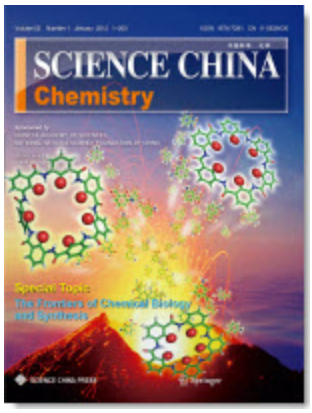


Turning Promiscuous Acyl Donors into Site-Selective Tools for Protein Modification

Journal:	*SCIENCE CHINA Chemistry*
Manuscript ID	SCC-2026-1048.R2
Manuscript Type:	Article
Date Submitted by the Author:	03-Jul-2026
Complete List of Authors:	Wang, Peng; South China University of Technology Liu, Junjie; Shenzhen Bay Laboratory Huang, Yang; Zunyi Medical University Lu, Anjing; Zunyi Medical University Wu, Chuanliu; Department of Chemistry, College of Chemistry and Chemical Engineering Tsai, Yu-Hsuan; Shenzhen Bay Laboratory He, Chunmao; Department of Chemistry and Chemical Engineering
Keywords:	Protein modification, Site-selective, Acyl donor, Live-cell labeling, Native chemical ligation
Speciality:	Chemical Biology
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Figures_and_Schemes.rar	

SCHOLARONE
MANUSCRIPTS

Turning Promiscuous Acyl Donors into Site-Selective Tools for Protein Modification

Peng Wang[a], Junjie Liu[b], Yang Huang[c], Anjing Lu[c], Chuanliu Wu[d*], Yu-Hsuan Tsai[b*], and Chunmao He[a*]

[a]School of Chemistry and Chemical Engineering, South China University of Technology, Guangzhou 510640, China

[b]Institute of Molecular Physiology, Shenzhen Bay Laboratory, Shenzhen 518132, China

[c]Key Laboratory of Biocatalysis & Chiral Drug Synthesis of Guizhou Province, Generic Drug Research Center of Guizhou Province, School of Pharmacy, Zunyi Medical University, Zunyi 563000, China

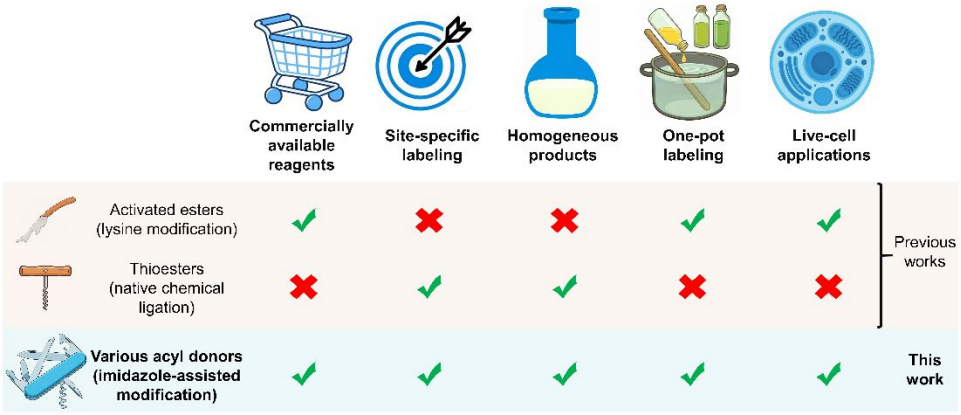
[d]The MOE Key Laboratory of Spectrochemical Analysis and Instrumentation, State Key Laboratory of Physical Chemistry of Solid Surfaces, Department of Chemistry, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China

*Corresponding authors: CW, chlwu@xmu.edu.cn; YHT, tsai.y-h@outlook.com; CH, hecm@scut.edu.cn

For Review Only

Abstract: The ability to modify proteins with precision has transformed modern biology and enabled a multi-billion-dollar biotherapeutics industry. Yet the most accessible conjugation reagents, such as N-hydroxysuccinimide (NHS) esters and other activated esters, remain fundamentally limited by the lack of site selectivity. As a result, they are typically used for random, heterogeneous labeling despite their fast kinetics and mild reaction conditions. In contrast, native chemical ligation offers exquisite selectivity for N-terminal cysteine, but it is often slow, requires cytotoxic additives, and relies on thioester reagents that can be laborious to prepare. Here we report an imidazole-assisted modification that merges the operational simplicity and live-cell compatibility of common acyl donors with the selectivity of native chemical ligation. This method enables efficient, site-selective formation of native amide bonds from readily available acyl donors under mild conditions. We demonstrate its utility in one-pot construction of multifunctional protein–drug conjugates, direct chemical modification of phage-displayed libraries without loss of infectivity, and fluorescent labeling of membrane proteins on live mammalian cells. By turning ubiquitous commercial acyl donors into plug-and-play reagents for site-selective bioconjugation, this strategy expands the practical toolkit for protein functionalization and is poised to streamline applications ranging from fundamental research to therapeutic development.

Keywords: Protein modification; Site-selective; Acyl donor; Live-cell labeling; Native chemical ligation



1 Introduction

The precise installation of synthetic probes on proteins is a cornerstone of modern chemical biology and pharmaceutical research [1–4]. It enables imaging of proteins in their native context, tracking their dynamics, and constructing therapeutic modalities such as antibody-drug conjugates. For decades, the workhorse reagents for protein conjugation have been simple and widely available acylating agents, most notably N-hydroxysuccinimide (NHS) esters and other activated esters, as well as anhydrides (There are more than 65,000 commercially available NHS esters, about 1,000 pentafluorophenyl (PFP) esters, about 7,000 *p*-nitrophenyl (PNP) esters, and more than 41,000 anhydrides, as shown in **Table S1**) [5]. These reagents react rapidly under mild conditions but lack site selectivity. As a result, they typically produce heterogeneous mixtures arising from modification of multiple lysine residues. This heterogeneity is a major liability in applications that require a defined drug-to-antibody ratio or a precisely positioned modification, where conjugation site and stoichiometry can directly affect efficacy, stability, and safety.

To address this, many highly selective bioorthogonal reactions have been developed, including the Staudinger ligation [6], azide-alkyne cycloaddition [7,8], and inverse electron-demand Diels-Alder reactions [9–11]. While these approaches offer excellent control, they require prior installation of a bioorthogonal handle. This often relies on chemical synthesis, genetic incorporation of unnatural amino acids, or other specialized steps [12,13], which can limit accessibility and throughput. In this context, the 1,2-aminothiol of N-terminal cysteine (NCys) is a particularly attractive handle because it can be targeted on canonical proteins without introducing unnatural functionalities [14–18]. This feature has motivated reagents that preferentially engage NCys, including thioesters [19,20], benzaldehyde derivatives [21–23], nitriles [24,25], and others [26–30]. Among these, thioesters enable native chemical ligation (NCL), which forms a native amide bond at NCys and has transformed protein chemical synthesis. However, NCL is often slow, commonly requires high concentrations of thiol additives that can be cytotoxic, and depends on thioester reagents that are not broadly available and are often laborious to prepare [31]. In fact, most

For Review Only

of the aforementioned NCys-selective reagents require nontrivial synthesis, which constrains adoption beyond specialist laboratories, particularly for biological applications.

These considerations highlight a persistent tradeoff among three desirable features for precision bioconjugation: broad commercial availability of reagents, site selectivity on canonical functionality, and mild conditions compatible with living systems. Recent advances that adapt NHS esters for selective labeling of NCys through an NCL-like pathway have moved toward this goal. However, they still require high concentrations of thiol additives such as sodium 2-mercaptoethanesulfonate (MesNa) or 4-mercaptophenylacetic acid (MPAA), which limits compatibility with live-cell settings [32,33].

Here we set out to combine the best attributes of these approaches. Inspired by previous reports on *N*-acyl azole-aided NCL [34–38], as well as our work on thiol-free, expressed *N*-acyl pyrazole and cysteine ligation (APCL) for chemical protein synthesis (CPS) [39], we reasoned that a simple, biocompatible small molecule like imidazole or 3-methylpyrazole could act as both an activator and a selectivity filter. We found that imidazole fulfils this role. With imidazole added to the reaction buffer, a broad range of commercially available or readily prepared acyl donors, including NHS esters, PFP esters, PNP esters, anhydrides, and hydrazides, become selective NCys-labeling agents. With a greatly expanded off-the-shelf substrate scope compared to that of both azole-aided NCL and APCL strategies where only bespoke peptide/protein thioesters or hydrazides are allowed, the reaction developed herein proceeds without thiol additives, operates under mild conditions, and is thus broadly compatible with phage and live cells. It also streamlines the assembly of complex conjugates, effectively bridging the gap between chemical accessibility and biological precision (**Figs. 1** and **S1**).

