HOKKAIDO UNIVERSITY

Title	Development of N-terminal specific dual modification strategies of proteins using triazolecarbaldehyde derivatives
Author(s)	張, 晏
Description	この博士論文の全文は利用可能日以降に公開されます。利用可能日以前の閲覧方法は https://www.lib.hokudai.ac.jp/dissertations/copy-guides/ をご参照ください。
Degree Grantor	北海道大学
Degree Name	博士(環境科学)
Dissertation Number	甲第16736号
Issue Date	2026-03-25
DOI	https://doi.org/10.14943/doctoral.k16736
Doc URL	https://hdl.handle.net/2115/99581
Type	doctoral thesis
File Information	ZHANG_Yan_summary.pdf, 全文の要約

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学 位 論 文 内 容 の 要 約/Dissertation Summary

博士（環境科学）　　氏　名/Applicant　張 晏

学 位 論 文 題 名/Title of Dissertation

Development of N-terminal specific dual modification strategies of proteins using triazolecarbaldehyde derivatives

(トリアゾールカルボアルデヒド誘導体を用いたタンパク質 N 末端特異的二重修飾法の開発)

Proteins play diverse roles in living systems, including the regulation of chemical reactions and the transport of substances. In addition to elucidating their structures and functions, there has been active research on applied studies that exploit their functional properties. For example, the construction of protein-based conjugates, such as enzyme electrodes with highly redox-active enzymes, has been investigated across a wide range of fields. Other example involves antibody–drug conjugates that links antibodies possessing antigen-recognition ability with therapeutic small molecules for medicine.

Rational design of protein conjugates requires consideration of several factors. The protein should contain a reactive group that enables selective attachment of the modifying reagent while minimizing undesired side reactions. Linker chemistry requires efficiency and compatibility with physiological conditions so that conjugation proceeds in mild aqueous environments without protein denaturation. The modification should preserve the native structure and biological function of the protein to maintain the desired activity and stability.

Proteins consist of 20 amino acids, and the side chains of these amino acids contain functional groups with different chemical properties. As a result, proteins have numerous potential reactive sites. Nucleophilic side-chain groups are commonly targeted in modification reactions, as their reactivity can be tuned by differences in pK_a values. Common examples of modification sites include the ε-amino group of lysine side chains and the thiol group of cysteine side chains. Because lysine residues are often abundant, they are commonly used when random labeling or multiple modification sites are acceptable. Meanwhile, cysteine side chains possess the most nucleophilic thiol group among amino acids in proteins and thus have been widely exploited in protein modification strategies. A representative example is the Michael addition to maleimide groups, which proceeds rapidly under mild conditions but often yields conjugates with limited stability.

The protein N-terminus has attracted considerable attention as a reactive site for site-specific modification reactions owing to its universality and uniqueness. In addition, the fact that the N-termini of many proteins are exposed on the molecular surface facilitates efficient reactions with modifying

agents, while minimizing the impact of modification on the structure and function of proteins. Although competition with the ε-amino group of lysine side chains often poses a problem, the pK_a value of the α-amino group at the protein N-terminus (6–8) is lower than that of the ε-amino group of lysine side chains (9–10). Consequently, adjusting the pH of the reaction solution can improve chemoselectivity.

Onoda and co-workers have reported that 1*H*-1,2,3-triazole-4-carbaldehyde (TA4C) derivatives react with the α-amino group at the protein N-terminus in a manner analogous to 2-pyridinecarbaldehyde (2-PCA), leading to the formation of an imidazolidinone ring. This N-terminal modification reaction exhibits a modification efficiency and N-terminal selectivity comparable to those of 2-pyridine carbaldehyde. TA4C derivatives can be synthesized in two steps from azides, including a copper(II)-catalyzed azide–alkyne cycloaddition (CuAAC) reaction, or in a single step from primary amines via a Dimroth rearrangement.

Site-specific dual modification of proteins is important for enhancing the functions of protein-based conjugates and for creating novel functional materials. Recent advances in site-specific dual modification have led to the development of diverse strategies, which can be broadly classified into three categories: unnatural amino acid–mediated methods, enzymatic methods, and purely chemical methods.

