Protein Engineering Strategies for Improved Pharmacokinetics

Aurélie Rondon, Sohaib Mahri, Francisco Morales, Mireille Dumoulin and Rita Vanbever*

Dr. A. Rondon, Dr. S. Mahri, Prof. R. Vanbever

Université catholique de Louvain (UCLouvain)

Louvain Drug Research Institute

Advanced Drug Delivery & Biomaterials

1200 Brussels, Belgium

E-mail : rita.vanbever@uclouvain.be

ORCIDs: 0000-0001-9981 (AR); 0000-0003-2663-3829 (SM); 2460; 0000-0002-0784-8771 (RV)

Dr. F. Morales, Dr. M. Dumoulin

Université de Liège

Center for Protein Engineering InBIOS

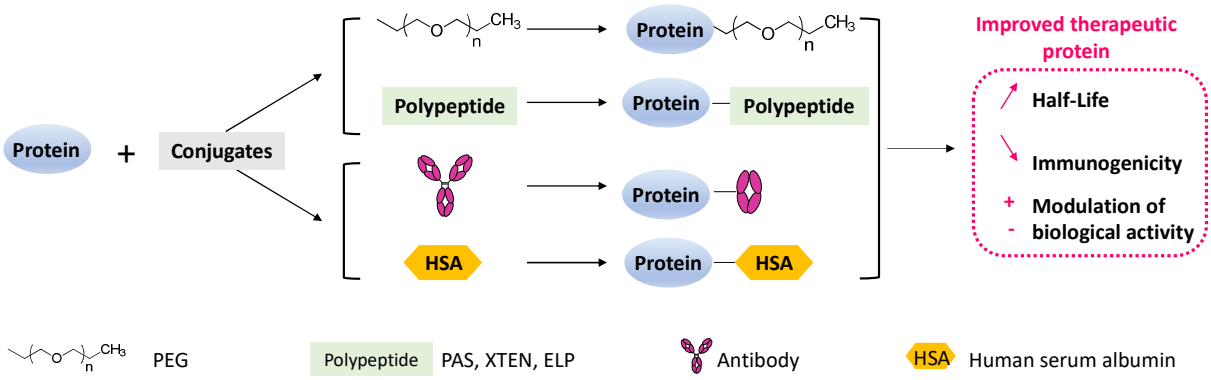
4000 Liège 1, Belgium

ORCIDs: 0000-0002-2297-1479 (FM); 0000-0002-4310-4622 (MD)

Keywords: pharmacokinetic, protein modifications, PEG, polypeptides, Fc-fusion, human serum albumin, genetic engineering, chemical conjugation, chemo-enzymatic reactions

Abstract

Protein therapeutics have gained momentum in recent years and become a pillar in treating many diseases and the only choice in several ailments. Protein therapeutics are highly specific, tunable and less toxic than conventional small molecules. However, reaping the full benefits of therapeutic proteins in the clinics is often hindered by issues of immunogenicity and short half-life due essentially to fast renal clearance and enzymatic degradation. Advances in polymer chemistry and protein engineering allowed overcoming some of these limitations. Strategies to prolong the half-life of proteins rely on increasing their size and stability and/or fusing them to endogenous proteins (albumin, Fc fragment of antibody) to hijack physiological pathways involved in protein recycling. On the downside, these modifications might alter therapeutic proteins structure and function. Therefore, a compromise between half-life and activity is sought. This review covers half-life extension strategies using natural and synthetic polymers as well as fusion to other proteins and sheds light on genetic engineering strategies and chemical and enzymatic reactions to achieve this goal. Promising strategies and successful applications in the clinics are highlighted.



Graphical abstract: In the last decades, breakthroughs in protein engineering strategies have revolutionized the development of long-acting medicines. Conjugation of polymers, polypeptides, Fc antibody fragments or human serum albumin to a parent therapeutic protein increases its half-life and modulate protein biological activity.

List of abbreviations

AGHD: Adult growth hormone deficiency

AUC: Area under the curve

CDR: Complementary determining region

CEA: Carcinoembryonic antigen

CPG2: Glutamate carboxypeptidase G2

CRS: Cytokine release syndrome

DARPin: Designed ankyrin repeat protein

EGFR: Epidermal growth factor receptor

ELP: Elastase-like protein

EMA: European Medicines Agency

EPO: Erythropoietin

EPR: Enhanced permeability retention effect

Fab: Antigen binding fragment

FcγR: Fc gamma receptor

FcRn: Neonatal Fc receptor

FDA: Food and Drug Administration

Fv: Variable fragment

G-CSF: Granulocyte-colony-stimulating factor

GFP: Green-fluorescent protein

GLK: Glucagon-like protein

GLP: Glucagon-like peptide

GLP-1: Glucagon-like peptide 1

GLP-1R: Glucagon-like peptide 1 Receptor

HAP: Glycine-rich polypeptide

HcAb: Heavy-chain antibodies

HFt: Human ferritin

hGH: Human growth hormone

IEDDA: Inverse-electron demand Diels-Alder cycloaddition

mAbs: Monoclonal antibodies

MP: Metalloproteinase

NHS: *N*-hydroxysuccinimide

NTA: Ni-nitrilotriacetic acid

PAS: Proline-alanine-serine

PBS : phosphate buffer saline

PD: Pharmacodynamic

PEG: Polyethylene glycol

PET: Positron emission tomography

PGHD: Pediatric growth hormone deficiency

PK: Pharmacokinetic

SA: Serum albumin

TCE: T-cell engagers

TCO: *Trans*-cyclooctene

TfR: Transferrin receptor

TNF: Tumour necrosis factor

Tz : Tetrazine

VHH: Variable heavy homodimer

VIP : Vasoactive intestinal peptide

VNARs: Variable domain of new antigen receptors

VWF: Von Willebrand factor

WT: Wild type

XPAT: XTENylated protease-activated T Cell Engager

XTEN: Polypeptides composed of alanine, glutamate, glycine, proline, serine, and threonine

Table of contents

1. Introduction

Since the market introduction of recombinant human insulin in 1982, biopharmaceuticals have gained momentum. More than 200 protein therapeutics are currently on the market and above 1,000 are in clinical development.[1] Antibodies continue to largely dominate biopharmaceutical approvals because their clinical applications are multiple in a wide range of diseases including cancer, cardiovascular disease, organ transplantation, autoimmunity, inflammation, and infection. Protein therapeutics present several advantages over small molecule drugs.[2] Proteins serve a highly specific and complex set of physiological functions that cannot be mimicked by simple chemical compounds. Since the action of proteins is highly specific, they barely interfere with normal biological processes and cause less adverse events. Protein therapeutics are frequently derived from proteins naturally produced by the body. These agents are therefore often well tolerated and poorly immunogenic.

However, proteins also suffer from significant limitations. Proteins with a molecular weight below the threshold for kidney filtration (67 kDa, the size of human serum albumin) are cleared from the systemic circulation within a day. Many proteins are even cleared within a few hours or a few minutes when metabolism contributes to elimination. Therefore, therapeutic proteins need to be injected to patients several times a week (*e.g.*, erythropoietin) or even several times a day (*e.g.*, glucagon-like peptide-1 or GLP-1), resulting in peaks and valleys in plasma concentrations with the alternate risks of systemic side effects and suboptimal therapeutic concentrations. Moreover, frequent administration of medication causes patient discomfort and reduces quality of life. A second limitation of proteins lies in protein immunogenicity. Foreign proteins from prokaryotes or animals might present interesting therapeutic properties in humans. However, intrinsic immunogenicity of non-human proteins hampers their therapeutic use in the clinic because specific antibodies generated against the foreign protein neutralize its activity and result in a loss of therapeutic efficacy over time. The unwanted immune response might even cause more serious general immune effects such as anaphylaxis.

Over the last three decades, protein engineering has largely demonstrated that it can provide solutions to the limitations of natural proteins. Breakthroughs in the field led to the development of biobetters, *i.e.*, conjugated or modified proteins with improved properties and modular functions over the original biologic. In particular, since the nineties, dozens of

biobetters with an improved pharmacokinetic profile have been approved for clinical use. Adagen®, the first biobetter marketed in 1990, is a chemical conjugate between bovine adenosine deaminase and polyethylene glycol (PEG). It is used for enzyme replacement therapy in severe combined immunodeficiency disease associated with a deficiency of adenosine deaminase. The about fifteen PEG chains attached to bovine adenosine deaminase decrease its immunogenicity and impart an elimination half-life of 5 days to the protein. This PEGylation strategy has then been applied to many other proteins with the main goal to place protein molecular weight above the threshold for kidney filtration and protect the protein from catabolism, thereby increasing serum half-life. Another major breakthrough has come from the discovery of the FcRn-mediated recycling pathway taken by immunoglobulins G (IgGs) and serum albumin (SA). IgGs and SA are internalized in endosomes of endothelial cells where they are protected from degradation by binding FcRn. Recycling through FcRn interactions result in the long serum half-lives of IgGs and SA and these serum proteins have been attached to many unrelated therapeutic proteins to increase their half-life. For instance, Albiglutide, marketed in 2013, is a fusion protein between a peptidase-resistant GLP 1 analog and SA. Albiglutide exhibits an extended half-life (5 days) compared with native GLP-1 (3 minutes).[3]

This review will detail the different protein engineering strategies that have been harnessed to extend protein half-life in the systemic circulation. These approaches comprise the preparation of polymer–protein conjugates and the exploitation of the long half-lives of IgGs and SA. In addition to PEGylation, XTENylation, PASylation and ELPylation have emerged. XTEN, PAS and ELP are natively disordered polypeptide polymers that have been used as alternatives to PEG. Polypeptide polymers offer the advantages of biodegradability and recombinant production as a single fusion product over PEG. IgG and SA have been used to prolong the half-life of proteins either by fusing the therapeutic protein to a FcRn binding protein (Fc domain of an IgG or SA) or, by conjugating the therapeutic protein to a molecule which non-covalently binds to SA. To end, this review will thoroughly describe the chemical and chemo-enzymatic methods used to modify proteins and graft peptide or prosthetic groups to them.

2. Polymer-protein conjugates

One of the best and widely investigated approaches to prolong serum half-life relies on the conjugation of highly soluble non-toxic polymers to bioactive proteins.[4] Polymer-protein conjugates present an increased hydrodynamic diameter impeding their clearance *via* kidney glomeruli filtration. Among polymers, PEG has been the first and most used, resulting in the commercialization of more than 15 biobetters (**Table *1***) and several others are still in clinical development. Yet, to overcome the non-biodegradability property of PEG, polypeptides have recently emerged as alternatives.

2.1. PEG

In 1977, the group of Frank Davis published a seminal scientific article on the impact of conjugation to PEG on the immunogenicity and blood circulating life of bovine liver catalase.[5] The initial goal of Davis was to minimize or eliminate the immunogenicity of therapeutic enzymes from non-human sources by covering antigenic determinants by a linear, flexible, uncharged and hydrophilic polymer. Thus, bovine liver catalase was randomly conjugated to 1.9 or 5 kDa PEG on its lysine residues. PEGylation of catalase was shown to decrease catalase-specific antibody production following injection in rabbits and to decrease the recognition of the protein by catalase-specific antibodies. The authors additionally observed that PEGylation greatly protected catalase from proteolysis and significantly prolonged its blood half-life in the rabbit.

2.1.1. The polymer

PEG is a linear or branched polyether with hydroxyl end groups with the general structure: $HO\text{-}(CH_2CH_2O)_n\text{-}CH_2CH_2\text{-}OH$.[6] Monomethoxy PEG, mPEG, $CH_3O\text{-}(CH_2CH_2O)_n\text{-}CH_2CH_2\text{-}OH$, is the most used for protein modification because its unique reactive group results in one-site attachment on the protein. PEG can present a wide range of molecular weights and some polydispersity (Mw/Mn is approximately 1.1), which is a drawback as it leads to undesired polydispersity of the conjugates. PEG is a neutral and amphiphilic polymer. The repeated ethylene moiety along the PEG chain is responsible for the polymer hydrophobicity, whereas the oxygen confers strong interactions with water: three water molecules are bound per monomer unit. Therefore, PEG is soluble in both organic and aqueous media and is highly hydrated. The carbon-carbon and carbon-oxygen bonds offer high flexibility to the whole polymer. The high mobility and hydration of PEG lead to a large and very effective exclusion

volume of approaching molecules. Accordingly, the polymer has a hydrodynamic volume five to ten times higher than that of a globular protein of the equivalent molecular weight.[7]

2.1.2. Immunogenicity

The repeated administration of therapeutic proteins can be highly immunogenic, especially in case of foreign proteins but also in case of human proteins. For instance, Vaisman *et al*. reported that chimeric monoclonal antibodies exhibited immunogenicity in up to 70% of patients and fully human monoclonal antibodies in up to 30%.[8] Protein immunogenicity generates anti-protein antibodies which can neutralize the therapeutic activity of the protein and cause allergic reactions.

PEGylation is able to decrease the immunogenicity of therapeutic proteins. Accordingly, PEGylation demonstrated a tremendous success and brought to market several proteins of non-human origin which might have never reached it as unconjugated versions due to their intrinsic immunogenicity. This is the case of Pegademase and its new recombinant version Elapegademase, Pegasparagase and its new longer-acting version Calasparagase pegol, and Pegvaliase. All these PEGylated proteins are enzymes from either prokaryotes (*E. coli*) or animals (beef, pig; **Table *1***). These are mainly used as enzyme replacement therapies in inherited enzyme deficiencies but also as a therapeutic treatment in leukemia.

Asparaginase is a critical component in the treatment of acute lymphoblastic leukemia.[9] Asparaginase hydrolyzes the amino acid L-asparagine to L-aspartic acid and ammonia. L-asparagine is synthesized in most human tissues from L-glutamine. However, acute lymphoblastic leukemia cells have very low levels of asparagine synthase and asparagine depletion causes their apoptosis. Asparaginase extracted from *E. coli* was approved by the FDA in 1978 and then withdrawn in 2012. Hypersensitivity was the most common adverse reaction to *E. coli* asparaginase and it occurred in up to one third of patients. Pegaspargase has been FDA-approved in 1994 and is the primary form of asparaginase in clinical use today. Pegaspargase shows a rate of allergic reactions of only 10% in naive patients.[10] PEGylation of asparaginase involves the random attachment of approximately 50 PEG chains of 5 kDa to its lysine residues and PEGylation increases asparaginase half-life from 24 h (unconjugated protein) to 5.5 days (Pegaspargase) and 13.5 days (Calasparagase pegol).[11] Calaspargase pegol uses the identical enzyme and polyethylene glycol moieties present in Pegaspargase. However, the succinimidyl carbamate linker used in Calaspargase pegol is more hydrolytically

stable than the succinimidyl succinate linker used in Pegaspargase which results in a longer half-life.

Pegademase bovine is an adenosine deaminase derived from bovine intestine and conjugated to 11 to 17 chains of 5 kDa PEG. Pegademase does not induce hypersensitivity reactions. However, there have been reports on neutralizing antibodies. In a clinical trial, two out of 17 patients showed an enhanced rate of clearance of plasma adenosine deaminase activity after 4 months of therapy.[12] Enhanced clearance was correlated with the appearance of an antibody that directly inhibited both the activity of unmodified adenosine deaminase and pegademase. Patients who previously received pegademase bovine may present antibodies to Elapegademase, a recombinant bovine adenosine deaminase manufactured in *E. coli*. Therefore, thorough plasma adenosine deaminase activity is monitored in patients for any persistent activity decline.

Pegvaliase is a PEGylated recombinant phenylalanine ammonia lyase derived from the cyanobacterium *Anabaena variabilis* and expressed in *E. coli*. Pegvaliase converts phenylalanine to ammonia and *trans*-cinnamic acid, and is indicated for the treatment of patients with phenylketonuria who have inadequate blood phenylalanine control. Although the enzyme is protected by 28 to 44 PEG chains of 20 kDa each, hypersensitivity reactions have been reported in 75% of patients treated with Pegvaliase and the acute systemic Type III (immune complex mediated) hypersensitivity reaction has been the most clinically significant and reached 6% of the patients. All patients treated with Pegvaliase developed a sustained anti-phenylalanine ammonia lyase and anti-PEG IgM and IgG response. Because antibodies bind to the PEG portion of Pegvaliase, binding with other PEGylated therapeutics and increased hypersensitivity to other PEGylated injectables might occur. Neutralizing antibodies capable of inhibiting the enzyme activity were detected in the majority of patients. Patients with higher antibody titers required higher doses to overcome clearance and achieve blood phenylalanine reduction.

The attachment of several small PEG chains to a protein better decreases its immunogenicity than the attachment of one large PEG chain because several PEG chains more widely shield the protein surface.[13] This PEGylation strategy has been followed in all the examples presented above. It is just the opposite approach to the one used to preserve protein activity

and only prolongs serum half-life where the conjugation to a unique large PEG chain on the protein side opposite to the active site is sought.

2.1.3. Prolongation of serum half-life and residence time in the lungs

PEGylation of fully human proteins aims to increase their serum half-life and thereby increase patient convenience by decreasing administration frequency. The serum half-life of proteins can increase up to 20-fold following PEGylation. For instance, Certolizumab pegol, a Fab antibody fragment conjugated to a 2-armed 40 kDa PEG in C-terminal, exhibits an elimination half-life of 14 days while unconjugated Fab antibody fragments show a half-life of 12-20 h.[14] The attachment of a PEG chain to a protein places its molecular weight above the threshold for kidney filtration and reduces renal clearance. In addition, PEG attachment to biopharmaceuticals can protect them from proteolysis. The half-lives of PEGylated protein conjugates increase with the molecular weight of the PEG and with the number of conjugated PEG chains.[15]

More recently, PEGylation has been shown to prolong the residence time of protein therapeutics in the lungs and to improve their local therapeutic efficacy in preclinical models of respiratory diseases.[16] PEGylation of recombinant human alpha1-antitrypsine with a 20 kDa PEG sustained the presence of the conjugate in the lungs of mice for 48 h, whereas the non-PEGylated counterpart was cleared within 24 h.[17] Recombinant human alpha1-antitrypsin conjugated to 20 kDa PEG protected mice against human leukocyte elastase-induced lung hemorrhage and the protection was sustained for 72 h. PEGylation of an anti-IL-17A Fab' antibody fragment with 2-armed 40 kDa PEG increased its residence time in the lungs of mice, rats and rabbits to more than 48 h while the unconjugated Fab' was cleared from the lungs within 24 h.[18] The prolonged pulmonary residency of the anti-IL-17A PEGylated antibody fragment translated in an improved efficacy in reducing lung inflammation in a murine model of house dust mite-induced lung inflammation.[19] Conjugation of PEG to recombinant human deoxyribonuclease I (rhDNase) resulted in an impressive extension of its residence time (≥ 15 days) in the murine lungs.[20] Moreover, one single dose of PEGylated rhDNase was as effective as 1 daily dose of unconjugated rhDNase during 5 days in decreasing the DNA content in the lungs of β-ENaC mice, a model of the cystic fibrosis lung disease. The lack of marketed PEGylated proteins for pulmonary delivery reflects the paucity of approved proteins for inhalation in the first place. However, in 2020, Bayer has initiated a phase 2 clinical

trial on an inhaled PEGylated peptide (PEGylated adrenomedullin or BAY1097761) for the treatment of acute respiratory distress syndrome (NCT04417036).

Several mechanisms might explain the sustained retention of PEGylated proteins within the lungs. The increase in molecular size decreases the protein transport across the alveolar-capillary barrier towards the systemic circulation. The steric hindrance created by PEG chains on the protein surface prevents proteases from degrading the protein.[21] Finally, the hydrophilic nature of PEG decreases protein interactions with the cell membrane and thereby, decreases protein endocytosis by epithelial cells and alveolar macrophages.[22]

2.1.4. Biological activity and impact on protein stability

PEGylation might result in a partial loss of the biological activity of the therapeutic protein. In order to avoid a reduction in biological activity, the PEGylation site should be distant from the active site. Indeed, the active site of the protein may be masked due to the steric hindrance of PEG. Accordingly, site-specific addition of 5 kDa-PEG to tumor necrosis factor resulted in a 20% decrease in the activity of the cytokine while random PEGylation led to a 90% activity loss.[23] In spite of reduced activity, several PEG-protein conjugates for injection are commercially available because of the tremendous increase in serum half-life. For instance, Pegasys® (40 kDa PEG-IFNα2a) only retains 7% of the wild-type interferon activity and is on the market since 2002.[24]

All marketed PEG-protein conjugates delivered by injection (Table 1) involve a permanent covalent link between the polymer chain and the protein and the PEGylated construct is the active entity. In contrast to conventional permanent PEGylation, a new technology called TransCon is currently developed by Ascendis Pharma where the PEG-protein conjugates are inactive prodrugs.[25] Accordingly, the protein is transiently bound to a four-arm 40 kDa PEG and the steric hindrance created by the polymer inactivates the protein. With the hydrolysis of the TransCon linker, the unmodified protein is gradually released in the body. The advantages of this technology are easily highlighted by presenting Ascendis Pharma flagship product, TransCon hGH for which a market authorization application has been submitted. [25] Human growth hormone (hGH) replacement therapy needs to achieve the same tissue distribution and receptor activation as endogenous hGH because hGH receptors are in essentially all tissues. Restricted access of protein-enlarged human GH into peripheral tissues led to unexpected outcome such as injection site lipoatrophy and to discontinuation of the

development of a permanently-PEGylated hGH. TransCon hGH leverages the known pharmacology and distribution of unmodified hGH with the properties of an inert PEG carrier molecule and avoids imbalances in organ distribution. It allows once-weekly dosing and will ease the lives of patients with hGH deficiency. TransCon PTH and TransCon C-type Natriuretic Peptide are other prodrug therapies in development by Ascendis Pharma.[26]

No critical changes to protein secondary and tertiary structures have been noted following PEGylation.[27] PEGylation generally increases the stability of proteins to aggregation. For instance, a Fab' antibody fragment conjugated to two PEG chains of 30 kDa presented higher resistance to protein aggregation than the unconjugated Fab' when exposed to heat and agitation.[28] The steric hindrance created by PEG likely prevents the association of unfolded proteins. However, there are cases where the propensity to aggregation increased.[27a] The number and size of the PEG chains as well as the type of the conjugation link can affect protein stability.[27a, 29]

2.1.5. Safety

Small molecular weight PEGs (< 10 kDa) are common excipients in oral, intravenous, nasal and inhalation formulations. However, larger PEGs (up to 40 kDa) are used in PEGylated protein therapeutics. PEG is non-biodegradable and its primary clearance mechanism is renal excretion of the intact molecule. Yet, above 30 kDa-PEG, renal ultrafiltration is markedly reduced and liver uptake and excretion through the bile take over.[30]

PEG is generally considered to have low toxicity whatever its molecular weight and route of administration. Complement activation and impact on coagulation have been observed at very high PEG concentrations (1-40 mg/mL).[31] However, these concentrations are largely exceeding the plasma concentrations reached after injection of PEGylated protein therapeutics that are rather in the ng to µg range.