For Review Only

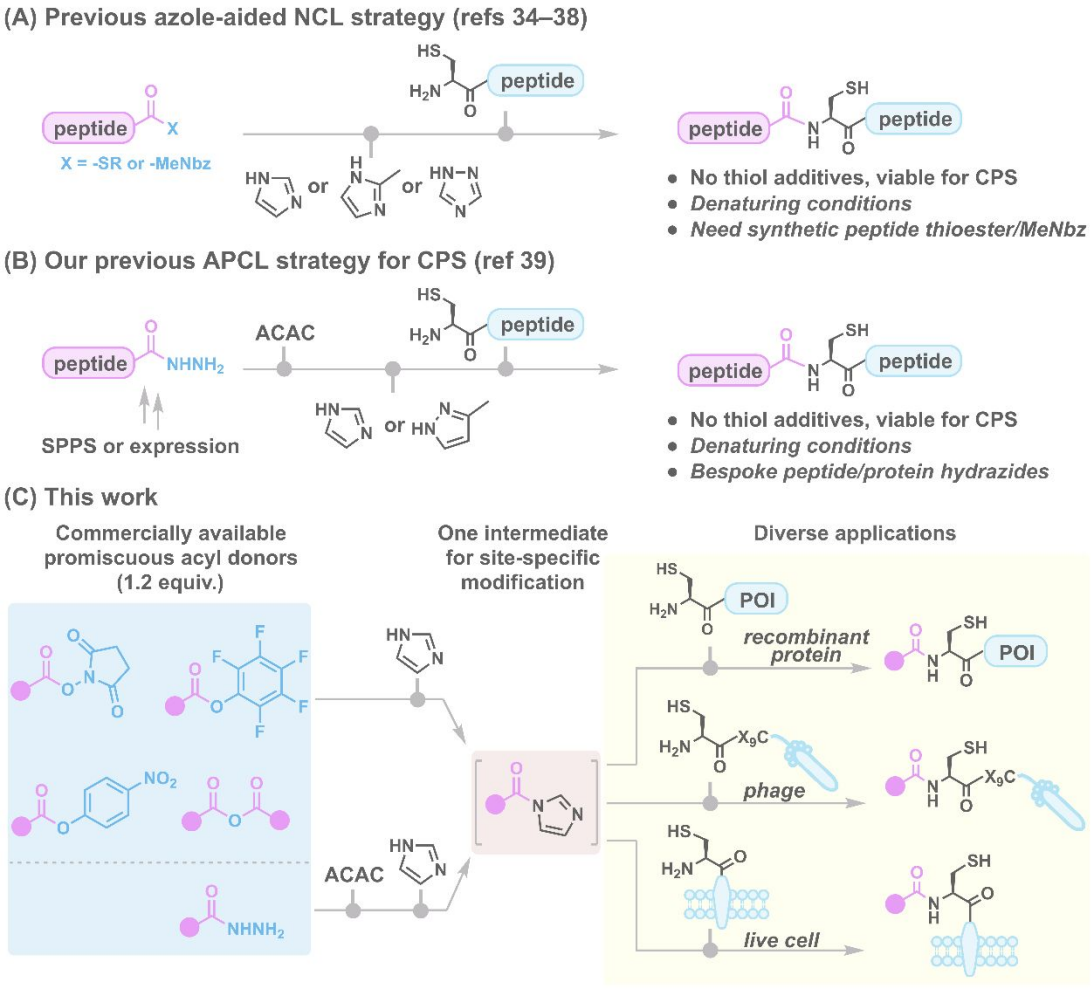


Fig. 1. Comparison of previous azole-aided NCL (A) and APCL (B) strategies with this work (C), which features a wide range of commercially available acyl donors for diverse applications, including phage-displayed peptide library modification and live-cell labeling (MeNbz: *N*-acyl-*N*'-methyl-benzimidazolinone).

2 Results and discussion

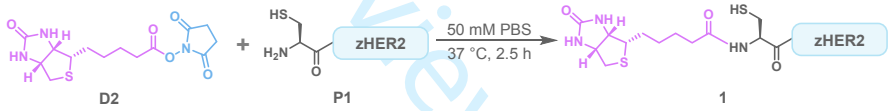
2.1 Imidazole: A simple solution for site-selective labeling

To test our hypothesis, we first targeted a folded protein, the HER2-targeting affibody zHER2 bearing an N-terminal cysteine (NCys-zHER2, **P1**). Biotin hydrazide (**D1**) was employed as the model substrate and reacted in 50 mM PBS (pH 6.5) with 20 mM tris(2-carboxyethyl)phosphine (TCEP) and either 2.5 M imidazole or 3-methylpyrazole. Imidazole consistently provided better and more reproducible outcomes and was thus adopted for subsequent investigations. As shown in **Fig. S3**, biotin hydrazide (**D1**, 8 equiv.) was activated by acetylacetone (ACAC) to its *N*-acyl 3,5-dimethylpyrazole derivative [40], which efficiently and selectively reacted with NCys on **P1** to form a single product, biotin-NCys-zHER2 (**1**), as confirmed by LC-MS/MS for selective NCys modification (**Fig. S4**). Reducing **D1** to 1.2 equiv. still yielded comparable conversion, underscoring the efficiency of the labeling reaction (**Fig. S5**). **D1** could also be

activated via $NaNO_2$ oxidation to its acyl azide derivative, a commonly employed intermediate in chemical protein synthesis [41,42]. The resulting acyl azide underwent selective NCys labeling in the presence of imidazole (**Fig. S6**).

Encouraged by these results, we evaluated the applicability of NHS esters—arguably the most frequently used protein labeling reagents but lacking site selectivity. When 8 equiv. of biotin-NHS (**D2**) were incubated with **P1** under the standard conditions (i.e., 2.5 M imidazole, pH 6.5), **1** was obtained as the major product with 81% HPLC yield after 2.5 hours (**Fig. S7**). In stark contrast, the same reaction without imidazole yielded a complex, intractable mixture of multiply labeled proteins (**Fig. S8**). This significant difference underscores the essential role of imidazole. It acts as a molecular translator, converting the highly reactive but promiscuous NHS ester into a more tempered acyl imidazole intermediate that is kinetically tuned to react selectively with the NCys thiol over the abundant lysine amines in the protein (see below). A systematic optimization confirmed the critical role of imidazole concentration and revealed that the reaction is efficient even with only 1.2 equiv. of the acyl donor **D2**, at low TCEP concentrations, and across a wide range of protein concentrations—conditions ideal for biological applications (**Table 1**).

Table 1. Screening of the reaction conditions using NCys-zHER2 (**P1**) and biotin-NHS (**D2**).



D2 P1 1

Entry	[**P1**] / mM	**D2** / equiv.	[TCEP] / mM	[imidazole] / M	pH	Yield of **1** (%)[a]
1	1	20	20	2.5	6.5	68
2	1	8	20	2.5	6.5	81
3	1	4	20	2.5	6.5	87
4	1	1.2	20	2.5	6.5	90
5	1	1.2	2	2.5	6.5	91
6	1	1.2	0	2.5	6.5	90
7	1	20	2	0	6.5	ND[b]
8	1	8	2	0	6.5	ND[b]
9	1	1.2	2	0	6.5	14[c]
10	1	1.2	2	0.1	6.5	55[c]
11	1	1.2	2	0.5	6.5	56[c]
12	1	1.2	2	1	6.5	63
13	1	1.2	2	2.5	7.2	82
14	2	1.2	2	2.5	6.5	82
15	0.1	1.2	2	2.5	6.5	77
16	0.05	1.2	2	2.5	6.5	66

For Review Only

Note: [a] HPLC yield; [b] ND: not determined, due to protein degradation; [c] detection of multi-labeled adducts in the reaction mixture.

2.2 A vast and accessible chemical space for bioconjugation

With the optimized conditions in hand, we explored the scope of compatible acyl donors to demonstrate the power of this approach (**Fig. 2**). We successfully and selectively labeled **P1** with a remarkable variety of cargos using their corresponding acyl donors. This included small-molecule tags (biotin, Cy5), structurally complex cyclic peptides [cR_{10}—a cell-penetrating peptide hydrazide with the sequence of K(r$R)_5$EGG-hydrazide (r: D-Arg, R: L-Arg), featuring a side-chain-to-side-chain cyclization between K and E], and another protein (ubiquitin) (**Figs. S12**, **S24–S40**). The acyl donors themselves were among the most accessible list of commercial reagents: NHS esters, PFP esters, PNP esters, and anhydrides (**D2**, **D6–D14**; **Tables S1** and **S2**). Intriguingly, for cR_{10}-hydrazide (**D3**), activation via $NaNO_2$ resulted in higher yield (**Fig. S27**), while no difference was observed for other hydrazides by activation through either ACAC condensation or $NaNO_2$ oxidation, offering flexibility. This broad substrate scope transforms our method from a single reaction into a versatile platform. Any researcher, regardless of synthetic skill, may now select a desired cargo, purchase its most common activated form, and conjugate it site-specifically to their target protein.

