




# A structural model for (GlcNAc)<sub>2</sub> translocation *via* a periplasmic chitooligosaccharide-binding protein from marine *Vibrio* bacteria

Received for publication, April 23, 2021, and in revised form, August 4, 2021 Published, Papers in Press, August 13, 2021,

<https://doi.org/10.1016/j.jbc.2021.101071>

Yoshihito Kitaoku<sup>1</sup>, Tamo Fukamizo<sup>1,\*</sup>, Sawitree Kumsaoad<sup>1</sup>, Prakayfun Ubonbal<sup>1</sup>, Robert C. Robinson<sup>1,2,\*</sup>, and Wipa Suginta<sup>1,\*</sup>

From the <sup>1</sup>School of Biomolecular Science and Engineering (BSE), Vidyasirimedhi Institute of Science and Technology (VISTEC), Rayong, Thailand; <sup>2</sup>Research Institute of Interdisciplinary Science (RIIS), Okayama University, Okayama, Japan

Edited by Gerald Hart

*Vh*CBP is a periplasmic chitooligosaccharide-binding protein mainly responsible for translocation of the chitooligosaccharide (GlcNAc)<sub>2</sub> across the double membranes of marine bacteria. However, structural and thermodynamic understanding of the sugar-binding/-release processes of *Vh*CBP is relatively less. *Vh*CBP displayed the greatest affinity toward (GlcNAc)<sub>2</sub>, with lower affinity for longer-chain chitooligosaccharides [(GlcNAc)<sub>3–4</sub>]. (GlcNAc)<sub>4</sub> partially occupied the closed sugar-binding groove, with two reducing-end GlcNAc units extending beyond the sugar-binding groove and barely characterized by weak electron density. Mutation of three conserved residues (Trp<sup>363</sup>, Asp<sup>365</sup>, and Trp<sup>513</sup>) to Ala resulted in drastic decreases in the binding affinity toward the preferred substrate (GlcNAc)<sub>2</sub>, indicating their significant contributions to sugar binding. The structure of the W513A–(GlcNAc)<sub>2</sub> complex in a ‘half-open’ conformation unveiled the intermediary step of the (GlcNAc)<sub>2</sub> translocation from the soluble CBP in the periplasm to the inner membrane–transporting components. Isothermal calorimetry data suggested that *Vh*CBP adopts the high-affinity conformation to bind (GlcNAc)<sub>2</sub>, while its low-affinity conformation facilitated sugar release. Thus, chitooligosaccharide translocation, conferred by periplasmic *Vh*CBP, is a crucial step in the chitin catabolic pathway, allowing *Vibrio* bacteria to thrive in oceans where chitin is their major source of nutrients.

*Vibrio* spp. are Gram-negative, ubiquitous bacteria that live in oceanic, estuarine, and fresh waters. Several *Vibrio* species are serious pathogens of humans, aquatic animals, and invertebrates, including crustaceans and zooplankton (1). In marine ecosystems, *Vibrio* spp. are the major consumers of chitin, the most abundant polysaccharide, and thus control the carbon and nitrogen balance on earth (2). Studies on chitin assimilation by *Vibrio* spp. revealed the expression profile of genes involved in chitin utilization and chitin-signal transduction (3–5). Chitin, a β-1,4-linked polysaccharide of

GlcNAc, is hydrolyzed by various types of chitinases (6), β-N-acetyl-glucosaminidase (7, 8) and chitin deacetylases (9). *Vibrio* spp. possess all of these enzymes, efficiently catabolizing chitin as their nutritional source. The enzymatic degradation of chitin produces chitooligosaccharides (GlcNAc)<sub>n</sub> (n = 2, 3, 4, 5, and 6), which are transported into the periplasmic space through an outer membrane porin, namely chitoporin (ChiP) (10–12). The transported (GlcNAc)<sub>n</sub> are further recognized by a periplasmic solute-binding protein (SBP) specific for (GlcNAc)<sub>n</sub>, namely chitooligosaccharide-binding protein (CBP) (13). CBP itself is a biological component of the chitooligosaccharide-specific ABC transporter localized to the inner membrane (5).

According to the classification proposed by Scheepers *et al.* (14), SBPs are structurally classified in seven clusters (A, B, C, D, E, F, and G), and cluster C is subdivided into five sub-clusters (I, II, III, IV, and V) (15). SBPs share a common structural feature comprising two domains connected by a flexible hinge region (16). The flexibility of the hinge enables them to admit specific solutes into the substrate-binding cavity between the two domains. The chitooligosaccharide-specific SBP is localized to the periplasmic compartment of the bacterial cell wall and serves as a mobile component of the (GlcNAc)<sub>2</sub>-specific ABC transporter (5, 11). Although the first CBP has been identified in *Vibrio cholerae* (*Vc*CBP), no functional/mechanistic details of *Vc*CBP were reported (5). Recently, Suginta *et al.* (13) reported the crystal structure of CBP from *Vibrio campbellii* (formerly *V. harveyi*) type strain ATCC BAA-1116 (*Vh*CBP) and the functional characterization indicated that *Vh*CBP strongly binds (GlcNAc)<sub>2,3,4</sub>. *Vc*CBP and *Vh*CBP belong to cluster C/subcluster IV (C-IV), members of which are specific for oligosaccharides, including manooligosaccharides, celooligosaccharides, and chitooligosaccharides (5).

Chitooligosaccharide binding to *Vibrio* CBPs has been proposed to trigger not only translocation of the bound chitooligosaccharides through the inner membrane by the (GlcNAc)<sub>2</sub>-specific ABC transporter but also activation of the gene cluster involved in the chitinolytic cascade, releasing CBPs from chitin-sensor kinase, ChiS, localized to

\* For correspondence: Wipa Suginta, [wipa.s@vistec.ac.th](mailto:wipa.s@vistec.ac.th); Tamo Fukamizo, [tamo0111fuka@gmail.com](mailto:tamo0111fuka@gmail.com); Robert C. Robinson, [br.okayama.u@gmail.com](mailto:br.okayama.u@gmail.com).

## Structural model of VhCBP in (GlcNAc)<sub>2</sub> translocation

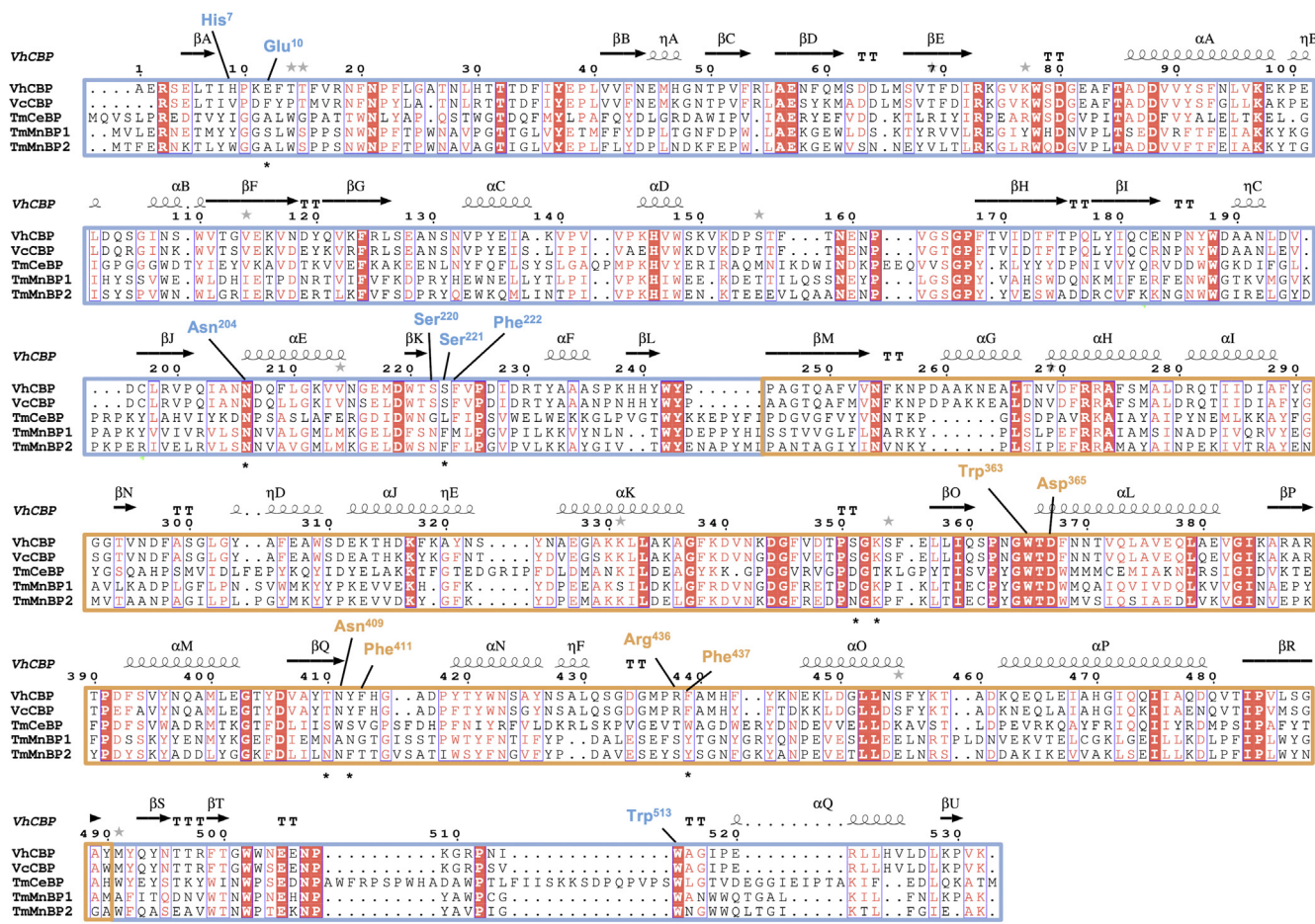
the inner membrane (17). Both biological processes have been shown to require chitooligosaccharide-induced conformational changes by the CBPs. Although the structure and mechanism of a ChiP from *Vibrio campbellii* (formerly *V. harveyi*) type strain ATCC BAA-1116 were intensively studied by crystallography in combination with physicochemical and electrochemical techniques (10, 12, 18–20), the molecular mechanism of the periplasmic CBP component involved in sugar translocation and control of gene activation is unknown. In this study, we have focused on the molecular basis of chitooligosaccharide binding and translocation by CBP, a periplasmic SBP specific for chitooligosaccharides. The specific interactions of VhCBP with chitooligosaccharides (GlcNAc)<sub>2–4</sub> were analyzed from the crystal structures. Furthermore, conserved amino acids identified as sugar-binding residues were mutated, and the mutant proteins were characterized with respect to their (GlcNAc)<sub>2</sub>-binding properties. Thermodynamic data obtained from isothermal titration calorimetry (ITC) binding experiments are discussed in the light of crystal structures of the individual VhCBP mutants. Based on these findings, we

here propose a ‘ping-pong’ model for binding/release during (GlcNAc)<sub>2</sub> translocation within the periplasmic space.

## Results

### Mutational design, recombinant expression, and purification

To identify important amino acid residues in SBPs belonging to the cluster C/subcluster IV, we carried out multiple sequence alignments of five orthologs, including VhCBP, VcCBP, celooligosaccharide-specific SBP from *Thermotoga maritima* (TmCeBP), and mannoooligosaccharide-specific SBPs from *T. maritima* (TmMnBP1 and TmMnBP2) (Fig. 1). Based on the crystal structure of VhCBP in complex with (GlcNAc)<sub>2</sub> (13), Glu<sup>10</sup>, Asn<sup>204</sup>, Ser<sup>221</sup>, Trp<sup>363</sup>, Asp<sup>365</sup>, Asn<sup>409</sup>, Phe<sup>411</sup>, Arg<sup>436</sup>, Phe<sup>437</sup>, and Trp<sup>513</sup> were found to interact with (GlcNAc)<sub>2</sub> and were therefore selected as targets for mutation. Four of these residues, Asn<sup>204</sup>, Trp<sup>363</sup>, Asp<sup>365</sup>, and Trp<sup>513</sup>, are completely conserved, and the others are partially conserved. VhCBP and the individual mutants were successfully expressed and purified, and all purified VhCBPs migrated as single protein bands on SDS-PAGE. The yields of



**Figure 1. Structure-based sequence alignment of the sugar-specific SBPs belonging to cluster C/subcluster IV.** The alignment was performed by MUSCLE and depicted with ESPrnt v.3.0. The sequences corresponding to the upper domain are boxed by the lines colored in marine and the lower domain in orange. Mutation sites are labeled with asterisks, and amino acid notations with the sequence number are colored in blue for the upper domain and in orange for the lower domain. The secondary structural elements of VhCBP (5YQW) are also aligned with the amino acid sequence of VhCBP. SBPs, solute-binding proteins; VhCBP, CBP from *V. campbellii* type strain ATCC BAA-1116.

the individual proteins from 1 l of the culture medium were 20 to 30 mg for WT and 9 to 15 mg for the mutants. We failed to produce R436A in quantities sufficient for further crystallographic and ITC binding experiments. Arg<sup>436</sup> may be involved in the correct folding of the protein because protein was mostly present in the 'inclusion bodies.'

