

Academic Dissertation for Doctor of Philosophy

**Study on Family-S46 Dipeptidyl Peptidases
for the Development of Novel Antimicrobics**

(新規抗菌薬開発を指向した Family-S46 ジペプチジルペプチダーゼに関する研究)

Akihiro Nakamura

(中村 彰宏)

Department of Science of Technology Innovation,
Nagaoka University of Technology, Nagaoka, Japan

March 2022

PREFACE

This is an academic dissertation for a Doctor of Philosophy of Engineering. This research was conducted under the supervision of Professor Wataru Ogasawara at the *Hakko* Science Laboratory, Nagaoka University of Technology. I would like to take this opportunity to express my sincere gratitude to Professor Ogasawara for his continuous support, constructive advice, in-depth discussion, and encouragement. He taught me how to be a good researcher and I am very glad that I chose his laboratory.

I am deeply grateful to Associate Professor Yosuke Shida and Specially Appointed Assistant Professor Yoshiyuki Suzuki for their support, guidance, and concrete suggestions. Dr. Suzuki taught me the basics of experimental techniques.

I want to express my most profound appreciation to Professor Nobutada Tanaka of Kitasato University as well as Associate Professor Yasumitsu Sakamoto and Assistant Professor Saori Roppongi of Iwate Medical University for their guidance and cooperation during the crystal structure analysis. Without Professor Tanaka's and Associate Professor Sakamoto's guidance and persistent help, this dissertation would not have been possible.

I am in debt to Professor Yuko Tsuda of Kobe Gakuin University and Part-time Lecturer Koushi Hidaka of Kobe University for their guidance and cooperation in the synthesis of peptide compounds. Without Dr. Hidaka's cooperation, the experiments outlined in Chapter 3 would not have been possible.

I want to thank Dr. Naohiro Noda of National Institute of Advanced Industrial Science and Technology for his cooperation in the emulsion experiment in conducting this research.

I am grateful to Professor Takashi Yamaguchi (Nagaoka University of Technology), Professor Hiroaki Takaku (Niigata University of Pharmacy and Applied Sciences), and Associate Professor Daisuke Kasai (Nagaoka University of Technology) for their helpful comments and review of the dissertation manuscript.

Special thanks also to the staff of the *Hakko* Science Laboratory, and my seniors, peers, and juniors, for their cooperation in this research. In particular, I would like to thank Mr. Honma and Mr. Tanaka for their support during the emulsion experiment.

Finally, I would like to thank my family for all their understanding and support so far.

Akihiro Hokamura March 2022

Table of Contents

Chapter 1	General introduction	4
1.1.	Background	4
1.2.	Non-fermenting Gram-negative bacteria	6
1.2.1.	Overview of non-fermenting Gram-negative bacteria	6
1.2.2.	Mechanism of peptide utilization in non-fermenting Gram-negative bacteria	6
1.2.3.	Representative pathogenic NFGNBs with family-S46 peptidase, a target of this study	8
1.2.3.1.	<i>Stenotrophomonas maltophilia</i>	8
1.2.3.2.	<i>Porphyromonas gingivalis</i>	8
1.3.	Clan PA family-S46 serine peptidase as a novel target molecule for antimicrobials	9
1.3.1.	Overview of clan PA family-S46 peptidases	9
1.3.2.	Structure of family-S46 peptidase	11
1.3.3.	Mechanisms of catalysis and N-terminus recognition in family-S46 peptidases	12
1.3.4.	Recognition mechanisms of the P1 side chain in family-S46 peptidases	14
1.4.	Exploratory technology for antimicrobials	15
1.4.1.	Overview of drug discovery	15
1.4.2.	Screening using computer-aided drug design (CADD)	16
1.4.3.	High throughput screening (HTS)	17
1.4.4.	Ultra-high-throughput screening using droplet microfluidics	18
1.5.	Purpose	21
Chapter 2	Elucidation of recognition mechanisms at the S2 subsite in family-S46 peptidases	22
2.1.	Introduction	22
2.2.	Materials and Methods	24
2.2.1.	Cloning	24
2.2.2.	Effect of b3-signal on heterologous expression of S46 peptidase in <i>E. coli</i>	24
2.2.3.	Overexpression and purification of SmDPP7 WT and mutants	25
2.2.4.	Overexpression and purification of SmDPP11	27
2.2.5.	Overexpression and purification of PgDPP7	29
2.2.6.	Overexpression and purification of PgDPP11	30
2.2.7.	Determination of substrate specificity toward dipeptidyl pNA	32
2.2.8.	Determination of kinetic parameters toward dipeptidyl MCA	32
2.2.9.	Inhibition assay using dipeptide library	33
2.2.10.	Crystallization of SmDPP7	33
2.2.11.	X-ray data collection	35
2.2.12.	Structure determination	36
2.2.13.	Isothermal Titration Calorimetry	37
2.3.	Results	38
2.3.1.	Determination of optimal pH	38
2.3.2.	P1 position preferences of SmDPP7	39
2.3.3.	P2 position preferences of SmDPP7	40
2.3.4.	Structure determination of SmDPP7	43
2.3.4.1.	The overall structure of SmDPP7	43
2.3.4.2.	Dipeptide complexes	48
2.3.5.	Site-directed mutagenesis studies on residues in the S2 subsite of SmDPP7	51
2.3.6.	Thermodynamic characterization of dipeptide bindings in SmDPP7	53
2.3.7.	P2 position preference of S46 peptidases	55
2.4.	Discussion	56

Chapter 3	Investigation of an ultra-high-throughput detection method for family-S46 peptidase.	62
3.1.	Introduction	62
3.2.	Materials and Methods	64
3.2.1.	Materials, bacterial strains, and medium	64
3.2.2.	Materials and general procedures for substrate synthesis	64
3.2.3.	Synthesis of coumarins and substrates	64
3.2.4.	Overexpression and purification of PmDAPBII	66
3.2.5.	Overexpression and purification of SmDPP7, SmDPP11, and PgDPP11	67
3.2.6.	WODL generation and sorting	67
3.2.7.	Retention time characterization in the WODL	67
3.2.8.	Image analysis	68
3.2.9.	Determination of kinetics parameter toward the dipeptidyl substrate	69
3.2.10.	Enzymatic reaction in WODLs	69
3.2.11.	WODL cultivation	69
3.2.12.	Preparation of 16s rDNA library	70
3.2.13.	Sequencing analysis	71
3.2.14.	Device fabrication and operation	72
3.2.15.	Graphical programs	72
3.3.	Results	73
3.3.1.	Examination of leakage of fluorescent substances from the WODL.	73
3.3.2.	Synthesis of fluorogenic substrates	76
3.3.3.	Evaluation of the detection of purified DPP activities in bulk	77
3.3.4.	Detection of purified DPP activities in the WODL	79
3.3.5.	Detection of DPP activities of bacterial cells in a WODL and screening by FADS	80
3.3.6.	Isolation of DPP-producing bacteria from the environment	86
3.3.7.	Comparative analysis of 16S rDNA	89
3.4.	Discussion	91
Chapter 4	General conclusion	95
Chapter 5	References	98
Chapter 6	Published paper	104

Chapter 1 General introduction

1.1. Background

The discovery of antimicrobials (antibiotics) produced by microorganisms has dramatically extended the average human life span, beginning with penicillin discovered by the British physician Alexander Fleming in the 1929,¹ followed by the discovery of the anti-tuberculosis agent streptomycin in 1943. In the first half of the 20th century, when medical technology was less advanced, antimicrobials were highly effective against infectious diseases, which were the leading cause of death. However, the medical infrastructure of antimicrobials for infectious diseases is collapsing due to the accelerated emergence of drug-resistant bacteria.²

Drug-resistant bacteria emerged in the 1940s,³ when antimicrobials began to be widespread. In 1993, drug-resistant strains of bacteria against carbapenems, which were said to be the last card, were confirmed.⁴ In fact, drug-resistant bacteria have been identified against almost all the current antimicrobial agents (including those not derived from microorganisms) mainly used in the treatment of infectious diseases. These antimicrobials have the following mechanisms of action: inhibition of nucleic acid synthesis, inhibition of cell wall synthesis, inhibition of cell membrane function, inhibition of nucleic acid synthesis, inhibition of protein synthesis, and inhibition of folate synthesis (Figure 1-1).⁵

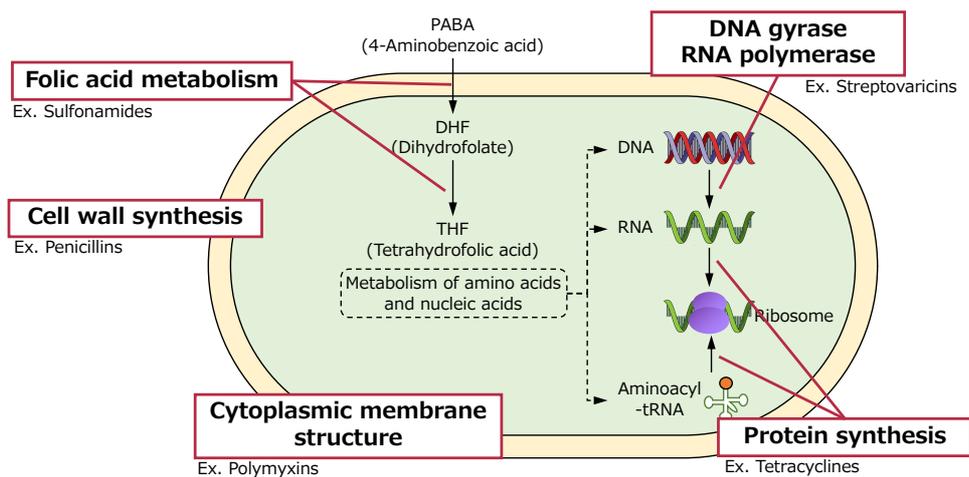


Figure 1-1 | Mechanisms of action of antimicrobials

The emergence of drug-resistant bacteria can be attributed to various factors, including the administration of broad-spectrum antimicrobial agents.⁶ Inappropriate or prolonged use of broad-spectrum antimicrobial agents enable the emergence of resistant bacteria.

Therefore, appropriate use of narrow-spectrum antimicrobial agents, which act only against the causative organism of the infection, is required whenever possible. In terms of antimicrobial development, the number of new antimicrobial approvals has been decreasing every year. In the U.S., for example, as many as 16 new antimicrobial agents were approved in 1983–1987, but the number has significantly decreased to zero in 2003–2007 (Figure 1-2).⁷⁻⁹ Although developments have increased in recent years, most target specific microorganisms such as *Mycobacterium tuberculosis* and *Clostridioides difficile*, and there is still a lack of potential treatment options for priority-resistant microorganisms, especially multidrug- and extensively drug-resistant Gram-negative pathogens. The World Health Organization (WHO) reported that only 6 out of 32 antimicrobial agents in clinical development

are classified as innovative medicines, and they warned that the shortage of antimicrobial agents would affect countries' health systems worldwide. Hence, it is desirable to develop narrow-spectrum antimicrobials with a different mechanism of action from existing antimicrobials.

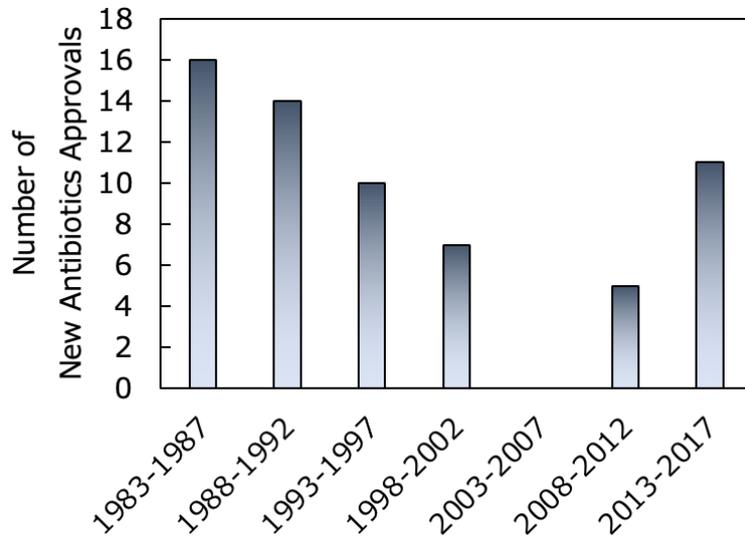


Figure 1-2 | Number of new antimicrobial approvals in the U.S.; modified from Spellberg, B. *et al.*, (2004)⁷ and Andrei, S. *et al.*, (2018).⁸

1.2. Non-fermenting Gram-negative bacteria

1.2.1. Overview of non-fermenting Gram-negative bacteria

Among bacterial pathogens, non-fermenting Gram-negative bacteria (NFGNB) are the main focus of the antimicrobial-resistance epidemic.¹⁰ NFGNB are unable to assimilate sugars, and they use proteins (peptides) for the energy needed for cell maintenance and growth.¹¹ NFGNB include opportunistic infectious organisms that cause infections in critically ill or immunocompromised patients due to other diseases; these organisms are ubiquitous in the soil and hydrosphere and are isolated in medical settings, such as respirators, catheters, and the skin of healthcare workers.^{12,13} Pathogenic NFGNB reported in the United States include *Pseudomonas aeruginosa*, *Acinetobacter baumannii*, *Stenotrophomonas* (formerly named *Xanthomonas*) *maltophilia*, and *Burkholderia cepacia*. *P. aeruginosa* and *B. cepacia* are pathogens that affect patients with cystic fibrosis,^{14,15} and *P. aeruginosa*, *S. maltophilia*, and *A. baumannii* are major causes of healthcare-associated infections in hospitalized and critically ill patients.^{13,16} In addition, *Porphyromonas gingivalis*, the causative agent of periodontal disease, is also classified as a pathogenic NFGNB.

NFGNB are prone to drug resistance because they possess various gene transduction mechanisms related to antibiotic resistance, such as overexpression of drug efflux pumps and possession of plasmids carrying antibiotic-resistance genes.¹⁷⁻¹⁹ Currently, strains resistant to all commonly used antimicrobials (penicillins, cephalosporins, aminoglycosides, tetracyclines, fluoroquinolones, trimethoprim-sulfamethoxazole, and carbapenems) have been identified. These NFGNB strains are considered dangerous in the medical field.¹⁰

1.2.2. Mechanism of peptide utilization in non-fermenting Gram-negative bacteria

The peptide uptake and degradation process is an essential mechanism for the growth of NFGNB, which use peptides rather than sugar as a nutrient source. The mechanism of peptide utilization will be reviewed using *P. gingivalis*, a relatively common NFGNB as an example (Figure 1-3).²⁰ Extracellular polypeptides (proteins) are degraded into oligopeptides by extracellular endopeptidases (lysine-gingipain protease, Kgp; and

arginine-gingipain protease, Rgp). Oligopeptides pass through the outer membrane and are degraded to dipeptides and tripeptides in the periplasmic space by various enzymes: dipeptidyl peptidase (DPP), prolyl tripeptidyl peptidase A (PtpA), and acyl peptidyl oligopeptidase (AOP). The dipeptides and tripeptides incorporated into the cell are finally hydrolyzed into amino acids by dipeptidases, aminopeptidases, and carboxyl peptidases in the cytoplasm and utilized as a nutrient source.

Since endocytic transporters preferentially take up dipeptides and tripeptides, the periplasmic exopeptidases—DPP4 (clan SC family-S9), DPP5 (clan SC family-S9), DPP7 (clan PA family-S46), DPP5 (clan SC family-S9), DPP7 (clan PA family-S46), and DPP11 (clan PA family-S46)—are thought to play important roles in both cell growth and pathogenicity. Nemoto *et al.* reported that growth of *dpp4-dpp5-dpp7-dpp11*-knockout *P. gingivalis* strain was significantly retarded.²¹ In addition, DPP7 and DPP11, which are family-S46 peptidases, are not found in humans but only in bacteria, thus making bacterial DPP7 and DPP11 promising target molecules for bacteriostatic antimicrobial agents with fewer side effects.

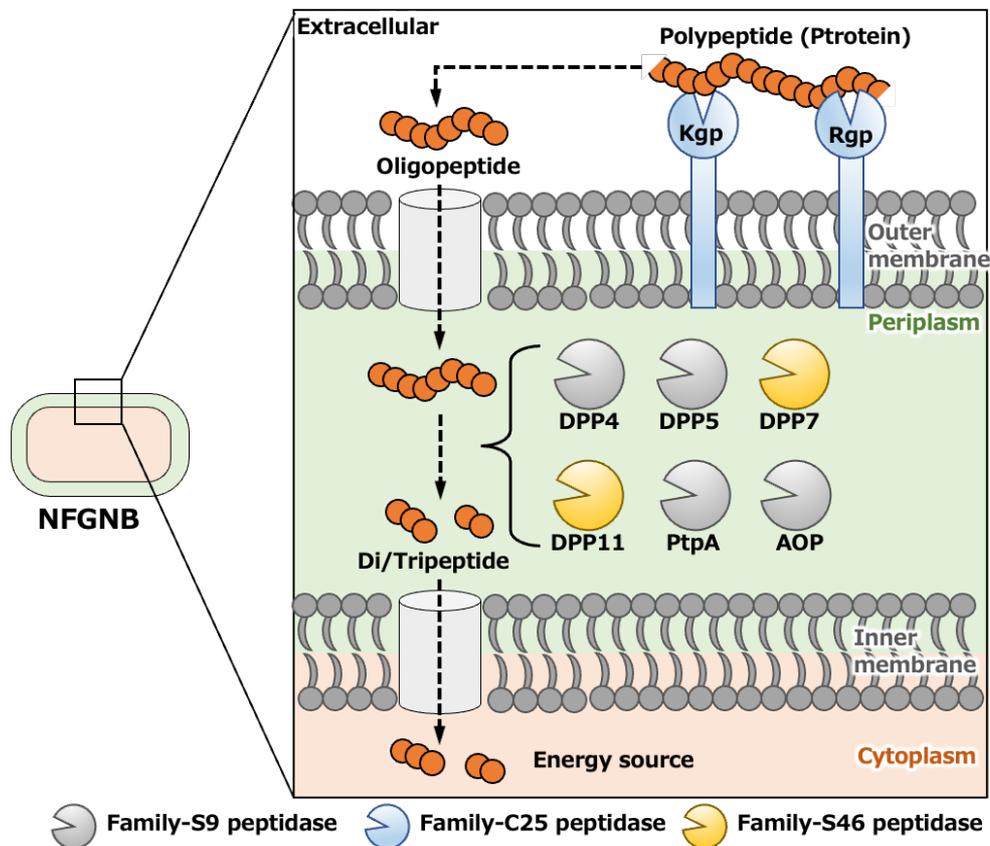


Figure 1-3 | Schematic diagram of degradation and uptake of extracellular peptides by non-fermenting Gram-negative bacteria; modified from Nemoto, T. K. & Ohara-Nemoto, *Jpn. Dent. Sci. Rev.* (2016).²⁰

1.2.3. Representative pathogenic NFGNBs with family-S46 peptidase, a target of this study

We introduce two NFGNBs known to possess S46 peptidase, be highly pathogenic, and infect many individuals. These two NFGNB were selected as pathogenic microorganisms to be addressed in this study.

1.2.3.1. *Stenotrophomonas maltophilia*

Stenotrophomonas maltophilia (formerly named *Pseudomonas maltophilia* or *Xanthomonas maltophilia*),²² a NFGNB, is a multiple-drug resistant bacterium responsible for opportunistic infections in immunocompromised patients.^{23–25} This bacterium exhibits antimicrobial resistance against the available carbapenem class of broad-spectrum antibiotics as it possesses multidrug-resistance pumps,^{17,18} plasmids harboring antibiotic resistance genes, and various gene transfer mechanisms involved in the acquisition of antimicrobial resistance.

The World Health Organization (WHO) recently classified *S. maltophilia* as one of the leading multidrug-resistant organisms (MDROs) in hospital settings.²⁶ A multitude of factors have been associated with acquiring *S. maltophilia* infections. These include underlying malignancy (especially hematologic malignancy), organ transplantation, human immunodeficiency virus (HIV) infection, cystic fibrosis, prolonged hospitalization, intensive care unit (ICU) admission, mechanical ventilation, indwelling catheters (vascular, urinary, and biliary), and corticosteroid or immunosuppressive therapy (treatments are concurrently given along with antibiotics administration).²⁷ Considering these risk factors, novel antimicrobial strategies are needed to combat *S. maltophilia* infections, and these antimicrobials must be highly specific to a particular group of bacteria to avoid the emergence of antimicrobial-resistant bacteria.

1.2.3.2. *Porphyromonas gingivalis*

Porphyromonas gingivalis is an anaerobic NFGNB and a major cause of chronic periodontitis.²⁸ Although there are more than 500 species in the oral cavity, *P. gingivalis*, *Treponema denticola*, and *Tannerella forsythia* have been called the “red complex”: pathogenic bacteria strongly associated with periodontal disease.²⁹ Clinical and

epidemiological studies have revealed an association between rheumatoid arthritis and periodontitis, and *P. gingivalis* causes the facilitation of collagen-induced arthritis.³⁰ In addition, it was reported that *P. gingivalis* is involved in Alzheimer's disease; its extracellular proteases have a negative effect on tau protein, which is needed for the normal function of nerve cells.³¹ It is noteworthy that *P. gingivalis*, an NFGNB collected from patients, is resistant to broad-spectrum antimicrobial agents such as tetracycline or erythromycin (macrolide).³² Thus, this bacterium has been implicated in periodontal disease and systemic diseases, so narrow-spectrum antibiotics with the new action of mechanisms against *P. gingivalis* are needed.

1.3. Clan PA family-S46 serine peptidase as a novel target molecule for antimicrobials

1.3.1. Overview of clan PA family-S46 peptidases

Family-S46 peptidases (hereafter referred to as S46 peptidases) are a group of exopeptidases with dipeptidyl peptidase (DPP) activity, which is the activity to release dipeptides from the N-terminus of oligopeptides. There are 14 families in clan PA serine peptidases (S), and all have endopeptidase activity except for S46 peptidase, which cleaves the interior of oligopeptides. However, in MEROPS,³³ family-S46 is the only clan PA (S) family with exo-activity (Table 1-1). To date, 1028 sequences have been registered in MEROPS family-S46. Of these, only 4 have been functionally identified: dipeptidyl aminopeptidase BII (PmDAP BII) from *Pseudoxansomonas mexicana* WO24, dipeptidyl peptidase 7 and 11 (PgDPP7 & PgDPP11) from the periodontal bacterium *Porphyromonas gingivalis*, and dipeptidyl peptidase 7 (BfDPP7) from *Bacteroides fragilis*. In addition, dipeptidyl peptidase 11 (SpDPP11) from the food poisoning bacterium *Shewanella putrefaciens*, dipeptidyl peptidase 11 (FpDPP11) from the bacterial coldwater disease bacterium *Flavobacterium psychrophilum*, and dipeptidyl peptidase 7 and 11 (PeDPP7 & PeDPP11) from the periodontal disease bacterium *Porphyromonas endodontalis* have been reported as known family-S46 peptidases not yet registered in MEROPS database (Figure 1-4).

S46 peptidases can be divided into two types, dipeptidyl peptidases 7 (DPP7) and dipeptidyl peptidase 11 (DPP11), according to the specificity at the P1 position (NH₂-P2-

P1-P1'-P2'-..., where the P1-P1' bond is the scissile bond).³⁴ DPP11s exhibit a strict specificity for acidic residues (Asp/Glu) at the P1 position, whereas DPP7s exhibit a broad specificity for aliphatic and aromatic residues at the P1 position.

Table 1-1 | List of clan PA serine peptidases

Family	Seq. ^a	Ident. ^b	Strc. ^c	Type of peptidase	Endo / Exo
S1	70904	702	141	Chymotrypsin A	Endo
S3	85	1	1	Togavirin	Endo
S6	1373	12	6	IgA1-specific serine peptidase	Endo
S7	360	2	3	Flavivirin	Endo
S29	236	2	2	Hepacivirin	Endo
S30	97	2	0	Potyvirus P1 peptidase	Endo
S31	32	1	1	Pestivirus NS3 polyprotein peptidase	Endo
S32	121	2	2	Equine arteritis virus serine peptidase	Endo
S39	72	2	2	Sobemovirus peptidase	Endo
S46	1028	4	2	Dipeptidyl peptidase 7	Exo
S55	338	1	0	SpoIVB peptidase	Endo
S64	40	1	0	Sey5 peptidase	Endo
S65	2	1	0	Picornain-like serine peptidase	Endo
S75	1	1	0	White bream virus serine peptidase	Endo

Extracted from the MEROPS database as of Sep. 28, 2021.³³ ^aThe number of all peptidase sequences registered in MEROPS, including peptidases that are expected to be classified into that peptidase family based on catalytic residues and sequence similarity. ^bThe number of peptidases among the registered peptidases whose function has been identified and for which a unique identifier has been issued by MEROPS. ^cThe number of peptidases registered in the Protein Data Bank (PDB).

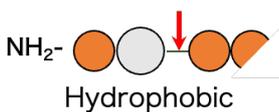
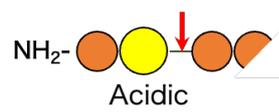
Enzyme	Dipeptidyl peptidase (DPP) 7	Dipeptidyl peptidase (DPP) 11
Substrate P1 specificity	 Hydrophobic	 Acidic
Type enzyme (Organism)	PmDAP BII (<i>Pseudoxanthomonas mexicana</i> WO24)	PgDPP11 (<i>Porphyromonas gingivalis</i>)
Other enzymes (Organism)	SmDPP7 (<i>Stenotrophomonas maltophilia</i>) PgDPP7 (<i>P. gingivalis</i>) PeDPP7 (<i>Porphyromonas endodontalis</i>) BfDPP7 (<i>Bacteroides fragilis</i>)	SmDPP11 (<i>S. maltophilia</i>) PeDPP11 (<i>P. endodontalis</i>) SpDPP11 (<i>Shewanella putrefaciens</i>) FpDPP11 (<i>Flavobacterium psychrophilum</i>)

Figure 1-4 | DPP7 and DPP11 belonging to S46 peptidase

1.3.2. Structure of family-S46 peptidase

S46 peptidases have two subunits forming a homodimer, with each subunit consisting of approximately 770 amino acid residues. The first three-dimensional structure of an S46 peptidase was determined for dipeptidyl aminopeptidase BII (DAP BII), a DPP7-type enzyme, from *P. mexicana* WO24 (Figure 1-5 A).³⁵ The study revealed that a protomer of DAP BII contains a peptidase domain, including a double β -barrel fold characteristic of the chymotrypsin superfamily and an unusual α -helical domain that regulates the exopeptidase activity of DAP BII (Figure 1-5 B-D). This structure is not present in any other enzyme except S46 peptidase. Subsequently, crystal structures of DPP11 from *P. gingivalis* (PgDPP11) were identified.^{36,37}

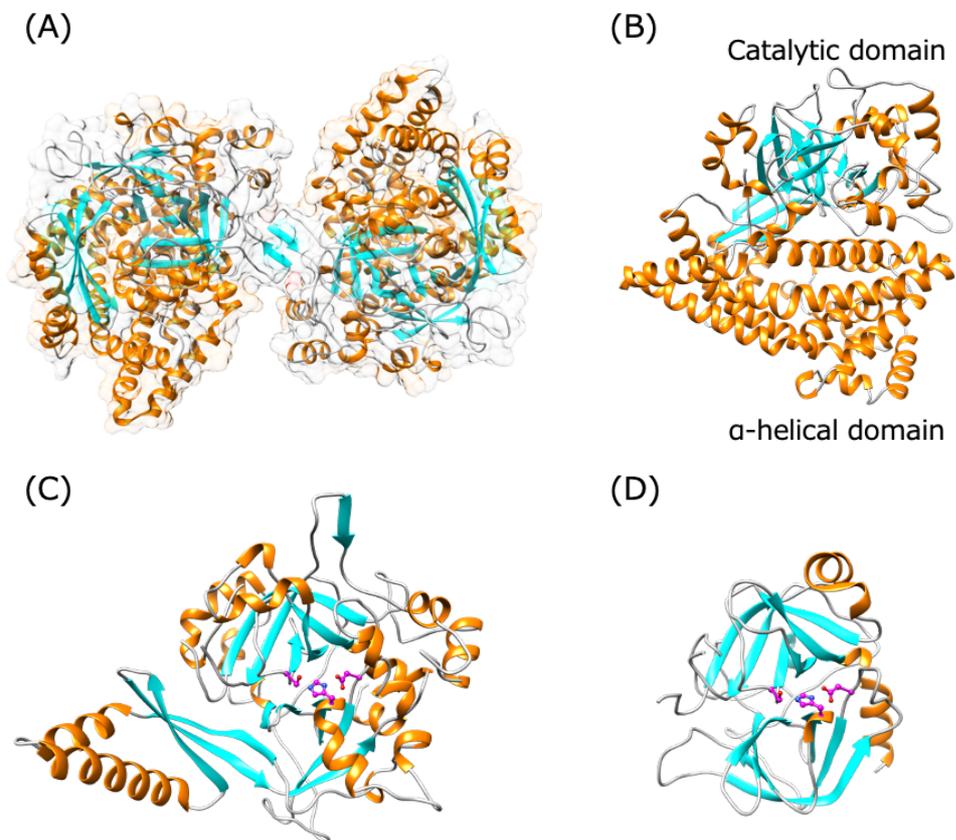


Figure 1-5 | Structure of PmDAPBII and clan PA peptidases (A) Overall structure of PmDAP BII (PDB, 3WOL). (B) Structure of protomers of PmDAP BII (PDB, 3WOL). (C) Structure of catalytic domains of PmDAP BII complexed with Val-Tyr dipeptide (PDB, 3WOL). Residues 25–276 and 574–720 of PmDAP BII are shown. Catalytic residues, His86, Asp224, and Ser657, are shown in magenta. (D) Structure of alpha-chymotrypsin from *Bos taurus* (PDB, 4CHA). Catalytic residues, His57, Asp102, and Ser195 are shown in magenta. The helix, sheet, and coil are shown in orange, cyan, and grey, respectively.

1.3.3. Mechanisms of catalysis and N-terminus recognition in family-S46 peptidases

Crystal structure analyses have unraveled the mechanisms of catalysis and N-terminus recognition of S46 peptidase. S46 peptidase is a serine protease with serine, histidine, and aspartate as catalytic residues. The reaction mechanism of S46 peptidase is represented in Figure 1-6.

The OH group of the catalytic serine residue nucleophilically attacks the carbonyl group of the substrate amide bond ($\text{NH}_2\text{-P2-P1-P1}'\text{-P2}'\text{-}\dots$, where the P1-P1' bond is the scissile bond). Normally, the OH group of serine residue is not nucleophilic enough to attack the amide bond. However, in the active center of the serine protease, the nitrogen atom of the imidazole side chain of the adjacent histidine residue attracts the hydroxy hydrogen, thereby increasing the nucleophilicity of the hydroxy oxygen atom. This hydroxyl activation of the histidine side chain's imidazole group is due to the hydrogen of the imidazole group by the aspartate side chain carboxylate anion adjacent to the histidine side chain residue. The substrate amide bond cleavage reaction by serine protease is initiated by the joint action of the aspartate, histidine, and serine side-chain functional groups. By nucleophilic attack of the serine hydroxyl group, the serine residue (catalytic residue) forms a covalent ester bond with the substrate and backbone NH groups of serine (catalytic residue) and glycine form hydrogen bond with the substrate (oxyanion hole), and serine protease forms an acyl-enzyme intermediate. Nucleophilic attack by water molecules activated by the same mechanism (as described above on the ester intermediate) leads to the hydrolysis of the ester, which completes the amide bond cleavage reaction of the substrate and regenerates the protease activity.

In addition, S46 peptidase regulates exo-activity by fixing the peptide to a specific conformation through hydrogen bonding with the amino terminus of the substrate. Asn residues play an important role in this mechanism (lower left at each scheme in Figure 1-6). A hydrogen bond is formed between the amide and carbonyl groups of the Asn side chain and the carbonyl and amino groups, respectively, of the main chain of the N-terminal P2 residue. Moreover, Asn residue and Asp residue is also involved in fixing N-terminal P2 residue. S46 peptidases undergo large-scale conformational changes, open to close conformation, associated with N-terminus recognition (Figure 1-7).

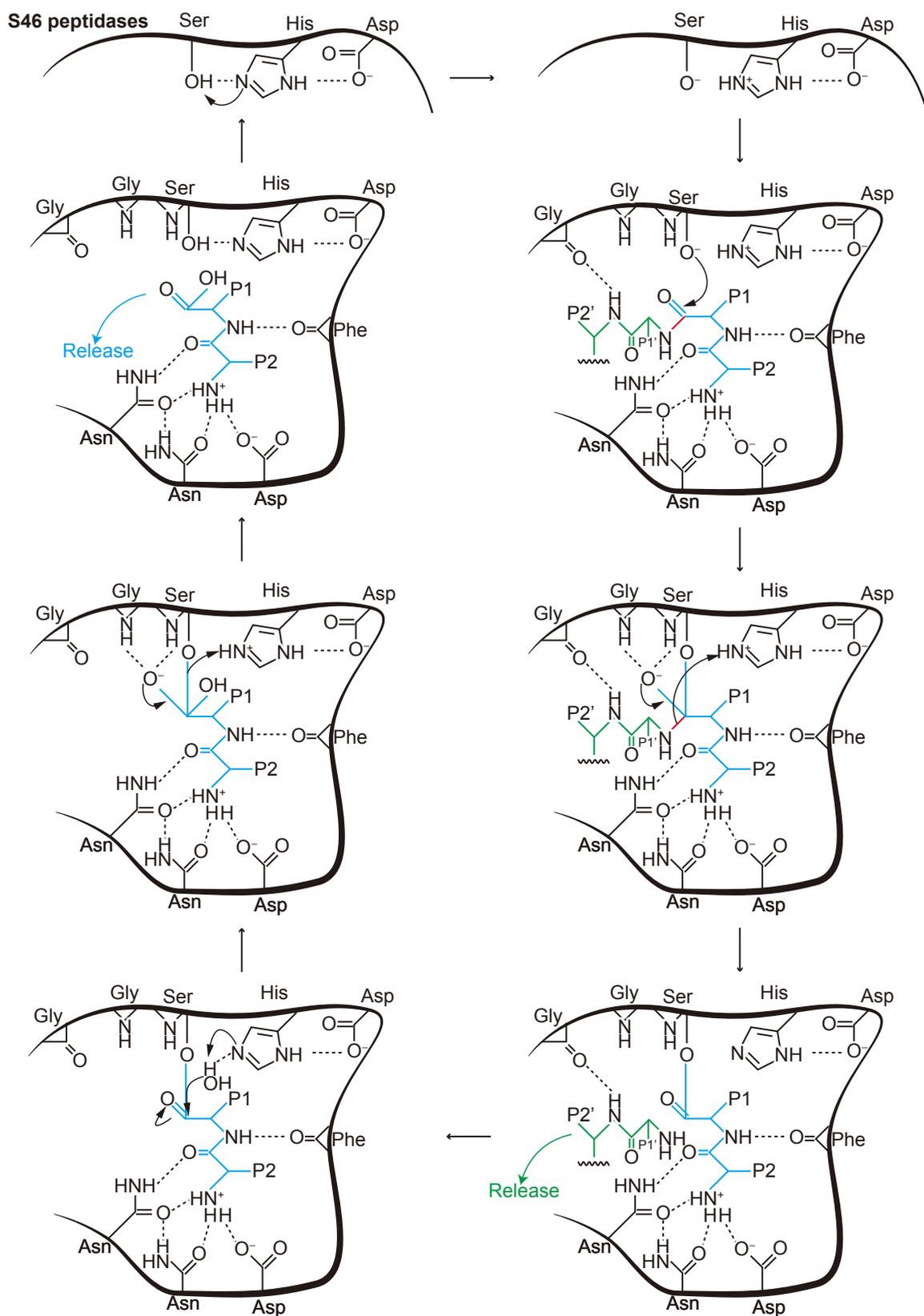


Figure 1-6 | Scheme of reaction mechanisms of S46 peptidase.³⁵ The scissile bond is shown in red. The main chain of P2-P1 and P1'-P2' amino acids are shown in blue and green, respectively.

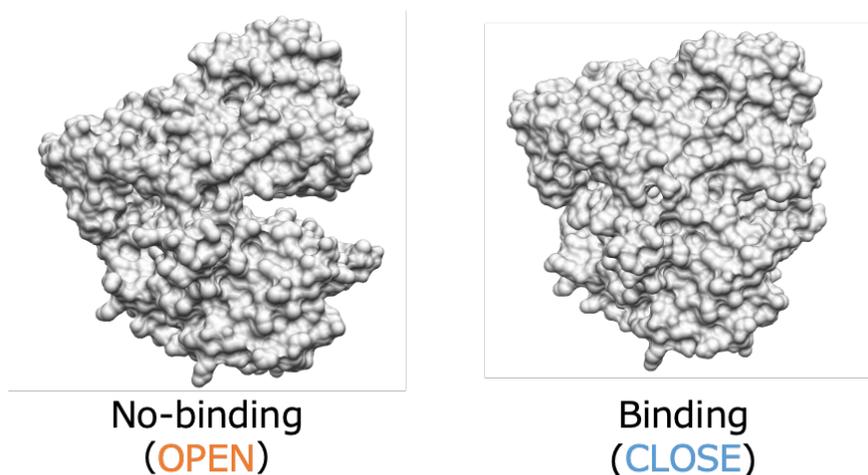


Figure 1-7 | Conformational change of PmDAPBII. The surface of PmDAP BII (PDB, 3WOL; closed conformation) and PmDAP BII (PDB, 3WOK; open conformation).

1.3.4. Recognition mechanisms of the P1 side chain in family-S46 peptidases

Crystal structure analyses and biochemical studies of PgDPP11^{34,38} have revealed that Arg673 in PgDPP11 is responsible for the strict Asp/Glu specificity of PgDPP11 at the P1 position of the substrate peptide. The P1 preference of S46 peptidases could be inferred according to the amino acid at position 673 of PgDPP11.³⁴ DPP7s have glycine, and DPP11s have arginine or serine at position 673 of PgDPP11 (Figure 1-8). As DPP7s have glycine at position 673, the S1 subsite of DPP7s can form a pocket deep enough to accommodate bulky hydrophobic residues. On the other hand, DPP11s exhibit specificity for Asp/Glu by the electrostatic interaction between the side chain of Arg673 and the carboxy group of the P1-side chain of Asp/Glu.³⁶

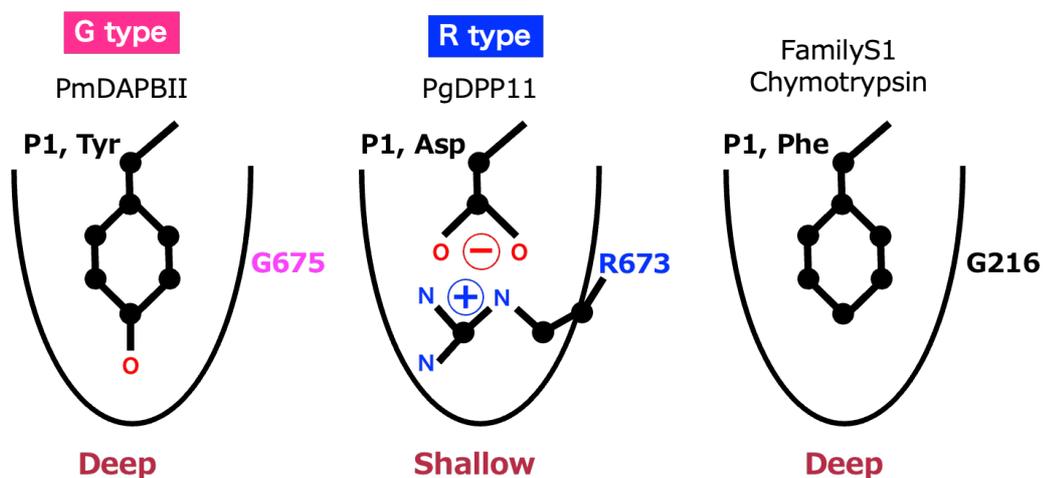


Figure 1-8 | Schematic diagrams of the S1 subsites of S46 peptidases and the S1 peptidase.

1.4. Exploratory technology for antimicrobials

1.4.1. Overview of drug discovery

Drug (antimicrobials) discovery is the process of discovering new drug candidates, which roughly follows the steps as Figure 1-9.³⁹ The first step in drug discovery is the selection of a target. The target is a naturally occurring cell or molecular structure involved in the pathology that the drug is intended to treat. Many target molecules for drug discovery are proteins. For instance, G-protein-coupled receptors (GPCRs) are well-known drug-targeting proteins. Approximately 34% of all drugs approved by the Food and Drug Administration (FDA) were targeted to 108 members of the GPCRs family.⁴⁰

The process of finding new drugs usually involves high-throughput screening (HTS), which finds compounds with potency against the target from an extensive library of chemicals. Another representative method in screening is computer-aided drug design (CADD).⁴¹ It is often used to identify new chemical moieties that may interact with target proteins or to screen active molecules in virtual space. HTS and CADD are typically used in the screening and evaluation stages. In this step, it is necessary to evaluate as many substances as possible. In other words, it is essential to have high efficiency. In addition, 974 small new chemical entities were developed between 1981 and 2006, of which 63% were naturally occurring or semi-synthetic derivatives of natural products.⁴² Hence, ideally, the chemical libraries should contain natural products such as organisms metabolite.⁴³ A series of lead compounds with target potency and selectivity, as well as desirable drug-like properties, can be optimized in terms of increasing activity on selected targets, thus reducing activity on irrelevant targets, and improving the ADME (absorption, distribution, metabolism, and excretion) properties of the molecule. Then, new drugs are born through non-clinical research, clinical trials, and certification.

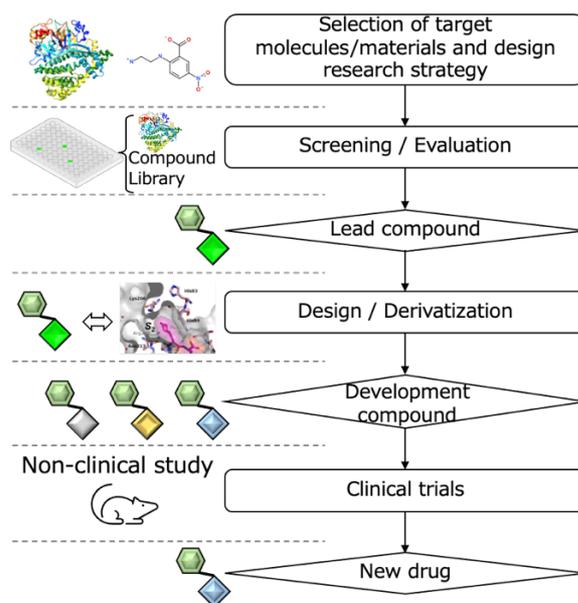


Figure 1-9 | Schematic overview of drug discovery.

