Doctorate Dissertation (Censored)
博士論文（要約）

Development of peptide-based covalent protein modifiers by means of *in vitro* selection

（標的タンパク質と共有結合を介して結合するペプチドの試験管内選択法による開発）

A Dissertation Submitted for Degree of Doctor of Philosophy
December 2018
平成 30 年 12 月 博士（理学）申請

Department of Chemistry, Graduate School of Science,
The University of Tokyo
東京大学大学院理学系研究科化学専攻

Naoya Ozawa
小澤 直也

Abstract

Molecules that covalently bind to a native protein of interest are useful for detection and regulation of the protein. Although many such protein modifiers have been developed by embedding a reactive group into a non-covalent ligand scaffold, the conventional methodology often required intense efforts to achieve both selectivity and reactivity. In this thesis, I describe a new methodology for development of covalent protein modifiers with readily tunable reactivity by means of *in vitro* selection. In addition to a model target protein used for the establishment of the methodology, it was applied to another target protein to demonstrate the generality. In addition to the work on covalent modifiers, I also report mathematical modeling of receptor activation induced by dimeric macrocyclic peptides.

Chapter 1 is the general introduction to this thesis. It first describes selective covalent modifiers of native proteins focusing on conventional design strategies. Then, the chapter describes non-standard macrocyclic peptides and the *in vitro* selection system based on which the new methodology was developed. Finally, the purpose of the research in this thesis is briefly explained.

Chapter 2 describes the development of the new methodology for discovery of peptide-based covalent protein modifiers. I constructed libraries of macrocyclic peptides bearing reactive groups and developed an *in vitro* selection scheme for covalent modifiers and applied it to a model target protein. I also demonstrated that the reactivity of a discovered covalent modifier could be readily modulated by rational design.

Chapter 3 describes application of the newly developed selection scheme to epidermal growth factor receptor (EGFR). A selective covalent modifier was successfully identified, demonstrating the generality of the selection scheme.

Chapter 4 describes mathematical modeling of Met activation induced by dimeric macrocyclic peptides. To test the hypothesis that bell-shaped dose-response curves observed for receptor activation induced by dimeric macrocyclic peptide ligands could be

explained by change in stoichiometry of the receptor-ligand complex, I constructed a mathematical model and compared it with the experimental data. The model could reproduce the observed profile and it was confirmed that the proposed stoichiometric change agreed with the model.

Chapter 5 is the general conclusion of this thesis. It summarizes achievements in this thesis and discusses perspectives.

Contents

Chapter 1

General introduction

1.1 Selective covalent modifiers of native proteins

1.1.1 Covalent drugs

Covalent drugs are drug molecules that covalently, and often irreversibly, bind to a target protein. Although reactive molecules had been associated with idiosyncratic toxicity induced by non-specific modification of endogenous molecules[1,2] and thus had been usually avoided in drug discovery programs[3], many successful drugs have been found to act in covalent mechanisms[4]. For example, an anti-inflammatory drug, aspirin, acetylates a serine residue near the active site of cyclooxygenase to block binding of the natural substrate, thereby irreversibly inhibiting the enzyme[5] (Figure 1.1a). Another example is penicillin antibiotics, which covalently bind to an active site serine of peptidoglycan transpeptidases to inhibit cell-wall synthesis[6] (Figure 1.1b).

Compared to conventional non-covalent drugs, covalent drugs have three major advantages. First, covalent binding allows exceptionally high affinity and thus high potency[7]. Irreversible inhibitors may even maintain activity against drug-resistant mutations which increase the affinity to a natural substrate, because, once the covalent bond is established, irreversible inhibitors do not compete with the reversible binding of the natural substrate[8]. Second, proteins inactivated by an irreversible inhibitor must be resynthesized to restore the activity, which enables prolonged duration of action and less frequent dosing[9]. Third, covalent binding can recognize the presence of an appropriate target residue at a specific position on the target protein, resulting the potential for high selectivity among a family of proteins[10,11].

Because of these advantages, a class of covalent drugs referred to as targeted covalent inhibitors (TCIs) has been developed recently[12,13]. In contrast to earlier covalent drugs which were discovered by serendipity, TCIs are specifically designed to covalently bind to a specific target protein thereby blocking its active site. Many such molecules have been developed against various target proteins[13,14], and some of them have been approved by FDA (Figure 1.1c).

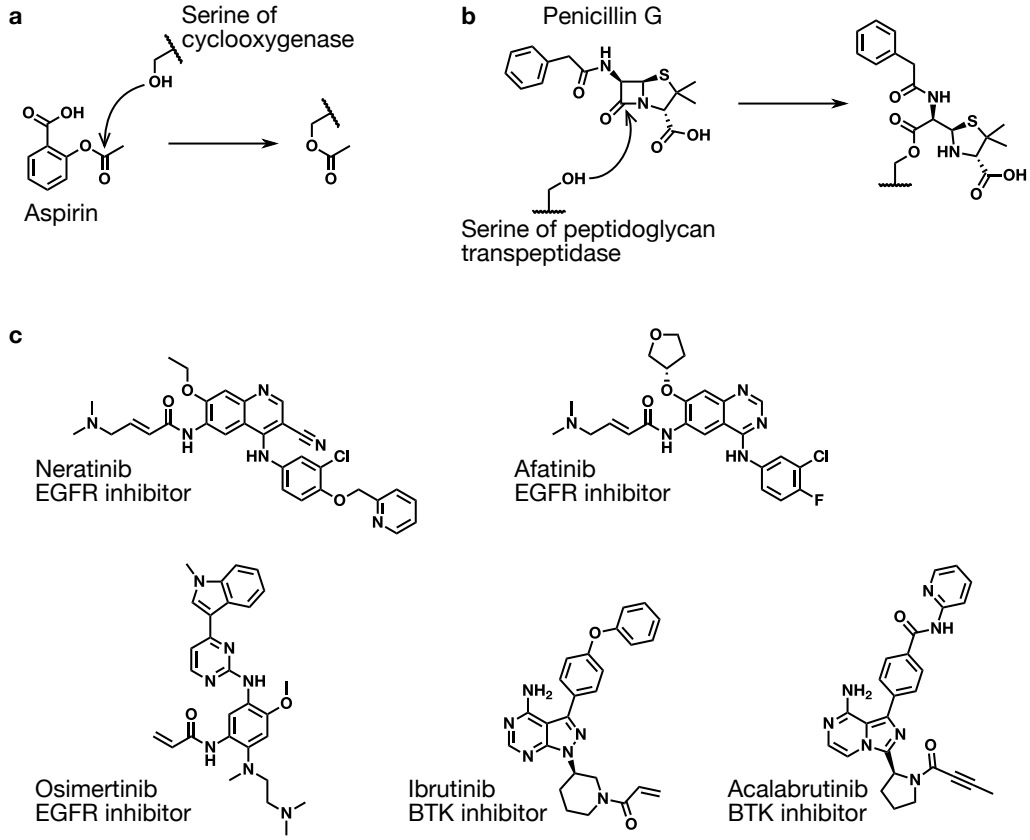


Figure 1.1 | Examples of covalent drugs
(a) Mechanism of action of aspirin.
(b) Mechanism of action of penicillin G.
(c) Examples of FDA-approved targeted covalent inhibitors.

1.1.2 Labeling of endogenous proteins

Selective labeling of a native protein of interest in complex mixtures such as live cells and cell lysates is useful for detection and functionalization of that protein. To this end, various techniques have been developed. The most widely used technique is expression of the protein as a fusion with a fluorescent protein such as GFP[15], which has been extensively used in live cell imaging. Another common technique is fusion with a protein-tag or peptide-tag which can be selectively labeled with a cognate reagent. These tags include SNAP-tag[16], CLIP-tag[17], HaloTag[18] and a tetracysteine motif for FlAsH labeling[19]. Although useful, these techniques require genetic manipulation, and cannot be applied to endogenous proteins.

As an alternative strategy, molecules that can recognize an endogenous protein of interest to selectively label it have been developed and application of some of them for detection and functionalization of a protein in live cells have been demonstrated. For example, an antibody equipped with a reactive group could selectively label cells expressing ErbB2 on the surface, making it detectable by fluorescence[20] (Figure 1.2a). Due to the irreversible covalent binding, the labeling could survive stringent washing in which labeling with normal antibody was lost. In another example, a reagent containing ^{19}F was developed for labeling of carbonic anhydrase in cells[21,22] (Figure 1.2b). The labeled protein worked as a biosensor for ligands of the protein whose response could be detected by the change in ^{19}F NMR chemical shift.

