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

Han-Ha Chai (hanha@korea.kr)
National Institute of Animal Science https://orcid.org/0000-0001-7752-3967
Young Ran Kim
College Of Pharmacy, Chonnam National University
Jun-Sang Ham
National Institute of Animal Science
Tae-Hun Kim
National Institute of Animal Science
Dajeong Lim
National Institute of Animal Science

Research

Keywords: ATP-binding cassette transporter, Bifidobacterium longum KACC91563, Oligopeptide-binding protein A, Structure-functional relationship, Structure-based molecular modeling

Posted Date: November 24th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-113331/v1

Abstract

Background: The OppA receptor as a ATP-binding cassette (ABC) transporter plays key roles in protecting host organism and transport nutrients across the intestine by the oligopeptide transporter from symbiotic bacteria 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. The extracellular OppA receptor from gram-positive *Bifidobacterium longum* KACC91563 translocate nutrients-peptides from the outside environment of intestinal tract to the inside of the symbiotic cell, as a peptide importer. Hence, it was attempting to explicate the relationship between the substrate's specificity from the OppA importer and the probiotic effects of *B. logum* KACC91563 in the host intestine. The probiotic effects of *B. logum* KACC91563 were attributed to the enhancement of the epithelial barrier by several different strain sepcific ways to prevent the strong adhesion of pathogens. The specialized structure-function relationship from the OppA importer could provide the abstract of substrate specificity into unique immunological properties of that the host organism.

Results: In the study, we characterized the extracellular OppA importer from *B. longum* KACC91563 of intestinal microbiome by various protein structure-based modelings in silico. Structural characterization by conserved 5 patches and 4 functional motifs from specific trace residues of the OppA importer. The hydrate surface of the binding site had been decipted by specific trace residues of the OppA that trace residues of Thr58, Lys185, Trp443, and Tyr447, which were characterized in highly exposed hydrophobic binding pocket by its aggregation prones. Therefore, the spatial aggregation propensity in the binding site of the extracellular OppA importer plays a vital role in the specific interaction determinant for peptide binding. In addition, alanine mutation energy values allowed for the virtual determination of the relationships between the energy effects of the peptide binding site mutation on the transporter structural stability, the peptide binding affinity, and the transporter-related peptide substrate selectivity from OppA importer. In particular, distinctive seven pharmacophoric features comprised of two H-bonding donor(**P1**), three H-bonding acceptor(**P8**), and two hydrophobic points (**P5** and **P8**) matched the the OppA receptor-peptide ligand interactions within their binding pocket structure. There are distinct interactions to fix the positions of the N(**P1**) and C(**P8**) termini of the complex of OppA-peptide from *B. longum* KACC91563 such as side chain-specific interactions with the OppA, compared to that of the *Lactococcus latis* (*L. lactics*) OppA.

Conclusions: The specialized structure-function relationship from the OppA import could provide the abstract of substrate specificity into unique immunological properties of the host organism by stucutre-based molecular modeling. In the current study, we attempted explication of 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. Moreover, functional characterization of solute-binding proteins (such as 15 cell wall proteins and 20 extracellular proteins) on the B. *longum* KACC91563 genome will lead to insight of key switch for substate's metabolism into reprogramming immune responses in the host intestine.

Background

Oligopeptide-binding protein A (OppA) from prokaryote acts as a peptide transporter into the cell that is a nutrient import receptor. The protein sequence identifies between the prokaryotic OppAs that are almost below 20%, but they share 3D-overall scaffolds [1, 2]. The protein structures of the prokaryotic OppAs osculate between two α/β domains (i.e., domain **I** and **II**) and a hinge region. Among their 3D-structural folds, the hinge region interconnected between the two α/β domains has dissimilar structural skeletons and arrangements between the prokaryotic OppAs. The peptide-binding pocket of the OppA receptor takes cavity up between the rigid α/β domains. With the mobility of its hinge-bending joined into both, the relative orientations and surrounding places from the α/β domains are reflected inside the voluminous cavity of OppAs. If a single hinge-link gets more bent, then the hinge bending motion is 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 more internal interactions between the α/β domains governed by OppA itself on the nonbinding of the peptide ligand. With the liganded receptor, this also promotes intermolecular interactions between the OppA receptor and its peptide ligand, which may help pack each other at their contactable interfaces. This was actually observed in crystallized prokaryotic OppA/peptide complex structures where the α/β domains were tightly put together with its peptide-ligand having suitable lengths buried at the binding pocket (e.g., 3FTO was compared in an open-unliganded conformation [3] to 3DRG in a closed-liganded conformer of OppA's crystal structures [4] from *Lactococcus latis*). The structural topology of the hinge-bending region allows the binding pocket to take open and closed OppA conformers in 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 binding kinetics of the "Venus flytrap" model [5]. 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 to note 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. Unfortunately, all of the structural states did not exist in their determined structures of the prokaryotic OppAs. The OppA receptor from *Lactococcus latis* (*L.latis*), 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 Figure S14 including the PDB ID of 3FTO, 3DRH, and 3DRG, respectively [4, 3].

In addition to the four structural states of the OppA receptor, even the prokaryotic OppAs with homogeneous structure alignments did not have all the same skeletons of the hinge region [2] and did not only bind to limited types of peptide-ligands. Therefore, the OppA receptors have different widths of cavities that can be accepting to varying ligands with broad specificity. Even if *Escherichia coil* OppA has a whole structural framework, it can transport different solutes (such as nickel [6] and dipeptide [7]) into the cell. From a structural point of view, the *E.coil* OppA contains both two connecting hinge-strands and an extra domain between its α/β domains (i.e., the PDB ID of 3TCG [22]). The extra domain of *E.coil* OppA maintains its ability to open the cleft spaces up between the α/β domains by the bulky volume raised

from the domain itself. Prokaryotic OppA receptors usually have a unique extra domain (e.g., domain **III** with regular shapes), the substructure of which may also be an important structural factor. This is because the extra domain can have fewer ligand selectivity for the prokaryotic OppA receptor than the other solute transport receptors (such as the ion channel) [2]. Therefore, some prokaryotic OppAs that have the same architecture as an extra domain and hinge region did not only accept peptides with similar lengths. For instance, OppA from *Salmonella typhimurium* accepts peptides with two to five amino acid residues [8], whereas *L.lactis* OppA can capture a wide range of peptides up to 35 residues long with preferences for a range of 9–11 residues [9, 10, 4]. As a result, the structural architectures and physical properties of an extra domain can also be associated with their transporter's functions by determining the peptide lengths and solute-ligand's types carried from the prokaryotic OppAs, as well as the hinge-bending motions. This extra domain is the structural and functional determinant from the prokaryotic OppAs.

In the open conformer of the OppA receptor, the two α/β domains move away from each other on the hydrated state where the peptide-binding pocket is well exposed to water. When any hydrophobic peptide binds, the ligand would push out water molecules from the hydrated cavity to keep on doing hydrophobic interactions with the counterpart residues that protruded into the OppA binding cleft. At the same time, binding the polar residues of the peptide leads to the 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 [11]. In particular, a broad specificity against its peptide-ligand had been highlighted in the *L.lactis* OppA receptor, which has a huge binding cavity (almost 4900$Å^3$ [4, 29]). Some residues of the peptide bound should indirectly interact with their counterpart residues on the outer contact surfaces of the hydrated pocket, which is in addition to also getting hands-on 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) seem to be more selective than the other side chains (i.e., positions 7, 8, 9 on the ligand) located on the outer contacts [10]. 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 instantaneous contacts with the OppA receptor, and highly ordered water molecules. Whether the open or closed complex conformer appears is dependent on its peptide-binding affinity to the OppA receptor [4, 3, 12] in conjunction with producing a structural hinge region and an extra domain. Then more energy compensation through binding both is required 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 much more from pushing out well-arranged water molecules from the hydrated pocket of the OppA receptor. Because of this, it should be an essential interaction determinant to characterize what are driving forces for satisfying the binding affinity from the prokaryotic OppA/peptide complex structures [10, 13].

In the current study, compressed information from prokaryotic OppAs was glued toward the OppA of *Bifidobacterium longum* KACC91563 and compared to *L.lactis* OppA. This is because the *B. longum* KACC91563, a subspecies of the *Bifidobacterium* genus, belongs to the *lactic acid bacteria* (LAB) family [14, 15]. Currently, the *Bifidobacterium* genus consists of 56 species and 9 subspecies [81] isolated from human and other mammalian intestinal contents. This includes the 311 bifidobacterial genome sequences that could be retrieved from the National Center for Biotechnology information (NCBI). *Bifidobacteria* are gram-positive bacteria that possess one lipid biayer memebrane enveloped by a cell well. Notably, the *Bifidobacterium longum* KACC91563 is a probiotic genus isolated from the feces of neonates (BioSample code SAMN02603656). It is also known to secrete 15 cell wall proteins (0.77%) and 20 extracellular proteins (1.02%) for a total of 1,856 proteins that have a genome size of 2.40 Mb [16] with 87.77% of the *Bifidobacterium longum* symmetrical identity constituting the bifidobacterial secretome database from the BioProject PRJNA66401. The best characterized member of the *Bifidobacteria* is *Bacillus subtills*, whose geneome of 4.21 Mb encodes the putative 77 ATP-binding transport proteins in its 4,100 protein-coding genes [17]. The most common extracellular protein of the LAB family has been identified in the OppA receptor belonging to the solute binding proteins of the ATP-binding cassette importers (as represented in Figure S10). *Bifidobacterium* directly degrades exogenous proteins (such as milk proteins) to peptides by its protease, and then the OppA receptor takes the peptide-ligand [14, 18] up as a nutrient transporter, which is similar to the other LAB members. Based on the structural colorations from both hinge-fragments and an extra domain, we had attempted to find conserved structural features of LAB OppAs from their determined X-ray structures before centering on sifting the OppA templates of *B. longum* KACC91563 among them. Therefore, we preformed 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 characteristicss distinctive from the LAB member's OppAs from previous experimental studies [14, 18, 82]. At the prediction stage of the OppA structure from *B. longum* KACC91563, we dicided not to couple equilibriums between the opened and closed complex conformers related to the ligand's binding affinity from the OppA templates of the LAB members. Practically, *L.lactis* OppAs were observed in structural differences 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 upon peptide-ligand bindings as RMSD (root-mean-squared) of 2.08 Å per 556 residues. Even so, 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 [19] rigid-body mode. This result indicates that there is only little structural difference between the open-unliganded and open-liganded conformers of the *L.lactis* OppA. This is because the open structures easily kept their open states from not only rotation hindrances consisting of their two hinge-strand fragments and an extra domain, but also bulky volume of the extra domain itself. Surprisingly, the backbones of the peptide-ligands apart from their binding affinities, were in fixed conformers (RMSD of 0.4 Å) while the forming H-bonding remained with counterpart residues (such as Ser472 and Ser474 residues of *L.lactis* OppA [4] represented in Figure S14) between the open and closed-liganded complexes. Aside from the addition of the different peptide

lengths, all backbones of the peptide-ligands were bound in the same structural patterns (shown in Figure S6).

The most notable structural feature of the *L.lactis* OppA is where the mobile loops on the outer surfaces of one α/β domain (residues 301–542 in domain **II**) are seen to pull up and down between the open and closed-liganded conformers (illustrated in Figure S1). Whereas in the other α/β domain (residues 32–82, 220–300, and 543–570 in domain **I**) and the extra domain (residues 83–219 in domain **III**), neither their flexible loops on the contact surfaces nor adjacent interfaces into the ligand binding sites are seen to significantly differ on the liganded state. As expected, the majority of the two hinge-strands were almost maintained between the complex conformers; the torsion angles of the center residues (Ala299 and Met542 [4]) in the two hinge strands changed little inside the 15° degree centered between the unliganded and liganded states. The functional determinants of the *L.lactis* OppA as a solute-binding protein (SBP) was therefore reflected into the large movements of the second α/β domain (in particular to residues 301–542 in domain **II**) that the OppA receptor closed the peptide-ligand up with high binding affinity, instead of incomplete sequence discernments of the peptide itself. As a result, the conformational changes in the peptide-binding sites (between two α/β domains) were less progressive than the mobility caused in the domain **II** of the *L.lactis* OppA when it changed from open-liganded to closed-liganded forms. There are still unfulfilled spaces by the peptide-ligand in the hydrated binding pocket. As if waters were filled instead of the ligand, the *L.lactis* OppA receptor covers over the bradykinin like peptide (i.e., RPPGFSPFA is only occupied by 30% of the available volumes 4900$Å^3$), which is known to bind with high affinity (dissociation constant K_D of 0.1 μM [4]) on the closed complex structure (PDB code of 3DRG). Because of this, the water molecules contained in the hydrated pocket from template OppAs should additionally be considered during the homology modeling of the query OppA from *B. longum* KACC91563.

To consider the ligand binding effects from water-mediated H-bonding interactions, water molecules were divided into two groups, in which the first group was filled into fixed water molecules without moving their coordinates in the *L.lactis* OppA binding pocket from any crystal structure. Then, the second water group collected moving waters by combining with the interaction residues to fit into the ligand-binding cavity. In the process of homology modeling for the *B. longum* KACC91563 OppA, the previous water group would have their initial positions consistent with those located at the place of the hydrated binding pocket through 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 mobile 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 the 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 induce the waters to stay away from the cavity. Indeed, the second water group has potential possibility of

contributing to the determinant factors against the specificity rather than the binding affinity of the peptide-ligand. The buried water effects were known for only the periplasmic OppA in *Salmonella typhimurium* [11, 20], which did not act as a physical barrier to its ligand binding, but seem to affect ligand selectivity in particular, when comparing relative binding affinities between its tripeptide-ligands. 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 to replenish lost H-bonding by arbitrating disrupted interactions within the binding pocket. Both roles of the buried water molecules will lead to the capitalization of a peptide-ligand with a broad binding affinity against the OppA receptor, which is unlike highly specific ligand recognition within other ion channel transporters.

Methods

All molecular modeling and optimization were performed in Discovery Studio, version 2017 R2, from BIOVIA (San Diego, USA) [21].

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 includ OppAs from *Burkholderia Pseudomallei* [22], *Escherichia coli* [23, 24], *Salmonella typhimurium* [11, 25, 26], and *Yersinia pestis* [27]). Otherwise, the other 11 structures of extracellular peptide binding proteins originated from gram-positive bacteria (such as AppA from *Basillus subtills* [28] and OppA from *Lactococcus latis* [4, 3, 29]). Among the known structures of gram-positive bacterial OppAs, biological relevance as template OppAs was determined by 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 [30]. The additional factor as peptide exporters from *L. lactis, B. subtills*, and *B. longum* KACC91563 is accepting proline-rich peptides [18, 4, 29]. However, the frequencies of the proline are not regular in peptide-ligands bound with both OppAs. Proline-rich peptides that include at least one hydrophobic residue will be bound with high affinity to both the template and the query OppAs from priority selection of peptide composition rather than their exact sequence [14, 4]. ii) The homologous to hydrophobicity and cavity size of their binding pockets have closely been related to the characteristic of trace residues derived from these conserved regions on the basis of structural similarities. For instance, templates (like *L. lactis* OppA) and the query of OppA from *B. longum* KACC91563 have been classified into cluster C of SBPs by R.P. –A. Berntsson et al. [2] based on crystalized structural alignments in PDB. The SBPs of cluster C were furthermore contemplated for their structural and biochemical features between the extracellular OppAs. This is because the applied template SBPs could bind with different solute-ligand types (for example nickel ion [6] or cellobiose [31], which is unlikely to be an oligopeptide within known ligands from the SBP of cluster C) to be dissimilarly functionally characterized. In contrast, all SBPs of cluster C have a unique extra domain, which has differences in its structural skeletons and sizes. Despite these quite little sequence identities (smaller than 20%), their 3D-structual architectures are highly conserved; In the study, the sequence identifies 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 [1, 2] via structural homology searches and sequence profiles of iterative protein PSI-BLAST within the UniProt knowledgebase of Swiss-Prot [32]. Among the truncated candidates, the DppA of *E. coli* [24] and two OppAs of S. *typhimurium* and of *Thermotoga maritima* [33] 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 being 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 though 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 [11] of PDB code of 1B9J) were categorized in cluster C of SBPs based on only their 3D-structural scaffolds (often, ≤ 25% as consequence of their protein sequence alignments to the query OppA). In contrast, the top 5 sorted template's 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 picked out, instead of exiguous sequence identities. iii) The query OppA from *B. longum* KACC 91563 is also an extracellular SBP in cluster C and a LAB member such as *L. lactis* and *B. substilis*.

The three considerations mentioned above lead to sorting out templates from the solved gram-positive bacterial OppA structures as being 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, in which the four templates were overlaid with characteristic determinants (as highlighted in Figure S2). There were high scoring matches of extracellular SBPs between the OppA from *Bacillus anthracis* str. *Ames* and AppA form *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 that exhibit 27% sequence identity and 2.60 Å of RMSD with 467 equivalent positions. This depends on whether the ligand is binding all of the whole structural differences between them, which is over 90% of mapping coverage in 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 known in detail (not published). AppA from *B. substilis* (PDB ID: 1XOC) [28, 34], OppA from *B. anthracis* str. *Ames* (PDB ID: 5U4O), and OppA from *L. lactis* (PDB ID: 3FTO, 3DRG) [3, 4] 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 plays 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 like that of the *L. lactis* OppA, too. Intriguingly, *L. lactis* OppA was also derived to be homologous to other SBPs that convey sex pheromones in the type of ABC transporters [1, 28, 35]. The homologous PreZ receptor not

only has the 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 [36]. The favored characteristics of hydrophobic pheromone ligands had 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 [36, 37, 38]. 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 by their protein sequence alignments.