Ivens *et al*. reviewed the preclinical safety data collected on PEGylated protein therapeutics administered by injection currently on the market.[32] Adverse effects observed in preclinical studies were usually related to the pharmacologically active drug component of the molecule, rather than to the PEG moiety. Cellular vacuolation in certain tissues and cell types has been observed for approximately half of the approved PEGylated drugs. Vacuolation was seen most frequently in macrophages. Cytoplasmic vacuolation probably reflects the body's normal

response to clear a foreign non-biodegradable body. No functional changes related to PEG for organs and tissues where cellular vacuolation was seen have been reported. Vacuolation was absent below a certain dose of PEG per month (0.4 µmol/kg/month) and it has been observed for PEG molecular weight of at least 30 kDa. Vacuolation was reversible provided sufficient recovery time was allowed.

PEG has been considered as a non-antigenic and non-immunogenic component. However, a number of reports have documented the presence of anti-PEG antibodies, such as IgG and IgM, following repeated injections of PEGylated proteins in the clinic (see above). The generation of anti-PEG antibodies is favored when the protein moiety is highly immunogenic. Anti-PEG antibodies were also found in 20 to 25% of 350 healthy blood donors who have not received PEGylated biopharmaceuticals.[7] Everyday use compounds such as cosmetics, food or household chemicals contain PEG and this might explain the occurrence of anti-PEG antibodies in a subset of the population. However, the neutralizing character of these anti-PEG antibodies has not been demonstrated. It should be noted that other neutral soluble polymers such as polyvinylpyrrolidone have been shown to decrease exogenous proteins immunogenicity but that PEG decreased it the most and generated the weakest anti-polymer antibody response.[33]

Year of approval	Commercial name	Generic name	Parent drug	Protein size (kDa)	PEG size and number	Bioconjugation method	Main site of Attachment	Application
1990	Adagen®	Pegademase	Adenosine deaminase	40	11-17 × 5 kDa	NHS ester ligation	Lysines	SCID
1994	Oncaspar®	Pegasparagase	Asparaginase	31	50 × 5 kDa	NHS ester ligation	Lysines	Leukemia
2000	PegIntron®	Peginterferon-α-2b	Interferon-α-2b	19.2	1 × 12 kDa	Urethane bond	Histidines	Hepatitis C
2001	Pegasys®	Peginterferon-α-2a	Interferon-α-2a	19.2	1 × 40 kDa	NHS ester ligation	Lysines	Hepatitis C
2002	Neulasta®	Pegfilgrastim	G-CSF	18.8	1 ×20 kDa	Aldehyde conjugation	N-terminal methionine	Neutropenia
2003	Somavert®	Pegvisomant	Human growth hormone	22	4-6 × 5 kDa	NHS ester ligation	Lysines	Acromegaly
2007	Mircera®	PEG-EPO	Erythropoietin	30	1 × 30 kDa	NHS ester ligation	Lysines	Anemia
2008	Cimzia®	Certolizumab Pegol	anti-TNFα Fab'	51	1 × 40 kDa	Maleimide conjugation	C-terminal cysteines	RA & Crohn disease
2010	Krystexxa®	Pegloticase*	Urate Oxidase	34	9 × 10 kDa	*p*-Nitrophenyl carbonate ester ligation	Lysines	Gout
2012	Omontys®	Peginesatide	Erythropoietin dimeric peptide	4.9	1 × 40 kDa	NHS ester ligation	N-terminal linker	ACKD

Year of approval	Commercial name	Generic name	Parent drug	Protein size (kDa)	PEG size and number	Bioconjugation method	Main site of Attachment	Application
2014	Plegridy®	Peginterferon beta-1a	Interferon β-1a	44	1 × 20 kDa	NHS ester ligation	Lysines	Multiple sclerosis
2016	Adynovate®	Antihemophilic pegylated factor	Coagulation factor VIII	280	1 × 20 kDa	NHS ester ligation	Lysines	Hemophilia A
2017	Refixia®	Nonacog beta pegol	Coagulation factor IX	50	1 × 40 kDa	Glycosylation	*N*-glycans	Hemophilia B
2018	Asparlas®	Calasparagase pegol	Asparagine enzyme	138	50 × 5 kDa	Urethane bond	Lysines	ALL
2018	Revcovi®	Elapegademase	Adenosine deaminase	115	11-17 × 5 kDa	NHS ester ligation	Alanines and lysines	ADA-SCID
2018	Jivi®	Damoctocog alfa pegol	Coagulation factor VIII	234	1 × 60 kDa	Maleimide conjugation	Cysteine	Hemophilia A
2018	Fulphila®	Pegfilgrastim-jmdb	G-CSF	40	1 × 20 kDa	Aldehyde conjugation	N-terminus methionines	Neutropenia
2018	Palynziq®	Pegvaliase	PAL enzyme	248	28-44 × 20 kDa	NHS ester ligation	Lysines	Phenylketonuria
2019	Esperoct®	Turoctocog alfa pegol	Coagulation factor VIII	166	1 × 40 kDa	*N*-glycan engineering	*O*-glycans	Hemophilia A

Table 1 : Chronological overview of the FDA-approved PEGylated protein drugs. ACKD: anemia associated chronic kidney disease; ADA: adenosine deaminase; ALL: acute lymphoblastic leukemia; AMD: age-mediated macular degeneration; Fab: antigen binding fragment; G-CSF: human granulocyte colony-stimulating factor; NHS: N-hydroxysuccinimide; RA: rheumatoid arthritis; SCID: severe combined immune deficiency; TNF: tumor necrosis factor; VEGF: vascular endothelial growth factor. *Market withdrawal in Europe in 2016 for economic reasons.

2.2. Polypeptides

Polypeptide polymers have been introduced in the hope of overcoming the limitations of PEG which suffers from several drawbacks including non-biodegradability and thereby possible risk of accumulation in cells and tissues and generation of antibodies against PEG and PEGylated proteins [34]

Similar to PEG, the action of polypeptides primarily relies on decreasing the clearance via increasing the hydrodynamic volume of the proteins they are fused to, but also on their shielding from proteolytic degradation, detection by the host immune system, and receptor-mediated clearance.[35] Yet, these polypeptides have not been used to decrease the immunogenicity of foreign proteins as PEG has and, up to now, their use has been limited to prolonging the serum half-life of human proteins. Polypeptides can be easily fused, *via* genetic engineering, to recombinant proteins and peptides. They are hydrophilic, stable, reportedly non-immunogenic, biodegradable, tuneable, do not alter the expression of proteins in bacterial systems, have large hydrodynamic volume thereby increasing the half-life of fused protein partners.[34b, 35-36] The two main polypeptides currently in active development are XTEN and PAS. HAPylation is the process of fusing a repeated sequence of a glycine-rich polypeptide; this strategy was developed by Schlapschy *et al*. for anti-HER2 Fab before discovering PAS but it is no longer pursued.[37]

2.2.1. XTEN

XTEN (loosely referred to as recombinant PEG) are genetically fused polypeptides composed of non-repetitive randomized segments of six chemically stable amino acids: alanine (A), glutamate (E), glycine (G), proline (P), serine (S) and threonine (T).[38] The selection of these amino acids is based on the idea of avoiding amino acids that might affect the solubility, activity, or stability of proteins. Therefore, positively charged amino acids (known to bind to the cell membrane) and amide-containing residues (could alter the stability of proteins) were excluded.[38] In addition, glycine and proline do not form secondary structures and provide a disordered conformation to XTEN.

XTEN can be expressed in *E. coli* system (> 8 mg/g wet-cell weight), making its production easy and cheap. Furthermore, the XTEN polypeptide length can easily be tuned, and the resulting conjugates are completely degradable into short peptides or amino acids. The proof-of-concept of XTEN using exenatide (GLP-1 receptor agonist, an antidiabetic peptide)

demonstrated that the 84 kDa-fused exenatide-XTEN is thermostable (up to 75 °C), has a large hydrodynamic radius and does not induce immunogenicity in mice.[38] Besides, exenatide-XTEN significantly improved the pharmacokinetics of the peptide by extending its half-life 65, 71, or 125-fold in rats, mice, or monkeys, respectively. Other peptides and proteins have been successfully XTENylated, such as glucagon, green fluorescent protein (GFP), factor VII, human growth hormone (hGH), teduglutide, a recombinant human Glucagon-like peptide 2 (GLP2-2G), annexin 5A, T-20 (antiretroviral peptide), and clotting factor IX.[38-39]

The encouraging *in vivo* results of XTEN fusion proteins have advanced three XTEN-conjugated proteins to clinical trials (**Table *2***). In phase I clinical trial, exenatide-XTEN (VRS-859) showed promising results in the glycemic control in patients suffering from type 2 diabetes mellitus[35]. The long half-life of ca. 5 days of exenatide-XTEN *vs* 2.4 h for unconjugated exenatide highlights the potential of a monthly administration Somavaratan (VRS-317) is a novel long-acting hGH for the treatment of hGH deficiency in children and adults.[40] Despite the 12-fold reduced potency of Somavaratan *in vitro* compared with hGH, the increased half-life of up to 60-fold resulted in an overall improved efficacy *in vivo*.[40b] Somavaratan demonstrated clinically significant improvements in the growth (height velocity and IGF-1) of prepubertal children in phase I clinical trials (NCT01718041). Adverse events following Somavaratan administration were similar to daily growth hormone in pediatric growth hormone deficiency and neutralizing antibodies were reported in 2 of the 64 children involved[25, 40b] However, twice-monthly SC injections of Somavaratan failed to meet the primary endpoint of non-inferiority compared to daily SC injection of reference drug rhGH (Genotropin®) in phase III VELOCITY clinical trials in children (NCT02339090), leading to the termination of phase II trials in adults (NCT02719990).

More promising results have been obtained in phase II clinical trials for the treatment of hemophilia A with BIVV001, a rFVIIIFc-VWF-XTEN construct (XTENylated recombinant coagulation Factor VIII Fc-von Willebrand Factor) (NCT03205163). BIVV001 was shown to be safe and to have a superior PK compared with the recombinant factor VIII.[41] In humans, a 3 to 4 increase in the half-life (9.1 h to 37.6 h and 13.2 h to 42.5 h for low and high doses, respectively) and up to a 7-fold increase in AUC were recorded compared with the recombinant factor VIII.[41a] Clinical trials have progressed to phase III to evaluate the long-

term safety and efficacy of weekly administration of BIVV001 in previously treated patients with severe hemophilia A (NCT04644575 and NCT04161495).

Name	Phase	Status	Protein conjugate	XTEN MW or a.a. number	Application	Ref
VRS-859	II	C	Exenatide-XTEN	~ 80 kDa (864 aa)	Type 2 diabetes	[40a]
BIVV001	III	R	rFVIIIFc-VWF-XTEN		Severe hemophilia A	NCT04161495
BIVV001	I/II	C	rFVIIIFc-VWF-XTEN	Two XTENs: 288 and 144 aa	Severe hemophilia A	NCT03205163
BIVV001	III	N	rFVIIIFc-VWF-XTEN		Severe hemophilia A	NCT04644575
Somavaratan (VRS-317)	II	T	hGH–XTEN		AGHD	NCT02719990
Somavaratan (VRS-317)	II	C	hGH–XTEN	Two XTENs: 83.6 kDa 13.3 kDa	AGHD	NCT02526420
Somavaratan (VRS-3017)	III	C	hGH–XTEN		PGHD	NCT02339090

Table 2 : XTEN-protein conjugates in clinical development. AA: amino acid; AGHD: adult growth hormone deficiency; C: completed; N: not yet recruiting; NCT: number clinical trial; PGDH: pediatric growth hormone deficiency; R: recruiting; rFVIIIFc-VWF: recombinant coagulation Factor VIII Fc-von Willebrand Factor; T: terminated.

XTENylation is also investigated in the field of anticancer therapeutics.[42] In particular, Amunix Pharmaceuticals develops XTEN-improved T-cell engagers to address some of their limitations in the treatment of solid tumors due to on-target off-tumor toxicity.[43] T-cell engagers are bispecific antibodies binding a target antigen on a tumor cell on one side and a CD3 on a T-cell on the other side to promote tumor cell apoptosis *via* the activation of the immune reaction. XTENylated Protease-Activated T-cell engagers (XPATs) can prolong the half-life and reduce off-target cytotoxicity induced by non-specific T cell activation (up to 15,000-fold). Once XTEN chains are cleaved by proteases in the tumor microenvironment, highly potent T-cell engagers are released. This strategy was used to target cancer cells-expressing HER2 or EGFR. Both

generated XPATs have a strong safety profile in cynomolgus monkeys (no cytokine release syndrome or systemic activation of T cells at high doses) and tumor regressions in murine tumor xenograft models.[42]

2.2.2. PAS

PASylation, introduced by XL-protein GmbH, is conceptually similar to XTENylation. However, it uses only three uncharged amino acids, namely, proline (P), alanine (A) and serine (S). The uncharged nature of PAS residues and their disordered conformation confers to PAS biophysical properties similar to those of PEG.[34a] PAS is biodegradable, hydrophilic, and reportedly non-immunogenic; its random coil conformation contributes to the expansion of the hydrodynamic volume, thereby increasing the serum half-life of the fused proteins.[34b, 44]

PAS was reported to be efficiently produced in bacterial systems as well as in eukaryotic cells with precise control over the composition of the sequence and its length. This latter spans from 100 to 1200 residues; PAS comprising 200 (18 kDa), 400 (35 kDa), and 600 (50 kDa) residues are however the most frequently used.[34a, 44] Pioneering work by Schlapschy and co-workers demonstrated the feasibility and usefulness of PASylation in improving the PK properties (half-life and area under the curve) of anti-HER2 Fab fragment (trastuzumab, 48 kDa), human interferon α2b (21 kDa) and hGH (22 kDa). Upon conjugation to PAS of 600 residues, the apparent molecular size of these proteins increased 22, 26, or 27-fold, respectively, translating into half-lives 21, 29, or 94-fold longer in mice compared with the unmodified proteins.[34a] Serum antibodies against the hGH or interferon moieties were detectable in mice. However, there was no cross-reactivity with unrelated proteins fused to PAS, indicating that the PAS polypeptide itself did not exhibit immunogenicity in these studies. Since the original publication by Schlapschy *et al*., the same group and others have applied this strategy to more than a dozen proteins. The list includes IFNβ superagonist YNSα8, IFN-β1b, IFNα, leptin and leptin antagonist, humanized anti-CD20 and anti-HER2 Fabs, coversin, erythropoietin (EPO), clotting factor VIII, uricase, exendin, IL-1Ra, Certolizumab Fab, anti-VEGFA nanobody, and Ankyrin Repeat Protein (DARPin) anti-EpCAM (epithelial cell adhesion molecule), FluoroCalins anti ED-B and VEGFR-3.[44-45] The increased hydrodynamic radius of the PASylated proteins improved their half-life by a factor of at least 10 in mice without compromising the biological activity of the fused proteins.[44] An increase in the biological activity (2-fold) was reported for PASylated IFN-β1b.[45b]

Most of the applications of PAS are for therapeutic proteins and are still in the preclinical stage. However, PASylation was shown to be valuable for *in vivo* imaging through extending the serum half-life of radiolabeled tracers allowing a better uptake in the tumor. For instance, PAS_{200} human CD98hcED-specific anticalin labeled with zirconium-89 was used for PET imaging of mice bearing-prostate cancer or B-cell lymphoma subcutanenous xenografts expressing the CD98 antigen.[46] Likewise, sensitive PET imaging of thyroid cancer was achieved using ^{89}Zr-Dfo-PAS_{200}-Gal3 Fab, a chimeric antigen-binding fragment directed against human Galectin-3 (Gal3 expressed in malignant thyroid nodules). The tracer accumulates selectively in the tumor-bearing thyroid lobe of xenograft mice giving strong contrast images 24 h post-injection.[47]

A first in-human study has been recently published for the imaging of HER2-positive metastatic breast cancer.[48] PAS-Fabs can be tailored to obtain a good compromise between the long half-life of full mAbs (which have major issues such as high toxicity due to slow clearance or low tumor penetration) and the short half-lives of Fab fragments limiting their accumulation in the tumor. Anti HER2-Fab was PASylated with PAS200 then radiolabeled with zirconium-89. PET imaging using ^{89}Zr-Dfo-PAS_{200}-HER2 Fab construct was thereby successful, well-tolerated, and represents a potential tool for diagnostic of HER2-positive breast cancer in patients.[48] The slower clearance from the blood allowed the accumulation of the tracer in both the primary tumor and metastases located in axillary lymph nodes at 24 h post-injection. However, the radiotracer was unsuccessful in detecting tumor metastases in the brain.

PASylation has also been exploited in the targeted delivery of small drugs and nucleic acids.[44] Flavo and co-workers have developed doxorubicin-loaded nanocage using PAS-modified ferritin protein nanocarrier (HFt-PAS).[49] Human ferritin (HFt) binds the transferrin receptor upregulated in many cancer cells. The modification of ferritin heavy chains with PAS of 40 or 75 residues (HFt-PAS_{40} and HFt-PAS_{75}, respectively) resulted in highly soluble and stable HFt-PAS nanocages with higher doxorubicin loading capacity (3 fold) compared with HFt. The half-life was also up to 5-fold longer in mouse blood compared with HFt and 56-fold longer compared with free doxorubicin.[49] The introduction of matrix metalloproteinase cleavable linker (MP) between the ferritin units and the PAS tags (HFt-MP-PAS) allowed targeting tumors more selectively by taking advantage of the high concentrations of matrix metalloproteinases in the tumor microenvironment.[50] This strategy was 4 and 8-times more efficient than

doxorubicin-albumin conjugate (Aldoxorubicin) and free doxorubicin in treating mice bearing xenogeneic PaCa-44 pancreatic tumor.[50] The authors inserted two glutamate residues in the PAS sequence to prevent nanocages aggregation by electrostatic repulsion of the negative surface charges. The new construct, termed HFt-MP-PASE, was shown to improve further the solubility and monodispersity of the nanocages.[51] Similarly, Tesarova *et al*. successfully encapsulated the cytostatic alkaloid ellipticine (Elli) in the cavity of PAS-modified ferritin with PAS of 10 residues (PAS_{10}-FRTElli). PAS_{10}-FRTElli exhibited a better accumulation in tumor tissue of mice bearing triple-negative breast cancer (MDA-MB-231) xenograft compared to free drug or FRTElli.[52] The higher accumulation of PAS_{10}-FRTElli in tumor tissue, likely due to the prolonged circulation time and EPR effect, was nonetheless no better than in free Elli or FRTElli in reducing the initial volume of the tumor.

XTEN and PAS of equivalent lengths (300, 600, 900 residues) were shown to increase the half-life of DARPin Ec1 (targeting epithelial cell adhesion molecule EpCAM) to the same extent (up to 114-fold).[53] DARPins (designed ankyrin repeats) are made of several 33-amino acid residue modules with alpha-helical structure engineered to bind a targeted protein with high specificity and affinity.[54] Interestingly, the charge difference between the two polymers had no effect on the biodistribution, clearance, or tumor accumulation of the fused protein in a xenograft model in mice. When cytotoxic maleimidocaproyl monomethyl auristatin F was conjugated, the largest DARPin (PAS 900 residues) did not have the highest anti-tumor response despite having the most prolonged half-life. The highest anti-tumor response was induced by intermediate size and half-life conjugates. Authors ascribed this interesting result to the balance between serum half-life and diffusion within the tumor.

