



OPEN Periplasmic chitooligosaccharide-binding protein requires a three-domain organization for substrate translocation

Takayuki Ohnuma^{1,2}, Jun Tsujii¹, Chikara Kataoka¹, Teruki Yoshimoto¹, Daijiro Takeshita³, Outi Lampela⁴, André H. Juffer^{4,5}, Wipa Suginta⁶ & Tamo Fukamizo^{1,6}

Periplasmic solute-binding proteins (SBPs) specific for chitooligosaccharides, (GlcNAc)_n (n = 2, 3, 4, 5 and 6), are involved in the uptake of chitinous nutrients and the negative control of chitin signal transduction in *Vibrios*. Most translocation processes by SBPs across the inner membrane have been explained thus far by two-domain open/closed mechanism. Here we propose three-domain mechanism of the (GlcNAc)_n translocation based on experiments using a recombinant VcCBP, SBP specific for (GlcNAc)_n from *Vibrio cholerae*. X-ray crystal structures of unliganded or (GlcNAc)₃-liganded VcCBP solved at 1.2–1.6 Å revealed three distinct domains, the Upper1, Upper2 and Lower domains for this protein. Molecular dynamics simulation indicated that the motions of the three domains are independent and that in the (GlcNAc)₃-liganded state the Upper2/Lower interface fluctuated more intensively, compared to the Upper1/Lower interface. The Upper1/Lower interface bound two GlcNAc residues tightly, while the Upper2/Lower interface appeared to loosen and release the bound sugar molecule. The three-domain mechanism proposed here was fully supported by binding data obtained by thermal unfolding experiments and ITC, and may be applicable to other translocation systems involving SBPs belonging to the same cluster.

Abbreviations

SBP	Periplasmic solute-binding protein
CBP	SBP specific to chitooligosaccharide
VcCBP	CBP from <i>Vibrio cholerae</i>
VhCBP	CBP from <i>Vibrio campbellii</i> (formerly, <i>Vibrio harveyi</i>)
TmCBP	SBP specific to cellobiose from <i>Thermotoga maritima</i>
GlcNAc	N-Acetylglucosamine
(GlcNAc) _n	β-1,4-Linked oligosaccharide of GlcNAc
CD	Circular dichroism
ITC	Isothermal titration calorimetry
RMSD	Root mean square deviations
ASA	Solvent-accessible surface area

¹Department of Advanced Bioscience, Kindai University, 3327-204 Nakamachi, Nara 631-8505, Japan. ²Agricultural Technology and Innovation Research Institute (ATIRI), Kindai University, 3327-204, Nakamachi, Nara 631-8505, Japan. ³Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba-Shi, Ibaraki 305-8566, Japan. ⁴Biocenter Oulu, University of Oulu, P.O. Box 5000, FI-90014 Oulu, Finland. ⁵Faculty of Biochemistry and Molecular Medicine, University of Oulu, P.O.Box 5000, FI-90014 Oulu, Finland. ⁶School of Biomolecular Science & Engineering, Vidyasirimedhi Institute of Science and Technology (VISTEC), Wangchan Valley 555 Moo 1 Payupnai, Wangchan, Rayong 21210, Thailand. ✉email: ohnumat@nara.kindai.ac.jp; tamo0111fuka@gmail.com

Chitin is a β -1,4-linked polysaccharide of *N*-acetylglucosamine (GlcNAc), and is degraded by chitinases¹, β -*N*-acetylglucosaminidases², chitin deacetylases³ and lytic polysaccharide monoxygenases⁴, to produce GlcNAc, chitin oligosaccharides, (GlcNAc)_n (n = 2, 3, 4, 5, and 6) and their de-*N*-acetylated derivatives. The products are utilized as structural components in various chitin-containing organisms⁵ or as carbon and nitrogen sources by chitin consumers⁶. *Vibrio* spp. live in the marine ecosystem and possess all these enzymes, enabling them to efficiently catabolize chitin as their nutritional source⁷. The enzymatic degradation of chitin by extracellular chitinases in *Vibrios* produces (GlcNAc)_n^{8–10}, which are taken up into the periplasm through a chitoporin localized in the outer membrane^{11–14}. (GlcNAc)_n transported into the periplasm undergo further enzymatic degradation^{15–17} and are trapped by a periplasmic solute-binding protein (SBP) specific for (GlcNAc)_n¹⁸. Various SBPs have been recognized to mediate the transport of the corresponding solutes into the cytoplasm through a specific transporter localized in the inner membrane¹⁹; furthermore, they interact with signal receptor proteins to control the signal response²⁰. SBPs specific to (GlcNAc)_n (referred to as CBP) also have such a dual functionality; (GlcNAc)_n-liganded or unliganded CBPs interact with a chitin sensing protein as well as a (GlcNAc)₂-specific ABC-transporter^{21,22}. A large conformational change of CBP from “open” to “closed”, or vice versa, induced by binding/release of (GlcNAc)_n¹⁹ triggers a series of interactions involving sensing or transporter proteins. It is therefore highly desirable to thoroughly analyze the structure, binding mechanism of CBP/(GlcNAc)_n and the CBP/(GlcNAc)_n/membrane protein interactions.

Sheepers et al.²³ classified SBPs into seven clusters, A, B, C, D, E, F and G, according to their crystal structures, and cluster C was subdivided into five subclusters (I, II, III, IV and V) in our previous review²⁴. Two CBPs from *Vibrios* have been investigated so far. One is from *V. cholerae* (VcCBP), a facultative anaerobe that causes a deadly disease to humans called cholera²⁵, and the second is from *V. campbelli*, (formerly *V. harveyi*, VhCBP), a bioluminescence marine bacterium that causes a fatal disease to aquatic animal called Vibriosis²⁶. Both VcCBP and VhCBP belong to cluster C/subcluster IV (C-IV), in which SBPs are specific for oligosaccharides, including (GlcNAc)_n, mannoooligosaccharides²⁷, and celloooligosaccharides²⁸. The first crystal structures of CBP were solved for VcCBP and deposited in the PDB database in 2006; PDB codes 1ZTY for unliganded (open) VcCBP and 1ZU0 for (GlcNAc)₂-liganded (closed) VcCBP, and in 2013; PDB codes, 4GF8 for unliganded (open) VcCBP and 4GFR for (GlcNAc)₂-liganded (closed) VcCBP. However, no functional or mechanistic information on this CBP was reported. Suginta et al.²⁹ solved the crystal structure of a closed form of (GlcNA)₂-liganded VhCBP, with an amino acid sequence that shares 83% identity with that of VcCBP as shown in Fig. 1. They regarded this protein as comprising two domains (the Upper and Lower domains) connected by a flexible hinge region, which enabled CBP to embrace (GlcNAc)_n in its substrate-binding groove lying between the two domains. Conserved Asp365, Phe437, Trp363 and Trp513 (Fig. 1) localized in the binding groove were found to contribute strongly to the interaction with (GlcNAc)₂³⁰. Most other studies on SBPs proposed that the proteins adopt a two-domain conformation, which is converted from the open state to the closed state upon capture of their substrates^{31–33}. Kitaoku et al.³⁰ found a suggestive conformation, “half-open” form, in the crystal structure of mutant VhCBP (W513A), which provided some clues for the elucidation of the mechanism of (GlcNAc)_n translocation by CBP. Moreover, a three-domain organization was proposed for the SBPs belonging to cluster C and referred to as domains I, II, and III¹⁹. Further studies are needed to gain more extensive insights into the domain organization of CBPs.

In this study, we conducted in-depth investigations of the structure and mechanism of CBP by means of crystallography, molecular dynamics simulation, thermal unfolding experiments and isothermal titration calorimetry (ITC), using VcCBP and (GlcNAc)_n (n = 2, 3, 4, 5 and 6). Based on the crystal structures and molecular dynamics simulation, we here propose that the three domains, Upper1, Upper2, and Lower domains, which correspond to domains I, II, and III, respectively¹⁹, are required to explain the translocation process of (GlcNAc)₂ by CBP. This proposal was well supported by the thermodynamic data for (GlcNAc)_n binding to VcCBP.

Results

Crystal structures of VcCBP

We solved the structure of unliganded protein (open form) at 1.6 Å (Fig. 2A), a higher resolution than the VcCBP structures formerly deposited (2.2–2.3 Å resolution). The closed structure of (GlcNAc)₃-liganded VcCBP was also solved at 1.22 Å resolution (Fig. 2B). The coordinates were deposited in the database with the PDB codes 8I5J and 8I5K, respectively. The statistics for the crystallization and refinement procedures are listed in Table 1. Superimposition of the unliganded structure (open form) obtained here with that of 1ZTY previously deposited revealed that these two structures are almost identical, with a root-mean-square deviation (RMSD) of 0.600 Å, although the space group of 8I5J (P2₁) was different from that of 1ZTY (P3₂21). The RMSD values obtained from the superimpositions of (GlcNAc)₃-liganded structure (8I5K) with those of (GlcNAc)₂-liganded VcCBP (1ZU0) and (GlcNAc)₃-liganded VhCBP (6LZQ³⁰) were 0.529 Å and 0.393 Å, respectively. Both structures obtained here for VcCBP showed three distinct domains, the Upper1, Upper2 and Lower domains (Fig. 2A,B). Compared with the unliganded structure, the angle of closure between Upper1 and Lower domains of the (GlcNAc)₃-liganded structure was found to be 54.7°. This was determined from the rotation angle of the Gln490 α -carbon atom (Upper1 domain) using the Ala486 α -carbon, which is located at the hinge, as the fixed rotation center. The Upper1 region is colored in deep green, Upper2 magenta, and Lower cyan. Upper1 (1–35, 140–239, and 489-C-terminal) comprises α 3- α 5/ α 16, β 1/ β 8- β 12/ β 19- β 21 and η 2- η 3, while Upper2 (36–139) comprises α 1- α 2, β 2- β 7 and η 1. The Lower domain (240–488) comprises α 6- α 15, β 13- β 18 and η 4- η 6 (Figs. 1, 2A,B). The division of the upper structural region into Upper1 and Upper2 domains appeared to be reasonable from the superimposition of (GlcNAc)₂-liganded (1ZU0) and (GlcNAc)₃-liganded VcCBP (8I5K) (Fig. 2C), where the Lower domain fully overlapped (RMSD, 0.306 Å), while the Upper1 domain partly deviated (RMSD, 1.203 Å)