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 Figure 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. In different loop segments, exposed hydrophobic loops to water molecules should affect the structural stability of the OppA receptor itself due to their aggregation prones. 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. There is because any homology restraints of templates cannot be applied in the loop segments of the query OppA models. These parts of loop segments 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 preferences of the different residue types and for the different side-chain rotamers in the regions (containing 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). Thence, it was 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 probability density function (PDF) [39] and discrete optimized protein energy (DOPE) [40] along the pre-structural alignment positions of the

templates (as depictured in Figure 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 were treated as rigid bodies in the OppA models without steric hindrances of the OppA model selves. The solvated OppA model did more optimiziation by conjugate gradient and simulated annealing optimization procedures using the CHARMm force field. As a result of this, the conserved structural characteristics are well reflected into spatial geometrical restraints beside 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, the 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 satisfies the homology-derived restraints better. 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. Then the final OppA model was optimized in 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 [41]. The total PDF energy of the best OppA model was determined to be -14,251.4 kJ/mol in terms of geometric restraints, which are 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 of -50,879.8 kJ/mol as a conformational energy that measures the relative stability of a conformation with respective to other conformations of the loops for optimizing local structures of the query OppA.

To ensure precise of its 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 [42–43]. 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 outliner region on the open-unliganded scaffolds. The OppA model was thus considered to be a good model structure for which there should be more than 90% of the residues in the favored region of the protein (shown in Figure S5).

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

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 [44], Z-score of ProSA [45], and Verify3D [46]) were then employed to measure the 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). The 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. The larger ERRAT2 value then points out the 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 track 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), unlike ERRAT2. Specially, the Z-score better captured the solvent exposed residues within the soluble globular proteins than protein structures containing transmembrane domains. As manifested in Figure S5, the folding features of the hydrated OppA model are well reflected in the Z-score, where there were get-togethers 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 appeared in only small differences of relative middle spots to induce the OppA conformation depending on open or closed states. The 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). On the open-unliganded structural state, the query of OppA model had lower ERRAT2 value and Z-score of ProSA than those of the OppA and AppA templates (OppAs from *B. anthracis* str. *Ames* and *L. lactis* as PDB code of 5U4O and 3FTO), as shown in Table S2.

In the OppA from *B. longum* KACC91563, a large part of the surface from the binding pocket (listed in Table 1.) was exposed to water rather than the hydrophobic environment like membrane interior. Then the effect of the shielding environment caused from the hydrated binding pocket can be indicated by the hydrophobicity plot (depicted in Figure S7), which shows the average hydrophobicity value of each residue with the neighboring four residues [47] instead of its own hydrophobicity. The fluctuations in the hydrophobicity index could be induced by the shielding effects excited from the water. Then the thickset wave was observed in the query OppA of *B. longum* KACC91563 than the template OppA of *L. lactis* (PDB ID: 3FTO). This indirectly explained the hydrophobic binding pocket of query OppA that was exposed to the surface too often, albeit it was not radical shielding effects drawn in water. Due to this, in the absence of ligand binding, an open structural state of the OppA model was more unstabile 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 0.72 of its neighboring 5-residues ran an average hydrophobicity. In contrast, Ala60 changed from 1.80 of the residue hydrophobicity to -0.64 of the neighboring 5-residues' running average

hydrophobicity. Indeed, the helix's secondary structure of the fluctuated region on the query OppA (predicted from such as DSC [48] was based on the solvent exposed patterns of globular proteins) changed into 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 gap of values in both hydrophobicity indexes of the residues becomes smaller in the closed-liganded state (in the case of Thr58 shown in Figure S4). The residues with leading deviations between their hydrophobic properties and the hydrated environment were temporally recognized in the binding pocket depending on whether or not 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 were 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 have its residue characteristics change as a result of the hydrated pocket, which is restricted by the structural compatibility of the query OppA model [49]. In the 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 (such as ERRAT2, Z-score of ProSA, and Verify3D) than those of the X-ray determined template OppA and AppA. However, the OppA query model must be a reliable model structure that has its passed standard qualities observed by other experimental protein structures (shown in Table S2 and Figure S5).

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	~ 2600Å^3 [36, 4]	~ 4900Å^3 [4]	~ 1600 Å^3 [36]	~ 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

	AppA from *B. substilis* (template)	OppA from *L. lactis* (template)	PreZ from *E. faecalis* (template)	OppA from *B. longum* KACC91563 (query)
The hydrophobic pocket for side chain 5 of nonapeptide	Thr41, Asp42, Asp43, Thr46, Arg119, Tyr268, Trp398, Leu401, Met405, Pro407, Trp419, Tyr442, Arg508 [28]	Asn55, Asp56, Ala57, Phe450, Trp453, Val454, Trp473, Tyr491, Phe493 [4]	Corresponding structurally to the hydrophobic pocket of PreZ for side chain 2 of cCF10: Val79, Asp80, Gln81, Met454, Met489, His491 [36]	Thr58, Glu59, Ala60, Gly61, Gly63, Ser136, Ser297, Ser423, Asp424, Glu425, Phe426, Leu427, Val430, Asp431, Gly442, Trp443, Gly444, Gly468, Asn469, Ser470, Gly536

In another aspect, the 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 were given low Verify3D scores. If the surface patches of the OppA model show low Verify3D scores, it may indicate that the patch is interacting with other proteins (such as the 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.

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

We applied the flexible docking [50], which allows for the flexibility of the open-unliganded OppA receptor from *B. longum* KACC91563 during the 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 the 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. The 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 Figure S1. For the varied residues of the OppA receptor, 2,734 of the 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 cystein 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 was coordinated from its X-ray structure (PDB code 3DRG). Then, 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 placed into the coordinates 34.57, 2.81, 12.28, and 14.11. The ensemble dockings of the 9-mer RPPGFSPFA peptide poses in the site sphere on 2,734 of OppA receptor conformational states were performed. After that, each docking pose was subjected to the simulated annealing molecular dynamics (heating to 700K for over 4,000 steps followed by cooling to 300K for over 6,000 steps) process under the CHARMm force field [51] before the complex poses were scored. For each final complex pose, the CDOCKER interaction energy as the CHARM energy (i.e. the interaction energy plus ligand strain) was 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 with 2.0 Å of each other. Then, the best docked pose of the closed-liganded OppA receptor from *B. longum* KACC91563 showed a top scoring complex pose (with the most negative thus favorable to binding) with a CDOCKER binding energy of -337.98 kcal/mol, as shown in Fig. 1.

Pharmacophores 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 tophological feature descriptions of the peptide ligand as well as their corresponding 3D location and direction constraints that are responsible for the peptide's specificity determinants that the OppA receptor from *B. longum* KACC91563 undergoes. The pharmalogical 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 centeroid 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 volumes. 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.

Residue alanine mutations and their impact on the 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 the 62 key residues (listed in Table 1) in the peptide binding site as 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 bining 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 were cacluated using the CHARMm force field and a generalized Born implicit solvent model. Van der Waals terms and electrostatic interactions as well as non-polar, surface-dependent solvation energy terms (i.e., the cavitation energy) were determined empirically. 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 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:

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

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

The mutation energy for the binding affinity was 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 added as the peptide binding site properties to the different complex variants. This allowed for the 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 Figure 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 the 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 relactive importance of hydrophobic interactions on the exposed hydrophobic surfaces of the pepide binding site across the 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 [52], which 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. Thus, the amino acids that were more hydrophobic than glycine 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:

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

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 Figure S4). Then an SAP map for the region was generated by red color-coding (in Figure 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 Figure 4). This can be expected, as most of the OppA receptor sufaces exposed to water are usually hydrophilic. The SAP valuse changes of the OppA receptor between the open-unliganded and the closed-liganded conformers might provide imformation 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.

Results

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

The OppA gene (Genbank code of AEI97628.1 [53]) 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 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 made up in 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. The OppA model is 59 kDa in theoretical protein size calculated from its sequence and is

acceptable to the other 5 clusters of extracellular SBPs in their size ranging from 55 to 70 kDa (with 493 to 543 total residues) [2, 13]. From the perspective of their 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. substilis* (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 of each other between three domains (two α/β domains and an extra domain) were closely aligned to the conformal mapping of the *L. lactis* OppA (PDB ID: 3FTO, 1.18 Å of RMSD with 490 equivalent positions) than the other template of OppA from *B. anthracis* str. *Ames* (PDB ID: 5U4O, 3.06 Å of RMSD with 468 equivalent positions); the similarities were overlaid by matching molecular filed (under condition of 50% steric and 50% electrostatic filed) between 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 closely mapped to the 3D spaces of the *L. lactis* OppA rather than OppA 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 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 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 as such in Figs. 5 and 6). Then exposed trace residues that were more likely to be responsible for binding activity could be distinguished from the buried trace residues that were more important for maintaining structural integrity on the hydrated binding sites of both the templates and the query OppA. Clustering functional trace residues were indeed applied to define the distinct patterns within the query of the OppA structure from those of the other template's structures. As shown in Fig. 3, the conserved patterns were mapped to the query OppA from *B. longum* KACC 91563 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 the query OppA OppA in Fig. 3 and Table S1) into five patches. The conserved distances of 21.0% should be depicted as discriminating molecular basis based on either the electrostatic or hydrophobic properties of subgrouping 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 Figure S13) were identified in these five 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 (in where x is any residue). These conserved patterns did not interlace from their common signature motif, as (LIVM)Ax$_2$(WI)x$_{1or2}$ (SN)(KE)Dx$_4$T(FY)x(LIV)Rx$_3$K were grouped into class 5 periplasmic and extracellular proteins (i.e., peptide and nickel-binding proteins) identified by *Tam et al* [1]. 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 of OppA. In view of their 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 they were preserved in their 3D-structural topologies and biological properties via the four pattern motifs. 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, Ala190 in domain **III** of its extra domain located in the interaction interfaces (highlighted in yellow in Fig. 3). Not the specific trace residues 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 for studying the OppA protein structure-functional relationship or as a target for structure-based pharmacophores by analyzing the interactions of a bound peptide-ligand in the OppA receptor from *B. longum* KACC91563.

The characterization of hydrophobic binding pocket by its aggregation prone

The OppA of *B. longum* KACC91563 is quite limited in the mobility of the two α/β domains (domains **I** and **II**) along with rotating the hinge-strands connecting both domains. In particular, the structural hindrances into the first α/β domain (domain **I**) is closely adjacent to the extra domain (domain **III**) and has a higher steric barrier than the second α/β domain (domain **II**), even in its open-unliganded conformer (as highlighted in Fig. 1). It was consistently observed that the domain **II** can be picked up instead of the domain **I** of the *L. lactis* OppA escaping from the structural hindrances on its open-unliganded crystal structure (PDB code of 3FTO) represented in Figure S1. Less steric hindrances of the domain **II** from the query OppA can easily access 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. Then the large extra domain (domain **III**) 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 4900Å^3 in the templates. This is beacuse they all have an extra large domain (domain **III**) extending the pocket, and also their hinge-region consistists of two β-strands, in which each β-strand is typically observed in 4–5 amino acids as a conserved structural feature (displayed in Figure S2). For these conserved structural traits, their peptide-ligands accommodated in the binding pocket have also 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 [2, 13]) in their closed-liganded states. Likewisely the OppA models from *B. longum* KACC91563 have been observed in the little reduced volume of the hydrophobic binding pocket from 2740 to 2200Å^3 by trapping a bradykinin-like peptide (as 9-mer peptide of RPPGFSPFA) may be possible (expressed in Fig. 2 and Fig. 4).

As another characteristic of the binding pocket, the OppA templates were presented as having discernible functional features by superposition of the specific hydrophobic cavity. The coordinated hydrophobic interfaces from the templates did distinguish their sizes and preferences of the counterpart's hydrophobic

residues from any location of the peptide ligands on the hydrated binding pocket (as shown in Fig. 5 and Fig. 6). In the nonbinding ligand, hydrophobic interactions on the uncovered interfaces of the hydrated pocket can be driven in the tendency to aggregate, which may decrease activity and open the structural stability of the OppA receptor's self. By predicting the OppA receptor surface sites that are liable to aggregation, 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 provide an overall tendency to aggregate in the OppA receptor's binding pocket between the open-unliganded and closed-liganded states (represented in Fig. 4 and Figure S4). Therefore, the hydrated surface of the binding site had been depicted by specific trace residues of the OppA that the trace residues of Thr58, Lys185, Trp443, and Tyr447 (in Figure 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 Figure S4, the 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. Then 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 an open-unligand state, the prone sites to aggregate in domain **II** showed broader areas and a greater number of localized aggregation sites on the OppA surfaces than the other 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 waters to the docked peptide. Upon ligand-binding, the OppA conformer changed to increase 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 had prominently decreased to act as an aggregation inclination changed in the two structural states, as captured in Figure S4; those 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.

Oligopeptide binding pocket presented in specific trace residues

Bradykinin (RPPGFSPFR, pfam code of PF06753 and InterPro code of IPR009608) is well documented as a pharmacologically therapy known as an ACE (angiotensin-converting-enzyme, CD143) inhibitor [54, 55] to reduce high blood pressure. Together, this bioactive peptide is an inflammatory mediator when kinin B1 receptor (BDKRB1) recruits neutrophil via the chemokine CXCL5 (CD185) [56]. Physiologically, the capability to bind bradykinin for the LAB microbiota (such as *L.latics, B.subtilis*, and *B.longum)* did elicit any health effect from a nutrient source, which catches up peptides with antihypertensive activity [57, 58].

Bradykinin is well-matched with the favored peptide-ligands from the templates of *L.latics* OppA and *B.subtilis* AppA. The oligopeptide ABC importers from the LAB microbiota select preferentially proline-rich peptide, containing at least one branched residue of leucine, isoleucine, and valine [14, 28, 4, 59] with no exact sequence preference. In that case, bradykinin was 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 [59]) in closed-ligand conformation. In essence, 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. On the basis of the wailing peptide's specificity, the 9-mer peptide of RPPGFSPFA has a fixed proline-rich peptide isoform, of which the backbone 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 surface is covered inside Glu47, Thr58, Glu59, Ala60, Gly61, Leu427, and Val430 residues (as presented in Table 1 and Fig. 5). The complex structure shows the 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 its structural stability and binding affinity within the OppA-peptide complex from *B.longum* KACC91563. Since the hydrophobic binding site underlies better stability to reduce aggregation, it switches from an open to a closed conformation of the OppA upon the 9-mer RPPGFSPFA binding. The Ala mutation effects tend to cause greater destabilizing structural stabilities of the complex than decreasing 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, the Ala mutant of Ser188, Asp446, and Asn452 did alter the stabilizing (mutation energy < -0.5 kcal/mol) complex structure by smaller mutation energies than the −1.2 kcal/mol. Intriguingly, the Ala mutation effects were well projected into the specific Thr58, Lys185, Trp443, and Tyr447 trace residues with being decreasing aggregative propensities (shown in Figure S4), wherein the intrinsic moving domain **II** of the OppA is the most defining feature of the switching conformation from a part (open conformation) to the two 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 the unsettling of its complex stability in the presence of 2- to 60-fold mutation energy effects more 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 on the 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. Taken together, the highly exposed residues of Ser186, Tyr191, and Met192 on the extra domain **III**, Asn294, and Pro445 on domain **II**, of which the counter partners are Arg(**P1**), Pro(**P3**), and Pro(**P7**) on the RPPGFSPFA peptide, are prompted by a descending aggregation factor in Figure S4. Among them, the biggest destabilizing Ala mutant effect of polar Tyr191 was envisaged as 5.87 kcal/mol by comparing that of hydrophobic Met191 as 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**) as -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 came from the favoritism of proline-rich (**P2**, **P3**, and **P7**) and hydrophobic Phe(**P5**), which also have impacted the OppA-RPPGFSPFA interactions for *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 are 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 prolin-rich and hydrophobic phenylalanine (**P5**) in turn lead to impute featuring conductors of the key trace residues in the interaction interfaces. Two H-bondings 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 it did not form with the OppA of *L. lactis* (as shown in Fig. 5.). The formation of two H-bondings 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(P1) 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. Amazingly, 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 the 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 serving as the complex's structural stabilization via decreassed aggregation poses. 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 grappled with their location and direction of two H-bondings 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*. The 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. These hydrophobic interactions may be the grassroot 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 was enforced rather than the projected hydrophobic point feature of the targeted Phe(**P5**) for making the interaction sites. In contrast to the 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 achieved by a hydrogen donor from the side chain of Arg440. Then the hydrophobic shape constant of Phe(**P8**) expanded to surrounding the Phe(**P8**) by Leu296, Phe334, Ala396, Trp399, and Trp515 residues on the OppA from *B. longum* KACC91563. The role of Arg440 was well superseded by the Arg416 residue on *L.latics* OppA. As a result, the configuration of pharmacophores on the complex structure-based 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. Then the shape constant 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 is physiologically broad peptide selectivity from the OppA transporter itself of the LAB microbiota, which is reflected by its selectivity score of the identified pharmacophores on the OppA from *B. longum* KACC91563. In Fig. 5, the *L.latics* OppA had exclusive H-bonds with the peptide backbone of RPPGFSPFA via the 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 the residues Ser472 and Ser474 of *L.latics* OppA did consistently interacts with other peptide (e.g., RDMPIQAF) backbones, irrespective of its binding affinity in even opened-ligand conformations [4], as shown in Figure S14. In contrast, the H-bond patterns of *L.latics* OppA compared to the H-bond register on the OppA from *B. longum* KACC91563 was 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 the 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 the Arg400 and the Phe(**P8**) to coincide with the shape constant of the targeted interaction sites. This would be caused by those different sizes of the peptide-binding cavity (4900Å^3 vs 2700Å^3) to explain the lower size limitation and the spacious specificity of the peptide-ligand in the *L.latics* 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 Figure 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.latics* OppA were known to be 1.2 μM and 4.2 μM for SLSQ**L**SSQS and SLSQS**L**SQS in the closed-ligand conformers, respectively [53]. Before screening the two peptides, we had 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). Then, validating the pharmacophore model as required features allows us to guide the knowledge that these pharmacophores are most 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-bonding 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 Figure 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.latics* 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 the bound bradykinin structure to *L.latics* OppA (see Figure S6 with RMSD 0.29–0.4 Å for their alpha carbons). The other five features comprised of two HBDs(**P1**), one HBA(**P8**), and two HY points (**P5** and **P8**) were not at all fitted in both peptides. These five features must be adapted to the specific binding determinants of the peptide-ligand to the OppA from *B. longum* KACC91563. The *L.latics* OppA was accepting to the five features on its peptide binding sites. This did shed light on how *L.latics* OppA accomplish broad specificity with its peptide variant length (4–35 residues) but not a striking trace coloration 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 [28] also favors hydrophobic nanopeptide substances with overlapping specificity to its OppA. Figure 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, Trp419 encloses Asn(**P5**) more through positioning two H-bonds between their backbones of the Asp43, Gly418, and Asn(**P5**). On the other hand, the Asp(**P2**) does not match the hydrophobic point constant but Trp(**P9**) does replace half occupancy in the conjugated hydrophobic shape constant on the side chain of the VDSKNTSSW peptide by the above pharmacophore fitness (shown in Figure S9). No proline residues on the VDSKNTSSW peptide were likely to maintain the backbone conformer like the RPPGFSPFA binding within the contact interfaces of AppA

through the H-bond network with both main chains. There seems to be no necessary requirement for the AppA to be proline-rich and the branched hydrophobic(**P5**) on the nonapeptide-ligand to have binding specificity. The AppA could bind to the bradykinin of RPPGFSPFR (the dissociation constants K_D binding to *B. subtilis* AppA is 50.3 μM [59]).

