Structural and thermodynamic insights into \( \beta \)-1,2-glucooligosaccharide capture by a solute-binding protein in *Listeria innocua*

Received for publication, December 18, 2017, and in revised form, April 18, 2018. Published, Papers in Press, April 20, 2018, DOI 10.1074/jbc.RA117.001536

Koichi Abe (阿部 紘一)***, Naoki Sunagawa (猿川 直輝)*, Tohru Terada (寺田 透)*, Yuta Takahashi (高橋 佑汰)**†, Takatoshi Arakawa (荒川 孝俊)*, Kiyohiko Igarashi (五十嵐 圭子)*****, Masahiro Samejima (斎島 正浩)**†, Hiroyuki Nakai (中井 博之)*, Hayao Taguchi (田口 俊男)**†, Masahiro Nakajima (中島 将博)**†, and Shinya Fushinobu (伏信 進矢)**†

From the **Department of Biotechnology, †Agricultural Bioinformatics Research Unit, and ‡Department of Biomolecular Sciences, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan, the **Faculty of Agriculture, Niigata University, Niigata 950-2181, Japan, the ***VT Technical Research Centre of Finland, Espoo FI-02044 VTT, Finland, and the ††Department of Applied Biological Science, Faculty of Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan

Edited by Gerald W. Hart

\( \beta \)-1,2-Glucans are bacterial carbohydrates that exist in cyclic or linear forms and play an important role in infections and symbioses involving Gram-negative bacteria. Although several \( \beta \)-1,2-glucan–associated enzymes have been characterized, little is known about how \( \beta \)-1,2-glucan and its shorter oligosaccharides (Sop\(_n\)s) are captured and imported into the bacterial cell. Here, we report the biochemical and structural characteristics of the Sop\(_n\)-binding protein (SO-BP, Lin1841) associated with the ATP-binding cassette (ABC) transporter from the Gram-positive bacterium *Listeria innocua*. Calorimetric analysis revealed that SO-BP specifically binds to Sop\(_n\)s with a degree of polymerization of 3 or more, with \( K_d \) values in the micromolar range. The crystal structures of SO-BP in an unliganded open form and in closed complexes with tri-, tetra-, and pentaoligosaccharides (Sop\(_n\)–2) were determined to a maximum resolution of 1.6 Å. The binding site displayed shape complementarity to Sop\(_n\), which adopted a zigzag conformation. We noted that water-mediated hydrogen bonds and stacking interactions play a pivotal role in the recognition of Sop\(_n\) by SO-BP, consistent with its binding thermodynamics. Computational free-energy calculations and a mutational analysis confirmed that interactions with the third glucose moiety of Sop\(_n\)s are significantly responsible for ligand binding. A reduction in unfavorable changes in binding entropy that were in proportion to the lengths of the Sop\(_n\)s was explained by conformational entropy changes. Phylogenetic and sequence analyses indicated that SO-BP ABC transporter homologs, glycoside hydrolases, and other related proteins are co-localized in the genomes of several bacteria. This study may improve our understanding of bacterial \( \beta \)-1,2-glucan metabolism and promote the discovery of unidentified \( \beta \)-1,2-glucan–associated proteins.

ATP-binding cassette (ABC)\(^2\)-type transporters are widely distributed in living organisms, forming one of the largest protein superfamilies. ABC transporters utilize the free energy obtained from ATP hydrolysis to import or export a wide variety of molecules across cellular membranes. They share a common architecture consisting of two transmembrane domains (TMDs) and intracellular nucleotide-binding domains (NBDs). In bacterial ABC importers, an additional domain, a solute (or substrate)-binding protein (SBP), serves as an initial receptor that specifically binds to ligands with high affinity, delivers them to TMDs, and stimulates the ATPase activity (1). SBPs from Gram-negative bacteria are located in the periplasm, whereas those from Gram-positive bacteria are anchored at the cell surface (2). SBPs are essential for the active transport of their ligands (3, 4).

The crystal structures of SBPs have revealed that the domain composition of SBPs is conserved despite their low sequence similarity and widely divergent molecular masses (25–70 kDa) (5). In general, the overall structure comprises two globular \( \alpha/\beta \) domains consisting of a central \( \beta \)-sheet flanked by \( \alpha \)-helices. The two domains are linked by a hinge, and ligand binding takes place between the two domains. In the absence of ligands, the two domains can move flexibly around the hinge, and an open conformation, in which the two domains are separated, is predominant (6). Upon ligand binding, the two domains get close to each other and are stabilized in a closed conformation. This open-close conformational transition has been called the “Venus Fly Trap” mechanism (7). Although many SBPs have been structurally and func-

This work was supported in part by the Platform Project for Support in Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from the Japan Agency for Medical Research and Development (AMED) and, by JSPS-KAKENHI Grants 15H02443 and 26660083 (to S. F.). The authors declare that they have no conflicts of interest with the contents of this article.

The atomic coordinates and structure factors (codes SYSB, SYSD, SYSE, and SYSF) have been deposited in the Protein Data Bank ([http://wwpdb.org/](http://wwpdb.org/))

1 To whom correspondence should be addressed. Tel: 81-3-5841-5151; Fax: 81-3-5841-5151; E-mail: asfushi@mail.ecc.u-tokyo.ac.jp

2 The abbreviations used are: ABC, ATP-binding cassette; TMD, transmembrane domain; NBD, nucleotide-binding domain; SBP, solute (or substrate)-binding protein; Sop\(_n\), sophorooligosaccharide; DP, degree of polymerization; LiSOGP, *L. innocua* 1,2-\( \beta \)-oligoglucan phosphorylase; LiBGL, *L. innocua* \( \beta \)-glucosidase; GH, glycoside hydrolase; ITC, isothermal calorimetry; MD, molecular dynamics; SO-BP, Sop\(_n\)-binding protein; r.m.s.d., root mean square deviation; PCA, principal component analysis; PDB, Protein Data Bank.

© 2018 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.
tionally characterized, full evaluations of many of them remain elusive because of the great diversity of their ligands.

β-1,2-Glucan is an extracellular polysaccharide predominantly found in a cyclic form in the periplasmic space of the α-2 subdivision of proteobacteria, containing the plant pathogen *Agrobacterium*, the plant symbiote *Rhizobium*, and the mamalian pathogen *Brucella* (8–10). These bacteria utilize cyclic β-1,2-glucans for infection (*Agrobacterium* and *Brucella*) or symbiosis (*Rhizobium*) by inhibiting host defense systems, whereas several other bacteria utilize them for adaptation to osmotic changes (8, 11). Linear β-1,2-glucans are also produced by several bacteria. *Escherichia coli* and *Pseudomonas aeruginosa* synthesize shorter β-1,2-glucans branched with β-1,6-glucosidic bonds for osmotic regulation (8, 12, 13). β-1,2-Gluco-oligosaccharides (sophorooligosaccharides; Sop*) where n denotes the degree of polymerization (DP) is found in nature as a disaccharide unit, such as sophorolipids and plant spheroids (14, 15). Sophorose (Sop2) is a well-established powerful cellulase inducer of *Trichoderma reesei* (16). Despite the importance of their physiological roles, little is known about metabolic enzymes and proteins for β-1,2-glucans and Sop*s because of their limited availability.

We previously discovered the gene cluster involved in the dissimilation of Sop*s in the Gram-positive bacteria *Listeria innocua* (Lin1838–1843) and characterized two cytosolic Sop*s-degrading enzymes: 1,2-β-oligoglucan phosphorylase (Lin1839 and LiSOGP) and β-glucosidase (Lin1840). LiSOGP belongs to glucoside hydrolase (GH) family 94, catalyzing reversible phosphorylization of Sop* (*n* ≥ 3) to release α-glucose 1-phosphate (17). Lin1840 is a GH3 β-glucosidase exhibiting marked specificity toward Sop* (18). A putative LacI family transcriptional regulator (Lin1838) is also co-localized in the cluster. Together with these gene products, putative ABC transporter components (Lin1841–1843) are encoded: Lin1841 and Lin1842–1843 encode an SNP and TMDs, respectively. However, it is unknown whether this transporter is responsible for uptake of Sop*s.

In this study, we focus on the Lin1841 protein to gain insights into the bacterial uptake of Sop*. We show that Lin1841 specifically binds to Sop*s with a DP of 3 or more and clarify its binding thermodynamics and its structural basis by X-ray crystallography, isothermal calorimetry (ITC), and molecular dynamics (MD) simulation.

