UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

Laboratoire de Conception et Application de Molécules Bioactives

THÈSE présentée par :

Charlotte SORNAY

soutenue le : **05 février 2021**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Chimie Biologique et Thérapeutique

Nouvelles stratégies de bioconjugaison de protéines natives

THÈSE dirigée par :
M WAGNER Alain   Directeur de recherche, UMR 7199, Université de Strasbourg

THÈSE encadrée par :
M CHAUBET Guilhem   Chargé de recherche, UMR 7199, Université de Strasbourg

RAPPORTEURS :
M BERNARDES Gonçalo   Professeur, University of Cambridge
M PAPOT Sébastien   Professeur, Université de Poitiers

EXAMINATEURS :
Mme CIANFERANI Sarah   Directrice de recherche, UMR 7178, Université de Strasbourg
M ZHU Jieping   Professeur, École Polytechnique Fédérale de Lausanne

"If you are going through hell,

keep going."

Winston Churchill

REMERCIEMENTS

Ces travaux de thèse ont été réalisés au laboratoire de Conception et Applications de Molécules Bioactives à la faculté de pharmacie d'Illkirch-Graffenstaden, au sein de l'équipe Chimie BioFonctionnelle (BFC) sous la direction du Dr Alain Wagner et la supervision du Dr Guilhem Chaubet.

Je tiens tout d'abord à remercier Alain pour m'avoir accueillie dans son équipe. Même si tu as eu un peu de mal à te souvenir de mon prénom au cours de ces trois années – on se souviendra particulièrement de « machine », Juliette ou même Tony – tu as toujours été à l'écoute dans mes moments de doute et su m'encourager et me soutenir dans mes différents projets.

Je pense que cette thèse se serait déroulée différemment si, Guilhem, tu n'avais pas été là pour l'encadrer. Du premier au dernier jour de cette thèse, tu as été la personne sur qui j'ai toujours pu compter, que ce soit pour me guider au quotidien dans les projets, ou bien pour m'amener à l'aéroport en catastrophe. Merci pour la confiance que tu m'as accordée au cours de ces trois années et j'espère que nous pourrons retravailler ensemble à l'avenir. Je te remercie également pour toutes les corrections que tu as apportées à ce manuscrit.

Je souhaite également remercier les différents membres du jury, Pr Gonçalo Bernardes, Pr Sébastien Papot, Pr Jieping Zhu et Dr Sarah Cianférani pour avoir accepté de juger ces travaux. Je tiens également à saluer Dr Alexandre Specht pour avoir, au côté du Dr Sarah Cianférani, participé aux deux suivis de thèse.

Un immense merci au Dr Sarah Cianférani et à son équipe, en particulier aux « bogosses du labo » - Stéphane, Steve, Anthony et Thomas - pour avoir, tout au long de cette thèse, accepté d'analyser beaucoup trop de conjugués par semaine. Comme je vous l'ai déjà dit plusieurs fois, sans vous, ces travaux n'auraient jamais pu avancer aussi vite et être réalisés. Alors, encore une fois, merci infiniment.

J'aimerais également saluer nos collaborateurs lausannois, le Pr Jérôme Waser et son équipe, qui nous ont fait confiance pour tester leurs iodes hypervalents sur protéines natives.

Direction de la **recherche**
et de la **valorisation**
Collège **doctoral**
Université de Strasbourg

Déclaration sur l'honneur
Declaration of Honour

J'affirme être informé que le plagiat est une faute grave susceptible de mener à des sanctions administratives et disciplinaires pouvant aller jusqu'au renvoi de l'Université de Strasbourg et passible de poursuites devant les tribunaux de la République Française.

Je suis conscient(e) que l'absence de citation claire et transparente d'une source empruntée à un tiers (texte, idée, raisonnement ou autre création) est constitutive de plagiat.

Au vu de ce qui précède,
j'atteste sur l'honneur que le travail décrit dans mon manuscrit de thèse est un travail original et que je n'ai pas eu recours au plagiat ou à toute autre forme de fraude.
I affirm that I am aware that plagiarism is a serious misconduct that may lead to administrative and disciplinary sanctions up to dismissal from the University of Strasbourg and liable to prosecution in the courts of the French Republic.

I am aware that the absence of a clear and transparent citation of a source borrowed from a third party (text, idea, reasoning or other creation) is constitutive of plagiarism.

*In view of the foregoing, **I hereby certify that the work described in my thesis manuscript is original work and that I have not resorted to plagiarism or any other form of fraud.***

Nom : Sornay Prénom : Charlotte

Ecole doctorale : ED222

Laboratoire : CAMB (UMR7199) - BFC

Date : 15.03.2021

Signature :

ABBREVIATIONS

2-EBA	2-ethynylbenzaldehyde		CUR	curtain gas
2-FPBA	2-formylphenyl boronic acid		Cy5	cyanine-5
2D	2 dimensions		Cys, C	cysteine
3e-4c	three-centre-four-electron		d	day
4 AMS	4 angstrom molecular sieves		Da	dalton
4CzIPN	2,4,5,6-tetra(9H-carbazol-9-yl)isophtalonitrile		DAR	drug-to-antibody ratio
5 AMS	5 angstrom molecular sieves		DBCO	dibenzocyclooctyne
ABX	1-(4-azidobut-1-yn-1-yl)-1λ^3-benzo[d][1,2]iodaoxol-3(1H)-one		DBU	1,8-diazabicyclo(5.4.0)undec-7-ene
Ac	acetyl		DCC	N,N'-dicyclohexylcarbodiimide
AcK	N^ε-acetyl-lysine		DCE	1,2-dichloroethane
ACL	aldehyde capture ligation		DCM	dichloromethane
ACN	acetonitrile		deg	degrees
ADC	antibody-drug conjugate		DFT	density functional theory
ADHP	2-amino-4,6-dihydroxy-pyrimidine		DhaA	*Rhodococcus* dehalogenase
ADPN	arylene dipropionitrile		DIBAC	dibenzoazacyclooctyne
ADT	2-aryl-4,5-dihydrothiazole		DIBO	dibenzocyclooctynol
AGT	O^6-alkylguanine-DNA alkyltransferase		DIFO	difluorinated cyclooctyne
Ala, A	alanine		DIPEA	N,N-diisopropylethylamine
AlkK	N^ε-(propargyloxycarbonyl)-L-lysine		DMAP	4-dimethylaminopyridine
Ar	aryl		DMEM	Dulbecco's modified eagle media
Arg, R	arginine		DMF	N,N-dimethylformamide
Asn, N	asparagine		DMSO	dimethylsulfoxide
Asp, D	aspartate		DNA	Deoxyribonucleic acid
ATG8	autophagy-related protein 8		DoC	degree of conjugation
BARAC	biarylazacyclooctynone		DP	declustering potential
BB	borate-buffered		DT	diphteria toxoid
BBS	borate-buffered saline		DTNB	5,5'-dithiosbis-(2-nitrobenzoic acid)
BCA	bicinchoninic acid		DTT	dithiothreitol
BCN	bicyclo[6.1.0]nonyne		DVP	divinylpyrimidine
BME	2-mercaptoethanol		DVT	divinyltriazine
Boc	*tert*-butyloxycarbonyl		e.g.	for example
BOP	benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate		EA	ethyl acetate
BSA	bovine serum albumin		EBX	ethynyl benziodoxol(on)e
BX	benziodoxol(on)e		EC50	half maximal effective concentration
c	centi		EDC	N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide
cal	calory		EDTA	ethylenediaminetetraacetic acid
CBT	2-cyanobenzothiazole		EFS	Etablissement Français du sang
CBX	cyano benziodoxol(on)e		EPFL	Ecole Polytechnique Fédérale de Lausanne
CD	circular dichroism		equiv.	equivalents
CD20	B-lymphocyte antigen		ESI	electrospray ionization
cHex	cyclohexyl		Et	ethyl
CMC	o-carboxymethylcellulose		eV	electron volt
CpK	N^ε-(1-methylcycloprop-2-enecarboxamido)lysine		Fab	Fragment antigen binding
CPs	capsular polysaccharides		FBDP	4-formylbenzene diazonium hexafluorophosphate
CuAAC	copper-catalyzed alkyne-azide cycloaddition		FBS	fetal bovine serum

FDA	Food and Drug Administration
FDNB	fluorodinitrobenzene
FGE	formylglycine generating enzyme
FITC	fluorescein isothiocyanate
FKBP	FK506 binding protein
Fmoc	fluorenylmethyloxycarbonyl
FPP	farnesyl pyrophosphate
g	gram
GFP	green fluorescent protein
GLP-1	glucagon-like protein 1
Glu, E	glutamate
Gln, Q	glutamine
Gly, G	glycine
GS1	ion source gas 1
GSH	glutathione
h	hour
HBTU	N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate
HC	heavy chain
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2	Human epidermal growth factor receptor 2
HEWL	Hen Egg-White lysosyme
His, H	histidine
HMBC	Heteronuclear Multiple Bond Correlation
HOMO	Highest Occupied Molecular Orbital
Hpg	homopropargylglycine
HPLC	High-performance liquid chromatography
HRMS	high-resolution mass spectra
HRP	horseradish peroxidase
HSA	Human serum albumin
HSQC	Heteronuclear single quantum coherence spectroscopy
Hz	hertz
i.e.	id est
IAA	iodoacetamide
IBA-OBz	1-benzoyloxy-1,2-benziodoxol-3-(1H)-one
IDA	information-dependent acquisition
IDT	Integrated DNA Technologies
iEDDA	inverse-electron demand Diels-Alder
IFN	interferon
IgG1	immunoglobulin G 1
IHT	interface heater temperature
Ile, I	isoleucine
IR	infrared
ISVF	ionspray voltage floating
J	coupling constant
k	kilo
keto-ABNO	9-Azabicyclo[3,3,1]nonan-3-one-9-oxyl
KetoK	2-amino-8-oxononanoic acid
L	liter
LAP	lithium phenyl-2,4,6-trimethylbenzoylphosphinate
LC	light chain
LC-MS	liquid chromatography mass spectrometry
Leu, L	leucine
Lys, K	lysine
m	milli
M	molar
m	meter
M	mega
m/z	mass-to-charge ratio
MCR	multicomponent reaction
Me	methyl
Met, M	methionine
min	minute
MMAE	monomethyl auristatin E
MO	molecular orbital
MP	melting point
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide
MWNTs	multiwall carbon nanotubes
n	nano
NAA	natural amino acid
NCL	native chemical ligation
NGM	new generation of maleimides
NHS	N-hydroxysuccimide
Ni(II)-NTA	Ni(II)-Nitrilotriacetic acid
NMR	nuclear magnetic resonance
ON	overnight
ON20	20-base oligonucleotide
p	pico
p-BTFP-iodosodilactone	6-(3,5-bis(trifluoromethyl)phenyl)-1H,4H-2aλ^3-ioda-2,3-dioxacyclopenta[hi]indene-1,4-dione
p-Tpa	p-(2-tetrazole)phenylalanine
PB	phosphate-buffered
PBS	phosphate-buffered saline
PD	dibromopyridazinedione
PEG	polyethylene glycol
pF-EBX	5-fluoro-1-((trimethylsilyl)ethynyl)-1λ^3-benzo[d][1,2]iodaoxol-3(1H)-one
PFTase	farnesyltransferase
pH	potential of hydrogen
Ph	phenyl
pH-EBX	1-((trimethylsilyl)ethynyl)-1λ^3-benzo[d][1,2]iodaoxol-3(1H)-one

Abbreviation	Definition
Phe, F	phenylalanine
PLP	pyridoxal-5-phosphate
PMA	phophomolybdic acid
pNO$_2$-EBX	5-nitro-1-((trimethylsilyl)ethynyl)-1λ^3-benzo[d][1,2]iodaoxol-3(1H)-one
POI	protein of interest
ppm	parts per million
Pro, P	proline
PYP	photoactive yellow protein
RML	*Rhizomucor miehei lipase*
RNase	bovine pancreatic ribonuclease A
RT	room temperature
s	second
SBTI	soybean trypsin inhibitor
SD	standard deviation
SDS-page	sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SEC	size-exclusion chromatography
Ser, S	serine
SPAAC	strain-promoted alkyne-azide cycloaddition
SPIEDAC	strain-promoted inverse electron-demande Diels-Alder cycloaddition
SrtA	Sortase A
t-Bu	*tert*-butyl
T-DM1	Trastuzumab emtansine
TAMM	2-((alkylthio)(aryl)methylene)malononitrile
TAMRA	tetramethylrhodamine
TAPS	[tris(hydroxymethyl)methylamino]propanesulfonic acid
TBAC	tetrabutylammonium chloride
TBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate
TCEP	Tris(2-carboxyethyl)phosphine
TCO	*trans*-cyclooctene
TEAA	triethylammonium acetate
Temp.	temperature
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
THPTA	tris-hydroxypropyltriazolylmethylamine
Thr, T	threonine
TIPS	triisopropylsilyl
TIPS-EBX	TIPS-ethynyl-benziodoxolone
TLC	thin layer chromatography
TM	transition metal
TMPP	tris(2,4,6-trimethoxyphenyl)phosphonium
TMS	trimethylsilyl
TMTH	tetramethylthiacycloheptyne
TNB	5-thio-2-nitrobenzoic
TNBS	trinitrobenzenesulfonate
TOF	time-of-flight
TPPTS	3,3′,3″-Phosphanetriyltris(benzenesulfonic acid) trisodium salt
Tris	tris(hydroxymethyl)aminomethane
Trp, W	tryptophan
TS	transition state
TT	tetanus toxoid
Tyr, Y	tyrosine
U-4C-3CR	Ugi four-component three-centre reaction
U-4CR	Ugi four-component reaction
UAA	unnatural amino acid
UPLC	Ultra Performance Liquid Chromatography
UV	ultraviolet
V	volt
Val, V	valine
VEGFR2	vascular endothelial growth factor receptor 2
W	watt
Zn(II)-DpaTyr	Zn(II)-bis((dipycolyamino)methyl)tyrosine
μ	micro

TABLE OF CONTENTS

INTRODUCTION

1.1. Bioconjugation chemistry
1.1.1. Early developments

Before becoming a topic of interest in chemical biology, the modification of proteins had long been present and used for practical purposes. Initially based on empirical observations and without a clear understanding of the processes taking place, first occurrence of intentional man-made protein modification using a chemical reagent could be dated back to the use of formaldehyde in the tanning industry – where it was used to avoid putrefaction of leather along with other reagents, such as chromium(III) sulfates – and for the production of toxoids.[1] In the latter case, formaldehyde was used for the modification of toxins related to a certain number of bacterial diseases, leading to a non-toxic protein that still retained its original antigenic determinants, and thus helped confer immunity to patients.

Over the 20th century, the improvement of analytical techniques for the characterization of proteins (cation-exchanger amino acid analyzer (50s), ion exchange and gel exclusion chromatography), the discovery of additional amino acids (Met 1922, Thr 1925) and the will for a better understanding of the biological activity of proteins has highly contributed to the development of new techniques for the modification of proteins.[1]

For the characterization of proteins, and in particular, for the identification of their amino acids sequence, several functional group-selective methods were described over the years (**Scheme 1**). Variations of the original Van Slyke procedure, consisting in measuring the development of gaseous dinitrogen from the reaction of primary amines with nitrous acid **R1**, were developed for the quantification of amine groups in proteins.[2] For instance, Fields proposed to use trinitrobenzenesulfonate **R3** (TNBS), leading to trinitroaniline complexes **R4** that were quantified by measuring their absorbance at 420 nm.[3]
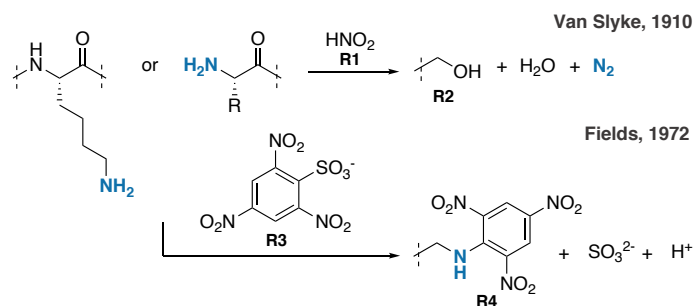
Scheme 1: Van slyke and Fields' methods for the detection of primary amines in proteins.

Similar to sodium nitroprusside that gave a red color in presence of thiols in protein, the Ellman's reagent, 5,5'-dithiobis-(2-nitrobenzoic acid) **R5** (DTNB), allowed to measure thiol concentration, via detection of the highly chromogenic 5-thio-2-nitrobenzoic acid **R7** (TNB) at 412 nm (**Scheme 2**).[4] For the quantification of carboxylic acids in proteins, Hoare and Koshland, in 1967, proposed to modify aspartate and glutamate side chains, and the *C*-terminal residues with a water soluble carbodiimide, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide **R8** (EDC), to form *O*-acylisourea **R9** that then reacted with the nucleophilic glycine methyl ester **R10**.[5] The protein was then hydrolyzed and the amino acid composition determined by an amino acid analyzer. For the determination of tryptophan residues in proteins, Barman and Koshland proposed to selectively label those residues with 2-hydroxyl-5-nitrobenzyl bromide **R12**, forming colored complexes whose absorbance at 410 nm was directly correlated to the number of indoles present in the protein.[6]
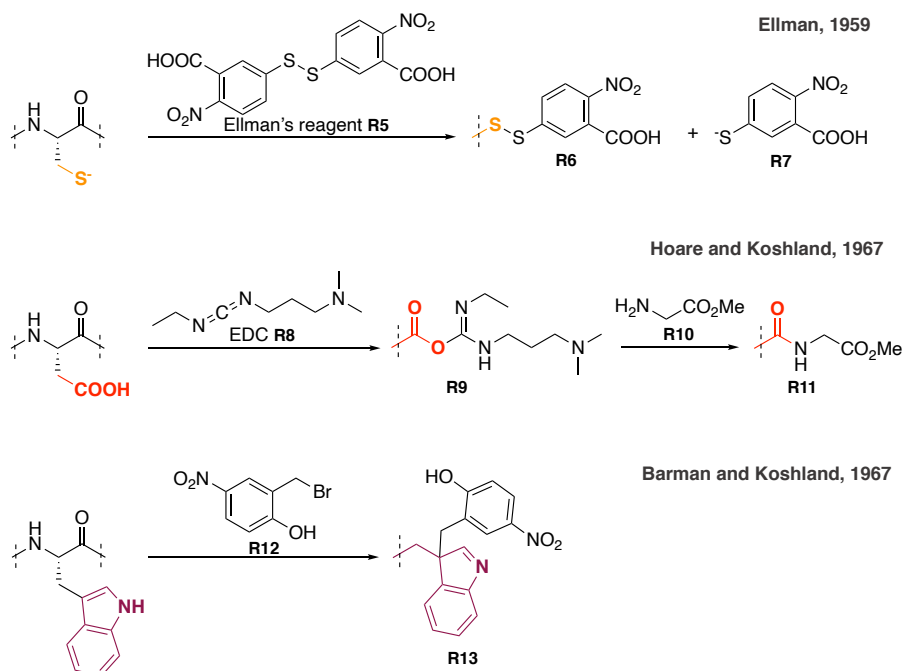
Scheme 2: Ellman, Hoare and Koshland, and, Barman and Koshland's methods for the detection of cysteine, aspartate/glutamate and tryptophan residues in proteins.

With the aim of discovering the amino acids responsible for the biological activity of proteins, the idea emerged that a particular residue in the active site of an enzyme might be identified on the basis of its reaction with a selective chemical reagent.[7,8] Based on this idea, in 1952, Balls and Jansen showed that diisopropyl fluorophosphate could inhibit several proteases by reacting with a serine present in the active site of enzymes, such as α-chymotrypsin, trypsin or cholinesterase.[9]

The innovations previously mentioned facilitated the characterization of modified proteins and led to a better understanding of the reagents' selectivity. With this knowledge and techniques in hand, the development of new conjugation strategies was hence facilitated, and in particular, side chain selective methods. Those labelling reactions, also called chemoselective, were defined as reactions that, under certain specified conditions, led to the modification of a single, or at least, a limited number of side chain groups in a protein. In general, those chemoselective reactions did not lead to complete conversion of all targeted functionalities.[1] While conversion could be increased under more vigorous conditions – e.g. longer reaction time, larger excess of reagent, higher reaction temperature –, this came with the risk of eroding the chemoselectivity and affecting the conformation of the protein, urging researchers to develop new families of reagents.

This explains why the number of chemoselective methods has exploded over the last three decades, in particular for the modification of lysine and cysteine side chain residues. The most classical reactions will be discussed in the following section.

1.1.2. Classical methods for the chemoselective modification of amino acids in biomolecules

Alongside with the improvement of analytical techniques for the characterization of proteins, the development of new reactions for the side chain selective modification of amino acids came. Those methods relied on the modification of the two most nucleophilic amino acids found in proteins, lysine and cysteine residues.

1.1.2.1. Lysine residues

Lysine is one of the most abundant α-amino acids in proteins (> 5% abundance).[10,11] Composed of a linear four-carbon chain terminated with a primary amino group, lysine, along

with arginine and histidine residues, contribute to the overall net positive charge of the protein. In comparison with the α-amino group of the *N*-terminus, the ε-amine of lysine possesses a slightly higher ionization point (pKa of 9.3 – 9.5 for lysine versus pKa of 7.6 – 8.0 for the *N*-terminus), resulting in the formation of positively charged ammonium groups at physiological pH. Due to the dominant ionic character of their side chain, lysine residues are generally exposed to solvent, at the surface of proteins, making them easily accessible for chemical derivatization. Several chemoselective methods were hence described for the modification of lysines in proteins, usually employing either carbonyls, activated esters or iso(thio)cyanate.[10,12–15]

In presence of aldehydes or ketones **R14**, primary amines react to form imines, or Schiff bases, the interaction being more efficient at high pH. Although, the condensation can proceed in aqueous media, the reaction is reversible, making the imine formed unstable. To circumvent this issue, the reductive amination of the imine is usually carried out with sodium cyanoborohydride **R15** (NaBH₃CN), a mild reducing agent that does not reduce aldehydes, to give stable secondary amines **R16** (**Scheme 3**).[16,17]

As an alternative to aldehydes and ketones, primary amines can be modified with activated carboxylic acids (**Scheme 3**). The activation can be performed with a carbodiimide to form *O*-acylisourea that then reacted with primary amines to form very stable amide linkages **R18**. A carbodiimide of choice for the modification of proteins would be the water-soluble EDC **R8**, even though hydrolysis of the resulting activated ester is a major competing reaction. To address this challenge, the *O*-acylisourea can be converted to a more stable succinimidyl ester **R17** (NHS ester).[18] Although the NHS ester **R17** is more stable, the hydrolysis of the reagent is still a major competing reaction.[19] Even though, the reactivity of the primary amine can be increased by increasing the pH of the reaction, the hydrolysis of the reagent is also accelerated. As an alternative to **R17**, acyl azide and fluorophenyl ester reagents were described for the functionalization of primary amine.[10,20–22] However, as for the NHS ester **R17**, those two classes of compounds might suffer from hydrolysis under alkaline conditions.

Primary amines were also found to react with isocyanates and isothiocyanates to form urea and thiourea products respectively (**Scheme 3**).[15,23,24] As isocyanates are very susceptible to hydrolysis, they are not often used in bioconjugation. On the contrary, isothiocyanates **R19** are much more resistant, making them a more common reagent for the labelling of biomolecules. One of the most known isothiocyanates is probably fluorescein isothiocyanate (FITC)

frequently used for the fluorescent labelling of proteins.[25,26] Even though, the rate of hydrolysis of isothiocyanates **R19** was found to be slower than the one of a NHS ester **R17**, the stability of the thiourea linkage **R20** was found to be weaker than the amide's bond **R18**.

Other reagents were also described for the chemoselective labelling of lysines, such as sulfonyl chlorides,[27–29] carbonates[10,15] or imido esters.[30,31] However, they are often highly sensitive to hydrolysis resulting to the limitation of their usage.
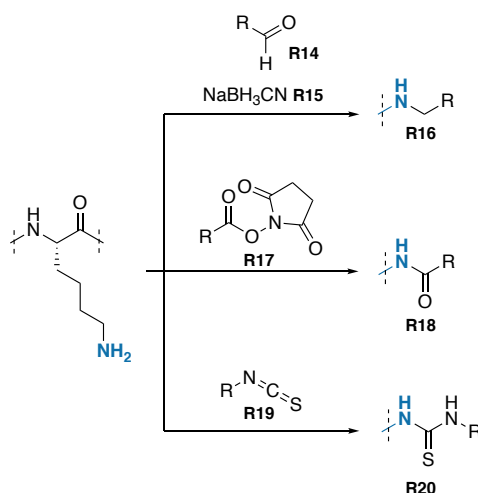
Scheme 3: Classical bioconjugation methods for the chemoselective labelling of lysine residues.

1.1.2.2. Cysteine residues

Because of their low abundance (1.5%) and the high nucleophilicity of their sulfhydryl group, cysteine residues are the most convenient targets in bioconjugation.[10,11,32] The ionization of the cysteine side chains occurs at high pH (pKa 8.8 – 9.1) resulting in a negatively charged thiolate residue. In nature, thiols do not often appear as free functional groups but under the form of disulfide bonds. The reduction of thiol bridges in a target protein can then give access to two free thiols, now available for protein derivatization. Maleimide **R21** and alkyl halide **R23** reagents consisted of the two most widely used reagents for the modification of cysteines and will be further discussed.[10,12–14,33]

In the presence of free thiols, maleimides **R21** undergo a Michael addition to form thioether bonds **R22** (**Scheme 4**).[34–36] Maleimide reagents **R21** were found to be very selective for thiols between pH 6.5 and 7.5, with rates of reaction a thousand times higher than primary amines, but this selectivity tends to diminish at higher pH (above 8.5), with competitive aza-Michael additions being observed.[37,38] Moreover, the formed thioether bond **R22** can be unstable due

to a possible retro-Michael addition and suffer from thiol-exchange. As for the reagents used for the chemoselective labelling of lysine, maleimides **R21** can be prone to hydrolysis, yielding an open maleamic acid that is unreactive towards thiol addition.

Alkyl halide and haloacetamide **R23** reagents were also employed for the labelling of thiols in biomolecules, leading also to thioether bonds **R24** (**Scheme 4**).[10,15] However, it was observed that those active halogen species were also able to react with less nucleophilic amino acids such as histidine, methionine, lysine and *N*-terminal residues.[39] Nevertheless, the selectivity and reactivity of those reagents can be tuned by either selecting the appropriate halogen – order of reactivity: I > Br > Cl > F – or the appropriate reaction conditions; cysteine selectivity has been demonstrated for iodoacetate derivatives used as limiting reagent under slightly alkaline conditions (pH 8.2). In the same perspective, halobenzene-type compounds were also described for the modification of thiols through aromatic nucleophilic substitution, generating aryl thioether species.[10] The reactivity of those reagents can be increased by introducing electron withdrawing groups on the aromatic ring and by choosing the appropriate leaving group – order of reactivity: F > Cl – Br > sulfonate.

Additional methods were described for the chemoselective labelling of cysteines, such as the previously discussed Ellman's reagent **R5** and the entire family of activated disulfide reagents **R25**, which are hydrolytically stable (**Scheme 4**).[10] The main drawback of such strategy comes from the sensitivity of the newly formed disulfide bond **R26** towards reduction, notably by intracellular glutathione.
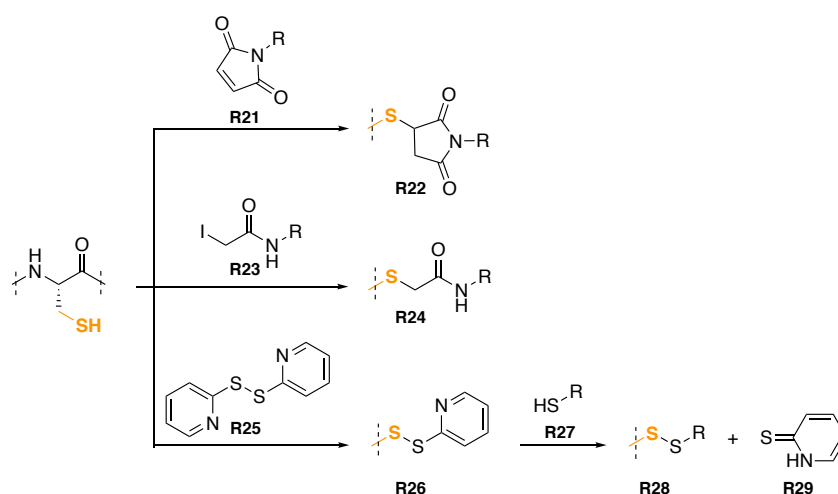
Scheme 4: Classical bioconjugation methods for the chemoselective labelling of cysteine residues.

The development of highly chemoselective, stable towards hydrolysis, and leading to stable conjugates reagents, was thus of high interest.

As classical maleimides **R21** were found to be unstable, due to competitive retro-Michael addition,[40] and suffered from thiol exchange, the development of more stable analogues was required. In 2010, a new generation of maleimide reagents, called dibromaleimides **R30**, was reported by Caddick and coworkers (**Scheme 5**).[41,42] The incorporation of two nucleofuges, either halogen or thiophenol, across the double bond of the maleimide reagent, allowed the consecutive addition of two free thiols with the subsequent elimination of two leaving groups. Nevertheless, the reagents afforded only a reversible modification on the protein, as a thiol exchange could take place with an excess of glutathione or 2-mercaptoethanol. Structurally related, mono- or di-bromopyridazinediones **R32** were described for cysteine conjugation in 2011 by Caddick *et al* (**Scheme 5**).[43] As observed with dibromaleimides **R30**, the conjugates formed were cleaved with an excess of glutathione but found stable in water.
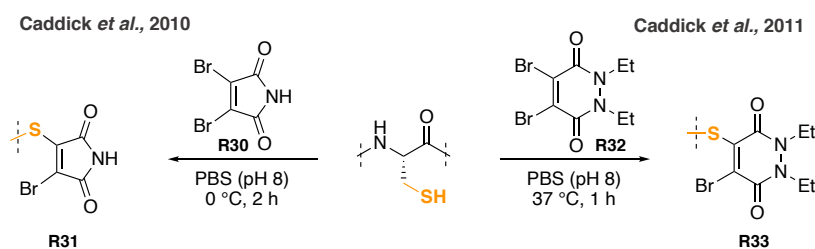
Scheme 5: New generation of maleimides for the chemoselective labelling of cysteine residues.

Kalia *et al.* showed that connecting a phenyl ring to the nitrogen of the maleimide **R34** enabled the formation of conjugates that could undergo a rapid ring hydrolysis leading to a stable payload that showed no sign of thiol exchange (**Scheme 6**).[44] They also demonstrated that exocyclic olefinic maleimides **R36** led to more stable conjugates over time than classical maleimides (**Scheme 6**).[45] Outside of maleimide reagents, carbonyl acrylic derivatives **R38** were identified as suitable Michael acceptors for the modification of cysteines residues (**Scheme 6**). These reagents demonstrated excellent chemoselectivity and afforded payloads in rates comparable to maleimide reagents **16** and possessed high stability.[46] Carbonyl acrylate reagents were later derivatized with either azido or diazo handles for the further functionalization of proteins.[47,48]
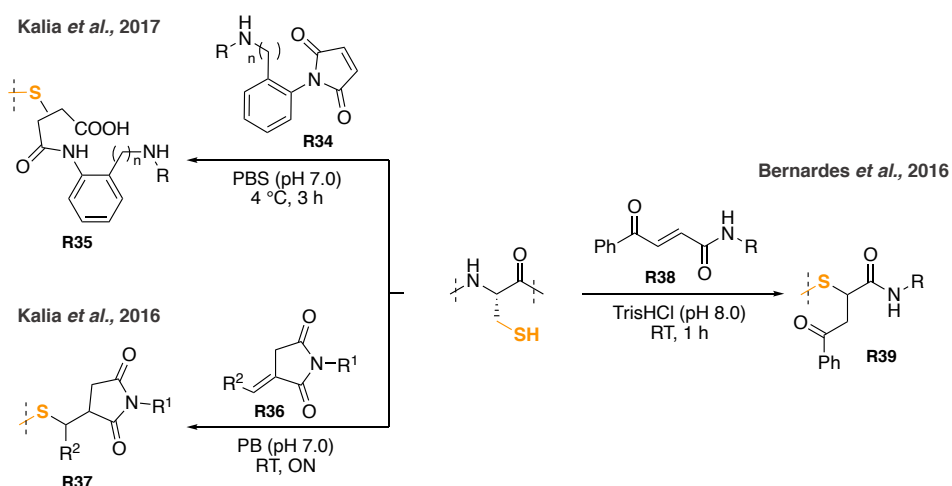
Scheme 6: New generation of Michael acceptors for the chemoselective labelling of cysteine residues.

Vinyl sulfone reagents **R40** were reported as a new class of compounds for the labelling of cysteines (**Scheme 7**).[49,50] These reagents were found to be highly selective for thiols under slightly alkaline conditions, and, led to the formation of β-thiosulfonyl bonds **R41**. As opposed to the previously mentioned reagents, vinyl sulfones **R40** were found to be stable in aqueous solution and not subjected to hydrolysis at neutral pH, thus making them promising reagents for the modification of cysteines in biomolecules.

Based on the thiol-yne and thiol-ene reactions, Hackenberger *et al.* reported two new compounds – an electron-poor ethynylphosphonamidate **R42** and a vinylphosphonothiolate **R44** – for the conjugation of cysteines (**Scheme 7**).[51,52] In both cases, the reagents showed excellent selectivity for cysteines and the conjugates formed demonstrated an excellent stability in buffer, cell lysate and human serum over several days, and showed high stability upon exposure with thiols.
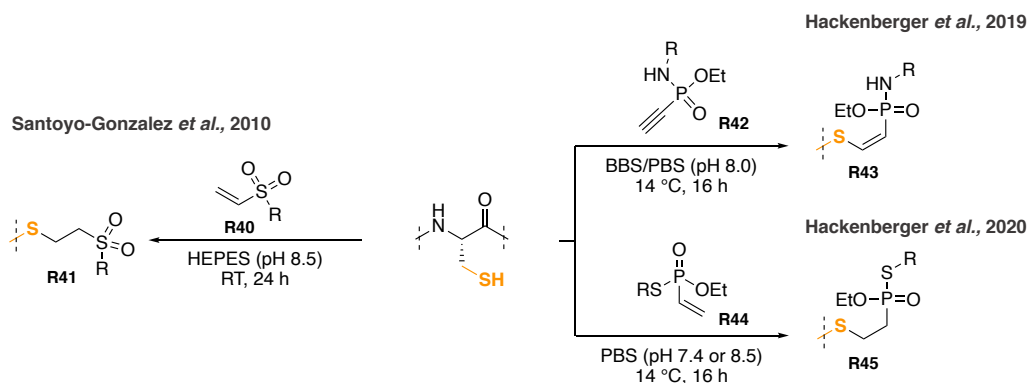
Scheme 7: Chemoselective labelling of cysteine residues with vinyl sulfone **R40**, ethynylphosphonamidate **R42** and vinylphosphonothiolate **R44** reagents.

More recently, Waser and coworkers described two types of hypervalent iodine reagents for the labelling of cysteines (**Scheme 8**). The first one, an azide-bearing EBX reagent **R46**, was described for the introduction of azides into proteins.[53,54] Attack of the free thiol onto the triple bond allowed to form stable vinylbenziodoxolone hypervalent iodine conjugates **R47** that could be sequentially modified with SPAAC and Suzuki-Miyaura cross-coupling reactions. The second hypervalent iodine reagent reported, a TMS-EBX **R48**, allowed the introduction of an acetylene group directly on the free thiol **R49**, which was then modified with the CuAAC reaction.[55] The method was proved to be efficient within few minutes on peptides and antibodies.
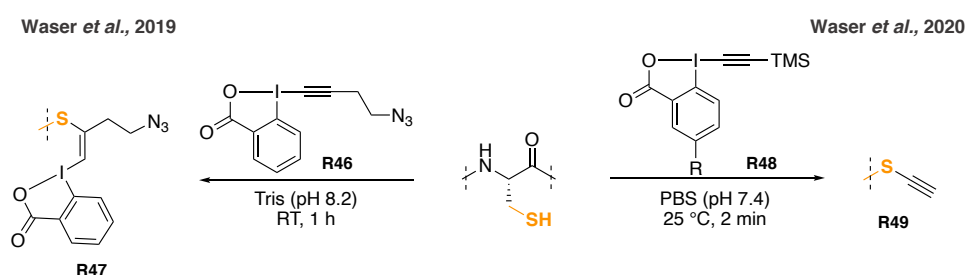
Scheme 8: Use of hypervalent iodine reagents for the chemoselective labelling of cysteine residues.

The number of methods for the chemoselective labelling of cysteines is not limited to the previously mentioned reagents. Additional reagents were described over the years, especially for the re-bridging of disulfide bonds, and will be described in more details in the second part of this introduction.

The development of chemoselective bioconjugation reactions opened the road to the generation of functionalized proteins. The incorporation of spectroscopic labels helped in the understanding of cellular mechanisms by analyzing the trafficking of labelled proteins;[56] polyethylene glycol chains functionalization yielded in less-immunogenic and more plasma-stable proteins;[57,58] and conjugation of a drug to a carrier protein helped in reducing its systemic toxicity and target its delivery.[59,60]

In the latter case, and in particular for antibody-drug conjugates (ADCs), the bioconjugation method was found to be one of the determinant parameters – along with the carrier protein, linker and cytotoxic drug – in the efficacy of the conjugate.[60,61] Indeed, these chemoselective bioconjugation methods, applied to large proteins, such as antibodies (150 kDa) possessing 84 reactive lysines and 8 accessible cysteines, logically led to the generation of highly heterogeneous mixtures of antibody-drug conjugates, all possessing different stabilities, pharmacokinetic and pharmacodynamic properties. Keeping the example of an antibody

possessing 88 reactive amine groups (84 lysines and 4 *N*-termini), Gautier and coworkers showed that up to 69 amine functionalities could be conjugated with a NHS-ester reagent.[62] With a single modification found on the antibody (degree of conjugation of 1), 69 different regioisomers would then be expected. Increasing the degree of conjugation would then enhance the heterogeneity of the conjugate's mixture, where more than 10 million of different conjugates would be expected with an average degree of conjugation above 2 obtained.

Therefore, the next logical step implied to move towards regioselective methods, now that chemoselectivity issues had been answered, in order to solve the heterogeneity issues. This request is then of prime importance as new therapeutics employing protein conjugates require a better control of the conjugation step. This is particularly the case in the field of antibody-drug conjugates where a better control in the conjugation site would then widen the therapeutic window and potentially improve the efficacy of the treatment against cancer.[60,61]

Those regioselective methods imply, most of the time, the generation of recombinant proteins, in which specific amino acids or sequences that are naturally absent from the protein, are introduced allowing a site-selective functionalization. Despite clear advantages, such strategies, often require many steps in addition to the synthesis of the recombinant proteins, and as such, can be tedious, time-consuming and costly, which currently prevents their industrial application.[63] A simpler and more efficient site-specific method would be to use native proteins, which can be easily produced on large scale and give access to a wider array of controlled DoC values. The two different strategies will then be discussed in the next sub-chapter.

1.2. Site-selective approaches for the modification of proteins
1.2.1. Single-site labelling of pre-engineered proteins

Genetic engineering is currently the method of choice to get homogeneous conjugates. The introduction of canonical or non-canonical amino acids within the structure of the protein of interest allowed a very precise labelling.

1.2.1.1. Natural amino acids

The first strategy explored consisted in the incorporation of poorly abundant natural amino acids at strategic positions on the protein of interest. To make the strategy efficient, the mutation introduced must then be located in an environment where a further modification is possible and must retain the structure and function of the protein. To avoid structural perturbation, the genetic manipulation is usually taking place at the *N*- or *C*-terminal positions, or, on a loop.[64]

Because of its high nucleophilicity and poor abundance in proteins, the engineering of free cysteines or selenocysteines was extensively employed.[64–66] As cysteines are usually found as disulfide bonds in proteins, the incorporation of a free thiol allowed the orthogonal and selective modification on the engineered protein. One of the most famous examples is the Thiomab®, where two or four cysteines were engineered on the antibody trastuzumab.

Alternative low abundant amino acids were engineered on proteins, such as tryptophan, tyrosine, methionine and histidine. However, because of their low reactivities, they have found less interest than cysteine.[66]

1.2.1.2. Unnatural amino acids

The incorporation of a variety of unnatural amino acids (UAA) provided unique chemical handles for the site-selective modification of proteins with bioorthogonal reactions.[63,64,66–68] Azide, alkyne, alkene, tetrazine and ketone amino acids were designed and incorporated into proteins with the amber codon suppression technique and then modified with appropriate reagents.

The Staudinger ligation between an azide and a triarylphosphine applied by Bertozzi *et al.* for the modification of azido-glycoproteins expressed on the surface of living cells via the sialic

acid biosynthetic and metabolic pathways was the pioneering work for the incorporation of UAA in proteins.[69] This technique and its variations ("traceless" Staudinger)[70] were then extended to the direct modification of azido-containing proteins that had incorporated azidoalanine **R50**,[71] azidohomoalanine **R51**[71] or *p*-azidophenylalanine **R52**[72,73] amino acids (**Figure 1**). Unfortunately, due to the slow kinetics associated with the Staudinger ligation, the retained triarylphosphine oxide appendage in the "non-traceless" variant, the problems related to phosphine oxidation and the possible side reactions of phosphine in the "traceless" variation, the use of the Staudinger ligation diminished over the years to profit of the Huisgen 1,3-dipolar cycloaddition.[74,75]

The 1,3-dipolar cycloaddition allows the formation of a triazole **R58** from an azide and a terminal alkyne **R57** and is accelerated in presence of copper(I) (**Figure 1**).[76,77] As the reaction was found to be orthogonal and highly efficient under physiological conditions, it was used for the modification of proteins. Also known as the copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC), this reaction was not only used for the modification of azido UAA[78–82] but also for the functionalization of alkyne UAA such as homopropargylglycine **R61** (Hpg),[78,83] *p*-(propargyloxy)phenylalanine **R62**[79] and *N*$^\varepsilon$-(propargyloxycarbonyl)-L-lysine **R63** (AlkK).[84] Despite an apparent cellular toxicity of copper, caused by the generation of reactive oxygen species, the CuAAC remains an invaluable tool for the modification of proteins due to its high specificity, fast kinetics and ease-of-use.[85]

As an alternative to the use of copper, Bertozzi and coworkers adapted the strain-promoted alkyne-azide cycloaddition (SPAAC), developed by Wittig and Krebs in 1961, to bioconjugation in 2004.[86,87] They found that a biotinylated cyclooctyne **R64** could react rapidly at room temperature with azido-glycoproteins without any catalyst and ligand. Because the kinetics of the SPAAC were initially found to be similar to the Staudinger ligation,[74] a number of strained alkynes have been developed for accelerating the rate of the reaction, such as difluorinated cyclooctyne **R65** (DIFO),[88–90] dibenzocyclooctynol **R66** (DIBO),[91] dibenzoazacyclooctyne **R67** (DIBAC),[92] biarylazacyclooctynone **R68** (BARAC),[93] tetramethylthiacycloheptyne **R69** (TMTH), and bicyclo[6.1.0]nonyne **R70** (BCN) (**Figure 1**).[94] Even if some of them were directly incorporated into proteins, their synthesis was complicated and a lack of stability was observed for those UAA.[95,96]
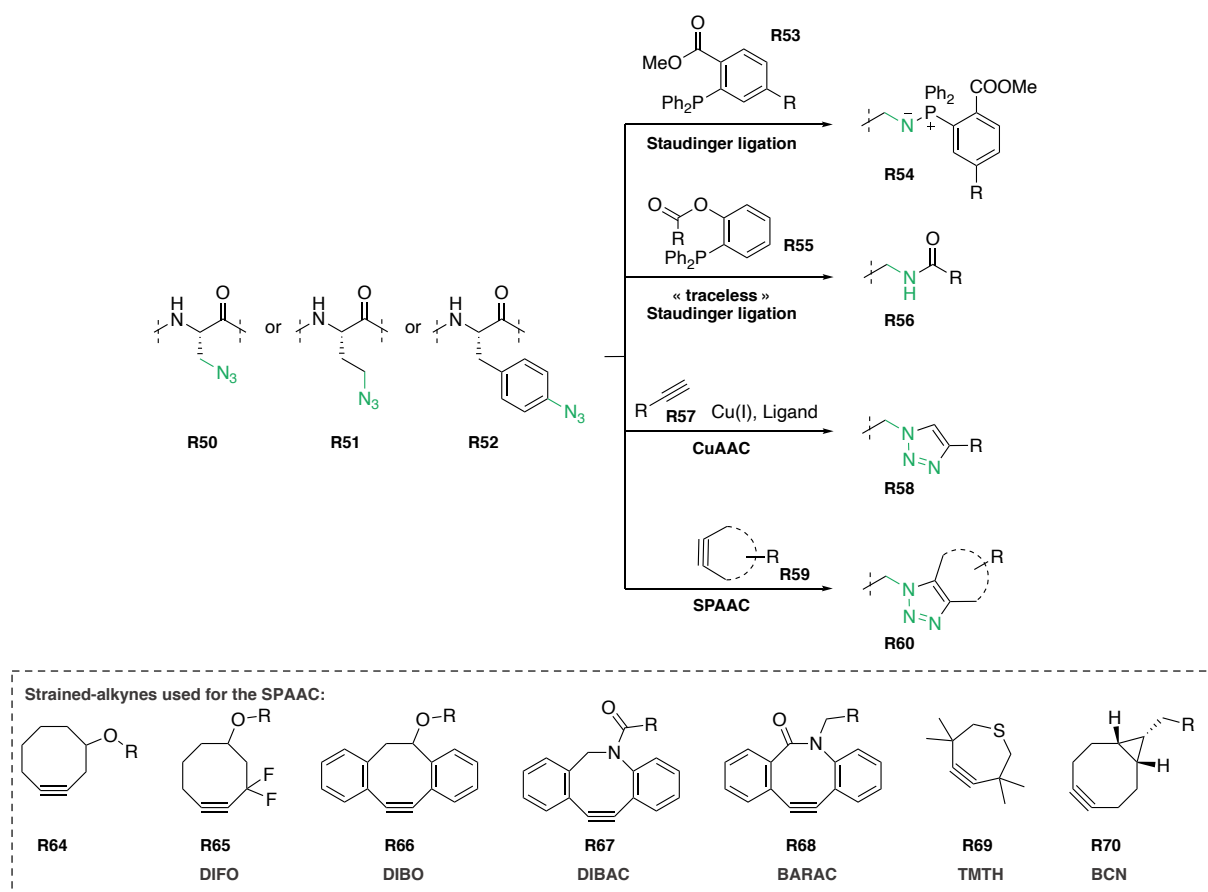
Figure 1: Incorporation of azido amino acids via genetic manipulation in proteins. The introduced azido group can then be modified with the Staudinger ligation, the "traceless" Staudinger ligation, the copper-catalyzed alkyne-azide cycloaddition (CuAAC) and the strain-promoted alkyne-azide cycloaddition (SPAAC).

The inverse-electron demand Diels-Alder (iEDDA) reaction emerged with the development of the SPAAC reaction as a new promising tool for the labelling of proteins (**Figure 2**). This reaction required reactive dienes, such as *trans*-cyclooctene **R71** (TCO) or norbornene **R72**, that reacted with tetrazine dienophiles **R73** with rates thousand times faster than what observed with SPAAC.[97–99] New reactive UAA containing tetrazine **R74**,[100] norbornene **R75**,[101–103] cyclooctene **R76**[96,103] or biscyclonene **R77**,[96] were thus developed and incorporated in proteins. Despite the fast kinetics of the reaction, some limitations remained, as the possible isomerization of the *trans*-cyclooctene or the possible instability of the tetrazine.[96,104]

As an alternative, the [3+2]-cycloaddition between a nitrile-imine **R78**, generated *in situ* by photo-irradiation of a tetrazole **R79**, and an unactivated alkene **R80/R81** was reported (**Figure 2**).[105] This new reaction led to the development of new alkenyl UAAs, *O*-allyl-tyrosine **R82**[106,107] and *N*ᵋ-(1-methylcycloprop-2-enecarboxamido)lysine **R83** (CpK),[107,108] and tetrazole UAA, *p*-(2-tetrazole)phenylalanine **R84** (*p*-Tpa).[109] While the rate of the [3+2]-cycloaddition was found to be comparable to that of the iEDDA reaction between norbornene and tetrazine, this reaction

can be spatially and temporally controlled, as UV irradiation is required to form the nitrile-imine reactive species.
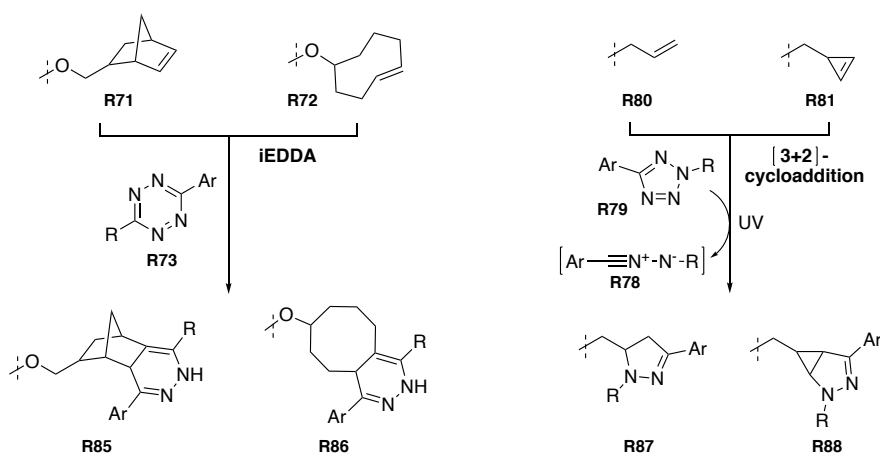
Figure 2: Modification of norbornene **R71** and *trans*-cyclooctene **R72** amino acid derivatives with the inverse-electron demand Diels-Alder cycloaddition (iEDDA). Modification of unactivated alkene **R80** and cyclopropene **R81** amino acid derivatives with the 1,3-dipolar cycloaddition.

As aldehydes and ketones are almost entirely absent from native proteins, their incorporation was found to be attractive for protein modification. The modification of aldehydes and ketones has found a particular use over the years as they can easily be modified with hydroxylamine and hydrazine reagents to form oximes and hydrazones, respectively.[110] Generation of a terminal aldehyde could be easily achieved on *N*-terminal serine or threonine by periodate oxidation.[111] As an alternative, it was also possible to incorporate aldehyde and ketone-containing amino acids,[112,113] such as *p*- and *m*-acetyl-L-phenylalanine (**R89** and **R90**),[114–120] *p*-benzoyl-L-phenylalanine **R91**,[116,118] N^ε-acetyl-lysine **R92** (AcK)[121] and 2-amino-8-oxononanoic acid **R93** (KetoK).[121] Nevertheless, some limitations remained with the hydrazone and the oxime linkages, their potential hydrolysis under acidic conditions limited their applications *in vivo*.

Transition metal catalysis emerged as an alternative strategy for the modification of UAAs on recombinant proteins (**Scheme 9**).[68,122] Associated with excellent chemoselectivity and high yields with reactive handles often inert in biological systems, palladium-catalyzed cross couplings between aryl/alkenyl halides **R94** and a variety of reactive partners such as boronic acids **R95** (Suzuki-Miyaura),[123–131] alkenes **R96** (Heck)[132–134] and alkynes **R97** (Sonogashira)[132,134,135] have been described. Nevertheless, the use of these reactions was restricted as they needed to proceed efficiently at low protein loading and in aqueous media, with non-specific binding of the catalyst on the protein nucleophilic residues. The Heck,

Sonogashira and Suzuki-Miyauri reactions were then associated to low conversions – 2%, 25% and 30%, respectively.[124,133,134] For the Suzuki-Miyaura reaction, the development of a water-and-air stable ligand (2-amino-4,6-dihydroxy-pyrimidine, ADHP) allowed to reach almost complete conversion (> 95%) under aqueous conditions (pH 8) and at 37 °C, and, the identification of a Pd scavenger allowed to the remove non-specific binding of palladium on the protein.[125,126] Modification of proteins was also found to be possible by olefin cross metathesis reaction.[136–140] Employing the Hoveyda-Grubbs II catalyst, incorporated allyl sulfides were successfully modified with olefins, with rates comparable to that of CuAAC reaction.
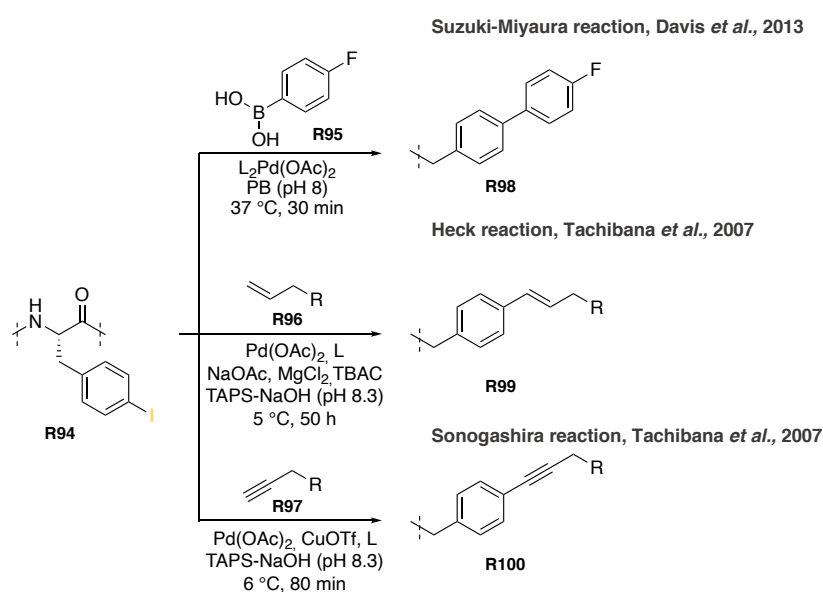
Scheme 9: Modification of the unnatural amino acid, 4-iodophenylalanine **R94**, with the Suzuki-Miyaura, the Heck or the Sonogashira reactions.

The different unnatural amino acids incorporated into proteins are summarized in the following figure and their functionalization are highlighted.
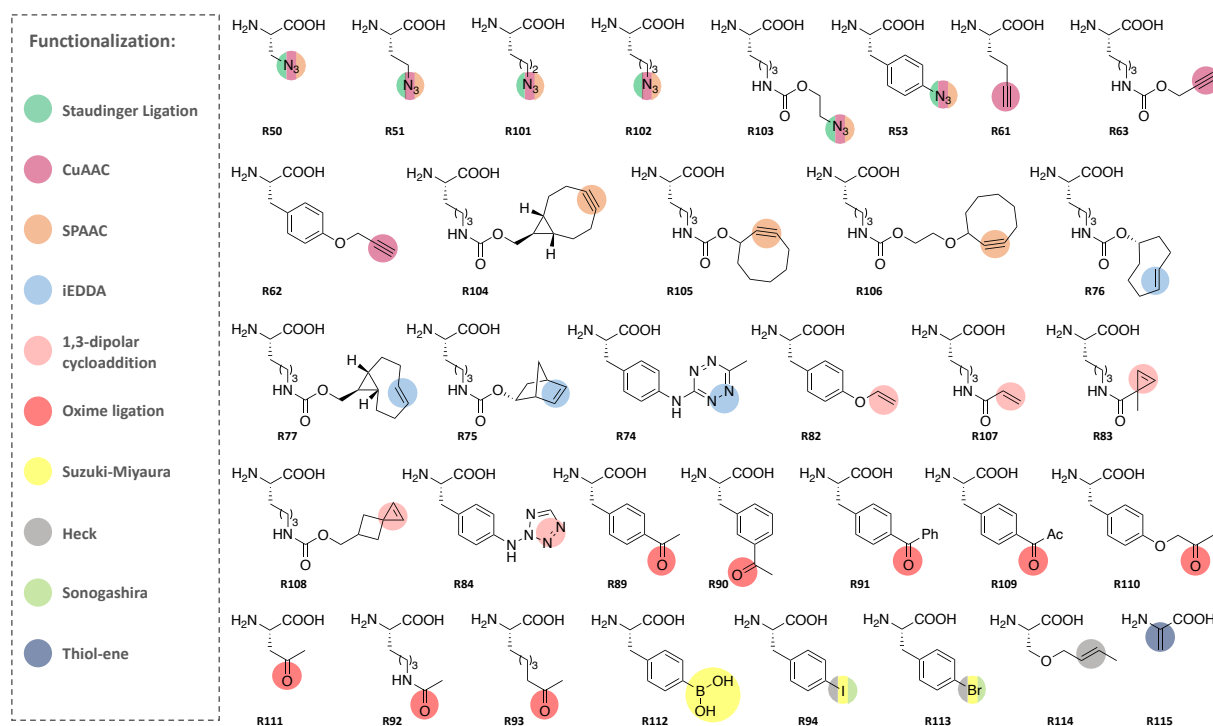
Figure 3: List of unnatural amino acids.

1.2.1.3. Motif insertion and enzymatic recognition sequence

Instead of modifying a particular natural or unnatural amino acid on the protein of interest (POI), it was thought to install either enzymes or proteins, or, a set of amino acids at the *N*- or *C*-terminal positions of the protein during its expression. The incorporation of enzymes or proteins allowed the proximity-driven modification of the POI by molecular recognition (**Figure 4**).[141–144] Different research groups relied on this strategy and fused their POI to a tag that can be covalently labelled with a small molecule. For example, the SNAP-tag, derived from DNA repair protein O^6-alkylguanine-DNA alkyltransferase (AGT), was covalently labelled with O^6-benzylguanine derivatives **R116** by undergoing an irreversible reaction in which the functionalized benzyl group of **R116** was transferred to an active cysteine.[145,146] Based on the SNAP tag, the CLIP tag was developed and functionalized with the benzyl group of benzylcytosine derivatives **R117** on the thiolate.[146] Based on a mutant of *Rhodococcus* dehalogenase (DhaA), the Halo tag, fused to a POI, formed a covalent ester bond between Asp106 – located in the enzyme – and a chloroalkane substrate **R118**.[147] The photoactive yellow protein (PYP), once fused to POI, was selectively functionalized on Cys69 by *trans*-thioesterification with a thioester-derivative of 7-hydroxycoumarin **R119**.[148] Although this strategy was found to be very powerful (fast labelling, flexibility of labelled functionalities, real-

time analysis of fast biological processes), the molecular weight of the proteins/enzymes appeared to be too large and, in some cases, altered the properties of the POI (structure, function, dynamics, localization).[63]
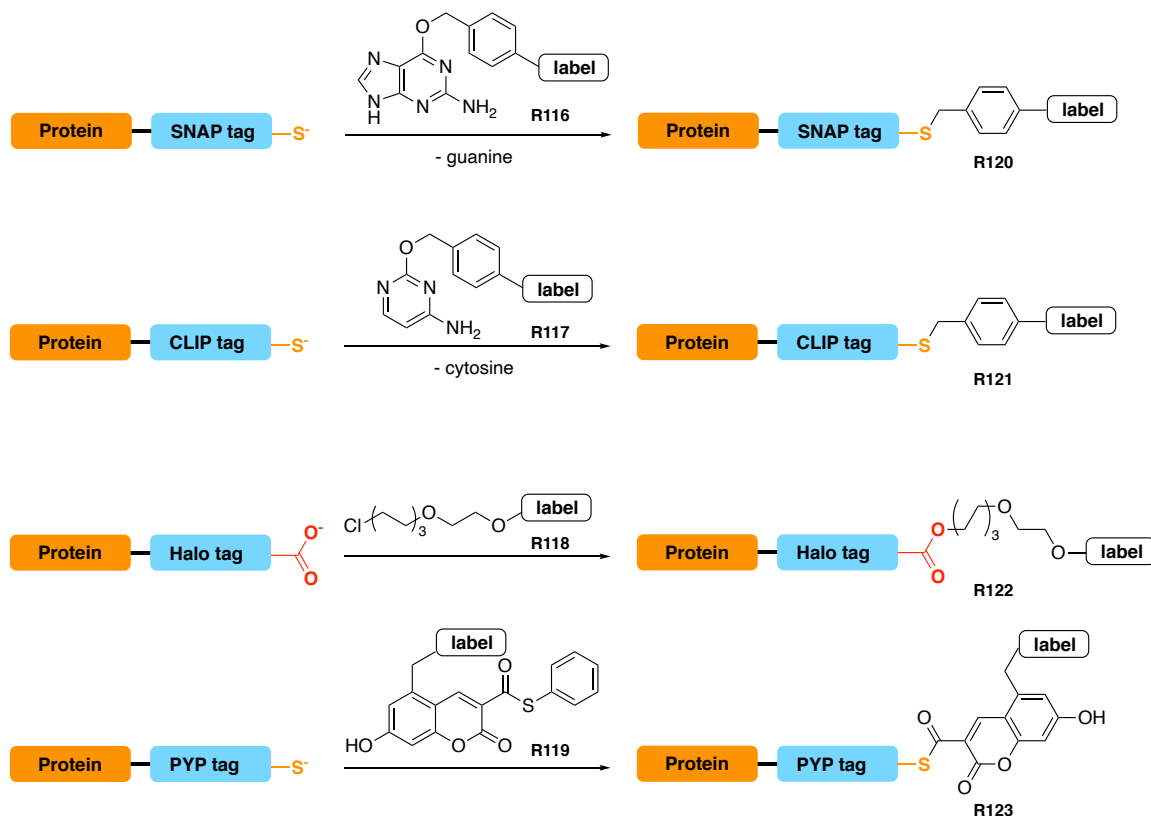
Figure 4: Installation of protein tags on the terminal positions of a protein of interest.

To bypass this limitation, smaller peptides (0.6-6 kDa) were designed and installed on the terminal positions of the POI (**Figure 5**). It was thought that their smaller size would limit their impact on the proteins' structure and function.[63] In a seminal work, Tsien *et al.* reported a tetracysteine motif **R124** (CCXXCC) able to bind bis-arsenic fluorophores **R125** with high selectivity and affinity.[149,150] Based on a similar design, a tetraserine motif (SSPGSS) was reported and described to bind bis-boronic dyes specifically.[151] Later, Hamachi and coworkers developed and improved the tetra-aspartate motif (DDDD or D4 tag) to selectively bind Zn(II)-bis((dipycolyamino)methyl)tyrosine (Zn(II)-DpaTyr).[152] In a similar approach, the interaction between a His tag (10 histidines) with Ni(II)-NTA facilitated the nucleophilic addition of one of the His tag imidazole to the tosylated Ni(II)-NTA probe.[153,154] More recently, Pentelute *et al.* reported the shortest reactive tag sequence to date, a tetrapeptide FCPF **R127** functioning as a π-clamp and allowing the tag cysteine to be specifically labelled with perfluoroaromatic reagents **R128**.[155]
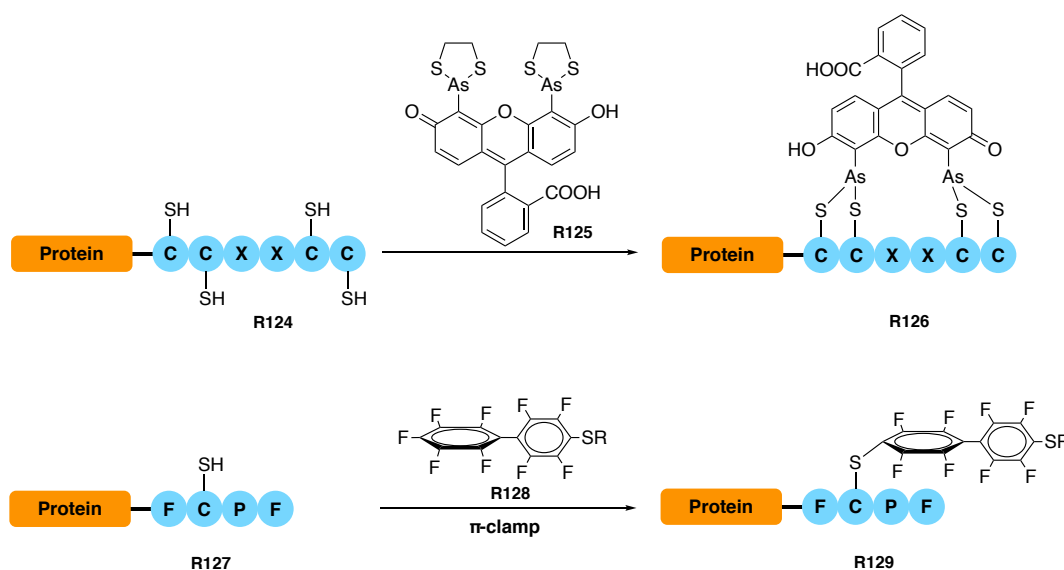
Figure 5: Installation of small peptides on the *C*-terminal position of a protein of interest. Tsien *et al.* modified the tetracysteine motif R124 (CCXXCC) with a bis-arsenic fluorophore R125 with high selectivity. The FCPF R127 motif (π-clamp), reported by Pentelute *et al.,* selectively reacted with perfluoroaromatic reagents R128. Adapted from the review: Rawale, D. G.; Thakur, K.; Adusumalli, S. R.; Rai, V. Chemical Methods for Selective Labelling of Proteins. *Eur. J. Org. Chem.* **2019**, *2019 (40)*, 6749–6763.

The utilization of enzymes for the modification of a defined peptide sequence was also reported (**Figure 6**).[156,157] As an example, in a CXPXR motif, formylglycine generating enzyme (FGE) was able to convert the cysteine residue into a formylglycine, which was then selectively functionalized with the oxime ligation.[158] Sortase A was found to cleave the amine bond between a threonine and glycine in the LPXTG motif, leading to reactive thioesters that then gave access to protein-peptide or protein-protein structures.[159,160] Transfer of a farnesyl group from farnesyl pyrophosphate (FPP) to cysteine on the CaaX motif (a = aliphatic amino acid) was described using the protein farnesyltransferase (PFTase) protein.[161]
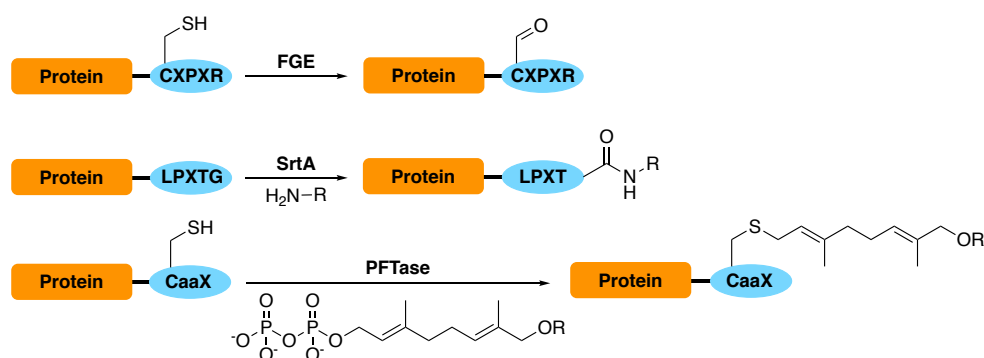
Figure 6: Labelling of engineered proteins with formylglycine generating enzyme (FGE), Sortase A and farnersyltransferase (PFTase). Adapted from the review: Rawale, D. G.; Thakur, K.; Adusumalli, S. R.; Rai, V. Chemical Methods for Selective Labelling of Proteins. *Eur. J. Org. Chem.* **2019**, *2019 (40)*, 6749–6763.

Prominent results have been obtained with recombinant proteins, for which site-specific conjugation was made possible by the incorporation of natural amino acids (NAA), unnatural amino acids (UAA), proteins/enzymes or peptide sequences. Unfortunately, the generation of those proteins can be a costly and time-consuming process, whilst the enzymatic functionalization step can be low yielding and have a limited scope.[63] Coupled to limited DoC values (1 to 2) and potential immunogenicity issues due to the artificial incorporation of amino acids or sequences, these limitations prevented the widespread use of this technique.[63] The development of site-selective methods on native proteins was then of interest.

1.2.2.   Single-site labelling of native proteins

As an alternative to recombinant proteins, over the past few years, the development of new regioselective strategies for the conjugation of native proteins has emerged, targeting various types of amino acids, such as cysteine, lysine, tryptophan, tyrosine, or histidine.

1.2.2.1.   Lysine residues

As one of the most abundant α-amino acid and due to the dominant ionic character of its side chain, lysines are generally exposed on the surface of proteins and can be easily derivatized.[10,11] Because of these two aspects, finding a site selective method for the modification of lysine has been challenging and only few approaches have been proposed relying on kinetic control of the conjugation reaction, on multicomponent reactions, or on template-directed approaches.

One of the most common strategy for the chemoselective modification of lysines is based on the use of activated esters. In 2012, Weil *et al.* showed that a regioselective variation of this approach was possible, with the successful modification of RNase A and lysozyme via a kinetically controlled labelling (**Scheme 10**).[162] Partial addition of only 0.5 equivalents of a NHS-biotin derivative **R130** over 100 minutes gave a single site of modification on the protein of interest, but with the expected detriment of reaching only partial conversion. Nevertheless, biotin incorporation helped at the separation of conjugates from the remaining native protein by affinity purification on an avidin column.

Moving on to bigger constructs, Geierstanger *et al.* developed conditions to favor the single labelling of Lys188 on the Fab fragment of trastuzumab (50 kDa) with a fluorophenyl ester **R132** (**Scheme 10**).[22] The authors attributed the unique reactivity of Lys188 to its local

environment. More specifically, they demonstrated that two neighboring amino acids, His189 and Asp151, were essential for the selective labelling of this lysine. It was found that the aspartate could activate the histidine to act as an acid-base catalyst that could accelerate the labelling of the lysine. It was also proposed that the acylation might occur first on the histidine with the payload being then transferred to the lysine via an acyl transfer. This method was unfortunately not extended to the modification of the full, intact antibody.
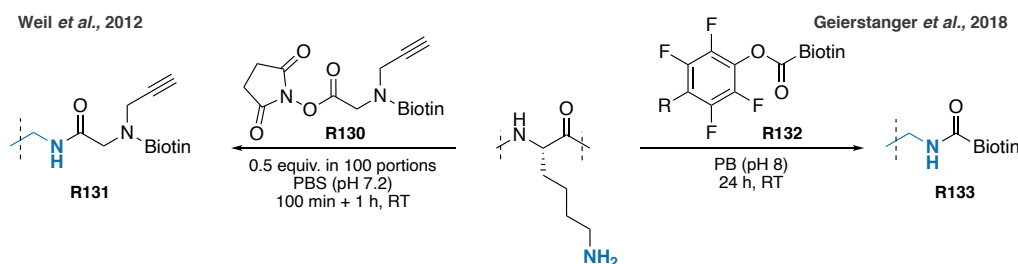
Scheme 10: Labelling of lysine residues with a NHS-ester or a fluorophenyl ester.

Barbas III *et al.* revealed that a single lysine modification on HSA was possible when employing a single equivalent of an α,β-unsaturated sulfonamide derivative **R134** (TAK242), also suggesting that certain lysines were more reactive than others (**Scheme 11**).[163] This was further supported by Rai *et al.* who also showed that certain lysine residues were more reactive than others under specific conditions on various proteins. For example, site-selective formylation of lysine with formate ester – generated from the auto-oxydation of an aldehyde **R136** – was achieved on RNase A, ubiquitin, lysozyme c, myoglobin, cytochrome c and histone with moderate yields (50-82%) (**Scheme 11**).[164] In 2018, Bernardes *et al.* employed a sulfonyl reagent **R138** and relied on kinetic control for the site-selective modification of reactive lysine (**Scheme 11**).[165] While regioselectivity was observed on small proteins, demonstrated by MS / MS analyses, no information was given about the selectivity on bigger proteins such as antibodies.
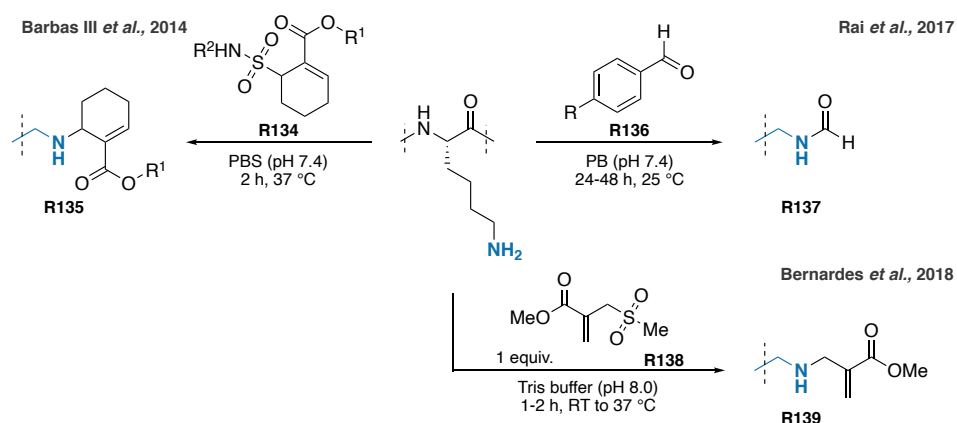
Scheme 11: Labelling of lysine residues with an *α,β*-unsaturated sulfonamide derivative **R134**, a formate ester generated from the auto-oxidation of an aldehyde **R136** or a sulfonyl acrylate **R138**.

Instead of using a single electrophile for the modification of lysine residues, it was suggested by Rai and coworkers that a gain in chemo- and regioselectivity could be achieved when multicomponent reactions (MCR) were employed. Two different three-component reactions were then evaluated. The first one allowed the introduction of an alkyne on lysine side chain by a Cu(II)-catalyzed A3 coupling between the amine group, formaldehyde **R140** and a terminal alkyne **R141** (**Scheme 12**).[166] For all proteins tested, a single site of modification was identified despite the excess of reagents used (100 equivalents of aldehyde, alkyne and metal). The phospho-Mannich reaction, conducted in the presence of an aldehyde **R143** and triethylphosphite **R144**, was the second MCR proposed for the regioselective labelling of proteins (**Scheme 12**).[167] Even though a single site of modification was identified on all proteins tested, conversions still remained low (< 40%), and the reaction could not be transposed to bigger proteins such as antibodies.
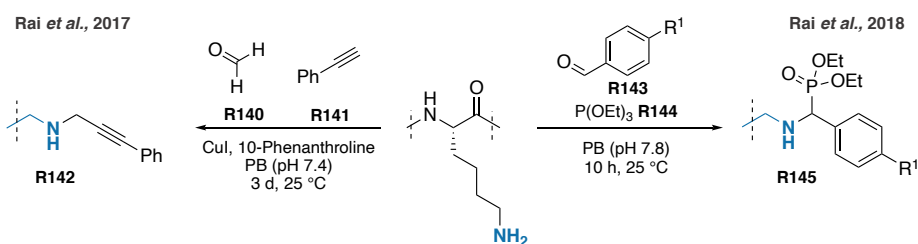
Scheme 12: Use of multicomponent reactions for the regioselective labelling of lysine residues.

Another strategy for the regioselective conjugation of native proteins employs template-directed approaches, using another amino acid to direct the modification of a neighboring lysine for example. Based on the native chemical ligation (NCL), Chudasama and coworkers showed that cysteine can act as a relay for the site-selective modification of lysines on the Fab

fragment of trastuzumab (**Figure 7**).[168] After reduction of the disulfide bond of the Fab fragment, a first *trans*-thioesterification with a reactive thioester **R146** occurred, followed by a "cysteine-to-lysine transfer" (or *S* to *N* acyl shift) with a proximal lysine – three sites of modification, all close to the disulfide bond, were identified. Following the formation of a stable amide bond **R147**, reformation of the disulfide bonds and payloads derivatization delivered functional Fab fragments of trastuzumab. Unfortunately, the method was not extended to the modification of the full, intact antibody.