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 10pM to the PreZ [9]) and the extracellular pheromone receptor (in Table S3 and Fig. 6). Specially, 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 widen binding cavity to accommodate the larger peptide-ligands from *E. facalis* PreZ to *L.latics* OppA is associated with a steric clash between the domain **II** and side-chain of the **P5** residue on the bound peptide by desceasing the water mediated H-bonding to bury the peptide in the binding pocket. Therefore, a 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 that shown in Fig. 4).

Discussion

Crosstalk functions of innate immune cells and beneficial microbiota for host defense by their amino acid substrate specificities

T-cell mediated immune responses are required to effectively protect host mammals from virus attacks. To defend the host against invading virus, native T-cells consummate massive clonal expension, differation, and migration to the infected tissues, allowing for the synthesis of cytockine and effector molecules to clear most targeting cells against the viral infection resulting in generating memory T-cells, all of which demand to fulfill more energy substrates by swift metabolic changes of host nutrients. Lymphocytes are profitable in generating energy in the form of ATP by oxidative glycolysis known as the Warburg effect, unlike most differentiated cells by employing the TCA cycle (tricarboxylic acid cycle) and mitochondrial respiration. In the resting state, lymphocytes became catabolic utilizing autophagy that originates materials essential for protein synthesis and energy. Upon sensing specific antigens by the T-cell receptor, the T-cell receptor markedly elevated the ability of T-cells to transport neutral amino acids as material sources into their plasma membrane. For example, leucine together with isoleucine and valine is abundant and accounts for more than 40% of the free ensential amino acids (at a protein-unbound state) in the blood plasma [61]. The activated T-cell pulled the trigger of metabolic switch by activating the glycolytic pathway, glutaminolysis, fatty acid synthesis, and pentose phosphate pathway while attenuating fatty acid oxidation [62–65]. 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 enkinding T-cell receptors and profusing leucines from CD98 (aliases LAT1 of L-leucine transpoter)

through animating the mTORC (mechanistic target of rapamycin complex) signaling pathway [66, 67]. Actually, the competence of mTORC can lead to control T-cell fate (denoted in Table S4), and the mTORC activity calls for the presence of amino acids. As such, the leucine antagonist of N-acetyl-leucine amide blocks mTORC activity and can inhibit T-cell function as an immunosuppressive agent. The dearth of leucine also attracts mTORC hindrance to foment infectious tolerance by producing more regulatory T-cells. If the T-cells have the defective CD98 of leucine transpoter, 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 CD98 null-$CD8^+$ T-cell also has the severe faulty capacity to respond to antigens for the production of cytotoxic effector cells. However, the T-cell receptor activated CD98 null-$CD8^+$ T-cell could upregulate the expression of IL-2 receptor-α (CD25), CD69, and CD44 (Pgp1) and thus discharge normal amounts of IL-2, albeit declined amounts of IFN-γ. Moreover, the activation of T-cell enhances CD28-mediated glutamine uptake as one of the leucine metabolities by switching on ASC2 (L-glutamine transporter) with activating mTORC signaling [68]. The T-cell activation is afforded from intracellular leucine uptake determined not by the extracellular leucine concentration, but by the expressed leucine transpoter 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 Figure S11 and S12(a)). Based on their substrate specificity, leucine or tryptophan requisites are quite 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 are not indespensible to IL-7-induced $CD8^+$ T-cell survival, but are 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 the leucine and tryptophan uptake of the T-cell are a key switch for metabolism reprogramming immune-activated T-cells, leading to an adjustment in adaptive immune responses. In this regard, in an infection resolution, antigen-specific $CD8^+$ T-cells deflate anabolism to the catabolism of nutrients from high to low mTOR activity, and thence from the effector T-cell to memory T-cell.

Similar to leucine, T-cell expansions crave tryptophan, by depleting which T-cell proliferation is blocked [69]. Tryptopan is one of the essential amino acids as a biosynthetic precursor of microbiota and host metabolities. This is because tryptopan is not biosynthesized by human and animal cells that rely on intake from exogenous dietary to catabolize bioactive components (such as kynurenine, serotonin, and metatonin). The tryptopan metabolism (described in Figure S12(b)) are generated in the gastrointestinal tract by (i) a kynurenine pathway in both immune cell and intestinal epithelial cells via two isoforms of indoleamine 2,3-dioxigenase (IDO1, IDO2), (ii) a serotonin production pathway to melatonin in intestinal enterochromaffin cells via tryptophan hydroxylase gene TpH1, and (iii) indole and its derivaties, including aryl hydrocarbon receptor (AhR) ligands by intestinal microorganisms [70]. When induced by proinflammatory cytokines, kynurine serves as an AhR agonist and acts on the dendritic cell to inhibit the

dendric cell mautation. There is more attunation to immune security via major antigen-presenting cells. Peculiarly, kynurenine can form the intestinal barrier and stimulate the proliferation of regulatory T-cell (Treg) to attenuate inflammation by enhancing the IL-10 generation. Then, decreasing the differentiation of highly inflammatory T_H17 cells reduces the production of IL-22. As the characterized neurotransmitter, serotonin (5-hydroxytryptamine, 5-HT) of another tryptophan metabolite plays neurobiologic functions and brain physiology such as mood stability, pain tolerances, or sleep patterns in the mammalian's central nervous system [71]. In spite of this, more than 90% of the body's serotonins are produced by the intestinal enterochromaffin cells of the digestive tract and are also under the influence of the intestinal microbiota by *Streptococcus spp.* and *E. coli*, which can synthesize serotonin themselves from tryptophan. Under the symptoms of serotonin dysbiosis, the peripheral serotonin does not cross the blood-brain barrier, but propagates signals from the intestine to intrinsic neurons. Then host immune cells are coupled with various serotonin receptors (at least fourteen members) on lymphocytes, monocytes, macropharge, T-cells, B-cells, and dendritic cells from the innate to adaptive immune system [72]. In addition, the serotonin-selective reuptake transpoter SERT (encoded by SLC6A4 gene) can remove serotonin from the intestinal epithelial cells to control serotonin bioavailability and to manage reuptake in the brain. Moreover, the intestinal commensal *Lactobascillus* can catabolize tryptophan to indole 3-aldehyde as an AhR ligand to protect against mucosal inflammation by its tryptophanse. Indole is also a signal molecule of bacterial physiology for antibiotic resistance and biofilm formation, whereas in non-indol producing bacteria, indole and its offshoot inhibit quorum sensing and modulate harmfuless factors in intestinal microenvironment. Of note, indol itself further vitalizes enteroendocine L cells to produce glucagon-like peptide-1 (GLP-1) into insulin secretion by pancreatic β-cells.

As an important organ in the host body, the intestine behaves in the shared junction of nutrient digestion and absorption, microbiota colonization, and 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 true characterization of the molecular interactions might provide valuable insights to decipher how mutualistic symbiosis-host's T-cell communication operates in amino acids for mediating intestinal immunity. In pigs, leucine and its metabolitie glutamine can be strength soruces to surpass mucosal barrier functions, the epithirial villus height in the intestine, and small intestinal growth [73–74]. In particular, only leucine achives in triggering an α-defensin secretion from Paneth cells among the 20 mammalian amino acids, thereby conducting surveillance in the intestine to maintain an intestinal homeostatis response to inner environement 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 keep up intestinal integrity. Therefore, both leucine and isoleucine stimulate the intestinal SIgA (secretory immunoglobulin A) secretion that is the most abundant antibody to improve the mucosal surface defense in the intestinal mucosa, thereby leading to inhibiting the pathogen introgression into the lamina propria. It is well known that the glutamate and glutamine as leucine metabolities (as Figure S12(a)) comprise almost 70% protein-unbound amino acids (account for 5 ~ 10% total amino acids) in human milk. In fact, the free glutamate level was 40-fold higher in milk compared to plasma [75]. The effects of the glutamate and glutamine on

immune parameters may diminish susceptibility relevant for allergic disease and infections in developing neonate over the course of lactation. Since at birth, the increased susceptibility in the neonatal immune system is derived from an immature intestinal barrier and incomplete microbial colonization compared to the adult. Fuctionally, the T-cell response to antigen was warped toward T_H2 immunity than innate $CD4^+$ effector cell functions. In turn, the T_H2 dominant immunity milieu augments the susceptibility of allergenic sensitization while a lowered T_H1 function is conducted to increase the neonatal vulnerability to infection. Conversely, the postnatal period supplementation of glutamine had reinforced the production of T_H1 cytokines IL-2 and IFN-γ, thereby promoting T_H1 cells, whereas the descended secretion of T_H2 cytokines IL-4 and IL-5 is accompained by suppressing T_H2 activity. The glutamate released from dendritic cells has dual immunorgulating roles. Upon specific antigen presentation, T-cell proliferation is enlivened by glutamate and the production of IL-2, IFN-γ on increasing T_H1 cells, and IL-10 on producing $T_{reg}1$ cells. Alternately, in non-specific antigen presentation, the T-cell activity is obstructed by glutamate. Neonates used glutamate and glutamine uptaked from the breastmilk by the intestines that have furnished the growth of intestinal epithelial cells and maturation of the intestinal barrier to support protective effects. In infants and weaning piglets, half of the dietary glutamate and glutamine is oxidized by intestinal epithelial cells and immune cells, ultimately leading to energy substrates for glutathione as an antioxidant in the preventive strategy against food allergies [76]. In addition, glutamine, but not gluatamate, can affect mucin synthesis as a precursor the attributable to the increase in the number of mucin-secreting globet cells safeguarding against allergenic sensitization.

As the other 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 bone marrow dendritic cells in the small intestine [77]. 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 virus and pathogenic bacteria. Besides the secretion of 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 [78]. The ecological success of *lactic acid bacteria* relies on the substate 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 [79–80, 78]. The metabolic substrates of the human milk oligosaccharides enhance the *Bifidobacterial* probiotic adhesions and colonisations to the host intestinal mucosa, especially during the host's early life [79–81]. In particular, members of the *Bifidobacterium* are among the key bacterial compontents (i.e., *B. longum*, *B. fidum*, and B. breve species) of the infant intestines. This is because a higher abundance of *Bifidobacteria* was observed in the infant *lactic acid bacteria* than in the adult (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 [82, 83]. The ability of *B. logum* KACC91563 was confirmed in the suppressed allergeic diarrhea from the mouse food allergy model induced by using ovalbumin and alum. Interestingly, *B. logum* KACC91563 had neither curtailing T_H2 cytokine levels nor influencing anti-inflammatory cytokine IL-10 by $Foxp3^+$ Treg cell [84]. As opposed to no effect on the T-cell immune responses, *B. logum* KACC91563 exerts influence on descending only mast cell numbers via increasing the Annexin V^+ apoptotic bone marrow-derived mast cell. This results in dramatically decreasing food allergies in the mouse model. Thereupon, *B. logum* KACC91563-derived extracellular vesicles did much to ameliorate portent food allergy by instigating the mast cell apoptosis [82]. Factually, the extracellular vesicles from *B. logum* KACC91563 are more internalized in mediating unknown receptors on bone marrow-derived mast cells than phargocyte by dendritic cells. Then the unknown receptor to the extracellular vesicles (with the majority being 60 nm) would be different from some exosome transporter CD63 and CD82 that convey vesicles with virus like properties. Therefore, the extracellular vesicles from *B. logum* 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. logum* KACC91563 are composed by an important component of OppA importer as a part of the family's 5 solute binding protein. This is because the OppA importer determined the substrate specificity of family's 5 solute binding protein through physiologic routes in the host intestine. The intestinal *L. lactis* populations and *B. logum* KACC91563 are also closely connected with selecting favored oligopeptides containing at least one branched residue of leucine, isoleucine, and valine via the OppA importer. This is because both organisms are auxotropic for branched amino acids, and they favor proline-rich caseins in milk media (9.8% of proline residues in milk casein [4]) as a nutrient source of these amino acids to fascilitate their growth. To satisfy the demands of the branched amino acids related to the physiology of both organisms, *L. lactis* and *B. logum* KACC91563 proteolytes the 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. Their interplay augments the augmenting immune balance in the nutritional amino acid substrate [85] by adjusting for the compotent assistants-host immune crosstalk. On the other hand, the fitted leucine or tryptophan in the C-terminal residue on the peptide substrate binding to host TAP redounded upon an allosteric crosstalk between the TMD and the NBD, which subsequently enkindles peptide trasmittion and ATP hydrolysis. In addition, the selected branched residues (e.g. leucine) are precursors of acetate in short-chain fatty acids (such as the acetoacetate shown in Figure S12(a)). The acetate products from *B. logum* assist the defense function of host epithelial cells [86], which are unliked to kill pathogenic enterobacteria via α-defensin secretion by butyric acid (one of the short chain fatty acids) and leucine in the Paneth cells [87]. The propionic and butyric acid metabolities were known to exert an antagonistic effect on colon cacinomacell proliferation, thereby inducing the apoptosis of the cancer cells [88]. Hence, it was attempting to explicate the relationship between the substrate's specificity from the OppA importer and the probiotic effects of *B. logum* KACC91563 in the host intestine, as illustrated in Fig. 9. Some factors such as pathogen invasions, proinflammatory cytokine, and toxins contribute to the lowering of the epithelial barrier functions in the host intestines. The probiotic effects of *B. logum* KACC91563 were

attributed to the enhancement of the epithelial barrier by several different strain sepcific ways to prevent the strong adhesion of pathogens.

Conclusions

Bifidobacterium longum KACC91563, a subspecies of *Bifidobacterium* geus belongs to the lactic acid bacteria (LAB), is a probiotic genus isolated from the faces of neonates. The extracellular vescies from B. *longum* KACC91563 are enriched for the solute binding proteins that more effectively transport oligopeptides. Specificlly, the extracellular vesicles of B. *longum* KACC91563 are composed by an important component of Oligopeptide-binding protein A (OppA) importer as a part of the solute binding protein. This is because the OppA importer as determined the substrate specificity of the solute binding protein through physiologic routes. OppA from B. *longum* KACC91563 acts as a peptide transporter into the cell that are a nutrient importer receptor. In the current study, we performed knowledhe-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 the LAB maber's OppAs from previous studies.

The open-unliganded OppA model from B. *longum* KACC91563 was made up in of integral main traces from Gly37 to Gln547, except for its N-terminal anchor frame (residues 1-36) by homology modeling with the template structures (PDB ID: 3FTO, 5U4O in open-unliganded forms). The OppA model is composed its two α/β domains (residues 37-83, 211-293, and 518-547 in domain I and 294-517 in domain II), and an extradomain (residues resdues 84-210 in domain III), wherein it is connected by two by two hinge-strand fragment between domain I and domain II. Bradykinin is well-matched with the favored peptide-ligands from the templates of *L.latics* OppA and *B.subtilis* AppA. The oligopeptide ABC importers from the LAB microbiota select preferentially proline-rich peptide, containing at least one branched residue of leucine, isoleucine, and valine with no exact sequence preference. In that case, bradykinin was known for being able to bind the OppA and AppA templates with high affinity (dissociation constant K_D of 0.10μM and 50.30μM, respectively [59]) in closed-ligand conformation. In essence, 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. The intrinsic moving domain II of the OppA is most defining feature of the switching comformation from a part (open conformation) to the two domains I and II together (closed conformation). Intriguingly this was well projected into the specific Thr58, Lys185, Trp443, and Tyr447 trace residues with bing 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 pocketof the OppA B. *longum* KACC91563. Taken together, the highly exposed residues of Ser186, Tyr191, and Met192 on the extradomain III, Asn 294, and Pro445 on domain II, of which the counter partners are Arg(**P1**), Pro(**P3**), and Pro(**P7**) on the RPPGFSPFA peptide, are prompted by a descending aggregation factor. The

stabilizing effects of induced fitness on the OppA came from the favoritism of proline-rich (**P2**, **P3**, and **P7**) and hydrophobic Phe(**P5**), which also have impacted the OppA-RPPGFSPFA interactions for B. *longum* KACC91563. In addition, the 3D-pharmacophore arrangement to interaction features of the OppA on the complex strucrure-based 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 was enforced as surrounding assembly of the peptide-ligand's pharmacophores with the excluded volume. In the only OppA of B. *longum* KACC91563, the H-bonding register with a side chain of Arg(**P1**) and Phe(**P8**) was specifically defined to a specific determinant for the 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 from 2700Å of the peptide-binding cavity. The intestinal B. *longum* KACC91563 is auxotropic for branched amino acids, and favors proline-rich caseins in milk media (9.8% proline residues in milk casein) as a nutrient source of these amino acid to fascilitate their growth. To satisfy the demands of the branched amino acids related to the physiology of B. *longum* KACC91563 proteolysis the 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. Their interplay arguments the augmenting immune blance in the nutritional amino acid substrate by adjusting for the compotent assistents-host immune crosstalk. The specialized structure-function relationship from the OppA import could provide the abstract of substrate specificity into unique immunological prpperties of the host organism. In the current study, we attempted explication of 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. Moreover, functional characterization of solute-binding proteins (such as 15 cell wall proteins and 20 extracellular proteins) on the B. *longum* KACC91563 genome will lead to insight of key switch for substate's metabolism into reprogramming immune responses in the host intestine.

Declarations

Ethics approval and consent to participate

Not applicable. The current study was permormed in computer aided molecular modeling in silico.

Consent for publication

Not applicable

Avalilability of data and material

All structure based modeling data were generated and analyzed during this study are available.