For Review Only

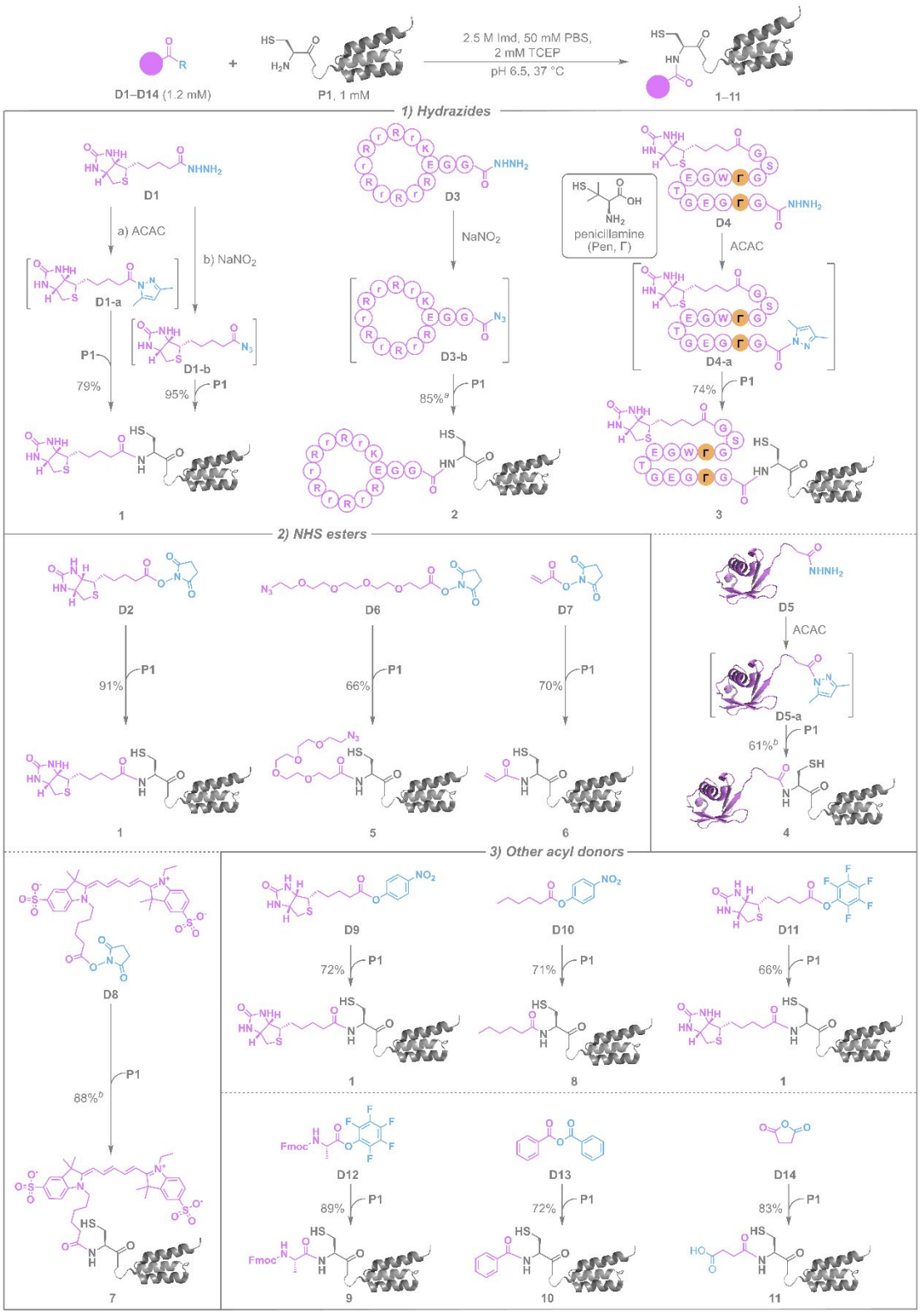


Fig. 2. Scope of the acyl donors used for the imidazole-assisted selective labeling (Imd = imidazole). Reaction conditions: 1 mM **P1**, 2.5 M imidazole, 2 mM TCEP, 50 mM PBS (pH 6.5), HPLC yields are shown. Note: [a] 2.4 equiv. of **D3**, and uppercase letters correspond to L-amino acids, lowercase letters correspond to D-amino acids; [b] 5 mM TCEP.

2.3 From simple peptides to complex proteins and dual-labeled conjugates

We further highlighted the utility of the selective labeling strategy with a panel of synthetic peptides and recombinant proteins of varying size and complexity (**Fig. 3**). Using biotin-NHS (**D2**) as the acyl donor, selective modification of NCys in peptide substrates **P2–P5** afforded conjugates **12–15** in 68% to 87% HPLC yields (**Figs. 3A** and **S41–S48**). Furthermore, the site-specific modification also proceeded smoothly with a number of NCys-bearing proteins, including Azurin (**P6**), SUMO (**P7**), and LmrR (**P8**). The single-modified protein conjugate was observed in all cases as the major product, with over 80% conversion to the resulting protein conjugates **16–18** (**Figs. S12** and **S49–S54**), and internal Lys and Cys residues remained untouched. This exquisite selectivity highlights the robustness of the imidazole-assisted mechanism in complex protein contexts.

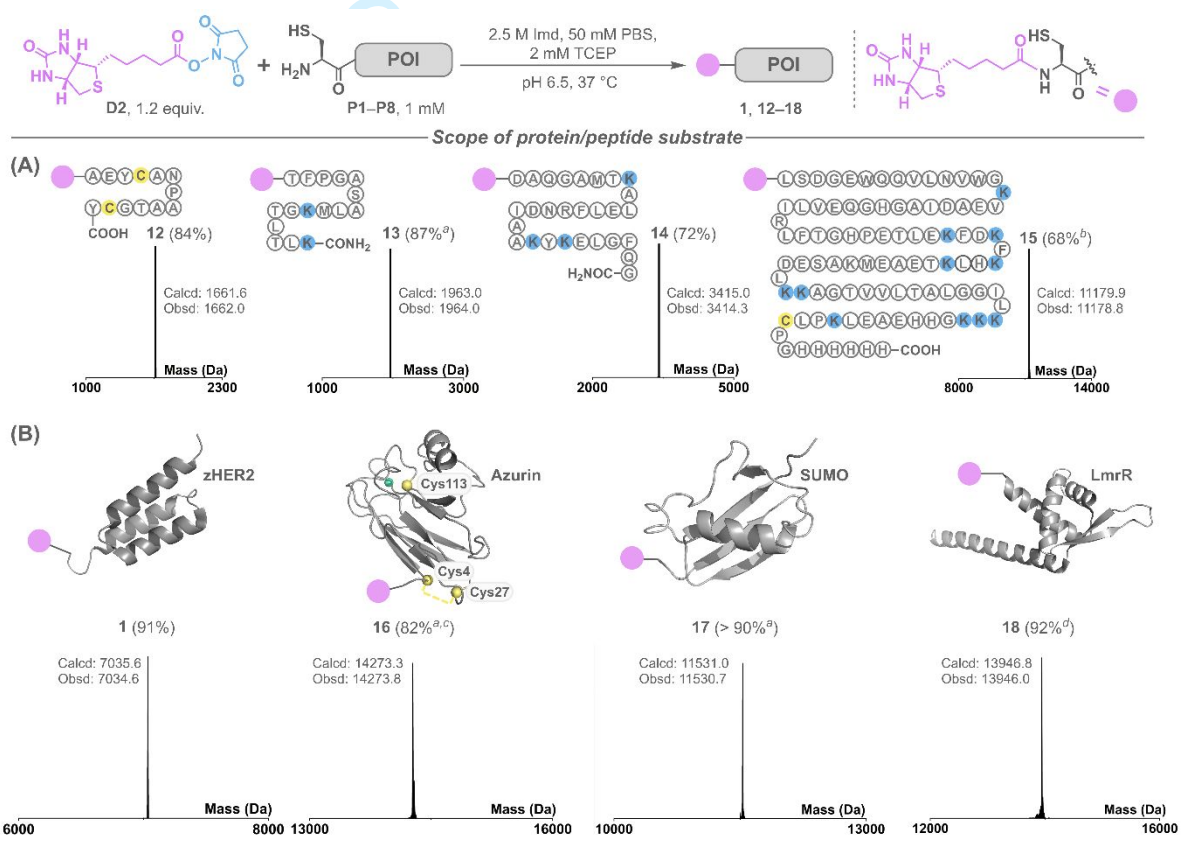


Fig. 3. Imidazole-assisted selective labeling of NCys peptides and proteins using **D2** as the acyl donor. Reaction conditions: 2.5 M imidazole, 2 mM TCEP, 50 mM PBS, pH 6.5 and HPLC yields are shown. Lys residues in the peptides are shown as blue circles. Cys residues in the peptides and proteins are shown as yellow circles or spheres, respectively. Disulfide bond in **16** between Cys4 and Cys27 is shown as a dashed line. Zinc ion in **16** is shown as a green-cyan sphere. Note: [a] substrate concentration, 0.5 mM; [b] 5 mM TCEP; [c] 1 M imidazole; [d] 0.1 mM **P8**.

