutant systems based on Model E, respectively. For Model F, the charge differences are 1.19 a.u. and 1.41 a.u. Accordingly, the Glu135Ala mutation may slightly enhance the polarity of the C1–O5 bond, which accounts for relatively lower barrier for the bond cleavage to some extent, compared to the Asn128Ala mutation. Beyond that, the resulting changes of the protein environment and the hydrogen-bond network around the active site from the mutation are also important for roles of the catalytic triad.

4. Conclusions

The dynamical behavior of the lid motif during the substrate binding by *Smu*NagB has been investigated by MM MD simulations, and the lid motif has been predicted to show totally different flexibility in acidic and basic solutions as the substrate enters into the active domain. In acidic solution, the lid motif would move from the open state to the close state with the substrate binding, which may be driven by the hydrogen bond interactions of Hip145···Asn134 and Thr150···Glu135, and Asn157···water···substrate. On the contrary, in basic solution, the lid motif always stays in the open state, which can be ascribed into the deprotonated state of His145. This key residue with different protonated states can mediate the formation of hydrogen bond network and modify the dynamics of the lid motif during the substrate

binding.

Extensive QM/MM MD simulations have been performed to explore the catalytic ring-opening of GlcN6P by *Smu*NagB, and two kinds of the proton-shuttle mechanisms dominated by the protonated state of the amino group in the substrate have been proposed, based on the present calculations. The neutral amino group of substrate will lead to a concerted ring-opening process while the protonated amino will give rise to a stepwise reaction. The stepwise process was predicted to more favorable than the concerted reaction dynamically. The presence and catalytic role of the residue triad have been characterized and validated theoretically, based on computational studies of the mutant systems. The present results provide a basis to understand the catalytic mechanism and activity of the NagB family.

Acknowledgments

This work was supported by the Ministry of Science and Technology (2011CB808504 and 2012CB214900,) and the National Science Foundation of China (21133007, 21203257, 21272289, and 21373164). We thank the National Supercomputing Center in Shenzhen and Guangzhou for providing the computational resource. We also thank Dr. Shenglong Wang at NYU-ITS for his kind help.

References

1 L. Warren, *Biosynthesis and Metabolism of Amino Sugars and Amino Sugar-containing Heterosaccharide: Glycoproteins*, (Gottschalk, A., ed)

Elsevier, Amsterdam, 1972.

2 R. Lara-Lemus, C. A. Libreros-Minotta, M. M. Altamirano and M. L. Calcagno, *Arch. Biochem. Biophys*., 1992, **297**, 213-220.

3 G. Oliva, M. R. Fontes, R. C. Garratt, M. M. Altamirano, M. L. Calcagno and E. Horjales, *Structure*, 1995, **3**, 1323-1332.

4. E. Rudino-Pinera, S. Morales-Arrieta, S. P. Rojas-Trejo and E. Horjales, *Acta Crystallogr. D, Biol. Crystallogr.*, 2002, **58**, 10-20.

5 R. Arreola, B. Valderrama, M. L. Morante and E. Horjales, *FEBS lett.*, 2003, **551**, 63-70.

6 F. Vincent, G. J. Davies and J. A. Brannigan, *J. Biol. Chem.*, 2005, **280**, 19649-19655.

7 C. Liu, D. Li, Y. H. Liang, L. F. Li and X. D. Su, *J. Mol. Biol.*, 2008, **379**, 73-81.

8 L. I. Álvarez-Añorve, M. L. Calcagno and J. Plumbridge, *J. Bacteriol.*, 2005, **187**, 2974-2982.

9 L. I. Álvarez-Añorve, I. Bustos-Jaimes, M. L. Calcagno and J. Plumbridge, *J. Bacteriol.*, 2009, **191**, 6401-6407.

10 I. Bustos-Jaimes, A. Sosa-Peinado, E. Rudino-Pinera, E. Horjales and M. L. Calcagno, *J. Mol. Biol.*, 2002, **319**, 183-189.

11 A. Uhde, J. W. Youn, T. Maeda, L. Clermont, C. Matano, R. Kramer, V. F. Wendisch, G. M. Seibold and K. Marin, *Appl. Microbiol. Biotechnol.*, 2013, **97**, 1679-1687.

12 M. Kawada-Matsuo, Y. Mazda, Y. Oogai, M. Kajiya, T. Kawai, S. Yamada, S. Miyawaki, T. Oho and H. Komatsuzawa, *PloS One*, 2012, **7**, e33382.

13 V. Shevchenko, M. Hogben, R. Ekong, J. Parrington and F. A. Lai, *Gene*, 1998, **216**, 31-38.

14 L. Liu, Y. Liu, H. -D. Shin, R. Chen, J. Li, G. Du and J. Chen, *Appl. Microbiol. Biotechnol.*, 2013, **97**, 6149-6158.

