

Title	Ligand recognition mechanism of dCACHE type chemoreceptors of <i>Vibrio cholerae</i>
Author(s)	高橋, 洋平
Citation	大阪大学, 2019, 博士論文
Version Type	VoR
URL	<a href="https://doi.org/10.18910/72673">https://doi.org/10.18910/72673</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University

**Ligand recognition mechanism of  
dCACHE type chemoreceptors of *Vibrio cholerae***

A Doctoral Thesis

by

**Yohei Takahashi**

Submitted to

the Graduate School of Science,

Osaka University

February, 2019

## Contents

<b>Chapter 1 General introduction</b>	<b>4</b>
1-1. Bacterial swimming mode	4
1-2. Bacterial chemotaxis	5
1-3. Chemoreceptor	7
1-4. <i>Vibrio cholerae</i>	11
1-5. Chemoreceptors of <i>Vibrio cholerae</i>	13
1-6. Outline of this study	14
1-7. References	16
<b>Chapter 2 Molecular mechanism of ligand recognition by Mlp37</b>	<b>21</b>
2-1. Introduction	21
2-2. Materials and Methods	23
2-2-1. Expression and purification of the periplasmic fragment of Mlp37	23
2-2-2. Crystallization	24
2-2-3. X-ray data collection and structure determination	26
2-3. Results	27
2-3-1. Overall structures of Mlp37p in complex with its ligand	27
2-3-2. Ligand recognition mechanism of Mlp37p	30
2-4. Discussion	35
2-4-1. Residues involved in ligand binding of Mlp37p	35
2-4-2. Conformational difference of Mlp37p CAD1 in various ligand binding structures	38
2-4-3. Role of CAD1 and CAD2	40
2-4-4. <i>In vivo</i> observation of ligand binding to Mlp37	41
2-5. References	43

<b>Chapter 3 Molecular mechanism of ligand recognition by Mlp24</b>	<b>45</b>
3-1. Introduction	45
3-2. Materials and Methods	48
3-2-1. Expression and purification of the periplasmic fragment of Mlp24	48
3-2-2. Crystallization	49
3-2-3. X-ray data collection and structure determination of Mlp24p	52
3-3. Results	53
3-3-1. Overall structures of ligand-free Mlp24p and Mlp24p in complex with various ligands	53
3-3-2. Ligand recognition mechanism of Mlp24p	59
3-3-3. Ligand binding induces the domain closure of CAD1.	62
3-4. Discussion	65
3-4-1. Calcium ion enhances the binding affinity of Mlp24p to ligands	65
3-4-2. Structural basis of the binding affinity of Mlp24p and Mlp37p for various amino acids	71
3-4-3. Conformational change of Mlp24p upon ligand binding	74
3-4-4. Implication for the signal transduction mechanism of Mlp24p	75
3-4-5. <i>Vibrio cholerae</i> may fine-tune its chemotactic behavior by calcium ion concentration	77
3-5. References	79

**Chapter 4 Structure of *V. cholerae* accessory colonisation factor, AcfC and its interaction with Mlp8** **83**

4-1. Introduction	83
4-2. Materials and Methods	87
4-2-1. Expression and purification of AcfC	87
4-2-2. Expression and purification of Mlp8p with C-terminal hexahistidine tag (Mlp8-His6)	88

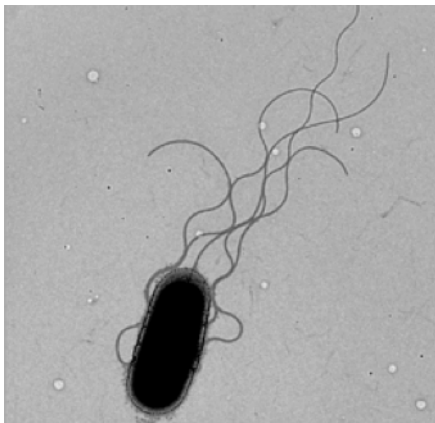
4-2-3. Crystallization of AcfC	89
4-2-4. X-ray data collection and structure determination of AcfC	90
4-2-5. Pull-down assay	92
4-2-6. Isothermal titration calorimetry (ITC)	92
4-3. Results	93
4-3-1. Structure of ligand-free AcfC	93
4-3-2. D-malate weakly binds to AcfC in solution	98
4-3-3. The AcfC binding to Mlp8p is promoted by D-malate but not by galactose-6-sulfate	98
4-4. Discussion	101
4-5. References	103
<b>Chapter 5 Summary</b>	<b>106</b>
<b>Acknowledgements</b>	<b>108</b>
<b>List of Publications</b>	<b>109</b>

## Chapter 1

### General introduction

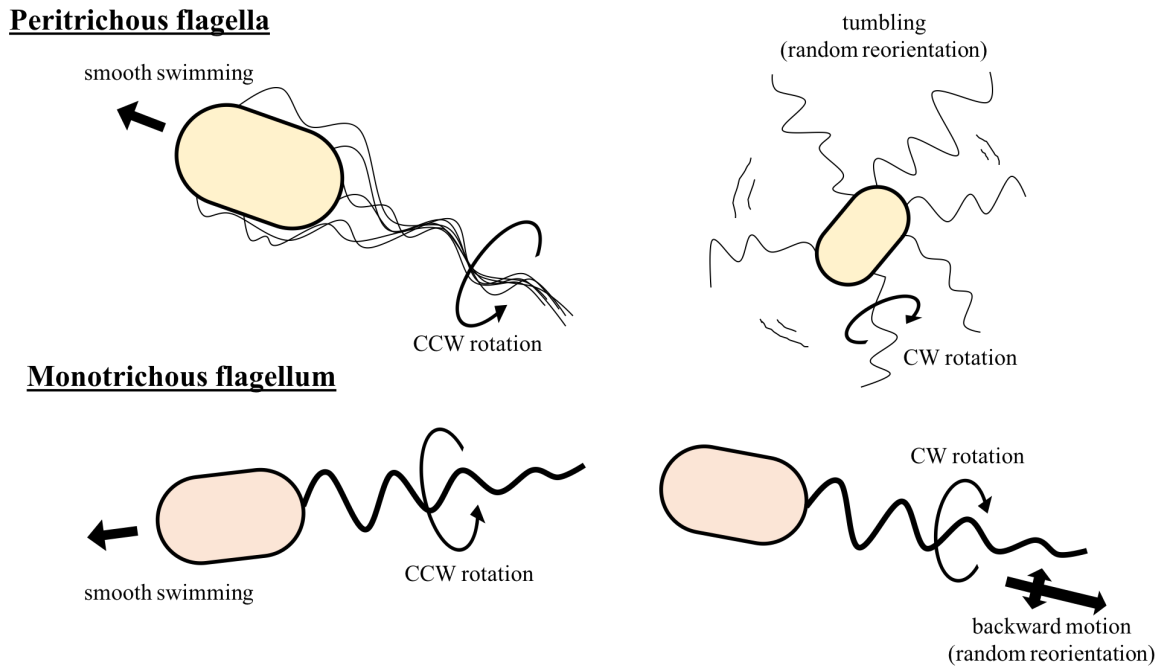
#### 1-1. Bacterial swimming mode

Many bacteria swim in liquid environments by rotating flagella (Fig. 1-1) (1). The flagellum is a fibrous organelle that consists of a long helical filament and a rotary motor located at the root of the filament. The flagellar motor rotates in either clockwise (CW) or counterclockwise (CCW). The CCW rotation mode of the motor results in smooth swimming of bacteria, whereas the CW rotation mode causes tumbling (peritrichous flagella) or backward motion (monotrichous flagella) of bacteria (Fig. 1-2) (2, 3). Both tumbling and backward motion result in random reorientation of bacterial cells (Fig. 1-2) (2, 3). Bacteria swim in solution by alternately repeating these two swimming modes. Many bacteria have only two swimming modes but they handle these two modes well to move to desired place.



**Figure 1-1. Bacterial flagellum**

Electron micrograph of *Salmonella enterica* serovar Typhimurium.



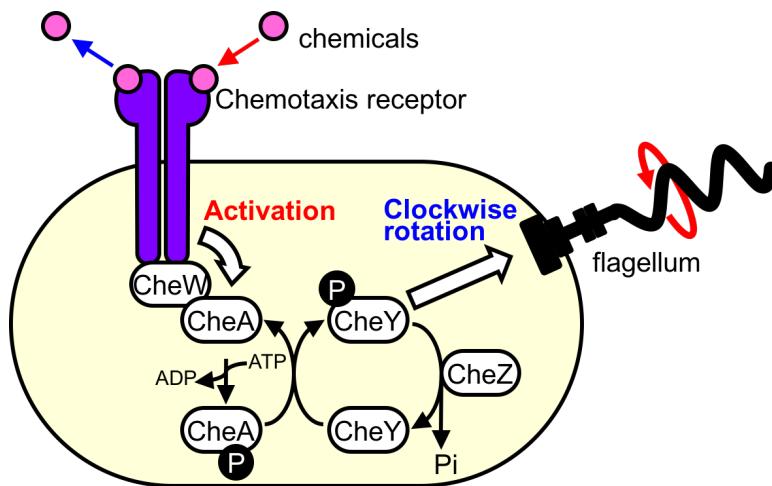
**Figure 1-2. Bacterial swimming**

Swimming modes of bacteria with peritrichous flagella or a monotrichous flagellum. Left/right side indicates the swimming mode of CCW/CW rotation of flagellar motor.

**1-2. Bacterial chemotaxis**

Bacteria randomly move in a uniform environment by randomly repeating straight swimming and reorientation. In the presence of chemical gradients, on the other hand, bacteria move to their favorable environments and run away from toxic matters by sensing the chemical gradients in their surroundings to assure survival. This behavior is called bacterial chemotaxis (4).

Bacterial chemotaxis is regulated by the His-Asp phosphorelay system. Chemoreceptor membrane proteins, chemotaxis signaling proteins (Che proteins) and the flagellar motor switching proteins are involved in the His-Asp phosphorelay signal transduction (Fig. 1-3). In the first step of the signal transduction, the environmental chemicals are sensed by periplasmic domains of chemoreceptor proteins embedded in the bacterial cytoplasmic membrane, and the binding signal is transferred to the cytoplasmic region of the chemoreceptor proteins, the mechanism of which is still unclear. The chemoreceptor proteins activate autophosphorylation of histidine kinase CheA with adenosine triphosphate (ATP). The phosphate group bound to CheA is, then, transferred to the messenger protein, CheY. The phosphorylated CheY binds to the flagellar motor to induce switching the direction of rotation from CCW to CW (Fig. 1-3) (5, 6).



**Figure 1-3. The model of bacterial chemotaxis**

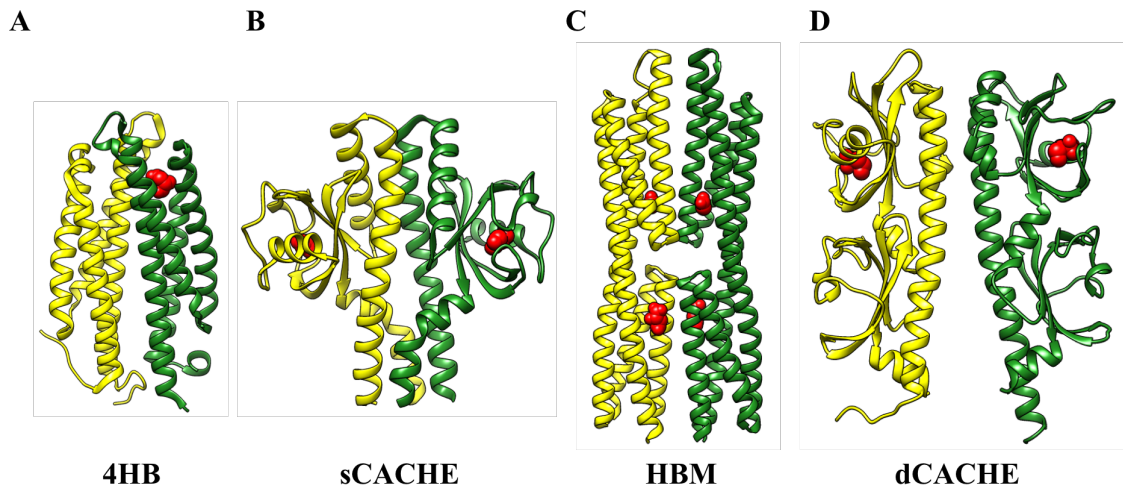
A cartoon model of the bacterial chemotaxis. A bacterium, chemoreceptors and environment chemicals are colored yellow, purple and magenta, respectively. Che proteins are indicated by white ellipses with their names.



### 1-3. Chemoreceptor

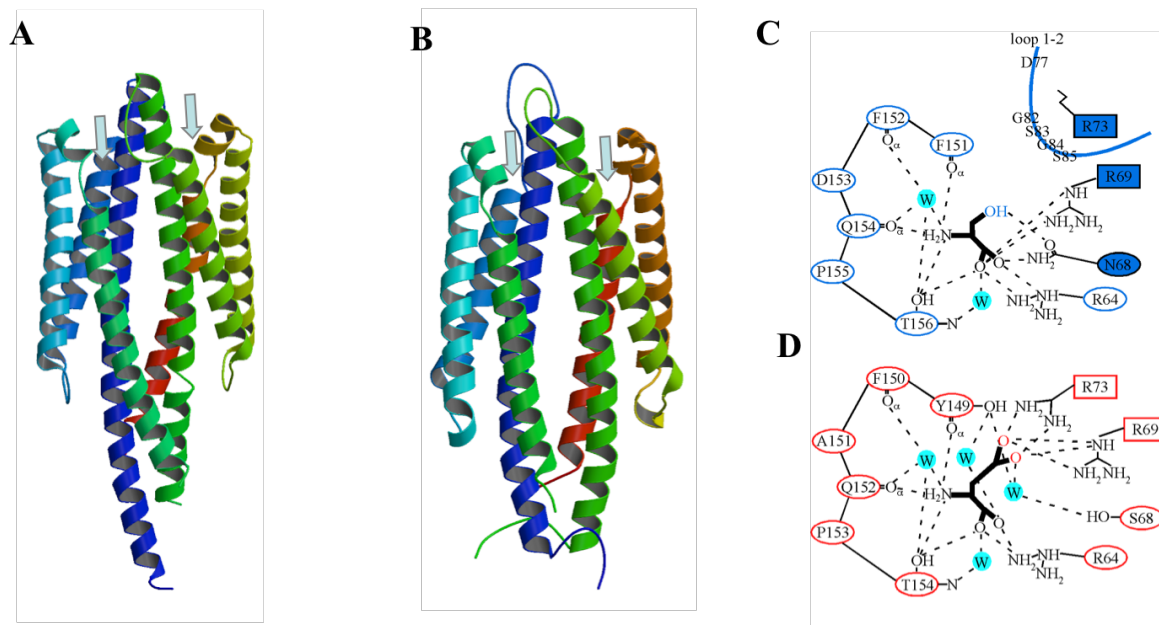
Two-transmembrane chemoreceptors are the most common type of bacterial chemoreceptors. They form dimer and each monomer consists of four regions, a periplasmic sensor domain, a transmembrane domain, a HAMP domain and a kinase control module. Their specific ligands are bound to their periplasmic sensor domains. The two-transmembrane chemoreceptors are classified into four groups by the periplasmic domain structure (7, 8); 4-helix bundle (4HB), helical bimodular (HBM), single CACHE (sCACHE) and double CACHE (dCACHE) (Fig. 1-4) (7). 4HB is the most well-studied chemoreceptor family. Tsr and Tar of *Escherichia coli* belong to this family (Fig. 1-5A and B) (9). HBM is composed of two 4-helix bundle domains tandemly arranged perpendicular to the membrane. McpS of *Pseudomonas putida* is a representative of this group (10). 4HB and HBM binds their specific ligands at the interface of the two monomers (Fig. 1-5C and D). sCACHE type chemoreceptors have a domain analogous to the PAS domain, called the CACHE domain. The CACHE domain is an extracellular ligand recognition domain composed of a five-stranded antiparallel  $\beta$ -sheet and several  $\alpha$ -helices. TlpB of *Helicobacter pylori* is a typical example of the sCACHE family chemoreceptor (11). dCACHE consists of two CACHE domains tandemly arranged perpendicular to the membrane.

Methyl-accepting Chemotaxis Proteins (MCPs) are the most popular type of the bacterial two-transmembrane chemoreceptors. Four chemoreceptors belong to MCPs (Fig. 1-4). The transmembrane and the cytoplasmic region of MCPs share a common structural architecture (Fig. 1-6) (12).



**Figure 1-4. Structures of bacterial two-transmembrane chemoreceptors**

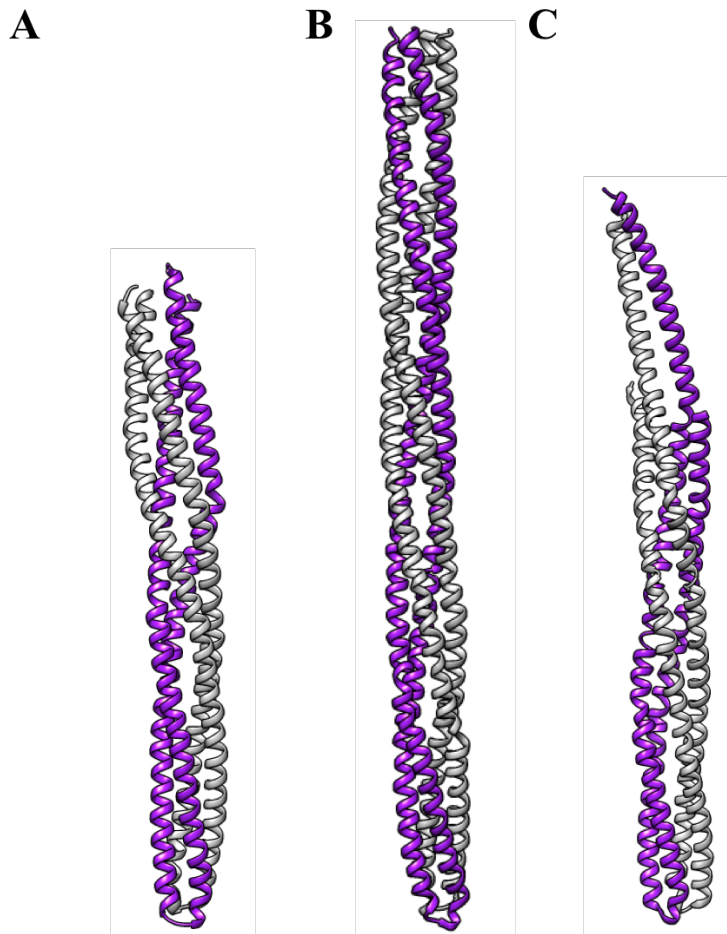
Periplasmic sensor domain structures of the bacterial two-transmembrane chemoreceptors. (A)-(D) shows the representative one of each groups; (A) *Escherichia coli* chemoreceptor Tar, (B) *Helicobacter pylori* chemoreceptor TlpB, (C) *Pseudomonas putida* chemoreceptor McpS and (D) *Vibrio cholerae* chemoreceptor Mlp37. Each monomer shows different colors (yellow and green). Ligands bound to the chemoreceptors are shown in red ball models.



**Figure 1-5. Ligand recognition mechanism of 4HB**

(A), (B) Three dimensional structures of *Escherichia coli* chemoreceptor Tsr (A) and Tar (B). Ligand binding sites are shown by pale blue arrows. (C), (D) Ligand binding mechanisms of Tsr (C) and Tar (D). The ligands (L-serine (C) and L-aspartate (D)) and the peptide chains of proteins are indicated by thick and thin black line. Possible hydrogen bonds are shown by black broken lines. The bound water molecules are shown as w in cyan circles.

(Tajima H *et al.* (2011) *J Biol Chem* **286**:42200-10. Figure 5.)



**Figure 1-6. Structure of the kinase control module of MCP**

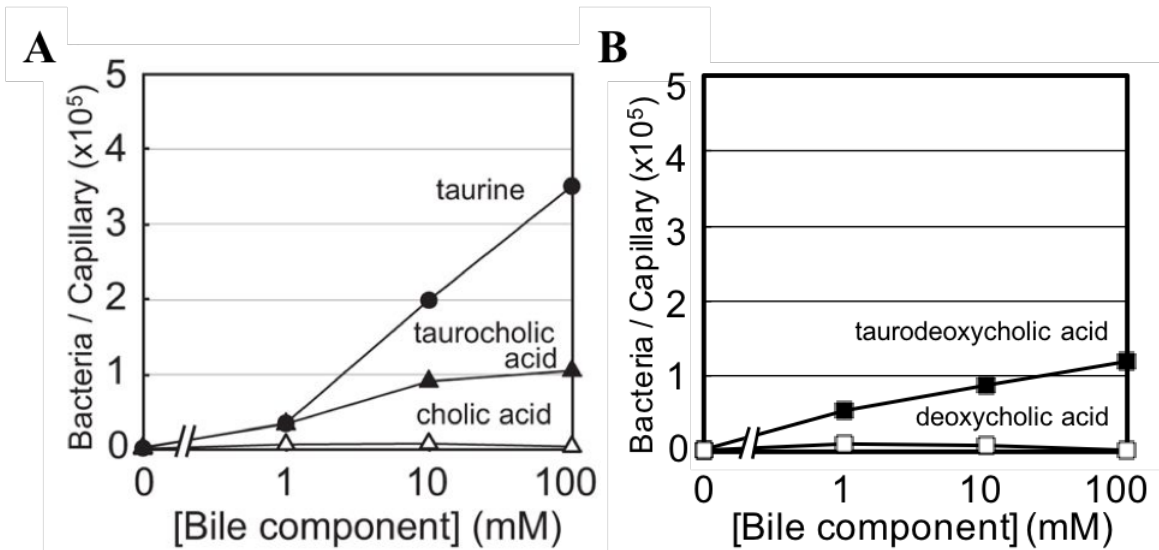
The structures of the kinase control module of *Ter motoga maritima* Tm14 (A), *T. maritima* Tm1143 (B) and *E. coli* Tsr (C). Each chain is colored in purple and gray. All three chemoreceptors belong to MCP.

#### **1-4. *Vibrio cholerae***

Cholera is an infection disease that causes severe diarrhea leading to serious dehydration and has spread all over the world to cause seven pandemics since 1817 (13). The infection brings high fatality rate if extraordinarily rapid treatment is not performed to the patient. In recent years, Cholera is rare in advanced countries because of development of public health and advanced water supply. However, cholera remains a severe disease especially in developing countries because health care is limited, sanitation is poor and water is not safe.

*Vibrio cholerae* is the etiological agent of cholera. *V. cholerae*, is a gram-negative bacterium, which has a single polar flagellum. The bacterium inhabits nutrient-poor aquatic environments like rivers, and occasionally intrudes into the lumen of the human gastrointestinal tract (14). *V. cholerae* produces toxins only in nutrient-rich environments. Therefore chemotaxis of *V. cholerae* to nutrients is critical for bacterial survival and pathogenicity (15-27).

Recent studies have revealed that *V. cholerae* is attracted by various L-amino acids, mucus, galactose-6-sulfate and several constituent of bile, such as taurine (2-aminoethylsulfonate), taurocholic acid and taurodeoxycholic acid, whereas other ingredients of bile such as cholic acid or deoxycholic acid do not attract to *V. cholerae* (Fig. 1-7) (28, 32, 33). These chemotactic responses are mediated by chemoreceptors of *V. cholerae*.



**Figure 1-7. Chemotactic responses of *V. cholerae* to bile components**

(A), (B) The cells of classical biotype strain of *V. cholerae* were subjected to capillary assays with several components of bile. Taurine elicited a strong attractant response (A). Taurocholic acid (A) and taurodeoxycholic acid (B) also served as weak attractants, whereas cholic acid (A) and deoxycholic acid (B) did not.

(Nishiyama S *et al.* (2016) *Sci Rep* 6:20866. Figure 1.)

### **1-5. Chemoreceptors of *Vibrio cholerae***

*V. cholerae* has at least 44 proteins homologous to MCP-Like Proteins (MLPs) (28). Among the MLPs, Mlp7, Mlp8, Mlp24 and Mlp30 are found to play a role in pathogenicity and Mlp37 is involved in biofilm formation (27-31), whereas the function of the other MLPs are unclear. In addition, Mlp8, Mlp24 and Mlp37 are involved in *V. cholerae* chemotaxis (28, 32, 33). These three MLPs are amino acid sequence homologous and they are predicted to belong to the dCACHE family of chemoreceptors. However, they recognize different ligands; Mlp24 and Mlp37 are involved in chemotaxis for various L-amino acids but only Mlp37 recognizes taurine besides the various L-amino acids (28, 32). On the other hand, Mlp8 is involved in chemotaxis for mucus and galactose-6-sulfate (33).