The ability to install a single cargo selectively is powerful, but many advanced applications—from mechanistic studies to targeted drug delivery—require the precise attachment of two different functionalities [43]. The free thiol generated from the NCys after our labeling reaction provides a second, orthogonal handle for conjugation. We leveraged this for a streamlined, one-pot, sequential dual-labeling strategy. In a model study, mono-labeled conjugate **1** was obtained as shown previously. Since the reaction mixture was rather clean, the second bioconjugation with *N*-methyl maleimide (**M1**) was carried out in one pot through simple reagent addition, and the dual-labeled protein conjugate **19** was obtained after only one purification step, with an isolated yield of 64% (**Figs. 4** and **S55**). Similarly, we achieved the dual functionalization of **P1** with a cell-penetrating cR_{10} peptide and the cytotoxic drug MMAE (**20**) in an isolated yield of 30% (**Figs. 4** and **S56**), and the dual functionalization of **P1**, affording a theranostic conjugate, with a Cy5 dye and MMAE (**21**) in an isolated yield of 55% (**Figs. 4** and **S57**). It is noteworthy that this one-pot workflow is not feasible for NCL, because the high concentrations of thiol additives required for routine NCL would compete with maleimide coupling. This operational simplicity therefore represents a significant advantage over multi-step strategies that require intermediate purification.

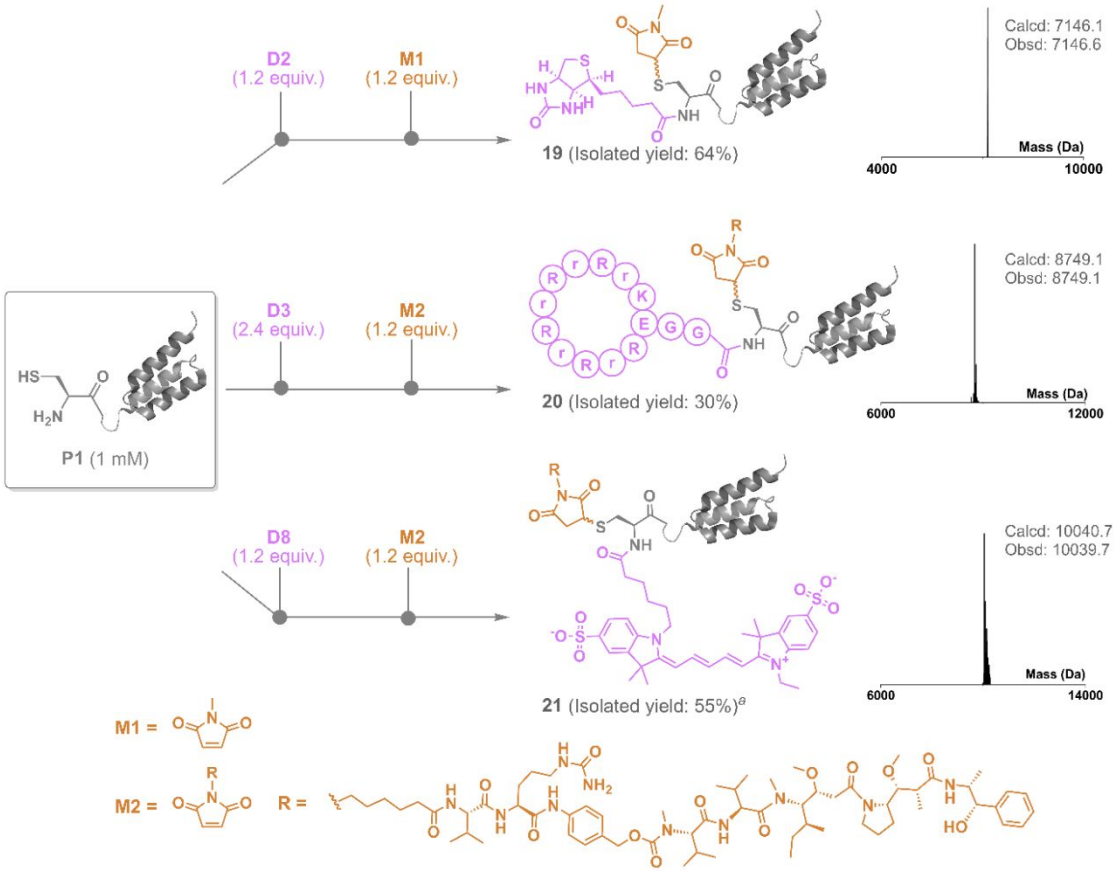
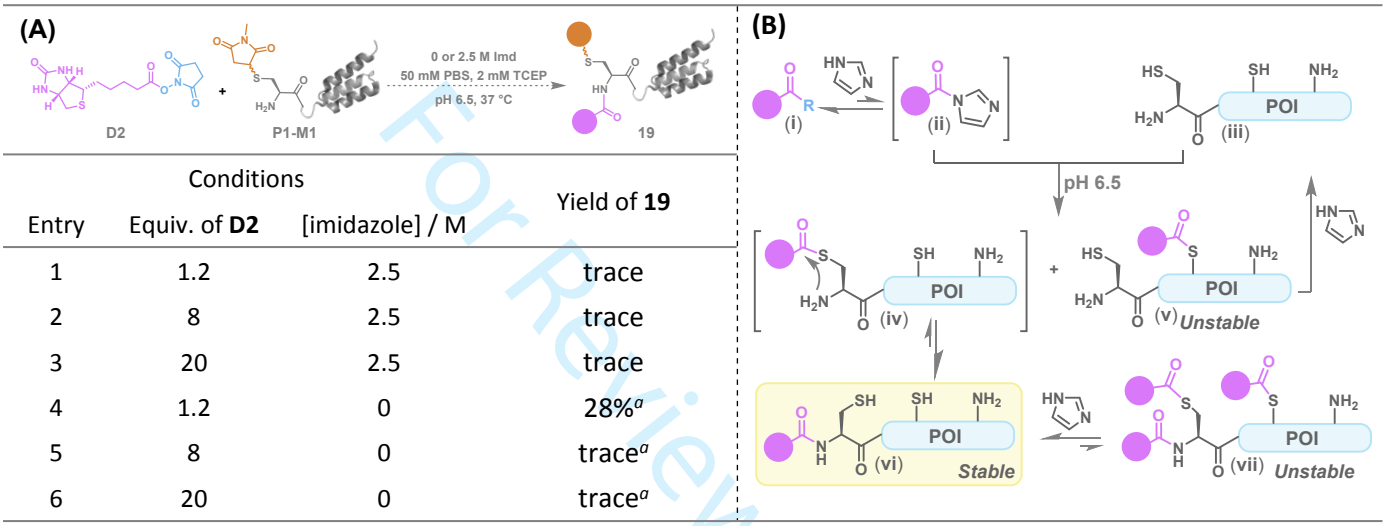


Fig. 4. Sequential one-pot dual labeling of **P1** using acyl donors and maleimides. Reaction conditions: 2.5 M imidazole, 2 mM TCEP, 50 mM PBS, pH 6.5. Note: [a] 5 mM TCEP.

2.4 A mechanistic insight: How imidazole enforces selectivity

To understand the origin of the selectivity, a series of experiments were performed (**Scheme 1**). When the thiol group of the NCys was capped by **M1** (**P1-M1**), reaction with 1.2 to 20 equiv. of **D2** only afforded a trace amount of the labeling product **19** (**Scheme 1A**, entries 1–3). This confirms that the NCys thiol is essential for the reaction. In the absence of imidazole, however, 1.2 equiv. of **D2** could initiate the labeling, albeit without selectivity as expected. Further increase of **D2** only led to a mixture of conjugates with varying numbers of modifications (**Scheme 1A**, entries 4–6), as zHER2 contains five lysine residues.

Scheme 1. Mechanistic investigation. (A) Control experiments. (B) Proposed reaction pathways.



Entry	Conditions: Equiv. of **D2**	Conditions: [imidazole] / M	Yield of **19**
1	1.2	2.5	trace
2	8	2.5	trace
3	20	2.5	trace
4	1.2	0	28%[a]
5	8	0	trace[a]
6	20	0	trace[a]

Note: [a] Lack of selectivity due to modification of lysine sidechains in **P1-M1**.