2.2.3. ELP

Elastin-like polypeptide (ELP) is a polymer constituted from randomly repeated motifs of valine (V), proline (P), glycine (G), X, and G where X represents any amino acid except proline.[55] ELPs are good candidates for conjugation as they are biocompatible and biodegradable due to their high similarity with elastin.[55-56] An additional attractive property of ELPs is that they exhibit a thermally-responsive reversible phase transition: when the temperature is increased above the phase transition temperature, ELPs transition from a soluble state to a gel-like state. The size and composition of ELP sequence can be tuned to aim a transition temperature below the physiological temperature; thus, once injected, soluble

ELPs at room temperature form a subcutaneous depot at body temperature and slowly release the drug into the circulation.[57] Beside sustained-release properties of ELPs, the half-life of ELP-fused proteins is also extended by increasing their hydrodynamic radius.[57-58]

Conrad *et al*. successfully fused an anti-TNF nanobody (VHH) to ELP resulting in an active construct with a half-life 24-fold longer than that of the non-ELPylated anti-TNF VHH in mice after IV injection (28 min to 11.4 h).[59] Other ELPyated therapeutics are being developed primarily by PhaseBio (phasebio.com), taking advantage of the dual mechanism of the sustained release of subcutaneous depots and extended circulation time of ELP fused proteins. Three products are currently in clinical trials for weekly subcutaneous injections.[42, 58] Glymera™ (PB1023), an ELP-GLP-1 (phase IIb, NCT01658501, now licensed to ImmunoForge, Co. Ltd) and PE0139, an ELP-insulin (phase 2a, NCT02581657), both in type 2 diabetes patients. PB1046, an ELP-VIP (vasoactive intestinal peptide), is in phase II clinical trial for pulmonary arterial hypertension (Pemziviptadil, NCT03556020 and NCT03795428) and COVID-19 patients with acute respiratory distress syndrome (NCT04433546). However, this latter indication is no longer pursued.

3. Fusion of therapeutic proteins to serum proteins

The abundance of IgGs and albumin in blood and their long serum half-life make them ideal tools for engineering therapeutic protein constructs with extended circulation time. SA is the most abundant protein in blood with a concentration of 45 g/L. The blood concentration of IgGs reaches 10 g/L. While IgGs have a serum half-life of 21 days, SA has a half-life of 19 days. Therefore, IgGs and SA have been used to prolong the half-life of protein therapeutics in two strategies: (i) by directly fusing the therapeutic protein to a FcRn binding region of IgG or SA or (ii) by fusing the therapeutic protein to a molecule (*e.g*. a peptide, a protein or a fatty acid) able to non-covalently bind IgG or SA.

3.1. FcRn recycling

FcRn (neonatal Fc receptor or Brambell receptor) is a heterodimer receptor widely expressed in mammalian cells including the endothelium, intestinal and respiratory epithelia, and macrophages. It is derived from the major histocompatibility complex class I receptor and comprises a transmembrane α-chain of 45 kDa with a 17 kDa β-2 macroglobulin chain involved

in folding, transport and FcRn functions. FcRn extends the half-life of IgG and SA by protecting them from degradation. IgG is internalized in endothelial cells by pinocytosis and is then found in the recycling endosome. The binding between IgG and FcRn is dependent on the acidic pH (< 6.5) of endosomes. It involves the α2 and β2 domains of FcRn and pH-dependent salt bridges mediated by two histidine residues located between the CH2-CH3 of the Fc domain. At physiological or higher pH, FcRn does not interact with those ligands. (**Figure *1***). Acidity in vesicles allows the strong binding of IgG to the endosomal FcRn, protection from endosomal degradation and translocation of IgG back to the cell surface, where they are released at the neutral pH (7.4) of blood. SA is internalized through macropinocytosis, mostly in macrophages, bind the FcRn —at the opposite site that binds the Fc— by involving a histidine residue, then follow the same recycling pathway as IgG.[60] While FcRn binding prolongs the serum half-life of IgG and SA, the non-recycled molecules or antibodies, as for example IgE or IgA, are transported to the lysosome for degradation.

	IgG_1	**IgG_2**	**IgG_3**	**IgG_4**
Average MW (kDa)	146	146	165	146
Mean Adult serum level (g/L)	9	3	1	0.5
Relative abundance (%)	60	32	4	4
Half-life (Days)	21	20	7	21
C1q binding	++	+	+++	-
FcγRI (K_D, nM)	10	No aff.	10	1
FcRn (K_D, nM)	20	20	0	80

Table 3 : Main properties of human IgG subclasses. C1q: complement 1q system; FcγRI: Fc gamma receptor I; MW: molecular weight. No aff.: No affinity. Data from Murphy *et al*.[61]

It is interesting to note that the IgG subclass influences the interaction with FcRn. IgGs comprise four subclasses, IgG_1, IgG_2, IgG_3, and IgG_4 differing in their constant C_{H2} domain and binding properties (**Table 3**). The choice of the most appropriate subclass depends on the desired half-life since the C_{H2} region is involved in binding to IgG-Fc receptors (**Erreur ! Source du renvoi introuvable.**).[62] While the subclasses IgG_1, IgG_2 and IgG_4 possess a high binding affinity for FcRn, the IgG_3 isotype lacks a functional FcRn binding domain resulting in a shorter

half-life compared to the other isotypes (**Erreur ! Source du renvoi introuvable.**).[62] The non-FcRn recycling IgG_3 can represent an asset for therapies involving mAbs as carriers for cytotoxic payloads —due to their short half-life limiting hematologic off-target toxicity.

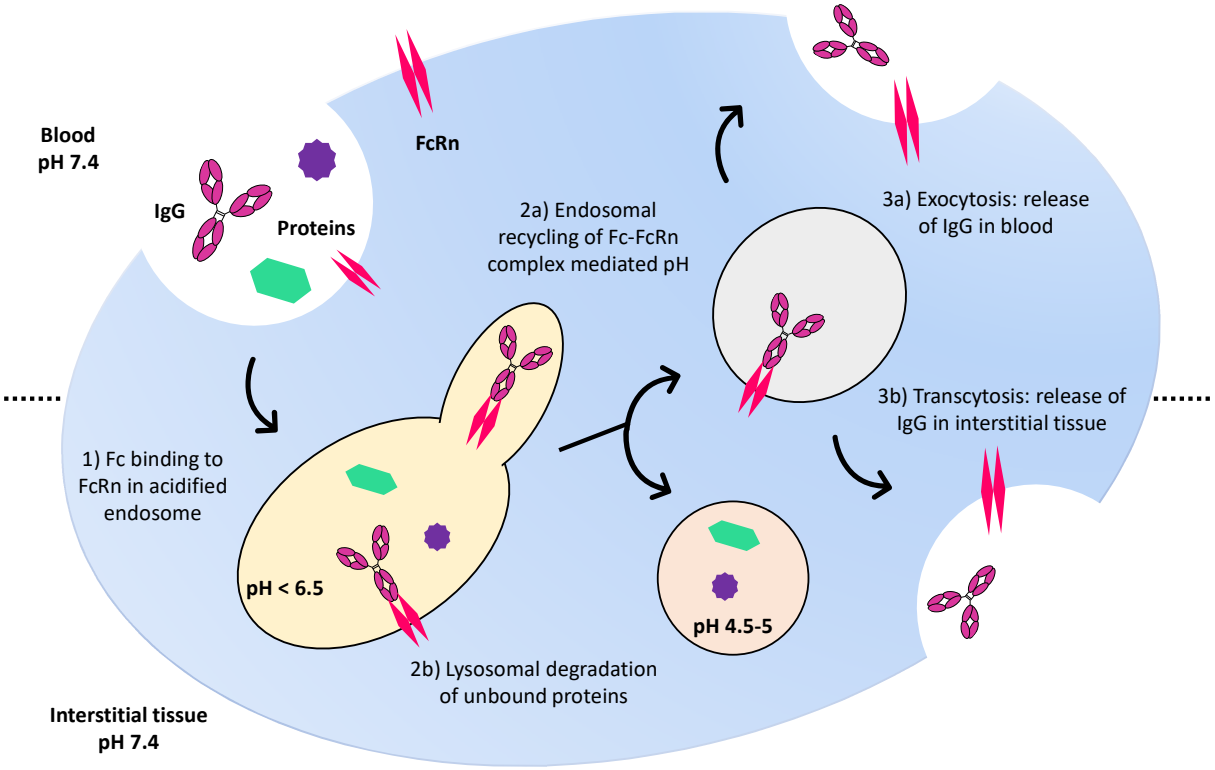


Figure 1: Recycling-mediated FcRn pathway after pinocytosis of serum IgG. IgGs are first internalized into cells *via* endocytosis. Acidity in vesicles allows the binding to FcRn while unbound plasma proteins undergo lysosomal degradation. The IgG bound FcRn is then translocated back to the cell surface. Due to neutral pH, the complex dissociates and the IgG is released in the blood after exocytosis or in interstitial tissue *via* transcytosis. The FcRn is then free to be involved in another cycle. Adapted from reference.[63]

Protein engineering strategies have been developed to increase the affinity of human IgG1 for FcRn in order to further increase its recycling and thus to obtain therapeutic antibodies with an even longer half-life.[64] MedImmune technology consists in inserting mutations in the CH_2 domain of the Fc region of IgG_1, for instance, a triple substitution (M252Y, S254T and T256E), referred as YTE (Patent US7658921B2). Proof-of-concept on palivizumab showed an *in vitro* 10-fold increase in binding affinity to human FcRn at low pH with an efficient release at pH 7.4.[65] Further experiments with palivizumab and several YTE variants of IgGs exhibited an

increased plasma half-life for up to three months.[66] Recently, YTE introduced in neutralizing HIV mAbs has been assessed in macaques infected with HIV. While the YTE sequence improved the PK profile following IV administration, it also induced immunogenicity, an issue limiting its clinical transfer for now.

One last characteristic of the Fc fragment is its immunoregulatory properties. IgG or Fc fusion proteins bound to the FcRn are diverted from antigen presentation compartments, obviating an immune response. The Fc is also able to enhance the suppressive activity of T-cells (Treg) involved in the maintenance of immunologic self-tolerance.[67] Two T-cells epitopes have been identified in the Fc region of IgG_1 that are capable of regulating immunogenicity by activating Tregs. Consequently, in addition to extending protein half-life, the Fc-fusion strategy also represents a great potential to reduce immunogenicity.

3.2. Fc fusion proteins

In most cases, Fc fusion proteins present an extended half-life compared with their parent protein or peptide. Most of the Fc-fusion proteins are produced by genetic engineering via the fusion of the C-terminus of the biological payload to the N-terminus of the IgG-CH3 domain resulting in a stable Fc-conjugate.[58] Currently, four groups of proteins or peptides are used for Fc-fusion partners, namely the extracellular domains of natural receptors (*e.g.*, etanercept, alefacept, etc.), cytokines (*e.g.*, aflibercept), peptides (*e.g.*, romiplostim, dulaglutide) and enzymes (*e.g.*, asfotase α, turoctocog α, etc.) as shown in **Table 4**. Etanercept® is the first Fc-fusion protein, approved in 1998, for the treatment of rheumatoid arthritis. Etanercept® is constructed by two 75 kDa-human TNFRII exodomain each linked to the Fc of a human IgG_1 (**Table 4**).[68] Due to the fusion to the Fc domain, etanercept® is eliminated slowly in patients, with a half-life of 70-100 h. It binds TNF-α and TNF-β with higher affinity than the endogenous TNF receptor, thereby preventing the proinflammatory cascade.[68]

Another good example of Fc-fusion protein is the well-established standard treatment for patients with hemophilia A: the factor VIII (FVIII).[69] Several strategies have been employed to extend the half-life of rFVIII products, including PEGylation, XTENylation and Fc-Fusion.[70] Recombinant FVIII Fc fusion protein (rFVIIIFc), constructed by fusion of a single molecule of rFVIII to the Fc region of human IgG1, has been the first FVIII approved with an extended half-life (Eloctate®, **Table 4**).[41c, 71] However, compared with the conventional rFVIII, the fused

rFVIIIFc has a half-life only 1.5-fold longer (19.0 h *vs* 12.4 h) and a slightly slower systemic clearance (2.0 *vs* 3.0 mL/h/kg).[72]

FDA approval	Generic name (US trade name)	MW (kDa)	Target	Protein format	Half-life (days)	Mechanism of action	Main application
1998	Etanercept (Enbrel®)	150	TNF-α	P75 TNFR exodomain IgG1 Fc fusion	2.8	Blocks TNF-α/TNFR interaction	Rheumatoid arthritis
2003	Alefacept (Amevive®)	92	CD2	CD58 (LFA-3) IgG1 Fc fusion	11	Blocks interaction of CD2 with LFA Inhibits T-cell activation	Moderate-severe psoriasis
2005	Abatacept (Orencia®)	92	CD80/CD86	CTLA-4 IgG1 Fc fusion	12-23	Blocks T-cell activation and cytokine production	Rheumatoid arthritis
2008	Rilonacept (Arcalyst®)	251	IL1A, IL1B, IL1RN	IL-1R1 & IL-1RAcP IgG1 Fc fusion	8.6	Blocks IL-1β signaling, reduces inflammation	Cryopyrin-associated periodic syndrome
2008	Romiplostim (Nplate®)	60	Thrombopoietin R (agonist)	Peptibody IgG 1 Fc fusion	3.5	Stimulates JAK2 et STAT5 pathways	Chronic immune thrombocytopenia
2011	Belatacept (Nulojix®)	90	CD80/CD86	CTLA-4 IgG1 Fc fusion	~ 8-10	Blocks T-cell activation and cytokine production	Prophylaxis after kidney transplant
2011	Aflibercept (Eylea®)	115	VEGF-A	VEGFR 1 & 2 IgG1 Fc fusion	~ 5-7	Inhibits angiogenesis	Age-related macular degeneration
2012	Ziv-Aflibercept (Zaltrap®)	115	VEGF-A	VEGFR 1 & 2 IgG1 Fc fusion	~ 5-7	Inhibits angiogenesis	Metastatic colorectal cancer
2013	Turoctocog alfa (NovoEight®)	166	Factor substitute	Truncated rhfactor VIII IgG1 Fc fusion	~ 0.4	Replaces factor VIII deficit	Hemophilia A

FDA approval	Generic name (US trade name)	MW (kDa)	Target	Protein format	Half-life (days)	Mechanism of action	Main applications
2014	Eftrenonacog alfa (Alprolix®)	98	Factor substitute	Monomeric Factor IX IgG1 Fc fusion	~ 3	Replaces factor IX	Hemophilia B
2014	Efmoroctocog alfa (Eloctate®)	170	Factor substitute	Monomeric factor VIII IgG1 Fc fusion	0.8	Replaces factor VIII	Hemophilia A
2014	Dulaglutide (Trulicity®)	60	GLP-1R (agonist)	GLP-1 analog IgG4 Fc fusion	3.75	Stimulates insulin production	Type 2 diabetes
2015	Asfotase alfa (Strensiq®)	180	Factor substitute	Human TNS-ALP IgG1 Fc fusion	~ 5	Replaces deficient ALP	Hyphosphatasia
2016	Etanercept-szzs (Erelzi®)	125	TNF-α	Dimeric TNFR IgG1 Fc fusion	2.8	Blocks TNF-α/TNFR interaction	Rheumatoid arthritis
2018	Damoctocog alfa pegol (Jivi®)	230	Factor substitute	Monomeric factor VIII-PEG IgG1 Fc Fusion	0.8	Replaces factor VIII	Hemophilia A
2019	Turoctocog alfa pegol (Esperoct®)	206	Factor substitute	Truncated factor VIII-PEG IgG1 Fc Fusion	0.8	Replaces factor VIII	Hemophilia A
2019	Luspatercept-aamt (Reblozyl®)	76	TGF-β	Activin RIIB IgG1 Fc fusion	11	Blocks TGF-β signaling	Beta thalassemia anemia

Table 4: Chronological overview of FDA-approved chimeric Fc-fusion proteins: structural characteristics, mechanisms of action and main applications. ALP: alkaline phosphatase; CTLA-4: cytotoxic T lymphocyte-associated antigen 4; FDA: Food and Drug Administration; GLP: glucagon-like protein; IgG: immunoglobulin G; IL: interleukin; JAK2: Janus kinase 2; LFA: lymphocyte function associated antigen; MW: molecular

weight; PEG: polyethylene glycol; R: receptor; STAT5: signal transducer and activator of transcription 5; TGF: transforming growth factor; TNF: tumor necrosis factor; TNS: tissue non-specific; VEGF: vascular endothelial growth factor.

3.3. Targeting serum albumin

Serum albumin, the most abundant protein in blood, is characterized by a long serum half-life and a broad tissue distribution. The long serum half-life of SA is due to: (i) its large size (*i.e.*, 67 kDa, 585 amino acids spanning three independently folding domains) which is higher than the cut-off for kidney filtration, and (ii) more importantly to its protection from intracellular lysosomal degradation via its binding to the FcRn (**Figure 1**). The binding of the therapeutic protein to SA (directly or indirectly) may further protect it by shielding it from proteolytic degradation. Along with this, SA exhibits a high capacity to extravasate from the bloodstream to reach the lymphatic system and thus it accumulates in cancerous or inflamed areas.[73] SA therefore constitutes a target of choice to which therapeutic proteins have been paired through mainly two different approaches in order to increase their serum half-life.

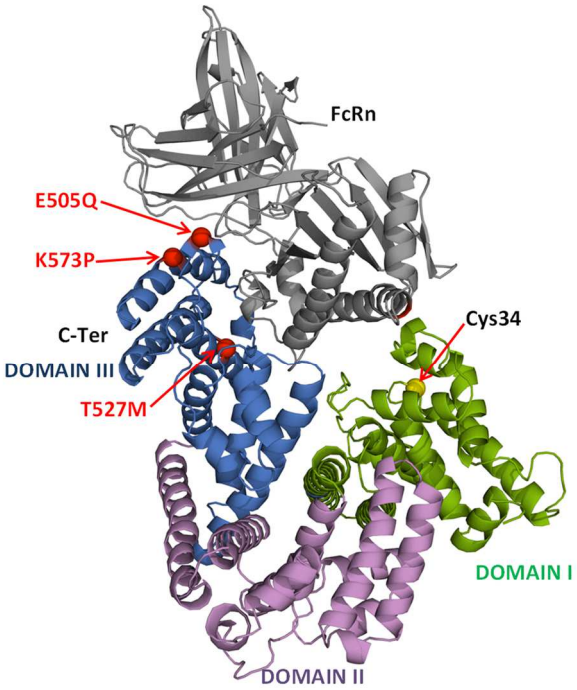


Figure 2 **:** Structure of SA in complex with FcRn. The three structural domains are highlighted as well as the position of the mutations discussed in the text and the free Cys34 (PDB 4K71).

3.3.1. Covalent conjugation of therapeutic peptides and proteins to serum albumin

In a first approach, the therapeutic peptide or protein is covalently linked to SA. The advantages of this approach, also referred to as albumination, are that SA, which is one of the best characterized proteins in the pharmaceutical field, is not immunogenic and has an excellent biocompatibility and degradability.[74] Moreover, this non-glycosylated protein, can be produced in large amounts in eukaryotic cells such as yeast or mammalian cells, alone or fused to therapeutic proteins.[75] SA was approved by the FDA as a therapeutic protein in 1982 which reduces the regulation considerations for the development of new SA-based therapeutics.

Two albuminated proteins have entered the market. In the first one, Albiglutide used for treatment of type 2 diabetes, SA is fused to GLP-1; while in the second, Albutrepenoncogalpha used to treat haemophilia, it is fused to recombinant coagulation factor IX. These fusion proteins are injected subcutaneously or intravenously, once weekly or up to once every other week, respectively. At least three other SA-fusion proteins have entered clinical trials (**Table 5**).

A number of factors affect the magnitude of half-life extension conferred by the fusion to SA including the affinity of SA for FcRn at neutral and acidic pHs, the size and nature of the therapeutic protein, and the site of attachment of the therapeutic protein on SA. In any case, the effects of the conjugation/fusion to SA on the binding, the stability, the therapeutic effects, the biodistribution and pharmacokinetics of a given therapeutic protein have to be evaluated to choose the best configuration.

In order to study the influence of the affinity of SA to FcRn, a number of mutational variants of SA with modified affinity for FcRn have been engineered.[76] For example, the replacement of Lys-573 with any amino acid resulted in enhanced binding to FcRn at acidic pH while minimally affecting the binding at neutral pH. In particular, the affinity of the variant K573P for human FcRn is more than 12-fold that of the wild type (WT) SA (10.3 versus 125.6 nM), resulting in extended serum half-life in WT-mice, mice transgenic for human FcRn, and cynomolgus monkeys (5.4 to 8.8 d)[76a]. Very recently, the same research group has generated a triple mutant, E505Q/ T527M/K573P (QMP-SA) of SA which can be expressed in similar amount than the WT-SA while its affinity for human FcRn is increased by about 160 fold. In mice, the half-time of QMP-SA fused to Factor VIIa (FVIIa) and administrated intravenously, is

almost 4-fold longer compared with the WT-SA fusion, without compromising the therapeutic properties of FVIIa. This enhanced efficiency can be rationalized by the fact that, compared to WT-SA fusion proteins, the higher affinity of the mutational variants of SA give them a competitive advantage over the endogenous SA for FcRn binding.[76b] Moreover, since the attachment of large cargo to SA often reduces the affinity of SA for FcRn, the use of engineered SA variants with improved affinity allows maintaining an affinity of the fusion protein above that of unmodified WT-SA. [77] The availability of a series of mutational variants of SA with different affinities constitutes an opportunity for optimizing the drug efficiency, tolerability and dosing by finely tuning the serum half-life of the therapeutic protein.[76a]

Trade name	Generic name	Parent drug	Position on SA	Molecular weight (kDa)	Application	Status
Albuferon®/Zalbin®/Joulefon®	Albinterferon	Interpheron alpha2	N-terminus	85.7	Chronic hepatitis C	NCT00964665*
Eperzan®/Tanzeum®	Albiglutide	GLP-1	C-terminus	72.9	Diabetes mellitus type II	FDA approval in 2014
Neugranin ®, Egranli®	Balugrastim	hG-CSF	N-terminus	85	Chemotherapy-induced neutropenia	Positive evaluation by EMA in 2014[#]
Idelvion®	Albutrepenonacog alfa	Coagulation factor IX	C-terminus	125	Hemophilia B	FDA approvai in 2016
CSL689	NA	Recombinant factor VIIa	C-terminus	120	Hemophilia A or B	NCT02484638

Table 5: Genetically-engineered serum albumin fusion proteins in the market or clinical trials. EMA: European Medicines Agency; FDA: Food and Drug Administration; GLP-1R: Glucagon-like peptide 1. hGCSF: human granulocyte colony stimulating factor. SA: serum albumin; NA: not available. *This phase 2 clinical trial was halted in 2010 due to severe adverse effects. [#]This application for approval has been withdrawn after phase 3 for commercial reasons.

