Enzymatic "Click" Ligation: Selective Cysteine Modification in Polypeptides Enabled by Promiscuous Glutathione S-Transferase

Citation: Zhang, Chi, Alexander M. Spokoyny, Yekui Zou, Mark D. Simon, and Bradley L. Pentelute. "Enzymatic 'Click' Ligation: Selective Cysteine Modification in Polypeptides Enabled by Promiscuous Glutathione S-Transferase." Angewandte Chemie International Edition 52, no. 52 (November 12, 2013): 14001–14005.

As Published: http://dx.doi.org/10.1002/anie.201306430

Publisher: Wiley Blackwell

Persistent URL: http://hdl.handle.net/1721.1/95629

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Published in final edited form as:
Angew Chem Int Ed Engl. 2013 December 23; 52(52): 14001–14005. doi:10.1002/anie.201306430.

Enzymatic "Click" Ligation: Selective Cysteine Modification in Polypeptides Enabled by Promiscuous Glutathione *S*-Transferase**

Chi Zhang, **Alexander M. Spokoyny**, **Yekui Zou**, **Mark D. Simon**, and **Bradley L. Pentelute***
Department of Chemistry Massachusetts Institute of Technology 77 Massachusetts Avenue, MA 02139, USA

Keywords

Bioconjugation; Click Chemistry; Cysteine Arylation; Enzyme Promiscuity; Peptide Macrocyclization

Post-translational modifications of biomolecules represent nature's versatile chemical toolbox, allowing for precise and on-demand placement of various functional moieties *in vivo*.[1] Synthetic chemists have strived to develop a similar reaction toolkit with chemical orthogonality, selectivity and reactivity comparable to those of natural systems.[2] Recently, "click" chemistry has produced several powerful transformations applicable to a wide range of synthetic settings.[3] While the efficiency and chemical orthogonality of such "click" processes are remarkable, regio-control for these transformations has been limited and remains a major challenge.[4] The difficulty of selectively differentiating between two chemically identical sites within a macromolecule is routinely surmounted in natural systems by recognition elements such as binding pockets, directing groups and allosteric components. Although various enzymes have been engineered to catalyze reactions for protein modification and engineering,[5] they often require a long recognition sequence in polypeptides, and their substrate scope is relatively restricted. In certain cases enzymes are capable of recognizing and carrying out efficient transformations on more than one specific substrate, a phenomenon known as enzyme promiscuity. Promiscuous enzymes have been previously engineered to catalyze several synthetically important organic transformations.[6]

Synthetic methods for modifying biomolecules provide opportunities for studying protein structure-function relationships as well as creating proteins with new properties and applications.[7] Cysteine (Cys) is often chosen for modification because of the unique reactivity of its thiol group [8] as well as the low abundance of Cys residues in the majority of naturally occurring proteins (ca. 1.7%).[9] Various chemical methods for Cys modification have been developed, such as alkylation[10], oxidation[11], and desulfurization[12]. Recently, we developed a perfluoroaryl-cysteine S_NAr chemistry

**The authors would like to thank Prof. Stephen L. Buchwald (S. L. B.) for encouragement and support. This research was generously sponsored by the National Institutes of Health (GM101762 for A.M.S. and GM046059 for S.L.B.), and MIT start-up funds and Damon Runyon Cancer Research Foundation award for B.L.P. The content is solely the responsibility oft he authors and does not necessarily represents the official views oft he NIH. We are indebted to Mr. Rocco Policarpo, Ms. Jingjing J. Ling, Dr. Xiaoli Liao, and Ms. Amy Rabideau for technical assistance and helpful discussion, and Prof. R. John Collier (Harvard) for commenting on the manuscript and contributing some of the laboratory equipment used in these studies. We also thank Prof. Stephen B. H. Kent (U. Chicago) for commenting on this manuscript.

*blp@mit.edu.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

approach to modifying unprotected peptides (Figure 1A).[13] While this method satisfies several requirements of "click" chemistry,[14] the insolubility and low reactivity of the perfluoroaromatic reagents in aqueous media restricts the general application of this process to a narrow range of biomolecules. In our search to render this chemistry suitable for aqueous conditions we turned our attention to glutathione *S*-transferase (GST).[15] Known as a class of promiscuous enzymes, GST catalyzes conjugation reactions between the Cys residue of glutathione (GSH, γ-Glu-Cys-Gly) and various electrophiles, thus allowing the cell to detoxify xenobiotics *in vivo* (Figure 1B).[15a]

Here we report a novel bioconjugation chemistry that combines a previously developed perfluoroaryl-cysteine S_NAr "click" reaction with GST enzyme catalysis (Figure 1C). The promiscuous nature of GST facilitates the bioconjugation with polypeptides other than GSH in aqueous media and provides previously unattainable chemo- and regioselective functionalization of a single cysteine thiol in the presence of other unprotected cysteine residues and reactive functional groups on the same polypeptide chain. Furthermore, we show that this process can be completed in seconds, thereby providing a new and efficient approach to peptide macrocyclization. This chemistry can be carried out over a broad range of temperatures (4-60°C) and is compatible with the addition of organic co-solvents (up to 20%).