These experiments support a model where excess imidazole nucleophile rapidly converts the acyl donor (**i**) into *N*-acyl imidazole (**ii**). This intermediate is only weakly reactive towards protonated amines at our working pH of 6.5 [34,35], but is highly reactive towards the more nucleophilic thiol of the NCys [36]. This initiates an NCL-like transthioesterification with **iii**, yielding a stable amide bond (**vi**) after an *S*→*N* acyl shift. Any transiently formed thioester modifications, either with internal thiols (**v**) or with the now free N-terminal thiol (**vii**), may be rapidly scavenged by the imidazole present in the reaction mixture, as reported by others [37,38], redirecting them to the correct path (**Scheme 1B**). Notably, the formation and removal of such transient thioester species have been observed in the labeling of **P2** with **D2**, as shown in **Fig. S42** (peaks **12a–c**), which is another reason why imidazole has to be added in large excess. Importantly, the moderately reactive *N*-acyl imidazole (**ii**), even at high dosage, can hardly react with internal amines that are most probably protonated at our working pH of 6.5, as shown in **Scheme 1A**, entries 1–3 and **Figs. S58–S60**; by contrast, the acyl donor itself (e.g., biotin-NHS **D2**) can modify almost any amine groups non-selectively (**Scheme 1A**, entries 4–6; **Figs. S61–S63**), which is a well-known characteristic. The excess imidazole serves multiple roles: it is an activator, a selectivity filter, and a repair agent, resembling those of MPAA in NCL [44], and is thus the key to facilitating selective labeling of NCys protein with a diverse range of "naturally non-selective" acyl donors (**Scheme 1B**).

For Review Only

2.5 *In vivo* imaging and cytotoxicity of the theranostic conjugate 21

With the dual-labeled theranostic conjugate Cy5-MMAE-zHER2 (**21**) in hand, we first evaluated its application for analyzing the abundance of surface HER2 in different cell lines [45]. Incubation of cells with 100 nM **21** for 30 minutes was sufficient to reveal cell surface HER2 level by flow cytometry (**Fig. 5A**) and confocal microscopy analysis (**Fig. 5B**). We then evaluated the application of **21** for *in vivo* imaging. Immunodeficient mice were subcutaneously grafted with $HER2^{high}$ SK-OV-3 cells at the right rear limb and $HER2^{low}$ A549 cells at the left rear limb (**Fig. 5C**). About three hours after intravenous injection of **21**, strong fluorescence was observed at the right (i.e., $HER2^{high}$ SK-OV-3) but not at the left (i.e., $HER2^{low}$ A549), demonstrating the sensitivity and specificity of **21**. This was further confirmed by tissue and organ samples, where the highest fluorescence signal was observed at the SK-OV-3 site and no detectable signal at the A549 site (**Fig. 5D**). Lastly, we assessed the cytotoxicity of **21** to $HER2^{high}$ and $HER2^{low}$ cells (**Fig. 5E**). While MMAE itself is cytotoxic to all tested cell lines regardless of surface HER2 level, at concentrations higher than 100 nM, **21** preferentially killed $HER2^{high}$ SK-OV-3 and OE19 cells rather than $HER2^{low}$ A549 cells, validating its potential as a targeted theranostic agent. It is also noteworthy that MMAE is more cytotoxic than **21** in all tested cell lines, likely due to higher cellular uptake of the small molecule than of the protein conjugate. Collectively, as opposed to traditional NCL-type labeling methods [32,33] where high doses of thiol reagents are required, the results presented herein highlight the utility of our strategy as a streamlined means of accessing protein–drug conjugates via commercially available acyl donors or readily synthesized peptide/protein hydrazides for theranostic applications.

For Review Only

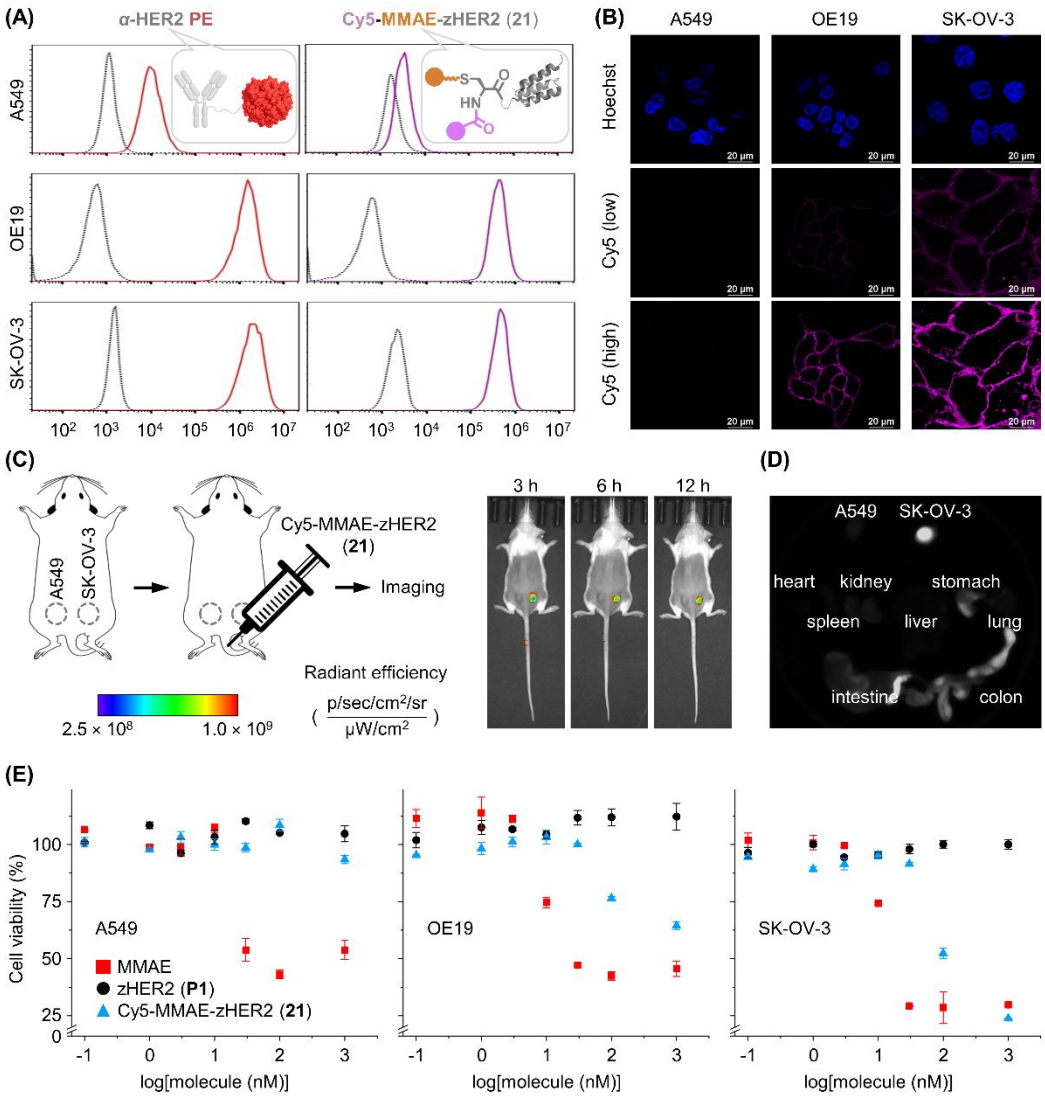


Fig. 5. Functional analysis of Cy5-MMAE-zHER2 (**21**) for HER2-specific imaging and cytotoxicity. (A) Flow cytometry analysis of surface HER2 on different cell lines by anti-HER2-phycoerythrin (*α*-HER2-PE) antibody conjugate (1:20 dilution, excitation = 561 nm, emission = 586 ± 7.5 nm, left) or **21** (100 nM, excitation = 640 nm, emission = 680 nm, right). Grey lines = non-treated cells. Colored lines = cells treated with *α*-HER2-PE (red) or **21** (purple). X axis = fluorescence intensity of PE (left) or Cy5 (right). Y axis = cell count. (B) Confocal microscopy analysis of cell surface HER2 on A549, OE19, and SK-OV-3 cells by staining with 100 nM **21** for 30 minutes. The fluorescence of Hoechst shows the nucleus. Cy5 fluorescence is shown in two settings (low: laser power 1.0, PMT HV 120; high: laser power 2.0, PMT HV 180). Scale bars of all images = 20 μm. (C) *In vivo* imaging of HER2^{high} SK-OV-3 cells by **21**. Female NCG mice (n=3) were subcutaneously grafted with A549 and SK-OV-3 cells on the left and right rear limbs of each mouse. After 10 days, 100 μL of 25 μM **21** was intravenously injected. Fluorescence images were acquired at 3, 6, and 12 h post injection. (D) Fluorescence image of the tumors and major organs of the mouse in (C) at 12 h post injection. Data of the other mice in (C) are shown in **Fig. S64**. (E) Cytotoxicity of **21** in A549, OE19, and SK-OV-3 cells, demonstrating its HER2-targeted toxicity.