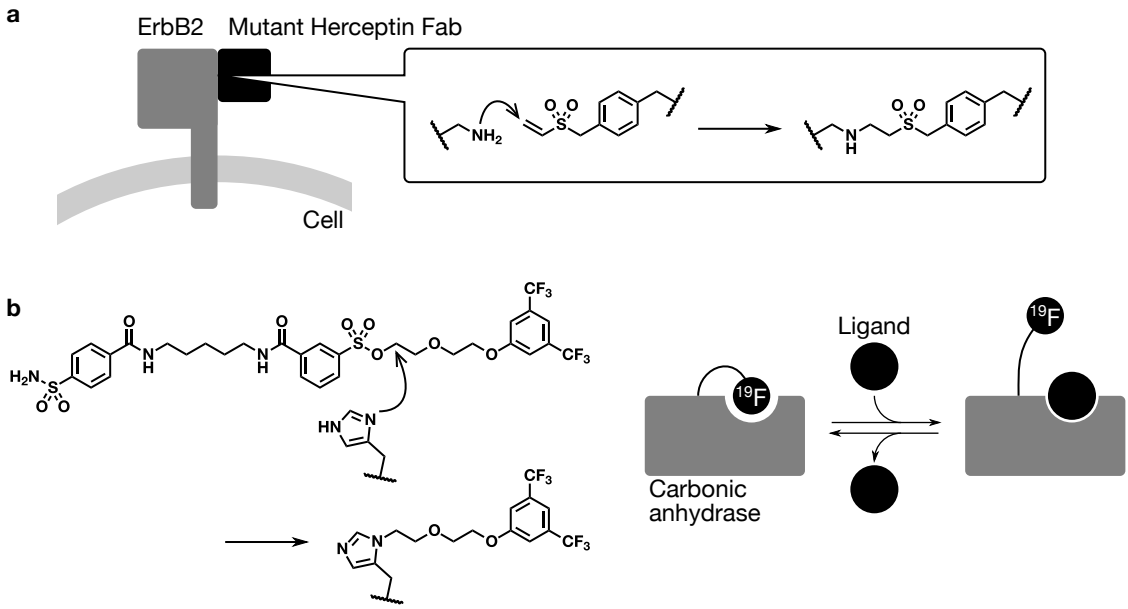


Figure 1.2 | Examples of covalent labeling
(a) Labeling of ErbB2 on cell surface by an antibody equipped with a reactive group.
(b) Biosensor created by labeling of carbonic anhydrase with ^{19}F containing reagent.

1.1.3 Development strategies

The most commonly used strategy involves two steps (Figure 1.3a). First, a selective non-covalent ligand scaffold is identified. Then, a reactive group of moderate reactivity (usually an electrophile) is introduced at an appropriate position of the scaffold. When the molecule is associated with the protein, the reactive group is positioned near an appropriate amino acid on the protein surface (usually a nucleophile) and covalent bond formation is induced by proximity effect. For example, development of a TCI against Bruton's tyrosine kinase, ibrutinib, started with identification of a non-covalent inhibitor of BTK[10]. Then, a three-dimensional structural model of the non-covalent complex was constructed based on known X-ray structure of a ligand bound state of a similar protein, and a cysteine poorly conserved across the kinase family was selected for the target of covalent modification. Finally, several candidate compounds were designed by embedding a reactive group in the scaffold so that the reactive group would react with the cysteine, and their activities were tested by *in vitro* and *in vivo* assays. In addition to small molecule ligands, this strategy has also been applied to various ligand scaffolds including stapled peptide ligands[23,24], an antibody[20] and an affibody[25].

Although the above strategy is useful for development of TCIs and labeling reagents, the target protein is usually inactivated because the active site is irreversibly blocked by the ligand scaffold. Thus, the strategy is not suitable for functional analysis of the target protein. To address this issue, an improved strategy called ligand-directed chemistry was developed[21,26,27] (Figure 1.3b). Like the previous strategy, the development starts with a selective non-covalent ligand scaffold. Instead of directly embedding a reactive group, a cleavable reactive group bearing a probe molecule is attached. When it reacts with a nucleophile on the target protein, the probe is covalently attached to the target protein, and the ligand is released as a leaving group, thereby restoring the activity of the target protein.

In addition to the strategies relying on known ligand scaffolds, there are several reports of screening-based discoveries of covalent modifiers. One example employed a technique called disulfide tethering[28] as the initial step for development of a TCI against K-Ras(G12C)[29] (Figure 1.3c). First, a library of disulfide-containing molecules was screened to identify a molecule that could label a cysteine on the protein. Then, using this molecule as a starting point, an irreversible inhibitor was developed. In another example,

mixtures of reactive molecules were incubated with papain and analyzed by mass spectrometry to identify covalent modifiers for the protein[30]. This strategy, named covalent tethering, has been applied to several other proteins[31]. There is also an isolated example where libraries of reactive small molecules encoded by peptide nucleic acids were employed to identify covalent modifiers for MEK2 and ErbB2[32].

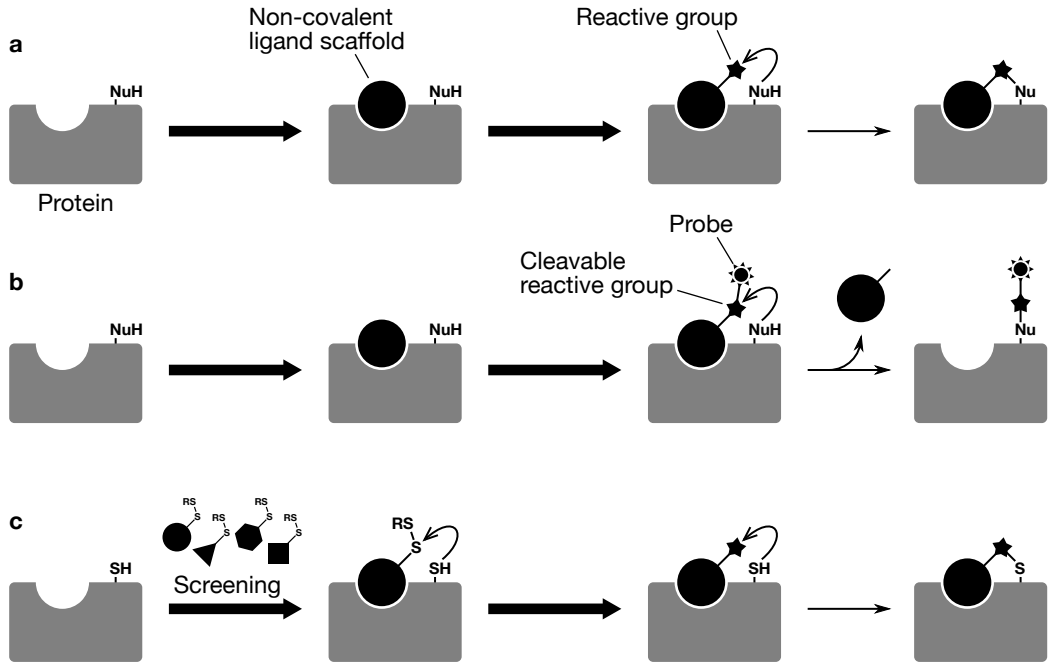


Figure 1.3 | Development strategies for covalent modifiers
(a) Identification of a non-covalent ligand scaffold followed by introduction of a reactive group.
(b) Ligand directed chemistry. Cleavable reactive group bearing a probe is attached onto a non-covalent ligand scaffold. Upon reaction with the target protein, the probe is covalently attached to the target protein and the ligand scaffold is released.
(c) Disulfide tethering. A disulfide-containing molecule which react with a cysteine on the target protein is converted to a covalent modifier.

1.2 Macrocyclic peptides

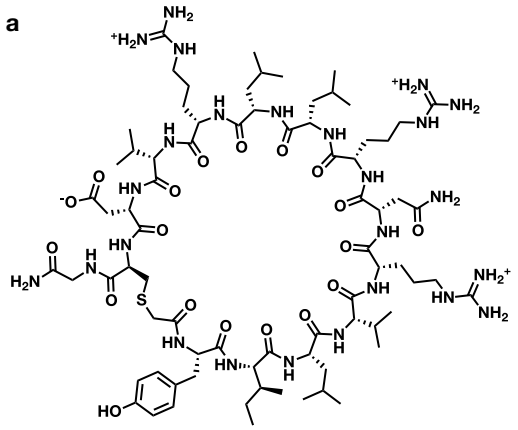
1.2.1 Macrocyclic peptides as ligand scaffold

Macrocyclic peptides are an attractive class of ligand scaffolds[33,34]. In contrast to conventional small molecules, whose small size limits the range of targetable proteins[35,36] and selectivity among proteins of related functionalities[37], peptides have larger interaction surface and thus can target wider range of proteins with high selectivity[38]. While antibodies can also target various proteins with high selectivity, they have higher production cost compared to chemically synthesized small molecules. On the other hand, peptides can be readily synthesized by established solid phase synthesis methods. In addition to these advantages arising from peptides, the macrocyclic structure offers two additional benefits. First, macrocyclization can increase the binding affinity by reducing the entropic cost of binding because the constrained structure limits the number of conformations available at the unbound state[39,40]. Second, macrocyclization can increase passive membrane permeability because the constrained structure facilitates the formation of intramolecular hydrogen bonding in hydrophobic media so that the energetic cost of desolvation can be reduced[41]. Some macrocyclic peptides are even orally bioavailable[42].

Because of these potential advantages, several methodologies for efficient screening of artificial macrocyclic peptides have been developed, and ligands for various target proteins have been successfully discovered[43]. For example, a macrocyclic peptide inhibitor with high selectivity was discovered by a screening against a kinase, Akt2 (Figure 1.4). The peptide could inhibit Akt2 with the IC_{50} value of 110 nM, while it showed ≥40-fold higher IC_{50} values for other members of the Akt family (Akt1 and Akt3) and other kinases (PKA and SGK).