The large size of SA may shield the therapeutic protein fused to it. Although the shielding can be beneficial to protect the therapeutic peptide or protein from proteolysis, it may also have detrimental effects on its functional properties. In order to reduce this eventual shielding effects, therapeutic proteins can be fused to the SA domain III (23 kDa), which is both necessary and sufficient for FcRn binding in a pH-dependent manner (Figure 2). Serum albumin domain III and mutational variants thereof have been fused for exemple to an ScFv and resulted in an improved half-life (i.e., up to 56.7 h versus 2.9 for the scFv alone).[78]. Reducing the size of the SA moiety was also shown to ensure a better tumor accumulation of the therapeutic protein.[78]

SA or its derivatives (i.e., mutational variants) can be conjugated to a therapeutic protein via different approaches. First, it can be site-selectively chemically coupled *via* its free cysteine residue at position 34 located in domain I and distant from the FcRn interface (**Figure 2**), *via* maleimide coupling. Recently, Bak et al, first labeled the cysteine 34 with a DBCO function that was further conjugated to a therapeutic peptide equipped with a clickable non-natural amino-acid-p-azido-1-phenyalanine (AzF) through strain-promoted azide-alkyne cycloaddition (SPAAC).[79] This approach was used to conjugate GLP-1 to SA at three different site-specific positions. Although, the half-life, in mice, of the three conjugated peptides was similar (i.e., 8 h compared to 3 min for the non-conjugated peptide), the potency of the peptide significantly depended on the site of conjugation.[79] We anticipate that in the future, further development of site-specific-bioorthogonal labelling will enable to increase the potency of albuminated pharmaceuticals. Finally, SA and its derivatives can be genetically fused to the therapeutic protein either at the C-terminus, N-terminus or both; and the chimeric proteins are expressed in the suitable host as a single polypeptide.[80]

3.3.2. Non-covalent binding of therapeutic proteins to serum albumin

The second approach consists in fusing or conjugating the therapeutic protein to molecules which bind to endogenous SA (**Table 6**). These later can be molecules that naturally bind SA (e.g., fatty acids or bacterial albumin domains), and proteins specifically engineered to bind SA.

3.4.2.1 Natural binders of serum albumin

Fatty acids. SA acts as a transporter of fatty acids (7 binding sites for long fatty acids and 2 for medium-size fatty acids). Thus, the conjugation of peptide or protein to fatty acids, referred to as lipidation, allows to extend their blood half-life and a series of lipidated peptides and proteins are on the market (**Table 7**). For example, the serum half-life of an insulin analog (desB30 human insulin) is increased from 4–6 min to 5–7 h by conjugating it to myristic acid (C14) through the Nε-amine of LysB29. This lipidated insulin, called Detemir, was approved in 2004 and the extension of half-life makes it suitable for a once-daily subcutaneous injection.[81] The prolongation of the blood half-life is due to a combination of two phenomena: the interaction of the C14 moiety with the fatty acid binding site on albumin and the prolongation of absorption via the oligomerization of the lipidated insulin. Indeed, the myristic acid is thought to stabilize both a hexamer−dihexamer equilibrium and hexamer−albumin complexes in the subcutis. Such complexes are likely to protract insulin absorption into the bloodstream. Then, upon dissociation and absorption into the bloodstream, the insulin Detemir monomers can bind to albumin through their fatty acids; more than 95% of circulating insulin Detemir is indeed albumin bound. Other examples of lipidated biopharmaceuticals are Liraglutide and Semaglitude. Liraglutide is a GPL-1 analog in which the lysine 34 has been mutated to an arginine, allowing it site-specific conjugation to palmitic acid through the Nε-amine of Lys 26 via a γGlu spacer. Liraglutide, which was approved in 2010, has a half-life of 8-10 h and 13-15 h following IV and SC injection, respectively. Such half-life extension makes it suitable for once daily administration.[82] Semaglutide consists in GPL-1 with two amino acid substitutions at positions 8 and 34, where alanine and lysine are replaced by 2-aminoisobutyric acid and arginine, respectively. It is conjugated to a C-18 fatty diacid (stearic acid) on the lysine 26. Its serum half-life is about 7 days and once-weekly injection is therefore enough.[82] While the first three commercialized lipidated biopharmaceuticals are peptides or small proteins, Somapacitan which has been recently approved by the FDA is derived from a larger protein, hGH (22 kDa, 191 aa). The leucine at position 101 of hGH is mutated to a cysteine residue that is used to conjugate a C16 fatty acid through a tetrazole linker by alkylation. Somapacitan long half-life allows a weekly injection for the treatment of adults with hGH deficiency.[83]

The advantages of lipidation are that the fatty acids are cheap to synthetize, biocompatible and non-immunogenic. However, the main drawbacks are the insolubility of fatty acids and

their lower affinity for SA than for example antibodies and fragments thereof.[84] Moreover, the linker might be immunogenic. An immunogenic response against lipidated biopharmaceuticals has indeed been reported. The levels of antibodies generated were however low and without clinical relevance.[81]

A number of factors affect the magnitude of half-life extension conferred by the conjugation to fatty acids including the length of the fatty acid and of the linker used to conjugate it to the protein and the size of the therapeutic protein. For lipidated GPL-1, a clear positive correlation between the length of fatty acids and the affinity for SA was observed; however, this was associated with a decreased potency of GPL-1 probably because only free GLP-1 can bind to the receptor. Therefore, a compromise should be made between these two parameters.[85]

Fatty acids can be conjugated to peptides and proteins via the lysines. However, this usually results in heterogenous labelling and loss of therapeutic efficacy. Site-specific conjugation via copper-catalyzed azide-alkyne cycloaddition (CuAAC) and strain-promoted azide-alkalyne cycloaddition (SPAAC) has been reported.[86] For example, in the latter case, a clickable non-natural amino acid, p-azido-L-phenylalanine (AzF) is introduced to a specific site of the target peptide or protein; then a fatty acid analog containing dibenzoclyclootyne group (DBCO-FA) is conjugated to AzF site of the target peptide/protein via SPAAC.[87] Such a protocol was used to site-specifically conjugate urate oxidase, a therapeutic protein for the treatment of tumour lysis syndrome, to palmitic acid at two positions. The conjugation resulted in high SA binding capacity and retained enzyme activity. Fu et al have developed a strategy, through the genetic encoding of ε-N-heptanoyl-l-lysine (HepoK), allowing the introduction a fatty-acid-containing amino acid into proteins with exquisite site-specificity and homogeneity. Using this approach, they produce a HepoK-incorporated GLP1 in *E. coli* which showed a more potent and long-lasting ability in decreasing blood glucose level in mice than WT GLP1.[88]

Although lipidation is effective in prolonging the blood half-life of peptides or small proteins (MW< 28kDa), it is less efficient to increase the half-life of larger proteins. Such proteins, when conjugated to fatty acids, are thought to compete with the binding to SA with FcRn due to the fact that the dominant fatty acid-binding sites partially overlap with the FcRn binding site.[89] For such large proteins, it was recently shown that increasing the linker length between the fatty acid and the target protein reduces the steric hindrance for the binding of FcRn to SA and results in longer serum half-life. For example, there is a linear correlation between the linker

length (from 0.25 nm up to 2.8 nm) and the serum half-life of urate oxidase (140kDa), a therapeutic protein used to treat hyperuricemia, conjugated to palmitic acid. The longer linker led to a 7-fold greater extension of serum half-life in mice.[89] Such a correlation was not observed for small proteins.[84]

Design	Molecule engaged	Advantages	Drawbacks	Site-specific conjugation
Genetic or chemical fusion to albumin	Full-length albumin (and mutational variants thereof)	Tunable effects via mutations Reduced regulatory considerations (non-immunogenic, biodegradable) Easy production	Shieding of the therapeutic protein	Yes
	Albumin domains	Reduced size Tunable effects via mutations Lower shieding effect Better tissue penetration Easy production		Yes
Non-covalent binding to albumin	Fatty acids	Low cost, biocompatibility	Poor solubility Lower affinity for SA Immunogenicity of the linker	Yes
	Bacterial ABD	Small size Tunable PK Easy production, high stability	Immunogenicity	Yes
	Antibody fragments (Fab, scFv, Fv, VH, VL, nanobody, & VNAR)	Small size Tunable PK Easy production	Immunogenicity if not from human Abs	Yes
	Artificial proteins (DARPIn & Aptide)	Small Size Tunable PK Easy production	Immunogenicity	Yes

Table 6 : Comparison of the different approaches to extend the serum half-life of proteins *via* targeting serum albumin (SA).

Bacterial serum albumin binding domains (ABD). The second class of SA naturally binding molecules are the bacterial proteins targeting SA including *Staphylococcus* protein A and *Streptococcus* protein G.[90] The SA binding domains of these proteins, composed of about 50 amino acids (~5 kDa) have been extensively engineered to further improve their half-life extension capability by a combination of combinatorial protein engineering, *in vitro* selection *via* phage display technology and rational design, leading to the selection of fentomolar affinity binders or of minimal size binders.[91] For example, Guo *et al*. have fused an ABD to a human epidermal growth factor receptor 2 (HER2)-specific immunotoxin ZHER2-PE38.[92] Compared with non-fused ZHER2-PE38, this new construct exhibited a clearly increased serum half-life (331 versus 13 min, approximately 25-fold extension) and remarkably improved antitumor effects in an NCI-N87 subcutaneous xenograft model. The Albumod™ platform developed by Affibody AB to improve the PK of biologics is based on such albumin binding domains (ABD). Izokibep (also known as IMG-020 or ABY-035) is a bispecific fusion protein made of two affibodies (i.e., 6.5 kDa artificial proteins derived from the Z domain of staphylococcal Protein A and structured as a triple α-helix bundle) with a high affinity for interleukin-17A (IL-17A) and one ABD domain (5kDa) with high affinity for SA.[93] It is in clinical trial to treat patients with ankylosing spondylitis (Table 7).

The advantages of ABD are their small size, high stability, easiness to be engineered and to be produced recombinantly. Their drawback is their potential immunogenicity since they derive from bacterial sources.

3.4.2.2 Non-natural binders of serum albumin

Finally, a number of proteins specifically binding SA have been generated. This includes various antibody fragments as well as artificial proteins (i.e., aptides and DARPins).[94]

Antibody fragments. A series of antibody fragments targeting SA have been generated including Fab, Fv, scFv, VH, VL derived from conventional IgG essentially from human, murine or rabbit origin, and VNAR and nanobodies which are derived from heavy-chain only antibodies.[14, 95] These later two are produced by sharks and camelids, respectively; they are devoid of light chains. Their binding site is therefore constituted by a single IgG domain referred to as VHH or nanobody when derived from camelid and VNAR when derived from

sharks.[96] Nanobodies are more and more used in various fields including human medicine imaging and therapy.[97] Despite their small size, nanobodies bind to their antigen with a high affinity. Due to their small size, they have a number of unique favorable properties including high stability, high solubility, easiness to be further engineered to adapt their properties to a given application and to create multi-domain constructs, easiness to produce and store, low immunogenicity due to the high sequence identity (~80%) with the human VH3 (variable domain of the heavy-chain of conventional antibodies) gene family and easiness to be humanized if necessary, ability to target unusual epitopes and capacity to work intracellularly.[96, 98] In February 2019, Caplacizumab, the first nanobody-derived drug, was approved by the FDA for acquired thrombotic thrombocytopenic purpura, a rare disease characterized by excessive blood clotting in small blood vessels. It consists in two identical nanobodies targeting the A1 domain of von Willebrand factor, linked by a linker made of three alanines.

A number of anti-SA nanobodies have been described; they extend the serum half-life of the protein they are fused to up to 376-fold in preclinical models.[99] Five nanobodies-based biopharmaceuticals involving a SA-binding moiety are in clinical trials (Table 7): Vobrarilizumab and Ozoralizumab to treat rheumatoid arthritis, M6495 to treat osteoarthritis, BI5508 to treat atherosclerosis and Sonelokimab to treat psoriasis. Vobrarilizumab consists in a bispecific nanobody targeting respectively IL6 and SA, and Ozoralizumab is a trivalent bispecific nanobody with two nanobodies targeting TNF-α and one targeting SA.[100] Sonelokimab (also known as M1095) is a trivalent nanobody; it is made of nanobodies specific to human IL-17A, IL-17F, and SA. M6495 is a bispecifc nanobody made of one anti-ADAMS5 protease (A Disintegrin And Metalloprotease with ThromboSpondin-motifs-5) nanobody and one nanobody anti-SA. In vitro, M6495 completely inhibits ADAMS5 which is involved in arthritic diseases. In an 8-week murine DMM (destabilization of the medial meniscus) model, it slowed progression of joint damage when administered prophylactically. Finally a bispecific nanobody binding to both SA and the chemokine receptor CX3CR1, referred to BI65088, is a potent antagonist to CX3CR1 that significantly inhibits plaque progression in a murine model of atherosclerosis.[101] After IV in cynomolgus monkey, its blood half life was 9.2 days. To reduce immunogenicity the sequence of the nanobodies have been humanized.[102]

Most of the anti-SA antibody fragments have been derived from immune librairies made from the blood of animals, essentially rabbits, mice and lamas immunized with the target protein.[14, 95a] Binders specific of SA can then be selected from these libraries by a panning strategy such as phage display. Inclusion of endogenous SA-binders during this procedure favors the selection of binders that will not interfere with the function of SA including the binding to FnRc. Moreover, the selection of binders that cross react with SA from different origins (e.g., mouse, rat, monkey and human) can be carried out by alternatively using one of these proteins as target in the consecutive rounds of selection.[95a] A combination of one round of phage display panning and next-generation DNA sequencing has also been used to identify cross-reactive nanobodies against SA.[95a]

One critical point with such non-human anti-SA antibody fragments is their potential immunogenicity. Thus, they should be humanized and this is usually carried out by CDR grafting onto human VH and VL frameworks for conventional antibody fragments.[14] Given their single-domain character, nanobodies can be easily humanized by mutating key residues in the sequence of their CDRs and various strategies to efficiently humanise nanobodies without significantly affecting their binding specificity, stability and solubility have been established.[95a, 103]

Note that an alternative strategy to increase the serum half-life of nanobodies, due to the easiness with which they can be used as building blocks is to fuse them to a subunit of a multimeric protein. Fan et al. have, for example, generated a platform they named fenobody, in which a nanobody developed against H5N1 virus is displayed on the surface of ferritin in the form of a 24mer. This overall affinity of the fenobody for H5N1 was drastically increased (i.e., by a factor 360) and its serum half-life in a murine model was extended by a factor 10 compared to the monovalent nanobody counterpart.[104]

The AlbudAb™ platform which is based on a drug fusion with a variable heavy or light chain domain derived from a human IgG that exhibits high albumin affinity was developed by GlaxoSmithKline. GSK2374697, a genetically engineered fusion protein of such a human domain antibody fragment to exendin-4 acts as a long-duration GLP-1 receptor agonist for the treatment of type 2 diabetes. The pharmacokinetic profile was prolonged, with estimated half-lives ranging from 6 to 10 days in humans (versus 2.5 h for exendin-4 alone).[105]

Artificial proteins. Several artificial proteins designed by combinatorial protein engineering have been selected against SA including DARPins and Aptides. MP0250 is a multi-domain consisting in four DARPin domains with the following successive specificities within a single polypeptide chain: SA, vascular endothelial factor-A, hepatocyte growth factor and SA.[106] This multidomain protein is in clinical phase 2 for the treatment of patients with solid tumors. MP0250 specifically inhibits both vascular endothelial growth factor (VEGF) and hepatocyte growth factor (HGF) with the aim of disrupting the tumour microenvironment in patients with solid tumours. Its serum half-life is about 2 weeks.[107] Thus, a dosing interval of 2, 3 or even 4 weeks can be envisioned. Aptides are structure-constrained peptides containing a randomizable binding region and a constant β-hairpin scaffold. An anti SA aptide with KD of 188 nM was isolated by phage and fused to exenatide. The serum half-life of the fusion protein was 4-fold longer compared with exenatide alone.[94e] The advantages and drawbacks of these proteins are very similar to those described above for nanobodies. Indeed, the advantages of artificial proteins are their small size, high stability, high yield of production, use as building block of multivalent/multifunctional constructs, ability to work intracellularly. Moreover, due to their high stability, artificial proteins can support the introduction of non-natural amino-acids in order to further increase their resistance to proteolysis or for their functionalization.[108] The essential drawback is the immunogenicity; this later can be evaluated and reduced *in silico* using for example Lonza's proprietary Epibase® *in silico* platform.

Trade name	Generic name	Parent drug	Molecular weight (kDa)	Conjugation reaction	Application	Status
Levemir®	Insulin detemir	Human insulin	5.9	Myristic acid on lysine at position B29	Diabetes mellitus I/II	FDA in 2005
Tresiba®	Insulin degludec	Human-insulin	5.9	Hexadecanedioic acid on lysine at position B29	Diabetes mellitus I/II	FDA in 2015
Victoza®/Saxenda®	Liraglutide	GPLP-1R (L27R)	3.7	Palmitic acid on lysine at position26	Type II diabetes	FDA in 2010
Ozempic®/Rybelsus®	Semaglutide	GPLP-1R (A8aminoisobutyric acid, L34R)	3.7	Stearic diacid (C18) on lysine at position26	Type II diabetes	FDA in 2017
Sogroya ®	Somapacitan	Human growth hormone (hGH) (L101C	23.3	C16 on position C101	Adults with growth hormone deficiency	FDA in 2020
NA	GSK2374697 (AlbudAb)	Exendin-4	16	Fusion of antibody albumin binding domain to exendin 4	Obesity	NCT02829307

Izokibep	**ABY-035 or IMG-020**		Two affibodies anti-II-17 and one ABD	Multiple autoimmune diseases	NCT04713072
NA	Vobrarilizumab	26	Bispecific nanobody anti-Il6 and anti SA domain	RA	NCT02101073 [109]
NA	Ozoralizumab (ATN103)	45	Trivalent bispecific nanobody anti-TNF- α and anti SA domain	RA	NCT01007175 NCT04077567
NA	Sonelokimab (M1095)		Trispecifc nanobody anti-human interleukin (IL)-17A, IL-17F, and anti-SA	Plaque psoriasis	NCT03384745
NA	M6495	28.1	Bisepcific nanobody anti-ADAMTS-5 and anti-SA	Osteoarthritis	NCT03224702 NCT03583346
NA	BI655088		Bispecific nanobdy anti chemokine receptor CX3CR1 and anti-SA	atherosclerosis	NCT02696616

NA	MP250	Tri specific DARPin, anti-VEGF, ani-HGF and anti-SA	Multiple myeloma	NCT03418532 NCT03136653

Table 7: Anti-serum albumin fusion proteins in clinical trials or in the market. FDA: Food and Drug Administration; hrPCA: Hormone refractory prostate cancer; GLP1R: Glucagon-like peptide 1 receptor. NA: not available. RA: Rheumatoid arthritis

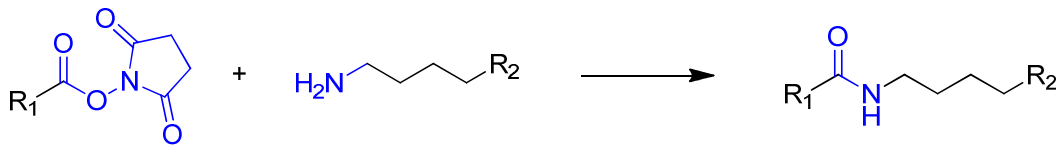
4. Methods for protein modification

The emergence of protein engineering in the last half century has led to several improvements in the design of attached groups and the approaches of conjugation. Covalent ligations can be performed on different accessible amino acid residues located on the surface of proteins, allowing a stable conjugation of several kinds of therapeutic payloads. In this section the different chemical and enzymatic reactions used either for a random or, site-specific conjugation of proteins are described. Bioconjugation strategies presented below are not limited to applications for extending therapeutic protein half-life.

4.1. Chemical reactions

4.1.1. NHS ester ligation

Random coupling to aliphatic lysine residues (*via* carbamate, urethane of amide linkage) has been the gold standard for protein conjugation since decades thanks to its ease and high yield in organic synthesis (**Table *8***).[110] Lysines are highly prevalent in proteins (*i.e.*, they represent from 6 to 10% of the sequence) and they stand as one of the most reactive amino-acid towards several reagents.[111] Activated carboxylic acid esters, such as *N*-hydroxysuccinimide (NHS), placed at one of the termini of the entity to attach, react in basic medium (pH: 7.9-8.5) with Ɛ-amine groups of lysine to form a peptide bond (**Table *8***).