## **1-6. Outline of this study**

To reveal the molecular mechanisms of ligand recognition by the three homologous chemoreceptors (Mlp8, Mlp24 and Mlp37) of *V. cholerae*, I have determined the structures of Mlp24 and Mlp37 in complex with their ligands. In addition, I have determined the structure of AcfC, which is a periplasmic binding protein involved in ligand recognition of Mlp8. I also carried out isothermal titration calorimetry (ITC) and pull-down assay to reveal the ligand recognition mechanism of the Mlp8/AcfC system.

In chapter 2, I describe the ligand recognition mechanism of Mlp37. Three ligand binding structures of the periplasmic domain of Mlp37 (Mlp37p) were determined. The structures of Mlp37p ligand complexes are very similar but the conformations of the ligand recognition domain are different from each other. Mlp37p adjusts the shape and size of the ligand recognition domain to the ligands so that Mlp37p recognizes various ligands with similar dissociation constants.

In chapter 3, molecular mechanism of ligand recognition of Mlp24 is shown. Five ligand complex structures of the periplasmic domain of Mlp24 (Mlp24p) were determined as well as the structure of ligand-free Mlp24p. I discuss differences in ligand recognition between Mlp24 and Mlp37. Comparison between the ligand-free and the ligand-bound Mlp24p structures suggest that the ligand binding to Mlp24 induces the ligand binding domain (CAD1) closure. A calcium ion binds to the CAD1 and the calcium ion is important for the ligand recognition of Mlp24.



In chapter 4, the ligand recognition mechanism of Mlp8 is described. A crystal structure of ligand-free AcfC, which is involved in the chemotactic response and/or the pathogenicity of *V. cholera* with Mlp8 was determined. The structural comparison between ligand-free AcfC and D-malate complexed AcfC indicates that D-malate binding to AcfC does not induce structural change. In addition, ITC and pull-down assays suggest that Mlp8 recognizes the ligands with, at least, two different mechanisms.

## 1-7. References

1. Berg HC, Anderson RA. 1973. Bacteria swim by rotating their flagellar filaments. *Nature* 245(5425):380-2.
2. Berg HC, Brown DA. 1972. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature* 239(5374):500-4.
3. Larsen SH, Reader RW, Kort EN, Tso WW, Alder J. 1974. Change in direction of flagellar rotation is the basis of the chemotactic response in *Escherichia coli*. *Nature* 249(452):74-7.
4. Parkinson JS, Hazelbauer GL, Falke JJ. 2015. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. *Trends Microbiol* 23:257-66.
5. Sourjik V. 2004. Receptor clustering and signal processing in *E. coli* chemotaxis. *Trends in Microbiol* 12(12):569-76.
6. Cluzel P, Surette M, Leibler S. 2000. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science* 287(5458):1652-5.
7. Matilla MA, Krell T. 2017. Chemoreceptor-based signal sensing. *Curr Opin Biotechnol* 45:8-14.
8. Upadhyay AA, Fleetwood AD, Adebali O, Finn RD, Zhulin IB. 2016. Cache Domains That are Homologous to, but Different from PAS Domains Comprise the Largest Superfamily of Extracellular Sensors in Prokaryotes. *PLoS Comput Biol* 12:e1004862.
9. Tajima H, Imada K, Sakuma M, Hattori F, Nara T, Kamo N, Homma M, Kawagishi I. 2011. Ligand specificity determined by differentially arranged common ligand-binding

- residues in bacterial amino acid chemoreceptors Tsr and Tar. *J Biol Chem* 286:42200-10.
10. Pineda-Molina E, Reyes-Darias JA, Lacal J, Ramos JL, Garcia-Ruiz JM, Gavira JA, Krell T. 2012. Evidence for chemoreceptors with bimodular ligand-binding regions harboring two signal-binding sites. *Proc Natl Acad Sci U S A* 109:18926-31.
  11. Emily GS, J. Nathan H, John G, Christopher W, Kevin GH, Jeneva KF, Raghuveer P, S. James R, Karen G. 2012. Structure and Proposed Mechanism for the pH-Sensing *Helicobacter pylori* Chemoreceptor TlpB. *Structure* 20:1-12.
  12. Abiola MP, Alecandrine MB, Brain RC. 2009. The structure of a Soluble Chemoreceptor Suggests a Mechanism for Propagating Conformational Signals. *Biochem* 48:1936-44.
  13. Wachsmuth K., Olsvik O, Evins GM, Popovic T. 1994. in *Vibrio cholerae* and Cholera: Molecular to global perspective 357-70 (ASM Press, Washington DC).
  14. Reidl J, Klose KE. 2002. *Vibrio cholerae* and cholera: out of the water and into the host. *FEMS Microbiol. Rev.* 26:125-39.
  15. Boin MA, Austin MJ, Hase CC. 2004. Chemotaxis in *Vibrio cholerae*. *FEMS Microbiol. Lett.* 239:1-8.
  16. Alm RA, Manning PA. 1990. Characterization of the *hylB* gene and its role in the production of the El Tor haemolysin of *Vibrio cholerae* O1. *Mol. Microbiol.* 4:413-25.
  17. Banerjee R, et al. 2002. Involvement of *in vivo* induced *cheY-4* gene of *Vibrio cholerae* in motility, early adherence to intestinal epithelial cells and regulation of virulence factors. *FEBS Lett.* 532:221-26.

18. Boin MA, Austin MJ, Hase CC. 2004. Chemotaxis in *Vibrio cholerae*. FEMS Microbiol. Lett. 239:1–8.
19. Everiss KD, Hughes KJ, Kovach ME, Peterson KM. 1994. The *Vibrio cholerae acfB* colonization determinant encodes an inner membrane protein that is related to a family of signal-transducing proteins. Infect. Immun. 62:3289–3298.
20. Everiss KD, Hughes KJ, Peterson KM. 1994. The accessory colonization factor and toxin-coregulated pilus gene clusters are physically linked on the *Vibrio cholerae* O395 chromosome. DNA Seq. 5:51–55.
21. Freter R, Allweiss B, O'Brien PC, Halstead SA, Macsai MS. 1981. Role of chemotaxis in the association of motile bacteria with intestinal mucosa: in vitro studies. Infect. Immun. 34:241–249.
22. Freter R, O'Brien PC. 1981. Role of chemotaxis in the association of motile bacteria with intestinal mucosa: chemotactic responses of *Vibrio cholerae* and description of motile nonchemotactic mutants. Infect. Immun. 34:215–221.
23. Freter R, O'Brien PC, Macsai MS. 1981. Role of chemotaxis in the association of motile bacteria with intestinal mucosa: in vivo studies. Infect. Immun. 34:234–240.
24. Gupta S, Chowdhury R. 1997. Bile affects production of virulence factors and motility of *Vibrio cholerae*. Infect. Immun. 65:1131–1134.
25. Hang L, et al. 2003. Use of in vivo-induced antigen technology (IVIAT) to identify genes uniquely expressed during human infection with *Vibrio cholerae*. Proc. Natl. Acad. Sci. U. S. A. 100:8508–8513.

26. Krukonis ES, DiRita VJ. 2003. From motility to virulence: sensing and responding to environmental signals in *Vibrio cholerae*. *Curr. Opin. Microbiol.* **6**:186–190.
27. Lee SH, Butler SM, Camilli A. 2001. Selection for *in vivo* regulators of bacterial virulence. *Proc Natl Acad Sci U S A* **98**(12): 6889-94.
28. Nishiyama S, Suzuki D, Itoh Y, Suzuki K, Tajima H, Hyakutake A, Homma M, Butler-Wu SM, Camilli A, Kawagishi I. 2012. Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infect Immun* **80**:3170-8.
29. Chaparro AP, Ali SK, Klose KE. 2010. The ToxT-dependent methyl-accepting chemoreceptors AcfB and TcpL contribute to *Vibrio cholerae* intestinal colonization. *FEMS Microbiol Lett* **302**(2):99-105.
30. Everiss KD, Hughes KJ, Kovach ME, Peterson KM. 1994. The *Vibrio cholerae acfB* colonization determinant encodes an inner membrane protein that is related to a family of signal-transducing proteins. *Infect Immun* **62**:3289-3298.
31. Alm RA, Manning PA. 1990. Characterization of the hlyB gene and its role in the production of the El Tor haemolysin of *Vibrio cholerae* O1. *Mol. Microbiol.* **4**(3):413-25.
32. Nishiyama S, Takahashi Y, Yamamoto K, Suzuki D, Itoh Y, Sumita K, Uchida Y, Homma M, Imada K, Kawagishi I. 2016. Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci Rep* **6**:20866.

33. Selvaral P, Gupta R, Peterson KM. 2015. The *Vibrio cholerae* ToxR regulon encodes Host-Specific chemotaxis proteins that function in intestinal colonization. SOJ Microbiol Infect Dis. 3(3).

## Chapter 2

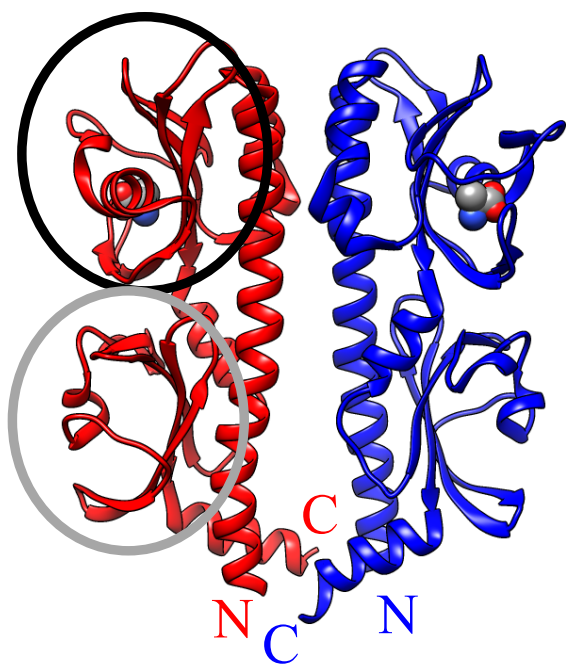
### Molecular mechanism of ligand recognition by Mlp37

#### 2-1. Introduction

Mlp37 mediates attractant responses to various L-amino acids and taurine, which is a major constituent of bile. Chemotactic responses of the  $\Delta mlp37$  mutant to L-serine, glycine, L-alanine, L-cysteine, L-arginine, L-asparagine, L-threonine, L-lysine and L-valine were considerably decreased (1). Isothermal titration calorimetry (ITC) with a periplasmic fragment of Mlp37 (residues 58-303; named Mlp37p) has indicated that the attractants directly bind to Mlp37. L-serine, L-alanine, L-arginine and taurine bind to Mlp37p with  $K_d$  values of 3.6  $\mu$ M, 2.7  $\mu$ M, 5.6  $\mu$ M and 3.2  $\mu$ M, respectively (1). In contrast, L-glutamate, which was not sensed by Mlp37 by capillary assay, does not appreciably bind to Mlp37p. These results demonstrated that Mlp37 is a chemoreceptor for taurine, L-serine, L-alanine and L-arginine (1).

The structure of Mlp37p (VCA0923) of an El Tor strain (PDB ID: 3C8C) has been determined by structural genomics group. Mlp37p forms a homodimer and each subunit has two CACHE domains (the membrane distal and proximal ones are hereafter referred to as CAD1 and CAD2, respectively), confirming that it is a dCACHE-type chemoreceptor (Fig. 2-1) (2). The 3C8C structure was solved as an L-alanine complex, in which alanine was bound to CAD1. Among the known ligands for Mlp37, only taurine is not an L-amino acid. Although the structure of taurine is quite distinct from the other ligands, its dissociation constant is comparable to L-arginine, L-serine and L-alanine. To elucidate the ligand recognition

mechanism of Mlp37, I solved the crystal structures of Mlp37p in complex with L-arginine, L-serine and taurine.



**Figure 2-1. Overall structure of Mlp37p L-alanine complex**

Ribbon representation of the structure of Mlp37p L-alanine complex (PDB ID: 3C8C). Each subunit is painted in different color (blue and red). Two CACHE domains, CAD1 and 2, are indicated with black and gray circles, respectively. The ligand L-alanine is shown as a ball model.



## **2-2. Materials and Methods**

### **2-2-1. Expression and purification of the periplasmic fragment of Mlp37**

Strain BL21(DE3) (Novagen) carrying pVCP37, which expresses Mlp37p with an N-terminal GST tag and a C-terminal hexahistidine tag, was cultured in LB broth Lennox (Nacalai Tesque, Inc., Kyoto Japan) containing 50 µg/ml of ampicillin at 37°C until the cell density had reached an OD<sub>600</sub> of about 0.8. Isopropyl β-d-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.1 mM to induce protein expression, and the culture was continued for 12 hours at 20°C. Cells were harvested by centrifugation, suspended in phosphate buffered saline (PBS) (137 mM of NaCl, 2.7 mM of KCl, 10 mM of Na<sub>2</sub>HPO<sub>4</sub>•12H<sub>2</sub>O, and 1.8 mM of KH<sub>2</sub>PO<sub>4</sub> at pH 7.4) and disrupted by sonication. After removing cell debris by centrifugation, the cell lysate was loaded on a Glutathione-Sepharose™ 4B column (GE Healthcare) followed by washing with PBS. Proteins were eluted with 50 mM Tris-HCl (pH 8.0) buffer containing 10 mM reduced glutathione. The N-terminal GST tag was then cleaved using PreScission™ Protease (GE Healthcare), and the reactant was dialysed against 1.0 L dialysis buffer (50 mM Tris-HCl pH 7.0, 150 mM NaCl, 1 mM DTT, and 1 mM EDTA) at 4°C for 12 hours using Spectra/por Dialysis Membrane MWCO 6,000-8,000 (Spectrum Laboratories, Inc.). The protein solution was loaded again on a Glutathione-Sepharose™ 4B column (GE Healthcare) to remove GST and unreacted proteins, and further purified by size exclusion chromatography with a High Load 26/60 Superdex 200 column in 20 mM Tris-HCl and 150 mM NaCl (pH 8.0). The peak fraction was collected and concentrated to 10 mg/mL. The purity of the purified proteins was examined by SDS-PAGE.

## 2-2-2. Crystallization

Initial crystallization screening was performed by sitting-drop vapor-diffusion technique. Crystallization drops were prepared by mixing 1  $\mu$ l of protein solution (10 mg/ml) with 1  $\mu$ l of reservoir solution. The ligand amino acids and taurine were added in the protein solution before preparing the drop to the final concentration of 10 mM. Initial screening was carried out using the following screening kit; Wizard Classic I and II (Rigaku Reagents, Inc.), Wizard Cryo I and II (Rigaku Reagents, Inc.), and Crystal Screen I and II (Hampton Research). The crystallization conditions were optimized by the hanging-drop vapor-diffusion method. Crystallization drops were prepared by mixing 1.5  $\mu$ l of protein solution with an equal volume of reservoir solution at 293K. All crystals appeared within 3 days after optimization. Their final crystallization conditions, space groups and unit cell dimensions are summarized in Table 2-1 and 2-2.

**Table 2-1 Summary of crystallization conditions of Mlp37p**

Ligands	Buffer	Precipitant	Additive
10 mM L-arginine	100 mM HEPES pH 7.5	2.0 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1.2% (v/v) PEG-400
10 mM L-serine	100 mM Phosphate-citrate pH 4.2	10 % (w/v) PEG-8,000	200 mM NaCl
10 mM taurine	100 mM MES pH 6.5	15 % (w/v) PEG-8,000	30 % (v/v) MPD

**Table 2-2 Summary of X-ray data collection and refinement statistics of Mlp37p**

Ligands	L-arginine	L-serine	taurine
<b>Data collection</b>			
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>C</i> 222 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions			
<i>a</i> (Å)	49.6	98.7	33.2
<i>b</i> (Å)	100.5	134.2	119.1
<i>c</i> (Å)	106.8	50.5	138.8
β (°)	90.0	90	90
Wavelength (Å)	1.0000	1.0000	1.0000
Resolution (Å)	47.2-2.35 (2.43-2.35)	32.8-1.80 (1.90-1.80)	33.3-1.95 (2.06-1.95)
<i>R</i> <sub>merge</sub> (%)	9.6 (42.9)	9.7 (41.1)	9.6 (37.1)
<i>I</i> /σ	14.2 (4.2)	11.4 (4.7)	13.7 (5.2)
Completeness (%)	98.8 (97.3)	99.5 (99.2)	99.9 (99.8)
Redundancy	6.3 (6.3)	7.2 (7.3)	7.2 (7.0)
Content per ASU	2	1	2
<b>Refinement statistics</b>			
Resolution (Å)	47.2-2.35 (2.41-2.35)	31.7-1.80 (1.86-1.80)	32.3-1.95 (2.00-1.95)
Number of reflections			
Working set	22,568 (1,391)	31,367 (2,650)	40,573 (2,584)
Test set	1,128 (133)	1,583 (160)	2,000 (134)
<i>R</i> <sub>w</sub> (%)	18.52 (20.91)	18.7 (25.8)	19.0 (22.2)
<i>R</i> <sub>free</sub> (%)	24.42 (26.82)	21.4 (31.5)	23.7 (31.2)
R.m. s. deviations			
Bond length (Å)	0.009	0.006	0.008
Bond angle (°)	1.08	1.01	1.02
<i>B</i> -factors			
Protein atoms	30.5	26.5	36.3
Ligand atoms	26.2	27.0	36.5
Solvent atoms	30.1	32.4	39.0
Ramachandran plot (%)			
In preferred regions	98.2	95.9	92.9
In allowed regions	1.8	4.1	7.1
Outliers	0.0	0.0	0.0
Determined regions (amino acid number and chain ID)	65-288 (A) 69-291 (B)	57-308	53-296 (A) 53-296 (B)
Number of atoms			
Protein	3,434	1,974	3,825
Ligand	24	7	14
Solvent	117	200	396

Values in parentheses are for the highest resolution shell.

$$R_w = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}, R_{free} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$$

### **2-2-3. X-ray data collection and structure determination**

X-ray diffraction data were collected at synchrotron beamline BL41XU in SPring-8 (Harima, Japan), with the approval of the Japan Synchrotron Radiation Research Institute (JASRI). The crystals were soaked in the 1 : 9 mixture of 2-Methyl-2, 4-pentanediol (MPD) and the reservoir solution for a few seconds before transfer into liquid nitrogen for freezing. All X-ray diffraction data were measured at 95K under nitrogen gas flow. The diffraction data were processed with MOSFLM (3) or XDS (4) and scaled with SCALA (5), Aimless (5) or XDS (4). The initial phases were determined by molecular replacement (MR) with Phenix (6). Molecular replacement was performed using the Mlp37p L-alanine complex structure (PDB ID: 3C8C). The atomic models were constructed with Coot (7) and refined with Phenix (6). Data collection and refinement statistics are summarized in Table 2-1 and 2-2.

## 2-3. Results

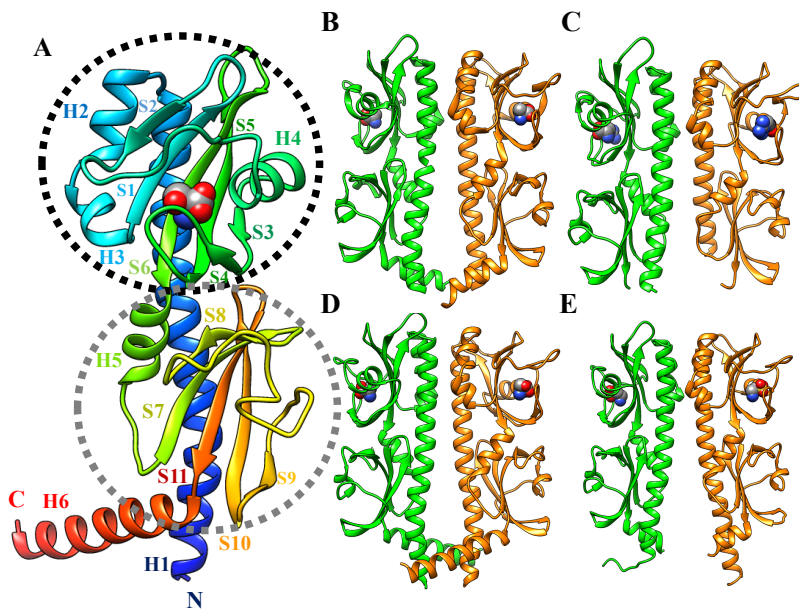
### 2-3-1. Overall structures of Mlp37p in complex with its ligand

The structures of Mlp37p in complex with L-arginine, L-serine and taurine were determined at 1.8-2.4 Å resolution (Fig. 2-2, Table 2-1 and 2-2). These ligand complex crystals were prepared by co-crystallization and grew in different space groups and unit cell parameters in distinct conditions. The molecular arrangements in these crystals, therefore, differ from one another (Table 2-2).

The tertiary and quaternary structures of Mlp37p are very similar to that of Mlp37p L-alanine complex (Fig. 2-1). Each Mlp37p subunit has a long N-terminal  $\alpha$ -helix (H1), two CACHE domains tandemly arranged along H1 and a C-terminal helix (H6), two CACHE domains tandemly arranged along H1 and a C-terminal helix (H6) (Fig. 2-2A). CAD1, the membrane-distal domain, consists of three  $\alpha$ -helices (H2-H4) and six  $\beta$ -strands (S1-S6), five of which (S1-S3, S5 and S6) form a central  $\beta$ -sheet (Fig. 2-2A). The domain has a mouth-like pocket with the upper jaw formed by the loop of 16 residues connecting S2 and H4 (UJ-loop) and the lower jaw formed by S4, the N-terminal half of S5 and the loop connecting S4 and S5 (LJ-segment). CAD2, the membrane-proximal domain, comprises an  $\alpha$ -helix (H5) and a  $\beta$ -sheet composed of five  $\beta$ -strands (S7-S11). Although CAD2 shows a similar fold to CAD1, all ligands bind only to CAD1 (Fig. 2-2B-E), like other known dCACHE-type chemoreceptors except for Tlp1 and TlpC (8-12).

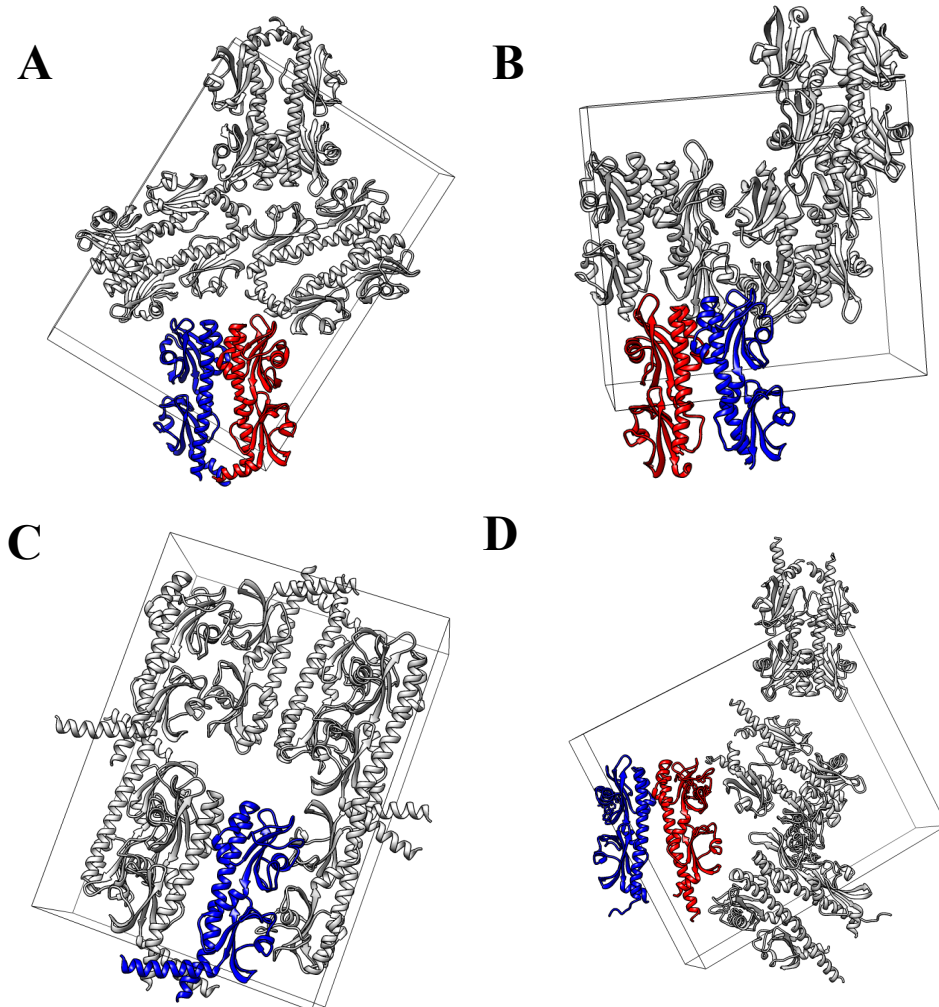
The L-arginine complex and the taurine complex forms a dimer with a neighbouring molecule in the asymmetric unit. In contrast, the L-serine complex forms a dimer with a molecule related by crystallographic two-fold symmetry (Fig. 2-3). In spite of these

differences in molecular packing, the dimer structures are basically identical (Fig. 2-2). H2 and the C-terminal region of H1 stabilize the dimer by forming an up-and-down four-helix bundle structure with its dimer mate.



