



Combination of multiple computational methods revealing specific sub-sectional recognition and hydrogen-bond dependent transportation of CKII peptide fragment in O-GlcNAc transferase



Xiao Zhang^a, Zhiyang Zhang^a, Jia Guo^b, Jing Ma^c, Songqiang Xie^c, Yuan Zhao^{a,*}, Chaojie Wang^{a,*}

^aThe Key Laboratory of Natural Medicine and Immuno-Engineering, Henan University, Kaifeng 475004, People's Republic of China

^bCollege of Chemistry and Chemical Engineering, Henan University, Kaifeng 475004, People's Republic of China

^cSchool of Pharmacy, Henan University, Kaifeng 475004, People's Republic of China

ARTICLE INFO

Article history:

Received 13 January 2021

Received in revised form 2 April 2021

Accepted 3 April 2021

Available online 8 April 2021

Keywords:

OGT

Peptide recognition

Conformation changes

Umbrella sampling

Hydrogen bond

ABSTRACT

O-linked β -N-acetyl-D-glucosamine (O-GlcNAc) transferase (OGT) is an essential enzyme in many cellular physiological catalytic reactions that regulates protein O-GlcNAcylation. Aberrant O-GlcNAcylation is related to insulin resistance, diabetic complications, cancer and neurodegenerative diseases. Understanding the peptide delivery in OGT is significant in comprehending enzymatic catalytic process, target-protein recognition and pathogenic mechanism. Herein extensive molecular dynamics (MD) simulations combined with various techniques are utilized to study the recognizing and binding mechanism of peptide fragment extracted from casein kinase II by OGT from atomic level. The residues of His496, His558, Thr633, Lys634, and Pro897 are demonstrated to play a dominant role in the peptide stabilization via hydrogen bonds and σ - π interaction, whose van der Waals and non-polar solvent effects provide the main driving force. In addition, two channels are identified. The delivery mode, mechanism together with thermodynamic and dynamic characterizations for the most favorable channel are determined. The peptide is more inclined to be recognized by OGT through the cavity comprised of residues 799–812, 893–899, and 865–871, and Tyr13-terminal is prior recognized to Met26-terminal. The transportation process is accompanied with conformation changes between the “spread” and “V” shapes. The whole process is strong exothermic that is highly dependent on the variation of hydrogen bond interactions between peptide and OGT as well as the performance of different subsections of peptide. Besides that, multiple computational methods combinations may contribute meaningfully to calculation of similar bio-systems with long and flexible substrate.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Many new biological functions are glycosylated to advanced glycation end products through the activation of sugar in organisms as a result of the catalysis of glycosyltransferases. This connects to different receptor molecules, such as proteins, nucleic acids, oligosaccharides, lipids and small molecules [1]. Sugar attaches to proteins or lipids through the act of enzymes, which is the process known as glycosylation. Glycosylation starts from the endoplasmic reticulum and ends at the Golgi apparatus [2]. O-GlcNAc glycosylation is a post-translational modification, which has been involved in many cell functions, such as regulation of gene expression [3–5], circadian

rhythm [6,7], vesicle transport, protein localization [8,9] and signal transduction [10–12]. Numerous diseases, including diabetes [13], cancer [14], neurodegenerative diseases [15], and cardiovascular disorders [16] (see Fig. 1) arise due to the mutations of O-linked glycosylation on a variety of protein substrates. Previous studies show that a wide variety of proteins can be O-GlcNAc glycosylation, such as RNA polymerase II, heat shock proteins, cytoskeletal proteins, etc [17–19].

The O-GlcNAc transferase (OGT) is the gene encoding glycosyltransferase in the organism, which catalyzes the connection of N-acetylglucosamine (GlcNAc) with serine or threonine residues of target protein through an O-linked β -glycosidic bond [20]. The gene is located on the X chromosome, which is essential for the survival and development of embryonic stem cells [21]. OGT includes two different regions as illustrated in Fig. 2: a series of superhelical tetratricopeptide repeat (TPR) [22] units that locates

* Corresponding authors.

E-mail addresses: zhaoyuan@henu.edu.cn (Y. Zhao), wcsjxq@henu.edu.cn (C. Wang).

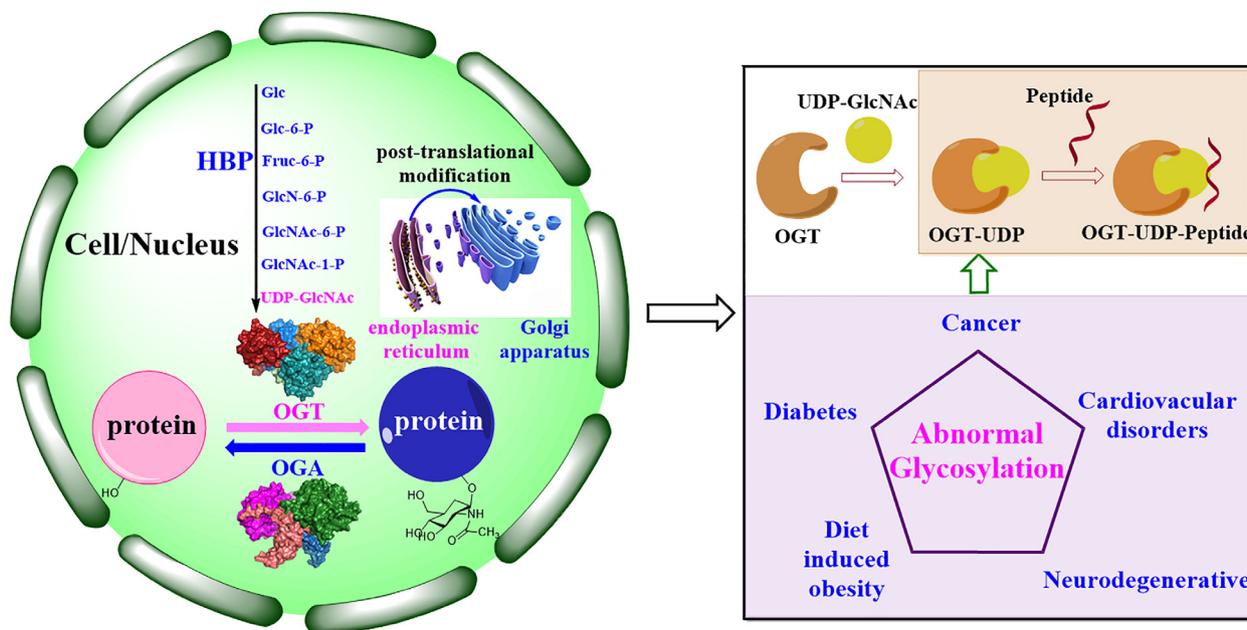


Fig. 1. The relationship of O-GlcNAcylation and human diseases, the role of peptide recognition of OGT as well as the overview of OGT and OGA.

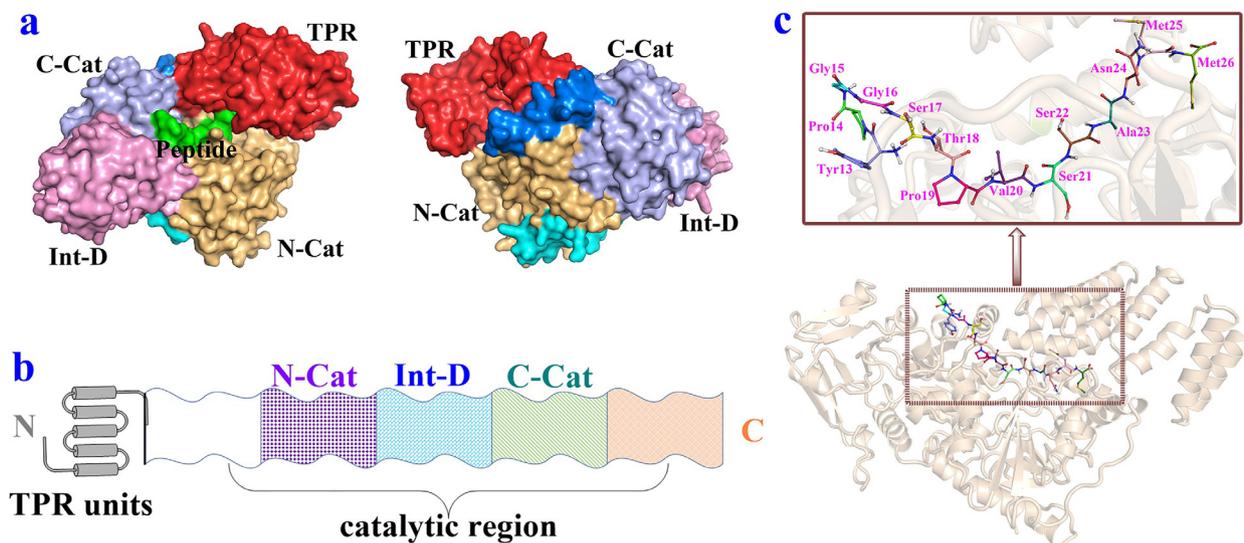


Fig. 2. (a) Different views of the crystal structures of OGT (PDB: 3TAX). (gold: N-Cat, purple: C-Cat, pink: Int-D, red: TPR units green: Peptide). (b) OGT composition. (c) The peptide from casein kinase II. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in N-terminal of OGT and the multi-domain catalytic region that lies in C-terminal. The former is mainly responsible for substrate recognition, and the repeat number is dependent on the species and type of alternating splicing isoforms. The latter is crucial in enzymatic catalysis, which is comprised of three parts: the N-terminal domain (N-cat), the intervening domain (Int-D), and the C-terminal domain (C-Cat) [23].