Please contact author for data requests.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was carried out 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.

Author’s contributions

HHC carried out all the OppAstructure-function based molecular modeling. YRK carried out the analysis of immunological biology for the OppA function. JSH and THK conceived of the study, and participated in its design and coordination and help to draft the manuscript. DL carried out the OppA importer receptor bioinformatics study and participated in the design of the study. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Tam R, Saier MH Jr. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiol Rev. 1993; 57;320-346.
2. Berntsson RP, Smits SH, Schmitt L, Slotboom, DJ, Poolman B. A structural classification of substrate-binding proteins. FEBS Letters. 2010; 584;2606-2617.
3. Berntsson RP, Alia Oktaviani N, Fusetti F, Thunnissen AM, Poolman B.; Slotboom DJ. Selenomethionine incorporation in proteins expressed in Lactococcus lactis. Protein Sci. 2009;18; 1121-1127.
4. Berntsson RP, Doeven MK, Fusetti F, Duurkens RH, Sengupta D, Marrink SJ, Thunnissen AM, Poolman B, Slotboom DJ. The structural basis for peptide selection by the transport receptor OppA. EMBO J. 2009;28;1332-1340.
5. Mao B, Pear MR, McCammon JA, Quiocho FA. Hinge-bending in L-arabinose-binding protein. The “Venus’s-flytrap” model. J Biol Chem. 1982;257;1131-1133.
6. Heddle J, Scott DJ, Unzai S, Park SY, Tame JR. Crystal structures of the liganded and unliganded nickel-binding protein NikA from Escherichia coli. J Biol. Chem. 2003;278;50322-50329.
7. Nickitenko A, Trakhanov S, Quiocho F. 2Å resolution structure of DPPA, a periplasmic dipeptide transport/chemosensory receptor. Biochemistry. 1995;34;16585-16595.

8. Tame JR, Murshudov GN, Dodeson EJ, Neil TK, Dodson GG, Higgins CF, Wilkinson AJ. The structural basis of sequence-independent peptide binding by OppA protein. Science. 1994;264;1578-1581.
9. Detmers FJ, Lanfermeijer FC, Abele R, Jack RW, Tampe R, Konings WN, Poolman B. Combinatorial peptide libraries reveal the ligand-binding mechanism of the oligopeptide receptor OppA of Lactococcus lactis. Proc Natl Acad Sci U.S.A. 2000;97;12487-12492.
10. Lanfermeijer FC, Detmers FJ, Konings WN, Poolman B. On the binding mechanism of the peptide receptor of the oligopeptide transport system Lactococcus lactis. EMBO J. 2000;19;3649-3656.
11. Van Veen HV, Putman M, Margolles A, Sakamoto K, Kornings WN. Multidrug resistance in lactic acid bacteria: molecular mechanisms and clinical relevance. Antonje Van leeuwenhoek. 1999;76; 347-352.
12. Sleigh SH, Seavers PR, Wilkinson AJ, Ladbury JE, Tame JR. Crystallographic and calorimetric analysis of peptide binding to OppA protein. J Mol Biol. 1999;291;393-415.
13. Monnet V. Bacterial oligopeptide-binding proteins. Cell Mol Life Sci. 2003;60;2100-2114.
14. Ha GE, Chang OK, Jo SM, Han GS, Park BY, Ham JS, Jeong SG. Identification of antihypertensive peptide derived from low molecular weight casein hydrolysates generated during fermentation of *Bifidobacterium longum* KACC91563. Korean J Food Anim Resour. 2015;35;738-747.
15. O'Callaghan A, van Sinderen D. Bifidobacteria and their role as members of the human gut microbiota. Frontiers in microbiology**.** 2016;7;925.
16. Ham JS, Lee T, Byun MJ, Kim MK, Han GS, Jeong SG, Oh MH, Kim DH, Kim H. Complete genome sequence of Bifidobacterium longum subsp. longum KACC91563. J Bacteriol*.* 2011;193; 5044.
17. Tosato V, Gjuracic K, Vlahovice K, Pongor S, Danchin A, Bruschi CV. The DNA secondary structure of the Bacillus substilis genome. FEMS Microbiol Lett 2003; 218; 23-30.
18. Chang OK, Seol KH, Jeong SG, Oh MH, Park BY, Perrin C, Ham JS. Casein hydrolysis by *Bifidobacterium longum* KACC91563 and antioxidant activities of peptides derived therefrom. J Dairy Sci. 2013;96;5544-5555.
19. Lanfermeijer FC, Picon A, Konings WN, Poolman B. Kinetics and consequences of binding protein of nona- and dodecapeptides to the oligopeptide binding protein (OppA) of Lactococcus lactis. Biochemistry. 1999;38;14440-14450.
20. Yoon HJ, Kim HJ, Mikami B, Yu YG, Lee HH. Crystal structure of a putative oligopeptide-binding periplasmic protein from a hypertherophile. Extremophiles. 2016;20;723-731.
21. Dassault Systémes BIOVIA, Discovery studio modeling environment, release R2 dassault systems, San Diago 2017.
22. Klepsch MM, Kovermann M, Löw C, Balbach J, Permentier HP, Flsetti F, de Gier JW, Slotboom DJ, Berntsson RP. Escherichia coli peptide binding protein OppA has a preference for positively charged peptides. J Mol Biol. 2011;414;75-85.
23. Davies TG, Hubbard RE, Tame J.R. Relating structure to thermodynamics: the crystal structures and binding affinity of eight OppA-peptide complexes. Protein Sci. 1999;8;1432-1444.

24. Cuneo MJ, Beese LS, Hellinga HW. Structural analysis of semi-specific oligosaccharide recognition by a cellulose-binding protein of *Thermotoga maritima* revelas adaptations for functional diversification of the oligopeptide periplasmic binding protein fold. J Biol Chem. 2009;33217-33223.
25. Sleigh SH, Tame JR, Dodson EJ, Wilkinson AJ. Peptide binding in OppA, the crystal structures of the periplasmic oligopeptide binding protein in the unliganded from and in complex with lysyllysine. Biochemistry 1997;36; 9747-9758.
26. Kozlowicz BK, Shi K, Gu ZY, Ohlendorf DH, Earhart CA, Dunny GM. Molecular basis for control of conjugation by bacterial pheromone and inhibitor peptides. Mol Microbiol. 2006;62;958-969.
27. Tanable M, Mirza O, Bertrand T, Atkins HS, Titball RW, Iwata S, Brown KA, Byrne B. Structure of OppA and PstS from Yersinia pestis indicate variability of interaction with transmembrane domains. Acta Crystallogr D Biol Crystallogr. 2007; D63;1185-1193.
28. Levdikov VM, Blagova EV, Brannigan JA, Wright L, Vagin AA, Wilkinson AJ. The structure of the oligopeptide-binding protein, AppA, form bacillus subtillis in complex with a nonpeptide. J Mol Biol. 2005;345;879-892.
29. Maurer M, de Beer SB, Oostenbrink C. Calculation of relative binding free energy in the water-filled active site of oligopeptide-binding protein A. Molecules**.** 2016;21;499-514.
30. Garai P, Chandra K, Chakravortty D. Bacterial peptide transporters: messengers of nutrition to virulence. Virulence 2017;8;297-309.
31. Felder CB, Graul RC, Lee AY, Merkle HP, Sadee W. The venus flytrap of periplasmic binding proteins: an ancient protein module presents in multiple drug receptors. AAPS PharmSci. 1999;1; E2.
32. Hinz U. UniProt Consortium. From protein sequences to 3D-structures and beyond: the example of the UniProt knowledgebase. Cell Mol Life Sci. 2010;67;1049-1064.
33. Lassaux P, Peri C, Ferrer-navarro M, Gourlay LJ, Gori A, Conchillo-Solé O, Rinchai D, Lertmemongkolchai G, Longhi R, Daura X, Colombo G, Bolognesi M. A structure-based strategy for epitope discovery in Burkholderia pseudomallei OppA antigen. Structure. 2013; 21;167-175.
34. Shin TS, Kim JH, Kim YS, Jeon SG, Zhu Z, Gho YS, Kim YK. Extracellular vesicles are key intercellular mediators in the development of immune dysfunction to allergens in the airways. Allergy. 2010;65;1256-1265.
35. Mackenzie AK, Valegård K, Iqbal A, Caines ME, Kershaw NJ, Jensen SE, Schofield CJ, Andersson I. Crystal structures of an oligopeptide-binding protein from the biosynthetic pathway of the β-lactamase inhibitor clavulanic acid. J Mol Biol. 2010;396;332-344.
36. Berntsoon RP, Schuurman-Wolters GK, Dunny G, Slotboom DJ, Poolman B. Structure and mode of peptide binding of pheromone receptor PrgZ. J Biol Chem. 2012;287;37165-37170.
37. Shi K, Brown CX, Gu ZY, Kozlowicz BK, Dunny GM, Ohlendorf DH, Earhart CA. Structure of peptide sex pheromone receptor PregX and PregX/phenome complex and regulation of conjugation in Enterococcus faecalis. Proc Natl Acard Sci U.S.A. 2005;102;18596-18601.
38. Kunji ER, Fang G, Jeronimus-Stratingh CM, Burins AP, Poolman B, Konings WN. Reconstruction of the proteolytic pathway for use of β-casein by *Lactococcus lactis*. Mol Microbiol. 1998;27;1107-1118.

39. Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci. 2006;15;2507-2524.
40. Dasqupta A, Sureka K, Mitra D, Saha B, Sanyal S, Das AK, Chakrabarti P, Jackson M, Gicquel B, Kundu M, Basu J. An oligopeptide transporter of mycobacterium tuberculosis regulates cytokine release and apoptosis of infected macrophages. PLoS One 2010;17;e12225.
41. Lovell SC, Davis IW, Arendale WBIII, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by C-alpha geometry: phi, psi, and C-beta deviation. Proteins 2003;50;437-450.
42. Dahl DB, Bohannan Z, Mo Q, Vannucci M; Tsai J. Assessing side-chain perturbations of the protein backbone: a knowledge-based classification of residue Ramachandran space. J Mol Biol. 2008; 378;749-58.
43. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 1993;2;1511-1519.
44. Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. Cell 2014;157;121-141.
45. Wiederstein M, Sippl M. ProSA-web: interaction web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35;W407-W410.
46. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 1997;277;396-404.
47. Kyte J, Doolittle RFA. Simple method for displaying the hydrophobicity character of a protein. J. Mol Biol. 1982;157;105-132.
48. King RD. Sagi M, Sayle R, Sternberg MJ. DSC: public domain protein secondary structure predication. Comput Appt Biosci. 1997;13;473-474.
49. Houben B, Michiels E, Ramakers M, Konstantoulea K, Louros N, Verniers J, van der Kant R, De Vleeschouwer M, Chicória N, Vanpoucke T, Gallardo R, Schymkowitz J, Rousseau F. Autonomous affregation suppression by acidic residues explains why chaperones favor basic residues. EMBO journal. 2020; e102864.
50. Koska JR, Spassov VZ, Maynard AJ, Yan L, Austin N, Flook PK, Venkatachalam CM. Fully automated molecular mechanics based induced fit protein-ligand docking method. J Chem Inf Mod 2008;48;1965-1973.
51. Best RB, Zhu X, Shim J, Lopes P, Mittal J, Feig M, Mackerell ADJr. Optimization of the additive CHARMm all-atom protein force field targeting improved sampling of the backbone φψ and side-chain χ_1 and χ_2 dihedral angles. *J Chem Theory Comput* 2012;8;3257-3273.
52. Black SD, Mould DR. Development of hydrophobicity parameters to analyze proteins which bear posttranslational or cotranslational modifications. Anal Biochem 1991;193;72-82.
53. Berntsson RP, Thunnissen AM, Poolman B, Slotboom DJ. Importance of a hydrophobic pocket for peptide binding in lactococcal OppA. J Bacteriol. 2011;193;4254-4256.
54. Eric BMD, Deog HO, Byong HL. Bioactive peptides. Foods 2017;6;32.

55. Macros-Contreras OA, Martinze de Lizarrondo S, Bardou I, Orset C, Pruvost M, Anfray A, Frigout Y, Hommet Y, Lebouvier L, Montaner J, Vivien D, Gauberti, M. Hyperfibrinolysis increases blood-brain barrier permeability by a plasmin- and bradykinin-dependent mechanism. Blood 2016;128;2423-2434.
56. Duchene J, Lecomte F, Ahmed S, Caryla C, Pesquero J, Bader M, Perretti M, Ahluwalia A. A novel inflammatory pathway involved in leukocyte recruitment: role for the kinin B1 receptor and the chemokine CXCL5. J immunol. 2007;179;4849-4856.
57. Leclerc PL, Gauthier SF, Bachelard H, Santure M, Rod, D. Antihypertensive activity of casein-enriched milk fermented by Lactobacillus helveticus. International Dairy Journal**.** 2002;12;995-1004.
58. Sharma S, Singh R, Rana, S. Bioactive peptides: A review. International journal Bioautomation 2011;15;223-250.
59. Picon A, van Wely KHM. Peptide binding to the *Bacillus subtilis* oligopeptide-binding proteins OppA and AppA. Mol Biol Today, 2001; 2; 21-25.
60. Chennamestty N, Voynov V, Kayser V, Helk B, Trout BL. Prediction of aggregation prone region of therapeutic proteins. J Phys Chem B. 2010;114; 6614-6624.
61. Ananieva EA, Powell JD, Huston SM. Leucine metabolism in T cell activation: mTOR signaling and beyond. Adv Nutr 2016;7;7985-8055.
62. Ren W, Liu G, Yin J, Tan B, Wu G, Bazer FW, Peng Y; Yin, Y. Amino-acid transporters in T-cell activation and differation. *Cell Death Dis* 2017;8; e2655.
63. Zhang S, Zeng X, Ren M, Mao X, Qiao, S. Novel metabolic and physiological functions of branched chain amino acids: a review. J Anim Sci Biotechnol. 2017; 8; 10.
64. Fang X, Bauer CE. Regulation of stringent factor by branched-chain amino acids. Proc Natl Acad Sci USA 2018;115;6446-6451.
65. Song W. Li D, Tao L, Luo Q, Chen, L. Solute carrier transporters: the metabolic gatekeepers of immune cells. Acta Pharm Sin B 2020;10;61-78.
66. Bröer S, Fairweather SJ. Amino acid transport across the mammalian intestine. Compr Physiol. 2018;9;343-373.
67. Sinclair LV, Rolf J, Emslie E, Shi YB, Taylor PM, Cantrell DA. Antigen receptor control of amino acid transporter coordinates the metabolic re-programing that is essential for T-cell differentiation. Nat Immunol. 2013;14; 500-508.
68. Powell JD, Pollizzi KN, Heikamp EB, Horton MR. Regulation of immune responses by mTOR. Annu Rev Immunol 2012;30;39-68.
69. Ma N, Guo P, Zhang J, He T, Kim SW, Zhang G, Ma, X. Nutrients mediate intestinal bacteria-mucosal immune crosstalk. Front immunol. 2018;9;5.
70. Agus A, Planchais J, Sokolo H. Gut Microbiota Regulation of tryptophan metabolism in helath and disease. Cell Host Microbe. 2018;23;716-724.

71. Quintero-Villegas A, Valdés-Ferrer SI. Role of 5-HT(7) receptors in the immune system in health and disease. Mol Med. 2019;26;2.
72. Kim JJ, Khan WI. 5-HT7 receptor signaling: improved therapeutic strategy in gut disorders. Front Behav Neurosci. 2014;8;396.
73. Ren M, Zhang SH, Zeng XF, Liu H, Qiao SY. Branched-chain amino acids are benetical to maintain growth performance and intestinal immune-related function in weaned piglets fed protein restricted diet. Asian-australas J Anim Sci. 2015;28;1742-1750.
74. Hu J, Nie Y, Chen S, Xie C, Fan Q, Wang Z, Long B, Yan G, Zhong Q, Yan, X. Leucine residues reactive oxygen species levels via an energy metabolism switch by activation of the mTOR-HIF-α pathway in porcine intestinal epithelial cells. Int J Biochem Cell Biol. 2017;89;42-56.
75. Zhang Z, Adelman AS, Rai D, Boettcher J, Lonnerdal B. Amino acid profiles in term and preterm human milk through lactation: a systematic review. *Nutrients.* 2013;26;4800-4821.
76. van Sadelhoff JHJ, Wiertsema SP, Garssen J, Hogenkamp, A. Free amino acids in human milk: a potential role for glutamine and glutamate in the protection against neonatal allergies and infections. Front Immunol 2020;11;1007.
77. Kawashima T, Kosaka A, Yan H, Guo Z, Uchiyama R, Fukui R, Kaneko D, Kumagai Y, You DJ, Carreras J, Uemastu S, Jang MH, Takeuchi O, Kaisho T, Akira S, Miyake K, Tsutsui H, Saito T, Nishimura I, Tsuji NM. Double-stranded RNA of intestinal commensal but not pathogenic bacteria triggers production of protective interferon-β. Immunity, 2013; 38;1187-1197.
78. Sarkar A, Mandal S. Bifidobacteria-insight into clinical outcomes and mechanisms of its probiotic action. Microbiol Res. 2016;192;159-171.
79. Sharma V, Mobeen F, Prakash T. Exploration of survival traits, probiotic determinants, host interactions, and functional evolution of bifidobacterial genomes using comparative genomics. Genes. 2018;9;477.
80. O'Callaghan A, Bottacini F, O'Connell Motherway M, van Sindreen D. Pangenome analysis of bifidobacterium longum and site-directed mutagenesis through by-pass of restriction-modification systems. BMC Genomics. 2015;16;832.
81. Lugli GA, Mancino W, Milani C, Duranti S, Turroni F, van Sinderen D, Ventura, M. Reconstruction of the bifidobacterial pan-secretome reveals the network of extracellular interactions between bifidobacteria and the infant gut. Appl Environ Microbiol**.** 2018;84;e00796-18.
82. Kim JH, Jeun EJ, Hong CP, Kim SH, Jang MS, Lee EJ, Moon SJ, Yun CH, lm SH, Jeong SG, Park BY, Kim KT, Seoh JY, Kim KT, Seoh JY, Kim YK, Oh SH, Ham JS, Yang BG, Jang MH. Extracellular vesicle-derived protein from *Bifidobacterium longum* alleviates food allergy through mast cell suppression. J Allergy Clin Immunol. 2016;137; 507-516.
83. Shanahan F. Probiotics in perspective. Gastroenterology, 2010;139;1808-1812.
84. Feng ZM, Li TJ, Wu L, Xiao DF, Blachier F, Yin YL. Monosodium l-gulatamte and dietary fat differently modify the composition of the intestinal microbiota in growing pigs. Obes Facts. 2015;8; 87-100.