We hypothesized that the broad scope of electrophiles accepted by GST isozymes might be sufficient for members of this enzyme family to mediate reactions between perfluoroaryl electrophiles and peptides containing GSH in an aqueous environment. To achieve the broadest electrophile scope, a mixture of GST isozymes was chosen for screening.[16] We first tested the GST-catalyzed conjugation of GSH to model peptides containing L-pentafluorophenylalanine (residue **I**, Figure 2A). Reacting **1** with GSH at 37 °C in aqueous solution at pH 8.0 with 2 mg/mL GST (ca. 5-10 mol% relative to **1**) for two hours generated conjugated product **2**, as confirmed by LCMS analysis (Figure S6), whereas no product was observed without the enzyme. Nucleophilic residues in model peptide **1**, such as Cys and Lys, were unreactive, indicating that the Cys of GSH can be selectively modified with pentafluorophenyl-based electrophiles under GST catalysis.

To examine the substrate scope of this reaction, we first tested whether the mixture of GST isozymes could catalyze the conjugation of L-pentafluorophenylalanine residue to peptides bearing N-terminal GSH (γ-Glu-Cys-Gly). However, a hexapeptide containing N-terminal glutathione sequence (γ-Glu-Cys-Gly-Gly-Leu-Leu) did not show reactivity towards **1** (Figure S7). We then hypothesized that increasing the electrophilicity of the perfluoroaryl moiety might improve the reactivity of the peptide-based substrate sufficiently to allow GST-mediated conjugation with peptides containing N-terminal GSH sequence. Our previous study showed that *para*-thioether substituent on the perfluoroaryl moiety can stabilize the negative charge of the S_NAr reaction intermediate thereby increasing the reaction rate.[13] We evaluated the enzymatic reactivity of peptides containing several *para*-thioether substituted electrophiles derived from cysteine. Importantly, reactions with these peptides showed enhanced reaction rates as compared to peptide containing L-pentafluorophenylalanine (residue **I**, Figure 2A and Figure S7). Specifically, peptide containing cysteine modified with perfluorophenyl residue (Cys-**II**) (Figure 2A and 2C, **3a**; Figure S1 for synthesis) reacted with GSH in the presence of GST at a significantly higher rate as compared to **1,** yielding 93% of GSH-conjugated product in less than four hours. Reactions with peptides containing Cys moiety functionalized with perfluorobiphenyl species (Cys-**III**) (Figure 2A and 2B, **3b**) and perfluorobiphenyl sulfide (Cys-**IV**) (Figure 2A, **3c**) proceeded with quantitative conversions in less than 30 minutes (Figure 2C).

Using more reactive electrophiles, we conducted studies beyond the GSH substrate. We focused on ligations between peptides containing Cys-**III** residue and hexamer peptides with an N-terminal γ-Glu-Cys-Gly. Our study commenced with the synthesis of glycyl-modified GSH-based peptides, where the first amino acid directly after GSH was varied (Figure 3, **5a-e**). All reactions proceeded quantitatively within two hours (Figure 3, entries 1-5), and reaction with **5a** showed high conversion within 30 minutes (Figure 3, entry 1). Decreased reaction rates with **5b** and **5e** as compared to **5a** were observed, suggesting this site may be important for interacting with GST. Changing the second amino acid in the sequence linked to the Gly site to a less bulky residue (Figure S13F) had no effect on the conjugation rate. Nucleophilic residues (His and Lys, Figure 3, entries 3 and 4) are compatible with the reaction. Surprisingly, the reaction with **5c** produced 98% conversion within 30 minutes, suggesting that His residue may favor the reaction.