#### Evaluation of binding affinity of WT VhCBP to (GlcNAc)<sub>2-4</sub>

We previously reported ITC data for (GlcNAc)<sub>2-4</sub> binding to VhCBP in 20 mM potassium phosphate, pH 8.0, at 4 °C (13). Here, we re-evaluated the ITC data by conducting binding experiments in Tris HCl, pH 8.0, at 25 °C using (GlcNAc)<sub>n</sub> (n = 2–4) as the ligands. The measurements are summarized in Table 1. The binding interactions for individual ligands were determined assuming a 1:1 stoichiometry. Favorable enthalpy changes (–3.9 to –4.8 kcal mol<sup>–1</sup>) and entropy changes (–6.4 to –4.0 kcal mol<sup>–1</sup>) contributed similarly to the interactions. From these data, we obtained the  $\Delta G^\circ$  values and the dissociation constants ( $K_d$ ) for (GlcNAc)<sub>2-4</sub>: the binding affinities were fairly dependent on the degree of polymerization of the ligands, with (GlcNAc)<sub>3</sub> exhibiting 3-fold and (GlcNAc)<sub>4</sub> 12-fold lower affinity than (GlcNAc)<sub>2</sub>.

#### Crystal structures of VhCBP in complex with (GlcNAc)<sub>3</sub> or (GlcNAc)<sub>4</sub>

In the previous article, we reported the mode of binding of (GlcNAc)<sub>2</sub> to VhCBP, based on the crystal structure (13). To further analyze the sugar residue interactions in the groove extending beyond the (GlcNAc)<sub>2</sub>-binding site, we solved the crystal structures of VhCBP in complex with (GlcNAc)<sub>3</sub> and (GlcNAc)<sub>4</sub>. Structural statistics are listed in Table 2. The resolutions of the structures range from 1.5 to 1.8 Å. Figure 2 shows the bound VhCBP structures with (GlcNAc)<sub>n</sub> (n = 2, 3, and 4) and the apo-VhCBP structure model, obtained by homology modeling using the apo-structure of VcCBP (1ZTY). The overall modeled structure of apo-VhCBP (Fig. 2A) was different from those of VhCBP with bound (GlcNAc)<sub>2</sub>, (GlcNAc)<sub>3</sub>, or (GlcNAc)<sub>4</sub> (Fig. 2, B–D, respectively), especially in the upper domain colored in blue. The three ligand-bound structures were essentially identical to each other; however, minor conformational differences between the (GlcNAc)<sub>2</sub>-bound and (GlcNAc)<sub>3</sub>/(GlcNAc)<sub>4</sub>-bound structures were found at the GlcNAc-binding site of the lower domain (colored in brown).

Close-up views of the bound (GlcNAc)<sub>n</sub> are shown in Figure 3. The arrangement of amino acid residues surrounding (GlcNAc)<sub>3</sub> or (GlcNAc)<sub>4</sub> was almost identical to that in the VhCBP–(GlcNAc)<sub>2</sub> complex (13). (GlcNAc)<sub>2</sub> was found at

affinity sites 1 and 2 in the protein's binding groove (Fig. 3A), where the nonreducing end of the dimer occupied the innermost affinity site 1, while the reducing end was found at the outer affinity site 2. The electron density of (GlcNAc)<sub>3</sub> (Fig. 3B) in the VhCBP–(GlcNAc)<sub>3</sub> complex was clearly defined at the affinity sites 1, 2, and 3. The main-chain carbonyl of Ala<sup>514</sup> was found to form a hydrogen bond with O6 of the reducing-end GlcNAc of the bound (GlcNAc)<sub>3</sub>. By contrast, the electron density of the reducing-end two GlcNAc residues of the bound (GlcNAc)<sub>4</sub> was weak and uninterpretable (Fig. 3C) at the imaginary sites 3 and 4.

The absence of electron density of pyranose rings may be caused by the disorder of the two GlcNAc residues on the reducing-end side. In fact, the 2Fo–Fc map of the reducing-end GlcNAc of the bound (GlcNAc)<sub>3</sub> was also patchy, allowing only partial occupancy of this sugar ring to be set (Fig. S1). Figure 3D shows the superimposition of the complex structures of VhCBP–(GlcNAc)<sub>2</sub>, VhCBP–(GlcNAc)<sub>3</sub>, and VhCBP–(GlcNAc)<sub>4</sub>. The GlcNAc residues at affinity sites 1 and 2 overlapped completely with each other, indicating identical conformations of the bound two GlcNAc residues at these sites. We noticed that the glycosidic link between the two GlcNAc residues at affinity sites 2 and 3 was twisted, indicating that the link may have a higher conformational energy than that of the neighboring glycosidic link between sites 1 and 2. As noted above, the aromatic side chain of Phe<sup>411</sup> is flipped away from the bound GlcNAc by 3.5 Å in the complexed structures with (GlcNAc)<sub>3</sub> or (GlcNAc)<sub>4</sub> with full occupancy.

#### Binding mode of (GlcNAc)<sub>4</sub> to VhCBP predicted by docking simulation

To predict the structure of the two missing reducing-end residues of bound (GlcNAc)<sub>4</sub>, we conducted docking simulation using AutoDock Vina. Among nine conformers, the most stable provided –10.9 kcal mol<sup>–1</sup> of binding energy. However, the dihedral bond angle, defined by atoms O5–C1–O1–C4 at the glycosidic bond between the GlcNAc residues at the affinity sites 3 and 4, was –23.3°. This value is far smaller than those of the most stable conformations reported previously (about –90°) (21). Therefore, we selected the second-most stable conformer, with –10.5 kcal mol<sup>–1</sup> of binding energy, which exhibited reasonable dihedral angles ranging from –96.4° to –73.5°. Figure 4 shows the stereo view of the simulated complex structure.

#### ITC binding experiments

To evaluate the contributions of the amino acid residues in (GlcNAc)<sub>2</sub> binding, we carried out site-directed mutagenesis to

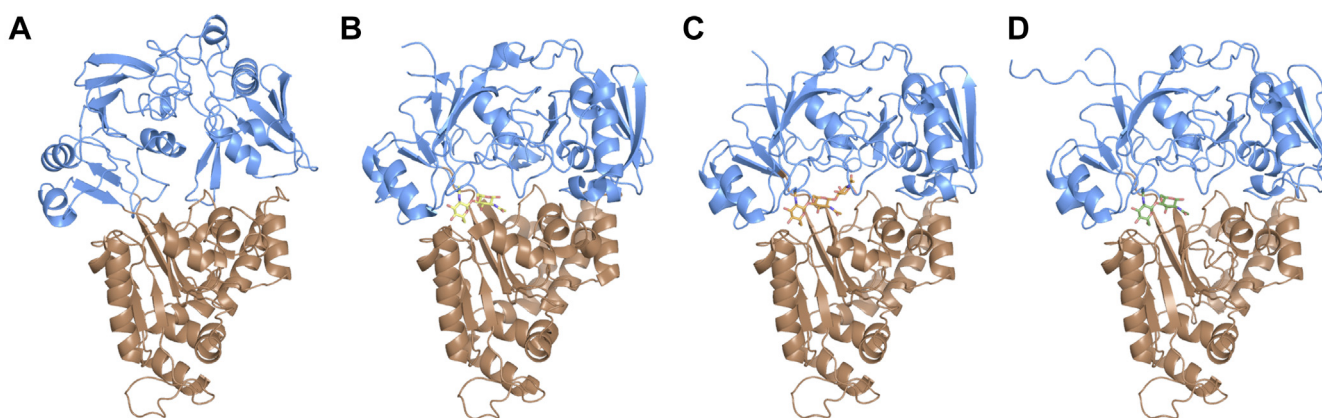
**Table 1**  
Thermodynamic parameters of VhCBP titrated with (GlcNAc)<sub>2-4</sub>

| Ligand                | n | $K_d$ (μM)  | $\Delta H^\circ$ (kcal.mol <sup>–1</sup> ) | $-T\Delta S^\circ$ (kcal.mol <sup>–1</sup> ) | $\Delta G^\circ$ (kcal.mol <sup>–1</sup> ) |
|-----------------------|---|-------------|--|--|--|
| (GlcNAc) <sub>2</sub> | 1 | 0.03 ± 0.01 | –3.9 ± 0.1                                 | –6.4 ± 0.2                                   | –10.3 ± 0.3                                |
| (GlcNAc) <sub>3</sub> | 1 | 0.09 ± 0.05 | –4.7 ± 0.1                                 | –5.0 ± 0.3                                   | –9.7 ± 0.3                                 |
| (GlcNAc) <sub>4</sub> | 1 | 0.40 ± 0.24 | –4.8 ± 0.6                                 | –4.0 ± 1.0                                   | –8.8 ± 0.3                                 |

Thermodynamic parameters for WT were obtained by ITC experiments. The titration conditions for the tested chitooligosaccharides and the VhCBP variants are described in text. All thermodynamic values presented in Figure 6A are the mean ± SD, which were obtained from three separate sets of ITC experiments.