The following sections describe one of the most successful methodologies for screening of artificial macrocyclic peptides.

a



b

Kinase	IC_{50} (nM)
Akt2	110
Akt1	25,000
Akt3	4,200
PKA	>10,000
SGK	>10,000

Figure 1.4 | Akt2-selective macrocyclic peptide inhibitor
(a) Structure of the Akt2 inhibitor.
(b) Inhibitory activities against members of Akt family and other kinases.

1.2.2 FIT system for ribosomal synthesis of macrocyclic peptides

Translation is a biosynthetic reaction of proteins, where peptides are synthesized by ribosome using mRNA as the template[44]. Each amino acid is encoded as a triplet of nucleotides (codon) on mRNA, and the recognition of codons is mediated by specific hybridization with the anticodons on the cognate tRNAs. Each tRNA carries a cognate amino acid which has been attached by an aminoacyl tRNA synthetase. Combination of these recognition points realizes translation of mRNA sequences into amino acid sequences.

Genetic code reprogramming is a technique that allows assigning noncanonical amino acids to arbitrary codons so that the noncanonical amino acids can be incorporated into ribosomally synthesized peptides[45–47]. This can be achieved by removing aminoacyl-tRNAs for these codons from the translation system and supplementing it with tRNAs charged with noncanonical amino acids so that the noncanonical aminoacyl-tRNAs are used in the translation. A translation system lacking some aminoacyl-tRNAs can be readily prepared using an *in vitro* translation system reconstituted with purified components[48] in which cognate amino acids are omitted.

As a facile and general methodology for preparation of aminoacyl-tRNA, ribozymes called flexizymes were developed[47,49]. Flexizymes are highly promiscuous catalyst which can aminoacylate tRNAs using amino acids activated by a leaving group (Figure 1.5a). Several flexizymes with different substrate scopes have been developed. eFx accepts cyanomethyl ester of amino acids with aromatic side chains, while dFx accepts 3,5-dinitrobenzyl ester of diverse amino acids. By introducing aminoacyl-tRNAs prepared by flexizymes into a customized *in vitro* translation system in which cognate canonical amino acids are omitted, noncanonical amino acids can be incorporated into peptides. This system, referred to as the flexible *in vitro* translation (FIT) system, have been successfully applied to various noncanonical amino acids including D-amino acids[50–52], β-amino acids[53], *N*-alkylated amino acids[54,55] and α-hydroxy acids[56,57].

The high substrate tolerance of the FIT system allows for ribosomal synthesis of macrocyclic peptides by incorporating amino acids capable of intramolecular reaction with another amino acid. One of the most convenient and reliable example is translation initiated with *N*-chloroacetylated amino acids, where the N-terminal chloroacetyl group spontaneously react with a downstream cysteine to form a macrocyclic structure closed by a

stable thioether linkage[58,59] (Figure 1.5b). Chloroacetyl group has also been incorporated into side chain to form an inter-side-chain bond[60]. Other cyclization chemistries employed by the FIT system include oxidative coupling[61], click chemistry[62] and Michael addition[63]. Even bicyclic structures can be produced by combining two cyclization chemistries[62].

Taken together, the FIT system can synthesize macrocyclic peptides, optionally containing multiple non-proteinogenic amino acids, in a template dependent manner. This feature facilitates construction of libraries of macrocyclic peptides with high diversity.

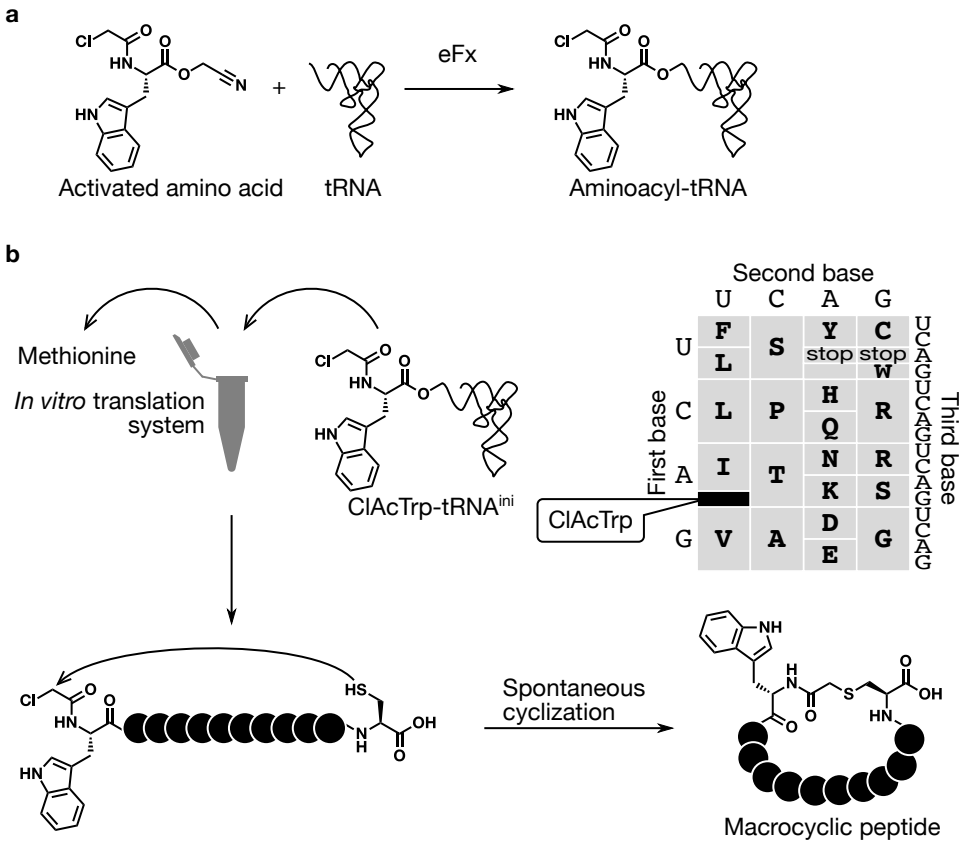


Figure 1.5 | Flexible *in vitro* translation system
(a) Flexizyme-mediated aminoacylation of tRNA.
(b) Ribosomal synthesis of macrocyclic peptide. Initiator tRNA ($tRNA^{ini}$) charged with chloroacetyl tryptophan (ClAcTrp) is added to methionine deficient *in vitro* translation system to assign ClAcTrp to the initiator AUG codon. Peptides synthesized by this reprogrammed translation system has chloroacetyl group on the N-terminus, and it spontaneously reacts with the downstream cysteine to form a macrocyclic structure closed by a thioether linkage.

1.2.3 RaPID system for screening of macrocyclic peptides

To efficiently screen macrocyclic peptides produced by the FIT system, it was combined with an *in vitro* selection technique known as mRNA display[64,65]. The resulting system, referred to as the random non-standard peptide integrated discovery (RaPID) system, allows discovery of macrocyclic peptides selectively binding to a given target protein from a library of trillions of peptides[66]. The system works as follows (Figure 1.6). First, a mixture of mRNA containing a random sequence is prepared, and a linker bearing a puromycin moiety is ligated to the 3′ end. The mRNA library is translated by the FIT system to prepare a library of macrocyclic peptides. At the end of the translation reaction, the puromycin moiety on the mRNA is incorporated into the peptide to form a covalent linkage between the peptide and the mRNA coding for the sequence. Then, the peptide-mRNA conjugate is converted into peptide-mRNA/cDNA conjugate by reverse transcription. The resulting conjugates are incubated with a target protein immobilized on magnetic beads, and the binding fraction is recovered. Finally, the cDNA in the recovered fraction is amplified by PCR and transcribed to generate an mRNA library enriched with mRNAs coding for target-binding peptides. Usually this selection cycle is repeated several times to obtain strong binders. After the selection, candidate peptides can be identified by sequencing the cDNA.

The RaPID system has been successfully applied to various target proteins to identify macrocyclic peptides with high affinity and selectivity. Successfully targeted proteins include a ubiquitin ligase[66], a deacetylace[67], a kinase[68], a demethylase[69], a mutase[70], a glycosidase[71], transporters[72,73] and receptors[74,75].

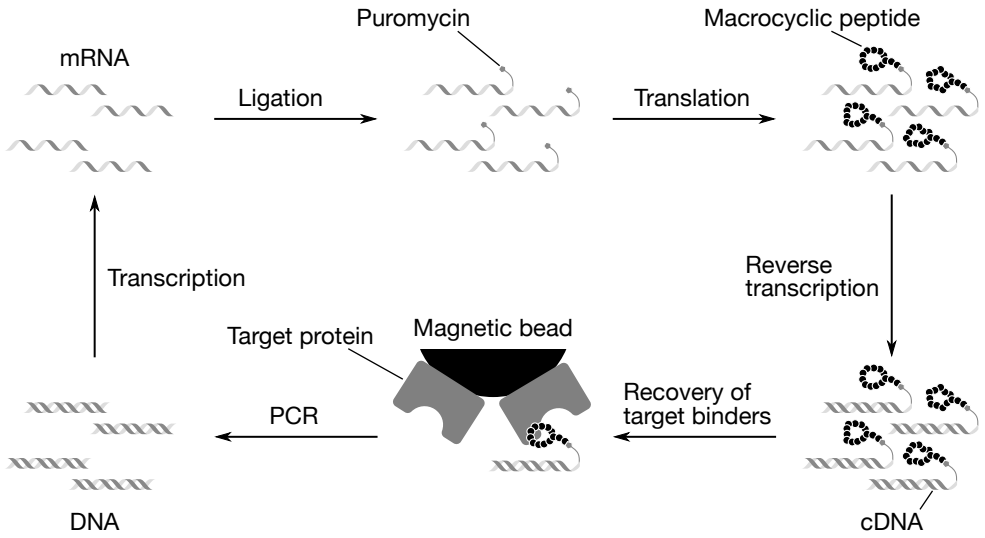


Figure 1.6 | Scheme of the RaPID system
Macrocyclic peptides are synthesized by the FIT system, and covalently linked with cognate mRNA through the puromycin moiety. mRNA is reverse transcribed to synthesize cDNA, and the target binders are recovered using target protein immobilized on magnetic beads. The cDNA of the binders is amplified by PCR and transcribed to generate mRNA for the next round of selection. Strong binders can be selected by repetition of this cycle.

