Combinatorially Screened Peptide as Targeted Covalent Binder: Alteration of Bait-Conjugated Peptide to Reactive Modifier

Shuta Uematsu, Yudai Tabuchi, Yuji Ito, Masumi Taki

Bioconjugate Chemistry, 2018, Volume 29, Issue 6, Pages 1866-1871

DOI: 10.1021/acs.bioconjchem.8b00301
Combinatorially screened peptide as targeted covalent binder: alteration of bait-conjugated peptide to reactive modifier

Shuta Uematsu¹, Yudai Tabuchi¹, Yuji Ito², and Masumi Taki*¹

1. Department of Engineering Science, Bioscience and Technology Program, The Graduate School of Informatics and Engineering, The University of Electro-Communications (UEC), 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan.
2. Department of Chemistry and Bioscience, Graduate School of Science and Engineering, Kagoshima University, 1-21-35 Korimoto, Kagoshima, Kagoshima 890-0065, Japan.
E-mail: taki@pc.uec.ac.jp; Tel: +81-42-443-5980

ABSTRACT: A peptide-type covalent binder for a target protein was obtained by combinatorial screening of fluoroprobe-conjugated peptide libraries on bacteriophage T7. The solvatochromic fluoroprobe works as a bait during the affinity selection process of phage display. To obtain the targeted covalent binder, the bait in the selected consensus peptide was altered into a reactive warhead possessing a sulfonyl fluoride. The reaction efficiency and site / position specificity of the covalent conjugation between the binder and the target protein were evaluated by liquid chromatography-tandem mass spectrometry (LC-MS/MS), and rationalized by a protein-ligand docking simulation.

INTRODUCTION

In contrast to conventional medicines, targeted covalent drugs can form permanent bonds to target proteins and eternally deactivate them.¹⁻⁵ This prolonged duration of inhibition⁶ would reduce dose frequency of drugs and improve quality-of-life of patients.⁷⁻⁸ The covalent drugs should be required to possess high target selectivity and less off-target reactions,⁷,⁹ to reduce irreversible side effects (i.e., toxicity).⁸,¹⁰ To retain maximum target selectivity with minimal off-target reactivity,¹⁰ numerous efforts for rational designing of novel covalent drugs have been paid by way of, for example, computational methods,³,¹¹ fragment-based drug discovery,¹² and covalent tethering.¹³

In many cases, optimizing and chemical synthesis of such ideal covalent drugs require additional time and steps. Moreover, if three-dimensional structures of target proteins are not well-known, the structure-based drug development will be hindered. Currently, the development should be performed by limited numbers of sophisticated
methodologies (e.g., chemoproteomics-based one\textsuperscript{14, 15}), and alternative general principles for obtaining target covalent binders are requested.

Contrary to the rational designing, expanded combinatorial screening\textsuperscript{16-19} of such targeted covalent binders would be an alternative way to solve these problems. Indeed, screening of the covalent binders from DNA-encoded chemical libraries has been successfully reported\textsuperscript{20}. For the screenings, displayed-peptide libraries would be also attractive mainly because the library diversity is large (up to $\sim 10^{12}$ and $\sim 10^{14}$ for phage and mRNA display, respectively\textsuperscript{18}) and preparation of the library is fairly easy\textsuperscript{21}, compared with the chemical ones. However, to the best of our knowledge, there are no reports to obtain targeted covalent binders from the displayed randomized peptide, presumably because 1) the introduction method of warhead (i.e., reactive group) into the displayed peptides with vast diversity is limited, and 2) exclusive enrichment of the targeted covalent binders from the vast library seems difficult: the warhead in target-unrelated peptides could unfavorably conjugate to the target protein during biopanning process (Fig. S1).