15 D. W. Rooklin, M. Lu and Y. Zhang, *J. Am. Chem. Soc.*, 2012, **134**, 15595-15603.

16 Z. Ke, G. K. Smith, Y. Zhang and H. Guo, *J. Am. Chem. Soc.*, 2011, **133**, 11103-11105.

17 R. Wu, Z. Lu, Z. Cao and Y. Zhang, *J. Am. Chem. Soc.*, 2011, **133**, 6110-6113.

18 R. Wu, S. Wang, N. Zhou, Z. Cao and Y. Zhang, *J. Am. Chem. Soc.*, 2010, **132**, 9471-9479.

19 R. Wu, H. Xie, Z. Cao and Y. Mo, *J. Am. Chem. Soc.*, 2008, **130**, 7022-7031.

20 Y. Zhou and Y. Zhang, *Chem. Comm.*, 2011, **47**, 1577-1579.

21 Y. Shi, Y. Zhou, S. Wang and Y. Zhang, *J. Phys. Chem. Lett.*, 2013, **4**, 491-495.

22 R. Wu, W. Gong, T. Liu, Y. Zhang and Z. Cao, *J. Phys. Chem. B*, 2012, **116**, 1984-1991.

23 Z. Ke, H. Guo, D. Xie, S. Wang and Y. Zhang, *J. Phys. Chem. B*, 2011, **115**, 3725-3733.

24 W. J. Loesche, *Microbiol. Rev.*, 1986, **50**, 353-380.

25 D. J. Lynch, S. M. Michalek, M. Zhu, D. Drake, F. Qian and J. A. Banas, *Oral*

Health Dent. Manag., 2013, **12**, 191-199.

26 J. A. Lemos, R. G. Quivey, Jr., H. Koo and J. Abranches, *Microbiology*, 2013, **159**, 436-445.

27 K. Cheon, S. A. Moser, H. W. Wiener, J. Whiddon, S. S. Momeni, J. D. Ruby, G. R. Cutter and N. K. Childers, *Eur. J. Oral Sci.*, 2013, **121**, 148-155.

28 C. Apel, A. Barg, A. Rheinberg, G. Conrads and I. Wagner-Dobler, *Dent. Mater.*, 2013, **29**, 1188-1199.

29 L. F. Espinosa-Cristóbal, G. A. Martínez-Castañón, E. J. Téllez-Déctor, N. Niño-Martínez, N. V. Zavala-Alonso and J. P. Loyola-Rodríguez, *Mater. Sci. Eng. C*, 2013, **33**, 2197-2202.

30 H. Li, A. D. Robertson and J. H. Jensen, *Proteins*, 2005, **61**, 704-721.

31 D. C. Bas, D. M. Rogers and J. H. Jensen, *Proteins*, 2008, **73**, 765-783.

32 M. H. M. Olsson, C. R. Sondergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525-537.

33 C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 2284-2295.

34 D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel and P. Kollman, *Comput. Phys. Commun.*, 1995, **91**, 1-41.

35 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G.

Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision A.1; Gaussian Inc.: Wallingford CT*, 2009.

36 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.

37 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179-5197.

38 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712-725.

39 J. M. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049-1074.

40 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.

41 D. A. Case, T. E. Cheatham, 3rd, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668-1688.

42 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577-8593.

43 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comp. Phys.*, 1977, **23**, 327-341.

44 S. Miyamoto and P. A. Kollman, *J. Comput. Phys.*, 1992, **13**, 952-962.

45 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 12*, *University of California: San Francisco, CA,* 2012.

46 L. Dong, Suying Ma, Limin Yang and Xiangde Sun, *Lishizhen medicine and material medica research*, 2008, **19**, 836-837

47 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.

48 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.

49 S. H. Vosko, L. Wilk and M. Nusair, *J. Phys.*, 1980, **58**, 1200-1211.

50 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*,

1994, **98**, 11623-11627.

51 Y. Zhang, *J. Chem. Phys.*, 2005, **12**, 024114.

52 Y. Zhang, *Theor. Chem. Acc.*, 2006, **116**, 43-50.

53 Y. Zhang, T. S. Lee and W. Yang, *J. Chem. Phys.*, 1999, **110**, 46-54.

54 Y. Zhang, H. Liu and W. Yang, *J. Chem. Phys.*, 2000, **112**, 3483-3492.

55 D. Beeman, *J. Comput. Phys.*, 1976, **20**, 131-139.

56 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.

57 S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman and J. M. Rosenberg, *J. Comput. Chem.*, 1992, **13**, 1011-1021.

58 M. Souaille and B. Roux, *Comput. Phys. Commun.*, 2001, **135**, 40-57.

59 A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, 1988, **61**, 2635-2638.

60 A. Grossfield, WHAM: the Weighted Histogram Analysis Method, Version 2.0.6, http://membrane.urmc.rochester.edu/content/wham.

61 Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, Jr., R. C. Lochan, T. Wang, G. J. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C. P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P.

A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, 3rd, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, 3rd, J. Kong, A. I. Krylov, P. M. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3172-3191.

62 J. W. Ponder, *TINKER*, *Software Tools for Molecular Design, Version 4.2*, *Washington University: Saint Louis, MO*, 2004.

