Exploring the structural factors related to function appearance of biomembrane-interacting factors

May 2016

Yuki Kawaguchi

CONTENTS

Chapter 1 Preface

Biological membrane mainly consists of phospholipids, proteins, and sugar chains at ratios that differ in anatomy- and organ-specific manners. Function of biomembranes are known as a not only physical boundary, but also as scaffolds for information transmission and transport carriers. Each constituent element of biomembranes is playing a role in their functions. Because biomembranes function as a reaction field for various physiological actions, and membrane disorders cause various illness, an analysis of membrane-acting factors seem to provide useful information for disease prevention and therapy. Membrane protein function as a receptor of a virus or a toxin had been known. Abnormalities in sugar chains and phospholipids have been implicated in the pathogenesis of various diseases (*e.g.* arterial sclerosis, tumor cell growth). Structure-function analyses on biomembrane component molecules have been performed with the aim of developing diagnostic techniques and treatments.

Proteins including membrane-associated proteins continuously change their conformations and make a large number of conformers under biological conditions. A specific conformation exists for many conformers, and the selected structural conformer interacts with a target molecule. An X-ray crystallographic analysis instantaneously provides a structural image, whereas an NMR analysis provides fluctuant structural information in a solvent. *In silico* molecular simulation can perform the structural analysis of target protein under various conditions. The molecular simulation provides useful information for the structural understanding of proteins (*e.g.* membrane-acting proteins, receptor molecules). A molecular simulation analysis for membrane-acting protein factors has also been discussed, with a focus on the following membrane-associated protein factors:

1) The interaction profile of bacterial cytolysin with a cell membrane; the interactive features of the membrane-acting 11mer regions of cholesterol-dependent cytolysins (CDCs) were examined using molecular modeling and molecular simulation. Human specific cytolysin (*e.g.* intermedilysin (ILY)) interacted with human CD59, and the interactive features were examined. 2) The molecular modification of *Streptococcus anginosus* derived-SagA (SagA1, SagA2); *S. anginosus* SagA molecules have 3 ~ 4 heterocycle formable sites, and the order of priority of heterocycle formation and the role of the heterocycle structure was examined. 3) The cytolysis of snake (*e.g. Naja naja siamensis*, *Naja naja atra*, *Naja naja*) venom-derived cardiotoxins (CTs); Cobra venom includes CT-I ~ IV and cardiotoxin-like basic protein (CLBP). Although CLBP has a similar structure to CTs, it does not exhibit any cytolytic activity. Difference in molecular features of CTs and CLBP were examined.

Chapter 2. Interactive Profiles of Bacterial Cytolysins

2-1. Background

Anginosus group streptococci (AGS) are opportunistic pathogens and form part of the normal flora in the human oral cavity. AGS have been classified into following three bacterial species including five subspecies: *Streptococcus anginosus* (*S. anginosus* subsp. *anginosus*, and *S. anginosus* subsp. *whiley*), *S. constellatus* (*S. constellatus* subsp. *constellatus*, *S. constellatus* subsp. *pharyngis,* and *S. constellatus* subsp. *viborgensis*), and *S. intermedius* (Whiley, R. A. *et al*. (1989); Whiley, R. A. *et al*. (1991); Whiley, R. A. *et al*. (1999); Jensen, A. *et al*. (2013)). Some types of the Lancefield antigen are distributed in *S. anginosus* (A, C, F, G, K, and L antigens), *S. constellatus* (A, C, and F antigens), and *S. intermedius* (K antigen). However, no classified species were found to exist in these streptococcal species using the Lancefield antigen type, and immunologically classified identification was unsuitable for the discrimination of AGS (Whiley, R. A. *et al*. (1990)). AGS are generally regarded as bacteria with weak pathogenicity that co-exist with the host in the oral cavity and female genitals. When host resistance decreases, AGS cause various opportunistic infections (Brook, I. *et al*. (1994); Murray, H. W. *et al*. (1978); Ruoff, K. L. *et al*. (1988−1); Ruoff, K. L. *et al*. (1988−2)).

Between 10 and 40% of *S. anginosus* and *S. constellatus* strains exhibit beta-hemolysis activity on blood agar plates. The *S. intermedius* UNS46 strain, which was isolated from a human liver abscess, showed human specific beta-hemolysis on blood agar plates. This hemolytic activity depended on the protein toxin secreted from the UNS46 strain. This toxin is called intermedilysin (ILY) and has a molecular weight of 55kDa. The

hemolytic activity of ILY is 100-fold stronger in human than in chimpanzees and crab-eating macaques, and is negligible in animals other than primates (Nagamune, H. *et al*. (1996)). ILY has been shown to universally exist in this bacterial strain, and was found to belong to the cholesterol dependent cytolysin (CDC) gene family by a gene analysis (Nagamune, H. *et al*. (2000); Nagamune, H. *et al*. (2004)).

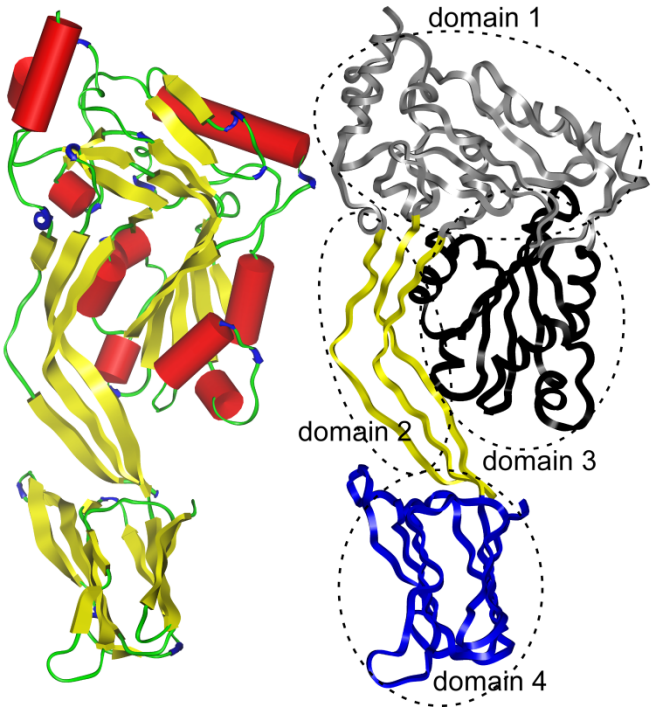


Figure 2-1. Structure of intermedilysin.
Intermedilysin consists of four domains, and domain 4 interacts with cell membrane.

ILY has four domains similar to *Clostridium perfringens* perfringolysin O (PFO), *Streptococcus pneumoniae* pneumolysin (PLY), and *Streptococcus pyogenes* streptolysin O (SLO) of group A streptococcus (Figure 2-1) (Nagamune, H. *et al*. (2004); Polekhina, G. *et al*. (2005)). ILY plays a role in pore formation on cell membrane similar to CDCs, and destroys

cells. Interestingly, ILY recognizes CD59 in the human cell membrane as a receptor (Gidding, K. S. *et al*. (2004)), whereas CDCs generally recognize cholesterol in the cell membrane as a receptor. Human CD59 (hCD59) is a GPI anchor-type cell membrane protein so-called protectin. hCD59 binds with complement element and disturbs the formation of the membrane attack complex (MAC), thereby protecting its own cell from an excessive immune response. ILY has been shown to recognize the C8/C9 binding site (F^{67} – E^{83}) of hCD59 (Gidding, K. S. *et al*. (2004)). This area forms the alpha-helix structure and is located near GPI anchor region (Figure 2-2). The homology of the amino acid sequence between hCD59 and another primate CD59 (*e.g.* baboon and African green monkey) is very low at 58.5 %. In this chapter, the interaction profile between ILY and hCD59 was introduced. Moreover, vaginolysin (VLY) and *Streptococcus mitis*-derived human platelet aggregation factor (Sm-hPAF) have been identified as the toxins with different species-specificities and the cholesterol-dependencies from ILY (Aroutcheva, A. A. *et al*. (2001); Gelber, S. E. *et al*. (2008); Ohkuni, H. *et al*. (2012)). The mode of reactions between hCD59 and these toxins were also analyzed and discussed.

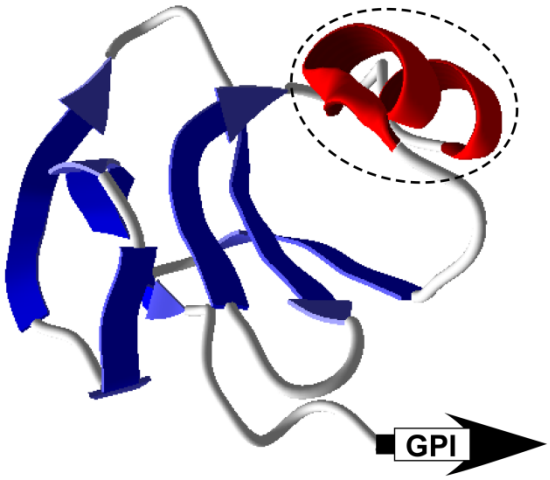


Figure 2-2. Molecular structure of human CD59.
Intermedilysin recognized the C8/C9 binding site (dotted circle) of hCD59.

2-2. Molecular Profiles of Bacterial Cytolysin 11mer Regions: Interactive Comparison with Membrane Cholesterol

2-2-1. Introduction

Some bacterial cytolysins, such as *Clostridium perfringens* perfringolysin O (PFO), *Streptococcus pneumoniae* pneumolysin (PLY), *Streptococcus pyogenes* streptolysin O (SLO), *Listeria monocytogenes* listeriolysin O (LLO), *Bacillus alvei* alveolysin (ALV), and *Streptococcus suis* suilysin (SLY) form ring-clusters on the target cell membranes and lyse cells through pore formation (Billington, S. J. *et al*. (2000); Olofsson, A. *et al*. (1993)). The cytolytic activities of these toxins are dependent on cholesterol, and thus, are called cholesterol-dependent cytolysins (CDCs). In general, members of the CDC family have a conserved undecapeptide (11mer) region, the amino acid sequence of which is ECTGLAWEWWR. The 11mer region of traditional CDCs plays a key role in cell membrane recognition (Korchev, Y. E. *et al*. (1998); Sekino-Suzuki, N. *et al*. (1996); Jones, S. *et al*. (1996)).

Intermedilysin (ILY) was found to be secreted from *Streptococcus intermedius*, and exhibited human-specific cell lytic activity (Nagamune, H. *et al*. (1996)). Mutation experiments on the 11mer region confirmed that the ILY 11mer region participated in species-specific cell lytic activity (Ohkura, K. *et al*. (2006–1)). *Arcanobacterium pyogenes* secreted a cholesterol-dependent cytolysin named pyolysin (PLO), which exhibits cholesterol-dependent cytolytic activity, but not human cell specificity (Funk, P. G. *et al*. (1996); Billington, S. J. *et al*. (1997)). Vaginolysin (VLY) is produced by *Gardnerella vaginalis* and is thought to be the offending bacterium in preterm birth (Aroutcheva, A. A. *et*

al. (2001); Gelber, S. E. *et al*. (2008)). VLY has 55% or more homology to ILY and exhibits human-specific cytolytic activity. Moreover, *Streptococcus mitis*-derived human platelet aggregation factor (Sm-hPAF) has been identified as a multifunctional factor that exhibits not only cytolytic activity, but also platelet aggregation activity (Ohkuni, H. *et al*. (2012)). These cytolysins (ILY, PLO, VLY, and Sm-hPAF) individually have unique 11mer region that may play a key role in controlling species-specificity and cholesterol dependency. In this chapter, the molecular models of these cytolysins were constructed based on the X-ray data of PFO (protein data bank ID=1PFO), and the structural features of 11mer regions were analyzed.

2-2-2. Materials and Methods

Structural construction of cytolysins.

Amino acid sequences of ILY, PLO, PLY, LLO, SLO, ALV, SLY, VLY, and Sm-hPAF were aligned using clustalX (Thompson, J. D. *et al*. (1997)). Homology model of ILY was constructed based on the X-ray crystal structure data of PFO (1PFO) using InsightII Homology module (Accelrys Inc., San Diego, CA, USA) on a Silicon Graphics Octane Workstation (Ohkura, K. *et al*. (2004)). The initial molecular model was built in two stages: [1] identification of significant regions of sequence identity between PFO and ILY, and assigning coordinates to ILY within these regions; [2] designation of ILY coordinates in the regions that showed less convincing regions of sequence identity to PFO using a database of peptide fragments. Since PFO (1PFO) included no hydrogen data in its X-ray crystallographic structure, the hydrogens were added using the Biopolymer module. The molecular model was visually checked to remove any obvious steric clashes, followed by

energy minimization using a Consistence Valence Forcefield (CVFF) with Discover3 module. The RMS deviation for all Cα positions on superposition between PFO and ILY was 0.165 angstrom. The quality of the final molecular model was assessed using PROCHECK and XPLOR (Laskowski, R. A. *et al*. (1993); Brunger, A. T. *et al*. (1987)). Homology models of other cytolytic proteins were designed as well as ILY.

Molecular features of 11mer regions of cytolysins.

The coordinate data of 11mer region were extracted from the minimized molecular models of ILY, PLO, PLY, LLO, SLO, ALV, SLY, VLY, Sm-hPAF and hydrogen atoms were added to each N– and C–terminus of the peptides. Energy calculations were performed with PM3 Hamiltonian using MOPAC (Fujitsu Limited, Tokyo, Japan), and the stable and transient structures were initially built with general parameters of bond length, bond angle and dihedral angle, and refined with the eigen-vector following (EF) optimization method (Ohkura, K. *et al*. (2006–1)).

Solvation free energy of 11mer region.

Solvation free energies (dGW) of 11mer regions were determined from MO parameters (Ohkura, K. *et al*. (1999)). The solvation free energy (dGW = dGWa – dGWm) of target molecule (region) was defined from the free-energy changes for association in the aqueous solution (dGWa) and in the cell membrane (dGWm; using the dGW in vacuum, because cell membrane have hydrophobic property) by the COSMO method (Lambardo, F. *et al*. (1996)). The dielectric constant (ε) of water and vacuum were 78.4 and 1.0, respectively.

Positional relationship between cholesterol and cytolysin 11mer region.

Docking simulation between cholesterol molecule and modeled cytolysins were performed using InsightII Discover module (Accelrys Inc., San Diego, CA, USA) under CVFF forcefield, and the binding energy was monitored during simulation period. Positional effect of cholesterol molecule on binding energy was examined, and effective orientation of 3-hyroxy group of cholesterol was determined.

2-2-3. Results

Molecular features of cytolysin 11mer regions.

Traditional CDCs which are previously called as 'cholesterol-binding cytolysins' have a conserved 11mer amino acid sequence (ECTGLAWEWWR), and amino acid substitutions at 11mer region were observed in other types of cytolysins (*e.g.* ILY, PLO, VLY, and Sm-hPAF) (Figure 2-3). The solvation free energy (dGW: lower dGW value means higher hydrophobicity), which is an index of structure-dependent hydrophobicity, of these 11mer regions was analyzed (Table 2-I). The dGW values of ILY and PLO 11mer regions were –440.6 and –457.3 kJ/mol, respectively. The dGW value range of traditional CDC 11mers (*e.g.* PFO, PLY, LLO, SLO, ALV, SLY) is –554.7 to –448.1 kJ/mol. The dGW of VLY 11mer region was –490.0 kJ/mol. Sm-hPAF had the most hydrophobic 11mer profile (–655.8 kJ/mol) among the tested CDCs.

```
      CDC : ECTGLAWEWWR
      ILY : GATGLAWEPWR
      PLO : EATGLAWDPWW
      VLY : EKTGLVWEPWR
   Sm-hPAF : EKTGLVWEPWR
```

Figure 2-3. Amino acid sequences of 11mer regions of cytolysins.
CDC: cholesterol dependent cytolysin; ILY: intermedilysin; PLO: pyolysin; VLY: vaginolysin; Sm-hPAF: *S. mitis*-derived human platelet aggregation factor.

Type of the dipole moment direction, which is an index of reactivity to target molecule, of 11mer regions in cytolysins are shown in Figure 2-4. Dipole moment of ILY (1.731 debye) was significantly smaller than that of traditional CDCs (Table 2-I), and the moment direction faced the intramolecular site crowded with other amino acid residues (type A in Figure 2-4). The dipole moment of PLO 11mer region (2.427 debye) was smaller than that of traditional CDCs, as well as that of the ILY moment, but it faced the outer side (type B). In traditional CDCs, dipole moment intensity is in the range of 8.028 to 9.022 debye, and it faces the outer side of the cytolysin molecule (type C). VLY (3.647 debye) and Sm-hPAF (8.540 debye) showed the fourth type of dipole moment direction, facing downward and to the left, on the far side of the surface of the space (type D).