Hanaya and co-workers reported an operationally simple Cu(II)-mediated aldol reaction using 2-PCA derivatives to achieve highly site-selective N-terminal modification of peptides and proteins under physiological pH and ambient temperature. Cu(II) coordinates with the Schiff base formed between the N-terminal amino group and 2-PCA, activating the α-carbon to generate a nucleophilic intermediate that reacts with another molecule of 2-PCA in this reaction. Thereby this method could install two identical functional units from the same reagent at a single N-terminus via stable C–C bonds. This method is compatible with various N-terminal residues and has been successfully applied to multiple proteins, including biopharmaceuticals such as filgrastim and trastuzumab. These results demonstrate the robustness and biocompatibility of the approach. Importantly, the modified trastuzumab retained its HER2 binding activity, indicating that the Cu(II)-mediated process does not compromise the native structure or biological function of the protein. However, this approach also exhibits several intrinsic limitations. The reaction outcome is highly dependent on substrate structure: short peptides predominantly afford dually modified products (~98 %), whereas proteins and long peptides mainly yield singly modified species (~74 %) with only a minor proportion of dual modification (~32 %).

Machida and Kanemoto developed a Cu(I)-catalyzed three-component [3+2] cycloaddition method that enables N-terminal–specific dual modification of Gly in a single step using aldehydes and maleimides. This method maintains high selectivity for the N-terminus of iminopeptides even in the presence of highly nucleophilic lysine ε-amines, affording the cycloadducts in excellent yields. The reaction exhibits broad functional group tolerance and can rapidly construct doubly functionalized peptides using readily available aldehyde and maleimide modules. It is not directly applicable to the

modification of water-soluble proteins, because the reaction requires organic solvent conditions. Similarly, Hanaya et al. reported a copper(II)-mediated [3+2] cycloaddition approach that enables one-step, site-specific dual functionalization of diverse protein N-termini using commercially available maleimides and 2-PCA derivatives. This method operates under mild, non-denaturing aqueous conditions (pH 6.0, 37 °C) and does not require genetic engineering. It enables efficient attachment of a wide range of functional groups to diverse N-terminal amino acids.

Protein–protein conjugates are hybrid biomolecular assemblies composed of two or more protein domains that are covalently or non-covalently linked together in a defined manner. Such conjugates enable the precise integration of distinct protein functionalities within a single macromolecular framework. Proteins with complementary or even mutually exclusive biological activities can exhibit emergent and synergistic behavior when they are placed in close proximity. These effects are often difficult or impossible to obtain when the proteins function independently. The design of protein–protein conjugates allow researchers to tailor multifunctional systems with enhanced specificity, stability, and biological performance. Representative examples include antibody–drug conjugates (ADCs), which covalently link monoclonal antibodies with cytotoxic small molecules through chemical linkers to achieve targeted drug delivery in cancer therapy. Another well-known example is the enzyme-linked immunosorbent assays (ELISAs), which exploit antibody–enzyme conjugation for highly sensitive biomolecular detection. These examples illustrate how rational protein–protein conjugation can bridge therapeutic and diagnostic applications, offering a versatile platform for the development of advanced bioconjugate technologies. The development of strategies that enable site-selective and stable protein–protein linkage is crucial. Such approaches support the rational design of complex biomolecular architectures with tunable structure and function.

Chemical modification of native amino acid side chains using crosslinkers enables protein conjugation without genetic engineering. Heterobifunctional linkers such as succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) could form covalent bridges between lysine and cysteine residues. SMCC exploits the high abundance of lysine residues and the strong thiol reactivity of cysteine to achieve crosslinking. However, because lysine residues are abundant on protein surfaces, achieving high site specificity with such linkers is inherently challenging. A widely used homo-bifunctional method is cysteine–maleimide conjugation, a ubiquitous method in chemical biology for covalently linking two thiols, as exemplified by bismaleimide crosslinkers. This approach benefits from the high nucleophilicity of cysteine thiols, and the rapid and efficient nature of the Michael addition reaction. Cysteine residues are generally present at low abundance in proteins, so site-specific modification often requires introducing an additional cysteine by genetic engineering. This step increases experimental complexity and preparation time.