The protein O-GlcNAcylation was first discovered by using bovine galactosyltransferase to transfer the [3H] galactose moiety of UDP-[3H] galactose onto glycoproteins which contains terminal GlcNAc residues. This process was conducted by Hart and Torres thirty five years ago [24]. More recently, Lazarus and his co-workers reported the co-crystallized X-ray structure of truncated human OGT, which contains 4.5 N-terminal TPR sequences, complete C-terminal catalytic domain, UDP and CKII polypeptides

[25] (see Fig. 2). Based on the crystal structure, they found that OGT adopts an ordered bi-bi mechanism, which is prior in combination with sugar substrate to polypeptide [25].

The binding mode of OGT-UDP binary complex and the delivery mechanism of UDP were previously studied in our group by applying molecular dynamic simulations. The research helps to understand the enzymatic process, and also improves the design of OGT inhibitor on a basis of UDP and UDP-GlcNAc [26]. Here the main focus is on the transportation of peptide, which is a vital step in the entire nucleotide sugar metabolism process. A fragment of casein kinase II (CKII) is chosen as the peptide in this study, and this peptide includes three important serines, which can be recognized by OGT. Selecting the peptide mainly due to the observed complete crystal structure of OGT-UDP-peptide complex (PDB ID: 3TAX) in previous experiment [27], whose bi-substrates are stabi-

lized in the active site of OGT. In addition, the complex of OGT-CKII peptide helps to focus on the interaction between OGT and adaptor protein such as p38 [28].

As mentioned above, several key issues need to be resolved: (i) what are the main factors that contribute the key residues of OGT in peptide recognition and binding; (ii) in the possible transport channels for peptide, which channel is the most competitive one? (iii) how about the understanding and resolution of thermodynamics and dynamic properties in the process; (iv) what is the most favourable mode for the transportation of peptide in OGT? (v) what are the main factors that effect the transportation of peptide? Is there any noticeable characteristic during the whole process? Therefore, extensive molecular dynamic simulations with various technologies are carried out to study the delivery mechanism of CKII-peptide from atomic view to clarify these issues. This research would help the deep understanding of the enzymatic catalysis for OGT and recognition mechanism of target protein by, especially for CKII, OGT from an atomic perspective. The result will able to improve the design of the inhibitor which influences the interaction of OGT and target protein. Overall, this research will provide directions and ideas for further potential theoretical and experimental studies.

2. Computational methods

2.1. Preparation of the simulated system

On the basis of crystal structure of OGT-UDP-peptide (PDB ID: 3TAX) from the protein data bank, the initial model of OGT complex with sugar substrate and peptide is built by removing the repeat segment (see Fig. 2). The missing fragments of 715–718 and 747–761 residues are added using homologous modeling technique in SYBYL-X 2.1.1 program [29]. The protonation states of the charged residues are determined using the PROPKA 3.1 [30,31] package at pH 7.0 and the surrounding environment. The sugar substrate is treated with AMBER GAFF force field [32], and the OGT and peptide are dealt with AMEBR99SB [33] force field. The complex is solvated into an $86 \times 98 \times 127$ Å cubic TIP3P water box [34] with 10 Å buffer distance between the box edge and the nearest solute atoms, which is neutralized by adding Na^+ ions. The protons are automatically added to gain the initial coordinates and the topology parameters using the tleap module in AMBER 12 package [35].

2.2. Molecular dynamics simulations

First of all, the model is minimized for 10,000 steps by the steepest descent method, followed by another 10,000 minimization steps using the conjugate gradient method to correct poor interatomic contacts. Next, the system is heated gradually from 0 K to 300 K under the NPT ensemble for 100 ps. After that, an additional 100 ps MD simulation is performed to loosen the system density to about 1.0 g/cm^3 . Ultimately, a standard 50 ns MD simulation under the NVT ensemble is undertaken with an integration time step of 1 fs based on the periodic boundary condition [36]. 12 Å cutoff is set to calculate van der Waals interactions and electrostatic interactions. The SHAKE [37,38] algorithm is employed to constrain all of the hydrogen-containing bonds with a tolerance of 10^{-5} . The Langevin method [39] is utilized to keep the temperature at 300 K. The stability of backbone for OGT is analyzed by root mean square deviation (RMSD). The hydrogen bonds are counted using the equilibrium trajectories in the MM MD simulations. All the computational simulations are executed with the AMBER 12 package [35].

2.3. MM/GBSA calculations

Molecular mechanics/generalized born surface area (MM/GBSA) method [40] is successfully used to calculate binding free energy [41], which shows good performance in ranking binding affinity. Here the modified generalized Born model (GB_{OBC1}) [42] developed by Onufriev and his co-workers [43] is selected to calculate the polar solvation energy according to our previous test [44]. The free energy decomposition is performed on the basis of 50 snapshots extracted from the last 5 ns equilibrium trajectory by MM/GBSA method. The internal dielectric constant of the molecule and the dielectric constant of the solvent are set as 1.0 and 78.5, respectively. The maximum distance between the atomic pairs involved in the Born radius is set to be 25 Å. The surface tension used to calculate the non-polar contribution of the solvation free energy is $0.005 \text{ kcal/mol/Å}^2$. All calculations in this part are implemented using the AMBER 12 package [35].

2.4. RAMD MD simulations

Elucidation of ligand binding and release channel is crucial in drug discovery and application of target protein. RAMD simulation [45,46] is a widely used technique that deals with the problem well [47,48]. In RAMD MD simulations, a constant force in a random direction is applied to the center of mass of the ligand. When this ligand moves out of the threshold distance, this direction will be preserved; otherwise, another new random direction will be set to determine the release route of the ligand. The ligand automatically moves around the active site to find a feasible escape channel without determining the direction. This is the advantage of this method. If the random force is excessive, the ligand may move towards the wrong direction, and thus the classical MD simulation can be carried out to relax the system to avoid the error. In this work, on the basis of the equilibrium configuration, RAMD MD is performed, and the force field is the same with MM MD simulations. The random accelerations of 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, and 0.20 kcal/Å/g and a threshold of 0.2, 0.3, 0.4, and 0.5 Å are applied to peptide, respectively. 112 RAMD MD trajectories are acquired. The simulations are carried out by employing NAMD package [49].

2.5. SMD simulations

Since the peptide is highly flexible, more than one release mode may be observed. Hence, for the most advantage channel, the constant velocity pulling (PCV) method in steer molecular dynamics with different schemes are performed. This is mainly for detection of the favorable mode. The head, middle and tail of peptide are set as pulling part, respectively. The Amber99SB force field [33] is used for OGT and peptide, and AMBER GAFF force field [32] is applied for UDP. Chen and his co-workers [50] described the determination of appropriate parameters by judging the stability and quantity of information in the curve to ensure a moderate separation between proteins and ligands. Here one pulling scheme is selected for parameter testing. The spring constant k is selected as 4, 5, 6, 7, 8, 10, 12, 15, 18 kcal/mol/Å², while keeping velocity v fixed at 5×10^{-3} Å/fs to achieve the requirement of the stiff-spring approximation. Then the v is set as 1×10^{-3} , 2×10^{-3} , 2.5×10^{-3} , 3×10^{-3} , 4×10^{-3} , 5×10^{-3} , 6×10^{-3} Å/fs while keeping k at 5 kcal/mol/Å² (Fig. S4). By evaluating the pulling detail along the simulation time and trajectory animation, the most favorable k and v are chosen as 5 kcal/mol/Å² and 5×10^{-3} Å/fs, which will be used for the other two schemes in SMD simulations. For more reliable results, ten paralleled 60 ps SMD simulations are carried out for each release mode to accommodate the departure of the peptide. The calculations are performed with NAMD package [49].

2.6. Umbrella sampling

Umbrella sampling (US) method is first proposed by Torrie and Valleau [51,52] in the 1970 s. According to the optimal delivery mode of the most favorable channel determined by RAMD MD and SMD simulations, the MM MD simulation combined with umbrella sampling method is performed to obtain the detail mechanism, thermodynamic and kinetic properties as well as key residues for process of peptide delivery. On the basis of the delivery direction of peptide, the reaction coordinate is defined by selecting reference atoms with one moved to the “door”. For each window, 20 ns MM MD simulation with harmonic bias potential is performed to guarantee the sampling arrives at reasonable overlap. The weighted histogram analysis (WHAM) [53] is utilized to produce the free energy profile. All the calculations in this part are performed by using AMBER 12 package [35].

3. Results and discussion

3.1. OGT-peptide complex

Estimation of the root mean square deviation (RMSD) of the backbone for OGT, the system is detected to stabilize after 8 ns (see Fig. 3). The equilibrium snapshot is consistent with two crystal structures obtained by previous experiments [23,27] both in key residues around the peptide and protein conformation (see Fig. S1 and Fig. S2). It implies that the model is reliable to some extent by MM MD simulations. Moreover, ten snapshots extracted from the last 5 ns are almost overlapped for both protein and peptide as described in Fig. 3, which also means that the peptide is stable in the active site of OGT. It is located below the peptide chain, forming a stable OGT-peptide complex. Therefore, the binding mode in the active site is analyzed on the basis of statistic snapshot from equilibrium states. Since the peptide is long and flexible, it is divided into three parts as head, middle, and tail that refer to

Tyr13-Ser17, Thr18-Ser21, and Ser22-Met26, respectively (see Fig. S3) to facilitate analysis more understandable. The residues of His496, His558, Thr633, Lys634 and Pro897 interact with polar amino acids of peptide through six hydrogen bonds. At the same time, the conjugate ring of His496, His558, Pro897 and Tyr13, Pro19, and Ala23 of CKII peptide form the σ - π interaction. To further understand the specific role of each residue, the binding free energy of the OGT-peptide complex is analyzed by the MM/GBSA method. The binding free energy decomposition (see Fig. 4) shows that His496, His558, Pro559, Thr633, Lys634, Phe868, and Pro897 have a greater contribution to the binding of OGT-peptide. The total binding free energy is -17.74 kcal/mol, which mainly comes from the non-polar interaction (-101.11 kcal/mol) that comprised of van der Waals and non-polar solvation energy. Meanwhile, the polar interactions (83.36 kcal/mol) included electrostatic interaction and polar solvation energy play a negative role in peptide binding. In addition, both side chain and main chain of these key residues of OGT are pivotal for the peptide binding as presented in Table S2. The side chain of His496, Pro559, Lys634, and Phe868 accounts for a larger proportion of the interaction, while for Pro897, Lys634, and Thr633, the main chain has a stronger contribution.