### Results

#### Binding thermodynamics of Lin1841

The binding thermodynamic parameters of Lin1841 for various carbohydrates (list of ligands under “Experimental procedures”) were determined by ITC (Table 1 and Fig. 1). Among the tested oligosaccharides, tight interactions were observed for Sop3, Sop4, and Sop5, and their isotherms exhibited typical sigmoidal curves (Fig. 1, B–D). The binding constants (*K*<sub>b</sub>) were on the order of 10<sup>4–6</sup> M<sup>−1</sup>, and the binding of Sop3 showed the highest affinity (*K*<sub>b</sub> = 1.72 × 10<sup>6</sup> M<sup>−1</sup> at 20 °C), which is in the same range as those of other SBPs (5). No heat pulses were observed for Sop2 (Fig. 1A), laminarintriose, cellotriose, gentio-oligosaccharides, or maltotriose (data not shown), demonstrating that the Lin1841 protein is specific for Sop*s with a DP ≥ 3. To date, no SNP that can bind to β-1,2-linked glucosides has been reported. The chain length specificity of Lin1841 is consistent with that of LiSOGP (17). Thus, we named Lin1841 the “Sop*′ binding protein” (SO-BP). SO-BP also bound to larger linear β-1,2-glucans (average DP 25, Fig. 1E). The binding isotherm was not clear, likely due to impurity of the linear β-1,2-glucan sample. The affinity was estimated to be ~10–1000-fold weaker than that of Sop3–5, indicating that SO-BP prefers oligosaccharides rather than polysaccharides. All binding interactions were enthalpy-driven with unfavorable entropy changes, and each of the stoichiometries (*n*) was almost 1:1. This slight deviation from a 1:1 stoichiometry may be due to impurities of Sop*s. The favorable enthalpy changes increased with an increase in temperature and were not able to sufficiently compensate for increases of the entropic penalties, resulting in slight

### Table 1

| Protein | Ligand | *K*<sub>b</sub> × 10<sup>6</sup> M<sup>−1</sup> | Δ*G*° | Δ*H* | Δ*S*° | Δ*C*<sub>p</sub> | n |
|---------|--------|------------------------|--------|-------|--------|----------------|---|
| WT      | Sop2   | 0.339 ± 0.003          | −7.29  | −23.3 | −16.0  | −288 ± 43      | 0.909 ± 0.001 |
|         |        | 0.174 ± 0.001          | −7.03  | −24.9 | −17.9  |                 | 0.901 ± 0.002 |
|         |        | 0.0883 ± 0.0007        | −6.75  | −25.0 | −18.3  |                 | 0.866 ± 0.002 |
|         |        | 0.0417 ± 0.0008        | −6.40  | −27.2 | −20.8  |                 | 0.937 ± 0.009 |
|         |        | 0.0193 ± 0.0007        | −6.04  | −29.3 | −23.3  |                 | 0.870 ± 0.020 |
| D193A   | Q197A  | 0.0201 ± 0.0005        | −5.87  | −24.8 | −18.9  |                 | 0.946 ± 0.007 |
|         |        | 0.00043 ± 0.00028      | −4.96  | −11.7 | −6.74  |                 | 1.000 |
| WT      | Sop3   | 1.72 ± 0.06            | −8.36  | −18.8 | −10.4  | −295 ± 20       | 1.150 ± 0.002 |
|         |        | 1.34 ± 0.06            | −8.36  | −20.2 | −11.8  |                 | 1.150 ± 0.002 |
|         |        | 0.822 ± 0.020          | −8.20  | −22.0 | −13.8  |                 | 1.100 ± 0.002 |
|         |        | 0.441 ± 0.009          | −7.96  | −23.1 | −15.1  |                 | 1.080 ± 0.002 |
| WT      | Sop4   | 1.53 ± 0.07            | −8.15  | −14.4 | −6.25  | −194 ± 22       | 1.390 ± 0.003 |
|         |        | 1.21 ± 0.04            | −8.16  | −15.0 | −6.84  |                 | 1.110 ± 0.002 |
|         |        | 0.762 ± 0.020          | −8.02  | −16.0 | −7.98  |                 | 1.130 ± 0.002 |
|         |        | 0.442 ± 0.010          | −7.83  | −17.3 | −9.47  |                 | 1.120 ± 0.002 |
|         |        | 0.232 ± 0.006          | −7.56  | −16.6 | −9.04  |                 | 1.210 ± 0.003 |
| WT      | β-1,2-Glucan<sup>b</sup> (average DP 25) | 0.00621 ± 0.00007 | (−5.17) | −174.2 | −15.1 | 0.867 ± 0.323 |

<sup>a</sup> Regression analysis was performed with the fixed ν value because the low affinity of the mutant protein limited the protein concentration setting of the titration experiment, resulting in the low c value (*ν* × protein concentration × *K*<sub>b</sub> < 1).

<sup>b</sup> The molar concentration was calculated based on the average DP.

<sup>c</sup> Uninterpretable heat pulses made it difficult to determine accurate values.
decreases of the Gibbs free energy change at higher temperatures (Table 1).

We also measured the change in the molar heat capacity with ligand binding ($\Delta C_p$) by plotting $\Delta H$ versus temperature to estimate the detailed binding thermodynamics of SO-BP (Fig. 1F). The $\Delta C_p$ values for Sop3–5 were approximately $-300$ to $-200$ cal mol$^{-1}$ K$^{-1}$. A previous study showed that stacking interactions between one aromatic residue and one sugar ring give $\Delta C_p$ values of $-150$ to $-100$ cal mol$^{-1}$ K$^{-1}$ (19). Therefore, it was predicted that two or three aromatic residues are involved in the stacking interactions with Sop3–5.

Overall structure of SO-BP

The crystal structures of SO-BP in a ligand-free form and in complexes with Sop3, Sop4, and Sop5 were determined at 2.2 to 1.6 Å resolutions (Table 2). The asymmetric units of the crystals contained two molecules. The protein construct used for crystallization consisted of residues 26–422 without the signal peptide. There were no disordered regions in all protein structures, except for the N and C termini; residues 35–420 (ligand-free, chain A), 36–420 (ligand-free, chain B), 34–420 (Sop3 complex, chain A), 34–422 (Sop3 and Sop5 complex, chain B), 35–421 (Sop4 complex, chain A), 36–422 (Sop4 complex, chain B), and 34–421 (Sop5 complex, chains A and B) were modeled. The molecules in an asymmetric unit of each structure were virtually identical, with a root mean square deviation (r.m.s.d.) for the Cα atoms of $<0.2$ Å. Hereafter, we focused on chain A. The overall structure of SO-BP adopts a typical SBP-fold consisting of two globular $\alpha/\beta$ domains as follows: N-terminal domain I (residues 34–145 and 301–349) and C-terminal domain II (residues 148–298 and 352–416) connected by hinge regions (residues 146–147, 299–300, and 350–351) (Fig. 2, A, B, and D). SO-BP is classified in structural cluster D-I of SBPs (5). The complex structures with Sop3–5 could be overlaid well,

Figure 1. Isothermal titration curves at 25 °C for binding of Lin1841 with Sop2–5 and β-1,2-glucan (A–E) and the binding enthalpy changes at different temperatures (F). A–E top and bottom panels indicate a thermogram and a binding isotherm, respectively. F, error bars represent the fitting errors of linear regression analysis. For linearity, the plot of the enthalpy change for Sop3 at 35 °C is excluded.
with a r.m.s.d. of <0.4 Å relative to each other. When r.m.s.d. values were calculated separately for domains I and II, the r.m.s.d. values of the Ca atoms between the ligand-free and the Sop4 complex were 0.9 and 0.3 Å, respectively. In contrast, because of a large hinge motion between the ligand-free (open, Fig. 2A) and complex (closed, Fig. 2B) structures, attempts to overlay the whole protein in these two states provided Ca r.m.s.d. values above 1.6 Å. Superimposition of domain II clearly showed a hinge motion (Fig. 2C). A domain movement analysis using the DynDom server (20) indicated that the inter-domain rotation angle was 26.1° (97.5% closure), and the translation movement was only 0.3 Å. Although a part of the first α-helix (residues 44–47) of the open state deforms due to the absence of interactions between Glu-45 and the ligands (Fig. 3), it is likely that the conformational change occurs through rigid body rotation, similar to other SBPs (21). DALI structural homology search revealed that the structure of SO-BP in complex with Sop4 was most similar to Xac-MaLE (a putative maltose/trehalose-binding protein, PDB code 3UOR) from Xanthomonas citri (Xanthomonas axonopodis pv. citri strain 306), with Z-score of 46.6, r.m.s.d. for the 386 Ca atoms of 3.2 Å, and a sequence identity of 34%. The other DALI hits included a trehalose/maltose-binding protein from Thermococcus litoralis (PDB code 1EU8; Z-score = 43.0, r.m.s.d. for the 373 Ca atoms = 2.3 Å, and sequence identity = 24%), and the solute-binding protein Lmo0181 from Listeria monocytogenes (PDB code 5F7V; Z-score = 41.8, r.m.s.d. for the 372 Ca atoms = 2.2 Å, and sequence identity = 20%). When the open form SO-BP was used for the homology search, similar hits were obtained (data not shown).

**MD simulation**

To explore possible ligand-free conformations of SO-BP that were not observed in the crystal structures, we performed 100-ns MD simulation. Principal component analysis (PCA) was also performed to examine the conformational distribution in the MD trajectory (Fig. 2, E and F). We found two representative structures (state 1 and 2) based on PCA. State 1 (Fig. 2) and state 2 (shown in orange) was similar to the open state crystal structure (Ca r.m.s.d. <0.8 Å), whereas state 2 (shown in blue) was subtly different from the crystal structure (Ca r.m.s.d. >1.3 Å). The Ca r.m.s.d. value and the difference in the interdomain bend angle between states 1 and 2 were 1.1 Å and 88°, respectively. Moreover, a domain movement analysis using DynDom indicated that the bend angle difference between the closed state crystal structure and state 2 was 32.5°, suggesting that the ligand-free SO-BP can adopt a more open conformation in aqueous solution.