To avoid the latter problem of unnecessary side reactions, here we take a detoured selection strategy to obtain a peptide-type covalent binder possessing target specificity (Fig. 1). First, we introduce several solvatochromic bait fragments, instead of the warhead, to designated cysteine on T7 phage-displayed library peptides\textsuperscript{22} via the gp10 based-thioetherification (10BASE\textsubscript{4}-T)\textsuperscript{23, 24}. Second, we obtain targeted non-covalent binders from each bait-conjugated peptide library. Third, when a consensus peptide sequence around the designated cysteine appears, no matter what kind of bait fragment is used, we alter the bait in the consensus peptide into a different-structured warhead, to obtain the targeted covalent binder.
**Fig. 1.** Targeted covalent binders selected by combinatorial screening of bait fragment-conjugated peptide library on bacteriophage T7, followed by an alteration to a reactive warhead possessing sulfonyl fluoride. The chemical structures of different solvatochromic bait fragments (i.e., Prodan, 4-DMN, and DBD), as well as that of an altered warhead, are shown in dashed inset. (A) Specific introductions of different solvatochromic bait fragments into a designated cysteine on displaying library peptides on a capsid protein (gp10) of bacteriophage T7 were performed. This gp10-based thioetherification (10BASEd-T) was carried out without side reactions or loss of phage infectivity. (B) From each bait-conjugated peptide library, target (i.e., GST) binders were selected by biopanning. Then, the peptide sequences were analyzed via a next generation sequencer (NGS), and consensus sequences for every bait-conjugated peptide around the designated cysteine was determined. For each bait-conjugated peptide, solvatochromic fluorescence-change upon GST-binding was also confirmed. (C) To obtain the targeted covalent binder, the bait fragments in the consensus peptide was altered into the warhead. The covalent binder was mixed with the target protein, and the modified position on the target was analyzed by LC-MS/MS.

**RESULTS AND DISCUSSION**

As a model target protein, we chose glutathione S-transferase (GST) because rationally-designed covalent GST binders in which a warhead is conjugated with artificial reporter tags\textsuperscript{25} or a natural ligand (i.e., glutathione)\textsuperscript{26} have been reported; their structure information as well as conjugation efficiency upon GST binding could be easily compared with our combinatorially-screened binder. For the bait fragments with different shapes, we have chosen several small or middle-sized solvatochromic fluorophores with neutral charge,\textsuperscript{22, 27} so that we would sense when the bait could be buried deeply into a pocket of the target protein through hydrophobic interaction (Fig. 1). These fragments were independently reacted with a designated cysteine on a T7-displayed randomized peptide library, and three rounds of biopanning were performed against biotinylated-GST.\textsuperscript{22} After the selection process, amino-acid sequences of the polyclonal binders possessing each fragment were directly analyzed by a next generation sequencer (NGS). As shown in Table 1, common consensus sequences for all of these three bait fragments were obtained. In the common consensus peptides, structural flexibility of the fragment which is dangling on the designated cysteine seemed to be widely allowed. Among them, we chose ZC*DGZ sequence (highlighted in blue, Table 1) for further experiments, because it does not contain any wobble amino acids (i.e., X).
Table 1. GST-specific binders selected after 3 rounds of biopanning. C*, Z, and X represent bait-fragment conjugated cysteine, hydrophobic amino acids, and any amino acids, respectively. The consensus sequences from each bait fragment were summarized by frequency of the top 100 occurrences obtained from NGS data analysis. The bold letter stands for common consensus sequence for all of the different-bait-conjugated peptides. The detailed sequences with the abundance ratio for each monoclone of ZC*DGZ were presented in Table S1.

(A) C*: DBD-conjugated cysteine

<table>
<thead>
<tr>
<th>Consensus sequence</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZC*DGZ</td>
<td>20/100</td>
</tr>
<tr>
<td>ZZC*DGZ</td>
<td>12/100</td>
</tr>
<tr>
<td>C*XDGZ</td>
<td>10/100</td>
</tr>
</tbody>
</table>

(B) C*: 4-DMN-conjugated cysteine

<table>
<thead>
<tr>
<th>Consensus sequence</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZC*DGZ</td>
<td>9/100</td>
</tr>
<tr>
<td>ZXYC*XDGZ</td>
<td>2/100</td>
</tr>
</tbody>
</table>

(C) C*: Prodan-conjugated cysteine

<table>
<thead>
<tr>
<th>Consensus sequence</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>C*XXXDDGZ</td>
<td>10/100</td>
</tr>
<tr>
<td>ZC*DDGZ</td>
<td>8/100</td>
</tr>
<tr>
<td>ZC*DGZ</td>
<td>7/100</td>
</tr>
<tr>
<td>C*XDDGZ</td>
<td>6/100</td>
</tr>
<tr>
<td>C*XXXXDGZ</td>
<td>6/100</td>
</tr>
<tr>
<td>C*XDGZ</td>
<td>5/100</td>
</tr>
<tr>
<td>C*XXXDGZ</td>
<td>4/100</td>
</tr>
</tbody>
</table>