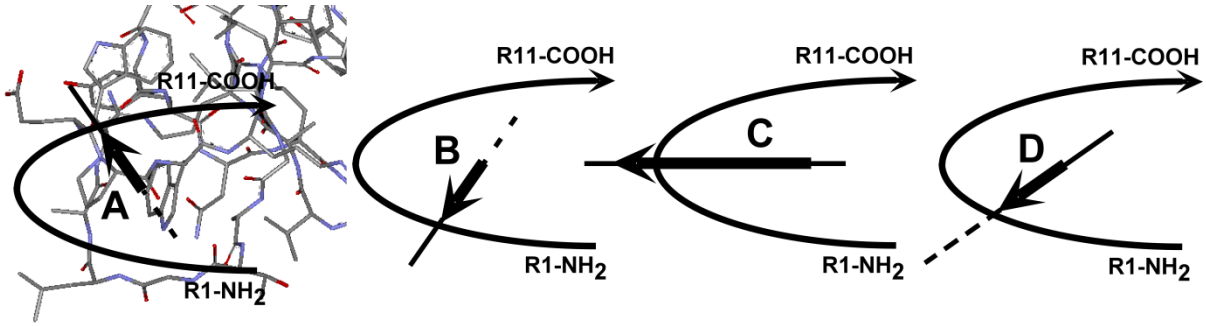


Figure 2-4. Dipole moment direction of 11mer regions.
Dipole moment direction of ILY (type A), PLO (type B), traditional CDCs (type C), VLY and Sm-hPAF (type D). R1-HN_2: N-terminal of 11mer region. R11-COOH: C-terminal of 11mer region. A broken line of the arrow context of A, B, D indicates the far side of the surface of space (*i.e.* plane of paper), and a solid line indicates the front side of the surface of space.

Table 2-I. Structural parameters of 11mer regions.

	dGW: Solvation free energy (kJ/mol)	Dipole moment (debye)	Moment type	HOMO energy (eV)	LUMO energy (eV)	HOMO−LUMO energy gap (eV)*
ILY	−440.6	1.731	A	−8.081	0.076	8.157
PLO	−457.3	2.427	B	−8.076	0.100	8.176
PFO	−554.7	9.022	C	−8.323	−0.298	8.025
PLY	−448.4	8.028	C	−8.426	−0.181	8.245
LLO	−498.3	8.367	C	−8.316	−0.166	8.150
SLO	−502.3	8.133	C	−8.426	−0.184	8.242
ALV	−448.1	8.297	C	−8.425	−0.181	8.244
SLY	−502.6	8.215	C	−8.416	−0.178	8.238
VLY	−490.0	3.647	D	−7.992	−1.351	6.641
Sm-hPAF	−655.8	8.540	D	−3.559	−1.202	2.358

Solvation free energy and MO-parameters were obtained from the data in vacuum condition. Moment types are shown in Figure 2-4. Intermedilysin (ILY), pyolysin (PLO), perfringolysin O (PFO), pneumolysin (PLY), listeriolysin O (LLO), streptolysin O (SLO), alveolysin (ALV), suilysin (SLY), vaginolysin (VLY), *Streptococcus mitis*-derived human platelet aggregation factor (Sm-hPAF).
*The absolute value of HOMO−LUMO energy gap.

Analysis of 11mer region interaction with cholesterol molecule.

In the interactive analysis of 11mer region mutated SLO and ILY, the position of 3-hydroxyl group (3-OH) in cholesterol molecule influenced non-binding energies (index of the interaction) between cytolysin and cholesterol (Ohkura, K. *et al*. (2006−1)). In this section, the interaction between 11mer region and cholesterol was analyzed in various cytolysins, and the results are shown in Table 2-II. The position of 3-hydoxyl group in cholesterol molecule was different in three binding patterns (cases 1, 2, and 3 in Figure 2-5).

When the 3-hydroxyl group of cholesterol was located near the glutamic acid (Glu492) of ILY (case 1), the non-binding energy (NBE) between cholesterol and ILY was significantly lower (–2.603 kcal/mol) than that of SLO (3.175 kcal/mol). The cholesterol dependency is different in ILY and SLO; therefore, it seemed that case 1 was appropriate as the interaction index model. When the ILY 11mer region was mutated to SLO type 11mer region [ILY(ECW)], NBE had a positive value (6.462 kcal/mol), as did that for SLO. A positive NBE value means an interactive relationship. In contrast, the 11mer mutated SLO [SLO(GAP)], which has an ILY type 11mer region, had a negative NBE value (–2.593 kcal/mol), as did that for ILY. In case 1 NBE analysis, the NBE of PLO (–1.616 kcal/mol), VLY (–3.345 kcal/mol), and Sm-hPAF (–1.972 kcal/mol) also had negative values. Although effect of cholesterol molecule on hemolysis was different between ILY and SLO, there was no difference in the NBE positions of ILY and SLO in case 2 and case 3 (Figure 2-5, Table 2-II); therefore, these positions should not be considered.

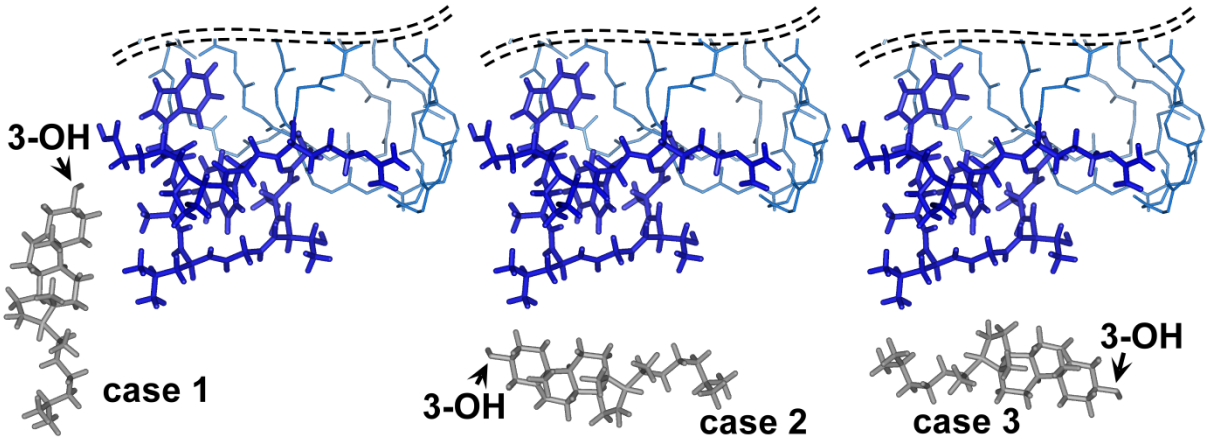


Figure 2-5. Interaction profile between cytolysin 11mer region and cholesterol molecule.

The 3-hydroxy (3-OH) position in cholesterol is different in each simulation condition (case 1 – case 3).

Table 2-II. Intermolecular parameters between cytolysin 11mers and cholesterol.

	Non-binding energy (kcal/mol)			Cholesterol effect on hemolysis*[1]
	case 1	case 2	case 3	
ILY	-2.603	-3.321	-4.813	−
ILY(ECW)	6.462	-3.470	-1.339	+++
SLO	3.175	-3.392	-4.835	+++++
SLO(GAP)	-2.593	-3.325	-5.81E-33	++++
PLO	-1.616	-1.842	-2.470	ND
VLY	-3.345	-3.899	-5.143	ND
Sm-hPAF	-1.972	-1.861	-2.380	+++*[2], − *[2]

Cholesterol and 11mer region of each cytolysin are located as shown by case 1, 2, 3 in Figure 2-5, and non-binding energies were determined using InsightII-Discover. ILY(ECW) is a ILY mutant, which has the traditional CDC type 11mer. SLO(GAP) is a SLO mutant, which has the ILY type 11mer. ND: not determined. *[1]Summarized from reference (Ohkura, K. *et al*. (2006–1)). *[2]Cholesterol effect on human erythrocyte : +++, rabbit erythrocyte : −.

2-2-4. Discussion

In this chapter, structural profiles of the 11mer regions of bacterial cytolysins were examined using the molecular modeling and molecular orbital (MO) analysis. The dipole moment directions of ILY and PLO differed respectively from that of traditional CDCs. It was interesting that VLY and Sm-hPAF had the fourth type of dipole moment profile (Figure 2-4). Although the dipole moment direction was the same in VLY and Sm-hPAF, the dipole moment strength of VLY was smaller (3.647 debye) than that of Sm-hPAF (8.540 debye). The dGW value of Sm-hPAF 11mer region (–655.8 kJ/mol) was smaller than that of VLY (–490.0 kJ/mol), and the 11mer region hydrophobicity of Sm-hPAF was higher than that of VLY. These results suggest that the dynamic structure of the 11mer region was influenced by the structure of the whole cytolysin molecule. Amino acid sequence of 11mer region was the same in VLY and Sm-hPAF, but differences in the three-dimensional 11mer structure might arise because structural differences of other parts affect the 11mer structure. The (HOMO–LUMO) energy gap (an index of chemical reactivity; the smaller absolute value of HOMO–LUMO energy gap indicates the higher reactivity) was estimated from the values of HOMO and LUMO obtained by MO analysis. The absolute energy gap value in the 11mer region of VLY (6.641) and Sm-hPAF (2.358) was smaller than in other CDCs [8.025 (PFO) to 8.245 (PLY)]. The 11mer regions of VLY and Sm-hPAF had higher reactivity than other cytolysins.

The 3-acetylated cholesterol analogue (3-hydroxyl group substituted with an acetyl group) did not inhibit hemolytic activity of SLO (data not shown), and the 3-hydroxyl group has an important role in the cytolytic activity of traditional CDCs. Membrane recognition of CDC is necessary for the cell membrane destruction process, and lipid raft domain takes part

in the membrane recognition of CDCs. In the lipid raft, it was assumed that the hydroxyl groups of cholesterol molecules would be positioned along the water-accessible outer side. This consideration does not contradict the result of NBE analysis in which the interaction energy between cholesterol and cytolysin (*e.g.* ILY, SLO) differs only under the case 1 condition. ILY is a member of CDC gene family, but ILY preferentially recognize the primary receptor, not cholesterol but human CD59, in lipid raft-dependent manner (Giddings, K. S. *et al*. (2004)). Thus, ILY can be regarded as an atypical CDC because of its lipid raft-dependency for interaction with cell membrane. From the NBE values in interactive analysis, VLY and Sm-hPAF seemed to have no cholesterol-binding features, as did ILY, or it was thought that they showed extremely low affinity to cholesterol compared with traditional CDCs (*e.g.* SLO). Cholesterol molecule affected the interaction between Sm-hPAF and human erythrocytes, but did not influence the interaction with rabbit erythrocytes (Ohkuni, H. *et al*. (2012)). As for Sm-hPAF, its behavior in cholesterol molecule was different from that of other cytolysins. Only the Sm-hPAF has domain 0 on the *N*-terminal side, and the dynamic structure behavior of domain 0 might have various influences on membrane and cholesterol recognition. These domain 0 effects are currently being analyzed. Moreover, Sm-hPAF is a multifunctional factor, and the confirmation of its roles in host-parasite interaction is at an advanced stage.

2-3. Interaction Profile between Human CD59 and Bacterial Cytolysins

2-3-1. Introduction

The interaction between 11mer region of the CDCs and cholesterol molecule was described in the previous chapter. This chapter focused on the interaction between human-specific cytolysins (*e.g.* ILY) and their receptor molecules (*e.g.* human CD59).

Some bacterial cytolytic proteins form clusters on the target cell membranes and lyse target cells through pore formation. Among these cytolytic proteins, cytolysins in a group bind to cholesterol on the cell membrane as a receptor and are called cholesterol-dependent cytolysins (CDCs). CDCs are produced by various bacterial strains [*e.g. Clostridium perfringens*, perfringolysin O (PFO); *Streptococcus pneumoniae*, pneumolysin (PLY); *Streptococcus pyogenes*, streptolysin O (SLO); *Listeria monocytogenes*, listeriolysin O (LLO)], and form the CDC gene family (Billington, S. L. *et al*. (2000); Olofsson, A. *et al*. (1993); Korchev, Y. E. *et al*. (1998); Sekino-Suzuki, N. *et al*. (1996); Jones, S. *et al*. (1996)).

Streptococcus intermedius intermedilysin (ILY) exhibits human-specific cell lytic activity, but not the cell-lytic activity directly based on cholesterol binding (Nagamune, H. *et al*. (1996); Ohkura, K. *et al*. (2006−1)). A previous study reported that ILY binds selectively to human CD59 (hCD59) as a receptor protein (Giddings, K. S. *et al*. (2004)). *Gardnerella vaginalis*-derived vaginolysin (VLY) and *Streptococcus mitis*-derived human platelet aggregation factor (Sm-hPAF) were recently identified as novel types of cell-lytic proteins (Gelber, S. E. *et al*. (2008); Ohkuni, H. *et al*. (2012)). Sm-hPAF has five domains (ILY-like domains 1−4 and the extra-domain; domain 0), and its cholesterol dependency is significantly weaker than that of traditional CDCs. Sm-hPAF is a multifunctional factor that exhibits

platelet-aggregatory and cell-lytic activities. These toxins, VLY and Sm-hPAF, exhibited similar behavior in the cell lytic analysis using human erythrocytes and appeared to interact with hCD59 as well as ILY. These CDCs were previously analyzed and classified into different types by the features of their 11mer regions: ILY (type A), pyolysin (PLO: type B), traditional CDC (type C), VLY (type D), and Sm-hPAF (type D) (Ohkura, K. *et al*. (2012)). In this chapter, the interaction profiles between these CDCs (*e.g.* ILY, VLY, and Sm-hPAF) and hCD59 were analyzed, and the behavior of the 11mer neighborhood region in the interaction with hCD59 was discussed.

2-3-2. Materials and Methods

Molecular modeling of bacterial cytolysins.

Molecular models of VLY and Sm-hPAF were constructed based on the X-ray data of ILY (protein data bank ID=1S3R) using the InsightII Homology module (Accelrys Inc., San Diego, CA, USA) (Ohkura, K. *et al*. (2004)). The molecular structures of ILY, VLY, and Sm-hPAF were overlapped at the same position in three-dimensional coordinates, and the interaction with hCD59 was analyzed.

Interactive analysis of cytolysin with hCD59.

Non-binding energies (NBE: an index of molecular interaction) between hCD59 (protein data bank ID=2J8B) and modeled cytolysins were examined using the InsightII-Discover (Accelrys Inc., San Diego, CA, USA) as previously described. The features of the 11mer regions [dipole moment, solvation free energy (dGW)] of cytolysins were analyzed using MOPAC (Fujitsu Inc., Tokyo, Japan) as previously described (Ohkura, K.

et al. (1999); Ohkura, K. *et al*. (2005); Ohkura, K. *et al*. (2003)).

Electrostatic potential field analysis.

The electrostatic potential (ESP) fields were calculated using Delphi module with InsightII-Discover (Accelrys, Inc., San Diego, CA, USA). The +1.0 kT/e (green) and –1.0 kT/e (red) contour was displayed (Ohkura, K. *et al*. (2006–2)).

2-3-3. Results

Non-binding energy between bacterial cytolysins and hCD59.

The van der Waals (Vdw) energy (an index of attracting force) between ILY and hCD59 was 234199 kcal/mol. This Vdw energy was significantly larger than that between VLY (433.3 kcal/mol), Sm-hPAF (351.9 kcal/mol) and hCD59 (Table 2-III). The electrical energy between ILY and hCD59 was −33.8 kcal/mol. This electrical energy was also larger than that between VLY (−227.8 kcal/mol), Sm-hPAF (−238.7 kcal/mol) and hCD59. The attractive force was generated between VLY, Sm-hPAF and hCD59, although it was weaker than that between ILY and hCD59.

Table 2-III. Non-binding energy between ILY, VLY, Sm-hPAF and human CD59.

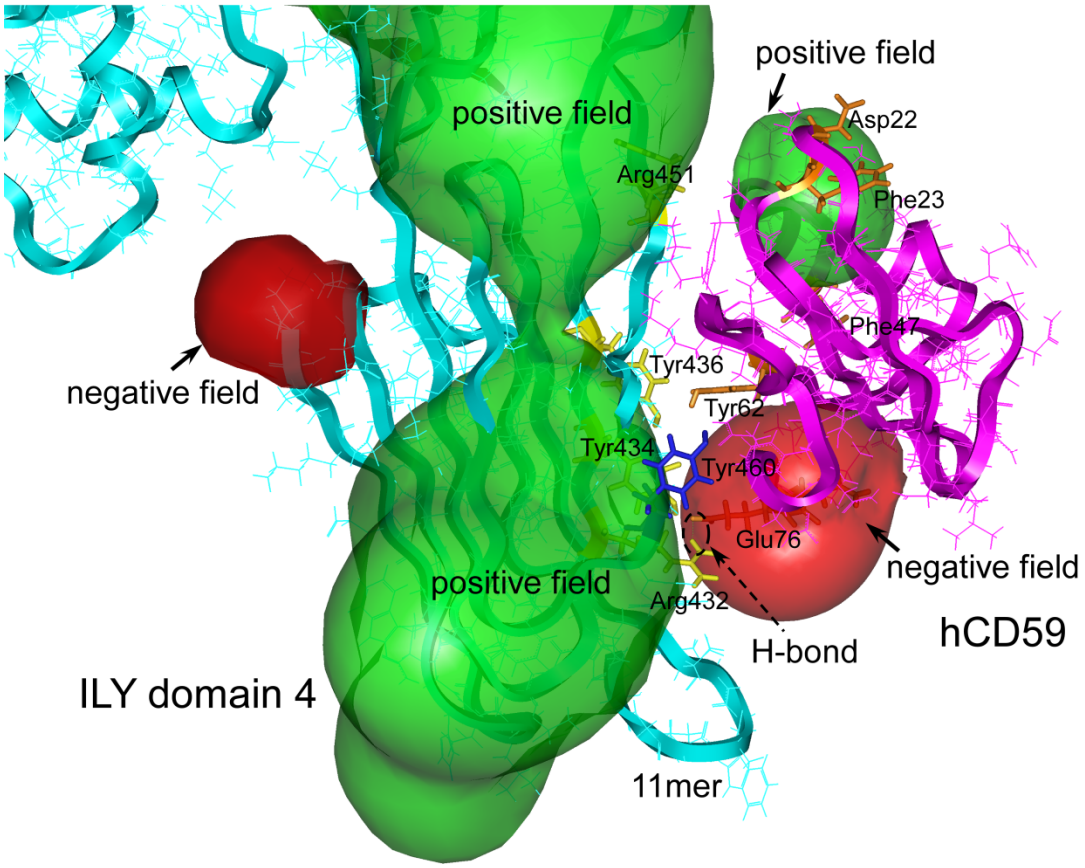
	ILY	VLY	Sm-hPAF
Vdw (kcal/mol)	234199	433.3	351.9
Elect (kcal/mol)	-33.8	-227.8	-238.7
Total (kcal/mol)	234165	205.5	113.2

Vdw: van der Waals energy. Elect: electric energy. Total: total energy = Vdw energy + Elect energy.

