Arabian Journal of Chemistry (2022) **15**, 104198

King Saud University

Arabian Journal of Chemistry

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Identifying ligand-binding specificity of the oligopeptide receptor OppA from *Bifidobacterium longum* KACC91563 by structure-based molecular modeling

Han-Ha Chai [a,b,*], **Jun-Sang Ham** [c], **Tae-Hun Kim** [a], **Dajeong Lim** [a,*]

[a] *Animal Genome & Bioinformatics Division, National Institute of Animal Science, RDA, Wanju 55365, Republic of Korea*
[b] *College of Pharmacy, Chonnam National University, Kwangju 61186, Republic of Korea*
[c] *Animal Products Research and Development Division, National Institute of Animal Science, RDA, Wanju 55365, Republic of Korea*

Received 23 January 2022; accepted 11 August 2022
Available online 17 August 2022

KEYWORDS

Bifidobacterium longum KACC91563;
Oligopeptide-binding protein A;
Importer;
Structure-based molecular modeling, Probiotics

Abstract As a ATP-binding cassette (ABC) transporter the OppA receptor plays key roles in protecting the host organism and transporting nutrients across the intestine by the oligopeptide transporter from symbiotic bacteria and directs maturation of the host immune system. Among lactic acid bacteria, *Bifidobacterium longum* KACC91563, isolated from fecal samples of healthy Korean neonates, has the capability to alleviate food allergy effects. Operating as a peptide importer, the extracellular OppA receptor from gram-positive *B. longum* KACC91563 translocates nutrients, specifically peptides, from the outside environment of the intestinal tract to the inside of symbiotic cells. In the present study we attempt to explicate the relationship between the substrate's specificity from the OppA importer and the probiotic effects of *B. longum* KACC91563 in the host intestine. It was first identified in this study the specialized structure–function relationship from the OppA importer of *B. longum* KACC91563 with its structural and functional determinants. This could provide insights into substrate specificity of unique immunological properties and a key switch for the substrate's metabolism to reprogramming immune responses in the host intestine by structure-based

Abbreviations: ABC, ATP-binding cassette; LAB, Lactic acid bacteria; OppA, Oligopeptide-binding protein A; B. longum, Bifidobacterium longum; L. lactis, Lactococcus lactis; B. substilis, Basillus subtills; SBP, Solute-binding protein; E. faecalis, Enterococcus faecalis; SAP, Spatial aggregation propensity; mTORC, mechanistic target of rapamycin complex

* Corresponding authors at: Animal Genome & Bioinformatics Division, National Institute of Animal Science, RDA, Wanju, 55365, and College of Pharmacy, Chonnam National University, Kwangju 61186, Republic of Korea (H.-H. Chai). Animal Genome & Bioinformatics Division, National Institute of Animal Science, RDA, Wanju 55365, Republic of Korea (D. Lim).
E-mail addresses: hanha@korea.kr (H.-H. Chai), hamjs@korea.kr (J.-S. Ham), thkim63@korea.kr (T.-H. Kim), lim.dj@korea.kr (D. Lim).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.arabjc.2022.104198

molecular modeling. The probiotic effects of oligopeptide substrate (such as a proline-rich peptide containing at least one branched residue of leucine, isoleucine, and valine) and its metabolism for the OppA from *B. longum* KACC91563 are attributed to enhancement of the epithelial barrier by several different strain specific pathways to prevent the strong adhesion of pathogens.

1. Introduction

Bifidobacteria are gram-positive bacteria that possess one lipid bilayer membrane enveloped by a cell wall. Currently, the *Bifidobacterium* genus consists of 56 species and 9 subspecies (Ventura et al., 2018) isolated from human and other mammalian intestinal contents. A total of 311 *Bifidobacterial* genome sequences could be retrieved from the National Center for Biotechnology Information (NCBI). *Bifidobacterium longum* KACC91563, a subspecies of the *Bifidobacterium* genus, belongs to the lactic acid bacteria (LAB) family (Jeong et al., 2015; van Sinderen, 2016) and one of the probiotics, which is often isolated from the feces of neonates (BioSample code SAMN02603656). The 2.40 Mb-sized genome of the strain consist of 1,856 protein-coding genes including 15 cell wall proteins (0.77 %) and 20 extracellular proteins (1.02 %) (Kim et al., 2011) with 87.77 % of the *Bifidobacterium longum* (*B. longum*) symmetrical identity constituting the *Bifidobacterial* secretome database from the BioProject PRJNA66401. The best characterized member of the *Bifidobacteria* is *Bacillus subtilis*, whose genome of 4.21 Mb encodes the putative 77 ATP-binding transport proteins in its 4,100 protein-coding genes (Bruschi et al., 2003). The most common extracellular protein of the LAB family has been identified in the oligopeptide-binding protein A (OppA) receptor as belonging to the solute binding proteins of the ATP-binding cassette (ABC) importers (as represented in Fig. S11). *Bifidobacterium*, similarly to the other LAB members, directly degrades exogenous proteins (such as milk proteins) to peptides using its protease, and the OppA receptor takes up the peptide-ligand (Jeong et al., 2015; Ham et al.,2013) as a nutrient transporter.

The OppA from prokaryotes acts as a peptide transporter into the cell, thereby serving as a nutrient import receptor. The prokaryotic OppAs share 3D scaffolds, while the % identity of amino acid sequence are <20 % (Saier, 1993; Poolman et al., 2010; Quiocho, 1995; Wilkinson et al., 1994; Lee, 2016). The scaffold is composed of two α/β domains (i.e., domain I and II) and a hinge region. The hinge region interconnecting both α/β domains displays diverse structural skeletons and arrangements between the prokaryotic OppAs. The peptide-binding pocket of the OppA receptor forms a cavity between the rigid α/β domains. With the mobility of its hinge-bending joined into both, the relative orientations and surrounding areas of the α/β domains are reflected inside the voluminous cavity of OppAs. If a single hinge-link becomes more bent, then the hinge bending motion becomes larger than the two hinge-strand fragments to intercross spaces from the two α/β domains. As a result, the hinge-bending motion can accelerate internal interactions between the α/β domains governed by OppA on the nonbinding of the peptide ligand. This also promotes intermolecular interactions between the OppA receptor and its peptide ligand, which may help mutual packing at their contactable interfaces. This was observed in crystallized prokaryotic OppA/peptide complex structures where the α/β domains were tightly located together with the peptide-ligand having suitable lengths buried in the binding pocket (e.g., 3FTO was compared in an open-unliganded conformation (Slotboom et al., 2009a) to 3DRG in a closed-liganded conformer of OppA's crystal structures (Slotboom et al., 2009b) from Lactococcus lactis). The structural topology of the hinge-bending region allows the binding pocket to take open and closed OppA conformers through substantial movements induced during peptide-binding. The hinge-bending motions should therefore be considered as one of the major structural features of the prokaryotic OppAs that permits the OppA receptor to attach to its peptide ligand. Specifically, there is a tendency to stabilize the OppA/peptide complex structure transferred into a closed conformer via the binding kinetics of the "Venus flytrap" model (Quiocho, 1982; Quiocho, 1995; Sadee 1999). The OppA receptors share 3D-structure folds and also have common characteristics with the ligand-binding mechanism of the Venus flytrap model. It is interesting that OppA receptors have four structural states (i.e., liganded, unliganded, open, and closed states) at particular coupling equilibriums on the basis of the Venus flytrap model. However, not all of the structural states exist as the determined structures of the prokaryotic OppAs. The OppA receptor from *Lactococcus lactis* (*L. lactis*), which is the best characterized member of the four structural states, is available in open-unliganded, open-liganded, and closed-liganded forms due to its crystallized structures, shown in Fig. S15, including the PDB ID of 3FTO, 3DRH, and 3DRG, respectively (Slotboom et al., 2009a,b).

In the open conformer of the OppA receptor, the two α/β domains move away from each other in the hydrated state where the peptide-binding pocket is well exposed to water. When any hydrophobic peptide binds, the ligand will push out water molecules from the hydrated cavity to maintain hydrophobic interactions with the counterpart residues that protrude into the OppA binding cleft. At the same time, binding the polar residues of the peptide leads to reordering of the water molecules via H-bonding networks between the polar residues and water. The use of water molecules by the OppA receptor is further linked to the filling of spaces not occupied by the peptide within its voluminous binding pocket. This is responsible for maintaining its H-bonding potentials before and after the ligand-binding (Kornings, 1999). In particular, a broad specificity against its peptide-ligand was highlighted in the *L. lactis* OppA receptor, which has a huge binding cavity (almost 4900 $Å^3$) (Slotboom et al., 2009b; Oostenbrink, 2016). Some residues of the bound peptide should indirectly interact with their counterpart residues on the outer contact surfaces of the hydrated pocket, in addition to direct interactions inside the pocket. If the *L. lactis* OppA interacts with the 9-mer peptides as an optimal binding preference, the side-chains of the 9-mer peptides positioned inside the binding pocket (i.e., positions from 1 to 6 side chains on the ligand fit into the hydrated pocket) appear to be more selective than the other side chains (i.e., positions 7, 8, and 9 on the ligand) located on the outer contacts (Poolman, 2000). On the basis of the side chain's size and characteristics, the peptide-ligand would be accepted in the hydrated pocket and be deeply fitted into the interaction interfaces rather than having instantaneous contact with the OppA receptor. Whether the open or closed complex conformer appears is dependent on its peptide-binding affinity to the OppA receptor (Slotboom et al., 2009a,b; Tame 1999) in conjunction with producing a structural hinge region and an extra domain III. More energy compensation through binding both is then required rather than losing the entropy effects from the OppA receptor and peptide-ligand, respectively. In particular, stronger hydrophobic interactions within the OppA/peptide complex are acquired from pushing out well-arranged water molecules from the hydrated pocket of the OppA receptor. The hydrophobic interactions thus should be an essential interaction determinant characterizing the driving forces for satisfying the binding affinity from the prokaryotic OppA/peptide complex structures (Plooman, 2000; Monnet, 2003).

At the prediction stage of the OppA structure from *B. longum* KACC91563, we not decided to couple equilibriums between the open and closed complex conformers related to the ligand's binding affinity from the OppA templates of the LAB members. Practically, structural

difference in *L. lactis* OppAs were observed between the open (PDB ID: 3DRH, bound to 6-mer peptide consisting of all Ala residues with low affinity) and closed (PDB ID: 3DRG, bound to 9-mer peptide being composed of bradykinin like RPPGFSPFA peptide sequence with high affinity) conformers (Slotboom et al., 2009b) upon peptide-ligand binding with a RMSD (root-mean-square) of 2.08 Å per 556 residues. It was found that the structural difference between unliganded (PDB ID: 3FTO) and liganded (PDB ID: 3DRH) in their open structural states is quite small with an RMSD of 0.34 Å per 552 residues by structural pair alignment of the jFATCAT (Poolman, 1999) rigid-body mode. This indicates that there is little structural difference between the open-unliganded and open-liganded conformers of the *L. lactis* OppA. This is because the open structures easily retained their open states induced in not only rotation hindrances consisting of their two hinge-strand fragments and an extra domain, but also the bulky volume of the extra domain III. Surprisingly, the backbones of the peptide-ligands, apart from their binding affinities, were in fixed conformers (RMSD of 0.4 Å) while the H-bonding remained with counterpart residues (such as Ser472 and Ser474 residues of *L. lactis* OppA (Slotboom et al., 2009b) represented in Fig. S15) between the open and closed-liganded complexes. Aside from addition of the different peptide lengths, all backbones of the peptide-ligands were bound in the same structural patterns (shown in Fig. S6).

Based on the structural colorations from both hinge-fragments and an extra domain III, we attempted to find conserved structural features of LAB OppAs from their determined X-ray structures before sifting the OppA templates of *B. longum* KACC91563 among them. To this end, we performed knowledge-based modeling for an unknown OppA structure from *B. longum* KACC91563 by projecting its biophysical information into conserved pictures on the OppAs of the LAB family and by providing its molecular characteristic markers distinctive from the LAB member's OppAs from previous experimental studies (Jeong, 2015; Ham et al., 2013; Jang et al., 2016). Especially, the distinguishing specificity of the peptide ligand will be demonstrated by the interaction factors of its binding cavity based on the closed-liganded OppA model docked in a bradykinin like peptide (as 9-mer peptide of RPPGFSPFA). In the current study, we first identified both structural and functional determinants of the OppA from *B. longum* KACC91563 by demonstrating how this specificity of the peptide ligand could be accepted.

We further attempted to explicate the relationship between the substrate's specificity from the OppA importer and the probiotic effects of *B. longum* KACC91563 in the host intestine based on the structure–function perspectives of the OppA importer. The specialized structure–function relationship from the OppA importer could provide an abstract of substrate specificity into unique immunological properties of the host organism. Moreover, functional characterization of solute-binding proteins (such as 15 cell wall proteins and 20 extracellular proteins) on the *B. longum* KACC91563 genome will provide insight into the key switch for the substrate's metabolism into reprogramming immune responses in the host intestine.

2. Materials and methods

All molecular modeling and optimization were performed in Discovery Studio, version 2017 R2, from BIOVIA (San Diego, USA) (Dassault Systémes BIOVIA Discovery studio modeling environment, release R2 dassault systems, 2017) and the steps of the research process used in the Discovery Studio are diagrammed in Fig. S17.

2.1. Selection of template OppAs

In total, bacterial OppAs have 50 crystal structures, of which 39 structures belong to periplasmic OppAs solved in gram-negative bacteria; they include OppAs from *Burkholderia Pseudomallei* (Berntsson et al., 2011; Klepsch et al., 2011), *Escherichia coli* (Tame, 1999; Hellinga, 2009), *Salmonella typhimurium* (Kornings, 1999; Wilkinson, 1997; Dunny et al., 2006), and *Yersinia pestis* (Byrne et al., 2007). Otherwise, the other 11 structures of extracellular peptide binding proteins originated from gram-positive bacteria (such as AppA from *Basillus subtills* (Wilkinson et al., 2005) and OppA from *Lactococcus lactis* (Slotboom et al., 2009a,b; Oostenbrink, 2016; Poolman, 2000). Among the known structures of gram-positive bacterial OppAs, biological relevance as template OppAs was determined by the following shared attributes, which are physiologically similar to OppA from *L. lactis* and AppA from *B. subtills*; i) the nature of peptide-ligand is fundamentally attendant on lipoprotein derivatives such that in gram-positive bacteria, the OppA transporter consists of a lipoprotein subcomponent distended beyond the extracellular faces of the cell membrane (Chakravortty, 2017). ii) The homologous hydrophobicity and cavity size of their binding pockets is closely related to the characteristic of trace residues derived from these conserved regions on the basis of structural similarities (Slotboom et al., 2009b; Jeong et al., 2015; Ham et al., 2013, Oostenbrink, 2016). For instance, templates (such as *L. lactis* OppA) and the query of OppA from *B. longum* KACC91563 have been classified into cluster C of SBPs by Poolman B (Poolman B. et al., 2010) based on crystalized structural alignments in PDB. In this study, the sequence identities of template candidate SBPs were excited in 15 - 28 % over their equivalent 450–520 residues by searching peptide transporters from the cluster C of SBPs (Saier, 1993; Poolman B. et al., 2010) via structural homology searches and sequence profiles of iterative protein PSI-BLAST (Consortium, 2010) within the UniProt knowledgebase of Swiss-Prot. Among the truncated candidates, the DppA of *E. coli* (Tame, 2003; Hellinga, 2009) and two OppAs of *S. typhimurium* and of *Thermotoga maritima* (Bolognesi et al., 2013) have no clear distinctions of structural features from the searched template SBPs (these also have the same 3D-scaffolds containing an extra domain). Despite this, they were differently marked in a biochemical overview from each other due to their discrete nature being accessible to the extracytoplasmic receptor OppAs from gram-negative bacteria. From a critical point, periplasmic OppAs in gram-negative bacteria have a limiting size of peptide-ligands (such as peptides from two to five resides) that can be transported through the outer membrane. In contrast, no such physical constraints exist in extracellular OppAs from gram-positive bacteria. The extracellular OppAs therefore transport longer peptide-ligands depending on the corporeal constraints of the binding pocket's size. Practically, some periplasmic OppAs from gram-negative bacteria (including the *S. typhimurium* OppA (Kornings, 1999) of PDB code of 1B9J) were categorized in cluster C of SBPs based on only their 3D-structural scaffolds (often $\leq$ 25 % as a consequence of their protein sequence alignments to the query OppA). In contrast, the top five sorted template' candidates were less frequently listed than in the OppAs of gram-positive bacteria. Even in the gram-positive bacteria of *B. substilis,* the extracellular AppA was selected into one of the templates, except for its periplasmic OppA. As a result, the OppA templates were carefully selected, instead of exiguous sequence identities. iii) The query OppA from *B. longum* KACC 91,563 is also an

extracellular SBP in cluster C and a LAB member such as *L. lactis* and *B. substilis*.