**Figure 2-2. Overall structures of Mlp37p**

(A) Ribbon representation of the Mlp37p L-serine complex is colored in rainbow from the N-terminus (blue) to the C-terminus (red). The membrane distal and proximal CACHE domain (CAD1 and CAD2) are indicated with black and gray broken circles, respectively. The secondary structure elements are labeled. The ligand L-serine is shown as a ball model. (B-E) The dimeric structures of Mlp37p L-alanine complex (PDB ID: 3C8C) (B), the L-arginine complex (C), the L-serine complex (D), the taurine complex (E) with ribbon model. The ligand molecules are shown as ball models. Each subunit is painted in different color (green and orange).



**Figure 2-3. Differences in molecular packing of the Mlp37p ligand complexes**

Molecular packing of the crystals of Mlp37p L-alanine complex (A), L-arginine complex (B), L-serine complex (C) and taurine complex (D). Molecules in an asymmetric unit is colored in red and blue and other symmetry related molecules are in gray. The unit cells are shown with gray line.

### **2-3-2. Ligand recognition mechanism of Mlp37p**

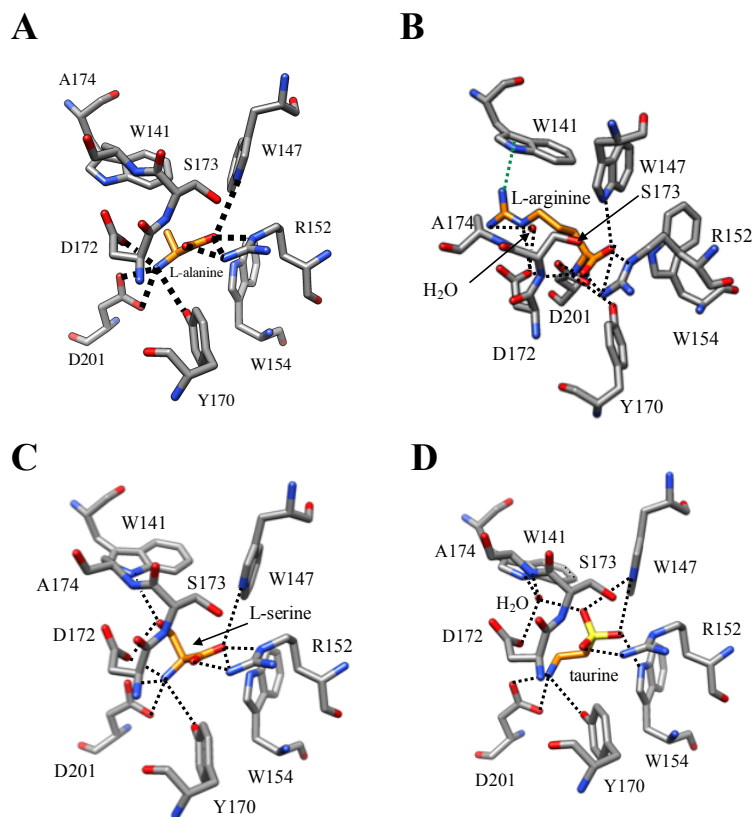
All ligands of Mlp37p are bound only in CAD1 in the crystal structures (Fig. 2-2B-E). The amino acid backbone atoms of the ligand L-arginine and L-serine are located at essentially the same positions as those of the bound L-alanine molecule in the 3C8C structure (Fig. 2-4A-C). The amino group of L-arginine and L-serine make hydrogen bonds with Y170, D172 and D201, which are triangularly arranged around it. The carboxy group of L-arginine and L-serine mainly hydrogen bonds with the guanidino group of R152. The indole NH of W147 and W154 also hydrogen bonds with the carboxy group. These interactions are exactly conserved in the L-alanine complex (Fig. 2-4A-C). The side chain guanidino group interacts with D172 and makes cation-pi interaction with the indole ring of W141 (Fig. 2-4B). The side chain hydroxy group of L-serine hydrogen bonds with D172 and the indole NH of W141 (Fig. 2-4C).

The binding manner of taurine is similar to that of L-amino acid, especially very close to L-serine (Fig. 2-4D). The amino group of taurine lies at the equivalent position of that of L-amino acids, and interacts with Y170, D201 and D172. Two oxygen atoms of the sulfonate group locate in the corresponding positions to the carboxy oxygen atoms of L-amino acids, but are out from the plane formed by the guanidino group of R152. Thus the indole NH of W147 seems to also contribute to the interaction in addition to the guanidino group of R152, the indole NH of W154 and the main chain amino group of S173. The remaining oxygen of the sulfonate is present at the position near the hydroxy oxygen of L-serine, and interacts with the indole NH of W147 and a water molecule bound to D172, the indole NH of W141, and the main chain amino group of A174. These structural features are essentially the same in the



two subunits of the Mlp37p taurine complex dimer (Fig. 2-5B). The residues in contact with taurine can be superimposed on to the corresponding residues of the L-serine complex (Fig. 2-5A) with the root mean square deviations of about 0.3 Å (0.314 for A-subunit and 0.265 for B-subunit), indicating that taurine well mimic L-serine in Mlp37.

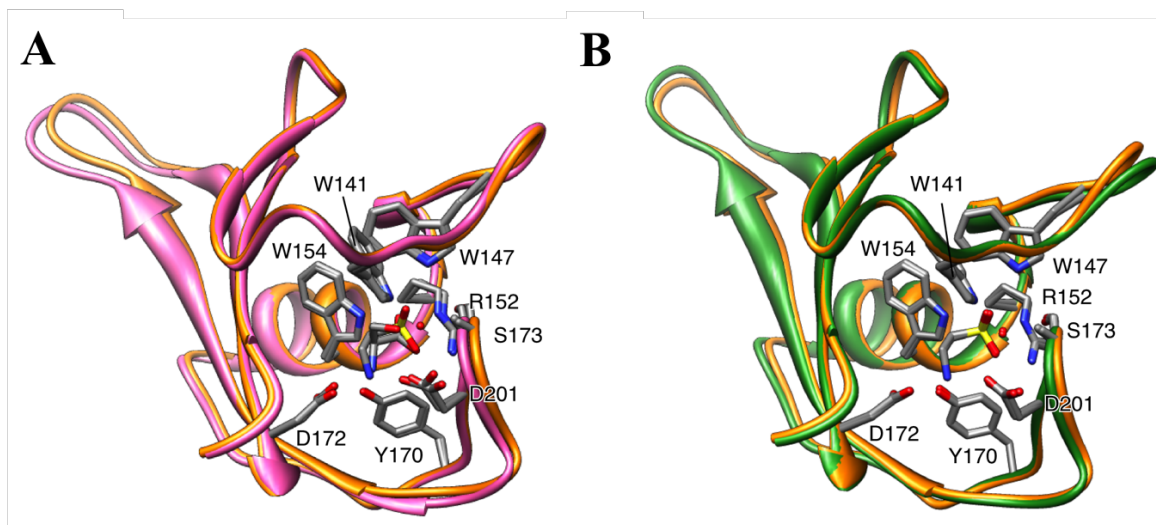
CAD1 of one subunit in the L-alanine complex dimer shows a fully closed conformation embracing the ligand (Fig. 2-6). By contrast, that of the other subunit has opening almost the same as the L-serine and the taurine complexes. The distance between the indole nitrogen of W141 and one of the carboxy oxygen atoms of D172 in the L-alanine complex is 3.2 Å, indicating that W141 in the UJ loop interacts directly with D172 in the LJ segment. In contrast, these residues in the other complexes are much further apart because W141 and D172 interact with the side chain guanidino group in the L-arginine complex, the hydroxy group in the L-serine complex and with the water molecule in the taurine complex. As a result, the entrances of CAD1 in these latter complexes are open. Such an opening would allow larger R groups of other amino acids to stick out from the pocket without much affecting the interactions of the amino acid backbone atoms, providing a basis for versatile sensing of amino acid attractants. The opening, however, is not wide enough to allow passage of an attractant molecule itself. Thus, ligand binding must involve some open-to-closed-type conformational change in CAD1.



**Figure 2-4. Ligand binding mechanism of Mlp37p**

Close up view of the ligand-binding pockets of Mlp37p in complex with L-alanine (A), L-arginine (B), L-serine (C) and taurine (D). The ligand molecules are colored in orange. The protein residues are represented by gray stick models. The nitrogen, oxygen and sulfur atoms are colored in blue, red and yellow, respectively. The bound water molecules are shown as a red ball. Possible hydrogen bonds are indicated by black broken lines. The cation pi interaction is represented by green broken line.

(Nishiyama S *et al.* (2016) *Sci Rep* 6:20866. Figure 4c-e.)

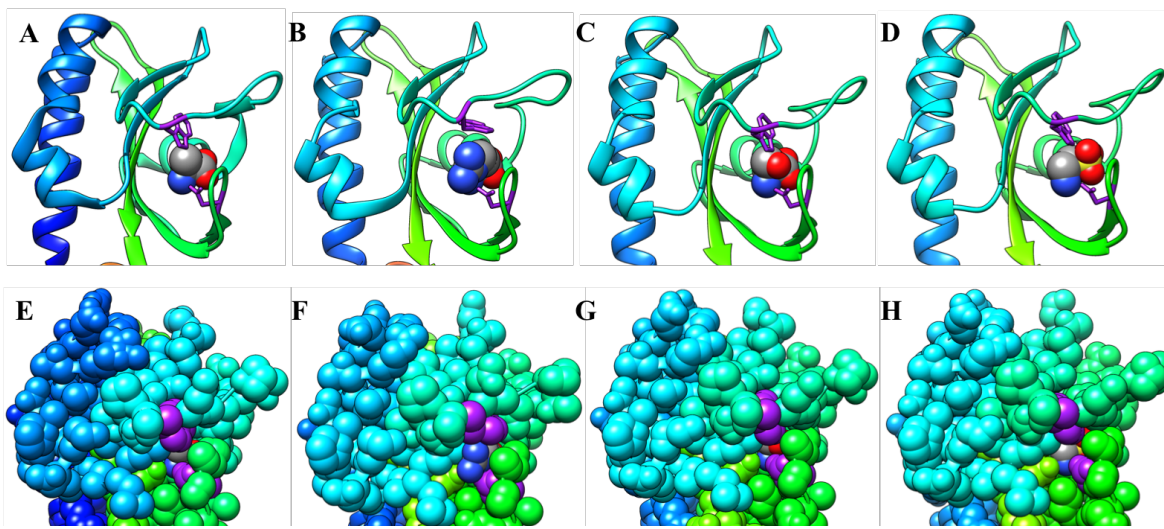


**Figure 2-5. Structural comparison of Mlp37p CAD1**

(A) CAD1 of the L-serine complex (pink) is superimposed onto that of the taurine complex subunit A (orange). The residues involved in the ligand recognition are shown in stick model.

(B) The CAD1 structures of the taurine complex subunit A (orange) and B (green) are superimposed. No significant structural difference is observed in the two subunits.

(Nishiyama S *et al.* (2016) *Sci Rep* 6:20866. Figure S5f, g.)



**Figure 2-6. Entrance of the ligand binding pocket of CAD1**

(A-D) The CAD1 structures of the L-alanine (A), the L-arginine (B), the L-serine(C) and the taurine complexes (D) are drawn in ribbon representation. The bound ligand molecules are shown as ball models, and W141 and D172 are as stick models. (E-H) The CAD1 structures of the L-alanine (E), the L-arginine (F), the L-serine (G), the taurine complex (H) are shown as ball models painted with the same color as (A-D). The CAD1 entrance of the L-arginine complex is widely open and the ligand atoms can be seen (B and F), whereas that of the L-alanine complex is closed (A and E). The CAD1 entrances of the L-serine and the taurine complexes are slightly open (C, D, G and H).

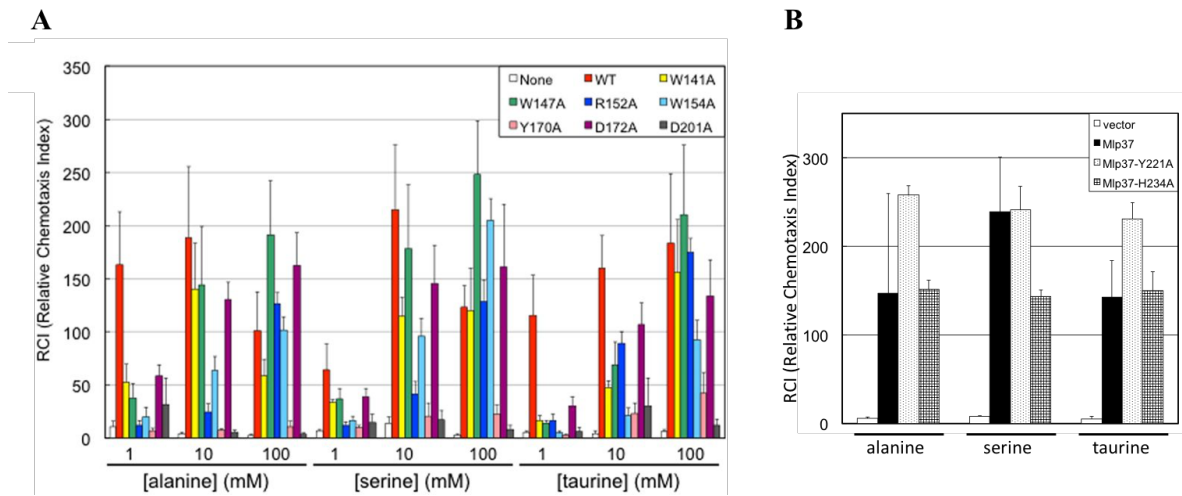
## **2-4. Discussion**

### **2-4-1. Residues involved in ligand binding of Mlp37p**

The crystal structures of Mlp37p suggest that the side chains of W141, W147, R152, W154, Y170, D172, S173 and D201 in CAD1 are important for the ligand recognition. The significance of these residues for ligand binding was confirmed by mutational analysis performed by Prof. Kawagishi's group. The *in vivo* roles of the residues in contact with the ligands (W141, W147, R152, W154, Y170, D172 or D201) were examined by mutational analysis. Mutant receptors with alanine substitution for each residue were expressed in strain Vmlp201, which lacks *mlp24* and *mlp37*, and therefore shows only residual tactic responses to amino acids. S173 was not tested because ligand recognition involves its main chain. Although expression levels of all the mutant proteins were essentially the same as that of the wild-type version, all showed defects in mediating taxis toward L-alanine, L-serine and taurine (Fig. 2-7A). The most profound effects were observed with the lowest attractant concentration (1 mM in capillary) tested (Fig. 2-7A). Y170A and D201A show severe effects on chemotaxis even at higher attractant concentrations (10 or 100 mM in capillary), suggesting that recognition of the amino group by these residues are essential for the ligand binding. It should be noted that bacteria respond to an attractant diffused out from the mouth of the capillary and therefore the attractant concentrations bacteria would encounter are much lower than those in capillary. Other mutants showed nearly wild-type performance at high attractant concentrations (10 or 100 mM in capillary), implying that these mutants have reduced affinities for the attractants. Each mutation had similar effects on responses to the

three attractants tested as expected from the crystal structures. These results reinforce the involvement of these residues in recognition of the attractants, L-alanine, L-serine and taurine.

The crystal structures suggest that CAD2 is not involved in attractant sensing as was shown for Mlp24 (1) and *B. subtilis* chemoreceptors (10) (13). To confirm this hypothesis, they substituted alanine for two of the potential ligand-binding residues in CAD2, Y221A and H234A that are well conserved among the related amino acid chemoreceptors of different species. In contrast to the CAD1 mutations, the CAD2 mutations had milder or no effects (Fig. 2-7B), demonstrating that recognition of taurine and amino acid attractants involves CAD1, but not CAD2.



**Figure 2-7. Effects of alanine substitutions of residues in CAD1 and CAD2 on attractant responses**

(A) (B) Capillary assays were carried out by my collaborative research group using *Vmlp201* which lacks *mlp24* and *mlp37* cells expressing wild-type or mutant Mlp37. Wild-type or each mutant Mlp37 are represented in different styles: (A) none, open bars; wild-type, red; W141A, yellow; W147A, green; R152A, blue; W154A, cyan; Y170A, pink; D172A, purple; D201A, grey bars and (B) none, open bars; wild-type, closed bars; Y221A, dotted bars; or H234A, cross-hatched bars. Capillaries were filled with TMN buffer containing the indicated concentrations (1, 10 or 100 mM) of each attractant.

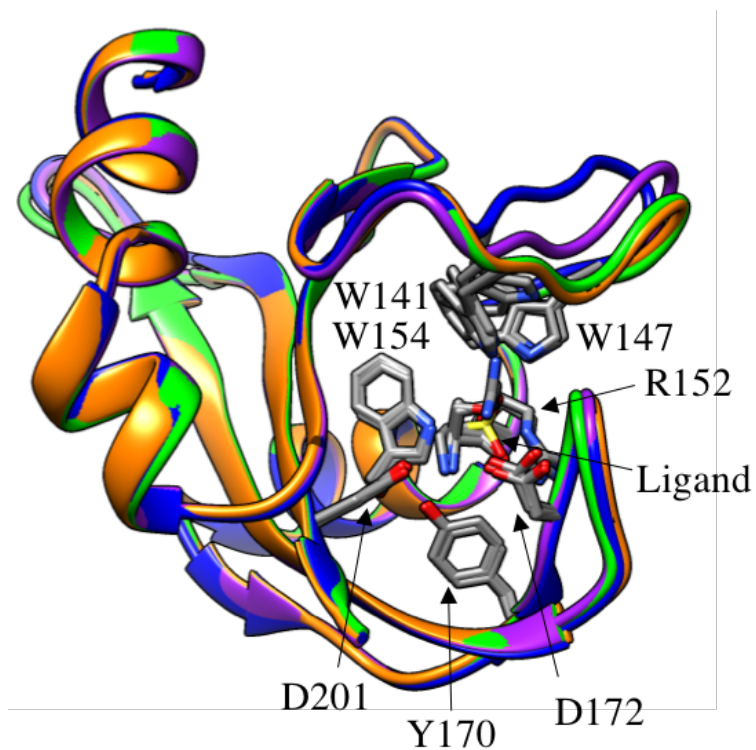
(Nishiyama S *et al.* (2016) *Sci Rep* 6:20866. Figure 4f)

#### **2-4-2. Conformational difference of Mlp37p CAD1 in various ligand binding structures**

The UJ loop of Mlp37p has different conformation as the binding ligand (Fig. 2-8). The UJ loop of Mlp37p L-serine and taurine complex locate closer to the LJ segment than that of L-alanine complex. In addition to this, that of L-arginine complex locates farther than that of L-alanine complex. The differences are occurred by the position of W141 included in the UJ loop. W141 is in contact with the hydroxy oxygen of ligand L-serine and the oxygen atom of the sulfonate group of ligand taurine as well as with the side chain guanidino group of ligand L-arginine, although their size and shape are different (Fig. 2-4B-D). This indicates that W141 of Mlp37p can move adequate positions to make interaction with the ligand and that is why Mlp37p recognizes a variety of ligand with similar affinities.

The taurine complex structure also provides a clue to the reason why taurocholic acid and taurodeoxycholic acid act as weak attractants, but cholic acid and deoxycholic acid do not. The amino group of taurine lies near the entrance of the binding pocket, and the entrance is slightly open (Fig. 2-6D). Taurocholic acid and taurodeoxycholic acid are compounds that cholic acid and deoxycholic acid are linked to the amino group of taurine. If the pocket entrance opens a little more, it is possible to accommodate the tauroconjugates by binding the taurine moiety in the pocket while leaving the cholic and deoxycholic moieties outside of the pocket. However, the binding affinity may be greatly reduced because some of the interactions with the amino group of taurine would be lost in this configuration.





**Figure 2-8. Comparison of the UJ loop structure of Mlp24p and Mlp37p**

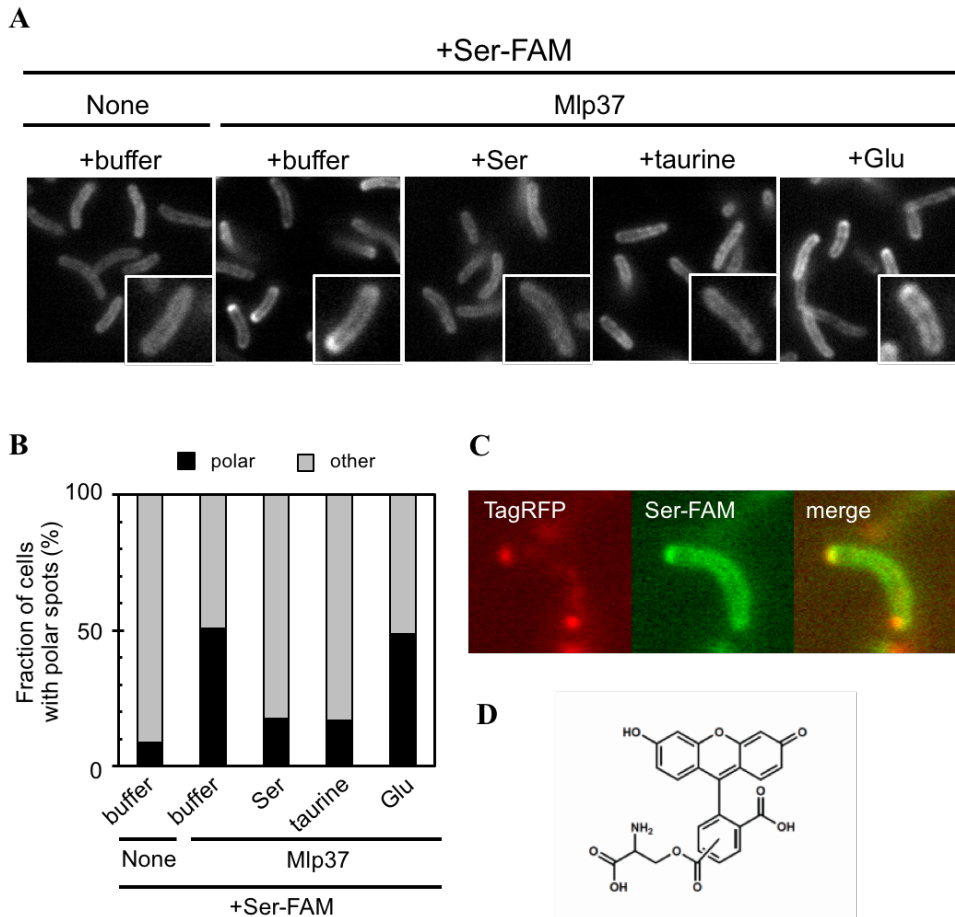
The CAD1 of the Mlp37p L-alanine complex chain A (purple), Mlp37p L-serine complex (green) and Mlp37p taurine complex chain A (orange) are superimposed onto that of the Mlp37p L-arginine complex chain A (blue). The residues involved in the ligand recognition are shown in stick models and labeled.

### **2-4-3. Role of CAD1 and CAD2**

The ligand complex structures and the mutational analyses demonstrate that only one of the two potential ligand-binding pockets (CAD1 in the membrane-distal CACHE domain) is involved in ligand binding. This is consistent with the previous reports on the related amino acid chemoreceptors, McpB of *B. subtilis* (10) and Mlp24 of *V. cholerae* (1). Some histidine kinases with two tandem PAS domains (PAS domain means cytosolic located CACHE domain) and other known dCACHE-type chemoreceptors except for Tlp1 and TlpC have also been reported to bind their ligands through membrane distal PAS domain or CAD1 (8-12, 14). Although the CAD2 of McpC interacts with lipid-anchored ligand-binding proteins (13), there has been no evidence that CAD2 in any MCP or histidine kinase with tandem PAS-like domains binds small ligands. It is possible that the CAD2 plays a role primarily in maintaining the receptor structure or facilitating receptor signaling.