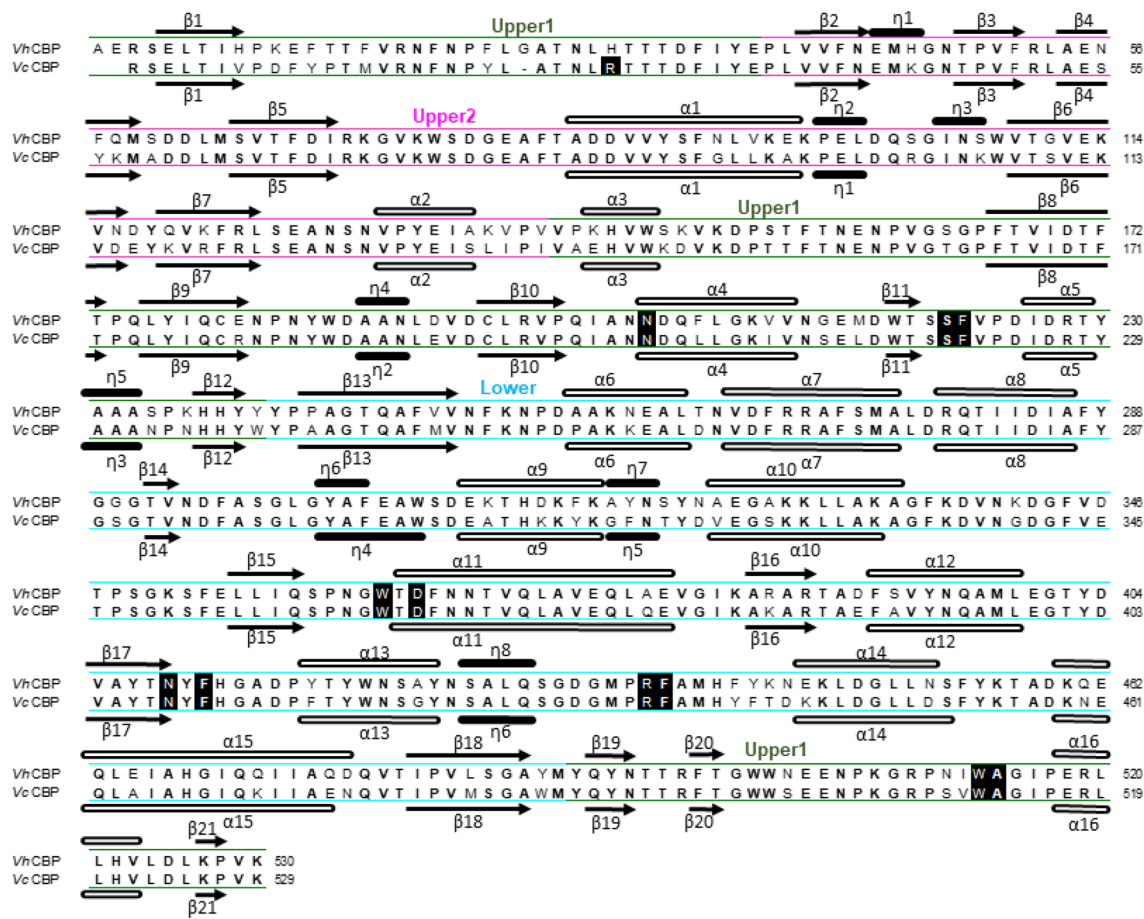


Figure 1. Sequence alignment of the periplasmic solute-binding proteins specific to (GlcNAc)₂ from *Vibrios*. Identical amino acids are written in bold. The secondary structures are designated α1, α2, α3, ... (α-helices, open bars), β1, β2, β3, ... (β-strands, arrows), and η1, η2, η3, ... (3₁₀-helices, filled bars) from the N-terminus. The sequence of Upper1 domain is underlined (VcCBP) or overlined (VhCBP) in dark green, Upper2 domain in magenta, and Lower domain in cyan. Amino acid residues, which interact with the ligand and are conserved between VhCBP and VcCBP, are highlighted on a black background.

and the Upper2 domain strongly deviated (RMSD, 1.930 Å). The structure liganded with (GlcNAc)₃ differed slightly from that with (GlcNAc)₂ especially in the Upper2 domain. Accommodating the reducing-end GlcNAc of the bound (GlcNAc)₃ appeared to affect more strongly the Upper2 domain.

Binding mode of (GlcNAc)₃

Figure 3A shows a 2F₀-F_c map of the bound (GlcNAc)₃. As in the case of VhCBP²⁹, electron density of (GlcNAc)₃ was identified in the (GlcNAc)₃-liganded VcCBP. Since Kitaoku et al.³⁰ designated these subsites as Site1, Site2 and Site3, the same subsite nomenclature was used in this study. In the (GlcNAc)₃-liganded VcCBP, the electron density of the Site3 GlcNAc was lower than those of the two GlcNAc residues at Site1 and Site2, and the occupancy of Site3 GlcNAc was set at 0.5. Two conformers of the Phe410 side chain were observed near the Site3 GlcNAc, and the individual occupancies of the side chain were also set at 0.5. It is most likely that the mobility of the Site3 GlcNAc is much higher than that of the other GlcNAc residues and that the Phe410 side chain flips alternatively back and forth in the liganded structure. Both GlcNAc units at Site1 and Site2 appeared to be hydrogen-bonded or in hydrophobic contact with Asp9, Asn203, Ser220, Phe221, Trp362, Asp364, Asn408, Phe410, Arg435, Phe436 and Trp512, of which all except Asp9 are conserved between VhCBP and VcCBP (highlighted on a black background in Figs. 1 and 3B), while the GlcNAc unit at Site3 makes only a couple of hydrogen bonds with the main chain carbonyl of Ala513 and the guanidyl nitrogen of Arg27, which are localized to the interface between the Upper1 and Upper2 domains (Fig. 3B, Supplementary Table S1).

B-factor values visualized in the main chain structure of (GlcNAc)₃-liganded VcCBP are shown in Fig. 4A, and were higher in the Upper2/Lower interface as well as the Site3-GlcNAc contact surface of Upper1 domain. Notably, the B-factors were significantly higher at the loop region immediately following the secondary structure η6 and the loop between α14 and α15 (red broken circles, Fig. 4A). These two loops are located at the Upper2-Lower interface. Furthermore, B-factor values of the individual GlcNAc residues were 9.86, 12.67 and 16.28 Å², respectively (Fig. 4D), indicating the higher mobility of the Site3 GlcNAc. These B-factor data were fully

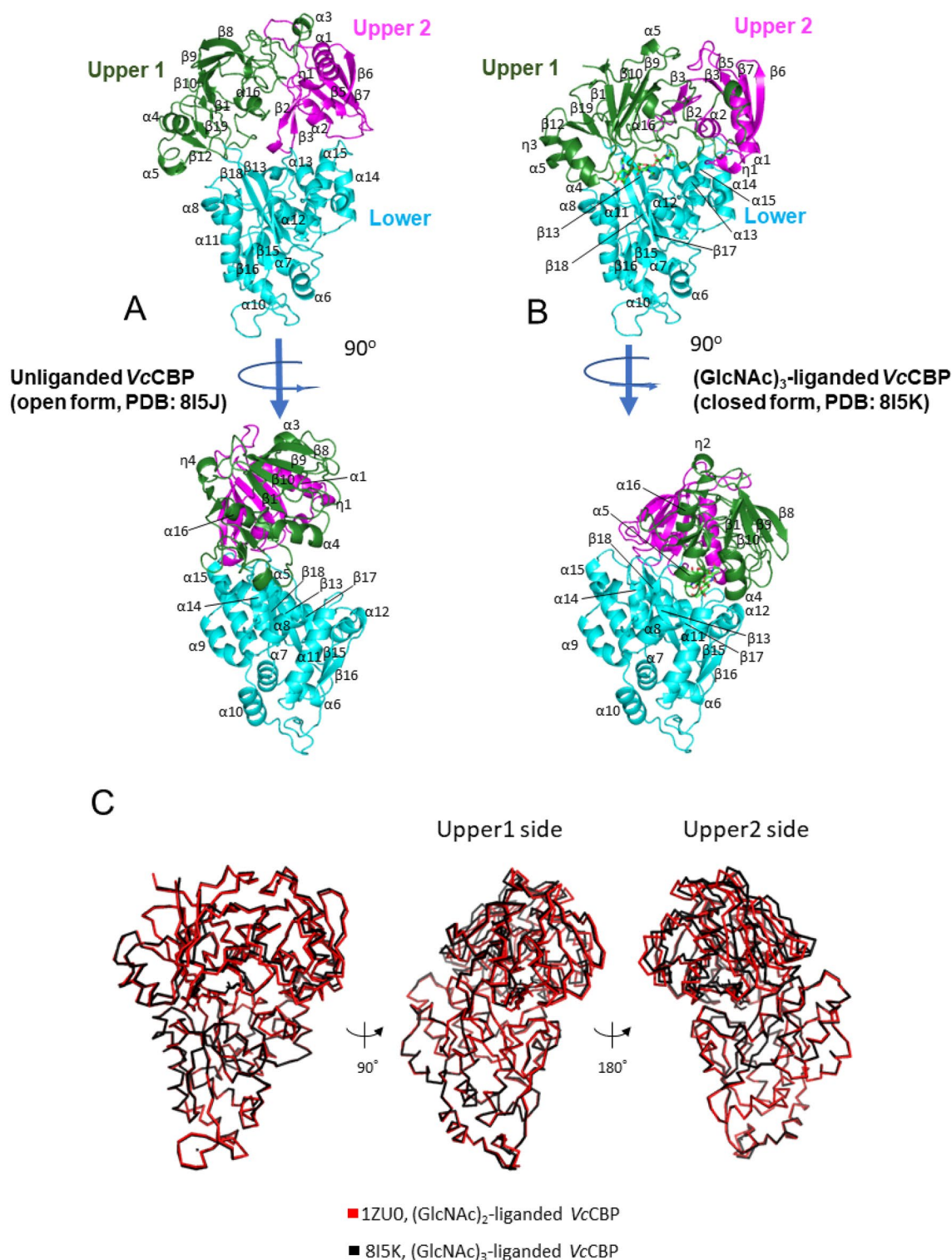


Figure 2. Crystal structures of unliganded VcCBP (A, open state; 8I5J) and (GlcNAc)₃-liganded VcCBP (B, closed state; 8I5K). The upper panels represent views from the front, and the lower panels show side views. Upper1 domain is colored dark green, Upper2 magenta, and Lower cyan. Bound (GlcNAc)₃ is represented by stick model colored light green. The secondary structures are designated as $\alpha 1, \alpha 2, \alpha 3, \dots, \beta 1, \beta 2, \beta 3, \dots$, and $\eta 1, \eta 2, \eta 3, \dots$, from the N-terminus, and correspond to those shown in Fig. 1. (C) Superimposition of the structures of (GlcNAc)₃-liganded (8I5K, black) and (GlcNAc)₂-liganded VcCBP (1ZU0, red). Left, view from the front of the binding groove; middle, view from the Upper1 side; right, view from the Upper2 side.

	Unliganded VcCBP	(GlcNAc) ₃ -liganded VcCBP
Data collection		
Wavelength (Å)	0.9800	0.9800
Space group	<i>P</i> 2 ₁	<i>C</i> 2
Unit-cell parameters		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	<i>a</i> = 81.12, <i>b</i> = 81.64, <i>c</i> = 87.27	<i>a</i> = 135.09, <i>b</i> = 51.68, <i>c</i> = 70.33
α , β , γ (°)	α = 90, β = 104.15, γ = 90	α = 90, β = 98.83, γ = 90
Resolution (Å)	20.0–1.60	20.0–1.22
	(1.69–1.60)	(1.28–1.22)
Completeness (%)	91.6 (61.7)	84.8 (41.0)
<i>R</i> _{sym}	0.097 (0.763)	0.079 (0.771)
<i>I</i> / σ (<i>I</i>)	8.75 (1.56)	8.82 (1.29)
Redundancy	3.3 (2.6)	3.4 (2.8)
<i>CC</i> _{1/2}	0.994 (0.519)	0.996 (0.507)
Refinement		
Resolution (Å)	19.8–1.60	19.3–1.22
No. of reflections	261,323	237,863
<i>R</i> _{work} / <i>R</i> _{free} (%)	16.2/19.1	16.9/18.7
No. of atoms		
Protein	8475	4244
Ligand		44
Ion	2	1
Water	773	563
<i>B</i> -factors (Å ²)		
Protein	22.6	17.4
Ligand		13.0
Ion	25.5	14.9
Water	34.2	28.5
R.m.s.d's		
Bond lengths (Å)	0.009	0.007
Bond angles (°)	0.97	0.94
Ramachandran plot		
Favored region (%)	96.7	97.3
Allowed region (%)	3.3	2.7
Disallowed region (%)	0	0

Table 1. Data-collection and refinement statistics. Values in parentheses are for the highest resolution shell.

consistent with the electron density data described above. The cooperative motions of the Site3 GlcNAc and the Phe410 side chain may be significant from the mechanistic viewpoint.

Molecular dynamics simulation

Figure 5 shows the time-dependent RMSD for unliganded, (GlcNAc)₂-liganded and (GlcNAc)₃-liganded VcCBP. The unliganded form exhibited the largest motion (RMSD, 0.8 nm; the top panel of Fig. 5A) of the entire protein molecule, while the motions of the individual domains remained at lower levels (RMSD, 0.1–0.3 nm; Fig. 5A). This clearly indicated a domain motion, in which the individual domains did not change their own conformations but changed their relative arrangement. On the other hand, the motions of the entire protein molecules as well as individual domains were very low in the (GlcNAc)₂-liganded and (GlcNAc)₃-liganded VcCBP, where RMSDs were only 0.1–0.2 nm (Fig. 5B,C).