63 J. C. Sinclair, J. Sandy, R. Delgoda, E. Sim and M. E. Noble, *Nat. Struct. Biol.*, 2000, **7**, 560-564.

64 P. Carter and J. A. Wells, *Nature*, 1988, **332**, 564-568.

65 E. Zeiler, A. List, F. Alte, M. Gersch, R. Wachtel, M. Poreba, M. Drag, M. Groll and S. A. Sieber, *PNAS*, 2013, **110**, 11302-11307.

66 J. J. Tesmer, T. J. Klem, M. L. Deras, V. J. Davisson and J. L. Smith, *Nat. Struct. Biol.*, 1996, **3**, 74-86.

67 E. L. Ash, J. L. Sudmeier, E. C. De Fabo and W. W. Bachovchin, *Science*, 1997, **278**, 1128-1132.

68 M. Arand, H. Wagner and F. Oesch, *J. Biol. Chem.*, 1996, **271**, 4223-4229.

69 L. Brady, A. M. Brzozowski, Z. S. Derewenda, E. Dodson, G. Dodson, S. Tolley, J. P. Turkenburg, L. Christiansen, B. Huge-Jensen, L. Norskov and L. M. Thim, *Nature*, 1990, **343**, 767-770.

70 A. C. Ruzzini, S. Ghosh, G. P. Horsman, L. J. Foster, J. T. Bolin and L. D. Eltis, *J. Am. Chem. Soc.*, 2012, **134**, 4615-4624.

Table 1 Statistical Mulliken charge populations of selected atoms in the active sites of Model E **(a)** and Model F **(b)** at their different reactive states.

(a)

	O1	Nε	H1	O5	C1	NH_2
R	-0.68±0.05	-0.32±0.10	0.40±0.07	-0.55±0.07	0.27±0.15	-0.22±0.03
TS	-0.74±0.06	-0.16±0.14	0.40±0.06	-0.86±0.07	0.40±0.09	-0.26±0.03
IM	-0.44±0.04	-0.26±0.13	0.41±0.09	-0.75±0.08	0.30±0.09	-0.24±0.04

(b)

	O1	Nε	H1	O5	C1	NH_3^+
R	-0.49±0.05	-0.29±0.14	0.30±0.08	-0.47±0.07	0.19±0.17	0.63±0.02
TS1	-0.72±0.06	-0.16±0.16	0.37±0.11	-0.59±0.07	0.37±0.15	0.64±0.03
IM1	-0.70±0.05	-0.10±0.17	0.37±0.10	-0.61±0.07	0.36±0.15	0.64±0.03
TS2	-0.38±0.03	-0.25±0.16	0.46±0.09	-0.76±0.10	0.29±0.06	0.62±0.03
IM2	-0.38±0.03	-0.30±0.13	0.36±0.15	-0.73±0.06	0.24±0.07	0.57±0.03

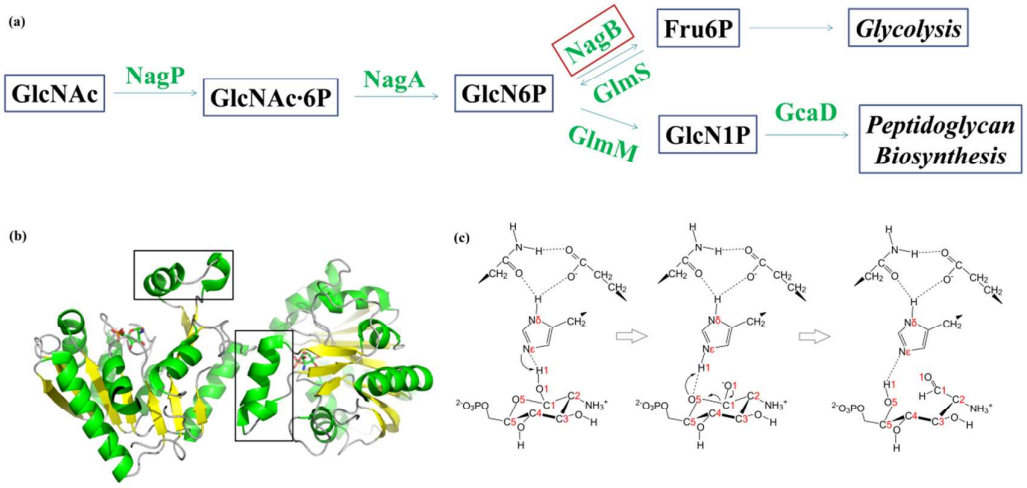


Fig. 1 (a) The pathway of GlcNAc metabolism. **(b)** The structures of SmuNagB-GlcN6P complex (PDB: 2RI1),[7] where the region in silver box is the lid motif. (**c)** The previously- suggested ring-opening catalytic mechanism.[7]