1.4.2. Screening using computer-aided drug design (CADD)

In general, CADD approaches can be categorized into structure-based and ligand-based methods (Figure 1-10).⁴⁴ Structure-based methods (structure-based drug design [SBDD]) consist of using the structure of the target (enzyme/receptor) to generate or screen for potential ligands, followed by synthesis, biological testing, and optimization.^{45–47} Therefore, selecting the drug target and obtaining the structure of the target are essential first steps in SBDD. X-ray crystallography, cryo-electron microscopy, nuclear magnetic resonance, homology modeling, and molecular dynamic have been used to generate protein structures.^{46,48} On the other hand, a ligand-based approach (ligand-based drug design [LBDD]) consists of constructing the theoretical predictive models compared with diverse structures and known potencies. Suppose structural information on the target protein is lacking. In that case, the selection of candidate compounds is based on the ligand's structural and physicochemical features that are known potencies against the target protein, assuming that structurally similar compounds exhibit similar biological responses and interactions against the target protein.^{49,50}

Taken together, CADD conducted virtually using a computer is a powerful method for discovering potential ligands because of its high efficiency and low cost. It should be noted that biochemical assays are a necessary step for screening and evaluating potential ligands

in CADD. In other words, biochemical assay methods with high throughput would accelerate the CADD approach.

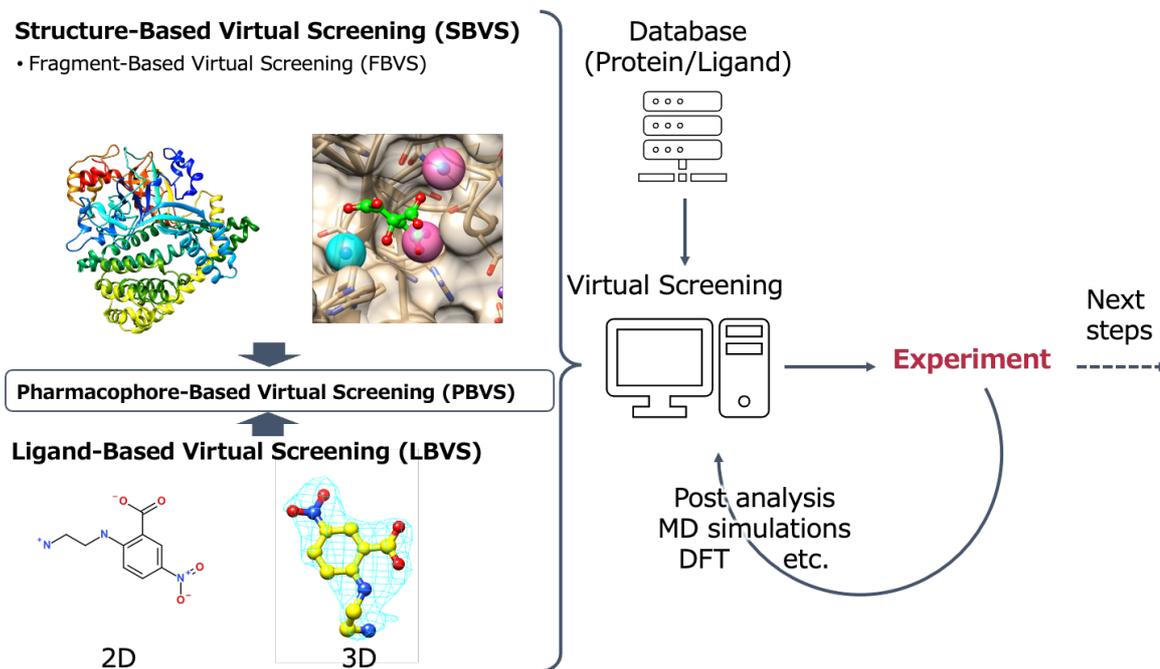


Figure 1-10 || Schematic overview of CADD.

1.4.3. High throughput screening (HTS)

High throughput screening (HTS) consists of screening an extensive chemical library for activity/function against targets utilizing automation, miniaturized assays, and large-scale data analysis. HTS ranges from 10,000 to 100,000 compounds to be tested per day, whereas ultra-high-throughput (uHTS) is defined as the number of screenings with more than 100,000 data points generated per day.⁵¹ Biological targets and chemicals to be screened (e.g., small molecules, siRNA, and miRNA) are introduced into microwell plates and evaluated their activity and function (Figure 1-11).⁵² The most common detection method is plate reader screening, which measures the fluorescence or absorbance of the entire microwell. In addition to this method, the high content screen method uses the morphological information of the target as an indicator,⁵³ and the kinetic imaging screen, which evaluates the changes in morphological information over time.⁵⁴ However, the number of candidate target molecules and chemical substances to be screened has increased with the expansion of gene resources and chemical libraries due to the rapid development of science and technology. Therefore, ultra-high-throughput screening

technologies that can measure target activity/functions more efficiently and directly are required.

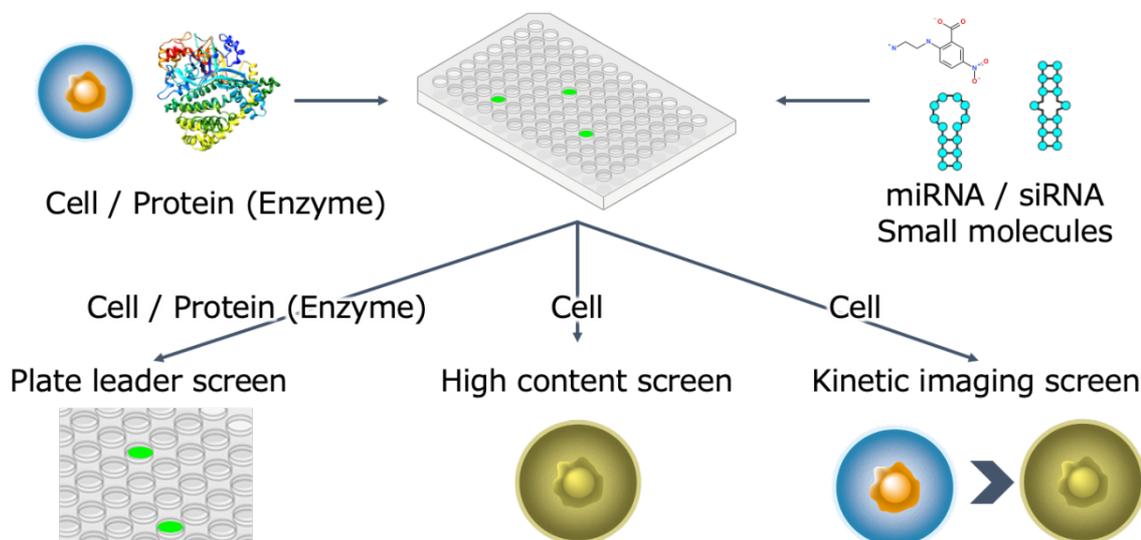


Figure 1-11 | Schematic overview of HTS

1.4.4. Ultra-high-throughput screening using droplet microfluidics

Droplet-based microfluidic systems using water-in-oil droplet (WODL) are a powerful tool for biological assays with ultra-high throughput, and the number of samples processed per day exceeds 1 million (Figure 1-12 A).^{55,56} Along with ultra-high throughput, compartmentalization and miniaturization are features specific to this technique,⁵⁷ which facilitate the screening for activity and function.⁵⁸ Stochastic compartmentalization effectively isolates a single cell from a mixed suspension of cell culture (drug target). Miniaturization also reduces the required amounts of medium, ligand, and substrate, which reduces costs. In addition, unlike bulk emulsions, this technology allows for WODL fusion, splitting, collection, and alignment control.^{59–62} In general, biological assays with droplet-based microfluidic systems using WODLs are carried out using the procedure shown in Figure 1-12 B.

Although many methods have been developed to generate WODLs, the flow-focusing method is one of the major methods for culturing cells and conducting biological assays in the WODL because it can produce small-sized droplets with high throughput (Figure 1-12 B [i]).⁶³ For the aqueous phase, a mixture of the medium, where the cells are suspended

and the indicator or substrate is dissolved, is used. Fluorinated oil or mineral oil is often used as the oil phase, and fluorinated oil is commonly used for cell culture using WODL because of its excellent gas permeability.⁶⁴ Since fluorinated oils are the mainstream of WODL cultivation, perfluorinated polyether surfactants, in which the carbon skeleton is replaced with fluorine, have been developed.⁶⁵ Surfactant-containing emulsions can lower the surface tension of the droplet interface and maintain robustness so that droplets do not merge. Moreover, electrowetting, which directs electrical control of the surface tension and increases surface area, enables droplets to merge and split.^{66–69} Another application of electrowetting is “pico-injection,” in which droplets with an electrical potential difference are injected with a small amount of aqueous solution by touching.^{70,71} The emulsion generated by the above operations was placed into a container and incubated (Figure 1-12 B [ii]). Droplet sorting involves detecting and collecting each droplet signal on a microfluidic device (Figure 1-12 B [iii]). Fluorescence is the most common signal, but absorbance, scattered light, and Raman spectroscopy are also used.^{72–74}

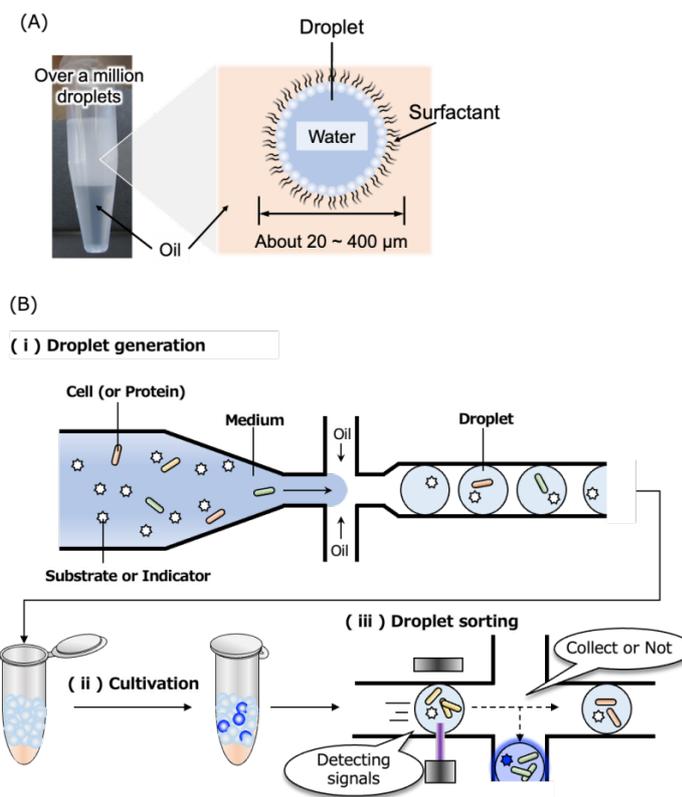


Figure 1-12 | Schematic of biological assays using droplet-based microfluidic systems. (A) Schematic of water-in-oil emulsion (droplet). (B) Schematic overview of the method of biological assays using droplet-based microfluidic systems.

Recently, due to its ultra-high-throughput, droplet-based microfluidic systems using WODL have been widely applied to biochemical assays and screening. To date, various enzyme activities have been detected and screened, such as amylase, glucosidase, esterase, lipase, cellulase, and laccase.^{75–81} Also, this technology is being applied in drug discovery.⁸² The compartment of a pico- to nanoliter size droplet with surfactant-containing oil phase, which is a feature of WODL, can confine proteins in its limited space. For instance, this feature is useful for screening antibodies involved in treating various diseases such as infectious diseases, cancer malignancies, and inflammations.^{83–86} Because of the very low fusion efficiency of the antibody-producing B-cells and myeloma cells, the efficiency of hybridoma cell screening is extremely low.⁸⁷ The use of WODL and a probe that captures secreted antibodies enables droplet sorting and cell enrichment of antibody-producing single mouse hybridoma cells at about 200 cells/sec.⁸⁸ Thus, droplet-based microfluidic systems are suitable for screening cells, microorganisms, and phages (with different genes) that produce screening targets such as enzymes, proteins, and antibodies. However, it is not easy to generate chemically distinct droplets. The previous report is a way to parallelize the inlet of the water phase. Brouzes *et al.* created eight different drug libraries and monitored their toxicity to human cells in WODL.⁸⁹ Thus, droplet-based microfluidic systems have recently been studied for their application in drug screening.

1.5. Purpose

Family-S46 peptidase is an important enzyme for the growth of pathogenic nonfermenting Gram-negative bacilli such as *S. maltophilia* and *P. gingivalis*. Indeed, Nemoto *et al.* reported that *dpp7* and *dpp11*-knockout strain growth was significantly retarded.²¹ In addition, since S46 peptidases are not found in humans, antimicrobial agents targeting S46 peptidases are promising as antimicrobial agents with slight side effects on humans and with a different mechanism of action. Thus, the final goal of this study is to find antimicrobials that target S46 peptidase.

However, strict differences in the properties of substrate P1 of S46 peptidases, DPP7 and DPP11, have hindered the development of antimicrobials. Previous studies have suggested the involvement of the position 673 amino acid residue (Gly/Arg/Ser) in PgDPP11 as the substrate P1 recognition residue.³⁴ On the other hand, although the specificity of the P2 amino acid has been reported, a detailed analysis combined with structural analysis has not been performed.⁹⁰

In our analysis, the amino acid residues (comprising the S2 subsite of S46 peptidase) are highly conserved, suggesting that substrate P2 specificity is conserved in DPP7 and DPP11. Considering that S46 peptidase does not recognize the side chain of amino acid on the substrate prime side, the analysis of the S2 subsite is essential for the development of a universal inhibitor for S46 peptidase. In this study, we aimed to elucidate the mechanism of substrate P2 recognition by analyzing mutant enzymes, enzymology, and structural biology of S46 peptidases (SmDPP7, SmDPP11, PgDPP7, and PgDPP11) from pathogenic non-fermenting Gram-negative bacilli.

In addition, we also aimed to establish a highly efficient method to detect the activity of S46 peptidase for the development of antimicrobials targeting S46 peptidase. The current mainstream is high-throughput screening using CADD and microplate readers, but further improvement of efficiency is desired while directly detecting the activity of target molecules. Therefore, we focused on droplet-based microfluidic systems using WODL, which can analyze more than 100 samples per second. However, the substrates and indicators that can be used are limited due to the unique environment of an oil phase surrounding the 1 nL droplets. This study develops new substrates for DPP7 and DPP11 that can be used in WODL and establish an ultra-high throughput detection method for S46 peptidase activity.

Chapter 2 Elucidation of recognition mechanisms at the S2 subsite in family-S46 peptidases

2.1. Introduction

Since S46 peptidases are important enzymes for the growth of bacteria,^{91,92} these peptidases are anticipated as a novel molecular target of antibiotics.⁹³ However, a clear difference in the specificity at the P1 position of the substrate between the two types of S46 peptidases has been an obstacle when designing a universal inhibitor. S46 peptidases can be divided into two types, dipeptidyl peptidases 7 (DPP7) and dipeptidyl peptidase 11 (DPP11) according to the specificity at the P1 position (NH₂-P2-P1-P1'-P2'-..., where the P1-P1' bond is the scissile bond).³⁴ DPP11s exhibit a strict specificity for acidic residues (Asp/Glu) at the P1 position, whereas DPP7s exhibit a broad specificity for both aliphatic and aromatic residues at the P1 position. The overall structure, the molecular basis of the exopeptidase activity, the catalytic mechanism, and the strict P1-residue recognition mechanisms of S46 peptidases have been unraveled by the crystal structure analyses of DAP BII and PgDPP11.^{35,36,38} In addition, with regards to the recognition of the prime side (P1'-P2'-...) of the substrate peptide, it has been reported that specific interactions between the prime side subsites (S1'-S2'-...) and the side chain of the substrate (P1'-P2'-...) were not observed in the crystal structure of DAP BII complexed with an octapeptide.³⁵ Because antibiotics targeting S46 peptidases need to inhibit both DPP7 and DPP11, this specificity difference between DPP7 and DPP11 at the P1 position can be a potential obstacle when designing a universal inhibitor of S46 peptidases. However, little is known about the mechanism of P2-residue recognition. Biochemical studies of DPP7 from *P. gingivalis* (PgDPP7) and PgDPP11 suggested that the side chain of Phe664 in PgDPP7 and of Phe671 in PgDPP11 are involved in the recognition of the hydrophobic P2 residue of the substrate peptide. This hydrophobic specificity at the P2 position is conserved between DPP7 and DPP11.⁹⁰ The conservation of specificity at the P2 position has provided access to design the universal inhibitor of S46 peptidases, however, determinants for the P2-residue specificity/preference of S46 peptidases at the atomic level remain to be fully elucidated. Therefore, understanding the substrate recognition mechanism at the S2 subsite could provide more insight for the development of a universal inhibitor of S46 peptidases. It should be noted that in mammals, a peptidase designated DPP7 is one of the two members of the S28 peptidase family^{94,95} and is distinct from bacterial DPP7s that belong to the S46

peptidase family. The S28 peptidases contain two major domains: a peptidase domain, including an α/β -hydrolase fold, and an α -helical bundle. Here, we attempted to elucidate substrate recognition mechanism at the S2 subsite of DPP7 from *S. maltophilia* (SmDPP7) using enzymatic and structural study for gaining novel insights into the design of a universal inhibitor of S46 peptidases.

2.2. Materials and Methods

2.2.1. Cloning

These S46 peptidases coding gene were used in this study: *dpp7* (UniProt accession number B4SLK2) and *dpp11* (UniProt accession number B4SNQ8) from multidrug-resistant bacterium, *Stenotrophomonas maltophilia* R551-3; *dpp7* (UniProt accession number Q7MWU6) and *dpp11* (UniProt accession number B2RID1) from the periodontal bacterium, *Porphyromonas gingivalis* W83. Synthetic genes coding S46 peptidases, codon-optimised for expression in *E. coli*, were purchased from GenScript (Piscataway, NJ, US). In addition, Synthetic genes in which the native signal peptide of each S46 peptidase was replaced with the signal peptide (b3 signal) of DAP BIII from *Pseudoxanthomonas mexicana* WO24 were purchased from GenScript (Piscataway, NJ, US). Using restriction enzymes NdeI and BamHI, the coding region containing the native signal peptide was cloned into the multi-cloning site (under the T7 promoter) of the pET-22b (+) expression plasmid (Merck, Germany). Subsequently, we constructed an expression construct (pET_b3_s46peptidase) in which the native signal peptide was replaced by the b3-signal using various restriction enzymes (Figure 2-1). DNA sequencing was performed by Eurofins Genomics (Japan).

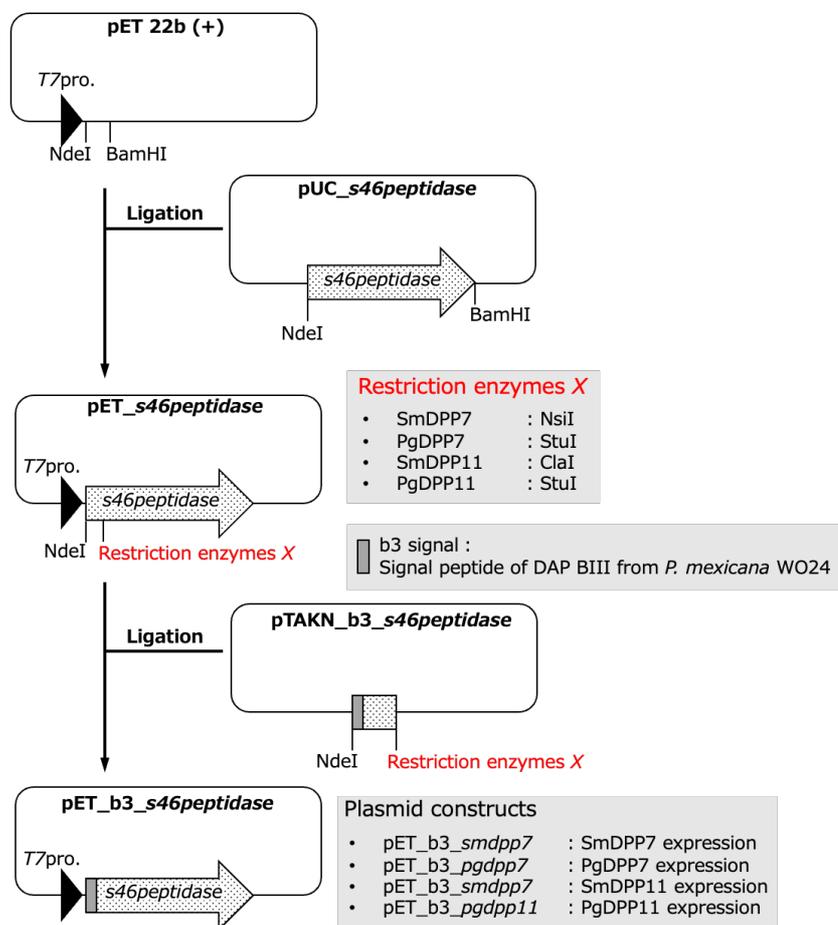


Figure 2-1 | Schematic diagram of the construction of a construct for S46 peptidase expression.

2.2.2. Effect of b3-signal on heterologous expression of S46 peptidase in *E. coli*

When PmDAP BII was expressed in *E. coli*, the solubilized expression level of the enzyme was increased in case of using plasmid construction containing b3-signal (data not disclosed). Considering this finding, we added the b3-signal to the expression of SmDPP7, PgDPP7, SmDPP11, and PgDPP11. In this section, we investigated the effect of the addition of the b3 signal on the expression level of the S46 peptidases in the

heterologous host expression using *E. coli*. The expression constructs with b3-signal or wild-type signal added upstream of each structural gene were introduced into *E. coli* BL21-gold (DE3) and cultured in 2×YT. IPTG was added to the culture medium at OD₆₀₀ 0.6 to achieve a final concentration of 0.1 mM. After the addition of IPTG, *E. coli* was incubated for 15 hours, and the enzyme was extracted from the *E. coli* by sonication, and the supernatant (cell extract) obtained by centrifugation. The amount of expression was quantified by measuring the activity against dipeptidyl peptidase substrate in the cell extract. The results showed that the b3 signal had no significant effect on the expression level of S46 peptidase from *S. maltophilia* in the phylum Proteobacteria (up to 1.76-fold), while it had a significant effect on the expression level of S46 peptidase from *P. gingivalis* in the phylum Bacteroidetes (up to 6.03-fold). This suggests that the addition of the b3 signal is useful for increasing the expression level of S46 peptidase from microorganisms belonging to a different phylum from that of *E. coli*, which belongs to the Proteobacteria phylum. For the subsequent heterologous host expression of S46 peptidase in *E. coli*, we used the expression construct containing the b3 signal.

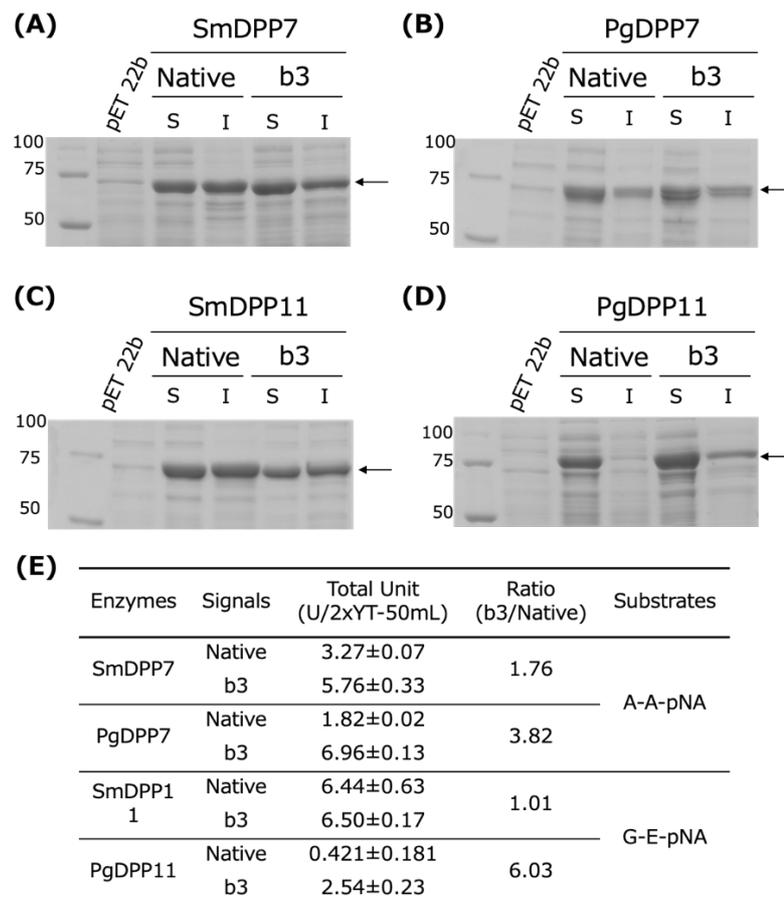
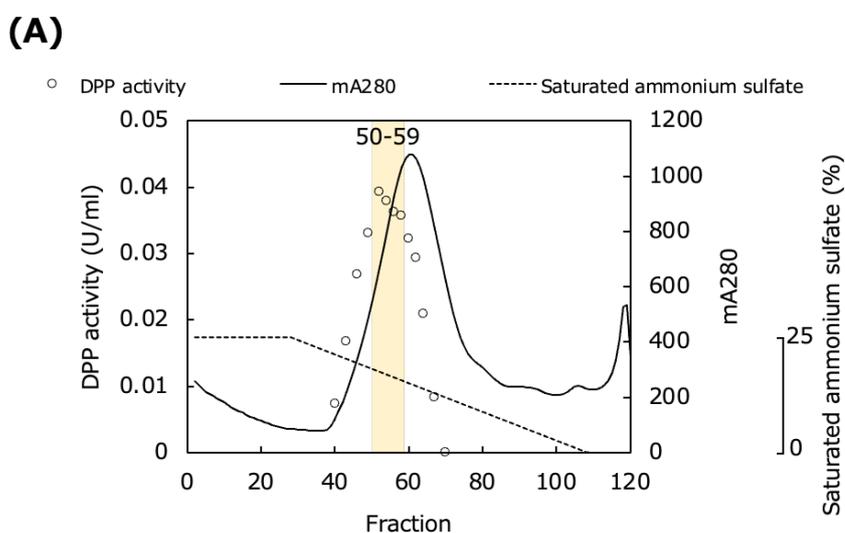


Figure 2-2| Effect of b3 signaling on heterologous host expression of S46 peptidase in *E. coli*. (A)-(D) Expression analysis using SDS-PAGE. 10% acrylamide gel was used and stained with CBB. 10 µl of cell extracts were applied as samples. Markers were Precision Plus protein uncolored standards (BioRad). pET22b indicates cell extracts from *E. coli* transfected with pET-22b (+) vector. b3 indicates extract sample from cells which transfected plasmid containing b3 signal peptide. “S” and “I” represent soluble and insoluble fractions, respectively. (A) SmDPP7. (B) PgDPP7. (C) SmDPP11. (D) PgDPP11. (E) Activity of cell extracts. enzyme activity was measured under the 50 mM sodium phosphate buffer pH 7.0 and 5 mM EDTA at 37 °C for 20 min. In SmDPP7 and PgDPP7, 0.3 mM Ala-Ala-pNA was used as substrate; in SmDPP11 and PgDPP11, 0.3 mM Gly-Glu-pNA was used as substrate.

2.2.3. Overexpression and purification of SmDPP7 WT and mutants

The mature SmDPP7 was composed of 698 amino acids (residues Ala23 to Lys720), with a theoretical molecular weight of 75673.1 and an isoelectric point (pI) of 8.07. Plasmids for expression of mutants,

K206A, R218A, R218Q, R218K, T220A, and F671A, were obtained with overlap extension PCR using wild type expression plasmid as a template (Table 2-1). *E. coli* BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, US) transformed with the pET_b3_smdpp7 WT and mutants expression plasmid were grown in TB media at 298 K to an OD₆₀₀ of 0.6. Overproductions of SmDPP7 WT and mutants were induced by adding 0.1 mM Isopropyl-β-D-thiogalactopyranoside (IPTG) for 15 h at 298 K. Thereafter, the cells were harvested by centrifugation at 6,000 x g. Cells were disrupted using sonication and the cell extract was obtained by centrifuging the lysate at 20,000 x g for 30 min. The SmDPP7 WT and mutants were purified by precipitation with 35 to 70% ammonium sulfate and hydrophobic column chromatography using a HiPrep 16/10 Butyl column (Cytiva, MA, US) (Figure 2-3). Gly-Phe-pNA hydrolysis activity was confirmed in the eluted protein peaks, and fraction numbers 50 to 59 were collected. The eluate was desalted using a HiPrep 26/10 desalting column (Cytiva) and finally subjected to anion-exchange column chromatography using a Mono Q 5/50 GL column (Cytiva). Since the pI of SmDPP7 is 8.07, it does not adsorb on anion chromatography at pH 9.0 of Buffer. The non-binding eluate containing SmDPP7 WT and mutants were pooled, buffer-exchanged to 80 mM Tris/HCl pH 8.5 and concentrated to 5 mg/ml using Vivaspin 20 concentrator (Cytiva). Purity and molecular mass of purified SmDPP7 were estimated by SDS-PAGE (Figure 2-4).



(B)

	Total Protein (mg)	Total act. (U)	Sp. act. (U/mg)	Yield (%)	Purification fold
Cell extract	1713	4.22	0.00389	100	1.00
35-70% ASP	1401	3.61	0.00258	85.5	0.663
HIC	123.9	1.16	0.00935	27.4	2.40
AEC	9.95	0.907	0.0912	21.5	23.5

Figure 2-3 | Purification of SmDPP7 WT
 (A) Hydrophobic interaction chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a linear concentration gradient of 25% to 0% saturated ammonium sulfate. Fractions shown in yellow were collected. (B) Protein amount and DPP activity at each purification step. The Bradford method was used for protein measurement. Gly-Phe-pNA was used as a substrate for measurement of DPP activity. ASP, HIC, and AEC represents ammonium sulfate precipitation, hydrophobic interaction chromatography, anion exchange chromatography, respectively.

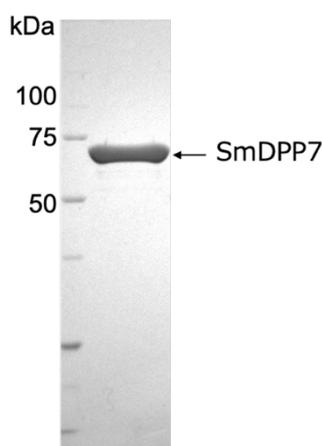


Figure 2-4 | Coomassie-stained 10% SDS-PAGE analysis of purified SmDPP7 6 µg protein loaded per lane. Molecular-weight markers of different masses (kDa) are labelled.

Table 2-1 | Primers used for construction of mutant SmDPP7s.

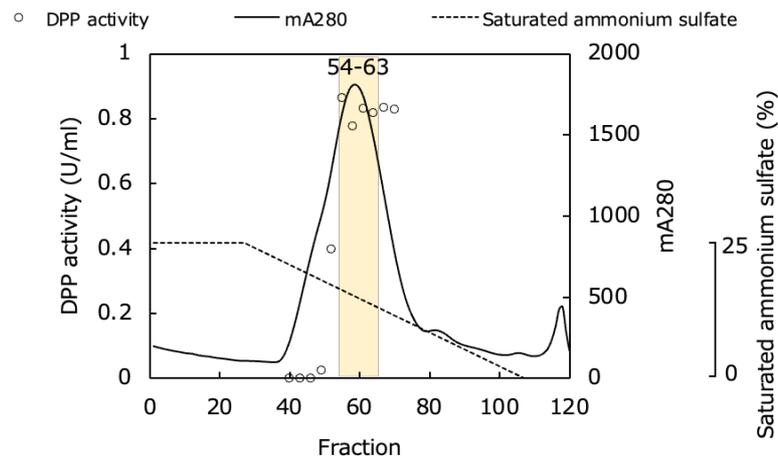
Primer	Sequence 5' → 3'
K206A-Fw	<u>GTCGGT</u> <u>gcg</u> TTCGGCGGTGATGTGGAC
K206A-Rv	GCCGA <u>Acgc</u> ACCGACGCTGCCCGGCGG
R218A-Fw	TGGCCG <u>gcg</u> CATACCGGTGATTTTTCT
R218A-Rv	GGTATG <u>gcg</u> CGGCCACATCCAGTTGTC
R218Q-Fw	TGGCCG <u>cag</u> CATACCGGTGATTTTTCT
R218Q-Rv	GGTATG <u>ctg</u> CGGCCACATCCAGTTGTC
R218K-Fw	TGGCCG <u>aaa</u> CATACCGGTGATTTTTCT
R218K-Rv	GGTATG <u>ttt</u> CGGCCACATCCAGTTGTC
T220A-Fw	CGTCAT <u>gcg</u> GGTGATTTTTCTTTCTAT
T220A-Rv	ATCAC <u>Cgc</u> ATGACGGCCACATCCA
F671A-Fw	GGTCTGGC <u>agc</u> GATGGTAACTGGGAAAGT
F671A-Rv	ATC <u>cgc</u> TGCCAGACCAACCAGTTT

Underlined area represents overlap regions and lower-case letters represent mutation points

2.2.4. Overexpression and purification of SmDPP11

The mature SmDPP11 was composed of 692 amino acids (residues Asp24 to Gln715), with a theoretical molecular weight of 75602.5 and an isoelectric point (pI) of 8.65. *E. coli* BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, US) transformed with the pET_b3_ *sm DPP11* expression plasmid were grown in TB media at 298 K to an OD₆₀₀ of 0.6. Overproduction of SmDPP11 was induced by adding 0.1 mM IPTG for 15 h at 298 K. Thereafter, the cells were harvested by centrifugation at 6,000 x g. Cells were disrupted using sonication and the cell extract was obtained by centrifuging the lysate at 20,000 x g for 30 min. The SmDPP11 was purified by precipitation with 35 to 70% ammonium sulfate and hydrophobic column chromatography using a HiPrep 16/10 Butyl column (Cytiva, MA, US) (Figure 2-5). Gly-Glu-pNA hydrolysis activity was confirmed in the eluted protein peaks, and fraction numbers 54 to 63 were collected. The eluate was desalted using a HiPrep 26/10 desalting column (Cytiva) and finally subjected to anion-exchange column chromatography using a Mono Q 5/50 GL column (Cytiva). Since the pI of SmDPP11 is 8.65, it does not adsorb on anion chromatography at pH 9.0 of Buffer. The non-binding eluate containing SmDPP11 was pooled, buffer-exchanged to 80 mM Tris/HCl pH 8.5 and concentrated to 5 mg/ml using Vivaspin 20 concentrator (Cytiva). Purity and molecular mass of purified SmDPP11 were estimated by SDS-PAGE (Figure 2-6).

(A)



(B)

	Total Protein (mg)	Total act. (U)	Sp. act. (U/mg)	Yield (%)	Purification fold
Cell extract	1432	1376	0.961	100	1.00
35-70% ASP	745	834	1.12	60.6	1.17
HIC	137	321	2.35	23.3	2.44
AEC	57.5	205	3.56	14.9	3.71

Figure 2-5 | Purification of SmDPP11
(A) Hydrophobic interaction chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a linear concentration gradient of 25% to 0% saturated ammonium sulfate. Fractions shown in yellow were collected. (B) Protein amount and DPP activity at each purification step. The Bradford method was used for protein measurement. Gly-Glu-pNA was used as a substrate for measurement of DPP activity. ASP, HIC, and AEC represents ammonium sulfate precipitation, hydrophobic interaction chromatography, anion exchange chromatography, respectively.

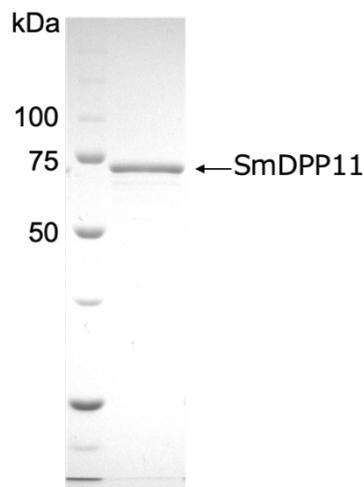
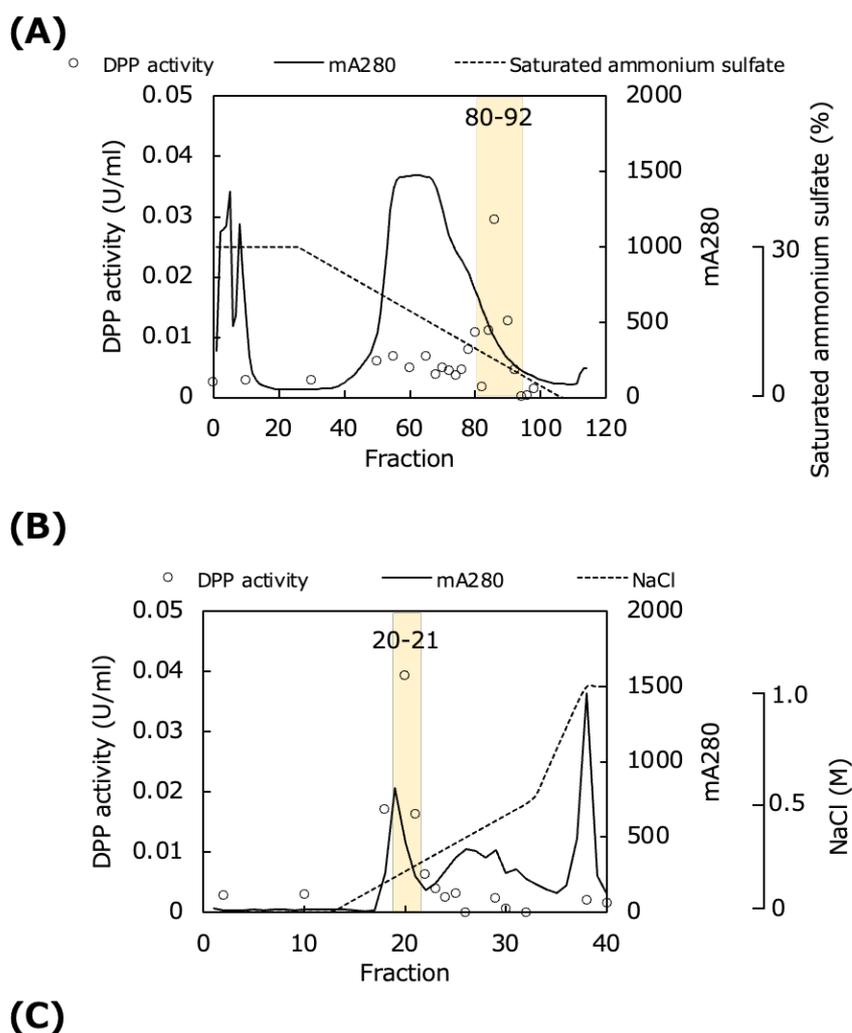


Figure 2-6 | Coomassie-stained 10% SDS-PAGE analysis of purified SmDPP11
3 µg protein loaded per lane. Molecular-weight markers of different masses (kDa) are labelled.

2.2.5. Overexpression and purification of PgDPP7

The mature PgDPP7 was composed of 688 amino acids (residues Asp25 to Ile712), with a theoretical molecular weight of 77690.5 and an isoelectric point (pI) of 7.23. *E. coli* BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, US) transformed with the pET_b3_pgdp7 expression plasmid were grown in TB media at 298 K to an OD₆₀₀ of 0.6. Overproduction of PgDPP7 was induced by adding 0.1 mM IPTG for 15 h at 298 K. Thereafter, the cells were harvested by centrifugation at 6,000 x g. Cells were lysed using BugBuster Protein Extraction Reagent (Merck), and the cell extract was obtained by centrifuging the lysate at 20,000 x g for 30 min. The PgDPP7 was purified by precipitation with 35 to 70% ammonium sulfate and hydrophobic column chromatography using a HiPrep 16/10 Butyl column (Cytiva, MA, US) (Figure 2-7 A). Gly-Phe-pNA hydrolysis activity was confirmed in the eluted protein peaks, and fraction numbers 82 to 92 were collected. The eluate was desalted using a HiPrep 26/10 desalting column (Cytiva) and finally subjected to anion-exchange column chromatography using a Mono Q 5/50 GL column (Cytiva) (Figure 2-7 B). The fractions containing PgDPP7 (fraction numbers 20 and 21) was pooled, buffer-exchanged to 80 mM Tris/HCl pH 8.5 and concentrated to 5 mg/ml using Vivaspin 20 concentrator (Cytiva). Purity and molecular mass of purified PgDPP7 were estimated by SDS-PAGE (Figure 2-8).



	Total Protein (mg)	Total act. (U)	Sp. act. (U/mg)	Yield (%)	Purification fold
Cell extract	276	1.76	0.00639	100	1.00
35-70% ASP	151	0.343	0.00227	19.4	0.360
HIC	21.9	0.298	0.0136	16.9	2.10
AEC	1.97	0.294	0.149	16.7	23.0

Figure 2-7 | Purification of PgDPP7
 (A) Hydrophobic interaction chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a linear concentration gradient of 30% to 0% saturated ammonium sulfate. Fractions shown in yellow were collected. (B) anion exchange chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a linear concentration gradient of 0 M to 0.5 M NaCl. Fractions shown in yellow were collected. (C) Protein amount and DPP activity at each purification step. The Bradford method was used for protein measurement. Gly-Phe-pNA was used as a substrate for measurement of DPP activity. ASP, HIC, and AEC represents ammonium sulfate precipitation, hydrophobic interaction chromatography, anion exchange chromatography, respectively.

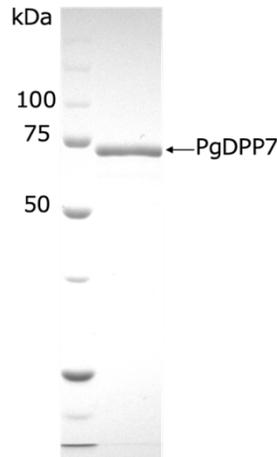
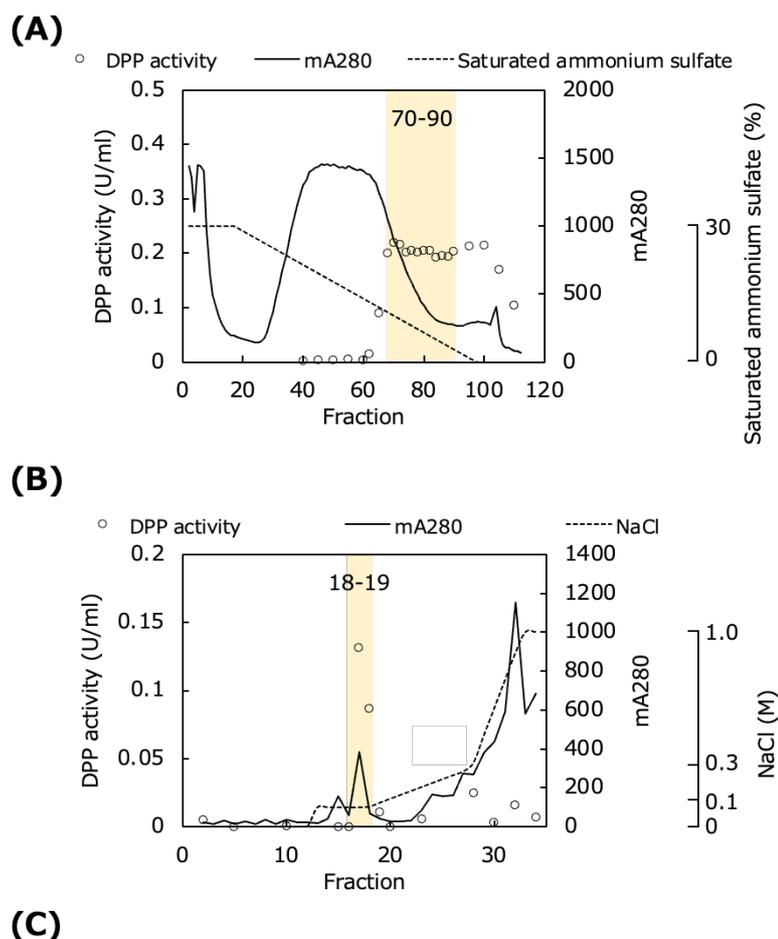


Figure 2-8 | Coomassie-stained 10% SDS-PAGE analysis of purified PgDPP7. 3 µg protein loaded per lane. Molecular-weight markers of different masses (kDa) are labelled.