**Table 2**  
X-ray data collection and refinement statistics for VhCBP variants in complex with (GlcNAc)<sub>2-4</sub>

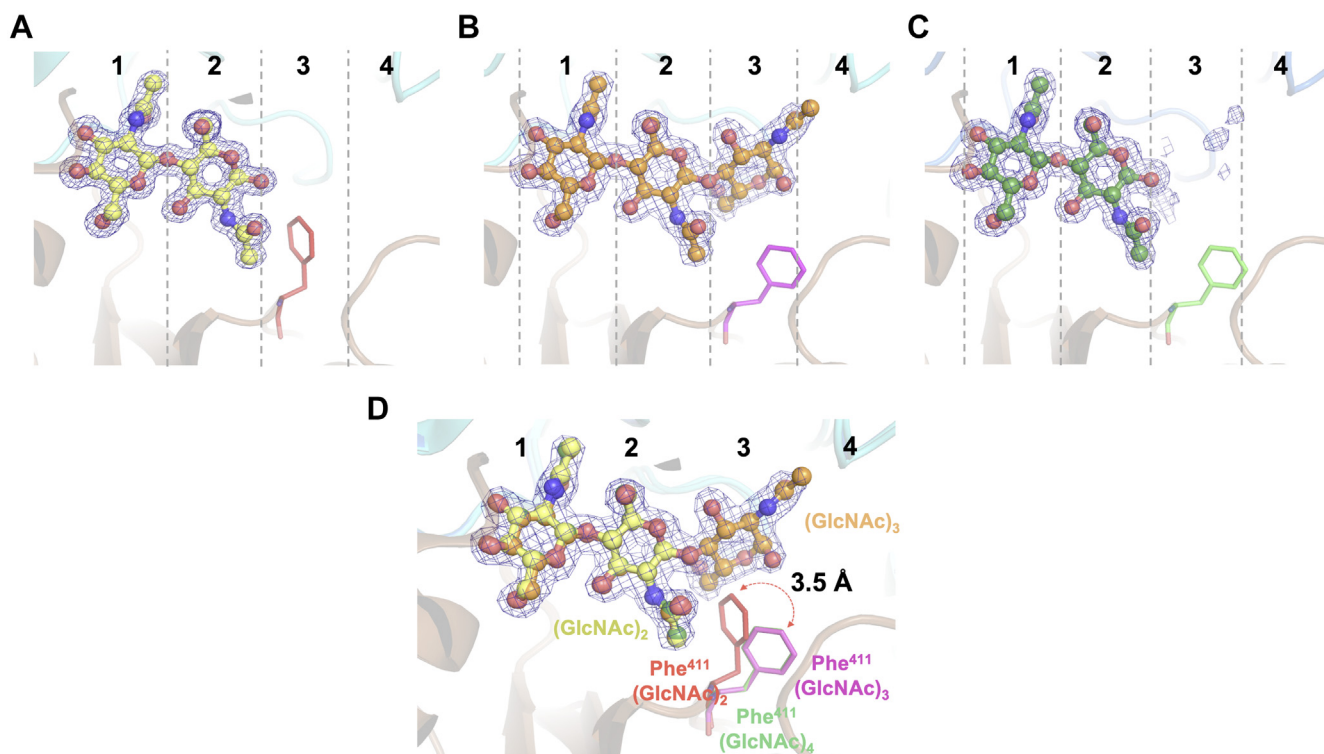
| VhCBP variant                                       | WT                         | WT                         | W363A                      | N409A                      | F411A                      | F437A                      | W513A                      |
|---|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| Ligand  | (GlcNAc) <sub>3</sub>      | (GlcNAc) <sub>4</sub>      | (GlcNAc) <sub>2</sub>      | (GlcNAc) <sub>2</sub>      | (GlcNAc) <sub>2</sub>      | (GlcNAc) <sub>2</sub>      | (GlcNAc) <sub>2</sub>      |
| PDB code  | 6LZQ                       | 7EBI                       | 7EBM                       | 6LZT                       | 6LZU                       | 6LZV                       | 6LZW                       |
| Data collection                                     |                            |                            |                            |                            |                            |                            |                            |
| Space group   | <i>P</i> 12 <sub>1</sub> 1 | <i>P</i> 12 <sub>1</sub> 1 | <i>P</i> 12 <sub>1</sub> 1 | <i>P</i> 12 <sub>1</sub> 1 | <i>P</i> 12 <sub>1</sub> 1 | <i>P</i> 12 <sub>1</sub> 1 | <i>P</i> 12 <sub>1</sub> 1 |
| Molecule/asymmetric unit                            | 1                          | 1                          | 1                          | 1                          | 1                          | 1                          | 1                          |
| Unit-cell parameters                                |                            |                            |                            |                            |                            |                            |                            |
| a, b, c (Å)   | 61.8, 56.8, 81.9           | 61.9, 57.1, 82.1           | 58.6, 56.0, 80.3           | 59.5, 56.4, 80.9           | 62.3, 57.2, 82.0           | 59.5, 57.0, 85.2           | 61.7, 57.1, 81.5           |
| α, β, γ (°)   | 90.0, 99.4, 90.0           | 90.0, 99.5, 90.0           | 90.0, 101.8, 90.0          | 90.0, 101.0, 90.00         | 90.0, 100.5, 90.0          | 90.0, 101.4, 90.0          | 90.0, 101.7, 90.0          |
| Wavelength (Å)                                      |                            |                            |                            | 0.99984                    |                            |                            |                            |
| Resolution range (Å)                                | 50.0–1.80 (1.83–1.80)      | 50.0–1.50 (1.53–1.50)      | 50.0–2.00 (2.03–2.00)      | 50.0–1.85 (1.88–1.85)      | 50.0–1.90 (1.93–1.90)      | 20.0–2.20 (2.24–2.20)      | 50.0–1.90 (1.93–1.90)      |
| Completeness (%)                                    | 92.2 (85.5)                | 97.8 (95.7)                | 99.9 (100.0)               | 99.8 (100.0)               | 98.2 (91.8)                | 99.6 (95.2)                | 95.0 (91.7)                |
| <i>R</i> <sub>merge</sub>                           | 0.128 (1.114)              | 0.065 (0.396)              | 0.184 (1.863)              | 0.112 (0.687)              | 0.105 (0.552)              | 0.066 (0.593)              | 0.105 (0.757)              |
| <i>R</i> <sub>pim</sub>                             | 0.085 (0.765)              | 0.040 (0.312)              | 0.075 (0.796)              | 0.067 (0.423)              | 0.068(0.444)               | 0.040 (0.376)              | 0.075 (0.630)              |
| Multiplicity  | 3.2 (2.9)                  | 3.4 (2.3)                  | 6.9 (5.8)                  | 3.7 (3.6)                  | 3.2 (2.1)                  | 3.7 (3.3)                  | 2.9 (2.5)                  |
| <i>I</i> / $\sigma$ ( <i>I</i> )                    | 8.8 (0.8)                  | 20.0 (2.6)                 | 20.4 (2.1)                 | 27.0 (1.9)                 | 10.8 (1.4)                 | 18.5 (1.8)                 | 9.8 (4.2)                  |
| Refinement  |                            |                            |                            |                            |                            |                            |                            |
| Resolution range (Å)                                | 36.5–1.80                  | 24.5–1.50                  | 23.7–1.90                  | 23.6–1.85                  | 24.1–1.90                  | 19.9–2.20                  | 26.9–1.90                  |
| No. of reflections used                             | 43,328                     | 87,087                     | 38,780                     | 44,261                     | 40,989                     | 26,982                     | 38,988                     |
| <i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> | 0.164/0.206                | 0.174/0.192                | 0.183/0.249                | 0.236/0.296                | 0.186/0.234                | 0.171/0.225                | 0.221/0.289                |
| No. of atoms  |                            |                            |                            |                            |                            |                            |                            |
| Protein   | 4301                       | 4338                       | 4250                       | 4289                       | 4158                       | 4301                       | 4156                       |
| Ligand  | 118                        | 81                         | 32                         | 29                         | 41                         | 33                         | 33                         |
| Ion   | 8                          | 8                          | 4                          | 2                          | 4                          | 6                          | 3                          |
| Water   | 337                        | 308                        | 261                        | 135                        | 256                        | 133                        | 134                        |
| RMSD  |                            |                            |                            |                            |                            |                            |                            |
| Bonds (Å)   | 0.007                      | 0.006                      | 0.008                      | 0.008                      | 0.008                      | 0.009                      | 0.008                      |
| Angles (°)  | 1.052                      | 1.067                      | 1.073                      | 1.115                      | 1.037                      | 1.162                      | 1.134                      |
| Ramachandran plot                                   |                            |                            |                            |                            |                            |                            |                            |
| Favored (%)   | 96.5                       | 97.8                       | 96.8                       | 95.3                       | 96.2                       | 94.04                      | 93.5                       |
| Allowed (%)   | 3.2                        | 2.2                        | 3.0                        | 4.5                        | 3.4                        | 5.59                       | 5.6                        |
| Disallowed (%)                                      | 0.4                        | 0.0                        | 0.2                        | 0.2                        | 0.4                        | 0.37                       | 1.0                        |



**Figure 2. Cartoon representations of the structures of VhCBP variants.** A, apo-VhCBP model and structures of VhCBP in complex with (B) (GlcNAc)<sub>2</sub>, (C) (GlcNAc)<sub>3</sub>, and (D) (GlcNAc)<sub>4</sub>. The upper domain is colored in blue and the lower domain in brown. The bound (GlcNAc)<sub>n</sub> (n = 2, 3, and 4) are represented by stick models colored in yellow, orange, and green, respectively. The apo structure model shown in A was obtained by homology modeling using the VcCBP structure (4GF8) as a template. VcCBP, CBP from *Vibrio cholerae*; VhCBP, CBP from *Vibrio campbellii* type strain ATCC BAA-1116.

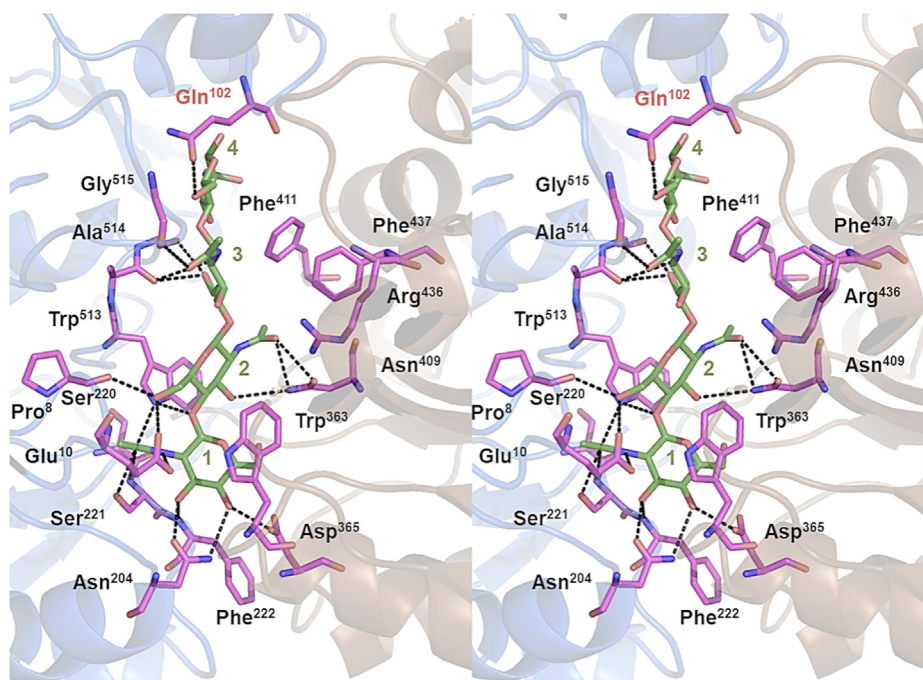
substitute the designated residues and successfully produced nine single mutants, E10A, N204A, S221A, W363A, D365A, N409A, F411A, F437A, and W513A, and one double mutant, W363A/W513A. All proteins were purified to homogeneity by Ni-NTA affinity chromatography, followed by gel filtration, and their purities were checked by SDS-PAGE analysis (Fig. S2). These proteins were used for ITC binding experiments with

(GlcNAc)<sub>2</sub> as the ligand. The ITC results are shown in Figure 5, and the thermodynamic parameters are listed in Table 3. The bar diagrams of the thermodynamic parameters,  $\Delta G^\circ$ ,  $\Delta H^\circ$ , and  $-T\Delta S^\circ$ , are also shown in Figure 6A. All binding interactions showed exothermic profiles (Fig. 5), but the heat generation varied in different mutants. For the double mutant W363A/W513A, (GlcNAc)<sub>2</sub> titration did not exhibit any clear heat



**Figure 3. Close-up view of bound (GlcNAc)<sub>n</sub> (n = 2, 3, and 4) in the binding groove of VhCBP.** A, (GlcNAc)<sub>2</sub> (5YQW), (B) (GlcNAc)<sub>3</sub> (6LZQ), and (C) (GlcNAc)<sub>4</sub> (7EBI). The molecules of (GlcNAc)<sub>n</sub> in the complex structures are represented by stick models colored yellow for (GlcNAc)<sub>2</sub>, orange for (GlcNAc)<sub>3</sub>, and forest green for (GlcNAc)<sub>4</sub>. The Phe<sup>411</sup> side chains are depicted in stick models and colored brown for the VhCBP-(GlcNAc)<sub>2</sub> complex, magenta for VhCBP-(GlcNAc)<sub>3</sub>, and green for (GlcNAc)<sub>4</sub>. The affinity subsites are designated by numbering from the inner (left)-most site (1, 2, 3, and 4). The 2F<sub>o</sub>-F<sub>c</sub> map at  $\sigma = 1.0$  (blue) around ligand molecules is depicted in a mesh representation. The main-chain structure of VhCBP is represented by a cartoon model colored marine blue for the upper domain and brown for the lower domain. D, superimposition of the structures of VhCBP in complex with (GlcNAc)<sub>2</sub>, (GlcNAc)<sub>3</sub>, or (GlcNAc)<sub>4</sub>. The Phe<sup>411</sup> side chain flips away from the bound GlcNAc by 3.5 Å upon binding of (GlcNAc)<sub>3</sub> and (GlcNAc)<sub>4</sub>, probably due to steric hindrance of the bound GlcNAc at affinity site 3. VhCBP, CBP from *Vibrio campbellii* type strain ATCC BAA-1116.