On the basis of the three considerations mentioned above, the templates from the resolved gram-positive bacterial OppA structures can be sorted to determine the most suitable for predicting the query OppA structure from the *B. longum* KACC91563. The hand-picked templates were well matched in their structural alignments, where four templates were overlaid with characteristic determinants (as highlighted in Fig. S2). There were high scoring matches of extracellular SBPs between the OppA from *Bacillus anthracis str. Ames* and AppA from *B. substilis* (the known structures 5U4O and 1XOC are open-unliganded and closed-liganded forms, respectively). Both similarities were consistently anticipated from their structures or sequences, which exhibit 27 % sequence identity and 2.60 Å of RMSD with 467 equivalent positions. This depends on whether the ligand binds all of the structural differences between them, is over 90 % mapping coverage with an e-value of $3e^{-50}$. In the case of the OppA structure from *B. anthracis str. Ames* (PDB ID: 5U4O), the structural features have not been identified in detail (not published). AppA from *B. substilis* (PDB ID: 1XOC) (Wilkinson et al., 2005; Kim et al., 2010), OppA from *B. anthracis str. Ames* (PDB ID: 5U4O), and OppA from *L. lactis* (PDB ID: 3FTO, 3DRG) (Slotboom et al., 2009a,b) were commonly found in ATP-binding cassette (ABC) transporters for oligopeptide uptake. In contrast, an unexpected template of PreZ from *Enterococcus faecalis* (*E. faecalis*) is a pheromone binding protein. The designated fourth template PreZ is a lipid-anchored extracellular SBP and serves as a pheromone receptor from *E. faecalis,* whose protein sequence is 19 % identical to that of the query OppA from *B. longum* KACC 91563. The *E. faecalis* PreZ has an external domain similar to that of the *L. lactis* OppA. Intriguingly, *L. lactis* OppA was also found to be homologous to other SBPs that convey sex pheromones in the type of ABC transporters (Saier, 1993; Wilkinson et al., 2005; Andersson et al., 2010). The homologous PreZ receptor not only has high selectivity for hydrophobic 7-mer peptide (present in Fig. 6), but also *E. faecalis* is a gram-positive bacterium that lives in the gastrointestinal tract of mammals (Poolman, 2012). The favored characteristics of hydrophobic pheromone ligands have been presented from two competitive 7-mer peptides. Both originated from processing cCF10 (LVTLVFV) of the *ccfA* gene product and the inhibitor peptide iCF10 (AITLIFI) encoded by the *icf10* gene. The 7-mer pheromone ligands compete for the same binding site (as shown in Fig. 6) of the *E. faecalis* PreZ (Poolman, 2012; Earhart et al., 2005; Konings et al., 1998). These templates all must accommodate their peptide-ligand in a manner of Venus-flytrap mode. We preferentially distinguished their open-unliganded structures (PDB ID: 3FTO, 5U4O) from their closed-liganded forms (PDB ID: 3DRG, 1XOC, 4FAJ), as if all ligands were to be wholly buried into the peptide-binding clefts. A comparison of the templates and the query is given in Table S1 on the basis of their protein sequence alignments.

2.2. Optimizing the OppA model structures from B. longum KACC91563 on the open-unliganded structural state

The templates and the query of OppAs have high structural similarity at the core framework (from domain **I** to domain **III**, including the segments of hinge strands shown in Fig. S2). However, they have different lengths and conformations for some loops (residues Phe205-Val210, Tyr352-Lys357, and Tyr534-Val539 on the query OppA from *B. longum* KACC91563) adjoining the ligand-binding sites. To consider the ligand binding effects from water-mediated H-bonding interactions, water molecules were divided into two groups, where the first group was filled into fixed water molecules without moving their coordinates in the *L. lactis* OppA binding pocket from any crystal structure. The second water group then collected moving water by combining with the interaction residues to fit into the ligand-binding cavity. The moving water group was picked in expending to 4.0 Å from the side chains of the specified amino acids in 5.0 Å of the binding cavity. This is considered on its explicit solvation in the second water boundary, where part of entire ligand binding site is for studying water-mediated H-binding interactions. The explicit solvation from the only second water group was optionally minimized to the Particle Mesh Ewald (PME) option in the solvated system to compute long range electrostatic interaction (Kutzman, T., et al, 2021), could be applied in prior-homology modeling of the the OppA of *B. longum* KACC91563. In the process of homology modeling for the OppA of *B. longum* KACC91563, the previous water group would have initial positions consistent with those located at the place of the hydrated binding pocket by overlaying the liganded conformers from the templates. As a result, the first water group had less movement within the binding site from the OppA model of *B. longum* KACC91563, regardless of the binding of its peptide-ligand (on liganded and unliganded states). However, the latter water group can be move into the designated cavity depending on special interaction properties (either electrostatic interactions or hydrophobic interactions) connected with their surrounding environment upon ligand binding. Despite the requirements to gather highly energetically compensation for water rearrangements of the second group, other strong hydrophobic interactions between the counterpart residues in the modelled OppA/peptide complex from the *B. longum* KACC91563 should repel the water from the cavity. Indeed, the second water group can contribute to the determinant factors against the specificity rather than the binding affinity of the peptide-ligand. The buried water effects were identified only for the periplasmic OppA in *Salmonella typhimurium* (Kornings, 1999; Lee, 2016), which did not act as a physical barrier to its ligand binding, but appear to affect ligand selectivity in particular, when comparing relative binding affinities between its tripeptide-ligands (Tame, 1999). Otherwise, the buried water-mediated interactions may mitigate unfavorable interactions (such as charge repulsions) by shielding between the OppA receptor and its peptide-ligand or replenish lost H-bond by arbitrating disrupted interactions within the binding pocket. Both roles of the buried water molecules will lead to capitalization of a peptide-ligand with a broad binding affinity to the OppA receptor, which is unlike highly specific ligand recognition within other ion channel transporters.

In different loop segments, hydrophobic loops exposed to water molecules should affect the structural stability of the OppA receptor due to their aggregation propensity. Loop segments of the query OppA align best to one template in one loop segment and another template in another loop segment. The loop sections at these positions could not align to those of template structures thoroughly, even if spatial restraints

were placed on the conformational similarities (as bonded terms of geometrical features) for residues within the vicinity of these loop spots. This is because homology restraints of templates cannot be applied in the loop segments of the query OppA models. These parts of loop segments expending into an entrance of the hydrated pocket were defined by the MODELER loop refinement based on the sequence-structure alignments between the query sequence and common local structures of the template OppA and AppA (PDB code of 3FTO and 5U4O). After defining an initial coordination of these loop segments in the OppA model structure, the local loop conformers were additionally optimized according to the CHARM-derived stereochemical and nonbonded restraints, respectively, for statistical preference of the different residue types and for the different side-chain rotamers in the regions (contained in the residues Phe205-Val210, Tyr352-Lys357, and Tyr534-Val539 within the query of OppA models).

By providing rough OppA models from *B. longum* KACC91563, we further optimize their local conformers of contiguous segments (i.e., loop and specific trace residues with high aggregation propensities). This indicated whether the residues of the segments are in the desired 3D environment of the open OppA structural states evaluated by the feature energy functions of the MODELER program. In the feature energy functions, geometric features (such as distances and dihedral angles) of the OppA and AppA templates are restrained by setting lower and upper bounds on their allowed values associated with each residue on the specific contiguous segments. This is in terms of the probability density function (PDF) (Sali, 2006) and discrete optimized protein energy (DOPE) (Basu et al., 2010) along the pre-structural alignment positions of the templates (as depictured in Fig. S2). Both feature energy functions of MODELER were also considered in solvation effects by adding the solvation energy term to other intramolecular energy terms. This occurred when water molecules in the best template of *L. lactis* OppA (PDB ID: 3FTO with open-unliganded folds) were reproduced into the OppA models and then treated as rigid bodies in the OppA models without steric hindrances of the OppA model self. The solvated OppA model was further optimized by conjugate gradient and simulated annealing optimization procedures using the CHARMm force field. As a result, the conserved structural characteristics are well reflected into spatial geometrical restraints in addition to the homology-derived restraints toward the OppA and AppA templates. If there are hydrophobic residues on the hydrated surface and polar residues in the hydrophobic cores of the query OppA, higher restraint violations of both features are given to the higher PDF and DOPE for each residue. A smaller PDF energy thus means that the OppA model better satisfies the homology-derived restraints. A lower DOPE also indicates a better model. The OppA model that had the lowest PDF and DOPE energies was chosen as the final model. The final OppA model was then optimized in terms of the relationships between the structural features of the templates, by the fitness of derived restraints from its current 3D environment and by studying Ramachandran plots (Richardson et al., 2003). The total PDF energy of the best OppA model was determined to be −14,251.4 kJ/mol in terms of geometric restraints, which is the sum of the scoring function values of all homology-derived pseudo-energy terms and stereochemical pseudo-energy terms. At the same time, the best model had the lowest DOPE score (−50,879.8 kJ/mol) as a conformational energy that measures the relative stability of a conformation with respect to other conformations of the loops for optimizing local structures of the query OppA.

2.3. Docking the bradykinin-like peptide ligand to the open-unliganded OppA model from B. longum KACC91563

We applied flexible docking (Venkatachalam et al., 2008), which allows for flexibility of the open-unliganded OppA receptor from *B. longum* KACC91563 during docking of the bradykinin-like peptide (as 9-mer peptide of RPPGFSPFA) ligand in the induced fit structure (as the closed-liganded OppA receptor from *L. lactis*). The side-chains of the specified amino acids in 5.0 Å of the active site (listed in Table 1) and in domain II (residues 294–517 shown in Fig. 1) were allowed to move during the peptide docking by generating ensembles of the OppA receptor conformations. Domain II of the OppA receptor was seen to reel up and down toward the peptide binding site between the open-unliganded and closed-ligand conformers from the template structures of the *L. lactis* OppA bounded to the 9-mer RPPGFSPFA peptide (PDB code 3DRG), as illustrated in Fig. S1. For the varied residues of the OppA receptor, 2,734 conformational states were optimized from the lowest energy of −106.49 kcal/mol to the highest energy of −77.95 kcal/mol by CHARM-based scoring functions.

However, the backbone and the side-chains of the OppA receptor, which were not specified, were fixed at their original positions. Also, in the specified residues of the OppA receptor, alanine, glycine, proline, and cysteine in disulfide bridges could not be optimized for the conformational ensembles due to the fewer rotatomers of their residues. Subsequently, the initial structure and pose of the 9-mer RPPGFSPFA peptide were coordinated from its X-ray structure (PDB code 3DRG). The sphere at the center of the active site where the peptide interactions were aligned to the site features of the OppA receptor (e.g. polar and apolar or hydrogen bond donor and acceptor) as hotspots was then placed into the coordinates 34.57, 2.81, 12.28, and 14.11. Ensemble dockings of the 9-mer RPPGFSPFA peptide poses in the site sphere on 2,734 of the OppA receptor conformational states were performed. Each docking pose was then subjected to the simulated annealing molecular dynamics (heating to 700 K for over 4,000 steps followed by cooling to 300 K for over 6,000 steps) process under the CHARMm force field (Mackerell et al., 2012) before the complex poses were scored. A final minimization of the 9-mer RPPGFSPFA peptide ligand in the rigid OppA receptor using non-softened potentials is performed. For each final complex pose, the CDOCKER interaction energy (Wu et al., 2003) as the CHARM energy (i.e. the interaction energy plus ligand strain) and the interaction energy alone were calculated and the top 10 scoring poses were retained. The presented top 10 docking poses (based on CDOCKER scoring) were likely to be the native docking conformation. The superposed main chains in the interaction interface between the best docked pose from *B. longum* KACC91563 and the reference complex structure from *L. lactis* (PDB code 3DRG) were observed within 2.0 Å of each other. The best docked pose of the closed-liganded OppA receptor from *B. longum* KACC91563 showed a top scoring complex pose (with the most negative

Table 1 Well-marked features of peptide-binding sites between three templates and query structures in 5.0 Å.

	AppA from *B. substilis* (template)	OppA from *L. lactis* (template)	PreZ from *E. faecalis* (template)	OppA from *B. longum* KACC91563 (query)
Binding pocket volume	~ 2600 Å^3 (Slotboom et al., 2009b; Pooltman, 2012)	~ 4900 Å^3 (Slotboom et al., 2009b)	~1600 Å^3 (Pooltman, 2012)	~2200 Å^3
PDB ID	1XOC	3DRG	4FAJ	Model structure
Peptide ligand	Val-Asp-Ser-Lys-Asn-Thr-Ser-Ser-Trp	Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Ala (bradykinin like peptide)	Leu-Val-Thr-Leu-Val-Phe-Val	Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Ala (bradykinin like peptide)
Binding pocket (Interaction residues with the peptide ligand)	Ile29, Gly30, Thr41, Asp42, Asp43, Ala44, Ser45, Thr46, Asn50, Thr59, Arg119, Thr122, Asn150, Asn151, Leu153, Asp154, Ser155, Ala157, Asn241, Ala263, Ser265, Val267, Leu301, Thr366, Asn367, Gly369, Asn370, Val372, Arg373, Ile376, Ala377, Trp398, Val402, Met405, Asn406, Val416, Gly418, Trp419, Ser420, Leu421, Ser422, Thr423, Gln427, Ile430, Tyr442, Tyr487, Pro489, Asn490, Asn491, Lys507, Arg508	Gln40, Ser41, Ser51, Asn55, Asp56, Ala57, Thr58, Phe59, Gly64, Thr75, Arg135, Ser139, Gln184, Ser185, Gly186, Asn187, Gly188, Tyr189, Leu191, Glu192, Thr193, Asn276, Gly277, Val279, Tyr301, Ser303, Ser350, Arg416, Gly418, Asn421, Ala422, Ile425, Ala426, Phe450, Trp453, Val454, Met457, Thr458, Asp470, Gly471, Ser472, Trp473, Ser474, Leu475, Ala476, Ser477, Asp483, Leu484, Tyr491, Phe493, Asn540, Met542, Asn544, Gly560, Ala561	Gly67, Thr68, Val79, Asp80, Gln81, Thr82, Ser83, Ile84, Ala88, Leu97, Leu161, Asp204, leu205, Ser207, Leu208, Thr209, Ala210, Tyr212, Ile280, Pro296, Leu297, Ala298, Asn318, Met320, Leu356, Ser418, Gly420, Phe422, Glu423, Ala426, Gly427, Ala450, Phe453, Met454, Leu457, leu466, Ser467, Gly468, Trp469, Gln470, Ala471, Asp472, Ser478, Met489, Phe531, Val533, Thr535, Ile551, Gly552	Ser46, Glu47, Pro54, Thr58, Glu59, Ala60, Gly61, Gly62, Gly63, Asp67, Tyr76, Ser136, Phe139, Pro156, Val184, Lys185, Ser186, Gly187, Ser188, His189, Ala190, Tyr191, Met192, Pro253, His269, Ala270, Ile271, Lys277, Gly292, Asn294, Leu296, Phe334, Asn391, Asp393, Thr395, Ala396, Trp399, Val400, Ser423, Phe426, Leu427, Val430, Asp431, Arg440, Ser441, Gly442, Trp443, Gly444, Pro445, Asp446, Tyr447, Pro448, Asn452, Leu457, Gly468, Ser470, Trp515, Gln517, Asn518, Ala519, Gly535, Gly536
The hydrophobic pocket for side chain 5 of nonapeptide	Thr41, Asp42, Asp43, Thr46, Arg119, Tyr268, Trp398, Leu401, Met405, Pro407, Trp419, Tyr442, Arg508 (Wilkinson et al., 2005)	Asn55, Asp56, Ala57, Phe450, Trp453, Val454, Trp473, Tyr491, Phe493 (Slotboom et al., 2009b)	Corresponding structurally to the hydrophobic pocket of PreZ for side chain 2 of cCF10: Val79, Asp80, Gln81, Met454, Met489, His491 (Pooltman, 2012)	Thr58, Glu59, Ala60, Gly61, Gly63, Ser136, Ser297, Ser423, Asp424, Glu425, Phe426, Leu427, Val430, Asp431, Gly442, Trp443, Gly444, Gly468, Asn469, Ser470, Gly536

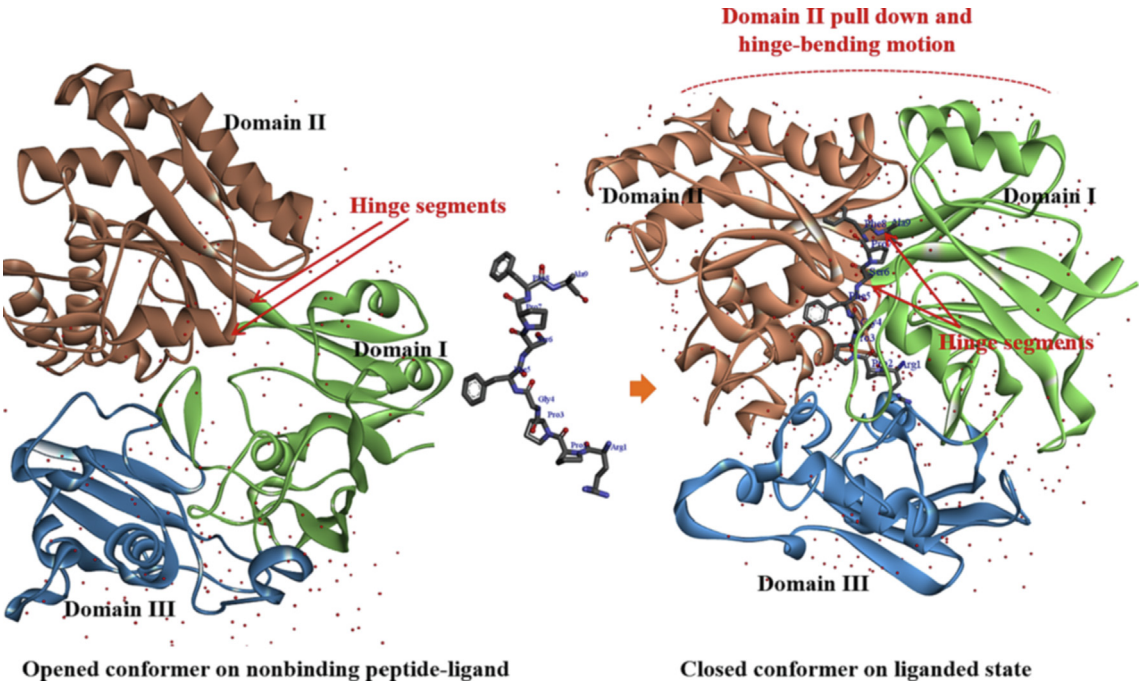

Fig. 1 The predicted OppA complex from *B. longum* KACC 91563 bounded with bradykinin-like peptide. On the bradykinin-like peptide binding, the equilibrium between open and closed conformations shifts toward the closed-liganded. The two α/β domains (domain **I** and **II** with inner β-sheets flanked by α-helices) are little affected by the rotation of its hinge region with two β-sheet (central residues Gly293 and Gln517 are green to orange transitions) owing to the presence of an extra domain (domain **III**) beneath the center of both α/β domains.

thus favorable to binding) with a CDOCKER binding energy of −337.98 kcal/mol, as shown in Fig. 1.

