BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Design and synthesis of bioactive molecules

Edited by Fengzhi Zhang and David R. Spring

Generated on 03 November 2026, 06:33

Imprint

Beilstein Journal of Organic Chemistry
www.bjoc.org
ISSN 1860-5397
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Organic Chemistry* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

The copyright to this document as a whole, which is published in the *Beilstein Journal of Organic Chemistry*, is held by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften. The copyright to the individual articles in this document is held by the respective authors, subject to a Creative Commons Attribution license.

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Design and synthesis of bioactive molecules

Fengzhi Zhang*[1] and David R. Spring*[2]

Editorial

Open Access

Address:
[1]School of Pharmacy, Hangzhou Medical College, Hangzhou, 310014, P. R. China and [2]Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW, Cambridge, United Kingdom

Email:
Fengzhi Zhang* - fengzhi_zhang@hotmail.com;
David R. Spring* - spring@ch.cam.ac.uk

* Corresponding author

Keywords:
bioactive molecules; design; synthesis

Beilstein J. Org. Chem. **2026,** *22,* 1382–1384.
https://doi.org/10.3762/bjoc.22.112

Received: 16 September 2026
Accepted: 28 September 2026
Published: 07 October 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Guest Editor: F. Zhang; Associate Editor: D. R. Spring



© 2026 Zhang and Spring; licensee Beilstein-Institut.
License and terms: see end of document.

The design and synthesis of bioactive molecules remains a central pillar of modern chemical science, underpinning advances in drug discovery, chemical biology and molecular medicine. The ability to create molecules that modulate biological systems with precision continues to evolve rapidly, driven by innovations in synthetic methodology, molecular design and interdisciplinary integration. In this thematic issue, we present 8 Full Research Papers, 2 Letters, 2 Reviews and 1 Perspective, reflecting the breadth and emerging directions of this dynamic field.

A prominent theme is the growing importance of targeted protein modulation, particularly through bifunctional molecules. The development of an erdafitinib-derived degrader demonstrates selective degradation of FGFR2 and highlights the potential of targeted protein degradation to overcome resistance and off-target effects associated with conventional inhibitors [1]. Complementing this, the design of CAM4066-based PROTACs targeting CK2 illustrates how structure-guided linker installation can preserve binding interactions while enabling E3 ligase recruitment [2]. These advances are contextualised by a comprehensive Review on PROTACs, which emphasises their ability to achieve high selectivity within homologous protein families through event-driven pharmacology [3].

The issue also highlights advances in chemical biology tools and reactive probe development. A rhodium-catalysed strategy enables the connective synthesis of diverse reactive probes bearing S(VI) electrophilic warheads, expanding the toolkit for covalent probe design and target engagement [4]. In parallel, the development of silicon-rhodamine derivatives via a practical CO_2-mediated carboxylation strategy provides efficient access to key intermediates for targeted imaging probes, facilitating applications in live-cell super-resolution microscopy [5]. Together with a work on nanogold-labelled HDAC probes [6], these studies demonstrate how synthetic innovation enables increasingly sophisticated interrogation of biological systems.

At the interface of peptide science and small-molecule design, a Perspective explores how generative artificial intelligence can transform peptide binders into small-molecule leads, spanning workflows from structure prediction to scaffold generation and candidate prioritisation [7]. This is complemented by experimental advances in peptide chemistry, including the efficient solid-phase synthesis and structural characterisation of segetalins A–H, J and K, which highlights the continued importance of synthetic access to structurally complex bioactive peptides [8]. A Review on site-specific labelling strategies further underscores the importance of selective and scalable modification of

native peptides and proteins, with broad implications for chemical biology and therapeutic development [9].

The enduring influence of natural-product-inspired and bioactive scaffold synthesis is reflected in contributions exploring structure–activity relationships and biological evaluation. The synthesis of brassinosteroid analogues demonstrates how subtle structural modifications influence receptor binding, signalling pathways and biological responses [10]. Through the functional group splicing strategy, the design of a novel FXR/ASK1 dual-target modulator Z8 displays potential promise for the treatment of metabolic dysfunction-associated steatohepatitis (MASH) [11]. The synthesis and evaluation of glycidyl esters of phosphorus acids highlights the development of novel bioactive scaffolds with potential anticancer activity, supported by both biological assays and electrochemical studies [12]. In addition, a prodrug approach yielded the L-alanine conjugate AKBA-Ala, which showed anti-inflammatory potency comparable to that of AKBA [13].

Across these contributions, several unifying themes emerge. First, there is a clear shift toward functional molecular design, where compounds are engineered not only to bind targets but to induce specific biological outcomes. Second, synthetic efficiency and practicality remain essential, with new methodologies enabling more rapid, scalable and sustainable access to functional molecules. Third, the integration of computational and experimental approaches, particularly through artificial intelligence, is reshaping how bioactive molecules are conceived and optimised.

We are grateful to all authors for their outstanding contributions and to the reviewers for their insightful evaluations. We also thank the editorial team for their support in assembling this thematic issue. We hope that this collection will serve not only as a snapshot of current advances but also as a source of inspiration for future developments in the design and synthesis of bioactive molecules.

As chemistry continues to converge with biology and data science, the opportunities for innovation are immense. The work presented in this issue reflects this momentum and highlights the central role of chemical synthesis in addressing complex biological challenges.

Fengzhi Zhang and David R. Spring

Hangzhou and Cambridge, September 2026

ORCID® iDs

Fengzhi Zhang - https://orcid.org/0000-0001-9542-6634
David R. Spring - https://orcid.org/0000-0001-7355-2824

Data Availability Statement

Data sharing is not applicable as no new data was generated or analyzed in this study.

References

1. Jin, Y.; Wang, S.; Pan, S.; Huang, S.; Zhou, W.; Huang, X.; Zheng, L.; Chen, L. *Beilstein J. Org. Chem.* **2026,** *22,* 583–591. doi:10.3762/bjoc.22.44
2. Day-Riley, S.; Krajcovicova, S.; Sokhal, A. R.; Venne, J. L.; Brear, P.; Hyvönen, M.; Whitehurst, B. C.; Carroll, J. S.; Spring, D. R. *Beilstein J. Org. Chem.* **2026,** *22,* 611–619. doi:10.3762/bjoc.22.47
3. Yang, L.; Mao, X.; Zhang, J.; Shu, J.; Huang, W.; Dong, X.; Chen, Y.; Wu, M. *Beilstein J. Org. Chem.* **2026,** *22,* 628–661. doi:10.3762/bjoc.22.49
4. Rice, S.; Chesti, J.; Mosedale, W. R. T.; Wright, M. H.; Marsden, S. P.; Smith, T. K.; Nelson, A. *Beilstein J. Org. Chem.* **2025,** *21,* 1924–1931. doi:10.3762/bjoc.21.150
5. Hou, D.; Wu, S.; Xu, N.; Bao, P.; Jia, W.; Qiao, Q.; Xu, Z. *Beilstein J. Org. Chem.* **2026,** *22,* 915–924. doi:10.3762/bjoc.22.72
6. Pytel, W. A.; Schwabe, J. W. R.; Hodgkinson, J. T. *Beilstein J. Org. Chem.* **2026,** *22,* 480–485. doi:10.3762/bjoc.22.35
7. Mills, J.; Lau, Y. H. *Beilstein J. Org. Chem.* **2026,** *22,* 672–679. doi:10.3762/bjoc.22.51
8. Liu, L.; Lu, W.; Guo, Q.; Xu, Z. *Beilstein J. Org. Chem.* **2025,** *21,* 2612–2617. doi:10.3762/bjoc.21.202
9. Angelastro, A.; Bargh, J.; Guria, S.; Laserna, V.; Luk, L. *Beilstein J. Org. Chem.* **2026,** *22,* 857–881. doi:10.3762/bjoc.22.67
10. Núñez, M.; Escobar, C.; Párraga, M.; Soto, M.; Espinoza-Catalán, L.; Díaz, K.; Olea, A. F. *Beilstein J. Org. Chem.* **2026,** *22,* 753–762. doi:10.3762/bjoc.22.57
11. Zhang, X.; Wang, J.; Zhao, Z.; Wang, C.; Ye, Z.; Ma, W.-Y.; Xu, J.-X.; Zhang, F. *Beilstein J. Org. Chem.* **2026,** *22,* 771–781. doi:10.3762/bjoc.22.59
12. Zagidullin, A. A.; Bulatov, E. R.; Khrizanforov, M. N.; Davletshin, D. R.; Gilyazova, E. M.; Strelkov, I. A.; Miluykov, V. A. *Beilstein J. Org. Chem.* **2025,** *21,* 1909–1916. doi:10.3762/bjoc.21.148
13. Peschel, J. M.; Jordan, P. M.; Friebe, B.; Görls, H.; Köhler, P.; Schubert, U. S.; Werz, O.; Gottschaldt, M. *Beilstein J. Org. Chem.* **2026,** *22,* 1229–1236. doi:10.3762/bjoc.22.99

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.112

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Synthesis, biological and electrochemical evaluation of glycidyl esters of phosphorus acids as potential anticancer drugs

Almaz A. Zagidullin[*1,2], Emil R. Bulatov[2], Mikhail N. Khrizanforov[3], Damir R. Davletshin[2], Elvina M. Gilyazova[2], Ivan A. Strelkov[1] and Vasily A. Miluykov[1]

Letter

Open Access

Address:
[1]Arbuzov Institute of Organic and Physical Chemistry, FRC Kazan Scientific Center of Russian Academy of Sciences, Arbuzov Street 8, Kazan, 420088, Russia, [2]Institute of Fundamental Medicine and Biology, Kazan Federal University, Kremlyovskaya Street 18, Kazan, 420008, Russia and [3]Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninsky prospekt 47, Moscow, 119991, Russia

Email:
Almaz A. Zagidullin[*] - almaz_zagidullin@mail.ru

* Corresponding author

Keywords:
alkylating agent; glycidyl ester; electrochemical evaluation; phosphorus-containing drug

Beilstein J. Org. Chem. **2025,** *21,* 1909–1916.
https://doi.org/10.3762/bjoc.21.148

Received: 15 April 2025
Accepted: 28 August 2025
Published: 15 September 2025

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring



© 2025 Zagidullin et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Organophosphorus compounds are important in synthetic organic chemistry and pharmaceutical applications due to their diverse biological activities. In this study, we synthesized three novel glycidyl esters of phosphorus acids **1**–**3** via the condensation of chlorophosphine oxides or phosphorus oxychloride with glycidol in the presence of a base, obtaining products with high purity and moderate to excellent yields. Their cytotoxic potential was evaluated using the MTT assay on human fibroblasts (HSF), prostate cancer (PC-3), and breast cancer (MCF7) cell lines, revealing moderate preferential cytotoxicity toward cancer cells, particularly in the case of MCF7. Additionally, linear sweep voltammetry (LSV) studies on human serum albumin (HSA) were conducted to investigate their alkylating properties. The electrochemical results suggest that these compounds effectively modify albumin, highlighting their potential as reactive anticancer agents. These findings provide important insights into the synthesis, cytotoxic activity, and biochemical reactivity of glycidyl esters of phosphorus acids, underscoring their potential as lead structures for further development in anticancer drug discovery and pharmaceutical research.

Introduction

Phosphorus-containing drugs represent a crucial category of therapeutic agents, extensively utilized in clinical practice due to their diverse pharmacological properties and applications [1-4]. These compounds have garnered considerable attention from both pharmaceutical companies and researchers, reflecting their significance in drug development and therapeutic innovation [5-7]. The structural diversity of phosphorus-containing molecules, which includes phosphotriesters, phosphonates, phosphinates, phosphine oxides, and bisphosphonates, allows for tailored modifications that enhance selectivity, bioavailability, and reduce potential side effects [8-13]. This versatility makes them valuable not only as drugs but also as intermediates in synthetic organic chemistry, facilitating access to a wide array of molecular targets [14-16]. The importance of phosphorus-containing drugs extends beyond their therapeutic applications; they also play a pivotal role in addressing specific medical conditions such as chronic kidney disease (CKD) [17,18].

The synthesis of organophosphorus compounds is a dynamic field of research, with numerous synthetic methodologies being explored to create novel phosphorus derivatives [19-21]. Recent studies have highlighted the increasing relevance of three-membered strained cycles containing phosphorus in various domains such as agrochemicals, synthetic chemistry, and medicine. This surge in interest has led to the development of innovative synthetic routes aimed at producing new members of these compounds [22]. For instance, fosfomycin [23] stands out as a broad-spectrum antibiotic currently employed in clinical settings, while thiotepa has been approved for treating several cancers, including gastrointestinal tumors and bladder cancer (Figure 1). Additionally, phosphoric triamides alkylating agents featuring aziridine rings are recognized for their role as nitrogen mustards in cancer therapy [24].

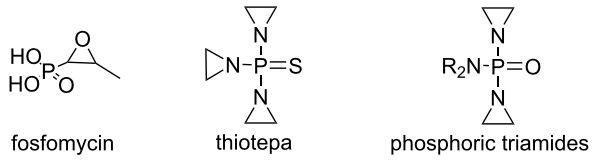


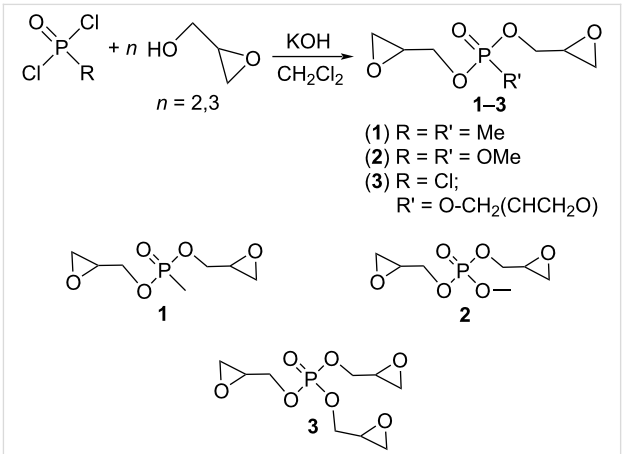
Figure 1: Structures of some pharmacological important phosphorus-containing molecules with oxirane or aziridine fragments.

Although there are numerous examples in the chemical literature regarding the biological activity (including anticancer properties) of phosphoric esters, reports on biological studies of systems based on the P=O fragment and oxirane skeletons are less common. Nevertheless, systems containing both of the mentioned structural motifs are rarely encountered in the literature. In this paper, we report the synthesis, biological activity, and electrochemical evaluation of glycidyl esters of phosphorus acids.

Results and Discussion

Synthesis of glycidyl esters of phosphorus acids 1–3

Glycidyl esters of phosphorus acids **1**–**3** were obtained by condensation of chlorophosphine oxides (methylphosphonic dichloride $MeP(O)Cl_2$; methyl dichlorophosphate $(MeO)P(O)Cl_2$) and phosphorus oxychloride $P(O)Cl_3$ with racemic glycidol in CH_2Cl_2 in the presence of KOH as basic agent (Scheme 1). Further filtration and final distillation at low pressure leads to the products as thick liquids with good yields (44–67%) and purity.



Scheme 1: Synthesis of glycidyl esters of phosphorus acids **1**–**3**.

The structures of glycidyl esters of phosphorus acids **1**–**3** were confirmed by ^{31}P, ^{1}H NMR, IR spectroscopy, and elemental analysis (see Experimental part for additional information). The $^{31}P\{^{1}H\}$ NMR spectrum of diglycidyl methylphosphonate (**1**) shows a singlet at +32 ppm; for diglycidyl methylphosphate (**2**) and triglycidyl phosphate (**3**) also a singlet in the region 0–1 ppm is observed, despite the presence of a chiral carbon atom in the oxirane fragment. In the ^{1}H NMR spectra of esters **1**–**3** the characteristic signals of the oxirane fragment at 2.41–3.24 ppm and the $POCH_2$- fragment at 3.66–4.36 ppm can be observed. The NMR data for the glycidyl esters of phosphorus acids **1**–**3** are comparable to those of related compounds.

Biological activity of glycidyl esters of phosphorus acids

To evaluate the biological activity of diastereomeric mixtures of glycidyl esters of phosphorus acids **1**–**3**, their cytotoxic effects

were assessed using the MTT assay on two tumor cell lines (PC-3 and MCF7) and one non-cancerous line (HSF). The assay measures the concentration of each compound required to inhibit cellular metabolic activity by 50% (IC_{50}). All experiments were performed in biological triplicates, and standard deviations were calculated to assess statistical reliability. The results are summarized in Figure 2 and Table 1.

Table 1: IC_{50} values (μM ± SD) for glycidyl esters of phosphorus acids **1**–**3** in human skin fibroblasts (HSF), prostate cancer cells (PC-3), and breast cancer cells (MCF7).[a]

	IC50 (μM)		
Cell line	**1**	**2**	**3**
HSF	394 ± 28	398 ± 33	254 ± 19
PC-3	355 ± 25	300 ± 21	257 ± 20
MCF7	216 ± 16	128 ± 10	182 ± 14

[a]Values represent the mean ± standard deviation (n = 3) of biological replicates, determined by MTT assay after 48 hours of treatment. IC_{50} indicates the concentration required to reduce cell viability by 50%; lower values correspond to higher cytotoxic potency.

Diglycidyl methylphosphonate (**1**) reduced cell viability by 50% at concentrations of 394 ± 28 μM, 355 ± 25 μM, and 216 ± 16 μM for HSF, PC-3, and MCF7 cell lines, respectively. Similarly, diglycidyl methylphosphate (**2**) achieved 50% inhibition at concentrations of 398 ± 33 μM, 300 ± 21 μM, and 128 ± 10 μM. Triglycidyl phosphate (**3**) exhibited IC_{50} values of 254 ± 19 μM for HSF, 257 ± 20 μM for PC-3, and 182 ± 14 μM for MCF7 cells.

Among the tested compounds, triglycidyl phosphate (**3**) demonstrated the highest overall cytotoxicity against HSF and PC-3 cell lines, while diglycidyl methylphosphate (**2**) showed the greatest potency toward MCF7 breast cancer cells. Although the IC_{50} values for compounds **1** and **2** were somewhat higher in normal fibroblasts (HSF) compared to cancer cells, the differences were moderate (less than twofold). These results suggest a modest preferential cytotoxicity toward cancer cells, particularly in the case of compound **2** against MCF7, though further studies are needed to establish meaningful selectivity.

Electrochemical studies

Alkylating agents are widely recognized for their ability to form covalent bonds with biological macromolecules (proteins, DNA). The literature discusses the interaction of small molecules with proteins, highlighting how linear sweep voltammetry (LSV) can be used to understand these interactions. The method provides insight into protein structures and functions using electrochemical methods that can also be applied to studies involving alkylating agents [25,26]. In this study, LSV was employed to investigate the interactions between human serum albumin (HSA) and the three prospective alkylating agents **1**–**3**. The motivation behind these experiments was to explore whether these compounds, which individually exhibit no appreciable redox activity in the potential window applied, can chemically modify (alkylate) serum albumin and thus suppress its characteristic oxidation peaks.

Human serum albumin was chosen as a model protein because of its well-characterized structure and the presence of reactive sites that are known to be susceptible to alkylation. In standard

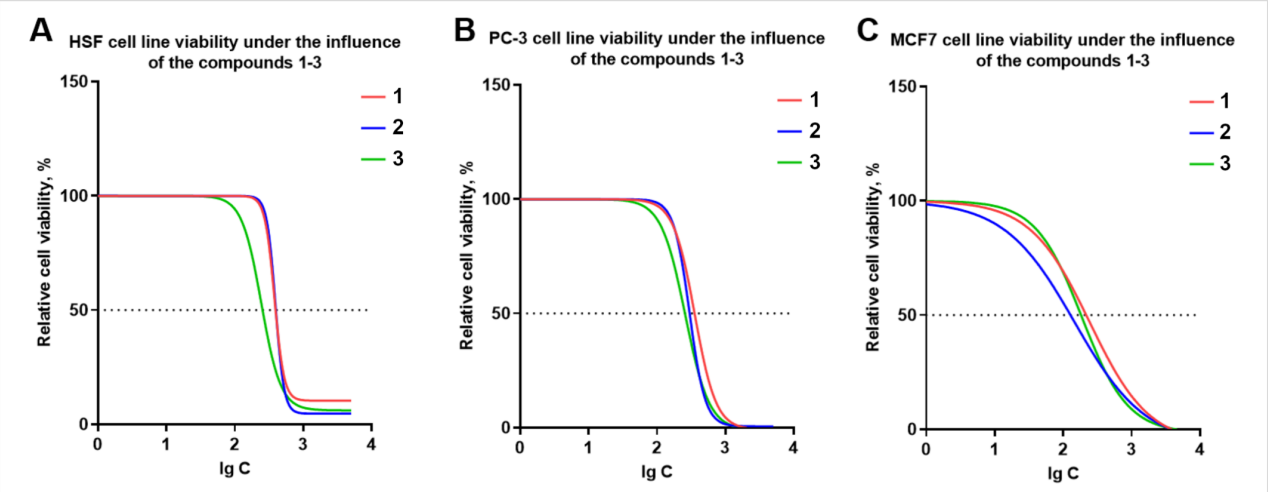


Figure 2: Dose–response curves for the cytotoxic effects of glycidyl esters of phosphorus acids **1**–**3** on human cell lines. (A) Human skin fibroblasts (HSF), (B) prostate cancer cells (PC-3), and (C) breast cancer cells (MCF7) were treated with diglycidyl methylphosphonate (**1**, red), diglycidyl methylphosphate (**2**, blue), and triglycidyl phosphate (**3**, green) for 48 hours. Cell viability was determined using the MTT assay. Data represent mean ± standard deviation (n = 3). The y-axis shows relative cell viability (%) compared to untreated control, and the x-axis indicates the logarithmic concentration of each compound. The dashed line marks the 50% viability threshold (IC_{50}).

aqueous media, the electrochemical oxidation of HSA can be observed via LSV as a broad wave, which is often attributed to the oxidation of amide and other amino acid side-chain fragments. By tracking changes in this oxidation signal upon addition of an alkylating agent, we can infer whether the agent has effectively reacted with (and thus structurally altered) the protein.

As illustrated by the black trace in the LSV plot, pure HSA in aqueous medium shows a characteristic oxidation wave that begins to rise around +0.5 V and significantly increases up to +1.2 V vs Ag/AgCl (Figure 3). This wave is attributed to oxidation processes at peptide bonds or specific side chains (such as cysteine, methionine, tyrosine, serine, tryptophan residues), as well as the overall structure of the protein. The peak intensity and shape can vary depending on pH, ionic strength, and protein conformation. However, under our conditions, the HSA oxidation was consistent, well-defined, and served as a clear baseline reference.

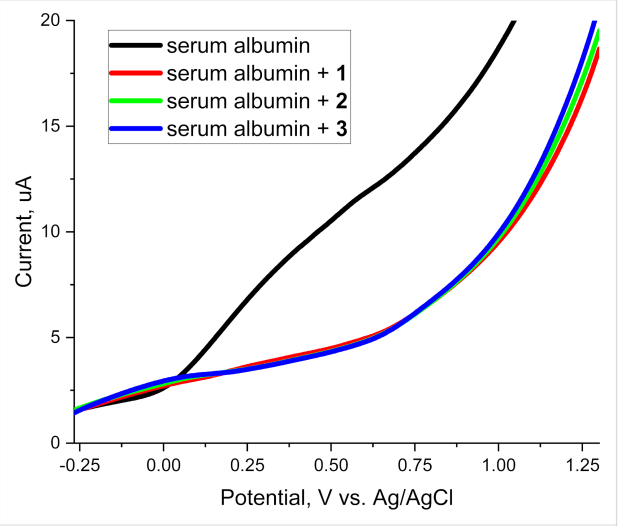


Figure 3: Linear sweep voltammograms of 1×10^{-4} M HSA (black) and HSA mixed with each of the three alkylating agents; **1** (red), **2** (green), and **3** (blue). Conditions: supporting electrolyte: 0.1 M Et_4NBF_4, working electrode: glassy carbon, scan rate: 0.1 V/s, pH 6.7.

Subsequent to acquiring the control LSV of HSA, 10 µL of each alkylating agent was introduced separately into the albumin solution. As soon as the alkylating agent was added, the characteristic oxidation wave of the albumin nearly vanished or became drastically reduced. Control experiments confirmed that compounds **1**–**3** themselves exhibit no discernible redox activity in this potential range when tested in the absence of HSA. Consequently, any changes in the recorded voltammogram could be attributed to the interaction (alkylation) of albumin rather than to new electrochemical processes arising directly from the compounds.

When these agents alkylate the HSA amino acid residues (particularly reactive sites like lysine, cysteine, serine NH_2, SH, OH-side chains, and possibly other nucleophilic groups), the resulting covalent modification can disrupt the electroactive centers responsible for the protein's oxidation peaks (Figure 4). In many alkylation scenarios, crosslinking or other structural rearrangements can render previously oxidizable moieties inaccessible or shift the protein's conformational state. This suppresses or altogether eliminates the characteristic oxidation wave of HSA.

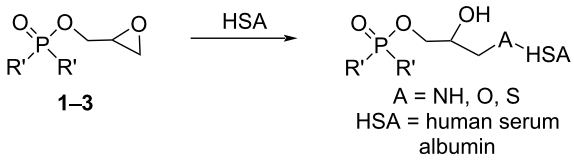


Figure 4: Chemical structure of alkylating fragment of **1**–**3** and the associated chemical pathway of its covalent attachment to HSA.

Based on established literature, the significant suppression or disappearance of the HSA oxidation peak upon addition of glycidyl esters **1**–**3** can be interpreted as evidence of covalent modification (alkylating) of nucleophilic sites on HSA, rather than non-specific binding or merely non-reactive association [27-29]. The observed disappearance of the albumin oxidation peak strongly suggests that all three investigated compounds can modify proteins under physiological conditions. Moreover, the fact that each agent was capable of this disruption aligns well with prior tests on the studied cell lines (PC-3, MCF-7, and HSF), where differences in IC_{50} values reflected the degree of alkylating potency and the selective toxicity toward cancer cells.

Conclusion

In this study, we synthesized and comprehensively characterized a series of glycidyl esters of phosphorus acids **1**–**3**, evaluating their structural features, cytotoxic potential, and electrochemical behavior. The compounds were efficiently obtained via the condensation of chlorophosphine oxides and phosphorus oxychloride with glycidol, affording high-purity products in moderate to good yields. Cytotoxicity studies revealed that all three compounds possess antiproliferative activity against cancer cell lines (PC-3 and MCF7), with diglycidyl methylphosphate (**2**) demonstrating the highest potency toward MCF7 cells. While all compounds exhibited some level of toxicity toward non-cancerous HSF cells, their IC_{50} values in fibroblasts were generally higher than those observed in tumor cells, indicating a trend toward moderate preferential cytotoxicity. These findings suggest that the glycidyl phosphorus esters,

particularly compound **2**, may serve as promising lead structures for further exploration as anticancer agents.

A key innovation in this work was the application of linear sweep voltammetry (LSV) to investigate the alkylating properties of the synthesized compounds. Unlike traditional biochemical assays, this electrochemical approach enabled real-time monitoring of protein modifications. The significant suppression of human serum albumin (HSA) oxidation peaks following exposure to compounds **1**–**3** strongly indicates their ability to covalently modify nucleophilic sites in proteins. This finding underscores the potential of LSV as a rapid and effective tool for assessing alkylating reactivity, with implications for future drug development.

Overall, this study offers meaningful insights into the synthesis, cytotoxic behavior, and biochemical reactivity of glycidyl esters of phosphorus acids. The results support their potential as reactive anticancer candidates and lay a foundation for future structure–activity relationship studies and further development in medicinal chemistry.

Experimental

General. All reactions and manipulations were carried out under dry pure N_2 in standard Schlenk apparatus. All solvents were distilled from sodium/benzophenone or phosphorus pentoxide and stored under nitrogen before use. The NMR spectra were recorded on a Bruker MSL-400 (^{1}H 400 MHz, ^{31}P 161.7 MHz, ^{13}C 100.6 MHz). ^{1}H and ^{13}C NMR data are reported with reference to solvent resonances, and ^{31}P NMR spectra were reported with respect to external 85% H_3PO_4 (0 ppm). All experiments were carried out using standard Bruker pulse programs. Infrared (IR) spectra were recorded on a Bruker Vector-22 spectrometer. The elemental analyses were carried out at the microanalysis laboratory of the Arbuzov Institute of Organic and Physical Chemistry, Russian Academy of Sciences.

Cell cultivation. Adherent cell lines HSF (human skin fibroblasts), MCF7 (breast adenocarcinoma), and PC-3 (prostate cancer) were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 5% fetal bovine serum (FBS), 1 mM L-glutamine, and antibiotics (penicillin 5,000 U/mL and streptomycin 5,000 µg/mL). Cells were incubated at 37 °C in a humidified atmosphere with 5% CO_2. For cytotoxicity assays, cells were seeded into 96-well flat-bottom plates at a density of 5×10^3 cells per well and allowed to adhere for 24 hours under standard culture conditions.

Preparation of compound solutions. Stock solutions of diglycidyl methylphosphonate (**1**), diglycidyl methylphosphate (**2**), and triglycidyl phosphate (**3**) were prepared by dissolving the compounds in the culture medium to a final concentration of 25 mM. These stock solutions were stored and used for subsequent treatments.

Cell treatment. 24 hours after seeding, cells were treated with the test compounds at final concentrations of 25 µM, 50 µM, 100 µM, 250 µM, 500 µM, 750 µM, 1,000 µM, 2,500 µM, and 5,000 µM. Each concentration was tested in triplicate. Control wells received an equivalent volume of culture medium without compounds and served as untreated negative controls.

Cytotoxicity analysis of compounds. After 48 hours of treatment, MTT reagent was added to each well at a final concentration of 0.5 mg/mL. Plates were incubated for 3 hours at 37 °C in a CO_2 incubator to allow for formazan crystal formation. Subsequently, 150 µL of dimethyl sulfoxide (DMSO) was added to each well to solubilize the formazan. Plates were shaken for 15 minutes, and absorbance was measured at 590 nm using an Infinite M200 microplate reader (Tecan, Switzerland). Cell viability was calculated relative to the untreated control (set at 100%). Data were processed and analyzed using GraphPad Prism 10 software.

Electrochemistry. Linear sweep voltammograms were recorded using a BASi Epsilon Eclipse potentiostat (USA). The device includes a measuring unit, a DellOptiplex 320 personal computer with Epsilon-EC-USB-V200 software. As supporting electrolyte 0.1 M Et_4NBF_4 was used. A glassy carbon electrode modified with carbon paste (surface area 1 mm^2) served as the working electrode. Ag/AgCl (0.01 M KCl) was used as a reference electrode. A platinum wire was used as an auxiliary electrode. The scanning rate was 100 mV s^{-1}. Measurements were carried out in a thermostatted electrochemical cell (volume 5 mL) in an inert gas atmosphere (N_2). Between measurements or before recording the voltammetric wave, the aqueous solution was actively stirred with a magnetic stirrer in an atmosphere of constant inert gas flow.

Starting materials. Methylphosphonic dichloride $MeP(O)Cl_2$ and methyl dichlorophosphate $(MeO)P(O)Cl_2$) [30] were prepared according to literature procedures. Phosphorus oxychloride $P(O)Cl_3$ and glycidol were purchased from suppliers and used without additional purification.

Synthesis of diglycidyl methylphosphonate (1). A 500 mL flask with 200 mL of dichloromethane, equipped with a mechanical stirrer, was cooled to −30 °C. Then, 2 equiv of glycidol (21.4 g, 0.289 mol) and 4.2 equiv of potassium hydroxide (34.0 g, 0.607 mol) were added to the flask. Methylphosphonic dichloride $MeP(O)Cl_2$ (1 equiv, 19.2 g, 0.144 mol) was added

dropwise to the mixture with constant stirring in 1 hour. The reaction mixture was additionally stirred at −25 to −30 °C for 2 h and precipitating for 12 h, while the temperature did not rise above 0 °C. After removal of the precipitate by filtration at 25 °C (filter consisted of layers of Celite, sodium sulfate and activated carbon), the filtrate was evaporated under reduced pressure to remove dichloromethane and excess glycidol. After vacuum distillation ($p = 1{\cdot}10^{-3}$ mbar, bp = 130–136 °C) the product diglycidyl methylphosphonate (**1**) was obtained as a thick liquid in 58% yield (17.5 g). ^{1}H NMR ($CDCl_3$, δ, ppm, *J*, Hz) 1.31 (d, $^2J_{PH}$ = 17.8, 3H, Me), 2.41–2.43 (m, 2H, CH_2-oxirane), 2.58–2.61 (m, 2H, CH_2-oxirane), 2.96–2.98 (m, 2H, CH-oxirane), 3.66–3.69 (m, 2H, OCH_2), 4.07–4.09 (m, 2H, OCH_2); $^{31}P\{^1H\}$ NMR ($CDCl_3$, δ, ppm, *J*, Hz) 32.04 (s); IR (liquid, cm^{-1}): 762 (m, oxirane), 858 (m,oxirane), 927 (m, oxirane), 1019 (m), 1139 (w), 1166 (w, P=O), 1240 (m, oxirane), 1316 (m, R-P(O)OR), 1349 (s, P=O), 1425 (w), 1455 (m), 1647 (m), 2932 (m), 3004 (m), 3066 (w); Anal. calcd for $C_7H_{13}PO_5$: C, 40.39; H, 6.30; O, 38.43; P, 14.88; found: C, 40.24; H, 6.52; P, 14.79.

Synthesis of diglycidyl methylphosphate (2). A 500 mL flask with 200 mL of dichloromethane, equipped with a mechanical stirrer, was cooled to −30 °C. Then, 2 equiv of glycidol (20.0 g, 0.27 mol) and 4.2 equiv of potassium hydroxide (31.8 g, 0.567 mol) were added to the flask. Methyl dichlorophosphate $(MeO)P(O)Cl_2$ (1 equiv, 20.1 g, 0.135 mol) was added dropwise to the mixture with constant stirring in 1 hour. The reaction mixture was additionally stirred at −25 to −30 °C for 2 h and precipitating for 12 h, while the temperature did not rise above 0 °C. After removal of the precipitate by filtration at 25 °C (filter consisted of layers of Celite, sodium sulfate and activated carbon), the filtrate was evaporated under reduced pressure to remove dichloromethane and excess glycidol. After two vacuum distillations ($p = 1{\cdot}10^{-3}$ mbar, bp 113–116 °C) the product diglycidyl methylphosphate (**2**) was obtained as a thick liquid in 44% yield (13.5 g). ^{1}H NMR ($CDCl_3$, δ, ppm, *J*, Hz) 2.49–2.55 (m, 2H, CH_2-oxirane), 2.67–2.72 (m, 2H, CH_2-oxirane), 3.06–3.14 (m, 2H, CH-oxirane), 3.65 (d, $^3J_{PH}$ = 11.4, 3H, Me), 3.75–3.84 (m, 2H, OCH_2), 4.13–4.23 (m, 2H, OCH_2); $^{31}P\{^1H\}$ NMR ($CDCl_3$, δ, ppm, *J*, Hz) −0.1 (s); IR (liquid, cm^{-1}): 598 (w), 763 (m, oxirane), 865 (m, oxirane), 921 (m, oxirane), 1021 (m, $P(O)(OR)_2$), 1140 (w, $P(O)(OR)_2$), 1168 (s, $P(O)(OR)_2$), 1185 (w), 1261 (m), 1350 (m, P=O), 1430 (w), 1455 (m), 1644 (w), 2858 (w), 2960 (w), 3008 (w), 3066 (w); Anal. calcd for $C_7H_{13}PO_6$: C, 37.51; H, 5.85; O, 42.83; P, 13.82; found: C, 37.50; H, 6.03; P, 13.97.

Synthesis of triglycidyl phosphate (3). Synthesis of **3** was carried out in a manner similar to [31], but without sodium sulfate as a drying agent. A 500 mL flask with 200 mL of dichloromethane, equipped with a mechanical stirrer, was cooled to −30 °C. Then, 3 equiv of glycidol (22.6 g, 0.305 mol) and 4.5 equiv of potassium hydroxide (25.6 g, 0.457 mol) were added to the flask. Phosphorus oxychloride $P(O)Cl_3$ (1 equiv, 15.6 g, 0.102 mol) was added dropwise to the mixture with constant stirring in 1 hour. The reaction mixture was additionally stirred at −25 to −30 °C for 3 h and precipitating for 12 h, while the temperature did not rise above 10 °C. After removal of the precipitate by filtration at 25 °C (filter consisted of layers of Celite, sodium sulfate and activated carbon), the filtrate was evaporated under reduced pressure to remove dichloromethane and excess glycidol. The product triglycidyl phosphate (**3**) was obtained as a yellowish thick liquid in 67% yield (18.1 g). ^{1}H NMR ($CDCl_3$, δ, ppm, *J*, Hz) 2.56–2.65 (m, 3H, CH_2-oxirane), 2.74–2.85 (m, 3H, CH_2-oxirane), 3.14–3.24 (m, 3H, CH-oxirane), 3.84–3.95 (m, 3H, OCH_2), 4.22–4.36 (m, 3H, OCH_2); $^{31}P\{^1H\}$ NMR ($CDCl_3$, δ, ppm, *J*, Hz) −1.2 (s); IR (liquid, cm^{-1}): 599 (w), 700 (w), 763 (m, oxirane), 797 (m, oxirane), 869 (m, oxirane), 918 (m, oxirane), 1024 (m), 1139 (w, P=O), 1166 (s, P=O), 1259 (m), 1349 (m, P=O), 1429 (w), 1454 (m), 1483 (w), 1520 (w), 1634–1644 (m), 2614 (w), 2899 (w), 2953 (m), 3006 (m), 3065 (m); Anal. calcd for $C_9H_{15}PO_7$: C, 40.61; H, 5.68; O, 42.07; P, 11.64; found: C, 40.85; H, 5.82; P, 11.97.

Supporting Information

Supporting Information File 1

^{1}H, ^{31}P NMR and IR spectra of compounds **1**–**3**.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-21-148-S1.pdf]

Acknowledgements

The measurements have been carried out using the equipment of the Assigned Spectral-Analytical Center of Shared Facilities for Study of Structure, Composition and Properties of Substances and Materials of the FRC Kazan Scientific Center of RAS. We thank Prof. Shamil K. Latypov for discussing NMR spectra.

Funding

This research activity was funded by the government assignment for the FRC Kazan Scientific Center of RAS.

Author Contributions

Almaz A. Zagidullin: investigation. Emil R. Bulatov: investigation. Mikhail N. Khrizanforov: investigation. Damir R. Davletshin: investigation. Elvina M. Gilyazova: investigation. Ivan A.

Strelkov: investigation. Vasily A. Miluykov: writing – original draft.

ORCID® iDs

Almaz A. Zagidullin - https://orcid.org/0000-0003-3125-7506
Vasily A. Miluykov - https://orcid.org/0000-0002-8069-457X

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Dormán, G.; Szalai, Z.; Keglevich, G. *ChemMedChem* **2024,** *19,* e202400370. doi:10.1002/cmdc.202400370
2. Finkbeiner, P.; Hehn, J. P.; Gnamm, C. *J. Med. Chem.* **2020,** *63,* 7081–7107. doi:10.1021/acs.jmedchem.0c00407
3. Rodriguez, J. B.; Gallo-Rodriguez, C. *ChemMedChem* **2019,** *14,* 190–216. doi:10.1002/cmdc.201800693
4. Kumar, A.; Mukhopadhyay, J.; Bhagat, S. *ChemistrySelect* **2024,** *9,* e202404258. doi:10.1002/slct.202404258
5. Yu, H.; Yang, H.; Shi, E.; Tang, W. *Med. Drug Discovery* **2020,** *8,* 100063. doi:10.1016/j.medidd.2020.100063
6. Zielonka, J.; Joseph, J.; Sikora, A.; Hardy, M.; Ouari, O.; Vasquez-Vivar, J.; Cheng, G.; Lopez, M.; Kalyanaraman, B. *Chem. Rev.* **2017,** *117,* 10043–10120. doi:10.1021/acs.chemrev.7b00042
7. Pradere, U.; Garnier-Amblard, E. C.; Coats, S. J.; Amblard, F.; Schinazi, R. F. *Chem. Rev.* **2014,** *114,* 9154–9218. doi:10.1021/cr5002035
8. Farrell, R. E.; Steele, H.; Middleton, R. J.; Skropeta, D.; Liu, G.-J. *RSC Med. Chem.* **2024,** *15,* 1973–1981. doi:10.1039/d4md00115j
9. Wu, J.; Xu, R.; Shao, M.; Zhao, L.; Xu, W.; Guo, Y. *ACS Appl. Nano Mater.* **2024,** *7,* 11022–11036. doi:10.1021/acsanm.4c00015
10. Gibadullina, E.; Neganova, M.; Aleksandrova, Y.; Nguyen, H. B. T.; Voloshina, A.; Khrizanforov, M.; Nguyen, T. T.; Vinyukova, E.; Volcho, K.; Tsypyshev, D.; Lyubina, A.; Amerhanova, S.; Strelnik, A.; Voronina, J.; Islamov, D.; Zhapparbergenov, R.; Appazov, N.; Chabuka, B.; Christopher, K.; Burilov, A.; Salakhutdinov, N.; Sinyashin, O.; Alabugin, I. *Int. J. Mol. Sci.* **2023,** *24,* 12637. doi:10.3390/ijms241612637
11. Tsepaeva, O. V.; Nemtarev, A. V.; Pashirova, T. N.; Khokhlachev, M. V.; Lyubina, A. P.; Amerkhanova, S. K.; Voloshina, A. D.; Mironov, V. F. *RSC Med. Chem.* **2023,** *14,* 454–469. doi:10.1039/d2md00363e
12. Ermolaev, V. V.; Arkhipova, D. M.; Miluykov, V. A.; Lyubina, A. P.; Amerhanova, S. K.; Kulik, N. V.; Voloshina, A. D.; Ananikov, V. P. *Int. J. Mol. Sci.* **2021,** *23,* 86. doi:10.3390/ijms23010086
13. Terekhova, N. V.; Lyubina, A. P.; Voloshina, A. D.; Sapunova, A. S.; Khayarov, K. R.; Islamov, D. R.; Usachev, K. S.; Evtugyn, V. G.; Tatarinov, D. A.; Mironov, V. F. *Bioorg. Chem.* **2022,** *127,* 106030. doi:10.1016/j.bioorg.2022.106030
14. Mironov, V. F.; Dimukhametov, M. N.; Nemtarev, A. V.; Pashirova, T. N.; Tsepaeva, O. V.; Voloshina, A. D.; Vyshtakalyuk, A. B.; Litvinov, I. A.; Lyubina, A. P.; Sapunova, A. S.; Abramova, D. F.; Zobov, V. V. *Nanomaterials* **2023,** *13,* 2840. doi:10.3390/nano13212840
15. Pashirova, T. N.; Nemtarev, A. V.; Souto, E. B.; Mironov, V. F. *Russ. Chem. Rev.* **2023,** *92,* RCR5095. doi:10.59761/rcr5095
16. Terekhova, N. V.; Khailova, L. S.; Rokitskaya, T. I.; Nazarov, P. A.; Islamov, D. R.; Usachev, K. S.; Tatarinov, D. A.; Mironov, V. F.; Kotova, E. A.; Antonenko, Y. N. *ACS Omega* **2021,** *6,* 20676–20685. doi:10.1021/acsomega.1c02909
17. Bacchetta, J.; Bernardor, J.; Garnier, C.; Naud, C.; Ranchin, B. *Calcif. Tissue Int.* **2021,** *108,* 116–127. doi:10.1007/s00223-020-00665-8
18. Lobo, L.; Kishore, N.; Sharma, M. *J. Drug Delivery Ther.* **2024,** *14,* 176–181. doi:10.22270/jddt.v14i1.6336
19. Zagidullin, A. A.; Sakhapov, I. F.; Miluykov, V. A.; Yakhvarov, D. G. *Molecules* **2021,** *26,* 5283. doi:10.3390/molecules26175283
20. Oshchepkova, E. S.; Zagidullin, A. A.; Miluykov, V. A.; Sinyashin, O. G. *Phosphorus, Sulfur Silicon Relat. Elem.* **2016,** *191,* 1530–1532. doi:10.1080/10426507.2016.1212350
21. Zagidullin, A.; Ganushevich, Y.; Miluykov, V.; Sinyashin, O.; Hey-Hawkins, E. *Phosphorus, Sulfur Silicon Relat. Elem.* **2013,** *188,* 238–242. doi:10.1080/10426507.2012.744017
22. Mayorquín-Torres, M. C.; Simoens, A.; Bonneure, E.; Stevens, C. V. *Chem. Rev.* **2024,** *124,* 7907–7975. doi:10.1021/acs.chemrev.4c00090
23. Falagas, M. E.; Vouloumanou, E. K.; Samonis, G.; Vardakas, K. Z. *Clin. Microbiol. Rev.* **2016,** *29,* 321–347. doi:10.1128/cmr.00068-15
24. van Maanen, M. J.; Smeets, C. J. M.; Beijnen, J. H. *Cancer Treat. Rev.* **2000,** *26,* 257–268. doi:10.1053/ctrv.2000.0170
25. Parker, V. D.; Roddick, A.; Seefeldt, L. C.; Wang, H.; Zheng, G. *Anal. Biochem.* **1997,** *249,* 212–218. doi:10.1006/abio.1997.2176
26. Sun, W.; Zhao, N.; Niu, X.; Wang, Y.; Jiao, K. *J. Chem. Sci.* **2009,** *121,* 217–223. doi:10.1007/s12039-009-0025-8
27. Zatloukalova, M.; Mojovic, M.; Pavicevic, A.; Kabelac, M.; Freeman, B. A.; Pekarova, M.; Vacek, J. *Redox Biol.* **2019,** *24,* 101213. doi:10.1016/j.redox.2019.101213
28. Attar, A. M.; Richardson, M. B.; Speciale, G.; Majumdar, S.; Dyer, R. P.; Sanders, E. C.; Penner, R. M.; Weiss, G. A. *ACS Appl. Mater. Interfaces* **2019,** *11,* 4757–4765. doi:10.1021/acsami.8b16071
29. Fu, L.; Liu, X.-f.; Zhou, Q.-x.; Zhang, J.-x.; Dong, J.-y.; Wang, J.-f. *J. Lumin.* **2014,** *149,* 208–214. doi:10.1016/j.jlumin.2014.01.023
30. Zagidullin, A. A.; Naileva, F. F.; Fayzullin, R. R.; Islamov, D. R.; Miluykov, V. A. *J. Mol. Struct.* **2025,** *1338,* 142323. doi:10.1016/j.molstruc.2025.142323
31. Amirov, R. R.; Akhmadiev, K.; Gaifutdinov, A. M.; Andrianova, K. A.; Shmelev, A.; Gatiatulin, A. K.; Zagidullin, A. A.; Milyukov, V. A.; Amirova, L. M. *Mater. Today Chem.* **2023,** *29,* 101464. doi:10.1016/j.mtchem.2023.101464

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.21.148

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Rhodium-catalysed connective synthesis of diverse reactive probes bearing S(VI) electrophilic warheads

Scott Rice[1,2], Julian Chesti[1,2], William R. T. Mosedale[3], Megan H. Wright[1,2], Stephen P. Marsden[1], Terry K. Smith[3] and Adam Nelson[*1,2]

Full Research Paper

Open Access

Address:
[1]School of Chemistry, University of Leeds, Leeds, LS2 9JT, UK, [2]Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, LS2 9JT, UK and [3]Schools of Biology and Chemistry, Biomedical Sciences Research Complex, University of St Andrews, St Andrews, KY16 9ST, UK

Email:
Adam Nelson[*] - a.s.nelson@leeds.ac.uk

* Corresponding author

Keywords:
covalent probes; molecular diversity; rhodium carbenoids

Beilstein J. Org. Chem. **2025,** *21,* 1924–1931.
https://doi.org/10.3762/bjoc.21.150

Received: 20 June 2025
Accepted: 08 September 2025
Published: 17 September 2025

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2025 Rice et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

The value of small molecules that chemically modify proteins is increasingly being recognised and utilised in both chemical biology and drug discovery. The discovery of such chemical tools may be enabled by screening diverse sets of reactive probes. Most existing sets of reactive probes are armed with cysteine-directed warheads, a limitation that we sought to address. A connective synthesis was developed in which α-diazoamide substrates, armed with a S(VI) warhead, were reacted with diverse co-substrates. A high-throughput approach was used to identify promising substrate/co-substrate/catalyst combinations which were then prioritised for purification by mass-directed HPLC to yield a total of thirty reactive probes. The structural diversity of the probe set was increased by the multiplicity of reaction types between rhodium carbenoids and the many different co-substrate classes, and the catalyst-driven selectivity between these pathways. The probes were screened for activity against *Trypanosoma brucei*, and four probes with promising anti-trypanosomal activity were identified. Remarkably, the synthetic approach was compatible with building blocks bearing three different S(VI) warheads, enabling the direct connective synthesis of diverse reactive probes armed with non-cysteine-directed warheads. Reactive probes that are synthetically accessible using our approach may be of value in the discovery of small molecule modifiers for investigating and engineering proteins.

Introduction

Diverse sets of reactive probes can facilitate the discovery of chemical tools and drugs that chemically modify protein targets [1-3]. Established sets of reactive probes are typically armed with electrophilic warheads that have the potential to target nucleophilic amino acid side chains. Most reactive probe sets bear cysteine-directed warheads [3-7], although sets have also been designed to target a wider range of amino acids [8-10]. Sets of reactive probes are generally prepared using robust reac-

tions, most usually amide formation, chosen from the toolkit that currently dominates medicinal chemistry [11] which may, in turn, limit probe structural diversity.

We have developed a unified connective approach for the synthesis of structurally diverse reactive probes bearing S(VI) electrophiles. Proteome-wide screens have shown that S(VI) electrophiles predominantly target lysine and tyrosine [12], although other residues (e.g. serine) may also be targeted within enzyme active sites [13]. It was envisaged that the reactive probes would be prepared by dirhodium-catalysed reactions between pairs of building blocks: an α-diazoamide **2** bearing a S(VI) electrophile and a suitable co-substrate (→ **3**) (Figure 1). Here, metal-catalysed carbenoid chemistry was chosen because of the wide range of potentially reactive functional groups that might be incorporated into co-substrates [14]. The richness of potential connective chemistry, and the availability of alternative dirhodium catalysts with distinctive reactivity, was expected to expand the structural diversity of accessible reactive probes. Herein, we describe the successful execution of this approach and the demonstration of biological function of the resulting reactive probes.

Results and Discussion

We prepared five α-diazoamide substrates bearing S(VI) electrophiles (Scheme 1 and Table 1) [15]. Initially, three amines – morpholine, 4-phenylpiperidine and isoindoline – were reacted with 2,2,6-trimethyl-4*H*-1,3-dioxin-4-one to give the corresponding β-ketoamides **4**. Treatment of the β-ketoamides **4** with 4-acetamidobenzenesulfonyl azide (*p*-ABSA) and triethylamine gave the α-diazo-β-ketoamides **5**. Subsequent KOH-mediated deacetylation yielded the corresponding α-diazoamides **1**. Finally, Pd-catalysed cross-coupling with warhead-substituted phenyl iodides gave, in low to moderate yield, the required α-diazoamide substrates **2** (referred to individually as **D1**–**5** below). Whilst the Pd-catalysed arylation of α-diazoamides and esters is known [15-19], its tolerance of pendant S(VI) electrophiles has not been previously explored and is notable.

Due to the relatively large size of the diazo substrates **D1**–**5**, it was decided to design a set of diverse co-substrates with 15 or fewer heavy (non-hydrogen) atoms. It was decided that the set should include co-substrates with the potential to react with metal carbenoids in many different ways [14], for example through O–H, N–H or formal C–H insertion, cyclopropanation, or oxazole [20] formation. The 16 co-substrates, selected from available compounds in our laboratory, are shown in Figure 2 (panel A). Many of these substrates had more than one potentially reactive site to enable, for example, O–H insertion (**C1**–**5**, **C8**, **C11** and **C14**), N–H insertion (**C3**, **C6**, **C12**, **C13** and **C15**), formal C–H insertion (**C1**, **C3, C4**, **C12**, **C15** and **C16**), oxazole formation (**C9** and **C10**) and cyclopropanation (**C7**, **C10**, **C14** and **C16**).

To start with, we investigated reactions of the α-diazoamide substrates **D1**, **D2** and **D3** with the 16 co-substrates **C1**–**16** catalysed by three diverse [21] dirhodium catalysts (Rh_2piv_4, Rh_2pfb_4 and Rh_2cap_4) i.e., an array of 144 reactions. An α-diazoamide substrate (20 μmol; 16 μL of a 1.25 M solution in

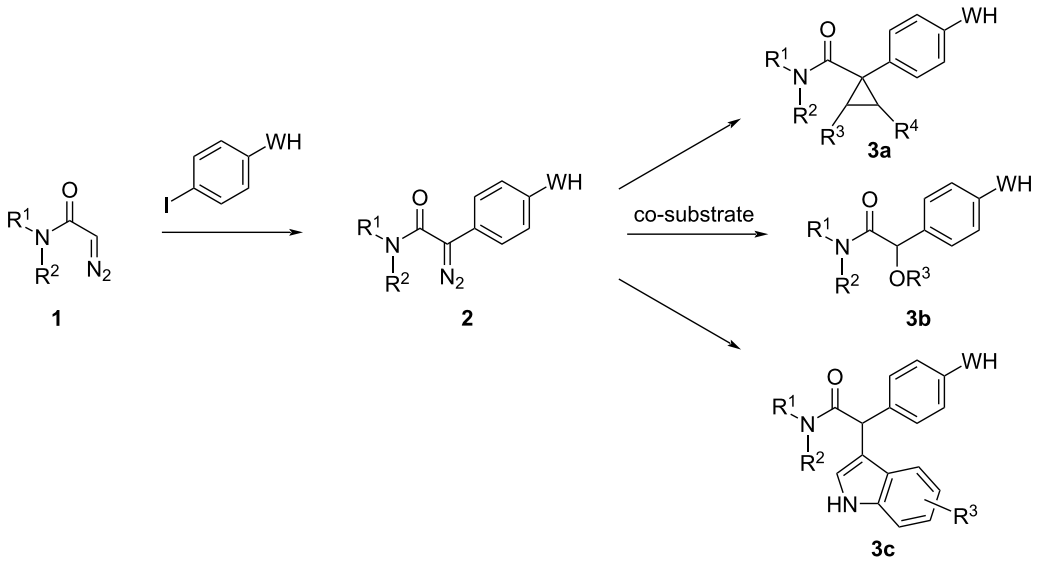
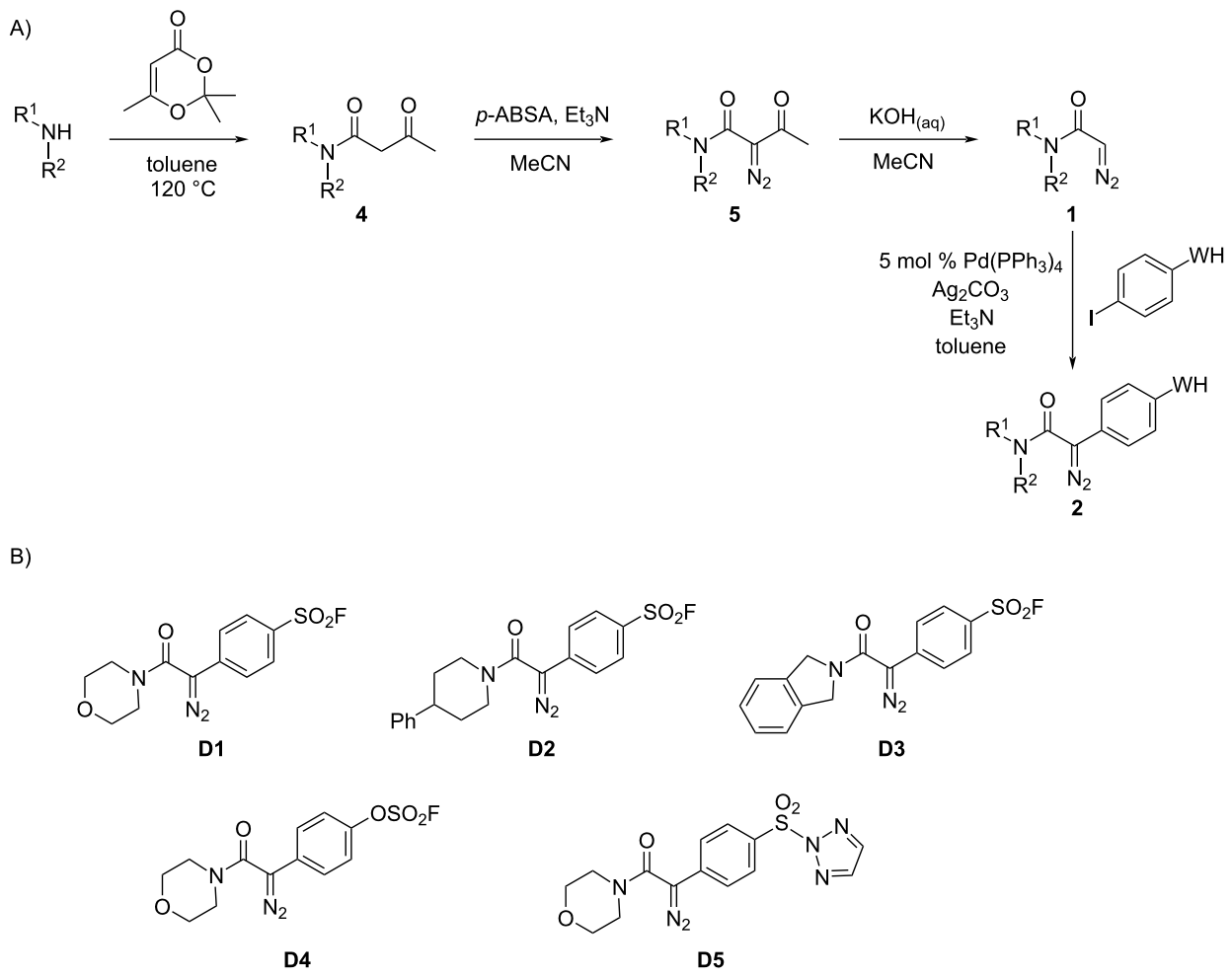
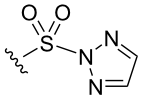


Figure 1: Envisaged connective synthesis of reactive probes **3** bearing S(VI) electrophilic warheads (WH). Diverse probes **3** might be accessible by functionalising α-diazoamide substrates **2** via alternative reaction modes.



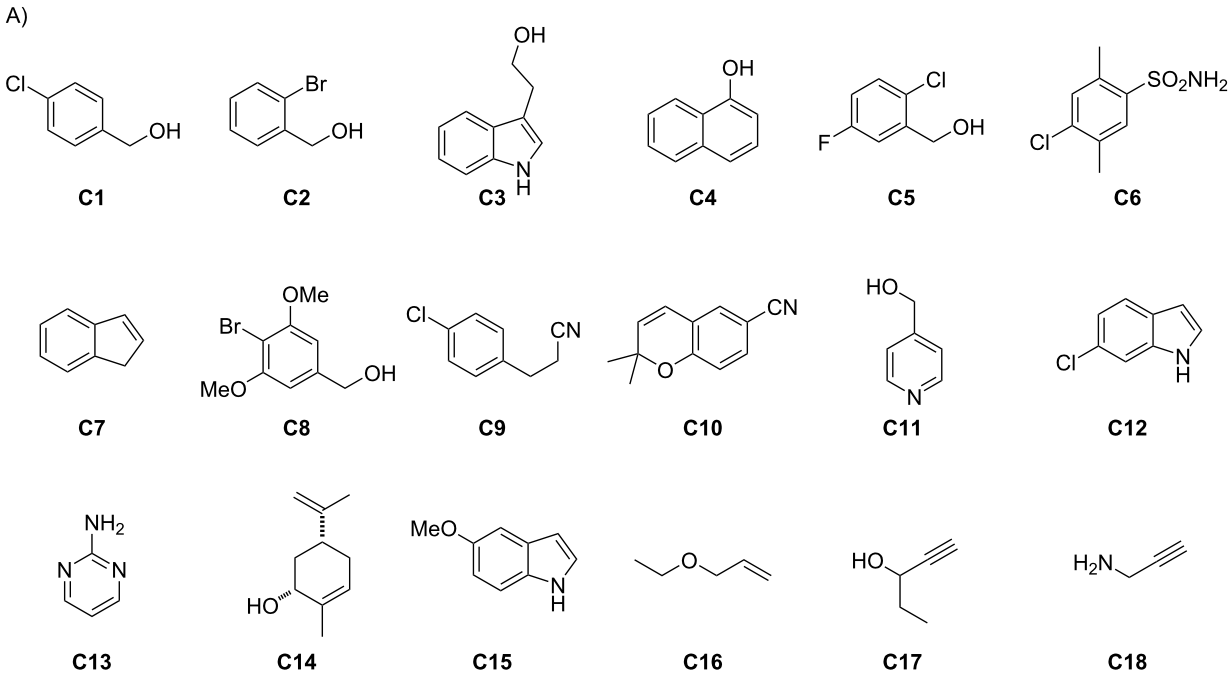
Scheme 1: Synthesis of α-diazoamide substrates **D1**–**5** of general structure **2** bearing S(VI) electrophiles. Panel A: Overview of synthesis (see Table 1 for details of synthesis of individual substrates). Panel B: Substrates that were prepared.

Table 1: Synthesis of α-diazoamide substrates of general structure **2** bearing S(VI) electrophiles (see Scheme 1).

Amine	Yield **4** (%)	Yield **5** (%)	Yield **1** (%)	WH	Substrate (yield, %)
morpholine	94	80	55	–SO_2F	**D1** (46)
				–OSO_2F	**D4** (26)
					**D5** (23)
4-phenylpiperidine	85	82	87	–SO_2F	**D2** (53)
isoindoline	88	86	99	–SO_2F	**D3** (12)

CH_2Cl_2) and a co-substrate (5 equiv; 16 µL of a 6.25 M solution in CH_2Cl_2) were added to glass vials in a 96-well reaction block, and the solvent left to evaporate after each addition. Subsequently, a dirhodium catalyst (1 mol %; 200 µL of a 1 mM solution in CH_2Cl_2) was also added to each vial. The final volume of each reaction was thus 200 µL, with final concentrations of 100 mM (for substrates), 500 mM (for co-substrates) and 1 mM (for catalysts).

A)



B)

	catalyst	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
D1	Rh_2piv_4	8	10	7	2	10	7	–	10	–	–	–	8[a]	9	–	15[a]	–
	Rh_2pfb_4	12	9	16	8	13	6	–	11	–	–	–	14[a]	–	6	24[a]	–
	Rh_2cap_4	16	14	10[a]	4	12	9	4	14	–	–	–	3	–	6	10	–
D2	Rh_2piv_4	10	11	11[a]	2	13	12	12	13	–	–	–	–	12	8	19[a]	–
	Rh_2pfb_4	18	13	18	13	17	5	7	14	–	–	–	17[a]	–	12	17[a]	–
	Rh_2cap_4	10	19	14[a]	10	20	13	2	18	–	–	–	4	–	15	15	–
D3	Rh_2piv_4	5	14	–	9	5	7	–	4	–	–	–	2	9	-	5[a]	–
	Rh_2pfb_4	3	8	9	8	3	5	–	9	–	–	–	5[a]	–	2	2	–
	Rh_2cap_4	3	4	7	10	4	2	5	6	–	–	–	3[a]	–	4	10[a]	–

Figure 2: Structures and reactions of co-substrates. Panel A: structures of the 16 selected co-substrates **C1–16**, together with two additional co-substrates **C17** and **C18** that were subsequently used. Panel B: yields, estimated by evaporative light scattering detection, of reactions involving combinations of substrates, co-substrates and catalysts (dash: <2% estimated yield). Highlighted combinations (green boxes) were selected for mass-directed purification. [a]Multiple intermolecular products observed by analytical HPLC.

After 48 h, the outcome of the reactions was determined by analytical UPLC–MS with, additionally, evaporative light-scattering detection [22,23] to enable estimation of the yield of each product (Figure 2, panel B). It was found that many reactions involving alcohol- (e.g., **C1–5**, **C8**, **C11** and **C14**) and indole- (e.g., **C3**, **C12** and **C15**) containing co-substrates yielded intermolecular products, whilst those involving nitrile-containing co-substrates (**C9** and **10**) and the allylic ether **C16** did not. It is remarkable that S(VI) electrophiles are tolerated. Eighteen substrate/co-substrate combinations gave, with at least one of the catalysts, an intermolecular product in >10% estimated yield (typically corresponding to >1 mg product). For all but one of these reactions, a product with molecular weight consistent with O–H insertion into water was also observed. For these 18 substrate/co-substrate combinations, the reaction with the highest estimated yield was selected for mass-directed purification (Table 2). In total, 23 intermolecular reaction products were isolated and structurally characterised (using, where appropriate, HMBC, COSY and nOe NMR methods; see Figure 3). In general, the yields of these products were rather low, which

Table 2: Outcomes of reactions between α-diazoamide substrates and co-substrates.

Diazo	Co-substrate	Catalyst	Product[a]	Yield[b]
D1	**C1**	Rh_2cap_4	**1-1**	14
D1	**C2**	Rh_2cap_4	**1-2**	12
D1	**C3**	Rh_2pfb_4	**1-3a** **1-3b**	15 1
D1	**C5**	Rh_2pfb_4	**1-5**	11
D1	**C8**	Rh_2cap_4	**1-8**	12
D1	**C15**	Rh_2piv_4	**1-15a** **1-15b**	6 8
D2	**C2**	Rh_2cap_4	**2-2**	14
D2	**C3**	Rh_2pfb_4	**2-3a** **2-3b**	13 1
D2	**C4**	Rh_2pfb_4	**2-4**	11
D2	**C5**	Rh_2cap_4	**2-5**	14
D2	**C6**	Rh_2cap_4	**2-6**	10
D2	**C7**	Rh_2piv_4	**2-7**	13[c]
D2	**C8**	Rh_2cap_4	**2-8**	13
D2	**C13**	Rh_2piv_4	**2-13**	12
D2	**C14**	Rh_2cap_4	**2-14**	10[d]
D2	**C15**	Rh_2cap_4	**2-15a** **2-15b**	11 1
D3	**C2**	Rh_2piv_4	**3-2**	13
D3	**C4**	Rh_2cap_4	**3-4a** **3-4b**	5[e] 5[e]
D4	**C1**	Rh_2pfb_4	**4-1**	56
D4	**C3**	Rh_2pfb_4	**4-3**	23
D4	**C5**	Rh_2cap_4	**4-5**	8
D4	**C13**	Rh_2piv_4	**4-13**	35
D4	**C17**	Rh_2pfb_4	**4-17**	11
D4	**C18**	Rh_2pfb_4	**4-18**	23
D5	**C1**	Rh_2pfb_4	**5-1**	26

[a]Reactions were performed in glass vials with an α-diazoamide substrate (20 μmol; limiting reactant), a co-substrate (5 equiv) and 1 mol % dirhodium catalyst. [b]Isolated yield of purified product. [c]dr: >95:<5. [d]dr: 51:49. [e]Obtained as a 50:50 mixture of inseparable products.

may stem from poor (co-)substrate solubility in some cases; and/or competitive O–H insertion into adventitious water.

On the basis of these results, additional reactions involving the α-diazoamide substrates **D4** (with a fluorosulfate warhead) and **D5** (with a sulfonyltriazole warhead) were also executed. In addition to using these two α-diazoamide substrates with different warheads, two additional co-substrates bearing an alkyne tag (**C17** and **C18**) were used. The reactions were assembled from stock solutions, with some variation in stock concentrations to improve solubility. After 24 h, the reaction products were analysed by LC–MS, and promising reactions selected for mass-directed purification. Seven additional intermolecular products were obtained (see Figure 3 and Table 2). The marked improvement in product yields, compared to those observed with **D1**–**3**, may reflect the change to the workflow, i.e., variation in stock concentration to improve solubility.

The diversity of the obtained products was increased by the multiple reaction modes of dirhodium carbenoids that were possible [14]. Overall, products were formed via O–H insertion into an alcohol (to give 14 products) or phenol (→ **2-4** and **3-4a**); N–H insertion into an indole (→ **1-3a**, **1-15b**, **2-3a**, **2-15b** and **4-3**), sulfonamide (→ **2-6**), aminopyrimidine (→ **2-13** and **4-13**) or amine (→ **4-18**); cyclopropanation (→ **2-7**); and formal C–H insertion into an indole (→ **1-15a** and **2-15a**) or naphthol (→ **2-4** and **3-4b**). In the case of **4** (2-naphthol) and **15** (5-methoxyindole), co-substrates containing functional groups with more than one potentially reactive site, two regioisomeric products were obtained. In the case of co-substrate **3**, which contains both an indole and an alcohol, thus raising chemoselectivity issues, products were observed from both O–H and N–H insertion. It is notable, however, that despite many of the co-substrates having multiple potentially reactive sites, one intermolecular reaction was generally dominant.

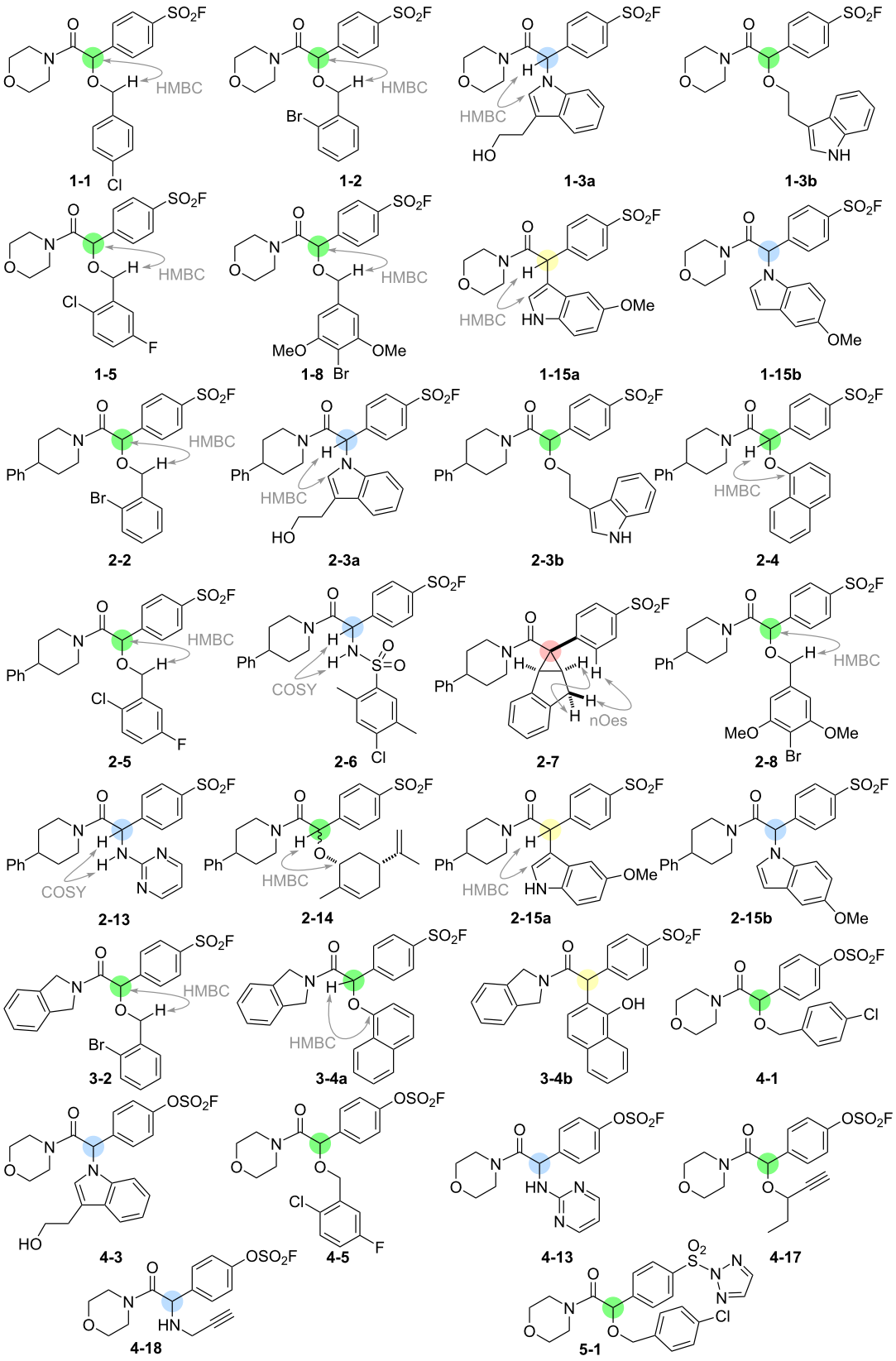


Figure 3: Structures and structure elucidation of intermolecular reaction products. The relevant reactivity modes are indicated by colour: O–H insertion (green); N–H insertion (blue); formal C–H insertion (yellow); and cyclopropanation (pink).

We have previously discovered sulfonyl fluoride probes with promising activity against *T. brucei,* a parasitic kinetoplastid that causes vector-borne African trypanosomiasis (sleeping sickness) [24]. We therefore screened the 23 sulfonyl fluoride probes (derived from diazo compounds **1**, **2** and **3**) against *T. brucei* in 96-well plate format (final concentrations: ≈2–50 μM). Four sulfonyl fluorides were found to have promising activity: **2-5** (EC_{50}: 9.38 ± 0.06 μM); **2-6** (EC_{50}: 6.81 ± 0.07 μM); **2-14** (EC_{50}: 9.26 ± 0.06 μM) and **2-15a** (EC_{50}: 11.9 ± 0.2 μM). It is notable that all of these active compounds are 4-phenylpiperidinyl amides derived from the same α-diazoamide **2**, suggesting that this feature is important for activity.

Conclusion

We have developed a connective synthesis of reactive probes bearing S(VI) electrophilic warheads. Each probe was prepared by rhodium-catalysed reaction between an α-diazo amide substrate bearing a warhead, and a co-substrate. The structural diversity of the probe set was increased by the multiple possible reaction modes of rhodium carbenoids, which enabled many different co-substrate classes and catalyst-driven selectivities to be exploited. A high-throughput synthetic approach was harnessed to identify substrate/co-substrate/catalyst combinations, which led to the productive formation of intermolecular reaction products. Overall, the approach enabled the synthesis of thirty diverse reactive probes. The probes were screened for activity against *T. brucei,* a parasitic kinetoplastid that causes vector-borne African trypanosomiasis, and four probes with promising anti-trypanosomal activity were identified. Remarkably, the synthetic approach was compatible with building blocks bearing three different S(VI) warheads, and enabled the direct connective synthesis of diverse reactive probes. We envisage that such probes may enable chemical modification of non-cysteine residues within proteins, and may be valuable in investigating and engineering the biology of proteins.

Supporting Information

Supporting Information File 1

Experimental part and NMR spectra of synthesised compounds.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-21-150-S1.pdf]

Acknowledgements

We thank Dan Cox, Samuel Liver, Chris Arter, Jeanine Williams, Mark Howard, Alex Heyam, Lawrence Collins and Stuart Warriner for their support and assistance in this work. This work is based on the doctoral theses of Scott Rice ("Synthesis of Novel Polyfunctional 3D Scaffolds for Drug Discovery", University of Leeds, 2021) and Julian Chesti ("Modular Synthesis and Biological Evaluation of Structurally-Diverse Reactive Fragment Libraries", University of Leeds, 2025).

Funding

We thank Redbrick Molecular, EPSRC (EP/N025652/1; EP/W002914/1) and the Leverhulme Trust (RPG-2018-030) for funding.

Author Contributions

Scott Rice: conceptualization; data curation; investigation; methodology; writing – review & editing. Julian Chesti: data curation; investigation; methodology; writing – review & editing. William R. T. Mosedale: data curation; investigation; methodology; writing – review & editing. Megan H. Wright: funding acquisition; supervision; writing – review & editing. Stephen P. Marsden: funding acquisition; supervision; writing – review & editing. Terry K. Smith: funding acquisition; supervision; writing – review & editing. Adam Nelson: conceptualization; funding acquisition; supervision; writing – original draft.

ORCID® iDs

Julian Chesti - https://orcid.org/0000-0001-5416-7648
Megan H. Wright - https://orcid.org/0000-0003-2731-4707
Stephen P. Marsden - https://orcid.org/0000-0002-2723-8954
Adam Nelson - https://orcid.org/0000-0003-3886-363X

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Drewes, G.; Knapp, S. *Trends Biotechnol.* **2018,** *36,* 1275–1286. doi:10.1016/j.tibtech.2018.06.008
2. Boike, L.; Henning, N. J.; Nomura, D. K. *Nat. Rev. Drug Discovery* **2022,** *21,* 881–898. doi:10.1038/s41573-022-00542-z
3. Lu, W.; Kostic, M.; Zhang, T.; Che, J.; Patricelli, M. P.; Jones, L. H.; Chouchani, E. T.; Gray, N. S. *RSC Chem. Biol.* **2021,** *2,* 354–367. doi:10.1039/d0cb00222d
4. Crowley, V. M.; Thielert, M.; Cravatt, B. F. *ACS Cent. Sci.* **2021,** *7,* 613–623. doi:10.1021/acscentsci.0c01336
5. Douangamath, A.; Fearon, D.; Gehrtz, P.; Krojer, T.; Lukacik, P.; Owen, C. D.; Resnick, E.; Strain-Damerell, C.; Aimon, A.; Ábrányi-Balogh, P.; Brandão-Neto, J.; Carbery, A.; Davison, G.; Dias, A.; Downes, T. D.; Dunnett, L.; Fairhead, M.; Firth, J. D.; Jones, S. P.; Keeley, A.; Keserü, G. M.; Klein, H. F.; Martin, M. P.; Noble, M. E. M.; O'Brien, P.; Powell, A.; Reddi, R. N.; Skyner, R.; Snee, M.; Waring, M. J.; Wild, C.; London, N.; von Delft, F.; Walsh, M. A. *Nat. Commun.* **2020,** *11,* 5047. doi:10.1038/s41467-020-18709-w

6. Resnick, E.; Bradley, A.; Gan, J.; Douangamath, A.; Krojer, T.; Sethi, R.; Geurink, P. P.; Aimon, A.; Amitai, G.; Bellini, D.; Bennett, J.; Fairhead, M.; Fedorov, O.; Gabizon, R.; Gan, J.; Guo, J.; Plotnikov, A.; Reznik, N.; Ruda, G. F.; Díaz-Sáez, L.; Straub, V. M.; Szommer, T.; Velupillai, S.; Zaidman, D.; Zhang, Y.; Coker, A. R.; Dowson, C. G.; Barr, H. M.; Wang, C.; Huber, K. V. M.; Brennan, P. E.; Ovaa, H.; von Delft, F.; London, N. *J. Am. Chem. Soc.* **2019,** *141,* 8951–8968. doi:10.1021/jacs.9b02822
7. St. Denis, J. D.; Chessari, G.; Cleasby, A.; Cons, B. D.; Cowan, S.; Dalton, S. E.; East, C.; Murray, C. W.; O'Reilly, M.; Peakman, T.; Rapti, M.; Stow, J. L. *J. Med. Chem.* **2022,** *65,* 12319–12333. doi:10.1021/acs.jmedchem.2c01044
8. Brulet, J. W.; Borne, A. L.; Yuan, K.; Libby, A. H.; Hsu, K.-L. *J. Am. Chem. Soc.* **2020,** *142,* 8270–8280. doi:10.1021/jacs.0c00648
9. Gilbert, K. E.; Vuorinen, A.; Aatkar, A.; Pogány, P.; Pettinger, J.; Grant, E. K.; Kirkpatrick, J. M.; Rittinger, K.; House, D.; Burley, G. A.; Bush, J. T. *ACS Chem. Biol.* **2023,** *18,* 285–295. doi:10.1021/acschembio.2c00633
10. Huang, T.; Hosseinibarkooie, S.; Borne, A. L.; Granade, M. E.; Brulet, J. W.; Harris, T. E.; Ferris, H. A.; Hsu, K.-L. *Chem. Sci.* **2021,** *12,* 3295–3307. doi:10.1039/d0sc06623k
11. Boström, J.; Brown, D. G.; Young, R. J.; Keserü, G. M. *Nat. Rev. Drug Discovery* **2018,** *17,* 709–727. doi:10.1038/nrd.2018.116
12. Zanon, P. R. A.; Yu, F.; Musacchio, P.; Lewald, L.; Zollo, M.; Krauskopf, K.; Mrdović, D.; Raunft, P.; Maher, T. E.; Cigler, M.; Chang, C.; Lang, K.; Toste, F. D.; Nesvizhskii, A. I.; Hacker, S. M. *ChemRxiv* **2021**. doi:10.26434/chemrxiv-2021-w7rss-v2
13. Narayanan, A.; Jones, L. H. *Chem. Sci.* **2015,** *6,* 2650–2659. doi:10.1039/c5sc00408j
14. Doraghi, F.; Baghershahi, P.; Ghasemi, M.; Mahdavi, M.; Al-Harrasi, A. *RSC Adv.* **2024,** *14,* 39337–39352. doi:10.1039/d4ra07010k See for a review.
15. Chow, S.; Green, A. I.; Arter, C.; Liver, S.; Leggott, A.; Trask, L.; Karageorgis, G.; Warriner, S.; Nelson, A. *Synthesis* **2020,** *52,* 1695–1706. doi:10.1055/s-0039-1690905
16. Peng, C.; Cheng, J.; Wang, J. *J. Am. Chem. Soc.* **2007,** *129,* 8708–8709. doi:10.1021/ja073010+
17. Ye, F.; Qu, S.; Zhou, L.; Peng, C.; Wang, C.; Cheng, J.; Hossain, M. L.; Liu, Y.; Zhang, Y.; Wang, Z.-X.; Wang, J. *J. Am. Chem. Soc.* **2015,** *137,* 4435–4444. doi:10.1021/ja513275c
18. Yamamoto, K.; Qureshi, Z.; Tsoung, J.; Pisella, G.; Lautens, M. *Org. Lett.* **2016,** *18,* 4954–4957. doi:10.1021/acs.orglett.6b02423
19. Fu, L.; Mighion, J. D.; Voight, E. A.; Davies, H. M. L. *Chem. – Eur. J.* **2017,** *23,* 3272–3275. doi:10.1002/chem.201700101
20. Moody, C. J.; Doyle, K. J. *Prog. Heterocycl. Chem.* **1997,** *9,* 1–16. doi:10.1016/s0959-6380(97)80003-7
21. Green, A. I.; Tinworth, C. P.; Warriner, S.; Nelson, A.; Fey, N. *Chem. – Eur. J.* **2021,** *27,* 2402–2409. doi:10.1002/chem.202003801
22. Squibb, A. W.; Taylor, M. R.; Parnas, B. L.; Williams, G.; Girdler, R.; Waghorn, P.; Wright, A. G.; Pullen, F. S. *J. Chromatogr. A* **2008,** *1189,* 101–108. doi:10.1016/j.chroma.2007.11.017
23. Griggs, S. D.; Piticari, A.-S.; Liver, S.; Arter, C.; Sievers, S.; Marsden, S. P.; Nelson, A. *Chem. Commun.* **2025,** *61,* 3528–3531. doi:10.1039/d4cc06605g
24. Mantilla, B. S.; White, J. S.; Mosedale, W. R. T.; Gomm, A.; Nelson, A.; Smith, T. K.; Wright, M. H. *Commun. Chem.* **2024,** *7,* 237. doi:10.1038/s42004-024-01327-8

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at: https://doi.org/10.3762/bjoc.21.150

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Efficient solid-phase synthesis and structural characterization of segetalins A–H, J and K

Liangyu Liu, Wanqiu Lu, Quanping Guo* and Zhaoqing Xu*

Full Research Paper

Open Access

Address:
Key Laboratory of Preclinical Study for New Drugs of Gansu Province, School of Basic Medical Sciences, Lanzhou University, 199 West Donggang Road, Lanzhou 730000, China

Email:
Quanping Guo* - guoqp@lzu.edu.cn; Zhaoqing Xu* - zqxu@lzu.edu.cn

* Corresponding author

Keywords:
Fmoc-solid-phase peptide synthesis (Fmoc-SPPS); head-to-tail cyclization; plant cyclopeptides; *Vaccaria segetalis*

Beilstein J. Org. Chem. **2025,** *21,* 2612–2617.
https://doi.org/10.3762/bjoc.21.202

Received: 04 September 2025
Accepted: 10 November 2025
Published: 27 November 2025

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2025 Liu et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

This study establishes an efficient solid-phase strategy for the total synthesis of segetalins A–H, J and K (**1**–**10**), bioactive cyclopeptides isolated from *Vaccaria segetalis*. Linear precursors were assembled on cost-effective 2-chlorotrityl chloride resin via Fmoc-SPPS, followed by PyBOP-mediated head-to-tail cyclization in DMF (10^{-3} M). After RP-HPLC purification, all cyclopeptides were obtained in 45–70% isolated yields. Structural identities were confirmed by HRESIMS, NMR, and HPLC (>95% purity). Circular dichroism (CD) spectroscopy revealed distinct secondary structures, including β-sheets (**1**, **2**, **3**, **4**, **7**, **8**, **10**) and α-helical elements (**5**, **6**). This scalable methodology overcomes limitations of prior syntheses, enabling biological evaluation.

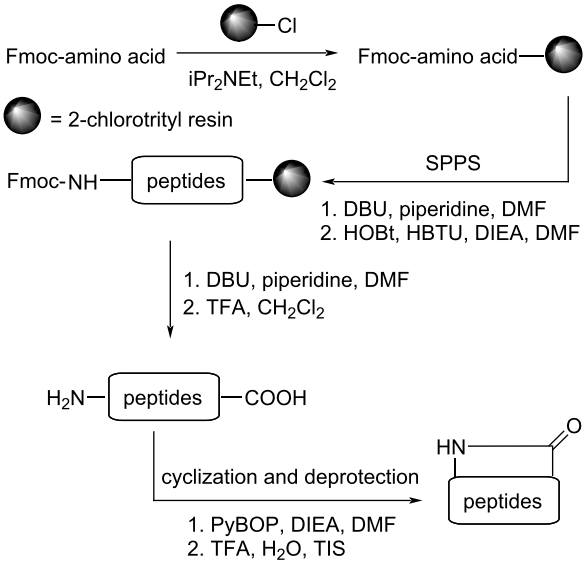
Introduction

Cyclopeptides have garnered significant research interest owing to their unique conformational constraints imposed by cyclization and diverse biological activities [1-3]. Specifically, plant-derived cyclopeptides represent a valuable source of potential lead compounds for drug discovery [4]. Segetalins A–H, J and K (**1**–**10**), isolated from the seeds of *Vaccaria segetalis* (Caryophyllaceae), are head-to-tail cyclic oligopeptides comprising 5–9 amino acid residues [5-13]. These natural products exhibit a significant diversity of pharmacological activities [14-16], including estrogen-like activity (**1**, **2**, **7**, **8**), antitumor effects (**5**), and antimicrobial properties (**3**). Given their unique structural features and pharmacological potential, segetalins have become important targets for both synthetic chemistry and drug development. However, efficient and general synthetic routes to access this family have remained limited over the past decades.

Previous synthetic approaches have encountered significant challenges. Sonnet et al. reported the first total synthesis of segetalin A (**1**) via Sasrin resin-based SPPS, followed by cyclization under highly dilute conditions (10^{-4} M) with diphenylphosphoryl azide (DPPA) [17]. While successful, this approach suffers from the high cost of the specialized resin and large solvent volumes required for dilution, coupled with DPPA's poor efficiency in forming sterically hindered peptide bonds involving residues like Val or Ile. Dahiya and Kaur synthesized sege-

talin C (**3**) via a solution-phase fragment coupling strategy, culminating in cyclization mediated by *N*,*N*'-dicyclohexylcarbodiimide (DCC)/*N*-methylmorpholine (NMM) at 0 °C for 7 days [18]. This method, however, is lengthy, operationally complex, difficult for product isolation, and carries a significant risk of racemization. Wong and Jolliffe synthesized segetalins B (**2**) and G (**7**) using a pseudoprolinic acid strategy to induce *cis*-amide bond formation, followed by desulfurization [19]. Despite achieving cyclization, this route involves intricate procedures, expensive starting materials, and has limited applicability to other segetalins.

Given the limitations of existing methodologies and the biological significance of the segetalins, we sought to develop an efficient, scalable, and generally applicable solid-phase synthesis strategy for the *Vaccaria segetalis* cyclopeptide family.



Scheme 1: Preparation of segetalins A–H, J and K.

Results and Discussion

Synthesis strategy and optimization

While both solution-phase and solid-phase approaches are viable for peptide synthesis, Fmoc-based SPPS offers distinct advantages in operational simplicity and efficiency for laboratory-scale production [20]. We therefore devised a streamlined solid-phase strategy for synthesizing the *Vaccaria segetalis* cyclopeptide family (Scheme 1).

Building upon previous work [17-19], we focused on optimizing key parameters: resin selection, Fmoc deprotection conditions, coupling reagents for linear assembly, and crucially, the cyclization step. Cost-effectiveness and commercial availability led us to select 2-chlorotrityl chloride resin as the solid support, enabling mild cleavage of the partially protected linear peptide precursor [21]. Efficient Fmoc deprotection was achieved using a solution of 1% pyridine and 1% 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in *N*,*N*-dimethylformamide (DMF) [22]. For the assembly of the linear sequences, coupling efficiency was significantly enhanced using a 1:1 mixture of 1-hydroxybenzotriazole (HOBt) and 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) in DMF [23]. Finally, we obtained crude linear peptides with 75% to 95% yields (Table 1).

The critical head-to-tail cyclization step proved challenging. Initial attempts using common coupling reagents such as 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU), HBTU, or HOBt alone in DMF failed to produce any detectable cyclized product [24-26]. Ultimately, successful macrocyclization was achieved by employing benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) as the coupling reagent in DMF

Table 1: Preparation of linear peptides for segetalins A–H, J and K.

	Structure	Yield[a]
1	Gly-Val-Pro-Val-Trp(Boc)-Ala	89%
2	Gly-Val-Ala-Trp(Boc)-Ala	77%
3	Gly-Leu-His(Trt)-Phe-Ala-Phe-Pro	93%
4	Gly-Leu-Ser(*t*-Bu)-Phe-Ala-Phe-Pro	91%
5	Gly-Tyr(*t*-Bu)-Val-Pro-Leu-Trp(Boc)-Pro	89%
6	Ala-Ser(*t*-Bu)-Tyr(*t*-Bu)-Ser(*t*-Bu)-Ser(*t*-Bu)-Lys(Boc)-Pro-Phe-Ser(*t*-Bu)	87%
7	Gly-Val-Lys(Boc)-Tyr(*t*-Bu)-Ala	95%
8	Gly-Tyr(*t*-Bu)-Arg(Pbf)-Phe-Ser(*t*-Bu)	94%
9	Phe-Gly-Thr(*t*-Bu)-His(Trt)-Gly-Leu-Pro-Ala-Pro	89%
10	Gly-Arg(Pbf)-Val-Lys(Boc)-Ala	87%

[a]Yield of crude linear peptide.

at a concentration of 10^{-3} M. After cleavage from the resin and global side-chain deprotection, the crude cyclic peptides were purified by preparative RP-HPLC. This optimized protocol afforded segetalins A–H, J and K (**1**–**10**) with 45% to 70% isolated yields (Figure 1).

Structural characterization

The synthetic compounds **1**–**10** were rigorously characterized to confirm their identity and purity (see Supporting Information File 1). High-resolution electrospray ionization mass spectrometry (HRESIMS) data for all compounds matched the calculated exact masses for their respective molecular formulas. NMR spectroscopic analysis in appropriate deuterated solvents (e.g., DMSO-d_6, D_2O) fully corroborated the amino acid sequence and cyclic connectivity, demonstrating unequivocal structural identity with the natural isolates. Analytical HPLC confirmed the high purity (>95%) of all synthetic segetalins. However, experimental data for segetalin C revealed the exis-

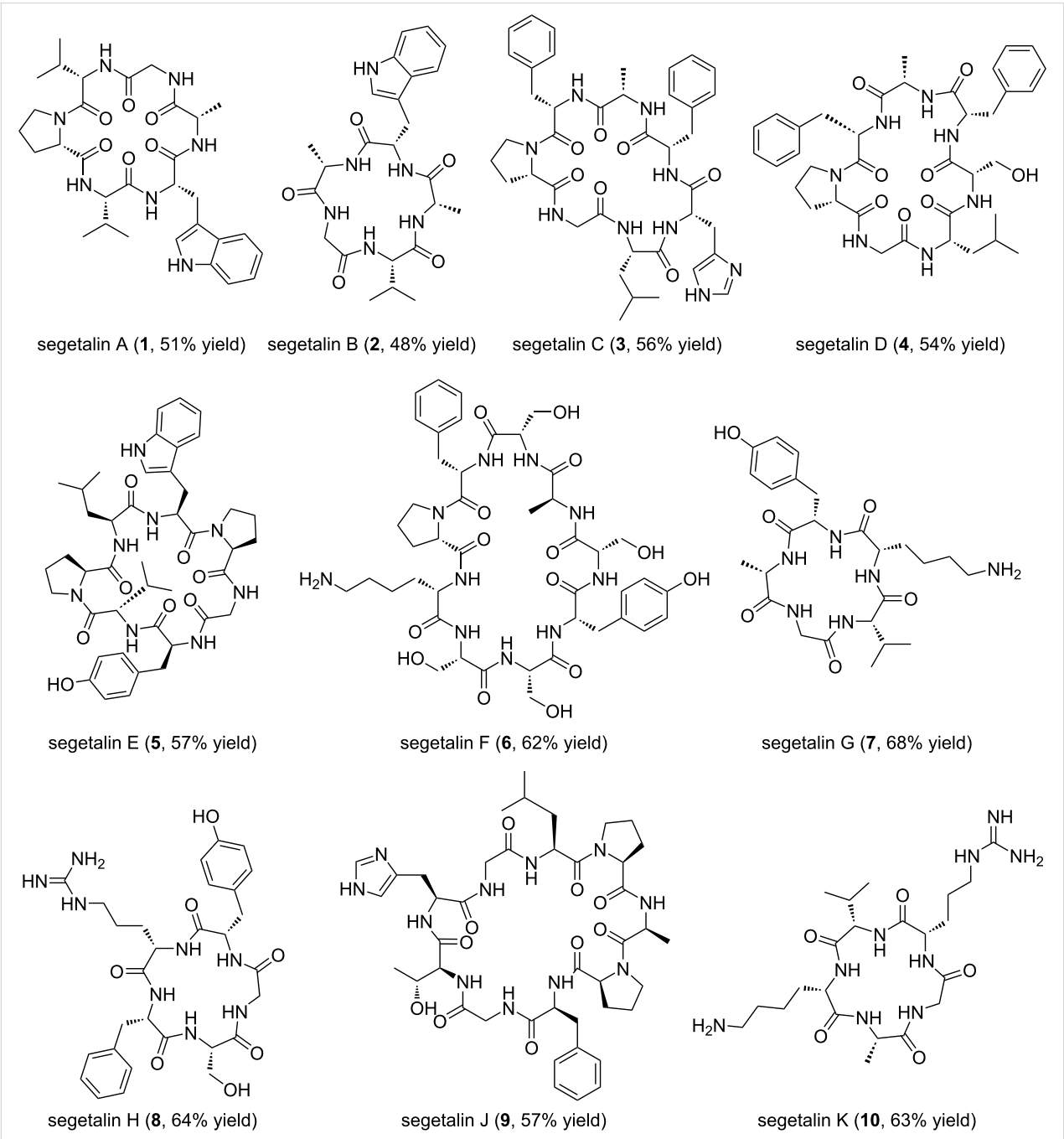


Figure 1: Cyclization reactions to segetalins A–H, J and K.

tence of a multi-state conformational equilibrium in solution, which is dependent on solvent polarity.