1.3 Purpose of this study

Although covalent modifiers are useful as drugs and labeling reagents, the conventional strategy for development of selective covalent modifiers suffers from three challenges, namely, identification of selective non-covalent ligand scaffolds, embedding reactive groups with appropriate orientation, and fine tuning of the reactivity. Because of these challenges, development of selective covalent inhibitors often requires intense efforts. To overcome these challenges, my research aimed at establishing a new general methodology for development of covalent modifiers by exploiting the high selectivity and the screening system of macrocyclic peptides. In Chapter 2, I developed the methodology using a model target protein. In Chapter 3, the methodology was applied to another target protein, EGFR, to demonstrate its generality.

In addition to the work on covalent modifiers, I also worked on theoretical consideration for activation of a receptor, Met, induced by dimeric peptides. The dimeric peptides had been developed by homodimerizing macrocyclic peptide ligands discovered by the RaPID system. Although the dimeric peptides could induce the dimerization and activation of the receptor, they showed bell-shaped dose response curves, which was counter-intuitive. To explain such phenomena, a mechanism had been proposed in which the stoichiometry of the receptor-ligand complex was affected by the concentration of the ligand. In Chapter 3, I constructed a mathematical model and compared it to the experimental data.

Chapter 2

Development of a methodology for discovery of peptide-based covalent protein modifiers

This chapter is not shown due to reasons involving patent applications and future publications.

Chapter 3

Discovery of a covalent modifier for the extracellular region of EGFR

This chapter is not shown due to reasons involving patent applications and future publications.

Chapter 4

Mathematical modeling of activation of Met induced by dimeric peptides

4.1 Introduction

Met, also known as hepatocyte growth factor receptor (HGFR) or c-Met, is a receptor tyrosine kinase, and it recognizes hepatocyte growth factor (HGF)[92]. When the extracellular region of Met binds to HGF, the receptor dimerizes, and the kinase activity of the intracellular region is activated by trans-phosphorylation. The activation of Met initiates several signal transduction pathways promoting cell proliferation, survival and migration. Because Met-mediated signaling plays an important role in regeneration and protection of tissues, molecules that can activate Met are potentially useful as a therapeutic for various diseases[93,94].

Previously, artificial Met agonists based on macrocyclic peptides was devloped[75]. The agonists were developed in two steps: first, macrocyclic peptides which can bind to the extracellular region of Met were discovered by the RaPID system; then, the peptides were homodimerized through linkers of appropriate lengths. The resulting dimers could induce dimerization and phosphorylation of Met[75,95]. However, the activation profiles showed counterintuitive dose response curves. Increasing concentrations of dimeric peptides increased the level of Met dimerization and phosphorylation of Met at low concentrations (below ~100 nM), whereas the level of dimerization and phosphorylation decreased at higher concentrations.

This bell-shaped profile has also been reported for several other combinations of a receptor and a bivalent ligand, and it has been proposed that it was caused by the change in stoichiometry of the receptor-ligand complex[96]. According to this model, the ligand-

induced receptor dimerization involves two steps: first, receptor (R) and ligand (L) form 1:1 complex (RL), and then the 1:1 complex binds with another receptor molecule to form 2:1 complex (R_2L). At high concentrations of the ligand, the formation of 1:1 complex is favored over 2:1 complex, resulting in the decrease of 2:1 complex (Figure 4.1).

To confirm that the above model could explain the bell-shaped response of Met phosphorylation by the dimeric peptides, I constructed a simple mathematical model and compared the predictions by the model with the experimental data.

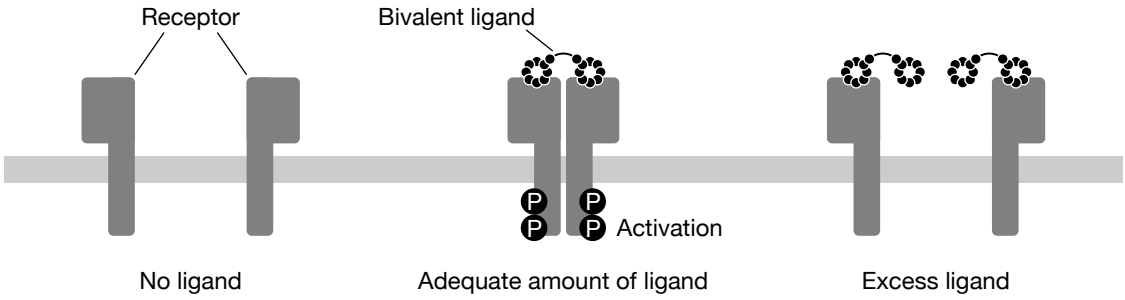


Figure 4.1 | Proposed mechanism of bell-shaped activation profile induced by bivalent ligands
Adequate amount of bivalent ligand induces the formation of 2:1 complex between the receptor and the ligand thereby activating the receptor, while excess ligand induces inactive 1:1 complex.

4.2 Results and discussion

4.2.1 The model

Although a mathematical model for complex formation based on reaction kinetics had been reported[97], I could obtain equivalent equations from a simple model based on equilibrium. The model assumed that the complex formation could be expressed by the following reactions:

$$\mathrm{R} + \mathrm{L} \rightleftharpoons \mathrm{RL} \quad K_1 \tag{4.1}$$

$$\mathrm{R} + \mathrm{RL} \rightleftharpoons \mathrm{R_2L} \quad K_2 \tag{4.2}$$

where K_1 and K_2 were the equilibrium constants specific to the ligand. In addition to the affinity between the peptide and the receptor, K_2 would be affected also by the orientation of the unbound side of the dimeric peptide bound to a receptor and thus should depend on the linker between the cyclic peptide units. At equilibrium, this model predicted that the concentration of R, RL and R_2L on cell surface ([R], [RL] and [R_2L]) could be expressed using the concentration of L ([L]) as follows:

$$[\mathrm{R}] = \frac{2[\mathrm{R}]_{\mathrm{total}}}{K_1[\mathrm{L}] + 1 + \sqrt{(K_1[\mathrm{L}] + 1)^2 + bK_1[\mathrm{L}]}}, \tag{4.3}$$

$$[\mathrm{RL}] = \frac{2K_1[\mathrm{L}][\mathrm{R}]_{\mathrm{total}}}{K_1[\mathrm{L}] + 1 + \sqrt{(K_1[\mathrm{L}] + 1)^2 + bK_1[\mathrm{L}]}}, \tag{4.4}$$

$$[\mathrm{R_2L}] = \frac{bK_1[\mathrm{L}][\mathrm{R}]_{\mathrm{total}}}{2\left[K_1[\mathrm{L}] + 1 + \sqrt{(K_1[\mathrm{L}] + 1)^2 + bK_1[\mathrm{L}]}\right]^2}, \tag{4.5}$$

where b ($= 8K_2[\mathrm{R}]_{\mathrm{total}}$) and $[\mathrm{R}]_{\mathrm{total}}$ were constants (see Materials and methods for details on the derivation).

To predict the phosphorylation level of the receptor, I hypothesized that the amount of phosphorylated receptor (y) would be proportional to the amount of the receptor dimer ([R_2L]). Because the phosphorylation is triggered by the dimerization of the receptor, larger amount of receptor dimer would cause larger amount of phosphorylated receptor. For the sake of simplicity, I assumed that y could be expressed as follows:

$$y = a[\mathrm{R_2L}], \tag{4.6}$$

where a was a constant. The value of a might depend on the ligand, because different ligands might induce different structures of the receptor dimer and thus different efficiencies of the phosphorylation reaction. From Equation 4.5, y could be expressed as a function of [L]:

$$y = \frac{cbK_1[\mathrm{L}]}{2\left[K_1[\mathrm{L}] + 1 + \sqrt{(K_1[\mathrm{L}] + 1)^2 + bK_1[\mathrm{L}]}\right]^2}, \tag{4.7}$$

where $c = a[\mathrm{R}]_{\mathrm{total}}$.

To confirm that the above model gave a bell-shaped curve, Equation 4.7 was differentiated:

$$\frac{\mathrm{d}y}{\mathrm{d}[\mathrm{L}]} = \frac{cb^2K_1[\mathrm{L}](1 - K_1[\mathrm{L}])}{2\left[K_1[\mathrm{L}] + 1 + \sqrt{(K_1[\mathrm{L}] + 1)^2 + bK_1[\mathrm{L}]}\right]^4\sqrt{(K_1[\mathrm{L}] + 1)^2 + bK_1[\mathrm{L}]}}. \tag{4.8}$$

The derivative was positive when $0 < [\mathrm{L}] < K_1$ and negative when $K_1 < [\mathrm{L}]$. This result indicated that the model gave a bell-shaped curve with a peak at $[\mathrm{L}] = 1/K_1$.