2.4. Pharmacophore generation from the OppA/peptide-ligand complex from B. longum KACC91563

The best complex poses of the OppA receptor/9-mer RPPGFSPFA peptide were utilized to explore the optimal intermolecular interactions with an ensemble of steric and electronic features (i.e., pharmacophores according to IUPAC definition). The docked RPPGFSPFA peptide poses were scanned within their binding pocket structure for distinctive pharmacophoric features that matched the OppA receptor-peptide ligand interactions. The pharmacophore ensemble was interpreted according to the topological feature descriptions of the peptide ligand as well as its corresponding 3D location and direction constraints, which are responsible for the peptide's specificity determinants that the OppA receptor from *B. longum* KACC91563 undergoes. The pharmacological interactions were generated by mapping H-bond acceptors, H-bond donors, and hydrophobic features within the binding site of the 9-mer RPPGFSPFA peptide by using the receptor-ligand pharmacophore generation protocol. The tying pairs of the H-bond donors and H-bond acceptors on the peptide directional features are adjacent to the surrounding OppA residues within a distance of 3.0 Å. On the other hand, the hydrophobic features on the scanned peptide ligand contain location constraints within 5.5 Å of the centroid of hydrophobic residues that have surface accessibility. At this time, the steric location of the OppA receptor near the binding site was reflected as excluded volume. We considered information regarding the Phe(**P5**) of the peptide ligand inserted into the hydrophobic cavity formed by Glu47, Thr58, Glu59, Ala60, Gly61, Leu427, and Val430 of the OppA from *B. longum* KACC91563 as criteria of a reasonable shape feature constraint to select and edit the best model among the top 10 pharmacophore models.

2.5. Residue alanine mutations and their impact on protein stability, binding affinity, and aggregation in the OppA/peptide-ligand complex from B. longum KACC91563

We used the structural knowledge of the OppA/9-mer RPPGFSPFA peptide complex from *B. longum* KACC91563 to consider the effect of the peptide binding site on the peptide substrate selectivity by focusing on the structural stability of the closed-liganded OppA and on the binding affinity of the 9-mer RPPGFSPFA peptide in the complex. We calculated the stability contributions of 62 key residues (listed in Table 1) in the peptide binding site on the basis of the difference between the folding free energy of the Ala mutated structure (i.e., single Ala mutation) and the wild type of the OppA receptor corresponding to the peptide binding site variants.

$$\Delta\Delta G_{mut} = \Delta\Delta G_{folding}^{(mutant)} - \Delta\Delta G_{folding}^{(wildtype)}$$

$$\Delta G_{folding} = \Delta G_{folded} - \Delta G_{unfolded}$$

All interaction energy terms of ΔG were calculated using the CHARMm force field and a generalized Born implicit solvent model. Van der Waals terms (E_{vdw}) and electrostatic interactions (ΔG_{elec}), an entropy contribution (-TSsc) related to changes in side-chain mobility, and a non-polar, surface-

dependent solvation energy terms (i.e., the cavitation energy, ΔG_{np}) were determined empirically. The total free energy, ΔG_{tot}, of folded or unfolded state was calculated as the following weighted sum of energy terms according to specific temperature (when T is 293 K).

$$\Delta G_{tot}(\mathrm{T}) = \mathrm{aEvdw} + \mathrm{b}\Delta G_{elec}(T) - c\mathrm{TSsc} + \Delta G_{np}$$

where a, b, and c are empirical scaling parameters. The optimal values of scaling factors a = b = 0.5 and c = 0.8 were estimated from the best fit to experimental mutation energies from the alanine scanning of the binding interface for temperature-dependent calculation of the mutation energy (stability and binding) (Yan, 2013). The energy function also contained terms for the side-chain and back-bone entropy to allow for room temperature-dependent calculations. Therefore, substituting a polar residue with a relatively non-polar alanine could result in changes in the conformation of neighboring residues. After building the OppA mutation structures of the alanine variants, the conformations of the mutated residues and neighbors were optimized further by using the MODELER protocol.

The mutation energy of binding was calculated as the free energy difference of the binding of the OppA receptor from *B. longum* KACC91563 and the 9-mer RPPGFSPFA peptide that contain the alanine variants of the 62 key residues in the mutated complex structure:

$$\Delta\Delta G_{mut} = \Delta\Delta G_{bind}^{(mutant)} - \Delta\Delta G_{bind}^{(wildtype)}$$

$$AB \leftrightarrow A + B, \Delta G_{bind} = \Delta G_{AB} - \Delta G_{A-B \text{ seperated}}$$

The mutation energy and total free energy for the binding affinity were calculated as the sum of the scaled van der Waals, electrostatic, non-polar, and entropy terms as well as the OppA structural stability. Mutation energy values and the corresponding effects of the complex alanine variants were evaluated, and these values were summated to give the peptide binding site properties to the different complex variants. This allowed for virtual determination of the relationships between the energy effects of the peptide binding site mutation on the OppA structural stability, the 9-mer RPPGFSPFA peptide binding affinity and the OppA receptor-related peptide substrate selectivity from *B. longum* KACC91563.

There was no self-aggregation of the active site on the open-unliganded OppA receptor from *B. longum* KACC91563 due to the steric effects from the presence of a large extra domain III (residues 84–210 shown in Fig. 1**)** beneath the center of both α/β domains. The result provides more support for the OppA/9-mer RPPGFSPFA peptide complex from *B. longum* KACC91563 in which self-aggregation cannot occur as a complement to the 9-mer RPPGFSPFA peptide due to direct hydrophobic interactions with the interaction site of the closed liganded OppA receptor. We predicted selective alanine mutation effects and the relative importance of hydrophobic interactions on the exposed hydrophobic surfaces of the peptide binding site across the OppA/9-mer RPPGFSPFA peptide complex from *B. longum* KACC91563 by calculating the spatial aggregation propensity (SAP) based on the pre-calculated solvent accessible area (SAA) of the fully exposed side-chain by the CHARMm force field. The SAP for the closed-liganded OppA receptor from the open-unliganded was obtained as the specified radii from the hydrophobicity scale of Black and Mould (Mould, 1991), and was added as atom and residue properties on the patches of the exposed hydrophobic residues. The hydrophobicity scale was normalized such that glycine had a hydrophobic value of zero. The amino acids that were more hydrophobic than glycine thus were positive, while more hydrophilic residues were negative than glycine. Therefore, the aggregation propensity of the OppA receptor conformers for the OppA receptor atom was defined as follows:

$$\sum[(\frac{\text{SAA of side chain atoms within radius R}}{\text{SAA of side chain atoms of fully exposed residues}}) \times \text{residue hydrophobicity}]$$

The SAP for each residue on the patches of the exposed hydrophobic residues was obtained as the average of its atomic aggregation scores. High aggregation scores (0.0 < SAP < 0.5) indicated highly exposed regions (in Fig. S4). An SAP map for the region was then generated by red color-coding (in Fig. 4), which allowed us to perform target mutations of the peptide binding site to enhance the peptide substrate specificity of the OppA receptor from *B. longum* KACC91563. Low SAP values (-0.5 < SAP < 0.0) indicated that the exposed surface was a hydrophilic region (blue in Fig. 4). This can be expected, as most of the OppA receptor surfaces exposed to water are usually hydrophilic. The changes of the SAP value of the OppA receptor between the open-unliganded and the closed-liganded conformers might provide information on the physicochemical properties of the substrate's specificity, based on the interaction site in the OppA/9-mer RPPGFSPFA peptide complex from *B. longum* KACC91563.

3. Results

3.1. Validation of the OppA model structures from ***B. longum*** *KACC91563 compared to experimentally determined template structures*

The OppA gene (Genbank code of AEI97628.1 (Slotboom, 2011)) from *B. longum* KACC91563 is encoded for protein residues 1 to 547. The query of OppA is composed of its *N*-terminal hydrophobic anchor (residues 1–36 of the extracellular OppA is bound to the membrane via the anchor region), two α/β domains (residues 37–83, 211–293, and 518–547 in domain I and 294–517 in domain II), and an extra domain (residues 84–210 in domain III), wherein it is connected by two hinge-strand fragments between domain I and domain II. The final OppA model was composed of integral main traces from Gly37 to C-terminal residue Gln547, except for its *N*-terminus anchor frame (residues 1–36), by homology modeling with the template structures (PDB ID: 3FTO, 5U4O in open-unliganded forms) to build its opened scaffold on nonbinding peptide-ligands. The topology of the OppA model structure from *B. longum* KACC91563 is shown in Fig. 1. To ensure precision of the backbone conformations in the OppA model, the residues on Ramachandran spaces were analyzed by comparing the steric effects of a residue's torsion angles (phi and psi angles) that were derived from how they fold (Tsai, 2008; Yeates 1993). Since the Ramachandran spaces could be classified into allowed and disallowed conformations, misfolded model structures were roughly defined

within the disallowed region with bad contacts. The ordered patterns of residues in the OppA model were found to have 469 residues (92.1 %) in the favored region, 26 residues (5.1 %) in the allowed region, and 14 residues (2.8 %) in the outlier region on the open-unliganded scaffolds. The OppA model was thus considered to be a good model structure for which there >90 % of the residues should be in the favored region of the protein (shown in Fig. S5).

The final OppA model (from Gly37 to Gln547 residues) was subjected to further validation of the compatibility modes of its own fold and sequence segments on the hydrate environment. Three methods (ERRAT2 (Hand, 2014), Z-score of ProSA (Sippl, 2007), and Verify3D (Bowie, 1997)) were then employed for an overall quality assessment of the OppA model structure for its fitness in its current 3D environment by comparing the characteristics of other experimental structures, including the OppA and AppA templates (PDB code of 3FTO and 5U4O in the open-unliganded structural state). ERRAT2 evaluates the structural errors of a model structure by distinguishing between the correct and incorrect fractions of nonbonded pairwise interactions (i.e., the six types of atom–atom interactions between carbon, nitrogen, and oxygen) within a specified distance limit (as predetermined distance cutoffs limits of 3.00 to 4.75 Å) determined from protein X-ray crystal structures. A larger ERRAT2 value then indicates a better refined model (ranges from 0 to 100 %) when compared with the score distribution of residues in correct protein structures. This is because error values of residues in the modelled structure higher than 95 % were not included in its overall quality. The Z-score of ProSA also evaluates the overall model quality, but is a recode of only the alpha carbons in the OppA model by measuring the deviations of total energy distributions for all determined proteins (X-ray, NMR in PDB), in contrast to ERRAT2. This protein structure analysis is frequently employed in the refinement and validation of experimental protein structures obtained X-ray analysis, NMR spectroscopy and in structure protein and modeling from theoretical calculations to check poteinal errors (Sippl, 2007). In particular, the Z-score better captured the solvent exposed residues within the soluble globular proteins than protein structures containing transmembrane domains. As presented in Fig. S5, the folding features of the hydrated OppA model are well reflected in the Z-score, where there were contacts in the middle region of the score scatterplot observed for the experimental protein structures. If the OppA model contained some errors, the Z-score would fall outside the range of characteristic values from known proteins. At the same, both Z-scores showed in only small differences of relative middle spots to induce the OppA conformation depending on open or closed states. Verify3D was additionally employed to inspect the validity of the OppA model structure by measuring the self-compatibility of its 3D structural profiles with its own protein sequence as the overall quality and local 3D-1D scores of the model structure in a fixed-length (typically about 5 to 20 residues) based on experimental data. An overall quality factor of 82.76 % was assigned by ERRAT2 for the OppA model, and the Z score of −8.64 in ProSA was within the range of native conformations, like other known proteins of similar size (refer to Table S2). For the open-unliganded structural state, the query of the OppA model had a lower ERRAT2 value and Z-score of ProSA than those of the OppA and AppA templates (OppA from *L. lactis* and AppA from *B. anthracis str. Ames* as PDB code of 3FTO and 5U4O), as shown in Table S2.

3.2. The shielding effect of hydrophobic binding pocket exposed to water on the open-unliganded OppA model

In the OppA from *B. longum* KACC91563, a large part of the surface from the binding pocket (listed in Table 1) is exposed to water rather than a hydrophobic environment such as the membrane interior. The effect of the shielding environment caused from the hydrated binding pocket can then be indicated by a hydrophobicity plot (depicted in Fig. S7), which shows the average hydrophobicity value of each residue with the neighboring four residues (Doolittle, 1982) instead of its own hydrophobicity. The fluctuations in the hydrophobicity index could be induced by shielding effects excited from the water. A thickset wave was observed in the query OppA of *B. longum* KACC91563 rather than the template OppA of *L. lactis* (PDB ID: 3FTO). This indirectly explained the hydrophobic binding pocket of query OppA that was extensively exposed to the surface, although complete shielding effects were not observed in water. Due to this, in the absence of ligand binding, an open structural state of the OppA model was more unstable compared to that of the template *L. lactis* OppA. The fluctuations of the hydrophobicity index further showed that the lower ERRAT2 value and Z-score of ProSA were not directly influenced by the quality of the OppA model, but by the distinct nature of the hydrophobic binding pocket from *B. longum* KACC91563 compared to that of the OppA and AppA templates. In the case of the open OppA model, the residue hydrophobicity of Thr58 was −0.70, but its neighboring 5-residues had an average hydrophobicity of 0.72. In contrast, Ala60 changed from 1.80 residue hydrophobicity to −0.64 for the neighboring 5-residues' average hydrophobicity (displayed in Fig. S4). Indeed, the helix's secondary structure of the fluctuated region on the query OppA (predicted from such as DSC (Sternberg, 1997) was based on the solvent exposed patterns of globular proteins) changed into a coil that corresponded to those of template PDB structures. If there is a significant difference in the hydrophobicity changes and in the underlying structure of residues that lie outside the binding pocket on an open-unliganded state, the difference in values in both hydrophobicity indexes of the residues becomes smaller in the closed-liganded state (in the case of Thr58 shown in Fig. S4). The residues with leading deviations between their hydrophobic properties and the hydrated environment were temporally recognized in the binding pocket depending on whether the ligand was bound. The significant residues were Thr58, Glu59, Ala60 (domain **I**), and Ala190 (domain **III**), which are consistent with the specific trace residues of the query OppA from *B. longum* KACC91563. Surprisingly, the significant residues of Thr58, Glu59, and Ala60 on the query OppA corresponded to a specific hydrophobic pocket (denoted Phe(**P5**) in Fig. 5) binding to the peptide-ligand. In contrast, the other residues of Phe334, Ala396, Val400, Phe426, Leu427, Val430 (domain **II**), Leu457, and Ala519 (domains **I** and **II**) were located in the conserved trace patches of the template and the query OppAs (expressed in Fig. 3). Such interaction residues showed changed characteristics as a result of the hydrated pocket, which is restricted by the structural compatibility of the query OppA model (Rousseau et al., 2020). In the

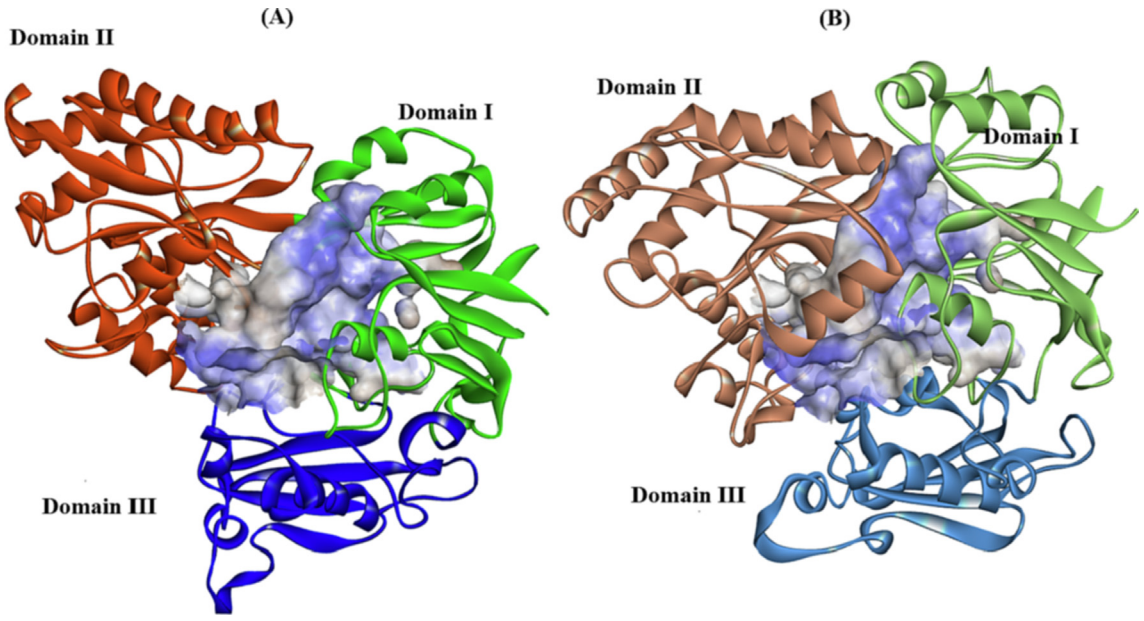

Fig. 2 Comparison of hydrophobic binding cavity volumes between opened (**A**) and closed (**B**) conformers of OppA from *B. longum* KACC 91563. The flexible volume was calculated from 2740 Å^3(**A**) to 2200 Å^3(**B**) by Delorme et al. (Delorme et al., 2001) when the OppA is closed to the bound peptide ligand.