In order to identify the role of specific residues, His496, His558, Thr633, Lys634 Phe868, and Pro897 are selected to take site-directed mutation to alanine, respectively. The mutant systems are named as His496Ala, Pro559Ala, Thr633Ala, Lys634Ala, Phe868Ala, and Pro897Ala. The alanine scanning in MM/GBSA method is used to calculate the binding free energy by 50 conformations that extracted from the last 5 ns MM MD trajectory. As demonstrated in Table 1, Table S1, and Fig. 4, the binding free energy for mutant system is decreased, among which the binding free energy of His496Ala decreases the most by 6.76 kcal/mol. With regard to Lys634Ala, Thr633Ala, and Pro897Ala, the extent of the decreased binding free energy is smaller than that of other three mutant systems. It is mainly because the interaction between

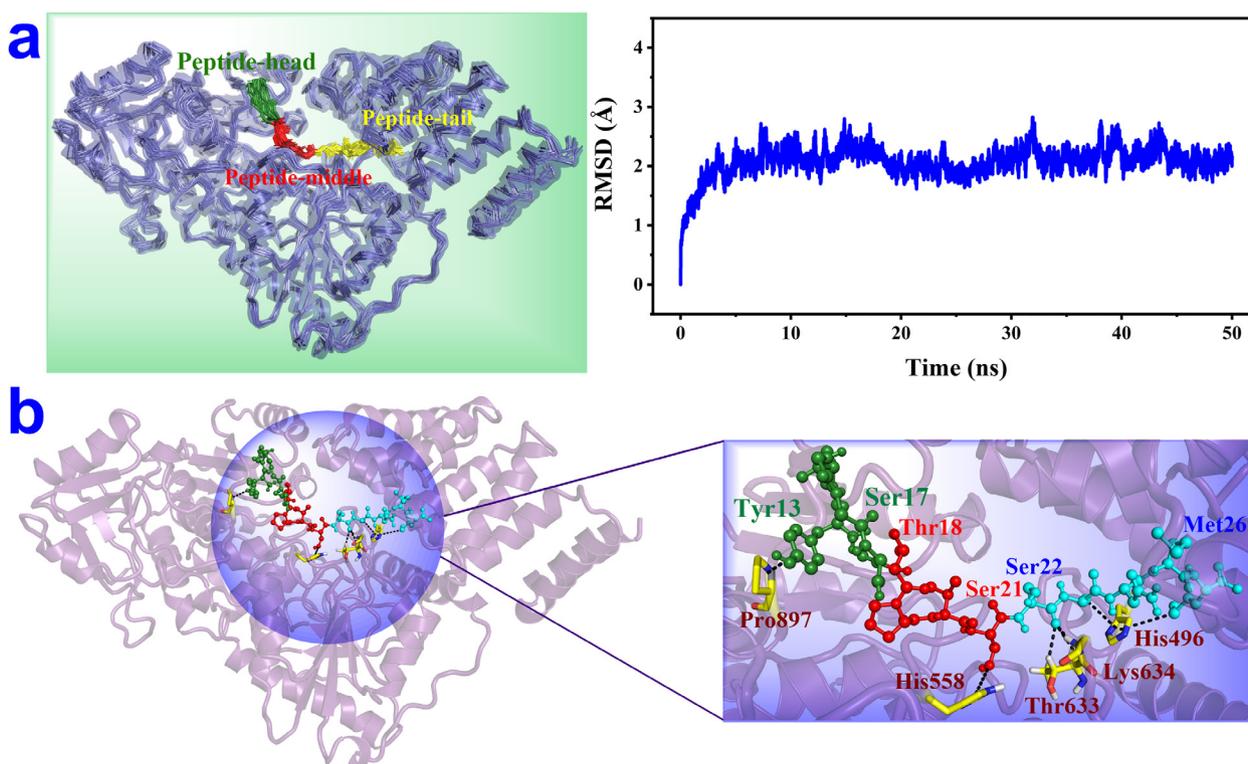


Fig. 3. (a) The last 5 ns overlap image of umbrella sampling and RMSD of backbone for OGT-peptide complex; (b) Binding mode of peptide in the active site of OGT.

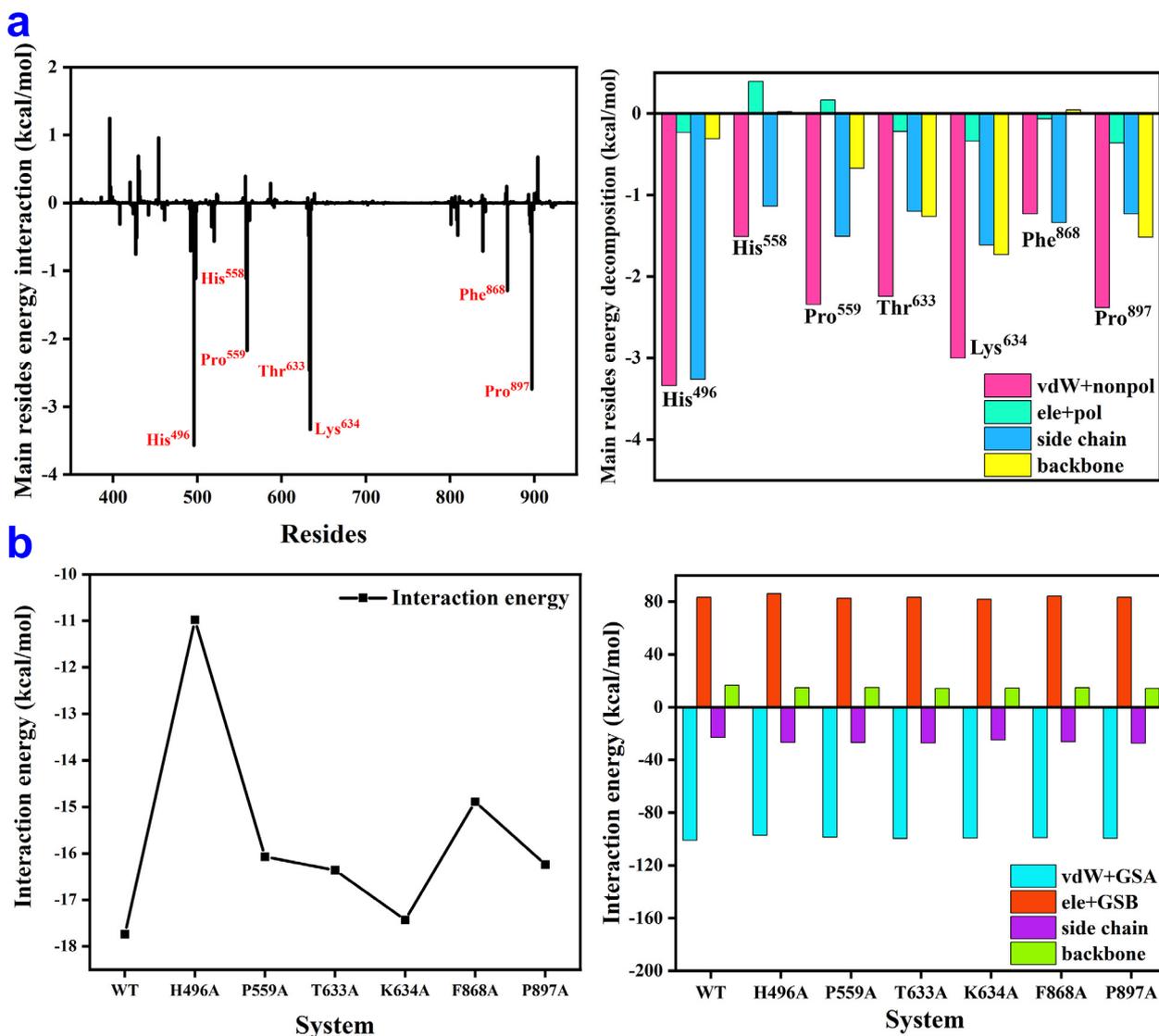


Fig. 4. (a) Free energy decomposition for the residues; (b) Interaction energy for the OGT and peptide complex in wild and mutant systems and binding free energy decomposition of the wild and six mutation systems.

Table 1

Binding free energy and corresponding standard deviation of the mutant systems (kcal/mol).

Mutated System	$\Delta G_{\text{wild}} - \Delta G_{\text{mutant}}$
H496A	-6.76 ± 1.94
P559A	-1.67 ± 0.80
T633A	-1.38 ± 0.59
K634A	-0.31 ± 0.57
F868A	-2.85 ± 0.43
P897A	-1.50 ± 0.39

main chain of these residues and peptide is dominant (see Table S3). The results are consistent with free energy decomposition and binding mode of wide system as above mentioned.

3.2. Explore the possible dissociated channel and mode of peptide

The dissociation of peptide plays a crucial role in the whole catalytic and recognized process of O-GlcNAc glycosylation. Thus, based on RAMD MD and SMD simulations, the possible channels and modes for peptide delivery are analyzed.