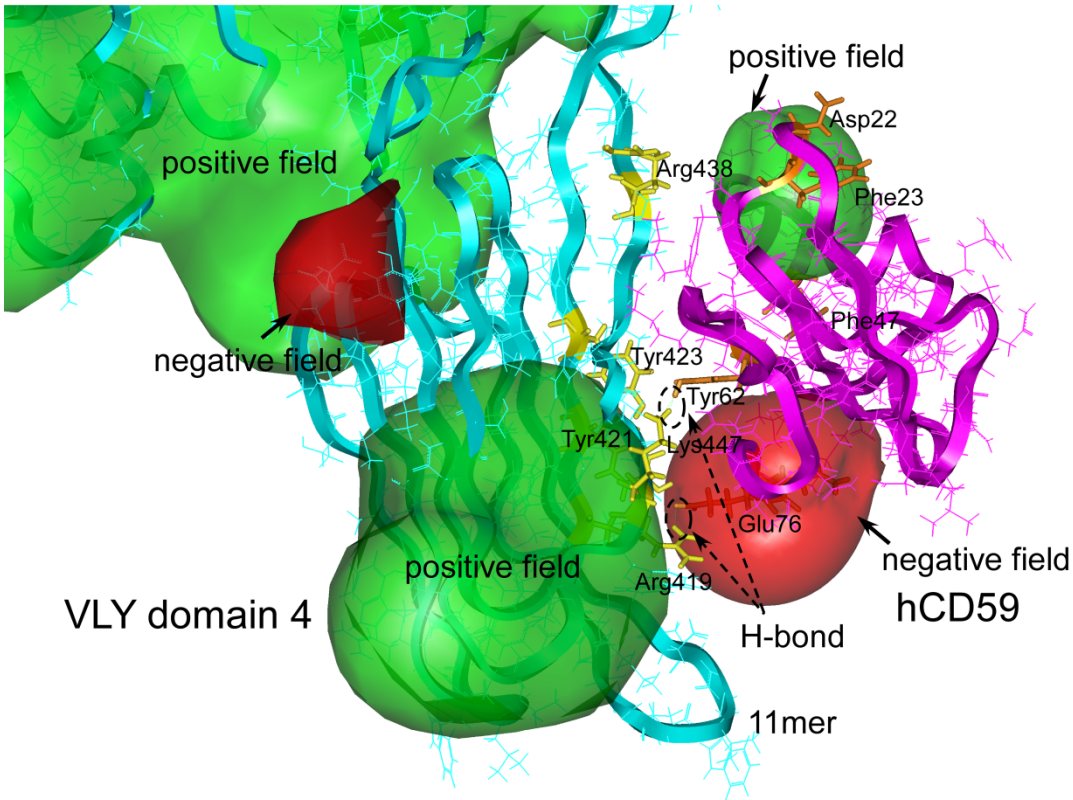
Interactive profiles of cytolysins with hCD59.

The amino acid residues (Arg451, Tyr436, Tyr434, Tyr460, Arg432) of ILY molecule are located at the point of interaction with hCD59 molecule (Figure 2-6A). The hydrogen atom of Arg432 in ILY formed a hydrogen bond with the oxygen atom of Glu76 in hCD59. Tyr460 of ILY domain 4 existed at a good position of regulating the interaction with hCD59 molecule. The expansion of the positive electrostatic potential (ESP) field at ILY domain 4 and the negative ESP field at lower area of hCD59 was observed. The ILY 11mer region was located outside the positive ESP field and did not interact directly with hCD59 molecule.

In molecular interactive analysis between VLY and hCD59, two hydrogen bonds were observed. The Lys447 (nitrogen atom) and Arg419 (hydrogen atom) of VLY domain 4 formed a hydrogen bond with Tyr62 (hydrogen atom) and Glu76 (oxygen atom) of hCD59, respectively (Figure 2-6B). A positive ESP field at VLY domain 4 and negative ESP field at lower area of hCD59 were observed, as well as the interaction between ILY and hCD59. Similarly, in Sm-hPAF domain 4, Lys403 (nitrogen atom) and Arg375 (hydrogen atom) formed a hydrogen bond with hCD59 Tyr62 (hydrogen atom) and Glu76 (oxygen atom), respectively (Figure 2-6C). Positive ESP field was observed at Sm-hPAF domain 4, and negative ESP field was observed at lower area of hCD59. Both 11mer regions of VLY and Sm-hPAF were located outside the positive ESP field and moved freely.


positive field
positive field
Asp22
Arg451
Phe23
Phe47
negative field
Tyr436
Tyr62
Tyr434
Tyr460
Glu76
negative field
positive field
Arg432
hCD59
H-bond
ILY domain 4
11mer

(A)


positive field
positive field
Asp22
Arg438
Phe23
Phe47
negative field
Tyr423
Tyr62
Tyr421
Lys447
Glu76
negative field
positive field
Arg419
hCD59
VLY domain 4
H-bond
11mer

(B)

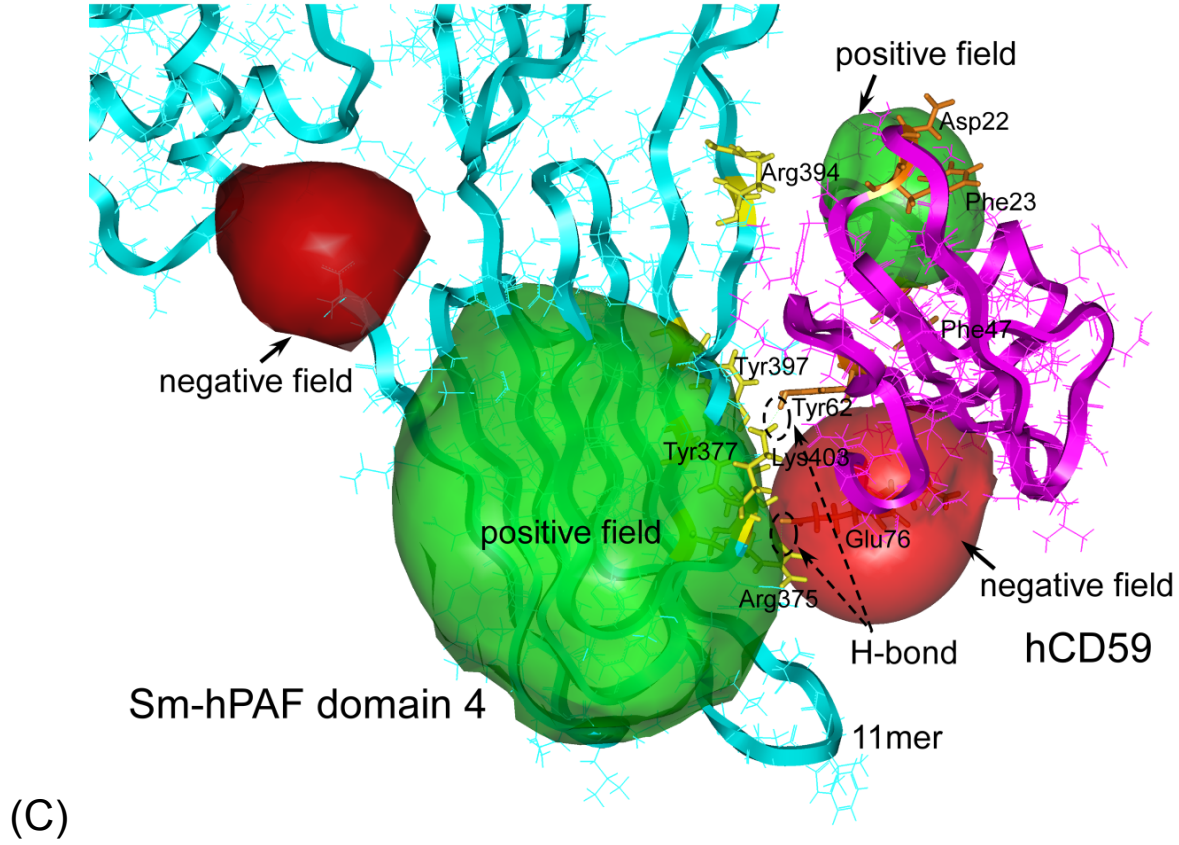


Figure 2-6. Profiles of interaction between cytolysins and hCD59.
Electrostatic potential field of intermedilysin (ILY) (A), vaginolysin (VLY) (B), and *Streptococcus mitis*-derived human platelet aggregation factor (Sm-hPAF) (C) domain 4 and hCD59 are shown as green (positive charged field) and red (negative charged field) clouds.

11mer region feature in interaction with hCD59.

The dipole moment direction of the ILY 11mer region was to the upper left (Figure 2-7) and was directed toward the positive electrostatic potential field at ILY domain 4 (Figure 2-6). This dipole moment intensity of ILY was 16.153 debye (Table 2-IV). The dipole moment directions of VLY and Sm-hPAF were to the lower left (Figure 2-7), and were directed to the outer side of domain 4 (Figure 2-6). Dipole moment intensities of VLY and

Sm-hPAF were 52.023 and 52.259 debye, respectively. The solvation free energy (an index of stereo-hydrophobicity (dGW): lower dGW value means higher hydrophobicity) of the ILY 11mer region was −871.7 kJ/mol, higher than those of VLY (−1286.2 kJ/mol) and Sm-hPAF (−1292.8 kJ/mol) (Table 2-IV).

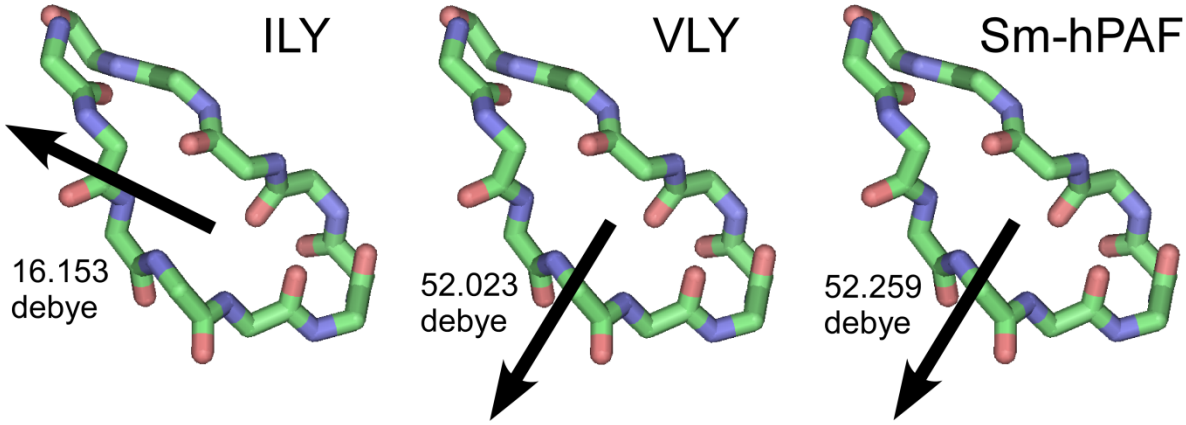


Figure 2-7. Dipole moment profiles of cytolysin 11mer regions.
Dipole moment direction (arrow) and intensity of ILY (16.153 debye), VLY (52.023 debye), and Sm-hPAF (52.259 debye) are shown.

Table 2-IV. Dipole moment intensity and solvation free energies of 11mer regions.

	dipole moment (debye)	dGW (kJ/mol)
ILY	16.153	-871.7
VLY	52.023	-1286.2
Sm-hPAF	52.259	-1292.8

2-3-4. Discussion

The amino acid residues (Asp22, Phe23, Phe47) of hCD59 are located at unrelated positions to the direct interaction with ILY. However, amino acid mutation of these hCD59 residues was found to reduce the reactivity to ILY (Wickham, S. E. *et al*. (2011)), and these residues seem to regulate the interaction between ILY and hCD59. Indeed, electrostatic repulsion by the positive ESP field was observed between ILY and the upper region of hCD59 (Figure 2-6A). One hydrogen bond was observed between ILY (Arg432) and hCD59 (Glu76), and these molecules interacted flexibly. On the other hand, two hydrogen bonds were observed between VLY and hCD59 (VLY Arg419−hCD59 Glu76, VLY Lys447−hCD59 Tyr62) in molecular dynamics analysis (Figure 2-6B), and these molecules interacted rigidly. From these results, I consider that the elaborate fitting of VLY with hCD59 was consequently disturbed compared with ILY, and non-binding energy was thought to be three order smaller than interaction between ILY and hCD59. In interactive analysis between Sm-hPAF and hCD59, two hydrogen bonds were observed (Sm-hPAF Arg375−hCD59 Glu76, Sm-hPAF Lys403−hCD59 Tyr62), and a similar non-binding energy profile and ESP field profile was obtained (Figure 2-6C) as the interactive profile of VLY–hCD59. In the X-ray analysis (PDB ID = 4BIK), positional relation of ILY with hCD59 was similar to the results of present ILY–hCD59 binding analysis, and Tyr436 neighborhood region of ILY related with hCD59 (Johnson, S. *et al.* (2013)). This X-ray data set of hCD59 bound ILY is incomplete, and including two data deletion sites (deleted at ILY Gly327–Thr334 and Gly351–Lys359), and the release of complete data set is waiting to compare with the interactive simulation data of ILY with hCD59. Moreover, no ESP field analysis was performed in this X-ray examination. ESP field is one of the important index

for molecular interaction, and positive (ILY) and negative (hCD59) ESP field function as attracting force in molecular interaction between ILY and hCD59 (Figure 2-6A).

The 11mer region of ILY moved freely during interactive analysis period with hCD59 and did not interact directly with the receptor molecule (*i.e.* hCD59). In the interactive analysis between CDC and cholesterol, the 11mer region directly interacted with cholesterol molecule, and the 3-hydroxy group of cholesterol was observed to be a key structure in their interaction (Ohkura, K. *et al*. (2012)). The structural features of cholesterol-dependent/independent bacterial cytolysin are involved in molecular pathogenesis and infectious disease therapy. Now the detailed interactive analyses of these cytolysins with hCD59 are performing. Moreover, it is interesting how cytolysins with new properties in cholesterol-independency, such as Sm-hPAF and VLY, have appeared.

Chapter 3. Molecular Analysis of *Streptococcus anginosus*-derived SagA Molecules

3-1. Background

The cytolysins introduced in the previous chapters were secreted from bacteria as matured structure without post-translational modification and exhibited cytolytic activity. On the other hand, structure-modifiable-type bacterial toxins, which were converted from the pro-active form to active form by post-translational modification, also exist. One typical mechanism is the extracellular activation by external factors from pro-form to an active form; however, another mechanism in which self-structure-modifiable bacterial protein toxins are intracellularly maturated as active form has also been known. In this chapter, the molecular features of bacterial cytolytic peptides intracellularly modified were discussed.

Streptococcus anginosus subsp. *anginosus* (SAA) expresses various types of Sag homolog molecules similar to *S. pyogenes* (SagA ~ SagI) (Victor, N. *et al*., 2000). SagA molecules (SagA1 and SagA2) exhibit cytolytic activity. SagB, SagC, and SagD molecules assume the participation in structural modification of *sagA1* and *sagA2* gene products. SagG, SagH, and SagI have been known to associate with the secretion (transport) control of the bacterial streptolysin S (SLS) homologue (Shaun, W. L. *et al*., 2008). SagA molecules have heterocycle structure-formable sites in their own structures. In this chapter, the relationship between heterocycle (*e.g.* oxazoline/thiazoline) ring formation and active form conversion in SagA molecules was discussed.

