

Review

# Exploiting Protein N-Terminus for Site-Specific Bioconjugation

Lucia De Rosa <sup>1</sup>, Rossella Di Stasi <sup>1</sup>, Alessandra Romanelli <sup>2</sup> and Luca Domenico D'Andrea <sup>3,\*</sup>

<sup>1</sup> Istituto di Biostrutture e Bioimmagini, CNR, Via Mezzocannone 16, 80134 Napoli, Italy; lucia.derosa@cnr.it (L.D.R.); rossella.distasi@cnr.it (R.D.S.)

<sup>2</sup> Dipartimento di Scienze Farmaceutiche, Università Degli Studi di Milano, Via Venezian 21, 20133 Milano, Italy; alessandra.romanelli@unimi.it

<sup>3</sup> Istituto di Scienze e Tecnologie Chimiche "Giulio Natta", CNR Via M. Bianco 9, 20131 Milano, Italy

\* Correspondence: luca.dandrea@cnr.it

**Abstract:** Although a plethora of chemistries have been developed to selectively decorate protein molecules, novel strategies continue to be reported with the final aim of improving selectivity and mildness of the reaction conditions, preserve protein integrity, and fulfill all the increasing requirements of the modern applications of protein conjugates. The targeting of the protein N-terminal alpha-amine group appears a convenient solution to the issue, emerging as a useful and unique reactive site universally present in each protein molecule. Herein, we provide an updated overview of the methodologies developed until today to afford the selective modification of proteins through the targeting of the N-terminal alpha-amine. Chemical and enzymatic strategies enabling the selective labeling of the protein N-terminal alpha-amine group are described.

**Keywords:** protein labeling; chemo-selective reaction; molecular probe; chemical ligation; click-chemistry; aldehyde protein; azide protein



**Citation:** De Rosa, L.; Di Stasi, R.; Romanelli, A.; D'Andrea, L.D.

Exploiting Protein N-Terminus for Site-Specific Bioconjugation.

*Molecules* **2021**, *26*, 3521. <https://doi.org/10.3390/molecules26123521>

26123521

Academic Editor: Martin Witte

Received: 10 May 2021

Accepted: 7 June 2021

Published: 9 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Site-specific protein bioconjugation refers to the controlled modification of the protein chemical structure through the introduction of functional handles at well-defined positions along the protein sequence, endowing such biomolecules with novel and useful properties. A wide set of site-specific protein bioconjugation chemistries have been developed up to today, allowing for the effective preparation of modified proteins with valuable applications that pave the way to relevant scientific innovations in many disciplines of the life sciences such as in chemical biology, biophysics, biotechnology, material sciences, and biomedicine [1–5]. In these research fields, the use of site-specific protein decoration strategies instead of stochastic labeling methodologies is strictly required in order to ensure the highest level of uniformity over the properties of the protein conjugates, especially in the case of protein pharmaceuticals. In this latter context, several protein molecules site-specifically decorated with imaging probes, drugs, or carriers have been developed as novel protein-based diagnostic and therapeutic tools and many of them are successfully proceeding in clinical trials or are still to reach the pharmaceutical market [6–8]. Conventional strategies used to modify a protein molecule in a site-specific manner target amino acid side-chains [2,9], exploiting the peculiar reactivity of unique functional groups such as the thiol of cysteine [10], the epsilon-primary amine of lysine [11,12], the thioether of the methionine [13–15], the indole [16,17], imidazole [18], and phenol ring [19] respectively of tryptophan, histidine, and tyrosine. However, such methodology strictly requires the presence of a single reactive amino acid along the protein sequence to avoid multiple and heterogeneous patterns of decoration.

A more sophisticated and laborious methodology affording site-specific protein labeling relies on the use of the expanded genetic code to introduce by recombinant expression an unnatural amino acid into a protein target, which can react with orthogonal probes

through chemo-selective reactions [20,21]. Although representing a noteworthy technological advance, such an approach is technically demanding and low-yielding and, therefore, does not provide a resolutive solution to the issue. Chemical ligation approaches, in particular native chemical ligation (NCL) [22] and expressed protein ligation (EPL) [23], consisting of the assembly of a protein from synthetic or recombinant peptide segments, have revolutionized the protein bioconjugation concept and scope, allowing us to virtually introduce any kind of modification with surgical site-specificity [24,25]. However, chemical ligation reactions are usually performed in denaturing conditions and thus the labeled protein requires a refolding step that is not always successful and high-yielding. The search for novel bioconjugation chemistries point to overcoming such limitations, aiming at developing methodologies that ensure (1) a high level of selectivity, (2) could be performed in mild conditions compatible with the preservation of the native protein structure and function; (3) could be easily scaled-up; and (4) could be generally applicable to any protein target. Protein labeling through the selective targeting of the protein N-terminal alpha-amino group has recently emerged as a convenient solution to this issue, allowing to potentially satisfy all the above-mentioned requirements [26]. The N-terminal alpha-amino group is universally present in each protein molecule and is a unique functionality within a polypeptide chain endowed with a peculiar reactivity that allows it to be chemically distinguished from the other nucleophiles harbored by a protein molecule (i.e., the epsilon-amine group of lysines and the thiol group of cysteines). In fact, the N-terminal amine group is the only alpha-amine present within a protein molecule. It possesses an adjacent amide bond that influences its reactivity by lowering its pKa value (6.0–8.0) with respect to that of the epsilon-amine group of lysine (~10.5) and to the thiol group of the cysteine (~8.3) [27]. Consequently, the alpha-amine group is the most reactive nucleophile within a protein molecule at pH around neutrality and, therefore, can be selectively targeted with electrophilic probes at physiological pH. However, it should be noted that the pKa value of a functional group within a native protein molecule could in some cases vary from the expected one due to the influence of the chemical environment. Protein N-terminal alpha-amine is usually exposed on the protein surface, thus resulting in being easily accessible to chemical modification [28]. Accordingly, protein modification at this position usually has minimal or no impact on protein structure and function. For all these reasons, N-terminal protein labeling has gained great interest, leading in recent years to the development of a portfolio of chemical and enzymatic strategies for selectively tagging such protein sites. Herein, we provide a comprehensive overview of the bioconjugation chemical strategies developed to selectively target protein N-terminus using direct or indirect labeling procedures (Table 1). Direct approaches provide N-terminal protein modification in one step, taking advantage of molecular probes able to selectively react with such protein sites, often through conjugation reaction mechanisms that involve the side chain of the first amino acid of the protein sequence. Indirect approaches are, instead, two-steps procedures in which the N-terminus is first converted into a reactive handle and then selectively targeted with an orthogonal probe through a second reaction step. Although one-step approaches for direct protein labeling may appear more convenient and rapid, indirect strategies are more versatile with respect to the kind of probe to be introduced as they may take advantage of the wide collection of commercially available molecular probes without requiring severe synthetic efforts. Enzymatic approaches affording N-terminal protein labeling are also summarized (Table 2).

**Table 1.** Chemical strategies developed for site-specific protein bioconjugation through the targeting of the N-terminal position.

Chemical Reagent	Target Function	Reaction	Junction	Experimental Conditions <sup>a</sup>	Yield of Conversion (%) <sup>a</sup>	Ref
benzaldehydes	$\alpha$ -amine	reductive amination	secondary amine	citric acid buffer pH 6.1, 2 eq of benzaldehyde, 5 eq <sup>b</sup> of NaBH <sub>3</sub> CN, 6–48 h, r.t.	30–70	[29]
2-ethynylbenzaldehydes (2-EBA)	$\alpha$ -amine	imine formation and intramolecular 6-endo-dig cyclization	isoquinolinium	PBS/DMSO 9:1 pH 6.5–7.4, 5–10 eq of 2-EBA, 1–16 h, 37 °C	10–92	[30]
N-hydroxy-phthalimide (NHP)	$\alpha$ -amine	phthalimidation	phthalimidoamine	phosphate buffer pH 7.0, 1 eq of NHP, 8 h, r.t.	45–50	[31]
selenobenzaldehyde (SBA) ester	$\alpha$ -amine	acylation through ACL	amide	dimethylformamide/PBS pH 7.0, 5 eq of SBA ester, 40 h, r.t.	70	[32]
thioesters	1,2-aminothiol (N-term Cys)	acylation through NCL	amide	phosphate buffer pH 7.4, 20 eq mercaptoethanesulfonate 4 eq of thioester probe, 5 h, r.t.	~100	[33]
2-cyanobenzothiazole (CBT)	1,2-aminothiol (N-term Cys)	condensation	thiazoline	PBS pH 7.4, 2 mM glutathione or TCEP, 5 eq of CBT, 2 h, r.t.	Not reported	[34]
aldehydes (e.g., glyoxalaldehyde)	1,2-aminothiol (N-term Cys) or 1,2-aminoalcohol (N-term Ser/Thr)	condensation	thiazolidine or oxazolidine	acetate buffer pH 4.5, 1 mM DTT or TCEP 16 h, then 40–200 eq aldehyde, 2.5–4 days, 4 °C	90	[35]
2-formyl phenylboronic acid (2-FPBA)	1,2-aminothiol (N-term Cys)	condensation	thiazolidino-boronate	phosphate buffer pH 7.0, 1 eq of 2-FPBA 30 min, r.t.	~100	[36]
aldehydes	N-term Trp	Pictet–Spengler condensation	tetrahydro-beta- carboline	Glacial acetic acid, 24 h reagents ratio not reported	~100	[37]
2-pyridinecarbaldehyde (2-PCA)	$\alpha$ -amine	imine condensation	imidazolidinone	phosphate buffer pH 7.5, 25 $\mu$ M protein substrate, 10 mM 2-PCA, 37 °C, 16 h	43–95	[38]

Table 1. Cont.

Chemical Reagent	Target Function	Reaction	Junction	Experimental Conditions <sup>a</sup>	Yield of Conversion (%) <sup>a</sup>	Ref
ortho-aminophenols	$\alpha$ -amine (especially N-term Pro)	oxidative coupling	secondary or tertiary amine <sup>c</sup>	phosphate buffer pH 7.5, 20 $\mu$ M protein substrate, 5 eq of ortho-aminophenol, 5 mM $K_3Fe(CN)_6$ , 30 min, r.t.	up to ~ 100 (on N-term prolyl-protein)	[39]
2-(2-formylphenoxy)acetic acid (2-FPOAA)	$\alpha$ -amine of N-term Gly	imine condensation + nucleophilic addition	amino-alcohol	bicarbonate buffer pH 7.8, 20 $\mu$ M protein substrate 10 mM 2-FPOAA, 24–48 h, 20 °C	40–71	[40]
4-methoxyphenyl esters (4-MOPE)	$\alpha$ -amine of Gly-His <sub>n</sub> N-term tag	acylation	amide	HEPES buffer pH 7.5, 20 eq of 4-MOPE, 24 h, 4 °C	45 (on Gly-His <sub>6</sub> tagged protein)	[41]
pyridoxal-5'-phosphate (PLP)	$\alpha$ -amine	transamination	aldehyde/ketone	phosphate buffer pH 6.5, 10 mM PLP, 18–20 h, 37–55 °C	30–80	[42]
Rapoport's salt (RS)	$\alpha$ -amine (especially N-term Glu)	transamination	aldehyde/ketone	phosphate buffer pH 6.5, 100 mM RS 1 h, 37 °C	67	[43]
sodium periodate (NaIO <sub>4</sub> )	1,2 amino-alcohol (N-term Ser or Thr)	oxidation	aldehyde	phosphate buffer/NaCl pH 7.0, 1.5 eq. of NaIO <sub>4</sub> , 1 h, r.t.	~100	[44]
imidazole-1-sulfonyl azide (I-1-SA)	$\alpha$ -amine	diazotransfer	azide	diethanolamine buffered solution pH 8.5, 17.5 eq of I-1-SA, overnight, r.t.	~100	[45]
phenyl ketene (PK)	$\alpha$ -amine	acylation	amide	phosphate buffer pH 6.3 or 9.2, 6–10 eq of PK, 15 min—overnight, 37 °C	23–38	[46]

<sup>a</sup> Yields of conversion and reaction conditions are reported according to the cited references. Of note, these parameters strongly rely on the protein targets and on the type of probe used as labeling reaction substrates and therefore are not universal values but should vary according to protein substrate and probe properties. <sup>b</sup> Reagent equivalents are calculated with respect to the protein substrate. <sup>c</sup> A tertiary amine is obtained if the reaction is performed on the N-terminal prolyl-protein. PBS, phosphate buffer saline; r.t., room temperature; GuHCl, guanidinium chloride; TCEP, tris-carboxyethylphosphine; DTT, dithiothreitol; ACL, aldehyde capture ligation; NCL, native chemical ligation;  $K_3Fe(CN)_6$ , potassium ferricyanide.