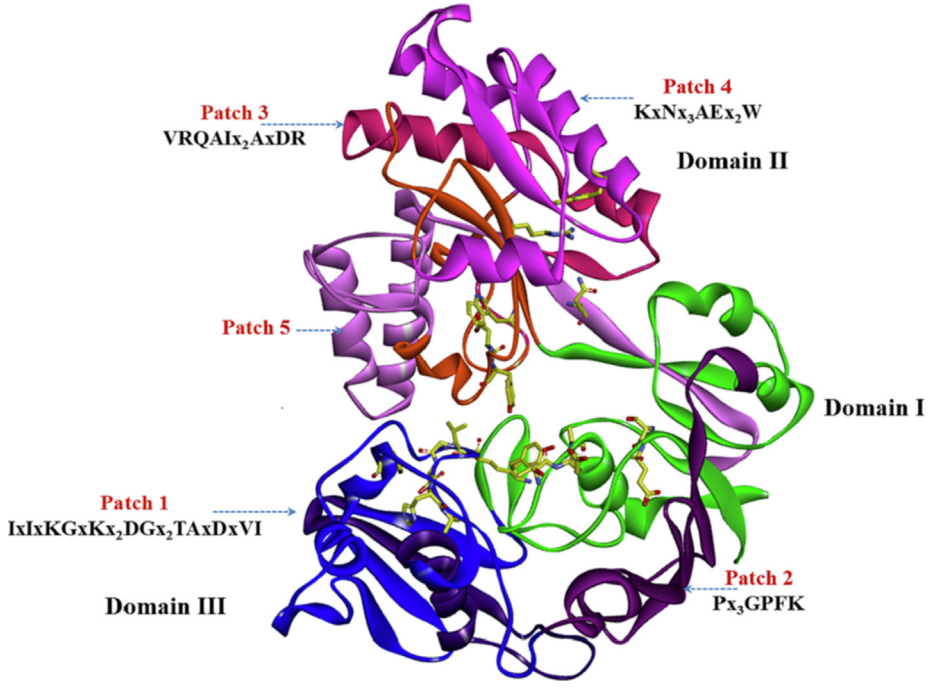

Fig. 3 Five conserved patches and specific trace residues for OppA model structure from *B. longum* KACC 91563. The specific trace residues shown in stick style highlighted in yellow, where the color scheme of the domains is the same as in Fig. 1. The conservation patterns of the patches represented by trace residues were analyzed based on the sequence group at a specified distance cut off of 21.0% between the templates and the query of OppAs shown in Fig. S14.

open states, the shielding effect of water into the OppA model was insufficient to reproduce the characterized structural stability of templates by only the homology-induced geometrical restrictions. This results in different patterns of hydrophilic and hydrophobic interaction residues in the hydrated binding pocket of the OppAs. Therefore, the OppA model structure had lower scores in the validated quality assessments (i.e. ERRAT2, Z-score of ProSA, and Verify3D) than those of the X-ray determined template OppA and AppA. However, the OppA query model is deemed a reliable model structure

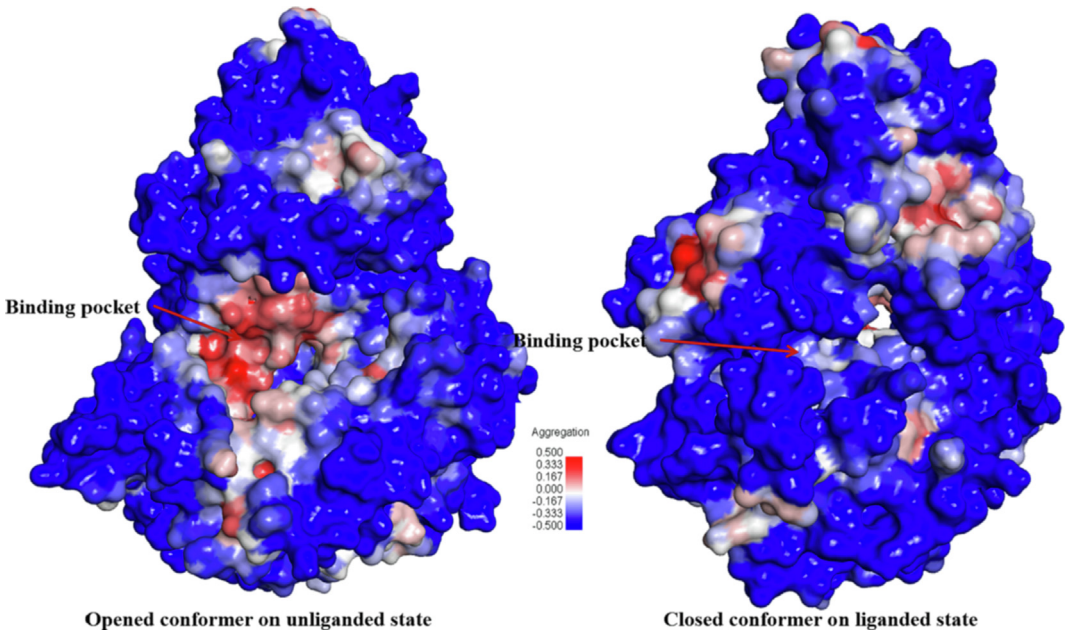

Fig. 4 Map of spatial aggregation propensity (SAP) for the OppA receptor from *B. longum* KACC 91563 in both open-unliganded and closed-liganded conformers. Positive SAP scores are red (hydrophobic) whereas negative SAP scores are blue (hydrophilic); therefore, a highly exposed hydrophobic fragment would be deep red and a highly exposed hydrophilic fragment would be deep blue.

as it passed standard qualities observed for other experimental protein structures (shown in Table S2 and Fig. S5).

Meanwhile, Verify3D allows us to reaffirm hydrophobic patches on surfaces of the hydrated binding pocket within the OppA model as well as the map of spatial aggregation propensity (in Fig. 4) by connoting whether its resides are in the desired 3D environment. In this context, the hydrophobic residues on the exposed OppA surfaces and the polar residues in the hydrophobic OppA cores received low Verify3D scores. If the surface patches of the OppA model show low Verify3D scores, this may indicate that the patch is interacting with other proteins (such as other solute binding protein transporters) and should be buried internally. This effect results from the incompatibility of the OppA interface regions with high aggregation propensity, but can be significant for the OppA transporter functions.

3.3. Structural characterization by conserved patches and specific trace residues of the OppA from B. longum KACC91563.

The theoretical protein size of the OppA model calculated from its sequence is 59 kDa and this is consistent with the other five clusters of extracellular SBPs in size ranging from 55 to 70 kDa (with 493 to 543 total residues) (Poolman et al., 2010; Monnet, 2003). From the perspective of its isoelectric point (pI) and the molecular weight (MW), the OppA from *B. longum* KACC91563 (pI:5.5, MW:59 kDa) is closer to AppA from *B. subtilis* (pI:6.0, MW:62 kDa) than that of the *L. lactis* OppA (pI:8.9, MW:66 kDa) when the biochemical properties were also calculated with those of OppA from *B. anthracis str. Ames* (pI:6.4, MW:58 kDa) and the *E. faecalis* PreZ (pI:8.1, MW:63 kDa). Nevertheless, the OppA model from *B. longum* KACC 91563 shows that the geometric topologies and orientations relative to each other between three domains (two α/β domains and an extra domain III) were more closely aligned to the conformal mapping of the *L. lactis* OppA (PDB ID: 3FTO, 1.18 Å of RMSD with 490 equivalent positions) than to the other template of AppA from *B. anthracis str. Ames* (PDB ID: 5U4O, 3.06 Å of RMSD with 468 equivalent positions); the similarities were overlaid by matching the molecular field (under conditions of 50 % steric and 50 % electrostatic field) between the OppA and AppA of the templates, and the query OppA were 0.51 and 0.44, respectively. There was a prediction that the conserved residues of OppA from *B. longum* KACC91563 were more closely mapped to the 3D spaces of the *L. lactis* OppA than were those of AppA from *B. anthracis str. Ames*.

Conserved functional patterns between four templates and the query OppA were represented by trace residues, which were identified from their sequences (Fig. S14) and mapped to their 3D-structures (shown in Fig. 3). The trace residues were further characterized by partitioning the conserved functional surface patches into subgroups according to the inferred roles of specific residues within the query OppA from their structures. In the present study, the trace residues were forced to make a direct connection between the conserved residues and their functional importance based on corresponding interaction residues (illustrated in Figs. 5, 6). Exposed trace residues that were more likely to be responsible for binding activity then could be distinguished from the buried trace residues that were more important for maintaining structural integrity at the hydrated binding sites of both the templates and the query OppA. Functional trace residues were clustered to define the distinct patterns within the query OppA structure from those of the other templates' structures. As shown in Fig. 3, the conserved patterns were mapped to the query OppA from *B. longum* KACC 91,563 by creating specific groups of trace residues at a special distance cutoff of 21.0 % (based on the protein sequence identities between the templates and

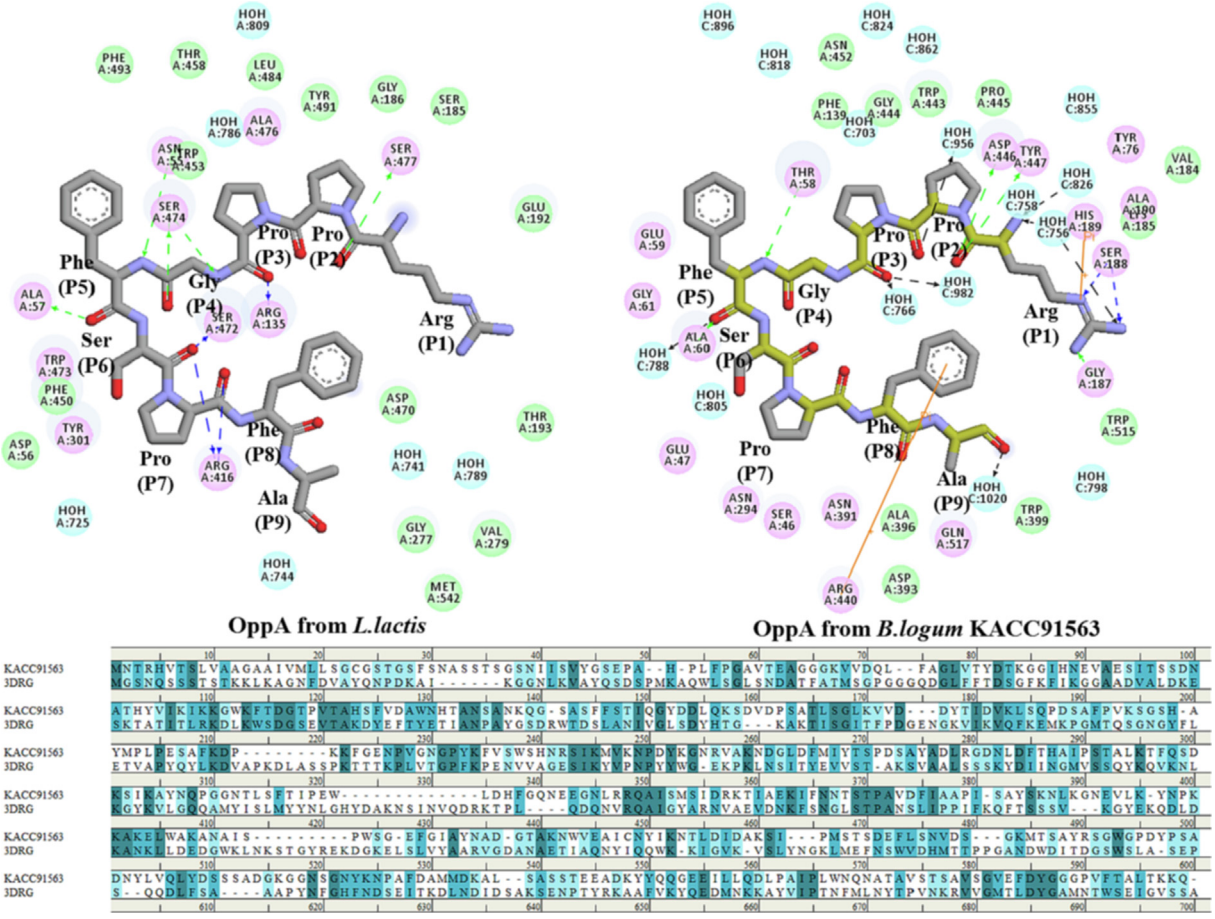

Fig. 5 Representation of bradykinin like peptide (RPPGFSPFA) binding to the OppAs from *L. lactis* (PDB ID: 3DRG) and the *B. logum* KACC91563 in different interaction registers.

the query OppA in Fig. 3, Fig. S14) into five patches. The conserved distance of 21.0 % should be considered a discriminating molecular basis based on either the electrostatic or hydrophobic properties of subgrouped trace residues from the four templates and the query OppA. The conserved trace patches were Asp94-Asn129 (in domain **III**), Asp201-Ser255 (in domains **I** and **III**), His306-Ala347 (domain **II**), Lys363-Ser432 (domain **II**), and Leu457-Ser528 (domains **I** and **II**) in order of the patch's numbers, respectively, inside the OppA from *B. longum* KACC91563. Moreover, four functional motifs (in Fig. S14) were identified in these conserved patches from both the templates and the query of OppA: IxIxKGxKx$_2$DGx$_2$TAxDxVI in the first patch, Px$_3$GPFK in the second patch, VRQAIx$_2$AxDR in the third patch, and KxNx$_3$AEx$_2$W in the fourth patch (where × is any residue). These conserved patterns did not match their common signature motif, as (LIVM)Ax$_2$(WI)x$_{1or2}$ (SN)(KE)Dx$_4$T(FY)x (LIV)Rx$_3$K was grouped into class 5 periplasmic and extracellular proteins (i.e., peptide and nickel-binding proteins) identified by *Saier et al* (Saier, 1993). This is because we only subjected class 5 external proteins from a few prokaryotes that were between the OppA and AppA templates and the query OppA. In view of its pattern motifs on the conserved patches, the pheromone binding protein of *E. faecalis* PreZ (PDB ID of 4FAJ) was not more distant from other templates and the query OppA, since its 3D-structural topologies and biological properties via the four pattern motifs were preserved. The specific trace residues of OppA from *B. longum* KACC91563 were Ser46, Glu47, Thr58, Glu59, Ala60, and Tyr76 in domain **I**, Trp399, Arg440, Trp443, Asp446, Tyr447, and Gln517 in domain **II**, and Phe139, Val184, Lys185, Gly187, Ser188, His189, and Ala190 in domain **III** of its extra domain located in the interaction interfaces (highlighted in yellow in Fig. 3). Note that the specific trace residues (in Table 2) in the interaction interfaces represent the four functional motifs in any conserved patch shown in Fig. 3. In particular, these specific trace residues would serve to guide the site-directed mutagenesis in silico (as in Fig. 7) for studying the OppA protein structure-functional relationship or as a target for structure-based pharmacophores (as seen in Fig. 8) by analyzing the interactions of a bound peptide-ligand in the OppA receptor from *B. longum* KACC91563.