This dissertation establishes a TA4C–centered platform for site-selective N-terminal dual modification of proteins under mild aqueous conditions. The intrinsic N-terminal reactivity of TA4C derivatives is exploited to enable the chemoselective installation of two functional groups at a single protein N-terminus, thereby generating stable and multifunctional protein conjugates. The strategies

developed herein integrate metal-mediated reactivity control, cycloaddition-based dual functionalization, and bifunctional linker design within a unified chemical framework. This integrated approach extends TA4C chemistry from homogeneous dual modification to hetero-dual labeling and further to covalent protein–protein conjugation, without the need for genetic manipulation.

In Chapter 2, the Ni(II)-mediated N-terminal dual modification of proteins using TA4C derivatives was investigated in detail with respect to reaction selectivity, condition dependence, substrate scope, product stability, and reaction mechanism.

The influence of different metal ions on the dual modification reaction was systematically evaluated using green fluorescent protein (GFP) and benzyl-substituted TA4C as the substrate and aldehyde derivative, respectively. Reactions were performed in the presence of 50 equivalents of Zn^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, Mn^{2+}, Sc^{3+}, V^{4+}, and Ni^{2+}. The results show that only Co^{2+} and Ni^{2+} were able to promote the reaction, whereas the other metal ions exhibited no catalytic effect. Ni^{2+} shows the highest yield, confirming its role as the optimal metal catalyst for the dual modification among them. A reasonable mechanistic framework suggests that the superior performance of Ni^{2+} may be attributed to its moderate coordination strength, which is conducive to the formation of a reactive Schiff base metal complex. Its Lewis acidity may further facilitate deprotonation at the α-position, rendering the intermediate enolate-like in character and thereby favoring the addition of a second TA4C molecule.

The author further refined the reaction conditions by examining the effect of Ni^{2+} concentration using 10, 20, 50, 100, and 200 equiv. The results indicate that the optimal range for Ni^{2+} lies between 20 and 50 equiv., since the conversion does not improve in higher concentrations of Ni. In addition, different nickel salts, including nickel acetate, nickel chloride, and nickel sulfate were compared at 50 equiv. No appreciable differences in dual modification yield were observed. These findings indicate that Ni^{2+} is the key catalytic species in TA4C-mediated dual modification.

All proton signals are assigned by ^{1}H NMR and COSY experiments in DMSO-d_6. Amide NH protons appear at 8.69 and 8.16 ppm, and aromatic protons from the benzene and triazole rings appear at 7.47–7.27 ppm. The benzyl CH_2 is detected at 5.65–5.59 ppm, and methylene protons adjacent to –NH appeared at 3.72 and 3.62 ppm. Resonances corresponding to Leu are identified as β-proton (2.66 ppm), χ-proton (1.97–2.02 ppm) and δ-proton (0.81–0.85 ppm). The β-proton of Leu correlates only with its χ-proton in the COSY spectrum, suggesting the presence of an adjacent quaternary carbon. Moreover, distinct –NH groups adjacent to methylene protons indicates that imidazolidine ring formation does not occur. A correlation is observed between the proton of the methylene directly attached to the amide NH of Leu and its corresponding carbon (–NH–CH_2) in the HSQC spectrum. The signals of aromatic carbons in the phenyl ring and triazole appear between 128.34–136.31 ppm, while amide and carboxylate carbons are observed at 161.57, 168.24 and 171.75 ppm in the HMBC spectrum. HMBC analysis further reveals a long-range correlation between the Leu methylene protons and a downfield carbon resonance at 198.67 ppm, confirming the presence of a ketone functionality in the dually modified structure.

The dual modification was first validated using biologically active peptides with azide-substituted

TA4C as the N-terminal modification reagent. Peptides including angiotensin I, bovine adrenal medullary decapeptide, delta sleep-inducing peptide (DSIP), and neurotensin-13 were tested. Dual modification was confirmed on each peptide, with conversion rates of 14.0%, 70.8%, 53.4%, and 11.1%, respectively, as determined by ESI-MS analysis.