While the residue C-terminal to GSH can affect the relative rates of the GST-catalyzed reaction, the observed product yield for reactions with Cys-**III** has thus far been independent of the peptide sequence employed, further suggesting that the electrophilicity of the perfluoroaryl substituents can dominate the GST-catalyzed S_NAr reaction with modified GSH (*vide supra*). These results show the promiscuity associated with GST is adequate to catalyze the selective bioconjugation of two unprotected polypeptide fragments and is unique, as it demonstrates the possibility of enhancing a click reaction via enzymatic catalysis.[17]

This reaction is highly selective for arylation of Cys thiols, as evidenced by the competition experiment with a large excess of exogenous thiol (4-mercaptophenylacetic acid - MPAA, 100-fold, Figure S8). No MPAA-arylated product was generated, while the GST-catalyzed product was exclusively produced. In addition, reacting peptide **3b** with another peptide **5k** containing C-terminal hydrazide moiety and N-terminal γ-Glu-Cys-Gly sequence gave solely S-arylated product **6k**, which could be further modified by hydrazone ligation (Figure S8).[18]

The unique chemo- and regioselectivity of the GST-catalyzed arylation reaction could be exploited to label one Cys residue in the presence of another on a fully unprotected peptide or protein. We first labeled the N-terminal GSH Cys with biotin and subsequently modified a separate Cys with a fluorophore maleimide. This provides the first example of orthogonal site-specific labeling of two unprotected Cys residues within the same unprotected peptide or protein. Biotin probe containing Cys-**III** moiety (Figure 4A, **11**, see Figure S3 for synthesis) was conjugated to the Cys position of N-terminal γ-Glu-Cys-Gly fragment in protein **7** (see Figure S4 for synthesis). The enzymatic reaction produced the mono-labeled product **8**, and the other cysteine was labeled with fluorescein-5-maleimide **10**, thereby producing the site-specific dual-labeled protein species **9** (Figure 4A). Using similar protocol, the biotin and fluorescein probes were regiospecifically attached to peptide **14** to produce the dual labeled species **16** (Figure S10). The authentic sample was prepared using conventional protecting-group chemistry to confirm the identity of mono-labeled peptide species **15** (Figure S16). Importantly, regiospecific and orthogonal modification of two chemically identical Cys sites was enabled by GST selective recognition of the N-terminal GSH moiety. This result indicates that the GST-catalyzed arylation could greatly expand the scope of previous cysteine modification methods, which necessitate the use of protecting groups or multiple steps to differentially functionalize two or more cysteine residues.[19]

Cyclic peptides constitute a very important class of medicinally relevant macrocycles.[20] Although various methods have been previously developed,[21] synthesis of macrocyclic peptide fragments remains challenging. We found intramolecular arylation catalyzed by GST of peptide **12** containing γ-Glu-Cys-Gly fragment and Cys-**III** site afforded

quantitative conversion to the cyclized product **13** in less than 30 seconds (Figure 4B). In contrast, the control experiment without GST showed no product formation as observed by LCMS analysis. Furthermore, the cyclization reaction was shown to be dominant even with GSH added as a competing substrate (Figure S11). The perfluorinated moiety used in this example can potentially enhance the cell-permeability of cyclic peptides as shown previously.[13]

This GST mediated transformation is not limited to specific temperatures, solvent conditions and peptides featuring γ-Glu-Cys-Gly sequence. For example, the reaction between **3b** and GSH under GST catalysis produces the desired arylated product **4b** at temperatures ranging from 4 to 60 °C (Figure 4C, right panel), as well as in the presence of up to 20% organic co-solvent (Figure 4C, left panel). Additionally, GSH analogues with the mutation of γ-Glu to a genetically encodable Glu or Asp (Glu-Cys-Gly or Asp-Cys-Gly) undergo facile S-arylation with peptide **3b** under the developed GST-catalyzed conditions (Fig. S17). Finally, our experiments show that N-terminal Glu-Cys-Gly or Asp-Cys-Gly sequences in peptides (Figure S12, **5g** and **5h**) can also be selectively S-arylated with peptide **3b** though with lower efficiency as compared to peptides with N-terminal γ-Glu-Cys-Gly moiety (Figure S13A, **5a**). These results are consistent with previous findings that certain GST isozymes are capable of catalyzing the S_NAr reaction between 1,4-dinitrochlorobenzene and glutathione analogues.[22] Together, our observation suggests that the requirement for an N-terminal γ-glutamic acid residue could be eliminated with engineered GST.