Next, a representative common peptide whose sequence is LNYCDGW (the common consensus sequence is underlined) was synthesized, because it is one of the most abundant sequences obtained from six rounds of biopanning against GST using 4-DMN-conjugated randomized peptide library. Then, its sulfhydryl group was conjugated with all of the bait fragments independently. When each bait-conjugated peptide was mixed
with GST, a remarkable solvatochromic fluorescence change was observed (Fig. S2). This means that the microenvironment around the bait changed after the addition of GST, and most probably, all of the bait fragments were located inside of the hydrophobic pocket of GST, no matter what kind of the chemical structures were used.

Encouraged by the solvatochromic fluorescence, next we altered the bait structures to a reactive warhead, and target-protein specific covalent binding of the altered peptide was evaluated. We chemically synthesized a targeted covalent binder peptide whose sequence is Fam–GGLNYC*DGW (Fig. 2A: Fam, GG, and C* represent carboxyfluorescein, glycylglycine spacer, and warhead-conjugated cysteine, respectively), and its appropriate covalent conjugation with GST was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; Fig. 2B). As the warhead, we chose sulfonyl fluoride (SO₂F),²⁵, ²⁶, ²⁸ because the reactive group possesses an exquisite balance of aqueous stability and moderate reactivity toward nucleophilic amino acids (i.e., serine, threonine, lysine, tyrosine, cysteine, and histidine residues in the target protein).²⁹, ³⁰ When 4-(2-bromoacetyl)benzene-1-sulfonyl fluoride was reacted with the sulfhydryl group of the precursor peptide, complex reaction products were formed and we could not separate the expected covalent binder peptide. In contrast, when ethenesulfonyl fluoride²⁶ was reacted, only a single component of the covalent binder peptide was successfully obtained (Fig. 2A; for details, see Fig. S3). The covalent binder peptide was mixed with GST in the presence of serum proteins, and the reaction mixture was analyzed by SDS-PAGE followed by fluorescence imaging / coomassie brilliant blue (CBB) staining. An intense fluorescent band could be seen at an appropriate molecular weight (ca. 29 kDa) of the peptide-fused GST (Fig. 2B). This means that we could successfully confirm the covalent-binding property, as well as the GST specificity, of the peptide.
Fig. 2. (A) Synthesis of a targeted covalent binder peptide: Fam stands for a fluorophore at the N-terminus. (B) Specific conjugation between the covalent binder and GST, confirmed by 15% SDS-PAGE / fluorescence imaging. Whole proteins were visualized by CBB staining (left panel), and a protein conjugated with the covalent binder was visualized by fluorescence in the same gel (right). Blue arrow represents the band of the target protein (i.e., GST). At this stage, the crosslinking reaction yield was less than few% and the structural information could not be obtained by LC-MS/MS analysis.
To avoid unfavourable steric hindrance upon target binding and non-specific hydrophobic interactions arising from the fluorophore, LNYC*DGW peptide without Fam–GG was once again synthesized. The chemical structure of the covalent binder was tentatively identified by combination of 1H NMR, MS, and MS/MS analyses (Fig. S4), and favorably, the warhead did not react with its intrinsic tyrosine residue. Then, the reaction efficiency and site / position specificity of the covalent conjugation between the Fam-eliminated covalent binder and GST were evaluated as the following. First, the covalent binder and GST were reacted, and SDS-PAGE was performed. Then, the GST band in the gel was excised and digested with trypsin. The resulting peptide fragments were analyzed by LC-MS/MS. An intense absorbance in the chromatogram was newly detected after the conjugation (Fig. 3A), and it was identified as the GST-conjugated covalent binder by MS (Fig. 3B) and tandem mass spectroscopy (Fig. 3C). The crosslinking site was determined to be the glutathione binding pocket; the crosslinked amino acid was a tyrosine which was located at the 111th position from the N-terminus. This position was exactly the same when the rationally-designed binders25, 26 were covalently conjugated (also see Fig. S5). From measurement of the abundance ratio of unreacted peptide fragment possessing the 111th tyrosine, the crosslinking reaction yield of our covalent binder was estimated to be 37%, which was much superior to that of the rationally-designed one (Fig. S6).
Fig. 3. Identification of fragments derived from the covalent-binder-conjugated GST peptide by LC-MS/MS analysis. (A) LC profiles of trypsin-digested peptide fragments derived from modified and unmodified GST. The absorbance above 270 nm was measured. (B) MS and (C) MS/MS spectra of the newly appeared peak (arrowed in Fig. 3A). All the detected fragments were consistent with theoretical m/z values of the represented structure. The peptide fragment of IAYSK was derived from a constituent of the glutathione binding pocket of GST protein. Y* means conjugated tyrosine.
Finally, the conjugation of the covalent binder was rationalized by a protein-ligand docking simulation using MolDesk Basic / myPresto. Thirty separate poses resulted in docking to the glutathione binding pocket of GST with free energy in the range -12.4 to -7.83 kcal/mol. As shown in Fig. 4, the docking model of the lowest energy suggested that the warhead was buried deep inside the hydrophobic region of the glutathione-binding pocket and located very close to the conjugated tyrosine of the 111th position. The geometry of the altered warhead is in good agreement with the solvatochromic fluorescence change of the bait-conjugated peptide upon GST binding (Fig. S2). We speculate this proximity between the warhead and the 111th tyrosine caused the site- and position-specific efficient crosslinking. In contrast, the lowest docking model of the rationally-designed one suggested that the warhead was located outside of the pocket (Fig. S7), which might result in the insufficient conjugation with the 111th tyrosine.