The scope was extended to proteins, encouraged by these results. Genetically expressed GFP-His_6 and Cut-His_6 were selected as model proteins, with C-terminal His-tags introduced for purification purposes. Dual modification with azide-substituted TA4C under standard conditions (37 °C, 16 h) afforded conversions of 33.6 % and 42.6 %, respectively, as confirmed by LC-MS. The impact of His-tag positioning and presence on dual modification was evaluated by examining two additional GFP variants, namely His_6-GFP with an N-terminal His-tag and GFP-Strep without a His-tag. Both variants undergo single modification in the absence of Ni^{2+} with conversion rates of 72.8 % (His_6-GFP) and 64.9 % (GFP-Strep). Ni^{2+} addition results in no detectable modification for His_6-GFP, whereas GFP-Strep shows single and dual modification ratios of 30.7 % and 41.3 %, respectively. These findings suggest that Ni^{2+} coordinates with the N-terminal His-tag of His_6-GFP, thereby preventing the formation of the Ni/TA4C/N-terminus complex required for modification. This inhibitory effect explains the complete suppression of dual modification and at the same time confirms that the reaction occurs specifically at the N-terminus.

Competition experiments provide further evidence supporting the suppressive role of His-tags. When GFP-Strep was reacted with azide-substituted TA4C and Ni^{2+} in the presence of 1 equiv. of His_6 peptide, the dual modification yield decreased to 12.5 %. In contrast, addition of a His-free tripeptide (Leu-Gly-Gly) increased the yield to 54.2 %. These results indicate that a His-tag strongly suppresses the yield of dual modification by binding to Ni^{2+}. They also show that the N-terminus is the exclusive reaction site for TA4C-mediated dual modification.

Dual modification was tested in different types of buffers and pH using GFP-His_6 as the substrate. Reactions were carried out in phosphate buffer at pH 6.0, 7.0, 7.5, and 8.0 (10 mM), in Tris–HCl buffer at pH 7.1, 8.0, and 8.9 (10 mM), and in phosphate-buffered saline (1 × PBS, pH 7.4). PBS gives one of the good outcomes among the examined buffers under the tested conditions.

The author investigated the reaction pathway leading to dual modification and evaluated the relative rates of single versus dual modification by reacting GFP-His_6 with 200 equivalents of TA4C and 100 equiv. of nickel acetate under different temperatures (4 °C, 25 °C and 37 °C) and time points (2, 4, 8, and 16 h). The dual modification product was observed at 4 °C by LC–MS analysis, suggesting that dual modification proceeds preferentially under low-temperature conditions and follows a pathway distinct from that of single modification. Divergent outcomes were observed at higher temperatures. The single modification reaction is more efficient at 25 °C than at 37 °C over 16 h, whereas the dual modification reaction is suppressed under the same conditions. These findings indicate that rapid progression of the single modification pathway interferes with dual modification, thereby revealing a competitive relationship between the two processes.

The author evaluated the effect of reagent stoichiometry by treating Cut-His_6 with increasing

amounts of benzyl-substituted TA4C (10–200 eq.) in the presence of $Ni(OAc)_2$ at pH 7.4 and 37 °C for 16 h. Dual modification is strongly favored over single modification under all tested conditions, as confirmed by LC–MS analysis. While the yield of singly modified species remained nearly constant (16–19 %), the proportion of dually modified Cut-His_6 increased significantly with higher equivalents of benzyl-substituted TA4C, reaching 71 % at 200 eq. These results demonstrate that Ni^{2+}-mediated dual modification is highly efficient and dominates over single modification as the reagent concentration increases.

N-terminal dual modification reactions were performed on GFP-His_6 and Cut-His_6 using a series of TA4C derivatives bearing either hydrophilic or hydrophobic substituents to evaluate the substrate scope and substituent tolerance of the developed method. All TA4C derivatives successfully reacted with both proteins under the optimized reaction conditions, affording the corresponding dually modified products as confirmed by LC-MS analysis. These results clearly demonstrate that the dual modification platform is broadly applicable and chemically compatible with a variety of hydrophilic and hydrophobic substituents on the triazolecarbaldehyde scaffold.