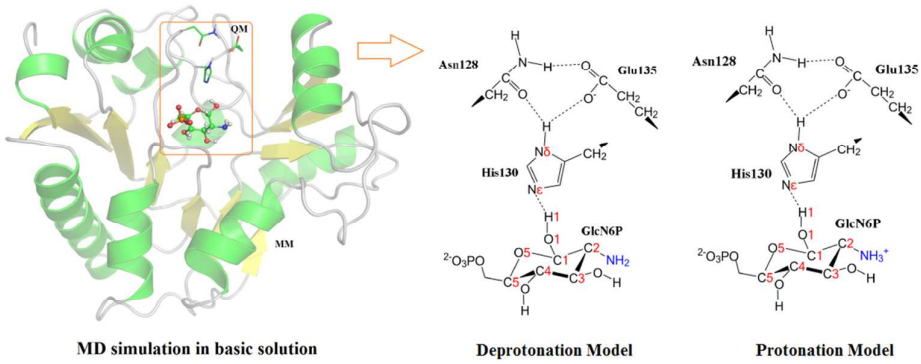


Fig. 2 The QM/MM models for ring-opening catalytic mechanism computations

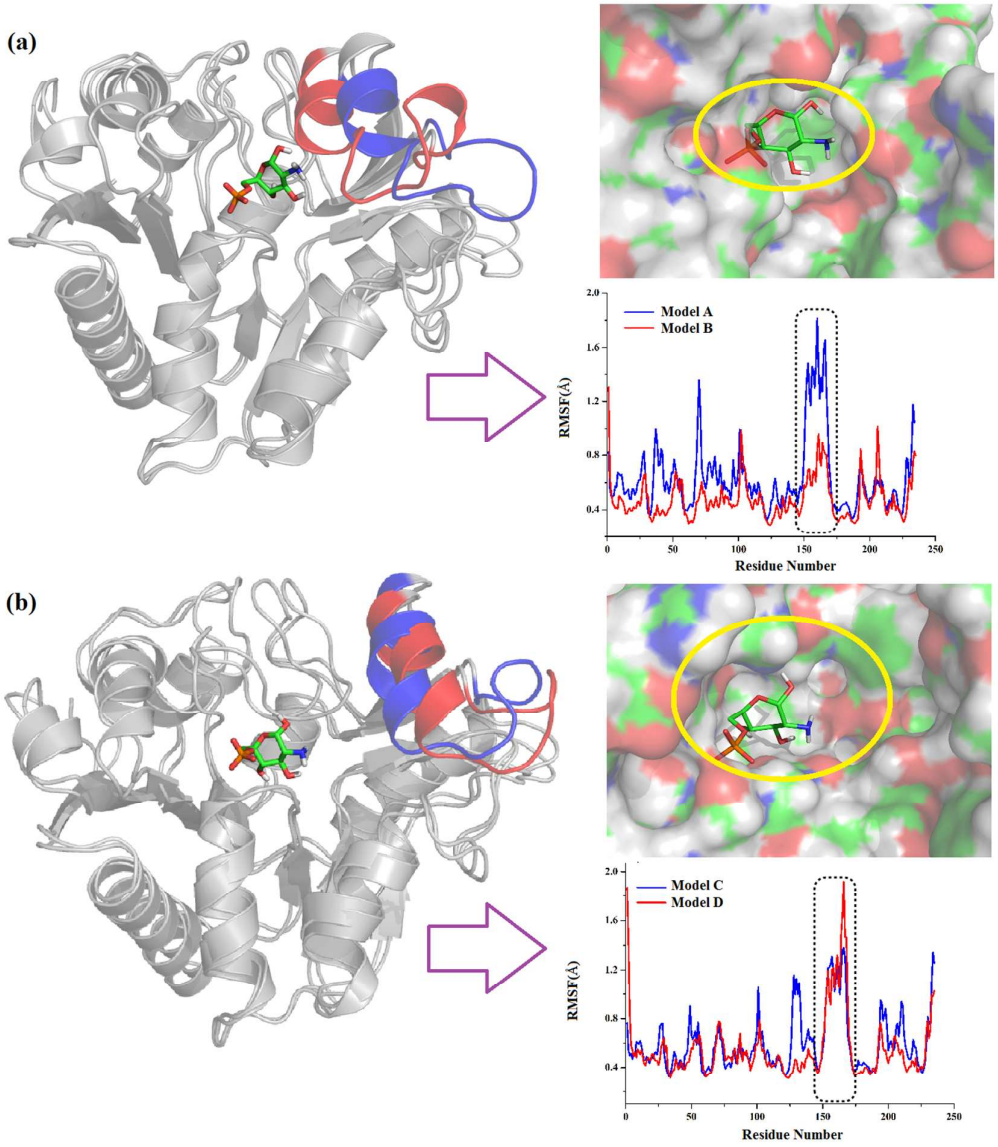


Fig. 3 The overlap for the apo state (Blue) and enzyme-substrate complex (Red) in acidic **(a)** and basic solutions **(b)** along with RMSF values of the backbone and surface of the active site pocket (The yellow circles refer to the pocket of the lid motif)