Fig. 4. Molecular docking simulation of the covalent binder (shown as a stick) to GST (PDB ID: 1UA5) using sievgene of myPresto: the best docking model with a lowest binding energy of -12.4 kcal/mol was presented. Fluorine atom in the warhead and conjugated tyrosine in GST were colored in cyan and dark red, respectively. GST was shown as a cartoon with side chains as a line description.
CONCLUSIONS
In conclusion, we have established a precise and efficient conjugation method between a hydrophobic pocket of a target protein and a targeted peptide. The peptide as a covalent binder has been obtained by combinatorial screening of a bait fragment-conjugated peptide library on phage, followed by an alteration of the bait to a reactive warhead. Although the example shown here is not completely representative, the choice of GST enables us to rationalize the target binding retrospectively by using a computational model based on the crystal structure. Current limitation of this methodology seems that the bait-conjugated peptide tend to bind to a promiscuous large binding pocket. Nevertheless, we believe the non-reactive bait-alteration concept will be useful in general for the future, to discover covalent drugs made of peptide as antibody-drug substitutes, especially aimed for cell-surface receptors.

MATERIALS AND METHODS
Biopanning against GST by using bait fragment-conjugated peptide library. Synthesis of each bait-conjugated peptide library on phage and biopanning against GST was performed as described previously.22 In brief, approximately $1.0 \times 10^{11}$ pfu of T7Select10 library (S-G-G-G-X$_5$-C-X$_5$-C-X$_5$; X represents any randomized amino acid) was modified with bromoacetamide (BA)-conjugated bait fragment via the 10BASE$_d$-T. After modification, the T7 phage library was dissolved in selection buffer (PBS supplemented with 0.1% v/v TritonX-100). To remove non-specific binders (e.g., beads and streptavidin binders), the modified T7 phage library was pre-incubated with streptavidin-coupled beads for 16 hours at 4 °C, and then the supernatant was further incubated with the GST-immobilized beads for 15 hours. The latter beads were washed three times for 14 minutes in total with 0.2 mL of the selection buffer. The entire binding and washing processes were performed using an automated machine (Target Angler 8, Tamagawa Seiki, Japan). GST-bound phage was directly infected and amplified with E. coli BLT5403 strain. Increasing stringent conditions such as shortening the binding time (e.g., 400 min at the final round) and increasing the washing frequency / time (up to five times / 70 min in total at the final round), were applied stepwise to each round. After 3 rounds of biopanning, a mixture of T7 phage polyclones was subjected to a next generation sequencer (NGS), to obtain the common consensus sequence.