Density functional theory (DFT) calculations were performed to elucidate the reaction mechanism of TA4C-mediated dual modification using Gaussian16 at the B3LYP-D3 level of theory (SDD for Ni; 6-31+G(d) for H, C, N, O, and Cl). Solvent effects were considered using the PCM model with water ($\varepsilon = 78.4$) and an SAS cavity, and thermodynamic corrections were evaluated at 310 K and 1 atm. The computational results demonstrate that Ni(II) stabilizes the Schiff base intermediate formed between TA4C and the N-terminal amino group, thereby lowering the free energy barrier for the subsequent C–C bond-forming step. This stabilization provides theoretical support for the thermodynamic favorability of the dual modification pathway.

The applicability of the dual modification strategy was examined through fluorescent labeling of Cut-His_6 and antibody trastuzumab. The reaction introduced an azide-substituted TA4C derivative at the N-terminus under Ni^{2+}-mediated conditions, generating a dually azide-modified intermediate that subsequently reacted with DBCO–TAMRA via SPAAC to incorporate two fluorescent dyes. The observations show that TA4C-based N-terminal dual modification offers a reliable route for installing multiple functional motifs and generates protein conjugates with markedly improved fluorescence intensity.

The author evaluated the stability of the dually modified species by selecting Cut-His_6 as a model protein and reacting it with hexyl-substituted TA4C (C6-TA4C) in the presence of Ni^{2+}. The reaction produced efficient N-terminal dual modification using 200 equiv. of C6-TA4C at 37 °C for 16 h. The reaction mixture was incubated at 37 °C for one week after removal of excess reagent, and the dual modification yield was monitored every 24 h by LC–MS. The modification levels at each time point were 68 %, 56 %, 48 %, 41 %, 37 %, 32 %, 29 %, and 26 %, corresponding to a calculated half-life of 121 h. These results demonstrate that dual-modified conjugates maintain substantial stability over several days under physiological conditions.

This study established a Ni^{2+}-mediated TA4C method that enables selective dual N-terminal

modification of proteins. The method worked across peptides, proteins, and antibodies and showed clear site specificity at the N-terminus. Reaction outcomes are influenced by buffer conditions, Ni^{2+} amount, and the presence of N-terminal His-tags. Dually modified proteins exhibit greater hydrolytic and thermal stability compared to singly modified species and support efficient fluorescent labeling. Mechanistic experiments and DFT calculations indicate that Ni^{2+} stabilizes the key imine intermediate and promotes a reaction pathway distinct from single modification. Together, these findings show that TA4C-based dual modification provides a practical and broadly applicable platform for protein functionalization.

In chapter 3, the author will describe the development and mechanistic investigation of copper-mediated TA4C/maleimide dual modification of proteins. Particular emphasis will be placed on comparison between TA4C and the reported 2-PCA chemistry, highlighting differences in reactivity, pH-dependence, and hydrolytic stability. Furthermore, applications of this dual modification method in protein–protein conjugation and construction of multifunctional protein assemblies will be discussed, underscoring its potential utility in bioconjugate chemistry and chemical biology.

Evaluation using bovine adrenal medulla peptide 12P (BAM-12P) showed that copper(II) is required for TA4C-based dual modification. The peptide was reacted with benzyl-substituted TA4C and N-ethylmaleimide under standard conditions (10 mM potassium phosphate buffer, pH 6.0, 37 °C, 16 h) in the presence of various metal ions (Zn^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, Mn^{2+}, Sc^{3+}, V^{4+}, Ni^{2+}).The dually modified product is obtained only in the presence of Cu^{2+}, whereas reactions with Ni^{2+}, Co^{2+}, Mn^{2+}, Zn^{2+}, or Fe^{3+} predominantly yielded the singly modified species. No reaction occurs with Sc^{3+} or V^{4+}. These results clearly demonstrate that Cu^{2+} acts as the most effective additive for promoting the [3+2] cycloaddition between TA4C and maleimide derivatives at protein N-terminus.