**Table 2.** Enzyme-catalyzed strategies for site-specific N-terminal protein labeling.

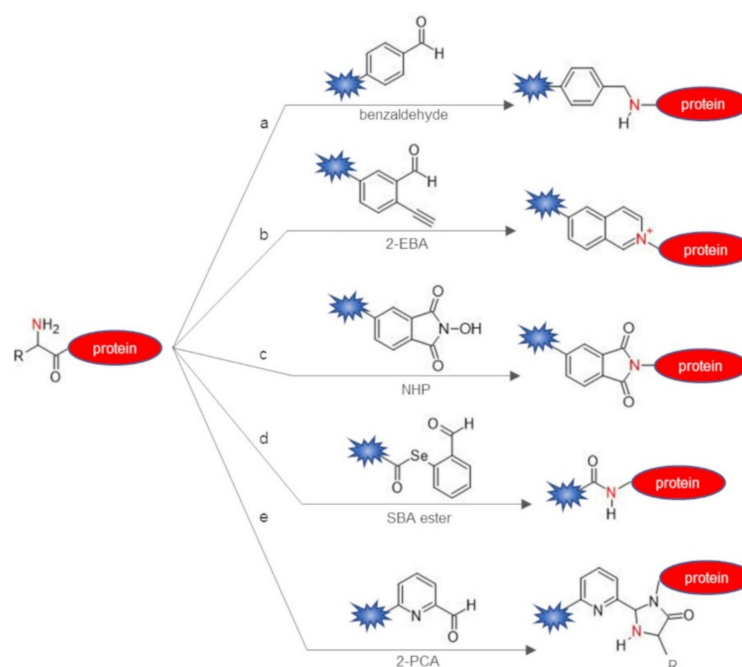
Enzyme	Reagent	N-terminal Target Sequence or Function	Reaction Conditions <sup>a</sup>	Yield <sup>a</sup>	Ref
Sortase A (SrtA)	Leu-Pro-X-Thr-Gly/Ala-X peptide (X any a.a.)	$\alpha$ -amine of oligo-Gly	0.5–1 mM probe-Leu-Pro-X-Thr-Gly/Ala-, 10–50 $\mu$ M target protein, 20–150 $\mu$ M SrtA, in 50 mM Tris buffer pH 7.5, 150 mM NaCl, 10 mM CaCl <sub>2</sub> 15 min—5 h, r.t. or 37 °C	up to 90	[47]
Subtiligase	C-terminal ester peptide	$\alpha$ -amine	5 mM peptide ester-probe, 10–50 $\mu$ M target protein, 1 $\mu$ M subtiligase, in 100 mM tricine pH 8.0, 1–2 h, 4 °C or r.t.	variable <sup>b</sup>	[48]
Butelase 1	Thiodepsipeptide Asn/Asp(S) Gly-Val	$\alpha$ -amine of X <sub>1</sub> -X <sub>2</sub> (X <sub>1</sub> any a.a. except Pro; X <sub>2</sub> Cys, Ile, Leu, Val)	100 $\mu$ M target protein, 0.1 $\mu$ M butelase 1, 400–500 $\mu$ M thiodepsipeptide (1 eq added every 30 min) 1 mM EDTA, 20 mM phosphate buffer pH 6.5, 100 min—2.5 h, temperature not specified	up to 95	[49]
N-Myristoyltransferase (NMT)	myristoyl-CoA derivatives	$\alpha$ -amine of Gly-X <sub>1</sub> -X <sub>2</sub> -X <sub>3</sub> -Ser/Thr (Lys) (X is any amino acid)	15 $\mu$ M target protein, 2 mM DTT 30 $\mu$ M myristoyl-CoA derivative in 30 mM Tris buffer, 0.5 mM EGTA, 0.1% TRITON <sup>®</sup> X-100, pH 7.4 18 h—37 °C	not reported	[50]

<sup>a</sup> Yields of conversion and reaction conditions are reported according to the cited references. These parameters are not universal values but could vary according to the substrates' properties. <sup>b</sup> Dependent on the N-terminal target protein sequence and structure. EDTA, ethylenediaminetetraacetic acid; a.a., amino acid; r.t., room temperature; DTT, dithiothreitol; EGTA, ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid.

## 2. Chemical Strategies Targeting Protein N-Terminus

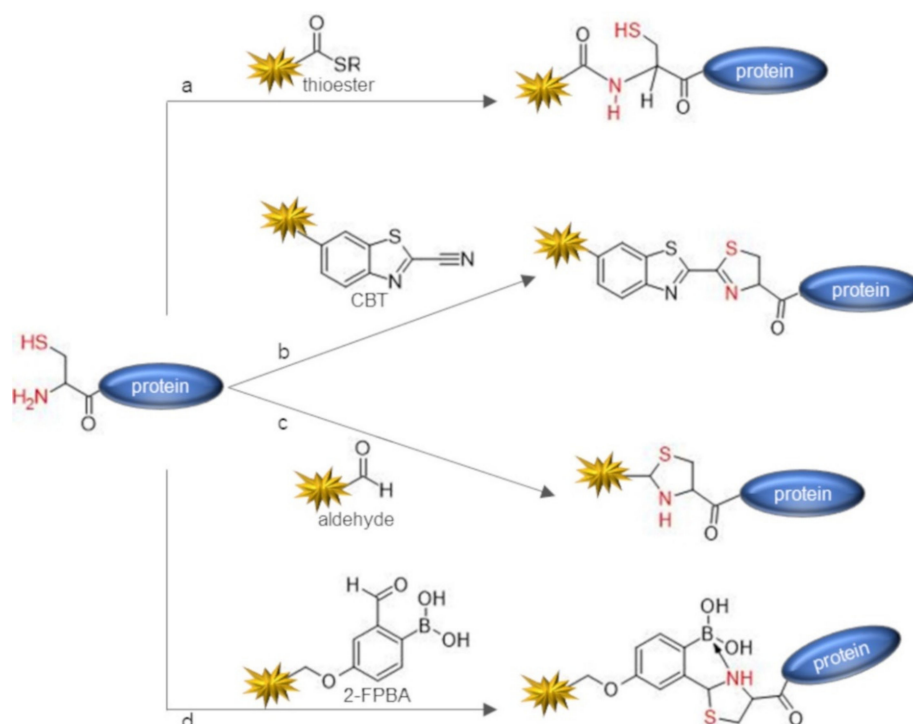
### 2.1. Direct Labeling of Protein N-Terminus

The simplest approach used to directly target protein N-terminus refers to the use of amine-reactive compounds such as activated ester, for example, N-hydroxysuccinimide (NHS) esters or perfluorophenyl esters, exploiting the different pKa value of the alpha-amine to discriminate between the N-terminal amine group and the other nucleophiles by performing the labeling reaction at low-to-neutral pH [51]. NHS ester probes react in mild conditions with the N-terminal amine groups through an acylation reaction, leading to the formation of an amide conjugate, as shown using unprotected peptides [52]. Aldehydes may also be used to selectively target N-terminal alpha-amine in a pH-controlled manner using different chemistries. Benzaldehyde derivatives have been used to selectively alkylate the protein N-terminus by reductive amination in the presence of sodium cyanoborohydride ( $\text{NaBH}_3\text{CN}$ ), leading to the formation of a secondary amine as the conjugate product [29] (Table 1, Figure 1a). Conveniently, protein modification by reductive amination allows the positive charge on the N-terminus to be preserved, which could be crucial to retain, in some cases, protein bioactivity. Similarly, 2-ethynylbenzaldehyde (2-EBA) derivatives were reported to selectively react with the protein N-terminal alpha-amine group in phosphate-buffered saline solution at pH 6.5–7.4 through the formation of an isoquinolinium conjugate in which the nitrogen atom is positively charged (Figure 1b) [30]. A phthalimidation protocol using N-hydroxy-phthalimide reagents and targeting protein N-terminus has also been reported, showing selectivity against alpha-amine in the presence of lysine residues at pH 7.0 (Figure 1c) [31]. The reaction afforded a phthalimidoamine product that can be conveniently converted back to the amine in the presence of hydrazine. Selenobenzaldehyde esters are also able to selectively react with the N-terminal alpha-amine in the presence of multiple Lys residues in a pH-controlled manner through aldehyde capture ligation (ACL), which consists of an acylation reaction that yields an amide bonded conjugate [32] (Figure 1d).



**Figure 1.** Selective protein labeling through the direct targeting of the N-terminal alpha-amine group. (a) Reductive amination with a benzaldehyde probe; (b) imine formation and intramolecular 6-endo-dig cyclization with a 2-ethynylbenzaldehyde (2-EBA) probe; (c) phthalimidation with a N-hydroxy-phthalimide (NHP) probe; (d) aldehyde capture ligation with a selenobenzaldehyde (SBA) ester probe; (e) imine condensation with a 2-pyridinecarbaldehyde (2-PCA) probe.