2.6 Expanding the chemical universe of the phage-displayed peptide library

Display-based evolution of disulfide-rich multicyclic peptide scaffolds is a powerful strategy for discovering protein binders for biological applications [45–48]. Encouraged by the success in peptide and recombinant protein modification, we envisaged the strategy could be applied for constructing novel phage display libraries by expanding their chemical spaces that are largely limited to the 20 canonical amino acids. Recently, enzyme-mediated ligations have been used to construct phage-displayed peptide macrocycles [49,50]. However, the specific enzyme-recognition sequences also remain in the ligation product, leaving a scar of several amino acid residues. In contrast, we envision that our biocompatible imidazole-assisted modification can afford scarless products directly onto the phage surface, while unnatural amino acids can be readily incorporated concomitantly. For example, penicillamine (hereafter abbreviated as Pen or Γ, refer to **Fig. 2** for its structure) forms a mixed disulfide with Cys preferentially over a Pen–Pen disulfide [51,52]. Thus, conjugation of a $\Gamma X_7 \Gamma G$ peptide to phages displaying CX_9C sequences would afford $\Gamma X_7 \Gamma G C X_9 C$ adducts [29,53]. Since disulfide bond formation within a CXC motif (e.g., ΓGC in this case) is not favorable [54], oxidation of $\Gamma X_7 \Gamma G C X_9 C$ would afford bicyclic peptides only as a single pre-defined, isomer with [1-3, 2-4] disulfide pairing pattern.

For proof of concept, we applied this strategy to identify bicyclic binders of Keap1, a key therapeutic target for cancers and other diseases [55]. We first used a synthetic peptide biotin-GSGΓWGETGEGΓG-$NHNH_2$ (**D4**) to identify a suitable condition for labeling phages. We were able to obtain a 75% modification yield when incubating phages with 1 mM biotinylated peptide (activated by ACAC condensation) in the presence of 2 mM TCEP and 2.5 M imidazole for 30 minutes (**Figs. 6A**, **B**). The modification efficiency is comparable to that of enzyme-mediated ligations (68% for OaAEP1 [50]; 73% for SrtA [49]). Next, peptide hydrazide GSGΓWGETGEGΓG-$NHNH_2$ (**D4'**) was synthesized by SPPS (**Fig. S65**) and used for phage labeling and screening. After three rounds of selection against immobilized Keap1, we randomly picked five colonies for Sanger sequencing (**Fig. 6C**). Phage pools after each round of selection were also subjected to next-generation sequencing (**Fig. 6D** and **Tables S3**–**S5**). KP1, KP2, and KP3 identified by Sanger sequencing were also among the top 20 abundant sequences after the 3rd round of screening. We then synthesized KP1 and the ligated product Pen2KP1, and characterized the binding of Keap1 to the oxidized (i.e., bicyclic o-Pen2KP1 and monocyclic o-KP1) or reduced (i.e., linear r-Pen2KP1 and r-KP1) sequences (**Figs. 6E**, **S66**–**S71**). Surface plasmon resonance (SPR) analysis confirmed o-Pen2KP1 as a stronger binder (K_D = 4.4 nM, an affinity comparable to the well-established Keap1 inhibitors reported in the literature [56]. **Figs. 6F** and **S71**) than r-Pen2KP1 (no detectable binding up to 10 μM, **Fig. S70**) or o-KP1 (K_D = 25 nM, **Fig. S68**). The complex model of Keap1 and o-Pen2KP1 by AlphaFold2 (**Figs. 6G** and **S72**) suggests that the second loop identified by screening (i.e., KP1 or RNEETGEEE) binds to Keap1 in a similar fashion as the ETGE motif on Nrf2, the physiological binder of Keap1 [57,58]. Indeed, mutation of Thr18, a critical interacting residue [59], to Ala in the second loop (o-Pen2KP1TA) lowered its binding affinity by more than 10-fold, whereas the Thr8Ala mutation (o-Pen2KP1AT) in the first loop showed negligible impact (**Figs. 6E**, **S73**–**S78**). Nevertheless, the Trp5 and Glu7 in the first loop (i.e., WGETGE) likely form hydrophobic and ionic interactions with Phe478 and

For Review Only

His436 of Keap1 (**Fig. S72**), respectively, conferring higher binding affinity to o-Pen2KP1 than o-KP1. This result is a powerful validation of our approach, which substantially broadens the chemical diversity of phage-displayed libraries, enabling the rapid discovery of a novel, potent, and structurally defined bicyclic peptide lead.

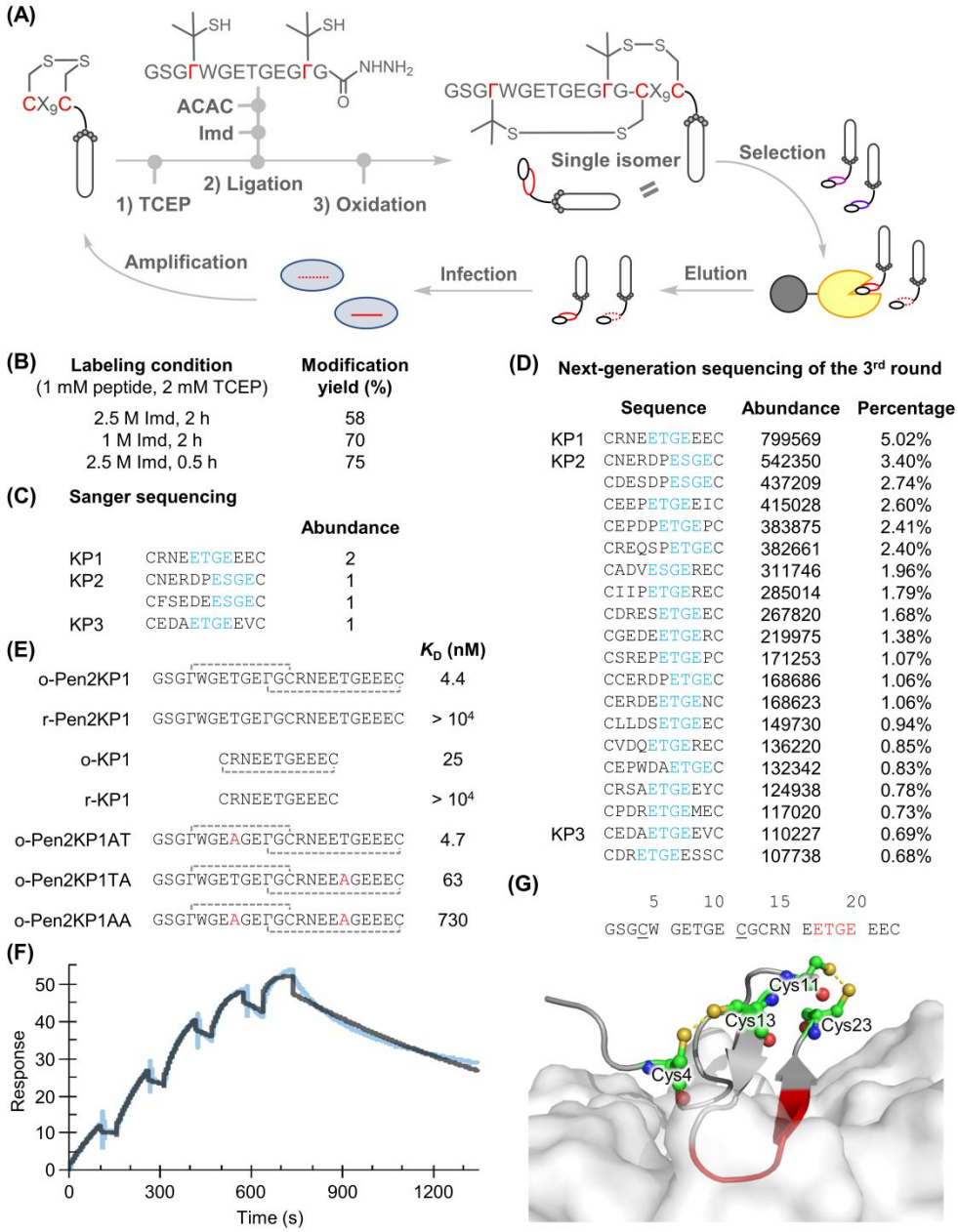


Fig. 6. Selective modification of phage proteins for screening bicyclic peptide binders. (A) Scheme for the experimental operation. Phages employed here display a CX_9C (C = Cys, X = any) sequence at the N-terminal of g3p proteins. During the selection step, phages were incubated with target proteins (yellow) immobilized on magnetic beads (grey circle), and the non-binding phages were removed. (B) Investigation of the labeling conditions. (C) Sequences of five randomly selected clones after the 3rd round of selection. (D) Next-generation sequencing results of the phage pool after the 3rd round of selection. (E) Sequences characterized for Keap1 binding affinity. (F) SPR result of o-Pen2KP1 to Keap1 using the single-cycle

kinetics mode. [o-Pen2KP1] = 12.5, 25, 50, 100, 200 nM. Y axis = relative response unit (RU). (G) Predicted model of the binding of o-Pen2KP1 to Keap1, highlighting [1-3,2-4] disulfide pairing in o-Cys2KP1. The sequence of o-Cys2KP1 is shown at the top with residues corresponding to Pen in o-Pen2KP1 underlined, which were changed into Cys to facilitate AlphaFold2 modeling. Side chains of the Cys residues in o-Cys2KP1 are shown in the sphere model. The key interacting residues ETGE are highlighted in red.