3.4. Characterization of hydrophobic binding pocket by its aggregation propensity

The OppA of *B. longum* KACC91563 is quite limited in its mobility of the two α/β domains (domains **I** and **II**) along with rotation of the hinge-strands connecting both domains. In particular, the structural hindrances into the first α/β domain (domain **I**) are closely adjacent to the extra domain (domain **III**) and have a higher steric barrier than the second α/β domain

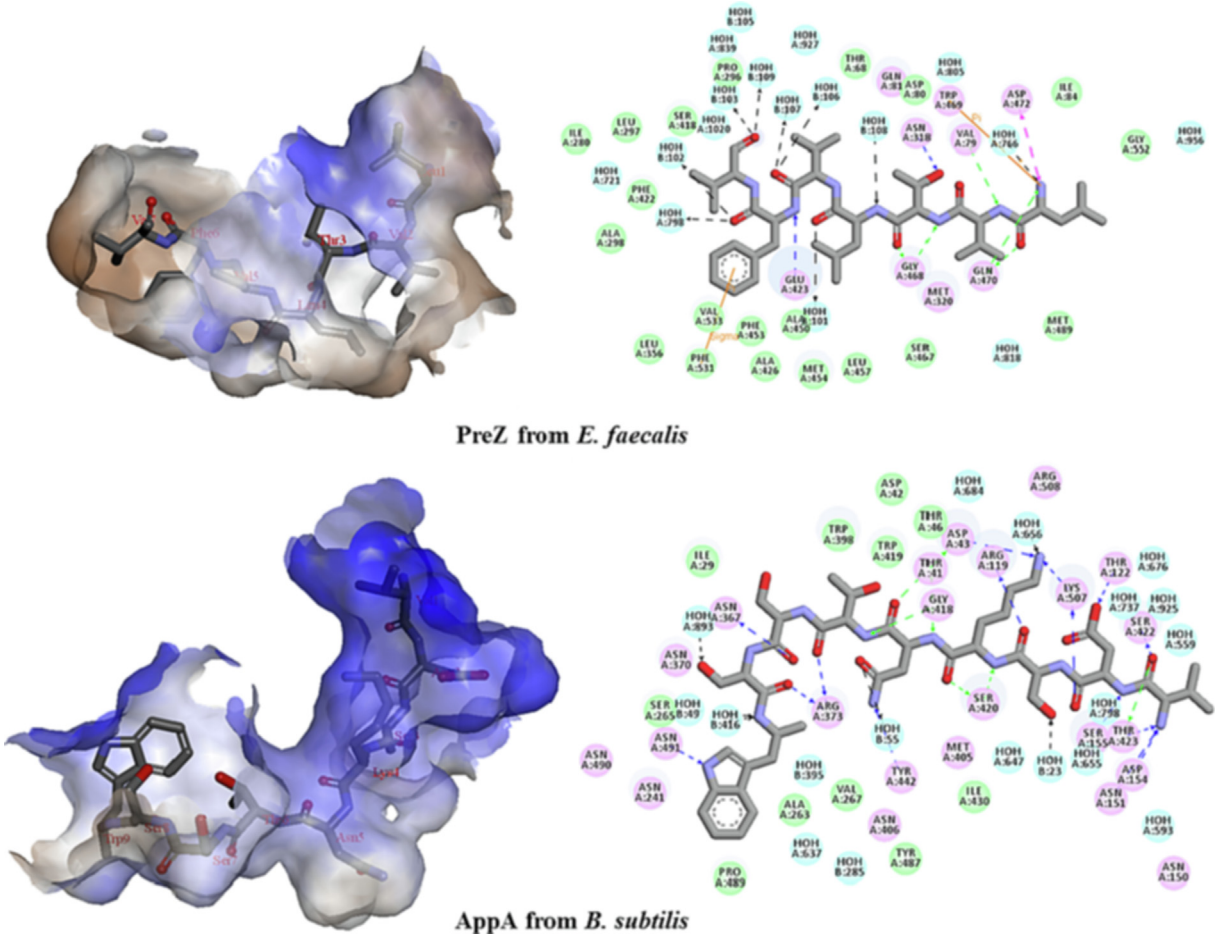

Fig. 6 The interaction contacts between the peptide ligands and two template receptors (PreZ and AppA) from their crystalized structures (PDB ID: 4FAJ and 1XOC). The interaction residues of PreZ from *E. faecalis* complexed with 7-mer peptide (LVTLVFV) and for 9-mer peptide (VDSKNTSSW) within AppA from *B. subtilis* are represented in their hydrophobic pockets.

Table 2 Strucutral important residues by identifying specific trace residues and highly exposed residues of the OppA from B. longum KACC91563.

Domain	Specific trace residues	Highly exposed residues
Domain I	Ser46, Glu47, Thr58, Glu59, Ala60, Tyr76	Pro51, Pro54, Ala56, Val57, Thr58, Phe70, Ala71, Val166
Domain II	Trp339, Arg440, Trp443, Asp446, Tyr447, Gln517	Asn294, Ile379, Ser380, Ser441, Trp443, Pro445, Tyr447, Pro448, Ser449, Ala450, Leu454, Gln456
Domain III	Phe139, Val184, Lys185, Gly187, Ser188, His189, Ala190	Lys185, Ser186, Tyr191, Met192, Lys200

(domain **II**), even in its open-unliganded conformer (as highlighted in Fig. 1). It was consistently observed that domain **II** can be picked up instead of domain **I** of the *L. lactis* OppA escaping the structural hindrances on its open-unliganded crystal structure (PDB code of 3FTO) represented in Fig. S1. Less steric hindrances of domain **II** from the query OppA allow easier access to binding sites than the other domains (domains **I** and **III**) in a closed structural state. Upon the open-unliganded conformer from the OppA of *B. longum* KACC91563, both α/β domains (domains **I** and **II**) were split into each other. The large extra domain (domain **III**) then exited in rigid movements by also twisting the interlinked β-strands in a hinge-region. In the process of binding the peptide-ligand, there may be movements in its two α/β domains (domains **I** and **II**) to spread out and to turn back toward the binding pocket together with the hinge-bending motions into the closed conformers. Specific volumes for binding pockets range from 1600 to 4900 $Å^3$ in the templates (summarized in Table 1 blow). This is because they all have an extra domain (domain **III**) extending the pocket, and also their hinge-region consists of two β-strands, where each β-strand is typically observed in 4–5 amino acids as a conserved structural feature (displayed in Fig. S2). For these conserved structural traits, their peptide-ligands accommodated in the binding pocket also have been shown to have similar preferences of 7-mer to 9-mer peptide length with a high affinity (dissociation constant K_D of μM to pM range (Poolman et al., 2010; Monnet, 2003) in their closed-liganded states. Likewise, the OppA models from *B. longum* KACC91563 have been observed in slightly reduced volume of the hydrophobic binding pocket from 2740 to 2200 $Å^3$ by trapping a bradykinin-like

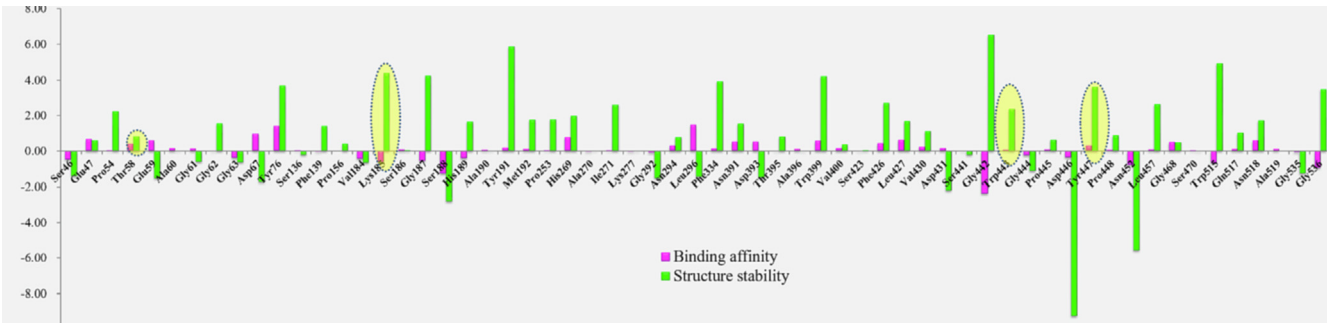

Fig. 7 Effect of single-point mutations on the OppA stability (under the OppA-RPPGFSPFA complex) and on the binding energy changes for the OppA-RPPGFSPFA binding by mutating each key residue in the binding site of the OppA-RPPGFSPFA complex to alanine. The mutation effect defined as follows: Stabilizing (mutation energy < -0.5 kcal/mol), neutral (-0.5 kcal/mol $<$ mutation energy < 0.5 kcal/mol), and destabilizing (> 0.5 kcal/mol).

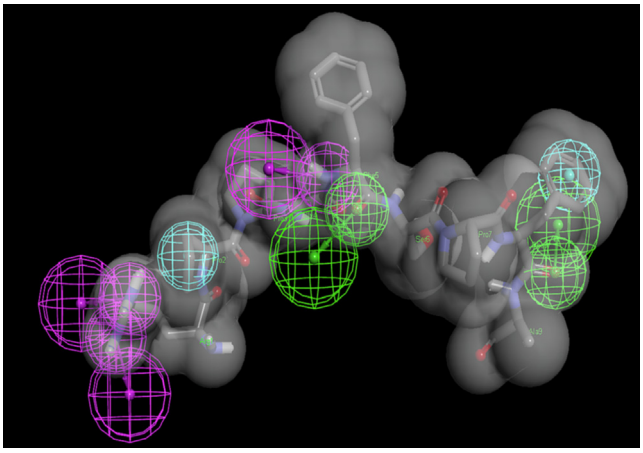

Fig. 8 The OppA-RPPGFSPFA pharmacophore generated based on the interaction in the complex. Receptor-ligand pharmacophore features essential to interact with key features on RPPGFSPFA. Pharmacophore features convert into a particular color-code (blue, hydrophobic; purple, H-bonding donor; green, H-bonding acceptor; gray shape, shape constraints in the binding pocket of the OppA).

peptide (as 9-mer peptide of RPPGFSPFA expressed in Figs. 2, 4). As another characteristic of the binding pocket, the OppA templates were shown to have discernible functional features by superposition of the specific hydrophobic cavity. The coordinated hydrophobic interfaces from the templates distinguish the sizes and preferences of the counterpart's hydrophobic residues from any location of the peptide ligands on the hydrated binding pocket (as shown in Figs. 5, 6). In the nonbinding ligand, hydrophobic interactions on the uncovered interfaces of the hydrated pocket can drive in the tendency to aggregate, which may decrease activity and the open-unliganded structural stability of the OppA receptor self. By predicting the OppA receptor surface sites that are likely to aggregate, we observed the aggregation propensity of the five conserved patches and of specific trace residues, into which the exposed hydrophobic residues were spatially closed. The aggregation propensity is an approximate indication of the equilibrium between the multiple structural states of the OppA receptor in the water solution and may not be an integer. In context, the aggregation propensity is likely to reflect an overall tendency to aggregate in the OppA receptor's binding pocket between the open-unliganded and closed-liganded states (represented in Fig. 4, Fig. S4). Therefore, the hydrated surface of the binding site has been depicted by specific trace residues of the OppA, Thr58, Lys185, Trp443, and Tyr447 (in Fig. S4), which were located on highly exposed hydrophobic regions in the binding pocket (residues with high aggregation propensity scores are colored red, while those with lower scores were colored blue, as shown in Fig. 4). In Fig. S4, specific trace residues (Thr58, Lys185, Trp443, and Tyr447) also had higher aggregation scores in the open-unliganded state than the closed-liganded state of the OppA models from *B. longum* KACC91563. Highly exposed regions (where the surface is red in Fig. 4) contain residues of Pro51, Pro54, Ala56, Val57, Thr58, Phe70, Ala71, and Val166 in domain **I**, and Asn294, Ile379, Ser380, Ser441, Trp443, Pro445, Tyr447, Pro448, Ser449, Ala450, Leu454, and Gln456 in domain **II**. The residues of Lys185, Ser186, Tyr191, Met192, and Lys200 in extra domain **III** also have a tendency to aggregate in a hydrated environment via internal hydrophobic interactions of the OppA from *B. longum* KACC91563 (in Table 2). In an open-unliganded state, the sites prone to aggregate in domain **II** showed broader areas and a greater number of localized aggregation sites on the OppA surfaces than domains **I** and **III**. This should be a driving force to move domain **II** into a binding site to have stronger hydrophobic interactions with domains **I** and **III** during the shifting of the pocket from some water to the docked peptide. Upon ligand-binding, the OppA conformer changed to increase the complex structural stability into the closed state and to allow the surface interfaces of the binding pocket to decrease the high aggregation propensity (as shown in Fig. 4). In particular, the specific Thr58, Lys185, Trp443, and Tyr447 trace residues on the binding pocket prominently decreased the likelihood of aggregation in the two structural states, as captured in Fig. S4; the aggregation prone scores for the four trace residues can be important for bradykinin-like peptide (RPPGFSPFA) binding, as they changed from 0.088, 0.167, 0.203, and 0.207 (at before the ligand-binding) to -0.225, -0.192, -0.058, and -0.097 (after the ligand-binding), respectively.

3.5. Oligopeptide binding pocket presented in specific trace residues

Bradykinin (RPPGFSPFR, pfam code of PF06753 and InterPro code of IPR009608) is well documented as a pharmacological therapy and is known as an ACE (angiotensin-converting-enzyme, CD143) inhibitor (Byong, 2017; Gauberti et al., 2016) to reduce high blood pressure. This bioactive peptide is an inflammatory mediator when kinin B1 receptor (BDKRB1) recruits neutrophil via the chemokine CXCL5 (CD185) (Ahluwalia et al., 2007). Physiologically, the capability to bind bradykinin for the LAB microbiota (such as *L. lactis, B. subtilis*, and *B. longum)* elicited a health effect from a nutrient source, which catches up peptides with antihypertensive activity (Rod, 2002; Rana 2011). Bradykinin is well-matched with the favored peptide-ligands from the templates of *L. lactis* OppA and *B. subtilis* AppA. The oligopeptide ABC importers from the LAB microbiota preferentially select proline-rich peptides, containing at least one branched residue of leucine, isoleucine, and valine (Slotboom et al., 2009b; Jeong et al., 2015; Wilkinson et al., 2005; van Wely, 2001) with no exact sequence preference. In this case, bradykinin is known for being able to bind the OppA and AppA templates with high affinity (dissociation constant K_D of 0.10 μM and 50.30 μM, respectively (van Wely, 2001)) in a closed-ligand conformation. The query OppA from *B. longum* KACC91563 not only has a similar overall structural architecture, but also operates in the organized hydrophobic binding pocket from the template of *L. lactis* OppA, as if bound with an equivalent bradykinin as a common oligopeptide-ligand. In this study, we predicted the complex structure of OppA from *B. longum* KACC91563 bound to a bradykinin like peptide (RPPGFSPFA) by docking the peptide (in Fig. 1). On the basis of the favor peptide's specificity, the 9-mer peptide of RPPGFSPFA has a fixed proline-rich peptide isoform, the backbone of which does not change like that of the X-ray determined conformer (PDB code of 3DRG) bound to *L. lactis* OppA. One of the binding factors was defined by a central hydrophobic pocket (i.e., the counterpart residues to the phenylalanine side chain at position 5) in the OppA interaction interfaces from *B. longum* KACC91563. The hydrophobic pocket's interior is covered with Glu47, Thr58, Glu59, Ala60, Gly61, Leu427, and Val430 residues (as presented in Table 1 and Fig. 5). The complex structure shows partitioned trace residues located in hydrophobic patches as a part of the binding pocket, which correspond to specific hydrophobic residues of RPPGFSPFA (listed in Table 1 and Fig. 3).