The substrate generality of the TA4C/maleimide dual modification method was next evaluated using structurally diverse peptides. Representative bioactive sequences, angiotensin I and delta sleep-inducing peptide (DSIP), were selected to probe the applicability of the method beyond model systems. The reactions proceed with remarkable efficiency, affording the corresponding dually modified products in > 99 % conversion as determined by LC–MS analysis. Extension to protein substrates proceeded efficiently, affording dual modification of Cutinase (Cut) and superfolder GFP (sfGFP) with conversions of 84% and >85%, respectively. LC–MS analysis confirmed the formation of the dually labeled products as the predominant species, indicating high selectivity. These results clearly demonstrate that the TA4C/maleimide dual modification method is broadly applicable across different peptide frameworks, underscoring its robustness and utility for constructing multifunctional bioconjugates.

TA4C derivatives containing different linker structures were evaluated to find out the compatibility in the Cu^{2+}-mediated dual modification of proteins. The derivatives cover alkyne and azide groups as bioorthogonal handles, a hydrophilic polyethylene glycol (PEG) chain, and a hydrophobic alkyl substituent. All derivatives exhibit high reactivity, affording dually modified protein products with conversions ranging from 80 % to 92 %. These results demonstrate that the TA4C

framework tolerates a wide range of steric environments, including hydrophobic, hydrophilic, and bioorthogonal substituents. Such versatility underscores the robustness of the TA4C scaffold as a broadly applicable platform for installing diverse chemical functionalities at protein N-terminus, enabling facile construction of functional and orthogonally addressable bioconjugates.

Dual modification significantly enhanced the hydrolytic stability of N-terminal adducts. The modification yield of the singly modified protein decreased sharply from approximately 80% to below 5% after incubation at 37 °C for 48 h, whereas the dually modified protein remained largely intact over the same period, exhibiting only a minor decrease. These results clearly demonstrate that dual modification significantly enhances the hydrolytic stability of the N-terminal linkage.

The pH dependence of dual modification using 2-PCA and TA4C was evaluated over pH 5.0–8.0. While 2-PCA afforded similar dual modification yields across all pH values, these reactions were consistently accompanied by substantial non-specific lysine modification. In contrast, TA4C exhibited strong pH dependence, achieving the highest dual modification efficiency under acidic conditions and maintaining high N-terminal selectivity without detectable lysine modification.

Time-course analysis revealed clear kinetic differences between the two methods. Whereas 2-PCA achieved near-complete modification within 1 h, TA4C showed a substantially lower conversion over the same period. These results indicate that 2-PCA is intrinsically more reactive but suffers from poor selectivity, while TA4C exhibits slower kinetics in exchange for superior chemoselectivity. The reduced reactivity of TA4C may be attributed to weaker metal coordination by the triazole N3 site compared to the pyridyl nitrogen of 2-PCA.

The practical utility of dual modification with TA4C/maleimide was demonstrated by constructing multivalent protein assemblies. Cut was subjected to dual modification with azide-TA4C and biotinylated maleimide 7in the presence of Cu^{2+} under mild aqueous conditions (37 °C, 4 h, 10 mM potassium phosphate, pH 6.0), affording the dually functionalized intermediate Cut_1 in 56 % yield. Cut_1 underwent copper-free strain-promoted alkyne–azide cycloaddition (SPAAC) with sulfo-Cy5-DBCO, generating fluorescently labeled Cut_2. Only the SPAAC-derived product Cut_2 exhibited a detectable Cy5 fluorescence signal, whereas unmodified Cut and control samples showed no labeling, demonstrating the high specificity and efficiency of the two-step modification strategy.

Assembly formation between Sav and Cut variants was evaluated by native gel electrophoresis. Sav alone migrated as a tetramer, and no interaction was observed upon mixing with unmodified Cut. In contrast, Sav formed higher-molecular-weight assemblies with biotin-bearing Cut_1, appearing as a characteristic ladder corresponding to binding of one to four Cut molecules per Sav tetramer, confirming efficient biotin–Sav recognition. No mobility shift was detected in control experiments lacking biotin, excluding nonspecific interactions. A similar ladder pattern was observed for Sav mixed with Cut_2, demonstrating that the biotin moiety remains functional after SPAAC labeling. Furthermore, identical assembly patterns were obtained when SPAAC was performed after Sav–Cut_1 assembly, indicating that the azide group retains full reactivity even in the assembled state.