3-2. Introduction

Streptococcus anginosus subsp. *anginosus* (SAA) is a member of Anginosus group streptococci (AGS) (Whiley, R. A. *et al*. (1991); Whiley, R. A. *et al*. (1999); Jensen, A. *et al*. (2013)). It is an opportunistic pathogen that forms part of the normal flora in the human oral cavity, genitourinary tract, and gastrointestinal tract (Whiley, R. A. *et al*. (1992)). It exhibits lower pathogenic potential than other streptococci, particularly members of Pyogenic group streptococci (PGS) such as *S. pyogenes* (SPy), which is also designated a member of Group A streptococci, GAS. SAA is increasingly recognized as a pathogen that is able to cause a wide range of purulent infections that commonly manifest in the formation of abscesses, and interestingly, it has also been detected in esophageal cancer (Sasaki, M. *et al*. (2005); McKenzie, T. J. *et al*. (2010)) and suspected to participate in the cancer. Although awareness of the clinical importance of SAA has increased, the molecular basis of the pathogenicity of this species has not been clearly determined. Several species of AGS, including SAA, exhibit beta-hemolytic activity on blood agar plates, and a beta-hemolytic reaction has been assumed to indicate the production of cell lytic factors considered to be important for their pathogenicity. However, the beta-hemolytic factor of AGS examined was only in the human-specific cytolysin, intermedilysin, secreted from *S. intermedius* (Nagamune, H. *et al*. (2000)). Previous studies have not described other factors conferring beta-hemolytic capabilities on beta-hemolytic SAA and other beta-hemolytic AGS, except for *S. intermedius*.

The beta-hemolysis factors in SAA-type strain NCTC10713^{T} had been investigated in random gene-knockout examinations (Tabata, A. *et al*. (2013)). The genes responsible for the production of beta-hemolytic factors were found to be a homologue of *sag* operon gene

clusters including *sagA* encoding streptolysin S (SLS) of PGS such as *S. pyogenes*. A significant difference in the *sag* operon homologue of beta-hemolytic SAA was observed around the *sagA* gene, and two *sagA* homologues (*sagA1* and *sagA2*) were found to tandemly exist at the upstream of the *sagB* gene. No such tandem structure was found in the PGS *sag* operon (Tabata, A. *et al*. (2013)). The alignment of the deduced amino acid sequences of *sagA1* and *sagA2* product (SagA1 and SagA2) shows that the primary structures of these products were highly conserved (Tabata, A. *et al*. (2013)). These SagA homologs have leader peptide and propeptide regions, and the amino acid sequence alignment of SagA1, SagA2, and SagA of *S. pyogenes* (SPySagA) revealed a conserved sequence, particularly in the leader peptides among these molecules (Figure 3-1) (Tabata, A. *et al*. (2013)). Regarding the amino acids potentially contributing to heterocycle formation, the number and location of candidate amino acids concerned with oxazoline/thiazoline formation are suspected to vary among these SagA molecules (Figure 3-1) (Tabata, A. *et al*. (2013)). In this chapter, the structural features of these SagA molecules that are the candidates as virulence factor of SAA were examined, and the role of heterocycle formation in their maturation and function were discussed.

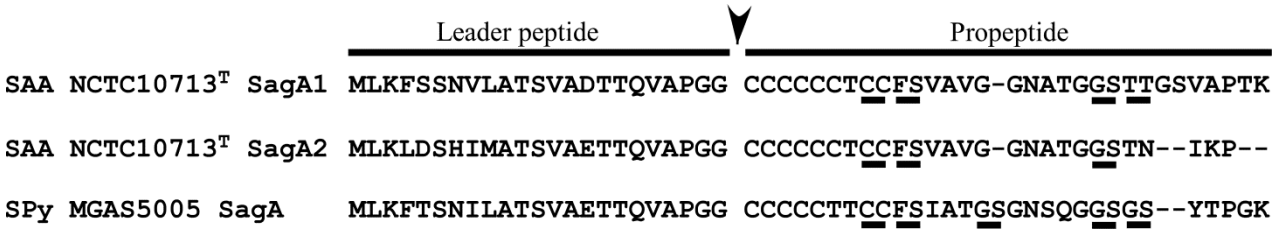


Figure 3-1. Alignment of amino acid sequences of SagA molecules for SagA1 and SagA2 of *Streptococcus anginosus* subsp. *anginosus* NCTC10713^T and SagA of *S. pyogenes* MGAS5005.

The alignment analyses were performed using ClustalX. Amino acids deduced to contribute to heterocycle formation in SagA1, SagA2, and SagA are underlined. Arrowhead indicates the cleavage site of SagA molecules for maturation, which divides the molecule into leader peptide and propeptide.

3-3. Materials and Methods

Molecular construction of SagA molecules.

Molecular models of SagA molecules from SAA NCTC10713^T (SagA1 and SagA2) and SagA from *S. pyogenes* MGAS5005 (SPySagA) were constructed using InsightII-Discover (Accelrys Inc., San Diego, CA, USA) as previously described (Ohkura, K. *et al*. (2004)). SagA1 molecule has four heterocycle formable sites (31C-32C, 33F-34S, 44G-45S, 46T-47T), SagA2 has three heterocycle formable sites (31C-32C, 33F-34S, 44G-45S) and SPySagA has five formable sites (31C-32C, 33F-34S, 38G-39S, 45G-46S, 47G-48S) (underlined in Figure 3-1). The heterocycle-formed models of SagA molecules were constructed using Builder module and their structures were optimized under Consistence Valence Forcefield (CVFF). The molecular mechanics (MM) and molecular dynamics (MD) analysis of modeled SagA molecules (with/without heterocycle formation) were

performed by Discover 3 module under CVFF (Ohkura, K. *et al*. (2006−1)).

Energy profile analysis of bacterial SagA molecules.

The kinetic energy and the potential energy of modeled SagA molecules (with/without heterocycles) during simulation period (500 ps) were monitored, and the total energy (= kinetic energy + potential energy) profile was determined (Ohkura, K. *et al*. (2006−1)).

Electrostatic potential field analysis of SagA molecules.

After MD simulation period (500 ps), the electrostatic potential (ESP) fields of modeled SagA molecules (with/without heterocycles) were examined. The ESP fields of SagA molecules were examined using Delphi module with InsightII-Discover (Accelrys, Inc., San Diego, CA, USA). The +1.0 kT/e (green) and −1.0 kT/e (red) contour was displayed (Ohkura, K. *et al*. (2006−2); Ohkura, K. *et al*. (2009)).

3-4. Results

Molecular features of bacterial SagA molecules.

The differences of structural features among SagA molecules from SAA NCTC10713^{T} (SagA1 and SagA2), and SPySagA from *S. pyogenes* MGAS5005 were examined especially for intramolecular heterocycle formation. Total energy of modeled SagA1 molecule gradually decreased during the whole molecular dynamics (MD) simulation period (500 ps) and the average of total energy was 1830.8 kcal/mol (Figure 3-2A). In the heterocycle-formed SagA1 molecule, the total energy decreased smoothly from the start of

MD simulation at 100 ps and the average of the total energy (1784.7 kcal/mol) (Figure 3-2D) was lower than that of pre-heterocycled SagA1. The total energy of SagA2 molecule gradually converged during the MD simulation period (500 ps) and the average of total energy was 1574.9 kcal/mol (Figure 3-2B). Total energy of heterocycle-formed SagA2 molecule decreased within 100 ps of the start of simulation and the average of energy (1538.0 kcal/mol) (Figure 3-2E) was lower than that of the pre-heterocycled SagA2 molecule. The SagA2 molecule was lower in total energy than SagA1, and the SagA2 molecule was more stable than SagA1. For SPySagA molecule, the convergence of total energy indicates a similar tendency between pre- and post-heterocycle formation, but the total energy of heterocycle formed molecule was lower than that before heterocycle formation. The total energy of SPySagA decreased with heterocycle formation and the average of total energy decreased from 1799.3 to 1777.5 kcal/mol (Figure 3-2C and 3-2F).

The molecular structure containing the leader peptide and propeptide in modeled SagA1, SagA2, and SPySagA molecules were compared pre- and post-heterocycle formation (Figure 3-3). Before heterocycle formation, the border between the leader peptide (dark blue line) and propeptide (greenish line) was unclear (Figure 3-3B, D and F). After heterocycle formation, conformation of these SagA1, SagA2, SPySagA molecules were significantly changed. In SagA1 and SagA2 molecule, the leader peptide domain (dark blue line) was enclosed in the inner part of the molecule (Figure 3-3A and 3-3C). In SPySagA, the leader peptide domain (dark blue) was bundled with the propeptide domain (Figure 3-3E). These results suggest that the formation of heterocyclic structure is involved not only in the cytolytic activity of these molecules but also in the proper processing to convert them into their active form.

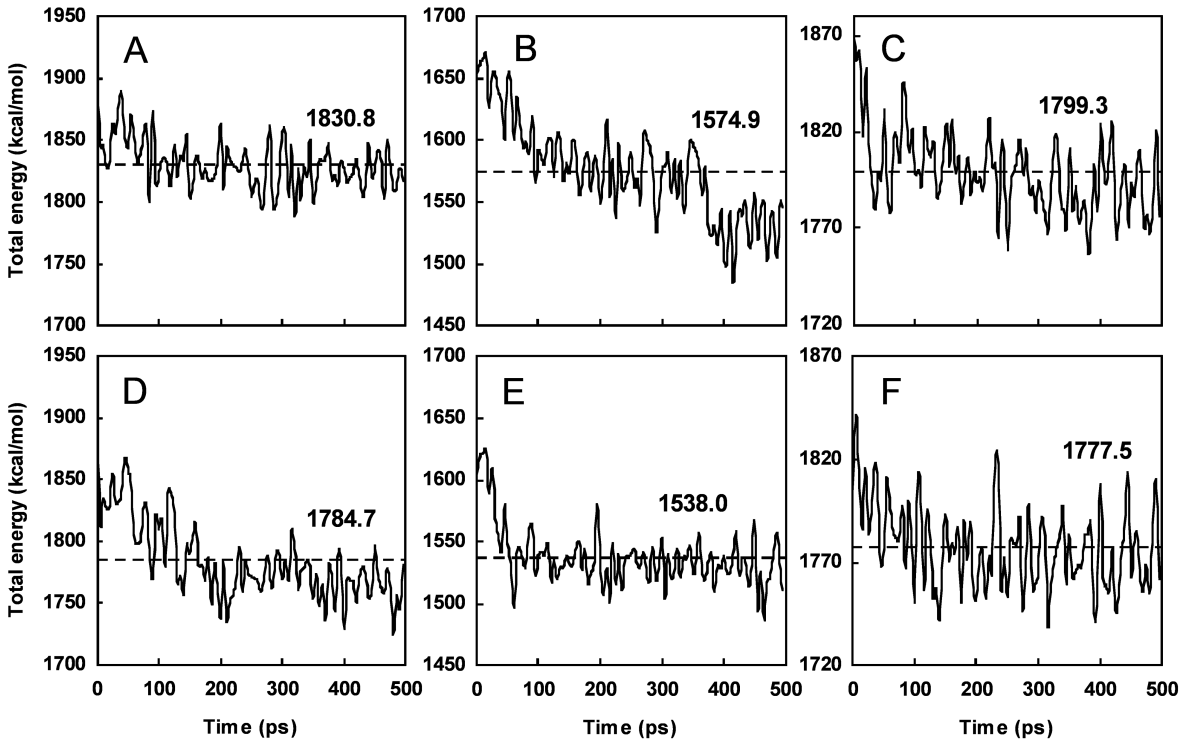


Figure 3-2. Total energy profile of SagA molecules.

Total energy of SagA1 (A, D) and SagA2 (B, E) from *S. anginosus* subsp. *anginosus* NCTC10713^{T}, and SagA (C, F) from *S. pyogenes* MGAS5005 were calculated. Before (A, B, C) and after (D, E, F) heterocycle formation. Broken lines indicate the average of total energy (kcal/mol) during the simulation period (500 ps), and the average values are shown.

Distribution of electrostatic potential (ESP) field (an index of reactivity with target molecule) also changed markedly due to heterocycle formation (Figure 3-3G, I and K). In pre-heterocycled SagA1 molecule, a negative ESP field (red cloud) covered the whole molecule (Figure 3-3H). After heterocycle formation, propeptide region of SagA1 molecule was covered with a negative ESP field (Figure 3-3G). For SagA2 molecule, positive (green cloud) and negative (ref cloud) ESP fields covered the whole molecule with and without heterocycle formation (Figure 3-3I and J). In heterocycled and pre-heterocycled SPySagA molecule, ESP distribution pattern was similar to that of SagA1 molecule (Figure 3-3K and L), respectively. These results indicate that the reactivity of these SagA molecules with their target molecule is extremely different in pre- and post-heterocycle formation.

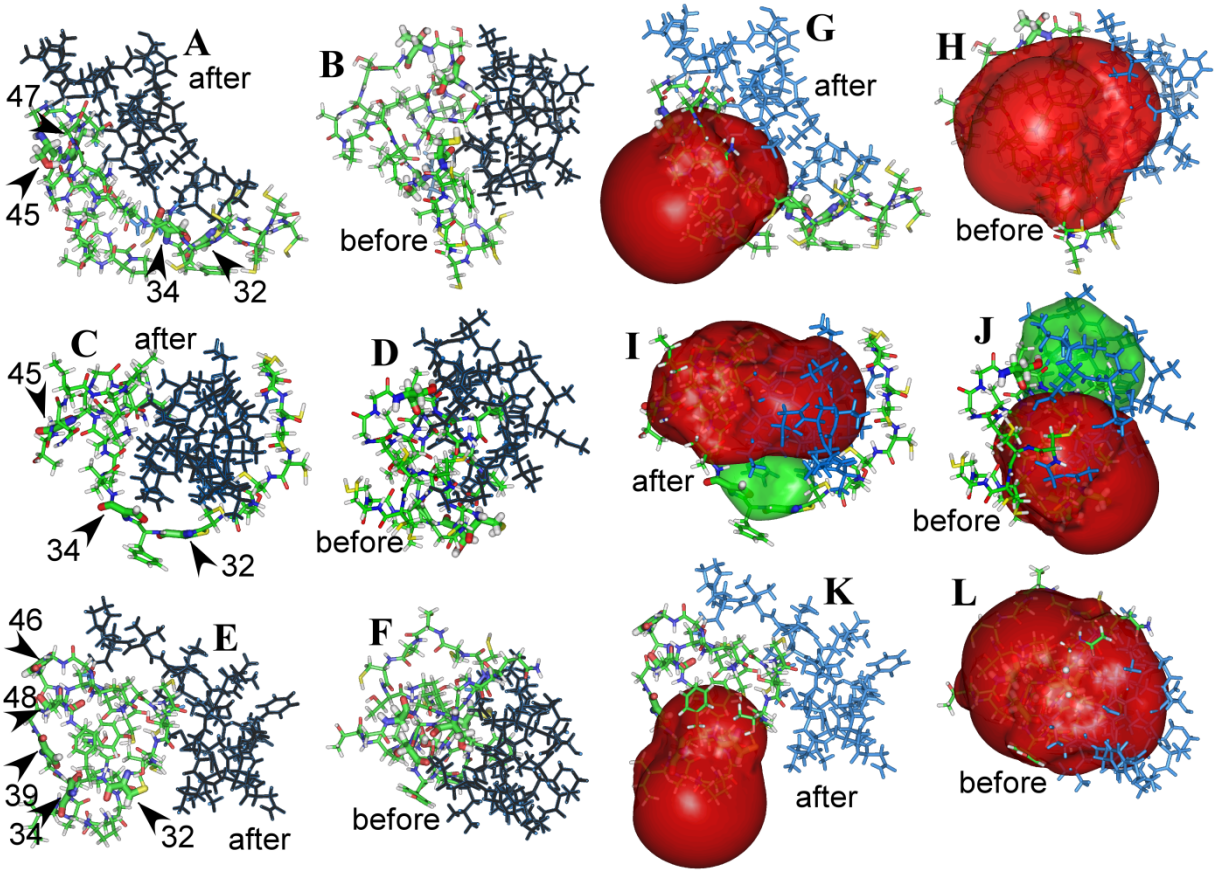


Figure 3-3. Molecular structure and electrostatic potential field of SagA molecules.

Molecular simulated structures of before (B, D, F)- and after (A, C, E)-heterocycle formation of SagA1 (A, B) and SagA2 (C, D) from *S. anginosus* subsp. *anginosus* NCTC10713^{T}, and SagA (E, F) from *S. pyogenes* MGAS5005 are shown. Dark blue lines indicate the leader peptide regions (A – F). Arrowheads indicate the heterocycle-formable sites (amino acid numbers) are shown. The distribution pattern of electrostatic potential field (negatively charged field: red cloud, positively charged field: green cloud) were also overlaid onto SagA1 (G, H), SagA2 (I, J), and SagA (K, L). In each electrostatic potential field figure (G – L), the results for before and after heterocycle formation are shown.

Table 3-I. Average of total energy of SagA molecules during heterocycle ring formation.

	Step					
Molecule	0	1st	2nd	3rd	4th	5th
SAA SagA1	1830.8	1772.3(33F-34S) 1789.3(31C-32C) 1801.7(46T-47T) 1828.0(44G-45S)	1762.1(46T-47T) 1805.1(31C-32C) 1828.8(44G-45S)	1755.0(31C-32C) 1766.6(44G-45S)	1784.7(44G-45S)	
SAA SagA2	1574.9	1617.6(31C-32C) 1630.8(33F-34S) 1635.9(44G-45S)	1623.9(33F-34S) 1624.8(44G-45S)	1538.0(44G-45S)		
S. p. SagA	1799.3	1811.9(31C-32C) 1815.8(33F-34S) 1818.1(45G-46S) 1823.9(47G-48S) 1827.5(38G-39S)	1796.9(33F-34S) 1817.9(45G-46S) 1824.1(38G-39S) 1830.8(47G-48S)	1803.8(38G-39S) 1806.2(47G-48S) 1815.4(45G-46S)	1798.0(45G-46S) 1803.6(47G-48S)	1777.5(47G-48S)

The values indicated the total energy (kcal/mol) of heterocycled SagA molecules in each step. Values in the parentheses indicated the amino acid positions (Figure 3-1) for heterocycle ring formation.