2.2.6. Overexpression and purification of PgDPP11

The mature PgDPP11 was composed of 699 amino acids (residues Asp22 to Pro720), with a theoretical molecular weight of 79549.3 and an isoelectric point (pI) of 5.88. *E. coli* BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, US) transformed with the pET_b3_pgdp11 expression plasmid were grown in TB media at 298 K to an OD₆₀₀ of 0.6. Overproduction of PgDPP11 was induced by adding 0.1 mM IPTG for 15 h at 298 K. Thereafter, the cells were harvested by centrifugation at 6,000 x g. Cells were disrupted using sonication and the cell extract was obtained by centrifuging the lysate at 20,000 x g for 30 min. The PgDPP11 was purified by precipitation with 35 to 70% ammonium sulfate and hydrophobic column chromatography using a HiPrep 16/10 Butyl column (Cytiva, Marlborough, MA, US) (Figure 2-9 A). Gly-Glu-pNA hydrolysis activity was confirmed in the eluted protein peaks, and fraction numbers 70 to 90 were collected. The eluate was desalted using a HiPrep 26/10 desalting column (Cytiva) and finally subjected to anion-exchange column chromatography using a Mono Q 5/50 GL column (Cytiva) (Figure 2-9 B). The fractions containing PgDPP11 (fraction numbers 18 and 19) was pooled, buffer-exchanged to 80 mM Tris/HCl pH 8.5 and concentrated to 5 mg/ml using Vivaspin 20 concentrator (Cytiva). Purity and molecular mass of purified PgDPP11 were estimated by SDS-PAGE (Figure 2-10).



	Total Protein (mg)	Total act. (U)	Sp. act. (U/mg)	Yield (%)	Purification fold
Cell extract	2160	22.1	0.0103	100	1.00
35-70% ASP	2390	14.4	0.00603	65.2	0.587
HIC	156	9.02	0.0579	40.7	5.63
AEC	2.94	0.302	0.103	1.36	10.0

Figure 2-9 | Purification of PgDPP11
 (A) Hydrophobic interaction chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a linear concentration gradient of 30% to 0% saturated ammonium sulfate. Fractions shown in yellow were collected. (B) anion exchange chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a multi-step concentration gradient of 0 M to 0.1 M and 0.1 M to 0.3 M NaCl. Fractions shown in yellow were collected. (C) Protein amount and DPP activity at each purification step. The Bradford method was used for protein measurement. Gly-Glu-pNA was used as a substrate for measurement of DPP activity. ASP, HIC, and AEC represents ammonium sulfate precipitation, hydrophobic interaction chromatography, anion exchange chromatography, respectively.

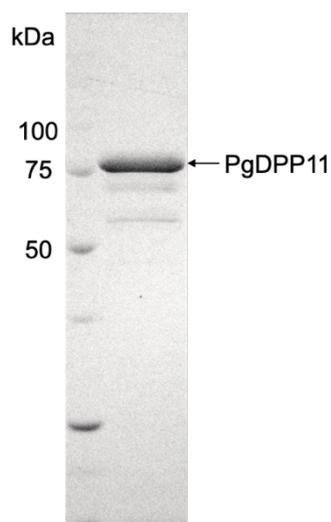


Figure 2-10 | Coomassie-stained 10% SDS-PAGE analysis of purified PgDPP11 5 μ g protein loaded per lane. Molecular-weight markers of different masses (kDa) are labelled.

2.2.7. Determination of substrate specificity toward dipeptidyl pNA

Dipeptidyl-*p*-nitroanilide (pNA) was used as a substrate at 300 μ M in a 200 μ L of reaction buffer consisting of 50 mM sodium phosphate buffer pH 7.0 and 5 mM EDTA. 10 μ L of purified SmDPP7, PgDPP7, SmDPP11, and PgDPP11 were added into reaction buffer. Enzyme reaction was conducted at 310 K for 20 min. Absorbance was measured at 385 nm using microplate reader (Infinite 200 PRO microplate reader, Tecan, Switzerland). One unit (U) is the amount of enzyme that released 1 μ mol of pNA from dipeptidyl-pNA substrate per 1 min. “n.d.” means not detected due to low activity. The standard deviations were obtained from three independent experiments.

2.2.8. Determination of kinetic parameters toward dipeptidyl MCA

Kinetic parameters were determined by fitting the experimental data to the Michaelis–Menten equation using Excel Solver (Microsoft, WA, US) by nonlinear least-squares fitting with different concentrations of glycyl-L-tyrosyl-4-methylcoumaryl-7-amide (3.91, 7.81, 15.6, 31.3, 62.5, 125, 250, and 500 μ M; Gly-Tyr-MCA; GenScript, NJ, US), L-tyrosyl-L-tyrosyl-4-methylcoumaryl-7-amide (1.56, 3.13, 6.25, 12.5, 25, 50, 100, and 200 μ M; Tyr-Tyr-MCA; GenScript, NJ, US), L-asparaginy-L-tyrosyl-4-methylcoumaryl-7-amide (1.56, 3.13, 6.25, 12.5, 25, 50, 100, and 200 μ M; Asn-Tyr-MCA; GenScript, NJ, US), and L-leuciny-L-asparaginic acid-4-methylcoumaryl-7-amide (0.781, 1.56, 3.13, 6.25, 12.5, 25, 50, and 100 μ M; Leu-Asp-MCA; Peptide Institute, Osaka, Japan) as a substrate. The enzyme reaction was performed in a reaction buffer consisting of 50 mM sodium phosphate buffer pH 7.0, 5 mM EDTA, and 0.005% Tween 20 at 298 K for 20 min and standard deviations were calculated from three independent experiments. Tween 20 was added into reaction buffer to prevent proteins adsorbing to the plastic surface of the 96well plate (Figure 2-11). For SmDPP7 (wild-type and mutants), concentrations of the purified enzymes were 10, 2, and 2 nM for the hydrolyses of Gly-Tyr-MCA, Tyr-Tyr-MCA, and Asn-Tyr-MCA, respectively. Concentration of purified PgDPP7 used for determining kinetic parameters of Tyr-Tyr-MCA hydrolysis was 25 nM. For SmDPP11 and PgDPP11, concentrations of the purified enzymes used for determining kinetic parameters of Leu-Asp-MCA hydrolysis were 0.5 nM and 1 nM, respectively. The fluorescence intensity of the released MCA was measured with excitation at 355 nm and emission at 460 nm using an Infinite 200 PRO microplate reader (Tecan, Switzerland).

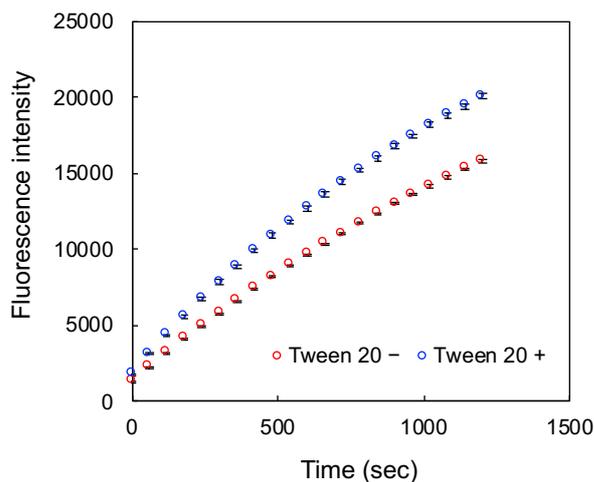


Figure 2-11 | Effect of the addition of Tween 20 on enzyme activity measurement of SmDPP7. Tyr-Tyr-MCA was used at 100 μ M as a substrate in a reaction buffer consisting of 50 mM sodium phosphate buffer pH 7.0 and 5 mM EDTA. For Tween 20 +, 0.005% Tween20 was contained in the reaction buffer. Purified SmDPP7 was used in reaction buffer with a concentration of 2 nM. Enzyme reaction was conducted at 298 K for 20 min. The fluorescence intensity of the released MCA was measured with excitation at 355 nm and emission at 460 nm using an Infinite 200 PRO microplate reader. The standard deviations were obtained from three independent experiments.

2.2.9. Inhibition assay using dipeptide library

The dipeptide library was purchased from AnaSpec (Fremont, CA, US). Ile-Tyr and Arg-Tyr were synthesized by Eurofins Genomics (Tokyo, Japan). Cys-Tyr and Cys-Asp could not be purchased commercially. Different concentrations (3.13, 6.25, 12.5, 25, 50, 100, and 200 μ M) of dipeptides as the inhibitor and 200 μ M Tyr-Tyr-MCA (GenScript, NJ, US) as the substrate for DPP7 were added into a reaction buffer consisting of 50 mM sodium phosphate buffer pH 7.0, 5 mM EDTA and 0.005% Tween 20. For DPP11s, different concentrations (1.56, 3.13, 6.25, 12.5, 25, 50, and 100 μ M) of dipeptides as the inhibitor and 100 μ M Leu-Asp-MCA (Peptide Institute, Osaka, Japan) as the substrate were used. The enzyme reaction was performed at 298 K for 20 min and standard deviations were calculated from three independent experiments. Half-maximal inhibitory concentration (IC_{50}) values were determined by fitting to a sigmoid curve (4-parameter logistic curve) using ImageJ software⁹⁶ and inhibition constants (K_i) were calculated using the Cheng-Prusoff equation.⁹⁷ The fluorescence intensity of the released MCA was measured with the same method of determination of kinetics parameters toward dipeptidyl MCA.

2.2.10. Crystallization of SmDPP7

(Collaborated with Prof. Nobutada Tanaka of Kitasato University and Assoc. Prof. Yasumitsu Sakamoto of Iwate Medical University.)

Crystals of the SmDPP7/dipeptide complexes were prepared as follows: Asn-Tyr and Val-Tyr-Pro were purchased from GenScript (NJ, US). Phe-Tyr and Tyr-Tyr-Tyr were purchased from Sigma-Aldrich (St Louis, MO, US) and Santa Cruz Biotech (Dallas, TX, SUA) respectively. The tripeptides, Val-Tyr-Pro and Tyr-Tyr-Tyr (tripeptides were used as a substrate, and the N-terminal dipeptide was observed as the bound product as described below), were dissolved in 80 mM Tris-HCl, pH 8.5 to concentrations of 30.0 mM and 20.0 mM, respectively. The 5-mg/ml SmDPP7 solution was mixed with aliquots of the respective ligand solutions at a volume ratio of 9:1 for the tripeptides Val-Tyr-Pro and Tyr-Tyr-Tyr, with final ligand concentrations of 3.0 mM and 2.0 mM, respectively. The samples were crystallized using the hanging-drop method; 1 μ l of protein solution was mixed with the same volume of reservoir solution (20% (w/v) PEG8000 and 0.2 M ammonium acetate) and incubated at 293 K. The drops were resuspended over 200 μ l of reservoir solution in 48-well plates. Asn-Tyr and Phe-Tyr were dissolved in purified water to concentrations of 100 mM. The 5-mg/ml SmDPP7

solution was mixed with aliquots of the respective ligand solutions at a volume ratio of 25:1 for the dipeptides Asn-Tyr and Phe-Tyr, with a final ligand concentration of 4.0 mM. The samples were crystallized using the hanging-drop method; 0.95 μ l of protein solution was mixed with the same volume of reservoir solution (20%(w/v) PEG8000 and 0.2 M Calcium acetate) and incubated at 293 K. The drops were suspended over 200 μ l of reservoir solution in 48-well plates.

A dipeptide complex, Tyr-Tyr complex, was obtained by co-crystallization of SmDPP7 with a tripeptide Tyr-Tyr-Tyr, because the tripeptide Tyr-Tyr-Tyr was commercially available at a lower cost as compared with custom peptide synthesis of the dipeptide Tyr-Tyr. For the Tyr-Tyr complex, clear continuous electron density was observed for the first two residues of the tripeptide (Figure 2-12A), and no clear electron density was observed for the last residue. Because the SmDPP7 enzyme reaction occurred in the solution used for crystallization, the Tyr-Tyr-Tyr (P2-P1-P1') tripeptide acted as the substrate, and the reaction products were the N-terminal Tyr-Tyr (P2-P1) dipeptide and the C-terminal Tyr (P1'). While the N-terminal product Tyr-Tyr remained at the active site, the C-terminal product Tyr dissociated from the active site. The asymmetric unit of the Tyr-Tyr complex was composed of two independent SmDPP7 subunits; in both subunits (Figure 2-17A), the hydrolyzed dipeptide product (NH₂-Tyr-Tyr-COOH), rather than a reaction intermediate, was observed (Figure 2-12A). Similarly, another dipeptide complex, Val-Tyr complex, was obtained by co-crystallization of SmDPP7 with a tripeptide Val-Tyr-Pro and the hydrolyzed dipeptide product (NH₂-Val-Tyr-COOH) was observed (Figure 2-12B). In this case, co-crystallization with Val-Tyr-Pro was performed to confirm the ability of imino-bond (X-Pro) hydrolysis by SmDPP7, though the dipeptide Val-Tyr was commercially available at a reasonable cost. The other dipeptide complexes (Phe-Tyr and Asn-Tyr complexes), were obtained by co-crystallization of SmDPP7 with dipeptides Phe-Tyr and Asn-Tyr, respectively (Figure 2-12C and D).

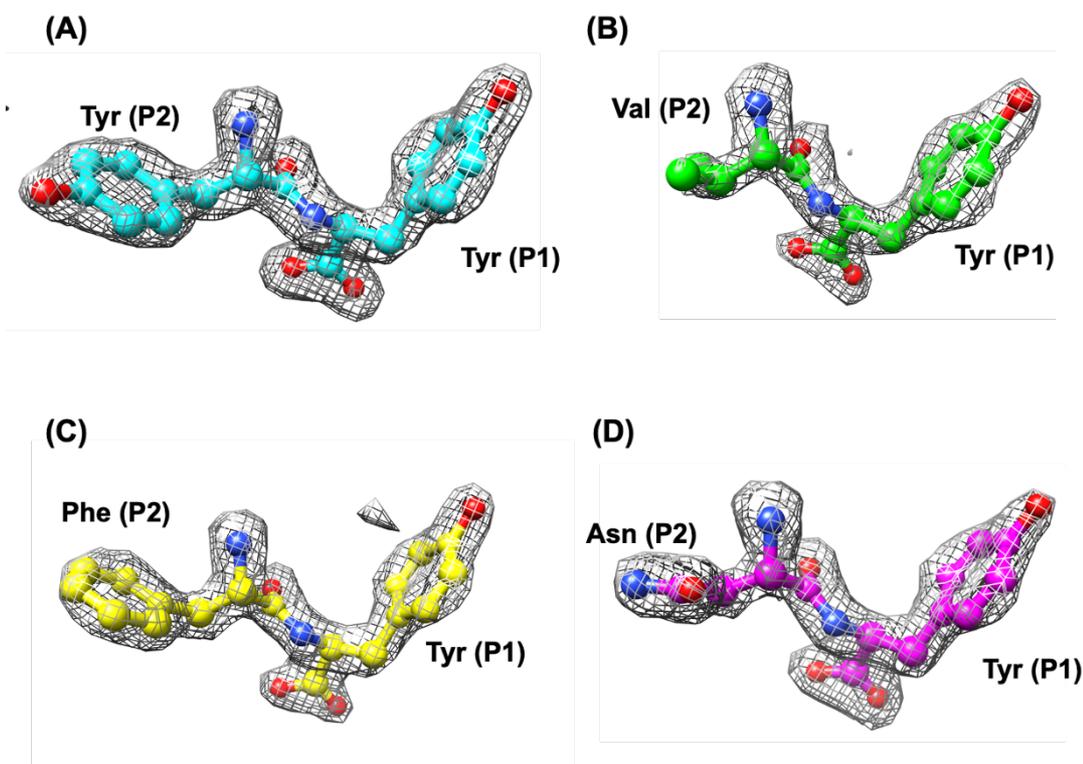


Figure 2-12 | Diagrams showing weighted $m|F_o|-D|F_c|$ omit maps of the bound ligand molecule in the active site of SmDPP7. The contour levels are 3.0σ (grey). (A) Tyr-Tyr dipeptide at a 1.86-Å resolution. (B) Val-Tyr dipeptide at a 2.03-Å resolution. (C) Phe-Tyr dipeptide at a 1.91-Å resolution. (D) Asn-Tyr dipeptide at a 1.92-Å resolution. These figures were produced using the program UCSF Chimera version 1.14.

2.2.11. X-ray data collection

(Collaborated with Prof. Nobutada Tanaka of Kitasato University and Assoc. Prof. Yasumitsu Sakamoto of Iwate Medical University.)

For data collection under cryogenic conditions, dipeptide-complex crystals in a droplet were directly transferred to harvesting solutions [16%(w/v) PEG8000, 0.16 M ammonium acetate and 20%(v/v) glycerol] and [16%(w/v) PEG8000, 0.16 M calcium acetate and 20%(v/v) glycerol] respectively for 10 seconds. Crystals were mounted in nylon loops or MicroMounts (MiTeGen, Ithaca, NY, US) and flash-cooled in a cold nitrogen gas stream at 100 K immediately before data collection. Data were collected by the rotation method at 100 K using a MAR300HE CCD detector or EIGER 16M detector with synchrotron radiation source on the beamline BL44XU at SPring-8. Laue group and unit-cell parameters were determined using the xia2/DIALS software package⁹⁸ with XDS⁹⁹ or MOSFLM.¹⁰⁰ The cell parameters and data-collection statistics are summarized in Table 2-2.

Table 2-2 | Data collection statistics for SmDPP7.

Data set	Val-Tyr	Tyr-Tyr	Asn-Tyr	Phe-Tyr
Facility	SPring-8	SPring-8	SPring-8	SPring-8
Beamline	BL44XU	BL44XU	BL44XU	BL44XU
Wavelength (Å)	0.90000	0.90000	0.90000	0.90000
Detector	MX300HE	MX300HE	EIGER16M	EIGER16M
Crystal-to-detector distance (mm)	325.4	289.3	310.0	250.0
Rotation angle per image (°)	0.3	0.3	0.1	0.1
Total rotation range (°)	369.6	289.3	360	360
Exposure time per image (sec)	0.1	3 (gradual move mode)	0.1	0.1
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions				
<i>a</i> (Å)	66.490	67.850	68.552	68.035
<i>b</i> (Å)	73.330	74.320	74.892	74.555
<i>c</i> (Å)	151.820	153.320	154.647	153.869
α (°)	90	90	90	90
β (°)	95.16	94.15	94.50	94.44
γ (°)	90	90	90	90
Number of molecules per ASU	2	2	2	2
Mosaicity (°)	0.289	0.302	0.175	0.194
Resolution (Å) (outer shell)	52.64-2.03 (2.07-2.03)	52.60-1.86 (1.89-1.86)	50.48-1.92 (1.95-1.92)	43.13-1.91 (1.94-1.91)
No. of observed reflections	700,660 (28,063)	531,519 (22,826)	823,378 (39,675)	816,227 (37,362)
No. of unique reflections	91,976 (3,919)	127,585 (6,238)	119,220 (5,629)	118,877 (5,580)
Completeness (%)	97.7 (83.9)	99.8 (98.0)	99.7 (94.8)	99.6 (93.9)
Redundancy	7.6 (7.2)	4.2 (3.7)	6.9 (7.0)	6.9 (6.7)
<i>I</i> / σ(<i>I</i>)	8.5 (2.0)	13.0 (2.0)	7.4 (2.0)	7.0 (2.0)
CC _{half}	0.991 (0.597)	0.998 (0.637)	0.989 (0.632)	0.991 (0.631)
<i>R</i> _{merge} (<i>I</i>)	0.251 (1.179)	0.077 (0.698)	0.186 (1.199)	0.157 (1.012)
<i>R</i> _{meas} (<i>I</i>)	0.270 (1.271)	0.088 (0.820)	0.201 (1.295)	0.170 (1.098)
<i>R</i> _{pim} (<i>I</i>)	0.098 (0.470)	0.043 (0.421)	0.076 (0.485)	0.064 (0.421)
Wilson <i>B</i> -factor (Å ²)	17.4	19.9	14.5	25.2

2.2.12. Structure determination

(Collaborated with Prof. Nobutada Tanaka of Kitasato University and Assoc. Prof. Yasumitsu Sakamoto of Iwate Medical University.)

The initial phase determination was performed for the Val-Tyr complex of SmDPP7 using the molecular replacement method. One protomer of PmDAP BII³⁵ (PDB code: 3WOL), which has approximately 78% amino-acid sequence identity to SmDPP7, was used as a search model. Cross-rotation and translation functions were calculated using the MOLREP program¹⁰¹ from CCP4 suite.¹⁰² Structure refinement was carried out with the program REFMAC5,¹⁰³ and further iterative manual model building and refinement were performed using the programs Coot¹⁰⁴ and REFMAC5.¹⁰³ The stereochemistry of the model was verified using RAMPAGE¹⁰⁵ and PROCHECK¹⁰⁶ programs. The refined structure of the Val-Tyr complex was then used for the structural determination of the Tyr-Tyr complex by the difference Fourier method. The refined structure of the Tyr-Tyr complex was used for the initial phase determination of Asn-Tyr and Phe-Tyr complexes. Cross-rotation and translation functions were calculated using the program PHASER¹⁰⁷ from the CCP4 suite.¹⁰² After the final round of refinement, the bound dipeptide molecules were removed from the model. Then, the amplitude $|Fc|$ and phase angles calculated from the partial structure were used to calculate a weighted $m|Fo|-D|Fc|$ omit map,¹⁰³ where ‘ m ’ is the figure of merit (approximately equal to the cosine of the phase error) and ‘ D ’ is the estimate of the coordinate error in the partial structure (Figure 2-12). The refinement statistics are summarized in Table 2-3.

Table 2-3 | Refinement statistics for SmDPP7

Dataset	Val-Tyr	Tyr-Tyr	Asn-Tyr	Phe-Tyr
PDB ID	7DKB	7DKC	7DKD	7DKE
Resolution range (Å)	40.00-2.03	40.00-1.86	40.00-1.92	40.00-1.91
Completeness (%)	97.62	99.78	99.80	99.68
No. of reflections working set test set	87,429 4,522	121,344 6,217	113,046 6,099	112,986 5863
<i>R</i> -factor	0.201	0.171	0.202	0.207
Free <i>R</i> -factor	0.245	0.210	0.248	0.251
No. of protein atoms (avg. <i>B</i> -factors (Å ²))	10,658 (31.1)	10,660 (28.9)	10,669 (31.88)	10,667 (35.13)
No. of ligand atoms (avg. <i>B</i> -factors (Å ²))	40 (2 x 20) (23.4)	50 (2 x 25) (22.4)	42 (2 x 21) (20.7)	48 (2 x 24) (23.9)
No. of glycerol atoms (avg. <i>B</i> -factors (Å ²))	0	36 (6 x 6) (44.7)	18 (6 x 3) (30.4)	30 (6 x 5) (39.6)
No. of water molecules (avg. <i>B</i> -factors (Å ²))	461 (31.8)	1,361 (35.1)	1,162 (31.2)	861 (36.9)
Ramachandran plot statistics				
favoured (%)	1,340 (96.4)	1,349 (97.1)	1,325 (95.3)	1,350 (97.1)
allowed (%)	48 (3.5)	40 (2.9)	61 (4.4)	39 (2.8)
outlier (%)	2 (0.1)	1 (0.1)	4 (0.3)	1 (0.1)
RMSD				
bonds (Å)	0.0097	0.0105	0.0090	0.0097
angles (°)	1.58	1.61	1.55	1.57

2.2.13. Isothermal Titration Calorimetry

The bindings were analyzed using a MICROCAL PEAQ-ITC microcalorimeter (Malvern, UK). The binding reactions were performed in 50 mM sodium phosphate pH 7.0 and 2.5% dimethyl sulfoxide at 25 °C and were stirred at 750 rpm. A single injection of 0.4 μl and 19 times injections of 2.0 μl of the dipeptide solution were injected into 350 μl of enzyme solution (wild-type SmDPP7). The wild-type SmDPP7 concentration was 25 μM and the concentration of each dipeptide solution was 250 μM in reaction buffer. Each injection was performed for 4 seconds with an interval of 150 sec between injections. The dissociation constant (K_d) and heats of binding (ΔH) were obtained using MICROCAL PEAQ-ITC Analysis software (Malvern, UK). Gibbs free energy (ΔG) values were calculated according to $\Delta G = -RT \ln K_a = RT \ln K_d$ ($K_a = 1/ K_d$, association constants).

2.3. Results

2.3.1. Determination of optimal pH

The optimal pH of the known S46 peptidases, PmDAP BII, PgDPP7 and PgDPP11, has been reported so far. The optimal pH of PmDAP BII on Gly-Phe-pNA substrate shows 8.0,¹⁰⁸ that of PgDPP7 on Ala-Phe-pNA substrate shows 6.5 - 9.0,¹⁰⁹ and that of PgDPP11 on Leu-Glu-MCA substrate shows 7.0.³⁸ In this section, in addition to the known S46 peptidases PgDPP7 and PgDPP11, we determined the optimal pH of SmDPP7 and SmDPP11. We determined the pH optimum for SmDPP7 and PgDPP7, which have specificity for the hydrophobic amino acid at substrate P1, using Gly-Phe-pNA as a substrate. As a result, the optimal pH for DPP7 against Gly-Phe-pNA was pH 7.0-7.5. Similarly, for SmDPP11 and PgDPP11, which have specificity for acidic amino acids at substrate P1, the optimal pH test was conducted using Gly-Glu-pNA as a substrate. As a result, the optimal pH of base PgDPP11 was 7.0. On the other hand, the optimal pH of SmDPP11 was found to be 7.5. Based on the previous studies and the present study, the optimum pH of S46 peptidase was found to be pH 7.0 - 8.0.

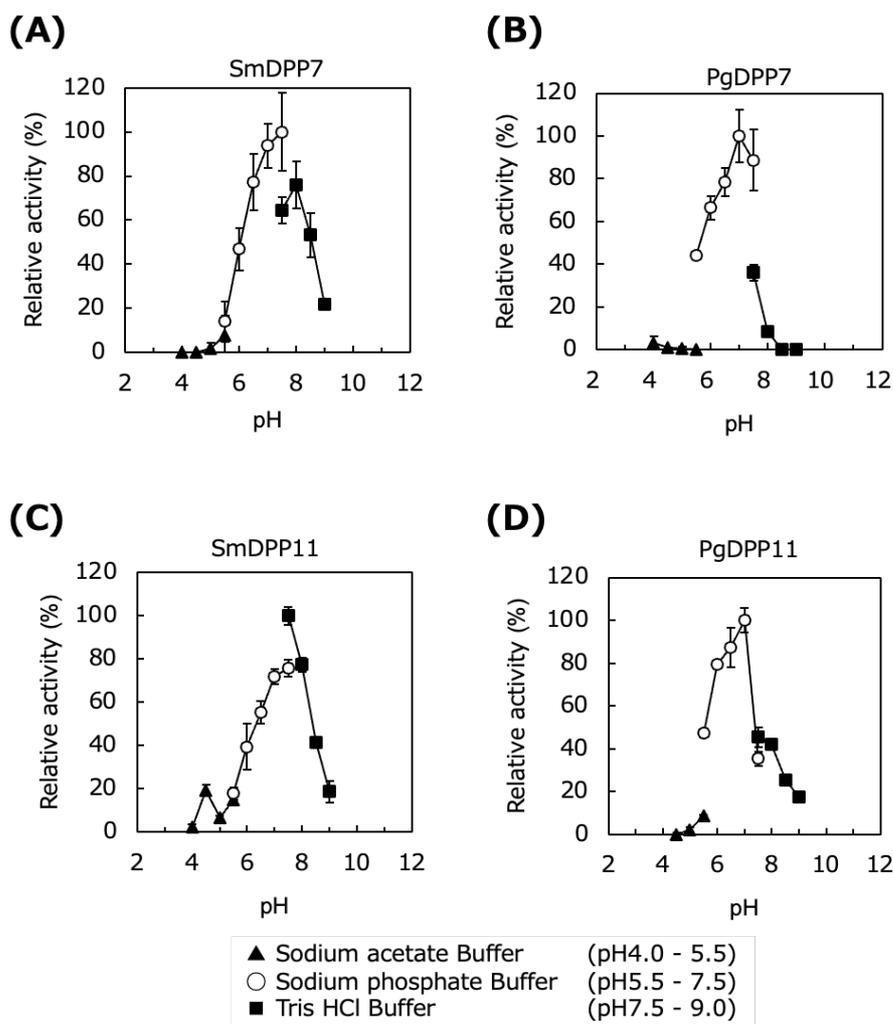
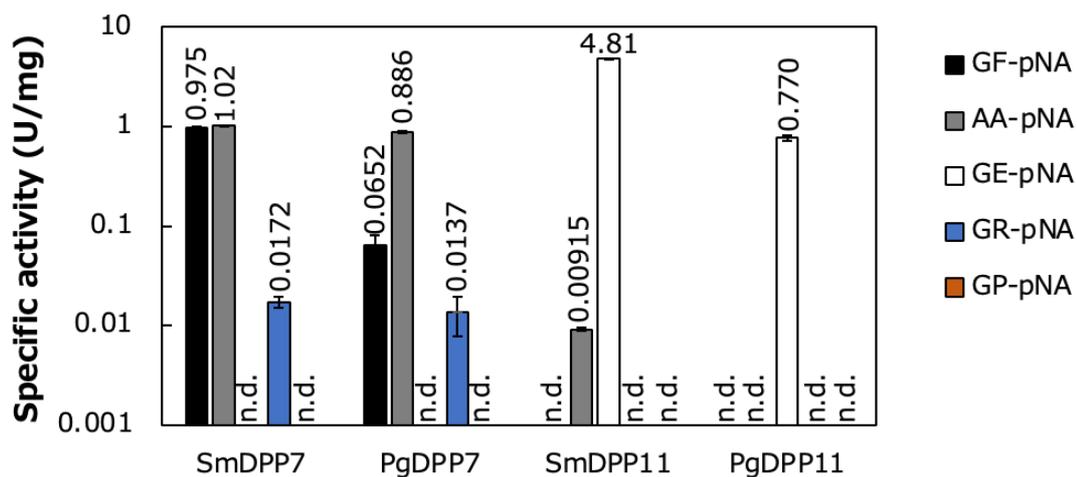


Figure 2-13 | Optimal pH of S46 peptidase. Sodium acetate buffer was used in the pH 4.0-5.5 range; sodium phosphate buffer was used in the pH 5.5-7.5 range; Tris-HCl buffer was used in the pH 7.5-9.0 range. The enzyme activity was measured in the presence of 50 mM of each buffer and 5 mM EDTA at 37 °C for 20 min. 0.3 mM Gly-Phe-pNA was used as substrate for SmDPP7 (A) and PgDPP7 (B). 0.3 mM Gly-Phe-pNA was used as substrate for SmDPP11 (C) and PgDPP11 (D). In SmDPP11 (C) and PgDPP11 (D), 0.3 mM Gly-Glu-pNA was used as substrate. Standard deviation was calculated from three independent measurements.

2.3.2. P1 position preferences of SmDPP7

First, we evaluated the P1 preferences of SmDPP7, SmDPP11, PgDPP7, and PgDPP11 by using synthetic substrates dipeptidyl-p-nitroanilide (Figure 2-14). SmDPP7 showed clear preferences for P1-hydrophobic amino acids and an ability to degrade a dipeptidyl substrate with basic amino acid at the P1 position. This result was in agreement with previous reports, in which PmDAP BII and PgDPP7 showed a preference against P1-hydrophobic amino acids and an ability to degrade dipeptidyl substrates consisting of basic amino acids at the P1 position.^{34,90,110,111}



Enzyme concentrations (mg/ml) added into reaction buffer with the volume of 10 μ L

		Substrates				
		Gly-Phe-pNA	Ala-Ala-pNA	Gly-Glu-pNA	Gly-Arg-pNA	Gly-Pro-pNA
Enzymes	SmDPP7	0.100	0.100	10.0	10.0	10.0
	PgDPP7	0.940	0.940	2.67	0.940	2.67
	SmDPP11	5.00	5.00	0.0142	5.00	5.00
	PgDPP11	3.70	3.70	0.161	3.70	3.70

Figure 2-14 | Specific activities of S46 peptidases toward synthetic substrates. Dipeptidyl-p-nitroanilide (pNA) was used as a substrate at 300 μ M in a 200 μ L of reaction buffer consisting of 50 mM sodium phosphate buffer pH 7.0 and 5 mM EDTA. 10 μ L of purified SmDPP7, PgDPP7, SmDPP11, and PgDPP11 were added into reaction buffer. Enzyme concentrations used in this assay are presented in the table above. Enzyme reaction was conducted at 310 K for 20 min. Absorbance was measured at 385 nm using GeneQuant 100 (Cytiva). One unit (U) is the amount of enzyme that released 1 μ mol of pNA from dipeptidyl-pNA substrate per 1 min. “n.d.” means not detected due to low activity. The standard deviations were obtained from three independent experiments.

2.3.3. P2 position preferences of SmDPP7

In order to quantitatively estimate the preference at the P2 position of the substrate, inhibitory effects of the dipeptides Xaa-Tyr against the hydrolytic activity of wild-type SmDPP7 were evaluated on a synthetic substrate, L-tyrosyl-L-tyrosyl-4-methylcoumaryl-7-amide (Tyr-Tyr-MCA). Here, Xaa indicates any amino acid except for Cys, because Cys-Tyr was not included in the commercially available dipeptide library (AnaSpec, Fremont, CA, US) and the synthetic service (Eurofins Genomics, Japan) that we used. The hydrolytic activity of SmDPP7 was markedly inhibited by Phe-Tyr, Leu-Tyr, Tyr-Tyr, Asn-Tyr, and Met-Tyr, the top five of the 19 dipeptides shown in descending order (Figure 2-15). The inhibitory constants (K_i) of Phe-Tyr, Leu-Tyr, Tyr-Tyr, and Met-Tyr were 1.27, 2.39, 7.66 and 11.2 μ M, respectively, and these were clearly aligned with the

hydrophobicity indexes (H.I.) of P2 (N-terminal) amino acids^{112,113} (Table 2-4). The correlation coefficient between the residual activity of SmDPP7 and H.I. of the P2 amino acid of dipeptides was -0.748 when the outlier dipeptides Asn-Tyr and Trp-Tyr (red and blue, respectively, in Figure 2-16A) were excluded, whereas the value was -0.491 when all the dipeptides were considered. The negative correlations between the residual activity and the H.I. of the P2 amino acid of dipeptides indicated that SmDPP7 prefers the hydrophobic amino acids at the P2 position of the substrate. Interestingly, Asn-Tyr (H.I. of Asn: -28) showed an exceptional inhibitory effect with K_i value of $7.80 \mu\text{M}$ against SmDPP7. In order to evaluate the characteristics of the Asn-Tyr as N-terminal part of the substrate of SmDPP7, the kinetic parameters of SmDPP7 for Asn-Tyr-MCA were determined (Table 2-5). The specificity constant (k_{cat}/K_m) for Asn-Tyr-MCA of wild-type SmDPP7 was $71.5 \text{ sec}^{-1} \text{ mM}^{-1}$, which was 1.5 times higher than that for Tyr-Tyr-MCA. This result indicated that Asn was accommodated into the S2 subsite of SmDPP7 not only as the N-terminal residue of the product dipeptide (Asn-Tyr) but also as the P2 residue of substrate peptide. We further examined the binding of amino acids at the S2 subsite of SmDPP7 by structural and site-directed mutagenesis studies, as described below.

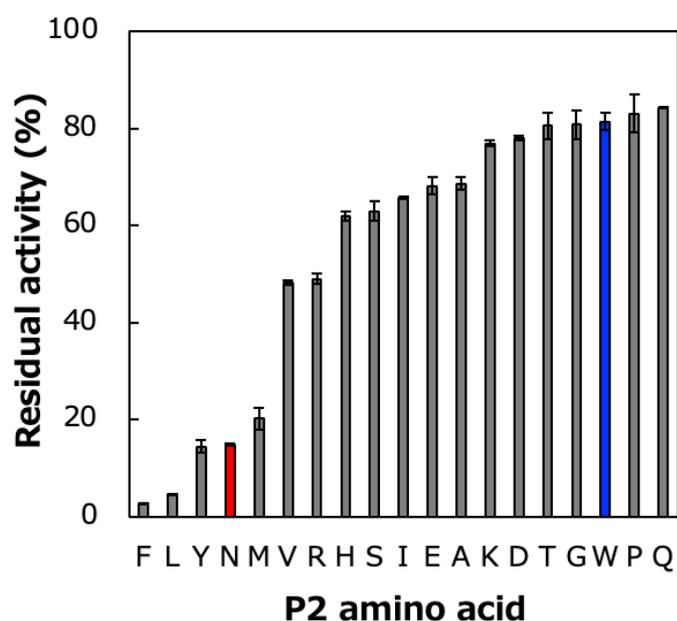


Figure 2-15 | Effect of Xaa-Tyr dipeptides on the hydrolytic activity of SmDPP7 towards synthetic substrate Tyr-Tyr-MCA.

Inhibition assay was performed as described in Materials and Methods. Residual activity was measured under conditions where $200 \mu\text{M}$ Xaa-Tyr dipeptide and $200 \mu\text{M}$ Tyr-Tyr-MCA (synthetic substrate) were added into the reaction solution. The outlier dipeptides Asn-Tyr and Trp-Tyr are colored in red and blue, respectively (see text and Figure 2-16A). The standard deviations were obtained from three independent experiments.

Table 2-4 | Inhibition constants (K_i) of Xaa-Tyr / Xaa-Asp dipeptides against the hydrolytic activities on synthetic substrates of S46 peptidases.

P2 amino acid (Xaa)	H.I. of P2 aa	K_i (μM)			
		SmDPP7	PgDPP7	SmDPP11	PgDPP11
Asp	-55	-	-	-	-
Pro	-46*	-	-	-	-
Glu	-31	-	-	-	-
Asn	-28	7.80 \pm 0.42	210 \pm 18	61.9 \pm 6.6	4.06 \pm 0.16
Lys	-23	-	-	-	-
Arg	-14	-	-	23.2 \pm 1.4	-
Gln	-10	-	-	-	-
Ser	-5	-	-	-	-
Gly	0	-	-	-	-
His	8	-	-	-	-
Thr	13	-	-	-	-
Ala	41	-	-	-	-
Tyr	63	7.66 \pm 1.27	-	5.66 \pm 0.23	-
Met	74	11.2 \pm 0.4	-	-	-
Val	76	55.2 \pm 0.5	-	-	-
Trp	97	-	30.6 \pm 0.9	6.27 \pm 0.18	1.36 \pm 0.06
Leu	97	2.39 \pm 0.08	61.0 \pm 1.2	5.68 \pm 0.07	3.35 \pm 0.23
Ile	99	-	-	-	-
Phe	100	1.27 \pm 0.06	150 \pm 2	7.01 \pm 0.27	5.11 \pm 0.06

Xaa-Tyr and Xaa-Asp were used as competitive inhibitors for DPP7s and DPP11s, respectively, and Tyr-Tyr-MCA and Leu-Asp-MCA were used as substrates for DPP7s and DPP11s, respectively. The hydrophobicity indexes (H.I.) of the amino acids were adopted from Sereda *et al.*, 1994 and Mohera *et al.*, 1995.^{112,113} *The hydrophobicity index of proline is normalized from Sereda *et al.*, 1994 under the condition at pH 2.0. “-“ means not determined due to low inhibitory activity. Standard deviations were obtained from three independent experiments.

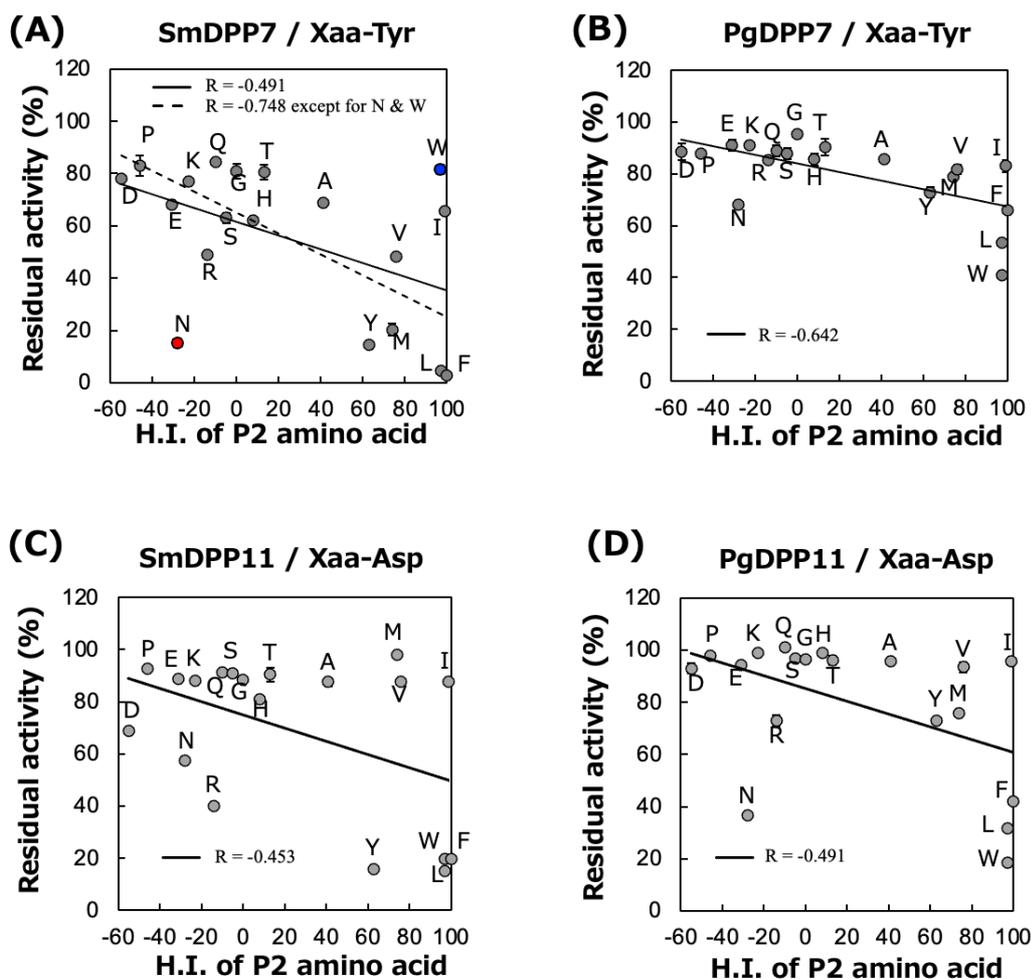


Figure 2-16 | Correlation diagrams between the hydrophobicity index (H.I.) of P2-amino acid and residual DPP activities of S46 peptidases. (A) SmDPP7. The outlier dipeptides Asn-Tyr and Trp-Tyr are colored in red and blue, respectively (see text). (B) PgDPP7. Tyr-Tyr-MCA hydrolysis of PgDPP7 was inhibited by Xaa-Tyr dipeptide. Residual activity was measured under conditions where 200 μ M dipeptides and 200 μ M Tyr-Tyr-MCA were added into the reaction solution. (C) DPP11 from *S. maltophilia* (SmDPP11). (D) PgDPP11. For DPP11s, Leu-Asp-MCA hydrolysis was inhibited by Xaa-Asp dipeptide. Residual activity was measured under conditions where 100 μ M dipeptides and 100 μ M Leu-Asp-MCA were added into the reaction solution. The hydrophobicity indices of the amino acids were adopted from Sereda *et al.*, 1994 and Mohera *et al.*, 1995.^{112,113} R is the correlation coefficient. The standard deviations were obtained from three independent experiments.