**Architecture of the ligand-binding site**

The SO-BP structures obtained from co-crystals with Sop3–5s delineated the positions of the corresponding sugars at

---

**Table 2**

| Data collection and refinement statistics | Data set |
|------------------------------------------|----------|
| **Ligand-free** | **Sop4 complex** | **Sop4 complex** | **Sop4 complex** |
| P3_2 | P2_1 | P1 | P2_1 |
| Space group | a = b = 74.9, c = 120.8 | a = 36.1, b = 125.6, c = 91.8 | a = 36.0, b = 63.3, c = 87.5 | a = 36.2, b = 126.0, c = 91.0 |
| Cell dimensions (Å) | | | | |
| Angles (°) | α = 101.6 | α = 90.0, β = 82.2, γ = 86.1 | β = 100.5 |
| Resolution (Å)* | 50.0–2.20 (2.24–2.20) | 50.0–2.10 (2.14–2.10) | 50.0–1.60 (1.63–1.60) | 50.0–1.90 (1.93–1.90) |
| Total reflections | 206,593 | 160,585 | 337,940 | 232,606 |
| Unique reflections# | 38,141 (1,912) | 44,909 (2,234) | 98,051 (4,453) | 61,562 (2,964) |
| Completeness (%)# | 98.8 (100.0) | 96.5 (96.5) | 96.0 (87.0) | 97.5 (95.6) |
| Redundancy* | 5.5 (5.7) | 3.6 (3.5) | 3.4 (3.3) | 3.8 (3.7) |
| Mean I/σ(I) | 21.0 (2.7) | 12.4 (1.7) | 20.7 (1.5) | 16.2 (2.1) |
| Rmerge (%) | 9.7 (57.1) | 11.6 (63.9) | 6.8 (73.7) | 9.9 (54.5) |
| CC1/2,* | (0.906) | (0.697) | (0.617) | (0.824) |

---

*Values in parentheses denote the highest resolution shell.
the center of the cleft enclosed by domains I and II. We defined glucose units of Sop\textsubscript{n} as A–E from the nonreducing end (Fig. 3, A, C and E). The complex structures showed that Sop\textsubscript{3–5}s adopt a zigzag conformation, bind to SO-BP in a similar manner, and share units from the nonreducing end. Units A and B are sequestered from solvent, whereas units at the reducing end (C–E in Sop\textsubscript{3–5}s) are exposed to solvent (Fig. 2B). Two tryptophan residues, Trp-71 and Trp-268, function in stacking with units A and B and the opposite side of unit B, respectively. This observation is in accordance with the $\Delta C_p$ values of Sop\textsubscript{n} binding, as revealed by ITC (Table 1). Trp-97 forms hydrophobic interactions with the C6 hydroxymethyl group of unit A.

Hydrogen bonds play a pivotal role in stabilizing the SO-BP–Sop\textsubscript{n} complexes (Fig. 3, B, D and F). Thr-95, Thr-96, Glu-147, and Gly-301 form hydrogen bonds with the hydroxyl groups of unit A. Glu-45 and Tyr-145 form hydrogen bonds with the hydroxyl groups of unit B. Asp-193 and Gln-197 form hydrogen bonds with the hydroxyl groups at unit C. These interactions are shared among the Sop\textsubscript{n}-bound structures. In particular, the
hydrogen bonds of Asp-193 and Gln-197 toward unit C appear to be indispensable for ligand binding because Sop2 was not able to bind to SO-BP (Fig. 1A). The anomeric hydroxyl group of the glucose moiety at unit D in the Sop4 complex forms a hydrogen bond with the Nε1 atom of Trp-268. The C4 hydroxyl group at unit E forms an additional hydrogen bond with the side chain of Ser-265. Eighteen-, 36-, and 21-ordered waters are found within 5 Å of Sop3–5s, respectively, and thus water-mediated hydrogen bonds also substantially contribute to the recognition of Sopn5. These water molecules are held in place by hydrogen bonds with Met-42, Asp-44, Glu-45, Trp-71, Thr-96, Glu-147, Arg-149, Asp-193, Gln-195, Glu-196, Glu-197, Met-264, Trp-268, Gly-300, and Glu-377 (Fig. 3, B, D, and F).

**Figure 3. Interactions of SO-BP with Sop3 (A and B), Sop4 (C and D), and Sop5 (E and F).** A, C, and E, nA-weighted mFo – DFc maps (countered at 3.0 σ) of the bound Sopn5s are delineated. Sop3–5s are colored as green, cyan, and yellow, respectively. Glucose units are defined as A–E from the nonreducing end of each sugar. B, D, and F, stereoviews of the ligand-binding sites. Dotted lines indicate polar interactions. Polar interactions are drawn using a threshold of 3.3 Å.

**Contribution of interactions at unit C to complex stability**

To estimate the energetic contributions of interactions at unit C to the binding to SO-BP, the difference in free energy changes of binding (ΔΔG) was calculated based on MD simulations (Fig. 4). We first calculated the ΔΔG value between Sop3 and Sop2. This calculation demonstrated that binding of Sop3 was more energetically favorable than that of Sop2, with a ΔΔG of −3.59 ± 0.59 kcal mol⁻¹ (Fig. 4, A and C). By subtracting the ΔΔG value from the binding free energy of Sop3 at 25 °C, the free energy change in Sop2 binding was estimated to be −3.16 kcal mol⁻¹ (Kd = 2.08 × 10² M⁻¹), which was below the detection limit of ITC. Indeed, no heat signal was observed in the titration of Sop2 (Fig. 1A). We also calculated the ΔΔG value between WT SO-BP and a Q197A mutant. To simply compare the effects of mutation, the ligand-free structures (WT and Q197A) were postulated to be in the closed form. The calculation showed that WT SO-BP exhibited more stable binding...
Binding protein specific for β-1,2-glucooligosaccharides

\[
\Delta G_1 + \Delta G_{\text{bind}(\text{Sop}_2)} = \Delta G_{\text{bind}(\text{Sop}_3)} + \Delta G_2 \\
\Delta G = \Delta G_{\text{bind}(\text{Sop}_3)} - \Delta G_{\text{bind}(\text{Sop}_2)} = \Delta G_1 - \Delta G_2
\]

\[
\Delta G_3 + \Delta G_{\text{bind}(\text{Q197A})} = \Delta G_{\text{bind}(\text{WT})} + \Delta G_4 \\
\Delta G = \Delta G_{\text{bind}(\text{WT})} - \Delta G_{\text{bind}(\text{Q197A})} = \Delta G_3 - \Delta G_4
\]

| Method                  | \( \Delta G \)                | kcal mol\(^{-1}\) |
|-------------------------|-------------------------------|-----------------|
| Computational           | \( \Delta G_{\text{bind}(\text{Sop}_3)} - \Delta G_{\text{bind}(\text{Sop}_2)} \) | -3.59 ± 0.59    |
| Experimental            | \( \Delta G_{\text{bind}(\text{WT})} - \Delta G_{\text{bind}(\text{D193A})} \)  | -0.88           |
| Computational           | \( \Delta G_{\text{bind}(\text{WT})} - \Delta G_{\text{bind}(\text{Q197A})} \)  | -2.34 ± 0.23    |
| Experimental            | \( \Delta G_{\text{bind}(\text{WT})} - \Delta G_{\text{bind}(\text{Q197A})} \)  | -1.79           |
than Q197A, with a $\Delta \Delta G$ of $-2.34 \pm 0.23$ kcal mol$^{-1}$ (Fig. 4, B and C). In the same way as described above, the free energy change of Q197A in Sop$_3$ binding was estimated to be $-4.41$ kcal mol$^{-1}$ ($K_a = 1.72 \times 10^4$ M$^{-1}$). The calculation for the D193A mutant was not performed because the free energy change for a mutation accompanying a change in the total charge was not calculated accurately.

In addition to the free energy calculations, we determined the energetic contributions of Asp-193 and Gln-197 in Sop$_3$ binding by ITC (Table 1 and Fig. 4, C–E). A mutation at Asp-193 (D193A) resulted in more than a 4-fold reduction in the $K_a$ value (2.01 $\times$ 10$^6$ M$^{-1}$) compared with WT. A mutation at Gln-197 (Q197A) gave the $K_a$ value of 4.34 $\times$ 10$^4$ M$^{-1}$, consistent with the $K_a$ value (1.72 $\times$ 10$^5$ M$^{-1}$) estimated by the above calculation. These computational and experimental results indicate that the interactions at unit C have a significant contribution to the binding free energy of Sop$_3$ ($-6.75$ kcal mol$^{-1}$ at 25 °C), and the contribution of Gln-197 is larger than that of Asp-193.