More recently, Rai *et al.* described a lysine-directed lysine modification method (**Figure 7**).[169] The reagent they designed bore two reactive groups, an aldehyde and a morpholine ester, separated by a spacer **R148**. The aldehyde first led to the formation of imines with multiple solvent-accessible lysines, helping to regulate the second – intramolecular – lysine modification by the morpholine ester electrophile. Imines hydrolysis in the final step allowed the isolation of mono-labelled conjugate. Regioselectivity was achieved on different proteins by tuning the size of the spacer.
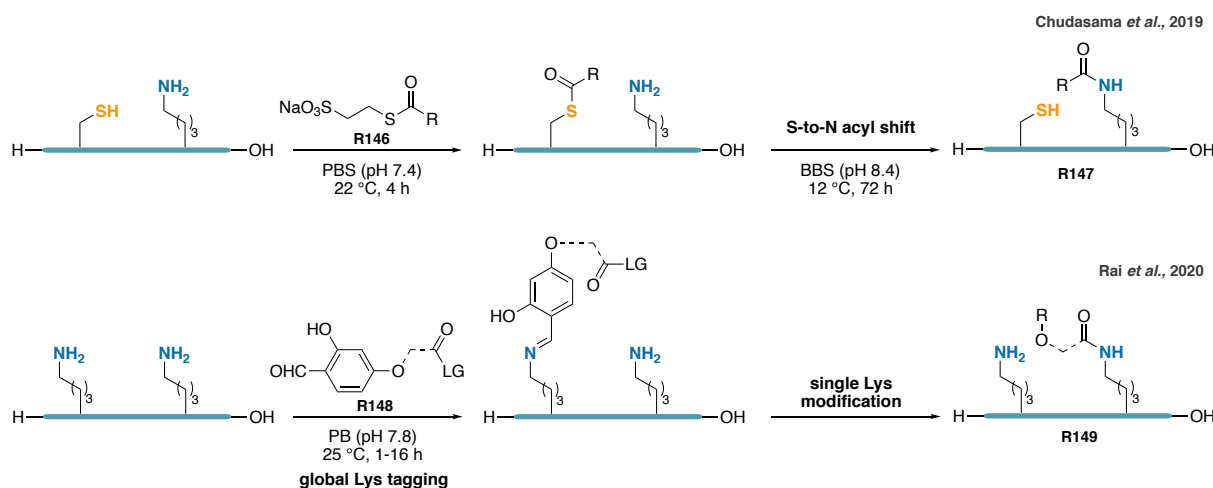
Figure 7: Amino acid-directed lysine modification strategy.

Protein-peptide interaction strategies were also reported as a successful approach for the site-selective modification of lysine.[170] Installation of an electrophilic group – isothiocyanate, fluorodinitrobenzene (FDNB) and phenyl ester – on a targeting peptide allowed the conjugation of the ε-NH$_2$ of the sole lysine residue present in the vicinity of the peptide binding site of the protein. The method was successfully applied to the single modification of three different proteins, ATG8, LC3 and Csk SH3. While this approach was considered as regioselective since the only lysine present in the binding site was modified, the authors still wondered whether the approach would be selective in the presence of multiple lysines.

The reagents reported for the site-selective modification of lysine residues are presented in the following **Table 1**.

Table 1: List of the reagents used for the site-selective modification lysine residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R130[162]	*(structure)*	*(structure)*	8**	100 min / RT	40%	CuAAC	9 steps	Yes
R132[22]	*(structure)*	*(structure)*	5	24 h / RT	-***	No	1 step	Yes on antibody fragment
R134[163]	*(structure)*	*(structure)*	1 or 2	2 h / 37 °C	-	CuAAC	3 steps	Yes on HSA
R136[164]	*(structure)*	*(structure)*	600	24-48 h / 25 °C	50-82%	No	1 step	yes
R138[165]	*(structure)*	*(structure)*	1	1-2 h / 37 °C	100%	Aza-Michael addition	1 step	unsure
R140 / R141[166]	*(structure)*	*(structure)*	100	72 h / 25 °C	40-99%	No	Commercially available	yes
R143 / R144[167]	*(structure)*	*(structure)*	300	30 min - 10 h / 25 °C	32-41%	Oxime ligation	1 step	no
R146[168]	*(structure)*	*(structure)*	100	4+72 h / 22 to 12 °C	-	CuAAC	1 step	Yes on antibody fragment
R148[169]	*(structure)*	*(structure)*	25	1-16 h / 25 °C	26-86%	Oxime ligation	3 steps	yes

* lysine residue in blue, functionalization site in red

** 8 equivalents were used for the NHS-ethynyl-biotin; for the NHS-biotin, only 0.5 equivalents were used.

*** information was not provided by the authors

1.2.2.2.   Cysteine and cystine residues

In Nature, free accessible thiols on cysteine residues are quite scarce, being usually present as disulfide groups; such as the cystine residues bridging heavy and light chains of IgG1 antibodies. While the reduction of the inter-chain disulfide bonds of IgG1 can give access to eight free cysteines, it can also lead to the loss of protein structure, shape, stability and activity.[171] Thus, we will focus mostly on conjugation methods enabling the reformation of those bonds, or of analogues thereof.

Brocchini and coworkers were the first ones to reconstruct intramolecular disulfide bonds in proteins – with two consecutive Michael additions – by forming a three-carbon bond between two sulfur atoms with a monosulfone reagent **R150** generated *in situ.*[172,173] This strategy was applied to the PEGylation of the interferon IFn α-2b. Bissulfone or monosulfone reagents, conjugated to a cytotoxic drug, MMAE, or to the bioorthogonal tag, tetrazine, were further applied as bridging reagents for trastuzumab (**Scheme 13**).[174,175]
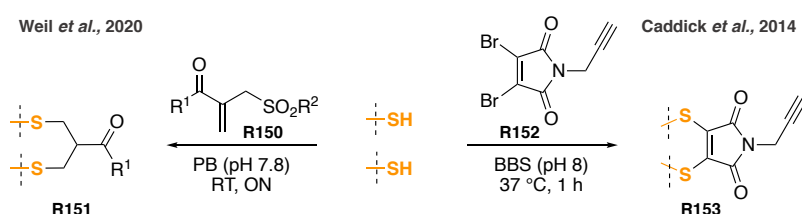
Scheme 13: Reconstruction of disulfide bonds with monosulfone **R150** and dibromomaleimide **R153** reagents.

The new-generation of maleimides (NGM), such as dibromomaleimides **R152**, has also found an interest in the reformation of disulfide bonds (**Scheme 13**).[176–179] Indeed, the incorporation of two nucleofuges, either bromide or thiophenol, across the double bond of the maleimide reagent, allowed the consecutive addition of two thiols. Described in 2011 by Caddick *et al.,* the dibromopyridazinediones **R32** (PDs) were later investigated as disulfide-bridging reagents. Even though, the structure of PDs is close to that of the new-generation of maleimides (NGM), they comprise four points of attachment: the first two, reacting with the free thiols through two consecutive addition-elimination steps, and two others containing different bioorthogonal tags allowing a dual labelling of the protein. In 2015, Chudasama and coworkers reported the use of PDs **R154** for the reconstruction of interchain bonds of the antibody trastuzumab (**Figure 8**).[180] These reagents allowed to recover almost completely the structure of the antibody while incorporating four sites of modifications as a whole. Since the reagent was designed to bear two orthogonal handles, a strained and a terminal alkyne, the antibody was functionalized first by SPAAC and then CuAAC with two different azide payloads per reconstructed disulfide bond. This led to the formation of a homogeneous antibody conjugate with a degree of conjugation equals to 8 with only four modifications introduced. Based on those results, new PD reagents were also designed to afford homogeneous antibody conjugates with a degree of conjugation of 2 and 4. DAR 4 conjugates were simply obtained by adding a single orthogonal tag on PDs **R155**.[181–183] Connecting two PDs together with a PEG linker **R156** allowed to link simultaneously the light chain with the heavy chain on the antibody and the two heavy chains together. Incorporating only one terminal alkyne on one of the DPs gave a single site of

functionalization instead of 2, thus giving them the possibility to generate very precisely DAR2 antibody conjugates.[184] The DP reagent was further optimized and it was shown in 2016 that the reduction and reoxidation of the disulfide bonds could be done in parallel if the reducing agent, TCEP, was directly incorporated to the DP derivative **R157**.[183,185] Finally, the DP molecule **R158** was further explored for the construction of bispecific antibody.[186] By attaching a strained alkyne onto one DP and a tetrazine on another one, two different Fab fragments could be linked together with the strain-promoted inverse electron-demand Diels-Alder cycloaddition (SPIEDAC). In addition, a terminal alkyne was added as a fourth point of attachment on the DP molecule, enabling a further functionalization of the bispecific antibody.
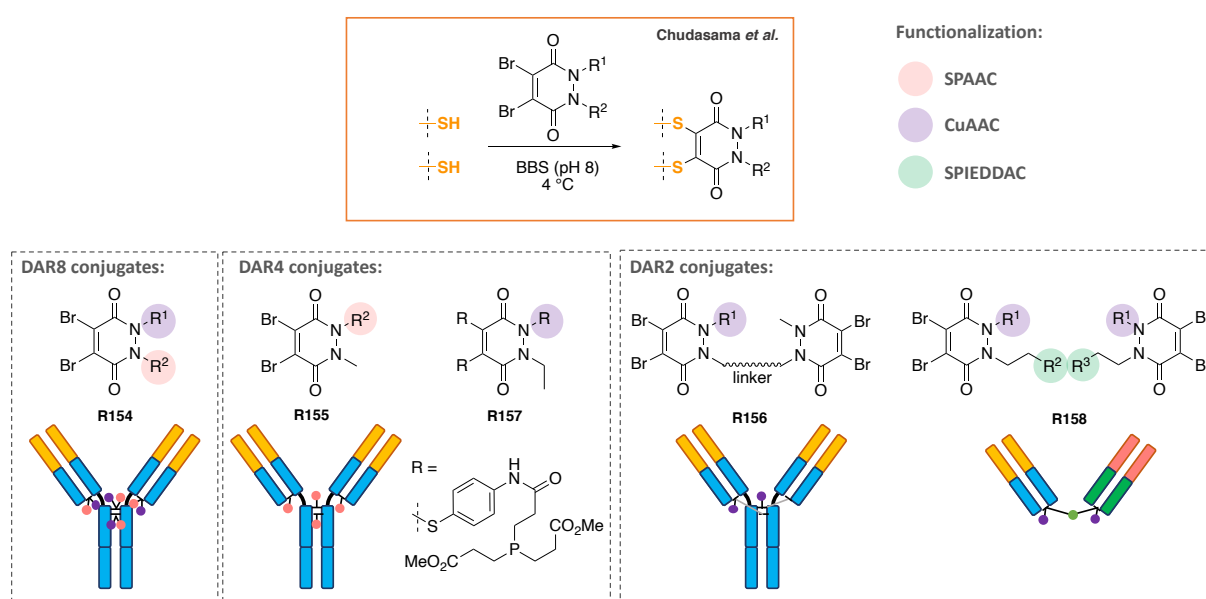
Figure 8: Use of dibromopyridazinediones for the reconstruction of disulfide bonds.

In 2017, Bernardes and coworkers showed that 3,3-bis(bromomethyl)oxetane **R159** was a suitable bridging reagent for the reconstruction of the disulfide bond on the Fab fragment of trastruzumab (**Scheme 14**).[187] It was shown that the oxetane bridge **R160** was more stable under reducing conditions and in human plasma than the native disulfide bond, and, could enhance receptor binding affinity. In the same perspective, Bernardes *et al.* proposed a new reagent derived from the oxetane, the isobutylene bromide **R161** (**Scheme 14**).[188] However, hundreds to thousands equivalents of reagent were necessary to modify a single cysteine and the reaction failed to elicit re-bridging of disulfide bonds.
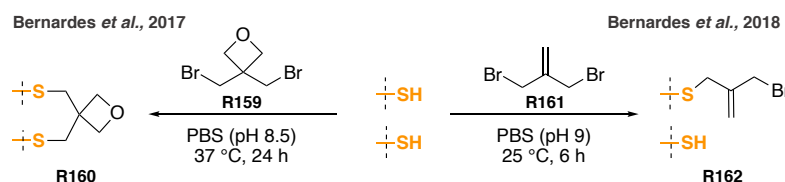
Scheme 14: Reconstruction of disulfide bonds with a 3,3-bis(bromomethyl)oxetane **R159** or an isobutylene bromide **R161** reagents.

The thiol-yne coupling reaction was proposed as a photochemical disulfide bridging strategy by Griebenow *et al* (**Scheme 15**). After reduction of the Fab-fragment of the antibody, M14-G07, the thiol-yne was carried out in the presence of 6-heptynoic acid **R163** and a radical initiator, lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP). The radical initiator would provide a free radical to the thiol that can subsequently react with the alkyne. In a second cycle, the second radical would be transferred to the other thiol that can then react with the alkene and allow the rebridging of the disulfide bond. Even though, a single equivalent of terminal alkyne was used, the conversion only reached 40% at very high antibody concentrations (46.1 mg/mL).[189]

Reported by our group and also based on a thiol-yne reaction, arylene dipropionitrile (ADPN) was described as a re-bridging reagent (**Scheme 15**).[190] The *meta* isomer of the ADPN **R165** has shown to be the most efficient reagent for cysteine conjugation on trastuzumab, despite incomplete re-bridging of the antibody being observed by gel electrophoresis.

Divinylpyrimidine **R167** (DVP) was found to be a successful thiol-ene reagent for the reconstruction of disulfide bonds (**Scheme 15**).[191,192] Derivatized with orthogonal tags such as a terminal alkyne, the resulting rebridged proteins were then functionalized by CuAAC to introduce cytotoxic drugs and fluorescent probes. Even though the reagent could efficiently label cysteines, an incomplete re-bridging of the antibody was observed. In the same perspective, Spring *et al.* recently proposed divinyltriazines **R168** (DVT) as rebridging reagents for antibodies.[193] As for the DVP reagents, it was not possible to fully reconstruct the antibody at the end of the reaction.

R163
radical initiator (LAP)
UV (365 nm)
PBS
4 °C, 4 h

R164

R165

PBS (pH 7.4)
25 °C, 12 h

R166

R168

Tris buffer (pH 8)
37 °C, 2 h

R169
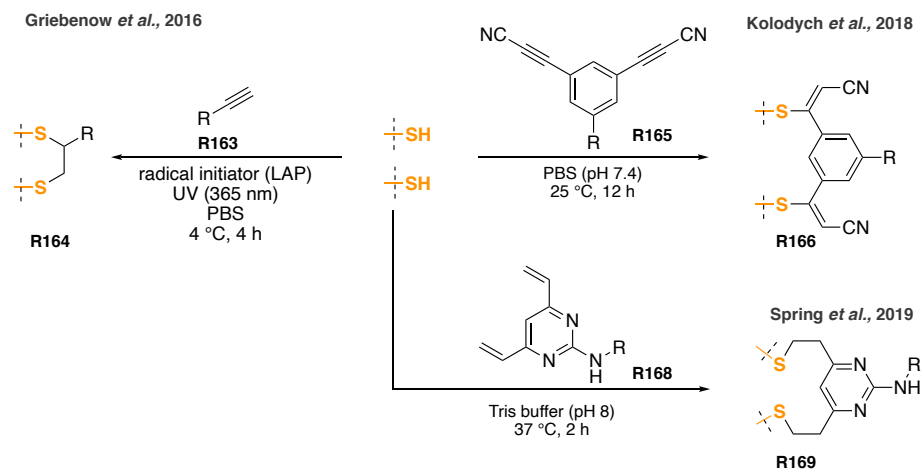
Scheme 15: Reconstruction of disulfide bonds with the thiol-yne and the thiol-ene reactions.

The reagents reported for the site-selective modification of cysteine residues are presented in the following **Table 2**.

Table 2: List of the reagents used for the site-selective modification cysteine residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R152 178			5	1 h / 37 °C	100%	CuAAC	2 steps	yes
R150 175			48	ON / RT	80%	iEDAA	4 steps	yes
R154 180			20	1 h / 4 °C	100%	SPAAC / CuAAC	8 steps	yes
R155 182			20	16 h / 4 °C	100%	SPAAC	<5 steps	yes
R156 184			16	16 h / 4 °C	100%	CuAAC	5 steps	yes
R159 187			20	24 h / 37 °C	-**	No	Commercially available	yes
R161 188			50-1000	1-6 h / 25 °C	100%	Alkylation with amines or thiols	Commercially available	unsure
R163 189			1	4 h / 4 °C	-	No	Commercially available	yes
R165 190			5	12 h / 25 °C	100%	SPAAC	5 steps	yes

R167[191]			10	2 h / 37 °C	90-95%	CuAAC	2 steps	yes
R168[193]			6.6	1 h / 37 °C	100%	CuAAC	4 steps	yes

* cysteine residue in orange, functionalization site in red

** information was not provided by the authors

ON stands for overnight

1.2.2.3. *N*-terminal residues

With a lower pKa than the ε-amino side chain of lysine (7.6-8.0 vs 9.3-9.5), the selective modification of the *N*-terminus of proteins have gained attention over the years.[10,11] General strategies have been described for the selective modification of all types of *N*-terminal residues over lysines, most of which employed aldehydes. Other general methods using sulfonamides, phthalimides or enzymes have also been published, while conjugation of specific *N*-terminal amino acids - notably *N*-terminal cysteine, glycine and proline – offered an extra level of selectivity.

1.2.2.3.1.   General methods

In 2006, Francis *et al.* were the first to report the use of aldehydes for the selective modification of the *N*-terminus.[194] By using pyridoxal-5-phosphate R170 (PLP), the α-NH$_2$ was converted to a pyruvamide R171 that was further functionalized with an oxyamine (**Scheme 16**). A single modification with nearly 70% conversion was observed on myoglobin. However, the reaction was found to be incompatible with *N*-terminal serine, threonine, cysteine and tryptophan. As an alternative, in 2013, they identified *N*-methyl pyridinium 4-carboxyaldehyde R172 (Rapoport's salt) as a new and more effective reagent, leading to the selective transamination of the *N*-terminus (**Scheme 16**).[195] Wild type trastuzumab was successfully conjugated on the heavy chain and a double modification of the antibody was achieved when *N*-terminal glutamate residue was introduced on the light chain. As the scope of *N*-terminal amino acids was found to be limited with the Rapoport's salt, 2-pyridinecarboxyaldehyde R174 was latter reported (**Scheme 16**).[196] This reagent allowed the formation of a stable imidazolidinone R175 on the α-NH$_2$, with the imine formed in a first step attacked by the amide backbone in the second step. Despite good conversions, the method could not be applied to proteins containing a *N*-terminal proline or glycine.
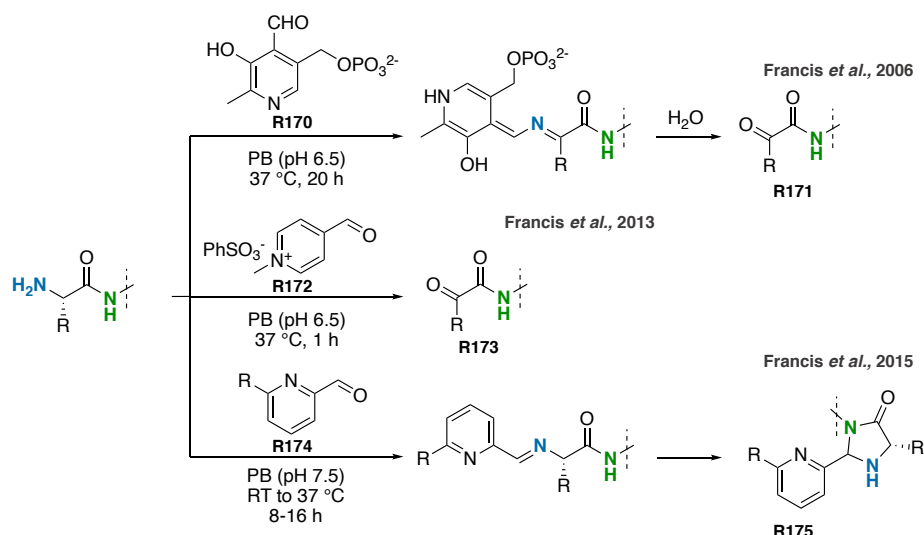
Scheme 16: Labelling of *N*-terminal residues with a pyridoxal-5-phosphate **R170** (PLP), a *N*-methyl pyridinium 4-carboxyaldehyde **R172** or 2-pyrdinecarboxyaldehyde **R174**.

Following the strategies developed by Francis and coworkers, Deng and Lai recently described 2-ethynylbenzaldehyde **R176** (2-EBA) as a new reagent for the selective conjugation of proteins' *N*-terminal residues (**Scheme 17**).[197] Following initial imine formation, a 6-*endo-dig* cyclisation delivered an isoquinolinium derivative as the final product **R177**. Even if the strategy was found to be suitable for the modification of three proteins – lysozyme, RNAse A and BCarg mutant – at pH 7.4 with moderate to good conversions, undesirable labelling of lysine was still observed.
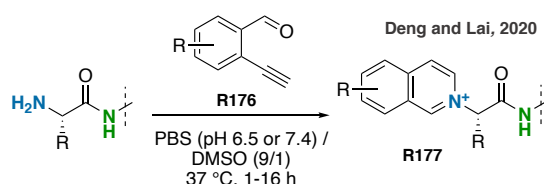
Scheme 17: Labelling of *N*-terminal residues with 2-alkynylarylaldehydes **R176**.

Based on the native chemical ligation (NCL), the aldehyde capture ligation (ACL) was described in 2015 by Arora *et al* (**Scheme 18**).[198] The *o*-benzaldehyde ester **R178** was able to condense with the primary *N*-terminal amine to either give an imine **III**, a hemiaminal ester **IV** or a hemiaminal intermediate **II**. Subsequent rearrangement of the three intermediates led to the formation of a stable amide product **R179** and an *o*-benzaldehyde auxiliary **VI**. With only 5 equivalents of an *o*-benzaldehyde ester derivative, 70% of mono labelled ubiquitin was obtained after 40 hours with only the α-NH$_2$ being modified.
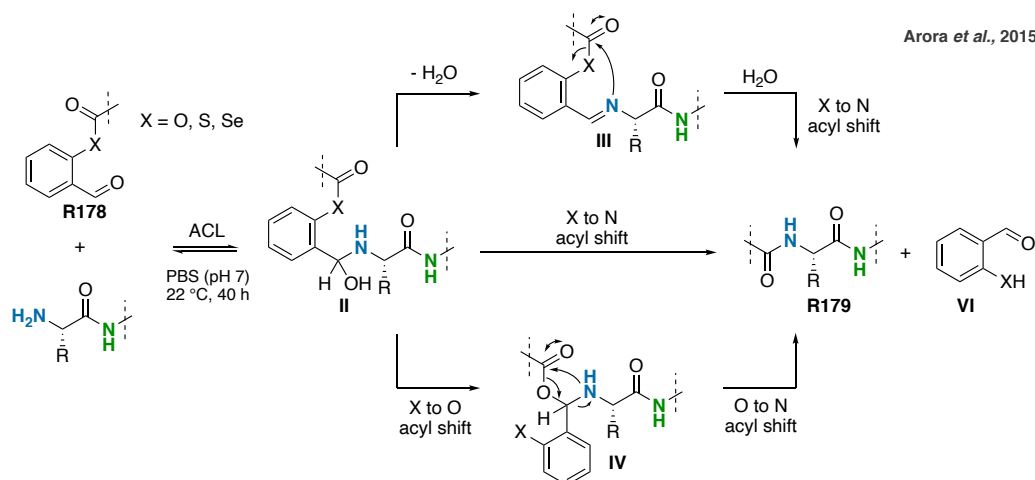
Scheme 18: Use of the aldehyde capture ligation (ACL) for the site-selective modification of *N*-terminus. Three mechanism were proposed by the authors to explain the mechanism of this transformation. Adapted from: Raj, M.; Wu, H.; Blosser, S. L.; Vittoria, M. A.; Arora, P. S. Aldehyde Capture Ligation for Synthesis of Native Peptide Bonds. *J. Am. Chem. Soc.* **2015**, *137 (21),* 6932–6940

Rai and coworkers showed in 2018 that site-selective modification of *N*-terminus with epoxides **R180** and sulfonate esters **R182** could be achieved only if they were combined with an aldehyde on a bifunctional reagent (**Scheme 19**).[199] It was suggested that the fast and reversible imine formation could first guide and enhance the rate of the slow and irreversible addition of the epoxide or sulfonate ester in a second step. This proved to be effective with epoxy-aldehyde reagents leading to 62% of mono-labeled RNAse A conjugates in 12 hours with a *N*-terminal selective modification, while sulfonate esters derivatives gave 48% of mono-labeled conjugate.

Chou and coworkers proposed to modify the *N*-terminal residues via reductive alkylation with commercially available functionalized benzaldehyde derivatives **R184** (**Scheme 19**).[200] In presence of a benzaldehyde derivatives **R184** and NaBH$_3$CN, different proteins (GLP-1, RNase A, lysozyme, aldolase, insulin, creatine phosphokinase) were mono-labelled on the α-NH$_2$ in one-pot with conversions going from 30% to 90%.
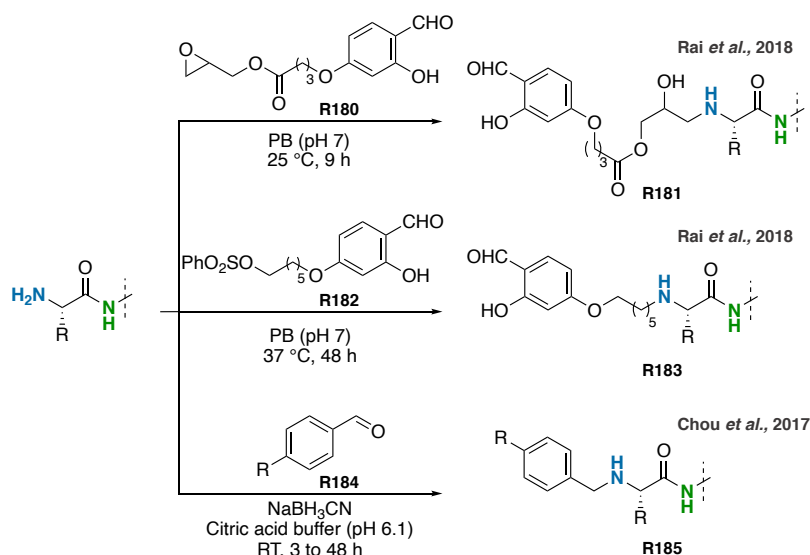
Scheme 19: Labelling of *N*-terminal residues with an epoxide **R180** or a sulfonate ester **R182** or use of the reductive amination reaction.

Moderate to high *N*-terminus selectivity could also be achieved with ketene reagents **R186** (**Scheme 20**).[201] Decreasing the pH could enhance the selectivity at the expense of conversion: adding the ketene in small portions helped to improve the conversion of the reaction, albeit it never exceeded 40%.

In 2020, Raj and Tang proposed an azolation strategy for the selective *N*-terminal labelling of proteins (**Scheme 20**).[202] The nucleophilic attack of the α-NH$_2$ on the oxazoline reagent **R188** led to the formation of a stable C-N bond under physiological conditions on both peptides and proteins. Good conversions (50 – 99%) and excellent *N*-terminal selectivity were achieved on five different proteins – myoglobin, cytochrome c, α-lactalbumin, ubiquitin and insulin – under physiological conditions. Moreover, the addition of the oxazoline group to peptides/proteins increased the ionization of their peptide fragments in mass spectrometry and enhanced the sensitivity of the method, making this conjugation strategy valuable for proteomics. Despite all the advantages found for this method – high selectivity, fast kinetics, high yield, high stability of the payload, mass-sensitivity booster – it was unfortunately not tested on proteins with molar mass above 17 kDa.
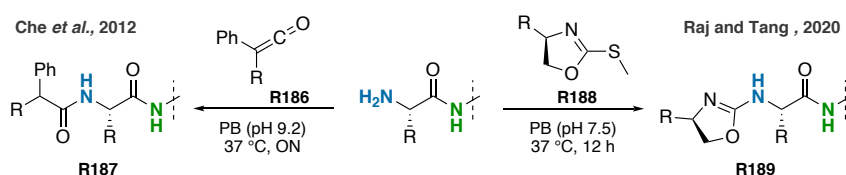
Scheme 20: Labelling of *N*-terminal residues with ketene derivatives **R186** and oxazoline reagents **R188**.

Francis and coworkers proposed an oxidative strategy for the efficient labelling of α-NH$_2$ (**Scheme 21**).[203] The strategy used *o*-aminophenols **R190** that were oxidized *in situ* with potassium ferricyanide. The resulting quinones then reacted with the *N*-terminal amino acid of different proteins, such as lysozyme, RNase A, myoglobin, GFP, MS2 and chymotrypsinogen. The reaction showed moderate reactivity with only 30 to 40% of mono-labeled conjugate obtained and did not tolerate free cysteine.
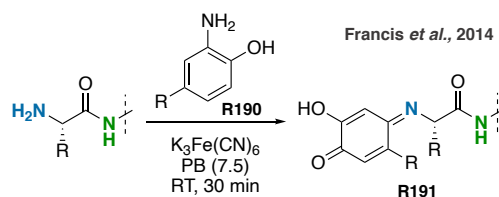
Scheme 21: Labelling of *N*-terminal residues with *o*-aminophenols **R190** and potassium ferricyanide

A new reagent was described in 2011 for the introduction of azides in proteins, imidazole-1-sulfonyl azide.[204] Even though, the compound was not *N*-terminus selective on lysozyme and led to side conjugation on lysines, this article paved the way for new classes of *N*-terminus selective reagents, such as vinylsulfonamides **R192**, which were capable of modifying selectively the α-NH$_2$ of somatostatin, lysozyme and RNase A in 5 hours at pH 6 with three equivalents (**Scheme 22**).[205]

Phthalimidation of primary amines was found to be effective for the selective modification of the *N*-terminus of RNase A (**Scheme 22**).[206] Site-selectivity was favored when a single equivalent of *N*-hydroxyphthalimide **R194** was used, even though poor conversion was obtained (50%).
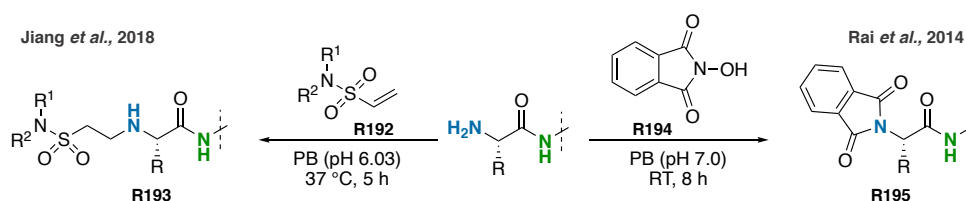
Scheme 22: Labelling of *N*-terminal residues with a vinylsulfonamide **R192** or a *N*-hydroxyphthalimide reagent **R194**.

The reagents reported for the site-selective modification of *N*-terminal residues are presented in the following **Table 3**.

Table 3: List of the reagents used for the site-selective modification of *N*-terminal residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R170 [194]			20-30	20 h / 37 °C	69-75%	Oxime ligation	Commercially available	yes
R172 [195]			-**	1 h / 37 °C	37%	Oxime ligation	Commercially available	yes
R174 [196]			10-400	16 h / RT to 37 °C	0-96%	No	4 steps	yes
R176 [197]			5-10	16 h / 37 °C	40-92%	CuAAC	3 steps	yes
R178 [198]			5	40 h / 22 °C	70%	No	< 2 steps	yes
R180 [199]			25	9 h / 25 °C	62%	Oxime ligation	2 steps	yes
R182 [199]			25	48 h / 37 °C	40-48%	Oxime ligation	2 steps	yes
R184 [200]			2	3-48 h / RT	30-90%	Oxime ligation / CuAAC	Commercially available	yes
R186 [201]			6-500	ON*** / 37 °C	<40%	CuAAC	3 steps	yes
R188 [202]			50 - 300	12 h / 37 °C	50-99%	SPAAC / CuAAC	4-6 steps	yes
R190 [203]			2-10	< 30 min / RT	30-40%	No	-	yes
R192 [205]			3	5 h / 37 °C	33-44%	CuAAC	5 steps	yes
R194 [206]			1-5	8-24 / RT	45-50%	No	Commercially available	yes

* *N*-terminus residue in blue, functionalization site in red

** information was not provided by the authors

*** ON stands for overnight

Rao and cowokers reported a water compatible condensation reaction for the labelling of *N*-terminal cysteines (**Scheme 23**).[207] Condensation between a 2-cyanobenzothiazole **R196** (CBT) and a *N*-terminal cysteine was achieved on peptides and proteins *in vitro* and at cell surface. Based on the CBT reagent, the condensation of 2-((alkylthio)(aryl)methylene)malononitrile **R198** (TAMM) with *N*-terminal cysteine on GFP was described by Wu and coworkers as new site-specific method (**Scheme 23**).[208] The vinyl sulfide group of TAMM **R198** was first substituted with the thiol of the cysteine, through a Michael-addition / elimination sequence. Subsequent thiazolidine formation by Michael addition of the α-NH$_2$ was followed by malonitrile elimination to give the final 2-aryl-4,5-dihydrothiazole (ADT) payload **R199**. While successfully applied to peptides and GFP, this method was not extended to other proteins.
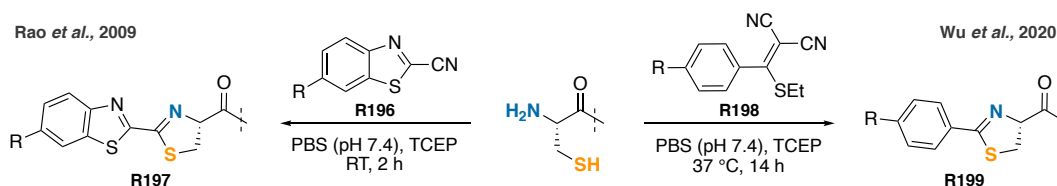
Scheme 23: Labelling of *N*-terminal cysteines with 2-cyanobenzothiazole **R196** (CBT) or 2-((alkylthio)(aryl)methylene)malononitrile **R198** (TAMM) reagents.

As an alternative method, Gois and Gao independently described 2-formylphenyl boronic acids **R200** (2-FPBA) for the fast modification of *N*-terminal cysteines at neutral pH with fast kinetics (rate constants greater than 10^3 M^{-1}s^{-1}) (**Scheme 24**).[209,210] The imine formed by condensation of the primary amine with the aldehyde was activated by the boronic acid facilitating the formation of thiazolidines. Even though, the reaction was highly efficient, the thiazolidine scaffold was found to be unstable in presence of others free cysteines, possibly due to the reversible reaction leading back to the linear imine. To overcome this problem, Gao and coworkers hypothesized that the acylation of the thiazolidine nitrogen would disfavor this pathway and hence give stable conjugates.[211] To do so, they developed an acetylated version of 2-FPBA **R202**, which, after rapid thiazolidine formation, gave stable protein conjugates thanks to an intramolecular acyl transfer.
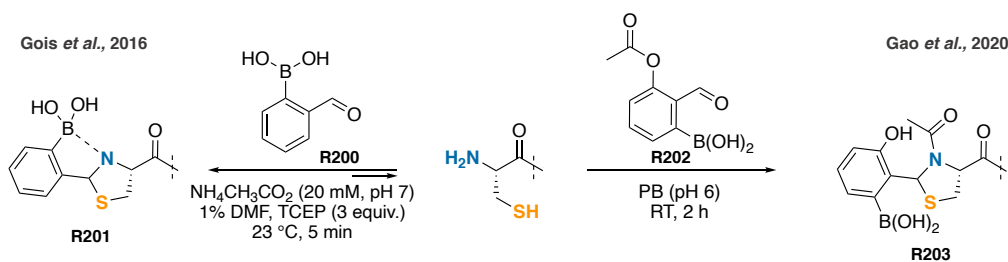
Scheme 24: Labelling of *N*-terminal cysteines with 2-formylphenyl boronic acid derivatives.

1.2.2.3.2.2.   *N*-terminal glycine residues

In 2019, Rai *et al.*, showed that *N*-terminal glycine could form amino alcohols **R205** when reacted with *ortho*-substituted benzaldehydes **R204** (**Scheme 25**).[212] The reaction was proved to be regio- and chemoselective for the *N*-terminal glycine as proteins not containing this residue were not labelled. Nevertheless, large excess of aldehyde (500 equivalents) was necessary to reach moderate conversions.
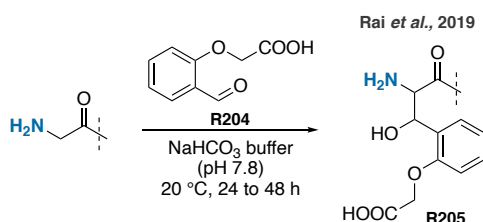
Scheme 25: Labelling of *N*-terminal glycine with *ortho*-substituted benzaldehyde derivatives **R204**.

1.2.2.3.2.3.   *N*-terminal proline residues

More recently, selective modification of *N*-terminal prolines with the Petasis reaction was shown by Raj and coworkers (**Scheme 26**).[213] The condensation of the secondary amine with an aldehyde **R206** led to the formation of an iminium ion that then reacted with a boronic acid **R207** to give the Petasis product **R208**. No reaction was observed on proteins not containing this *N*-terminal residue, thus confirming the chemo- and regioselectivity of the reaction.

In a similar approach, Francis *et al.* proposed in 2019 to use a tyrosinase enzyme to oxidize selectively phenol groups **R209** to highly reactive *o*-quinone that then reacted with *N*-terminal prolines (**Scheme 26**).[214] Even though, the reaction proceeded with other *N*-terminal amino acids, the conversions obtained were below 40%, thus confirming the necessity to have proline

residues to achieve excellent conversions (> 90%). The enzyme used also showed no side reactivity towards the native tyrosines present in the protein, as no background oxidation was observed.
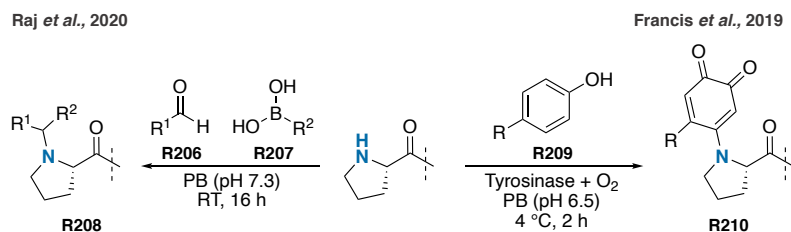
Scheme 26: Labelling of *N*-terminal proline residues with the Petasis reaction or with phenol derivatives and tyrosinase.

The reagents reported for the site-selective modification of *N*-terminal cysteine, glycine and proline residues are presented in the following **Table 4**.

Table 4: List of the reagents used for the site-selective modification of *N*-terminal cysteine, glycine and proline residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R196 207			-**	2 h / RT	100%	no	3 steps	yes
R198 208			10	14 h / 37 °C	100%	no	> 2 steps	yes
R200 210			-	5 min / 23 °C	100%	no	Commercially available	yes
R202 211			1-5	2 h / RT	100%	no	3 steps	yes
R204 212			500	24-48 h / 20 °C	40-71%	Oxime ligation	2 steps	yes
R206 / R207 213			5:50	16 h / RT	-	no	Commercially available	yes
R209 214			10	2 h / 4 °C	90%***	no	1 step	yes

* *N*-terminus residue in blue, functionalization site in red

** information was not provided by the authors

*** conversion for Pro-sfGFP; conversions were not given for the other proteins tested.

1.2.2.4. Tyrosine residues

As tyrosine residues are less often observed on the surface of proteins and moderately abundant (around 3%), they have become an interesting target for site-selective modification of proteins.[10,11] Over the years, different reagents and methods were developed to target this amino acid, such as diazonium reagents, or multicomponent, radical and palladium-catalyzed reactions.

Barbas III and coworkers were one of the first researchers to show interest in the modification of this amino acid. In 2010 and 2012, they developed diazonium derivatives that could selectively target tyrosines, such as cyclic diazocarboxamides **R211 (Scheme 27)**.[215] Despite some side reactivity observed on lysine and tryptophan in peptides, the method was still found to be regioselective on chymotrypsinogen A and BSA with two and four sites of conjugation identified, respectively. However, this strategy turned to be inefficient and not chemoselective on myoglobin (4% conversion).

In 2012, Barbas III *et al.* described 4-formylbenzene diazonium hexafluorophosphate **R213** (FBDP) for the conjugation of tyrosine residues (**Scheme 27**).[216] Excellent chemoselectivity was demonstrated; however, no information was given concerning the regioselectivity. Based on diazonium salts, Hulme *et al.* developed a catch-and-release protein-tagging strategy for the selective modification of tyrosines.[217] Based on surface-surface interaction between the protein of interest and a resin tethered with a chemoselective electrophile such as diazonium salt, only solvent accessible tyrosine residues on the protein were successfully conjugated. The protein, attached to the resin with an azobenzene group, can be released under mild conditions using dithionate. The cleaved protein bore an *o*-aminophenol modification which was functionalized by oxidative coupling with a fluorophore, fluoresceinamine. This solid-phase approach was found to be more selective than the other approaches with one or two sites of modification being identified on diverse proteins (RNase A, SBTI and HEWL).
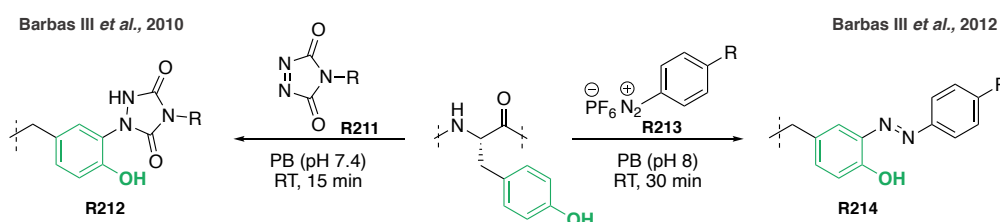
Scheme 27: Labelling of tyrosine residues with cyclic diazocarboxamide **R211** or 4-formylbenzene diazonium hexafluorophosphate **R213** (FBDP) reagents.

Multicomponent reactions, and most notably the Mannich reaction, have emerged as a new strategy for the regioselective modification of tyrosine. In 2004, Francis and coworkers were the first to report a three-component Mannich-type reaction for the selective conjugation of tyrosine on proteins (**Scheme 28**).[218] At pH 6.5, in the presence of formaldehyde **R215** and aniline **R216**, proteins bearing solvent accessible tyrosines, such as chymotrypsinogen A, lysozyme and RNase A, were successfully modified. Although, the chemoselectivity of the reaction was confirmed, it was not possible to draw conclusions regarding the regioselectivity. Indeed, the native mass spectra showed three modifications on the proteins when only a single site of modification was identified. As an extension of the Mannich reaction, cyclic imines **R218** were reported for the labelling of tyrosine on proteins (lysozyme, myoglobin, cytochrome c and chymotrypsinogen A) (**Scheme 28**).[219] However, they showed poor reactivity: formation of a mono adduct was only observed for lysozyme after four days of incubation at pH 6 and 37 °C. Another three-component reaction was described by Ball *et al.* for the selective modification of tyrosine through transmetalation of a boronic reagent (**Scheme 28**).[220] Peptides were successfully modified with boronate **R220** (10 equiv.) and a rhodium(III) complex **R221** (10 equiv.) in buffered medium at pH 9.4. Additional equivalents of boronate and Rh(III) (50 equiv.) were required for the modification of proteins (chymotrypsinogen A, myoglobin, ovalbumin, RNase A, BSA and trastuzumab). Because the labelling was only proved by fluorescence, the method could not be called regioselective, since no MS analyses were performed.
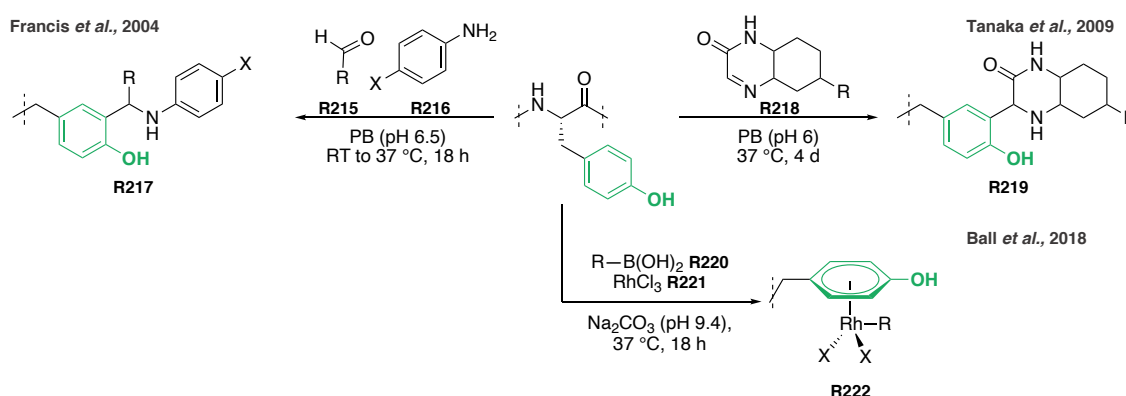
Scheme 28: Use of the Mannich reaction and a three-component reaction for the labelling of tyrosine residues.

Palladium-catalyzed reactions were described as a new strategy for the *O*-alkylation of tyrosines (**Scheme 29**).[221] With only five equivalents of allylic acetate **R223** and a catalytic amount of palladium acetate, tyrosines of chymotrypsinogen A and α-chymotrypsin were efficiently modified in less than an hour. No modification was observed on myoglobin since

there was no tyrosine accessible for the modification, thus confirming the chemoselectivity. However, it was not possible to conclude regarding the regioselectivity since only one site of modification was observed by MS / MS but two adducts formed by native MS.

The selective generation of radicals on tyrosines was the last regioselective method described. In 2011, Kamiya *et al.* used this strategy for the cross-linking of BAP proteins.[222] In presence of horseradish peroxidase (HRP) and H_2O_2, tyrosines underwent one-electron oxidation reaction and formed *C*-centered tyrosyl radicals that reacted with other phenolic moieties, including tyrosine residues of other proteins. The reaction was assumed to be site selective by computational study, with only one accessible identified site. Based on a very similar strategy, Sato and Nakamura recently demonstrated that it was possible to selectively modify tyrosines with HRP, H_2O_2 and *N*-methylated luminol derivative **R225** (**Scheme 29**).[223] MS / MS analysis revealed a unique site of modification on the antibody trastuzumab (heavy chain: Tyr57) whereas four sites were identified on the antibody rituximab (heavy chain: Tyr32, Tyr52, Tyr101 and Tyr102). This difference was explained by the number of residues being solvent accessible, which are preferentially modified on this reaction.
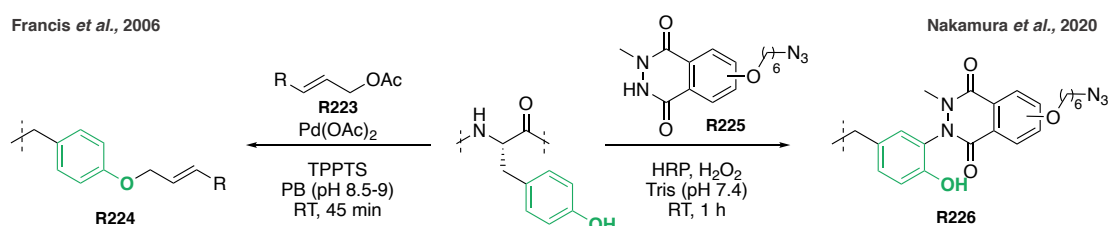
Scheme 29: Labelling of tyrosine residues with a palladium-catalyzed reaction or a single-electron transfer catalyzed reaction.

The reagents reported for the site-selective modification of tyrosine residues are presented in the following **Table 5**.

Table 5: List of the reagents used for the site-selective modification of tyrosine residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R211 [215]			-**	15 min / RT	4-96%	Oxime ligation	4 steps	yes
R213 [216]			3-10	30 min / RT	-	Oxime ligation	1 step	Unsure (yes[217])

	Reagent	Product		Time / Temp	Yield		Steps	
R215 / R216 218			1250	18-24 h / RT to 37 °C	60%	no	1 step	unsure
R218 219			-	4 d / 37 °C		no	7 steps	unsure
R220 / R221 220			10-50	18 h / 37 °C	-	no	3 steps	unsure
R223 221			5	<1 h / RT	50-65%	no	2 steps	unsure
R225 223			60	14 h / RT	80%	SPAAC	-	yes

* tyrosine residue in green, functionalization site in red

** information was not provided by the authors

1.2.2.5. Tryptophan residues

With a natural abundance of 1.1%, tryptophan is the least abundant amino acid in proteins and hence, very interesting for site selective modification.[11] Over the years, different research groups have tried to modify this amino acid, albeit with limited success on proteins, harsh conditions being often necessary to achieve the modification. Three types of reagents were envisioned for the selective conjugation of tryptophan residues: metallocarbenes, hypervalent iodines and radicals.

In 2004, Francis *et al.* proposed to use a metallocarbene **R227** (generated from $Rh_2(OAc)_4$ and a diazo compound) in the presence of hydroxylamine to obtain a selective modification of tryptophans on myoglobin (**Scheme 30**).[224] Due to the poor accessibility of tryptophan residues, the reaction had to be run under denaturating conditions, at pH 1.5, which caused protein denaturation. However, the protein was re-folded after the conjugation and found to be competent to bind the heme group again. Nevertheless, only two tryptophans were modified with this method and 60% conversion was achieved when 100 equivalents of metallocarbene were used. By choosing *tert*-butylhydroxylamine, as an additive instead of hydroxylamine, the same reaction proceeded under less acidic conditions on lysozyme (pH 6.0) (**Scheme 30**).[225] However, due to the poor accessibility of tryptophans in lysozyme, the reaction had to be conducted at high temperatures (85 °C), thus leading to the protein denaturation. However, when a mutant of FKBP protein was designed to bear an accessible tryptophan, its modification was found to be possible at room temperature and pH 6.0 over 15 hours.
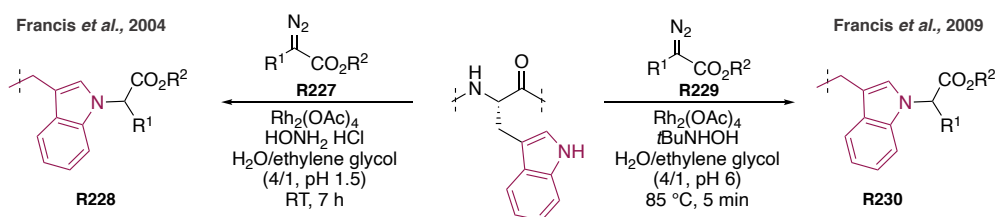
Scheme 30: Labelling of tryptophan residues with metallocarbenes **R227/R229**.

Combination of hypervalent iodines, such as TIPS-EBX **R231**, and a gold catalyst, AuCl(SMe$_2$), was used for the modification of small molecules and small proteins (mellitin and apomyoglobin) (**Scheme 31**).[226] The method was proved to be regio- and chemoselective on apomyoglobin with only two tryptophans being modified. However, even if 5 equivalents of hypervalent iodine and stoichiometric amount of catalyst were used, the reaction had to be performed in organic solvent (3/1 acetonitrile/water) which was the main limitation of the method.

Radicals were also found to be suitable new reagents for the modification of tryptophans as they showed fast reactivity and selectivity for this residue under physiological conditions and did not require the use of toxic transition metals. Single modification of tryptophan on various proteins was achieved by Kanai *et al.* by mixing a single equivalent of keto-ABNO **R233** with NaNO$_2$ (**Scheme 31**).[227] This resulted in a mixture of mono-adduct and non-modified protein with conversions oscillating between 11 and 64% on 5 different proteins – lysozyme (64%), myoglobin (53%), concanavalin (11%), BSA (25%) and β2-microglobulin (15%). Site selectivity was only confirmed on a single tryptophan, Trp62, on lysozyme.
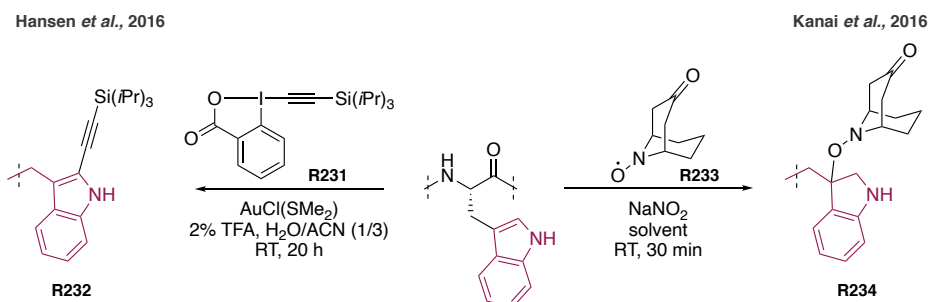
Scheme 31: Labelling of tryptophan residues with an hypervalent iodine, TIPS-EBX **R231**, and a gold catalyst, or, with a keto-ABNO **R233**.

Radical trifluoromethylation of tryptophans was described by Davis *et al.* (**Scheme 32**).[228] By generating a radical CF$_3$ from sodium trifluoromethanesulfinate **R235** (NaTFMS or Langlois reagent) and *tert*-butyl hydroperoxide (*t*BuOOH) under aqueous conditions, tryptophan

residues of different proteins (myoglobin, lysozyme, panthonetate synthetase) were selectively modified. This direct trifluoromethylation technique allowed the analysis fluorinated protein constructs by ^{19}F NMR.

Taylor and coworkers proposed a photochemically driven bioconjugation reaction for the modification of tryptophan residues based on *N*-substituted pyridinium salt **R237** and UV-B irradiation (**Scheme 32**).[229] The photoexcitation process allowed generating tryptophan and pyridinium positive radical species that could combine to give a carbamate-modified tryptophan conjugate. This process was found to work best when glutathione (GSH) was added as an additive, and, excellent levels of selectivity for tryptophan residues were noticed on peptides and lysozyme.
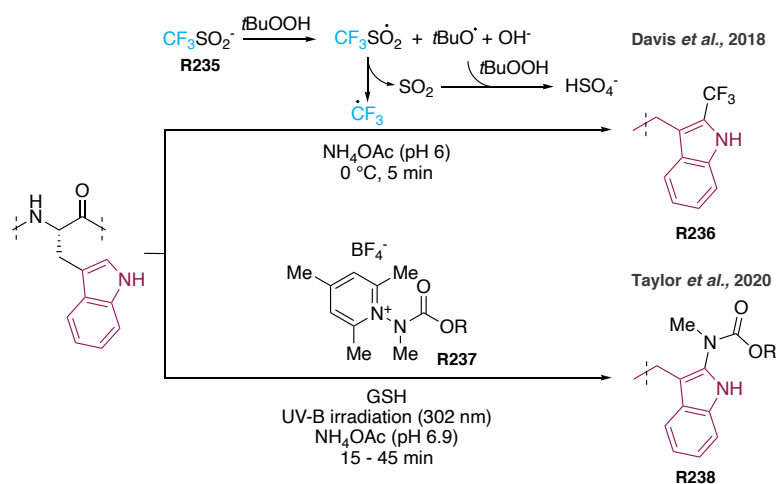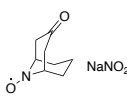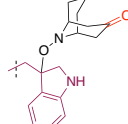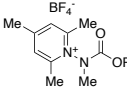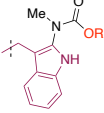
Scheme 32: Labelling of tryptophan residues with radical species.

The reagents reported for the site-selective modification of tryptophan residues are presented in the following **Table 6**.

Table 6: List of the reagents used for the site-selective modification of tryptophan residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R227 224	[M] R[1] CO_2R[2] + NH_2OH	(structure) CO_2R[2] R[1]	100	7 h / RT	60%	no	2 steps	yes
R229 225	[M] R[1] CO_2R[2] + tBuNHOH	(structure) CO_2R[2] R[1]	200	5 min / 85 °C	62%	no	2 steps	yes
R231 226	(structure) O—I≡—Si(iPr)_3 O AuCl(SMe_2)	(structure) N≡—Si(iPr)_3	5	20 h / RT	75%	TIPS removal then CuAAC	Commercially available	yes

	*	*						
R233 227	NaNO$_2$	(structure)	5	30 min / RT	0-64%	Oxime ligation	Commercially available	yes
R235 228	CF$_3$SO$_2^-$ *t*BuOOH	CF$_3$ (structure)	200 : 25	5 min / 0 °C	100%	no	Commercially available	yes
R237 229	BF$_4^-$ Me, Me, Me, Me, Me, OR (structure)	Me, OR (structure)	5-25	45 min / **	77-94%	CuAAC	4 steps	yes

* tryptophan residue in purple, functionalization site in red

** information regarding the temperature is not provided

1.2.2.6.　Histidine residues

The chemoselective modification of histidine has been challenging due to the presence of more reactive groups on proteins, such as thiols and amines, that can outcompete the imidazole ring for nucleophilic substitutions with different electrophiles, as observed with iodomethane for example.[10] Development of new electrophiles that can permit chemo- and regioselective modification of histidine were thus developed.

Rai and coworkers screened several electrophiles on the model protein ubiquitin, which contains only a single histidine residue.[230] Lack of chemoselectivity was observed for iodoacetamide, maleimide and α,β-unsaturated sulfone reagents, even though the latter exhibited very good reactivity. In contrast, 2-cyclohexenone **R239** showed modest reactivity but good chemoselectivity for histidine (**Scheme 33**). Even though a single site of modification was identified on several proteins, the method required between 150 and 250 equivalents of 2-cyclohexenone for a single histidine and gave only moderate conversions (from 25 to 52%, with the exception of lysozyme c: 80%).
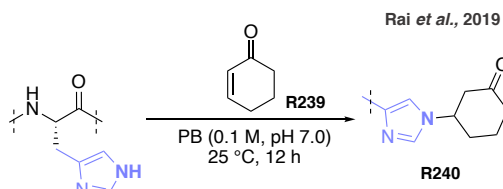
Scheme 33: Labelling of histidine residues with 2-cyclohexanone **R239**.

Based on the lysine-directed lysine method they had previously developed (see page 22, **Figure 7**), Rai *et al.* relied again on a linchpin to direct the installation of an electrophile on

histidines (**Scheme 34**).[231] To do so, the authors designed a new reagent **R241** bearing two reactive parts: a first part corresponding to an aromatic aldehyde that formed stable Schiff bases with lysines present at the surface of the protein; and a second part, containing an epoxide that could further react with a proximal histidine. In a last step, the Schiff base was hydrolyzed to recover the aldehyde that could then be further functionalized with an oxyamine. By tuning the size of the linker between the two reactive groups, regioselectivity and protein selectivity was achieved in a mixture of proteins.
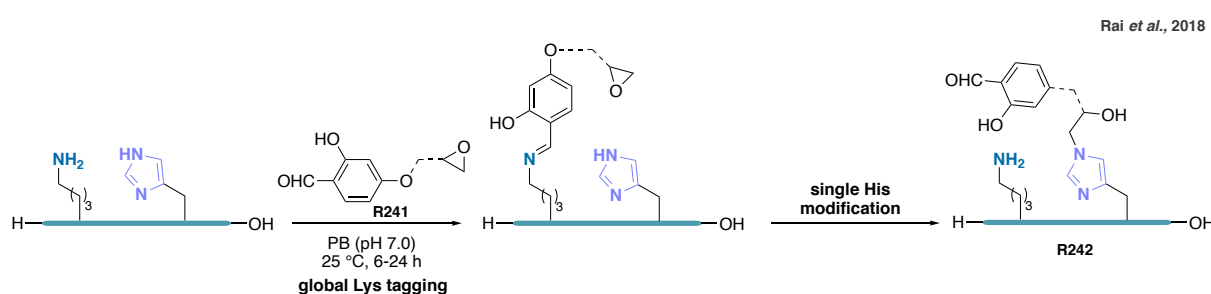
Scheme 34: Lysine-directed histidine modification strategy.

The reagents reported for the site-selective modification of histidine residues are presented in the following **Table 7**.

Table 7: List of the reagents used for the site-selective modification of histidine residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R239 230			150-250	12 h / 25 °C	25-80%	Oxime ligation	Commercially available	yes
R241 231			25	6-24 h / 25 °C	34-57%	Oxime ligation	2-7 steps	yes

*histidine residue in purple, functionalization site in red

1.2.2.7.   Methionine residues

In contrast to the cysteine residue that has been extensively studied over the years, its cousin, methionine, has been left in the shadow. Methionine is the second rarest amino acids in proteins and often buried within the interior of protein cores and not accessible for modifications.[232] Exploiting the poor solvent accessibility of this residue was thus envisioned as a good strategy for the development of a regioselective method.

Site selective conjugation of methionine residues on peptides and small proteins (5 to 14 kDa) was achieved with hypervalent iodines (iodonium salt) **R243** in presence of thiourea, TEMPO and formic acid (**Scheme 35**).[233] Selectivity was achieved on small proteins with one or two sites being modified, however, it is not known if a similar selectivity could be achieved on proteins bigger than 15 kDa.
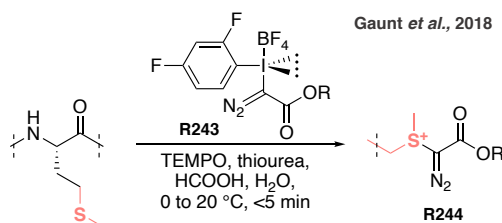
Scheme 35: Labelling of methionine residues with an hypervalent iodine reagent **R243**.

The reagents reported for the site-selective modification of methionine residues are presented in the following **Table 8**.

Table 8: List of the reagents used for the site-selective modification of methionine residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R243 [233]			500	<5 min / 0 to 20 °C	84-95%	Photoredox radical cross coupling reaction	3 steps	yes

*methionine residue in corail, functionalization site in red

1.2.2.8. Arginine residues

Arginine is a relatively abundant amino acid in proteins (> 5%).[11] However, compared to lysines, they have a lower tendency to be located at the protein surface consisting an interesting feature for the development of site-selective methods.

Klok *et al.* showed that arginine residues present on lysozyme were modified with glyoxal derivatives **R245** with a good regioselectivity (four arginines modified out of eleven) despite the excess of reagent used (**Scheme 36**).[234,235] Even though side reactivity with lysine amines occurred, a final treatment with a solution of hydroxylamine allowed to regenerate the lysine residues at the end of the reaction. The main limitation of this reaction was attributed to very slow kinetics; 14 days of reaction at pH 9 with 50 equivalents of oxo-aldehyde were required to reach a complete conversion.

Scheme 36: Labelling of arginine residues with an oxo-aldehyde derivative R245.

The reagents reported for the site-selective modification of arginine residues are presented in the following **Table 9**.

Table 9: List of the reagents used for the site-selective modification of arginine residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R245 234,235			1.25-4.5 / arg	14 d / **	100%	no	3 steps	yes

*arginine residue in blue

** information regarding the temperature is not provided

1.2.2.9.   Serine residues
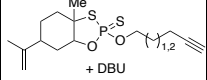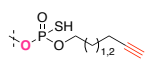
Due to the high abundance (> 6%) of serine and the poor nucleophilicity of alcohols compared to other functionalities, such as thiols and amines, the development of chemoselective and site-selective methods for serine conjugation has been a long-time challenge only met recently by Baran and coworkers **(Scheme 37)**.[11,236] They showed that phosphorus(V)-based electrophiles **R247** reacted preferentially with alcohols. In the presence of DBU, those electrophiles modified serine, as free amino acid or in peptides, in high yield and great selectivity. Unfortunately, when the reaction was transposed to ubiquitin, the conversion was significantly lowered (20 to 40%). Nevertheless, site-selectivity was achieved with this reagent as only mono-labeled conjugate was identified with only Ser65 being modified, out of two other serines. Despite the novelty of the method, the authors claimed that the phosphorus(V) electrophiles still needed to be optimized to make it more reactive on proteins, more stable in aqueous conditions (reaction is done in 80% of organic solvent) and compatible with cysteine residues.

Scheme 37: Labelling of serine residues with phosphorus(V)-based electrophiles **R247**.

The reagents reported for the site-selective modification of serine residues are presented in the following **Table 10**.

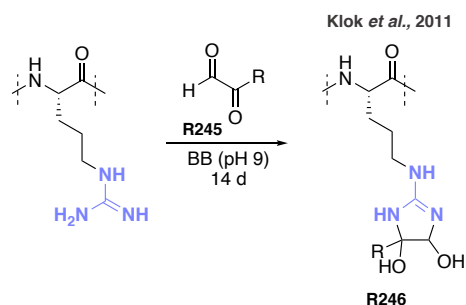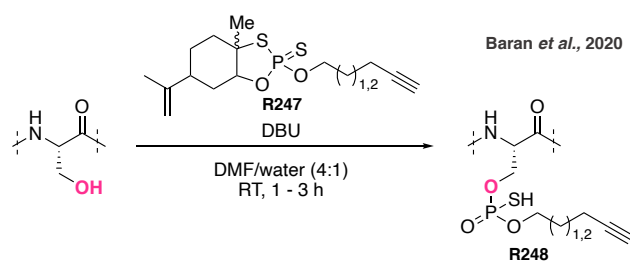**Table 10:** List of the reagents used for the site-selective modification of serine residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R248 236			50 + 50	1- 3 h / RT	20 – 40%	CuAAC	-**	yes

*serine residue in pink, functionalization site in red

** information was not provided by the authors

1.2.2.10. Aspartate and Glutamate residues

Even if aspartate and glutamate residues are quite abundant in proteins (> 5%), only few bioconjugation methods targeting them selectively have been developed.[11] The most known utilizes carbodiimides such as EDC to form activated esters.[237] However, the presence of water and nucleophilic groups on the protein compete with the nucleophile of interest for attacking the activated esters. Even though, this problem was overcame by using diazo reagents, to date, no method were proposed for the site-selective modification of aspartate and glutamate residues on native proteins.[238]

1.2.2.11. *C*-terminal residues

Based on the difference in oxidation potentials between the side chain alkyl carboxylates of aspartate and glutamate and the *C*-terminal α-amino carboxylate, Macmillan *et al.* proposed that the *C*-terminus could selectively be modified over aspartate and glutamate via single electron transfer (**Scheme 38**).[239] With the water compatible photocatalyst, flavin, the *C*-

terminus of three different proteins was efficiently modified with α,β-unsaturated carbonyls **R249** via photoredox chemistry. Even though, the method selectively targeted the *C*-terminal residues of all the peptides and proteins tested, with as little as 10 equivalents of starting material and 30 mol% of catalyst, the conversions remained medium (49% conversion) and the method was not applied to proteins bigger than 10 kDa.
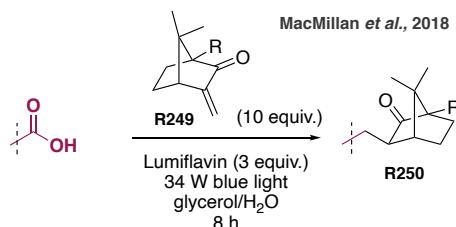
Scheme 38: Labelling of *C*-terminal residues with α,β-unsaturated carbonyls R249 via photoredox chemistry.

The reagents reported for the site-selective modification of *C*-terminal residues are presented in the following Table 11.

Table 11: List of the reagents used for the site-selective modification of *C*-terminal residues.

	Reagent	Conjugate structure*	Equiv.	Time / Temp.	Conversion	Further functionalization	Accessibility of the reagent	Site-selectivity
R249 239			10	8 h / **	49%	CuAAC	> 4 steps	yes

C-terminal residue in purple, functionalization site in red

** information regarding the temperature is not provided

1.3.   Conclusion

Two main strategies are thus coexisting for the site-specific modification of proteins. The first one relies on engineered, artificial proteins, in which the introduction, via genetic manipulation, of natural or unnatural amino acids, or peptide tags allows to access regioselectivity and produce homogeneous conjugates. In parallel, research has also focused on a second strategy aiming at the development of new regioselective bioconjugation reactions on native proteins.

Achieving regioselective modification on native proteins is not an easy task, as the reagent has to be chemoselective for a particular functional group and be able to discriminate between its multiple copies present on the protein. Over the last decade, different ways were investigated to achieve this goal. They either relied on physicochemical properties of residues

and of their microenvironment – e.g. solvent accessibility, hydrophobicity, pKa values – or on the presence of neighboring functional groups to direct the conjugation. Currently, the easiest way to achieve site-selective modification of native proteins is to target cysteine residues as they are poorly abundant in their reduced form and are highly reactive. Reaching this regioselectivity on other amino acids is more complicated and is often done at the cost of conversion, explaining why regioselective methods seem to be limited essentially to small and medium-size proteins (< 20 kDa). Thus, limitations in this field still exist, which would undoubtedly fuel the development of exciting and innovative strategies in the following years. Essential features of such strategies should be high kinetics and conversion, use of limited amounts of reagent, applicability to all sorts of proteins (from 5 kDa to > 150 kDa), and easy functionalization of the resulting adducts to introduce any given type of payload.

With the aim of pursuing the efforts in this field, different approaches for the selective modification of native proteins were attempted during this thesis (**Figure 9**). Hypervalent iodines were first explored for the labelling of cysteine residues on antibodies (Part I). Based on the native chemical ligation, we tried to achieve a regioselective labelling of lysines following a pre-conjugation on cysteines (Part II). Finally, based on multicomponent reactions, we aimed to target two vicinal amino acids simultaneously, lysine and aspartate or glutamate, in order to reduce the number of modification sites (Part III).
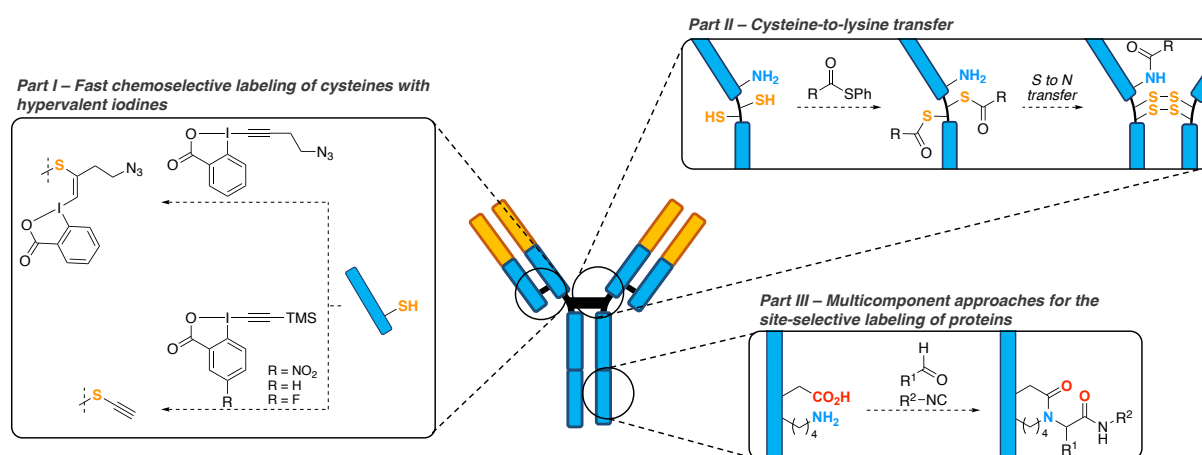
Figure 9: The three project aims

CHEMOSELECTIVE LABELLING OF CYSTEINE RESIDUES WITH HYPERVALENT IODINES

2.1. Introduction

Among the 20 natural proteogenic α-amino acids, cysteine is one of the most convenient targets for bioconjugation because of its low abundance (< 2%) and the nucleophilicity of the sulfhydryl group.[10,11,32] Various strategies and reagents have already been described and discussed for the selective labelling of cysteine residues in proteins, such as the well-known maleimide compound and its derivates, but also monosulfone or bisulfone reagents, and, dibromopyradizinediones.[33,65,240] Despite, the large number of methods already described, the need for highly reactive and chemoselective reagents in bioconjugation is still of interest. Very recently, the hypervalent iodine compounds were described as a suitable alternative to the maleimide reagent for the chemoselective labelling of cysteines.

2.1.1. Hypervalent iodines

First isolated in 1811 from the ash of seaweed by B. Courtois, iodine was named after a Greek term meaning "violet color" to illustrate the dark purple color of crystalline iodine. Present at trace levels, the iodine element is essential in many biological organisms, including humans. Over the last two centuries, iodine compounds have been extensively used in medicine, chemistry and photography to cite few.[241,242]

This chemical element is most often found in its monovalent form, where its oxidation state is -1. Amongst the common halogens, iodine is the largest atom, making it the most polarizable but also the least electronegative. Therefore, it can form stable multiple bonds with other atoms, forming compounds that belong to the class of hypervalent molecules.[242] Iodane can then adopt +3, +5 and +7 oxidation states. These hypervalent compounds hence have a central iodine linked to either three, five or seven groups, respectively corresponding to λ^3-, λ^5- and λ^7-iodanes. Employing Martin-Arduengo N-X-L designation of hypervalent molecules – where N describes the number of valence electrons around the central atom X and L the number of ligands – λ^3-iodanes can adopt two structures: 8-I-2 corresponding to iodonium salts and 10-I-3, to organic iodosyl compound, λ^5-iodanes can exist under the form of 10-I-4 and 12-I-5, and, λ^7-iodanes as 14-I-6 and 14-I-7 (**Figure 10**).[241–244]
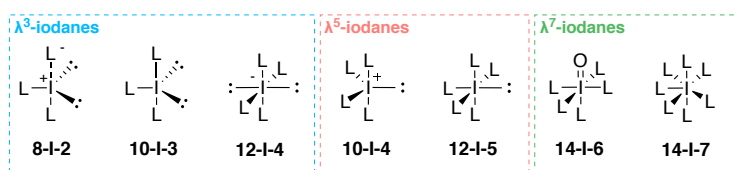
Figure 10: Structures of polyvalent iodine compounds.

The concept of "hypervalent molecules" was introduced in 1969 by Musher to define molecules containing atoms of the groups 15-18 possessing more electrons than the octet in their valence shell.[241,245] This hypervalent bonding could be explained by the valence bond model developed by Pimentel and Rundle.[246,247] According to this model, only the non-hybridized 5p orbitals are involved in the bonding, with the singly occupied equatorial 5p orbital bounding to the least electronegative ligand and giving a normal covalent bond. The interaction of the two filled axial 5p orbitals of iodine with the half-filled orbitals of the two ligands L, *trans* to each other, leads to the formation of a three-center-four-electron (3c-4e) bond (**Figure 11**). This bond contains three molecular orbitals (MO): bonding, nonbonding and antibonding. The highest occupied molecular orbital (HOMO) – i.e. nonbonding MO – has a node at the central iodine, making the bond highly polarized. Indeed, the ligands contain the high electron density while the iodine is depleted and becomes electrophilic. This behavior explains why more electronegative atoms tend to occupy those axial positions. This bond is hence called "hypervalent" and is longer than regular covalent bond. The geometry adopted by a λ^3-iodane is then pseudotrigonal bipyramid with the two most electronegative ligands in the axial position, when the least electronegative one and the two electron pairs are in equatorial positions.[244]
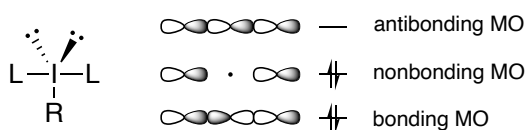
Figure 11: Molecular orbital description of the three-center-four-electron bond in hypervalent iodine (III) molecules. Adapted from: Zhdankin, V. V. *Hypervalent Iodine Chemistry: Preparation, Structure, and Synthetic Applications of Polyvalent Iodine Compounds*; John Wiley & Sons, Inc: Chichester, West Sussex, 2014.