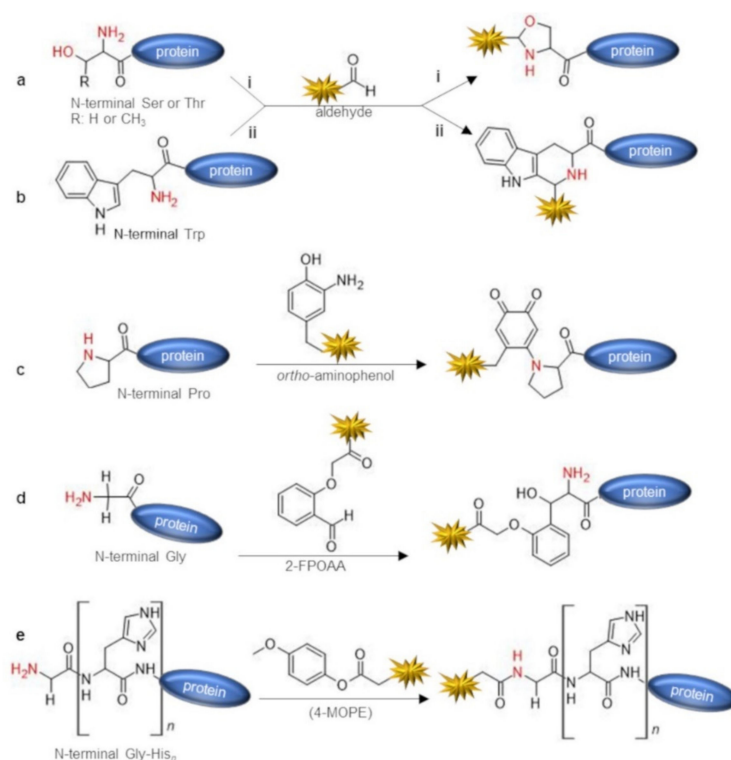
However, labeling approaches based on the control of the pH to target protein N-terminus do not ensure complete selectivity and off-target labeling may frequently occur, leading to heterogeneous protein mixtures. Most sophisticated methodologies developed to directly target protein N-terminus exploit the selective reactivity of small organic molecules against the N-terminal alpha-amine by involving the participation of the side-chain of the first amino acid in the conjugation reaction mechanism. Several approaches have been developed to target the 1,2-aminothiol function exposed by N-terminal cysteinyl proteins (Figure 2).



**Figure 2.** Selective targeting of N-terminal cysteinyl proteins. (a) Native chemical ligation with a thioester probe; (b) condensation with a 2-cyanobenzothiazole (CBT) probe; (c) thiiazolidine ligation with an aldehyde probe; (d) condensation with a 2-formyl phenylboronic acid (2-FPBA) probe.

Cysteines are rarely present at the N-terminal position in natural proteins and therefore should be introduced by protein engineering [53]. However, the additional Cys residue may hamper the correct pairing of native Cys residues into disulfide bridges, and hence may lead to protein misfolding. Besides, during recombinant expression and protein purification, the N-terminal cysteine may undergo the cyclization reaction by condensation with aldehyde catabolites present within the cell, leading to the formation of a thiazolidine derivative that masks the terminal 1,2-aminothiol function [54]. Although in this case the N-terminal cysteine could be restored by treatment with O-methylhydroxylamine [54,55], the most convenient strategy to prepare N-terminal cysteinyl proteins refers to the exposition of the N-terminal cysteine after protein expression and purification through proteolytic cleavage with Factor Xa, tobacco etch virus (TEV), and thrombin proteases [56–58] or using self-cleavable fusion partners such as intein [59]. The most powerful and widespread chemical approach exploited to target N-terminal cysteinyl proteins refers to the reaction with a thioester probe through NCL [33,54,60–65] (Figure 2a). The NCL reaction proceeds through a reversible trans-thioesterification mediated by the thiol group of the cysteine, followed by a spontaneous intramolecular S→N acyl shift that translocates the probe on the alpha-amine group, affording an amide bonded product [22]. Notably, the thiol group of the N-terminal cysteine mediates NCL reaction but is restored in the free form in the final conjugated product. Therefore, such a thiol group could be exploited for

pursuing second site-specific labeling using a thiol reactive probe, for instance, through the thiol–maleimide chemistry [64,66]. Conveniently, thioester probes can be obtained from commercially available NHS-ester derivatives by treatment with thiols [65,67]. N-terminal cysteine can also be directly tagged using 2-cyanobenzothiazole (CBT) through a water-compatible condensation reaction [34] that mimics the last reaction step in the synthesis of D-luciferin in firefly [68] (Figure 2b). The reaction is fast, requires slight excess of the CBT-probe, leads to a stable conjugate, and efficiently proceeds in aqueous solution at physiological conditions (PBS pH 7.4), also being compatible with in cells and in vivo imaging applications [34,69,70]. The ability of N-terminal cysteine to react with aldehydes through a condensation reaction forming a thiazolidine derivative has also been explored as a N-terminal labeling strategy [71,72] (Figure 2c). N-terminal serine and threonine may similarly react with aldehyde probes, providing an oxazolidine conjugate [73] (Figure 3a).



**Figure 3.** Site-specific protein labeling through the targeting of N-terminal Ser (or Thr) (a), Trp (b), Pro (c), Gly (d), or Gly-His<sub>n</sub> (e).

Although thiazolidine ligation has been successfully applied to the synthesis of protein conjugates [35,74], thiazolidine ligation suffers of severe limitations as the reaction requires a slightly acidic pH (pH 4–5), is not always well tolerated by the protein target, and shows a slow kinetic requiring several days of incubation to proceed, even in the presence of large excess of the aldehyde reactant. Besides, the thiazolidine product is not stable and may hydrolyze [26]. The use of *ortho*-boronic acid substituted benzaldehydes such as reagents such as 2-formyl phenylboronic acid (2-FPBA) allows us to overcome the drawbacks of thiazolidine ligation. 2-FPBA promotes the fast formation of a thiazolidino-boronate (TzB) conjugate by reacting with the N-terminal cysteine at neutral pH and in equimolar ratio between the protein and the probe derivative [36,75] (Figure 2d). Notably, the TzB product was found to exhibit superior stability at physiological pH with respect to thiazolidine conjugates due to boron coordination by the thiazolidine ring. Conveniently, TzB may dissociate in the presence of benzyl hydroxylamine [75] and at acidic pH [36]. This latter feature makes the TzB chemistry interesting for the development of pH responsive drug-conjugates such as antibody–drug conjugates designed to selectively release the cytotoxic payload in the endosomes. Aldehyde probes may also be exploited to target



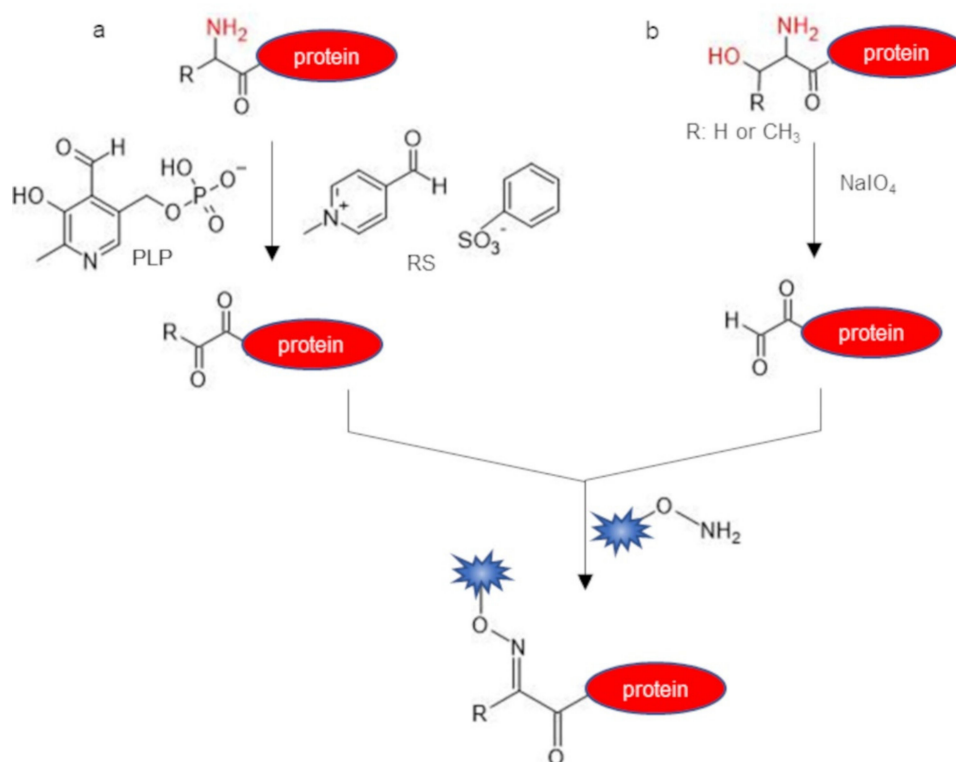
an N-terminal tryptophan-containing protein through the Pictet–Spengler reaction that generates a carbon–carbon bonded conjugate [37,76] (Figure 3b). Such a reaction was performed to ligate peptide fragments in acid conditions and is not compatible with the preservation of the native protein structure. However, the Pictet–Spengler reaction was also performed in milder conditions, conjugating an indole probe on the N-terminal aldehyde protein [77,78]. Francis et al. reported a versatile and straightforward method that used 2-pyridinecarbaldehyde (2-PCA) based probes to selectively label protein N-terminal alpha-amine without requiring the assistance of the first amino acid side-chain [38] (Figure 1e). 2-PCA reacts with the protein N-terminal alpha-amine through an imine condensation reaction, after which the amide nitrogen of the neighboring amino acid cyclizes on the imine to form an imidazolidinone conjugate. As they lack the nearby amide group, the amine group of lysine side chains are not able to form stable products, thus ensuring high selectivity of the chemistry against the N-terminus. Apart from the N-terminal prolyl protein, which cannot react with 2-PCA, the method is general and proceeds smoothly in a variety of common biological buffers at pH 7.5. The imidazolidinone conjugate showed high stability over a pH range from 3 to 11 but was prone to hydrolysis at 37 °C. The repertoire of bioconjugation chemistries for the direct labeling of protein N-terminus also includes an oxidative coupling reaction performed at pH 7.5 with ortho-aminophenols catalyzed by potassium ferricyanide ( $K_3Fe(CN)_6$ ), affording a stable bioconjugate [39] (Figure 3c). The reaction requires a slight excess of the ortho-aminophenol (2–5 eq). Although the majority of N-terminal amino acids showed good-to-high levels of conversion (60–90%), the reaction worked more efficiently on N-terminal prolyl peptides and proteins, reaching in this case nearly complete modification (90–100%) and reacting in a shorter time (less than 30 min). Reaction of an ortho-aminophenol with an N-terminal prolyl protein gives an ortho-quinone product that resulted stable to reductant, nucleophiles, and acid or basic pH. Disadvantageously, cysteine side-chains also react with o-aminophenols, limiting the applicability of the strategy to protein without cysteines or to their temporary protection with 5, 5'-dithiobis (2-nitrobenzoic acid) (DTNB; Ellman's reagent). Very recently, an effective chemical platform for the targeting of N-terminal glycyl proteins has also been reported [40] (Figure 3d). The great advantage of this methodology is that natural proteins often possess a N-terminal glycine and such residue can be easily exposed at the protein N-terminus during recombinant expression, resulting in a versatile and general labeling methodology [79,80]. In fact, the N-terminal formyl-Met (fMet) residue is promptly removed during recombinant expression of glycyl-proteins by the intracellular Met-aminopeptidase, which efficiently acts on the nascent polypeptide chain to expose the N-terminal Gly residue. Besides, the protease cleavage sites introduced at the N-terminus of a recombinant protein, which allow for the removal of the affinity tag by site-directed proteolysis after purification such as TEV or Factor Xa cleavage sites, are usually designed to leave a Gly residue as the +1 amino acid of the protein target. The N-terminal glycine directed bioconjugation chemistry exploits ortho-substituted benzaldehydes such as 2-(2-formylphenoxy) acetic acid (2-FPOAA) with an appropriately designed carbonyl substituent as a hydrogen bond acceptor that results in the exclusive labeling of the N-terminal glycine residue, affording a stable amino-alcohol as a bioconjugation product through the intermediate formation of an imine. The reaction proceeds under mild conditions, being performed in a carbonate buffer at pH 7.8 at room temperature and was demonstrated to exclusively target the N-terminal glycine even on protein targets exposing lysine or cysteine residues. Using a symmetric bis-aldehyde derivative of the reagent, various chemical tags can be selectively introduced on the protein N-terminus exploiting the second aldehyde function for protein functionalization. Notably, such additional aldehyde function was also exploited for protein purification through hydrazone ligation using a hydrazide derivatized resin. The immobilized protein could be effectively released from the solid support by treatment with O-hydroxylamine derivatives through trans-oximization. Inspired by an undesired chemical modification occurring at the N-terminus of His-tagged recombinant protein [81], D-gluconic acid  $\delta$ -lactone (GDL) was