Secondary structure analysis by circular dichroism (CD)

The secondary structures of compounds **1**–**10** were investigated using circular dichroism (CD) spectroscopy in aqueous buffer (0.01×PBS), deionized H_2O, and 30% TFE (2,2,2-trifluoroethanol) (Figure 2). CD spectra in the far-UV region (190–250 nm) provide signatures of peptide backbone conformation [27-29]: a random coil typically shows a negative band near 198 nm and a positive band near 218 nm; an α-helix exhibits characteristic minima at 208 nm and 222 nm and a maximum near 192 nm; β-sheet structures are often indicated

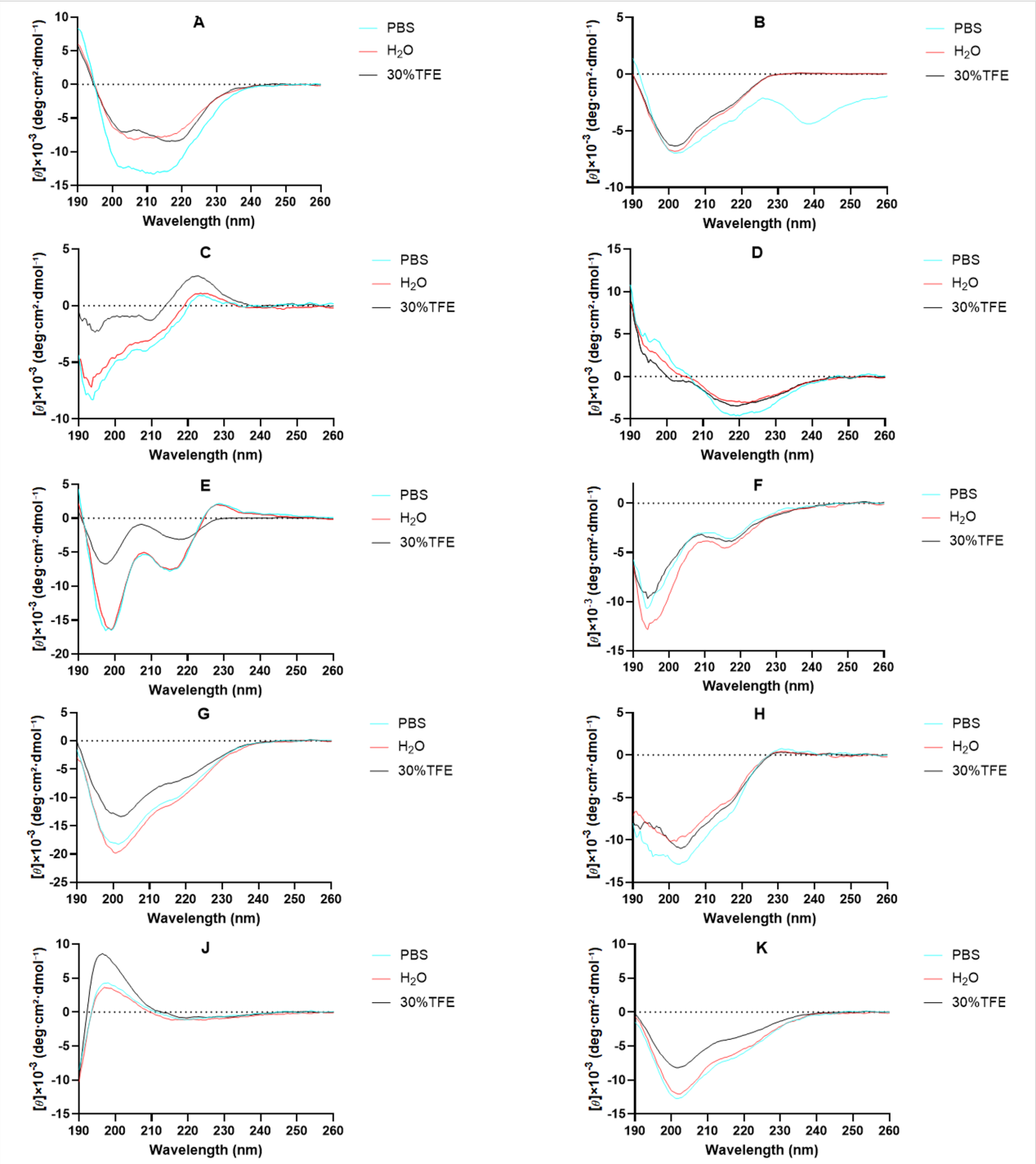


Figure 2: The CD spectra of segetalins A–H, J and K.

by a single minimum between 210–220 nm and a maximum below 200 nm.

CD spectroscopy revealed that cyclic topology shifts characteristic peaks and stabilizes secondary structures through cooperative ring size/sequence/solvent effects [30]. TFE disrupts the hydrogen-bonding network of water, reduces solvent polarity, and enhances the stability of intramolecular hydrogen bonds in the peptide backbone, thereby promoting the formation of stable secondary structures (such as α-helices or β-sheets) in cyclic peptides. Spectra acquired in H_2O, 0.01×PBS, and 30% TFE (Table S1, Supporting Information File 1) demonstrate: (i) definitive β-sheet signatures (217 nm minima) in **1**, **3**, **4**; (ii) enhanced β-sheet stability from constrained macrocycles in **2**, **7**, **8**, **10**; (iii) universal α-helix induction by TFE in **1**–**10**, with **5** and **6** retaining helicity in aqueous buffers.

Conclusion

We have developed an efficient and reliable solid-phase synthesis strategy for the cyclopeptide family of segetalins A–H, J and K from *Vaccaria segetalis*. Key optimizations include the use of a cost-effective 2-chlorotrityl chloride resin, efficient Fmoc deprotection and linear coupling conditions (HOBt/HBTU), and the identification of PyBOP as a highly effective coupling reagent for the challenging head-to-tail macrocyclization step under moderate dilution (10^{-3} M). This protocol afforded the target cyclic peptides in practical isolated yields (45–70%) and high purity. Comprehensive structural characterization by HRESIMS, NMR, and HPLC confirmed the identity and high purity of the synthetic segetalins. CD spectroscopy provided insights into their secondary structural preferences. This robust and scalable methodology overcomes significant limitations of previous synthetic approaches, providing ample quantities of these bioactive cyclopeptides for detailed biological evaluation and structure–activity relationship studies. The systematic investigation of the their key biological activities, including estrogenic activity (assessed via breast cell proliferation assays), antitumor activity (evaluated through HeLa cell inhibition assays), and antibacterial activity (evaluated against Gram-positive bacteria), will be conducted in our laboratory.

Supporting Information

Supporting Information File 1

Experimental section, characterization and copies of spectra.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-21-202-S1.pdf]

Funding

We are grateful to National Natural Science Foundation of China (U23A20524 & 22271126), Gansu Provincial Science and Technology Leading Talents Program (24RCKB005) and Gansu Science and Technology Program (25JRRA695).

Author Contributions

Liangyu Liu: investigation; methodology; writing – original draft. Wanqiu Lu: investigation; methodology. Quanping Guo: writing – original draft. Zhaoqing Xu: conceptualization; funding acquisition; project administration; supervision; writing – original draft.

ORCID® iDs

Zhaoqing Xu - https://orcid.org/0000-0001-7663-6249

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Ji, X.; Nielsen, A. L.; Heinis, C. *Angew. Chem., Int. Ed.* **2024,** *63,* e202308251. doi:10.1002/anie.202308251
2. Jin, K. *Future Med. Chem.* **2020,** *12,* 1687–1690. doi:10.4155/fmc-2020-0171
3. Nielsen, D. S.; Shepherd, N. E.; Xu, W.; Lucke, A. J.; Stoermer, M. J.; Fairlie, D. P. *Chem. Rev.* **2017,** *117,* 8094–8128. doi:10.1021/acs.chemrev.6b00838
4. Zhang, J.; Yuan, J.; Li, Z.; Fu, C.; Xu, M.; Yang, J.; Jiang, X.; Zhou, B.; Ye, X.; Xu, C. *Med. Res. Rev.* **2021,** *41,* 3096–3117. doi:10.1002/med.21792
5. Condie, J. A.; Nowak, G.; Reed, D. W.; Balsevich, J. J.; Reaney, M. J. T.; Arnison, P. G.; Covello, P. S. *Plant J.* **2011,** *67,* 682–690. doi:10.1111/j.1365-313x.2011.04626.x
6. Morita, H.; Sook Yun, Y.; Takeya, K.; Itokawa, H.; Shiro, M. *Tetrahedron* **1995,** *51,* 5987–6002. doi:10.1016/0040-4020(95)00277-f
7. Morita, H.; Yun, Y. S.; Takeya, K.; Itokawa, H.; Shirota, O. *Phytochemistry* **1996,** *42,* 439–441. doi:10.1016/0031-9422(95)00911-6
8. Yun, Y. S.; Morita, H.; Takeya, K.; Itokawa, H. *J. Nat. Prod.* **1997,** *60,* 216–218. doi:10.1021/np960617n
9. Morita, H.; Yun, Y. S.; Takeya, K.; Itokawa, H. *Tetrahedron Lett.* **1994,** *35,* 9593–9596. doi:10.1016/0040-4039(94)88519-2
10. Itokawa, H.; Yun, Y.; Morita, H.; Takeya, K.; Yamada, K. *Planta Med.* **1995,** *61,* 561–562. doi:10.1055/s-2006-959373
11. Morita, H.; Yun, Y. S.; Takeya, K.; Itokawa, H.; Yamada, K. *Tetrahedron* **1995,** *51,* 6003–6014. doi:10.1016/0040-4020(95)00278-g
12. Morita, H.; Eda, M.; Iizuka, T.; Hirasawa, Y.; Sekiguchi, M.; Yun, Y. S.; Itokawa, H.; Takeya, K. *Bioorg. Med. Chem. Lett.* **2006,** *16,* 4458–4461. doi:10.1016/j.bmcl.2006.06.083
13. Barber, C. J. S.; Pujara, P. T.; Reed, D. W.; Chiwocha, S.; Zhang, H.; Covello, P. S. *J. Biol. Chem.* **2013,** *288,* 12500–12510. doi:10.1074/jbc.m112.437947
14. Houshdar Tehrani, M. H.; Gholibeikian, M.; Bamoniri, A.; Mirjalili, B. B. F. *Front. Endocrinol.* **2021,** *11,* 600856. doi:10.3389/fendo.2020.600856

15. Dahiya, R.; Dahiya, S.; Shrivastava, J.; Fuloria, N. K.; Gautam, H.; Mourya, R.; Fuloria, S. *Arch. Pharm. (Weinheim, Ger.)* **2021,** *354,* 2000446. doi:10.1002/ardp.202000446
16. Tian, M.; Huang, Y.; Wang, X.; Cao, M.; Zhao, Z.; Chen, T.; Yuan, C.; Wang, N.; Zhang, B.; Li, C.; Zhou, X. *Front. Chem. (Lausanne, Switz.)* **2021,** *9,* 666280. doi:10.3389/fchem.2021.666280
17. Sonnet, P.; Petit, L.; Marty, D.; Guillon, J.; Rochette, J.; Brion, J.-D. *Tetrahedron Lett.* **2001,** *42,* 1681–1683. doi:10.1016/s0040-4039(00)02354-6
18. Dahiya, R.; Kaur, K. *Arzneim. Forsch.* **2008,** *58,* 29–34. doi:10.1055/s-0031-1296463
19. Wong, M. S. Y.; Jolliffe, K. A. *Pept. Sci.* **2018,** *110,* e24042. doi:10.1002/pep2.24042
20. Behrendt, R.; White, P.; Offer, J. *J. Pept. Sci.* **2016,** *22,* 4–27. doi:10.1002/psc.2836
21. In this study, solid-phase synthesis was accomplished using the cost-effective dichlorotrityl chloride (DCL) resin. Its commercial price is approximately $2.1 per gram (from suppliers such as Shanghai Acmec Biochemical Technology Co., Ltd), and it is readily available in multi-kilogram quantities. This price stands in sharp contrast to other resins commonly used in analogous syntheses: for instance, Wang resin and Sasrin resin are priced at about $100 per gram and these resins are not available in kilogram quantities (BACHEM). The significant cost advantage of DCL resin makes it an economically viable option for large-scale combinatorial library synthesis and industrial applications.
22. Wang, M.; Pan, D.; Zhang, Q.; Lei, Y.; Wang, C.; Jia, H.; Mou, L.; Miao, X.; Ren, X.; Xu, Z. *J. Am. Chem. Soc.* **2024,** *146,* 6675–6685. doi:10.1021/jacs.3c12879
23. Qi, R.; Chen, Q.; Liu, L.; Ma, Z.; Pan, D.; Wang, H.; Li, Z.; Wang, C.; Xu, Z. *Nat. Commun.* **2023,** *14,* 3295. doi:10.1038/s41467-023-38871-1
24. White, C. J.; Yudin, A. K. *Nat. Chem.* **2011,** *3,* 509–524. doi:10.1038/nchem.1062
25. Fagundez, C.; Sellanes, D.; Serra, G. *ACS Comb. Sci.* **2018,** *20,* 212–219. doi:10.1021/acscombsci.7b00154
26. Ma, B.; Litvinov, D. N.; He, L.; Banerjee, B.; Castle, S. L. *Angew. Chem., Int. Ed.* **2009,** *48,* 6104–6107. doi:10.1002/anie.200902425
27. Ranjbar, B.; Gill, P. *Chem. Biol. Drug Des.* **2009,** *74,* 101–120. doi:10.1111/j.1747-0285.2009.00847.x
28. Kuril, A. K.; Vashi, A.; Subbappa, P. K. *J. Pept. Sci.* **2025,** *31,* e3648. doi:10.1002/psc.3648
29. Fiametti, L. O.; Franco, C. A.; Nunes, L. O. C.; de Castro, L. M.; Santos-Filho, N. A. *Int. J. Mol. Sci.* **2025,** *26,* 2017. doi:10.3390/ijms26052017
30. Chan, L. Y.; Zhang, V. M.; Huang, Y.-h.; Waters, N. C.; Bansal, P. S.; Craik, D. J.; Daly, N. L. *ChemBioChem* **2013,** *14,* 617–624. doi:10.1002/cbic.201300034

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.21.202

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Synthesis of a HDAC inhibitor–nanogold probe for cryo-EM visualization in class I HDAC co-repressor complexes

Wiktoria A. Pytel[1,2], John W. R. Schwabe[*1,2] and James T. Hodgkinson[*1,3]

Full Research Paper

Open Access

Address:
[1]Institute for Structural and Chemical Biology, University of Leicester, Leicester LE1 7RH, UK, [2]Department of Molecular and Cell Biology, University of Leicester, Leicester LE1 7RH, UK and [3]School of Chemistry, University of Leicester, Leicester LE1 7RH, UK

Email:
John W. R. Schwabe[*] - john.schwabe@leicester.ac.uk;
James T. Hodgkinson[*] - jthodgkinson@le.ac.uk

* Corresponding author

Keywords:
CI-994; co-repressor complex; CoREST; cryo-EM; gold nanoparticle; HDAC

Beilstein J. Org. Chem. **2026,** *22,* 480–485.
https://doi.org/10.3762/bjoc.22.35

Received: 14 January 2026
Accepted: 03 March 2026
Published: 17 March 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2026 Pytel et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Class I histone deacetylases (HDACs 1–3) serve as catalytic subunits within seven multiprotein co-repressor complexes, each of which has distinct functions in the cell. We report the synthesis of a HDAC inhibitor–nanogold probe, derived from the class I HDAC inhibitor CI-994, for cryo-electron microscopy (cryo-EM) visualization of the HDAC catalytic domain within class I HDAC co-repressor complexes. The nanogold probe retained HDAC inhibitory activity comparable to CI-994 against the HDAC1-LSD1-CoREST complex in vitro. In cryo-EM studies, 2D class averages revealed the bi-lobed architecture of the CoREST complex and partial localization of the gold nanoparticle probe to the CoREST complex. However, the probe was not observed in classes showing the side-view of the CoREST complex, limiting unambiguous identification and positioning of the HDAC catalytic domain within the CoREST complex.

Introduction

Histone deacetylase (HDAC) enzymes catalyze the hydrolysis of acetyl groups from *N*-acetylated lysine residues in histone proteins. HDACs are also capable of the deacetylation of non-histone proteins [1], and the hydrolysis of other acyl functional groups [2]. The human genome encodes 18 histone deacetylases (HDACs), which are divided into two main groups based on their catalytic mechanisms [3]. Eleven HDACs are zinc-dependent enzymes, while the remaining seven, known as sirtuins (SIRT1–7), require nicotinamide adenine dinucleotide (NAD^+) as a cofactor [3]. HDACs are further classified into four classes: class I (HDAC1–3 and HDAC8), class IIa (HDAC4, 5, 7, and 9), class IIb (HDAC6 and 10), class IV (HDAC11), and the NAD^+-dependent sirtuins are grouped separately as class III [3].

HDAC1, HDAC2, and HDAC3 of the class I HDACs exist in multiprotein co-repressor complexes in vivo [4]. HDAC1 and HDAC2 exist interchangeably in the CoREST, MIDAC, SIN3, NuRD, MIER, and RERE complexes, while HDAC3 exists in

the SMRT/NCoR complex [4]. The protein complex partners govern the nucleosomal substrate specificity [5,6], and each complex has distinct cellular functions [7].

Cryo-electron microscopy (cryo-EM) has revolutionized structural biology by enabling high-resolution, three-dimensional visualization of macromolecular multiprotein complexes in differing functional states. Examples of structure elucidation utilizing cryo-EM for class I HDAC complexes include the MiDAC and SIN3 complexes [8,9]. However, despite these advances, obtaining high-resolution structures of flexible multiprotein complexes can still prove challenging. One such example of this includes the tripartite CoREST complex that encompasses HDAC1/2, the co-repressor of REST (CoREST) and the lysine-specific demethylase 1 (LSD1). Cryo-EM and small angle X-ray scattering revealed that the CoREST complex exists as a bi-lobed structure [10]. Enzyme kinetics studies showed that HDAC1 and LSD1 do not act independently and that their activities and modulation by inhibitors and activators are closely coupled. Both enzymes exist in at least two distinguishable states that differ in their kinetic properties, consistent with the two distinct structural states of the complex observed in the cryo-EM maps [10]. However, the intrinsic flexibility of the CoREST complex has limited the achievable resolution in cryo-EM reconstructions, making it difficult to confidently assign the orientation of the HDAC catalytic domain relative to LSD1 [10]. Understanding the spatial positioning of the HDAC active site within the complex is critical, as it may significantly influence how CoREST engages with nucleosomal substrates.

Nanogold particles (1–5 nm in diameter) are highly electron-dense and provide high contrast in microscopy. They have been successfully applied in immunogold electron microscopy to study protein localization within a cell, as well as in cellular cryo-electron tomography [11,12]. More recently, gold nanoparticles have been used as labeling tools in vitrified cells to locate cellular compartments in cellular cryo-EM [13]. Given the considerable utility of nanogold particles in microscopy, we aimed to synthesize a nanogold-conjugated HDAC inhibitor and evaluate its applicability in single-particle cryo-EM to unambiguously determine the positioning and orientation of the HDAC active site within the CoREST complex.

Results and Discussion

Design and synthesis of a HDAC inhibitor–nanogold probe

For the basis of the HDAC inhibitor–nanogold probe we utilized the class I HDAC inhibitor CI-994 (Figure 1). CI-994 is an inhibitor of HDAC1–3 and it inhibits HDAC1-CoREST-LSD1, HDAC2-CoREST-LSD1, and HDAC3-SMRT complex with IC_{50} values of 0.53 μM, 0.62 μM, and 0.14 μM, respectively [14,15]. Additionally, CI-994, like other benzamide HDAC inhibitors, exhibits slow on/off binding kinetics, hence once bound to the HDAC within the complex it should not readily dissociate [16]. A crystal structure of HDAC2 bound to an analogue of CI-994 (PDB: 4LY1) revealed that the acetamide moiety is oriented outside the HDAC active site [17]. Hence, we decided to functionalize this position with an alkyl linker consisting of 9 carbon atoms to mitigate any steric clashes between the HDAC inhibitor and the nanogold particle, which could be detrimental to the probe binding affinity (Figure 1). Further to this, we previously functionalized this position with linkers for the development of HDAC1–3 proteolysis targeting chimeras (PROTACs) [14,18]. Alkyl-linker lengths of approximately 12 atoms and greater were the most effective degraders [18]. We chose the commercially available amine functionalized nanogold particles (Au–NH_2, specifically Monoamino-Nanogold® 1.4 nm purchased from Nanoprobes) for our probe design. Au–NH_2 consists of a gold cluster of eleven gold atoms coordinated by trisarylphosphine ligands with a diameter of 1.4 nm. One of the trisarylphosphine ligands contains the 3-aminopropylamido group, allowing for stoichiometric conjugation with suitable substrates. The small size of Au–NH_2 minimizes steric hindrance and allows for enhanced spatial resolution, relative to colloidal nanogold particles,

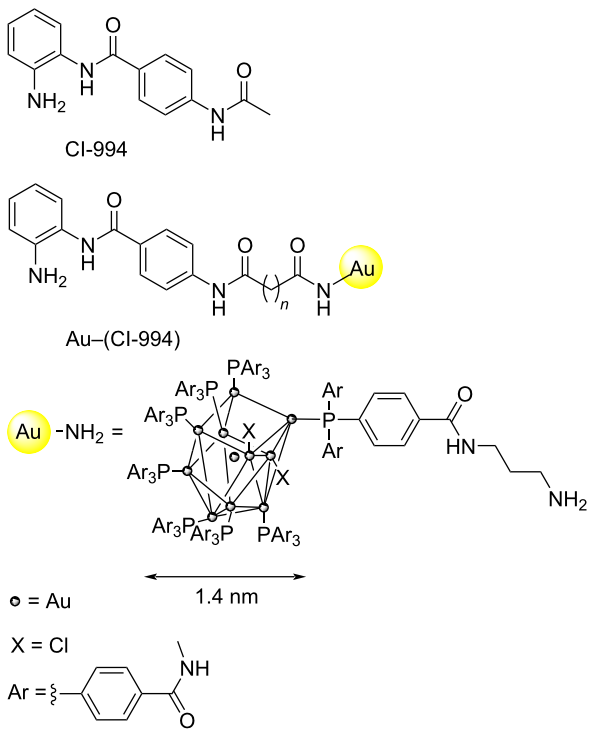
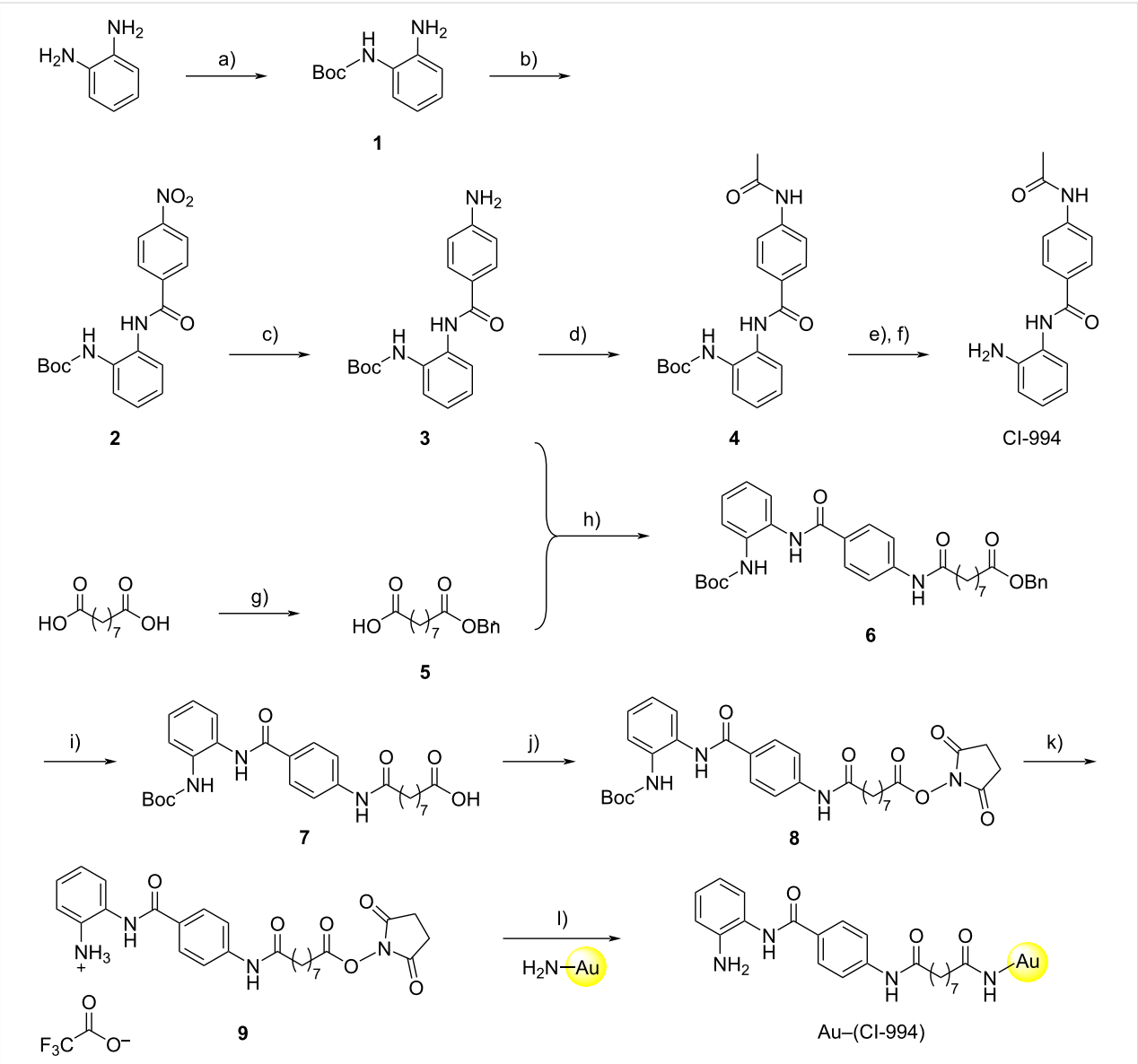


Figure 1: HDAC1–3 inhibitor CI-994; HDAC inhibitor–nanogold probe Au–(CI-994); structure of amine-functionalized 1.4 nm nanogold particle (Au–NH_2).

thereby facilitating high precision labeling of the HDAC active site.

CI-994 was synthesized using previously established routes (Scheme 1) [14,15,18]. Intermediates **5**–**7** were prepared in a manner analogous to Smalley et al. [14]. The first step in the linker synthesis for Au–(CI-994) involved a monoprotection of nonanedioic acid with a benzyl group to give **5** which proceeded in moderate yield due to the formation of the dibenzylated by-product. Compound **5** was then coupled to the CI-994 intermediate **3** via HATU-mediated amide bond formation to produce **6** in good yield. Removal of the benzyl protecting group was performed by catalytic hydrogenation and acid **7** was obtained in near quantitative yield. Intermediate **7** was converted to its corresponding NHS ester **8** and residual starting material was removed by column chromatography. The Boc-protecting group was removed under anhydrous conditions in good yield and the structure of **9** was confirmed by NMR spectroscopy. Compound **9** was maintained as the TFA salt for the final step, and was stable for days stored at −20 °C (stability determined by ^{1}H NMR, Figures S1–S4 in Supporting Information File 1).



Scheme 1: Synthesis of CI-994 and the Au–(CI-994) conjugate. Conditions: a) Boc_2O, NEt_3, THF, 0 °C to rt, 19 h, 70%; b) 4-nitrobenzoyl chloride, DIPEA, DCM, 0 °C to rt, 16 h, 74%; (c) H_2, 10% Pd/C, MeOH/THF 1:1, rt, 17 h, 95%; d) AcCl, NEt_3, THF, 0 °C to rt, 21 h, 67%; e) TFA, DCM, 0 °C to rt, 20 h; (f) MP-carbonate resin, MeOH, rt, 3 h, 95%; g) benzyl bromide, 1,4-dioxane/DMF 1:1, 90 °C, 19 h, 45%; h) **3**, **5**, HATU, DIPEA, DMF, 0 °C to rt, 40 h, 59%; i) H_2, 10% Pd/C, THF, rt, 19 h, 99%; j) *N*-hydroxysuccinimide, EDC, DMF, 0 °C to rt, 16 h, 90%; k) TFA, DCM, 0 °C to rt, 18 h, 92%; l) Au–NH_2, DMSO, rt, 20 h.

For the conjugation of **9** to Au–NH_2, an excess of **9** was used to drive the conjugation reaction to completion. Unreacted **9** was then removed by repeated washing with water and buffer, followed by centrifugation (see Supporting Information File 1 for full details). The resulting Au–(CI-994) conjugate was concentrated to a final volume of 1 mL. The concentration of Au–(CI-994) was determined by UV–vis spectroscopy, using absorbance at 420 nm – a characteristic wavelength for nanogold [19]. Based on the Beer–Lambert law, the concentration was calculated to be 2.69 μM. Attempts at further concentration led to precipitation, indicating limited solubility at higher concentrations.

The HDAC inhibitor–nanogold probe inhibits HDAC enzymatic activity in the CoREST complex

We first wanted to confirm that the conjugation of CI-994 to the nanogold particle did not significantly affect HDAC inhibition. The HDAC1-LSD1-CoREST complex, incorporating a FLAG tag in CoREST, was expressed and purified from HEK293F cells as previously reported [10]. Fluorescent deacetylase assays were carried out using Boc-Lys-(Ac)-AMC as the HDAC substrate [15,20], which on cleavage by HDAC releases fluorescent 7-amino-4-methylcoumarin. Au–(CI-994) and several controls were evaluated under the HDAC assay conditions in the presence and absence of the CoREST complex (Figure 2). In the absence of the CoREST complex Au–NH_2 and Au–(CI-994) did not affect fluorescence under the HDAC assay conditions, implying that the gold nanoparticle does not interfere with the HDAC assay conditions. As expected, 20 μM of CI-994 completely inhibited HDAC activity of the CoREST complex. Notably, Au–(CI-994) also exhibited near-complete inhibition of the HDAC activity in the CoREST complex, even at 0.54 μM. Surprisingly, Au–NH_2 was found to reduce the HDAC activity of the CoREST complex by nearly 50%. One plausible explanation for this effect is a direct interaction between the gold nanoparticles and solvent-accessible cysteine residues. Cysteines possess thiol side chains that exhibit strong affinity for gold, enabling them to form stable bonds with metal surfaces [21]. Such interactions may disrupt the native conformation of cysteine-containing proteins or peptides, potentially impairing the structural integrity of the CoREST complex and diminishing its deacetylase function. However, the maximal HDAC inhibition by Au–NH_2 was considerably less compared to Au–(CI-994) and CI-994, suggesting Au–(CI-994), is inhibiting HDAC enzymatic activity by direct competition for the HDAC active site. We next determined the IC_{50} of Au–(CI-994) against the CoREST complex. Au–(CI-994) exhibited an IC_{50} value of 0.19 ± 0.08 μM, this is directly comparable to the IC_{50} value for CI-994, IC_{50} = 0.53 μM for the CoREST complex [14,15]. We noted HDAC activity did not reach 0% with Au–(CI-994) which we speculate was due to precipitation of

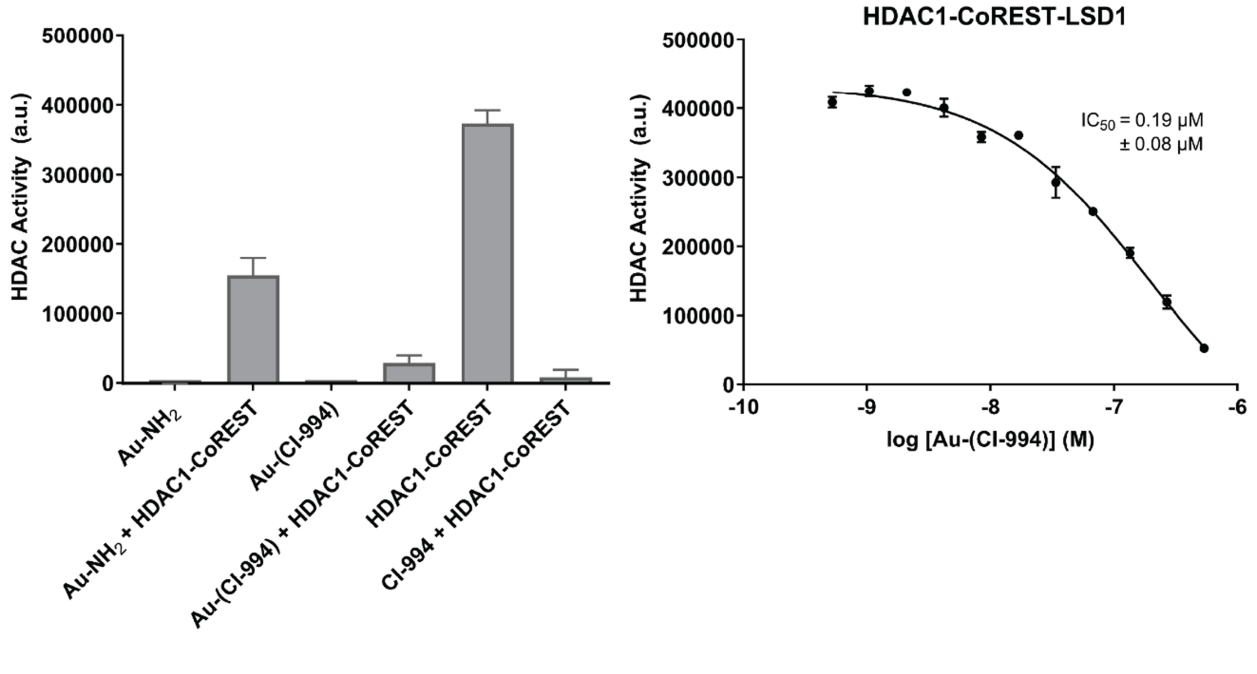


Figure 2: Left: Single point concentration HDAC assay. Concentration of Au–NH_2 and Au–(CI-994) = 0.54 μM. Concentration of CI-994 = 20 μM. Right: Dose response inhibition of HDAC activity in the CoREST complex with Au–(CI-994). Boc-(Ac)Lys-AMC substrate = 100 μM; CoREST complex = 12.5 nM. Error bars represent ± S.E.M. (n = 3). For full assay conditions see Supporting Information File 1.

Au–(CI-994) at higher concentrations. The Au–(CI-994) probe was further characterized using electron microscopy (Figure S5 in Supporting Information File 1). The electron-dense gold nanoparticles appeared prominently in the micrographs due to their intense black contrast. The particles exhibited uniform size distribution, were evenly dispersed across the grid, and showed no signs of aggregation. These studies confirmed the suitability of Au–(CI-994) probe for structural studies with the CoREST complex.

The HDAC inhibitor–nanogold probe localizes with the CoREST complex in cryo-EM

To stabilize the HDAC1-LSD1-CoREST complex on the EM grid, glutaraldehyde cross-linking was performed, with successful cross-linking confirmed by SDS-PAGE. The cross-linked ternary complex was incubated with the Au–(CI-994) probe for 2 hours in 2:1 molar ratio to minimize non-specific binding. Screening and data collection were carried out on a Titan Krios microscope. A range of protein particles resembling the CoREST ternary complex were observed, with several displaying a distinct "black dot" indicative of the electron-dense gold nanoparticle (Figure S6 in Supporting Information File 1). However, as expected, Au–(CI-994) was not observed in all protein particles, indicating that the CoREST complex was not fully saturated with the probe. Automated particle picking using TOPAZ identified 370,670 particles from 1,965 micrographs, averaging 189 particles per micrograph (Figure S7 in Supporting Information File 1). Both Au–(CI-994)-bound and unbound particles were picked and subsequently separated during data analysis. Particles were extracted and subjected to multiple rounds of 2D classification to remove artefacts and junk particles. The final 2D classes (Figure 3) revealed a bi-lobed architecture consistent with the published CoREST ternary complex structure [10]. The Au–(CI-994) probe was visible in some classes due to its distinct, brighter contrast. While various orientations of the complex were captured, the probe was absent in side-view classes. These views are especially important for visualizing the active sites of HDAC1 and LSD1, and thus the spatial relationship of the Au–(CI-994) probe to the LSD1 active site. Hence, although the probe localized to the CoREST complex, unfortunately, the positioning and spatial orientation of the HDAC catalytic domain could not be unambiguously determined.

Conclusion

In summary, we report the synthesis and validation of Au–(CI-994) as a nanogold-conjugated HDAC inhibitor for cryo-EM visualization in class I HDAC co-repressor complexes. Au–(CI-994) effectively inhibits HDAC activity within the CoREST complex in vitro, showing comparable potency to CI-994. The probe was clearly visualized in association with the CoREST complex by cryo-EM; however, its absence in side-view 2D classes prevented precise localization of the HDAC catalytic domain. We speculate that the flexibility of the 9-carbon alkyl linker, and additional flexible linker to the nanogold particles, may have resulted in signal averaging, obscuring the probe's exact position. Furthermore, although probe localization was evident, complete saturation of the CoREST complex with the probe was not achieved. Future studies will focus on rigidifying the linker and enhancing HDAC binding affinity. Nonetheless, nanogold-labeled HDAC inhibitors could serve as effective fiducial markers in cryo-EM, facilitating the localization of distinct subunits or binding sites within large and flexible multiprotein complexes.

Supporting Information

Supporting Information File 1

Chemical protocols and characterization data for compounds, biological protocols including cryo-EM grid prep, data collection, and images of micrographs.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-35-S1.pdf]

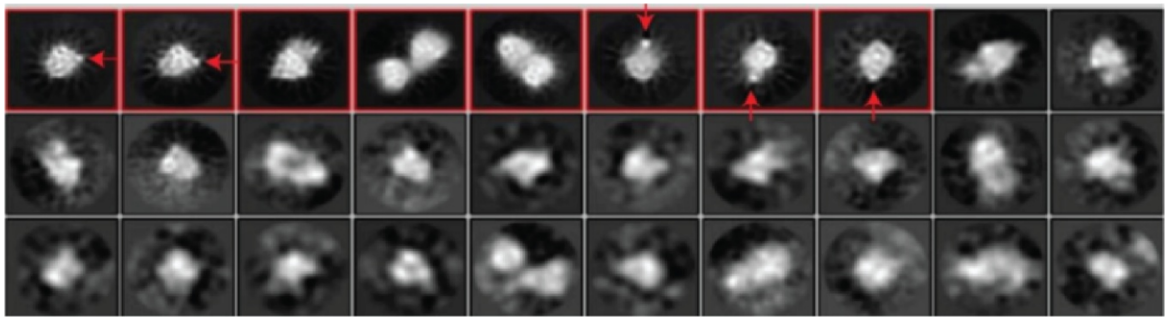


Figure 3: Final round of 2D classification of the cross-linked CoREST complex in the presence of Au–(CI-994) (87,649 particles). Red arrows indicate the position of the Au–(CI-994) probe localized with the CoREST complex. The probe was absent in side-view classes.

Acknowledgements

We thank Dr. Christos Savva for assistance and guidance in Cryo-EM experiments. We thank Dr. Vanessa M. Timmermann and Dr. Rebecca Hawker for assistance in NMR analysis and Dr. Sharad C. Mistry for assistance in mass spectrometry.

Funding

We thank the following funders: EPSRC (EP/S030492/1 to JTH; EP/W02151X/1 NMR facility); MRC (MRC MR/Z506059/1 to JTH and JWRS; MC_PC_17136 Cryo-EM facility).

ORCID® iDs

James T. Hodgkinson - https://orcid.org/0000-0001-9978-7322

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Glozak, M. A.; Sengupta, N.; Zhang, X.; Seto, E. *Gene* **2005,** *363,* 15–23. doi:10.1016/j.gene.2005.09.010
2. Kutil, Z.; Novakova, Z.; Meleshin, M.; Mikesova, J.; Schutkowski, M.; Barinka, C. *ACS Chem. Biol.* **2018,** *13,* 685–693. doi:10.1021/acschembio.7b00942
3. Asmamaw, M. D.; He, A.; Zhang, L.-R.; Liu, H.-M.; Gao, Y. *Biochim. Biophys. Acta, Rev. Cancer* **2024,** *1879,* 189150. doi:10.1016/j.bbcan.2024.189150
4. Millard, C. J.; Watson, P. J.; Fairall, L.; Schwabe, J. W. R. *Trends Pharmacol. Sci.* **2017,** *38,* 363–377. doi:10.1016/j.tips.2016.12.006
5. You, S.-H.; Lim, H.-W.; Sun, Z.; Broache, M.; Won, K.-J.; Lazar, M. A. *Nat. Struct. Mol. Biol.* **2013,** *20,* 182–187. doi:10.1038/nsmb.2476
6. Wang, Z. A.; Millard, C. J.; Lin, C.-L.; Gurnett, J. E.; Wu, M.; Lee, K.; Fairall, L.; Schwabe, J. W. R.; Cole, P. A. *eLife* **2020,** *9,* e57663. doi:10.7554/elife.57663
7. Hayakawa, T.; Nakayama, J.-i. *J. Biomed. Biotechnol.* **2011,** 129383. doi:10.1155/2011/129383
8. Guo, Z.; Chu, C.; Lu, Y.; Zhang, X.; Xiao, Y.; Wu, M.; Gao, S.; Wong, C. C. L.; Zhan, X.; Wang, C. *Nat. Struct. Mol. Biol.* **2023,** *30,* 753–760. doi:10.1038/s41594-023-00975-z
9. Turnbull, R. E.; Fairall, L.; Saleh, A.; Kelsall, E.; Morris, K. L.; Ragan, T. J.; Savva, C. G.; Chandru, A.; Millard, C. J.; Makarova, O. V.; Smith, C. J.; Roseman, A. M.; Fry, A. M.; Cowley, S. M.; Schwabe, J. W. R. *Nat. Commun.* **2020,** *11,* 3252. doi:10.1038/s41467-020-17078-8
10. Song, Y.; Dagil, L.; Fairall, L.; Robertson, N.; Wu, M.; Ragan, T. J.; Savva, C. G.; Saleh, A.; Morone, N.; Kunze, M. B. A.; Jamieson, A. G.; Cole, P. A.; Hansen, D. F.; Schwabe, J. W. R. *Cell Rep.* **2020,** *30,* 2699–2711.e8. doi:10.1016/j.celrep.2020.01.091
11. Dahan, I.; Sorrentino, S.; Boujemaa-Paterski, R.; Medalia, O. *Structure* **2018,** *26,* 1408–1413.e3. doi:10.1016/j.str.2018.06.009
12. Melo, R. C. N.; Morgan, E.; Monahan-Earley, R.; Dvorak, A. M.; Weller, P. F. *Nat. Protoc.* **2014,** *9,* 2382–2394. doi:10.1038/nprot.2014.163
13. Groysbeck, N.; Hanss, V.; Donzeau, M.; Strub, J.-M.; Cianférani, S.; Spehner, D.; Bahri, M.; Ersen, O.; Eltsov, M.; Schultz, P.; Zuber, G. *Small Methods* **2023,** *7,* 2300098. doi:10.1002/smtd.202300098
14. Smalley, J. P.; Baker, I. M.; Pytel, W. A.; Lin, L.-Y.; Bowman, K. J.; Schwabe, J. W. R.; Cowley, S. M.; Hodgkinson, J. T. *J. Med. Chem.* **2022,** *65,* 5642–5659. doi:10.1021/acs.jmedchem.1c02179
15. Pytel, W. A.; Patel, U.; Smalley, J. P.; Millard, C. J.; Brown, E. A.; Pavan, A. R.; Wang, S.; Kalin, J. H.; dos Santos, J. L.; Cole, P. A.; Hodgkinson, J. T.; Schwabe, J. W. R. *J. Am. Chem. Soc.* **2025,** *147,* 36044–36052. doi:10.1021/jacs.5c08929
16. Becher, I.; Dittmann, A.; Savitski, M. M.; Hopf, C.; Drewes, G.; Bantscheff, M. *ACS Chem. Biol.* **2014,** *9,* 1736–1746. doi:10.1021/cb500235n
17. Lauffer, B. E. L.; Mintzer, R.; Fong, R.; Mukund, S.; Tam, C.; Zilberleyb, I.; Flicke, B.; Ritscher, A.; Fedorowicz, G.; Vallero, R.; Ortwine, D. F.; Gunzner, J.; Modrusan, Z.; Neumann, L.; Koth, C. M.; Lupardus, P. J.; Kaminker, J. S.; Heise, C. E.; Steiner, P. *J. Biol. Chem.* **2013,** *288,* 26926–26943. doi:10.1074/jbc.m113.490706
18. Smalley, J. P.; Adams, G. E.; Millard, C. J.; Song, Y.; Norris, J. K. S.; Schwabe, J. W. R.; Cowley, S. M.; Hodgkinson, J. T. *Chem. Commun.* **2020,** *56,* 4476–4479. doi:10.1039/d0cc01485k
19. Hainfeld, J. F.; Powell, R. D. *J. Histochem. Cytochem.* **2000,** *48,* 471–480. doi:10.1177/002215540004800404
20. Waldecker, M.; Kautenburger, T.; Daumann, H.; Busch, C.; Schrenk, D. *J. Nutr. Biochem.* **2008,** *19,* 587–593. doi:10.1016/j.jnutbio.2007.08.002
21. Sasaki, Y. C.; Yasuda, K.; Suzuki, Y.; Ishibashi, T.; Satoh, I.; Fujiki, Y.; Ishiwata, S. *Biophys. J.* **1997,** *72,* 1842–1848. doi:10.1016/s0006-3495(97)78830-1

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.35

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Design and synthesis of an erdafitinib-based selective FGFR2 degrader

Yumeng Jin, Shidong Wang, Sihan Pan, Shuqi Huang, Weichen Zhou, Xiaohao Huang, Lei Zheng* and Lingfeng Chen*

Full Research Paper

Open Access

Address:
School of Pharmacy, Hangzhou Medical College, Hangzhou 310012 Zhejiang, China

Email:
Lei Zheng* - 18267735533@163.com; Lingfeng Chen* - lfchen@hmc.edu.cn

* Corresponding author

Keywords:
CRBN; erdafitinib; FGFR2; selective degrader

Beilstein J. Org. Chem. **2026,** *22,* 583–591.
https://doi.org/10.3762/bjoc.22.44

Received: 14 January 2026
Accepted: 31 March 2026
Published: 15 April 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2026 Jin et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

This study aimed to develop a novel degrader capable of selectively degrading fibroblast growth factor receptor 2 (FGFR2) to overcome the issues of drug resistance and adverse reactions associated with traditional inhibitors in the treatment of FGFR2-driven tumors. Erdafitinib was employed as the targeting ligand, and its aliphatic amine site was conjugated with a CRBN E3 ligase ligand to design and synthesize a series of PROTAC molecules with different linkers. Screening was performed in KATO III cells with high FGFR2 expression, leading to the identification of LC-JD-6 as a potent degrader. Experimental results demonstrated that LC-JD-6 effectively induced FGFR2 protein degradation with a half-maximal degrading concentration (DC_{50}) of 121.4 nM, and this effect exhibited time- and concentration-dependence. Assessed at the cellular level, LC-JD-6 has a half-maximal inhibitory concentration in the KATO III (IC_{50}) of 96.0 nM and showed low inhibitory activity in normal cells. Selectivity analysis revealed that LC-JD-6 specifically degraded FGFR2 with minimal impact on other FGFR subtypes. Further studies confirmed that LC-JD-6 also efficiently reduced the expression of FGFR2 on the cell membrane surface. In conclusion, this study successfully developed LC-JD-6, a novel FGFR2-selective degrader, and for the first time confirmed its ability to degrade the membrane-bound form of FGFR2. This work provides an innovative targeted protein degradation strategy for the treatment of FGFR2-driven tumors and holds significant potential for clinical application.

Introduction

Fibroblast growth factor receptors (FGFR) are a family of single-pass transmembrane receptor tyrosine kinases (RTKs) localized on the cell surface that bind to fibroblast growth factors [1-3]. Dimerization and autophosphorylation of FGFRs are induced by their binding to ligands, which trigger intracellular signaling cascades to activate downstream substrates and

pathways [4]. This process involves PLC γ-mediated PKC activation and pFRS2-induced activation of the PI3K/AKT and MAPK/ERK pathways, thereby regulating cell proliferation, differentiation, migration, and survival [5-8]. FGF/FGFR dysregulation is associated with various diseases, including cancers, skeletal disorders, and metabolic syndromes [9,10]. Among family members, tumorigenesis driven by FGFR2 activation is attributed to genetic amplification, mutations, and gene rearrangements or fusions [11-13]. Genetic amplification of FGFR2 leads to an excessive number of gene copies, resulting in over-expression of the FGFR2 protein on the cell surface. This over-abundance can cause the receptor to be constitutively active, continuously sending growth signals to the cell even without proper ligand binding, thus promoting uncontrolled cell proliferation [11,14]. Mutations in FGFR2 may alter its structure, enabling it to bypass normal regulatory mechanisms and activate downstream oncogenic pathways independently of external stimuli, which also contributes to tumor development. Gene rearrangements or fusions involving FGFR2 can create novel chimeric proteins with enhanced or aberrant signaling capabilities, further driving cells towards a malignant phenotype [13,15]. Notably, FGFR2 amplification and fusions are frequently detected in patients with advanced gastric cancer and cholangiocarcinoma [16-18]. Accordingly, research into therapeutics targeting FGFR2 has emerged as a major focus.

To treat FGFR2-driven tumors, current research primarily focuses on two distinct classes of inhibitors, namely FGFR pan-inhibitors and FGFR2-selective inhibitors. Among these in Figure 1, futibatinib, infigratinib, and erdafitinib, as pan-FGFR inhibitors, exhibit therapeutic efficacy against tumors driven by FGFR2. For advanced cholangiocarcinoma driven by FGFR2, infigratinib and futibatinib are targeted drugs specifically approved for this indication [19]. In contrast, erdafitinib is primarily effective in the treating of urothelial carcinoma with FGFR2/3 alterations [20]. However, due to the risk of adverse effects associated with pan-inhibitors, an increasing number of FGFR2-selective inhibitors have been reported in recent years [21-25]. Lirafugratinib (RLY-4008), an FGFR2-selective inhibitor, is the first highly selective FGFR2 inhibitor. Although substantial advancements have been made in developing FGFR2-targeted therapeutics, the sustained clinical efficacy of these inhibitors in oncology remains constrained by the emergence of acquired resistance mechanisms.

Proteolysis-targeting chimera (PROTAC) is a chemical molecule which induce the target protein to approach the ubiquitin protein through the ubiquitin proteasome system, then it can be ubiquitinated and degraded [26-28]. This drives to form degraders through the unique properties of their own degradation functions and overcome a series of critical issues in cancer treatment.

The numerous advantages of PROTAC technology in targeted therapy have prompted us to conduct further research into FGFR2 degraders. In this work, a series of degraders conjugating erdafitinib with a CRBN binder was synthesized and screened, leading to the identification of LC-JD-6, a potent and FGFR2-selective degrader. This study establishes a fundamental framework for the rational design of therapeutic modalities targeting FGFR2 degradation, offering novel insights into protein homeostasis-based cancer intervention.

Results and Discussion

FGFR2-targeted degrader design

PROTAC molecules are composed of three key elements: a specific ligand targeting the protein of interest (POI), a recruiter for an E3 ubiquitin ligase, and a linker that covalently connects these two functional moieties into a single molecular entity [29]. Herein, we purpose the clinically validated inhibitor erdafitinib as a POI binder that selectively targets FGFR2. Based on the analysis of the co-crystal structure of erdafitinib bound to the FGFR kinase domain (PDB: 5EW8, https://doi.org/10.2210/pdb5EW8/pdb) [30], we observed that the aliphatic amine group is positioned in the solvent-exposed region of the molecule. Given this observation, we selected the aliphatic amine group as a suitable conjugation site (Figure 2a). Previous studies have shown that different linkers exhibit distinct selectivity profiles [31]. Therefore, in this study, novel FGFR2 degraders were designed by conjugating the aliphatic amine in erdafitinib to two sites on CRBN (Figure 2b and Scheme 1), and the length and type of the linker were modified to achieve the selective degradation of FGFR2. Compound **5** was prepared according to the previously reported procedure [32,33], and then condensed with the reported compound 6 [31] under EDCI to afford the final novel FGFR2 degrader.

Screening and identification of LC-JD-6 as a potent FGFR2 degrader

Targeted protein degradation molecules in degrading FGFR2 were tested in KATO III cells (high basal FGFR2), identifying LC-JD-6 as the optimal candidate. Western blot analysis showed that LC-JD-6 reduced FGFR2 by 90% after 12 hours at 500 nM (Figure 3a). However, for structural modifications, neither long-chain alkyl linkers (LC-JD-1/2) nor PEG-based linkers (LC-JD-3/4) demonstrated significant degradation of FGFR2 protein levels in KATO III gastric cancer cells. Although LC-JD-5, which is the homoisomer of LC-JD-6, is still lacking compared to LC-JD-6. To further verify the potency, LC-JD-6 was prescribed at an extensive concentration range (0.4–10,00 nM) in KATO III cells after 6 hours, with the DC_{50}

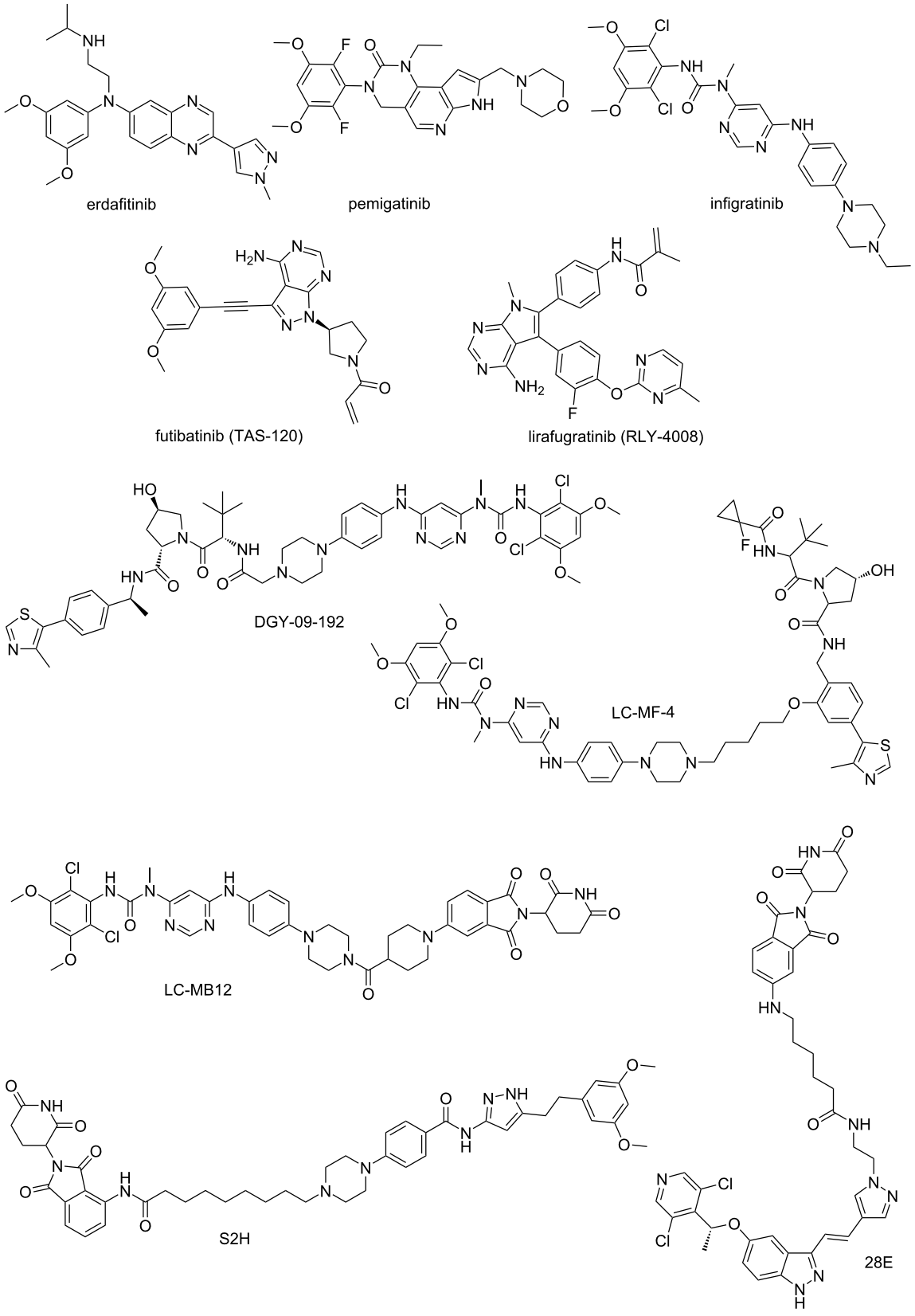


Figure 1: Chemical structures of severel FGFR inhibitors and degraders.

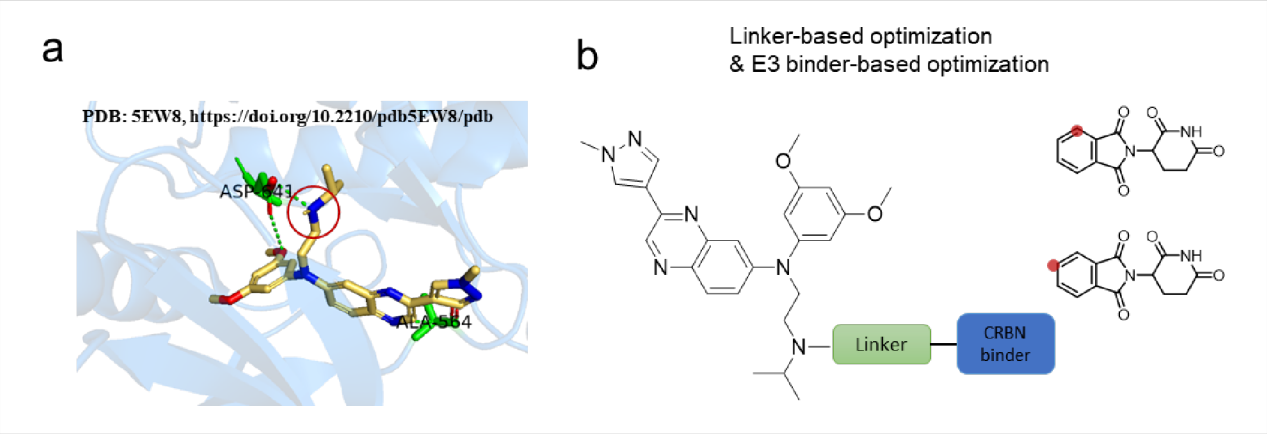


Figure 2: Rationale of the FGFR2 degrader design. a) Structure of the erdafitinib–FGFR complex. The figure was generated by PyMOL (PDB: 5EW8, https://doi.org/10.2210/pdb5EW8/pdb) [30]. b) The FGFR2 design strategy.

of 121.4 nM and a maximal degradation (D_{max}) after 12 hours of treatment (Figure 3b and Figure 3d). As shown in Figure 3e, LC-JD-6 have potency in the KATO III proliferation assay with an IC_{50} of 96.0 nM and showed no significant inhibitory activity in HEK293T cells with low FGFR2 expression. Those results showed LC-JD-6 is indeed a potent degrader.

LC-JD-6 as a selective degrader for FGFR2

Pharmacologically, the capacity of PROTACs to discriminate between intended and unintended targets represents a pivotal attribute that mitigates adverse off-target effects. Therefore, LC-JD-6 underwent in vitro profiling against diverse FGFR subtypes to evaluate its potential as a selective degrader. The selective induction of FGFR2 degradation by LC-JD-6 was demonstrated in cell lines endogenously expressing four FGFR subtypes (NCI-H1581, KATO III, RT112, and Hep3B). As shown in Figure 4a, treatment with 100 nM LC-JD-6 for 12 hours resulted in a reduction of over 80% in FGFR2 protein levels, with negligible effects on the other subtypes. These findings establish LC-JD-6 as a highly selective FGFR2 degrader.

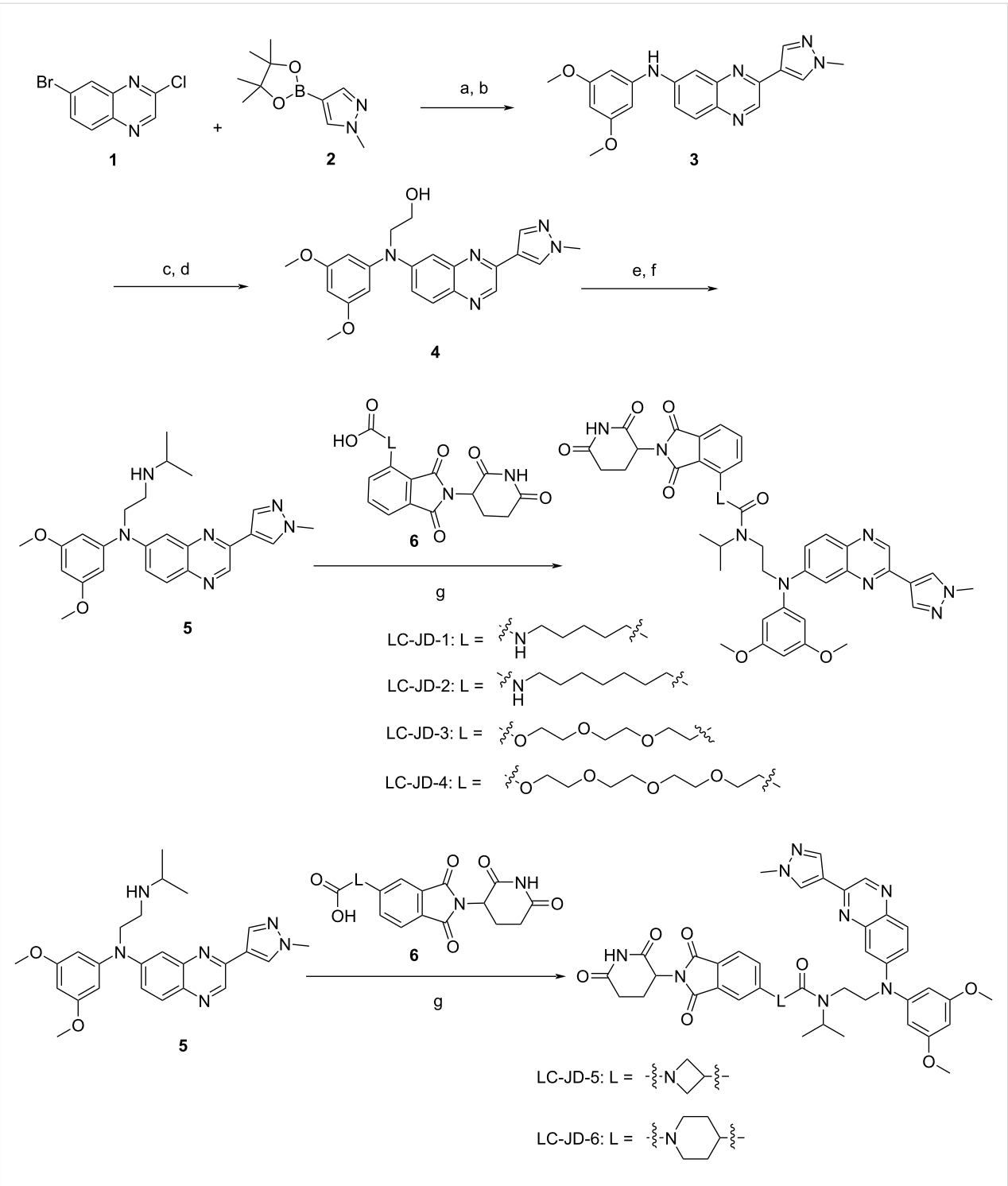
LC-JD-6 reduced the expression of membrane-bound FGFR2

Since most PROTACs reported to date act on intracellular targets, we examined whether LC-JD-6 could directly degrade FGFR2 from the plasma membrane. KATO III gastric cancer cells were incubated with 500 nM LC-JD-6 for 12 hours, followed by flow cytometric quantification of cell surface FGFR2 expression. As shown in Figure 4b and Figure 4c, LC-JD-6 treatment led to a marked decrease in cell surface FGFR2 expression, supporting its role as a membrane-specific degrader.

Conclusion

Over the past decade, with the identification of FGFR2 as a key oncogenic driver in multiple malignancies [34], significant progress has been made in elucidating the molecular mechanisms underlying the FGFR2 signaling pathway [35,36]. These advances have driven the development of several selective FGFR2 inhibitors, including erdafitinib, infiglatinib, and FGF401 [37-39]. However, issues such as acquired drug resistance and on-target toxicity emerging in clinical applications have limited the long-term efficacy of such inhibitors, and some of these agents remain in clinical development or have been discontinued [21,22]. Against this backdrop, the PROTAC technology has offered a novel strategy to overcome these challenges. Previously reported degraders, such as LC-MB12 [28], have been shown to effectively degrade FGFR2 via the proteasomal pathway, while the VHL E3 ligase-based degrader DGY-09-192 is capable of degrading both FGFR1 and FGFR2 simultaneously [40]. These findings suggest that different E3 ligase ligands may influence the selectivity and efficiency of degradation [41]. Nevertheless, due to limitations in experimental conditions, the specific regulatory mechanisms by which E3 ligase components modulate receptor selectivity remain to be further elucidated.

In general, the PROTAC technology has opened up a new avenue for addressing the drug resistance of traditional inhibitors by harnessing the intracellular endogenous protein degradation machinery. However, this technology still faces challenges in degrading membrane proteins such as FGFR2, partly due to the cell-type-specific expression differences of E3 ligases. Through systematic medicinal chemistry design and pharmacodynamic evaluation, this study successfully developed LC-JD-6, a CRBN E3 ligase-based PROTAC molecule that potently and selectively degrades FGFR2. Experiments demonstrated that LC-JD-6 exhibited a DC_{50} of 121.4 nM in KATO III cells, degraded more than 90% of FGFR2 protein within 12 hours, and displayed favorable subtype selectivity and the ability to degrade membrane-bound



Scheme 1: Synthesis of PROTACs towards FGFR2. Reagents and conditions: (a) K_2CO_3, Pd (dppf)Cl_2, 1,4-dioxane/H_2O 4:1, 100 °C, 5 h; (b) 3,5-dimethoxyaniline, $Pd_2(dba)_3$, BINAP, Cs_2CO_3, toluene, 100 °C, 12 h; (c) (2-bromoethoxy)-*tert*-butyldimethylsilane, NaH, DMF, rt, 12 h; (d) tetrabutylammonium fluoride, THF rt, 12 h; (e) methanesulfonic anhydride, TEA, DCM, rt, 3 h; (f) isopropylamine, DIPEA, MeCN, 100 °C, 12 h; (g) EDCI, HOBt, DIPEA, DMF, rt, 12 h.

FGFR2. This molecule successfully repurposed erdafitinib, a broad-spectrum FGFR inhibitor, into a highly selective degrader. It not only provides a new chemical tool for investigating FGFR2 function but also lays a crucial foundation for the development of novel therapeutic strategies against FGFR2-driven tumors.

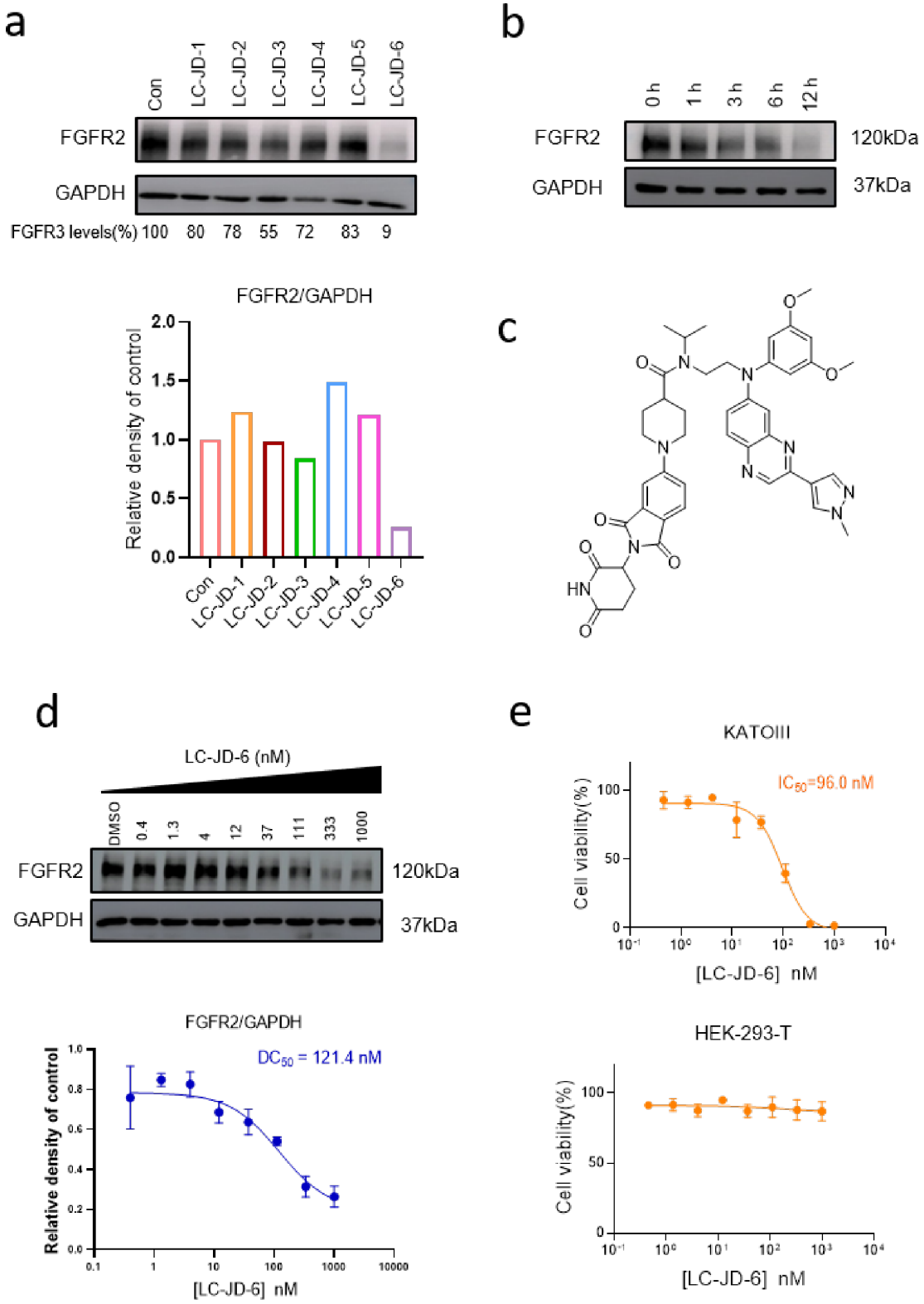


Figure 3: a) Representative western blots evaluating the total FGFR2 levels in KATO III cells following treatment using the indicated PROTAC. b) Time-course of FGFR2 degradation. c) Chemical structure of LC-JD-6. d,) Dose-course of FGFR2 degradation. e) Cell viability in KATO III and HEK293T cells.

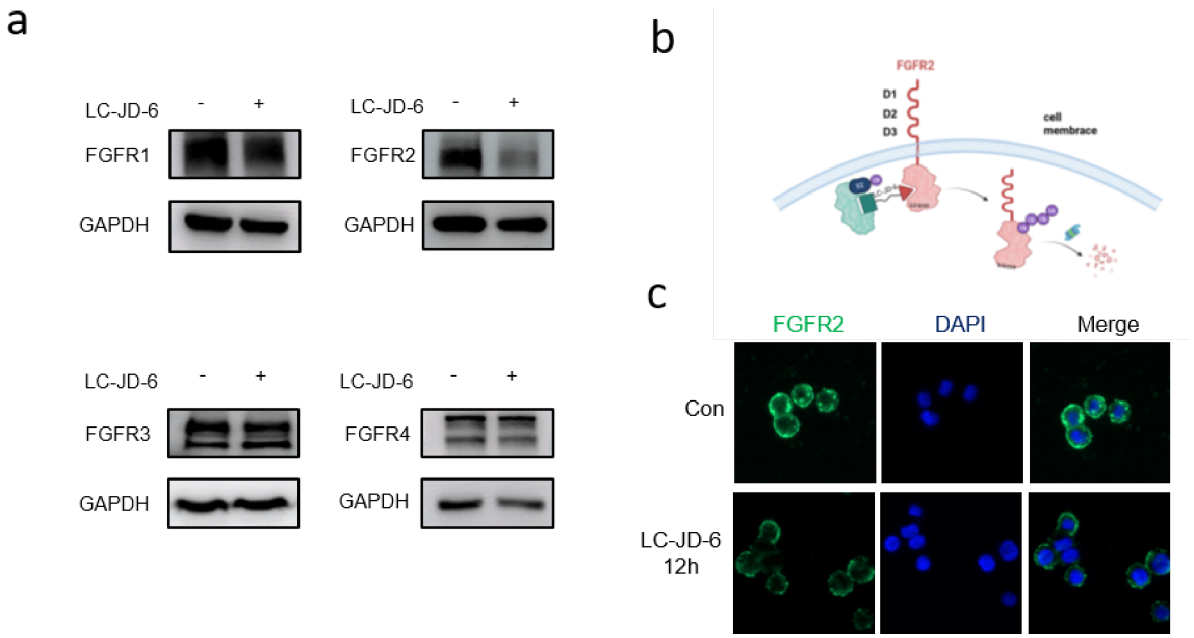


Figure 4: a) FGFR1, FGFR2, FGFR3, and FGFR4 levels in cells after treatment. b) The mechanism of PROTAC. It was created in BioRender (Chen, Lingfeng https://BioRender.com/7td2yot). This content is not subject to CC BY 4.0. c) Cellular localization of FGFR2 after treatment with LC-JD-6.

Supporting Information

Supporting Information File 1

Experimental details and spectral data for all compounds.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-44-S1.pdf]

Acknowledgements

The graphical abstract was created in BioRender (Chen, Lingfeng https://BioRender.com/7td2yot). This content is not subject to CC BY 4.0.

Funding

This work was supported by the Natural Science Funding of Zhejiang Province (LDG25H300001 and LR24H300001 to L.C.), and Zhejiang Medical and Health Science Project (2023KY622 to L.Z.).

Author Contributions

Yumeng Jin: investigation; validation; writing – original draft. Shidong Wang: validation; writing – review & editing. Sihan Pan: formal analysis; software; supervision. Shuqi Huang: formal analysis; software; validation. Weichen Zhou: resources; validation. Xiaohao Huang: methodology; validation; visualization. Lei Zheng: data curation; funding acquisition; investigation. Lingfeng Chen: conceptualization; data curation; methodology; resources; writing – original draft.

ORCID® iDs

Lingfeng Chen - https://orcid.org/0000-0003-0089-6559

Data Availability Statement

Data generated and analyzed during this study is available from the corresponding author upon reasonable request.

References

1. Beenken, A.; Mohammadi, M. *Nat. Rev. Drug Discovery* **2009,** *8,* 235–253. doi:10.1038/nrd2792
2. Itoh, N.; Ornitz, D. M. *J. Biochem.* **2011,** *149,* 121–130. doi:10.1093/jb/mvq121
3. Katoh, M.; Nakagama, H. *Med. Res. Rev.* **2014,** *34,* 280–300. doi:10.1002/med.21288
4. Chen, L.; Marsiglia, W. M.; Chen, H.; Katigbak, J.; Erdjument-Bromage, H.; Kemble, D. J.; Fu, L.; Ma, J.; Sun, G.; Zhang, Y.; Liang, G.; Neubert, T. A.; Li, X.; Traaseth, N. J.; Mohammadi, M. *Nat. Chem. Biol.* **2020,** *16,* 267–277. doi:10.1038/s41589-019-0455-7
5. Eswarakumar, V. P.; Lax, I.; Schlessinger, J. *Cytokine Growth Factor Rev.* **2005,** *16,* 139–149. doi:10.1016/j.cytogfr.2005.01.001
6. Babina, I. S.; Turner, N. C. *Nat. Rev. Cancer* **2017,** *17,* 318–332. doi:10.1038/nrc.2017.8
7. Faure, M.; Hollebecque, A.; Bahleda, R.; Soria, J.-C.; Dyevre, V. *Ann. Oncol.* **2017,** *28,* v575. doi:10.1093/annonc/mdx390.006

8. Katoh, M. *Trends Pharmacol. Sci.* **2016,** *37,* 1081–1096. doi:10.1016/j.tips.2016.10.003
9. Scheffler, M.; Chanra, T.; Kron, A.; Koleczko, S.; Abdulla, D. S. Y.; Ihle, M. A.; Holzem, A.; Riedel, R.; Michels, S.; Fischer, R. N.; Merkelbach-Bruse, S.; Büttner, R.; Nogova, L.; Wolf, J. *Ann. Oncol.* **2019,** *30,* v640. doi:10.1093/annonc/mdz260.076
10. Jogo, T.; Nakamura, Y.; Shitara, K.; Bando, H.; Yasui, H.; Esaki, T.; Terazawa, T.; Satoh, T.; Shinozaki, E.; Nishina, T.; Sunakawa, Y.; Komatsu, Y.; Hara, H.; Oki, E.; Matsuhashi, N.; Ohta, T.; Kato, T.; Ohtsubo, K.; Kawakami, T.; Okano, N.; Yamamoto, Y.; Yamada, T.; Tsuji, A.; Odegaard, J. I.; Taniguchi, H.; Doi, T.; Fujii, S.; Yoshino, T. *Clin. Cancer Res.* **2021,** *27,* 5619–5627. doi:10.1158/1078-0432.ccr-21-1414
11. Helsten, T.; Elkin, S.; Arthur, E.; Tomson, B. N.; Carter, J.; Kurzrock, R. *Clin. Cancer Res.* **2016,** *22,* 259–267. doi:10.1158/1078-0432.ccr-14-3212
12. Hudkins, R. L.; Allen, E.; Balcer, A.; Hoffman, I. D.; Iyer, S.; Neal, M.; Nelson, K. J.; Rideout, M.; Ye, Q.; Starrett, J. H.; Patel, P.; Harris, T.; Swanson, R. V.; Bensen, D. C. *J. Med. Chem.* **2024,** *67,* 16737–16756. doi:10.1021/acs.jmedchem.4c01531
13. Sia, D.; Losic, B.; Moeini, A.; Cabellos, L.; Hao, K.; Revill, K.; Bonal, D.; Miltiadous, O.; Zhang, Z.; Hoshida, Y.; Cornella, H.; Castillo-Martin, M.; Pinyol, R.; Kasai, Y.; Roayaie, S.; Thung, S. N.; Fuster, J.; Schwartz, M. E.; Waxman, S.; Cordon-Cardo, C.; Schadt, E.; Mazzaferro, V.; Llovet, J. M. *Nat. Commun.* **2015,** *6,* 6087. doi:10.1038/ncomms7087
14. Touat, M.; Ileana, E.; Postel-Vinay, S.; André, F.; Soria, J.-C. *Clin. Cancer Res.* **2015,** *21,* 2684–2694. doi:10.1158/1078-0432.ccr-14-2329
15. Lamarca, A.; Barriuso, J.; McNamara, M. G.; Valle, J. W. *J. Hepatol.* **2020,** *73,* 170–185. doi:10.1016/j.jhep.2020.03.007
16. Lin, C.-C.; Melo, F. A.; Ghosh, R.; Suen, K. M.; Stagg, L. J.; Kirkpatrick, J.; Arold, S. T.; Ahmed, Z.; Ladbury, J. E. *Cell* **2012,** *149,* 1514–1524. doi:10.1016/j.cell.2012.04.033
17. Ahmed, Z.; Lin, C.-C.; Suen, K. M.; Melo, F. A.; Levitt, J. A.; Suhling, K.; Ladbury, J. E. *J. Cell Biol.* **2013,** *200,* 493–504. doi:10.1083/jcb.201204106
18. Saborowski, A.; Lehmann, U.; Vogel, A. *Ther. Adv. Med. Oncol.* **2020,** *12,* 1758835920953293. doi:10.1177/1758835920953293
19. Loriot, Y.; Matsubara, N.; Park, S. H.; Huddart, R. A.; Burgess, E. F.; Houede, N.; Banek, S.; Guadalupi, V.; Ku, J. H.; Valderrama, B. P.; Tran, B.; Triantos, S.; Kean, Y.; Akapame, S.; Deprince, K.; Mukhopadhyay, S.; Stone, N. L.; Siefker-Radtke, A. O. *N. Engl. J. Med.* **2023,** *389,* 1961–1971. doi:10.1056/nejmoa2308849
20. Loriot, Y.; Necchi, A.; Park, S. H.; Garcia-Donas, J.; Huddart, R.; Burgess, E.; Fleming, M.; Rezazadeh, A.; Mellado, B.; Varlamov, S.; Joshi, M.; Duran, I.; Tagawa, S. T.; Zakharia, Y.; Zhong, B.; Stuyckens, K.; Santiago-Walker, A.; De Porre, P.; O'Hagan, A.; Avadhani, A.; Siefker-Radtke, A. O. *N. Engl. J. Med.* **2019,** *381,* 338–348. doi:10.1056/nejmoa1817323
21. Subbiah, V.; Verstovsek, S. *Cell Rep. Med.* **2023,** *4,* 101204. doi:10.1016/j.xcrm.2023.101204
22. Kommalapati, A.; Tella, S. H.; Borad, M.; Javle, M.; Mahipal, A. *Cancers* **2021,** *13,* 2968. doi:10.3390/cancers13122968
23. Mahipal, A.; Tella, S. H.; Kommalapati, A.; Yu, J.; Kim, R. *Crit. Rev. Oncol. Hematol.* **2020,** *155,* 103091. doi:10.1016/j.critrevonc.2020.103091
24. Facchinetti, F.; Hollebecque, A.; Braye, F.; Vasseur, D.; Pradat, Y.; Bahleda, R.; Pobel, C.; Bigot, L.; Déas, O.; Florez Arango, J. D.; Guaitoli, G.; Mizuta, H.; Combarel, D.; Tselikas, L.; Michiels, S.; Nikolaev, S. I.; Scoazec, J.-Y.; Ponce-Aix, S.; Besse, B.; Olaussen, K. A.; Loriot, Y.; Friboulet, L. *Cancer Discovery* **2023,** *13,* 1998–2011. doi:10.1158/2159-8290.cd-22-1441
25. Katoh, M.; Loriot, Y.; Brandi, G.; Tavolari, S.; Wainberg, Z. A.; Katoh, M. *Nat. Rev. Clin. Oncol.* **2024,** *21,* 312–329. doi:10.1038/s41571-024-00869-z
26. Burslem, G. M.; Crews, C. M. *Cell* **2020,** *181,* 102–114. doi:10.1016/j.cell.2019.11.031
27. Sun, D.; Zhang, J.; Dong, G.; He, S.; Sheng, C. *J. Med. Chem.* **2022,** *65,* 14276–14288. doi:10.1021/acs.jmedchem.2c01159
28. Chirnomas, D.; Hornberger, K. R.; Crews, C. M. *Nat. Rev. Clin. Oncol.* **2023,** *20,* 265–278. doi:10.1038/s41571-023-00736-3
29. Ma, L.; Li, Y.; Luo, R.; Wang, Y.; Cao, J.; Fu, W.; Qian, B.; Zheng, L.; Tang, L.; Lv, X.; Zheng, L.; Liang, G.; Chen, L. *J. Med. Chem.* **2023,** *66,* 7438–7453. doi:10.1021/acs.jmedchem.3c00150
30. Heuveling, J.; Landmesser, H.; Schneider, E. *J. Bacteriol.* **2018,** *22,* e00092-18. doi:10.1128/jb.00092-18
31. Fischer, E. S.; Böhm, K.; Lydeard, J. R.; Yang, H.; Stadler, M. B.; Cavadini, S.; Nagel, J.; Serluca, F.; Acker, V.; Lingaraju, G. M.; Tichkule, R. B.; Schebesta, M.; Forrester, W. C.; Schirle, M.; Hassiepen, U.; Ottl, J.; Hild, M.; Beckwith, R. E. J.; Harper, J. W.; Jenkins, J. L.; Thomä, N. H. *Nature* **2014,** *512,* 49–53. doi:10.1038/nature13527
32. Wang, S.; Peng, N.; Yin, L.; Wang, W.; Wu, Y.; Zhang, D.; Gan, Z. *Synth. Commun.* **2024,** *54,* 2120–2129. doi:10.1080/00397911.2024.2423897
33. Singhania, V.; Nelson, C. B.; Reamey, M.; Morin, E.; Kavthe, R. D.; Lipshutz, B. H. *Org. Lett.* **2023,** *25,* 4308–4312. doi:10.1021/acs.orglett.3c01380
34. Chen, L.; Fu, L.; Sun, J.; Huang, Z.; Fang, M.; Zinkle, A.; Liu, X.; Lu, J.; Pan, Z.; Wang, Y.; Liang, G.; Li, X.; Chen, G.; Mohammadi, M. *Nature* **2023,** *618,* 862–870. doi:10.1038/s41586-023-06155-9
35. Chen, G.; Chen, L.; Li, X.; Mohammadi, M. *Nat. Rev. Drug Discovery* **2025,** *24,* 335–357. doi:10.1038/s41573-024-01125-w
36. Chen, L.; Zhang, Y.; Yin, L.; Cai, B.; Huang, P.; Li, X.; Liang, G. *J. Exp. Clin. Cancer Res.* **2021,** *40,* 345. doi:10.1186/s13046-021-02156-6
37. Kim, R.; Sharma, S.; Meyer, T.; Sarker, D.; Macarulla, T.; Sung, M.; Choo, S. P.; Shi, H.; Schmidt-Kittler, O.; Clifford, C.; Wolf, B.; Llovet, J. M. *Eur. J. Cancer* **2016,** *69* (Suppl. 1), S41. doi:10.1016/s0959-8049(16)32704-6
38. Konecny, G. E.; Kolarova, T.; O'Brien, N. A.; Winterhoff, B.; Yang, G.; Qi, J.; Qi, Z.; Venkatesan, N.; Ayala, R.; Luo, T.; Finn, R. S.; Kristof, J.; Galderisi, C.; Porta, D. G.; Anderson, L.; Shi, M. M.; Yovine, A.; Slamon, D. J. *Mol. Cancer Ther.* **2013,** *12,* 632–642. doi:10.1158/1535-7163.mct-12-0999
39. Fairhurst, R. A.; Knoepfel, T.; Buschmann, N.; Leblanc, C.; Mah, R.; Todorov, M.; Nimsgern, P.; Ripoche, S.; Niklaus, M.; Warin, N.; Luu, V. H.; Madoerin, M.; Wirth, J.; Graus-Porta, D.; Weiss, A.; Kiffe, M.; Wartmann, M.; Kinyamu-Akunda, J.; Sterker, D.; Stamm, C.; Adler, F.; Buhles, A.; Schadt, H.; Couttet, P.; Blank, J.; Galuba, I.; Trappe, J.; Voshol, J.; Ostermann, N.; Zou, C.; Berghausen, J.; Del Rio Espinola, A.; Jahnke, W.; Furet, P. *J. Med. Chem.* **2020,** *63,* 12542–12573. doi:10.1021/acs.jmedchem.0c01019

40. Du, G.; Jiang, J.; Wu, Q.; Henning, N. J.; Donovan, K. A.; Yue, H.; Che, J.; Lu, W.; Fischer, E. S.; Bardeesy, N.; Zhang, T.; Gray, N. S. *Angew. Chem., Int. Ed.* **2021,** *133,* 16041–16047. doi:10.1002/anie.202101328
41. Pike, A.; Williamson, B.; Harlfinger, S.; Martin, S.; McGinnity, D. F. *Drug Discovery Today* **2020,** *25,* 1793–1800. doi:10.1016/j.drudis.2020.07.013

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.44

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Towards the targeted protein degradation of CK2: design and synthesis of CAM4066-based PROTACs

Sophie Day-Riley[1], Sona Krajcovicova[1,2], Aryaman Raj Sokhal[1], Jan L. Venne[1], Paul Brear[3], Marko Hyvönen[3], Benjamin C. Whitehurst[4], Jason S. Carroll[5] and David R. Spring[*1]

Letter

Open Access

Address:
[1]Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW, Cambridge, United Kingdom, [2]Department of Organic Chemistry, Faculty of Science, Palacky University, Tr. 17. Listopadu 12, Olomouc, 77900, Czech Republic, [3]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1GA, United Kingdom, [4]Hit Discovery, Discovery Sciences, R&D, AstraZeneca, Cambridge, CB2 0AA, United Kingdom and [5]Cancer Research UK Cambridge Institute, Robinson Way, Cambridge, CB2 0RE, United Kingdom

Email:
David R. Spring[*] - spring@ch.cam.ac.uk

* Corresponding author

Keywords:
CAM4066; casein kinase 2 (CK2); PROTACs; targeted protein degradation

Beilstein J. Org. Chem. **2026,** *22,* 611–619.
https://doi.org/10.3762/bjoc.22.47

Received: 10 January 2026
Accepted: 10 April 2026
Published: 22 April 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: J. D. Rudolf



© 2026 Day-Riley et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Human protein kinase CK2 is a constitutively active serine/threonine kinase implicated in numerous cancers. Although ATP-competitive inhibitors such as CX-4945 show therapeutic potential, they are limited by off-target effects and incomplete or transient CK2 suppression. PROTACs offer an alternative strategy by inducing proteasome-mediated degradation, with potential advantages in potency, selectivity, and duration of action. Herein, a series of CK2-targeting PROTACs has been designed and synthesised. By conjugating a CAM4066-derived warhead to CRBN or VHL ligands, four VHL-recruiting PROTACs, were prepared using PEG and alkyl linkers, alongside two CRBN-recruiting analogues featuring constrained linkers. A ligand–linker analogue in which a linker is projected from the solvent-exposed region of CK2α retained binding affinity comparable to CAM4066, confirming that linker installation is tolerated and preserves key interactions in the αD and ATP sites.

Introduction

Protein kinases form a large family of more than 500 enzymes that regulate key cellular processes through the phosphorylation of protein substrates. Dysregulation of kinase activity, often through mutation or altered expression, is associated with numerous malignancies [1]. CK2 (previously called casein kinase 2), a constitutively active serine/threonine kinase, functions as a

heterotetramer composed of two catalytic (α or α′) and a dimeric regulatory (β) subunit [2]. Unlike most kinases, CK2 does not require upstream activation [3], a property that contributes to its pleiotropic role in cell signalling. Elevated CK2 expression correlates with enhanced cell survival across several cancer types, and its downregulation impairs viability in models of glioblastoma, medulloblastoma, cholangiocarcinoma, breast and renal cancers [4].

These features have established CK2 as a compelling therapeutic target. In 2016, structural studies revealed a cryptic αD pocket adjacent to the ATP-binding site of CK2α [5,6]. Fragment-based ligand discovery subsequently enabled the development of CAM4066, a selective inhibitor that simultaneously engages the αD and ATP sites. CAM4066 validated the αD region as a tractable and selective binding pocket for CK2 kinase inhibition. Despite successes with ATP-competitive inhibitors, small-molecule kinase inhibitors can be limited by reversible target engagement, off-target effects, and susceptibility to resistance mechanisms [7]. These challenges have motivated the exploration of alternative therapeutic modalities that move beyond occupancy-driven pharmacology.

Proteolysis-targeting chimeras (PROTACs) represent a rapidly advancing strategy for induced protein degradation. By transiently engaging both a target protein and an E3 ubiquitin ligase, PROTACs promote proteasomal elimination of the target rather than sustained binding [8]. This event-driven mode of action enables catalytic turnover, can reduce off-target toxicities associated with high inhibitor doses, and does not require a deep or well-defined binding pocket, allowing potential access to targets considered “undruggable” by conventional small molecules [8]. In 2018, a PROTAC targeting CK2 was reported in the literature. This used CX-4945, a potent CK2 inhibitor, as the CK2 warhead [9,10]. However, as CX-4945 targeted the ATP-binding site of CK2, it also displayed nanomolar affinity for the ATP-binding sites of other kinases, like CLK2 [11]. Therefore, finding a potent and selective PROTAC targeting CK2 might be valuable in the evaluation of specific degradation of CK2 as a therapeutic approach.

Herein, we report the design and synthesis of a series of CK2-targeting PROTACs. CAM4066 was selected as the warhead to enable CK2-specific target engagement. We conjugated CAM4066-derived ligands to both CRBN or VHL E3-ligase recruiters. We created VHL-recruiting PROTACs for CK2 using PEG and alkyl linkers, while the CRBN-recruiting analogues featured more constrained linkers. A ligand–linker analogue bearing a linker projected from the solvent-exposed region of the CK2 ligand retained binding affinity comparable to CAM4066, demonstrating that linker installation is well tolerated and preserves key interactions in both the αD and ATP sites.

Results and Discussion

CAM4066 (**1**) was selected as the CK2-targeting ligand for PROTAC development owing to its high selectivity and well-characterised bivalent binding mode across the ATP and αD pockets. Because a successful degrader must first engage the protein of interest with sufficient affinity, a PROTAC derived from CAM4066 was expected to retain CK2 specificity and achieve kinase-directed degradation rather than broader kinome perturbation. The first design objective was therefore to identify a suitable exit vector for linker attachment that would allow productive E3-ligase recruitment without compromising the established binding geometry of CAM4066.

Inspection of the co-crystal structure of **1** bound to CK2α revealed a small solvent-accessible channel adjacent to the amide linking the ATP- and αD-binding fragments (Figure 1). This channel represented a promising region from which to project a linker, as it was oriented toward solvent and spatially removed from residues critical for **1** binding to CK2α. The amide closest to this opening was therefore identified as the optimal site for structural modification. To allow more flexible functionalisation, the linking region was redesigned to introduce a secondary amine adjacent to the original amide position. This amine maintained the length and orientation of the native framework while serving as a more accessible synthetic handle for downstream coupling chemistry. As **1** itself was known not to have any effect on cell viability while its uncharged methyl ester analogue did, the neutralised pro-drug analogue of the ATP-site binder was employed to prevent such problems [6]. The ATP-site binder and αD-site binder fragments were prepared using literature procedures from the original development of **1**. The modified linking portion containing the new secondary amine was incorporated into the general assembly **2** to afford a CAM4066-derived scaffold suitable for PROTAC construction.