Figure 6 shows the distances between the α -carbons of the two numbered amino acid residues (amino acids labeled in red, Fig. 3B). The distance 203–361, which reflects the separation between Upper1 and Lower domains (red broken line between Asn203 and Gly361 in Fig. 3B; the top panels of Fig. 6A,B), did not significantly differ between (GlcNAc)₂-liganded and (GlcNAc)₃-liganded VcCBP. For the distance 24–391 (red broken line between Thr24 and Ala391 in Fig. 3B; second panels from the top of Fig. 6A,B), which also reflects the separation between Upper1 and Lower domains at another site, no significant difference was observed between (GlcNAc)₂-liganded and (GlcNAc)₃-liganded VcCBP. Furthermore, the distance 135–517 (red broken line between Ser135 and Glu517 in Fig. 3B; the bottom panels of Fig. 6A,B) reflecting the separation between the Upper1 and Upper2 domains, did not exhibit any differences between the two liganded VcCBP. However, the distances 101–432 (Gln101–Gly432) and 24–101 (Thr24–Gln101) reflecting the separations of Upper2–Lower and Upper1–Upper2, respectively, fluctuated more in (GlcNAc)₃-liganded VcCBP than in (GlcNAc)₂-liganded (the third and fourth

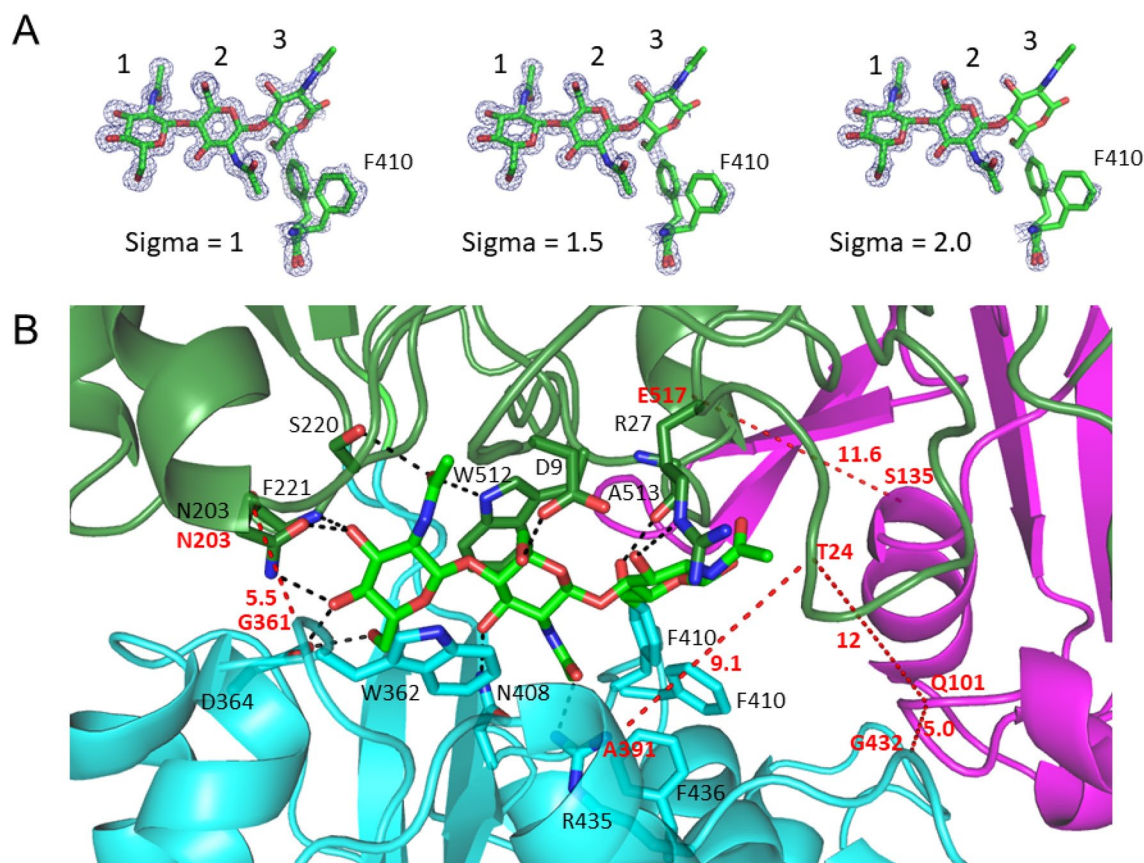


Figure 3. State of (GlcNAc)₃ bound to VcCBP. (A) $2F_0 - F_c$ map of the bound (GlcNAc)₃ and the Phe410 side chain. The occupancies of the Site3 GlcNAc and two conformers of the Phe410 side chain were each assumed to be 0.5. (B) Binding mode. The color system is identical to that in Fig. 2. The possible hydrogen bonds are shown by black broken lines. The red broken lines represent the distances between the α -carbons monitored in the molecular dynamics simulation shown in Fig. 6.

panels from the top of Fig. 6A,B). The movements are more intensive in the (GlcNAc)₃-liganded VcCBP than in the (GlcNAc)₂-liganded in this region. Elongation of the chain of the bound oligosaccharide from (GlcNAc)₂ to (GlcNAc)₃ was found to enhance the molecular motion around amino acid residue Gln101 located in the Upper2 domain. The larger conformational change in (GlcNAc)₃-liganded VcCBP was also confirmed from Fig. 7, which shows the 2-dimensional projection of the trajectories of principal components, eigenvector1 and eigenvector2. In unliganded VcCBP, essential subspaces were widely extended as compared with the liganded VcCBP (Fig. 7A; eigenvector1/eigenvector2, -7 ~ 15 nm/-22 ~ 9 nm) revealing a large conformational change from the open state to the closed state (Fig. 2A,B). Although the conformational spaces of the liganded states were much less extensive (Fig. 7B; -6 ~ 3.5 nm/-2.5 ~ 3.5 nm and Fig. 7C; -4.5 ~ 6.3 nm/-5.5 ~ 4.0 nm), (GlcNAc)₃-liganded VcCBP was found to occupy a larger conformational space than (GlcNAc)₂-liganded, indicating that the structural variations induced by (GlcNAc)₃ binding to Site1, Site2 and Site3 were more abundant than those induced by (GlcNAc)₂ binding to Site1 and Site2.

A cross-correlation heat map was generated based on the Ca-atom positions using the ProDy python package. The results are shown in Fig. 8. In unliganded VcCBP (Fig. 8A), strong self-positive correlations (red) were found in the region of residue No. 1-239, which corresponds to Upper1 + Upper2 domains, and also in residue No. 350-440, which includes amino acid residues interacting with the ligand from Lower domain (Fig. 3B and Supplementary Table S1). However, strong inter-negative correlations (blue) were found between 1 and 239 (Upper1 + Upper2) and 350-440 (the interacting region of Lower domain). This clearly indicated a large domain motion from open to closed, vice versa, shown in Fig. 2A and B. Although the correlations were weaker in (GlcNAc)₃-liganded VcCBP (Fig. 8C) than in the unliganded state, we observed self-positive correlations (red) in residue No. 36-139 (Upper2 domain); however, the inter-correlations between 36 and 139 (Upper2) and 140-239 (inner Upper1) and between 36 and 139 (Upper2) and 1-35 (N-terminal Upper1) were rather negative (blue) as compared with the unliganded state. Furthermore, inter-correlations between 140 and 239 (inner Upper1) and 1-35 (N-terminal Upper1) were rather positive (red). It was clear that Upper1 and Upper2 domains move unitedly in the unliganded state, but that these two domains move independently in the (GlcNAc)₃-liganded state. For the (GlcNAc)₂-liganded VcCBP (Fig. 8B), although correlations were even weaker than (GlcNAc)₃-liganded state, similar trends were observed in the region corresponding to 1-239 (Upper1 + Upper2). Weak but significant inter-negative correlations (blue) between Upper1 + Upper2 (1-239) and the ligand-interacting region of Lower domain (350-470) were also found in the two liganded states (Fig. 8B,C). However, the negative correlations were

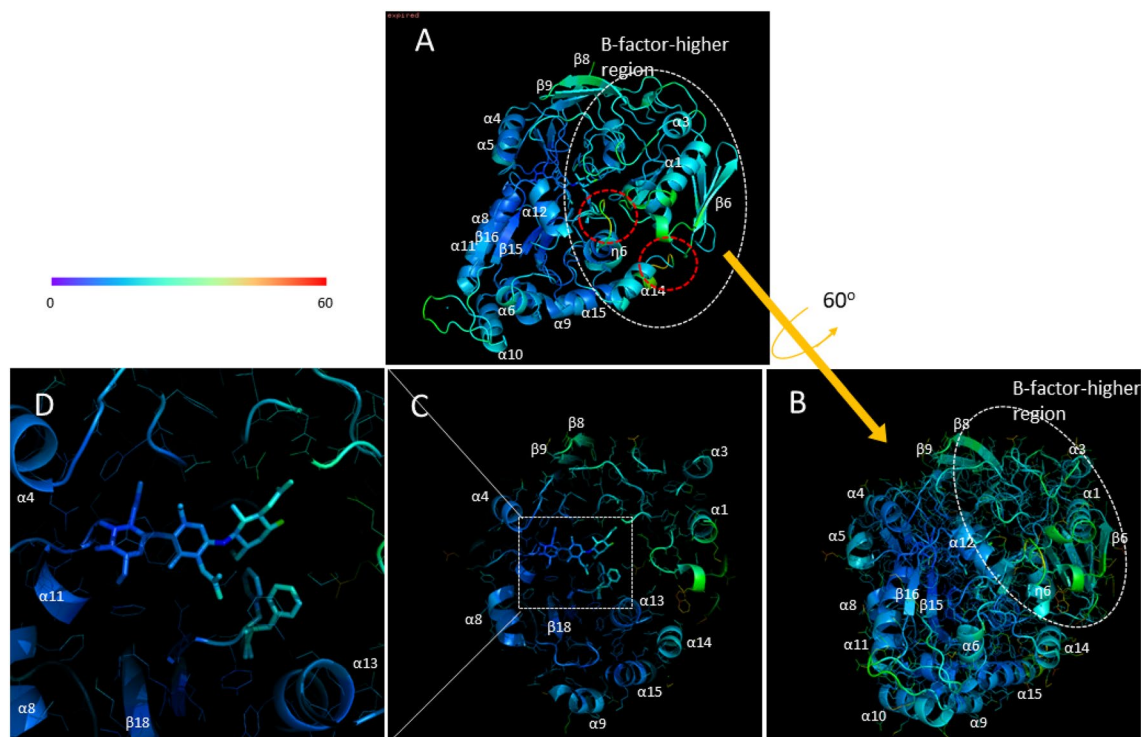


Figure 4. Visualization of crystallographic B-factors for the VcCBP structure. (A) and (B) entire VcCBP structure in complex with (GlcNAc)₃; (C) Cross-section of the (GlcNAc)₃-binding site; and (D), close-up view of the (GlcNAc)₃-binding site. The colors from violet to blue, green, yellow, orange and red indicate B factors from small to large. Broken circles colored white represent B-factor-higher regions, while those colored red represent loop regions with the highest B-factors.

more intensive in the (GlcNAc)₃-liganded state than in the (GlcNAc)₂-liganded, indicating that the movements were enhanced in (GlcNAc)₃-liganded VcCBP (Fig. 8C) as compared with (GlcNAc)₂-liganded state (Fig. 8B).

Thermal unfolding experiments with VcCBP in the presence of (GlcNAc)_n

A typical dataset of the thermal unfolding transitions of VcCBP in the absence or presence of (GlcNAc)_n was shown in Supplementary Fig. S1. The unfolding transition was highly cooperative, with the fraction unfolded increasing sharply at the transition temperature (T_m). The individual unfolding experiments were conducted twice, and the averaged values of T_m were calculated as listed in Table 2. Thermal unfolding of unliganded VcCBP took place at T_m of 43.2 °C. When (GlcNAc)₂, (GlcNAc)₃ or (GlcNAc)₄ was added to VcCBP, T_m was elevated from 43.2 to 54.4 °C ($\Delta T_m = 11.2$ °C), 54.8 °C ($\Delta T_m = 11.6$ °C) and 53.8 °C ($\Delta T_m = 10.6$ °C), respectively. As in the case of VhCBP²⁹, these three oligosaccharides bound strongly to VcCBP. However, ΔT_m values induced by the addition of (GlcNAc)₅ and (GlcNAc)₆ were moderate, 6.3 °C for both, indicating that these two oligosaccharides bind to VcCBP with significantly lower affinities³⁴. ΔT_m for GlcNAc was only 2.8 °C, suggesting that the monomer bound to VcCBP with much lower affinity.