#### **2-4-4. *In vivo* observation of ligand binding to Mlp37**

The open conformation of CAD1 in the L-arginine or the L-serine complex led me to test for *in vivo* visualization of ligand binding with a fluorescently labeled amino acid: L-serine 5(6)-carboxyfluorescein ester (Fig. 2-9D, hereafter referred to as Ser-FAM), in which carboxyfluorescein is linked to the hydroxyl group of L-serine through an ester bond. Prof. Kawagishi's group observed fluorescent spots in cells with a fluorescence microscopy. When treated with Ser-FAM, fluorescent spots were observed at poles of  $\Delta mlp24 \Delta mlp37$  (Vmlp201) cells expressing *mlp37* from pMlp37, but not those carrying the empty vector (Fig. 2-9). Pre-incubation with a 10-fold molar excess of L-serine or taurine before Ser-FAM labeling dramatically decreased the occurrence of polar fluorescent spots. By contrast, pre-incubation with L-glutamate, one of the weakest attractants, did not affect Ser-FAM labeling (Figs. 2-9A, B). These results suggest that the former attractants and Ser-FAM compete for binding to Mlp37. To further confirm that polar fluorescent spots result from Mlp37-bound Ser-FAM, they constructed a plasmid encoding Mlp37-TagRFP (named pKRB116). This fluorescent fusion protein expressed in strain Vmlp201 ( $\Delta mlp24 \Delta mlp37$ ) retained essentially the wild-type function to mediate taxis toward serine and showed polar localization (Fig. 2-9C). When treated with Ser-FAM, cells expressing Mlp37-TagRFP (Vmlp201/pKRB116) showed Ser-FAM spots that coincided with Mlp37-TagRFP (Fig. 2-9C). The results not only confirm the mechanism of ligand recognition by Mlp37, but also provide, for the first time, a tool to visualize ligand binding to a bacterial chemoreceptor *in vivo*.



**Figure 2-9. Imaging of Mlp37 with a fluorescently labelled ligand**

(A) Competitive labeling of Mlp37 with carboxyfluorescein-labeled L-serine [Ser-FAM].

Vmlp201 ( $\Delta mlp24 \Delta mlp37$ ) cells carrying the vector (none) or pMlp37 (Mlp37) were incubated with 100  $\mu$ M Ser-FAM and TMN buffer (+ buffer), 1 mM serine (+ Ser), 1 mM taurine (+ taurine), or 1 mM glutamate (+ Glu). (B) Fraction of cells with polar Ser-FAM spots in the absence or presence of an attractant. (C) Co-localization of Ser-FAM with Mlp37. Vmlp201 cells expressing Mlp37-TagRFP were treated with Ser-FAM: left, Mlp37-TagRFP; middle, Ser-FAM; right, merge. (D) The structural formula of Ser-FAM.

(Nishiyama S *et al.* (2016) *Sci Rep* 6:20866. Figure 5a-c and S7d.)

## 2-4. References

1. Nishiyama S, Suzuki D, Itoh Y, Suzuki K, Tajima H, Hyakutake A, Homma M, Butler-Wu SM, Camilli A, Kawagishi I. 2012. Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infect Immun* 80:3170-8.
2. Matilla MA, Krell T. 2017. Chemoreceptor-based signal sensing. *Curr Opin Biotechnol* 45:8-14.
3. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67:271-81.
4. Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* 66:125-32.
5. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235-42.
6. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-21.
7. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.

8. Machuca MA, Liu YC, Beckham SA, Gunzburg MJ, Roujeinikova A. 2016. The crystal structure of the tandem-PAS sensing domain of *Campylobacter jejuni* chemoreceptor Tlp1 suggests indirect mechanism of ligand recognition. *J Struct Biol* 194:205-13.
9. Gavira JA, Ortega A, Martin-Mora D, Conejero-Muriel MT, Corral-Lugo A, Morel B, Matilla MA, Krell T. 2018. Structural Basis for Polyamine Binding at the dCACHE Domain of the McpU Chemoreceptor from *Pseudomonas putida*. *J Mol Biol* 430:1950-1963.
10. Glekas GD, Foster RM, Cates JR, Estrella JA, Wawrzyniak MJ, Rao CV, Ordal GW. 2010. A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. *J Biol Chem* 285:1870-8.
11. Hothorn M, Dabi T, Chory J. 2011. Structural basis for cytokinin recognition by *Arabidopsis thaliana* histidine kinase 4. *Nat Chem Biol* 7:766-8.
12. Machuca MA, Johnson KS, Liu YC, Steer DL, Ottemann KM, Roujeinikova A. 2017. *Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate. *Sci Rep* 7:14089.
13. Glekas GD, Mulgem BJ, Kroc A, Dueffer KA, Lei V, Rao CV, Ordal GW. 2012. The *Bacillus subtilis* chemoreceptor McpC senses multiple ligands using two discrete mechanisms. *J Biol Chem* 287:39412-39418.
14. Zhang Z, Hendrickson WA. 2010. Structural characterization of the predominant family of histidine kinase sensor domains. *J Mol Biol* 400:335-53.

## Chapter 3

### Molecular mechanism of ligand recognition by Mlp24

#### 3-1. Introduction

Mlp24 and Mlp37 have been identified as major chemoreceptors involved in taxis to various amino acids in *V. cholera* (1, 2). Mlp37 mediates chemotactic responses to various amino acids, such as L-serine, L-alanine, L-cysteine, L-arginine, L-asparagine, L-threonine, L-lysine, L-valine and glycine, and taurine. Mlp24 is involved in chemotactic responses to L-serine, L-alanine, L-cysteine, L-arginine, L-asparagine, L-threonine, L-lysine, glycine, L-histidine, L-glutamine and L-proline. In addition to the chemotactic response, Mlp24 and Mlp37 have other distinct roles; Mlp24 is required for production of cholera toxin (1) and Mlp37 is involved in biofilm formation (2, 3). Although their cellular roles are different, the periplasmic region of Mlp24 and Mlp37 (Mlp24p and Mlp37p, respectively) shows about 44 % of amino acid sequence identity (Fig. 3-1). Mlp24 and Mlp37 bind various types of L-amino acids, but the binding affinities for their ligands are different from each other (1, 2). Mlp37p binds L-alanine, L-arginine, L-serine and taurine with similar affinity. In contrast to Mlp37p, Mlp24p shows different affinities for ligands. The dissociation constants of Mlp24p to L-arginine, L-asparagine, L-serine and L-proline are 2.5, 42.9, 71.9, 74.6  $\mu\text{M}$ , respectively (1). Moreover, Mlp24p does not bind taurine (2). These differences may be related to the distinct cellular functions of Mlp24 and Mlp37, but details are not known.

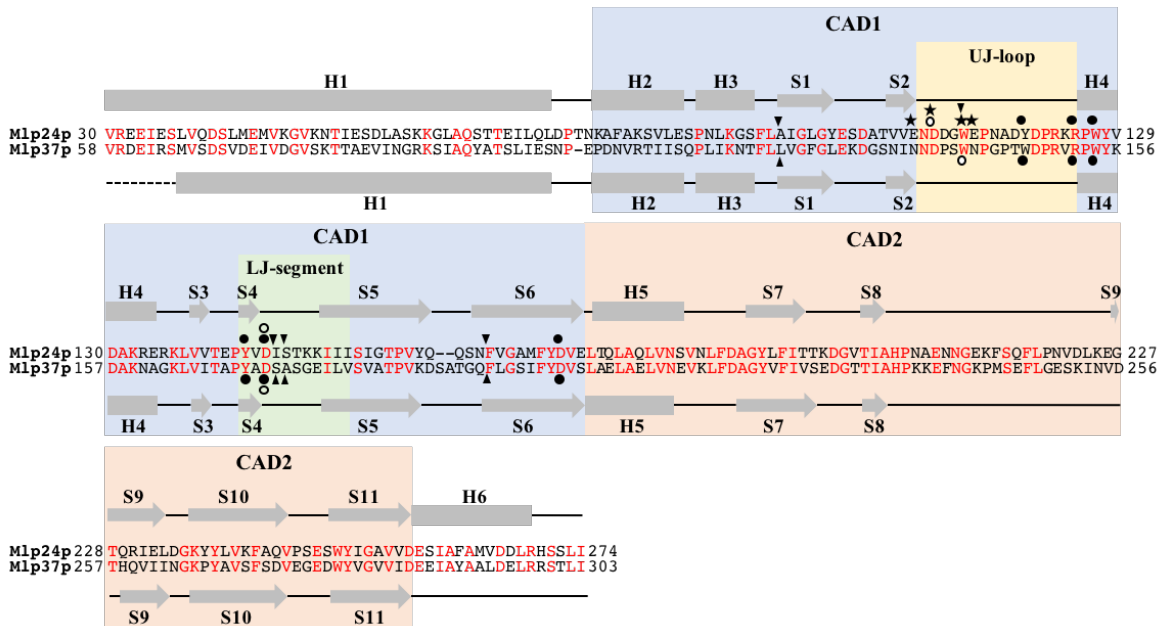
The structures of Mlp37p in complex with L-alanine, L-serine and taurine have been shown in chapter 2-2. The residues involved in the ligand binding of Mlp37p are almost fully

conserved in Mlp24p (Fig. 2-1), although the affinity and specificity of Mlp24p to the ligands differ from those of Mlp37p. Therefore, the molecular mechanism that produces the different specificity between Mlp24 and Mlp37 remains a mystery.

To elucidate the molecular mechanisms of ligand recognition by Mlp24 and Mlp37, I have determined a series of ligand complex structures of Mlp24p (the L-arginine, L-asparagine, glycine, L-proline and L-serine complexes) and compared them with the ligand binding structures of Mlp37p. The ligand-free form of Mlp24p has also determined.

In this chapter, I describe the structures of Mlp24p and the ligand recognition mechanism of Mlp24p. I discuss the molecular basis of the difference of ligand recognition between Mlp24 and Mlp37.





**Figure 3-1. Amino sequence alignment of Mlp24p and Mlp37p**

The secondary structures are shown above (Mlp24p) or below (Mlp37p) the sequences. Gray rectangles denote  $\alpha$ -helices and gray arrows  $\beta$ -strands. Residues involved in interaction with the ligand amino acid backbone and the side chain are marked with filled and open circles, respectively. The residues coordinating  $\text{Ca}^{2+}$  are indicated by black stars above the sequence of Mlp24p. Conserved residues are colored in red. CAD1, CAD2, the UJ-loop and the LJ-segment are shaded in pale blue, pale orange, yellow and green, respectively.

## **3-2. Materials and Methods**

### **3-2-1. Expression and purification of the periplasmic fragment of Mlp24**

Mlp24p was expressed and purified by an almost all similar method to Mlp37p. Strain BL21(DE3) (Novagen) carrying pVCP24 which expresses N-terminus GST and C-terminus His $\times$ 6-fused Mlp24p was cultured in LB broth Lennox (Nacalai Tesque, Inc., Kyoto Japan) containing 50  $\mu$ g/ml of ampicillin at 37°C until the cell density had reached an OD<sub>600</sub> of about 0.8. Isopropyl  $\beta$ -d-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.1 mM to induce protein expression, and the culture was continued for 12 hours at 20°C. Cells were harvested by centrifugation, suspended in phosphate buffered saline (PBS) (137 mM of NaCl, 2.7 mM of KCl, 10 mM of Na<sub>2</sub>HPO<sub>4</sub>•12H<sub>2</sub>O, and 1.8 mM of KH<sub>2</sub>PO<sub>4</sub> at pH 7.4) and disrupted by sonication. After removing cell debris by centrifugation, the cell lysate was loaded on a Glutathione-Sepharose<sup>TM</sup> 4B column (GE Healthcare) followed by washing with PBS. Proteins were eluted with 50 mM Tris-HCl (pH 8.0) buffer containing 10 mM reduced glutathione. The N-terminal GST tag was then cleaved using PreScission<sup>TM</sup> Protease (GE Healthcare), and the reactant was dialysed against 1.0 L dialysis buffer (50 mM Tris-HCl pH 7.0, 150 mM NaCl, 1 mM DTT, and 1 mM EDTA) at 4°C for 12 hours using Spectra/por Dialysis Membrane MWCO 6,000-8,000 (Spectrum Laboratories, Inc.). The protein solution was loaded again to a Glutathione-Sepharose<sup>TM</sup> 4B column (GE Healthcare) to remove GST and unreacted protein, and further purified by size exclusion chromatography with a High Load 26/60 Superdex 200 gel filtration column in 20 mM Tris-HCl and 150 mM NaCl (pH 8.0). The peak fraction was collected and concentrated to 10 mg/mL. The purity of the purified proteins was examined by SDS-PAGE.

### **3-2-2. Crystallization**

Initial crystallization screening was performed by sitting-drop vapor-diffusion technique. Crystallization drops were prepared by mixing 1  $\mu$ l of protein solution (10 mg/ml) with 1  $\mu$ l of reservoir solution. The ligand amino acids and taurine were added in the protein solution before preparing the drop to the final concentration of 10 mM except for L-arginine (100 mM).  $\text{CaCl}_2$  was also added in the protein solution to the final concentration of 10 mM for crystallization of the Mlp24p glycine and the L-serine complexes, because the calcium binding to the UJ-loop has been revealed before initial screening of these complexes. Initial screening was carried out using the following screening kit; Wizard Classic I and II (Rigaku Reagents, Inc.), Wizard Cryo I and II (Rigaku Reagents, Inc.), and Crystal Screen I and II (Hampton Research). The crystallization conditions were optimized by the hanging-drop vapor-diffusion method. Crystallization drops were prepared by mixing 1.5  $\mu$ l of protein solution with an equal volume of reservoir solution at 293K. All crystals appeared within 3 days after optimization. Their final crystallization conditions, space groups and unit cell dimensions are summarized in Table 3-1 and 3-2.

**Table 3-1 Summary of crystallization conditions of Mlp24p**

<b>Ligands</b>	<b>Buffer</b>	<b>Precipitant</b>	<b>Additive</b>
none	100 mM MES pH 6.0	42.5 % (v/v) PEG-400	100 mM Ca(OAc) <sub>2</sub> 3 % (v/v) DMSO
100 mM L-arginine	100 mM MES pH 6.0	45 % (v/v) PEG-400	100 mM Ca(OAc) <sub>2</sub>
10 mM L-asparagine	100 mM MES pH 6.0	50 % (v/v) PEG-400	100 mM Ca(OAc) <sub>2</sub>
10 mM glycine	100 mM Tris pH 8.5	20 % (w/v) PEG-8,000	200 mM MgCl <sub>2</sub>
10 mM L-proline	100 mM Tris pH 8.5	50 % (v/v) PEG-400	100 mM Ca(OAc) <sub>2</sub>
10 mM L-serine	100 mM Tris pH 8.5	35 % (w/v) PEG-4,000	none

**Table 2-2 Summary of X-ray data collection and refinement statistics of Mlp24p**

Ligands	none	L-arginine	L-asparagine	glycine	L-proline	L-serine
<b>Data collection</b>						
Space group	<i>C</i> 2	<i>P</i> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>I</i> 222	<i>I</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub>
Cell dimensions						
<i>a</i> (Å)	104.1	64.5	91.9	53.4	59.1	52.7
<i>b</i> (Å)	88.4	102.8	101.4	64.6	111.2	62.3
<i>c</i> (Å)	64.5	93.4	128.4	136.6	113.3	73.2
$\beta$ (°)	105.0	94.5	90.0	90.0	90.0	105.6
Wavelength (Å)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Resolution (Å)	49.9-2.30 (2.38-2.30)	48.5-2.70 (2.83-2.70)	47.0-2.50 (2.59-2.50)	41.6-2.14 (2.27-2.14)	56.6-2.35 (2.44-2.35)	39.4-2.10 (2.16-2.10)
<i>R</i> <sub>merge</sub> (%)	10.7 (32.1)	10.3 (33.7)	12.1(55.8)	7.1 (38.4)	10.9 (42.3)	9.1 (39.5)
<i>I</i> / $\sigma$ <i>I</i>	7.2 (3.3)	8.7 (3.7)	8.5(2.9)	13.7 (3.0)	14.5 (4.9)	10.2 (3.0)
Completeness (%)	98.3 (99.9)	97.1 (96.2)	100(100)	98.1 (95.1)	100 (100)	100 (100)
Redundancy	3.0 (3.0)	3.1 (3.0)	4.8(4.8)	3.1 (2.9)	7.3 (7.3)	3.6 (3.7)
Content per ASU	2	4	4	1	1	2
<b>Refinement statistics</b>						
Resolution (Å)	36.1-2.30 (2.39-2.30)	48.5-2.70 (2.78-2.70)	68.1-2.50 (2.56-2.50)	41.6-2.14 (2.27-2.14)	39.7-2.35 (2.50-2.35)	39.4-2.10 (2.19-2.10)
Number of reflections						
Working set	23,909 (2,669)	30,709 (2,490)	40,116 (2,648)	12,832 (2,473)	15,167 (2,473)	25,526 (2,796)
Test set	1,228 (131)	1,689 (160)	2,050 (125)	676 (130)	783 (144)	1,288 (151)
<i>R</i> <sub>w</sub> (%)	20.59 (25.10)	21.50 (24.67)	19.54 (23.30)	19.55 (22.87)	21.44 (21.05)	19.82 (21.86)
<i>R</i> <sub>free</sub> (%)	26.62 (35.45)	25.54 (28.02)	25.54 (29.78)	24.45 (26.94)	25.49 (28.91)	25.57 (26.74)
R.m. s. deviations						
Bond length (Å)	0.008	0.003	0.009	0.008	0.008	0.009
Bond angle (°)	1.064	0.668	1.106	1.067	1.180	1.167
<i>B</i> -factors						
Protein atoms	34.52	36.55	38.54	40.62	30.74	30.51
Ligand atoms	42.56	37.43	43.99	30.83	22.00	24.42
Solvent atoms	35.66	34.06	39.18	37.39	32.05	28.44
Ramachandran plot (%)						
In preferred regions	96.19	95.16	96.44	96.58	96.17	97.02
In allowed regions	3.81	4.73	3.56	2.99	3.83	2.13
Outliers	0.00	0.11	0.00	0.43	0.00	0.85
Determined regions (amino acid number and chain ID)	28-266 (A) 28-265 (B)	28-265 (A) 27-265 (B) 28-268 (C) 25 -265 (D)	28-268 (A) 28-268 (B) 28-267 (C) 28-267 (D)	28-263 (A)	30-266 (A)	33-269 (A) 33-269 (B)
Number of atoms						
Protein	3,738	7,511	7,540	1,850	1,863	3,726
Ligand	6	48	36	5	8	14
Solvent	104	82	128	45	112	127

Values in parentheses are for the highest resolution shell.

$$R_w = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}, R_{free} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$$

### **3-2-3. X-ray data collection and structure determination of Mlp24p**

X-ray diffraction data were collected at synchrotron beamline BL41XU in SPring-8 (Harima, Japan), with the approval of the Japan Synchrotron Radiation Research Institute (JASRI). The crystals of Mlp24p glycine and L-serine complexes were soaked in the 1 : 9 mixture of 2-Methyl-2, 4-pentanediol (MPD) and the reservoir solution for a few seconds before transfer into liquid nitrogen for freezing. All X-ray diffraction data were measured at 95K under nitrogen gas flow. The diffraction data were processed with MOSFLM (4) or XDS (5) and scaled with SCALA (6), Aimless (6) or XDS (5). The initial phases were determined by molecular replacement (MR) with Phenix (7). A poly-alanine model derived from the Mlp37p L-alanine complex structure (PDB ID: 3C8C) was divided into four regions (38 to 55, 59 to 73, 77 to 174 and 178 to 258) and was used as a separate search model for the phasing of Mlp24p ligand free. Molecular replacement for the other Mlp24p crystal data was performed using the Mlp24p structure as a search model. The atomic models were constructed with Coot (8) and refined with Phenix (7). Data collection and refinement statistics are summarized in Table 3-1 and 3-2.

### 3-3. Results

#### 3-3-1. Overall structures of ligand-free Mlp24p and Mlp24p in complex with various ligands

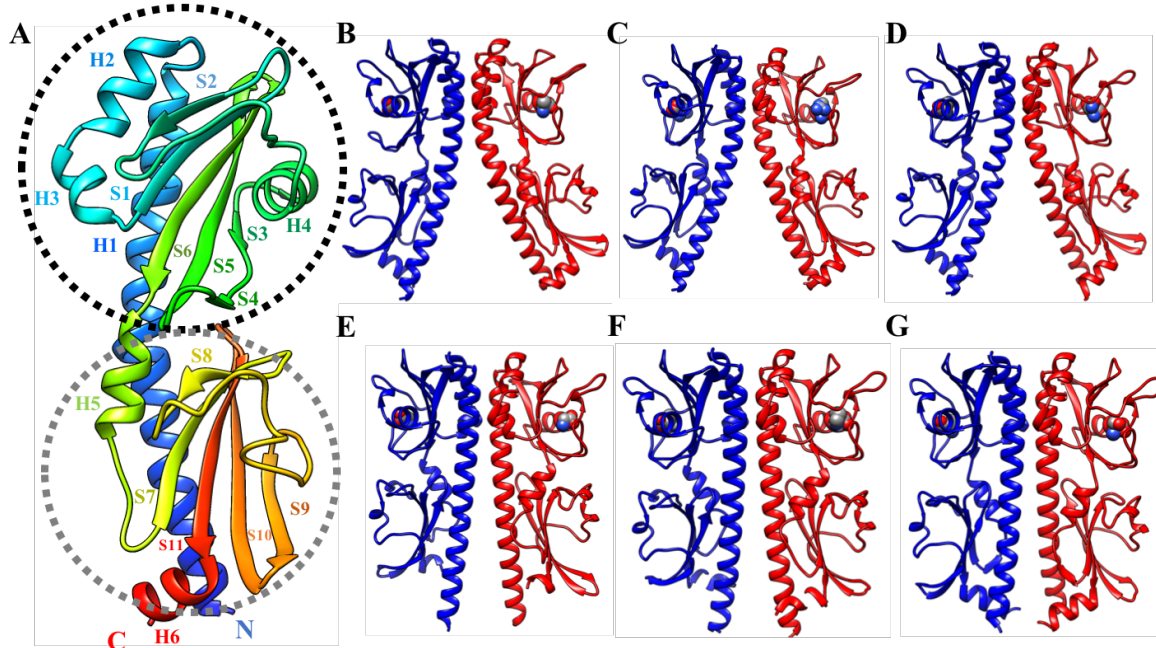
I determined the crystal structures of Mlp24p and its complexes with L-arginine, L-asparagine, glycine, L-proline and L-serine at 2.10 - 2.35 Å resolution (Fig. 3-2, Table 3-1 and 3-2). The ligand complex crystals were prepared by co-crystallization and grew in different space groups and unit cell parameters in distinct conditions. The molecular arrangements in these crystals, therefore, differ from one another (Table 3-2). The structures of Mlp24p are very similar to the Mlp37p structures. Each Mlp24p subunit is composed of a long N-terminal  $\alpha$ -helix (H1), two CACHE domains (CAD1 and CAD2) tandemly arranged along H1 and a C-terminal helix (H6) (Fig. 3-2A). CAD1, the membrane-distal domain, consists of three  $\alpha$ -helices (H2-H4) and six  $\beta$ -strands (S1-S6), five of which (S1-S3, S5 and S6) form a central  $\beta$ -sheet (Fig. 3-2A). The domain has a mouth-like pocket with the upper jaw formed by the loop of 16 residues connecting S2 and H4 (UJ-loop) and the lower jaw formed by S4, the N-terminal half of S5 and the loop connecting S4 and S5 (LJ-segment). CAD2, the membrane-proximal domain, comprises an  $\alpha$ -helix (H5) and a  $\beta$ -sheet composed of five  $\beta$ -strands (S7-S11). Although CAD2 shows a similar fold to CAD1, all ligands bind only to CAD1 (Fig. 3-2B-G), like other known dCACHE-type chemoreceptors except for Tlp1 and TlpC (9-13).

The ligand-free and the ligand complexes of Mlp24p form homo-dimers in crystals (Fig. 3-2). The C-terminal region of H1, H2 and the N-terminal region of H3 contribute to the dimer interaction (Fig. 3-3). L63, I70 and L73 in H1, F80 and V84 in H2, and L90 in H3 form a

hydrophobic interaction core on the dimer interface. The dimer is further stabilized by hydrophilic interactions surrounding the hydrophobic interface. The side chain of N89 in H3 hydrogen bonds to S66' in H1' (prime denotes the partner subunit) and the main chain nitrogen atom of N89 hydrogen bonds to the side chain of E69' in H1'. These dimer interactions were conserved in all the Mlp24p structures.