In conclusion, we demonstrated a powerful glutathione *S*-transferase catalyzed S_NAr "click" process for site-specific cysteine modification. This method allowed us to selectively modify cysteine in an N-terminal γ-Glu-Cys-Gly sequence within peptide and protein chains. The unique chemical orthogonality of the discovered arylation enabled by GST provides a route to modifying multiple cysteine sites with different chemical probes or biomolecules while avoiding the use of protecting groups and additional synthetic procedures. By showing how a naturally occurring enzyme catalyst improves and significantly broadens the scope of an established "click" synthetic transformation, this discovery paves a frontier in discovering new enzyme mediated reactivity modes suited for chemoselective modification of biomolecules.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Walsh CT, Garneau-Tsodikova S, Gatto GJ. Angew. Chem. Int. Ed. 2005; 44:7342.
2. a Davis BG. Science. 2004; 303:480. [PubMed: 14739446] b Marchetti L, Levine M. ACS Catalysis. 2011; 1:1090.
3. a Kolb HC, Finn MG, Sharpless KB. Angew. Chem. Int. Ed. 2001; 40:2004.b Jewett JC, Bertozzi CR. Chem. Soc. Rev. 2010; 39:1272–1279.c Moses JE, Moorhouse AD. Chem. Soc. Rev. 2007; 36:1249.d Becer CR, Hoogenboom R, Schubert US. Angew. Chem. Int. Ed. 2009; 48:4900.e Blackman ML, Royzen M, Fox JM. J. Am. Chem. Soc. 2008; 130:13518. [PubMed: 18798613] f Hoyle CE, Bowman CN. Angew. Chem. Int. Ed. 2010; 49:1540.
4. Wong C-H, Zimmerman SC. Chem. Commun. 2013; 49:1679.
5. a Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohana RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J, Darzins A, Klaubert DH, Bulleit RF, Wood KV. ACS Chem. Biol. 2008; 3:373.b Watanabe S, Mizukami S, Hori Y, Kikuchi K. Bioconjugate Chem. 2010; 21:2320.c George N, Pick H, Vogel H, Johnsson N, Johnsson K. J. Am. Chem. Soc. 2004; 126:8896.d Chen I, Howarth M, Lin W, Ting AY. Nat. Methods. 2005; 2:99. [PubMed: 15782206] e Uttamapinant C, White KA, Baruah H, Thompson S,

Fernandez-Subrez M, Puthenveetil S, Ting AY. Proc. Natl. Acad. Sci. U. S. A. 2010; 107:10914. [PubMed: 20534555] f Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K. Nat. Biotechnol. 2003; 21:86.g Witus LS, Francis M. Curr. Protoc. Chem. Biol. 2010; 2:125. [PubMed: 23836553] h Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL. Nat. Chem. Biol. 2007; 3:707. [PubMed: 17891153]

6. a Babtie A, Tokuriki N, Hollfelder F. Curr. Opin. Chem. Biol. 2010; 14:200. [PubMed: 20080434] b Humble MS, Berglund P. Eur. J. Org. Chem. 2011; 2011:3391.c Khersonsky O, Tawfik DS. Annu. Rev. Biochem. 2010; 79:471–505. [PubMed: 20235827] d Reetz MT. Angew. Chem. Int. Ed. 2011; 50:138.e Coelho PS, Brustad EM, Kannan A, Arnold FH. Science. 2013; 339:307.