To understand the key residues that contribute to spatial proximity in the contact interface of 5.0 Å, we first evaluated the mutation energy effects of single Ala mutants on their structural stability and binding affinity within the OppA-peptide complex from *B. longum* KACC91563 (shown in Fig. 7 and Table S5). Since the hydrophobic binding site underlies better stability to reduce aggregation, it switches from an open to a closed conformation of the OppA upon 9-mer RPPGFSPFA binding. The Ala mutation effects tend to cause greater structural destabilization of the complex than decreasing the binding affinities to the molecular partners, as shown in Fig. 7. The highest energy of Ala mutations was located at residues of Pro54, Tyr76, Lys185, Gly187, Tyr191, Ile271, Phe334, Phe426, Gly442, Trp443, Tyr447, Leu457, Trp515, and Gly536, all of which were designated as structural destabilizing effects (mutation energy > 0.5 kcal/mol). They have mutation energies above 2.0 kcal/mol. In contrast, Ala mutants of Ser188, Asp446, and Asn452 altered the stabilizing (mutation energy < -0.5 kcal/mol) complex structure by smaller mutation energies than the −1.2 kcal/mol (in Table S5). Intriguingly, the Ala mutation effects were well projected into the specific Thr58, Lys185, Trp443, and Tyr447 trace residues with decreasing aggregative propensities (shown in Fig. S4), wherein the intrinsic moving domain **II** of the OppA is the defining feature of the switching conformation (open conformation) to domains **I** and **II** together (closed conformation). The four residues were also involved in the interaction interfaces of OppA with the RPPGFSPFA peptide. Thus, their Ala mutants could greatly lead to unsettling of its complex stability in the presence of 2- to 60-fold greater mutation energy effects than its binding affinity. It is noteworthy that a unifying feature in the OppA-RPPGFSPFA complex has been linked to its surface-exposed and non-conserved trace residues at the contact interfaces, which show the largest difference in complex stability from the patches of exposed hydrophobic residues (highlighted yellow residues in Fig. 3) and its specificity dispensed for peptide-ligand compared to the template complex of *L. lactis* OppA. There are specific trace residues of Gly61, Asp67, His269, Ala270, Leu296, Gly444, Asn452, Gly535, and Gly536 in the binding pocket (indicated in Table 1 and Fig. 5) of the OppA from *B. longum* KACC91563. The highly exposed residues of Ser186, Tyr191, and Met192 on the extra domain **III**, and Asn294 and Pro445 on domain **II**, the counter partners of which are Arg(**P1**), Pro(**P3**), and Pro (**P7**) on the RPPGFSPFA peptide, are prompted by a descending aggregation factor in Fig. S4. Among them, the biggest destabilizing Ala mutant effect of polar Tyr191 was envisaged as 5.87 kcal/mol by comparing that of hydrophobic Met192 at 1.77 kcal/mol from other exposed residues on extra domain **III.** In contrast, the lowest stabilizing Ala mutants were acidic Asp446 and polar Asn452 residues (in domain **II**), i.e. −9.24 kcal/mol and −5.57 kcal/mol, relative to Pro(**P2**) on the corresponding peptide-ligand. This strongly suggests that the stabilizing effects of induced fitness on the OppA derived from the favoritism of proline-rich (**P2**, **P3**, and **P7**) and hydrophobic Phe(**P5**), which also impacted the OppA-RPPGFSPFA interactions for *B. longum* KACC91563.

3.6. Identification of key functional residues by oligopeptide binding pocket based on structure-based pharmacophore models from OppA/RPPGFSPFA complex from B. longum KACC91563

If any trace residues have a linked 3D-pharmacophore arrangement to common features of the OppAs between *L. lactis* and *B. longum* KACC91563, they provide a knowledge-based description of interaction constraints with the RPPGFSPFA. This is more reliable when pharmacophores are depicted as an assemble of essential features (such as hydrophobic, H-bonding acceptor, and H-bonding donor) to trigger the closed-ligand binding OppA. Moreover, the pharmacophores corresponding to the 3D location of key residues from the OppA of *B. longum* KACC91563 are required for optimal intermolecular interactions with the RPPGFSPFA. They also signify the functional specificity of the trace residues as a binding-site characterization via virtual site-directed Ala

mutation against that of the wild type. These pharmacophores of major proline-rich and hydrophobic phenylalanine (**P5**) in turn impute featuring conductors of the key trace residues in the interaction interfaces. Two H-bonds are formed in the N atoms of Arg(**P1**) to the side chain of Gly187 and Ser188 within the OppA of *B. longum* KACC91563, while the H-bonds did not form with the OppA of *L. lactis* (as shown in Fig. 5). The formation of two H-bonds at Arg(**P1**) is noted along with the Ala mutation effects of Gly187 and Ser188 matched to −0.57 kcal/mol and −1.27 kcal/mol with stabilizing binding affinity. This is because the mutation effects lead to stronger H-bonding between the N atoms of Arg(*P*1) with the backbone of those Ala mutants. The five-membered ring of Pro(**P2**) has a hydrophobic point feature (represented as a pharmacophore blue color-code in Fig. 8) in neighboring Asp446 and Tyr447 of the OppA from *B. longum* KACC91563 without direct hydrophobic interactions between them. Remarkably, mutants Asp446 and Tyr447 to Ala have the biggest mutation effects (as −9.24 kcal/mol and 3.63 kcal/mol) on the complex structural stability, as shown in Fig. 5. The niche from Asp446 to Ala can be achieved within the hydrophobic constraints of Pro(**P2**) to complex structural stabilization. For other reasons, the Tyr447 residue is a fundamental indicator of the domain **II** conformer to close domain **I** on the OppA to provide structural stabilization to the complex via decreased aggregation poses (in Fig. S4). This primary determinant could not be replaced by the Ala mutant of Tyr447. Therefore, the Pro(**P2**) is limited in the rotating degree of the peptide backbone of RPPGFSPFA in conjugation with the other prolines (**P3** and **P7**). The fitted geometry seized with the location and direction of two H-bonds on the backbone amide and carbonyl group of Phe(**P5**) to the Thr58 and Ala60 residues of the OppA from *B. longum* KACC91563, compared to Asn55 and Ala57 in the OppA of *L. lactis*. Phe(**P5**) is inserted into the hydrophobic cavity formed by Glu47, Thr58, Glu59, Ala60, Gly61, Leu427, and Val430 of the OppA from *B. longum* KACC91563 (in Table 1). These hydrophobic interactions may be the original cause of the favorable features with shape constraints (gray colored shape in Fig. 8), which add to the surrounding assembly of the peptide-ligand's pharmacophores, as the excluded volume is defined by the cavity shape of the peptide binding site of the OppA. The leading hydrophobic shape constraints to branched hydrophobic residues were enforced rather than the projected hydrophobic point feature of the targeted Phe(**P5**) for making the interaction sites. In contrast to Phe(**P5**), the phenyl ring of Phe(**P8**) was projected as a hydrophobic interaction point into the Arg440 residues with π-π interaction (shown in Fig. 5). At the same time, the backbone carbonyl group of Phe(**P8**) as a H-bonding acceptor was realized by a hydrogen donor from the side chain of Arg440. The hydrophobic shape constant of Phe(**P8**) then expanded to surround Phe(**P8**) by Leu296, Phe334, Ala396, Trp399, and Trp515 residues on the OppA from *B. longum* KACC91563. The role of Arg440 was superseded by the Arg416 residue on *L. lactis* OppA. As a result, the configuration of pharmacophores on the structure-based complex is an arrangement of seven chemical features of two H-bond acceptors (HBA) and the features of three H-bond donor (HBD) vectors as well as two hydrophobic (HY) point features with location constraints that represent locations in space within a given radius of 1.6 Å. The shape constant then adds to the surrounding pharmacophores that were generated from the binding site expanding to 3.5 Å, which corresponds to the OppA-RPPGFSPFA interactions in the docked complex model from *B. longum* KACC91563. The created pharmacophore model with a 12.42 selectivity score value is shown in Fig. 8. This reflects physiologically broad peptide selectivity from the OppA transporter of the LAB microbiota, which is reflected by selectivity score of the identified pharmacophores on the OppA from *B. longum* KACC91563. In Fig. 5, the *L. lactis* OppA had exclusive H-bonds with the peptide backbone of RPPGFSPFA via residues Asn55, Ala57, Arg135, Arg416, Ser472, Ser474, and Ser477 of the OppA, with the exclusion of H-bonding to the side chain of the peptide. When residues Ser472 and Ser474 of *L. lactis* OppA consistently are interact with other peptide (e.g., RDMPIQAF) backbones, irrespective of their binding affinity in even opened-ligand conformations (Slotboom et al., 2009b), as shown in Fig. S15. In contrast, the H-bond patterns of *L. lactis* OppA in contrast to the H-bond register on the OppA from *B. longum* KACC91563 formed between the residues Gly187, Ser188, and Arg440 of the OppA and the side chain of Arg(**P1**) and Phe(**P8**), in addition to achieving H-bonding to the peptide backbone for the residues Thr58, Ala60, Asp446, and Tyr447 of the OppA. In only OppA of *B. longum* KACC91563, the H-bonding register with a side chain of Arg(**P1**) and Phe(**P8**) was specially defined to a specific determinant for the RPPGFSPFA binding, together with a hydrophobic interaction point (π-π interaction) between Arg400 and Phe(**P8**) to coincide with the shape constant of the targeted interaction sites. This is caused by the different sizes of the peptide-binding cavity (4900 $Å^3$ vs 2700 $Å^3$), which explains the lower size limitation and the more spacious specificity of the peptide-ligand in the *L. lactis* OppA than in the OppA from *B. longum* KACC91563.

To confirm the key features of the OppA from *B. longum* KACC91563, we screened two similar 9-mer peptides with a central hydrophobic residue of leucine at positions 5 or 6 (Leu(**P5**) or Leu(**P6**) represented in Fig. S3) by focusing on the assembly of seven pharmacophores based on the interactions with the RPPGFSPFA. The dissociation constants K_D of binding to *L. lactis* OppA are known to be 1.2 μM and 4.2 μM for SLSQ**L**SSQS and SLSQS**L**SQS in the closed-ligand conformers, respectively (Slotboom, 2011). Before screening the two peptides, we docked both peptides into the OppA binding site from *B. longum* KACC91563, which was similar to docking the RPPGFSPFA peptide from the initial locations of their X-ray crystal structures (PDB code of 3YRA and 3RYB) shown in Fig.S16. This validates the pharmacophore model as required features allows us to determine that these pharmacophores are likely to be involved in common features related to the peptide binding affinity and contribute to interference fit through their peptide specificity. It was observed as a common feature that the geometrical pharmacophores of two H-bonds with the peptide backbone of Leu (**P5**) on the SLSQ**L**SSQS peptide to the key residues of Thr58 and Ala60 on the OppA. In the case of the SLSQS**L**SQS peptide, Leu(**P6**) was perfectly reproduced into the H-bonding record to complement Thr58 and Ala60, as shown in Fig. S8. In both peptides, the central leucine residues at positions 5 and 6 were well fitted into the hydrophobic shape constant, which was nailed into Phe(**P5**) from the RPPGFSPFA bound to the OppA from *B. longum* KACC91563. This shows that both crystalized complex structures (PDB code of 3RYA and 3RYB) bound to *L. lactis* OppA, where the backbone of

Ser(**P1**) to Gln(**P8**) on the SLSQ**L**SSQS peptide was superimposed on the backbone of Leu(**P2**) to Ser(**P9**) from the SLSQS**L**SQS. Indeed, the conformers of other bound peptide-ligands with different lengths were very similar in terms of the bound bradykinin structure to *L. lactis* OppA (see Fig. S6 with RMSD 0.29–0.4 Å for their alpha carbons). The other five features comprised two HBDs(**P1**), one HBA (**P8**), and two HY points (**P5** and **P8**) that were not at all fitted in either peptide. These five features must be adapted to the specific binding determinants of the peptide-ligand to the OppA from *B. longum* KACC91563. The *L. lactis* OppA accepted the five features on its peptide binding sites. This shed light on how *L. lactis* OppA accomplishes broad specificity with its peptide variant length (4–35 residues) but not a striking trace feature against the OppA from *B. longum* KACC91563. On the other hand, there are distinct interactions to fix the positions of the N(**P1**) and C(**P8**) termini of the complex of OppA-RPPGFSPFA from *B. longum* KACC91563 such as side chain-specific interactions with the OppA.

In *B. subtilis*, the AppA of another oligopeptide importer (Wilkinson et al., 2005) also favors hydrophobic nanopeptide substances with overlapping specificity to its OppA. Fig. 6 shows the interaction interfaces of the template structure (PDB code of 1XOC) from *B. subtilis* AppA with its bound nonapeptide-ligand as the VDSKNTSSW peptide. In comparison of the binding constants in the *B. subtilis* AppA and the OppA from *B. longum* KACC91563 (Fig. 5 and Table S3), the peptide backbone of Val(**P1**) forms two H-bonds to the side chains of Asp154, Thr423, and Asn151 on the AppA. Indeed, the hydrophobic shape constant that surrounds the residues of Trp398, Leu401, Met405, Met407, and Trp419 encloses Asn(**P5**) more by positioning two H-bonds between their backbones of Asp43, Gly418, and Asn(**P5**). On the other hand, the Asp(**P2**) does not match the hydrophobic point constant but Trp(**P9**) replaces half occupancy in the conjugated hydrophobic shape constant on the side chain of the VDSKNTSSW peptide by the above pharmacophore fitness (shown in Fig. S9). No proline residues on the VDSKNTSSW peptide are likely to maintain the backbone conformer, similar to the RPPGFSPFA binding within the contact interfaces of AppA through the H-bond network with both main chains. There appears to be no necessary requirement for AppA to be proline-rich or for the branched hydrophobic(**P5**) on the nonapeptide-ligand to have binding specificity. The AppA could bind to the bradykinin of RPPGFSPFR (the dissociation constant K_D binding to *B. subtilis* AppA is 50.3 μM (van Wely, 2001)).

Moreover, the peptide selectivity from *E. facalis* PreZ like other OppA and AppA templates, mainly originated from the hydrophobic shape interactions between the side chain of cCF10 (7-mer pheromone peptide of LVTLVFV with a high binding affinity of 10 pM to the PreZ (Pooltman et al., 2000)) and the extracellular pheromone receptor (in Table S3 and Fig. 6). In particular, the side chain of Leu(**P1**) is fixed via a salt bridge with the Asp472 residue in the PreZ's hydrophobic pocket corresponding to Val(**P2**). The side chain of hydrophilic Thr(**P3**) is lined by the neighboring hydrophilic residues of Asn328, Gly468, and Gln470 on the PreZ with distinct binding constants from the effects of aggregation pores against other OppA and AppA templates. As a common binding constant of the central hydrophobic pore (in Table 1), the widened binding cavity to accommodate the larger peptide-ligands from *E. facalis* PreZ to *L. lactis* OppA is associated with a steric clash between the domain **II** and side chain of the **P5** residue on the bound peptide by decreasing the water mediated H-bondings to bury the peptide in the binding pocket (shown in Fig.S10). Therefore, the spatial aggregation propensity (SAP) in the binding site of the extracellular SBP importer plays a vital role in the specific interaction determinant for peptide binding (as shown in Fig. 4).

4. Discussion

The ecological success of *lactic acid bacteria* relies on the substrate specificities of the extracellular proteins represented by the solute-binding proteins (e.g. OppA) of the ABC transport system and the metabolic activities of the glycoside hydrolase enzymes. This is possibly due to the exogenous protective effects accelerating the survival rate of the probiotic cells within the host gastrointestinal tract (Mandal, 2016; Prakash, 2018; van Sinderen, 2015). The metabolic substrates of the human milk oligosaccharides enhance the *Bifidobacterial* probiotic adhesion and colonization to the host intestinal mucosa, especially during the host's early life (Prakash, 2018; van Sinderen, 2015; Ventura et al., 2018). In particular, members of the *Bifidobacterium* are among the key bacterial components (i.e., *B. longum*, *B. bifidum*, and *B. breve* species) of the infant intestines. Notably, a higher abundance of *Bifidobacteria* was observed in infant *lactic acid bacteria* than in adults (from the BioProject PRJNA33914 aimed at exploring the vertical transmission of the microbiota from mothers to corresponding infants). Among *lactic acid bacteria*, *Bifidobacterium longum* KACC91563, isolated from fecal samples of healthy Korean neonates, has the capability to alleviate food allergy effects (Jang et al., 2016; Shanahan, 2010). The ability of *B. longum* KACC91563 was confirmed in suppressing allergic diarrhea from a mouse food allergy model induced by using ovalbumin and alum. Interestingly, *B. longum* KACC91563 neither curtailed T_H2 cytokine levels nor influenced anti-inflammatory cytokine IL-10 by $Foxp3^+$ Treg cells (Yin et al., 2015). As opposed to the absence of an effect on the *T*-cell immune responses, *B. longum* KACC91563 exerts an influence on decreasing mast cell numbers via increasing the Annexin V^+ apoptotic bone marrow-derived mast cells. This results in dramatically decreased food allergies in a mouse model. Thereupon, *B. longum* KACC91563-derived extracellular vesicles substantially ameliorated potent food allergy by instigating mast cell apoptosis (Jang et al., 2016). Notably, the extracellular vesicles from *B. longum* KACC91563 are more internalized in mediating unknown receptors in bone marrow-derived mast cells than in phagocytes by dendritic cells. The unknown receptor to the extracellular vesicles (with the majority being 60 nm) would be different from exosome transporters CD63 and CD82 that convey vesicles with virus-like properties. Therefore, the extracellular vesicles from *B. longum* KACC91563 are enriched for the family's 5 solute binding protein, DNA, and lipids inside their lipid bilayer that more effectively transport these molecules than their soluble forms. Specifically, the extracellular vesicles of *B. longum* KACC91563 are composed of an important component of OppA importer as a part of the family's 5 solute binding protein. This is because the OppA importer determines the substrate specificity of family's 5 solute binding protein through physiological routes in the host

intestine. The intestinal *L. lactis* populations and *B. longum* KACC91563 are also closely connected with selection of favored oligopeptides containing at least one branched residue of leucine, isoleucine, and valine via the OppA importer. This is because both organisms are auxotrophic for branched amino acids, and they favor proline-rich caseins in milk media (9.8 % of proline residues in milk casein (Slotboom et al., 2009b)) as a nutrient source of these amino acids to facilitate their growth. To satisfy the demands of the branched amino acids related to the physiology of both organisms, *L. lactis* and *B. longum* KACC91563 proteolyze exogenous proteins (e.g. α-, β-, and κ-caseins in milk) to oligopeptides, which are transported by the OppABCDF membrane-spanning channel and further metabolized in the cytoplasm (shown in Fig. S11).