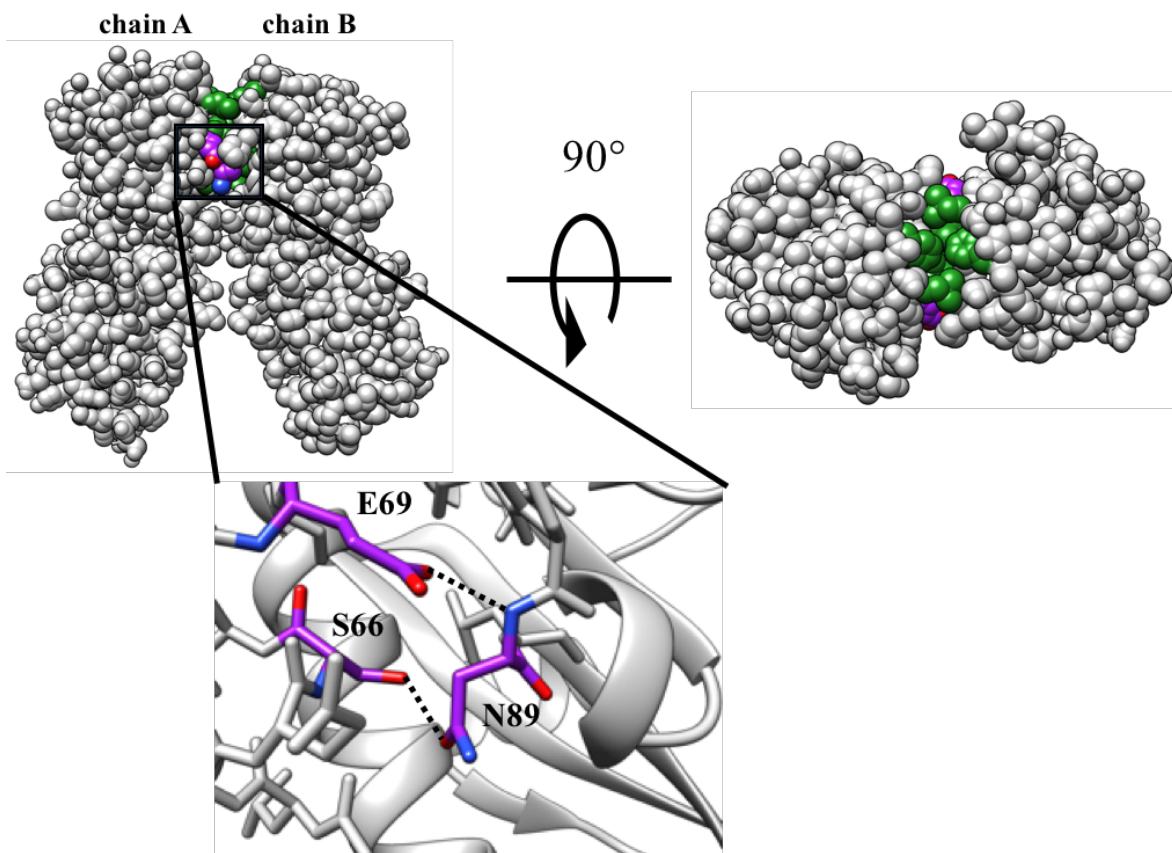
Although the crystal structures show almost identical dimer interactions, they show different dimeric structures that can be classified into three groups (Fig. 3-4A). The ligand-free Mlp24p, the L-arginine and the L-asparagine complexes show similar dimeric structures, in each of which the proximal ends of the two subunits are apart (Fig. 3-2E-G). They adopt very similar bending conformations in H1 (Fig. 3-4A), which is bent outward at S55, and therefore the N-terminal ends of H1 in the dimer are apart from each other to form a proximal open dimer conformation. In contrast, H1 in the glycine and the L-proline complexes has no bending and extends almost straight (Fig. 3-2E and F). Therefore, they adopt nearly parallel dimer conformations. The conformation of H1 in the L-serine complex resembles those in the glycine and L-proline complexes (Fig. 3-2G). However, the dimer structure is slightly different (Fig. 3-4A), because the second subunit of the dimer tilts at the dimer interface with 12.5 degree around the axis perpendicular to the dimer axis. These differences of the dimer structures seem to be caused by the different molecular packing in the crystals (Table 3-2). The N-terminal regions of two H1 helices of the dimer cross those of the neighboring dimer in the crystals of the proximal open dimers, but not in the other crystals (Fig. 3-4B-G).





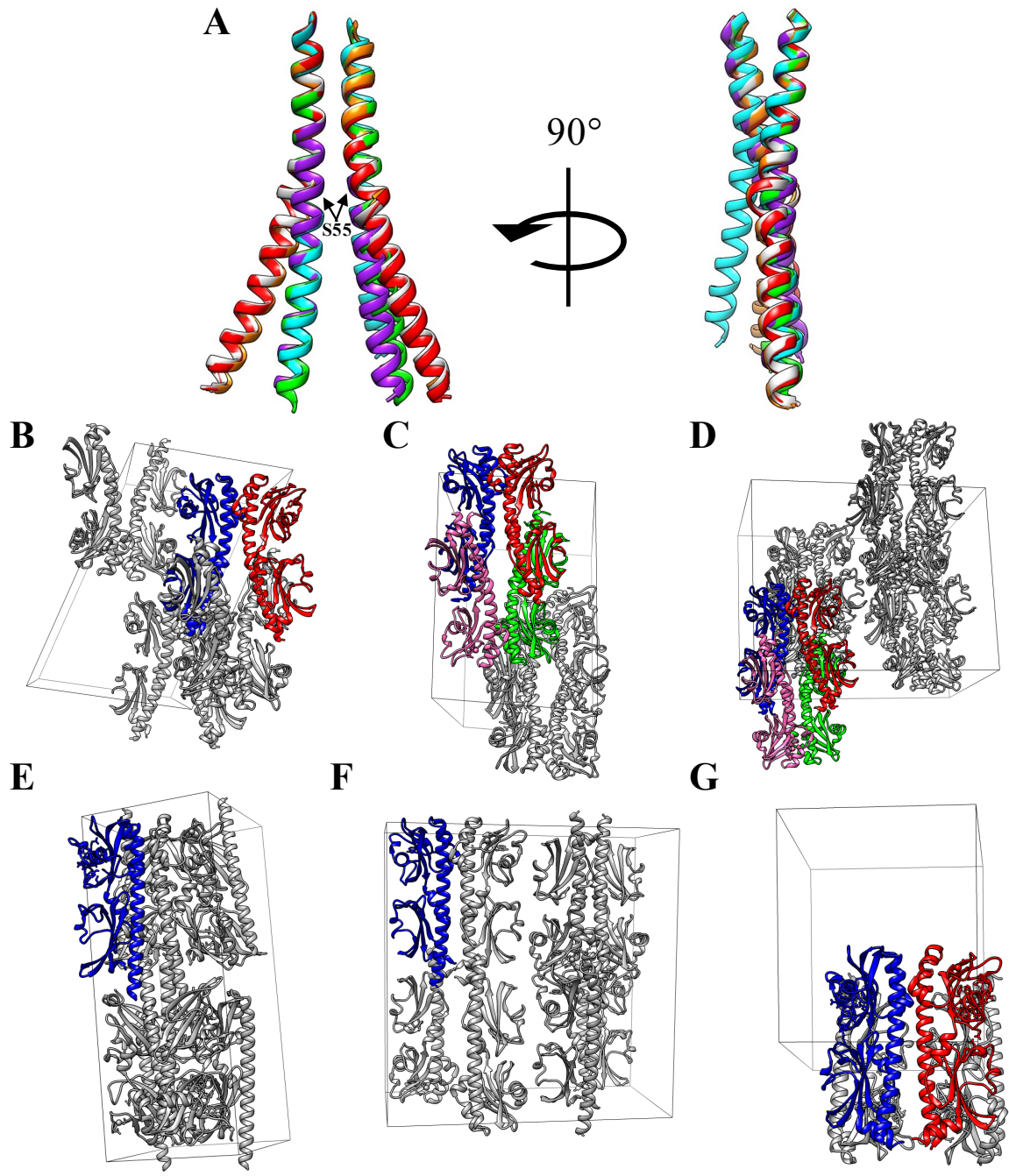
**Figure 3-2. Overall structures of Mlp24p**

(A) Ribbon model of the Mlp24p glycine complex is colored in rainbow from the N-terminus (blue) to the C-terminus (red). The membrane distal and proximal CACHE domain (CAD1 and CAD2) are indicated with black and gray broken circles, respectively. The secondary structure elements are labeled. The ligand glycine is shown as a ball model. (B-G) Ribbon representation of the dimeric structures of Mlp24p (B), the Mlp24p L-arginine complex (C), the Mlp24p L-asparagine complex (D), the Mlp24p glycine complex (E), the Mlp24p L-proline complex (F) and the Mlp24p L-serine complex (G). The ligand molecules are shown as ball models. Each subunit is painted in different color (blue and red).



**Figure 3-3. Dimer interface of Mlp24p**

The Mlp24p dimer is represented by ball model. The hydrophobic residues on the subunit interface (L63, I70, L73, F80 and L90) are colored in green. The residues that form the intersubunit hydrogen bonds (S66, E69 and N89) are colored in purple for carbon, red for oxygen and blue for nitrogen atoms.



**Figure 3-4. The H1 conformation correlates the crystal packing**

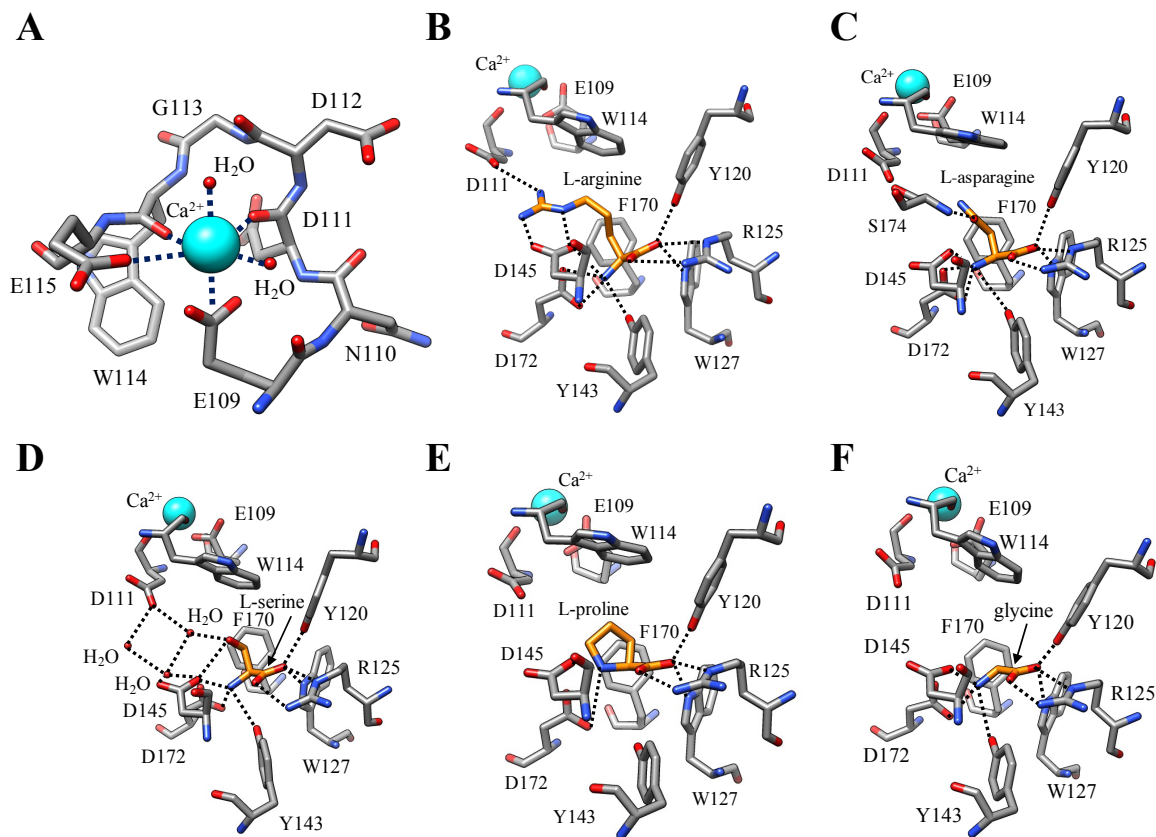
(A) Comparison of the H1 conformation of the Mlp24p structures. The H1 helices in the dimer of Mlp24p in complex with L-arginine (red), L-asparagine (orange), glycine (green), L-proline (purple) and L-serine (cyan) are superimposed onto those of apo-Mlp24p (light gray). Apo-Mlp24p, the L-arginine and the L-asparagine complexes show very similar open conformation which is bent outward at S55 and the glycine and the L-proline complexes adopt a closed conformation. The L-serine complex also adopts a closed conformation but slightly tilted from the glycine and the L-proline complexes. (B-G) Molecular packing of the crystals of apo-Mlp24p (B), the L-arginine complex (C), the L-asparagine complex (D), the glycine complex (E), the L-proline complex (F) and the L-serine complex (G). Molecules in an asymmetric unit is colored in red, blue, hot pink and green and other symmetry related molecules are in gray. The unit cells are shown with gray line.

### 3-3-2. Ligand recognition mechanism of Mlp24p

The ligand amino acid is bound in the mouth-like pocket of CAD1. The pocket is composed of the UJ loop, H4, the LJ-segment and the central  $\beta$ -sheet. A calcium ion binds to the N-terminal region of the UJ loop and defines the orientation of the loop by forming a unique turn conformation. The calcium ion is coordinated by the side chain carboxy groups of E109 and E115, the main chain carbonyl oxygen atoms of D111 and W114 and two water molecules with octahedral geometry (Fig. 3-5A). These interactions bend the UJ loop around the calcium to form the turn structure. This loop conformation is further stabilized by the hydrogen bond between the side chain carboxy group of D111 and the main chain NH group of G113 (Fig. 3-5A).

All the ligand amino acids except for L-proline bind to the pocket in the same orientation, and their  $\alpha$ -amino and  $\alpha$ -carboxy groups are recognized in the same manner by Mlp24p. The  $\alpha$ -carboxy group forms hydrogen bonds with Y120 in the UJ loop, and with R125 and W127 in H4. The  $\alpha$ -amino group hydrogen bonds with D172 in S6, and with Y143 and D145 in the LJ segment (Fig. 3-5B-E). These three residues are triangularly arranged below the amino group. These interactions are very similar to those in the Mlp37p ligand complexes, although Y120 is replaced to W147 in Mlp37p. L-proline binds to the pocket of Mlp24p in the similar manner to the other L-amino acids. The  $\alpha$ -carboxy group hydrogen bonds with Y120, R125 and W127 and the  $\alpha$ -amino group with D145 and D172 (Fig. 3-5F). These arrangement is reasonable because the amino group is positively charged under our experimental conditions (pH < 9).

In contrast to the ligand main chain, the ligand side chains are recognized in different manners, although the conformation of the residues involved in the ligand side chain recognition is almost identical in all cases. The guanidino group of the ligand L-arginine directly interacts with D111 in the UJ loop and D145 in the LJ segment, and the hydrophobic side chain arm of L-arginine interacts with F170 and W114 (Fig. 3-5B). The side chain of the ligand L-asparagine is recognized by main chain amide group in the LJ segment. The side chain carbonyl oxygen of L-asparagine hydrogen bonds to the main chain NH group of S147, but no specific hydrophilic interaction with the side chain NH<sub>2</sub> group of L-asparagine was found. The β-methylene group of L-asparagine is in contact with F170 (Fig. 3-5C). Unlike L-arginine and L-asparagine, the side chain hydroxy group of the ligand L-serine is indirectly recognized by D111 and D145 through the hydrogen bonding network mediated by three water molecules. The β-methylene group of L-serine interacts with F170 as is the case of L-asparagine (Fig. 3-5D). The side chain atoms of the ligand L-proline are in hydrophobic contact with F170 and W114 (Fig. 3-5E). The ligand-binding pocket of the glycine complex has a space between the bound glycine and W114 (Fig. 3-5F). The conformation of the UJ-loop is almost identical in all the ligand complexes, and therefore the small ligands do not tightly contact with the UJ-loop in the pocket.



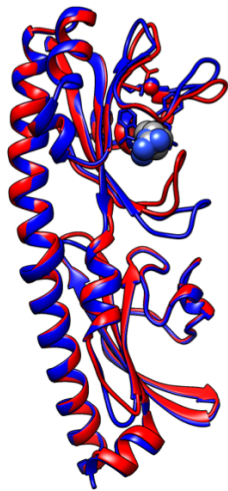
**Figure 3-5. Close up view of the ligand binding pocket of Mlp24p ligand complexes**

(A)  $\text{Ca}^{2+}$  coordination to the UJ loop. (B-F) The ligand-binding pockets of Mlp24p in complex with L-arginine (B), L-asparagine (C), L-serine (D), L-proline (E), glycine (F). The ligand molecules are colored in orange. The protein residues are represented by gray stick models. The nitrogen and oxygen atoms are colored in blue and red, respectively. The bound water molecules are shown as a red ball.  $\text{Ca}^{2+}$  is shown as a cyan ball. Possible hydrogen bonds are indicated by black broken lines.

### **3-3-3. Ligand binding induces the domain closure of CAD1.**

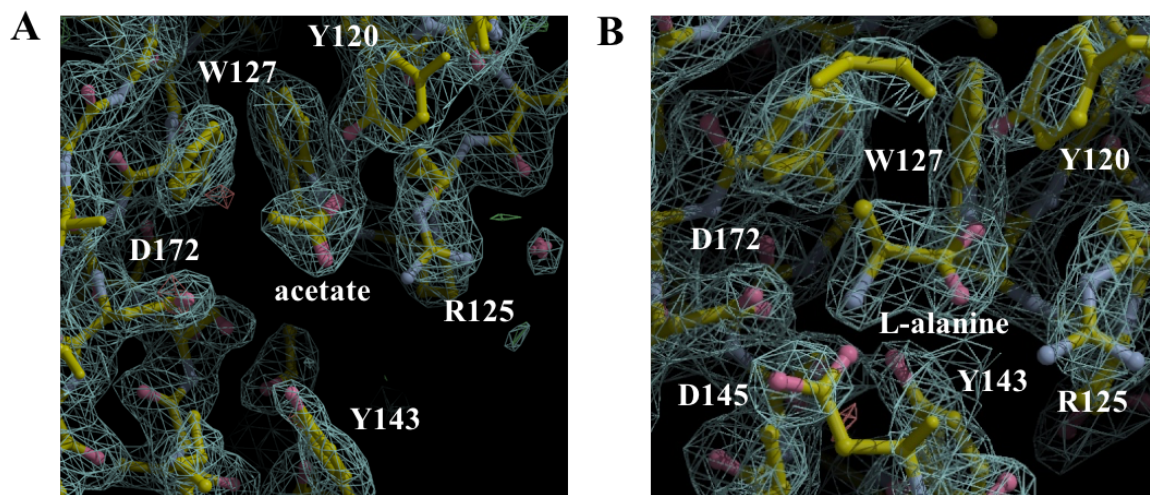
The two CAD1 domains of the ligand-free Mlp24p dimer adopt different conformations. The LJ-segment of CAD1 in one subunit of the Mlp24p dimer projects outward and hence the entrance of the pocket is wide open (Fig. 3-6). A small density is present in the CAD1 pocket (Fig. 3-7A), and is assigned as acetate because the shape of the density resembles acetate and the crystallization buffer contained acetate. The carboxy group of the acetate interacts with Y120, R125 and W127 in the same manner as the  $\alpha$ -carboxy group of the ligand L-amino acids (Fig. 3-8B). CAD1 in the other subunit of the Mlp24p dimer adopts a closed conformation. The LJ-segment bends upward to close the entrance of the pocket. A clear density that can be assigned as L-alanine is observed in the pocket (Fig. 3-7B and 3-8A), although all solutions used in purification and crystallization contain no L-alanine. Such unexpected L-alanine binding also occurred in the structure analysis of Mlp37p (PDB ID: 3C8C). The density corresponding to L-alanine was found in CAD1 in both subunits of the Mlp37p dimer without adding L-alanine in the crystallization solution. The CAD1 domains of all the ligand complex structures of Mlp24p show closed conformation. These observations suggest that L-amino acid binding to the CAD1 domain pocket induces bending of the LJ segment to close the entrance of the pocket. A similar conformational change has been observed in Tlp3 (14).



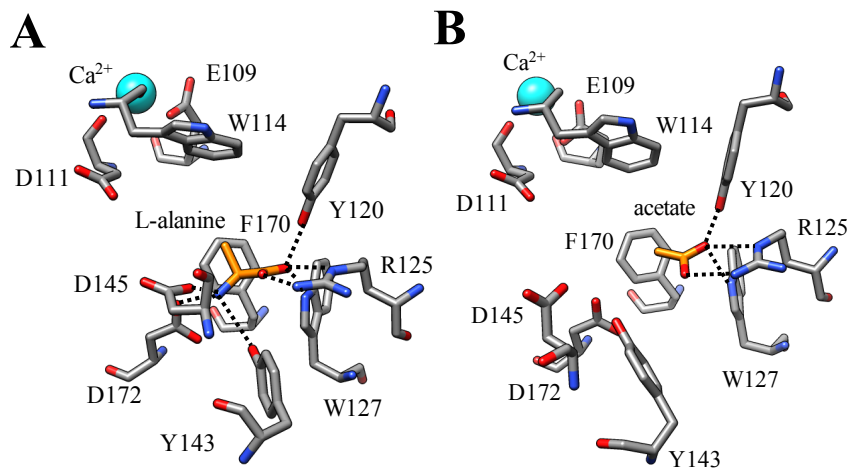


**Figure 3-6. The ligand induced structural change of Mlp24p**

The structure of the L-arginine complex chain A (red) is superimposed onto that of Mlp24p chain A (blue). L-arginine is shown in ball model.



**Figure 3-7.  $2F_o - F_c$  electron density map of the ligand binding pocket of apo-Mlp24p**  
 $2F_o - F_c$  maps of the ligand binding pocket of apo-Mlp24p chain A (A) and B (B) are drawn at 1.1 sigma with the molecular models. Clear density is observed in the ligand binding pocket of both chains. The residues involved in the ligand recognition are labelled.



**Figure 3-8. The ligand binding pocket of ligand-free Mlp24p**

The ligand binding pocket of ligand-free Mlp24p chain A (A) and chain B (B) are shown in gray stick models. The bounding L-alanine and acetate are colored in orange. The nitrogen and oxygen atoms are colored in blue and red, respectively.  $\text{Ca}^{2+}$  is shown as a cyan ball. Possible hydrogen bonds are indicated by black broken lines.

### **3-4. Discussion**

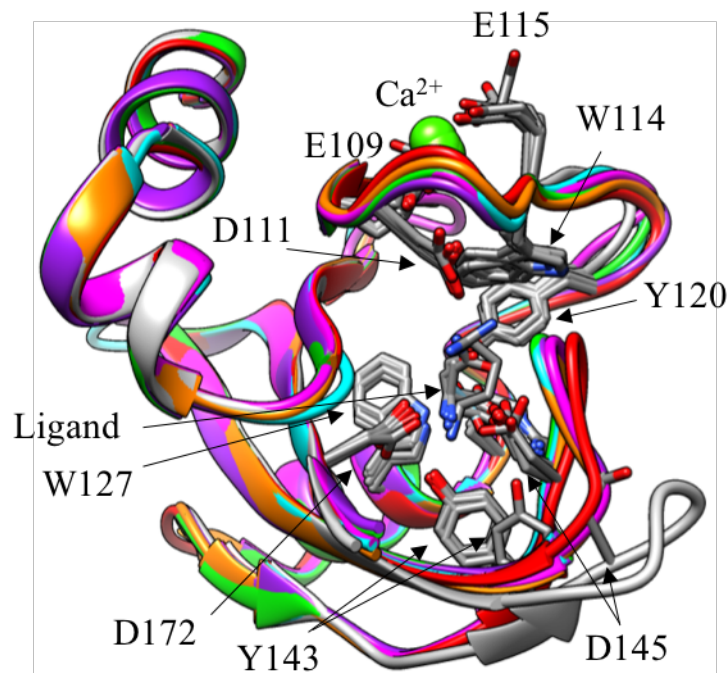
#### **3-4-1. Calcium ion enhances the binding affinity of Mlp24p to ligands**

The crystal structures of the Mlp24p indicated that the N-terminal region of the UJ loop of Mlp24 has a unique turn structure stabilized by a calcium ion (Fig. 2-8, 3-5 and 3-9). This turn structure restricts the conformation of the N-terminal region of the UJ loop and determines the orientation of the UJ loop (Fig. 3-5, 3-9). Therefore calcium ion is thought to be important for ligand binding.