**Conformation of bound ligands**

Distributions in the conformations of Sop$_n$s during MD simulations were analyzed to investigate the difference in the conformations between the Sop$_n$s bound to SO-BP and those free in aqueous solution (Fig. 5). Fig. 5 maps the glycosidic bond dihedral angles ($\phi$ and $\psi$) of Sop$_n$s. Hereafter, the Sop$_n$s moieties of Sop$_n$s are referred to as Sop$_n$ (xy), where xy denotes two of the monomer units in aqueous solution or in the binding site of SO-BP. The $\phi$ and $\psi$ of the bound Sop$_3$ (AB) converged around the conformation in the co-crystal (shown as black circle), with $\phi = -120$ to $-80$° and $\psi = 120$ to 180°. The average $\phi$ and $\psi$ angles were $-105 \pm 8$ and $145 \pm 8$°, respectively. These results indicate that the interactions between SO-BP and units A and B stabilized the Sop$_3$ moiety during the MD simulations in a similar conformation as the co-crystal structure (Fig. 5A). In contrast, those of the free Sop$_3$ (AB) were distributed in wide ranges, with $\phi = -180$ to $-60$° and $\psi = 60$ to 180°. The average $\phi$ and $\psi$ angles were $-91 \pm 27$ and $120 \pm 22$°, respectively, which are almost identical to the average values for the free Sop$_3$ reported previously by simulations and experiments ($\phi = -83.8$ to $-72$° and $\psi = 110$ to 128.5°) (23–27). Thus, the bound Sop$_3$ (AB) and the free Sop$_3$ (AB) were modestly different in their conformations. However, the bound Sop$_3$ (AB) adopts a sufficiently low free energy conformation according to the free energy map of Sop$_3$, obtained from the local elevation umbrella sampling (LEUS) method (27). This suggests that the affinity of SO-BP is not significantly influenced by the conformation of Sop$_3$. Similar tendencies were observed in the Sop$_{4, 5}$ (AB) (Fig. 5, B and C).

The $\phi$ and $\psi$ of the bound Sop$_3$ (BC) varied but were populated around the Sop$_3$ (BC) in the co-crystal ($\phi = -150$ to $-90$° and $\psi = 60$ to 120°) with the average angles of $\phi = -113 \pm 13$ and $\psi = 84 \pm 10$° (Fig. 5A). The free Sop$_3$ (BC) diverged in a similar manner as the free Sop$_3$ (AB) with the average angles of $\phi = -84 \pm 18$ and $\psi = 97 \pm 27$°. These angle distributions are also in the stable range shown in the LEUS method. In contrast to the bound Sop$_3$ (BC), the conformations of the bound Sop$_{4, 5}$ (BC) were only populated around the co-crystal structure (Fig. 5, B and C). The $\phi$ and $\psi$ of the free Sop$_{4, 5}$ (BC) showed the same conformational properties but did not populate around $\phi = -75$ and $\psi = 75$° like those of the free Sop$_3$ (BC) (Fig. 5, A–C).

The bound Sop$_{4, 5}$ (CD) showed similar conformational properties as the free Sop$_{4, 5}$ (CD) and were distributed around the structures of free Sop$_3$ (Fig. 5, B and C). In contrast to Sop$_{3, 5}$ (AB, BC, and CD), the conformations of the bound Sop$_3$ (DE) were extremely divergent and showed a similar conformation to the free Sop$_3$ (DE). These results suggest that SO-BP restricts the conformations of Sop$_n$s via units A–D site-specific interactions, which makes Sop$_n$s more stable than when free in solution.

Conformations of the hydroxymethyl groups of Sop$_n$s in the co-crystal structures and MD simulations were also examined (Table 3). Almost all the hydroxymethyl groups of the bound Sop$_n$s (at units A and B both in the co-crystals and simulations) adopted the gg conformation, whereas those of the free Sop$_n$s (A and B) were mainly in the gt conformation. The gg conformation of unit A was fixed by the water-mediated hydrogen bonds and that of unit B was stabilized by the hydrogen bond to the Oe1 atom of Glu-45 (Fig. 3, B, D, and F). The hydroxymethyl group of the unit C glucose moiety of the bound Sop$_{3, 4}$ mainly adopted the gt conformation, which is generally consistent with the free Sop$_{3, 4}$. In contrast, the hydroxymethyl group of unit C of the bound Sop$_3$ adopted primarily gg (in MD simulation) and gt (in the co-crystal) conformations. During part of the simulation time, the gg conformation was stabilized by a hydrogen bond with the Oδ2 atom of Asp-193 (data not shown). The hydroxymethyl groups of units D and E of the bound Sop$_{4, 5}$ mostly adopted a gt conformation in the MD simulation. In the co-crystal structures, the hydroxymethyl group of unit D of the Sop$_4$ adopted a gg conformation and that of unit E of the Sop$_5$ adopted an intermediate conformation between gt and gg. These hydroxymethyl groups did not form any direct interactions with SO-BP in the crystal structure (Fig. 3, B, D, and F) or during the MD simulations. These differences in the conformation of hydroxymethyl groups appeared to not affect the changes in binding free energy because the free energy difference between gt and gg conformations in glucose is estimated to be 0.1 kcal mol$^{-1}$ (28), but it may affect the differences in binding enthalpy and entropy changes for Sop$_n$s.

**Distribution of SO-BP homologs and β-1,2-glucan utilization loci**

We examined the distribution of homologous sequences of SO-BP (amino acid sequence identity >30%) and found that

---

*Figure 4. Energetic contributions of unit C to ligand binding. A and B, thermodynamic cycles for free energy calculations. The calculations of free energy ($\Delta G$) were based on the Bennett acceptance ratio method (see "Experimental procedures" for details). A, Sop$_3$ in solution and in complex with protein was transformed into Sop$_3$ to calculate $\Delta G_1$ and $\Delta G_2$, respectively. $\Delta G_3$ was calculated as the subtraction of $\Delta G_2$ from $\Delta G_1$. B, Gln-197 of the closed SO-BP in ligand-free and complexed form was transformed into alanine to calculate $\Delta G_2$ and $\Delta G_3$, respectively. $\Delta G_4$ was calculated as the subtraction of $\Delta G_3$ from $\Delta G_2$. C and D, ITC analysis of D193A and Q197A mutants at 25 °C. Top and bottom panels indicate a thermogram and a binding isotherm, respectively. In the bottom panels, plots deviated from a sigmoidal curve were removed for regression analysis. E, $\Delta G$ values estimated by a computational or an experimental method.*
orthologs of SO-BP are distributed largely in Firmicutes, Proteobacteria, and Actinobacteria with distinct homology (Fig. 6A). These bacteria are ubiquitously distributed in the environment, such as in soil, food, the sea, and the intestinal tract of animals. SO-BP homologs in Firmicutes are found in various genera, including mainly *Listeria*, *Bacillus*, and *Paenibacillus*. The SO-BP homologs in Actinobacteria are predominantly derived from *Streptomyces*, and those in Proteobacteria are distributed in almost all γ-Proteobacteria, including *Xanthomonas* and its related species.

Furthermore, we examined gene clusters containing putative SO-BP genes to understand the role of SO-BP in β-1,2-glucan or Sop$_n$s metabolism. Six representative gene loci were selected

![Figure 5. Distributions of the glycosidic bond dihedral angles during MD simulations.](image)

**A** Sop$_3$ (AB) and Sop$_3$ (BC)

**B** Sop$_4$ (AB), Sop$_4$ (BC), and Sop$_4$ (CD)

**C** Sop$_5$ (AB) and Sop$_5$ (BC)

**Table 3**

Conformations of the hydroxymethyl groups of Sop$_n$s

| Ligand | Unit | In the co-crystal | MD simulation $gg$/$tg$/$gt$ (complex/free) |
|--------|------|-------------------|-------------------------------------------|
| Sop$_3$ A | $gg$ | 95:0:5/11:1:85 | |
| B | $gg$ | 100:0:0/23:1:76 | |
| C | $gt$ | 62:0:38/20:1:79 | |
| Sop$_4$ A | $gg$ | 97:0:3/28:1:71 | |
| B | $gg$ | 100:0:0/23:1:76 | |
| C | $gt$ | 27:0:73/32:2:66 | |
| D | $gg$ | 22:1:77/33:1:66 | |
| E | $gg$ | 33:1:66 | |
| Sop$_5$ A | $gg$ | 95:0:5/11:1:85 | |
| B | $gg$ | 100:0:0/23:1:76 | |
| C | $gt$ | 62:0:38/20:1:79 | |
Figure 6. Distribution of SO-BP homologs. A, bootstrap consensus phylogenetic tree of SO-BP homologs. The phylogenetic tree was constructed using the neighbor-joining method. The bootstrap analysis was carried out with 100 resamplings of the data set. Species possessing SO-BP homologs are divided into four clusters based on phyla and encompassed with solid lines. L. innocua is denoted by a closed circle. The numbers denoted correspond to the gene organizations shown in B. B, /H9252-1,2-glucan utilization loci in various organisms aligned with an ABC transporter. ABC transporter components are shown with a black background and white letters. GHs are white, and LacI transcriptional regulators are light gray. Hypothetical membrane proteins are dark gray. According to SignalP 4.1 (67) and LipoP (68), GH94, GH3, and GH1 enzymes and LacI regulators do not have signal peptides, whereas the hypothetical proteins have type I signal peptides. GH144 enzymes appear to possess the type II signal peptides.
to cover all groups in the phylogenetic tree. Two TMDs adjacent to SO-BP are strictly conserved in the gene clusters (Fig. 6B). In the Firmicutes clade, GH94 SOGP and GH3 BGL homologs tend to co-occur in the cluster. In the Actinobacteria and Proteobacteria clades, GH1 and GH144 enzymes frequently co-occur in the cluster, and the GH144 enzymes are likely to have endo-β-1,2-glucanase activity (29). LacI transcriptional regulators also tend to co-localize together with these gene clusters.