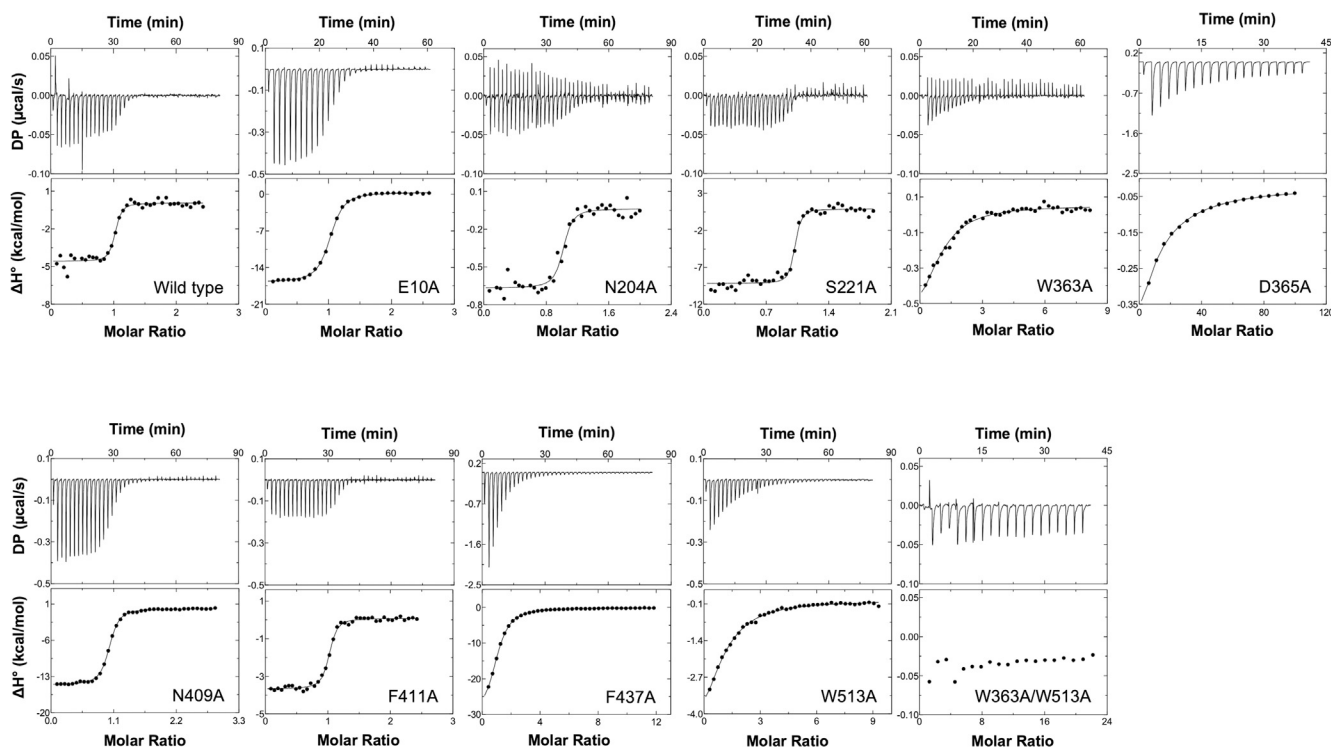
## Structural model of VhCBP in (GlcNAc)<sub>2</sub> translocation



**Figure 4. Stereo view of the modeled structure of VhCBP in complex with (GlcNAc)<sub>4</sub>.** The (GlcNAc)<sub>4</sub> structure was docked into the crystal structure of VhCBP in complex with (GlcNAc)<sub>4</sub> (7EBI) using AutoDock Vina. Upper (marine) and lower (brown) domains of the VhCBP molecule are differently shown as a cartoon model. The bound (GlcNAc)<sub>4</sub> (forest green) is shown in a ball-and-stick model, and amino acid residues directly interacting with (GlcNAc)<sub>4</sub> (magenta) are shown in stick models. An amino acid interacting with the reducing-end GlcNAc of the bound (GlcNAc)<sub>4</sub> is labeled red. VhCBP, CBP from *Vibrio campbellii* type strain ATCC BAA-1116.

release. Exceptional affinity changes ( $\Delta\Delta G^\circ$ ) were observed for the W363A, W513A, and D365A mutants. F437A also showed a relatively large change in the binding affinity. The effects on  $\Delta G^\circ$

in the other mutants (E10A, N204A, S221A, N409A, and F411A) were moderate, although one mutation had been introduced into an amino acid with a high degree of



**Figure 5. ITC profiles obtained from (GlcNAc)<sub>2</sub> titration using VhCBP WT and mutants.** Thermograms (upper) and binding isotherms with theoretical fits (lower) of the individual titrations are presented. Titration experiments were conducted in 20 mM Tris HCl, pH 8.0, at 25 °C. ITC, isothermal titration calorimetry. VhCBP, CBP from *Vibrio campbellii* type strain ATCC BAA-1116.

**Table 3**  
Thermodynamic parameters obtained for (GlcNAc)<sub>2</sub> binding to VhCBP variants

| VhCBP variant | n  | K <sub>d</sub> (μM) | ΔH° (kcal.mol <sup>-1</sup> ) | -TΔS° (kcal.mol <sup>-1</sup> ) | ΔG° (kcal.mol <sup>-1</sup> ) | ΔΔG° (kcal.mol <sup>-1</sup> ) |
|---------------|----|---------------------|-------------------------------|---------------------------------|-------------------------------|--------------------------------|
| WT            | 1  | 0.03 ± 0.01         | -3.9 ± 0.1                    | -6.4 ± 0.2                      | -10.3 ± 0.3                   | -                              |
| E10A          | 1  | 0.71 ± 0.07         | -8.8 ± 0.1                    | 0.4 ± 0.2                       | -8.4 ± 0.1                    | +1.9 ± 0.2                     |
| N204A         | 1  | 0.83 ± 0.17         | -0.75 ± 0.05                  | -7.6 ± 0.2                      | -8.3 ± 0.1                    | +2.1 ± 0.3                     |
| S221A         | 1  | 0.07 ± 0.06         | -5.3 ± 0.7                    | -4.9 ± 1.4                      | -10.2 ± 1.0                   | +0.1 ± 1.0                     |
| W363A         | 1  | 80.20 ± 19.9        | -1.9 ± 1.8                    | -3.7 ± 2.0                      | -5.6 ± 0.1                    | +4.7 ± 0.4                     |
| D365A         | 1  | 65,033 ± 4550       | -11.3 ± 0.4                   | 9.7 ± 0.4                       | -1.6 ± 0.0                    | +8.6 ± 0.3                     |
| N409A         | 1  | 0.45 ± 0.1          | -9.6 ± 0.6                    | 1.0 ± 0.7                       | -8.7 ± 0.1                    | +1.6 ± 0.2                     |
| F411A         | 1  | 0.40 ± 0.07         | -3.3 ± 0.1                    | -5.4 ± 0.1                      | -8.7 ± 0.1                    | +1.6 ± 0.3                     |
| F437A         | 1  | 29.73 ± 0.5         | -11.0 ± 0.4                   | -4.8 ± 0.4                      | -6.2 ± 0.0                    | +4.1 ± 0.3                     |
| W513A         | 1  | 36.83 ± 7.4         | -5.8 ± 0.0                    | -0.3 ± 0.1                      | -6.1 ± 0.1                    | +4.2 ± 0.3                     |
| W363A/W513A   | ND | -                   | -                             | -                               | -                             | -                              |

ND represents no detectable binding.

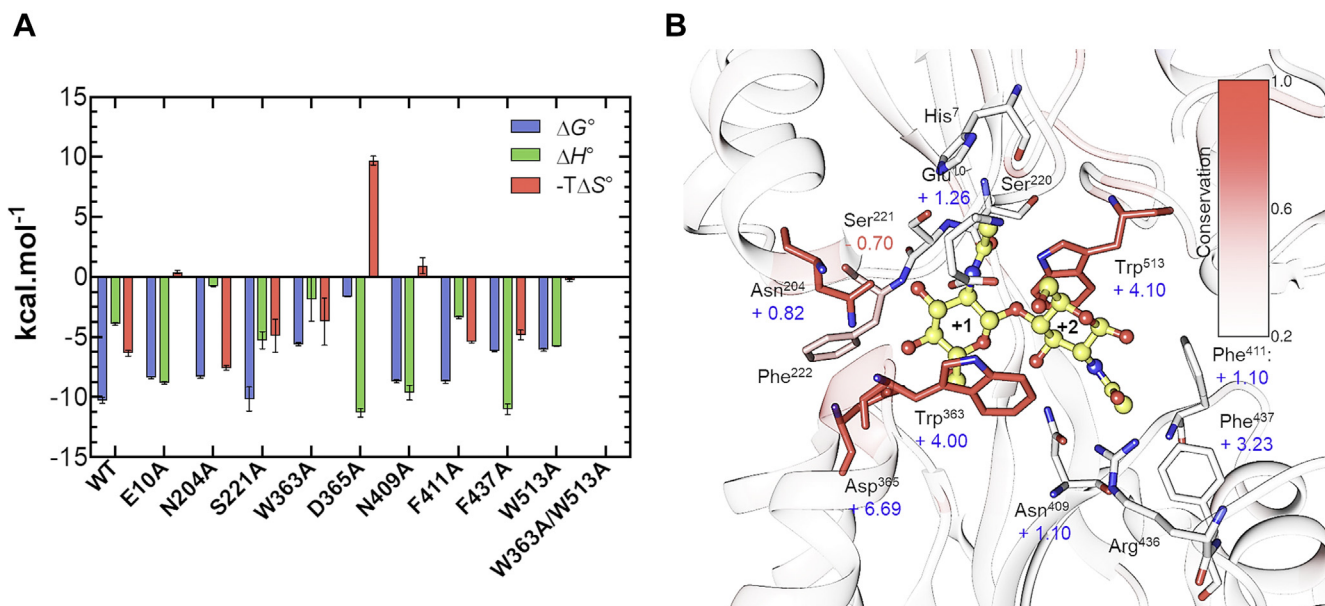
VhCBP variants were titrated with (GlcNAc)<sub>2</sub> under the conditions given in [Experimental procedures](#). Thermodynamic values (K<sub>d</sub>, ΔG°, ΔH°, -TΔS°) are derived from the corresponding thermograms shown in [Figure 5](#) and are presented as the mean ± SD, which were obtained from three separate sets of experiments.

conservation (Asn<sup>204</sup>). Changes in ΔG° (ΔΔG°) resulting from all the mutations were mapped on the structure of VhCBP in complex with (GlcNAc)<sub>2</sub> as shown in [Figure 6B](#). Degrees of conservation among five members of cluster/subcluster C-IV SBPs (VhCBP, VcCBP, TmCeBP, TmMnBP1, and TmMnBP2) are also represented by coloring dark red (high degree of conservation) to pale red (moderate degree of conservation) to white (low degree of conservation). As seen from this figure, the effects of mutations on the changes in the binding energy (ΔΔG°) were not always consistent with the degrees of conservation, especially in Asn<sup>204</sup> and Phe<sup>437</sup>.

#### Crystal structures of the VhCBP mutants in complex with (GlcNAc)<sub>2</sub>

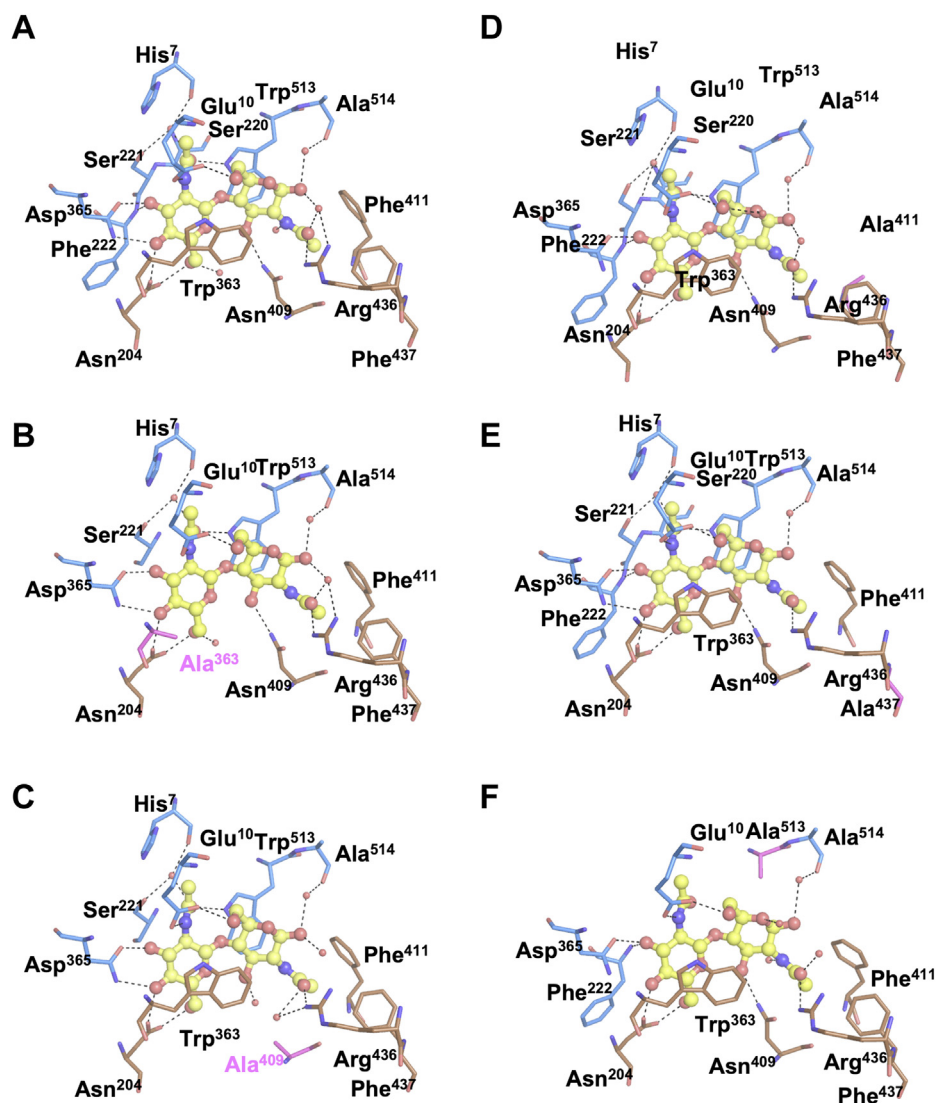
To explore the effects of mutations on the structure, the mutated VhCBPs were co-crystallized with (GlcNAc)<sub>2</sub> and