The author established an N-terminal dual functionalization method mediated by copper ion using

TA4C and maleimide under mild aqueous conditions. The method is broadly compatible with peptides and proteins in excellent selectivity and efficiency. TA4C derivatives bearing different substituents are well tolerated, allowing installation of useful functional groups. In addition, the dual modification is also remarkable stability. The method is able to be used for the formation of proteins assemblies. TA4C-based dual modification therefore offers a straightforward and adaptable approach for preparing multifunctional protein conjugates.

In chapter 4, The author envisioned a new bioconjugation reagent in which two TA4C moieties are bridged by an oligoethylene glycol linker, synthesized via a Dimroth rearrangement between a diamine reagent and a TA4C precursor containing a 2,3,5,6-tetrafluoro-4-cyanophenyl group. Here the author presents bisTA4C as a versatile reagent for N-to-N one-pot peptide–peptide conjugation and stepwise peptide–protein or protein–protein conjugation. The successful conjugations were verified by LC-MS and SDS-PAGE analyses.

BisTA4C was synthesized in a single step from a TA4C–CFP derivative and an oligoethylene glycol diamine via a Dimroth rearrangement. The reaction proceeded smoothly in THF at 40 °C for 12 h, affording the desired linker in good yield. The structure was confirmed by NMR and ESI–MS analysis. The reactivity of bisTA4C toward LGG exhibited a pronounced concentration dependence, as revealed by LC–MS analysis. Intermolecular peptide–peptide conjugation was favored at low linker loadings, whereas intermediate concentrations predominantly yielded mono-substituted LGG. At higher bisTA4C concentrations, dimer formation was again promoted. This non-linear, bimodal behavior reflects the interplay between linker availability and aldehyde hydrolysis, highlighting the tunability of bisTA4C in peptide–peptide assembly.

RNaseA was used as a model protein to examine N-terminal selectivity, affording a mono-modified product in 69% isolated yield. Subsequent coupling of bisTA4C-modified RNaseA with various peptides proceeded efficiently under mild aqueous conditions, with LC–MS confirming the formation of protein–peptide conjugates in all cases and minimal sequence dependence. Occasional formation of peptide–peptide dimers suggest partial hydrolysis of the bisTA4C–protein intermediate, highlighting the dual reactivity of bisTA4C and providing insight into its conjugation behavior.

BisTA4C was applied to heterologous protein–protein conjugation using GFP and RNaseA as model substrates. SDS–PAGE analysis confirmed the formation of a GFP–RNaseA heterodimer, appearing as a new band at ~42.7 kDa, with an estimated conjugation yield of approximately 7.5%. Although lower than peptide-based conjugation efficiencies, these results demonstrate that bisTA4C enables intermolecular crosslinking between large protein partners under mild aqueous conditions, with steric hindrance and limited N-terminal accessibility likely contributing to the reduced yield.

This chapter established bisTA4C as a modular linker for site-selective N-terminal bioconjugation. The linker enables the assembly of peptide–peptide, peptide–protein, and protein–protein conjugates under mild aqueous conditions using its two aldehyde groups. This method provides a simple and unified approach for constructing hybrid biomolecular structures. Some limitations are noted. Hydrolysis of N-terminal intermediates reduced efficiency when unmodified substrates were added,

and steric hindrance lowered yields with large protein partners. Even so, bisTA4C shows practical utility and good substrate compatibility. Its site-selectivity and operational simplicity highlight its potential for applications in chemical biology and protein engineering.

This thesis describes the development of a series of site-selective N-terminal dual modification methods for peptides and proteins based on triazole carbaldehyde (TA4C) chemistry. This work established a unified framework for introducing multiple functionalities at a single protein N-terminus with high selectivity and stability through the systematic design of metal-mediated and metal-free conjugation methods. The author established a practical and adaptable platform for N-terminal dual functionalization based on TA4C chemistry. This thesis also demonstrates the potential of TA4C chemistry in bioconjugation, materials construction, and chemical biology. The methods developed here provide new routes for creating proteins with specific functions and lay the foundation for future studies on site-selective reactions.