Synthesis of the bait-conjugated peptides and confirmation of their solvatochromic fluorescence-change upon GST-binding. Each bait-conjugated peptide possessing
LNYCDGW sequence was synthesized by reaction of each BA-conjugated bait fragment independently, according to the reported procedure. Then, each bait-conjugated peptide (0.07 mM) was mixed with equal molar of GST or streptavidin in PBS. It was incubated for 10 minutes at room temperature, and fluorescence spectrum was measured (Fig. S2).

Conjugation between GST and covalent binders:

(A) Preliminary covalent modification at the glutathione-binding pocket using a reported covalent binder. A rationally-designed covalent binder (abbreviated as E'C*G), which is known to irreversibly bind to the glutathione-binding pocket of human GST, was synthesized and identified according to the reported procedure in combination with MS/MS analysis (Fig. S5B). The binder (5.2 mM) was mixed with S. japonicum GST (2.7 mM) in 20 mM phosphate buffer (pH 7.5), and incubated for 21 hours at 37 °C in the dark. It was mixed with 1×sample buffer, denatured at 95 °C, separated by 15% SDS-PAGE followed by trypsinization / LC-MS/MS analysis. The reported covalent binder successfully conjugated to a tyrosine located at deep inside the glutathione binding pocket of S. japonicum GST (Fig. S5C).

(B) Specific conjugation between combinatorially-screened covalent binder and GST confirmed by SDS-PAGE / fluorescence imaging. Fluorescent peptide (Fam–GGLNYCDGW, 0.01 M) was dissolved in 20 mM phosphate buffer (pH 7.4) / 10% acetonitrile, and reacted with ethenesulfonyl fluoride (ESF, 0.1 M) or 4-(2-bromoacetyl)benzenesulfonyl fluoride (BBSF, 5 mM) in the presence of neutralized tris(2-carboxyethyl)phosphine (TCEP, 2 mM). The mixture was reacted for more than 4 hours at room temperature in the dark with vigorous shaking, and the reaction was monitored by HPLC (Fig. S3). Only the ESF-conjugated covalent binder could be purified by reverse-phase HPLC, and it was used for further experiments. The purified covalent binder (0.10 mM) was mixed with GST (0.31 mM) in phosphate buffer (pH 7.5) in the presence of 40% (v/v) human serum (Sigma, H4522), and incubated for 21 hours at 37 °C in the dark. It was mixed with 1×sample buffer, denatured at 95 °C, and separated by 15% SDS-PAGE. The binder-conjugated proteins were detected by fluorescence imaging, and whole proteins were visualized by CBB staining (Fig. 2B).

(C) Detailed identification of the conjugated site / position on GST after the reaction with combinatorially-screened covalent binder. Peptide (LNYCDGW, 0.24 M) was dissolved in 20 mM phosphate buffer (pH 7.4) / 10% acetonitrile, and reacted with ESF (0.26 M) in the presence of neutralized tris(2-carboxyethyl)phosphine (TCEP, 2 mM). The mixture was reacted for more than 4 hours at room temperature in the dark with vigorous shaking, and purified by reverse-phase HPLC (yield 60%), and identified with 1H NMR,
MS, and MS/MS analyses (Fig. S4). Covalent conjugation reaction with GST, and SDS-PAGE followed by trypsinization / LC-MS/MS analysis were performed under the same conditions described above (A), except using the combinatorially-screened covalent binder (i.e., LNYC*DGW) instead of the glutathione derivative (i.e., E’C*G).

■ ASSOCIATED CONTENT
Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: ***.
Additional detailed materials and methods; supplemental figures and tables (PDF)

■ AUTHOR INFORMATION
Corresponding Author
*E-mail: taki@pc.uec.ac.jp.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS
This work was supported by a JSPS KAKENHI grant (#17K05925) to M.T. We appreciate Dr. L. Nelson (MD Anderson), Prof. T. Yamashita (Keio Univ.), and Prof. M. Tanaka (UEC) for critical reading of this manuscript. We would like to acknowledge Ms. M. Yonekura and Mrs. Y. Kato as assistant researchers for performing NGS analysis and data mining in silico. We also thank to Dr. H. Nakamura (Biomodeling Research Inc.) for fruitful discussions about molecular modelling. Finally, we greatly appreciate a (blinded) peer reviewer who read our old manuscript so carefully, and gave us a lot of fruitful questions and comments to let us aware of deep understandings.

■ REFERENCES


Table of Contents (TOC) entry.