Their interplay augments the immune balance in the nutritional amino acid substrate (Ma et al., 2018) by adjusting for competent assistant-host immune crosstalk. Leucine and its metabolite glutamine can be strength sources to surpass the mucosal barrier functions, the epithelial villus height in the intestine, and small intestinal growth (Qiao, 2015; Yin et al., 2017). In particular, only leucine triggers α-defensin secretion from Paneth cells among the 20 mammalian amino acids, thereby conducting surveillance in the intestine to maintain an intestinal homeostatic response to inner environment facts. In contrast, the isoleucine is prominent in the order of lymphocytes, eosinophils, and neutrophils on the host by inducing β-defensin with a downward trend in TNF-α and IL-10. Isoleucines also promote mucosal immunity and maintain intestinal integrity. Therefore, both leucine and isoleucine stimulate secretion of intestinal SIgA (secretory immunoglobulin A), which is the most abundant antibody to improve the mucosal surface defense in the intestinal mucosa, thereby leading to inhibition of pathogen introgression into the lamina propria. It is well known that glutamate and glutamine as leucine metabolites (see Fig. S13(a)) comprise almost 70 % of protein-unbound amino acids (accounting for 5 ∼ 10 % of total amino acids) in human milk. In fact, the free glutamate level is 40-fold higher in milk compared to plasma (Lönnerdal, 2013). The effects of glutamate and glutamine on immune parameters may diminish susceptibility relevant to allergic disease (Hogenkamp, 2020) and infections in a developing neonate over the course of lactation, since at birth, increased susceptibility in the neonatal immune system is derived from an immature intestinal barrier and incomplete microbial colonization compared to adults.

As an important organ in the host body, the intestine acts as a shared junction of nutrient digestion and absorption, as well as microbiota colonization, and also locates immune cells. The geographical proximate of the intestinal tract contributes a high level of immunity activity to eliminate ingested pathogens from its motility. A characterization of the molecular interactions might provide valuable insights to decipher how mutualistic symbiosis-host' *T*-cell communication operates in amino acids for mediating intestinal immunity. In particular, the activated $CD8^+$ *T*-cells take up tryptophan from TAT1 (aromatic amino acid as Phe, Tyr, Trp transporter also known as SLC16A10) into the cells by enkindlling *T*-cell receptors and profusing leucine from CD98 (aliases LAT1 of L-leucine transporter) through animation of the mTORC (mechanistic target of rapamycin complex) signaling pathway (Huston, 2016; Yan et al., 2017; Qiao, 2017; Bauer, 2018; Chen 2020; Faierweather, 2018; Cantrell et al., 2013). Notably, the competence of mTORC can control *T*-cell fate (denoted in Table S4), and mTORC activity calls for the presence of amino acids. As such, the leucine antagonist of *N*-acetyl-leucine amide blocks mTORC activity (Yin et al., 2017) and can inhibit *T*-cell function as an immunosuppressive agent. A dearth of leucine also promotes mTORC hindrance to promote infectious tolerance by producing more regulatory *T*-cells. If the *T*-cells have a defective CD98 of the leucine transporter, the CD98 null-$CD4^+$ *T*-cell has no effect on the antigen receptor ligation without differentiating into T_H1 or T_H17 cells of *T*-helper cells. The *T*-cell activation is afforded by intracellular leucine uptake determined not by the extracellular leucine concentration, but by the expressed leucine transporter CD98 and glutamine transporter ASC2 under the control of the *T*-cell receptor. In this context, it is noteworthy that CD98 and ASC2 coupled with the transport of leucine are positively associated with *T*-cell activation as well as T_H1 and T_H17 cell differentiation to produce IL-2 and IFN-γ and secrete IL-17, IL-21, and IL-22, respectively (designated in Fig. S12, S13).

Based on their substrate specificity, leucine or tryptophan requisites are connected to the fixed sequential selectivity of only C-terminal residue from the antigenic peptide on the host TAP-mediated MHC class I presentation for $CD8^+$ *T*-cell immunity. Alternately, encountering the *T*-cell receptor by the host self-peptide/MHC complex, leucine, and tryptophan is not indispensable to IL-7-induced $CD8^+$ *T*-cell survival, but is required for sustaining native $CD8^+$ *T*-cell size and $CD8^+$ *T*-cell growth by interaction determinants of the IL-7 and IL-7 receptor (CD127). Consequently, both leucine and tryptophan uptake of the *T*-cells (Ma et al., 2018; Sokolo, 2018; Valdés-Ferrer, 2019; Khan, 2014) are key switches for metabolism reprogramming of immune-activated *T*-cells, leading to an adjustment in adaptive immune responses. In this regard, in an infection resolution, antigen-specific $CD8^+$ *T*-cells decrease anabolism to the catabolism of nutrients from high to low mTOR activity, and thence from the effector *T*-cells to memory *T*-cells.

On the other hand, the fitted leucine or tryptophan in the C-terminal residue on the peptide substrate binding to host TAP redounded upon allosteric crosstalk between the TMD and the NBD, which subsequently triggers peptide transition and ATP hydrolysis. Moreover, the intestinal commensal *Lactobacillus* can catabolize tryptophan to indole 3-aldehyde as an AhR (aryl hydrocarbon receptor) ligand to protect against mucosal inflammation by its tryptophanase (see Fig. S13(b)). Indole is also a signal molecule of bacterial physiology for antibiotic resistance and biofilm formation, whereas in non-indole producing bacteria, indole and its offshoot inhibit quorum sensing and modulate harmfulness factors in intestinal microenvironment. Of note, indole itself further vitalizes enteroendocrine L cells to produce glucagon-like peptide-1 (GLP-1) into insulin secretion by pancreatic β-cells.

Functionally, neonates used glutamate and glutamine uptook from breastmilk by the intestines that have furnished the growth of intestinal epithelial cells and maturation of the intestinal barrier to support protective effects. As another causative factor, except for shifting immune responses from T_H2 to T_H1 cells, double-stranded RNA from intestinal commensal *lactic acid bacteria* (including *Lactobacillus* and CRISPR family), but not pathogenic bacteria, is a natural ligand for Toll-like receptors TLR3 (CD283) and TLR9 (CD289) as a sensor of the commensal bacteria to trigger IFN-β production by

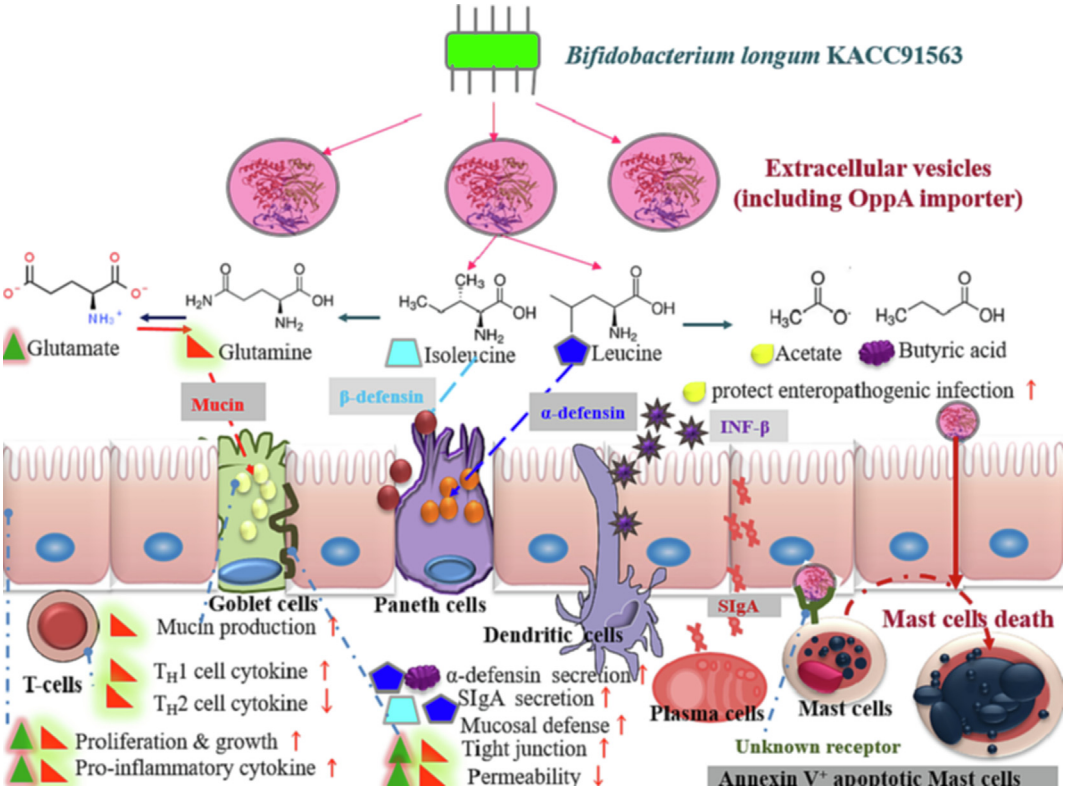

Fig. 9 The host immunity effects of amino acid substrate from probiotic *Bifidobacterium longum* KACC91563 in the intestine. The OppA importer in *B. longum* KACC91563-derived extracellular vesicles then plays a pivotal role in determining the substrate specificity that supports the probiotic effect.

bone marrow dendritic cells in the small intestine (Tsuji et. al, 2013). In light of this, $CD11c^+$ dendric cells in lamina propia or Peyer's patches may also induce IFN-β production to stabilize protective immunity against viruses and pathogenic bacteria. Besides secreting IFN-β, *Lactic acid bacteria* keep their intestinal community by producing lactic acid and by lowering pH related to detergent overgrowth harm of pathogenic bacteria (Mandal, 2016).

In addition, the selected branched residues (e.g. leucine) are precursors of acetate in short-chain fatty acids (such as the acetoacetate shown in Fig. S13(A, B)). The acetate products from *B. longum* assist the defense function of host epithelial cells (Ohno et. al, 2011), which are unlikely to kill pathogenic enterobacteria via α-defensin secretion by butyric acid (one of the short chain fatty acids) and leucine in the Paneth cells (Ayabe, 2019). The propionic and butyric acid metabolites are known to exert an antagonistic effect on colon carcinoma cell proliferation, thereby inducing apoptosis of cancer cells (Stenico, 2014). Hence, we attempted to explicate the relationship between the substrate's specificity from the OppA importer and the probiotic effects of *B. longum* KACC91563 in the host intestine, as illustrated in Fig. 9. Some factors such as pathogen invasions, proinflammatory cytokines, and toxins contribute to lowering of the epithelial barrier functions in the host intestine. The probiotic effects of *B. longum* KACC91563 were attributed to enhancement of the epithelial barrier by several different strain specific pathways to prevent the strong adhesion of pathogens.

5. Conclusions

Bifidobacterium longum KACC91563, a subspecies of *Bifidobacterium* genus, belongs to the *lactic acid bacteria* (LAB), a probiotic genus isolated from the feces of healthy Korean neonates. The intestinal *B. longum* KACC91563 is auxotrophic for branched amino acids, and favors proline-rich caseins in milk media as a nutrient source of these amino acids to facilitate their growth. Oligopeptide-binding protein A (OppA) determines the substrate specificity of oligopeptides through OppABCDF membrane spanning channel where acts as an oligopeptide transporter into the cell. OppAs from gram positive bacteria have a broad substrate specificity with their oligopeptide variant length (4–35 residues), instead of incomplete sequence discernments of the peptide itself. In the current study, we performed homology modeling for an unknown OppA structure (the open-unliganded conformation) from *B. longum* KACC91563 and docked the bradykinin-like peptide ligand (as 9-mer peptide of RPPGFSPFA) to the open-unliganded OppA model. As a result, we provide both model structures of the open-unliganded and the closed-liganded OppA from *B. longum* KACC91563 for its substrate specificity with clear distinctions of structural features from the searched template OppAs (*L. latics* OppA and *B. subtilis* AppA with high affinity of dissociation constant K_D of 0.10 μM and 50.30 μM, respectively in a closed-ligand conformation). Further, we performed knowledge-based modeling for an unknown OppA structure from B. *longum* KACC91563 by projecting its biophysical information into conserved pictures of the OppAs of the LAB family and by making its molecular masking characteristics distinctive from those of the LAB member's OppAs from previous studies. In essence, the query OppA from *B. longum* KACC91563 not only has a similar overall structural architecture (with two α/β domains connected by two hinge-strand fragments and an extra domain III), but

also operates in the organized hydrophobic binding pocket from the template of *L. lactis* OppA, as if bound with an equivalent bradykinin as a common oligopeptide-ligand. In the study, we note that the intrinsic moving domain II of the OppA is the defining feature of the switching conformation (open conformation) to the two domains I and II together (closed conformation) by representing the hydrophobicity value of each residue and the map of spatial aggregation propensity (SAP) in the substrate-binding pocket. Intriguingly, this was well projected into the specific Thr58, Lys185, Trp443, and Tyr447 trace residues with aggregative propensities and their alanine mutation effects. There are specific trace residues of Gly61, ASP67, His269, Ala270, Leu296, Gly444, Asn452, Gly535, and Gly536 in the binding pocket of the OppA B. *longum* KACC91563. The highly exposed residues of Ser186, Tyr191, and Met192 on the extra domain III, Asn 294, and Pro445 on domain II, the counter partners of which are Arg(**P1**), Pro(**P3**), and Pro(**P7**) on the RPPGFSPFA peptide, are prompted by a decreasing aggregation factor. The stabilizing effects of induced fitness on the OppA came from preference of proline-rich (**P2**, **P3**, and **P7**) and hydrophobic Phe(**P5**), which also impacted the OppA-RPPGFSPFA interactions for B. *longum* KACC91563. In addition, the 3D-pharmacophore arrangement of interaction features of the OppA on the structure-based complex is an arrangement of seven chemical features of two H-bond acceptors (HBA) and the features of three H-bond donors (HBD) vectors as well as two hydrophobic (HY) point features with location constraints when the leading hydrophobic shape constraints to branched hydrophobic residues were enforced as a surrounding assembly of the peptide-ligand's pharmacophores with the excluded volume. In only OppA of *B. longum* KACC91563, the H-bonding register with a side chain of Arg(**P1**) and Phe(**P8**) was defined to a specific determinant for RPPGFSPFA binding, together with a hydrophobic interaction point (π-π interaction) between the Arg400 and the Phe(**P8**) to coincide with the shape constant of the target interaction sites with a 2700 Å peptide-binding cavity. The specialized structure–function relationship from the OppA importer of *B. longum* KACC91563 with its structural and functional determinants was first identified by this study. This could provide an abstract of substrate specificity of the OppA importer from *B. longum* KACC91563 from into unique immunological properties of the host organism. Our results may expand the perspective in the screening process of the immunological peptides (e.g. bradykinin) and in the diagnostic fields from the OppA and AppA antigens from the gram-positive pathogen (including *Clostridium difficile*).

CRediT authorship contribution statement

Han-Ha Chai: Conceptualization, Data Curation, Methodology, Investigation, Project administration, Writing-original draft, Writing-review & editing. **Jun-Sang Ham:** Resources, Validation. **Tae-Hun Kim:** Funding acquisition, Supervision. **Dajeong Lim:** Formal analysis, Writing-review & editing.

Declaration of Competing Interest

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

Acknowledgments

This research was carried out with the support of "Molecular characterization of structural and functional variations of the cluster of differentiation family genes in pigs", Project No. PJ01268802 Rural Development Administration, Republic of Korea.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.arabjc.2022.104198.

References

Agus, A., Planchais, J., Sokolo, H., 2018. Gut Microbiota Regulation of tryptophan metabolism in health and disease. Cell Host Microbe. 23, 716–724.

Ananieva, E.A., Powell, J.D., Huston, S.M., 2016. Leucine metabolism in T cell activation: mTOR signaling and beyond. Adv Nutr. 7, 7985–8055.

Belkaid, Y., Hand, T.W., 2014. Role of the microbiota in immunity and inflammation. Cell 157, 121–141.

Berntsoon, R.P., Schuurman-Wolters, G.K., Dunny, G., Slotboom, D. J., Poolman, B., 2012. Structure and mode of peptide binding of pheromone receptor PrgZ. J. Biol. Chem. 287, 37165–37170.

Berntsson, R.P., Oktaviani, N.A., Slotboom, D.J., et al, 2009a. Selenomethionine incorporation in proteins expressed in Lactococcus lactis. Protein Sci. 18, 1121–1127.

Berntsson, R.P., Doeven, M.K., Slotboom, D.J., et al, 2009b. The structural basis for peptide selection by the transport receptor OppA. EMBO J. 28, 1332–1340.

Berntsson, R.P., Smits, S.H., Poolman, B., et al, 2010. A structural classification of substrate-binding proteins. FEBS Lett. 584, 2606–2617.

Berntsson, R.P., Thunnissen, A.M., Poolman, B., Slotboom, D.J., 2011. Importance of a hydrophobic pocket for peptide binding in lactococcal OppA. J. Bacteriol. 193, 4254–4256.