85. Nie C, He T, Zhang W, Zhang G, Ma X. Branched chain amino acids: Beyond nutrition metabolism. Int J Mol Sci. 2018;19;954.
86. Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, Tobe T, Clarke J M, Topping DL Suzuki T, Taylor TD, Itoh K, Kikuji J, Morita H, Hattori M, Ohno H. Bifidobacteria can protect from emteropathogenic infection through production of acetate. Nature. 2011;469;543-547.
87. Takakuwa A, Nakamura K, Kikuchi M, Sugimoto R, Ohira S, Yokoi Y, Ayabe T. Butyric acid and leucine induce α-defensin secretion from small intestinal paneth cells. *Nutrients.* **2019**, 11, 2817.
88. Gioa DD, Gaggía F, Baffoni L, Stenico V. Benefical microbe in fermented and functional foods In: Rai VR, Bai J A. (Eds), Role of Bifidobacteria in the production of bioactive compounds and detoxification of harmful compounds. **2014** CRC press, New York (Chapter 16).
89. Guédon E, Serror P, Ehrlich SD, Renault P, Delorme C. Pleiotropic transcriptional repressor CodY senses the intracellular pool of branched-chain amino acids in Lactococcus lactis. Mol Microbiol.

Figures

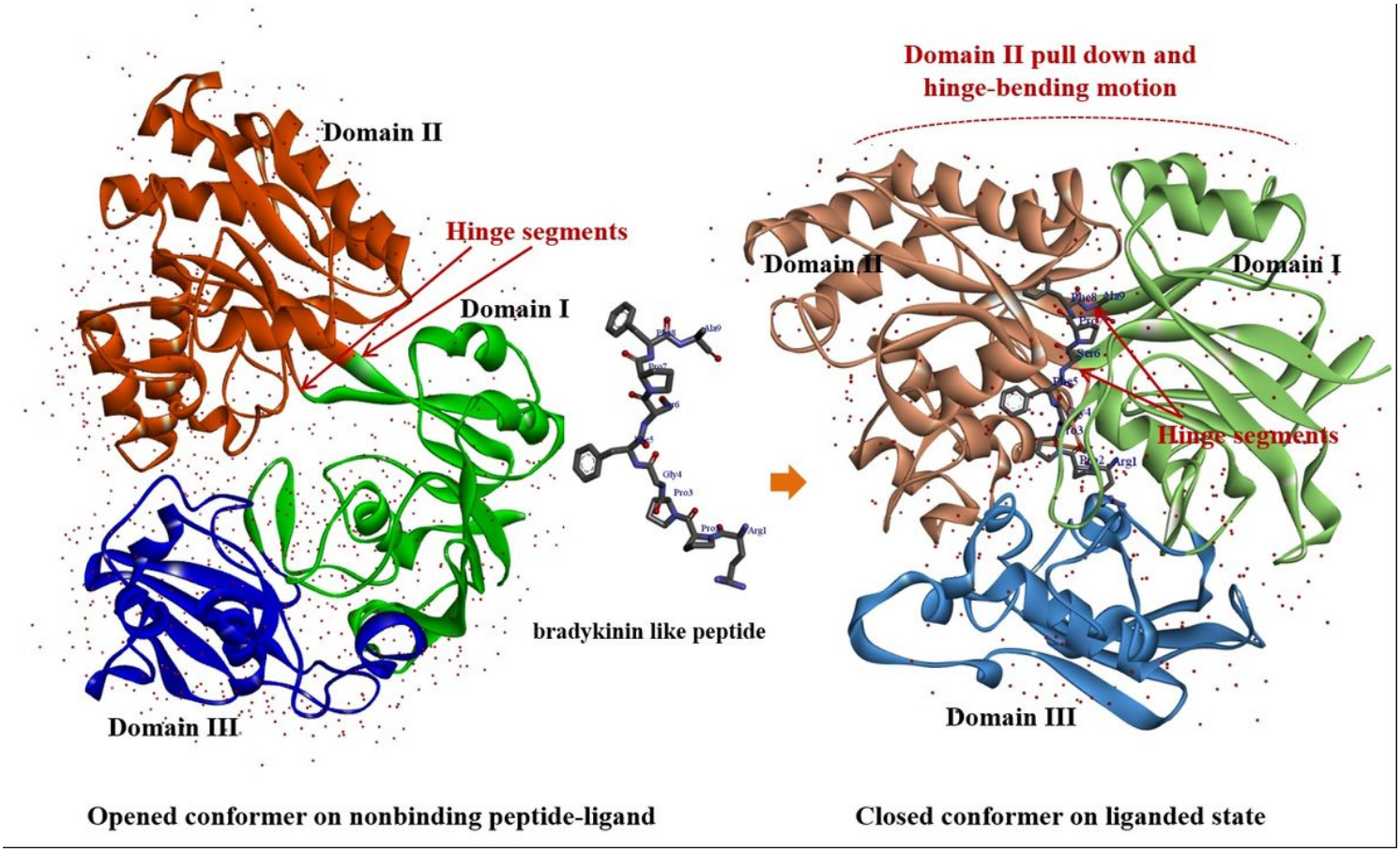


Figure 1

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 rotating 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.

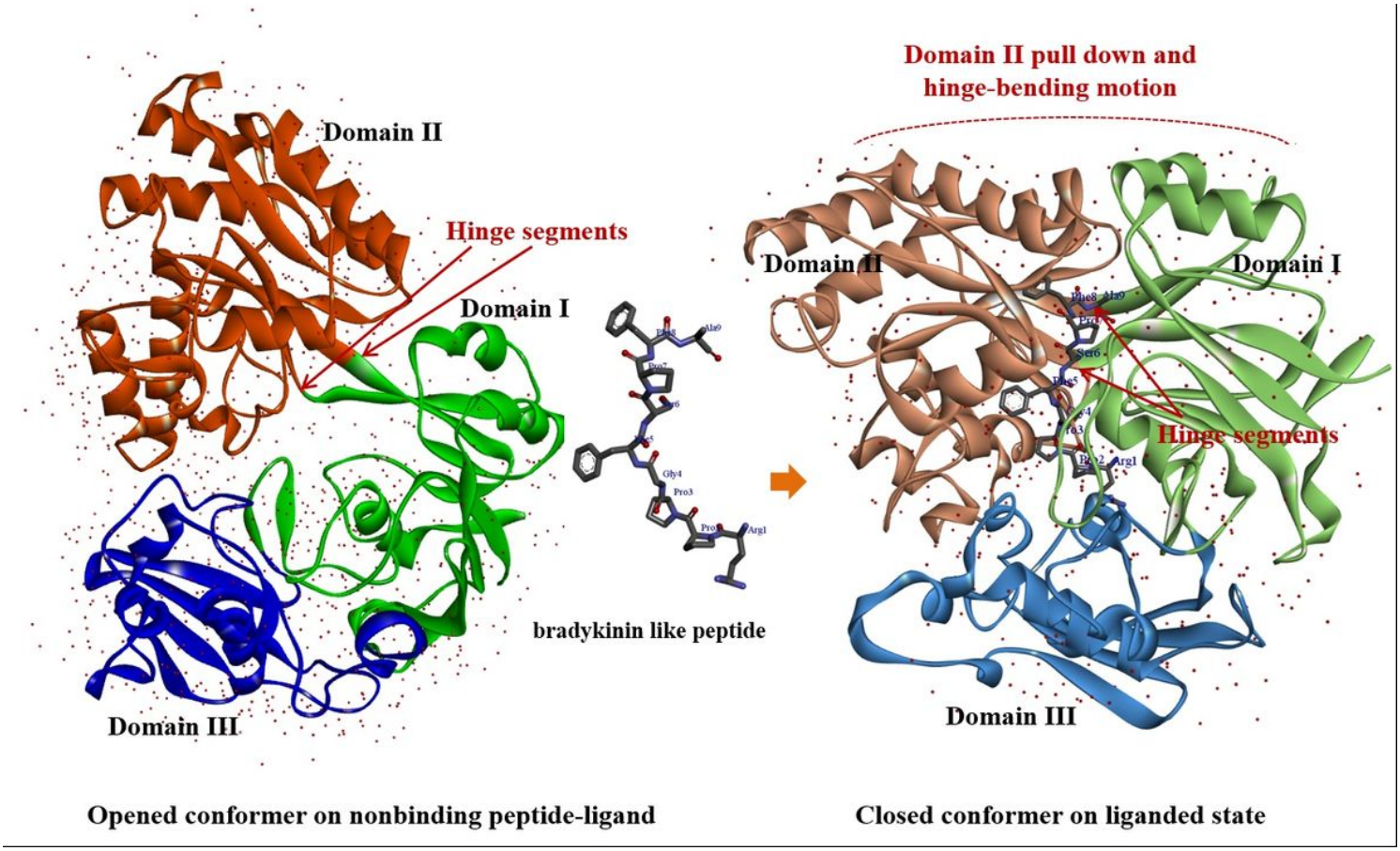


Figure 1

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 rotating 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.

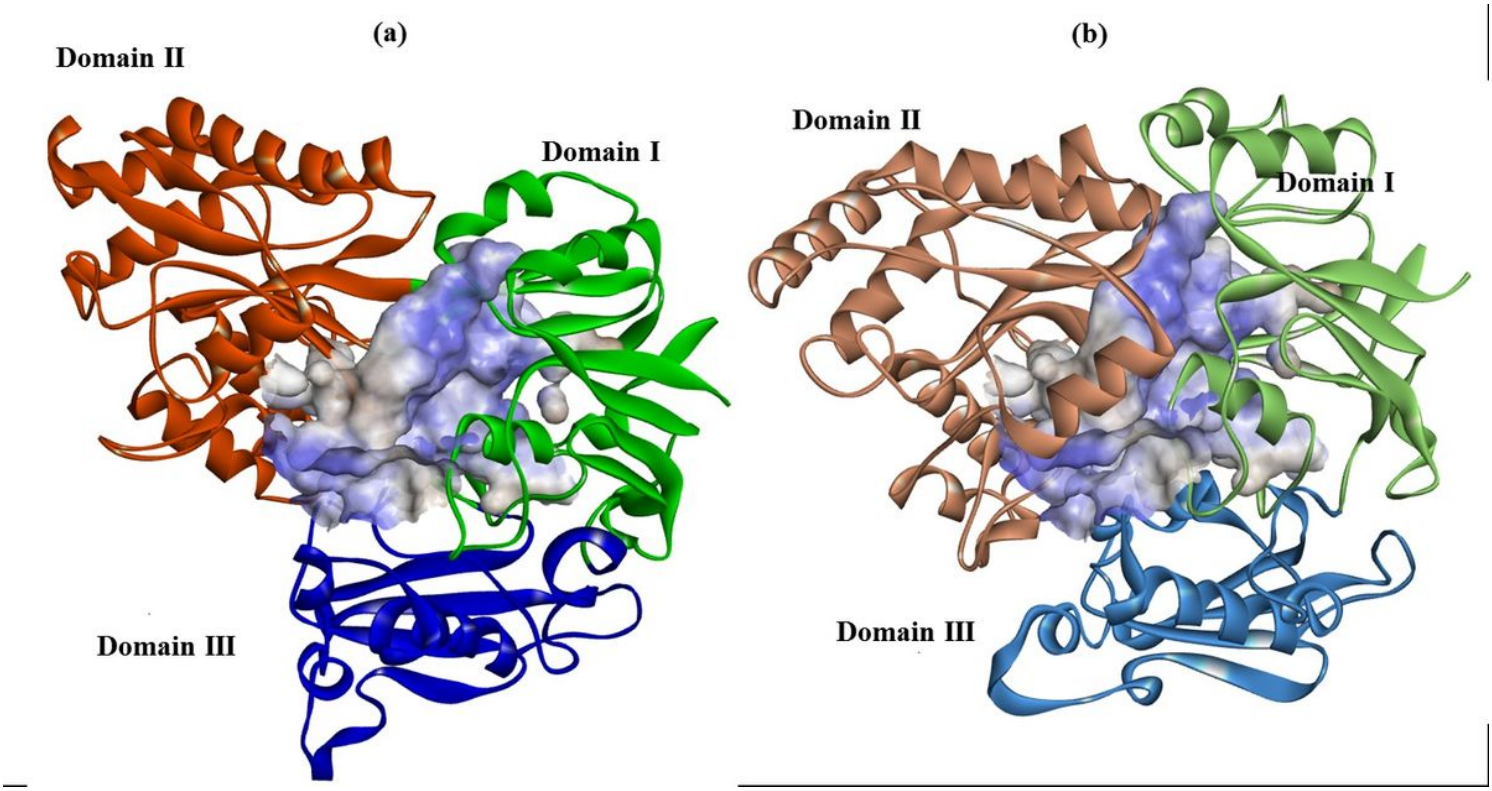


Figure 2

Comparison of hydrophobic binding cavity volumes changed between opened (a) and closed (b) conformers of OppA from B. longum KACC 91563. The flexible volume was calculated from 2740Å3(a) to 2200Å3(b) by Venkatachalam et al [89] when the OppA is close to the bound peptide ligand.

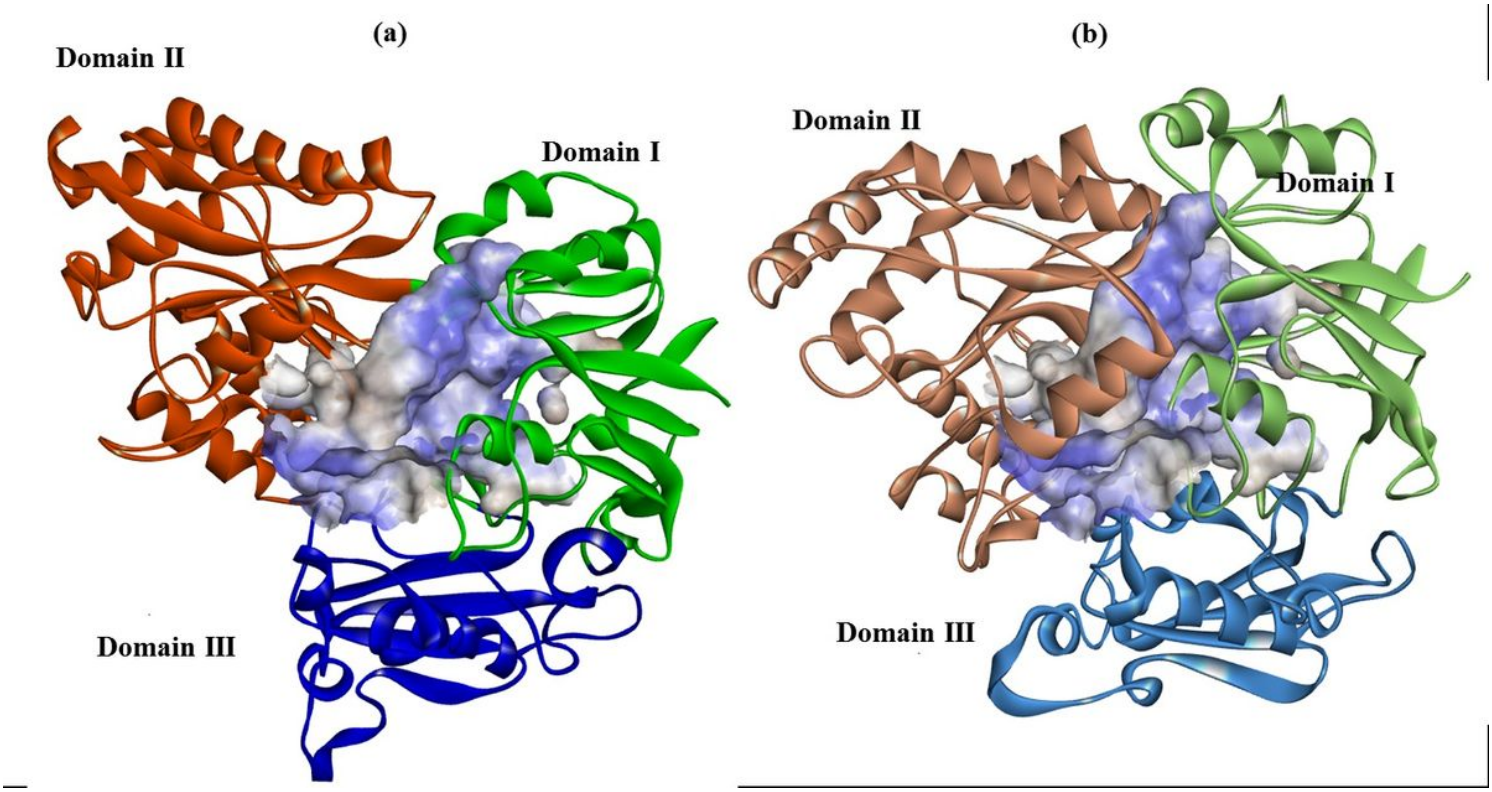


Figure 2

Comparison of hydrophobic binding cavity volumes changed between opened (a) and closed (b) conformers of OppA from B. longum KACC 91563. The flexible volume was calculated from 2740Å3(a) to 2200Å3(b) by Venkatachalam et al [89] when the OppA is close to the bound peptide ligand.