demonstrated to effectively acylate the alpha-amine group of protein targets harboring the on Gly-His<sub>n</sub> N-terminal amino acid stretch [41]. A higher yield of N-terminal acylation was obtained for tag sequence containing at least three His residues. The reaction selectively proceeded in HEPES buffer at pH 7.5 with high yield but appeared to be reversible. Other esters resulted in being able to selectively and irreversibly acylate the alpha-amine of the on Gly-His<sub>n</sub> tag, in particular, the 4-methoxyphenyl ester was selected as the optimal reagent and conveniently modified with functional handles such as a biotin or an azide group (Figure 3e). The reason for the selective acylation of the N-terminal alpha-amine on Gly-His<sub>n</sub>-tagged protein resides in the base catalysis of the reaction mechanism in which the His side-chain assists the deprotonation during the direct acylation of the Gly alpha-amine. Conveniently, Gly-His<sub>6</sub> tag, harboring six His residues, shows a dual application, as the His<sub>6</sub>-tag affords recombinant protein purification by affinity chromatography on Ni<sup>2+</sup>-nitriloacetic resins and can be exploited for N-terminal protein functionalization with GDL or 4-methoxyphenyl ester probes.

## 2.2. Indirect Labeling of Protein N-Terminus

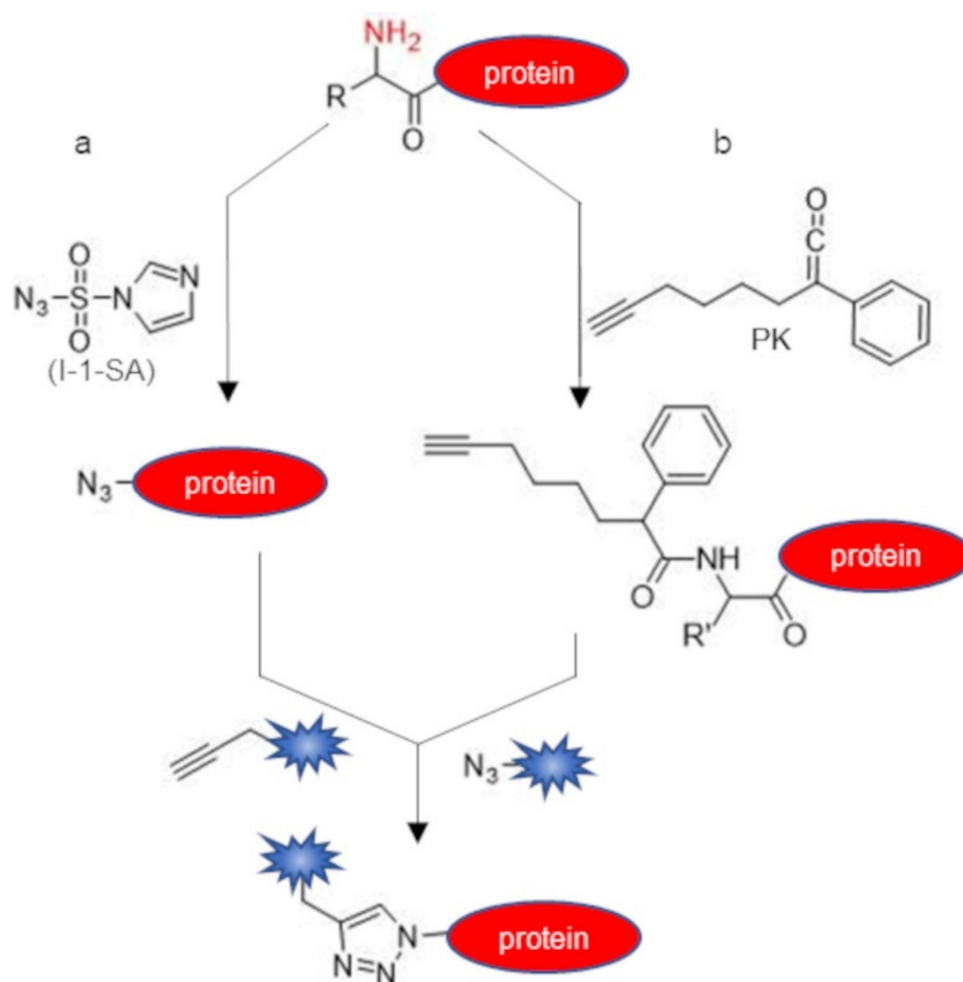
A collection of labeling methodologies developed to selectively install a molecular probe on the N-terminal position of a protein target exploits indirect approaches that require two consecutive reaction steps (Figure 4). The first reaction step affords the selective modification of the N-terminal alpha-amine group into a not naturally occurring functionality (i.e., an aldehyde, a ketone, an azide, or an alkyne group). The second step is the actual conjugation reaction that provides the introduction of the desired probe through a chemo-selective reaction between the functionality installed on the protein N-terminus and a molecular probe bearing an orthogonal group. The aldehyde function is the reactive handle most frequently used to afford the indirect labeling of the protein N-terminus because such function is versatile and can be targeted with a wide range of orthogonal probes through several chemistries. A convenient strategy exploited to selectively oxidize the N-terminal alpha-amine group into an aldehyde or a ketone function is the pyridoxal 5'-phosphate (PLP)-mediated transamination reaction [42] (Figure 4a). The PLP is an aldehyde able to condensate with all the amine groups of a protein molecule (i.e., the N-terminal alpha-amine and the epsilon-amine of lysines), leading to the formation of imine derivatives. However, only the imine formed on the N-terminus can tautomerize due to the lower value of the pKa of the alpha-amine with respect to that of the lysine amine group. The resulting glyoxyl imine hydrolyzes, affording the formation of an aldehyde or a ketone group specifically at the N-terminus. These carbonyl handles can be further modified with aldehyde reactive probes such as alkoxyamine or hydrazide probes that react with aldehydes through oxime and hydrazone ligation, respectively. However, the PLP transamination reaction yield is not always high and shows a strong sequence dependence. Besides, the reaction is incompatible with some N-terminal amino acids. For instance, His, Trp, Lys, and Pro generate adducts with PLP while Cys and Ser undergo a beta-elimination side-reaction [82]. The reaction was used to prepare antibody conjugates functionalized at the N-terminal position with biotin, PEG, or Alexa dye alkoxyamine derivatives [83]. Protein labeling with PLP was carried out in an aqueous buffer at pH 6.5 at a temperature of 37–50 °C for 18–20 h in the presence of 10 mM PLP using the target antibody at low concentration. However, the yields of labeling were low and elevated temperatures were required to promote the reaction, limiting the general applicability of the approach. An alternative transamination reagent, the N-methylpyridinium-4-carboxaldehyde benzenesulfonate salt (Rapoport's Salt) [84], resulted in being particularly effective for the labeling of glutamate-terminal proteins and thereby results in being useful for the site-selective modification of wild-type human IgG1 naturally possessing an N-terminal glutamic acid residue [43] (Figure 4a). N-terminal glutamate residues may also be exposed on recombinant proteins by proteolytic cleavage with Factor Xa [85]. A straightforward and rapid method for the installation of an N-terminal aldehyde function on proteins harboring a 1,2-aminoalcohol function (i.e., a protein harboring an N-terminal Ser or Thr residue) refers

to an oxidation reaction performed by treatment with sodium periodate ( $\text{NaIO}_4$ ) [72,86] (Figure 4b). The reaction is fast and requires mild conditions as it rapidly proceeds in an aqueous solution at neutral pH in the presence of a slight excess of  $\text{NaIO}_4$  and using very diluted reactant concentrations. However, a prolonged treatment with  $\text{NaIO}_4$  may lead to the undesired oxidation of other amino acids such as Met and Cys [86,87], hence the aldehyde protein should be quickly purified after the treatment with periodate in order to avoid side-reactions. The aldehyde function can be selectively targeted using a chemo-selective reaction such as oxime ligation. We recently exploited such methodology to site-specifically modify the domain 2 of the vascular endothelial growth factor receptor 1 (VEGFR1D2) by introducing a functional probe at the protein N-terminus [44,88]. The serine at the N-terminus was selectively converted into an N-terminal glyoxalamide group by periodate mediated oxidation and subsequently reacted by oxime ligation with an oxyamine-biotin. N-terminal Ser oxidation with periodate was also combined to a strain-promoted alkyne-nitrone cycloaddition (SPANC) bioorthogonal reaction for the N-terminal site-specific modification of peptides and proteins, leading to N-alkylated isoxazolines as conjugation products [89]. In this three-step protocol, the N-terminal Ser protein substrate was subjected to oxidation with  $\text{NaIO}_4$  (1.1 equiv, 1 h) and, after treatment with p-methoxybenzenethiol, N-methylhydroxylamine and p-anisidine, was finally reacted with a cyclooctynol probe, yielding, after 24 h of incubation, the complete conversion of the protein target into the isoxazoline derivative via nitrone formation. The protocol was successfully used to prepare N-terminally pegylated Interleukin-8 using a PEG-cyclooctynol as the reagent. Indole probes able to react by the Pictet–Spengler reaction in mild conditions with the aldehyde protein have also been successfully used for N-terminal protein labeling [77,78].



**Figure 4.** Two-step N-terminal protein labeling. (a) Transamination reaction with pyridoxal-5'-phosphate (PLP) or with N-methylpyridinium-4-carboxaldehyde benzenesulfonate salt (Rapoport's salt, RS), leading to a N-terminal ketone protein; (b) oxidation with sodium periodate ( $\text{NaIO}_4$ ) of 1,2 amino alcohol (N-terminal Ser or Thr), affording a N-terminal aldehyde protein. The ketone/aldehyde function installed at the protein N-terminus can be further reacted with an aminoxy-probe via oxime ligation.