Thermodynamic parameters for binding of (GlcNAc)_n (n = 2, 3, and 4)

ITC experiments were conducted at 25 °C by titrating a 1.0 mM solution of GlcNAc or (GlcNAc)_n (n = 2, 3, 4, 5 or 6) into VcCBP solution (50 μM). As shown in Fig. 9A–F, titrations with GlcNAc, (GlcNAc)₅, or (GlcNAc)₆ did not result in any heat release/absorption, whereas (GlcNAc)₂, (GlcNAc)₃ and (GlcNAc)₄ exhibited clear heat release or absorption. This was consistent with the experimental data obtained from thermal unfolding experiments (Supplementary Fig. S1 and Table 2). It is noteworthy that titration with (GlcNAc)₂ resulted in heat release but (GlcNAc)₃ and (GlcNAc)₄ resulted in absorption. The thermodynamic mechanism of the interaction with (GlcNAc)₂ appears to be basically different from those with (GlcNAc)₃ and (GlcNAc)₄. Thermodynamic parameters were obtained based on the experimental data as listed in Table 3. For (GlcNAc)₂, the enthalpic contribution (ΔH°) and the association constant (K_{assoc}) were obtained based on the theoretical fit and were -2.40 kcal mol⁻¹ and 4.08×10^6 M⁻¹, respectively. Thus the binding free energy (ΔG°) was -8.85 kcal mol⁻¹. The entropic contribution ($-T\Delta S^\circ$) was calculated to be -6.45 kcal mol⁻¹. The interaction was entropy-driven with a smaller enthalpy gain. When (GlcNAc)₃ or (GlcNAc)₄ was used instead of (GlcNAc)₂, the entropy gain was even greater but this was compensated by the loss of enthalpy, resulting in lower binding affinities (ΔG°) of -7.54 and -8.36 kcal mol⁻¹, respectively.

A: unliganded VcCBP

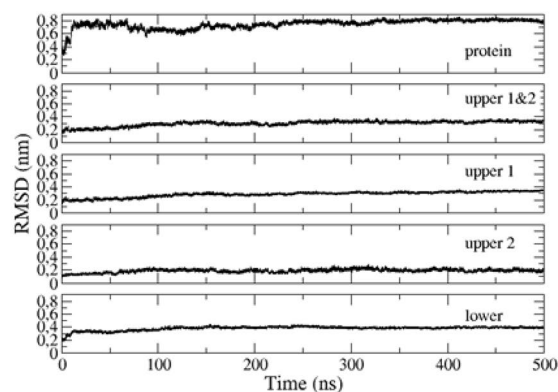
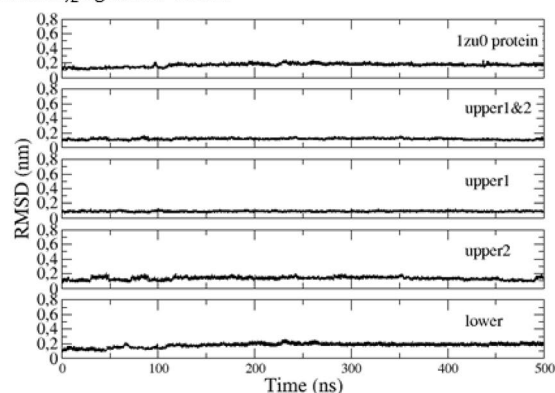
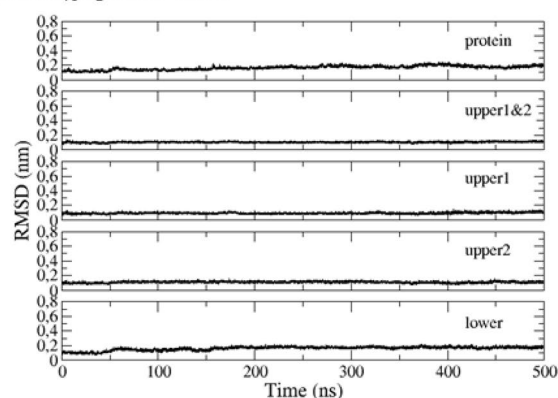
B: (GlcNAc)₂-liganded VcCBPC: (GlcNAc)₃-liganded VcCBP

Figure 5. Time-dependent backbone RMSDs calculated by molecular dynamics simulation. Unliganded VcCBP (**A**, 8I5J), (GlcNAc)₂-liganded VcCBP (**B**, 1ZU0) and (GlcNAc)₃-liganded VcCBP (**C**, 8I5K). From top to bottom: the entire protein structure, Upper1 + Upper2, Upper1, Upper2, and Lower domains.

Temperature dependences of thermodynamic parameters for (GlcNAc)_n binding

ITC profiles of (GlcNAc)₃ titration into VcCBP at various temperatures (10–35 °C) are shown in Supplementary Fig. S2. Heat absorption was observed up to 25 °C and converted to heat release at 30–35 °C. The binding stoichiometry of (GlcNAc)₃/VcCBP remained almost constant at approximately 1.0. The interactions were more exothermic at higher temperature. The temperature dependence was then examined for (GlcNAc)₂ and (GlcNAc)₄ at various temperatures from 5 to 35 °C, and the thermodynamic parameters obtained were plotted against individual temperatures. The results are shown in Supplementary Figs. S3A–S3C. The individual thermodynamic parameters (ΔG° , ΔH° and $-T\Delta S^\circ$) for the binding of (GlcNAc)₂, (GlcNAc)₃ and (GlcNAc)₄ were linearly related to the temperature. Based on the slope of ΔH° , we calculated heat capacity changes (ΔC_p°), which are listed in Table 4. Large negative values of ΔC_p° were observed for the interactions of individual (GlcNAc)_n with VcCBP, indicating significant hydrophobic interactions³⁵. The highest ΔC_p° (−406 cal K^{−1} mol^{−1}) was observed for the VcCBP-(GlcNAc)₃ interaction, while lower ΔC_p° values were obtained for (GlcNAc)₂ (−291 cal K^{−1} mol^{−1}) and (GlcNAc)₄ (−307 cal K^{−1} mol^{−1}). From the ΔC_p° values thus obtained, we calculated the solvation entropy change

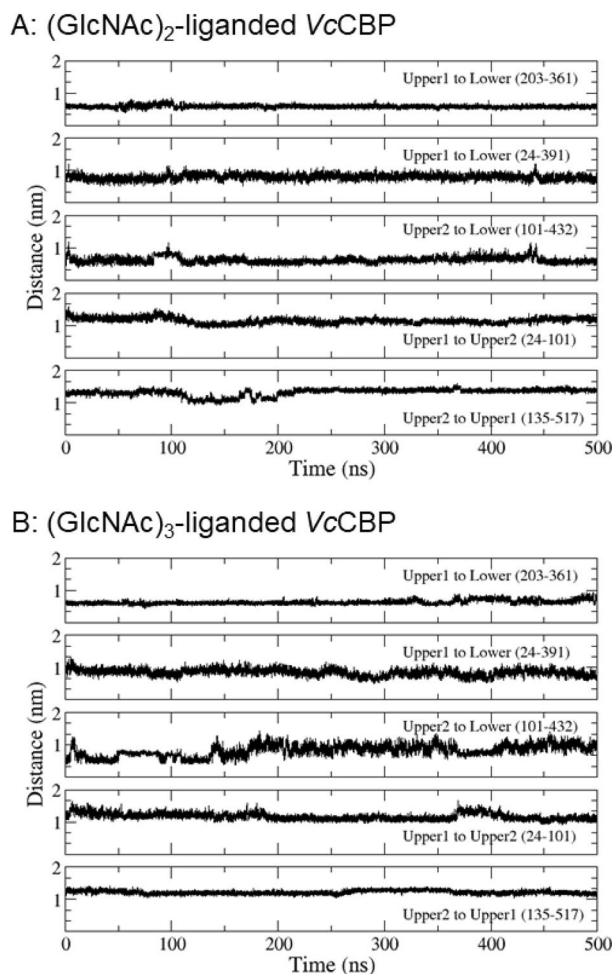


Figure 6. The selected, time-dependent interatomic distances between α -carbons of the amino acids. From top to bottom: Asn203-Gly361 (Upper1-Lower), Thr24-Ala391 (Upper1-Lower), Gln101-Gly432 (Upper2-Lower), Thr24-Gln101 (Upper1-Upper2) and Ser135-Glu517 (Upper2-Upper1), for (GlcNAc)₂-liganded VcCBP (A, 1ZU0) and (GlcNAc)₃-liganded VcCBP (B, 8I5K).

($\Delta S^{\circ}_{\text{solv}}$) according to Eq. (2), and subsequently obtained the conformational entropy change ($\Delta S^{\circ}_{\text{conf}}$) based on the Eq. (3). The results are also listed in Table 4. Both favorable $\Delta S^{\circ}_{\text{solv}}$ and unfavorable $\Delta S^{\circ}_{\text{conf}}$ contributed to the interaction with (GlcNAc)_n. The favorable $\Delta S^{\circ}_{\text{solv}}$ can be explained by the exclusion of bound water molecules upon ligand binding, while the unfavorable $\Delta S^{\circ}_{\text{conf}}$ by the reduction in the fluctuation of the VcCBP structure as seen in Fig. 5, which shows the fluctuations of RMSDs in unliganded and liganded states.

Binding experiments using VcCBP_R27A

To define the contribution of the Arg27 side chain to (GlcNAc)₃ binding, we conducted ITC analysis of the interaction with (GlcNAc)₃ using VcCBP_R27A. As in the case of VcCBP, heat absorption was observed up to 20 °C, and converted to heat release at 25–35 °C (Supplementary Fig. S4). The thermodynamic parameters were obtained for the individual temperatures, and the values were plotted as shown in Supplementary Fig. S3D. At all temperatures tested, the binding affinities (ΔG°) were found to be reduced by about 1 kcal/mol in VcCBP-R27A (Supplementary Fig. S3D, Table 3). Thus Arg27 participates in the binding of (GlcNAc)₃ but its contribution is only moderate. The temperature dependence of ΔH° of VcCBP_R27A indicated that both the solvation entropy gain and conformational entropy loss were reduced ($\Delta S^{\circ}_{\text{solv}}$, 110.9 \rightarrow 93.5 cal K⁻¹ mol⁻¹; $\Delta S^{\circ}_{\text{conf}}$, -70.3 \rightarrow -59.9 cal K⁻¹ mol⁻¹), and the reduction in the total entropic contribution ($-T\Delta S^{\circ}$, -9.56 \rightarrow -7.30 kcal mol⁻¹) was compensated by the reduction in enthalpy loss (ΔH° , 2.02 \rightarrow 0.92 kcal mol⁻¹) (Tables 3 and 4).