The synthesis of the ATP-site binder began with the alkylation of commercially available methyl 3-aminobenzoate (**3**) using bromoacetyl bromide. Although initial purification via silica gel chromatography led to substantial product loss, isolation by simple aqueous workup afforded the intermediate **4** cleanly and in quantitative yield. This intermediate enabled straightforward elaboration of the ATP-binding fragment, as *N*-Boc-1,4-diaminobutane was subsequently alkylated with intermediate **4** to generate the chain-extended linker precursor **5**. The use of DIPEA instead of triethylamine suppressed undesired alkylation of the base and significantly improved the isolated yield up to 84%. Introduction of an acetic acid handle was then

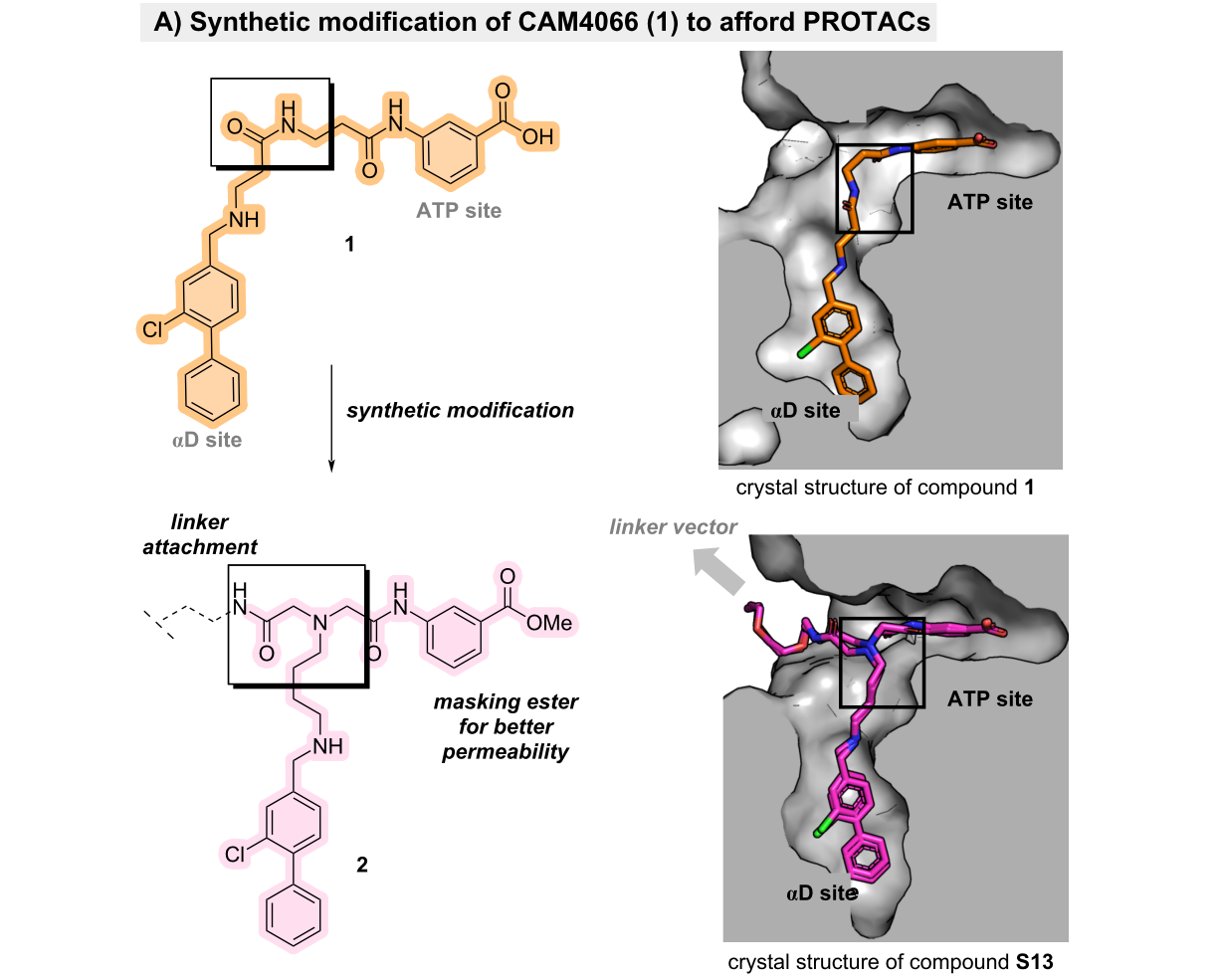
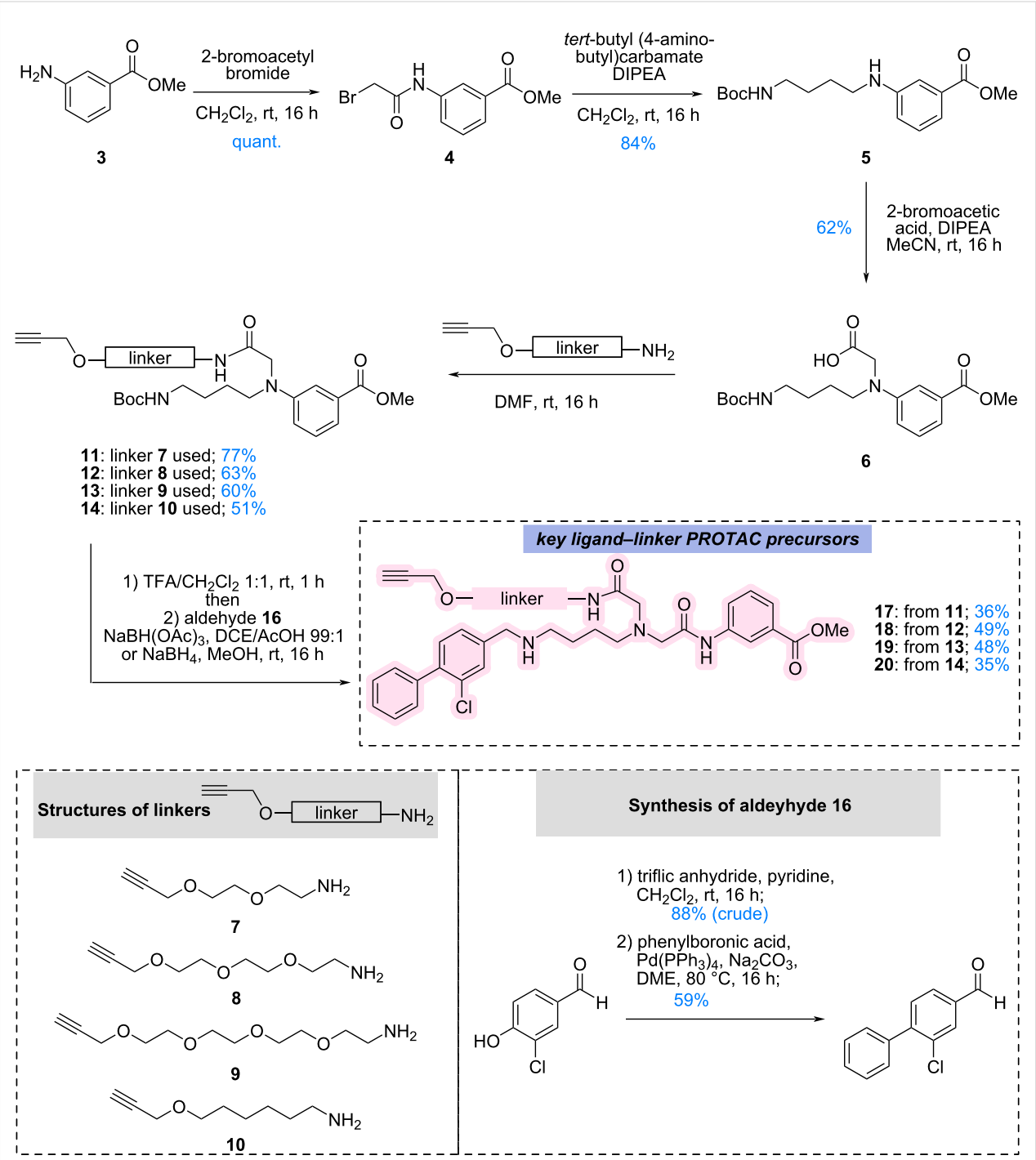


Figure 1: Design strategy and validation. A) Structure of CAM4066 (**1**) that served as a model design for the development of CK2 targeted PROTACs with a general structure **2**. Crystal structures showcased on **1** (PDB: 5CU4) and **S13** (PDB: 9TTA; please see Supporting Information File 1, section 1.4.18 for the full structure). B) Crystal structure of **S13**. The map is Fo-Fc contoured at 1.5 σ. C) Isothermal titration calorimetry (ITC) reveals the suitable position for linker vector to attach the E3 ligases (showcased on compound **S13**).

achieved via alkylation with bromoacetic acid under DIPEA mediation, giving intermediate **6**, which served as the common attachment point for linkers **7**–**10** (for full synthesis see Supporting Information File 1, section 1.3). In parallel, PEG-based linkers bearing terminal alkyne handles were synthesised for modular CuAAC-based PROTAC assembly. The αD binder **16** had been synthesised via a two-step reaction. First, the phenolic group of **15** was converted into the corresponding triflate using triflic anhydride, followed by Suzuki–Miyaura cross-coupling with phenylboronic acid to give **16** in 59% yield (Scheme 1).



Scheme 1: Synthesis of the key ligand-linker PROTAC precursors. For compound **14**, HATU/DIPEA has been used instead of HOBt/DIC. Abbreviations: DCE = 1,2-dichloroethane; DIC = *N*,*N*′-diisopropylcarbodiimide; DIPEA = *N*,*N*-diisopropylethylamine; DME = 1,2-dimethoxyethane; DMF = dimethylformamide; HOBt = 1-hydroxybenzotriazole; TFA = trifluoroacetic acid.

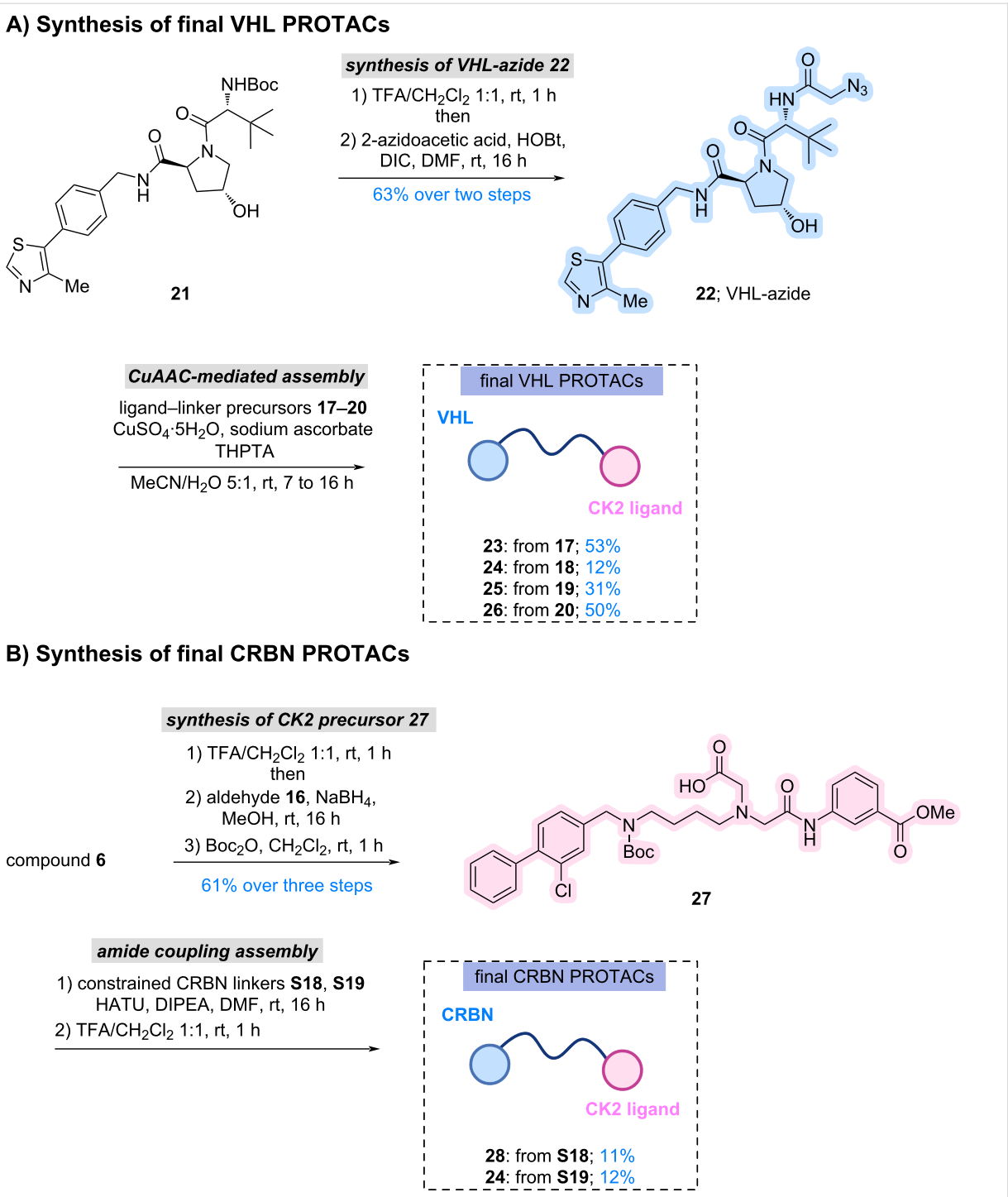
Boc removal from intermediates **11**–**14** enabled attachment of the αD binder aldehyde **16** via reductive amination. Interestingly, $NaBH(OAc)_3$ had performed well in the synthesis of the ligand–linker analogue **17** using intermediate **11**, however, use of intermediates **12**–**14** were consistently giving low yields. Therefore, for those, methanolic reductive amination with $NaBH_4$ was pursued. Successful optimisation of reductive amination conditions yielded the key ligand–linker PROTAC precursors **17**–**20** in moderate to good yields (35–49%; Scheme 1).

Before progressing to full PROTACs, it was essential to confirm that the modified CAM4066 scaffold tolerated linker installation. For this, the methyl ester of the key ligand–linker precursor **17** was hydrolysed using aqueous LiOH to yield acid **S13**. Isothermal titration calorimetry (ITC) revealed that **S13** bound CK2α with a dissociation constant (K_d = 600 nM) comparable to that of CAM4066 (350 nM; Figure 1C).

This indicated that the exit vector modification and linker installation were well tolerated. Co-crystallisation of **S13** with CK2α further confirmed, that both the ATP-binding and αD-binding motifs retained their expected hydrogen-bonding and hydrophobic interactions, and the linker projected cleanly toward solvent-exposed region (Figure 1). Importantly, the absence of the terminal alkyne substituent in the electron density map revealed that it remained solvent-exposed and sterically accessible, validating the design for subsequent attachment of E3-ligase ligands.

With the exit vector validated, the synthesis of full PROTACs was pursued. The VH032-derived VHL ligand **21** (for full synthesis see Supporting Information File 1, section 1.4) was prepared according to literature and functionalised with 2-azidoacetic acid to introduce an azido-handle suitable for CuAAC-mediated triazole formation. CuAAC coupling of these alkyne precursors **17**–**20** with the azido-VHL ligand **22** afforded four final VHL-based PROTACs **23**–**26** in moderate to good yields (Scheme 2A, Figure 2). It is important to mention that LCMS analysis indicated complete consumption of the alkyne-bearing intermediates and clean formation of the triazole-linked products after 16 hours, although isolated yields varied due to product loss during preparative RP-HPLC. Two CRBN-based degraders **28** and **29** were synthesised using constrained linker–ligand constructs **S18** and **S19**. Because the steric environment imposed by these linkers differed from that of the VHL system, a modified synthetic sequence was adopted. The αD-site binder **16** was first introduced onto the intermediate **6**, which upon Boc protection yielded intermediate **27**. Installation of CRBN linker–ligands proceeded smoothly under standard coupling conditions using HATU/DIPEA, and Boc removal afforded CRBN PROTACs **28** and **29** (Scheme 2B, Figure 2). Their low isolated yields were ultimately traced to poor solubility in $MeCN/H_2O$ systems used for preparative purification; the compounds dissolved only in DMSO, explaining the diminished recovery. Nevertheless, the synthetic route successfully delivered two structurally distinct CRBN-recruiting CK2 PROTACs.

All synthesised PROTACs **23**–**26**, **28**, and **29** were evaluated for their ability to induce degradation of CK2α and CK2α′ in cell-based assays. Western blot profiling across six colorectal cancer cell lines revealed heterogeneous expression of both CK2 isoforms. HCT116 cells exhibited robust expression of CK2α and CK2α′, with CK2α predominating, and were therefore selected as the primary model for degradation studies. The MDA-MB-231 breast cancer line was also included due to its prior use in published evaluations of a CK2-targeting PROTAC [10]. The four VHL-based PROTACs **23**–**26** were tested at concentrations up to 10 μM. However, no measurable degradation of CK2α or CK2α′ under these conditions was observed, as indicated by unchanged band intensities relative to controls. At the highest concentration (10 μM), PROTAC **24** caused a loss of both CK2α and vinculin bands, suggesting cytotoxicity rather than target-specific degradation. The two CRBN-based PROTACs **28** and **29** were tested under identical conditions but likewise failed to induce degradation of CK2α or CK2α′ in HCT116 or MDA-MB-231 cells (see Supporting Information File 1, section 2.4). Upon biological evaluation, ITC analysis of **24**, selected as a representative PROTAC, revealed no detectable productive binding event (see Supporting Information File 1, section 2.1). This may have resulted from several factors, including the limited solubility of **24** under the conditions required for the assay and the requirement for cooperative ternary complex formation with the E3 ligase, as the isolated warhead–linker construct **S13** displayed measurable binary binding in the absence of the degrader architecture. Although the ligand–linker analogue showed encouraging biophysical and structural properties, the six full PROTACs did not induce detectable CK2 degradation under the conditions tested. Several mechanistic factors may underlie this outcome. PROTACs frequently fail because the protein of interest and E3 ligase cannot adopt a productive orientation for ubiquitin transfer. Steric constraints, linker vector misalignment, or insufficient linker flexibility may prevent formation of a catalytically competent ternary complex. The azide–alkyne-derived triazole linkage may also impose a rigid orientation not compatible with CK2's topology. Moreover, PROTACs typically display reduced cellular permeability owing to their high molecular weight and polar surface area [12]. This is particularly relevant given that CAM4066 itself is a moderately polar molecule, and linker addition further increases polarity.



Scheme 2: (A) Synthesis of final VHL PROTACs; (B) Synthesis of final CRBN PROTACs. Abbreviations: Boc = *tert*-butoxycarbonyl; CRBN = cereblon; DIPEA = *N*,*N*-diisopropylethylamine; DMF = dimethylformamide; HATU = *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate; THPTA = tris(3-hydroxypropyltriazolylmethyl)amine; TFA = trifluoroacetic acid; VHL = von Hippel–Lindau.

Collectively, these proof-of-concept results highlight the complexity of designing bivalent molecules for CK2 degradation and underscore the need for further optimisation of warhead potency, linker length, exit vector geometry, and E3 ligase selection.

Conclusion

In summary, we have developed a modular synthetic platform for the construction of CK2-targeting degraders **23**–**26**, **28**, and **29** based on the selective bivalent inhibitor CAM4066. Structure-guided design enabled identification of a solvent-acces-

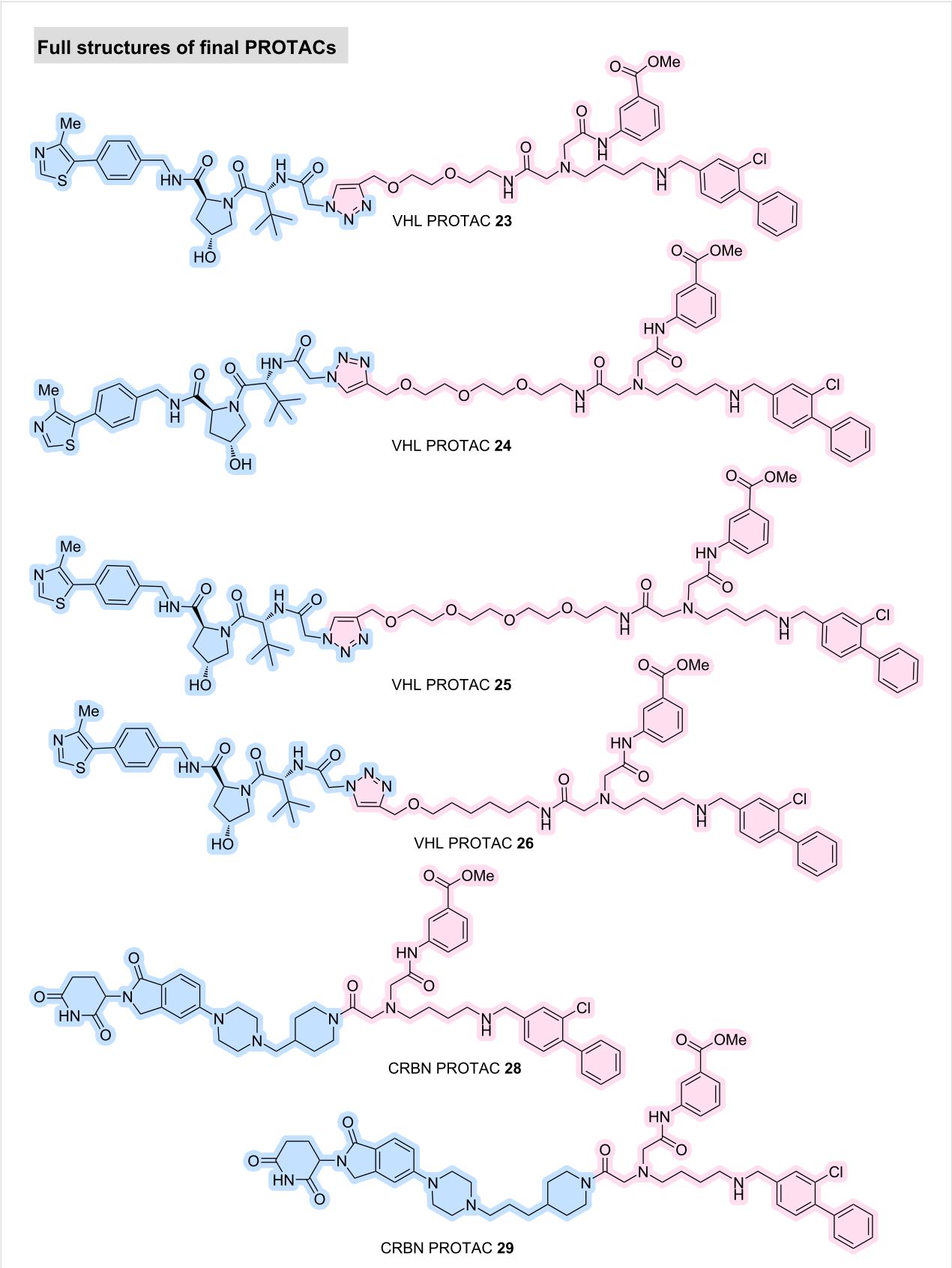


Figure 2: Full structures of final VHL PROTACs **23**–**26** and CRBN PROTACs **28** and **29**.

sible exit vector suitable for linker installation, and introduction of a secondary amine at this position provided a versatile handle for downstream functionalisation. The resulting CAM4066-derived ligand–linker analogue **S13** retained high-affinity binding to CK2α and preserved the characteristic bivalent binding mode in co-crystal structures, confirming that the exit vector and linker trajectory were well tolerated. These data established a robust foundation for the subsequent development of PROTACs.

Using this scaffold, we successfully synthesised a set of PROTAC precursors and completed six full degraders recruiting either VHL or CRBN. Although none of the prepared PROTACs induced degradation of CK2α or CK2α′ in cellular assays, the work provides valuable insights into the challenges associated with degrading compact kinases such as CK2. The modest potency of the modified CAM4066 warheads, potential geometric incompatibility within the ternary complex, and the inherent permeability limitations of high-molecular-weight degraders likely contributed to the lack of observed activity.

Despite these limitations, the study offers important lessons for future design. The validated exit vector and structural understanding of warhead tolerance provide a strong basis for further optimisation, including exploration of alternative warheads with improved potency, such as recently developed APL-5125 [13], a larger set of varied linker architectures [14,15], and engagement of additional E3 ligases [16]. More broadly, the synthetic strategy and the determination of tolerance of exit vector position developed here will support continued efforts to achieve targeted degradation of CK2.

Taken together, this work establishes the key design principles, synthetic routes, and structural requirements necessary for next-generation CK2 degraders and lays the groundwork for future optimisation toward achieving effective and selective CK2 degradation.

Supporting Information

Supporting Information File 1

Experimental part and copies of NMR spectra.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-47-S1.pdf]

Acknowledgements

The authors are grateful to Kristina Kostadinova and Dr Marc de la Roche for the help with the Western Blot analysis, to Dr Eleanor Atkinson and Dr Jessica Iegre for their prior work on CK2 inhibitors, and to Dr Michal Malon for his help with several NMR assignments. We are grateful for the Biophysical and X-ray crystallographic facilities at the Department of Biochemistry for access to instrumentation. We thank Diamond Light Source for access to beamline (proposal ID mx25402).

Funding

S. Day-Riley is grateful AstraZeneca Cambridge for a studentship and financial support. S. Krajcovicova is grateful to the Czech Science Foundation (GA CR 22-07138O) for their financial support. A. R. Sokhal is grateful to the Gates Cambridge Trust for their financial support (https://www.gatescambridge.org). The Spring lab acknowledges support from the EPSRC, BBSRC, MRC, and Cystic Fibrosis Trust UK. B. C. Whitehurst was an employee of AstraZeneca at the time this work was conducted.

Author Contributions

Sophie Day-Riley: data curation; formal analysis; investigation; methodology; writing – original draft. Sona Krajcovicova: conceptualization; formal analysis; investigation; methodology; supervision; validation; visualization; writing – review & editing. Aryaman Raj Sokhal: investigation; methodology; visualization; writing – review & editing. Jan L. Venne: formal analysis; investigation; writing – review & editing. Paul Brear: data curation; formal analysis; software; supervision; visualization; writing – review & editing. Marko Hyvönen: conceptualization; data curation; supervision; writing – review & editing. Benjamin C. Whitehurst: conceptualization; funding acquisition; resources; writing – review & editing. Jason S. Carroll: resources; supervision. David R. Spring: conceptualization; funding acquisition; project administration; supervision; writing – review & editing.

ORCID® iDs

Paul Brear - https://orcid.org/0000-0002-4045-0474
Benjamin C. Whitehurst - https://orcid.org/0000-0002-5485-1762

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Hou, Z.; Liu, H. *Cells* **2023,** *12,* 925. doi:10.3390/cells12060925
2. Niefind, K.; Guerra, B.; Ermakowa, I.; Issinger, O.-G. *EMBO J.* **2001,** *20,* 5320–5331. doi:10.1093/emboj/20.19.5320
3. Turowec, J. P.; Duncan, J. S.; French, A. C.; Gyenis, L.; St. Denis, N. A.; Vilk, G.; Litchfield, D. W. *Methods Enzymol.* **2010,** *484,* 471–493. doi:10.1016/b978-0-12-381298-8.00023-x

4. Chen, Y.; Wang, Y.; Wang, J.; Zhou, Z.; Cao, S.; Zhang, J. *J. Med. Chem.* **2023,** *66,* 2257–2281. doi:10.1021/acs.jmedchem.2c01523
5. De Fusco, C.; Brear, P.; Iegre, J.; Georgiou, K. H.; Sore, H. F.; Hyvönen, M.; Spring, D. R. *Bioorg. Med. Chem.* **2017,** *25,* 3471–3482. doi:10.1016/j.bmc.2017.04.037
6. Brear, P.; De Fusco, C.; Hadje Georgiou, K.; Francis-Newton, N. J.; Stubbs, C. J.; Sore, H. F.; Venkitaraman, A. R.; Abell, C.; Spring, D. R.; Hyvönen, M. *Chem. Sci.* **2016,** *7,* 6839–6845. doi:10.1039/c6sc02335e
7. Cromm, P. M.; Crews, C. M. *Cell Chem. Biol.* **2017,** *24,* 1181–1190. doi:10.1016/j.chembiol.2017.05.024
8. Burslem, G. M.; Crews, C. M. *Cell* **2020,** *181,* 102–114. doi:10.1016/j.cell.2019.11.031
9. Siddiqui-Jain, A.; Drygin, D.; Streiner, N.; Chua, P.; Pierre, F.; O'Brien, S. E.; Bliesath, J.; Omori, M.; Huser, N.; Ho, C.; Proffitt, C.; Schwaebe, M. K.; Ryckman, D. M.; Rice, W. G.; Anderes, K. *Cancer Res.* **2010,** *70,* 10288–10298. doi:10.1158/0008-5472.can-10-1893
10. Chen, H.; Chen, F.; Liu, N.; Wang, X.; Gou, S. *Bioorg. Chem.* **2018,** *81,* 536–544. doi:10.1016/j.bioorg.2018.09.005
11. Kim, H.; Choi, K.; Kang, H.; Lee, S.-Y.; Chi, S.-W.; Lee, M.-S.; Song, J.; Im, D.; Choi, Y.; Cho, S. *PLoS One* **2014,** *9,* e94978. doi:10.1371/journal.pone.0094978
12. Matsson, P.; Kihlberg, J. *J. Med. Chem.* **2017,** *60,* 1662–1664. doi:10.1021/acs.jmedchem.7b00237
13. Glossop, P. A.; Brear, P.; Wright, S.; Flanagan, N.; Glossop, M. S.; Lane, C. A. L.; Butt, R. P.; Spring, D. R.; Hyvönen, M.; Cawkill, D. *J. Med. Chem.* **2025,** *68,* 21587–21614. doi:10.1021/acs.jmedchem.5c01807
14. Troup, R. I.; Fallan, C.; Baud, M. G. J. *Explor. Targeted Anti-Tumor Ther.* **2020,** *1,* 273–312. doi:10.37349/etat.2020.00018
15. Chen, Q.-H.; Munoz, E.; Ashong, D. *Cancers* **2024,** *16,* 663. doi:10.3390/cancers16030663
16. Campos, M. A.; Riha, I. A.; Zhang, C.; Mozes, C.; Scheidt, K. A.; Zhang, X. *ACS Chem. Biol.* **2025,** *20,* 479–488. doi:10.1021/acschembio.4c00799

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.47

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Advantages of PROTACs in achieving selective degradation of homologous protein families

Luxi Yang[‡1,2], Xinfei Mao[‡3], Jingyi Zhang[1], Jing Shu[1], Wenhai Huang[1,2], Xiaowu Dong[*3], Yinqiao Chen[*1] and Mingfei Wu[*1]

Review

Open Access

Address:
[1]Affiliated Yongkang First People's Hospital, School of Pharmaceutical Sciences, Hangzhou Medical College, Zhejiang, Hangzhou, 311399, China, [2]Center of Safety Evaluation and Research, Hangzhou Medical College, Hangzhou 310053, China and [3]College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, 310058, China

Email:
Xiaowu Dong[*] - dongxw@zju.edu.cn; Yinqiao Chen[*] - chenyinqiao668@163.com; Mingfei Wu[*] - wmfayd@163.com.

* Corresponding author ‡ Equal contributors

Keywords:
homologous protein; PROTAC; protein–protein interaction; selectivity; ubiquitination

Beilstein J. Org. Chem. **2026,** *22,* 628–661.
https://doi.org/10.3762/bjoc.22.49

Received: 09 January 2026
Accepted: 08 April 2026
Published: 27 April 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2026 Yang et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Proteolysis-targeting chimeras (PROTACs) have emerged as a promising therapeutic modality and now represent an important addition to the toolkit of medicinal chemists. Compared with conventional small-molecule inhibitors, PROTACs exhibit notable advantages, particularly in achieving selectivity within highly homologous proteins. The ability to discriminate among members of such families holds broad implications for future disease treatment. In this review, we primarily summarize the advantages of PROTACs in conferring selectivity toward highly homologous proteins. This focus will provide a feasible strategy for developing PROTACs that selectively target highly homologous proteins and will ultimately support future therapeutic applications.

Introduction

The cell is the fundamental unit of structure and function in the human body [1,2]. More than 20,000 proteins act in concert to regulate the entire cellular life process [1]. To date, dysregulated protein function has been recognized as one of the leading causes of human diseases [3]. Although small-molecule inhibitors [4] remain the most widely used therapeutic agents for the treatment of disorders associated with abnormal protein activity, nearly all disease-relevant scaffolding proteins [5], transcription factors [6], and other non-enzymatic proteins [7] are essentially undruggable using conventional inhibitor-based approaches [4,8]. Consequently, protein–protein interaction (PPI)-based targeted protein degradation (TPD) strategies have attracted increasing attention over the past two decades [9,10]. Among them, PROTACs, one of the most extensively studied and promising TPD approaches, are reshaping the paradigm of small-molecule drug development [11].

The PROTAC technology was initially conceptualized by Crews et al. in 2001 [11]. As a heterobifunctional molecule, a PROTAC comprises two distinct ligands: one targeting an E3 ubiquitin ligase and the other binding to a protein of interest (POI), covalently joined by a flexible linker [12]. Upon cellular entry, the PROTAC molecule facilitates the formation of a ternary complex by simultaneously recruiting the POI and the E3 ligase (Figure 1). This proximity enables the proteasome to recognize the complex, subsequently degrading the target protein via the endogenous ubiquitin–proteasome system [13,14]. Notably, this mechanism does not necessitate prolonged occupancy of the POI binding site. Instead, the transient formation of the ternary complex is sufficient to trigger rapid ubiquitination [15,16]. Consequently, a PROTAC operates through an event-driven pharmacological paradigm, distinguishing itself from the traditional occupancy-driven model of small-molecule inhibitors [17]. Furthermore, PROTACs function catalytically, requiring only substoichiometric doses to achieve effective protein degradation [18].

Recent reviews have extensively detailed the principles, advancements, and therapeutic prospects of TPD, with a particular focus on PROTACs [15,19]. As understanding of this field matures, medicinal chemists are increasingly evaluating the comparative advantages and limitations of PROTACs relative to traditional small-molecule inhibitors [20]. A defining strength of the PROTAC technology lies in its superior selectivity for target proteins, especially within highly homologous proteins [21-23]. Selectivity remains a cornerstone of pharmacology; inhibitors that lack specificity may interfere with off-target biological functions, resulting in diminished therapeutic efficacy and unacceptable toxicity [24]. Consequently, maximizing compound selectivity is a primary objective in drug discovery [25,26]. In the context of PROTAC design, it is imperative to minimize the degradation of non-targeted proteins.

While conventional inhibitors often exhibit lacking selectivity towards proteins with high sequence similarity, PROTACs leverage their heterobifunctional structure to achieve exquisite selectivity [24,27]. This is accomplished by fine-tuning the conformation of the ternary complex through the strategic modification of the POI ligand, the linker, and the E3 ligase ligand. Several studies have successfully utilized this strategy to enhance disease treatment outcomes via the PROTAC technology [22,28,29].

The rapid progression of the PROTAC technology from an academic concept to clinical application underscores its transformative potential. In recent years, numerous PROTAC candidates have entered clinical trials, marking a significant milestone for the field [23]. Notably, multiple degraders targeting the androgen receptor (AR), such as ARV-110, ARV-766, and CC-94676, are currently being evaluated in patients with prostate cancer, demonstrating that the event-driven pharmacology of PROTACs translates into durable target knockdown in humans [30-32]. Beyond the AR, the clinical landscape is expanding to include targets discussed extensively in this review. For instance, selective degraders of CDK isoforms are advancing through early-phase trials [33], with molecules like

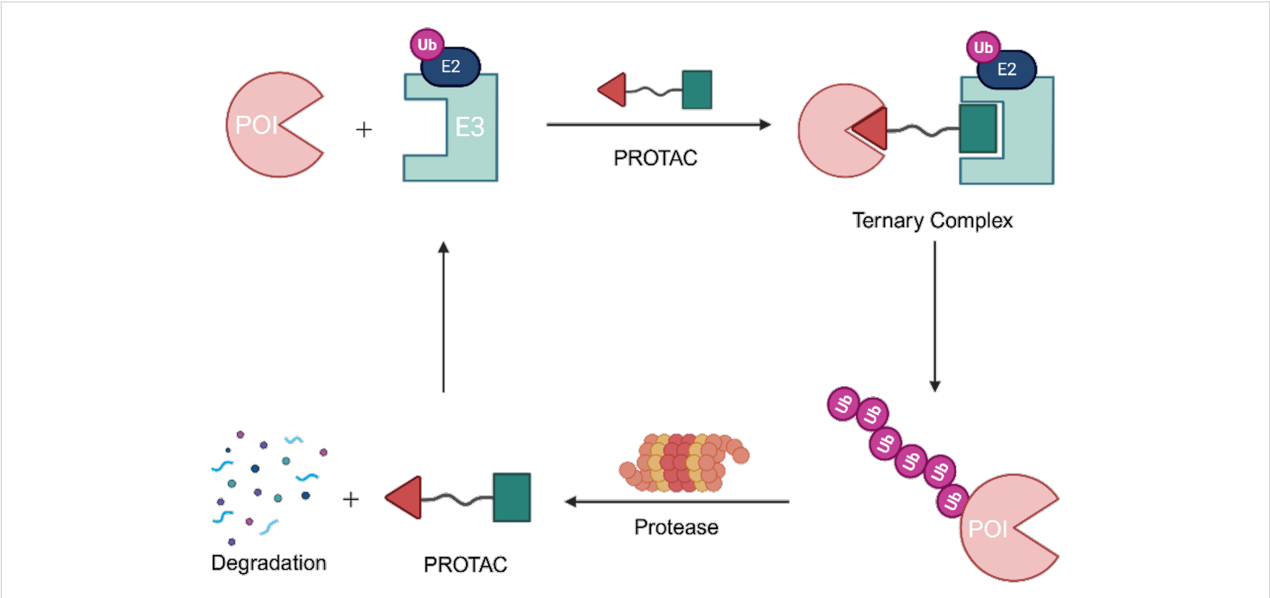


Figure 1: Mechanism of a PROTAC-mediated targeted protein degradation. Created in BioRender. Wu, M. (2026) https://BioRender.com/xig7tb2. This content is not subject to CC BY 4.0.

BSJ-03-123 and others showing promise in hematologic malignancies by exploiting the subtle structural differences among the highly homologous CDK family members to achieve precision targeting [34]. Similarly, efforts to clinically translate PROTACs against epigenetic targets, such as BRD4, are gaining momentum, aiming to overcome the isoform selectivity challenges that have historically hindered small-molecule inhibitors in the BET family [35]. This wave of clinical validation not only confirms the unique advantages of PROTACs in addressing previously undruggable targets but also highlights the critical importance of understanding how linker composition, E3 ligase choice, and protein–protein interactions govern the exquisite selectivity required for clinical success [24,36].

Accordingly, this review emphasizes the advantages of PROTACs in achieving high selectivity among homologous proteins. We systematically summarize the correlations between PROTAC structural components, specifically the POI ligand, the linker, and the E3 ligand, and their role in dictating isoform-specific selectivity. Furthermore, we examine the influence of PPIs, a factor increasingly recognized for its critical contribution to the refined selectivity of PROTACs. The review also explores potential selectivity variations arising from the spatial distribution of ubiquitination-accessible lysine residues within homologous proteins. In conclusion, by synthesizing current knowledge and offering future perspectives on PROTAC selectivity, this work aims to provide a robust framework and feasible strategies for the rational design of degraders specifically targeting homologous proteins.

Review

Influence of linkers on the selectivity of PROTACs in highly homologous protein families

As previously established [37], a PROTAC molecule comprises three essential components: a POI ligand, a linker, and an E3 ligand. In drug design, POI ligands are typically derived from FDA-approved small-molecule inhibitors, candidates in clinical trials, or their structural analogs. Similarly, E3 ligands are engineered to target specific E3 ubiquitin ligases. For instance, pomalidomide and lenalidomide are widely employed to recruit the CRBN ligase [38]. Given the established nature of these ligands, initial optimization often prioritizes the linker fragment, as its length and composition are critical determinants of target selectivity [39]. Recent evidence underscores the pivotal role of the linker in promoting the assembly of "positive cooperative" ternary complexes [40], wherein the linker participates in specific molecular interactions within the complex [41,42]. These findings have profound implications for the rational design of PROTACs that exhibit isoform-specific selectivity among structurally related POIs [43]. Consequently, this section first examines the influence of linker architecture on the overall selectivity of PROTACs, utilizing several high-profile targets as illustrative examples.

CDK

Cyclin-dependent kinases (CDKs) are members of the serine/threonine (S/T) kinase subfamily comprising 21 CDK enzymes [44]. CDK1, 2, 3, 4, and 6 play a key role in cell cycle regulation. CDK7, 8, 9, and 11 are mainly involved in transcriptional regulation. However, the biological functions of CDK10, 11, 14–18, and 20 have not been fully elucidated [45,46]. CDKs are essential in cell proliferation, transcriptional activity, and neuronal activity [44]. In addition, disorders of these protein kinases often exist in cancer and nervous system diseases [47]. Palbociclib was the first small-molecule inhibitor approved by the FDA in 2015 to target CDK4/6 in the CDK family [48]. Soon after, ribociclib and abemaciclib, small-molecule inhibitors targeting CDK4/6, were also approved by the FDA for breast cancer treatment [49,50]. However, due to the high homology of the CDK family (for example, the homology of CDK4/6 is 71%, the homology of CDK2/3 is 76% [51], the homology of CDK8/19 is 97%) [52], small-molecule inhibitors may often inhibit several CDK members, thus, to a certain degree, cause off-target toxicity and reduce the therapeutic effect [53].

Since PROTACs have been proven to demonstrate significant advantages over small-molecule inhibitors in selectivity, more and more research has focused on how to design PROTACs that can target the degradation of the CDK family with high specificity in recent years [12]. The selectivity of PROTACs toward CDK4/6 over other highly homologous CDKs can be attributed to their event-driven pharmacology and the formation of stable ternary complexes that exploit subtle structural differences among the family members [54]. Unlike traditional kinase inhibitors that rely on sustained occupancy of conserved ATP-binding pockets, PROTACs induce transient yet specific protein–protein interactions that dictate ubiquitination efficiency. By fine-tuning the length and composition of the linker, as well as the choice of E3 ligase and its attachment geometry, PROTACs can achieve isoform-specific degradation. This approach not only minimizes off-target toxicity but also enables the targeting of CDK6 in contexts where CDK4 is functionally redundant, such as in Ph+ ALL.

Currently, small-molecule inhibitors for the CDK family approved by the FDA all target CDK4/6. CDK4/6 has high homology and has been shown to play a crucial role in the development of breast cancer [26]. In recent years, it has been proved that expression of CDK6, but not of the highly related

CDK4, is required for the proliferation and survival of Ph+ acute lymphoblastic leukemia cells. According to this, some researchers consider that CDK6 may be a more attractive target than CDK4 in some diseases [55]. Therefore, how to distinguish between CDK4 and CDK6, to achieve better treatment effects for certain specific diseases, such as Ph+ALL, has become a problem of concern to many pharmaceutical chemists. In 2019, Gray, Zhang and co-workers designed a PROTAC-1 (**1**) to degrade CDK4/6 [21]. To construct compound **1** they used the CDK4/6 inhibitor palbociclib as the POI ligand, pomalidomide, which combines with CRBN as the E3 ligand and a 4-carbon alkyl linker to connect them. The compound can simultaneously degrade CDK4 (the degradation rate exceeds 50% at 0.1 μM in Jurkat cells) and CDK6 (the degradation rate exceeds 95% at 0.1 μM in Jurkat cells), and it shows a certain degree of selectivity towards CDK6. In the same year, Gray, Winter et al. developed another PROTAC named BSJ-03-123 (**2**), which can selectively degrade CDK6 [56]. Compared to compound **1**, the POI ligand and the E3 ubiquitin ligase combined with the E3 ligand of compound **2** have not changed. The only difference is that the linker of compound **2** is longer and comprises a PEG chain (Figure 2). The results showed that

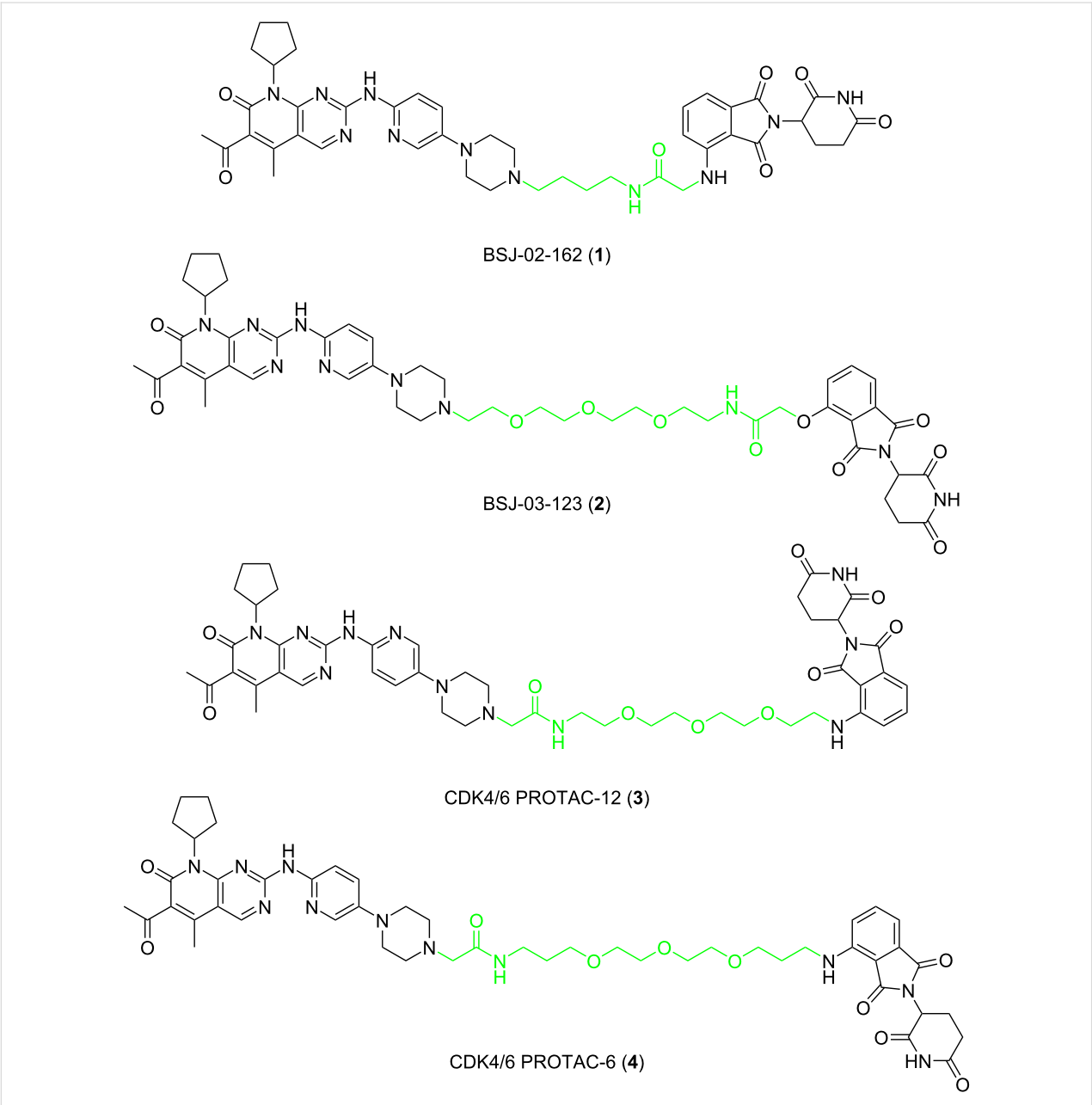


Figure 2: CDK4/6 PROTACs with alkyl or PEG chains as linkers.

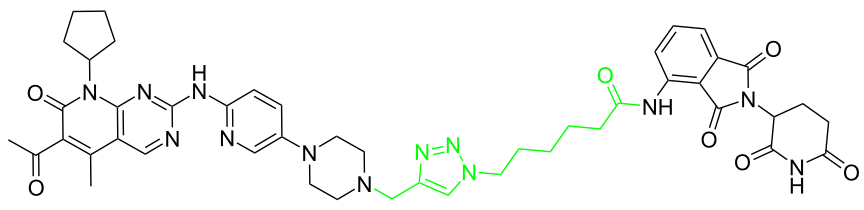
compound **2** has a better ability to target the degradation of CDK6, and proteomics results also show that compound **2** has no degradation effect on CDK4. In 2020, Benowitz and co-workers carried out a project to explore the impact of different E3 ligands on the selective degradation of CDK4/6 [57]. In their research, they designed various PROTACs consisting of different linkers and E3 ligands, and compound **3** (Figure 3) is one of them.

Interestingly, when comparing compounds **3** and **2**, the POI ligand, the binding E3 ubiquitin ligase and the linker are identical with the only difference being the direction the linker connects the POI ligand and E3 ligand. Similarly, the research results show that compound **3** can also selectively degrade CDK6. Meanwhile, Rana, Natarajan and co-workers developed a PROTAC-6 (**4**) based on palbociclib and pomalidomide [58] (Figure 2). In the study, the researchers evaluated the degradation effect of compound **4** in HPNE and MiaPaCa2 cells, respectively. It was found that compound **4** could selectively degrade CDK6 without affecting other members of the CDK family (CDK2, 4, 5, 7, and 9).

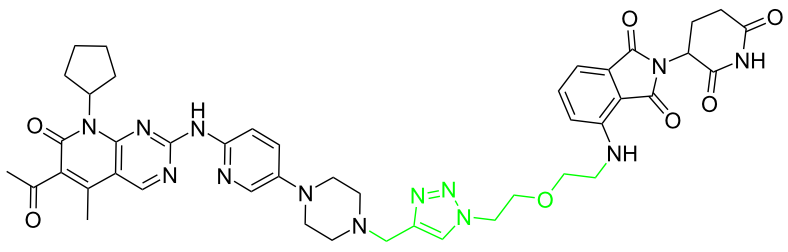
Based on the above studies, it seems an effective development strategy to use a longer linker to connect the POI ligand and E3 ligand when designing PROTACs with high selectivity to target the degradation of CDK6. However, some studies did not support this hypothesis. When considering the influence of the linker on PROTAC selectivity, attention must be paid to the composition of the linker in addition to its length.

In 2019, Burgess and Zhao synthesized various different PROTACs, based on palbociclib and ribociclib as POI ligands, pomalidomide as E3 ligand, and linkers with triazole fragments to connect them [59] (Figure 3). The results showed that the PROTAC-pal-pom (**5**), which used palbociclib as the POI ligand, could induce the degradation of CDK4 and CDK6 in MDA-MB-231 cells. Its DC_{50} (concentration of 50% protein degradation) was 13 and 34 nmol/L, respectively. That is, compound **5** has a more pronounced ability to degrade CDK4 than CDK6. The same year, Rao, Wu et al. published their work on the design of PROTACs for the selective degradation of CDK6 [60]. They also constructed a PROTAC-CP 10 (**6**) containing a triazole-containing linker to connect palbociclib and pomalidomide. Compound **6**, a selective CDK6 degrader, showed good selectivity in U251 cells (DC_{50}(CDK6) = 2.1 nM; DC_{50}(CDK4) > 100 nM) and the compound did not degrade CDK1, 2, 5, 9. Analyzing these two studies, we can notice that when the linker used contains triazole fragments, maybe its length does not need to be as long as the linker, which contains a flexible chain to increase the selectivity for CDK6. On the contrary, shorter linkers seem more conducive to the selective degradation of CDK6.

In addition to CDK4/6, CDK9 is another family member that is getting a lot of attention. Many studies have shown that CDK9 is closely related to many types of cancer and it is crucial to cancer cells' maintenance, growth, metastasis, and chemical resistance [61]. Although many small-molecule inhibitors have also been developed for CDK9 so far, due to the high homology



pal-pom (**5**)
DC_{50}(CDK4) = 13 nmol/L, DC_{50}(CDK6) = 34 nmol/L



CP 10 (**6**)
DC_{50}(CDK6) = 2.1 nmol/L, DC_{50}(CDK4) > 100 nmol/L

Figure 3: CDK4/6 PROTACs with triazole-containing linkers.

of the CDK family, inhibitors always tend to target a variety of CDK members, thereby reducing their specificity as therapeutic agents and chemical probes and results in many adverse effects [62,63]. Based on such background, to target CDK9 with high selectivity to treat diseases related to CDK9 abnormal levels, PROTACs gradually came into view and began to be widely concerned and studied.

In 2020, Chen et al. found that the length of the linker will significantly affect the selectivity of PROTACs to CDK9 [64]. They used pan-CDK inhibitors AT-7519 (**7**) and FN-1501 (**8**) as POI ligands and connected them with CRBN ligands through various linkers. First, the results showed that PROTACs with molecule **7** as POI ligand did not achieve a good degradation effect on CDK9, whereas most PROTACs with molecule **8** as POI ligands showed an excellent ability to degrade CDK9 (**9**–**14**, Figure 4). More importantly, as for the PROTACs with molecule **8** as the POI ligand, linkers containing 8–10 atoms (**9**–**11**) resulted in CDK2/9 dual degraders. When the linker length expands to 11 or 12 atoms (compounds **12**–**14**), the CDK2/9 dual degraders become a selective CDK9 degrader.

Similar to the design of PROTACs for the targeted degradation of CDK6, in addition to the length of linkers, the composition also greatly influences selectivity. In 2018, Bian, Li and co-workers developed PROTACs based on wogonin to degrade CDK9 selectively [65]. In their research, they designed two types of linkers to connect the POI ligand and the E3 ligand. One type is based on alkane chains (Figure 5), and the other additionally includes a triazole fragment (Figure 5). By varying the length of the linker, they synthesized 8 compounds in total. The results showed that PROTACs containing triazole fragments (**16a**–**d**) were more effective in degradation than PROTACs comprising alkane chains only (**15a**–**d**) and they showed higher selectivity towards CDK9 between CDK9 and CDK5. The 10-atoms linker in **16c** showed the best CDK9 degradation activity among these compounds and the selectivity and degradation effect decreased when the linker length was more than 10 atoms.

HDAC

Epigenetics plays a vital role in the occurrence, development, and metastasis of cancer [66-68]. The epigenetics process mainly includes DNA methylation, histone acetylation, histone methylation, and histone phosphorylation [69-71]. Histone acetylation is the most common type, and it is important for regulating the physiological process of normal cells. This process is controlled by histone deacetylases (HDACs) and histone acetyltransferases (HATs) [72]. Now, researchers have found 18 kinds of HDACs in total. They are divided into four categories: HDAC1, 2, 3, and 8 for class I, HDAC4, 5, 7, and 9 for class IIa, HDAC6 and 10 for class IIb, seven sirtuin proteins for class III, and HDAC11 for class IV [73]. Studies have confirmed that the expression and activity of HDACs are closely related to the occurrence of many types of cancers [74-76]. Abnormal expression of HDACs will significantly affect the occurrence, development, and migration of cancer [77-80]. Although an inhibition of HDACs by inhibitors can induce apoptosis and restrain cancer occurrence, the effectiveness of developing HDAC inhibitors to treat cancer is limited because these inhibitors often lack specificity and have toxic side effects [66,81,82]. For example, HDAC6, a class IIB HDAC isoenzyme, is involved in several signaling pathways associated with multiple neurological disorders, various cancers, and immune diseases [83]. A great deal of effort has been devoted to developing HDAC6 inhibitors. However, so far, only two compounds, ACY-1215 (ricolinostat, **17**) and ACY-241 (citarinostat, **18**) have reached clinical trials (Figure 6) [83]. Research has revealed that a variety of HDAC6 inhibitors are not suitable to effectively treat diseases and enter the clinical trials due to a lack of selectivity, which can cause serious side effects [84]. Consequently, developing drugs targeting HDACs specifically to treat cancer is urgent. In this context, PROTACs have become a straw to clutch to target HDACs specifically. In recent years, the design of various PROTACs targeting the HDAC family has become a hotspot. Besides, using what kind of linker to connect the POI ligand and the E3 ligand to improve the selectivity of PROTACs has also become the core problem in designing such molecules.

In 2018, Tang et al. developed the first PROTAC targeting HDAC6 degradation [85]. The study used a pan-HDAC inhibitor as the POI ligand and thalidomide as the E3 ligand to design four PROTACs (Figure 6, **19a**–**d**). These used linkers in these PROTACs all contained triazole fragments, but their lengths differed. Having synthesized the PROTACs, the authors explored the effect of linker length on the selectivity. The results showed that when the linker contained three ethyleneoxy fragments (**19c**), the resulting PROTAC demonstrated the best selective degradation effect on HDAC6 (by detecting the expression level of HDAC1 and HDAC2 as typical examples of class I HDACs, and HDAC4 and HDAC6 as typical examples of class II HDACs). The selectivity of PROTACs with other linker lengths (**19a**, **19b**, **19d**) was not as good as that of compound **19c**.

In 2019, Tang et al. continued on designing PROTACs based on CRBN for selective HDAC6 degradation [22]. They used the selective inhibitor nexturastat A of HDAC6 with the CRBN E3 ligand pomalidomide to design 18 PROTACs. According to the different amino sites on the phthalimide ring of pomalidomide, these PROTACs were divided into two series: a C4-linker series

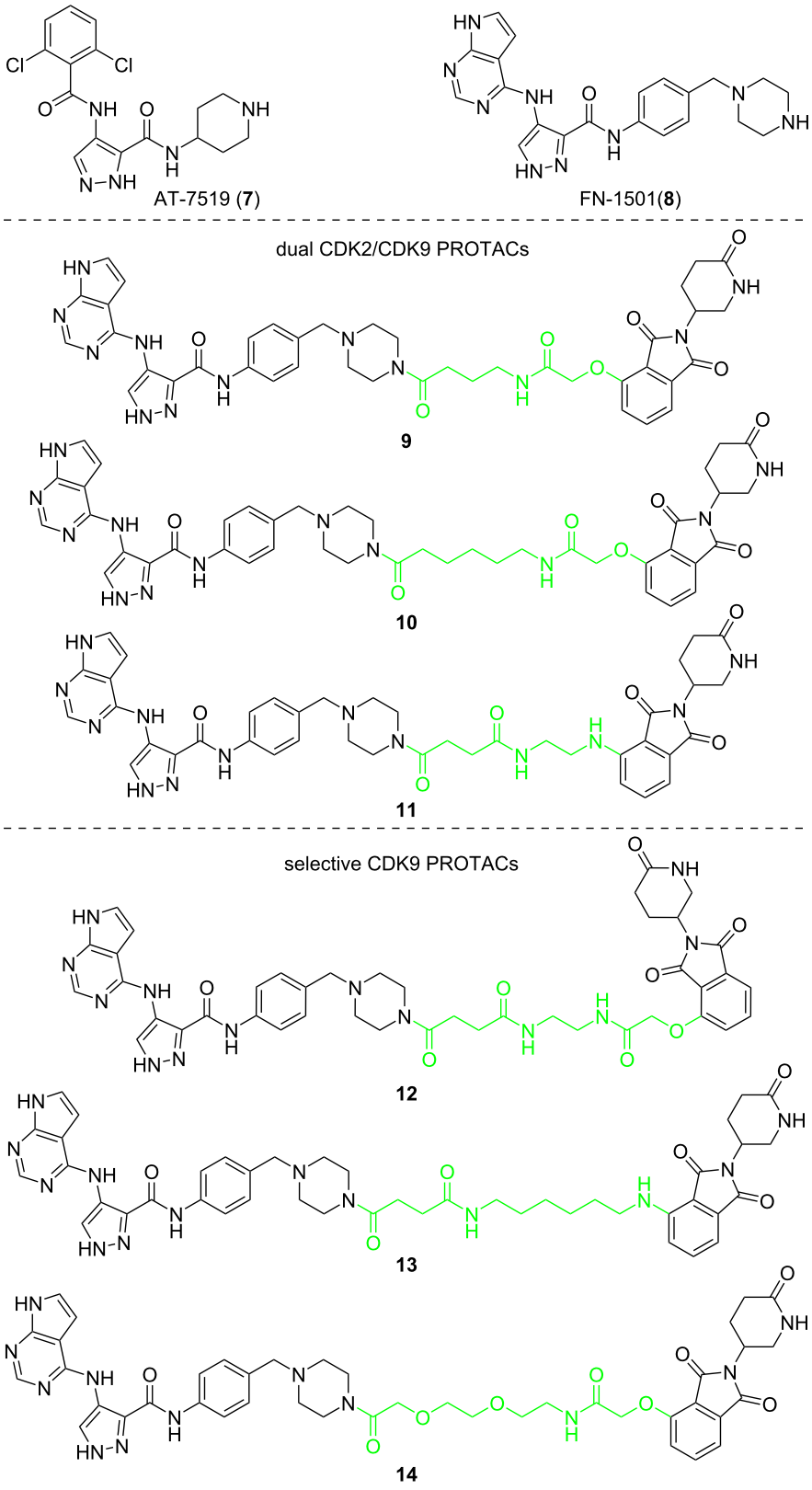
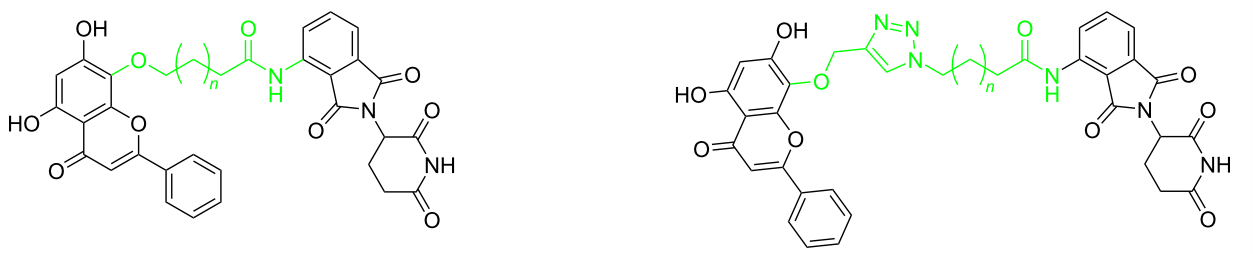


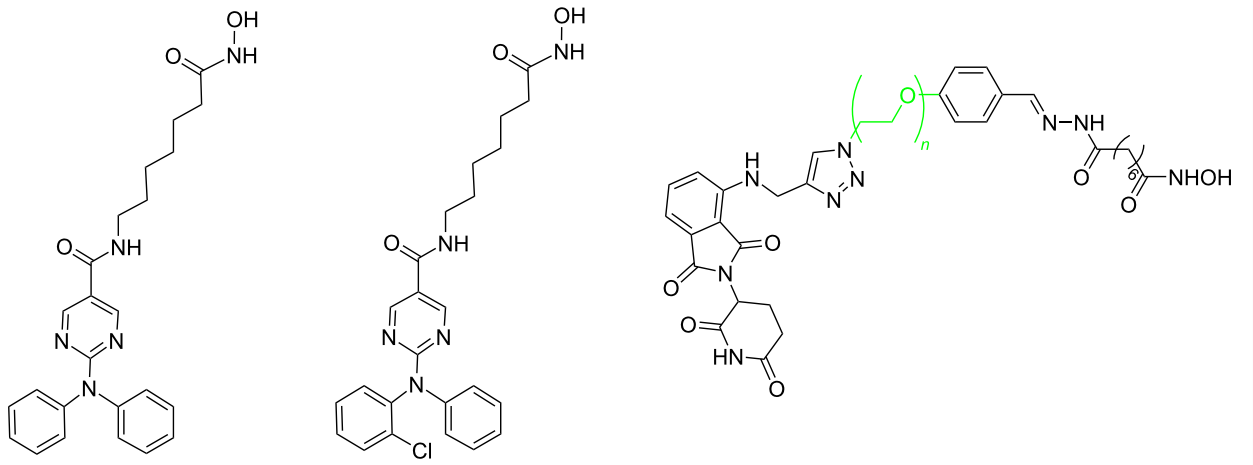
Figure 4: Structures of AT-7519 (**7**) and FN-1501 (**8**), and CDK9 degrading PROTACs based on compound **8** with varied lengths of linkers.



15a–d, *n* = 3, 4, 5, 7

16a–d, *n* = 3, 4, 5, 7

Figure 5: CDK9 PROTACs with alkane chain as linkers or triazole linkers.



ACY-1215 (**17**)

ACY-241 (**18**)

19a–d, *n* = 1, 2, 3, 4

Figure 6: Structures of HDAC6 inhibitors ACY1215 (**17**) and ACY-241 (**18**), as well as the structure of PROTACs **19a–d** synthesized to target HDAC6 degradation.

and a C5-linker series (Figure 7). As for these compounds, the difference is the number of carbon atoms (*n*) between nexturastat A and the triazole fragment or the number of carbon atoms (*m*) between pomalidomide and the triazole fragment.

It was found that compound **20d** (DC_{50} = 1.64 ± 0.24 nM) with a medium-length linker (*n* = 5, *m* = 1) within the C4-linker series achieved the most potent selective degradation (compared to HDAC1, 3, 4) in the subseries **20a–e** (*m* = 1, *n* = 2–6). For the subseries **20f–h** (*m* = 2, *n* = 2–4), when the linker length increased, the molecular degradation efficiency increases, and **20h** (*n* = 4, *m* = 2) was the best (at 100 nM degradation was 78.3–80.1% of HDAC6). In the subseries **20i–k** (*m* = 3, *n* = 2–4), compound **20i** with the shortest linker (*n* = 2, *m* = 3) has been the most effective degrader (at 100 nM degraded 82.1–84.1% of HDAC6) (Figure 7). In the C5-linker series, the selective degradation of the **20l–p** (*m* = 1, *n* = 2–6) series increased with the increase of linker length. Compounds **20o** (*n* = 5, *m* = 1) and **20p** (*n* = 6, *m* = 1) demonstrated similar ability to degrade HDAC6 (**20o** at 100 nM degraded 83.7–85.7% of HDAC6 and **20p** at 100 nM degraded 83.0–84.0% of HDAC6). Members of the subseries **20q** and **20r** (*m* = 3, *n* = 2 and 3) showed relatively low effects on HDAC6 degradation. These results showed that the optimal total number of methylene units in the linker is about 6, and the C4-linker series is slightly stronger than the C5-linker series. Therefore, the difference in degradation selectivity may be related to the ubiquitination sites available to HDAC6 by E3 ligase recruited by the degrader and both, distance and connection sites maybe affect it.

In 2020, Tang et al. continued with reporting studies on the development of PROTACs for the selective degradation of HDAC6 [86]. This time they used VHL as the E3 ubiquitin ligase. They also designed several series of compounds and used linkers with different types and lengths to connect the POI ligand and VHL ligand. Among them, compounds **21a–k** showed good activity. The number of methylene units (*n*) between the POI ligand and the triazole fragment ranged from 2 to 12, and there were four methylene units between the triazole

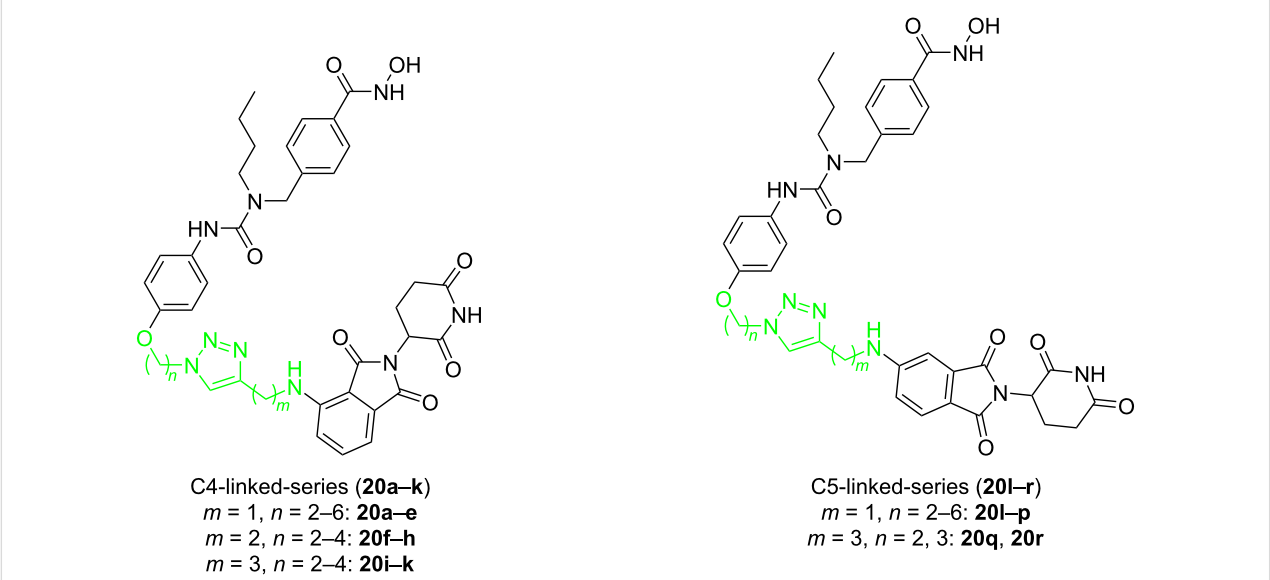


Figure 7: C4-linked-series and C5-linked-series of HDAC6 PROTACs.

ring and the VHL ligand. So these compounds were called the "n + 4" series. Although they also tested the "n + 2", "n + 3" series of compounds, and the "n + 4" series of compounds with the ethyleneoxy unit instead of the methylene unit, they found that these PROTACs were usually not active degraders, so no further research in this direction was carried out. When testing all 11 "n + 4" VHL-based potential degraders, it was found that the molecular selectivity almost steadily increased with the linker length starting from compound **21a** (n = 2) to **21k** (n = 12) (Figure 8). Among them, compound **21j** showed the best selectivity and degradation effect of HDAC6 compared to HDAC1, 2, 3, 4, 7, 8 (DC_{50} = 7.1 nM). A comparison of the above two studies, one can find that the linker length required by degraders based on the VHL ligand is much longer than that of PROTACs based on the CRBN ligand. Thus, it is evident that the linker's effect on PROTAC selectivity is not independent, and it is often closely related to the E3 ligand and even the POI ligand.

In 2021, Fischer and colleagues provided a detailed study on PROTACs for selectively targeting the HDAC protein family [87]. Some previous studies suggested that modifying the length of the linker is an effective strategy to change the selectivity of degraders, which was also confirmed in the study by Fischer et al. [23,43,88]. To systematically explore the effect of the linker length on HDAC degradation selectivity, they synthesized two series of multitarget degraders based on dacinostat that can recruit different E3 ubiquitin ligases (CRBN and VHL). In both series, they designed molecules containing 1–5 ethyleneoxy units as the linker. Then they compared the differences of each series to reveal the key impact caused by the change in linker length. Remarkably, for PROTACs that recruit VHL XY-07-093 (**23**, Figure 9), using shorter or longer linkers can afford PROTACs that selectively degrade HDAC3. However, when the linker contains 4 ethyleneoxy fragments, HDAC degradation activity is completely lost. This is surprising because compounds with similar linker lengths (containing 3 ethyleneoxy

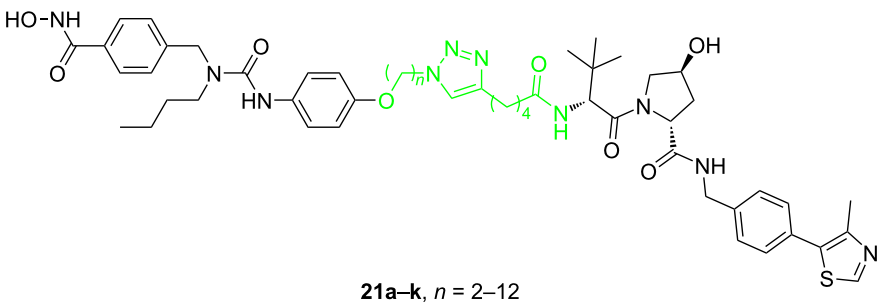


Figure 8: Structures of VHL-based degraders.

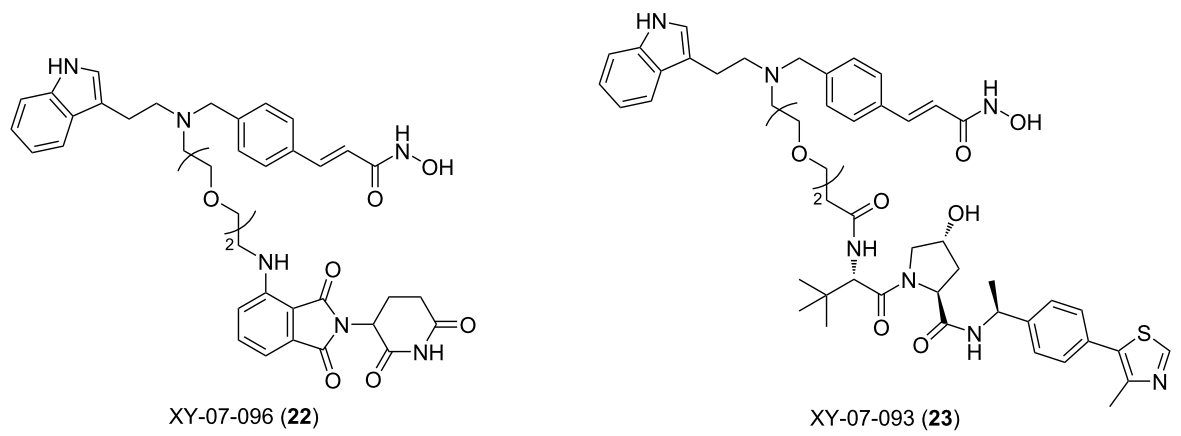


Figure 9: Structures of VHL-based and CRBN-based selective HDAC PROTACs.

fragments and 5 ethyleneoxy fragments) are effective and active degraders. Similarly, when using CRBN as the E3 ligand, they designed the PROTAC molecule XY-07-096 (**22**). The degradation activity of HDAC is completely lost with the increase of the linker (containing 5 ethyleneoxy fragments). In contrast, the PROTAC with a shorter linker is a molecule that can selectively degrade HDAC6 and HDAC8.

p38 MAPK

The p38 mitogen-activated protein kinase (MAPK) family comprises four members: p38α (MAPK14), p38β (MAPK11), p38γ (MAPK12), and p38δ (MAPK13), with p38α being widely expressed and the most abundant member in almost all cells [89]. There is a high degree of homology among the p38MAPK family members with p38α and p38β having 75% homology, p38γ and p38δ having 62% and 61% homology with p38α, respectively, and p38γ and p38δ having 70% homology [90-92]. The p38MAPK family has been widely studied as a potential therapeutic target for the past two decades. Many studies have shown that p38α, p38β, p38γ, and p38δ all play a crucial role in cellular processes related to cancer and inflammatory diseases [93-96]. So developing corresponding small-molecule drugs with the p38MAPK family as a key target is necessary. However, up to now, no small-molecule inhibitors targeting the p38MAPK family have been approved by the FDA, since every effort has failed in clinical trials [97-99]. The reasons for the failure are rarely discussed but one of the factors is believed to be due to the inhibition of several p38MAPK proteins [98]. Therefore, how to specifically target p38MAPK family members has become a vital issue in the development of drugs. In this regard, PROTAC again shows its unique advantages in selectivity for highly homologous proteins compared with small-molecule inhibitors [100].

In 2019, Crews et al. developed PROTACs, able to specifically degrade p38α and p38δ [23]. They used foretinib as the POI ligand and two E3 ligands with different structures to target VHL to explore the influence of the direction of VHL recruitment of PROTACs on selectivity (termed the “amide series” and “phenyl series” individually). Simultaneously, they used four linkers of different lengths to connect the E3 ligand and the POI ligand to explore the influence of linker fragments on molecular selectivity (Figure 10).

As for the “amide series”, the authors discovered that in MDA-MB-231 human breast cancer cells, the 10-atom and 11-atom

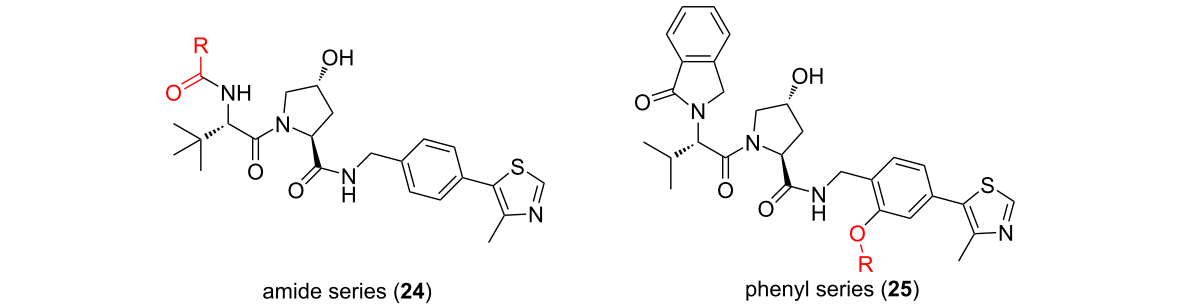


Figure 10: Structures of the “amide series” and “phenyl series” PROTACs studied by Crews et al.

linker PROTACs (SJF-6693 (**26**) and SJF-6690 (**27**)) were not selective and degraded both p38α and p38δ nonspecifically with DC_{50} < 100 nM (Figure 11). However, the 12-atom and 13-atom linker PROTACs SJF-8240 (**28**) and SJF-α (**29**) can selectively degrade p38α (Figure 11). Compound **29** degraded p38α with a DC_{50} of 7.16 ± 1.03 nM and maximum degradation, D_{max} of 97.4%, but the effect on the degradation of p38δ was much worse (D_{max} = 18%, DC_{50} = 299 nM). Other p38 isoforms (β and γ) were not degraded when the concentration was up to 2.5 μM.

On the contrary, PROTACs with the long linker in the "phenyl series" showed almost no ability to degrade the p38 subtype (**30**, **31**, **32**, **33**, Figure 12). When the linker length in the "phe-

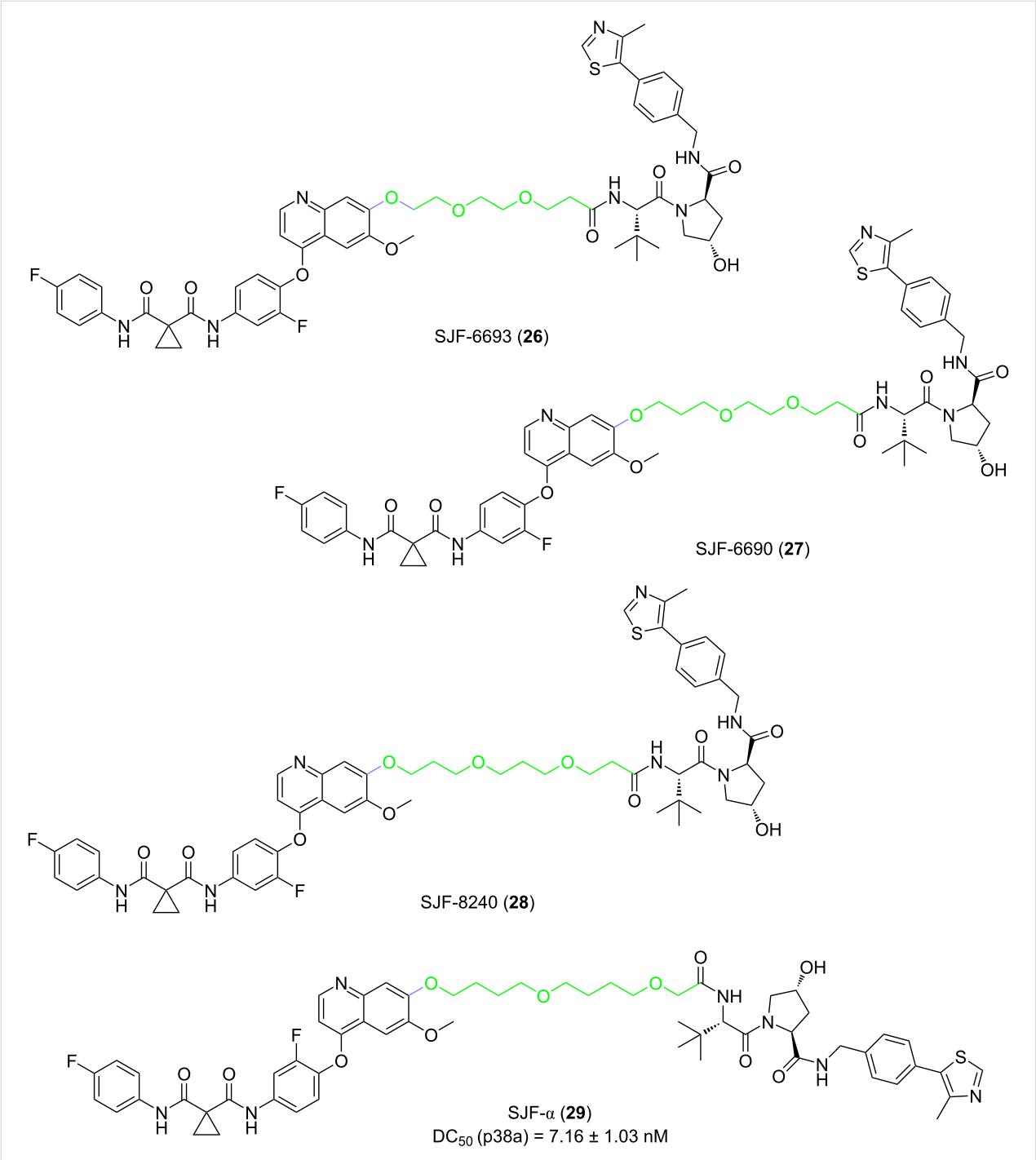


Figure 11: Structures of the “amide series” PROTACs.

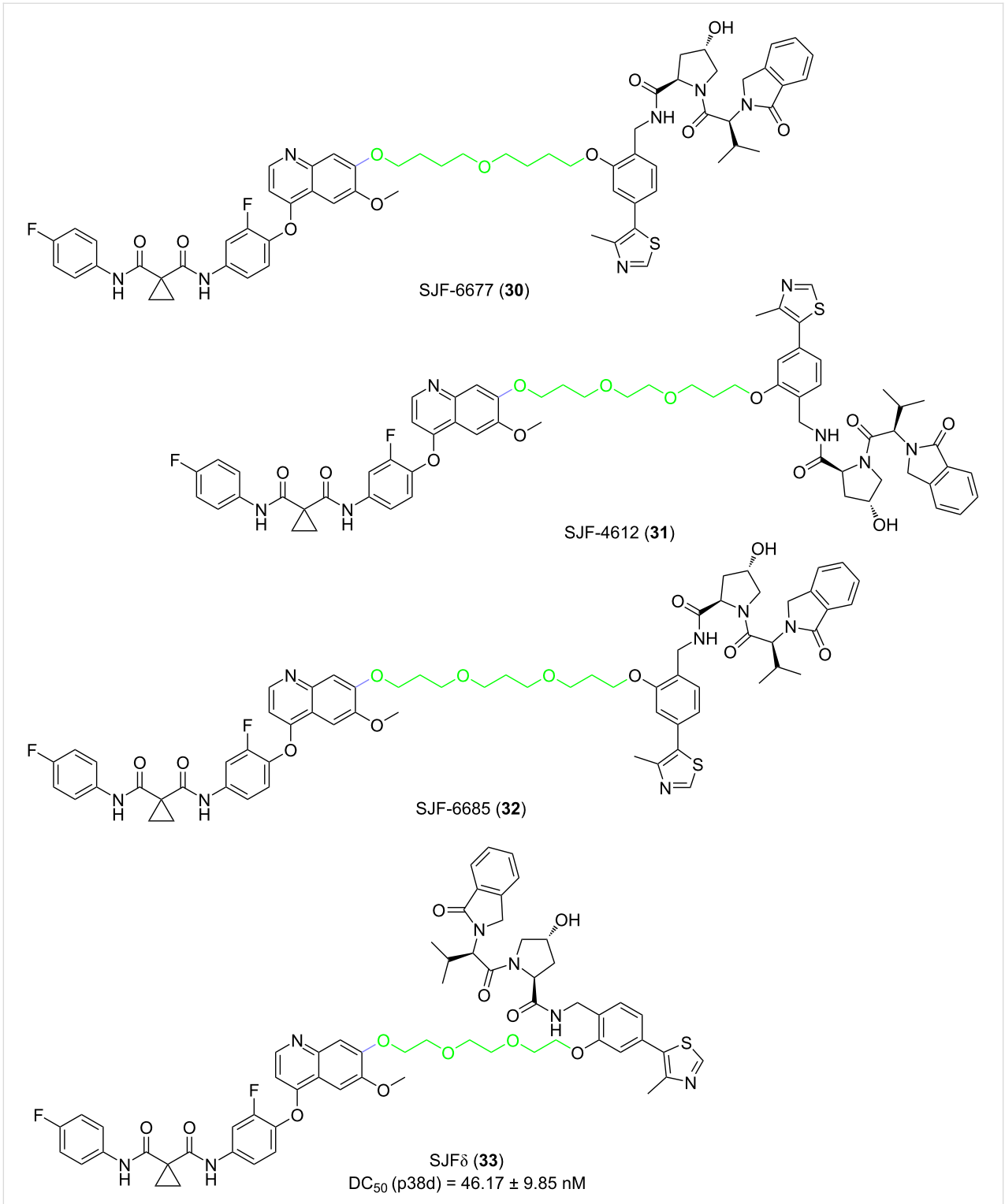


Figure 12: Structures of the “phenyl series” PROTACs.

nyl series" was reduced to 10 atoms, the resulting PROTAC showed a robust selective degradation of p38δ. Compound **33** degraded p38δ with a DC_{50} of 46.17 ± 9.85 nM and D_{max} of 99.41 ± 3.31% but did not degrade α, β, or γ at all.

In some cases, even changing single carbon atoms of the linker segment may cause a surprising change in the degradation selectivity of PROTACs. For example, compound **28** shows submicromolar degradation of two p38 subtypes. Once the

linker contains an additional carbon atom, the resulting PROTAC **29** degrades p38α at a nanomolar concentration and p38δ at a micromolar concentration. Similarly, the 10-atom-linker-containing compound **33** can almost completely degrade p38δ, but in the degradation of p38α, it is restricted. In contrast, the structure of compound **30** having additional carbon atoms added to the linker, resulted in less than 50% degradation of both subtypes at the maximum efficiency. It can be seen that the length of the linker is crucial for the selective degradation of PROTACs. The proper linker length enables PROTACs to distinguish protein families with more than 60% homology. Nevertheless, this characteristic is unimaginable for small-molecule inhibitors.

BET

Bromodomain and extraterminal domain (BET) proteins are epigenetic readers. They comprise the ubiquitously expressed BRD2, BRD3, and BRD4 and the testicular-specific expressed BRDT [101]. The function of the BET protein family is mainly to regulate gene transcription by recognizing acetylated lysine residues on histones [102,103]. The imbalance of BET protein activity, especially BRD4, is closely related to cancer and inflammatory diseases, so the BET protein family has become an attractive drug target [102]. Because BET proteins play an essential role in various diseases, many small-molecule inhibitors against BET have been developed [104]. However, the clinical studies for the BET inhibitors were not successful because the compounds did not show good antitumor activity [105]. One possible reason for the failure of these trials is that most of the BET inhibitors that entered the clinical trials are pan-BET inhibitors so they can bind multiple BET protein members [106]. However, due to the varying expression levels of BET protein members in different tumor types, these inhibitors failed in displaying good therapeutic effects. Moreover, due to the lack of highly selective BET inhibitors, the mechanism of cancer signaling pathways related to BET proteins is not yet clear. Therefore, it is necessary to develop drugs with high selectivity towards specific BET members.

In order to solve this problem, Ciulli et al. developed a molecule that selectively degrades BRD4 using the PROTAC technology in 2015 [107]. They used JQ1 (**34**), a non-selective BET protein family inhibitor, as the POI ligand and selected linkers of different lengths composed of PEG chains with three or four ethyleneoxy units to connect **34** with the VHL ligand, thus designing four PROTACs. The results showed that MZ1 (**35**, Figure 13), whose linker is composed of three ethyleneoxy units in the structure, has an excellent ability to degrade BRD4 selectively (compared to BRD2 and BRD3). This indicates that the linker affects the selectivity of PROTACs to degrade BET protein family members. The evidence that more directly proves this view comes from another study reported by Ciulli et al. in 2017 [41]. In this study, they analyzed the ternary composite structure of PROTACs for the first time and explained the selectivity of compound **35** in detail.

In 2020, Ciulli et al. continued to develop a PROTAC with a large ring linker (macro-PROTAC-1, **36**, Figure 14) based on their previous research [108]. This PROTAC also selectively degrades BRD4, whereas BRD2 and BRD3 can only be degraded at high concentrations. By comparing the structures of the two molecules **35** (Figure 13) and **36**, it becomes clear, that the main difference between them is that in compound **36** the POI ligand and VHL ligand are connected via a conformation-restricting macrocyclic linker. Although this seems to be a minor change, the comparison of these two studies reveals that compound **36** is able to cooperatively form ternary complexes with the BD2 domains of BRD4, BRD2, and BRD3. The cooperativity factor α, a factor reflecting effects of PPIs on ternary complex formation, was found to be $\alpha BRD2^{BD2}$ = 10.5, $\alpha BRD3^{BD2}$ = 9.5, and $\alpha BRD4^{BD2}$ = 4.0. Meanwhile, the study did not observe a synergy between compound **36** and the

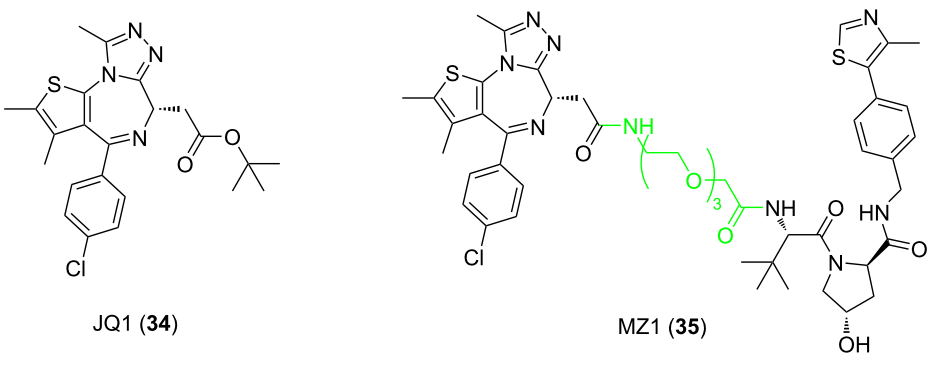


Figure 13: Structures of JQ1 (**34**) and MZ1 (**35**).

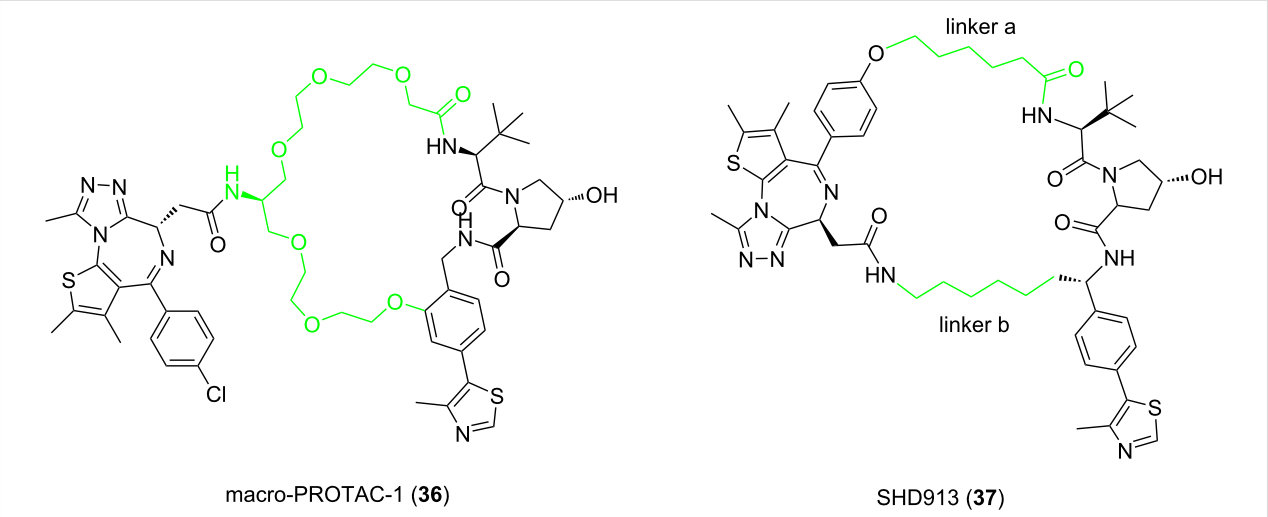


Figure 14: Structures of **macro-PROTAC-1** and **SHD913**.

first bromodomain (BD1) of BRD2, BRD3, and BRD4 ($\alpha BRD2^{BD2}$ = 0.7, $\alpha BRD3^{BD2}$ = 0.9, $\alpha BRD4^{BD2}$ = 0.8). This result is in sharp contrast to compound **35**, which forms synergistic complexes with all bromodomains of the BET protein family ($\alpha BRD2^{BD1}$ = 2.9, $\alpha BRD2^{BD2}$ = 2.3; $\alpha BRD3^{BD1}$ = 3.5, $\alpha BRD3^{BD2}$ = 10.7, $\alpha BRD4^{BD1}$ = 2.3, $\alpha BRD4^{BD2}$ = 17.6). These results are promising because all members of the BET protein family contain two highly homologous bromodomains (BD1 and BD2) with 49% homology [88,109]. From here, it is clear, that the change of linker fragments can not only play a decisive role in the differentiation of BET protein family subtypes but also allows distinguishing BD1 and BD2 domains in the different subtypes. This discovery shows that the PROTAC technology offers substantial advantages in terms of selectivity and that the design of suitable linkers for the application of this technology to the selective degradation of highly homologous protein families is of great significance.

In 2024, Ding et al. reported the "head-to-tail" macrocyclic PROTAC, SHD913 (**37**) [110]. Based on the crystal structure of the $BRD4^{BD2}$:MZ1:VHL ternary complex, the authors designed two short-chain linkers for macrocyclization, targeting two critical distances (4.6 Å and 7.1 Å) between the warhead molecule **34** and the VH032 ligand. Through systematic optimization of linker lengths, it was found that the molecule achieved optimal activity when the linker a consisted of six methylene groups and linker b consisted of three methylene groups. The resulting products **37** exhibited DC_{50} values of 7.7 nM and 5.0 nM for the long and short isoforms of BRD4 in PC-3 cells, respectively (Figure 14). Regarding selectivity, compound **37** demonstrated a significant degradation advantage for BRD4 over other BET family members: Western blot analysis showed that BRD4 was effectively degraded at a 30 nM treatment concentration, while BRD2/3 remained largely unaffected. Global proteomic analysis confirmed that BRD4 was the most significantly downregulated protein; the downregulation of BRD3 was markedly lower, and no significant change was observed for BRD2. Furthermore, while the DC_{50} values for BRD4 were in the single-digit nanomolar range, the DC_{50} values for BRD2/3 were both well above 100 nM. Due to the conformational constraint imposed by the macrocyclic linker, compound **37** did not exhibit a "hook effect" even at concentrations as high as 60 μM, whereas the linear control molecule **35** showed a decline in activity at 30 μM. NanoBRET and ITC experiments indicated that compound **37** induces a stable protein–protein interaction interface between $BRD4^{BD2}$ and VHL through linker-mediated conformational locking, achieving a cooperativity factor (α) of 117, which is significantly higher than those of compounds **35** and **36**. Its co-crystal structure revealed that linker b forms hydrophobic interactions with the ZA loop of $BRD4^{BD2}$ and stabilizes the PPI interface via a conserved salt-bridge network, while linker a acts as a "hinge" to bring the two proteins into proximity. This study systematically elucidated how macrocyclic linker length and composition regulate PROTAC selectivity, providing a reference design strategy for the precise targeting of highly homologous protein families through linker engineering.

The above highlighted examples summarized how to use the PROTAC technology to achieve highly selective degradation for some selected highly homologous proteins by changing the linker. Based on the results, it can be concluded that for different highly homologous proteins, there is no fixed conclusion about what kind of linker can better improve the selective degradation ability. What's more, it is difficult to answer whether more attention should be paid to the length of the linker or dif-

ferent linker compositions when designing PROTACs. Facing these problems, it is often needed to analyze, design, and explore on a case by case basis. Even, when using different E3 ligands or POI ligands to explore the influence of the linker fragments, completely opposite conclusions may be drawn. This indicates that although the linker is vital for the design of PROTACs, the impact of the POI ligand and E3 ligand need to be considered. PROTACs function as a whole, not just one or two parts and therefore, the POI ligands, linker, and E3 ligands cannot be analyzed separately. Of course, this does not mean that examinations and discussions of individual linkers are pointless. A detailed knowledge of the SAR of a linker fragment of PROTACs for different highly homologous protein families allows to quickly and specifically select the linker segment with the best selectivity and activity to synthesize PROTACs in the future. It also reduces the time required for exploration during the development of highly selective PROTACs. In addition, it can give full play to the remarkable selectivity of the PROTAC technology compared with small-molecule inhibitors.

Influence of E3 ligands on the selectivity of PROTACs in highly homologous protein families

It is well known that PROTAC is a technology for the specific degradation of proteins through the ubiquitin-proteasome system [111]. Ubiquitination refers to the process of covalently binding ubiquitin to target proteins under the catalysis of a series of enzymes [112]. The ubiquitinated protein can be recognized explicitly by the proteasome to achieve degradation. The whole process requires the participation of three enzymes: ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin-ligase enzymes (E3) [113-115]. So far, two E1 enzymes, about 40 E2 enzymes, and more than 600 E3 enzymes have been found in the human proteome [116]. As the specificity of the substrate protein in the ubiquitination is determined by E3 ubiquitin ligase, the E3 enzyme is a colossal protein family compared with the E1 enzyme and E2 enzyme. However, although there are many kinds of E3 enzymes, the most targeted combination of E3 ligands when designing PROTACs are Von-Hippel-Lindau (VHL), cereblon (CRBN), "inhibitor of apoptosis" protein (IAP), and MDM2 [117]. Among these four E3 ubiquitin ligases, VHL and CRBN are the most commonly used E3 ubiquitin ligases when designing PROTACs. This is because VHL ligands and CRBN ligands often have several favorable characteristics: (1) specific, strong, biophysically validated binding affinity for their target E3 ligases; (2) acceptable physicochemical characteristics such as molecular weight, solubility, lipophilicity, lack of metabolic hot spots; and (3) well characterized structural information of their binding modes [118,119].