Best, R.B., Zhu, X., Mackerell Jr., A.D., et al, 2012. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain $\chi_{(1)}$ and $\chi_{(2)}$ dihedral angles. J. Chem. Theory Comput. 8, 3257–3273.

Black, S.D., Mould, D.R., 1991. Development of hydrophobicity parameters to analyze proteins which bear posttranslational or cotranslational modifications. Anal. Biochem. 193, 72–82.

Bröer, S., Fairweather, S.J., 2018. Amino acid transport across the mammalian intestine. Compr Physiol. 9, 343–373.

Castrignanó, T., de Meo, P.D., Cozzetto, D., Talamo, I.G., Tramontano, A., 2006. The PMDB protein model database. Nucleic Acids Res. 34, D306–D309.

Chang, O.K., Seol, K.H., Ham, J.S., et al, 2013. Casein hydrolysis by Bifidobacterium longum KACC91563 and antioxidant activities of peptides derived therefrom. J. Dairy Sci. 96, 5544–5555.

Chen, L., Cruz, A., Roe, D.R., Kutzman, T., et al, 2021. Thermodynamic decomposion of solvation free energies with Particle Mesh Ewald and long-range lennard-jones interaction in grid inhomogeneous solvation theory. J. Chem. Theory Comput. 17, 2714–2724.

Chennamestty, N., Voynov, V., Kayser, V., Helk, B., Trout, B.L., 2010. Prediction of aggregation prone region of therapeutic proteins. J. Phys. Chem. B.114, 6614–6624.

Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 2, 1511–1519.

Cuneo, M.J., Beese, L.S., Hellinga, H.W., 2009. Structural analysis of semi-specific oligosaccharide recognition by a cellulose-binding protein of thermotoga maritima reveals adaptations for functional diversification of the oligopeptide periplasmic binding protein fold. J. Biol. Chem. 284, 33217–33223.

Dahl, D.B., Bohannan, Z., Mo, Q., Vannucci, M., Tsai, J., 2008. Assessing side-chain perturbations of the protein backbone: a knowledge-based classification of residue Ramachandran space. J. Mol. Biol. 378, 749–758.

Dasgupta, A., Sureka, K., Basu, J., et al, 2010. An oligopeptide transporter of mycobacterium tuberculosis regulates cytokine

release and apoptosis of infected macrophages. PLoS ONE. 5, e12225.

Dassault Systémes BIOVIA, Discovery studio modeling environment, release R2 dassault systems, San Diago 2017.

Davies, T.G., Hubbard, R.E., Tame, J.R., 1999. Relating structure to thermodynamics: the crystal structures and binding affinity of eight OppA-peptide complexes. Protein Sci. 8, 1432–1444.

Detmers, F.J., Lanfermeijer, F.C., Poolman, B., et al, 2000. Combinatorial peptide libraries reveal the ligand-binding mechanism of the oligopeptide receptor OppA of Lactococcus lactis. Proc. Natl. Acad. U. S. A. 97, 12487–12492.

Duchene, J., Lecomte, F., Ahluwalia, A., et al, 2007. A novel inflammatory pathway involved in leukocyte recruitment: role for the kinin B1 receptor and the chemokine CXCL5. J. Immunol. 179, 4849–4856.

Eisenberg, D., Lüthy, R., Bowie, J.U., 1997. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 277, 396–404.

Eric, B.M., Deog, H.O., Byong, H.L., 2017. Bioactive peptides. Food 6, 32.

Fang, M., Bauer, C.E., 2018. Regulation of stringent factor by branched-chain amino acids. Proc. Natl. Acad. Sci. U. S. A. 115, 6446–6451.

Felder, C.B., Graul, R.C., Lee, A.Y., Merkle, H.P., Sadee, W., 1999. The Venus flytrap of periplasmic binding proteins: an ancient protein module presents in multiple drug receptors. AAPS. PharmSci. 1, E2.

Feng, Z.M., Li, T.J., Yin, Y.L., et al, 2015. Monosodium L-gulatamate and dietary fat differently modify the composition of the intestinal microbiota in growing pigs. Obes Facts. 8, 87–100.

Fukuda, S., Toh, H., Ohno, H., et al, 2011. Bifidobacteria can protect from emteropathogenic infection through production of acetate. Nature 469, 543–547.

Garai, P., Chandra, K., Chakravortty, D., 2017. Bacterial peptide transporters: messengers of nutrition to virulence. Virulence 8, 297–309.

Gioa, D.D., Gaggía, F., Baffoni, L., Stenico, V., 2014. Beneficial microbe in fermented and functional foods. In: Rai, V.R., Bai, J.A. (Eds.), Role of Bifidobacteria in the Production of Bioactive Compounds and Detoxification of Harmful Compounds. CRC Press, New York (Chapter 16).

Guédon, E., Serror, P., Ehrlich, S.D., Renault, P., Delorme, C., 2001. Pleiotropic transcriptional repressor CodY senses the intracellular pool of branched-chain amino acids in Lactococcus lactis. Mol. Microbiol. 40, 1227–1239.

Ha, G.E., Chang, O.K., Jeong, S.G., et al, 2015. Identification of antihypertensive peptides derived from low molecular weight casein hydrolysates generated during fermentation of Bifidobacterium longum KACC91563. Korean J. Food sci. Anim. Resour. 35, 738–747.

Ham, J.S., Lee, T., Kim, H., et al, 2011. Complete genome sequence of Bifidobacterium longum subsp. longum KACC91563. J. Bacteriol. 193, 5044.

Heddle, J., Scott, D.J., Unzai, S., Park, S.Y., Tame, J.R., 2003. Crystal structures of the liganded and unliganded nickel-binding protein NikA from Escherichia coli. J. Biol. Chem. 278, 50322–50329.

Hinz, U., Consortium, U., 2010. From protein sequences to 3D-structures and beyond: the example of the UniProt knowledgebase. Cell Mol. Life Sci. 67, 1049-1064.

Houben, B., Michiels, E., Rousseau, F., et al, 2020. Autonomous aggregation suppression by acidic residues explains why chaperones favour basic residues. EMBO J. 39, e102864.

Hu, J., Nie, Y., Yan, X., et al, 2017. Leucine residues reactive oxygen species levels via an energy metabolism switch by activation of the mTOR-HIF-1α pathway in porcine intestinal epithelial cells. Int. J. Biochem. Cell Biol. 89, 42–56.

Kawashima, T., Kosaka, A., Tsuji, N.M., et al, 2013. Double-stranded RNA of intestinal commensal but not pathogenic bacteria triggers production of protective interferon-β. Immunity 38, 1187–1197.

Kim, J.H., Jeun, E.J., Jang, M.H., et al, 2016. Extracellular vesicle-derived protein from Bifidobacterium longum alleviates food allergy through mast cell suppression. J. Allergy Clin. Immunol. 137, 507–516.

Kim, J.J., Khan, W.I., 2014. 5-HT_7 receptor signaling: improved therapeutic strategy in gut disorders. Front. Behav. Neurosci. 8, 396.

King, R.D., Saqi, M., Sayle, R., Sternberg, M.J., 1997. DSC: public domain protein secondary structure predication. Comput. Appl. Biosci. 13, 473–474.

Klepsch, M.M., Kovermann, M., Berntsson, R.-P.-A., et al, 2011. Escherichia coli peptide binding protein OppA has a preference for positively charged peptides. J. Mol. Biol. 414, 75–85.

Koska, J., Spassov, V.Z., Venkatachalam, C.M., et al, 2008. Fully automated molecular mechanics based induced fit protein-ligand docking method. J. Chem. Inf. Model. 48, 1965–1973.

Kozlowicz, B.K., Shi, K., Dunny, G.M., et al, 2006. Molecular basis for control of conjugation by bacterial pheromone and inhibitor peptides. Mol. Microbiol. 62, 958–969.

Kunji, E.R., Fang, G., Konings, W.N., et al, 1998. Reconstruction of the proteolytic pathway for use of β-casein by Lactococcus lactis. Mol. Microbiol. 27, 1107–1118.

Kyte, J., Doolittle, R.F., 1982. A Simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.

Lanfermeijer, F.C., Picon, A., Konings, W.N., Poolman, B., 1999. Kinetics and consequences of binding of nona- and dodecapeptides to the oligopeptide binding protein (OppA) of Lactococcus lactis. Biochemistry 38, 14440–14450.

Lanfermeijer, F.C., Detmers, F.J., Konings, W.N., Poolman, B., 2000. On the binding mechanism of the peptide receptor of the oligopeptide transport system of Lactococcus lactis. EMBO J. 19, 3649–3656.

Lassaux, P., Peri, C., Bolognesi, M., et al, 2013. A structure-based strategy for epitope discovery in Burkholderia pseudomallei OppA antigen. Structure 21, 167–175.

Leclerc, P.L., Gauthier, S.F., Bachelard, H., Santure, M., Rod, D., 2002. Antihypertensive activity of casein-enriched milk fermented by Lactobacillus helveticus. Inter. Dairy J. 12, 995–1004.

Levdikov, V.M., Blagova, E.V., Wilkinson, A.J., et al, 2005. The structure of the oligopeptide-binding protein, AppA, form Bacillus subtilis in complex with a nonpeptide. J. Mol. Biol. 345, 879–892.

Lovell, S.C., Davis, I.W., Richardson, D.C., et al, 2003. Structure validation by Calpha geometry: phi, psi, and C-beta deviation. Proteins 50, 437–450.

Lugli, G.A., Mancino, W., Ventura, M., et al, 2018. Reconstruction of the Bifidobacterial pan-secretome reveals the network of extracellular interactions between Bifidobacteria and the infant gut. Appl. Environ Microbiol. 84, e00796–e818.

Ma, N., Guo, P., Ma, X., et al, 2018. Nutrients mediate intestinal bacteria-mucosal immune crosstalk. Front Immunol. 9, 5.

Mackenzie, A.K., Valegård, K., Andersson, I., et al, 2010. Crystal structures of an oligopeptide-binding protein from the biosynthetic pathway of the β-lactamase inhibitor clavulanic acid. J. Mol. Biol. 396, 332–344.

Macros-Contreras, O.A., Gauberti, M., et al, 2016. Hyperfibrinolysis increases blood-brain barrier permeability by a plasmin- and bradykinin-dependent mechanism. Blood 128, 2423–2434.

Mao, B., Pear, M.R., McCammon, J.A., Quiocho, F.A., 1982. Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. J. Biol. Chem. 257, 1131–1133.

Maurer, M., de Beer, B.A., Oostenbrink, C., 2016. Calculation of relative binding free energy in the water-filled active site of oligopeptide-binding protein A. Molecules 21, 499.

Monnet, V., 2003. Bacterial oligopeptide-binding proteins. Cell Mol. Life Sci. 60, 2100–2114.

Nickitenko, A.V., Trakhanov, S., Quiocho, F.A., 1995. 2Å resolution structure of DppA, a periplasmic dipeptide transport/chemosensory receptor. Biochemistry 34, 16585–16595.

Nie, C., He, T., Zhang, W., Zhang, G., Ma, X., 2018. Branched chain amino acids: Beyond nutrition metabolism. Int. J. Mol. Sci. 19, 954.
O'Callaghan, A., van Sinderen, D., 2016. Bifidobacteria and their role as members of the human gut microbiota. Front Microbiol. 7, 925.
O'Callaghan, A., Bottacini, F., O'Connell, Motherway, M, van Sinderen, D., 2015. Pangenome analysis of Bifidobacterium longum and site-directed mutagenesis through by-pass of restriction-modification systems. BMC Genomics 16, 832.
Picon, A., van Wely, K.H.M., 2001. Peptide binding to the Bacillus subtilis oligopeptide-binding proteins OppA and AppA. Mol. Biol. Today 2, 21–25.
Powell, J.D., Pollizzi, K.N., Heikamp, E.B., Horton, M.R., 2012. Regulation of immune responses by mTOR. Annu. Rev. Immunol 30, 39–68.
Quintero-Villegas, A., Valdés-Ferrer, S.I., 2019. Role of 5-HT_7 receptors in the immune system in health and disease. Mole. Med. 26, 2.
Ren, W., Liu, G., Yin, Y., et al, 2017. Amino-acid transporters in T-cell activation and differation. Cell Death Dis. 8, e2655.
Ren, M., Zhang, S.H., Zeng, X.F., Liu, H., Qiao, S.Y., 2015. Branched-chain amino acids are beneficial to maintain growth performance and intestinal immune-related function in weaned piglets fed protein restricted diet. Asian-Australas J. Anim. Sci. 28, 1742–1750.
Sarkar, A., Mandal, S., 2016. Bifidobacteria-insight into clinical outcomes and mechanisms of its probiotic action. Microbiol Res. 192, 159–171.
Shanahan, F., 2010. Probiotics in perspective. Gastroenterology 139, 1808–1812.
Sharma, V., Mobeen, F., Prakash, T., 2018. Exploration of survival traits, probiotic determinants, host interactions, and functional evolution of Bifidobacterial genomes using comparative genomics. Genes 9, 477.
Sharma, S., Singh, R., Rana, S., 2011. Bioactive peptides: A review. Inter. J. Bioautomation 15, 223–250.
Shen, M.Y., Sali, A., 2006. Statistical potential for assessment and prediction of protein structures. Protein Sci. 15, 2507–2524.
Shi, K., Brown, C.V., Earhart, C.A., et al, 2005. Structure of peptide sex pheromone receptor PrgX and PrgX/pheromone complexes and regulation of conjugation in Enterococcus faecalis. Proc. Natl. Acad. Sci. U. S. A. 102, 18596–18601.
Shin, T.S., Kim, J.H., Kim, Y.K., et al, 2010. Extracellular vesicles are key intercellular mediators in the development of immune dysfunction to allergens in the airways. Allergy 65, 1256–1265.
Sinclair, L.V., Rolf, J., Cantrell, D.A., et al, 2013. Control of amino acid transport by antigen receptors coordinates the metabolic reprogramming essential for T-cell differentiation. Nat. Immunol. 14, 500–508.
Sleigh, S.H., Tame, J.R., Dodson, E.J., Wilkinson, A.J., 1997. Peptide binding in OppA, the crystal structures of the periplasmic oligopeptide binding protein in the unliganded form and in complex with lysyllysine. Biochemistry 36, 9747–9758.
Sleigh, S.H., Seavers, P.R., Wilkinson, A.J., Ladbury, J.E., Tame, J. R., 1999. Crystallographic and calorimetric analysis of peptide binding to OppA protein. J. Mol. Biol. 291, 393–415.
Song, W., Li, D., Tao, L., Luo, Q., Chen, L., 2020. Solute carrier transporters: the metabolic gatekeepers of immune cells. Acta Pharm. Sin. B 10, 61–78.
Spassov, V.Z., Yan, L., 2013. pH-Selective mutagenesis of protein-protein interfaces: *in silico* design of therapeutic antibodies with prolonged half-life. Proteins 81, 704–714.
Takakuwa, A., Nakamura, K., Ayabe, T.T., 2019. Butyric acid and leucine induce α-defensin secretion from small intestinal paneth cells. Nutrients 11, 2817.
Tam, R., Saier Jr., M.H., 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiol Rev. 57, 320–346.
Tame, J.R., Murshudov, G.N., Wilkinson, A.J., et al, 1994. The structural basis of sequence-independent peptide binding by OppA protein. Science 264, 1578–1581.
Tanable, M., Mirza, O., Byrne, B., et al, 2007. Structures of OppA and PstS from Yersinia pestis indicate variability of interactions with transmembrane domains. Acta Crystallogr. D Biol. Crystallogra. 63, 1185–1193.
Tosato, V., Gjuracic, K., Bruschi, C.V., et al, 2003. The DNA secondary structure of the Bacillus subtilis genome. FEMS Microbiol Lett. 218, 23–30.
Van Sadelhoff, J.H.J., Wiertsema, S.P., Garssen, J., Hogenkamp, A., 2020. Free amino acids in human milk: a potential role for glutamine and glutamate in the protection against neonatal allergies and infections. Front Immunol. 11, 1007.
Van Veen, H.W., Margolles, A., Putman, M., Sakamoto, M.K., Kornings, W.N., 1999. Multidrug resistance in lactic acid bacteria: molecular mechanisms and clinical relevance. Antonie Van Leeuwenhoek 76, 347–352.
Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interaction web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35, W407–W410.
Wu, G., Robertson, D.H., Brooks, C.L.I.I.I., Vieth, M., 2003. Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm. J. Comput. Chem. 24, 1549–1562.
Yoon, H.J., Kim, H.J., Mikami, B., Yu, Y.G., Lee, H.H., 2016. Crystal structure of a putative oligopeptide-binding periplasmic protein from a hyperthermophile. Extremophiles 20, 723–731.
Zhang, Z., Adelman, A.S., Rai, D., Boettcher, J., Lönnerdal, B., 2013. Amino acid profiles in term and preterm human milk through lactation: a systematic review. Nutrients 5, 4800–4821.
Zhang, S., Zeng, X., Ren, M., Mao, X., Qiao, S., 2017. Novel metabolic and physiological functions of branched chain amino acids: a review. J. Anim. Sci. Biotechnol. 8, 10.