The height of the peak (i.e. the maximal phosphorylation level that could be achieved with the ligand) was

$$y_{\max} = y_{|[\mathrm{L}]=1/K_1} = \frac{cb}{2\left(2 + \sqrt{4 + b}\right)^2} = \frac{aK_2([\mathrm{R}]_{\mathrm{total}})^2}{\left(1 + \sqrt{1 + 2K_2[\mathrm{R}]_{\mathrm{total}}}\right)^2}. \tag{4.9}$$

The fraction of phosphorylated receptor at the peak was

$$\frac{y_{\max}}{[\mathrm{R}]_{\mathrm{total}}} = \frac{aK_2[\mathrm{R}]_{\mathrm{total}}}{\left(1 + \sqrt{1 + 2K_2[\mathrm{R}]_{\mathrm{total}}}\right)^2}. \tag{4.10}$$

Plotting them as a function of $[\mathrm{R}]_{\mathrm{total}}$ revealed that higher $[\mathrm{R}]_{\mathrm{total}}$ gave higher $y_{\max}$ and $y_{\max}/[\mathrm{R}]_{\mathrm{total}}$ (Figure 4.2), suggesting that cells with higher density of the receptor would be more susceptible to the ligand.

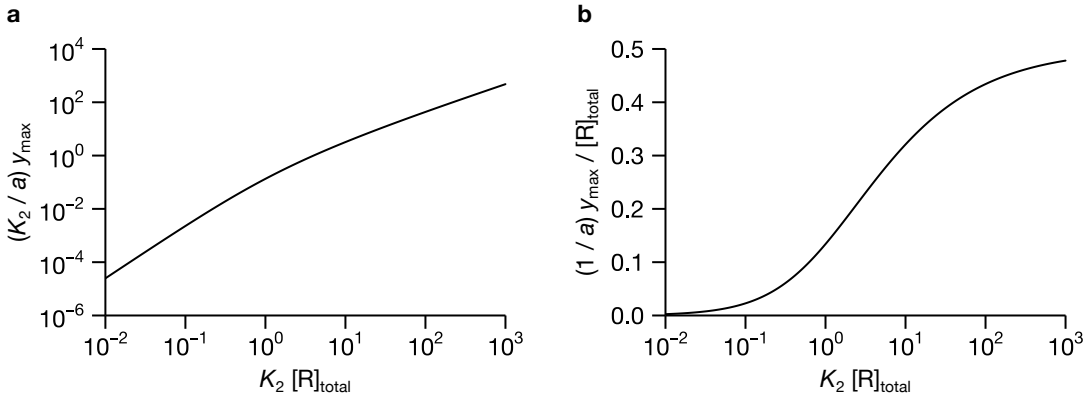


Figure 4.2 | Predicted maximal phosphorylation level of Met
The amount (a) and the fraction (b) of phosphorylated Met predicted by Equation 4.9 and Equation 4.10 was plotted as a function of receptor density on cell surface.

4.2.2 Comparison of experimental data and model

To confirm that the model could explain the observed bell-shaped phosphorylation profiles, appropriate values for parameters K_1, b and c were determined by least squares fitting to experimentally obtained y values (Figure 4.3 left). The experimental values were obtained by incubating cells with various concentrations of dimeric peptides (aML5-PEG3, aMD4-PEG11 and aMD5-PEG11) for 10 min followed by quantification of the amount of phosphorylated Met by ELISA. Although the K_1 values determined by the fitting (3.8, 5.0 and 4.2 μM^{-1}) did not match the values expected from the K_D values determined by surface plasmon resonance (SPR)[75] ($1/K_D$ = 530, 1100 and 480 μM^{-1}), the disagreement might be caused by the environmental difference between the cell surface and the SPR sensor surface. Using the parameters determined by fitting, theoretical curves for y were calculated, and it was confirmed that the model could reproduce the observed bell-shaped profiles (Figure 4.3 left).

The model also allowed estimation of the proportions of receptor states ([R] : [RL] : [R_2L]) (Figure 4.3 right). At lower concentration of L, increasing [L] caused increase of [R_2L], while at higher concentration of L, [R_2L] decreased and [RL] increased. These results agreed with the proposed mechanism in which 1:1 complex is favored at higher concentration of L. This observation can also be interpreted qualitatively. When the equilibriums for the reactions described by Equation 4.1 and Equation 4.2 are established, a combination of the reactions,

$$2\,RL \rightleftharpoons R_2L + L \tag{4.11}$$

is also in equilibrium. According to Le Châtelier's principle, increasing the concentration of L induces the leftward reaction, making RL more favored over R_2L. Therefore, with excess concentration of L, the active R_2L complex is converted to inactive RL, suppressing the formation of phosphorylated receptor.

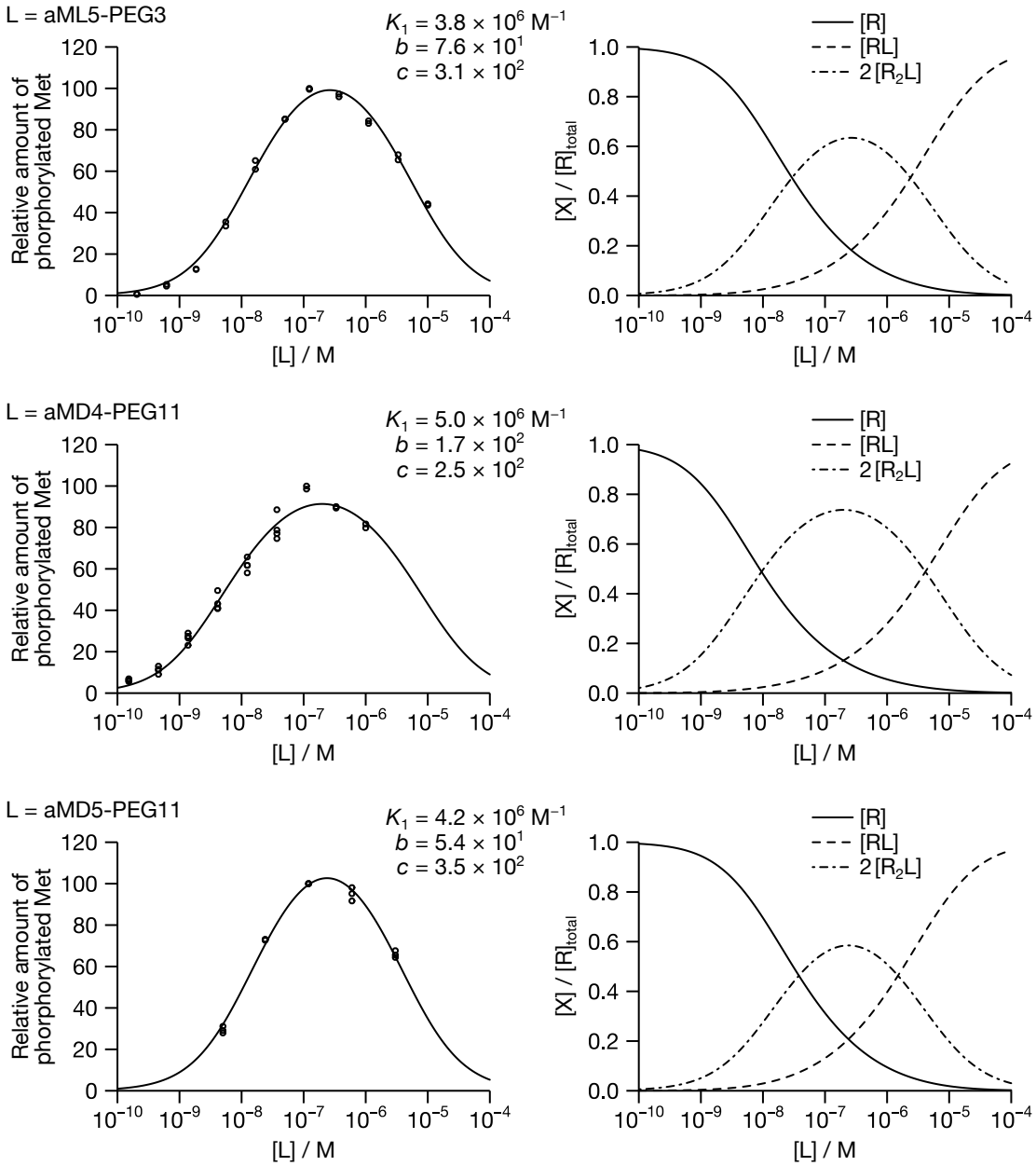


Figure 4.3 | Comparison of the model and experimental data for Met phosphorylation
The theoretical curve of Met phosphorylation (Equation 4.7) as a function of the concentration of dimeric peptide ([L]) was fitted to experimental values (open circles) by least squares method, and the parameters K_1, b and c were determined (left). Using these parameters, the relative abundance of R, RL and R_2L were calculated (right).