7. a Hermanson, GT. Bioconjugate Techniques. second ed.. Elsevier Inc.; 2008. b Carrico IS. Chem. Soc. Rev. 2008; 37:1423.c Antos JM, Francis MB. Curr. Opin. Chem. Biol. 2006; 10:253. [PubMed: 16698310] d Hackenberger CPR, Schwarzer D. Angew. Chem. Int. Ed. 2008; 47:10030.e Basle E, Joubert N, Pucheault M. Chem. Biol. 2010; 17:213. [PubMed: 20338513] f Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Science. 1994; 266:776. [PubMed: 7973629] g Sato S, Nakamura H. Angew. Chem. Int. Ed. 2013; 52:8681.h Ling JJ, Policarpo RL, Rabideau AE, Liao X, Pentelute BL. J. Am. Chem. Soc. 2012; 134:10749. [PubMed: 22686546]
8. a Chalker JM, Bernardes GJL, Lin YA, Davis BG. Chem.--Asian J. 2009; 4:630. [PubMed: 19235822] b Crankshaw, MWG.; G. A.. Modification at Cysteine. Wiley; 1996.
9. Fodje MN, Al-Karadaghi S. Protein Eng. 2002; 15:353.
10. Weerapana E, Simon GM, Cravatt BF. Nat. Chem. Biol. 2008; 4:405.
11. van Kasteren SI, Kramer HB, Jensen HH, Campbell SJ, Kirkpatrick J, Oldham NJ, Anthony DC, Davis BG. Nature. 2007; 446:1105.
12. a Yan LZ, Dawson PE. J. Am. Chem. Soc. 2001; 123:526. [PubMed: 11456564] b Crich D, Banerjee A. J. Am. Chem. Soc. 2007; 129:10064. [PubMed: 17658806] c Wan Q, Danishefsky SJ. Angew. Chem. Int. Ed. 2007; 46:9248.d Haase C, Rohde H, Seitz O. Angew. Chem. Int. Ed. 2008; 47:6807.e Chen J, Wan Q, Yuan Y, Zhu J, Danishefsky SJ. Angew. Chem. Int. Ed. 2008; 47:8521.f Bernardes GJ, Chalker JM, Errey JC, Davis BG. J. Am. Chem. Soc. 2008; 130:5052.g Pentelute BL, Kent SBH. Org. Lett. 2007; 9:687.
13. Spokoyny AM, Zou Y, Ling JJ, Yu H, Lin Y-S, Pentelute BL. J. Am. Chem. Soc. 2013; 135:5946. [PubMed: 23560559]
14. Becer CR, Babiuch K, Pilz D, Hornig S, Heinze T, Gottschaldt M, Schubert US. Macromolecules. 2009; 42:2387.
15. a Hayes JD, Flanagan JU, Jowsey IR. Annu. Rev. Pharmacol. Toxicol. 2005; 45:51.b Armstrong RN. Chem. Res. Toxicol. 1997; 10:2.
16. GST from equine liver, Sigma Aldrich catalog number: G6511. See SI for a complete description of experimental details regarding the choice and purification of GST utilized in this work.
17. Suzuki T, Ota Y, Kasuya Y, Mutsuga M, Kawamura Y, Tsumoto H, Nakagawa H, Finn MG, Miyata N. Angew. Chem. Int. Ed. 2010; 49:6817.
18. Dirksen A, Dawson PE. Bioconjugate Chem. 2008; 19:2543.
19. a Kurpiers T, Mootz HD. Angew. Chem. Int. Ed. 2007; 46:5234.b Nathani RI, Moody P, Chudasama V, Smith MEB, Fitzmaurice RJ, Caddick S. Chem. Sci. 2013; 4:3455.
20. Driggers EM, Hale SP, Lee J, Terrett NK. Nat. Rev. Drug Discovery. 2008; 7:608.
21. White CJ, Yudin AK. Nat. Chem. 2011; 3:509. [PubMed: 21697871]
22. a Adang AEP, Brussee J, Meyer DJ, Coles B, Ketterer B, Van d. G. A. Mulder GJ. Biochem. J. 1988; 255:721.b Adang AEP, Brussee J, Van d. G. A. Mulder GJ. Biochem. J. 1990; 269:47.