Order of heterocycle formation in SagA molecules.

SagA1 molecule has four heterocycle formable sites (31C-32C, 33F-34S, 44G-45S, 46T-47T), and the order of heterocycle ring formation was examined using the MM-MD energy simulation based on an assumption that the most stable intermediate structure will be preferentially and sequentially formed to reach to final form. Total energy of the first ring-formed SagA1 was compared for each four ring-formable sites, and the energy of the 33F-34S heterocycled molecule was the lowest (1772.3 kcal/mol in Table 3-I) among these

sites [1789.3 kcal/mol (31C-32C), 1801.7 kcal/mol (46T-47T), 1828.0 kcal/mol (44G-45S)]. The lowest total energy of the second, third, and fourth ring-formed SagA1 molecule was 1762.1 kcal/mol (46T-47T heterocycled SagA1), 1755.0 kcal/mol (31C-32C heterocycled SagA1), and 1784.7 kcal/mol (44G-45S heterocycled SagA1), respectively. SagA2 molecule has three heterocycle formable sites (31C-32C, 33F-34S, 44G-45S). The order of heterocycle formation in SagA2 was analyzed as well as SagA1, and the first ring was suggested to form at 31C-32C (1617.6 kcal/mol). The second and third rings were suggested to form at 33F-34S (1623.9 kcal/mol) and 44G-45S (1538.0 kcal/mol), respectively. SPySagA molecule has five heterocycle formable sites (31C-32C, 33F-34S, 38G-39S, 45G-46S, 47G-48S), and the order of heterocycle formation was determined by the MM-MD energy simulation as follows: first (31C-32C: 1811.9 kcal/mol), second (33F-34S: 1796.9 kcal/mol), third (38G-39S: 1803.8 kcal/mol), fourth (45G-46S: 1798.0 kcal/mol), and fifth (47G-48S: 1777.5 kcal/mol).

3-5. Discussion

The beta-hemolytic peptide called streptolysin S is well known and has been investigated in Pyogenic group streptococci (PGS) such as *S. pyogenes*, and this beta-hemolytic peptide was found to be encoded by the *sagA* gene in *sag* operon. Recently, the genes encoding twin streptolysin S-homologous peptides (SagA1 and SagA2) were found to exist in *sag* operon homolog of beta-hemolytic SAA strains, and each SagA molecule is responsible for the beta-hemolysis of beta-hemolytic SAA (Tabata, A. *et al*. (2013)). In this chapter, these SagA molecules from beta-hemolytic SAA and from *S. pyogenes* were modeled using molecular modeling technique and their molecular features were analyzed. These

SagA molecules had 3−5 heterocycle formable sites (Figure 3-1), and their structural features were expected to change according to the heterocycle formation. Total energy of SagA1, SagA2 and SPySagA molecules decreased with heterocycle formation during molecular dynamics (MD) simulation period (Figure 3-2), and these molecules were thought to be stabilized by the intramolecular heterocycle formation. Moreover, in the heterocycle-formed SagA molecules, total energy converged in a short time (100 ps in Figure 3-2D-F). From these results, it was also thought that the flexibility of SagA molecules changed by the intramolecular heterocycle formation. Subsequently, the change in this SagA molecular flexibility seems to affect the separation process occurred between leader peptide and propeptide (protoxin) region. The heterocycle formation is an important event not only for the functional appearance of SagA molecules but also for their molecular maturation. Indeed, in the heterocycle-formed SagA molecules, the boundary between the leader peptide and propeptide was plain (Figure 3-3A, C and E). For instance, it existed with leader peptide region contained in propeptide region in SagA1 (Figure 3-3A) and SagA2 (Figure 3-3C) molecule. The heterocycle formation of SagA molecules is also suggested to contribute to the compartmentalization of leader peptide and propeptide and divide the leader peptide from propeptide region. In this way, the later process for maturation of SagA molecules might be enhanced by the heterocycle formation.

The whole molecules of SagA1 and SPySagA were covered with the negatively-charged ESP field before heterocycle formation (Figure 3-3H and L). After heterocycle formation, the negatively-charged ESP field was observed in propeptide region of SagA1 and SPySagA molecules (Figure 3-3G and K). In SagA2 molecule, distribution pattern of the positive and negative ESP fields was significantly changed by heterocycle formation (Figure 3-3I: heterocycled, 3-3J: pre-heterocycled). Change of distribution of the

ESP field according to heterocycle formation might play a role in dividing the leader peptide region. ESP field distribution changed the functional appearance of general transport carriers, such as mitochondrial ATP/ADP carrier (Ohkura, K. *et al*. (2009)). The idea of a functional control mechanism by the heterocycle formation seen in SagA family can be applied to the structural analysis of *sag* operon products (*e.g.* SagB, SagC, and SagD) with the intra-molecular ring formation. Moreover, control of the molecular mechanism by intramolecular heterocycle ring formation can be also applied to the analysis of drug-excreting molecules (*e.g.* SagG, SagH, and SagI) related to the antibacterial and the anticancer drug resistance.

Chapter 4. Cytolytic Activity and Molecular Feature of Cardiotoxin and Cardiotoxin-like Basic Protein: Electrostatic Potential Field Is an Important Factor for Cell Lysis

4-1. Background

Poisonous snakes have been roughly divided into 4 types: vipers (Viperidae), mamushi (Crotalidae), cobras (Elapidae), and sea snakes (Hydrophiidae). A large number of proteins are categorized into the venom of cobra family snakes, and various studies have been performed on cobra toxins (Vidal, N. *et al.*, 2007). Elapidae front-fanged snakes, such as the cobra and sea snake, are very dangerous. The ED_{50} value of their venom to a mouse is tens of μg/kg ~ several mg/kg. Large-size cobras can accumulate up to 300 mg of poison, which has the ability kill 2.5 humans weighing 60 kg.

Cobra venom is composed of various types of proteins, such as short neurotoxins, long neurotoxins, cytotoxins (cardiotoxins (CTs)), and phospholipases. Neurotoxins consisting of 60 - 62 amino acid residues are called short neurotoxins, while those consisting of 66 - 74 amino acid residues are called long neurotoxins. CTs are basic polypeptides that exhibit cytolytic activity. In the previous chapter, electrostatic potential (ESP) was shown to be a useful index in analyses of molecular features. In this chapter, ESP is introduced as one of the key indexes for the cytolytic activities of snake venom-derived toxins and actually CTs derived from Thailand cobra (*Naja naja siamensis*) venom and the cardiotoxin-like basic proteins (CLBPs) derived from Thailand cobra (*Naja naja siamensis*), Taiwan cobra (*Naja naja atra*), and Indian cobra (*Naja naja*) venom were compared.

4-2. Introduction

The ESP field is one of the factors affecting the interactive property of proteins (Weiner, P. K. *et al*. (1982)). It can produce an attractive or repulsive force between proteins and target molecules. The distribution pattern of ESP in insulin molecules has been applied to control the medicinal effects of insulin preparations (*e.g.* rapid-acting, intermediate-acting, and long-acting insulin) (Inaba, T. *et al*. (2005); Inaba, T. *et al*. (2007)). In this chapter, the molecular features (*e.g.* ESP field distribution, stereo-hydrophobicity, and dihedral angle) participating in the cytolytic activities of CT and CLBP were examined. CT is a well-known cytolytic protein with a molecular weight of 7 kDa obtained from snake (*e.g. Naja naja siamensis*) venom. CT molecules consist of 60 amino acid residues and various types (CT-I, CT-II, CT-III and CT-IV) have been identified (Ohkura, K. *et al*. (1988)). These CT molecules exhibit cytolytic activities against various cells (*e.g.* human amnion Fogh-Lund cell). CLBP has also been purified from snake venom and consists of 62 amino acid residues (Inoue, S. *et al*. (1987); Inoue, S. *et al*. (1989)). The amino acid sequence of CLBP is very homologous to that of CT, with two amino acid residues being inserted into CT molecule. However, CLBP does not express cytolytic activity in cultured cells. Molecular models of the four CTs (CT-I, -II, -III, and -IV) from *Naja naja siamensis* and three CLBPs from *Naja naja siamensis* (CLBP-si), *Naja naja atra* (CLBP-at), and *Naja naja* (CLBP-na) were constructed based on the X-ray data (Protein data bank ID=1CDT) of *Naja mossambica mossambica* CT VII4. The molecular features of these CTs and CLBPs were analyzed and their relationships with cytolytic activity were discussed.

4-3. Materials and Methods

Isolation of toxins from cobra venom.

Lyophilized cobra venom was obtained from Miami Serpentarium (USA). Cobra toxins were separated by gel filtration chromatography using a Sephadex G-75 column and by following ion-exchange chromatography using a CM-cellulose column, described previously (Kaneda, N. *et al*. (1977)). Briefly, venom sample was dissolved in 1% acetic acid and applied to a Sephadex G-75 column equilibrated with the 1% acetic acid. The fractions containing proteins of 5-10 kDa was lyophilized and subjected to a CM-cellulose column equilibrated with 0.1 M sodium acetate (pH 6.0). Sample proteins were eluted with a linear gradient from 0.1 M sodium acetate (pH 6.0) to 0.5 M sodium acetate buffer (pH 6.5). Each protein fraction was collected, and desalted using Sephadex G-25 column and lyophilized. Homogeneity of obtained protein was confirmed by reversed-phase HPLC equipped with a SynChropak RP-8 column (SynChrom, USA) with the linear gradient from 0% to 45% acetonitrile solvent containing 0.1 % trifluoroacetic acid. The S-carboxymethylated (Cm) derivatives of obtained cytotoxin and cytotoxin-like basic protein were prepared as previously reported (Crestfield, A. M. *et al*. (1963)).

Cytotoxicity toward FL cells.

Fogh-Lund (FL) cells were suspended in phosphate-buffered saline (PBS) at a population of 2.5 x 10^6 cells/ml. Each obtained protein fraction at various concentrations was separately added to cell suspensions, and the mixtures were incubated at 37 ^oC for 30 min. Cytotoxic activity was measured by trypan-blue exclusion test (Ohkura, K. *et al*. (1988); Inoue, S. *et al*. (1987)). The cytotoxicity was expressed as ED_{50} (protein concentration

required to cause 50 % cell lysis).

Molecular analysis of cardiotoxins and cardiotoxin-like basic proteins.

The molecular models of CTs and CLBPs were constructed based on X-ray data of *Naja mossambica mossambica* CT VII4 (Protein Data Bank ID=1CDT) using InsightII-Discover with Homology module (Accelrys Inc., San Diego, CA, USA) (Ohkura, K. *et al*. (2004); Ohkura, K. *et al*. (2006–1)). Energy minimization of molecular models was performed using a Consistent Valence Forcefield (CVFF). The electrostatic potential fields of CTs and CLBPs were calculated and the +1.0 kT/e (green) and –1.0 kT/e (red) contour was displayed (Ohkura, K. *et al*. (2006–2)). The molecular dynamics (MD) simulation (500 ps) of these model proteins was performed using Discover module. The parameters of molecular orbital analysis (*e.g.* solvation free energy, which is an index of stereo-hydrophobicity) were determined for every 15 amino acid fragments, which were derived from N-terminal of simulated protein (Ohkura, K. *et al*. (1999)). The dihedral angles of peptide-bond (-CONH-) at the insertion sites (CTs: Cys^{3}-Asn^{4}, Lys^{23}-Met^{24}, CLBPs: Cys^{3}-His^{4}, His^{4}-Asn^{5}, Lys^{24}-Ala^{25}, Ala^{25}-Thr^{26}) were determined from MD analysis trajectory.

4-4. Results

Cytolytic activity of cardiotoxins and cardiotoxin-like basic proteins.

Protein concentration of CTs and CLBPs required to lyse 50% (ED_{50}) of Fogh-Lund (FL) cells are summarized in Table 4-I. ED_{50} values of CT-I, CT-II, CT-III and CT-IV were 5.4, 10.6, 16.1 and 12.6 μg/ml, respectively. CLBP-si, CLBP-at and CLBP-na had no cytotoxicity (>100 μg/ml) toward FL cells.

Table 4-I. 50% Cytotoxicity (ED_{50}) of cardiotoxins and cardiotoxin-like basic proteins.

Protein	ED_{50} (μg/ml)
CT-I	5.4
CT-II	10.6
CT-III	16.1
CT-IV	12.6
CLBP-si	>100
CLBP-at	>100
CLBP-na	>100

Fogh-Lund (FL) cells were suspended in phosphate-buffered saline (PBS) at a concentration of 2.5 x 10^6 cells/ml. Each protein fraction at various concentrations was separately added to the cell suspensions and the mixtures were incubated at 37 ^oC for 30 min. Cytotoxicity was expressed as ED_{50}: the protein concentration required to cause 50 % cell lysis. CT-I, -II, -III, and -IV were prepared from *Naja naja siamensis*. CLBPs were prepared from the venom of *Naja naja siamensis* (CLBP-si), *Naja naja atra* (CLBP-at), *Naja naja* (CLBP-na).

Molecular feature of cardiotoxins and cardiotoxin-like basic proteins.

The cardiotoxins consist of 60 amino acid residues and two amino acid insertions (His^4 and Ala^{25}) were detected in the CLBP molecules (Figure 4-1). CTs and CLBPs were previously shown to possess four S-S bridges (CT: C^3-C^{21}, C^{14}-C^{38}, C^{42}-C^{53}, C^{54}-C^{59}, CLBP: C^3-C^{22}, C^{15}-C^{40}, C^{44}-C^{55}, C^{56}-C^{61}) (Bhaskaran, R. *et al*. (1994)). Molecular configuration of these proteins collapsed when MD analysis was performed without these four S-S bridges. Structural collapse was absent in MD examination that was tempered with four S-S bridges. Total energies of CT-I, CT-II, CT-III and CT-IV during MD simulation (500 ps) converged within the range of 2518.5 to 3081.7 kcal/mol. The range of the total energies in CLBP

simulations converged from 3106.3 to 3166.4 kcal/mol and the total energies of CLBPs were higher than those of CT molecules (Figure 4-2). Stable structures of CTs and CLBPs were extracted from MD analysis trajectory data and divided from the N-terminal into 10 peptides every 15 amino acids. Solvation free energies (dGWs) of these divided peptides were determined using MO analysis and their stereo-hydrophobicities were examined. The dGW values of CT-I peptides ranged from –1686.2 to –799.7 kJ/mol (average –1141.7 kJ/mol) (Figure 4-3). In CT-II, CT-III and CT-IV-derived peptide analysis, the dGW profiles were essentially the same as that of CT-I and their averages were –1111.7, –1088.5 and –1013.1 kJ/mol, respectively. CLBP-derived peptides had similar dGW profiles with average dGW value of –1243.9 (CLBP-si), –1115.7 (CLBP-na), –1179.4 (CLBP-at) kJ/mol. No significant difference was observed in stereo-hydrophobicities of these CTs and CLBPs.

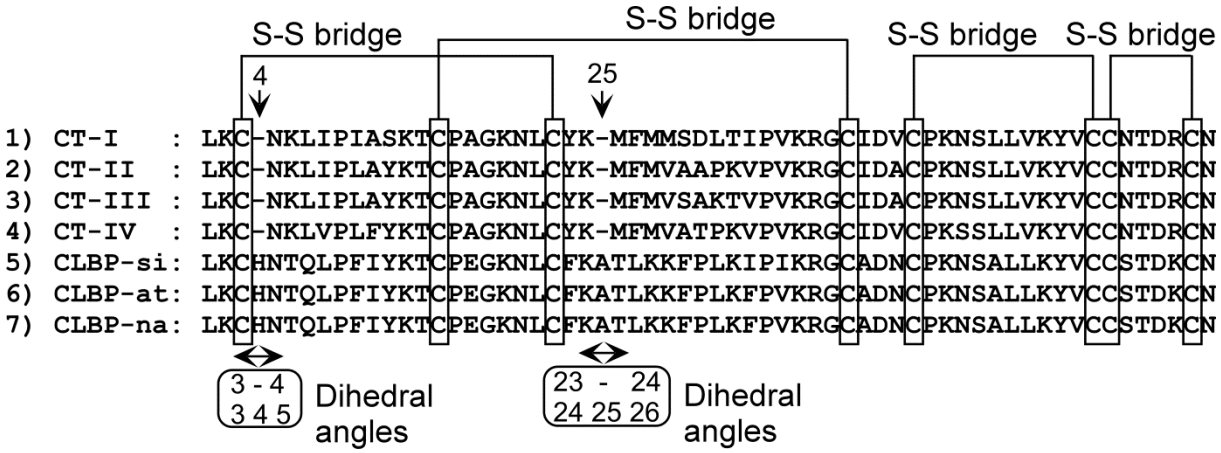


Figure 4-1. Amino acid sequence of cardiotoxin (CT) and cardiotoxin-like basic protein (CLBP).