4.3 Conclusion

In this chapter, I have described my work on mathematical modeling of the bell-shaped profiles observed in phosphorylation level of Met induced by dimeric peptides. Curve-fitting to the experimental data confirmed that the model could reproduce the observed behavior. Numerical calculations also confirmed that, according to the model, the decline of the active 2:1 complex between the receptor and ligand at excess amount of the ligand is accompanied by increase of inactive 1:1 complex. This behavior can be interpreted qualitatively based on Le Châtelier's principle. The model also predicted that cells with higher density of Met is more susceptible to dimeric peptides. Although detailed experimental confirmation remains to be performed, the model described in this chapter provides deeper understanding of the receptor dimerization induced by homodimeric ligands.

4.4 Materials and methods

Derivation of the equations

The model for dimerization of receptor (R) caused by bivalent ligand (L) is described by the two reactions,

$$\mathrm{R} + \mathrm{L} \rightleftharpoons \mathrm{RL} \quad K_1, \tag{4.12}$$

$$\mathrm{R} + \mathrm{RL} \rightleftharpoons \mathrm{R_2L} \quad K_2, \tag{4.13}$$

where K_1 and K_2 are the equilibrium constants. Assuming that equilibriums are established for these reactions,

$$\frac{[\mathrm{RL}]}{[\mathrm{R}][\mathrm{L}]} = K_1, \qquad \frac{[\mathrm{R_2L}]}{[\mathrm{R}][\mathrm{RL}]} = K_2, \tag{4.14}$$

where [R], [RL] and [R_2L] are two-dimensional concentrations on the cell surface, and [L] is three-dimensional concentration in the solution. The unit for, the three-dimensional concentration, K_1 and K_2 are mol dm^{-2}, mol dm^{-3}, mol^{-1} dm^3 and mol^{-1} dm^2, respectively. Note that the equilibrium constants. From Equation 4.14, the concentration of RL and R_2L can be expressed using that of R and L as follows:

$$[\mathrm{RL}] = K_1[\mathrm{R}][\mathrm{L}], \qquad [\mathrm{R_2L}] = K_1K_2[\mathrm{R}]^2[\mathrm{L}]. \tag{4.15}$$

Because the number of the receptor molecules stays constant in the model,

$$[\mathrm{R}] + [\mathrm{RL}] + 2[\mathrm{R_2L}] = [\mathrm{R}]_{\mathrm{total}}, \tag{4.16}$$

where $[\mathrm{R}]_{\mathrm{total}}$ denotes the concentration of R in the absence of the ligand. Using Equation 4.15, Equation 4.16 can be rewritten as:

$$[\mathrm{R}] + K_1[\mathrm{R}][\mathrm{L}] + 2K_1K_2[\mathrm{R}]^2[\mathrm{L}] = [\mathrm{R}]_{\mathrm{total}}. \tag{4.17}$$

Solution of this quadratic equation for [R] under the condition of $[\mathrm{R}] \geq 0$ gives:

$$\begin{aligned}[\mathrm{R}] &= \frac{-(K_1[\mathrm{L}] + 1) + \sqrt{(K_1[\mathrm{L}] + 1)^2 + 8K_1K_2[\mathrm{R}]_{\mathrm{total}}[\mathrm{L}]}}{4K_1K_2[\mathrm{L}]} \\ &= 2\frac{-(K_1[\mathrm{L}] + 1) + \sqrt{(K_1[\mathrm{L}] + 1)^2 + bK_1[\mathrm{L}]}}{bK_1[\mathrm{L}]}[\mathrm{R}]_{\mathrm{total}},\end{aligned} \tag{4.18}$$

where $b = 8K_2[\mathrm{R}]_{\mathrm{total}}$. Because subtraction operation in the numerator causes loss of significance in numerical analysis, the formula was further transformed to avoid subtraction:

$$\begin{aligned}[\mathrm{R}] &= 2\frac{-(K_1[\mathrm{L}]+1)+\sqrt{(K_1[\mathrm{L}]+1)^2+bK_1[\mathrm{L}]}}{bK_1[\mathrm{L}]}[\mathrm{R}]_{\mathrm{total}} \\ &\quad\cdot\frac{K_1[\mathrm{L}]+1+\sqrt{(K_1[\mathrm{L}]+1)^2+bK_1[\mathrm{L}]}}{K_1[\mathrm{L}]+1+\sqrt{(K_1[\mathrm{L}]+1)^2+bK_1[\mathrm{L}]}} \\ &= \frac{2[\mathrm{R}]_{\mathrm{total}}}{K_1[\mathrm{L}]+1+\sqrt{(K_1[\mathrm{L}]+1)^2+bK_1[\mathrm{L}]}}. \end{aligned} \tag{4.19}$$

From Equation 4.15,

$$[\mathrm{RL}] = \frac{2K_1[\mathrm{L}][\mathrm{R}]_{\mathrm{total}}}{K_1[\mathrm{L}]+1+\sqrt{(K_1[\mathrm{L}]+1)^2+bK_1[\mathrm{L}]}}, \tag{4.20}$$

$$[\mathrm{R_2L}] = \frac{bK_1[\mathrm{L}][\mathrm{R}]_{\mathrm{total}}}{2\left[K_1[\mathrm{L}]+1+\sqrt{(K_1[\mathrm{L}]+1)^2+bK_1[\mathrm{L}]}\right]^2}. \tag{4.21}$$

If the phosphorylation level (y) of the receptor is proportional to $[\mathrm{R_2L}]$,

$$\begin{aligned} y &= a[\mathrm{R_2L}] \\ &= \frac{bcK_1[\mathrm{L}]}{2\left[K_1[\mathrm{L}]+1+\sqrt{(K_1[\mathrm{L}]+1)^2+bK_1[\mathrm{L}]}\right]^2}, \end{aligned} \tag{4.22}$$

where a is a constant and $c = a[\mathrm{R}]_{\mathrm{total}}$.

Numerical analysis

Numerical analysis was performed with the R statistical environment[98]. Least-squares fitting was performed with the nlm function in the standard library of R.

Quantification of phosphorylated Met on cells

This experiment was performed by Prof. Matsumoto's group (Kanazawa University) as described in literatures[75,95].

EHMES-1 cells were treated with various concentrations of dimeric peptide for 10 min and fixed with paraformaldehyde. After washing and blocking, the cells were

incubated with anti-phospho-MET (Tyr1234/1235) (D26) XP rabbit mAb followed by horseradish peroxidase-conjugated goat anti-rabbit antibody. Tyrosine phosphorylated Met was detected with ImmunoStar LD reagent (Wako) and measured using ARVO MX (Perkin Elmer).

Chapter 5

General conclusion

In this thesis, I have reported a new methodology for discovery of peptide-based covalent protein modifiers by means of *in vitro* selection. In addition to a model target protein used for its development, the selection scheme was also applied to another target protein to demonstrate the generality. I have also reported mathematical modeling of receptor activation induced by bivalent ligands derived from peptides discovered by *in vitro* selection.

In Chapter 2, a series of non-natural amino acids were designed, and they were introduced into an *in vitro* translation system by genetic code reprogramming technique to prepare libraries of reactive macrocyclic peptides. Then I developed an *in vitro* selection scheme of covalent protein modifiers and succeeded in identifying covalent modifiers of a model target protein, Avi-TIM. I have also demonstrated that the modification efficiency of the discovered covalent modifier could be easily modulated by rationally changing the reactive group. I also applied the newly developed selection scheme to EGFR and succeeded in identifying a selective covalent modifier as described in Chapter 3. Because the methodology does not require known ligand scaffold or structural information, it should be readily applicable to various target proteins which cannot be targeted by conventional strategies. The facile and general methodology developed in this thesis would provide useful molecular tools for drug discovery and chemical biology, such as targeted covalent inhibitors for complete and long-lasting inhibition of a drug target, and labeling reagents for detection of a protein of interest.

In Chapter 4, I have demonstrated that a simple mathematical model based on chemical equilibrium could explain the bell-shaped profile observed in phosphorylation level of Met induced by dimeric peptides. The model fitted well with the experimental data, and it predicted that most of the receptor existed as inactive 1:1 complex with excess

amount of the ligand. The mathematical model should be applicable to other receptor-ligand pairs and would facilitate mechanistic understanding of the receptor dimerization induced by homodimeric ligands.

References

1. Walgren, J. L., Mitchell, M. D. & Thompson, D. C. Role of Metabolism in Drug-Induced Idiosyncratic Hepatotoxicity. *Crit. Rev. Toxicol.* **35**, 325–361 (2005).
2. Erve, J. C. L. Chemical toxicology: reactive intermediates and their role in pharmacology and toxicology. *Expert Opin. Drug Metab. Toxicol.* **2**, 923–946 (2006).
3. Lipinski, C. & Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **432**, 855–861 (2004).
4. Potashman, M. H. & Duggan, M. E. Covalent Modifiers: An Orthogonal Approach to Drug Design. *J. Med. Chem.* **52**, 1231–1246 (2009).
5. Vane, J. R. & Botting, R. M. The mechanism of action of aspirin. *Thromb. Res.* **110**, 255–258 (2003).
6. Waxman, D. J. & Strominger, J. L. Penicillin-Binding Proteins and the Mechanism of Action of Beta-Lactam Antibiotics. *Annu. Rev. Biochem.* **52**, 825–869 (1983).
7. Smith, A. J. T., Zhang, X., Leach, A. G. & Houk, K. N. Beyond picomolar affinities: Quantitative aspects of noncovalent and covalent binding of drugs to proteins. *J. Med. Chem.* **52**, 225–233 (2009).
8. Yun, C.-H. *et al.* The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. USA* **105**, 2070–5 (2008).
9. Sachs, G., Shin, J. M. & Howden, C. W. Review article: the clinical pharmacology of proton pump inhibitors. *Aliment. Pharmacol. Ther.* **23**, 2–8 (2006).
10. Pan, Z. *et al.* Discovery of Selective Irreversible Inhibitors for Bruton's Tyrosine Kinase. *ChemMedChem* **2**, 58–61 (2007).