CRBN, one of the most commonly used E3 ligand to target E3 enzymes in PROTAC molecular design, has been successfully used to target and degrade more than 30 different proteins [119] (Figure 15). These proteins include proteins related to various

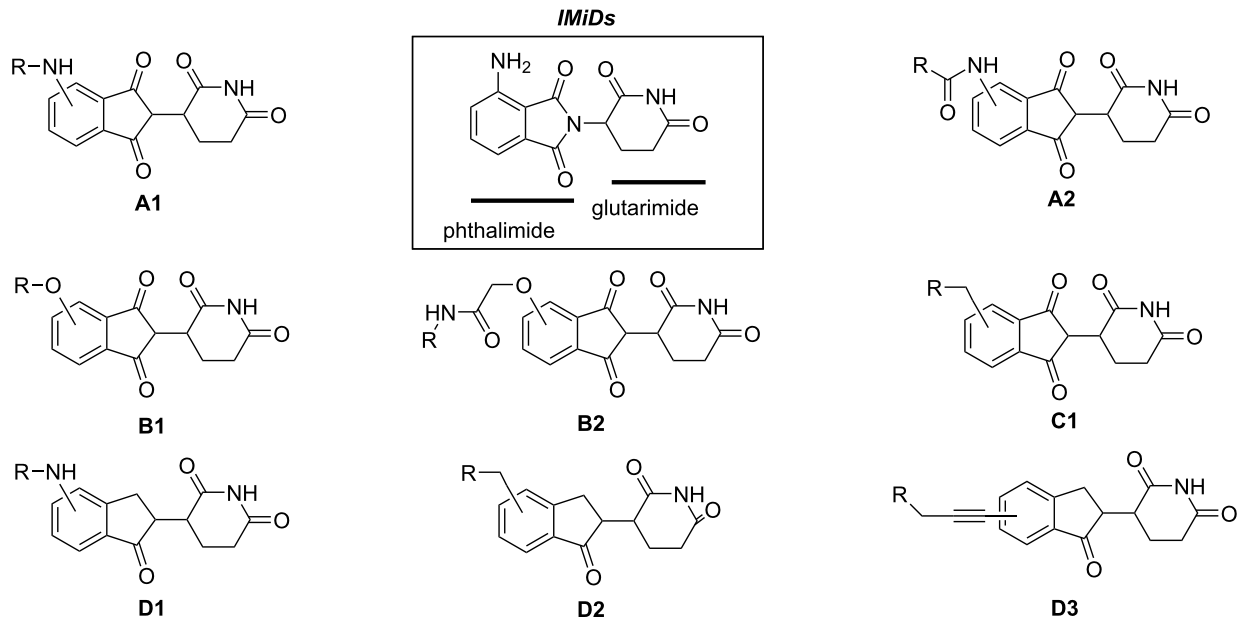


Figure 15: Commonly utilized thalidomide-derived CRBN ligands and possible linker attachment styles. **A1**, **A2**: pomalidomide derivatives; **B1**, **B2**: 4-hydroxythalidomide derivatives; **C1**: alkyl-type attachment to thalidomide; **D1**–**D3**: lenalidomide derivatives.

cancers, immune diseases, and even neurodegenerative diseases such as Tau [120-122]. VHL (Figure 16), another E3 ubiquitin ligase frequently used in PROTAC molecular design, has also been successfully used to target and degrade more than 20 different proteins [123].

It is precisely because more and more E3 ligands are being developed that the design of PROTACs has become more rich and diverse. With the deepening of research, it is found that different E3 ligands in the structure of PROTACs containing the same POI ligand and linker fragments will also induce protein degradation at different levels [117]. What is more important, there will be some differences in the degradation selectivity of POI. Therefore, this part of the review focusses on the influence of different E3 ligands in PROTACs on the overall molecular degradation selectivity.

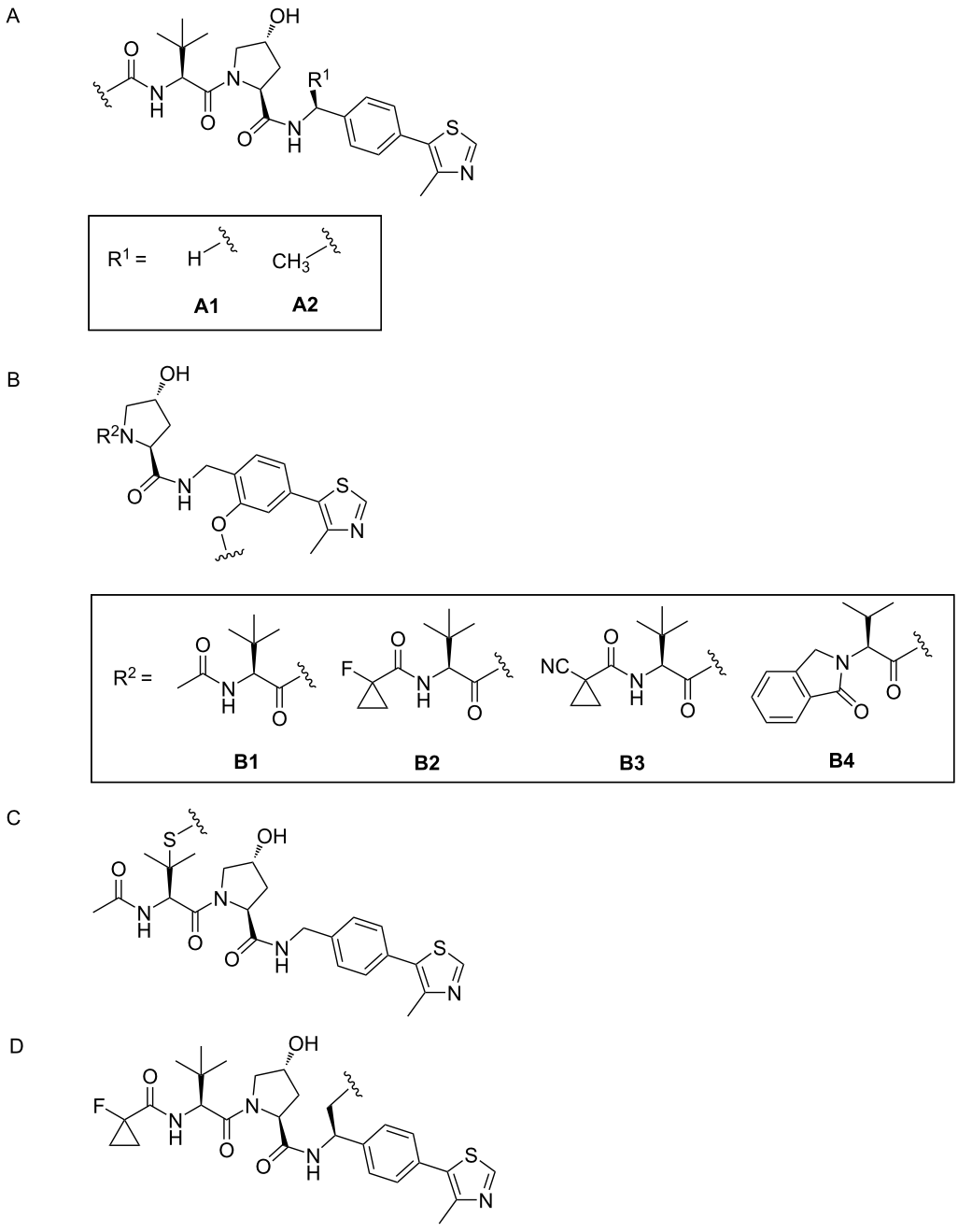


Figure 16: VHL ligands frequently used in PROTACs. Linker attachment options are represented with curly bonds and are: (A) via an amide bond after *tert*-leucine; (B) phenolic linkage point at the benzene ring; (C) via a thioether at the left-hand side amino acid; and (D) via the benzylic methylene group.

E3 ligands target different E3 ubiquitin ligases

During the design of PROTACs, the selected E3 ligands often target CRBN or VHL, however, in current research there is a preferential use of CRBN as E3 ubiquitin ligase instead of VHL. This is because ligands targeting CRBN have better drug-likeness properties, including lower molecular weight, fewer hydrogen bond donors in their structures, and fewer rotatable bonds [124]. However, in addition to influencing the physical and chemical properties of molecules, according to many relevant research results, it was found that when the selected E3 ubiquitin ligases are different, the selectivity of PROTACs indeed show significant differences. As a result, different E3 ligands targeting different E3 enzymes will have a vital impact on the selectivity of PROTACs in highly homologous protein families.

BCR-ABL and c-ABL: Currently, the vast majority of chronic myeloid leukemia and about 20–30% of acute lymphoblastic leukemia are caused by chromosome translocation between chromosomes 9 and 22 [125]. The gene rearrangement leads to the expression of the oncogenic fusion protein, BCR-ABL and the loss of autoinhibition of the c-ABL kinase domain in BCR-ABL is the main cause of cancer [126]. Imatinib mesylate was the first tyrosine kinase inhibitor (TKI) targeting BCR-ABL [127,128]. It can competitively bind to the ATP binding site of c-ABL to inhibit the functions of c-ABL and BCR-ABL, thereby inhibiting cell proliferation [125]. Because long-term treatment with imatinib will lead to drug resistance in CML patients, the second generation of TKI was subsequently developed for treatment [129]. Although TKIs targeting BCR-ABL have achieved significant therapeutic effects, CML patients often need to take such drugs for life to control the disease's further deterioration. Based on these circumstances, if the PROTAC technology can be used to target the degradation of the BCR-ABL protein, it may maintain a good therapeutic effect for CML and reduce the time for CML patients to take medicine. Therefore, to achieve good degradation of BCR-ABL by PROTACS has been the focus of intensive pharmaceutical chemistry research. Interestingly, it was found that using different POI ligands and E3 ligands combinations can achieve efficient selectivity for BCR-ABL and c-ABL. This discovery is of great significance for the subsequent development of PROTAC targeting other proteins with high homology.

In 2016, Crews and colleagues designed different PROTACs by connecting BCR-ABL TKI dasatinib (**38**) and bosutinib (**39**), which target the binding of the c-ABL kinase domain, with cereblon (CRBN) ligand **40** or Von Hippel-Lindau (VHL) E3 ligand **41** through a linker (Figure 17) [126]. Interestingly, it was found that when molecule **38** was combined with pomalidomide to recruit CRBN, the resulting dasatinib-CRBN PROTAC not only retained its ability to induce c-ABL degradation (1 μM > 85%) but also induced BCR-ABL degradation (1 μM > 60%). However, when molecule **38** was combined with

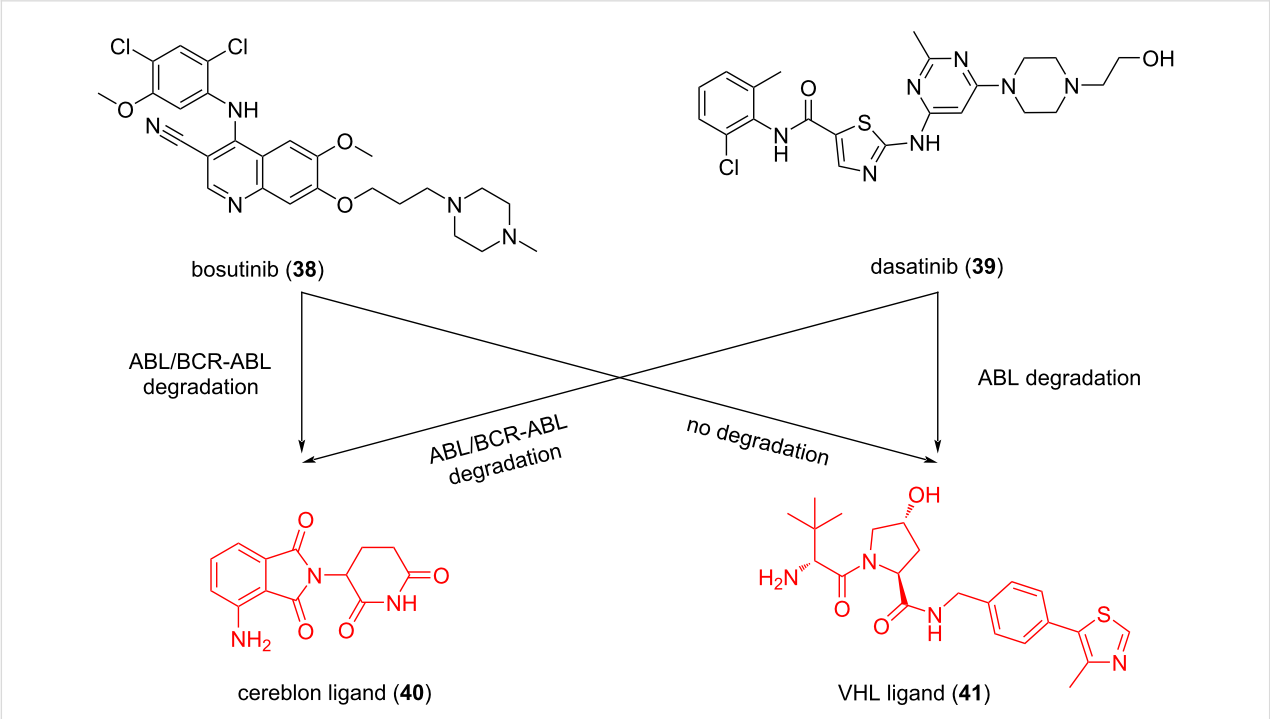


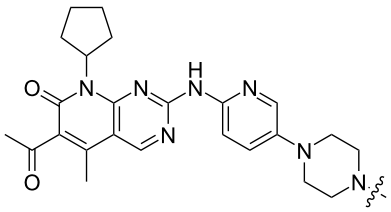
Figure 17: Varying the inhibitor warhead and the recruited E3 ligase permits targets to be accessed for degradation.

a VHL ligand to recruit VHL for degradation, the dasatinib-VHL PROTAC only degrades c-ABL and has no degradation effect on BCR-ABL. Meanwhile, the study also found that the PROTACs targeting VHL can still effectively bind and inhibit c-ABL and BCR-ABL in cell culture. This shows that the failure of these PROTACs to achieve degradation cannot be attributed to the loss of binding affinity. In addition, when molecule **39** is combined with pomalidomide to recruit CRBN, the formed PROTACs can degrade c-ABL and BCR-ABL. However, when molecule **39** is combined with a VHL ligand to recruit VHL for degradation, the bosutinib-VHL PROTAC does not display any degradation function. The authors believed that the proximity between the E3 ubiquitin ligase and the target protein is important for specific lysine residues required for ubiquitination degradation. They speculated that to achieve the selective degradation of BCR-ABL, specific E3 ligands, linkers, and POI ligands are needed when designing PROTACs to make POI and E3 ubiquitin ligases spatially close to each other to achieve ubiquitination. Thus, although many of the current PROTAC optimizations focus on linker fragments, the E3 ligands are also crucial for PROTACs to play an overall role in drug efficacy and selectivity.

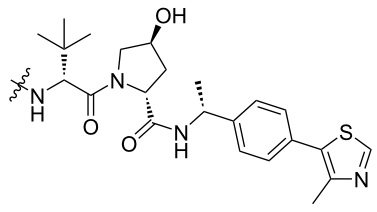
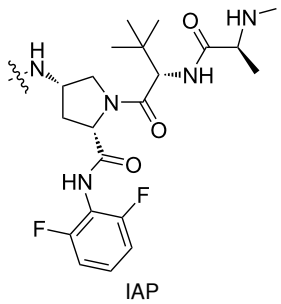
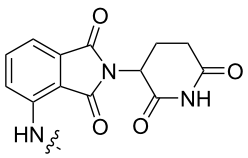
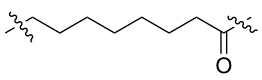
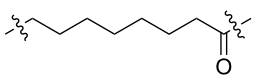
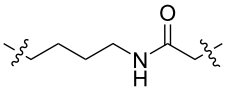
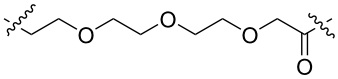
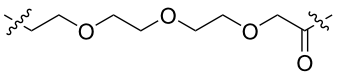
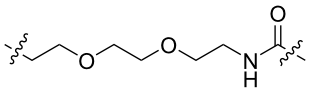
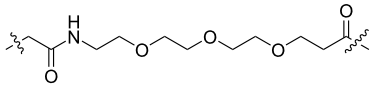
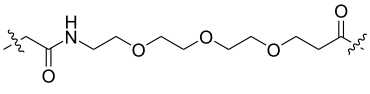
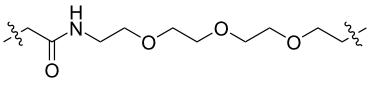
CDK: CDK is a protein family with many highly homologous members. In recent years, PROTAC technology has been widely used in the specific degradation of CDK proteins to treat related diseases. What's more, as mentioned above, achieving selectivity in the CDK protein family is crucial for certain diseases, thus some studies have explored the influence of E3 ligands differences on the selective degradation of CDK protein family members.

In 2020, Benowitz et al. designed a series of PROTACs, including VHL ligands, CRBN ligands, and IAP ligands, to explore the difference between the E3 ligands of PROTACs for the selective degradation of CDK4 and CDK6 (Table 1) [57]. The results showed that PROTACs based on CRBN, VHL, and IAP ligands can degrade CDK4 and CDK6 and show a specific

Table 1: CDK4/6 PROTACs based on VHL, IAP, and CRBN ligands.



palbociclib-based target binder

ligand	E3 ligase ligand VHL	E3 ligase ligand IAP	E3 ligase ligand cereblon (CRBN)
	 VHL	 IAP	 CRBN
linker			
linker			
linker			

selectivity for CDK6. Moreover, the PROTACs using CRBN ligands have better selectivity, while the PROTACs using VHL and IAP ligands display poor selectivity.

This conclusion has also been demonstrated by Wu, Rao et al. [60]. They found that PROTACs designed with other E3 ligands (VHL, cIAP, and MDM2) except CRBN could not effectively degrade CDK6 at 1 μM. In contrast, PROTACs designed with CRBN ligands were effective degraders of CDK6. However, unfortunately, only such a phenomenon was mentioned in their study and no detailed research or discussion was conducted. We look forward to conducting more relevant research on this interesting phenomenon in the future.

In the same year, Calabretta, Salvino and co-workers also reported their work on developing PROTACs for the selective degradation of CDK6 [130]. In order to treat Philadelphia chromosome-positive acute lymphoblastic leukemia (Ph^+ALL), they designed a series of PROTACs based on different E3 ligands and linkers. The results showed that PROTAC YX-2-107 (**43**), which recruited CRBN to degrade CDK6, had the best effect (DC_{50} = 4 nM, IC_{50} = 4.4 nM in BV173 cells). Compared with YX-2-107, YX-2-233 (**42**) is structurally different from compound **43** only in the E3 ligand fragment. The E3 ligand used in **42** is MDM2 (Figure 18). The results showed that the degradation ability of compound **42** decreased significantly, and it became a CDK4/6 dual degrader, which did not show high selectivity for CDK6. It can be seen that the selection of appropriate E3 ligand fragments may not only result in good selectivity for highly homologous target proteins but also effectively enhance the degradation effect of PROTACs.

Based on the above results, CRBN ligands seem to be a better choice when developing highly selective PROTACs targeting CDK6 degradation. However, in 2020, Krönke, Gütschow and colleagues reported a study denying this conclusion [131]. They selected compound **44** as the lead compound and replaced the E3 ligand CRBN with VHL to synthesize compound **45**. The compounds were tested in MM.1s cells to evaluate their activity. The results showed that although the degradation efficiency

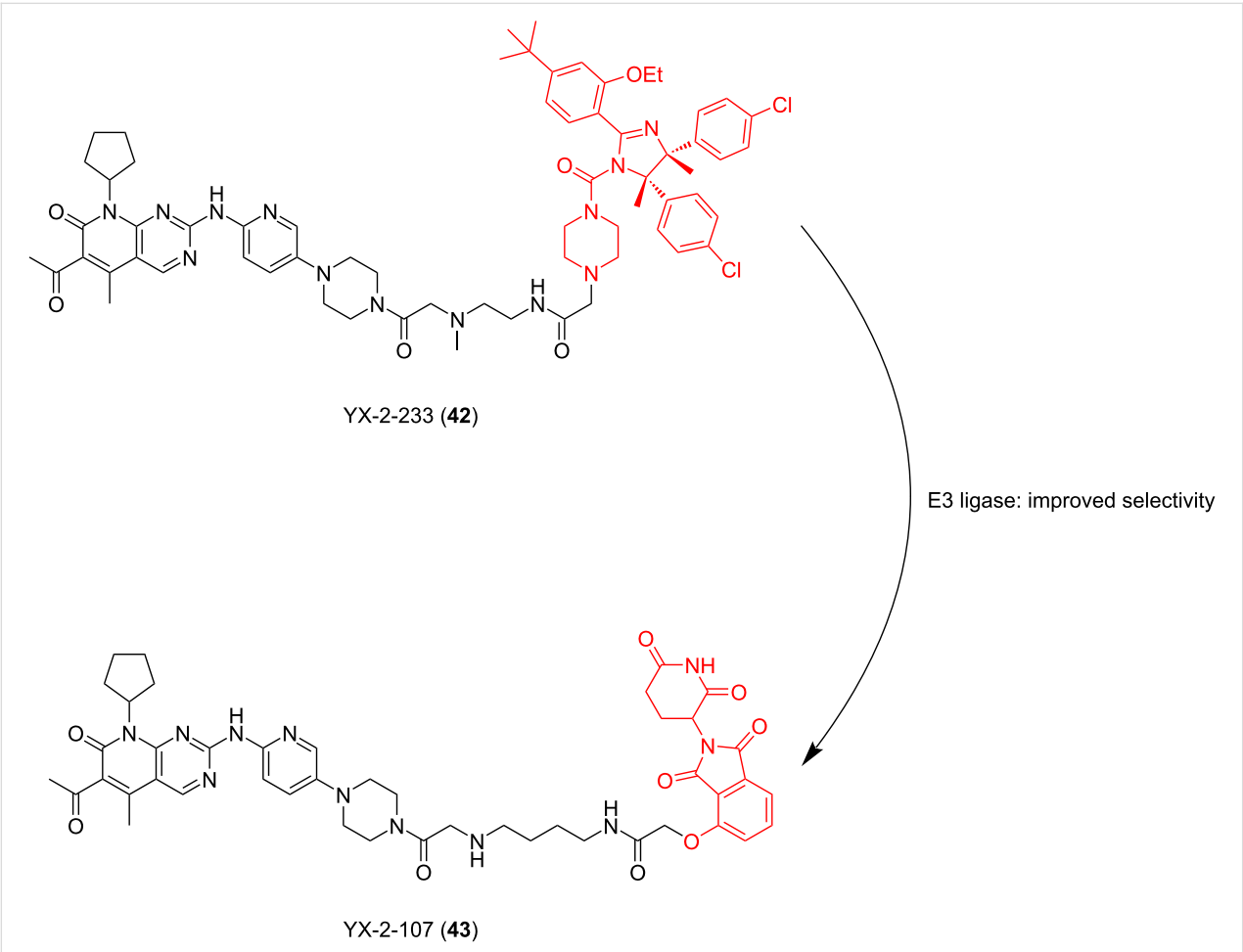


Figure 18: Structures of YX-2-233 (**42**) and YX-2-107 (**43**).

of compound **45** was close to that of compound **44**, compound **45** showed significant advantages in selectivity for CDK4 and CDK6 (compound **45**: $DC_{50\ for\ CDK4}/DC_{50\ for\ CDK6\ (at\ 0.1\ \mu M)}$ = 19; compound **44**: $DC_{50\ for\ CDK4}/DC_{50\ for\ CDK6\ (at\ 0.1\ \mu\ M)}$ = 4.9). Form these results, it can be seen that when designing PROTACs for different highly homologous target proteins, it is often necessary to consider the whole molecule rather than only one segment (Figure 19). Maybe when different linkers are used to connect the POI and E3 ligands, the difference in the E3 ligands may lead to the opposite result of selectivity.

SGK: Serum-glucocorticoid-induced protein kinase (SGK) plays a key role in mediating resistance to phosphoinositide 3-kinase (PI3K)/Akt inhibition in breast cancer cells [132]. It has been reported that different ATP competitive inhibitors have similar affinity for all SGK isoforms [133,134]. Due to the high homology and structural similarity of catalytic domains between different SGK subtypes, no specific inhibitors have been developed [135]. To cope with this challenge and to better treat corresponding diseases, Alessi, Ciulli et al. used the PROTAC technology to develop effective and highly selective SGK3-PROTACs [136] (Figure 20). In the study, they first designed two types of POI ligands, Sanofi 308-R (**46**) and Sanofi 290-R (**47**), which can inhibit the kinase activity of SGK3. However, they lack specificity and have a particular inhibitory effect on S6K1. Then they connected the two POI ligands to VHL ligand VH032 (**48**) and cereblon ligand pomalidomide (**49**) through a medium-length linker composed of three ethyleneoxy units to form PROTACs. The results showed that only DAT1 (**50**) could significantly reduce the expression of SGK3 and did not affect SGK1, SGK2, and S6K1. Other PROTACs had no good degradation effect and high selectivity for highly homologous SGK subtypes. Although there was no further study on the effect of E3 ligands on the selectivity of SGK subtypes, it can be concluded that when the POI ligands and linkers of PROTACs are uniform, the E3 ligands may play a decisive role in the high selectivity of the molecules.

HDAC: Regarding the influence of linkers on the selectivity of PROTACs in highly homologous protein families, HDAC has been also a widely studied highly homologous protein family. Up to now, about 100 HDAC-PROTACs targeting CRBN, VHL, or IAP E3 ubiquitin ligases have been reported [87,137]. In 2021, Fischer et al. reported their studies on the difference between PROTACs based on CRBN, VHL, and IAP ligands for HDAC degradation (Figure 21). In addition, when designing PROTACs based on VHL ligands, they also studied the influence of two attachment points of VHL ligands on the molecular degradation effect [87]. It has been previously reported that the attachment point of the VHL ligand has a significant impact on

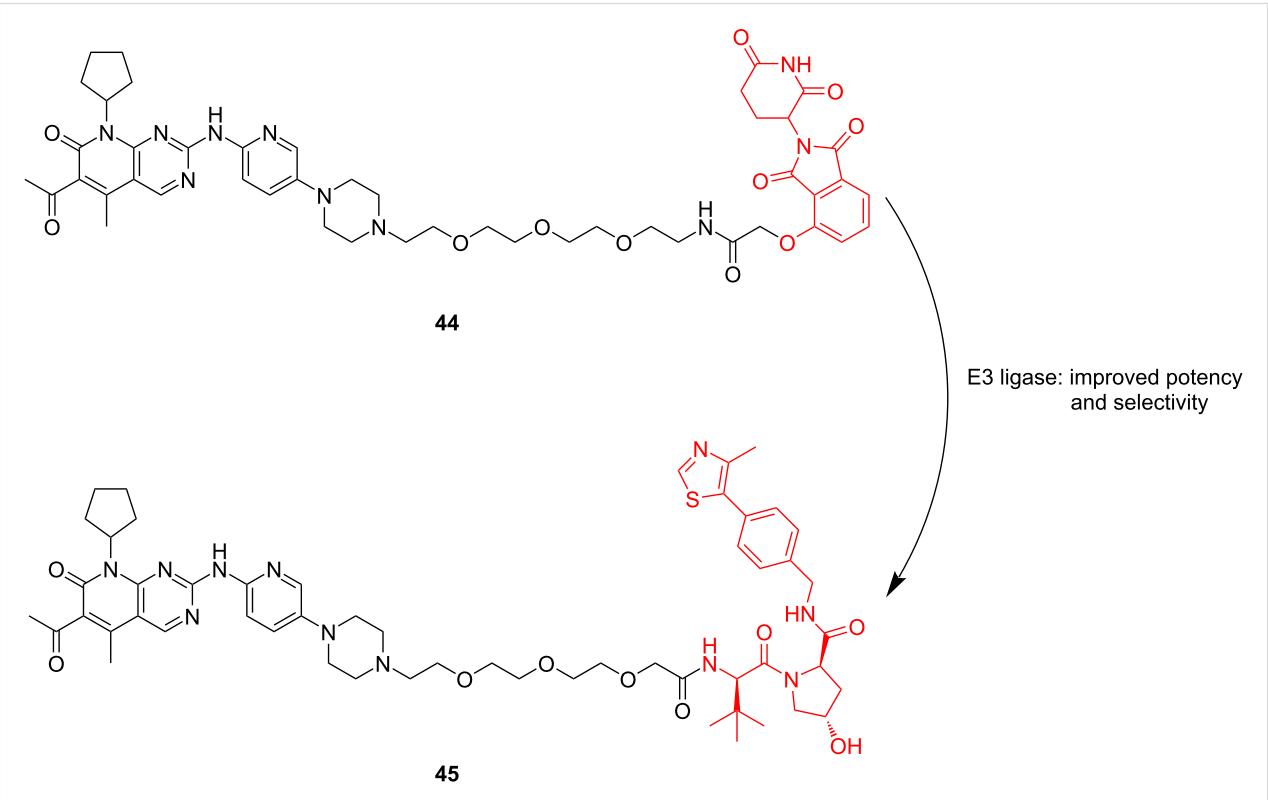
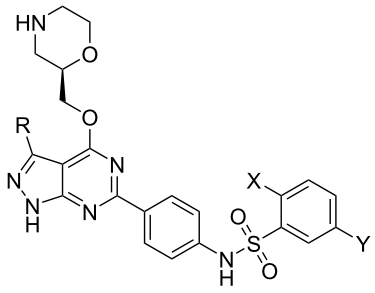
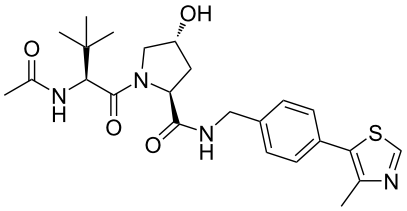


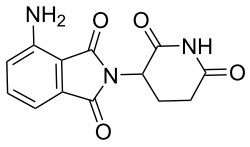
Figure 19: Structures of compounds **44** and **45**.



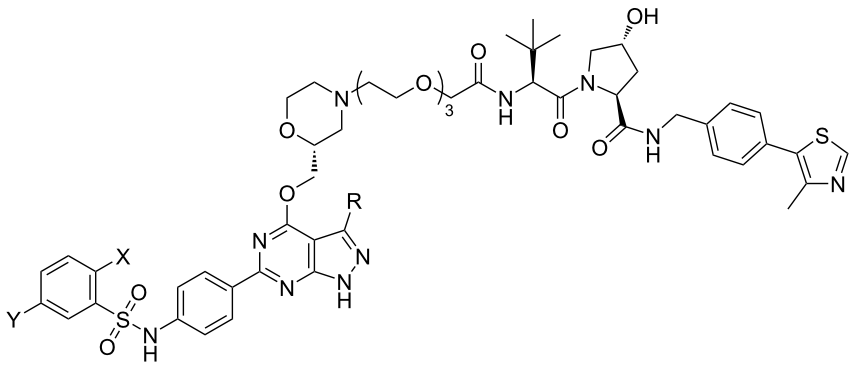
Sanofi 308-R (**46**): R = H, X = F, Y = Me
Sanofi 290-R (**47**): R = Me, X = Cl, Y = OMe



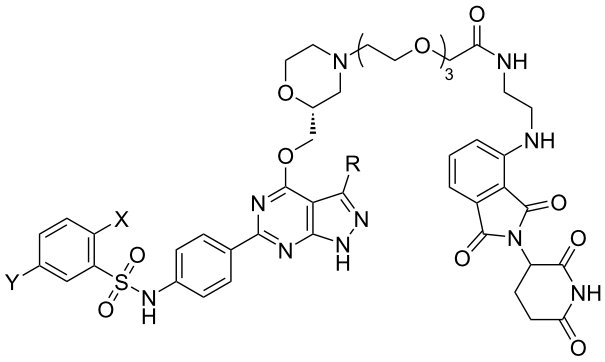
VH032 (**48**)



pomalidomide (**49**)



DAT1 (**50**): R = H, X = F, Y = Me
DAT2 (**51**): R = Me, X = Cl, Y = OMe



DAT3 (**52**): R = H, X = F, Y = Me
DAT4 (**53**): R = Me, X = Cl, Y = OMe

Figure 20: Design of the SGK3 PROTACs.

the degradation characteristics of PROTACs [23]. In their study, by using the same HDAC inhibitors and linkers to focus only on the role of the recruited E3 ubiquitin ligase, they clarified that there are differences in degradation efficiency and substrate selectivity between PROTACs. They designed a group of PROTACs based on dacinostat. The results showed that PROTACs recruiting CRBN prefer to degrade HDAC6 and HDAC8 (**54**), while PROTACs based on VHL ligands strongly prefer to degrade HDAC3 (**55** and **56**). Differences in degradation efficiency were observed between two VHL-based degraders with only different attachment points of VHL the ligands, but degradation selectivity showed only slight differences. Interestingly, degraders based on IAP ligands showed weak but selective HDAC6 degradation (**57**).

The above is a brief list of recent studies on the vital influence of E3 ligands on the selectivity of PROTACs in highly homologous protein families. These studies showed that VHL ligands

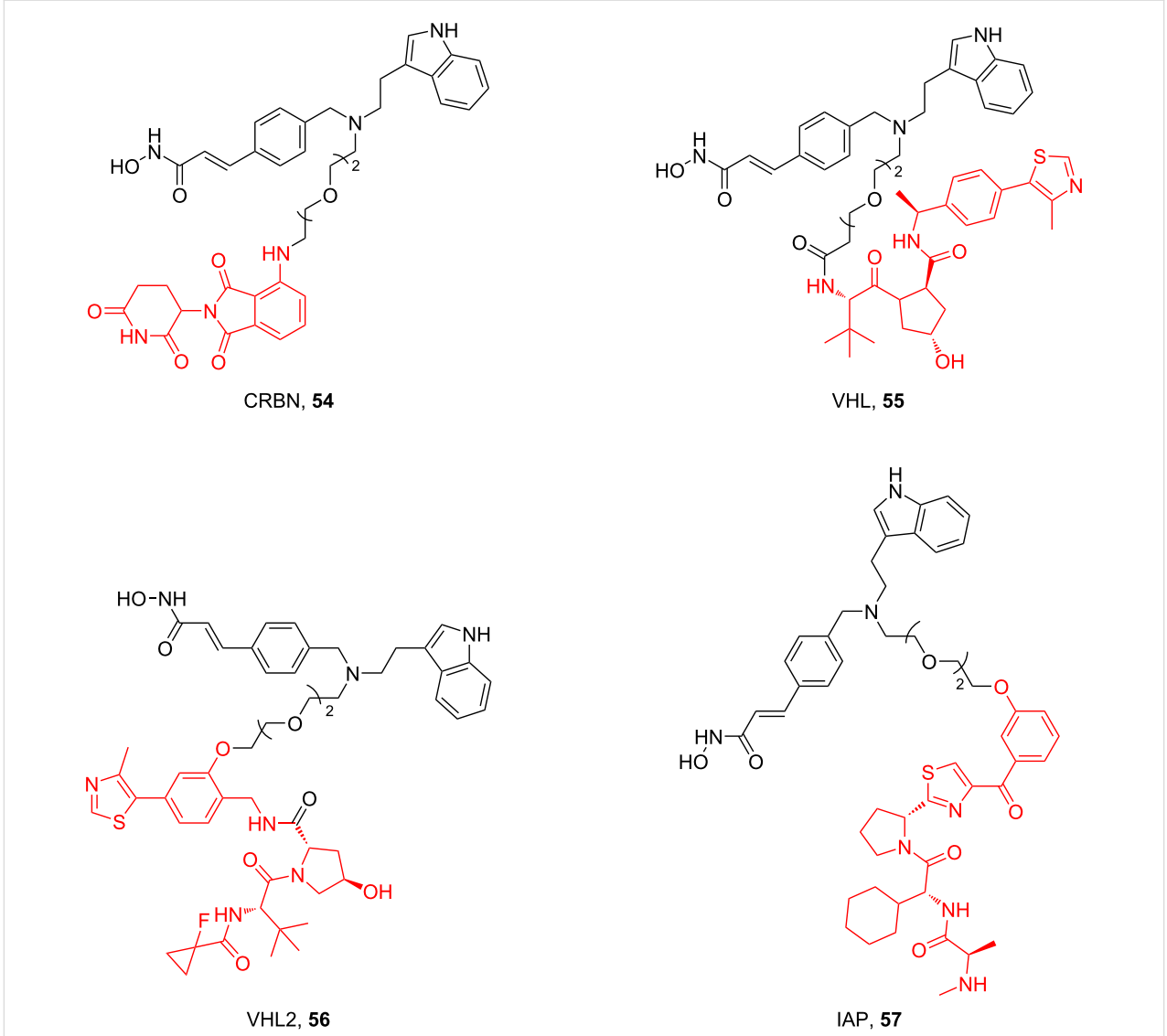


Figure 21: CRBN, VHL, and IAP ligands used when designing HDAC-PROTACs.

maybe have better selectivity than CRBN ligands in most cases. This view was put forward by Crews et al. as early as 2018 [138]. At that time, some studies showed that several PROTACs recruited CRBN could bind and degrade all BRD proteins bound by POI ligands [139,140]. However, when the PROTACs are designed with VHL ligands, it was proved that they can selectively degrade the target proteins [41,107]. To confirm this view, Crews et al. designed a series of PROTACs based on the c-Met tyrosine kinase inhibitor foretinib. Foretinib is a hybrid kinase inhibitor and proteomics has revealed that it can bind more than 130 kinases. When using foretinib as the POI ligand to connect with VHL ligand or CRBN ligand, respectively, through linkers, proteomic results show that the PROTACs recruiting CRBN and VHL have a significantly improved specificity compared with the individual POI ligand. The results showed that PROTACs recruiting CRBN could degrade 14 kinases, while PROTACs recruiting VHL could degrade 9 kinases. This shows that PROTAC technology can significantly improve the selectivity for target proteins compared with small-molecule inhibitors, and it seems that VHL ligands have better selectivity than CRBN ligands. Perhaps VHL ligands are an excellent choice for the development of PROTACs that specifically target highly homologous proteins in the future. However, this view still awaits further verification through additional research in the future.

E3 ligands with different structures target the same E3 ubiquitin ligases

As established by the above discussed studies utilizing different E3 ligands targeting different E3 ubiquitin ligases in

PROTACs design can facilitate degradation selectivity. However, when considering the impact of E3 ligands on the selectivity of PROTACs, except considering the different types of E3 enzymes recruited by E3 ligands, it is also necessary to pay additional attention to the impact of changes in attachment sites on the selectivity of PORTACs when targeting the same E3 ubiquitin ligase.

In 2020, Gürschow, Krönke and co-workers conducted a study to explore the effect of E3 ligands on the selective degradation of CDK4 and CDK6 by PROTACs [131]. In this study, in addition to the study of different E3 ligands targeting CRBN and VHL, respectively, they also modified the structure of the VHL ligand and replaced its junction site with a linker, and finally synthesized compounds **58** and **59** (Figure 22). The results showed that compound **59** ($DC_{50\ for\ CDK4}$/$DC_{50\ for\ CDK6\ (at\ 0.1\ \mu M)}$ = 31) further enhanced the degradation activity compared with compound **58** ($DC_{50\ for\ CDK4}$/$DC_{50\ for\ CDK6\ (at\ 0.1\ \mu M)}$ = 7.4) and improved the ability to selectively degrade CDK6 compared with CDK4.

When talking about the HDAC protein family, we have introduced that Tang and co-workers investigated the design of PROTACs for selectively degrading HDAC based on CRBN ligands in 2019 [22]. They used nexturastat A as the selective inhibitor of HDAC6 with the CRBN E3 ligand pomalidomide to design 18 PROTACs. According to the different amino sites on the phthalimide ring of pomalidomide, these PROTACs were divided into two series. As for these compounds, the difference is the number of carbon atoms (*n*) between nexturastat A and the triazole ring or the number of carbon atoms (*m*) between pomalidomide and the triazole ring.

The results showed that the degradation effect of the C4-linker series PROTACs (Figure 7) was slightly more potent than that of the C5-linker series molecules (Figure 7), but the C5-linker series compounds demonstrated better selectivity than the C4-linker series compounds. This may be related to the different ubiquitination sites of HDAC6 by E3 ubiquitin ligase. Thus, as can be seen, in addition to the influence of the spatial distance caused by different linkers on the degradation efficiency and selectivity of PROTACs, a slight change in the connection position of the E3 ligands and linkers may also significantly affect the two properties.

In addition, as mentioned above, Crews et al. conducted a study to develop PROTACs that specifically degrade p38α and p38δ of the p38MAPK protein family [23]. In this study, they used

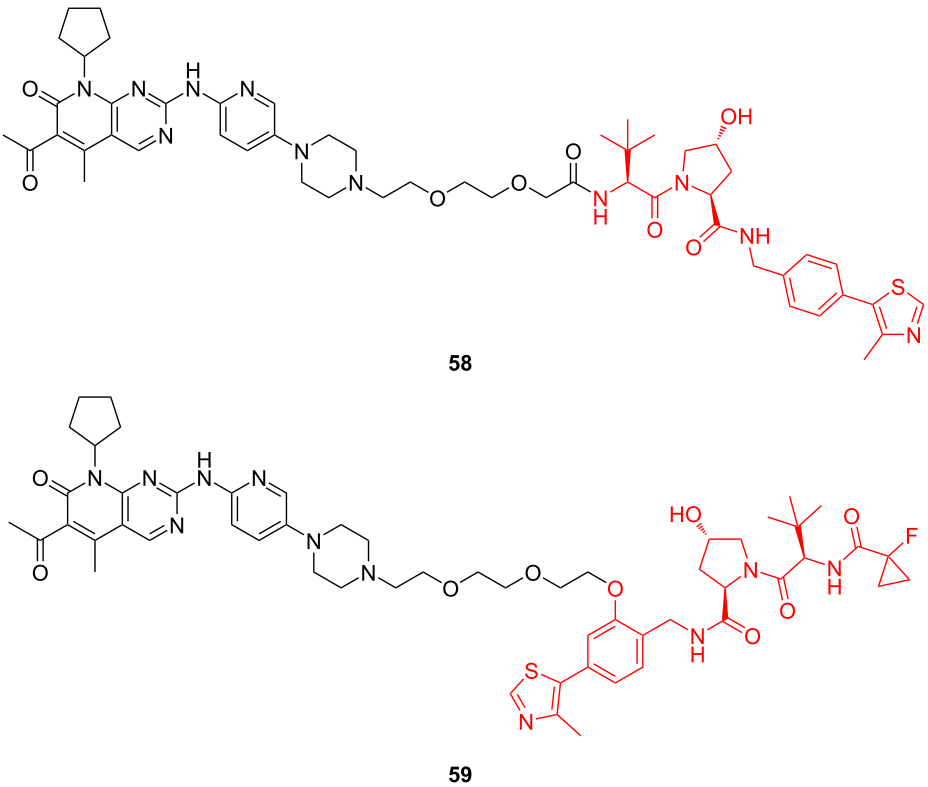


Figure 22: CDK4/6-PROTACs with different E3 ligands targeting the same E3 ubiquitin ligase.

foretinib as the POI ligand and two E3 ligands with different structures to target VHL together to explore the influence of the direction of VHL recruitment of PROTACs on selectivity (termed the "amide series" and "phenyl series" individually). Simultaneously, they used four linkers of different lengths to connect the E3 ligand and the POI ligand to explore the influence of linker fragments on molecular selectivity.

According to the results, it was found that in the two series of PROTACs, when the linker and the POI ligand are consistent, the molecular selectivity will also be widely different. For example, compound **26**, as example for the "amide series" PROTAC, shows no selectivity for p38MAPK protein family subtypes. In contrast, SJF-δ (**33**), as member of the "phenyl series" PROTAC, has a solid ability to degrade p38δ selectively (Figure 12).

At present, more studies still focus on the effect of linkers or different E3 ligands targeting different E3 ubiquitin ligases to design PROTACs with high degradation selectivity. There is still very little research on the selectivity of E3 ligands with different structures targeting the same E3 ubiquitin ligases to design different PROTACs. However, based on the above examples, it could be meaningful to study the effects of different structural types and connection sites targeting the same E3 ubiquitin ligases on the selectivity of PROTACs. When PROTACs are designed with the E3 ligands targeting the same E3 ubiquitin ligase, the replacement of the junction sites of E3 ligands and linker and the modification of E3 ligands may have unexpected effects on the selectivity of PROTACs. In the future, more research is needed in this field to better promote the development of PROTACs with good selectivity for members of highly homologous protein families.

Influence of POI ligands on the selectivity of PROTACs in highly homologous protein families

As it is well-known, PROTACs are composed of three parts: POI ligand, linker, and E3 ligand. When designing PROTACs, research mainly focusses on the kind of linkers and E3 ligands and often ignores the POI ligands. Many studies use already approved small-molecule inhibitors or compounds under clinical research as POI ligands. This will lead to the effect of the POI ligands on the selectivity of PROTACs in highly homologous protein families, which is mainly determined by the selectivity of the small-molecule inhibitors used by POI ligands (Figure 17). However, multiple literature reports suggest that pairing the E3 ligase with the target protein is one of the most critical factors in generating potent and selective PROTACs [87,131,138,141]. This is primarily driven by the differential ability of varying E3 ligases to form a favorable ternary complex with a target POI [24]. Therefore, according to this view and after comparison and summary of a large number of studies, we are more convinced that the POI ligands are crucial to the selectivity of PROTACs.

In 2016, Crews et al. reported that the optimal pairing between the POI ligand and the recruited E3 ubiquitin ligase to achieve potent and selective PROTAC-induced target POI degradation is critical [126]. The important influence of E3 ligands on the selectivity of PROTACs highlighted in this study was already introduced above. This time, we look at this study from the perspective of POI ligands. In this study, the authors designed different PROTACs by connecting BCR-ABL TKI, which targets the binding of the c-ABL kinase domain, with a VHL ligand or CRBN ligand through the linker. Interestingly, according to the results, it is found that when dasatinib and bosutinib combine pomalidomide to recruit CRBN, the formed PROTACs can degrade both c-ABL and BCR-ABL. However, when they combined the VHL ligand to form PROTACs, the dasatinib-VHL PROTAC only degrades c-ABL and has no degradation effect on BCR-ABL. At the same time, the bosutinib-VHL PROTAC does not display any degradation activity.

Based on this study, it can be noticed that POI ligands can genuinely affect the selectivity of PROTACs. Although there are not many literature reports on the impact of POI ligands on the selectivity of PROTACs, many research groups have started to try to improve the selectivity of PROTACs through the optimization of POI ligands. In addition, it is worth noting that the effect of POI ligands on the selectivity of PROTACs is closely related to E3 ligands, which was well confirmed in this study. Moreover, it is believed that the impact of POI ligands is also inseparable from linkers. POI ligands, linkers, and E3 ligands, as the three parts of PROTACs, will influence and restrict each other to jointly regulate the selectivity of PROTACs.

Influence of protein–protein interaction on the selectivity of PROTACs in highly homologous protein families

The mechanism of PROTAC-mediated target protein degradation intrinsically relies on the formation of a ternary complex [142]. This is a transient three-component assembly composed of the target protein, the PROTAC molecule, and the recruited E3 ubiquitin ligase. This process goes beyond the simple additive effects of the binary affinities between the POI ligand and its target, and between the E3 ligand and the ligase. A critical and often decisive factor governing the efficiency, stability, and selectivity of the ternary complex is the protein–protein interaction that arises between the POI and the E3 ligase under PROTAC-induced proximity [143]. The induced PPI interface can confer strong positive cooperativity, meaning that the

affinity of ternary complex formation greatly exceeds that predicted from the individual binding events. In highly homologous protein families, whose members share highly conserved active or binding sites that limit the discriminative capacity of conventional small-molecule inhibitors, the geometric and chemical complementarity of the induced PPI interface becomes a powerful discriminator. Subtle conformational or surface electrostatic differences among homologous proteins can be amplified in the context of the ternary complex, enabling PROTACs to selectively engage and degrade one member rather than others. This PPI-driven selectivity represents a paradigm shift from traditional occupancy-based inhibition and offers a unique solution to long-standing challenges in drug discovery [144].

Molecular matching and spatial orientation

The ability of PROTACs to distinguish proteins with high sequence and structural similarity depends on precise molecular matching within the ternary complex [145]. Homologous proteins, such as kinases from the same family or bromodomains within the BET family, typically share conserved core folds and ligand-binding pockets. However, surface regions outside these canonical pockets that participate in protein–protein interactions may exhibit greater variability [146]. Through its specific linker length, rigidity, and attachment points, a PROTAC positions the POI and the E3 ligase in a unique spatial orientation. This orientation favors the formation of a specific PPI interface that may be optimal only for one member within a homologous family. For example, the surface of one subtype may present a cluster of complementary charges or hydrophobic residues that perfectly matches a region on an E3 ligase (such as VHL or CRBN), thereby forming a stable ternary complex [147,148]. A closely related subtype, although binding the POI ligand with similar affinity, may display a slightly different surface topology or charge distribution at this interface, resulting in weaker or non-productive PPIs, reduced ternary-complex stability, and consequently inefficient ubiquitination and degradation.

CDK

For the CDK family, PROTAC molecules leverage the combinatorial design of their linkers and E3 ligands to shape finely differentiated PPI interfaces between homologous isoforms, thereby achieving a level of selectivity unattainable by traditional small-molecule inhibitors [149]. Taking CDK4 and CDK6 as examples, despite their 71% sequence homology, the PROTAC molecule **2** selectively degrades CDK6 while having no effect on CDK4 [34]. The fundamental reason is that this molecule, via a specific PEG linker, recruits CDK6 and the CRBN E3 ligase into a unique spatial conformation, forming a stable and catalytically active PPI interface. Although CDK4 can also bind to the same POI ligand (palbociclib), its surface cannot form a similarly compatible PPI interface with CRBN under the geometric configuration induced by molecule **2**, thus precluding effective ubiquitination (Figure 23).

Similarly, in the differentiation of CDK9 from CDK2, FN-1501-based PROTACs shift the degradation profile from dual CDK2/9 targeting to CDK9-specific selectivity by extending the linker length from 8–10 atoms to 11–12 atoms.

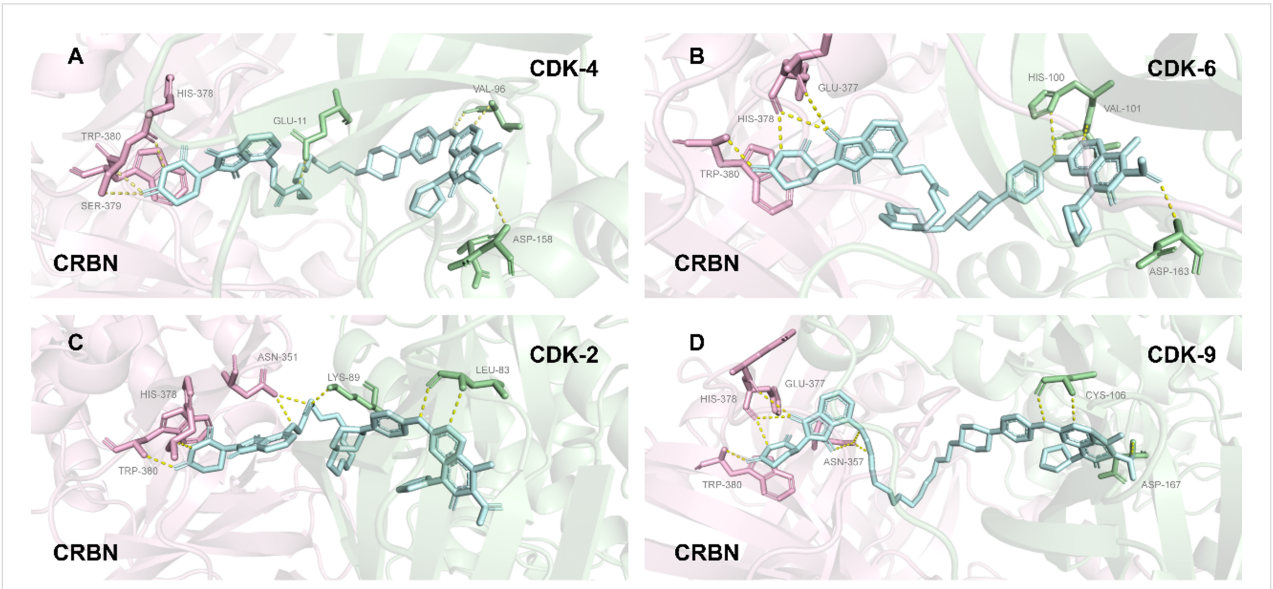


Figure 23: Structural basis for the selective degradation of CDK6 over CDK4 by PROTAC **2**. (A) CDK4–molecule **2**–CRBN complex. (B) CDK6–molecule **2**–CRBN complex. (C) CDK2–molecule **2**–CRBN complex. (D) CDK9–molecule **2**–CRBN complex.

This transition also stems from changes in the spatial orientation of the ternary complex caused by the variation in linker length: only the surface characteristics of CDK9 can form a complementary PPI interface with the E3 ligase under the new distance requirements, while CDK2 is excluded due to interface mismatch. These cases fully demonstrate that within homologous protein families, the spatial geometry regulated by the linker in a PROTAC determines the isoform-specific PPI interface, which in turn dictates selective degradation.

p38

Within the highly homologous p38 MAPK family, the selective degradation capacity of PROTACs primarily depends on the stability and geometric configuration of PPIs within the induced ternary complexes. In 2019, Crews et al. utilized the promiscuous kinase inhibitor foretinib as a warhead [23], combined with two VHL ligands featuring distinct linkage modes (amide vs phenyl series) and varying linker lengths, to develop two PROTAC molecules with orthogonal selectivity: molecule **29** (13-atom linker, amide linkage) and molecule **33** (10-atom linker, phenyl linkage). Although both molecules can bind to both p38α and p38δ, molecule **29** selectively degrades p38α (DC_{50} = 7.16 nM, D_{max} = 97.4%) while exhibiting almost no degradation of p38δ (D_{max} = 18%). Conversely, molecule **33** selectively degrades p38δ (DC_{50} = 46.17 nM, D_{max} = 99.4%) and has no effect on p38α. To investigate the mechanism underlying this differential selectivity, the authors performed in vitro GST-pulldown assays, which revealed that molecule **29** efficiently enriched p38α to form a ternary complex while molecule **33** did not. However, for p38δ, both PROTACs were able to induce ternary complex formation, suggesting that the mere presence of a ternary complex is not a sufficient condition for degradation. Further determination of the affinity and kinetics of the p38δ–PROTAC–VHL ternary complexes using surface plasmon resonance (SPR) showed that the complex induced by SJF-δ (**33**) possessed stronger affinity (K_d = 436 nM) and a longer dissociation half-life ($t_{1/2}$ = 38 s). In contrast, the complex induced by SJF-α exhibited weaker affinity (K_d = 1.2 μM) and a shorter half-life ($t_{1/2}$ = 8 s). Ternary complex capture assays in cell lysates also confirmed that SJF-δ could stably enrich endogenous p38δ, whereas SJF-α could not. Molecular dynamics simulations indicated that the two PROTACs guide VHL to dock with p38δ in different conformations: in the presence of molecule **33**, Arg108 of VHL forms favorable electrostatic interactions with Glu49/Glu160 of p38δ; however, in the presence of molecule **29**, Arg108 makes unfavorable contacts with Lys220/Thr221 of p38δ. To validate this prediction, the authors constructed a K220E/T221E double point mutant of p38δ. When expressed in cells, the mutant p38δ, which was originally resistant to molecule **29**, became effectively degradable by molecule **29**. This result directly demonstrates that subtle differences in the PPI interface within the ternary complex dictate the selectivity of PROTACs for highly homologous proteins, providing a theoretical foundation for the future design of precision degraders based on PPI optimization (Figure 24).

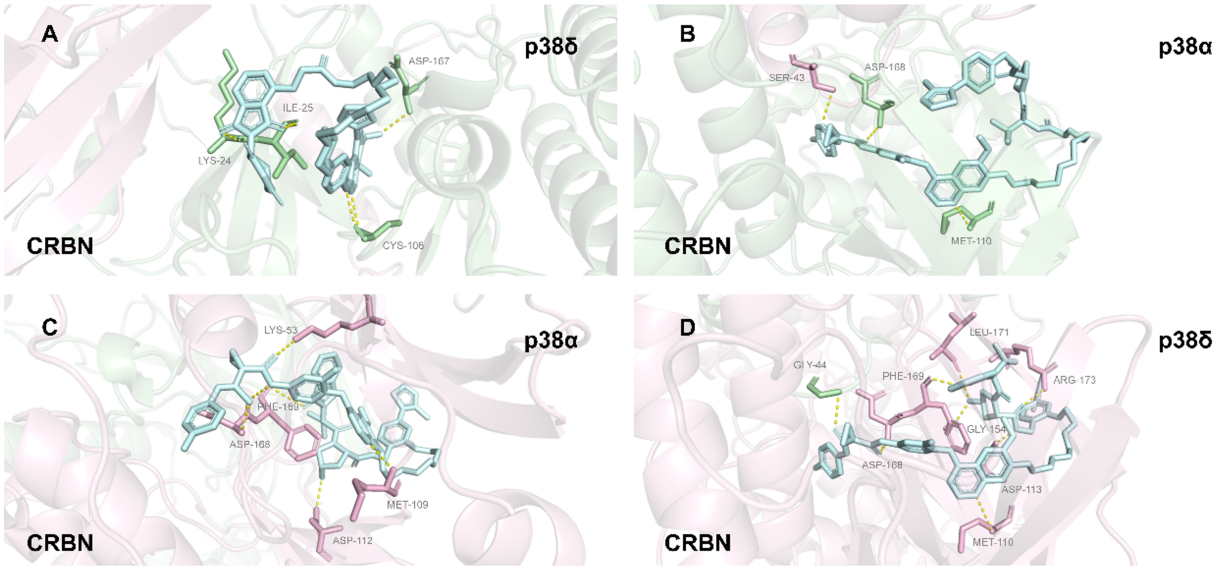


Figure 24: Regarding the PPI-driven selectivity mechanism of two PROTAC molecules, SJF-α and SJF-δ, for p38α and p38δ MAP kinases. (A) p38δ–molecule **29**–CRBN complex. (B) p38α–molecule **29**–CRBN complex. (C) p38α–molecule **33**–CRBN complex. (D) p38δ–molecule **33**–CRBN complex.

STAT

The members of the STAT (signal transducer and activator of transcription) family exhibit high structural homology, particularly within the SH2 domain [150].

In 2024, Wang and co-workers reported the discovery of AK-1690, the first potent and highly selective STAT6 PROTAC degrader [151]. The breakthrough of this study lies first in the optimization and acquisition of the ligand AK-068, which possesses an extremely high affinity for STAT6 (K_i = 6 nM) and >85-fold selectivity over STAT5. Based on this ligand, the researchers designed the PROTAC molecule AK-1690. Experimental data demonstrated that AK-1690 effectively induces STAT6 protein degradation in cells (with a DC_{50} as low as 1 nM), while showing almost no degradation effect on other STAT family members, including STAT1, STAT2, STAT3, and STAT5, at concentrations as high as 10 μM. The researchers successfully resolved the first co-crystal structure of STAT6 in complex with AK-1690, providing a structural basis for understanding its ultra-high selectivity. This structure revealed a precise complementary relationship at the binding interface between AK-1690 and STAT6. Although the ligand binding pocket of STAT6 shares some similarities with other STAT members, the ternary complex interface formed by AK-1690 and STAT6 contains numerous unique, non-conservative interactions (Figure 25). These PROTAC-induced interprotein contacts render the entire complex highly sensitive to the surface topology and amino acid residue characteristics of STAT6, thereby excluding other family members and achieving exceptional selectivity.

The protein–protein interactions within PROTAC-induced ternary complexes constitute a foundational mechanism for achieving unprecedented selectivity against highly homologous protein families. It goes beyond simple target engagement and instead exploits subtle biophysical complementarity between two proteins [145]. This selectivity arises from the cooperative interplay among the POI ligand, the linker, and the E3-ligase ligand. Quantitative frameworks for cooperativity, the explanatory power of structural biology, and the predictive capacity of computational modeling are transforming PROTAC design from empirical screening into a more rational, structure-guided endeavor. Even when using the same or similar POI ligands, the degradation profile of PROTACs against homologous proteins can be significantly altered by optimizing the choice of E3 ligase ligand, the length of the linker, and the attachment site (Table 2). Future advances will depend on discovering new E3-ligase ligands to expand the "PPI toolbox," obtaining more ternary-complex structures across diverse target families, and improving computational algorithms to accurately predict cooperative binding [152]. By harnessing the principles of PPI-driven selectivity, the PROTAC paradigm holds promise for delivering highly precise therapeutics capable of discriminating between "sibling" proteins, thereby minimizing off-target effects and opening new therapeutic avenues for diseases driven by specific members of redundant protein families [153].

Conclusion

To sum up, this review mainly introduces the advantages of PROTAC over small-molecule inhibitors in the selectivity of highly homologous protein families. According to the system-

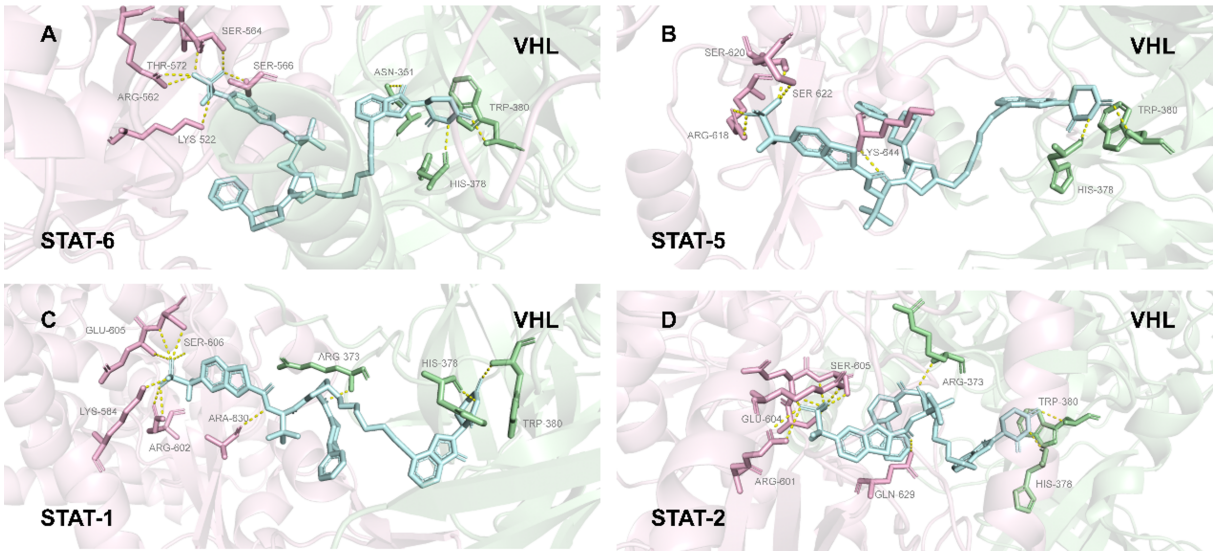


Figure 25: Co-crystal structure and ultra-high selectivity of the STAT6 PROTAC degrader AK-1690. (A) STAT6–AK-1609–CRBN complex. (B) STAT5–AK-1609–CRBN complex. (C) STAT1–AK-1609–CRBN complex. (D) STAT2–AK-1609–CRBN complex.

Table 2: Summary of representative PROTAC molecules and their degradation activities.

Molecule name (compound number)	E3 ligase	Target (POI)	Data values
PROTAC-1 (**1**)	CRBN	CDK4/6	in Jurkat cells: CDK4 degradation rate >50% (0.1 μM); CDK6 degradation rate >95% (0.1 μM)
BSJ-03-123 (**2**)	CRBN	CDK6	it exhibits superior degradation capability towards CDK6, while proteomic results indicate no degradation effect on CDK4
PROTAC-6 (**4**)	CRBN	CDK6	selectively degrading CDK6 in HPNE and MiaPaCa2 cells does not affect CDK2, 4, 5, 7, 9
PROTAC-pal-pom (**5**)	CRBN	CDK4/6	MDA-MB-231cells: DC_{50}(CDK4) = 13 nmol/L; DC_{50}(CDK6) = 34 nmol/L
PROTAC-CP 10 (**6**)	CRBN	CDK6	U251 cells: DC_{50}(CDK6) = 2.1 nmol/L; DC_{50}(CDK4) > 100 nmol/L
compounds **9**–**11**	CRBN	CDK2/9	it is a dual degrader for CDK2/9
compounds **12**–**14**	CRBN	CDK9	selective degradation of CDK9
compound **16c**	CRBN	CDK9	it exhibits optimal CDK9 degradation activity and high selectivity towards CDK9 compared to CDK5
compound **19c**	CRBN	HDAC6	it exhibits the best selective degradation effect on HDAC6 compared to HDAC1, 2, and 4
compound **20d**	CRBN	HDAC6	DC_{50} = 1.64 ± 0.24 nM
compound **20h**	CRBN	HDAC6	at a concentration of 100 nM, it can degrade 78.3–80.1% of HDAC6
compound **20i**	CRBN	HDAC6	at a concentration of 100 nM, it can degrade 82.1–84.1% of HDAC6
compound **21j**	VHL	HDAC6	DC_{50} = 7.1 nM (exhibiting the best selectivity compared to HDAC1, 2, 3, 4, 7, 8)
MZ1 (**35**)	VHL	BRD4	it exhibits excellent selective degradation capability for BRD4 (compared to BRD2 and BRD3)
SJF-8240/SJF-α (**29**)	VHL	p38α	DC_{50}(p38α) = 7.16 ± 1.03 nM; D_{max} = 97.4%
compound **33**	VHL	p38δ	DC_{50}(p38δ) = 46.17 ± 9.85 nM; D_{max} = 99.41 ± 3.31%
SHD913 (**37**)	VHL	BRD4	in PC-3 cells: BRD4 long isoform DC_{50} = 7.7 nM, short isoform DC_{50} = 5.0 nM

atic summary, we can clearly find that both linker and E3 ligand are crucial to the selectivity of PROTAC design (Table 3). Whether it is an E3 ligand or a linker fragment, it may ultimately affect the overall structure of the PORTAC to achieve highly selective degradation through the subtle differences between PPI and ubiquitinated lysine sites. Meanwhile, the selective degradation of highly homologous protein families by PROTACs underscores the fundamental pharmacological shift from occupancy-driven inhibition to event-driven target elimination. This shift allows selectivity to emerge from ternary complex cooperativity, induced protein–protein interactions, and spatial orientation, rather than solely from binding affinity differences.

The difference between protein–protein interaction and ubiquitinated lysine sites is currently widely accepted as two factors affecting the selectivity of PROTAC molecules in highly homologous protein families. It is generally accepted that the degradation effect of PROTACs requires the formation of ternary complexes with POI and E3 ubiquitin ligase. However, before Cullin et al. published their research in 2017, few people believed that when PROTAC formed ternary complexes, the protein–protein interaction between E3 ubiquitin ligase and POI PPI would have a huge impact on the degradation of PROTAC molecules, and even affect the selectivity of PROTAC molecules to highly homologous protein families. The existence of the PPI effect makes the formation of ternary complexes not only depend on the affinity between POI ligand and POI but also between E3 ligand and E3 ubiquitin ligase, which further confirms that there is no inevitable relationship between the binding affinity of PROTAC molecules and the degradation effect. The existence of the PPI effect makes it easier or more difficult to form ternary complexes among members of the highly homologous protein family with only slight structural

Table 3: Factors influencing the selectivity of PROTACs in highly homologous protein families.

protein level	surface lysine availability	the presence and spatial accessibility of specific lysine residues on the protein surface that are required for ubiquitination
	surface topology & residues	unique biophysical features and amino acid residues (even in highly homologous proteins) that facilitate or hinder specific protein-protein interactions (PPIs)
	tissue and cell expression	differential abundance of the target protein or the recruited E3 ligase across different tissues and cell types
	subcellular localization	the specific cellular compartment where a protein resides can determine its accessibility to the PROTAC and the degradation machinery
	protein–protein complexes	target proteins existing within multiprotein complexes can have different "degradability" compared to their monomeric forms
PROTAC level	POI warhead selection	the choice of the ligand for the target protein (POI) and its specific binding orientation or "exit vector"
	E3 ligase recruitment	selecting different E3 ligases (e.g., VHL vs CRBN) can drastically alter the degradation profile of the same target
	linker engineering	optimization of the linker's length, flexibility, and chemical composition (e.g., PEG vs alkyl chains) to allow for productive orientation
	linker attachment points	the specific positions on the ligands where the linker is attached, which determines the spatial orientation of the recruited proteins
	ternary complex stability	the ability of the PROTAC to induce stable and **cooperative** de novo protein–protein interactions between the POI and the E3 ligase

differences, thus enabling PROTAC molecules to achieve selectivity between different subtypes. This is a milestone for the future development of highly specific PROTAC molecules [154]. On this basis, in order to explore more about the formation of ternary complexes, it is essential to analyze the crystal structure of ternary complexes as much as possible to explore the mechanism by which a PROTAC plays its high selectivity in highly homologous protein families. Therefore, in recent years, many researchers have devoted themselves to building corresponding prediction models using computer simulations to accurately predict the crystal structure of PROTAC ternary composites, so as to better develop PROTAC molecules for various targets in the future [155]. Furthermore, following ternary complex formation, PROTACs promote the ubiquitination of the POI by recruiting the E3 ubiquitin ligase, thereby triggering its proteasomal degradation. Despite the high sequence similarity within certain protein families, differences in the accessibility and positioning of lysine residues result in distinct ubiquitination patterns. This provides an additional layer of selectivity for PROTAC-mediated degradation. Although this concept has been proposed, systematic investigations validating the relationship between lysine site distribution and degradation selectivity remain limited. Future studies in this area are therefore warranted [154].

Looking forward, stimulus-responsive degraders such as light-controllable PROTACs (opto-PROTACs) may further expand the capabilities of event-driven pharmacology by integrating structural precision with spatial and temporal regulation. Conditional activation strategies could ultimately enhance selectivity among highly homologous proteins by restricting degradation to defined biological contexts.

Acknowledgements

We would like to acknowledge the researchers whose relevant works are cited in this review and all co-authors for their support. The graphical abstract was created in BioRender. Wu, M. (2026) https://BioRender.com/kd4u2nm. This content is not subject to CC BY 4.0.

Funding

The authors are grateful to the financial support from the grant from the National Natural Science Foundation of China (No. 82504580), Young Elite Scientists Sponsorship Program by ZAST (to Mingfei Wu), and Jinhua Science and Technology Project (2023-3-060).

Author Contributions

Luxi Yang: visualization; writing – original draft. Xinfei Mao: visualization; writing – original draft. Jingyi Zhang: visualization. Jing Shu: data curation; investigation. Wenhai Huang: conceptualization; data curation; investigation; supervision. Xiaowu Dong: conceptualization; supervision; visualization; writing – review & editing. Yinqiao Chen: conceptualization; funding acquisition; supervision; visualization; writing – review

& editing. Mingfei Wu: funding acquisition; project administration; visualization; writing – review & editing.

ORCID® iDs

Mingfei Wu - https://orcid.org/0000-0002-5051-5788

Data Availability Statement

Data sharing is not applicable as no new data was generated or analyzed in this study.

References

1. Boija, A.; Klein, I. A.; Young, R. A. *Cancer Cell* **2021,** *39,* 174–192. doi:10.1016/j.ccell.2020.12.003
2. Min, K.-W.; Lee, S.-H.; Baek, S. J. *Cancer Lett.* **2016,** *370,* 108–116. doi:10.1016/j.canlet.2015.09.022
3. Goel, B.; Tripathi, N.; Bhardwaj, N.; Jain, S. K. *Curr. Top. Med. Chem.* **2020,** *20,* 1535–1563. doi:10.2174/1568026620666200516152756
4. Shua, Y.; Wang, Q.; Wang, Y.; Jiang, C. *Eur. J. Med. Chem.* **2025,** *300,* 118171. doi:10.1016/j.ejmech.2025.118171
5. Wang, J.; Lisanza, S.; Juergens, D.; Tischer, D.; Watson, J. L.; Castro, K. M.; Ragotte, R.; Saragovi, A.; Milles, L. F.; Baek, M.; Anishchenko, I.; Yang, W.; Hicks, D. R.; Expòsit, M.; Schlichthaerle, T.; Chun, J.-H.; Dauparas, J.; Bennett, N.; Wicky, B. I. M.; Muenks, A.; DiMaio, F.; Correia, B.; Ovchinnikov, S.; Baker, D. *Science* **2022,** *377,* 387–394. doi:10.1126/science.abn2100
6. Becskei, A. *Molecules* **2020,** *25,* 1902. doi:10.3390/molecules25081902
7. Gaffney, D. O.; Jennings, E. Q.; Anderson, C. C.; Marentette, J. O.; Shi, T.; Schou Oxvig, A.-M.; Streeter, M. D.; Johannsen, M.; Spiegel, D. A.; Chapman, E.; Roede, J. R.; Galligan, J. J. *Cell Chem. Biol.* **2020,** *27,* 206–213.e6. doi:10.1016/j.chembiol.2019.11.005
8. Ding, Y.; Fei, Y.; Lu, B. *Trends Pharmacol. Sci.* **2020,** *41,* 464–474. doi:10.1016/j.tips.2020.04.005
9. Zou, Y.; Ma, D.; Wang, Y. *Cell Biochem. Funct.* **2019,** *37,* 21–30. doi:10.1002/cbf.3369
10. Soini, L.; Leysen, S.; Davis, J.; Ottmann, C. *Curr. Opin. Chem. Biol.* **2022,** *69,* 102169. doi:10.1016/j.cbpa.2022.102169
11. Sakamoto, K. M.; Kim, K. B.; Kumagai, A.; Mercurio, F.; Crews, C. M.; Deshaies, R. J. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98,* 8554–8559. doi:10.1073/pnas.141230798
12. Li, X.; Song, Y. *J. Hematol. Oncol.* **2020,** *13,* 50. doi:10.1186/s13045-020-00885-3
13. Bond, M. J.; Crews, C. M. *RSC Chem. Biol.* **2021,** *2,* 725–742. doi:10.1039/d1cb00011j
14. Sakamoto, K. M. *Pediatr. Res.* **2010,** *67,* 505–508. doi:10.1203/pdr.0b013e3181d35017
15. Békés, M.; Langley, D. R.; Crews, C. M. *Nat. Rev. Drug Discovery* **2022,** *21,* 181–200. doi:10.1038/s41573-021-00371-6
16. Jang, E. R.; Lee, W.; Kim, K. B. *Curr. Protoc. Chem. Biol.* **2010,** *2,* 71–87. doi:10.1002/9780470559277.ch090242
17. Li, K.; Crews, C. M. *Chem. Soc. Rev.* **2022,** *51,* 5214–5236. doi:10.1039/d2cs00193d
18. Salami, J.; Crews, C. M. *Science* **2017,** *355,* 1163–1167. doi:10.1126/science.aam7340
19. Zhao, L.; Zhao, J.; Zhong, K.; Tong, A.; Jia, D. *Signal Transduction Targeted Ther.* **2022,** *7,* 113. doi:10.1038/s41392-022-00966-4
20. Koroleva, O. A.; Dutikova, Y. V.; Trubnikov, A. V.; Zenov, F. A.; Manasova, E. V.; Shtil, A. A.; Kurkin, A. V. *Russ. Chem. Bull.* **2022,** *71,* 2310–2334. doi:10.1007/s11172-022-3659-z
21. Jiang, B.; Wang, E. S.; Donovan, K. A.; Liang, Y.; Fischer, E. S.; Zhang, T.; Gray, N. S. *Angew. Chem., Int. Ed.* **2019,** *58,* 6321–6326. doi:10.1002/anie.201901336
22. Wu, H.; Yang, K.; Zhang, Z.; Leisten, E. D.; Li, Z.; Xie, H.; Liu, J.; Smith, K. A.; Novakova, Z.; Barinka, C.; Tang, W. *J. Med. Chem.* **2019,** *62,* 7042–7057. doi:10.1021/acs.jmedchem.9b00516
23. Smith, B. E.; Wang, S. L.; Jaime-Figueroa, S.; Harbin, A.; Wang, J.; Hamman, B. D.; Crews, C. M. *Nat. Commun.* **2019,** *10,* 131. doi:10.1038/s41467-018-08027-7
24. Guenette, R. G.; Yang, S. W.; Min, J.; Pei, B.; Potts, P. R. *Chem. Soc. Rev.* **2022,** *51,* 5740–5756. doi:10.1039/d2cs00200k
25. Zhang, N.; Wang, Z.; Zhao, Y. *Cytokine Growth Factor Rev.* **2020,** *55,* 80–85. doi:10.1016/j.cytogfr.2020.03.002
26. Yuan, K.; Wang, X.; Dong, H.; Min, W.; Hao, H.; Yang, P. *Acta Pharm. Sin. B* **2021,** *11,* 30–54. doi:10.1016/j.apsb.2020.05.001
27. Lu, P.; Cheng, Y.; Xue, L.; Ren, X.; Xu, X.; Chen, C.; Cao, L.; Li, J.; Wu, Q.; Sun, S.; Hou, J.; Jia, W.; Wang, W.; Ma, Y.; Jiang, Z.; Li, C.; Qi, X.; Huang, N.; Han, T. *Cell* **2024,** *187,* 7126–7142.e20. doi:10.1016/j.cell.2024.10.015
28. Khan, S.; Zhang, X.; Lv, D.; Zhang, Q.; He, Y.; Zhang, P.; Liu, X.; Thummuri, D.; Yuan, Y.; Wiegand, J. S.; Pei, J.; Zhang, W.; Sharma, A.; McCurdy, C. R.; Kuruvilla, V. M.; Baran, N.; Ferrando, A. A.; Kim, Y.-m.; Rogojina, A.; Houghton, P. J.; Huang, G.; Hromas, R.; Konopleva, M.; Zheng, G.; Zhou, D. *Nat. Med.* **2019,** *25,* 1938–1947. doi:10.1038/s41591-019-0668-z
29. Xiang, W.; Wang, S. *J. Med. Chem.* **2020,** *63,* 14560–14561. doi:10.1021/acs.jmedchem.0c01947
30. Snyder, L. B.; Neklesa, T. K.; Willard, R. R.; Gordon, D. A.; Pizzano, J.; Vitale, N.; Robling, K.; Dorso, M. A.; Moghrabi, W.; Landrette, S.; Gedrich, R.; Lee, S. H.; Taylor, I. C. A.; Houston, J. G. *Mol. Cancer Ther.* **2025,** *24,* 511–522. doi:10.1158/1535-7163.mct-23-0655
31. Rathkopf, D. E.; Patel, M. R.; Choudhury, A. D.; Rasco, D.; Lakhani, N.; Hawley, J. E.; Srinivas, S.; Aparicio, A.; Narayan, V.; Runcie, K. D.; Emamekhoo, H.; Reichert, Z. R.; Nguyen, M. H.; Wells, A. L.; Kandimalla, R.; Liu, C.; Suryawanshi, S.; Han, J.; Wu, J.; Arora, V. K.; Pourdehnad, M.; Armstrong, A. J. *Ann. Oncol.* **2025,** *36,* 76–88. doi:10.1016/j.annonc.2024.09.005
32. Zhang, Y.; Ming, A.; Wang, J.; Chen, W.; Fang, Z. *Pharmacol. Res.* **2024,** *205,* 107234. doi:10.1016/j.phrs.2024.107234
33. Schmalzbauer, B. S.; Thondanpallil, T.; Heller, G.; Schirripa, A.; Sperl, C.-M.; Mayer, I. M.; Knab, V. M.; Nebenfuehr, S.; Zojer, M.; Mueller, A. C.; Fontaine, F.; Klampfl, T.; Sexl, V.; Kollmann, K. *Cancers* **2022,** *14,* 1554. doi:10.3390/cancers14061554
34. Ataei-Nazari, S.; Amoushahi, M.; Madsen, J.; Jensen, J.; Heuck, A.; Mohammadi-Sangcheshmeh, A.; Lykke-Hartmann, K. *Front. Cell Dev. Biol.* **2022,** *10,* 1036917. doi:10.3389/fcell.2022.1036917
35. Raina, K.; Lu, J.; Qian, Y.; Altieri, M.; Gordon, D.; Rossi, A. M. K.; Wang, J.; Chen, X.; Dong, H.; Siu, K.; Winkler, J. D.; Crew, A. P.; Crews, C. M.; Coleman, K. G. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113,* 7124–7129. doi:10.1073/pnas.1521738113

36. Mannion, J.; Gifford, V.; Bellenie, B.; Fernando, W.; Ramos Garcia, L.; Wilson, R.; John, S. W.; Udainiya, S.; Patin, E. C.; Tiu, C.; Smith, A.; Goicoechea, M.; Craxton, A.; Moraes de Vasconcelos, N.; Guppy, N.; Cheung, K.-M. J.; Cundy, N. J.; Pierrat, O.; Brennan, A.; Roumeliotis, T. I.; Benstead-Hume, G.; Alexander, J.; Muirhead, G.; Layzell, S.; Lyu, W.; Roulstone, V.; Allen, M.; Baldock, H.; Legrand, A.; Gabel, F.; Serrano-Aparicio, N.; Starling, C.; Guo, H.; Upton, J.; Gyrd-Hansen, M.; MacFarlane, M.; Seddon, B.; Raynaud, F.; Roxanis, I.; Harrington, K.; Haider, S.; Choudhary, J. S.; Hoelder, S.; Tenev, T.; Meier, P. *Immunity* **2024,** *57,* 1514–1532.e15. doi:10.1016/j.immuni.2024.04.025
37. Zeng, S.; Huang, W.; Zheng, X.; Cheng, L.; Zhang, Z.; Wang, J.; Shen, Z. *Eur. J. Med. Chem.* **2021,** *210,* 112981. doi:10.1016/j.ejmech.2020.112981
38. Jan, M.; Sperling, A. S.; Ebert, B. L. *Nat. Rev. Clin. Oncol.* **2021,** *18,* 401–417. doi:10.1038/s41571-021-00479-z
39. Bemis, T. A.; La Clair, J. J.; Burkart, M. D. *J. Med. Chem.* **2021,** *64,* 8042–8052. doi:10.1021/acs.jmedchem.1c00482
40. Troup, R. I.; Fallan, C.; Baud, M. G. J. *Explor. Targeted Anti-Tumor Ther.* **2020,** *1,* 273–312. doi:10.37349/etat.2020.00018
41. Gadd, M. S.; Testa, A.; Lucas, X.; Chan, K.-H.; Chen, W.; Lamont, D. J.; Zengerle, M.; Ciulli, A. *Nat. Chem. Biol.* **2017,** *13,* 514–521. doi:10.1038/nchembio.2329
42. Roy, M. J.; Winkler, S.; Hughes, S. J.; Whitworth, C.; Galant, M.; Farnaby, W.; Rumpel, K.; Ciulli, A. *ACS Chem. Biol.* **2019,** *14,* 361–368. doi:10.1021/acschembio.9b00092
43. Zorba, A.; Nguyen, C.; Xu, Y.; Starr, J.; Borzilleri, K.; Smith, J.; Zhu, H.; Farley, K. A.; Ding, W.; Schiemer, J.; Feng, X.; Chang, J. S.; Uccello, D. P.; Young, J. A.; Garcia-Irrizary, C. N.; Czabaniuk, L.; Schuff, B.; Oliver, R.; Montgomery, J.; Hayward, M. M.; Coe, J.; Chen, J.; Niosi, M.; Luthra, S.; Shah, J. C.; El-Kattan, A.; Qiu, X.; West, G. M.; Noe, M. C.; Shanmugasundaram, V.; Gilbert, A. M.; Brown, M. F.; Calabrese, M. F. *Proc. Natl. Acad. Sci. U. S. A.* **2018,** *115,* E7285–E7292. doi:10.1073/pnas.1803662115
44. Rana, S.; Mallareddy, J. R.; Singh, S.; Boghean, L.; Natarajan, A. *Cancers* **2021,** *13,* 5506. doi:10.3390/cancers13215506
45. Malumbres, M.; Barbacid, M. *Nat. Rev. Cancer* **2009,** *9,* 153–166. doi:10.1038/nrc2602
46. Sonawane, Y. A.; Taylor, M. A.; Napoleon, J. V.; Rana, S.; Contreras, J. I.; Natarajan, A. *J. Med. Chem.* **2016,** *59,* 8667–8684. doi:10.1021/acs.jmedchem.6b00150
47. Kour, S.; Rana, S.; Contreras, J. I.; King, H. M.; Robb, C. M.; Sonawane, Y. A.; Bendjennat, M.; Crawford, A. J.; Barger, C. J.; Kizhake, S.; Luo, X.; Hollingsworth, M. A.; Natarajan, A. *Mol. Pharmacol.* **2019,** *96,* 419–429. doi:10.1124/mol.119.116855
48. Finn, R. S.; Dering, J.; Conklin, D.; Kalous, O.; Cohen, D. J.; Desai, A. J.; Ginther, C.; Atefi, M.; Chen, I.; Fowst, C.; Los, G.; Slamon, D. J. *Breast Cancer Res.* **2009,** *11,* R77. doi:10.1186/bcr2419
49. Rader, J.; Russell, M. R.; Hart, L. S.; Nakazawa, M. S.; Belcastro, L. T.; Martinez, D.; Li, Y.; Carpenter, E. L.; Attiyeh, E. F.; Diskin, S. J.; Kim, S.; Parasuraman, S.; Caponigro, G.; Schnepp, R. W.; Wood, A. C.; Pawel, B.; Cole, K. A.; Maris, J. M. *Clin. Cancer Res.* **2013,** *19,* 6173–6182. doi:10.1158/1078-0432.ccr-13-1675
50. Corona, S. P.; Generali, D. *Drug Des., Dev. Ther.* **2018,** *12,* 321–330. doi:10.2147/dddt.s137783
51. Ye, X.; Zhu, C.; Harper, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98,* 1682–1686. doi:10.1073/pnas.98.4.1682
52. Zarrin, A. A.; Bao, K.; Lupardus, P.; Vucic, D. *Nat. Rev. Drug Discovery* **2021,** *20,* 39–63. doi:10.1038/s41573-020-0082-8
53. Zhang, M.; Zhang, L.; Hei, R.; Li, X.; Cai, H.; Wu, X.; Zheng, Q.; Cai, C. *Am. J. Cancer Res.* **2021,** *11,* 1913–1935.
54. Wang, H.; Yuan, S.; Zheng, Q.; Zhang, S.; Zhang, Q.; Ji, S.; Wang, W.; Cao, Y.; Guo, Y.; Yang, X.; Geng, H.; Yang, F.; Xi, S.; Jin, G.; Zhang, J.; Gao, Q.; Bernards, R.; Qin, W.; Wang, C. *Gastroenterology* **2024,** *166,* 1130–1144.e8. doi:10.1053/j.gastro.2024.01.025
55. Tigan, A.-S.; Bellutti, F.; Kollmann, K.; Tebb, G.; Sexl, V. *Oncogene* **2016,** *35,* 3083–3091. doi:10.1038/onc.2015.407
56. Brand, M.; Jiang, B.; Bauer, S.; Donovan, K. A.; Liang, Y.; Wang, E. S.; Nowak, R. P.; Yuan, J. C.; Zhang, T.; Kwiatkowski, N.; Müller, A. C.; Fischer, E. S.; Gray, N. S.; Winter, G. E. *Cell Chem. Biol.* **2019,** *26,* 300–306.e9. doi:10.1016/j.chembiol.2018.11.006
57. Anderson, N. A.; Cryan, J.; Ahmed, A.; Dai, H.; McGonagle, G. A.; Rozier, C.; Benowitz, A. B. *Bioorg. Med. Chem. Lett.* **2020,** *30,* 127106. doi:10.1016/j.bmcl.2020.127106
58. Rana, S.; Bendjennat, M.; Kour, S.; King, H. M.; Kizhake, S.; Zahid, M.; Natarajan, A. *Bioorg. Med. Chem. Lett.* **2019,** *29,* 1375–1379. doi:10.1016/j.bmcl.2019.03.035
59. Zhao, B.; Burgess, K. *Chem. Commun.* **2019,** *55,* 2704–2707. doi:10.1039/c9cc00163h
60. Su, S.; Yang, Z.; Gao, H.; Yang, H.; Zhu, S.; An, Z.; Wang, J.; Li, Q.; Chandarlapaty, S.; Deng, H.; Wu, W.; Rao, Y. *J. Med. Chem.* **2019,** *62,* 7575–7582. doi:10.1021/acs.jmedchem.9b00871
61. Mandal, R.; Becker, S.; Strebhardt, K. *Cancers* **2021,** *13,* 2181. doi:10.3390/cancers13092181
62. Asghar, U.; Witkiewicz, A. K.; Turner, N. C.; Knudsen, E. S. *Nat. Rev. Drug Discovery* **2015,** *14,* 130–146. doi:10.1038/nrd4504
63. Morales, F.; Giordano, A. *Cell Cycle* **2016,** *15,* 519–527. doi:10.1080/15384101.2016.1138186
64. Zhou, F.; Chen, L.; Cao, C.; Yu, J.; Luo, X.; Zhou, P.; Zhao, L.; Du, W.; Cheng, J.; Xie, Y.; Chen, Y. *Eur. J. Med. Chem.* **2020,** *187,* 111952. doi:10.1016/j.ejmech.2019.111952
65. Bian, J.; Ren, J.; Li, Y.; Wang, J.; Xu, X.; Feng, Y.; Tang, H.; Wang, Y.; Li, Z. *Bioorg. Chem.* **2018,** *81,* 373–381. doi:10.1016/j.bioorg.2018.08.028
66. Dawson, M. A.; Kouzarides, T. *Cell* **2012,** *150,* 12–27. doi:10.1016/j.cell.2012.06.013
67. Kouzarides, T. *Cell* **2007,** *128,* 693–705. doi:10.1016/j.cell.2007.02.005
68. Strahl, B. D.; Allis, C. D. *Nature* **2000,** *403,* 41–45. doi:10.1038/47412
69. Bannister, A. J.; Kouzarides, T. *Methods Enzymol.* **2003,** *376,* 269–288. doi:10.1016/s0076-6879(03)76018-2
70. Pogo, B. G.; Allfrey, V. G.; Mirsky, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **1966,** *55,* 805–812. doi:10.1073/pnas.55.4.805
71. Fraga, M. F.; Ballestar, E.; Villar-Garea, A.; Boix-Chornet, M.; Espada, J.; Schotta, G.; Bonaldi, T.; Haydon, C.; Ropero, S.; Petrie, K.; Iyer, N. G.; Pérez-Rosado, A.; Calvo, E.; Lopez, J. A.; Cano, A.; Calasanz, M. J.; Colomer, D.; Piris, M. Á.; Ahn, N.; Imhof, A.; Caldas, C.; Jenuwein, T.; Esteller, M. *Nat. Genet.* **2005,** *37,* 391–400. doi:10.1038/ng1531
72. Ho, T. C. S.; Chan, A. H. Y.; Ganesan, A. *J. Med. Chem.* **2020,** *63,* 12460–12484. doi:10.1021/acs.jmedchem.0c00830
73. Li, Z.; Zhu, W.-G. *Int. J. Biol. Sci.* **2014,** *10,* 757–770. doi:10.7150/ijbs.9067