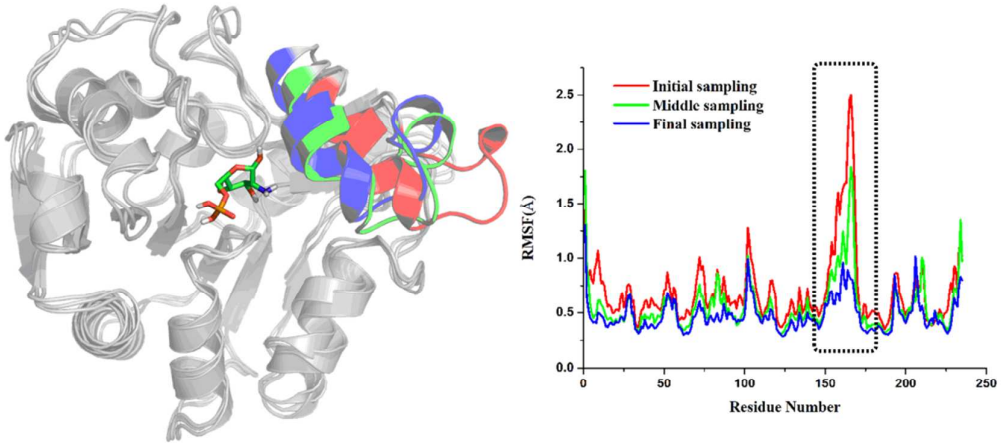


Fig. 4 The overlap of enzyme-substrate complex in acidic solution (Model B) in different time periods along with RMSF values of the backbone

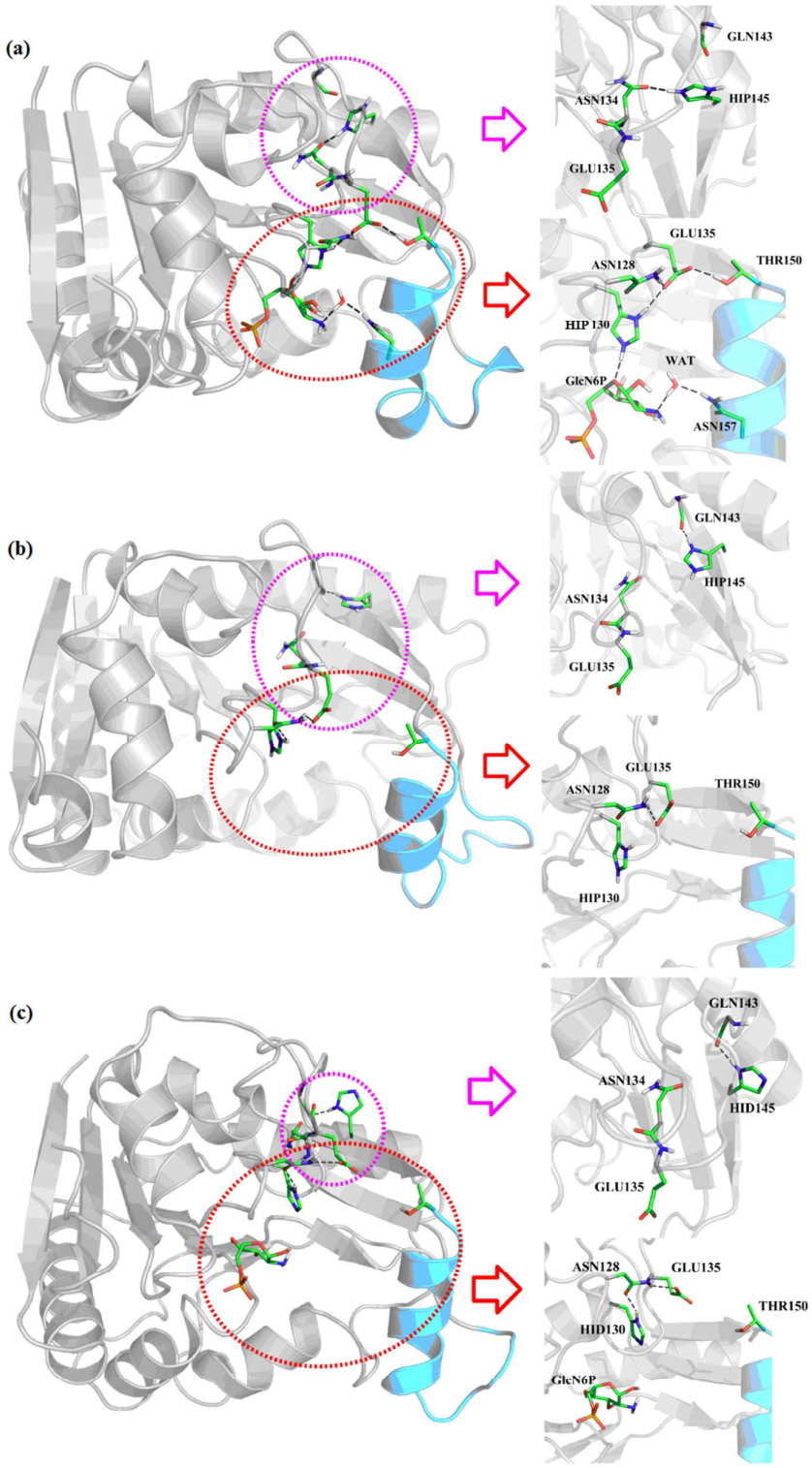


Fig. 5 The hydrogen bond networks around the active site and lid region of Model B **(a)**, Model A **(b)**, and Model D **(c)**. (HID means that the Nδ position is protonated)