For λ^5-iodanes molecules, a square bipyramidal structure is adopted as two hypervalent 3c-4e bonds are present and contains the four most electronegative ligands, which are then all in equatorial position, while the least electronegative substituent and the lone pair of electrons are in axial positions.

To tame their high reactivity, cyclic hypervalent iodines such as benziodoxol(on)e (BX) derivatives were described and found to be more stable than their acyclic counterpart, leading to new families of reagents: trifluoromethyl benziodoxol(on)e (Togni reagent),[248] ethynyl benziodoxol(on)e (EBX),[249,250] azido benziodoxol(on)e (ABX)[251] and cyano benziodoxol(on)e (CBX) (**Figure 12**).[252] In the context of bioconjugation, the first three BX reagents were utilized in new umpolung strategy for the functionalization of biomolecules over the last decade.

Hypervalent iodines for reactivity Umpolung – X⁺

δ^+ Y—I—X

δ^+ Y—⁺I—X

δ^+ O—I—X

Z = O, (CH$_3$)$_2$, (CF$_3$)$_2$

X = ═══R
     —CF$_3$ —N$_3$ —CN

λ^3-iodanes

iodonium salts

benziodoxol(on)es

Low stability

Enhanced stability
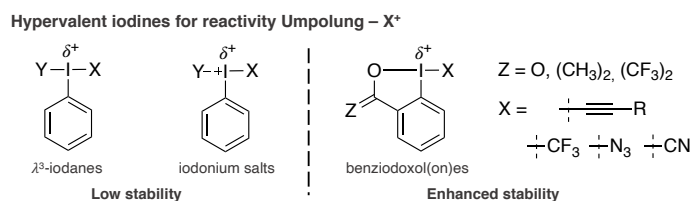
Figure 12: Reactivity of hypervalent iodines. Adapted from: Hari, D. P.; Caramenti, P.; Waser, J. Cyclic Hypervalent Iodine Reagents: Enabling Tools for Bond Disconnection via Reactivity Umpolung. *Acc. Chem. Res.* **2018**, *51* (12), 3212–3225. https://doi.org/10.1021/acs.accounts.8b00468.

2.1.2. Hypervalent iodines in bioconjugation chemistry

With more than 30 000 reactions employing hypervalent iodines described in literature, it was expected to find a plethora of bioconjugation reactions using these reagents. However, to date, only very few reactions were reported for the labelling of amino acids, peptides and proteins. Those few articles described the hypervalent iodines as coupling reagents for peptides, reagents for the photoredox catalyzed decarboxylative alkynylation of *C*-terminus or for the labelling of tryptophan, methionine and cysteine residues.

2.1.2.1   Modification of cysteine residues
2.1.2.1.1.   Fluoroalkylation of thiols

Thiols were the first nucleophiles to be trifluoromethylated with the hypervalent iodines **R251, R252, R253** and **R254** (**Figure 13**).[248,253–255]

R = ArO, PhS, *N*-Het.
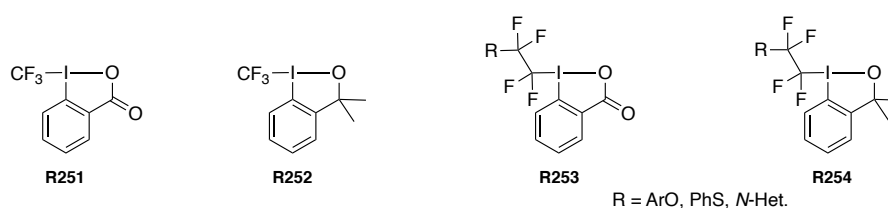
Figure 13: Structure of 1-(trifluoromethyl)-1,2-benziodoxol-3(1H)-one **R251**, trifluoromethyl-1,3-dihydro-3,3-dimethyl-1,2-benziodoxole **R252** and derivates.

The reaction of hypervalent iodines with thiols was found to be fast and selective, as a large number of functional groups – amines, amides, carboxylic acids, thioacetals, alcohols and alkynes – did not interfere in the reaction.[256] From a mechanistic point of view, the hypervalent iodine **I** was first protonated by either the thiol or the solvent (**Scheme 39**). A radical CF₃ **IV** was then generated from the homolytic cleavage of the I-CF₃ bond and a thiyl radical **III** was formed in parallel by either hydrogen transfer with the *O*-radical or by reduction of the *I* radical. Those two radicals recombined to give a sulfonium intermediate **V**. The catalytic cycle was finally closed by the protonation of a new hypervalent iodine with the sulfonium to give the final trifluoromethylated thiol **VI**.[256]

Scheme 39: Mechanistic proposal for the trifluoromethylation of thiols. Adapted from: Charpentier, J.; Früh, N.; Togni, A. Electrophilic Trifluoromethylation by Use of Hypervalent Iodine Reagents. *Chem. Rev.* **2015**, *115* (2), 650–682. https://doi.org/10.1021/cr500223h.

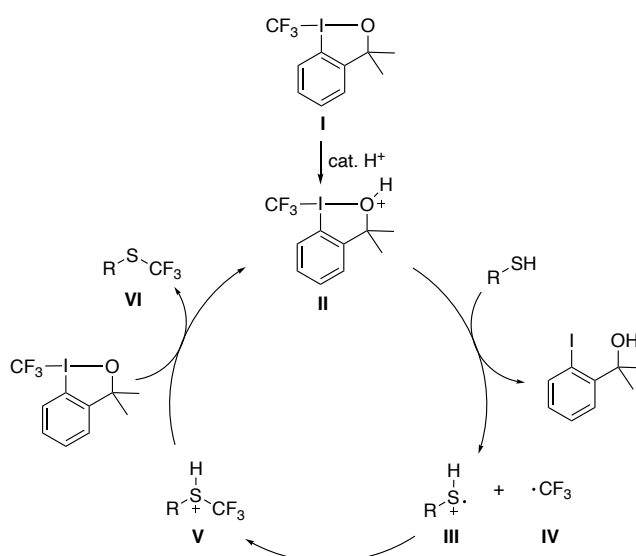As the reaction proceeded very quickly under mild conditions at low temperatures without any competing side reactions, the method was found to be promising for the functionalization of biomolecules. The introduction of fluoroalkyl groups hence gave the possibility to use ^{19}F NMR spectroscopy for the reaction monitoring and the determination of protein structure and dynamics.[257–259]

2.1.2.1.2.   Alkynylation of thiols

In 2013, Frei and Waser reported the first method for thiols alkynylation using the hypervalent iodine reagent, TIPS-ethynyl-benziodoxolone **R231** (TIPS-EBX).[260] They showed that the reaction was complete in less than five minutes in an open flask and was tolerant to a broad range of functional groups. Based on those results, the authors were later able to extend the

scope of hypervalent iodines suitable for this reaction and investigate the mechanism of this transformation (**Scheme 40**).[261] It was first proposed that the thiolate would attack the iodine on the EBX to give the intermediate **II** which would undergo a reductive elimination to give the final alkynylated product. Unfortunately, intermediate **II** was not observed in computational studies, suggesting that another mechanism was taking place. It was then proposed that the sulfhydryl group would undergo a conjugate addition to the EBX to give a vinyl benziodoxolone **III**, which would rearrange to the final thioalkyne via an α-elimination followed by a 1,2-shift. Another concerted pathway requiring less energy (10.8 kcal/mol vs 23 kcal/mol) was also identified and proposed; in which the thiolate, the iodine and the $\boldsymbol{\alpha}$-carbon of the acetylene arrange into a quasi-triangular atom structure – 3-atom transition state **V** – resulting in the direct addition of the sulfur to the $\boldsymbol{\alpha}$-carbon of the acetylene with the simultaneous cleavage of the C-I bond to give the desired alkynylated thiol.[261,262]
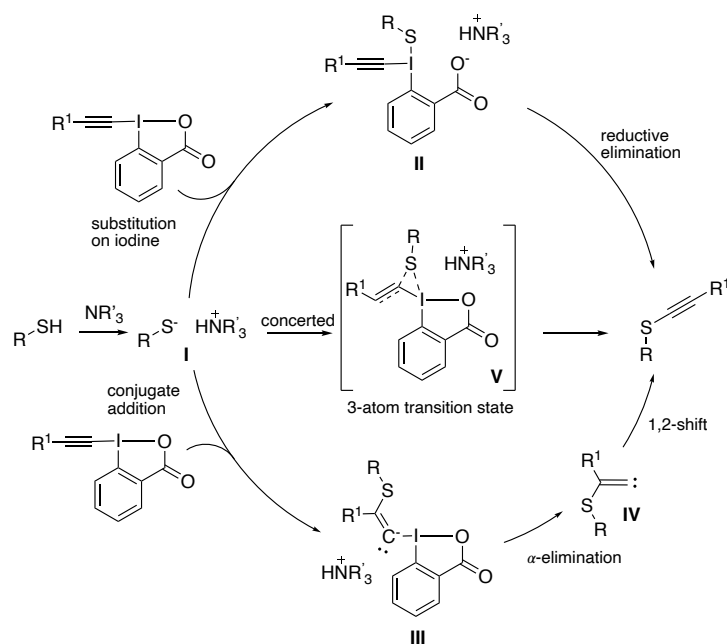
Scheme 40: Proposed mechanism for the alkynylation of thiols. Adapted from: Hari, D. P.; Caramenti, P.; Waser, J. Cyclic Hypervalent Iodine Reagents: Enabling Tools for Bond Disconnection via Reactivity Umpolung. *Acc. Chem. Res.* **2018**, *51* (12), 3212–3225. https://doi.org/10.1021/acs.accounts.8b00468.

The very fast and selective reactivity of EBX reagents for thiols under mild conditions encouraged Waser and coworkers to explore the use of EBX reagents in chemical biology. Indeed, in collaboration with the Adibekian group in 2015, they reported the use of an azide-functionalized alkynyl benziodoxolone **R46** (JW-RF-010) for the proteomic profiling of enzymes with highly reactive cysteines in living cells.[53] This azide-substituted EBX reagent **R46** was latter employed for the introduction of azido moieties into peptides and proteins (**Scheme 41**).[54] As opposed to the mechanism previously proposed where the hypervalent iodine gave access

to an alkylated thiol in organic solvent, a stable vinylbenziodoxolone hypervalent iodine conjugate **R47** was formed in water. Indeed, it was suspected that intermediate **III** would proceed through a β-addition to give a vinylic carbanion intermediate that could be protonated in water. The vinylbenziodoxolone hypervalent iodine conjugate formed in this reaction was sequentially modified by Suzuki-Miyaura cross-coupling **R254** and SPAAC **R255**, thus allowing a double functionalization of the payload. Even though the reaction was well described on peptides and histones, the method was unfortunately not applied to other proteins.
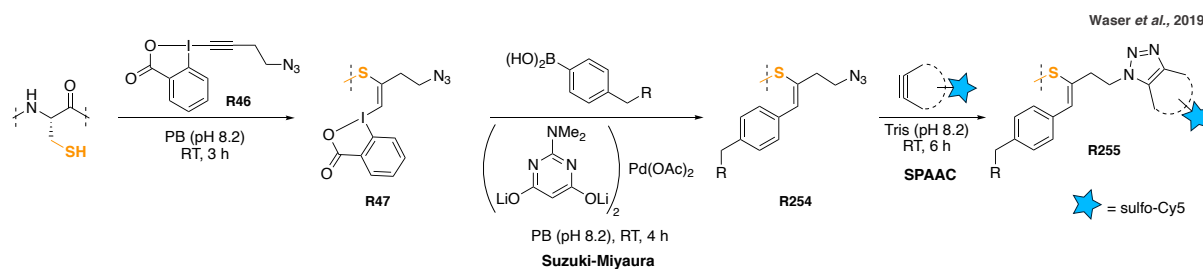
Scheme 41: Modification of cysteine residues with an azido-EBX **R46.**

2.1.2.2.   Alkynylation of tryptophan residues

After demonstrating that hypervalent iodines could modify simple nucleophiles, as thiols, Waser and coworkers showed that the direct alkynylation on indoles **R256** was also possible (**Scheme 42**).[263] Indeed, in presence of a gold(I) catalyst, AuCl, indoles were modified in good yields on the C3 position **R257**. Based on this result, in 2016, Waser *et al.* and Hansen *et al.* simultaneously proposed the same approach for the chemoselective labelling of tryptophans in peptides and proteins respectively.[226,264] Using this time AuCl(SMe)$_2$, as gold(I) catalyst, tryptophan residues were modified with TIPS-EBX **R231** on the C2 position of the indole. This strategy was explored by Hansen et *al.* on small molecules and small proteins (mellitin and apomyoglobin). Even though the reaction was found to be chemo- and regioselective on apomyoglobin with only two tryptophans modified, stoichiometric amount of catalyst was needed, and, the reaction had to be done in organic solvent (75% of acetonitrile in water), thus limiting the application of the method.
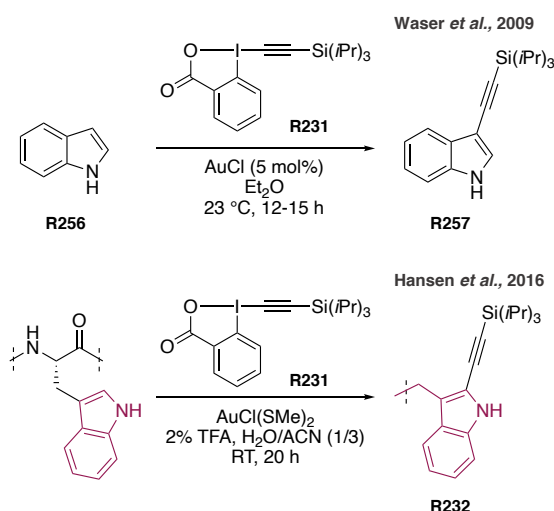
Scheme 42: Modification of indoles **R256** and tryptophan residues with TIPS-EBX **R231**.

2.1.2.3. Modification of the *C*-terminal positions

Few methods have been reported for the modification of *C*-terminus with hypervalent iodine reagents. The first one, proposed by Zhang *et al.,* showed that a direct condensation between carboxylic acids and alcohols or amines was possible, delivering esters, macrocyclic lactones, amides and peptides without racemization (**Scheme 43**).[265–267] Three generations of hypervalent iodines were described over the years to promote this transformation: iodosodilactone,   6-(3,5-bis(trifluoromethyl)phenyl)-1*H*,4*H*-2aλ³-ioda-2,3-dioxacyclopenta[*hi*]indene-1,4-dione   (*p*-BTFP-iodosodilactone)   and   1-benzoyloxy-1,2-benziodoxol-3-(1*H*)-one **R258** (IBA-OBz). Despite good yields obtained for the construction of peptides, the method was unfortunately not adapted to proteins.

In another perspective, Waser *et al*. developed a decarboxylative alkynylation on the *C*-terminus of peptides using photoredox catalysis and EBX reagents **R260**.[268,269] Instead of using classical iridium photocatalysts for their transformation, they proposed to use organic dyes, and in particular, 2,4,5,6-tetra(9*H*-carbazol-9-yl)isophtalonitrile (4CzIPN). This dye possessed the advantage of making this metal-free reaction selective for the *C*-terminus of peptides, even in presence of aspartate and glutamate residues. Despite the good conversions obtained on small peptides (< 4 amino acids), the reaction was unfortunately not effective on *C*-terminal tryptophan and tyrosine residues and was not applied to longer peptides.
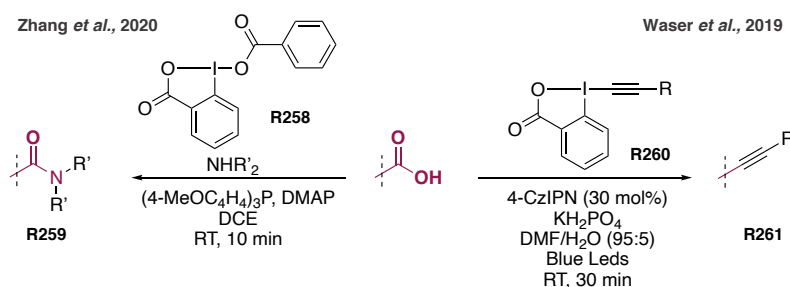
Scheme 43: Modification of *C*-terminus with hypervalent iodine reagents.

2.1.2.4. Modification of methionine residues

As opposed to cysteine, tryptophan and *C*-terminal residues, methionine residues in peptides and small proteins were found to be selectively modified with an acyclic hypervalent iodine **R243**, as shown by Gaunt and coworkers in 2018 (**Scheme 44**).[233] The use of an iodonium salt functionalized with a diazo group, thiourea, TEMPO and formic acid gave an excellent chemo- and regioselectivity on small proteins, with only one or two sites of modification being identified. The incorporation of the diazo group onto the protein allowed its further functionalization in a second step by photoredox radical cross coupling reaction with a C-4 substituted Hantzsch ester.
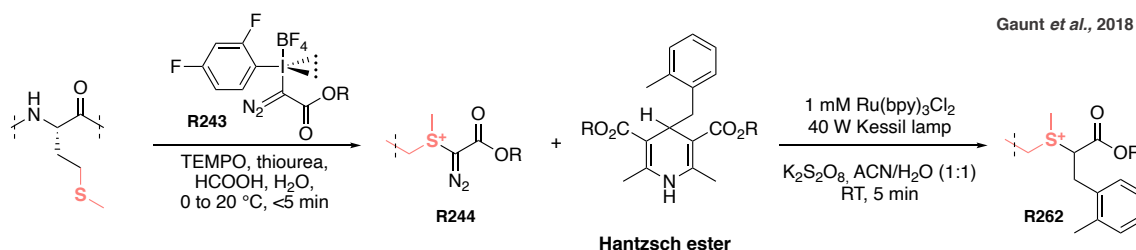
Scheme 44: Modification of methionine residues in proteins with an iodonium salt **R243**.

2.1.3. Conclusion and aim of the project

Over the last decade, hypervalent iodines have shown some attractivity in chemical biology because of their high reactivity, selectivity and stability. To date, very few methods using hypervalent iodines were proposed for the functionalization of peptides or proteins. Indeed, the modification of four residues – cysteine, tryptophan, methionine and *C*-terminal – were reported with only four types of hypervalent iodine reagents – CF$_3$-EBX, azido-EBX, ethynyl-EBX and iodonium salt. Despite being extremely efficient on short peptides, half of the methods

proposed are not transposable to proteins mostly due to the need of organic solvents. Building up on these literature precedents, and with the aim of pursuing the efforts in this field, we investigated the use of hypervalent iodines for the chemoselective modification of cysteine residues in antibodies in collaboration with the Waser group (EPFL, Switzerland). Two hypervalent iodines, azido-EBX (JW-RF-010) and TMS-EBX, previously developed by the Waser group, were thus investigated for the incorporation of vinyl azide and ethynyl groups, respectively, on cysteines (**Scheme 45**).

Part I – Chemoselective labeling of cysteine with azido-EBX

Part II – Chemoselective labeling of cysteines with TMS-EBX
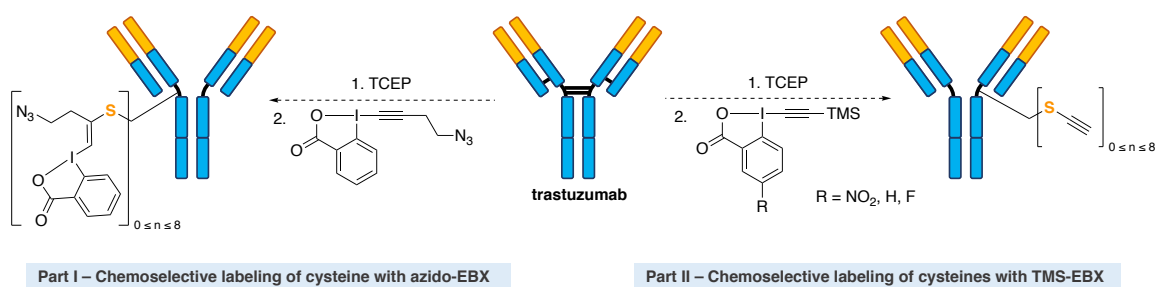
Scheme 45: Aim of the project

2.2.   Chemoselective labelling of cysteine with an azide-bearing EBX (ABX)

Following the results obtained by the Waser group on peptides, bioconjugation with 1-(4-azidobut-1-yn-1-yl)-1λ^3-benzo[*d*][1,2]iodaoxol-3(1*H*)-one **1** (ABX) was evaluated on the model antibody trastuzumab. Initial experiments were conducted with different equivalents of hypervalent iodine at pH 7.5 or 8.5 in a PBS buffer at 25 °C for 1 hour, followed by SPAAC with a bicyclonyne (BCN) strained alkyne equipped with a cyanine-5 fluorophore **37a.** Average degrees of conjugation (DoC) were first measured by fluorescence on a Nanodrop, and then confirmed by native mass spectrometry (native MS).

As shown in **Figure 14**, an average DoC of 1.9 was obtained with 5 equivalents of **1** on reduced trastuzumab. Increasing the pH or the amount of **1** only led to a slight increase of this value, up to 2.5 at pH 8.5 when 20 equivalents of **1** were employed. To evaluate the chemoselectivity of this transformation, similar experiments were also conducted on non-reduced trastuzumab. Unfortunately, labelling of the non-reduced antibody was consistently observed under all conditions tested, especially under basic conditions (pH 8.5), indicating a lack of cysteine selectivity.
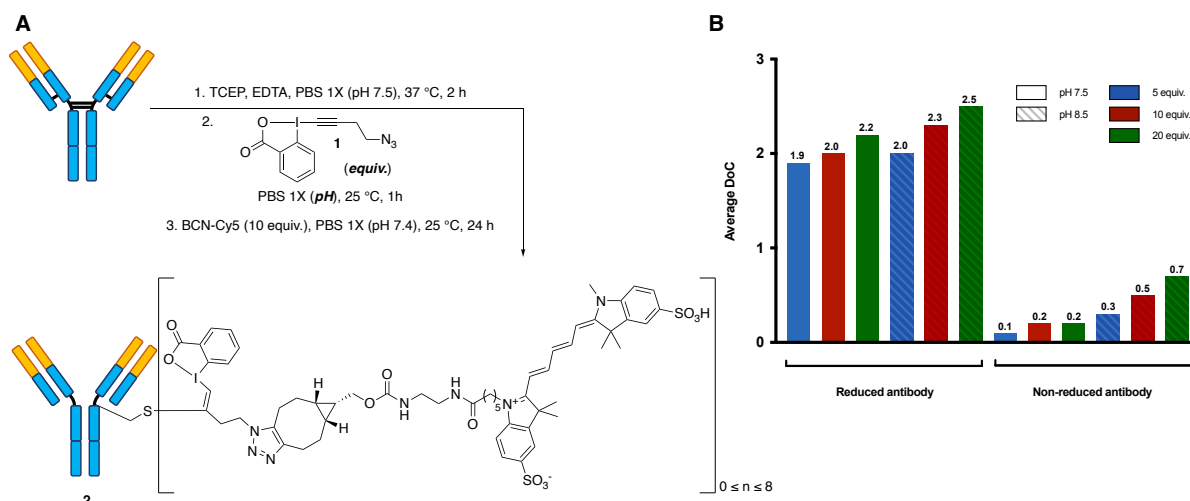
Figure 14: Labelling of trastuzumab with **1**. **A.** Bioconjugation conditions. **B.** influence of the pH and the equivalents of ABX on the average DoC values for reduced and non-reduced trastuzumab. The samples were set in a PBS 1X buffer and incubated at 25 °C for 1 hour.

As the presence of several cyanine-5 fluorophores on the antibody led to a substantial amount of protein aggregation (≥ 70%), samples were characterized by native MS prior to SPAAC, in order to determine the average DoC more accurately. Similar profiles were obtained for all the

conditions tested on reduced trastuzumab, with average DoC going from 4.0 to 5.8 with a distribution of conjugate species starting from 0 (D0; naked antibody) up to 10 (D10) when the reaction was set with 20 equivalents of **1** at pH 8.5 (25 °C, 1 h), demonstrating here again the lack of cysteine-selectivity of the reagent. Unexpectedly, additional peaks were identified on MS profiles (**Figure 15**A). They were correlated to the partial decomposition of the payload via the elimination of an iodobenzoic acid moiety **3**, with peaks correlating with a mass loss of 250 Da (D') and 500 Da (D'') from the corresponding DoC value D (**Figure 15**B). As increasing the cone voltage during the analysis did not lead to the apparition of additional peaks, this degradation was assumed to depend on the reaction conditions and not on the analytical experiments. This hypothesis was later confirmed with the apparition of more degradation peaks when the conjugate was incubated three days at 37 °C prior to the analysis by native MS (see page 137, **Figure 44** in the experimental part).
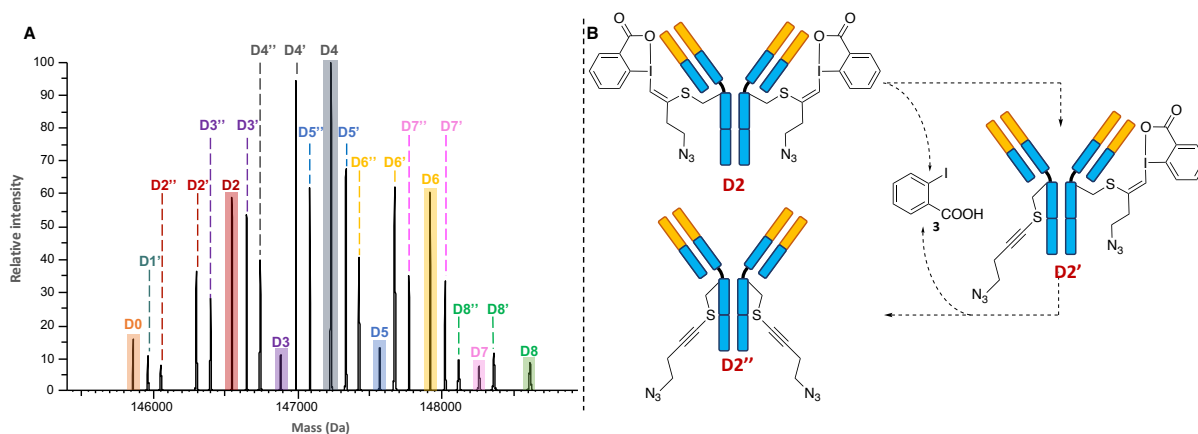
Figure 15: A. Deconvoluted mass spectrum of reduced trastuzumab labelled with 5 equivalents of **1** during 1 hour at 25 °C in PBS 1X (pH 7.5) (average DoC = 4.0). **B.** Potential rationale behind the presence of additional peaks.

To improve the thiol-selectivity of the reagent and prevent the degradation of the payload formed, we engaged into a second round of optimization. Two strategies were then envisioned in order to obtain more homogeneous conjugates: either varying the number of equivalents of **1**, or changing the incubation time (5, 15 or 30 min) (**Figure 16**). While, non-specific labelling was still observed after 15 and 30 minutes of incubation with 8 equivalents of **1**, decreasing the incubation time to 5 minutes had minimal interference on non-reduced antibody. Applied to reduced antibody, these conditions – 8 equivalents of **1** in PBS 1X (pH 7.5), 25 °C, 5 min – led to a diminution of the conjugated species (from D0 to D3) observed by native MS. Despite being less important than for the previous experiment (limited only to D'), the degradation of the payload was still present.
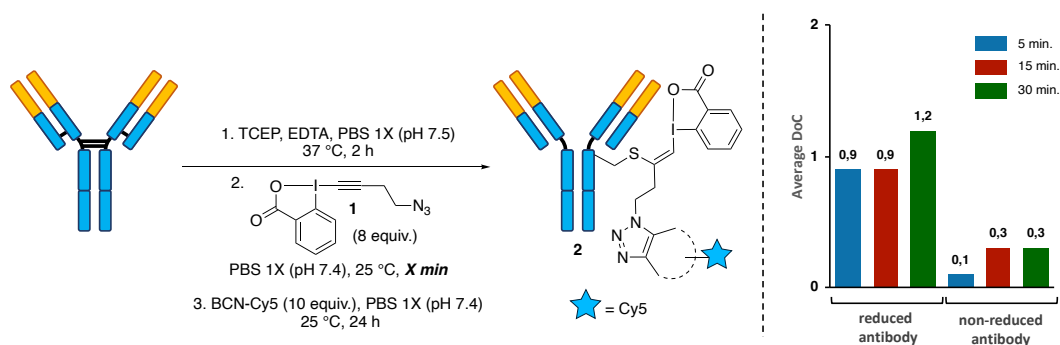
Figure 16: Labelling of trastuzumab with **1**. Influence of the incubation time on the average DoC values for reduced and non-reduced trastuzumab. The samples were set in a PBS 1X buffer (pH 7.5) with 8 equivalents of **1** and incubated at 25 °C.

Unfortunately, tuning the reactions conditions could neither prevent the degradation of the payload nor suppress the non-specific labelling on non-reduced trastuzumab. The use of the ABX reagent **1** for the modification of cysteine residues in antibodies was not investigated further. Switching to a structurally related yet different reagent, the reactivity and selectivity of TMS-bearing EBX was evaluated for the modification of cysteine residues.

## 2.3. Ethynylation of cysteine residues in antibodies with a TMS-bearing EBX.

2.3.1. Alkynylation of thiols with pNO2-EBX 4.

As previously done with the ABX reagent **1**, bioconjugation with 5-nitro-1-((trimethylsilyl)ethynyl)-1λ^3-benzo[*d*][1,2]iodaoxol-3(1*H*)-one **4** (*p*NO$_2$-EBX) was also evaluated on the model antibody trastuzumab. Initial experiments were conducted with different equivalents of **4** at pH 7.5 or 8.5 in a PBS buffer (1X) at 25 °C for 30 minutes. Because of its low molecular weight, the alkyne payload had to be further functionalized via copper-catalyzed alkyne-azide cycloaddition (CuAAC) reaction with an azide-containing TAMRA fluorophore **5** to allow detection by mass spectrometry and analysis by fluorescence spectroscopy. To evaluate the thiol-selectivity of the reagent, experiments were again performed on both reduced and non-reduced antibody.

Pleasingly, the first results showed that the reactions done on non-reduced trastuzumab gave almost no fluorescence signal, suggesting an excellent cysteine-selectivity of the *p*NO$_2$-EBX reagent **4**. However, important precipitation of the antibody ($\geq 80\%$) was observed upon CuAAC. Suspecting that this could be due to the large amount of copper used, we decided to repeat the reaction conditions with lowered amounts of copper sulfate CuSO$_4$ (1, 5 or 10 equivalents), tris-hydroxypropyltriazolylmethylamine THPTA (2, 10 or 20 equivalents) and sodium ascorbate (3, 15 or 30 equivalents). Whilst varying the amount of ascorbate had no impact on the antibody precipitation, the amount of copper and ligand was critical. With 5 or 10 equivalents of copper, nearly 20% of the antibody precipitated in 15 minutes, whereas no antibody precipitation was observed after the 20 hours of reaction necessary to reach completion when 1 equivalent of CuSO$_4$ and 2 equivalents of THPTA were used. Several washing at the end of the reaction with a buffer containing 1% of ethylenediaminetetraacetic acid (EDTA) was also found to be crucial to remove the residual copper in conjugates and obtain interpretable MS spectra.

Despite these optimized conditions developed for the CuAAC reaction, intricate native MS spectra were still obtained, suggesting a partial decomposition of the antibody or side reactions occurring during the ethynylation step (**Table 12**). Unfortunately, varying the buffer composition or decreasing the incubation time, the number of equivalents of **4**, or the reaction temperature, did not prevent the antibody degradation, urging us to switch to the less reactive EBX reagents, *p*H-EBX **7** and *p*F-EBX **8**.

Table 12: Conjugation of trastuzumab with *p*NO₂-EBX **4**.
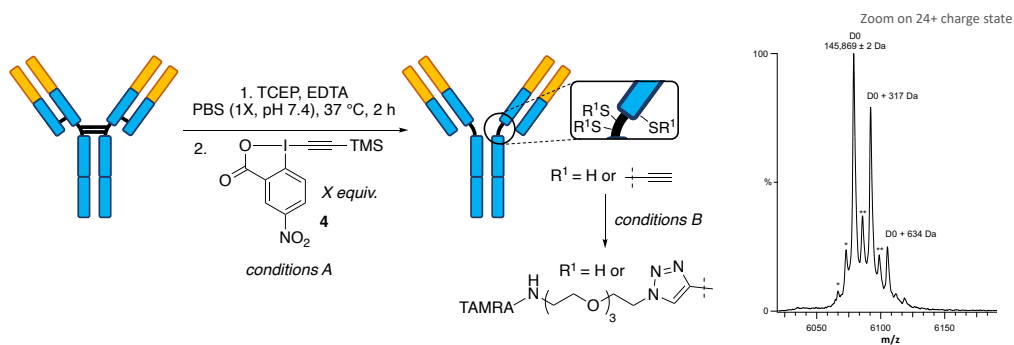
Entry	Conditions A	MS spectrum
1	20 equiv., PBS buffer, pH 8.5, 25 °C, 30 min	Degradation
2	20 equiv., PBS buffer, *pH 7.5*, 25 °C, 30 min	Degradation
3	20 equiv., PBS buffer, pH 7.5, 25 °C, *5 min*	Degradation
4	*8 equiv.,* PBS buffer, pH 8.5, 25 °C, 30 min	Degradation
5	8 equiv., PBS buffer, *pH 7.5*, 25 °C, 30 min	Degradation
6	8 equiv., PBS buffer, pH 7.5, 25 °C, *5 min*	Degradation
7	8 equiv., PBS buffer, pH 7.5, *37 °C*, 5 min	Degradation
8	*2 equiv.,* PBS buffer, pH 7.5, 25 °C, 30 min	Degradation
9	2 equiv., PBS buffer, pH 7.5, 25 °C, *5 min*	Degradation
10[a]	2 equiv., *PB buffer (200 mM), pH 8.2, 37 °C, 15 min*	Degradation
11	2 equiv., PB buffer (200 mM), pH 8.2, 37 °C, *5 min*	Degradation
12	2 equiv., PB buffer (200 mM), pH 8.2, 5 °C, *5 min*	Degradation
13	2 equiv., PB buffer (200 mM), pH 8.2, *5 °C*, 15 min	Degradation

[a] Conditions developed by the Waser group on peptides.
Conditions B: TAMRA-N₃ **5** (10 equiv.), CuSO₄ (1 equiv.), THPTA (2 equiv.), sodium ascorbate (3 equiv.), PBS (1X, pH 7.5), 25 °C, 24 h.

2.3.2. Alkynylation of thiols with *p*H-EBX **7** and *p*F-EBX **8**.

Following conditions developed by the Waser group on peptides (phosphate buffer, pH 8.2, 37 °C, 5 to 15 min), bioconjugation with p*H*-EBX **7** was also evaluated on the model antibody trastuzumab. As already observed with **4**, **7** showed some side reactivity on non-reduced trastuzumab under these conditions. However, it was possible to obtain clean native MS spectra this time, allowing us to determine average DoCs and conversion rates (**Figure 17**A).

Incubating reduced trastuzumab with 2 equivalents of **7** during 5 or 15 minutes had no influence on the average DoC and conversion rate (table 2, entry 1-2). However, increasing the number of equivalents from 2 to 8 and up to 16 gave increased average DoC – from 0.7 to 3.7 and 4.4 – and conversion rates – from 39% to 94 and 97% (entry 2-4). Switching to a phosphate-buffered saline (PBS) solution had a limited impact on the reaction (entry 5). On

the contrary, influences of pH and temperature were found to be critical on the conversion rates but also on the side-reactivity of the reaction. Poor reactivity was observed at pH 6.5, resulting in low average DoC and conversions compared to pH ≥ 7.5, with the highest reactivity found at pH 8.5 (entry 6-8). As expected, decreasing the incubation temperature to 25 °C diminished both the average DoC and the conversion values (entry 9), but it also helped in suppressing the side-reactivity observed on non-reduced trastuzumab. Finally, reducing the reaction time to 2 minutes allowed us to recover intact, non-reduced trastuzumab. Applied to reduced trastuzumab, these conditions gave 60% conversion and delivered conjugates with an average DoC of 1.2 (entry 10), and were deemed to be optimal for this chemoselective transformation. Interestingly, conjugation could still be observed after only 1 minute or at lowered temperatures, highlighting the high reactivity of EBX reagents for antibody modification (entry 11-14).

To pursue our study on hypervalent iodine reagents, the reactivity of *p*F-EBX **8** on cysteine residues was investigated (**Figure 17**B). To our delight, a similar pattern of reactivity was observed with the less reactive *p*F-EBX **8** reagent. Indeed, similar conversion and average DoC values were obtained under the same conditions tested. As observed for **7**, no side-reactivity on non-reduced trastuzumab was detected when the conjugation was performed at either 5 or 25 °C for 1-2 minutes, or, at pH 6.5 (entry 10-14 and 4). Interestingly, the reaction was also found to be chemoselective at 37 °C when 8 equivalents of **8** was incubated for 1 or 2 minutes, and, delivered conjugates with an average DoC of 0.6 and 2.1 respectively (entry 7-8).

Finally, the Adibekian group at the Scripps Research Institute evaluated different TMS-EBX reagents for the proteome-wide labelling of cysteines on living HeLa cells. They showed that with a fine tuning of EBX reagents, an improved cysteinome coverage was obtained compared to the iodoacetamide reagent – the current gold standard – while maintaining a high selectivity towards cysteines (90.2% of cysteines labeled relative to other nucleophilic amino acids). These results highlighted the potential of hypervalent iodine reagents as cysteine-reactive probes in chemoproteomics.

In summary, we, along with Waser and Adibekian groups, reported a new procedure for the ethynylation of cysteine residues of the monoclonal antibody trastuzumab, using TMS-EBX reagents.[55] We highlighted the strong influence of these reagents' stereoelectronic properties on the outcome of the bioconjugation reaction. In particular, two EBX reagents, *p*H-EBX **7** and

*p*F-EBX **8**, were shown to give access to good conversion and average DoC values with an apparent high chemoselectivity for thiols, as demonstrated by an absence of conjugation on non-reduced antibody under the optimized conditions.

2.4. Conclusion

With the aim of evaluating two new reagents for the labelling of cysteine residues, the reactivity of two classes of hypervalent iodine compounds, azido-EBX **1** and TMS-EBX, was investigated on the antibody trastuzumab. While ABX reagent **1** was found to be reactive on the antibody, two major issues were encountered: an apparent lack of thiol-selectivity, as demonstrated by the labelling of non-reduced trastuzumab, and the degradation of the newly-formed payload leading to a highly heterogeneous mixture of conjugates. As it was not possible to overcome those issues by tuning the reaction conditions, the ABX reagent **1** was not further investigated for the modification of cysteine residues. Switching to a structurally related yet different reagent, the reactivity and selectivity of TMS-EBX reagents was then evaluated. Using less reactive TMS-EBX reagents, *p*H-EBX **7** and *p*F-EBX **8**, allowed the smooth conjugation of trastuzumab with good conversions and average DoCs values. Most notably, under our optimized conditions, those reagents showed a perfect chemoselectivity.

Having developed a new bioconjugation method suitable for the fast labelling of cysteines in antibodies, we aimed in the next chapter to achieve a regioselective labelling of lysines with the help of a pre-conjugation of cysteines residues.

A
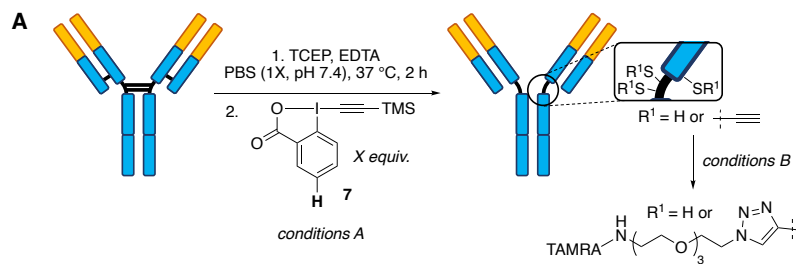
Entry	Conditions A	Side-reactivity	Av. DoC	Conv. (%)
1	2 equiv., PB buffer, pH 8.2, 37 °C, 15 min	Y	0.6	36
2	2 equiv., PB buffer, pH 8.2, 37 °C, *5 min*	Y	0.7	39
3	*8 equiv.*, PB buffer, pH 8.2, 37 °C, 5 min	Y	3.7	94
4	*16 equiv.*, PB buffer, pH 8.2, 37 °C, 5 min	Y	4.4	97
5	8 equiv., *PBS buffer*, pH 8.2, 37 °C, 5 min	Y	4.0	96
6	8 equiv., PBS buffer, *pH 6.5*, 37 °C, 5 min	N	0.9	50
7	8 equiv., PBS buffer, *pH 7.5*, 37 °C, 5 min	Y	3.4	93
8	8 equiv., PBS buffer, *pH 8.5*, 37 °C, 5 min	Y	4.4	97
9	8 equiv., PBS buffer, pH 7.5, *25 °C*, 5 min	Y	1.8	74
10	8 equiv., PBS buffer, pH 7.5, 25 °C, *2 min*	N	1.2	60
11	8 equiv., PBS buffer, pH 7.5, 25 °C, *1 min*	N	0.7	37
12	8 equiv., PBS buffer, pH 7.5, *5 °C*, 5 min	N	0.1	13
13	8 equiv., PBS buffer, pH 7.5, 5 °C, *2 min*	N	0	0
14	8 equiv., PBS buffer, pH 7.5, 5 °C, *1 min*	N	0	0

Conditions B: TAMRA-N$_3$ **5** (10 equiv.), CuSO$_4$ (1 equiv.), THPTA (2 equiv.), sodium ascorbate (3 equiv.), PBS (1X, pH 7.5), 25 °C, 24 h.

B
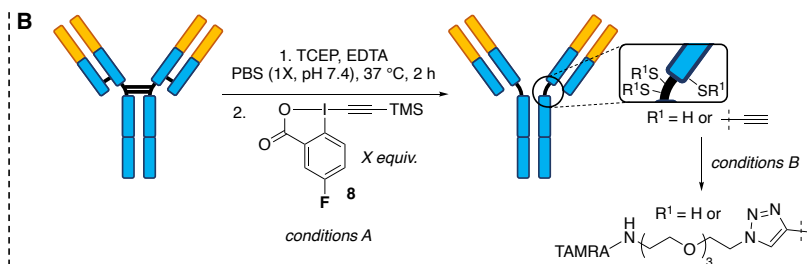
Entry	Conditions A	Side-reactivity	Av. DoC	Conv. (%)
1	2 equiv., PB buffer, pH 8.2, 37 °C, 15 min	Y	0.4	27
2	2 equiv., PB buffer, pH 8.2, 37 °C, *5 min*	Y	0.4	28
3	*8 equiv., PBS buffer*, pH 8.2, 37 °C, 5 min	Y	2.9	88
4	8 equiv., PBS buffer, *pH 6.5*, 37 °C, 5 min	N	1.3	62
5	8 equiv., PBS buffer, *pH 7.5*, 37 °C, 5 min	Y	2.7	88
6	8 equiv., PBS buffer, *pH 8.5*, 37 °C, 5 min	Y	3.0	90
7	8 equiv., PBS buffer, pH 7.5, 37 °C, *2 min*	N	2.1	78
8	8 equiv., PBS buffer, pH 7.5, 37 °C, *1 min*	N	0.6	48
9	8 equiv., PBS buffer, pH 7.5, *25 °C*, 5 min	Y	2.3	81
10	8 equiv., PBS buffer, pH 7.5, 25 °C, *2 min*	N	1.1	55
11	8 equiv., PBS buffer, pH 7.5, 25 °C, *1 min*	N	0.5	32
12	8 equiv., PBS buffer, pH 7.5, *5 °C*, 5 min	N	0	0
13	8 equiv., PBS buffer, pH 7.5, 5 °C, *2 min*	N	0	0
14	8 equiv., PBS buffer, pH 7.5, 5 °C, *1 min*	N	0	0

Conditions B: TAMRA-N$_3$ **5** (10 equiv.), CuSO$_4$ (1 equiv.), THPTA (2 equiv.), sodium ascorbate (3 equiv.), PBS (1X, pH 7.5), 25 °C, 24 h.

Figure 17: Conjugation of trastuzumab with p*H*-EBX **7** and p*F*-EBX **8**.
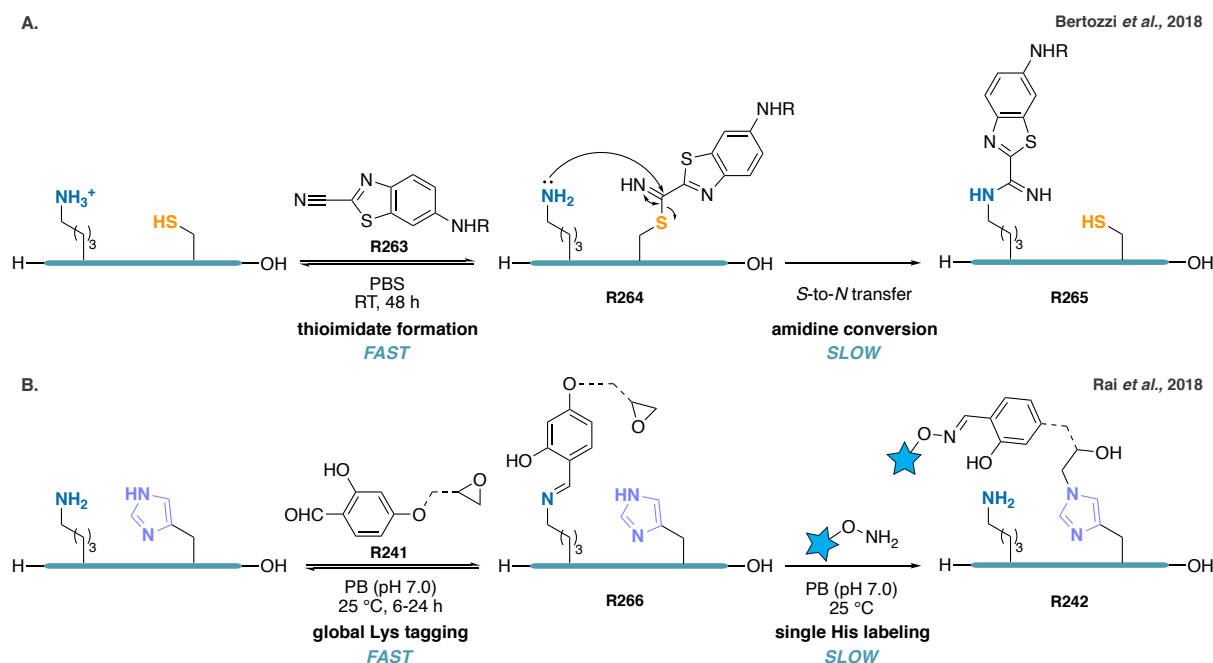
CYSTEINE-TO-LYSINE TRANSFER

3.1.   Introduction

3.1.1. Concept of an amino acid relay for the regioselective labelling of proteins

In 2018, with the aim of developing a site-selective strategy for the modification of proteins, the idea of using a more reactive or a more abundant amino acid as a "relay" for the labelling of a vicinal amino acid has emerged. Bertozzi and coworkers were the first ones to develop the concept of a cysteine-lysine relay.[270] The authors showed that the nucleophilic cysteine thiol of an 11-mer peptide was able to react rapidly and reversibly with the nitrile group of 2-cyanobenzothiazole (CBT) derivatives **R263** to a form a thioimidate intermediate **R264** that was then slowly transferred to a vicinal lysine residue via a final, irreversible *S*-to-*N* shift (**Figure 18**A).

Instead of using a poorly abundant amino acid, like cysteine, as a "relay amino acid", Rai and coworkers proposed to adopt a global lysine tagging strategy for the subsequent modification of a single histidine residue (**Figure 18**B).[231] To do so, reagents bearing two reactive groups **R241** were designed: on the one end, an aldehyde reacting with multiple solvent accessible lysines to give imines at pH 7.0 in a reversible manner; on the other end, an epoxide reacting slowly – but irreversibly – with histidine residues selectively. At the end of the reaction, the imines were hydrolyzed to free the lysine residues. As the two reactive handles are separated with a linker, regioselectivity and protein selectivity could be accessed by adjusting the size of this spacer.

Figure 18: The "relay amino acid" strategy. **A.** Cysteine to lysine transfer strategy proposed by Bertozzi *et al.*[270] The cysteine attacks an electrophile **R263** and then transfers it to a lysine nearby via a *S*-to-*N* transfer. **B.** Lysine to histidine transfer strategy proposed by Rai *et al.*[231] A global and reversible labelling of lysine residues directs the selective labelling of a histidine in a second step.

Inspired by this concept, we thought it could be possible to achieve a site-selective modification of lysine residues on antibodies, via a cysteine relay, by using a variation of the native chemical ligation.

3.1.2. The native chemical ligation (NCL)

First described by Kent and coworkers in 1994 for the direct synthesis of interleukin IL-8, the native chemical ligation (NCL) is commonly used nowadays for the construction of long peptides or proteins from shorter peptides.[271,272] This reaction employs a *N*-terminal cysteine **R268** and a *C*-terminal thioester **R269**, reacting together in a *trans*-thioesterification reaction to give a new thioester bond linking the two peptides **R270** (**Scheme 46**). Intramolecular nucleophilic attack of the α-NH$_2$ to the thioester results in a *S*-to-*N* acyl transfer, via a 5-membered ring intermediate, leading to the formation of a new peptide bond **R271**, in a traceless manner.[273,274]
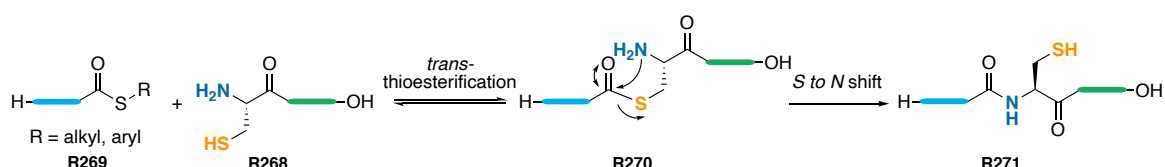
Scheme 46: Mechanism of the native chemical ligation (NCL). A *trans*-thioesterification is taking place when the *N*-terminal cysteine **R268** on a peptide attacks the *C*-terminal thioester **R269** on another. An irreversible *S*-to-*N* shift occurs in a last step and leads to the formation of an amide bond **R271**.

The NCL was found to be chemoselective for the ligation of two peptides, high yielding, tolerant to aqueous conditions (neutral pH) and to the twenty proteogenic amino acids in the *C*-terminal position, even though residues with bulky side-chains – valine, isoleucine, proline – have been correlated with slower reactions.[273] The reactivity of the thioester was also found to be crucial in this reaction. Indeed, aryl thioester should be privileged over alkyl thioester as they are better leaving groups, thus making the *C*-terminus more reactive.[273] When only alkyl thioester can be accessed in *C*-terminus, addition of an aryl thiol, such as thiophenol, to the reaction mixture favors the formation of a more reactive thioester *in situ* via thiol exchange and helps to increase the rate of the reaction.[275]

While NCL is considered the reaction of choice for the chemical synthesis of peptides and proteins, the need for a *N*-terminal cysteine was often considered to be the main limitation of the method. Indeed, because cysteines are poorly abundant in proteins and randomly distributed throughout a peptide sequence, the use of the NCL often requires the incorporation of additional cysteines at precise location. To overcome this limitation, it was possible to desulfurize the additional cysteines introduced to give an alanine residue with Raney nickel. Based on this strategy, it was later proposed to incorporate thiols on other amino acids – such as valine, alanine or phenylalanine – allowing then the use of alternative *N*-terminal residues in the NCL.[276–279] A final desulfurization allowed the generation of natural amino acids at the end of the process, a synthetic advancement that helped NCL to revolutionize further the chemical synthesis of peptides and proteins.[280,281]

3.1.3. Aim of the project

Inspired by these precedents, we wondered if an analogous NCL approach, applied to an antibody, could be regioselective, as only the lysine residues close to the eight solvent accessible cysteines would participate in the conjugation. In theory, after reduction of the disulfide bonds of the antibody, the free thiols could react with a thioester-containing molecule

through a *trans*-thioesterification reaction resulting in the formation of a new thioester. The nucleophilic attack of the ε-NH$_2$ of a vicinal lysine would result in a *S*-to-*N* acyl transfer to give a final amide compound. The remaining thioesters, that would not have participated in the reaction, would be hydrolysed, allowing in a final step the reformation of the disulfide bonds of the antibody.
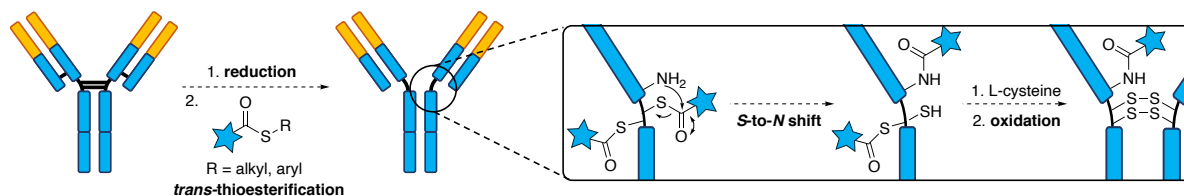

Figure 19: Application of the native chemical ligation for the site-selective modification of antibody.

3.2. Cysteine-to-lysine transfer strategy for the site-selective labelling of antibodies.

To apply NCL to the model antibody, trastuzumab, five steps were necessary, which needed optimization. In a first step, the four accessible disulfide bonds of the antibody needed to be reduced, allowing the *trans*-thioesterification reaction between the thiols of the antibody and a thioester-containing molecule to occur in a second step. In a third step, the nucleophilic attack of the side-chain amine of a vicinal lysine would result in a *S*-to-*N* acyl transfer, yielding an amide bond. Thioesters that did not partake in the transfer would have to be hydrolyzed in a fourth step in order to reconstruct the disulfide bonds of the antibody in a fifth step. Each of the five steps was optimized independently before all the steps were applied sequentially on the full antibody.

To facilitate reaction monitoring by fluorescence spectroscopy, we opted to use a TAMRA-thioester **9**, which was easily accessed from free TAMRA **10** and thiophenol **11**, reacted in presence of DCC.
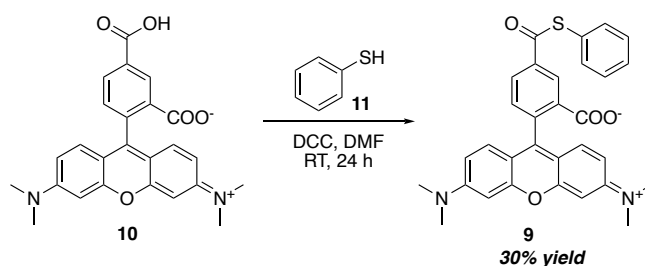
Scheme 47: Synthesis of the TAMRA-thioester **9**.

3.2.1. Optimization of the five steps of NCL bioconjugation reaction on trastuzumab.

3.2.1.1. Steps one and five: reduction and re-oxidation of the antibody.

As complete reduction of the four antibody's interchain disulfide bonds was necessary to free cysteines, trastuzumab was reacted with 10 equivalents of the reducing agent *tris*(2-carboxyethyl)phosphine (TCEP), in presence of ethylenediaminetetraacetic acid (EDTA, 1% v/v in PBS 1X (pH 7.5)) for 2 hours at 37 °C. Complete re-oxidation to the full antibody was achieved after 3 hours of incubation at 25 °C with 15 equivalents of L-dehydroascorbic acid (**Figure 20**).
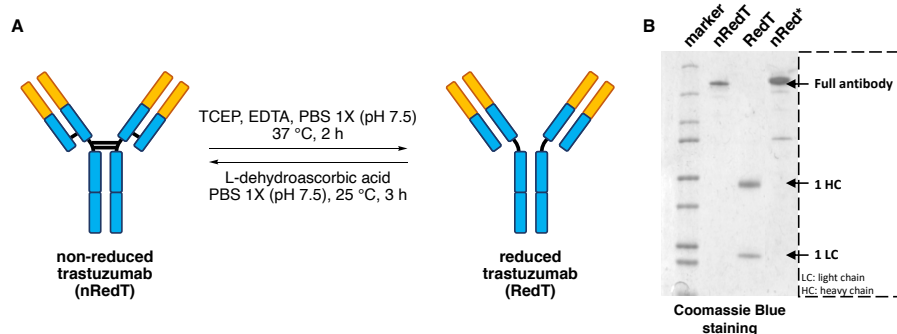
Figure 20: A: Conditions for the reduction and the oxidation of the disulfide bonds of the antibody trastuzumab. **B:** SDS page gel. Aliquots were taken before the reduction of trastuzumab (nRedT), after the reduction of trastuzumab (RedT) and after the reoxidation of its disulfide bonds (nRed*). The gel was stained with Coomassie blue.

3.2.1.2. Step two: *trans*-thioesterification reaction.

The *trans*-thioesterification step was initially evaluated with different equivalents of **9** and thiophenol **11** – to potential increase the kinetics of the thiol exchange – at 25 °C in a phosphate buffer (pH 7.5) and the reaction was stopped at different time intervals. To evaluate the thiol-selectivity of the reagent, experiments were performed on both reduced and non-reduced

antibody. Average DoC were first evaluated by fluorescence on a Nanodrop and later validated by native mass spectrometry (native MS).

The first results showed that the reactions done on non-reduced trastuzumab gave fluorescent signal when 10 equivalents of **9** were incubated at 25 °C for more than 15 minutes (DoC ≥ 0.2) (**Figure 21**). Decreasing the number of equivalents to 5 and the incubation time to 5 or 15 minutes prevented the side-reactivity of the reagent (DoC ≤ 0.1) on non-reduced trastuzumab. Applied to reduced trastuzumab, the conditions gave access to fluorescent conjugates with an average DoC of 1.1 or 1.2 for 5 or 15 minutes of incubation at 25 °C, respectively. Incubating reduced trastuzumab with **9** longer did not have a profound effect on DoC values (1.3 after 15 minutes of incubation). In the first experiments, thiophenol **11** was added to increase the kinetics of the thiol exchange, additional experiments later demonstrated that this additive had no influence on the reaction efficiency. The *trans*-thioesterification reaction on reduced trastuzumab was then found to be optimal with 5 equivalents of **9** incubated at 25 °C for 15 min in a PBS 1X (pH 7.5), with minimal side-reactivity on non-reduced antibody.
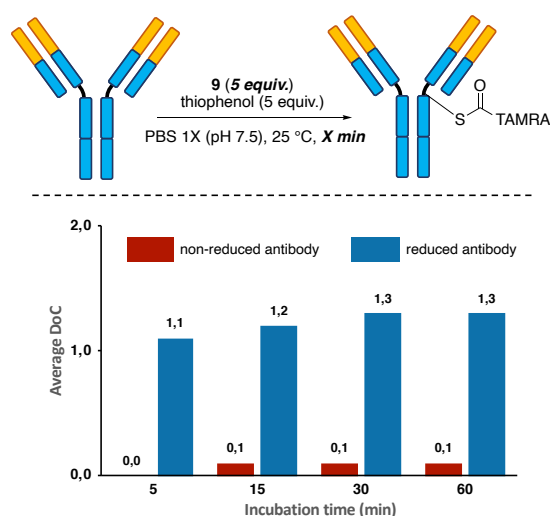
Figure 21: Optimization studies for the *trans*-thioesterification reaction. Influence of the incubation time on the average DoC values for reduced and non-reduced trastuzumab.

3.2.1.3.   Step three: *S*-to-*N* acyl shift.

After having optimized the reaction conditions for the *trans*-thioesterification step, conditions for effective *S*-to-*N* shift were then investigated. To establish which conditions led to the best transfer, the fluorescence intensity of the full antibody was evaluated by SDS page, which was also a good indication on the amount of full antibody recovered.

Switching to more alkaline conditions (pH 8.5) and increasing the incubation temperature to 37 °C, resulted in an increased rate of transfer. Indeed, most of the antibody was fully reconstructed and fluorescent by SDS-page, demonstrating that the transfer took place. Leaving the reaction for additional days (2 or 5 days) at 37 °C did not improve the transfer efficiency.

As the antibody could not be completely re-oxidized, due to the potential presence of remaining, unreacted thioesters blocking the reformation of the disulfide bonds, conditions to hydrolyze these thioesters were investigated.

3.2.1.4. Step four: removing of the remaining thioesters.

Two strategies were explored for the getting rid of the remaining thioesters. A first one relied on the use of various nucleophiles, such as hydroxylamine **12**, hydrazine **13** and guanidine **14**, to trigger an intramolecular *S*-to-*N* transfer; a second one relied instead on the use of a free L-cysteine **15** in order to mimic the NCL – thiol exchange followed by a intramolecular *S*-to-*N* shift.

After the formation of TAMRA-thioesters on trastuzumab, 50 to 730 equivalents of **12**, **13**, **14** or **15** were added, and the samples were incubated at 37 °C for 24 hours in a PBS 1X buffer (pH 8.5) before being purified by size exclusion chromatography. After 24 hours, DoC values had decreased with all the candidates tested (**Figure 22**). Addition of hydroxylamine **12** and hydrazine **13** led to a decrease in DoC values of 50 to 65%, when guanidine **14** led only to a diminution of 30%. Incubating the conjugates an additional day with the three candidates did not improve this step further. Fortunately, addition of L-cysteine **15** resulted in a complete disappearance of thioesters, as demonstrated by an absence of fluorescence after 24 hours of incubation at 37 °C. This result was also confirmed by SDS-page analysis, which also showed that the full antibody was completely recovered even without using L-dehydroascorbic acid for the reconstruction of the disulfide bonds.

In conclusion, free thiols of the antibody were completely recovered after incubating the thioester conjugate with 730 equivalents of **15** in a saline phosphate buffer (pH 8.5) at 37 °C for 24 hours, which concluded our optimization efforts.
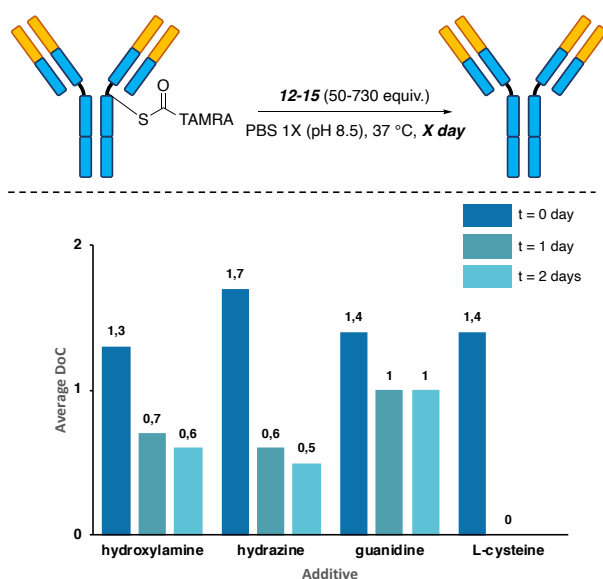
Figure 22: Optimization studies for the removing of the thioesters on the antibody trastuzumab. Influence of the additive and the incubation time on the removing of the thioesters. The efficiency of the reaction was determined with the evolution of the average DoC values.

3.2.2. NCL reaction on trastuzumab.

After optimizing each step individually, the full NCL sequence was applied to our model antibody, trastuzumab (**Figure 23**). After complete reduction of the four disulfide bonds of the antibody with TCEP and EDTA (37 °C, 2 h), the thiol exchange with **9** was conducted under our optimized conditions – i.e. 15 minutes at 25 °C, 8 equivalents of **9**. The pH was then increased to 8.5 in order to perform the *S*-to-*N* acyl shit and the conjugates were incubated at 37 °C for 24 hours. Thioesters that did not partake in the transfer were reacted with an excess of L-cysteine **15** over 24 hours at 37 °C, thus regenerating the antibody's free thiols and allowing their oxidative rebridging with L-dehydroascorbic acid (in PBS 1X (pH 7.5), 25 °C, 3 h). At the end of the reaction, an average DoC of 1.5 was determined by native MS with a DoC distribution going from 0 to 5 (**Figure 23**A).
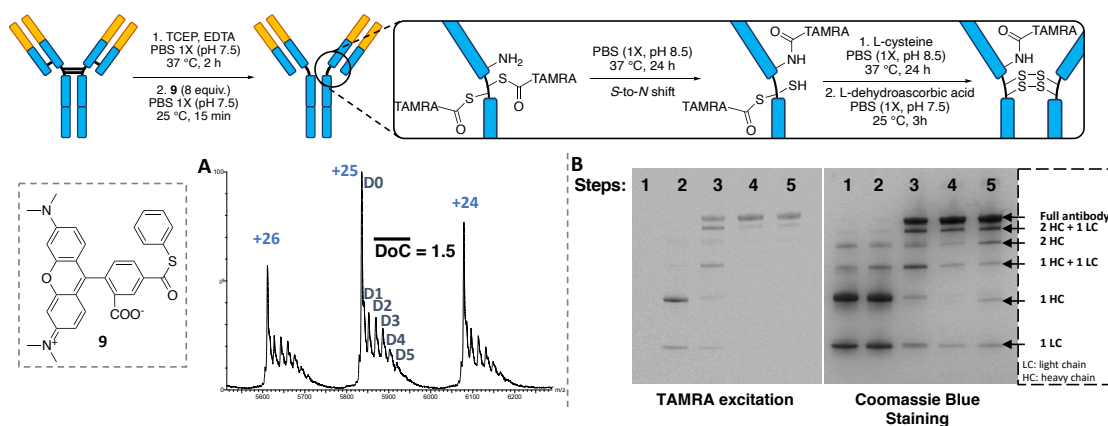
Figure 23: Application of the native chemical ligation on trastuzumab. **A**: deconvoluted mass spectrum of trastuzumab labelled with 8 equivalents of **9** after the NCL sequence (average DoC = 1.5). **B**: SDS page gel. Aliquots were taken after each step and analyzed by SDS page. The gel was revealed by fluorescence and Coomassie Blue staining.

Even though all the steps previously described had been optimized independently, combining them into a whole sequence did not allow us to recover a fully reoxidized antibody at the end of the process, as shown by Coomassie Blue staining of SDS page gel (**Figure 23**B). Three hypotheses were proposed to explain this result: either the reagent designed and used (**9**) was not adapted for this transformation, either some light chains of the antibody were lost in the several purification steps thus preventing to recover a non-reduced antibody, or the excess of cysteine used to remove the remaining thioester also led to the formation of disulfide bond between the L-cysteine **15** added and the cysteine residue of the antibody.

3.2.2.1.   Design of new thioester-containing reagents.

It was thought that the first reagent designed, **9**, was not adapted for the NCL conjugation strategy. To address this hypothesis, two smaller reagents, bearing an azido group for a further functionalization with SPAAC, were synthesized – *S*-phenyl 2-azidoethanethioate **16** and *S*-phenyl 4-azidobenzothioate **17**.

Starting from the coupling reaction between 1-azidoacetic acid **18** and thiophenol **11** in presence of BOP, *S*-phenyl 2-azidoethanethioate **16** was obtained in 29% yield. Under the same conditions, *S*-phenyl 4-azidobenzothioate **17** was obtained in 54% yield from 4-azidobenzoic acid **19** and thiophenol **11**.
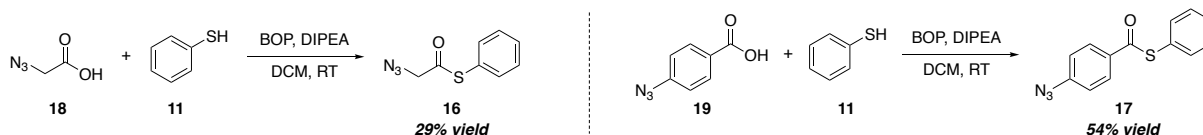
Scheme 48: Synthesis of *S*-phenyl 2-azidoethanethioate **16** and *S*-phenyl 4-azidobenzothiate **17**.

The thiol-selectivity of the two newly synthesized reagents was first evaluated on non-reduced antibody. Both reagents (4 or 8 equivalents) were then incubated at pH 7.5 and 25 °C for 15 minutes on non-reduced trastuzumab, before the resulting conjugates were reacted by SPAAC with a BCN equipped with a cyanine-5 fluorophore **37a**.

An excellent cysteine-selectivity was observed for **17**, as demonstrated by an absence of fluorescence. On the contrary, poor chemoselectivity was observed for **16**, which led to an average DoC of 1.4 on non-reduced trastuzumab. Decreasing the pH of the reaction or reducing the incubation time could not prevent this lack of cysteine-selectivity.

The efficiency of *S*-phenyl 4-azidobenzothioate **17** was then tested on reduced antibody and DoCs of 0.5 and 0.6 were obtained respectively when 4 or 8 equivalents of **17** were simply incubated at 25 °C for 15 minutes. Having thus proved that thioester formation with free cysteines was possible, our previously developed conditions for NCL with **9** were then applied to this new reagent. After *trans*-thioesterification, pH of the solution was increased to 8.5 to transfer the payload to a vicinal lysine and the conjugates were incubated at 37 °C for 24 hours. The remaining thioesters were then removed by the addition of L-cysteine and the antibody reoxidized with L-dehydroascorbic acid, as already reported. An aliquot was taken after each of the five steps and further analyzed by SDS-page (**Figure 24**). Unfortunately, this showed that the antibody was not fully reoxidized after the reaction, irrespective of the number of equivalents of **17** employed.

76
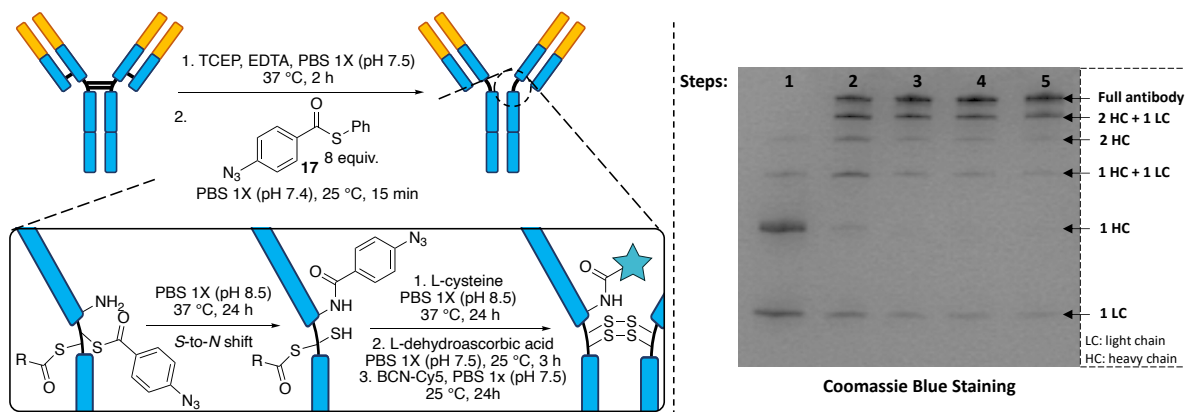
Figure 24: Application of the Native Chemical Ligation on trastuzumab with 8 equivalents of *S*-phenyl 4-benzothiole reagent **17**. Aliquots were taken after each step and analyzed by SDS page. The gel was stained with Coomassie Blue.

3.2.2.2.   Purification optimization

To explain the uncomplete reoxidation of the antibody after the NCL, it was also supposed that some light chains could be lost during the purification steps. As the light chains of the antibody have a molecular weight of 25 kDa, it is possible that they were adsorbed on the gel of a 30 kDa size exclusion chromatography (SEC) column. The different steps of the NCL process were then reproduced with 8 equivalents of **9** on reduced trastuzumab and purification was conducted with size exclusion columns with a cutoff of 6 kDa instead (**Figure 25**A). Unfortunately, as for the previous experiments, the full antibody was not recovered after the conjugation. Removing all the purification steps in the process did not help to recover the full antibody either (**Figure 25**B). In conclusion, the purification steps did not seem to be responsible for the incomplete reoxidation of trastuzumab after the NCL modification.
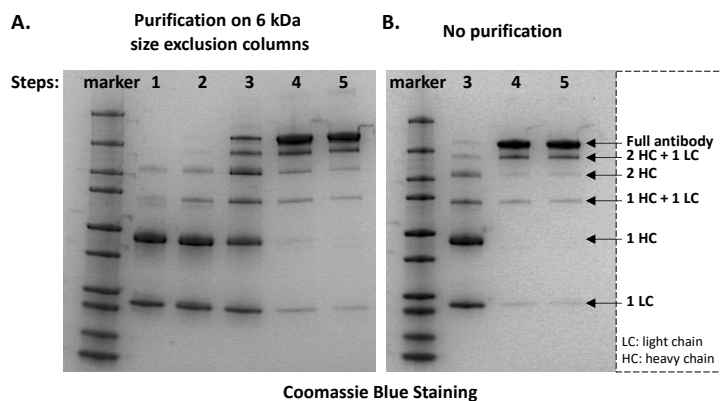

Figure 25: SDS page. **A**. The NCL conjugation reaction was applied on reduced trastuzumab with 8 equivalents of **9**, size exclusion columns with a 6 kDa cutoff were used for the purifications. Aliquots were taken after each step

and analyzed by SDS page. **B.** The NCL conjugation reaction was applied on reduced trastuzumab with 8 equivalents of **9** and no purifications were done during the process. Aliquots were taken after steps 3 (*S*-to-*N* shift), 4 (removing of thioesters with L-cysteine **15**) and 5 (oxidation) and analyzed by SDS page. The gels were stained with Coomassie Blue.