Here 112 trajectories for peptide dissociation are obtained by using RAMD MD simulations. Among them, nine trajectories show that peptide is not successfully released from OGT, accounting for 8.0% of all release trajectories. The remaining 103 trajectories are divided into two channels based on the release direction, which named as P1 and P2. The details of the amino acid fragment and the corresponding release probability are shown in Fig. 5 and Table S5. Located among Loop 1, 2 and 3 is the P1 channel, which accounts for a significant proportion of 82.1% in all release trajectories. It is the most statistical advantage channel, and we will conduct a detailed analysis for it.

To observe the possible release modes of the most dominant channel P1, PCV method in SMD simulations is carried out for departure of peptide from OGT. In the equilibrium state of OGT-peptide complex, the peptide is located in the exposed cavity formed by the fragments of Loop 1, 2 and 3. Due to the large cavity and the peptide's flexible feature, there may be a string of conformation changes such as turning, folding, and bending for peptide in the cavity. For this reason, the stretched atoms are selected from the main chain of residues located in head, middle and tail parts of the peptide, respectively. The most advantages release mode

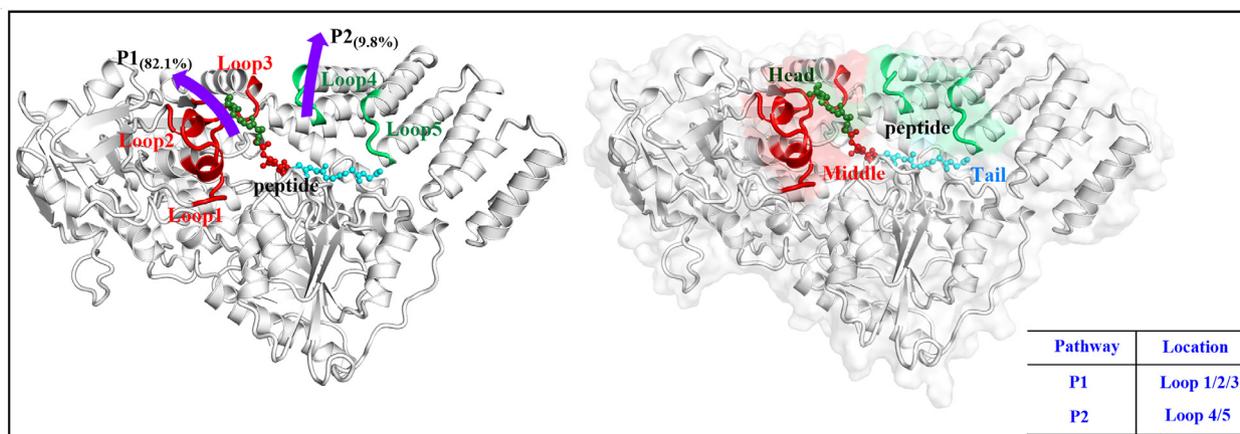


Fig. 5. Key segments in the possible channels for the release of peptide from OGT determined by RAMD MD simulations. (Loop 1: 799–812; Loop 2: 893–899; Loop 3: 865–871; Loop 4: 431–436; Loop 5: 398–402). The red shade shows the fragments of Loop 1, 2 and 3. The green shade denotes the fragments of Loop 4 and Loop 5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

will be obtained by comparing pulling force of the process. The detail stretched coordinates of three pulling modes are listed in Fig. S5 and Table S4.

As presented in Fig. 6, ten parallel trajectories are obtained for each mode. The corresponding maximum pulling forces are 1500–1800 pN, 1800–2000 pN, and 1800–2200 pN for Mode 1 to 3, respectively. The smaller the pulling force is, the easier the peptide releases from the cavity. Based on tension minimization, Mode 1 is the best choice for peptide release. To further identify the main factors influencing the peptide release, the binding modes of OGT and peptide at the location with the largest pulling forces are analyzed. In Mode 1, when peptide moves 11 ps, the pulling force reaches maximum, and the distance between the two reference atoms is 19.1 Å. The Asn805, Gln839, Lys634, Lys396, and Asn393 form five hydrogen bonds with peptide. Among them, Gln839, Lys396, Lys634, and Asn393 will hinder the departure of peptide, while Asn805 will facilitate the departure of peptide. As a result, the peptide needs energy to overcome the hydrogen bond interactions formed by preventing residues in its escaping. With regard to Mode 2, the strongest force mainly occurs at 23.2 Å between CA atom of His920 and CA19 atom of peptide after 25 ps of SMD simulations. Four strong hydrogen bonds form between peptide and residues of Lys634, Asn805, and Asp431, which prevent peptide fleeing away from the cavity. It is short of benefit driving force for peptide release, and thus the Mode 2 is inferior to the Mode 1 with longer time and bigger pulling force. As to Mode 3, the peptide experiences a conformation change with head prior to tail part. Therefore, it will take longer time to reach the maximum force, and this mode is not favorable for peptide release.

The SMD method provides a great convenience for selecting the best pulling mode on the basis of evaluating pulling force. With the above analysis, the maximum pulling force is about 1800 pN when pulling the tail, which is less than the pulling force required for the head and the middle parts, implying that this pulling mode is the most favorable one. Moreover, based on pulling trait, if pulling force is applied to the tail of the peptide, the hydrogen bonds at the head of the peptide gradually disappear as the tail approaches the exit, which relax for escaping from exit freely. However, if the force acts on the head part, there is an additional constrain applied for limitation of the mobile direction for peptide. If the force performs on the middle part, colossal conformation changes for the peptide will be occurred. Consequently, the latter two schemes are not optimal choice.

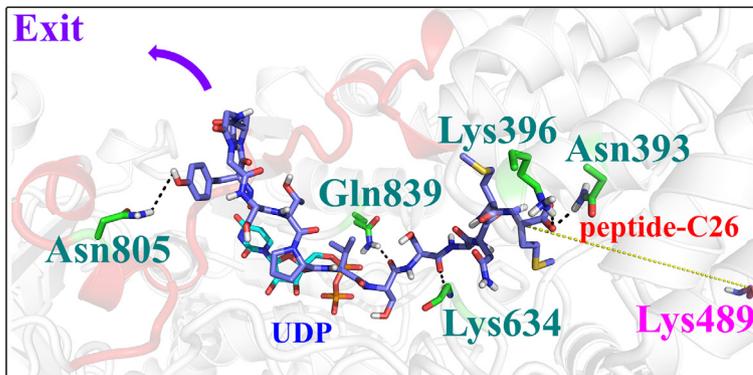
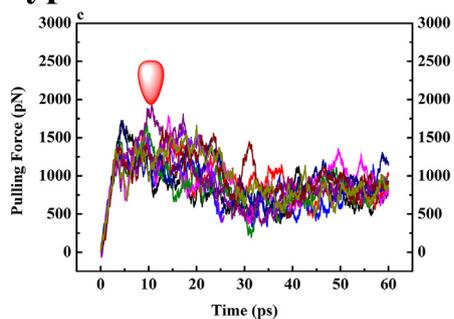
According to RAMD MD and SMD simulations, the most favorable channel and best pulling mode for the peptide dissociation are obtained. The process is preliminarily identified that tail part of the peptide flees away first along P1. Its inverse process is the peptide recognition and binding, which is of great concern to us. Since the external position of the peptide is uncertain, it is difficult to build the model for entering of peptide from outside to the active site. It should be noted that the process of substrate enters the active site has been discovered similar with the product release in the possible pathways and delivery mechanism [54]. Therefore, in order to further comprehend the mechanism, key residues together with corresponding thermodynamic and dynamic properties for the best channel and release mode in the recognition and binding process of CKII peptide, the classical MM MD simulations combined with umbrella sampling technology will be performed.

3.3. Thermodynamic and kinetic characteristics of the recognition mechanism of peptide.

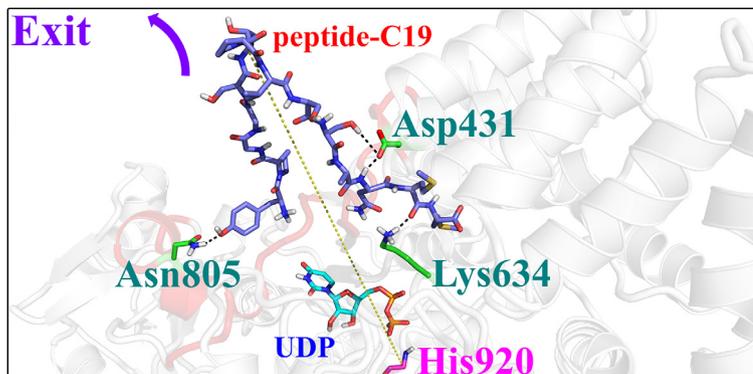
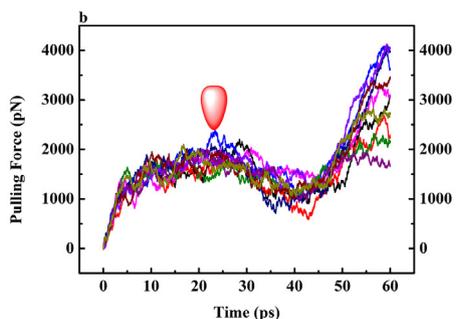
The distance between the CA atom of Lys498 and the C26 atom of peptide is selected as the reaction coordinate (RC) in umbrella sampling calculations (see Fig. 7). A total of 1340 ns MM MD simulations for 67 windows with appropriate biasing harmonic potentials from 17.0 Å to 50.0 Å with a 0.5 Å resolution are performed to map the free energy profile and explore the peptide delivery mechanism. For each window, the root mean square deviation (RMSD) is used to verify the stability of the trajectory, and a series of biasing harmonic potentials are tested to ensure full sampling and sufficient overlap between adjacent windows (see Fig. S6 and Fig. S7). Different time periods and districts are considered on the basis of previous studies [55,56] to prove the convergence of free energy curves (see Fig. S8). It can be seen that the free energy curves arrive at convergence after 15 ns, and thus the sampling during time durations of 16–18 ns, 18–20 ns, and 16–20 ns is adopted for getting reliable free energy files, whose maximal SD is 0.6 kcal/mol. According to the conformation changes and recognition characteristics of protein and peptide as well as the trend of free energy, the entire recognizing and binding process can be divided into five stages (Fig. 8).