Amino acid residue conservation

The degree of amino acid residue conservation of the homologs (Fig. 6A) was mapped on the surface model of SO-BP (Fig. 7A). Highly conserved residues (Fig. 7A, shown in red) are located not only in the binding site but also on the surface of domain I. The conserved patch in domain I corresponds to the region of the maltose/maltodextrin-binding protein from E. coli to interact with the transmembrane MalG (30), implying that these residues may also be responsible for interactions with the associated TMDs of the ABC transporter.

Conservation of the binding site is shown in Fig. 7B. The residues with high conservation scores are mainly located on the side of domain I, and the less-conserved residues are located on the opposite side. Notably, Asp-193 and Gln-197, which are important for ligand binding, are variable in distant homologs (Asp-193 is typically substituted with Pro or Gly), but they are mostly conserved in Listeria and its related species (data not shown).

Discussion

Because β-1,2-glucans are not abundant in nature, characterization of β-1,2-glucan–associated proteins has not progressed compared with other glucan–associated proteins. Our previous studies overcame this challenge by enzymatically synthesizing linear β-1,2-glucan using LiSOGP, leading to identification of β-1,2-glucan–degrading enzymes (18, 29, 31). However, how β-1,2-glucan or Sop₅s are captured and imported inside the cells remains unknown. In this study, we focused on the Lin1841 protein in the Sop₅ utilization locus of L. innocua and revealed its thermodynamic characteristics based on ligand binding and structural analysis. MD simulations supported the structural and thermodynamics data.

Specificity of SO-BP

Our free energy calculations and ITC analysis demonstrate that a key factor determining the specificity for chain length of ligands is the polar interactions toward unit C (Figs. 3, B, D, and F, and 4). This binding specificity of SO-BP is consistent with that of LiSOGP (GH94 phosphorylase), but the mechanism is different. SO-BP binds to Sop₅s with energetically favorable conformations, whereas SOGP from Lachnoclostridium phytofermentans (LiSOGP homolog) is considered to bind to a disaccharide unit of Sop₅ with an unfavorable conformation at subsites −1 and +1, which is compensated by favorable binding of the third glucose unit at subsite +2 (31).

Binding energetics of SO-BP

Binding of Sop₅s is driven by favorable enthalpy changes that are accompanied by unfavorable entropy changes and negative heat capacity changes. Such a calorimetric behavior is generally observed among carbohydrate-interacting proteins (32–35). This behavior is also supported by the present observations that the Sop₃₋₅s are fixed by ordered water molecules and two stacker tryptophan residues (Fig. 3, B, D, and F).

The stacking of the two tryptophan residues is expected to give similar ΔCₚ values (≈−150 to −100 cal mol⁻¹ K⁻¹) per a sugar-aromatic residue pair) (19). However, the ΔCₚ values for Sop₃₋₄ were approximately −300 cal mol⁻¹ K⁻¹, whereas that of Sop₅ was approximately −200 cal mol⁻¹ K⁻¹ (Table 1). ΔCₚ is a sensitive parameter for changes in the solvent environment, and solvation of polar groups causes negative ΔCₚ values (36). In the SO-BP–Sop₅ complex, the reducing end glucose moiety of Sop₅ appears to interfere with the water-mediated hydrogen bond networks around Asp-44 and Asp-193, which is also observed in the Sop₅ complex (Fig. 3, D and F). These polar interactions would contribute to the more negative ΔCₚ values for Sop₃₋₄ than for Sop₅.
The ΔH and ΔS values at each temperature increased linearly with DP; therefore, each glucose unit extended from Sop₃ must contribute to the unfavorable ΔH and the favorable ΔS (Table 1). Although differences in ΔS values according to the lengths of Sop₃ are dependent on ΔS_conf, a component describing the conformational freedom (Fig. 8), further molecular mechanisms are unclear. It is possible that ΔH and ΔS are easily influenced by the effects arising in the bulk solvent, as longer Sop₃ would expose their reducing ends more to the outside SO-BP (Fig. 7A). These trends in ΔH, ΔS, and ΔC_p were similar to the family 17 carbohydrate-binding module (35). In that study, binding of each glucose unit extended from cellotriose yielded an unfavorable ΔH (+1.8 kcal mol⁻¹ per glucose unit) and a favorable −ΔS (−2.8 kcal mol⁻¹ per glucose unit). In addition, the ΔC_p value for cellohexaose was 44 cal mol⁻¹ K⁻¹ higher than that for cellopentaose.

Comparison with other SBPs

The structural homology search of both the open and closed forms of SO-BP revealed that the best DALI hit is Xac-MalE from X. citri. SO-BP resembles Xac-MalE in not only primary and ternary structures but also residues on one side of the ligand-binding site (Met-42, Glu-45, Thr-96, Trp-97, Tyr-145, Arg-149, Asn-195, Glu-196, Ser-265, Trp-268, Gly-300, Gly-301, and Glu-377) (Fig. 7B). To date, Xac-MalE has been classified as a putative maltose/trehalose-binding protein (37). However, the Xac-MalE gene (XAC2310) is located adjacent to a GH144 endo-β-1,2-glucanase homolog (XAC2311) (SBP and GH144 in Fig. 6B, discussed below), strongly suggesting that Xac-MalE is capable of binding to Sop₃.

Biological implications

Phylogenetic analysis revealed that the ABC uptake system associated with SO-BP is conserved in a number of bacterial species in Firmicutes, Actinobacteria, and Proteobacteria (Fig. 6A). The co-occurrence of GHs and related proteins with the SO-BP homolog and ABC transporter suggest the course of β-1,2-glucan dissimilation. A representative degradation system (L. innocua) is schematically shown in Fig. 9. Because SO-BP can discriminate between Sop₃ and β-1,2-glucan based on its affinities, this system is highly likely to be dedicated to dissimilation of Sop₃. Other Firmicutes species appear to share this system, as exemplified in Paenibacillus peoriæ (Fig. 6B). The co-occurrence of intracellular GH1 enzymes with a SO-BP homolog and an ABC transporter (e.g. Streptomyces and Bifidobacterium in Fig. 6B) is widely found in Actinobacteria species. Considering that the GH1 enzymes mostly act on a β-glycosidic bond in exo-mode, these GH1 enzymes adjacent to the SO-BP homolog likely cleave a β-1,2-glucosidic bond from the nonreducing end. The absence of an endo-type glycosidase in these loci suggests that each of the loci also targets Sop₃. In the Proteobacteria group (e.g. X. citri and Stenotrophomonas maltophilia), an extracellular GH144 enzyme, which is likely an endo-β-1,2-glucanase, is present adjacent to a SO-BP ABC transporter homolog (Fig. 6B). In addition, a hypothetical membrane protein (Hypo) co-occurs with the GH144 and SO-BP ABC transporter homolog. This hypothetical membrane protein is predicted to be an outer membrane receptor (TonB_dep_Rec or PF00593 in Pfam), and thus it may assist with translocation of Sop₃ across the outer membrane.

In the case of the maltose/maltodextrin utilization gene locus in E. coli, an NBD (MalK) is present with a SBP (MalE) and TMDs (MalF and MalG). MalK is an ATPase responsible for energy coupling to the transport system. However, no NBD gene has been found in the β-1,2-glucan utilization loci examined so far. In the L. innocua genome, an NBD gene (Lin0304) is found at a distant locus, and it probably energizes the SO-BP ABC uptake system (Fig. 9).

From these findings, an ABC transporter associated with SO-BP, GH enzyme(s), and a hypothetical membrane protein appear to orchestrate dissimilation of β-1,2-glucan or Sop₃. However, we do not know exactly where the above bacteria encounter β-1,2-glucan or Sop₃. Several bacteria belonging to the Firmicutes and Actinobacteria groups have genes for an SO-BP homolog and TMDs but not a GH144 gene (Fig. 6B). Therefore, these bacteria may rely on degraded products supplied from other bacteria that have extracellular GH144 enzymes.

Conclusions

This study provides the first structural and biochemical insights into a Sop₃ transport protein and will facilitate improved understanding of β-1,2-glucan metabolism and the discovery of unidentified β-1,2-glucan metabolic proteins. SBPs have the potential to be utilized for biosensor exploitation
Therefore, SO-BP may be applicable as a biosensor for Sopₙs (n ≥ 3) that are rare in nature.