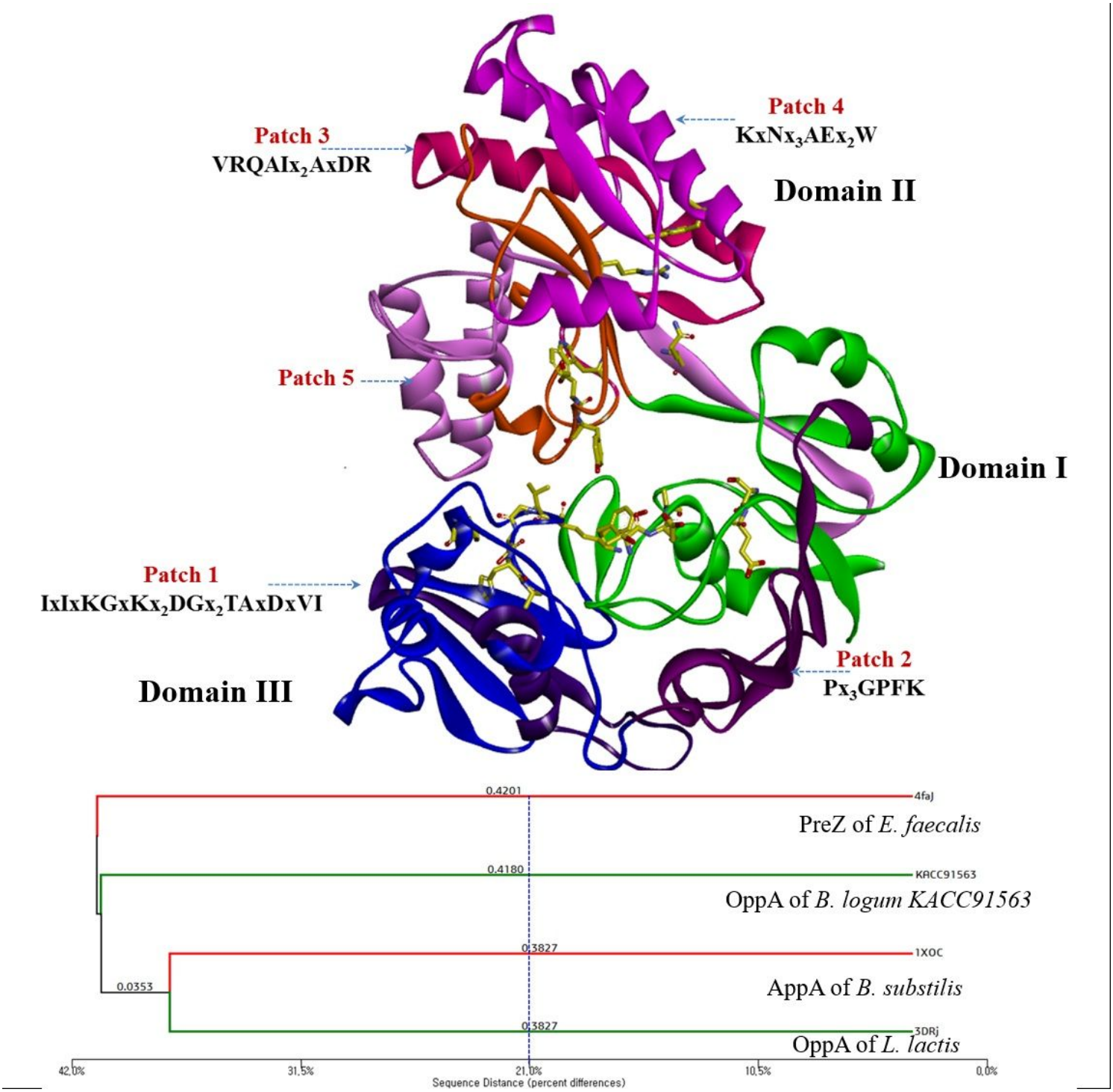


Figure 3

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, when the color scheme of 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.

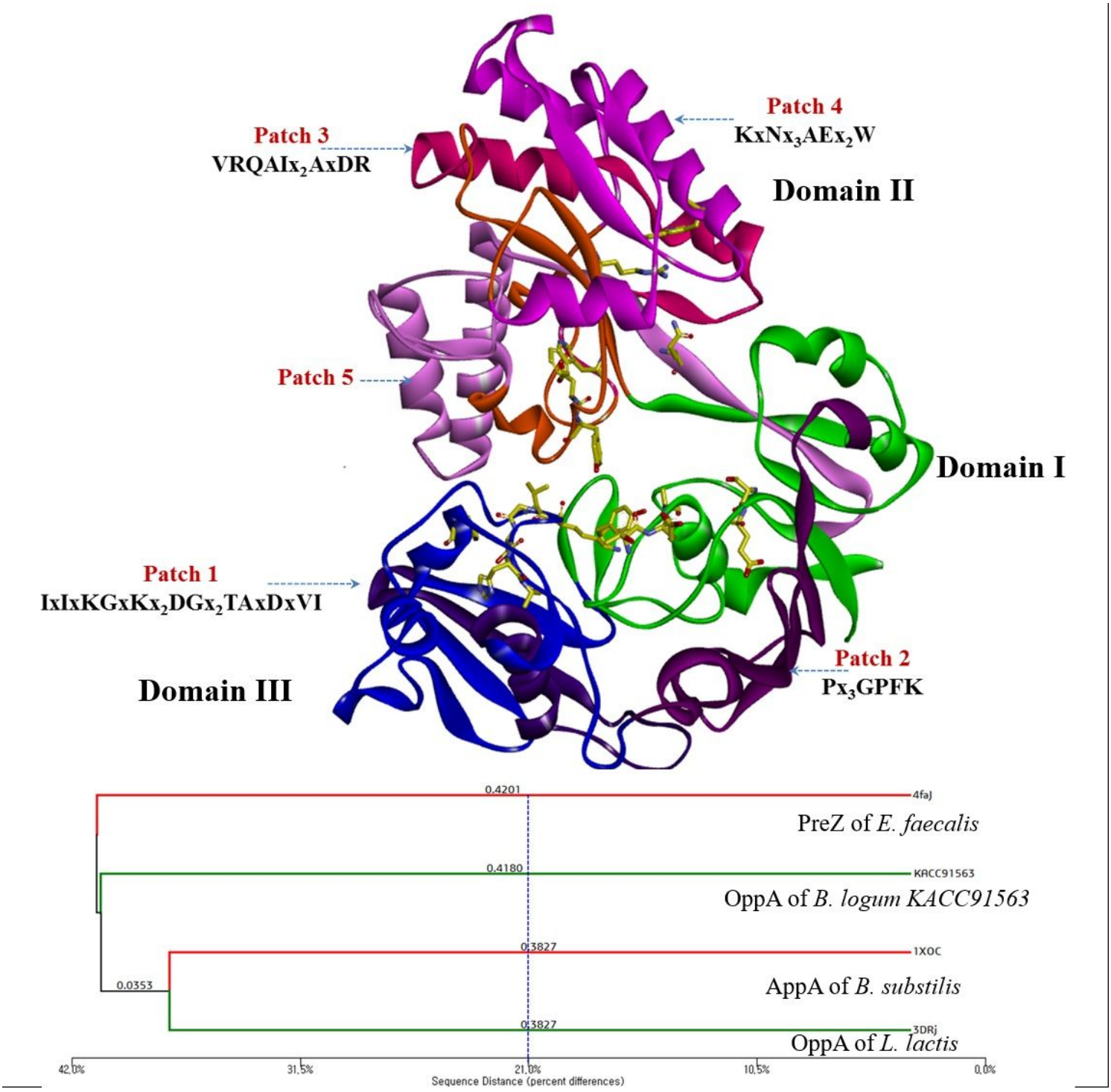


Figure 3

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, when the color scheme of 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.

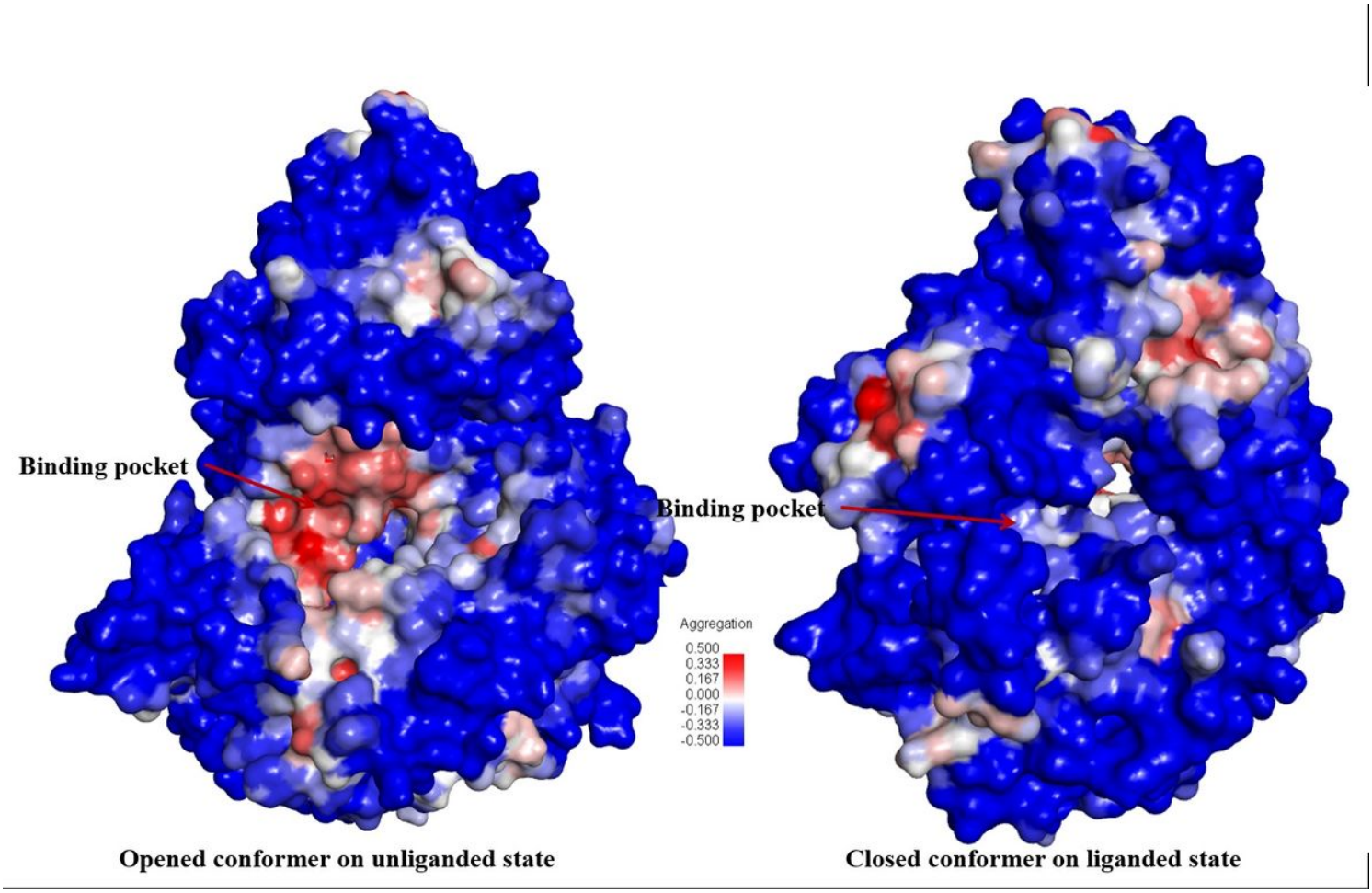


Figure 4

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.

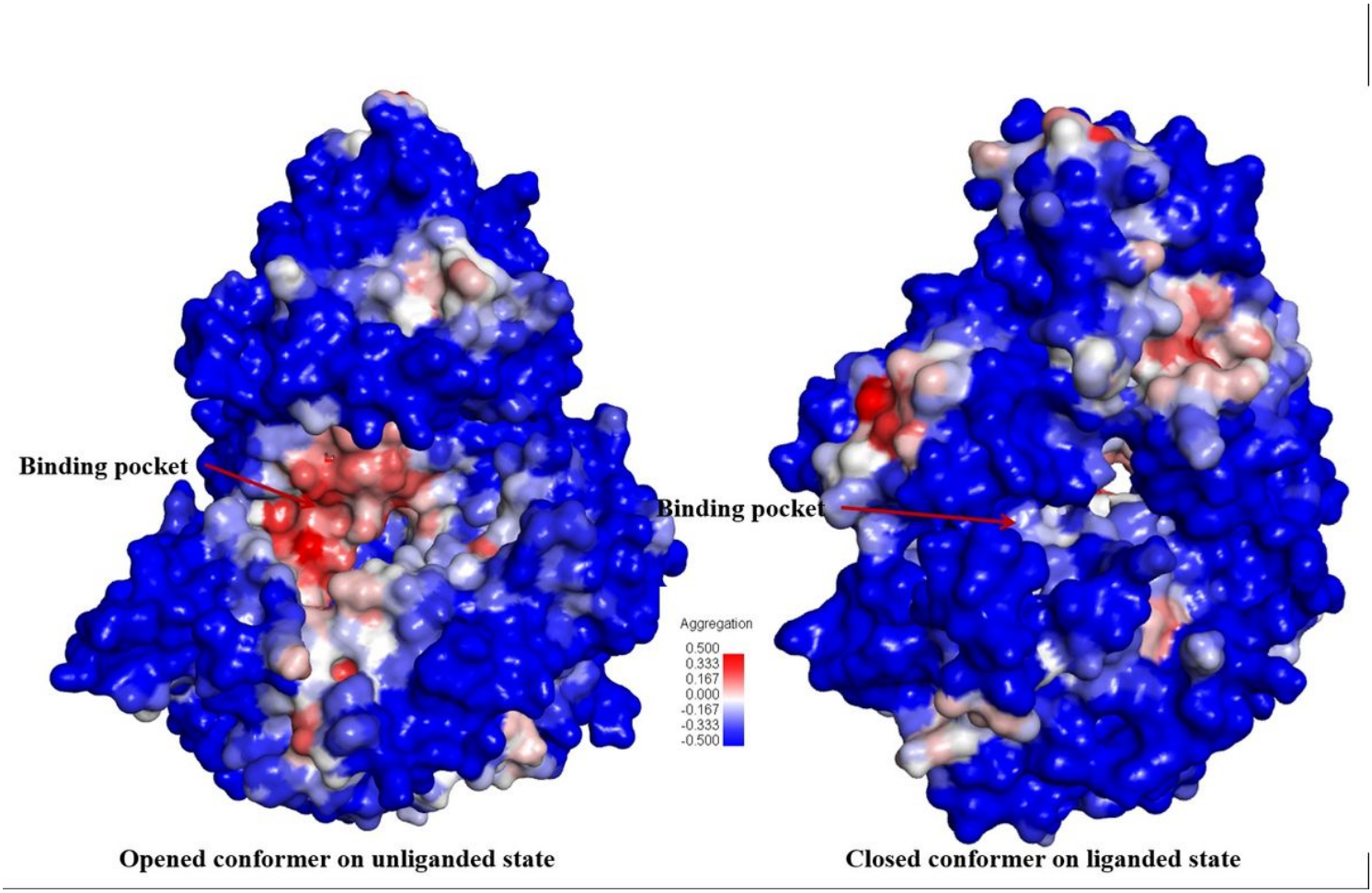


Figure 4

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.

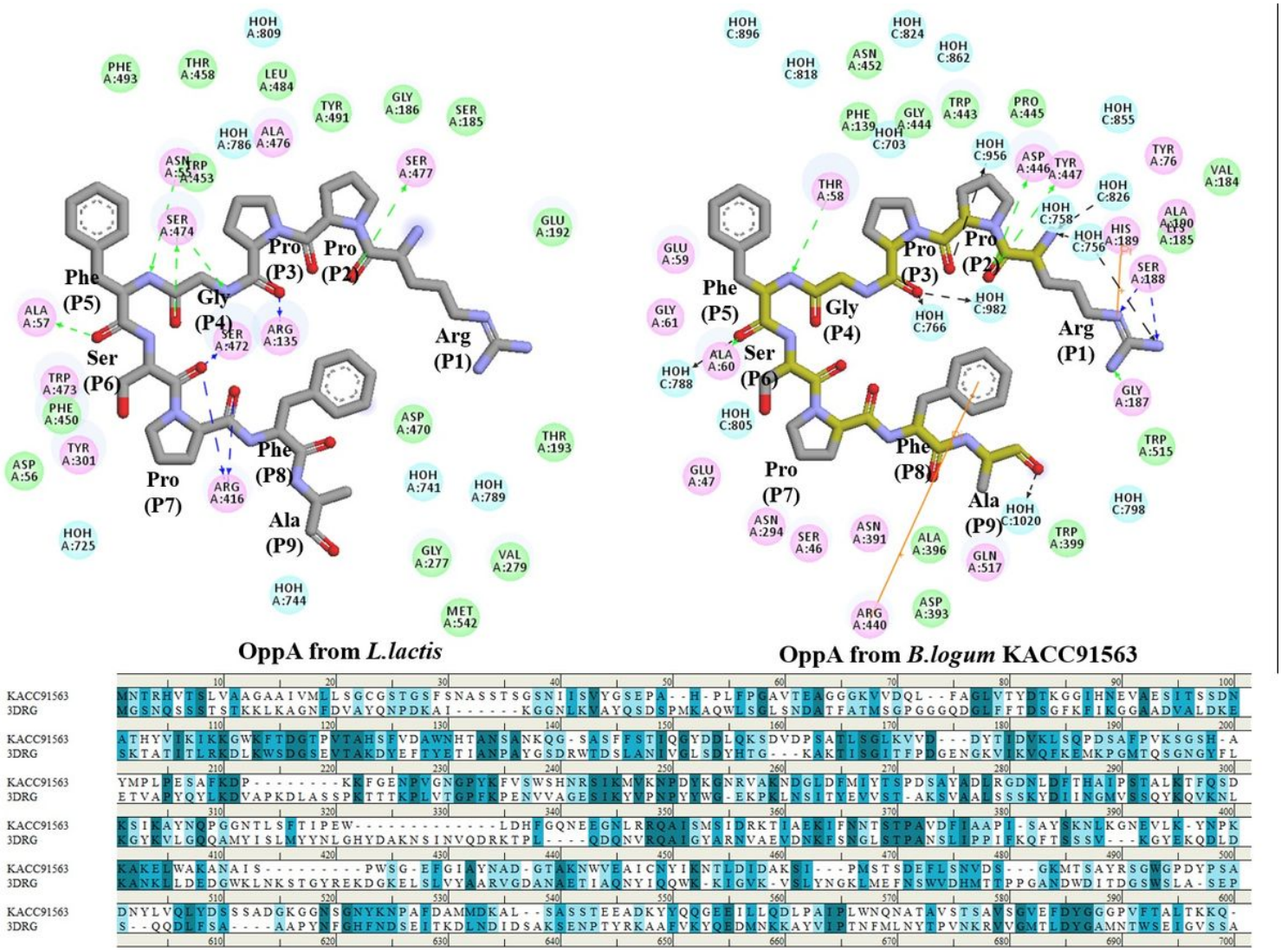


Figure 5

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.