2.7 Live-cell labeling with the reagent Cy5-NHS (D8)

The success of modifying surface peptides on phages prompted us to explore this reaction for selective fluorescence labeling of surface proteins on mammalian cells with Cy5-NHS (**D8**). HEK293T cells were transiently transfected with a plasmid encoding a model protein C-HA-Nlgn3, which contains an N-terminal signal peptide, followed by a cysteine residue (C), an HA tag for immunoblotting, and the cell-surface protein neuroligin 3 (Nlgn3) [30]. As the signal peptide is removed, the subsequent Cys residue becomes the first amino acid residue of the mature protein on the cell surface (**Fig. 7A**). In-gel fluorescence analysis confirmed both DTT and TCEP can be used as the reducing agents to ensure the free thiol form of the NCys, and 1 M imidazole was sufficient to facilitate the selective labeling in live cells. Critically, incubation with 1 M imidazole had a negligible impact on HEK293T cell viability (**Fig. 7B**). It is noteworthy that upon completion of the labeling, prompt removal of imidazole might further help to avoid any potential perturbations to the cells (see the SI for experimental details). Negligible Cy5 fluorescence was observed in non-transfected cells (i.e., lanes 1 and 3 of **Fig. 7C**), confirming the specificity of the reaction. Selectivity and specificity of the labeling were also confirmed by confocal microscopy analysis (**Fig. 7D**), where labeling was only observed on transfected cells (i.e., solid arrows) but not non-transfected cells (i.e., hollow arrows). The fluorescence of Cy5 faithfully recapitulates the membrane-localized mCherry fluorescence. As the negatively charged Cy5-NHS (**D8**) is cell-impermeable, the trace amount of intracellular Cy5 fluorescence likely resulted from trafficking and endocytosis of the labeled proteins, instead of direct labeling of intracellular proteins [30]. Overall, it demonstrates that our strategy can turn a simple, promiscuous NHS-ester dye into a selective, live-cell-compatible labeling tool, opening new avenues for studying membrane protein dynamics.

For Review Only

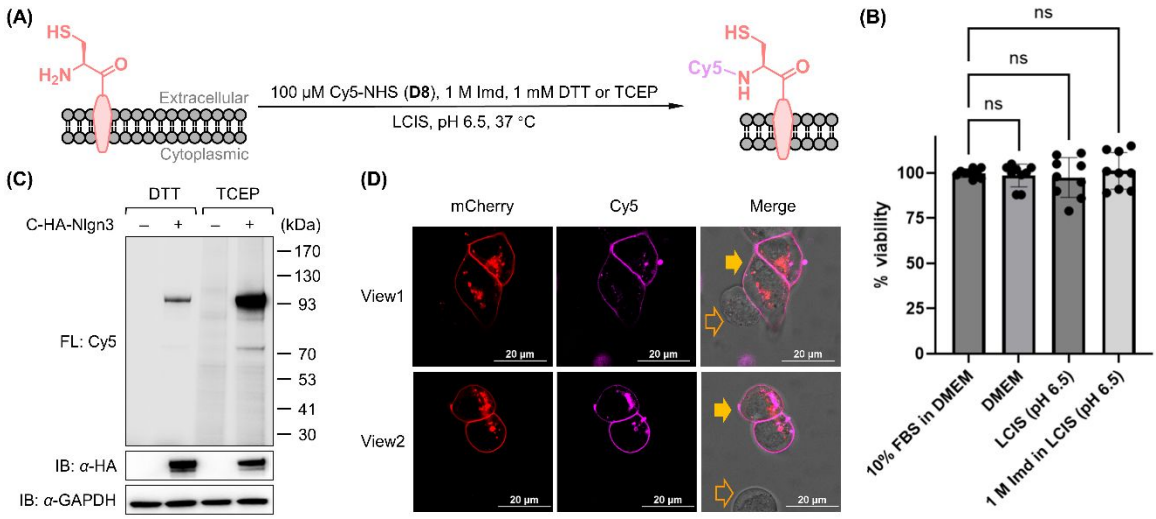


Fig. 7. Selective labeling of surface proteins on live mammalian cells. (A) Schematic presentation of the experimental procedure. Briefly, cells were transiently transfected with a plasmid encoding a cell-surface protein with an N-terminal Cys residue. Labeling of the target protein was performed under the indicated conditions for 30–60 minutes. LCIS = Live cell imaging solution. (B) Effect of 1 M Imd on HEK293T cell viability. Data represent mean ± standard deviation from three independent experiments, each performed in triplicate. No statistically significant differences (ns) were observed for any condition compared with complete medium. (C) In-gel fluorescence analysis of HEK293T cells labeled with Cy5-NHS for one hour. (D) Representative confocal microscopy images of HEK293T cells expressing C-HA-Nlgn3-mCherry labeled with Cy5-NHS for 30 minutes. Labeled (transfected) and unlabeled (non-transfected) cells are shown in solid and hollow yellow arrows, respectively. Scale bars of all images = 20 μm.

3 Conclusions

We report an imidazole-assisted chemical ligation that enables efficient, site-selective protein and peptide modification at N-terminal cysteine using readily available acyl donors. By adding imidazole to the reaction buffer, common activated esters that are otherwise used for random lysine acylation, such as NHS and related activated esters, are converted into plug-and-play reagents for selective NCys conjugation. This strategy merges the operational simplicity and fast kinetics of off-the-shelf acyl donors with the selectivity of native chemical ligation, while avoiding the high concentrations of thiol additives that often limit NCL in biological settings.

The scope of this method is broad. We demonstrate site-selective installation of diverse cargos, including small molecules, peptides, and proteins, and we show that the retained thiol handle supports one-pot sequential dual functionalization to access multifunctional conjugates, including protein–drug and theranostic constructs. We also extend the chemistry to direct modification of phage-displayed libraries without loss of infectivity, enabling seamless introduction of unnatural elements without enzymatic scars and supporting discovery of multicyclic peptide ligands. Finally, the mild, additive-free conditions

For Review Only

are compatible with labeling membrane proteins on live mammalian cells, providing a practical alternative to approaches that require preinstalled bioorthogonal handles.

By converting ubiquitous commercial acyl donors into selective reagents for native amide bond formation, this work removes a key bottleneck in precision bioconjugation. We anticipate that the simplicity and generality of this imidazole-assisted platform will accelerate applications ranging from fundamental chemical biology to discovery workflows and therapeutic conjugate development.

Conflict of interest

The authors declare no conflict of interest. A Chinese patent (No. ZL202511292123.1) has been granted covering the method described herein.

For Review Only

Acknowledgements

We are grateful for the financial support from the National Natural Science Foundation of China (22277029 and 22577031 to CH and 22277079 to YHT), Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515012078). Research funds were also provided by the Hubei Key Laboratory of Bioinorganic Chemistry and Materia Medica (BCMM202502) and Shenzhen Bay Laboratory (SZBL). We appreciate support from the Core Facilities at SZBL. We are grateful to Prof. Yi-Ming Li from Hefei University of Technology for providing the ubiquitin plasmid, Dr. Peisi Liao for helpful discussions, and Dr. Yi Zhou, Ms. Yuanmei Sun, Mr. Pengpeng Li, Ms. Shuting Lin, Ms. Jieyu Liao and Mr. Gaocan Dai for technical assistance.