Discussion

(GlcNAc)₂ binds primarily to the Upper1/Lower interface (Site1/Site2)

As listed in Table 3, the interaction of (GlcNAc)₂ with VcCBP was entropy-driven ($-T\Delta S^{\circ} = -6.45$ kcal mol⁻¹) with a smaller enthalpy gain ($\Delta H^{\circ} = -2.40$ kcal mol⁻¹). This is consistent with the data reported for VhCBP ($\Delta H^{\circ} = -3.9$ kcal mol⁻¹ and $-T\Delta S^{\circ} = -6.4$ kcal mol⁻¹)²⁸. Since the amino acid residues involved in (GlcNAc)₂ binding in VhCBP are almost conserved in VcCBP as shown in Fig. 1, the mechanism of (GlcNAc)₂ binding in

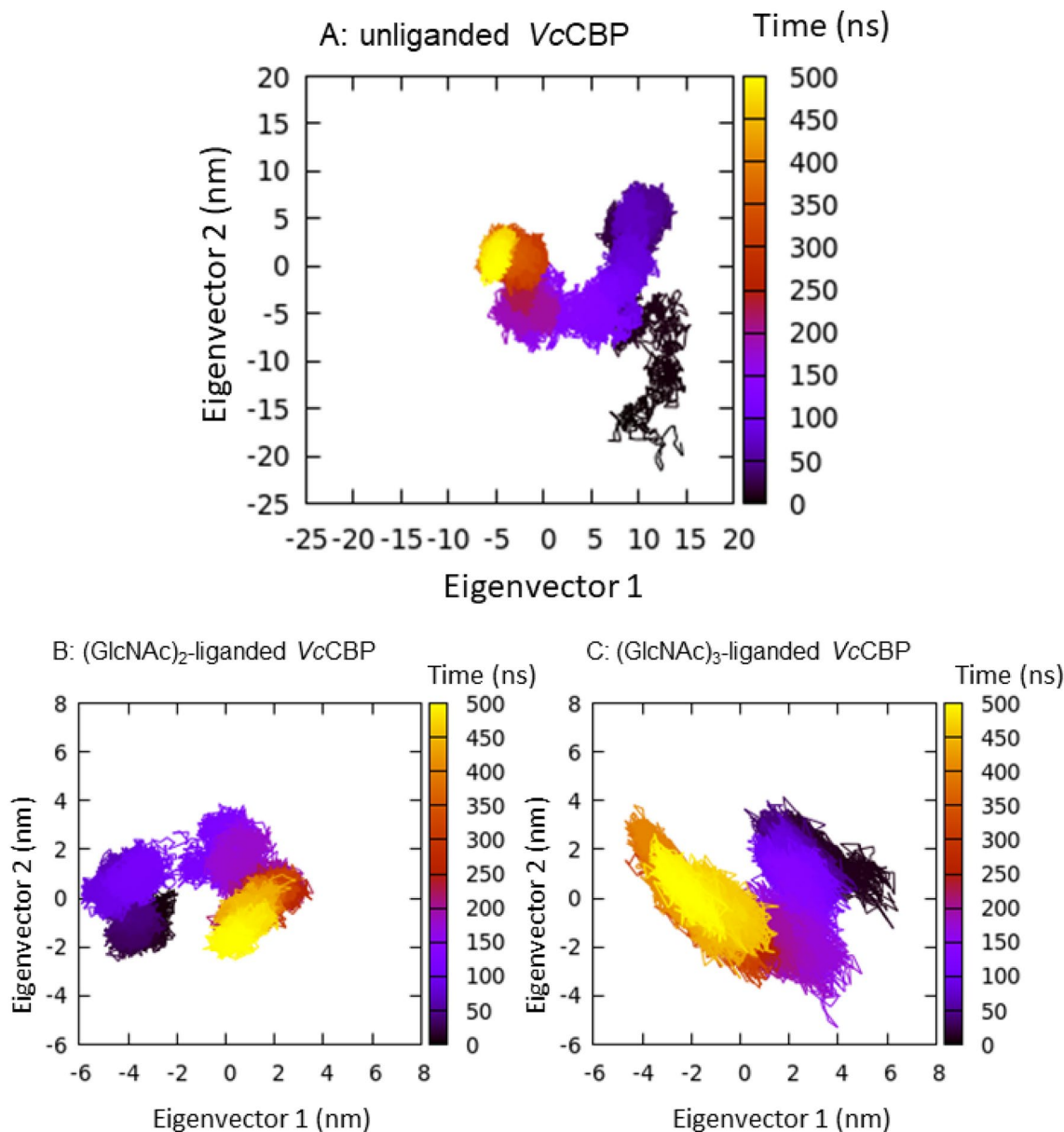


Figure 7. Two-dimensional projections of eigenvector1 and eigenvector2 showing the time-dependent conformational changes of unliganded (A), (GlcNAc)₂-liganded (B) and (GlcNAc)₃-liganded VcCBP (C). The color changes from black, violet, purple, magenta, red, orange and to yellow with progress of the simulation time.

VcCBP is similar to that in VhCBP from the viewpoints of structure and thermodynamics. As shown in Fig. 3B and Supplementary Table S1, only the Upper1/Lower interface appears to be involved in the binding of two GlcNAc residues from the non-reducing end. Since VcCBP has been regarded as being specific for (GlcNAc)₂ translocation to the ABC-transporter¹⁸, (GlcNAc)₂ binds primarily to the interface between the Upper1 and Lower domains (Site1/Site2) before translocation.

Enthalpic Site1/Site2 and entropic Site3/Site4

From the thermal unfolding and ITC experiments, we found that the binding affinities of (GlcNAc)_n (n = 2, 3, and 4) were comparable with each other (Supplementary Fig. S1 and Fig. 9; Tables 2 and 3), but that favorable contributions of entropic term ($-T\Delta S^\circ$) to the binding affinities were enhanced in the order (GlcNAc)₂ < (GlcNAc)₃ < (GlcNAc)₄ (Table 3). As the entropic contributions were enhanced, compensations were clearly found in the enthalpic terms. As seen from the crystal structures of (GlcNAc)₂-liganded (1ZU0) and (GlcNAc)₃-liganded VcCBP (Fig. 3B), reducing-end GlcNAc of bound (GlcNAc)₃ exists beyond the Upper1/Lower interface and that of bound (GlcNAc)₄ may be in contact with the Upper2/Lower interface (Site3 and Site4). From the temperature dependence of ΔH° , we found that favorable solvation entropy ($\Delta S^\circ_{\text{solv}}$) predominated over conformational entropy changes ($\Delta S^\circ_{\text{conf}}$) as listed in Table 4, suggesting that Site3 and Site4 bind a relatively large number

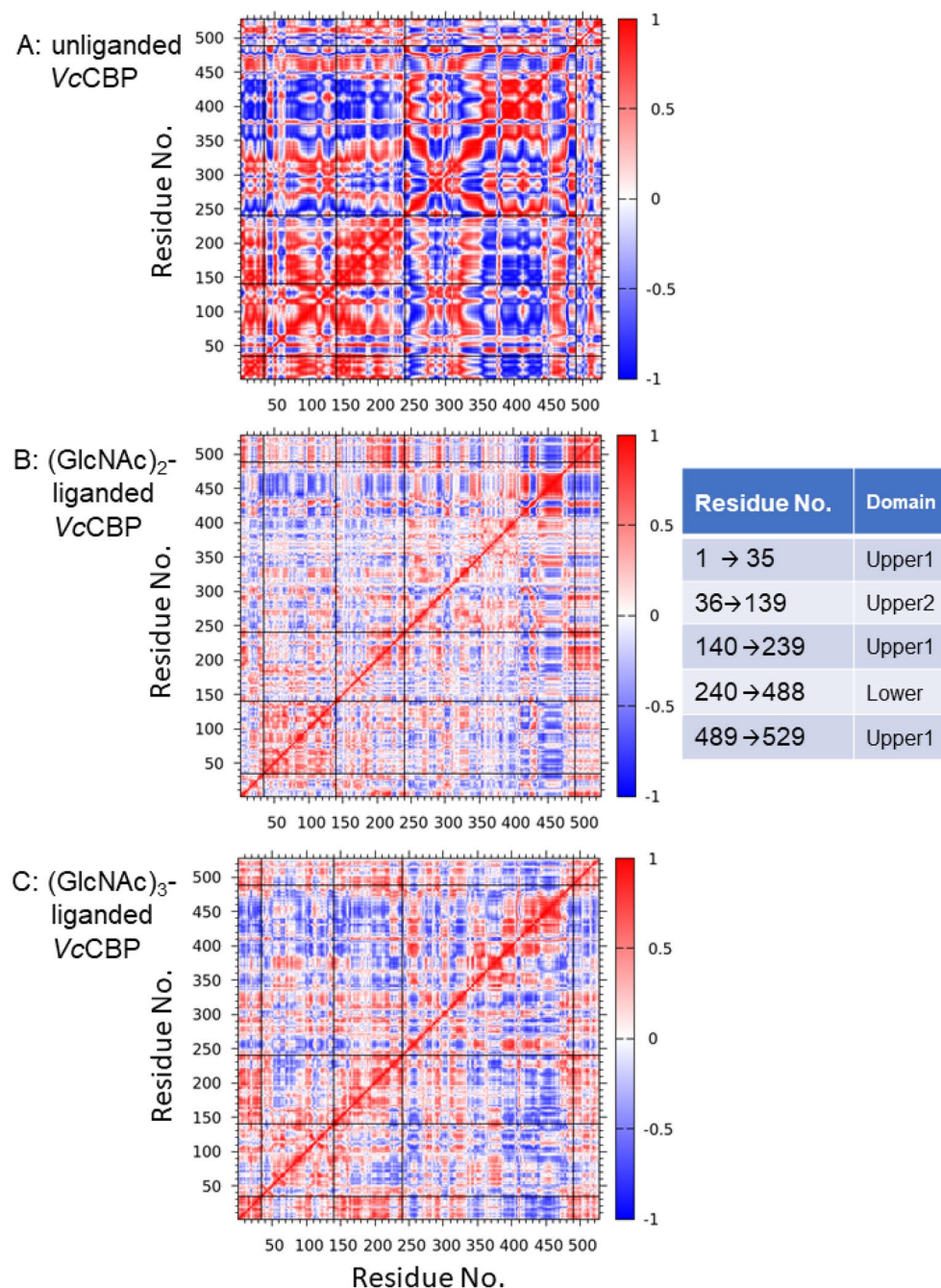


Figure 8. Cross correlation heat maps of Ca atoms around their mean positions for the entire simulation period. **(A)** Unliganded VcCBP (8I5J); **(B)** (GlcNAc)₂-liganded VcCBP (1ZU0); and **(C)** (GlcNAc)₃-liganded VcCBP (8I5K). The color gradation from red to blue corresponds to extents of correlated motions (from 1 to 0, positive correlations) and anti-correlated motions (from 0 to -1, negative correlations). The horizontal and vertical axes represent the amino acid residue No. of VcCBP. The thin guided lines in the figures represent the boundaries of the individual domains, Upper1, Upper2, and Lower. The correspondence table on the right shows which domain is located where.

of water molecules in the open form, and that the bound water molecules may be excluded upon (GlcNAc)₃ or (GlcNAc)₄ binding. We calculated the solvent-accessible surface areas (ASA) for unliganded, (GlcNAc)₂-liganded, and (GlcNAc)₃-liganded VcCBPs based on their crystal structures³⁶, and the data were presented in Supplementary Table 2. We found that the apolar solvent accessible surface area (ASA_{apolar}) was reduced by 3.3 or 4.2% upon binding of (GlcNAc)₂ or (GlcNAc)₃, respectively. This indicated that (GlcNAc)₃ binding to VcCBP excludes more water molecules from the apolar surface than (GlcNAc)₂ binding. This is consistent with the positive values of ΔS_{sol} , which is greater in (GlcNAc)₃ than in (GlcNAc)₂ (Table 4). The thermodynamic data

	T_m (°C)	ΔT_m (°C)
VcCBP alone	43.2 ± 1.2	
+ GlcNAc	46.0 ± 1.0	2.8
+ (GlcNAc) ₂	54.4 ± 0.1	11.2
+ (GlcNAc) ₃	54.8 ± 0.3	11.6
+ (GlcNAc) ₄	53.8 ± 0.0	10.6
+ (GlcNAc) ₅	49.6 ± 0.3	6.3
+ (GlcNAc) ₆	49.6 ± 0.2	6.3

Table 2. Transition temperatures of thermal unfolding of VcCBP in the presence of GlcNAc and (GlcNAc)_n (n = 2, 3, 4, 5, or 6). The unfolding experiments were conducted in 20 mM Tris–HCl buffer pH 8.0, monitoring the CD value at 222 nm using a Jasco J-720 spectropolarimeter (0.1 cm of cell length). Final concentrations of the protein and (GlcNAc)_n were 8 μM and 8 mM, respectively. Two repeated experiments were conducted to obtain the individual T_m values.