Amino acid sequences of CTs and CLBPs were compared and their S-S bridges were indicated. CT-I, -II, -III, -IV: cardiotoxins from *Naja naja siamensis*. CLBP-si, -at, -na: CLBPs from *Naja naja siamensis*, *Naja naja atra*, *Naja naja*, respectively. Amino acid insertion was observed at His^4 and Ala^{25} in CLBP molecule. The dihedral angles at insertion regions (CTs: 3/4, 23/24, CLBPs: 3/4, 4/5, 24/25, 25/26) were monitored during MD simulation.

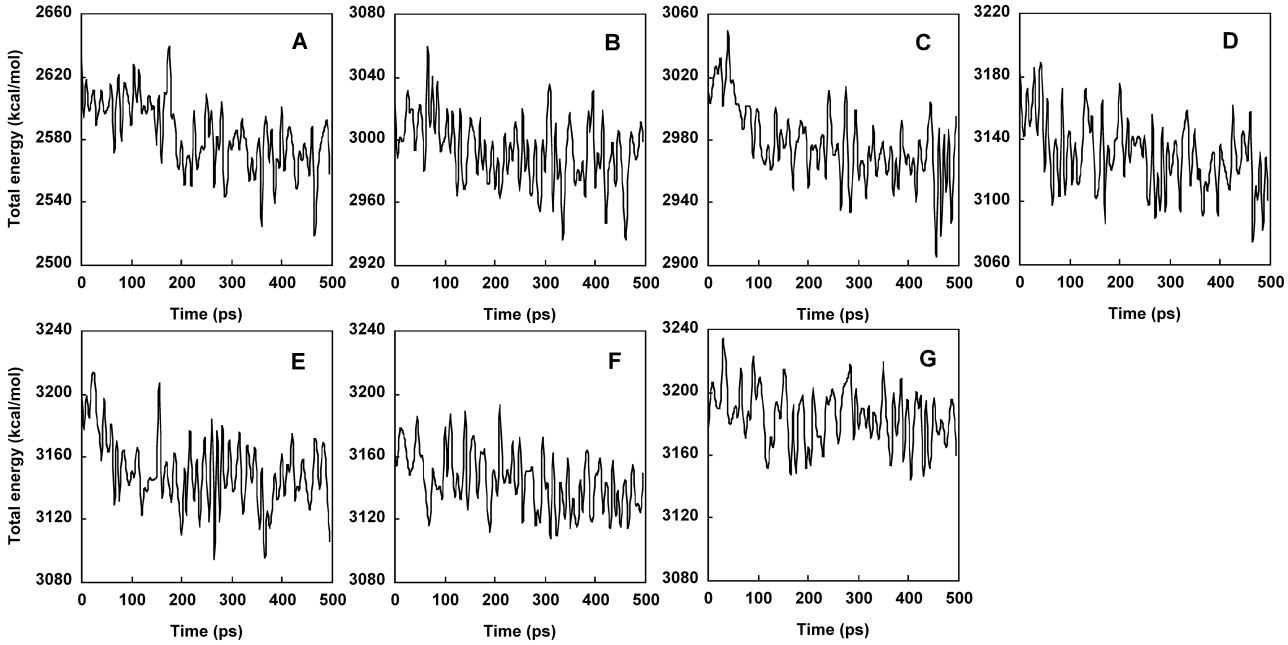
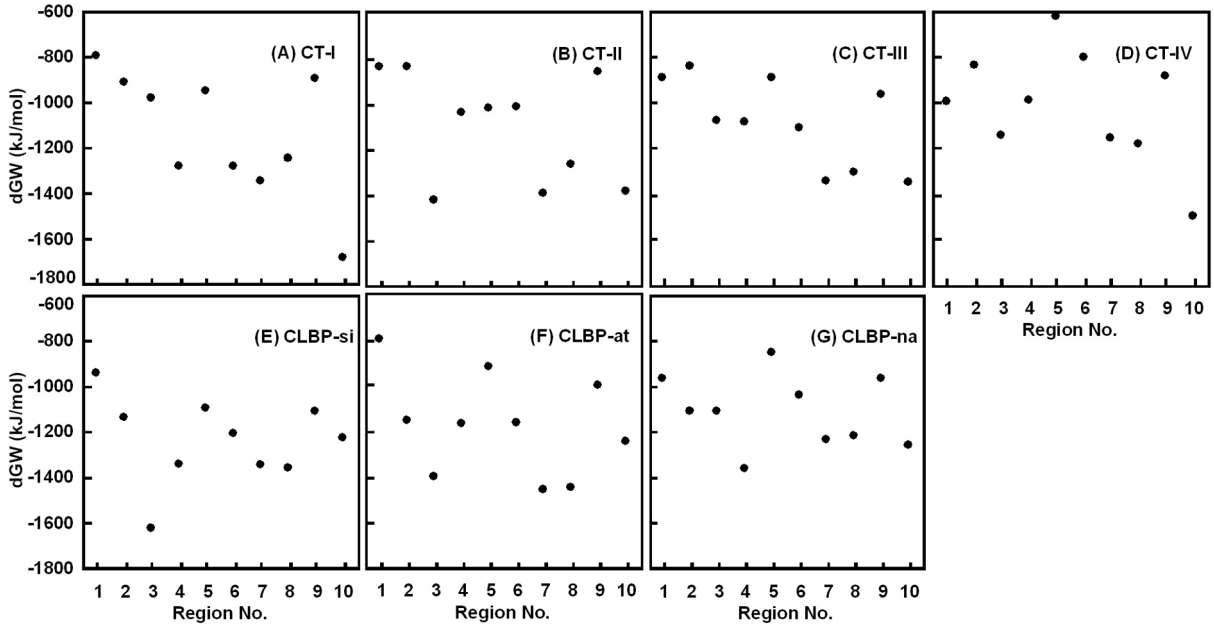


Figure 4-2. Total energy profiles of CTs and CLBPs during MM/MD analysis period (500 ps).

CT-I (A), -II (B), -III (C), -IV (D): cardiotoxins from *Naja naja siamensis*. CLBP-si (E), -at (F), -na (G): CLBPs from *Naja naja siamensis*, *Naja naja atra*, *Naja naja*, respectively.



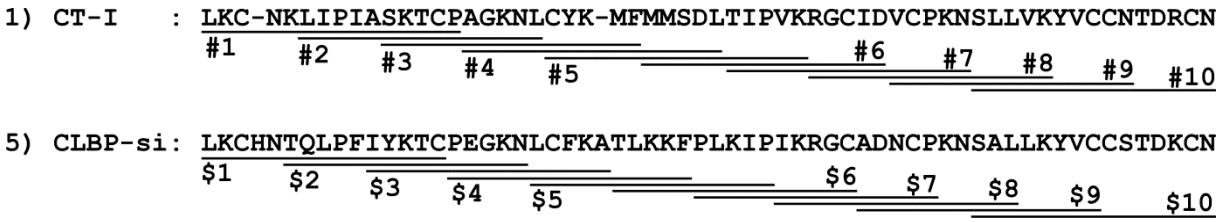


Figure 4-3. Solvation free energy profile of CTs and CLBPs.

Cardiotoxin-I (A: CT-I), -II (B: CT-II), -III (C: CT-III), -IV (D: CT-IV) from *Naja naja siamensis*. Cardiotoxin-like basic protein from *Naja naja siamensis* (E: CLBP-si), *Naja naja atra* (F: CLBP-at), *Naja naja* (G: CLBP-na), respectively. Stable structures of CTs and CLBPs were extracted from MD trajectory data and divided from N-terminal into 10 peptides every 15 amino acids. The region number of cardiotoxin (*e.g.* CT-I (#1 − #10)) and cardiotoxin-like basic protein (*e.g.* CLBP-si ($1 − $10)) peptides were shown according to the amino acid sequence of Figure 4-1. Solvation free energies (dGWs) of these divided peptides were determined using a molecular orbital (MO) analysis.

To analyze the molecular distortion at two insertion sites (*e.g.* His^4 and Ala^{25} in CLBP), dihedral angle changes in the insertion regions (CTs: 3/4, 23/24, CLBPs: 3/4, 4/5, 24/25, 25/26; Figure 4-1) were monitored during MD analysis. The biases of dihedral angles were not observed in CT or CLBP MD examination and the angles were approximately –180 or +180 degrees. Average of dihedral angles of CT and CLBP molecule did not exhibit bias to either plus or minus values (Figure 4-4). No significant differences were observed in structural features of the insertion regions in CT and CLBP molecule. Dipole moment intensity of His^4-inserted region in CLBP (CLBP-si: Leu^1-Cys^{15}) was 20.236 debye and its direction was the same as that of the CT-I dipole moment (18.327 debye, CT-I: Leu^1-Pro^{15}). In Ala^{25}-inserted region in CLBP (CLBP-si: Pro^{16}-Phe^{30}), intensity of dipole moment was 66.277 debye and was in the opposite direction to that of CT-I (60.725 debye, Ala^{16}-Leu^{30}) (Figure 4-5).

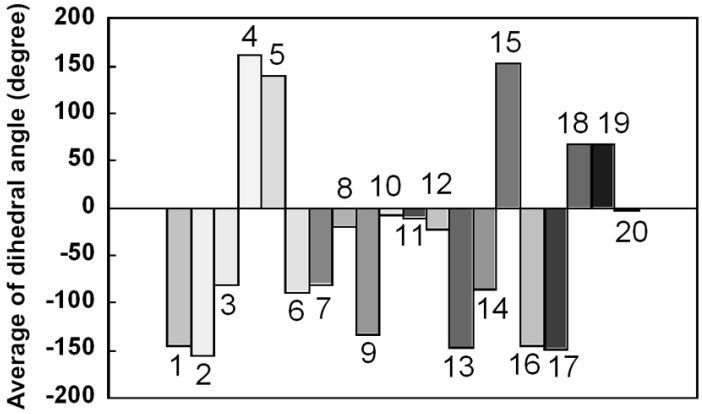


Figure 4-4. Average of dihedral angles of MD-simulated CTs and CLBPs at insertion regions.

Column numbers indicate the dihedral angle monitored insertion sites of CTs and CLBPs. Columns 1, 2: site 3/4, 23/24 of CT-I; columns 3, 4: site 3/4, 23/24 of CT-II; columns 5, 6: 3/4, 23/24 of CT-III; columns 7, 8: 3/4, 23/24 of CT-IV; columns 9−12: 3/4, 4/5, 24/25, 25/26 of CLBP-si; columns 13−16: 3/4, 4/5, 24/25, 25/26 of CLBP-at; columns 17−20: 3/4, 4/5, 24/25, 25/26 of CLBP-na, respectively.

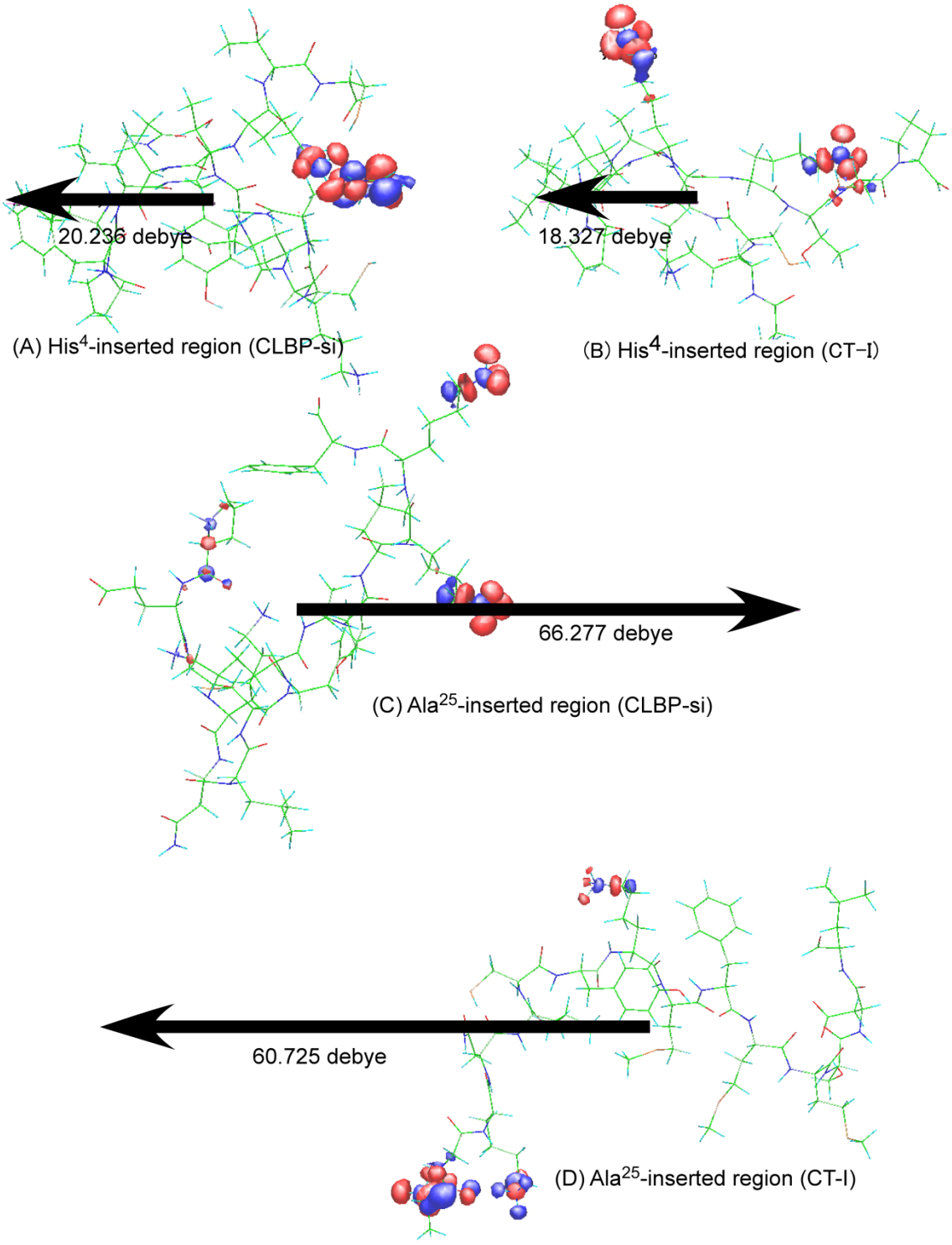


Figure 4-5. Dipole moment of amino acid inserted regions of CLBP-si and CT-I. Dipole moment direction and intensity of insertion site including peptides (A: CLBP-si Leu^{1}-Cys^{15}, B: CT-I Leu^{1}-Pro^{15}, C: CLBP-si Pro^{16}-Phe^{30}, D: CT-I Ala^{16}-Leu^{30}) at HOMO states are shown.

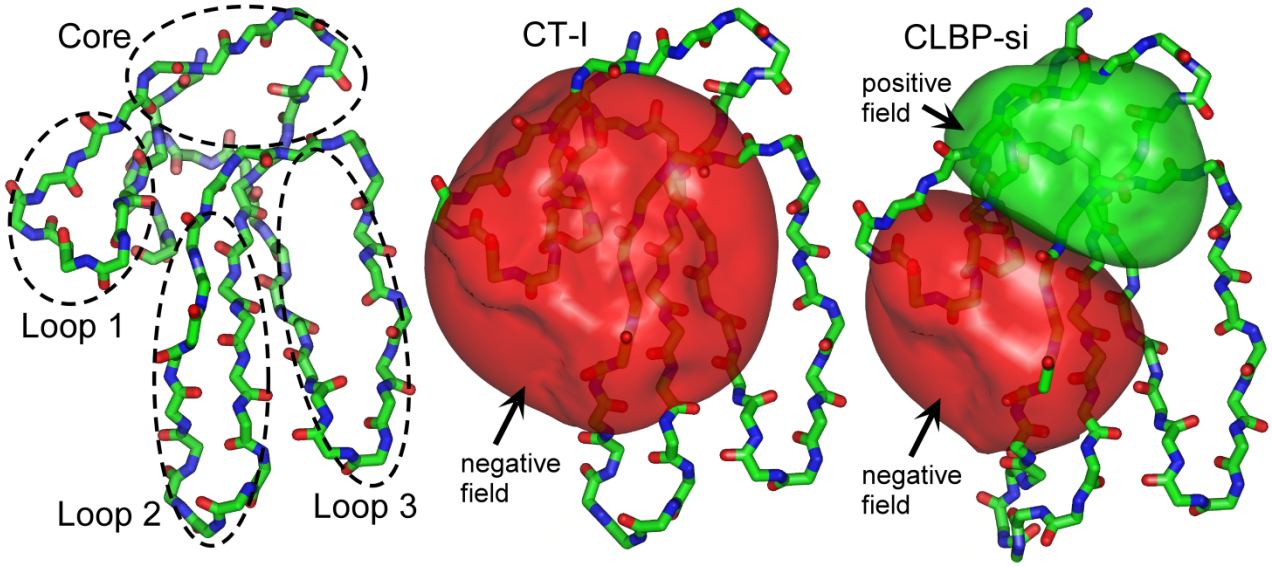


Figure 4-6. Electrostatic potential field of CT and CLBP.
CTs and CLBPs have one core region and three loops. CT-I molecule was covered with negative charged (red) ESP field. CLBP-si was covered with positive (green) and negative charged (red) ESP field.

Electrostatic potential fields of CT and CLBP.