Scheme 1 : Reaction between N-hydroxysuccinimide (NHS) ester bearing moiety and the aliphatic NH_2 function of a protein lysine residue.

While the NHS ester is predominantly used to perform conjugation on lysines, it can also react with the imidazole group of histidines or the hydroxyl group of tyrosines depending on the pH and the temperature of the reaction. If fixed ratios of NHS to proteins lead to a quite reproducible grafting in the average number of conjugated moieties, their distribution is widely dispersed with an heterogenous mixtures of unconjugated protein and protein

conjugated with one or more moieties. By contrast, imidazole-1-sulfonyl azide can convert histidine amines to azides at pH 8.5 which can then react with an alkyne group bearing payload (Staudinger reaction or strain-promoted azide-alkyne cycloaddition), resulting in a more limited number of conjugates.[112] Random conjugation can potentially alter protein activity or antigen-binding affinity in the case of immunoglobulins if conjugation is made on CDR regions.

Despite product heterogeneity, the NHS ligation has been successfully applied in a wide range of therapeutic applications and particularly for the preparation of PEGylated conjugates of foreign proteins with reduced immunogenicity (*cf*. 2.1.2. Immunogenicity).

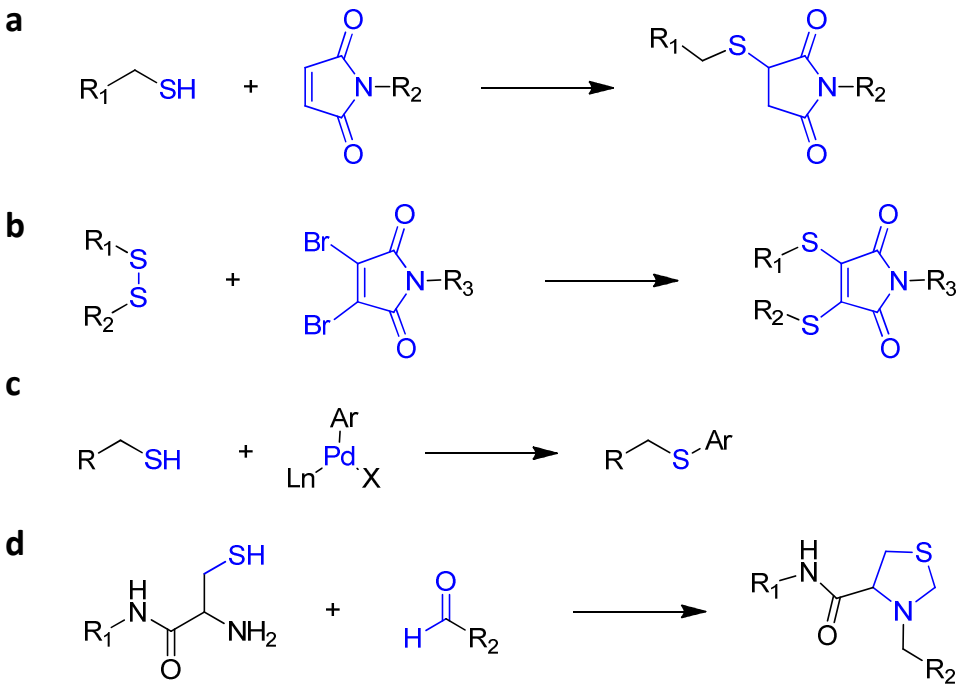
To circumvent the drawback of heterogeneity, numerous further strategies have been investigated to perform site-directed conjugation on sites not involved in the protein function with controlled stoichiometry and minimal structural consequences. One attractive site to perform selective conjugation is the α-amino group of the N-terminal amino-acid.[113] The N-terminal amine is basic and charged at physiological pH. It is solvent-exposed in about 80% of the cases as demonstrated by *in silico* studies performed on 425 monomeric proteins.[114] The N-terminal α-amine has a basic pK_a (~7.8) lower than that of lysine amino groups (pK_a ~ 10.5), due to the inductive effect of the carbonyl group situated nearby. At acidic pH, the proportion of NH_2 to NH_3^+ for both the alpha-amino and Epsilon-amino groups is low although higher for the alpha-amino than for the Epsilon-amino groups. Nucleophilic attack of the alpha-amino group on NHS esters (acylation) or aldehydes (alkylation) is then favored at acidic pH. However, the pH of the reaction should not be too low to avoid degradation of the protein and a too low reactivity of the alpha-amino group.[115] Chemoselective targeting of the N-terminal α-amine represents an asset to obtain reproducible well-defined protein mono-conjugates with more uniform PK/PD properties. One of the best examples of the efficiency of such conjugation is the approval of Pegfilgrastim (Neulasta®), a human granulocyte colony-stimulating factor conjugated to a linear 20 kDa-PEG *via* reductive alkylation at slightly acidic pH (4.5-4.8; **Table *1***).[116]

Scheme entry	Chemical conjugation	Sites of modification	Advantages	Drawbacks	Site-specific
1	NHS Ester	Lysine N-terminus (acid pH)	Simple and reliable	Heterogenous labeling Risk of decreasing protein functionality	No Yes
2a	Maleimide	Reduced cysteine Cysteine C-terminus Engineered cysteine THIOmabs®	Simple Can increase protein stability Homogeneous labeling	Heterogeneous labeling Requires genetic engineering Requires genetic engineering	No Yes Yes
2b	bis-Thiol maleimide	Cysteine disulfide bridge	Homogeneous labeling Increase structural stability	Risk of disulfide scrambling	Yes
2c	Aryl palladium complexes	Cysteine	Homogeneous labeling Preserves functionality Stable towards oxidation	Toxicity of palladium	Yes
2d	Aldehyde	Cysteine N-terminus	Stable Mono-labeling	Requires first-step genetic engineering	Yes
3	Hydrazine	Oligosaccharides	pH-dependent cleavage	Heterogenous labeling Limited to glycoproteins Risk of undesired cleavage	No
4	NTA	Histidine tag N- or C-terminus	Mono-labeling	Toxicity of nickel Requires genetic engineering to add HisTag on the protein	Yes
5	IEDDA	Lysines, cysteines or post NNAA	Covalent, quick, highly specific, non-toxic	2-steps approach Risks of TCO isomerization	No*/Yes
-	NNAA incorporation	N- or C-terminus	Homogeneous labeling	Difficult engineering process Risk of immunogenicity	Yes

Table 8 : Overview of the main chemical reactions used for protein bioconjugation. IEDDA: Inverse-electron-demand Diels-Alder cycloaddition; NNAA: non-natural amino acid; NTA: Ni-nitrilotriacetic acid; TCO: *trans*-cyclooctene. *If random labeling on lysines.

4.1.2. Reactions on cysteine residues

To circumvent the randomness of lysine conjugation, other approaches have been developed for coupling on cysteine, tyrosine, arginine, or histidine residues either via chemical reduction of amino acid side chain functions or addition of reactive amino acids by genetic engineering.[117] The most predominant reactions, summarized in **Table *8***, involved maleimides, diazonium salts, or metallocarbenoid reagents.[118] Maleimide coupling occurs between the high nucleophilic cysteine residue and thiol-reactive groups (**Scheme *2*a**). A thioether bond is formed through the hetero Michael addition, attaching covalently the thiol group to the payload. Maleimide reaction above pH 8 should be avoided because reaction with amino groups may take place.



Scheme 2 : Main reactions for the bioconjugation of moieties on cysteine residues. (a) Reaction between maleimide and a free cysteine. (b) Bis-thiol maleimide reaction on protein-bearing reduced cysteine. (c) Aryl-palladium complex reaction on free cysteine. (d) Aldehyde reaction on N-terminus free cysteine.

However, cysteine is one of the three less abundant amino acids in proteins, along with tryptophan and methionine, with a preponderance of about 1.7% of free cysteines on the protein surface.[119] Most of protein cysteines are involved in disulfide bridges, catalytic

residues or enzymatic and nucleophilic functions.[120] This limitation implies the need for a first genetic engineering step to add a solvent accessible free-cysteine, generally located at the C-terminus of the protein which can easily react with maleimides. This approach has gained high interest since the past decade due to the increasing need for homogenous conjugates for biomedical applications. Single cysteines at the C-terminus easily oxidize to form protein dimers or glutathione adducts, which should first be reduced by reducing reagents —such as tris(2-carboxyethyl)phosphine or 2-mercaptoethylamine— to obtain reactive sites.[120-121] After reduction of the disulfide bridge, free cysteines can be alkylated with payloads functionalized with maleimide function. This alkylation allows the covalent attachment of the payload to the protein.[122] For instance Certolizumab Pegol (Cimzia®) is produced by conjugation of a single cysteine added at the hinge region of the anti-TNFα Fab antibody fragment to a 2-armed 40 kDa PEG.

Thioether bonds are however prone to thiol-exchange reactions and can undergo a retro-Michael exchange process, particularly towards serum albumin or glutathione at physiological pH.[123] Hydrolysis of the thioether bond may influence PK or have consequences on PD and the non-desired cleavage of the payload from the protein can induce off-target toxicity. Alternative maleimide constructs and reactions to cysteine residues have therefore been investigated to generate physiologically stable conjugates *via* the formation of irreversible thioether linkage. Next-generation maleimides have been developed, with functionalization of some leaving groups in position 3 or 4 of the maleimide to obtain covalent linkages and preserve the maleimide scaffold for a second thiol addition.[123b, 124] This strategy consists first in reducing the intramolecular disulfide bridge to obtain two free reactive cysteines and then, reforming an intermolecular bridge with a bifunctional payload such as dibromomaleimide (**Scheme *2*b**).[125] Rebridging with a dibromopyridazinedione construct itself carrying two orthogonal reactive handles could be utilized to introduce two distinct functionalities on the disulfide bridge.[126] This conjugation method provides a homogeneous labeling as well as the precise control of stoichiometry associated with a higher stability of conjugated proteins thereby leading in enhanced PK/PD properties.[123a] Furthermore, dibromomaleimide and dibromopyridazinedione cross-linking reagents are both highly stable towards hydrolysis and highly reactive with cysteine residues.[127] Comprehensive reviews about the chemical mechanisms underlying the conjugation with maleimides are available here.[123]

Recent approaches to avoid retro-Michael addition have been published. One of them consists in performing a transcyclization reaction between a maleimide moiety and an N-terminal cysteine to obtain a 6-member ring locking the thioether moiety. The trapping of the thioether can be an elegant tool for synthesizing more stable maleimide protein conjugates.[128] Other teams suggested to perform self-hydrolyzing of the maleimide, right after conjugation, to make them lose their reactivity towards thiols.[129] An interesting work also suggested that maleimides thiol adducts can be stabilized easily through stretching by mechanical force *via* mild ultrasonication.[130] However, those promising up-to-date maleimide strategies are still at the stage of proof-of-concept and have not been applied in products in clinical trials yet.

The environment surrounding the conjugation site is involved in thiol-exchange or hydrolysis of maleimide payloads. Junutula and co-workers demonstrated the negative influence of the conjugation site on solvent accessibility, charge, propensity to maleimide exchange and downstream impacts on PK and pharmacology.[131] By comparing an antibody conjugated to monomethyl auristatin E (MMAE) at different sites in the heavy and light chains of the Fab or Fc regions the authors observed disparities in the therapeutic activities. The conjugation of MMAE on the light or heavy chain of the Fab region were the most appropriate sites for inhibiting tumor growth and improving survival in mice. The observed differences were correlated with the respective pharmacokinetic properties of the conjugates, the Fab-light chain conjugates showing the highest stability in blood. Mass analysis suggested that both the maleimide exchange from the antibody conjugate and the hydrolysis of succinimide ring in the linker influenced conjugate stability and therapeutic activity. To avoid such deleterious effects, the same team developed the THIOmabs® technology by engineering two cysteines in the constant domains of Fab fragment of antibodies, one in each arm, thus generating two known sites of bioconjugation for stable thioether linkages.[132] These sites have been carefully selected using a phage display-based method to avoid alteration of domains involved in antigen binding functions. Such genetic engineering strategy revolutionized protein bioconjugation with its ability to induce minimal conformational changes in the antibody structure and functionality.

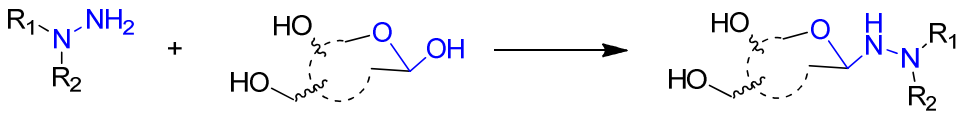
Some alternatives to maleimide reaction for selective conjugation to cysteine residues have also been investigated, among them, the use of metal complexes, such as palladium (**Scheme *2*c**).[133] Briefly, a biarylphosphine bearing palladium reagent is used to transfer an aryl group

onto a cysteine residue then form a covalent thioether bond. Proof of concept of aryl palladium (II) conjugation has been demonstrated on different proteins such as antibodies, protein A, and sortase A.[133-134] The thiol arylation with such metal complexes is interesting as it results in highly stable bioconjugates towards oxidation and acid degradation. However, applications in biological systems remain difficult due to the toxicity of palladium.

A cysteine residue positioned at the N-terminus of a protein can be selectively conjugated due to its distinct 1,2-aminothiol functionality.[135] Thereby, aldehydes-bearing payloads can react with the N-terminal cysteine to form thiazolidines without interfering with other nucleophilic residues such as serines, lysines or other internal cysteines (**Scheme *2*d**). However, the reaction occurs in acidic conditions (*i.e.*, pH 4-5), exhibits slow kinetics (up to 2 days) and requires high concentrations of reactants.[136] A recent study demonstrated that the reaction induces quick dissociation of about 10 to 25% of the payload depending on the pH. These observations suggest that this approach could be more appropriate for pro-drugs or antibody-drug conjugates —for which the detachment of the payload in endosomes could be an asset for therapeutic efficacy— than for the conjugation of long-acting nanomedicines.[137]

4.1.3. Glycosylation of proteins

Glycosylation is generally achieved *via* posttranslational modification through an enzymatic process allowing the conjugation of glycans to the chains on asparagine (*i.e*. *N*-glycosylation), serine or threonine (*i.e*. *O*-glycosylation) (**Scheme *3***).[138] Protein glycosylation is strongly dependent on the expression system used as the expression system induces variations in the nature and number of glycans added, which can affect the biodistribution of the protein.



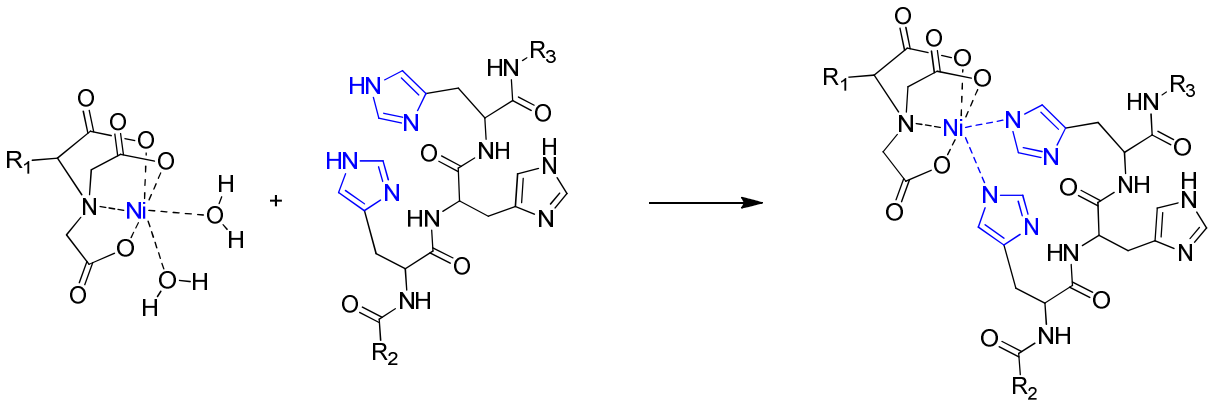
Scheme 3: Bioconjugation on oligosaccharides using a hydrazine-bearing moiety.

Selective conjugation of glycans on N-glycoproteins is performed using hydrazine reactive moieties, based on the oxidation of hydroxyl to aldehyde groups in oligosaccharides using sodium periodate. A covalent bond is thus formed between the aldehyde and the hydrazine group allowing the conjugation of a wide variety of functionalized materials.[139] However, the

hydrazide reaction of oxidized glycans as well as the purification methods to remove the unconjugated moieties highly influence the efficiency, stability, and functionality of the N-glycoprotein conjugates.[140] Such bioconjugation strategy remains mainly used for *in vitro* protein immobilization, purification or *in vivo* identification and quantification in diagnostic applications more than for developing nanomedicines.[141]

4.1.4. N- or C-term conjugation *via* Ni-nitrilotriacetic acid (NTA)

Nickel (II) chelate complex of nitrilotriacetic acid (Ni^{2+}-NTA) was originally developed for the immobilization of His-tagged proteins on a surface, as required for BIACORE analysis systems.[142] However, due to the complementary interaction between the His-Tag and Ni(II)-NTA this approach is also attractive for site-specific bioconjugation at the His-tag of recombinant proteins (**Scheme *4***).[143] Complexation of NTA-bearing moiety with protein N- or C-term His tag exhibits an affinity of about 10^{-6} M. The reaction is fast and selective and generally allows to preserve protein activities.



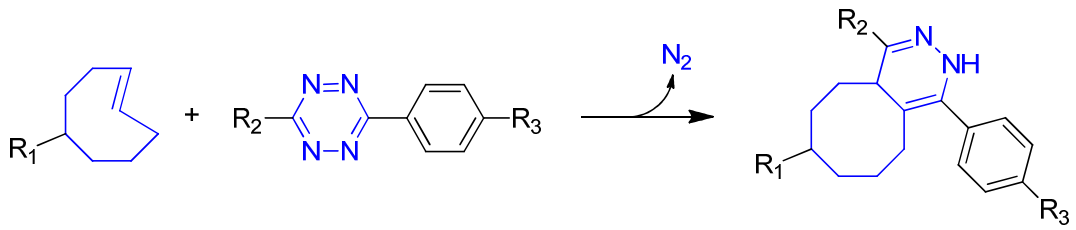
Scheme 4 Bioconjugation at N- or C-term of histidine residues using Ni-nitrilotriacetic acid (NTA).

NTA demonstrated interesting results with the possibility to either conjugate fluorescent probes to proteins or to directly target specific His-Tag proteins in cells (with cognate cell ligand-bearing protein) thereby offering great opportunities for specific protein tracking in living systems with minimal alterations on the protein structure, function, or localization.[144] Other metals such as Cobalt (II) or (III) have also been used with the NTA system to target the protein His-Tag and demonstrate high stability as well as inertia towards ligand exchange.[145] However, to date, NTA was mainly applied to nanoparticles engineering and did not succeed

in *in vivo* experiments yet, an issue resulting from its versatility, reversibility, and metal toxicity.

4.1.5. Bioorthogonal click chemistry

Bioorthogonal click chemistry encompasses different highly-specific and rapid chemical reactions that occur in living systems, at physiological pH and temperature, without the need of a catalyzer and where the chemical moieties used are totally inert towards biological molecules which ensures biocompatibility.[146] A decade ago, bioorthogonal click chemistry started to be investigated for in vivo protein labeling especially using polymers, toxins, radionuclides or cytotoxic molecules.[147] Bioorthogonal chemistry conjugation is performed in a two-steps approach where the protein is first conjugated to a highly reactive chemical and then the counterpart chemical bearing the moiety to attach is added in a separate step. One of the most popular bioorthogonal reactions is the inverse-electron demand [4+2] Diels-Alder cycloaddition (IEDDA), which occurs between a poor electronic diene, such as 1,2,4,5-tetrazine (Tz) and a dienophile alkene such as trans-cyclooctene (TCO) to form a covalent bond (**Scheme *5***).[146-147]



Scheme 5: Inverse-electron demand Diels-Alder (IEDDA) cycloaddition between *trans*-cyclooctene (TCO) and 1,2,4,5-tetrazine (Tz). The covalent bioorthogonal reaction is associated with a release of N_2. Note: inverse reaction with R1-linked Tz followed by the addition of R2-TCO works also but is less described.

TCO (or Tz, as the invert reaction is also feasible) is conjugated to proteins through different ways.[148] The most common method remains the random conjugation on lysine residues. However, site-specific conjugation, *via* engineered cysteine residues, C-term residues or other amino acids incorporated by genetic engineering (*cf*. 4.1.6. Non-natural amino-acid incorporation), has been used in the last few years to control reaction stoichiometry and avoid modifications on the functional part of the protein. Thanks to a high reaction kinetic rate (*i.e.*,

k_2 up to 10^6 $M^{-1}s^{-1}$), IEDDA allows conjugating proteins to theranostic payloads either before *in vivo* administration (direct protein modification) or by delaying administrations of protein and payloads by a few minutes to several days (indirect protein modification) for a binding occurring directly *in vivo*.[148] The latter is especially appropriate for pretargeting in a 2-steps strategy developed to reduce off-target toxicity induced by long-acting antibody carrying cytotoxic payloads.[146-147] However, the IEDDA reaction is also an interesting tool to conjugate PEG or SA to protein for the development of long-acting medicines.