single crystals selected for X-ray diffraction experiments. It should be mentioned that the quality of the crystals of E10A, N204A, S221A, D365A, and E363A/W513A in complex with (GlcNAc)<sub>2</sub> was insufficient for X-ray diffraction experiments, and we failed to obtain high-quality datasets for structural determination. Although crystals for all the other mutants were grown under similar conditions, some variations in the crystal packing were observed ([Table 2](#)). Individual structures were solved with the space group of P12<sub>1</sub>1. Close-up views of the (GlcNAc)<sub>2</sub>-binding site of the WT and mutated VhCBPs are shown in [Figure 7](#). Compared with the WT VhCBP ([Fig. 7A](#)), W363A not only eliminated the CH-π stacking interaction with the nonreducing end GlcNAc ([Fig. 7B](#)) but also slightly shifted the Trp<sup>513</sup> side chain and the bound (GlcNAc)<sub>2</sub> ([Fig. S3](#)). These shifts may have altered the interaction of GlcNAc with amino acid side chains from the upper



**Figure 6. Thermodynamic signature and degrees of conservation of the sugar-binding residues.** A, bar diagrams representing ΔG°, ΔH°, and -TΔS° for individual VhCBP mutants. Individual parameter values were obtained from the ITC data shown in [Figure 5](#) and are depicted in blue, green, and red, respectively. B, relationship between reductions in binding affinity by mutations (ΔΔG°) and degree of conservation of individual mutated amino acids. ΔΔG° values are stated near the individual mutated amino acids. The mutated amino acids are colored according to the degree of conservation with brightness of red, with color labeled 0.2 for modestly conserved to 1 for totally conserved. The degree of amino acid conservation was calculated by dividing the number of the conserved residues in the sequence alignment by the total number of the corresponding residues in the aligned sequences ([Fig. 1](#)). The indicator of the degree of conservation is shown on the upper-right region of this figure. Values presented in [Figure 6A](#) are the mean ± SD, which were obtained from three separate sets of ITC experiments. ITC, isothermal titration calorimetry; VhCBP, CBP from *Vibrio campbellii* type strain ATCC BAA-1116.

## Structural model of VhCBP in (GlcNAc)<sub>2</sub> translocation



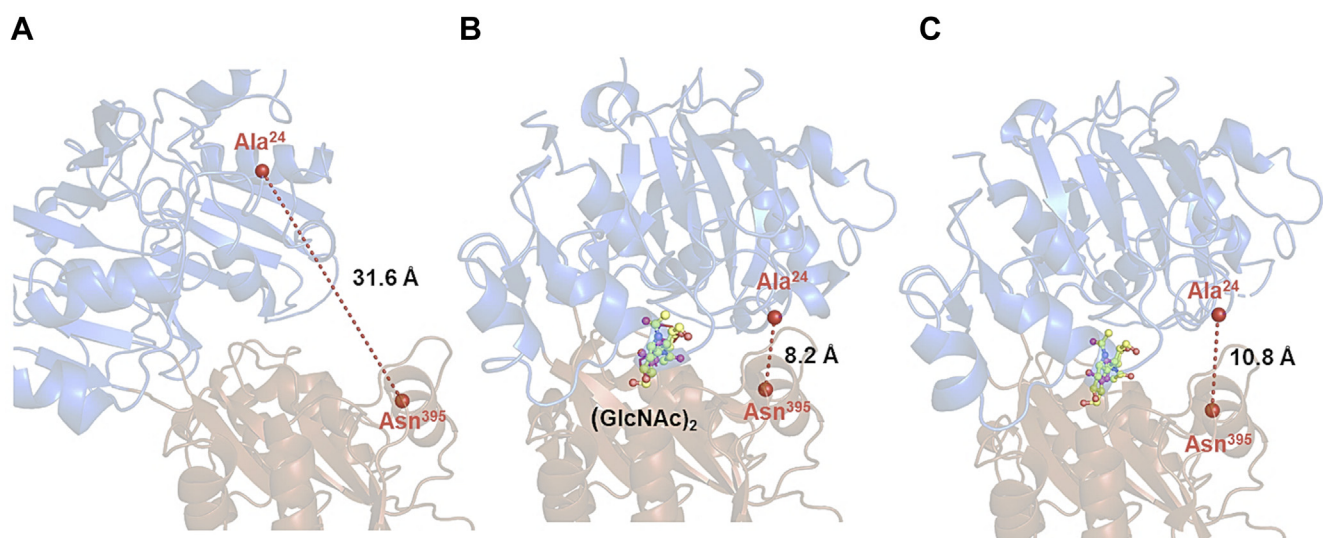
**Figure 7. Close-up views of the (GlcNAc)<sub>2</sub>-bound crystal structures.** A, WT, (B) W363A, (C) N409A, (D) F411A, (E) F437A, and (F) W513A. (GlcNAc)<sub>2</sub> is shown in a ball-and-stick model and colored yellow. The sugar-binding residues from the upper domain are shown as sticks and colored marine blue, while those from the lower domain are colored brown. Broken lines represent possible hydrogen bonds.

domain, including Trp<sup>513</sup>. The N409A mutation eliminated the hydrogen bond with the reducing-end GlcNAc, but simultaneously, two water molecules fit into the corresponding position (Fig. 7C). One water molecule interacts with O3 of the reducing-end GlcNAc, whereas the other bridges the *N*-acetyl carbonyl group and N $\eta$  of Arg<sup>436</sup>. The F411A mutation did not affect the hydrogen bonds but eliminated the hydrophobic interaction with the *N*-acetyl methyl group of the reducing-end GlcNAc (Fig. 7D). A similar effect was observed in F437A, that is, the hydrophobic interaction with the *N*-acetyl methyl group of the reducing-end GlcNAc was likely eliminated by the F437A mutation (Fig. 7E); however, in this case, the hydrogen bonding network formed by the Arg<sup>436</sup> side chain was partly affected by the Ala mutation of Phe<sup>437</sup>. As seen in Figure 7F, the W513A mutant was found to eliminate the hydrogen bond with the acetamido -NH<sub>2</sub> of the nonreducing end GlcNAc and the CH- $\pi$  stacking interaction with the pyranose ring of the reducing-end GlcNAc, most likely

causing the strong reduction of the binding affinity ( $\Delta\Delta G^\circ$ ,  $+4.2 \pm 0.3$  kcal mol<sup>-1</sup>).

A remarkable structure was seen with W513A, in which the substrate-binding groove was partially open at the reducing-end side (site 2). The distance between C $\alpha$  carbons of Ala<sup>24</sup> and Asn<sup>395</sup> located on the upper and lower domains of the apo WT VhCBP (homology structure) was 31.6 Å ('fully open', Fig. 8A), while in WT VhVBP in complex with (GlcNAc)<sub>2</sub>, it was 8.2 Å ('fully closed', Fig. 8B) and in W513A in complex with (GlcNAc)<sub>2</sub>, it was 10.8 Å ('half-open', Fig. 8C). The domain movement of W513A released bound water molecules trapped by several hydrogen bonds with the main-chain carbonyl of His<sup>7</sup>, the main-chain -NH of Ser<sup>220</sup>, the side chain O $\gamma$  of Ser<sup>221</sup>, and the *N*-acetyl carbonyl group of the nonreducing end GlcNAc, while the water molecules were retained in the other mutant structures. This conformational state of VhCBP was designated a 'half-open' conformation.





**Figure 8. Structural movement of the WT and W513A mutant in complex with (GlcNAc)<sub>2</sub>.** A, apo-WT VhCBP model, (B) (GlcNAc)<sub>2</sub>-bound WT VhCBP, and (C) (GlcNAc)<sub>2</sub>-bound W513A structures. The upper domain is colored blue and the lower domain brown. The (GlcNAc)<sub>2</sub> molecule is depicted in a ball-and-stick model (yellow). Amino acid residues of interest are labeled and depicted as stick models. Broken lines represent the distances between the corresponding C $\alpha$  atoms, which are represented as bigger red spheres. Numerals 1 and 2 represent the affinity sites 1 and 2, respectively. VhCBP, CBP from *Vibrio campbellii* type strain ATCC BAA-1116.

## Discussion

### The mode of binding of (GlcNAc)<sub>3</sub> and (GlcNAc)<sub>4</sub> to VhCBP

In most carbohydrate-related proteins that have long and extended binding grooves, the binding affinities depend strongly on the degree of polymerization of the oligosaccharide substrates (22); the higher the polymerization degree of the oligosaccharide, the greater the binding affinity. However, we report here that VhCBP binds (GlcNAc)<sub>2</sub>, (GlcNAc)<sub>3</sub>, and (GlcNAc)<sub>4</sub> with  $\Delta G^\circ$  values of  $-10.3 \pm 0.3$ ,  $-9.7 \pm 0.3$ , and  $-8.8 \pm 0.3$  kcal mol<sup>-1</sup>, respectively (Table 1). The binding affinity did not correspond to the degree of polymerization and was the lowest in the longest ligand, (GlcNAc)<sub>4</sub>. To rationalize the unique dependency on the degree of polymerization of ligands, we analyzed the crystal structures of VhCBP in complex with (GlcNAc)<sub>3</sub> and (GlcNAc)<sub>4</sub> (Figs. 2 and 3). A twist of the glycosidic linkage was observed between the GlcNAc residues at the affinity sites 2 and 3 of the bound (GlcNAc)<sub>3</sub> (Fig. 3B), suggesting a potential loss in the stabilization energy of the complexed structure. Furthermore, the Phe<sup>411</sup> side chain was found to flip away from the bound GlcNAc (3.5 Å) in the complex with (GlcNAc)<sub>3</sub> (Fig. 3D). A water molecule interacting with the carbonyl group of Ala<sup>514</sup> through a hydrogen bond is replaced by O6 of the reducing-end GlcNAc (affinity site 3) in the VhCBP–(GlcNAc)<sub>3</sub> complex. This can be seen from the modeled structure of the VhCBP–(GlcNAc)<sub>4</sub> complex (Fig. 4). No other interactions were found in the third GlcNAc, indicating only a weak interaction of the GlcNAc at site 3. This was confirmed from the patchy electron density of the reducing-end GlcNAc (Fig. S1), suggesting flexibility of this GlcNAc. The enthalpy/entropy gain derived from the additional hydrogen bond and the flexible GlcNAc residue at site 3 appear to be almost compensated by the disadvantages incurred by the flipping of the Phe<sup>411</sup> side chain and twisting of

the glycosidic bond, resulting in affinity for (GlcNAc)<sub>3</sub> that is similar to that for (GlcNAc)<sub>2</sub>.