Supporting Information

The supporting information is available online at http://chem.scichina.com and http://link.springer.com/journal/11426. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Krall N, Da Cruz FP, Boutureira O, Bernardes GJL. *Nat Chem*, 2016, 8: 103–113
2 Hoyt EA, Cal PMSD, Oliveira BL, Bernardes GJL. *Nat Rev Chem*, 2019, 3: 147–171
3 Beck A, Goetsch L, Dumontet C, Corvaïa N. *Nat Rev Drug Discovery*, 2017, 16: 315–337

4 Lin S, He C. *Chin Chem Lett*, 2018, 29: 1017–1021
5 Chauhan P, V R, Kumar M, Molla R, Mishra SD, Basa S, Rai V. *Chem Soc Rev*, 2024, 53: 380–449
6 Saxon E, Bertozzi CR. *Science*, 2000, 287: 2007–2010
7 Kolb HC, Finn MG, Sharpless KB. *Angew Chem Int Ed*, 2001, 40: 2004–2021
8 Sletten EM, Bertozzi CR. *Angew Chem Int Ed*, 2009, 48: 6974–6998
9 Blackman ML, Royzen M, Fox JM. *J Am Chem Soc*, 2008, 130: 13518–13519
10 Oliveira BL, Guo Z, Bernardes GJL. *Chem Soc Rev*, 2017, 46: 4895–4950
11 Chen J, Dai G, Duan S, Huang Y, Wu Y, Xie Z, Tsai Y-H. *Chem – Asian J*, 2024, 19: e202400824
12 Hayes HC, Luk LYP, Tsai Y-H. *Org Biomol Chem*, 2021, 19: 3983–4001
13 Chen J, Tsai Y-H. *J Mol Biol*, 2022, 434: 167424
14 Witus LS, Netirojjanakul C, Palla KS, Muehl EM, Weng C-H, Iavarone AT, Francis MB. *J Am Chem Soc*, 2013, 135: 17223–17229
15 Rosen CB, Francis MB. *Nat Chem Biol*, 2017, 13: 697–705
16 MacDonald JI, Munch HK, Moore T, Francis MB. *Nat Chem Biol*, 2015, 11: 326–331
17 Chan AO-Y, Ho C-M, Chong H-C, Leung Y-C, Huang J-S, Wong M-K, Che C-M. *J Am Chem Soc*, 2012, 134: 2589–2598
18 Chen D, Disotuar MM, Xiong X, Wang Y, Chou DH-C. *Chem Sci*, 2017, 8: 2717–2722
19 Dawson PE, Muir TW, Clark-Lewis I, Kent SB. *Science*, 1994, 266: 776–779
20 Dong S, Zheng J-S, Li Y, Wang H, Chen G, Chen Y, Fang G, Guo J, He C, Hu H, Li X, Li Y, Li Z, Pan M, Tang S, Tian C, Wang P, Wu B, Wu C, Zhao J, Liu L. *Sci China: Chem*, 2024, 67: 1060–1096
21 Bandyopadhyay A, Cambray S, Gao J. *Chem Sci*, 2016, 7: 4589–4593
22 Faustino H, Silva MJSA, Veiros LF, Bernardes GJL, Gois PMP. *Chem Sci*, 2016, 7: 5052–5058
23 Li K, Wang W, Gao J. *Angew Chem Int Ed*, 2020, 59: 14246–14250
24 Ren H, Xiao F, Zhan K, Kim Y-P, Xie H, Xia Z, Rao J. *Angew Chem Int Ed*, 2009, 48: 9658–9662
25 Zheng M, Haeffner F, Gao J. *Chem Sci*, 2022, 13: 8349–8354
26 Sun H, Huang Y, Tsai Y-H. Genetically encoded 1,2-aminothiol for site-specific modification of a cellular membrane protein via TAMM condensation. In: Tsai Y-H, Elsässer SJ, Eds. Methods in Molecular Biology. Volume 2676: Genetically Incorporated Non-Canonical Amino Acids. Methods in Molecular Biology. New York: Humana, 2023. 191–199
27 Wang L, Dai G, Tsai Y-H, Yao S. Chapter five - 1,2-aminothiol-specific conjugation for dual-color fluorescent labeling via ultrafast TAMM conjugates. In: Urbach AR, Ed. Bioorthogonal Labeling. Volume 731: Methods in Enzymology. Academic Press, 2026. 99–118
28 Liu J, Duan S, Huang Y, Jian M-Y, Lee HS, Dai G, Wu C, Wu Y-L, Tsai Y-H. *J Am Chem Soc*, 2026, 148: 18020–18029
29 Zheng X, Li Z, Gao W, Meng X, Li X, Luk LYP, Zhao Y, Tsai Y-H, Wu C. *J Am Chem Soc*, 2020, 142: 5097–5103
30 Huang Y, Wu C, Lu A, Wang J, Liang J, Sun H, Yang L, Duan S, Berezin AA, Wu C, Zhang B, Wu Y-L, Tsai Y-H. *J Am Chem Soc*, 2025, 147: 1612–1623
31 Agouridas V, El Mahdi O, Diemer V, Cargoët M, Monbaliu J-CM, Melnyk O. *Chem Rev*, 2019, 119: 7328–7443
32 Dempsey DR, Jiang H, Kalin JH, Chen Z, Cole PA. *J Am Chem Soc*, 2018, 140: 9374–9378
33 Novak A, Kersaudy F, Berger S, Morisset-Lopez S, Lefoulon F, Pifferi C, Aucagne V. *Eur J Med Chem*, 2023, 260: 115747
34 Mino T, Sakamoto S, Hamachi I. *Biosci Biotechnol Biochem*, 2021, 85: 53–60

For Review Only

35 Zaramella S, Strömberg R, Yeheskiely E. *Eur J Org Chem*, 2002, 2002: 2633–2639
36 Diemer V, Firstova O, Agouridas V, Melnyk O. *Chem – Eur J*, 2022, 28: e202104229
37 Sakamoto K, Tsuda S, Mochizuki M, Nohara Y, Nishio H, Yoshiya T. *Chem – Eur J*, 2016, 22: 17940–17944
38 Chisholm TS, Clayton D, Dowman LJ, Sayers J, Payne RJ. *J Am Chem Soc*, 2018, 140: 9020–9024
39 Liao P, He C. *Chem Sci*, 2024, 15: 7965–7974
40 Flood DT, Hintzen JCJ, Bird MJ, Cistrone PA, Chen JS, Dawson PE. *Angew Chem Int Ed*, 2018, 57: 11634–11639
41 Fang G-M, Li Y-M, Shen F, Huang Y-C, Li J-B, Lin Y, Cui H-K, Liu L. *Angew Chem Int Ed*, 2011, 50: 7645–7649
42 Zheng J-S, Tang S, Qi Y-K, Wang Z-P, Liu L. *Nat Protoc*, 2013, 8: 2483–2495
43 Xu L, Kuan SL, Weil T. *Angew Chem Int Ed*, 2021, 60: 13757–13777
44 Johnson ECB, Kent SBH. *J Am Chem Soc*, 2006, 128: 6640–6646
45 Meng X, Fu K, Liu Y, Liu Y, Zhang JZ, Kang X, Wu C, Tsai Y-H. *J Am Chem Soc*, 2026, 148: 400–413
46 Liu H, Song L, Meng X, Li J, Fan S, Dong H, Wang X, Li M, Yu H, Tsai Y-H, Yin Y, Wu C. *J Am Chem Soc*, 2025, 147: 24870–24883
47 Liu J, Sun X, Zhuang J, Liu Z, Xu C, Wu D, Wu C. *Sci China Chem*, 2025, 68: 1434–1444
48 Cao X, Chen H, Wang X, Yang X, Wang T, Zhou L, Shi C, Liu S, Tian C, Sun D. *Sci China Chem*, 2025, 68: 4276–4284
49 Zhang Y-N, Wan X-C, Tang Y, Chen Y, Zheng F-H, Cui Z-H, Zhang H, Zhou Z, Fang G-M. *Chem Sci*, 2024, 15: 9649–9656
50 Wan X-C, Zhang Y-N, Zhang H, Chen Y, Cui Z-H, Zhu W-J, Fang G-M. *Org Lett*, 2024, 26: 2601–2605
51 Wu C. *Acc Chem Res*, 2025, 58: 1620–1631
52 Zheng Y, Zhai L, Zhao Y, Wu C. *J Am Chem Soc*, 2015, 137: 15094–15097
53 Zha M, Lin P, Yao H, Zhao Y, Wu C. *Chem Commun*, 2018, 54: 4029–4032
54 Wu C, Leroux J-C, Gauthier MA. *Nat Chem*, 2012, 4: 1044–1049
55 Cuadrado A, Rojo AI, Wells G, Hayes JD, Cousin SP, Rumsey WL, Attucks OC, Franklin S, Levonen A-L, Kensler TW, Dinkova-Kostova AT. *Nat Rev Drug Discovery*, 2019, 18: 295–317
56 Chen J, Feng D, Zhu R, Li H, Chen L. *Redox Biol*, 2025, 85: 103781
57 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. *Nature*, 2021, 596: 583–589
58 Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. *Nat Methods*, 2022, 19: 679–682
59 Lo S, Li X, Henzl MT, Beamer LJ, Hannink M. *EMBO J*, 2006, 25: 3605–3617

For review only