**Experimental procedures**

**Ligands**

Sop₂₋₅s and linear β-1,2-glucan (average DP 25) were prepared as described previously (17, 29, 39). Sopₙs were separated by Toyopearl HW-40S resin in XK 50/100 columns (50 mm × 1000 nm, 2 columns were used in tandem; Tosoh, Tokyo, Japan) using an ÄKTA system (GE Healthcare, Buckinghamshire, UK), and then by a custom-made Asahipak NH₂P-90 20F column (20 mm × 300 mm; Showa Denko, Tokyo, Japan) using a Prominence HPLC system (Shimadzu, Kyoto, Japan). Laminartriose was purchased from Megazyme (Wicklow, Ireland). Celletriose was purchased from Seikagaku Corp. (Tokyo, Japan). Gentioooligosaccharides (mixture of gentiobiose, gentiotriose, gentiotetraose, and others) and maltotriose were purchased from Wako Pure Chemicals (Osaka, Japan).

**Cloning, overexpression, and purification**

The gene encoding the Lin1841 protein (GenBank ID: CAC97072.1) lacking its signal peptide (amino acid residues 27–414) was amplified by PCR from the genomic DNA of *L. innocua* Clip11262 using KOD plus (Toyobo, Osaka, Japan) and the following forward and reverse oligonucleotide primers: 5'-TGTGGTGGGcatatgGATGATGCAAATTCC-3' and 5'-CCTTTTATctcgagTTTTTTAAGAAGTGC-3', respectively. The lowercase letters in the forward and reverse primers indicate NdeI and XhoI sites, respectively. The genomic DNA was extracted from the cell pellet of *L. innocua* with InstaGene Matrix (Bio-Rad). The amplified gene was purified, digested by Ndel and Xhol, and inserted into pET30a(+) (Novagen, Madison, WI) to encode a His₆-tag fusion protein at the C terminus (pET30a(+) -lin1841). *E. coli* BL21 (DE3) cells (Novagen) were transformed with the constructed plasmid.

The transformant was cultured in Luria-Bertani medium containing 30 mg/liter kanamycin at 37 °C until the absorbance at 600 nm reached 0.6. The protein production was induced with 0.1 mM isopropyl-β-D-1-thiogalactopyranoside at 30 °C for 6 h. The transformant cells were collected by centrifugation at 3,900 g for 10 min and then suspended in 5 ml of 20 mM MOPS-NaOH buffer (pH 7.0) containing 500 mM NaCl (buffer A) per 1 g of the cells. The suspended cells were disrupted by sonication, and the cell debris was removed by centrifugation at 27,000 × g for 30 min to obtain cell extracts. The cell extract was applied to a HisTrap FF crude column (5 ml; GE Healthcare) pre-equilibrated with buffer A. The unabsorbed proteins were removed by washing with buffer A containing 10 mM imidazole, and then the absorbed proteins were eluted with buffer A containing 400 mM imidazole. The eluate was concentrated and buffer-exchanged to 10 mM MOPS-NaOH (pH 7.0) with an Amicon Ultra 10,000 molecular weight cutoff (Millipore). The target protein was applied to a Mono Q 10/100 GL column (GE Healthcare) and eluted with a linear gradient of 0–500 mM...
NaCl in 10 mM MOPS–NaOH (pH 7.0) using an ÄKTA purifier (GE Healthcare). The purity was estimated by SDS–PAGE. The molecular mass estimated by SDS–PAGE (42 kDa) corresponded to the theoretical molecular mass (44,531 Da). For ITC experiments, the fractions showing a single band were collected, concentrated, and buffer-exchanged to 20 mM sodium phosphate buffer (pH 7.0). For crystallographic experiments, the fractions showing a single band on SDS–PAGE and a single peak in the chromatogram were collected, concentrated, and buffer-exchanged to 10 mM MOPS–NaOH (pH 7.0). Protein concentration was determined by measurement of the absorbance at 280 nm and calculation from the theoretical extinction coefficients of Lin1841 lacking its signal peptide (72,880 M−1 cm−1).

**ITC**

ITC experiments were performed at 15–35 °C using MicroCal VP-ITC (Malvern Instruments Ltd., Malvern, Worcestershire, UK). The protein dissolved in 20 mM sodium phosphate buffer (pH 7.0) was used for experiments. Ligand solutions were prepared by dilution with the same buffer obtained from the filtrate of the buffer exchange. The protein solution (0.04–0.3 μl of a ligand (0.5–3.0 mM, except for gentio-oligosaccharides and β-1,2-glucan; 0.67 and 4.05 mg/ml were used, respectively) was stirred at 307 rpm in a 1.44-ml cell and titrated with 5 μl of a ligand (0.5–3.0 mM, except for gentio-oligosaccharides and β-1,2-glucan; 0.67 and 4.05 mg/ml were used, respectively) at 25 °C; intervals were set to 600 s (except for the Sop4 titration at 25 °C; intervals were set to 600 s). The heat of dilution of the oligosaccharides was determined to be negligible based on control experiments in which the ligand was titrated into buffer solution. Calorimetric data were analyzed using Origin 7.0 software. Thermodynamic parameters, such as the association constant (K0), the binding enthalpy (ΔH), and the number of binding sites (n), were determined by fitting data into a one-site binding model. It was difficult to determine the accurate ΔH value toward β-1,2-glucan because of uninterpretable heat pulses detected in the β-1,2-glucan–titration experiments. The binding Gibbs free energy change (ΔG0), the dissociation constant (K0), and the binding entropy change (ΔS0) were calculated from the equations ΔG0 = −RT lnK0 = RT lnK0 and ΔG0 = ΔH − TΔS0, where R is the gas constant, and T is the absolute temperature. It is assumed that ΔH values determined from ITC are equal to the standard enthalpy change (ΔHθ). The heat capacity change (ΔCp = ΔH/ΔT) was calculated from linear regression analysis of ΔH values at different temperatures. The total entropy change is expressed as the sum of entropy changes in solvent released upon ligand binding (ΔS0,solv), conformational freedom around torsion angles of proteins and ligands (ΔS0,conf), and the mixing of solute and solvent molecules (ΔS0,mix) (ΔS0 = ΔS0,solv + ΔS0,conf + ΔS0,mix) (40). ΔS0,solv and ΔS0,mix are calculated from the equation ΔS0,solv = −ΔCp ln(298.15/385.15) and ΔS0,mix = R ln(1/55.5), respectively (40).

**Crystallization and structure determination**

All crystals were obtained by the sitting-drop or hanging-drop vapor diffusion method at 25 °C. Initial crystallization screening was established using JCSG core suite I–IV and JCSG+ suite (Qiagen, Hilden, Germany) based on the sitting-drop vapor diffusion method. To obtain crystals in ligand-free form of Lin1841, the crystallization drops were prepared by mixing 0.5 μl of 19.3 mg/ml Lin1841 solution with an equal volume of the screening kit solution and equilibrated against 70 μl of the same solution. Ligand-free form crystals were generated in a drop of solution containing 0.2 M sodium acetate and 20% (w/v) PEG 3350. After optimizing the conditions, suitable crystals were obtained by mixing 1 μl of the protein solution with an equal volume of the reservoir solution containing 0.15 M sodium acetate and 15% (w/v) PEG 3350 and equilibrated against 500 μl of the same solution using the hanging-drop vapor diffusion method. These crystals completely grew in 2–3 days.

To obtain crystals of Lin1841 in complex with Sop3–5s, the crystallization drops were first prepared by mixing 0.5 μl of 27.2 mg/ml Lin1841 solution containing 10 mM Sop4 in 9 mM MOPS–NaOH (pH 7.0) with an equal volume of the screening kit solution and equilibrated against 70 μl of the same solution. Co-crystals with Sop4 were generated in a drop of solution containing 0.1 M MES–NaOH (pH 6.0) and 40% (v/v) MPD. After optimizing the conditions, suitable crystals were obtained by mixing 1 μl of the protein solution with an equal volume of the reservoir solution containing 0.1 M MES–NaOH (pH 5.3) and 42% (v/v) MPD using the sitting-drop vapor diffusion method for 6 days. Co-crystals with Sop3 were generated in a similar manner as described above, except that Sop4 was used instead of Sop3, and the reservoir solution consisted of 0.15 M MES–NaOH (pH 5.5) and 50% (v/v) MPD. Co-crystals with Sop5 were obtained using the streak seeding method as follows. A drop was equilibrated for a day in a similar manner as described above, except that 28.1 mg/ml protein containing 5 mM Sop5 was used, and the reservoir solution consisted of 0.15 M MES–NaOH (pH 5.3) and 42% (v/v) MPD. The co-crystals were generated in the same drop that was streaked with microseeds of the Sop5 co-crystal for a day.