The CT and CLBP molecules consist of one core region and three loops (Figure 4-6). A negative ESP field developed with whole CT-I molecule. A negative ESP field was observed in CT-II, CT-III and CT-IV, similar to CT-I. On the other hand, a positive ESP field existed in core region of CLBP-si molecule, while a negative ESP field existed in loop 1 and loop 2 region. Similar ESP distribution was confirmed in CLBP-at and CLBP-na molecules. A significant difference was observed in distribution pattern of ESP between CTs and CLBPs. When His^{4} and Ala^{25} residues of CLBP were deleted, distribution pattern of ESP changed to a CT-like pattern (Figure 4-7). In His^{4}-deleted CLBP model, a change was observed in distribution of ESP and the ESP pattern was altered to a CT-like pattern.

No changes were noted in ESP pattern in Ala^{25}-deleted CLBP model. Furthermore, ESP pattern did not change to a CT-like pattern in the CLBP model in which Leu^{1} and Lys^{2} residues were removed to more simply arrange the N-terminal length.

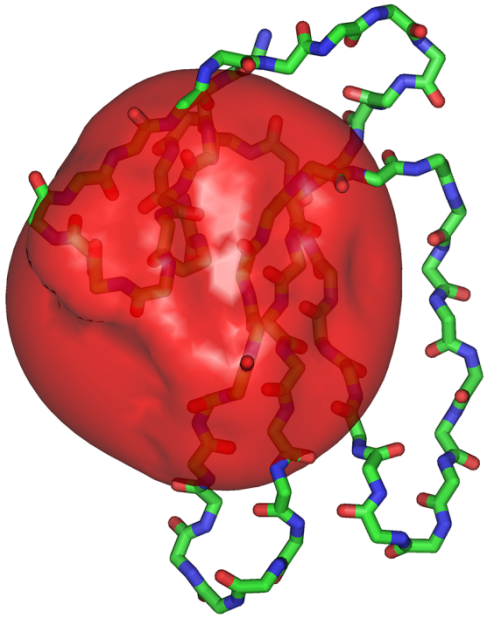


Figure 4-7. ESP of His^{4}, Ala^{25}-deleted CLBP-si.
The distribution of ESP of His^{4} and Ala^{25} deleted CLBP-si changed to a CT-like pattern. Negative charged ESP field was observed.

4-5. Discussion

Total molecular dynamics (MD) energies of CLBPs (3106.3 to 3166.4 kcal/mol) were higher than those of CTs (2518.5 to 3081.7 kcal/mol). These results suggest that the mobility and the reactivity of CLBPs are higher than those of CTs. However, CLBPs did not

exhibit any cytotoxicity that reflected reactivity with cell membrane. An appropriate balance between the molecular mobility and the MD energy may exist to show reactivity with cell membrane. No significant difference was observed in stereo-hydrophobicities of CTs and CLBPs (*e.g.* dGW profiles) (Figure 4-3). In molecular torsion analysis at the amino acid insertion regions of CT and CLBP molecules, no significant difference was noted in dihedral angles (one of the index of molecular structure) (Figure 4-4). These results indicate that other factors, besides molecular hydrophobicity and structural torsion, regulate the cytolytic activity.

Significant difference was observed in distribution of ESP field of CLBPs and CTs, while CLBPs had positive and negative charged ESP fields. These ESP field features may be related to loss of the cell lytic activity of CLBP molecules. Analysis of the protein function previously revealed that ESP field pattern is an important index (Weiner, P. K. *et al*. (1982)). Because distribution pattern of ESP on insulin molecules regulates its efficiency, various types of insulin molecules have been prepared, including rapid-acting, short-acting, intermediate-acting and long-acting insulins (Inaba, T. *et al*. (2007)). The stable insulin structures with hexameric, dimeric and monomeric forms have been identified (Derewenda, U. *et al*. (1989); Smith, G. D. *et al*. (1984)). Insulin monomer molecule has the ability to react with the receptor proteins and is the main bioactive form. Insulin hexamer is converted into the monomer type insulin through the dimer leading to the reappearance of activity. The amount of time required to express medicinal effect is dependent on conversion time to the monomer insulin form. ESP field of rapid-acting insulin is designed so that it cannot easily interact with itself and exists in the monomer insulin form. In long-acting insulin, ESP field has been modified to easily form hexameric structure of insulin. A time lag in alteration from the hexamer form to the monomer form is expected due to this ESP modification

resulting in a persistent medicinal effect. The modes of inter-molecular interactions of CTs and CLBPs are considered to differ because their distribution patterns of ESP field were found to be significantly different. CLBPs had both positive and negative charged ESP field and could easily interact with other CLBP molecules (Figure 4-6). This inter-molecular interaction of CLBPs appears to affect the titer of cytotoxicity. Amino acid sequence of CLBP was homologous with blood coagulation factor X-binding protein (X-bp) (Atoda, H. *et al*. (1998)). X-bp is a component of cobra venom and X-bp dimer has been shown to interact with platelet membrane-binding Gla domain of the coagulation factor X, thereby inhibiting blood coagulation system (Figure 4-8). CLBP dimer model had an appropriate ESP field and a good interactive profile with factor X Gla domain was observed during molecular simulation (Figure 4-9). These results indicate that ESP profile of protein drugs (*e.g.* insulin) is one of the key factors determining function (titer) control. In order to examine ESP-dependent function control of protein drugs, I have been designed a neo-CLBP molecule, which has a CT-like ESP field. Moreover, the amino acid insertion at 25th Ala of CLBP affect the dipole moment, and the moment direction of CT-derived peptide (Ala^{16}-Leu^{30}) and CLBP-derived peptide (Pro^{16}-Phe^{30}) are opposite. Because an intermolecular interaction occur when mutual dipole moments exist parallely or anti-parallely, and dipole moment can use as the control tool of molecular interaction. ESP field distribution and dipole moment direction are significant index for molecular reactivity. Appropriately designed membrane-acting probes, which have interactive features (*e.g.* ESP field distribution, dipole moment direction), appear useful tool of drug delivery by membrane modification.

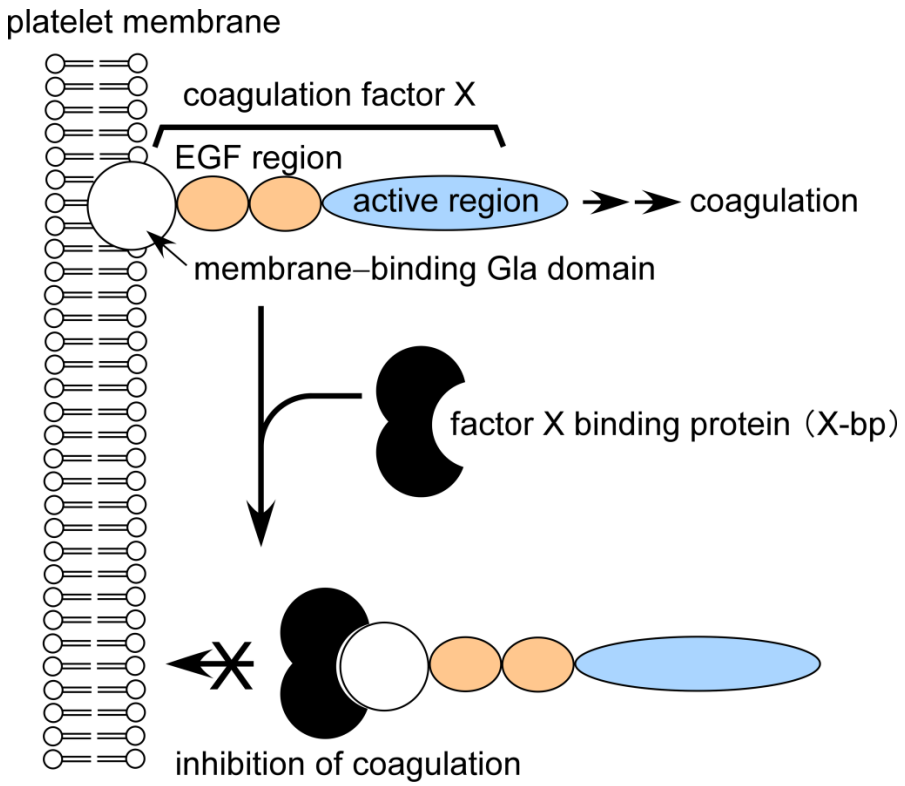


Figure 4-8. Inhibition mechanism of coagulation by X-bp.
Cobra venom derived blood coagulation factor X-binding protein (X-bp) inhibits the coagulation activity of factor X.

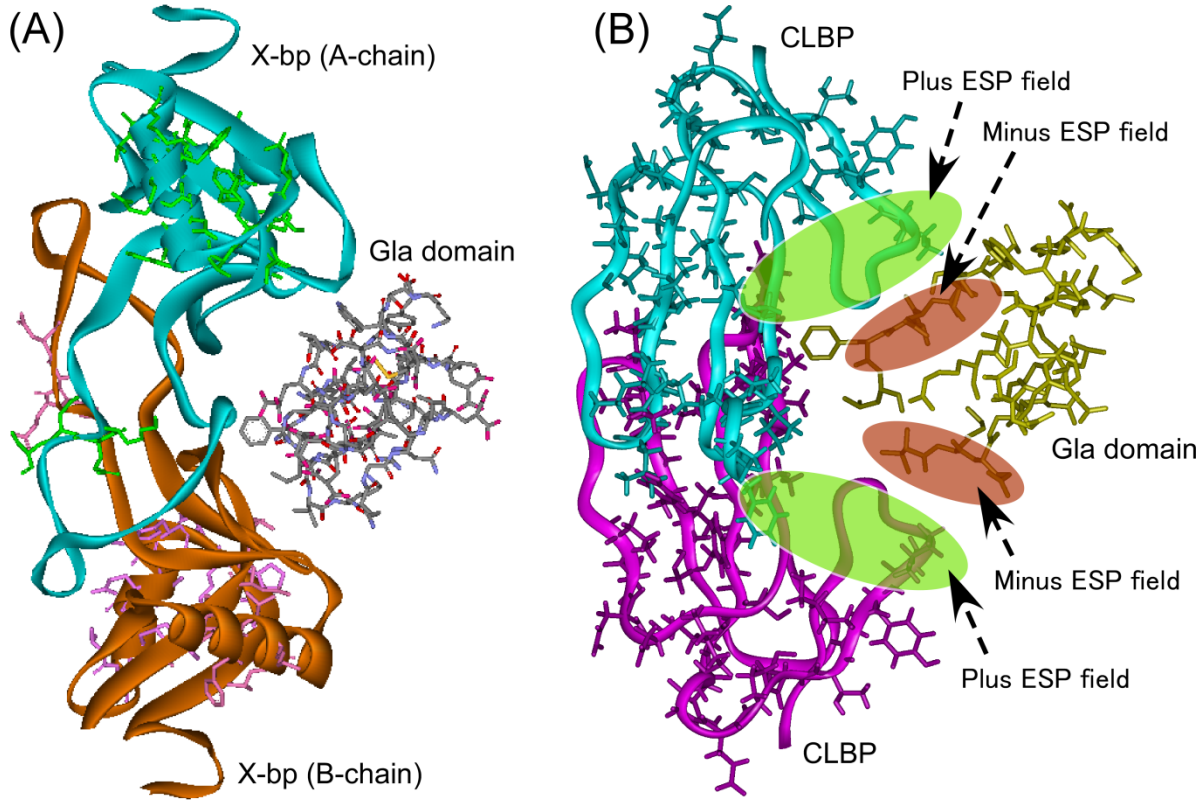


Figure 4-9. Interaction with Gla domain of coagulation factor X.

A: X-bp is a component of cobra venom and the X-bp dimer has been shown to interact with the platelet membrane-binding Gla domain of coagulation factor X, thereby inhibiting the blood coagulation system.

B: The CLBP dimer model had an appropriate ESP field and a good interactive profile with the factor X Gla domain was observed during the molecular simulation analysis.

Chapter 5 Overview

Membrane-acting proteins participate in an infection and other diseases, and the structure analyses of these proteins are useful for interpretation of pathogenesis and development of therapy. Structural analyses of membrane-acting proteins elucidate the pharmacophore, and can lead to the development of medical technologies and tools. For such purpose, this research introduced several cases focusing on the *in silico* molecular analyses of membrane-acting proteins.

In traditional CDC, conserved 11mer region plays important role on interaction with cell membrane. ILY, PLO, VLY, and Sm-hPAF have different cholesterol dependency from traditional CDC. These cell lytic proteins exhibit diagnostic membrane-acting features, and their 11mer regions classified into four types. ILY interacted with hCD59 and formed single hydrogen bond, and arranged their location each other. ILY 11mer region does not directly affect the interaction process between ILY and hCD59. However ILY 11mer region may coordinate the membrane behavior, and future verification have been waited. Structural feature of 11mer region can apply to design of membrane-acting probes, and lead to development of useful tools (*e.g.* membrane-attached drug delivery system). SagA peptides are small molecules relatively, and they seem to arrive to cellular mitochondrial inner membrane easily. Modification of mitochondrial membrane involve the apoptosis, and understanding of SagA peptide structural peculiarity may lead to the development of anticancer tools. Membrane-acting probes, which are including SagA motif structure, have been designed. Cytolytic activity of cobra CT was conducted by ESP field, and CLBP which had distinctive ESP field, exhibited no cytolytic action. CLBP showed the X-bp like structural features, and the control capability to coagulation system of CLBP have been

examined.

Structural simulation of low-molecular-weight compounds may be performed easily in current workstations. The analytical technique for low-molecular-weight compounds may be partially applied to protein structural analyses. Actually, such analyses of membrane-acting proteins are possible using currently available powerful central processing unit. Moreover, interaction analyses of proteins with equally ordered lipid bilayer are considered to be challenging. In such cases, because a large number of non-specific interactive factors co-exist in the phospholipid bilayer system, a vast amount of calculation is necessary. It may be similar to a simulation of the occurrence of a tidal wave and its course.

When an endogenous substance, such as a hormone, expresses biological activity in the living body, binding to a target receptor is the first step. In this binding reaction, the structures of attacking molecule (*e.g.* a hormone or enzyme) and receptive molecule (*e.g.* a receptor) change dynamically. The receptor protein is maintained by the cell membrane, and, thus, the condition of the membrane (*e.g.* fluidity) affects the interaction between the attacking molecule and its receptor. The proteins described herein exhibit membrane-acting natures, and their molecular weights are 55kDa (CDC), 3kDa (SagA), 7kDa (CT), respectively. The electrical charges of these proteins vary as follows: acidic (SLY; isoelectric point (IP) 4.74), neutral (SagA1; IP 7.45), and basic (ILY; IP 9.97, CT-II; IP 9.42) at physiological pH. Even if these proteins have different features (*e.g.* molecular weights, charges), they still interact with the cell membrane and modify the structure of the membrane. These protein factors utilize appropriate properties (*e.g.* stereo-hydrophobicity, electrostatic potential, dipole moment) to express an individual membrane-acting mechanism. No special nature is necessary for interacting with the cell membrane; a flexible approach/fitting may be important.

Hydrophobicity is one of the motive forces for protein holding, protein-protein interactions, and the lipid bilayer formation. The non-polar regions of membrane-acting proteins assume the hydrophobic interaction with target regions of membrane components. Therefore, hydrophobicity is playing an important role in expression of the molecular features of a membrane-acting proteins. In the early phase of molecular interaction, membrane-acting proteins and their acceptor molecules (*e.g.* receptor molecules) exist independently and have individual entropy. As the interaction progresses, the membrane-acting proteins and their acceptor molecules interact cooperatively, and their mutual individuality disappears. In other words, the interaction between the membrane-acting proteins and their acceptor molecules is called the disappearance of entropy. Protein-ligand (inhibitor) interactions are examined using Isothermal Titration Calorimetry (ITC), and molecular parameters (*e.g.* enthalpy, entropy) are determined (Klebe, G. 2015). Enthalpy (ΔH) and entropy (ΔS) are used as a constituent of binding free energy (ΔG), and these parameters are applied to drug design. An optimized drug, which is highly efficient and has few side effects, exhibits more enthalpy-dominated features (Freire, E. 2008). This dominant analysis of enthalpy is available for obtaining a clearer understanding on the interactive force of membrane-acing factors with target molecules. Two types of membrane-acting proteins were introduced herein. The first type is clustering factors (*e.g.* PFO and ILY) and the second type is the non-clustering factors (*e.g.* SagA1 and CT). A dominant analysis of enthalpy with clustering/non-clustering factors will be performed in the future in order to obtain a better understanding of the interaction mechanism used by membrane-acting factors.

In some cases, the intervention of a water molecule is necessary in the protein-protein interactions. In analyses involving membranes, the direct participation of

water molecule is difficult because biomembranes are hydrophobic. When the interaction features of biomembrane and membrane-acting proteins are interpreted, information from the simulated analyses on water molecule-intervening interaction may be useful. The intervention of water and ion (*e.g.* Ca^{2+}, Cl^{-}) molecules in the interaction between membrane-acting proteins and biomembranes may affect the functional expression of membrane-acting proteins. A novel analysis concept is needed in order to obtain a better understanding of the role of solvent and mineral molecules in membrane-involved reaction fields.

Acknowledgments

The author thanks to Profs. Yoshiro Uto and Hitoshi Matsuki, Department of Biological Science and Technology, Life System, Institute of Technology and Science, Tokushima University Graduate School, for helpful review and advice.

The author is sincerely grateful to Prof. Hideaki Nagamune, Department of Biological Science and Technology, Life System, Institute of Technology and Science, Tokushima University Graduate School, who has kindly given valuable suggestions and continuing encouragement through this study.