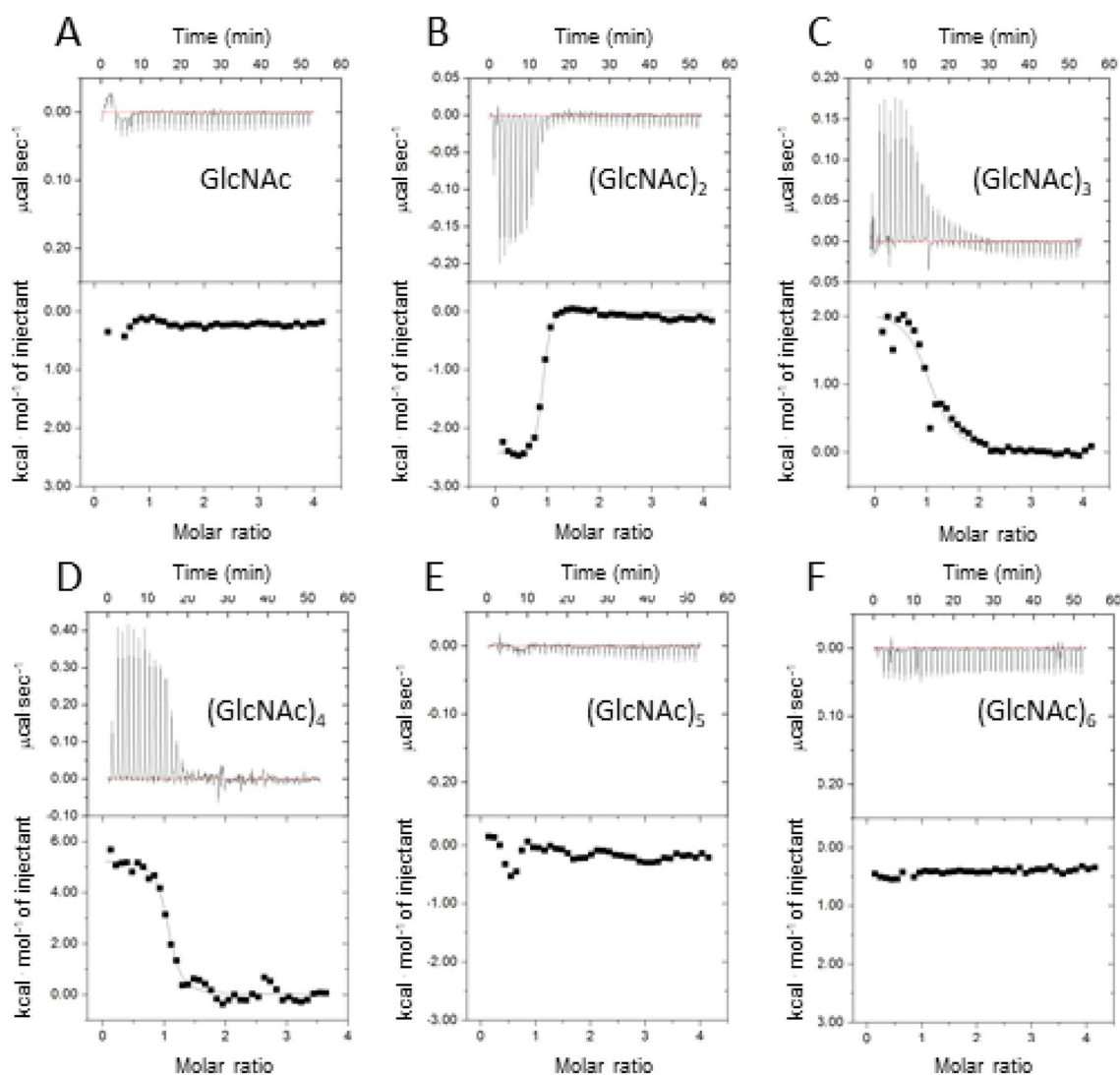


Figure 9. Thermograms and the theoretical fits obtained by ITC analysis. Titrations of VcCBP with GlcNAc (A), (GlcNAc)₂ (B), (GlcNAc)₃ (C), (GlcNAc)₄ (D), (GlcNAc)₅ (E) and (GlcNAc)₆ (F), were conducted in 20 mM Tris–HCl buffer, pH 8.0 at 25 °C. An iTC200 system (MicroCal Co.) was used to collect and analyze the experimental data. Other reaction conditions are described in the “Methods” section.

		<i>N</i>	<i>K_a</i> (10 ⁶ M ⁻¹)	ΔH° (kcal·mol ⁻¹)	$-T\Delta S^\circ$ (kcal·mol ⁻¹)	ΔG° (kcal·mol ⁻¹)
VcCBP	(GlcNAc) ₂	0.9 ± 0.0	4.08 ± 1.03	- 2.40 ± 0.03	- 6.45 ± 0.11	- 8.85 ± 0.14
VcCBP	(GlcNAc) ₃	1.1 ± 0.01	0.64 ± 0.14	2.02 ± 0.09	- 9.56 ± 0.03	- 7.54 ± 0.12
VcCBP	(GlcNAc) ₄	1.0 ± 0.05	1.70 ± 0.04	5.47 ± 0.19	- 13.8 ± 0.16	- 8.36 ± 0.02
VcCBP_R27A	(GlcNAc) ₃	0.7 ± 0.18	0.083 ± 0.061	0.92 ± 0.30	- 7.30 ± 0.55	- 6.38 ± 0.55

Table 3. Thermodynamic parameters obtained for the VcCBP/(GlcNAc)_n (n = 2, 3, and 4) interactions by means of ITC. The individual ligands (1.0 mM) were titrated into the VcCBP or VcCBP_R27A solution (50 μM) in 20 mM Tris–HCl buffer, pH 8.0, in the sample cell (0.2028 mL). The experiments were performed at 25 °C using an iTC200 system (Microcal Northampton, MA, USA). Three repeated experiments were conducted to obtain individual sets of the thermodynamic parameters.

Protein	Ligand	ΔC_p° (cal K ⁻¹ mol ⁻¹)	ΔS_r° (cal K ⁻¹ mol ⁻¹)	ΔS_{mix}° (cal K ⁻¹ mol ⁻¹)	ΔS_{solv}° (cal K ⁻¹ mol ⁻¹)	ΔS_{conf}° (cal K ⁻¹ mol ⁻¹)
VcCBP	(GlcNAc) ₂	- 291 ± 1	22.0 ± 0.4	- 8	81.6 ± 0.2	- 51.6 ± 0.6
VcCBP	(GlcNAc) ₃	- 406 ± 6	32.6 ± 0.1	- 8	110.9 ± 1.8	- 70.3 ± 1.9
VcCBP	(GlcNAc) ₄	- 307 ± 10	47.2 ± 0.6	- 8	84.0 ± 2.8	- 28.9 ± 3.3
VcCBP_R27A	(GlcNAc) ₃	- 343 ± 1	25.6 ± 0.0	- 8	93.5 ± 0.2	- 59.9 ± 0.2

Table 4. Dissection of entropic contribution to the VcCBP/(GlcNAc)_n (n = 2, 3, and 4) interactions. The individual ligands (1.0 mM) were titrated into the VcCBP or VcCBP_R27A solution (50 μM) in 20 mM Tris–HCl buffer, pH 8.0, in the sample cell (0.2028 mL) at various temperatures. Individual entropic contributions were obtained from Supplementary Fig. S3, and Eqs. (2) and (3).

obtained for (GlcNAc)₃ binding to VcCBP_R27A (Tables 3 and 4) well supported the contribution of solvation entropy in (GlcNAc)₃ binding. The favorable entropy change derived from solvation was suppressed by mutation of Arg27 to alanine, indicating that the bound water molecules were at least partly removed by this mutation. For the interactions of Site3 GlcNAc involving Arg27, it appears that the hydrogen bond is less significant but entropic contribution caused by the release of bound water molecules is more significant. Thus, we propose that Site1/Site2 corresponding to the Upper1/Lower interface and Site3/Site4 corresponding to the Upper1/Upper2/Lower interface can be respectively regarded as enthalpic- and entropic-interaction sites.

Upper1 and Upper2 domains play different roles in (GlcNAc)₂ translocation

In VcCBP, we found three structural domains, the Upper1, Upper2 and Lower domains (Fig. 2), which correspond to domains I, II and III of cluster C SBPs reported by Chandravanshi et al.¹⁹. The three-domain organization appears to be significant from a functional viewpoint. One of the reasons for this significance derives from the clear distinction found in the binding thermodynamics between Site1/Site2 and Site3/Site4; although both enthalpy and entropy contributions were involved in the interaction with the former site, the entropic contribution predominates in the latter site (Tables 3 and 4). The second reason is the larger RMSD of the Upper2 domain upon superimposition of (GlcNAc)₂-liganded and (GlcNAc)₃-liganded VcCBP structures (Fig. 2C). The larger RMSD is likely to produce the space for accommodating the additional GlcNAc unit at the interface between Upper1 and Upper2 domains. The third reason is the distinction in the molecular movements between the Upper1/Lower and Upper2/Lower interfaces. The distance Gln101–Gly432 fluctuated more in (GlcNAc)₃-liganded than in (GlcNAc)₂-liganded VcCBP, whereas the distances, Asn203–Gly361 and Thr24–Ala391 were similar (Fig. 6). In (GlcNAc)₃-liganded VcCBP, the fluctuations were more intensive in the Upper2/Lower interface (Gln101–Gly432) than in the Upper1/Lower interface (Asn203–Gly361/Thr24–Ala391) as shown in Fig. 6B. The larger fluctuation in agreed well with the highest B-factor values in the loop immediately following η6 and in the loop between α14 and α15 (Fig. 4). A “half-open” conformation observed in Trp513-mutated VhCBP³⁰ may correspond to a snapshot of the largely fluctuated state. Cross-correlation heat maps of (GlcNAc)₂-liganded and (GlcNAc)₃-liganded VcCBP (Fig. 8) also showed the independency of the Upper1 and Upper2 domains in their molecular movements; thus, it is most likely that Upper1 and Upper2 domains play different roles in the (GlcNAc)₂ translocation process.

Hypothetical releasing site, Site5/Site6

As shown in Fig. 9, titrations of VcCBP with (GlcNAc)₅ or (GlcNAc)₆ did not result in any heat release/absorption; the binding affinities are too low to obtain the thermodynamic parameters for these oligosaccharides by ITC. Nevertheless, thermal unfolding data (Supplemental Fig. S1 and Table 2) revealed a significant elevation of the transition temperature of thermal unfolding ($\Delta T_m = 6.3$ °C), suggesting a significant interaction of (GlcNAc)₅ or (GlcNAc)₆ with VcCBP. In the interaction of (GlcNAc)₆ with VcCBP, the two GlcNAc residues of the non-reducing end interact with the interface between the Upper1/Lower (Site1/Site2) and the neighboring (GlcNAc)₂ unit also interacts with the Upper1/Upper2/Lower interface (Site3/Site4). However, the additional (GlcNAc)₂ unit of the reducing-end side may be repelled from the protein surface (hypothetical Site5/Site6). This situation may

bring about the lower but significant ΔT_m in Table 2, accounting for the lower binding affinity of (GlcNAc)₅ and (GlcNAc)₆. We propose here that a specific substrate (GlcNAc)₂ primarily binds Site1/Site2 with both enthalpy- and entropy-driven interactions, and is subsequently translocated to Site3/Site4, where the binding interaction is looser, leading to release of the sugar molecule from Site5/Site6 to a (GlcNAc)₂-specific ABC transporter. The loosening of the interaction at Site3/Site4 may be caused by the higher mobility of the Upper2/Lower interface observed in the (GlcNAc)₃-liganded structure (Figs. 4, 6 and 7). Thus, all translocation processes are conducted by the cooperative action of the three domains, Upper1, Upper2 and Lower.