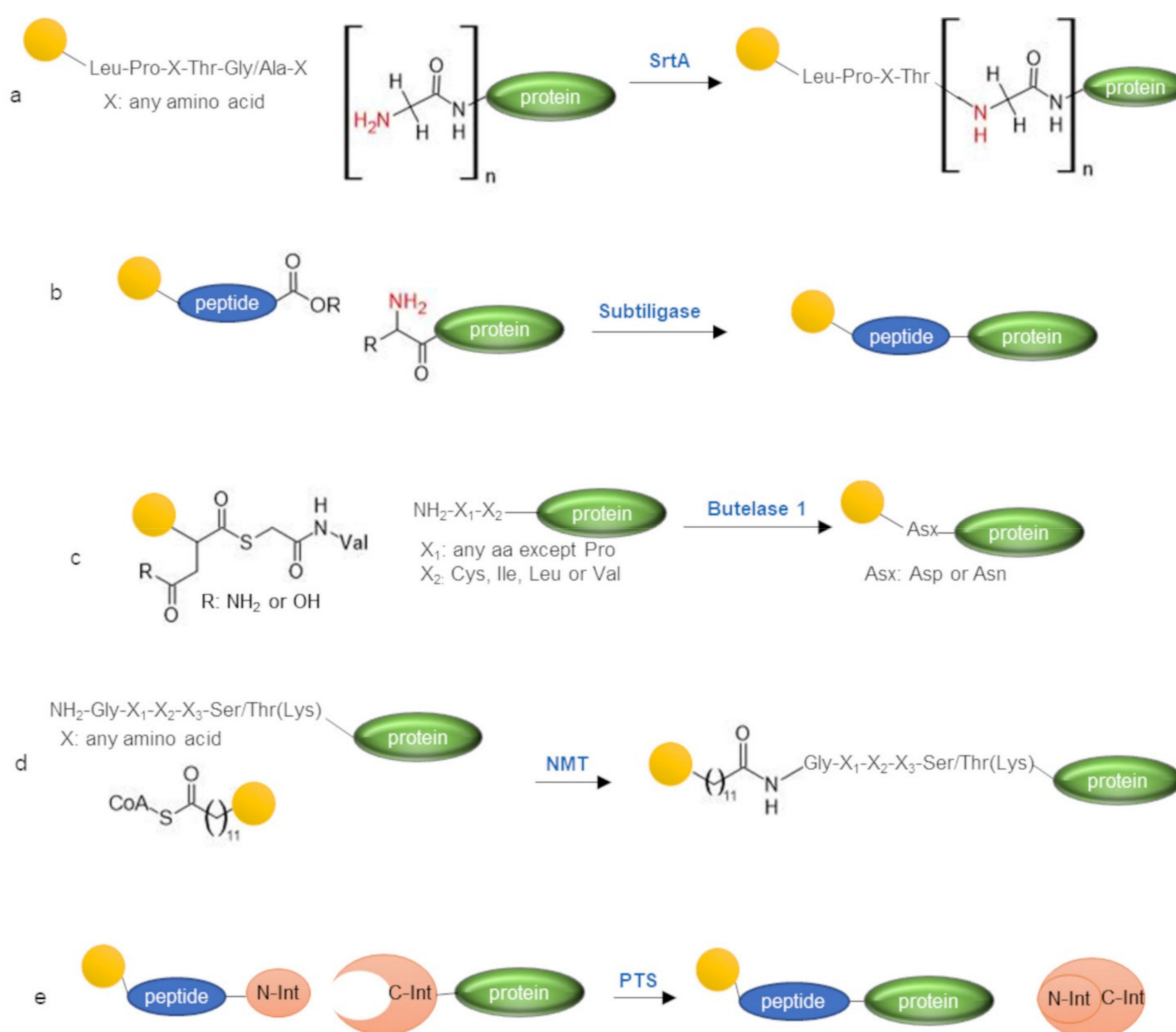
Besides aldehydes and ketones, the conversion of the N-terminal alpha-amine into an azide group is another straightforward way to introduce a targetable handle into proteins, as azides can be conveniently reacted using an alkyne-probe and click-chemistry reaction [90,91] (Figure 5a). N-terminal alpha-amine can be efficiently converted into azides through aqueous diazotransfer using the reagent imidazole-1-sulfonyl azide [45]. Performing the reaction at pH 8.5 allows for the selective azidation of the N-terminus over lysine residues. An alkyne function can also be installed at the protein N-terminus through the convenient use of a phenyl ketene reactive probe harboring an alkyne handle [46] (Figure 5b). The ketene probe developed was able to react with the alpha-amine group, leading to an amide conjugate and was demonstrated to show high selectivity against the N-terminal alpha amine of the majority of the natural amino acid residues. The strategy appeared compatible with the presence of other reactive groups such as the  $-OH$  group of Ser and Thr and  $-NH_2$  of Lys. However, the thiol group of Cys reacts with the ketene probe, leading to a thioester derivative that can be hydrolyzed by treatment with hydroxylamine. As proof of concept, the alkyne modified ketene probe was used to selectively label insulin, lysozyme, RNase A, and BCArg at the N-terminal position in aqueous solution. The copper-catalyzed [3 + 2] cycloaddition (CuAAC) reaction with a dansyl azide probe afforded protein labeling.



**Figure 5.** Two-step N-terminal protein labeling. (a) N-terminal alpha-amine group can be converted into an azide group by diazotransfer with imidazole-1-sulfonyl azide (I-1-SA); (b) acylation with phenyl ketene (PK) allows for the introduction of a N-terminal alkyne group. After the selective introduction of an alkyne or an azide function at the protein N-terminus, the protein can be further modified through copper-catalyzed [3 + 2] cycloaddition (CuAAC) click chemistry.

### 3. Enzymatic Labeling of Protein N-Terminus

Enzymatic approaches for N-terminal protein labeling represent a convenient alternative to the use of chemical reagents (Table 2, Figure 6). Enzymes ensure a uniquely high level of selectivity, perform in very mild reaction conditions, and need short reaction times. However, the use of enzymes in protein labeling may be expensive and consequently limited to a small scale. A number of enzymes, mainly transferases, have been exploited to selectively target protein N-terminus. They are used to catalyze the appending of a peptide, comprising the enzyme consensus sequence and site-specifically modified with the label of choice (such as fluorophores, biotin, lipids, nucleic acids, carbohydrates, and so on) at the N-terminus of the protein target.



**Figure 6.** Enzyme-catalyzed N-terminal protein labeling. N-terminal protein modification mediated by (a) sortase A (SrtA); (b) subtiligase; (c) N-Myristoyltransferase (NMT); (d) butelase 1; (e) protein trans-splicing (PTS). Aa, amino acid; CoA, coenzyme-A; N-Int, N-terminal domain of a split intein; C-Int, C-terminal domain of a split intein.

#### 3.1. Sortase A

Sortase A (SrtA) is one of the most widely explored enzymes for protein labeling [92] (Figure 6a). SrtA is a bacterial transpeptidase able to catalyze the ligation of a synthetic peptide reproducing the SrtA recognition sequence (Leu-Pro-X-Thr-Gly/Ala, where X is any amino acid, the fifth amino acid is Gly or Ala depending on the specific SrtA used) to an N-terminal oligo-Gly stretch introduced upstream of the protein target sequence [47,93,94]. SrtA exhibits a catalytic Cys able to cleave the peptide bond between the Thr and the

Gly residues within the pentapeptide consensus sequence and form a thioacyl intermediate complex, which ultimately reacts with the N-terminal alpha-amine of an oligo-Gly protein, resulting in the formation of a new peptide bond between the protein and the peptide tag. SrtA can also be exploited to label protein C-terminus [92,95,96], internal loop regions [97,98], and for protein head-to-tail cyclization [99].

### 3.2. Subtiligase

Subtiligase is another powerful tool for the selective enzymatic modification of protein N-terminus [48,100] (Figure 6b). Subtiligase is a peptide ligase derived from the serine protease subtilisin through rational mutagenesis. By introducing two amino acid substitutions (Ser221Cys and Pro225Ala) into the subtilisin sequence, peptide ligation activity of the enzyme was effectively enhanced with respect to the parent protease activity, successfully yielding a powerful peptide ligation catalyst [101]. Subtiligase catalyzes the ligation of a C-terminal ester peptide to the N-terminal alpha-amine group of a protein target, with absolute selectivity over other amino acid nucleophiles. Using specifically modified C-terminal ester peptides, functional handles can be selectively introduced at the protein N-terminus. The C-terminal thioester peptide can also be exploited as subtiligase substrates, performing even better than ester peptides [102]. Reaction efficiency strongly relies on the N-terminal sequence and structure of the target protein. A set of subtiligase mutants with different sequence specificities is available and the most convenient subtiligase mutant can be selected according to the first two N-terminal amino acids of the protein substrate using the ALPINE ( $\alpha$ -Amine Ligation Profiling Informing N-terminal modification Enzyme selection, <https://wellslab.ucsf.edu/alpine>) web application. Alternatively, a mixture of subtiligase mutants, showing a broad sequence specificity, can be conveniently employed [103]. Subtiligase can modify loop and  $\beta$ -sheet protein regions more efficiently than  $\alpha$ -helices [104]. Therefore, if the N-terminus of the target protein is structurally organized into a helix, the protein target should be modified, extending the N-terminal sequence with residues endowed of low helical propensity [104]. The subtiligase mutant stabiligase can instead be adopted to label protein targets with a structurally inaccessible N-terminal region, being able to retain enzymatic activity in denaturant conditions in the presence of 0.1% sodium dodecyl sulfate or 4 M guanidinium chloride [104].

### 3.3. Butelase 1

Butelase 1 is the fastest known Asn/Asp-specific peptide ligase isolated from the medicinal plant *Clitoria ternatea* and mediates, in vivo, the backbone cyclization in the biosynthesis of cyclotides [105]. Butelase 1 recognizes the tripeptide Asn/Asp-His-Val at the C-terminus of the target and mediates backbone cyclization by cleaving the consensus sequence between Asn/Asp and His and ligating the Asn/Asp residue to the N-terminal alpha-amine to form a macrocycle. Although mainly proposed for protein macrocyclization [106], butelase 1 was also efficiently adopted for peptide ligation and N-terminal protein labeling [49]. Conveniently, butelase 1 accepts any amino acid at the N-terminal protein target position, except Pro, to form a new Asn/Asp-X peptide bond. However, it exhibits a more stringent amino acid requirement at the second position, showing a strong preference for Cys and the hydrophobic amino acids Ile, Leu, and Val [105]. Butelase 1-mediated intramolecular cyclization proceeds with high efficiency and is an irreversible reaction. In contrast, the intermolecular peptide ligation reaction catalyzed by butelase 1 is a reversible reaction that requires an excess of substrate to reach completion [105]. Using Asn/Asp-thiopeptide as a substrate, which releases a thiol that is a poor competing nucleophile and is not recognizable by butelase 1, the ligation reaction becomes irreversible, allowing the effective use of butelase 1 in peptide ligation and the N-terminal protein labeling reaction (Figure 6c) [49]. A similar behavior was also observed for SrtA-mediated labeling, in which case, the use of depsipeptide substrates enabled the ligation reaction to efficiently proceed using equimolar quantities of substrates and substoichiometric quantities of SrtA [107].

### 3.4. N-Myristoyltransferase

N-Myristoyltransferase (NMT) is an eukaryotic enzyme devoted to catalyzing the acylation with myristic acid (using myristoyl-CoA as the activated reagent) of the alpha-amine group of proteins exposed at the N-terminus of the NMT recognition sequence Gly-X1-X2-X3-Ser/Thr (Lys) (X is any amino acid) (Figure 6d). NMT can also use surrogates of the myristic acid as a substrate [108–110] such as azido or alkyne containing fatty acids, allowing the N-terminal site-specific introduction of an orthogonal reactive handle that can be conveniently exploited for protein decoration using Staudinger ligation [111], CuAAC, or strain-promoted azide-alkyne cycloaddition (SPAAC) click-chemistries [50,112,113]. A protocol for the recombinant co-expression in *Escherichia coli* of NMT and the protein target for the in vivo N-terminal labeling with a clickable myristic acid analog was also provided [114].