In the first stage ($50.0 \text{ \AA} \geq RC > 42.5 \text{ \AA}$), the head and middle part of peptide is priority recognized by the residues around the entrance cavity of OGT. The two residues of Val895 and Asn805 are most importantly among these residues for binding head part of peptide through two hydrogen bonds, whose probability of

Hypothetical 1



Hypothetical 2



Hypothetical 3

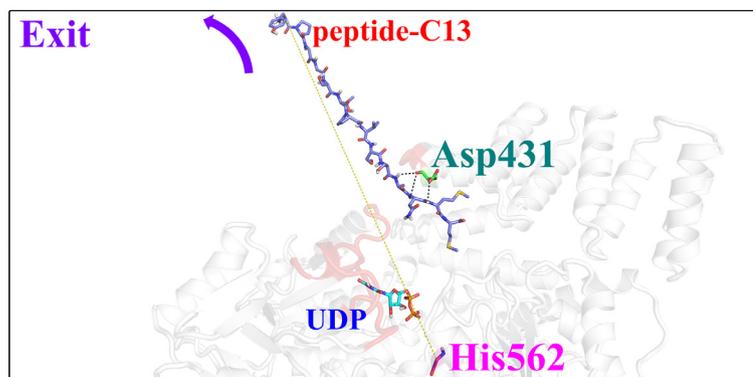
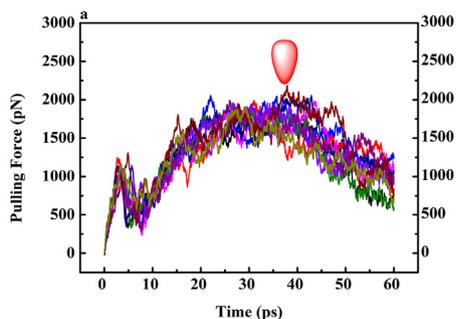


Fig. 6. Pulling forces and residues around peptide at maximum pulling force in SMD.

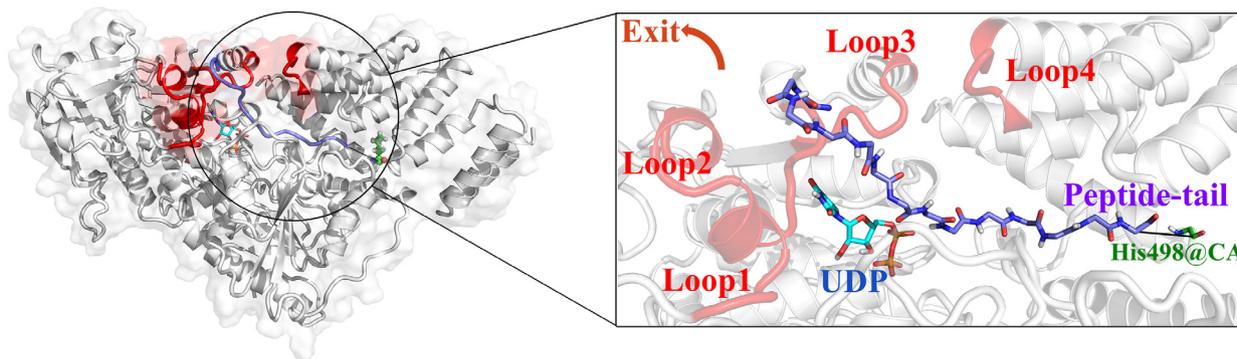


Fig. 7. Reaction coordinate definition in umbrella sampling for the most favorable pulling mode along P1.

forming the interactions approaches 90% with the head part of peptide during the stage. Furthermore, Asp431 and Gln839 are extremely essential by one or two hydrogen bonds with the middle

part of peptide alternately, which also has a high probability as the peptide moving. Besides that, Asn557 also assists the recognition of peptide, which mainly acts on the head part. However, as to

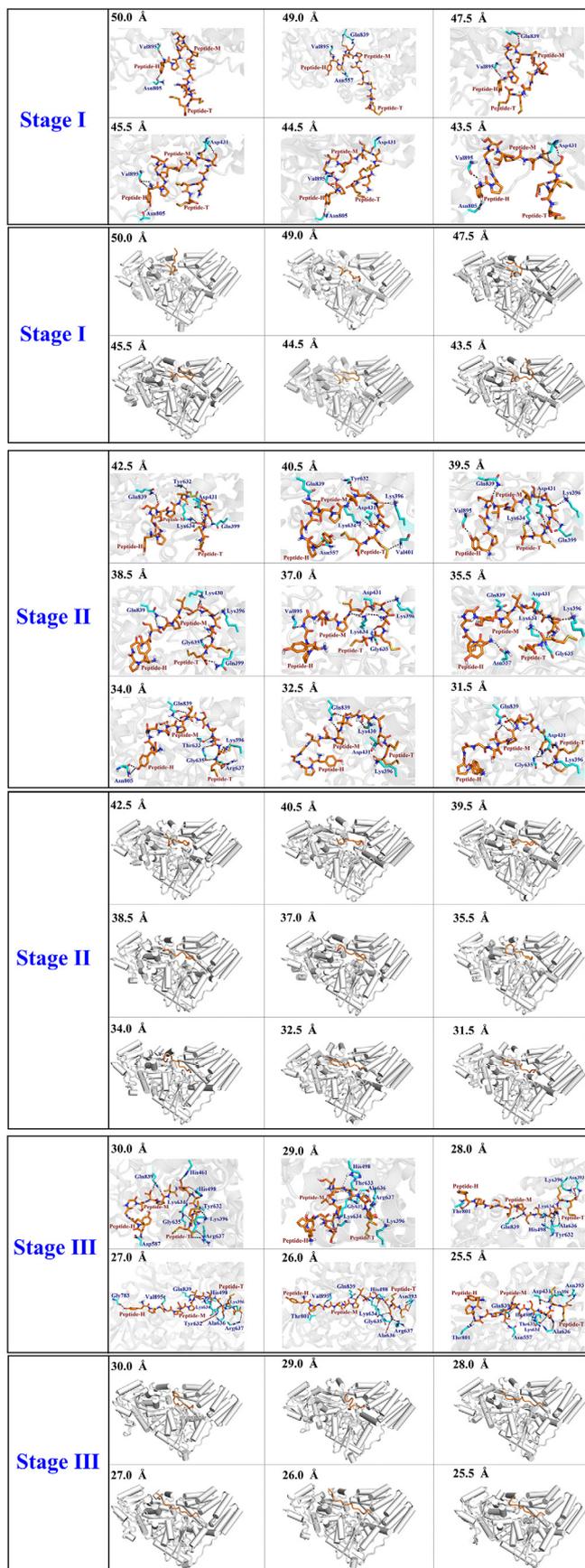


Fig. 8. The conformation change and location of the peptide recognize along P1 in MM MD simulations combined with umbrella sampling technology in five stages.

the tail part, it is freely floating around the interacted cavity. Here it should be noted that the conformation of the peptide presents a curly “V” shape during the whole stage.

In the second stage ($42.5 \text{ \AA} \geq RC > 30.0 \text{ \AA}$), a great deal of hydrogen bonds between OGT and peptide gradually increase to further bind the peptide, which are mainly located in the middle and tail part of the peptide. They provide a driving force for recognition of the peptide, especially for Lys396, exceeding 90% probability of occurrence. Moreover, other residues, such as Val401, Lys430, Tyr632, Lys634, Gln399, and Gly635 also contribute to the peptide delivery. The peptide continues to show a “V” shape, and most of the snapshots along the reaction coordinate present more visible crimp for the peptide than the first stage. It is probably caused by the recognition of tail part of peptide. First notably, Lys634 mostly forms hydrogen bonds with both middle and tail part by high probability simultaneously, which may be the main reason for the noticeable crimp for peptide. In addition, the hydrogen bond between peptide and residues of Asp431 is going from the middle to the tail gradually. Moreover, the Gly635 arises with high probability during the later period and the residues such as Gln399, Val401, and Arg637 appear around the tail part. All of them form hydrogen bonds with peptide. These factors influence the degree of crook for peptide. The Gln839 sostenuto forms hydrogen bonds with the head or middle part by stability probability in the snapshots that with deep bending for peptide, which also implies that the interaction came from the residue is important for the conformation changes of peptide.

In the third stage ($30.0 \text{ \AA} \geq RC > 25.0 \text{ \AA}$), the peptide is still wavering to the active site with an obvious folding, whose structure experiences “V → spread” conformation changes. It is worth noting that the residues around the head part of the peptide become unstable, particularly as Asp587, Gly783, Val895, and Thr801. When the latter three residues present around the head part, the peptide shows “spread” shape, otherwise, the “V” shape can be observed. On the contrary, if Asp587 interacts with head part of peptide, “V” shape can be identified, otherwise, the peptide shows “spread” shape. Therefore, it may be one of the reasons for the conformation perturbation of peptide, which guides mobile direction for the peptide. Moreover, more residues of OGT than those of the second stage appear to form hydrogen bonds with middle and tail part of peptide alternately with erratic probability, such as Gln839, His498, His461, Asp431, Lys396, Asn393, contributing to the obvious conformation of peptide. Among these residues, the new hydrogen bonds formed by His498, His461, and Asn393 are useful for the delivery of peptide towards the active site. In addition, the hydrogen bonds formed between the fragment of residues Tyr632–Arg637 and peptide can be detected, which is also conducive for the peptide moving.