11. Kwarcinski, F. E., Fox, C. C., Steffey, M. E. & Soellner, M. B. Irreversible Inhibitors of c-Src Kinase That Target a Nonconserved Cysteine. *ACS Chem. Biol.* **7**, 1910–1917 (2012).
12. Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **10**, 307–317 (2011).
13. Baillie, T. A. Targeted covalent inhibitors for drug design. *Angew. Chemie Int. Ed.* **55**, 13408–13421 (2016).
14. Bauer, R. A. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. *Drug Discov. Today* **20**, 1061–1073 (2015).
15. Day, R. N. & Davidson, M. W. The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.* **38**, 2887–2921 (2009).
16. Keppler, A. *et al.* A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89 (2003).
17. Gautier, A. *et al.* An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chem. Biol.* **15**, 128–136 (2008).
18. Los, G. V. *et al.* HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).
19. Griffin, B. A., Adams, S. R. & Tsien, R. Y. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **281**, 269–272 (1998).
20. Furman, J. L. *et al.* A Genetically Encoded aza-Michael Acceptor for Covalent Cross-Linking of Protein–Receptor Complexes. *J. Am. Chem. Soc.* **136**, 8411–8417 (2014).
21. Tsukiji, S., Miyagawa, M., Takaoka, Y., Tamura, T. & Hamachi, I. Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat. Chem. Biol.* **5**, 341–343 (2009).
22. Takaoka, Y. *et al.* Quantitative comparison of protein dynamics in live cells and in vitro by in-cell 19F-NMR. *Chem. Commun.* **49**, 2801–2803 (2013).
23. Huhn, A. J., Guerra, R. M., Harvey, E. P., Bird, G. H. & Walensky, L. D. Selective Covalent Targeting of Anti-Apoptotic BFL-1 by Cysteine-Reactive Stapled Peptide Inhibitors. *Cell Chem. Biol.* **23**, 1123–1134 (2016).

24. Hoppmann, C. & Wang, L. Proximity-enabled bioreactivity to generate covalent peptide inhibitors of p53–Mdm4. *Chem. Commun.* **52**, 5140–5143 (2016).
25. Chen, X. H. *et al.* Genetically encoding an electrophilic amino acid for protein stapling and covalent binding to native receptors. *ACS Chem. Biol.* **9**, 1956–1961 (2014).
26. Hughes, C. C. *et al.* Marinopyrrole A Target Elucidation by Acyl Dye Transfer. *J. Am. Chem. Soc.* **131**, 12094–12096 (2009).
27. Yamaguchi, T. *et al.* Turn-ON fluorescent affinity labeling using a small bifunctional O-nitrobenzoxadiazole unit. *Chem. Sci.* **5**, 1021–1029 (2014).
28. Erlanson, D. a *et al.* Site-directed ligand discovery. *Proc. Natl. Acad. Sci. USA* **97**, 9367–9372 (2000).
29. Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
30. Kathman, S. G., Xu, Z. & Statsyuk, A. V. A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Cysteine Proteases. *J. Med. Chem.* **57**, 4969–4974 (2014).
31. Kathman, S. G. & Statsyuk, A. V. Covalent tethering of fragments for covalent probe discovery. *Med. Chem. Commun.* **7**, 576–585 (2016).
32. Zambaldo, C., Daguer, J. P., Saarbach, J., Barluenga, S. & Winssinger, N. Screening for covalent inhibitors using DNA-display of small molecule libraries functionalized with cysteine reactive moieties. *Med. Chem. Commun.* **7**, 1340–1351 (2016).
33. Leenheer, D., ten Dijke, P. & Hipolito, C. J. A current perspective on applications of macrocyclic-peptide-based high-affinity ligands. *Biopolymers* **106**, 889–900 (2016).
34. Zorzi, A., Deyle, K. & Heinis, C. Cyclic peptide therapeutics: past, present and future. *Curr. Opin. Chem. Biol.* **38**, 24–29 (2017).
35. Russ, A. P. & Lampel, S. The druggable genome: an update. *Drug Discov. Today* **10**, 1607–1610 (2005).
36. Hopkins, A. L. & Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727–730 (2002).

37. Davis, M. I. *et al.* Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **29**, 1046–1051 (2011).
38. Craik, D. J., Fairlie, D. P., Liras, S. & Price, D. The Future of Peptide-based Drugs. *Chem. Biol. Drug Des.* **81**, 136–147 (2013).
39. Khan, A. R. *et al.* Lowering the Entropic Barrier for Binding Conformationally Flexible Inhibitors to Enzymes. *Biochemistry* **37**, 16839–16845 (1998).
40. Mallinson, J. & Collins, I. Macrocycles in new drug discovery. *Future Med. Chem.* **4**, 1409–1438 (2012).
41. Rezai, T., Yu, B., Millhauser, G. L., Jacobson, M. P. & Lokey, R. S. Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J. Am. Chem. Soc.* **128**, 2510–2511 (2006).
42. Nielsen, D. S. *et al.* Orally Absorbed Cyclic Peptides. *Chem. Rev.* **117**, 8094–8128 (2017).
43. Passioura, T., Katoh, T., Goto, Y. & Suga, H. Selection-Based Discovery of Druglike Macrocyclic Peptides. *Annu. Rev. Biochem.* **83**, 727–752 (2014).
44. Schmeing, T. M. & Ramakrishnan, V. What recent ribosome structures have revealed about the mechanism of translation. *Nature* **461**, 1234–1242 (2009).
45. Forster, A. C. *et al.* Programming peptidomimetic syntheses by translating genetic codes designed de novo. *Proc. Natl. Acad. Sci. USA* **100**, 6353–6357 (2003).
46. Josephson, K., Hartman, M. C. T. & Szostak, J. W. Ribosomal Synthesis of Unnatural Peptides. *J. Am. Chem. Soc.* **127**, 11727–11735 (2005).
47. Murakami, H., Ohta, A., Ashigai, H. & Suga, H. A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* **3**, 357–359 (2006).
48. Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751–755 (2001).
49. Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **6**, 779–790 (2011).
50. Goto, Y., Murakami, H. & Suga, H. Initiating translation with D-amino acids. *RNA* **14**, 1390–1398 (2008).

51. Fujino, T., Goto, Y., Suga, H. & Murakami, H. Reevaluation of the d-amino acid compatibility with the elongation event in translation. *J. Am. Chem. Soc.* **135**, 1830–1837 (2013).
52. Katoh, T., Tajima, K. & Suga, H. Consecutive Elongation of D-Amino Acids in Translation. *Cell Chem. Biol.* **24**, 46–54 (2017).
53. Fujino, T., Goto, Y., Suga, H. & Murakami, H. Ribosomal Synthesis of Peptides with Multiple β-Amino Acids. *J. Am. Chem. Soc.* **138**, 1962–1969 (2016).
54. Kawakami, T., Murakami, H. & Suga, H. Ribosomal Synthesis of Polypeptoids and Peptoid−Peptide Hybrids. *J. Am. Chem. Soc.* **130**, 16861–16863 (2008).
55. Kawakami, T., Murakami, H. & Suga, H. Messenger RNA-Programmed Incorporation of Multiple N-Methyl-Amino Acids into Linear and Cyclic Peptides. *Chem. Biol.* **15**, 32–42 (2008).
56. Ohta, A., Murakami, H., Higashimura, E. & Suga, H. Synthesis of Polyester by Means of Genetic Code Reprogramming. *Chem. Biol.* **14**, 1315–1322 (2007).
57. Ohta, A., Murakami, H. & Suga, H. Polymerization of α-Hydroxy Acids by Ribosomes. *ChemBioChem* **9**, 2773–2778 (2008).
58. Goto, Y. *et al.* Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **3**, 120–129 (2008).
59. Iwasaki, K., Goto, Y., Katoh, T. & Suga, H. Selective thioether macrocyclization of peptides having the N-terminal 2-chloroacetyl group and competing two or three cysteine residues in translation. *Org. Biomol. Chem.* **10**, 5783 (2012).
60. Sako, Y., Goto, Y., Murakami, H. & Suga, H. Ribosomal synthesis of peptidase-resistant peptides closed by a nonreducible inter-side-chain bond. *ACS Chem. Biol.* **3**, 241–249 (2008).
61. Yamagishi, Y., Ashigai, H., Goto, Y., Murakami, H. & Suga, H. Ribosomal Synthesis of Cyclic Peptides with a Fluorogenic Oxidative Coupling Reaction. *ChemBioChem* **10**, 1469–1472 (2009).
62. Sako, Y., Morimoto, J., Murakami, H. & Suga, H. Ribosomal Synthesis of Bicyclic Peptides via Two Orthogonal Inter-Side-Chain Reactions. *J. Am. Chem. Soc.* **130**, 7232–7234 (2008).