This idea is supported by ITC assay performed by Nishiyama and Kawagishi. They constructed a plasmid encoding a His-tag-free fragment of Mlp24p (Mlp24pNH) to avoid undesired binding of ions to the His-tag and purified it for ITC measurements. The results demonstrated that Mlp24pNH binds  $\text{Ca}^{2+}$ , although the affinity is relatively low ( $K_d$  is in the sub-millimolar range, Fig.3-10A and Table 3-3). ITC analyses of Mlp24pNH with various amino acids in the presence and absence of 10 mM  $\text{CaCl}_2$  in EDTA-free buffer indicated that  $\text{Ca}^{2+}$  significantly decreased  $K_d$  values of Mlp24pNH for amino acids with smaller side chains (Table 3-3). For example, the  $K_d$  values of Mlp24pNH for L-asparagine in the presence and absence of  $\text{Ca}^{2+}$  were 6.3  $\mu\text{M}$  and 49.5  $\mu\text{M}$ , respectively, and those for serine were 18.3  $\mu\text{M}$  and 184.2  $\mu\text{M}$  (Fig. 3-10B, C and Table 3-3). On the other hand, the  $K_d$  values for amino acids with long, positive side chains, such as L-arginine and L-lysine, were low even in the absence of  $\text{Ca}^{2+}$ , and were slightly decreased by the addition of  $\text{Ca}^{2+}$  (Table 3-3). L-glutamate, a negatively charged amino acid, did not bind to Mlp24pNH. These results indicate that  $\text{Ca}^{2+}$  increases affinities of Mlp24 for the ligands, especially for the amino acids with smaller side

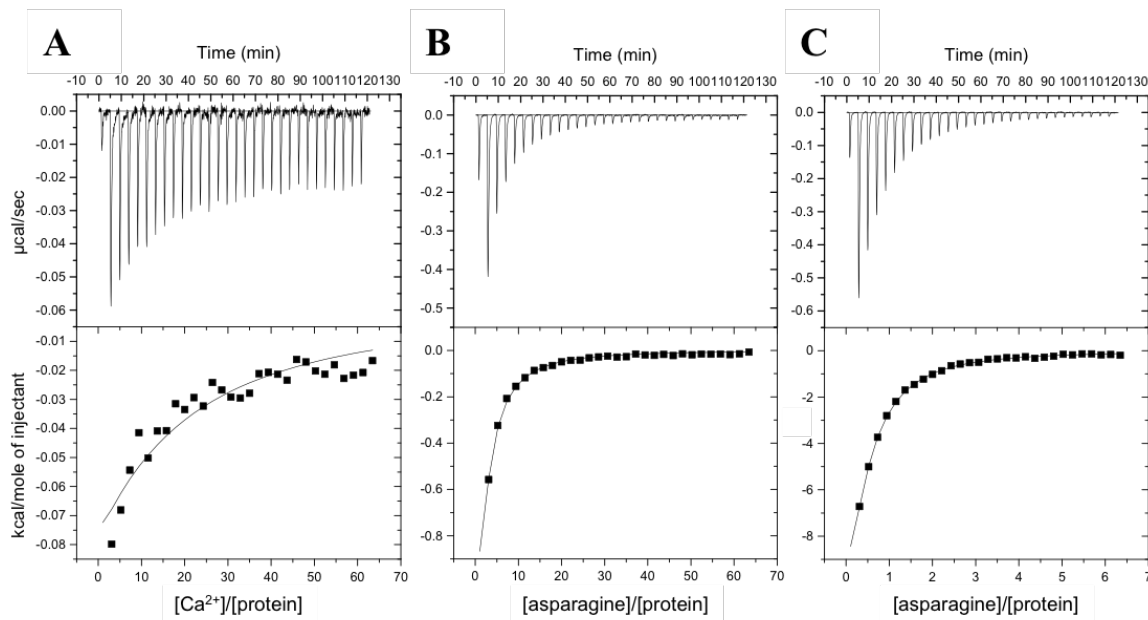
chains. However, Mlp24 still shows different affinities for ligands to some extent in the presence of 10 mM  $\text{Ca}^{2+}$ .

Capillary assay (1) of cells expressing Mlp24 demonstrated that  $\text{Ca}^{2+}$  enhanced Mlp24-mediated chemotaxis *in vivo*. A strain (Vmlp201) lacking Mlp24 as well as Mlp37 (2) transformed with an *mlp24*-expressing plasmid (pMlp24). The mutant cells overexpressing Mlp24 were not attracted to Tris-MgCl<sub>2</sub>-NaCl buffer containing up to 100 mM CaCl<sub>2</sub> (Fig. 3-11A), suggesting that  $\text{Ca}^{2+}$  binding to Mlp24 did not elicit an attractant signal. Responses to 0.1 to 10 mM L-asparagine were enhanced by 10 mM CaCl<sub>2</sub> (Fig. 3-11B).  $\text{Ca}^{2+}$  showed a milder effect on responses to 10 mM L-serine (Fig. 3-11C). In contrast, no significant effects were observed for L-arginine response (Fig. 3-11D). These results indicate that calcium ion binding to Mlp24 potentiates chemotaxis to amino acids with smaller side chains by increasing its ligand-binding affinities.



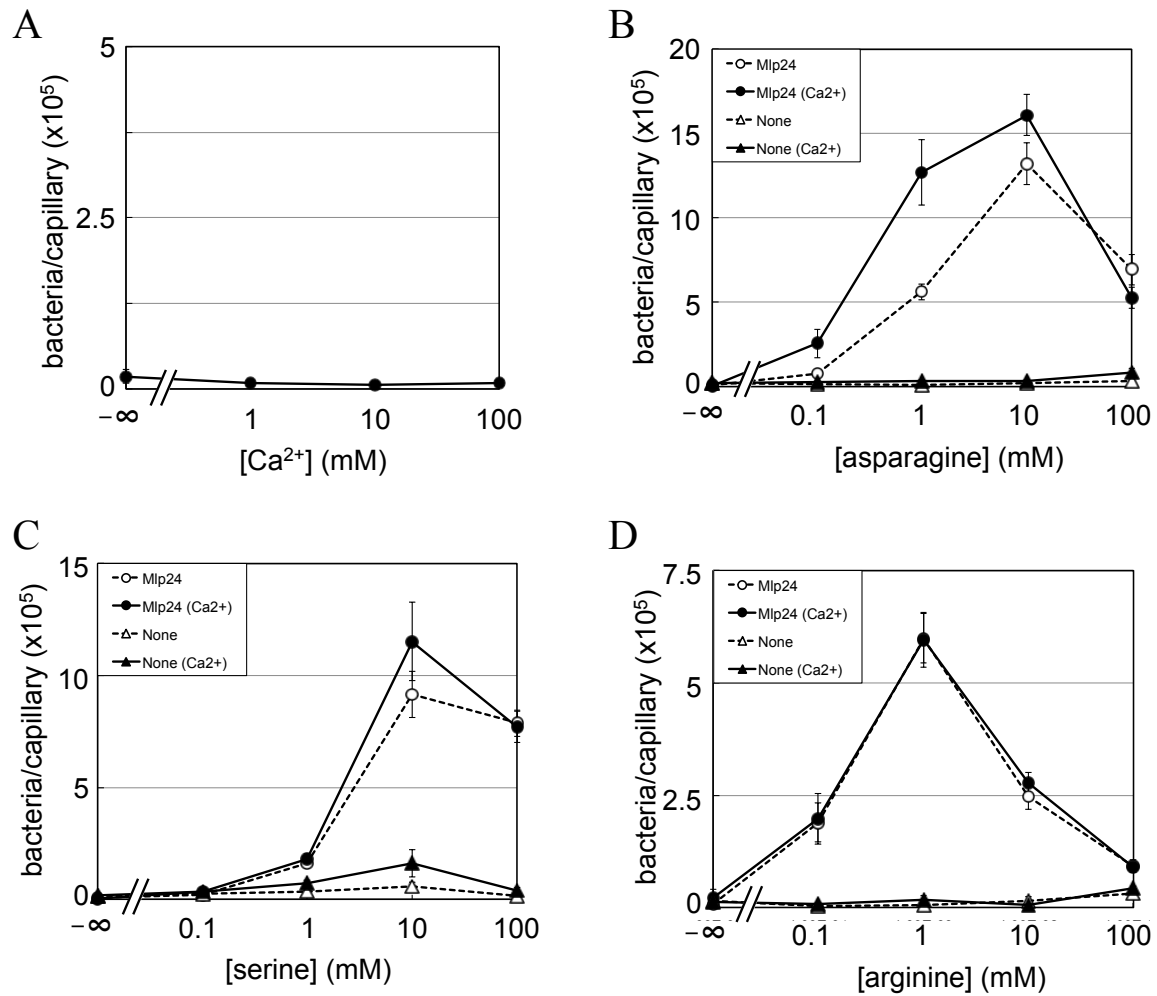
**Figure 3-9. Comparison of the CAD1 structure of Mlp24p**

The CAD1 of the Mlp24p chain B (pink), Mlp24p L-arginine complex chain A (red), Mlp24p L-asparagine complex chain A (orange), Mlp24p glycine complex (green), Mlp24p L-proline complex (purple) and Mlp24p L-serine complex chain A (cyan) are superimposed onto that of Mlp24p chain A (gray). The residues involved in the ligand recognition or the Ca<sup>2+</sup> coordination are shown in stick models and labeled. Ca<sup>2+</sup> is shown as a cyan ball.



**Figure 3-10. Binding of  $\text{Ca}^{2+}$  and asparagine to the periplasmic fragment of Mlp24 (Mlp24pNH)**

ITC measurements with  $10 \mu\text{M}$  Mlp24pNH were carried out with  $10 \text{ mM}$   $\text{CaCl}_2$  (A) or  $10 \text{ mM}$  L-asparagine in the absence (B) or presence (C) of  $10 \text{ mM}$   $\text{CaCl}_2$ . Enthalpy changes per mol are plotted as a function of the molar ratio of  $\text{CaCl}_2$  or L-asparagine to Mlp24pNH.



**Figure 3-11. Chemotactic responses of *Vibrio cholerae* cells overexpressing to various amino acids**

Vmlp201 ( $\Delta mlp24 \Delta mlp37$ ) cells carrying pMlp24 (the Mlp24-FLAG-expressing plasmid, circles) or pAH901 (the empty vector, triangles) were subjected to capillary assays with calcium (A) and with L-asparagine (B), L-serine (C) and L-arginine (D) in the absence (open symbols, dotted lines) or presence (closed symbols, solid line) of 10 mM CaCl<sub>2</sub>.

**Table 3-3 Ligand binding affinities of the Mlp24 periplasmic fragment to various amino acids and divalent cations**

Ligands	$K_d$ ( $\mu\text{M}$ )	
	$- \text{Ca}^{2+}$	$+ \text{Ca}^{2+}$
L-serine	184.2	18.3
L-asparagine	49.5	6.3
L-arginine	5.7	4.8
L-alanine	40.0	5.1
glycine	92.6	9.0
L-proline	79.4	23.8
L-methionine	24.9	4.7
L-valine	212.8	22.1
L-glutamine	28.7	5.5
L-lysine	9.8	6.8
L-glutamate	NT <sup>1)</sup>	no binding
$\text{Ca}^{2+}$	452.5	NT
$\text{Mg}^{2+}$	no binding	NT

<sup>1)</sup>NT, not tested.



### **3-4-2. Structural basis of the binding affinity of Mlp24p and Mlp37p for various amino acids**

I have determined a series of the ligand complex structures of Mlp24p and Mlp37p and revealed the molecular basis of their affinity and specificity for various ligands. Mlp24p and Mlp37p recognize the backbone of ligand amino acids in the same manner (Fig. 2-3, 3-5 and 3-8). However, they recognize the ligand side chains in different ways, although the residues involved in the ligand recognition are conserved in both proteins (Fig. 3-1, 2-3, 3-5 and 3-8). The most remarkable structural difference between Mlp24p and Mlp37p was found in the UJ loop, which forms the upper wall of the ligand binding pocket (Fig. 2-8, 3-9).

The crystal structures of the Mlp37p ligand complexes revealed that Mlp37p adjusts the size and shape of the ligand-binding pocket to accommodate each ligand, thereby tightly binds the ligands with similar dissociation constants. The UJ-loop of Mlp37p changes its conformation according to each ligand side chain and directly interacts with the side chain atoms (Fig. 2-8). W141 in the UJ-loop moves to adequate positions to make a cation- $\pi$  interaction with L-arginine or a hydrogen bond to L-serine (Fig. 2-3). This conformational flexibility of the UJ-loop of Mlp37p is provided by P143, G144 and P145. The dihedral angles of these residues vary depending on ligands, but those of other residues do not show any significant changes. These three residues are not conserved in Mlp24 and may play a key role in recognizing ligand side chains by Mlp37 with similar dissociation constants.

The crystal structures of the Mlp24p indicated that the N-terminal region of the UJ loop has a unique turn structure stabilized by a calcium ion (Fig. 3-9). This turn structure restricts the conformation of the N-terminal region of the UJ loop and determines the orientation of the

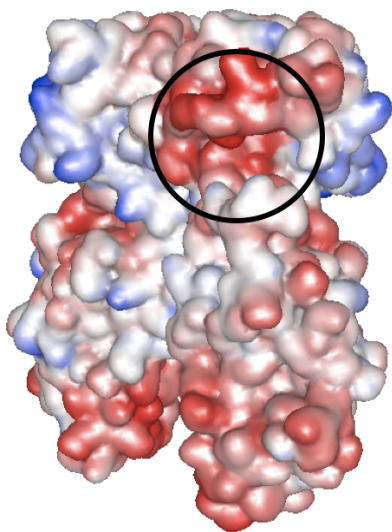
UJ-loop (Fig. 2-8 and 3-9). Therefore, the calcium ion is thought to be important for ligand binding.

The ITC measurements confirmed that  $\text{Ca}^{2+}$  affects the binding affinity of Mlp24 for the ligands. The binding affinities for small amino acids are low without  $\text{Ca}^{2+}$ , but significantly increased by the addition of  $\text{Ca}^{2+}$  (Table 3-3). G113 is located at the corner of the turn, and the conformational flexibility of the glycine residue is essential to form the turn structure around the calcium ion (Fig. 3-5A). Without the bound  $\text{Ca}^{2+}$ , the turn structure may be disrupted and the orientation of the UJ loop may not be well determined because of the high conformational flexibility of G113. Consequently, the upper wall of the pocket is not properly formed and the ligands do not stably bind to Mlp24p. L-arginine, however, strongly binds to Mlp24p even in the absence of  $\text{Ca}^{2+}$ . This is probably because that the guanidino group of arginine is recognized mainly by an electrostatic interaction with D145 and, therefore,  $\text{Ca}^{2+}$  would not significantly alter the affinity for L-arginine. By contrast, the N-terminal conformation of the UJ loop is restricted by P139 in Mlp37. Therefore the upper wall of the pocket of Mlp37p is formed properly without  $\text{Ca}^{2+}$ .

ITC measurements have indicated that Mlp24p still shows different affinities for the ligands in the presence of 10 mM  $\text{Ca}^{2+}$ . In contrast to Mlp37p, the UJ-loop of Mlp24p adopts almost the same conformation in all complex structures. The size and the shape of the pocket of the complex structures are very similar (Fig. 3-9). The pocket of the Mlp24p nicely accommodates L-arginine but is rather big for smaller amino acid. Therefore the binding affinities of Mlp24p for smaller amino acids are relatively low except for alanine.

Taurine binds to Mlp37 but not to Mlp24 (2). The sulfonate group of taurine in Mlp37p hydrogen bonds to W141 through water molecules (Fig. 2-3D). W114 of Mlp24p, which corresponds to W141 of Mlp37p, is too distant to form hydrogen bond to the sulfate group of taurine, and hence Mlp24p does not bind taurine.

Electrostatic surface charge distribution of Mlp24p indicates that the entrance of the binding pocket is negatively charged (Fig. 3-12). The negatively charged surface may attract positively charged amino acids, such as L-arginine and L-lysine, and repel negatively charged amino acids, such as L-glutamate and L-aspartate. This is probably another reason why Mlp24 strongly binds amino acids with basic side chains and do not bind those with acidic side chains (Table 3-3).



**Figure 3-12. Electro static potential surface of the apo-Mlp24p dimer**

Positively charged regions are colored in blue and negatively charged regions in red. The entrance of the ligand binding pocket is indicated by a circle.

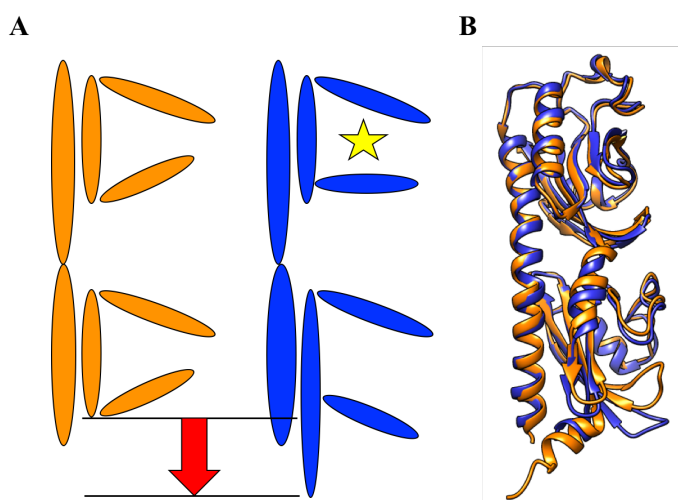
### **3-4-3. Conformational change of Mlp24p upon ligand binding**

The ligand-free and ligand-bound structures of Mlp24p indicate that the ligand binding induces the domain closure of CAD1. Similar conformational changes by ligand binding have been observed in other dCACHE type sensors, such as Tlp3 (14) and sensor kinase vpHK1<sub>s</sub>-Z8 (15). On the basis of the structures of Mlp24p, I propose a plausible mechanism of ligand binding by Mlp24. The ligand amino acid enters the pocket of CAD1 in open conformation. The carboxy group of the ligand is bound to R125 in H4, which is the first step of ligand binding, and the interaction is further stabilized by Y120 and W127. This idea is supported by the acetate binding to CAD1 of the subunit in open conformation. Then the amino group of the ligand backbone interacts with D172 in S6 of the central  $\beta$ -sheet to adjust the orientation of the ligand. The positive charge of the ligand amino group induces the conformational change of the LJ segment to bind to Y143 and D145. Finally, the entrance of the pocket is closed and the ligand side chain atoms are recognized by the residues in the pocket. Essentially the same mechanism may be exerted by Mlp37.

#### **3-4-4. Implication for the signal transduction mechanism of Mlp24p**

The ligand-binding signal should be transmitted into the cytoplasmic domain, but the mechanism is still under debate. Structural studies of the ligand-free and -bound structures of various dCACHE-type receptors have proposed possible structural changes involved in transmembrane signaling. Piston-like movements along the N-terminal helix have been proposed to be important for the signal transduction from the structures of Z2-Bistris complex and apo Z3 (Fig. 3-13) (15). However, such movement was not observed in the ligand-free and ligand-bound structures of Mlp24p. The relative movement of the two CACHE domains caused by the bending of H1 is proposed to be another candidate involved in the signal transduction (15). However, the H1 bending in Mlp24p occurs independent of ligand binding. The ligand free- and L-arginine and L-asparagine complex structures of Mlp24p show almost the same bending conformation in H1 and the same relative position of the two CACHE domains. On the contrary, other ligand-bound structures, such as glycine, L-proline and L-serine complexes, show different relative positions of the two CACHE domains and different H1 conformations. Therefore, ligand binding to Mlp24p does not necessarily induce the movement of the CACHE domain. However, the ligand free-Mlp24p subunit actually forms a hetero dimer with a subunit that binds L-alanine in CAD1. Thus, the hetero dimer can be regarded as a ligand-bound state. A recent study on Tlp3 (14) proposed that ligand binding to the distal domain induces opening of the proximal domain followed by the displacement of the C-terminal helix towards the membrane, which may cause a piston-like movement of the TM helix. This is, however, not the case in Mlp24. Mlp24p does not show any structural changes in CAD2 and the C-terminal helix by binding of ligands (Fig. 3-6). No significant

structural difference was observed between the ligand free-Mlp24p and the L-arginine or the L-asparagine complex except for the LJ segment of CAD1. In contrast, the ligand complex structures show large conformational variation in the dimer structures. This structural variation may be artifact caused by the crystal packing or the lack of membrane tethering. To elucidate the signal transduction mechanism across the membrane, both ligand-free and ligand-bound structures of the full-length protein or a larger fragment containing the membrane spanning region are needed.



**Figure 3-13. Advocated signal transduction mechanism of dCACHE-type chemoreceptor**

(A) Cartoon model of conformational changes (piston-like movement) associated with ligand binding. The left side (orange model) shows the apo-state and the right side (blue model) shows the ligand binding-state. The movement is shown in red arrow. The ligand is labeled as a yellow star. (B) The stick model of the apo-state structure (orange) superposed onto the ligand binding-state structures (blue).

### **3-4-5. *Vibrio cholerae* may fine-tune its chemotactic behavior by calcium ion concentration.**

Mlp24 binds calcium ion with the  $K_d$  values of about 0.45 mM. The concentration of  $\text{Ca}^{2+}$  is about 10 mM in seawater and in gastrointestinal tract (16). Therefore 96 % of the Mlp24 molecules bind  $\text{Ca}^{2+}$  in seawater and in gastrointestinal tract, whereas 50 to 80 % of the Mlp24 molecules release  $\text{Ca}^{2+}$  in fresh water ( $\text{Ca}^{2+}$  concentration: 0.025–0.5 mM).  $\text{Ca}^{2+}$  increases the binding affinity of Mlp24 to smaller amino acids to the levels close to those of L-arginine and L-lysine and, as a result, Mlp24 exerts broader substrate specificity. The chemotactic behavior of *V. cholerae* mediated by Mlp24 is also affected by calcium ion. Therefore, it might be possible that *V. cholerae* fine-tunes its chemotactic behavior in response to environmental calcium concentrations by modulating the ligand specificity of Mlp24, although it is still unclear that the physiological role of the calcium dependency of the amino acid sensing. It has been known that calcium affects biofilm formation, cell adhesion and other virulence activities in *V. cholerae* as well as other *Vibrio* species (17-22).

I compared the amino acid sequence of Mlp24 with its closely related homologues around the N-terminal region of the UJ-loop. Interestingly, the residues important for  $\text{Ca}^{2+}$  binding (E109 and E115, which coordinate to  $\text{Ca}^{2+}$  and G113, which is important to form the turn structure) are conserved only in putative MLPs in marine bacteria, such as VPA0511 of *Vibrio parahaemolyticus* and N646\_4651 of *Vibrio alginolyticus* (Fig. 3-14). As described above, seawater contains 10 mM of calcium ion. The MLPs of marine bacteria may bind calcium ion to form a specific loop structure like Mlp24.

			▼	*	▼	
<b>VcMlp24</b>	<b>98</b>	<b>GLGYESDATVV</b>	<b>ENDD</b>	<b>GW</b>	<b>EPNA</b>	<b>118</b>
VcMlp37C1	97	GFGLEKDGSNINNDPSWNP	GP			117
VcMlp37E1	125	GFGLEKDGSNINNDPSWNP	GP			145
VaN646_4651	98	GFGYESNGFVI	ENDD	GW	DAGP	118
VpVPA0511	98	GFGYESNGFVI	ENDD	GW	DAGP	118
PaPctA	99	YLGQOD-GVFTMRPDSPMP	-A			119
PaPctB	99	YLGQTD-GTYTARPTSDLP	-A			119
PaPctC	102	YLGEAASGTFTMRPYDAMP	-E			122
VcMlp8	100	GYADERTGEIITNDPNFKVPT				120
BsMcpB	111	IYGGADNGTYVQAPKEKLP	-E			131

**Figure 3-14. Sequence alignment of the variety of MCPs and MLPs**

The sequence around the UJ loop of Mlp24 is compared with other MCPs and MLPs. Residues that are identical and homological to that of Mlp24 are highlighted in grey. Residues that are important for Ca<sup>2+</sup> coordinating (E109 and E115) and forming the turn structure (G113) are highlighted in red and indicated by arrowhead and asterisk respectively. Cl, classical biotypes; El, El Tor biotypes; Vc, *Vibrio cholerae*; Vp, *Vibrio parahaemolyticus*; Pa, *Pseudomonas aeruginosa*; Bs, *Bacillus subtilis*.