In the fourth stage ($25.0 \text{ \AA} \geq RC > 20.0 \text{ \AA}$), due to the softness of peptide and specific interactions between key residues and peptide, the conformation of the peptide is back to “V” shape, then fluctuates to “spread” shape again, and then back again to relaxed “V” shape. The hydrogen bond environment is generally similar to that of the third stage, which is also alternately by comparing each window. The head part of the peptide keeps swinging up and down, while the middle and tail parts are relatively stable. That is why a folding or bending change for the peptide can be detected, which is similar with that of last stage. The obvious alternated changes of hydrogen bond interactions between residue of Tyr13 located in the head part of peptide and residues of Asp587, Asp554, Ser432, and Val895 of OGT are the main reason for wobble of head part of peptide. If the first three residues form interactions with head part of peptide, “V” shape can be detected, while if Val895 instead of them that appears in the head part, the peptide shows “spread” shape. The fragment comprised of Tyr632–Arg637 of OGT as well as the residues of His496, His461, and Lys396 alter-

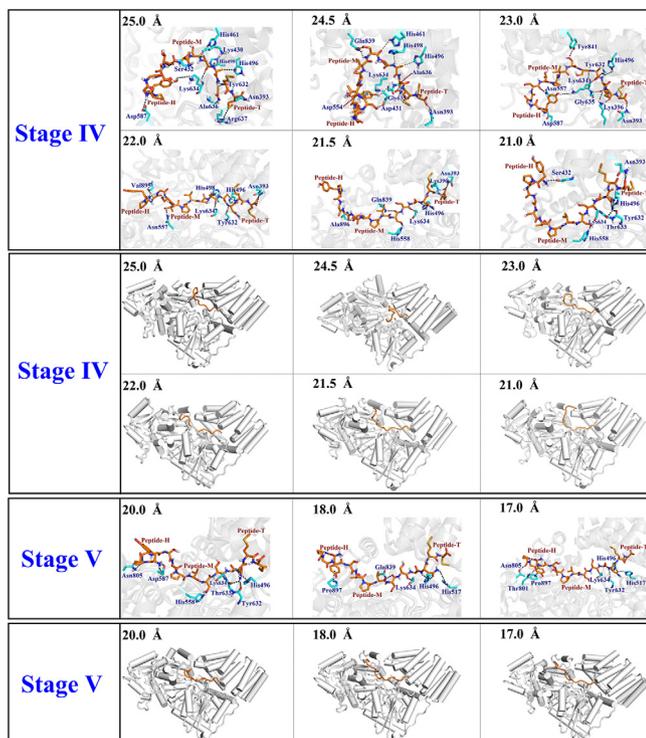


Fig. 8 (continued)

nately interact with the middle and tail part of peptide through hydrogen bonds. They can stabilize peptide in the cavity, which are benefit to peptide binding. It is worth mentioning that the probability of occurrence for Asn393 and His496 is obviously increased than the above stage, which plays a pivotal role in attracting peptide moving to the active site.

In the fifth stage ($20.0 \text{ \AA} \geq RC \geq 17.0 \text{ \AA}$), peptide is stabilized at the active site of OGT gradually, and its conformation shows “spread” shape with little fluctuation. Thr801, Asn805, and Pro897 form hydrogen bonds with the head part, which is benefit for peptide moving to the active site and maintaining stretched mode. Moreover, Lys634 and His496 are identified to play an important role in constraining the middle and tail part of peptide with high probability by hydrogen bond interactions. For the mobility of peptide, other residues, such as His517, Tyr632, Thr633, and Gln839 also form key interactions with the substrate.

What is noteworthy is that the release mechanism is strongly related with the interactions between OGT and different subsection of the peptide. Typically, we find that if the peptide is divided into two segments according to the characteristic of interactions around the peptide and the conformation changes of the peptide, more particular and deep comprehending can be acquired. The two segments are the residues of Tyr13-Pro19 and Val20-Met26, which are named as first part and second part, respectively, as Fig. 9 and Fig. 10 depicted. For the first part, it swings strongly because of the flexible property of peptide and alternated residues around them. It should take major responsibility for guidance of mobile direction and three times of “V → spread” conformation changes of peptide. It is related with the region comprised with residues Thr315-Ala650 of OGT. As to the second part, great majority hydrogen bond interactions are formed with it, which mainly provide the driving force for peptide binding and locating. It is relevant to the part involved with residues of Met651-Lys1028. In order to clearly clarify the feature of two segments, the key points during the whole process are extracted with new present mode

shown in Fig. 9. We can see clearly that in the first stage, the first part is prior recognized and the second part is freely outside the protein. Along the further binding of the peptide, the second part is recognized, and then the two parts are both binding gradually. During the second to fourth stage, the first part guides the binding direction by swinging that influences the conformation changes of peptide, and the second part is relatively stable with growing hydrogen bond interactions. The behavior of both sections is benefit for further transportation of peptide to the inward of OGT. In the fifth stage, the first part finds the most comfortable direction, and both segments are stabilized in the active site with stable conformation and interactions. Consequently, it can be further proved that the peptide recognition and binding process are highly dependent on the features of surroundings for the different subsection of peptide.

Let us go back to the free energy profile, which is depicted in Fig. 11. In the first stage, the peptide is recognized primitively by OGT with head and middle part prior binding. Meanwhile, the tail part floats freely outside the entrance without any constrains. For the high occurred probability of key residues, stable hydrogen bond interactions and the similar conformations of each window, the free energy exhibits gently decreased trend with about 4.0 kcal/mol. As to the second stage, the new hydrogen bond interactions increased significantly by comparing that of first stage, especially for the location around the middle and tail part of peptide. It implies that the middle and tail part is gradually further recognized, which provides the driving force for the peptide delivery towards the interior of OGT. Moreover, the quantity of broken interactions between OGT and peptide is obviously less than the formation interactions. In addition, the peptide continues to show the curly “V” shape with weak conformation changes. Accordingly, the free energy shows a gradual exothermic trend by gathering up the threads for about 15.4 kcal/mol. With regard to the third stage, the number of hydrogen bond interactions is further increased along the peptide moving toward the active site. Here more interactions arisen from residues of Asn393, His461 and His498. The fragment of Tyr632-Arg637 distinctly appeared around the middle and tail part of peptide that offers main driving force for the peptide binding. Furthermore, it should be noted that obvious conformation changes have been detected from “V” to “spread” shape, and the peptide shows a comfortable stretch state by its delivery. Accordingly, by the above points, the free energy falls off speedily by 15.1 kcal/mol with higher rate than that of the second stage. For the fourth stage, the overall number of hydrogen bond interactions is similar with the third stage, while the appeared probability of Asn393 and His496 is apparently increased than the third stage, which plays a major role in attracting peptide further binding. In addition, the peptide gradually transitions to a relaxed state by its moving. Both reasons imply that the free energy continues to exothermic by 4.7 kcal/mol with mild pattern compared with the third stage. During the fifth stage, the peptide experiences a relaxation process with a stretched state. It is stabilized in the active site gradually by reposeful interaction from high occurred probability residues. Hence the free energy shows a smooth convergence pattern. As a result, the whole process of peptide recognition and binding by OGT is strongly exothermic with about 39.2 kcal/mol, which mainly comes from the hydrogen bond interactions formation that serves as main driving force and conformation relax of “V → spread” for peptide.

4. Conclusions

OGT is related to the life activities of more than 1000 proteins, and it has an extremely important relationship with the occurrence of many diseases. Recently, it is a hot topic to study the entire