The major drawback of this reaction arises from the isomerization of TCO into its inactive isomer *cis*-cyclooctene which occurs over time after protein-TCO has been administered in blood due to non-specific interactions with transition metals.[149] Indeed, *cis*-cyclooctene is the most favorable conformation as the *cis* position of the C=C in the cycle is more stable. TCO is thus synthesized from *cis*-cyclooctene by photochemical conversion and has to undergo quality check before any coupling as it is prone to isomerize.[150] The risk of isomerization has to be particularly considered for pretargeting strategies and implies limiting the delay between the injection of the protein-TCO and the payload-bearing Tz. However, in the case of direct coupling before *in vivo* administration of the complex, the risk of isomerization remains negligible, TCO being highly stable in PBS during several weeks.[149b]

Click chemistry progressively becomes a very useful tool in biochemistry, especially for protein bioconjugation due to its ease of use. Another bioorthogonal reaction, the strain-promoted azide-alkyne cycloaddition (SPAAC) has recently demonstrated efficacy for site specific PEGylation of fibroblast growth factor 2 with a slightly improved half-life in mice.[151] Today, about a dozen of bioorthogonal click reactions have been developed for *in vitro* or *in vivo* applications. The choice of the appropriate reaction is highly dependent on the reaction rate, the type of protein targeted, the moiety to attach and the desired application.[152] While those reactions are specific, catalyst free and non-immunogenic, their application in living systems are recent and still challenging. The TCO/Tz IEDDA cycloaddition was the first bioorthogonal reaction to enter clinical trials last October 2020 for application in oncology as an antibody-drug conjugate (NCT04106492).

4.1.6. Non-natural amino-acid incorporation

Genetic engineering can be used to insert non-natural amino acids (NNAA) in proteins to allow site-specific chemical conjugation at known sites without involving natural amino acids from

the original protein sequence.[153] NNAA are incorporated genetically by reprogramming a rarely used codon or a non-sense codon in the gene of interest achieved by an orthogonal pair aminoacyl tRNA synthetase (aaRS) that loads a specific tRNA with a NNAA. The selected NNAA-specific synthetase must not recognize any host tRNAs or cognate amino acids while the orthogonal tRNA anticodon must not be aminoacylated by any host aaRS and should be mutated to recognize a stop codon or a non-sense codon (**Figure 3**).[154] The amber stop codon TAG is frequently used for NNAA incorporation as this codon is rarely used in *E. coli* and is one of the lowest used in humans (*i.e.*, occurrence of TAG in humans: 23%).[155] Today, over 80 NNAAs have been encoded into proteins in several prokaryotic and mammalian *in vivo* expression systems leading to proteins functionalized either with PEGs, chemicals or radionuclides.[156]

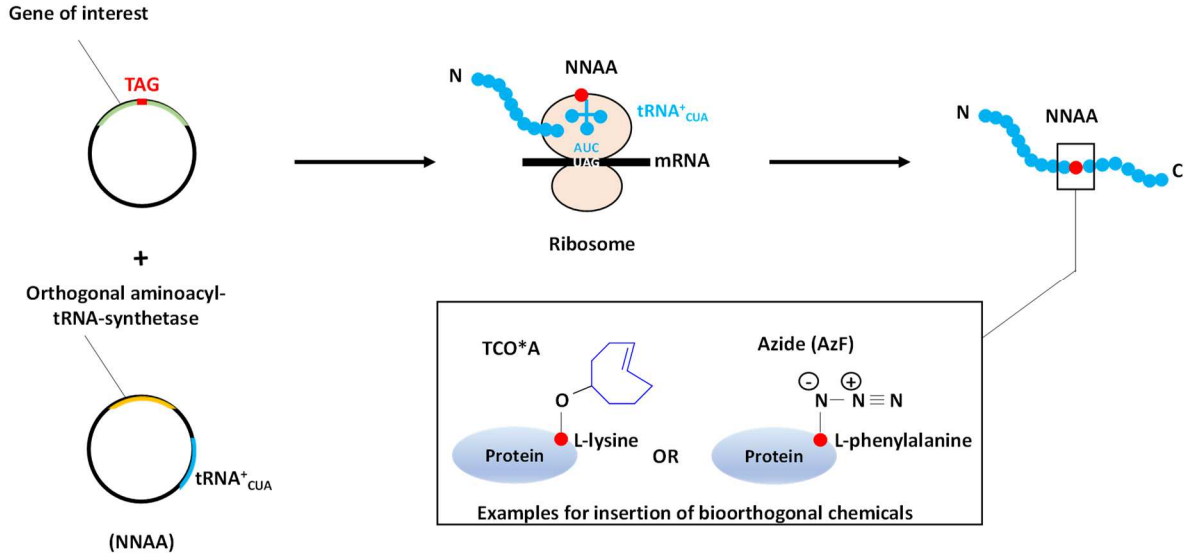


Figure 3 : Incorporation of non-natural amino acids (NNAA) bearing bioorthogonal moiety into a protein *via* the Amber stop codon (TAG) approach. Examples with the incorporation of *trans*-cyclooctene-L-lysine (TCO*A) and *p*-azido-L-phenylalanine (AzF). Adapted from reference [157].

Recently, cell free synthesis has been developed as a rapid, cost-effective and virus-free process for manufacturing protein bearing NNAA. Employing an *M. jannaschii* TyrRS-derived synthetase/tRNA pairs, Otting and co-workers succeeded to incorporate a variety of NNAA, such as para-acetyl-L-phenylalanine, bipyridyl-phenylalanine or L-(7-hydroxycoumarin-4-yl)

ethylglycine in *E. coli*-based cell-free expression system but failed to incorporate para-azido-L-phenylalanine required for further click chemistry reaction.[158] Later, another team succeeded to incorporate para-azidomethyl-L-phenylalanine for performing strain-promoted azide-alkyne cycloaddition.[153a] Thereby, NNAA has been effective in human cells for the incorporation of amino-acid bearing bioorthogonal components such as azide moiety (for Staudinger or strain-promoted azide-alkyne cycloaddition) or cyclooctene moiety (for IEDDA cycloaddition) into eGFP expression plasmid (**Figure 3**).[159] P-azido-L-phenylalanine was incorporated in eGFP *via* the orthogonal *M. jannaschii* TyrRS-derived synthetase/tRNA pairs and the *E. coli* aaRS encoded on two separate plasmids while TCO-L-lysine incorporation was mediated by an orthogonal tRNA/aaRS pair from *Methanosarcina mazei* co-encoded on a single plasmid.

Since the last decade, NNAA incorporation has gained increasing interest as it enables the attachment of various organic probes to proteins without inducing deleterious modifications of the functional region of proteins. Compared to classical chemical reactions on solvent free amino acid residues, the NNAA approach is advantageous as it allows selective conjugation of moieties while preserving the moiety linkage from the risk of unspecific cleavage. However, the need of genetic engineering on proteins is a constraint, in terms of cost, materials and time of accomplishment, which can be a major drawback in comparison to easy and quick chemical conjugation on lysines or cysteines. To date, there is no protein-conjugated NNAA in the market yet as this approach is in its early stage preclinical development.

4.2. Chemo-enzymatic reactions

Chemo-enzymatic reactions are an excellent way to selectively conjugate payloads to proteins; the main approaches are summarized in **Table *9***. Enzymes can be used either to tag the protein or as a catalyzer for chemical conjugation. Chemo-enzymatic reactions can be used on WT or genetically engineered proteins, thereby combining different conjugation approaches to obtain optimal conjugates with tightly controlled stoichiometry.[160] The choice of the method is critical to obtain proteins conjugated with optimal PK/PD and stability. A wide variety of molecules can be conjugated to proteins, each with specific purposes, ranging from polymers to antibodies, enzymes, toxins, drugs, cytokines or other proteins.[160-161] It is also interesting to note that combinations of different moieties can be performed to multiply different desired effects.

Chemo-enzymatic conjugation	Sites of modification	Advantages	Drawbacks	Site-specific
Biotin	Lysine	Simple	Heterogeneous labeling (Strept)avidin toxicity	Yes/No
Sortase	LPXTG glycine tag	Homogeneous labeling	Engineering of LPXTG motif Recreates original sequence motif Need high molar excess	Yes
*Oa*AEP1	Glycine-Valine	Homogenous labeling Does not recreate the original sequence motif	Engineering the modified motif	Yes
Transglutaminase	Glutamine	Highly homogeneous labeling	Requires a first-step of genetic engineering	Yes
GalT/SialT	Asparagine	Homogeneous labeling	Multi-steps engineering Oxidation of methionine residues	Yes
GalT/GalNAz	Asparagine	High homogeneous labeling	Multi-steps engineering	Yes
GlycoConnect	Serine or threonine	Homogeneous labeling	Multi-steps engineering	Yes
Formylglycine generating enzyme	Cysteine	Homogeneous labeling	Engineering of CXPXR sequence	Yes

Table 9: Overview of the main chemo-enzymatic reactions used for protein conjugation. GalT: β-1,4-galactosyltransferase; GalNAz: galastosyltransferase N-azidoacetyl-galactosamine; LPXTG: lysine-proline-X (any amino acid)-tyrosine-glycine tag; SialT: α-2,6-sialytransferase.

4.2.1. Biotinylation

Biotinylation is one of the easiest ways to modify proteins with enzymes (**Table *9***). Biotin (also known as vitamin H, vitamin B7 or coenzyme R) is a small cofactor of 244 Da for carboxylase enzymes present in all living organisms. Avidin is a positively-charged tetrameric protein of about 60 kDa naturally found in white eggs of birds, reptiles and amphibians; its bacterial analog, streptavidin, being produced by *Streptomyces avidinii*.[162] The interaction between biotin and (strept)avidin is one of the strongest known non-covalent biological reaction with a K_D of about 10^{-15} M at pH 5.0. The ligation is also quick, highly specific and highly stable, allowing the reaction to resist the action of protease or denaturant agents and occurs even under harsh conditions such as high temperature or pH, an asset for stable attachment of polymers, proteins, fluorophores or other payloads.[162]

Chemical conjugation of biotin, often performed on lysine residues generates heterogenous products. However, biotin ligase can also be used as a catalyzer for the enzymatic conjugation of biotin derivatives onto proteins thereby leading to homogeneous products.[160, 163] This technology is ATP-dependent and requires a biotin ligase, such as BirA which is produced in *E. coli*, and a previous step of genetic engineering on the protein to add a polypeptide sequence (*e.g.*, AviTag peptide as specific substrate of BirA enzyme) specific for the enzyme.[164] Indeed, Schatz and coworkers have found a sequence of 13 amino acids to be the minimal substrate for BirA enzyme (*i.e*. LXYIFEAQKIEWR, where X = any amino acid and Y = any except L, V, I, W, F or Y). In order to improve the rate of biotinylation, the sequence has then been improved several times resulting in AviTag (GLNDIFEAQKIEWHE), BioTag (ALNDIFEAQKIEWHA) and other derivatives. Briefly, the Tag peptide is genetically inserted either at the N- or C-terminus or even in exposed loops of the target protein. The co-addition of BirA, biotin and ATP in the reaction medium forms an intermediate biotinoyl-5'-AMP stuck in BirA ligase until the enzyme recognizes the Tag sequence. Metabolization of the Tag substrate by BirA release the biotinoyl-5'-AMP intermediate which reacts with the proximal lysine residue contained in the sequence of the Tag fusion protein for covalent bioconjugation (**Figure 4**).

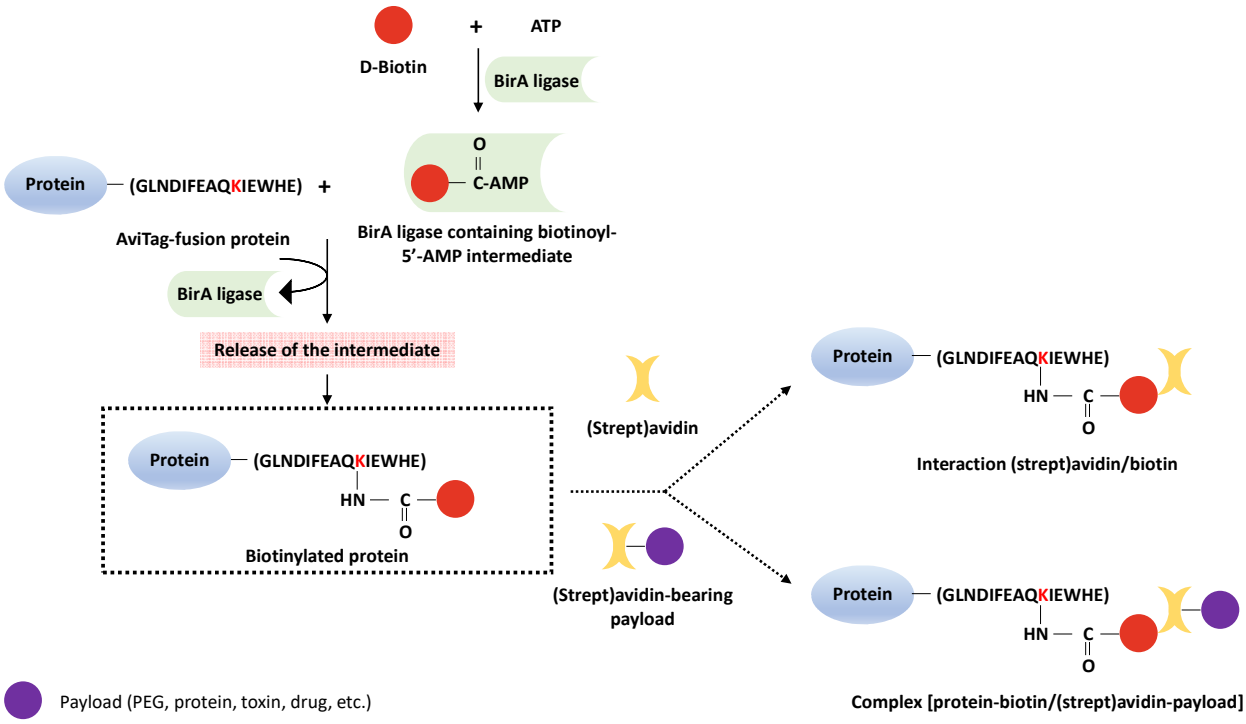


Figure 4: Site-specific enzymatic biotinylation using BirA ligase. BirA recognizes specifically biotin and the 15-amino acids AviTag sequence genetically inserted in the N- or C-terminus of the target protein. The ATP-dependent reaction results in the formation of biotinoyl-5'-AMP intermediate stuck into BirA until the ligase recognizes the lysine of the AviTag peptide. The acylation of the lysine forms a covalent bond between the AviTag-fused protein and biotin. Biotin can then react with wild-type (strept)avidin for *in vitro* protein detection or be conjugated with (strept)avidin-bearing payload for direct *in vivo* administration or for indirect pretargeting strategies.

Site-specific functionalization of proteins using biotin ligase has been demonstrated for numerous applications, as for example for the *in vivo* quantification of protein-protein proximity of Sox2 and Oct4 transcription factors, the labeling of cell surface proteins or to conjugate polymers, other proteins or chemical payloads.[165] However, if such strategy brings an interesting prospective, (strept)avidin immunogenicity still impedes further clinical development.[166]

4.2.2. Transpeptidation

A number of transpeptidases, including sortase A from *Staphylococcus aureus*, (the most extensively used), butelase-1 from the tropical plant *Clitoria ternatea* and OaAEP1 from the cyclotide-producing plant *Oldenlandia affinis* have been used to modify proteins site specifically (**Table *9***). These enzymes cleave the peptide bond of a specific motif (cleavage motif highlighted in red in **Figure 5**) and form a new peptide bond with an incoming

nucleophile composed of specific amino acids (the receiving motif highlighted in green in **Figure 5**) to which a payload of choice is attached. The payload is often another peptide or protein but importantly, non-genetically encoded functional groups can be fused to proteins according to this approach given that they are synthetized and chemically grafted to the appropriate cleavage or incoming motif (**Figure 5**). Such non-genetically encoded functional groups include PEG, dyes, biotin, oligonucleotides, radioisotope, unnatural amino-acid, lipids, or carbohydrates.[167] Since the reaction is carried out in organic solvent-free mild conditions, most proteins can be modified according to this approach without affecting their structure and stability. The unique requirement is that they can first be equipped, at the appropriate location, with a cleavage sequence and/or an incoming sequence by genetic engineering and produced recombinantly. The specific design features required depend on which terminus of the protein the modification is desired and on which enzyme or combination of enzymes is used. A short, flexible linker often composed of Gly_4Ser repeats is usually added between the target protein and the engineered cleavage site.[168]

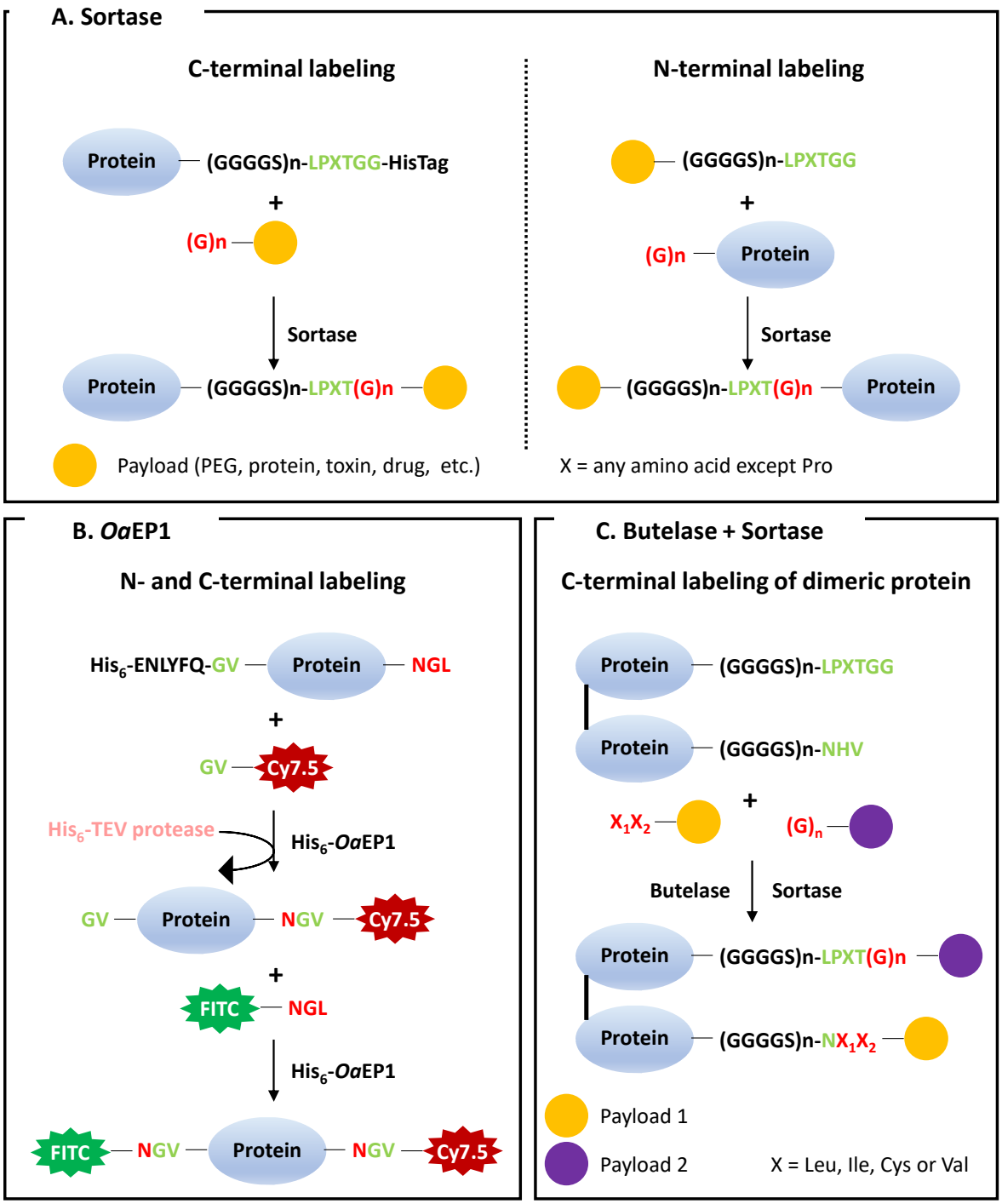


Figure 5: Site-specific modification of proteins using transpeptidases. In A-C, the cleavage motif and the incoming nucleophile motif are shown in red and green, respectively. (GGGGS)n constitutes the linker. A) Sortase A-mediated modification of a protein at its C- or N-terminus: Sortase A specifically recognizes the LPXTG (X= any amino acid) motif. It mediates the bioconjugation by cleaving the bond between threonine (T) and glycine (G) for binding the desired moiety, fused -via an oligoglycine. B) Sequential double labelling of a protein with two fluorescence probes via OaAEP1. C) Simultaneous labelling of a dimeric protein (with an intermolecular disulfide bond) using butelase 1 and sortase A.