3.2.2.3. Reduction of the undesired disulfide bonds

Finally, a last hypothesis to account for the partial reoxidation of the antibody was that the formation of disulfide bonds between free L-cysteines **15**, added in large excess in the fourth step, and the cysteine residues on the antibody was preventing interchain rebridging. An extra reducing step was thus added after the removal of the excess L-cysteine, in order to reduce these potential undesired disulfide bonds. The now 7-step procedure for the NCL reaction was then performed with *S*-phenyl 4-azidobenzothiole **15** as a labelling reagent, and an aliquot was taken after each step and analyzed by SDS page (**Figure 26**).
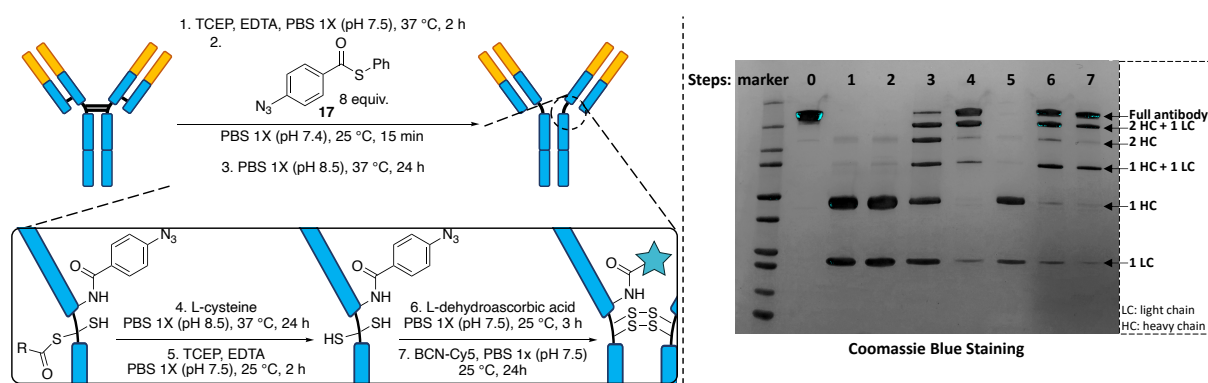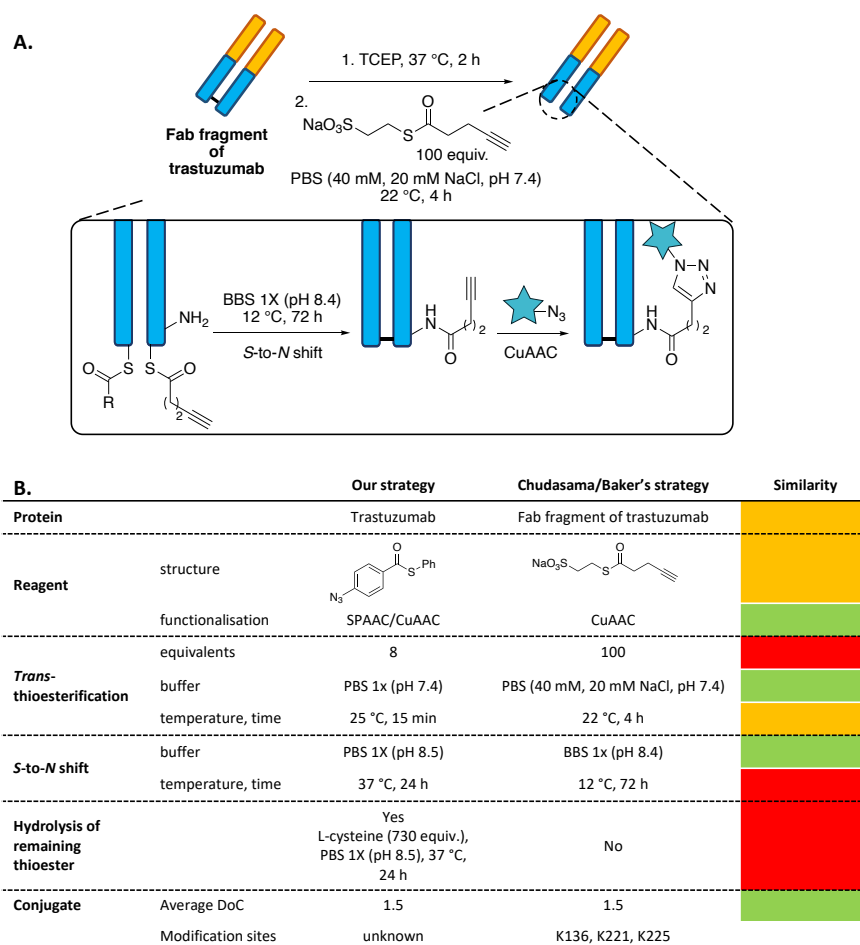

Figure 26: Application of the Native Chemical Ligation on trastuzumab with 8 equivalents of *S*-phenyl 4-benzothiole reagent **17**. Aliquots were taken after each step and analyzed by SDS page. The gel was stained with Coomassie Blue.

Despite the incorporation of this additional reduction step, it was still not possible to reoxidize the antibody completely at the end of the reaction.

Unfortunately, before any additional conditions were tested to optimize this new conjugation reaction, a very similar strategy was published by Chudasama and coworkers, forcing us to abandon this project.[168] To achieve selectivity, they proposed to use a cysteine residue to deliver acylating agents specifically to proximal lysine residues, with a process based on a *trans*-thioesterification followed by a *S*-to-*N* acyl transfer (**Figure 27**A). The conditions proposed were found to be very similar to the conditions we have developed, as shown in **Figure 27**B. As opposed to our strategy, the method was developed on the Fab fragment of

trastuzumab. They also demonstrated that alkyl thioesters were more suitable in the trans-thioesterification step than aryl thioesters, as no background lysine conjugation was occurring with a large excess of reagents (up to 100 equivalents), and, that the *S*-to-*N* acyl shift was optimal at 12 ° C for 72 h (pH 8.4 as found in our strategy), as the transfer was found to be complete and the formed thioester on the Fab fragment suffered less from hydrolysis than at 37 °C. Under their conditions, an average DoC value of 1.5 was obtained on the Fab with three lysine residues modified and located at 8.4 to 13.4 Å from the cysteine relay.

B.		Our strategy	Chudasama/Baker's strategy	Similarity
Protein		Trastuzumab	Fab fragment of trastuzumab	
Reagent	structure			
	functionalisation	SPAAC/CuAAC	CuAAC	
***Trans*-thioesterification**	equivalents	8	100	
	buffer	PBS 1x (pH 7.4)	PBS (40 mM, 20 mM NaCl, pH 7.4)	
	temperature, time	25 °C, 15 min	22 °C, 4 h	
***S*-to-*N* shift**	buffer	PBS 1X (pH 8.5)	BBS 1x (pH 8.4)	
	temperature, time	37 °C, 24 h	12 °C, 72 h	
Hydrolysis of remaining thioester		Yes L-cysteine (730 equiv.), PBS 1X (pH 8.5), 37 °C, 24 h	No	
Conjugate	Average DoC	1.5	1.5	
	Modification sites	unknown	K136, K221, K225	

Figure 27: A. Strategy proposed by Chudasama, Baker and coworkers. **B.** Comparison of our strategy with Chudasama/Baker's approach.

After publication of the method, the authors were contacted to know if they were able to apply their conditions to the full antibody, trastuzumab. They have tried the conjugation on trastuzumab with their alkyl thioester at the optimized conditions for the Fab. After clicking with a fluorophore, a promising average DoC of 4.4 was obtained, with the fluorophore shown to be only on the heavy chain, as observed by SDS page. However, analytically, the distribution of conjugates was found to be broad – from lower loaded species to a maximum degree of

conjugation of 8 – that was explained by thioester hydrolysis. Regarding the reconstruction of the full antibody, they did not have this issue, but they suggested to use 2-mercaptoethanol, instead of L-cysteine, to cleave any non-transferred thioesters and facilitate reoxidation.

3.3.   Conclusion

With the aim of developing a new regioselective strategy, an analogous NCL approach was investigated on the antibody, trastuzumab. Based on a first *trans*-thioesterification between a thioester-containing molecule and a cysteine residue, followed then by an acyl transfer to a lysine residue nearby, the reaction was expected to be highly regioselective. The different steps of this process were first optimized independently, before the full sequence was applied on the antibody. Unfortunately, with the conditions developed, the antibody was never completely reconstructed at the end of the process, and the DoC distribution obtained was broad with a lot of naked antibody still present. At this stage, the reasons behind the non reoxidation of the antibody at the end of the reaction are still unknown, and before corrective measures were applied and tested, a competing publication forced us to stop this project.

Having investigated a first regioselective strategy for the labelling of lysine residues, we aimed in the next chapter to achieve site-selectivity by targeting the site-chains of two vicinal amino acids with a multicomponent reaction.

MULTICOMPONENT APPROACHES FOR THE SITE-SELECTIVE LABELLING OF PROTEINS

## 4.1. Multicomponent reactions in bioconjugation chemistry

4.1.1. History of the multicomponent reactions (MCR)

Multicomponent reactions (MCRs) are defined as reactions in which three or more starting materials react to form a single product.[282] Classically, various steps are required for the synthesis of complex molecules. On the contrary, MCRs allow their syntheses in a single step, in which all the reagents are added simultaneously, with most of the atoms contributing to the formation of the new molecule, making for straightforward and cost-efficient reactions.[283] Generally considered to be the first MCR, the Strecker reaction was reported in 1850 for the synthesis of α-amino acids **R275** from aldehyde **R272**, ammonia **R273** and hydrogen cyanide **R274** (**Scheme 49**).[284] Although the reaction is now known as the Strecker reaction, Laurent and Gerhardt isolated the product of this MCR by reacting bitter almond oil with ammonia, twelve years before Strecker.[285] Eighty years later, Bucherer and Bergs extended this reaction to a four-component version by adding carbon dioxide **R277** (**Scheme 49**). In contrast to the Strecker reaction that was in equilibrium and gave α-amino acids in unsatisfactory yields, the Bucherer-Bergs reaction was found to be irreversible upon addition of CO_2.[282]
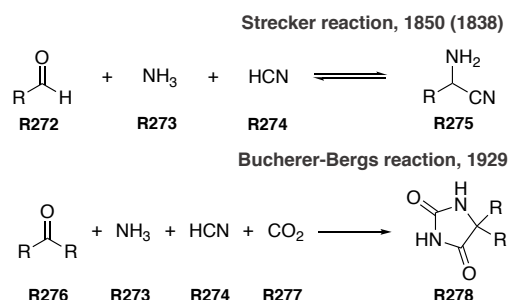

Scheme 49: Strecker and Bucherer-Bergs reactions

In 1882, the Hantzsch 1,4-dihydropyridine synthesis was reported (**Scheme 50**).[286] This four-component reaction, which requires an aldehyde **R279**, ammonia **R273** and two β-keto esters **R280**, was later employed by Bayer AG for the synthesis of Nifedipin, a drug used for cardiovascular disease.[287] Following the synthesis of dihydropyridine, Hantzsch proposed a three-component version for the synthesis of pyrroles in 1890.[288] Indeed, the formation of pyrroles **R284** is made possible by the reaction of α-halomethyl ketones **R281** with β-keto esters **R282** and amines **R283**. Also based on β-keto esters, the Biginelli reaction was described one year later for the synthesis of pyrimidones **R288** from urea **R285**, β-keto esters **R286** and aryl aldehydes **R287**.[289,290]
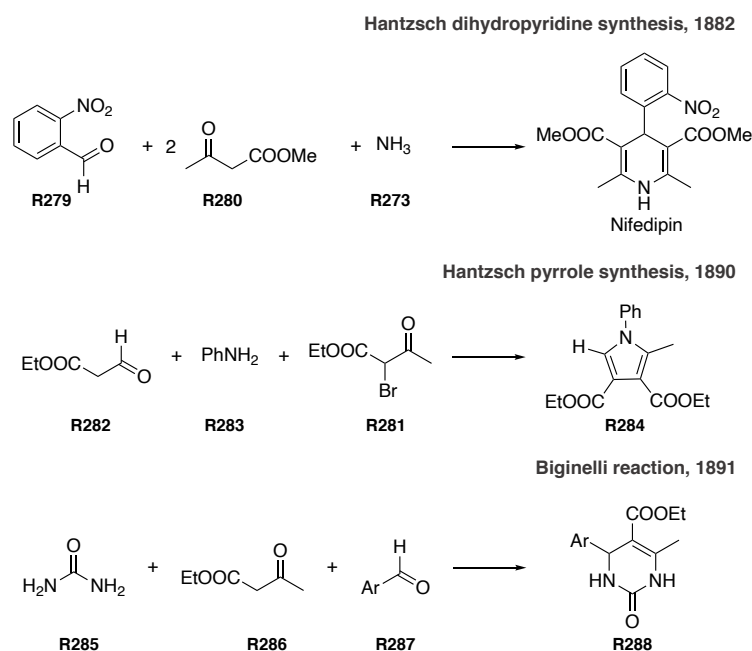
Scheme 50: Hantzsch dihydropyridine and pyrrole syntheses and Biginelli reaction

In the early 20th century, the Mannich reaction was reported, giving access to β-amino enone derivatives **R292** from non-enolizable aldehydes **R289**, primary or secondary amines **R290** and enolizable carbonyls **R291** (**Scheme 51**).[291] Following the results obtained by Carl Mannich, the Robinson condensation was later described in 1917, demonstrating the possibility of obtaining, in a single step, tropinone alkaloids from a dialdehyde, a primary amine and a β-keto ester.[292]
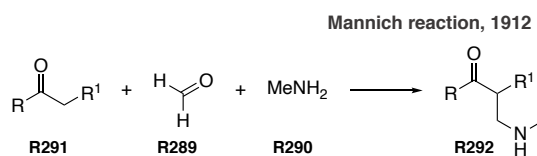
Scheme 51: Mannich reaction

The first multicomponent reaction employing isocyanides was described by Mario Passerini in 1921 and allowed the formation of α-acyloxy amides **R296** from three components – an aldehyde **R293**, a carboxylic acid **R294** and an isocyanide **R295** – at or below room temperature, in apolar solvent at high concentration of reagents.[293] Using DFT calculation, Morokuma and coworkers proposed a mechanism for this transformation (**Scheme 52**).[294] The reaction started with the generation of a H-bond cluster **I** between the aldehyde **R293** and the carboxylic acid **R294**. The isocyanide **R295** then reacted with this H-bond cluster **I** to give

intermediate **II** with small energy of activation. The rearrangement of intermediate **II** into the final product **IV** was associated with a high energy barrier. To lower this barrier, the coordination of an extra carboxylic acid, used as a fourth component, was found to be necessary, allowing a first rearrangement to give intermediate **III** and release a carboxylic acid component. The released carboxylic acid coordinated with intermediate **III**, allowing a final rearrangement to give the α-acyloxycarboxamide product **IV / R296**.
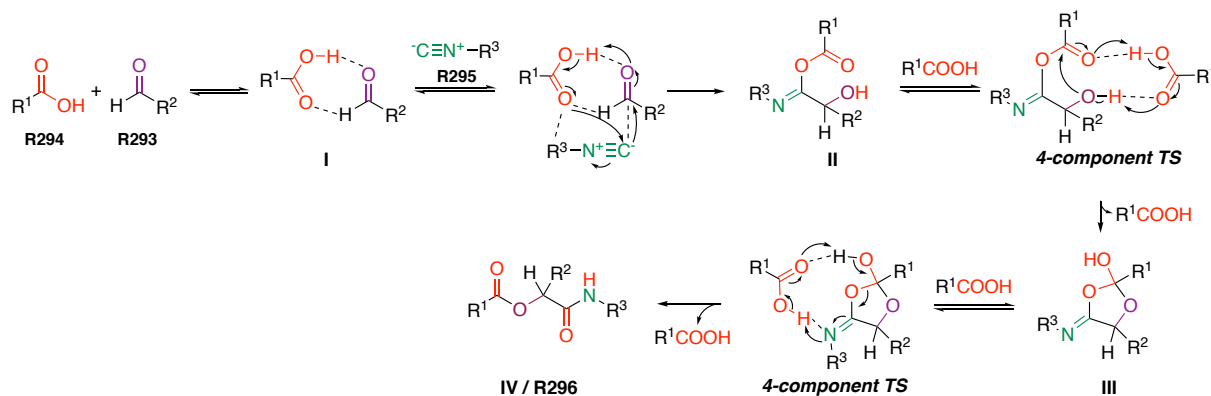
Scheme 52: Mechanism proposed by Morokuma and coworkers for the Passerini reaction.[294]

Surprisingly, it was only 40 years later (1959) that Ivar Karl Ugi had the idea of adding an amine **R297** to the mixture, giving then access to the bis-amide compound **R298** as a final product.[282,295] From a mechanistic point of view (**Scheme 53**), the reaction started with the formation of an imine **I**, from an aldehyde **R294** and an amine **R297**, which was then stabilized by hydrogen bonding **II** with an acid **R294**. An isocyanide **R295** was then inserted on the imine, creating a new C-C bond, while the proton of the acid was transferred to the nitrogen, leading to the irreversible formation of the nitrilium carboxylate ion intermediate **III** which in turn led to the imidate intermediate **IV**. In a last step, the imidate **IV** underwent a Mumm rearrangement to give the bis-amide product **R298**. An alternative path was proposed for the formation of the imidate **IV**. By calculation, it was found that a first fragmentation of the hemiaminal followed by the addition of the isocyanide would lead to the imidate. However, the energy barriers associated with this path were found by the authors to be too high, making it unlikely to occur. The only rate-determining step found for the Ugi reaction was the isocyanide insertion.[296]
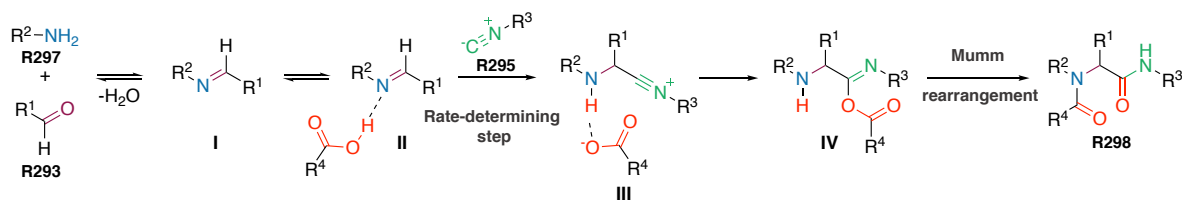
Scheme 53: Mechanism proposed by Fleurat-Lessard for the Ugi reaction.[296]

Variations of the Ugi four-component reaction (U-4CR) were described over the years, such as the Ugi-Smiles reaction where a phenol replaced the carboxylic acid, or, the Ugi four-center three-component reaction (U-4C-3CR), in which two of the four components necessary for the reaction were borne by the same molecule, as it is the case with preformed imines for example.

As some MCRs, such as the Mannich or the Ugi reaction, can be done in water, they have gained considerable attention over the past two decades, as they could be used as a new tools for protein's modification.

4.1.2. Multicomponent reactions in bioconjugation
4.1.2.1. The Mannich reaction

Francis and coworkers were the first to use the Mannich reaction for the chemoselective modification of tyrosine residues in 2004, using the side-chain phenol group as a surrogate for the traditional enolate nucleophile (**Scheme 54**).[218,297] With a large excess of formaldehyde **R215** and aniline **R216**, surface-accessible tyrosine of different proteins – chymotrypsinogen A, lysozyme and RNase A – were successfully modified at pH 6.5 within 18 hours. The authors were able to remove the undesired imines formed during the reaction by adding a solution of hydroxylamine. The method described by Francis was later used for the attachment of biologically active peptide sequences on small proteins and for the production of pesticide-conjugates used in immunoassays.[297,298]
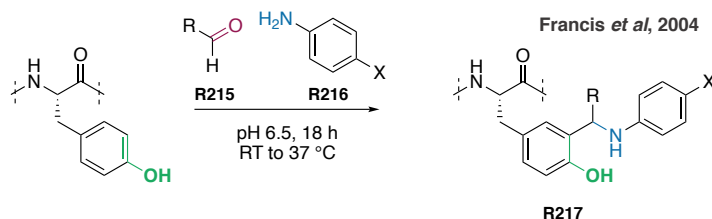

Scheme 54: Chemoselective modification of tyrosine residues on proteins with the Mannich reaction

In order to target other amino acids, such as lysine or *N*-terminal proline residues, the Mannich reaction was derivatized.

Rai *et al.* proposed a phospho-Mannich reaction for the chemoselective modification of lysine in proteins using triethylphosphite **R144** as the nucleophile component (**Scheme 55**).[167] A single-site labelling was achieved with 40% conversion on average from several proteins upon reaction with triethylphosphite **R144** and benzaldehyde derivatives **R143**.

Rai and coworkers were later able to label a single lysine on various proteins based on this reaction (**Scheme 55**).[166] Single modification was achieved on lysine in presence of formaldehyde **R140** and a terminal alkyne **R141**. However, to proceed, the reaction required a large excess of Cu(I) (100 equiv.).
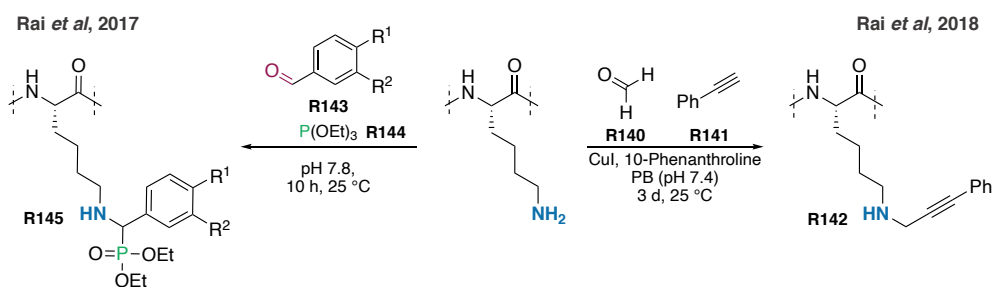
**Scheme 55:** Rai and coworkers approaches for the single conjugation of lysine residues in proteins.

As an alternative, the Petasis reaction – a variation of the Mannich reaction where boronates **R207** are used as nucleophiles – was also reported for the conjugation of proteins (**Scheme 56**). As for the Mannich reaction, an iminium was first formed from the condensation of a secondary amine **R299** with an aldehyde **R206**, before the nucleophilic addition of the organic ligand from a boronate species led to a final tertiary amine. Raj *et al* showed that the Petasis reaction was efficient for the *N*-terminal proline labelling of tripeptides, bioactive peptides and small proteins.[213]
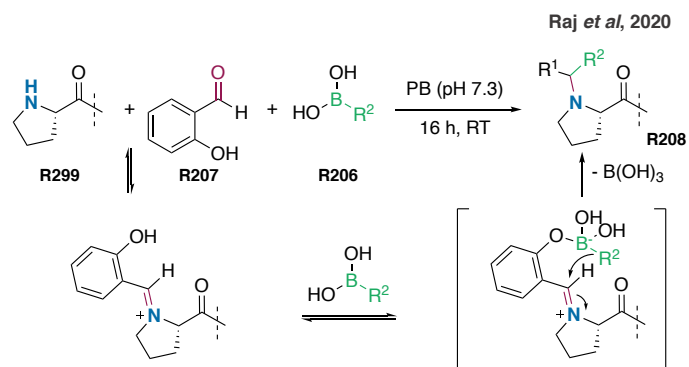
Scheme 56: Use of the Petasis reaction for the labelling of the *N*-terminal proline in proteins.

Encouraged by the results obtained with the Mannich reaction as a site-selective approach for the conjugation of proteins, other multicomponent reactions were described.

Ball and coworkers developed a three-component organometallic reaction for the selective modification of tyrosine (**Scheme 57**).[220] Indeed, in the presence of Rh(III) salts **R221**, aryl boronic acid **R220** were selectively linked to tyrosines through an arene complex **R222** on various proteins.
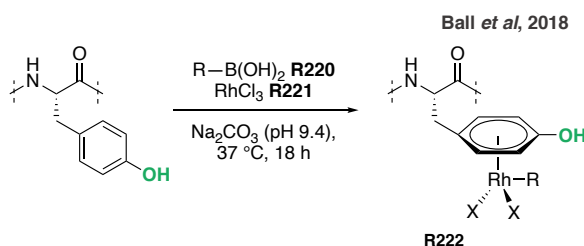

Scheme 57: Multicomponent reaction developed by Ball and coworkers for the chemoselective labelling of tyrosine residues in proteins.

4.1.2.2.   The Ugi and Passerini reactions

The U-4CR and its variations have for long been useful tools for the construction,[299–303] stapling[304,305] and macrocyclization of peptides.[283,305–311] However, conditions used for the modification of peptides are not often compatible with proteins. Indeed, the reactions are usually done in methanol, and often in presence of organic co-solvent at high concentrations of starting materials. This notwithstanding, a few methods for the modification of proteins with the Ugi reaction have been reported, mostly for the immobilization of enzymes on different materials.

The first report dates back to 1974, when the immobilization of three proteins – trypsin, succinyl trypsin and pepsin – on the surface of a polyisocyanide-nylon **R300** was reported (**Scheme 58**).[312] Both the amino and the acid functional groups of the proteins were used for their immobilizations. Mixed with acetaldehyde **R301** and an excess of acetate **R302**, trypsin was bound to isocyanide-containing nylon fibers via its lysines' side chains **R303** in an aqueous buffer at neutral pH. Conversely, succinyl trypsin and pepsin were covalently attached to the nylon through their carboxylate groups in presence of acetaldehyde and a primary amine, 2-amino-2-hydroxymethylpropane-1,3-diol (Tris). However, immobilization had a detrimental impact on the enzymes' activities: trypsin proteins lost 40% of their activities, while only 7% of pepsin activity was recovered. Two hypotheses were given to explain the poor enzymatic activity of the immobilized pepsin: either the conjugation site was too close to the enzymatic active site, leading to its inactivation, or the enzyme was inactivated due to a prolonged exposure to acidic pH (5). Despite a loss of their activities, their thermal and pH stabilities was found to be improved compared to native enzymes. Goldstein *et al* further applied this method in 1978 for trypsin and succinyl trypsin's immobilizations on a polyethylene polymer grafted with isocyanide groups.[313]
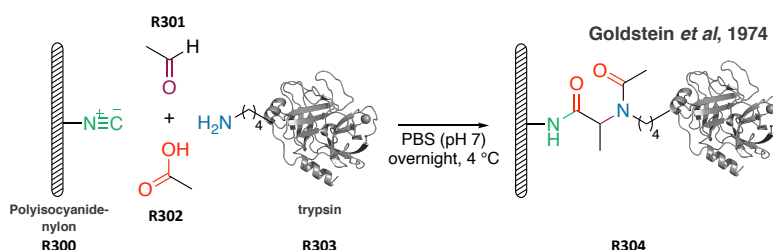
Scheme 58: Immobilization of trypsin on a polyisocyanide-nylon **R300**. Mixed with acetaldehyde **R301** and acetate **R302**, the protein **R303** was covalently attached to the surface through its lysine sidechains via the U-4CR.

Using the same approach, a biosensor device for hydrogen peroxide (H_2O_2) quantification was developed in 2007, by attaching horseradish peroxidase **R305** (HRP) to a polysaccharide-coated gold electrode **R306** (**Scheme 59**).[314] The carboxylic acids of the polysaccharide reacted with HRP's lysine side chains or *N*-terminus in the presence of formaldehyde **R140** and *tert*-butyl isocyanide **R307**. The efficiency of the HRP-modified electrode **R308** was evaluated by cyclic voltammetry. In presence of HRP, H_2O_2 oxidized hydroquinone **R309** into benzoquinone **R310**, which was then subsequently reduced at the surface of the electrode. The enzymatic electrode showed good sensitivity and selectivity towards H_2O_2 and retained its

full activity after one month of storage at 4 °C, making the Ugi reaction a good tool for the preparation of amperometric enzyme biosensor.
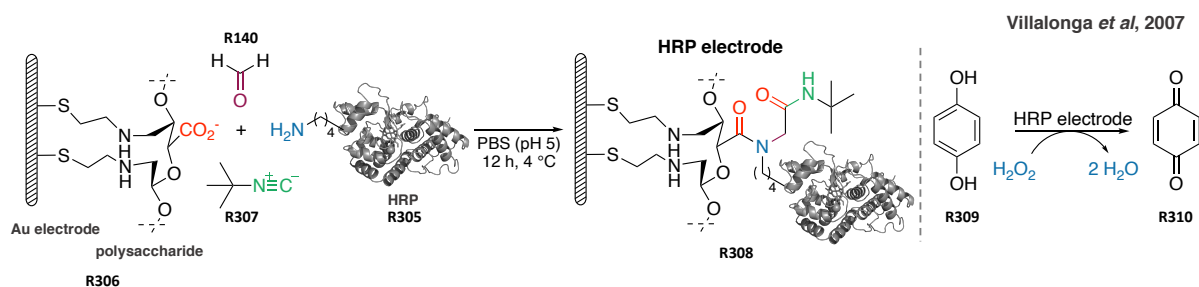
Scheme 59: Development of a biosensor device for the hydrogen peroxide quantification. HRP **R305** was immobilized through its lysine sidechains on a gold electrode coated with polysaccharides **R306** (containing the acid component) in presence of formaldehyde **R140** and *tert*-butyl isocyanide **R307**. The immobilized protein **R308** could efficiently convert hydroquinone **R309** into benzoquinone **R310** in presence of hydrogen peroxide.

Based on this method, Villalonga and coworkers described a novel synthetic procedure for the preparation of trypsin-polysaccharide neoglycoenzymes.[315] By using the Ugi reaction as a coupling strategy, free amino groups of trypsin were covalently attached to *o*-carboxymethylcellulose (CMC) or sodium alginate in presence of acetone and *tert*-butyl isocyanide. After this immobilization process, trypsin was found to be more stable towards thermal treatment and autolytic degradation compared to the native enzyme. However, the esterase activity of the enzyme was found to be decreased upon immobilization (61-69% activity with respect to native trypsin).

Due to high cost and poor stability of a lipase enzyme, *Rhizomucor miehei lipase* (RML), there was a high demand for new immobilization methods that could improve the catalytic properties of the biocatalyst – stability, activity and recycling properties – without requiring harsh conditions. To respond to this demand, Mohammadi and its coworkers immobilized the lipase with the U-4C-3CR or the U-4CR on different supports coated with different functional groups (**Scheme 60**).

For the development of a new biocatalyst that can produce fatty acid methyl esters from colza oil and methanol, RML was immobilized on a silica support **R311**.[316] In this example, the enzyme supplied either the amino and/or the carboxylic acid component, while the silica support provided the aldehyde group; *tert*-butyl isocyanide was used as the fourth component. As two reactions – Ugi or Passerini reactions – could compete with each other, the real mechanism of this immobilization was investigated. Two experiments were then designed. In a first one, the free amines of the lipase were blocked by reaction with an aldehyde to form imines. As a consequence, only a negligible portion of the protein was grafted on the aldehyde-

coated support, thus suggesting that the protein's carboxylic acids did not partake in the Passerini reaction. In a second experiment, the carboxylic acids were coupled with ethylenediamine, thus making them inaccessible to partake in the immobilization. As observed for the first experiment, the immobilization was found to be unsuccessful. From those two experiments, the authors conclude that the U-4CR took place between aspartate/glutamate and lysine residues of the enzyme. The immobilization of the enzyme on silica materials improved its thermal stability at 55 °C; 60-69% of the immobilized enzyme activity was retained after 90 minutes, in comparison, the free enzyme was completely inactivated under the same conditions. The biocatalyst catalyzed the transesterification of colza oil and methanol to produce fatty acid methyl esters with moderate yields and was recycled up to four times.

In the same perspective, the immobilization of RML via its carboxylate groups was investigated by Mohammadi *et al.* in 2018 on amino-modified silica support **R312** in the presence of cyclohexyl isocyanide and acetaldehyde.[317] As for the previous example, a gain in thermal and co-solvent stability was observed upon immobilization. The resulting immobilized preparation was used as a biocatalyst in the enantioselective esterification of (*R,S*)-ibuprofen with 1-propanol with only the *S* compound being modified.

Based on the same approach, they were able to graft RML on multiwall carbon nanotubes (MWNTs) and graphene nanosheets **R313** using the U-4CR.[318] Those supports supplied the carboxylic acid groups, and the protein of interest, the amino groups. Reacting it with acetaldehyde and cyclohexyl isocyanide, lipases were successfully attached to both supports. The thermal stability of the protein was also improved (94% activity after 24 hours at 55 °C) and its activity was maintained in presence of organic solvents (retains 71-100% of its activity after 24 hours of incubation in presence of 20% of an organic solvent).
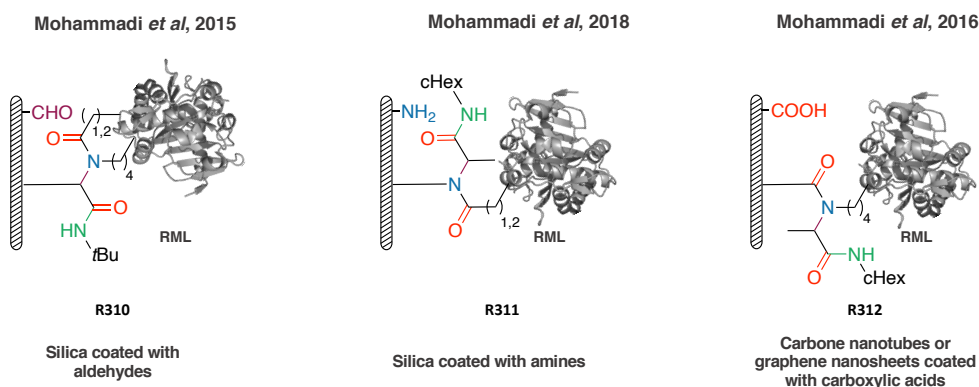
Scheme 60: Immobilization of the RML enzyme on supports coated either with aldehydes, amine or carboxylic acids with the U-4C-3CR or the U-4CR.

Although different enzymes were successful coupled to various materials, no information was given concerning the site of modification on the protein.

In parallel, the Ugi reaction has also been applied to the functionalization of proteins. In 2000, Ziegler and coworkers reported the preparation of bioconjugates with the Ugi reaction (**Scheme 61**).[319] Two proteins were then selected, BSA and HRP, and mixed with different carboxylic acids, isocyanides, carbonyls and amines. The two proteins reacted either via their amino or carboxylic groups, in their native form, or via aldehyde groups after oxidation of the HRP's carbohydrates with sodium periodate, depending on the conditions used. To achieve good conversions, up to 4000 equivalents of reagents were necessary, with incubation times of 1 to 4 days at 25 °C. Under the conditions used, a side reaction would have been possible, the Passerini reaction, however, the authors made no comment regarding this possibility.
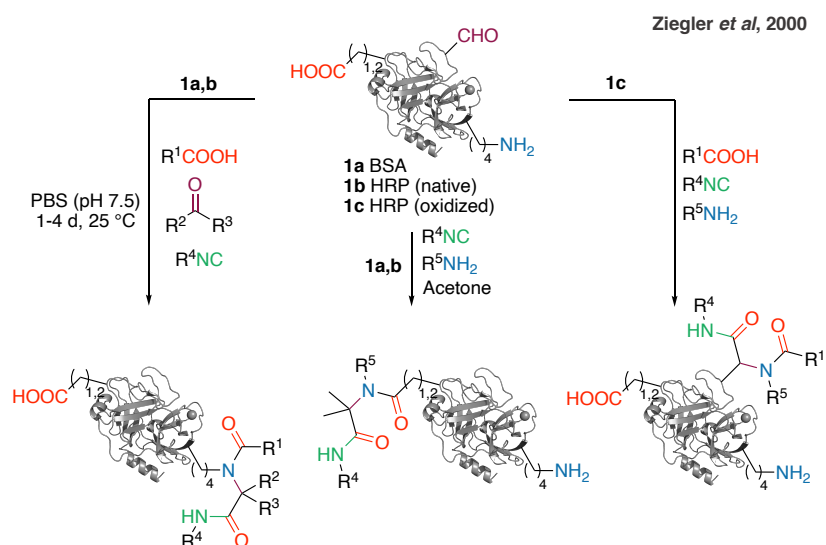
Scheme 61: Labelling of BSA, HRP (native) and HRP (oxidized) with the U-4CR.

For the development of glycoconjugate vaccine candidates, Rivera *et al.* proposed to use the U-4CR (**Scheme 62**).[320] For the construction of their vaccines, two bacterial capsular polysaccharides (CPs) from *Streptotoccus Pneumoniae* and *Streptotoccus Typhi* were oxidized and chosen as the aldehyde component **R313** and conjugated to carrier proteins **R314** – tetanus toxoid (TT) and diphteria toxoid (DT) – acting here as the amine source in the multicomponent reaction. Acetic acid **R302** and *tert*-butyl isocyanide **R307** were chosen as the two other components for the reaction. Because only 41% conversion was observed after 48 hours, a more nucleophilic hydrazide group was introduced on the proteins, via the modification of aspartate / glutamate residues, allowing for complete conversion after 4 hours. Two years later, based on the same strategy, they were able to increase the complexity of their vaccines by making multivalent glycoconjugates from four different capsular polysaccharides.[321]
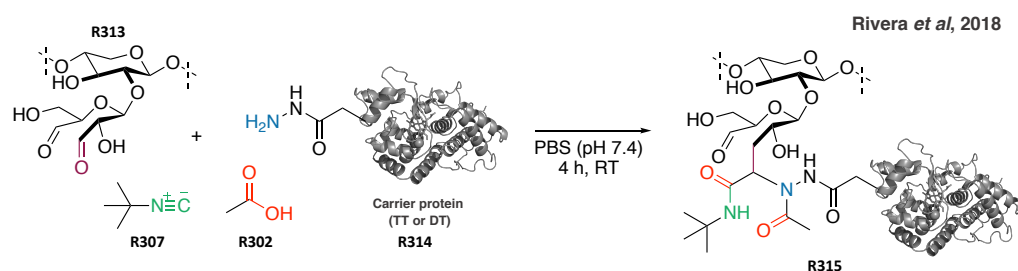
Scheme 62: Development of antibacterial glycoconjugate vaccine candidates. Carriers proteins **R314** (TT or DT) were covalently attached to bacterial capsular polysaccharides **R313** (CP) in presence of acetate **R302** and *tert*-butyl isocyanide **R307** with the U-4CR.

To summarize, most of the strategies reported so far rely on the four-component version of the Ugi reaction (U-4CR) and aim to modify the protein on either lysine or aspartate / glutamate residues. Most notably, all these strategies tend to require a large excess of reagents and long reaction times. Due to the absence of thorough analytical investigation, it is not possible to confirm that the Ugi reaction, and not another uncontrolled side reaction such as the Passerini reaction, was responsible for the bioconjugation. Similarly, no conclusion can be drawn regarding the regioselectivity of the method.

4.1.3. Aim of the project

Building up on these literature precedents, we started pondering over the idea of using the Ugi four-centre three-component (U-4C-3CR) version of the reaction to target lysine / aspartate –

or lysine / glutamate – duets in close spatial proximity. As the vast majority of protein conjugation strategies focus on the modification of a unique residue, we, along with the Cianférani group at the University of Strasbourg, hypothesized that targeting two different amino acid side chains simultaneously should give higher chances of developing a site-selective strategy.

Following the general mechanism described previously in **Scheme 53**, the U-4C-3CR – conducted in the presence of a carbonyl compound and an isocyanide – should give access to stapled proteins **R318**, via a macrolactamization between the two residues' side chains (**Scheme 63**). Even if all amine functionalities could theoretically react with the carbonyl compound to form imines, only the ones close to a carboxylate functional group should lead to the formation of a stable final product. Remaining, unreacted imines could then be simply reacted with a hydroxylamine solution at the end of the reaction to regenerate the free amino functional groups.
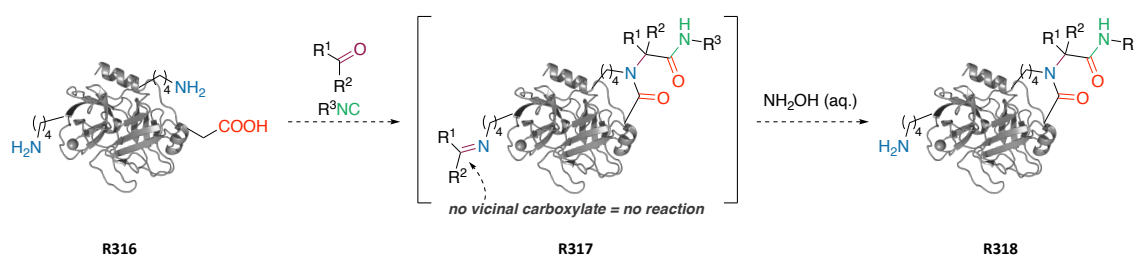
**Scheme 63:** Application of the Ugi reaction to the bioconjugation of native proteins.

In order to do so, the reaction was explored on the model antibody trastuzumab (150 kDa). In addition, thorough analytical experiments were conducted to identify the sites modified on the antibody with the U-4C-3CR and to conclude whether the method is site-selective.

4.2.   Investigating Ugi / Passerini multicomponent reactions for the site-selective conjugation of native proteins.[322]

4.2.1. Design of the reagents

For practical purpose, an isocyanide-functionalized fluorophore was initially designed to easily monitor the reaction on biomolecules. TAMRA **10** was coupled with *tert*-butyl (3-aminopropyl)carbamate **20** in presence of HBTU as a coupling reagent. The Boc-protecting group was then removed with 20% HCl in 1,4-dioxane to afford compound **21** in 40% yield over two steps.[323] Potassium 2-isocyanoacetate **22** was finally coupled with **21.** Even though the isocyanide-containing fluorophore **23** was obtained under these conditions, it was unfortunately not possible to recover a pure fraction of the desired product. Purification of the crude on reverse phase (with or without TFA) or normal phase chromatography led to the hydrolysis of the isocyanide, yielding the corresponding formamide.
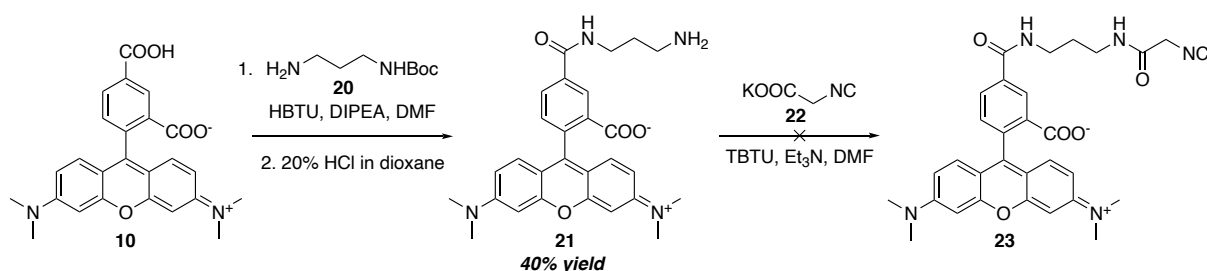

Scheme 64: Attempt to make the TAMRA-isocyanide **23.**

To avoid the hydrolysis of the isocyanide, it was thought of mimicking the structure of the *tert*-butyl isocyanide **33a**, known to be stable under aqueous conditions – and already used for the functionalization of proteins[314,316,320] – by introducing two methyl groups at the α-position of the isocyanide. From potassium 2-isocyanoacetate **22**, potassium 2-isocyano-2-methylpropanoate **24** was obtained in 49% yield over three steps and was then coupled with **21**.[324,325] Unfortunately, as in the previous attempt, the isocyanide functionality of **27** was hydrolyzed during the purification, yielding also the formamide product.
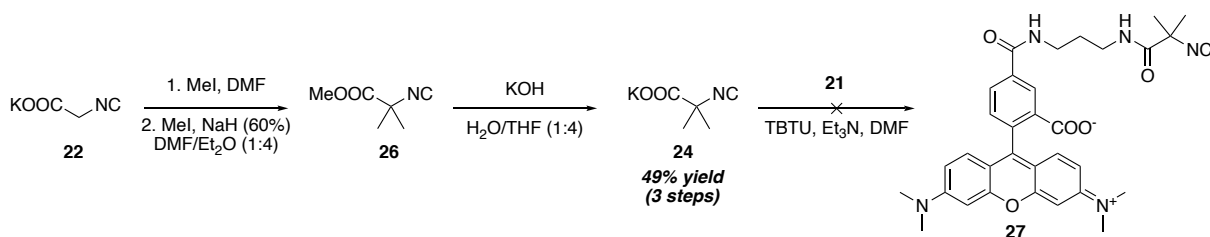
Scheme 65: Attempt to make the TAMRA-isocyanide **27.**

As the isocyanide functionality could not be introduced onto TAMRA, it was attempted onto a coumarin molecule.[326] Starting from 7-amino-4-methylcoumarin **28**, formamide **29** was obtained by refluxing **28** in a solution of formic acid and toluene, which was then dehydrated with phosphorus oxychloride to give 7-isocyano-4-methylcoumarin **30** in poor yield. Unfortunately, **30** was found to be unstable in phosphate buffer.
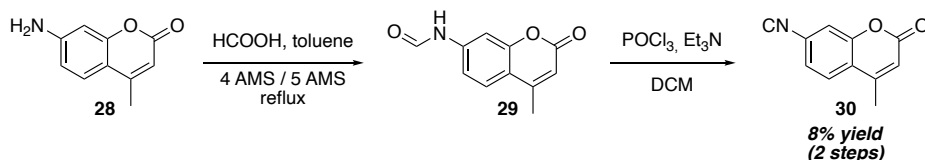
Scheme 66: Synthesis of 7-isocyano-4-methylcoumarin **30**.

As it was not possible to obtain a stable isocyanide-containing fluorophore, we turned our attention to aldehyde-containing fluorophores instead.[327] Starting from TAMRA **10**, 3,3-diethoxypropan-1-amine **31** was coupled within few hours. The hydrolysis of the acetal during the purification by preparative HPLC directly gave pure TAMRA-CHO **32a**.
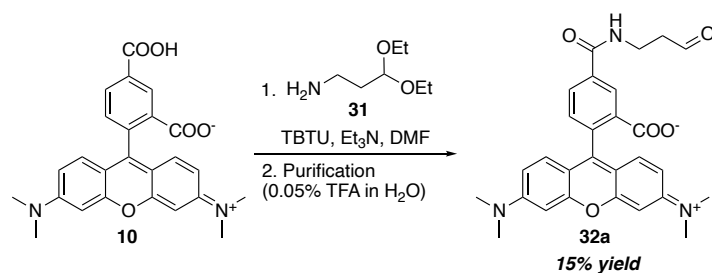

Scheme 67: Synthesis of the aldehyde-containing TAMRA **32a**.

94

4.2.2. Ugi reaction on proteins

4.2.2.1. Optimization of the multicomponent reaction on trastuzumab

The feasibility of the Ugi reaction on proteins was initially tested on the monoclonal antibody trastuzumab, in presence of various amounts of the model aldehyde, TAMRA-CHO **32a**, and *tert*-butyl isocyanide **33a**. To avoid the use of amine or carboxylate-containing buffers that might compete with the antibody for the Ugi reaction, experiments were set up in a phosphate-buffered saline (PBS) buffer (1X, pH 7.5, 137 mM NaCl). At the end of the reaction, a solution of hydroxylamine **12** was added to remove the imines that did not partake in the reaction. The efficiency of the reaction – average degree of conjugation (DoC) – was then measured by fluorescence before being confirmed by native mass spectrometry (native MS).

Pleasingly, even with few equivalents of **32a** and **33a** (from 10 to 40 equivalents), conversions were observed on the antibody in just 16 hours at 25 °C, without the addition of catalyst or imine preformation (**Figure 28**A). Average DoCs and conversion rates were improved by increasing the amounts of both aldehyde and isocyanide; however, a plateau was reached when more than 60 equivalents of **32a** and 120 equivalents of **33a** were used, giving access to a maximum average DoC of 2.5 and a conversion rate of 90%. This maximum in conversion and average DoC values could indicated a limited number of functionalization sites accessible on the antibody, as the maximum DoC value reached was 6 and only represented less than 5% of all conjugates for the highest conversions ($\geq$ 80%) (**Figure 28**B). Overall, better conversions and DoC distributions were obtained when equimolar or near equimolar ratios of **32a** and **33a** were used. The best ratio of aldehyde/isocyanide was found to be 45 equivalents of both **32a** and **33a**, excellent conversions with a narrow DoC distribution were obtained, with adducts with DoC > 3 (ie. D4 and D5) representing a negligible fraction of all conjugate species (**Figure 28**B). Control experiments where the antibody was reacted with either **32a** (100 equivalents) or **33a** (200 equivalents) led to no conjugation in either case, thus confirming the multicomponent aspect of the reaction.
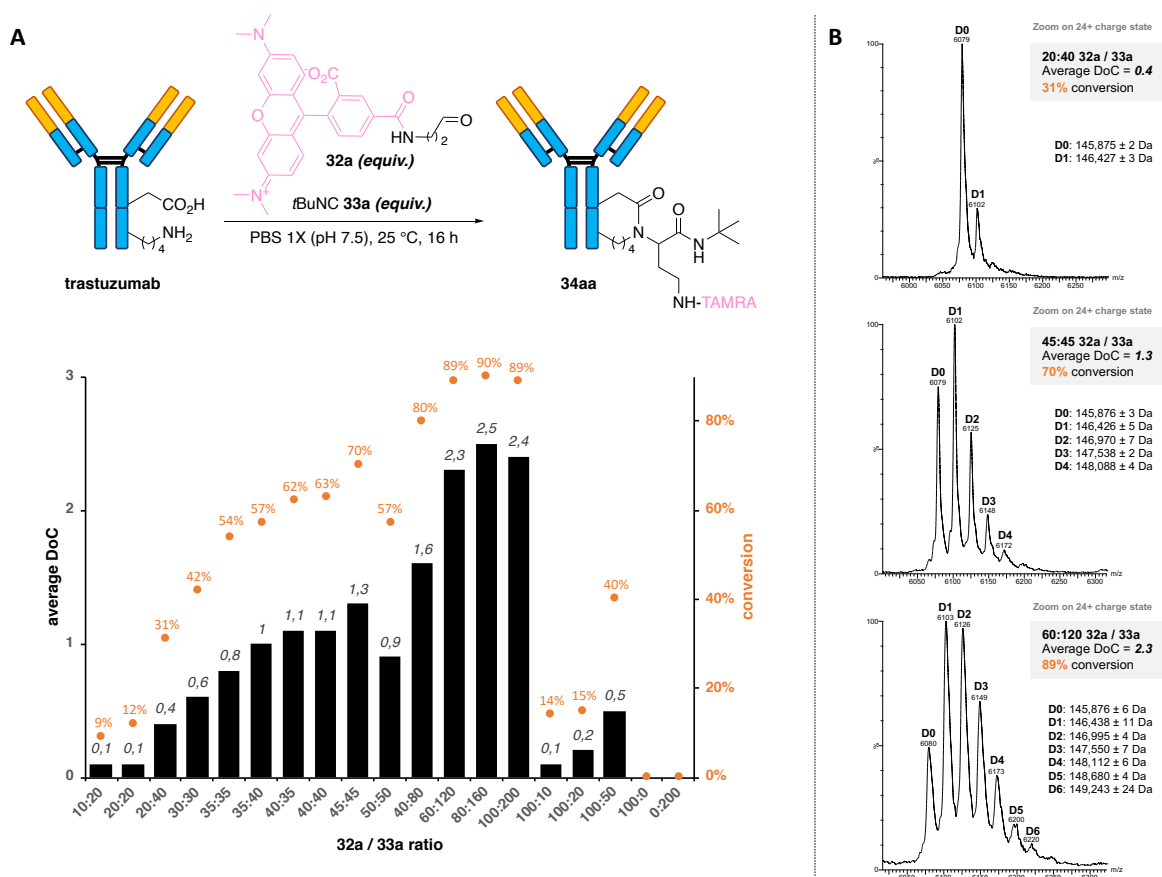
Figure 28: Ugi four-centre three-component reaction on trastuzumab. **A**. Influence of the aldehyde **32a** to the isocyanide **33a** ratio on average DoC and conversion values on trastuzumab, determined by native MS. **B**. Native mass spectra of the resulting functionalized trastuzumab after deglycosylation (zoom on the 24+ charge state); reaction conditions: **32a** (20, 45 or 60 equiv.), **33a** (40, 45 or 120 equiv.), PBS 1X (pH 7.5, 137 mM NaCl), 25 °C, 16 h; conversion (%) = 100 % - D0.

Incubating the antibody with **32a** and **33a** for longer time (2.5 or 5 days) or increasing the incubation temperature to 37 °C led to higher DoCs and conversion rates but also broadened the DoC distribution (up to D4 when 20:40 equivalents of **32a**:**33a** was used). As expected, decreasing the reaction temperature to 4 °C dramatically decreased the average DoC – 0.09 when 20:40 ratio of **32a**:**33a** was used. On the other hand, pH and concentration of NaCl in the buffer were found to have little to no impact (**Figure 29**). The optimal conditions for the conjugation of trastuzumab, offering the best compromise between conversion and DoC distribution, were found to be 45 equivalents of both **32a** and **33a** in a PBS buffer (1X, pH 7.5, 137 mM of NaCl), with a reaction time of 16 hours at 25 °C.
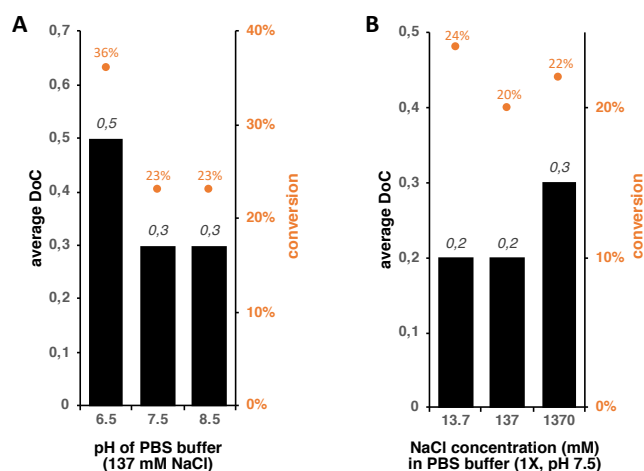
Figure 29: Influence of pH (**A**) and buffer composition (**B**) on average DoC and conversions values of the Ugi reaction.

4.2.2.2.   "Plug-and-play" strategy

Building on the knowledge acquired from this optimization phase, we then aimed to develop a "plug-and-play" version of our strategy, by first incorporating an azide group onto the antibody, via the multicomponent reaction, that could be functionalized in a second step with SPAAC. For practical synthetic reasons, it was found easier to incorporate the azide functional group onto the aldehyde component. From the N_3-PEG4-OH **35** – obtained in three steps from the commercially available tetraethylene glycol **36** and generously provided by Dr Sylvain Ursuegui and Michel Mosser – the aldehyde, N_3-PEG4-CHO **32b**, was obtained in 66% yield by applying the Swern oxidation conditions.[328,329]
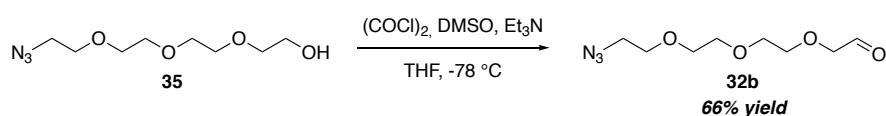

Scheme 68: Synthesis of N_3-PEG4-CHO **32b**

4.2.2.2.1.   Optimization of the "plug-and-play" reaction on trastuzumab

The Ugi four-centre three-component reaction was then evaluated with the aldehyde, N_3-PEG4-CHO **32b** and *tert*-butyl isocyanide **33a** on trastuzumab under our optimal conditions. The resulting conjugate **34ba** was then purified by size-exclusion chromatography and subjected to SPAAC with the bicyclononyne (BCN) derivative **37a**, bearing a cyanine-5 fluorophore, to give the functionalized adduct **38**. It is important to note that functionalization of conjugate **34ba** with **37a** led to partial precipitation of the adduct – around 40% – likely due

to the hydrophobic nature of the cyanine-5 payload, thus slightly impacting average DoC and conversion values.

Different aldehyde-to-isocyanide ratios – from 40:40 to 90:90 equivalents with respect to 1 equivalent of antibody – were then screened (**Figure 30**). In general, average DoC and conversion values were found slightly lower than with **32a**. However, the DoC distribution turned out to be better, with D4 being the highest conjugated species observed across all conducted experiments and representing <8% of all conjugates.
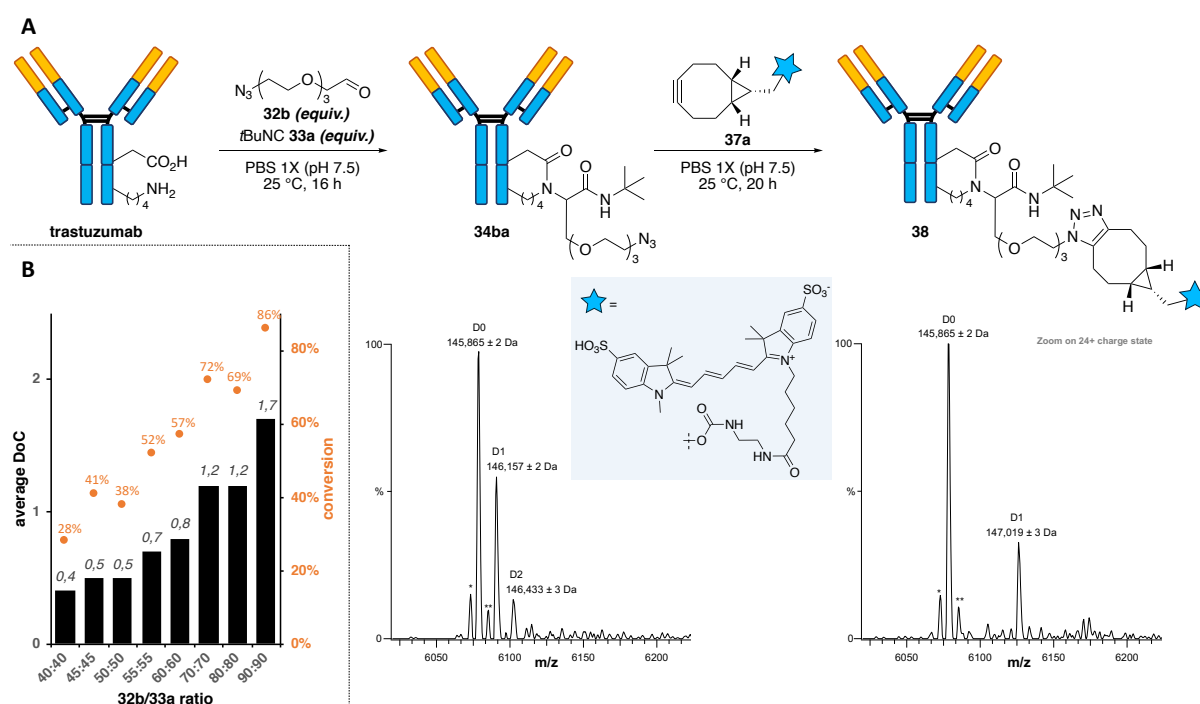
Figure 30: "Plug-and-play" strategy. **A**. Ugi four-centre three-component reaction on trastuzumab with the azide-containing aldehyde **32b** and the isocyanide **33a** followed by the functionalization with the strained alkyne **37a** with the native MS spectra of the different adducts after deglycosylation (zoom on the 24+ charge state). **B.** Optimization of the "plug step": various aldehyde-to-isocyanide ratios were screened, and their corresponding average DoC and conversion values were determined after deglycosylation by native MS.

4.2.2.2.2.   Versatility of the "plug-and-play" reaction

4.2.2.2.2.1.   Isocyanide scope

Having developed a set of robust and modular conditions, the versatility of the U-4CR-3C reaction was investigated at the "plug stage" by screening commercially available or easily accessible isocyanides (**Figure 31**).

The reaction proceeded smoothly with commercial isocyanides. Better average DoC and conversion values were obtained with isopropyl isocyanide **33b**, cyclohexyl isocyanide **33c** and ethyl isocyanoacetate **33e**. In contrast, benzyl isocyanide **33d** showed moderate reactivity and isocyanomethylphosphonate **33f** led to almost no conjugation.

Isocyanides bearing different biorthogonal groups were also synthesized: **33g** was obtained in a single step from **33e** and propargylamine **39** in 60% yield,[330] while isocyanides **33h-l** were easily generated from the *p*-nitrophenol carbonate isocyanide **33m**.[331] Contrasted results were obtained with this set of isocyanides for the conjugation of trastuzumab. Indeed, isocyanides **33h** and **33i** were the only ones giving acceptable average DoC (around 0.6) and conversions rates (around 50%), while the others led to almost no conjugation. Small and apolar isocyanides seemed to be more suitable for this transformation, which could possibly indicate that the reactive location sites are located in hydrophobic pockets.

As cyclohexyl isocyanide **33c** afforded conjugates with increased DoC and conversion values than **33a**, with clean native MS spectra, it was then used as the isocyanide partner for the next conjugation experiments.
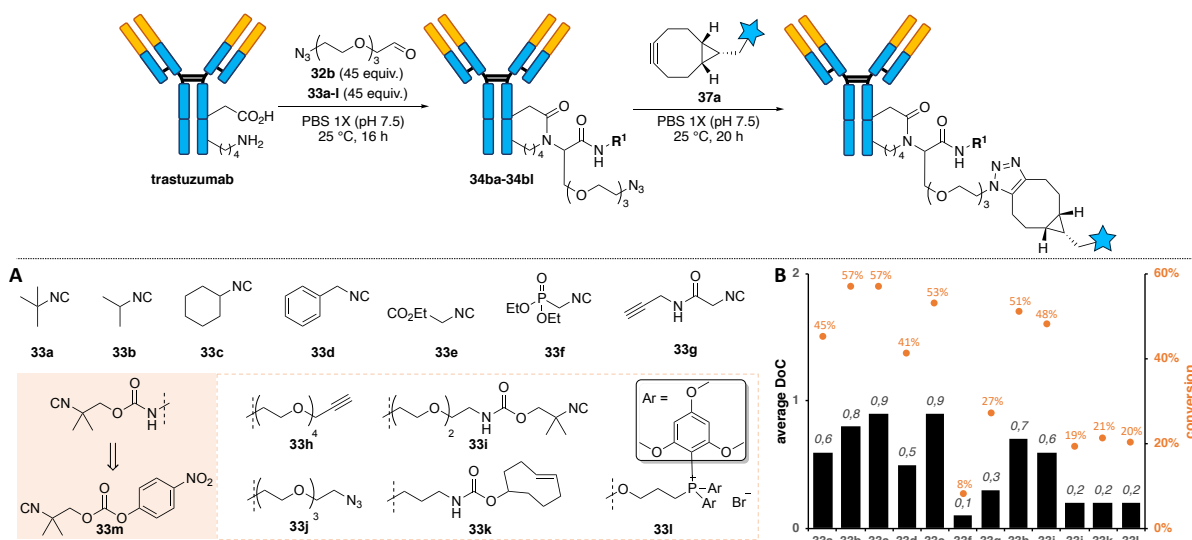
Figure 31: A. Isocyanide scope. Reactions conditions: "plug step": **32b** (45 equiv.), **33a-l** (45 equiv.), PBS 1X (pH 7.5, 137 mM NaCl), 25 °C, 16 h. "play step": **37a** (10 equiv.), PBS 1X (pH 7.5, 137 mM NaCl), 25 °C, 20 h. **B**. Average DoC and conversion rates obtained with **33a-l**.

As done for **33a**, different aldehyde-to-isocyanide ratio (**32b** and **33c**) were again screened for the conjugation of trastuzumab under our optimized conditions (PBS 1X (pH 7.5, 137 mM NaCl), 25 °C, 16 h) (**Figure 32**). As observed with **33a**, low concentration of reagents led to positive results and a plateau was reached when a large excess of aldehydes and isocyanides was used (i.e, ≥ 70:70 equiv. **32b/33c**), with DoC values oscillating around 1.6 and 1.9 and conversion rates above 70%. Good conversion with a narrow DoC distribution – from DoC 0 to DoC 3 – were obtained when 45 equivalents of both **32b** and **33c** were employed. Moving to an aldehyde-to-isocyanide ratio of 60:60 did not significantly improve average DoC and conversion values, and, using a large excess of reagents (i.e., ≥ 70:70 equiv. of **32b/33c**) led to a consequent precipitation of conjugate **38a**, thus making the 45:45 ratio optimal for the "plug-and-play" version of the Ugi reaction.
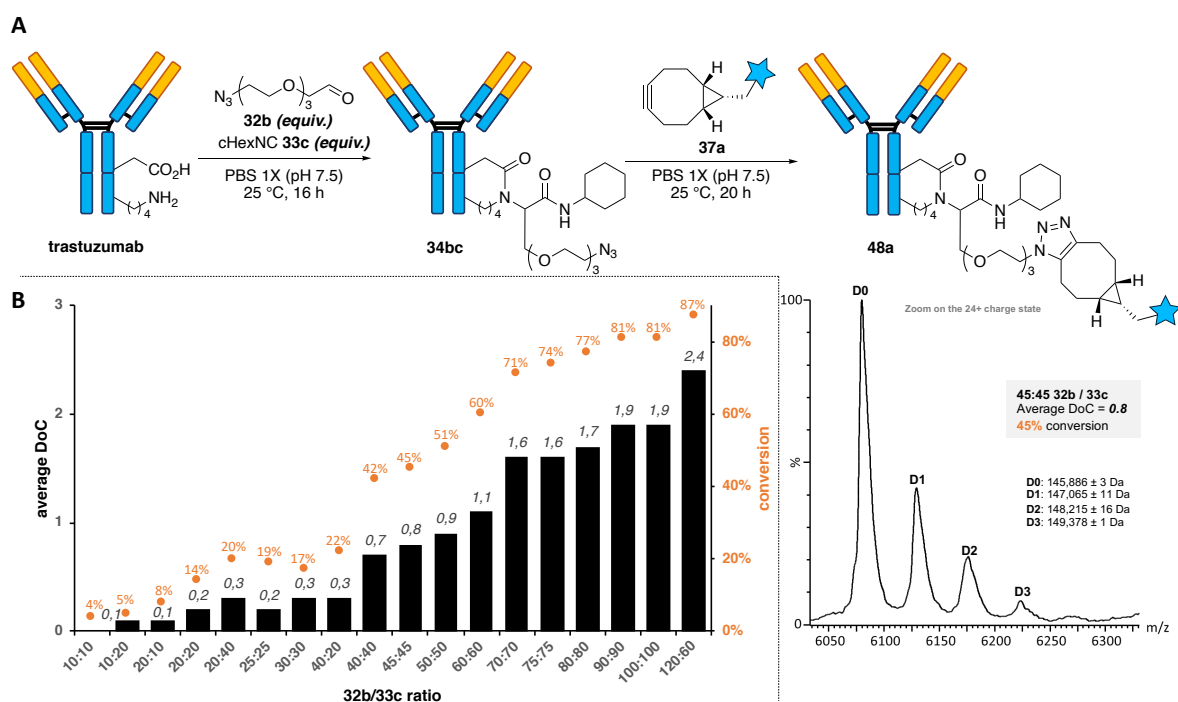

Figure 32: A. Optimization of the "plug-and-play" strategy on trastuzumab with azide-containing aldehyde **32b** and the cyclohexyl isocyanide **33c** followed by the functionalization with strained alkyne **37a** and native MS spectrum of trastuzumab, modified with 45 equivalents of both aldehyde and isocyanide, after deglycosylation (zoom on the 24+ charged state). **B.** Optimization of the "plug step" (average DoC and conversion rates determined by native MS analysis of the corresponding conjugates **37a** after deglycosylation).

In parallel, stability studies were conducted to demonstrate that the payload formed was stable at different pHs and in human plasma (**Figure 33**A-B). Stability studies of conjugate **38a** was

then carried out for five days in different media at 37 °C. Incubated in PBS buffer (1X) at different pHs – 5, 6, 7, 8 and 9 – for 5 days at 37 °C, **38a** showed good stability. Indeed, as observed by native MS, the DoC distribution remained the same even though the average DoC values had slightly decreased compared to the control conjugate (average DoC = 0.8). Incubated in human plasma, no apparent decomposition of **38a** was observed, as demonstrated by conserved fluorescent intensity of the conjugate **38a** by SDS page. Circular dichroism spectroscopy experiments were also conducted on a fresh batch of conjugate **34bc** and demonstrated no influence of the chemical conjugation on the secondary structure of the antibody (**Figure 33**C).
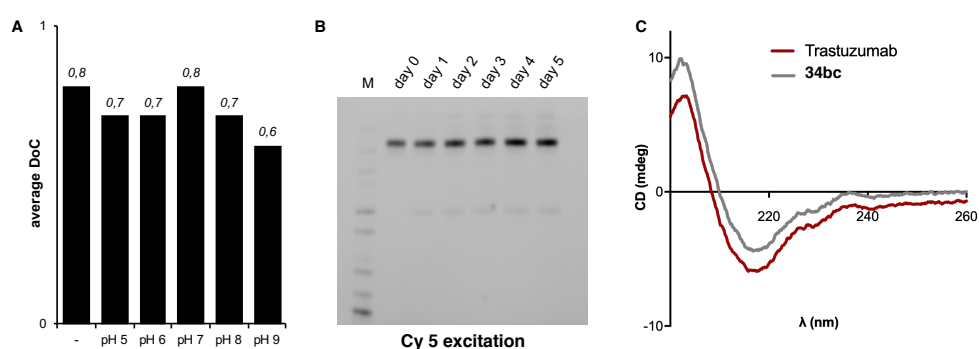
Figure 33: A. stability studies of conjugate **38a**. Average DoC values of **38a** after 5 days at 37 °C in PBS 1X at various pH. **B.** Denaturating SDS page analysis of **38a**'s plasma stability over 5 days at 37 °C (M: molecular weight marker). The gel was revealed by fluorescence (Cy5 excitation). **C.** Circular dichroism spectra of native trastuzumab and conjugate **34bc**

4.2.2.2.2.3. Strained-alkyne scope and dual functionalization

The versatility of the "play step" was investigated with different strained alkynes (**Figure 34**). Starting from the conjugate **34bc**, obtained with 45 equivalents of both **32b** and **33c**, complete conversions for SPAAC were observed with BCN motifs, allowing the introduction of various payloads onto the antibody: fluorophore (**37a**), mass spectrometry and NMR tags (**37b** and **37d,** respectively), iminobiotin (**37c**) or 20-mer oligonucleotides (**37e** and **37f**). With the dibenzocyclooctyne (DBCO) motif, introduction of a cytotoxic drug, monomethyl auristatin E (MMAE) connected to the strained alkyne scaffold by a cleavable valine-citrulline linker **37g**, allowed to generate the antibody-drug conjugate (ADC) **38g**. However, SPAAC was found to be incomplete, and an average drug-to-antibody ratio (DAR) of 1.4 was obtained. This DAR was found to be higher than the ones previously observed for **38a** and could be explained by a lack of precipitation during SPAAC with **37g**.
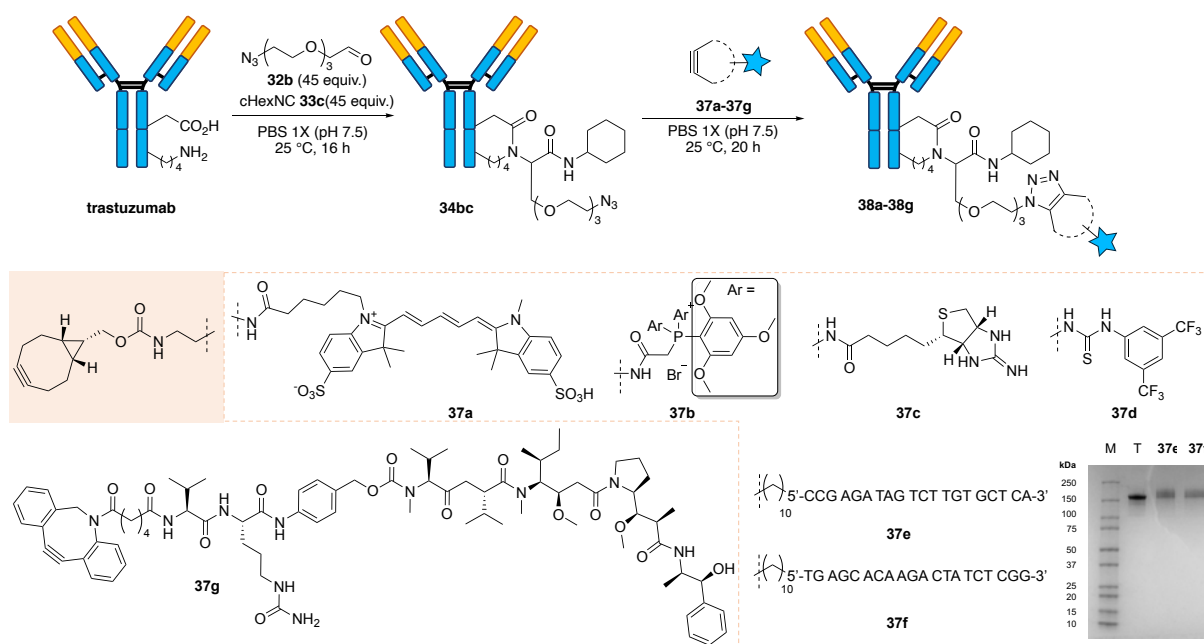
Figure 34: Strained-alkyne scope. Reactions conditions: "plug step": **32b** (45 equiv.), **33c** (45 equiv.), PBS 1X (pH 7.5, 137 mM NaCl), 25 °C, 16 h. "play step": **37a-37g** (10 equiv.), PBS 1X (pH 7.5, 137 mM NaCl), 25 °C, 20 h.

The cytotoxicity of our ADC **38g** was then evaluated by Dr Igor Dovgan on a HER2-positive cell line (SKBR-3) and a HER2-negative cell line (MDA-MB-231) and compared to the FDA-approved trastuzumab-emtansine (T-DM1) (**Figure 35**). The two cell lines were incubated at different concentrations of **38g** and T-DM1 – from 10 pM to 10 nM – for 96 hours before the cell viability was assessed by MTT assay. Active mitochondria allowed cleavage of a yellow tetrazolium salt (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide or MTT) into a formazan product, the absorbance of which was measured and correlated to the metabolic activity, and thus, viability of cells. Pleasingly, our ADC **38g** was found to be at least as active and selective as the benchmark T-DM1, with comparable EC$_{50}$ values on HER2-positive cells (66 pM for **38g** and 174 pM for T-DM1) and lack of apparent toxicity on HER2-negative cells at the tested concentrations.
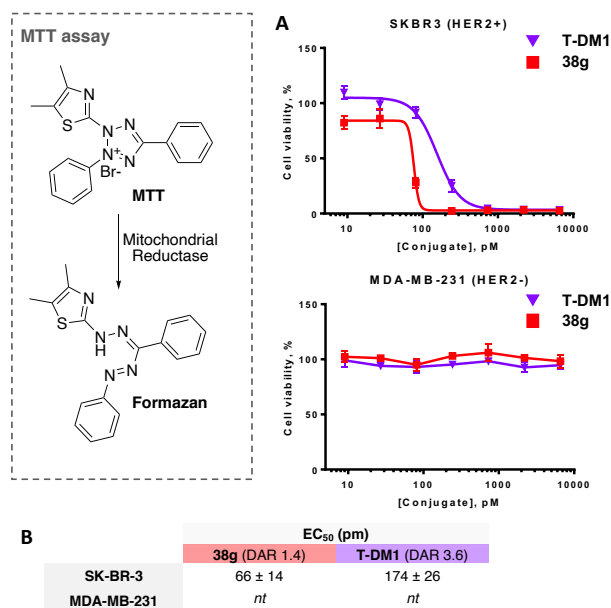
Figure 35: Cytotoxic assay. **A.** ADCs cytotoxicity assay on SKBR-3 and MDA-MB-231 cell lines. **B.** EC$_{50}$ values of ADCs **38g** and T-DM1 in SKBR-3 and MDA-MB-231 cancer cell lines (nt = not toxic at assayed concentration range; EC$_{50}$ ± standard deviations from two independent experiments).