The electron density of (GlcNAc)<sub>4</sub> on the reducing-end side was too weak to place the two GlcNAc moieties at the imaginary affinity sites 3 and 4 (Fig. 3C). This suggests that the mobility of the fourth sugar has a knock-on effect to weaken the binding of the third sugar. Flipping of the Phe<sup>411</sup> side chain was also observed in the complex with (GlcNAc)<sub>4</sub>, as in the complex with (GlcNAc)<sub>3</sub>, and appears to widen the space at sites 3 and 4. It is most likely that GlcNAc residues are present at sites 3 and 4 and are disordered because of the large space on these two sites. Docking simulation of the structure of VhCBP in complex with (GlcNAc)<sub>4</sub> revealed that interactions are formed mostly with the two GlcNAc units at the affinity sites 1 and 2, whereas on sites 3 and 4, only two hydrogen bonds are formed, by the main-chain carbonyl of Ala<sup>514</sup> and the O $\delta$  side chain of Gln<sup>102</sup>, respectively (Fig. 4). The middle glycosidic linkage of the bound (GlcNAc)<sub>4</sub> was also twisted, as in (GlcNAc)<sub>3</sub> bound to VhCBP (Fig. 3B). The disadvantage derived from the twist of the glycosidic linkage appears to predominate over the advantages derived from additional interactions/disorder of the GlcNAc residues at sites 3 and 4, resulting in the lower affinity for (GlcNAc)<sub>4</sub>. A similar observation was reported for TmCeBP in complex with cellopentaose, where cellopentaose was bound through water molecules, but not by direct interaction at the same affinity sites (23). A longer and more versatile binding cleft was observed for maltose-binding protein (MBP) from *Escherichia coli* (24, 25). MBP was identified to be able to bind maltooligomers of 2 to 7 glucose units. The crystal structure of MBP in complex with maltodextrin indicated that the strong affinity sites of MBP are more extended to sites S1, S2, and S3, enabling maltotriose to be tightly accommodated (25).

## Structural model of VhCBP in (GlcNAc)<sub>2</sub> translocation

### Major contributions of Asp<sup>365</sup>, Phe<sup>437</sup>, Trp<sup>363</sup>, and Trp<sup>513</sup>

Asp<sup>365</sup>, Trp<sup>363</sup>, and Trp<sup>513</sup> are completely conserved among the cluster C-VI SBPs (Figs. 1 and 6B), suggesting their importance in the binding of sugars to SBPs. We already reported the importance of the two tryptophan residues<sup>16</sup>, which appear to sandwich the oligosaccharide ligand (Fig. 6B). The mutational effects of Trp<sup>363</sup> and Trp<sup>513</sup> on  $\Delta G^\circ$  were similar ( $\Delta\Delta G^\circ = +4.7 \pm 0.4$  and  $+4.2 \pm 0.3$  kcal mol<sup>-1</sup>, respectively), but the effects of mutations on enthalpy/entropy contributions are significantly different between the two Trp mutants. The W363A mutation had a large effect on the enthalpic term, while in W513A, a large change was observed in the entropic contribution (Table 3). This can be attributed to the difference in the interaction modes between the two residues. Trp<sup>363</sup> interacts purely through CH- $\pi$  stacking, while Trp<sup>513</sup> interacts through both CH- $\pi$  stacking and hydrogen bonding with the *N*-acetyl group of the GlcNAc in site 1. As demonstrated from the crystal structure of W363A (Fig. S3), the Trp<sup>513</sup> side chain is shifted by the mutation of Trp<sup>363</sup> by 0.6 Å, and the bound (GlcNAc)<sub>2</sub> also shifted slightly. These rearrangements probably affect the binding interactions provided by amino acids in the upper domain, including Glu<sup>10</sup>, Asn<sup>204</sup>, Ser<sup>221</sup>, and Trp<sup>513</sup>. Similar phenomena may take place on mutation of Trp<sup>513</sup>. These effects may enhance  $\Delta\Delta G^\circ$  caused by the mutations of the two important tryptophan residues.

The greatest impact on  $\Delta G^\circ$  was observed in the Asp<sup>365</sup> mutation ( $\Delta\Delta G^\circ = +8.6 \pm 0.3$  kcal mol<sup>-1</sup>). Specifically, this mutation affected the entropic term much more than the enthalpic term, not only eliminating the two hydrogen bonds that were observed in the crystal structure (13) but also affecting the state of solvation. The number of bound water molecules around Asp<sup>365</sup> was strongly reduced by the mutation to alanine; hence, the entropy gain due to the desolvation upon GlcNAc residue binding was much lower in the D365A mutant than in the WT. This enhanced  $\Delta\Delta G^\circ$  in the D365A mutant. By contrast, the reduction in hydrophobicity in the F437A mutation may not only have weakened hydrophobic interaction with the *N*-acetyl methyl group of the reducing-end GlcNAc at the site 2 but also added a conformational restraint to Phe<sup>411</sup>. The F437A mutation may have also affected the hydrogen bond network formed by the nearby amino acid, Arg<sup>436</sup>, with the *N*-acetyl carbonyl group of the same GlcNAc and with a bound water molecule. These subtle changes resulted in a rather large impact on  $\Delta G^\circ$  ( $\Delta\Delta G^\circ = +4.1 \pm 0.3$  kcal mol<sup>-1</sup>). It is notable that Phe<sup>437</sup> is conserved only among the chitin-specific SBPs, that is, VcCBP and VhCBP, and Arg<sup>436</sup> and Phe<sup>437</sup> recognize the *N*-acetyl group of the ligand. The specificity toward GlcNAc may be mainly attributed to these two amino acids, indicating that mutational study of Arg<sup>436</sup> and Phe<sup>437</sup> may lead to specificity-engineered sugar-specific SBPs. Here, we concluded that the major contributors, Asp<sup>365</sup>, Phe<sup>437</sup>, Trp<sup>363</sup>, and Trp<sup>513</sup>, do not simply interact with GlcNAc but control the state of the surrounding amino acids or bound water molecules.

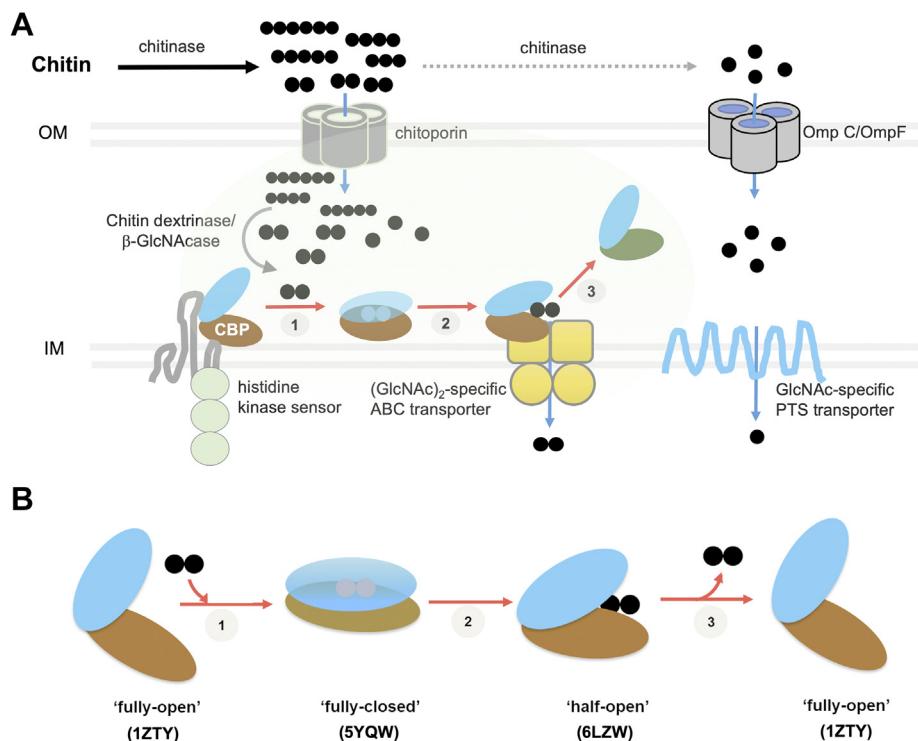
### Minor contributions of Glu<sup>10</sup>, Asn<sup>204</sup>, Ser<sup>221</sup>, Asn<sup>409</sup>, and Phe<sup>411</sup>

The other mutational effects (of E10A, N204A, S221A, N409A, and F411A) on  $\Delta G^\circ$  were moderate, although one mutation involved the amino acid with the highest degree of conservation, Asn<sup>204</sup> (Fig. 6B). Glu<sup>10</sup>, Asn<sup>204</sup>, and Ser<sup>221</sup> interact with the GlcNAc residue at site 1 with a complicated hydrogen bonding network in the narrow binding cavity. The effect of eliminating the interactions of one of these amino acids may be compensated by the interactions of the other amino acids located close to the bound (GlcNAc)<sub>2</sub>; for example, the interaction of O4 of the nonreducing GlcNAc is shared by Asn<sup>204</sup> and Asp<sup>365</sup>, while O3 of the same GlcNAc is shared by Asn<sup>204</sup> and the main-chain NH of Phe<sup>222</sup>. Thus, the mutational effect of Asn<sup>204</sup> is compensated by Asp<sup>365</sup> and Phe<sup>222</sup>. Similar observations were made for Glu<sup>10</sup> and Ser<sup>221</sup>. A hydrogen bond of Asn<sup>409</sup> with O3 of the GlcNAc at site 2 is eliminated by the N409A mutation; however, two water molecules were introduced into the VhCBP-(GlcNAc)<sub>2</sub> interface. This may be the cause of the decreased enthalpic term and increased entropic term listed in Table 3. Finally, the N409A mutation resulted in a small decrease in the binding affinity. Mutation of Phe<sup>411</sup> lowered the favorable contribution of the enthalpy change, suggesting a role of Phe<sup>411</sup> in masking the *N*-acetyl group from water molecules, avoiding direct contact with water and forming a hydrophobic interaction with the *N*-acetyl methyl group.

### A 'ping-pong' model for (GlcNAc)<sub>2</sub> translocation by CBPs from *Vibrio* spp.

The chitin catabolic cascade of *Vibrio* spp. was previously proposed by Li and Roseman (5) to be rigorously controlled by the environment signal (GlcNAc)<sub>2</sub>. Figure 9A illustrates how (GlcNAc)<sub>2</sub> plays a central role in the rigorous control of the chitin catabolic cascade. (GlcNAc)<sub>2</sub> and other chitooligosaccharides are transiently generated from chitin degradation by chitinase (26, 27) and then transported across the outer membrane through ChiP. The final degradation products (GlcNAc)<sub>2</sub> and GlcNAc are generated by periplasmic enzymes, chitin dextrinase and  $\beta$ -GlcNAcase, respectively (5). Binding of (GlcNAc)<sub>2</sub> releases the CBP from a membrane-bound histidine kinase sensor (ChiS), which essentially activates the gene cluster that is involved in the downstream cascade. CBP translocates (GlcNAc)<sub>2</sub> to the (GlcNAc)<sub>2</sub>-specific ABC transporter, for further energy production. An alternative route of chitin catabolism involves GlcNAc transport through a general diffusion porin, followed by (GlcNAc)-specific phosphoenolpyruvate:carbohydrate phosphotransferase system permease. In the chitin catabolic cascade, (GlcNAc)<sub>2</sub> acts as the intrinsic signaling molecule and a key controller of the chitin catabolic cascade.

Although we failed to obtain the crystal structure of apo-VhCBP, the crystal structure of the apo-VcCBP from *V. cholerae* (1ZTY), which is highly homologous to VhCBP (83% identity) was observed in a 'fully open' form (Fig. 8A), in



**Figure 9. The structural model for (GlcNAc)<sub>2</sub> translocation assisted by CBP in the chitin catabolism of *Vibrio* spp.** *A*, (GlcNAc)<sub>2</sub> and other chitooligosaccharides derived from the degradation of environmental chitin by chitinase are transported across outer membrane through chitoporin. (GlcNAc)<sub>2</sub> and GlcNAc are further generated in the periplasm by chitin dextrinase/β-GlcNAcase (5). Binding of (GlcNAc)<sub>2</sub> releases CBP from an inner membrane-bound histidine kinase sensor. (GlcNAc)<sub>2</sub> is translocated to the (GlcNAc)<sub>2</sub>-specific ABC transporter. GlcNAc, on the other hand, is transported across outer membrane through OmpC/OmpF and then into the cytoplasm by a carbohydrate phosphotransferase system-driven permease. The (GlcNAc)<sub>2</sub> translocation by CBP is proposed as a three-step process (oval, shaded area); step 1, (GlcNAc)<sub>2</sub> binds the 'fully open' form of CBP and induces the 'fully closed' form of CBP; step 2, CBP translocates (GlcNAc)<sub>2</sub> to the (GlcNAc)<sub>2</sub>-specific transporter; step 3, the CBP releases (GlcNAc)<sub>2</sub> to the transporter. *B*, the 'ping-pong' model for the (GlcNAc)<sub>2</sub> translocation suggested by the crystal structures of CBP in the absence or presence of (GlcNAc)<sub>2</sub>. The 'fully open' form was observed for the ligand-free CBP; the 'fully closed' form for the CBP tightly bound to (GlcNAc)<sub>2</sub>; and the 'half open' form for the CBP loosely bound to (GlcNAc)<sub>2</sub>. Blue indicates the upper lobe and dark brown the lower lobe. GlcNAc and chitooligosaccharides are presented in black, filled circles. CBP, chitooligosaccharide-binding protein.