2.3.4. Structure determination of SmDPP7

2.3.4.1. The overall structure of SmDPP7

Crystal structures of SmDPP7 in complexes with dipeptides Val-Tyr, Phe-Tyr, Tyr-Tyr, and Asn-Tyr were determined at resolutions of 2.03 to 1.86 Å (Table 2-2 and Table 2-3). Representative electron density maps of the bound dipeptides are shown in Figure 2-12. The SmDPP7 enzyme forms a homodimer, with each subunit consisting of 697 residues (Ala23-Ala719) and a molecular weight of approximately 155 kDa (Figure 2-17). A dimer

of SmDPP7 is situated in the asymmetric unit (Figure 2-17A). Two protomers of SmDPP7 are related by a non-crystallographic two-fold axis and form a dimer. Dimerization have also been observed in the crystal structures of other S46 peptidases, PmDAP BII and PgDPP11 S46 peptidases are the first exopeptidases.³⁵⁻³⁷ The protruding β -hairpin dimerization interface reported in the above S46 peptidases is also observed for SmDPP7. Each subunit contains a catalytic double β -barrel domain harboring the Asp-His-Ser catalytic triad (Figure 2-17B, top domain) and an α -helical domain that caps the active site (Figure 2-17B, bottom domain). The assignment of the secondary structural elements is provided in Figure 2-18.

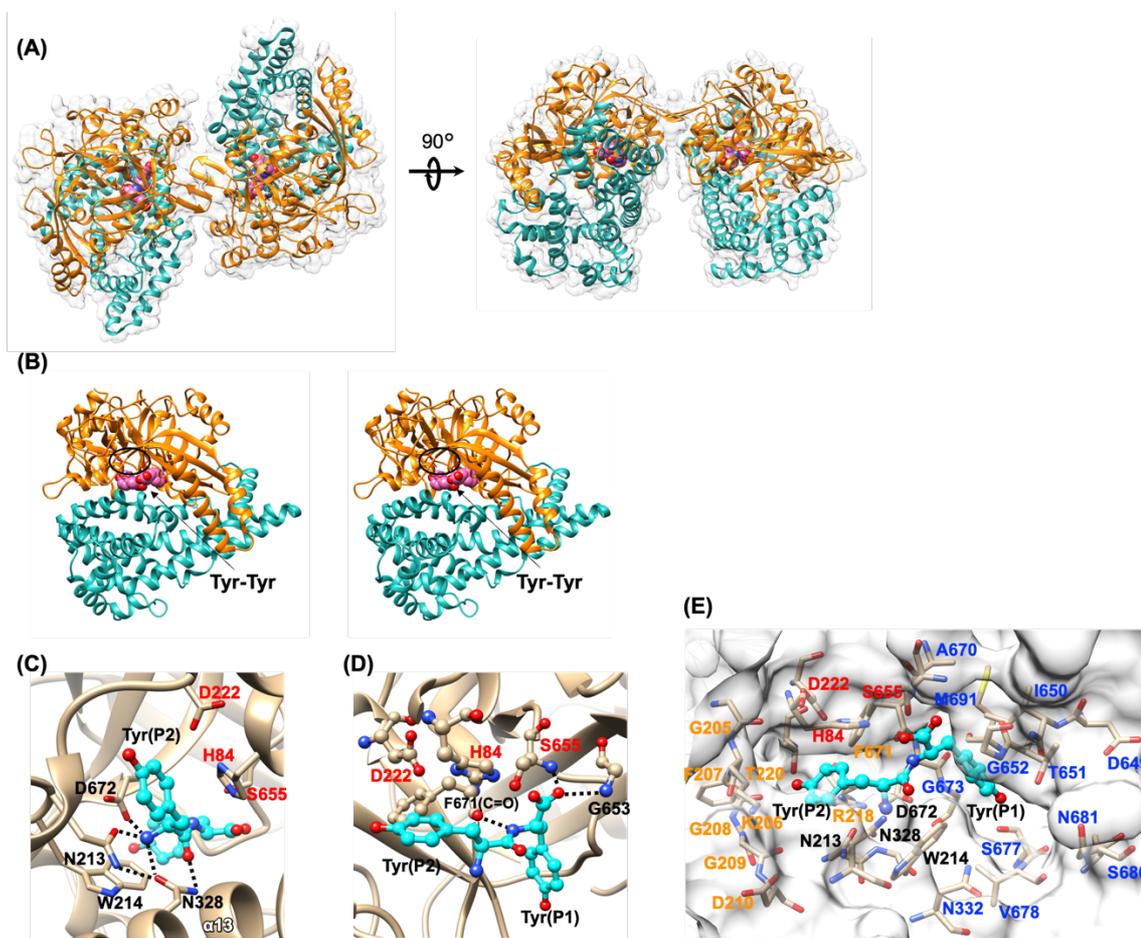


Figure 2-17 | Three-dimensional structure of SmDPP7 complexed with Tyr-Tyr dipeptide. (A) Three-dimensional structure of dimeric SmDPP7 complexed with Tyr-Tyr dipeptide (PDB, 7DKC). Orange and sea green represent catalytic domain (residues 23-274 and 572-720) and α -helical domain (residues 275-571), respectively. The bound Tyr-Tyr dipeptide is colored in pink. (B) Wall-eye stereo view of monomeric SmDPP7. The catalytic triad “Asp222-His84-Ser655” (red) is marked by an ellipsoid. The other color codes are the same as in (A). (C) N-terminus recognition. Catalytic residues and N-terminal recognition residues are represented in red and black, respectively. (D) Oxyanion hole. Catalytic residues are represented in red. (E) S1 and S2 subsite. S1 subsite residues and S2 subsite residues are represented in blue and orange, respectively. The other color codes are the same as in (C). These figures were produced using the program UCSF Chimera version 1.14.¹¹⁴

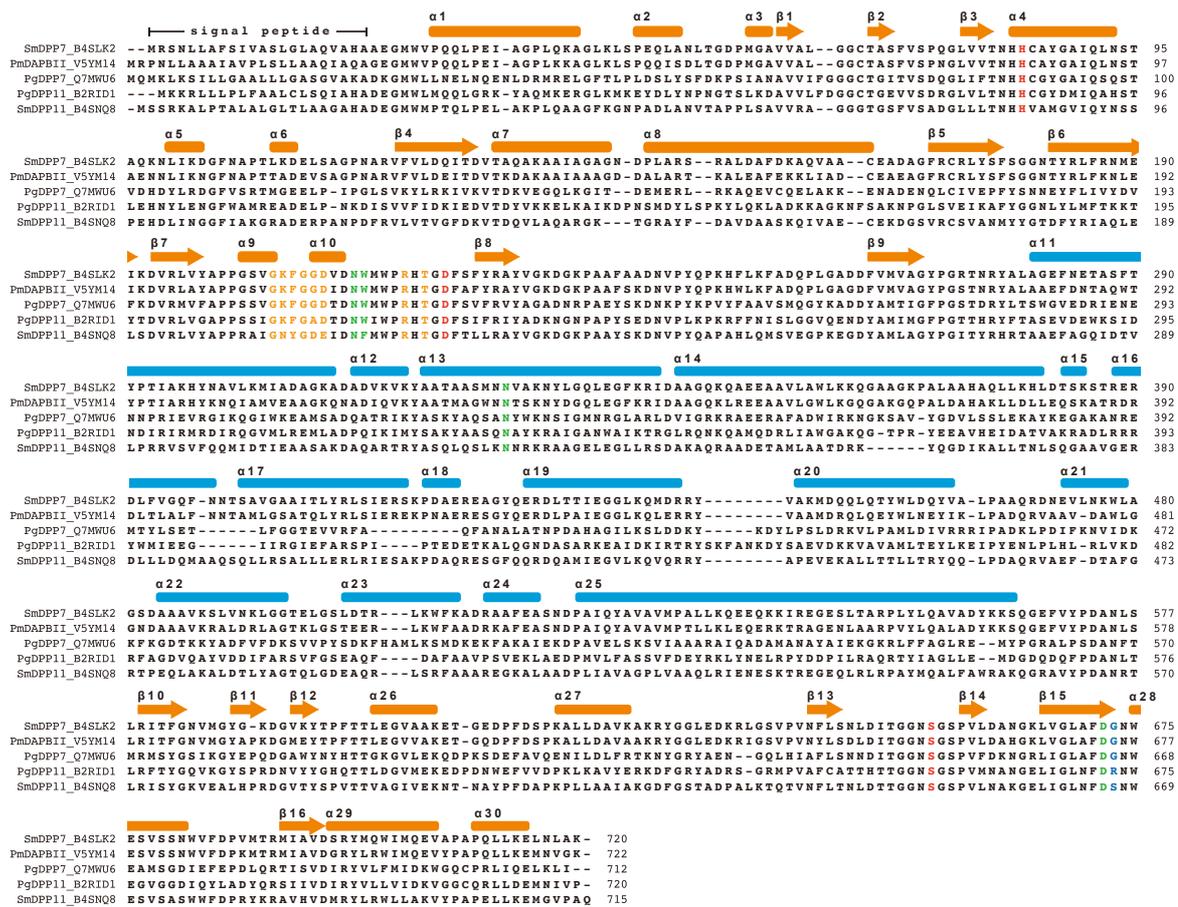


Figure 2-18 | Amino acid sequences of S46 peptidases mentioned in this study. The characters following the enzyme names are the UniProt accession numbers. Secondary structural elements of SmDPP7 are shown on top. α -helices (α 1- α 30) and β -sheets (β 1- β 16) are shown in round square and arrow, respectively. Colors represent different domains: orange is the catalytic domain and blue is the helical domain. Catalytic residues, N-terminus recognition residues, and S2 subsite residues are shown in red, green, and orange, respectively. Blue represents the P1 recognition residues corresponding to position 673 amino acid of PgDPP11.

The catalytic domain includes residues 23–274 and 572–720 and contains a double β -barrel structure. The β -barrel structure is a characteristic of the chymotrypsin superfamily. The catalytic domain of SmDPP7 can be superimposed on those of PmDAP BII and PgDPP11 (Figure 2-19A). The serine peptidase catalytic triads, His84, Asp222 and Ser655 in SmDPP7 and His86, Asp224 and Ser657 in PmDAP BII, are almost completely superimposable, with a root mean square (rms) deviation between the two structures of 0.355 Å for 397 structurally equivalent C α atoms that had 84.0% sequence identity for that region. Similarly, the rms deviation between the catalytic domains of SmDPP7 and PgDPP11 is 1.01 Å for 308 structurally equivalent C α atoms that had 41.9% sequence identity for that region. Although the sequence identity of the catalytic domains between SmDPP7 and PgDPP11 is low, the catalytic triad (His, Asp, and Ser) of both enzymes can

be superposed (Figure 2-19A, right-side).

The α -helical domain is inserted between strands β 9 and β 10 of the catalytic domain and spans residues 275–571 (Figure 2-18). The domain consists of 15 helices packed into a helical bundle that caps the catalytic triad of the catalytic domain. No structural homologue of this domain is found in the DALI database, except for the α -helical domains of the other structurally characterized S46 peptidases, PmDAP BII and PgDPP11.^{35,36} Thus, the α -helical domain is absolutely restricted to S46 peptidases. The α -helical domain of SmDPP7 can be superimposed on that of PmDAP BII and PgDPP11 (Figure 2-19B). The rms deviation between the α -helical domains of SmDPP7 and PmDAP BII is 0.636 Å for 279 structurally equivalent C α atoms across all 297 pairs, which had 68.5% sequence identity for that region. The rms deviation between the α -helical domains of SmDPP7 and PgDPP11 is 1.38 Å for 66 structurally equivalent C α atoms across all 285 pairs, which had 19.1% sequence identity for that region. In particular, the α -helical domain of PgDPP11 exhibits different inter-domain orientation (Figure 2-19B, blue) as compared with those of SmDPP7 and PmDAP BII (Figure 2-19B, right). Thus, we found that the structure of the α -helical domain of SmDPP7 is similar to that of PmDAP BII but is difficult to superpose onto that of PgDPP11.

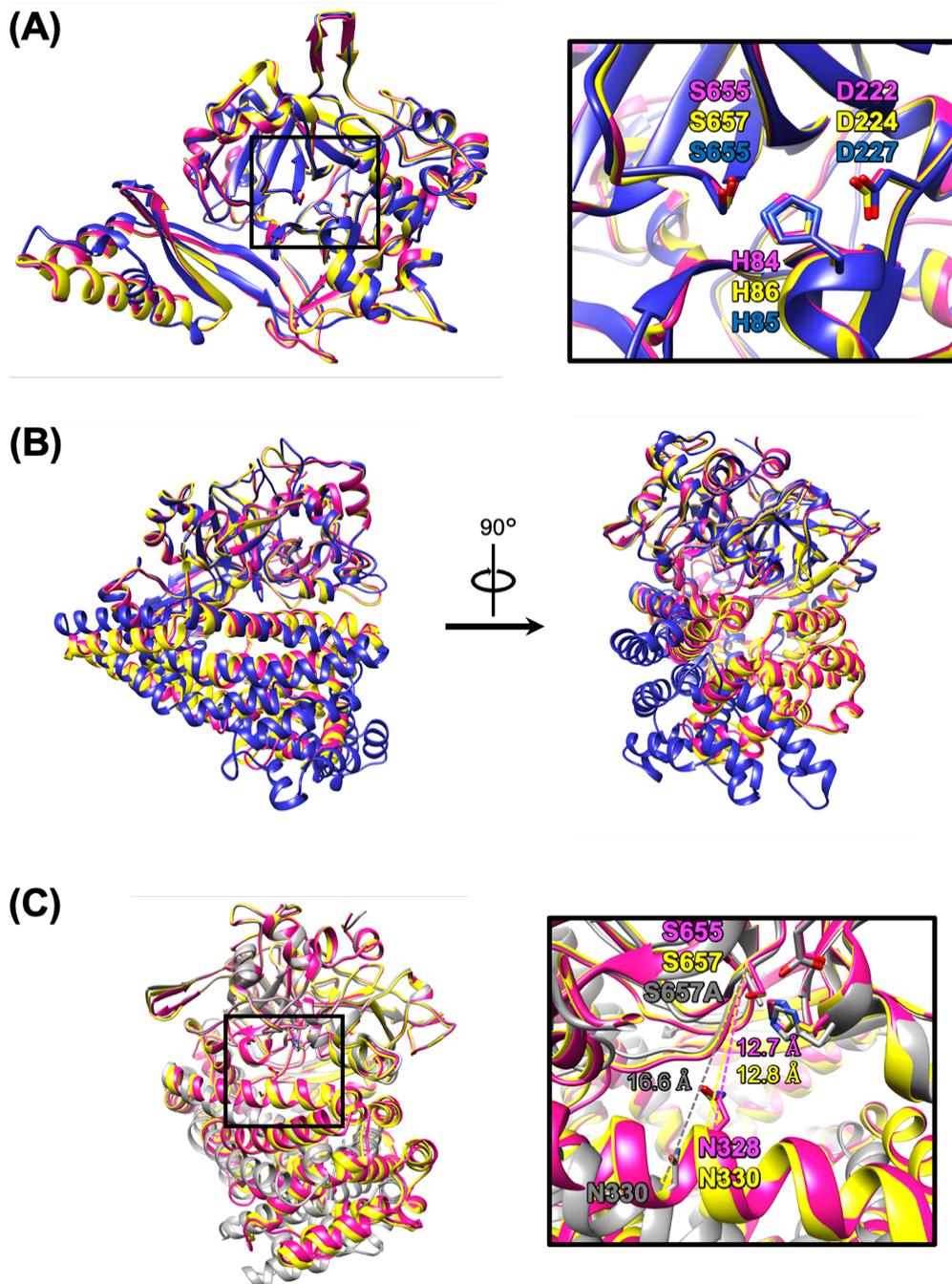


Figure 2-19 | Superpositions of the protomers of SmDPP7, PmDAP BII, and PgDPP11. (A) Superposition of catalytic domains. Ribbon diagrams of SmDPP7 complexed with Tyr-Tyr dipeptide (PDB, 7DKC), PmDAP BII complexed with Val-Tyr dipeptide (PDB, 3WOL), and PgDPP11 complexed with citrate ion (PDB, 6JTB) are shown in magenta, yellow, and blue, respectively. Residues 23–274 and 572–720 of SmDPP7 are shown as a catalytic domain. Residues 25–276 and 574–720 of PmDAP BII and residues 22–279 and 572–720 of PgDPP11, corresponding to the catalytic domain of SmDPP7, are shown. (B) The superposition of the protomers of SmDPP7 (PDB, 7DKC), PmDAP BII (PDB, 3WOL), and PgDPP11 (PDB, 6JTB). The color coding is the same as (A). (C) The superposition of the protomers of SmDPP7 and PmDAP BII. Ribbon diagrams of SmDPP7 (PDB, 7DKC), PmDAP BII (PDB, 3WOL; closed conformation), and PmDAP BII (PDB, 3WOK; open conformation) are shown in magenta, yellow, and grey, respectively. These figures were produced using the program UCSF Chimera version 1.14.¹¹⁴

2.3.4.2. Dipeptide complexes

The four dipeptide complexes clearly show the molecular basis for peptide recognition mechanism at the S2 and S1 sites of SmDPP7. For simplicity, the following description refers primarily to subunit A of the 1.86-Å-resolution structure of the Tyr-Tyr complex of SmDPP7. The bound dipeptide was found in the active site cleft of the catalytic domain (Figure 2-17B) and was covered by the α -helical domain. SmDPP7 hydrolyses peptides from the N-terminus of oligopeptides, cleaving the dipeptide units (NH₂-P2-P1-COOH) when the second P1 residue is a hydrophobic amino acid. To act as a dipeptidyl aminopeptidase, SmDPP7 must fix the N-terminus of the substrate peptide in position. The N-terminal amino group recognition residues of SmDPP7 are Asn213, Trp214, and Asp672 from the catalytic domain and Asn328 from the α -helical domain (Figure 2-17C). For PmDAP BII, a large-scale conformational change, from open to closed, was observed upon peptide binding.³⁵ The distance between the C α atoms of Asn330 belonging to the α -helical domain and catalytic Ser657 belonging to the catalytic domain of PmDAP BII observed in the peptide-free, open conformation (PDB ID: 3WOK, Figure 2-19C, grey) was 16.6 Å, whereas that observed in the dipeptide-bound, closed conformation (PDB ID: 3WOL, Figure 2-19C, yellow) was 12.8 Å. The corresponding distance between the C α atoms of Asn328 and Ser655 of SmDPP7 observed in the current dipeptide-bound conformation was 12.7 Å (Figure 2-19C, pink). This result suggests that the present structure of SmDPP7 corresponds to the closed conformation of PmDAP BII and that the active site cleft of SmDPP7 is closed upon peptide binding. The catalytic triad of SmDPP7 is composed of His84, Asp222, and Ser655. The hydroxy group of Ser655 is hydrogen bonded to the imidazole group of His84 (OG(Ser655)---NE2(His84): 2.9 Å). One of the oxygen atoms of the carboxy group of Asp222 forms a hydrogen-bond with His84 and completes the catalytic triad (ND1(His84)---OD2(Asp222): 2.7 Å). The oxyanion hole is formed by the backbone amide nitrogen atoms of Ser655 and Gly653, and the backbone NH group of P1 residue of the bound peptide is recognized by a hydrogen bond with the carbonyl oxygen of Phe671 (Figure 2-17D).

The S1 subsite of SmDPP7 is observed adjacent to the catalytic Ser655 and the oxyanion hole (Figure 2-17E). The S1 subsite consists of Asn335, Asp649, Ile650, Thr651, Gly652, Ala670, Gly673, Ser677, Val678, Ser680, Asn681, and Met691. Gly673 in SmDPP7, corresponds to Arg673 in PgDPP11, which is a crucial residue for the P1 residue specificity/preference of S46 peptidases (relatively loose preference of DPP7 and Asp/Glu

specificity of DPP11),^{34,36,38} and is located in the wall of the S1 subsite. Thus, the S1 subsite of SmDPP7 is large and deep enough to accommodate any amino acid as the P1 residue of the substrate peptide. The aromatic ring of the bound Tyr(P1) has hydrophobic contacts with the side chains of Trp214 and Ile650. The hydroxy group of the bound Tyr(P1) points toward the bottom of the S1 subsite. There are four buried water molecules at the bottom of the S1 subsite and they constitute a hydrogen bond network with the main-chain and side-chain atoms in the S1 subsite. The hydroxy group of the bound Tyr(P1) is involved in the hydrogen bond network of the water molecules (not shown in Figure 2-17E).

The S2 subsite of SmDPP7 composed of His84, Gly205(C=O), Lys206, Phe207(C=O), Gly208, Gly209, Asp210, Arg218, Thr220, Asp222, and Phe671 is sufficiently wide enough to accommodate a bulky side chain (Figure 2-17E). The side chain of bound Tyr(P2) is accommodated in the S2 subsite and has aliphatic contacts with the side chains of His84, Lys206, and Phe671 (Figure 2-20A). The hydroxy group of Tyr(P2) forms hydrogen bonds to the carbonyl oxygens of Phe207 and Gly209, and the side chain of Arg218 via a water molecule. Thus, specific interaction between the S2 subsite and the side chain of Tyr(P2) is limited, whereas the main-chain atoms of Tyr(P2) are tightly fixed by the side chains of Asn213, Trp214, Asn328, and Asp672 as described above. This is consistent with the lack of a strict specificity of SmDPP7 for the P2 residue.

For the Val-Tyr complex, some differences were observed in the intermolecular interactions around the S2 subsite as compared to those for the Tyr-Tyr complex, while intermolecular interactions between the side chain of Tyr(P1) and the residues in the S1 subsite were almost identical between the Val-Tyr and Tyr-Tyr complexes (not shown in Figure 2-20B). The unique features of the Val-Tyr complex are that the bottom space of the S2 subsite is occupied by four water molecules (Figure 2-20B) and the side chain of Lys206 is less ordered. In the Tyr-Tyr complex, one of the four water molecules at the bottom of the S2 subsite in the Val-Tyr complex is replaced by the hydroxy group of Tyr(P2) and the side chain of Lys206 is well ordered due to aliphatic contacts with the side chain of Tyr(P2). For the Phe-Tyr complex, intermolecular interactions around the active site are almost conserved as compared with that of the Tyr-Tyr complex. The side chain of Lys206 is well ordered and the five water molecules observed at the bottom of the S2 subsite of the Tyr-Tyr complex are also observed for the Phe-Tyr complex (Figure 2-20C). For the Asn-Tyr complex, an extensive hydrogen-bond network was formed at the bottom

of the S2 subsite (Figure 2-20D). Interestingly, a pentagonal hydrogen-bond network consisting of HOH35, HOH63, HOH67, HOH1112, and the ND2 atom of P2-Asn was observed (Figure 2-21). Thus, the exceptional Asn(P2) preference of SmDPP7 could be explained by the water-mediated hydrogen-bond network at the bottom of the S2 subsite of SmDPP7.

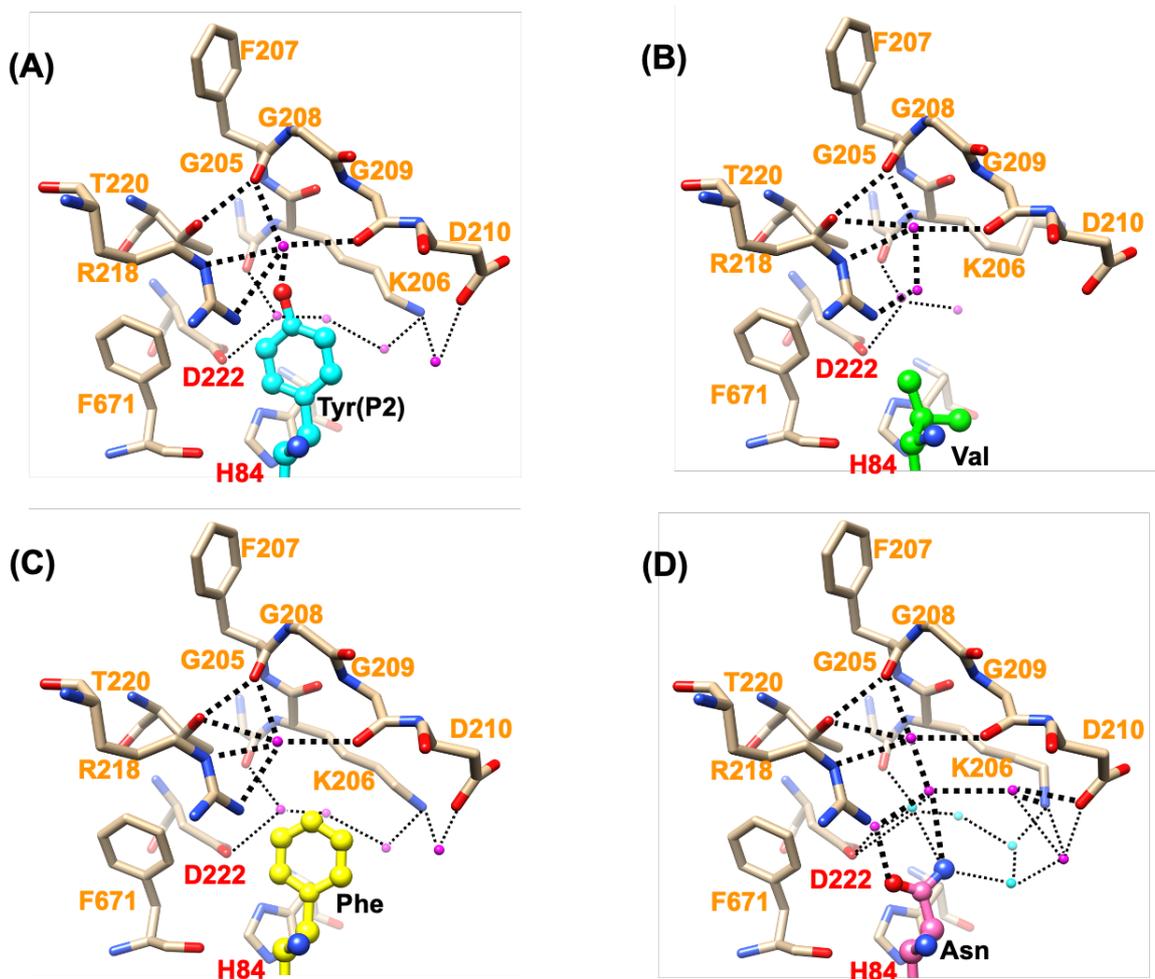


Figure 2-20 | The S2 subsite of SmDPP7/dipeptide complexes. Catalytic residues and S2 subsite residues are denoted in red and orange, respectively. Magenta and cyan spheroidal denote a water molecule. (A) Tyr-Tyr dipeptide (cyan) complex at a 1.86-Å resolution. (B) Val-Tyr dipeptide (green) complex at a 2.03-Å resolution. (C) Phe-Tyr dipeptide (yellow) complex at a 1.91-Å resolution. (D) Asn-Tyr dipeptide (pink) complex at a 1.92-Å resolution. Cyan spheroids denote a water molecule associated with a pentagonal hydrogen-bond network consisting of HOH35, HOH63, HOH67, HOH1112, and the ND2 atom of P2-Asn (see Figure 2-21). These figures were produced using the program UCSF Chimera version 1.14.¹¹⁴

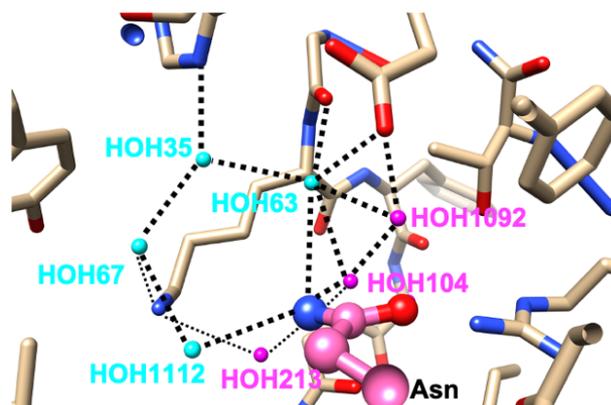


Figure 2-21 | A pentagonal hydrogen-bond network observed in the S2 subsite of SmDPP7/Asn-Tyr complex. Magenta and cyan spheroidal denote water molecules. Cyan spheroids represent water molecules associated with a pentagonal hydrogen-bond network consisting of HOH35, HOH63, HOH67, HOH1112, and the ND2 atom of P2-Asn. Asn-Tyr dipeptide complex at a 1.92-Å resolution (PDB 7DKD). This figure was produced using the program UCSF Chimera version 1.14.¹¹⁴

2.3.5. Site-directed mutagenesis studies on residues in the S2 subsite of SmDPP7

To test the roles of the residues located in the S2 subsite for P2 residue recognition by SmDPP7, we replaced the following residues with alanine and analysed the enzymatic activities of the mutant enzymes on synthetic substrates, Gly-Tyr-MCA, Tyr-Tyr-MCA, and Asn-Tyr-MCA (Table 2-5). The mutated residues and their estimated roles were: Lys206, Arg218, Thr220, and Phe671 for interaction with the P2 side chain. The Arg218 to Ala mutant (R218A) showed complete loss of activity, whereas the F671A mutant resulted in a significant loss of activity (approximately 7 – 20% of the k_{cat}/K_m value of the wild-type enzyme). For the F671A mutant, specificity constant (k_{cat}/K_m) values for the Gly-Tyr-MCA, Tyr-Tyr-MCA, and Asn-Tyr-MCA substrates were significantly decreased. This result indicates that Phe671 in SmDPP7 plays an important role in the accommodation of the substrate peptide into the catalytic domain. The complete loss of the activity of R218A indicates that this residue is important for fixing the carboxy group of Asp672 in position, which is involved in the recognition of the N-terminus of the substrate peptide. The R218Q and R218K mutants retain slight dipeptidyl peptidase activities. Comparison of R218Q and R218K mutants revealed that R218K shows higher activity for all the dipeptidyl substrates examined than R218Q. The length of the Gln side chain of the R218Q mutant appears to be insufficient to fix the side chain of Asp672. The T220A mutant showed moderate loss of activity, whereas the K206A mutant retained activity comparable to that of the wild-type

enzyme. For the T220A mutant, k_{cat}/K_m values for all the substrates examined decreased to approximately 16–45% of the k_{cat}/K_m value of the wild-type enzyme, which is attributed to the decreased k_{cat} value. It is noteworthy that k_{cat}/K_m values of the K206A mutant for Tyr-Tyr-MCA and Asn-Tyr-MCA were significantly changed from that of the wild-type enzyme as compared with that for Gly-Tyr-MCA. Considering that Gly-Tyr-MCA possessed no side chain at the P2 position, the change in k_{cat}/K_m values for Tyr-Tyr-MCA and Asn-Tyr-MCA of K206A suggested that the Lys206 residue is involved in an interaction with the P2 side chain.

Table 2-5 | Kinetic parameters of wild-type and mutant SmDPP7s toward each synthetic substrate.

Variation	Substrate	K_m (μM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($\text{sec}^{-1} \text{mM}^{-1}$)
Wild type	Gly-Tyr-MCA	183 ± 4	0.702 ± 0.022	3.85 ± 0.03
	Tyr-Tyr-MCA	81.7 ± 3.4	3.83 ± 0.09	46.9 ± 1.2
	Asn-Tyr-MCA	51.3 ± 1.6	3.67 ± 0.12	71.5 ± 0.3
K206A	Gly-Tyr-MCA	201 ± 10	0.751 ± 0.015	3.73 ± 0.13
	Tyr-Tyr-MCA	55.8 ± 1.5	3.52 ± 0.04	63.2 ± 1.5
	Asn-Tyr-MCA	54.1 ± 2.1	3.47 ± 0.02	64.4 ± 2.6
R218A	Gly-Tyr-MCA	-	-	-
	Tyr-Tyr-MCA	-	-	-
	Asn-Tyr-MCA	-	-	-
R218Q	Gly-Tyr-MCA	717 ± 63	0.0365 ± 0.0018	0.0510 ± 0.0020
	Tyr-Tyr-MCA	275 ± 11	0.0403 ± 0.0004	0.146 ± 0.005
	Asn-Tyr-MCA	390 ± 14	0.198 ± 0.006	0.509 ± 0.020
R218K	Gly-Tyr-MCA	586 ± 40	0.0238 ± 0.0011	0.0407 ± 0.0017
	Tyr-Tyr-MCA	251 ± 6	0.240 ± 0.003	0.955 ± 0.012
	Asn-Tyr-MCA	277 ± 6	0.183 ± 0.008	0.660 ± 0.017
T220A	Gly-Tyr-MCA	215 ± 10	0.454 ± 0.010	2.12 ± 0.06
	Tyr-Tyr-MCA	46.4 ± 2.4	1.82 ± 0.07	39.2 ± 0.7
	Asn-Tyr-MCA	66.6 ± 1.5	2.60 ± 0.04	39.0 ± 0.3
F671A	Gly-Tyr-MCA	688 ± 38	0.191 ± 0.009	0.278 ± 0.003
	Tyr-Tyr-MCA	463 ± 88	1.54 ± 0.18	3.36 ± 0.23
	Asn-Tyr-MCA	17.4 ± 1.4	0.230 ± 0.012	13.7 ± 0.4

Kinetic parameters were determined by fitting the experimental data to the Michaelis–Menten equation. “-” means not determined due to low activity. Standard deviations were obtained from three independent experiments.

2.3.6. Thermodynamic characterization of dipeptide bindings in SmDPP7

Crystal structure analysis of SmDPP7 complexed with Asn-Tyr indicated that a hydrogen bond network consisting of water molecules and the side chain of bound Asn contributes to the exceptional P2-Asn preference in the S2 subsite of SmDPP7, as described above. The thermodynamic parameters of dipeptide bindings in SmDPP7 were determined using isothermal titration calorimetry (ITC) (Figure 2-22, Figure 2-23, and Table 2-6). Here, we analyzed bindings of five dipeptides, Asn-Tyr, Tyr-Tyr, Val-Tyr, Leu-Tyr, and Phe-Tyr. The bindings with all of the dipeptides examined in SmDPP7 were an enthalpy-driven process with an unfavorable entropic contribution. Because S46 peptidases involve conformational changes upon dipeptide binding, which reduce the degrees of freedom of the polypeptide chain (Figure 2-17C),³⁵ the transition from open to closed state is entropically unfavorable. For Val-Tyr, thermodynamic parameters were not determined due to low-avidity ($K_d > 40 \mu\text{M}$).

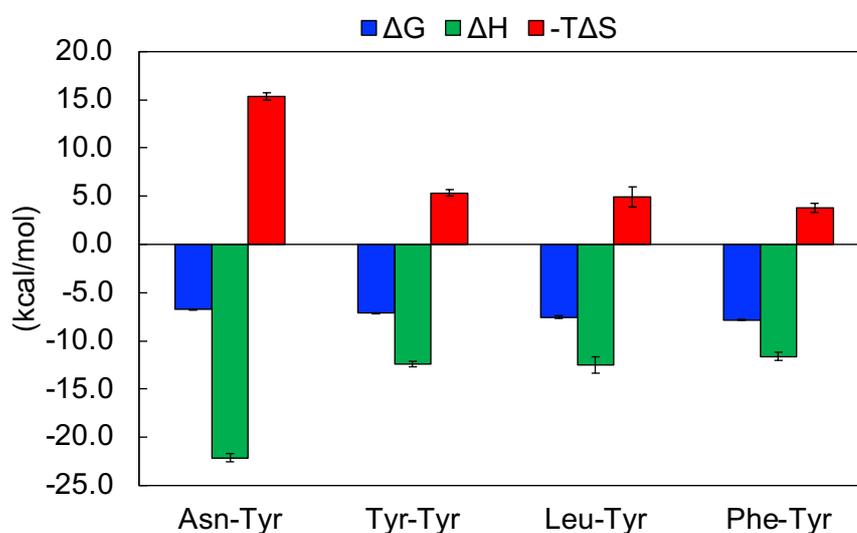


Figure 2-22 | Thermodynamic parameters of dipeptide binding of SmDPP7. The calorimetric data for the respective dipeptide bindings are shown in Figure 2-23. The value of thermodynamic parameters is represented in Table 2-6. The dissociation constant (K_d) and heats of binding (ΔH) were obtained using MICROCAL PEAQ-ITC Analysis software (Malvern, UK). Gibbs free energy (ΔG) and entropy energy (ΔS) were calculated according to the equation, $\Delta G = -RT \ln K_a = RT \ln K_d$ ($K_a = 1/K_d$, association constants). Standard deviations were obtained from three independent experiments.

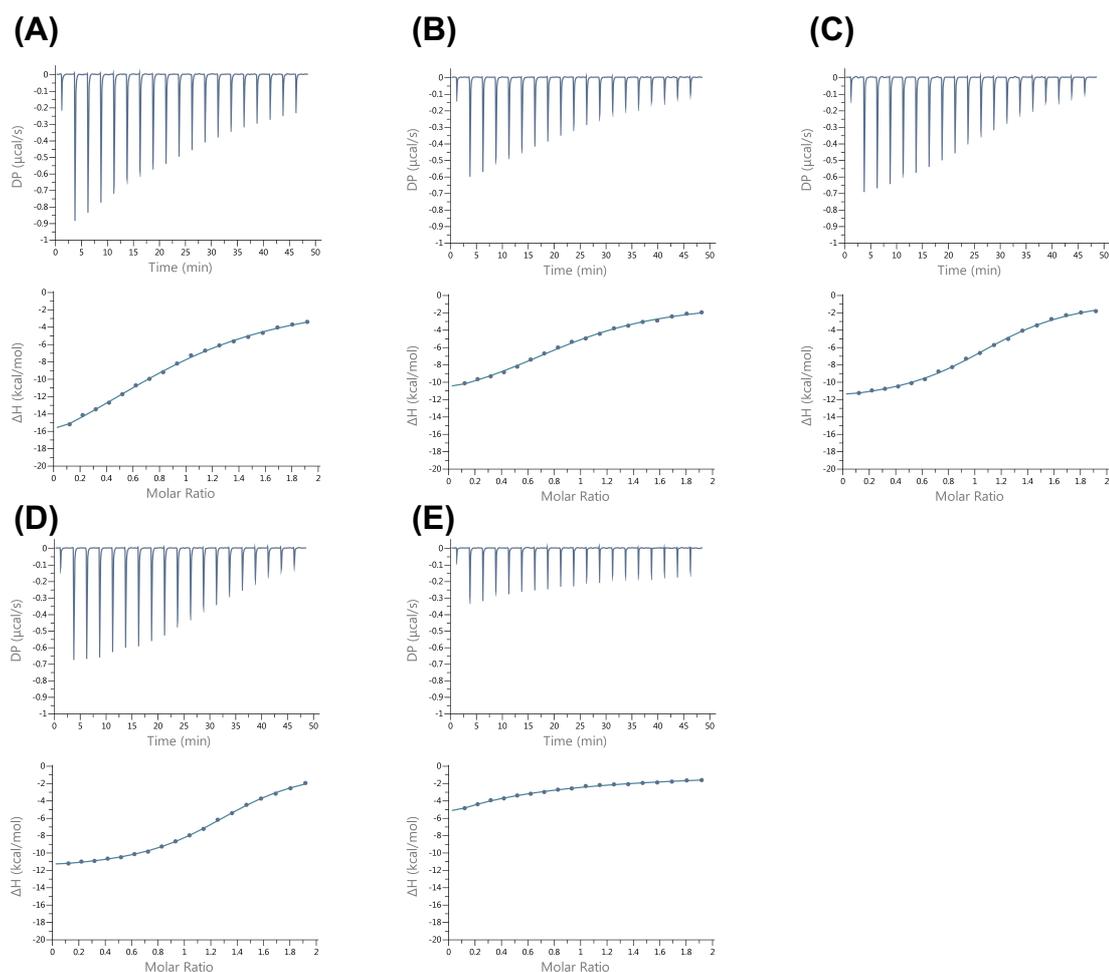


Figure 2-23 | Isothermal titration calorimetry (ITC) analyses of dipeptide binding to SmDPP7. Raw data of ITC experiment depicting differential power (DP) as a function of time (top), and the graphs of fitted data using a fitting model, one-site binding model (bottom). The diagram and value of thermodynamic parameters are represented in Figure 2-22 and Table 2-6, respectively. Data from one of the three measurements are shown. (A) Asn-Tyr dipeptide. (B) Tyr-Tyr dipeptide. (C) Leu-Tyr dipeptide. (D) Phe-Tyr dipeptide. (E) Val-Tyr dipeptide.

Table 2-6 | Thermodynamic parameters of dipeptide binding to SmDPP7.

Ligand	H.I. of P2 a.a.	K_d (μM)	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
Asn-Tyr	-28	11.3 ± 0.2	-22.1 ± 0.5	15.3 ± 0.4	-6.70 ± 0.01
Tyr-Tyr	63	6.17 ± 0.39	-12.4 ± 0.3	5.32 ± 0.34	-7.11 ± 0.04
Leu-Tyr	97	3.01 ± 0.78	-12.5 ± 0.9	4.91 ± 1.02	-7.55 ± 0.15
Phe-Tyr	100	1.90 ± 0.23	-11.6 ± 0.4	3.79 ± 0.52	-7.81 ± 0.07
Val-Tyr*	76	> 40	-	-	-

The calorimetric data for the respective dipeptide bindings are shown in Figure 2-23. The diagram of thermodynamic parameters is represented in Figure 2-22. Standard deviations were obtained from three independent experiments. *The thermodynamic parameters were not determined due to low affinity (see Figure 2-23E).

2.3.7. P2 position preference of S46 peptidases

To clarify whether the P2-preferences at the S2 subsite of S46 peptidases are conserved, we analyzed P2 residue preference for other S46 peptidases, PgDPP7, SmDPP11 (DPP11 from *S. maltophilia*), and PgDPP11 (Figure 2-16B, C, and D). The correlation coefficients between the residual activities of the peptidases and the H.I. of the P2 residue of dipeptides used as competitive inhibitors were -0.642, -0.453, and -0.491, for PgDPP7, SmDPP11, and PgDPP11, respectively. The S46 peptidases showed a preference for hydrophobic amino acids at the P2 position of the substrate, this tendency was particularly obvious for DPP7s. The preferences of Trp(P2) for SmDPP7, and Met(P2) and Arg(P2) for SmDPP11 were out of trend and exceptional. Interestingly, Asn(P2) preference was conserved among the S46 peptidases analyzed. A dipeptide harboring Asn at the P2 position (Asn-Tyr or Asn-Asp) inhibited PgDPP7, PgDPP11, and SmDPP11 with K_i values of 210, 4.06, and 61.9 μM , respectively (Table 2-4). These results indicated that substrate recognition mechanisms at the S2 subsite are partly conserved between DPP7s and DPP11s.