### 3.5. Protein Trans-Splicing

Protein trans-splicing (PTS) can also be employed to afford N-terminal protein labeling [24,115,116] (Figure 6e). PTS takes advantage of a particular class of inteins, called split inteins, that, if divided into two domains, can reassemble non-covalently in solution into a functional intein able to catalyze protein splicing by mediating its self-cleavage and ligation of the two exteins. Some inteins such as Ssp DnaB can be split at an amino acid position along the intein sequence that is very close to the N-terminus [117], rendering the N-terminal fragment of the split intein so short that it can be prepared by the chemical route using solid-phase peptide synthesis (SPPS), thus allowing the incorporation of any desired modification such as the introduction of a functional label on the N-terminal alpha-amine group.

## 4. Conclusions

In summary, a wide collection of chemical reactions are available for the site-specific modification of a protein through the targeting of the N-terminal alpha-amine group. The plethora of methodologies available, that have been described herein with their advantages and disadvantages, allow for the decoration of protein targets with very different properties. The labeling strategies can be grouped in direct and indirect approaches, respectively requiring one or two reaction steps to afford N-terminal protein decoration. Indirect strategies appear to be more versatile methodologies with respect to the kind of probe to be introduced as they allow for the functionalization of the protein target with a reactive handle that can be selectively targeted with a wide collection of commercially available molecular probes. In contrast, direct approaches, although being more rapid as the protein modification is accomplished in a single reaction step, require a greater synthetic effort for the preparation of the molecular probe that should be specifically tailored. The more appropriate chemical strategy to be adopted among all those available should be chosen according to the properties of the protein target (especially the type of native N-terminal amino acid) and of the handle to be introduced. Of note, methods based on the control of the reaction pH should be carefully evaluated as, in a protein, the pKa of a functional group could be, sometimes, very different from the expected one because of the influence of the chemical environment.

**Author Contributions:** Conceptualization, L.D.R. and L.D.D.; writing—original draft preparation, L.D.R. and R.D.S.; writing—review and editing, L.D.R., A.R., L.D.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research did not receive any funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Spicer, C.D.; Davis, B.G. Selective chemical protein modification. *Nat. Commun.* **2014**, *5*, 1–14. [[CrossRef](#)] [[PubMed](#)]
2. Koniev, O.; Wagner, A. Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev.* **2015**, *44*, 5743. [[CrossRef](#)] [[PubMed](#)]
3. Shadish, J.A.; DeForest, C.A. Site-selective protein modification: From functionalized proteins to functional biomaterials. *Matter-Us* **2020**, *2*, 50–77. [[CrossRef](#)]
4. Xue, L.; Karpenko, I.A.; Hiblot, J.; Johnsson, K. Imaging and manipulating proteins in live cells through covalent labeling. *Nat. Chem. Biol.* **2015**, *11*, 917–923. [[CrossRef](#)]
5. Stephanopoulos, N.; Francis, M.B. Choosing an effective protein bioconjugation strategy. *Nat. Chem. Biol.* **2011**, *7*, 876–884. [[CrossRef](#)]
6. Turecek, P.L.; Bossard, M.J.; Schoetens, F.; Ivens, I.A. PEGylation of biopharmaceuticals: A review of chemistry and nonclinical safety information of approved drugs. *J. Pharm. Sci.* **2016**, *105*, 460–475. [[CrossRef](#)]
7. Tsuchikama, K.; An, Z.Q. Antibody-drug conjugates: Recent advances in conjugation and linker chemistries. *Protein Cell* **2018**, *9*, 33–46. [[CrossRef](#)]
8. Xu, L.J.; Kuan, S.L.; Weil, T. Contemporary approaches for site-selective dual functionalization of proteins. *Angew. Chem. Int. Ed.* **2021**, *60*, 13757–13777. [[CrossRef](#)]
9. Basle, E.; Joubert, N.; Pucheault, M. Protein chemical modification on endogenous amino acids. *Chem. Biol.* **2010**, *17*, 213–227. [[CrossRef](#)]
10. Chalker, J.M.; Bernardes, G.J.L.; Lin, Y.A.; Davis, B.G. Chemical modification of proteins at cysteine: Opportunities in chemistry and biology. *Chem. Asian J.* **2009**, *4*, 630–640. [[CrossRef](#)]
11. Chen, X.; Muthoosamy, K.; Pfisterer, A.; Neumann, B.; Weil, T. Site-selective lysine modification of native proteins and peptides via kinetically controlled labeling. *Bioconjugate Chem.* **2012**, *23*, 500–508. [[CrossRef](#)]
12. Matos, M.J.; Brown, L.; Bernardim, B.; Guerreiro, A.; Jimenez-Oses, G.; Bernardes, G.J.L. Sequential dual site-selective protein labelling enabled by lysine modification. *Bioorgan. Med. Chem.* **2020**, *28*, 115783. [[CrossRef](#)]
13. Taylor, M.T.; Nelson, J.E.; Suero, M.G.; Gaunt, M.J. A protein functionalization platform based on selective reactions at methionine residues. *Nature* **2018**, *562*, 563–568. [[CrossRef](#)]
14. Xie, L.J.; Liu, L.; Cheng, L. Modifying methionine on proteins. *ChemBioChem* **2020**, *21*, 461–463. [[CrossRef](#)]
15. Lin, S.X.; Yang, X.Y.; Jia, S.; Weeks, A.M.; Hornsby, M.; Lee, P.S.; Nichiporuk, R.V.; Iavarone, A.T.; Wells, J.A.; Toste, F.D.; et al. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **2017**, *355*, 597–602. [[CrossRef](#)]
16. Antos, J.M.; McFarland, J.M.; Iavarone, A.T.; Francis, M.B. Chemoselective tryptophan labeling with rhodium carbenoids at mild pH. *J. Am. Chem. Soc.* **2009**, *131*, 6301–6308. [[CrossRef](#)]
17. Chen, Z.; Popp, B.V.; Bovet, C.L.; Ball, Z.T. Site-specific protein modification with a dirhodium metallopeptide catalyst. *ACS Chem. Biol.* **2011**, *6*, 920–925. [[CrossRef](#)]
18. Joshi, P.N.; Rai, V. Single-site labeling of histidine in proteins, on-demand reversibility, and traceless metal-free protein purification. *Chem. Commun.* **2019**, *55*, 1100–1103. [[CrossRef](#)]
19. Sato, S.; Matsumura, M.; Kadosono, T.; Abe, S.; Ueno, T.; Ueda, H.; Nakamura, H. Site-selective protein chemical modification of exposed tyrosine residues using tyrosine click reaction. *Bioconjugate Chem.* **2020**, *31*, 1417–1424. [[CrossRef](#)]
20. Lee, K.J.; Kang, D.; Park, H.S. Site-specific labeling of proteins using unnatural amino acids. *Mol. Cells* **2019**, *42*, 386–396.
21. Chin, J.W. Expanding and reprogramming the genetic code of cells and animals. *Ann. Rev. Biochem.* **2014**, *83*, 379–408. [[CrossRef](#)] [[PubMed](#)]
22. Dawson, P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776–779. [[CrossRef](#)] [[PubMed](#)]
23. Muir, T.W.; Sondhi, D.; Cole, P.A. Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705–6710. [[CrossRef](#)] [[PubMed](#)]
24. De Rosa, L.; Russomanno, A.; Romanelli, A.; D’Andrea, L.D. Semi-synthesis of labeled proteins for spectroscopic applications. *Molecules* **2013**, *18*, 440–465. [[CrossRef](#)]
25. D’Andrea, L.; Romanelli, A. *Chemical Ligation: Tools for Biomolecule Synthesis and Modification*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2017.
26. Rosen, C.B.; Francis, M.B. Targeting the N terminus for site-selective protein modification. *Nat. Chem. Biol.* **2017**, *13*, 697–705. [[CrossRef](#)]
27. Sereda, T.J.; Mant, C.T.; Quinn, A.M.; Hodges, R.S. Effect of alpha-amino group on peptide retention behavior in reversed-phase chromatography—Determination of the Pk(a) values of the alpha-amino group of 19 different N-terminal amino-acid-residues. *J. Chromatogr.* **1993**, *646*, 17–30. [[CrossRef](#)]
28. Jacob, E.; Unger, R. A tale of two tails: Why are terminal residues of proteins exposed? *Bioinformatics* **2007**, *23*, E225–E230. [[CrossRef](#)]
29. Chen, D.; Disotuar, M.M.; Xiong, X.C.; Wang, Y.X.; Chou, D.H.C. Selective N-terminal functionalization of native peptides and proteins. *Chem. Sci.* **2017**, *8*, 2717–2722. [[CrossRef](#)]