Structure triggering the (GlcNAc)₂ unit translocation

Kitaoku et al.³⁰ observed electron density of the Phe411 side chain with a full occupancy (1.0) close to the *N*-acetyl methyl group of the Site2 GlcNAc in (GlcNAc)₂-liganded *Vh*CBP. However, in (GlcNAc)₃-liganded *Vh*CBP, the same side chain was found 3.5 Å away from the Site2 GlcNAc with a full occupancy (1.0). Here, in (GlcNAc)₃-liganded *Vc*CBP we observed the side chain of the corresponding phenylalanine residue (Phe410) at both positions with individual occupancies of 0.5 (Fig. 3A), indicating a significant flipping of the Phe410 side chain. The density of the Site3 GlcNAc was also observed with an occupancy of 0.5, indicating the higher mobility of the Site3 GlcNAc (Figs. 3A and 4D), which appeared to be coordinated with that of the Phe410 side chain. The coordinated motions of the phenylalanyl side chain and the Site3 GlcNAc suggested that Phe410 may be involved in the translocation of bound (GlcNAc)₂ from Site1/Site2 to Site3/Site4. Perhaps the translocation process is triggered by interaction with the corresponding ABC transporter, which may further translocate (GlcNAc)₂ to the hypothetical release site, Site5/Site6, located in the Upper2/Lower interface.

Conclusion

Taken together, the three domains, Upper1, Upper2, and Lower domains, found in the crystal structure of *Vc*CBP play different roles and function cooperatively in translocation of (GlcNAc)₂. The mechanism proposed here was fully supported by binding data obtained by thermal unfolding and ITC experiments and may be applicable to other translocation systems involving SBPs belonging to the same cluster.

Materials and methods

Materials

(GlcNAc)_n (n = 2–6) oligosaccharides were produced by acid hydrolysis of chitin³⁷ and purified by gel filtration on Cellufine Gcl-25 m (JNC Co., Tokyo, 3.5 × 180 cm). Ni-NTA Agarose was purchased from QIAGEN (Tokyo, Japan). Q-Sepharose Fast Flow and HiPrep 16/60 Sephacryl S-100 were from GE Healthcare (Tokyo, Japan) and TOYOPEARL Butyl-650 M was from Tosoh (Tokyo, Japan). Other reagents were of analytical grade and commercially available.

Construction of expression plasmid for *Vc*CBP and *Vc*CBP_R27A

Synthetic genes encoding *Vc*CBP and *Vc*CBP_R27A, in which Arg27 was mutated to alanine, were obtained from Invitrogen (Carlsbad, CA, USA). The nucleotide sequences of the genes were optimized to increase expression in *E. coli* without changing the amino acid sequences of these proteins. The expression vectors, pRham-*Vc*CBP and pRham-*Vc*CBP_R27A, were constructed by the Expresso^(R) Rhamnose Cloning and Expression System, N-His (Lucigen, UK).

Protein expression and purification

The expression vector, pRham-*Vc*CBP or pRham-*Vc*CBP_R27A was transformed into *E. coli* C43(DE3). Induction with 0.2% α-L(+) -rhamnose was conducted according to the supplier's instruction. After induction, the culture was incubated at 15 °C for 40 h, then the cells were harvested and disrupted by sonication in 20 mM Tris-HCl buffer, pH 8.0. The sonicated extract was centrifuged at 12,000×g for 15 min at 4 °C. The soluble fraction was dialyzed against 20 mM Tris-HCl buffer pH 8.0, and applied to a Ni-NTA column equilibrated with the same buffer. After washing the column with the Tris buffer, the bound protein fractions were eluted with a linear gradient of 0–0.2 M imidazole. The fractions containing a protein of molecular mass 60 kDa, which corresponds to that of *Vc*CBP, were collected, and ammonium sulfate was added to the protein solution to a final concentration of 1 M. The solution was applied to the TOYOPEARL Butyl 650 M column equilibrated with 20 mM Tris-HCl buffer pH 8.0 containing 1 M ammonium sulfate. The adsorbed fraction was eluted with a linear gradient of 1–0 M ammonium sulfate in 20 mM Tris-HCl buffer pH 8.0. The protein fractions containing *Vc*CBP were pooled and applied to a Q-Sepharose column previously equilibrated with the same buffer. The protein was eluted stepwise with 0.15 M NaCl in the same buffer. The fractions containing *Vc*CBP were pooled and further applied to a Sephacryl S-100 HR gel-filtration column equilibrated with the same buffer containing 0.1 M NaCl. Fractions exhibiting a single protein band on SDS-PAGE³⁸ (Supplementary Fig. S5) were pooled and stored at 4 °C.

Protein concentration

Protein concentrations were determined by reading absorbance at 280 nm, using an extinction coefficient of *Vc*CBP (110,365 M⁻¹ cm⁻¹) calculated from the equation proposed by Pace et al.³⁹.

Thermal unfolding experiments

To obtain the thermal unfolding curve of *Vc*CBP, the CD value at 222 nm was monitored using a Jasco J-720 spectropolarimeter (cell length 0.1 cm), while the solution temperature was raised at a rate of 1 °C min⁻¹ using a temperature controller (PTC-423L, Jasco). To facilitate comparison of unfolding curves, the experimental

data (molar ellipticities) were normalized to obtain unfolded fractions at individual temperatures. To assess the binding ability of GlcNAc and (GlcNAc)_n (n = 2, 3, 4, 5, and 6), the unfolding experiments of VcCBP were conducted in the presence or absence of (GlcNAc)_n. Individual unfolding experiments were repeated twice under the same conditions. The transition temperature of thermal unfolding (T_m) was elevated when the ligand was added to the VcCBP solution. The elevation of T_m (ΔT_m) indicated the binding of the ligand⁴⁰. The solvent condition was 20 mM Tris-HCl buffer pH 8.0. Final concentrations of the protein and (GlcNAc)_n were 8 μ M and 8 mM, respectively.

Isothermal titration calorimetry (ITC)

Solutions of 50 μ M VcCBP or VcCBP_R27A in 20 mM Tris-HCl buffer, pH 8.0, were degassed and loaded into the sample cell (0.2028 mL). The individual ligands (1.0 mM) were dissolved in the same buffer, degassed and loaded into a syringe. Calorimetric titration was performed at 25 °C using an iTC200 system (Microcal Northampton, MA, USA). In the titrations, 2.5 μ L of a ligand was injected into the sample cell at 180-s intervals with a stirring speed of 1000 rpm. The dilution heat caused by each titration was measured by titrating ligand to buffer solution without protein under identical conditions. The dilution heat thus obtained was subtracted from the heat change that was observed in the presence of protein. Individual titration experiments were repeated three times to obtain reliable values of thermodynamic parameters. The Origin software installed in the ITC instrument was used to analyze the ITC data. Using the One-set of Sites model, individual datasets obtained from the titration experiments fitted well to the theoretical curves, providing the stoichiometries (n), equilibrium association constants (K_a) and enthalpy changes (ΔH°) of the protein-ligand interactions. The binding free energy change (ΔG°) and entropy change (ΔS°) were calculated from the relationship as follows,

$$\Delta G^\circ = -RT \cdot \ln K_a = \Delta H^\circ - T\Delta S^\circ \quad (1)$$

The accuracy of the thermodynamic parameters obtained was assessed from the c -values calculated from the equation $c = n \cdot K_a \cdot [M]_p$, where $[M]_p$ is the total concentration of protein⁴¹. The ITC measurements were conducted in 20 mM Tris-HCl buffer pH 8.0, at various temperatures from 5 to 35 °C. The ΔH° values obtained for various temperatures were plotted against temperatures, and the slope of a straight line fitted to the experimental points corresponds to the heat capacity change (ΔC_p°). As the entropy of solvation is regarded as zero for proteins near 385 K, ΔC_p° was converted to the solvation entropy change ($\Delta S_{\text{solv}}^\circ$) at 25 °C (298 K) according to the following relationship,

$$\Delta S_{\text{solv}}^\circ = \Delta C_p^\circ \ln \left(\frac{298.15 \text{ K}}{385.15 \text{ K}} \right) \quad (2)$$

The conformational entropy change ($\Delta S_{\text{conf}}^\circ$) was calculated from ΔS° obtained from Eq. (1), the solvation entropy change ($\Delta S_{\text{solv}}^\circ$) and the mixing entropy change ($\Delta S_{\text{mix}}^\circ$, $-8 \text{ cal K}^{-1} \text{ mol}^{-1}$), based on the following Eq. ⁴²,

$$\Delta S^\circ = \Delta S_{\text{solv}}^\circ + \Delta S_{\text{mix}}^\circ + \Delta S_{\text{conf}}^\circ \quad (3)$$

Crystallization and data collection

Crystallization conditions for VcCBP were screened using the sparse-matrix sampling method by sitting drop vapor diffusion at 20 °C. Under optimized crystallization conditions, 1 μ L of protein solution (5 mg/ml in water) was mixed with 1 μ L of 0.1 M sodium citrate containing 0.2 M ammonium acetate with 30% w/v polyethylene glycol 4000, pH 5.6. Rod-like crystals of VcCBP grew within 3 weeks to a size of up to 0.1 \times 0.1 \times 0.5 mm³. To prepare the crystals of (GlcNAc)₃-liganded VcCBP, the unliganded crystals were transferred to the crystallization well solution containing 26 mM (GlcNAc)₃ and incubated at 20 °C for 3.5 h. The crystals were successfully grown in the presence of (GlcNAc)₃. For data collection, the crystals were transferred into the cryoprotectant solution containing 0.01 M zinc sulfate, 0.1 M MES (pH 6.5) and 30% PEG MME550, and then flash-cooled in a nitrogen stream at 95 K. The diffraction data were collected at the beam-line BL-17A of Photon Factory (Ibaraki, Japan), using an EIGER X 16 M (Dectris), at a cryogenic temperature (95 K). The data were integrated and scaled with XDS⁴³. The processing statistics are summarized in Table 1.

Structural determination and refinement

The structures of unliganded and (GlcNAc)₃-liganded VcCBP were solved by the molecular replacement method using the program PHASER⁴⁴, where the structures of unliganded VcCBP (PDB code, 1ZTY) and (GlcNAc)₂-liganded VhCBP (PDB code, 5YQW) served as search models, respectively. For unliganded VcCBP, two protein molecules were located in the crystallographic asymmetric unit. The model was improved by several rounds of refinement with PHENIX⁴⁵ and COOT programs⁴⁶. Occupancies of the Site3 GlcNAc and two conformers of the Phe410 side chain were set at 0.5. The structure of unliganded VcCBP was refined to an $R_{\text{work}}/R_{\text{free}}$ of 16.2/19.1% at a resolution of 1.6 Å. The final model contains two protein molecules that include residues 1–532 for each molecule and 773 water molecules. The stereochemistry of the model was verified using MolProbity⁴⁷, showing 96.7%, 3.3% and 0% of protein residues in the favored, allowed and disallowed regions of the Ramachandran plot, respectively. For (GlcNAc)₃-liganded VcCBP, one protein molecule was located in the crystallographic asymmetric unit. The structure of (GlcNAc)₃-liganded VcCBP was refined to an $R_{\text{work}}/R_{\text{free}}$ of 16.9/18.7% at a resolution of 1.22 Å. The final model contained one protein molecule that includes residues 1–532 and 563 water molecules. The stereochemistry verification showed 97.3%, 2.7% and 0% of protein residues in the corresponding regions, respectively. Molecular graphics were illustrated with PyMol software (<http://www.pymol.org/>). The refinement statistics are summarized in Table 1.