which the upper and lower lobes are widened. Binding of (GlcNAc)<sub>2</sub> induced domain movement, as observed in the crystal structure of VhCBP in complex with (GlcNAc)<sub>2</sub> (5YQW)<sup>16</sup> and agreeing with the structure of VcCBP in complex with the same substrate (1ZU0). Such a 'sugar-induced-fit' produced a 'fully closed' state. The most interesting finding in our study was made with W513A in complex with (GlcNAc)<sub>2</sub>. Trp<sup>513</sup> is located at the mouth of the GlcNAc-binding groove, and its position is thought to regulate the translocation of (GlcNAc)<sub>2</sub> to the (GlcNAc)<sub>2</sub>-specific ABC transporter. In fact, the crystal structure of the W513A mutant showed a 'half-open' conformation (Fig. 8C), which effectively provides a snapshot of the sugar-releasing process that facilitates translocation. An uptake system for an acidic polysaccharide, alginate, in *Sphingomonas* sp. A1 was thoroughly investigated by crystallography (28, 29). The crystal structure of alginate-specific SBP (AlgQ2) in complex with an alginate-specific ABC transporter (AlgM1M2SS) was solved in the presence of alginate oligosaccharide. AlgQ2 and AlgM1M2SS form a tunnel-like binding cavity at their interface, and the nonreducing end of bound alginate oligosaccharide interacts strongly with the inner part (closed end, affinity sites 1 and 2) of the binding cavity in the closed conformation of AlgQ2. However, the reducing-end residue is exposed from the exit of

the tunnel-like binding cavity (open end, affinity sites 3, 4, 5...) and the oligosaccharide appears to be translocated into the charged surface of the transmembrane domain of AlgM1M2SS (29). Although the sequence similarity is quite low, the 'half-open' conformation of W513A obtained for VhCBP is analogous to AlgQ2 in complex with AlgM1M2SS. VhCBP interacts strongly with (GlcNAc)<sub>n</sub> at the nonreducing end, while the reducing end interacts loosely or unfavorably with VhCBP. After binding to the corresponding transporter, VhCBP in complex with (GlcNAc)<sub>n</sub> may undergo a subtle change in its conformation from 'closed' to 'half-open' by eliminating the interaction with Trp<sup>513</sup>. This half-opening may further enhance the looseness of the interaction, facilitating the translocation of (GlcNAc)<sub>n</sub>. The bound (GlcNAc)<sub>n</sub> may then be released from the substrate-binding groove, like a ball bouncing off a racket. Similar observations were reported for other oligosaccharide-specific SBPs, such as TmCeBP (23) and *E. coli* MBP (24, 25). This common feature of binding grooves in SBPs suggests an energetic advantage of the 'half-open' binding groove on the reducing-end side on the sliding motion of oligosaccharides during sugar translocation.

The structural data obtained from our previous report 16 and current studies together suggest the 'ping-pong' model for (GlcNAc)<sub>2</sub> binding/release as depicted in Figure 9B. Such a

## Structural model of VhCBP in (GlcNAc)<sub>2</sub> translocation

model involves three steps: step 1, the apo-VhCBP in the ‘fully open’ form binds (GlcNAc)<sub>2</sub>, which enters the substrate-binding groove, its binding inducing a ‘fully closed’ conformation of VhCBP. Step 2, the fully closed conformation subsequently undergoes transition to the intermediate ‘half-open’ conformation, and in step 3, (GlcNAc)<sub>2</sub> leaves the binding sites, triggering the return of VhCBP to its original open conformation. The crystal structure of the W513A–(GlcNAc)<sub>2</sub> complex observed in the present report is the first experimental evidence capturing an intermediary conformation between fully closed and fully open conformations in the sugar translocation pathway. Our ITC data suggested the closed (sugar-bound) state has high affinity to the (GlcNAc)<sub>2</sub> substrate ( $K_d$  of 0.03  $\mu$ M), while the semiopen state has 1200 lower affinity than the closed state ( $K_d$  of  $\sim$ 37  $\mu$ M). We propose that this decrease in the binding affinity facilitates the sugar release from the VhCBP to its membrane-transporting partner. In conclusion, the proposed structural model is thought to be a general mechanism in the *Vibrio* system, facilitating the translocation of small chitooligosaccharides from the periplasm to inner membrane-transporting components, with (GlcNAc)<sub>2</sub> being the preferred molecule, allowing the bacteria to utilize chitin as their source of energy.

### Experimental procedures

#### Multiple sequence alignment of SBPs

VhCBP was found to belong to cluster C/subcluster IV (C-IV) (15). Four other SBP members occur in the same cluster/subcluster, including VcCBP from *Vibrio cholera* (PDB code, 4GFR), a TmCeBP (PDB code, 2O71), and TmMnBP1 and TmMnBP2 (PDB codes, 4PFY and 4PFW, respectively). Amino acid sequences of VhCBP (PDB code 5YQW) and the four well-defined sugar-specific SBPs mentioned above were aligned by MUSCLE (29, 30). The secondary structural elements of VhCBP were integrated and displayed on top of the aligned sequences using the program ESPript v. 3.0 (31) (Fig. 1).

#### Construction of the mutated proteins of VhCBP

Mutant plasmids were constructed using the previously generated pET23d(+)-cbp from *V. harveyi* as a DNA template (13). Mutation was introduced into the WT gene using a site-directed mutagenesis kit (Stratagene). Primers used for site-directed mutagenesis are listed in Table S1. Mutation sites are shown in bold, and genetic codons corresponding to the mutated residues are indicated by single underlining. For the double mutant W363A/W513A, the W363A mutation was introduced using the plasmid encoding W513A mutant as a DNA template. Correct mutations were verified by DNA sequencing service (Macrogen, Inc).

#### Expression and purification of WT VhCBP and mutants

The competent cells of *E. coli* Origami (DE3) (Novagen) were transformed with the recombinant plasmids harboring the WT or mutated VhCBP genes. Each transformant was inoculated into the LB medium containing 100  $\mu$ g ml<sup>-1</sup>

ampicillin and shaken at 37 °C until  $OD_{600}$  reached 0.8. Protein expression was induced by the addition of 0.05 mM IPTG to the cell culture that had been precooled on ice for 10 min. The IPTG-induced cell culture was shaken at 18 °C for a further 24 h, after which the cell pellet was harvested by centrifugation at 4500 rpm, 4 °C for 45 min, and the separated culture medium was discarded. The cell pellet was resuspended in 20 mM sodium phosphate buffer, pH 7.4, containing 150 mM NaCl, 10% (v/v) glycerol, 1 mM PMSF, and 0.1% Triton X-100, and the cells in the suspension were disrupted with an ultrasonic processor (Cole-Parmer Instrument Company, LLC). The cell lysate was harvested by centrifugation at 12,000 rpm at 4 °C for 45 min. For protein purification, the crude cell lysate was applied to a Ni-NTA column equilibrated with 20 mM sodium phosphate, pH 7.4, containing 150 mM NaCl, then washed twice with the buffer containing 10 mM imidazole, and eluted with the buffer containing 100 mM imidazole. Eluted protein fractions were collected and diluted three-fold with the same buffer and applied to a HiTrap Q FF column 5 ml (GE Healthcare). The column was washed twice with 20 mM sodium phosphate buffer, pH 7.4, containing 50 mM NaCl and adsorbed protein then eluted with a linear gradient of 50 to 500 mM NaCl. Protein fractions containing VhCBP or the mutants were collected, concentrated by Amicon Ultra-15 (Millipore), and applied to a HiPrep 16/60 Sephacryl S-100 column (GE Healthcare) equilibrated with 20 mM sodium phosphate buffer, pH 7.4, containing 150 mM NaCl. Proteins were isocratically eluted and fractions collected and analyzed by SDS-PAGE. Fractions containing VhCBP or the mutant were pooled and stored at 4 °C for further analysis. Protein concentration was determined from absorbance at 280 nm, using a molar extinction coefficient calculated from the equation proposed by Pace *et al.* (32).

#### ITC binding experiments

ITC experiments were conducted by titrating 200  $\mu$ M (GlcNAc)<sub>n</sub> (n = 2, 3, or 4) solutions from the injection syringe into 20  $\mu$ M WT VhCBP in the sample cell. The titrations were performed in 20 mM Tris-HCl buffer, pH 8.0, by a PEAQ-ITC (Malvern Instruments) using the operational software installed in the equipment. Experimental conditions are summarized in Table S2. Briefly, the temperature in the sample cell was maintained at 25 °C, and the difference in temperature ( $\Delta T$ ) between the reference cell and the sample cell was compensated by heaters fitted to the cells. The compensation energy utilized per second was plotted against experimental time. Peak integrations of the individual troughs were replotted against the molar ratio of the ligand to protein. The binding isotherm thus obtained was used for data-fitting analysis using the ‘One-Set of Sites’ model defined in MicroCal PEAQ-ITC Analysis Software to obtain  $K_d$ ,  $\Delta G^\circ$ ,  $\Delta H^\circ$ , and  $\Delta S^\circ$  terms. Stoichiometry  $n$  was fixed at 1.0 because the crystal structure of VhCBP in complex with (GlcNAc)<sub>n</sub> clearly indicated a 1:1 (protein:ligand) interaction. Titrations of (GlcNAc)<sub>2</sub> to the individual VhCBP mutants were performed as described for the WT protein. Because affinity and heat generation changed

depending on the protein used, ligand and protein concentrations used for the experiments were varied (Table S2) to fulfill the criteria proposed by Turnbull and Daranas (33).

### Crystallization and structure determination of VhCBP mutants in complex with (GlcNAc)<sub>2</sub>

Crystallization conditions for individual mutants were screened using Morpheus (Molecular Dimensions). Protein solution (10 mg ml<sup>-1</sup>) was freshly prepared in 10 mM Hepes buffer, pH 7.0, containing 100 mM NaCl and mixed with an equal volume of the precipitant solution to make a crystallization drop. Crystal screens were set up by a microbatch-under-oil technique using 100% (v/v) silicone oil (Hampton Research) at 20 °C. Single crystals were harvested from crystallization drops by a cryoloop and flash-frozen and stored under liquid nitrogen. VhCBP was successfully cocrystallized with (GlcNAc)<sub>3</sub> and (GlcNAc)<sub>4</sub>. Although cocrystallization with (GlcNAc)<sub>2</sub> was attempted with all the mutants produced, only five mutants, W363A, N409A, F411A, F437A, and W513A, were cocrystallized with (GlcNAc)<sub>2</sub> under the conditions described in Table S3. With the other mutants, crystallization was unsuccessful.

X-ray diffraction experiments were performed at the beamline, TPS-05A, in NSRRC (National Synchrotron Radiation Research Center, Taiwan). An automated system controlled by the software installed in the Linux machines at the beamline handled the crystals during experiments. Briefly, crystals were mounted onto the goniometer automatically by a robotic arm and irradiated with X-ray at  $\lambda = 0.999840$  Å, and the diffraction images were collected by MX300-HS (Rayonix, LLC), integral to the beamline. Diffraction images were analyzed by HKL2000 (34). The resulting intensity file was transformed to structural factor file and phased by molecular replacement using the complex structure of VhCBP and (GlcNAc)<sub>2</sub> (PDB code, 5YQW) as the template by PHASER MR in CCP4i program suite v. 7.0 (35, 36). The modeled structures were refined by a combination of Coot (37) and Phenix.Refine (38–41) in Phenix 1.9 (42). The statistics of data collection and structural refinement are shown in Table 2. The structures were displayed by PyMOL Molecular Graphics System v. 2.4.0. Structures were successfully solved with the highest resolutions below 2.2 Å and deposited in the PDB database under the PDB codes 6LZQ for VhCBP–(GlcNAc)<sub>3</sub>, 7EBI for VhCBP–(GlcNAc)<sub>4</sub>, 7EBM for W363A–(GlcNAc)<sub>2</sub>, 6LZT for N409A–(GlcNAc)<sub>2</sub>, 6LZU for F411A–(GlcNAc)<sub>2</sub>, 6LZV for F437A–(GlcNAc)<sub>2</sub>, and 6LZW for W513A–(GlcNAc)<sub>2</sub>.