Sortase. Sortase, as the other transpeptidases, can be used to modify proteins at their C-ter and/or N-ter.[169] In order to modify a protein at its C-terminus, a specific LPXTGG cleavage motif is added by genetic engineering to the C-terminus of a protein of interest, while an incoming nucleophile G(n) motif is added -by genetic engineering or chemical synthesis- to the N-terminus of the payload to be ligated. If the protein has to be modified at its N-terminus, then the receiving nucleophile motif should be genetically engineered at this position, while the cleavage motif is added to the payload. An illustration of these two scenarios is given for sortase A-mediated modification in **Figure 5 A**. With a specific design of the target protein, sortase has also been used to prepare unnatural C-to-C fusions of two different proteins or to label the protein at the internal position of its sequence.[167a, 168] The completion of the sortase reaction recreates the sequence motif originally recognized by the enzyme, so that the reaction is reversible. In the absence of an adequate concentration of incoming nucleophile motif, the acylenzyme intermediate is remarkably stable, and the reaction is driven to completion only in the presence of a significant molar excess of the incoming nucleophile.[167a] The reconstitution of the cleavage motif in the ligation product prevents, therefore, the use of WT-Sortase A to successively label a protein at both its N- and C- termini. One of the first applications of sortase to protein engineering was the conjugation of cytokines with PEG to extend the cytokine half-life via direct coupling with a GG-PEG payload.[170] Sortase is also capable of protein cyclization to improve the stability of a protein. To achieve this the nucleophile GG is added to the N-terminus of the protein, while the LPTXGG motif is attached to the C-terminus.[170] To circumvent the poor kinetic parameters of WT sortase, mutational variants of sortase A, with at least 3-fold improved catalytic efficiency compared to the WT enzyme, have been produced leading to a better coupling yields.[171] Moreover, recent studies with a series of sortase mutational variants showed, that different variants can lead to substantially better results according to the protein modification that is desired. Thus, there is not such a thing like a unique sortase able to perform all kinds of reactions efficiently, but a set of variants each one dedicated for a specific application. A comprehensive comparison of different variants is available in the work carried out by Li et al.[172] Moreover, new transpeptidases, with specific characteristics, have been recently identified such as VyPAL2. [173] In addition, from the 3D-structure of different transpeptidases and their mutational variants, the molecular bases underlying efficient ligase activity start to be understood and could be used for the rational engineering of enzymes with the desired catalytic activity.[173]

Moreover, optimisation of transpeptidation protocols involving for example the use of immobilized enzymes allows to reduce significantly the amount of enzyme needed and it should allow to scale-up reactions for an industrial use.[174]

Recently, the sortase reaction has been a matter of extensive research in the field of imaging and radiology. One example is the labelling of nanobodies with ^{111}In for single-photon emission computed tomography (SPECT).[175] In this approach, a nanobody is tagged with the motif LPTXG at its C-terminal and the peptide GGGKY is functionalized with the chelating agents CHX-A"-DTPA and NOTA to bind to the radioactive isotope. The sortase catalyzes the site-specific incorporation of the chelating agent at the C-terminal of the nanobody yielding a stoichiometric and homogenous product that can be easily separated by size-exclusion chromatography. Among the numerous applications of labelling antibodies in radiology, the image-guided surgery is one of the most innovative and has a high potential to become a standard practice in oncologic surgery in the future.[176] It however requires a precise and rapid delimitation of the tumor to be excised in the course of a surgery with exquisite contrast to normal tissues. Given their versatility, specificity, and the short time needed to attain high contrast, nanobodies are very promising is this field.[177]

Butelase. The cysteine-transpeptidase butelases show a substantially higher catalytic efficiency than sortases and a shorter recognition motif (D/N-HV), leaving an N residue after cleavage. The recent peptide specificity characterization of the recombinant enzyme in *E. coli* using an affibody model has allowed the synthesis of cyclic hydrophilic doxorubicine conjugates with conserved affinity towards EGFR-overexpressing A431 cells.[178] Given the superior catalytic efficiency of this enzyme, and the possibility of *E. coli* expression, we expect substantial progress in protein engineering strategies based on butelases in the coming future.

OaAEP1. In the case of OaAEP1, by screening different nucleophile motives, Rehm and coworkers identified a GV dipeptide that readily served as a nucleophile in the ligation reaction, but the product of that reaction (NGV) is poorly recognized by the enzyme.[167b] Thus, the ligation product is resistant to the reverse reaction and this allows a straightforward efficient site-specific sequential modification of a protein of interest both at the C and N terminus using the same enzyme (**Figure *5* B**). Such an approach was efficiently used to prepare a nanobody functionalized with two fluorescent probes: Cy7,5 at the C-Ter and fluorescein at the N-ter.[167b] The nanobody was first equipped with (i) a TEV protease

recognition sequence (ENLYFQ) at the N-terminus, followed by the GV sequence.[167b] The TEV sequence initially protects the GV sequence from the attack of the enzyme, and (ii) a C-terminal NGL. In a first reaction, the C-terminus of the nanobody was fused with a Cy7.5-labeled GV nucleophile, generating the NGV sequence at the C-terminus. Then, TEV was added to the reaction mixture to remove the ENLYFQ sequence and thus expose the GV nucleophile at the N-terminus. Finally, the N-terminus was fused with a fluorescein-labeled NGL probe, while leaving the C-terminal NGV sequence generated in the first ligation intact. The labeling at each step was higher than 90% as determined by MS and the product could be purified easily using an IMAC chromatography (i.e., the His-tagged transpeptidase and TEV bind to the column) and a cut-off filtration to remove enzymes and probe between steps. A similar strategy was used to create polymerized proteins step by step in a rationally-controlled sequence.[179] WT OaAEP1 has low kinetic parameters that limit its utilization in bioconjugation of proteins, nevertheless the mutation C247A substantially increases the catalytic efficiency of the enzyme.[180]

Dual site-specific coupling using a combination of transpeptidases. Interestingly, since each transpeptidase possesses a specific cleavage and receiving motives, it is possible to use a combination of two of them to label a protein at two distinct sites in a one-pot reaction and thus creating multimodal proteins (**Figure 5 C**). For example, sortase A and butelase 1 were used to label an IgG at both the C-terminus of its light and heavy chain.[168] The IgG was first modified genetically to add a sortase LPETGG cleavage motif at the C-terminus of the light chain and a butelase NHV cleavage motif at the C-terminus of the heavy chain. Two different fluorescent probes bearing an appropriate incoming nucleophile motif were synthetized: (ii) an oligo-glycine peptide bearing 5,6-carboxyfluorescein (GGG-FAM) and an alanine-leucine peptide bearing the AlexaFluor 647 fluorescent (AL-Alexa). The modified IgG was incubated with GGG-FAM, AL-Alexa, sortase A and butelase 1 at 4 °C for 15 h followed by incubation at 37 °C for 4 h. The yield of labelling for each fluorescent probe was higher than 95%. A simple centrifugation-based size exclusion was used to remove unincorporated dyes and to obtain pure dually modified IgG. Another dual labeling involving two transpeptidases (i.e. butelase and VyPAL2) in one single pot allowed to modify an EGFR-targeting affibody with a fluorescein tag and a mitochondrion-lytic peptide at its respective N- and C-terminal ends.[178] Moreover, butelase was used in conjunction with a π-clamp conjugation to conjugate the C-terminus of

a scFv-Fc of the 4B3monoclonal antibody (i.e., a mAb specific for the human EphA2 receptor which is overexpressed in glioblastoma) simultaneously to a rhodamine and a fluorinated biphenyl-PEG11-fluorescein.[181]

4.2.3. *N*- and *O*-glycan engineering

N- and *O*-glycan engineering is a recent tool for embedding chemical reporters within glycans *via* the *N*-glycan terminus of asparagine residues using multi-steps chemo-enzymatic reactions (**Table *9***). At first, this approach was developed to label glycans directly in cells with chemical reporters for diagnostic detection, in the same way as GFP has been for proteins. Using endogenous biosynthetic cell-surface pathways, chemical reporters have been embedded in glycans *via* the metabolic introduction of sialic acid, *N*-acetylgalactosamine or *N*-acetylglucosamine monosaccharides bearing azides (**Figure 6 *A***).[159] Selective glycoconjugation has proven to be effective for PEGylation of FVIII (Esperoct®) and rFIX protein (Refixia®) commercialized for the treatment of hemophilia A and B, respectively.[182] CMP-activated sialic acid-6'-40 kDa PEG was first transferred using sialyltransferase to terminal galactoses of *O*-linked glycan of N8 peptide (FVIII protein) or *N*-glycans of N9 peptide (rFIX protein) and then sialylated by addition of excess unmodified CMP-sialic acid. The reaction achieved a product yield of about 99% and glycoPEGylation preserved the biological activity of the protein.[183]

Recent advances in post-translational remodeling of native glycan located at specific asparagine residues (e.g., Asn-297) —*via* the combination of β-1,4-galactosyltransferase (GalT) and α-2,6-sialyltransferase (SialT)— within the Fc domain of mAbs allow to incorporate sialic acid moieties.[184] The latter is then mildly oxidized to introduce functional aldehyde groups which can subsequently be conjugated with aldehyde-reactive aminooxy bearing payload (**Figure 6 *B***).[185] However, the oxidation of sialic acid for the conversion into aldehyde also induces the oxidation of methionine residues, an issue requiring new optimizations. Thereby, a recent chemoenzymatic bioconjugation strategy, called GlycoConnect, uses endoglycosidase S to hydrolyze the chitobiose core of Asn-297-linked heavy chain Fc glycans (**Figure 6 *C***). Then, β-1,4-galactosyltransferase T (Y289L) and *N*-azidoacetylgalactosamine are both added to the reaction medium of the immunoconjugate to incorporate azides into the residual glycan chains.[186] The azido residue finally allows to perform click chemistry to conjugate several kinds of conjugates, ranging from biological (peptides, proteins, toxins,

oligonucleotide) to chemicals (cytotoxic drug, radionuclide, fluorescent probe, nanoparticle). To improve the therapeutic effect, dendronized chemical structures can be clicked to the azide moiety thereby providing several binding moieties with minimal modifications on the antibody. This recent strategy was efficient to conjugate 8 reactive TCO (4 on each part of the IgG_1 Fc) with only two sites of modification for pretargeting purposes. The dendrimeric scaffold demonstrated specific interaction and showed a significantly higher uptake in tumors than the non-dendronized one.[186b]

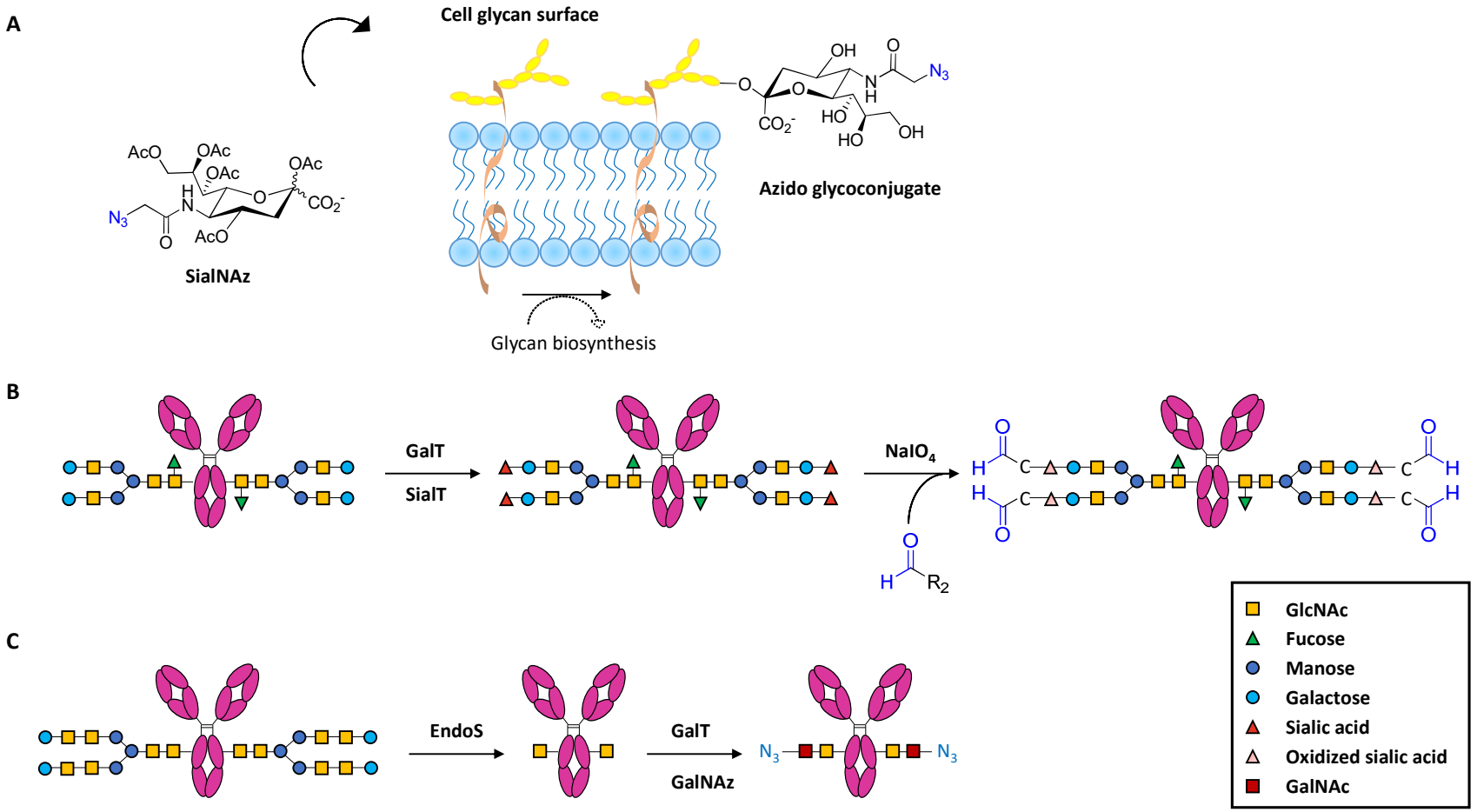


Figure 6 : Different strategies for N-glycan engineering. (A) Direct conjugation of azide-chemical reporter on cell surface glycans via endogenous biosynthetic pathways. (B) General antibody *N*-glycan engineering using the combination of β-1,4-galactosyltransferase (GalT) and α-2,6-sialyltransferase (SialT). NaIO4 is then added to oxidize sialic acid into reactive aldehyde. (C) Simplified GlycoConnect strategy developed by Zeglis and co-workers. Endonuclease S first hydrolyzes Asn-297 then β-1,4-galactosyltransferase T (Y289L) and *N*-azidoacetylgalactosamine are both added to introduce an azide moiety for further click chemistry reaction. GalNAc: *N*-acetylgalactosamine; GlcNAc: *N*-acetylglucosamine.

4.2.4. Formylglycine generating enzyme

In aerobic conditions, the formylglycine generating enzyme oxidizes cysteine and serine side chains into formylglycine (**Table *9***).[187] Physiologically, formylglycine-generating enzyme activates type-I sulfatases by post-translational generation of the catalytic C^{α}-formylglycine leading to the conversion of a cysteine (C) or serine (S) residue embedded in a highly conserved C (or S) XPXR minimal motif (X = any amino acid, P = proline, R = arginine), into an aldehyde-bearing formylglycine, with a conversion rate of 75% to 90%. Formylglycine-generating enzyme is an interesting strategy for site-specific bioconjugation as this system not only recognizes the required motif in sulfatases but also in recombinant proteins engineered with the minimal motif CXPXR, called “aldehyde tag”, allowing the introduction of bioorthogonal reactive chemicals in the protein of interest (**Figure 7**). Indeed, aldehydes are a highly reactive electrophile suitable for bioorthogonal reactions such as aminooxy or hydrazide reagents to form oxime and hydrazine conjugates, respectively, as previously described in the sections “4.1.2. Reactions on cysteine residues” and “4.1.5. Bioorthogonal click chemistry”.[188]

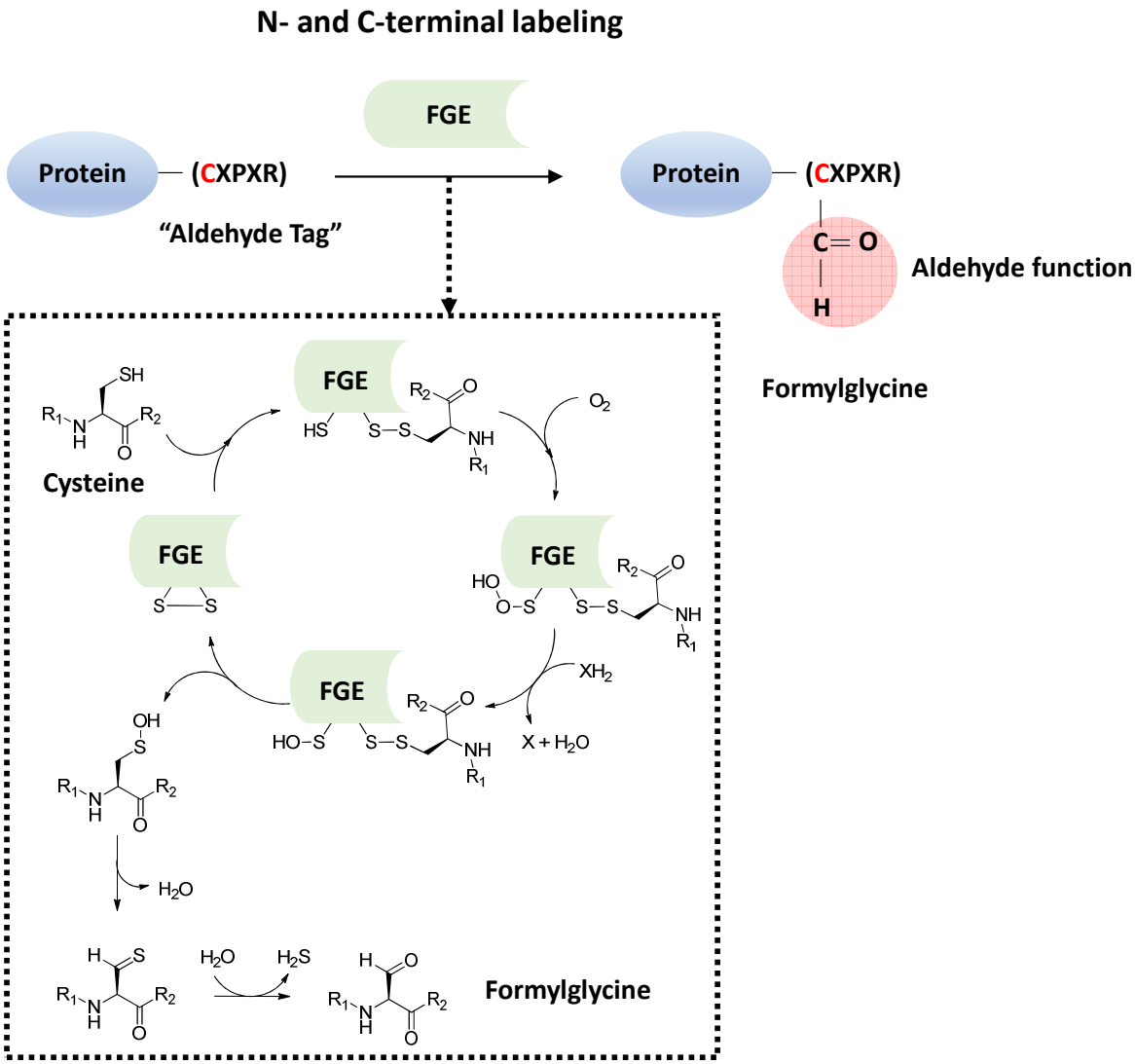


Figure 7 : Conversion of cysteine embedded in CXPXR motif into formylglycine-bearing reactive aldehyde using formylglycine-generating enzyme (FGE). The CXPXR (C: cysteine, P: proline, R: arginine and X: any amino acid) motif called "aldehyde tag" is first genetically encoded in the target protein. FGE catalyzes the conversion of the cysteine into formylglycine-bearing reactive aldehyde through the proposed catalytic mechanism (aerobic condition). The newly formed aldehyde then allows bioorthogonal reactions with various hydroxylamines or hydrazines-bearing payloads.

The formylglycine generating enzyme method is quite simple, only requiring to encode the CXPXR motif at the desired site of the gene of interest in an appropriate expression vector to be inserted at the N- or C-terminus of the protein or within internal regions to facilitate a protein design without critical alterations of protein stability or functionality.[189]

Thereby, the formylglycine generating enzyme is efficient for site-specific PEGylation of proteins or for the conjugation of fluorogenic probes and payloads. However, its main

application deals with oncology for the construction of antibody-drug-conjugates.[188, 190] To improve the conversion rate, recent attempts have been made with copper(II)-containing media as a catalyst of the reaction. Despite improving the *in vitro* activity 20-fold, the toxicity of copper makes application in living systems challenging.[191]

5. Conclusions and perspectives

The high specificity of their biological activities renders proteins particularly attractive drugs for the treatment of a wide range of diseases. Yet, many proteins are rapidly cleared from the systemic circulation and need to be injected several times a week or even a day to patients. This high therapy burden jeopardizes patient quality of life, patient adherence and therapeutic outcomes of medications. Therefore, a major medical need for biologics with a prolonged half-life has arisen and protein engineering and chemistry have been able to answer this demand.

The discovery that conjugation to PEG was able to decrease protein immunogenicity and to prolong blood circulating half-life has been the first significant milestone in the field. PEGylation resulted in the marketing of several biobetters with prolonged half-life comprising a few PEGylated conjugates of prokaryote and animal proteins. These latter are particularly remarkable because they are able to relieve severe human diseases such as leukemia and severe combined immunodeficiency disease. Yet, the non-biodegradability of PEG represents a limitation of this polymer and scientists have turned to natively disordered polypeptides such as XTEN, PAS, and ELP as PEG alternatives. In addition to biodegradability, polypeptide polymers are genetically fused to the parent biologic and can be expressed in *E. coli*. Although no polypeptide fusion proteins have reached the market yet, several are in clinical development.