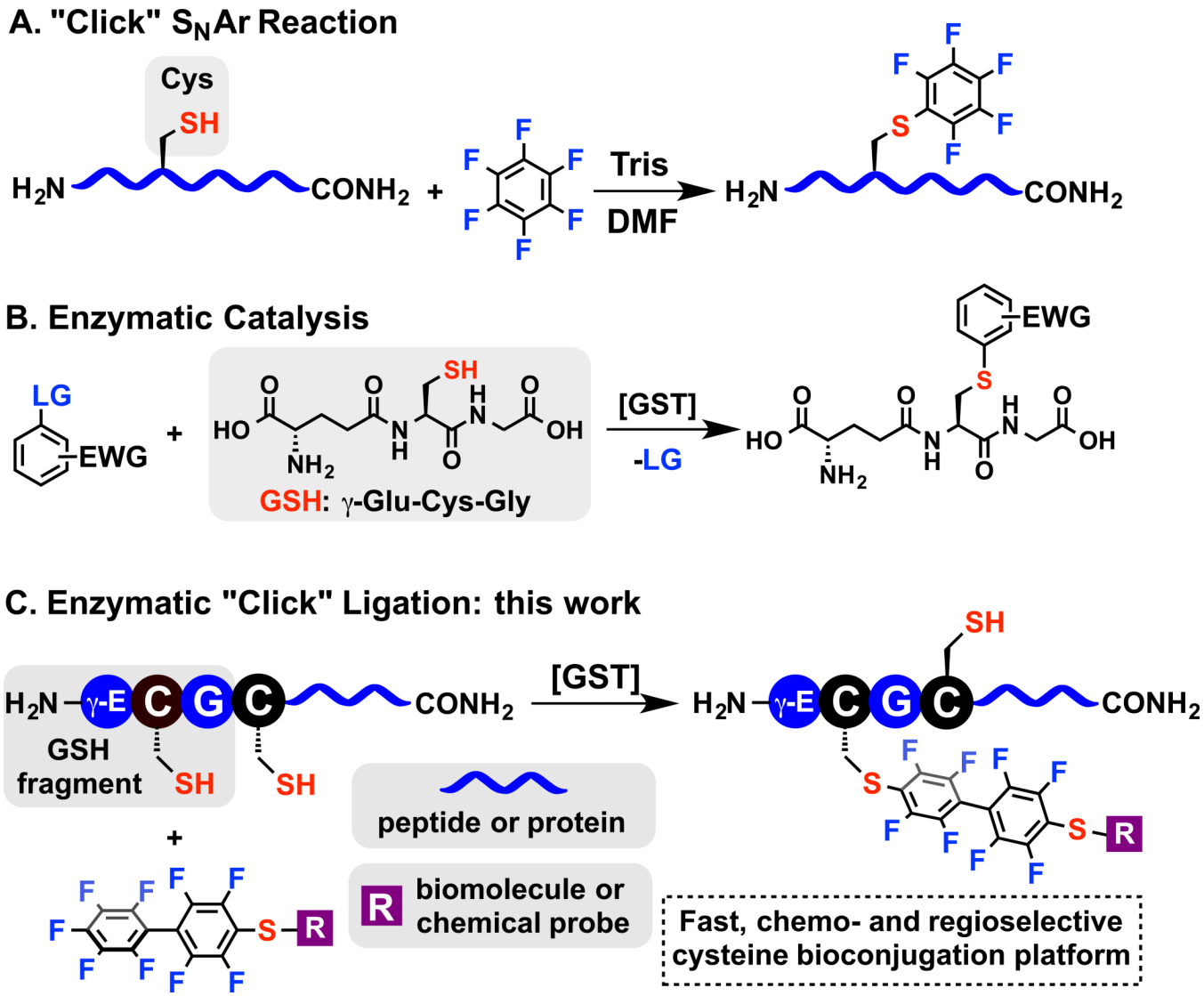


Figure 1.

A) Peptide “click” modification via perfluoroaryl-cysteine S_NAr chemistry. B) GST-catalyzed S_NAr reaction. GST catalyzes the conjugation of activated aromatic electrophiles to GSH cysteine thiol. EWG: electron withdrawing group, LG: leaving group. C) GST - catalyzed cysteine arylation. GST catalyzes the conjugation of probes bearing 4-mercaptoperfluoro-biphenyl moiety (Cys-**III**) to the N-terminal γ-Glu-Cys-Gly sequence of a peptide or a protein. Amino acids are shown in single-letter code. γ-E stands for γ-glutamyl.