2.4. Discussion

In this study, we evaluated the P2 residue preference of DPP7 from *S. maltophilia* (SmDPP7), which is distinct from mammalian DPP7s that belong to the S28 peptidase family, to better understand the determinants for P2 residue specificity/preference of S46 peptidases. We solved the crystal structures of SmDPP7 in the presence of four kinds of dipeptides, Val-Tyr, Phe-Tyr, Tyr-Tyr, and Asn-Tyr. SmDPP7 showed a preference for bulky hydrophobic amino acids (except for Trp) and Asn at the P2 position of the substrate (Figure 2-15, Figure 2-16A, and Table 2-4). It should be noted that inhibition assays with dipeptides consisting of the same P1 amino acids reflect P2 preference of DPPs. The preference for bulky hydrophobic amino acids at the P2 position is reasonable considering the structure of the S2 subsite, which is sufficiently wide enough to accommodate the large side chain (Figure 2-17E). Preference for hydrophobic residues at the P2 position by SmDPP7 is in agreement with previous reports.^{90,111} Rouf *et al.* reported that PgDPP7 and PgDPP11 exhibited a preference for hydrophobic amino acids at the P2 position of various dipeptidyl substrates and that the Phe664 (PgDPP7 numbering) is involved in the recognition of P2 hydrophobic amino acids.⁹⁰ In this study, the F671A mutant of SmDPP7 (which corresponds to the Phe664 in PgDPP7) showed a remarkably decreased specificity constant toward dipeptidyl substrates with hydrophobic and hydrophilic amino acids at the P2 position (Table 2-5). This indicated that the side chain of Phe671 (numbering in SmDPP7) appears to be important for maintaining the conformation of the S2 subsite rather than for the recognition of P2 hydrophobic amino acids. However, the specificity for Asn(P2) residue by DPPs has never been reported for bacterial DPPs or mammalian DPPs.³³ To confirm whether the preference for bulky hydrophobic amino acids and Asn at the P2 position was conserved in S46 peptidases, we performed a multiple sequence alignment analysis against 4931 genes coding S46 peptidases. We observed that the residues in the S2 subsite are highly conserved among S46 peptidases as compared with those in the S1 subsite (Table 2-7), indicating that S46 peptidases would share a common preference at the P2 position of the substrate. Indeed, P2 preference for bulky hydrophobic amino acids and Asn are conserved among S46 peptidases examined in this study (Figure 2-16). In contrast, the specificity at the P1 position of the substrate is distinct between DPP7 and DPP11 in S46 peptidases according to the lower conservation in S1 subsite residues (Figure 2-14 and Table 2-7). SmDPP7 and PgDPP7 showed a preference for P1-hydrophobic amino acids and an ability to degrade the dipeptidyl substrate with a basic

amino acid at the P1 position, whereas SmDPP11 and PgDPP11 prefer the acidic amino acids at the P1 position. This indicates that the S2 subsite can be an influential target to develop the universal inhibitor of S46 peptidases.

Table 2-7 | Conservation of S2 subsite residues in S46 peptidases.

Residues of SmDPP7		Amino acid residue	Residue conservation (%)
S2 subsite	His83	His	100
	Gly205	Gly	89.6
		Ala	10.1
		Other	0.3
	Lys206	Lys	68.7
		Asn	11.5
		Other	19.8
	Phe207	Phe	78.9
		Tyr	20.6
		Other	0.5
	Gly208	Gly	99.9
		Other	0.1
	Gly209	Gly	70.9
		Ser	10.2
		Other	18.9
	Asp210	Asp	79.3
		Glu	19.8
		Other	0.9
Arg218	Arg	98.8	
	Other	1.2	
Thr220	Thr	79.2	
	Other	20.8	
Asp222	Asp	100	
Phe671	Phe	99.3	
	Other	0.7	
S1 subsite	Asn332	Asn	49.6
		Lys	21.6
		Other	28.8
	Asp649	Asp	80.8
		His	18.8
		Other	0.3
	Ile650	Ile	56.3
		Thr	36.4
		Other	7.3
	Gly652	Gly	98.1
		Other	1.9
	Ala670	Ala	48.5
		Asn	32.3
		Other	19.2
	Gly673	Gly	60.8
Arg		25.5	
Ser		12.4	
Other		1.4	
Ser677	Ser	43.3	

		Ala	30.1
		Gly	24.7
		Other	1.8
	Val678	Leu	29.0
		Val	20.8
		Met	20.2
		Thr	16.5
		Ile	12.3
		Other	1.1
	Ser680	Gly	46.1
		Ser	36.3
		Other	17.6
	Asn681	Asp	76.9
		Other	23.1
	Met691	Thr	24.5
		Ala	21.3
Asn		16.2	
Cys		15.6	
Ser		15.4	
Other		7.0	
N-ter. recognition	Asn213	Asn	99.9
		Other	0.1
	Trp214	Trp	77.1
		Phe	18.8
		Other	4.1
	Asn328	Asn	99.5
		Other	0.5
	Asp672	Asp	99.8
Other		0.2	

Amino acid sequence variations were analyzed for 4931 peptidases belonging to the S46 family. Amino acid sequence of PmDAP BII (UniProt accession numbers: V5YM14) was used as a query sequence. BLAST (Basic Local Alignment Search Tool) was used for searching the amino acid sequences of S46 peptidases. Multiple sequence alignment was conducted with the *Muscle* program and CLC genomics workbench 11 application. Residues located in the S2 subsite, S1 subsite, and N-terminus recognition of SmDPP7 are shown on the left. Dark green, green, and light green mean $\geq 80\%$, $80\% > 60\%$, and $\leq 60\%$ of conservation, respectively. “Other” represents the sum of the residues with conservation degree of under 10%.

The structures of four dipeptide complexes and results of ITC experiments clarify the dipeptide binding of SmDPP7 (Figure 2-20 and Figure 2-22). The binding modes of all dipeptide (reaction products) to SmDPP7 were enthalpy driven and were stabilized through the hydrogen bonds for N-terminus recognition by the side chains of Asn213, Trp214, Asp672 and Asn328 (Figure 2-17C) and P1-NH group recognition by the carbonyl oxygen of Phe671 (Figure 2-17D). In addition, an open-to-closed conformational change could occur for SmDPP7 when it recognizes dipeptides as observed in PmDAP BII (Figure 2-19C), which has 78.4% amino acid sequence identity with that of SmDPP7. Generally, recognition of a ligand molecule by strong hydrogen bonds concomitant with large-scale conformational change is mainly enthalpy driven and unfavorable in entropy.¹¹⁵ In addition,

the loss of conformational degrees of freedom is also entropically unfavorable for the bound ligand molecule. This is the case for SmDPP7 — it recognizes dipeptides by a tight hydrogen bond network (Figure 2-17C and D) associated with open-to-closed conformational change. The order of dissociation constants (K_d) of dipeptide (P2 a.a.; $V > N > Y > L > F$) was consistent with the result of competitive inhibition assay (Figure 2-15 and Table 2-6). Among the dipeptide complexes, observations, such as strong enthalpic interactions of SmDPP7 with Asn-Tyr can be explained by an extensive hydrogen-bonding network with water molecules (Figure 2-20D and Figure 2-22). On the other hand, trapped water molecules in the S2 subsite are not favorable in entropy, so the Asn-Tyr binding exhibits highly unfavorable entropic contribution in ITC data (Figure 2-22 and Table 2-6). Bindings with Tyr-Tyr, Leu-Tyr, and Phe-Tyr showed favorable enthalpic and unfavorable entropic contributions. The magnitudes of both parameters are smaller than those of Asn-Tyr binding. These thermodynamic parameters are attributed to hydrogen bonds for the dipeptide binding associated with open-to-closed conformational change. For the binding of Tyr-Tyr, Leu-Tyr, and Phe-Tyr, the ejecting water molecules from the S2 subsite as an entropy reservoir is entropically favorable. Thus totally unfavorable entropic contributions of the bindings of Tyr-Tyr, Leu-Tyr, and Phe-Tyr are explained by the fact that the favorable entropic contribution by the elimination of water molecules in the S2 subsite is smaller than the unfavorable entropic contributions by the open-to-closed conformational change and the loss of flexibility of the bound ligand molecule, which is at the same level for the four dipeptide used in this study. Indeed, ejecting water molecules were observed for Tyr-Tyr and Phe-Tyr complexes as compared with Asn-Tyr complex (Figure 2-20). Binding with Leu-Tyr, which has a hydrophobic side chain at the P2 position, could also eject water molecules from the S2 subsite. In order to clarify the detail of entropic contribution, it is necessary to measure the heat capacity changes (ΔC_p) and deconvolute total binding entropy (ΔS_{tot}). Although the structure of the Tyr-Tyr complex showed hydrogen bonds to the carbonyl oxygens of Phe207 and Gly209 and the side chain of Arg218 via a water molecule (Figure 2-20A), the enthalpic contribution showed little difference as compared with Phe-Tyr binding. This suggested that these hydrogen bonds do not significantly contribute to Tyr(P2) binding. To summarize, the dipeptide binding mode focused on the S2 subsite of SmDPP7, the P2-Asn is recognized by the establishment of hydrogen-bonding network at the bottom of the S2 subsite, and the bulky hydrophobic amino acids are accommodated in the S2 subsite by the hydrophobic interaction associated

with the ejection of water molecules.

For DPP11 from *Porphyromonas endodontalis* (PeDPP11), the Leu-Asp (product) binding mode is energetically favorable both in enthalpy and entropy and the enthalpic contribution is dominant.³⁷ The binding mode includes a hydrogen-bond network involving the N-terminus and acidic residue(P1) recognition, and the ejection of solvent molecules from the inter-domain cleft concomitant with domain motion.³⁷ For the entropic contribution in the Leu-Asp binding of PeDPP11, the unfavorable entropic contribution concomitant with the structural change is completely offset by the favorable entropic contribution attributed to the ejecting water molecules, and the total entropic contribution is favorable. Although the substrate main chain recognition mechanism of PeDPP11 is similar to that of SmDPP7, the thermodynamic parameters observed for dipeptide binding in ITC measurements are distinct from that of SmDPP7. For SmDPP7, the unfavorable entropic contribution in dipeptide binding indicated that an unfavorable entropic contribution by conformational changes is more dominant than the favorable entropic contribution by ejecting water molecules from inter-domain cleft as an entropy reservoir. It is possible to classify DPP7 and DPP11 according to the entropic contribution in substrate binding other than the specificity at the P1 position of the substrate. The DPP7 type of S46 peptidases show an unfavorable entropic contribution that is attributed mainly to an open-to-closed conformational change, and the DPP11 type of S46 peptidases show a favorable entropic contribution that is attributed to ejecting solvent molecules from the inter-domain cleft.

As described above, a universal inhibitor of S46 peptidases is desirable to optimize for the S2 subsite. In this study, we observed that the decrease in dissociation constant values correlated with the hydrophobicity of the residue at the P2 position (Table 2-6). This indicates that the replacement of water molecules at the S2 subsite contributes to increasing the binding affinity to ligands associated with favorable enthalpic and entropic contributions. Enthalpically optimized inhibitors avoid bacteria with emerging resistance caused by mutation of targeted protein.^{116,117} For NFGNB, mutational resistance to most antibiotics classes can arise easily.¹¹⁸ Given these reports, compounds with favorable enthalpic contribution are convenient for the development of NFGNB antibiotics targeting S46 peptidases. In addition, the bacterial proton-dependent oligopeptide transporter (POT), a transport protein existing on the inner membrane, reportedly has substrate specificity against hydrophobic di- or tri-peptides.^{119,120} We infer that hydrophilic compounds might

be effective in inhibiting S46 peptidases that generally exist in the periplasmic space. Therefore, a compound that replaces water molecules in the S2 subsite with low $\log P$ values may be a valuable universal inhibitor of S46 peptidase for antibiotics of NFGNB such as *S. maltophilia* and *P. gingivalis*.

In this study, we discovered the exceptional bonding of Asn as the P2 amino acid residue to SmDPP7, a serine peptidase from the family-S46, and unraveled the constituents of its S2 and S1 subsites, which could be largely responsible for the substrate recognition mechanisms of the S46 peptidases. Asn(P2) is recognized by a hydrogen-bonding network, and hydrophobic residues are accepted by hydrophobic interactions associated with ejecting water molecules from the S2 subsite. Our findings contribute toward the development of a dipeptidyl universal inhibitor of S46 peptidases, which could potentially serve as NFGNB antibiotics.

Chapter 3 Investigation of an ultra-high-throughput detection method for family-S46 peptidase.

3.1. Introduction

In droplet-based microfluidic systems, the enzyme assays are typically performed inside nanoliter-sized or picoliter-sized water-in-oil droplets (WODLs).^{73,75,121} To date, various enzymes have been detected/screened such as amylase, glucosidase, esterase, lipase, cellulase, and laccase.^{75–81} However, the scope of application of WODL to droplet-based microfluidic systems is limited to some enzyme activities. This is because the most significant feature of the WODL, which is not found in other enzyme assay methods, is the existence of the oil phase surrounding the nanoliter or picoliter of aqueous solution. This feature is sometimes disadvantageous in detecting enzyme activity in a WODL, especially when used with fluorogenic substrates, which are generally used to detect enzyme activity in a WODL with fluorescence-activated droplet sorting (FADS).^{79,122} Because the hydrophobic fluorogenic substrate leaks from the aqueous solution to the oil phase and diffuses throughout WODLs, the evaluation of enzyme activity is difficult.^{123–125} Hence, the fluorogenic substrate used to detect enzyme activity in a WODL needs hydrophilicity. To prevent compound leakage from a WODL, methods, such as addition of hydrophilic functional groups (phosphate, carboxyl, and sulfo groups) to compounds^{123,124,126} and alteration of types and concentrations of surfactants,^{125,127} have been investigated.

Coumarins are commonly used as fluorescent probes of the substrate to detect enzyme activity.¹²⁸ Unsubstituted coumarins show almost no fluorescence, but when an electron-donating group is introduced at position 7, they exhibit strong luminescence.¹²⁹ In fact, by introducing a hydroxy group or amino group at position 7, they can be bound to various substrates such as amino acids, lipids, sugars, and other molecules.^{128,130} 7-Amino-4-methylcoumarin (AMC), with the addition of the amino group at position 7, has blue fluorescence and has been widely used as a probe for the fluorogenic substrate, in particular protease activities also S46 peptidases.¹²⁸ But, a previous study by Woronoff *et al.* showed that AMC leaked into the oil phase and could not be used in a WODL.¹²⁴ They added a sulfo group to AMC to prevent compound leakage from WODL and succeeded in detecting the acylase activity of *E. coli* in the WODL.¹²⁴ Therefore, coumarin derivatives retained in the WODL can facilitate high-throughput detecting/screening of enzyme using droplet-based microfluidic systems. Here, we attempted to develop an AMC-derived

substrate retained in a WODL that can be used in droplet-based microfluidic systems and demonstrated the availability of an AMC-derived substrate to detect/screen for DPP activities, S46 peptidases.

3.2. Materials and Methods

3.2.1. Materials, bacterial strains, and medium

A fluorescent nucleic acid probe for bacterial sorting (FNAP-sort),¹²² which is an RNA probe labeled with Alexa 488 at the 5' end and with Black Hole Quencher1 (BHQ1) at the 3' end was synthesized by Eurofins Genomics (Japan). 7-Amino-4-methyl-3-coumarinylacetic acid (AMCA-H), 7-amino-4-trifluoromethylcoumarin (AFC), and sulforhodamine B were purchased from FUJIFILM Wako Pure Chemical Corporation (Japan). AMC and ACA were purchased from Peptide Institute, Inc. (Japan) and Fluorochem Ltd. (UK), respectively. 7-Aminocoumarin-4-methanesulfonic acid (ACMS) was synthesized as described in the supplementary information. L-Methionyl-L-leucyl-7-aminocoumarin-4-acetic acid (Met-Leu-ACA), L-methionyl-L-leucyl-7-aminocoumarin-4-methanesulfonic acid (Met-Leu-ACMS), and L-leucyl-L-aspartyl-7-aminocoumarin-4-acetic acid (Leu-Asp-ACA) were synthesized in this study (see details in the Supporting Information). L-Methionyl-L-leucyl-7-amino-4-methylcoumarin (Met-Leu-AMC) and L-leucyl-L-aspartyl-7-amino-4-methylcoumarin (Leu-Asp-AMC) were purchased from Peptide Institute, Inc. (Japan). Two types of bacteria were used in this study: *Escherichia coli* DH5 α (Takara Bio Inc., Japan) derived from the K12 strain was transfected with a plasmid (p005-RFP-strong, Addgene, MA, USA) containing a red fluorescent protein (RFP)- coding gene and used as a DPP-non-producing bacterium. *Pseudoxanthomonas mexicana* WO24, isolated by W.O.,¹³¹ was used as a DPP-producing bacterium. *Tofu* (soybean curd) factory waste fluid was sampled from *Imai-tofu-ten* at Nagaoka city, Niigata, Japan. The casitone medium comprised: 1% (w/v) Bacto Casitone, 0.2% (w/v) Bacto Yeast extract, and 4 mM MgSO₄. The LB agar medium comprised: 1% (w/v) Bacto Tryptone, 0.5% (w/v) Bacto Yeast extract, 1% (w/v) NaCl, and 1.5% (w/v) agar.

3.2.2. Materials and general procedures for substrate synthesis

Reagents and solvents used were purchased from BLD Pharm (CN), FUJIFILM Wako Pure Chemical Co., Ltd. (Japan), Nacalai Tesque (Japan), and Watanabe Chemical Industries, Ltd. (Japan), and used without further purification. All synthetic compounds were purified by preparative HPLC with > 95% purity and identified by Time-of-Flight Mass Spectrometry (TOF-MS). Analytical HPLC was performed using a C18 reversed-phase column (4.6 \times 250 mm; Cosmosil 5C₁₈-AR-II) with binary solvent systems: a linear gradient of CH₃CN 10–90% in 0.05% aqueous TFA for 40 min at a flow rate of 1.0 mL/min, detected at 220 nm. Preparative HPLC was carried out on a C18 reversed-phase column (20 \times 250 mm; Cosmosil 5C₁₈-AR-II) with a binary solvent system: a linear gradient of CH₃CN in 5 mM aqueous hydrochloric acid with a flow rate of 5.0 mL/min and detection at 220 nm. Proton nuclear magnetic resonance (¹H NMR) spectra were recorded on DPX400 (Bruker, Co., MA, USA) spectrometers with tetramethylsilane (TMS) as an internal standard. Mass spectra with electrospray ionization, with 50% aqueous methanol as the mobile phase, were obtained from a micrOTOF-Q II spectrometer (Bruker, Co., MA, USA).

3.2.3. Synthesis of coumarins and substrates

(Collaborated with Part-time Lecturer Dr. Koushi Hidaka of Kobe University.)

2-(7-(((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-2-oxo-2H-chromen-4-yl)acetic acid (Fmoc-ACA, 1): To a solution of 2-(7-amino-2-oxo-2H-chromen-4-yl)-acetic acid (241 mg, 1.1 mmol) and sodium bicarbonate (92.4 mg, 1.1 mmol) in water (5.5 mL) with acetone was added 9-fluorenylmethyl chloroformate (259 mg, 1.0 mmol). The mixture was stirred overnight at room temperature. After removal of the solvent in vacuo, the residue was added into 1 M hydrochloric acid. The precipitate was filtered and washed with water and hexane to give the crude product. Precipitation with ether, filtration, and additional washing with 1 M hydrochloric acid and hexane yielded compound 1 with the following characteristics: a yellowish-white solid; yield 71%; ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 12.80 (br. s., 1 H), 10.21 (s, 1 H), 7.92 (d, *J* = 7.46 Hz, 2 H), 7.76 (d, *J* = 7.46 Hz, 2 H), 7.62 (d, *J* = 8.68 Hz, 1 H), 7.55 (br. s., 1 H), 7.46–7.34 (m, 5 H), 6.39–6.32 (m, 1 H), 4.57 (d, *J* = 6.36 Hz, 2 H), 4.34 (t, *J* = 6.42 Hz, 1 H), 3.86 (s, 2 H); MS (electrospray ionization (ESI)-TOF) *m/z*: calculated for C₂₆H₂₀NO₆ [M + H]⁺ 442.1285; found 442.1268.

Dipeptidyl ACA substrates: Peptides containing ACA were prepared by conventional solid phase peptide synthesis using Fmoc chemistry. A mixture of 2-chlorotriyl chloride resin (354 mg, 0.54 mmol), 1 (309 mg, 0.70 mmol), *N,N*-diisopropylethylamine (177 μ L, 1.0 mmol) in dry dichloromethane (2 mL) was vortexed

overnight (loading yield, 60%). The resin (173 mg, 0.11 mmol) was treated with 20% piperidine in *N,N'*-dimethylformamide (DMF) for 20 min at room temperature. Fmoc-amino acid (0.54 mmol), 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU) (207 mg, 0.54 mmol), and 2,4,6-collidine (72 μ L, 0.54 mmol) in DMF (1.5 mL) were added to the resin, and vortexed overnight at room temperature. The Fmoc group was deprotected with 20% piperidine in DMF for 20 min. A mixture of 1-hydroxybenzotriazole (HOBt) (50 mg, 0.33 mmol), Fmoc-amino acid (0.33 mmol), and *N,N'*-diisopropylcarbodiimide (DIC) (51 μ L, 0.33 mmol) in DMF (1 mL) was added to the resin and vortexed for 2 h, and the subsequent deprotection was performed with 20% piperidine in DMF. The resins were stirred with trifluoroacetic acid (TFA)/triisopropylsilane (TIS)/H₂O/3,6-dioxo-1,8-octane-dithiol (DODT) (92.5:2.5:2.5:2.5) for peptide 2 or TFA/TIS/H₂O (95:2.5:2.5) for peptide 3 for 2 h at room temperature. Ether precipitation resulted in a crude white solid. Purification using preparative RP-HPLC gave the desired peptides as white powders.

2-(7-((*S*)-2-((*S*)-2-Amino-4-(methylthio)butanamido)-4-methylpentanamido)-2-oxo-2*H*-chromen-4-yl)acetic acid hydrochloride (Met-Leu-ACA, 2): Yield 10.6 mg (19%); purity 98.3%; ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 10.71 (s, 1 H), 8.80 (d, *J* = 7.46 Hz, 1 H), 8.13 (br. s., 3H), 7.81–7.77 (m, 1 H), 7.66–7.76 (m, 1 H), 7.53–7.47 (m, 1 H), 6.38 (s, 1 H), 4.52 (ddd, *J* = 9.63, 7.31, 5.20 Hz, 1 H), 4.00–3.72 (m, 3 H), 2.57–2.53 (m, 1 H), 2.53 (br. s., 1 H), 2.46 (dd, *J* = 3.67, 1.83 Hz, 1 H), 2.40 (d, *J* = 1.22 Hz, 1 H), 2.12–1.92 (m, 5 H), 1.75–1.52 (m, 3 H), 0.94 (t, *J* = 6.85 Hz, 6 H); MS (ESI-TOF) *m/z*: calculated for C₂₂H₃₀N₃O₆S [M + H]⁺ 464.1850; found 464.1833.

(*S*)-3-((*S*)-2-Amino-4-methylpentanamido)-4-((4-(carboxymethyl)-2-oxo-2*H*-chromen-7-yl)amino)-4-oxobutanoic acid hydrochloride (Leu-Asp-ACA, 3): Yield 6.3 mg (12%), purity 95.8%; ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 10.70 (s, 1 H), 8.96 (d, *J* = 7.09 Hz, 1 H), 8.20 (br. s., 3H), 7.82–7.75 (m, 1 H), 7.75–7.64 (m, 1 H), 7.54–7.44 (m, 1 H), 6.38 (s, 1 H), 6.29 (d, *J* = 1.34 Hz, 1 H), 4.80–4.71 (m, 1 H), 3.88 (s, 2 H), 3.77 (dd, *J* = 8.31, 5.99 Hz, 1 H), 2.88–2.81 (m, 1 H), 2.70–2.62 (m, 2 H), 2.52 (br. s., 1 H), 2.40 (d, *J* = 1.22 Hz, 1 H), 1.76–1.49 (m, 4 H), 0.94–0.84 (m, 6 H); MS (ESI-TOF) *m/z*: calculated for C₂₁H₂₆N₃O₈ [M + H]⁺ 448.1714; found 448.1705.

(*S*)-2-((*S*)-2-((*tert*-Butoxycarbonyl)amino)-4-(methylthio)butanamido)-4-methylpentanoic acid (Boc-Met-Leu-OH, 4): To a solution of leucine methyl ester hydrochloride (73.9 mg, 0.4 mmol) in DMF (5 mL), *N,N,N*-triethylamine (55.9 μ L, 0.4 mmol), HOBt monohydrate (67.4 mg, 0.44 mmol), Boc-Met-OH (110 mg, 0.44 mmol), and *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) were added. The mixture was stirred overnight at room temperature. After removal of the solvent in vacuo, the residue was added into ethyl acetate, washed sequentially with 10% citric acid, 5% sodium bicarbonate, and saturated sodium chloride, dried with sodium sulfate, and filtered. Removal of the ethyl acetate in vacuo gave a white solid as an intermediate (137 mg). To a solution of the filtered solid in methanol (2 mL), 1 M NaOH aq. (2 mL) was added. It was stirred for two hours at room temperature. After removal of the methanol in vacuo, the solution was acidified by citric acid, then extracted with ethyl acetate, washed with brine, dried with sodium sulfate, and filtered. After removal of the solvent in vacuo, precipitation with ether gave the titled compound as a white solid (143 mg); yield 143 mg 99%; ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 12.44 (br. s., 1 H), 8.00 (d, *J* = 7.95 Hz, 1 H), 6.94 (d, *J* = 8.19 Hz, 1 H), 4.22 (ddd, *J* = 9.75, 7.98, 5.26 Hz, 1 H), 4.09–3.93 (m, 1 H), 2.77–2.67 (m, 1 H), 2.67–2.57 (m, 1 H), 2.44 (t, *J* = 7.89 Hz, 2 H), 2.03 (s, 3 H), 1.88–1.71 (m, 2 H), 1.71–1.60 (m, 1 H), 1.60–1.43 (m, 2 H), 1.37 (s, 9 H), 0.89 (d, *J* = 6.60 Hz, 3 H), 0.83 (d, *J* = 6.48 Hz, 3 H); MS (ESI-TOF) *m/z*: calculated for C₁₆H₃₁N₂O₅S [M + H]⁺ 363.1948; found 363.1938.

(7-((*S*)-2-((*S*)-2-Amino-4-(methylthio)butanamido)-4-methylpentanamido)-2-oxo-2*H*-chromen-4-yl)methanesulfonic acid hydrochloride (Met-Leu-ACMS, 5): To a solution of intermediate 4 (121 mg, 0.30 mmol) in dry DMF (4 mL), *N*-methylmorpholine (33.0 μ L, 0.30 mmol), and isobutyl chloroformate (38.9 μ L, 0.30 mmol) were added at –10 °C in an ice-salt bath. The mixture was stirred for 30 min. A solution of (7-amino-2-oxo-2*H*-chromen-4-yl)methanesulfonic acid² (69.7 mg, 0.27 mmol) and *N*-methylmorpholine (150 μ L, 1.37 mmol) in dry DMF (1 mL) was added and stirred overnight at room temperature. After removal of the solvent in vacuo, the residue was added into ethyl acetate, washed sequentially with 1M HCl aq. and brine, dried with sodium sulfate, filtered, and evaporated to give an intermediate as a slightly yellow-colored oil. The solid was dissolved in 4M HCl/ethyl acetate (2 mL) with anisole (65.2 μ L, 0.6 mmol) and stirred for one hour at room temperature. After removal of the solvent in vacuo, the residue was added to ether. The precipitate was filtered, dissolved in water, and lyophilized. Purification using preparative RP-HPLC gave the desired peptides as white powders: yield 0.64 mg; purity 96.3%; ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 10.65 (s, 1 H), 8.89 (br. s., 1 H), 7.91 (d, *J* = 8.80 Hz, 1 H), 7.82 (d, *J* = 2.08 Hz, 1 H), 7.42–7.40 (m, 1 H), 6.27 (s, 1 H), 4.63–4.54 (m, 1 H), 4.08–3.93 (m, 2 H), 3.90–3.82 (m, 1 H), 2.62–2.59 (m, 2 H), 2.09–1.87 (m, 5 H), 1.73–1.48 (m, 3 H), 0.94 (d, *J* = 6.11 Hz, 3 H), 0.91 (d, *J* = 6.24 Hz, 3 H); MS (ESI-TOF) *m/z*: calculated for C₂₁H₃₀N₃O₇S₂ [M + H]⁺ 500.1520; found 500.1489.

3.2.4. Overexpression and purification of PmDAPBII

For the plasmid construction, expression, and purification of PmDAPBII, we refer to previous studies.¹³² The mature PmDAPBII was composed of 698 amino acids (residues Gly25 to Lys722), with a theoretical molecular weight of 76251.2 and an isoelectric point (pI) of 6.31. *E. coli* BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, US) transformed with the pET_b3_pmdap2 expression plasmid were grown in TB media at 298 K to an OD₆₀₀ of 0.6. Overproduction of PmDAPBII was induced by adding 0.1 mM IPTG for 15 h at 298 K. Thereafter, the cells were harvested by centrifugation at 6,000 x g. Cells were disrupted using sonication and the cell extract was obtained by centrifuging the lysate at 20,000 x g for 30 min. The PmDAPBII was purified by precipitation with 35 to 70% ammonium sulfate and hydrophobic column chromatography using a HiPrep 16/10 Butyl column (Cytiva, Marlborough, MA, US) (Figure 3-1 A). Gly-Phe-pNA hydrolysis activity was confirmed in the eluted protein peaks, and fraction numbers 48 to 54 were collected. The eluate was desalted using a HiPrep 26/10 desalting column (Cytiva) and finally subjected to anion-exchange column chromatography using a Mono Q 5/50 GL column (Cytiva) (Figure 3-1 B). The fractions containing PmDAPBII (fraction numbers 20 and 25) was pooled, buffer-exchanged to 80 mM Tris/HCl pH 8.5 and concentrated to 5 mg/ml using Vivaspin 20 concentrator (Cytiva). Purity and molecular mass of purified PmDAPBII were estimated by SDS-PAGE (Figure 3-2).

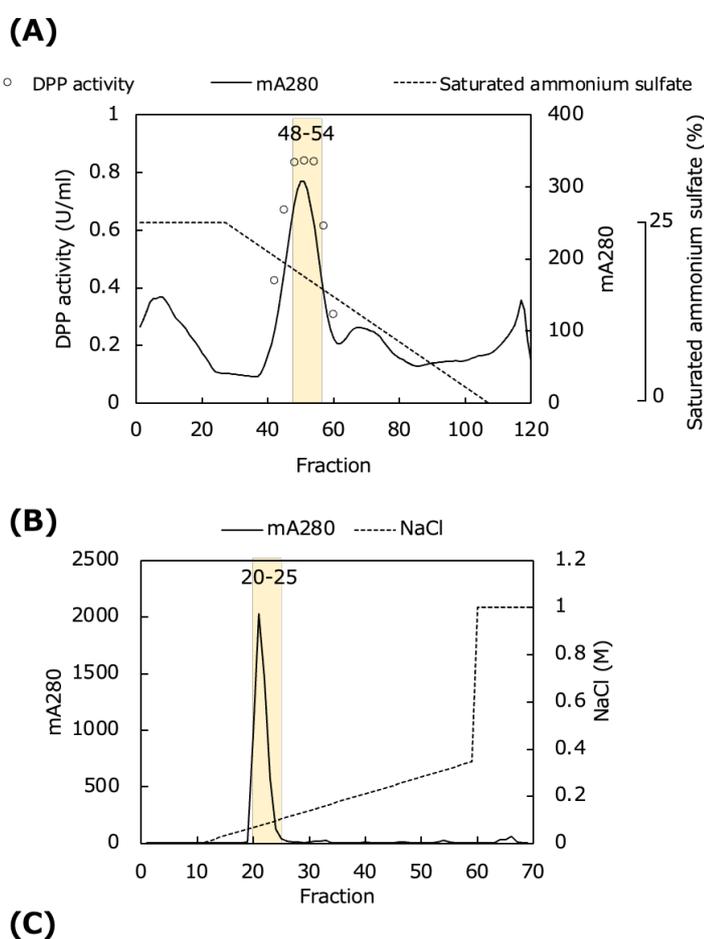


Figure 3-1 | Purification of PmDAPBII
 (A) Hydrophobic interaction chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a linear concentration gradient of 25% to 0% saturated ammonium sulfate. Fractions shown in

yellow were collected. (B) anion exchange chromatogram. 50 mM Tris-HCl pH 9.0 was used as buffer, and proteins were eluted by a linear concentration gradient of 0 M to 0.35 M NaCl. Fractions shown in yellow were collected. (C) Protein amount and DPP activity at each purification step. The Bradford method was used for protein measurement. Gly-Phe-pNA was used as a substrate for measurement of DPP activity. ASP, HIC, and AEC represents ammonium sulfate precipitation, hydrophobic interaction chromatography, anion exchange chromatography, respectively.

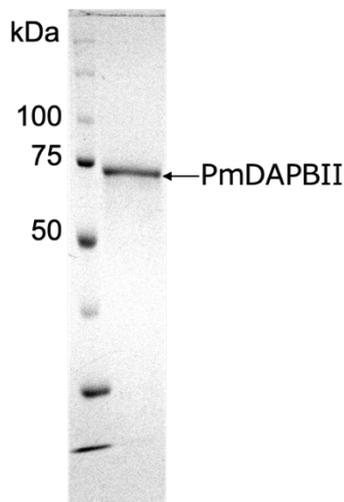


Figure 3-2 | Coomassie-stained 10% SDS-PAGE analysis of purified PmDAPBII 3 μ g protein loaded per lane. Molecular-weight markers of different masses (kDa) are labelled.

3.2.5. Overexpression and purification of SmDPP7, SmDPP11, and PgDPP11

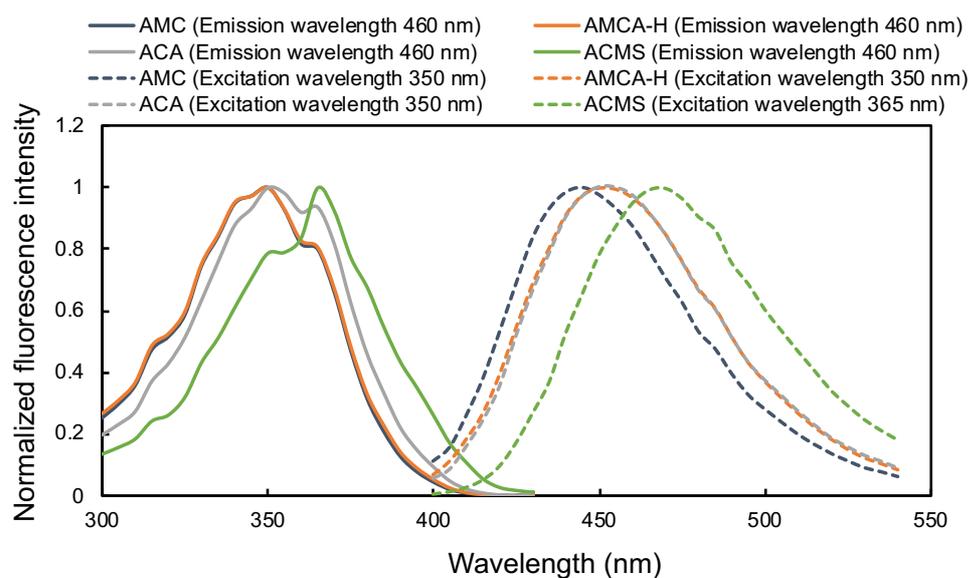
Expression and purification of SmDPP7, SmDPP11, and PgDPP11 were described in section 2.2.3. , 2.2.4. , and 2.2.6. , respectively.

3.2.6. WODL generation and sorting

WODLs with a diameter of about 120 μ m were produced using an On-Chip droplet generator (On-Chip Biotechnologies, Japan) at a rate of about 400,000 droplets/min. The oil phase was HFE-7500 3M Novec Engineered Fluid (HFE-7500) containing 2% (w/w) 008-FluoroSurfactant (RAN Biotechnologies, MA, USA) as a surfactant. FADS was performed using an On-Chip Sort (On-Chip Biotechnologies, Japan) at a rate of about 300 droplets/sec. A quantity of 0.1% (w/w) 008-FluoroSurfactant in HFE-7500 was used as sheath solution.

3.2.7. Retention time characterization in the WODL

Each fluorescent substance was diluted to 100 μ M with 50 mM sodium phosphate buffer containing 5 mM EDTA after being dissolved in dimethyl sulfoxide at 10 mM. The excitation and emission light spectra of each coumarin derivative are shown in Figure 3-3. A WODL with fluorescent substances (positive WODL) and a WODL without fluorescent substances (negative WODL) were prepared and equal volumes of both WODLs were mixed. The negative WODLs contained sulforhodamine B. Blue and red fluorescent intensities were measured by confocal microscope and analyzed by ImageJ, as described below. MiLogP values were calculated by Molinspiration (<http://www.molinspiration.com/cgi-bin/properties>).



Fluorescent substances	Maximum Ex. (nm)	Maximum Em. (nm)
AMC	350	445
ACA	350	450
ACMS	365	470
AMCA-H	350	450

Figure 3-3 | Excitation and emission spectra of coumarins used in this study. The solid and dotted lines represent spectra of excitation light and emission light, respectively. The fluorescence intensity was measured using an Infinite 200 PRO microplate reader (Tecan, Switzerland).

3.2.8. Image analysis

Micrographs were obtained by a laser-scanning confocal microscope system A1 (Nikon, Japan) operated by NIS-Elements software (Nikon, Japan), under 100× magnification. The WODLs were placed into a μ -Slide VI flat microscopy chamber (Ibidi, Germany) prefilled with 0.1% (w/w) 008-FluoroSurfactant in HFE-7500. A fluorescence intensity analysis was performed with ImageJ.⁹⁶ The calibration curve of the fluorescence intensity is shown in Figure 3-4.

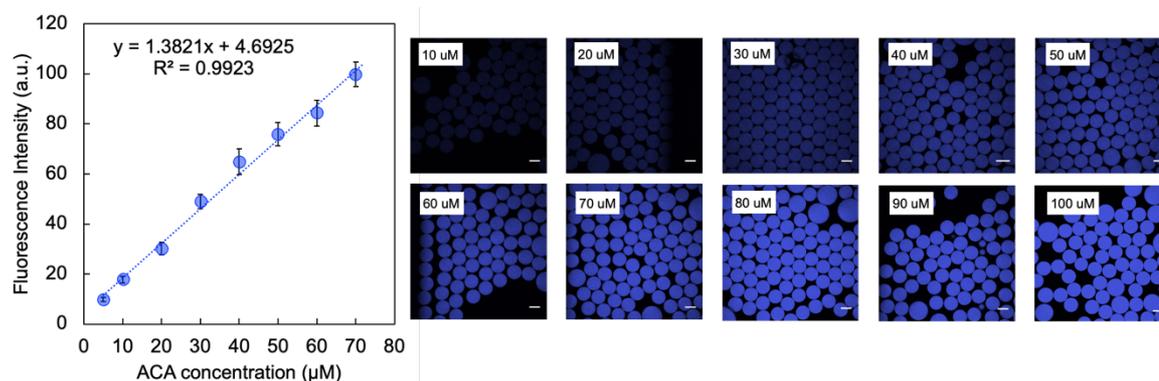


Figure 3-4 | Calibration curve of ACA concentration in image analysis. Droplets containing each ACA concentration were prepared, and the fluorescence intensity was analyzed using ImageJ.⁹⁶ Standard deviations were obtained from 20 droplets. The scale bar represents 100 μ m.

3.2.9. Determination of kinetics parameter toward the dipeptidyl substrate

Kinetic parameters were determined by fitting the experimental data to the Michaelis–Menten equation using Excel Solver (Microsoft, WA, USA) by nonlinear least-squares fitting with various substrate concentrations: Met-Leu-AMC (0.781, 1.56, 3.13, 6.25, 12.5, 25, 50, and 100 μM for 2.5 nM PmDAPBII and 2.5 nM SmDPP7); Met-Leu-ACA (0.195, 0.391, 0.781, 1.56, 3.13, 6.25, 12.5, and 25 μM for 2 nM PmDAPBII, and 0.391, 0.781, 1.56, 3.13, 6.25, 12.5, 25, and 50 μM for 4 nM SmDPP7); Met-Leu-ACMS (0.781, 1.56, 3.13, 6.25, 12.5, 25, 50, and 100 μM for 500 nM PmDAPBII, and 100 nM SmDPP7); Leu-Asp-AMC (0.781, 1.56, 3.13, 6.25, 12.5, 25, 50, and 100 μM for 0.5 nM PgDPP11, and 0.2 nM SmDPP11); and Leu-Asp-ACA (0.781, 1.56, 3.13, 6.25, 12.5, 25, 50, and 100 μM for 0.5 nM PgDPP11, and 0.1 nM SmDPP11). The enzyme reaction was performed in a reaction buffer consisting of 50 mM sodium phosphate buffer pH 7.0, 5 mM EDTA, and 0.005% Tween 20 at 25 °C for 20 min. Standard deviations were calculated from three independent experiments. The fluorescence intensities of the released AMC, ACA, and ACMS were measured with excitation at 355 nm, 350 nm, and 365 nm, respectively, and emission at 460 nm, 450 nm, and 470 nm, respectively, using an Infinite 200 PRO microplate reader (Tecan, Switzerland).

3.2.10. Enzymatic reaction in WODLs

A 50 mM sodium phosphate buffer containing 5 mM EDTA was used as a reaction buffer; 50 nM PmDAPBII was used for 100 μM Met-Leu-ACA hydrolyzed; and 5 nM PgDPP11 was used for 100 μM Leu-Asp-ACA hydrolyzed. These reaction solutions were encapsulated as WODLs using the On-Chip droplet generator. Enzyme reactions were performed at room temperature.

3.2.11. WODL cultivation

The bacterial culture medium and *tofu* factory waste fluid were centrifuged at $6,000 \times g$, and the pellet was washed in 0.9% NaCl solution before suspension in the casitone medium. In both *E. coli* and *P. mexicana*, stochastic encapsulation of microorganisms followed a Poisson distribution⁸⁸ due to the previously investigated relationship between OD_{600} and a Colony Forming Unit (Figure 3-5). Because the number of bacterial cells in the *tofu* factory waste fluid was not determined, bacterial suspension in the casitone medium was directly used for WODL generation. WODLs containing microorganism cells and a 100 μM ACA substrate were statically cultivated at 30 °C for one day.