74. Huang, M.; Zhang, J.; Yan, C.; Li, X.; Zhang, J.; Ling, R. *Bioorg. Chem.* **2019,** *91,* 103184. doi:10.1016/j.bioorg.2019.103184
75. Li, T.; Zhang, C.; Hassan, S.; Liu, X.; Song, F.; Chen, K.; Zhang, W.; Yang, J. *J. Hematol. Oncol.* **2018,** *11,* 111. doi:10.1186/s13045-018-0654-9
76. Eckschlager, T.; Plch, J.; Stiborova, M.; Hrabeta, J. *Int. J. Mol. Sci.* **2017,** *18,* 1414. doi:10.3390/ijms18071414
77. Kalin, J. H.; Bergman, J. A. *J. Med. Chem.* **2013,** *56,* 6297–6313. doi:10.1021/jm4001659
78. Kanno, K.; Kanno, S.; Nitta, H.; Uesugi, N.; Sugai, T.; Masuda, T.; Wakabayashi, G.; Maesawa, C. *Oncol. Rep.* **2012,** *28,* 867–873. doi:10.3892/or.2012.1898
79. Bradbury, C. A.; Khanim, F. L.; Hayden, R.; Bunce, C. M.; White, D. A.; Drayson, M. T.; Craddock, C.; Turner, B. M. *Leukemia* **2005,** *19,* 1751–1759. doi:10.1038/sj.leu.2403910
80. Bazzaro, M.; Lin, Z.; Santillan, A.; Lee, M. K.; Wang, M.-C.; Chan, K. C.; Bristow, R. E.; Mazitschek, R.; Bradner, J.; Roden, R. B. S. *Clin. Cancer Res.* **2008,** *14,* 7340–7347. doi:10.1158/1078-0432.ccr-08-0642
81. West, A. C.; Johnstone, R. W. *J. Clin. Invest.* **2014,** *124,* 30–39. doi:10.1172/jci69738
82. McClure, J. J.; Li, X.; Chou, C. J. *Adv. Cancer Res.* **2018,** *138,* 183–211. doi:10.1016/bs.acr.2018.02.006
83. Pulya, S.; Amin, S. A.; Adhikari, N.; Biswas, S.; Jha, T.; Ghosh, B. *Pharmacol. Res.* **2021,** *163,* 105274. doi:10.1016/j.phrs.2020.105274
84. He, X.; Li, Z.; Zhuo, X.-T.; Hui, Z.; Xie, T.; Ye, X.-Y. *Recent Pat. Anti-Cancer Drug Discovery* **2020,** *15,* 32–48. doi:10.2174/1574892815666200217125419
85. Yang, K.; Song, Y.; Xie, H.; Wu, H.; Wu, Y.-T.; Leisten, E. D.; Tang, W. *Bioorg. Med. Chem. Lett.* **2018,** *28,* 2493–2497. doi:10.1016/j.bmcl.2018.05.057
86. Yang, K.; Wu, H.; Zhang, Z.; Leisten, E. D.; Nie, X.; Liu, B.; Wen, Z.; Zhang, J.; Cunningham, M. D.; Tang, W. *ACS Med. Chem. Lett.* **2020,** *11,* 575–581. doi:10.1021/acsmedchemlett.0c00046
87. Xiong, Y.; Donovan, K. A.; Eleuteri, N. A.; Kirmani, N.; Yue, H.; Razov, A.; Krupnick, N. M.; Nowak, R. P.; Fischer, E. S. *Cell Chem. Biol.* **2021,** *28,* 1514–1527.e4. doi:10.1016/j.chembiol.2021.07.002
88. Nowak, R. P.; DeAngelo, S. L.; Buckley, D.; He, Z.; Donovan, K. A.; An, J.; Safaee, N.; Jedrychowski, M. P.; Ponthier, C. M.; Ishoey, M.; Zhang, T.; Mancias, J. D.; Gray, N. S.; Bradner, J. E.; Fischer, E. S. *Nat. Chem. Biol.* **2018,** *14,* 706–714. doi:10.1038/s41589-018-0055-y
89. Cuadrado, A.; Nebreda, A. R. *Biochem. J.* **2010,** *429,* 403–417. doi:10.1042/bj20100323
90. Yokota, T.; Wang, Y. *Gene* **2016,** *575,* 369–376. doi:10.1016/j.gene.2015.09.030
91. Yurtsever, Z.; Scheaffer, S. M.; Romero, A. G.; Holtzman, M. J.; Brett, T. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015,** *71,* 790–799. doi:10.1107/s1399004715001212
92. Escós, A.; Risco, A.; Alsina-Beauchamp, D.; Cuenda, A. *Front. Cell Dev. Biol.* **2016,** *4,* 31. doi:10.3389/fcell.2016.00031
93. Madkour, M. M.; Anbar, H. S.; El-Gamal, M. I. *Eur. J. Med. Chem.* **2021,** *213,* 113216. doi:10.1016/j.ejmech.2021.113216
94. Roche, O.; Fernández-Aroca, D. M.; Arconada-Luque, E.; García-Flores, N.; Mellor, L. F.; Ruiz-Hidalgo, M. J.; Sánchez-Prieto, R. *Int. J. Mol. Sci.* **2020,** *21,* 7524. doi:10.3390/ijms21207524
95. Xu, W.; Liu, R.; Dai, Y.; Hong, S.; Dong, H.; Wang, H. *Int. J. Biol. Sci.* **2021,** *17,* 4036–4046. doi:10.7150/ijbs.63537
96. Anton, D. B.; Ducati, R. G.; Timmers, L. F. S. M.; Laufer, S.; Goettert, M. I. *Cancers* **2021,** *13,* 2077. doi:10.3390/cancers13092077
97. Wu, P.; Nielsen, T. E.; Clausen, M. H. *Trends Pharmacol. Sci.* **2015,** *36,* 422–439. doi:10.1016/j.tips.2015.04.005
98. Haller, V.; Nahidino, P.; Forster, M.; Laufer, S. A. *Expert Opin. Ther. Pat.* **2020,** *30,* 453–466. doi:10.1080/13543776.2020.1749263
99. Hammaker, D.; Firestein, G. S. *Ann. Rheum. Dis.* **2010,** *69,* i77–i82. doi:10.1136/ard.2009.119479
100. Martín-Acosta, P.; Xiao, X. *Eur. J. Med. Chem.* **2021,** *210,* 112993. doi:10.1016/j.ejmech.2020.112993
101. Stathis, A.; Bertoni, F. *Cancer Discovery* **2018,** *8,* 24–36. doi:10.1158/2159-8290.cd-17-0605
102. Belkina, A. C.; Denis, G. V. *Nat. Rev. Cancer* **2012,** *12,* 465–477. doi:10.1038/nrc3256
103. Dawson, M. A.; Kouzarides, T.; Huntly, B. J. P. *N. Engl. J. Med.* **2012,** *367,* 647–657. doi:10.1056/nejmra1112635
104. Doroshow, D. B.; Eder, J. P.; LoRusso, P. M. *Ann. Oncol.* **2017,** *28,* 1776–1787. doi:10.1093/annonc/mdx157
105. Cochran, A. G.; Conery, A. R.; Sims, R. J., III. *Nat. Rev. Drug Discovery* **2019,** *18,* 609–628. doi:10.1038/s41573-019-0030-7
106. Fu, Y.; Zhang, Y.; Sun, H. *Eur. J. Med. Chem.* **2021,** *226,* 113853. doi:10.1016/j.ejmech.2021.113853
107. Zengerle, M.; Chan, K.-H.; Ciulli, A. *ACS Chem. Biol.* **2015,** *10,* 1770–1777. doi:10.1021/acschembio.5b00216
108. Testa, A.; Hughes, S. J.; Lucas, X.; Wright, J. E.; Ciulli, A. *Angew. Chem., Int. Ed.* **2020,** *59,* 1727–1734. doi:10.1002/anie.201914396
109. Cheung, K. L.; Kim, C.; Zhou, M.-M. *Front. Mol. Biosci.* **2021,** *8,* 728777. doi:10.3389/fmolb.2021.728777
110. Li, C.; Chen, Y.; Huang, W.; Qiu, Y.; Huang, S.; Zhou, Y.; Zhou, F.; Xu, J.; Ren, X.; Zhang, J.; Wang, Z.; Ding, M.; Ding, K. *JACS Au* **2024,** *4,* 4866–4882. doi:10.1021/jacsau.4c00831
111. Chen, Y.; Jin, J. *Acta Biochim. Biophys. Sin.* **2020,** *52,* 776–790. doi:10.1093/abbs/gmaa053
112. Shaid, S.; Brandts, C. H.; Serve, H.; Dikic, I. *Cell Death Differ.* **2013,** *20,* 21–30. doi:10.1038/cdd.2012.72
113. Hershko, A.; Ciechanover, A. *Annu. Rev. Biochem.* **1992,** *61,* 761–807. doi:10.1146/annurev.bi.61.070192.003553
114. Pickart, C. M. *Annu. Rev. Biochem.* **2001,** *70,* 503–533. doi:10.1146/annurev.biochem.70.1.503
115. Zheng, N.; Shabek, N. *Annu. Rev. Biochem.* **2017,** *86,* 129–157. doi:10.1146/annurev-biochem-060815-014922
116. Kannt, A.; Đikić, I. *Cell Chem. Biol.* **2021,** *28,* 1014–1031. doi:10.1016/j.chembiol.2021.04.007
117. He, S.; Dong, G.; Cheng, J.; Wu, Y.; Sheng, C. *Med. Res. Rev.* **2022,** *42,* 1280–1342. doi:10.1002/med.21877
118. Ishida, T.; Ciulli, A. *SLAS Discovery* **2021,** *26,* 484–502. doi:10.1177/2472555220965528
119. Wang, C.; Zhang, Y.; Wu, Y.; Xing, D. *Eur. J. Med. Chem.* **2021,** *225,* 113749. doi:10.1016/j.ejmech.2021.113749
120. Sun, X.; Gao, H.; Yang, Y.; He, M.; Wu, Y.; Song, Y.; Tong, Y.; Rao, Y. *Signal Transduction Targeted Ther.* **2019,** *4,* 64. doi:10.1038/s41392-019-0101-6

121. Bassi, Z. I.; Fillmore, M. C.; Miah, A. H.; Chapman, T. D.; Maller, C.; Roberts, E. J.; Davis, L. C.; Lewis, D. E.; Galwey, N. W.; Waddington, K. E.; Parravicini, V.; Macmillan-Jones, A. L.; Gongora, C.; Humphreys, P. G.; Churcher, I.; Prinjha, R. K.; Tough, D. F. *ACS Chem. Biol.* **2018,** *13,* 2862–2867. doi:10.1021/acschembio.8b00705
122. Silva, M. C.; Ferguson, F. M.; Cai, Q.; Donovan, K. A.; Nandi, G.; Patnaik, D.; Zhang, T.; Huang, H.-T.; Lucente, D. E.; Dickerson, B. C.; Mitchison, T. J.; Fischer, E. S.; Gray, N. S.; Haggarty, S. J. *eLife* **2019,** *8,* e45457. doi:10.7554/elife.45457
123. Bricelj, A.; Steinebach, C.; Kuchta, R.; Gütschow, M.; Sosič, I. *Front. Chem. (Lausanne, Switz.)* **2021,** *9,* 707317. doi:10.3389/fchem.2021.707317
124. Xi, J.-Y.; Zhang, R.-Y.; Chen, K.; Yao, L.; Li, M.-Q.; Jiang, R.; Li, X.-Y.; Fan, L. *Bioorg. Chem.* **2022,** *125,* 105848. doi:10.1016/j.bioorg.2022.105848
125. Rossari, F.; Minutolo, F.; Orciuolo, E. *J. Hematol. Oncol.* **2018,** *11,* 84. doi:10.1186/s13045-018-0624-2
126. Lai, A. C.; Toure, M.; Hellerschmied, D.; Salami, J.; Jaime-Figueroa, S.; Ko, E.; Hines, J.; Crews, C. M. *Angew. Chem., Int. Ed.* **2016,** *55,* 807–810. doi:10.1002/anie.201507634
127. Druker, B. J.; Tamura, S.; Buchdunger, E.; Ohno, S.; Segal, G. M.; Fanning, S.; Zimmermann, J.; Lydon, N. B. *Nat. Med.* **1996,** *2,* 561–566. doi:10.1038/nm0596-561
128. Zimmermann, J.; Buchdunger, E.; Mett, H.; Meyer, T.; Lydon, N. B. *Bioorg. Med. Chem. Lett.* **1997,** *7,* 187–192. doi:10.1016/s0960-894x(96)00601-4
129. An, X.; Tiwari, A. K.; Sun, Y.; Ding, P.-R.; Ashby, C. R., Jr.; Chen, Z.-S. *Leuk. Res.* **2010,** *34,* 1255–1268. doi:10.1016/j.leukres.2010.04.016
130. De Dominici, M.; Porazzi, P.; Xiao, Y.; Chao, A.; Tang, H.-Y.; Kumar, G.; Fortina, P.; Spinelli, O.; Rambaldi, A.; Peterson, L. F.; Petruk, S.; Barletta, C.; Mazo, A.; Cingolani, G.; Salvino, J. M.; Calabretta, B. *Blood* **2020,** *135,* 1560–1573. doi:10.1182/blood.2019003604
131. Steinebach, C.; Ng, Y. L. D.; Sosič, I.; Lee, C.-S.; Chen, S.; Lindner, S.; Vu, L. P.; Bricelj, A.; Haschemi, R.; Monschke, M.; Steinwarz, E.; Wagner, K. G.; Bendas, G.; Luo, J.; Gütschow, M.; Krönke, J. *Chem. Sci.* **2020,** *11,* 3474–3486. doi:10.1039/d0sc00167h
132. Bruhn, M. A.; Pearson, R. B.; Hannan, R. D.; Sheppard, K. E. *Growth Factors* **2010,** *28,* 394–408. doi:10.3109/08977194.2010.518616
133. Halland, N.; Schmidt, F.; Weiss, T.; Saas, J.; Li, Z.; Czech, J.; Dreyer, M.; Hofmeister, A.; Mertsch, K.; Dietz, U.; Strübing, C.; Nazare, M. *ACS Med. Chem. Lett.* **2015,** *6,* 73–78. doi:10.1021/ml5003376
134. Sherk, A. B.; Frigo, D. E.; Schnackenberg, C. G.; Bray, J. D.; Laping, N. J.; Trizna, W.; Hammond, M.; Patterson, J. R.; Thompson, S. K.; Kazmin, D.; Norris, J. D.; McDonnell, D. P. *Cancer Res.* **2008,** *68,* 7475–7483. doi:10.1158/0008-5472.can-08-1047
135. Gong, G. Q.; Wang, K.; Dai, X.-C.; Zhou, Y.; Basnet, R.; Chen, Y.; Yang, D.-H.; Lee, W.-J.; Buchanan, C. M.; Flanagan, J. U.; Shepherd, P. R.; Chen, Y.; Wang, M.-W. *Acta Pharmacol. Sin.* **2018,** *39,* 1902–1912. doi:10.1038/s41401-018-0087-6
136. Tovell, H.; Testa, A.; Zhou, H.; Shpiro, N.; Crafter, C.; Ciulli, A.; Alessi, D. R. *ACS Chem. Biol.* **2019,** *14,* 2024–2034. doi:10.1021/acschembio.9b00505
137. Weng, G.; Shen, C.; Cao, D.; Gao, J.; Dong, X.; He, Q.; Yang, B.; Li, D.; Wu, J.; Hou, T. *Nucleic Acids Res.* **2021,** *49,* D1381–D1387. doi:10.1093/nar/gkaa807
138. Bondeson, D. P.; Smith, B. E.; Burslem, G. M.; Buhimschi, A. D.; Hines, J.; Jaime-Figueroa, S.; Wang, J.; Hamman, B. D.; Ishchenko, A.; Crews, C. M. *Cell Chem. Biol.* **2018,** *25,* 78–87.e5. doi:10.1016/j.chembiol.2017.09.010
139. Remillard, D.; Buckley, D. L.; Paulk, J.; Brien, G. L.; Sonnett, M.; Seo, H.-S.; Dastjerdi, S.; Wühr, M.; Dhe-Paganon, S.; Armstrong, S. A.; Bradner, J. E. *Angew. Chem., Int. Ed.* **2017,** *56,* 5738–5743. doi:10.1002/anie.201611281
140. Winter, G. E.; Buckley, D. L.; Paulk, J.; Roberts, J. M.; Souza, A.; Dhe-Paganon, S.; Bradner, J. E. *Science* **2015,** *348,* 1376–1381. doi:10.1126/science.aab1433
141. Donovan, K. A.; Ferguson, F. M.; Bushman, J. W.; Eleuteri, N. A.; Bhunia, D.; Ryu, S.; Tan, L.; Shi, K.; Yue, H.; Liu, X.; Dobrovolsky, D.; Jiang, B.; Wang, J.; Hao, M.; You, I.; Teng, M.; Liang, Y.; Hatcher, J.; Li, Z.; Manz, T. D.; Groendyke, B.; Hu, W.; Nam, Y.; Sengupta, S.; Cho, H.; Shin, I.; Agius, M. P.; Ghobrial, I. M.; Ma, M. W.; Che, J.; Buhrlage, S. J.; Sim, T.; Gray, N. S.; Fischer, E. S. *Cell* **2020,** *183,* 1714–1731.e10. doi:10.1016/j.cell.2020.10.038
142. Zhang, J.; Zhang, Y.; Wang, J.; Xia, Y.; Zhang, J.; Chen, L. *Signal Transduction Targeted Ther.* **2024,** *9,* 211. doi:10.1038/s41392-024-01911-3
143. Nguyen, K. Q. T.; Nguyen, H. H.; Phung, H. T. T.; Chung, K. L.; Vu, T. Y. *J. Biomol. Struct. Dyn.* **2025,** *43,* 5260–5267. doi:10.1080/07391102.2024.2308778
144. Domostegui, A.; Nieto-Barrado, L.; Perez-Lopez, C.; Mayor-Ruiz, C. *Chem. Soc. Rev.* **2022,** *51,* 5498–5517. doi:10.1039/d2cs00197g
145. Simonetta, K. R.; Taygerly, J.; Boyle, K.; Basham, S. E.; Padovani, C.; Lou, Y.; Cummins, T. J.; Yung, S. L.; von Soly, S. K.; Kayser, F.; Kuriyan, J.; Rape, M.; Cardozo, M.; Gallop, M. A.; Bence, N. F.; Barsanti, P. A.; Saha, A. *Nat. Commun.* **2019,** *10,* 1402. doi:10.1038/s41467-019-09358-9
146. Zhao, Y.; Guan, Y.-Y.; Zhao, F.; Yu, T.; Zhang, S.-J.; Zhang, Y.-Z.; Duan, Y.-C.; Zhou, X.-L. *Eur. J. Med. Chem.* **2022,** *231,* 114144. doi:10.1016/j.ejmech.2022.114144
147. Yamanaka, S.; Horiuchi, Y.; Matsuoka, S.; Kido, K.; Nishino, K.; Maeno, M.; Shibata, N.; Kosako, H.; Sawasaki, T. *Nat. Commun.* **2022,** *13,* 183. doi:10.1038/s41467-021-27818-z
148. Cho, H.; Lee, S. *Eur. J. Med. Chem.* **2026,** *303,* 118485. doi:10.1016/j.ejmech.2025.118485
149. Kargbo, R. B. *ACS Med. Chem. Lett.* **2025,** *16,* 523–525. doi:10.1021/acsmedchemlett.5c00111
150. Morris, R.; Kershaw, N. J.; Babon, J. J. *Protein Sci.* **2018,** *27,* 1984–2009. doi:10.1002/pro.3519
151. Kaneshige, A.; Yang, Y.; Bai, L.; Wang, M.; Xu, R.; Mallik, L.; Chinnaswamy, K.; Metwally, H.; Wang, Y.; McEachern, D.; Tošović, J.; Yang, C.-Y.; Kirchhoff, P. D.; Meagher, J. L.; Stuckey, J. A.; Wang, S. *J. Med. Chem.* **2025,** *68,* 5125–5151. doi:10.1021/acs.jmedchem.4c01009
152. Liu, P.-K.; Wang, Z.; Li, L. *Curr. Opin. Chem. Biol.* **2025,** *87,* 102605. doi:10.1016/j.cbpa.2025.102605
153. Bushweller, J. H. *Nat. Rev. Cancer* **2019,** *19,* 611–624. doi:10.1038/s41568-019-0196-7

154. Popow, J.; Farnaby, W.; Gollner, A.; Kofink, C.; Fischer, G.; Wurm, M.; Zollman, D.; Wijaya, A.; Mischerikow, N.; Hasenoehrl, C.; Prokofeva, P.; Arnhof, H.; Arce-Solano, S.; Bell, S.; Boeck, G.; Diers, E.; Frost, A. B.; Goodwin-Tindall, J.; Karolyi-Oezguer, J.; Khan, S.; Klawatsch, T.; Koegl, M.; Kousek, R.; Kratochvil, B.; Kropatsch, K.; Lauber, A. A.; McLennan, R.; Olt, S.; Peter, D.; Petermann, O.; Roessler, V.; Stolt-Bergner, P.; Strack, P.; Strauss, E.; Trainor, N.; Vetma, V.; Whitworth, C.; Zhong, S.; Quant, J.; Weinstabl, H.; Kuster, B.; Ettmayer, P.; Ciulli, A. *Science* **2024,** *385,* 1338–1347. doi:10.1126/science.adm8684
155. Sun, Q.; Wang, H.; Xie, J.; Wang, L.; Mu, J.; Li, J.; Ren, Y.; Lai, L. *Chem. Rev.* **2025,** *125,* 6309–6365. doi:10.1021/acs.chemrev.4c00969

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.49

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Using generative AI to transform peptide hits into small molecule leads

Joshua Mills*[1,2] and Yu Heng Lau*[1,2]

Perspective

Open Access

Address:
[1]School of Chemistry, The University of Sydney, Camperdown NSW 2006, Australia and [2]ARC Centre for Innovations in Peptide and Protein Science, The University of Sydney, Australia

Email:
Joshua Mills* - joshua.mills@sydney.edu.au; Yu Heng Lau* - yuheng.lau@sydney.edu.au

* Corresponding author

Keywords:
diffusion models; drug discovery; generative AI; peptides; small molecules

Beilstein J. Org. Chem. **2026,** *22,* 672–679.
https://doi.org/10.3762/bjoc.22.51

Received: 13 February 2026
Accepted: 17 April 2026
Published: 30 April 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2026 Mills and Lau; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

There is a wealth of structural data on peptides that bind potential drug targets, serving as rich inspiration for designing small molecule inhibitors which can recapitulate key binding interactions while improving pharmacokinetic properties. With the rapid advancement of artificial intelligence capabilities including powerful generative models, there is significant potential for new AI-based tools to expedite the structure-based transformation of peptide hits into small molecule leads. In this Perspective, we highlight how AI-enabled prediction and design tools can potentially span the entire workflow from peptide to small molecule: target protein structure prediction, de novo peptide binder generation, diffusion models for generating novel small molecule scaffolds, and deep-learning predictors of binding affinity to rapidly triage candidates.

Introduction

In drug discovery, peptides serve as accessible starting points for designing molecules that bind and inhibit a chosen protein target. Many enzyme targets (e.g., proteases) natively process peptides as their substrates. Peptide domains are also common motifs in protein–protein interactions, mediating the binding between protein partners that form functional complexes [1-3]. Furthermore, powerful high-throughput library screening technologies such as mRNA and phage display have revolutionised our ability to identify novel peptide binders, frequently generating binders with nanomolar affinity directly from screening campaigns [4-6].

While there has been increasing recognition of the translational potential of peptides in drug discovery, significant optimisation is typically required to convert a naïve peptide into a bona fide drug candidate [7]. Many promising peptide hits fail to translate beyond academic research, hampered by pharmacokinetic liabilities such as poor metabolic stability, limited membrane

permeability, and low oral bioavailability. In some cases, peptides may be suitable for therapeutic use after installing a small number of chemical modifications (e.g., lipidation of peptide hormones such as GLP-1 receptor agonists). In other cases, however, challenging medicinal chemistry is required to achieve clinical efficacy, as exemplified by the complex development of the heavily modified tricyclic peptide MK-0616, a PCSK9 inhibitor in phase III clinical trials that was originally derived from an mRNA display screen of monocyclic peptides [8].

As an alternative to optimising peptide-based molecules, knowledge of key binding interactions exhibited by potent peptide binders can be leveraged to inform small-molecule design. The field of peptidomimetics seeks to design small molecules that mimic the binding mode of a peptide, while mitigating the pharmacokinetic liabilities associated with unmodified peptides. This peptide-first approach to drug discovery is particularly valuable for challenging targets such as protein–protein interactions, where traditional small molecule high-throughput screening often fails to yield hits [9]. While the term peptidomimetics has been used to describe a variety of molecular classes, including modified peptides and molecular scaffolds for display of amino acid side-chains, this Perspective focuses on non-peptidic small molecules with the appropriate shape and functionality to recapitulate the pharmacophore of the original peptide (Class D according to the classification proposed by Grossmann and co-workers [10]).

The challenge of transforming a peptide binder into a small molecule mimic is not new [11]. Indeed, there are well-known historical examples of blockbuster drugs derived from native peptide substrates. A classic example is the ACE inhibitor captopril, an analogue of a snake venom peptide, the development of which has been cited as an early success story for structure-based rational drug design [12,13]. Despite the long history, there is still no straightforward and generalisable workflow with a reasonable success rate for achieving the transformation from peptide to small molecule inhibitor.

With the rapid maturation of computational tools based on artificial intelligence (AI), this Perspective highlights selected nascent examples where AI has been used for small molecule design by leveraging data on peptide binders, and proposes potential opportunities where generative AI and machine learning (ML) tools may augment various stages throughout the pipeline from peptide hit discovery to small molecule lead.

Perspective

Current workflows for designing small molecules that mimic peptide pharmacophores

Illustrative example of a non-AI workflow

To understand where generative AI may play a role in transforming peptides into small molecules, we first briefly outline how traditional non-AI tools are typically used in the field. Starting from an experimental co-crystal structure of the bound complex, standard medicinal chemistry principles are used to probe structure–activity relationships (SAR) and determine the key interactions that form a minimal pharmacophore, supported by classical physics-based molecular modelling methods such as molecular dynamics (MD) and docking (Figure 1).

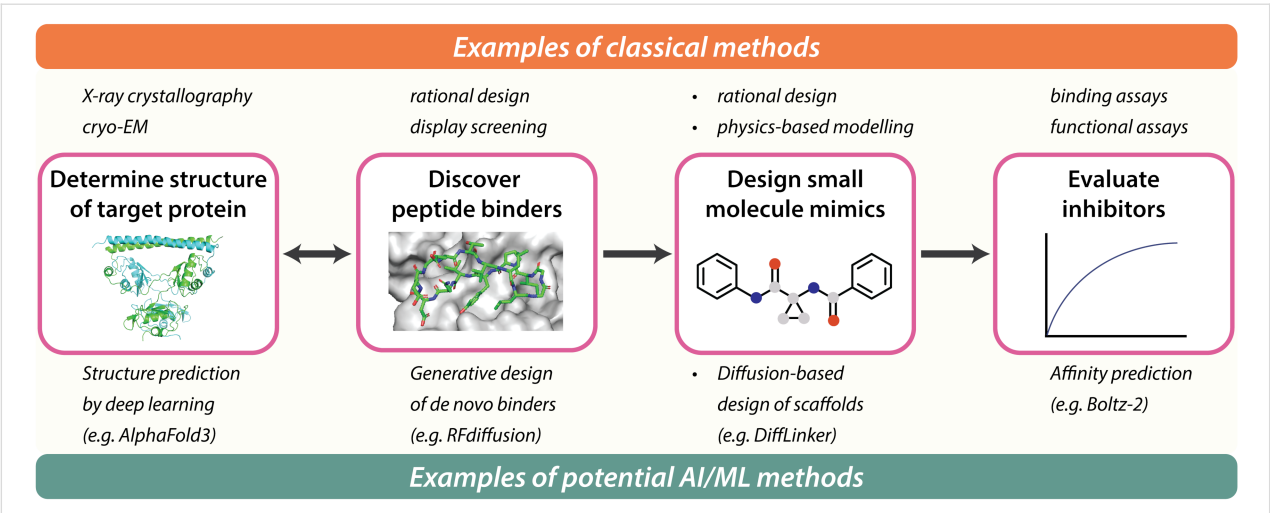


Figure 1: An example of a typical workflow from peptide hit discovery to small molecule evaluation. The top half shows typical methods used in each stage of the workflow, while the bottom half shows potential AI-based methods to augment the workflow. The left two images were generated using PyMOL (The PyMOL Molecular Graphics System, Version 3.1 Schrödinger, LLC.)

As an illustrative example, Yoshida and co-workers at Shionogi reported the design of small molecule inhibitors of nicotinamide *N*-methyltransferase [14] and β-herpesvirus proteases [15], starting from cyclic peptide hits obtained by mRNA display with flexible in vitro translation (also known as RaPID, random non-standard peptides integrated discovery [4]). Key interactions that comprise the pharmacophores were experimentally determined by alanine scanning and other amino acid substitutions to explore SAR, coupled with analysis of co-crystal structures obtained for the bound complexes, including computational analysis of the amino acid interactions (SiteMap [16]) and hydration in the binding pocket (grid inhomogeneous solvation theory [17] in AmberTools [18]). The pharmacophores were then used as the basis for virtual screening with docking (Glide [19,20]) to obtain candidate small molecules for experimental validation in functional assays.

Nascent examples of ML approaches

In contrast to physics-based models for extracting pharmacophores and predicting small molecule binders, ML methods are now emerging as alternatives for accomplishing these tasks.

An early application of ML for small molecule design based on peptide datasets was reported in 2023 by Hou and co-workers, where models were trained to classify ligands as ghrelin receptor binders [21]. The complete dataset for training and evaluation consisted of 548 peptides/peptidomimetics and 2193 small molecules spanning known binders, non-binders, and random compounds, each represented using one-hot encoding. A variety of well-established machine-learning classifiers were tested, including random forests, support vector machines, and gradient boosted decision trees. Strategies that used a larger proportion of the dataset for training showed reasonable predictive ability when evaluated on the remaining compounds as the validation set, as well as on an external set of 30 binders that was not in the original dataset. This result hints at the potential of ML models in predicting small molecule binders based on key peptide attributes, although the classifier was trained on more small molecules than peptides/peptidomimetics, making it unclear if peptide and peptidomimetic data alone would provide sufficient training to predict small molecule binders. The generality of this approach to other targets with no known small molecule binders has also yet to be evaluated.

Also in 2023, a deep learning-driven molecular generator for designing peptidomimetics was reported by Nakamura, Bajorath and co-workers [22]. Named DeepCubist, the method superimposes scaffolds from an sp^3-rich skeleton library onto the target peptide to identify spatially aligned structures. Once a scaffold is selected, a transformer model converts the three-dimensional scaffolds into complete chemical structures by decorating the backbone with heteroatoms and unsaturated bonds. The resulting small molecules are designed to be synthetically tractable, although synthetic accessibility was noted to be somewhat restricted by the highly sp^3-rich nature of the generated molecules. New iterations of these methods have since been reported by the authors [23], though experimental validation of predictions remains to be reported at the time of writing this Perspective.

Opportunities for using generative AI and ML tools

Here, we highlight how multiple stages of the typical peptide to small molecule drug discovery workflow may be amenable to AI-enabled structure-based tools (Table 1). Starting from a target protein structure (which may be modelled by AI), generative models can be used to design de novo peptide binders. Upon identifying the minimal peptide pharmacophore, fragment-linking diffusion-based models can be used to recapitulate key binding features on a novel molecular scaffold. Subsequently, machine learning models can assist in triaging generated candidates prior to more resource-intensive modelling, laboratory synthesis, and evaluation in functional assays.

Deep learning for protein structure prediction

The growing wealth of structural data in the Protein Data Bank, along with ongoing improvements in computational processing power, has fuelled the success of protein structure prediction tools based on deep learning algorithms. Most notably, DeepMind's AlphaFold2 heralded a breakthrough in protein structure prediction with its exemplary performance in the 14th Critical Assessment of protein Structure Prediction (CASP14) experiment in 2020 [24]. Since then, numerous models with advanced functionality have been released (Table 1), including AlphaFold3 which incorporates diffusion-based components, providing a range of options for predicting the structure of protein targets (including protein complexes and interactions with other biomolecules) where no experimental structure has been reported. While the accuracy of these models continues to increase with time, protein structures modelled on experimental data (X-ray crystallography, cryo-electron microscopy) still remain the gold standard as starting points for structure-based drug design.

Generative AI for peptide design

Building off the success of structure prediction tools, a major area of growth has been in new AI-based tools for designing de novo protein binders, many of which can be adapted to peptides. The release of RFdiffusion [42] and ProteinMPNN [35] from David Baker and co-workers at the Institute of Protein Design in 2022–23 popularised the use of diffusion models to build de novo protein backbones followed by sequence

Table 1: Selected examples of structure-based AI/ML tools for potential end-to-end coverage of the peptide to small molecule pipeline. Tools are listed in alphabetical order. Year corresponds to the final publication date if peer-reviewed, or preprint date otherwise.

Capability	Tools	Developer
predicting structure of target protein (and complexes)	AlphaFold3 (2024) [25]	Google DeepMind
	Boltz-1 (2024) [26]	MIT Jameel Clinic
	Chai-1 (2024) [27]	Chai Discovery
	ESMFold (2023) [28]	Meta AI
	RoseTTAFold (2021) [29]	Institute for Protein Design, UW
generative design of peptide binders (including cyclic peptides)	AfCycDesign (2025) [30]	Institute for Protein Design, UW
	BindCraft (2025) [31]	Correia Lab, EPFL
	BoltzGen (2025) [32]	MIT
	EvoBind2 (2025) [33]	Bryant Lab, Stockholm
	PepMimic (2025) [34]	Ma Lab, Tsinghua University
	ProteinMPNN (2022) [35]	Institute for Protein Design, UW
	RFpeptides (2025) [36]	Institute for Protein Design, UW
fransforming peptides hits into small molecules	DiffLinker (2024) [37]	Correia Lab, EPFL
	Peptide2Mol (2025) [38]	Ma Lab, Tsinghua University
	ShEPhERD (2024) [39]	Coley Research Group, MIT
predicting affinities	Boltz-2 (2025) [40]	MIT and Recursion
	RosettaVS (2024) [41]	Institute for Protein Design, UW

design. As with structure prediction, there has since been an explosion of new methods improving upon these initial methods (Table 1), including RFdiffusion2 [43] and RFdiffusion3 [44], along with integrated pipelines such as BindCraft [31] which are built upon tools such as ProteinMPNN and AlphaFold2. Other design tools that are specific for peptide binders also include the evolution-based EvoBind2 that only requires the protein target sequence as an input [33], and the diffusion-based PepMimic that uses structural information from binding interfaces [34]. Ongoing evaluation by independent groups suggests that reliable one-shot prediction of binders is currently target-dependent [45], and continued improvements may see these methods reach parity with traditional peptide discovery methods in the future.

Diffusion models to generate small molecule mimics

The reliable design of synthesisable small molecules that can accurately mimic a peptide pharmacophore remains the core challenge in the overall peptide to small molecule workflow. There are many AI tools for general ligand-based drug design under active development, spanning a range of generative architectures (e.g., variational autoencoders, generative adversarial networks). Here, we specifically highlight the potential of using diffusion-based models for generating small molecules that recapitulate key components of a peptide pharmacophore.

In general, diffusion models function by progressively adding Gaussian noise to training data over a series of timesteps (forward diffusion). A neural network is then trained to reverse this process, iteratively reconstructing the original data structure through predicting and removing noise (reverse diffusion). Once trained, the model initialises and iteratively denoises random Gaussian noise to generate novel outputs (Figure 2a). Diffusion models have been particularly adept at generating new visual outputs such as images and videos, hence their suitability for de novo molecular design is not unexpected. In the case of small molecule design, the chemical inputs can be represented in different formats depending on the level of information and complexity required [46]. One-dimensional SMILES strings are the simplest form of molecular representation that lacks spatial information, while two- and three-dimensional molecular graph and point cloud representations are less compact but carry additional spatial information that is likely to be critical for small molecule binder design (Figure 2b).

In the context of transforming peptides into small molecules, diffusion models can be applied using a workflow reminiscent of fragment-based drug design. Starting from the key elements of a peptide pharmacophore (e.g., critical side-chains or backbone functional groups that interact with the target), diffusion models can generate candidate small molecules that spatially link or merge all these elements together. To preserve three-dimensional spatial information, the following examples of diffusion models all use three-dimensional atomic point clouds and an E(3)-equivariant (or SE(3)-equivariant) graph neural network for symmetry-consistent modelling of molecular structures [47].

DiffLinker is an example of a conditional diffusion model developed by Correia and co-workers that enables linking of multiple fragments in a predefined orientation [37]. After pretraining on linkers extracted from molecules originating from several compound databases (ZINC [48], CASF [49], GEOM [50]), input fragments were represented as three-dimen-

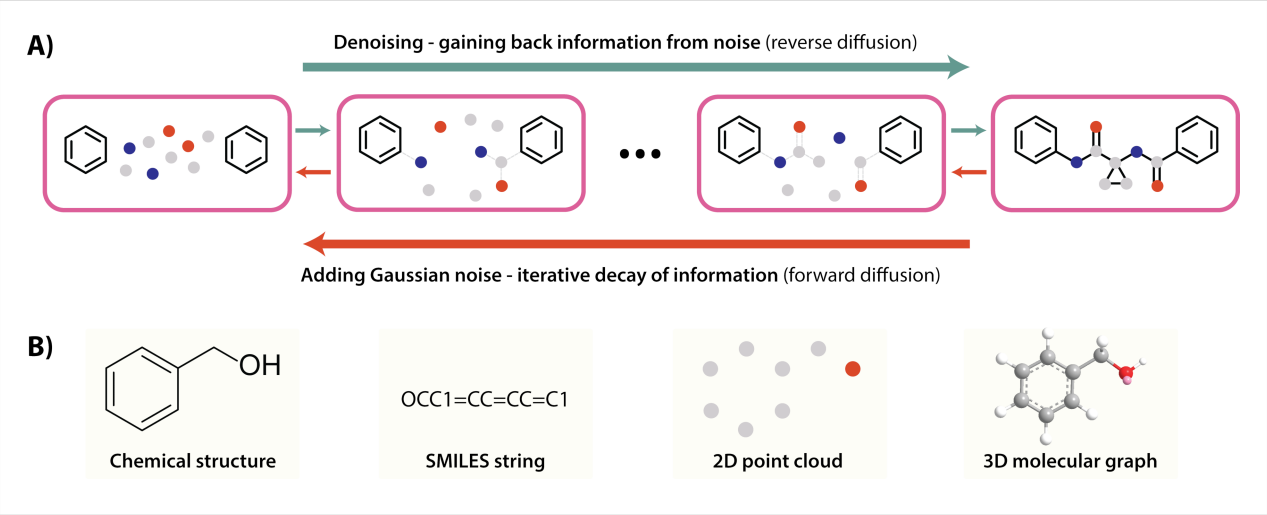


Figure 2: A) A non-technical schematic of a diffusion model, in which random noise is iteratively removed to reconstruct chemically valid molecules, based on training with a supplied dataset. B) Models can use different chemical representations as inputs, such as SMILES strings, point clouds, or molecular graphs.

sional point clouds, while the appropriately-sized linker was iteratively generated and refined by sampling atom types and positions using Gaussian noise as the starting point. The model could also be trained on a protein–ligand dataset (from Binding MOAD [51]) to factor in the constraints of typical binding pockets.

ShEPhERD is an interaction-aware diffusion model developed by Coley and co-workers that is capable of bioisosteric fragment merging [39]. For fragment merging, the model was trained on a subset from the MOSES database [52]. Then starting with a set of published fragment data from an experimental screen, ShEPhERD generated small molecule candidates with high similarity in electrostatic and pharmacophore scoring metrics.

Meanwhile, Peptide2Mol is a diffusion model developed by Ma and co-workers that leverages peptide-binder structural data to generate small molecules within the protein pocket. It is trained on a range of target-bound complexes, including small molecule ligands as well as peptides and protein binders [38]. Similar to DiffLinker, small molecule generation is initiated from Gaussian noise, then after iterative denoising within the target pocket, an additional pocket-aware refinement step is used to remove steric clashes and shape complementary issues.

Development of this class of generative molecular design tools is still in its infancy. Outputs typically require manual curation or integration with traditional physics-based molecular modelling to filter out unstable functional groups, strained rings and other unfavourable conformations. Experimental validation of predicted small molecules also remains to be conducted. Nevertheless, continued improvements in model design and quality input data may see these tools mature in future, potentially mirroring the trajectory of the aforementioned structure prediction and binder design tools.

Deep learning for predicting binding affinities

As a brief note, AI tools are also available for preliminary evaluation of small molecule candidates, providing potential alternatives to traditional computationally expensive physics-based workflows such as docking, MD, and free-energy perturbation (FEP) calculations for affinity estimation. For example, Boltz-2 features a strong emphasis on affinity prediction in addition to its core capability of structure prediction [40]. Other tools integrate AI methods into virtual screening pipelines, such as RosettaVS which trains a target-specific neural network during physics-based docking to prioritise candidates for more expensive computational modelling [41].

Outlook

There is still significant progress that must be made before a complete AI-assisted computational workflow for transforming peptide hits into small molecule leads can be reliably achieved. While the early stages of structure prediction and peptide binder generation are rapidly maturing, small molecule design and evaluation methods are far less advanced, likely reflecting the more diverse and complex nature of small molecule chemical space, and the relatively sparse coverage of this space in training datasets. Nevertheless, early indications suggest that there could be a significant future role for generative AI in peptide to small molecule drug discovery workflows, even though ultimately, all predictions require validation in an experimental laboratory setting.

An open question is whether a peptide-first approach to drug discovery will continue to hold advantages in the future. Indeed, many target-based computational drug discovery tools that are ligand-agnostic are under development, along with small molecule prediction methods that do not require experimental data on target-specific peptide binders. Nevertheless, we envisage that given the ease of binder design and screening, synthetic accessibility, and in vitro assay validation, peptides will continue to play a significant role in supporting the small molecule drug discovery pipeline.

Just as the Critical Assessment of Structure Prediction (CASP) competition supported the development of protein structure prediction methods, the emergence of small molecule hit prediction competitions, such as the CACHE (Critical Assessment of Computational Hit-finding Experiments) challenge [53], are driving a pivotal shift toward rigorous, blind, and prospectively validated benchmarking. As methods improve, we expect such competitions to raise the bar and focus on increasingly challenging targets, such as protein–protein interactions without any known small molecule binders. Additionally, a focus on 'make-on-demand' constraints would allow generative models to move beyond the commercially available chemical space and into the realm of truly novel, de novo chemical space.

One potential barrier to rapid development and uptake of AI in drug discovery is the restricted accessibility of some models and datasets. While the development of proprietary AI-based platforms in private organisations such as large pharmaceutical companies and small start-up enterprises is driving local innovation within their respective organisations, such platforms are rarely shared for public use and may only be accessible through licensing at substantial cost. At the same time, one could argue that these platforms may have never been developed without the corporate structure that private enterprise offers. Although such debates are not new to drug discovery, we anticipate that the balance between proprietary and open-source development will have a disproportionate influence on the future of AI-based drug discovery, due to the data-intensive nature of the discipline.

Finally, if the full breadth of potential applications of AI in drug discovery come to fruition, one could imagine a fully-automated pipeline from target identification and virtual screening through to retrosynthetic planning, and robotics for hit-to-lead optimisation and functional evaluation in cellular models. The examples highlighted in this Perspective represent only a few of the many ongoing developments in the broader field. With the field moving at such a rapid pace, it is difficult to forecast which AI tools will become the next field standard akin to AlphaFold2 for structure prediction, and which others will fall short of more traditional computational or experimental methods. Only time will tell whether AI optimism in drug discovery is well-founded or not.

Funding

JM acknowledges funding from an Australian Government Research Training Program scholarship, and a Centre for Drug Discovery Innovation and Sydney Precision Data Science Centre supplementary scholarship. YHL acknowledges funding from the National Health and Medical Research Foundation Investigator grant (2041692), Synergy grant (2045600), and funding from the Cancer Institute of New South Wales Career Development Fellowship (2024/CDF1222).

ORCID® iDs

Joshua Mills - https://orcid.org/0009-0009-3790-9772
Yu Heng Lau - https://orcid.org/0000-0001-6560-8410

Data Availability Statement

Data sharing is not applicable as no new data was generated or analyzed in this study.

References

1. Bullock, B. N.; Jochim, A. L.; Arora, P. S. *J. Am. Chem. Soc.* **2011,** *133,* 14220–14223. doi:10.1021/ja206074j
2. Watkins, A. M.; Arora, P. S. *ACS Chem. Biol.* **2014,** *9,* 1747–1754. doi:10.1021/cb500241y
3. Gavenonis, J.; Sheneman, B. A.; Siegert, T. R.; Eshelman, M. R.; Kritzer, J. A. *Nat. Chem. Biol.* **2014,** *10,* 716–722. doi:10.1038/nchembio.1580
4. Passioura, T.; Suga, H. *Chem. Commun.* **2017,** *53,* 1931–1940. doi:10.1039/c6cc06951g
5. Huang, Y.; Wiedmann, M. M.; Suga, H. *Chem. Rev.* **2019,** *119,* 10360–10391. doi:10.1021/acs.chemrev.8b00430
6. Deyle, K.; Kong, X.-D.; Heinis, C. *Acc. Chem. Res.* **2017,** *50,* 1866–1874. doi:10.1021/acs.accounts.7b00184
7. Gare, C. L.; White, A. M.; Malins, L. R. *Trends Biochem. Sci.* **2025,** *50,* 467–480. doi:10.1016/j.tibs.2025.01.009
8. Li, H.; Thaisrivongs, D. A.; Shang, G.; Chen, Y.; Chen, Q.; Tan, L.; Xiao, K.-J.; Larson, R. T.; Kuethe, J. T.; Lee, J.; Deprez, N. R.; Nolting, A. F.; Poirier, M.; Bulger, P. G.; Regalado, E. L.; Biba, M.; Tsay, F.-R.; DaSilva, J.; Prier, C. K.; Strulson, C. A.; Zawatzky, K.; Liu, Z.; Newman, J. A.; Sokolowsky, K.; Tang, W.; Hullen, K.; Thakur, N.; Welch, C.; Patel, S.; He, Y.; Xu, J.; Variankaval, N.; Klapars, A.; Kong, J.; Desmond, R.; Varsolona, R.; Maligres, P. E.; Pons Siepermann, C. A.; Robison, L.; Piou, T.; Hartmanshenn, C.; Chandra, A.; Patel, A.; Becker, M. R.; Liu, G.; Duan, J.; Wan, B.; Xiao, C.; Yuan, Y.; Cao, X.; Chen, L.; Yi, R.; Wu, Z.; Feng, M.; Li, D.; Song, Z.; Dong, Y.; Sun, J.; Li, B.; Shao, G.; Campeau, L.-C.; Yin, J. *J. Am. Chem. Soc.* **2025,** *147,* 11036–11048. doi:10.1021/jacs.4c15966
9. Wells, J. A.; McClendon, C. L. *Nature* **2007,** *450,* 1001–1009. doi:10.1038/nature06526

10. Pelay-Gimeno, M.; Glas, A.; Koch, O.; Grossmann, T. N. *Angew. Chem., Int. Ed.* **2015,** *54,* 8896–8927. doi:10.1002/anie.201412070
11. Gante, J. *Angew. Chem., Int. Ed. Engl.* **1994,** *33,* 1699–1720. doi:10.1002/anie.199416991
12. Nemec, K.; Schubert-Zsilavecz, M. *Pharm. Unserer Zeit* **2003,** *32,* 11–16. doi:10.1002/pauz.200390001
13. Cushman, D. W.; Ondetti, M. A. *Hypertension* **1991,** *17,* 589–592. doi:10.1161/01.hyp.17.4.589
14. Yoshida, S.; Uehara, S.; Kondo, N.; Takahashi, Y.; Yamamoto, S.; Kameda, A.; Kawagoe, S.; Inoue, N.; Yamada, M.; Yoshimura, N.; Tachibana, Y. *J. Med. Chem.* **2022,** *65,* 10655–10673. doi:10.1021/acs.jmedchem.2c00919
15. Yoshida, S.; Sako, Y.; Nikaido, E.; Ueda, T.; Kozono, I.; Ichihashi, Y.; Nakahashi, A.; Onishi, M.; Yamatsu, Y.; Kato, T.; Nishikawa, J.; Tachibana, Y. *ACS Med. Chem. Lett.* **2023,** *14,* 1558–1566. doi:10.1021/acsmedchemlett.3c00359
16. Halgren, T. A. *J. Chem. Inf. Model.* **2009,** *49,* 377–389. doi:10.1021/ci800324m
17. Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. *J. Chem. Phys.* **2012,** *137,* 044101. doi:10.1063/1.4733951
18. Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T. J.; Götz, A. W.; Gohlke, H.; Izadi, S.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kurtzman, T.; Lee, T.-S.; Li, P.; Liu, J.; Luchko, T.; Luo, R.; Manathunga, M.; Machado, M. R.; Nguyen, H. M.; O'Hearn, K. A.; Onufriev, A. V.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Risheh, A.; Schott-Verdugo, S.; Shajan, A.; Swails, J.; Wang, J.; Wei, H.; Wu, X.; Wu, Y.; Zhang, S.; Zhao, S.; Zhu, Q.; Cheatham, T. E., III; Roe, D. R.; Roitberg, A.; Simmerling, C.; York, D. M.; Nagan, M. C.; Merz, K. M., Jr. *J. Chem. Inf. Model.* **2023,** *63,* 6183–6191. doi:10.1021/acs.jcim.3c01153
19. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. *J. Med. Chem.* **2004,** *47,* 1750–1759. doi:10.1021/jm030644s
20. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004,** *47,* 1739–1749. doi:10.1021/jm0306430
21. Liu, W.; Hopkins, A. M.; Yan, P.; Du, S.; Luyt, L. G.; Li, Y.; Hou, J. *Mol. Diversity* **2023,** *27,* 2239–2255. doi:10.1007/s11030-022-10555-w
22. Umedera, K.; Yoshimori, A.; Chen, H.; Kouji, H.; Nakamura, H.; Bajorath, J. *J. Comput.-Aided Mol. Des.* **2023,** *37,* 107–115. doi:10.1007/s10822-022-00493-y
23. Yoshimori, A.; Bajorath, J. *Eur. J. Med. Chem. Rep.* **2025,** *13,* 100249. doi:10.1016/j.ejmcr.2025.100249
24. 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, S. A. A.; Ballard, A. J.; 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, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. *Nature* **2021,** *596,* 583–589. doi:10.1038/s41586-021-03819-2
25. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlin, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. *Nature* **2024,** *630,* 493–500. doi:10.1038/s41586-024-07487-w
26. Wohlwend, J.; Corso, G.; Passaro, S.; Getz, N.; Reveiz, M.; Leidal, K.; Swiderski, W.; Atkinson, L.; Portnoi, T.; Chinn, I.; Silterra, J.; Jaakkola, T.; Barzilay, R. *bioRxiv* **2025**. doi:10.1101/2024.11.19.624167
27. Discovery, C.; Boitreaud, J.; Dent, J.; McPartlon, M.; Meier, J.; Reis, V.; Rogozhnikov, A.; Wu, K. *bioRxiv* **2024**. doi:10.1101/2024.10.10.615955
28. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. *Science* **2023,** *379,* 1123–1130. doi:10.1126/science.ade2574
29. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; Millán, C.; Park, H.; Adams, C.; Glassman, C. R.; DeGiovanni, A.; Pereira, J. H.; Rodrigues, A. V.; van Dijk, A. A.; Ebrecht, A. C.; Opperman, D. J.; Sagmeister, T.; Buhlheller, C.; Pavkov-Keller, T.; Rathinaswamy, M. K.; Dalwadi, U.; Yip, C. K.; Burke, J. E.; Garcia, K. C.; Grishin, N. V.; Adams, P. D.; Read, R. J.; Baker, D. *Science* **2021,** *373,* 871–876. doi:10.1126/science.abj8754
30. Rettie, S. A.; Campbell, K. V.; Bera, A. K.; Kang, A.; Kozlov, S.; Bueso, Y. F.; De La Cruz, J.; Ahlrichs, M.; Cheng, S.; Gerben, S. R.; Lamb, M.; Murray, A.; Adebomi, V.; Zhou, G.; DiMaio, F.; Ovchinnikov, S.; Bhardwaj, G. *Nat. Commun.* **2025,** *16,* 4730. doi:10.1038/s41467-025-59940-7
31. Pacesa, M.; Nickel, L.; Schellhaas, C.; Schmidt, J.; Pyatova, E.; Kissling, L.; Barendse, P.; Choudhury, J.; Kapoor, S.; Alcaraz-Serna, A.; Cho, Y.; Ghamary, K. H.; Vinué, L.; Yachnin, B. J.; Wollacott, A. M.; Buckley, S.; Westphal, A. H.; Lindhoud, S.; Georgeon, S.; Goverde, C. A.; Hatzopoulos, G. N.; Gönczy, P.; Muller, Y. D.; Schwank, G.; Swarts, D. C.; Vecchio, A. J.; Schneider, B. L.; Ovchinnikov, S.; Correia, B. E. *Nature* **2025,** *646,* 483–492. doi:10.1038/s41586-025-09429-6
32. Stark, H.; Faltings, F.; Choi, M.; Xie, Y.; Hur, E.; O'Donnell, T.; Bushuiev, A.; Uçar, T.; Passaro, S.; Mao, W.; Reveiz, M.; Bushuiev, R.; Pluskal, T.; Sivic, J.; Kreis, K.; Vahdat, A.; Ray, S.; Goldstein, J. T.; Savinov, A.; Hambalek, J. A.; Gupta, A.; Taquiri-Diaz, D. A.; Zhang, Y.; Hatstat, A. K.; Arada, A.; Kim, N. H.; Tackie-Yarboi, E.; Boselli, D.; Schnaider, L.; Liu, C. C.; Li, G.-W.; Hnisz, D.; Sabatini, D. M.; DeGrado, W. F.; Wohlwend, J.; Corso, G.; Barzilay, R.; Jaakkola, T. *bioRxiv* **2025**. doi:10.1101/2025.11.20.689494
33. Li, Q.; Vlachos, E. N.; Bryant, P. *Commun. Chem.* **2025,** *8,* 211. doi:10.1038/s42004-025-01601-3
34. Kong, X.; Jiao, R.; Lin, H.; Guo, R.; Huang, W.; Ma, W.-Y.; Wang, Z.; Liu, Y.; Ma, J. *Nat. Biomed. Eng.* **2025,** 1507. doi:10.1038/s41551-025-01507-4

35. Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. *Science* **2022,** *378,* 49–56. doi:10.1126/science.add2187
36. Rettie, S. A.; Juergens, D.; Adebomi, V.; Bueso, Y. F.; Zhao, Q.; Leveille, A. N.; Liu, A.; Bera, A. K.; Wilms, J. A.; Üffing, A.; Kang, A.; Brackenbrough, E.; Lamb, M.; Gerben, S. R.; Murray, A.; Levine, P. M.; Schneider, M.; Vasireddy, V.; Ovchinnikov, S.; Weiergräber, O. H.; Willbold, D.; Kritzer, J. A.; Mougous, J. D.; Baker, D.; DiMaio, F.; Bhardwaj, G. *Nat. Chem. Biol.* **2025,** *21,* 1948–1956. doi:10.1038/s41589-025-01929-w
37. Igashov, I.; Stärk, H.; Vignac, C.; Schneuing, A.; Satorras, V. G.; Frossard, P.; Welling, M.; Bronstein, M.; Correia, B. *Nat. Mach. Intell.* **2024,** *6,* 417–427. doi:10.1038/s42256-024-00815-9
38. He, X.; Zhang, Y.; Lin, H.; Peng, X.; Kong, X.; Li, M.; Ma, J. *arXiv* **2025,** 2511.04984. doi:10.48550/arxiv.2511.04984
39. Adams, K.; Abeywardane, K.; Fromer, J.; Coley, C. W. *arXiv* **2025,** 2411.04130. doi:10.48550/arxiv.2411.04130
40. Passaro, S.; Corso, G.; Wohlwend, J.; Reveiz, M.; Thaler, S.; Somnath, V. R.; Getz, N.; Portnoi, T.; Roy, J.; Stark, H.; Kwabi-Addo, D.; Beaini, D.; Jaakkola, T.; Barzilay, R. *bioRxiv* **2025**. doi:10.1101/2025.06.14.659707
41. Zhou, G.; Rusnac, D.-V.; Park, H.; Canzani, D.; Nguyen, H. M.; Stewart, L.; Bush, M. F.; Nguyen, P. T.; Wulff, H.; Yarov-Yarovoy, V.; Zheng, N.; DiMaio, F. *Nat. Commun.* **2024,** *15,* 7761. doi:10.1038/s41467-024-52061-7
42. Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Hanikel, N.; Pellock, S. J.; Courbet, A.; Sheffler, W.; Wang, J.; Venkatesh, P.; Sappington, I.; Torres, S. V.; Lauko, A.; De Bortoli, V.; Mathieu, E.; Ovchinnikov, S.; Barzilay, R.; Jaakkola, T. S.; DiMaio, F.; Baek, M.; Baker, D. *Nature* **2023,** *620,* 1089–1100. doi:10.1038/s41586-023-06415-8
43. Ahern, W.; Yim, J.; Tischer, D.; Salike, S.; Woodbury, S. M.; Kim, D.; Kalvet, I.; Kipnis, Y.; Coventry, B.; Altae-Tran, H. R.; Bauer, M. S.; Barzilay, R.; Jaakkola, T. S.; Krishna, R.; Baker, D. *Nat. Methods* **2026,** *23,* 96–105. doi:10.1038/s41592-025-02975-x
44. Butcher, J.; Krishna, R.; Mitra, R.; Brent, R. I.; Li, Y.; Corley, N.; Kim, P. T.; Funk, J.; Mathis, S.; Salike, S.; Muraishi, A.; Eisenach, H.; Thompson, T. R.; Chen, J.; Politanska, Y.; Sehgal, E.; Coventry, B.; Zhang, O.; Qiang, B.; Didi, K.; Kazman, M.; DiMaio, F.; Baker, D. *bioRxiv* **2025**. doi:10.1101/2025.09.18.676967
45. Filius, M.; Patsos, T.; Minnee, H.; Turco, G.; Chong, H. E.; Liu, J.; Gnatzy, M.; Rooth, R. S. M.; Liu, A. C. H.; Ta, R. D. T.; Rijk, I. H. A.; Ziani, S.; Boxman, F. J.; Pomplun, S. J. *ACS Chem. Biol.* **2025,** *20,* 2991–2998. doi:10.1021/acschembio.5c00774
46. Alakhdar, A.; Poczos, B.; Washburn, N. *J. Chem. Inf. Model.* **2024,** *64,* 7238–7256. doi:10.1021/acs.jcim.4c01107
47. Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. *Nat. Commun.* **2022,** *13,* 2453. doi:10.1038/s41467-022-29939-5
48. Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.* **2005,** *45,* 177–182. doi:10.1021/ci049714+
49. Su, M.; Yang, Q.; Du, Y.; Feng, G.; Liu, Z.; Li, Y.; Wang, R. *J. Chem. Inf. Model.* **2019,** *59,* 895–913. doi:10.1021/acs.jcim.8b00545
50. Axelrod, S.; Gómez-Bombarelli, R. *Sci. Data* **2022,** *9,* 185. doi:10.1038/s41597-022-01288-4
51. Hu, L.; Benson, M. L.; Smith, R. D.; Lerner, M. G.; Carlson, H. A. *Proteins: Struct., Funct., Bioinf.* **2005,** *60,* 333–340. doi:10.1002/prot.20512
52. Polykovskiy, D.; Zhebrak, A.; Sanchez-Lengeling, B.; Golovanov, S.; Tatanov, O.; Belyaev, S.; Kurbanov, R.; Artamonov, A.; Aladinskiy, V.; Veselov, M.; Kadurin, A.; Johansson, S.; Chen, H.; Nikolenko, S.; Aspuru-Guzik, A.; Zhavoronkov, A. *Front. Pharmacol.* **2020,** *11,* 565644. doi:10.3389/fphar.2020.565644
53. Ackloo, S.; Al-awar, R.; Amaro, R. E.; Arrowsmith, C. H.; Azevedo, H.; Batey, R. A.; Bengio, Y.; Betz, U. A. K.; Bologa, C. G.; Chodera, J. D.; Cornell, W. D.; Dunham, I.; Ecker, G. F.; Edfeldt, K.; Edwards, A. M.; Gilson, M. K.; Gordijo, C. R.; Hessler, G.; Hillisch, A.; Hogner, A.; Irwin, J. J.; Jansen, J. M.; Kuhn, D.; Leach, A. R.; Lee, A. A.; Lessel, U.; Morgan, M. R.; Moult, J.; Muegge, I.; Oprea, T. I.; Perry, B. G.; Riley, P.; Rousseaux, S. A. L.; Saikatendu, K. S.; Santhakumar, V.; Schapira, M.; Scholten, C.; Todd, M. H.; Vedadi, M.; Volkamer, A.; Willson, T. M. *Nat. Rev. Chem.* **2022,** *6,* 287–295. doi:10.1038/s41570-022-00363-z

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.51

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Synthesis and biological evaluation of new brassinosteroid analogs with C-22 benzoate function

María Núñez[1], Camila Escobar[2], Mario Párraga[2], Mauricio Soto[1], Luis Espinoza-Catalán[1], Katy Díaz[*1] and Andrés F. Olea[*3]

Full Research Paper

Open Access

Address:
[1]Departamento de Química, Universidad Técnica Federico Santa María, Avenida España 1680, Valparaíso 2340000, Chile, [2]Centro Interdisciplinario de Investigación Biomédica e Ingeniería para la Salud, MEDING. Laboratorio de Biología molecular. Escuela de Medicina, Universidad de Valparaíso, Chile and [3]Instituto de Ciencias Aplicadas, Facultad de Ingeniería, Universidad Autónoma de Chile, Av. del Valle Sur 534, Santiago 8580640, Chile

Email:
Katy Díaz[*] - katy.diaz@usm.cl; Andrés F. Olea[*] - andres.olea@uautonoma.cl

* Corresponding author

Keywords:
biological activity; brassinosteroids; BRI1-EMS-Suppressor1; ligand–receptor interaction; structure–activity relationship

Beilstein J. Org. Chem. **2026,** *22,* 753–762.
https://doi.org/10.3762/bjoc.22.57

Received: 29 January 2026
Accepted: 16 April 2026
Published: 18 May 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2026 Núñez et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

The synthesis and structural characterization of new brassinosteroid (BR) analogs, in which substituents with different electronegativities and molecular sizes have been attached to a C-22 benzoate function, are described. The biological activities of all new compounds were evaluated by using the rice lamina inclination test (RLIT) and inhibition of root growth of *Arabidopsis thaliana*. The RLIT data is compared with those previously reported for two series of compounds having the same substitution pattern at C-22 but different structure in ring A. This comparison revealed that a 2α,3α-dihydroxy configuration is more active than a 3-carbonyl or 3β-hydroxy function in this ring. Additionally, the accumulation of the dephosphorylated form of the BES1 protein, which is part of the BRs signaling pathway and control their activity, has been evaluated as well. The results are analyzed in terms of BR analog's structure and compared with binding energies obtained from a docking study. In this way, it is intended to assess the effect of chemical structure on the initial and one intermediate step, and on the final plant response. Our results show that the binding of BR analogs to the active site, which initiate the signaling process, and dephosphorylation of BES1 depend on the structure of BR analogs in a similar way. However, the relationship between the BR analog's structure and the final plant response is different. The dependence on unknown factors which are able to activate or repress genes associated with growth and development is discussed.

Introduction

Brassinosteroids (BRs) are phytohormones that are widespread in the plant kingdom, and are found in very low concentrations, i.e., in the of nano- to micromolar range [1-4]. BRs play various important roles, such as, cell development, vascular differentiation, reproduction, modulation of gene expression [5,6], and other developmental processes [7,8]. Their natural occurrence

and biological activities have been reviewed by different groups, and brassinolide (**1**) and castasterone (**2**) are the most active and widely distributed [4,9,10], whereas teasterone (**3**) and 3-dehydroteasterone (**3a**), which are intermediates in the biosynthetic pathway, exhibit much lower activities [11-14] (Figure 1).

It has been established that BRs are sensed by the extracellular part of a receptor kinase, NRI1, and then the activated BRI1 initiates a series of processes comprising the transduction pathway of BRs [3,15]. The main consequence of this mechanism is the accumulation of two transcription factors: BZR1 (Brassinazole Resistant1) and BES1 (BRI1-EMS-Suppressor1) in the nucleus, which regulates the expression of a large number of genes involved in different physiological processes in *Arabidopsis thaliana* [15-17]. At the same time, BZR1 and BES1 repress thousands of genes associated to BRs biosynthesis, helping to maintain BRs homeostasis [18]. As consequence, BRs responses are inhibited.

Various bioassays have been developed to evaluate the effects of different concentrations of exogenous application of BRs on plant growth. The most widely used are bean second internode bioassay (BSIB) [19], rice lamina inclination test (RLIT) [20,21], and inhibition of root and hypocotyl elongation in *Arabidopsis thaliana* seedlings, among other bioassays. The results show that for the same BRs, the measured effects vary with the bioassay and concentration. For example, in the RLIT the growing effect increases with increasing concentration, whereas on root growth BRs act as root growth promoter at low concentration and the opposite effect is observed at high concentrations [22].

As sensing of BRs by BRI1 is the starting point of the mechanism by which plant responses are produced, this binding process has been extensively studied and it has been shown that bioactivity can be correlated with the binding of BRs to BRI1 [23,24], and some structure–activity relationships (SAR) have been proposed [25,26].

The low abundance of BRs present in plants has prompted the quest of synthetic pathways to obtain these compounds. However, to synthesize molecules with the same stereochemical configuration of natural BRs, both in the rings and side alkyl chain, is an enormous synthetic challenge [27]. Consequently, synthetic BRs and BR analogs are not affordable for large-scale commercial applications and therefore, the search for more accessible and bioactive analogs has become a matter of current interest. For example, derivatives of teasterone (**3**), compounds **4**–**11**, and castasterone, compounds **12**–**14**, with benzoyl function at C-22 (Figure 2) have been synthesized, and their bioactivities have been evaluated by BSIB, RLIT, and inhibition of root and hypocotyl elongation in *A. thaliana* seedlings [28,29].

The obtained results indicate that compounds **10** and **11** are the most active of teasterone analogs [28], whereas fluorinated castasterone analogs **13** and **14** with a fluoro substituent in the *ortho*- and *para*-positions, show interesting activities at intermediate concentrations [29].

Based on the preliminary and promissing results obtained for castasterone analogs **13** and **14**, it was decided to expand this study by considering the 2α,3α diol function in ring A, a ketone function at C-6, and C-22 benzoate-function with various substituents, of different electronegativities and sizes in the aromatic ring. In this way, we intend to assess the effect of these properties on the biological activity. Additionally, by comparing RLIT results obtained for this new series of castasterone (**2**) with those obtained for teasterone (**3**) [30] and 3-dehydroteasterone (3-DT, **3a**) [31] analogs, it would be possible to evaluate the contribution of different functions attached to ring A. Thus, herein the synthesis of analogs **17**–**22**, starting from the precur-

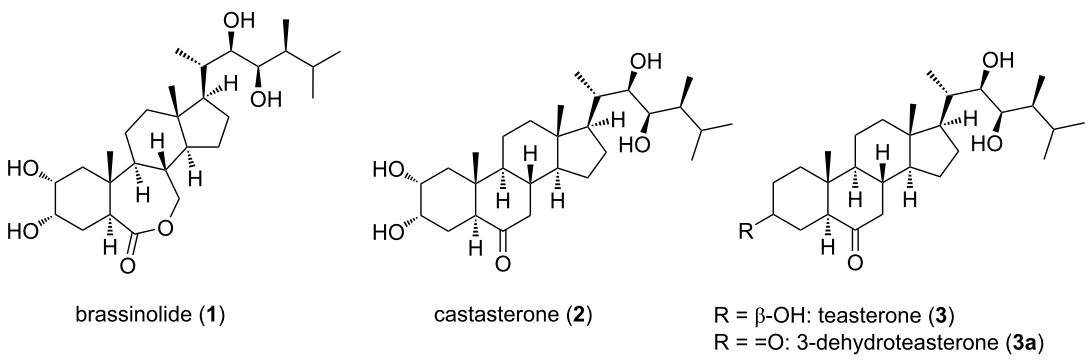


Figure 1: Structures of natural occurring brassinosteroids.

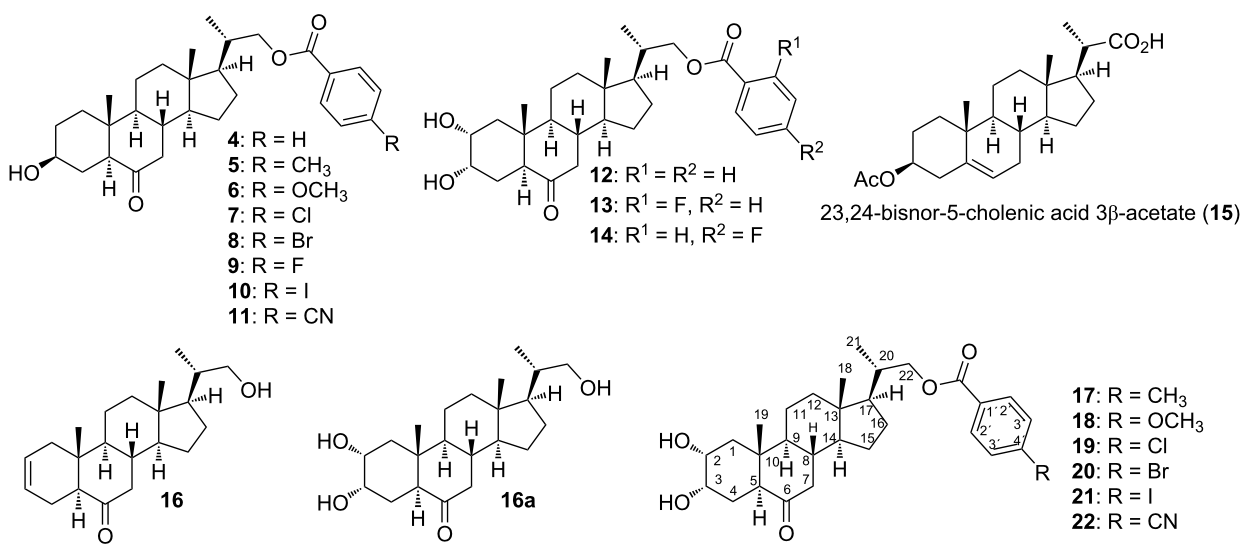


Figure 2: Structures of C-22 benzoate-functionalized brassinosteroid analogs **4–14** and **17–22** prepared from the known precursor **16** [29], analog **16a** [29] and structure of starting material 23,24-bisnor-5-cholenic acid 3β-acetate (**15**).

sor **16**, is described. All compounds have been fully characterized by spectroscopic techniques (IR, 1D, 2D NMR and HRMS).

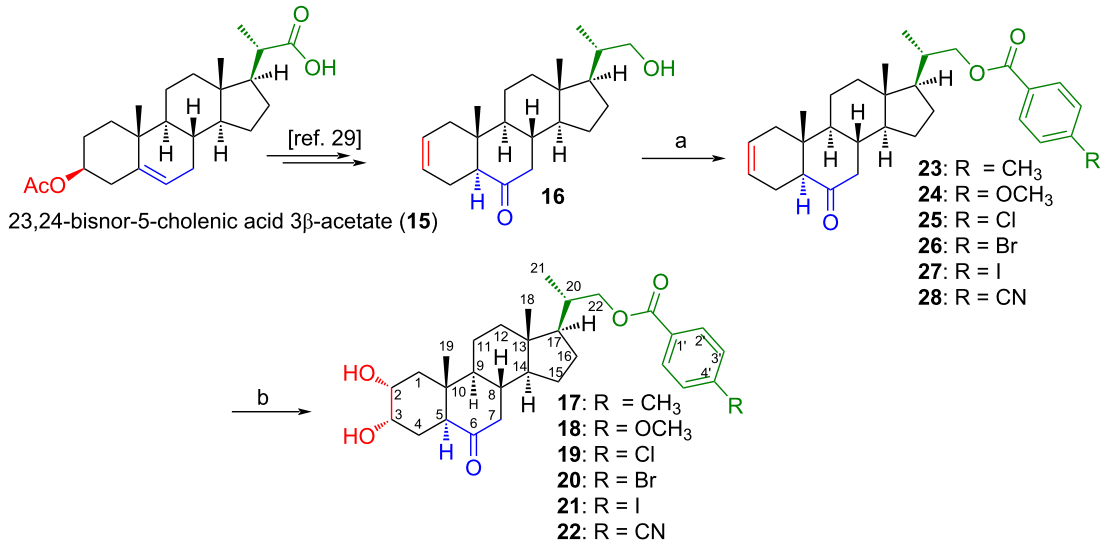
The biological activities have been assessed by RLIT and inhibition of root growth of *A. thaliana*. Additionally, the Western blot technique has been applied for the detection and identification of BES1 in *A. thaliana* seedlings, which has been used as a model organism for this type of study [32].

Results and Discussion

Chemistry

The synthesis of new BR analogs **17–22** (Figure 2) and compounds **23–28** has been carried out using the olefinic precursor **16** according to previously described procedures (Scheme 1) [28,29].

Treatment of olefine **16** with *p*-R-PhCOCl in DMAP/py/CH_2Cl_2, leads to compounds **23–28** with high yields



Scheme 1: Synthesis of compounds **23–28** and new BR analogs **17–22**, from 23,24-bisnor-5-cholenic acid 3β-acetate (**15**). Conditions: a) *p*-R-PhCOCl/CH_2Cl_2/py, DMAP, rt, 2 h, 96.7%, 78.0%, 91.2%, 95.3%, 95.6%, and 93.6% yields for **23–28**, respectively. b) DHQD-CLB, $CH_3SO_2NH_2$, K_2CO_3, $K_3[Fe(CN)_6]$, OsO_4, *t*-BuOH/H_2O (1:1 v/v), rt, 36 h, 77.8%, 67.9%, 65.8%, 84.9%, 77.5%, and 63.4% yields for **17–22**, respectively. The overall yields for compounds **17–22** are 75.2%, 53.0%, 60.0%, 80.9%, 74.1%, and 59.3%, respectively.

(78.0–96.7%). All compounds have been characterized by IR, NMR (1D and 2D) and HRMS spectroscopy techniques (see Supporting Information File 1, Figures S1–S30). Thus, in the ^{1}H NMR spectrum of compound **23**, the presence of signals at δ = 7.91 ppm (2H, d, *J* = 8.1 Hz) and 7.21 ppm (2H, d, *J* = 8.1 Hz) are assigned to aromatic hydrogens H-2' and H-3', respectively. While the observed signal at δ = 2.38 ppm (3H, s) corresponds to CH_3-Ar. Additionally, the signals observed downfield (by the electron-withdrawing effect of the ester function) at δ = 4.27 (1H, dd, *J* = 10.8 and 3.4 Hz, H-22a) and 4.01 (1H, dd, *J* = 10.8 and 7.2 Hz, H-22b), are assigned to carbinolic hydrogens H-22a and H-22b, respectively. In the ^{13}C NMR spectrum, the signals observed at δ = 166.57, 143.31, 129.40, 128.93, 127.65, and 21.52 ppm, are assigned to carbon ArCOO, C-4', C-2', C-3', C-1', and CH_3-Ar of the benzoate group, respectively. All these signals show direct correlation ($^1J_{HC}$ coupling) in the 2D HSQC and $^2J_{HC}$ or $^3J_{HC}$ on the 2D HMBC spectra (Figures S1–S5, Supporting Information File 1). The structural determination of compounds **24**–**28** has been performed in a similar manner to that used for compound **23**.

Next, Sharpless dihydroxylation of double bond on C-2 of compounds **23**–**28**, following a described procedure [29], leads to BR analogs **17**–**22** with good reaction yields (63.4–84.9%) (Scheme 1).

Structural characterization of BR analogs **17**–**22** was accomplished by using ^{13}C NMR data, combined with 2D NMR techniques (HSQC and HMBC). For example, in the ^{1}H NMR spectrum of compound **17**, the three signals appearing at δ = 4.26 (1H, dd, *J* = 10.4 and 3.4 Hz), 4.02–3.97 (2H, m), and 3.71 (1H, bd, *J* = 10.2 Hz) are assigned to carbinolic hydrogens H-22a, H-2, H-22b, and H-3, respectively (Figure S32, Supporting Information File 1). While from the ^{13}C NMR spectrum, the signals observed at δ = 69.65, 68.38, and 68.27, are assigned to carbinolic carbons C-22, C-3, and C-2, respectively (Figure S33, Supporting Information File 1). In this way, the carbinol signals of H and C were confirmed by correlations to $^1J_{HC}$, $^2J_{HC}$, and $^3J_{HC}$ from the ^{1}H,^{13}C HSQC NMR and ^{1}H,^{13}C HMBC NMR spectra data, respectively (Figures S35 and S36, Supporting Information File 1).

Previous studies have shown that Sharpless dihydroxylation of 2,3-steroidal alkenes with OsO_4/DHQD-CLB system, leads to formation of "*cis*" glycols with C-2(*R*) and C-3(*S*) stereochemistry in the A ring of A/B *trans*-fused (5α steroids). This is the consequence of the steric hindrance imposed by the 19-methyl group, which prevents OsO_4 from approaching the top face of the A ring, and thereby favoring the reaction from the lower face of the double bond [29,33,34].

The structural determination of compounds **18**–**22** has been performed in a similar way to that used to characterize compound **17** (Figures S37–S66, Supporting Information File 1). In summary, BRs analogs **17**–**22** have been efficiently synthesized in a two-step pathway starting from **16**. The overall yields for **17**–**22** are 75.2%, 53.0%, 60.0%, 80.9%, 74.1%, and 59.3%, respectively.

Biological activity

Rice lamina inclination test (RLIT)

RLIT is a standard, sensitive, and specific method widely used to evaluate BRs activity [20,21]. BRs induce a greater cell expansion of adaxial cells relative to abaxial cells in the attachment regions, causing laminar inclination in a concentration-dependent manner. Although the molecular mechanism of this phenomenon remains elusive [35,36], this bioassay has been used with a large number of BR analogs and the obtained results have allowed the establishment of structure–activity relationships [37,38]. Interestingly, the interpretation of the obtained data is commonly carried out in terms of the binding energy of BRs analogs to the active site of *A. thaliana*, which is calculated from molecular docking studies [28].

Herein, the difference of bending angles induced by analogs **12**, **14** and **17**–**22** have been measured at different concentrations (Table S1, Supporting Information File 1). The activities, calculated as the difference of bending angles obtained for the analogs and the negative control, and normalized respect to the latter, are shown in Table 1.

Table 1: RLIT activities of BR analogs of 2α,3α-dihydroxy-5α-cholan-6-oxo-23,24-dinor-22-(4-substituted)-benzoate-22-yl type, on seedlings after 72 h of incubation.

	RLIT activities ± (standard error)[a]		
Compounds	1×10^{-8} M	1×10^{-7} M	1×10^{-6} M
1	4.58 ± 1.33	6.25 ± 1.92	7.83 ± 2.03
12	4.50 ± 1.67^{ns}	0.92 ± 0.40***	0.58 ± 0.28***
17	6.83 ± 1.78*	4.08 ± 1.13*	2.58 ± 0.73**
18	2.17 ± 0.65**	6.33 ± 1.49^{ns}	7.00 ± 1.71^{ns}
19	4.58 ± 1.43^{ns}	4.17 ± 1.02*	3.67 ± 1.28**
14	3.25 ± 0.89*	3.33 ± 0.88**	1.83 ± 0.79***
20	2.08 ± 0.92**	2.75 ± 0.82**	7.08 ± 2.10^{ns}
21	1.50 ± 0.68***	5.08 ± 1.83^{ns}	3.83 ± 1.54**
22	2.00 ± 0.81**	2.75 ± 1.01**	5.83 ± 1.63*

[a]Values are the mean of leaf–sheath bending angle, normalized to the average of negative control (12 ± 2.4). Data are expressed as mean ± standard deviation of two independent experiments, each with at least six replicates (*n* = 12). Statistical significance was determined using the Student's t-test comparing each treatment to the positive control (**1**). ns = not significant; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Brassinolide (**1**) was used as positive control.

The data show that the plant response to exogenous application of brassinolide (**1**) (positive control) or synthetic BR analogs strongly depends on the chemical structure of the BRs. For example, the activity increases with increasing concentration of **1**, and analogs **18**, **20** and **22**, whereas the opposite effect is obtained by applying analogs **12** and **17**. On the other hand, the bioactivity of synthetic analogs **19**, **14** and **21** does not vary with increasing concentration. It is well established that lamina inclination is promoted by brassinolide and that other phytohormones participate in the regulation of this process as well [35,39]. Thus, the concentration dependence of lamina inclination activity could be attributed to multiple factors such as inhibition or activation of factors affecting BRs activity [40]. This effect becomes more important at highest concentrations and therefore, the relationship between activity and chemical structures should be made at the lowest tested concentration [28,41].

Comparison of activities at 1×10^{-8} M show that the activity of analogs with benzoate function at C-22 change with the nature of the substituent in the *para* position. The unsubstituted derivative **12** and compound **19**, substituted with a Cl atom, are as active as the positive control **1**, whereas compound **17**, with a *p*-methyl group, is more active than **1**. On the other hand, analogs **18**, **20**, **21** and **22**, which are substituted with OCH_3, Br, I and CN groups, are half as active as **12**. These results indicate that small substituents with no electronegative effects on the aromatic ring enhance the analogs' bioactivity in the RLIT.

A comparison of results obtained for compounds with the same substituent at C-22, but having just one hydroxy group at C-3 (**3**) (TE analogs) [28], or only a carbonyl group in this position (**3a**) (3-DT analogs) [31] is given in Table 2.

Table 2: RLIT relative activity at 1×10^{-8} M, using brassinolide (**1**) as reference compound, for different series of BRs analogs with the same substituent at C-22 and structural differences in ring A, i.e., castasterone (CAT), teasterone (TE) and 3-dehydroxyteasterone (3-DT).

Analogue	CAT	TE	3-DT
12 (*p*-H)	0.98	0.95	0.72
17 (*p*-CH_3)	1.49	1.05	0.80
18 (*p*-OCH_3)	0.47	1.50	0.36
19 (*p*-Cl)	1.00	0.55	0.45
14 (*p*-F)	0.71	0.73	0.77
20 (*p*-Br)	0.45	0.68	–
21 (*p*-I)	0.33	1.73	0.77

The data indicate that derivatives with a OH group at C-3 (TE analogs) are much more active than derivatives with a carbonyl group in this position (3-DT analogs). However, the addition of another OH group at C-2 (CAT analogs) increases the activity of analogs substituted with *p*-methyl (**17**) or *p*-Cl (**19**). The opposite effect is observed for those analogs substituted with *p*-OCH_3 (**18**) or *p*-I (**21**).

Inhibition of root growth in *Arabidopsis thaliana* seedlings assay

The root of *A. thaliana* has been used as a model to study cell division and elongation processes [42]. It has been shown that BRs affect root growth of *Oryza sativa* and *A. thaliana* in a dose-dependent manner. At low concentrations a slight root elongation is observed, whereas at higher concentrations a marked inhibition of growth is obtained [43,44]. These observations are consistent with reports showing that BRs modulate cell division and elongation processes in a concentration-depending manner [45]. The inhibitory effects on root growth of *A. thaliana* seedlings (Columbia ecotype, Col-0), caused by exogenous application of brassinolide (**1**) and 2α,3α-dihydroxylated analogs, **14**, **17**–**20** are shown in Figure 3.

The root lengths measured for *A. thaliana* grown in presence of **1**, **12**, **16a**, analogs **17**–**22** at 1×10^{-6} M are summarized in Table 3.

The activities displayed by the unsubstituted analog **12**, and that with a Cl atom at the *para*-position, **19**, are comparable to that shown by brassinolide (**1**). Interestingly, those analogs containing F and OCH_3 at the *para*-position, **14** and **18**, respectively, exhibit activities that are higher than that measured for **1** at the tested concentration (1×10^{-6} M). On the other hand, the compound with no benzoyl function in the side chain, **16a**, and analogs with I and CN substituents, **21** and **22**, respectively, show the lowest activities at 1×10^{-6} M. These results are quite similar to those previously reported for compounds **4**–**11**, in which ring A has one hydroxy group at C-3 [28]. However, analogs with I and CN substituents in that series, **10** and **11**, are much more active than analogs **21** and **22**.

In summary, for this type of dihydroxylated analogs the biological root inhibitory activity follows the order: **14** > **18** > **17** > **20** > **1** > **19** > **12** > **21** = **22**. The highest activity observed for both mono- and dihydroxylated analogs is compound **14** having a small and very electronegative substituent, i.e., F atom at the *para*-position.

Western blotting analyses of compounds 2α,3α-dihydroxy-5α-cholan-6-oxo-23,24-dinor-22-(4-substituted)-benzoate-type treatment on BES1 protein

Western blot is a very useful, sensitive, specific and efficient technique for the detection and identification of specific proteins from a complex mixture of proteins. In addition, it is

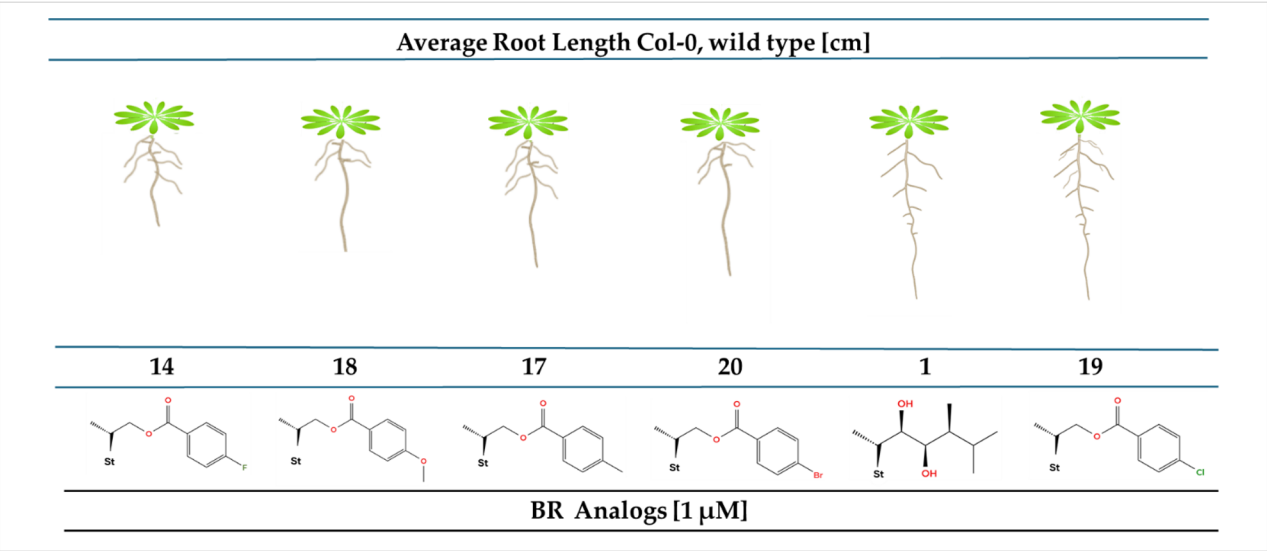


Figure 3: Effect of brassinolide (**1**), and BR analogs **14**, **17–20**) and analog **1** on root system inhibition of *A. thaliana* Col-0 wild type. Seedlings were grown on medium containing **1** and analogs (1×10^{-6} M) for 6 days. DMSO was used as negative control. For each treatment, more than 20 seedlings were analyzed. St = steroidal skeleton of the structure.

Table 3: Effect of brassinolide (**1**), and BR analogs with 23,24-bisnor-5α-cholane-type side chains (**12**, **14** and **17–22**) and analog **16a,** applied at 1×10^{-6} M, on root growth of *A. thaliana* Col-0 wild type[a].

Treatments	Root length [cm]	Treatments	Root length [cm]
control	0.93 ± 0.05†‡	**19**	0.98 ± 0.16†‡§
1	0.93 ± 0.12†‡§	**14**	0.49 ± 0.11*
16a	1.13 ± 0.22‡§	**20**	0.87 ± 0.08†
12	1.07 ± 0.23†‡§	**21**	1.16 ± 0.05§
17	0.86 ± 0.14†	**22**	1.16 ± 0.18§
18	0.67 ± 0.24*†		

[a]Five days old *A. thaliana* seedlings (Columbia ecotype, Col-0) treated with 1×10^{-6} M of **1** (positive control) and BR analogs, whereas 0.1% of DMSO was used as negative control. For each treatment, more than 20 seedlings were analyzed. Means followed by the same symbol (†, ‡, §, and *) do not differ significantly according to Duncan's multiple range test ($p \leq 0.05$)".

possible to quantify their expression by using suitable primary and secondary antibodies to visualize it.

It has been established that the dephosphorylated forms of BZR1 and BES1, and their concomitant accumulation in the nucleus, control the activity in the BRs signaling pathway [32,46], whereas their phosphorylated forms remain inactive and unstable and can be degraded by the proteasome under normal conditions [46]. On the other hand, accumulation of non-phosphorylated BES1 protein is specific to BRs, as other plant hormones such as auxin, cytokinin, abscisic acid and gibberellin do not cause BES1 accumulation [47], therefore this assay is accurate and complements the biological tests performed above.

Thus, in this study, the Western blot analysis was used to detect BES1 in presence of different exogenously applied BR analogs (**12**, **14** and **17–22**). The main idea is to correlate the amount and form of BES1 with the activity exhibited by these analogs on the root growth assay.

The results of the immunoblot analysis of proteins with an antibody specific for BES1, in roots of *A. thaliana* ecotype Col-0 plants are shown in Figure S67, Supporting Information File 1. The results are expressed as percentage of dephosphorylated BES1 (dBES1) for a negative control, brassinolide (used as positive control), and exogenously applied BR analogs. In Figure S68, Supporting Information File 1 values are represented of experimental percentages relative to both the negative control and positive control. From comparison of data with the negative control relative values of 1.33, 2.56 and 3.70, have been obtained for treatment with **1**, **12** and **19**, respectively, at 1×10^{-6} M . So, by considering the negative control it becomes clear that these compounds induce significantly high dBES1 overexpression. The following discussion has been made using the values relative to **1**, i.e., the expression values shown above the bold columns are relative to brassinolide, the positive control. The data show that the highest relative accumulation in dBES1 overexpression, 1.93 and 2.79, are observed for analogs **12** and **19**, respectively. The relative values follow the order **19** > **12** > **1** > **21** > **22** > **14** > **18** > **20** > **17**. These values suggest that these compounds should follow a similar trend in activities related to the BRs signaling process. However, this assumption is not validated by the results obtained in the root inhibition assay which gave a different order of activities (Figure 3, Table 2). In this assay the most active

analogs are **14** and **18**. However, even though RLIT was performed at much lower concentrations and in rice instead of *Arabidopsis*, the dBES1 overexpression data and the RLIT results obtained at 1×10^{-8} M follow a similar pattern. This result is probably a mere coincidence, but we believe it is worth to mention it.

Molecular docking study

Molecular docking is a computational tool that is widely used to evaluate the association process of BR analogs to the active site of *A. thaliana*. The results provide information on the most probable configuration, the main interactions, and the binding energy as a function of the chemical structure of BR analogs. The lower binding energy is commonly associated to the most active BR analog. Thus, docking results have allowed to explain experimental correlations between activity and chemical structure of BR analogs [34,48,49]. For example, previous molecular docking studies have predicted similar or better binding energies than brassinolide (**1**) for compounds with a phenyl ring with small groups such as fluorine, chlorine, or methyl [34]. In this work, molecular docking simulations were performed for all dihydroxylated analogs using the crystal structure of the BRI1–BAK1 complex (PDB ID: 4M7E) (Figure S69 and Table S2, Supporting Information File 1). This heterodimeric complex is commonly used because it represents the most structurally and functionally comprehensive model system resolved by X-ray crystallography. Unlike BRL1 and BRL3, for which only monomeric ectodomain structures are available, the 4M7E structure enables modeling of the full receptor–co-receptor assembly required for the "molecular glue" mechanism involving BAK1 [50,51]. Furthermore, BRL2 was also excluded from the analysis because it does not exhibit high-affinity binding to brassinolide and is therefore considered non-functional in ligand perception within this signaling context. The calculated binding energy values are presented in Table 4.

Table 4: Average binding energies calculated over at least 10 poses, for brassinolide (**1**) and synthetic analogs (**12**, **14**, **16a** and **17**–**22**).

Compounds	ΔE_b, kcal/mol	Compounds	ΔE_b, kcal/mol
1	−13.0	**19**	−14.0
16a	−11.6	**14**	−13.1
12	−13.2	**20**	−13.1
17	−12.9	**21**	−12.9
18	−12.9	**22**	−13.0

Interestingly, the estimated binding energy of compound **19**, −14.0 kcal/mol, is lower than that obtained for brassinolide (**1**) (−13.0 kcal/mol). As the predicted conformations are similar, this difference is attributed to six H-bond interactions: two of them involve the ArCO carbonyl group, which interacts with Ser647 (1.89 Å) and Tyr597 (2.4 Å), other two are established between the hydroxy group at C-2 and residues Tyr642 (2.7 Å) and Val62 (2.8 Å), and finally two extra H-bonds are formed between the hydroxy group at C-3 and residues His61 (2.1 Å) and Asn705 (2.4 Å). Additionally, the aromatic ring of compound **19** generates two π–π stacking interactions with Trp564 and Tyr597. Finally, there are also van der Waals interactions with residues Phe60, Phe681, Ile682, Ile706, Tyr642, Tyr599, Pro648, Ile540, Ile563, and Trp564 (Figure 4).

The calculated binding energies suggest the following order of activity **19** > **12** > **14** = **20** > **22** = **1** > **21** > **17** > **18** > **16a**.

Considering that binding to the active site is the initial step in BRs signaling process it is interesting to verify if this calculated pattern is in line with the experimental results obtained for dephosphorylation of BES1 and root inhibitory activity.

From a comparison of binding energy and relative accumulation in dBES1 overexpression it might be expected that there is

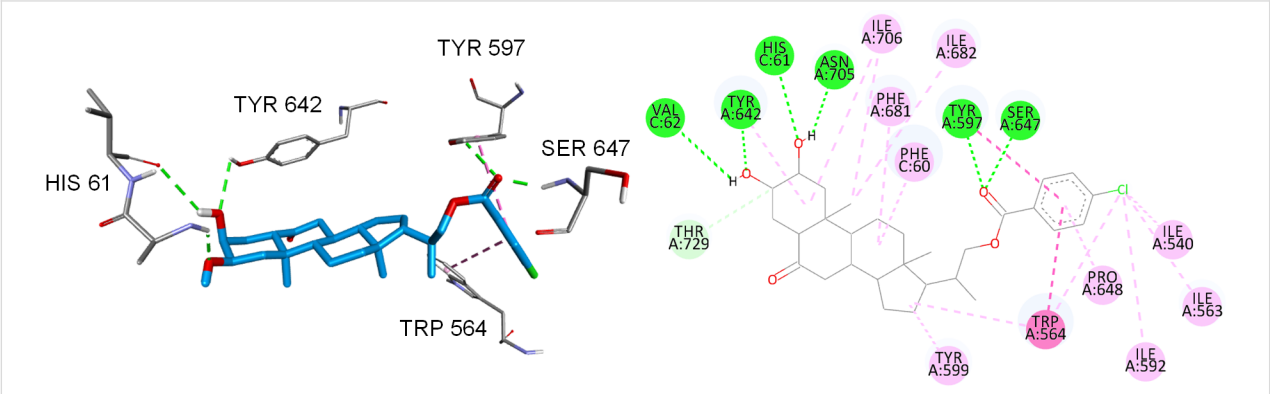


Figure 4: Binding modes of compound **19** into BRI1–BAK1 heterodimer. Hydrogen bonds are represented in green segmented lines. π–π stacking is represented by dark pink segmented lines. Hydrophobic interactions are represented by pink segmented lines.

an acceptable correlation between both parameters. For example, compound **19**, possessing a chlorine substituent on the phenyl ring, has an outstanding molecular coupling in BRI1, and at the same time shows a significant increase of BES1 dephosphorylation. However, an increase in dBES1 should induce a greater root inhibition and lower cell growth at high concentrations of exogenously applied BRs. However, in the presence of compound **19** the root inhibition is as high as that observed for brassinolide, but much lower than that obtained for compounds **14** and **18**. The latter is poorly docked into the binding site and its dBES1 is equally poor. But surprisingly, its root inhibition activity is one of the highest. Similar results are found for compounds **14** and **20**.

Consequently, the data suggest that the initial step of the signaling process and the subsequent dephosphorylation of BES1 depend on the structure of BRs analogs in the same way. In other words, the interaction between BRs analogs and the binding site determines the response of the plant and one intermediate step. However, it seems that on the signaling cascade this structural effect is lost, and the outcome do not follow the initial pattern.

It is known that plant cells can synthesize most hormones and, therefore their regulation is decentralized. In addition, each hormone modulates its own synthesis, degradation, conjugation and oxidation via dynamic interaction with different phytohormones and on homeostasis mechanisms [52]. Consequently, exogenous application of a growth regulator initiates not only the process leading to a phenotypic response but also all those associated to its own regulation. Thus, only a fraction of the whole exogenous applied BRs effectively reaches the receptors and once bound, the active form can result in a positive or negative feedback on its own biosynthesis, adding an additional level of regulatory complexity.

For these reasons, establishing a direct correlation between one of the possible phenotypic responses (inhibition of root growth) and the level of activation of BR-dependent transcription factors (such as dBES1) is particularly difficult to achieve. In other words, the intensity of the signal does not necessarily reflect the amount applied nor does it translate linearly into the accumulation of the transcription factors responsible for activating or repressing the genes associated with growth and development. Other factors such as tissue type, plant development stage and environmental conditions might also influence the different steps of the signaling process.

Conclusion

In this study, a series of new BR analogs with C22 benzoate-function, **17**–**22**, has been synthesized. The synthesis was carried out in two steps with high yields. The bioactivity of these analogs has been evaluated by two bioassays, rice lamina inclination test, and inhibition of root growth in *A. thaliana* seedlings. In the former test the best biological effects were observed for the analogs with 2α,3α-dihydroxy function (castasterone derivatives) as compared to those obtained for analogs with 3β-hydroxy (teasterone derivatives) and 3-carbonyl function. This result confirms that the best activity results in RLIT are given by analogs having the 2α,3α-dihydroxy configuration in ring A. Additionally, western blot analyses were performed to determine the amount of dBES1 produced in *A. thaliana* seedlings by each exogenously added analog. The results indicate that the measured bioactivity depends on the analogs' structure and on the used bioassay. Interestingly, the plant response, western blot, and docking results obtained in *A. thaliana* do not follow the same structure–activity pattern. For example, analogs **14** and **18**, are more active than brassinolide (**1**) in the inhibition of root growth, whereas the highest relative accumulation in dBES1, are exhibited by analogs **12** and **19**. On the other hand, analogue **19** forms the most stable complex in the active site. From these results it can be concluded that the efficiency of BR analogs to initiate the signaling process is not a determining factor in the final plant response. Intermediate processes such as dephosphorylation of BES1 depend on the structure of BR analogs in the same way, but it becomes obvious that on following steps this relation is changed. Thus, additional research is needed to evaluate the BRs structure effect on additional subsequent steps of the signaling process.

Supporting Information

Supporting Information File 1

NMR spectra of compounds, biological bioassays and protein–ligand interactions (molecular docking).
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-57-S1.pdf]

Supporting Information File 2

Experimental section.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-57-S2.pdf]

Funding

Luis Espinoza-Catalán thanks FONDEQUIP (Fondo de Equipamiento Científico y Tecnológico, Grants EQM190025 and EQM240189). This research was funded by ANID, FONDECYT Regular (Fondo Nacional de Desarrollo Científico y Tecnológico), Grant N° 1231502 and N° 1230368.

Conflicts of Interest

There are no conflicts to declare.

Author Contributions

María Núñez: data curation; investigation; methodology; validation; visualization; writing – original draft; writing – review & editing. Camila Escobar: formal analysis; methodology. Mario Párraga: formal analysis; methodology; resources. Mauricio Soto: data curation; formal analysis; methodology; software; writing – original draft. Luis Espinoza-Catalán: conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; writing – original draft; writing – review & editing. Katy Díaz: conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; writing – original draft; writing – review & editing. Andrés F. Olea: conceptualization; formal analysis; investigation; writing – original draft; writing – review & editing.