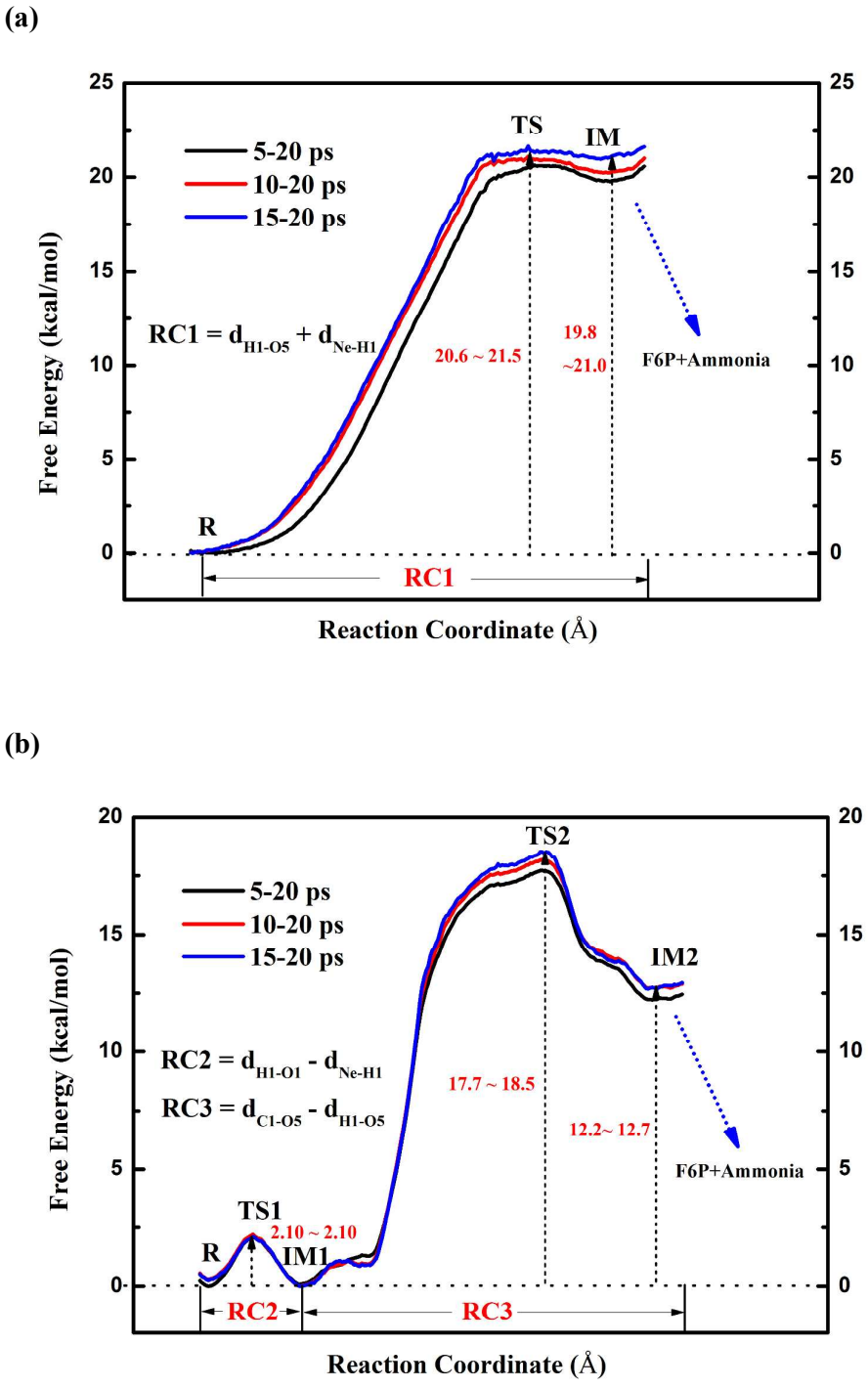


Fig. 6 Relative free energy profiles of the ring-opening reaction along the reaction coordinate determined by *ab initio* QM/MM MD simulations for Model E **(a)** and Model F **(b)**. (The further reaction to the final product will be followed by the ring-opening process, with the release of notable energy by the primary QM calculation)