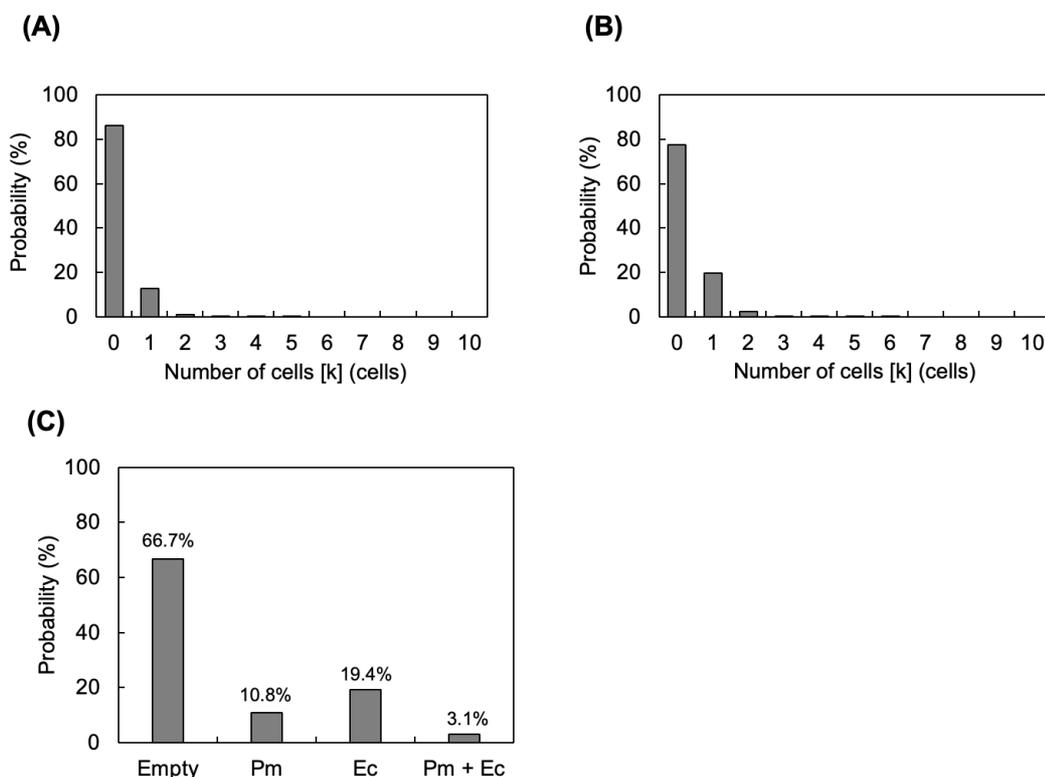


Figure 3-5 | Poisson distribution. 88
 (A) (B) The probability $P(X = x)$ of finding x entities (bacterial cells) was calculated by the equation $P(X = x) = e^{-\lambda} (\lambda^x / x!)$, where e is Napier's constant and λ is the number of bacterial cells in each droplet volume, set at 0.150 for *P. mexicana* WO24 (A) and at 0.255 for *E. coli* (B). (C) The probability of co-encapsulation $P(X_1 = x_1, X_2 = x_2)$ of x_1 cells (*P. mexicana* WO24; Pm) and x_2 cells (*E. coli*; Ec) in a single droplet was calculated by the equation $P(X_1 = x_1, X_2 = x_2) = [e^{-\lambda_1} (\lambda_1^{x_1} / x_1!)] \times [e^{-\lambda_2} (\lambda_2^{x_2} / x_2!)]$. Here, λ was set at 0.125 for *P. mexicana* WO24 and at 0.213 for *E. coli*. 'Empty' refers to an empty droplet with no microorganisms enclosed. 'Pm' is a droplet in which only *P. mexicana* WO24 was enclosed in one or more cells. 'Ec' is a droplet in which only *E. coli* was enclosed in one or more cells. 'Pm + Ec' is a droplet in which *P. mexicana* WO24 and *E. coli* were enclosed in one or more cells of each microorganism.

3.2.12. Preparation of 16s rDNA library

For metagenomic analysis, genomic DNA from each sample except the WODL sample was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals, Santa Ana, CA). Genomic DNA from the WODL sample was extracted by an original method as follows: The WODLs were broken using Pico-Break (Sphere Fluidics Limited, UK) and centrifuged at $6,000 \times g$. The pellet was washed by 0.9% NaCl solution and suspended in 700 μ l of 6 mg/ml lysozyme in a lysis buffer consisting of 100 mM Tris-HCl (pH 8.0), 100 mM EDTA, and 1.5 M NaCl. The suspension was incubated overnight at 37 °C. Next, 10 μ l of 10 mg/mL proteinase K (Nacalai Tesque, Inc., Kyoto, Japan), 10 μ l of 10 mg/ml RNase (Sigma Aldrich, USA), and 70 μ l of 10% (w/v) sodium dodecyl sulfate (Nacalai Tesque, Inc., Kyoto, Japan) were added to the suspension and shaken at 55 °C for 30 min. After centrifuging at $20,000 \times g$, 700 μ l of supernatant was mixed with phenol/chloroform (Nacalai Tesque, Inc., Kyoto, Japan). After centrifugation at $20,000 \times g$, 500 μ l of supernatant was collected in a new Eppendorf tube, and 50 μ l of 3 M CH₃COONa and 1,250 μ l of 100% ethanol were added to the supernatant. The supernatant was removed after centrifugation at $20,000 \times g$, and the pellet was washed by 1 mL of 70% (v/v) ethanol. After centrifugation at $20,000 \times g$, the pellet was dried and dissolved in a TE buffer consisting of 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA. A forward universal bacterial primer 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and a reverse universal primer 806R (5'-GGACTACHVGGGTWTCTAAT-3') were used to amplify the bacterial 16S rRNA genes. Polymerase chain reaction (PCR) was performed using KOD FX Neo (Toyobo Co., Ltd., Osaka, Japan) as a DNA polymerase. Conditions were set as follows: one cycle of 94 °C for 2 min, 40 cycles of 98 °C for 10 sec, and 68 °C for 30 sec. PCR products were purified using a MinElute PCR purification kit (Qiagen, Germany) (Figure S15), and the PCR product concentrations were measured with a Qubit2.0 Fluorometer (Thermo Fisher Scientific, MA, USA). The 16S rRNA gene sequencing was performed as described in the literature.¹³³ DNA was sequenced using a MiSeq Reagent Kit v2 and the MiSeq System (Illumina Inc., CA, USA). For isolated microorganisms, between 27 bp and 1,500 bp of 16S rRNA genes were amplified using forward primer 27F (5'-AGAGTTTGTATCATGGCTCAG-3') and reverse primer 1,500R (5'-TACCTTGTTACGACTT-3'), and each colony was used as a template. PCR was performed using KOD FX Neo (Toyobo Co., Ltd., Osaka, Japan) as a DNA polymerase. Conditions were set as follows: one cycle of 94 °C for 2 min, 35 cycles of 98 °C for 10 sec, 55 °C for 30 sec, and 68 °C for 2 min. PCR products were sequenced by the Sanger method carried out by Eurofins Genomics (Tokyo, Japan).

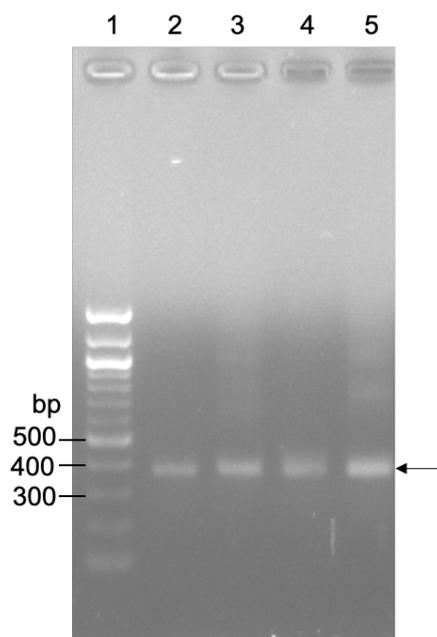


Figure 3-6 | Agarose-gel electrophoresis of an ante-sequencing 16S rDNA sample. A 1.5% agarose gel was used, and 2 μ L of amplified and purified 16S rDNA fragments were applied. Lane 1: 100 bp DNA Ladder (New England Biolabs, Japan). Lane 2: tofu factory waste fluid sample. Lane 3: sample after WODL cultivation. Lane 4: FADS (M-L-ACA activity) sample. Lane 5: FADS (L-D-ACA activity) sample.

3.2.13. Sequencing analysis

The 16s rDNA library preparation, sequencing, and data analysis were carried out based on a previous report.¹³⁴ Metagenomic sequencing data were analyzed with the Quantitative Insights Into Microbial Ecology software (version 1.9.1).¹³⁵ Operational taxonomic units (OTUs) were selected at 97% identity with UCLUST. Taxonomic classification was assigned with the Basic Local Alignment Search Tool (BLAST) based on the Greengenes database, version 13_8. The number of reads after quality filtering and OTUs is shown in Table 3-1. BLAST in the NCBI database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to search for bacterial species corresponding to the 16S rDNA sequence determined by Sanger method sequencing. Sanger sequencing was carried out at Eurofins Genomics (Tokyo, Japan). The metagenomic sequencing data in this study may be obtained from the DNA Data Bank of Japan (DDBJ) database under accession number DRA012262.

Table 3-1 | The number of reads after quality filtering and OTUs.

Sample	Reads (after quality filtering)	OTUs
Tofu factory waste fluid	27 081	534
After WODL cultivation	32 906	349
FADS (M-L-ACA activity)	28 855	252
FADS (L-D-ACA activity)	31 344	318

3.2.14. Device fabrication and operation

A mold of device was made using a 3D printer (Hunter, APPLE TREE, Japan). Then, degassed polydimethylsiloxane (PDMS, Sylgard 184) (Dow, MI, US) solution was poured into the mold and heat-cured (70 °C, 5 h). A detailed design of the device is shown in Figure 3-23. The surfaces of the PDMS and glass slides were treated with oxygen plasma using a plasma cleaner (PDC-32G, Harrick Plasma, NY, US), and then each treated surface was bonded together to fabricate the microfluidic device. After that, WODL was introduced into the channel of the microfluidic device at 10 ul/min using a syringe pump (CX07200, Tokyo Garasu Kikai Japan). The excess WODL remaining in the channel was removed with 10 ul/min of fluorine oil. the WODL trapped in the wells was observed by a laser-scanning confocal microscope system A1 (Nikon, Japan) operated by NIS-Elements software (Nikon, Japan) under 100x magnification.

3.2.15. Graphical programs

Histograms and plots of FADS were produced by FlowJo software, version 10.7.1 (Becton, Dickinson & Company, NJ, USA). Chemical structural formulas were depicted by ChemDraw Std, version 14.0.0.117 (PerkinElmer, MA, USA) and MarvinSketch, version 21.13.0 (ChemAxon, Hungary).

3.3. Results

3.3.1. Examination of leakage of fluorescent substances from the WODL.

Hydrophobic fluorescent substances such as AMC leak from WODLs because of the fluorophore exchange between WODLs.^{123–125} Here, we evaluated four coumarin-derived fluorescent substances, AMC, ACA,¹³⁶ AMCA-H,¹³⁷ and ACMS¹²⁴ to assess their ability to remain in WODL (Figure 3-7).

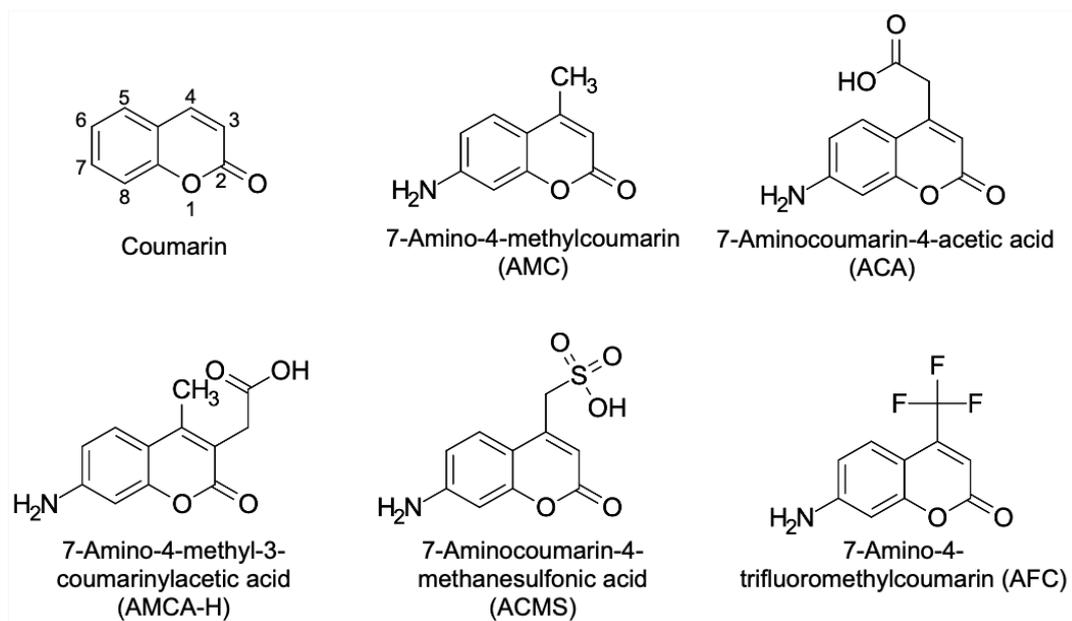
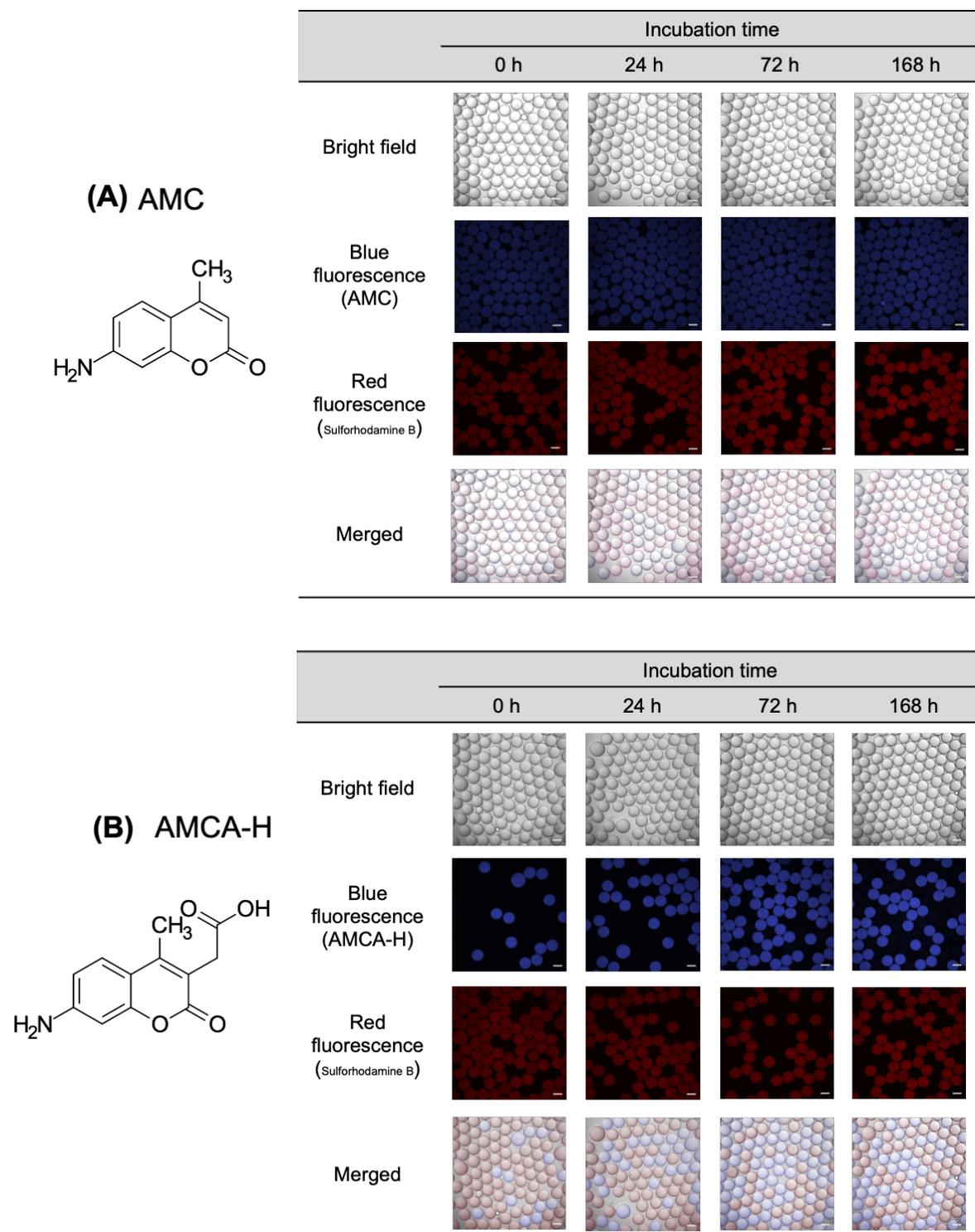


Figure 3-7 | Structural formula of coumarins used in this study.

A WODL with a fluorescent substance (positive WODL) and one with sulforhodamine B instead of the fluorescent substances (negative WODL) were prepared. Fluorescent intensity was measured after positive and negative WODLs were mixed. Since ACMS was not commercially available, it was synthesized in this study (as shown in the Supporting Information). AMC leaked into the oil immediately after mixing, and blue fluorescence was uniform in all WODLs (Figure 3-8A). AMCA-H showed no leakage immediately after mixing, but the blue fluorescence increased in the negative (non-AMCA-H) WODLs after 168 hours (Figure 3-8B). Indeed, the signal-to-background (S/B) ratio of positive (containing AMCA-H) and negative WODLs of blue fluorescence decreased as time passed (Table 3-2). On the other hand, ACA was retained in the WODL even after 168 hours (Figure 3-8C), and the S/B ratio had a similar value compared to that seen

immediately after mixing. As reported previously,¹²⁴ ACMS remained in the WODL even after 168 hours due to increased hydrophilicity (Figure 3-8D), and it had the highest S/B ratio.



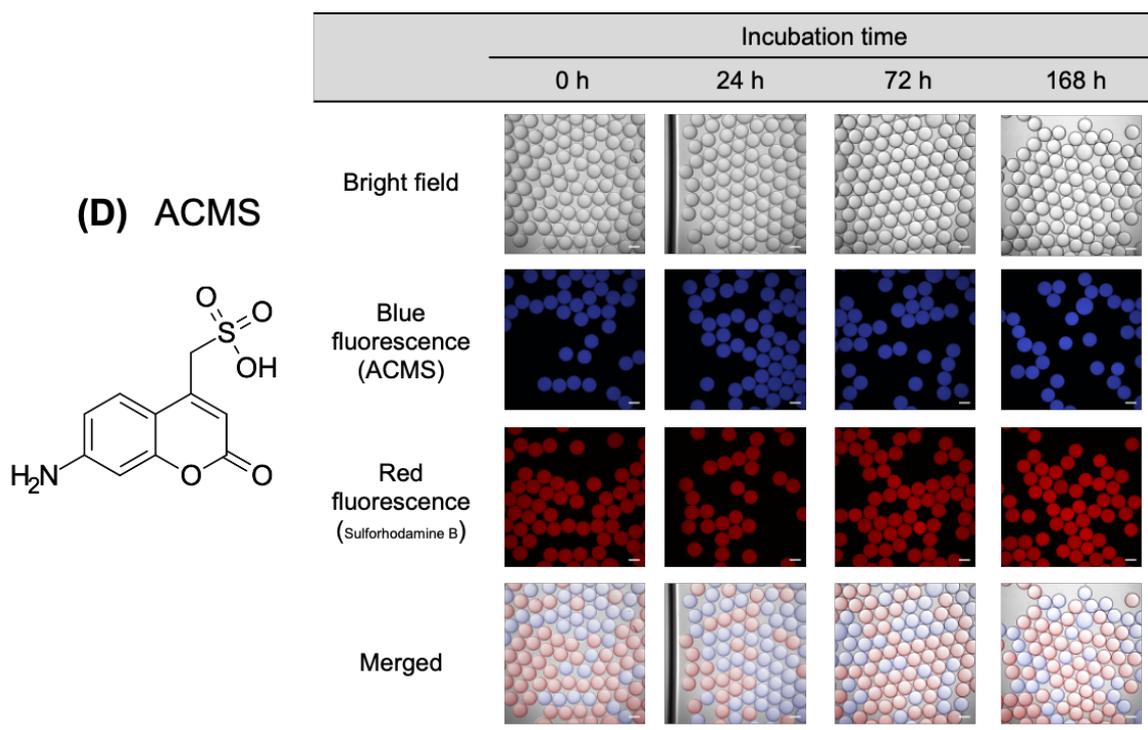
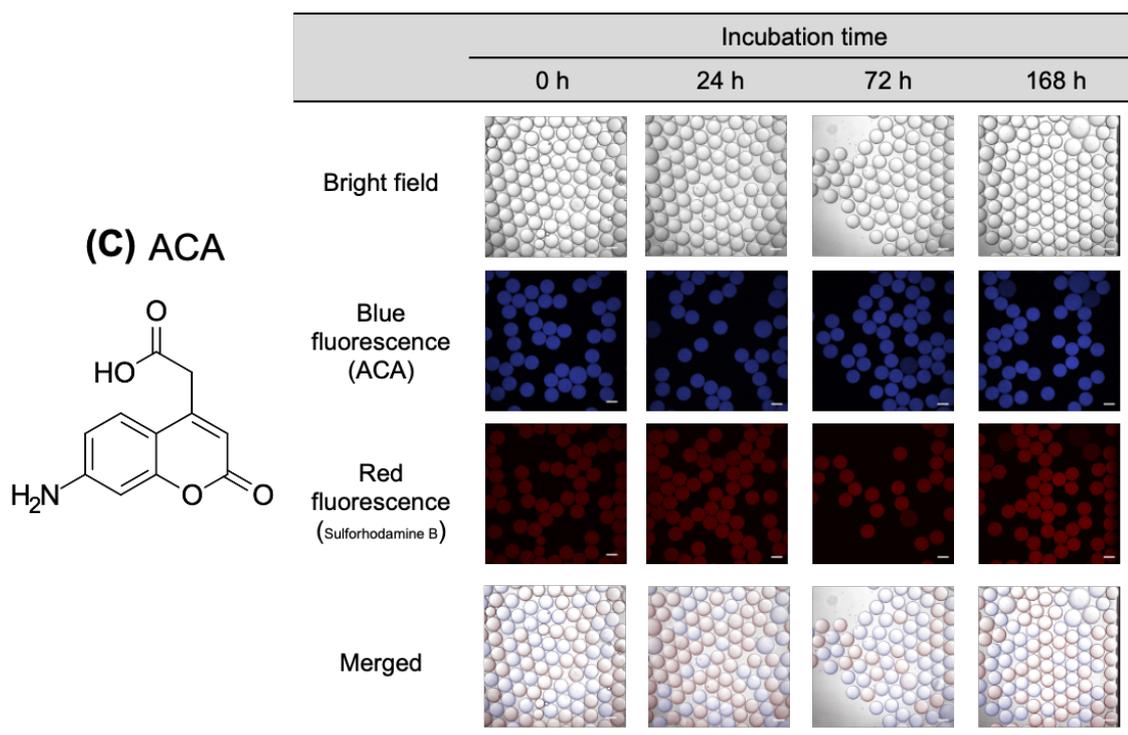


Figure 3-8 | Leakage examination of the fluorescent compounds. (A) AMC. (B) AMCA-H. (C) ACA. (D) ACMS. A quantity of 100 μM of each fluorescent substance was encapsulated in fluorescent WODLs and 10 μM of sulforhodamine B was contained in non-fluorescent substance WODLs. 0 h refers to the time immediately after mixing WODLs. Mixed WODLs were incubated at 30 $^{\circ}\text{C}$ for seven days. Micrographs were obtained by the confocal microscope A1 HD25 under 100x magnification. The scale bar represents 100 μm .

Table 3-2 | S/B ratios and MiLogP values of the fluorescent substances.

Fluorescent substance	S/B ratio of each incubation time				MiLogP
	0 h	24 h	72 h	168 h	
AMC	1.01	1.00	1.00	1.01	1.44
AMCA-H	18.8	15.6	11.2	6.57	0.80
ACA	22.8	22.8	26.7	26.4	0.42
ACMS	64.9	68.8	66.2	73.4	-1.95

The S/B ratio was calculated by dividing the average value of the blue fluorescence of 10 positive WODLs and that of 10 negative WODLs from micrographs. '0 h' refers to microscopic observation immediately after mixing positive and negative WODLs. Micrographs are shown in Figure 3-8. Blue fluorescence intensities were analyzed by ImageJ.⁹⁶

3.3.2. Synthesis of fluorogenic substrates

Since ACMS and ACA were retained in the WODL, they were considered as candidates for a fluorescent probe to detect DPP activities in WODL. Toward this, we first synthesized Met-Leu-ACMS and Met-Leu-ACA, both of which would be degradable by bacterial DPP7. It should be noted that the dipeptidyl ACA substrate is a novel substrate that has not been reported yet. Fluorogenic substrates that contain ACA were synthesized using solid-phase peptide synthesis (Figure 3-9). The Fmoc group was introduced to ACA using Fmoc-Cl, as reported by Harris *et al.*¹³⁶ The Fmoc-ACA was loaded onto 2-chlorotrityl chloride resin and the Fmoc group was removed by 20% piperidine. The first Fmoc amino acid was coupled using 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate with collidine, and after deprotection of the Fmoc, the second amino acid was incorporated using *N,N'*-diisopropylcarbodiimide with 1-hydroxybenzotriazole. After the deprotection, the extended peptides were cleaved with a trifluoroacetic acid cocktail. Reverse-phase high performance liquid chromatography (HPLC) purification was performed to obtain the dipeptidyl ACA substrates. Their synthetic yields were low, presumably because of the difficulty in coupling with the first amino acid and the amino group of ACA loaded on the resin. On the other hand, the substrate containing ACMS¹²⁴ was synthesized as shown in Figure 3-10. Leucine methyl ester was coupled with Boc-methionine and the subsequent saponification yielded the dipeptide intermediate. We chose a mixed anhydride method for coupling with ACMS. The final deprotection of the Boc group and reverse-phase HPLC purification yielded the

ACMS substrate. The synthesis methods are described in detail in materials and methods.

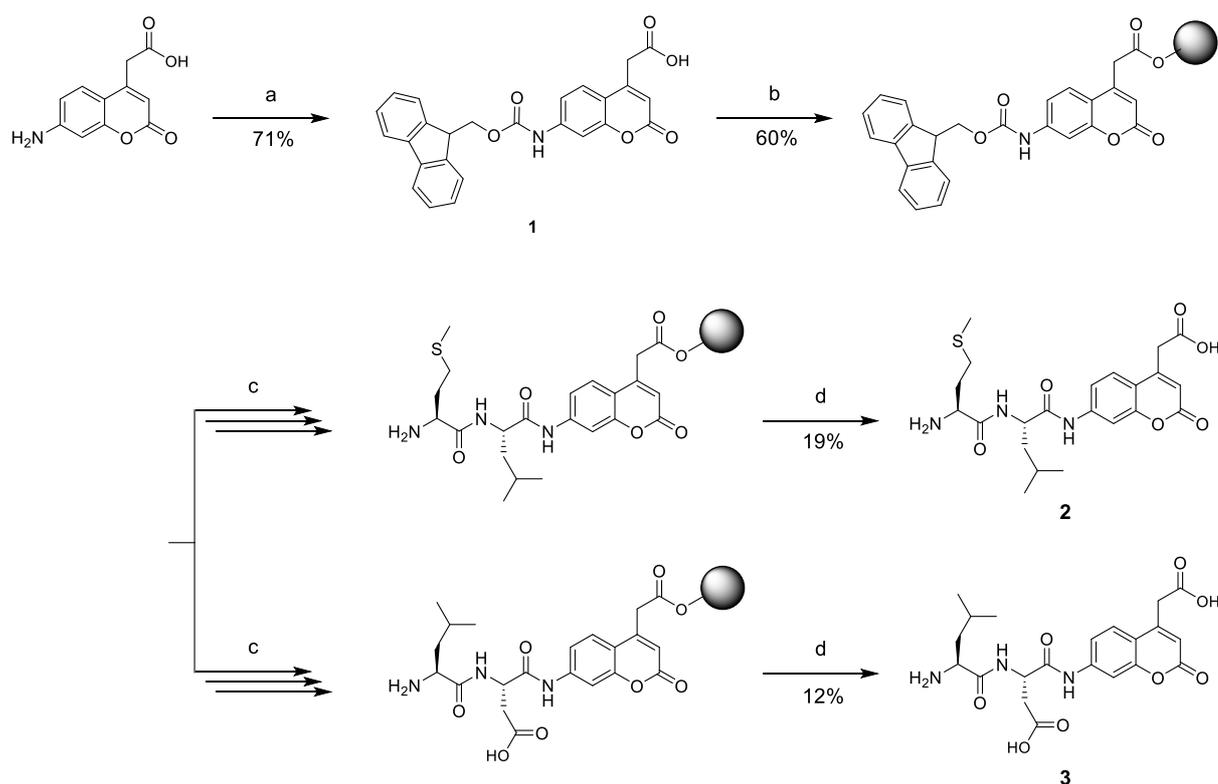


Figure 3-9 | Synthesis of dipeptidyl-ACA 1 and 2.^a
^a (a) Fmoc-Cl, NaHCO₃ aq. 18 h; (b) 2-chlorotrityl chloride resin, DIEA, dichloromethane, 17 h; (c) i. 20% piperidine/DMF, 20 min; ii. Fmoc-Leu-OH or Fmoc-Asp(tBu)-OH, HATU, 2,4,6-collidine/DMF, 24 h; iii., 20% piperidine/DMF, 20 min; iv. Fmoc-Met-OH or Fmoc-Leu-OH, DIC, HOBT/DMF, 2h; v. 20% piperidine/DMF, 20 min; (d) i. TFA/TIS/H₂O/DODT (92.5:2.5:2.5:2.5) or TFA/TIS/H₂O (95:2.5:2.5), 2 h; ii. Prep. RP-HPLC purification.

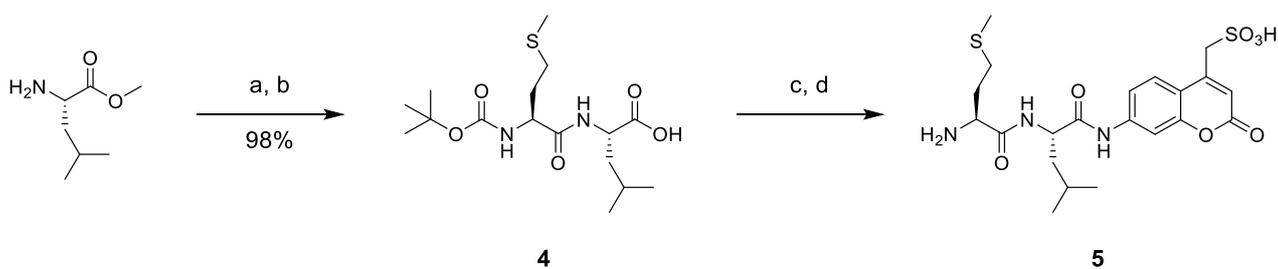


Figure 3-10 | Synthesis of dipeptidyl-ACMS 5.^a
^a (a) Boc-Met-OH, Et₃N, EDC, HOBT/DMF, overnight; (b) 1M NaOH aq., MeOH, 2 h; (c) i. isobutyl chloroformate, 4-methylmorpholine/DMF, 30 min; ii. ACMS, 4-methylmorpholine/DMF, overnight; (d) i. 4M HCl/EtOAc, anisole, 1 h; ii. Prep. RP-HPLC purification.

3.3.3. Evaluation of the detection of purified DPP activities in bulk

Synthesized Met-Leu-ACMS and Met-Leu-ACA were evaluated to establish whether they could detect purified DPP activity in bulk (non-compartment) and compared with the AMC substrate. PmDAP BII (currently known as bacterial DPP7) and SmDPP7, which are

typical bacterial DPP7s showing substrate specificity for Met-Leu-AMC, were used for demonstration purposes.^{35,93} We found that the dipeptidyl ACMS substrate was completely undetectable for DPP activities (Figure 3-11). For Met-Leu-ACA, this substrate was able to detect for DPP activities. The specificity constant (k_{cat}/K_m) of DPP7s against this substrate increased approximately four-fold, which is attributed to a lower K_m value, compared to that of Met-Leu-AMC (Table 3-3). Therefore, we synthesized Leu-Asp-ACA and examined whether it could detect DPP11 activity, along with DPP7. SmDPP11 and PgDPP11, which possess substrate specificity against Leu-Asp-AMC, were used for verification.^{36,93} DPP activities were successfully detected using Leu-Asp-ACA (Table 3-3). For SmDPP11, the value of k_{cat}/K_m increased compared to that of Leu-Asp-AMC substrate, was similar to DPP7s, but not in PgDPP11. It is noteworthy that the specificity constants increased for three of the four family S46 DPPs measured in this study. In addition, the slight increase in Stokes shift due to change in the maximum fluorescence wavelength to the long wavelength side suggests that ACA is a fluorescent group with a better signal-to-noise ratio than AMC (Figure 3-4). Therefore, we note that ACA could become a standard fluorescent substance for DPPs replacing AMC.

Table 3-3 | kinetic parameters of the bacterial DPP7 and DPP11 toward synthetic substrates measured in bulk.

Enzymes	Substrates	V_{max} (IU mg ⁻¹)	K_m (μM)	k_{cat} (sec ⁻¹)	k_{cat}/K_m (μM ⁻¹ sec ⁻¹)	Ratio of k_{cat}/K_m to AMC
PmDAPBII	Met-Leu-AMC	0.637 ± 0.012	5.30 ± 0.12	0.811 ± 0.016	0.153 ± 0.004	1.00
	Met-Leu-ACA	0.483 ± 0.002	1.11 ± 0.02	0.614 ± 0.003	0.552 ± 0.005	3.60
SmDPP7	Met-Leu-AMC	0.672 ± 0.006	21.1 ± 0.5	0.849 ± 0.007	0.0402 ± 0.0007	1.00
	Met-Leu-ACA	0.512 ± 0.003	4.09 ± 0.13	0.647 ± 0.004	0.158 ± 0.004	3.94
PgDPP11	Leu-Asp-AMC	3.93 ± 0.06	10.6 ± 0.2	5.22 ± 0.08	0.492 ± 0.007	1.00
	Leu-Asp-ACA	2.94 ± 0.07	10.1 ± 0.5	3.90 ± 0.10	0.386 ± 0.008	0.785
SmDPP11	Leu-Asp-AMC	14.3 ± 0.4	45.0 ± 1.9	18.0 ± 0.5	0.401 ± 0.007	1.00
	Leu-Asp-ACA	18.1 ± 0.3	12.2 ± 0.4	22.8 ± 0.3	1.88 ± 0.03	4.68

Standard deviations were obtained from three independent experiments. Michaelis–Menten plots are shown in Figure 3-11.

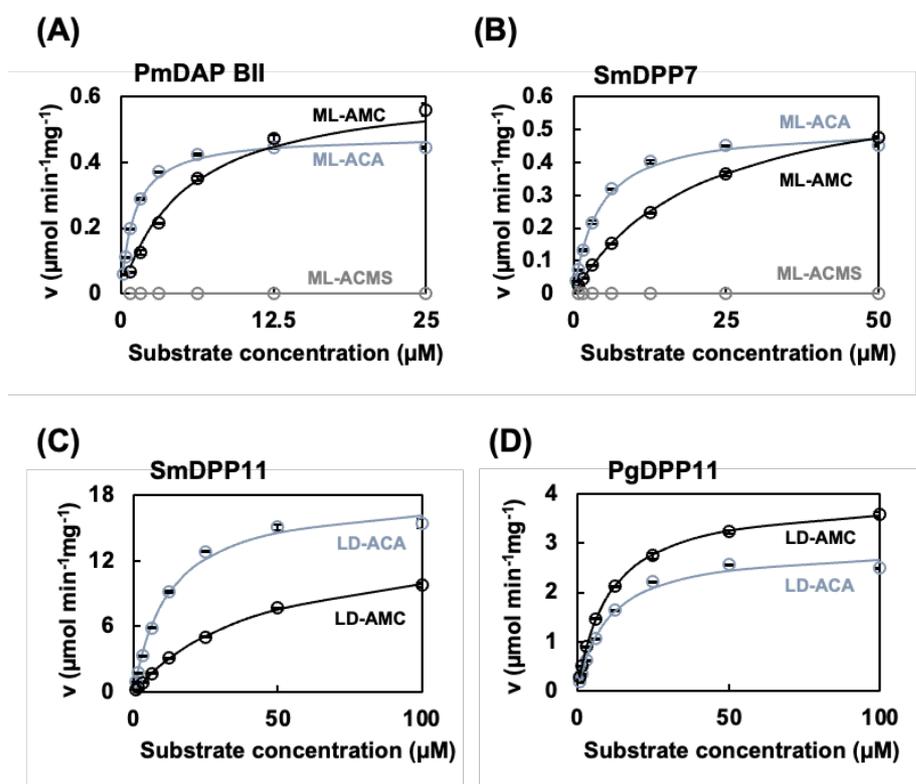


Figure 3-11 | Michaelis–Menten plots of PmDAP BII (A), SmDPP7 (B), SmDPP11 (C), and PgDPP11 (D). Enzymatic reaction was performed in a 50 mM sodium phosphate buffer pH 7.0 containing 5 mM EDTA and 0.005% Tween 20 at 25 °C for 20 min. Standard deviations were obtained from three independent experiments.

3.3.4. Detection of purified DPP activities in the WODL

The dipeptidyl fluorogenic substrates Met-Leu-ACA and Leu-Asp-ACA were capable of detecting dipeptidyl peptidase activity in bulk (non-compartment). Subsequently, we verified whether the detection of purified enzyme activity was possible in the WODL (compartment). We used 50 nM PmDAPBII and 5 nM PgDPP11 as degradable enzymes for Met-Leu-ACA and Leu-Asp-ACA, respectively. An increase in blue fluorescence intensity was confirmed to be time-dependent in both ACA substrates using microscopy (Figure 3-12 A and B). The specific activities of PmDAPBII and PgDPP11 against Met-Leu-ACA and Leu-Asp-ACA were 0.327 IU/mg and 1.41 IU/mg, respectively. The lower specific activity than that measured by the bulk method was at a lower temperature than in the bulk method (25 °C) because the temperature was not controlled on the microscope stage. Moreover, the WODLs at zero and 60 minutes after the reaction were analyzed using FADS and an increase in blue fluorescence intensity was detected in the WODL at 60 minutes after the reaction (Figure 3-12C and D). This suggests that rapid screening using

FADS based on DPP activity is possible when using the ACA substrate. To the best of our knowledge, the dipeptidyl ACA substrate is the first fluorogenic substrate to be shown as capable of detecting and quantifying DPP activity in a WODL.

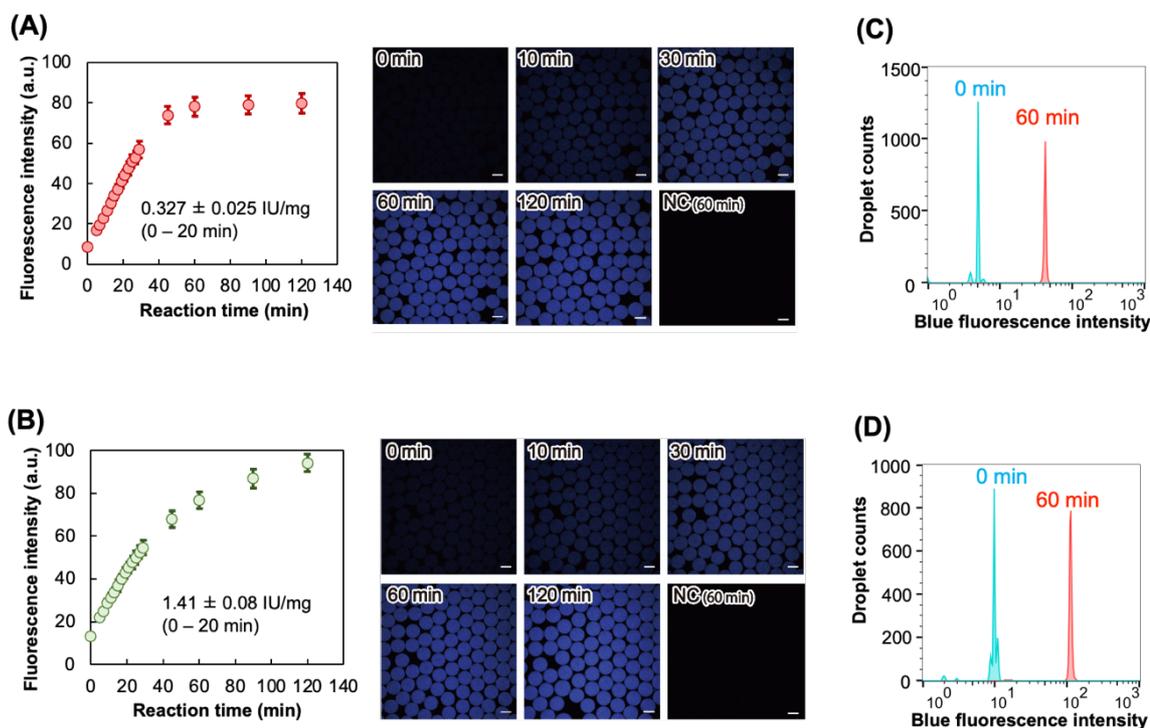


Figure 3-12 | Detection of purified DPP activity in WODLs using dipeptidyl ACA substrates. (A) and (C): PmDAP BII hydrolyzed Met-Leu-ACA. (B) and (D): PgDPP11 hydrolyzed Leu-Asp-ACA. (A) and (B): Enzyme reaction curves. Fluorescence intensity was measured from microscopy images with ImageJ.⁹⁶ Micrographs were obtained by a confocal microscope under 100x magnification. ‘NC’ means no enzyme, only substrate and buffer. The scale bar represents 100 μ m. Standard deviations were obtained from the fluorescence intensity of 10 droplet images. (C) and (D): FADS histogram. ‘0 min’ means analyzing immediately after generating the WODLs. A total of about 10,000 droplets were analyzed by FADS.

3.3.5. Detection of DPP activities of bacterial cells in a WODL and screening by FADS

Here, we validated that rapid screening of microorganisms based on DPP activity using WODL is possible. Incidentally, ACA substrates were detectable for DPP activities in the casitone medium, including many peptides and amino acids from casein (Figure 3-13).