ORCID® iDs

María Núñez - https://orcid.org/0009-0002-0771-5351
Camila Escobar - https://orcid.org/0009-0008-5164-6393
Mario Párraga - https://orcid.org/0000-0002-3365-3070
Mauricio Soto - https://orcid.org/0009-0001-1864-7561
Luis Espinoza-Catalán - https://orcid.org/0000-0002-4054-7487
Katy Díaz - https://orcid.org/0000-0002-7345-1123

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Mitchell, J. W.; Mandava, N.; Worley, J. F.; Plimmer, J. R.; Smith, M. V. *Nature* **1970,** *225,* 1065–1066. doi:10.1038/2251065a0
2. Clouse, S. D. *J. Plant Growth Regul.* **2015,** *34,* 828–844. doi:10.1007/s00344-015-9540-7
3. Clouse, S. D.; Sasse, J. M. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998,** *49,* 427–451. doi:10.1146/annurev.arplant.49.1.427
4. Bajguz, A. Brassinosteroids – occurence and chemical structures in plants. In *Brassinosteroids: A Class of Plant Hormone;* Hayat, S.; Ahmad, A., Eds.; Springer: Dordrecht, Netherlands, 2011; pp 1–27. doi:10.1007/978-94-007-0189-2_1
5. Oklestkova, J.; Rárová, L.; Kvasnica, M.; Strnad, M. *Phytochem. Rev.* **2015,** *14,* 1053–1072. doi:10.1007/s11101-015-9446-9
6. Oh, M.-H.; Honey, S. H.; Tax, F. E. *Int. J. Mol. Sci.* **2020,** *21,* 1743. doi:10.3390/ijms21051743
7. Müssig, C. *Plant Biol.* **2005,** *7,* 110–117. doi:10.1055/s-2005-837493
8. Sasse, J. M. *J. Plant Growth Regul.* **2003,** *22,* 276–288. doi:10.1007/s00344-003-0062-3
9. Fujioka, S. Natural Occurrence of Brassinosteroids in the Plant Kingdom. In *Brassinosteroids: Steroidal Plant Hormones;* Sakurai, A.; Yokota, T.; Clouse, S. D., Eds.; Springer: Tokyo, Japan, 1999; pp 21–45.
10. Zullo, M. A. T.; Adam, G. *Braz. J. Plant Physiol.* **2002,** *14,* 143–181. doi:10.1590/s1677-04202002000300001
11. Suzuki, H.; Fujioka, S.; Takatsuto, S.; Yokota, T.; Murofushi, N.; Sakurai, A. *J. Plant Growth Regul.* **1994,** *13,* 21–26. doi:10.1007/bf00210703
12. Fujioka, S.; Inoue, T.; Takatsuto, S.; Yanagisawa, T.; Yokota, T.; Sakurai, A. *Biosci., Biotechnol., Biochem.* **1995,** *59,* 1543–1547. doi:10.1271/bbb.59.1543
13. Fujioka, S.; Inoue, T.; Takatsuto, S.; Yanagisawa, T.; Yokota, T.; Sakurai, A. *Biosci., Biotechnol., Biochem.* **1995,** *59,* 1973–1975. doi:10.1271/bbb.59.1973
14. Bajguz, A. *Plant Physiol. Biochem.* **2007,** *45,* 95–107. doi:10.1016/j.plaphy.2007.01.002
15. Clouse, S. D. *Plant Cell* **2011,** *23,* 1219–1230. doi:10.1105/tpc.111.084475
16. Sun, Y.; Fan, X.-Y.; Cao, D.-M.; Tang, W.; He, K.; Zhu, J.-Y.; He, J.-X.; Bai, M.-Y.; Zhu, S.; Oh, E.; Patil, S.; Kim, T.-W.; Ji, H.; Wong, W. H.; Rhee, S. Y.; Wang, Z.-Y. *Dev. Cell* **2010,** *19,* 765–777. doi:10.1016/j.devcel.2010.10.010
17. Guo, H.; Li, L.; Aluru, M.; Aluru, S.; Yin, Y. *Curr. Opin. Plant Biol.* **2013,** *16,* 545–553. doi:10.1016/j.pbi.2013.08.002
18. He, J.-X.; Gendron, J. M.; Sun, Y.; Gampala, S. S. L.; Gendron, N.; Sun, C. Q.; Wang, Z.-Y. *Science* **2005,** *307,* 1634–1638. doi:10.1126/science.1107580
19. Takatsuto, S. *J. Chromatogr. A* **1994,** *658,* 3–15. doi:10.1016/0021-9673(94)85202-2
20. Wada, K.; Marumo, S.; Abe, H.; Morishita, T.; Nakamura, K.; Uchiyama, M.; Mori, M. *Agric. Biol. Chem.* **1984,** *48,* 719–726. doi:10.1080/00021369.1984.10866208
21. Cao, H.; Chen, S. *Plant Growth Regul.* **1995,** *16,* 189–196. doi:10.1007/bf00029540
22. Kim, T.-W.; Lee, S. M.; Joo, S.-H.; Yun, H. S.; Lee, Y.; Kaufman, P. B.; Kirakosyan, A.; Kim, S.-H.; Nam, K. H.; Lee, J. S.; Chang, S. C.; Kim, S.-K. *Plant, Cell Environ.* **2007,** *30,* 679–689. doi:10.1111/j.1365-3040.2007.01659.x
23. Kinoshita, T.; Caño-Delgado, A.; Seto, H.; Hiranuma, S.; Fujioka, S.; Yoshida, S.; Chory, J. *Nature* **2005,** *433,* 167–171. doi:10.1038/nature03227
24. Thompson, M. J.; Meudt, W. J.; Mandava, N. B.; Dutky, S. R.; Lusby, W. R.; Spaulding, D. W. *Steroids* **1982,** *39,* 89–105. doi:10.1016/0039-128x(82)90129-5
25. Takatsuto, S.; Yazawa, N.; Ikekawa, N.; Takematsu, T.; Takeuchi, Y.; Koguchi, M. *Phytochemistry* **1983,** *22,* 2437–2441. doi:10.1016/0031-9422(83)80135-6
26. Liu, J.; Zhang, D.; Sun, X.; Ding, T.; Lei, B.; Zhang, C. *Steroids* **2017,** *124,* 1–17. doi:10.1016/j.steroids.2017.05.005
27. Back, T. G. Stereoselective Synthesis of Brassinosteroids. In *Stereoselective Synthesis (Part J);* Atta-ur-Rahman, Ed.; Studies in Natural Products Chemistry; Elsevier: Amsterdam, Netherlands, 1995; pp 321–364. doi:10.1016/s1572-5995(06)80056-6
28. Nuñez, M.; Wang, Y.; Russinova, E.; Estévez-Braun, A.; Amesty, A.; Olea, A. F.; Mellado, M.; Díaz, K.; Espinoza Catalán, L. *Int. J. Mol. Sci.* **2024,** *25,* 10158. doi:10.3390/ijms251810158
29. Aitken, V.; Diaz, K.; Soto, M.; Olea, A. F.; Cuellar, M. A.; Nuñez, M.; Espinoza-Catalán, L. *Int. J. Mol. Sci.* **2024,** *25,* 419. doi:10.3390/ijms25010419
30. Núñez, G.; Llovera, L.; Arrieche, D.; Berrios, R.; Soto, M.; Osorio-Olivares, M.; Olea, A. F.; Sarmiento, E.; González, A.; Carrasco, H.; Taborga, L. *Plants* **2025,** *14,* 3277. doi:10.3390/plants14213277

31. Valdés, E.; Díaz, K.; Núñez, M.; Olea, A. F.; Moral, J. F. Q. d.; Carvajal, R.; Cuellar, M. A.; Espinoza-Catalán, L. *Int. J. Mol. Sci.* **2025,** *26,* 8710. doi:10.3390/ijms26178710
32. Vert, G.; Nemhauser, J. L.; Geldner, N.; Hong, F.; Chory, J. *Annu. Rev. Cell Dev. Biol.* **2005,** *21,* 177–201. doi:10.1146/annurev.cellbio.21.090704.151241
33. Korinkova, P.; Bazgier, V.; Oklestkova, J.; Rarova, L.; Strnad, M.; Kvasnica, M. *Steroids* **2017,** *127,* 46–55. doi:10.1016/j.steroids.2017.08.010
34. Kvasnica, M.; Oklestkova, J.; Bazgier, V.; Rárová, L.; Korinkova, P.; Mikulík, J.; Budesinsky, M.; Béres, T.; Berka, K.; Lu, Q.; Russinova, E.; Strnad, M. *Org. Biomol. Chem.* **2016,** *14,* 8691–8701. doi:10.1039/c6ob01479h
35. Zhang, F.; Fang, C.; Liang, W. *Agronomy (Basel, Switz.)* **2024,** *14,* 1562. doi:10.3390/agronomy14071562
36. Zhou, L.-J.; Xiao, L.-T.; Xue, H.-W. *Plant Physiol.* **2017,** *174,* 1728–1746. doi:10.1104/pp.17.00413
37. Back, T. G.; Pharis, R. P. *J. Plant Growth Regul.* **2003,** *22,* 350–361. doi:10.1007/s00344-003-0057-0
38. Baron, D. L.; Luo, W.; Janzen, L.; Pharis, R. P.; Back, T. G. *Phytochemistry* **1998,** *49,* 1849–1858. doi:10.1016/s0031-9422(98)00367-7
39. Zhao, B.; Li, J. *J. Integr. Plant Biol.* **2012,** *54,* 746–759. doi:10.1111/j.1744-7909.2012.01168.x
40. Tong, H.; Xiao, Y.; Liu, D.; Gao, S.; Liu, L.; Yin, Y.; Jin, Y.; Qian, Q.; Chu, C. *Plant Cell* **2014,** *26,* 4376–4393. doi:10.1105/tpc.114.132092
41. Jorquera, S.; Soto, M.; Díaz, K.; Nuñez, M.; Cuellar, M. A.; Olea, A. F.; Espinoza-Catalán, L. *Int. J. Mol. Sci.* **2024,** *25,* 7515. doi:10.3390/ijms25147515
42. Petricka, J. J.; Winter, C. M.; Benfey, P. N. *Annu. Rev. Plant Biol.* **2012,** *63,* 563–590. doi:10.1146/annurev-arplant-042811-105501
43. Clouse, S. D.; Langford, M.; McMorris, T. C. *Plant Physiol.* **1996,** *111,* 671–678. doi:10.1104/pp.111.3.671
44. González-García, M.-P.; Vilarrasa-Blasi, J.; Zhiponova, M.; Divol, F.; Mora-García, S.; Russinova, E.; Caño-Delgado, A. I. *Development* **2011,** *138,* 849–859. doi:10.1242/dev.057331
45. Chaiwanon, J.; Wang, Z.-Y. *Curr. Biol.* **2015,** *25,* 1031–1042. doi:10.1016/j.cub.2015.02.046
46. Ryu, H.; Kim, K.; Cho, H.; Hwang, I. *Mol. Cells* **2010,** *29,* 291–296. doi:10.1007/s10059-010-0034-y
47. Yin, Y.; Wang, Z.-Y.; Mora-Garcia, S.; Li, J.; Yoshida, S.; Asami, T.; Chory, J. *Cell* **2002,** *109,* 181–191. doi:10.1016/s0092-8674(02)00721-3
48. Díaz, K.; Espinoza, L.; Carvajal, R.; Conde-González, M.; Niebla, V.; Olea, A. F.; Coll, Y. *Int. J. Mol. Sci.* **2020,** *21,* 1832. doi:10.3390/ijms21051832
49. Lei, B.; Liu, J.; Yao, X. *Steroids* **2015,** *104,* 111–117. doi:10.1016/j.steroids.2015.09.002
50. Caño-Delgado, A.; Yin, Y.; Yu, C.; Vafeados, D.; Mora-Garcia, S.; Cheng, J.-C.; Nam, K. H.; Li, J.; Chory, J. *Development* **2004,** *131,* 5341–5351. doi:10.1242/dev.01403
51. Fàbregas, N.; Lozano-Elena, F.; Blasco-Escámez, D.; Tohge, T.; Martínez-Andújar, C.; Albacete, A.; Osorio, S.; Bustamante, M.; Riechmann, J. L.; Nomura, T.; Yokota, T.; Conesa, A.; Alfocea, F. P.; Fernie, A. R.; Caño-Delgado, A. I. *Nat. Commun.* **2018,** *9,* 4680. doi:10.1038/s41467-018-06861-3
52. Hagihara, S.; Yamada, R.; Itami, K.; Torii, K. U. *Curr. Opin. Plant Biol.* **2019,** *47,* 32–37. doi:10.1016/j.pbi.2018.09.002

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.57

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Design, synthesis, and biological evaluation of FXR/ASK1 dual-target modulators

Xi Zhang*,‡, Jingyan Wang‡, Ziqiang Zhao, Caiyi Wang, Zenghui Ye, Wei-Yuan Ma, Jian-Xing Xu and Fengzhi Zhang*

Full Research Paper

Open Access

Address:
School of Pharmacy, Hangzhou Medical College, Hangzhou, 310014, P. R. China

Email:
Xi Zhang* - zhangxi@hmc.edu.cn; Fengzhi Zhang* - zhangfengzhi@hmc.edu.cn

* Corresponding author ‡ Equal contributors

Keywords:
ASK1 inhibitors; FXR agonists; FXR/ASK1 dual-target modulators; MASH

Beilstein J. Org. Chem. **2026,** *22,* 771–781.
https://doi.org/10.3762/bjoc.22.59

Received: 12 January 2026
Accepted: 07 May 2026
Published: 20 May 2026

Associate Editor: N. Sewald

© 2026 Zhang et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

The dual modulation of FXR and ASK1 is considered a promising therapeutic strategy for metabolic dysfunction-associated fatty liver disease (MAFLD) and its progressive form, metabolic dysfunction-associated steatohepatitis (MASH). GW4064 and GS-4997 are effective regulators for FXR and ASK1, respectively. Through the effective functional group splicing strategy, a new dual-target modulator is designed. Compound **Z8**, which acts on both targets, was found to more potently reduce intracellular lipid droplet accumulation in OA-treated HepG2 cells than the FXR agonist GW4064 and the ASK1 inhibitor selonsertib (GS-4997).

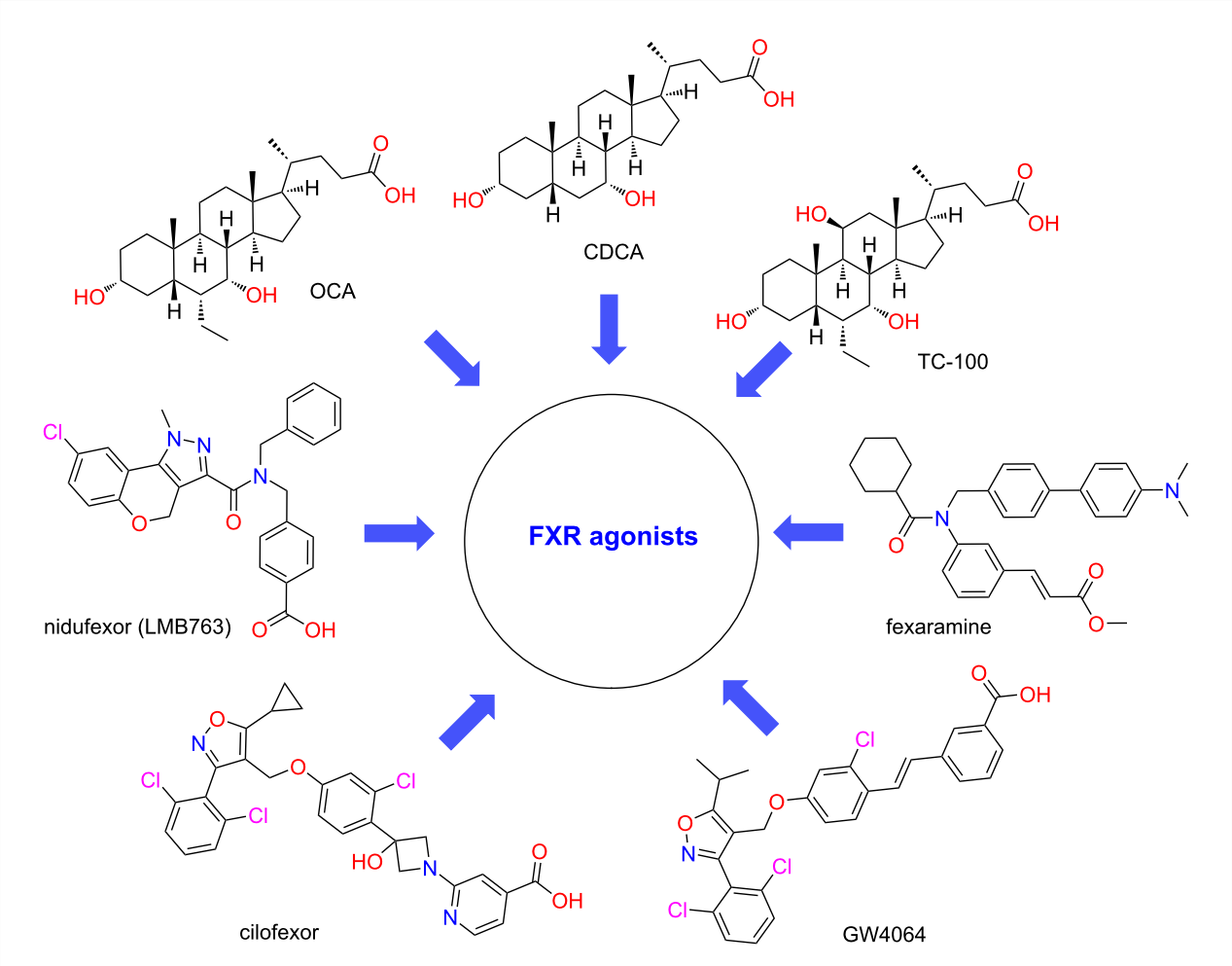
Introduction

The pathogenesis of metabolic dysfunction-associated fatty liver disease (MAFLD) is characterized by multifactorial interactions between environmental, genetic, extrahepatic, and intrahepatic factors. These factors can act in a parallel or sequential manner to drive the development of steatosis, inflammation, and fibrosis [1]. MAFLD is a clinical entity that encompasses a spectrum of disease, with metabolic dysfunction-associated steatohepatitis (MASH) denoting a more advanced form [2]. The condition is characterized by steatosis, hepatocyte injury, and inflammation [3], which can progress to fibrosis, cirrhosis and hepatocellular carcinoma (HCC) [4]. Although isolated hepatic steatosis often follows a benign and non-progressive course, untreated MASH carries a significant risk of severe sequelae [5]. These conditions constitute a substantial and growing burden on healthcare systems, yet treatment options remain limited. Presently, resmetirom, a THR-β agonist, stands as the sole pharmacotherapy to have received approval, thus underscoring the critical and as yet unmet medical need in this domain [6].

The farnesoid X receptor (FXR), a bile acid-activated nuclear receptor highly expressed in the liver and intestine, is a key regulator of genes involved in cholesterol and bile acid homeostasis, hepatic gluconeogenesis, lipogenesis, inflammation, and

fibrosis. It has been demonstrated that the substance under discussion also helps to maintain intestinal barrier integrity, to prevent bacterial translocation, and to support a balanced gut microbiota [7]. In view of this pleiotropic role, FXR activation has emerged as a well-established pharmacological target for MASH [8]. Consequently, a diverse range of FXR agonists – categorized as bile acid derivatives, non-bile-acid steroidal agonists, non-steroidal agonists, and partial agonists – are in advanced clinical development [9,10]. CDCA is the most pharmacologically active FXR agonist in natural bile acids. OCA and TC-100 were derived by a series of derivations using CDCA as a lead compound. Among the non-steroidal agonists, GW 4064, a classical isoxazole FXR agonist, served as the lead compound for a series of derivatives, such as cilofexor (GS-9674). Unfortunately, these agonists have failed in clinical trials because of side effects such as itching or a failure to meet primary endpoints. Despite the failure of several FXR agonists to achieve the desired outcomes, recent research has shifted towards investigating the potential of dual FXR agonists. A number of dual-targeting drugs, which engage FXR in conjunction with other pathways, have demonstrated considerable potential in enhancing the efficacy of MASH treatment outcomes [11,12] (Scheme 1).

Apoptosis signal-regulating kinase 1 (ASK1), a member of the MAP3K family, is known to be activated by various cellular stressors. These can be induced by diverse factors, spanning from reactive oxygen species (ROS) and endoplasmic reticulum stress to calcium influx and inflammatory agents (e.g., TNF and LPS) [13,14]. One such factor is the activation of ASK1, a complex process that involves homodimerization and autophosphorylation. Two major aspects characterize the role of ASK1 in liver pathophysiology. First, its phosphorylation initiates a signaling relay that activates both JNK and p38 MAPK pathways, resulting in apoptotic cell death, production of proinflammatory cytokines, and fibrogenic gene activation. Second, clinical data reveal that this same signaling node is overactive in the livers of individuals with obesity or MASH, where its func-

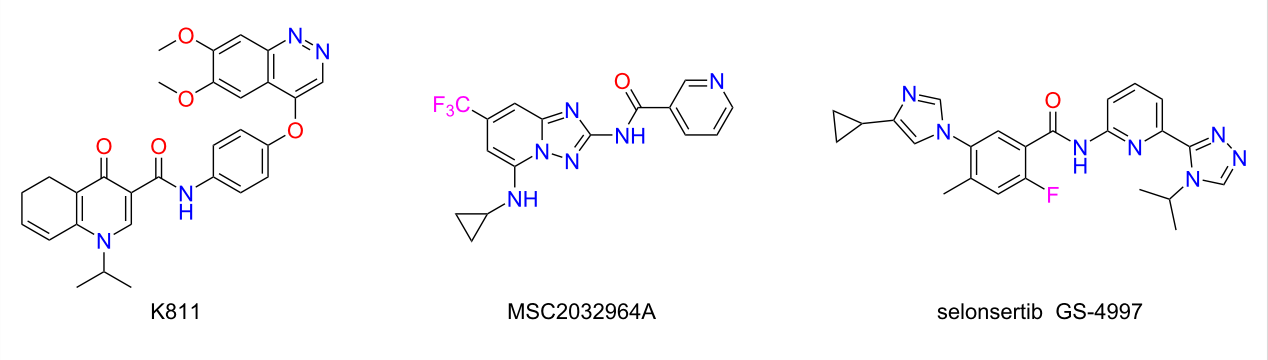


Scheme 1: Steroidal and nonsteroidal FXR agonists.

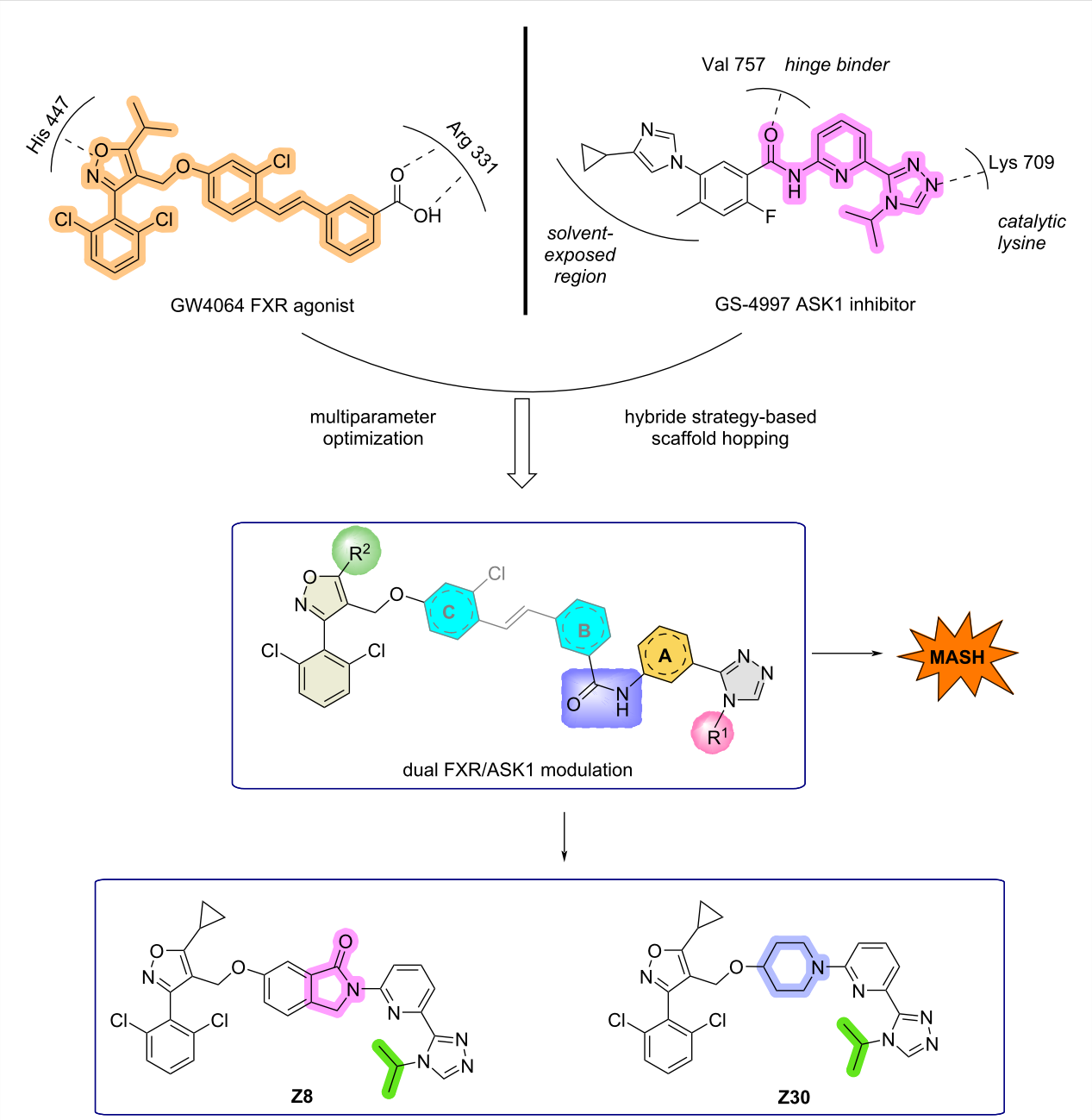
tional output correlates directly with the severity of insulin resistance, inflammation, and hepatic steatosis [15]. The most advanced ASK1 inhibitor in clinical development is the compound celonsertib (Scheme 2). Following this, a series of new ASK1 inhibitors were derived using selonsertib as a lead compound in a variety of structure-optimized ways. Given its central role, there are several dual-target drugs being developed that inhibit ASK1 alongside other pathways as antifibrotic agents for MASH treatment [16,17].

ASK1 activation has been demonstrated to drive key pathological processes in the liver, including hepatocyte dysfunction, apoptosis, inflammation, and myofibroblast activation [18-20]. Conversely, intestinal FXR agonism has been demonstrated to promote the release of fibroblast growth factor 19 (FGF19) [21,22], which in turn has been shown to downregulate hepatic bile acid synthesis, lipogenesis, and gluconeogenesis [23-25]. A previous study demonstrated that in a liver fibrosis model and a chronic fast food diet (FFD)-induced MASH mouse mode [26,27], the combination of an ASK1 inhibitor and an FXR agonist led to a more pronounced reduction in hepatic fibrosis than either monotherapy [28].

Recently, several publications reported many dual-target drug molecules based on FXR and other relevant therapeutic targets, such as FXR/sEH (soluble epoxide hydrolase), FXR/PPARδ (peroxisome proliferators activated receptor δ), FXR/LTA4H (leukotriene A4 hydrolase), and FXR/FABP 1 (fatty acid binding protein 1) [12,29-31]. In this study, we present the rational design and characterization of a balanced dual modulator targeting the nuclear receptors FXR and ASK1 [32], with the objective of advancing this experimental therapeutic concept (Scheme 3) [33,34]. The conception of this class of dual ligands was achieved by the merging of the minimal pharmacophores of the selective FXR agonist GW4064 and the selective ASK1 inhibitor selonsertib (GS-4997) [35-37]. Published structure–activity relationship (SAR) studies have identified the isoxazole moiety and terminal carboxylic acid of GW4064 [38], along with the 4-aliphatically substituted 4*H*-1,2,4-triazolopyridine biaryl scaffold of selonsertib [31,39,40], as key pharmacophores. In the co-crystal structure of hFXR-LBD with GW4064 [41], the isoxazole core coordinates the π-cation interaction between His 447 and Trp 469; the 5-isopropyl group is embedded into a hydrophobic pocket formed by Phe 284, Leu 287, Trp 454, and Phe 461; the dichlorophenyl moiety engages in π–π stacking with Phe 329; and the *meta*-positioned carboxy group forms a strong electrostatic interaction with Arg 331. The modification strategies for GW4064 are as follows: The stilbene moiety can be replaced through conformational restriction to reduce potential toxicity. The hydrophobic spacer element offers some flexibility in shape and size, but the linker length must remain conserved. Furthermore, to address the issue of poor aqueous solubility, heteroatoms can be introduced into the aromatic ring systems or the phenyl ring at the 3-position of the isoxazole to enhance hydrophilicity and improve pharmacokinetic properties [42]. Co-crystallization of GS-4997 with ASK1 unveiled key stabilizing interactions in the active site. Specifically, the amide carbonyl forms a hydrogen bond with the backbone amide of Val 757, and concurrently, a triazole nitrogen engages in a hydrogen bond with the amine moiety of Lys 709 [43]. By merging the key pharmacophores of the two drugs and systematically optimizing the resulting bipharmacophoric structure, we ultimately obtained balanced dual-target modulators, including the highly active component **Z8**. Despite a significant decrease in solubility due to its relatively high molecular weight, the compound's overall performance remains excellent. It was demonstrated that, in the absence of receptor overexpression, **Z8** effectively activated FXR and inhibited ASK1 signaling. Furthermore, treatment of hepatic cell lines with **Z8** validated its dual-target engagement in native cellular environments. In view of this favorable profile, **Z8** emerges as a highly promising lead compound for further investigation of dual FXR/ASK1 modulation as a polypharmacological strategy for MASH.



Scheme 2: ASK1 inhibitors.



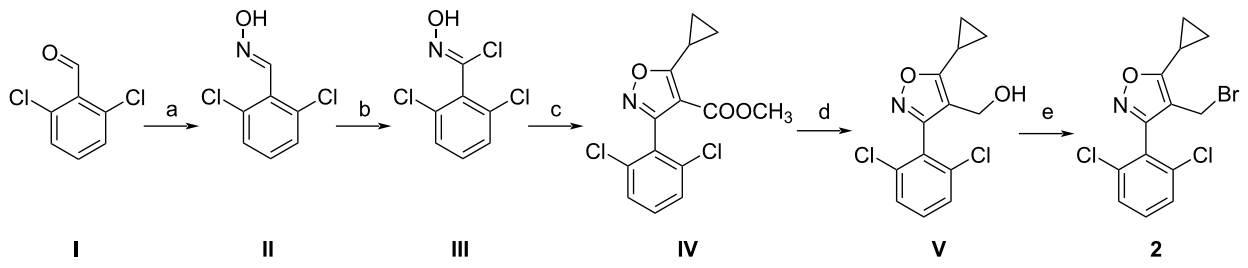
Scheme 3: Dual FXR/ASK1 modulation strategy for MASH.

Results and Discussion

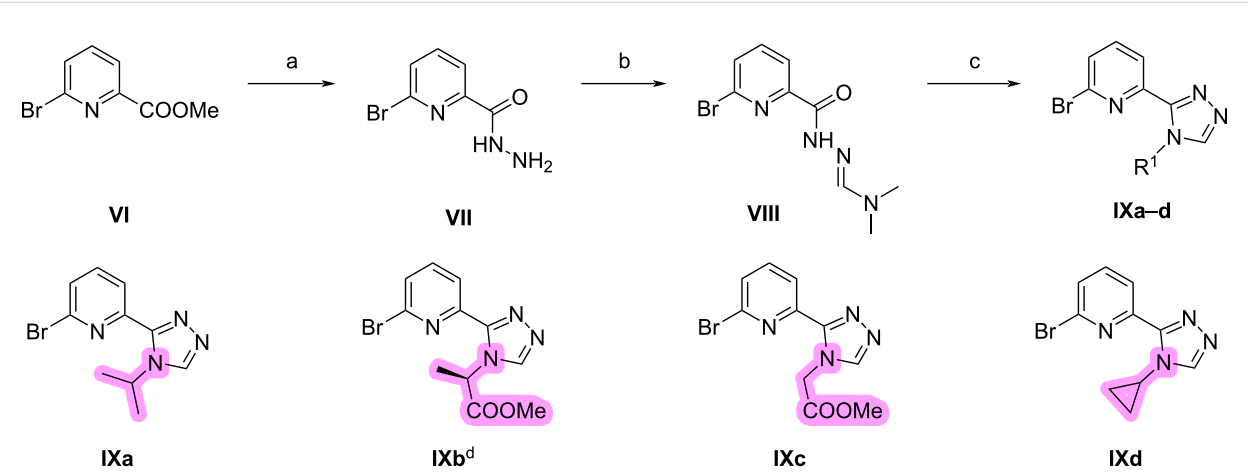
Chemistry

The synthetic routes to the target compounds are outlined in Schemes 4–8. The synthesis of isoxazole intermediate **2** commenced with the condensation of commercially available benzaldehyde **I** with hydroxylamine to afford oxime intermediate **II**. Subsequent treatment of **II** with NCS, followed by cyclization with methyl 3-cyclopropyl-3-oxopropionate, furnished isoxazole **IV**. Reduction of **IV** with lithium aluminum hydride gave alcohol **V**, which was then brominated with phosphorus tribromide to yield the desired isoxazole intermediate **2** in five steps (Scheme 4) [31].

The preparation of triazole intermediates **IXa–d** is shown in Scheme 5. Firstly, compound **VI** was reacted with hydrazine hydrate in methanol to obtain hydrazide **VII** which was mixed with DMF/DMA and heated to provide intermediate **VIII**. The obtained **VIII** was not purified and directly reacted with different amines to afford triazoles **IXa**,**c**,**d** or, in case of product **IXb** with (*R*)-alanine methyl ester. The absolute stereochemistry of



Scheme 4: Synthesis of compound **2**. Conditions: (a) $NH_2OH·HCl$, NaOH, $EtOH/H_2O$, 70 °C, 12 h; (b) NCS, DMF, 40 °C, 2 h; (c) methyl 3-cyclopropyl-3-oxopropanoate, triethylamine, EtOH, rt, 12 h, 64% yield; (d) $LiAlH_4$, THF, 0 °C, 2 h, 68% yield; (e) PBr_3, DCM, 0 °C, 12 h, 79% yield.



Scheme 5: Synthesis of compounds **IXa–d**. Conditions: (a) $N_2H_4·H_2O$, MeOH, rt, 12 h; (b) DMF–DMA, 80 °C, 12 h; (c) amines (for **IXa**,**c**,**d**) or (*R*)-alanine methyl ester (for **IXb**), AcOH, MeCN, 90 °C, 16 h, 66–80% yield; (d) absolute stereochemistry is *R*.

IXb is *R*. All compounds with defined absolute configuration described below were synthesized from **IXb**, thereby preserving the absolute stereochemistry and affording the corresponding *R* enantiomer.

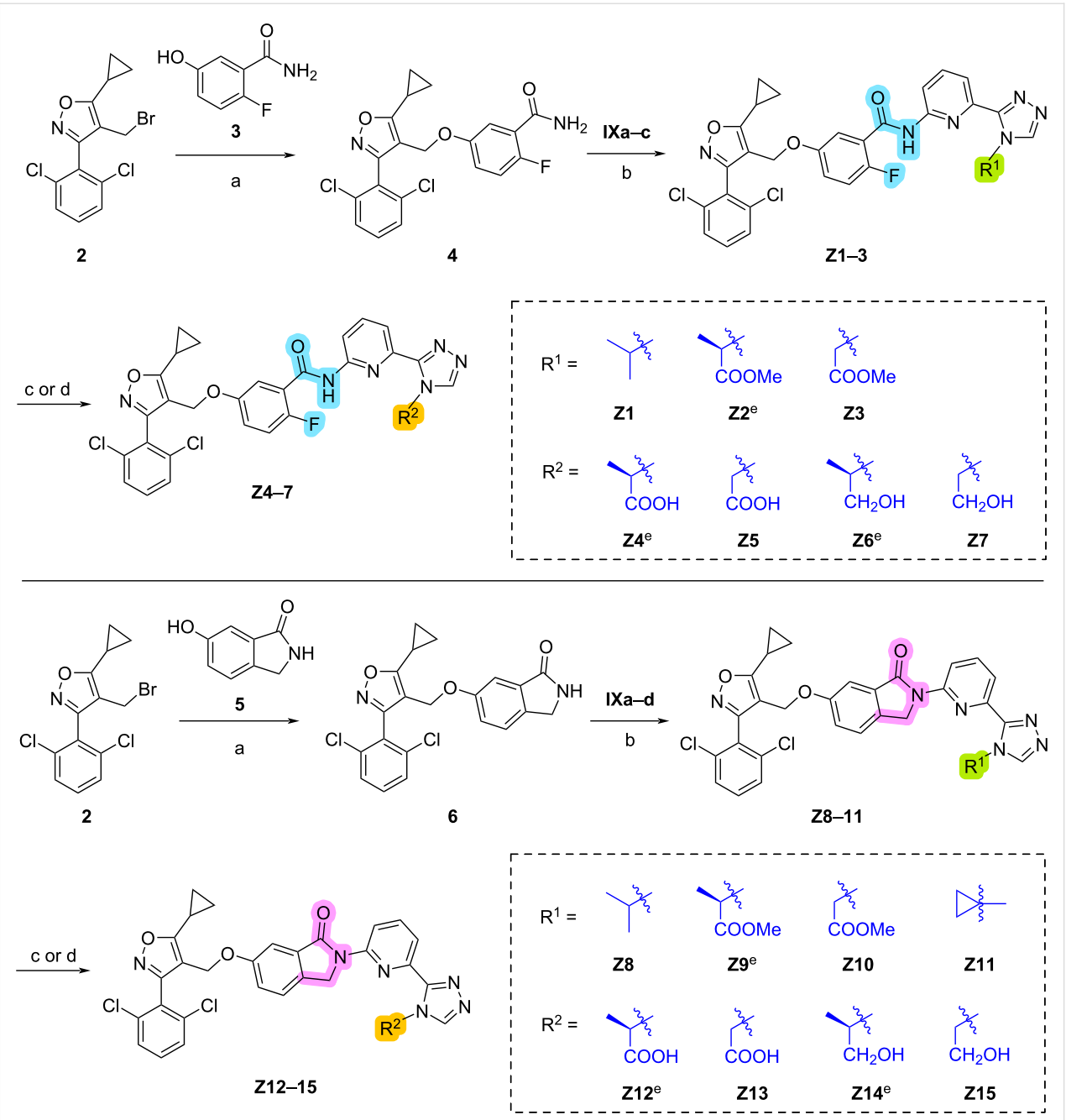
The synthetic routes of compounds **Z1**–**15** are shown in Scheme 6. First, intermediates **2** and **3** were reacted under basic conditions in MeCN to obtain intermediate **4** which was subsequently reacted with bromides **IXa–c** to get compounds **Z1–3** in 80–91% yield. Then, compounds **Z1**–**3** were treated with with LiOH or lithium aluminum hydride to obtain compounds **Z4**–**7**. The synthesis of compounds **Z8**–**15** followed a similar preparation procedure as for **Z1**–**7**.

The synthesis of target compounds **Z16**–**29** is illustrated in Scheme 7. Intermediates **8** and **10** were synthesized by the Suzuki cross-coupling reaction using dioxaborolane **7** or boronic acid **9** and **IXa**–**c**. Next, **Z16**–**18** and **Z23**–**25** were obtained by Williamson ether synthesis from the intermediates **8**, **10** and isoxazole intermediates **2**. Finally, hydrolysis in the presence of LiOH and reduction in the presence of lithium aluminum hydride provided target compounds **Z19**–**22** and **Z26**–**29**.

The synthesis of target compound **Z30** is illustrated in Scheme 8. Treatment of intermediate **2** with *N*-Boc-4-hydroxypiperidine (**11**) in the presence of sodium hydride produced ether **12**, followed by removal of the Boc-protecting group, provided intermediate **13**. Coupling of intermediate **13** with **IXa** under the catalytic conditions of $Pd_2(dba)_3$ provided compound **Z30**.

Molecular simulation

To elucidate the binding modes of compound **Z8** with FXR and ASK1, molecular docking simulations were performed as shown in Scheme 9. The results revealed that within the FXR binding pocket, the isoxazole and triazole moieties of **Z8** formed hydrogen bonds with the key residues ARG-331 and HIS-447, respectively. Similarly, in the ASK1 binding pocket, the isoxazole, carbonyl, and triazole groups of **Z8** engaged in hydrogen bonding with residues GLY-759, VAL-757, and LYS-709, respectively. Notably, the binding sites of **Z8** over-



Scheme 6: Synthesis of compounds **Z1**–**15**. Conditions: (a) K_2CO_3, KI, MeCN, 50 °C, 6 h, 86–96% yield; (b) $Pd_2(dba)_3$, Xantphos, Cs_2CO_3, 1,4-dioxane, 80 °C, 12 h, 80–91% yield; (c) LiOH·H_2O, THF/MeOH/H_2O, rt, 4 h, 55–76% yield; (d) $LiAlH_4$, THF, 0 °C, 4 h, 43–89% yield; (e) absolute stereochemistry is *R*.

lapped with those of the reference compounds, GW 4064 and selonsertib, indicating that **Z8** likely engages with FXR and ASK1 through a similar binding mode [45-47].

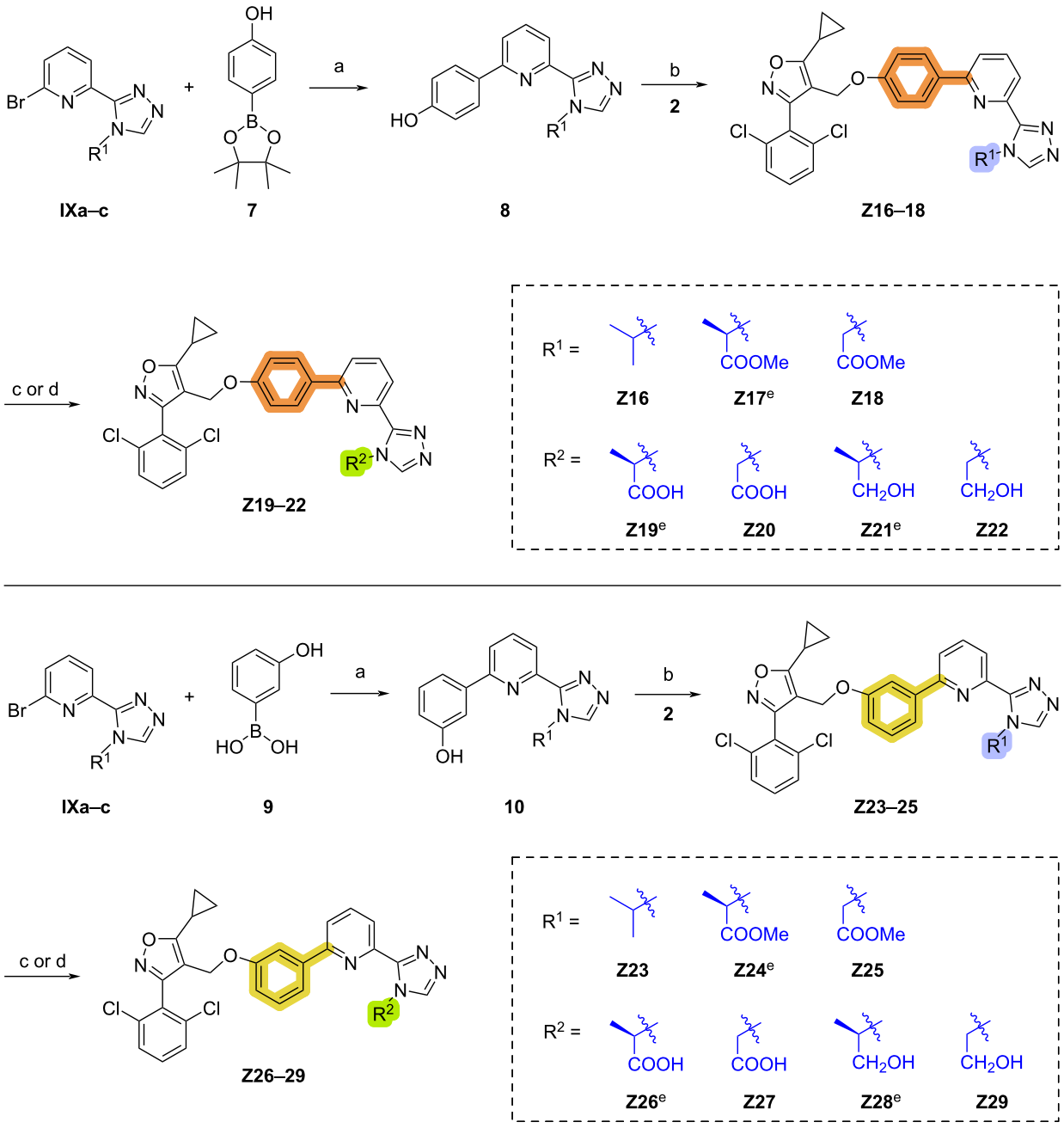
Biological evaluations

FXR agonist activity

The evaluation of the FXR agonist activity was conducted through the implementation of a dual-luciferase reporter gene assay. As demonstrated in Figure 1, among the designed dual-target modulators, compounds **Z8**, **Z11**, **Z17**, and **Z30** exhibited the most potent effects at a concentration of 10 μM.

ASK1 Inhibition

The inhibitory activity against ASK1 was evaluated using the ADP-Glo™ kinase assay. A screen of the designed dual-target modulators identified 17 compounds that inhibited ASK1

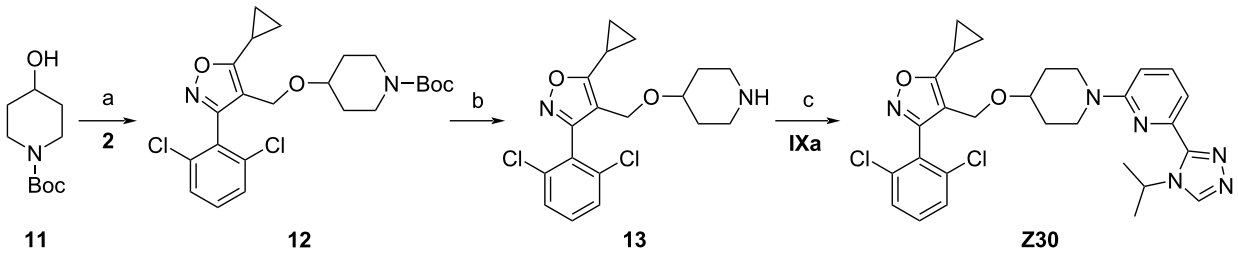


Scheme 7: Synthesis of compounds **Z16–29**. Conditions: (a) $Pd(PPh_3)_4$, Na_2CO_3, 1,4-dioxane/H_2O, 80 °C, 18 h, 64–90% yield; (b) K_2CO_3, KI, MeCN, 50 °C, 6 h, 53–95% yield; (c) LiOH·H_2O, THF/MeOH/H_2O, rt, 4 h, 55–76% yield; (d) $LiAlH_4$, THF, 0 °C, 4 h, 43–89% yield; (e) absolute configuration is *R*.

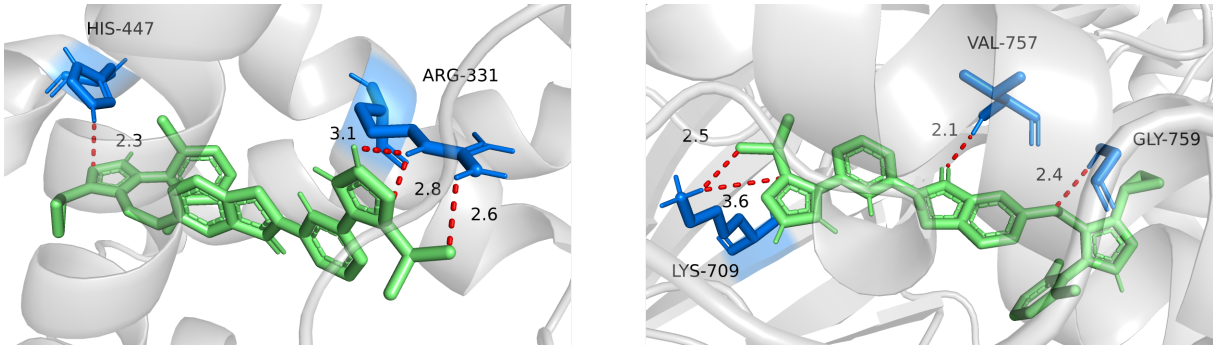
(Figure 2). Among the compounds examined, **Z8** demonstrated the greatest potency, exhibiting an inhibition rate of 84.59% at a concentration of 10 μM.

Lipid-modifying activity

In light of the prevalence of the reduction of lipid levels as a therapeutic objective in MASH, the present study sought to evaluate the lipid-modifying potential of modulators **Z8** and **Z30**. Utilizing an in vitro model of oleic acid (OA)-induced steatosis in HepG2 cells, the study evaluated the effects of these compounds on hepatic lipid accumulation. The presence of lipid droplets was confirmed through direct visualization using Oil Red O staining (Figure 3A). The combined treatment with GW4064 and GS4997 exerted a stronger lipolytic effect than



Scheme 8: Synthesis of compound **Z30**. Conditions: (a) NaH, DMF, rt, 16 h; (b) TFA, DCM, 0 °C, 3 h; (c) $Pd_2(dba)_3$, Xantphos, Cs_2CO_3, 1,4-dioxane, 80 °C, 18 h, 21% yield.



Scheme 9: Molecular docking of dual-target modulator **Z8** to the ligand binding sites of FXR (PDB ID: 3DCT, https://doi.org/10.2210/pdb3DCT/pdb) [41] and ASK1 (PDB ID: 6OYT, https://doi.org/10.2210/pdb6OYT/pdb) [44].

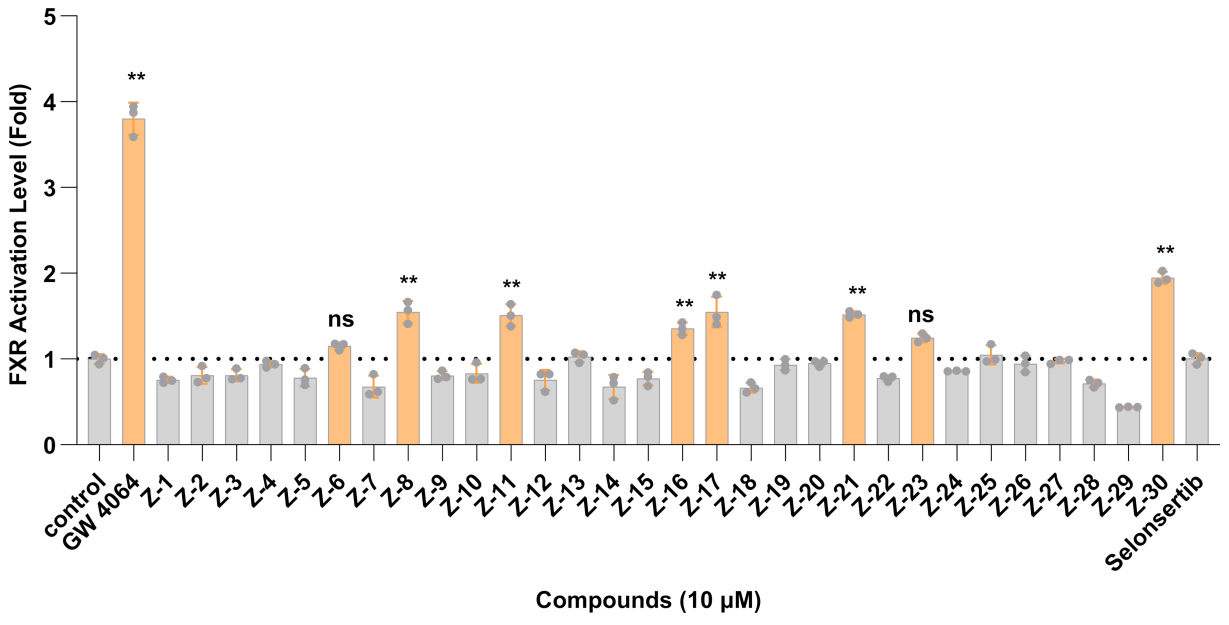


Figure 1: FXR agonist activity of all compounds. Results of the dual-luciferase reporter assay in CHO cells cultured for 24 hours following addition of GW4064 and test compounds. Compared to the control group, $^{**}p < 0.01$ and ns indicate no significant difference, with control; DMSO serving as the control group.

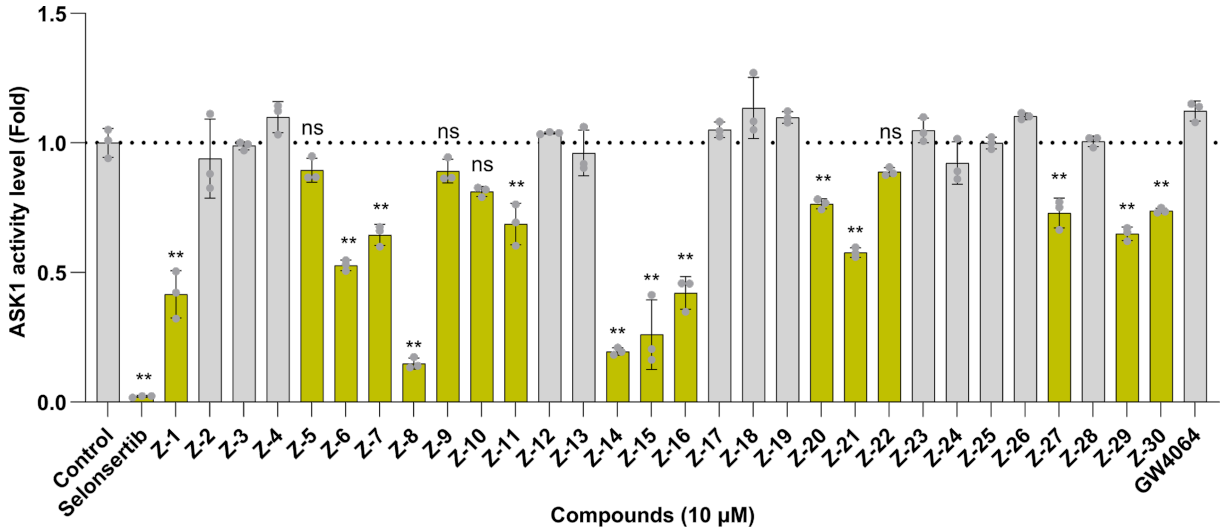


Figure 2: ASK1 inhibition of all compounds**.** Results of the ADP-Glo™ kinase assay after treating ASK1 kinase with selonsertib and test compounds for 1 hour. Compared to the control group: $^{**}p < 0.01$ and ns indicates no significant difference; control: DMSO control group.

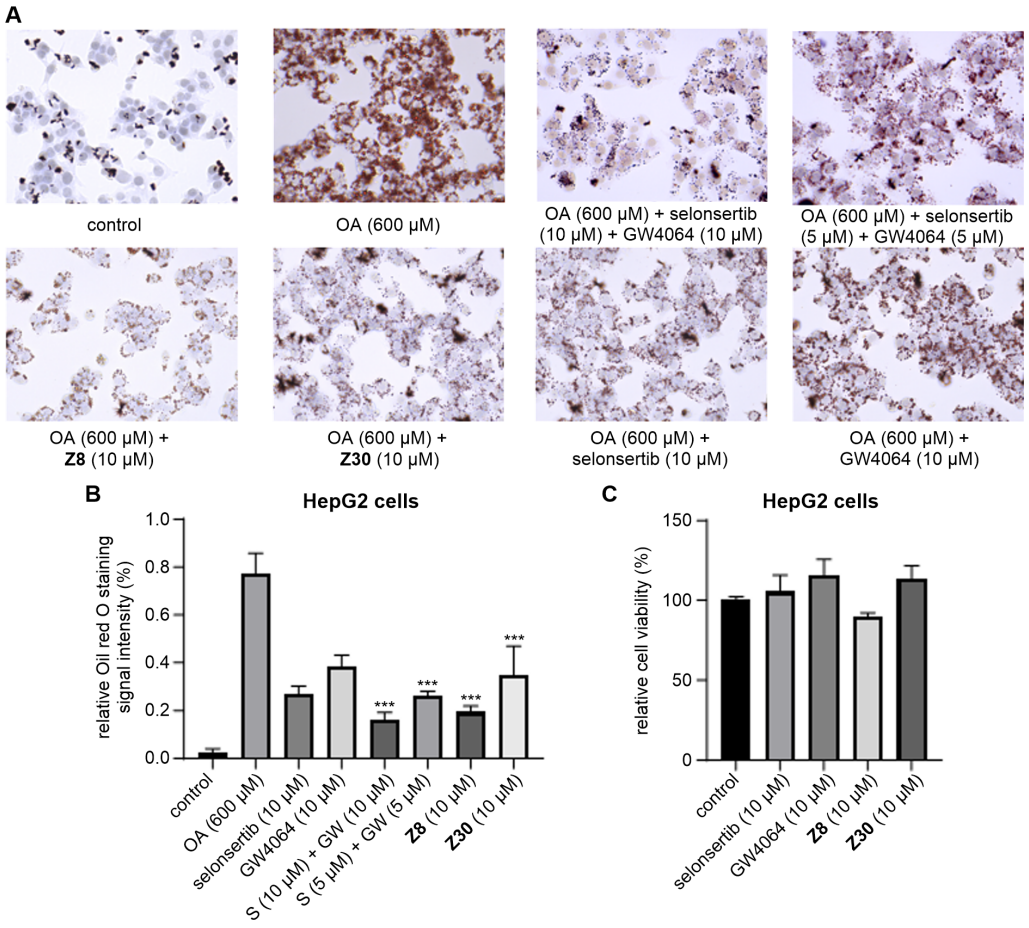


Figure 3: Effects of **Z8** and **Z30** on OA-induced lipid accumulation in HepG2 cells. Cells were treated with GW4064, selonsertib, **Z8**, **Z30**, or the combination of GW4064 and selonsertib for 2 h prior to OA stimulation (600 μM). (A) Representative Oil Red O staining images (40×). (B) Quantification of staining intensity by ImageJ. Data are presented as mean ± SD ($n = 3$). Statistical significance was analyzed by one-way ANOVA followed by the appropriate post hoc test; $^{***}p < 0.001$ versus OA control. Additional pairwise comparisons between **Z8** and the positive-control groups are described in the text. (C) Cell viability at 10 μM under the same treatment conditions.

either positive control alone, and its activity was comparable to that of **Z8**. As anticipated, both **Z8** and **Z30** demonstrated a dose-dependent attenuation of OA-induced lipid accumulation. As illustrated in Figure 3B, **Z8** demonstrated significantly higher activity in comparison to the positive controls selonsertib and GW4064. To ascertain whether the viability of the cells was impaired by the compounds, an MTT assay was conducted. The results of this assay demonstrated that neither compound caused a decrease in cell viability at concentrations of up to 10 μM (Figure 3C).

Conclusion

Given the respective pharmacological roles of FXR and ASK1 in MASH, we hypothesized that a dual modulator targeting both could potentially serve as an excellent therapeutic agent, concurrently activating FXR and inhibiting ASK1. In this work, a new modulator was designed by hybridizing GW4064 with GS-4997, aiming to overcome the limitations of GW4064 and potentiate the inhibitory activity against ASK1. Following comprehensive multiparameter optimization, compound **Z8** emerged as a lead FXR/ASK1 dual modulator, demonstrating suitable liver microsome stability and high target selectivity. Collectively, compound **Z8** holds great promise for the treatment of MASH.

Supporting Information

Supporting Information File 1

Experimental, characterization data and copies of spectra.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-59-S1.pdf]

Funding

We thank Hangzhou Medical College for financial support.

ORCID® iDs

Fengzhi Zhang - https://orcid.org/0000-0001-9542-6634

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Pipitone, R. M.; Ciccioli, C.; Infantino, G.; La Mantia, C.; Parisi, S.; Tulone, A.; Pennisi, G.; Grimaudo, S.; Petta, S. *Ther. Adv. Endocrinol. Metab.* **2023,** *14,* 20420188221145549. doi:10.1177/20420188221145549
2. Do, A.; Zahrawi, F.; Mehal, W. Z. *Nat. Rev. Drug Discovery* **2025,** *24,* 171–189. doi:10.1038/s41573-024-01084-2
3. Tao, L.; Yang, X.; Ge, C.; Zhang, P.; He, W.; Xu, X.; Li, X.; Chen, W.; Yu, Y.; Zhang, H.; Chen, S.-D.; Pan, X.-Y.; Su, Y.; Xu, C.; Yu, Y.; Zheng, M.-H.; Min, J.; Wang, F. *Cell Metab.* **2024,** *36,* 2190–2206.e5. doi:10.1016/j.cmet.2024.07.013
4. Wang, S.; Friedman, S. L. *Sci. Transl. Med.* **2023,** *15,* eadi0759. doi:10.1126/scitranslmed.adi0759
5. Wang, X.; Zhang, L.; Dong, B. *Hepatology (Hoboken, NJ, U. S.)* **2025,** *82,* 1303–1324. doi:10.1097/hep.0000000000000786
6. Keam, S. J. *Drugs* **2024,** *84,* 729–735. doi:10.1007/s40265-024-02045-0
7. Rausch, M.; Samodelov, S. L.; Visentin, M.; Kullak-Ublick, G. A. *Int. J. Mol. Sci.* **2022,** *23,* 13967. doi:10.3390/ijms232213967
8. Wang, K.; Zhang, Y.; Wang, G.; Hao, H.; Wang, H. *Med. Res. Rev.* **2024,** *44,* 568–586. doi:10.1002/med.21991
9. Jiang, L.; Zhang, H.; Xiao, D.; Wei, H.; Chen, Y. *Comput. Struct. Biotechnol. J.* **2021,** *19,* 2148–2159. doi:10.1016/j.csbj.2021.04.029
10. Adorini, L.; Trauner, M. *J. Hepatol.* **2023,** *79,* 1317–1331. doi:10.1016/j.jhep.2023.07.034
11. Wei, C.; Liu, Y.; Yan, Z.; Han, J.; Jiang, N. *J. Med. Chem.* **2025,** *68,* 6104–6107. doi:10.1021/acs.jmedchem.5c00539
12. Schierle, S.; Brunst, S.; Helmstädter, M.; Ebert, R.; Kramer, J. S.; Steinhilber, D.; Proschak, E.; Merk, D. *ChemMedChem* **2021,** *16,* 2366–2374. doi:10.1002/cmdc.202100118
13. Brys, R.; Gibson, K.; Poljak, T.; Van Der Plas, S.; Amantini, D. *Prog. Med. Chem.* **2020,** *59,* 101–179. doi:10.1016/bs.pmch.2020.02.001
14. Schuster, S.; Feldstein, A. E. *Nat. Rev. Gastroenterol. Hepatol.* **2017,** *14,* 329–330. doi:10.1038/nrgastro.2017.42
15. Sarkar, H. S.; Sen, A.; Hoque, I.; Ghosh Dastidar, U.; Pal, S.; Sarkar, D.; Chakrabarti, P.; Talukdar, A. *RSC Med. Chem.* **2025,** *16,* 3469–3494. doi:10.1039/d5md00252d
16. Shang, S.; Wan, Q.; Chen, F.; Hu, J. *Biochem. Biophys. Res. Commun.* **2024,** *705,* 149739. doi:10.1016/j.bbrc.2024.149739
17. Qin, J.; Cao, M.; Hu, X.; Tan, W.; Ma, B.; Cao, Y.; Chen, Z.; Li, Q.; Hu, G. *Eur. J. Med. Chem.* **2023,** *247,* 115057. doi:10.1016/j.ejmech.2022.115057
18. Schuster-Gaul, S.; Geisler, L. J.; McGeough, M. D.; Johnson, C. D.; Zagorska, A.; Li, L.; Wree, A.; Barry, V.; Mikaelian, I.; Jih, L. J.; Papouchado, B. G.; Budas, G.; Hoffman, H. M.; Feldstein, A. E. *JCI Insight* **2020,** *5,* e123294. doi:10.1172/jci.insight.123294
19. Wang, Y.; Wen, H.; Fu, J.; Cai, L.; Li, P.-L.; Zhao, C.-L.; Dong, Z.-f.; Ma, J.-p.; Wang, X.; Tian, H.; Zhang, Y..; Liu, Y.; Cai, J.; Shi, Z.-G.; Huang, Z.; Li, W.; Li, H. *Hepatology (Hoboken, NJ, U. S.)* **2020,** *71,* 93–111. doi:10.1002/hep.30822
20. Chen, Q.; Guo, J.; Qiu, T.; Zhou, J. *J. Gastroenterol. Hepatol.* **2023,** *38,* 378–385. doi:10.1111/jgh.16087
21. Katafuchi, T.; Makishima, M. *Int. J. Mol. Sci.* **2022,** *23,* 6046. doi:10.3390/ijms23116046
22. Che, T.; Roth, B. L. *Cell* **2023,** *186,* 5203–5219. doi:10.1016/j.cell.2023.10.029
23. Li, Y.; Wang, L.; Yi, Q.; Luo, L.; Xiong, Y. *Front. Nutr.* **2024,** *11,* 1447878. doi:10.3389/fnut.2024.1447878
24. Chiang, J. Y.; Ferrell, J. M. *Gene Expression* **2018,** *18,* 71–87.
25. Sun, R.; Kong, B.; Yang, N.; Cao, B.; Feng, D.; Yu, X.; Ge, C.; Feng, S.; Fei, F.; Huang, J.; Lu, Z.; Xie, Y.; Yang, C. S.; Guo, G. L.; Wang, G.; Aa, J. *Drug Metab. Dispos.* **2021,** *49,* 276–286. doi:10.1124/dmd.120.000215

26. Choi, Y.-J.; Johnson, J. D.; Lee, J.-J.; Song, J.; Matthews, M.; Hellerstein, M. K.; McWherter, C. A. *Am. J. Physiol.: Gastrointest. Liver Physiol.* **2024,** *326,* G120–G132. doi:10.1152/ajpgi.00158.2023
27. van den Hoek, A. M.; Verschuren, L.; Worms, N.; van Nieuwkoop, A.; de Ruiter, C.; Attema, J.; Menke, A. L.; Caspers, M. P. M.; Radhakrishnan, S.; Salic, K.; Kleemann, R. *Cells* **2020,** *9,* 2014. doi:10.3390/cells9092014
28. Schwabl, P.; Hambruch, E.; Budas, G. R.; Supper, P.; Burnet, M.; Liles, J. T.; Birkel, M.; Brusilovskaya, K.; Königshofer, P.; Peck-Radosavljevic, M.; Watkins, W. J.; Trauner, M.; Breckenridge, D. G.; Kremoser, C.; Reiberger, T. *Biomedicines* **2021,** *9,* 60. doi:10.3390/biomedicines9010060
29. Schierle, S.; Neumann, S.; Heitel, P.; Willems, S.; Kaiser, A.; Pollinger, J.; Merk, D. *J. Med. Chem.* **2020,** *63,* 8369–8379. doi:10.1021/acs.jmedchem.0c00618
30. Helmstädter, M.; Kaiser, A.; Brunst, S.; Schmidt, J.; Ronchetti, R.; Weizel, L.; Proschak, E.; Merk, D. *J. Med. Chem.* **2021,** *64,* 9525–9536. doi:10.1021/acs.jmedchem.1c00831
31. Ren, Q.; Chen, Y.; Zhou, Z.; Cai, Z.; Jiao, S.; Huang, W.; Wang, B.; Chen, S.; Wang, W.; Cao, Z.; Yang, Z.; Deng, L.; Hu, L.; Zhang, L.; Li, Z. *J. Med. Chem.* **2023,** *66,* 6082–6104. doi:10.1021/acs.jmedchem.2c01918
32. Meijer, F. A.; Leijten-van de Gevel, I. A.; de Vries, R. M. J. M.; Brunsveld, L. *Mol. Cell. Endocrinol.* **2019,** *485,* 20–34. doi:10.1016/j.mce.2019.01.022
33. Gibson, T. S.; Johnson, B.; Fanjul, A.; Halkowycz, P.; Dougan, D. R.; Cole, D.; Swann, S. *Bioorg. Med. Chem. Lett.* **2017,** *27,* 1709–1713. doi:10.1016/j.bmcl.2017.02.079
34. Wang, T.; Pang, L.; He, M.; Wang, Z. *Eur. J. Med. Chem.* **2023,** *262,* 115889. doi:10.1016/j.ejmech.2023.115889
35. Schmidt, J.; Rotter, M.; Weiser, T.; Wittmann, S.; Weizel, L.; Kaiser, A.; Heering, J.; Goebel, T.; Angioni, C.; Wurglics, M.; Paulke, A.; Geisslinger, G.; Kahnt, A.; Steinhilber, D.; Proschak, E.; Merk, D. *J. Med. Chem.* **2017,** *60,* 7703–7724. doi:10.1021/acs.jmedchem.7b00398
36. Makri, E. S.; Xanthopoulos, K.; Mavrommatis Parasidis, P.; Makri, E.; Pettas, S.; Tsingotjidou, A.; Cheva, A.; Ballaouri, I.; Gerou, S.; Goulas, A.; Polyzos, S. A. *Endocrine* **2024,** *85,* 704–716. doi:10.1007/s12020-024-03769-5
37. Chen, S.; Xie, J.; Ye, R.; Xu, D. D.; Yang, Y. *Chem. Sci.* **2024,** *15,* 10366–10380. doi:10.1039/d4sc00094c
38. Jung, Y.-H.; Salmaso, V.; Wen, Z.; Bennett, J. M.; Phung, N. B.; Lieberman, D. I.; Gopinatth, V.; Randle, J. C. R.; Chen, Z.; Salvemini, D.; Karcz, T. P.; Cook, D. N.; Jacobson, K. A. *J. Med. Chem.* **2021,** *64,* 5099–5122. doi:10.1021/acs.jmedchem.1c00164
39. Mathur, V.; Alam, O.; Siddiqui, N.; Jha, M.; Manaithiya, A.; Bawa, S.; Sharma, N.; Alshehri, S.; Alam, P.; Shakeel, F. *Molecules* **2023,** *28,* 5860. doi:10.3390/molecules28155860
40. Kinzel, O.; Steeneck, C.; Schlüter, T.; Schulz, A.; Gege, C.; Hahn, U.; Hambruch, E.; Hornberger, M.; Spalwisz, A.; Frick, K.; Perović-Ottstadt, S.; Deuschle, U.; Burnet, M.; Kremoser, C. *Bioorg. Med. Chem. Lett.* **2016,** *26,* 3746–3753. doi:10.1016/j.bmcl.2016.05.070
41. Akwabi-Ameyaw, A.; Bass, J. Y.; Caldwell, R. D.; Caravella, J. A.; Chen, L.; Creech, K. L.; Deaton, D. N.; Jones, S. A.; Kaldor, I.; Liu, Y.; Madauss, K. P.; Marr, H. B.; McFadyen, R. B.; Miller, A. B.; Navas, F., III; Parks, D. J.; Spearing, P. K.; Todd, D.; Williams, S. P.; Wisely, G. B. *Bioorg. Med. Chem. Lett.* **2008,** *18,* 4339–4343. doi:10.1016/j.bmcl.2008.06.073
42. Maloney, P. R.; Parks, D. J.; Haffner, C. D.; Fivush, A. M.; Chandra, G.; Plunket, K. D.; Creech, K. L.; Moore, L. B.; Wilson, J. G.; Lewis, M. C.; Jones, S. A.; Willson, T. M. *J. Med. Chem.* **2000,** *43,* 2971–2974. doi:10.1021/jm0002127
43. Wang, T.; Pang, L.; He, M.; Wang, Z. *Eur. J. Med. Chem.* **2023,** *262,* 115889. doi:10.1016/j.ejmech.2023.115889
44. Himmelbauer, M. K.; Xin, Z.; Jones, J. H.; Enyedy, I.; King, K.; Marcotte, D. J.; Murugan, P.; Santoro, J. C.; Hesson, T.; Spilker, K.; Johnson, J. L.; Luzzio, M. J.; Gilfillan, R.; de Turiso, F. G.-L. *J. Med. Chem.* **2019,** *62,* 10740–10756. doi:10.1021/acs.jmedchem.9b01206
45. Gege, C.; Kinzel, O.; Steeneck, C.; Schulz, A.; Kremoser, C. *Curr. Top. Med. Chem.* **2014,** *14,* 2143–2158. doi:10.2174/1568026614666141112094430
46. Li, J.; Liu, M.; Li, Y.; Sun, D.-d.; Shu, Z.; Tan, Q.; Guo, S.; Xie, R.; Gao, L.; Ru, H.; Zang, Y.; Liu, H.; Li, J.; Zhou, Y. *J. Med. Chem.* **2020,** *63,* 12748–12772. doi:10.1021/acs.jmedchem.0c01065
47. Sepe, V.; Marchianò, S.; Finamore, C.; Baronissi, G.; Di Leva, F. S.; Carino, A.; Biagioli, M.; Fiorucci, C.; Cassiano, C.; Monti, M. C.; del Gaudio, F.; Novellino, E.; Limongelli, V.; Fiorucci, S.; Zampella, A. *ACS Med. Chem. Lett.* **2019,** *10,* 407–412. doi:10.1021/acsmedchemlett.8b00423

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.59

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Site-specific labelling of native peptides and proteins: chemical and enzymatic strategies

Antonio Angelastro*[1], Jonathan Bargh[2], Subhajit Guria[3], Victor Laserna[1] and Louis Luk*[4]

Review

Open Access

Address:
[1]Biologics Engineering, Oncology R&D, Cambridge Biomedical Campus, AstraZeneca, 1 Francis Crick Ave, Trumpington, Cambridge, CB2 0AA, UK, [2]Hit Discovery, Discovery Sciences, R&D, AstraZeneca, Pepparedsleden 1, Mölndal, Gothenburg, 43183, Sweden, [3]Interdisciplinary Research Center, Division of Interdisciplinary Science, Institute for Radiation Sciences, The University of Osaka, 1-2 Machikaneyama-cho, Toyonaka 560-0043, Osaka, Japan and [4]School of Chemistry and Cardiff Catalysis Institute, Cardiff University Main Building, Park Pl, Cardiff, CF10 3AT, UK

Email:
Antonio Angelastro* - antonio.angelastro@astrazeneca.com;
Louis Luk* - lukly@cardiff.ac.uk

* Corresponding author

Keywords:
affinity-guided chemistry; disulfide-rebridging; enzyme catalysis; glycan modifications; native-sequence; protein labelling; small-molecule modifications; terminal

Beilstein J. Org. Chem. **2026,** *22,* 857–881.
https://doi.org/10.3762/bjoc.22.67

Received: 28 January 2026
Accepted: 22 April 2026
Published: 03 June 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2026 Angelastro et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Site-specific modifications of native-sequence proteins are technologies that underpin progresses in chemical biology, diagnostics and next-generation biotherapeutics. However, the pursuit of site-specificity has often come at the expense of scalability and usability, ultimately limiting translational potential of a modification tool. This review critically compares current key strategies, including terminal, disulfide-rebridging, small-molecule, glycan and enzyme-based modifications. Finally, we present a decision tree for method selection and highlight opportunities for innovation in next-generation native-sequence protein modification technologies.

Introduction

Protein and peptide modifications introduce functionalities not achievable through recombinant methods alone, making them vital for fundamental research and biomedical applications [1-4]. Such chemistry underpins wide-ranging biological investigations including microscopy, proteomic analyses and tissue sampling. Their translational power is evident from the success of biotherapeutics: antibody–drug conjugates (ADCs), in which cytotoxic payloads covalently linked to targeting antibodies, are at the forefront of cancer therapy, delivering tumour-selective killing for improved patient outcomes; Neulasta, an *N*-terminal

PEGylated granulocyte colony-stimulating factor (G-CSF), is administered to reduce risk of neutropenia in patients receiving chemotherapy. Despite clinical successes, unsolved technical challenges remain. Taking ADC development as an example, a single payload can be insufficient to overcome resistance, highlighting the need for complementary site-specific antibody labelling strategies. Meanwhile, advances in high-throughput screening and AI-guided design have been creating new scaffolds, including nanobodies, de novo binders and peptoids, where tailored modification chemistries have strong potential to generate novel biotherapeutics.

As the development of protein-based therapeutics and diagnostics continues to expand to address global health challenges, the demand for tailored protein-modification strategies is increasing in parallel. Indeed, protein-modification chemistry is a key area of chemical biology in which academic discoveries can rapidly advance from fundamental research (technology readiness level, TRL 1) to patentable, preclinical-stage technologies (TRL 4–5), with some ultimately reaching real-world applications approaching TRL 9. The ADC market alone is projected to reach USD 32 billion by 2034 (https://www.precedenceresearch.com/antibody-drug-conjugates-market), with advanced modalities such as bispecific antibody–drug conjugates (BsADC) gaining traction [2]. Leading suppliers, including ThermoFisher Scientific (Molecular Probes/Invitrogen), LI-COR Biosciences, Roche, Jena Bioscience, Merck (Sigma-Aldrich), New England Biolabs, Promega, Agilent and Cytiva (Danaher), offer a plethora of reagents and kits, from fluorescent tags, crosslinkers to enzymatic tools. Meanwhile, companies such as EnzyTag, Shiru, and Revvity offer bespoke conjugation services. Beyond therapeutics, advances in protein modification have the potential to partially replace solid-phase peptide synthesis (SPPS), reducing environmental impact by minimising protecting groups and harmful solvents. Driven by scientific innovation, clinical demand and sustainability goals, the global protein-modification market is expected to reach USD 6.02 billion by 2034 (https://www.towardshealthcare.com/insights/protein-labeling-market-sizing).

Review

Key criteria for effective native-sequence modifications

Effective protein modification relies on tools that balance three interlinked factors translating basic research to commercial applications: site-specificity, usability and scalability.

Site-specificity. The ideal modification chemistry targets a single, well-defined site without stochastic off-site reactions. Product homogeneity enhances safety by mitigating toxic by-products that require excessive chromatography, while also offering well-defined pharmacokinetic and pharmacodynamic profiles. Homogeneous batches of labelled proteins thus improve therapeutic indices and reproducibility, key determinants of clinical trial outcomes.

Usability. Effective strategies act directly on native proteins, avoiding substrate engineering and minimising risks such as aggregation and immunogenicity, and hence site-specific modifications of commonly found surface residues are desired. Tools are also expected to be easy to prepare and apply, yet remain reliably functional in vitro and in vivo settings. The use of harsh reagents, organic solvents or chaotropic agents must be avoided to preserve protein stability, preventing folding heterogeneity and supporting sustainability.

Scalability. For therapeutic and industrial applications, modification methods should minimise waste and curtail costs, reflected in low process mass intensity (PMI). Reagents should be near-stoichiometric, purification steps minimal, and catalysts effective in substoichiometric quantities. Compatibility with one-pot system and automation, such as flow chemistry and robotic synthesis, further enhance scalability.

Existing tools and techniques

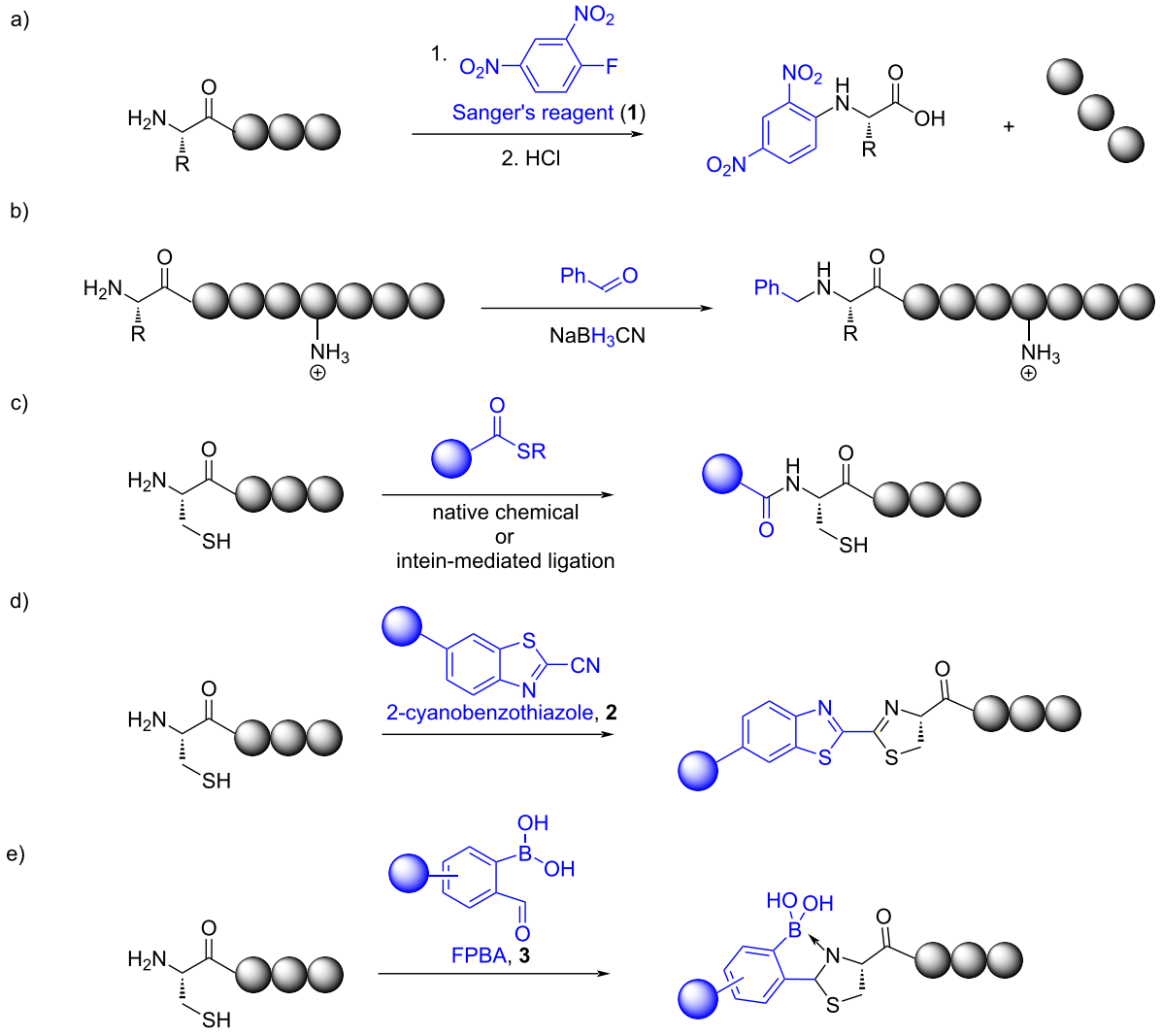
While numerous native-sequence protein-modification tools exist with many being site-specific, few can fulfil all the aforementioned criteria simultaneously. Cysteine modification is among the most effective strategies due to the high reactivity of the thiolate, enabling selective chemical transformations [3,4]. However, the limited abundance of free cysteine residues in proteins has driven the development of alternative approaches targeting other chemical motifs. This review summarizes current strategies for site-specific modifications of these alternative sites, followed by a comparative analysis and outlook on future developments. The strengths, limitations and opportunities are also discussed. Given the diversity and breadth of protein-modification chemistry, more detailed reviews on specific topics are referenced within relevant subsections.

Protein terminal modifications

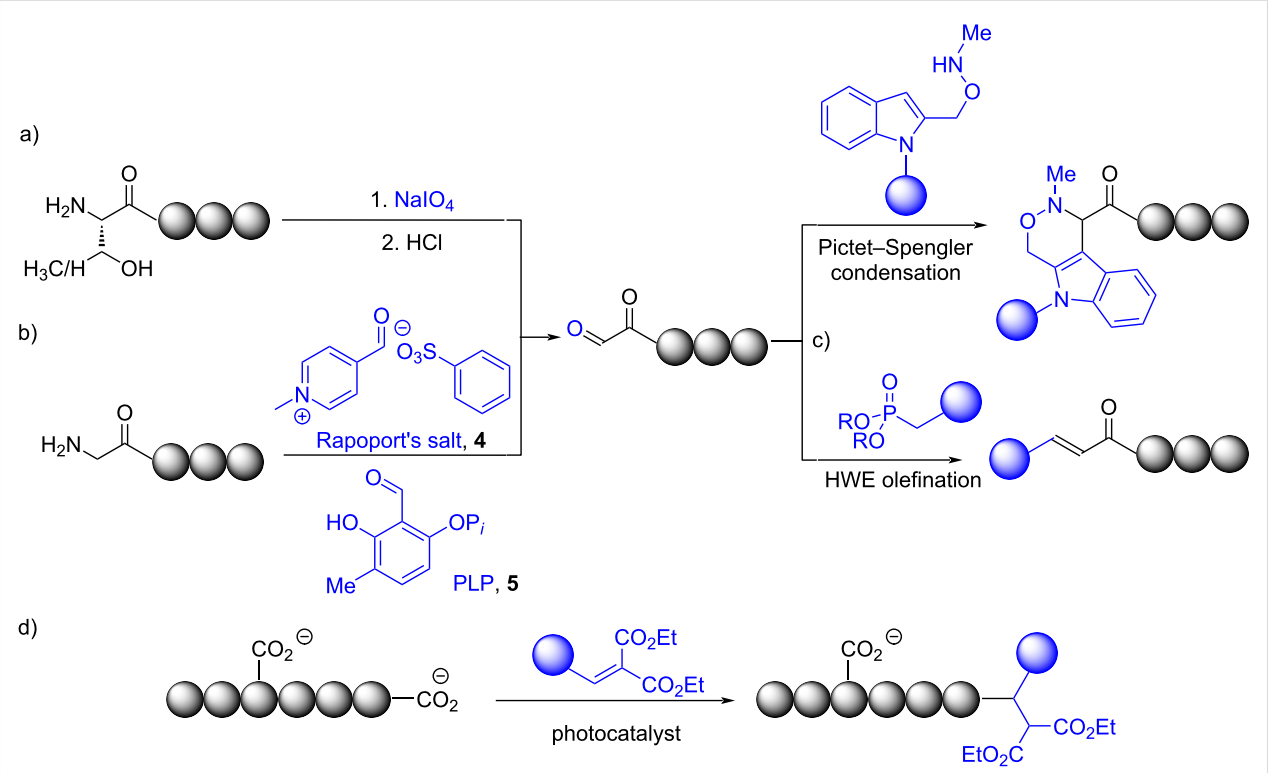
The majority of proteins are linear polypeptide chains, possessing a single *N*- and *C*-terminus and hence presenting these regions as attractive targets for site-specific modification. Indeed, over 80% of protein termini are solvent-exposed and chemically accessible [5]. Notably, the *N*-terminal amine and *C*-terminal carboxyl group exhibit subtle biophysical differences from those of internal residues (lysine, aspartate and glutamate), allowing development of tailored chemistries.

The first chemical modification of an *N*-terminus was reported by Frederick Sanger in 1945 [6], when 1-fluoro-2,4-dinitrobenzene (Sanger's reagent, **1**) was used to label the α-amino group of insulin during sequence determination (Scheme 1a). Its selectivity over lysine's ε-amine originates from the low pK_a of the *N*-terminus, influenced by the adjacent carbonyl group and systematically characterized at ≈6–7 (Scheme 1b) [7]. Building on this principle, pH-controlled *N*-terminal PEGylation has been applied to the labelling of filgrastim, producing Neulasta®, a long-acting chemotherapy support drug [8].

Enhanced *N*-terminal modification selectivity can be achieved with dinucleophilic residues. Native chemical ligation recruits *N*-terminal cysteine to form amide bonds via *trans*-thioesterification, though often under denaturing conditions (Scheme 1c) [9]. Intein-mediated ligation is a recombinant alternative but often requires linker sequences. Newer reagents such as 2-cyanobenzothiazole **2** (CBT) react with *N*-terminal cysteine to form luciferin-like linkages (≈9 M^{-1} s^{-1}; Scheme 1d) [10], whereas 2-formylphenylboronic acid **3** (FPBA) generates thiazolidino boronate with noticeably improved kinetics (10^3 M^{-1} s^{-1}; Scheme 1e) [11,12]. Aldehydes can be introduced at protein *N*-termini through several strategies: oxidation of serine or threonine with sodium periodate (Scheme 2a) [13], selective modification of glutamate using Rapoport's salt (**4**) which has generated modified trastuzumab in moderate yield (67%; Scheme 2b) [14], or pyridoxal phosphate (PLP, **5**)-mediated carbonyl formation at residues such as serine, threonine, glycine, alanine, or cysteine [15]. These inter-



Scheme 1: a) Sanger's reagent (**1**), b) reductive amination by sodium cyanoborohydride, c) native chemical ligation, d) cysteine-reactive 2-cyanobenzothiazole **2** and e) formylphenylboronic acid **3** (FPBA) reacting with *N*-terminal cysteine.



Scheme 2: Introduction of carbonyl functionality at protein *N*-termini via a) oxidation by periodate, b) Rapoport's salt (**4**) or pyridoxal phosphate (PLP, **5**), c) payload attached via Pictet–Spengler or Horner–Wadsworth–Emmons (HWE) reaction, and d) *C*-terminal modification through photoredox catalysis.

mediates enable downstream chemical conjugations such as Pictet–Spengler condensation [16-18] and Horner–Wadsworth–Emmons (HWE) olefination [19] (Scheme 2c).

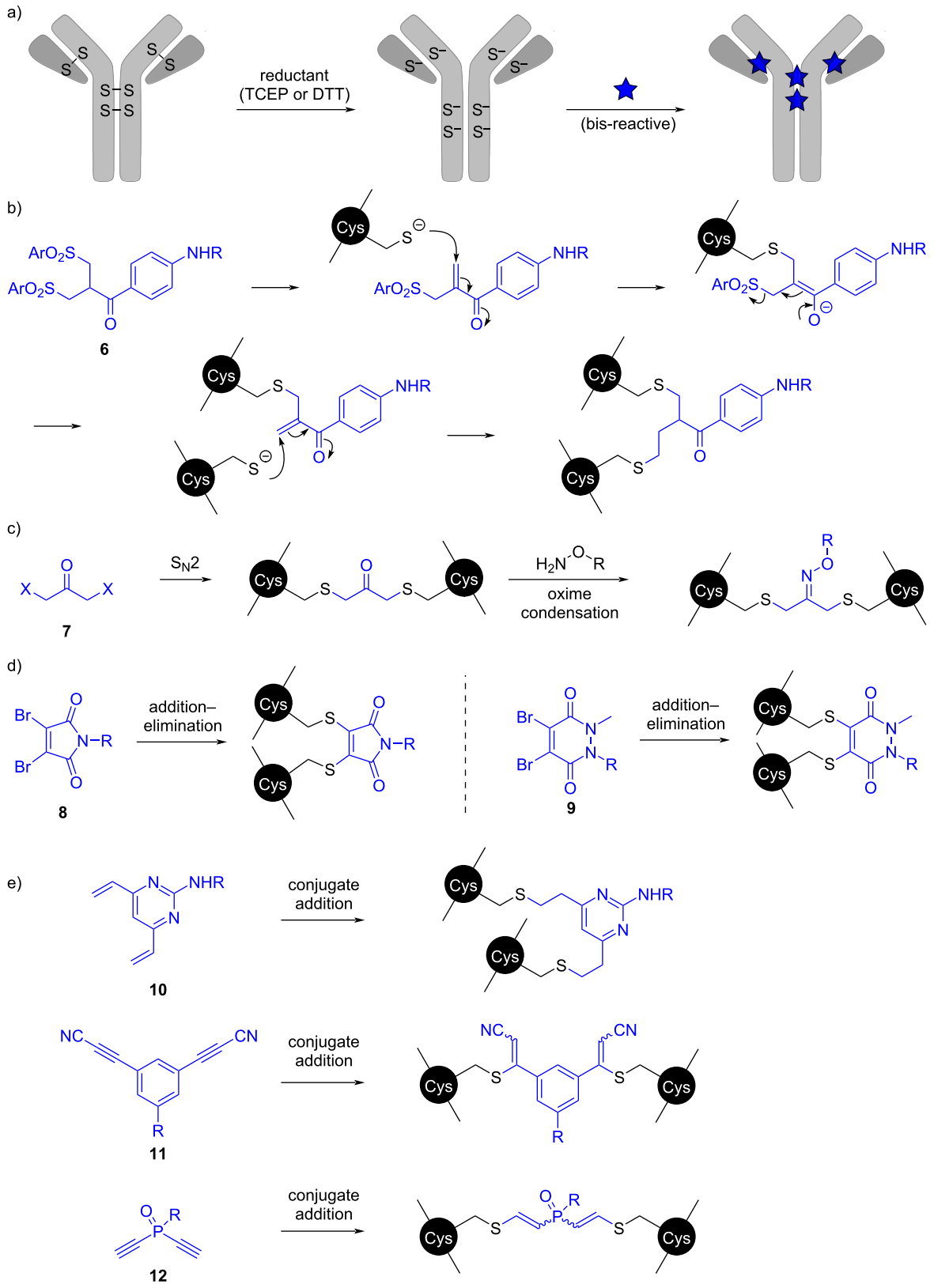
With respect to *C*-terminal modification, one of the earliest strategies involved the use of intein-mediated protein splicing to generate a *C*-terminal thioester from recombinantly expressed proteins [20-22]. One distinctive feature of the *C*-terminal carboxylate is its lower oxidation potential compared to the side chains of aspartate or glutamate, making it more amenable to selective redox-based modifications. This has led to the development of labelling techniques, wherein the *C*-terminal carboxylate can be functionalized via visible-light photoredox decarboxylative catalysis (Scheme 2d) [23]. Similarly, decarboxylative alkylation has been applied to complex proteomes to enable proteome-wide *C*-terminal profiling [24].

Despite significant advances, major challenges persist. Current *N*-terminal labelling strategies often require engineered residues, operate only under narrow pH conditions, or involve multistep procedures that hinder practical application. Efficiencies often depend on methionine excision, linker or cleavable leader peptides. *C*-Terminal labelling typically suffers from modest conversions (≈50% for photoredox strategies). While selectivity has been improved, few approaches are fully specific. For further discussion, see reviews [25,26].

Rebridging agents

Protein sidechains frequently form complementary pairs, such as disulfide bonds (cystine), salt bridges, hydrogen bonds, and cation–π interactions. Among these, disulfide bonds between cysteine residues are the most widely used for native-sequence protein labelling, due to their unique nucleophilicity and low abundance. Therapeutically relevant proteins, including somatostatin, immunoglobulins (across all isotypes and subclasses), insulin and interleukin-10 naturally contain disulfide bridges. These can be selectively reduced to yield two nucleophilic thiols for reaction with bis-electrophilic reagents to rebridge the cysteines, while introducing a chemical label or crosslinking functionality (Scheme 3a).

One of the earliest rebridging strategies employed bis-sulfone reagents **6** which, upon in situ activation, generate monosulfone Michael acceptors that selectively react with a reduced cysteine pair (Scheme 3b) [27]. This approach was later adapted for native-sequence ADC design on IgG_1, where four accessible interchain disulfide bonds can be reduced and rebridged to



Scheme 3: a) Concept of disulfide rebridging. Reagents include b) sulfones **6**, c) dihaloacetones **7**, d) dibromomaleimides **8** and dibromopyridazinediones **9**, e) divinylpyrimidines **10**, arylenedipropiolonitriles **11**, and diethynylphosphonites **12**.

yield conjugates with a drug-to-antibody ratio (DAR) of 4, offering optimal pharmacokinetics and reduced aggregation versus a stochastic maleimide-conjugated antibody. Incorporating cytotoxic payloads into bis-sulfone reagents yielded functional conjugates that exhibited stability in the presence of albumin in vivo [28]. Notably, this technology offers one of the few established routes to generate highly homogeneous ADC candidates, with a moderate level of drug loading, from native antibody sequences.