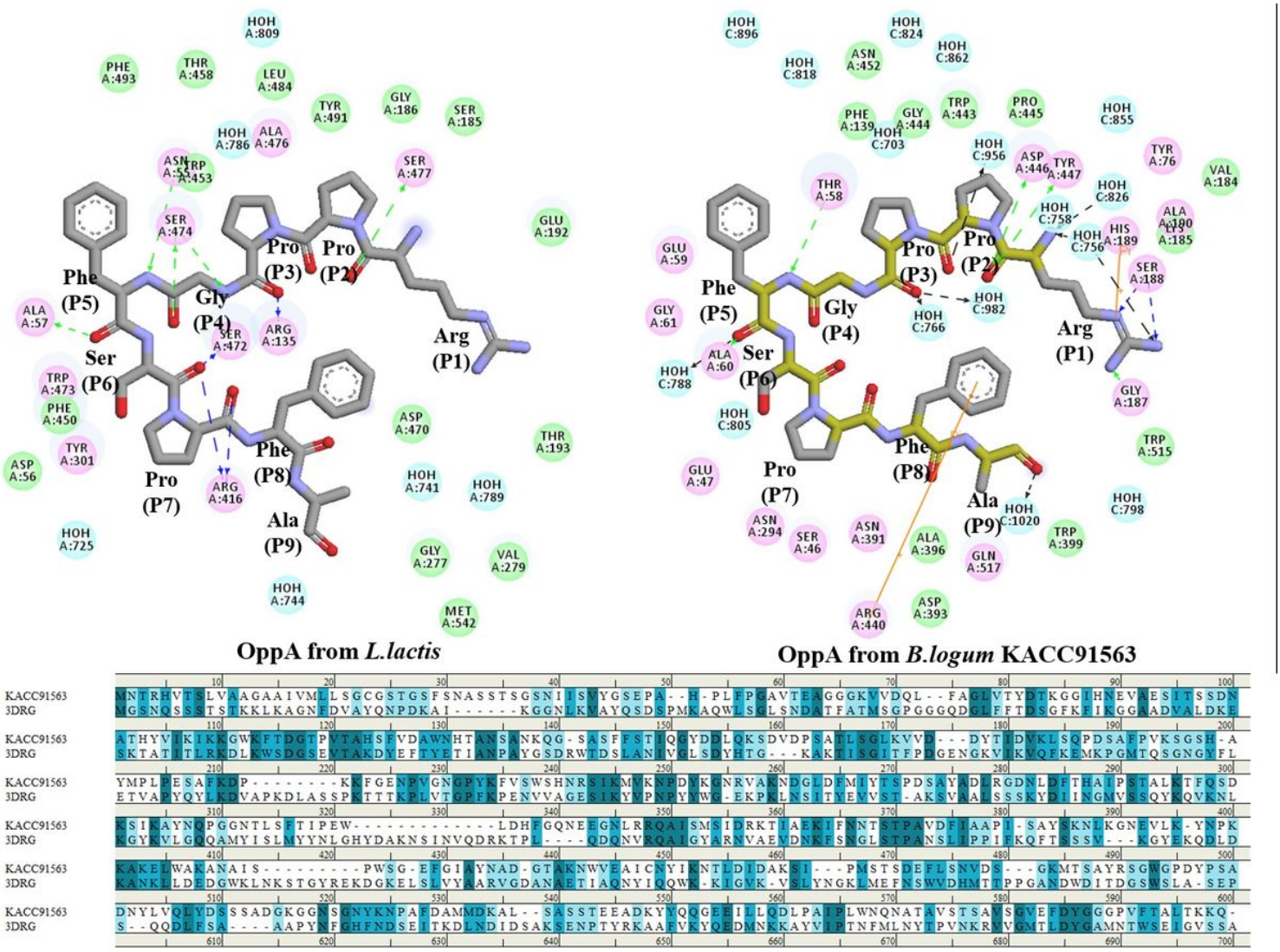


Figure 5

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.

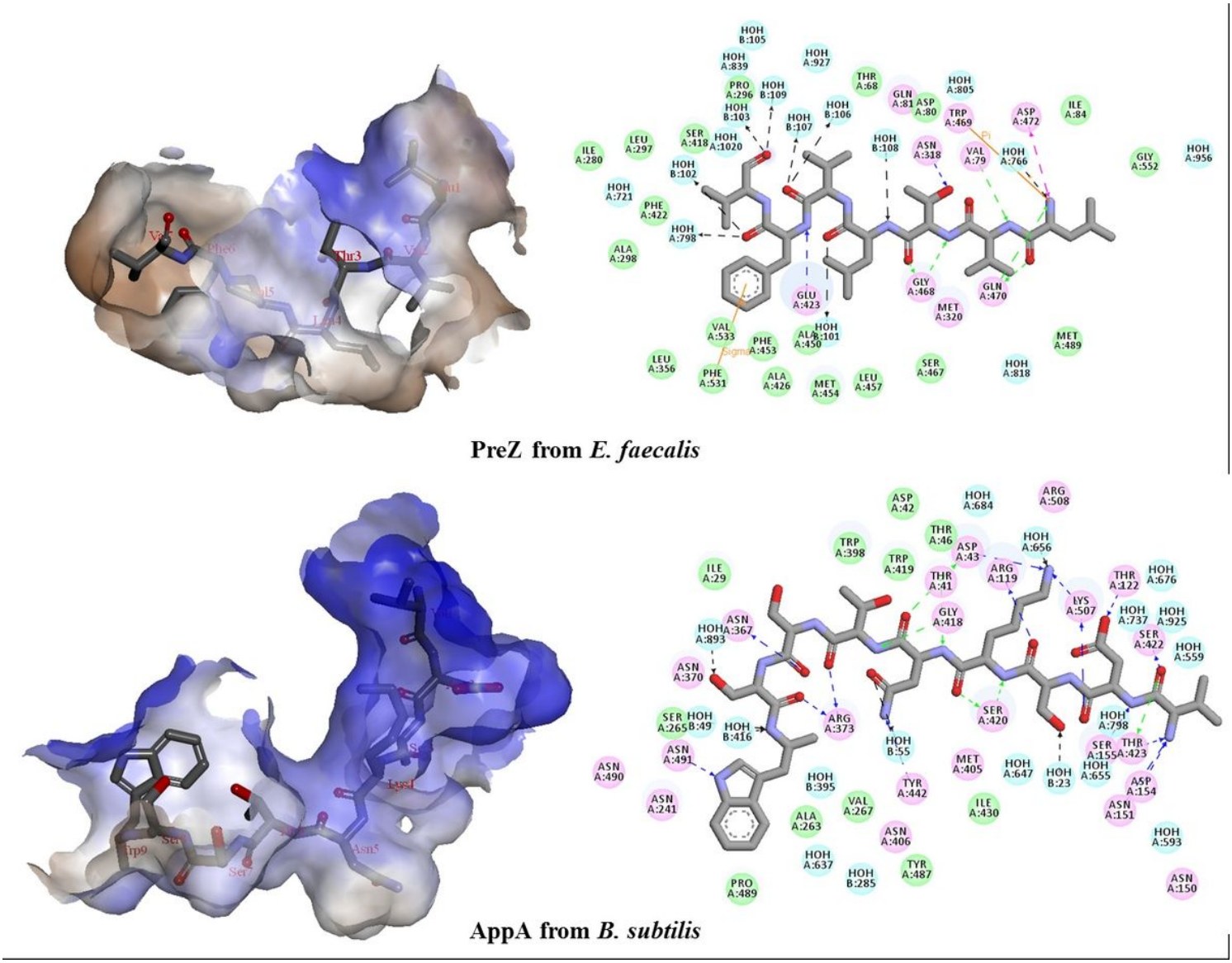


Figure 6

The interaction contacts of 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.

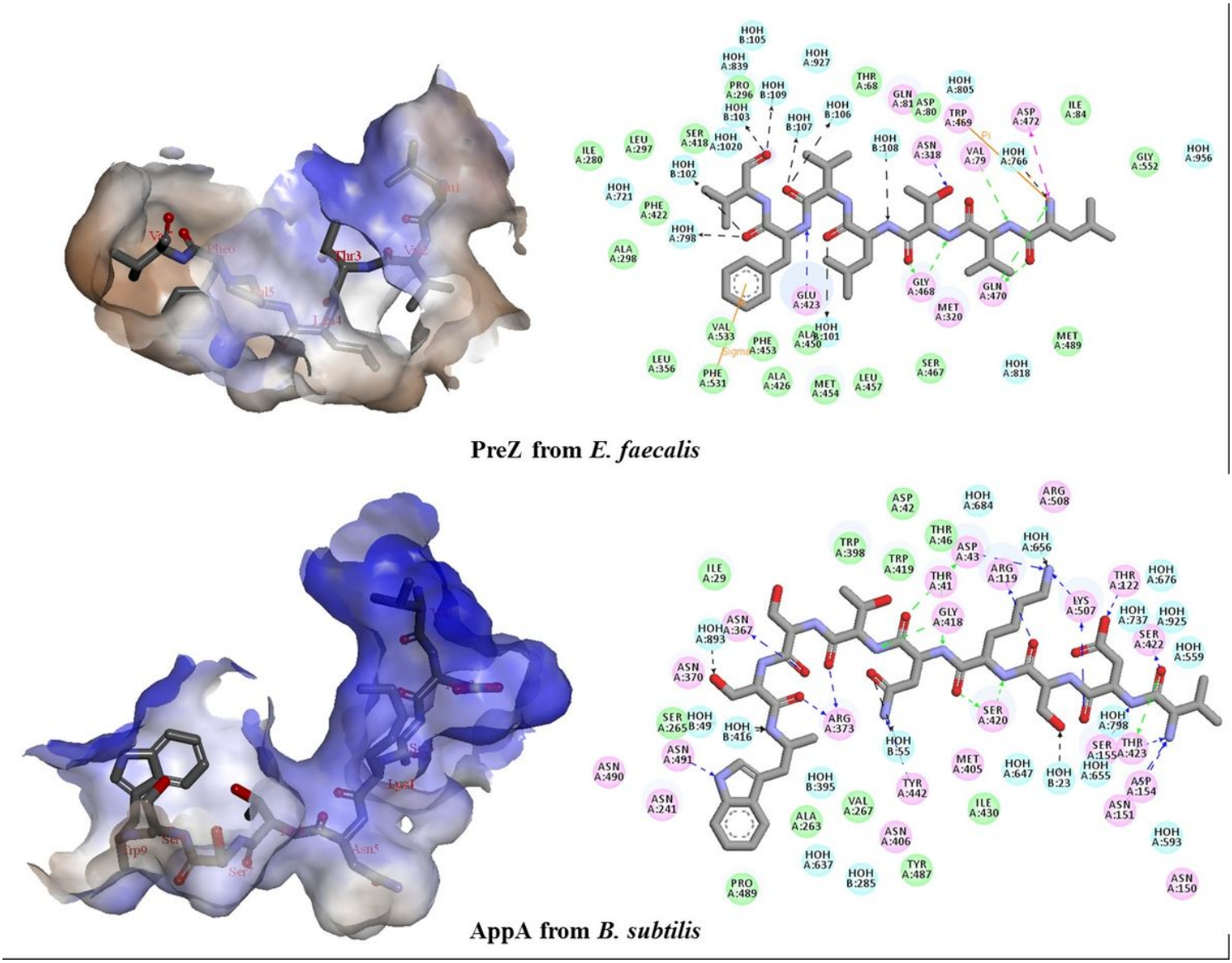
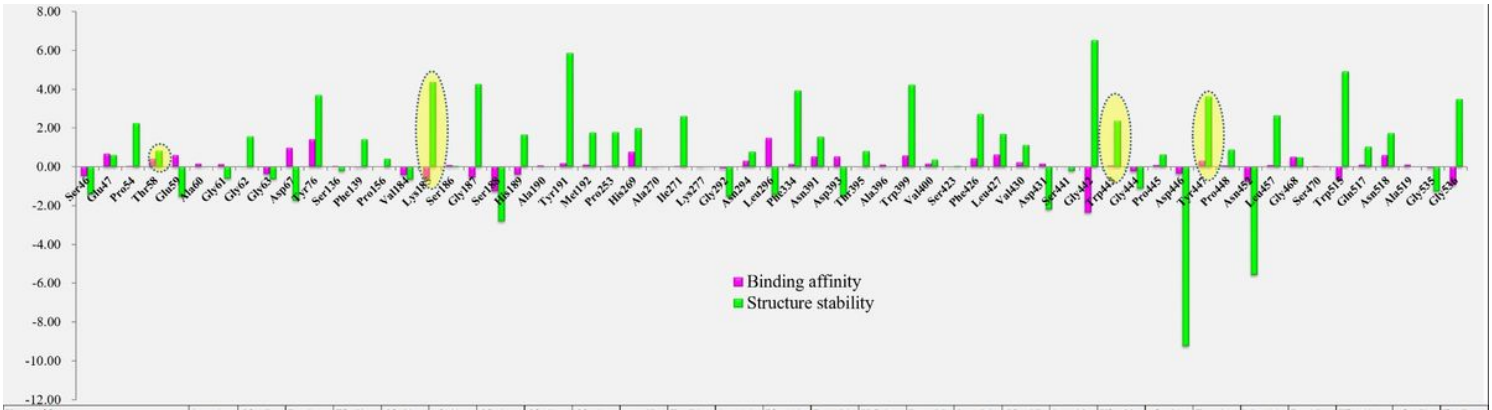


Figure 6

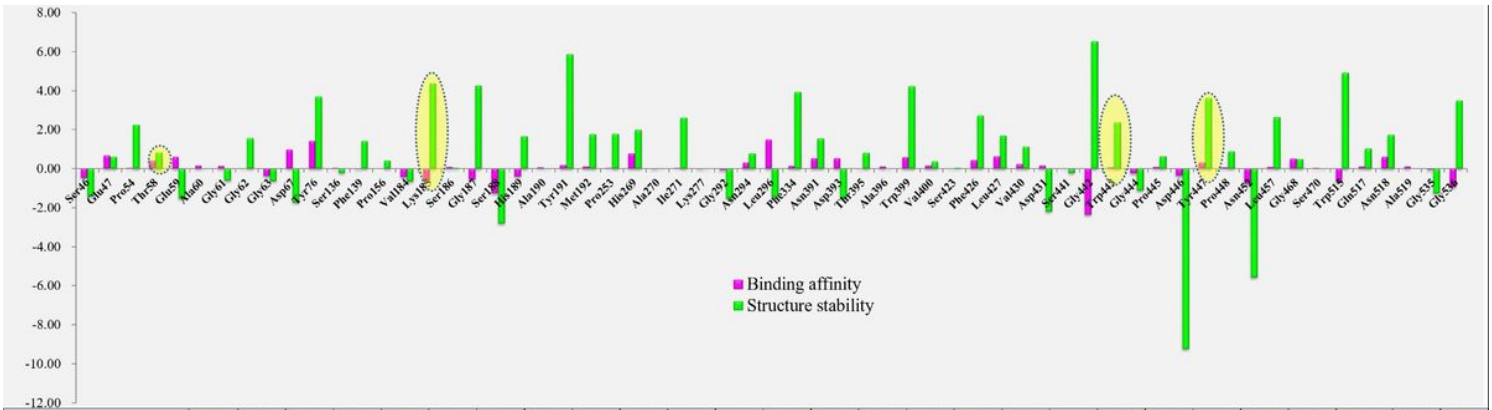
The interaction contacts of between the peptide ligands and two template receptors (PreZ and AppA) from their crystalized structures (PDB ID: 4FAJ and 1XOC). The interaction residues of PreZ from E. faecalis complexed with 7-mer peptide (LVTLVFV) and for 9-mer peptide (VDSKNTSSW) within AppA from B.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



Key residues	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
Binding affinity(kcal/mol)	-0.48	0.67	0.02	0.39	0.59	0.15	0.12	-0.02	-0.39	0.96	1.39	0.02	-0.05	0.00	-0.43	-0.63	0.08	-0.57	-1.27	-0.42	0.04	0.18	0.11	0.01	0.76	-0.04	0.01
Structure stability(kcal/mol)	-1.42	0.62	2.25	0.83	-1.54	0.01	-0.62	1.58	-0.64	-1.73	3.70	-0.21	1.42	0.43	-0.65	4.37	0.06	4.26	-2.81	1.67	0.01	5.87	1.77	1.80	1.98	0.01	2.61
Key residues	Lys277	Gly292	Asn294	Leu296	Phe334	Asn391	Asp393	Thr395	Ala396	Trp399	Val400	Ser423	Phe426	Leu427	Val430	Asp431	Ser441	Gly442	Trp443	Gly444	Pro445	Asp446	Tyr447	Pro448	Asn452	Leu457	Gly468
Binding affinity(kcal/mol)	-0.05	-0.09	0.29	1.47	0.13	0.52	0.51	-0.03	0.10	0.56	0.14	-0.01	0.41	0.61	0.23	0.15	-0.01	-2.36	0.04	-0.24	0.08	-0.37	0.29	0.05	-0.79	0.08	0.49
Structure stability(kcal/mol)	-0.03	-1.54	0.79	-1.45	3.95	1.54	-1.48	0.82	-0.01	4.23	0.37	0.05	2.71	1.69	1.14	-2.20	-0.22	6.54	2.38	-1.12	0.64	-9.24	3.63	0.88	-5.57	2.65	0.48
Key residues	Gly468	Ser470	Trp515	Gln517	Asn518	Ala519	Gly535	Gly536																			
Binding affinity(kcal/mol)	0.49	0.01	-0.67	0.10	0.59	0.09	-0.06	-0.95																			
Structure stability(kcal/mol)	0.48	-0.02	4.93	1.04	1.74	0.00	-1.28	3.50																			

Figure 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 < -0.5kcal/mol), neutral (-0.5kcal/mol < mutation energy < 0.5kcal/mol), and destabilizing (>0.5kcal/mol).