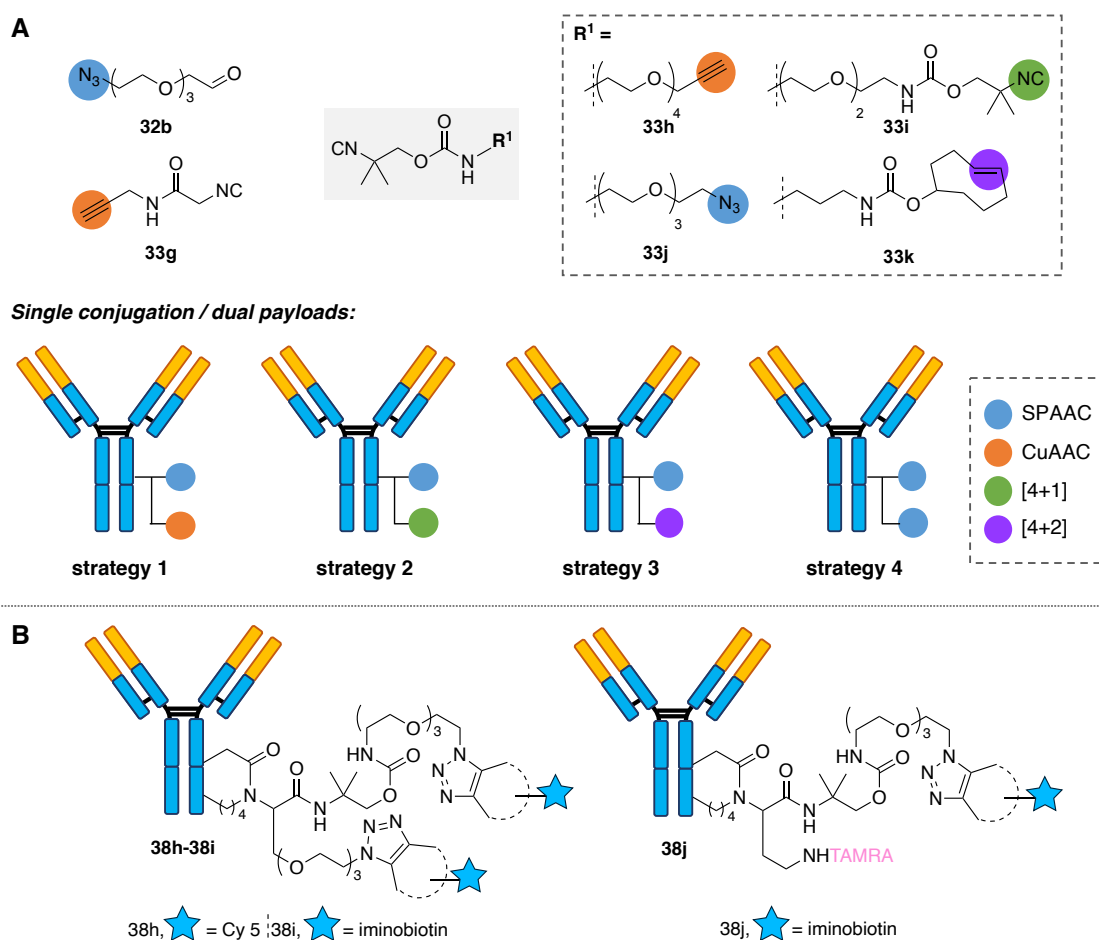#### 4.2.2.2.2.4.   Dual labelling of trastuzumab

While our "plug-and-play" strategy offered a simple and convenient way for the post-derivatization of trastuzumab conjugate **34bc**, the use of the isocyanides **33g-k**, bearing another reactive group, gave an excellent opportunity for dual, and potentially orthogonal, functionalization. To achieve this goal, four strategies were envisioned (**Figure 36**A): 1) SPAAC followed by the CuAAC, by combining aldehyde **32b** with isocyanides **33g** and **33h**; 2) SPAAC followed by the [4+1]-cycloaddition with **33i**; 3) the SPAAC followed by the [4+2]-cycloaddition with **33k**; or 4) in a non-orthogonal manner, with a double SPAAC, using **32b** and **33j**, by adding two equivalents of the same strained alkyne payload.

In the latter case, double labelling of conjugate **34bj** with complete conversion was achieved with alkynes **37a** and **37c**, allowing the incorporation of two identical payloads per conjugation sites (**Figure 36**B – left). Despite numerous attempts, the first three strategies only met with failure.

In the first strategy, while alkynes **37a** and **37c** were successfully incorporated by SPAAC, the terminal alkyne could not be functionalized with CuAAC, either because of partial decomposition (**34bg**) or lack of reaction (**34bh**). In the second strategy, two tetrazines, 4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzoic acid **40** and dimethyl 1,2,4,5-tetrazine-3,6-dicarboxylate

41, were employed with the aim of functionalizing the isocyanide moiety in conjugate **34bi** after SPAAC with the [4+1] retro-[4+2]-cycloaddition cascade leading to *4H*-pyrazole products.[331–333] Unfortunately, both tetrazines proved to be unreactive, even when the [4+1] cycloaddition was attempted prior to SPAAC. The third strategy also employed tetrazine **41** for the modification of *trans*-cyclooctene (TCO) moiety on conjugate **34bk** with the retro-[4+2]-cycloaddition to give 1,4-dihydropyridazine products. In this case, the poor reactivity of the TCO-containing isocyanide **33k** in the Ugi reaction led to too small quantity of resulting conjugate **34bk** to attempt the iEDDA step after the first SPAAC functionalization.

To access dually modified conjugates, an alternative route was proposed. By performing the multicomponent conjugation step with TAMRA-CHO **32a**, which acted as a first functionalization point, and isocyanide **33j**, the resulting conjugate **34aj** was then successfully functionalized by SPAAC with the alkyne **37c**. The resulting conjugate **38j** was labelled with two distinct functional groups: a fluorescent probe and an iminobiotin (**Figure 36**B – right).

Figure 36: A. Envisioned strategies for the dual labelling of trastuzumab. **B.** Homo-functionalization by double SPAAC to give conjugates **38h** or **38i** or heterofunctionalization by the incorporation in **32a** and functionalization with SPAAC to give **38j** conjugate.

4.2.2.2.2.5.   Aldehyde scope

Having screened various isocyanides and strained alkynes, the scope of aldehyde was investigated, using **33i** as a model isocyanide (**Figure 37**). TAMRA-CHO **32a** and commercially available aldehydes **32c-r** were thus evaluated on trastuzumab. Degrees and patterns of substitution were found to play an important role, as demonstrated by linear valeraldehyde **32c** showing poor reactivity compared to its branched analogue **32d**, or by pivaldehyde **32g** giving poor average DoCs as opposed to isobutylraldehyde **32e** or methyl isopropyl ketone **32f**. As already observed with cyclohexyl isocyanide **32c**, cyclohexanecarboxaldehyde **32h** gave better conversion and narrower DoC distribution (DoC 1 to DoC 5) than **32e**. More polar aldehydes **32i-l** gave mixed results: both glucose **32k** and glutaraldehyde **32l** failed to modify trastuzumab, while formaldehyde **32i** and glyceraldehyde **32j** performed very well. However, for the latter, mass spectra of poor quality were obtained, possibly indicating some side reactions. The 2-pyridinecarboxaldehyde **32m** was the only aromatic aldehyde giving positive results on trastuzumab. The five other aromatic aldehydes tested, **32n-r**, led at best to mediocre results (maximum average DoC of 0.2). For **32q** and **32r**, those results could be explained by the lack of the solubility of the compounds in buffers. However, for **32n** and **32o**, their inefficacies were disconcerting as they had been reported for the functionalization of proteins.[209,210,334] Conversion rates of **32n** and **32p** were respectively increased to 34% (average DoC of 0.3) and 62% (average DoC of 0.8) when 100 equivalents of both **32n-32p** and **33i** were used. Taken together, these results suggested that small and flexible apolar carbonyls and isocyanides were more appropriate for this reaction.
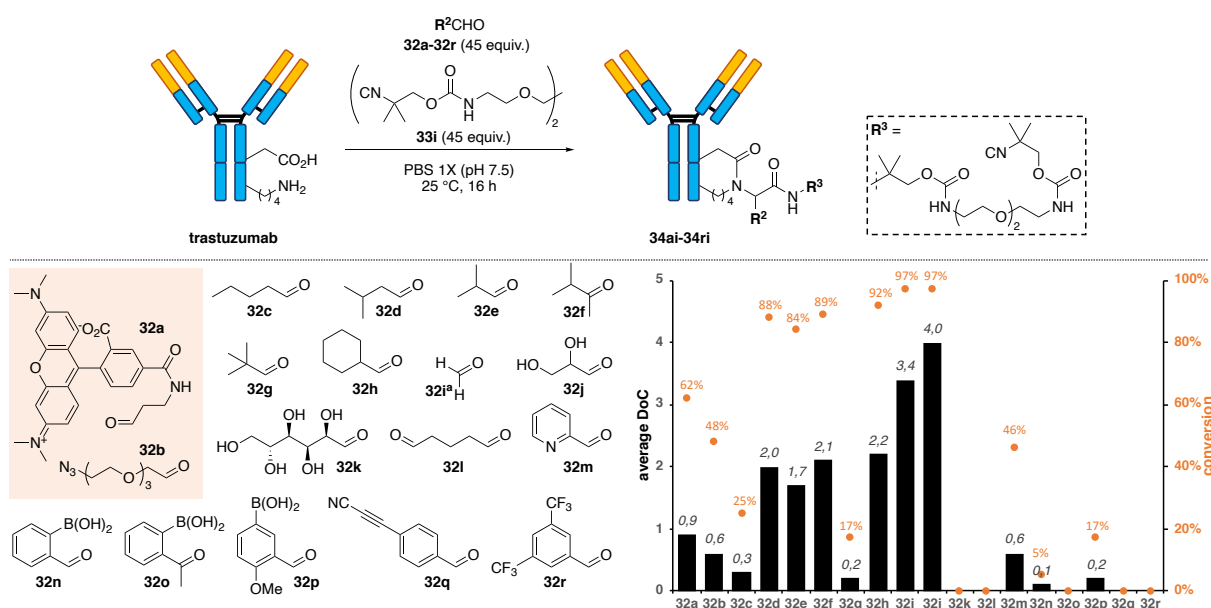
Figure 37: A. Aldehyde scope of the multicomponent bioconjugation reaction. Reaction conditions: **32a-r** (45 equiv.), **33i** (45 equiv.), PBS 1X (pH 7.5, 137 mM NaCl), 25 °C, 16 h; [a] 5 equiv. used. **B.** Average DoC and conversion rates obtained with **32a-r**.

4.2.2.2.2.6.   The Ugi reaction on proteins

Having screened various aldehydes and isocyanides, the conditions of our multicomponent reaction were applied to other proteins (**Figure 38**).

The reaction was first investigated on smaller proteins – ubiquitin **42a** (8.5 kDa), cytochrome c **42b** (12.0 kDa), lysozyme **42c** (14.3 kDa), myoglobin **42d** (17.2 kDa), α-chymotrypsin **42e** (25.0 kDa) and human serum albumin **42f** (HSA, 66.5 kDa) – which were incubated separately with 45 equivalents of both **32a** and **33c** in PBS 1X (pH 7.5, 137 mM NaCl) for 16 hours at 25 °C. All the proteins tested – except α-chymotrypsin **42e** – showed conjugation with an apparent lack of correlation between the size of the proteins and the average DoC and conversion values. Interestingly, the conjugation of HSA **42f** gave only monofunctionalized adducts (DoC 1 species), a remarkable feature considering the size of the protein and the conversion obtained. LC-MS/MS studies are currently performed to identify the sites of modification on this protein.

The multicomponent strategy was then tested on the anti-VEGFR2 monoclonal antibody, ramacirumab **42g**, with 45 equivalents of both **32b** and **33c** under our optimized conditions. Analyses by native MS showed conjugation for ramacirumab **42g**, with an average DoC of 0.6 and a conversion of 53% observed – values that were close to those reported for trastuzumab.
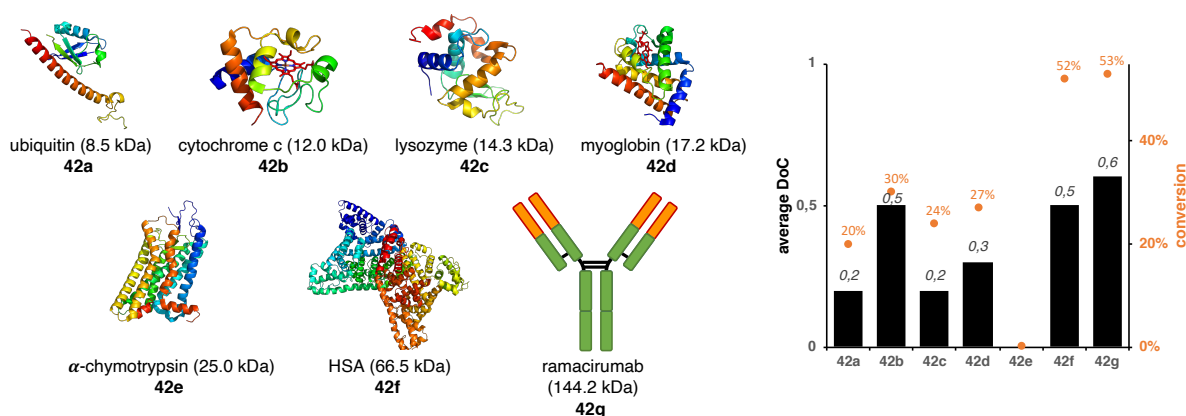
Figure 38: Protein scope. The average DoC and conversion rates were reported and obtained by native MS.

Given that some proteins were more prone to conjugation than others, we decided to evaluate a potential protein selectivity. Trastuzumab was incubated with one or two proteins with 45 equivalents of both **32a** and **33c** in PBS 1x (pH 7.5, 137 mM NaCl) for 16 hours at 25 °C. After 16 hours, hydroxylamine **12** was added to hydrolyze remaining imines and samples were then analyzed by SDS page (**Figure 39**). By fluorescence, in all cases, trastuzumab was labelled, however in some cases it was not possible to conclude if the second or third protein added was conjugated. Altogether, these results suggested that a protein selectivity could not be achieved with this reaction as fluorescent signal was detected for all proteins.
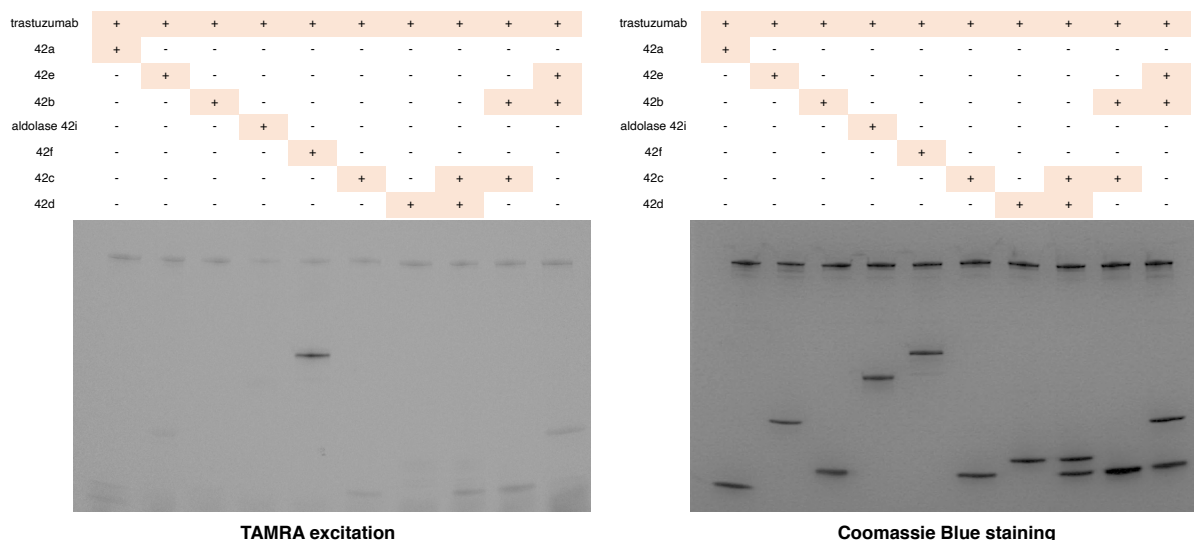

Figure 39: Application of the Ugi reaction in a mixture of two or three proteins – trastuzumab, ubiquitin **42a**, α-chymotrypsin **42e**, cytochrome c **42b**, aldolase **42i**, HSA **42f**, lysozyme **42c** and myoglobin **42d**. Reaction conditions: **32a** (45 equiv.), **33c** (45 equiv.), PBS 1x (pH 7.5, 137 mM NaCl), 25 °C, 16 h. The samples were analyzed by SDS page (0.2 mg/mL concentration of proteins). The gel was revealed by fluorescence and Coomassie blue.

4.2.2.3. Mechanistic investigations

Having developed an efficient, operationally simple and quick multicomponent bioconjugation strategy that tolerated various functional groups and substrates, the mechanism taking place in this strategy and the identification of the conjugation sites on trastuzumab were explored. To answer these questions, three experiments were conducted: a competition experiment to identify the types of amino acids involved in the conjugation; a NMR study with the aim of differentiating between the Ugi and Passerini payloads, and; a LC-MS/MS study to find the conjugation sites and also discriminate the multicomponent reactions taking place.

4.2.2.3.1. Competition experiments

With the will to better understand which amino acids participated in the multicomponent reaction, we added free amino acids to the reaction mixture on trastuzumab. Indeed, if a decrease in average DoC were to be observed by fluorescence, it might indicate that the amino acid added partook in the reaction. Eight amino acids, all protected in *C*- and *N*-terminus, were selected – Ac-Lys-NHMe **43**, Ac-Glu-NHMe **44**, Ac-Trp-NHMe **45**, Ac-Leu-NHMe **46**, Ac-Phe-NHMe **47**, Ac-Gly-NHMe **48**, Ac-His-NHMe **49** and Ac-Tyr-NHMe **50** – and added in large excess (1000 equiv.) to trastuzumab before it was reacted with 45 equivalents of both **32b** and **33c** under the optimized conditions. Surprisingly, Ac-Glu-NHMe **44** was the only amino acid moderately competing in the reaction (**Figure 40**). The seven others, including lysine, did not compete, as observed with the conservation of fluorescence. Adding simultaneously 1000 equivalents of **43** and **44** was also not found to be successful. Altogether, these results did not provide clear evidence of the amino acids involved in our reaction.
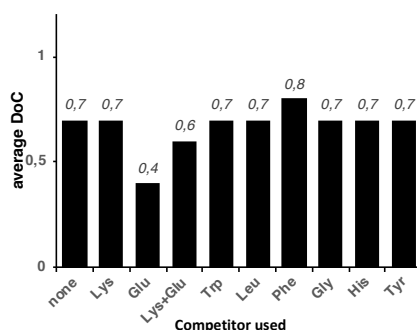
Figure 40: Competition experiments. Influence of a competitor on the average DoC. Trastuzumab was reacted with 45 equivalents of both **32b** and **33c** in presence of 1000 equivalents of a competitor in PBS 1X (pH 7.5), 25 °C, 16 h, and further functionalized with **37c**.

4.2.2.3.2.   NMR studies

We then wondered if we could use NMR spectroscopy to identify the preferred payload formed – Ugi or Passerini – in the multicomponent reaction on trastuzumab. Indeed, a small difference in ^{13}C chemical shift is expected on the carbon brought by the isocyanide moiety, for each conceivable product's structures – amide bond formed for the Ugi pathway versus ester bond formed for the Passerini reaction. To improve the sensitivity of this carbon, cyclohexyl isocyanide enriched with ^{13}C **51** was synthesized and used for the modification of trastuzumab. Trastuzumab was then labelled with 45 equivalents of both **32b** and **51** under the optimized conditions and a similar average DoC value (0.8) was obtained with the enriched isocyanide, confirming that moving to ^{13}C did not have an impact on the reaction. Prior to NMR experiment, the conjugated and enriched antibody was desalted with water to prevent a decrease in sensitivity of the NMR probe and concentrated to 2.10^{-4} M. Similarly, under the same conditions, the ^{13}C NMR spectrum of the native antibody was recorded, to be able to compare the spectra of the two experiments. Despite the use of **51**, no additional peaks were observed probably due to the high abundance of amide bonds on the antibody.

To circumvent this issue, it was decided to use a smaller protein, ubiquitin **42a** (8.5 kDa). It was expected that with a smaller protein – that possess much less amide bonds – there would be greater chance to observe a signal by NMR spectroscopy. Ubiquitin **42a** (8 mg) was then modified with 5 equivalents of ^{13}C/D-formaldehyde – to prevent the precipitation of the protein – and 45 equivalents of **51**. By performing a 2D ^{13}C-^{13}C INADEQUATE experiment, it was expected to differentiate more easily which of the payloads was formed, as a difference of 10 to 20 ppm was expected if the ^{13}C, coming from aldehyde, was located next to an amide (i.e. Ugi payload) or an ester (i.e. Passerini payload). Unfortunately, as for the previous experiment, it was not possible to observe enriched peaks coming from the payload.

As NMR studies could not provide information regarding the mechanism of our multicomponent reaction on proteins, we turned to LC-MS/MS studies.

4.2.2.3.3.   LC-MS/MS studies

To validate the mechanism of the bioconjugation reaction and localize the exact conjugation sites, and finally establish whether our multicomponent strategy can be called site-specific, LC-MS/MS studies were performed. The peptide mapping was then performed on three separate batches that were prepared and purified in different manners. A first batch of **38a** with an average DoC of 0.6 was generated and analyzed without further purification (protein

sequence coverage: 100%). Due to the high abundance of unconjugated trastuzumab (60%), thus giving unconjugated peptides that outcompeted the conjugated fragments for mass detection, two enriched samples were produced by two similar but distinct routes (protein sequence coverage: 31-48%). A first strategy used conjugate **38c**, obtained from **34bc** – itself obtained with 100 equivalents of both **32b** and **33c** – and BCN-iminobiotin **37c**. Trypsin digestion followed by affinity purification of the digestate on streptavidin gave a first enriched batch. The second strategy inverted the order of events: conjugate **34bc** was digested with trypsin before functionalization with **37c**. The resulting mixture of peptides was enriched in the same manner as for the first strategy (more details in the experimental part: **Figure 49**).

The peptide mapping analyses highlighted six conjugation sites, only two of which were detected in all three batches: the *N*-terminal glutamate (E1) of the heavy chain and the *N*-terminal aspartate (D1) of the light chain (**Figure 41**). Interestingly, those two residues partook in two different types of multicomponent reaction: an intramolecular U-4C-3CR, giving β- and γ-lactams (from D1 and E1 respectively) involving the α-NH_2 and the side chain carboxylate, and a Passerini reaction, involving only the side chain of the amino acid (detected only in the two enriched samples for D1, but in all three batches for E1). Four other conjugation sites were detected, but only in a single batch and thus could not be cross validated. Three of them corresponded to glutamate residues modified by the Passerini reaction and were either located on the light chain E195 or on the heavy chain E236 and E348. The last conjugation site matched with the expected inter-residue U-4C-3CR and was located on the light chain and involved the side chains of K126 and either D122 or E123 (undistinguishable from each other). Interestingly, the *C*-terminal lysine (K478) located on the heavy chain did not react in either a U-4C-3CR or a Passerini reaction and was found to be unaffected.
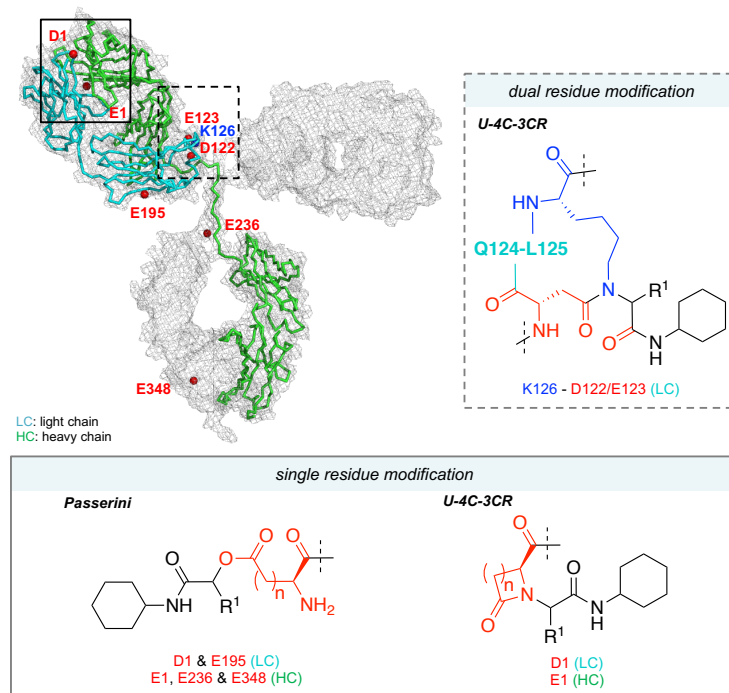
Figure 41: Trastuzumab conjugation sites determined by peptide mapping.

4.3.   Conclusion

With the aim of developing a new site-specific method for the conjugation of proteins, we reported a multicomponent strategy based on the Ugi four-component three centre reaction. By combining commercially available – or easily accessible – aldehydes and isocyanides, this approach gave a rapid and straightforward access to antibody – and protein – conjugates with different payloads, such as the ADC **38g**, that retained its selectivity towards HER2-positive cancer cell lines and was found to be as active as the benchmark T-DM1. Two multicomponent reactions were found to compete with each other in the conjugation process, the expected Ugi four-component three-centre reaction and the Passerini reaction. While both reactions occurred at the *N*-terminal aspartate and glutamate position of trastuzumab, the expected inter-residue U-4C-3CR was identified on the light chain of the antibody, and, the Passerini reaction was found to label three other glutamate residues. Most notably, this strategy represented the first example of an aspartate/glutamate selective conjugation method that did not require activated agents for their conversion into activated esters. These results offered a new approach towards a site-specific bioconjugation strategy on native proteins and highlighted the potential of multicomponent reactions.

To improve the chemoselectivity of our multicomponent strategy towards aspartate and glutamate residues, and hence the global selectivity of the reaction – i.e. only the Ugi four-component reaction –, it was proposed to mix our antibody with imine or oxime derivatives in combination with an isocyanide to solely modify its aspartate / glutamate residues. This strategy – still under investigation in the laboratory – would then limit the type of amino acid residues being modified as well as the type of MCR being now possible.

GENERAL CONCLUSION AND PERSPECTIVES

During this thesis, three different strategies for the modification of proteins were explored. The first approach aimed for the chemoselective labelling of cysteine residues. The other two aimed for regioselective modification by targeting either lysines in close spatial proximity to pre-conjugated cysteine residues, or pairs of residues that can be simultaneously involved in a multicomponent reaction.

Hypervalent iodine reagents, and in particular, ethynyl benziodoxolone (EBX) derivatives were first investigated for the chemoselective conjugation of cysteine residues on the antibody trastuzumab. While azido-EBX reagents **1** were found to be reactive on the antibody, they also demonstrated a poor thiol-selectivity and the newly-formed payload showed some signs of degradation, making them unsuitable to generate homogeneous antibody-conjugates. Switching to the analogous reagent, TMS-EBX allowed ethynylation of cysteine residues. Through a fine tuning of the reactivity of this reagent, antibody-conjugates were obtained with good conversion and DoC values, and with excellent thiol-selectivity.

To achieve a site-selective modification of lysine residues on antibodies, via a cysteine relay, an approach analogous to the native chemical ligation was explored. Cysteine residues were first labelled with thioester-containing molecule through a *trans*-thioesterification. The payload was then transferred to a vicinal lysine residue via a *S*-to-*N* acyl shift. The thioesters that did not partake in the transfer were hydrolyzed with L-cysteine before the antibody was reoxidized in a last step. The different steps involved were first optimized independently, before the full sequence was applied on the antibody. Unfortunately, a fully reconstructed antibody was never obtained, and before the reasons behind this non-reoxidation were explored, a competing publication forced us to stop this project.

In a last part, a multicomponent reaction – the Ugi four-component three-centre reaction (U-4C-3CR) – was explored as a new regioselective labelling strategy. After a short optimization, this approach gave a rapid and facile access to protein conjugates with different payloads by employing various commercially available – or easily accessible – aldehydes and isocyanides. Using LC-MS / MS studies, six conjugation sites were identified on the antibody trastuzumab, with two multicomponent reactions found to compete with each other in the conjugation

process: the expected U-4C-3CR and the Passerini reaction. Both reactions were found to occur on the *N*-terminal aspartate and glutamate position of trastuzumab, while the expected inter-residue U-4C-3CR was identified on the light chain of the antibody. Three additional glutamate residues were also found to be modified via a Passerini reaction. In addition to providing conjugates with high stability, this reaction allowed us to generate an ADC which was as active as the benchmark T-DM1. Even though this method is not strictly speaking site-selective, since six conjugation sites were identified, it still highlights the potential of multicomponent reactions for bioconjugation and opens the road to new projects aiming to modify aspartate and glutamate residues.

By combining these last two projects, we believe a highly regioselective modification on trastuzumab to be reachable. Indeed, when trastuzumab was modified with U-4CR-3CR, one site of modification – K126 and D122/E123 – was found to be close to a cysteine residue involved in the bridging between the light and heavy chain of the antibody. It would thus be conceivable to selectively modify this site by pre-orienting the aldehyde – or isocyanide – component of the U-4C-3CR from this cysteine.

As developed in the second project, the aldehyde – or isocyanide – installed onto a thioester-containing compound, would first be transferred to the cysteine residues via a *trans*-thioesterification (**Figure 42**). In a second step, addition of an isocyanide – or aldehyde – would lead to an Ugi reaction, modifying only the site K126-D122/E123 thanks to the cysteine pre-orientation. The thioester would then be hydrolyzed using free L-cysteine, to allow for the antibody re-oxydation, as well as the introduction of a free thiol on the payload. The site-selectively installed handle would then be modified by thiol-selective methods described in this manuscript yielding a functionalized antibody. As all these steps have already been optimized over this thesis, the main challenge of this strategy would come from the optimization spacer's length as the cysteine relay step will rely on the distance between the thioester and the aldehyde – isocyanide – groups. To conclude, this strategy would give access to Thiomab® related conjugates made from native antibody.
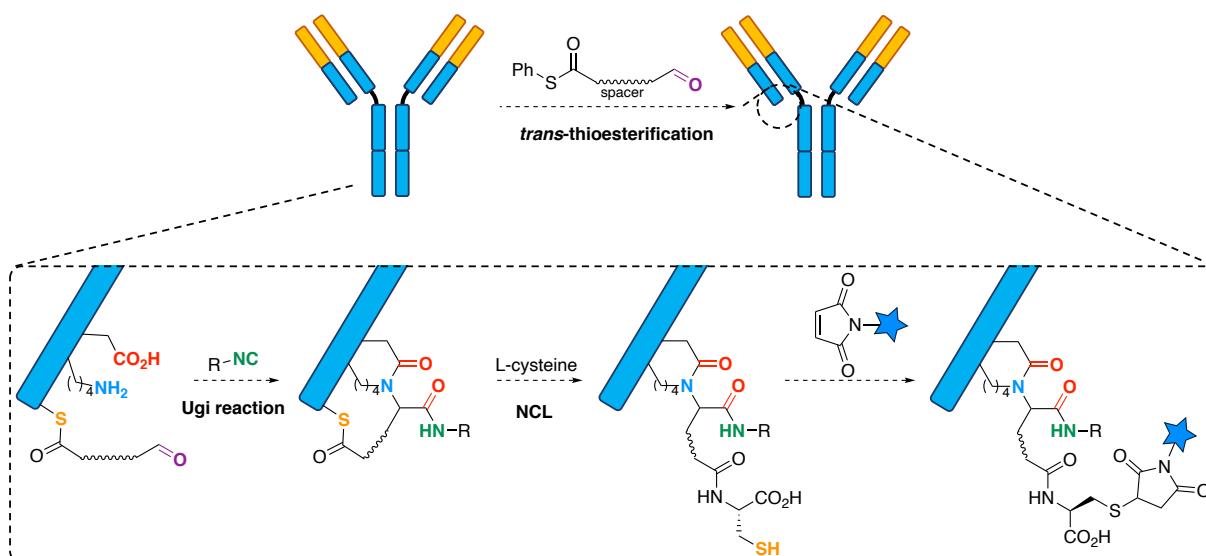
Figure 42: Aim of the "2 in 1 project": towards the labelling of K126 and D122 or E123 residues on trastuzumab with the Ugi reaction following the pre-orientation of an aldehyde via a cysteine residue.

EXPERIMENTAL PART

6.1.   Material and methods

Synthetic chemistry

All reagents were obtained from commercial sources and used without any further purifications. Amino-modified oligonucleotides were purchased from IDT. Dry solvents were obtained from Sigma-Aldrich. Column chromatography was carried out as "Flash Chromatography" using silica gel G-25 (40-63 µm) from Macherey-Nagel. Thin layer chromatography (TLC) was performed using plates cut in aluminum sheets (ALUGRAM Xtra SIL G/UV254) purchased from Macherey-Nagel. Visualization was achieved under a 254 or 365 nm UV light and by using an appropriate TLC stain.

Spectroscopy and spectrometry

¹H and ¹³C NMR spectra were recorded at 23 °C on Bruker Avance III - 400 MHz / 500 MHz spectrometers. Recorded shifts are reported in parts per million (δ) and calibrated using residual non-deuterated solvent. Data are represented as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad, app = apparent), coupling constant (*J*, Hz), integration and assignment in case of ¹H NMR data.

Analytical LC-MS analyses were carried out on Waters 2695 separations module equipped with Waters 2487 UV detector, Waters Acquity QDa mass detector and CORTECS, 2.7 µm, C18, 50 x 4.6 mm column. The flow rate was 1 mL/min and the solvent system was composed as follows: solvent A: 0.05% HCOOH in water; solvent B: acetonitrile. The gradient run was: 0 to 5 min. – 5% to 95% B; 5 to 6 min. – 95% B; 6 to 7.8 min. – 5% B. Mass detector was operated in positive MS Scan mode with 600 °C probe temperature, 1.5 kV capillary voltage and 10 V cone voltage.

High-resolution mass spectra (HRMS) were obtained using an Agilent Q-TOF 6520.

IR spectra were recorded in a Thermo-Nicolet FT/IR-380 spectrometer. Spectra were interpreted with OMNIC 9 software and are reported in cm^{-1}. The abbreviations used are: w (weak), m (medium), s (strong) and br (broad).

Melting points were carried out on a melting point apparatus SMP3 from Stuart Scientific.

Native mass spectrometry (native MS) analyses of intact proteins and conjugates were performed using a mass spectrometer (described below) coupled to an automated chip-based nanoESI infusion source (Triversa Nanomate, Advion, Ithaca, NY) both operating in positive ion mode. Electrospray ionization was conducted at a capillary voltage of 1.75 kV and nitrogen nanoflow of 0.75 psi. Samples were directly infused after manual desalting step at a concentration of 10 μM. Average DoC values were calculated using equation (1):

$$DoC = \frac{\left(\sum_{k=0}^{8} k \times \text{intensity DoC}k\right)}{\sum_{k=0}^{8} \text{intensity DoC}k} \quad (1)$$

These results were derived from the relative peak intensities measured from deconvoluted mass spectra, following a method developed and validated by the Sarah Cianférani group.[335]

Circular dichroism (CD) measurements were recorded using a Jasco J-820 spectrophotometer equipped with Jasco PTC-4232 temperature control unit (25 °C). Spectra were acquired in a 0.1 cm quartz suprasil cell (Hellma) in a wavelength range from 200 to 260 nm at a scanning speed of 20 nm/min, with a bandwidth of 1.0 nm, a per-point acquisition delay of 2 s, a data pitch of 0.1 nm and a sensitivity going from -10 mdeg to +10 mdeg. Spectra were averaged over three scans and the spectrum from a blank sample containing only the buffer was subtracted from the averaged data.

List of mass spectrometers used for native MS:

● **LCT (Waters, Manchester, UK).** The extraction cone value was set to 5 V or 50 V and the cone voltage was set to 120 V or 180 V for cysteine-conjugates and lysine-conjugates, respectively. The pressure in the interface region was fixed at 6 mbar. Acquisitions were performed in the m/z range 1,000–10,000 with a 4 s scan time. External calibration was performed using singly charged ions produced by a 2 g/L solution of cesium iodide in 2-propanol/water (50/50 v/v). MS data interpretations were performed using Mass Lynx V4.1 (Waters, Manchester, UK).

● **Synapt G2 HDMS (Waters, Manchester, UK).** The extraction cone value was set to 5 V and the cone voltage was set to 120 V or 180 V for cysteine-conjugates and lysine-conjugates, respectively. The pressure in the interface region was fixed at 6 mbar. Acquisitions were performed in the m/z range 1,000–10,000 with a 4 s scan time. External calibration was performed using singly charged ions produced by a 2 g/L solution of cesium iodide in 2-propanol/water (50/50 v/v). MS data interpretations were performed using Mass Lynx V4.1 (Waters, Manchester, UK).

● **Exactive Plus EMR (Thermo Fisher, Bremen, Germany).** The in-source collision-induced dissociation (CID) was set to 75 eV for cysteine-conjugates and optimised between 75 - 150 eV for lysine-conjugates. The higher-energy collisional dissociation (HCD) cell was set to 10 eV for each analyses to improve the desolvation. The trapping gas pressure was set to 7 a.u. (which corresponds to an Ultra High Vacuum of 1.10^{-9} mbar). To improve the transmission of the high mass species, the voltages on the injection-, inter-, and bent-flatapoles were fixed to 8, 7, and 6 V, respectively. Acquisitions were performed in the m/z range 1,000–10,000 with a 3 s scan time and a resolution of 17,500 at 200 m/z with an automatic gain control (AGC target) fixed to 1.10^{6} and a maximum injection time set to 100 ms. External calibration was performed using singly charged ions produced by a 2 g/L solution of cesium iodide in 2-propanol/water (50/50 v/v). Orbitrap MS data interpretation was performed using BioPharmaFinder 2.0 (Thermo Fisher Scientific, Bremen, Germany).

Biomolecules

Protein concentration of protein solutions was determined by UV absorbance using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Illkirch, France). The concentration of antibody conjugates was measured using a BCA Protein Assay Kit (Ref. 23225, Thermo Fisher Scientific, Illkirch, France). Bio-spin P-30 and P-6 columns were obtained from Bio-rad (Hercules, U.S.A) and micro-concentrators (Vivaspin 500 μL, 50 kDa, 30 kDa, 10 kDa and 3 kDa cutoff) from Sartorius (Gottingen, Germany). Antibody deglycosylation was achieved with Remove-iT® Endo S (New England Biolabs, Ipswich, U.S.A). Purifications of modified oligonucleotides were carried out on a Shimadzu system (pumps: LC 20-AD, detector: SPD 20-A, autosampler: SIL 20-A) using a SunFire™ C18 5 μM 4.6 × 150 mm column (Waters) with a flow rate of 1 mL/min and a gradient going from 10 to 40% of mobile phase B over 30 min. Composition of mobile phases were as follow: phase A was triethylammonium acetate (TEAA)

in water (50 mM); phase B was TEAA in acetonitrile (50 mM). The detection was done at 260 nm. Human plasma was supplied by *Établissement Français du Sang* (EFS Strasbourg).

In vitro ADC cytotoxicity assay

SK-BR-3 (HER2-positive) and MDA-MB-231 (HER2-negative) cell lines were grown in DMEM (Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% fetal bovine serum (FBS), Penicillin (100 units/mL), and Streptomycin (100 µg/mL). Cell lines were maintained in a 5% CO_2 humidified atmosphere at 37 °C. The day before experiment, cells were seeded in 96-well plates at 6000 cells/well in 100 µL fresh cell medium. The day of experiment, serial three-fold dilution of conjugates with cell medium was performed. The medium was carefully removed from the plated cells and the cells were incubated with conjugates (100 μL/well, in triplicate) for 96 h. The MTT reagent (Sigma-Aldrich, 100 µL, 1 mg/mL in cell media) was then added into each well and cells were incubated for 1.5 h at 37 °C. The medium was removed carefully and the blue precipitate was solubilized by adding 100 µL of DMSO. Cell viability was measured by quantifying absorbance at 570 nm using a 96-well plate reader (Flx-Xenius XM, Safas, Monaco). EC50 values were determined using four-parameter logistic fitting in GraphPad Prism 7.0. EC50 ± SDs were calculated from two independent experiments.

6.2. General procedures

Reduction of trastuzumab:

To a solution of trastuzumab (5 mg/mL, 1 equiv., 100 µL in PBS 1X with 1% EDTA, pH 7.5) was added TCEP (20 mM in H_2O, 10 equiv., 1.71 µL). The reaction mixture was then incubated for 3 hours at 37 °C, before the excess of reagent was removed by gel filtration chromatography on Bio-Spin P-30 columns pre-equilibrated with PBS 1X (pH 7.5) to give a solution of reduced trastuzumab.

Procedure for the SDS-PAGE analysis:

Non-reducing glycine-SDS-PAGE was performed on 4 – 15% Mini-PROTEAN® TGXTM gel (Bio-Rad, Hercules, U.S.A) following standard lab procedures. To the samples containing antibody conjugates (24 μL, 0.1 mg/mL solution in H_2O) was added 8 μL of loading buffer

(Laemmli SDS sample buffer) and heated at 95 °C for 3 min. The gel was run at constant voltage (200 V) for 35 min using TRIS 0.25 M – Glycine 1.92 M – SDS 1% as a running buffer. The fluorescence was visualized on GeneGenius bio-imaging system prior to staining with Coomassie blue.

Samples preparation for native mass spectrometry:

Prior to native MS analyses, antibody conjugates were deglycosylated by incubating (37 °C – 2 h) 0.4 unit of Remove-iT® Endo S per microgram of antibody conjugates. The antibody conjugates were then desalted against 150 mM ammonium acetate solution buffered at pH 7.4 using ten cycles of concentration/dilution on Vivaspin centrifugal concentrators (500 μL, 50 kDa). Protein concentration was determined by UV absorbance using a NanoDrop spectrophotometer.

Trypsin digestion protocol:

Trastuzumab conjugate was concentrated to 40 μL to reach a final concentration of 21 mg/mL. The solution was then diluted with 200 μL of a 50 mM NH_4HCO_3 buffer containing 0.1% of *Rapi*Gest™. Dithiothreitol (DTT; 0.1 M solution in 50 mM aqueous NH_4HCO_3, 8 μL) was added and the sample was incubated 30 min at 60 °C. After reduction of the disulfide bonds, the sample was cooled to room temperature and iodoacetamide (IAA; 0.1 M solution in 50 mM aqueous NH_4HCO_3, 16 μL) was added to alkylate the free thiols. The sample was then placed in the dark for 40 min at 25 °C. Trypsin was resuspended in 100 μL of trypsin buffer and heated at 30 °C for 15 min before 32 μL were added to the conjugate. The resulting sample was then incubated overnight at 37 °C before being concentrated to dryness the next day.

Peptide enrichment by affinity chromatography:

The streptavidin column was pre-equilibrated with 10 mL of binding buffer (PBS 1/20x, pH 10; flow rate: 1 mL/min). The digestion sample diluted in 1 mL of binding buffer was loaded onto the streptavidin column, which was then incubated 30 min at 25 °C. The streptavidin column was then washed 2 times with 10 mL of H_2O (flow rate: 1 mL/min) and left overnight at 5 °C. The next day, a washing was done every hour with 10 mL of buffers (flow rate: 1 mL/min; 2 x PBS 1/20x (pH 7.5), 2 x PBS 1x (pH 7.5), 1 x H_2O). The conjugated fragments were finally recovered by washing 2 times the streptavidin column with 10 mL of elution buffer (50 mM

solution of NH$_4$OAc in H$_2$O, pH 4; flow rate: 1 mL/min). The first 5 mL were concentrated to dryness and then resuspended in 50 μL in a buffer containing H$_2$O/ACN/HCOOH (98/2/0.1) for LC-MS/MS analysis.

6.3.   Chemical syntheses and characterizations

**2-(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)-5-
((phenylthio)carbonyl)benzoate 9**
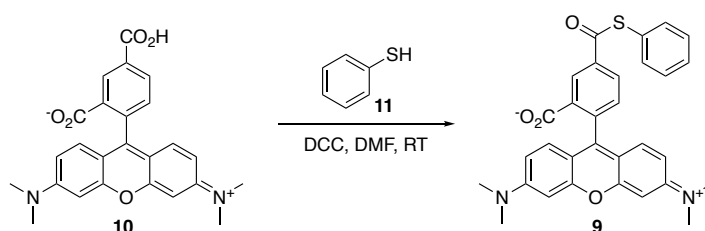
To a solution of TAMRA-CO$_2$H **10** (50 mg, 0.12 mmol, 1 equiv.) and thiophenol **11** (13 μL, 0.13 mmol, 1.1 equiv.) in DMF (3 mL) was added DCC (36 mg, 0.17 mmol, 1.5 equiv.) The reaction mixture was stirred 24 hours at room temperature before being dried under reduced pressure. The thioester **9** was purified by the reverse phase chromatography and was obtained as purple solid with 23% yield (14 mg, 0.027 mmol).

¹H NMR (500 MHz, DMSO-*d*$_6$) δ 8.43 (d, *J* = 1.7 Hz, 1H), 8.32 (dd, *J* = 8.1, 1.7 Hz, 1H), 7.62 – 7.55 (m, 5H), 7.45 (d, *J* = 8.0 Hz, 1H), 6.60 (d, *J* = 8.7 Hz, 2H), 6.53 – 6.49 (m, 4H), 2.96 (s, 12H).

HRMS (ESI⁺) calcd for C$_{31}$H$_{26}$N$_2$O$_4$S [M+H⁺] 523.1686; found 523.1698.

S-phenyl 2-azidoethanethioate 16
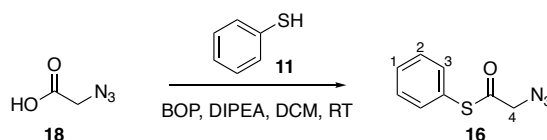
2-Azidoacetic acid **18** (50 mg, 0.49 mmol, 1 equiv.) and thiophenol **11** (50 μL, 0.49 mmol, 1 equiv.) were stirred with BOP (229 mg, 0.52 mmol, 1.05 equiv.) and DIPEA (0.16 mL, 0.95 mmol, 1.93 equiv.) overnight in DCM (2 mL). The next day, the reaction was quenched

with NaHCO$_3$ (5 mL, aq., sat.). The layers were separated and the organic layer was sequentially washed with NH$_4$Cl (5 mL, aq., sat.) and brine (5 mL). The organic layer was then dried over Na$_2$SO$_4$, filtered and concentrated under reduced pressure. The desired product **16** was obtained with 29% yield (28 mg, 0.14 mmol) after being purified by chromatography column (eluent: from 100% cyclohexane to 9/1 cyclohexane/EA).

Rf = 0.4 (eluent: 9/1 cyclohexane/EA)

¹H NMR (400 MHz, CDCl$_3$) δ 7.48 – 7.41 (m, 5H, H1-2-3), 4.12 (s, 2H, H4).

¹³C NMR (100 MHz, CDCl$_3$) δ 194.0, 134.8, 130.1, 129.6, 126.2, 58.0.

ν_{max} (thin film) /cm^{-1} 2921 (br), 2100 (s), 1700 (m), 1477 (w), 1440 (w), 1267 (br), 1066 (w), 997 (w), 896 (w), 745 (m), 688 (m).

S-phenyl 4-azidobenzothioate 17
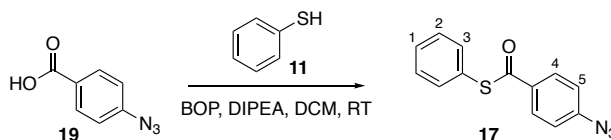
4-Azidobenzoic acid **19** (200 mg, 1.23 mmol, 1 equiv.) and thiophenol **11** (0.13 mL, 1.23 mmol, 1 equiv.) were stirred with BOP (596 mg, 1.35 mmol, 1.1 equiv.) and DIPEA (0.41 mL, 2.45 mmol, 2 equiv.) overnight in DCM (10 mL). The next day, the reaction was quenched with NaHCO$_3$ (15 mL, aq., sat.). The layers were separated and the organic layer was sequentially washed with NH$_4$Cl (15 mL, aq., sat.) and brine (15 mL). The organic layer was then dried over Na$_2$SO$_4$, filtered and concentrated under reduced pressure. The desired product **17** was obtained with 54% yield (170 mg, 0.67 mmol) after being purified by chromatography column (eluent: from 100% cyclohexane to 9/1 cyclohexane/EA).

Rf = 0.7 (eluent: 9/1 cyclohexane/EA)

¹H NMR (400 MHz, CDCl$_3$) δ 8.06 – 8.02 (m, 2H, H4), 7.54 – 7.49 (m, 2H, H3), 7.49 – 7.44 (m, 3H, H1, H2), 7.14 – 7.09 (m, 2H, H5).

¹³C NMR (100 MHz, CDCl$_3$) δ 188.8, 145.7, 135.2, 133.3, 129.7, 129.5, 129.4, 127.3, 119.3.

ν_{max} (thin film) /cm^{-1} 2088 (m), 1663 (m), 1593 (w), 1500 (w), 1286 (s), 1212 (s), 1173 (m), 904 (s), 837 (m), 705 (m), 689 (w), 637(w), 598 (w).
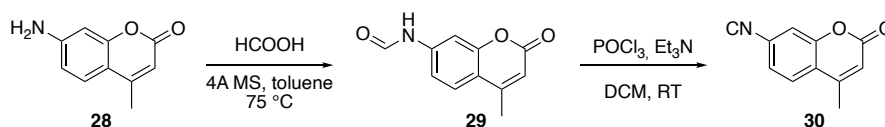
HRMS (ESI⁺) calcd for C$_{13}$H$_{10}$N$_3$OS$^+$ [M+H⁺] 256.0539; found 256.0525.

5-((3-aminopropyl)carbamoyl)-2-(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)benzoate 21
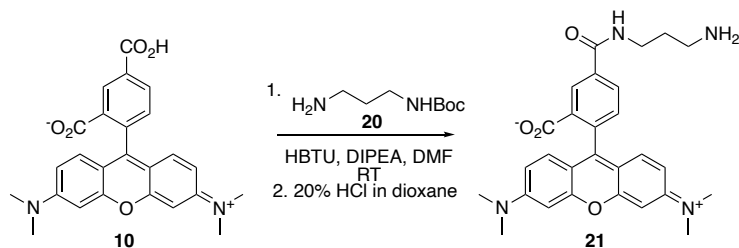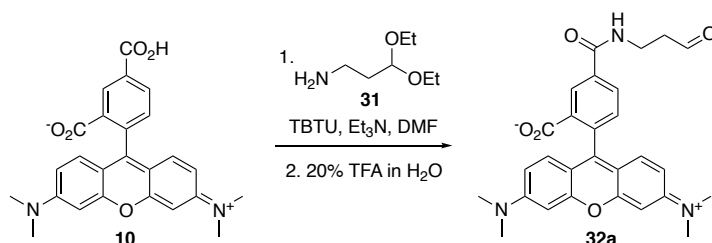

21 was synthesized according to reported procedures.[336]
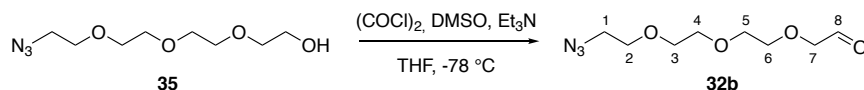
7-isocyano-4-methylcoumarin 30

30 was synthesized according to reported procedures.[337]

2-(6-(Dimethylamino)-3-(dimethyliminio)-3H-xanthen-9-yl)-5-((3-oxopropyl)carbamoyl)benzoate 32a

32a was synthetized according to reported procedures.[338]

2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)acetaldehyde 32b

Oxalyl chloride (0.08 mL, 0.92 mmol, 2.0 equiv.) was added dropwise to a solution of DMSO (0.10 mL, 1.37 mmol, 3.0 equiv.) in dry THF (3 mL) at -78 °C under N_2 atmosphere, and the

mixture was stirred for 15 min. **N₃-PEG4-OH 35** (100 mg, 0.46 mmol, 1.0 equiv.) in dry THF (3 mL) was added dropwise to the reaction mixture, which was stirred for 30 min at -78 °C. Triethylamine (0.40 mL, 2.74 mmol, 6.0 equiv.) was then added dropwise, and the reaction mixture was stirred for 15 min at -78 °C before being warmed to room temperature. The precipitate formed was removed by filtration and the filtrate was concentrated under reduced pressure. The crude was solubilized in DCM (5 mL) and washed three times with water (5 mL). The aqueous layer was finally extracted three times with DCM (10 mL). The organic layer was dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude product was purified by column chromatography (eluent: cyclohexane 3/7 ethyl acetate) and **1b** was obtained as a yellow oil with 66% yield (65 mg, 0.30 mmol).

Rf = 0.20 (eluent: 3/7 cyclohexane/EA; stained with PMA)

¹H NMR (400 MHz, CDCl₃) δ 9.73 (brs, 1H, H8), 4.16 (d, *J* = 0.9 Hz, 2H, H7), 3.76 – 3.65 (m, 10H, H2, H3, H4, H5, H6), 3.39 (t, *J* = 5.1 Hz, 2H, H1).

¹³C NMR (100 MHz, CDCl₃) δ 201.1 (CH), 77.0 (CH₂), 71.4 (CH₂), 71.0 (CH₂), 70.9 (CH₂), 70.8 (CH₂), 70.2 (CH₂), 50.8 (CH₂).

ν_max (thin film) /cm⁻¹ 3431 (br), 2869 (m), 2098 (s), 1733 (m), 1444 (w), 1285 (m), 1118 (s), 938 (w).

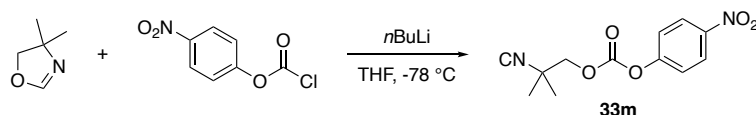**HRMS (ESI⁺)** calcd for C₈H₁₅N₃O₄Na⁺ [M+Na⁺] 240.0955; found 240.0956.

2-Isocyano-*N*-(prop-2-yn-1-yl)acetamide 33g
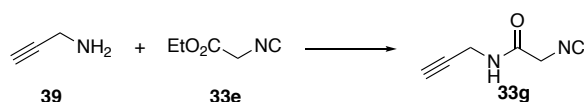
33g was synthetized according to reported procedures.[339]
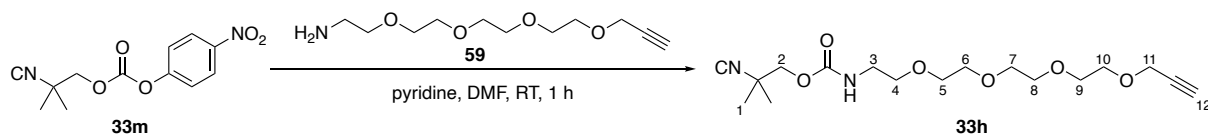

2-Isocyano-2-methylpropyl (4-nitrophenyl) carbonate 33m

33m was synthetized according to reported procedures.[340]

2-Isocyano-2-methylpropyl (3,6,9,12-tetraoxapentadec-14-yn-1-yl)carbamate 33h

2-Isocyano-2-methylpropyl (4-nitrophenyl) carbonate **33m** (17 mg, 0.064 mmol, 1.0 equiv.), 3,6,9,12-tetraoxapentadec-14-yn-1-amine **59** (15 mg, 0.064 mmol, 1.0 equiv.) and pyridine (5.6 μL, 0.07 mmol, 1.1 equiv.) were stirred in DMF (1 mL) for 1 h. After 1 h, the reaction was stopped and diluted in ethyl acetate. The organic layer was washed with HCl (5 mL, aq., 0.1 M) and brine (5 x 5 mL). The organic layer was then dried over Na_2SO_4, filtered and concentrated under reduced pressure. The desired product **33h** was obtained as a colourless oil with 35% yield (8 mg, 0.022 mmol) after being purified by chromatography column (eluent: from 4/1 cyclohexane/ EA to 100% EA).

Rf = 0.10 (eluent: 4/1 cyclohexane/EA; stained with vanillin)

^{1}H NMR (400 MHz, CDCl$_3$) δ 4.20 (d, J = 2.4 Hz, 2H, H11), 4.06 – 4.01 (m, 2H, H2), 3.72 – 3.61 (m, 12H), 3.58 (t, J = 5.1 Hz, 2H, H4), 3.39 (app q, J = 5.3 Hz, 2H, H3), 2.43 (t, J = 2.4 Hz, 1H, H12), 1.47 – 1.39 (m, 6H, H1).

^{13}C NMR (126 MHz, CDCl$_3$) δ 156.0 (C), 155.3 (t, $J_{C\text{-}N}$ = 3.8 Hz, C), 125.5 (C), 79.8 (CH), 74.7 (CH$_2$), 70.8 (CH$_2$), 70.7 (CH$_2$), 70.6 (CH$_2$), 70.5 (CH$_2$), 70.1 (CH$_2$), 70.0 (CH$_2$), 69.3 (CH$_2$), 58.5 (CH$_2$), 56.8 (t, $J_{C\text{-}N}$ = 6.3 Hz, C), 41.2 (CH$_2$), 26.0 (CH$_3$).

ν_{max} (thin film) /cm^{-1} 3289 (br), 2922 (br), 2356 (w), 2135 (m), 1720 (s), 1529 (m), 1463 (w), 1349 (w), 1242 (m), 1095 (s), 921 (w), 668 (br), 510 (br).

HRMS (ESI$^+$) calcd for $C_{17}H_{28}N_2O_6Na^+$ [M+Na$^+$] 379.1840; found 379.1838.

Bis(2-isocyano-2-methylpropyl) ((ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))dicarbamate 33i
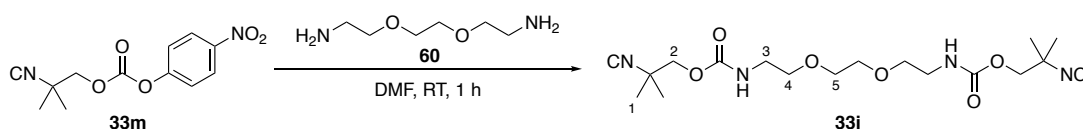
2-Isocyano-2-methylpropyl (4-nitrophenyl) carbonate **33m** (50 mg, 0.19 mmol, 2.2 equiv.) and 2,2′-(ethylenedioxy)bis(ethylamine) **60** (12 μL, 0.086 mmol, 1.0 equiv.) were stirred in DMF (2 mL) for 1 h. After 1 h, the reaction was stopped and diluted in ethyl acetate. The organic

layer was washed with HCl (10 mL, aq., 0.1 M) and brine (5 x 10 mL). The organic layer was then dried over Na_2SO_4, filtered and concentrated under reduced pressure. The desired product **33i** was obtained as a colourless oil with 55% yield (19 mg, 0.047 mmol) after being purified by chromatography column (eluent: from 4/1 cyclohexane/ EA to 100% EA).

**Rf =** 0.10 (eluent: 4/1 cyclohexane/EA; stained with vanillin)

¹H NMR (400 MHz, CDCl₃) δ 5.41 (t, J = 5.8 Hz, 2H, NH), 4.09 − 4.01 (m, 4H, H2), 3.63 (s, 4H, H5), 3.58 (t, J = 5.2 Hz, 4H, H4), 3.41 (app q, J = 5.4 Hz, 4H, H3), 1.45 − 1.40 (m, 12H, H1).

¹³C NMR (100 MHz, CDCl₃) δ 155.8 (C), 155.2 (t, J_{C-N} = 3.8 Hz, C), 70.3 (CH₂), 69.9 (CH₂), 56.7 (t, J_{C-N} = 5.7 Hz, C), 40.9 (CH₂), 25.8 (CH₃).

ν_{max} (thin film) /cm⁻¹ 3327 (br), 2873 (br), 2361 (w), 2135 (m), 1708 (s), 1528 (m), 1471 (w), 1238 (m), 1101 (m), 1047 (m), 773 (w).

HRMS (ESI⁺) calcd for $C_{18}H_{31}N_4O_6^+$ [M+H⁺] 399.2245; found 399.2237.

2-Isocyano-2-methylpropyl (2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)carbamate 33j
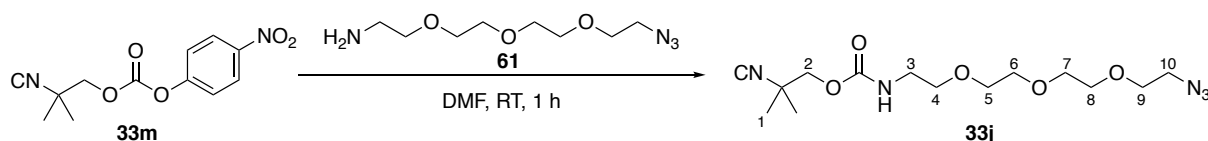
2- Isocyano-2-methylpropyl (4-nitrophenyl) carbonate **33m** (50 mg, 0.19 mmol, 1.0 equiv.) and 1-amino-11-azido-3,6,9-trioxaundecane **61** (45.4 mg, 0.21 mmol, 1.1 equiv.) were stirred in DMF (2 mL) for 1 h. After 1 h, the reaction was stopped and diluted in ethyl acetate. The organic layer was washed with HCl (10 mL, aq., 0.1 M) and brine (5x 10 mL). The organic layer was then dried over Na_2SO_4, filtered and concentrated under reduced pressure. The desired product **33j** was obtained as a colourless oil with 63% yield (41 mg, 0.12 mmol) after being purified by chromatography column (eluent: from 4/1 cyclohexane/ EA to 100% EA).

**Rf =** 0.10 (eluent: 4/1 cyclohexane/EA; stained with vanillin)

¹H NMR (400 MHz, CDCl₃) δ 4.03 − 4.00 (m, 2H, H2), 3.69 − 3.60 (m, 10H, H4, H5, H6, H7, H8), 3.56 (dd, J = 5.5, 4.6 Hz, 2H, H9), 3.40 − 3.34 (m, 4H, H3, H10), 1.45 − 1.35 (m, 6H, H1).

¹³C NMR (100 MHz, CDCl₃) δ 155.9 (C), 155.3 (C), 70.8 (CH₂), 70.7 (CH₂), 70.7 (CH₂), 70.4 (CH₂), 70.1 (CH₂), 69.9 (CH₂), 56.8 (C), 50.8 (CH₂), 41.1 (CH₂), 25.9 (CH₃).

ν_{max} (thin film) /cm⁻¹ 3331 (br), 2868 (br), 2099 (m), 1721 (s), 1525 (m), 1470 (m), 1240 (w), 1101 (s), 929 (w), 555 (w).

HRMS (ESI⁺) calcd for $C_{14}H_{26}N_5O_5^+$ [M+H⁺] 344.1936; found 344.1929.

(*E*)-Cyclooct-4-en-1-yl (2-isocyano-2-methylpropyl) propane-1,3-diyldicarbamate 33k
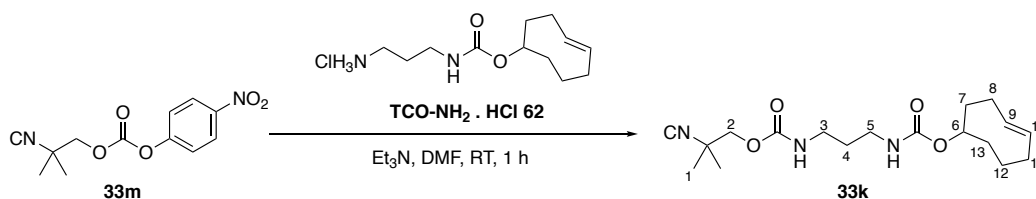
2- Isocyano-2-methylpropyl (4-nitrophenyl) carbonate **33m** (5 mg, 0.019 mmol, 1.0 equiv.), **TCO-NH₂.HCl 62** (5.5 mg, 0.021 mmol, 1.2 equiv.) and triethylamine (2.0 μL, 0.023 mmol, 1.2 equiv.) were stirred in DMF (0.5 mL) for 1 h. After 1 h, the reaction was stopped and diluted in ethyl acetate. The organic layer was washed with HCl (2 mL, aq., 0.1 M) and brine (5 x 2 mL). The organic layer was then dried over Na_2SO_4, filtered and concentrated under reduced pressure. The desired product **33k** was obtained as a colourless oil with 75% yield (5 mg, 0.014 mmol) after being purified by chromatography column (eluent: from 4/1 cyclohexane/ EA to 7/3 cyclohexane/ EA).

**Rf =** 0.10 (eluent: 4/1 cyclohexane/EA, stained with $KMnO_4$)

¹H NMR (400 MHz, CDCl₃) δ 5.61 – 5.44 (m, 2H, H9/10), 5.30 (d, *J* = 10.0 Hz, 1H, NH), 4.89 (s, 1H, NH), 4.33 (dd, *J* = 10.7, 5.9 Hz, 1H, H6), 4.05 – 4.01 (m, 2H, H2), 3.27 – 3.16 (m, 5H), 2.39 – 2.29 (m, 3H), 2.04 – 1.87 (m, 3H), 1.78 – 1.51 (m, 5H), 1.46 – 1.40 (m, 6H, H1).

¹³C NMR (126 MHz, CDCl₃) δ 156.9 (C), 156.2 (C), 155.4 (t, J_{C-N} = 3.8 Hz, C), 135.1 (CH), 133.1 (CH), 80.8 (CH), 70.0 (CH₂), 56.8 (t, J_{C-N} = 7.6 Hz, C), 41.3 (CH₂), 38.8 (CH₂), 37.9 (CH₂), 37.6 (CH₂), 34.4 (CH₂), 32.7 (CH₂), 31.1 (CH₂), 30.6 (CH₂), 26.0 (CH₃).

ν_{max} **(thin film) /cm⁻¹** 3328 (br), 2934 (m), 2361 (w), 2135 (w), 1697 (s), 1540 (m), 1457 (w), 1257 (m), 1145 (w), 993 (w).

HRMS (ESI⁺) calcd for $C_{18}H_{29}N_3O_4Na^+$ [M+Na⁺] 374.2050; found 374.2053.

(3-(((((2-Isocyano-2-methylpropoxy)carbonyl)amino)oxy)propyl)tris(2,4,6-trimethoxyphenyl)phosphonium bromide 33l
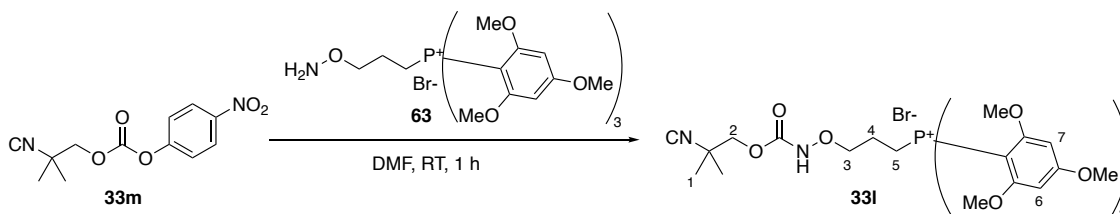
2-Isocyano-2-methylpropyl (4-nitrophenyl) carbonate **33m** (20 mg, 0.076 mmol, 1.0 equiv.) and (3-(aminooxy)propyl)tris(2,4,6-trimethoxyphenyl)phosphonium bromide **63** (51.9 mg, 0.076 mmol, 1.0 equiv.) were stirred in DMF (1 mL) for 1 h. After 1 h, the reaction was stopped and diluted in ethyl acetate. The organic layer was washed with HCl (5 mL, aq., 0.1 M) and brine (5 x 5 mL). The organic layer was then dried over Na_2SO_4, filtered and concentrated under reduced pressure. The desired product **33l** was obtained as a colorless oil with 5% yield (3 mg, 0.004 mmol) after being purified by chromatography column (eluent: from 4/1 cyclohexane/ EA to 100% EA).

Rf = 0.10 (eluent: 4/1 cyclohexane/EA)

¹H NMR (400 MHz, CDCl₃) δ 6.09 (d, J = 4.6 Hz, 6H, H6/7), 4.07 (s, 2H, H2), 3.99 (t, J = 6.0 Hz, 2H, H3), 3.86 (s, 9H, -pOMe), 3.62 (s, 18H, -oOMe), 3.26 – 3.17 (m, 2H, H5), 1.68 – 1.59 (m, 2H, H4), 1.45 (t, J_{H-N} = 1.8 Hz, 6H, H1).

¹³C NMR (100 MHz, CDCl₃) δ 165.7 (d, J_{CP} = 1.4 Hz, C), 163.9 (d, J_{CP} = 1.6 Hz, C), 156.8 (C), 147.7 (C), 92.4 (d, J_{CP} = 104.8 Hz, C), 91.1 (d, J_{CP} = 7.1 Hz, CH), 76.0* (CH₂), 70.0 (CH₂), 57.1** (C), 56.2 (CH₃), 55.8 (CH₃), 26.1 (CH₃), 25.5* (CH₂), 23.0 (CH₂).
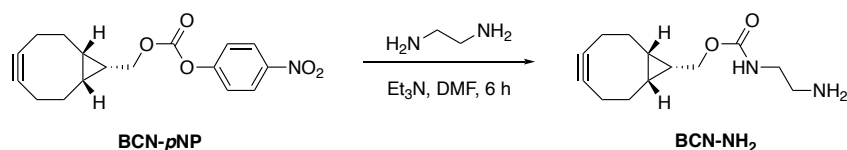
ν_{max} (thin film) /cm⁻¹ 3328 (br), 2934 (m), 2361 (w), 2135 (w), 1697 (s), 1540 (s), 1457 (w), 1257 (m), 1145 (w), 993 (w).

HRMS (ESI⁺) calcd for $C_{36}H_{48}N_2O_{12}P^+$ [M⁺] 731.2939; found 731.2926.

* identified by ¹H-¹³C HSQC

** identified by ¹H-¹³C HMBC

((1R,8S,9S)-Bicyclo[6.1.0]non-4-yn-9-yl)methyl (2-aminoethyl)carbamate (BCN-NH₂)

<div align="center">**BCN-pNP**   **BCN-NH₂**</div>

BCN-pNP and **BCN-NH₂** were synthetized according to reported procedures.[341,342]

1-(6-((2-((((((1R,8S,9s)-Bicyclo[6.1.0]non-4-yn-9-yl)methoxy)carbonyl)amino)ethyl)amino)-6-oxohexyl)-3,3-dimethyl-2-((1E,3E)-5-((E)-1,3,3-trimethyl-5-sulfoindolin-2-ylidene)penta-1,3-dien-1-yl)-3H-indol-1-ium-5-sulfonate (BCN-Cy5) 37a
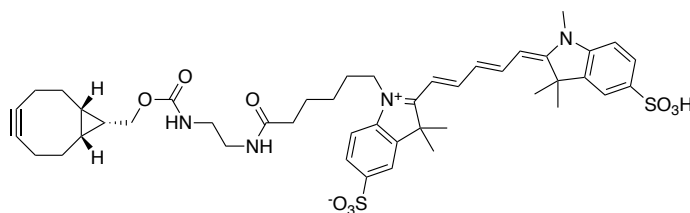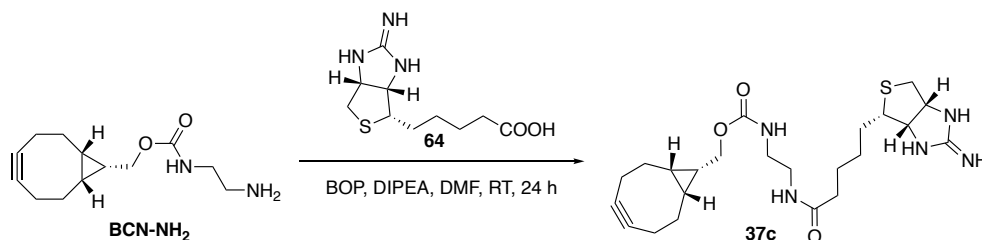

37a was synthetized according to reported procedures.[343]

(2-((2,5-Dioxo-1-pyrrolidinyl)oxy)-2-oxoethyl)(tris(2,4,6-trimethoxyphenyl))phosphonium bromide 37b
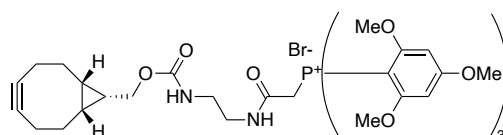
37b was synthetized according to reported procedures.[344]

((1*R*,8*S*,9*s*)-Bicyclo[6.1.0]non-4-yn-9-yl)methyl (2-(5-((4*S*)-2-iminohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamido)ethyl)carbamate 37c

BCN-NH₂ (21.4 mg, 0,090 mmol, 1.1 equiv.), 2-iminobiotin **64** (20 mg, 0.082 mmol, 1.0 equiv.), DIPEA (40 μL, 0.246 mmol, 3 equiv.) and BOP (54.6 mg, 0.122 mmol, 1.5 equiv.) were stirred in DMF (2 mL) for 24 h at room temperature. After 24 h, the reaction was dried under reduced pressure. The crude was then diluted with ethyl acetate (5 mL) and the organic layer was sequentially washed with NH₄Cl (5 mL, aq., sat.), NaHCO₃ (5 mL, aq., sat.) and brine (3 x 5 mL). The organic layer was then dried over Na₂SO₄, filtered and concentrated under reduced pressure. The crude was purified by chromatography column (eluent: 100% EA then 8/2 DCM/MeOH to recover the desired product) and the desired product **37c** was obtained with 16% yield as a colourless oil (6 mg, 0.005 mmol).

HRMS (ESI⁺) calcd for $C_{23}H_{36}N_5O_3S^+$ [M+H⁺] 462.2541; found 462.2542.

((1*R*,8*S*,9*s*)-Bicyclo[6.1.0]non-4-yn-9-yl)methyl (2-(3-(3,5-bis(trifluoromethyl)phenyl)thioureido)ethyl)carbamate 37d
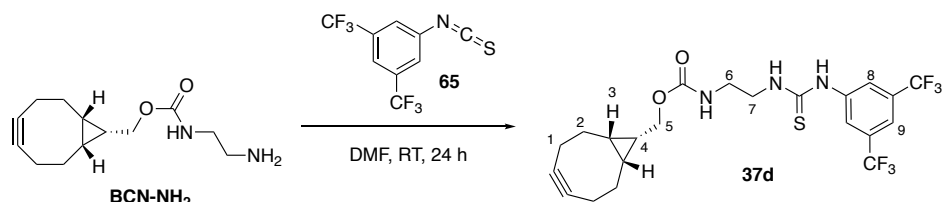
BCN-NH₂ (10.0 mg, 0.042 mmol, 1 equiv.) and 1-isothiocyanato-3,5-bis(trifluoromethyl)benzene **65** (11.4 mg, 0.042 mmol, 1.0 equiv.) were stirred in DMF (0.5 mL) for 24 h at room temperature. After 24 h, the reaction was dried under reduced pressure and the crude was purified by chromatography column (eluent: from 100% cyclohexane to 100% EA to recover **37d**). The desired product **37d** was obtained as a colourless oil with 64% yield (13.8 mg, 0.027 mmol).

^{1}H NMR (400 MHz, CDCl₃) δ 7.93 (s, 2H, H8), 7.69 (s, 1H, H9), 4.16 – 3.99 (m, 2H, H5), 3.46 – 3.35 (m, 2H, H7), 2.33 – 2.14 (m, 6H, H1, H2), 1.45 – 1.43 (m, 2H, H2), 1.32 – 1.19 (m, 3H, H4, H6), 0.98 – 0.81 (m, 2H, H3).

^{13}C NMR (126 MHz, CDCl₃) δ 181.4 (C), 158.7 (C), 132.5 (C), 124.5 (CH), 124.2 (-CF₃), 122.0 (-CF₃), 119.2 (CH), 98.9 (C), 63.9 (CH₂), 40.4 (CH₂), 29.9 (CH₂), 29.2 (CH₂), 29.1 (CH₂), 21.5 (CH₂), 20.3 (CH), 17.7 (CH).

***ν*$_{max}$ (thin film) /cm^{-1}** 3304 (br), 2923 (w), 1689 (m), 1527 (m), 1381 (m), 1275 (s), 1173 (m), 1130 (s), 885 (w), 733 (w).

HRMS (ESI⁺) calcd for $C_{22}H_{23}N_3O_2F_6Na^+$ [M+Na⁺] 530.1317; found 530.1313.

BCN-oligonucleotide derivatives 37e and 37f
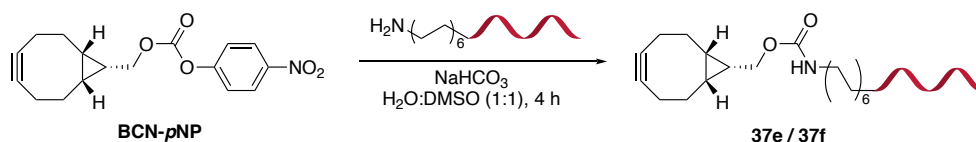
Commercial amino-modified 20-base oligonucleotide (**ON20**) solution (10-50 nmol) was diluted with water to the final volume of 300 μL and **ON20** was precipitated by addition of acetone (900 μL) and LiClO₄ (20 μL, 3.0 M in water). The sample was centrifuged at 15000 g for 8 min and supernatant was discarded. The precipitate was dissolved in 300 μL of water and this

procedure was repeated again two times. The precipitate obtained after the final round was dissolved in water to a final concentration of 1.0 mM.