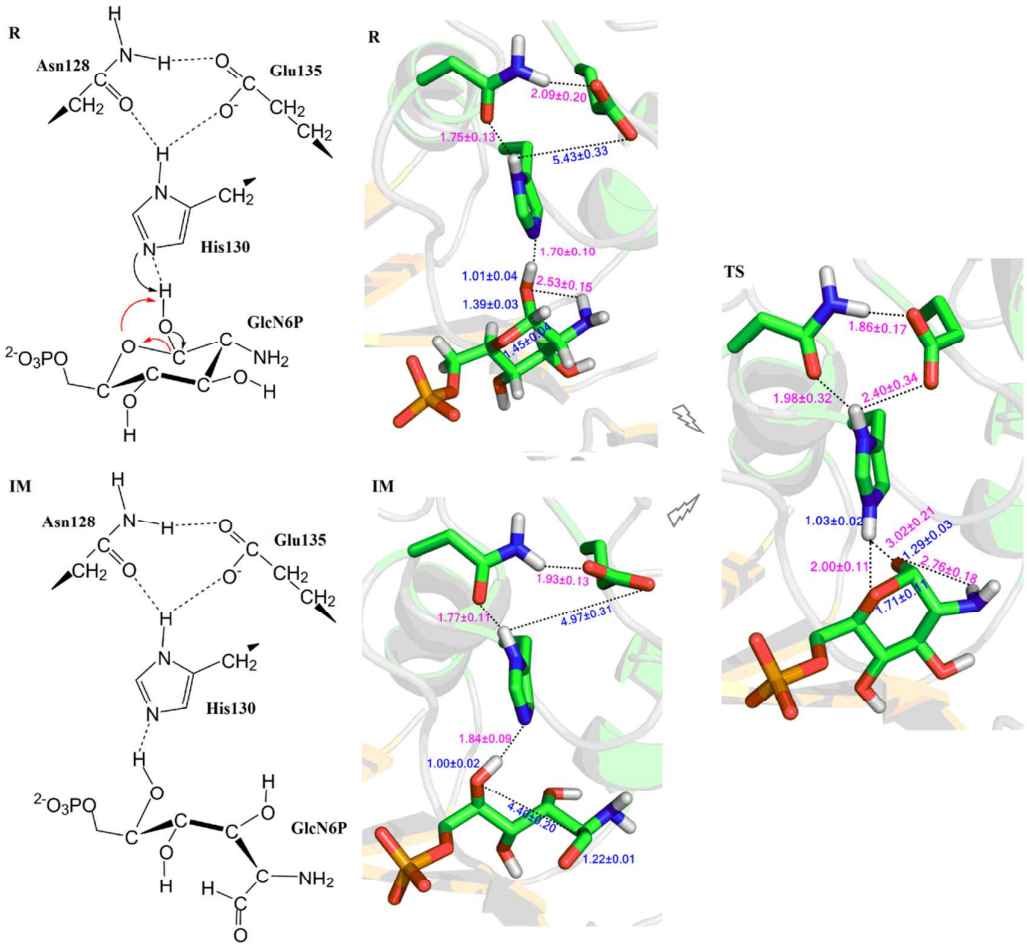


Fig. 7 The structures of reactant, transition-state, intermediate configurations for Model E along the reaction path, with average distances (Å) and standard deviations for the main atoms from the QM/MM-MD sampling.