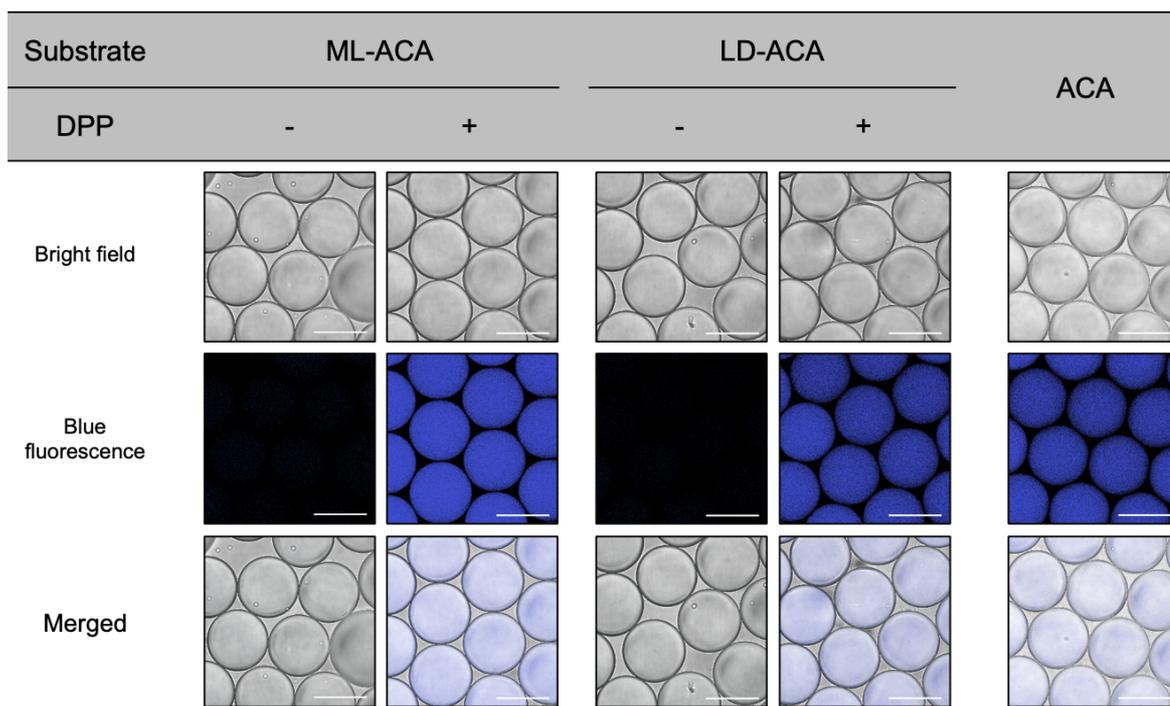


Figure 3-13 | Detection of DPP activity using ACA substrate in casitone medium. Micrographs were obtained by the confocal microscope A1 HD25 under 200 \times magnification. The scale bar represents 100 μ m. Blue fluorescence shows DPP activity. For Met-Leu-ACA, 1 nM PmDAP BII and 20 μ M Met-Leu-ACA were mixed in casitone medium. For Leu-Asp-ACA, 1 nM PgDPP11 and 20 μ M Leu-Asp-ACA were combined in casitone medium. The control was 20 μ M ACA in casitone medium. These reaction solutions were encapsulated as WODLs using HFE-7500 3M Novec Engineered fluid (HFE-7500) with a fluorine oil containing 2% (w/w) 008-FluoroSurfactant (RAN Biotechnologies, MA, USA). Enzyme reaction was performed at 30 $^{\circ}$ C for 24 h in the WODL.

As a model microorganism, a DPP-producing bacterium *P. mexicana* WO24¹³¹ and a DPP-non-producing bacterium *E. coli* DH5 α were used. In addition, *E. coli* was transfected with a plasmid containing an RFP coding gene to distinguish between *E. coli* and *P. mexicana* WO24. Both microorganisms (not mixture) were each compartmentalized into WODLs along with ACA substrates and casitone medium. For *P. mexicana* WO24, an increase in blue fluorescence showing DPP activity with the growth of microorganisms was observed using both Met-Leu-ACA and Leu-Asp-ACA substrates (Figure 3-14A and B). The WODLs with increased blue fluorescence could be separated by FADS. For *E. coli*, although red fluorescence associated with their growth was detected, blue fluorescence was not increased with either of the ACA substrates (Figure 3-15). We attempted to separate *P. mexicana*- and *E. coli*-encapsulated WODLs based on DPP activity. FADS sorted 11.2% (with Met-Leu-ACA) and 7.1% (with Leu-Asp-ACA) of the WODLs' elevated blue fluorescence values due to degradation of the ACA substrates (Figure 3-14C and D). Microscopic analysis showed that only *P. mexicana* WO24-encapsulated WODLs were isolated according to DPP activity using FADS on each substrate. These results

demonstrate that DPP-producing microorganisms could be rapidly screened by FADS using the dipeptidyl ACA substrates.

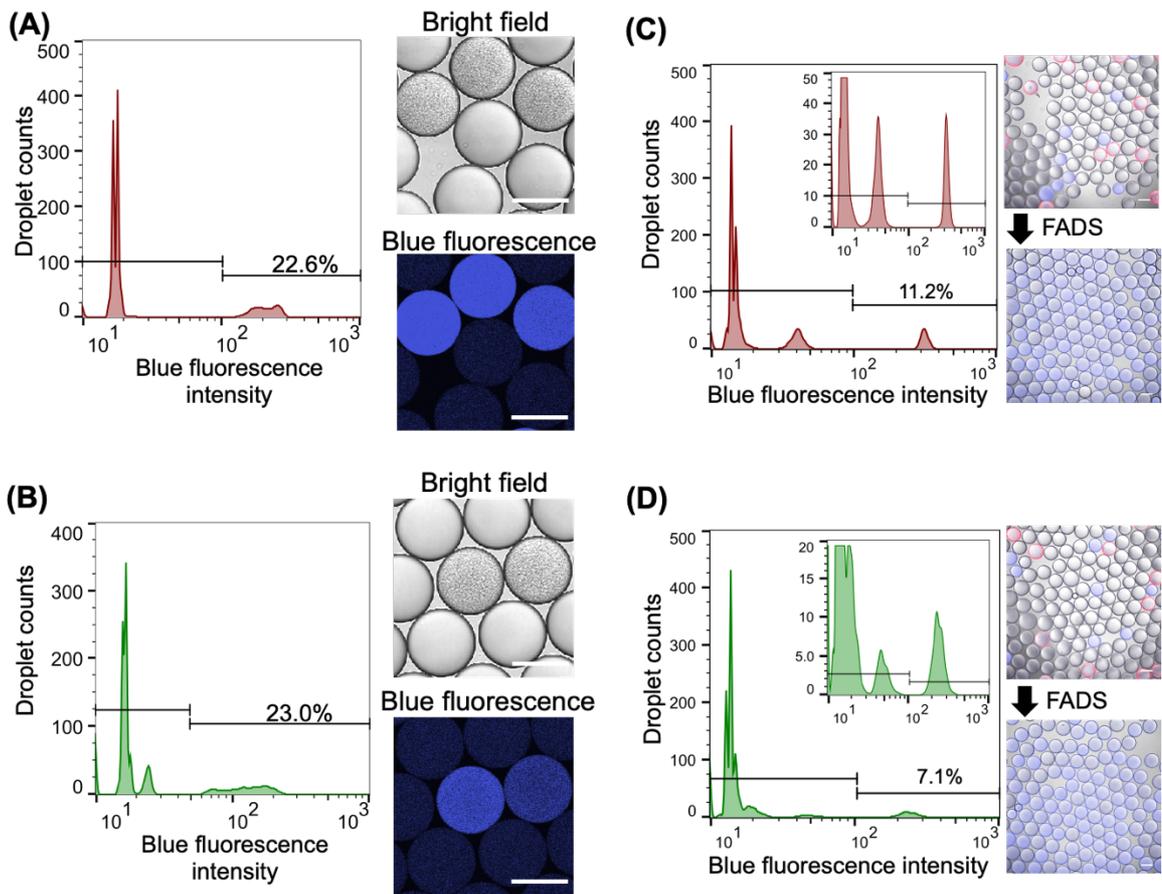


Figure 3-14 | High-throughput isolation of model microorganisms using Met-Leu-ACA (A, C, and E) and Leu-Asp-ACA (B, D, and F). (A) and (B): Detection of DPP activity of the bacterial cell (*P. mexicana* WO24) in WODLs. A total of about 3,000 droplets were analyzed by FADS. (B) and (C): High-throughput sorting of model microorganisms, *P. mexicana* WO24 (DPP-producing) and *E. coli* (DPP non-producing). *E. coli* expressed RFP. A total of about 3,000 droplets were analyzed by FADS. Micrographs before merging are shown in Figure 3-16.

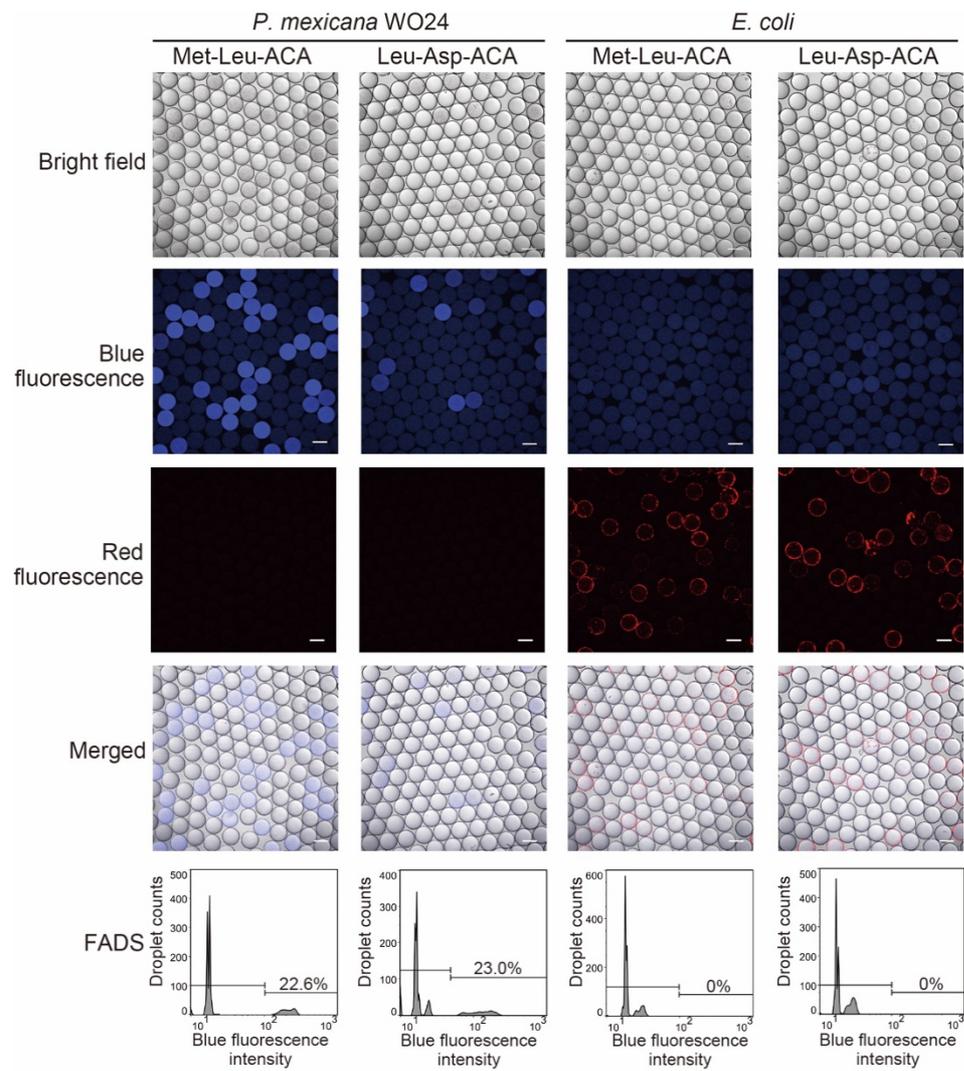


Figure 3-15 | Detection of DPP-producing and non-producing bacteria in WODLs using dipeptidyl ACA substrates.

P. mexicana WO24 and *E. coli* were used as model microorganisms for DPP-producing and non-producing bacteria, respectively. Each microorganism was encapsulated in WODLs together with ACA substrates and cultured for 24 hours at 30 °C. *E. coli* was transfected with a plasmid containing an RFP-coding gene. FADS was performed using On-Chip Sort, and a total of about 3,000 droplets were analyzed. Micrographs were obtained by the confocal microscope A1 HD25 under 100× magnification. Blue fluorescence and red fluorescence indicate DPP activity and *E. coli* cells, respectively. The scale bar represents 100 μm.

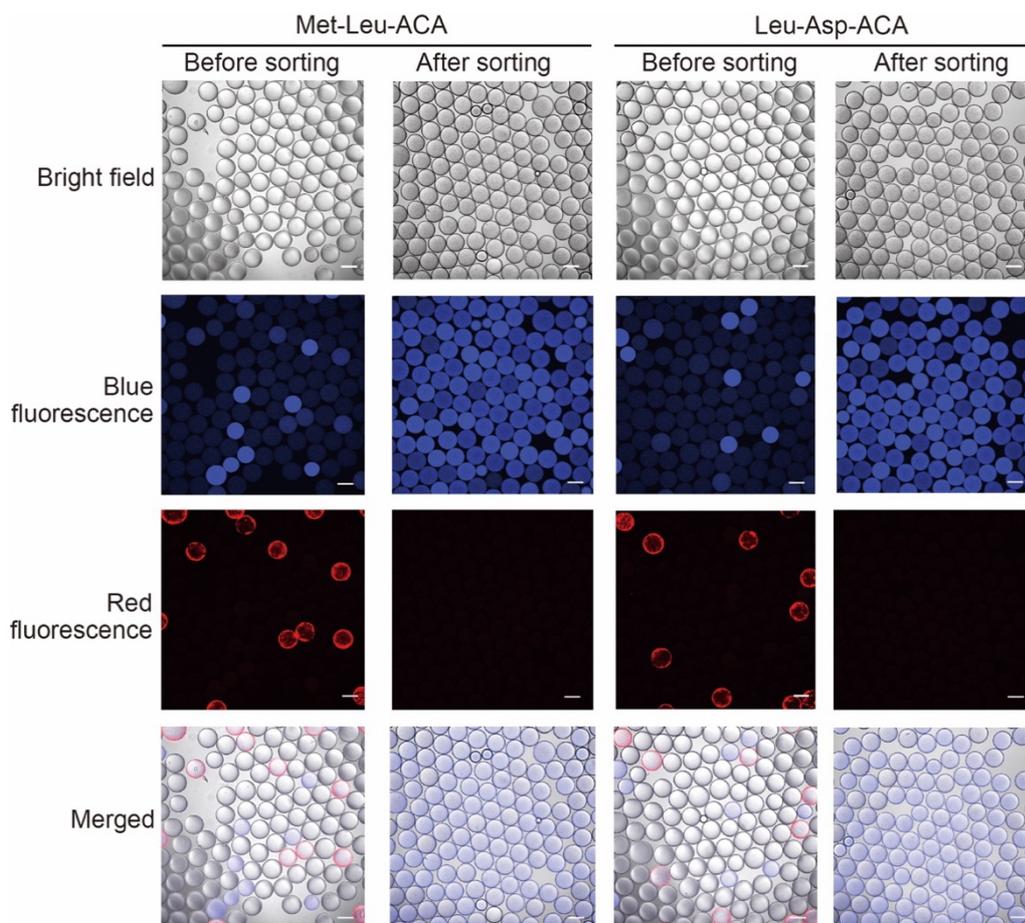


Figure 3-16 | Micrographs of high-throughput sorting of model microorganisms using dipeptidyl ACA substrates.

Droplet images were obtained by the confocal microscope A1 HD25 under 100× magnification. Blue fluorescence and red fluorescence indicate DPP activity and *E. coli* cells, respectively. The scale bar represents 100 μm.

One of the advantages of the WODL for microorganism screening is compartmentalization. Thus, *E. coli* and *P. mexicana* were mixed before formation of WODLs and only the *P. mexicana*-encapsulated WODL was isolated from the two kinds of bacteria mixed in solution based on DPP activity. In addition, FNAP-sort was included in the WODLs to distinguish the WODLs in which bacteria grew.¹²² FNAP-sort is a fluorescent nucleic acid substrate that emits green fluorescence when degraded by RNase produced by microorganisms. Figure 3-17A and B show the result of FADS after cultivation at 30 °C for one day. The WODLs were divided into three areas by green and blue fluorescence intensities as follows: A1) This area with low green and blue fluorescence showed empty WODLs with no microorganisms and no DPP activity; A2) This area with high green fluorescence and low blue fluorescence showed *E. coli*-

encapsulated WODLs, which grew the DPP non-productive microorganism; A3) This area with high green and blue fluorescence showed *P. mexicana*-encapsulated WODLs because microorganisms grew and produced DPP. In practice, only the *P. mexicana*-encapsulated WODLs could be isolated after WODLs distributed in A3 were sorted. However, RFP fluorescence derived from *E. coli* was detected in *P. mexicana*-encapsulated WODLs. This is because both *P. mexicana* and *E. coli* are simultaneously encapsulated with a certain probability when a WODL is encapsulated according to Poisson distribution. In this connection, the abundance ratio of each WODL follows approximately equal numbers of the Poisson distribution (Figure 3-5). Consequently, high throughput isolation of a DPP-producing bacterium based on DPP activity from a microorganism mixture is possible by using dipeptidyl ACA substrate and WODL.

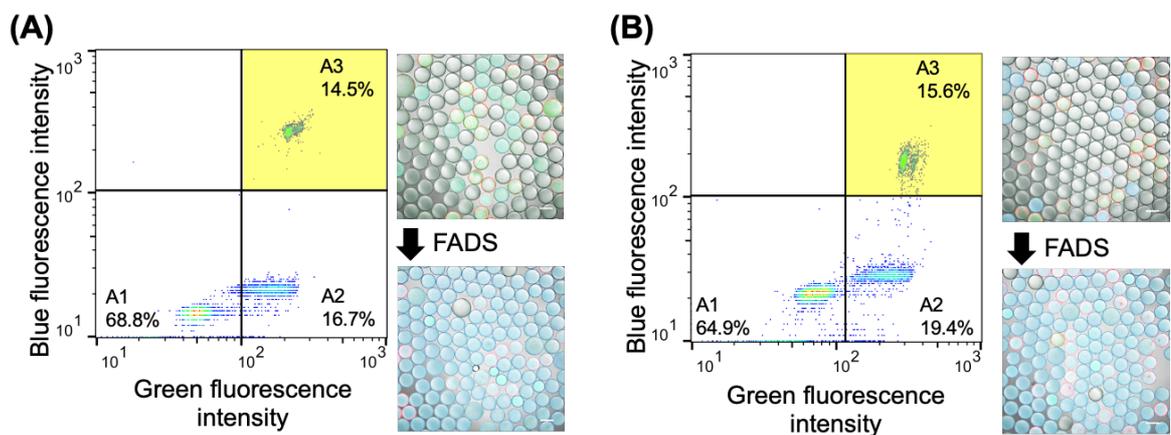


Figure 3-17 | High-throughput isolation based on DPP activity from suspensions containing two kinds of bacteria, *P. mexicana* WO24 and *E. coli* using Met-Leu-ACA (A) and Leu-Asp-ACA (B). In the pseudocolor plot of FADS, blue and green correspond to areas of lower cell density, red and orange to areas of high cell density, and yellow is mid-range. A total of about 8,000 droplets were analyzed by FADS. Micrographs before merging are shown in Figure 3-18. All scale bars in micrographs represent 100 μ m.

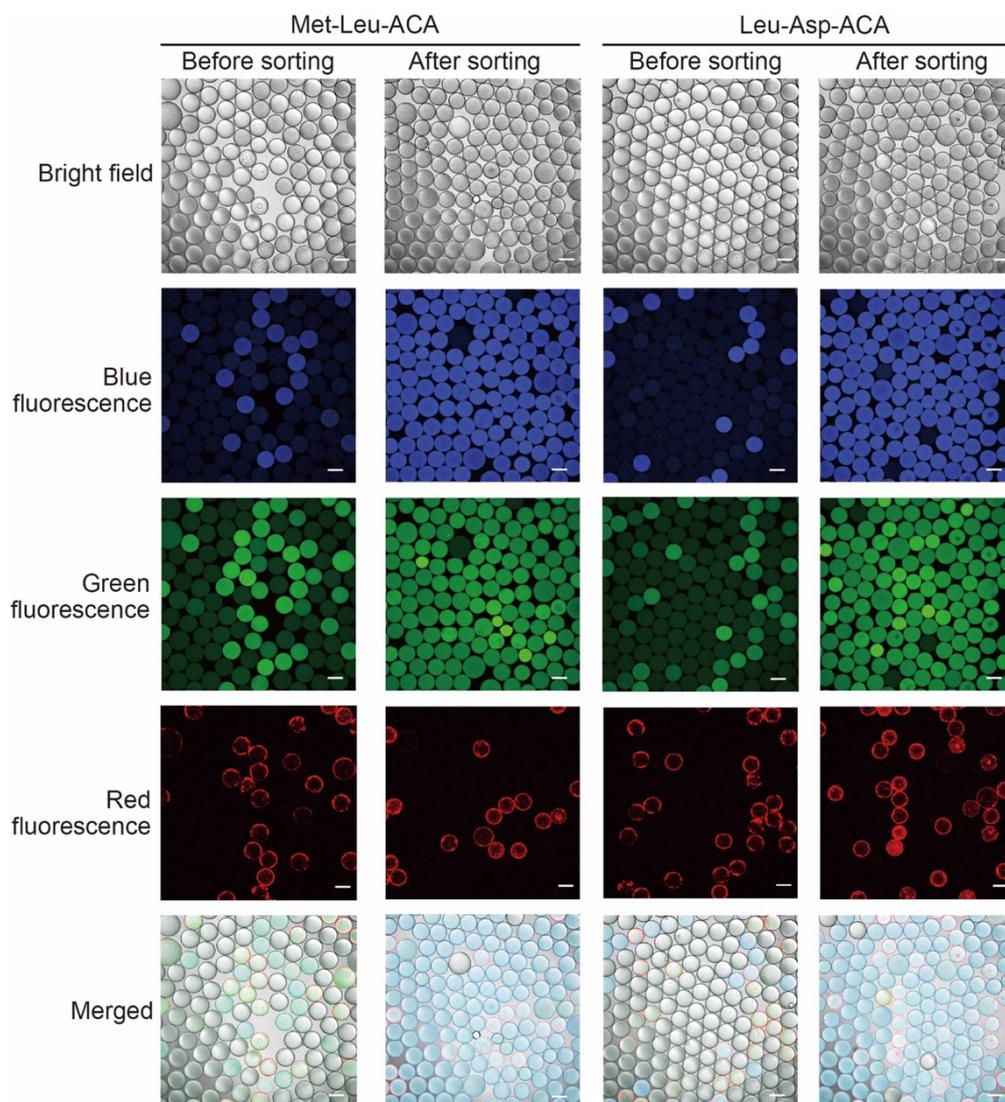


Figure 3-18 | Micrographs of high-throughput isolation based on DPP activity from suspensions containing two kinds of bacteria. Droplet images were obtained by the confocal microscope A1 HD25 under 100× magnification. Blue fluorescence, green fluorescence, and red fluorescence indicate DPP activity, growth of microorganisms (RNase activity), and *E. coli* cells, respectively. The scale bar represents 100 μm.

3.3.6. Isolation of DPP-producing bacteria from the environment

To demonstrate high-throughput screening of environmental microorganisms using ACA substrate, microorganism screening was performed based on DPP activity from *tofu* factory waste fluid, which is expected to be rich in a protein substrate (Figure 3-19A). A bacterial cell suspension was enclosed in a WODL with each of the ACA substrates and an FNAP-sort. After cultivation at 30 °C for 24 hours some WODLs showed blue fluorescence from DPP activity and green fluorescence from bacterial growth, and the WODLs in the A3 area were sorted at about 300 droplets/sec by FADS (Figure 3-19B and C). The sorted WODLs

were dispersed by surfactant-free fluorine oil, and only the aqueous phase containing the microorganism was spread on the LB agar medium. A total of 22 colonies screened based on the ACA substrate activity were randomly isolated. The species were identified by 16S rRNA gene amplicon sequencing. A total of 30 species were identified in screening using each ACA substrate (Table 3-4). In particular, *Stenotrophomonas rhizophila*, *Sphingomonas yabuuchiae*, *Pseudoduganella danionis*, *Flavobacterium chilense*, *Chryseobacterium camelliae*, and *Aeromonas hydrophila* were isolated based on both Met-Leu-ACA and Leu-Asp-ACA hydrolysis activity. A homology search was performed using BLAST^{138,139} to determine whether each species carries the DPP genes.¹⁴⁰ Among the DPP derived from microorganisms, DPP3, DPP5, DPP7, and DPP11, which have the potential to hydrolyze Met-Leu-ACA and/or Leu-Asp-ACA, were selected. In total, 17 strains possessed DPP genes, most of them carrying the S46 peptidases DPP7 and DPP11 (Table 3-4). This result shows that it is possible to isolate microorganisms that possess DPP from the environment using the dipeptidyl ACA substrate with high throughput.

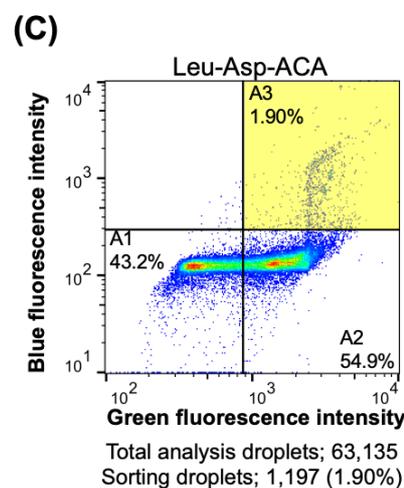
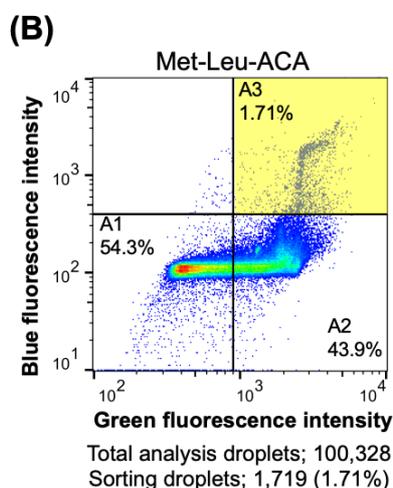
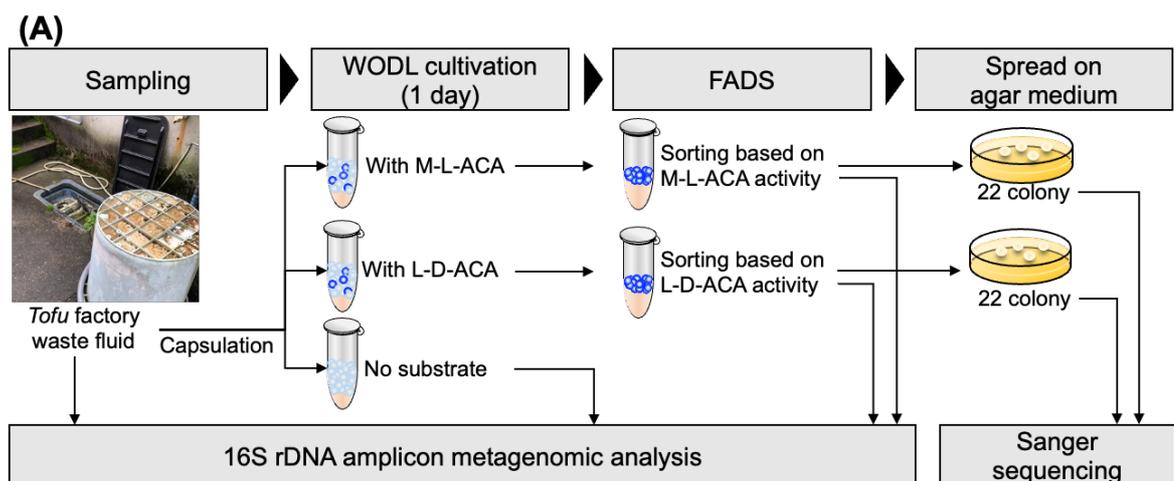


Figure 3-19 | High-throughput screening of environmental microorganisms based on DPP activity. (A): Scheme overview of application of dipeptidyl ACA substrate for high-throughput screening of DPP-producing bacteria. (B) and (C): Pseudocolor plot of FADS for high-throughput WODLs, sorted using Met-Leu-ACA (B) and Leu-Asp-ACA (C). Blue and green correspond to areas of lower cell density, red and orange are areas of high cell density, and yellow is mid-range.

Table 3-4 | Isolated microorganisms using dipeptidyl ACA substrates.

Sample specificity	No.	Determined strains using BLAST				Homology search using BLAST			
		Strains	Identify (%)	Accession No.	Taxid	Predicted substrate degradability			
						M-L-ACA L-D-ACA	M-L-ACA DPP3	M-L-ACA DPP5	M-L-ACA L-D-ACA
Common	1	<i>Sphingomonas yabuuchiae</i> strain A1-18	99.7	NR_028634.1	172044	-	-	-	-
	2	<i>Pseudoduganella danionis</i> strain E3/2	98.2	NR_152711.1	1890295	-	○	○	○
	3	<i>Stenotrophomonas rhizophila</i> strain e-p10	99.0	NR_121739.1	216778	-	○	○	○
	4	<i>Chryseobacterium camelliae</i> strain THG C4-1	95.2	NR_133724.1	1265445	-	○	○	○
	5	<i>Flavobacterium chilense</i> strain LM-09-Fp	99.6	NR_108512.1	946677	-	○	○	○
	6	<i>Aeromonas hydrophila</i> strain ATCC 7966	99.7	NR_074841.1	644	-	○	-	-
M-L-ACA specific	7	<i>Enterobacter soli</i> ATCC BAA-2102 strain LF7	99.4	NR_117547.1	885040	-	-	-	-
	8	<i>Falsarthrobacter (Arthrobacter) nasiphocae</i> strain M597/99/10	98.1	NR_025424.1	1663	_*	○*	_*	_*
	9	<i>Janthinobacterium agaricidamnosum</i> strain NBRC 102515	99.4	NR_114134.1	55508	-	○	○	○
	10	<i>Stenotrophomonas terrae</i> strain R-32768	99.0	NR_042569.1	405446	-	○	○	○
	11	<i>Lactococcus lactis</i> strain NBRC 100933	99.8	NR_113960.1	1358	-	-	-	-
	12	<i>Pantoea agglomerans</i> strain DSM 3493	99.2	NR_041978.1	549	-	-	-	-
	13	<i>Klebsiella grimontii</i> strain SB73	99.0	NR_159317.1	2058152	-	-	-	-
	14	<i>Cronobacter dublinensis</i> subsp. lausannensis strain E515	99.6	NR_044058.1	413497	-	-	-	-
	15	<i>Moraxella osloensis</i> strain DSM 6998	99.1	NR_113392.1	34062	-	-	-	-
	16	<i>Pseudomonas otitidis</i> strain MCC10330	99.8	NR_043289.1	319939	-	-	-	-
L-D-ACA specific	17	<i>Brevundimonas faecalis</i> strain CS20.3	99.1	NR_117187.1	947378	-	-	-	-
	18	<i>Massilia violacea</i> strain CAVIO	99.8	NR_148592.1	1715466	-	○	○	○
	19	<i>Aeromonas allosaccharophila</i> strain CECT 4199	99.7	NR_025945.2	656	-	-	-	-
	20	<i>Pseudomonas protegens</i> strain CHA0	99.8	NR_114749.1	380021	-	-	-	-
	21	<i>Chryseobacterium gallinarum</i> strain 100	99.5	NR_133726.1	1324352	-	○	○	○
	22	<i>Paenarthrobacter nitroguajacolicus</i> strain G2-1	99.8	NR_027199.1	211146	-	-	-	-
	23	<i>Chryseobacterium taihuense</i> strain THMBM1	97.6	NR_109542.1	1141221	-	○	○	○
	24	<i>Flavobacterium ginsengiterrae</i> strain DCY55	96.4	NR_132661.1	237	○*	○*	○*	○*
	25	<i>Pararheinheimera arenilitoris</i> strain J-MS1	98.1	NR_134151.1	152089	-	-	-	-
	26	<i>Chryseobacterium arachidis</i> strain 91A-593	98.5	NR_133723.1	1416778	-	○	○	○

27	<i>Sphingomonas trueperi</i> strain NBRC 100456	99.9	NR_113897.1	53317	-	○	○	-
28	<i>Pseudomonas japonica</i> NBRC 103040	98.9	NR_114192.1	256466	-	○	-	-
29	<i>Cloacibacterium haliotis</i> strain WB5	98.6	NR_125655.1	501783	-*	○*	○*	○*
30	<i>Brevundimonas terrae</i> strain KSL-145	99.4	NR_043726.1	363631	-	○	○	○

The strains were identified from the sequences containing the variable regions V1–V5 of the 16S rRNA gene, using the Basic Local Alignment Search Tool (BLAST).¹³⁸ The 16S ribosomal RNA sequences (Bacteria and Archaea) database was used, and top hit strains are listed in this table. A homology search was performed using BLAST^{138,139} to determine whether each species (not strains) carries the DPP gene, based on the following amino acid sequences. DPP3 from *Porphyromonas gingivalis* (UniProtKB: B2RLB9), DPP5 from *P. gingivalis* (UniProtKB: B2RIT0), DAP BII (UniProtKB: V5YM14), and PgDPP11 (UniProtKB: B2RID1) were used as the query sequence for DPP3, DPP5, DPP7, and DPP11, respectively. The specificity of the P1 position is as follows. DPP3: arginine, aliphatic, and acidic (degradable but low activity) amino acids.²¹ DPP5: aromatic and aliphatic amino acid.²¹ Family-S46 peptidases DPP7: aromatic, aliphatic, and acidic (degradable but low activity) amino acid.¹⁴¹ Family-S46 peptidases DPP11: aliphatic and acidic amino acid.^{36,142} A circle means possessing DPP homolog genes with over 30% of sequence homology. *Since there were no taxonomy IDs for species in BLAST, we performed homology searches using taxonomy IDs for genera.

3.3.7. Comparative analysis of 16S rDNA

The abundance of bacteria in the *tofu* factory waste fluid sample, WODL cultivation sample, and post-sorting samples was analyzed using a 16S rRNA amplicon metagenomic approach to confirm whether DPP-producing microorganisms are concentrated according to DPP activity screening using the ACA substrate. The abundance ratios of each sample's top 10 OTUs are shown in Figure 3-20. Prevotellaceae, Microcolaceae, and Veillonellaceae were relatively abundant in the *tofu* factory waste fluid sample. However, they were in low abundance after WODL cultivation. In contrast, some microorganisms with low abundance in the *tofu* factory waste fluid sample were increased by WODL cultivation. For instance, the abundance ratios of Moraxellaceae (denovo2810), Enterobacteriaceae (denovo1993), and Aeromonadaceae (denovo1771), whose abundance ratio for the *tofu* factory waste fluid sample were 1% or less, were increased by 93.6-fold, 39.7-fold, and 17.3-fold, respectively by WODL cultivation. Comparison of the *tofu* factory waste fluid sample to FADS based on Met-Leu-ACA catalyze showed that the concentration ratio of Xanthomonadaceae (denovo2489) was the highest among post-sorting samples and was 76.6-fold. The abundance ratios of Moraxellaceae (denovo2810), Weeksellaceae (denovo1186), and Oxalobacteraceae (denovo1645) in the FADS (Met-Leu-ACA) sample were increased by 52.8-fold, 46.4-fold, and 43.8-fold, respectively, compared to the *tofu* factory waste fluid sample. Some species classified as Xanthomonadaceae, Moraxellaceae, Oxalobacteraceae, or Weeksellaceae, which belong to Proteobacteria or Bacteroidetes, possess bacterial DPP7

and DPP11.^{33,110,143} For the FADS (Leu-Asp-ACA) sample, the abundance ratio of Aeromonadaceae (denovo1771) increased the most, by 55.1-fold. The abundance ratios of Weeksellaceae (denovo1186), Oxalobacteraceae (denovo1645), and Pseudomonadaceae (denovo3742) increased 48.7-fold, 31.1-fold, and 23.3-fold, respectively, compared to the *tofu* factory waste fluid sample. Similar to the FADS (Met-Leu-ACA) sample, some species classified as Aeromonadaceae, Oxalobacteraceae, Weeksellaceae, or Pseudomonadaceae belong to the Proteobacteria or Bacteroidetes phylum, and possess the S46 peptidases.^{33,110,143} Therefore, these results support that the screening of microorganisms based on DPP activity using the dipeptidyl ACA substrates succeeded in enriching the abundance of microorganisms possibly through bacterial DPP7 or DPP11.

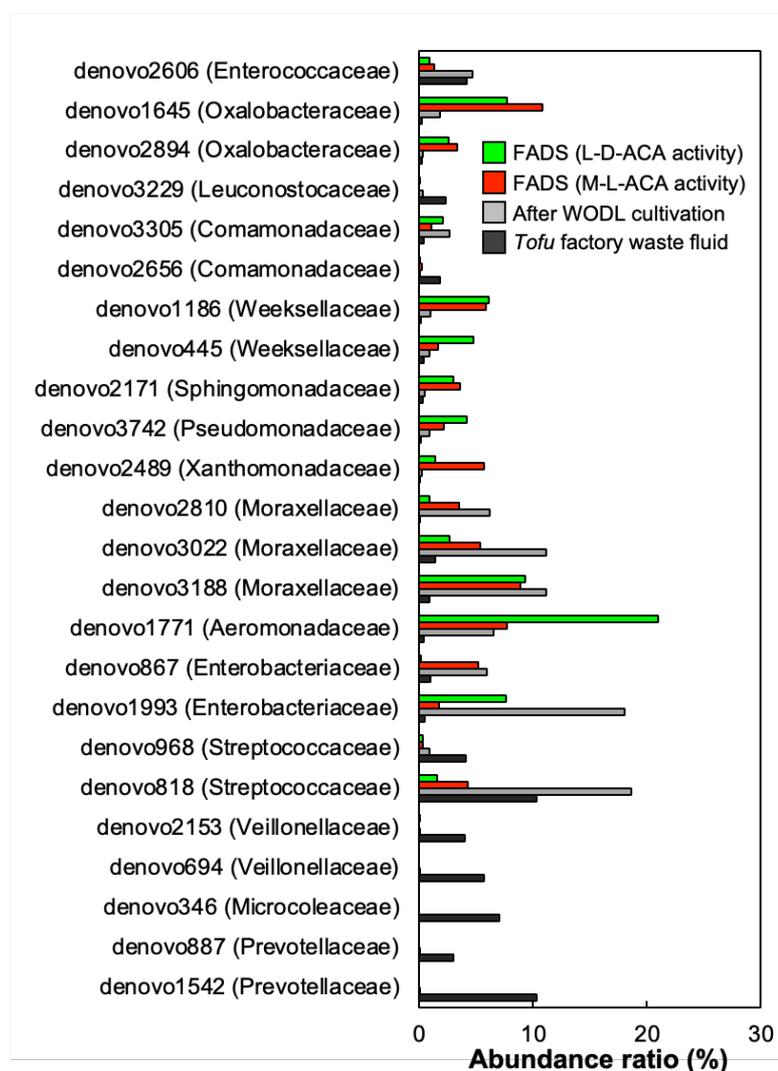


Figure 3-20 | The abundance ratio of the top 10 OTUs of each sample. FADS (M-L-ACA activity) and FADS (L-D-ACA activity) refer to post-sorting samples based on Met-Leu-ACA and Leu-Asp-ACA hydrolysis activity, respectively. The name in parentheses is the family to which each OTU belongs. The abundance ratio was calculated based on 16S rDNA sequence data performed with the MiSeq System.

3.4. Discussion

Here, we developed an AMC-derived substrate retained in a WODL that can be used in droplet-based microfluidic systems and demonstrated the availability of an AMC-derived substrate to screen for DPP activities. We synthesized a dipeptidyl 7-aminocoumarin-4-acetic acid (ACA) substrate in which the carboxyl group of the dipeptide forms a peptide bond with the amino group at position 7 of ACA.¹³⁶ ACA was retained in the WODL for more than seven days. To investigate the relationship between hydrophobicity and fluorophore exchange, the $\log P$ value of each compound was calculated by Molinspiration (Table 3-2). Incidentally, we sought to evaluate AFC, but 100 μM AFC was insoluble in a 50 mM sodium phosphate buffer (pH 7.0). The $\log P$ value indicated a correlation between hydrophobicity and leakage from the WODL, in agreement with previous reports.¹²⁶ Compared to ACA (MiLog P , 0.42) and AMCA-H (MiLog P , 0.80), the retentive properties of fluorescent probes in the WODL appeared to vary between $\log P$ values of 0.42 and 0.80 (Figure 3-21). Fluorescent probes may be necessary to establish MiLog P values below 0.42 when detecting enzyme activity in the WODL. However, since fluorophore exchange is caused by transport with surfactant instead of diffusion,¹²⁵ the $\log P$ value alone would not be enough to determine whether the fluorescent substance is retained in the WODL rigorously or not. Moreover, although AFC (MiLog P , 1.89) is insoluble in water, not all compounds with an MiLog P value above 1.89 are insoluble. Indeed, it is possible to dissolve fluorescein (MiLog P , 2.56) and resorufin (MiLog P , 2.14) in organic solvents diluted with water and encapsulated in the WODL.^{123,125,127} Thus, it is necessary to consider not only the hydrophobicity but also the polarity and the molecular weight of compounds.

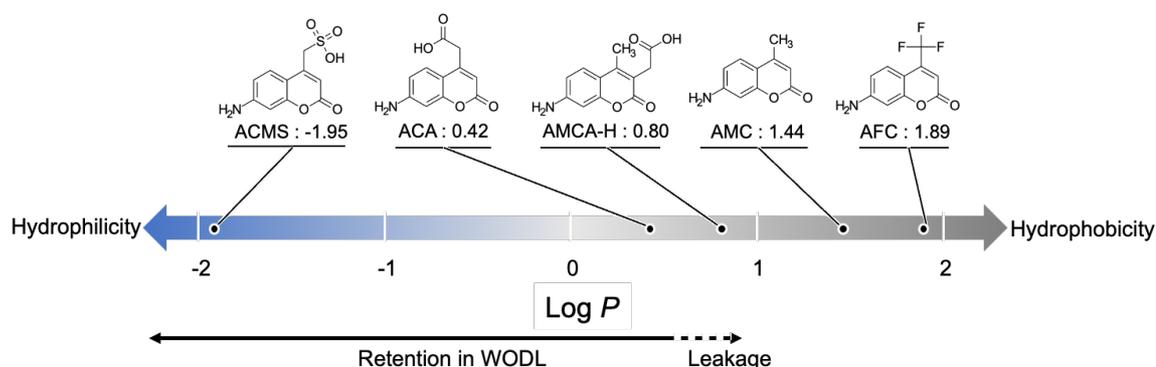


Figure 3-21 | Schematic diagram of the relationship between $\log P$ of coumarins used in this study and retentive properties in the WODL.

Improving the substrate hydrophilicity by adding the sulfo group is an effective way to perform enzyme assay in a WODL. This is supported by the report of Woronoff *et al.*, who succeeded in synthesizing phenylacetyl-ACMS and detecting the activity of *E. coli*-derived acylase in the WODL.¹²⁴ Moreover, phosphotriesterase and cellulase activities could be detected in WODL by synthesizing a substrate with a sulfo group attached.^{126,144} However, this study revealed that the functional group of the fluorescent substance affects the enzyme reaction rate (Table 3-3). Although family-S46 peptidases such as PmDAP BII do not recognize an amino acid side chain at the prime side subsite³⁵ and have a wide space on the prime side (Figure 3-22), the dipeptidyl ACMS substrate was undetectable for DPP activities. It predicted that a hydrophobic surface of S46 peptidases would reject a substrate with its high polarity at the entrance of the substrate-binding pocket. This indicated that excessive polarization, as in adding the sulfo group to the fluorescent probe, could negatively affect not only DPPs, but also the enzyme activity. On the other hand, ACA binding a carboxyl group improves hydrophilicity and has no negative effect on detecting DPP activity. The present study suggests that adding the sulfo group to the fluorescent probe is not the best way to improve the substrate hydrophilicity for use in the WODL. The functional group should be selected according to the target enzymes.