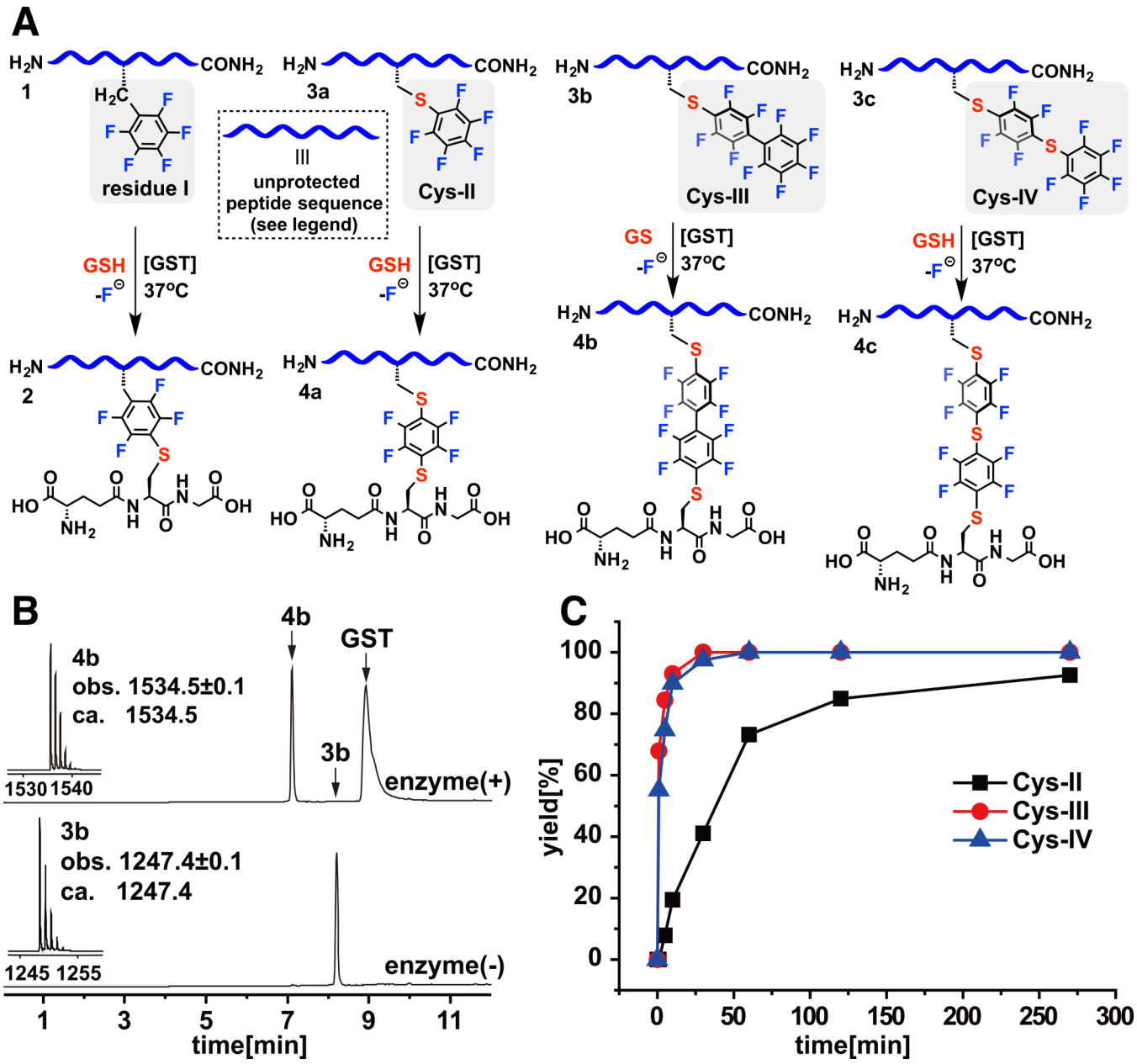
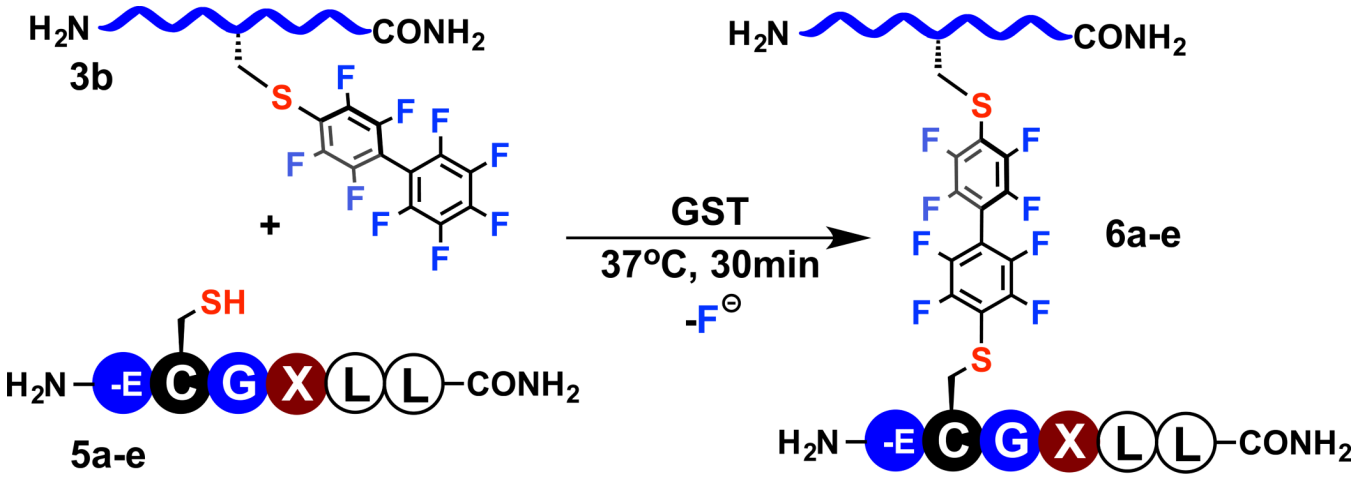


Figure 2.

A) GST-catalyzed conjugation of GSH with peptides containing perfluoroaromatic electrophilic residues. Peptide sequence of **1**: NH_2-ITPCNLLF*YYGKKK-$CONH_2$, F* stands for L-pentafluorophenylalanine; peptide sequences of **3a-c**: H_2NVTLPSTC*GAS-$CONH_2$, C* refers to the modified cysteine. Reaction conditions: 1 mM **1** or **3a-c**, 1 mM GSH, 2 mg/mL GST, 20 mM *tris*(2-carboxylethyl)phosphine hydrochloride (TCEP·HCl), 0.1 M phosphate buffer, pH 8.0, 37 °C. B) LCMS analysis of crude reaction with peptide containing Cys-**III** residue after 30 minutes. C) Rates of formation of GSH-conjugated product with different electrophiles (see Figure S14 for complete LCMS analysis). Yields were determined by LCMS analysis (see SI).



Entry	X		Yield (%)
1	Gly	**5a**	99
2	Leu	**5b**	86 (99*)
3	His	**5c**	98 (99*)
4	Lys	**5d**	68 (99*)
5	Pro	**5e**	67 (99*)

Figure 3.