### 3-5. References

1. Nishiyama S, Suzuki D, Itoh Y, Suzuki K, Tajima H, Hyakutake A, Homma M, Butler-Wu SM, Camilli A, Kawagishi I. 2012. Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infect Immun* 80:3170-8.
2. Nishiyama S, Takahashi Y, Yamamoto K, Suzuki D, Itoh Y, Sumita K, Uchida Y, Homma M, Imada K, Kawagishi I. 2016. Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci Rep* 6:20866.
3. Hung DT, Zhu J, Sturtevant D, Mekalanos JJ. 2006. Bile acids stimulate biofilm formation in *Vibrio cholerae*. *Mol Microbiol* 59:193-201.
4. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67:271-81.
5. Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* 66:125-32.
6. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235-42.
7. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a

- comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-21.
8. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.
  9. Machuca MA, Liu YC, Beckham SA, Gunzburg MJ, Roujeinikova A. 2016. The crystal structure of the tandem-PAS sensing domain of *Campylobacter jejuni* chemoreceptor Tlp1 suggests indirect mechanism of ligand recognition. *J Struct Biol* 194:205-13.
  10. Gavira JA, Ortega A, Martin-Mora D, Conejero-Muriel MT, Corral-Lugo A, Morel B, Matilla MA, Krell T. 2018. Structural Basis for Polyamine Binding at the dCACHE Domain of the McpU Chemoreceptor from *Pseudomonas putida*. *J Mol Biol* 430:1950-1963.
  11. Glekas GD, Foster RM, Cates JR, Estrella JA, Wawrzyniak MJ, Rao CV, Ordal GW. 2010. A PAS domain binds asparagine in the chemotaxis receptor McpB in *Bacillus subtilis*. *J Biol Chem* 285:1870-8.
  12. Hothorn M, Dabi T, Chory J. 2011. Structural basis for cytokinin recognition by *Arabidopsis thaliana* histidine kinase 4. *Nat Chem Biol* 7:766-8.
  13. Machuca MA, Johnson KS, Liu YC, Steer DL, Ottemann KM, Roujeinikova A. 2017. *Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate. *Sci Rep* 7:14089.
  14. Liu YC, Machuca MA, Beckham SA, Gunzburg MJ, Roujeinikova A. 2015. Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-

- Sim (tandem PAS) chemoreceptor sensory domains. *Acta Crystallogr D Biol Crystallogr* 71:2127-36.
15. Zhang Z, Hendrickson WA. 2010. Structural characterization of the predominant family of histidine kinase sensor domains. *J Mol Biol* 400:335-53.
  16. Dintzis FR, Laszlo JA, Nelsen TC, Baker FL, Calvert CC. 1995. Free and total ion concentrations in pig digesta. *J Anim Sci* 73:1138-1146.
  17. Bilecen K, Yildiz FH. 2009. Identification of a calcium-controlled negative regulatory system affecting *Vibrio cholerae* biofilm formation. *Environ Microbiol* 11:2015-29.
  18. Garrison-Schilling KL, Grau BL, McCarter KS, Olivier BJ, Comeaux NE, Pettis GS. 2011. Calcium promotes exopolysaccharide phase variation and biofilm formation of the resulting phase variants in the human pathogen *Vibrio vulnificus*. *Environ Microbiol* 13:643-54.
  19. Gode-Potratz CJ, Chodur DM, McCarter LL. 2010. Calcium and iron regulate swarming and type III secretion in *Vibrio parahaemolyticus*. *J Bacteriol* 192:6025-38.
  20. Hay AJ, Yang M, Xia X, Liu Z, Hammons J, Fenical W, Zhu J. 2017. Calcium Enhances Bile Salt-Dependent Virulence Activation in *Vibrio cholerae*. *Infect Immun* 85.
  21. Jones GW, Abrams GD, Freter R. 1976. Adhesive properties of *Vibrio cholerae*: adhesion to isolated rabbit brush border membranes and hemagglutinating activity. *Infect Immun* 14:232-39.

22. Kierek K, Watnick PI. 2003. The *Vibrio cholerae* O139 O-antigen polysaccharide is essential for Ca<sup>2+</sup>-dependent biofilm development in sea water. Proc Natl Acad Sci U S A 100:14357-62.

## Chapter 4

### Structure of *V. cholerae* accessory colonisation factor, AcfC, and its interaction with Mlp8

#### 4-1. Introduction

As described in section 1-5, Mlp7, Mlp8, Mlp24 and Mlp30 are related to pathogenicity of *V. cholerae* and Mlp37 is involved in biofilm formation. Mlp24 and Mlp37 are dCACHE-type chemoreceptors for various amino acids (2). Mlp8 is also considered to be a dCACHE-type chemoreceptor because the periplasmic domain of Mlp8 (Mlp8p) shows amino acid sequence identities of 25% to Mlp24p and 26% to Mlp37p. Mlp8 is involved in chemotactic response to mucus and galactose-6-sulfate (Table 4-1), but not to amino acids, such as serine (3). Mlp8 is therefore expected to recognize ligands in a different manner from Mlp24 and Mlp37. DNA sequence analysis of *V. cholerae* has revealed that *mlp8* forms an operon with *acfC*. AcfC protein has a signal peptide at the N-terminus and does not have a membrane-spanning domain. The protein subcellular localization prediction analysis for Gram-native bacteria showed that AcfC is localized in periplasm (4). Thus AcfC is expected to be a periplasmic protein. AcfC has high amino acid sequence similarity to Porcine attaching-effacing associated (Paa) proteins of *Escherichia coli*, Paa proteins of the O157:H7 EDL933 Sakai *E. coli* strains and PEB3, an adhesin of *Campylobacter jejuni* (Fig. 4-1) (5). In addition, a recent research has revealed that AcfC is required for *V. cholerae* chemotaxis towards mucin but does not affect chemotaxis towards chitin (Fig. 4-2) (4). These results suggest that AcfC is involved in the pathogenicity and/or chemotaxis together with Mlp8.

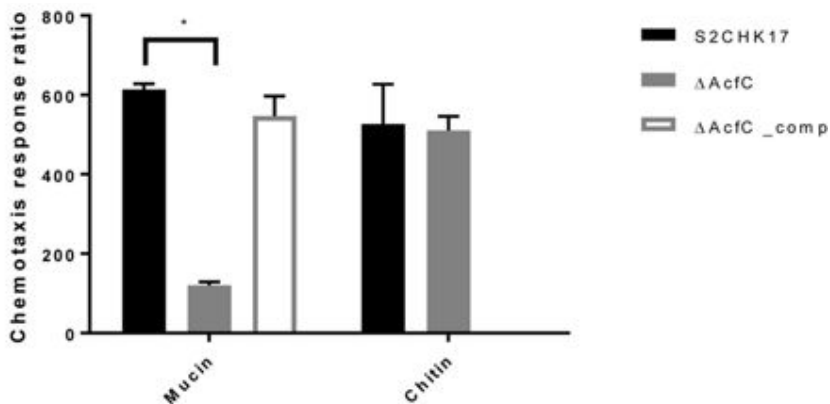
The structure of AcfC in complex with D-malate has been determined (PDB ID: 4JB7) by a structural genomics group (Fig. 4-3), but the ligand recognition mechanism of AcfC and the interaction with Mlp8 remain unclear. To reveal these questions, I determined the structure of ligand-free AcfC at 2.31 Å, and performed isothermal titration calorimetry (ITC) and pull-down experiments.

**Table 4-1 Chemotactic response of *V. cholerae* to chemoattractants**

Strain	Serine	Mucus	Galactose-6-sulfate
Wild type	18	12	12
<i>Δmlp8</i>	17	2	4

The ratio of vibrio accumulation in chemoattractant-containing capillaries compared with capillaries containing buffer alone.

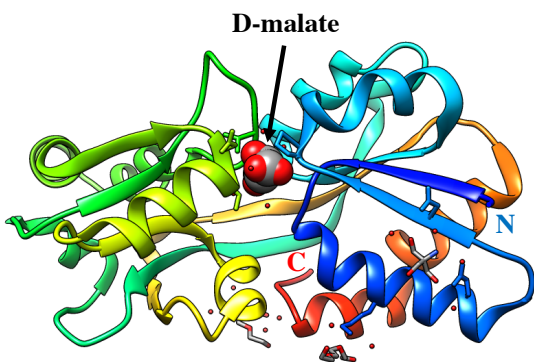




**Figure 4-2. *V. cholerae* chemotaxis assay (4)**

Chemotaxis assay with *V. cholerae* S2CHK17,  $\Delta acfC$  and AcfC complement strain towards mucin and chitin. The values are the means  $\pm$  standard deviation (n = 3). \*p<0.05, as determined by Tukey's test.

(Valiente E *et al.* (2018) *Sci Rep* 8:8390. Figure 3b.)



**Figure 4-3. The structure of AcfC in complex with D-malate**

Ribbon representation of the AcfC D-malate complex structure (PDB ID: 4JB7) is colored in rainbow form from the N-terminus (blue) to the C-terminus (red). D-malate is shown as a ball model. The nitrogen and oxygen atoms of D-malate are colored in blue and red, respectively. The bound water molecules are shown in small red balls.



#### **4-2-1. Expression and purification of AcfC**

*E. coli* strain BL21(DE3) (Novagen) carrying a plasmid expressing AcfC (23-253) with an N-terminal His×6 tag followed by a thrombin cleavage site was cultured in LB broth Lennox (Nacalai Tesque, Inc., Kyoto Japan) containing 50 µg/ml of ampicillin at 37°C until the cell density had reached an OD<sub>600</sub> of about 0.5. Isopropyl β-d-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.3 mM to induce protein expression, and the culture was continued for 16 hours at 18°C. Cells were harvested by centrifugation, suspended in Suspend buffer (50 mM of Tris and 300 mM NaCl at pH 8.0) and disrupted by sonication. After removing cell debris by centrifugation, the cell lysate was loaded onto a His Trap HP 5 mL column (GE Health care) followed by washing with Wash buffer (50 mM of Tris, 300 mM NaCl and 20 mM of imidazole at pH 8.0). Proteins were eluted with a 50-300 mM imidazole gradient in wash buffer and fractions containing AcfC were collected. The His×6 tag was removed by thrombin (GE Healthcare) to the AcfC solution. The reaction mixture was dialyzed in 1.0 L dialysis buffer (50 mM Tris and 300 mM NaCl at pH 8.0) at 4°C for 12 hours using Spectra/por Dialysis Membrane MWCO 6,000-8,000 (Spectrum Laboratories, Inc.). The protein solution was loaded onto a Ni-NTA agarose resin (QIAGEN) to remove the N-terminally His-tagged polypeptide, non-cleaved His×6-fused AcfC and thrombin, and further purified by size exclusion chromatography with a High Load 26/60 Superdex 200 column in 20 mM Tris and 150 mM NaCl at pH 8.0. The peak fraction was collected and concentrated. The purity of the purified proteins was examined by SDS-PAGE.

#### **4-2-2. Expression and purification of Mlp8p with C-terminal hexahistidine tag (Mlp8-His6)**

Strain BL21(DE3) (Novagen) carrying pVCP8, which expresses Mlp8p (28-277) with N-terminal GST tag and C-terminal His $\times$ 6 tag, was cultured in LB broth Lennox (Nacalai Tesque, Inc., Kyoto Japan) containing 50  $\mu$ g/ml of ampicillin at 37°C until the cell density had reached an OD<sub>600</sub> of about 0.6. Isopropyl  $\beta$ -d-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.1 mM to induce protein expression, and the culture was continued for 12 hours at 20°C. Cells were harvested by centrifugation, suspended in phosphate buffered saline (PBS) (137 mM of NaCl, 2.7 mM of KCl, 10 mM of Na<sub>2</sub>HPO<sub>4</sub>•12H<sub>2</sub>O, and 1.8 mM of KH<sub>2</sub>PO<sub>4</sub> at pH 7.4) and disrupted by sonication. After removing cell debris by centrifugation, the cell lysate was loaded on a Glutathione-Sepharose<sup>TM</sup> 4B column (GE Healthcare) followed by washing with PBS. Proteins were eluted with 50 mM Tris-HCl (pH 8.0) buffer containing 10 mM reduced glutathione. The N-terminal GST tag was then cleaved using PreScission<sup>TM</sup> Protease (GE Healthcare), and the reactant was dialysed against 1.0 L dialysis buffer (50 mM Tris-HCl pH 7.0, 150 mM NaCl, 1 mM DTT, and 1 mM EDTA) at 4°C for 12 hours using Spectra/por Dialysis Membrane MWCO 6,000-8,000 (Spectrum Laboratories, Inc.). The protein solution was loaded again to a Glutathione-Sepharose<sup>TM</sup> 4B column (GE Healthcare) to remove GST and unreacted protein, and further purified by size exclusion chromatography with a High Load 26/60 Superdex 200 column in 20 mM Tris-HCl and 150 mM NaCl (pH 8.0). The peak fraction was collected and concentrated to 10 mg/mL. The purity of the purified proteins was examined by SDS-PAGE.

### 4-2-3. Crystallization of AcfC

Crystallization of AcfC was performed by the hanging-drop vapor-diffusion technique. Crystallization drops were prepared by mixing 1.0  $\mu$ L of AcfC solution (15.3 mg/ml) with equal volume of a reservoir solution. Screening was carried out using the following screening kit: Wizard Classic I and II (Rigaku Reagents, Inc.), Cryo I and II (Rigaku Reagents, Inc.) and Crystal Screen I and II (Hampton Research). Crystals suitable for X-ray analysis were obtained from Crystal Screen I No. 22 (30% (w/v) PEG-4,000, 100 mM Tris and 200 mM sodium acetate at pH 8.6) within a week. This crystal belongs to an orthorhombic space group of  $P 2_12_12_1$  with unit cell dimensions of  $a = 58.2 \text{ \AA}$ ,  $b = 85.1 \text{ \AA}$  and  $c = 102.8 \text{ \AA}$ . The crystallization condition, the space group and the unit cell dimensions of AcfC are different from that of AcfC in complex with D-malate (Table 4-2).

**Table 4-2 The crystallization conditions of AcfC**

	AcfC	AcfC D-malate complex (PDB ID: 4JB7)
Protein Buffer	10 mM Tris pH 8.5 250 mM NaCl	10 mM Tris pH 8.0 250 mM NaCl 5 mM BME
Protein concentration	15.3 mg/mL	7.3 mg/mL
Reservoir	Buffer	100 mM Tris pH 8.5
	Precipitant	30 % (w/v) PEG 4,000
	Additive	200 mM NaAc
Unit cell dimension	$a = 58.21 \text{ \AA}$ , $b = 85.12 \text{ \AA}$ , $c = 102.82 \text{ \AA}$	$a = 72.34 \text{ \AA}$ , $b = 72.34 \text{ \AA}$ , $c = 87.88 \text{ \AA}$
Space group	$P 2_12_12_1$	$P 4_12_12$

#### **4-2-4. X-ray data collection and structure determination of AcfC**

The crystals were soaked in solution containing 90% (v/v) of the reservoir solution and 10% (v/v) MPD for a few seconds and transferred into liquid nitrogen for freezing. X-ray diffraction data were collected at 95 K under nitrogen gas flow at the synchrotron beamline BL41XU of SPring-8 (Harima, Japan), with the approval of the Japan Synchrotron Radiation Research Institute (JASRI). The data were processed and scaled with XDS (7). Initial phase was calculated by molecular-replacement using the Phenix programme suite (8) using the AcfC D-malate complexed structure (PDB ID: 4JB7) as a search model. The atomic model was constructed with Coot (9) and refined with Phenix to 2.31 Å. During the refinement process, iterative manual modifications were performed. The final refinement R factor and the free R factor were 19.1% and 23.5%, respectively. The Ramachandran plot indicated that 95.1%, 4.5% and 0.4% residues were located in the preferred, allowed and outliers regions, respectively. Data collection and refinement statistics are summarized in Table 4-3.

**Table 4-3 Summary of X-ray data collection and refinement statistics**

<b>Data collection</b>		<b>Refinement statistics</b>	
Wavelength (Å)	1.0000	Resolution (Å)	44.0-2.31 (2.40-2.31)
Resolution (Å)	44.0-2.31 (2.45-2.31)	Number of reflections	
$R_{\text{merge}}$ (%)	14.8 (57.3)	Working set	22,142 (2,703)
$I/\sigma I$	1.40 (4.81)	Test set	1,165 (142)
Completeness (%)	90.4 (89.3)	$R_w$ (%)	19.13 (23.42)
Redundancy	2.1 (2.0)	$R_{\text{free}}$ (%)	23.48 (28.63)
Content per ASU	2	R. m. s. deviations	
		Bond length (Å)	0.003
		Bond angle (°)	0.708
		B-factors	
		Protein atoms	32.46
		Solvent atoms	26.94
		Ramachandran plot (%)	
		In preferred regions	95.09
		In allowed regions	4.49
		Outliers	0.43
		Number of atoms	
		Protein	3,700
		Solvent	52

Values in parentheses are for the highest resolution shell.

$$R_w = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}, R_{\text{free}} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$$

#### **4-2-5. Pull-down assay**

Binding of Mlp8p-His6 with AcfC was examined by a pull down assay with Ni-NTA agarose resin (QIAGEN). Mlp8p-His6 (1.0 mg/mL) and AcfC (1.0 mg/mL) were mixed with 2.0 M of D-malate, 1.0 M of galactose-6-sulfate or pull-down buffer containing 20 mM Tris pH 8.0 and 150 mM NaCl. The sample solutions were diluted with pull-down buffer to a protein concentration of 0.48 mg/mL each. The final concentration of D-malate was 95mM and that of galactose-6-sulfate was 48 mM. Each sample solution was mixed with 1.0 ml of the Ni-NTA agarose resin and incubated at 4°C for 1h. The mixture was loaded onto a Poly-Prep® column (Bio-Rad), and non-bound sample was removed by gravity fall and following washing with 10 ml of the pull-down buffer. After washing, the bound sample was eluted with elution buffer containing 20 mM Tris pH 8.0, 150 mM NaCl and 500 mM Imidazole. The sample was analysed by SDS PAGE with Coomassie blue staining.

#### **4-2-6. Isothermal titration calorimetry (ITC)**

Calorimetric titration experiments were performed using a MicroCal™ iTC<sub>200</sub> System (Malvern Panalytical Ltd.). During titration, the sample was held at 25°C in a stirred (750 rpm) reaction cell (200 µl). An injection series (0.4 µl for the first injection and 2.0 µl eac for later injections) was carried out using a syringe filled with D-malate solution (19 injections with 150 seconds intervals). Data points were averaged and stored at 5 seconds intervals. The data were analyzed using an NanoAnalyze™ Software (TA Instruments, Inc.).

## 4-3. Results

### 4-3-1. Structure of ligand-free AcfC

The crystal structure of ligand-free AcfC was determined at 2.31 Å resolution (Fig. 4-4 and Table 4-3). The final model includes AcfC (23-253) with N-terminal leftover residues (GSHM) of His tag. The crystallographic asymmetric unit contains two AcfC molecules, which are very similar with the root mean square deviations of 0.371 Å.

AcfC is composed of ten  $\alpha$ -helices (H1-H10), nine  $\beta$ -strands (S1-S9) including a long  $\beta$ -strand (S4) and loops (Fig. 4-4). The overall structure of ligand-free AcfC is very similar to the structure of the AcfC D-malate complex (PDB ID: 4JB7) (10). AcfC folds into two discontinuous domains, domain 1 (G19-M100 and Y214-K253) and domain 2 (R101-I213), and a cleft is formed between the two domains. The two domains are structurally similar to each other. Both domains have a central five-stranded  $\beta$ -sheet. The long  $\beta$ -strand (S4) contributes to the formation of the central  $\beta$ -sheets of both domains. The strand order of domain 1 is S2-S1-S3-S9-S4 with S9 running antiparallel to the other four strands, and that of domain 2 is S6-S5-S7-S4-S8 with S4 running antiparallel to the other four strands. These structural features indicate that AcfC belongs to the classic type II periplasmic-binding protein family (11).

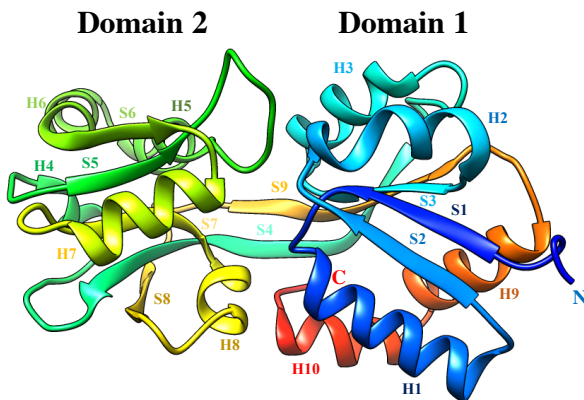
The crystal structure of ligand-free AcfC formed homo-dimer like other classic type II periplasmic-binding protein, PEB3, and the AcfC D-malate complex structure (Fig. 4-5) (10). The dimer interaction is essentially identical to the AcfC D-malate complex structure. The AcfC dimer is stabilized by 14 hydrogen bonds between the monomers. The side chain of E76 in H3 hydrogen bonds to R162' in H6' (prime denotes the partner subunit) and the main

chain of E76 hydrogen bonds to E146' in H5'. The main chain oxygen atom of K92 make hydrogen bond to the side chain of R162' in H6' and the main chain oxygen atom of P97 in S4 to the side chain nitrogen atom of Q159' in H6'. The side chain of E146 in H5 hydrogen bonds to the main chain nitrogen atom of E76' and H77' in H3' and the side chain of E146 in H5 to 215R' in S9'.

The crystal of the AcfC D-malate complex (PDB ID: 4JB7) was prepared in MMT buffer, which contains D-malate (Table 4-2), and clear electron density corresponding to D-malate was appeared in the cleft between domain 1 and 2 (Fig. 4-6B). In this study, on the other hand, all solutions used in purification and crystallization of AcfC contained no D-malate (Table 4-2), and no electron density except for water molecules was present at the D-malate binding site (Fig. 4-6A), so I succeeded in solving the structure of ligand-free AcfC.

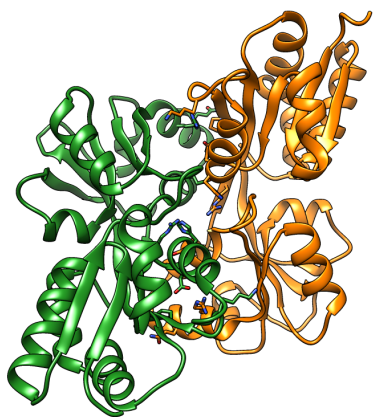
Although ligand binding induces a structural change and a type of a binding ligand influence a conformation of the ligand binding domain in other periplasmic-binding proteins such as PEB3 (10, 12-14), AcfC does not change the structure upon binding of D-malate (Fig. 4-7A). The structure of ligand-free AcfC is very close to that of the AcfC D-malate complex with the root mean square deviations of 0.471 Å for C $\alpha$  atoms.





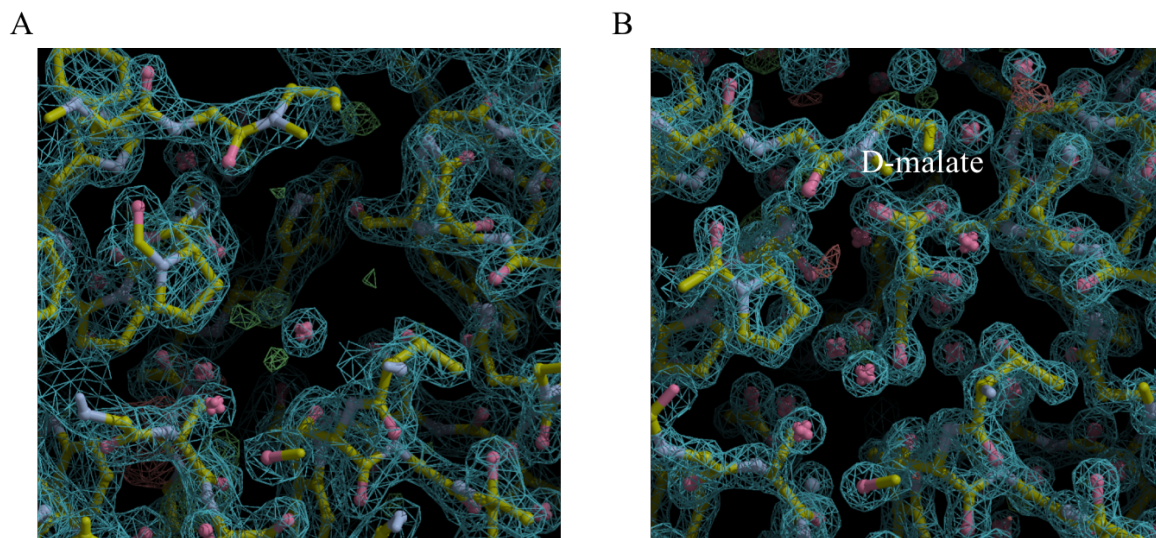
**Figure 4-4. Overall structure of ligand-free AcfC**

Ribbon model of ligand-free AcfC is colored in rainbow from the N-terminus (blue) to the C-terminus (red). The secondary structure elements and the two domains (Domain 1 and Domain 2) are labeled.