The author thanks to Associate Prof. Toshifumi Tomoyasu, Department of Biological Science and Technology, Life System, Institute of Technology and Science, Tokushima University Graduate School, for helpful advice.

The author is a special thanks to Prof. Kazuto Ohkura, Faculty of Pharmaceutical Sciences, Suzuka University of Medical Science, for experimental suggestion and helpful discussion on this study. The author also thanks to Prof. Hitoshi Hori, Department of Biological Science and Technology, Life System, Institute of Technology and Science, Tokushima University Graduate School, for helpful discussion.

The author thanks to Assistant Prof. Atsushi Tabata, Department of Biological Science and Technology, Life System, Institute of Technology and Science, Tokushima University Graduate School, for experimental advice and helpful discussion.

Finally I would like to thank my own family for their support and encouragement.

References

Aroutcheva, A. A., Simoes, J. A., Behbakht, K., Faro, S. (2001): *Gardnerella vaginalis* isolated from patients with bacterial vaginosis and from patients with healthy vaginal ecosystems. *Clin. Infect. Dis.* **33**, 1022–1027.

Atoda, H., Ishikawa, M., Mizuno, H., Morita, T. (1998): Coagulation factor X-binding protein from *Deinagkistrodon acutus* venom is a Gla domain-binding protein. *Biochemistry* **37**, 17361–17370.

Bhaskaran, R., Huang, C-C., Tsai, Y-C., Jayaraman, G., Chang, D-K., Yu, C. (1994): Cardiotoxin II from Taiwan cobra venom, *Naja naja* atra. *J. Biol. Chem.* **269**, 23500–23508.

Billington, S. J., Jost, B. H., Cuevas, W. A., Bright, K. R. (1997): The *Arcanobacterium* (*Actinomyces*) *pyogenes* hemolysin, pyolysin, is a novel member of the thiol-activated cytolysin family. *J. Bacteriol.* **179**, 6100–6106.

Billington, S. J., Jost, B. H., Songer, J. G. (2000): Thiol-activated cytolysins: structure, function and role in pathogenesis. *FEMS Microbiol. Lett.* **182**, 197–205.

Brook, I., Frazier, E. H. (1994): Microaerophilic streptococci as a significant pathogen: A twelve-year review. *J. Med.* **25**, 129–144.

Brunger, A. T., Kuriyan, J., Karplus, M. (1987): Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460.

Crestfield, A. M., Moore,S., Stein, W. H. (1963): The preparation and enzymatic hydrolysis

of reduced and S-carboxymethylated proteins. *J. Biol. Chem.* **238**, 622–627.

Derewenda, U., Derewenda, Z., Dodson, G. G., Hubbard, R. E., Korber, F. (1989): Molecular structure of insulin: The insulin monomer and its assembly. *Br. Med. Bull.* **45**, 4–18.

Freire, E. (2008): Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov. Today* **13**, 869–874.

Funk, P. G., Staats, J. J., Howe, M., Nagaraja, T. G., Chengappa, M. M. (1996): Identification and partial characterization of an *Actinomyces pyogenes* hemolysin. *Vet. Microbiol.* **50**, 129–142.

Gelber, S. E., Aguilar, J. L., Lewis, K. L. T., Ratner, A. J. (2008): Functional and phylogenetic characterization of vaginolysin, the human-specific cytolysin from *Gardnerella vaginalis*. *J. Bacteriol.* **190**, 3896–3903.

Gidding, K. S., Zhao, Ji, Sims, R. J., Tweten, R. K. (2004): Human CD59 is a receptor for the cholesterol-dependent cytolysin intermedilysin. *Nat. Struct. Mol. Biol.* **11**,1173–1178.

Inaba, T., Tahara, S., Nishikawa, N., Kashiwagi, H., Sato, F. (2005): All-electron density functional calculation on insulin with quasi-canonical localized orbitals. *J. Comp. Chem.* **26**, 987–993.

Inaba, T., Tsunekawa, N., Hirano, T., Yoshihiro, T., Kashiwagi, H., Sato, F. (2007): Density functional calculation of the electronic structure on insulin hexamer. *Chem. Phys. Lett.* **434**, 331–335.

Inoue, S., Ohkura, K., Ikeda, K., Hayashi, K. (1987): Amino acid sequence of a cytotoxin-like basic protein with low cytotoxic activity from the venom of the Thailand cobra *Naja naja siamensis*. *FEBS Lett.* **218**, 17–21.

Inoue, S., Okumura, K., Tsujino, M., Ohkura, K., Ikeda, K., Takechi, M., Tanaka, Y., Hayashi, K. (1989): Amino acid sequences of cytotoxin-like basic proteins derived from cobra venoms. *FEBS Lett.* **257**, 319–323.

Jensen, A., Hoshino, T., Kilian, M. (2013): Taxonomy of the Anginosus group of the genus *Streptococcus* and description of *Streptococcus anginosus* subsp. *whileyi* subsp. nov. and *Streptococcus constellatus* subsp. *viborgensis* subsp. nov. *Int. J. Syst. Evol. Microbiol.* **63**, 2506–2519.

Johnson, S., Brooks, N. J., Smith, R. A. G., Lea, S. M., Bubeck, D. (2013): Structural Basis for Recognition of the Pore-Forming Toxin Intermedilysin by Human Complement Receptor CD59. *Cell Rep.* **3**, 1369-1377.

Jones, S., Preiter, K., Portnoy, D. (1996): Conversion of an extracellular cytolysin into a phagosome-specific lysin which supports the growth of an intracellular pathogen. *Mol. Microbiol.* **21**, 1219–1225.

Kaneda, N., Sasaki, T., Hayashi, K. (1977) : Primary structures of cardiotoxin analogues II and IV from the venom of *Naja naja atra*. *Biochim. Biophys. Acta* **491**, 53–66.

Klebe, G. (2015): Applying thermodynamic profiling in lead finding and optimization. *Nat. Rev. Drug Discov.* **14**, 95–110.

Korchev, Y. E., Bashford, C. L., Pederzolli, C., Pasternak, C. A., Morgan, P. J., Andrew, P. W., Mitchell, T. J. (1998): A conserved tryptophan in pneumolysin is a determinant of the characteristics of channels formed by pneumolysin in cells and planar lipid bilayers. *Biochem. J.* **329**, 571–577.

Lambardo, F., Blanke, J. F., Curatolo, W. J. (1996): Computation of brain-blood partitioning of organic solutes via free energy calculations. *J. Med. Chem.* **39**, 4750–4755.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993): PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.

McKenzie, T. J., Lillegard, J. B., Grotz, T. E., Moir, C. R., Ishitani, M. B. (2010): Pyogenic liver abscess secondary to *Streptococcus anginosus* in an adolescent. *J. Pediatr. Surg.* **45**, E15–E17.

Murray, H. W., Gross, K. C., Masur, H., Roberts, R. B. (1978): Serious infections caused by Streptococcus milleri. *Amer. J. Med.* **64**, 759–764.

Nagamune, H., Ohnishi, C., Katsuura, A., Fushitani, K., Whiley, R. A., Tsuji, A., Matsuda, Y. (1996): Intermedilysin, a novel cytotoxin specific for human cells, secreted by Streptococcus intermedius UNS46 isolated from a human liver abscess. *Infect. Immun.* **64**, 3093–3100.

Nagamune, H., Whiley, R. A., Goto, T., Inai, Y., Maeda, T., Hardie, J.M., Kourai, H. (2000): Distribution of the intermedilysin gene among the anginosus group streptococci and correlation between intermedilysin production and deep-seated infection with Streptococcus intermedius. *J. Clin. Microbiol.* **38**, 220–226.

Nagamune, H., Ohkura, K., Sukeno, A., Cowan, G., Mitchell, T. J., Ito, W., Ohnishi, O., Hattori, K., Yamato, M., Hirota, K., Miyake, Y., Maeda, T., Kourai, H. (2004): The human-specific action of intermedilysin, a homolog of streptolysin O, is dictated by domain 4 of the protein. *Microbiol. Immunol.* **48**, 677–692.

Ohkuni, H., Nagamune, H., Ozaki, N., Tabata, A., Todome, Y., Watanabe, Y., Takahashi, H., Ohkura, K., Kourai, H., Ohtsuka, H, Fischetti, V. A., Zabriskie, J. B. (2012): Characterization of recombinant *Streptococcus mitis*-derived human platelet aggregation factor. *Acta Pathol. Microbiol. Immunol. Scand.* **120**, 56–71.

Ohkura, K., Inoue, S., Ikeda, K., Hayashi, K. (1988): Amino-acid sequences of four cytotoxins (cytotoxins I, II, III and IV) purified from the venom of the Thailand cobra, *Naja naja siamensis*. *Biochim. Biophys. Acta* **954**, 148–153.

Ohkura, K., Hori, H. (1999): Analysis of structure-permeability correlation of nitrophenol analogues in newborn rat abdominal skin using semiempirical molecular orbital calculation. *Bioorg. Med. Chem.* **7**, 309–314.

Ohkura, K., Sukeno, A., Yamamoto, K., Nagamune, H., Maeda, T., Kourai, H. (2003): Analysis of structural features of bis-quaternary ammonium antimicrobial agents 4,4'-(α,ω-polymethylenedithio) bis (1-alkylpyridinium iodide)s using computational simulation. *Bioorg. Med. Chem.* **11**, 5035–5043.

Ohkura, K., Nagamune, H., Kourai, H. (2004): Structural analysis of human specific cytolysin intermedilysin aiming application to cancer immunotherapy. *Anticancer Res.* **24**, 3343–3354.

Ohkura, K., Sukeno, A., Nagamune, H., Kourai, H. (2005): Bridge-linked bis-quaternary ammonium anti-microbial agents: Relationship between cytotoxicity and anti-bacterial activity of 5,5'-[2,2'-(tetramethylenedicarbonyldioxy)-diethyl] bis (3-alkyl-4-methylthiazonium iodide)s. *Bioorg. Med. Chem.* **13**, 2579–2587.

Ohkura, K., Hori, H., Nagamune, H. (2006–1): Molecular dynamics of human-specific cytolysin: analysis of membrane binding motif for therapeutic application. *Anticancer Res.* **26**, 4055–4062.

Ohkura, K., Nagasawa, H., Uto, Y., Okamura, N., Murakami, A., Hori, H. (2006–2): The role of Gc protein oligosaccharide structure: a Risk factor for COPD. *Anticancer Res.* **26**, 4073–4078.

Ohkura, K., Hori, H., Shinohara, Y. (2009): Role of C-terminal region of yeast ADP/ATP carrier 2 protein: dynamics of flexible C-terminal arm. *Anticancer Res.* **29**, 4897–4900.

Ohkura, K., Yamamoto, A., Tabata, A., Shinohara, Y., Nagamune, H., Hori, H. (2012): Molecular profiles of cholesterol-dependent cytolysin family-derived 11mer regions. *Anticancer Res.* **32**, 2343–2346.

Olofsson, A., Herbert, H., Thelestam, M. (1993): The projection structure of perfringolysin-O (*Clostridium perfringens* theta-toxin). *FEBS Lett.* **319**, 125–127.

Polekhina, G., Giddings, K. S., Tweten, R. K., Parker, M. W. (2005): Insights into the action of the superfamily of cholesterol-dependent cytolysins from studies of intermedilysin. *Pro. Natl. Acad. Sci. USA* **102**, 600–605.

Ruoff, K. L. (1988−1): *Streptococcus anginosus* ("*Streptococcus milleri*''): The unrecognized pathogen. *Clin. Microbiol. Rev.* **1**, 102−108.

Ruoff, K. L. (1988−2): *Streptococcus anginosus* ("*Streptococcus milleri*"). *Clin. Microbiol. Newslett.* **10**, 65−68.

Sasaki, M., Yamaura, C., Ohara-Nemoto, Y., Tajika, S., Kodama, Y., Ohya, T., Harada, R., Kimura, S. (2005): *Streptococcus anginosus* infection in oral cancer and its infection route. *Oral Dis.* **11**, 151−156.

Sekino-Suzuki, N., Nakamura, M., Mitsui, K., O-Iwashita, Y. (1996): Contribution of individual tryptophan residues to the structure and activity of θ-toxin (perfringolysin O), a cholesterol-binding cytolysin. *Eur. J. Biochem.* **241**, 941−947.

Shaun W. L., Douglas A. M., Andrew L. M., Mary E. H., David G., Aaron W., Pieter C. D., Victor N., Jack E. D. (2008): Discovery of a widely distributed toxin biosynthetic gene cluster. *Proc. Natl. Acad. Sci. USA* **105**, 5879−5884.

Smith, G. D., Swenson, D. C., Dodson, E. J., Dodson, G. G., Reynolds, C. D. (1984): Structural stability in the 4-zinc human insulin hexamer. *Proc. Natl. Acad. Sci. USA* **81**, 7093−7097.

Tabata, A., Nakano, K., Ohkura, K., Tomoyasu, T., Kikuchi, K., Whiley, R.A., Nagamune, H. (2013): Novel twin Streptolysin S-Like peptides encoded in the sag operon homologue of beta-hemolytic *Streptococcus anginosus*. *J. Bacteriol.* **195**, 1090−1099.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., Higgins, D. G. (1997): The

ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acid Res.* **24**, 4876–4882.

Victor N., Bernard B., Darrin J. B., Vivekananda D., Laurie K., Donald E. L, Joyce C. S. De A. (2000): Genetic locus for streptolysin S production by group A streptococcus: *Infect. Immun.* **68**, 4245–4254.

Vidal N., Delmas A. -S., David P., Cruaud C., Couloux A., Hedges, S. B. (2007): The phylogeny and classification of caenophidian snakes inferred from seven nuclear protein-coding genes: *C.R. Biol.* **330**, 182–187.

Weiner, P. K., Langridge, R., Blaney, J. M., Schaefer, R., Kollman, P. A. (1982): Electrostatic potential molecular surfaces. *Proc. Natl. Acad. Sci. USA* **79**, 3754–3758.

Whiley, R. A., Hardie, J. M. (1989): DNA-DNA hybridization studies and phenotypic characteristics of strains within the 'Streptococcus milleri group'. *J. Gen. Microbiol.* **135**, 2623–2633.

Whiley, R. A., Fraser, H., Hardie, J. M., Beighton, D. (1990): Phenotypic differentiation of Streptococcus intermedius, Streptococcus constellatus, and Streptococcus anginosus strains within the “Streptococcus milleri group”. *J. Clin. Microbiol.* **28**, 1497–1501.

Whiley, R. A., Beighton, D. (1991): Emended descriptions and recognition of Streptococcus constellatus, Streptococcus intermedius, and Streptococcus anginosus as distinct species. *Int. J. Syst. Bacteriol.* **41**, 1–5.

Whiley, R. A., Beighton, D., Winstanley, T. G., Fraser, H. Y., Hardie, J. M. (1992):

Streptococcus intermedius, *Streptococcus constellatus*, and *Streptococcus anginosus* (the *Streptococcus* milleri group): association with different body sites and clinical infections. *J. Clin. Microbiol.* **30**, 243–244.

Whiley, R. A., Hall, L. M. C., Hardie, J. M., Beighton, D. (1999): A study of small-colony, β-hemolytic, Lancefield group C streptococci within the anginosus group: description of Streptococcus constellatus subsp. pharyngis. nov., associated with the human throat and pharyngitis. *Int. J. Syst. Bacteriol.* **49**, 1443–1449.

Wickham, S. E., Hotze, E. M., Farrand, A. J., Polekhina, G., Nero, T. L., Tomlinson, S., Parker, M. W., Tweten, R. K. (2011): Mapping the intermedilysin-human CD59 receptor interface reveals a deep correspondence with the binding site on CD59 for complement binding proteins C8α and C9. *J. Biol. Chem.* **286**, 20952–20962.

List of Publications

Flexible structure of cytochrome P450: Promiscuity of ligand binding in the CYP3A4 heme pocket.

Kazuto Ohkura, Yuki Kawaguchi, Yasuko Watanabe, Yasuhiro Masubuchi, Yasuo Shinohara, Hitoshi Hori.

Anticancer Research 2009 Mar; **29**(3): 935–942.

Molecular profiles of cholesterol-dependent cytolysin family-derived 11mer regions.

Kazuto Ohkura, Yuki Kawaguchi, Atsushi Tabata, Atsushi Yamamoto, Yasuo Shinohara, Hideaki Nagamune, Hitoshi Hori.

Anticancer Research 2012 Jun; **32**(6): 2343–2346.

Profiles of ILY, VLY and Sm-hPAF interaction with human CD59.

Yuki Kawaguchi, Atsushi Tabata, Hideaki Nagamune, Kazuto Ohkura.

Anticancer Research 2013 Jul; **33**(7): 2901–2904.

Molecular analysis of *Streptococcus anginosus*-derived SagA peptides.

Yuki Kawaguchi, Atsushi Tabata, Hideaki Nagamune, Kazuto Ohkura.

Anticancer Research 2014 Aug; **34**(8): 4627–4631.

Cytolytic activity and molecular feature of cardiotoxin and cardiotoxin-like basic protein: The electrostatic potential field is an important factor for cell lytic activity.

Yuki Kawaguchi, Yohei Tatematsu, Atsushi Tabata, Hideaki Nagamune, Kazuto Ohkura.

Anticancer Research 2015 Aug; **35**(8): 4515–4519.