30. Deng, J.R.; Lai, N.C.H.; Kung, K.K.Y.; Yang, B.; Chung, S.F.; Leung, A.S.L.; Choi, M.C.; Leung, Y.C.; Wong, M.K. N-Terminal selective modification of peptides and proteins using 2-ethynylbenzaldehydes. *Commun. Chem.* **2020**, *3*, 1. [[CrossRef](#)]
31. Singudasa, R.; Adusumalli, S.R.; Joshi, P.N.; Rai, V. A phthalimidation protocol that follows protein defined parameters. *Chem. Commun.* **2015**, *51*, 473–476. [[CrossRef](#)]
32. Raj, M.; Wu, H.B.; Blosser, S.L.; Vittoria, M.A.; Arora, P.S. Aldehyde capture ligation for synthesis of native peptide bonds. *J. Am. Chem. Soc.* **2015**, *137*, 6932–6940. [[CrossRef](#)]
33. Busch, G.K.; Tate, E.W.; Gaffney, P.R.J.; Rosivatz, E.; Woscholski, R.; Leatherbarrow, R.J. Specific N-terminal protein labelling: Use of FMDV 3C(pro) protease and native chemical ligation. *Chem. Commun.* **2008**, *29*, 3369–3371. [[CrossRef](#)]
34. Ren, H.J.; Xiao, F.; Zhan, K.; Kim, Y.P.; Xie, H.X.; Xia, Z.Y.; Rao, J. A biocompatible condensation reaction for the labeling of terminal cysteine residues on proteins. *Angew. Chem. Int. Ed.* **2009**, *48*, 9658–9662. [[CrossRef](#)]
35. Casi, G.; Huguenin-Dezot, N.; Zuberbuhler, K.; Scheuermann, J.; Neri, D. Site-specific traceless coupling of potent cytotoxic drugs to recombinant antibodies for pharmacodelivery. *J. Am. Chem. Soc.* **2012**, *134*, 5887–5892. [[CrossRef](#)]
36. Bandyopadhyay, A.; Cambray, S.; Gao, J.M. Fast and selective labeling of N-terminal cysteines at neutral pH via thiazolidino boronate formation. *Chem. Sci.* **2016**, *7*, 4589–4593. [[CrossRef](#)]
37. Li, X.F.; Zhang, L.S.; Hall, S.E.; Tam, J.P. A new ligation method for N-terminal tryptophan-containing peptides using the Pictet-Spengler reaction. *Tetrahedron Lett.* **2000**, *41*, 4069–4073. [[CrossRef](#)]
38. MacDonald, J.I.; Munch, H.K.; Moore, T.; Francis, M.B. One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. *Nat. Chem. Biol.* **2015**, *11*, 326–331. [[CrossRef](#)]
39. Obermeyer, A.C.; Jarman, J.B.; Francis, M.B. N-Terminal modification of proteins with o-aminophenols. *J. Am. Chem. Soc.* **2014**, *136*, 9572–9579. [[CrossRef](#)]
40. Purushottam, L.; Adusumalli, S.R.; Singh, U.; Unnikrishnan, V.B.; Rawale, D.G.; Gujrati, M.; Mishra, R.K.; Rai, V. Single-site glycine-specific labeling of proteins. *Nat. Commun.* **2019**, *10*, 1–9. [[CrossRef](#)]
41. Martos-Maldonado, M.C.; Hjuler, C.T.; Sorensen, K.K.; Thygesen, M.B.; Rasmussen, J.E.; Villadsen, K.; Midtgaard, S.R.; Kol, S.; Schoffelen, S.; Jensen, K.J. Selective N-terminal acylation of peptides and proteins with a Gly-His tag sequence. *Nat. Commun.* **2018**, *9*, 3307. [[CrossRef](#)]
42. Gilmore, J.M.; Scheck, R.A.; Esser-Kahn, A.P.; Joshi, N.S.; Francis, M.B. N-terminal protein modification through a biomimetic transamination reaction. *Angew. Chem. Int. Ed.* **2006**, *45*, 5307–5311. [[CrossRef](#)]
43. Witus, L.S.; Netirojanakul, C.; Palla, K.S.; Muehl, E.M.; Weng, C.H.; Iavarone, A.T.; Francis, M.B. Site-Specific protein transamination using N-methylpyridinium-4-carboxaldehyde. *J. Am. Chem. Soc.* **2013**, *135*, 17223–17229. [[CrossRef](#)]
44. De Rosa, L.; Di Stasi, R.; Longhitano, L.; D’Andrea, L.D. Labeling of VEGFR1D2 through oxime ligation. *Bioorg. Chem.* **2019**, *91*, 103160. [[CrossRef](#)]
45. Schoffelen, S.; van Eldijk, M.B.; Rooijackers, B.; Raijmakers, R.; Heck, A.J.R.; van Hest, J.C.M. Metal-free and pH-controlled introduction of azides in proteins. *Chem. Sci.* **2011**, *2*, 701–705. [[CrossRef](#)]
46. Chan, A.O.Y.; Ho, C.M.; Chong, H.C.; Leung, Y.C.; Huang, J.S.; Wong, M.K.; Che, C.M. Modification of N-Terminal alpha-Amino Groups of Peptides and Proteins Using Ketenes. *J. Am. Chem. Soc.* **2012**, *134*, 2589–2598. [[CrossRef](#)]
47. Theile, C.S.; Witte, M.D.; Blom, A.E.M.; Kundrat, L.; Ploegh, H.L.; Guimaraes, C.P. Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **2013**, *8*, 1800–1807. [[CrossRef](#)]
48. Weeks, A.M.; Wells, J.A. N-Terminal modification of proteins with subtiligase specificity variants. *Curr. Protoc. Chem. Biol.* **2020**, *12*, e79. [[CrossRef](#)]
49. Nguyen, G.K.T.; Cao, Y.; Wang, W.; Liu, C.F.; Tam, J.P. Site-specific N-terminal labeling of peptides and proteins using butelase 1 and thiodipeptide. *Angew. Chem. Int. Ed.* **2015**, *54*, 15694–15698. [[CrossRef](#)]
50. Heal, W.P.; Wickramasinghe, S.R.; Bowyer, P.W.; Holder, A.A.; Smith, D.F.; Leatherbarrow, R.J.; Tate, E.W. Site-specific N-terminal labelling of proteins in vitro and in vivo using N-myristoyl transferase and bioorthogonal ligation chemistry. *Chem. Commun.* **2008**, *4*, 480–482. [[CrossRef](#)]
51. Nanda, J.S.; Lorsch, J.R. Labeling a protein with fluorophores using NHS ester derivitization. *Lab. Methods Enzymol. Protein A* **2014**, *536*, 87–94.
52. Selo, I.; Negroni, L.; Creminon, C.; Grassi, J.; Wal, J.M. Preferential labeling of alpha-amino N-terminal groups in peptides by biotin, application to the detection of specific anti-peptide antibodies by enzyme immunoassays. *J. Immunol. Methods* **1996**, *199*, 127–138. [[CrossRef](#)]
53. Berezovsky, I.N.; Kilosanidze, G.T.; Tumanyan, V.G.; Kisselev, L.L. Amino acid composition of protein termini are biased in different manners. *Protein Eng.* **1999**, *12*, 23–30. [[CrossRef](#)] [[PubMed](#)]
54. Gentle, I.E.; De Souza, D.P.; Baca, M. Direct production of proteins with N-terminal cysteine for site-specific conjugation. *Bioconjugate Chem.* **2004**, *15*, 658–663. [[CrossRef](#)] [[PubMed](#)]
55. Villain, M.; Vizzavona, J.; Rose, K. Covalent capture: A new tool for the purification of synthetic and recombinant polypeptides. *Chem. Biol.* **2001**, *8*, 673–679. [[CrossRef](#)]
56. Romanelli, A.; Shekhtman, A.; Cowburn, D.; Muir, T.W. Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6397–6402. [[CrossRef](#)] [[PubMed](#)]

57. Tolbert, T.J.; Wong, C.H. Conjugation of glycopeptide thioesters to expressed protein fragments: Semisynthesis of glycosylated interleukin-2. *Methods Mol. Biol.* **2004**, *283*, 255–266. [[PubMed](#)]
58. Liu, D.S.; Xu, R.; Dutta, K.; Cowburn, D. N-terminal cysteinyl proteins can be prepared using thrombin cleavage. *FEBS Lett.* **2008**, *582*, 1163–1167. [[CrossRef](#)]
59. Xu, M.Q.; Evans, T.C. Intein-mediated ligation and cyclization of expressed proteins. *Methods* **2001**, *24*, 257–277. [[CrossRef](#)]
60. Wissner, R.F.; Batjargal, S.; Fadzen, C.M.; Petersson, E.J. Labeling proteins with fluorophore/thioamide forster resonant energy transfer pairs by combining unnatural amino acid mutagenesis and native chemical ligation. *J. Am. Chem. Soc.* **2013**, *135*, 6529–6540. [[CrossRef](#)]
61. Xiao, J.P.; Tolbert, T.J. Synthesis of N-Terminally linked protein dimers and trimers by a combined native chemical ligation-CuAAC click chemistry strategy. *Org. Lett.* **2009**, *11*, 4144–4147. [[CrossRef](#)]
62. Xiao, J.P.; Hamilton, B.S.; Tolbert, T.J. Synthesis of N-terminally linked protein and peptide dimers by native chemical ligation. *Bioconjugate Chem.* **2010**, *21*, 1943–1947. [[CrossRef](#)]
63. Xiao, J.P.; Burn, A.; Tolbert, T.J. Increasing solubility of proteins and peptides by site-specific modification with betaine. *Bioconjugate Chem.* **2008**, *19*, 1113–1118. [[CrossRef](#)]
64. Hawala, I.; De Rosa, L.; Aime, S.; D’Andrea, L.D. An innovative approach for the synthesis of dual modality peptide imaging probes based on the native chemical ligation approach. *Chem. Commun.* **2020**, *56*, 3500–3503. [[CrossRef](#)]
65. Dempsey, D.R.; Jiang, H.J.; Kalin, J.H.; Chen, Z.; Cole, P.A. Site-specific protein labeling with N-hydroxysuccinimide-esters and the analysis of ubiquitin ligase mechanisms. *J. Am. Chem. Soc.* **2018**, *140*, 9374–9378. [[CrossRef](#)]
66. De Rosa, L.; Cortajarena, A.L.; Romanelli, A.; Regan, L.; D’Andrea, L.D. Site-specific protein double labeling by expressed protein ligation: Applications to repeat proteins. *Org. Biomol. Chem.* **2012**, *10*, 273–280. [[CrossRef](#)]
67. Jiang, H.; Cole, P.A. N-terminal protein labeling with N-hydroxysuccinimide esters and microscale thermophoresis measurements of protein-protein interactions using labeled protein. *Curr. Protoc.* **2021**, *1*, e14. [[CrossRef](#)]
68. White, E.H.; McCapra, F.; Field, G.F. The structure and synthesis of firefly luciferin. *J. Am. Chem. Soc.* **1963**, *85*, 337–343. [[CrossRef](#)]
69. Liang, G.L.; Ren, H.J.; Rao, J.H. A biocompatible condensation reaction for controlled assembly of nanostructures in living cells. *Nat. Chem.* **2010**, *2*, 54–60. [[CrossRef](#)]
70. Jeon, J.; Shen, B.; Xiong, L.Q.; Miao, Z.; Lee, K.H.; Rao, J.; Chin, F.T. Efficient method for site-specific F-18-labeling of biomolecules using the rapid condensation reaction between 2-cyanobenzothiazole and cysteine. *Bioconjugate Chem.* **2012**, *23*, 1902–1908. [[CrossRef](#)]
71. Liu, C.F.; Tam, J.P. Chemical ligation approach to form a peptide-bond between unprotected peptide segments—Concept and model study. *J. Am. Chem. Soc.* **1994**, *116*, 4149–4153. [[CrossRef](#)]
72. Zhang, L.S.; Tam, J.P. Thiazolidine formation as a general and site-specific conjugation method for synthetic peptides and proteins. *Anal. Biochem.* **1996**, *233*, 87–93. [[CrossRef](#)]
73. Liu, C.F.; Tam, J.P. Peptide segment ligation strategy without use of protecting groups. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6584–6588. [[CrossRef](#)]
74. Zhao, Z.G.; Im, J.S.; Lam, K.S.; Lake, D.F. Site-specific modification of a single-chain antibody using a novel glyoxylyl-based labeling reagent. *Bioconjugate Chem.* **1999**, *10*, 424–430. [[CrossRef](#)]
75. Faustino, H.; Silva, M.J.S.A.; Veiros, L.F.; Bernardes, G.J.L.; Gois, P.M.P. Iminoboronates are efficient intermediates for selective, rapid and reversible N-terminal cysteine functionalization. *Chem. Sci.* **2016**, *7*, 6280. [[CrossRef](#)]
76. Pulka, K.; Slupska, M.; Puzsko, A.; Misiak, M.; Wilczek, M.; Kozminski, W.; Misicka, A. Peptides and peptidoaldehydes as substrates for the Pictet-Spengler reaction. *J. Pept. Sci.* **2013**, *19*, 433–440. [[CrossRef](#)]
77. Agarwal, P.; van der Weijden, J.; Sletten, E.M.; Rabuka, D.; Bertozzi, C.R. A Pictet-Spengler ligation for protein chemical modification. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 46–51. [[CrossRef](#)]
78. Sasaki, T.; Kodama, K.; Suzuki, H.; Fukuzawa, S.; Tachibana, K. N-terminal labeling of proteins by the Pictet-Spengler reaction. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4550–4553. [[CrossRef](#)]
79. Waugh, D.S. An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr. Purif.* **2011**, *80*, 283–293. [[CrossRef](#)] [[PubMed](#)]
80. Hirel, P.H.; Schmitter, J.M.; Dessen, P.; Fayat, G.; Blanquet, S. Extent of N-terminal methionine excision from escherichia-coli proteins is governed by the side-chain length of the penultimate amino-acid. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 8247–8251. [[CrossRef](#)] [[PubMed](#)]
81. Geoghegan, K.F.; Dixon, H.B.F.; Rosner, P.J.; Hoth, L.R.; Lanzetti, A.J.; Borzilleri, K.A.; Marr, E.S.; Pezzullo, L.H.; Martin, L.B.; LeMotte, P.K.; et al. Spontaneous alpha-N-6-phosphogluconoylation of a “His tag” in Escherichia coli: The cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **1999**, *267*, 169–184. [[CrossRef](#)] [[PubMed](#)]
82. Scheck, R.A.; Dedeo, M.T.; Lavarone, A.T.; Francis, M.B. Optimization of a biomimetic transamination reaction. *J. Am. Chem. Soc.* **2008**, *130*, 11762–11770. [[CrossRef](#)]
83. Scheck, R.A.; Francis, M.B. Regioselective labeling of antibodies through N-terminal transamination. *ACS Chem. Biol.* **2007**, *2*, 247–251. [[CrossRef](#)]
84. Buckley, T.F.R.H. Mild and simple biomimetic conversion of amines to carbonyl compounds. *J. Am. Chem. Soc.* **1982**, *104*, 4446–4450. [[CrossRef](#)]