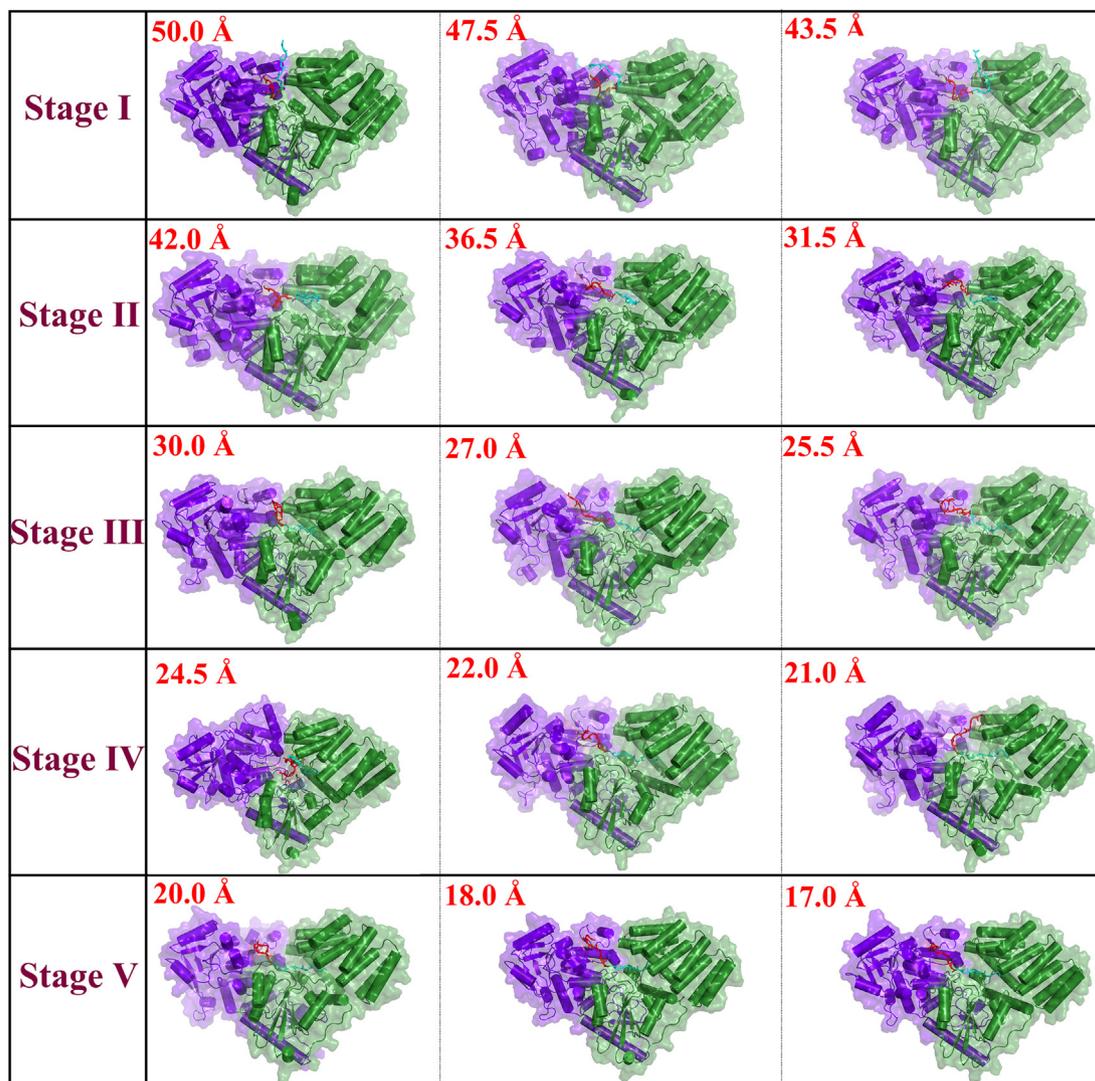


Fig. 9. The location and conformation changes of peptide recognize along P1 from subsection view. For peptide, Tyr13-Pro19 is chosen as the head part that colored with cyan, and Val20-Met26 is defined as the tail part that colored with red; for protein, Thr315-Ala650 is selected as the left part that colored with purple, and Met651-Lys1028 is used as the right part that colored with forest. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

enzymatic catalysis, protein recognized mechanisms, and design of inhibitors for OGT. In this paper, MM MD simulations combined with MM/GBSA, RAMD MD, SMD and umbrella sampling are carried out to study the recognition process of CKII peptide in OGT protein. His496, His558, Thr633, Lys634, and Pro897 stabilize the peptide at the active site through hydrogen bonds on the basis of the equilibrium state of OGT-peptide complex from MM MD simulations. Energy decomposing analysis and calculation of six mutation models indicate that the binding free energy is mainly composed by Van der Waals and electrostatic interaction, and residues of Lys634, Pro559, Thr633, Pro897, Phe868, His496, and His558 contribute most to the binding of peptide. As to His496, Pro559, and Phe868, the side chain plays an extremely significant role in peptide binding, and for Thr633, Lys634, and Pro897, the main chain has a great effect. Apart from this, two possible release channels of P1 and P2 for the peptide are discovered, and P1 located among residues of Loop 1, Loop 2 and Loop 3 is preponderance. Moreover, the head part of peptide is recognized prior to the tail part for P1. The whole recognition and binding process is strongly exothermic by 39.2 kcal/mol. The main driving force probably mainly comes from the increasing formed of the hydrogen bonds between peptide and OGT and the relaxed of peptide with

conformation changes. More notably, three distinct properties significantly attribute the entire process. Firstly, the hydrogen bond interactions are identified as the protagonist for driving the recognition and delivery of peptide. Here Val895 and Asn805 are the most important residues for peptide initial recognition through hydrogen bonds with high occurred probability. The hydrogen bonds from Asn393, His461, His496, His498, and Lys634 are key contributor to drive the further movement of peptide. Secondly, the conformation of OGT is not changed significantly, while the peptide experiences particular changes of three times of “V → spread” for its flexible feature and various interactions around its head part, which bring important effect for the peptide delivery process. If Ser432, Asp554, and Asp587 interacts with head part of peptide, the “V” shape can be observed; while when the Gly783, Thr801, and Val895 appear around the head part, the “spread” shape can be detected. Finally, different parts of the peptide have different performance. The recognition and binding mechanism particularly depends on the hydrogen bond interactions between OGT and different subsections of the peptide. Tyr13-Pro19 of peptide leads its moving direction, which directly influences the conformation changes; meanwhile interactions mainly focus on Val20-Met26 are responsible for the location

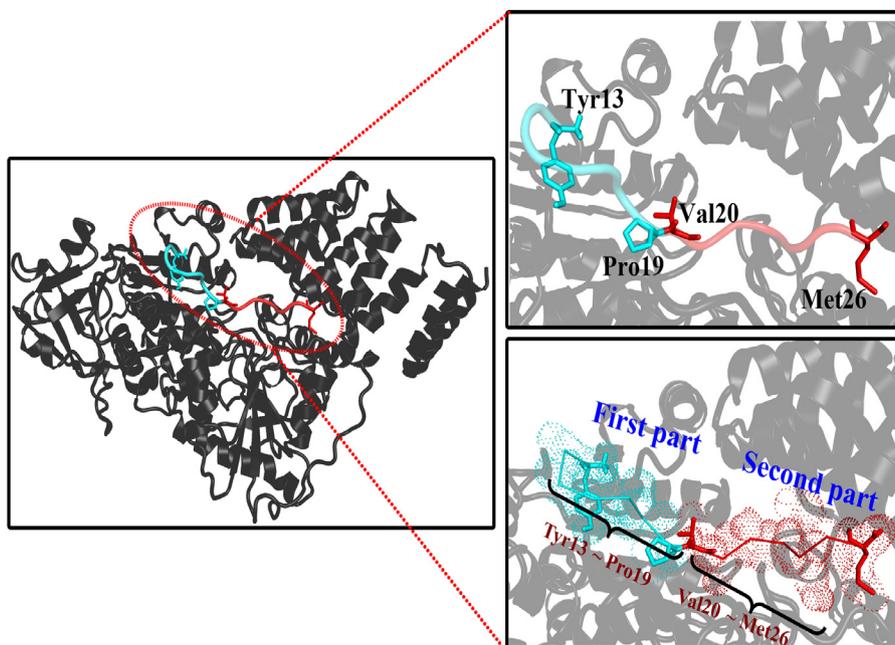


Fig. 10. Two segments of peptide, first part is colored as cyan that refers Tyr13-Pro19, and the second part is colored as red that refers Val20-Met26. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

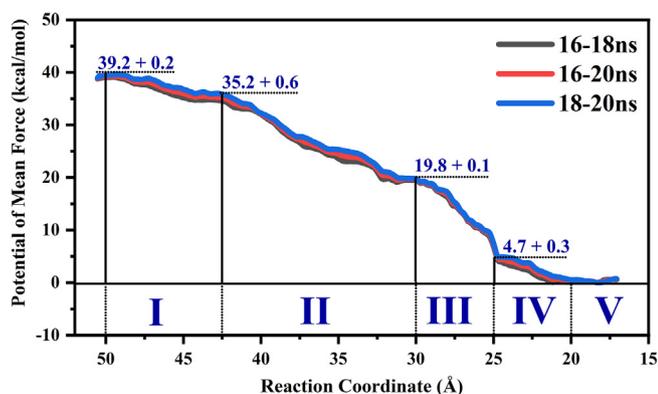


Fig. 11. Potential of mean force (PMF) for peptide recognizing and binding along the P1 channel by using MM MD simulations combined with umbrella sampling technology.

and stability of peptide. Although the CKII peptide recognition and binding mechanism studied here is typical, it also offers important useful theoretical clues for understanding performance of the other peptide with similar residues and features. For one thing, similar residues probably have similar interactions, which are closely related to recognition and binding mechanisms. For another, with regards to the long and flexible properties of peptide, the different performance of different parts may appear. In addition, what is also noteworthy is that various methods combination provide a general framework. It can be applied in the transported mechanism calculation for long substrate with flexible properties in enzyme system. Therefore, our results provide important theoretical clues for the peptide recognition and delivery process in enzymatic catalysis at the atomic level, which is also crucial for protein identification. Furthermore, it is probably useful for the inhibitor design that targets the interaction between OGT and target proteins. We hope this work will accelerate the progress of the related research both in theory and experiment.

CRediT authorship contribution statement

Xiao Zhang: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft. **Zhiyang Zhang:** Methodology, Formal analysis. **Jia Guo:** Resources, Investigation. **Jing Ma:** Writing - original draft. **Songqiang Xie:** Conceptualization. **Yuan Zhao:** Conceptualization, Supervision, Methodology, Formal analysis, Writing - original draft, Funding acquisition. **Chaojie Wang:** Conceptualization, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funding Sources: This work was supported by the National Science Foundation of China (Nos. 22073023 and 21603057), the Foundation of Science and Technology Department of Henan Province (No. 212102310238), the Project funded by the China Post-doctoral Science Foundation (Nos. 2017M622324 and 2018T110721), and the Project funded by the Henan Postdoctoral Science Foundation (No. 001702017). We thank Supercomputer Center in Wuhan University and East China Normal University as well as National Supercomputer Center in Changsha and Guangzhou for providing the computational resources.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.04.009>.

References

- [1] Julka S, Regnier F. Quantification in proteomics through stable isotope coding: a review. *J Proteome Res* 2004;3(3):350–63.