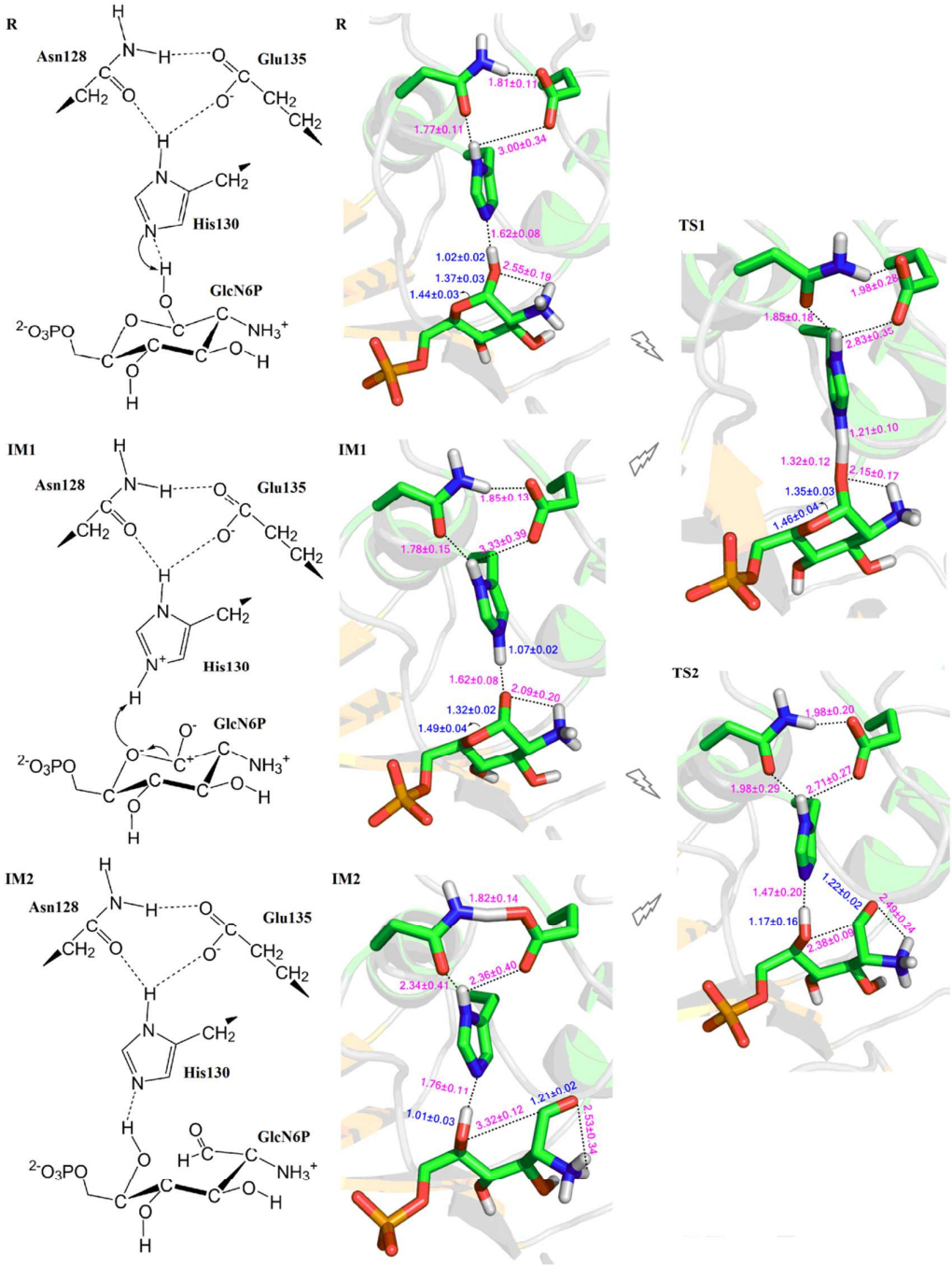


Fig. 8 The structures of reactant, transition-state, and intermediate configurations of Model F along the reaction path, with average distances (Å) and standard deviations for main atoms from the QM/MM-MD sampling.

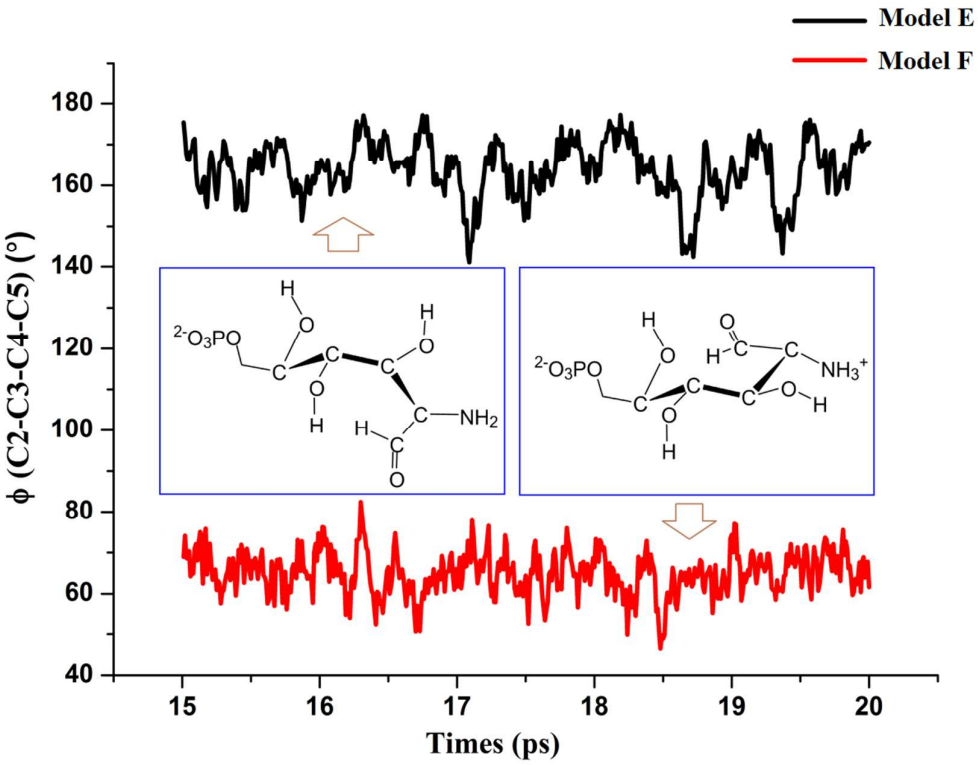
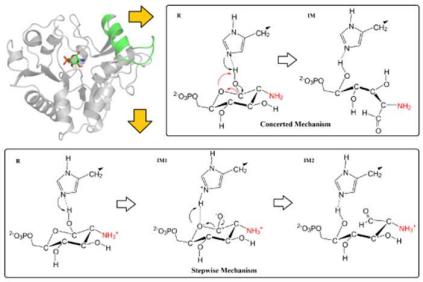


Fig. 9 The fluctuation of ϕ(C2-C3-C4-C5) in Models E and F from the last 5 ps QM/MM-MD sampling.

Graphical Abstract



QM/MM MD and MM MD simulations reveal pH-dependent proton-shuttle ring-opening mechanisms of GlcN6P and dynamical behavior of the lid motif in *Smu*NagB.