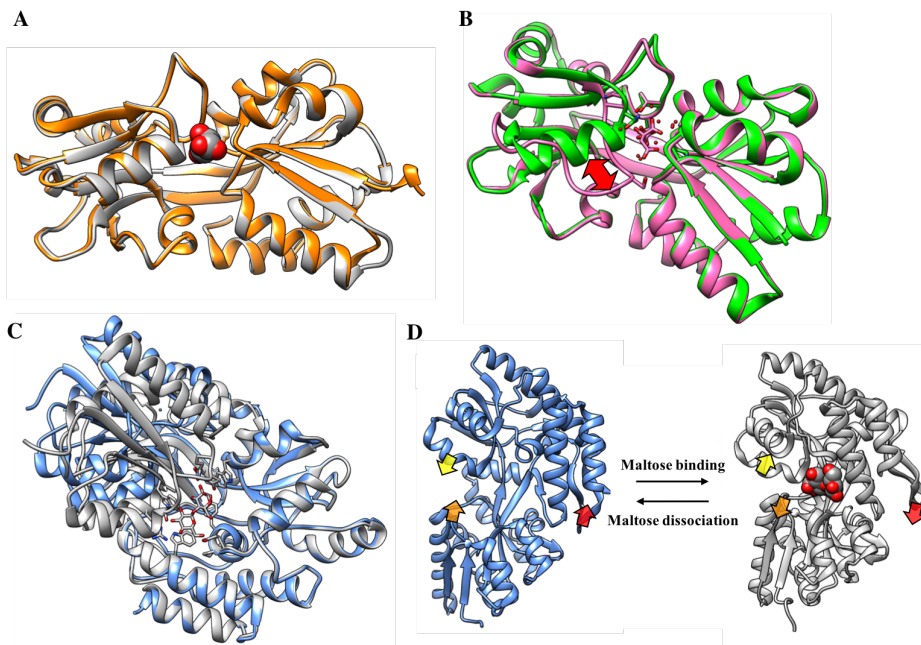
**Figure 4-5. Structure of the ligand-free AcfC dimer**

The dimer structure of ligand-free AcfC is presented by ribbon model. The subunits are colored in Green and Orange. Residues involved in the dimer interaction are showed in stick model and colored in red for oxygen and blue for nitrogen atoms.



**Figure 4-6.  $2Fo - Fc$  electron density maps of the D-malate binding site**

$2Fo - Fc$  maps of the D-malate binding sites of ligand-free AcfC (A) and of AcfC D-malate complex (B) are drawn at 1.1 sigma with the molecular models. The bound D-malate is labelled in (B). No significant density is observed in the D-malate binding site in (A).



**Figure 4-7. Structural comparison of the periplasmic-binding proteins with/without ligand**

(A) AcfC (orange) is superposed onto AcfC D-malate complexed (gray). D-malate is shown in ball model and colored in red for oxygen atoms. No significant difference is observed in the structure including D-malate binding site. (B) Conformational difference between the PEB3 phosphate binding structure (green) (PDB ID: 2HXW) and the citrate binding structure (pink) (PDB ID: 3FJM). The bound phosphate and citrate are shown in stick models and colored in red for oxygen atoms and orange for phosphorus atom. The conformational difference is shown in red arrow. (C) Structural difference between maltose binding protein with (736 to 1106) or without maltose (gray or blue, respectively) (PDB ID: 3MP6 or 1LLS). The bound maltose is shown in stick model and colored in red for oxygen atoms. (D) Maltose-mediated conversion from open to closed conformation. Maltose binding or dissociation induces structural change indicated with colored arrows (red, orange and yellow).

#### **4-3-2. D-malate weakly binds to AcfC in solution**

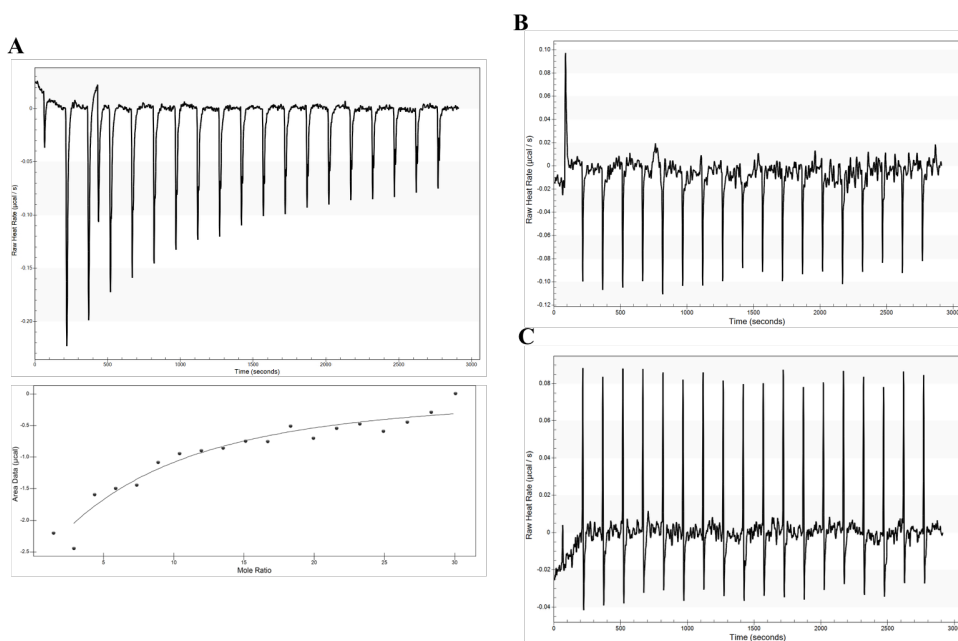
The previous structural study of AcfC has revealed that AcfC binds D-malate. However, D-malate has never been identified as a chemotaxis reagent of *V. cholerae*. The crystallization reagent of the AcfC D-malate complex contained significant amount of D-malate. Therefore I investigated the affinity of D-malate to AcfC using ITC. The titration curve of AcfC with D-malate show direct binding with a dissociation constant ( $K_d$ ) of 664  $\mu$ M, calculated by assuming that an AcfC monomer binds one D-malate molecule (Fig. 4-8A). In contrast, D-malate did not bind to Mlp8 (Fig. 4-8B). The  $K_d$  value of AcfC to D-malate is much higher than those of Mlp24 to ligand amino acids, indicating that binding affinity of D-malate to AcfC is rather low.

#### **4-3-3. The AcfC binding to Mlp8p is promoted by D-malate but not by galactose-6-sulfate.**

*AcfC* and *Mlp8* form an operon and both are involved in *V. cholerae* chemotaxis. Mlp8 is a chemoreceptor and AcfC belongs to the classic type II periplasmic-binding protein family. These facts raise a possibility that AcfC that binds the ligand binds to the periplasmic region of Mlp8. To detect the interaction between AcfC and Mlp8p, I performed a pull-down assay. AcfC did not co-elute with Mlp8p, However, in the presence of 95 mM of D-malate, a small amount of AcfC bound to Mlp8p (Fig. 4-9). These results suggest that Mlp8p recognizes D-malate with a help of AcfC.

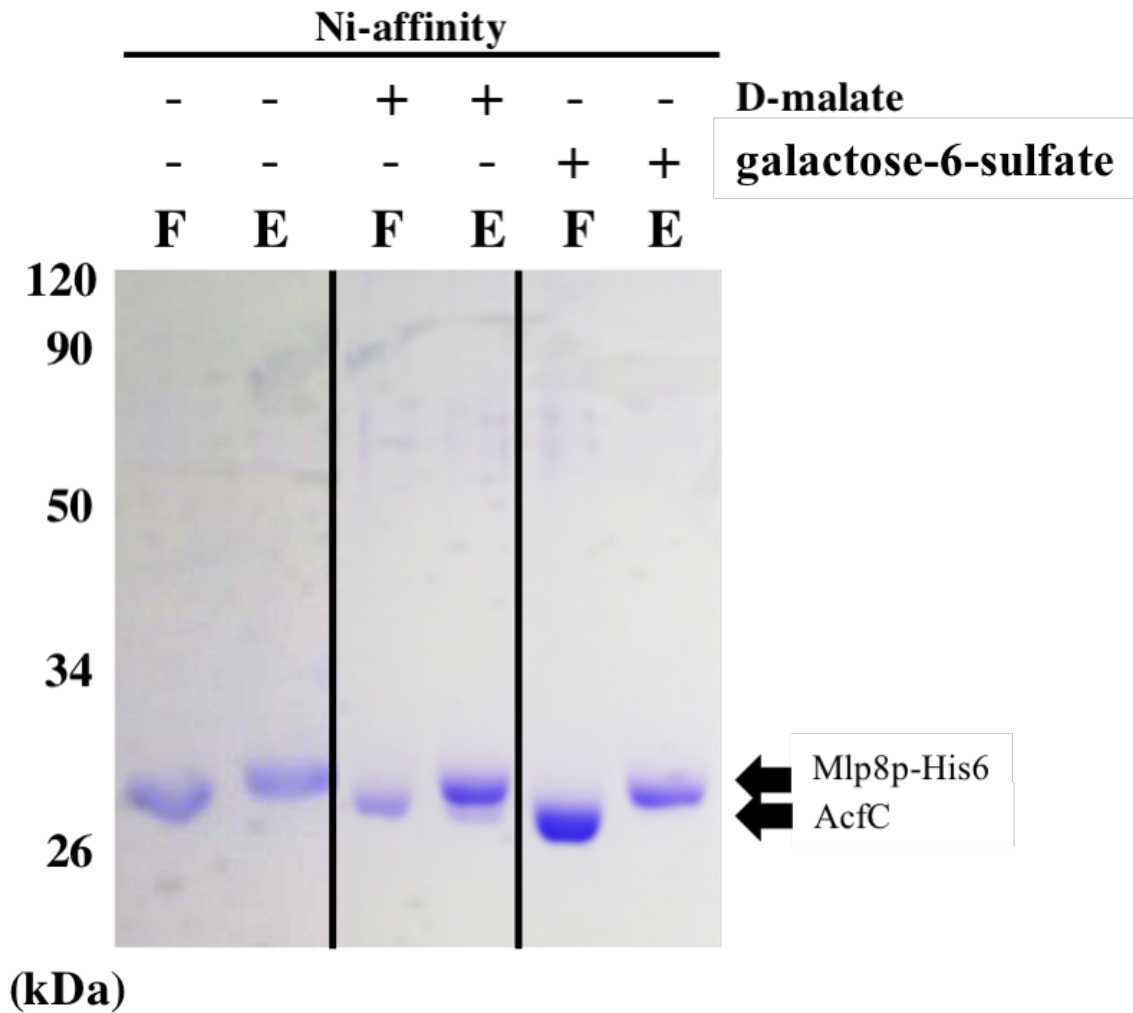
A recent study has revealed that Mlp8 is involved in chemotactic response to galactose-6-sulfate (Table 4-1) (3), but the recognition mechanism of galactose-6-sulfate is unclear.

Galactose-6-sulfate is expected to be recognized by Mlp8 by the same manner as D-malate. To examine this idea, I studied the binding of galactose-6-sulfate to AcfC using ITC. Galactose-6-sulfate, however, did not bind to AcfC (Fig. 4-8C). In agreement with this result, the pull-down experiments revealed that AcfC did not bind to Mlp8p in the presence of 47 mM of galactose-6-sulfate (Fig. 4-9), suggesting that Mlp8 recognizes galactose-6-sulfate in a way different from D-malate.



**Figure 4-8. Binding of D-malate or galactose-6-sulfate to AcfC or Mlp8-His6**

ITC measurements with 35  $\mu$ M AcfC were carried out with 5 mM D-malate (A) or 0.1 mM galactose-6-sulfate (B) and the measurement with 35  $\mu$ M Mlp8-His6 were carried out with 5 mM D-malate (C). The upper panel of (A) shows isotherms of enthalpic changes in AcfC upon D-malate binding. The lower panel of (A) shows the enthalpy changes per mol as a function of the molar ration of D-malate to AcfC.



**Figure 4-9. Interaction between Mlp8-His6 and AcfC**

*In vitro* pull down assays using Mlp8p-His6, AcfC and their ligands was carried out. Flow through fractions (F) and Elution fractions (with 500 mM imidazole) were analyzed by CBB (Coomassie brilliant blue) staining.

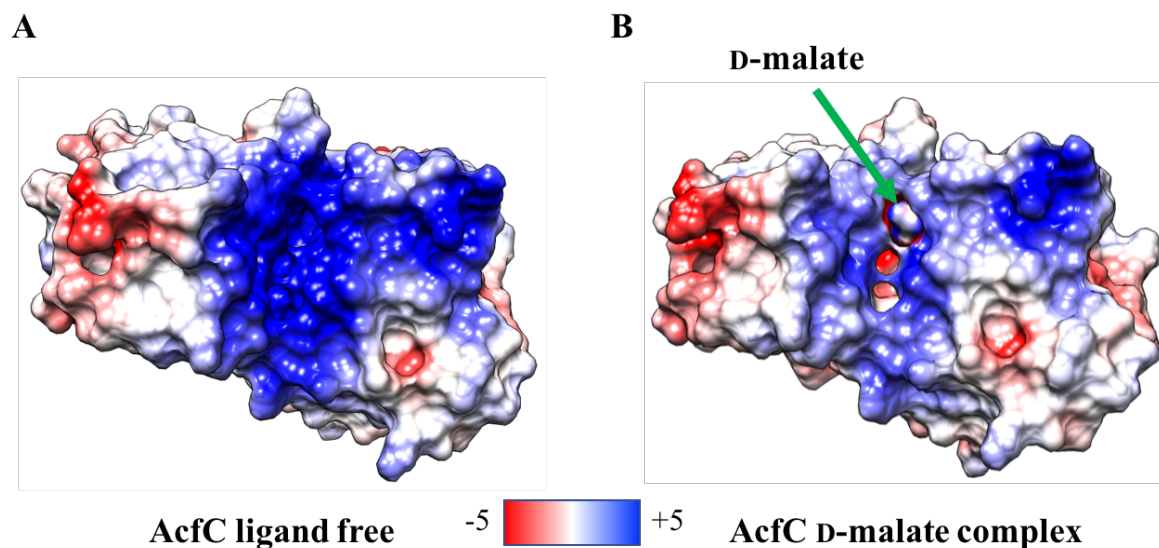
#### 4-4. Discussion

The previous structures of periplasmic binding proteins are classified into 3 states, open1, open2 and closed states. The open1 state structure has wide cleft between domain 1 and 2. PEB3 with phosphate (PDB ID: 2HXW) and the ligand-free maltose binding protein (PDB ID: 1LLS) adopt the open1 conformation (Fig. 4-7B and C). The open2 state is similar to the open1 state, but H7 is collapsed. The PEB3 citrate complex structure is in this state (Fig. 4-7B). The closed state is clearly different from the open1 and open2 states in terms of the angle between domain 1 and 2. The entrance of the ligand binding site is closed in this state. The structure of maltose binding protein in complex with maltose (PDB ID: 3MP6) shows this conformation (Fig. 4-7C). Previous studies of periplasmic binding proteins suggest that the structural change from the open1 to the closed state enhances the affinity to receptor proteins (Fig. 4-7D) (15). The D-malate binding to AcfC, however, enhances the binding affinity of AcfC to Mlp8 without any conformational change of AcfC. How D-malate binding to AcfC increases the binding affinity of AcfC to Mlp8? The D-malate binding site of ligand-free AcfC is highly positively charged (Fig. 4-10A). The positive charge is neutralized by D-malate binding. (Fig. 4-10B). This surface charge change may promote the binding of AcfC to Mlp8.

The result of pull-down assay indicated that the addition of galactose-6-sulfate did not change the binding affinity of AcfC to Mlp8p, although previous study suggests that Mlp8 is involved in the chemotactic response of *V. cholerae* to galactose-6-sulfate (Table 4-1) (3). This suggests that Mlp8 directly recognizes galactose-6-sulfate or with a help of another

partner protein. Since Mlp8 is expected to be a dCACHE-type chemoreceptor, Mlp8 may recognize galactose-6-sulfate directly like Mlp24 and Mlp37.

As described above, my results imply that Mlp8p recognizes ligands with, at least, two different mechanisms. Mlp8p recognizes D-malate together with AcfC, but recognizes galactose-6-sulfate without AcfC. The reason why Mlp8 has two different mechanisms of ligand recognition remains unclear. To answer this question, we need structural information of Mlp8p with and without AcfC, as well as more *in vivo* and *in vitro* studies of Mlp8 and AcfC.



**Figure 4-10. Electro static potential surface of AcfC**

Surface charge of ligand-free AcfC (A) and D-malate complexed AcfC (B) are shown. Positively charged regions are colored in blue and negatively charged regions in red. The positively charged region in the D-malate binding site is indicated with a white broken circle.



#### 4-5. References

1. Nishiyama S, Suzuki D, Itoh Y, Suzuki K, Tajima H, Hyakutake A, Homma M, Butler-Wu SM, Camilli A, Kawagishi I. 2012. Mlp24 (McpX) of *Vibrio cholerae* implicated in pathogenicity functions as a chemoreceptor for multiple amino acids. *Infect Immun* 80:3170-8.
2. Nishiyama S, Takahashi Y, Yamamoto K, Suzuki D, Itoh Y, Sumita K, Uchida Y, Homma M, Imada K, Kawagishi I. 2016. Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci Rep* 6:20866.
3. Selvaral P, Gupta R, Peterson KM. 2015. The *Vibrio cholerae* ToxR regulon encodes Host-Specific chemotaxis proteins that function in intestinal colonization. *SOJ Microbiol Infect Dis.* 3(3).
4. Esmeralda V, Cadi D, Dominic CM, Maria G, Jennifer MR, Brendan WW. 2018. *Vibrio cholerae* accessory colonisation factor AcfC: a chemotactic protein with a role in hyperinfectivity. *Sci. Rep.* 8:8390.
5. Isabelle B, Marie-Pierre G, Francis G, Hongyan A, Chengru Z, Eric O, John MF, Mario J, Josee H. 2003. Characterization of the Novel Factor Paa involved in the early steps of the Adhesion mechanism of attaching and effacing *Escherichia coli*. *Infect Immun* 71(8):4516-25.
6. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics.* 23:2947-48.

7. Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* 66:125-32.
8. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-21.
9. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.
10. Erumbi SR, Smita B, David CW, Christine M, Mirosław C, Allan M, N. Marin Y. 2007. Structural context for protein *N*-glycosylation in bacteria: The structure of PEB3, an adhesin from *Campylobacter jejuni*. *Protein Sci.* 16(5):990-95.
11. Dwyer MA, Hellinga HW. 2004. Periplasmic binding proteins: a versatile superfamily for protein engineering. *Curr Opin Struct Biol.* 14(4):495-504
12. Marvin JS, Corcoran EE, Hattangadi NA, Zhang JV, Gere SA, Hellinga HW. 1997. The rational design of allosteric interactions in a monomeric protein and its applications to the construction of biosensors. *Proc Natl Acad Sci U S A* 94(9):4366-71.
13. M. Jack Borrok, Laura LK, Katrina TF. 2007. Conformational changes of glucose/galactose-binding protein illuminated by open, unliganded, and ultra-resolution ligand-bound structures. *Protein Sci.* 16(6):1032-41.
14. Brandon S, Shauna MC, Kirandeep B, Christine CL, Calvin T, Yogesh H, Charies C, Husain K, Caixia M, John HB, Anthony BS, Bruce AV, Trevor FM. 2015. Active

transport of phosphorylated carbohydrates promotes intestinal colonization and transmission of a bacterial pathogen. *PLoS Pathog.* 11(8):e1005107.

15. David. EW. 2003. The energetics of structural change in maltose-binding protein. *Proc Natl Acad Sci U S A* 100(22):12529-30.

## Chapter 5 Summary

I focused on three *Vibrio cholerae* chemoreceptors, Mlp8, Mlp24 and Mlp37, which are homologues each other but the cell functions are different. I carried out X-ray crystal structure analysis, isothermal titration calorimetry (ITC) and pull-down assay. The results demonstrate that these three chemoreceptors have different ligand recognition mechanisms in spite of close similarity on their structure of their ligand binding domains.

In chapter 2, I describe the ligand recognition mechanism of Mlp37. I determined crystal structures of the periplasmic domain of Mlp37 (Mlp37p) in complexed with a variety of their ligands (L-arginine, L-serine and taurine). Structural comparison of these structures and the Mlp37p L-alanine complex structure (PDB ID: 3C8C) indicated that Mlp37p recognizes ligands in the membrane distal domain, CAD1. Mlp37p recognizes various ligands by adjusting the shape and size of CAD1 to the ligands. The entrance of CAD1 shows an opening that allows accommodation of larger R groups to stick out from the pocket without affecting the interactions of the amino acid backbone atoms, providing a basis for versatile sensing of amino acid attractants. These structural insights are consistent with ITC assay, chemotaxis assay and *in vivo* visualization experiments with a fluorescent ligand performed by Prof. Kawagishi's group.

In chapter 3, molecular mechanism of ligand recognition of Mlp24 is shown. Five ligand complex structures of the periplasmic domain of Mlp24 (Mlp24p) were determined as well as the structure of ligand-free Mlp24p. The overall structures of Mlp24p are very similar each other and are similar to that of Mlp37p, whereas the ligand recognition mechanism of Mlp24p is different from that of Mlp37p. In contrast to Mlp37p, the size and the shape of the ligand

binding pocket of CAD1 are not optimized to each ligand. Therefore the binding affinities of Mlp24p for smaller amino acids are relatively low. The ligand binding to Mlp24 induces the ligand binding domain (CAD1) closure. A calcium ion binds to the CAD1 and the calcium ion is important for the ligand recognition of Mlp24, and following *in vivo* experiments performed by Prof. Kawagishi group suggests that *V. cholerae* fine-tunes its chemotactic behavior in response to environmental calcium concentrations by modulating the ligand specificity of Mlp24

In chapter 4, a crystal structure of AcfC, which is involved in the pathogenicity and/or chemotaxis together with Mlp8, is shown. Structural comparison with the AcfC D-malate complex indicates that AcfC does not change its conformation upon D-malate binding, although ligand binding induces a structural change in other periplasmic-binding proteins. The results of pull-down assay and ITC analysis together with the published results suggest that Mlp8 recognizes the ligands with, at least, two different mechanisms; direct ligand recognition and indirect ligand recognition with AcfC.

## Acknowledgements

This study has been carried out in the Laboratory of Macromolecular Structure, Graduate school of Macromolecular Science, Osaka University from 2013 to 2019.

Here I deep gratitude to my supervisor, Prof. Katsumi Imada for his professional supervising throughout my Ph. D work.

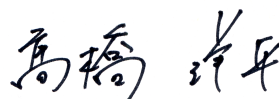
I am deeply grateful to Prof. Ikuro Kawagishi of Hosei University, So-ichiro Nishiyama of Niigata University of Pharmacy and Kazumasa Sumita for extraordinary contribution to accomplishing this work.

I thank the SPring-8 beamline staff for technical help with data collection and Hiroyuki Terashima, Norihiro Takekawa, Setsuko Kobashi and Yumiko Uchida for technical support.

This thesis was reviewed by Prof. Katsumi Imada, Prof. Atsushi Nakagawa and Prof. Genji Kurisu. I appreciate for their careful reviewing and helpful comments on this thesis.

Finally, I would say many thanks to every single colleague, friend and family who has stood in my life for their kindest support.

February, 2019



---

**Takahashi Yohei**

Department of Macromolecular Science  
Graduate School of Science, Osaka University

## List of Publications

1. So-ichiro Nishiyama\*, **Yohei Takahashi\***, Kentaro Yamamoto\*, Daisuke Suzuki, Yasuaki Itoh, Kazumasa Sumita, Yumiko Uchida, Michio Homma, Katsumi Imada and Ikuro Kawagishi. 2016, Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. Scientific Reports **6**; 20866.

**\*These authors contributed equally to this work.**

2. **Yohei Takahashi\***, So-ichiro Nishiyama\*, Kazumasa Sumita, Ikuro Kawagishi and Katsumi Imada. Calcium ion modulates amino acid sensing of the chemoreceptor Mlp24 of *Vibrio cholerae*. Journal of Bacteriology (in press).

**\*These authors contributed equally to this work.**

3. **Yohei Takahashi\***, So-ichiro Nishiyama\*, Kazumasa Sumita, Ikuro Kawagishi and Katsumi Imada. Structural basis of ligand affinity of Mlp24 and Mlp37, chemoreceptor proteins of *Vibrio cholerae*. Biophysics and Physicobiology (in preparation)

**\*These authors contributed equally to this work.**

4. **Yohei Takahashi\***, So-ichiro Nishiyama\*, Ikuro Kawagishi and Katsumi Imada. Surface charge neutralization by ligand binding induces the AcfC binding to Mlp8. (in preparation)

**\*These authors contributed equally to this work.**