Molecular dynamics simulation

Molecular dynamics simulations were started from three different conformations of VcCBP, unliganded, (GlcNAc)₂-liganded, and (GlcNAc)₃-liganded VcCBP (PDB codes: 8I5J, 1ZU0, and 8I5K, respectively). Molecular dynamics package Gromacs⁴⁸, version 2019 and 2020.7, were used for the simulations and analysis of the simulated data. Protein topologies were generated with Amber99SB force field⁴⁹. The ligand topologies were generated with ACPYPE server⁵⁰. Amber GAFF force field was used for generation of the parameters and partial charges with AM1-BCC model to correspond to HF/6-31G* RESP charges⁵¹. The structures were placed in a rectangular box 1 nm from the protein and solvated in a pre-equilibrated water configuration with TIP3 water model⁵². Counter ions were added to neutralize the system. The structure was energy-minimized with steepest descent algorithm. To equilibrate the system into constant temperature, pressure and density 100 ps NVT and NPT simulations constraining the heavy atoms and 100 ps NPT simulation without constraints were performed before the actual NPT production runs. Berendsen pressure coupling was used in the constrained NPT simulation⁵³. In the production runs modified Berendsen method, velocity rescaling⁵⁴ was used for temperature coupling and Parinello and Rahman⁵⁵ for pressure coupling. Production runs were 500 ns with time step of 2 fs and the conformations and energies were collected every 10 ps. RMSDs of the whole protein and the domains from the starting conformation of each simulation were calculated based on the backbone atoms. Principal component analysis (PCA) or essential dynamics⁵⁶ based on the variations of Ca atom positions were used to extract the largest motions in each simulation. The cross-correlation data were obtained through ProDy server⁵⁷ via interface from VMD⁵⁸.

Solvent accessibility surface area (ASA)

ASA was calculated using the method reported by Fraczekiewicz and Braun³⁶ (GetArea 1.1, TX, United States), based on the crystal structures of unliganded, (GlcNAc)₂-liganded, and (GlcNAc)₃-liganded VcCBPs.

Data availability

The atomic co-ordinates and structural factor data of VcCBP have been deposited in the PDB database under the accession codes 8I5J and 8I5K for the unliganded VcCBP and (GlcNAc)₃-liganded VcCBP, respectively.

Received: 4 July 2023; Accepted: 10 November 2023

Published online: 23 November 2023

References

- Gooday, G. W. Aggressive and defensive roles for chitinases. *EXS* **87**, 157–169 (1999).
- Zhang, R., Zhou, J., Song, Z. & Huang, Z. Enzymatic properties of β -N-acetylglucosaminidases. *Appl. Microbiol. Biotechnol.* **102**, 93–103 (2018).
- Bonin, M., Sreekumar, S., Cord-Landwehr, S. & Moerschbacher, B. M. Preparation of defined chitosan oligosaccharides using chitin deacetylases. *Int. J. Mol. Sci.* **21**, 7835 (2020).
- Loose, J. S., Forsberg, Z., Fraaije, M. W., Eijssink, V. G. & Vaaje-Kolstad, G. A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monoxygenase. *FEBS Lett.* **588**, 3435–3440 (2014).
- Muthukrishnan, S., Merzendorfer, H., Arakane, Y. & Yang, Q. Chitin organizing and modifying enzymes and proteins involved in remodeling of the insect cuticle. *Adv. Exp. Med. Biol.* **1142**, 83–114 (2019).
- Nazari, B. *et al.* Chitin-induced gene expression in secondary metabolic pathways of *Streptomyces coelicolor* A3(2) grown in soil. *Appl. Environ. Microbiol.* **79**, 707–713 (2013).
- Zhang, X., Lin, H., Wang, X. & Austin, B. Significance of *Vibrio* species in the marine organic carbon cycle: A review. *Sci. China Earth Sci.* **61**, 1357–1368 (2018).
- Bassler, B. L., Yu, C., Lee, Y. C. & Roseman, S. Chitin utilization by marine bacteria: Degradation and catabolism of chitin oligosaccharides by *Vibrio furnissii*. *J. Biol. Chem.* **266**, 24276–24286 (1991).
- Suginta, W., Robertson, P. A. W., Austin, B., Fry, S. C. & Fothergill-Gilmore, L. A. Chitinases from *Vibrio*: Activity screening and purification of chiA from *Vibrio carchariae*. *J. Appl. Microbiol.* **89**, 76–84 (2000).
- Suginta, W. *et al.* An endochitinase A from *Vibrio carchariae*: Cloning, expression, mass and sequence analyses, and chitin hydrolysis. *Arch. Biochem. Biophys.* **424**, 171–180 (2004).
- Keyhani, N. O., Li, X. B. & Roseman, S. Chitin catabolism in the marine bacterium *Vibrio furnissii*. Identification and molecular cloning of a chitoporin. *J. Biol. Chem.* **275**, 33068–33076 (2000).
- Suginta, W. *et al.* Molecular uptake of chitooligosaccharides through chitoporin from the marine bacterium *Vibrio harveyi*. *PLoS ONE*. **8**, e55126 (2013).
- Suginta, W., Chumjan, W., Mahendran, K. R., Schulte, A. & Winterhalter, M. Chitoporin from *Vibrio harveyi*, a channel with exceptional sugar specificity. *J. Biol. Chem.* **288**, 11038–11046 (2013).
- Aunkham, A. *et al.* Structural basis for chitin acquisition by marine *Vibrio* species. *Nat. Commun.* **9**, 220 (2018).
- Keyhani, N. O. & Roseman, S. The chitin catabolic cascade in the marine bacterium *Vibrio furnissii*. Molecular cloning, isolation, and characterization of a periplasmic chitodextrinase. *J. Biol. Chem.* **271**, 33414–33424 (1996).
- Keyhani, N. O. & Roseman, S. The chitin catabolic cascade in the marine bacterium *Vibrio furnissii*. Molecular cloning, isolation, and characterization of a periplasmic β -N-acetylglucosaminidase. *J. Biol. Chem.* **271**, 33425–33432 (1996).
- Suginta, W., Chuenark, D., Mizuhara, M. & Fukamizo, T. Novel β -N-acetylglucosaminidases from *Vibrio harveyi* 650: Cloning, expression, enzymatic properties, and subsite identification. *BMC Biochem.* **11**, 40 (2010).
- Keyhani, N. O., Wang, L. X., Lee, Y. C. & Roseman, S. The chitin catabolic cascade in the marine bacterium *Vibrio furnissii*. Characterization of an N, N'-diacetyl chitobiose transport system. *J. Biol. Chem.* **271**, 33409–33413 (1996).
- Chandravanshi, M., Tripathi, S. K. & Kanaujia, S. P. An updated classification and mechanistic insights into ligand binding of the substrate-binding proteins. *FEBS Lett.* **595**, 2395–2409 (2021).
- Ortega, Á., Matilla, M. A. & Krell, T. The repertoire of solute-binding proteins of model bacteria reveals large differences in number, type, and ligand range. *Microbiol. Spectr.* **10**, e0205422 (2022).
- Li, X. & Roseman, S. The chitinolytic cascade in *Vibrios* is regulated by chitin oligosaccharides and a two-component chitin catabolic sensor/kinase. *Proc. Natl. Acad. Sci. USA.* **101**, 627–631 (2004).
- Klancher, C. A., Yamamoto, S., Dalia, T. N. & Dalia, A. B. ChiS is a noncanonical DNA-binding hybrid sensor kinase that directly regulates the chitin utilization program in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. USA.* **117**, 20180–20189 (2020).

23. Scheepers, G. H., Lycklama, J. A. & Poolman, B. An updated structural classification of substrate-binding proteins. *FEBS Lett.* **590**, 4393–4401 (2016).
24. Fukamizo, T., Kitaoku, Y. & Suginta, W. Periplasmic solute-binding proteins: Structure classification and chitoooligosaccharide recognition. *Int. J. Biol. Macromol.* **128**, 985–993 (2019).
25. Shears, P. Cholera. *Ann. Trop. Med. Parasitol.* **88**, 109–122 (1994).
26. Austin, B. & Zhang, X. H. *Vibrio harveyi*: A significant pathogen of marine vertebrates and invertebrates. *Let. Appl. Microbiol.* **43**, 119–214 (2006).
27. Chekan, J. R. *et al.* Structural and biochemical basis for mannan utilization by *Caldanaerobius polysaccharolyticus* strain ATCC BAA-17. *J. Biol. Chem.* **289**, 34965–34977 (2014).
28. Cuneo, M. J., Beese, L. S. & Hellinga, H. W. 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 (2009).
29. Suginta, W. *et al.* Structure and function of a novel periplasmic chitoooligosaccharide-binding protein from marine *Vibrio* bacteria. *J. Biol. Chem.* **293**, 5150–5159 (2018).
30. Kitaoku, Y. *et al.* A structural model for (GlcNAc)₂ translocation via a periplasmic chitoooligosaccharide-binding protein from marine *Vibrio* bacteria. *J. Biol. Chem.* **297**, 101071 (2021).
31. Chaudhuri, B. N., Ko, J., Park, C., Jones, T. A. & Mowbray, S. L. Structure of D-allose binding protein from *Escherichia coli* bound to D-allose at 1.8 Å resolution. *J. Mol. Biol.* **286**, 1519–1531 (1999).
32. Borrok, M. J., Kiessling, L. L. & Forest, K. T. Conformational changes of glucose/galactose-binding protein illuminated by open, unliganded, and ultra-high-resolution ligand-bound structures. *Protein Sci.* **16**, 1032–1041 (2007).
33. Anamizu, K. *et al.* Substrate size-dependent conformational changes of bacterial pectin-binding protein crucial for chemotaxis and assimilation. *Sci. Rep.* **12**, 12653 (2022).
34. Pace, C. N. & McGrath, T. Substrate stabilization of lysozyme to thermal and guanidine hydrochloride denaturation. *J. Biol. Chem.* **255**, 3862–3865 (1980).
35. Zolotnitsky, G. *et al.* Mapping glycoside hydrolase substrate subsites by isothermal titration calorimetry. *Proc. Natl. Acad. Sci. USA.* **101**, 11275–11280 (2004).
36. Fraczkiewicz, R. & Braun, W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem.* **19**, 319–333 (1998).
37. Rupley, J. A. The hydrolysis of chitin by concentrated hydrochloric acid, and the preparation of low-molecular weight substrates for lysozyme. *Biochim. Biophys. Acta* **83**, 245–255 (1964).
38. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
39. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423 (1995).
40. Honda, Y., Fukamizo, T., Boucher, I. & Brzezinski, R. Substrate binding to the inactive mutants of *Streptomyces* sp. N174 chitosanase: indirect evaluation from the thermal unfolding experiments. *FEBS Lett.* **411**, 346–350 (1997).
41. Turnbull, W. B. & Daranas, A. H. On the value of c : Can low affinity systems be studied by isothermal titration calorimetry?. *J. Am. Chem. Soc.* **125**, 14859–14866 (2003).
42. Ohnuma, T. *et al.* Chitin oligosaccharide binding to a family GH19 chitinase from the moss *Bryum coronatum*. *FEBS J.* **278**, 3991–4001 (2011).
43. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
44. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
45. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D.* **68**, 352–367 (2012).
46. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D.* **60**, 2126–2132 (2004).
47. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
48. van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
49. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
50. Sousa Da Silva, A. W. & Vrancan, W. F. ACPYPE: AnteChamber PYthon Parser interfacE. *BMC Res. Not.* **5**, 367 (2012).
51. Jakalian, A., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II Parameterization and validation. *J. Comput. Chem.* **23**, 1623–1641 (2002).
52. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
53. Berendsen, H. J. C., Postma, J. P. M., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
54. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
55. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
56. Amadei, A., Linssen, A. B. M. & Berendsen, H. J. C. Essential dynamics of proteins. *Proteins* **17**, 412–425 (1993).
57. Bakan, A., Meireles, L. M. & Bahar, I. ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575–1577 (2011).
58. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

Acknowledgements

This study was supported in part by a grant from the Agricultural Technology and Innovation Research Institute (ATIRI), Kindai University to T.O. We thank the beam-line staff at BL-17A of the Photon Factory (Ibaraki, Japan) for technical assistance during data collection. Thanks are also due to the CSC-IT Center for Science, Espoo, Finland, for computational resources. Finally, we are grateful to Dr. David Apps, The University of Edinburgh, UK, for his kindly reviewing our manuscript.

Author contributions

T.O., W.S. and T.F. proposed the research plan, and J.T., C.K., T.Y., D.T. and O.L. conducted the experiments. T.O., D.T., O.L., A.H.J., and T.F. analyzed and interpreted the experimental data, and T.O., W.S. and T.F. wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-47253-y>.

Correspondence and requests for materials should be addressed to T.O. or T.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023