85. Palla, K.S.; Witus, L.S.; Mackenzie, K.J.; Netirojjanakul, C.; Francis, M.B. Optimization and expansion of a site-selective N-methylpyridinium-4-carboxaldehyde-mediated transamination for bacterially expressed proteins. *J. Am. Chem. Soc.* **2015**, *137*, 1123–1129. [[CrossRef](#)]
86. Geoghegan, K.F.; Stroh, J.G. Site-directed conjugation of nonpeptide groups to peptides and proteins via periodate-oxidation of a 2-amino alcohol—Application to modification at N-terminal serine. *Bioconjugate Chem.* **1992**, *3*, 138–146. [[CrossRef](#)]
87. El-Mahdi, O.; Melnyk, O. Alpha-oxo aldehyde or glyoxyl group chemistry in peptide bioconjugation. *Bioconjugate Chem.* **2013**, *24*, 735–765. [[CrossRef](#)]
88. Di Stasi, R.; Diana, D.; Capasso, D.; Palumbo, R.; Romanelli, A.; Pedone, C.; Fattorusso, R.; D’Andrea, L.D. VEGFR1(D2) in drug discovery: Expression and molecular characterization. *Biopolymers* **2010**, *94*, 800–809. [[CrossRef](#)]
89. Ning, X.; Temming, R.P.; Dommerholt, J.; Guo, J.; Ania, D.B.; Debets, M.F.; Wolfert, M.A.; Boons, G.J.; van Delft, F.L. Protein modification by strain-promoted alkyne-nitrone cycloaddition. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 3065–3068. [[CrossRef](#)]
90. Tornøe, C.W.; Christensen, C.; Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057–3064. [[CrossRef](#)]
91. Rostovtsev, V.V.; Green, L.G.; Fokin, V.V.; Sharpless, K.B. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599. [[CrossRef](#)]
92. Popp, M.W.L.; Ploegh, H.L. Making and breaking peptide bonds: Protein engineering using sortase. *Angew. Chem. Int. Ed.* **2011**, *50*, 5024–5032. [[CrossRef](#)] [[PubMed](#)]
93. Antos, J.M.; Chew, G.L.; Guimaraes, C.P.; Yoder, N.C.; Grotenbreg, G.M.; Popp, M.W.L.; Ploegh, H.L. Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity. *J. Am. Chem. Soc.* **2009**, *131*, 10800–10801. [[CrossRef](#)] [[PubMed](#)]
94. Yamamoto, T.; Nagamune, T. Expansion of the sortase-mediated labeling method for site-specific N-terminal labeling of cell surface proteins on living cells. *Chem. Commun.* **2009**, *9*, 1022–1024. [[CrossRef](#)] [[PubMed](#)]
95. Antos, J.M.; Miller, G.M.; Grotenbreg, G.M.; Ploegh, H.L. Lipid modification of proteins through sortase-catalyzed transpeptidation. *J. Am. Chem. Soc.* **2008**, *130*, 16338–16343. [[CrossRef](#)]
96. Guimaraes, C.P.; Carette, J.E.; Varadarajan, M.; Antos, J.; Popp, M.W.; Spooner, E.; Brummelkamp, T.R.; Ploegh, H.L. Identification of host cell factors required for intoxication through use of modified cholera toxin. *J. Cell Biol.* **2011**, *195*, 751–764. [[CrossRef](#)]
97. Popp, M.W.; Artavanis-Tsakonas, K.; Ploegh, H.L. Substrate filtering by the active site crossover loop in UCHL3 revealed by sortagging and gain-of-function mutations. *J. Biol. Chem.* **2009**, *284*, 3593–3602. [[CrossRef](#)]
98. Guimaraes, C.P.; Witte, M.D.; Theile, C.S.; Bozkurt, G.; Kundrat, L.; Blom, A.E.M.; Ploegh, H.L. Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **2013**, *8*, 1787–1799. [[CrossRef](#)]
99. Popp, M.W.; Dougan, S.K.; Chuang, T.Y.; Spooner, E.; Ploegh, H.L. Sortase-catalyzed transformations that improve the properties of cytokines. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3169–3174. [[CrossRef](#)]
100. Weeks, A.M.; Wells, J.A. Subtiligase-catalyzed peptide ligation. *Chem. Rev.* **2020**, *120*, 3127–3160. [[CrossRef](#)]
101. Abrahmsen, L.; Tom, J.; Burnier, J.; Butcher, K.A.; Kossiakoff, A.; Wells, J.A. Engineering subtilisin and its substrates for efficient ligation of peptide-bonds in aqueous-solution. *Biochemistry* **1991**, *30*, 4151–4159. [[CrossRef](#)]
102. Tan, X.H.; Yang, R.L.; Liu, C.F. Facilitating subtiligase-catalyzed peptide ligation reactions by using peptide thioester substrates. *Org. Lett.* **2018**, *20*, 6691–6694. [[CrossRef](#)]
103. Weeks, A.M.; Wells, J.A. Engineering peptide ligase specificity by proteomic identification of ligation sites. *Nat. Chem. Biol.* **2018**, *14*, 50. [[CrossRef](#)]
104. Chang, T.K.; Jackson, D.Y.; Burnier, J.P.; Wells, J.A. Subtiligase—A tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12544–12548. [[CrossRef](#)]
105. Nguyen, G.K.T.; Wang, S.J.; Qiu, Y.B.; Hemu, X.; Lian, Y.L.; Tam, J.P. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **2014**, *10*, 732–738. [[CrossRef](#)]
106. Nguyen, G.K.T.; Kam, A.; Loo, S.; Jansson, A.E.; Pan, L.X.; Tam, J.P. Butelase 1: A versatile ligase for peptide and protein macrocyclization. *J. Am. Chem. Soc.* **2015**, *137*, 15398–15401. [[CrossRef](#)]
107. Williamson, D.J.; Fascione, M.A.; Webb, M.E.; Turnbull, W.B. Efficient N-terminal labeling of proteins by use of sortase. *Angew. Chem. Int. Ed.* **2012**, *51*, 9377–9380. [[CrossRef](#)]
108. Devadas, B.; Lu, T.B.; Katoh, A.; Kishore, N.S.; Wade, A.C.; Mehta, P.P.; Rudnick, D.A.; Bryant, M.L.; Adams, S.P.; Li, Q.; et al. Substrate-specificity of *Saccharomyces cerevisiae* myristoyl-coa-protein N-myristoyltransferase—Analysis of fatty-acid analogs containing carbonyl groups, nitrogen heteroatoms, and nitrogen-heterocycles in an in vitro enzyme assay and subsequent identification of inhibitors of human immunodeficiency virus-I replication. *J. Biol. Chem.* **1992**, *267*, 7224–7239.
109. Kishore, N.S.; Lu, T.B.; Knoll, L.J.; Katoh, A.; Rudnick, D.A.; Mehta, P.P.; Devadas, B.; Huhn, M.; Atwood, J.L.; Adams, S.P.; et al. The substrate specificity of *Saccharomyces cerevisiae* myristoyl-CoA:protein N-myristoyltransferase. Analysis of myristic acid analogs containing oxygen, sulfur, double bonds, triple bonds, and/or an aromatic residue. *J. Biol. Chem.* **1991**, *266*, 8835–8855. [[CrossRef](#)]
110. Lu, T.; Li, Q.; Katoh, A.; Hernandez, J.; Duffin, K.; Jackson-Machelski, E.; Knoll, L.J.; Gokel, G.W.; Gordon, J.I. The substrate specificity of *Saccharomyces cerevisiae* myristoyl-CoA: Protein N-myristoyltransferase. Polar probes of the enzyme’s myristoyl-CoA recognition site. *J. Biol. Chem.* **1994**, *269*, 5346–5357. [[CrossRef](#)]

111. Hang, H.C.; Geutjes, E.J.; Grotenbreg, G.; Pollington, A.M.; Bijlmakers, M.J.; Ploegh, H.L. Chemical probes for the rapid detection of fatty-acylated proteins in mammalian cells. *J. Am. Chem. Soc.* **2007**, *129*, 2744–2745. [[CrossRef](#)]
112. Heal, W.P.; Wickramasinghe, S.R.; Leatherbarrow, R.J.; Tate, E.W. N-Myristoyl transferase-mediated protein labelling in vivo. *Org. Biomol. Chem.* **2008**, *6*, 2308–2315. [[CrossRef](#)]
113. Kulkarni, C.; Kinzer-Ursem, T.L.; Tirrell, D.A. Selective functionalization of the protein N terminus with N-myristoyl transferase for bioconjugation in cell lysate. *ChemBioChem* **2013**, *14*, 1958–1962. [[CrossRef](#)]
114. Heal, W.P.; Wright, M.H.; Thinon, E.; Tate, E.W. Multifunctional protein labeling via enzymatic N-terminal tagging and elaboration by click chemistry. *Nat. Protoc.* **2011**, *7*, 105–117. [[CrossRef](#)] [[PubMed](#)]
115. Wood, D.W.; Camarero, J.A. Intein applications: From protein purification and labeling to metabolic control methods. *J. Biol. Chem.* **2014**, *289*, 14512–14519. [[CrossRef](#)]
116. Aranko, A.S.; Wlodawer, A.; Iwai, H. Nature's recipe for splitting inteins. *Protein Eng. Des. Sel.* **2014**, *27*, 263–271. [[CrossRef](#)]
117. Ludwig, C.; Pfeiff, M.; Linne, U.; Mootz, H.D. Ligation of a synthetic peptide to the N terminus of a recombinant protein using semisynthetic protein trans-splicing. *Angew. Chem. Int. Ed.* **2006**, *45*, 5218–5221. [[CrossRef](#)]