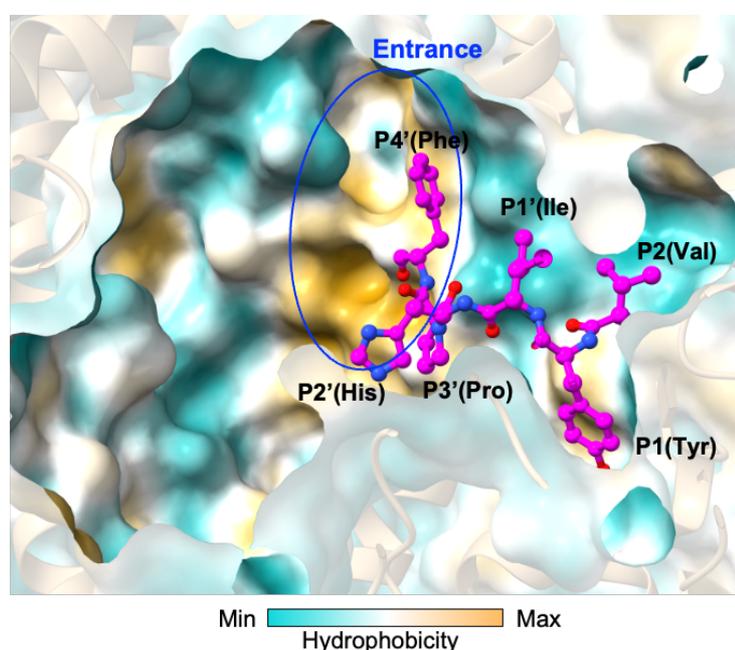


Figure 3-22 | Structure of the substrate binding pocket with surface hydrophobicity of PmDAP BII (PDB; 3WOQ).

Protein structural model produced using the program UCSF Chimera X.¹⁴⁵ Map values for surface '3woq.pdb_A SES surface': minimum -28.39, mean -5.109, maximum 22.83. MLPP: A program for the calculation of molecular lipophilicity potential in proteins.¹⁴⁶

The development of substrates detectable in the WODL enables high-throughput screening and leads to a reduction in the amount of substances used, and also facilitates detection of activity with high sensitivity on activity-based screening: 0.5 nM PgDPP11, which shows an activity of 2.94 IU/mg against Leu-Asp-ACA (in bulk assay), was detectable in a WODL by using 100 μ M Leu-Asp-ACA. The number of molecules that a 5 nM enzyme encapsulated in a 1 nL droplet was about 5×10^{-18} mol (1.17×10^{-15} IU for Leu-Asp-ACA). Moreover, since the WODL shrink allows 100,000-fold enrichment, ultra-trace amounts of released ACA could be detected.¹⁴⁷ Indeed, in our research, 1 pmol ACA was detected in WODLs by shrinking and microscopy (Figure 3-23). This indicates that a tiny amount of DPP would be detected with high sensitivity using the ACA substrate and WODL. Further, because a 100 μ M substrate solution is a tiny amount in a droplet of about 1 nL, a combination of an ACA substrate and a WODL becomes a powerful tool for drug screening that requires a large number of samples, reducing the amount of substances used.

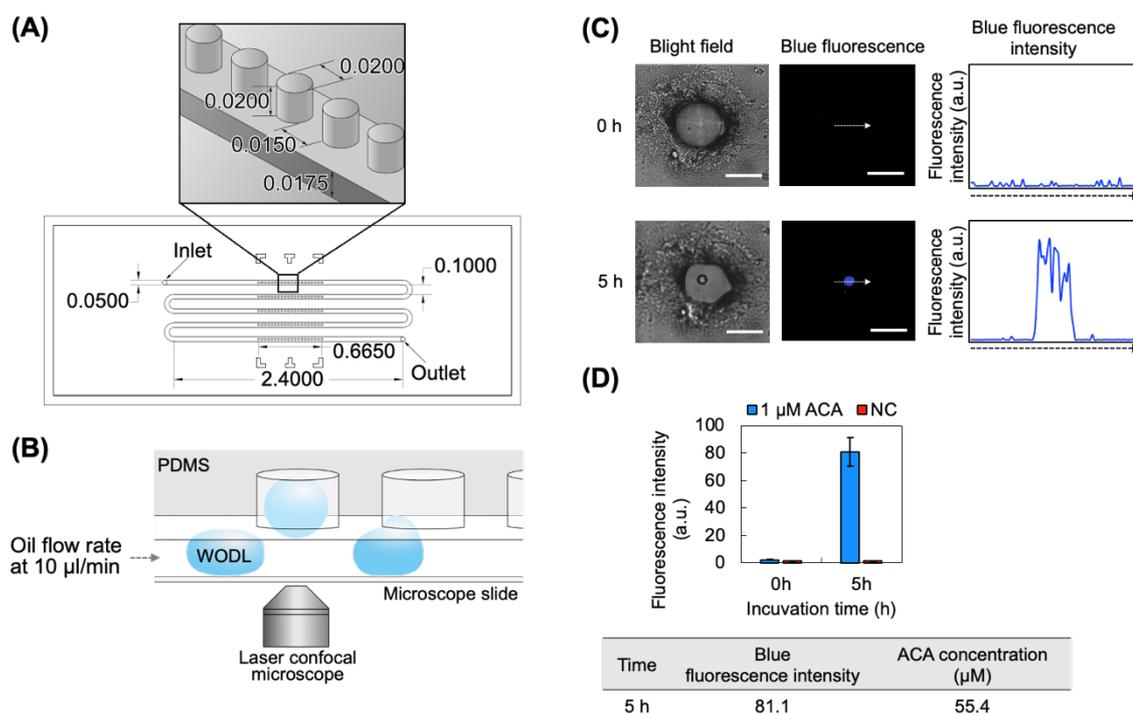


Figure 3-23 | Shrinking analysis using WODL trapped devices. (A) Design of the device. The numbers represent millimeters. Methods of device creation was described in methods. (B) Schematic diagram of Shrinking analysis using WODL trapped devices. WODLs were introduced into device with oil phase flow at 10 μ l/min. (C) Shrinkage analysis. 1 nL WODLs contained 1 μ M ACA were fixed in device. WODLs were observed by confocal microscope under 100x magnification. (D) The Blue fluorescence intensity was measured by ImageJ.⁹⁶ NC is 50 mM sodium phosphate buffer. The ACA concentration was calculated using calibration curve of ACA (Figure 3-4).

This is the first study demonstrating that an ACA substrate can be used to detect enzyme activity of microorganisms in WODLs. Interestingly, the dipeptidyl ACA substrates detect bacterial DPP activity with more sensitivity than AMC substrates, which is attributed to decrease in the K_m values. Moreover, we were able to detect bacterial cell growth and DPP activity suggests that it is possible to screen biological targets while simultaneously measuring their DPP inhibitory capacity and antimicrobial activity. In particular, considering that droplet-based microfluidic systems are suitable for screening cells and microorganisms (possessed different genes) but it is not easy to generate chemically distinct droplets, this technology could be useful in the search for microbial-derived antimicrobials (antibiotics). Since ACA can modify substrates such as peptides, its range of applications will extend to peptidases other than S46 peptidases. For instance, dipeptidyl peptidase IV (DPP IV, DPP4, or DAP IV), which possesses proline-specific DPP activity, is responsible for reducing incretins involved in glucose homeostasis in mammals.^{148,149} So, DPP IV is a well-known target of oral hypoglycemics. If glycyl-L-prolyl-7-aminocoumarin-4-acetic acid (Gly-Pro-ACA) became a substrate for DPP IV, dipeptidyl ACA substrate contributes to high-throughput screening of hypoglycemic agents using droplet-based microfluidic systems. This work would be the base for a highly efficient screening platform for detecting enzyme activity using the WODL, and will not be limited to only DPPs.

Chapter 4 General conclusion

Antimicrobial resistance is a serious worldwide problem.² Moreover, as bacteria are able to circumvent the effects of all the currently available antibiotics, there is an urgent need for the development of new antibiotics with novel modes of action. However, the development of new antimicrobial agents is declining worldwide. Among bacterial pathogens, non-fermenting Gram-negative bacteria (NFGNB) are the main focus of the antimicrobial resistance epidemic.¹⁰ Inappropriate or prolonged use of broad-spectrum antimicrobial agents induce the emergence of resistant bacteria. Thus, appropriate use of narrow-spectrum antimicrobial agents, which act only against the causative organism of the infection as NFGNB, is required.

Family-S46 peptidases are serine proteases showing dipeptidyl peptidase activity and are distributed in NFGNB (such as Bacteroidetes and Proteobacteria), but are not found in mammals.^{110,143} Since S46 peptidases are important enzymes for the growth of bacteria and distributed in NFGNB,^{91,92} these peptidases could be a novel molecular target of antibiotics for NFGNB.⁹³ Therefore, the final goal of this study was to find antimicrobials for NFGNB that target S46 peptidase. In this thesis, we worked on the following two studies:

Chapter 2 – Elucidation of recognition mechanisms at the S2 subsite in family-S46 peptidases

Here, we evaluated the P2-residue preference of DPP7 from *S. maltophilia* (SmDPP7) and determined the crystal structures of SmDPP7 in complexes with four kinds of dipeptides (Val-Tyr, Phe-Tyr, Tyr-Tyr, and Asn-Tyr) at resolutions of 2.03 to 1.86 Å. Biochemical studies showed that SmDPP7 prefers hydrophobic amino acids at the S2 subsite in general, except for asparagine. Crystal structure analysis and isothermal titration calorimetry (ITC) analysis of the Asn-Tyr bindings with SmDPP7 revealed that the exceptionally strong preference for asparagine residue is facilitated by a water-mediated hydrogen bond network in the S2 subsite. An extensive amino acid sequence comparison also revealed that residues in the S2 subsite of S46 peptidases are better conserved than those in the S1 subsite. Indeed, competitive inhibition assay using dipeptides against other S46 peptidases from pathogenic bacteria showed a conserved preference for hydrophobic residues and asparagine at the S2 subsite. These observations

provide novel insights into the design of a universal inhibitor of S46 peptidases.

Chapter 3 – Investigation of an ultra-high-throughput detection method for family-S46 peptidase

Here, we developed an AMC-derived substrate retained in a WODL that can be used in droplet-based microfluidic systems, and demonstrated the availability of an AMC-derived substrate to detect/screen for DPP (S46 peptidases) activities. We synthesized a dipeptidyl 7-aminocoumarin-4-acetic acid (ACA) substrate in which the carboxyl group of the dipeptide forms a peptide bond with the amino group at position 7 of ACA. ACA was retained in the WODL for more than 7 days, and the dipeptidyl ACA substrate detected DPP activity of the bacterial cell in the WODL in addition to detecting purified DPP activity. Since ACA can be applied not only to S46 peptidases but to various other substrates, detectable microbial enzyme activity for droplet-based microfluidic systems can be greatly expanded.

As described in Chapter 2, a universal inhibitor of S46 peptidases is desirable to optimize for the S2 subsite, and a compound that replaces water molecules in the S2 subsite with low $\log P$ values may be a valuable universal inhibitor of S46 peptidase for antibiotics of NFGNB such as *S. maltophilia* and *P. gingivalis*. In general, $\log P \leq 5$ is proposed as a structural property with excellent oral absorption.¹⁵⁰ The ACA-based droplet microfluidics system developed in chapter 3 is useful for screening compounds with these low $\log P$ values. This is because the compounds with lower $\log P$ values would retain in WODL, whereas leakage (not detected) of hydrophobic substances would occur in compounds with high $\log P$ values.

In particular, due to the fact that droplet-based microfluidic systems would be suitable for screening cells, microorganisms, and phage possessed different genes, but as it is not easy to generate chemically distinct droplets, this technology could be helpful in the search for microbial-derived antimicrobials (antibiotics). By using ACA substrates and sensor cells expressing target enzymes (e.g., *E. coli*), we expect to be able to construct an innovative and ultra-high-throughput screening method for narrow-spectrum antibiotics for

NFGNB targeting S46 peptidases with lower $\log P$ values from microorganisms, as shown in Figure 4-1.

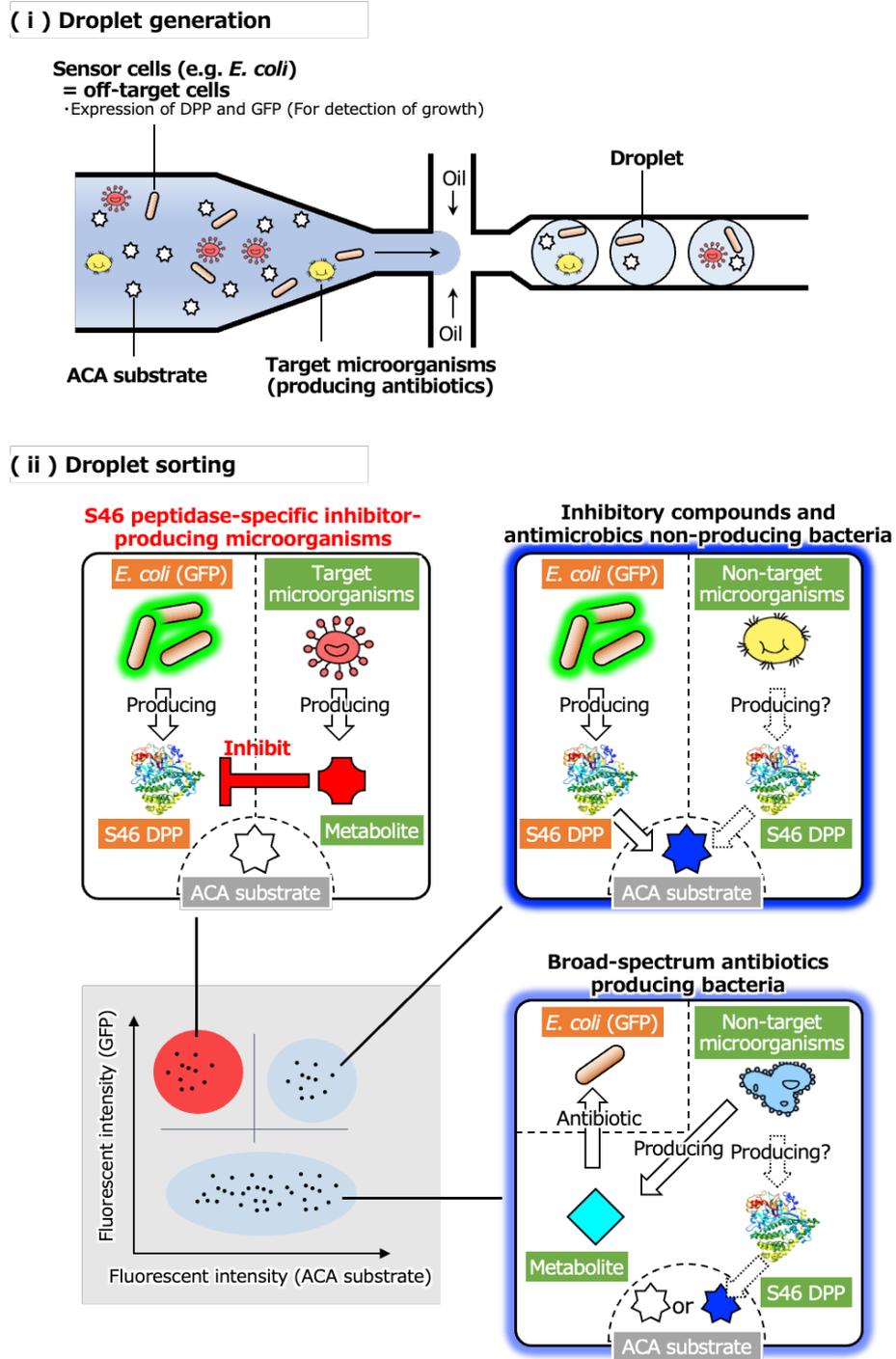


Figure 4-1 | Ultra high-throughput screening methods expected to be constructed in the future. The key point of this screening method is “how to efficiently evaluate and select the inhibitory ability of an inhibitor compound that inhibits DPP7/11 activity without acting on sensor cells (off-target cells)”. Here, we select *E. coli* as a sensor cell and off-target cell because it does not have DPP7/11 and it is easy to construct a transformant expressing S46 peptidases and not NFGNB. *E. coli* is co-transfected with DPP7/11 and GFP expression plasmids. In this way, we can evaluate the inhibition ability of DPP7/11 activity of the target microorganism and select DPP7/11 inhibitor-producing bacteria that do not inhibit the growth of *E. coli* using the fluorescence of GFP as an indicator.

Chapter 5 References

1. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
2. Farha, M. A. & Brown, E. D. Drug repurposing for antimicrobial discovery. *Nat. Microbiol.* **4**, 565–577 (2019).
3. Zhu, Y.-G. *et al.* Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat. Microbiol.* **2**, 16270 (2017).
4. Naas, T. & Nordmann, P. Analysis of a carbapenem-hydrolyzing class A β -lactamase from *Enterobacter cloacae* and of its LysR-type regulatory protein. *Proc. Natl. Acad. Sci.* **91**, 7693 – 7697 (1994).
5. CRC group, Bacteriological tests and bacterial isolate statistics. <http://www.crc-group.co.jp/crc/ff/syukken.html>.
6. Rasmussen, B. A. & Bush, K. Carbapenem-hydrolyzing beta-lactamases. *Antimicrob. Agents Chemother.* **41**, 223–232 (1997).
7. Boucher, H. W. *et al.* 10 × '20 progress - Development of new drugs active against gram-negative bacilli: An update from the infectious diseases society of America. *Clin. Infect. Dis.* **56**, 1685–1694 (2013).
8. Spellberg, B., Powers, J. H., Brass, E. P., Miller, L. G. & Edwards Jr., J. E. Trends in antimicrobial drug development: implications for the future. *Clin. Infect. Dis.* **38**, 1279–1286 (2004).
9. Andrei, S., Valeanu, L., Chirvasuta, R. & Stefan, M.-G. New FDA approved antibacterial drugs: 2015-2017. *Discov. (Craiova, Rom)*. **6**, e81 (2018).
10. El Chakhtoura, N. G. *et al.* Therapies for multidrug resistant and extensively drug-resistant non-fermenting gram-negative bacteria causing nosocomial infections: a perilous journey toward 'molecularly targeted' therapy. *Expert Rev. Anti. Infect. Ther.* **16**, 89–110 (2018).
11. McGowan, J. E. Resistance in nonfermenting gram-negative bacteria: multidrug resistance to the maximum. *Am. J. Infect. Control* **34**, S29–S37(2006).
12. Higgins, C. S. *et al.* Resistance to antibiotics and biocides among non-fermenting Gram-negative bacteria. *Clin. Microbiol. Infect.* **7**, 308–315 (2001).
13. Quinn, J. P. Clinical problems posed by multiresistant nonfermenting Gram-negative pathogens. *Clin. Infect. Dis.* **27**, S117–S124 (1998).
14. Lisa, S. & Jane, S. Infection Control in Cystic Fibrosis. *Clin. Microbiol. Rev.* **17**, 57–71 (2004).
15. Jones, A. M. *et al.* *Burkholderia cenocepacia* and *Burkholderia multivorans*: influence on survival in cystic fibrosis. *Thorax* **59**, 948–951 (2004).
16. Vidal, F. *et al.* Bacteraemia in adults due to glucose non-fermentative Gram-negative bacilli other than *P. aeruginosa*. *QJM An Int. J. Med.* **96**, 227–234 (2003).
17. Spencer, R. C. The emergence of epidemic, multiple-antibiotic-resistant *Stenotrophomonas (Xanthomonas) maltophilia* and *Burkholderia (Pseudomonas) cepacia*. *J. Hosp. Infect.* **30**, 453–464 (1995).
18. Fujita, J. *et al.* Clinical features of *Stenotrophomonas maltophilia* pneumonia in immunocompromised patients. *Respir. Med.* **90**, 35–38 (1996).
19. Varela, M. F. *et al.* Bacterial resistance to antimicrobial agents. *Antibiotics* **10**, 593 (2021).
20. Nemoto, T. K. & Ohara-Nemoto, Y. Exopeptidases and gingipains in *Porphyromonas gingivalis* as prerequisites for its amino acid metabolism. *Jpn. Dent. Sci. Rev.* **52**, 22–29 (2016).
21. Ohara-Nemoto, Y. *et al.* Identification and characterization of prokaryotic dipeptidyl-peptidase 5 from *Porphyromonas gingivalis*. *J. Biol. Chem.* **289**, 5436–5448 (2014).
22. Palleroni, N. J. & Bradbury, J. F. *Stenotrophomonas*, a new bacterial genus for *Xanthomonas maltophilia* (hugh 1980) swings *et al.* 1983. *Int. J. Syst. Bacteriol.* **43**, 606–609 (1993).
23. Brooke, J. S. *Stenotrophomonas maltophilia*: An emerging global opportunistic pathogen. *Clin. Microbiol. Rev.* **25**, 2–41 (2012).
24. Chang, Y. T., Lin, C. Y., Chen, Y. H. & Hsueh, P.-R. Update on infections caused by *Stenotrophomonas maltophilia* with particular attention to resistance mechanisms and therapeutic options. *Frontiers in Microbiology* **6**, 893 (2015).
25. Falagas, M. E. *et al.* Attributable mortality of *Stenotrophomonas maltophilia* infections: a systematic review of the literature. *Future Microbiol.* **4**, 1103–1109 (2009).
26. Brooke, J. S. New strategies against *Stenotrophomonas maltophilia*: a serious worldwide intrinsically drug-resistant opportunistic pathogen. *Expert Rev. Anti. Infect. Ther.* **12**, 1–4 (2014).
27. Al-Anazi, K. A. & Al-Jasser, A. M. Infections caused by *Stenotrophomonas maltophilia* in recipients of hematopoietic stem cell transplantation. *Frontiers in Oncology* **4**, 232 (2014).

28. Mysak, J. *et al.* *Porphyromonas gingivalis*: major periodontopathic pathogen overview. *J. Immunol. Res.* **2014**, 476068 (2014).
29. Bodet, C., Chandad, F. & Grenier, D. Pathogenic potential of *Porphyromonas gingivalis*, *Treponema denticola* and *Tannerella forsythia*, the red bacterial complex associated with periodontitis. *Pathol. Biol.* **55**, 154–162 (2007).
30. Adamowicz, K. *et al.* A6.1 “Intratracheal inoculation of *porphyromonas gingivalis* induces inflammatory changes in the joints of DBA/1 mice.” *Ann. Rheum. Dis.* **73**, A70–A71 (2014).
31. Dominy, S. S. *et al.* *Porphyromonas gingivalis* in Alzheimer’s disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Sci. Adv.* **5**, eaau3333 (2019).
32. Sanai, Y. *et al.* Presence and antibiotic resistance of *Porphyromonas gingivalis*, *Prevotella intermedia*, and *Prevotella nigrescens* in children. *J. Clin. Periodontol.* **29**, 929–934 (2002).
33. Rawlings, N. D. *et al.* The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **46**, D624–D632 (2017).
34. Rouf, S. M. A. *et al.* Discrimination based on Gly and Arg/Ser at position 673 between dipeptidyl-peptidase (DPP) 7 and DPP11, widely distributed DPPs in pathogenic and environmental gram-negative bacteria. *Biochimie* **95**, 824–832 (2013).
35. Sakamoto, Y. *et al.* S46 peptidases are the first exopeptidases to be members of clan PA. *Sci. Rep.* **4**, 4977 (2014).
36. Sakamoto, Y. *et al.* Structural and mutational analyses of dipeptidyl peptidase 11 from *Porphyromonas gingivalis* reveal the molecular basis for strict substrate specificity. *Sci. Rep.* **5**, 11151 (2015).
37. Bezerra, G. A. *et al.* Bacterial protease uses distinct thermodynamic signatures for substrate recognition. *Sci. Rep.* **7**, 1–12 (2017).
38. Ohara-Nemoto, Y. *et al.* Asp- and Glu-specific novel dipeptidyl peptidase 11 of *Porphyromonas gingivalis* ensures utilization of proteinaceous energy sources. *J. Biol. Chem.* **286**, 38115–38127 (2011).
39. Hughes, J. P., Rees, S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *Br. J. Pharmacol.* **162**, 1239–1249 (2011).
40. Hauser, A. S. *et al.* Pharmacogenomics of GPCR drug targets. *Cell* **172**, 41–54.e19 (2018).
41. Rollinger, J. M., Stuppner, H. & Langer, T. Virtual screening for the discovery of bioactive natural products. *Nat. Compd. as drugs* **65**, 211–249 (2008).
42. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **70**, 461–477 (2007).
43. Feher, M. & Schmidt, J. M. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **43**, 218–227 (2003).
44. Macalino, S. J. Y., Gosu, V., Hong, S. & Choi, S. Role of computer-aided drug design in modern drug discovery. *Arch. Pharm. Res.* **38**, 1686–1701 (2015).
45. Grinter, S. Z. & Zou, X. Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design. *Molecules* **19**, 10150–10176 (2014).
46. Lionta, E., Spyrou, G., Vassilatis, K. D. & Cournia, Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current Topics in Medicinal Chemistry* **14**, 1923–1938 (2014).
47. Lavecchia, A. & Giovanni, D. C. Virtual screening strategies in drug discovery: a critical review. *Current Medicinal Chemistry* **20**, 2839–2860 (2013).
48. Anderson, A. C. The process of structure-based drug design. *Chem. Biol.* **10**, 787–797 (2003).
49. Abel, R., Young, T., Farid, R., Berne, B. J. & Friesner, R. A. Role of the active-site solvent in the thermodynamics of factor Xa ligand binding. *J. Am. Chem. Soc.* **130**, 2817–2831 (2008).
50. Melo-Filho, C. C., Braga, C. R. & Andrade, H. C. 3D-QSAR approaches in drug design: perspectives to generate reliable CoMFA models. *Current Computer-Aided Drug Design* **10**, 148–159 (2014).
51. Mayr, L. M. & Bojanic, D. Novel trends in high-throughput screening. *Curr. Opin. Pharmacol.* **9**, 580–588 (2009).
52. Mercola, M., Colas, A. & Willems, E. Induced pluripotent stem cells in cardiovascular drug discovery. *Circ. Res.* **112**, 534–548 (2013).
53. Jentsch, C. *et al.* A phenotypic screen to identify hypertrophy-modulating microRNAs in primary cardiomyocytes. *J. Mol. Cell. Cardiol.* **52**, 13–20 (2012).
54. Cerignoli, F. *et al.* High throughput measurement of Ca²⁺ dynamics for drug risk assessment in human stem cell-derived cardiomyocytes by kinetic image cytometry. *J. Pharmacol. Toxicol. Methods* **66**, 246–256 (2012).
55. Guo, M. T., Rotem, A., Heyman, J. A. & Weitz, D. A. Droplet microfluidics for high-throughput

- biological assays. *Lab Chip* **12**, 2146–2155 (2012).
56. Rosenfeld, L., Lin, T., Derda, R. & Tang, S. K. Y. Y. Review and analysis of performance metrics of droplet microfluidics systems. *Microfluid. Nanofluidics* **16**, 921–939 (2014).
 57. Basova, E. Y. & Foret, F. Droplet microfluidics in (bio)chemical analysis. *Analyst* **140**, 22–38 (2015).
 58. Kaminski, T. S., Scheler, O. & Garstecki, P. Droplet microfluidics for microbiology: techniques, applications and challenges. *Lab Chip* **16**, 2168–2187 (2016).
 59. Takinoue, M., Onoe, H. & Takeuchi, S. Fusion and fission control of picoliter-sized microdroplets for changing the solution concentration of microreactors. *Small* **6**, 2374–2377 (2010).
 60. Sugiura, H. *et al.* Pulse-density modulation control of chemical oscillation far from equilibrium in a droplet open-reactor system. *Nat. Commun.* **7**, 10212 (2016).
 61. Tsuji, Y. *et al.* Droplet split-and-contact method for high-throughput transmembrane electrical recording. *Anal. Chem.* **85**, 10913–10919 (2013).
 62. Lv, T. *et al.* Superhydrophobic surface with shape memory micro/nanostructure and its application in rewritable chip for droplet storage. *ACS Nano* **10**, 9379–9386 (2016).
 63. Autour, A. & Ryckelynck, M. Ultrahigh-throughput improvement and discovery of enzymes using droplet-based microfluidic screening. *Micromachines* **8**, 128 (2017).
 64. Lowe, K. C., Davey, M. R. & Power, J. B. Perfluorochemicals: Their applications and benefits to cell culture. *Trends Biotechnol.* **16**, 272–278 (1998).
 65. Scanga, R. *et al.* Fluorinated amphiphilic block copolymers to stabilize water-in-fluorocarbon emulsions. *Polym. Prepr.* **50**, 148–149 (2009).
 66. Pollack, M. G., Fair, R. B. & Shenderov, A. D. Electrowetting-based actuation of liquid droplets for microfluidic applications. *Appl. Phys. Lett.* **77**, 1725–1726 (2000).
 67. Niu, X., Gulati, S., Edel, J. B. & deMello, A. J. Pillar-induced droplet merging in microfluidic circuits. *Lab Chip* **8**, 1837–1841 (2008).
 68. Park, J. *et al.* On-demand acoustic droplet splitting and steering in a disposable microfluidic chip. *Lab Chip* **18**, 422–432 (2018).
 69. Dong, C. *et al.* A 3D microblade structure for precise and parallel droplet splitting on digital microfluidic chips. *Lab Chip* **17**, 896–904 (2017).
 70. Lyu, F. *et al.* Phenotyping antibiotic resistance with single-cell resolution for the detection of heteroresistance. *Sensors Actuators B Chem.* **270**, 396–404 (2018).
 71. Mahler, L. *et al.* Detection of antibiotics synthesized in microfluidic picolitre-droplets by various actinobacteria. *Sci. Rep.* **8**, 1–11 (2018).
 72. Vallejo, D., Nikoomezar, A., Paegel, B. M. & Chaput, J. C. Fluorescence-activated droplet sorting for single-cell directed evolution. *ACS Synth. Biol.* **8**, 1430–1440 (2019).
 73. Gielen, F. *et al.* Ultrahigh-throughput-directed enzyme evolution by absorbance-activated droplet sorting (AADS). *Proc. Natl. Acad. Sci.* **113**, E7383–E7389 (2016).
 74. Kim, H. S. *et al.* Raman spectroscopy compatible PDMS droplet microfluidic culture and analysis platform towards on-chip lipidomics. *Analyst* **142**, 1054–1060 (2017).
 75. Nakamura, K. *et al.* Culture-independent method for identification of microbial enzyme-encoding genes by activity-based single-cell sequencing using a water-in-oil microdroplet platform. *Sci. Rep.* **6**, 22259 (2016).
 76. Sjostrom, S. L. *et al.* High-throughput screening for industrial enzyme production hosts by droplet microfluidics. *Lab Chip* **14**, 806–813 (2014).
 77. Beneyton, T. *et al.* High-throughput screening of filamentous fungi using nanoliter-range droplet-based microfluidics. *Sci. Rep.* **6**, 1–10 (2016).
 78. Ma, F., Xie, Y., Huang, C., Feng, Y. & Yang, G. An improved single cell ultrahigh throughput screening method based on in vitro compartmentalization. *PLoS One* **9**, e89785 (2014).
 79. Qiao, Y. *et al.* Fluorescence-activated droplet sorting of lipolytic microorganisms using a compact optical system. *Lab Chip* **18**, 190–196 (2018).
 80. He, R., Ding, R., Heyman, J. A., Zhang, D. & Tu, R. Ultra-high-throughput picoliter-droplet microfluidics screening of the industrial cellulase-producing filamentous fungus *Trichoderma reesei*. *J. Ind. Microbiol. Biotechnol.* **46**, 1603–1610 (2019).
 81. Beneyton, T., Coldren, F., Baret, J.-C., Griffiths, A. D. & Taly, V. CotA laccase: high-throughput manipulation and analysis of recombinant enzyme libraries expressed in *E. coli* using droplet-based microfluidics. *Analyst* **139**, 3314–3323 (2014).
 82. Shembekar, N., Chaipan, C., Utharala, R. & Merten, C. A. Droplet-based microfluidics in drug discovery, transcriptomics and high-throughput molecular genetics. *Lab Chip* **16**, 1314–1331 (2016).
 83. Casadevall, A., Dadachova, E. & Pirofski, L. Passive antibody therapy for infectious diseases. *Nat. Rev. Microbiol.* **2**, 695–703 (2004).
 84. Qiu, X. *et al.* Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. *Nature*

- 514**, 47–53 (2014).
85. Carter, P. Improving the efficacy of antibody-based cancer therapies. *Nat. Rev. Cancer* **1**, 118–129 (2001).
 86. Strait, R. T. *et al.* IgG1 protects against renal disease in a mouse model of cryoglobulinaemia. *Nature* **517**, 501–504 (2015).
 87. Kodituwakku, A. P., Jessup, C., Zola, H. & Robertson, D. M. Isolation of antigen-specific B cells. *Immunol. Cell Biol.* **81**, 163–170 (2003).
 88. Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **8**, 870–891 (2013).
 89. Brouzes, E. *et al.* Droplet microfluidic technology for single-cell high-throughput screening. *Proc. Natl. Acad. Sci.* **106**, 14195–14200 (2009).
 90. Rouf, S. M. A. *et al.* Phenylalanine 664 of dipeptidyl peptidase (DPP) 7 and Phenylalanine 671 of DPP11 mediate preference for P2-position hydrophobic residues of a substrate. *FEBS Open Bio* **3**, 177–183 (2013).
 91. Ohara-Nemoto, Y. *et al.* Identification and characterization of prokaryotic dipeptidyl-peptidase 5 from *Porphyromonas gingivalis*. *J. Biol. Chem.* **289**, 5436–5448 (2014).
 92. Takahashi, N., Sato, T. & Yamada, T. Metabolic pathways for cytotoxic end product formation from glutamate- and aspartate-containing peptides by *Porphyromonas gingivalis*. *J. Bacteriol.* **182**, 4704–4710 (2000).
 93. Sakamoto, Y. *et al.* Fragment-based discovery of the first nonpeptidyl inhibitor of an S46 family peptidase. *Sci. Rep.* **9**, 13587 (2019).
 94. Soisson, S. M. *et al.* Structural definition and substrate specificity of the S28 protease family: the crystal structure of human prolylcarboxypeptidase. *BMC Struct. Biol.* **10**, 16 (2010).
 95. Bezerra, G. A. *et al.* Structures of human DPP7 reveal the molecular basis of specific inhibition and the architectural diversity of proline-specific peptidases. *PLoS One* **7**, e43019 (2012).
 96. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
 97. Yung-Chi, C. & Prusoff, W. H. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
 98. Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D* **74**, 85–97 (2018).
 99. Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **66**, 125–132 (2010).
 100. Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Jt. CCP4 +ESF-EAMCB Newsl. Protein Crystallogr.* **26**, (1992).
 101. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. Sect. D* **66**, 22–25 (2010).
 102. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**, 760–763 (1994).
 103. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D* **67**, 355–367 (2011).
 104. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* **60**, 2126–2132 (2004).
 105. Lovell, S. C. *et al.* Structure validation by Ca geometry: ϕ , ψ and C β deviation. *Proteins Struct. Funct. Bioinforma.* **50**, 437–450 (2003).
 106. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
 107. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
 108. Ogasawara, W., Kobayashi, G., Okada, H. & Morikawa, Y. Two types of novel dipeptidyl aminopeptidases from *Pseudomonas* sp. strain WO24. *J. Bacteriol.* **178**, 6288–6295 (1996).
 109. Banbula, A. *et al.* *Porphyromonas gingivalis* DPP-7 represents a novel type of dipeptidylpeptidase. *J. Biol. Chem.* **276**, 6299–6305 (2001).
 110. Suzuki, Y. *et al.* Identification of the catalytic triad of family S46 exopeptidases, closely related to clan PA endopeptidases. *Sci. Rep.* **4**, 1–10 (2014).
 111. Nemoto, T. K., Ono, T. & Ohara-Nemoto, Y. Establishment of potent and specific synthetic substrate for dipeptidyl-peptidase 7. *Anal. Biochem.* **548**, 78–81 (2018).
 112. Sereda, T. J., Mant, C. T., Sönnichsen, F. D. & Hodges, R. S. Reversed-phase chromatography of synthetic amphipathic α -helical peptides as a model for ligand/receptor interactions Effect of changing hydrophobic environment on the relative hydrophilicity/hydrophobicity of amino acid side-chains. *J. Chromatogr. A* **676**, 139–153 (1994).

113. Monera, O. D., Sereda, T. J., Zhou, N. E., Kay, C. M. & Hodges, R. S. Relationship of sidechain hydrophobicity and α -helical propensity on the stability of the single-stranded amphipathic α -helix. *J. Pept. Sci.* **1**, 319–329 (1995).
114. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
115. Ferenczy, G. G. & Keserú, G. M. Thermodynamics guided lead discovery and optimization. *Drug Discov. Today* **15**, 919–932 (2010).
116. Freire, E. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov. Today* **13**, 869–874 (2008).
117. Ohtaka, H., Freire, E., Ã, E. F. & Freire, E. Adaptive inhibitors of the HIV-1 protease. *Prog. Biophys. Mol. Biol.* **88**, 193–208 (2005).
118. Hancock, R. E. W. & Speert, D. P. Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and impact on treatment. *Drug Resist. Updat.* **3**, 247–255 (2000).
119. Newstead, S. *et al.* Crystal structure of a prokaryotic homologue of the mammalian oligopeptide–proton symporters, PepT1 and PepT2. *EMBO J.* **30**, 417–426 (2011).
120. Doki, S. *et al.* Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter POT. *Proc. Natl. Acad. Sci.* **110**, 11343–11348 (2013).
121. Markel, U. *et al.* Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem. Soc. Rev.* **49**, 233–262 (2020).
122. Ota, Y. *et al.* Fluorescent nucleic acid probe in droplets for bacterial sorting (FNAP-sort) as a high-throughput screening method for environmental bacteria with various growth rates. *PLoS One* **14**, e0214533 (2019).
123. Courtois, F. *et al.* Controlling the retention of small molecules in emulsion microdroplets for use in cell-based assays. *Anal. Chem.* **81**, 3008–3016 (2009).
124. Woronoff, G. *et al.* New generation of amino coumarin methyl sulfonate-based fluorogenic substrates for amidase assays in droplet-based microfluidic applications. *Anal. Chem.* **83**, 2852–2857 (2011).
125. Gruner, P. *et al.* Controlling molecular transport in minimal emulsions. *Nat. Commun.* **7**, 10392 (2016).
126. Ma, F. *et al.* Substrate engineering enabling fluorescence droplet entrapment for IVC-FACS-based ultrahigh-throughput screening. *Anal. Chem.* **88**, 8587–8595 (2016).
127. Sandoz, P. A., Chung, A. J., Weaver, W. M. & Di Carlo, D. Sugar additives improve signal fidelity for implementing two-phase resorufin-based enzyme immunoassays. *Langmuir* **30**, 6637–6643 (2014).
128. Goddard, J.-P. & Reymond, J.-L. Recent advances in enzyme assays. *Trends Biotechnol.* **22**, 363–370 (2004).
129. Seixas de Melo, J. S., Becker, R. S. & Macanita, A. L. Photophysical behavior of coumarins as a function of substitution and solvent: experimental evidence for the existence of a lowest lying 1 (n, π^*) state. *J. Phys. Chem.* **98**, 6054–6058 (1994).
130. Goddard, J.-P. & Reymond, J.-L. Enzyme assays for high-throughput screening. *Curr. Opin. Biotechnol.* **15**, 314–322 (2004).
131. Ogasawara, W. *et al.* A novel dipeptidyl aminopeptidase from *Pseudomonas* sp. strain WO24. *J. Bacteriol.* **178**, 1283–1288 (1996).
132. Sakamoto, Y. *et al.* Crystallization and preliminary X-ray crystallographic studies of dipeptidyl aminopeptidase BII from *Pseudoxanthomonas mexicana* WO24. *Acta Crystallogr. Sect. F Structural Biol. Commun.* **70**, 221–224 (2014).
133. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).
134. Takimoto, Y., Hatamoto, M., Ishida, T., Watari, T. & Yamaguchi, T. Fouling development in A/O-MBR under low organic loading condition and identification of key bacteria for biofilm formations. *Sci. Rep.* **8**, 11427 (2018).
135. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
136. Harris, J. L. *et al.* Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc. Natl. Acad. Sci.* **97**, 7754–7759 (2000).
137. Khalfan, H., Abuknesha, R., Rand-Weaver, M., Price, R. G. & Robinson, D. Aminomethyl coumarin acetic acid: a new fluorescent labelling agent for proteins. *Histochem. J.* **18**, 497–499 (1986).
138. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
139. Altschul, S. F. *et al.* Protein database searches using compositionally adjusted substitution matrices.

- FEBS J.* **272**, 5101–5109 (2005).
140. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
 141. Nishimata, H. *et al.* Identification of Dipeptidyl-peptidase (DPP)5 and DPP7 in *Porphyromonas endodontalis*, distinct from those in *Porphyromonas gingivalis*. *PLoS One* **9**, e114221 (2014).
 142. Nakamura, A. *et al.* Structural basis for an exceptionally strong preference for asparagine residue at the S2 subsite of *Stenotrophomonas maltophilia* dipeptidyl peptidase 7. *Sci. Rep.* **11**, 7929 (2021).
 143. Nemoto, T. K. *et al.* Identification of a new subtype of dipeptidyl peptidase 11 and a third group of the S46-family members specifically present in the genus *Bacteroides*. *Biochimie* **147**, 25–35 (2018).
 144. Najah, M. *et al.* New glycosidase substrates for droplet-based microfluidic screening. *Anal. Chem.* **85**, 9807–9814 (2013).
 145. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
 146. Laguerre, M., Saux, M., Dubost, J. P. & Carpy, A. MLPP: A program for the calculation of molecular lipophilicity potential in proteins. *Pharm. Pharmacol. Commun.* **3**, 217–222 (1997).
 147. Kopp, M. R. G. *et al.* Microfluidic shrinking droplet concentrator for analyte detection and phase separation of protein solutions. *Cite This Anal. Chem* **92**, 5803–5812 (2020).
 148. Barnett, A. DPP-4 inhibitors and their potential role in the management of type 2 diabetes. *Int. J. Clin. Pract.* **60**, 1454–1470 (2006).
 149. Juillerat-Jeanneret, L. Dipeptidyl peptidase IV and its inhibitors: therapeutics for type 2 diabetes and what else? *J. Med. Chem.* **57**, 2197–2212 (2014).
 150. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).

Chapter 6 Published paper

1. Nakamura, A.; Suzuki, Y.; Sakamoto, Y.; Roppongi, S.; Kushibiki, C.; Yonezawa, N.; Takahashi, M.; Shida, Y.; Gouda, H.; Nonaka, T.; Tanaka, N.; Ogasawara, W. Structural basis for an exceptionally strong preference for asparagine residue at the S2 subsite of *Stenotrophomonas maltophilia* Dipeptidyl Peptidase 7. *Sci. Rep.* **2021**, 11 (1), 7929. <https://doi.org/10.1038/s41598-021-86965-x>.
2. Nakamura, A.; Honma, N.; Tanaka, Y.; Suzuki, Y.; Shida, Y.; Tsuda, Y.; Hidaka, K.; Ogasawara, W. 7-Aminocoumarin-4-acetic acid as a fluorescent probe for detecting bacterial dipeptidyl peptidase activities in water-in-oil droplets and in bulk. *Anal. Chem.* **In press**. <https://doi.org/10.1021/acs.analchem.1c04108>