63. Goto, Y., Iwasaki, K., Torikai, K., Murakami, H. & Suga, H. Ribosomal synthesis of dehydrobutyrine- and methyllanthionine-containing peptides. *Chem. Commun.* **0**, 3419 (2009).
64. Nemoto, N., Miyamoto-Sato, E., Husimi, Y. & Yanagawa, H. In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **414**, 405–408 (1997).
65. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* **94**, 12297–12302 (1997).
66. Yamagishi, Y. *et al.* Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem. Biol.* **18**, 1562–1570 (2011).
67. Morimoto, J., Hayashi, Y. & Suga, H. Discovery of Macrocyclic Peptides Armed with a Mechanism-Based Warhead: Isoform-Selective Inhibition of Human Deacetylase SIRT2. *Angew. Chemie Int. Ed.* **51**, 3423–3427 (2012).
68. Hayashi, Y., Morimoto, J. & Suga, H. In Vitro Selection of Anti-Akt2 Thioether-Macrocyclic Peptides Leading to Isoform-Selective Inhibitors. *ACS Chem. Biol.* **7**, 607–613 (2012).
69. Kawamura, A. *et al.* Highly selective inhibition of histone demethylases by de novo macrocyclic peptides. *Nat. Commun.* **8**, 14773 (2017).
70. Yu, H. *et al.* Macrocycle peptides delineate locked-open inhibition mechanism for microorganism phosphoglycerate mutases. *Nat. Commun.* **8**, 14932 (2017).
71. Jongkees, S. A. K. *et al.* Rapid Discovery of Potent and Selective Glycosidase-Inhibiting De Novo Peptides. *Cell Chem. Biol.* **24**, 381–390 (2017).
72. Tanaka, Y. *et al.* Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. *Nature* **496**, 247–251 (2013).
73. Kodan, A. *et al.* Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. *Proc. Natl. Acad. Sci. USA* **111**, 4049–4054 (2014).
74. Matsunaga, Y., Bashiruddin, N. K., Kitago, Y., Takagi, J. & Suga, H. Allosteric Inhibition of a Semaphorin 4D Receptor Plexin B1 by a High-Affinity Macrocyclic Peptide. *Cell Chem. Biol.* **23**, 1341–1350 (2016).

75. Ito, K. *et al.* Artificial human Met agonists based on macrocycle scaffolds. *Nat. Commun.* **6**, 6373 (2015).
76. Lonsdale, R. & Ward, R. A. Structure-based design of targeted covalent inhibitors. *Chem. Soc. Rev.* **47**, 3816–3830 (2018).
77. Chen, G. *et al.* Reactivity of functional groups on the protein surface: Development of epoxide probes for protein labeling. *J. Am. Chem. Soc.* **125**, 8130–8133 (2003).
78. Lanning, B. R. *et al.* A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat. Chem. Biol.* **10**, 760–767 (2014).
79. Niessen, S. *et al.* Proteome-wide Map of Targets of T790M-EGFR-Directed Covalent Inhibitors. *Cell Chem. Biol.* **24**, 1388–1400.e7 (2017).
80. Xuan, W., Shao, S. & Schultz, P. G. Protein Crosslinking by Genetically Encoded Noncanonical Amino Acids with Reactive Aryl Carbamate Side Chains. *Angew. Chemie Int. Ed.* **56**, 5096–5100 (2017).
81. González, M., Argaraña, C. E. & Fidelio, G. D. Extremely high thermal stability of streptavidin and avidin upon biotin binding. *Biomol. Eng.* **16**, 67–72 (1999).
82. Plumb, R. S. *et al.* UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun. Mass Spectrom.* **20**, 1989–1994 (2006).
83. Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* **2**, 127–137 (2001).
84. Harris, R. C., Chung, E. & Coffey, R. J. EGF receptor ligands. *Exp. Cell Res.* **284**, 2–13 (2003).
85. Ferguson, K. M. Structure-Based View of Epidermal Growth Factor Receptor Regulation. *Annu. Rev. Biophys.* **37**, 353–373 (2008).
86. Garrett, T. P. J. *et al.* Crystal Structure of a Truncated Epidermal Growth Factor Receptor Extracellular Domain Bound to Transforming Growth Factor α. *Cell* **110**, 763–773 (2002).
87. Ogiso, H. *et al.* Crystal Structure of the Complex of Human Epidermal Growth Factor and Receptor Extracellular Domains. *Cell* **110**, 775–787 (2002).
88. Ferguson, K. M. *et al.* EGF Activates Its Receptor by Removing Interactions that Autoinhibit Ectodomain Dimerization. *Mol. Cell* **11**, 507–517 (2003).

89. Normanno, N. *et al.* Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* **366**, 2–16 (2006).
90. Scaltriti, M. & Baselga, J. The Epidermal Growth Factor Receptor Pathway: A Model for Targeted Therapy. *Clin. Cancer Res.* **12**, 5268–5272 (2006).
91. Zhen, Y., Caprioli, R. M. & Staros, J. V. Characterization of Glycosylation Sites of the Epidermal Growth Factor Receptor. *Biochemistry* **42**, 5478–5492 (2003).
92. Trusolino, L., Bertotti, A. & Comoglio, P. M. MET signalling: principles and functions in development, organ regeneration and cancer. *Nat. Rev. Mol. Cell Biol.* **11**, 834–848 (2010).
93. Funakoshi, H. & Nakamura, T. Hepatocyte growth factor: from diagnosis to clinical applications. *Clin. Chim. Acta* **327**, 1–23 (2003).
94. Sakai, K., Aoki, S. & Matsumoto, K. Hepatocyte growth factor and Met in drug discovery. *J. Biochem.* **157**, 271–284 (2015).
95. Miao, W. *et al.* Cellular signaling and gene expression profiles evoked by a bivalent macrocyclic peptide that serves as an artificial MET receptor agonist. *Sci. Rep.* **8**, 16492 (2018).
96. Whitty, A. & Borysenko, C. W. Small molecule cytokine mimetics. *Chem. Biol.* **6**, R107–R118 (1999).
97. Perelson, A. S. & DeLisi, C. Receptor clustering on a cell surface. I. theory of receptor cross-linking by ligands bearing two chemically identical functional groups. *Math. Biosci.* **48**, 71–110 (1980).
98. R Core Team. R: A Language and Environment for Statistical Computing. (2018). Available at: https://www.r-project.org/.

List of accomplishments

Publications

1. Ito, K., Sakai, K., Suzuki, Y., Ozawa, N., Hatta, T., Natsume, T., Matsumoto, K., Suga, H. Artificial human Met agonists based on macrocycle scaffolds. *Nat. Commun.* **6**, 6373 (2015).
2. 広瀬久昭, 小澤直也 & 菅裕明. 特殊環状ペプチドを中核とした革新的次世代バイオ医薬品の開発. *Med. Sci. Dig.* **42**, 455–457 (2016).
3. Ozawa, N., Goto, Y. & Suga, H. *In vitro* selection of macrocyclic peptides that bind to epidermal growth factor receptor. *Peptide Science 2017*, 165–166 (2018).
4. Miao, W., Sakai, K., Ozawa, N., Nishiuchi, T., Suzuki, Y., Ito, K., Morioka, T., Umitsu, M., Takagi, J., Suga, H. & Matsumoto, K. Cellular signaling and gene expression profiles evoked by a bivalent macrocyclic peptide that serves as an artificial MET receptor agonist. *Sci. Rep.* **8**, 16492 (2018).

Poster presentations

1. Ozawa, N., Katoh, T. & Suga, H. Development of an *in vitro* selection system for D-peptides that bind to a microRNA precursor. 24th ZESTY Network seminar, Tokyo, Feb. 2017.
2. 小澤直也, 後藤佑樹 & 菅裕明. 共有結合を介して標的タンパク質と結合する特殊ペプチドの探索. 第 5 回バイオ関連化学シンポジウム若手フォーラム, Tokyo, Sep. 2017.
3. Ozawa, N., Goto, Y. & Suga, H. *In vitro* selection of macrocyclic peptides binding covalently to a target protein. 第 54 回ペプチド討論会, Osaka, Nov. 2017.
4. Ozawa, N., Goto, Y. & Suga, H. Facile Development of Peptide-based Covalent Protein Modifiers by In Vitro Selection. 10th International Peptide Symposium, Kyoto, Dec. 2018.

Acknowledgements

This thesis was completed under the supervision of Prof. Hiroaki Suga. I would like to express my sincere gratitude to Prof. Suga for his kind guidance, valuable suggestions and encouragement throughout this work. I would like to thank Assoc. Prof. Yuki Goto for valuable discussions and teaching me various concepts and techniques for organic chemistry and biochemistry. I also thank Assoc. Prof. Takayuki Kato, Assoc. Prof. Naokazu Kano and Asst. Prof. Toby Passioura for their helpful advice and discussions. I thank all members in Suga Lab.

I thank Prof. John P. Richard in the University at Buffalo for a gift of the target protein used in Chapter 2. I also appreciate the work done by Prof. Kunio Matsumoto's group in Kanazawa University for biological assays in Chapter 4.

I thank Materials Education program for the future leaders in Research, Insudtry, and Technology (MERIT) for financial support. I also thank Japan Society for the Promotion of Science (JSPS) for a grant and financial support.

Finally, I would like to thank my family for invaluable assistance and cheering words.