Key residues	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
Binding affinity(kcal/mol)	-0.48	0.67	0.02	0.39	0.59	0.15	0.12	-0.02	-0.39	0.96	1.39	0.02	-0.05	0.00	-0.43	-0.63	0.08	-0.57	-1.27	-0.42	0.04	0.18	0.11	0.01	0.76	-0.04	0.01
Structure stability(kcal/mol)	-1.42	0.62	2.25	0.83	-1.54	0.01	-0.62	1.58	-0.64	-1.73	3.70	-0.21	1.42	0.43	-0.65	4.37	0.06	4.26	-2.81	1.67	0.01	5.87	1.77	1.80	1.98	0.01	2.61
Key residues	Lys277	Gly292	Asn294	Leu296	Phe334	Asn391	Asp393	Thr395	Ala396	Trp399	Val400	Ser423	Phe426	Leu427	Val430	Asp431	Ser441	Gly442	Trp443	Gly444	Pro445	Asp446	Tyr447	Pro448	Asn452	Leu457	Gly468
Binding affinity(kcal/mol)	-0.05	-0.09	0.29	1.47	0.13	0.52	0.51	-0.03	0.10	0.56	0.14	-0.01	0.41	0.61	0.23	0.15	-0.01	-2.36	0.04	-0.24	0.08	-0.37	0.29	0.05	-0.79	0.08	0.49
Structure stability(kcal/mol)	-0.03	-1.54	0.79	-1.45	3.95	1.54	-1.48	0.82	-0.01	4.23	0.37	0.05	2.71	1.69	1.14	-2.20	-0.22	6.54	2.38	-1.12	0.64	-9.24	3.63	0.88	-5.57	2.65	0.48
Key residues	Gly468	Ser470	Trp515	Gln517	Asn518	Ala519	Gly535	Gly536																			
Binding affinity(kcal/mol)	0.49	0.01	-0.67	0.10	0.59	0.09	-0.06	-0.95																			
Structure stability(kcal/mol)	0.48	-0.02	4.93	1.04	1.74	0.00	-1.28	3.50																			

Figure 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 < -0.5kcal/mol), neutral (-0.5kcal/mol < mutation energy < 0.5kcal/mol), and destabilizing (>0.5kcal/mol).

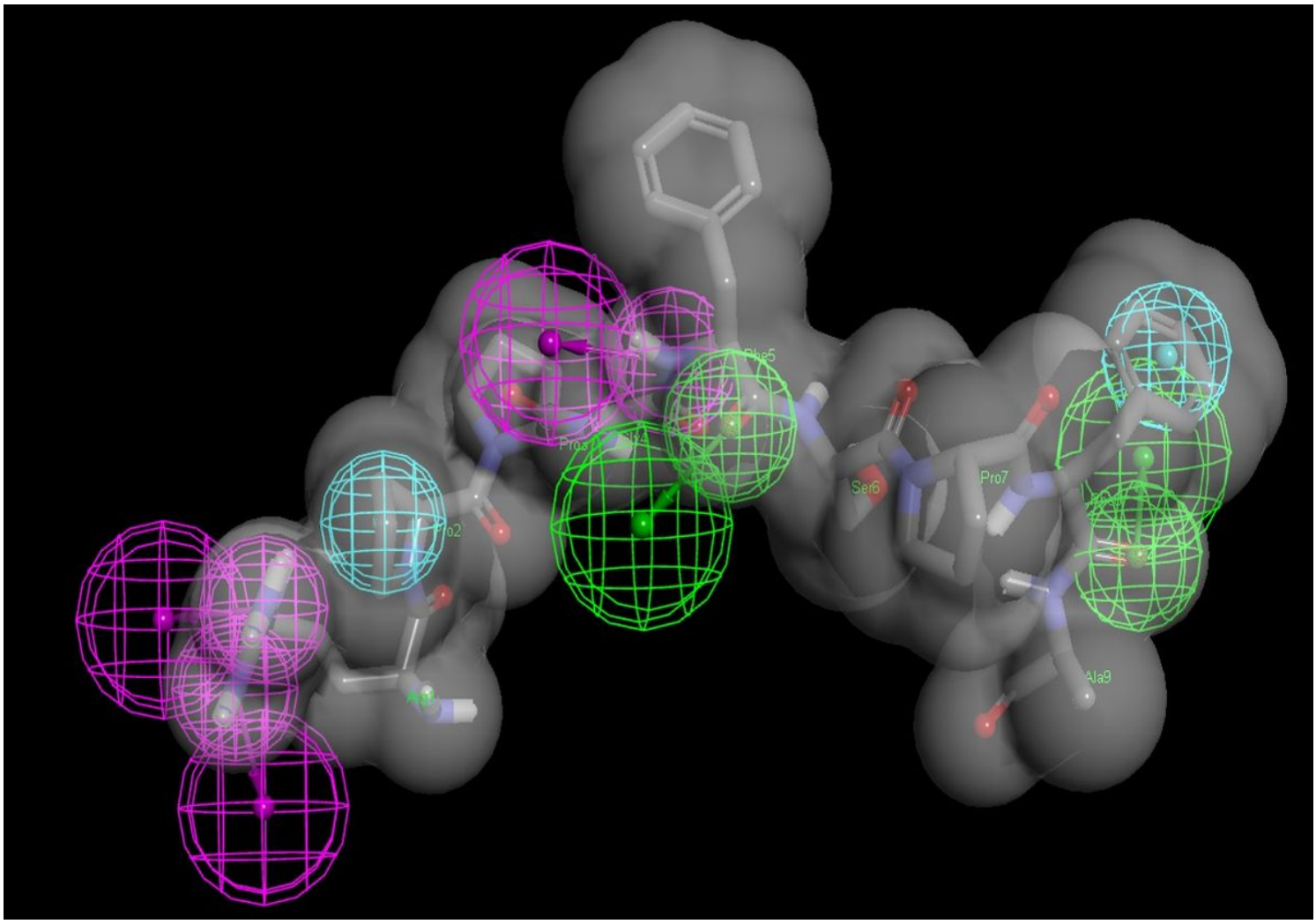


Figure 8

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).

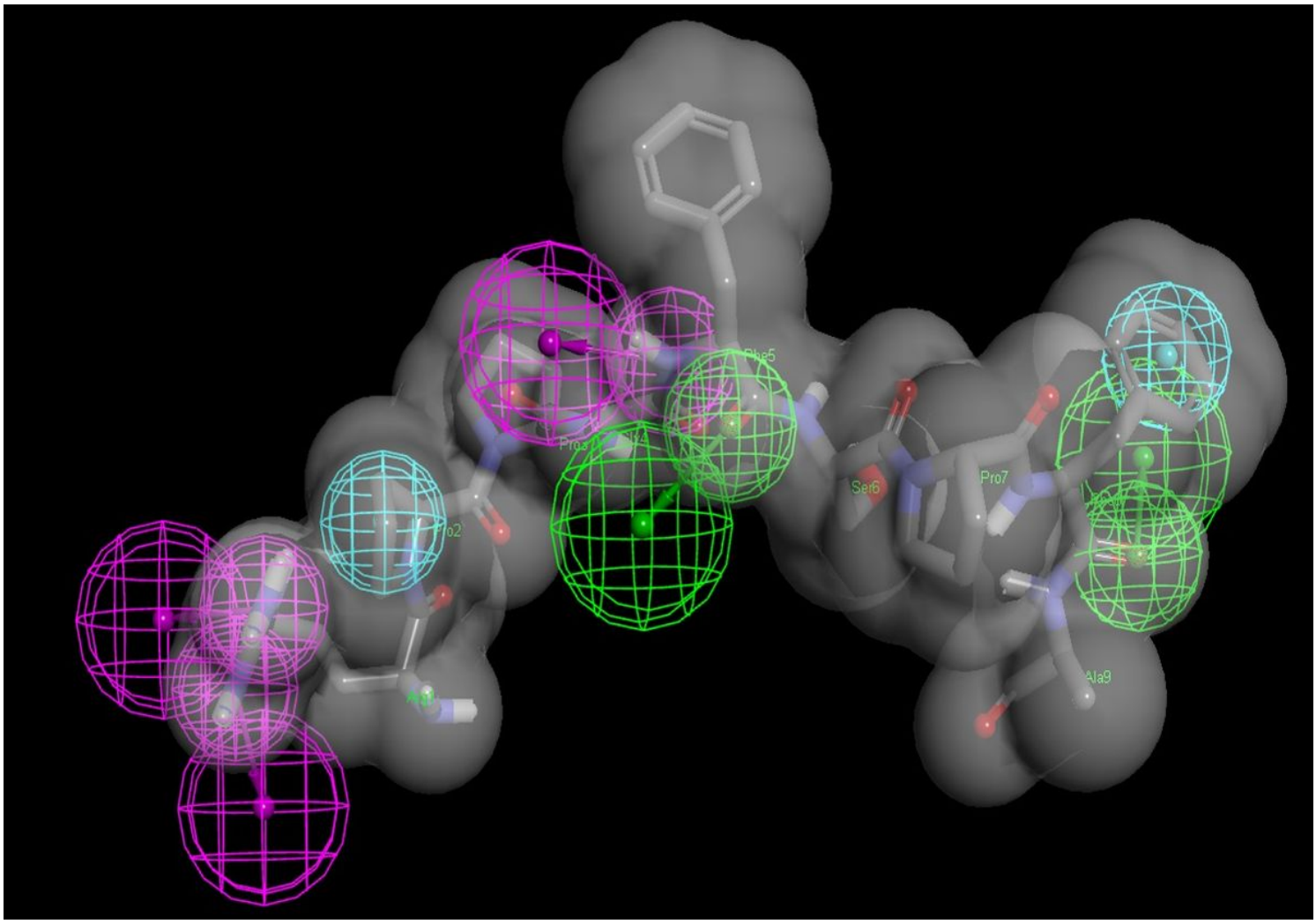


Figure 8

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).

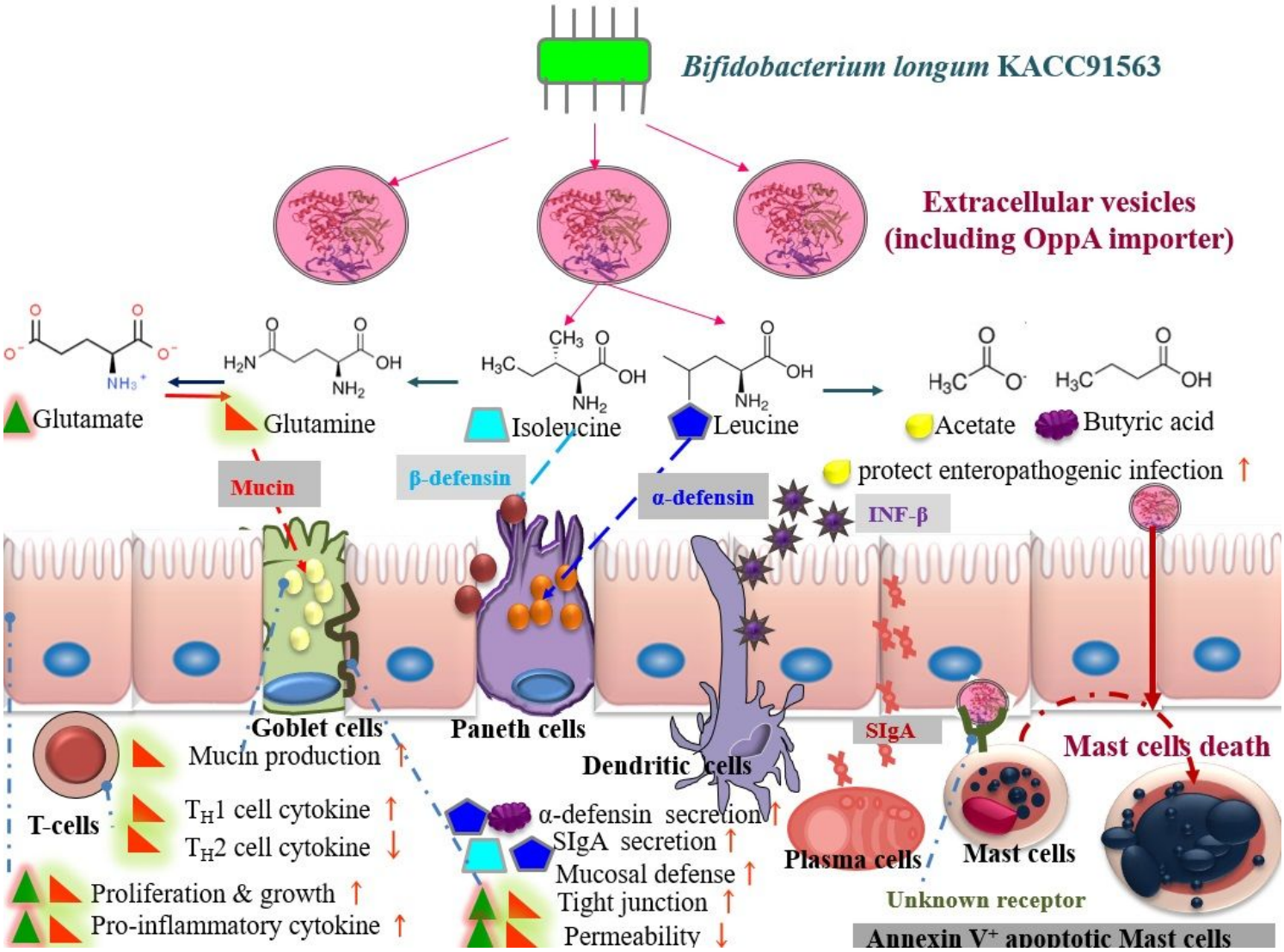


Figure 9

The host immunity effects of amino acid substrate from probiotic Bifidobacterium longum KACC91563 in the intestine. Then, the OppA importer in B. logum KACC91563-derived extracellular vesicles is a pivotal role in determining the substrate specificity that supports the probiotic effect.

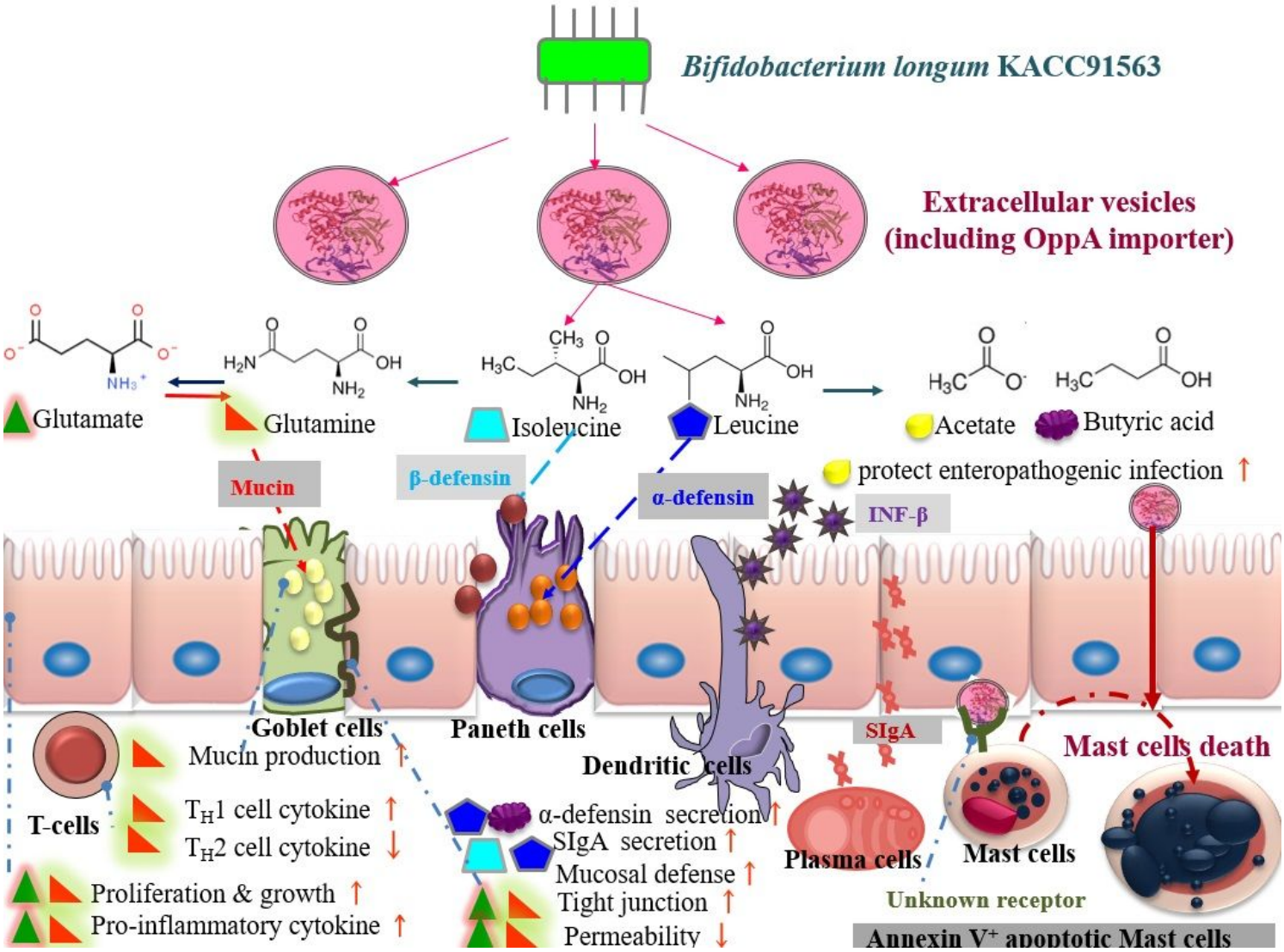


Figure 9

The host immunity effects of amino acid substrate from probiotic Bifidobacterium longum KACC91563 in the intestine. Then, the OppA importer in B. logum KACC91563-derived extracellular vesicles is a pivotal role in determining the substrate specificity that supports the probiotic effect.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- peptideABCtransportersofOppAsupplementary.docx
- peptideABCtransportersofOppAsupplementary.docx