The bis-sulfone ADC discovery spurred the development of a wide range of rebridging systems, with the apex goal of creating ADCs with a DAR of 4. For example, dihaloacetones **7** have been utilized for rebridging by alkylation, whereby the ketone-containing rebridged antibody product can be reacted with hydroxylamine-containing payloads to afford oxime-linked conjugates (Scheme 3c) [29,30]. Additionally, dibromomaleimides **8** [31,32] and dibromopyridazinediones **9** (Scheme 3d) [33] bearing payloads at the *N*-position can rebridge cysteines through an addition–elimination mechanism. These reagents have been extensively explored for protein conjugates, especially as ADCs [34]. Similarly, divinylpyrimidines **10** [35] and arylenedipropiolonitriles **11** [36] have also been investigated for analogous purposes but via a conjugate addition mechanism (Scheme 3e). Exploiting the unique oxidation state of phosphorous, diethynylphosphonites **12** also utilize conjugate addition to provide a versatile three-point conjugation platform [37].

Related technologies have been adapted to link multiple disulfide residues simultaneously. Using bis-dibromopyridazinediones **bis-9** [38] or bis/tetra-divinylpyrimidines **bis/tetra-10** [39], the toolbox of native sequence modification has been expanded to enable the construction of homogeneous conjugates bearing 1 or 2 payloads (Scheme 4). In the case of tetra-divinylpyrimidines **tetra-10**, antibody conjugates have been constructed that bear a single fluorophore, cytotoxin or SpyCatcher protein [40]. Despite the considerable modification of the chemical space around the antibody's hinge region, antigen affinity and antibody stability are largely retained.

Translating disulfide chemistries into therapeutic applications requires optimisation of key parameters including reaction kinetics, minimisation of reaction steps, conjugate stability and solubility. Furthermore, disulfide rebridging carries an inherent risk of scrambling, compromising homogeneity. In the case of DAR 4 ADCs, this approach enables conjugation with fairly homogeneous loading, yet heterogeneity within populations of the same DAR species exists, with bis-reactive reagents forming a mixture of half- and full-antibody conjugates through intra- and interchain rebridging, respectively (Scheme 5). Nevertheless, the two halves of the half-antibody are still held tightly together by non-covalent interchain interactions. It remains unclear whether half-antibody structures are less stable or efficacious than their full counterparts [41], however, it can be posited that the technology has not yet entirely fulfilled its original promise of completely homogeneous conjugate generation.

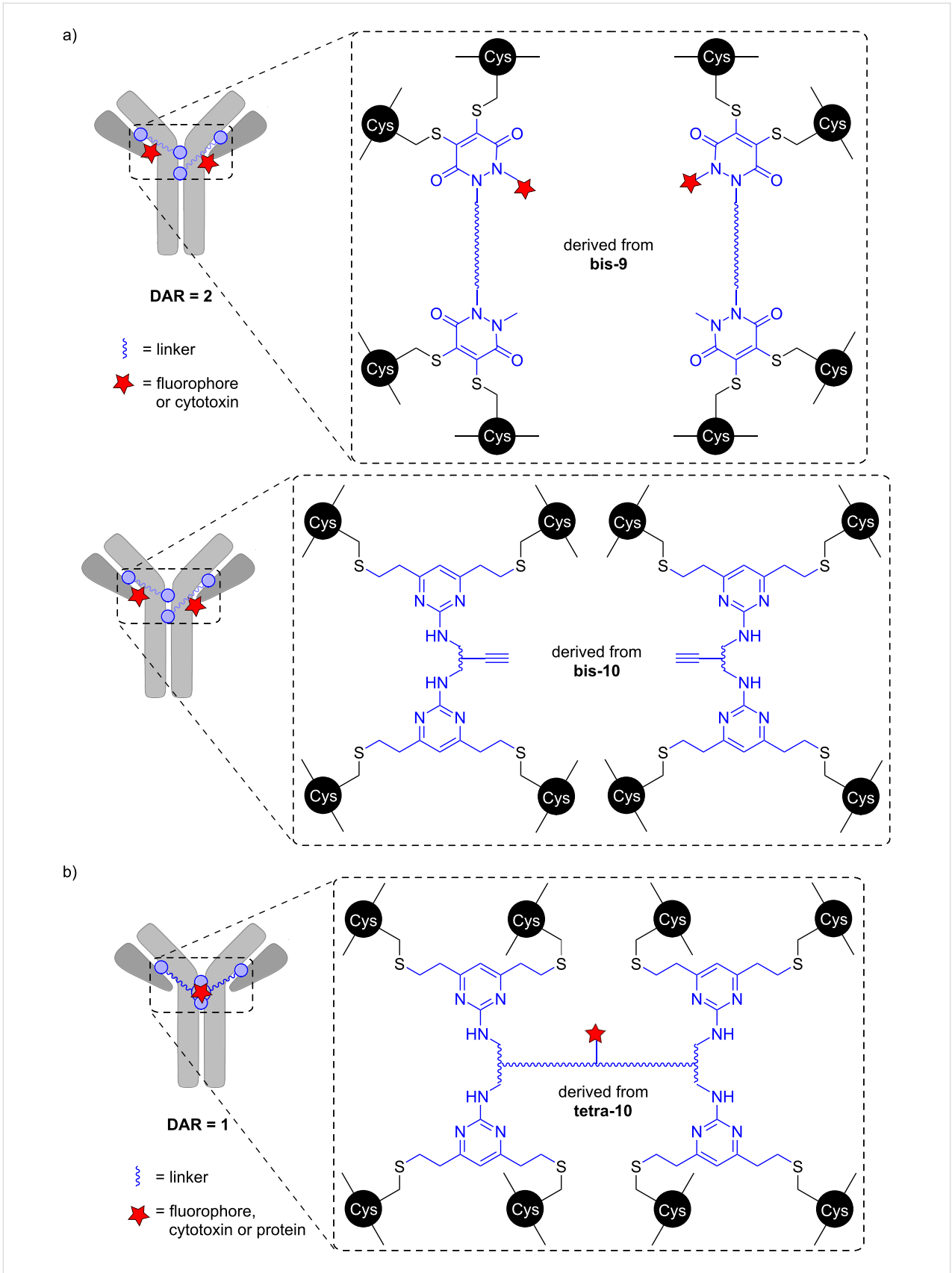
More broadly, this strategy is limited by its reliance on native cystine residues, making it unsuitable for proteins that lack disulfide bonds or that aggregate upon reduction. Its application to endogenous protein labelling is also restricted, as disulfide reduction may destabilise native protein structures, limiting uses in biological research. Perhaps, to expand applicability, future efforts may target alternative functional pairs, such as aromatic–π systems and the *N*-/*C*-termini (which are often in close proximity). One related example is cross-linking histidine residues by use of monosulfone [42]. For in-depth discussions of rebridging technologies, see references [34,43-45].

Affinity-guided protein modifications

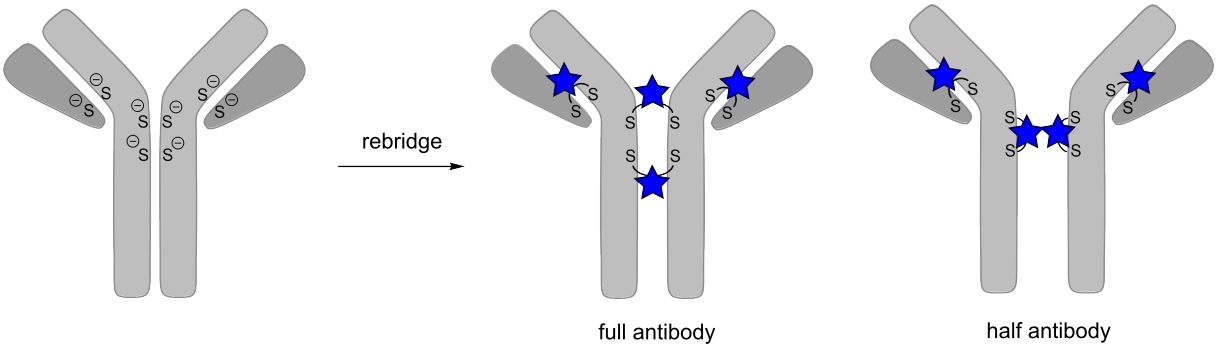
Affinity-guided protein labelling, also known as affinity-directed conjugation, is a native-sequence modification strategy founded on supramolecular recognition. It employs a bifunctional probe composed of (1) an affinity ligand that binds to a localised region of a target protein, and (2) a reactive moiety, either a pre-installed electrophile or a catalytic unit that generates one. Once bound, the reactive group is transferred to the target protein, enabling covalent modification. This concept has become a cornerstone of biological research and therapeutic development, with applications including in vitro systems, live cells and animal models.

The earliest form of affinity-guided labelling traces back to mechanism-based inhibitors, which exploit heightened reactivity of active-site residues for enzyme modification [46]. This concept evolved into activity-based protein profiling (ABPP), pioneered by the Cravatt group, where peptides or small molecules bearing electrophilic groups covalently react with active-site residues, enabling phenotype-driven functional profiling in a complex biological environment [46]. In a related concept, the Chin group reported the genetic incorporation of the unnatural amino acid 2,3-diaminopropionic acid (DAP, **13**), a nucleophilic amine that has a relatively low pK_a (6–8) and is structurally analogous to serine or cysteine, enabling identification of native-sequence protease substrates through site-specific hydrolysis and isopeptide-bond formation (Scheme 6) [47].

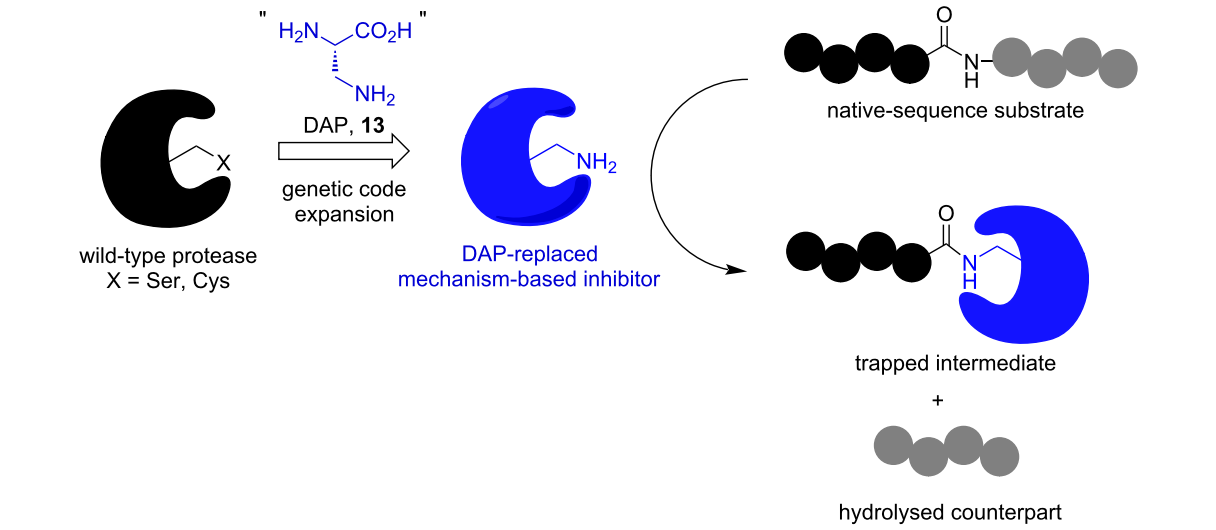
Labelling residues on protein surfaces is noticeably more challenging due to a lack of residues with heightened reactivity. Substantial progress has been made to address this issue, with current strategies broadly classified into non-catalytic and catalytic approaches.



Scheme 4: Multi-disulfide rebridging technologies.



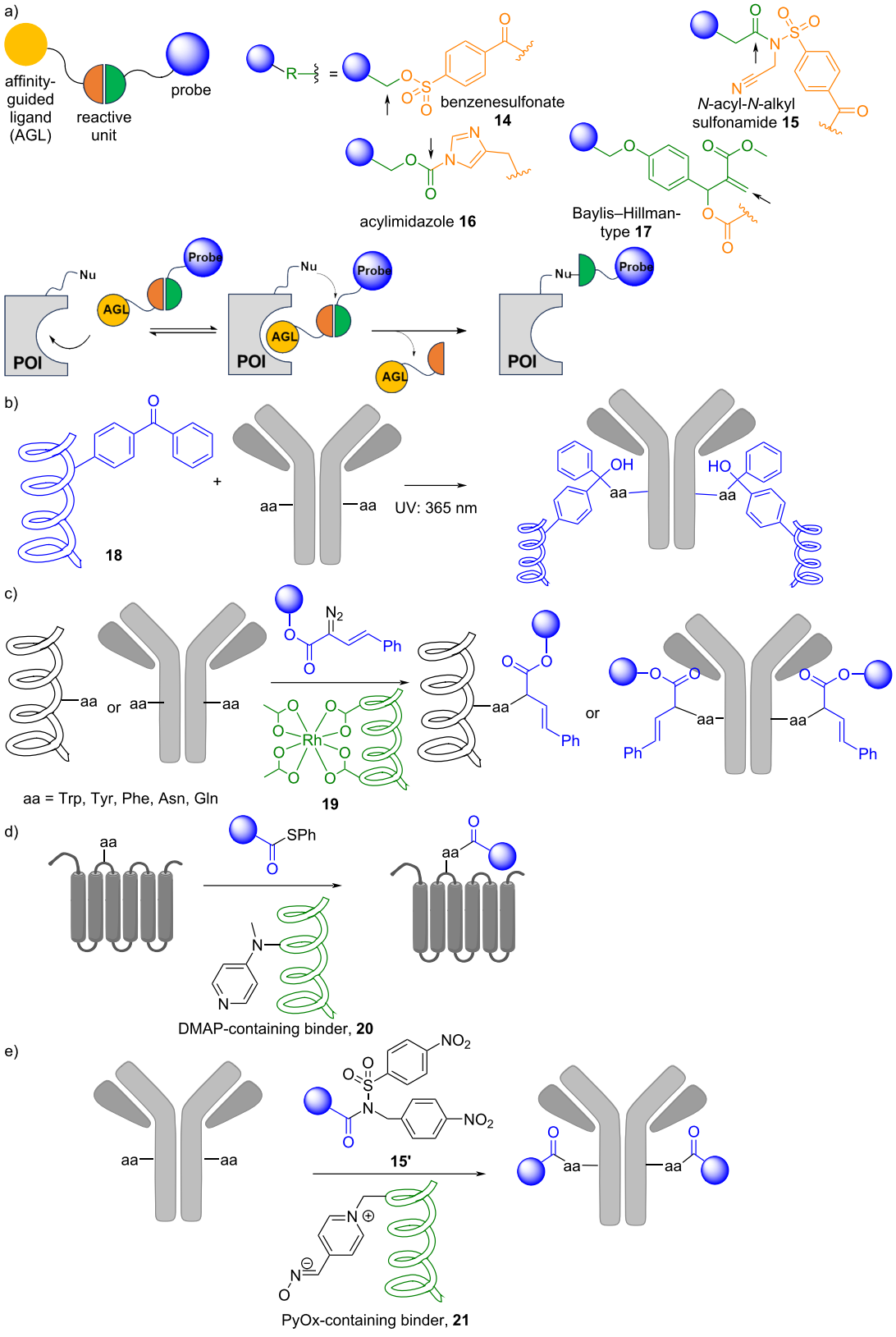
Scheme 5: Formation of full- and half-antibody as a result of disulfide rebridging.



Scheme 6: Use of 2,3-diaminopropionic acid (DAP, **13**) for native-sequence substrate profiling.

Non-catalytic approaches often involve tethering a reactive electrophile to a protein-binding ligand, thereby positioning the electrophile near a neighbouring nucleophilic residue. Example electrophiles include 4-methylbenzenesulfonates **14**, *N*-acyl-*N*-alkyl sulfonamides **15**, acylimidazoles **16**, Baylis–Hillman-type **17** reagents, among others (see reviews [48,49] for a comprehensive summary; Scheme 7a). Highlighting the challenges in pursuing both reactivity and stability, the *N*-acyl-*N*-alkyl sulfonamide possesses superior kinetics [50], whereas tosylation is one of the few procedures compatible with red blood cells and animal labelling [51]. Alternatively, benzophenone **18**, which can be incorporated in the form of an unnatural amino acid, can form a UV-induced diradical for protein modifications (Scheme 7b) [52-55]. Proteins, including IgG (trastuzumab) [56], EGFR [57], the adenosine A1 receptor [58] and P2X7 [59], have been labelled in vitro and/or in cellular systems.

Catalytic approaches generate reactive species in situ, thus enhancing both selectivity and efficiency. A notable example involves rhodium (Rh)-based catalysis, which mediates carbene transfer to modify electron-rich residues (Scheme 7c). This was first demonstrated in a coil dimer (Jun/c-Fos), where a dirhodium complex on one subunit enabled selective labelling of tryptophan, tyrosine and phenylalanine residues on the partner subunit [60]. This Rh-based method (**19**) was later adapted for modification of an asparagine (Asn) residue in the Fc domain of trastuzumab [61]. Rhodium has also been used in visible-light-triggered single-electron transfer (SET) reactions, enabling labelling under mild conditions with spatiotemporal control. Another strategy employs a 4-dimethylaminopyridine-containing binder **20**, in which dimethylaminopyridine (DMAP) acts as nucleophilic base catalyst mediating acyl cation transfer, that was successfully applied to label lectins and cell-surface re-



Scheme 7: Affinity-guided labelling. a) Reagents and catalysts appended to the binder include electrophilic motifs **14**–**17** as well as b) benzophenone **18**, c) rhodium complexes **19**, d) 4-dimethylaminopyridine **20** (DMAP), and e) pyridinium aldoxime **21** (PyOx).

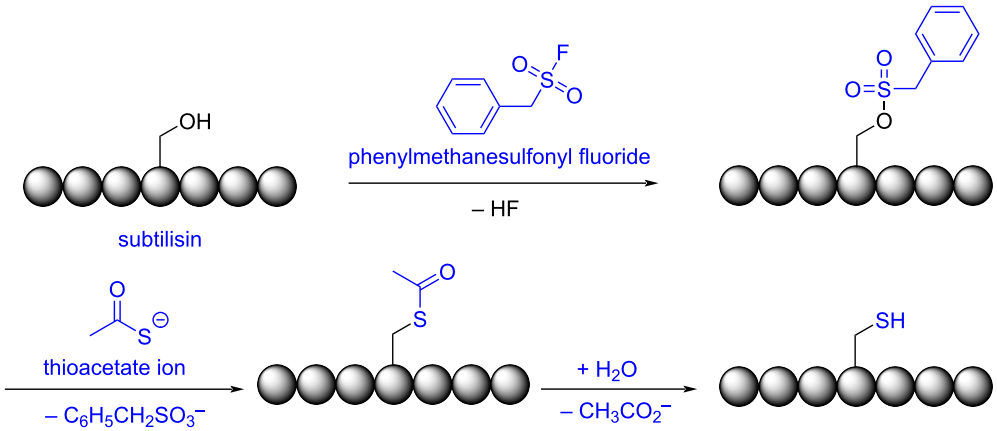
ceptors (HER2 and EGFR; Scheme 7d) [62,63]. More recently, genetic code expansion has enabled pyridinium aldoxime **21** (PyOx) incorporation into protein scaffolds (Scheme 7e) [64]. Acting as a catalytic centre for acyl cation transfer, PyOx reacts with a selected *N*-acyl-*N*-alkyl sulfonamide **15'** to generate a reactive ester intermediate. This strategy has successfully modified lysine residues in model proteins Z-DM, Fc domains and trastuzumab.

Affinity-guided protein labelling has proven highly valuable for both biological and translational research. Approaches involving genetic code expansion are especially promising, as this technique is now widely implemented in most molecular biology laboratories. The diversity of binder templates has also expanded rapidly with advances in AI-based and high-throughput technologies [65-77]. Nevertheless, design considerations remain critical; for example, benzophenone could cross-link an entire binder to its target protein [52-55], whereas incorporation of DAP requires photodeprotection [47]. Furthermore, all current proximity-directed reagents, including catalytic systems such as the reactions mediated by DMAP and PyOx (Scheme 7d and 7e), still require extensive screening and substantial stoichiometric quantities, which significantly limits their up-scaling potential. For further discussion, including non-specific labelling for protein interactome analyses, see reviews [48,49].

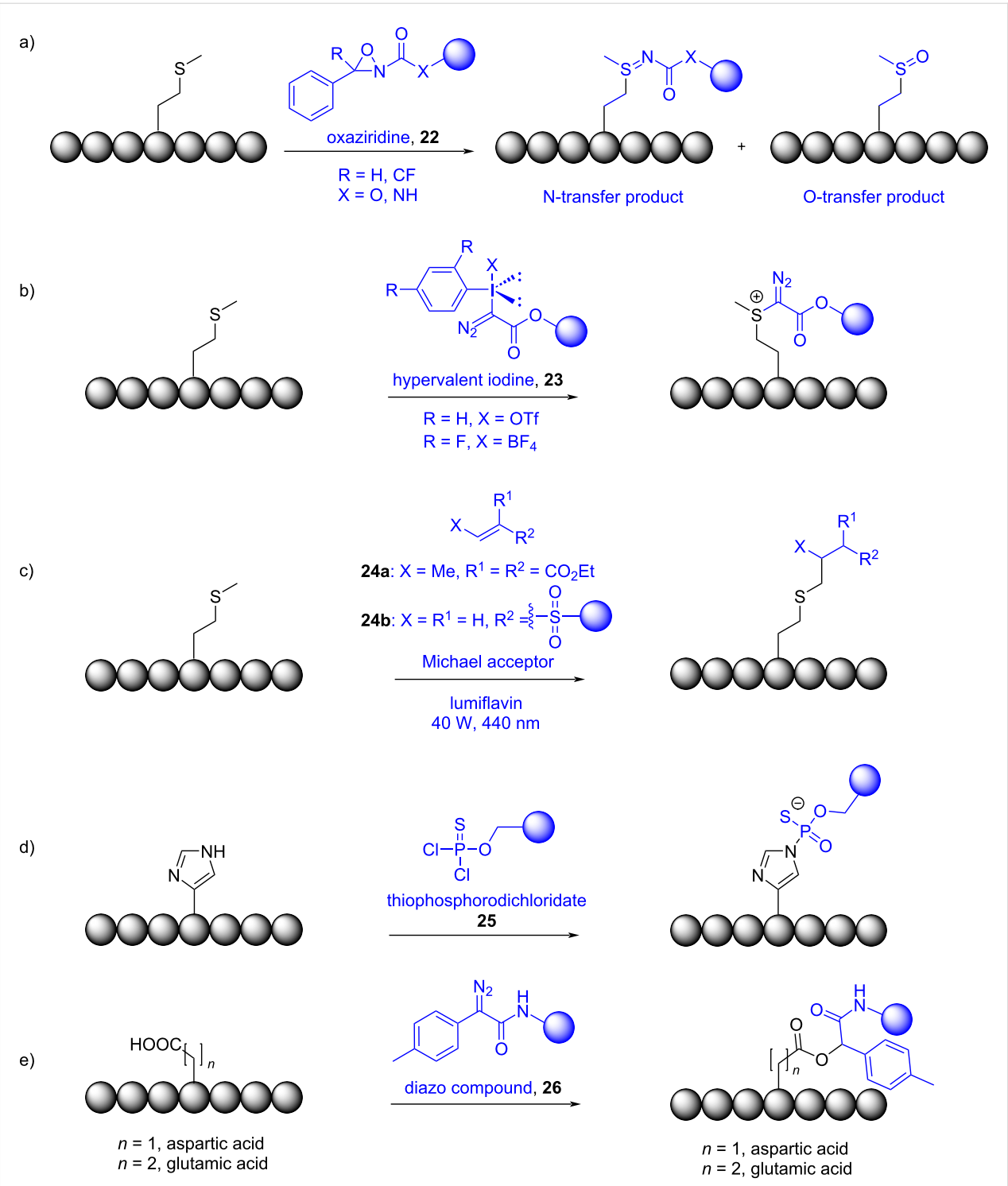
Small molecule residue modifications

Small-molecule reagents are generally user friendly and scalable, with many already commercially available. Their small size makes them atom-economical but achieving site-specificity remains challenging since identical residues are often distributed across protein surfaces. The first example was Sanger's reagent (1-fluoro-2,4-dinitrobenzene, **1**) mentioned above [6]. Other early examples include cysteine modification of ribonuclease A (RNase A) using iodoacetic acid [78]. Early enzyme engineering also explored single-atom modifications, converting subtilisin's serine residue into a cysteine analogue through a three-step chemical sequence (sulfonylation → thioacetate substitution → hydrolysis; Scheme 8) [79], laying the groundwork for ligase development [80,81]. Lysine residues are abundant and typically solvent-exposed [82], hyper-reactive variants have been reported, and have led to the development of tools such as "K-lock" [83-85]. Unfortunately, the reactivity difference is often not sufficient for site-specific modifications. Consequently, recent effort has shifted towards targeting alternative less-frequent amino acid residues [82], as well as novel mechanisms such as linchpin-directed modifications summarized below.

Various chemoselective methods for alternative residues have recently emerged. Methionine has gained traction as a labelling site due to its electron-rich thioether sulfur atom, which possesses versatile bond-forming capacity. The redox-activated chemical tagging (ReACT) platform employs oxaziridines **22** to form stable sulfimide (S=N) linkages, with applications in ADC development, radiolabelling and ROS/Ca^{2+} imaging (Scheme 9a) [86-91]. Methionine can also be modified via sulfonium conjugates using hypervalent iodine (**23**, Scheme 9b) [92] or photoredox approaches (**24a**/**b**, Scheme 9c) [93]. Histidine has been targeted using thiophosphorodichloridate reagents **25**, as shown with hexahistidine-tagged proteins (Scheme 9d) [94]. Aspartate and glutamate, while weakly nucleophilic, can undergo *O*-alkylation via diazo compounds **26** (Scheme 9e) [95,96]. Arginine, traditionally difficult to modify due to its high pK_a values, can be targeted with dielectrophilic glyoxal-based amidine formation (**27**, Scheme 10a) [97-100], aldehyde generation via phenanthrenequinones **28** (Scheme 10b) [101], diketopinic acid (DKPA, **29**) conjugation (Scheme 10c) [102], and methylglyoxal (**30**)-mediated arginine–lysine cyclisation (Scheme 10d) [103].

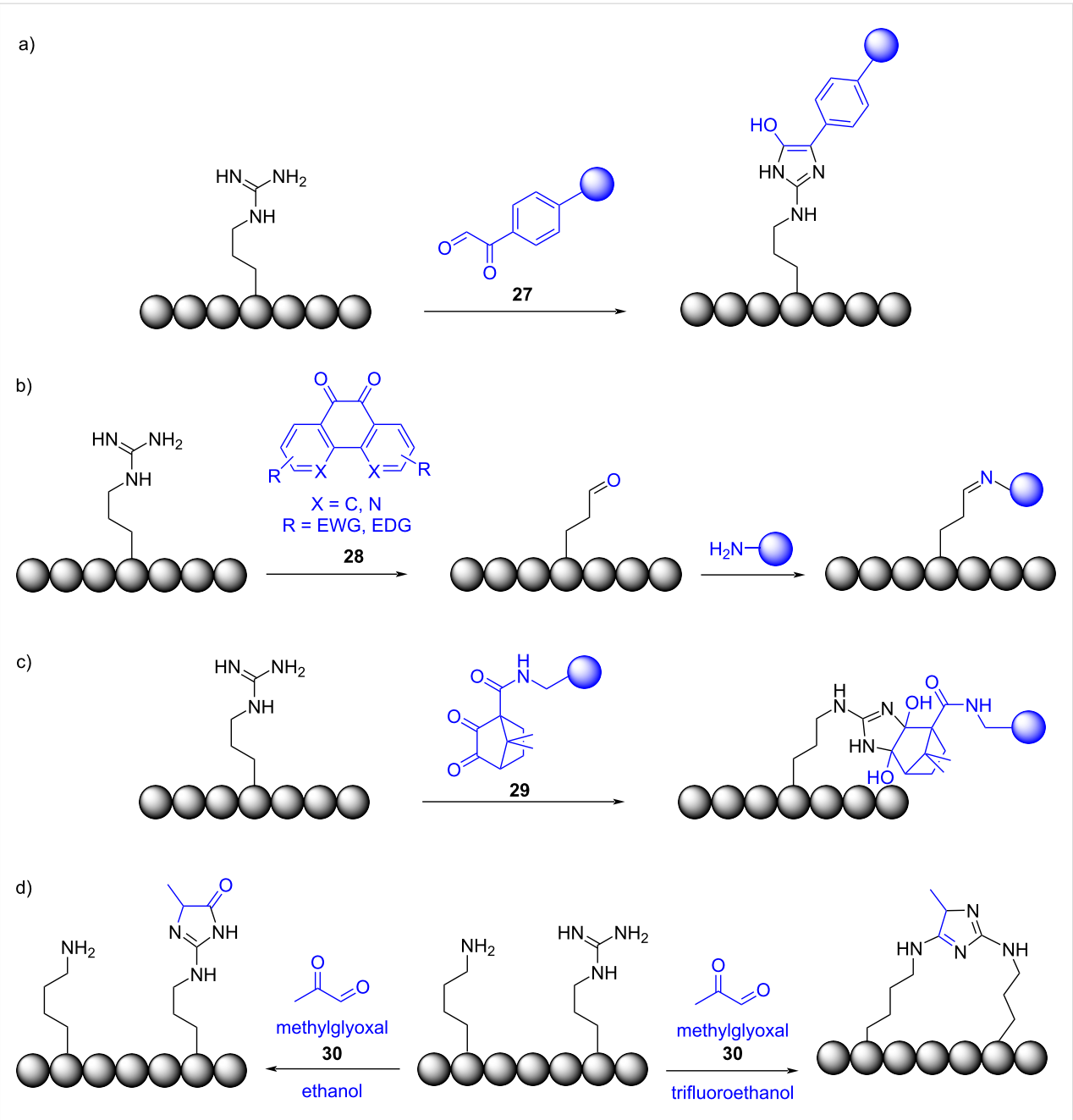


Scheme 8: Chemical sequence to convert subtilisin serine residue into a cysteine analogue.



Scheme 9: Examples of small molecule reagents. Modifications of a–c) methionine by oxaziridine **22**, hypervalent iodine **23**, Michael acceptors (**24a**/**b**), d) histidine by thiophosphorodichoridate **25,** e) aspartic and glutamic acid by diazo-containing compounds **26**.

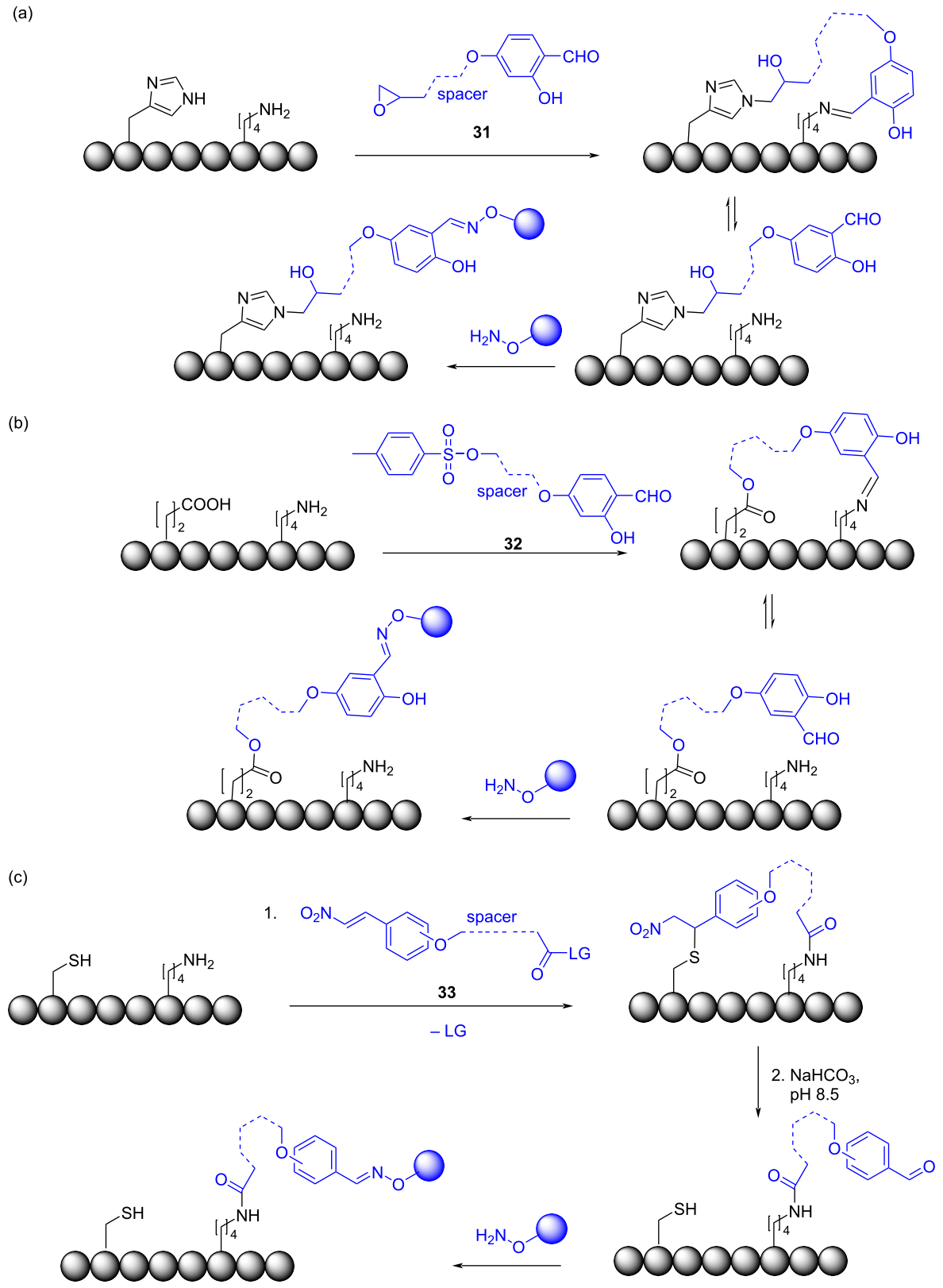
A different approach has been the development of linchpin-directed modification (LDM) [104]. These bifunctional reagents contain two reactive groups separated by a spacer: one engages in a reversible interaction with a common surface residue, while the other is an irreversible chemical handle that covalently modifies a nearby target residue such as histidine (Scheme 11a) [105], aspartate (Scheme 11b) [106] and lysine (Scheme 11c) [107]. For example, reagents containing *o*-hydroxybenzalde-



Scheme 10: Labelling of arginine by reagents carrying a) glyoxal **27**, b) phenanthrenequinones **28**, c) DKPA **29** groups and d) methylglyoxal (**30**).

hyde (**31**, **32**) and β-nitroalkene (**33**) can reversibly engage lysine and cysteine residues, respectively, while a proximal electrophile, such as an epoxide, sulfonate ester or ester, enables targeted covalent modification of amino acids [105-107]. In addition to the choice of electrophiles, selectivity can be refined by tuning the length and flexibility of the linker, which controls the spatial proximity of the reactive groups. While simple, the LDM approach, with appropriate use of linker, could label exclusively cytochrome C out of a mixture of proteins (ubiquitin, RNaseA, myoglobin, ubiquitin, lactalbumin, lysozyme C, α-chymotrypsinogen A). This approach can be extended to native-sequence labelling of Trastuzumab, enabling site-specific histidine modification of the Fc region, albeit at a modest yield (37%). Another related and noteworthy technology is the cysteine-to-lysine transfer reaction, which operates via a mechanism analogous to native chemical ligation in Fab fragments [108].

Together the advances of small-molecule modifications provide powerful strategies for site-selective modification of otherwise



Scheme 11: Examples of linchpin-directed modification targeting a) histidine, b) aspartic acid, and c) lysine residues via transient interaction with neighbouring residues, including lysine-reacting *o*-hydroxybenzaldehyde reagents **31**–**32** and cysteine-reacting β-nitroalkene reagent **33**.

intractable residues, as summarised in a comprehensive review [109]. However, most new conjugation reagents are not yet commercially available, requiring in-house synthesis, and the resulting conjugates often exhibit uncertain and incompletely characterised stability. Moreover, depending on the conjugated epitope, modifications can cause loss of function or aggregation, underscoring the need for further innovation.

Post-translational modification (PTM) chemical labelling

Post-translational modifications (PTMs) refer to chemical changes to a protein in-cell after it is synthesised by the ribosome, encompassing over 650 types of transformations, including phosphorylation, methylation, acylation, ubiquitination and glycosylation [110]. While numerous naturally occurring PTMs have been repurposed for bioconjugation chemistry [111-113], manipulating the composition of native protein glycans, known as glycoengineering, has become a privileged strategy for site-specific modification without altering the protein's native amino-acid sequence. This approach requires the target protein to be *N*- or *O*-glycosylated (Scheme 12a). *N*-Linked glycans, prominently present on antibodies, have been extensively exploited for antibody–drug conjugate (ADC) development, while *O*-linked glycans remain less explored but represent an emerging opportunity.

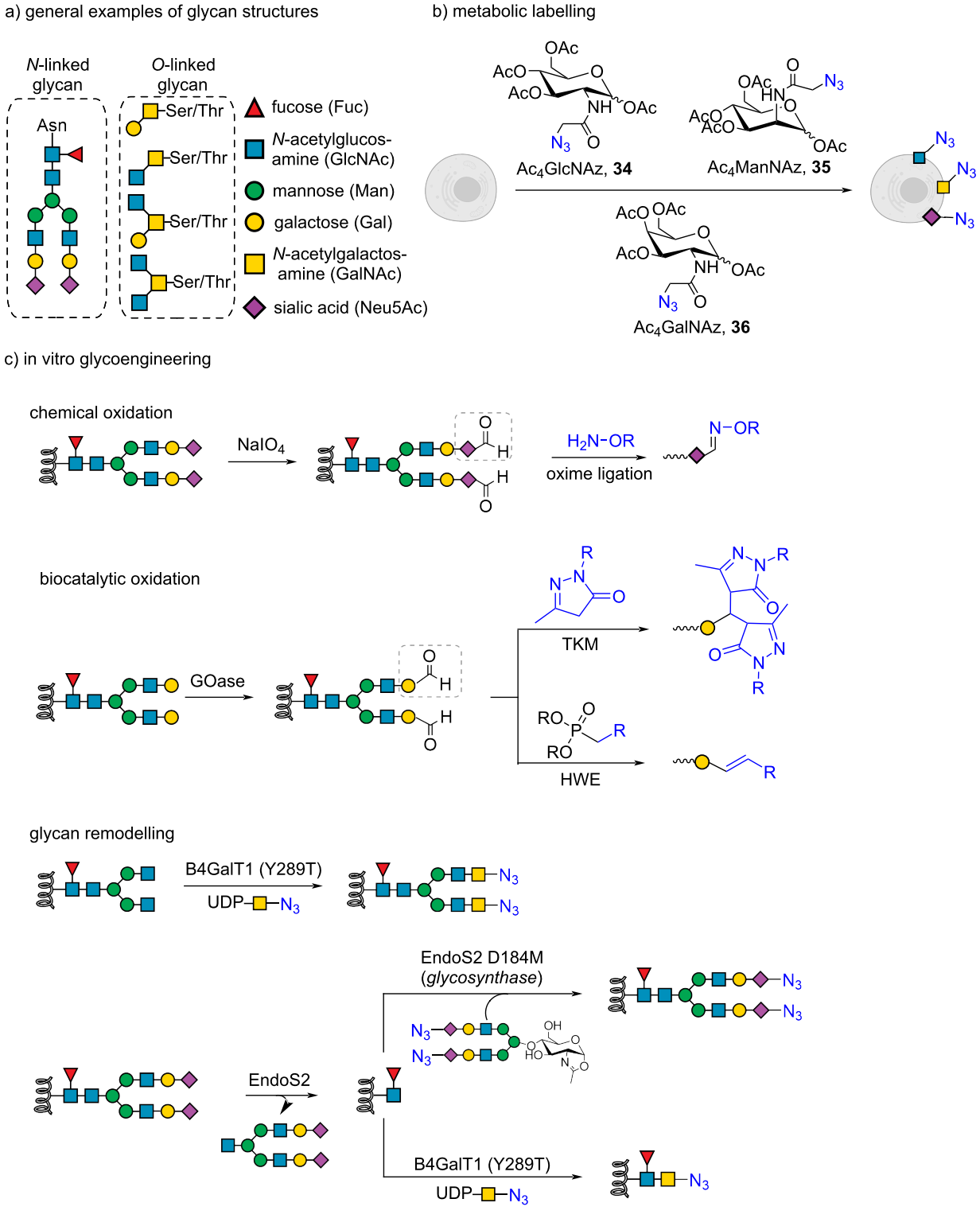
Glycan remodelling strategies are typically dissected into two main branches: metabolic and in vitro glycoengineering. Metabolic glycoengineering introduces a non-natural sugar precursor, such as Ac_4GlcNAz (**34**), Ac_4ManNAz (**35**) and Ac_4GalNAz (**36**, Scheme 12b) into living cells. These analogues enter the cellular glycosylation pathway, thereby incorporating azide motifs into native glycans for chemical attachment through bioorthogonal chemistry (e.g. strain-promoted azide–alkyne cycloaddition/SPAAC) [114-119]. While this strategy enables visualisation and derivatisation of native glycoproteins in their natural environment, metabolic labelling rarely achieves complete functionalisation due to competition from natural substrates, limiting its utility for recombinant protein production. GALaXy technology is a notable exception [120,121]; however, as the technology requires a tag sequence, it is outside the scope of this review.

In vitro glycoengineering, a technology that uses engineered glycoenzymes such as glycosyltransferases and glycosidases to modify glycans, has been widely adopted for native protein labelling, with particular emphasis on antibodies (Scheme 12c). This section aims to give an overview of general strategies adopted for in vitro glycoengineering, and the reader is directed to specialised reviews for in-depth discussions [122-124]. Periodate oxidation of the non-reducing end of *N*-glycans is perhaps the oldest and most straightforward approach to generate an aldehyde biorthogonal handle on a glycoprotein, which can then be subjected to conjugation such as hydrazone or oxime ligation [125-128]. However, while this approach is still a valid strategy for protein functionalisation, potential undesired side reactions such as amino acid oxidation need to be carefully taken into consideration. Use of galactose oxidase (GOase) to site-specifically oxidise galactose to galactoaldehyde on the non-reducing end of a glycan represents a mild alternative, as showcased by its combination with Horner–Wadsworth–Emmons (HWE) chemistry in a one-pot oxidation olefination strategy [19], as well as its implementation with tandem Knoevenagel–Michael addition (TKM) conjugation to enable production of glycoengineered high DAR ADCs [129].

Qasba's engineered bovine galactosyltransferase (β4GalT1(Y289L)) is an elegant example of using glycosyltransferases to remodel glycans of target protein with non-natural sugars (e.g., GalNAz), leading to its wide adoption in academic glycan remodelling research and the development of commercial kits [130-132]. Based on structural data, Tyr289 of bovine β4GalT1 was found to be important for determining UDP-Gal specificity and its mutation to a less hindered leucine residue broadens the scope of azido-carrying nucleotide sugar analogues (UDP-GalNAz), thereby introducing the bioorthogonal handle on the glycan [131,132]. Similarly, sialyltransferases, featuring remarkable substrate promiscuity, have also been exploited to install non-natural sugars on glycans [115-119].

Glycosynthases are engineered variants of glycosidases widely used for glycan remodelling. By substituting the catalytic nucleophile with a non-nucleophilic residue, the engineered enzyme can catalyse glycosidic bond formation with abolished hydrolase activity [133-137]. One such example is the mutation of endoglycosidase S2 (EndoS2) catalytic Asp184 residue into Met (EndoS2 D184M), generating a variant capable of transferring a pre-assembled azido-glycan oxazoline donor substrate onto an antibody with high homogeneity [137]. However, while the elegance of this strategy is undeniable, synthesis of glycan oxazolines remains challenging thus limiting its wider adoption. In an alternative strategy, wild-type EndoS2 can be used for antibody glycan trimming, before the aforementioned β4GalT1(Y289L) can be used to mediate GalNAz transfer generating homogenised *N*-glycan, a patented technology called (GlycoConnect®) used in ADC manufacturing [138,139].

In summary, glycan-directed labelling excels by omitting engineering or denaturation of the protein substrates. Often, the protein function is mostly preserved, for example, the antigen-



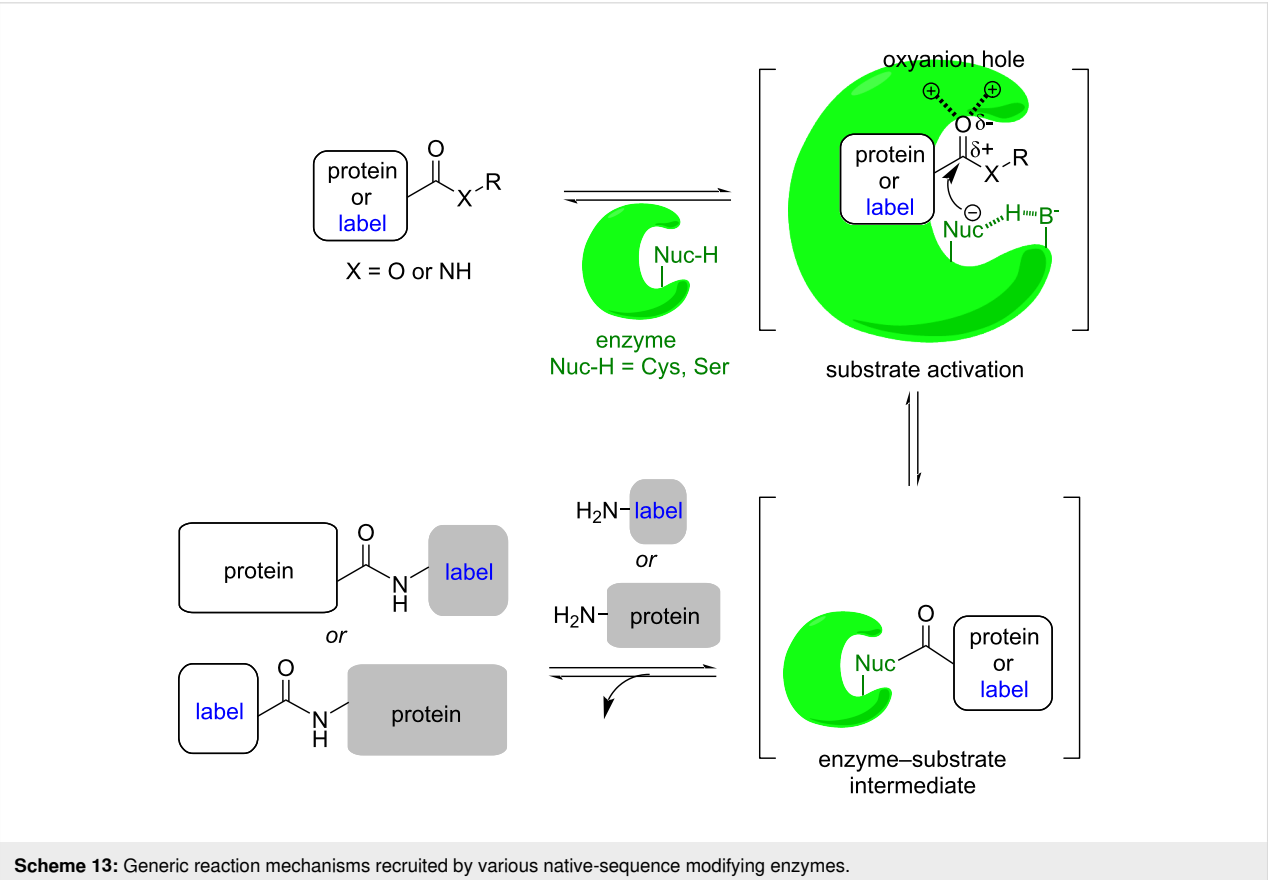
Scheme 12: Overview of glycan-directed strategies for site-specific protein modification. a) Structural motifs of *N*-linked and *O*-linked glycans commonly targeted in bioconjugation; b) metabolic glycoengineering introduces azido-functionalised sugars such as Ac_4GlcNAz (**34**), Ac_4ManNAz (**35**) and Ac_4GalNAz (**36**) into cellular glycosylation pathways for bioorthogonal labelling; c) in vitro glycoengineering employs chemical (e.g., periodate) and/or enzymatic remodelling including the use of engineered oxidases (galactose oxidase), glycosyltransferases (β4GalT1 Y289L) or glycosynthases to install reactive handles on native glycans, enabling precise conjugation for applications such as antibody–drug conjugates (ADCs).

binding (Fab) region of antibodies is free of glycan modification. The chemical or enzymatic steps are generally mild (room temperature, aqueous buffer) yet highly specific. However, the main limitation is that only glycosylated proteins (especially IgGs) can be targeted. Intrinsic glycan heterogeneity can also lead to incomplete functionalisation. Biological activity of the modified protein, recyclability and manufacturing cost of the enzymes also need to be carefully considered, as costs can easily balloon in process development. Nonetheless, the use of glycan remodelling is expanding for applications such as site-specific protein labelling, with ADC development exemplifying its significant potential.

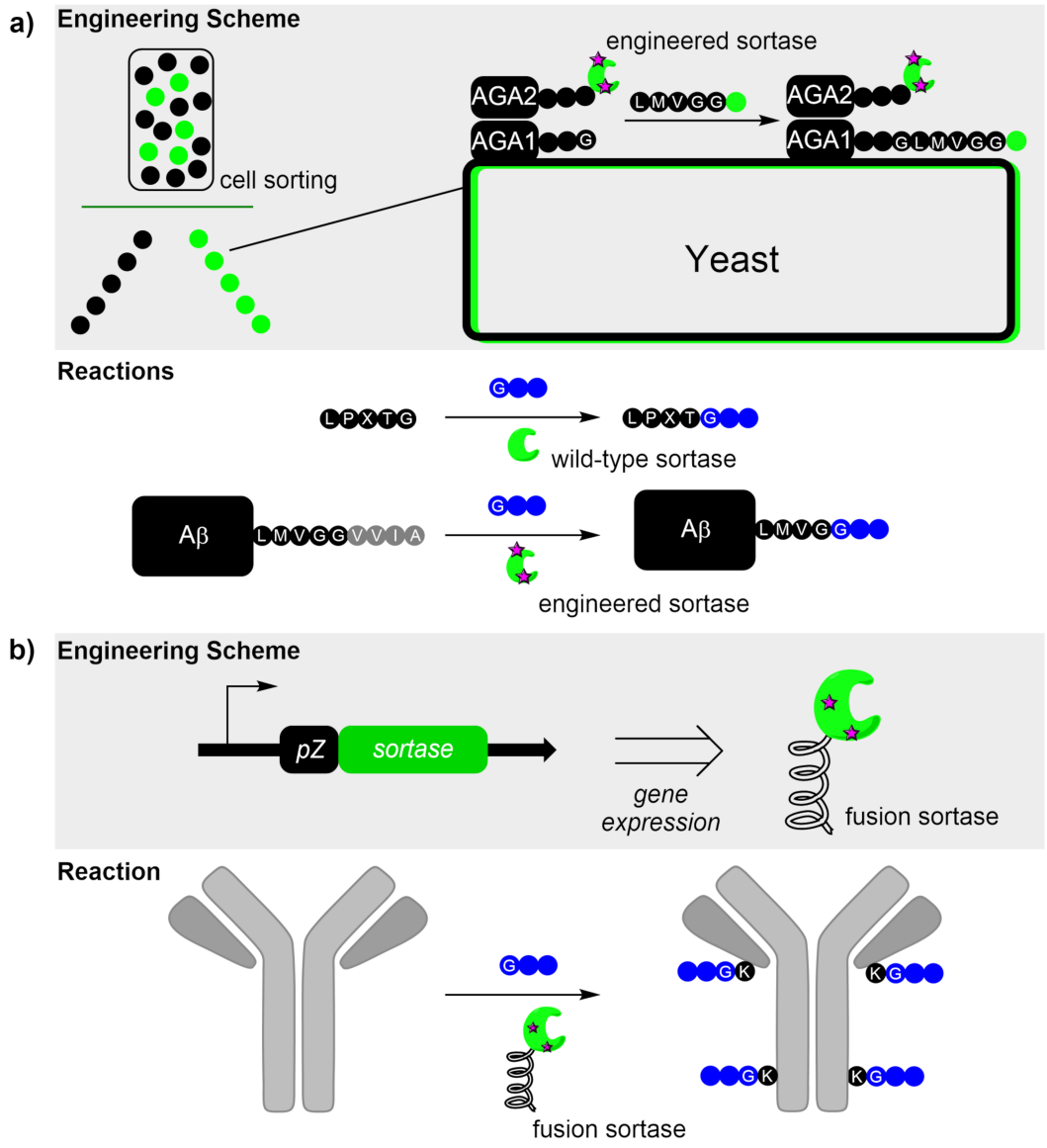
Enzyme-mediated modifications

Enzymes are highly attractive tools for native-sequence protein modification, offering exquisite spatial precision under mild, biocompatible conditions and easy access via recombinant technologies. Their catalytic activity supports scalability, making them suitable for both academic and industrial applications. Moreover, enzymes enable a broad spectrum of chemistries, targeting both amino acid sidechains and the protein backbone [140-142]. Despite these advantages, the repertoire of enzymes applied for native-sequence modifications remains limited. The following examples illustrate recent advances in recruiting enzymes for native-sequence labelling, all of which rely on reaction mechanisms that activate carbonyl groups (Scheme 13).

One of the most studied protein-modifying enzymes is sortase A. Recognising the LPXTG motif, the enzyme cleaves the Thr–Gly peptide bond to form an acyl intermediate that is ligated to an incoming primary amine nucleophile, such as an *N*-terminal glycine or lysine ε-amine (Scheme 14a) [143]. Native proteins with an *N*-terminal glycine and sufficient flexibility can thus be directly modified by sortase. Optimisations using depsipeptides or coupling strategies have further improved conversion efficiency and reduced reagent consumption [144]. Due to its ease of preparation, large libraries of sortase variants have been generated and screened to expand substrate scope [143]. In one example, sortase variants and *N*-terminal polyglycine tags were fused to yeast surface proteins AGA-1 and -2, respectively, while peptides carrying GG alternative motifs and reporter tags were introduced as substrates. Active variants covalently attached the reporter to the yeast surface, allowing isolation via FACS (Scheme 14a). After 16 rounds of selection from a library of 10^7 variants, enzymes capable of recognising LMVGG, a sequence absent from wild-type specificity but present in endogenous amyloid protein Aβ in human cerebrospinal fluid, were obtained [145].



Scheme 13: Generic reaction mechanisms recruited by various native-sequence modifying enzymes.



Scheme 14: a,b) Engineered biocatalysts including sortase. Proteins AGA1 and AGA2 are surface proteins naturally found on yeast surface.

An affinity-guided strategy was used to evolve sortase A variants for antibody conjugation [146]. Exploiting the ability of sortase to accept the ε-amino group of lysine, the enzyme was fused to an antibody Fc-binding domain protein (protein Z/pZ), positioning it near surface-exposed lysine on the target antibody for isopeptide bond modification (Scheme 14b). These fusion constructs, developed based on the concept of affinity-guided labelling, were immediately active and served as starting points for further engineering. After only one to three rounds of mutagenesis via colony screening with glycosylated trastuzumab and cetuximab, variants 1M (N127K) and 3M (T156S/D170E/D176E) with enhanced labelling activity were isolated. Although this system offers switchable activity by removing calcium ions from the buffer, it achieves only moderate site-specificity; labelling was found in K5, K123, K135, K292, and K441 on the heavy chain and K207 on the light chain of cetuximab, out of ≈80 lysine residues. However, antigen binding was largely unaffected, and a range of payloads including TAMRA, biotin and the cytotoxic drug MMAE could be incorporated to yield a DAR ratio of 2–3 [146].

Microbial transglutaminase (MTG), which forms isopeptide bonds between a target protein and glutamine (or other small primary amides) is another frequently used enzyme for biocon-

jugation. The primary amide in glutamine is converted into an acyl intermediate that is subsequently attacked by amine nucleophiles such as lysine ε-amines and glycine α-amines. Although MTG can modify Gln295 on the antibody Fc, this residue is close to the Asn297 glycan and hence normally requires deglycosylation or N297A mutation for enzyme access, both of which reduce structural stability and diminish antibody-dependent cell-mediated cytotoxicity (ADCC) [147,148]. To enable labelling of fully native antibodies, two complementary strategies have been developed (Scheme 15a). First, rational engineering of MTG yielded active-site mutants, with modifications located in a loop projecting into the catalytic pocket (G250S), which enhanced reactivity towards native antibody Asn297 [149]. Further optimisation delivered double variants G250S/E300D ("S2") and G250S/K327T ("S13"), both capable of generating ADCs with DAR close to 2.0. Second, substrate engineering revealed that MTG preferentially processes positively charged peptides, with an RAK-containing motif supporting efficient ligation to fully glycosylated antibodies [150]. Collectively, these enzyme- or substrate-engineering approaches enabled native-state labelling of multiple antibodies (avelumab, matuzumab, cetuximab, and rituximab), producing stable conjugates with favourable biodistribution [149,150].

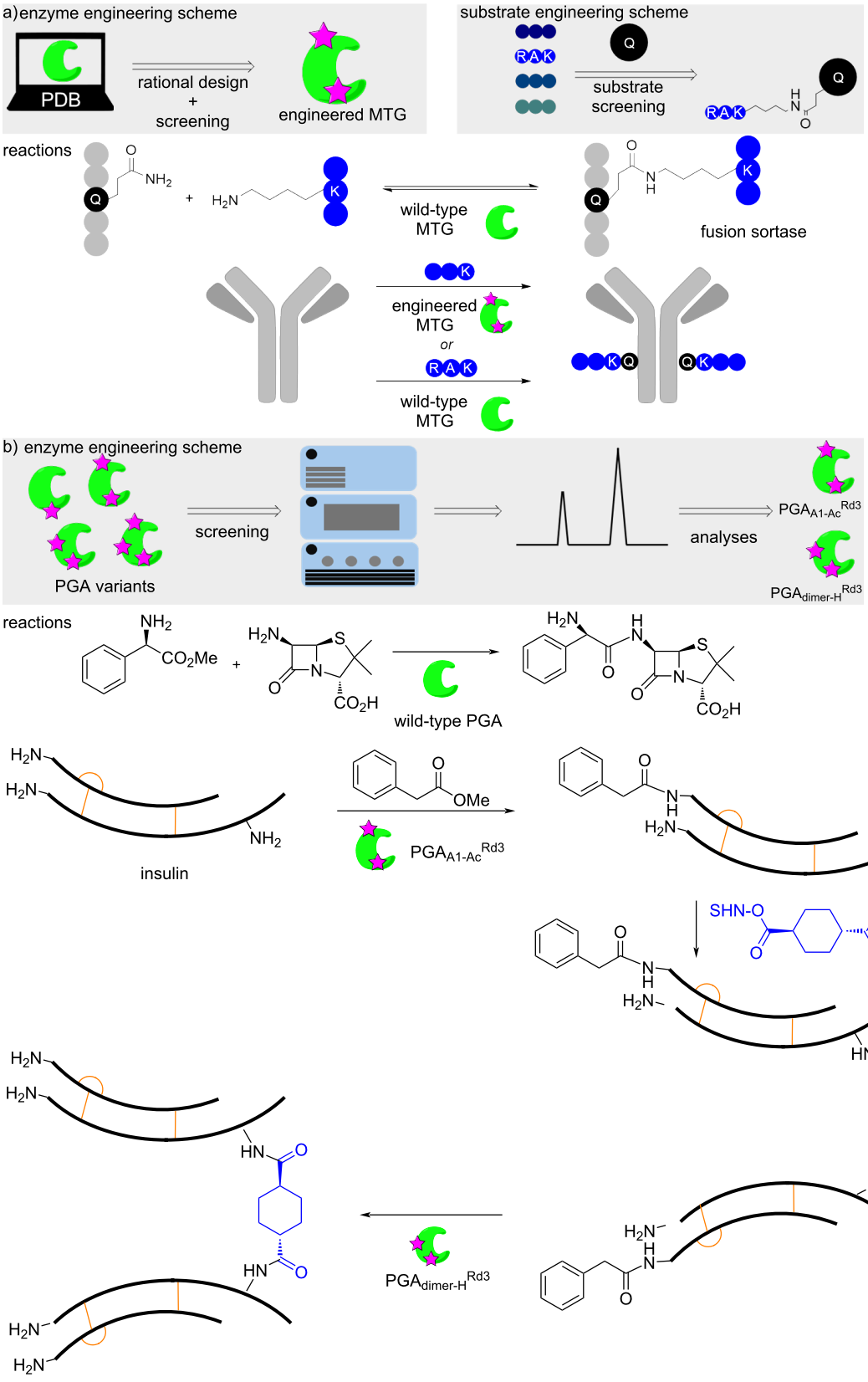
Another example is the engineering of the hydrolase penicillin G acylase (PGA) for the selective modification of insulin [151]. After screening 100 commercially available PGA homologs and initial optimisation for expression and handling, *Kluyvera cryocrescens* PGA was chosen for directed evolution (Scheme 15b). The goal was to selectively protect or deprotect key amines on insulin, including the *N*-terminal glycine of chain A, the *N*-terminal phenylalanine of chain B and the ε-amino group of lysine B29. Using a colony picker and UPLC analysis, mutations were introduced within the binding pocket (to enhance localised recognition of insulin) and at distal sites (to improve stability and solvent tolerance). The engineered enzyme $PGA_{A1\text{-}Ac}^{Rd3}$ selectively introduced a benzoyl group at the *N*-terminus of chain A, whereas $PGA_{dimer\text{-}H}^{Rd3}$ catalysed deprotection enabling formation of dimerised insulin candidates under mild chemoenzymatic conditions. Chromatography-free precipitation gave yields up to >90%, and scalability was demonstrated at >120 grams.

These enzyme-based examples are elegant but not without limitations, largely due to the labour-intensive workflows and the intrinsic design constraints associated with the use of naturally derived enzymes. The sortase variant engineered to recognise amyloid proteins was highly selective for the altered sequence (K_M = 128 μM) but achieved only modest catalytic efficiency (143 M^{-1} s^{-1}, compared with 10^3–10^4 M^{-1} s^{-1} in wild-type), likely because the FACS-based screening was based on single-turnover events. The sortase A-protein Z system has limited site specificity; it also requires equimolar enzyme-to-antibody ratios and additional purification steps. MTG, similarly, requires near-stoichiometric enzyme to function, restricting scalability; this enzyme is also produced as a zymogen and requires activation post-purification, complicating the engineering workflow. PGA optimisation benefited from the simplicity of insulin, which contains only three primary amines whose labelling can be readily resolved by UPLC, but such strategies become impractical for larger, lysine-rich proteins. For a broader overview, Zhu et al. [141] provide a comprehensive summary of recent advances in engineering amide-bond-forming enzymes for modifications, including asparaginyl endopeptidase (AEP), omniligase and other emerging biocatalysts.

Comparative Analysis, Outlook and Future Directions

In Table 1, we compile a bird's-eye overview of current site-specific native-sequence protein modification strategies. The check marks represent best-case scenarios reported in the literature. Native chemical ligation offers high precision for *N*-terminal cysteine modifications and synthesis of diverse proteins, whereas disulfide-rebridging approaches provide scalability due to the relative ease of reagent preparation (see above for detailed discussion). Overall, however, the success of any given method is often case-dependent, reflecting the diversity of protein and peptide sequences and structures. Factors such as protein preparation can also influence labelling efficiency; this is often the case for *N*-terminal and glycan modifications where composition of the protein may vary with the recombinant host. Protein precipitation and unintended side reactions, which are often underreported, can further complicate the anticipated outcome. The limitations listed in Table 1 reflect intrinsic constraints of the tool design, such as the requirement for accessible cysteine or disulfide motifs in the corresponding technology. For modifications by naturally derived enzymes, a balance often needs to be struck between substrate promiscuity and selectivity. In general, highly substrate-promiscuous enzymes require more complex preparation procedures, which can hinder their engineering, whereas highly selective enzymes are easier to handle but depend on defined recognition sequences. Ultimately, we found that there is no single strategy simultaneously fulfilling all criteria, and trade-off(s) are inevitable. Hence, we hope this overview will help readers identify opportunities for developing new strategies that address these current limitations.

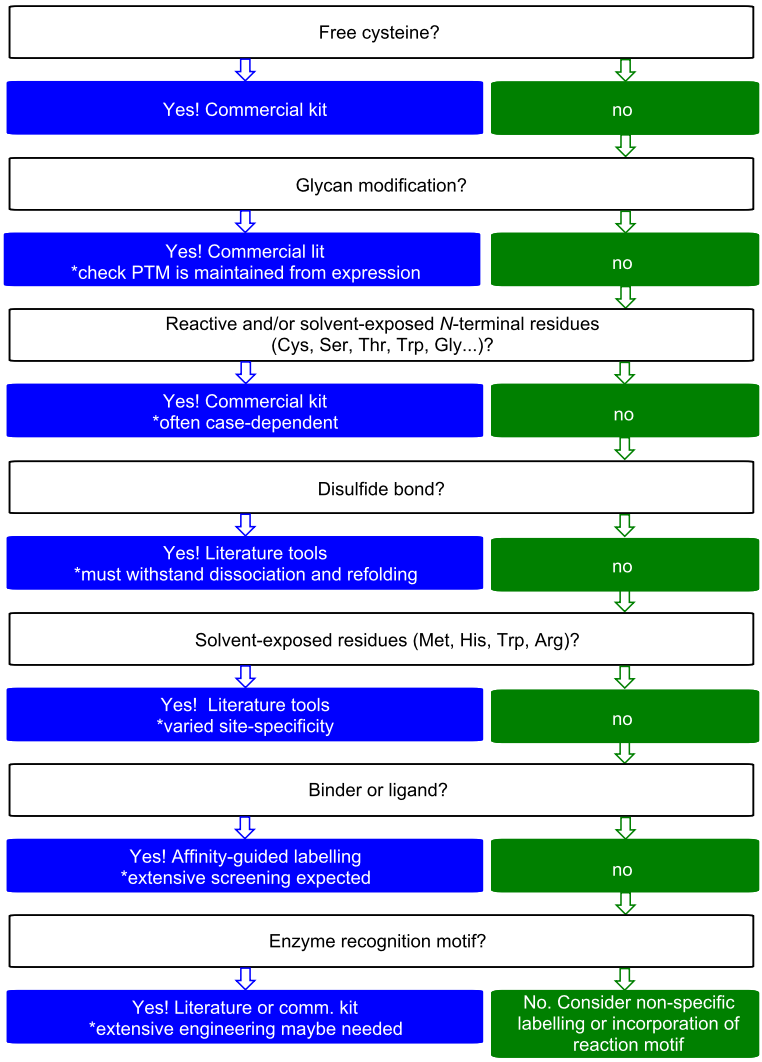
For readers seeking to adapt existing technologies for protein modifications, we present a generic decision framework for selecting a site-specific native-sequence protein modification strategy (Scheme 16). The first step is to evaluate cysteine content, as these residues are the most established targets with a



Scheme 15: Engineered biocatalysts include a) microbial transglutaminase (MTG) and b) penicillin G acylase (PGA).

Table 1: Comparison of existing site-specific, native-sequence protein modification strategies.

	Cysteine modification	*N*-Terminal labelling	Affinity-guided chemistry	Disulfide-rebridging technology	Small molecule	Glycan modification	Existing enzymes
precision	✓	✓	✓	✓	limited positional control, though often residue specific	✓	✓
scalability	✓	✓	complex reagents used in stoichiometric portion	✓	✓	✓	may require stoichiometric portion
usability	only for low-abundance cysteine	highly sequence dependent	✓	only for low abundance disulfide motif	✓	glycan must be present	may require recognition motif may require extensive engineering



Scheme 16: Selection scheme of a suitable native-sequence labelling tool.

wide range of stable, commercially available reagents. Similarly, if the target protein is glycosylated, numerous glycoengineering strategies and commercial tools can be considered. In the absence of free cysteines or glycans, the *N*-terminus can be chemically targeted, particularly when it contains dinucleophilic residues such as cysteine, serine, threonine, or tryptophan. Enzymes such as sortase (for glycine-containing motifs), omniligase and AEP can be also considered as tools when the *N*-terminal modification site is sufficiently solvent exposed. Alternatively, disulfide bonds may provide modification sites through rebridging, provided that the protein tolerates reduction-and-rebridging reaction sequences. If these options are unsuitable, other solvent-exposed nucleophilic residues (e.g., methionine, histidine, tryptophan, tyrosine or arginine) can be targeted, although such reagents are often prepared in-house and may lack positional control. When high-affinity ligands or binders are available, affinity-guided conjugation strategies can be implemented, though probe design and optimisation are often required. Finally, if an enzyme recognition motif is available, biocatalytic approaches can be pursued, particularly for commercially accessible enzymes, although at times they require in-house screening and/or engineering to achieve the desired specificity. As a general guideline, the final choice depends on multiple factors, including protein properties (sequence, flexibility, post-translational modification status, stability) and reagents availability (commercial reagents, in-house preparations, or enzyme accessibility).

In summary, a fundamental barrier associated with native-sequence protein modifications arises from the presence of multiple identical residues within a protein, each complicated by its specific structural context. Moreover, all chemistries must operate under biocompatible conditions, restricting the scope of viable options. Nevertheless, continued innovations at the interface of synthetic chemistry and biology are expected to address these barriers, enabling the development of broadly adaptable platforms for the efficient production of modified proteins.

Funding

We would like to thank for the financial support provided by BBSRC (BB/T015799/1), Leverhulme Trust (RPG-2017-195) and Royal Society (RG170187) for Louis Y. P. Luk.

Author Contributions

Antonio Angelastro: conceptualization; formal analysis; investigation; writing – original draft; writing – review & editing. Jonathan Bargh: conceptualization; formal analysis; writing – original draft; writing – review & editing. Subhajit Guria: conceptualization; investigation; writing – original draft; writing – review & editing. Victor Laserna: conceptualization; formal analysis; writing – original draft; writing – review & editing. Louis Luk: conceptualization; investigation; writing – original draft; writing – review & editing.

ORCID® iDs

Antonio Angelastro - https://orcid.org/0000-0002-4023-7411
Jonathan Bargh - https://orcid.org/0000-0003-3023-2569
Subhajit Guria - https://orcid.org/0000-0002-5754-6157
Victor Laserna - https://orcid.org/0000-0003-2951-9327
Louis Luk - https://orcid.org/0000-0002-7864-6261

Data Availability Statement

Data sharing is not applicable as no new data was generated or analyzed in this study.

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint:
https://chemrxiv.org/engage/chemrxiv/article-details/694bdef36e8e74030e67d865

References

1. Journeaux, T.; Bernardes, G. J. L. *Nat. Chem.* **2024,** *16,* 854–870. doi:10.1038/s41557-024-01507-y
2. Gu, Y.; Wang, Z.; Wang, Y. *Acta Pharm. Sin. B* **2024,** *14,* 1965–1986. doi:10.1016/j.apsb.2024.01.009
3. Chen, F.-J.; Gao, J. *Chem. – Eur. J.* **2022,** *28,* e202201843. doi:10.1002/chem.202201843
4. Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. *Chem. – Asian J.* **2009,** *4,* 630–640. doi:10.1002/asia.200800427
5. Jacob, E.; Unger, R. *Bioinformatics* **2007,** *23,* e225–e230. doi:10.1093/bioinformatics/btl318
6. Sanger, F. *Biochem. J.* **1945,** *39,* 507–515. doi:10.1042/bj0390507
7. Sereda, T. J.; Mant, C. T.; Quinn, A. M.; Hodges, R. S. *J. Chromatogr. A* **1993,** *646,* 17–30. doi:10.1016/s0021-9673(99)87003-4
8. Lindsley, C. W. *ACS Chem. Neurosci.* **2015,** *6,* 505–506. doi:10.1021/acschemneuro.5b00098
9. Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. *Chem. Rev.* **2019,** *119,* 7328–7443. doi:10.1021/acs.chemrev.8b00712
10. Cui, L.; Rao, J. *Methods Mol. Biol. (N. Y., NY, U. S.)* **2015,** *1266,* 81–92. doi:10.1007/978-1-4939-2272-7_5
11. Bandyopadhyay, A.; Cambray, S.; Gao, J. *Chem. Sci.* **2016,** *7,* 4589–4593. doi:10.1039/c6sc00172f
12. Faustino, H.; Silva, M. J. S. A.; Veiros, L. F.; Bernardes, G. J. L.; Gois, P. M. P. *Chem. Sci.* **2016,** *7,* 6280. doi:10.1039/c6sc90045c
13. Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.* **1992,** *3,* 138–146. doi:10.1021/bc00014a008
14. Witus, L. S.; Netirojjanakul, C.; Palla, K. S.; Muehl, E. M.; Weng, C.-H.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2013,** *135,* 17223–17229. doi:10.1021/ja408868a
15. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed.* **2006,** *45,* 5307–5311. doi:10.1002/anie.200600368

16. Agarwal, P.; van der Weijden, J.; Sletten, E. M.; Rabuka, D.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 46–51. doi:10.1073/pnas.1213186110
17. Agarwal, P.; Kudirka, R.; Albers, A. E.; Barfield, R. M.; de Hart, G. W.; Drake, P. M.; Jones, L. C.; Rabuka, D. *Bioconjugate Chem.* **2013,** *24,* 846–851. doi:10.1021/bc400042a
18. Pomplun, S.; Mohamed, M. Y. H.; Oelschlaegel, T.; Wellner, C.; Bergmann, F. *Angew. Chem., Int. Ed.* **2019,** *58,* 3542–3547. doi:10.1002/anie.201814200
19. Angelastro, A.; Barkhanskiy, A.; Journeaux, T.; Sivapalan, R.; King, T. A.; Rodríguez Pérez, L.; Goundry, W. R. F.; Barran, P.; Flitsch, S. L. *Nat. Synth.* **2024,** *3,* 976–985. doi:10.1038/s44160-024-00563-z
20. Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 6705–6710. doi:10.1073/pnas.95.12.6705
21. Shah, N. H.; Muir, T. W. *Chem. Sci.* **2014,** *5,* 446–461. doi:10.1039/c3sc52951g
22. Thompson, R. E.; Muir, T. W. *Chem. Rev.* **2020,** *120,* 3051–3126. doi:10.1021/acs.chemrev.9b00450
23. Bloom, S.; Liu, C.; Kölmel, D. K.; Qiao, J. X.; Zhang, Y.; Poss, M. A.; Ewing, W. R.; MacMillan, D. W. C. *Nat. Chem.* **2018,** *10,* 205–211. doi:10.1038/nchem.2888
24. Zhang, L.; Floyd, B. M.; Chilamari, M.; Mapes, J.; Swaminathan, J.; Bloom, S.; Marcotte, E. M.; Anslyn, E. V. *ACS Chem. Biol.* **2021,** *16,* 2595–2603. doi:10.1021/acschembio.1c00631
25. Rosen, C. B.; Francis, M. B. *Nat. Chem. Biol.* **2017,** *13,* 697–705. doi:10.1038/nchembio.2416
26. Spicer, C. D.; Davis, B. G. *Nat. Commun.* **2014,** *5,* 4740. doi:10.1038/ncomms5740
27. del Rosario, R. B.; Wahl, R. L.; Brocchini, S. J.; Lawton, R. G.; Smith, R. H. *Bioconjugate Chem.* **1990,** *1,* 51–59. doi:10.1021/bc00001a006
28. Badescu, G.; Bryant, P.; Bird, M.; Henseleit, K.; Swierkosz, J.; Parekh, V.; Tommasi, R.; Pawlisz, E.; Jurlewicz, K.; Farys, M.; Camper, N.; Sheng, X.; Fisher, M.; Grygorash, R.; Kyle, A.; Abhilash, A.; Frigerio, M.; Edwards, J.; Godwin, A. *Bioconjugate Chem.* **2014,** *25,* 1124–1136. doi:10.1021/bc500148x
29. Assem, N.; Ferreira, D. J.; Wolan, D. W.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2015,** *54,* 8665–8668. doi:10.1002/anie.201502607
30. Hu, Q. Y.; Imase, H. Methods for making conjugates from disulfide-containing proteins. Int. Pat. Appl. WO/2014/083505 A1, June 5, 2014.
31. Smith, M. E. B.; Schumacher, F. F.; Ryan, C. P.; Tedaldi, L. M.; Papaioannou, D.; Waksman, G.; Caddick, S.; Baker, J. R. *J. Am. Chem. Soc.* **2010,** *132,* 1960–1965. doi:10.1021/ja908610s
32. Schumacher, F. F.; Nunes, J. P. M.; Maruani, A.; Chudasama, V.; Smith, M. E. B.; Chester, K. A.; Baker, J. R.; Caddick, S. *Org. Biomol. Chem.* **2014,** *12,* 7261–7269. doi:10.1039/c4ob01550a
33. Chudasama, V.; Smith, M. E. B.; Schumacher, F. F.; Papaioannou, D.; Waksman, G.; Baker, J. R.; Caddick, S. *Chem. Commun.* **2011,** *47,* 8781–8783. doi:10.1039/c1cc12807h
34. Forte, N.; Chudasama, V.; Baker, J. R. *Drug Discovery Today: Technol.* **2018,** *30,* 11–20. doi:10.1016/j.ddtec.2018.09.004
35. Walsh, S. J.; Omarjee, S.; Galloway, W. R. J. D.; Kwan, T. T.-L.; Sore, H. F.; Parker, J. S.; Hyvönen, M.; Carroll, J. S.; Spring, D. R. *Chem. Sci.* **2019,** *10,* 694–700. doi:10.1039/c8sc04645j
36. Koniev, O.; Dovgan, I.; Renoux, B.; Ehkirch, A.; Eberova, J.; Cianférani, S.; Kolodych, S.; Papot, S.; Wagner, A. *MedChemComm* **2018,** *9,* 827–830. doi:10.1039/c8md00141c
37. Stieger, C. E.; Franz, L.; Körlin, F.; Hackenberger, C. P. R. *Angew. Chem., Int. Ed.* **2021,** *60,* 15359–15364. doi:10.1002/anie.202100683
38. Lee, M. T. W.; Maruani, A.; Richards, D. A.; Baker, J. R.; Caddick, S.; Chudasama, V. *Chem. Sci.* **2017,** *8,* 2056–2060. doi:10.1039/c6sc03655d
39. Dannheim, F. M.; Walsh, S. J.; Orozco, C. T.; Hansen, A. H.; Bargh, J. D.; Jackson, S. E.; Bond, N. J.; Parker, J. S.; Carroll, J. S.; Spring, D. R. *Chem. Sci.* **2022,** *13,* 8781–8790. doi:10.1039/d2sc02198f
40. Krajcovicova, S.; Wharton, T.; Driscoll, C. L.; King, T. A.; Howarth, M. R.; Spring, D. R. *Chem. Sci.* **2025,** *16,* 10602–10609. doi:10.1039/d5sc02286j
41. Bahou, C.; Love, E. A.; Leonard, S.; Spears, R. J.; Maruani, A.; Armour, K.; Baker, J. R.; Chudasama, V. *Bioconjugate Chem.* **2019,** *30,* 1048–1054. doi:10.1021/acs.bioconjchem.9b00174
42. Peciak, K.; Laurine, E.; Tommasi, R.; Choi, J.-w.; Brocchini, S. *Chem. Sci.* **2019,** *10,* 427–439. doi:10.1039/c8sc03355b
43. Wu, K.-L.; Yu, C.; Lee, C.; Zuo, C.; Ball, Z. T.; Xiao, H. *Bioconjugate Chem.* **2021,** *32,* 1947–1959. doi:10.1021/acs.bioconjchem.1c00342
44. Walsh, S. J.; Bargh, J. D.; Dannheim, F. M.; Hanby, A. R.; Seki, H.; Counsell, A. J.; Ou, X.; Fowler, E.; Ashman, N.; Takada, Y.; Isidro-Llobet, A.; Parker, J. S.; Carroll, J. S.; Spring, D. R. *Chem. Soc. Rev.* **2021,** *50,* 1305–1353. doi:10.1039/d0cs00310g
45. Bahou, C.; Chudasama, V. *Org. Biomol. Chem.* **2022,** *20,* 5879–5890. doi:10.1039/d2ob00310d
46. Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annu. Rev. Biochem.* **2008,** *77,* 383–414. doi:10.1146/annurev.biochem.75.101304.124125
47. Tang, S.; Beattie, A. T.; Kafkova, L.; Petris, G.; Huguenin-Dezot, N.; Fiedler, M.; Freeman, M.; Chin, J. W. *Nature* **2022,** *602,* 701–707. doi:10.1038/s41586-022-04414-9
48. Sakamoto, S.; Hamachi, I. *Isr. J. Chem.* **2023,** *63,* e202200077. doi:10.1002/ijch.202200077
49. Mortensen, M. R.; Skovsgaard, M. B.; Gothelf, K. V. *ChemBioChem* **2019,** *20,* 2711–2728. doi:10.1002/cbic.201900157
50. Tamura, T.; Ueda, T.; Goto, T.; Tsukidate, T.; Shapira, Y.; Nishikawa, Y.; Fujisawa, A.; Hamachi, I. *Nat. Commun.* **2018,** *9,* 1870. doi:10.1038/s41467-018-04343-0
51. Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. *Nat. Chem. Biol.* **2009,** *5,* 341–343. doi:10.1038/nchembio.157
52. Hui, J. Z.; Tsourkas, A. *Bioconjugate Chem.* **2014,** *25,* 1709–1719. doi:10.1021/bc500305v
53. Kanje, S.; Hober, S. *Biotechnol. J.* **2015,** *10,* 564–574. doi:10.1002/biot.201400808
54. Kanje, S.; von Witting, E.; Chiang, S. C. C.; Bryceson, Y. T.; Hober, S. *Bioconjugate Chem.* **2016,** *27,* 2095–2102. doi:10.1021/acs.bioconjchem.6b00346
55. Hui, J. Z.; Tamsen, S.; Song, Y.; Tsourkas, A. *Bioconjugate Chem.* **2015,** *26,* 1456–1460. doi:10.1021/acs.bioconjchem.5b00275
56. von Witting, E.; Hober, S.; Kanje, S. *Bioconjugate Chem.* **2021,** *32,* 1515–1524. doi:10.1021/acs.bioconjchem.1c00313
57. Hayashi, T.; Yasueda, Y.; Tamura, T.; Takaoka, Y.; Hamachi, I. *J. Am. Chem. Soc.* **2015,** *137,* 5372–5380. doi:10.1021/jacs.5b02867

58. Comeo, E.; Goulding, J.; Lin, C.-Y.; Groenen, M.; Woolard, J.; Kindon, N. D.; Harwood, C. R.; Platt, S.; Briddon, S. J.; Kilpatrick, L. E.; Scammells, P. J.; Hill, S. J.; Kellam, B. *J. Med. Chem.* **2024,** *67,* 12099–12117. doi:10.1021/acs.jmedchem.4c00835
59. Arnould, B.; Martz, A.; Belzanne, P.; Peralta, F. A.; Cevoli, F.; Hovhannisyan, V.; Goumon, Y.; Hosy, E.; Specht, A.; Grutter, T. *eLife* **2026,** *14,* RP106096. doi:10.7554/elife.106096.3
60. Popp, B. V.; Ball, Z. T. *Chem. Sci.* **2011,** *2,* 690–695. doi:10.1039/c0sc00564a
61. Ohata, J.; Ball, Z. T. *J. Am. Chem. Soc.* **2017,** *139,* 12617–12622. doi:10.1021/jacs.7b06428
62. Koshi, Y.; Nakata, E.; Miyagawa, M.; Tsukiji, S.; Ogawa, T.; Hamachi, I. *J. Am. Chem. Soc.* **2008,** *130,* 245–251. doi:10.1021/ja075684q
63. Wang, H.; Koshi, Y.; Minato, D.; Nonaka, H.; Kiyonaka, S.; Mori, Y.; Tsukiji, S.; Hamachi, I. *J. Am. Chem. Soc.* **2011,** *133,* 12220–12228. doi:10.1021/ja204422r
64. Kim, S.; Kim, S.; Kim, S.; Kim, N.; Lee, S. W.; Yi, H.; Lee, S.; Sim, T.; Kwon, Y.; Lee, H. S. *Adv. Sci.* **2024,** *11,* 2306401. doi:10.1002/advs.202306401
65. Yeh, A. H.-W.; Norn, C.; Kipnis, Y.; Tischer, D.; Pellock, S. J.; Evans, D.; Ma, P.; Lee, G. R.; Zhang, J. Z.; Anishchenko, I.; Coventry, B.; Cao, L.; Dauparas, J.; Halabiya, S.; DeWitt, M.; Carter, L.; Houk, K. N.; Baker, D. *Nature* **2023,** *614,* 774–780. doi:10.1038/s41586-023-05696-3
66. Baek, M.; Baker, D. *Nat. Methods* **2022,** *19,* 13–14. doi:10.1038/s41592-021-01360-8
67. Dauparas, J.; Anishchenko, I.; Bennett, N.; Bai, H.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Courbet, A.; de Haas, R. J.; Bethel, N.; Leung, P. J. Y.; Huddy, T. F.; Pellock, S.; Tischer, D.; Chan, F.; Koepnick, B.; Nguyen, H.; Kang, A.; Sankaran, B.; Bera, A. K.; King, N. P.; Baker, D. *Science* **2022,** *378,* 49–56. doi:10.1126/science.add2187
68. Wang, J.; Lisanza, S.; Juergens, D.; Tischer, D.; Watson, J. L.; Castro, K. M.; Ragotte, R.; Saragovi, A.; Milles, L. F.; Baek, M.; Anishchenko, I.; Yang, W.; Hicks, D. R.; Expòsit, M.; Schlichthaerle, T.; Chun, J.-H.; Dauparas, J.; Bennett, N.; Wicky, B. I. M.; Muenks, A.; DiMaio, F.; Correia, B.; Ovchinnikov, S.; Baker, D. *Science* **2022,** *377,* 387–394. doi:10.1126/science.abn2100
69. Lovelock, S. L.; Crawshaw, R.; Basler, S.; Levy, C.; Baker, D.; Hilvert, D.; Green, A. P. *Nature* **2022,** *606,* 49–58. doi:10.1038/s41586-022-04456-z
70. Chevalier, A.; Silva, D.-A.; Rocklin, G. J.; Hicks, D. R.; Vergara, R.; Murapa, P.; Bernard, S. M.; Zhang, L.; Lam, K.-H.; Yao, G.; Bahl, C. D.; Miyashita, S.-I.; Goreshnik, I.; Fuller, J. T.; Koday, M. T.; Jenkins, C. M.; Colvin, T.; Carter, L.; Bohn, A.; Bryan, C. M.; Fernández-Velasco, D. A.; Stewart, L.; Dong, M.; Huang, X.; Jin, R.; Wilson, I. A.; Fuller, D. H.; Baker, D. *Nature* **2017,** *550,* 74–79. doi:10.1038/nature23912
71. Huang, P.-S.; Boyken, S. E.; Baker, D. *Nature* **2016,** *537,* 320–327. doi:10.1038/nature19946
72. Anishchenko, I.; Pellock, S. J.; Chidyausiku, T. M.; Ramelot, T. A.; Ovchinnikov, S.; Hao, J.; Bafna, K.; Norn, C.; Kang, A.; Bera, A. K.; DiMaio, F.; Carter, L.; Chow, C. M.; Montelione, G. T.; Baker, D. *Nature* **2021,** *600,* 547–552. doi:10.1038/s41586-021-04184-w
73. Fallas, J. A.; Ueda, G.; Sheffler, W.; Nguyen, V.; McNamara, D. E.; Sankaran, B.; Pereira, J. H.; Parmeggiani, F.; Brunette, T. J.; Cascio, D.; Yeates, T. R.; Zwart, P.; Baker, D. *Nat. Chem.* **2017,** *9,* 353–360. doi:10.1038/nchem.2673
74. Brunette, T. J.; Parmeggiani, F.; Huang, P.-S.; Bhabha, G.; Ekiert, D. C.; Tsutakawa, S. E.; Hura, G. L.; Tainer, J. A.; Baker, D. *Nature* **2015,** *528,* 580–584. doi:10.1038/nature16162
75. Gidley, F.; Parmeggiani, F. *Curr. Opin. Struct. Biol.* **2021,** *68,* 208–214. doi:10.1016/j.sbi.2021.02.002
76. Yeh, C.-T.; Obendorf, L.; Parmeggiani, F. *Front. Bioeng. Biotechnol.* **2020,** *8,* 568318. doi:10.3389/fbioe.2020.568318
77. Yeh, C.-T.; Brunette, T. J.; Baker, D.; McIntosh-Smith, S.; Parmeggiani, F. *J. Struct. Biol.* **2018,** *201,* 100–107. doi:10.1016/j.jsb.2017.09.001
78. Vithayathil, P. J.; Richards, F. M. *J. Biol. Chem.* **1961,** *236,* 1386–1389. doi:10.1016/s0021-9258(18)64182-1
79. Polgar, L.; Bender, M. L. *J. Am. Chem. Soc.* **1966,** *88,* 3153–3154. doi:10.1021/ja00965a060
80. Polgar, L.; Bender, M. L. *Biochemistry* **1967,** *6,* 610–620. doi:10.1021/bi00854a032
81. Neet, K. E.; Nanci, A.; Koshland, D. E., Jr. *J. Biol. Chem.* **1968,** *243,* 6392–6401. doi:10.1016/s0021-9258(18)93152-2
82. White, A. D.; Nowinski, A. K.; Huang, W.; Keefe, A. J.; Sun, F.; Jiang, S. *Chem. Sci.* **2012,** *3,* 3488–3494. doi:10.1039/c2sc21135a
83. Haque, M.; Forte, N.; Baker, J. R. *Chem. Commun.* **2021,** *57,* 10689–10702. doi:10.1039/d1cc03976h
84. Zhang, J.; Liu, R.; Gao, S.; Li, W.; Chen, Y.; Meng, Y.; Liu, C.; Jin, W.; Wu, J.; Wang, Y.; Hao, Y.; Yi, S.; Qing, Y.; Ge, J.; Hu, X. *npj Breast Cancer* **2023,** *9,* 28. doi:10.1038/s41523-023-00522-5
85. Xue, T. T.; Miao, Z. W.; Wang, J.; Chen, G.; Qing, Y.; Zhu, Y.; Xiao, L.; Zhang, H.; Yang, Q. Y.; Deng, D. D.; Liu, L. P.; Zeng, H.; Yin, L.; Shi, Q. F.; Song, H. M.; Zhao, X.; Wang, L. C.; Wang, J. Y. Anti-erbb2 antibody-drug conjugate and composition thereof, preparation method therefor, and application there of. Int. Pat. Appl. WO/2017/088734 A1, June 1, 2017.
86. Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; Chang, C. J. *Science* **2017,** *355,* 597–602. doi:10.1126/science.aal3316
87. Christian, A. H.; Jia, S.; Cao, W.; Zhang, P.; Meza, A. T.; Sigman, M. S.; Chang, C. J.; Toste, F. D. *J. Am. Chem. Soc.* **2019,** *141,* 12657–12662. doi:10.1021/jacs.9b04744
88. Elledge, S. K.; Tran, H. L.; Christian, A. H.; Steri, V.; Hann, B.; Toste, F. D.; Chang, C. J.; Wells, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2020,** *117,* 5733–5740. doi:10.1073/pnas.1920561117
89. Hashad, R. A.; Jap, E.; Casey, J. L.; Candace Ho, Y. T.; Wright, A.; Thalmann, C.; Sleeman, M.; Lupton, D. W.; Hagemeyer, C. E.; Cryle, M. J.; Robert, R.; Alt, K. *Chem. – Eur. J.* **2023,** *29,* e202202491. doi:10.1002/chem.202202491
90. Lin, D.; Wallace, M.; Allentoff, A. J.; Donnelly, D. J.; Gomes, E.; Voronin, K.; Gong, S.; Huang, R. Y.-C.; Kim, H.; Caceres-Cortes, J.; Bonacorsi, S., Jr. *Bioconjugate Chem.* **2020,** *31,* 1908–1916. doi:10.1021/acs.bioconjchem.0c00256
91. Ohata, J.; Krishnamoorthy, L.; Gonzalez, M. A.; Xiao, T.; Iovan, D. A.; Toste, F. D.; Miller, E. W.; Chang, C. J. *ACS Cent. Sci.* **2020,** *6,* 32–40. doi:10.1021/acscentsci.9b01038
92. Taylor, M. T.; Nelson, J. E.; Suero, M. G.; Gaunt, M. J. *Nature* **2018,** *562,* 563–568. doi:10.1038/s41586-018-0608-y
93. Kim, J.; Li, B. X.; Huang, R. Y.-C.; Qiao, J. X.; Ewing, W. R.; MacMillan, D. W. C. *J. Am. Chem. Soc.* **2020,** *142,* 21260–21266. doi:10.1021/jacs.0c09926
94. Jia, S.; He, D.; Chang, C. J. *J. Am. Chem. Soc.* **2019,** *141,* 7294–7301. doi:10.1021/jacs.8b11912

95. McGrath, N. A.; Andersen, K. A.; Davis, A. K. F.; Lomax, J. E.; Raines, R. T. *Chem. Sci.* **2015,** *6,* 752–755. doi:10.1039/c4sc01768d
96. Mix, K. A.; Raines, R. T. *Org. Lett.* **2015,** *17,* 2358–2361. doi:10.1021/acs.orglett.5b00840
97. Gauthier, M. A.; Klok, H.-A. *Biomacromolecules* **2011,** *12,* 482–493. doi:10.1021/bm101272g
98. Gauthier, M. A.; Ayer, M.; Kowal, J.; Wurm, F. R.; Klok, H.-A. *Polym. Chem.* **2011,** *2,* 1490–1498. doi:10.1039/c0py00422g
99. Nilchan, N.; Alburger, J. M.; Roush, W. R.; Rader, C. *Biochemistry* **2021,** *60,* 1080–1087. doi:10.1021/acs.biochem.0c00955
100. Hwang, D.; Nilchan, N.; Nanna, A. R.; Li, X.; Cameron, M. D.; Roush, W. R.; Park, H.; Rader, C. *Cell Chem. Biol.* **2019,** *26,* 1229–1239.e9. doi:10.1016/j.chembiol.2019.05.010
101. Prosser, L.; Emenike, B.; Sihag, P.; Shirke, R.; Raj, M. *J. Am. Chem. Soc.* **2025,** *147,* 10139–10150. doi:10.1021/jacs.4c14476
102. Bertelsen, M. B.; Tsang, E.; Palmfeldt, J.; Kristoffersen, C. H.; Nisavic, M.; Gothelf, K. V. *Bioconjugate Chem.* **2024,** *35,* 1755–1761. doi:10.1021/acs.bioconjchem.4c00317
103. Galindo, A. V.; Raj, M. *Org. Lett.* **2024,** *26,* 8356–8360. doi:10.1021/acs.orglett.4c03101
104. Bal, A.; Singh, S. K.; Kashyap, T.; Rai, V. *Methods Enzymol.* **2022,** *675,* 383–396. doi:10.1016/bs.mie.2022.07.016
105. Adusumalli, S. R.; Rawale, D. G.; Singh, U.; Tripathi, P.; Paul, R.; Kalra, N.; Mishra, R. K.; Shukla, S.; Rai, V. *J. Am. Chem. Soc.* **2018,** *140,* 15114–15123. doi:10.1021/jacs.8b10490
106. Rawale, D. G.; Thakur, K.; Sreekumar, P.; T. K., S.; A., R.; Adusumalli, S. R.; Mishra, R. K.; Rai, V. *Chem. Sci.* **2021,** *12,* 6732–6736. doi:10.1039/d1sc00335f
107. Reddy, N. C.; Molla, R.; Joshi, P. N.; T. K., S.; Basu, I.; Kawadkar, J.; Kalra, N.; Mishra, R. K.; Chakrabarty, S.; Shukla, S.; Rai, V. *Nat. Commun.* **2022,** *13,* 6038. doi:10.1038/s41467-022-33772-1
108. Forte, N.; Benni, I.; Karu, K.; Chudasama, V.; Baker, J. R. *Chem. Sci.* **2019,** *10,* 10919–10924. doi:10.1039/c9sc03825f
109. Chauhan, P.; V., R.; Kumar, M.; Molla, R.; V. B., U.; Rai, V. *ACS Cent. Sci.* **2023,** *9,* 137–150. doi:10.1021/acscentsci.2c01455
110. Zhong, Q.; Xiao, X.; Qiu, Y.; Xu, Z.; Chen, C.; Chong, B.; Zhao, X.; Hai, S.; Li, S.; An, Z.; Dai, L. *MedComm* **2023,** *4,* e261. doi:10.1002/mco2.261
111. Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods* **2005,** *2,* 99–104. doi:10.1038/nmeth735
112. Plaks, J. G.; Falatach, R.; Kastantin, M.; Berberich, J. A.; Kaar, J. L. *Bioconjugate Chem.* **2015,** *26,* 1104–1112. doi:10.1021/acs.bioconjchem.5b00161
113. Dozier, J. K.; Khatwani, S. L.; Wollack, J. W.; Wang, Y.-C.; Schmidt-Dannert, C.; Distefano, M. D. *Bioconjugate Chem.* **2014,** *25,* 1203–1212. doi:10.1021/bc500240p
114. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004,** *126,* 15046–15047. doi:10.1021/ja044996f
115. Agarwal, P.; Beahm, B. J.; Shieh, P.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2015,** *54,* 11504–11510. doi:10.1002/anie.201504249
116. Shajahan, A.; Parashar, S.; Goswami, S.; Ahmed, S. M.; Nagarajan, P.; Sampathkumar, S.-G. *J. Am. Chem. Soc.* **2017,** *139,* 693–700. doi:10.1021/jacs.6b08894
117. Cheng, B.; Dong, L.; Zhu, Y.; Huang, R.; Sun, Y.; You, Q.; Song, Q.; Paton, J. C.; Paton, A. W.; Chen, X. *ACS Chem. Biol.* **2019,** *14,* 2141–2147. doi:10.1021/acschembio.9b00556
118. Dold, J. E. G. A.; Wittmann, V. *ChemBioChem* **2021,** *22,* 1243–1251. doi:10.1002/cbic.202000715
119. Almaraz, R. T.; Aich, U.; Khanna, H. S.; Tan, E.; Bhattacharya, R.; Shah, S.; Yarema, K. J. *Biotechnol. Bioeng.* **2012,** *109,* 992–1006. doi:10.1002/bit.24363
120. Murray, T. V.; Kozakowska-McDonnell, K.; Tibbles, A.; Taylor, A.; Higazi, D.; Rossy, E.; Rossi, A.; Genapathy, S.; Tamburrino, G.; Rath, N.; Tigue, N.; Lindo, V.; Vaughan, T.; Papworth, M. A. *mAbs* **2021,** *13,* 1992068. doi:10.1080/19420862.2021.1992068
121. Journeaux, T.; Geeson, M. B.; Murray, T. V.; Papworth, M. A.; Gothard, M.; Kettle, J. G.; Vasco, A. V.; Bernardes, G. J. L. *Angew. Chem., Int. Ed.* **2025,** *64,* e202417620. doi:10.1002/anie.202417620
122. Wang, L.-X.; Tong, X.; Li, C.; Giddens, J. P.; Li, T. *Annu. Rev. Biochem.* **2019,** *88,* 433–459. doi:10.1146/annurev-biochem-062917-012911
123. Cheng, B.; Xie, R.; Dong, L.; Chen, X. *ChemBioChem* **2016,** *17,* 11–27. doi:10.1002/cbic.201500344
124. Kufleitner, M.; Haiber, L. M.; Wittmann, V. *Chem. Soc. Rev.* **2023,** *52,* 510–535. doi:10.1039/d2cs00764a
125. O'Shannessy, D. J.; Dobersen, M. J.; Quarles, R. H. *Immunol. Lett.* **1984,** *8,* 273–277. doi:10.1016/0165-2478(84)90008-7
126. Wolfe, C. A. C.; Hage, D. S. *Anal. Biochem.* **1995,** *231,* 123–130. doi:10.1006/abio.1995.1511
127. Zuberbühler, K.; Casi, G.; Bernardes, G. J. L.; Neri, D. *Chem. Commun.* **2012,** *48,* 7100–7102. doi:10.1039/c2cc32412a
128. Mukai, K.; Cost, R.; Zhang, X. S.; Condiff, E.; Cotton, J.; Liu, X.; Boudanova, E.; Niebel, B.; Piepenhagen, P.; Cai, X.; Park, A.; Zhou, Q. *mAbs* **2024,** *16,* 2415333. doi:10.1080/19420862.2024.2415333
129. Angelastro, A.; Barkhanskiy, A.; Mattey, A. P.; Pallister, E. G.; Spiess, R.; Goundry, W.; Barran, P.; Flitsch, S. L. *ChemSusChem* **2022,** *15,* e202102592. doi:10.1002/cssc.202102592
130. Ramakrishnan, B.; Qasba, P. K. *J. Biol. Chem.* **2002,** *277,* 20833–20839. doi:10.1074/jbc.m111183200
131. Pasek, M.; Ramakrishnan, B.; Boeggeman, E.; Manzoni, M.; Waybright, T. J.; Qasba, P. K. *Bioconjugate Chem.* **2009,** *20,* 608–618. doi:10.1021/bc800534r
132. Qasba, P. K. *Bioconjugate Chem.* **2015,** *26,* 2170–2175. doi:10.1021/acs.bioconjchem.5b00173
133. Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. *J. Am. Chem. Soc.* **1998,** *120,* 5583–5584. doi:10.1021/ja980833d
134. Ou, C.; Li, C.; Zhang, R.; Yang, Q.; Zong, G.; Dai, Y.; Francis, R. L.; Bournazos, S.; Ravetch, J. V.; Wang, L.-X. *Bioconjugate Chem.* **2021,** *32,* 1888–1897. doi:10.1021/acs.bioconjchem.1c00314
135. Tang, F.; Shi, W.; Huang, W. *Methods Mol. Biol. (N. Y., NY, U. S.)* **2019,** *2033,* 221–238. doi:10.1007/978-1-4939-9654-4_15
136. Tong, X.; Li, T.; Orwenyo, J.; Toonstra, C.; Wang, L.-X. *Bioorg. Med. Chem.* **2018,** *26,* 1347–1355. doi:10.1016/j.bmc.2017.07.053
137. Li, T.; Tong, X.; Yang, Q.; Giddens, J. P.; Wang, L.-X. *J. Biol. Chem.* **2016,** *291,* 16508–16518. doi:10.1074/jbc.m116.738765
138. van Geel, R.; Wijdeven, M. A.; Heesbeen, R.; Verkade, J. M. M.; Wasiel, A. A.; van Berkel, S. S.; van Delft, F. L. *Bioconjugate Chem.* **2015,** *26,* 2233–2242. doi:10.1021/acs.bioconjchem.5b00224
139. de Bever, L.; Popal, S.; van Schaik, J.; Rubahamya, B.; van Delft, F. L.; Thurber, G. M.; van Berkel, S. S. *Bioconjugate Chem.* **2023,** *34,* 538–548. doi:10.1021/acs.bioconjchem.2c00611
140. Müller, M. M. *Biochemistry* **2018,** *57,* 177–185. doi:10.1021/acs.biochem.7b00861
141. Zhu, T.; Wu, B. *ACS Catal.* **2024,** *14,* 15811–15826. doi:10.1021/acscatal.4c04145

142. Zhang, Y.; Park, K.-Y.; Suazo, K. F.; Distefano, M. D. *Chem. Soc. Rev.* **2018,** *47,* 9106–9136. doi:10.1039/c8cs00537k
143. Freund, C.; Schwarzer, D. *ChemBioChem* **2021,** *22,* 1347–1356. doi:10.1002/cbic.202000745
144. Williamson, D. J.; Fascione, M. A.; Webb, M. E.; Turnbull, W. B. *Angew. Chem., Int. Ed.* **2012,** *51,* 9377–9380. doi:10.1002/anie.201204538
145. Podracky, C. J.; An, C.; DeSousa, A.; Dorr, B. M.; Walsh, D. M.; Liu, D. R. *Nat. Chem. Biol.* **2021,** *17,* 317–325. doi:10.1038/s41589-020-00706-1
146. Yu, W.; Gillespie, K. P.; Chhay, B.; Svensson, A.-S.; Nygren, P.-Å.; Blair, I. A.; Yu, F.; Tsourkas, A. *Bioconjugate Chem.* **2021,** *32,* 1058–1066. doi:10.1021/acs.bioconjchem.1c00099
147. Dennler, P.; Chiotellis, A.; Fischer, E.; Brégeon, D.; Belmant, C.; Gauthier, L.; Lhospice, F.; Romagne, F.; Schibli, R. *Bioconjugate Chem.* **2014,** *25,* 569–578. doi:10.1021/bc400574z
148. Jeger, S.; Zimmermann, K.; Blanc, A.; Grünberg, J.; Honer, M.; Hunziker, P.; Struthers, H.; Schibli, R. *Angew. Chem., Int. Ed.* **2010,** *49,* 9995–9997. doi:10.1002/anie.201004243
149. Dickgiesser, S.; Rieker, M.; Mueller-Pompalla, D.; Schröter, C.; Tonillo, J.; Warszawski, S.; Raab-Westphal, S.; Kühn, S.; Knehans, T.; Könning, D.; Dotterweich, J.; Betz, U. A. K.; Anderl, J.; Hecht, S.; Rasche, N. *Bioconjugate Chem.* **2020,** *31,* 1070–1076. doi:10.1021/acs.bioconjchem.0c00061
150. Wehrmüller, J. E.; Frei, J. C.; Hechler, T.; Kulke, M.; Pahl, A.; Béhé, M.; Schibli, R.; Spycher, P. R. *ChemBioChem* **2025,** *26,* e202400511. doi:10.1002/cbic.202400511
151. Fryszkowska, A.; An, C.; Alvizo, O.; Banerjee, G.; Canada, K. A.; Cao, Y.; DeMong, D.; Devine, P. N.; Duan, D.; Elgart, D. M.; Farasat, I.; Gauthier, D. R.; Guidry, E. N.; Jia, X.; Kong, J.; Kruse, N.; Lexa, K. W.; Makarov, A. A.; Mann, B. F.; Milczek, E. M.; Mitchell, V.; Nazor, J.; Neri, C.; Orr, R. K.; Orth, P.; Phillips, E. M.; Riggins, J. N.; Schafer, W. A.; Silverman, S. M.; Strulson, C. A.; Subramanian, N.; Voladri, R.; Yang, H.; Yang, J.; Yi, X.; Zhang, X.; Zhong, W. *Science* **2022,** *376,* 1321–1327. doi:10.1126/science.abn2009

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.67

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

A practical CO_2-mediated synthesis of 5,6-carboxylated silicon-rhodamines for targeted probe development

Dongjie Hou[1,2], Shaowei Wu[1,2], Ning Xu[1], Pengjun Bao[1,2], Wenhao Jia[1,2], Qinglong Qiao[*1] and Zhaochao Xu[*1,2]

Full Research Paper

Open Access

Address:
[1]State Key Laboratory of Chemical Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China and [2]University of Chinese Academy of Sciences, Beijing 100049, China

Email:
Qinglong Qiao[*] - qqlqiao@dicp.ac.cn; Zhaochao Xu[*] - zcxu@dicp.ac.cn

* Corresponding author

Keywords:
carboxylation; CO_2; fluorescent probes; silicon-rhodamine; super-resolution imaging

Beilstein J. Org. Chem. **2026,** *22,* 915–924.
https://doi.org/10.3762/bjoc.22.72

Received: 29 March 2026
Accepted: 29 May 2026
Published: 10 June 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring

© 2026 Hou et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Silicon-rhodamine (SiR) dyes are among the most important fluorophores for super-resolution imaging owing to their far-red emission, high photostability, and tunable lactone–zwitterion equilibrium. However, their broader application in targeted probe development has been limited by the challenging synthesis of 5,6-carboxylated SiR derivatives, which are essential intermediates for bioconjugation. Here, we report a practical CO_2-mediated strategy for the synthesis of 5,6-carboxylated SiRs from brominated SiR precursors via lithium–halogen exchange and direct carboxylation. Using *n*-BuLi and readily available CO_2, this method delivers carboxylated SiR derivatives in 60–93% yields while avoiding the use of *t*-BuLi, toxic CO, and expensive palladium catalysts. In addition, the crude carboxylation mixtures can be directly subjected to amide coupling without chromatographic purification, enabling one-pot access to HaloTag-targeted SiR probes. The resulting probes were successfully applied to long-term live-cell super-resolution imaging, allowing visualization of filopodial dynamics and mitochondrial remodeling. This work establishes a concise and efficient route to functionalized SiR dyes and facilitates the rapid development of targeted SiR probes.