Conjugation of peptide **3b** containing Cys-**III** residue (H_2N-VTLPSTC*GAS-$CONH_2$, C* refers to the modified cysteine) to peptides with N-terminal γ-Glu-Cys-Gly sequence featuring varied neighboring amino acid residues under GST-catalysis. Yields were determined by LCMS analysis at λ=280 nm (see SI) after 30 minutes. Reaction conditions: 2 mM **3b**, 1 mM **5a-e**, 2 mg/mL GST, 20 mM TCEP·HCl, 0.1 M phosphate, pH 8.0, 37 °C. [*] Yields for reactions at 120 minutes.

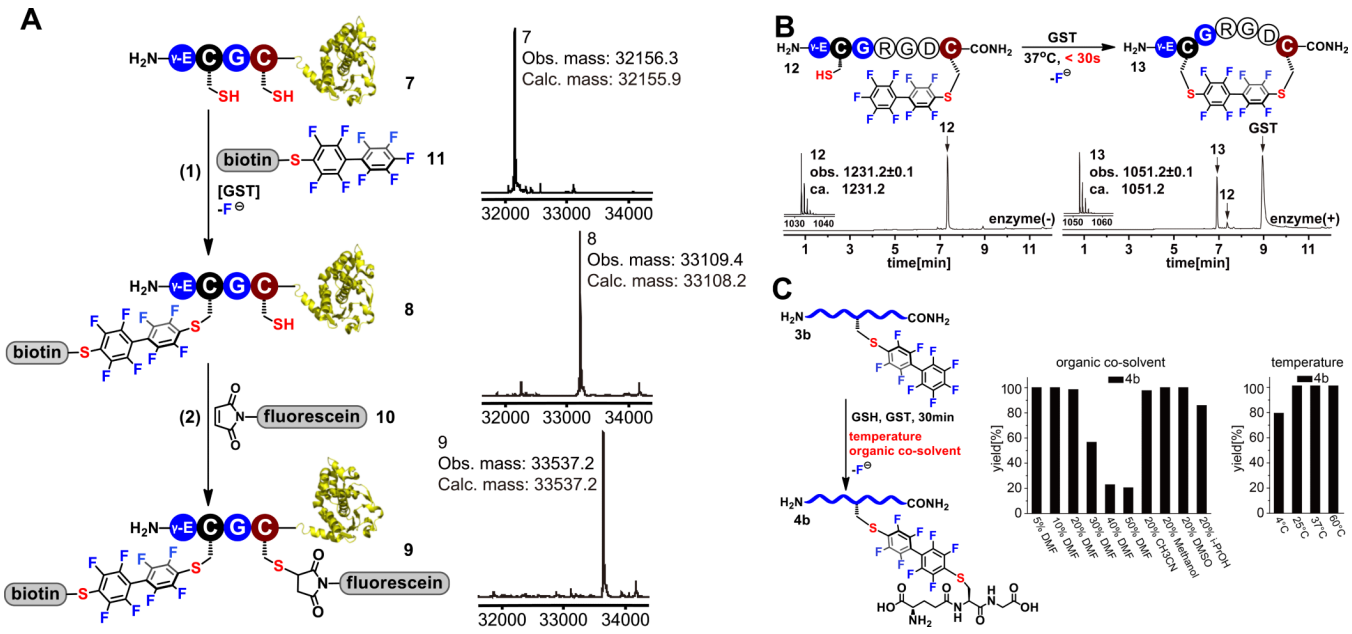


Figure 4.

Synthetic utility of GST-catalyzed cysteine modification. LCMS traces are shown as total ion current. Yields were determined by LCMS analysis at λ=280 nm (see SI). A) Sequential labeling of protein **7** containing two Cys residues with biotin and fluorescein probes. Crystal structure shown is the N-terminal domain of anthrax toxin lethal factor 1-263, PDB ID: 1J7N. Reaction conditions: (1) 0.5 mM **11**, 26 μM **7**, 2 mg/mL GST, 20 mM TCEP·HCl, 0.1 M phosphate, pH 8.0, 37 °C, 2 hours; (2) 13 μM **8**, 100 μM **10**, 0.1 M phosphate, pH 6.0, room temperature, 10 minutes. B) Peptide macrocyclization catalyzed by GST. Reaction conditions: 1 mM **12**, 2 mg/mL GST, 20 mM TCEP·HCl, 0.1 M phosphate, pH 8.0, 37 °C. C) GST-catalyzed conjugation of GSH to peptide containing Cys-**III** residue with mixed solvent at variable conditions and temperatures. Reaction conditions were the same as those in Figure 1 except for solvents (Left panel, percentages shown as volume ratio) and temperatures (right panel), see Figure S15 for complete LCMS analysis. DMF: dimethylformamide, DMSO: dimethylsulfoxide, i-PrOH: isopropanol.