### Docking of (GlcNAc)<sub>4</sub> into VhCBP

Because the electron density of two sugar residues on the reducing-end side of bound (GlcNAc)<sub>4</sub> was not observed, (GlcNAc)<sub>4</sub> was modeled into the VhCBP structure extracted from the VhCBP–(GlcNAc)<sub>4</sub> complex (7EBI), generating the most plausible structure of the VhCBP–(GlcNAc)<sub>4</sub> complex. AutoDock Vina was used to simulate the complex structure (43, 44). The ligand was prepared using GLYCAM and

PRODRG (45, 46). The ligand was docked into the VhCBP molecule using Iterated Local Search global optimizer in AutoDock Vina. Conformers were searched within a grid box with 40, 40, and 30 Å for each dimension covering the entire binding pocket of VhCBP. Finally, the conformation was converged to nine conformers with the energies calculated, -10.9, -10.5, -10.1, -9.3, -9.0, -8.8, -8.6, -8.5, and -8.5 kcal mol<sup>-1</sup>, respectively, from the most stable to unstable conformer.

### Data availability

The atomic co-ordinates and structural factor data of VhCBP have been deposited in the PDB database under the accession codes 6LZQ for WT VhCBP–(GlcNAc)<sub>3</sub>, 7EBI for WT VhCBP–(GlcNAc)<sub>4</sub>, 7EBM for W363A–(GlcNAc)<sub>2</sub>, 6LZT for N409A–(GlcNAc)<sub>2</sub>, 6LZU for F411A–(GlcNAc)<sub>2</sub>, 6LZV for F437A–(GlcNAc)<sub>2</sub>, and 6LZW for W513A–(GlcNAc)<sub>2</sub>. The ITC data are available upon request to Wipa Suginta ([wipa.s@vistec.ac.th](mailto:wipa.s@vistec.ac.th)).

**Supporting information**—This article contains supporting information with supplementary experiments (Figs. S1, and S2) and full wwPDB X-ray structure validation reports for all the crystal structures presented in this study.

**Acknowledgments**—We thank Linh Thuy Tran for X-ray data collection and the technical services provided by the Synchrotron Radiation Protein Crystallography Facility of the National Core Facility Program for Biotechnology, Ministry of Science and Technology, and the National Synchrotron Radiation Research Center (NSRRC), Taiwan. We also thank the beamline TPS-05A staff who supported in X-ray diffraction data collection and data processing. We greatly appreciate Dr David Apps, the University of Edinburgh Medical School, U.K., for critical reading and English improvement of the manuscript. This research was supported by Vidyasirimedhi Institute of Science and Technology (300/111100/1711111000030) and Thailand Research Fund through The Basic Research Grant (BRG610008). This grant is partially supported by Thailand Science Research and Innovation through Global Partnership Grant (Contract no: PMUB-P5-63-B20PIC\_WIS\_CHU-PMB010).

**Author contributions**—Y. K., S. K., P. U., and R. C. R. data curation; Y. K. formal analysis; Y. K., R. C. R., and W. S. validation; Y. K. and T. F. investigation; Y. K., S. K., and W. S. visualization; Y. K., S. K., P. U., and R. C. R. methodology; Y. K., T. F., and W. S. writing—original draft; T. F., R. C. R., and W. S. supervision; T. F., R. C. R., and W. S. writing—review and editing; R. C. R. resources; R. C. R. software; W. S. conceptualization; W. S. funding acquisition; W. S. project administration.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: CBP, chitoooligosaccharide-binding protein; ChiP, chitoporin; ITC, isothermal titration calorimetry; SBP, solute-binding protein; TmCeBP, cellobiosaccharide-specific SBP from *Thermotoga maritima*; TmMnBP, mannooligosaccharide-specific SBP from *T. maritima*;

## Structural model of VhCBP in (GlcNAc)<sub>2</sub> translocation

VcCBP, CBP from *Vibrio cholerae*; VhCBP, CBP from *Vibrio campbellii* (formerly *V. harveyi*) type strain ATCC BAA-1116.

### References

1. Austin, B., and Zhang, X. H. (2006) *Vibrio harveyi*: A significant pathogen of marine vertebrates and invertebrates. *Let. Appl. Microbiol.* **43**, 119–124
2. Zhang, X., Lin, H., Wang, X., and Austin, B. (2018) Significance of *Vibrio* species in the marine organic carbon cycle—a review. *Sci. China Earth Sci.* **61**, 1357–1368
3. Meibom, K. L., Li, X. B., Nielsen, A. T., Wu, C., Roseman, S., and Schoolnik, G. K. (2004) The *Vibrio cholerae* chitin utilization program. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 2524–2529
4. Bassler, B. L., Yu, C., Lee, Y. C., and Roseman, S. (1991) Chitin utilization by marine bacteria: Degradation and catabolism of chitin oligosaccharides by *Vibrio furnissii*. *J. Biol. Chem.* **266**, 24276–24286
5. Li, X., and Roseman, S. (2004) The chitinolytic cascade in *Vibrios* is regulated by chitin oligosaccharides and a two-component chitin catabolic sensor/kinase. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 627–631
6. Suginta, W., Robertson, P. A. W., Austin, B., Fry, S. C., and Fothergill-Gilmore, L. A. (2000) Chitinases from *Vibrio*: Activity screening and purification of chiA from *Vibrio carchariae*. *J. Appl. Microbiol.* **89**, 76–84
7. Chitlaru, E., and Roseman, S. (1996) Molecular cloning and characterization of a novel  $\beta$ -N-acetyl-D-glucosaminidase from *Vibrio furnissii*. *J. Biol. Chem.* **271**, 33433–33439
8. Suginta, W., Chuenark, D., Mizuhara, M., and Fukamizo, T. (2010) Novel  $\beta$ -N-acetylglucosaminidases from *Vibrio harveyi* 650: Cloning, expression, enzymatic properties, and subsite identification. *BMC Biochem.* **11**, 40
9. Li, X., Wang, L. X., Wang, X., and Roseman, S. (2007) The chitin catabolic cascade in the marine bacterium *Vibrio cholerae*: Characterization of a unique chitin oligosaccharide deacetylase. *Glycobiology* **17**, 1377–1387
10. Suginta, W., Chumjan, W., Mahendran, K. R., Janning, P., Schulte, A., and Winterhalter, M. (2013) Molecular uptake of chitooligosaccharides through chitoporin from the marine bacterium *Vibrio harveyi*. *PLoS One* **8**, e55126
11. Suginta, W., Chumjan, W., Mahendran, K. R., Schulte, A., and Winterhalter, M. (2013) Chitoporin from *Vibrio harveyi*, a channel with exceptional sugar specificity. *J. Biol. Chem.* **288**, 11038–11046
12. Soysa, H. S. M., Aunkham, A., Schulte, A., and Suginta, W. (2020) Single-channel properties, sugar specificity, and role of chitoporin in adaptive survival of *Vibrio cholerae* type strain O1. *J. Biol. Chem.* **295**, 9421–9432
13. Suginta, W., Sritho, N., Ranok, A., Bulmer, D. M., Kitaoku, Y., van den Berg, B., and Fukamizo, T. (2018) Structure and function of a novel periplasmic chitooligosaccharide-binding protein from marine *Vibrio* bacteria. *J. Biol. Chem.* **293**, 5150–5159
14. Scheepers, G. H., Lycklama, J. A., and Poolman, B. (2016) An updated structural classification of substrate-binding proteins. *FEBS Lett.* **590**, 4393–4401
15. Fukamizo, T., Kitaoku, Y., and Suginta, W. (2019) Periplasmic solute-binding proteins: Structure classification and chitooligosaccharide recognition. *Int. J. Biol. Macromol.* **128**, 985–993
16. Wilkinson, A. J., and Verschuere, K. H. G. (2003). In: Holland, I. B., Cole, S. P. C., Kuchler, K., Higgins, C. F., eds. *ABC Proteins from Bacteria to Man*, Academic Press, London
17. Klancher, C. A., Yamamoto, S., Dalia, T. N., and Dalia, A. B. (2020) ChiS is a noncanonical DNA-binding hybrid sensor kinase that directly regulates the chitin utilization program in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 20180–20189
18. Keyhani, N. O., Li, X. B., and Roseman, S. (2000) Chitin catabolism in the marine bacterium *Vibrio furnissii*. Identification and molecular cloning of a chitoporin. *J. Biol. Chem.* **275**, 33068–33076
19. Chumjan, W., Winterhalter, M., Schulte, A., Benz, R., and Suginta, W. (2015) Chitoporin from the marine bacterium *Vibrio harveyi*: Probing the essential roles of Trp<sup>136</sup> at the surface of the constriction zone. *J. Biol. Chem.* **290**, 19184–19196
20. Aunkham, A., Zahn, M., Kesireddy, A., Pothula, K. R., Schulte, A., Baslé, A., Kleinekathöfer, U., Suginta, W., and van den Berg, B. (2018) Structural basis for chitin acquisition by marine *Vibrio* species. *Nat. Commun.* **9**, 220
21. Franca, E. F., Lins, R. D., Freitas, L. C. G., and Straatsma, T. P. (2008) Characterization of chitin and chitosan molecular structure in aqueous solution. *J. Chem. Theory Comput.* **4**, 2141–2149
22. Ohnuma, T., Sørli, M., Fukuda, T., Kawamoto, N., Taira, T., and Fukamizo, T. (2011) Chitin oligosaccharide binding to a family GH19 chitinase from the moss *Bryum coronatum*. *FEBS J.* **278**, 3991–4001
23. Cuneo, M. J., Beese, L. S., and Hellings, H. W. (2009) Structural analysis of semi-specific oligosaccharide recognition by a cellulose-binding protein of *Thermotoga maritima* reveals adaptations for functional diversification of the oligopeptide periplasmic binding protein fold. *J. Biol. Chem.* **284**, 33217–33223
24. Ferenci, T. (1980) The recognition of maltodextrins by *Escherichia coli*. *Eur. J. Biochem.* **108**, 631–636
25. Spurlino, J. C., Lu, G. Y., and Quijcho, F. A. (1991) The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266**, 5202–5219
26. Suginta, W., Vongsuwan, A., Songsiririthigul, C., Svasti, J., and Prinz, H. (2005) Enzymatic properties of wild-type and active site mutants of chitinase A from *Vibrio carchariae*, as revealed by HPLC-MS. *FEBS J.* **272**, 3376–3386
27. Songsiririthigul, C., Pantoom, S., Aguda, A. H., Robinson, R. C., and Suginta, W. (2008) Crystal structures of *Vibrio harveyi* chitinase A complexed with chitooligosaccharides: Implications for the catalytic mechanism. *J. Struct. Biol.* **162**, 491–499
28. Kaneko, A., Uenishi, K., Maruyama, Y., Mizuno, N., Baba, S., Kumasaka, T., Mikami, B., Murata, K., and Hashimoto, W. (2017) A solute-binding protein in the closed conformation induces ATP hydrolysis in a bacterial ATP-binding cassette transporter involved in the import of alginate. *J. Biol. Chem.* **292**, 15681–15690
29. Maruyama, Y., Itoh, T., Kaneko, A., Nishitani, Y., Mikami, B., Hashimoto, W., and Murata, K. (2015) Structure of a bacterial ABC transporter involved in the import of an acidic polysaccharide alginate. *Structure* **23**, 1643–1654
30. Edgar, R. C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797
31. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324
32. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423
33. Turnbull, W. B., and Daranas, A. H. (2003) On the value of c: Can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **125**, 14859–14866
34. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
35. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
36. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 1131–1137
37. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
38. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
39. Headd, J. J., Echols, N., Afonine, P. V., Grosse-Kunstleve, R. W., Chen, V. B., Moriarty, N. W., Richardson, D. C., Richardson, J. S., and Adams, P. D. (2012) Use of knowledge-based restraints in phenix.refine to improve

- macromolecular refinement at low resolution. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 381–390
40. Afonine, P. V., Grosse-Kunstleve, R. W., Adams, P. D., and Urzhumtsev, A. (2013) Bulk-solvent and overall scaling revisited: Faster calculations, improved results. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 625–634
  41. Afonine, P. V., Grosse-Kunstleve, R. W., Urzhumtsev, A., and Adams, P. D. (2009) Automatic multiple-zone rigid-body refinement with a large convergence radius. *J. Appl. Crystallogr.* **42**, 607–615
  42. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877
  43. Trott, O., and Olson, A. J. (2009) AutoDock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461
  44. Morris, G. M., Ruth, H., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791
  45. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) GLYCAM06: A generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **29**, 622–655
  46. van Aalten, D. M., Bywater, R., Findlay, J. B., Hendlich, M., Hooft, R. W., and Vriend, G. (1996) PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput. Aided. Mol. Des.* **10**, 255–262