Introduction

Rhodamine dyes, which enable fluorescence imaging in living cells – a powerful technique for studying biological systems in vivo [1,2] – have emerged as one of the most important fluorophores for live-cell bioimaging [3-7] by virtue of their high brightness, excellent photostability [8-10], and, most importantly, tunable lactone–zwitterion (L–Z) equilibrium [11-13]. This distinctive feature enables controllable fluorescence switching and fluorogenic responses, making rhodamines particularly valuable for advanced imaging beyond the optical diffraction limit [14-19]. Replacement of the conventional bridging oxygen atom with silicon shifts the emission of rhodamines from red to the far-red region (typically $\lambda_{em} \approx 650$–670 nm) while retaining

high brightness and excellent photostability [20-24]. As a result, silicon-rhodamines (SiRs) have become the most practical rhodamine scaffolds for live-cell imaging, benefiting from reduced cellular autofluorescence and phototoxicity, together with favorable cell permeability [25,26]. Johnsson and co-workers further showed that SiRs can exist in a nonfluorescent spirolactone form, whereas target binding shifts the equilibrium toward the fluorescent zwitterionic form, thereby generating a strong fluorogenic response (Figure 1a) [27-29]. This behavior makes SiRs highly attractive as fluorogenic probes for

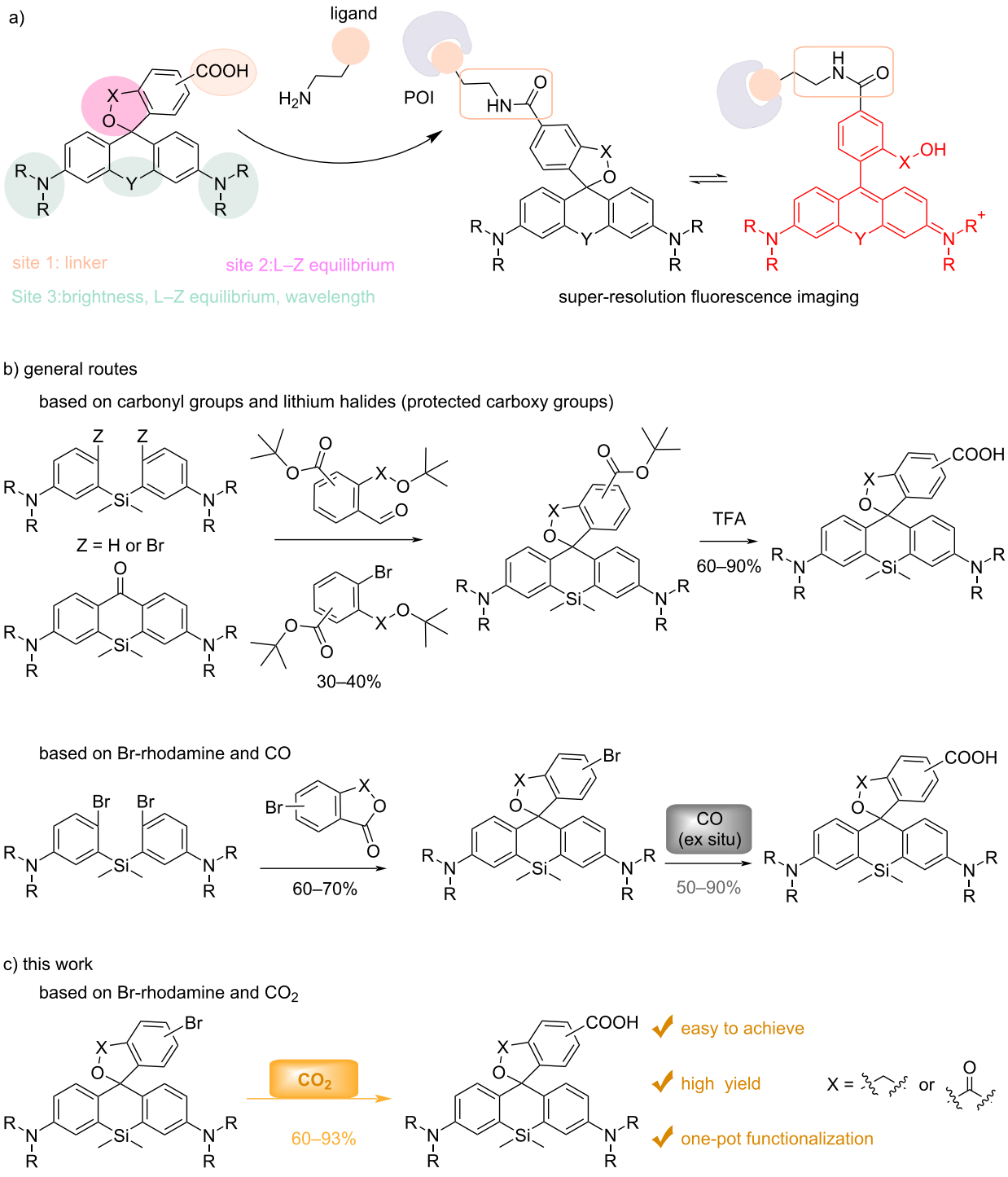


Figure 1: (a) Modular structure of spirocyclic rhodamines. (b) Conventional carboxylation strategies. (c) CO_2-based direct carboxylation method developed in this work.

precise super-resolution imaging by enabling target-specific recognition and significantly improving signal-to-noise ratio (SNR). In addition, hydroxymethyl-substituted SiR derivatives, including **HMSiR** [30,31] and **Aze-HMSiR** [32-34], can undergo spontaneous blinking and have become indispensable tools for live-cell single-molecule localization microscopy (SMLM), highlighting the versatility of SiR dyes in live-cell imaging [35,36].

Building on this foundation, structural modification of SiR dyes has evolved into a well-established molecular design framework (Figure 1a). Conformational restriction and optimization of amino substituents have been widely employed to enhance brightness and photostability [8,37-39], whereas extension of π-conjugation and electronic tuning have expanded the emission of SiR dyes from the far-red to the near-infrared region, reaching up to 740 nm [21,22,40]. In addition, modulation of the L–Z equilibrium through tuning of N-substituents has enabled the development of fluorogenic and spontaneously blinking SiR probes with finely controlled off–on behavior [11,26,28,41]. However, despite the maturity of photophysical optimization, the broader expansion of SiR chemistry is increasingly constrained by the limited accessibility of 5,6-carboxylated SiR derivatives. The carboxy groups provide a general and efficient handle for amide coupling with targeting ligands and are therefore central to the construction of SiR-based probes, including HaloTag, SNAP-tag, peptide-, and small-molecule-conjugated derivatives [18,35]. The synthetic accessibility of 5,6-carboxylated SiRs has thus become a central bottleneck in the diversification and broader application of SiR chemistry.

Traditional approaches to the synthesis of 5,6-carboxylated SiRs generally rely on the prior installation of protected carboxy-containing precursors (Figure 1b), followed by cyclization to construct the SiR framework and subsequent deprotection [41-44]. In these routes, the strong electron-withdrawing effect of the carboxy substituent reduces substrate reactivity, thereby diminishing the efficiency of the key ring-closing step and often resulting in low overall yields [45]. Moreover, these methods commonly rely on strongly basic conditions, particularly the use of highly reactive *tert*-butyllithium, which further reduces their practical utility. In addition, the preparation of protected carboxy-group containing substrates is often cumbersome [46], especially in the synthesis of (hydroxymethyl)rhodamine intermediates, thereby further increasing synthetic complexity [41]. To overcome these limitations, Butkevich developed an improved strategy based on brominated precursors [45]. Owing to the weaker electron-withdrawing effect of bromine relative to carboxy groups, these precursors exhibit higher reactivity and enable the construction of brominated SiR scaffolds in improved yields (60–70%) under milder conditions. Subsequent Pd-catalyzed carbonylation using in situ-generated CO then furnished the corresponding carboxylated SiRs in 50–90% yields. Although this approach significantly improved synthetic efficiency, it still relies on specialized setups for CO generation and expensive palladium catalysts, thereby increasing operational complexity, safety concerns, and overall cost.

Herein, inspired by the limitations of existing methods, we report a direct CO_2-mediated carboxylation strategy for the efficient synthesis of 5,6-carboxylated SiR derivatives from brominated SiR precursors (Figure 1c). By introducing the classical carboxylation of metallated aryl halides with CO_2 into SiR chemistry, this method enables the efficient conversion of brominated SiRs into the corresponding carboxylated products using *n*-BuLi and readily available CO_2 under standard Schlenk conditions. The target products are obtained in 60–93% yields under relatively mild conditions. Compared with previous approaches, this strategy avoids the need for specialized apparatus for in situ CO generation as well as expensive palladium catalysts, thereby significantly reducing experimental complexity and synthetic cost. Importantly, the resulting carboxylated SiR derivatives can be directly used in subsequent amidation reactions with targeting ligands without column chromatography, facilitating the rapid construction of high-performance targeted fluorescent probes. Overall, this strategy provides a straightforward, efficient, and practical synthetic route for the functional derivatization of SiR dyes and the development of targeted probes.

Results and Discussion

Synthesis of carboxy-substituted silicon-rhodamine

Inspired by Butkevich's work [45], we envisioned that carboxylated SiRs might be accessed through brominated SiR intermediates, thereby avoiding the traditional multistep synthesis from protected carboxylated precursors. Butkevich convincingly demonstrated the generality of Pd-catalyzed carbonylation conditions employing ex situ CO for the conversion of bromo-oxorhodamines, carbo-rhodamines, and silicon-rhodamines into their corresponding carboxylated forms. However, the operational complexity of these conditions may limit their practical utility. In particular, the generation and handling of CO remain major concerns because of its toxicity and limited accessibility in standard synthetic laboratories. To address this limitation, we explored a CO-free carboxylation strategy using CO_2 as an inexpensive, readily available, and non-toxic C1 source. In principle, CO_2 can serve as a carbonyl source through nucleophilic addition followed by protonation to furnish the corresponding carboxylic acid.

On the basis of this design (Figure 2a), brominated SiR precursors **1a–c** were treated with 2.5 equiv of *n*-BuLi in anhydrous THF at −78 °C under a nitrogen atmosphere to generate the corresponding aryllithium intermediates, followed by introduction of CO_2 via a balloon. The reaction mixture was then allowed to warm to room temperature and stirred overnight to ensure complete trapping of the aryllithium intermediates by CO_2. After aqueous quenching and extraction, the crude products were directly subjected to HPLC analysis.

To evaluate the generality of this strategy, three brominated SiR precursors **1a–c**, all accessible by reported procedures, were selected as model substrates (Figure 3). Direct HPLC analysis of the crude reaction mixtures at 640 nm revealed clear separation between substrates and products. The retention times of **1a**, **1b**, and **1c** were 8.5, 7.5, and 7.4 min, respectively, whereas those of the corresponding carboxylated products **3a**, **3b**, and **3c** were 5.2, 4.4, and 3.2 min, allowing unambiguous distinction between starting materials and products. On the basis of peak

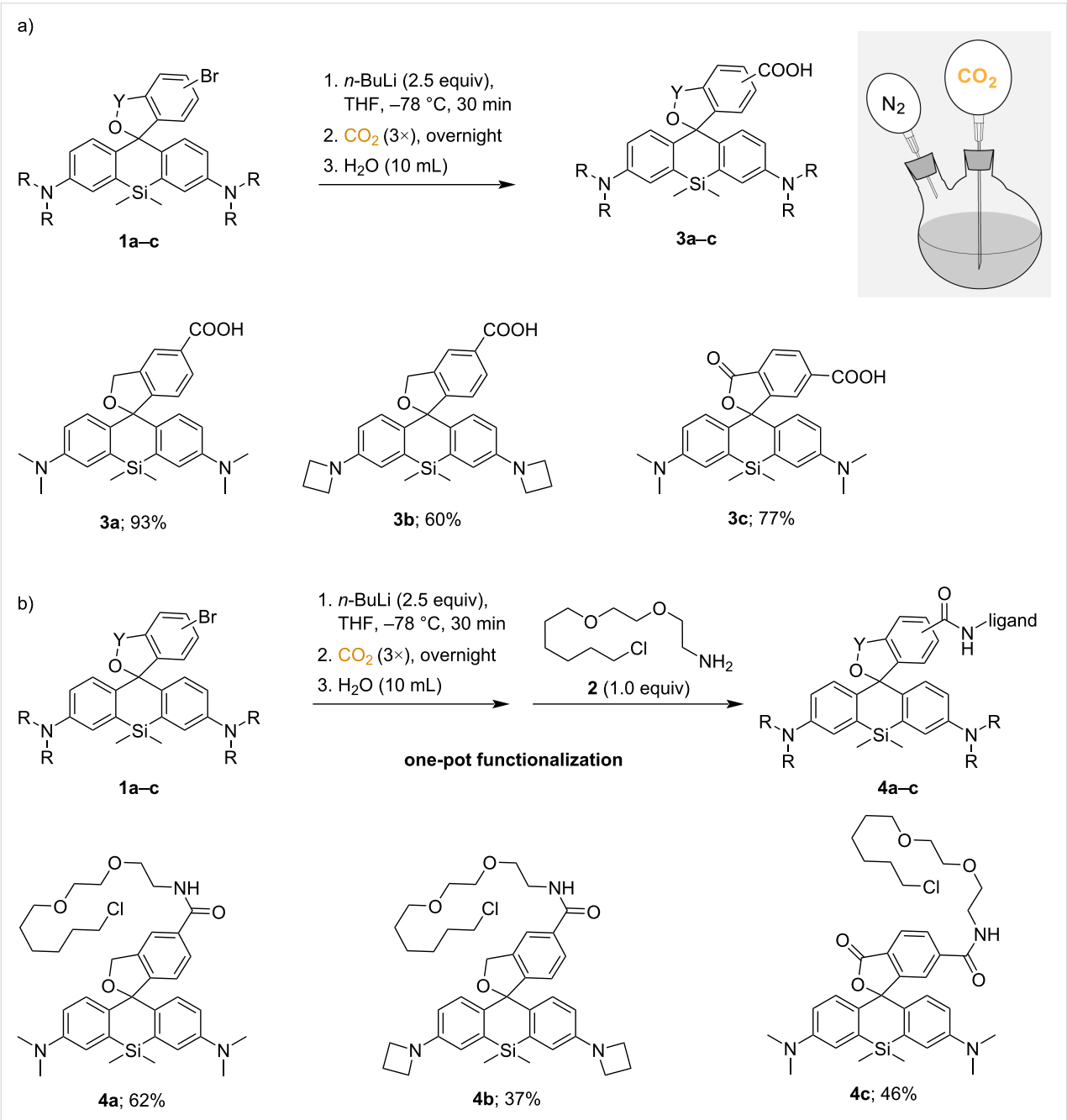


Figure 2: (a) Direct carboxylation of brominated silicon-rhodamines **1a–c** under *n*-BuLi/CO_2. (b) Strategy for the one-pot functionalization of brominated silicon-rhodamines **1a–c** to construct SiR–halo fluorescent probes **4a–c**.

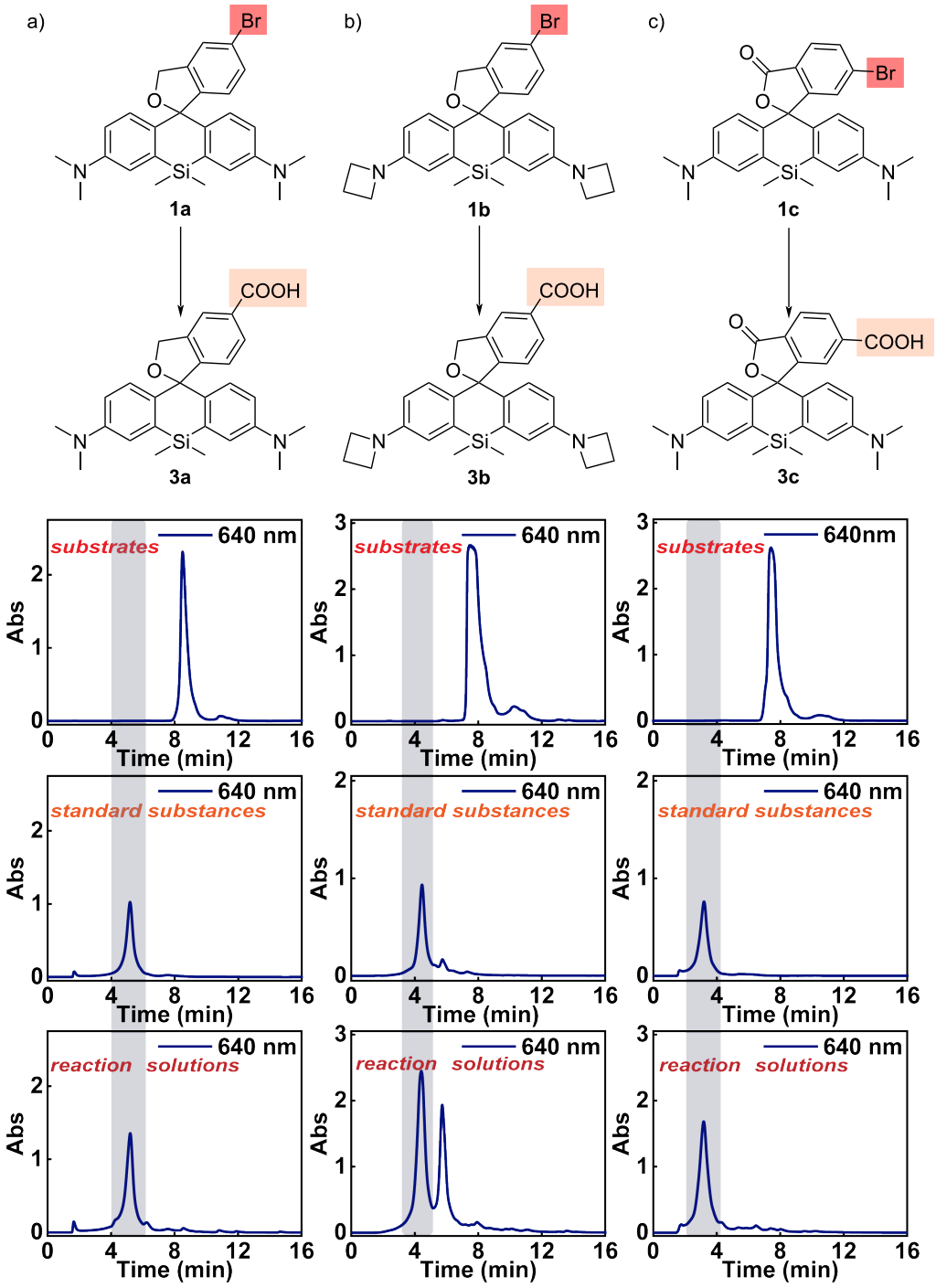


Figure 3: (a–c) Conversion of three bromo-SiR precursors **1a–c** in the CO_2 carboxylation reaction and their corresponding HPLC results.

integration, products **3a**, **3b**, and **3c** were obtained in 93%, 60%, and 77% yields, respectively. Compared with previously reported routes, this protocol provides a simpler and safer approach to carboxylated SiR derivatives while delivering good to excellent yields.

Sequential one-pot functionalization of brominated silicon-rhodamine

In the carboxylation of **1b**, a prominent side peak was observed in addition to the desired product peak (Figure 3b). NMR (Supporting Information File 1, Figure S17) and HRMS (Figure

S24) analysis identified this side product as the uncarboxylated rhodamine formed by lithium–halogen exchange followed by protonation. This byproduct likely arises from protonation of the aryllithium intermediate before efficient trapping by CO_2, possibly due to trace moisture or other proton sources in the reaction system. Improving the CO_2-trapping efficiency and maintaining strictly anhydrous conditions may help suppress this side reaction. Similar uncarboxylated rhodamine byproducts were also present in the reaction mixtures of the other two precursors, although in lower amounts. Importantly, these species lack a carboxy group and therefore cannot participate in subsequent amide coupling reactions. Moreover, although small amounts of these byproducts were formed, they did not significantly interfere with the subsequent reactions. Thus, the reaction mixtures could be extracted with ethyl acetate and concentrated before being directly subjected to amide coupling without chromatographic purification.

To demonstrate the utility of this strategy for probe construction, we selected three representative SiR scaffolds **1a–c**, which were widely used in super-resolution imaging, for direct functionalization with HaloTag ligands through a one-pot sequence (Figure 2b). After treatment of compounds **1a–c** with *n*-BuLi and CO_2, the reaction mixtures were extracted with ethyl acetate and concentrated, and the resulting crude carboxylated intermediates were directly subjected to amide coupling with the HaloTag amine using HATU/DIPEA as the activation system. This strategy afforded three high-performance fluorescent probes **4a**, **4b**, and **4c**, in overall yields of 62%, 37%, and 46%, respectively. The absorption and emission spectra of **4a**, **4b**, and **4c** (Figures S1–S3 in Supporting Information File 1) showed characteristic SiR fluorescence with emission maxima around 640 nm, making them suitable for far-red fluorescence imaging. These results demonstrate that the present CO_2-mediated strategy provides a simple, safe, and efficient route to carboxylated SiR derivatives and enables their rapid conversion into targeted SiR probes.

Long-term live-cell super-resolution imaging using one-pot-functionalized silicon-rhodamine probes

Encouraged by the successful one-pot functionalization of silicon-rhodamine scaffolds, we next evaluated the performance of the resulting HaloTag-targeted SiR probes in long-term super-resolution imaging. Among them, probes **4a** and **4b**, corresponding to **HMSiR-Halo** and **AzeHMSiR-Halo** originally reported by the Urano and Xu groups [30,32-34], respectively, are widely used fluorophores for single-molecule imaging because of their excellent spontaneous blinking properties. Accordingly, the two probes obtained through our one-pot strategy were first examined by single-molecule localization microscopy (SMLM) imaging (Figure 4).

Lifeact-HaloTag, H_2B-HaloTag, and TOMM20-HaloTag fusion proteins were transiently expressed in HeLa cells to label actin filaments, the nucleus, and the mitochondrial outer membrane, respectively. Both **4a** and **4b** enabled high-quality super-resolution imaging of these distinct subcellular structures (Figure 4a), with localization precisions ranging from 24 to 46 nm (Figure 4b). In the reconstructed SMLM images, both probes clearly resolved microfilament structures with high localization precision. Probe **4b** exhibited a localization precision of 24.9 nm due to its high brightness, superior to that of **4a** (45.9 nm). In addition, filopodia displayed a width of 69.4 nm in the SMLM images, whereas the corresponding width measured by wide-field (WF) imaging was 348.8 nm (Figure S4 in Supporting Information File 1), further demonstrating the excellent performance of these probes in super-resolution imaging.

On the basis of its superior localization precision, compound **4b** was further used to investigate microfilament dynamics by long-term live-cell SMLM imaging (Figure 4c). Segmental reconstruction using continuously acquired image sequences enabled real-time tracking of filopodial growth and wiggling dynamics (Figure 4d). During the first 0–160 s, the filopodium indicated by the blue arrow underwent slight retraction, followed by gradual extension. Its length increased from 6.03 μm to 11.16 μm over the subsequent approximately 240 s (Figure S5 in Supporting Information File 1). In addition, the filopodium labeled by the orange arrow exhibited a pronounced wiggling process, with the angle between the filopodium and the cell membrane edge increasing from 50° to 95° (Figure S6 in Supporting Information File 1). This demonstrates that compound **4b** is suitable for long-term dynamic super-resolution imaging.

Compound **4c** is a practical fluorogenic HaloTag probe for live-cell imaging. In previous studies, HaloTag-compatible rhodamine probes have been used to visualize a variety of intracellular proteins, including histone H_2B, Sox2, vimentin, clathrin, and ensconsin [15,16,39], underscoring the broad utility of this labeling strategy for protein-specific imaging. Subsequently, the **4c** probe obtained through our one-pot strategy was applied to SIM imaging of multiple subcellular structures, including actin filaments, nucleus, and mitochondria (Figure 5a). Long-term live-cell imaging of mitochondria over 20 min was successfully performed, revealing pronounced temporal remodeling of the mitochondrial network, including recurrent twine-unfasten events (Figure 5b). Specifically, at 20 s, two neighboring mitochondria marked by the white arrow

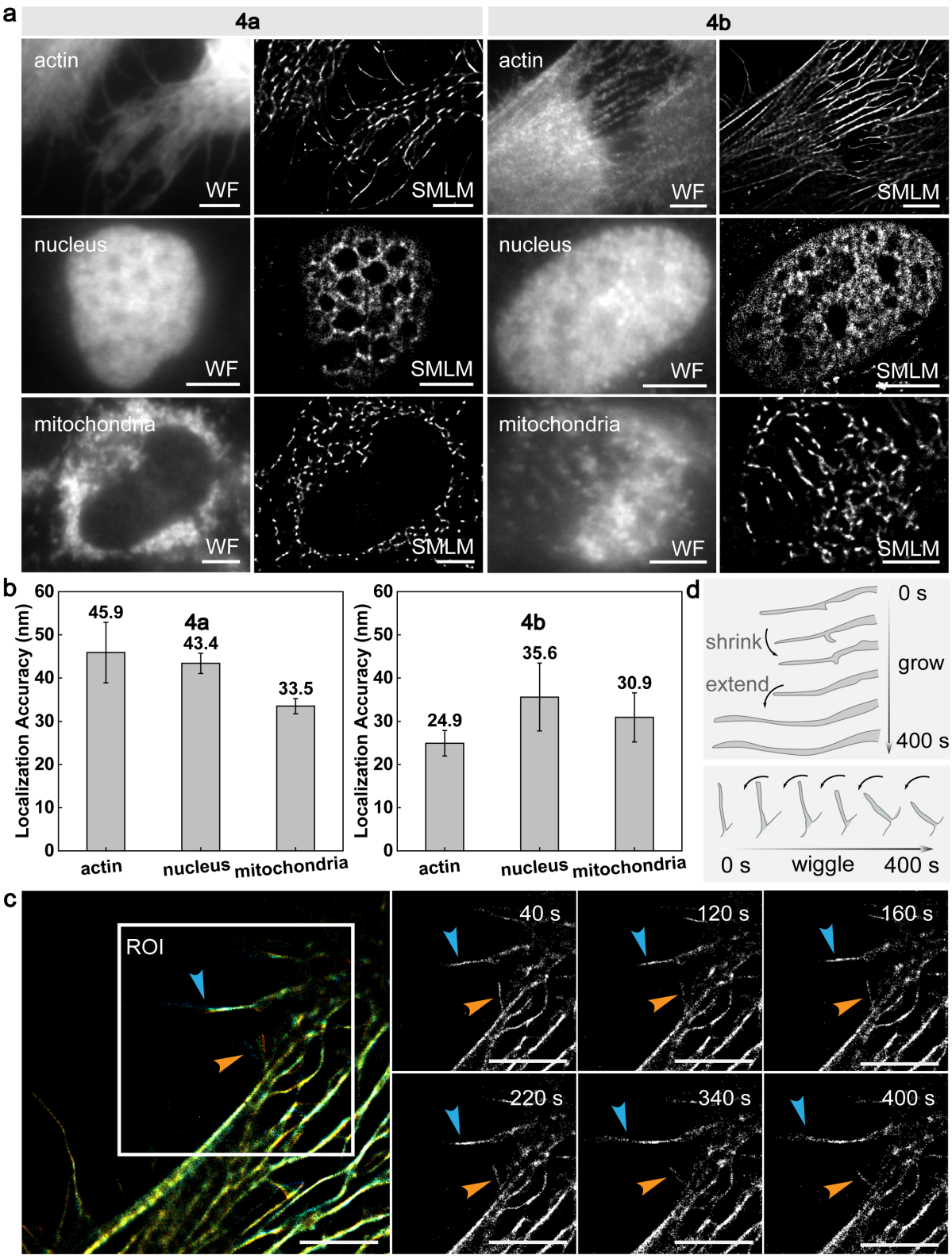


Figure 4: SMLM imaging of different organelles in live HeLa cells stained with **4a** and **4b** (150 nM). (a) SMLM and wide-field images of actin labeled with Lifeact-HaloTag fusion protein, nucleus labeled with H_2B-HaloTag fusion protein, and mitochondria labeled with TOMM20-HaloTag fusion protein. (b) Average localization precisions of **4a** and **4b** for different organelles. (c) Long-term SMLM imaging of actin with **4b** in live HeLa cells. (d) Schematic diagram of filopodia growth and wiggling process in (c). Scale bars: 5.0 μm.

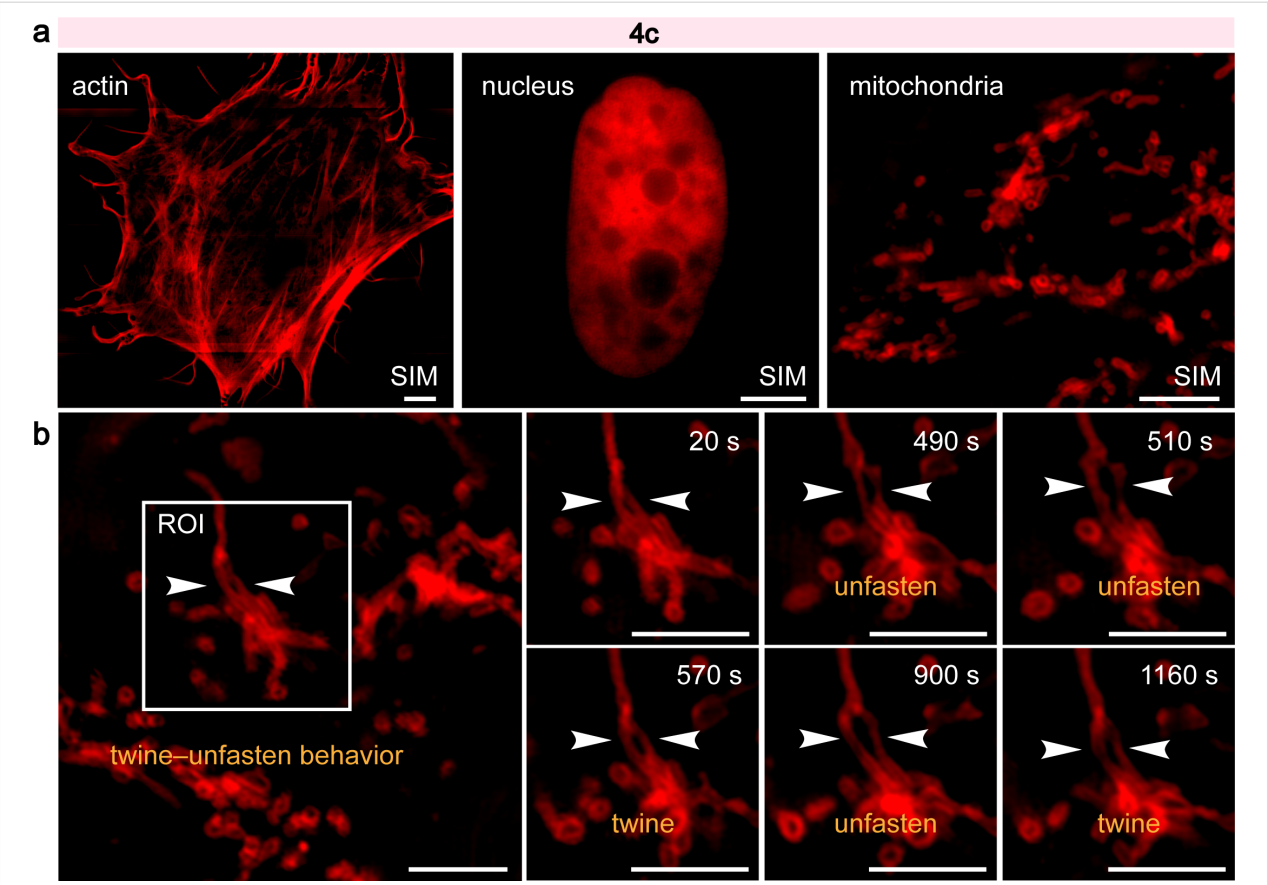


Figure 5: SIM imaging of different organelles in live HeLa cells stained with **4c** (200 nM). (a) SIM images of actin labeled with Lifeact-HaloTag fusion protein, nucleus labeled with H_2B-HaloTag fusion protein, and mitochondria labeled with TOMM20-HaloTag fusion protein. (b) Long-term SIM imaging of mitochondria with **4c** in live HeLa cells. Scale bars: 5.0 μm.

were in a twined state. By 490 s, they gradually unfastened, followed by further separation at 510 s. At 570 s, the twined state reappeared, whereas at 900 s, a clearly unfastened state was again observed, followed by gradual re-twining. These observations demonstrate the reversible and highly dynamic nature of mitochondrial twine-unfasten behavior and highlight the utility of **4c** for long-term dynamic super-resolution imaging.

Conclusion

In conclusion, we have developed a practical CO_2-mediated strategy for the synthesis of 5,6-carboxylated silicon-rhodamine (SiR) derivatives from brominated SiR precursors via lithium–halogen exchange and direct carboxylation. This method enables efficient access to carboxylated SiRs under mild conditions while avoiding the use of *t*-BuLi, toxic CO, and expensive palladium catalysts. The crude carboxylation mixtures can be directly advanced to amide coupling without chromatographic purification, allowing streamlined one-pot construction of functionalized SiR probes. Using this approach, several high-performance probes were readily prepared and successfully applied to long-term live-cell super-resolution imaging. Overall, this work improves the synthetic accessibility of functionalized SiR dyes and provides a concise and versatile platform for the rapid development of targeted super-resolution probes.

Supporting Information

Supporting Information File 1

Experimental details and characterization of compounds by NMR and ESIMS.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-72-S1.pdf]

Funding

This work is supported by the National Natural Science Foundation of China (22541705, U25A20620, 22225806, 22522816, 22578443, 22378385, 22278394, 22507121, and 22508383); the Chinese Academy of Sciences Project for Young Scientists in Basic Research Grant No.YSBR-104; the Energy Revolution

S&T Program of Yulin Innovation Institute of Clean Energy, Grant No.E412050705; the Liaoning Binhai Laboratory (LBLD-2024-07); the Dalian Institute of Chemical Physics (DICPI202436, DICP&SIA UN202502, DMU-1&DICP UN202301, DMU-1&DICP UN202301, and DMU-2&DICP UN202502); the Natural Science Foundation of Liaoning (2025-MS-061 and 2025-BS-0147); and the Dalian Science and Technology Innovation Fund Program (2022JJ11CG007).

Author Contributions

Dongjie Hou: conceptualization; data curation; formal analysis; investigation; writing – original draft. Shaowei Wu: data curation; investigation. Ning Xu: data curation; investigation. Pengjun Bao: data curation; investigation. Wenhao Jia: data curation. Qinglong Qiao: conceptualization; funding acquisition; project administration; supervision; writing – original draft; writing – review & editing. Zhaochao Xu: conceptualization; funding acquisition; project administration; supervision; writing – review & editing.

ORCID® iDs

Zhaochao Xu - https://orcid.org/0000-0002-2491-8938

Data Availability Statement

All data that supports the findings of this study is available in the published article and/or the supporting information of this article.

References

1. Kim, H. N.; Lee, E.-H.; Xu, Z.; Kim, H.-E.; Lee, H.-S.; Lee, J.-H.; Yoon, J. *Biomaterials* **2012,** *33,* 2282–2288. doi:10.1016/j.biomaterials.2011.11.073
2. Qiao, Q.; Zhao, M.; Lang, H.; Mao, D.; Cui, J.; Xu, Z. *RSC Adv.* **2014,** *4,* 25790–25794. doi:10.1039/c4ra03725a
3. Wang, L.; Du, W.; Hu, Z.; Uvdal, K.; Li, L.; Huang, W. *Angew. Chem., Int. Ed.* **2019,** *58,* 14026–14043. doi:10.1002/anie.201901061
4. Zhang, J.; Shi, H.; Huang, C.; Mei, L.; Guo, Q.; Cheng, K.; Wu, P.; Su, D.; Chen, Q.; Gan, S.; Chan, C. K. W.; Shi, J.; Chen, J. L.; Choi, C. H. J.; Yao, S. Q.; Chen, X.-K.; Tang, B. Z.; He, J.; Sun, H. *ACS Nano* **2023,** *17,* 3632–3644. doi:10.1021/acsnano.2c10467
5. Zeng, S.; Liu, X.; Kafuti, Y. S.; Kim, H.; Wang, J.; Peng, X.; Li, H.; Yoon, J. *Chem. Soc. Rev.* **2023,** *52,* 5607–5651. doi:10.1039/d2cs00799a
6. Tao, Y.; Pan, J.; Yin, W.; Qiao, Q.; Xu, Z. *Chin. Chem. Lett.* **2026,** *37,* 112611. doi:10.1016/j.cclet.2026.112611
7. Liu, W.; Chen, J.; Qiao, Q.; Liu, X.; Xu, Z. *Chin. Chem. Lett.* **2022,** *33,* 4943–4947. doi:10.1016/j.cclet.2022.03.121
8. Grimm, J. B.; English, B. P.; Choi, H.; Muthusamy, A. K.; Mehl, B. P.; Dong, P.; Brown, T. A.; Lippincott-Schwartz, J.; Liu, Z.; Lionnet, T.; Lavis, L. D. *Nat. Methods* **2016,** *13,* 985–988. doi:10.1038/nmeth.4034
9. Ye, Z.; Yang, W.; Wang, C.; Zheng, Y.; Chi, W.; Liu, X.; Huang, Z.; Li, X.; Xiao, Y. *J. Am. Chem. Soc.* **2019,** *141,* 14491–14495. doi:10.1021/jacs.9b04893
10. Jiang, G.; Ren, T.-B.; D'Este, E.; Xiong, M.; Xiong, B.; Johnsson, K.; Zhang, X.-B.; Wang, L.; Yuan, L. *Nat. Commun.* **2022,** *13,* 2264. doi:10.1038/s41467-022-29547-3
11. Wang, L.; Tran, M.; D'Este, E.; Roberti, J.; Koch, B.; Xue, L.; Johnsson, K. *Nat. Chem.* **2020,** *12,* 165–172. doi:10.1038/s41557-019-0371-1
12. Chi, W.; Qi, Q.; Lee, R.; Xu, Z.; Liu, X. *J. Phys. Chem. C* **2020,** *124,* 3793–3801. doi:10.1021/acs.jpcc.9b11673
13. Deng, F.; Sun, D.; Yang, S.; Huang, W.; Huang, C.; Xu, Z.; Liu, L. *Spectrochim. Acta, Part A* **2022,** *268,* 120662. doi:10.1016/j.saa.2021.120662
14. Wang, Y.; Wang, X.; Ma, W.; Lu, R.; Zhou, W.; Gao, H. *Chemosensors* **2022,** *10,* 399. doi:10.3390/chemosensors10100399
15. Si, D.; Li, Q.; Bao, Y.; Zhang, J.; Wang, L. *Angew. Chem., Int. Ed.* **2023,** *62,* e202307641. doi:10.1002/anie.202307641
16. Wang, N.; Wang, T.; Fan, M.; Li, C.; Tian, Y.; Cui, X. *Aggregate* **2025,** *6,* e70077. doi:10.1002/agt2.70077
17. Pham, L. D.; Hight, M. O.; Wang, G.; Pimentel, A. E.; Haque, L.; Su, T. A. *J. Am. Chem. Soc.* **2025,** *147,* 39340–39350. doi:10.1021/jacs.5c11150
18. Wu, S.; Hou, D.; Qiao, Q.; Xu, Z. *Chem. Soc. Rev.* **2026,** *55,* 1935–1953. doi:10.1039/d5cs01304f
19. Zhou, W.; Fang, X.; Qiao, Q.; Jiang, W.; Zhang, Y.; Xu, Z. *Chin. Chem. Lett.* **2021,** *32,* 943–946. doi:10.1016/j.cclet.2021.02.003
20. Koide, Y.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T. *ACS Chem. Biol.* **2011,** *6,* 600–608. doi:10.1021/cb1002416
21. Koide, Y.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T. *J. Am. Chem. Soc.* **2011,** *133,* 5680–5682. doi:10.1021/ja111470n
22. Koide, Y.; Urano, Y.; Hanaoka, K.; Piao, W.; Kusakabe, M.; Saito, N.; Terai, T.; Okabe, T.; Nagano, T. *J. Am. Chem. Soc.* **2012,** *134,* 5029–5031. doi:10.1021/ja210375e
23. Lesiak, L.; Dadina, N.; Zheng, S.; Schelvis, M.; Schepartz, A. *ACS Cent. Sci.* **2024,** *10,* 19–27. doi:10.1021/acscentsci.3c01173
24. Zhou, X.; Miao, L.; Zhou, W.; Qiao, Q.; Xu, Z. *Chin. Chem. Lett.* **2025,** *36,* 110984. doi:10.1016/j.cclet.2025.110984
25. Myochin, T.; Hanaoka, K.; Iwaki, S.; Ueno, T.; Komatsu, T.; Terai, T.; Nagano, T.; Urano, Y. *J. Am. Chem. Soc.* **2015,** *137,* 4759–4765. doi:10.1021/jacs.5b00246
26. Zheng, Q.; Ayala, A. X.; Chung, I.; Weigel, A. V.; Ranjan, A.; Falco, N.; Grimm, J. B.; Tkachuk, A. N.; Wu, C.; Lippincott-Schwartz, J.; Singer, R. H.; Lavis, L. D. *ACS Cent. Sci.* **2019,** *5,* 1602–1613. doi:10.1021/acscentsci.9b00676
27. Wang, L.; Frei, M. S.; Salim, A.; Johnsson, K. *J. Am. Chem. Soc.* **2019,** *141,* 2770–2781. doi:10.1021/jacs.8b11134
28. Lardon, N.; Wang, L.; Tschanz, A.; Hoess, P.; Tran, M.; D'Este, E.; Ries, J.; Johnsson, K. *J. Am. Chem. Soc.* **2021,** *143,* 14592–14600. doi:10.1021/jacs.1c05004
29. Kompa, J.; Bruins, J.; Glogger, M.; Wilhelm, J.; Frei, M. S.; Tarnawski, M.; D'Este, E.; Heilemann, M.; Hiblot, J.; Johnsson, K. *J. Am. Chem. Soc.* **2023,** *145,* 3075–3083. doi:10.1021/jacs.2c11969
30. Uno, S.-n.; Kamiya, M.; Yoshihara, T.; Sugawara, K.; Okabe, K.; Tarhan, M. C.; Fujita, H.; Funatsu, T.; Okada, Y.; Tobita, S.; Urano, Y. *Nat. Chem.* **2014,** *6,* 681–689. doi:10.1038/nchem.2002
31. Werther, P.; Yserentant, K.; Braun, F.; Kaltwasser, N.; Popp, C.; Baalmann, M.; Herten, D.-P.; Wombacher, R. *Angew. Chem., Int. Ed.* **2020,** *59,* 804–810. doi:10.1002/anie.201906806
32. Qiao, Q.; Song, A.; An, K.; Xu, N.; Jia, W.; Ruan, Y.; Bao, P.; Tao, Y.; Zhang, Y.; Wang, X.; Xu, Z. *Angew. Chem., Int. Ed.* **2025,** *64,* e202417469. doi:10.1002/anie.202417469

33. Qiao, Q.; Song, A.; Jiang, G.; Zhou, Y.; Ruan, Y.; Jia, W.; Liu, X.; Xu, Z. *Angew. Chem., Int. Ed.* **2025,** *64,* e202503177. doi:10.1002/anie.202503177
34. Song, A.; Qiao, Q.; Xu, N.; Ruan, Y.; Jia, W.; Wang, X.; Xu, Z. *Chin. Chem. Lett.* **2025,** *36,* 110643. doi:10.1016/j.cclet.2024.110643
35. Chi, W.; Tan, D.; Qiao, Q.; Xu, Z.; Liu, X. *Angew. Chem., Int. Ed.* **2023,** *62,* e202306061. doi:10.1002/anie.202306061
36. Ohno, H.; Sasaki, E.; Yamada, S.; Hanaoka, K. *Org. Biomol. Chem.* **2024,** *22,* 3099–3108. doi:10.1039/d4ob00130c
37. Grimm, J. B.; Muthusamy, A. K.; Liang, Y.; Brown, T. A.; Lemon, W. C.; Patel, R.; Lu, R.; Macklin, J. J.; Keller, P. J.; Ji, N.; Lavis, L. D. *Nat. Methods* **2017,** *14,* 987–994. doi:10.1038/nmeth.4403
38. Fang, X.; Qiao, Q.; Li, Z.; Li, H.-K.; Huang, Y.; Hou, D.; Chen, J.; Xu, N.; An, K.; Jiang, W.; Tao, Y.; Bao, P.; Zhang, Y.; Wu, Z.; Liu, X.; Xu, Z. *J. Am. Chem. Soc.* **2025,** *147,* 22253–22267. doi:10.1021/jacs.5c08833
39. Xu, N.; Qiao, Q.; Wang, C.; Zhou, W.; Bao, P.; Li, J.; Wu, S.; Liu, X.; Xu, Z. *Angew. Chem., Int. Ed.* **2025,** *64,* e202505579. doi:10.1002/anie.202505579
40. Deng, F.; Xu, Z. *Chin. Chem. Lett.* **2019,** *30,* 1667–1681. doi:10.1016/j.cclet.2018.12.012
41. Chi, W.; Qiao, Q.; Wang, C.; Zheng, J.; Zhou, W.; Xu, N.; Wu, X.; Jiang, X.; Tan, D.; Xu, Z.; Liu, X. *Angew. Chem., Int. Ed.* **2020,** *59,* 20215–20223. doi:10.1002/anie.202010169
42. Wang, B.; Chai, X.; Zhu, W.; Wang, T.; Wu, Q. *Chem. Commun.* **2014,** *50,* 14374–14377. doi:10.1039/c4cc06178k
43. Butkevich, A. N.; Belov, V. N.; Kolmakov, K.; Sokolov, V. V.; Shojaei, H.; Sidenstein, S. C.; Kamin, D.; Matthias, J.; Vlijm, R.; Engelhardt, J.; Hell, S. W. *Chem. – Eur. J.* **2017,** *23,* 12114–12119. doi:10.1002/chem.201701216
44. Bucevičius, J.; Gilat, T.; Lukinavičius, G. *Chem. Commun.* **2020,** *56,* 14797–14800. doi:10.1039/d0cc06759h
45. Butkevich, A. N. *Org. Lett.* **2021,** *23,* 2604–2609. doi:10.1021/acs.orglett.1c00512
46. Lee, S.; Chung, C. Y.-S.; Liu, P.; Craciun, L.; Nishikawa, Y.; Bruemmer, K. J.; Hamachi, I.; Saijo, K.; Miller, E. W.; Chang, C. J. *J. Am. Chem. Soc.* **2020,** *142,* 14993–15003. doi:10.1021/jacs.0c05727

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.72

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Molecular structure of 3-*O*-acetyl-11-keto-β-boswellic acid and its L-alanine conjugate (AKBA-Ala) and their impact on lipoxygenases

Jan M. Peschel[1,2], Paul M. Jordan[2,3], Benjamin Friebe[1,2], Helmar Görls[4], Phil Köhler[4], Ulrich S. Schubert[1,2], Oliver Werz[2,3] and Michael Gottschaldt[*1,2]

Full Research Paper

Open Access

Address:
[1]Laboratory of Organic and Macromolecular Chemistry (IOMC), Friedrich Schiller University Jena, Humboldtstraße 10, 07743 Jena, Germany, [2]Jena Center for Soft Matter (JCSM), Friedrich Schiller University Jena, Philosophenweg 7, 07743 Jena, Germany, [3]Department of Pharmaceutical/Medicinal Chemistry, Institute of Pharmacy, Friedrich Schiller University Jena, 07743 Jena, Germany and [4]Institute of Inorganic and Analytical Chemistry (IAAC), Friedrich Schiller University Jena, Humboldtstraße 8, 07743 Jena, Germany

Email:
Michael Gottschaldt[*] - michael.gottschaldt@uni-jena.de

* Corresponding author

Keywords:
AKBA; amino acid functionalization; boswellic acids; 5-lipoxygenase; specialized pro-resolving mediators

Beilstein J. Org. Chem. **2026,** *22,* 1229–1236.
https://doi.org/10.3762/bjoc.22.99

Received: 12 May 2026
Accepted: 18 August 2026
Published: 03 September 2026

This article is part of the thematic issue "Design and synthesis of bioactive molecules".

Associate Editor: D. Spring



© 2026 Peschel et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

3-*O*-Acetyl-11-keto-β-boswellic acid (AKBA) is an anti-inflammatory constituent of frankincense. Its potency is attributed to its ability to interact with enzymes in the inflammation cascade, such as lipoxygenase and cyclooxygenase, to decrease formation of pro-inflammatory products while increasing the production of pro-resolving mediators. However, the availability of AKBA in vivo may be limited by factors such as poor solubility and protein binding in serum. A prodrug approach, wherein AKBA is conjugated to amino acids through its carboxylate moiety, could be a suitable way to improve solubility and half-life in blood while maintaining anti-inflammatory efficacy. In this study, we describe a robust protocol for the preparation of high-purity amino acid-AKBA conjugates, exemplified by synthesis of the L-alaninate derivative AKBA-Ala. The alanine conjugate was characterized by NMR spectroscopy, ESI-MS, HPLC, and X-ray crystallography, representing the first reported solid-state structure of an amino acid-AKBA conjugate. In comparison to AKBA, biological investigations of AKBA-Ala demonstrated a reduced inhibition (40% weaker at 10 μM) of 5-lipoxygenase (5-LOX) product formation in human neutrophils, but slightly enhanced production (120% higher at 10 μM) of specialized pro-resolving mediators (SPM) in human M2 macrophages. Overall, the anti-inflammatory potency of AKBA-Ala remains comparable to that of AKBA.

Introduction

Frankincense (also known as olibanum) is the oleogum resin of several tree species in the *Boswellia* genus, mostly native to arid climates in northeastern Africa, southern Arabia, and central India [1]. It has held great cultural significance since antiquity, from its use as incense in religious ceremonies to perfumes and various applications in traditional medicine [2]. Clinical trials of frankincense extracts have suggested promising efficacy in the treatment of inflammatory diseases such as asthma, rheumatoid arthritis, osteoarthritis, Crohn's disease, and collagenous colitis [3-5]. These favorable properties of the complex mixtures have prompted detailed investigation of the anti-inflammatory effects of the individual components and are largely attributed to their boswellic acid content [6,7]. Boswellic acids are a group of bioactive pentacyclic triterpenes found in *Boswellia* resin. The resin contains approximately 30% organic acids, and more than half of this fraction consists of boswellic acids, though the exact composition varies depending on species [8].

The most prominent boswellic acid derivative is 3-*O*-acetyl-11-keto-β-boswellic acid (AKBA), which interferes with several enzymes involved in the production of lipid mediators (LMs) that regulate the inflammatory response [4]. AKBA was identified as a potent inhibitor of arachidonate 5-lipoxygenase (5-LOX) [9] as well as cyclooxygenase-1 (COX-1) [10], which convert arachidonic acid (AA) into pro-inflammatory leukotrienes (LT) and prostaglandins (PG), respectively. Recent findings indicate a more complex interaction with LOXs, in which AKBA acts as an allosteric modulator of 5-LOX, promoting the production of specialized pro-resolving mediators (SPMs) over LTs by altering the regiospecificity of the enzyme [11,12]. AKBA also acts as an allosteric activator of 15-LOX-1, which contributes to the biosynthesis of SPM [13]. SPMs include a range of LMs derived from polyunsaturated fatty acids (PUFA, namely resolvins (Rv), protectins, and maresins) and lipoxins (LX), which contribute to the resolution of inflammation and tissue regeneration through different receptor-mediated pathways [14-16]. AKBA has also been found to block the nuclear factor kappa beta (NF-κB), resulting in reduced expression of associated pro-inflammatory, anti-apoptotic, proliferative, and angiogenic gene products [17,18]. Tumor-promoting inflammation has been dubbed one of the essential hallmarks of cancer [19]. Consequently, the anti-inflammatory properties of AKBA, in particular the NF-κB inhibition, render it a promising candidate for cancer therapy, with numerous studies demonstrating antiproliferative and pro-apoptotic effects, particularly in colorectal and prostate cancers [20,21].

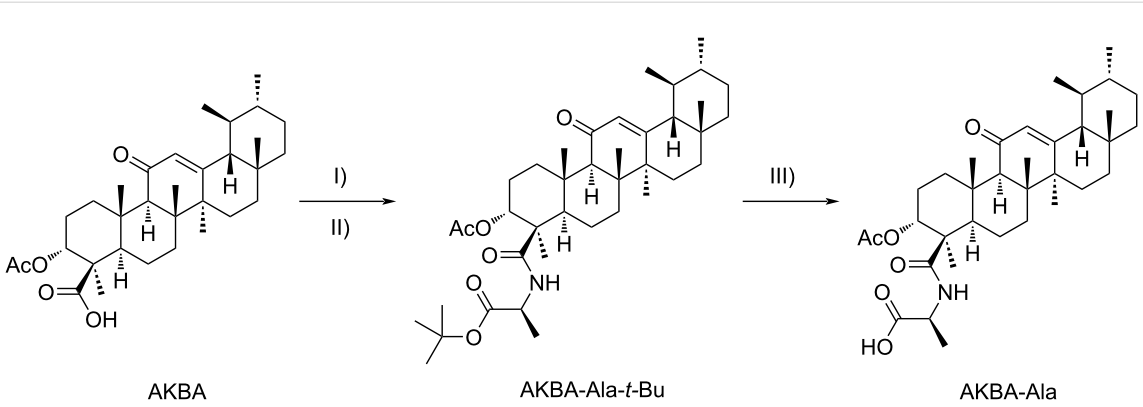
While AKBA strongly inhibits 5-LOX in various cells, the inhibitory effect is greatly diminished in the presence of whole blood or serum due to albumin binding [9]. Moreover, AKBA has a poor bioavailability due to low solubility, unfavorable absorption, and short half-life after oral intake [4]. To overcome barriers in drug delivery, a prodrug approach could be a simple solution that does not require sophisticated drug delivery systems. Conjugation of AKBA, e.g., to peptides or hydrophobic residues via click chemistry has been found to alter solubility, influence biodistribution, and modulate anti-inflammatory activity [22,23]. Notably, it has been demonstrated that peptide-AKBA conjugates with PepH3 and Angiopep2 connected through a cleavable ester are capable of brain-targeted drug delivery and release after intravenous administration, without the need for additional nanocarriers [23,24]. These findings make peptide-AKBA conjugates particularly interesting for targeted drug delivery. In comparison to conjugation via an ester linker, the direct functionalization of AKBA's carboxylate group with amino acids may increase its half-life in blood where amides are typically stable, and facilitate its release in protease-rich environments, such as macrophage lysosomes, tumor tissue, liver, or colon [25-27]. While the direct conjugation of AKBA to amino acids has been previously claimed by Naeini et al. [28], to us the synthetic procedure proposed therein appears to be inadequate for the synthesis and isolation of pure amino acid-AKBA conjugates. The analytical data provided suggest that little to no conjugates were formed, and no efforts were undertaken to purify the products.

Herein, we report a method for the preparation of high purity amino acid-AKBA conjugates exemplified by the synthesis of the L-alanine conjugate AKBA-Ala. The acquired conjugate was characterized by ESI-MS, NMR spectroscopy, and HPLC. Additionally, the solid-state molecular structures of AKBA and its alanine conjugate were determined by X-ray crystallography. While the molecular structure of AKBA has been previously reported (including a reported structure of a mixed crystal of AKBA and brominated AKBA) [29,30], to the best of our knowledge this constitutes the first reported solid-state molecular structure of any amino acid-AKBA conjugate. Finally, the capacity of AKBA-Ala to inhibit 5-LOX product formation in human neutrophils and to induce SPM biosynthesis in primary human M2 macrophages was evaluated in comparison to AKBA.

Results and Discussion

Synthesis of AKBA-Ala

The synthetic route to AKBA-Ala is displayed in Scheme 1. To introduce the amino acid, we chose commercially available *tert*-butyl-protected L-alanine hydrochloride. Protection of the carboxylate group minimizes side reactions in the amide coupling step and the *tert*-butyl group can be readily removed by



Scheme 1: Schematic representation of the synthesis of AKBA-Ala. I) $SOCl_2$, toluene, 20 °C, 18 h. II) L-Alanine *tert*-butyl ester hydrochloride, 4-(dimethylamino)pyridine, NEt_3, CH_2Cl_2, 0 °C → 20 °C, 72 h. III) Trifluoroacetic acid, CH_2Cl_2, 0 °C → 20 °C, 24 h.

trifluoroacetic acid under anhydrous conditions without risking hydrolysis of the C3α-acetate group.

For the amide coupling step, AKBA was first converted into its acid chloride by reaction with thionyl chloride in toluene. After removal of excess thionyl chloride and solvent, AKBA-Cl was obtained. The full conversion to the acid chloride was confirmed by an upfield shift of the carboxyl signal (C23: 182.15 ppm to 176.94 ppm) and a downfield shift of the nearby *ipso*-carbon signal (C4: 46.50 ppm to 56.91 ppm, Figure S7 in Supporting Information File 1). A downfield shift of signals belonging to nearby hydrogen atoms was also observed (Figure 1), and the carboxyl hydrogen signal disappeared (Supporting Information File 1, Figure S1). Due to air exposure, some hydrolysis occurred before the next reaction, as observed by ESI-MS and thin-layer chromatography (TLC).

AKBA-Cl was allowed to react with an excess (2 equiv) of *tert*-butyl-protected L-alanine hydrochloride in the presence of triethylamine as the base and catalytic amounts of 4-(dimethylamino)pyridine. The acylation proceeds slowly at room temperature owing to the steric bulk of AKBA-Cl, but complete conversion of AKBA-Cl was accomplished after several days. After chromatography some amount of contamination was present in the product fraction, most likely AKBA due to its similar polarity (R_f 0.59 for AKBA and 0.57 for product in 1:1 *n*-pentane/diethyl ether with 1% acetic acid). As the deprotected AKBA-Ala is substantially more polar any leftover AKBA should be removable by column chromatography after the next step. The successful introduction of the *t*-Bu-Ala moiety was confirmed by the presence of corresponding signals in the ^{1}H NMR spectrum (Figure 1) and ESI-MS (Supporting Information File 1, Figure S9). The ^{13}C-APT-NMR spectrum likewise shows new signals belonging to *t*-Bu-Ala (Supporting Information File 1, Figure S10). One notable concern during the coupling of chiral amino acids such as L-alanine is the possibility of racemization at the α-carbon. In the present case, the conditions for acid chloride coupling of *t*-Bu-Ala were unlikely to cause racemization of L-alanine. Only mild organic bases were employed, the reaction temperature was kept low, and the carboxyl group of L-alanine was not activated, thus avoiding the formation of more CH-acidic reactive intermediates such as oxazolones [31]. Furthermore, the ^{1}H and ^{13}C NMR spectra showed no additional signals that could be attributed to the alanine moiety beyond what is expected for a single diastereomer, suggesting that if any D-alanine conjugate was present, it was below the limit of detection for these techniques.

For the final step, AKBA-Ala-*t*-Bu was treated with 35% (v/v) trifluoroacetic acid in dry dichloromethane. The *tert*-butyl ester can be readily cleaved even in anhydrous conditions via an $A_{AL}1$ mechanism, while the 3-*O*-acetate group of AKBA remains intact. NMR analysis confirmed the complete removal of the *t*-Bu group through the absence of the corresponding signals in the ^{1}H and ^{13}C NMR spectra (Figure 1 and Figure S13 in Supporting Information File 1). In the extended ^{1}H NMR spectrum (Figure S1, Supporting Information File 1), the carboxylic acid signal appeared as a broad singlet at 9.0–7.5 ppm.

X-ray crystallography of AKBA and AKBA-Ala

Molecular structure in the crystal

The solid-state molecular structures of AKBA and AKBA-Ala are displayed in Figure 2 and Figure 3, the unit cell measurements of both compounds are compared in Supporting Information File 1, Table S1.

The configuration of AKBA matches literature results [29,30]. Briefly, the four cyclohexane rings containing C1–C18 feature an all equatorial arrangement with only the C17–C22 ring in

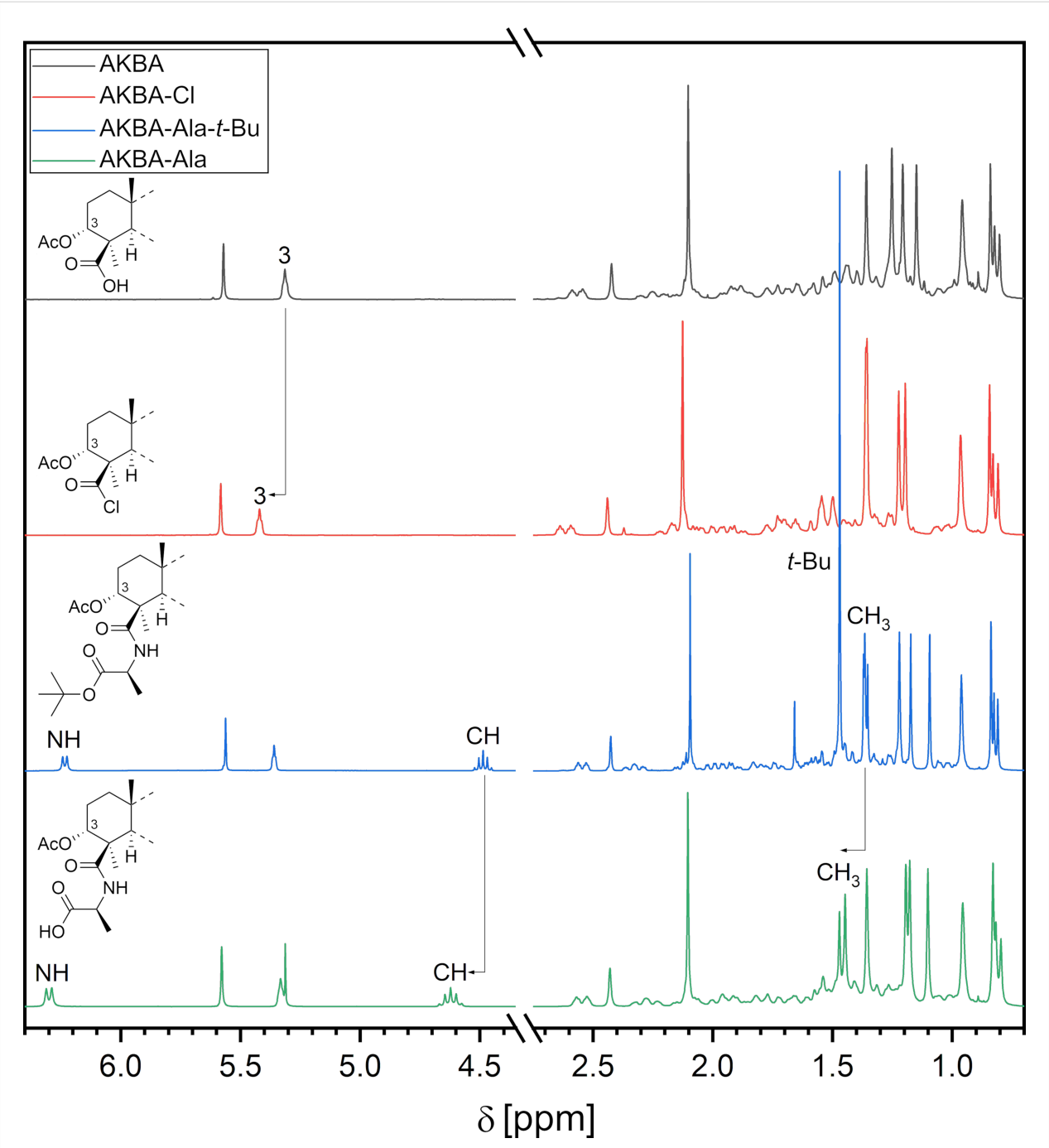


Figure 1: Stacked ^{1}H NMR spectra of all products (AKBA and AKBA-Cl: 300 MHz; AKBA-Ala-*t*-Bu and AKBA-Ala: 400 MHz, $CDCl_3$). For clarity, only a section of the molecules is displayed. For shift range 6.6 to 15 ppm see Supporting Information File 1, Figure S1.

axial position. The α,β-unsaturated ketone enforces a nearly planar arrangement of C9 and C11–14 in the third ring, all other rings feature chair conformation. The 4-carboxyl group and the 3-acetoxy group are both in axial position sticking out from the sheet-like structure of the ring system. The AKBA molecules are arranged in linear chains connected by hydrogen bonds between the hydroxy group of the carboxylic acid and the C=O of the 3-*O*-acetate moiety (O–O-distance 2.692 Å, Supporting Information File 1, Figure S16). Adjacent molecules in the chain adopt alternating orientations related by an approximately 180° rotation.

The solid-state molecular structure of AKBA-Ala (Figure 3) shows the newly introduced L-alanine moiety sticking out of the

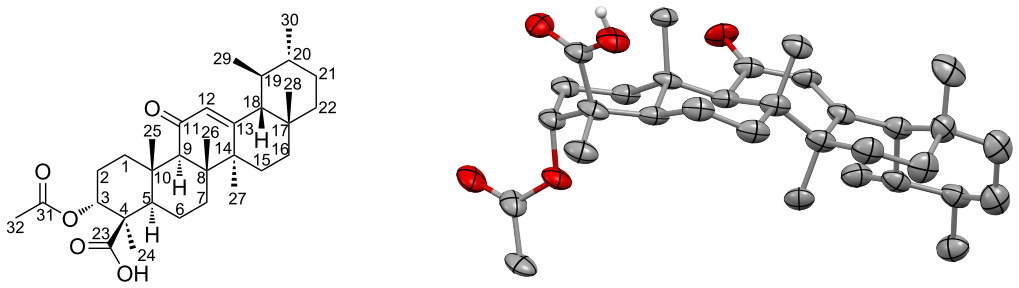


Figure 2: Displacement ellipsoid plot and atom-numbering scheme of AKBA (drawn at 30% probability; one of two molecules in the asymmetric unit). Hydrogen atoms attached to C-atoms are omitted for clarity. Colors: C (grey), H (white), O (red).

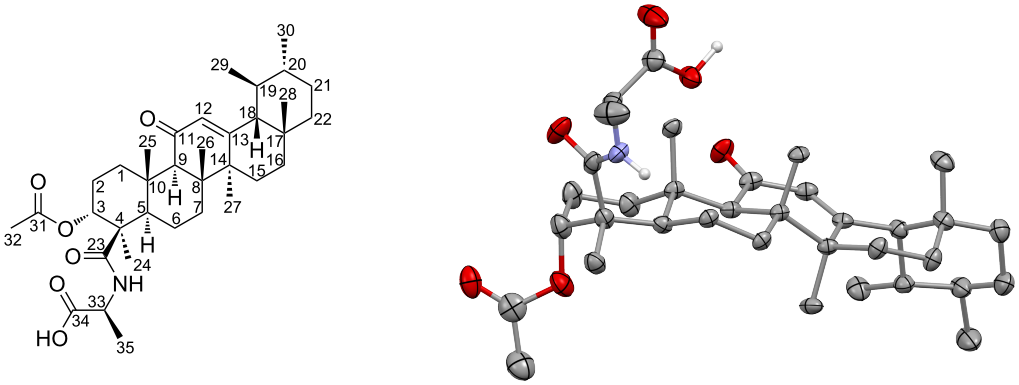


Figure 3: Displacement ellipsoid plot and atom-numbering scheme of AKBA-Ala (drawn at 50% probability). Hydrogen atoms attached to C-atoms are omitted for clarity. Colors: C (grey), H (white), O (red), and N (blue).

plane formed by the AKBA cyclohexane rings. Only minor changes in the conformation of AKBA were observed as a result of L-alanine introduction. The increased steric bulk was reflected in a slightly increased distance between the axial C25 and C23 (3.149 Å in AKBA-Ala, 3.092 Å in AKBA), pushing the 4-amide group outward (O–C3–C4–C23 torsion angle 164.3° in AKBA, 163° in AKBA-Ala). The AKBA-Ala molecules form hydrogen-bonded zigzag chains with alternating orientations of adjacent molecules. Unlike in AKBA, hydrogen bonds of the COOH group of the L-alanine moiety connect to the ketone at C11 rather than 3-*O*-acetate (O–O-distance 2.665 Å, Supporting Information File 1, Figure S17).

The wavelength (Mo-K_α = 0.71073 Å) did not allow distinction between AKBA and its enantiomer as no heavy atoms are present, so the absolute structure of AKBA could not be determined. In AKBA-Ala, because the configuration of the introduced alanine moiety was known to be the L-isomer, it could act as a chiral auxiliary revealing the configuration of the attached AKBA to be as displayed in Figure 3.

Biological evaluations of anti-inflammatory properties of AKBA-Ala

AKBA exhibits anti-inflammatory properties by allosterically inhibiting 5-LOX product formation [12]. Therefore, we first assessed the potency of AKBA-Ala to inhibit 5-LOX product formation in Ca^{2+}-ionophore A23187-stimulated human neutrophils in comparison with AKBA. At a concentration of 1 µM, neither AKBA-Ala nor AKBA affected 5-LOX product formation. However, at 10 µM, both compounds significantly reduced 5-LOX activity (Figure 4). Notably, AKBA demonstrated a stronger inhibitory effect, reducing 5-LOX product formation by approximately 85%, whereas AKBA-Ala inhibited 5-LOX activity by about 50% (Figure 4).

Besides the inhibition of pro-inflammatory LTs, AKBA induces the formation of 15-LOX-1 products and SPM in human primary M2 macrophages, leading to a faster resolution of the inflammatory process [13]. We therefore investigated whether AKBA-Ala similarly stimulates pro-resolving LM formation. At a concentration of 10 µM, AKBA-Ala increased the formation of eicosapentaenoic acid (EPA)-derived 15-HEPE

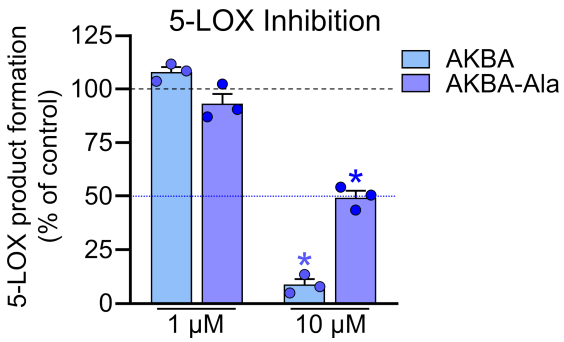


Figure 4: Inhibition of 5-LOX product formation by AKBA and AKBA-Ala in human neutrophils. Neutrophils were preincubated with AKBA or AKBA-Ala at 1 and 10 µM or control (0.1% DMSO) for 15 min at 37 °C and then stimulated with 2.5 µM A23187. After 10 min, the reaction was stopped, and 5-LOX products were extracted via solid-phase extraction (SPE) and analyzed with HPLC. Values are given as 5-LOX products (leukotriene B4 (LTB_4), trans-LTB_4, epi-trans-LTB_4 and 5-hydroxyeicosatetraenoic acid (5-HETE)) in percentage of control. For statistical analysis a paired t-test versus vehicle was used, * < 0.05; *n* = 3 individual experiments.

(15-hydroxyeicosapentaenoic acid) as well as docosahexaenoic acid (DHA)-derived 7-HDHA (7-hydroxydocosahexaenoic acid) and 17-HDHA (17-hydroxydocosahexaenoic acid) in human M2-MDM (monocyte-derived macrophages) to a comparable extent as AKBA (Figure 5). Interestingly, AKBA-Ala induced slightly higher levels of the SPM resolvin D5 (RvD5) compared with AKBA (Figure 5). Overall, these findings demonstrate that the anti-inflammatory biological potency of AKBA-Ala is comparable to that of AKBA, with only minor differences.

Conclusion

AKBA was functionalized with an L-alanine moiety at the carboxylate group through a series of three simple reactions. The successful synthesis and isolation of AKBA-Ala was verified by NMR spectroscopy, ESI-MS, and HPLC. Furthermore, the solid-state molecular structures of AKBA and AKBA-Ala were determined by X-ray crystallography. The use of a *tert*-butyl protecting group at the carboxyl group of the amino acid enables clean conjugation and subsequent deprotection by TFA while leaving the 3-*O*-acetyl group of AKBA intact. In direct comparison to AKBA, the conjugate AKBA-Ala revealed slightly decreased potency in human neutrophils to inhibit 5-LOX, but in human M2-MDM it increased the formation of inflammation-resolving LM, suggesting a strong anti-inflammatory potential.

This protocol should be applicable for the conjugation of any amino acid (or amino-acid sequence, including proteins) as long as its side chain tolerates the presence of acyl chloride. For amino acids with protic side chains (e.g., serine, lysine or cysteine) additional TFA-cleavable protecting groups at the side chain would be necessary, similar to what is universally applied in Fmoc-based solid-phase peptide synthesis [32].

Supporting Information

Supporting information features a list of abbreviations, detailed experimental procedures, ^{1}H and ^{13}C NMR spectra, ESI mass spectra, and HPLC data. Crystallographic data (excluding structure factors) has been deposited with the Cambridge Crystallographic Data Centre as supplementary publication CCDC-2517232 for AKBA, and CCDC-2517233 for AKBA-Ala. Copies of the data can be obtained free of charge on application to CCDC, 12 Union Road, Cambridge CB2 1EZ, UK [Email: deposit@ccdc.cam.ac.uk].

Supporting Information File 1

Experimental part and analytical data.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-22-99-S1.pdf]

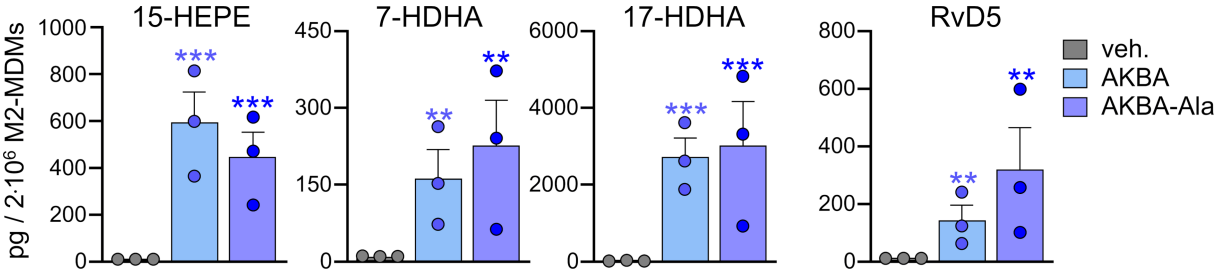


Figure 5: Induction of pro-resolving 15-LOX products and RvD5 by AKBA and AKBA-Ala in human M2-MDMs. M2-MDM (2 × 10^6 cells) were incubated with 10 µM AKBA, 10 µM AKBA-Ala, or vehicle (0.1% DMSO) for 180 min at 37 °C. The formed LM were quantified in the supernatants using UPLC-MS/MS. 15-HEPE, 7-HDHA, 17-HDHA, and RvD5 are shown as pg/2 × 10^6 M2-MDMs in mean ± S.E.M and in single values. Data were log-transformed for statistical analysis; one-way ANOVA with Dunnett's multiple comparisons test, $^{**}p < 0.01$; $^{***}p < 0.001$ versus vehicle.

Funding

This study was funded by the AiF Projekt GmbH (ZiM-FuE-Kooperationsprojekt KK5064604SK4). Further financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project No. 316213987, SFB 1278 PolyTarget (Project No. A04) is acknowledged.

Author Contributions

Jan M. Peschel: investigation; visualization; writing – original draft. Paul M. Jordan: investigation; visualization; writing – review & editing. Benjamin Friebe: investigation; writing – review & editing. Helmar Görls: investigation; writing – review & editing. Phil Köhler: investigation; writing – review & editing. Ulrich S. Schubert: funding acquisition; writing – review & editing. Oliver Werz: conceptualization; funding acquisition; investigation; writing – review & editing. Michael Gottschaldt: conceptualization; funding acquisition; supervision; writing – review & editing.

ORCID® iDs

Jan M. Peschel - https://orcid.org/0009-0002-4787-2757
Paul M. Jordan - https://orcid.org/0000-0002-8364-4236
Benjamin Friebe - https://orcid.org/0009-0003-9769-6314
Phil Köhler - https://orcid.org/0000-0002-4660-1691
Ulrich S. Schubert - https://orcid.org/0000-0003-4978-4670
Michael Gottschaldt - https://orcid.org/0000-0003-4210-2956

Data Availability Statement

Data generated and analyzed during this study are available from the corresponding author upon reasonable request.

References

1. Report on Boswellia distribution; Center for Middle Eastern Plants; Frankincense Resource Portal; Royal Botanic Garden Edinburgh. https://padme.rbge.org.uk/boswellia/files/03%20Boswellia%20DISTRIBUTION%20.pdf.
2. Al-Harrasi, A.; Khan, A. L.; Asaf, S.; Al-Rawahi, A. Frankincense and Human Civilization: A Historical Review. In *Biology of Genus Boswellia;* Al-Harrasi, A.; Khan, A. L.; Asaf, S.; Al-Rawahi, A., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp 1–9. doi:10.1007/978-3-030-16725-7_1
3. Ernst, E. *BMJ [Br. Med. J.]* **2008,** *337,* a2813. doi:10.1136/bmj.a2813
4. Abdel-Tawab, M.; Werz, O.; Schubert-Zsilavecz, M. *Clin. Pharmacokinet.* **2011,** *50,* 349–369. doi:10.2165/11586800-000000000-00000
5. Al Lawati, A.; Waladwadi, T.; Alhabsi, A.; Al Rawahi, N.; Al Breiki, F.; Al Lawati, M.; Das, S. *Nat. Prod. J.* **2026,** *16,* e22103155347241. doi:10.2174/0122103155347241241217042413
6. Efferth, T.; Oesch, F. *Semin. Cancer Biol.* **2022,** *80,* 39–57. doi:10.1016/j.semcancer.2020.01.015
7. Peng, C.; Yang, Y.; Wang, Y.; Gong, B.; Sun, X.; Yang, X. *Front. Pharmacol.* **2025,** *16,* 1692443. doi:10.3389/fphar.2025.1692443
8. Paul, M. Chemotaxonomic investigations on resins of the frankincense species Boswellia papyrifera, Boswellia serrata and Boswellia sacra, respectively, Boswellia carterii: A qualitative and quantitative approach by chromatographic and spectroscopic methodology. Ph.D. Thesis, Saarland University, Germany, Dec 3, 2012. doi:10.22028/d291-22839
9. Siemoneit, U.; Pergola, C.; Jazzar, B.; Northoff, H.; Skarke, C.; Jauch, J.; Werz, O. *Eur. J. Pharmacol.* **2009,** *606,* 246–254. doi:10.1016/j.ejphar.2009.01.044
10. Siemoneit, U.; Hofmann, B.; Kather, N.; Lamkemeyer, T.; Madlung, J.; Franke, L.; Schneider, G.; Jauch, J.; Poeckel, D.; Werz, O. *Biochem. Pharmacol.* **2008,** *75,* 503–513. doi:10.1016/j.bcp.2007.09.010
11. Gilbert, N. C.; Newcomer, M. E.; Werz, O. *Biochem. Pharmacol.* **2021,** *193,* 114759. doi:10.1016/j.bcp.2021.114759
12. Gilbert, N. C.; Gerstmeier, J.; Schexnaydre, E. E.; Börner, F.; Garscha, U.; Neau, D. B.; Werz, O.; Newcomer, M. E. *Nat. Chem. Biol.* **2020,** *16,* 783–790. doi:10.1038/s41589-020-0544-7
13. Börner, F.; Pace, S.; Jordan, P. M.; Gerstmeier, J.; Gomez, M.; Rossi, A.; Gilbert, N. C.; Newcomer, M. E.; Werz, O. *Adv. Sci.* **2023,** *10,* 2205604. doi:10.1002/advs.202205604
14. Serhan, C. N.; Levy, B. D. *J. Clin. Invest.* **2018,** *128,* 2657–2669. doi:10.1172/jci97943
15. Sansbury, B. E.; Spite, M. *Circ. Res.* **2016,** *119,* 113–130. doi:10.1161/circresaha.116.307308
16. Recchiuti, A.; Mattoscio, D.; Isopi, E. *Front. Pharmacol.* **2019,** *10,* 252. doi:10.3389/fphar.2019.00252
17. Takada, Y.; Ichikawa, H.; Badmaev, V.; Aggarwal, B. B. *J. Immunol.* **2006,** *176,* 3127–3140. doi:10.4049/jimmunol.176.5.3127
18. Xu, F.; Li, W.; Zheng, X.-J.; Hao, Y.; Yang, Y.-H.; Yang, H.; Zhang, S.; Cao, W.-X.; Li, X.-X.; Zhang, X.; Du, G.-H.; Ji, T.-F.; Wang, J.-H. *Oncol. Res.* **2025,** *33,* 1969–1989. doi:10.32604/or.2025.062386
19. Hanahan, D. *Cancer Discovery* **2022,** *12,* 31–46. doi:10.1158/2159-8290.cd-21-1059
20. Roy, N. K.; Deka, A.; Bordoloi, D.; Mishra, S.; Kumar, A. P.; Sethi, G.; Kunnumakkara, A. B. *Cancer Lett.* **2016,** *377,* 74–86. doi:10.1016/j.canlet.2016.04.017
21. Trivedi, V. L.; Soni, R.; Dhyani, P.; Sati, P.; Tejada, S.; Sureda, A.; Setzer, W. N.; Faizal Abdull Razis, A.; Modu, B.; Butnariu, M.; Sharifi-Rad, J. *Front. Pharmacol.* **2023,** *14,* 1187181. doi:10.3389/fphar.2023.1187181
22. Meyiah, A.; Avula, S. K.; Al-Harrasi, A.; Elkord, E. *Pharmaceuticals* **2024,** *17,* 1650. doi:10.3390/ph17121650
23. Cui, N.; Fan, T.; Li, L.; Liu, W.; Li, M.; Huang, M.; Wu, Q.; Wang, J.; Zhang, W.; Wang, S.; Ding, Y. *Chem. Eng. J.* **2025,** *509,* 160664. doi:10.1016/j.cej.2025.160664
24. Meng, Q.; Huang, C.; Bai, M.; Liao, Y.; Zhang, K.; Wang, Y.; Wang, J.; Yang, L.; Ding, Y. *Chem. Eng. J.* **2025,** *526,* 170989. doi:10.1016/j.cej.2025.170989
25. Vale, N.; Ferreira, A.; Matos, J.; Fresco, P.; Gouveia, M. J. *Molecules* **2018,** *23,* 2318. doi:10.3390/molecules23092318
26. Shah, K.; Gupta, J. K.; Chauhan, N. S.; Upmanyu, N.; Shrivastava, S. K.; Mishra, P. *Open Med. Chem. J.* **2017,** *11,* 146–195. doi:10.2174/1874104501711010146
27. Poreba, M. *FEBS J.* **2020,** *287,* 1936–1969. doi:10.1111/febs.15227
28. Esmaeili, S.; Naeini, F. B.; Nezhati, M. N.; Afrasiabi, A.; Modaresi, S. M. S.; Dadras, A.; Riazi, G. H. *Int. J. Pharm. Sci. Rev. Res.* **2014,** *28,* 200–206.
29. Schweizer, S.; Eichele, K.; Ammon, H. P. T.; Safayhi, H. *Planta Med.* **2000,** *66,* 781–782. doi:10.1055/s-2000-9614

30. Ito, M.; Misao, K.; Suzuki, H.; Noguchi, S. *Chem. Pharm. Bull.* **2025,** *73,* 314–317. doi:10.1248/cpb.c24-00819
31. Kemp, D. S. Chapter 7 - Racemization in Peptide Synthesis. In *Major Methods of Peptide Bond Formation;* Gross, E.; Meienhofer, J., Eds.; Academic Press: New York, U.S.A., 1979; Vol. 1, pp 315–383. doi:10.1016/b978-0-12-304201-9.50013-8
32. Isidro-Llobet, A.; Álvarez, M.; Albericio, F. *Chem. Rev.* **2009,** *109,* 2455–2504. doi:10.1021/cr800323s

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (https://www.beilstein-journals.org/bjoc/terms), which is identical to the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
https://doi.org/10.3762/bjoc.22.99