In a 2-mL Eppendorf tube were sequentially introduced **ON20** (1.0 equiv., 1.0 mM in water), **BCN-*p*NP** (10.0 equiv., 10.0 mM in DMSO), and NaHCO$_3$ (100.0 equiv.,1.0 M in water). The resulting mixture was incubated at 25 °C for 4 h under argon atmosphere before the BCN-modified ON was precipitated by acetone and purified by preparative HPLC (see 'Material & methods'). After lyophilization, **37e/37f** was dissolved in water and concentration was measured by spectrophotometry, giving a final yield for the BCN-modified ON ranging between 53% and 56%. Purity of **37e/37f** was confirmed by HPLC after SPAAC reaction with a slight excess of a commercially available TAMRA azide **5**, as shown on **Figure 43**.
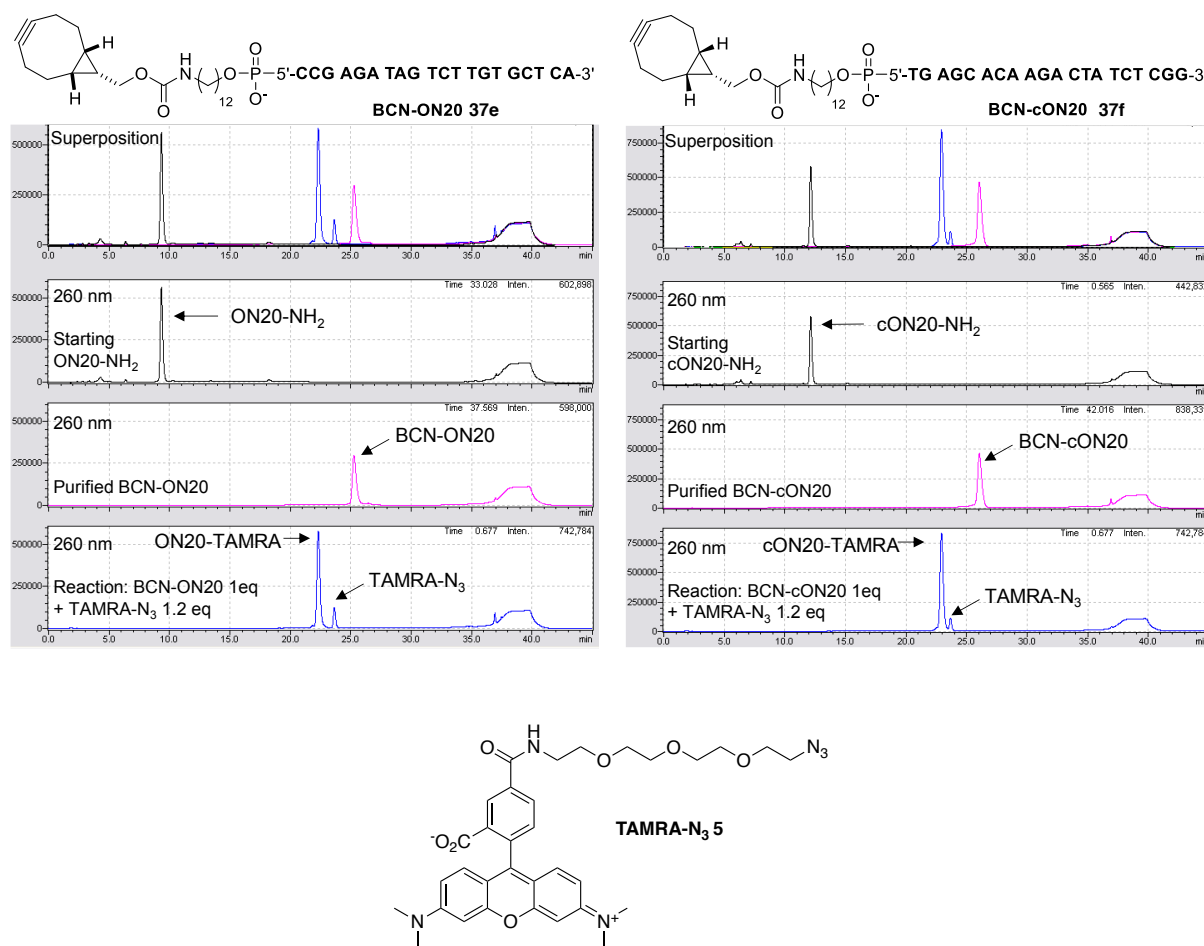
Figure 43: Structures and HPLC profiles of BCN-ON **37e** and **37f.**

4-(6-Methyl-1,2,4,5-tetrazin-3-yl)benzoic acid 40 and dimethyl 1,2,4,5-tetrazine-3,6-dicarboxylate 41

$R^1 = R^2 = CO_2Me$, **40**
$R^1 = Me$, $R^2 = 4\text{-}CO_2H\text{-}(C_6H_4)$, **41**
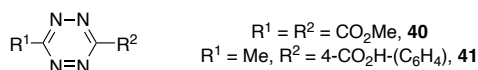
Tetrazine **40** is commercially available (CAS N° 2166-14-5).

Tetrazine **41** was synthetized according to reported procedures.[345]

Fmoc-Glu(O*t*Bu)-NHMe 52

Fmoc-Glu(O*t*Bu)-OH

MeNH$_2$ (2 M in THF)
BOP, DIPEA
—————————→
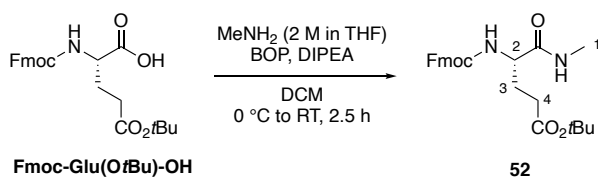DCM
0 °C to RT, 2.5 h

52

BOP (0.55 g, 1.23 mmol, 1.05 equiv.) and DIPEA (0.4 mL, 2.27 mmol, 1.93 equiv.) were sequentially added to a solution of **Fmoc-Glu(O*t*Bu)-OH** (0.50 g, 1.18 mmol, 1 equiv.) in DCM (10 mL) at 0°C. Methylamine (2 M in THF, 1.2 mL, 2.35 mmol, 2 equiv.) was added dropwise at 0°C. The reaction mixture was then warmed to room temperature and stirred for 2.5 hours. After 2.5 hours, the organic layer was washed with NaHCO$_3$ (25 mL, aq., sat.), NH$_4$Cl (25 mL, aq., sat.) and brine (25 mL). The organic layer was then dried over Na$_2$SO$_4$, filtered and concentrated under reduced pressure. The compound **52** was purified by chromatography column (eluent: cyclohexane 1/1 ethyl acetate) and afforded with 97% yield (0.50 g, 1.14 mmol) as a white solid.

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.74 (d, J = 7.5 Hz, 2H, Fmoc ArH), 7.57 (d, J = 7.6 Hz, 2 H, Fmoc ArH), 7.38 (t, J = 7.5 Hz, 2H, Fmoc ArH), 7.29 (t, J = 7.6 Hz, 2H, Fmoc ArH), 6.56 – 6.48 (m, 1H, Fmoc CH), 5.94 (d, J = 7.9 Hz, 1H, Fmoc CH$_2$), 4.36 (d, J = 7.2 Hz, 2H, NH), 4.25 – 4.13 (m, 1H, H2), 2.79 (d, J = 4.7 Hz, 3H, H1), 2.47 – 2.35 (m, 1H, H4), 2.35 – 2.24 (m, 1H, H4), 2.15 – 2.01 (m, 1H, H3), 2.00 – 1.87 (m, 1H, H3), 1.44 (s, 9H, *t*Bu).

^{13}C NMR (100 MHz, CDCl$_3$) δ 172.8, 172.0, 156.4, 143.8, 143.8, 141.3, 127.7, 127.1, 125.1, 120.0, 81.0, 67.1, 54.4, 47.1, 31.7, 28.1, 26.3.
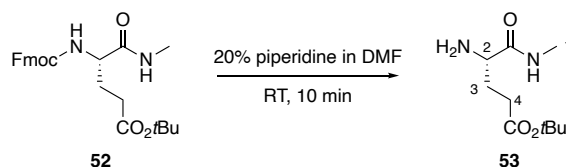
v_{max} (thin film) /cm^{-1} 3297 (m), 2976 (w), 2359 (s), 2339 (m), 1731 (w), 1690 (w), 1653 (m), 1539 (w), 1159 (w), 1050 (w), 739 (w), 669 (w).

HRMS (ESI$^+$) calcd for C$_{25}$H$_{31}$N$_2$O$_5$$^+$ [M+H$^+$] 439.2241; found 439.2233.

Mp = 147 °C

H-Glu(O*t*Bu)-NHMe 53

52   **53**

Fmoc-Glu(O*t*Bu)-NHMe **52** (0.50 g, 1.14 mmol, 1 equiv.) was stirred 10 minutes in a solution of 20% piperidine in DMF (10 mL). After 10 minutes, the reaction mixture was dried under reduced pressure. The compound **53** was purified by chromatography column (eluent: cyclohexane 7/3 ethyl acetate, then 100% ethyl acetate, then DCM 9/1 MeOH) and afforded as a light yellow oil with 80% yield (0.20 g, 0.93 mmol).
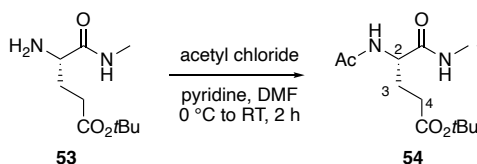
^{1}H NMR (400 MHz, CDCl$_3$) δ 3.43 (dd, J = 7.4, 5.3 Hz, 1H, H2), 2.77 (d, J = 4.9 Hz, 3H, H1), 2.38 – 2.23 (m, 2H, H3), 2.10 – 1.99 (m, 1H, H4), 1.85 – 1.74 (m, 1H, H4), 1.39 (s, 9H, *t*Bu).

^{13}C NMR (100 MHz, CDCl$_3$) δ 174.6, 172.8, 80.6, 54.6, 32.0, 30.1, 28.1, 25.9.

v_{max} (thin film) /cm^{-1} 3304 (br), 2976 (w), 2932 (w), 2359 (w), 1721 (m), 1650 (m), 1539 (w), 1410 (w), 1392 (w), 1366 (m), 1252 (w), 1148 (s), 846 (w).

HRMS (ESI$^+$) calcd for $C_{10}H_{21}N_2O_3^+$ [M+H$^+$] 217.1553; found 217.1548

Ac-Glu(O*t*Bu)-NHMe 54

53   **54**

The amino acid, H-Glu(O*t*Bu)-NHMe **53** (200 mg, 0.93 mmol, 1 equiv.) and pyridine (0.45 mL, 5.60 mmol, 6 equiv.) were solubilized in DMF (5 mL). The reaction mixture was cooled to 0 °C and acetyl chloride (0.13 mL, 1.85 mmol., 2 equiv.) added dropwise. The reaction was allowed to stir at room temperature for 2 hours. The reaction was quenched with ethanol (2 mL) and then dried under reduced pressure. Once dried, the crude was dissolved in DCM (25 mL). The organic layer was washed with 0.1 HCl (15 mL, aq.) and brine (2 x 15 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered and concentrated under reduced pressure. The compound **54** was purified by chromatography column (eluent: cyclohexane 7/3 ethyl acetate, then 100% ethyl acetate, then DCM 9/1 MeOH) and afforded with 54% yield (130 mg, g, 0.43 mmol) as a light yellow solid.

¹H NMR (400 MHz, CDCl₃) δ 4.38 (dd, J = 7.6, 5.6 Hz, 1H, H2), 2.82 (d, J = 4.9 Hz, 3H, H1), 2.50 – 2.38 (m, 1H, H3), 2.31 – 2.17 (m, 1H, H3), 2.09 – 2.01 (m, 1H, H4), 2.01 (s, 3H, Ac), 1.97 – 1.87 (m, 1H, H4), 1.45 (s, 9H, *t*Bu).

¹³C NMR (100 MHz, CDCl₃) δ 173.2, 172.1, 170.7, 81.2, 52.8, 31.9, 28.2, 27.8, 26.4, 23.3.

ν_max (thin film) /cm⁻¹ 3281 (w), 2976 (w), 1727 (m), 1633 (s), 1542 (m), 1367 (m), 1256 (m), 1153 (s), 731 (w), 603 (w).

HRMS (ESI⁺) calcd for $C_{12}H_{23}N_2O_4^+$ [M+H⁺] 259.1660; found 259.1646.

MP = 134 °C

Ac-Glu-NHMe 44

54 → **44**
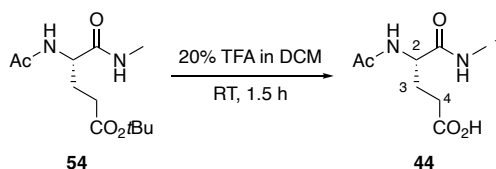
Ac-Glu(O*t*Bu)-NHMe **54** (0.11 g, 0.43 mmol, 1 equiv.), was stirred in a solution of 20% TFA in DCM (with few drops of water) (10 mL). After 1.5 hours, the reaction mixture was dried under reduced pressure. The compound was triturated with Et₂O (3 x 20 mL). The compound **44** was afforded as a light yellow solid with quantitative yield (0.10 g, 0.43 mmol).

¹H NMR (400 MHz, MeOD) δ 4.32 (dd, J = 9.0, 5.3 Hz, 1H, H2), 2.73 (s, 3H, H1), 2.37 (dd, J = 8.4, 7.1 Hz, 2H, H3), 2.13 – 2.01 (m, 1H, H4), 2.01 (s, 3H, Ac), 1.94 – 1.82 (m, 1H, H4).

¹³C NMR (100 MHz, MeOD) δ 176.4, 174.4, 173.5, 54.2, 31.2, 28.4, 26.3, 22.5.

ν_max (thin film) /cm⁻¹ 3294 (br), 1629 (s), 1543 (m), 1413 (w), 1377 (w), 1295 (w), 1157 (s), 812 (w), 703 (w), 597 (w).

HRMS (ESI⁺) calcd for $C_8H_{15}N_2O_4^+$ [M+H⁺] 203.1034; found 203.1028.

Cyclohexyl formamide labelled with ¹³C 55

55
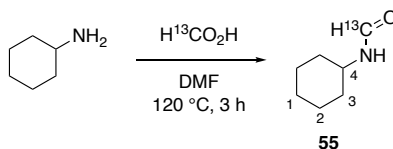
In a sealed tube, formic acid-¹³C (0.25 mL, 6.58 mmol, 1.5 equiv.) was added dropwise to a solution of cyclohexanamine (0.5 mL, 4.39 mmol, 1 equiv.) in DMF (2.5 mL) at 0 °C. After complete addition of formic acid, the tube was sealed and the reaction mixture heated at

120 °C for 3 hours. After 3 hours, the reaction was diluted in water (2.5 mL) and DCM (10 mL). The layers were separated and the organic layer was sequentially washed with NaHCO$_3$ (10 mL, aq., sat.) and brine (5 x 10 mL). The organic layer was dried over Na$_2$SO$_4$, filtered and concentrated under reduced pressure. The compound **55** was directly used in the next step without purification.

¹H NMR (400 MHz, CDCl$_3$) δ 7.88 (d, J = 7.8 Hz, 1H, ^{13}CH), 6.70 (d, J = 91.9 Hz, 1H, NH), 3.62 (tdt, J = 11.1, 8.3, 3.8 Hz, 1H, CH, H4), 1.70 (dt, J = 12.6, 4.1 Hz, 2H, CH2, H1), 1.48 (ddt, J = 45.4, 12.7, 4.0 Hz, 4H, CH2, H3), 1.21 – 0.95 (m, 4H, CH2, H2).

¹³C NMR (100 MHz, CDCl$_3$) δ 160.6 (^{13}C), 51.1, 46.9, 34.4, 32.7, 25.3, 24.7.

Cyclohexyl isocyanide labelled with ¹³C 51

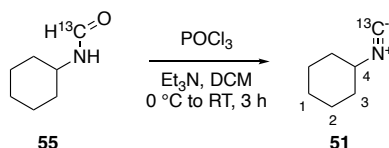POCl$_3$ (0.17 mL, 1.82 mmol, 1.1 equiv.) was added dropwise to a solution of **55** (211 mg, 1.65 mmol, 1.0 equiv.) and triethylamine (4.59 mL, 33.02 mmol, 20 equiv.) in DCM (15 mL) at 0 °C. After complete addition of POCl$_3$, the reaction was warmed to RT and the reaction was stirred for 3 hours. After 3 hours, the reaction was quenched with water (15 mL). The layers were separated and the organic layer was sequentially washed with HCl (2 x 15 mL, aq., 0.1 M) and brine (15 mL). The organic layer was dried over Na$_2$SO$_4$, filtered and concentrated under reduced pressure. The crude compound was first purified by column chromatography (eluent: cyclohexane 4/1 ethyl acetate) to afford 60 mg of **51** (0.55 mmol, 33% yield). Product **51** was purified a second time by distillation to afford 2 mg of pure product.

¹H NMR (400 MHz, DMSO-d_6) δ 3.81 – 3.74 (m, 1H, CH, H4), 1.86 – 1.77 (m, 2H, CH2, H1), 1.64 – 1.54 (m, 4H, CH2, H3), 1.45 – 1.26 (m, 4H, CH2, H2).

¹³C NMR (100 MHz, CDCl$_3$) δ 154.1 (t, J = 5.0 Hz), 51.8 (t, J = 5.5 Hz), 27.0, 25.1, 22.9.

6.4. Bioconjugation
6.4.1. Chemoselective labelling of cysteine residues with hypervalent iodines
6.4.1.1. Labelling of cysteine residues with ABX **1**

Bioconjugation procedures:
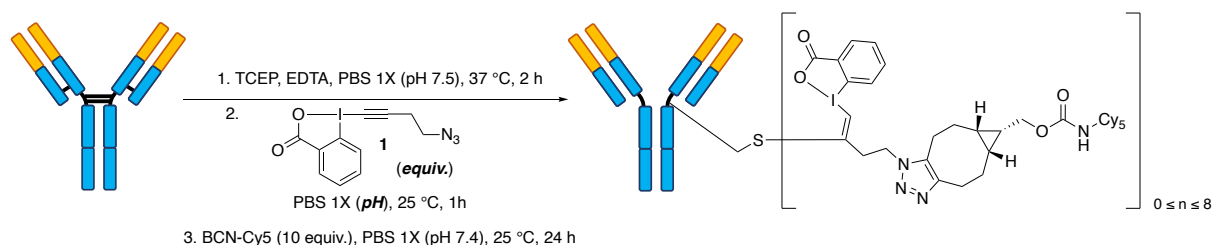
Reduction of trastuzumab:

See protocol in General procedure section.

Conjugation step:

To a solution of reduced trastuzumab (1 equiv., 5 mg/mL 100 μl in PBS 1X, 3.33 nmol, at either pH 7.5 or 8.5) was added ABX reagent **1** (pipetted from a 10 mM solution in DMSO, so as to get a volume of DMSO < 10% to that of the final solution). The resulting solution was then incubated for a given time at 25 °C, before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 columns pre-equilibrated with PBS 1X (pH 7.5) to a give a solution of trastuzumab-azide which was further used in the functionalization step.

Functionalization with SPAAC:

To the solution of trastuzumab-azide in PBS 1X (pH 7.5) was added BCN-Cy5 **37a** (as a 13 mM or 0.1 M solution in DMSO, 10 equiv.). The resulting solution was incubated for 20 hours at 25 °C, before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 columns pre-equilibrated with PBS 1/20X (pH 7.4) to give a solution of conjugated trastuzumab.

Sample preparation for native mass spectrometry:

See protocol in General procedure section.

Stability studies:

To evaluate the stability of the payload, trastuzumab-azide conjugate was incubated 3 days at 37 °C in a PBS buffer (1X, 137 mM of NaCl, pH 7.5). After three days, the resulting probe was analyzed by native mass spectrometry and compared to a control (i.e. non incubated conjugate), which confirmed the poor stability of the payload as more degradation peaks appeared.
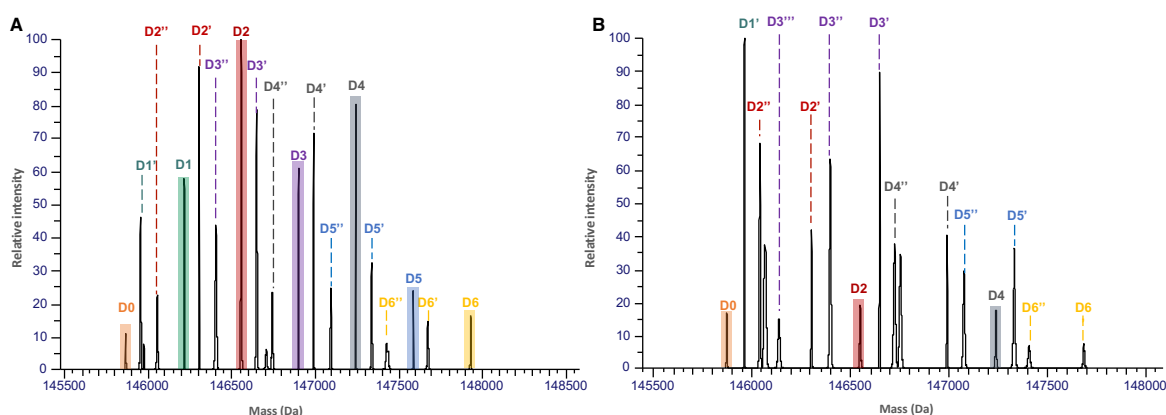
Figure 44: Trastuzumab-azide conjugate was obtained from 8 equivalents of **1** incubated 15 minutes at 25 °C on trastuzumab, and then incubated 3 days at 37 °C in PBS 1X (pH 7.5). **A**. Native MS spectrum of the trastuzumab-azide conjugate before the incubation at 37 °C for 3 days. **B**. Native MS spectrum of the trastuzumab-azide conjugate after the incubation at 37 °C for 3 days.

6.4.1.2.   Ethynylation of cysteine residues with a TMS-bearing EBX

Bioconjugation procedures:
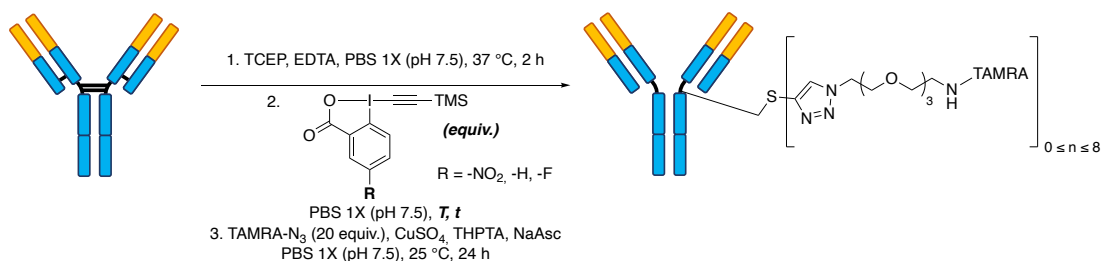
Reduction of trastuzumab:

See protocol in General procedure section

Conjugation step:

To a solution of reduced trastuzumab (1 equiv., 5 mg/mL, 100 µL in PBS 1X, 3.33 nmol, pH 7.5) was added the TMS-EBX reagent **4/7/8** (as a 10 mM solution in DMSO). The resulting solution was then incubated for a given time at either 4 °C, 25 °C or 37 °C, before the excess of reagent was removed by gel filtration chromatography using Bio-Spin P-30 columns pre-equilibrated with PBS 1X (pH 7.5) to give a solution of conjugated trastuzumab.

Functionalization step with CuAAC:

To a solution of conjugated trastuzumab (1 equiv., 50 μL in PBS 1X, pH 7.5) was added TAMRA- N$_3$ **5** (as a 100 mM solution in DMSO, 20 equiv.). A pre-mixed solution of CuSO$_4$ (as a 10 mM solution in H$_2$O, 1 equiv.) and THPTA (as a 10 mM solution in H$_2$O, 2 equiv.) was added to the reaction mixture, followed by sodium ascorbate (as a 10 mM solution in H$_2$O, 3 equiv.). The resulting mixture was then incubated for 24 hours at 25 °C before the excess of reagent was removed by gel filtration chromatography using Bio-Spin P-30 columns pre-equilibrated with PBS 1/20X (pH 7.5). The resulting conjugate was then washed six times with PBS 1/20X (pH 7.5) containing 1% EDTA on Vivaspin centrifugal concentrators (500 µL, 10 kDa) to remove chelated copper.

Sample preparation for native mass spectrometry:
See protocol in General procedure section.

6.4.2. Cysteine-to-lysine transfer

Bioconjugation procedures:

1. TCEP, EDTA, PBS 1X (pH 7.5), 37 °C, 2 h
2. **9 (equiv.),** PBS 1X (pH 7.5), 25 °C, 15 min
3. PBS 1X (pH 8.5), 37 °C, 24 h
4. L-cysteine (730 equiv.), PBS 1X (pH 8.5), 37 °C, 24 h
5. L-dehydroascobic acid (15 equiv.), PBS 1X (pH 7.5), 25 °C, 3 h

$0 \leq n \leq 8$

Reduction of trastuzumab:
See protocol in General procedure section

To a solution of reduced trastuzumab (1 equiv., 5 mg/mL, 100 μL in PBS 1X, 3.33 nmol, pH 7.5) was added thioester **9** (as a 10 mM solution in DMSO). The reaction mixture was then incubated at 25 °C for 15 min, before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 columns pre-equilibrated with PBS 1X (pH 8.5) to give a solution of conjugated trastuzumab that was then directly engaged in a *S*-to-*N* shift. The conjugated trastuzumab (1 equiv., 100 μL in PBS 1X, pH 8.5) was then incubated at 37 °C for 24 hours. After 24 hours, an excess of L-cysteine (as a 0.1 M solution in DMSO, 730 equiv.) was added to the antibody solution, and the resulting solution was further incubated for 24 hours at 37 °C, before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 columns pre-equilibrated with PBS 1X (pH 7.5) to give a solution of conjugated trastuzumab. The antibody (1 equiv., 100 μL in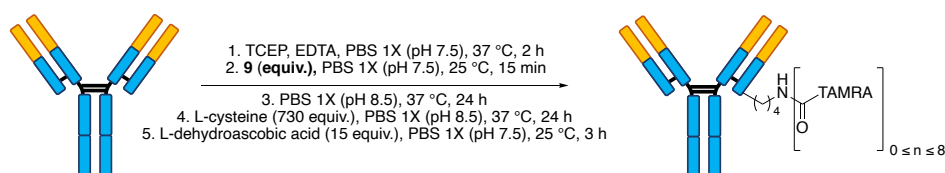 PBS 1X, pH 7.5) was then reoxidized with L-dehydroascorbic acid (as a 10 mM solution in H_2O, 15 equiv.). The resulting solution was incubated for 3 hours at 25 °C before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 columns pre-equilibrated with PBS 1/20X (pH 7.5) to afford the final antibody conjugate.

Sample preparation for native mass spectrometry:
See protocol in General procedure section.

6.4.3. Multicomponent approaches for the site-selective labelling of proteins

Structures of conjugates 34aa and 34ai to 34ri
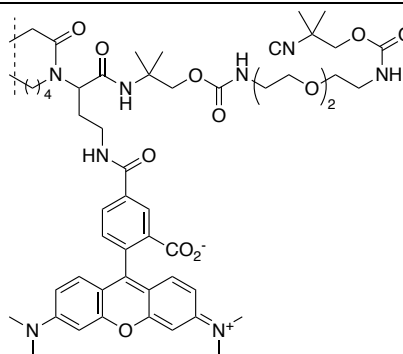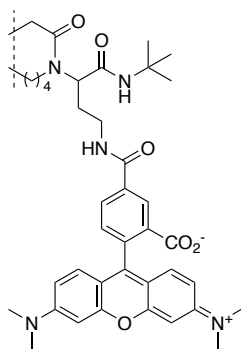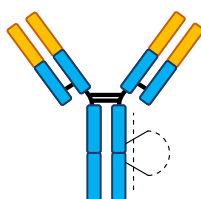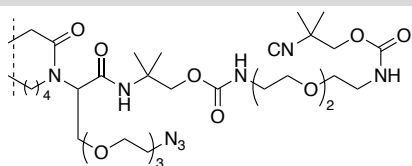
34aa	34ai
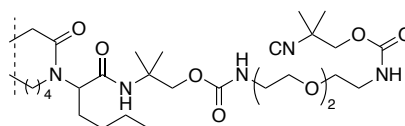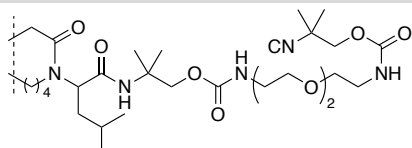
34bi	34ci
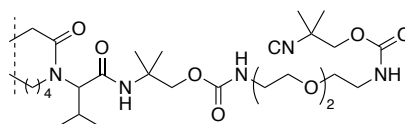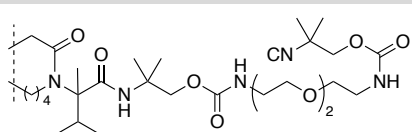
34di	34ei
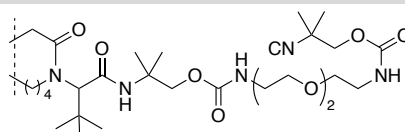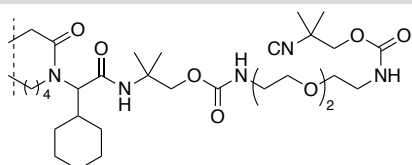
34fi	34gi
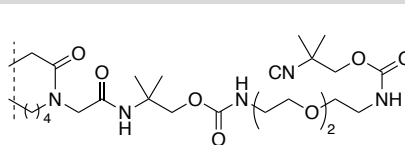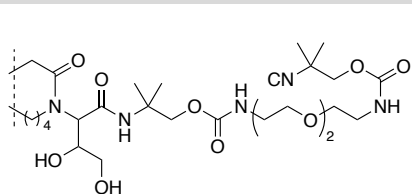
34hi	34ii
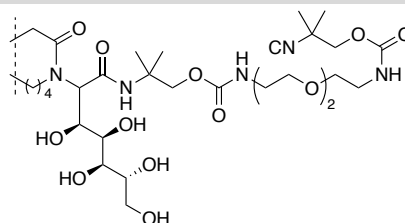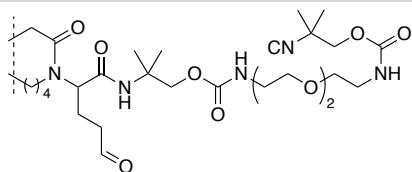
34ji	34ki
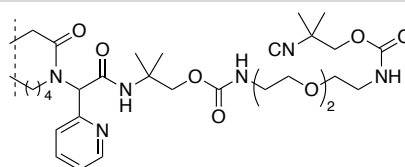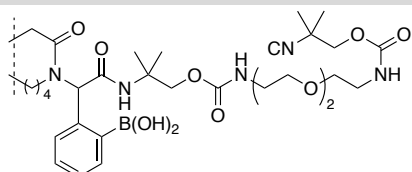
34li	34mi
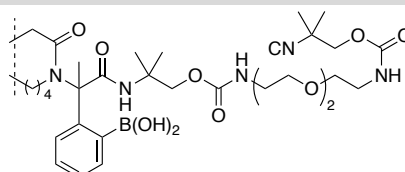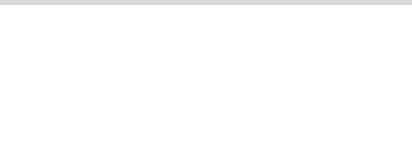
34ni	34oi
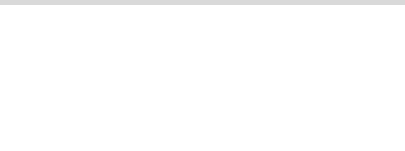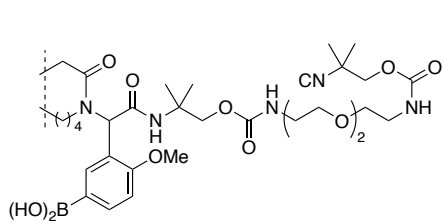
34pi
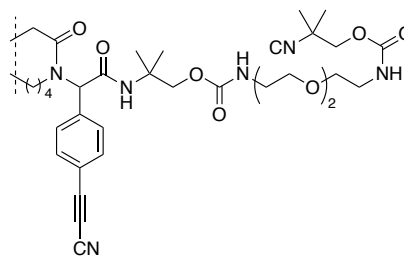
34qi
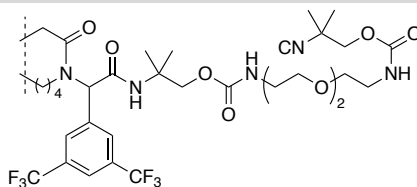
34ri

Structures of conjugates 34ba to 34bl
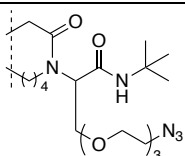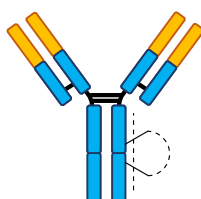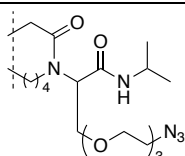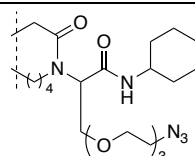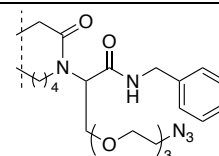
34ba **34bb** **34bc** **34bd**
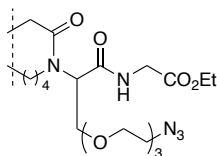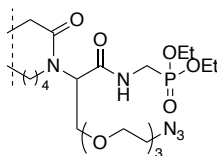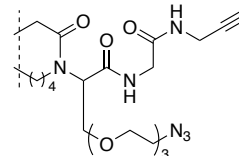
34be **34bf** **34bg**
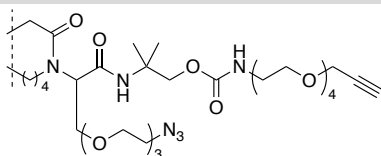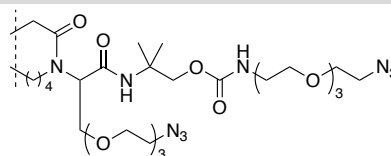
34bh **34bj**
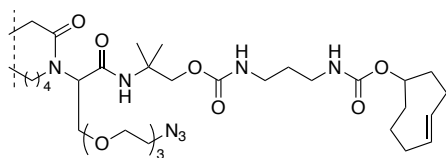
34bk
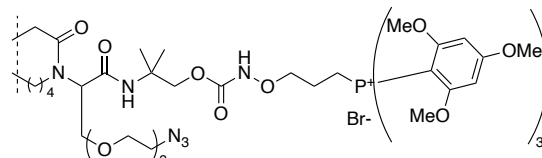
34bl

Structures of conjugates 38 to 38j

38

38a

38b

38c

38d

38e

5'-CCG AGA TAG TCT TGT GCT CA-3'

38f

5'-TG AGC ACA AGA CTA TCT CGG-3'
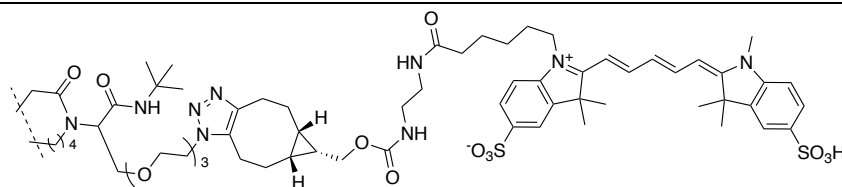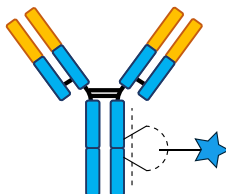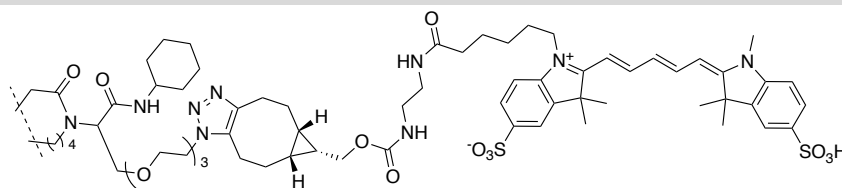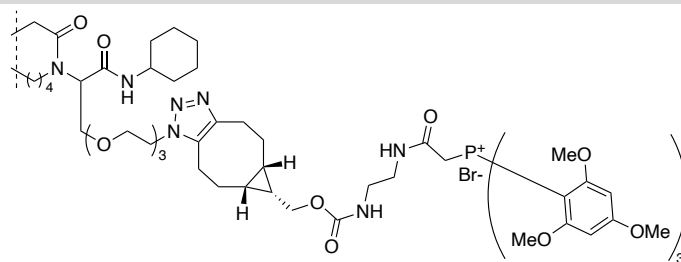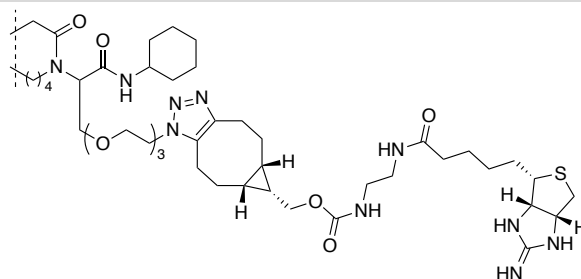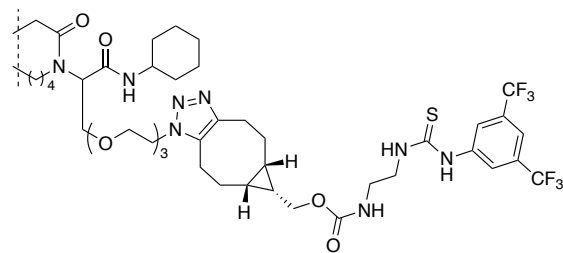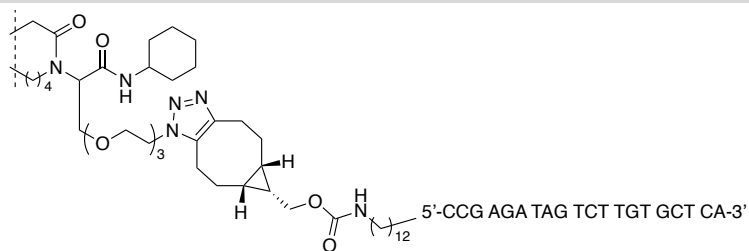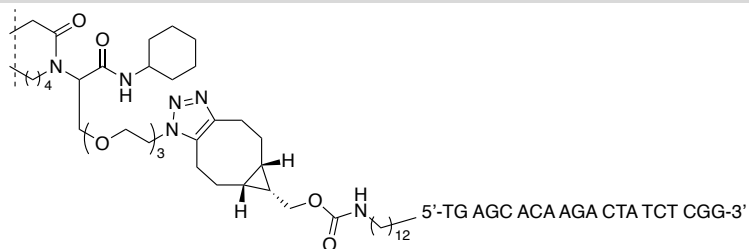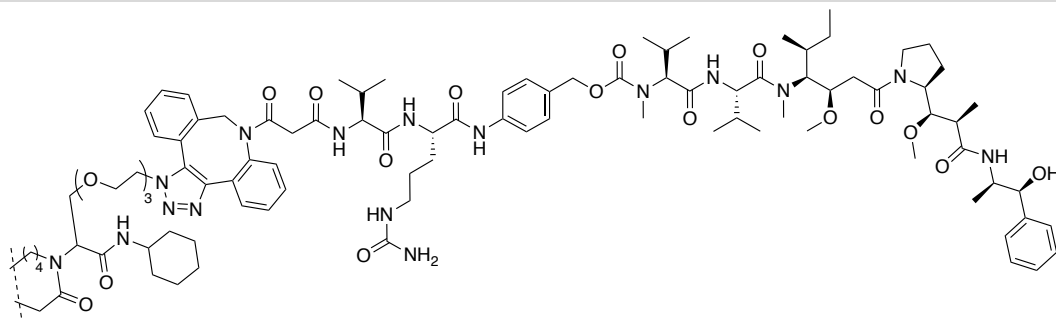
38g
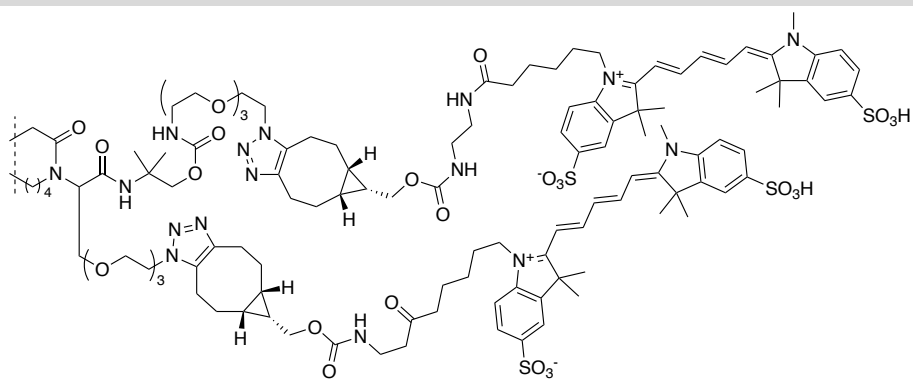
38h
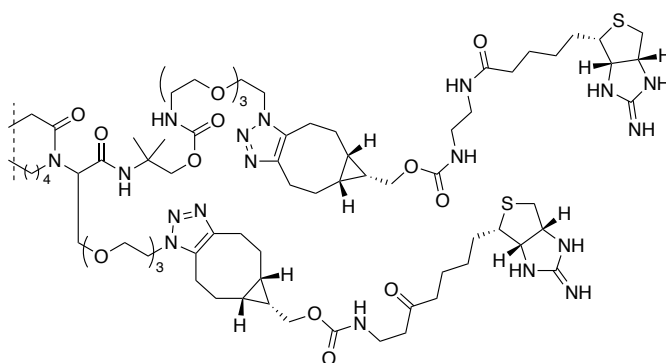
38i
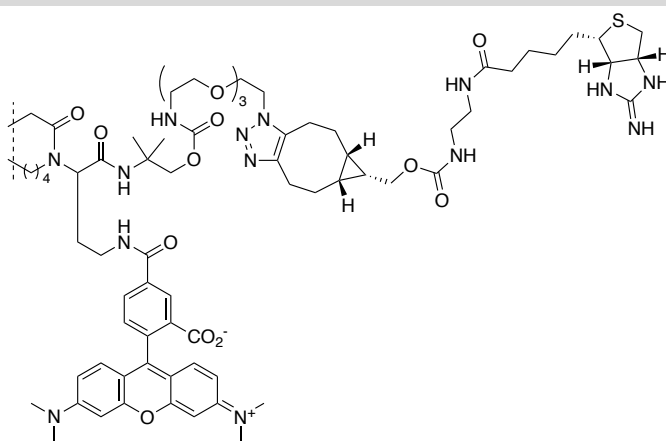
38j

Bioconjugation procedures:

Multicomponent conjugation reaction (Procedure A)

To a solution of trastuzumab (1 equiv., 10 mg/mL, 50 μL in PBS 1x, 3.33 nmol, pH 7.5) was added aldehyde **32a** and *tert*-butyl isocyanide **33a** – both reagents were pipetted from either 0.01 M or 0.1 M stock solutions in DMSO, so as to get a volume of DMSO <10% to that of the final solution. The reaction mixture was then incubated for 16 h at 25 °C, after which a solution of hydroxylamine hydrochloride in PBS 1x (0.5 M, pH 7.5, 2 μL, 292 equiv.) was added. The resulting solution was incubated for 1 h at 25 °C, before the excess of reagent was removed by gel filtration chromatography on Bio-spin P-30 Columns pre-equilibrated with PBS 1/20x (pH 7.5) to give a solution of conjugated trastuzumab that was further analyzed by native mass spectrometry.

'Plug-and-play' conjugation reaction (Procedure B)

Plug step: To a solution of trastuzumab (1 equiv., 10 mg/mL, 50 μL in PBS 1x (pH 7.4)) was added aldehyde **32b** (as a 0.01 M or 0.1 M solution in DMSO – see Procedure A) and **33a-33l** (as a 0.01 M or a 0.1 M solution in DMSO – see Procedure A). The reaction mixture was then incubated for 16 h at 25 °C, after which a solution of hydroxylamine hydrochloride in PBS 1x (0.5 M, pH 7.5, 2 μL, 292 equiv.) was added. The resulting solution was incubated for 1 h at 25 °C, before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 Columns pre-equilibrated with PBS 1x (pH 7.5) to give a solution of trastuzumab-azide which was further used in the play step.

Play step: To the solution of trastuzumab-azide in PBS 1x was added a BCN-derivative **37a-g** (as a 13 mM or 0.1 M solution in DMSO or water, 10 equiv.). The resulting solution was incubated for 20 h at 25 °C, before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 Columns pre-equilibrated with PBS 1/20x (pH 7.4) to give a solution of conjugated trastuzumab that was further analyzed by native mass spectrometry.

Stability of conjugates in human plasma

Conjugate **38a** (25 μL, 1 mg/mL) was mixed with human plasma (25 μL), filled with nitrogen and incubated at 37 °C. Every day at certain time points, aliquots (2 μL) were taken, diluted with water (48 μL), frozen in liquid nitrogen and stored at -20 °C. The resulting samples (50 μL) were then subjected to SDS PAGE analysis (Figure S4; see 'General procedures').

pH stability of the conjugates in buffer

Conjugate **38a** (100 μL, 1 mg/mL) was incubated 5 days at 37 °C in a PBS buffer (1x, 137 mM of NaCl) adjusted to either pH 5, 6, 7, 8 or 9 using HCl (0.1 M solution in H_2O) and NaOH (0.1 M solution in H_2O). After 5 days, the samples were purified by gel filtration chromatography using Bio-spin P-30 columns pre-equilibrated with PBS 1/20x (pH 7.4). The resulting probes were then analyzed by native mass spectrometry and compared to a control (i.e. non-incubated **38a**), which showed a comparable stability in all media.

In vitro cytotoxicity assay

The cytotoxicity of antibody-drug conjugate (ADC) **38g** was evaluated on both HER2-positive and HER2-negative cell lines and compared to that of FDA-approved ADC T-DM1 (**Table 13**).

Table 13: EC_{50} values of ADCs **38g** and **T-DM1** in SK-BR-3 and MDA-MB-231 cancer cell lines (nt = non-toxic at assayed concentration range)

Cell line	38g DAR 1.4	T-DM1 DAR 3.6
SK-BR-3 (HER2+)	56 pM	192 pM
	76 pM	155 pM
Average ± StdDev	66 ± 14 pM	174 ± 26 pM
MDA-MB-231	nt	nt
(HER2-)	nt	nt

Dual functionalization studies

Initially, four types of dual functionalization strategies were envisioned:

i) SPAAC (blue spheres) followed by copper(I)-catalyzed azide alkyne cycloaddition (CuAAC, orange spheres), using conjugates **34bg** and **34bh**;

ii) SPAAC (blue spheres) followed by a [4+1] cycloaddition (green sphere), using conjugate **34bi**;

iii) SPAAC (blue spheres) followed by a [4+2] cycloaddition (purple sphere), using conjugate **34bk**;

iv) Double SPAAC using conjugate **34bj;**

v) Alternatively, we also imagined using TAMRA aldehyde **32a** as a first functionalization point while incorporating a second reactive handle via isocyanide **33j** that could further be functionalized by SPAAC.
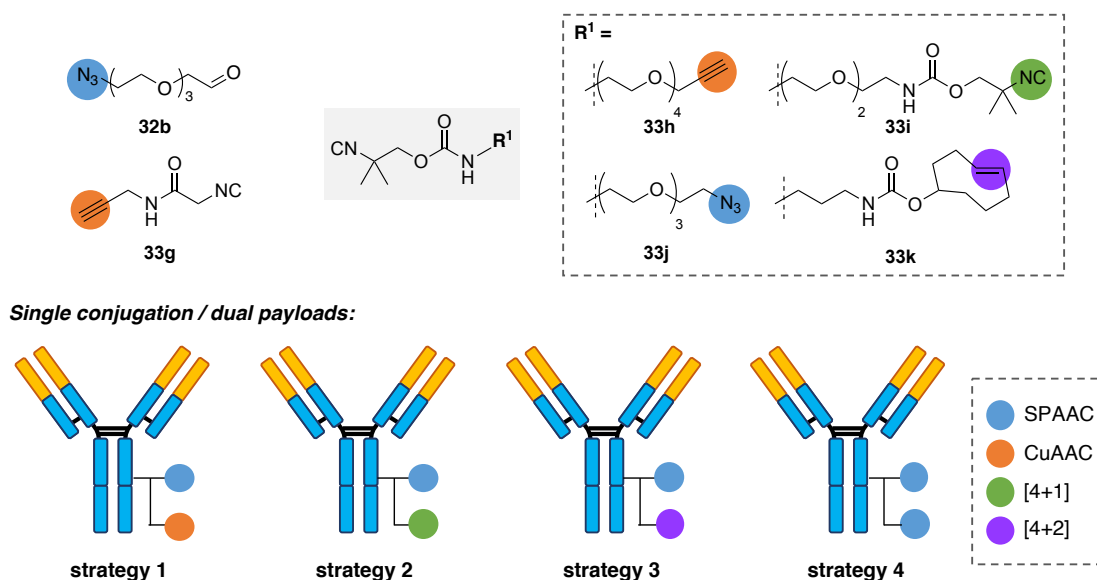
Figure 45: Envisioned strategies for the dual-labelling of trastuzumab.
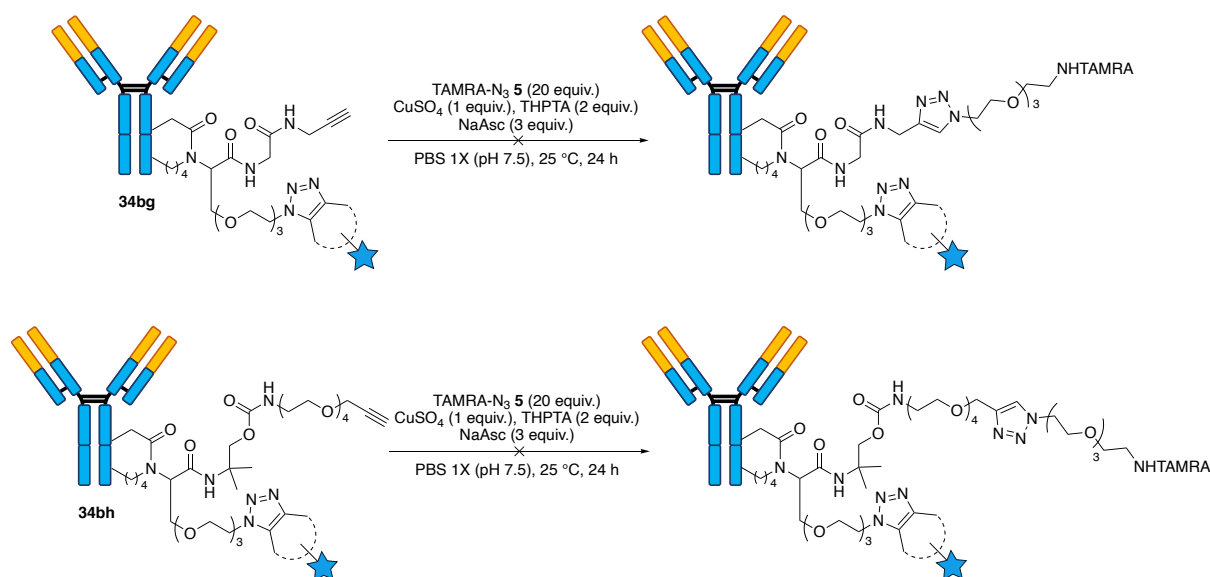
In this first strategy, dual functionalization of trastuzumab conjugates **34bg** and **34bh** was attempted with SPAAC followed by CuAAC.

While BCN-Cy5 **37a** and BCN-iminobiotin **37c** were successfully incorporated by SPAAC (see 'General procedure for the "Plug-and-Play" conjugation of proteins – Play stage'), the terminal alkyne could not be functionalized by CuAAC, either because of partial decomposition (**34bg**) or because no reaction was observed (**34bh**).

Procedure for the CuAAC reaction*:* to a solution of trastuzumab-alkyne **34bg** / **34bh** in PBS 1X (pH 7.5) was added **TAMRA-N₃ 5** (as a 0.1 M solution in DMSO, 20 equiv.). Then, was added $CuSO_4$ (as a 0.01 M solution in H_2O, 1 equiv.) pre-mixed with THPTA (as a 0.01 M solution in H_2O, 2 equiv.) and sodium ascorbate (as a 0.01 M solution in H_2O, 3 equiv.). The resulting solution was incubated for 24 hours at 25 °C before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 Columns pre-equilibrated with PBS 1/20X (pH 7.4) to give a solution of conjugated trastuzumab. The resulting conjugates were then washed 6 times with PBS 1/20X (pH 7.4, 450 µL) containing 1% EDTA on Vivaspin centrifugal concentrators (500 µL, 50 kDa) to remove chelated copper. The conjugates were then analyzed by native mass spectrometry.

We envisioned as a second strategy to perform the Ugi reaction with aldehyde **32b** and diisocyanide **33i** with the aim of functionalizing the resulting conjugate **34bi** by SPAAC followed by [4+1]-cycloaddition with tetrazines **40** and **41**:
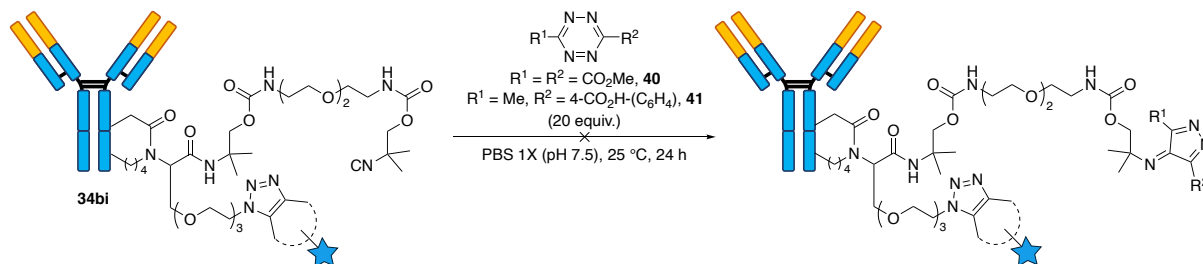
Tetrazines **40** and **41** were employed with the aim of functionalizing the isocyanide moiety in **34bi**, after SPAAC, with a well-documented [4+1] retro-[4+2]-cycloaddition cascade leading to 4*H*-pyrazole products. Unfortunately, both tetrazines proved to be unreactive, even when the [4+1] cycloaddition was attempted prior to SPAAC.

Procedure for the [4+1] cycloaddition*:* To a solution of trastuzumab-isocyanide conjugate in PBS 1X (pH 7.5) was added either tetrazine **40** or **41** (20 equiv., 0.1 M solution in DMSO). The reaction mixture was incubated for 24 hours at 25 °C before the excess of reagent was removed by gel filtration chromatography using Bio-spin P-30 Columns pre-equilibrated with PBS 1/20X (pH 7.4) to give a solution of conjugated trastuzumab which was further analyzed by native mass spectrometry.

iii) SPAAC followed by [4+2]-cycloaddition

We envisioned as a third strategy to perform the Ugi reaction with aldehyde **32b** and TCO-containing isocyanide **33k** with the aim of functionalizing the resulting conjugates **34bk**, respectively, by SPAAC followed by a [4+2]-cycloaddition employing tetrazine **41**.
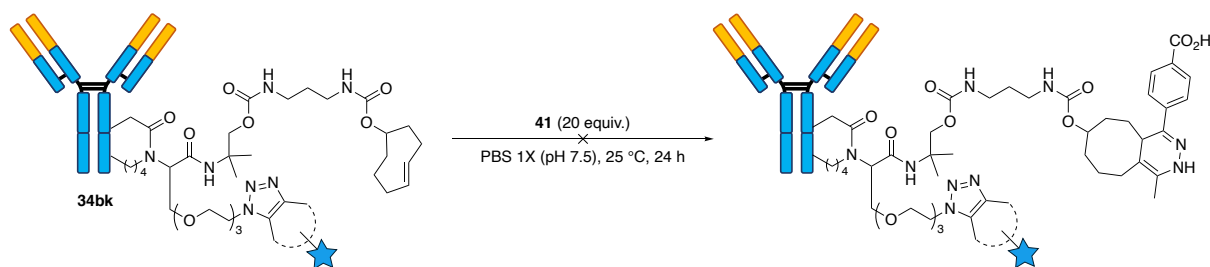
In this case however, the mediocre reactivity of the TCO-containing isocyanide **33k** in the Ugi reaction led to too small quantity of the resulting conjugate **34bk** to attempt the iEDDA step after the first SPAAC functionalization.

iv) Double SPAAC

Conjugate **34bj**, containing two azide groups, was subjected to a double SPAAC with either alkyne **37a** or **37c**, leading to conjugates **38h** and **38i** respectively.
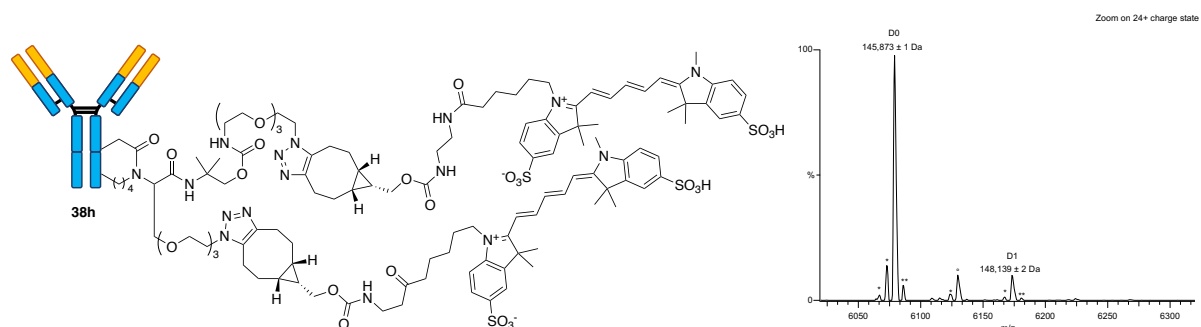
Figure 46: Structure of adduct **38h** and corresponding native mass spectrum, zoomed on the 24+ charge state. The presence of hexose (+162 Da) is represented by a double asterisk, the loss of fucose (-146 Da) is represented by a single asterisk, the species represented by an empty circle has a mass shift of +1,218 Da compared to the D0 species, and could correspond to fragmentation of the D1 payloads.
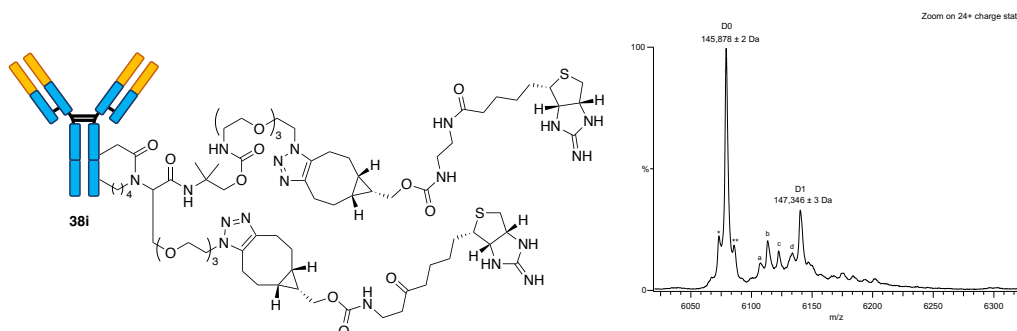
Figure 47: Structure of adduct **38i** and corresponding native mass spectrum, zoomed on the 24+ charge state. The presence of hexose (+162 Da) is represented by a double asterisk, the loss of fucose (-146 Da) is represented by

a single asterisk. The other species have a mass of (a) 146,559 ± 7 Da, (b) 146,703 ± 5 Da, (c) 146,918 ± 6 Da and (d) 147,187 ± 8 Da, and could correspond to fragmentations of the D1 payloads.

See 'General procedure for the "Plug-and-Play" conjugation of proteins (Procedure B) – Play stage' for the SPAAC procedure.

v) Alternative route

We finally used an alternative route to access dually modified conjugates, by performing the multicomponent conjugation step with TAMRA aldehyde **32a**, which acted as a first functionalization point, and isocyanide **33j** that could then be functionalized by SPAAC with alkyne **37c**. The resulting conjugate **38j** thus bears two distinct functional groups: a fluorophore probe and an iminobiotin.
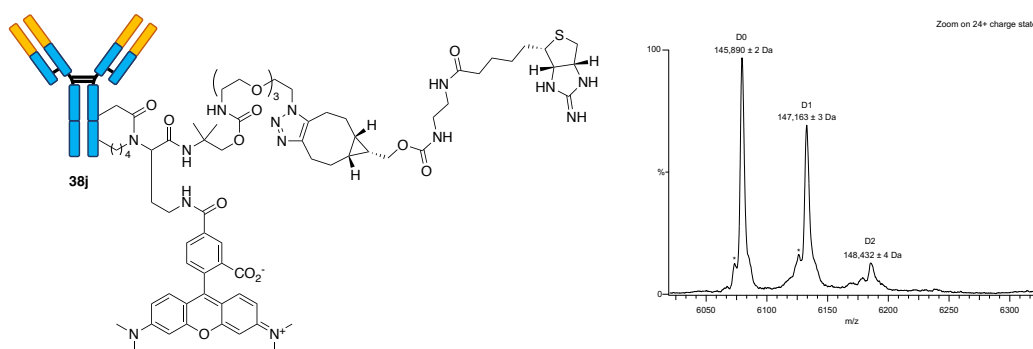
Figure 48: Structure of adduct **38j** and corresponding native mass spectrum, zoomed on the 24+ charged state. The loss of fucose (-146 Da) is represented by a single asterisk.

See 'General procedure for the "Plug-and-Play" conjugation of proteins (Procedure B) – Play stage' for the SPAAC procedure.

Peptide mapping analyses

Sample preparation for peptide mapping analysis

Fifteen micrograms of deglycosylated antibody conjugate **38a** were solubilized in 150 mM NH_4HCO_3, 0.1% RapiGest™ (Waters, Milford, USA) at pH 7.4. Disulfide reduction was performed by incubating the resulting solution with 5 mM DTT for 30 min at 60 °C. Alkylation was performed with 15 mM IAA for 30 min in the dark. After these steps, the samples were split in two for enzymatic digestion using trypsin or pepsin.

Trypsin digestion was performed by adding trypsin (Promega, Madison, USA) to a 1:50 enzyme / substrate ratio. Samples were incubated overnight at 37 °C. The reaction was quenched by adding 1% of TFA. RapiGest™ was eliminated by centrifugation at 10 000 g for 5 min.

For pepsin digestion, pH was decreased to 2.0 prior to pepsin (Promega, Madison, USA) addition. Digestion was performed by adding pepsin at a 1:50 enzyme / substrate ratio. Samples were incubated at 37 °C for 3 h. The reaction was stopped by heating at 95 °C for 10 min. RapiGest™ was eliminated by a centrifugation at 10 000 g for 5 min.

Peptide enrichment

Two strategies were envisioned for the production of enriched batches of conjugated fragments of Trastuzumab (**Figure 49**). A first strategy utilized the conjugate **38c**, obtained from **34bc** and the BCN-iminobiotin derivative **37c**. The conjugate is first digested with trypsin and then loaded on a streptavidin column at pH above 10. The non-modified fragments are then eluted with PBS and water before the conjugated fragments are eluted with an acidic buffer and then analyzed by LCMS-MS. The second strategy inverted the order of events. The conjugate **34bc** is first digested with trypsin. The resulting fragments are then clicked with the BCN-iminobiotin derivative **37c** before being loaded on the streptavidin column. As for the first strategy, the non-conjugated are first eluted with PBS and water, then the conjugated fragments are recovered under acidic conditions and then analyzed by LC-MS / MS.
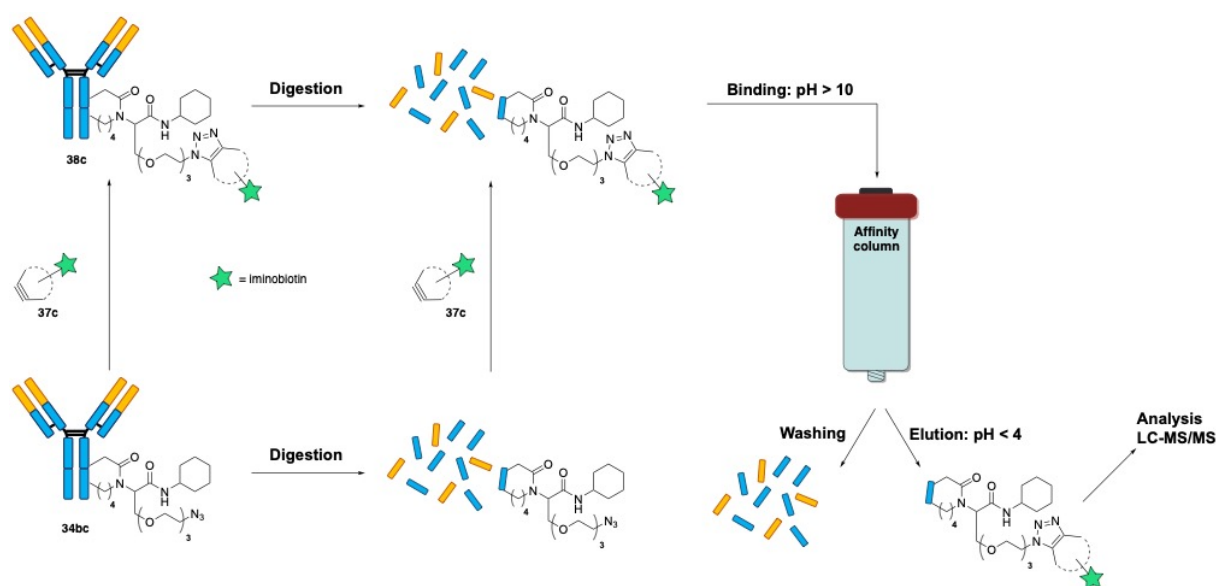
Figure 49: Schematic representation of peptide enrichment strategies (see 'General procedures' for digestion and enrichment procedures)

NanoLC-MS/MS analyses were performed using a nanoAcquity Ulta-Performance-LC (Waters, Wilmslow, UK) coupled to the TripleTOF 5600 mass spectrometer (Sciex, Ontario, Canada). The samples were trapped on a nanoACQUITY UPLC precolumn (C18, 180 μm x 20 mm, 5 μm particle size), and the peptides were separated on a nanoACQUITY UPLC column (C18, 75 μm x 250 mm with 1.7 μm particle size, Waters, Wilmslow, UK). Mobile phase A was 0.1% (v/v) formic acid in water and mobile phase B was 0.1% (v/v) formic acid in acetonitrile. A gradient (3% B for 35 min, 3-85% B for 1 min, 85-3% B for 1 min, maintained 3% B for 13 min) was used at a flow rate of 300 nL/min. The TripleTOF 5600 was operated in the positive mode, with the following settings: ionspray voltage floating (ISVF) 2350 V, curtain gas (CUR) 25, interface heater temperature (IHT) 75, ion source gas 1 (GS1) 8, declustering potential (DP) 80 V. Information-dependent acquisition (IDA) mode was used with Top 10 MS/MS scans. The MS scan accumulation time was set to 250 ms on m/z range [300;1250] and the MS/MS scans to 100 ms on the m/z range [100;1800] in the high sensitivity mode. Switching criteria were set to ions with charge state of 2-4 and an abundance threshold of more than 250 counts, exclusion time was set at 8 s. IDA rolling collision energy script was used for automatically adapting the CE. Mass calibration of the analyser was achieved using peptides from digested BSA. The complete system was fully controlled by AnalystTF 1.7.1 (Sciex).

Conjugation sites identification

Raw data collected were processed and converted into .mgf format. For the **38a** sample, the mgf files of trypsin and pepsin digestions were merged using Mass Spectrometry Data Analysis 2.7.3 (MSDA).[346] The MS/MS data were interpreted using a local Mascot server with MASCOT 2.5.0 algorithm (Matrix Science, London, UK). Spectra were searched with a mass tolerance of 15 ppm for MS and 0.07 Da for MS/MS data, using none (for **38a**) or trypsin (for enriched samples **38c / 34bc**) as enzyme. Carbamidomethylation of cysteine residues and oxidation of methionine residues were specified as variable modifications. Both chemical bioconjugation reactions were set as variable modifications on targeted amino acids: for the Ugi reaction, + 1169 Da for **38a**, +769 Da for enriched samples **38c / 34bc** were added to lysine, glutamic and aspartic acid residues while for the Passerini linker, +1187 Da for **38a**, +787 Da for enriched samples **38c / 34bc** were added to glutamic and aspartic acid of the mAb. Protein identifications were validated with Mascot ion score above 25. Each conjugation sites was

manually validated based on the presence of y-ion and b-ion series and the peak intensity observed on the MS/MS spectra, using Proline 1.5 software.[347] The presence of a characteristic diagnostic fragment of the payload (685 m/z for **38a** and 286 m/z for enriched samples **38c / 34bc**) on the MS/MS spectrum was used as last criteria in order to validate the conjugation sites (see Tables 14-15 and Figure 50).