- [2] Schwarz F, Aebi M. Mechanisms and principles of N-linked protein glycosylation. *Curr Opin Struc Biol* 2011;21:576–82.
- [3] Fujiki R, Chikanishi T, Hashiba W, Ito H, Takada I, Roeder RG, et al. GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis. *Nature* 2009;459(7245):455–9.
- [4] Hanover JA, Krause MW, Love DC. Bittersweet memories: linking metabolism to epigenetics through O-GlcNAcylation. *Nat Rev Mol Cell Biol* 2012;13(5):312–21.
- [5] Rexach JE, Clark PM, Mason DE, Neve RL, Peters EC, Hsieh-Wilson LC. Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nat Chem Biol* 2012;8(3):253–61.
- [6] Durgan DJ, Pat BM, Laczby B, Bradley JA, Tsai J-Y, Grenett MH, et al. O-GlcNAcylation, novel post-translational modification linking myocardial metabolism and cardiomyocyte circadian clock. *J Biol Chem* 2011;286(52):44606–19.
- [7] Kim EY, Jeong EH, Park S, Jeong H-J, Edery I, Cho JW. A role for O-GlcNAcylation in setting circadian clock speed. *Genes Dev* 2012;26(5):490–502.
- [8] Geng F, Zhu W, Anderson RA, Leber B, Andrews DW. Multiple post-translational modifications regulate E-cadherin transport during apoptosis. *J Cell Sci* 2012;125(Pt 11):2615–25.
- [9] Lefebvre T, Ferreira S, Dupont-Walloyis L, Bussi re T, Dupire M-J, Delacourte A, et al. Evidence of a balance between phosphorylation and O-GlcNAc glycosylation of Tau proteins—a role in nuclear localization. *Biochim Biophys Acta* 2003;1619(2):167–76.
- [10] Yang X, Ongusaha PP, Miles PD, Havstad JC, Zhang F, So WV, et al. Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. *Nature* 2008;451(7181):964–9.
- [11] Pathak S, Borodkin VS, Albarbarawi O, Campbell DG, Ibrahim A, van Aalten DMF. O-GlcNAcylation of TAB1 modulates TAK1-mediated cytokine release. *EMBO J* 2012;31(6):1394–404.
- [12] Ku N-O, Toivola DM, Strnad P, Omary MB. Cytoskeletal keratin glycosylation protects epithelial tissue from injury. *Nat Cell Biol* 2010;12:876–85.
- [13] Ma J, Hart GW. Protein O-GlcNAcylation in diabetes and diabetic complications. *Expert Rev Proteomics* 2013;10(4):365–80.
- [14] de Queiroz RM, Carvalho E, Dias WB. O-GlcNAcylation: The Sweet Side of the Cancer. *Front Oncol* 2014;4:132.
- [15] Yuzwa SA, Vocadlo DJ. O-GlcNAc and neurodegeneration: biochemical mechanisms and potential roles in Alzheimer's disease and beyond. *Chem Soc Rev* 2014;43(19):6839–58.
- [16] Laczby B, Hill BG, Wang K, Paterson AJ, White CR, Xing D, et al. Protein O-GlcNAcylation: a new signaling paradigm for the cardiovascular system. *Am J Physiol Heart Circ Physiol* 2009;296(1):H13–28.
- [17] Kelly WG, Dahmus ME, Hart GW. RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *J Biol Chem* 1993;268(14):10416–24.
- [18] Roquemore EP, Chevrier MR, Cotter RJ, Hart GW. Dynamic O-GlcNAcylation of the small heat shock protein alpha B-crystallin. *Biochemistry* 1996;35(11):3578–86.
- [19] Hagmann J, Grob M, Burger MM. The cytoskeletal protein talin is O-glycosylated. *J Biol Chem* 1992;267(20):14424–8.
- [20] Holt GD, Snow CM, Senior A, Haltiwanger RS, Gerace L, Hart GW. Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine. *J Cell Biol* 1987;104(5):1157–64.
- [21] Shafi R, Iyer SP, Ellies LG, O'Donnell N, Marek KW, Chui D, et al. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc Natl Acad Sci USA* 2000;97(11):5735–9.
- [22] Kreppel LK, Blomberg MA, Hart GW. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *J Biol Chem* 1997;272(14):9308–15.
- [23] Lazarus MB, Nam Y, Jiang J, Sliz P, Walker S. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* 2011;469(7331):564–7.
- [24] Lubas WA, Hanover JA. Functional expression of O-linked GlcNAc transferase. Domain structure and substrate specificity. *J Biol Chem* 2000;275(15):10983–8.
- [25] Torres CR, Hart GW. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J Biol Chem* 1984;259(5):3308–17.
- [26] She N, Zhao Y, Hao J, Xie S, Wang C. Uridine diphosphate release mechanism in O-N-acetylglucosamine (O-GlcNAc) transferase catalysis. *Biochim Biophys Acta Gen Subj* 1863;2019:609–22.
- [27] Jiang J, Lazarus MB, Pasquina L, Sliz P, Walker S. A neutral diphosphate mimic crosslinks the active site of human O-GlcNAc transferase. *Nat Chem Biol* 2012;8(1):72–7.
- [28] Cheung WD, Hart GW. AMP-activated protein kinase and p38 MAPK activate O-GlcNAcylation of neuronal proteins during glucose deprivation. *J Biol Chem* 2008;283(19):13009–20.
- [29] Sybyl-X 2.1, Sybyl-X 2.1, Tripos, St. Louis, MO, USA (2013).
- [30] Bas DC, Rogers DM, Jensen JH. Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins* 2008;73(3):765–83.
- [31] S ndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J Chem Theory Comput* 2011;7(7):2284–95.
- [32] J. Wang, J.W. Wolf Rm Fau – Caldwell, P.A. Caldwell Jw Fau – Kollman, D.A. Kollman Pa Fau – Case, D.A. Case, Development and testing of a general amber force field. *J Comput Chem* 25(0192-8651) (2004) 1157-74.
- [33] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 2006;65(3):712–25.
- [34] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79(2):926–35.
- [35] Case DA, Darden TA, Cheatham III TE, Simmerling CL, Wang J, Duke RRE, Luo, Walker RC, Zhang W, Merz KM, Roberts B, Hayik S, Roitberg A, Seabra G, Swails J, G tz AW, Kolossv ry I, Wong KF, Paesani F, Vanicek J, Wolf RM, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Salomon-Ferrer R, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA. AMBER 12, AMBER 12. San Francisco: University of California; 2012.
- [36] Jung J, Mori T, Kobayashi C, Matsunaga Y, Yoda T, Feig M, et al. GENESIS: a hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations. *Wires Comput Mol Sci* 2015;5(1759-0876):310–23.
- [37] Ryckaert J-P, Ciccotti G, Berendsen H. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints – Molecular-Dynamics of N-Alkanes. *J Comput Phys* 1977;23:327–41.
- [38] Miyamoto S, Kollman P. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* 1992;13:952–62.
- [39] Pastor R, Brooks B, Szabo A. An analysis of the accuracy of Langevin and molecular dynamics algorithm. *Mol Phys* 1988;65:1409–19.
- [40] Chen F, Sun H, Wang J, Zhu F, Liu H, Wang Z, et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein-RNA complexes. *RNA* 2018;24:1183–94.
- [41] Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* 2011;51(1):69–82.
- [42] Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* 2004;55(2):383–94.
- [43] Xia B, Tsui V, Case DA, Dyson HJ, Wright PE. Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. *J Biomol NMR* 2002;22(4):317–31.
- [44] M. Feig, M.S. Onufriev A Fau - Lee, W. Lee Ms Fau - Im, D.A. Im W Fau - Case, C. L. Case Da Fau - Brooks, 3rd, C.L. Brooks, 3rd, Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures, *J Comput Chem* 25(0192-8651) (2004) 265-284.
- [45] L demann SK, Lounnas V, Wade RC. How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J Mol Biol* 2000;303(5):797–811.
- [46] Vashisth H, Abrams CF. Ligand escape pathways and (un)binding free energy calculations for the hexameric insulin-phenol complex. *Biophys J* 2008;95(9):4193–204.
- [47] Zhao Y, She N, Zhang X, Wang C, Mo Y. Product release mechanism and the complete enzyme catalysis cycle in yeast cytosine deaminase (yCD): a computational study. (BBA) – Proteins and Proteomics 2017;1865.
- [48] Klvana M, Pavlova M, Koudelakova T, Chaloupkova R, Dvorak P, Prokop Z, et al. Pathways and mechanisms for product release in the engineered haloalkane dehalogenases explored using classical and random acceleration molecular dynamics simulations. *J Mol Biol* 2009;392:1339–56.
- [49] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem* 2005;26(16):1781–802.
- [50] Chen X, Wu T, Wang Q, Shen J-W. Shield effect of silicate on adsorption of proteins onto silicon-doped hydroxyapatite (100) surface. *Biomaterials* 2008;29(15):2423–32.
- [51] Torrie G, Valleau J. Monte Carlo free-energy estimates using non-Boltzmann sampling—application to subcritical Lennard-Jones fluid. *Chem Phys Lett* 1974;28:578–81.
- [52] Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* 1977;23(2):187–99.
- [53] Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 1992;13(8):1011–21.
- [54] Zhao Y, Chen N, Wang C, Cao Z. A comprehensive understanding of enzymatic catalysis by hydroxynitrile lyases with S stereoselectivity from the α/β -hydrolase superfamily: revised role of the active-site lysine and kinetic behavior of substrate delivery and sequential product release. *ACS Catal* 2016;6(4):2145–57.
- [55] Magistrato AA-O, Sgrignani JA-O, Krause R, Cavalli A. Single or multiple access channels to the CYP450s active site? An answer from free energy simulations of the human aromatase enzyme. *J Phys Chem Lett* 2017;8(1948-7185):2036–42.
- [56] Zhang Z, Fan F, Luo W, Zhao Y, Wang C. Molecular dynamics revealing a detour-forward release mechanism of tacrine: implication for the specific binding characteristics in butyrylcholinesterase. *Front Chem* 2020;8:730.