Crystals in ligand-free form were grown in a drop of reservoir solution supplemented with 30% (v/v) PEG 400 for cryoprotection. There was no need to supply co-crystals with cryoprotectants because of the high concentration of MPD. The crystals were flash-cooled at 100 K in a stream of nitrogen gas. X-ray diffraction data were collected using a charge-coupled device (ADSC Quantum 270) on an NW12A station at the Photon Factory Advanced Ring, High Energy Accelerator Research Organization (KEK), Tsukuba, Japan (λ = 1.000 Å), and processed using HKL2000 (41).

The initial phase of the ligand-free form Lin1841 was determined by the molecular replacement method using MOLREP (42), and the structure of Xac-MaLE from *X. citri* (PDB code 3UOR) was used as a search model. The phase improvement was performed using Morph model on Phenix (43). The initial phase of the Lin1841 in complex with Sop4 was also determined using MOLREP, and the structure of the ligand-free form was used as a search model. The phase improvement and the automated model building were performed using ARW/ARP (44). Sop4 and Sop5 complex structures were solved by the molecular replacement method using MOLREP, and the structure of Sop4 complex was used as a search model. Initial model structures of α-Sop4, α-Sop3, and β-Sop4 were built with JLigand (45). Manual model building was carried out using Coot (46). Crystallo-


**Binding protein specific for β-1,2-glucooligosaccharides**

graphic refinement was performed using REFMAC5 (47) with the TLS parameters generated by the TLSMD server (48). The refined structures were validated using Molprobity (49) and Rampage (50). The molecular graphic figures were prepared using PyMOL (DeLano Scientific, Palo Alto, CA).

**MD simulation**

The crystal structures of SO-BP (chain A) in ligand-free form and in complex with Sop3–5s were used to construct the initial structures for MD simulations. The N and C termini of proteins were capped with acetyl and N-methyl groups, respectively. The protonation states of histidine were assigned using PROPKA 3.1 (51), and pH 7.0 was used for the calculation. A His-75 of the Sopn complex was protonated on the N61 and Nε2 atoms, and the other histidines were protonated only on the Nε2 atom. The ligand-free structure and complex structures with Sop3–5s were first immered in cubic water boxes where the distance between protein atoms and the closest boundary was at least 10 Å. Sodium ions were added to the systems for neutralization. The LEaP module of AmberTools 16 (52) was used to produce the initial structures. Amber ff14SB (53) and GLYCAM06j (54) force-field parameters and the TIP3P model (55) were used for the protein, carbohydrate, and water, respectively. The systems were gradually heated to 300 K during 200-ps constant NVT-MD simulations with position restraints on the nonhydrogen atoms of the proteins, and the force constants were set to 10 kcal mol⁻¹ Å⁻². During subsequent 800-ps constant NPT-MD simulations, the pressure was adjusted to 1.0 × 10⁵ pascal, and the force constants of the position restraints were gradually decreased to 0 kcal mol⁻¹ Å⁻². Finally, unrestrained MD simulations were carried out for 100 ns.

The ligand-free SO-BP in the closed form and the free Sop3–5s were also equilibrated in aqueous solution. The ligand-free closed SO-BP was prepared by removing the Sop3 from the SO-BP–Sop3 complex structure. The structures of the free Sop3–5s were prepared using the LEaP module. The simulation procedures were the same as described above, except that the unrestrained MD simulation was performed at 10 ns for the system of the free Sop3–5s. The final coordinates were employed as the initial structures of the MD simulations for free energy calculations. All the MD simulations were performed using GROMACS version 5.0.5 (56). In the simulations, the temperature and pressure were controlled by the velocity rescaling method (57) and the weak coupling method (58), respectively. The bond lengths involving the hydrogen atoms were constrained using LINCS algorithm (59), allowing the use of 2-fs time steps. The electrostatic interaction was calculated using the particle mesh Ewald method (60).

PCA was carried out for the MD trajectory of the open form of SO-BP using the method described previously (61). Deviation of the Cα atoms of the MD snapshots of the trajectory from those of the average structure was analyzed.

**Free energy calculation**

The contributions of the unit C glucose and Gln-197 of SO-BP to the binding free energy were calculated using alchemical thermodynamic cycles illustrated in Fig. 4, A and B. ∆G₁ and ∆G₂ were defined as the difference in the free energy between the free Sop3 and the free Sop₂ and that between Sop₃ and Sop₂ in the complex, respectively. The contribution of the unit C glucose moiety was calculated as ∆ΔG = ∆G₁ - ∆G₂ (Fig. 4A). ∆G₁ and ∆G₂ were defined as the difference in the free energy between the ligand-free closed forms of SO-BP and its Q197A mutant and between their ligand-bound forms, respectively. The contribution of Gln-197 was calculated as ∆ΔG = ∆G₃ - ∆G₄ (Fig. 4B). These free energy calculations were based on the Bennett acceptance ratio method (62). In the calculations, a part of the structure was alchemically transformed from one to the other (Sop₃ to Sop₂ or glutamine to alanine) in a stepwise manner, considering 39 intermediate states in each transformation process. In each step of the transformation process, 1-ns MD simulations were performed sequentially, and the last 500 ps of each simulation were used for the free-energy calculation with the Bennett acceptance ratio method. The calculations were performed with the g_bar module of GROMACS.

**Site-directed mutagenesis**

Mutants of SO-BP (D193A and Q197A) were constructed using a PrimeSTAR mutagenesis basal kit (Takara) and the pET30a(+)–lin1841 was used as a template. The following primer pairs were used for amplification of mutant proteins (mutation sites are underlined): 5’-GCAATTCTCCACAAACGC-AACCAACTACT-3’ and 5’-GGTTGAGCTTTGGAAACGGTGATCAT-3’ (D193A); 5’-AACGAGCAACTCTGTTTACACATTTCC-3’ and 5’-AGTAGTGGCTTTGGATCAATTC-3’ (Q197A). The mutant proteins were produced and purified using the same way as described above.

**Other analyses**

Sequences for phylogenetic analyses were retrieved from the RefSeq database with sequence identities of >30% (E-value of <10⁻⁶⁰) after the BLAST search. The retrieved 57 sequences were aligned using MUSCLE (63), and the phylogenetic tree was constructed using MEGA7 (64) based on the neighboring method. Conservation scores of each residues of SO-BP were calculated and colored with Consurf (65, 66) using the above phylogenetic tree and the SO-BP structure in complex with Sop₃.

**Author contributions**—K. A., T. T., M. N., and S. F. conceptualization; K. A. data curation; K. A., N. S., T. T., T. A., and S. F. formal analysis; K. A., T. T., A., K. I., M. S., H. N., H. T., and M. N. supervision; K. A., N. S., and S. F. validation; K. A., N. S., T. T., T. A., M. N., and S. F. investigation; K. A. visualization; K. A. and S. F. writing—original draft; K. A., N. S., T. T., K. I., H. N., H. T., M. N., and S. F. writing—review and editing; N. S., T. T., K. I., and M. N. methodology; Y. T., K. I., H. N., and M. N. resources; M. S., H. T., M. N., and S. F. project administration; H. T., M. N., and S. F. funding acquisition.

**Acknowledgments**—We thank the staff of the Photon Factory and SPing-8 for the X-ray data collection. We thank Dr. A. Miyana for helpful discussion.
References

1. Davidson, A. L., Shuman, H. A., and Nikaido, H. (1992) Mechanism of maltose transport in *Escherichia coli*: transmembrane signaling by periplasmic binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2360–2364 CrossRef Medline

2. van der Heide, T., and Poolman, B. (2002) ABC transporters: one, two or four? *Extracytoplasmic substrate-binding sites? EMBO Rep.* **3**, 938–943 CrossRef Medline

3. Shuman, H. A. (1982) Active transport of maltose in *Escherichia coli* K12. Role of the periplasmic maltose-binding protein and evidence for a substrate recognition site in the cytoplasmic membrane. *J. Biol. Chem.* **257**, 5455–5461 Medline

4. Ames, G. F. (1986) Bacterial periplasm transport systems: structure, mechanism, and evolution. *Annu. Rev. Biochem.* **55**, 397–425 CrossRef Medline

5. Berntsson, R. P., Smits, S. H., Schmitt, L., Slotboom, D. J., and Poolman, B. (2010) A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606–2617 CrossRef Medline

6. Tang, C., Schwieters, C. D., and Clore, G. M. (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **449**, 1078–1082 CrossRef Medline

7. Mao, B., Pear, M. R., McCammon, J. A., and Quiocho, F. A. (1982) Hinge bending in *i*-arabinobiose-binding protein. The “Venus’s-flytrap” model. *J. Biol. Chem.* **257**, 1131–1133 Medline

8. Bohn, J.-P. (2000) Osmoregulated periplasmic glucans in proteobacteria. *FEBS Microbiol. Lett.* **186**, 11–19 CrossRef Medline

9. Arellano-Reynoso, B., Lapaque, N., Salcedo, S., Briones, G., Ciocchini, A. E., Ugalde, R., Moreno, E., Moriyón, I., and Gorvel, J.-P. (2005) Cyclic β-1,2-glucan is a *Brucella* virulence factor required for intracellular survival. *Nat. Immunol.* **6**, 618–625 CrossRef Medline

10. Breedveld, M. W., and Miller, K. J. (1994) Cyclic β-glucans of members of the family Rhizobiaceae. *Microbiol. Rev.* **58**, 145–161 Medline