Table 14: Conjugation sites observed by peptide mapping for the Ugi reaction

			Conjugation sites		
			E1 (HC)	D1 (LC)	K126-D122/E123 (LC)
samples		38a	x	x	x
	enriched	38c	x	x	
		34bc	x	x	

Table 15: Conjugation sites observed by peptide mapping for the Passerini reaction

			Conjugation sites				
			E1 (HC)	E236 (HC)	E348 (HC)	D1 (LC)	E195 (LC)
samples		38a	x	x		x	
	enriched	38c	x			x	x
		34bc	x		x		

Query 7711 - DIQMTQSPSSLSASVGDR

B

286.16 m/z

Query 14425 - PSDEQLKSGTASVVCLLNNFYPR

C

685.30 m/z

Query 7975 - EVQLVESGGGLVQPGGSLR

D

286.16 m/z

Query 15079 - THTCPPCPAPELLGGPSVFLFPPKPK

E

685.30 m/z

Query 2951 - EPQVYTLPPSR

F

286.16 m/z

Query 4609 - DIQMTQSPSSLSASVGDR

G

286.16 m/z
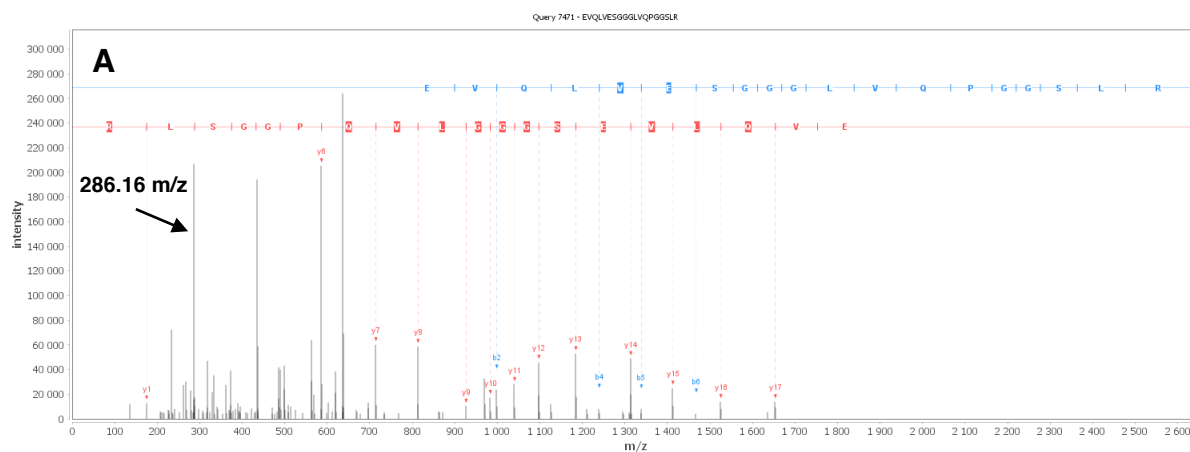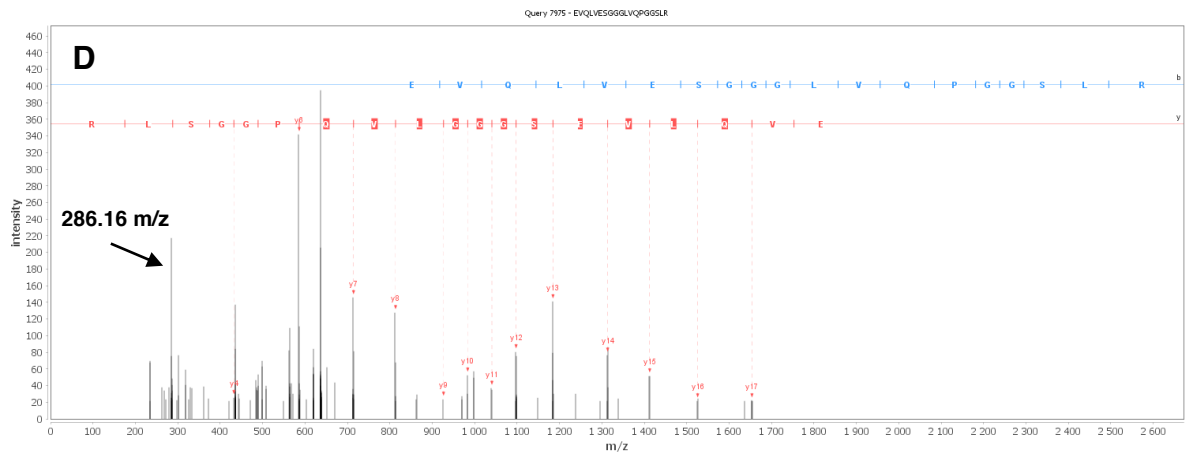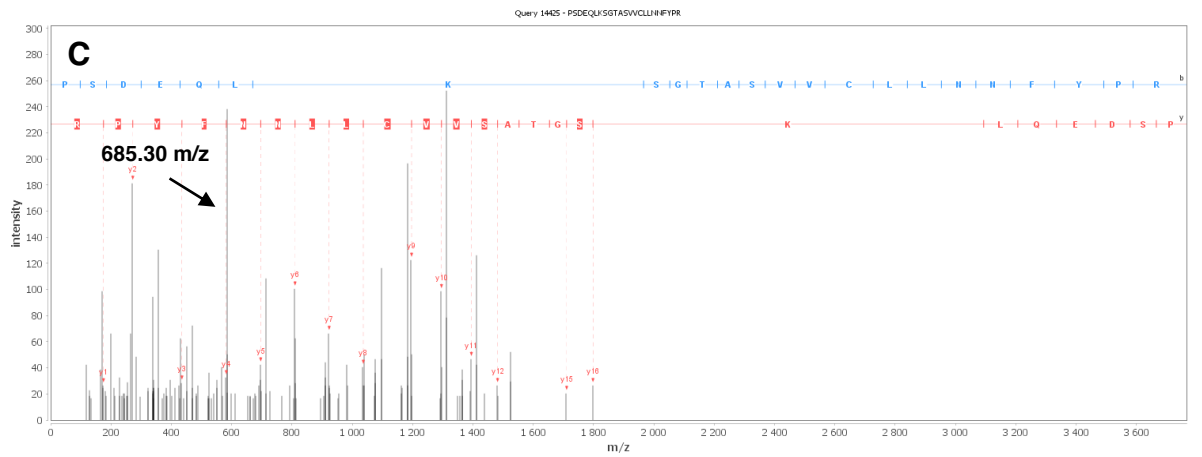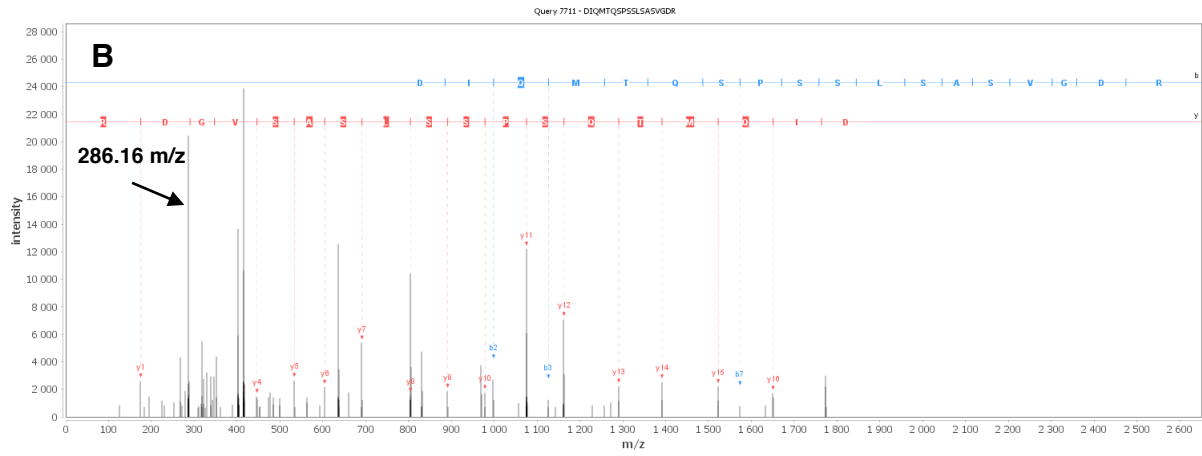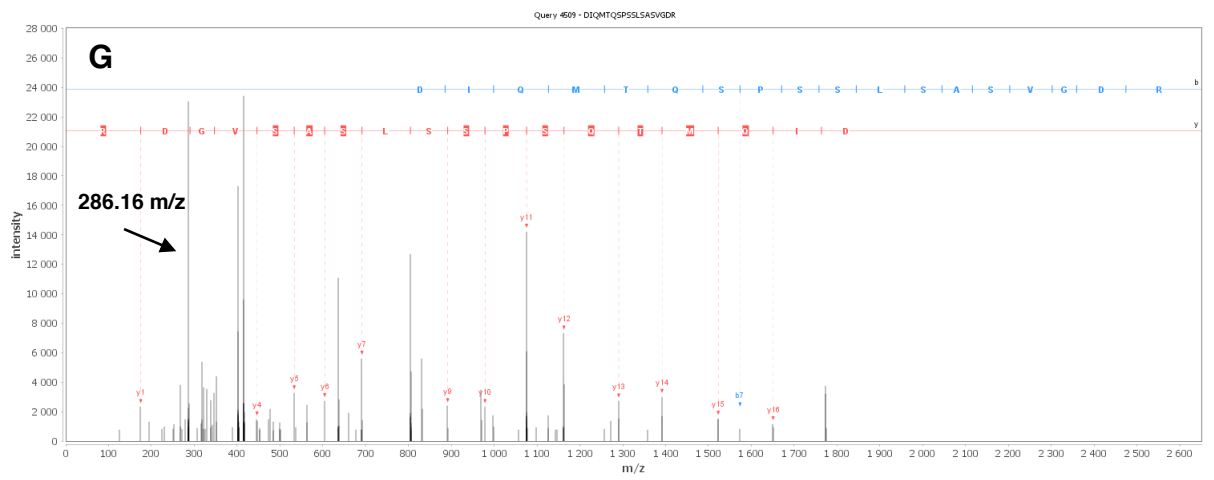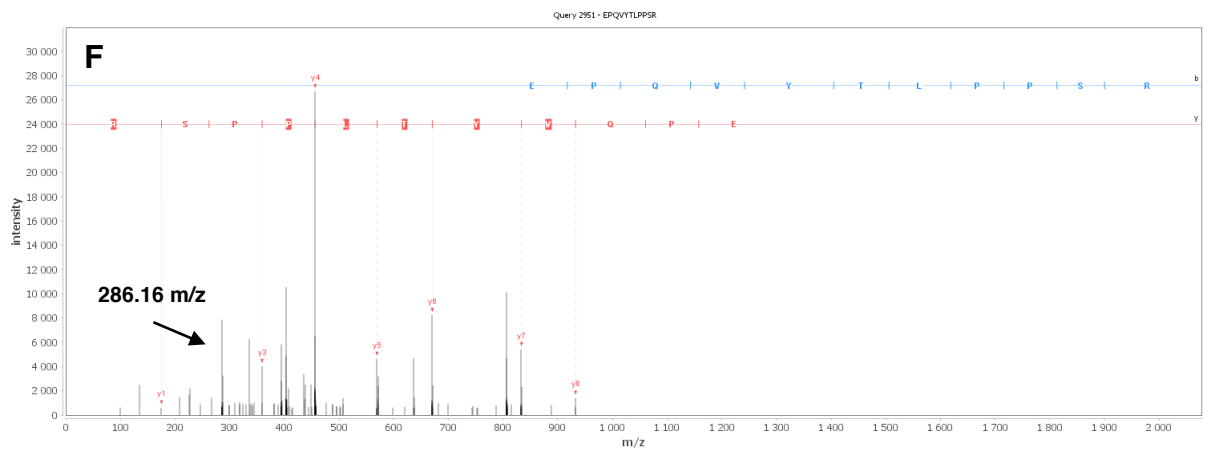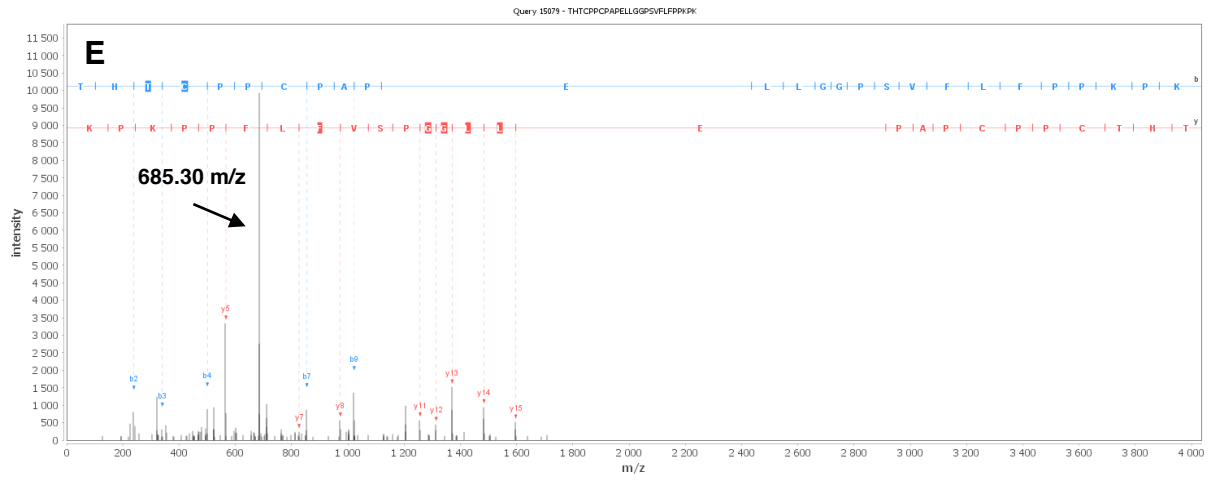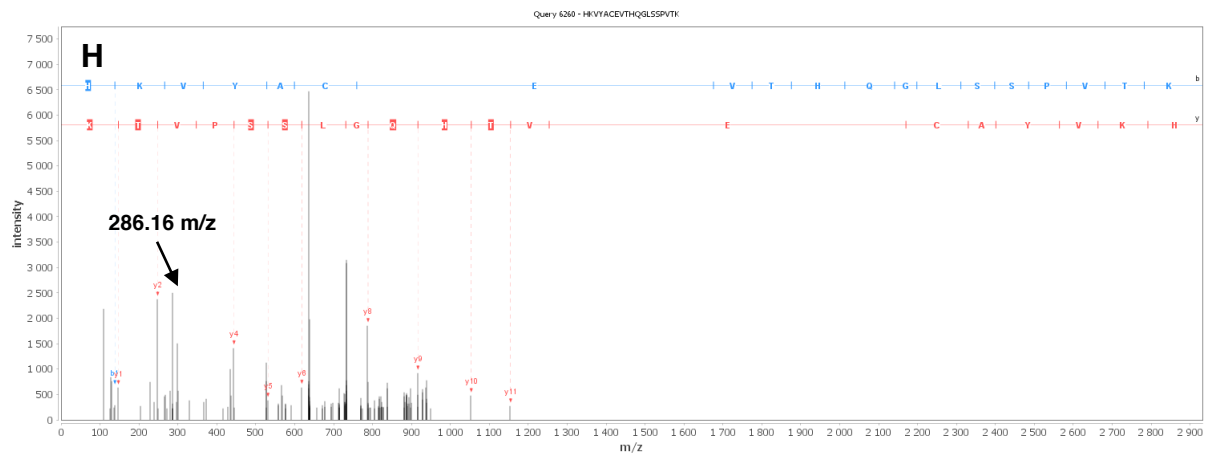
Figure 50: MS/MS spectra of each conjugated peptides validated by peptide mapping. MS/MS spectrum of conjugated peptide from (**A**) Ugi reaction on heavy-chain *N*-ter E1 validated in enriched sample **34bc**. (**B**) Ugi reaction on light-chain *N*-ter D1 validated in enriched sample **38c**. (**C**) Ugi reaction between light-chain K126 and D122 or E123 validated in sample **38a**. (**D**) Passerini reaction on heavy-chain E1 validate in enriched sample **38c**. (**E**) Passerini reaction on heavy-chain E236 validated in sample **38a**. (**F**) Passerini reaction on heavy-chain E348 validated in enriched sample **34bc**. (**G**) Passerini reaction on light-chain D1 validated in enriched sample **34bc**. (**H**) Passerini reaction on light-chain E195 validated in enriched sample **38c**.

RESUME

La modification site-sélective de protéines a fait l'objet de nombreuses études au cours des dernières décennies, notamment dans le domaine thérapeutique. Des résultats probants ont été obtenus avec des protéines recombinantes, pour lesquelles la conjugaison site-spécifique est rendue possible par l'incorporation d'acides aminés naturels (NAA), non-naturels (UAA) ou de séquences peptidiques.[64,66,67] Cependant, en raison du coût élevé de la méthode, les chercheurs se sont aussi tournés vers le développement de méthodes de marquage site-sélectif sur protéines natives. Au cours des dernières années, le développement de nouvelles stratégies régio-sélectives a émergé sur différents types d'acides aminés, tels que la cystéine, la lysine, le tryptophane, la tyrosine ou l'histidine, et s'est avérée efficace sur peptides et protéines.[33,64] Dans le but de poursuivre les efforts déjà menés dans ce domaine, différentes approches pour la modification site-sélective des protéines ont été menées au cours de cette thèse.

Projet 1 : Modification chimiosélective des résidus de cystéines par des iodes hypervalents

Parmi les vingt acides α-aminés naturels protéinogènes, la cystéine est l'une des cibles les plus étudiées en bioconjugaison en raison de sa faible abondance (1.5%) et de la nucléophilie du thiol.[10,11,32] Divers réactifs ont déjà été utilisés pour modifier ce résidu, comme le célèbre maléimide et ses dérivés, les monosulfones ou bissulfones, et les dibromopyridazinediones.[33,65,240] Dans le but de poursuivre les efforts déjà menés dans ce domaine, en collaboration avec le groupe de Jérôme Waser (EPFL, Suisse), nous avons étudié l'utilisation d'iodes hypervalents pour la modification chimiosélective des cystéines. Par le passé, des réactifs à base d'iode hypervalents ont déjà été décrits pour le marquage de cystéines, tryptophanes et méthionines (**Figure 51**). En effet, un alkynyl benziodoxolone fonctionnalisé par un azoture (**1**) a été rapporté pour le profilage protéomique des cystéines, mais n'a pas été étudié en tant que nouvelle méthode chimiosélective pour la modification des protéines ; le 1-[(triisopropylsilyl)éthynyl]-1,2-benziodoxol-3(1H)-one (TIPS-EBX, **2**) en présence d'un catalyseur à base d'or, AuCl(SMe)$_2$, a été décrit pour le marquage sélectif des tryptophanes, et, un sel d'iodonium, le tétrafluoroborate de (1-diazo-2-éthoxy-2-oxoéthyle)(2,4-difluorophényl)iodonium (**3**), peut conjuguer sélectivement les résidus de méthionine en présence de thiourée, de TEMPO et d'acide formique.[53,226,233]
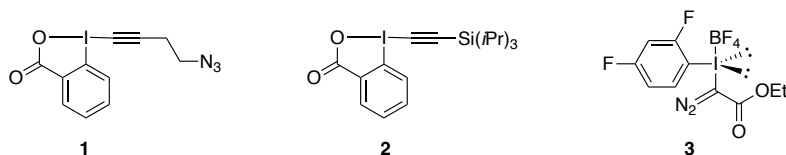
Figure 51 : Structures des iodes hypervalents préalablement utilisés pour la modification des protéines.

Pour la modification chimiosélective des cystéines, deux EBX, un azido-EBX (**1**) et un TMS-EBX (**4**), déjà développés par le groupe de Jérôme Waser sur peptides, ont été testés sur un anticorps modèle, le trastuzumab.

En se basant sur les conditions optimales déjà développées sur peptides, les premières expériences ont été menées avec différents équivalents d'azido-EBX (**1**) dans un tampon PBS à différents pH, températures, temps d'incubation et équivalents. Les expériences ont toujours été effectuées en parallèle sur anticorps réduits et non réduits afin d'évaluer la chimiosélectivité du réactif. Malheureusement, un marquage non spécifique de l'anticorps et une décomposition partielle du produit de la réaction au fil du temps ont été observés. En effet, lorsque les conjugués obtenus ont été caractérisés par une analyse MS native, des pics supplémentaires ont été observés, et ont été corrélés à la décomposition du produit via l'élimination d'acide iodobenzoïque, des pics correspondant à une perte de masse de 250 Da (D') et 500 Da (D'') étant toujours détectés pour une valeur D de DoC donnée (**Figure 52**). Les différentes conditions testées n'ont pu prévenir la dégradation du conjugué, de sorte que l'utilisation de l'azido-EBX (**1**) n'a pas été plus étudiée pour la modification des cystéines.
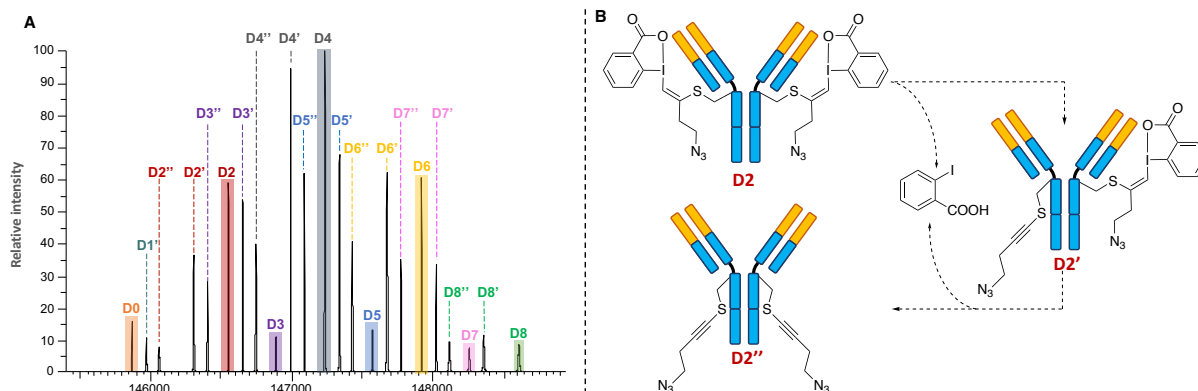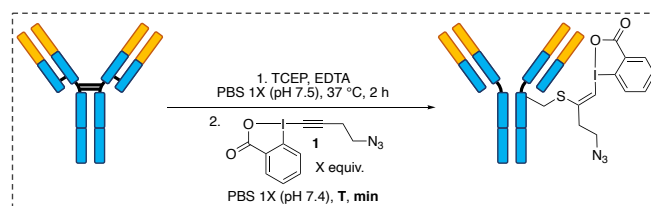
Figure 52 : Modification chimiosélective du trastuzumab avec l'azido-EBX (1). A: spectre de masse déconvolué de l'anticorps réduit et conjugué avec 5 équivalents de 1 à pH 7.5 (DoC moyen = 4.00). B. Justification possible pour la présence de pics supplémentaires.

Passant à un réactif apparenté mais structurellement différent, la réactivité et la sélectivité du *p*-NO$_2$-TMS-EBX (**4**) ont ensuite été évaluées sur le trastuzumab. Comme ce réactif porte un groupe triméthylsilyle labile, il est possible d'incorporer facilement un simple acétylène sur le thiol de la cystéine. Le traitement du trastuzumab réduit avec **4** a conduit à l'obtention d'un mélange complexe de conjugués, suggérant une décomposition partielle de l'anticorps ou des réactions secondaires se produisant lors de l'étape d'éthynylation. La variation des conditions de réaction (tampon, pH, température, temps de réaction, équivalents) sur l'anticorps réduit et non réduit ne pouvant empêcher cette dégradation, il a donc été décidé d'utiliser des EBX moins réactifs, tels que le TMS-EBX (**5**) et le *p*-F-TMS-EBX (**6**). La modification de l'anticorps avec ces nouveaux réactifs a empêché sa dégradation et permis d'obtenir des spectres MS natifs interprétables. Les conditions ont ensuite été optimisées avec les réactifs **5** et **6** afin d'obtenir la conversion la plus élevée possible sans aucune réaction parasite (**Schéma 1**). Ces résultats ont donné lieu à une publication conjointe dans le journal *Angewandte Chemie*.[55]
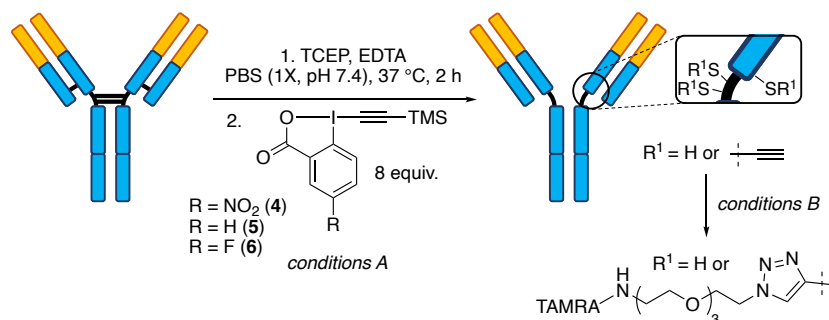
Schéma 1 : Modification chimiosélective du trastuzumab avec les TMS-EBX **4,5** et **6. Conditions A**: **4, 5** ou **6** (8 equiv.), PBS 1X (137 mM NaCl, pH 7.4), 25 °C, 2 min. **Conditions B**: TAMRA-N$_3$ (20 equiv.), CuSO$_4$ (1 equiv.), THPTA (2 equiv.), NaAsc (3 equiv.), PBS 1X (137 mM NaCl, pH 7.4), 25 °C, 24 h.

Après avoir développé de nouvelles méthodes de bioconjugaison adaptées à la modification des cystéines, nous avons cherché à obtenir un marquage régiosélectif des lysines à l'aide d'un pré-marquage sur une cystéine.

Projet 2 : Transfert de cystéine à lysine

La ligation chimique native (NCL) est couramment utilisée pour la construction de longs peptides à partir de peptides plus courts (**Schéma 2**).[274] En effet, un thioester *C*-terminal (**7**) peut réagir avec une cystéine *N*-terminale (**8**) par réaction de *trans*-thioestérification conduisant à la formation d'un nouveau thioester (**9**) reliant les deux peptides. L'attaque nucléophile de l'amine primaire *N*-terminale sur le thioester entraîne un transfert d'acyle *S→N* intramoléculaire, donnant naissance à une nouvelle liaison peptidique (**10**).
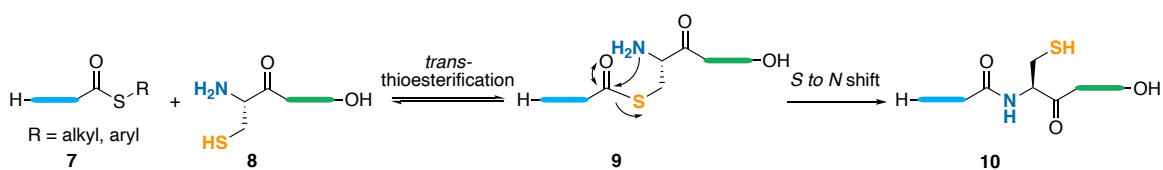
Schéma 2 : Ligation chimique native (NCL)

Appliquée à un anticorps, cette réaction serait en théorie hautement régiosélective car seules les lysines proches des thiols libres pourraient être conjuguées. Comme pour les projets précédents, le trastuzumab a été utilisé comme anticorps modèle et un fluorophore contenant un thioester, le TAMRA-SR (**11**), a été synthétisé et utilisé pour optimiser les conditions de réaction.

Pour réaliser cette réaction sur un anticorps, cinq étapes consécutives sont nécessaires. Dans une première étape, les quatre ponts disulfures inter-chaines de l'anticorps sont complètement réduits en présence de TCEP et d'EDTA. Ensuite, la *trans*-thioestérification a lieu entre le réactif **11** et les thiols libres de l'anticorps, les conditions de réaction s'avérant optimales à 25 °C en 10 minutes avec seulement 5 équivalents de **11**. Pour augmenter la vitesse et l'efficacité du transfert d'acyle, le pH de la réaction a été augmenté à 8.5, et les conjugués incubés à 37 °C pendant 24 heures. Les thioesters qui n'ont pas participé au transfert d'acyle, en raison de l'absence de lysine à proximité, réagissent ensuite avec un excès de L-cystéine, régénérant ainsi les thiols libres de l'anticorps et permettant leur réoxydation dans une dernière étape (**Figure 53**).
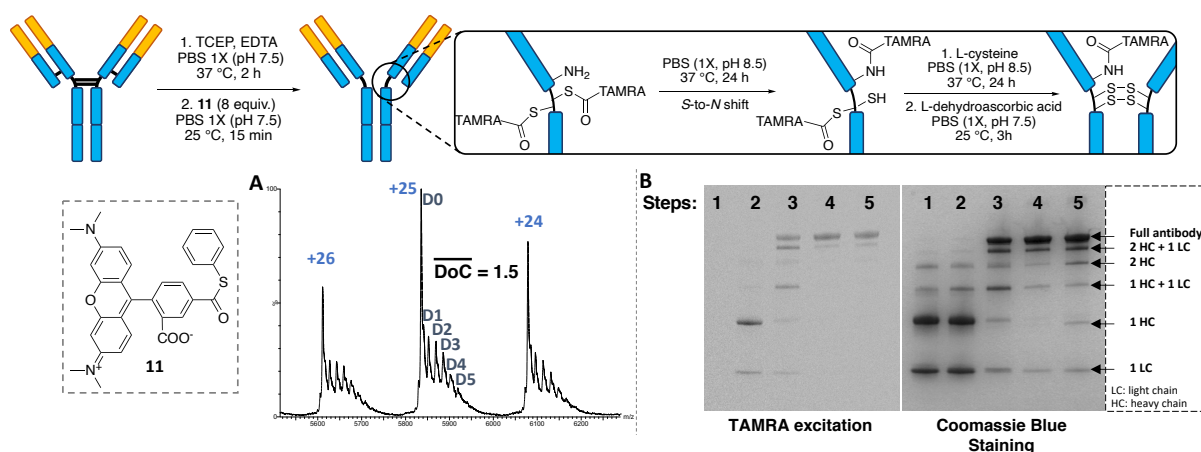
Figure 53 : Application de la ligation chimique native au trastuzumab. **A :** spectre de masse déconvolué du trastuzumab conjugué avec 5 équivalents de 11 (DoC moyen = 1.5). **B :** gel SDS page. Des échantillons ont été récupérés à chacune des étapes et analysés par SDS page. Le gel est révélé par fluorescence et bleu de Coomassie.

Même si toutes les étapes décrites précédemment ont été optimisées indépendamment, il n'a pas été possible de récupérer un anticorps entièrement oxydé à la fin de ces cinq étapes. Deux hypothèses ont été proposées pour expliquer ce résultat : soit le réactif **11** n'est pas adapté pour cette réaction, soit l'excès de cystéine utilisé pour éliminer le thioester restant entraîne également la formation d'un pont disulfure entre la L-cystéine ajoutée et le résidu cystéine de l'anticorps.

Malheureusement, avant que des mesures correctives ne soient appliquées et testées, une stratégie très similaire a été publiée par le groupe de Vijay Chudasama, nous obligeant à abandonner ce projet.[168]

Comme alternative à la modification régiosélective des lysines à l'aide d'un pré-marquage sur les cystéines, nous avons ensuite étudié la modification site-sélective des lysines et aspartates / glutamates par une approche multi-composante.

Projet 3 : Approches multi-composantes pour la conjugaison site-sélective de protéines natives.

Publiée pour la première fois en 1959 par Ivar Karl Ugi, la réaction de Ugi est une réaction multi-composantes entre un carbonyle (**12**), un isonitrile (**13**), une amine (**14**) et un acide (**15**) donnant un composé bis-amide comme produit final (**16**).[283]

Schéma 3 : Réaction de Ugi

La réaction de Ugi a souvent été utilisée par le passé comme outil pour la construction, l'accroche et la macrocyclisation de peptides. Cependant, les conditions généralement utilisées pour la modification des peptides ne sont souvent pas compatibles avec la modification des protéines. En effet, la réaction se fait généralement en présence de méthanol et d'un autre co-solvant organique à haute concentration de réactifs. Néanmoins, quelques méthodes ont déjà été rapportées pour la modification de protéines par cette réaction, et notamment pour l'immobilisation d'enzymes sur différents matériaux.[348–350] Cependant, ces modifications, s'appuyant sur la version à quatre composantes de la Ugi (U-4CR), ne visent qu'un seul acide aminé à la surface des protéines (lysine ou aspartate/glutamate) et nécessitent des quantités excessives de réactifs, allant jusqu'à 4000 équivalents, ainsi que des temps de réaction très longs (jusqu'à 4 jours).[319,320]

En nous basant sur des études précédentes, nous avons proposé d'utiliser la version à quatre centres et trois composantes de la réaction de Ugi (U-4C-3CR) afin de cibler uniquement les résidus lysines et aspartates/glutamates étant spatialement proches. L'U-4C-3CR, réalisée en présence d'un carbonyle (**12**) et d'un isonitrile (**13**), ne devrait ainsi conduire qu'à une macrolactamisation entre les chaînes latérales de ces deux résidus. Même si toutes les

fonctions amines pourraient en théorie réagir avec le carbonyle et conduire à la formation d'imines, seules celles étant proches d'un carboxylate devraient former un produit final stable (**16**). L'addition finale d'une solution d'hydroxylamine devraient permettre de régénérer les amines primaires depuis les imines résiduelles n'ayant pas réagi (**Schéma 4**).
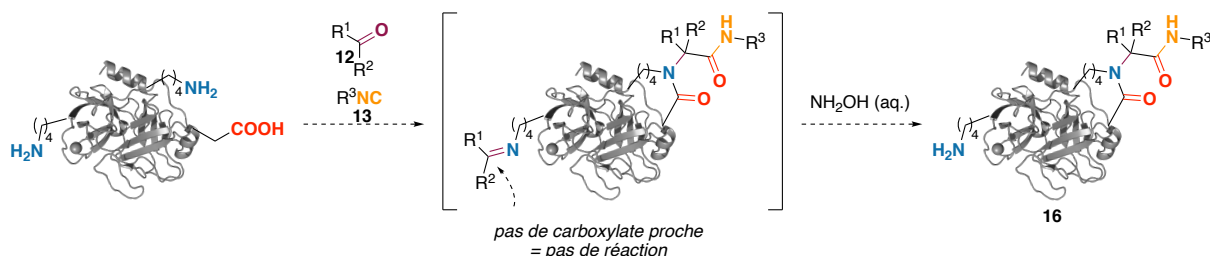
Schéma 4 : Modification des protéines avec la réaction de Ugi

Pour étudier la faisabilité de notre réaction, le trastuzumab, un anticorps de 150 kDa, a été sélectionné comme protéine modèle et modifié avec divers équivalents de TAMRA-aldéhyde (**12a**) et d'isonitrile de *tert*-butyle (**13a**). Le produit de la Ugi a été obtenu rapidement après les premières conditions testées, et l'étude de différents paramètres – tels que le nombre d'équivalents, le temps d'incubation, la température, le pH et le tampon – nous ont permis d'obtenir les conditions optimales suivantes : 45 équivalents de **12a** et **13a**, tampon PBS 1X (137 mM de NaCl), pH 7.5, 16 h, 25 ºC. Différentes expériences contrôles ont été également réalisées, confirmant que la réaction ne pouvait avoir lieu qu'en présence de l'aldéhyde et de l'isonitrile (**Figure 54**).
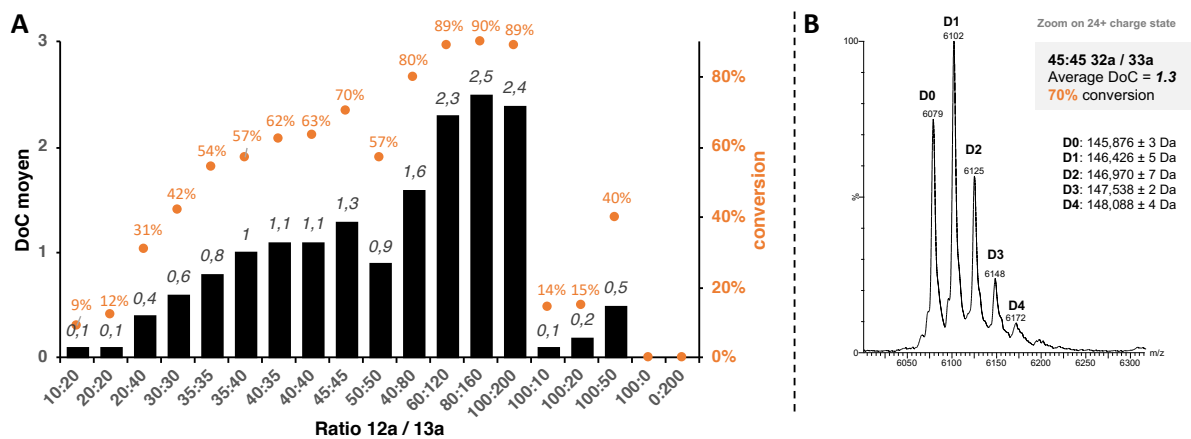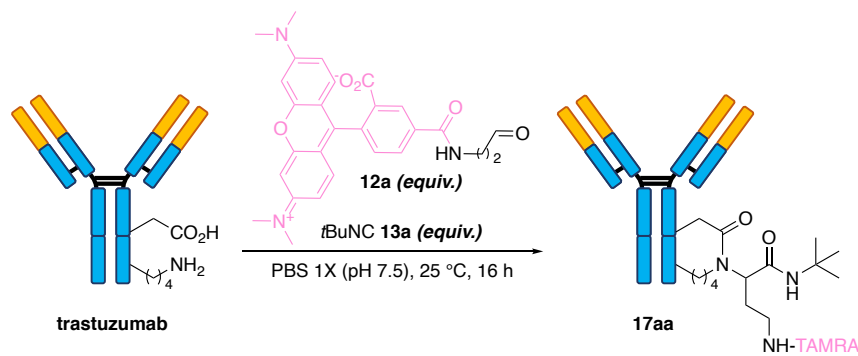
Figure 54 : Réaction de Ugi sur trastuzumab. A : Influence du ratio aldéhyde 12a et isonitrile 13a sur le degré de conjugaison (DoC), déterminé par spectrométrie de masse (MS) native. B : Spectre de masse native de l'anticorps conjugué après déglycosylation (zoom sur l'état de charge 24+). Conditions réactionnelles : 12a (45 equiv.), 13a (45 equiv.), PBS 1X (137 mM NaCl, pH 7.5), 25 °C, 16 h.

Cette réaction ne se limite pas qu'à **12a** et **13a**, plusieurs aldéhydes et isonitriles, comportant divers groupes fonctionnels, ont en effet donné des résultats satisfaisants. Par exemple, la réaction de Ugi a pu être adaptée à une stratégie dite « plug-and-play » (**Figure 55**), qui permet l'introduction d'un azoture dans une première étape (« plug step ») et ensuite sa fonctionnalisation via une cycloaddition avec des alcynes tendus liés à des fluorophores, des molécules cytotoxiques ou des oligonucléotides.
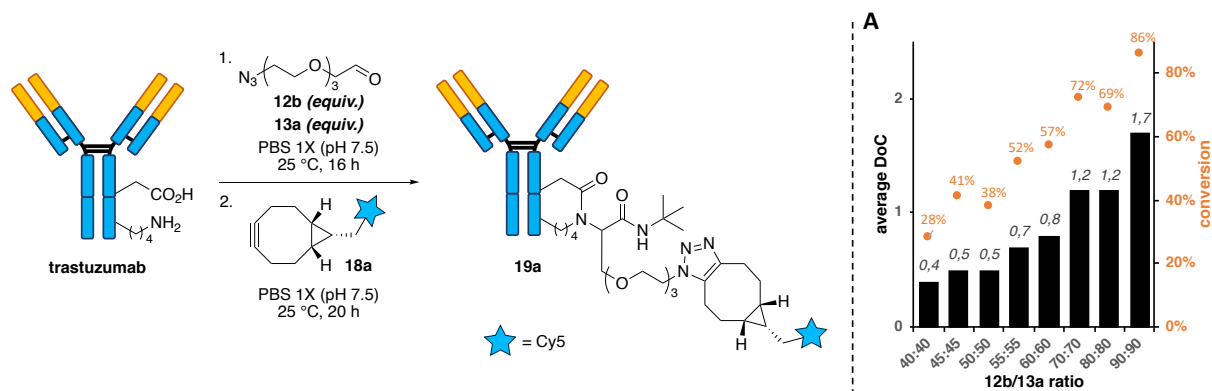
Figure 55 : Stratégie « plug-and-play » : modification du trastuzumab avec un aldéhyde contenant un azoture **12b** et l'isonitrile **13a**, suivi d'une fonctionnalisation avec un alcyne tendu **18a. A** : Optimisation de la « plug step » (DoCs moyen et conversions déterminés by MS native du conjugué 19a). Conditions réactionnelles : « Plug step » : **12b** (X equiv.), **13a** (X equiv.), PBS 1X (137 mM NaCl, pH 7.5), 25 °C, 16 h. « Play step » : **18a** (10 equiv.), PBS 1X (137 mM NaCl, pH 7.5), 25 °C, 16 h.

Ayant développé une stratégie multi-composante efficace, simple d'utilisation et rapide, tolérant divers réactifs et substrats, nous nous sommes tournés vers des études LCMS/MS pour valider le mécanisme de la réaction de bioconjugaison et localiser les sites de modification sur l'anticorps trastuzumab. Par cartographie peptidique, six sites de conjugaison ont été identifiés et trois mécanismes réactionnels validés (**Figure 56**). Le glutamate *N*-terminal (E1) sur la chaîne lourde de l'anticorps et l'aspartate *N*-terminal (D1) sur la chaîne légère ont tous deux été modifiés et ont participé à deux types de réactions multi-composantes : une réaction intramoléculaire U-4C-3CR donnant un β- ou γ- lactame (provenant respectivement de D1 et E1), et une réaction de Passerini n'impliquant que les chaînes latérales carboxylate. Trois sites supplémentaires, correspondant à la modification des glutamates par la réaction de Passerini, ont été observés (chaîne légère E195, et chaînes lourdes E236 et E348) et la réaction inter-résidus attendue, la U-4C-3CR, a été identifiée comme étant responsable de la sixième modification, située sur la chaîne légère et impliquant les chaînes latérales de K126 et D122 ou E123.
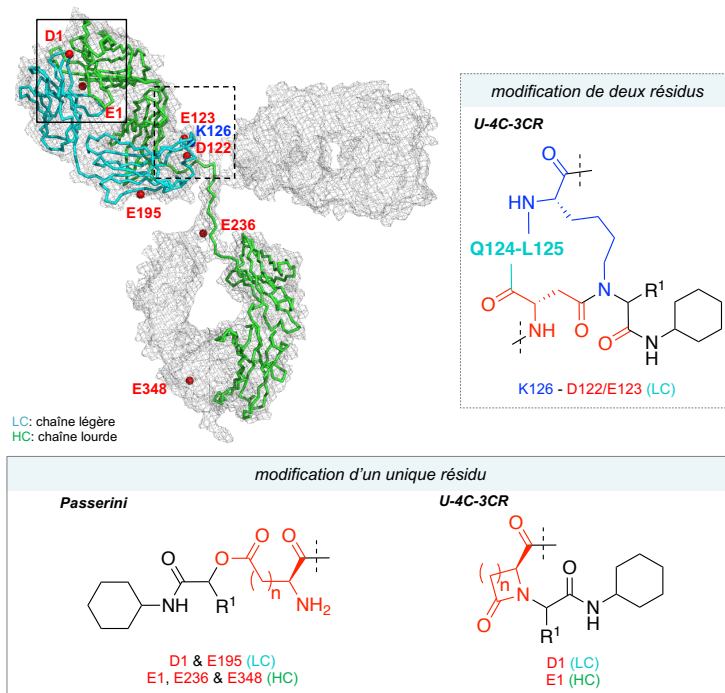
Figure 56 : Sites de modification identifiés sur le trastuzumab par cartographie peptidique.

En conclusion, nous avons démontré que cette stratégie peut donner un accès facile et rapide à des conjugués portant différents groupes fonctionnels à partir d'aldéhydes et d'isonitriles disponibles dans le commerce ou facilement accessibles, ce qui limite le nombre de sites de conjugaison à seulement six sur l'anticorps modèle trastuzumab.

Pour améliorer la chimiosélectivité de la réaction vis-à-vis des résidus aspartate et glutamate, et donc améliorer la sélectivité globale de la réaction, il a été proposé de faire réagir notre anticorps avec une imine préformée ou une oxime et un isonitrile. Cette stratégie limiterait ainsi le type d'acides aminés modifiés et ne laisserait plus qu'une seule réaction avoir lieu (la réaction de Ugi dans ce cas). Les conditions de réaction sont actuellement en cours d'optimisation.

Conclusion

Dans le but de trouver de nouvelles méthodes régiosélectives pour la modification des protéines, trois stratégies ont été étudiées au cours du doctorat. Une première méthode, développée en collaboration avec le groupe de Jérôme Waser et reposant sur l'utilisation des iodes hypervalents a été explorée pour la modification chimiosélective des cystéines. Sur la base de la ligation chimique native, nous avons cherché à modifier sélectivement des lysines

grâce à un pré-marquage sur des cystéines dans un deuxième temps. Finalement, une méthode basée sur la réaction de Ugi visait à cibler simultanément deux acides aminés proches afin de réduire le nombre de sites de modification

REFERENCES

(1)   Means, G. E.; Feeney, R. E. Chemical Modifications of Proteins: History and Applications. *Bioconjug. Chem.* **1990**, *1* (1), 2–12. https://doi.org/10.1021/bc00001a001.

(2)   van Slyke, D. D. Eine Methode zur quantitativen Bestimmung der aliphatischen Aminogruppen; einige Anwendungen derselben in der Chemie der Proteine, des Harns und der Enzyme. *Berichte Dtsch. Chem. Ges.* **1910**, *43* (3), 3170–3181. https://doi.org/10.1002/cber.19100430382.

(3)   Fields, R. The Rapid Determination of Amino Groups with TNBS. In *Methods in Enzymology*; Elsevier, 1972; Vol. 25, pp 464–468. https://doi.org/10.1016/S0076-6879(72)25042-X.

(4)   Ellman, G. L. Tissue Sulfhydryl Groups. *Arch. Biochem. Biophys.* **1959**, *82* (1), 70–77. https://doi.org/10.1016/0003-9861(59)90090-6.

(5)   Hoare, D. G.; Koshland, D. E. A Method for the Quantitative Modification and Estimation of Carboxylic Acid Groups in Proteins. *J. Biol. Chem.* **1967**, *242* (10), 2447–2453.

(6)   Barman, T. E.; Koshland, D. E. A Colorimetric Procedure for the Quantitative Determination of Tryptophan Residues in Proteins. *J. Biol. Chem.* **1967**, *242* (23), 5771–5776.

(7)   Olcott, H. S.; Fraenkel-Conrat, Heinz. Specific Group Reagents for Proteins. *Chem. Rev.* **1947**, *41* (1), 151–197. https://doi.org/10.1021/cr60128a004.

(8)   Herriott, R. M. Reactions of Native Proteins with Chemical Reagents. In *Advances in Protein Chemistry*; Elsevier, 1947; Vol. 3, pp 169–225. https://doi.org/10.1016/S0065-3233(08)60080-7.

(9)   Balls, A. K.; Jansen, E. F. Stoichiometric Inhibition of Chymotrypsin. *Adv. Enzymol. Relat. Subj. Biochem.* **1952**, *13*, 321–343. https://doi.org/10.1002/9780470122587.ch8.

(10)   Hermanson, G. T. *Bioconjugate Techniques*, 2. ed.; Elsevier Acad. Press: Amsterdam, 2008.

(11)   Gauthier, M. A.; Klok, H.-A. Peptide/Protein–Polymer Conjugates: Synthetic Strategies and Design Concepts. *Chem. Commun.* **2008**, No. 23, 2591. https://doi.org/10.1039/b719689j.

(12)   Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem. Rev.* **2015**, *115* (5), 2174–2195. https://doi.org/10.1021/cr500399p.

(13)   Algar, W. R. A Brief Introduction to Traditional Bioconjugate Chemistry. In *Chemoselective and Bioorthogonal Ligation Reactions*; Algar, W. R., Dawson, P. E., Medintz, I. L., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2017; pp 1–36. https://doi.org/10.1002/9783527683451.ch1.

(14)   Baslé, E.; Joubert, N.; Pucheault, M. Protein Chemical Modification on Endogenous Amino Acids. *Chem. Biol.* **2010**, *17* (3), 213–227. https://doi.org/10.1016/j.chembiol.2010.02.008.

(15)   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* (15), 5495–5551. https://doi.org/10.1039/C5CS00048C.

(16)   Jentoft, N.; Dearborn, D. G. Labeling of Proteins by Reductive Methylation Using Sodium Cyanoborohydride. *J. Biol. Chem.* **1979**, *254* (11), 4359–4365.

(17)   Gildersleeve, J. C.; Oyelaran, O.; Simpson, J. T.; Allred, B. Improved Procedure for Direct Coupling of Carbohydrates to Proteins via Reductive Amination. *Bioconjug. Chem.* **2008**, *19* (7), 1485–1490. https://doi.org/10.1021/bc800153t.

(18)   Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. The Use of Esters of N-Hydroxysuccinimide in Peptide Synthesis. *J. Am. Chem. Soc.* **1964**, *86* (9), 1839–1842. https://doi.org/10.1021/ja01063a037.

(19)   Smith, G. P. Kinetics of Amine Modification of Proteins. *Bioconjug. Chem.* **2006**, *17* (2), 501–506. https://doi.org/10.1021/bc0503061.

(20)   Berndt, D. C.; Faburada, A. L. Reaction of Acyl Azide and Amines. Kinetics and Mechanism. *J. Org. Chem.* **1982**, *47* (21), 4167–4169. https://doi.org/10.1021/jo00142a033.

(21)   Ślósarczyk, A. T.; Ramapanicker, R.; Norberg, T.; Baltzer, L. Mixed Pentafluorophenyl and O-Fluorophenyl Esters of Aliphatic Dicarboxylic Acids: Efficient Tools for Peptide and Protein Conjugation. *RSC Adv* **2012**, *2* (3), 908–914. https://doi.org/10.1039/C1RA00530H.

(22)   Pham, G. H.; Ou, W.; Bursulaya, B.; DiDonato, M.; Herath, A.; Jin, Y.; Hao, X.; Loren, J.; Spraggon, G.; Brock, A.; Uno, T.; Geierstanger, B. H.; Cellitti, S. E. Tuning a Protein-Labeling Reaction to Achieve Highly Site Selective Lysine Conjugation. *ChemBioChem* **2018**, *19* (8), 799–804. https://doi.org/10.1002/cbic.201700611.

(23)   Jorbágy, A.; Király, K. Chemical Characterization of Fluorescein Isothiocyanate-Protein Conjugates. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1966**, *124* (1), 166–175. https://doi.org/10.1016/0304-4165(66)90325-4.

(24)   Lowe, A. The Chemistry of Isocyanates. *Proc. R. Soc. Med.* **1970**, *63* (4), 367–368.

(25)   Riggs, J. L.; Seiwald, R. J.; Burckhalter, J. H.; Downs, C. M.; Metcalf, T. G. Isothiocyanate Compounds as Fluorescent Labeling Agents for Immune Serum. *Am. J. Pathol.* **1958**, *34* (6), 1081–1097.

(26)   Tuls, J.; Geren, L.; Millett, F. Fluorescein Isothiocyanate Specifically Modifies Lysine 338 of Cytochrome P-450scc and Inhibits Adrenodoxin Binding. *J. Biol. Chem.* **1989**, *264* (28), 16421–16425.

(27)   Lin, J.-Kun.; Chang, J.-Yoa. Chromophoric Labeling of Amino Acids with 4-Dimethylaminoazobenzene-4'-Sulfonyl Chloride. *Anal. Chem.* **1975**, *47* (9), 1634–1638. https://doi.org/10.1021/ac60359a007.

(28)   Evangelista, R. A.; Pollak, A.; Allore, B.; Templeton, A. F.; Morton, R. C.; Diamandis, E. P. A New Europium Chelate for Protein Labelling and Time-Resolved Fluorometric Applications. *Clin. Biochem.* **1988**, *21* (3), 173–178. https://doi.org/10.1016/0009-9120(88)90006-9.

(29)   Andreoni, A.; Bottiroli, G.; Colasanti, A.; Giangarè, M. C.; Riccio, P.; Roberti, G.; Vaghi, P. Fluorochromes with Long-Lived Fluorescence as Potential Labels for Pulsed Laser Immunocytofluorometry: Photophysical Characterization of Pyrene Derivatives. *J. Biochem. Biophys. Methods* **1994**, *29* (2), 157–172. https://doi.org/10.1016/0165-022X(94)90052-3.

(30)   Bozler, H.; Jany, K. D.; Pfleiderer, G. Synthesis and Application of a Fluorescent Imido Ester for Specific Labelling of Amino Groups in Proteins. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1983**, *749* (3), 238–243. https://doi.org/10.1016/0167-4838(83)90230-3.

(31)   Lee, Y. C.; Stowell, C. P.; Krantz, M. J. 2-Imino-2-Methoxyethyl 1-Thioglycosides: New Reagents for Attaching Sugars to Proteins. *Biochemistry* **1976**, *15* (18), 3956–3963. https://doi.org/10.1021/bi00663a008.

(32)   Poole, L. B. The Basics of Thiols and Cysteines in Redox Biology and Chemistry. *Free Radic. Biol. Med.* **2015**, *80*, 148–157. https://doi.org/10.1016/j.freeradbiomed.2014.11.013.

(33)   Cal, P. M. S. D.; Bernardes, G. J. L.; Gois, P. M. P. Cysteine-Selective Reactions for Antibody Conjugation. *Angew. Chem. Int. Ed.* **2014**, *53* (40), 10585–10587. https://doi.org/10.1002/anie.201405702.

(34)   Smyth, D.; Blumenfeld, O.; Konigsberg, W. Reactions of N-Ethylmaleimide with Peptides and Amino Acids. *Biochem. J.* **1964**, *91* (3), 589–595. https://doi.org/10.1042/bj0910589.

(35)   Partis, M. D.; Griffiths, D. G.; Roberts, G. C.; Beechey, R. B. Cross-Linking of Protein by w-Maleimido Alkanoyl N-Hydroxysuccinimido Esters. *J. Protein Chem.* **1983**, *2* (3), 263–277. https://doi.org/10.1007/BF01025358.

(36)   Ravasco, J. M. J. M.; Faustino, H.; Trindade, A.; Gois, P. M. P. Bioconjugation with Maleimides: A Useful Tool for Chemical Biology. *Chem. – Eur. J.* **2019**, *25* (1), 43–59. https://doi.org/10.1002/chem.201803174.

(37)   Gorin, G.; Martic, P. A.; Doughty, G. Kinetics of the Reaction of N-Ethylmaleimide with Cysteine and Some Congeners. *Arch. Biochem. Biophys.* **1966**, *115* (3), 593–597.

https://doi.org/10.1016/0003-9861(66)90079-8.

(38)   Brewer, C. F.; Riehm, J. P. Evidence for Possible Nonspecific Reactions between N-Ethylmaleimide and Proteins. *Anal. Biochem.* **1967**, *18* (2), 248–255. https://doi.org/10.1016/0003-2697(67)90007-3.

(39)   Chin, C. C. Q.; Wold, F. Some Chemical Properties of Carboxymethyl Derivatives of Amino Acids. *Arch. Biochem. Biophys.* **1975**, *167* (2), 448–451. https://doi.org/10.1016/0003-9861(75)90486-5.

(40)   Renault, K.; Fredy, J. W.; Renard, P.-Y.; Sabot, C. Covalent Modification of Biomolecules through Maleimide-Based Labeling Strategies. *Bioconjug. Chem.* **2018**, *29* (8), 2497–2513. https://doi.org/10.1021/acs.bioconjchem.8b00252.

(41)   Smith, M. E. B.; Schumacher, F. F.; Ryan, C. P.; Tedaldi, L. M.; Papaioannou, D.; Waksman, G.; Caddick, S.; Baker, J. R. Protein Modification, Bioconjugation, and Disulfide Bridging Using Bromomaleimides. *J. Am. Chem. Soc.* **2010**, *132* (6), 1960–1965. https://doi.org/10.1021/ja908610s.

(42)   Schumacher, F. F.; Nobles, M.; Ryan, C. P.; Smith, M. E. B.; Tinker, A.; Caddick, S.; Baker, J. R. In Situ Maleimide Bridging of Disulfides and a New Approach to Protein PEGylation. *Bioconjug. Chem.* **2011**, *22* (2), 132–136. https://doi.org/10.1021/bc1004685.

(43)   Chudasama, V.; Smith, M. E. B.; Schumacher, F. F.; Papaioannou, D.; Waksman, G.; Baker, J. R.; Caddick, S. Bromopyridazinedione-Mediated Protein and Peptide Bioconjugation. *Chem. Commun.* **2011**, *47* (31), 8781. https://doi.org/10.1039/c1cc12807h.

(44)   Kalia, D.; Pawar, S. P.; Thopate, J. S. Stable and Rapid Thiol Bioconjugation by Light-Triggered Thiomaleimide Ring Hydrolysis. *Angew. Chem. Int. Ed.* **2017**, *56* (7), 1885–1889. https://doi.org/10.1002/anie.201609733.

(45)   Kalia, D.; Malekar, P. V.; Parthasarathy, M. Exocyclic Olefinic Maleimides: Synthesis and Application for Stable and Thiol-Selective Bioconjugation. *Angew. Chem. Int. Ed.* **2016**, *55* (4), 1432–1435. https://doi.org/10.1002/anie.201508118.

(46)   Bernardim, B.; Cal, P. M. S. D.; Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Albuquerque, I. S.; Perkins, E.; Corzana, F.; Burtoloso, A. C. B.; Jiménez-Osés, G.; Bernardes, G. J. L. Stoichiometric and Irreversible Cysteine-Selective Protein Modification Using Carbonylacrylic Reagents. *Nat. Commun.* **2016**, *7* (1), 13128. https://doi.org/10.1038/ncomms13128.

(47)   Ariyasu, S.; Hayashi, H.; Xing, B.; Chiba, S. Site-Specific Dual Functionalization of Cysteine Residue in Peptides and Proteins with 2-Azidoacrylates. *Bioconjug. Chem.* **2017**, *28* (4), 897–902. https://doi.org/10.1021/acs.bioconjchem.7b00024.

(48)   Bernardim, B.; Dunsmore, L.; Li, H.; Hocking, B.; Nuñez-Franco, R.; Navo, C. D.; Jiménez-Osés, G.; Burtoloso, A. C. B.; Bernardes, G. J. L. Precise Installation of Diazo-Tagged Side-Chains on Proteins to Enable In Vitro and In-Cell Site-Specific Labeling. *Bioconjug. Chem.* **2020**, *31* (6), 1604–1610. https://doi.org/10.1021/acs.bioconjchem.0c00232.

(49)   Morales-Sanfrutos, J.; Lopez-Jaramillo, J.; Ortega-Muñoz, M.; Megia-Fernandez, A.; Perez-Balderas, F.; Hernandez-Mateo, F.; Santoyo-Gonzalez, F. Vinyl Sulfone: A Versatile Function for Simple Bioconjugation and Immobilization. *Org Biomol Chem* **2010**, *8* (3), 667–675. https://doi.org/10.1039/B920576D.

(50)   Gil de Montes, E.; Jiménez-Moreno, E.; Oliveira, B. L.; Navo, C. D.; Cal, P. M. S. D.; Jiménez-Osés, G.; Robina, I.; Moreno-Vargas, A. J.; Bernardes, G. J. L. Azabicyclic Vinyl Sulfones for Residue-Specific Dual Protein Labelling. *Chem. Sci.* **2019**, *10* (16), 4515–4522. https://doi.org/10.1039/C9SC00125E.

(51)   Kasper, M.; Glanz, M.; Stengl, A.; Penkert, M.; Klenk, S.; Sauer, T.; Schumacher, D.; Helma, J.; Krause, E.; Cardoso, M. C.; Leonhardt, H.; Hackenberger, C. P. R. Cysteine-Selective Phosphonamidate Electrophiles for Modular Protein Bioconjugations. *Angew. Chem. Int. Ed.* **2019**, *58* (34), 11625–11630. https://doi.org/10.1002/anie.201814715.

(52)   Baumann, A. L.; Schwagerus, S.; Broi, K.; Kemnitz-Hassanin, K.; Stieger, C. E.; Trieloff, N.; Schmieder, P.; Hackenberger, C. P. R. Chemically Induced Vinylphosphonothiolate

Electrophiles for Thiol-Thiol Bioconjugations. *J. Am. Chem. Soc.* **2020**, jacs.0c03426. https://doi.org/10.1021/jacs.0c03426.

(53)   Abegg, D.; Frei, R.; Cerato, L.; Prasad Hari, D.; Wang, C.; Waser, J.; Adibekian, A. Proteome-Wide Profiling of Targets of Cysteine Reactive Small Molecules by Using Ethynyl Benziodoxolone Reagents. *Angew. Chem. Int. Ed.* **2015**, *54* (37), 10852–10857. https://doi.org/10.1002/anie.201505641.

(54)   Tessier, R.; Ceballos, J.; Guidotti, N.; Simonet-Davin, R.; Fierz, B.; Waser, J. "Doubly Orthogonal" Labeling of Peptides and Proteins. *Chem* **2019**, *5* (8), 2243–2263. https://doi.org/10.1016/j.chempr.2019.06.022.

(55)   Tessier, R.; Nandi, R. K.; Dwyer, B. G.; Abegg, D.; Sornay, C.; Ceballos, J.; Erb, S.; Cianferani, S.; Wagner, A.; Chaubet, G.; Adibekian, A.; Waser, J. Ethynylation of Cysteines from Peptides to Proteins in Living Cells. *Angew. Chem. Int. Ed.* **2020**. https://doi.org/10.1002/anie.202002626.

(56)   Tamura, T.; Hamachi, I. Recent Progress in Design of Protein-Based Fluorescent Biosensors and Their Cellular Applications. *ACS Chem. Biol.* **2014**, *9* (12), 2708–2717. https://doi.org/10.1021/cb500661v.

(57)   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* (2), 460–475. https://doi.org/10.1016/j.xphs.2015.11.015.

(58)   Nischan, N.; Hackenberger, C. P. R. Site-Specific PEGylation of Proteins: Recent Developments. *J. Org. Chem.* **2014**, *79* (22), 10727–10733. https://doi.org/10.1021/jo502136n.

(59)   Lambert, J. M.; Berkenblit, A. Antibody–Drug Conjugates for Cancer Treatment. *Annu. Rev. Med.* **2018**, *69* (1), 191–207. https://doi.org/10.1146/annurev-med-061516-121357.

(60)   Beck, A.; Goetsch, L.; Dumontet, C.; Corvaïa, N. Strategies and Challenges for the next Generation of Antibody–Drug Conjugates. *Nat. Rev. Drug Discov.* **2017**, *16* (5), 315–337. https://doi.org/10.1038/nrd.2016.268.

(61)   Schumacher, D.; Hackenberger, C. P. R.; Leonhardt, H.; Helma, J. Current Status: Site-Specific Antibody Drug Conjugates. *J. Clin. Immunol.* **2016**, *36* (S1), 100–107. https://doi.org/10.1007/s10875-016-0265-6.

(62)   Gautier, V.; Boumeester, A. J.; Lössl, P.; Heck, A. J. R. Lysine Conjugation Properties in Human IgGs Studied by Integrating High-Resolution Native Mass Spectrometry and Bottom-up Proteomics. *PROTEOMICS* **2015**, *15* (16), 2756–2765. https://doi.org/10.1002/pmic.201400462.

(63)   Sakamoto, S.; Hamachi, I. Recent Progress in Chemical Modification of Proteins. *Anal. Sci.* **2019**, *35* (1), 5–27. https://doi.org/10.2116/analsci.18R003.

(64)   Rawale, D. G.; Thakur, K.; Adusumalli, S. R.; Rai, V. Chemical Methods for Selective Labeling of Proteins. *Eur. J. Org. Chem.* **2019**, *2019* (40), 6749–6763. https://doi.org/10.1002/ejoc.201900801.

(65)   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* (5), 630–640. https://doi.org/10.1002/asia.200800427.

(66)   Hoyt, E. A.; Cal, P. M. S. D.; Oliveira, B. L.; Bernardes, G. J. L. Contemporary Approaches to Site-Selective Protein Modification. *Nat. Rev. Chem.* **2019**, *3* (3), 147–171. https://doi.org/10.1038/s41570-019-0079-1.

(67)   Spicer, C. D.; Davis, B. G. Selective Chemical Protein Modification. *Nat. Commun.* **2014**, *5* (1), 4740. https://doi.org/10.1038/ncomms5740.

(68)   Isenegger, P. G.; Davis, B. G. Concepts of Catalysis in Site-Selective Protein Modifications. *J. Am. Chem. Soc.* **2019**, jacs.8b13187. https://doi.org/10.1021/jacs.8b13187.

(69)   Saxon, E. Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **2000**, *287* (5460), 2007–2010. https://doi.org/10.1126/science.287.5460.2007.

(70)   Bednarek, C.; Wehl, I.; Jung, N.; Schepers, U.; Bräse, S. The Staudinger Ligation.

Chem. Rev. **2020**, acs.chemrev.9b00665. https://doi.org/10.1021/acs.chemrev.9b00665.

(71)   Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. Incorporation of Azides into Recombinant Proteins for Chemoselective Modification by the Staudinger Ligation. *Proc. Natl. Acad. Sci.* **2002**, *99* (1), 19–24. https://doi.org/10.1073/pnas.012583299.

(72)   Tsao, M.-L.; Tian, F.; Schultz, P. G. Selective Staudinger Modification of Proteins Containing P-Azidophenylalanine. *ChemBioChem* **2005**, *6* (12), 2147–2149. https://doi.org/10.1002/cbic.200500314.

(73)   Serwa, R.; Wilkening, I.; Del Signore, G.; Mühlberg, M.; Claußnitzer, I.; Weise, C.; Gerrits, M.; Hackenberger, C. P. R. Chemoselective Staudinger-Phosphite Reaction of Azides for the Phosphorylation of Proteins. *Angew. Chem. Int. Ed.* **2009**, *48* (44), 8234–8239. https://doi.org/10.1002/anie.200902118.

(74)   Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. A Comparative Study of Bioorthogonal Reactions with Azides. *ACS Chem. Biol.* **2006**, *1* (10), 644–648. https://doi.org/10.1021/cb6003228.

(75)   Bernardes, G. J. L.; Linderoth, L.; Doores, K. J.; Boutureira, O.; Davis, B. G. Site-Selective Traceless Staudinger Ligation for Glycoprotein Synthesis Reveals Scope and Limitations. *ChemBioChem* **2011**, *12* (9), 1383–1386. https://doi.org/10.1002/cbic.201100125.

(76)   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 Engl.* **2002**, *41* (14), 2596–2599. https://doi.org/10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO;2-4.

(77)   Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **2002**, *67* (9), 3057–3064. https://doi.org/10.1021/jo011148j.

(78)   Beatty, K. E.; Tirrell, D. A. Two-Color Labeling of Temporally Defined Protein Populations in Mammalian Cells. *Bioorg. Med. Chem. Lett.* **2008**, *18* (22), 5995–5999. https://doi.org/10.1016/j.bmcl.2008.08.046.

(79)   Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. Adding Amino Acids with Novel Reactivity to the Genetic Code of *Saccharomyces Cerevisiae*. *J. Am. Chem. Soc.* **2003**, *125* (39), 11782–11783. https://doi.org/10.1021/ja0370037.

(80)   Link, A. J.; Tirrell, D. A. Cell Surface Labeling of *Escherichia c Oli* via Copper(I)-Catalyzed [3+2] Cycloaddition. *J. Am. Chem. Soc.* **2003**, *125* (37), 11164–11165. https://doi.org/10.1021/ja036765z.

(81)   Link, A. J.; Vink, M. K. S.; Tirrell, D. A. Presentation and Detection of Azide Functionality in Bacterial Cell Surface Proteins. *J. Am. Chem. Soc.* **2004**, *126* (34), 10598–10602. https://doi.org/10.1021/ja047629c.

(82)   Link, A. J.; Vink, M. K. S.; Agard, N. J.; Prescher, J. A.; Bertozzi, C. R.; Tirrell, D. A. Discovery of Aminoacyl-TRNA Synthetase Activity through Cell-Surface Display of Noncanonical Amino Acids. *Proc. Natl. Acad. Sci.* **2006**, *103* (27), 10180–10185. https://doi.org/10.1073/pnas.0601167103.

(83)   Beatty, K. E.; Liu, J. C.; Xie, F.; Dieterich, D. C.; Schuman, E. M.; Wang, Q.; Tirrell, D. A. Fluorescence Visualization of Newly Synthesized Proteins in Mammalian Cells. *Angew. Chem. Int. Ed.* **2006**, *45* (44), 7364–7367. https://doi.org/10.1002/anie.200602114.

(84)   Krogager, T. P.; Ernst, R. J.; Elliott, T. S.; Calo, L.; Beránek, V.; Ciabatti, E.; Spillantini, M. G.; Tripodi, M.; Hastings, M. H.; Chin, J. W. Labeling and Identifying Cell-Specific Proteomes in the Mouse Brain. *Nat. Biotechnol.* **2018**, *36* (2), 156–159. https://doi.org/10.1038/nbt.4056.

(85)   Kennedy, D. C.; McKay, C. S.; Legault, M. C. B.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P. Cellular Consequences of Copper Complexes Used To Catalyze Bioorthogonal Click Reactions. *J. Am. Chem. Soc.* **2011**, *133* (44), 17993–18001. https://doi.org/10.1021/ja2083027.

(86)   Wittig, G.; Krebs, A. Zur Existenz niedergliedriger Cycloalkine, I. *Chem. Ber.* **1961**, *94*

(12), 3260–3275. https://doi.org/10.1002/cber.19610941213.

(87)   Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J. Am. Chem. Soc.* **2004**, *126* (46), 15046–15047. https://doi.org/10.1021/ja044996f.

(88)   Codelli, J. A.; Baskin, J. M.; Agard, N. J.; Bertozzi, C. R. Second-Generation Difluorinated Cyclooctynes for Copper-Free Click Chemistry. *J. Am. Chem. Soc.* **2008**, *130* (34), 11486–11493. https://doi.org/10.1021/ja803086r.

(89)   Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. Copper-Free Click Chemistry for Dynamic in Vivo Imaging. *Proc. Natl. Acad. Sci.* **2007**, *104* (43), 16793–16797. https://doi.org/10.1073/pnas.0707090104.

(90)   Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. In Vivo Imaging of Membrane-Associated Glycans in Developing Zebrafish. *Science* **2008**, *320* (5876), 664–667. https://doi.org/10.1126/science.1155106.

(91)   Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G.-J. Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. *Angew. Chem. Int. Ed.* **2008**, *47* (12), 2253–2255. https://doi.org/10.1002/anie.200705456.

(92)   Debets, M. F.; van Berkel, S. S.; Schoffelen, S.; Rutjes, F. P. J. T.; van Hest, J. C. M.; van Delft, F. L. Aza-Dibenzocyclooctynes for Fast and Efficient Enzyme PEGylation via Copper-Free (3+2) Cycloaddition. *Chem Commun* **2010**, *46* (1), 97–99. https://doi.org/10.1039/B917797C.

(93)   Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R. Rapid Cu-Free Click Chemistry with Readily Synthesized Biarylazacyclooctynones. *J. Am. Chem. Soc.* **2010**, *132* (11), 3688–3690. https://doi.org/10.1021/ja100014q.

(94)   Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J. A.; Rutjes, F. P. J. T.; van Hest, J. C. M.; Lefeber, D. J.; Friedl, P.; van Delft, F. L. Readily Accessible Bicyclononynes for Bioorthogonal Labeling and Three-Dimensional Imaging of Living Cells. *Angew. Chem. Int. Ed.* **2010**, *49* (49), 9422–9425. https://doi.org/10.1002/anie.201003761.

(95)   Plass, T.; Milles, S.; Koehler, C.; Schultz, C.; Lemke, E. A. Genetically Encoded Copper-Free Click Chemistry. *Angew. Chem. Int. Ed.* **2011**, *50* (17), 3878–3881. https://doi.org/10.1002/anie.201008178.

(96)   Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. Genetic Encoding of Bicyclononynes and *Trans* -Cyclooctenes for Site-Specific Protein Labeling in Vitro and in Live Mammalian Cells via Rapid Fluorogenic Diels–Alder Reactions. *J. Am. Chem. Soc.* **2012**, *134* (25), 10317–10320. https://doi.org/10.1021/ja302832g.

(97)   Blackman, M. L.; Royzen, M.; Fox, J. M. Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels–Alder Reactivity. *J. Am. Chem. Soc.* **2008**, *130* (41), 13518–13519. https://doi.org/10.1021/ja8053805.

(98)   Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A. Tetrazine-Based Cycloadditions: Application to Pretargeted Live Cell Imaging. *Bioconjug. Chem.* **2008**, *19* (12), 2297–2299. https://doi.org/10.1021/bc8004446.

(99)   Taylor, M. T.; Blackman, M. L.; Dmitrenko, O.; Fox, J. M. Design and Synthesis of Highly Reactive Dienophiles for the Tetrazine– *Trans-* Cyclooctene Ligation. *J. Am. Chem. Soc.* **2011**, *133* (25), 9646–9649. https://doi.org/10.1021/ja201844c.

(100) Seitchik, J. L.; Peeler, J. C.; Taylor, M. T.; Blackman, M. L.; Rhoads, T. W.; Cooley, R. B.; Refakis, C.; Fox, J. M.; Mehl, R. A. Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific *in Vivo* Bioorthogonal Ligation with *Trans* -Cyclooctenes. *J. Am. Chem. Soc.* **2012**, *134* (6), 2898–2901. https://doi.org/10.1021/ja2109745.

(101) Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W. Genetically Encoded Norbornene Directs Site-Specific Cellular Protein Labelling via a Rapid Bioorthogonal Reaction. *Nat. Chem.* **2012**, *4* (4), 298–304. https://doi.org/10.1038/nchem.1250.

(102) Kaya, E.; Vrabel, M.; Deiml, C.; Prill, S.; Fluxa, V. S.; Carell, T. A Genetically Encoded Norbornene Amino Acid for the Mild and Selective Modification of Proteins in a Copper-Free Click Reaction. *Angew. Chem. Int. Ed.* **2012**, *51* (18), 4466–4469. https://doi.org/10.1002/anie.201109252.

(103) Plass, T.; Milles, S.; Koehler, C.; Szymański, J.; Mueller, R.; Wießler, M.; Schultz, C.; Lemke, E. A. Amino Acids for Diels-Alder Reactions in Living Cells. *Angew. Chem. Int. Ed.* **2012**, *51* (17), 4166–4170. https://doi.org/10.1002/anie.201108231.

(104) Karver, M. R.; Weissleder, R.; Hilderbrand, S. A. Bioorthogonal Reaction Pairs Enable Simultaneous, Selective, Multi-Target Imaging. *Angew. Chem. Int. Ed.* **2012**, *51* (4), 920–922. https://doi.org/10.1002/anie.201104389.

(105) Song, W.; Wang, Y.; Qu, J.; Madden, M. M.; Lin, Q. A Photoinducible 1,3-Dipolar Cycloaddition Reaction for Rapid, Selective Modification of Tetrazole-Containing Proteins. *Angew. Chem. Int. Ed.* **2008**, *47* (15), 2832–2835. https://doi.org/10.1002/anie.200705805.

(106) Song, W.; Wang, Y.; Qu, J.; Lin, Q. Selective Functionalization of a Genetically Encoded Alkene-Containing Protein via "Photoclick Chemistry" in Bacterial Cells. *J. Am. Chem. Soc.* **2008**, *130* (30), 9654–9655. https://doi.org/10.1021/ja803598e.

(107) Ramil, C. P.; Lin, Q. Photoclick Chemistry: A Fluorogenic Light-Triggered in Vivo Ligation Reaction. *Curr. Opin. Chem. Biol.* **2014**, *21*, 89–95. https://doi.org/10.1016/j.cbpa.2014.05.024.

(108) Yu, Z.; Pan, Y.; Wang, Z.; Wang, J.; Lin, Q. Genetically Encoded Cyclopropene Directs Rapid, Photoclick-Chemistry-Mediated Protein Labeling in Mammalian Cells. *Angew. Chem. Int. Ed.* **2012**, *51* (42), 10600–10604. https://doi.org/10.1002/anie.201205352.

(109) Wang, J.; Zhang, W.; Song, W.; Wang, Y.; Yu, Z.; Li, J.; Wu, M.; Wang, L.; Zang, J.; Lin, Q. A Biosynthetic Route to Photoclick Chemistry on Proteins. *J. Am. Chem. Soc.* **2010**, *132* (42), 14812–14818. https://doi.org/10.1021/ja104350y.

(110) Dirksen, A.; Dawson, P. E. Rapid Oxime and Hydrazone Ligations with Aromatic Aldehydes for Biomolecular Labeling. *Bioconjug. Chem.* **2008**, *19* (12), 2543–2548. https://doi.org/10.1021/bc800310p.

(111) 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. *Bioconjug. Chem.* **1992**, *3* (2), 138–146. https://doi.org/10.1021/bc00014a008.

(112) Cornish, V. W.; Hahn, K. M.; Schultz, P. G. Site-Specific Protein Modification Using a Ketone Handle. *J. Am. Chem. Soc.* **1996**, *118* (34), 8150–8151. https://doi.org/10.1021/ja961216x.

(113) Zeng, H.; Xie, J.; Schultz, P. G. Genetic Introduction of a Diketone-Containing Amino Acid into Proteins. *Bioorg. Med. Chem. Lett.* **2006**, *16* (20), 5356–5359. https://doi.org/10.1016/j.bmcl.2006.07.094.

(114) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. Addition of the Keto Functional Group to the Genetic Code of Escherichia Coli. *Proc. Natl. Acad. Sci.* **2003**, *100* (1), 56–61. https://doi.org/10.1073/pnas.0234824100.

(115) Zhang, Z.; Smith, B. A. C.; Wang, L.; Brock, A.; Cho, C.; Schultz, P. G. A New Strategy for the Site-Specific Modification of Proteins *in Vivo* †. *Biochemistry* **2003**, *42* (22), 6735–6746. https://doi.org/10.1021/bi0300231.

(116) Chin, J. W. An Expanded Eukaryotic Genetic Code. *Science* **2003**, *301* (5635), 964–967. https://doi.org/10.1126/science.1084772.

(117) Liu, H.; Wang, L.; Brock, A.; Wong, C.-H.; Schultz, P. G. A Method for the Generation of Glycoprotein Mimetics. *J. Am. Chem. Soc.* **2003**, *125* (7), 1702–1703. https://doi.org/10.1021/ja029433n.

(118) Ye, S.; Köhrer, C.; Huber, T.; Kazmi, M.; Sachdev, P.; Yan, E. C. Y.; Bhagat, A.; RajBhandary, U. L.; Sakmar, T. P. Site-Specific Incorporation of Keto Amino Acids into Functional G Protein-Coupled Receptors Using Unnatural Amino Acid Mutagenesis. *J. Biol.*

Chem. **2008**, *283* (3), 1525–1533. https://doi.org/10.1074/jbc.M707355200.

(119)   Hutchins, B. M.; Kazane, S. A.; Staflin, K.; Forsyth, J. S.; Felding-Habermann, B.; Schultz, P. G.; Smider, V. V. Site-Specific Coupling and Sterically Controlled Formation of Multimeric Antibody Fab Fragments with Unnatural Amino Acids. *J. Mol. Biol.* **2011**, *406* (4), 595–603. https://doi.org/10.1016/j.jmb.2011.01.011.

(120)   Kim, C. H.; Axup, J. Y.; Dubrovska, A.; Kazane, S. A.; Hutchins, B. A.; Wold, E. D.; Smider, V. V.; Schultz, P. G. Synthesis of Bispecific Antibodies Using Genetically Encoded Unnatural Amino Acids. *J. Am. Chem. Soc.* **2012**, *134* (24), 9918–9921. https://doi.org/10.1021/ja303904e.

(121)   Huang, Y.; Wan, W.; Russell, W. K.; Pai, P.-J.; Wang, Z.; Russell, D. H.; Liu, W. Genetic Incorporation of an Aliphatic Keto-Containing Amino Acid into Proteins for Their Site-Specific Modifications. *Bioorg. Med. Chem. Lett.* **2010**, *20* (3), 878–880. https://doi.org/10.1016/j.bmcl.2009.12.077.

(122)   Antos, J.; Francis, M. Transition Metal Catalyzed Methods for Site-Selective Protein Modification. *Curr. Opin. Chem. Biol.* **2006**, *10* (3), 253–262. https://doi.org/10.1016/j.cbpa.2006.04.009.

(123)   Ojida, A.; Tsutsumi, H.; Kasagi, N.; Hamachi, I. Suzuki Coupling for Protein Modification. *Tetrahedron Lett.* **2005**, *46* (19), 3301–3305. https://doi.org/10.1016/j.tetlet.2005.03.094.

(124)   Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. A Genetically Encoded Boronate-Containing Amino Acid. *Angew. Chem. Int. Ed.* **2008**, *47* (43), 8220–8223. https://doi.org/10.1002/anie.200803240.