11. Gay-Fraret, J., Ardissone, S., Kambara, K., Broughton, W. J., Deakin, W. J., and Le Quéré, A. (2012) Cyclic β-glucans of *Rhizobium* (*Sinorhizobium*) sp. strain NGR234 are required for hypo-osmotic adaptation, motility, and efficient symbiosis with host plants. *FEBS Microbiol. Lett.* **333**, 28–36 CrossRef Medline

12. Miller, K. J., Kennedy, E. P., and Reinhold, V. N. (1986) Osmotic adaptation by Gram-negative bacteria: possible role for periplasmic oligosaccharides. *Science* **231**, 48–51 CrossRef Medline

13. Talaga, P., Fournet, B., and Bohn, J. P. (1994) Periplasmic glucans of *Pseudomonas syringae* pv. *syringae*. *J. Bacteriol.* **176**, 6538–6544 CrossRef Medline

14. Van Bogaert, I. N., Saerens, K., De Muynck, C., Develter, D., Soetaert, W., and Vandamme, E. J. (2007) Microbial production and application of sophorolipids. *Appl. Microbiol. Biotechnol.* **76**, 23–34 CrossRef Medline

15. Schmidt, S., Zietz, M., Schreiner, M., Rohn, S., Kroh, L. W., and Krumbein, V. (2014) Mechanistic insight into the substrate specificity of 1,2-β-glucan phosphorylase from *Lachnoclostridium phytofermentans*. *Appl. Environ. Microbiol.* **80**, 2419–2424 CrossRef Medline

16. Mandels, M., Parrish, F. W., and Reese, E. T. (1962) Sophorose as an extracytoplasmic substrate-binding site in the family Rhizobiaceae. *J. Biol. Chem.* **233**, 292–300 Medline

17. Nakajima, M., Tanaka, N., Furukawa, N., Nihira, T., Kodutsumi, Y., Takahashi, Y., Sugimoto, N., Miyanaga, A., Nakai, H., Arakawa, T., Fushinobu, S., and Taguchi, H. (2017) Biochemical and structural analyses of a bacterial endo-β-1,2-glucanase reveal a new glycoside hydrolase family. *J. Biol. Chem.* **292**, 7487–7506 CrossRef Medline

18. Oldham, M. L., and Chen, J. (2011) Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science* **332**, 1202–1205 CrossRef Medline

19. Nakajima, M., Tanaka, N., Furukawa, N., Nihira, T., Kodutsumi, Y., Takahashi, Y., Sugimoto, N., Miyanaga, A., Fushinobu, S., and Taguchi, H. (2017) Mechanistic insight into the substrate specificity of 1,2-β-glucan phosphorylase from *Lachnoclostridium phytofermentans*. *Sci. Rep.* **7**, 42671 CrossRef Medline

20. Abe, K., Nakajima, M., Yamashita, T., Kaminou, K., Nihira, T., Takahashi, Y., Sugimoto, N., Miyanaga, A., Nakai, H., Arakawa, T., Fushinobu, S., and Taguchi, H. (2017) Biochemical and structural analyses of a bacterial endo-β-1,2-glucanase reveal a new glycoside hydrolase family. *J. Biol. Chem.* **292**, 7487–7506 CrossRef Medline

21. Geh, M., Federlund, F., Andersen, J. M., Vujčić Žagar, A. V., Henriksen, J. R., Andersen, T. L., Svensson, B., Slotboom, D. J., and Abou Hachem, M. (2016) An ATP binding cassette transporter mediates the uptake of α-(1,6)-linked dietary oligosaccharides in *Bifidobacterium* and correlates with competitive growth on these substrates. *J. Biol. Chem.* **291**, 20220–20231 CrossRef Medline

22. Boraston, A. B., Creagh, A. L., Alam, M. M., Kormos, J. M., Tomme, P., Haynes, C. A., Warren, R. A., and Kilburn, D. G. (2001) Binding specificity and thermodynamics of a family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A. *Biochemistry* **40**, 6240–6247 CrossRef Medline

23. Notenboom, V., Boraston, A. B., Chiu, P., Freelove, A. C., Kilburn, D. G., and Rose, D. R. (2001) Recognition of cello-oligosaccharides by a family 17 carbohydrate-binding module: an X-ray crystallographic, thermodynamic and mutagenic study. *J. Mol. Biol.* **314**, 797–806 CrossRef Medline

24. Prabhoo, N. V., and Sharp, K. A. (2005) Heat capacity in proteins. *Annu. Rev. Phys. Chem.* **56**, 521–548 CrossRef Medline
Binding protein specific for \(\beta\)-1,2-glucooligosaccharides

37. Medrano, F. I., de Souza, C. S., Romero, A., and Balan, A. (2014) Structure determination of a sugar-binding protein from the phytopathogenic bacterium Xanthomonas citri. Acta Crystallogr. F Struct. Biol. Commun. 70, 564–571 CrossRef Medline

38. Medrano, F. J., de Souza, C. S., and Balan, A. (2014) Crystal structure of a sugar-binding protein from the phytopathogenic bacterium Xanthomonas citri. Acta Crystallogr. F Struct. Biol. Commun. 70, 564–571 CrossRef Medline

39. Abe, K., Nakajima, M., Kitaoka, M., Toyoizumi, H., Taguchi, Y., Sugimoto, N., Nakai, H., and Taguchi, H. (2015) Large-scale preparation of 1,2-\(\beta\)-glucan using 1,2-\(\beta\)-oligoglucan phosphorylase. J. Appl. Glycosci. 62, 47–52

40. Baker, B. M., and Murphy, K. P. (1997) Dissecting the energetics of a protein–protein interaction: the binding of ovomucoid third domain to elastase. J. Mol. Biol. 268, 557–569 CrossRef Medline

41. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 CrossRef Medline

42. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022–1025 CrossRef

43. Terwilliger, T. C., Read, R. I., Adams, P. D., Brunger, A. T., Afonine, P. V., and Hung, L. W. (2013) Model morphing and sequence assignment after molecular replacement. Acta Crystallogr. D Biol. Crystallogr. 69, 2244–2250 CrossRef Medline

44. Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/WARP version 7. Nat. Protoc. 3, 1171–1179 CrossRef Medline

45. Lebedev, A. A., Young, P., Isupov, M. N., Moroz, O. V., Vagin, A. A., and Murshudov, G. N. (2012) Ligand: a graphical tool for the CCP4 template-restraint library. Acta Crystallogr. D Biol. Crystallogr. 68, 431–440 CrossRef Medline

46. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 CrossRef Medline

47. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367 CrossRef Medline

48. Painter, J., and Merritt, E. A. (2006) TLSMD web server for the generation of multi-group TLS models. J. Appl. Crystallogr. 39, 109–111 CrossRef

49. Chen, V. B., Arendall, W. B., 3rd., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 CrossRef Medline

50. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Cα geometry: \(\psi\), \(\phi\) and Cβ deviation. Proteins 50, 437–450 CrossRef Medline

51. Olsson, M. H., Sondergaard, C. R., Rostkowski, M., and Jensen, J. H. (2011) PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. J. Chem. Theory Comput. 7, 525–537 CrossRef Medline

52. Case, D. A., Cerutti, D. S., Cheatham, T. E., 3rd., Darden, T. A., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Greene, D., Homeyer, N., Izadi, S., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., et al. (2017) AMBER 2017. University of California, San Francisco

53. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) f14SB: Improving the accuracy of protein side chain and backbone parameters from f99SB. J. Chem. Theory Comput. 11, 3696–3713 CrossRef Medline

54. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeirito, 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 CrossRef Medline

55. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935 CrossRef

56. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1, 19–25 CrossRef

57. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. J. Chem. Phys. 126, 14101 CrossRef

58. Berendsen, H. J., Postma, J. P., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81, 3684–3690 CrossRef

59. Hess, B. (2008) P-LINCS: A parallel linear constraint solver for molecular simulation. J. Chem. Theory Comput. 4, 116–122 CrossRef Medline

60. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. J. Chem. Phys. 103, 8577–8593 CrossRef

61. Terada, T., and Kidera, A. (2012) Comparative molecular dynamics simulation study of crystal environment effect on protein structure. J. Phys. Chem. B 116, 6810–6818 CrossRef Medline

62. Bennett, C. H. (1976) Efficient estimation of free energy differences from Monte Carlo data. J. Comput. Phys. 22, 245–268 CrossRef

63. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797 CrossRef Medline

64. Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874 CrossRef Medline

65. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) ConSurf 2005: the projection of evolutionary conservation in macromolecules. Mol. Biol. Evol. 22, 525–537 CrossRef

66. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 44, W299–W302 CrossRef Medline

67. Nielsen, H. (2017) Predicting secretory proteins with SignalP. Methods Mol. Biol. 1611, 59–73 CrossRef Medline

68. Rahman, O., Cummings, S. P., Harrington, D. J., and Sutcliffe, I. C. (2008) Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. World J. Microbiol. Biotechnol. 24, 2377–2382 CrossRef