(125)   Chalker, J. M.; Wood, C. S. C.; Davis, B. G. A Convenient Catalyst for Aqueous and Protein Suzuki–Miyaura Cross-Coupling. *J. Am. Chem. Soc.* **2009**, *131* (45), 16346–16347. https://doi.org/10.1021/ja907150m.

(126)   Spicer, C. D.; Davis, B. G. Palladium-Mediated Site-Selective Suzuki–Miyaura Protein Modification at Genetically Encoded Aryl Halides. *Chem. Commun.* **2011**, *47* (6), 1698. https://doi.org/10.1039/c0cc04970k.

(127)   Wang, Y.-S.; Russell, W. K.; Wang, Z.; Wan, W.; Dodd, L. E.; Pai, P.-J.; Russell, D. H.; Liu, W. R. The de Novo Engineering of Pyrrolysyl-TRNA Synthetase for Genetic Incorporation of l-Phenylalanine and Its Derivatives. *Mol. Biosyst.* **2011**, *7* (3), 714. https://doi.org/10.1039/c0mb00217h.

(128)   Spicer, C. D.; Triemer, T.; Davis, B. G. Palladium-Mediated Cell-Surface Labeling. *J. Am. Chem. Soc.* **2012**, *134* (2), 800–803. https://doi.org/10.1021/ja209352s.

(129)   Spicer, C. D.; Davis, B. G. Rewriting the Bacterial Glycocalyx via Suzuki–Miyaura Cross-Coupling. *Chem. Commun.* **2013**, *49* (27), 2747. https://doi.org/10.1039/c3cc38824g.

(130)   Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G. Self-Liganded Suzuki-Miyaura Coupling for Site-Selective Protein PEGylation. *Angew. Chem. Int. Ed.* **2013**, *52* (14), 3916–3921. https://doi.org/10.1002/anie.201208626.

(131)   Gao, Z.; Gouverneur, V.; Davis, B. G. Enhanced Aqueous Suzuki–Miyaura Coupling Allows Site-Specific Polypeptide [18] F-Labeling. *J. Am. Chem. Soc.* **2013**, *135* (37), 13612–13615. https://doi.org/10.1021/ja4049114.

(132)   Li, J.; Lin, S.; Wang, J.; Jia, S.; Yang, M.; Hao, Z.; Zhang, X.; Chen, P. R. Ligand-Free Palladium-Mediated Site-Specific Protein Labeling Inside Gram-Negative Bacterial Pathogens. *J. Am. Chem. Soc.* **2013**, *135* (19), 7330–7338. https://doi.org/10.1021/ja402424j.

(133)   Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. Regioselective Carbon–Carbon Bond Formation in Proteins with Palladium Catalysis; New Protein Chemistry by Organometallic Chemistry. *ChemBioChem* **2007**, *8* (2), 159–159. https://doi.org/10.1002/cbic.200790002.

(134)   Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. Site-Specific

Functionalization of Proteins by Organopalladium Reactions. *ChemBioChem* **2007**, *8* (2), 232–238. https://doi.org/10.1002/cbic.200600432.

(135) Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q. Copper-Free Sonogashira Cross-Coupling for Functionalization of Alkyne-Encoded Proteins in Aqueous Medium and in Bacterial Cells. *J. Am. Chem. Soc.* **2011**, *133* (39), 15316–15319. https://doi.org/10.1021/ja2066913.

(136) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. Allyl Sulfides Are Privileged Substrates in Aqueous Cross-Metathesis: Application to Site-Selective Protein Modification. *J. Am. Chem. Soc.* **2008**, *130* (30), 9642–9643. https://doi.org/10.1021/ja8026168.

(137) Lin, Y. A.; Davis, B. G. The Allylic Chalcogen Effect in Olefin Metathesis. *Beilstein J. Org. Chem.* **2010**, *6*, 1219–1228. https://doi.org/10.3762/bjoc.6.140.

(138) Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. Enabling Olefin Metathesis on Proteins: Chemical Methods for Installation of S-Allyl Cysteine. *Chem. Commun.* **2009**, No. 25, 3714. https://doi.org/10.1039/b908004j.

(139) Lin, Y. A.; Chalker, J. M.; Davis, B. G. Olefin Cross-Metathesis on Proteins: Investigation of Allylic Chalcogen Effects and Guiding Principles in Metathesis Partner Selection. *J. Am. Chem. Soc.* **2010**, *132* (47), 16805–16811. https://doi.org/10.1021/ja104994d.

(140) Lin, Y. A.; Boutureira, O.; Lercher, L.; Bhushan, B.; Paton, R. S.; Davis, B. G. Rapid Cross-Metathesis for Reversible Protein Modifications via Chemical Access to *Se* -Allyl-Selenocysteine in Proteins. *J. Am. Chem. Soc.* **2013**, *135* (33), 12156–12159. https://doi.org/10.1021/ja403191g.

(141) Li, C.; Tebo, A.; Gautier, A. Fluorogenic Labeling Strategies for Biological Imaging. *Int. J. Mol. Sci.* **2017**, *18* (7), 1473. https://doi.org/10.3390/ijms18071473.

(142) Thorn, K. Genetically Encoded Fluorescent Tags. *Mol. Biol. Cell* **2017**, *28* (7), 848–857. https://doi.org/10.1091/mbc.e16-07-0504.

(143) Dean, K. M.; Palmer, A. E. Advances in Fluorescence Labeling Strategies for Dynamic Cellular Imaging. *Nat. Chem. Biol.* **2014**, *10* (7), 512–523. https://doi.org/10.1038/nchembio.1556.

(144) Crivat, G.; Taraska, J. W. Imaging Proteins inside Cells with Fluorescent Tags. *Trends Biotechnol.* **2012**, *30* (1), 8–16. https://doi.org/10.1016/j.tibtech.2011.08.002.

(145) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A General Method for the Covalent Labeling of Fusion Proteins with Small Molecules in Vivo. *Nat. Biotechnol.* **2003**, *21* (1), 86–89. https://doi.org/10.1038/nbt765.

(146) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K. An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chem. Biol.* **2008**, *15* (2), 128–136. https://doi.org/10.1016/j.chembiol.2008.01.007.

(147) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **2008**, *3* (6), 373–382. https://doi.org/10.1021/cb800025k.

(148) Hori, Y.; Ueno, H.; Mizukami, S.; Kikuchi, K. Photoactive Yellow Protein-Based Protein Labeling System with Turn-On Fluorescence Intensity. *J. Am. Chem. Soc.* **2009**, *131* (46), 16610–16611. https://doi.org/10.1021/ja904800k.

(149) Griffin, B. A. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **1998**, *281* (5374), 269–272. https://doi.org/10.1126/science.281.5374.269.

(150) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. *J. Am. Chem. Soc.* **2002**, *124* (21), 6063–6076. https://doi.org/10.1021/ja017687n.

(151) Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. Selective

Recognition of Protein Tetraserine Motifs with a Cell-Permeable, Pro-Fluorescent Bis-Boronic Acid. *J. Am. Chem. Soc.* **2009**, *131* (2), 438–439. https://doi.org/10.1021/ja807872s.

(152) Nonaka, H.; Tsukiji, S.; Ojida, A.; Hamachi, I. Non-Enzymatic Covalent Protein Labeling Using a Reactive Tag. *J. Am. Chem. Soc.* **2007**, *129* (51), 15777–15779. https://doi.org/10.1021/ja074176d.

(153) Uchinomiya, S.; Nonaka, H.; Fujishima, S.; Tsukiji, S.; Ojida, A.; Hamachi, I. Site-Specific Covalent Labeling of His-Tag Fused Proteins with a Reactive Ni(Ii)–NTA Probe. *Chem. Commun.* **2009**, No. 39, 5880. https://doi.org/10.1039/b912025d.

(154) Martos-Maldonado, M. C.; Hjuler, C. T.; Sørensen, 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* (1), 3307. https://doi.org/10.1038/s41467-018-05695-3.

(155) Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. π-Clamp-Mediated Cysteine Conjugation. *Nat. Chem.* **2016**, *8* (2), 120–128. https://doi.org/10.1038/nchem.2413.

(156) Rashidian, M.; Dozier, J. K.; Distefano, M. D. Enzymatic Labeling of Proteins: Techniques and Approaches. *Bioconjug. Chem.* **2013**, *24* (8), 1277–1294. https://doi.org/10.1021/bc400102w.

(157) Li, C.; Wang, L.-X. Chemoenzymatic Methods for the Synthesis of Glycoproteins. *Chem. Rev.* **2018**, *118* (17), 8359–8413. https://doi.org/10.1021/acs.chemrev.8b00238.

(158) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. Introducing Genetically Encoded Aldehydes into Proteins. *Nat. Chem. Biol.* **2007**, *3* (6), 321–322. https://doi.org/10.1038/nchembio878.

(159) Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. Sortase-Mediated Protein Ligation: A New Method for Protein Engineering. *J. Am. Chem. Soc.* **2004**, *126* (9), 2670–2671. https://doi.org/10.1021/ja039915e.

(160) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. Sortagging: A Versatile Method for Protein Labeling. *Nat. Chem. Biol.* **2007**, *3* (11), 707–708. https://doi.org/10.1038/nchembio.2007.31.

(161) Duckworth, B. P.; Zhang, Z.; Hosokawa, A.; Distefano, M. D. Selective Labeling of Proteins by Using Protein Farnesyltransferase. *ChemBioChem* **2007**, *8* (1), 98–105. https://doi.org/10.1002/cbic.200600340.

(162) Chen, X.; Muthoosamy, K.; Pfisterer, A.; Neumann, B.; Weil, T. Site-Selective Lysine Modification of Native Proteins and Peptides via Kinetically Controlled Labeling. *Bioconjug. Chem.* **2012**, *23* (3), 500–508. https://doi.org/10.1021/bc200556n.

(163) Asano, S.; Patterson, J. T.; Gaj, T.; Barbas, C. F. Site-Selective Labeling of a Lysine Residue in Human Serum Albumin. *Angew. Chem. Int. Ed.* **2014**, *53* (44), 11783–11786. https://doi.org/10.1002/anie.201405924.

(164) Purushottam, L.; Adusumalli, S. R.; Chilamari, M.; Rai, V. Chemoselective and Site-Selective Peptide and Native Protein Modification Enabled by Aldehyde Auto-Oxidation. *Chem. Commun.* **2017**, *53* (5), 959–962. https://doi.org/10.1039/C6CC09555K.

(165) Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Guerreiro, A.; Cal, P. M. S. D.; Bertoldo, J.; Maneiro, M.; Perkins, E.; Howard, J.; Deery, M. J.; Chalker, J. M.; Corzana, F.; Jiménez-Osés, G.; Bernardes, G. J. L. Chemo- and Regioselective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **2018**, *140* (11), 4004–4017. https://doi.org/10.1021/jacs.7b12874.

(166) Chilamari, M.; Purushottam, L.; Rai, V. Site-Selective Labeling of Native Proteins by a Multicomponent Approach. *Chem. - Eur. J.* **2017**, *23* (16), 3819–3823. https://doi.org/10.1002/chem.201605938.

(167) Chilamari, M.; Kalra, N.; Shukla, S.; Rai, V. Single-Site Labeling of Lysine in Proteins through a Metal-Free Multicomponent Approach. *Chem. Commun.* **2018**, *54* (53), 7302–7305. https://doi.org/10.1039/C8CC03311K.

(168) Forte, N.; Benni, I.; Karu, K.; Chudasama, V.; Baker, J. R. Cysteine-to-Lysine Transfer

Antibody Fragment Conjugation. *Chem. Sci.* **2019**, *10* (47), 10919–10924. https://doi.org/10.1039/C9SC03825F.

(169)   Adusumalli, S. R.; Rawale, D. G.; Thakur, K.; Purushottam, L.; Reddy, N. C.; Kalra, N.; Shukla, S.; Rai, V. Chemoselective and Site-Selective Lysine-Directed Lysine Modification Enables Single-Site Labeling of Native Proteins. *Angew. Chem. Int. Ed.* **2020**, anie.202000062. https://doi.org/10.1002/anie.202000062.

(170)   Zhang, Y.; Liang, Y.; Huang, F.; Zhang, Y.; Li, X.; Xia, J. Site-Selective Lysine Reactions Guided by Protein–Peptide Interaction. *Biochemistry* **2019**, *58* (7), 1010–1018. https://doi.org/10.1021/acs.biochem.8b01223.

(171)   Akkapeddi, P.; Azizi, S.-A.; Freedy, A. M.; Cal, P. M. S. D.; Gois, P. M. P.; Bernardes, G. J. L. Construction of Homogeneous Antibody–Drug Conjugates Using Site-Selective Protein Chemistry. *Chem. Sci.* **2016**, *7* (5), 2954–2963. https://doi.org/10.1039/C6SC00170J.

(172)   Shaunak, S.; Godwin, A.; Choi, J.-W.; Balan, S.; Pedone, E.; Vijayarangam, D.; Heidelberger, S.; Teo, I.; Zloh, M.; Brocchini, S. Site-Specific PEGylation of Native Disulfide Bonds in Therapeutic Proteins. *Nat. Chem. Biol.* **2006**, *2* (6), 312–313. https://doi.org/10.1038/nchembio786.

(173)   Balan, S.; Choi, J.; Godwin, A.; Teo, I.; Laborde, C. M.; Heidelberger, S.; Zloh, M.; Shaunak, S.; Brocchini, S. Site-Specific PEGylation of Protein Disulfide Bonds Using a Three-Carbon Bridge. *Bioconjug. Chem.* **2007**, *18* (1), 61–76. https://doi.org/10.1021/bc0601471.

(174)   Badescu, G.; Bryant, P.; Bird, M.; Henseleit, K.; Swierkosz, J.; Parekh, V.; Tommasi, R.; Pawlisz, E.; Jurlewicz, K.; Farys, M.; Camper, N.; Sheng, X.; Fisher, M.; Grygorash, R.; Kyle, A.; Abhilash, A.; Frigerio, M.; Edwards, J.; Godwin, A. Bridging Disulfides for Stable and Defined Antibody Drug Conjugates. *Bioconjug. Chem.* **2014**, *25* (6), 1124–1136. https://doi.org/10.1021/bc500148x.

(175)   Xu, L.; Raabe, M.; Zegota, M. M.; Nogueira, J. C. F.; Chudasama, V.; Kuan, S. L.; Weil, T. Site-Selective Protein Modification *via* Disulfide Rebridging for Fast Tetrazine/ *Trans* -Cyclooctene Bioconjugation. *Org. Biomol. Chem.* **2020**, *18* (6), 1140–1147. https://doi.org/10.1039/C9OB02687H.

(176)   Schumacher, F. F.; Nunes, J. P. M.; Maruani, A.; Chudasama, V.; Smith, M. E. B.; Chester, K. A.; Baker, J. R.; Caddick, S. Next Generation Maleimides Enable the Controlled Assembly of Antibody–Drug Conjugates *via* Native Disulfide Bond Bridging. *Org Biomol Chem* **2014**, *12* (37), 7261–7269. https://doi.org/10.1039/C4OB01550A.

(177)   Nunes, J. P. M.; Morais, M.; Vassileva, V.; Robinson, E.; Rajkumar, V. S.; Smith, M. E. B.; Pedley, R. B.; Caddick, S.; Baker, J. R.; Chudasama, V. Functional Native Disulfide Bridging Enables Delivery of a Potent, Stable and Targeted Antibody–Drug Conjugate (ADC). *Chem. Commun.* **2015**, *51* (53), 10624–10627. https://doi.org/10.1039/C5CC03557K.

(178)   Bryden, F.; Maruani, A.; Savoie, H.; Chudasama, V.; Smith, M. E. B.; Caddick, S.; Boyle, R. W. Regioselective and Stoichiometrically Controlled Conjugation of Photodynamic Sensitizers to a HER2 Targeting Antibody Fragment. *Bioconjug. Chem.* **2014**, *25* (3), 611–617. https://doi.org/10.1021/bc5000324.

(179)   Behrens, C. R.; Ha, E. H.; Chinn, L. L.; Bowers, S.; Probst, G.; Fitch-Bruhns, M.; Monteon, J.; Valdiosera, A.; Bermudez, A.; Liao-Chan, S.; Wong, T.; Melnick, J.; Theunissen, J.-W.; Flory, M. R.; Houser, D.; Venstrom, K.; Levashova, Z.; Sauer, P.; Migone, T.-S.; van der Horst, E. H.; Halcomb, R. L.; Jackson, D. Y. Antibody–Drug Conjugates (ADCs) Derived from Interchain Cysteine Cross-Linking Demonstrate Improved Homogeneity and Other Pharmacological Properties over Conventional Heterogeneous ADCs. *Mol. Pharm.* **2015**, *12* (11), 3986–3998. https://doi.org/10.1021/acs.molpharmaceut.5b00432.

(180)   Maruani, A.; Smith, M. E. B.; Miranda, E.; Chester, K. A.; Chudasama, V.; Caddick, S. A Plug-and-Play Approach to Antibody-Based Therapeutics via a Chemoselective Dual Click Strategy. *Nat. Commun.* **2015**, *6* (1), 6645. https://doi.org/10.1038/ncomms7645.

(181)   Robinson, E.; Nunes, J. P. M.; Vassileva, V.; Maruani, A.; Nogueira, J. C. F.; Smith, M. E. B.; Pedley, R. B.; Caddick, S.; Baker, J. R.; Chudasama, V. Pyridazinediones Deliver Potent,

Stable, Targeted and Efficacious Antibody–Drug Conjugates (ADCs) with a Controlled Loading of 4 Drugs per Antibody. *RSC Adv.* **2017**, *7* (15), 9073–9077. https://doi.org/10.1039/C7RA00788D.

(182)   Bahou, C.; Richards, D. A.; Maruani, A.; Love, E. A.; Javaid, F.; Caddick, S.; Baker, J. R.; Chudasama, V. Highly Homogeneous Antibody Modification through Optimisation of the Synthesis and Conjugation of Functionalised Dibromopyridazinediones. *Org. Biomol. Chem.* **2018**, *16* (8), 1359–1366. https://doi.org/10.1039/C7OB03138F.

(183)   Lee, M. T. W.; Maruani, A.; Baker, J. R.; Caddick, S.; Chudasama, V. Next-Generation Disulfide Stapling: Reduction and Functional Re-Bridging All in One. *Chem. Sci.* **2016**, *7* (1), 799–802. https://doi.org/10.1039/C5SC02666K.

(184)   Lee, M. T. W.; Maruani, A.; Richards, D. A.; Baker, J. R.; Caddick, S.; Chudasama, V. Enabling the Controlled Assembly of Antibody Conjugates with a Loading of Two Modules without Antibody Engineering. *Chem. Sci.* **2017**, *8* (3), 2056–2060. https://doi.org/10.1039/C6SC03655D.

(185)   Morais, M.; Nunes, J. P. M.; Karu, K.; Forte, N.; Benni, I.; Smith, M. E. B.; Caddick, S.; Chudasama, V.; Baker, J. R. Optimisation of the Dibromomaleimide (DBM) Platform for Native Antibody Conjugation by Accelerated Post-Conjugation Hydrolysis. *Org. Biomol. Chem.* **2017**, *15* (14), 2947–2952. https://doi.org/10.1039/C7OB00220C.

(186)   Maruani, A.; Szijj, P. A.; Bahou, C.; Nogueira, J. C. F.; Caddick, S.; Baker, J. R.; Chudasama, V. A Plug-and-Play Approach for the *De Novo* Generation of Dually Functionalized Bispecifics. *Bioconjug. Chem.* **2020**, *31* (3), 520–529. https://doi.org/10.1021/acs.bioconjchem.0c00002.

(187)   Martínez-Sáez, N.; Sun, S.; Oldrini, D.; Sormanni, P.; Boutureira, O.; Carboni, F.; Compañón, I.; Deery, M. J.; Vendruscolo, M.; Corzana, F.; Adamo, R.; Bernardes, G. J. L. Oxetane Grafts Installed Site-Selectively on Native Disulfides to Enhance Protein Stability and Activity In Vivo. *Angew. Chem. Int. Ed.* **2017**, *56* (47), 14963–14967. https://doi.org/10.1002/anie.201708847.

(188)   Lee, B.; Sun, S.; Jiménez-Moreno, E.; Neves, A. A.; Bernardes, G. J. L. Site-Selective Installation of an Electrophilic Handle on Proteins for Bioconjugation. *Bioorg. Med. Chem.* **2018**, *26* (11), 3060–3064. https://doi.org/10.1016/j.bmc.2018.02.028.

(189)   Griebenow, N.; Dilmaç, A. M.; Greven, S.; Bräse, S. Site-Specific Conjugation of Peptides and Proteins via Rebridging of Disulfide Bonds Using the Thiol–Yne Coupling Reaction. *Bioconjug. Chem.* **2016**, *27* (4), 911–917. https://doi.org/10.1021/acs.bioconjchem.5b00682.

(190)   Koniev, O.; Dovgan, I.; Renoux, B.; Ehkirch, A.; Eberova, J.; Cianférani, S.; Kolodych, S.; Papot, S.; Wagner, A. Reduction–Rebridging Strategy for the Preparation of ADPN-Based Antibody–Drug Conjugates. *MedChemComm* **2018**, *9* (5), 827–830. https://doi.org/10.1039/C8MD00141C.

(191)   Walsh, S. J.; Omarjee, S.; Galloway, W. R. J. D.; Kwan, T. T.-L.; Sore, H. F.; Parker, J. S.; Hyvönen, M.; Carroll, J. S.; Spring, D. R. A General Approach for the Site-Selective Modification of Native Proteins, Enabling the Generation of Stable and Functional Antibody–Drug Conjugates. *Chem. Sci.* **2019**, *10* (3), 694–700. https://doi.org/10.1039/C8SC04645J.

(192)   Walsh, S. J.; Iegre, J.; Seki, H.; Bargh, J. D.; Sore, H. F.; Parker, J. S.; Carroll, J. S.; Spring, D. R. General Dual Functionalisation of Biomacromolecules *via* a Cysteine Bridging Strategy. *Org. Biomol. Chem.* **2020**, 10.1039.D0OB00907E. https://doi.org/10.1039/D0OB00907E.

(193)   Counsell, A. J.; Walsh, S. J.; Robertson, N. S.; Sore, H. F.; Spring, D. R. Efficient and Selective Antibody Modification with Functionalised Divinyltriazines. *Org. Biomol. Chem.* **2020**, *18* (25), 4739–4743. https://doi.org/10.1039/D0OB01002B.

(194)   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* (32), 5307–5311. https://doi.org/10.1002/anie.200600368.

(195)   Witus, L. S.; Netirojjanakul, 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* (45), 17223–17229. https://doi.org/10.1021/ja408868a.

(196)   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* (5), 326–331. https://doi.org/10.1038/nchembio.1792.

(197)   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), 67. https://doi.org/10.1038/s42004-020-0309-y.

(198)   Raj, M.; Wu, H.; Blosser, S. L.; Vittoria, M. A.; Arora, P. S. Aldehyde Capture Ligation for Synthesis of Native Peptide Bonds. *J. Am. Chem. Soc.* **2015**, *137* (21), 6932–6940. https://doi.org/10.1021/jacs.5b03538.

(199)   Adusumalli, S. R.; Rawale, D. G.; Rai, V. Aldehydes Can Switch the Chemoselectivity of Electrophiles in Protein Labeling. *Org. Biomol. Chem.* **2018**, *16* (48), 9377–9381. https://doi.org/10.1039/C8OB02897D.

(200)   Chen, D.; Disotuar, M. M.; Xiong, X.; Wang, Y.; Chou, D. H.-C. Selective N-Terminal Functionalization of Native Peptides and Proteins. *Chem. Sci.* **2017**, *8* (4), 2717–2722. https://doi.org/10.1039/C6SC04744K.

(201)   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 α-Amino Groups of Peptides and Proteins Using Ketenes. *J. Am. Chem. Soc.* **2012**, *134* (5), 2589–2598. https://doi.org/10.1021/ja208009r.

(202)   Raj, M.; Tang, K. C. One Step Azolation Strategy for Site- and Chemo-selective Labeling of Proteins with Mass Sensitive Probes. *Angew. Chem. Int. Ed.* **2020**, anie.202007608. https://doi.org/10.1002/anie.202007608.

(203)   Obermeyer, A. C.; Jarman, J. B.; Francis, M. B. N-Terminal Modification of Proteins with *o* -Aminophenols. *J. Am. Chem. Soc.* **2014**, *136* (27), 9572–9579. https://doi.org/10.1021/ja500728c.

(204)   Schoffelen, S.; van Eldijk, M. B.; Rooijakkers, 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* (4), 701. https://doi.org/10.1039/c0sc00562b.

(205)   Huang, R.; Li, Z.; Ren, P.; Chen, W.; Kuang, Y.; Chen, J.; Zhan, Y.; Chen, H.; Jiang, B. *N* -Phenyl- *N* -Aceto-Vinylsulfonamides as Efficient and Chemoselective Handles for N-Terminal Modification of Peptides and Proteins: *N* -Phenyl- *N* -Aceto-Vinylsulfonamides as Efficient and Chemoselective Handles for N-Terminal Modification of Peptides and Proteins. *Eur. J. Org. Chem.* **2018**, *2018* (6), 829–836. https://doi.org/10.1002/ejoc.201701715.

(206)   Singudas, R.; Adusumalli, S. R.; Joshi, P. N.; Rai, V. A Phthalimidation Protocol That Follows Protein Defined Parameters. *Chem. Commun.* **2015**, *51* (3), 473–476. https://doi.org/10.1039/C4CC08503E.

(207)   Ren, H.; Xiao, F.; Zhan, K.; Kim, Y.-P.; Xie, H.; Xia, Z.; Rao, J. A Biocompatible Condensation Reaction for the Labeling of Terminal Cysteine Residues on Proteins. *Angew. Chem. Int. Ed.* **2009**, *48* (51), 9658–9662. https://doi.org/10.1002/anie.200903627.

(208)   Zheng, X.; Li, Z.; Gao, W.; Meng, X.; Li, X.; Luk, L. Y. P.; Zhao, Y.; Tsai, Y.-H.; Wu, C. Condensation of 2-((Alkylthio)(Aryl)Methylene)Malononitrile with 1,2-Aminothiol as a Novel Bioorthogonal Reaction for Site-Specific Protein Modification and Peptide Cyclization. *J. Am. Chem. Soc.* **2020**, *142* (11), 5097–5103. https://doi.org/10.1021/jacs.9b11875.

(209)   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 Functionalisation. *Chem. Sci.* **2016**, *7* (8), 5052–5058. https://doi.org/10.1039/C6SC01520D.

(210)   Bandyopadhyay, A.; Cambray, S.; Gao, J. Fast and Selective Labeling of N-Terminal

Cysteines at Neutral PH via Thiazolidino Boronate Formation. *Chem. Sci.* **2016**, *7* (7), 4589–4593. https://doi.org/10.1039/C6SC00172F.

(211)   Gao, J.; Li, K.; Wang, W. Fast and Stable N-Terminal Cysteine Modification via Thiazolidino Boronate Mediated Acyl Transfer. *Angew. Chem. Int. Ed.* **2020**, anie.202000837. https://doi.org/10.1002/anie.202000837.

(212)   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), 2539. https://doi.org/10.1038/s41467-019-10503-7.

(213)   Sim, Y. E.; Nwajiobi, O.; Mahesh, S.; Cohen, R. D.; Reibarkh, M. Y.; Raj, M. Secondary Amine Selective Petasis (SASP) Bioconjugation. *Chem. Sci.* **2020**, *11* (1), 53–61. https://doi.org/10.1039/C9SC04697F.

(214)   Maza, J. C.; Bader, D. L. V.; Xiao, L.; Marmelstein, A. M.; Brauer, D. D.; ElSohly, A. M.; Smith, M. J.; Krska, S. W.; Parish, C. A.; Francis, M. B. Enzymatic Modification of N-Terminal Proline Residues Using Phenol Derivatives. *J. Am. Chem. Soc.* **2019**, *141* (9), 3885–3892. https://doi.org/10.1021/jacs.8b10845.

(215)   Ban, H.; Gavrilyuk, J.; Barbas, C. F. Tyrosine Bioconjugation through Aqueous Ene-Type Reactions: A Click-Like Reaction for Tyrosine. *J. Am. Chem. Soc.* **2010**, *132* (5), 1523–1525. https://doi.org/10.1021/ja909062q.

(216)   Gavrilyuk, J.; Ban, H.; Nagano, M.; Hakamata, W.; Barbas, C. F. Formylbenzene Diazonium Hexafluorophosphate Reagent for Tyrosine-Selective Modification of Proteins and the Introduction of a Bioorthogonal Aldehyde. *Bioconjug. Chem.* **2012**, *23* (12), 2321–2328. https://doi.org/10.1021/bc300410p.

(217)   Allan, C.; Kosar, M.; Burr, C. V.; Mackay, C. L.; Duncan, R. R.; Hulme, A. N. A Catch-and-Release Approach to Selective Modification of Accessible Tyrosine Residues. *ChemBioChem* **2018**, *19* (23), 2443–2447. https://doi.org/10.1002/cbic.201800532.

(218)   Joshi, N. S.; Whitaker, L. R.; Francis, M. B. A Three-Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation. *J. Am. Chem. Soc.* **2004**, *126* (49), 15942–15943. https://doi.org/10.1021/ja0439017.

(219)   Guo, H.-M.; Minakawa, M.; Ueno, L.; Tanaka, F. Synthesis and Evaluation of a Cyclic Imine Derivative Conjugated to a Fluorescent Molecule for Labeling of Proteins. *Bioorg. Med. Chem. Lett.* **2009**, *19* (4), 1210–1213. https://doi.org/10.1016/j.bmcl.2008.12.071.

(220)   Ohata, J.; Miller, M. K.; Mountain, C. M.; Vohidov, F.; Ball, Z. T. A Three-Component Organometallic Tyrosine Bioconjugation. *Angew. Chem. Int. Ed.* **2018**, *57* (11), 2827–2830. https://doi.org/10.1002/anie.201711868.

(221)   Tilley, S. D.; Francis, M. B. Tyrosine-Selective Protein Alkylation Using π-Allylpalladium Complexes. *J. Am. Chem. Soc.* **2006**, *128* (4), 1080–1081. https://doi.org/10.1021/ja057106k.

(222)   Minamihata, K.; Goto, M.; Kamiya, N. Site-Specific Protein Cross-Linking by Peroxidase-Catalyzed Activation of a Tyrosine-Containing Peptide Tag. *Bioconjug. Chem.* **2011**, *22* (1), 74–81. https://doi.org/10.1021/bc1003982.

(223)   Sato, S.; Matsumura, M.; Kadonosono, T.; Abe, S.; Ueno, T.; Ueda, H.; Nakamura, H. Site-Selective Protein Chemical Modification of Exposed Tyrosine Residues Using Tyrosine Click Reaction. *Bioconjug. Chem.* **2020**, *31* (5), 1417–1424. https://doi.org/10.1021/acs.bioconjchem.0c00120.

(224)   Antos, J. M.; Francis, M. B. Selective Tryptophan Modification with Rhodium Carbenoids in Aqueous Solution. *J. Am. Chem. Soc.* **2004**, *126* (33), 10256–10257. https://doi.org/10.1021/ja047272c.

(225)   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* (17), 6301–6308. https://doi.org/10.1021/ja900094h.

(226)   Hansen, M. B.; Hubálek, F.; Skrydstrup, T.; Hoeg-Jensen, T. Chemo- and Regioselective Ethynylation of Tryptophan-Containing Peptides and Proteins. *Chem. - Eur. J.* **2016**, *22* (5), 1572–1576. https://doi.org/10.1002/chem.201504462.

(227) Seki, Y.; Ishiyama, T.; Sasaki, D.; Abe, J.; Sohma, Y.; Oisaki, K.; Kanai, M. Transition Metal-Free Tryptophan-Selective Bioconjugation of Proteins. *J. Am. Chem. Soc.* **2016**, *138* (34), 10798–10801. https://doi.org/10.1021/jacs.6b06692.

(228) Imiołek, M.; Karunanithy, G.; Ng, W.-L.; Baldwin, A. J.; Gouverneur, V.; Davis, B. G. Selective Radical Trifluoromethylation of Native Residues in Proteins. *J. Am. Chem. Soc.* **2018**, *140* (5), 1568–1571. https://doi.org/10.1021/jacs.7b10230.

(229) Tower, S. J.; Hetcher, W. J.; Myers, T. E.; Kuehl, N. J.; Taylor, M. T. Selective Modification of Tryptophan Residues in Peptides and Proteins Using a Biomimetic Electron Transfer Process. *J. Am. Chem. Soc.* **2020**, *142* (20), 9112–9118. https://doi.org/10.1021/jacs.0c03039.

(230) 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* (8), 1100–1103. https://doi.org/10.1039/C8CC08733D.

(231) Adusumalli, S. R.; Rawale, D. G.; Singh, U.; Tripathi, P.; Paul, R.; Kalra, N.; Mishra, R. K.; Shukla, S.; Rai, V. Single-Site Labeling of Native Proteins Enabled by a Chemoselective and Site-Selective Chemical Technology. *J. Am. Chem. Soc.* **2018**, *140* (44), 15114–15123. https://doi.org/10.1021/jacs.8b10490.

(232) Stipanuk, M. H. SULFUR AMINO ACID METABOLISM: Pathways for Production and Removal of Homocysteine and Cysteine. *Annu. Rev. Nutr.* **2004**, *24* (1), 539–577. https://doi.org/10.1146/annurev.nutr.24.012003.132418.

(233) 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* (7728), 563–568. https://doi.org/10.1038/s41586-018-0608-y.

(234) Gauthier, M. A.; Ayer, M.; Kowal, J.; Wurm, F. R.; Klok, H.-A. Arginine-Specific Protein Modification Using α-Oxo-Aldehyde Functional Polymers Prepared by Atom Transfer Radical Polymerization. *Polym. Chem.* **2011**, *2* (7), 1490. https://doi.org/10.1039/c0py00422g.

(235) Gauthier, M. A.; Klok, H.-A. Arginine-Specific Modification of Proteins with Polyethylene Glycol. *Biomacromolecules* **2011**, *12* (2), 482–493. https://doi.org/10.1021/bm101272g.

(236) Vantourout, J. C.; Adusumalli, S. R.; Knouse, K. W.; Flood, D. T.; Ramirez, A.; M. Padial, N.; Istrate, A.; Maziarz, K.; deGruyter, J. N.; Merchant, R. R.; Qiao, J. X.; Schmidt, M. A.; Deery, M. J.; Eastgate, M. D.; Dawson, P. E.; Bernardes, G. J. L.; Baran, P. S. Serine-Selective Bioconjugation. *J. Am. Chem. Soc.* **2020**, jacs.0c05595. https://doi.org/10.1021/jacs.0c05595.

(237) Totaro, K. A.; Liao, X.; Bhattacharya, K.; Finneman, J. I.; Sperry, J. B.; Massa, M. A.; Thorn, J.; Ho, S. V.; Pentelute, B. L. Systematic Investigation of EDC/SNHS-Mediated Bioconjugation Reactions for Carboxylated Peptide Substrates. *Bioconjug. Chem.* **2016**, *27* (4), 994–1004. https://doi.org/10.1021/acs.bioconjchem.6b00043.

(238) McGrath, N. A.; Andersen, K. A.; Davis, A. K. F.; Lomax, J. E.; Raines, R. T. Diazo Compounds for the Bioreversible Esterification of Proteins. *Chem. Sci.* **2015**, *6* (1), 752–755. https://doi.org/10.1039/C4SC01768D.

(239) Bloom, S.; Liu, C.; Kölmel, D. K.; Qiao, J. X.; Zhang, Y.; Poss, M. A.; Ewing, W. R.; MacMillan, D. W. C. Decarboxylative Alkylation for Site-Selective Bioconjugation of Native Proteins via Oxidation Potentials. *Nat. Chem.* **2018**, *10* (2), 205–211. https://doi.org/10.1038/nchem.2888.

(240) Forte, N.; Chudasama, V.; Baker, J. R. Homogeneous Antibody-Drug Conjugates via Site-Selective Disulfide Bridging. *Drug Discov. Today Technol.* **2018**, *30*, 11–20. https://doi.org/10.1016/j.ddtec.2018.09.004.

(241) Wirthlin, T.; Kita, Y.; Koser, G. F.; Ochiai, M. *Hypervalent Iodine Chemistry: Modern Developments in Organic Synthesis.*; Springer: New York, 2003.

(242) Stang, P. J.; Zhdankin, V. V. Organic Polyvalent Iodine Compounds. *Chem. Rev.* **1996**, *96* (3), 1123–1178. https://doi.org/10.1021/cr940424+.

(243) Perkins, C. W.; Martin, J. C.; Arduengo, A. J.; Lau, W.; Alegria, A.; Kochi, J. K. An

Electrically Neutral .Sigma.-Sulfuranyl Radical from the Homolysis of a Perester with Neighboring Sulfenyl Sulfur: 9-S-3 Species. *J. Am. Chem. Soc.* **1980**, *102* (26), 7753–7759. https://doi.org/10.1021/ja00546a019.

(244) Zhdankin, V. V. *Hypervalent Iodine Chemistry: Preparation, Structure, and Synthetic Applications of Polyvalent Iodine Compounds*; John Wiley & Sons, Inc: Chichester, West Sussex, 2014.

(245) Musher, J. I. The Chemistry of Hypervalent Molecules. *Angew. Chem. Int. Ed. Engl.* **1969**, *8* (1), 54–68. https://doi.org/10.1002/anie.196900541.

(246) Pimentel, G. C. The Bonding of Trihalide and Bifluoride Ions by the Molecular Orbital Method. *J. Chem. Phys.* **1951**, *19* (4), 446–448. https://doi.org/10.1063/1.1748245.

(247) Rundle, R. E. The Implications of Some Recent Structures for Chemical Valence Theory. In *Survey of Progress in Chemistry*; Elsevier, 1963; Vol. 1, pp 81–131. https://doi.org/10.1016/B978-1-4832-0003-3.50008-9.

(248) Eisenberger, P.; Gischig, S.; Togni, A. Novel 10-I-3 Hypervalent Iodine-Based Compounds for Electrophilic Trifluoromethylation. *Chem. - Eur. J.* **2006**, *12* (9), 2579–2586. https://doi.org/10.1002/chem.200501052.

(249) Ochiai, M.; Masaki, Y.; Shiro, M. Synthesis and Structure of 1-Alkynyl-1,2-Benziodoxol-3(1H)-Ones. *J. Org. Chem.* **1991**, *56* (19), 5511–5513. https://doi.org/10.1021/jo00019a007.

(250) Zhdankin, V. V.; Kuehl, C. J.; Krasutsky, A. P.; Bolz, J. T.; Simonsen, A. J. 1-(Organosulfonyloxy)-3(1 *H*)-1,2-Benziodoxoles: Preparation and Reactions with Alkynyltrimethylsilanes. *J. Org. Chem.* **1996**, *61* (19), 6547–6551. https://doi.org/10.1021/jo960927a.

(251) Zhdankin, V. V.; Krasutsky, A. P.; Kuehl, C. J.; Simonsen, A. J.; Woodward, J. K.; Mismash, B.; Bolz, J. T. Preparation, X-Ray Crystal Structure, and Chemistry of Stable AzidoiodinanesDerivatives of Benziodoxole. *J. Am. Chem. Soc.* **1996**, *118* (22), 5192–5197. https://doi.org/10.1021/ja954119x.

(252) Zhdankin, V. V.; Kuehl, C. J.; Krasutsky, A. P.; Bolz, J. T.; Mismash, B.; Woodward, J. K.; Simonsen, A. J. 1-Cyano-3-(1H)-1,2-Benziodoxols: Stable Cyanoiodinanes and Efficient Reagents for Direct N-Alkyl Cyanation of N,N-Dialkylarylamines. *Tetrahedron Lett.* **1995**, *36* (44), 7975–7978. https://doi.org/10.1016/0040-4039(95)01720-3.

(253) Kieltsch, I.; Eisenberger, P.; Togni, A. Mild Electrophilic Trifluoromethylation of Carbon- and Sulfur-Centered Nucleophiles by a Hypervalent Iodine(III)–CF3 Reagent. *Angew. Chem. Int. Ed.* **2007**, *46* (5), 754–757. https://doi.org/10.1002/anie.200603497.

(254) Sala, O.; Santschi, N.; Jungen, S.; Lüthi, H. P.; Iannuzzi, M.; Hauser, N.; Togni, A. *S* -Trifluoromethylation of Thiols by Hypervalent Iodine Reagents: A Joint Experimental and Computational Study. *Chem. - Eur. J.* **2016**, *22* (5), 1704–1713. https://doi.org/10.1002/chem.201503774.

(255) Matoušek, V.; Václavík, J.; Hájek, P.; Charpentier, J.; Blastik, Z. E.; Pietrasiak, E.; Budinská, A.; Togni, A.; Beier, P. Expanding the Scope of Hypervalent Iodine Reagents for Perfluoroalkylation: From Trifluoromethyl to Functionalized Perfluoroethyl. *Chem. - Eur. J.* **2016**, *22* (1), 417–424. https://doi.org/10.1002/chem.201503531.

(256) Charpentier, J.; Früh, N.; Togni, A. Electrophilic Trifluoromethylation by Use of Hypervalent Iodine Reagents. *Chem. Rev.* **2015**, *115* (2), 650–682. https://doi.org/10.1021/cr500223h.

(257) Klimánková, I.; Hubálek, M.; Matoušek, V.; Beier, P. Synthesis of Water-Soluble Hypervalent Iodine Reagents for Fluoroalkylation of Biological Thiols. *Org. Biomol. Chem.* **2019**, *17* (47), 10097–10102. https://doi.org/10.1039/c9ob02115a.

(258) Charkoudian, L. K.; Liu, C. W.; Capone, S.; Kapur, S.; Cane, D. E.; Togni, A.; Seebach, D.; Khosla, C. Probing the Interactions of an Acyl Carrier Protein Domain from the 6-Deoxyerythronolide B Synthase. *Protein Sci. Publ. Protein Soc.* **2011**, *20* (7), 1244–1255. https://doi.org/10.1002/pro.652.

(259) Capone, S.; Kieltsch, I.; Flögel, O.; Lelais, G.; Togni, A.; Seebach, D. Electrophilic *S* -

Trifluoromethylation of Cysteine Side Chains in *α* - and *β* -Peptides: Isolation of Trifluoro-Methylated *Sandostatin* ® (Octreotide) Derivatives. *Helv. Chim. Acta* **2008**, *91* (11), 2035–2056. https://doi.org/10.1002/hlca.200890217.

(260) Frei, R.; Waser, J. A Highly Chemoselective and Practical Alkynylation of Thiols. *J. Am. Chem. Soc.* **2013**, *135* (26), 9620–9623. https://doi.org/10.1021/ja4044196.

(261) Frei, R.; Wodrich, M. D.; Hari, D. P.; Borin, P.-A.; Chauvier, C.; Waser, J. Fast and Highly Chemoselective Alkynylation of Thiols with Hypervalent Iodine Reagents Enabled through a Low Energy Barrier Concerted Mechanism. *J. Am. Chem. Soc.* **2014**, *136* (47), 16563–16573. https://doi.org/10.1021/ja5083014.

(262) Hari, D. P.; Caramenti, P.; Waser, J. Cyclic Hypervalent Iodine Reagents: Enabling Tools for Bond Disconnection via Reactivity Umpolung. *Acc. Chem. Res.* **2018**, *51* (12), 3212–3225. https://doi.org/10.1021/acs.accounts.8b00468.

(263) Brand, J.; Charpentier, J.; Waser, J. Direct Alkynylation of Indole and Pyrrole Heterocycles. *Angew. Chem. Int. Ed.* **2009**, *48* (49), 9346–9349. https://doi.org/10.1002/anie.200905419.

(264) Tolnai, G. L.; Brand, J. P.; Waser, J. Gold-Catalyzed Direct Alkynylation of Tryptophan in Peptides Using TIPS-EBX. *Beilstein J. Org. Chem.* **2016**, *12*, 745–749. https://doi.org/10.3762/bjoc.12.74.

(265) Tian, J.; Gao, W.-C.; Zhou, D.-M.; Zhang, C. Recyclable Hypervalent Iodine(III) Reagent Iodosodilactone as an Efficient Coupling Reagent for Direct Esterification, Amidation, and Peptide Coupling. *Org. Lett.* **2012**, *14* (12), 3020–3023. https://doi.org/10.1021/ol301085v.

(266) Zhang, C.; Liu, S.-S.; Sun, B.; Tian, J. Practical Peptide Synthesis Mediated by a Recyclable Hypervalent Iodine Reagent and Tris(4-Methoxyphenyl)Phosphine. *Org. Lett.* **2015**, *17* (16), 4106–4109. https://doi.org/10.1021/acs.orglett.5b02045.

(267) Qiu, L.-J.; Liu, D.; Zheng, K.; Zhang, M.-T.; Zhang, C. A Benziodoxole-Based Hypervalent Iodine(III) Compound Functioning as a Peptide Coupling Reagent. *Front. Chem.* **2020**, *8*, 183. https://doi.org/10.3389/fchem.2020.00183.

(268) Le Vaillant, F.; Courant, T.; Waser, J. Room-Temperature Decarboxylative Alkynylation of Carboxylic Acids Using Photoredox Catalysis and EBX Reagents. *Angew. Chem. Int. Ed.* **2015**, *54* (38), 11200–11204. https://doi.org/10.1002/anie.201505111.

(269) Garreau, M.; Le Vaillant, F.; Waser, J. C-Terminal Bioconjugation of Peptides through Photoredox Catalyzed Decarboxylative Alkynylation. *Angew. Chem. Int. Ed.* **2019**, *58* (24), 8182–8186. https://doi.org/10.1002/anie.201901922.

(270) Keyser, S. G. L.; Utz, A.; Bertozzi, C. R. Computation-Guided Rational Design of a Peptide Motif That Reacts with Cyanobenzothiazoles via Internal Cysteine–Lysine Relay. *J. Org. Chem.* **2018**, *83* (14), 7467–7479. https://doi.org/10.1021/acs.joc.8b00625.

(271) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266* (5186), 776–779. https://doi.org/10.1126/science.7973629.

(272) Burke, H. M.; McSweeney, L.; Scanlan, E. M. Exploring Chemoselective S-to-N Acyl Transfer Reactions in Synthesis and Chemical Biology. *Nat. Commun.* **2017**, *8* (1), 15655. https://doi.org/10.1038/ncomms15655.

(273) Johnson, E. C. B.; Kent, S. B. H. Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction. *J. Am. Chem. Soc.* **2006**, *128* (20), 6640–6646. https://doi.org/10.1021/ja058344i.

(274) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **2019**, *119* (12), 7328–7443. https://doi.org/10.1021/acs.chemrev.8b00712.

(275) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives. *J. Am. Chem. Soc.* **1997**, *119* (19), 4325–4329. https://doi.org/10.1021/ja962656r.

(276) Yan, L. Z.; Dawson, P. E. Synthesis of Peptides and Proteins without Cysteine

Residues by Native Chemical Ligation Combined with Desulfurization. *J. Am. Chem. Soc.* **2001**, *123* (4), 526–533. https://doi.org/10.1021/ja003265m.

(277) Crich, D.; Banerjee, A. Native Chemical Ligation at Phenylalanine. *J. Am. Chem. Soc.* **2007**, *129* (33), 10064–10065. https://doi.org/10.1021/ja072804l.

(278) Chen, J.; Wan, Q.; Yuan, Y.; Zhu, J.; Danishefsky, S. J. Native Chemical Ligation at Valine: A Contribution to Peptide and Glycopeptide Synthesis. *Angew. Chem. Int. Ed.* **2008**, *47* (44), 8521–8524. https://doi.org/10.1002/anie.200803523.

(279) Dawson, P. E. Native Chemical Ligation Combined with Desulfurization and Deselenization: A General Strategy for Chemical Protein Synthesis. *Isr. J. Chem.* **2011**, *51* (8–9), 862–867. https://doi.org/10.1002/ijch.201100128.

(280) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. Protein Synthesis by Native Chemical Ligation: Expanded Scope by Using Straightforward Methodology. *Proc. Natl. Acad. Sci.* **1999**, *96* (18), 10068–10073. https://doi.org/10.1073/pnas.96.18.10068.

(281) Torbeev, V. Yu.; Kent, S. B. H. Convergent Chemical Synthesis and Crystal Structure of a 203 Amino Acid "Covalent Dimer" HIV-1 Protease Enzyme Molecule. *Angew. Chem. Int. Ed.* **2007**, *46* (10), 1667–1670. https://doi.org/10.1002/anie.200604087.

(282) Dömling, A.; Ugi, I. K. Multicomponent Reactions with Isocyanides. *Angew. Chem. Int. Ed Engl.* **2000**, *39* (18), 3168–3210. https://doi.org/10.1002/1521-3773(20000915)39:18<3168::AID-ANIE3168>3.0.CO;2-U.

(283) *Multicomponent Reactions*, 1st ed.; Zhu, J., Bienaymé, H., Eds.; Wiley, 2005. https://doi.org/10.1002/3527605118.

(284) Strecker, A. Ueber die künstliche Bildung der Milchsäure und einen neuen, dem Glycocoll homologen Körper; *Ann. Chem. Pharm.* **1850**, *75* (1), 27–45. https://doi.org/10.1002/jlac.18500750103.

(285) Laurent, A.; Gerhardt, C. F. Sur Diverses Combinaisons Azotées Du Benzoïle. *Ann Chim Phys* **1838**, *66*, 181–195.

(286) Hantzsch, A. Ueber die Synthese pyridinartiger Verbindungen aus Acetessigäther und Aldehydammoniak. *Justus Liebigs Ann. Chem.* **1882**, *215* (1), 1–82. https://doi.org/10.1002/jlac.18822150102.

(287) Bossert, F.; Meyer, H.; Wehinger, E. 4-Aryldihydropyridines, a New Class of Highly Active Calcium Antagonists. *Angew. Chem. Int. Ed. Engl.* **1981**, *20* (9), 762–769. https://doi.org/10.1002/anie.198107621.

(288) Hantzsch, A. Neue Bildungsweise von Pyrrolderivaten. *Berichte Dtsch. Chem. Ges.* **1890**, *23* (1), 1474–1476. https://doi.org/10.1002/cber.189002301243.

(289) Biginelli, P. Ueber Aldehyduramide des Acetessigäthers. *Berichte Dtsch. Chem. Ges.* **1891**, *24* (1), 1317–1319. https://doi.org/10.1002/cber.189102401228.

(290) Biginelli, P. Ueber Aldehyduramide des Acetessigäthers. II. *Berichte Dtsch. Chem. Ges.* **1891**, *24* (2), 2962–2967. https://doi.org/10.1002/cber.189102402126.

(291) Mannich, C.; Krösche, W. Ueber ein Kondensationsprodukt aus Formaldehyd, Ammoniak und Antipyrin. *Arch. Pharm. (Weinheim)* **1912**, *250* (1), 647–667. https://doi.org/10.1002/ardp.19122500151.

(292) Robinson, R. LXXV.—A Theory of the Mechanism of the Phytochemical Synthesis of Certain Alkaloids. *J Chem Soc Trans* **1917**, *111* (0), 876–899. https://doi.org/10.1039/CT9171100876.

(293) Passerini, M.; Simone, L. Sopra Gli Isonitrili (I). Composto Del p-Isonitril-Azobenzolo Con Acetone Ed Acido Acetico. *Gazz Chim Ital* **1921**, *51* (II), 126–129.

(294) Maeda, S.; Komagawa, S.; Uchiyama, M.; Morokuma, K. Finding Reaction Pathways for Multicomponent Reactions: The Passerini Reaction Is a Four-Component Reaction. *Angew. Chem. Int. Ed.* **2011**, *50* (3), 644–649. https://doi.org/10.1002/anie.201005336.

(295) Ugi, I.; Meyr, R.; Fetzer, U.; Steinbrückner, C. Versuche mit Isonitrilen. *Angew. Chem.* **1959**, *71* (11), 386. https://doi.org/10.1002/ange.19590711110.

(296) Chéron, N.; Ramozzi, R.; Kaïm, L. E.; Grimaud, L.; Fleurat-Lessard, P. Challenging 50

Years of Established Views on Ugi Reaction: A Theoretical Approach. *J. Org. Chem.* **2012**, *77* (3), 1361–1366. https://doi.org/10.1021/jo2021554.

(297) Romanini, D. W.; Francis, M. B. Attachment of Peptide Building Blocks to Proteins Through Tyrosine Bioconjugation. *Bioconjug. Chem.* **2008**, *19* (1), 153–157. https://doi.org/10.1021/bc700231v.

(298) Feng, Q.; Xu, Y.; Zhou, Y.; Lu, L.; Chen, F.; Wang, X. Preparation of Dichlorvos–Protein Complete Antigen by Mannich-Type Reaction. *J. Mol. Struct.* **2010**, *977* (1–3), 100–105. https://doi.org/10.1016/j.molstruc.2010.05.020.

(299) Cristau, P.; Vors, J.-P.; Zhu, J. A Rapid Access to Biaryl Ether Containing Macrocycles by Pairwise Use of Ugi 4CR and Intramolecular S $_N$ Ar-Based Cycloetherification. *Org. Lett.* **2001**, *3* (25), 4079–4082. https://doi.org/10.1021/ol0168420.

(300) Cristau, P.; Vors, J.-P.; Zhu, J. Solid-Phase Synthesis of Natural Product-like Macrocycles by a Sequence of Ugi-4CR and SNAr-Based Cycloetherification. *Tetrahedron Lett.* **2003**, *44* (30), 5575–5578. https://doi.org/10.1016/S0040-4039(03)01378-9.

(301) Kazmaier, U.; Hebach, C.; Watzke, A.; Maier, S.; Mues, H.; Huch, V. A Straightforward Approach towards Cyclic Peptides via Ring-Closing Metathesis—Scope and Limitations. *Org Biomol Chem* **2005**, *3* (1), 136–145. https://doi.org/10.1039/B411228H.

(302) Wessjohann, L. A.; Voigt, B.; Rivera, D. G. Diversity Oriented One-Pot Synthesis of Complex Macrocycles: Very Large Steroid-Peptoid Hybrids from Multiple Multicomponent Reactions Including Bifunctional Building Blocks. *Angew. Chem. Int. Ed.* **2005**, *44* (30), 4785–4790. https://doi.org/10.1002/anie.200500019.

(303) Hartweg, M.; Edwards-Gayle, C. J. C.; Radvar, E.; Collis, D.; Reza, M.; Kaupp, M.; Steinkoenig, J.; Ruokolainen, J.; Rambo, R.; Barner-Kowollik, C.; Hamley, I. W.; Azevedo, H. S.; Becer, C. R. Ugi Multicomponent Reaction to Prepare Peptide–Peptoid Hybrid Structures with Diverse Chemical Functionalities. *Polym. Chem.* **2018**, *9* (4), 482–489. https://doi.org/10.1039/C7PY01953J.

(304) Vasco, A. V.; Méndez, Y.; Porzel, A.; Balbach, J.; Wessjohann, L. A.; Rivera, D. G. A Multicomponent Stapling Approach to Exocyclic Functionalized Helical Peptides: Adding Lipids, Sugars, PEGs, Labels, and Handles to the Lactam Bridge. *Bioconjug. Chem.* **2019**, *30* (1), 253–259. https://doi.org/10.1021/acs.bioconjchem.8b00906.

(305) Reguera, L.; Rivera, D. G. Multicomponent Reaction Toolbox for Peptide Macrocyclization and Stapling. *Chem. Rev.* **2019**, *119* (17), 9836–9860. https://doi.org/10.1021/acs.chemrev.8b00744.

(306) Vasco, A. V.; Pérez, C. S.; Morales, F. E.; Garay, H. E.; Vasilev, D.; Gavín, J. A.; Wessjohann, L. A.; Rivera, D. G. Macrocyclization of Peptide Side Chains by the Ugi Reaction: Achieving Peptide Folding and Exocyclic *N* -Functionalization in One Shot. *J. Org. Chem.* **2015**, *80* (13), 6697–6707. https://doi.org/10.1021/acs.joc.5b00858.

(307) Wessjohann, L. A.; Morejón, M. C.; Ojeda, G. M.; Rhoden, C. R. B.; Rivera, D. G. Applications of Convertible Isonitriles in the Ligation and Macrocyclization of Multicomponent Reaction-Derived Peptides and Depsipeptides. *J. Org. Chem.* **2016**, *81* (15), 6535–6545. https://doi.org/10.1021/acs.joc.6b01150.

(308) Morejón, M. C.; Laub, A.; Westermann, B.; Rivera, D. G.; Wessjohann, L. A. Solution- and Solid-Phase Macrocyclization of Peptides by the Ugi–Smiles Multicomponent Reaction: Synthesis of *N* -Aryl-Bridged Cyclic Lipopeptides. *Org. Lett.* **2016**, *18* (16), 4096–4099. https://doi.org/10.1021/acs.orglett.6b02001.

(309) Puentes, A. R.; Morejón, M. C.; Rivera, D. G.; Wessjohann, L. A. Peptide Macrocyclization Assisted by Traceless Turn Inducers Derived from Ugi Peptide Ligation with Cleavable and Resin-Linked Amines. *Org. Lett.* **2017**, *19* (15), 4022–4025. https://doi.org/10.1021/acs.orglett.7b01761.

(310) Ricardo, M. G.; Vasco, A. V.; Rivera, D. G.; Wessjohann, L. A. Stabilization of Cyclic β-Hairpins by Ugi-Reaction-Derived *N* -Alkylated Peptides: The Quest for Functionalized β-Turns. *Org. Lett.* **2019**, *21* (18), 7307–7310. https://doi.org/10.1021/acs.orglett.9b02592.

(311) Abdelraheem, E.; Khaksar, S.; Dömling, A. Concise Synthesis of Macrocycles by Multicomponent Reactions. *Synthesis* **2018**, *50* (05), 1027–1038. https://doi.org/10.1055/s-0036-1590946.

(312) Goldstein, L.; Freeman, A.; Sokolovsky, M. Chemically Modified Nylons as Supports for Enzyme Immobilization. Polyisonitrile-Nylon. *Biochem. J.* **1974**, *143* (3), 497–509. https://doi.org/10.1042/bj1430497a.

(313) Blassberger, D.; Freeman, A.; Goldstein, L. Chemically Modified Polyesters as Supports for Enzyme Immobilization: Isocyanide, Acylhydrazide, and Aminoaryl Derivatives of Poly(Ethylene Terephthalate). *Biotechnol. Bioeng.* **1978**, *20* (2), 309–316. https://doi.org/10.1002/bit.260200216.

(314) Camacho, C.; Matías, J. C.; García, D.; Simpson, B. K.; Villalonga, R. Amperometric Enzyme Biosensor for Hydrogen Peroxide via Ugi Multicomponent Reaction. *Electrochem. Commun.* **2007**, *9* (7), 1655–1660. https://doi.org/10.1016/j.elecom.2007.03.013.

(315) García, A.; Hernández, K.; Chico, B.; García, D.; Villalonga, M. L.; Villalonga, R. Preparation of Thermostable Trypsin–Polysaccharide Neoglycoenzymes through Ugi Multicomponent Reaction. *J. Mol. Catal. B Enzym.* **2009**, *59* (1–3), 126–130. https://doi.org/10.1016/j.molcatb.2009.02.001.

(316) Mohammadi, M.; Ashjari, M.; Dezvarei, S.; Yousefi, M.; Babaki, M.; Mohammadi, J. Rapid and High-Density Covalent Immobilization of Rhizomucor Miehei Lipase Using a Multi Component Reaction: Application in Biodiesel Production. *RSC Adv.* **2015**, *5* (41), 32698–32705. https://doi.org/10.1039/C5RA03299G.

(317) Mohammadi, M.; Habibi, Z.; Gandomkar, S.; Yousefi, M. A Novel Approach for Bioconjugation of Rhizomucor Miehei Lipase (RML) onto Amine-Functionalized Supports; Application for Enantioselective Resolution of Rac-Ibuprofen. *Int. J. Biol. Macromol.* **2018**, *117*, 523–531. https://doi.org/10.1016/j.ijbiomac.2018.05.218.

(318) Mohammadi, M.; Ashjari, M.; Garmroodi, M.; Yousefi, M.; Karkhane, A. A. The Use of Isocyanide-Based Multicomponent Reaction for Covalent Immobilization of Rhizomucor Miehei Lipase on Multiwall Carbon Nanotubes and Graphene Nanosheets. *RSC Adv.* **2016**, *6* (76), 72275–72285. https://doi.org/10.1039/C6RA14142K.

(319) Ziegler, T.; Gerling, S.; Lang, M. Preparation of Bioconjugates through an Ugi Reaction. *Angew. Chem. Int. Ed Engl.* **2000**, *39* (12), 2109–2112. https://doi.org/10.1002/1521-3773(20000616)39:12<2109::AID-ANIE2109>3.0.CO;2-9.

(320) Méndez, Y.; Chang, J.; Humpierre, A. R.; Zanuy, A.; Garrido, R.; Vasco, A. V.; Pedroso, J.; Santana, D.; Rodríguez, L. M.; García-Rivera, D.; Valdés, Y.; Vérez-Bencomo, V.; Rivera, D. G. Multicomponent Polysaccharide–Protein Bioconjugation in the Development of Antibacterial Glycoconjugate Vaccine Candidates. *Chem. Sci.* **2018**, *9* (9), 2581–2588. https://doi.org/10.1039/C7SC05467J.

(321) Humpierre, A. R.; Zanuy, A.; Saenz, M.; Garrido, R.; Vasco, A. V.; Perez-Nicado, R.; Soroa-Milán, Y.; Santana-Mederos, D.; Westermann, B.; Vérez-Bencomo, V.; Méndez, Y.; García-Rivera, D.; Rivera, D. G. Expanding the Scope of Ugi Multicomponent Bioconjugation to Produce Pneumococcal Multivalent Glycoconjugates as Vaccine Candidates. *Bioconjug. Chem.* **2020**, acs.bioconjchem.0c00423. https://doi.org/10.1021/acs.bioconjchem.0c00423.

(322) Sornay, C.; Hessmann, S.; Erb, S.; Dovgan, I.; Ehkirch, A.; Botzanowski, T.; Cianférani, S.; Wagner, A.; Chaubet, G. Investigating Ugi/Passerini Multicomponent Reactions for the Site-Selective Conjugation of Native Trastuzumab**. *Chem. – Eur. J.* **2020**, *26* (61), 13797–13805. https://doi.org/10.1002/chem.202002432.

(323) Egloff, C.; Jacques, S. A.; Nothisen, M.; Weltin, D.; Calligaro, C.; Mosser, M.; Remy, J.-S.; Wagner, A. Bio-Specific and Bio-Orthogonal Chemistries to Switch-off the Quencher of a FRET-Based Fluorescent Probe: Application to Living-Cell Biothiol Imaging. *Chem Commun* **2014**, *50* (70), 10049–10051. https://doi.org/10.1039/C4CC03548H.

(324) Qiu, G.; Sornay, C.; Savary, D.; Zheng, S.-C.; Wang, Q.; Zhu, J. From Isonitrile to Nitrile via Ketenimine Intermediate: Palladium-Catalyzed 1,1-Carbocyanation of Allyl Carbonate by

α-Isocyanoacetate. *Tetrahedron* **2018**, *74* (49), 6966–6971. https://doi.org/10.1016/j.tet.2018.10.032.

(325) Stairs, S.; Neves, A. A.; Stöckmann, H.; Wainman, Y. A.; Ireland-Zecchini, H.; Brindle, K. M.; Leeper, F. J. Metabolic Glycan Imaging by Isonitrile-Tetrazine Click Chemistry. *ChemBioChem* **2013**, *14* (9), 1063–1067. https://doi.org/10.1002/cbic.201300130.

(326) Freedman, D. A.; Keresztes, I.; Asbury, A. L. Metal–Coumarin Complexes: Synthesis and Characterization of 7-Isocyanocoumarin Ligands and Mo(CO)4(7-Isocyanocoumarin)2 Complexes. X-Ray Crystal Structure of Mo(CO)4(7-Isocyano-4-Trifluoromethylcoumarin)2. *J. Organomet. Chem.* **2002**, *642* (1–2), 97–106. https://doi.org/10.1016/S0022-328X(01)01186-X.

(327) Leriche, G.; Nothisen, M.; Baumlin, N.; Muller, C. D.; Bagnard, D.; Remy, J.-S.; Jacques, S. A.; Wagner, A. Spiro Diorthoester (SpiDo), a Human Plasma Stable Acid-Sensitive Cleavable Linker for Lysosomal Release. *Bioconjug. Chem.* **2015**, *26* (8), 1461–1465. https://doi.org/10.1021/acs.bioconjchem.5b00280.

(328) Ursuegui, S.; Schneider, J. P.; Imbs, C.; Lauvoisard, F.; Dudek, M.; Mosser, M.; Wagner, A. Expedient Synthesis of Trifunctional Oligoethyleneglycol-Amine Linkers and Their Use in the Preparation of PEG-Based Branched Platforms. *Org. Biomol. Chem.* **2018**, *16* (44), 8579–8584. https://doi.org/10.1039/C8OB02097C.

(329) Saarbach, J.; Masi, D.; Zambaldo, C.; Winssinger, N. Facile Access to Modified and Functionalized PNAs through Ugi-Based Solid Phase Oligomerization. *Bioorg. Med. Chem.* **2017**, *25* (19), 5171–5177. https://doi.org/10.1016/j.bmc.2017.05.064.

(330) Wang, Y.; Patil, P.; Dömling, A. Easy Synthesis of Two Positional Isomeric Tetrazole Libraries. *Synthesis* **2016**, *48* (21), 3701–3712. https://doi.org/10.1055/s-0035-1562435.

(331) Stöckmann, H.; Neves, A. A.; Stairs, S.; Brindle, K. M.; Leeper, F. J. Exploring Isonitrile-Based Click Chemistry for Ligation with Biomolecules. *Org. Biomol. Chem.* **2011**, *9* (21), 7303. https://doi.org/10.1039/c1ob06424j.

(332) Tu, J.; Xu, M.; Parvez, S.; Peterson, R. T.; Franzini, R. M. Bioorthogonal Removal of 3-Isocyanopropyl Groups Enables the Controlled Release of Fluorophores and Drugs in Vivo. *J. Am. Chem. Soc.* **2018**, *140* (27), 8410–8414. https://doi.org/10.1021/jacs.8b05093.

(333) Tu, J.; Svatunek, D.; Parvez, S.; Liu, A. C.; Levandowski, B. J.; Eckvahl, H. J.; Peterson, R. T.; Houk, K. N.; Franzini, R. M. Stable, Reactive, and Orthogonal Tetrazines: Dispersion Forces Promote the Cycloaddition with Isonitriles. *Angew. Chem. Int. Ed.* **2019**, *58* (27), 9043–9048. https://doi.org/10.1002/anie.201903877.

(334) Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V.; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P. Iminoboronates: A New Strategy for Reversible Protein Modification. *J. Am. Chem. Soc.* **2012**, *134* (24), 10299–10305. https://doi.org/10.1021/ja303436y.

(335) Debaene, F.; Bœuf, A.; Wagner-Rousset, E.; Colas, O.; Ayoub, D.; Corvaïa, N.; Van Dorsselaer, A.; Beck, A.; Cianférani, S. Innovative Native MS Methodologies for Antibody Drug Conjugate Characterization: High Resolution Native MS and IM-MS for Average DAR and DAR Distribution Assessment. *Anal. Chem.* **2014**, *86* (21), 10674–10683. https://doi.org/10.1021/ac502593n.

(336) Egloff, C.; Jacques, S. A.; Nothisen, M.; Weltin, D.; Calligaro, C.; Mosser, M.; Remy, J.-S.; Wagner, A. Bio-Specific and Bio-Orthogonal Chemistries to Switch-off the Quencher of a FRET-Based Fluorescent Probe: Application to Living-Cell Biothiol Imaging. *Chem Commun* **2014**, *50* (70), 10049–10051. https://doi.org/10.1039/C4CC03548H.

(337) Freedman, D. A.; Keresztes, I.; Asbury, A. L. Metal–Coumarin Complexes: Synthesis and Characterization of 7-Isocyanocoumarin Ligands and Mo(CO)4(7-Isocyanocoumarin)2 Complexes. X-Ray Crystal Structure of Mo(CO)4(7-Isocyano-4-Trifluoromethylcoumarin)2. *J. Organomet. Chem.* **2002**, *642* (1–2), 97–106. https://doi.org/10.1016/S0022-328X(01)01186-X.

(338) Leriche, G.; Nothisen, M.; Baumlin, N.; Muller, C. D.; Bagnard, D.; Remy, J.-S.; Jacques, S. A.; Wagner, A. Spiro Diorthoester (SpiDo), a Human Plasma Stable Acid-Sensitive

Cleavable Linker for Lysosomal Release. *Bioconjug. Chem.* **2015**, *26* (8), 1461–1465. https://doi.org/10.1021/acs.bioconjchem.5b00280.

(339) Wang, Y.; Patil, P.; Dömling, A. Easy Synthesis of Two Positional Isomeric Tetrazole Libraries. *Synthesis* **2016**, *48* (21), 3701–3712. https://doi.org/10.1055/s-0035-1562435.

(340) Stöckmann, H.; Neves, A. A.; Stairs, S.; Brindle, K. M.; Leeper, F. J. Exploring Isonitrile-Based Click Chemistry for Ligation with Biomolecules. *Org. Biomol. Chem.* **2011**, *9* (21), 7303–7305. https://doi.org/10.1039/C1OB06424J.

(341) Schieber, C.; Bestetti, A.; Lim, J. P.; Ryan, A. D.; Nguyen, T.-L.; Eldridge, R.; White, A. R.; Gleeson, P. A.; Donnelly, P. S.; Williams, S. J.; Mulvaney, P. Conjugation of Transferrin to Azide-Modified CdSe/ZnS Core–Shell Quantum Dots Using Cyclooctyne Click Chemistry. *Angew. Chem. Int. Ed.* **2012**, *51* (42), 10523–10527. https://doi.org/10.1002/anie.201202876.

(342) Ursuegui, S.; Recher, M.; Krężel, W.; Wagner, A. An in Vivo Strategy to Counteract Post-Administration Anticoagulant Activity of Azido-Warfarin. *Nat. Commun.* **2017**, *8* (1), 1–8. https://doi.org/10.1038/ncomms15242.

(343) Dovgan, I.; Ursuegui, S.; Erb, S.; Michel, C.; Kolodych, S.; Cianférani, S.; Wagner, A. Acyl Fluorides: Fast, Efficient, and Versatile Lysine-Based Protein Conjugation via Plug-and-Play Strategy. *Bioconjug. Chem.* **2017**, *28* (5), 1452–1457. https://doi.org/10.1021/acs.bioconjchem.7b00141.

(344) Baatarkhuu, Z.; Chaignon, P.; Borel, F.; Ferrer, J.-L.; Wagner, A.; Seemann, M. Synthesis and Kinetic Evaluation of an Azido Analogue of Methylerythritol Phosphate: A Novel Inhibitor of E . Coli YgbP/IspD. *Sci. Rep.* **2018**, *8* (1), 1–10. https://doi.org/10.1038/s41598-018-35586-y.

(345) Tu, J.; Xu, M.; Parvez, S.; Peterson, R. T.; Franzini, R. M. Bioorthogonal Removal of 3-Isocyanopropyl Groups Enables the Controlled Release of Fluorophores and Drugs in Vivo. *J. Am. Chem. Soc.* **2018**, *140* (27), 8410–8414. https://doi.org/10.1021/jacs.8b05093.

(346) Carapito, C.; Burel, A.; Guterl, P.; Walter, A.; Varrier, F.; Bertile, F.; Dorsselaer, A. V. MSDA, a Proteomics Software Suite for in-Depth Mass Spectrometry Data Analysis Using Grid Computing. *PROTEOMICS* **2014**, *14* (9), 1014–1019. https://doi.org/10.1002/pmic.201300415.

(347) Bouyssié, D.; Hesse, A.-M.; Mouton-Barbosa, E.; Rompais, M.; Macron, C.; Carapito, C.; de Peredo, A. G.; Couté, Y.; Dupierris, V.; Burel, A.; Menetrey, J.-P.; Kalaitzakis, A.; Poisat, J.; Romdhani, A.; Burlet-Schiltz, O.; Cianférani, S.; Garin, J.; Bruley, C. Proline: An Efficient and User-Friendly Software Suite for Large-Scale Proteomics. *Bioinformatics* **2020**. https://doi.org/10.1093/bioinformatics/btaa118.

(348) Reguera, L.; Méndez, Y.; Humpierre, A. R.; Valdés, O.; Rivera, D. G. Multicomponent Reactions in Ligation and Bioconjugation Chemistry. *Acc. Chem. Res.* **2018**, *51* (6), 1475–1486. https://doi.org/10.1021/acs.accounts.8b00126.

(349) Afshari, R.; Shaabani, A. Materials Functionalization with Multicomponent Reactions: State of the Art. *ACS Comb. Sci.* **2018**, *20* (9), 499–528. https://doi.org/10.1021/acscombsci.8b00072.

(350) Rocha, R. O.; Rodrigues, M. O.; Neto, B. A. D. Review on the Ugi Multicomponent Reaction Mechanism and the Use of Fluorescent Derivatives as Functional Chromophores. *ACS Omega* **2020**, *5* (2), 972–979. https://doi.org/10.1021/acsomega.9b03684.

Charlotte SORNAY

Nouvelles stratégies de bioconjugaison de protéines natives

Résumé

Au cours des dernières décennies, les méthodes permettant la modification chémo- et site-sélectives de protéines ont été les objets de nombreux travaux, en particulier dans le domaine thérapeutique. Bien que les modifications site-sélective aient été jusqu'alors limitées aux protéines d'ingénierie, réaliser de telles modifications sur des protéines natives devient possible grâce à des méthodes émergentes permettant la fonctionnalisation contrôlée de différents acides aminés. Les travaux présentés dans cette thèse ont pour objet le développement de nouvelles méthodes chimiques pour la modification de protéines natives. La première partie décrit l'utilisation d'iodes hypervalents pour le marquage des cystéines. L'optimisation minutieuse du réactif a permis de réaliser des conjugaisons extrêmement rapides et sélectives des thiols. Dans la seconde partie, nous nous sommes inspirés des méthodes dites de ligature chimique pour conjuguer des résidus lysine de manière régiosélective à la suite d'une étape de pré-conjugaison de résidus cystéine. Enfin, nous avons développé une nouvelle méthode de conjugaison site-sélective mettant en œuvre la réaction de Ugi pour conjuguer simultanément les groupements amine et carboxylate des chaînes latérales de résidus lysine et aspartate/glutamate voisins.

Mots-clés : anticorps, bioconjugaison, délivrance de médicaments, iodes hypervalents, modification de protéines, réactions multi-composantes.

Abstract

Over the past decades, chemoselective and in particular, site-selective methods for the modification of proteins have been the subjects of intense studies, with a special emphasis on therapeutic applications. While protein engineering techniques have facilitated site-selective modification of proteins, methods to achieve the same results from native proteins are starting to bear fruits, allowing the controlled functionalization of various types of amino acids. This work is devoted to the development of new chemical strategies for the modification of native proteins. The first part describes the use of hypervalent iodines for the labelling of cysteines residues. Fine tuning of the reagent enabled extremely fast conjugation, with excellent thiol-selectivity. In a second part, we describe a strategy inspired by native chemical ligation that allows regioselective lysine-conjugation after pre-conjugation to cysteine residues. Finally, we set to employ the Ugi reaction as a novel site-selective method, with the aim of simultaneously conjugating the side-chain amine and carboxylate groups of two neighboring lysines and aspartate/glutamate residues.

Keywords: antibodies, bioconjugation, drug-delivery, hypervalent iodines, multicomponent reactions, protein modifications.