In parallel to conjugation to polymers, scientists have exploited the FcRn-mediated recycling pathway to design proteins with prolonged half-life. Therefore, IgG and albumin have been harnessed to prolong the half-life of proteins by either fusing the therapeutic protein to a FcRn binding region or by fusing the therapeutic protein to a molecule that non covalently binds to albumin. Etanercept, the first biobetter based on FcRn recycling, has been marketed in 1998. It is a fusion protein between the TNFα receptor and the Fc of an IgG1 antibody. Most of the

Fc- and albumin-fusion proteins have been produced by genetic engineering. The second approach where the therapeutic protein is fused to molecules binding endogenous SA has also resulted in several marketed biobetters and insulin Detemir is certainly the most known. However, research and development are still on-going in this area and molecules binding albumin are sought in bacterial proteins as well as antibody fragments.

Several biobetters have been easily produced by genetic fusion. However, others have required more complex chemical or chemo-enzymatic reactions. Initial reactions have used the abundance of lysine residues in proteins to randomly ligate polymers. Although random conjugation causes a significant loss in protein activity, it is also the most effective to cover antigenic sites and to decrease the immunogenicity of non-human proteins. Therefore, the interest in random conjugation is expected to persist in the coming years. Yet, the chemistries mostly investigated lately have involved site-specific ligation in order to avoid grafting peptides or prosthetic groups too close to the active site and losing protein activity. In this regard, chemo-enzymatic reactions are an excellent way to selectively conjugate payloads to proteins. Selective glycoconjugation has proven to be effective for PEGylation of coagulation factors with the recent marketing of Nonacog beta pegol and Turoctocog alfa pegol in 2017 and 2019, respectively.

In this review, we emphasized the intense activity and the constant development of the protein engineering field to provide biobetters with optimized PK profiles to patients. Yet, challenges and opportunities remain. Up to now, only PEGylation has been investigated as an approach to decrease protein immunogenicity. Because PEG is non-biodegradable, substitute biodegradable polymers should be identified with this goal. The abundance of IgGs and albumin in blood and their long serum half-life have made them ideal target carriers for engineering therapeutic protein constructs with extended circulation time. The future is likely to witness the use of immunoglobulin-binding domains, as alternatives to albumin-binding domains, to prolong the half-life of proteins.

Acknowledgements

This work was supported by research grants of the Fonds National de la Recherche Scientifique (FNRS, Belgium; grant number: T023819F) and Belgian Walloon Region (WALInnov grant

number: 1710173). Mireille Dumoulin and Rita Vanbever are Research Associate and Research Director of the FNRS, respectively.

Bibliography

[1] G. Walsh, *Nat Biotechnol* **2018**, 36, 1136.
[2] B. Leader, Q. J. Baca, D. E. Golan, *Nat Rev Drug Discov* **2008**, 7, 21.
[3] J. E. Matthews, M. W. Stewart, E. H. De Boever, R. L. Dobbins, R. J. Hodge, S. E. Walker, M. C. Holland, M. A. Bush, G. Albiglutide Study, *J Clin Endocrinol Metab* **2008**, 93, 4810.
[4] I. Ekladious, Y. L. Colson, M. W. Grinstaff, *Nat Rev Drug Discov* **2019**, 18, 273.
[5] A. Abuchowski, J. R. McCoy, N. C. Palczuk, T. van Es, F. F. Davis, *J Biol Chem* **1977**, 252, 3582.
[6] J. M. Harris, DOI: 10.1007/978-1-4899-0703-5_1, Springer, Boston, MA, 1992.
[7] F. M. Veronese, *PEGylated Protein Drugs: Basic Science and Clinical Applications*, **2009**.
[8] A. Vaisman-Mentesh, M. Gutierrez-Gonzalez, B. J. DeKosky, Y. Wine, *Front Immunol* **2020**, 11, 1951.
[9] G. Lew, *Clin Cancer Res* **2020**, 26, 325.
[10] M. L. Graham, *Adv Drug Deliv Rev* **2003**, 55, 1293.
[11] A. L. Angiolillo, R. J. Schore, M. Devidas, M. J. Borowitz, A. J. Carroll, J. M. Gastier-Foster, N. A. Heerema, T. Keilani, A. R. Lane, M. L. Loh, G. H. Reaman, P. C. Adamson, B. Wood, C. Wood, H. W. Zheng, E. A. Raetz, N. J. Winick, W. L. Carroll, S. P. Hunger, *J Clin Oncol* **2014**, 32, 3874.
[12] S. Chaffee, A. Mary, E. R. Stiehm, D. Girault, A. Fischer, M. S. Hershfield, *J Clin Invest* **1992**, 89, 1643.
[13] L. Wu, J. Chen, Y. Wu, B. Zhang, X. Cai, Z. Zhang, Y. Wang, L. Si, H. Xu, Y. Zheng, C. Zhang, C. Liang, J. Li, L. Zhang, Q. Zhang, D. Zhou, *J Control Release* **2017**, 249, 84.
[14] R. Adams, L. Griffin, J. E. Compson, M. Jairaj, T. Baker, T. Ceska, S. West, O. Zaccheo, E. Dave, A. D. Lawson, D. P. Humphreys, S. Heywood, *MAbs* **2016**, 8, 1336.
[15] A. P. Chapman, *Adv Drug Deliv Rev* **2002**, 54, 531.
[16] M. J. Guichard, T. Leal, R. Vanbever, *Current Opinion in Colloid & Interface Science* **2017**, 31, 43.
[17] A. M. Cantin, D. E. Woods, D. Cloutier, E. K. Dufour, R. Leduc, *Am J Respir Cell Mol Biol* **2002**, 27, 659.
[18] D. Freches, H. P. Patil, M. Machado Franco, C. Uyttenhove, S. Heywood, R. Vanbever, *Int J Pharm* **2017**, 521, 120.
[19] S. J. Koussoroplis, G. Paulissen, D. Tyteca, H. Goldansaz, J. Todoroff, C. Barilly, C. Uyttenhove, J. Van Snick, D. Cataldo, R. Vanbever, *J Control Release* **2014**, 187, 91.
[20] M.-J. Guichard, T. Wilms, S. Mahri, H. P. Patil, D. Hoton, B. Ucakar, K. Vanvarenberg, P. Cheou, M. Beka, E. Marbaix, T. Leal, R. Vanbever, *Advanced Therapeutics* n/a, 2000146.
[21] S. Mahri, A. Rondon, T. Wilms, C. Bosquillon, R. Vanbever, *J Control Release* **2020**, DOI: 10.1016/j.jconrel.2020.10.034.
[22] S. Mahri, E. Hardy, T. Wilms, H. De Keersmaecker, K. Braeckmans, S. De Smedt, C. Bosquillon, R. Vanbever, *Int J Pharm* **2021**, 593, 120107.
[23] Y. Yamamoto, Y. Tsutsumi, Y. Yoshioka, T. Nishibata, K. Kobayashi, T. Okamoto, Y. Mukai, T. Shimizu, S. Nakagawa, S. Nagata, T. Mayumi, *Nature Biotechnology* **2003**, 21, 546.
[24] F. M. Veronese, G. Pasut, *Drug Discov Today* **2005**, 10, 1451.
[25] K. Sprogoe, E. Mortensen, D. B. Karpf, J. A. Leff, *Endocr Connect* **2017**, 6, R171.
[26] a) L. Holten-Andersen, S. Pihl, C. E. Rasmussen, J. Zettler, G. Maitro, J. Baron, S. Heinig, E. Hoffmann, T. Wegge, M. Krusch, F. Faltinger, S. Killian, K. Sprogoe, D. B. Karpf, V. M. Breinholt, F. Cleemann, *J Bone Miner Res* **2019**, 34, 2075; b) V. M. Breinholt, C. E. Rasmussen, P. H.

Mygind, M. Kjelgaard-Hansen, F. Faltinger, A. Bernhard, J. Zettler, U. Hersel, *J Pharmacol Exp Ther* **2019**, 370, 459.
[27] a) F. M. Veronese, A. Mero, F. Caboi, M. Sergi, C. Marongiu, G. Pasut, *Bioconjug Chem* **2007**, 18, 1824; b) M. J. Guichard, H. P. Patil, S. J. Koussoroplis, R. Wattiez, T. Leal, R. Vanbever, *Int J Pharm* **2017**, 524, 159.
[28] C. Roque, A. Sheung, N. Rahman, S. F. Ausar, *Mol Pharm* **2015**, 12, 562.
[29] B. Treetharnmathurot, C. Ovartlarnporn, J. Wungsintaweekul, R. Duncan, R. Wiwattanapatapee, *Int J Pharm* **2008**, 357, 252.
[30] a) P. Caliceti, F. M. Veronese, *Adv Drug Deliv Rev* **2003**, 55, 1261; b) T. Yamaoka, Y. Tabata, Y. Ikada, *J Pharm Sci* **1994**, 83, 601.
[31] a) I. Hamad, A. C. Hunter, J. Szebeni, S. M. Moghimi, *Molecular Immunology* **2008**, 46, 225; b) M. Imran ul-haq, B. F. Lai, R. Chapanian, J. N. Kizhakkedathu, *Biomaterials* **2012**, 33, 9135.
[32] I. A. Ivens, W. Achanzar, A. Baumann, A. Brandli-Baiocco, J. Cavagnaro, M. Dempster, B. O. Depelchin, A. R. Rovira, L. Dill-Morton, J. H. Lane, B. M. Reipert, T. Salcedo, B. Schweighardt, L. S. Tsuruda, P. L. Turecek, J. Sims, *Toxicol Pathol* **2015**, 43, 959.
[33] P. Caliceti, O. Schiavon, F. M. Veronese, *Bioconjug Chem* **2001**, 12, 515.
[34] a) M. Schlapschy, U. Binder, C. Borger, I. Theobald, K. Wachinger, S. Kisling, D. Haller, A. Skerra, *Protein Eng Des Sel* **2013**, 26, 489; b) M. Gebauer, A. Skerra, *Bioorg Med Chem* **2018**, 26, 2882.
[35] V. N. Podust, S. Balan, B. C. Sim, M. P. Coyle, U. Ernst, R. T. Peters, V. Schellenberger, *J Control Release* **2016**, 240, 52.
[36] a) F. Unverdorben, F. Richter, M. Hutt, O. Seifert, P. Malinge, N. Fischer, R. E. Kontermann, *MAbs* **2016**, 8, 120; b) R. Anand, J. Vallooran, in *Engineering of Biomaterials for Drug Delivery Systems*, DOI: https://doi.org/10.1016/B978-0-08-101750-0.00011-8 (Ed: A. Parambath), Woodhead Publishing **2018**, p. 299.
[37] M. Schlapschy, I. Theobald, H. Mack, M. Schottelius, H. J. Wester, A. Skerra, *Protein Eng Des Sel* **2007**, 20, 273.
[38] V. Schellenberger, C. W. Wang, N. C. Geething, B. J. Spink, A. Campbell, W. To, M. D. Scholle, Y. Yin, Y. Yao, O. Bogin, J. L. Cleland, J. Silverman, W. P. Stemmer, *Nat Biotechnol* **2009**, 27, 1186.
[39] a) N. C. Geething, W. To, B. J. Spink, M. D. Scholle, C. W. Wang, Y. Yin, Y. Yao, V. Schellenberger, J. L. Cleland, W. P. Stemmer, J. Silverman, *PLoS One* **2010**, 5, e10175; b) S. E. Alters, B. McLaughlin, B. Spink, T. Lachinyan, C. W. Wang, V. Podust, V. Schellenberger, W. P. Stemmer, *PLoS One* **2012**, 7, e50630; c) A. Haeckel, F. Appler, L. Figge, H. Kratz, M. Lukas, R. Michel, J. Schnorr, M. Zille, B. Hamm, E. Schellenberger, *J Nucl Med* **2014**, 55, 508; d) S. Ding, M. Song, B. C. Sim, C. Gu, V. N. Podust, C. W. Wang, B. McLaughlin, T. P. Shah, R. Lax, R. Gast, R. Sharan, A. Vasek, M. A. Hartman, C. Deniston, P. Srinivas, V. Schellenberger, *Bioconjug Chem* **2014**, 25, 1351; e) A. Haeckel, F. Appler, A. Ariza de Schellenberger, E. Schellenberger, *PLoS One* **2016**, 11, e0157193.
[40] a) J. L. Cleland, N. C. Geething, J. A. Moore, B. C. Rogers, B. J. Spink, C. W. Wang, S. E. Alters, W. P. Stemmer, V. Schellenberger, *J Pharm Sci* **2012**, 101, 2744; b) W. V. Moore, H. J. Nguyen, G. B. Kletter, B. S. Miller, D. Rogers, D. Ng, J. A. Moore, E. Humphriss, J. L. Cleland, G. M. Bright, *J Clin Endocrinol Metab* **2016**, 101, 1091.
[41] a) B. A. Konkle, A. D. Shapiro, D. V. Quon, J. M. Staber, R. Kulkarni, M. V. Ragni, E. S. Chhabra, S. Poloskey, K. Rice, S. Katragadda, J. Fruebis, C. C. Benson, *N Engl J Med* **2020**, 383, 1018; b) E. Seth Chhabra, T. Liu, J. Kulman, S. Patarroyo-White, B. Yang, Q. Lu, D. Drager, N. Moore, J. Liu, A. M. Holthaus, J. M. Sommer, A. Ismail, D. Rabinovich, Z. Liu, A. van der Flier, A. Goodman, C. Furcht, M. Tie, T. Carlage, R. Mauldin, T. M. Dobrowsky, Z. Liu, O. Mercury, L. Zhu, B. Mei, V. Schellenberger, H. Jiang, G. F. Pierce, J. Salas, R. Peters, *Blood* **2020**, 135, 1484; c) B. Nolan, J. Mahlangu, D. Perry, G. Young, R. Liesner, B. Konkle, S. Rangarajan, S. Brown, H. Hanabusa, K. J. Pasi, I. Pabinger, S. Jackson, L. M. Cristiano, X. Li, G. F. Pierce, G. Allen, *Haemophilia* **2016**, 22, 72.
[42] A. Varanko, S. Saha, A. Chilkoti, *Adv Drug Deliv Rev* **2020**, 156, 133.

[43] P. Hoffmann, R. Hofmeister, K. Brischwein, C. Brandl, S. Crommer, R. Bargou, C. Itin, N. Prang, P. A. Baeuerle, *Int J Cancer* **2005**, 115, 98.
[44] U. Binder, A. Skerra, *Current Opinion in Colloid & Interface Science* **2017**, 31, 10.
[45] a) D. Harari, N. Kuhn, R. Abramovich, K. Sasson, A. Zozulya, P. Smith, M. Schlapschy, R. Aharoni, M. Köster, R. Eliam, A. Skerra, G. Schreiber, *Journal of Biological Chemistry* **2014**, 289, 29014; b) E. A. Zvonova, A. V. Ershov, O. A. Ershova, M. A. Sudomoina, M. B. Degterev, G. N. Poroshin, A. V. Eremeev, A. P. Karpov, A. Y. Vishnevsky, I. V. Goldenkova-Pavlova, A. V. Petrov, S. V. Ruchko, A. M. Shuster, *Appl Microbiol Biotechnol* **2017**, 101, 1975; c) F. Bolze, A. Bast, S. Mocek, V. Morath, D. Yuan, N. Rink, M. Schlapschy, A. Zimmermann, M. Heikenwalder, A. Skerra, M. Klingenspor, *Diabetologia* **2016**, 59, 2005; d) F. Bolze, V. Morath, A. Bast, N. Rink, M. Schlapschy, S. Mocek, A. Skerra, M. Klingenspor, *Endocrinology* **2016**, 157, 233; e) V. Morath, F. Bolze, M. Schlapschy, S. Schneider, F. Sedlmayer, K. Seyfarth, M. Klingenspor, A. Skerra, *Mol Pharm* **2015**, 12, 1431; f) C. T. Mendler, A. Feuchtinger, I. Heid, M. Aichler, C. D'Alessandria, S. Pirsig, B. Blechert, H. J. Wester, R. Braren, A. Walch, A. Skerra, M. Schwaiger, *J Nucl Med* **2016**, 57, 1971; g) C. T. Mendler, L. Friedrich, I. Laitinen, M. Schlapschy, M. Schwaiger, H. J. Wester, A. Skerra, *MAbs* **2015**, 7, 96; h) C. T. Mendler, T. Gehring, H. J. Wester, M. Schwaiger, A. Skerra, *J Nucl Med* **2015**, 56, 1112; i) M. J. Harder, N. Kuhn, H. Schrezenmeier, B. Hochsmann, I. von Zabern, C. Weinstock, T. Simmet, D. Ricklin, J. D. Lambris, A. Skerra, M. Anliker, C. Q. Schmidt, *Blood* **2017**, 129, 970; j) N. Kuhn, C. Q. Schmidt, M. Schlapschy, A. Skerra, *Bioconjug Chem* **2016**, 27, 2359; k) M. H. Hedayati, D. Norouzian, M. Aminian, S. Teimourian, R. Ahangari Cohan, S. Sardari, M. R. Khorramizadeh, *Protein J* **2017**, 36, 36; l) J. Breibeck, A. Skerra, *Biopolymers* **2018**, 109; m) N. E. Powers, B. Swartzwelter, C. Marchetti, D. M. de Graaf, A. Lerchner, M. Schlapschy, R. Datar, U. Binder, C. K. Edwards, 3rd, A. Skerra, C. A. Dinarello, *J Biol Chem* **2020**, 295, 868; n) S. Mazaheri, Y. Talebkhan, F. Mahboudi, L. Nematollahi, R. A. Cohan, E. Mirabzadeh Ardakani, E. Bayat, M. Sabzalinejad, S. Sardari, F. Torkashvand, *Sci Rep* **2020**, 10, 18464; o) F. Khodabakhsh, D. Norouzian, B. Vaziri, R. Ahangari Cohan, S. Sardari, F. Mahboudi, M. Behdani, K. Mansouri, A. Mehdizadeh, *Artif Cells Nanomed Biotechnol* **2018**, 46, 1402; p) F. Khodabakhsh, M. Salimian, A. Mehdizadeh, M. S. Khosravy, A. Vafabakhsh, E. Karami, R. A. Cohan, *J Pharmacol Exp Ther* **2020**, 375, 69; q) E. Eggenstein, A. Richter, A. Skerra, *Protein Eng Des Sel* **2019**, 32, 289.
[46] F. C. Deuschle, V. Morath, A. Schiefner, C. Brandt, S. Ballke, S. Reder, K. Steiger, M. Schwaiger, W. Weber, A. Skerra, *Theranostics* **2020**, 10, 2172.
[47] E. Peplau, F. De Rose, S. Reder, M. Mittelhauser, G. Scafetta, M. Schwaiger, W. A. Weber, A. Bartolazzi, A. Skerra, C. D'Alessandria, *Thyroid* **2020**, 30, 1314.
[48] A. Richter, K. Knorr, M. Schlapschy, S. Robu, V. Morath, C. Mendler, H. Y. Yen, K. Steiger, M. Kiechle, W. Weber, A. Skerra, M. Schwaiger, *Nucl Med Mol Imaging* **2020**, 54, 114.
[49] E. Falvo, E. Tremante, A. Arcovito, M. Papi, N. Elad, A. Boffi, V. Morea, G. Conti, G. Toffoli, G. Fracasso, P. Giacomini, P. Ceci, *Biomacromolecules* **2016**, 17, 514.
[50] G. Fracasso, E. Falvo, G. Colotti, F. Fazi, T. Ingegnere, A. Amalfitano, G. B. Doglietto, S. Alfieri, A. Boffi, V. Morea, G. Conti, E. Tremante, P. Giacomini, A. Arcovito, P. Ceci, *Journal of Controlled Release* **2016**, 239, 10.
[51] E. Falvo, F. Malagrino, A. Arcovito, F. Fazi, G. Colotti, E. Tremante, P. Di Micco, A. Braca, R. Opri, A. Giuffre, G. Fracasso, P. Ceci, *J Control Release* **2018**, 275, 177.
[52] B. Tesarova, S. Dostalova, V. Smidova, Z. Goliasova, Z. Skubalova, H. Michalkova, H. David, P. Michalek, H. Polanska, M. Vacuovicova, J. Hacek, T. Eckschlager, M. Stiborova, A. S. Pires, A. R. M. Neves, A. M. Abrantes, T. Rodrigues, P. Matafome, M. F. Botelho, P. Teixeira, F. Mendes, Z. Heger, *Applied Metarials Today* **2020**, 18, 100501.
[53] F. Brandl, H. Merten, M. Zimmermann, M. Behe, U. Zangemeister-Wittke, A. Pluckthun, *J Control Release* **2019**, 307, 379.
[54] A. P. Caputi, P. Navarra, *Curr Opin Pharmacol* **2020**, 51, 93.
[55] D. M. Floss, K. Schallau, S. Rose-John, U. Conrad, J. Scheller, *Trends Biotechnol* **2010**, 28, 37.
[56] C. E. Mills, Z. Michaud, B. D. Olsen, *Biomacromolecules* **2018**, 19, 2517.

[57] A. Varanko, S. Saha, A. Chilkoti, *Advanced drug delivery reviews* **2020**, 156, 133.
[58] W. R. Strohl, *BioDrugs* **2015**, 29, 215.
[59] U. Conrad, I. Plagmann, S. Malchow, M. Sack, D. M. Floss, A. A. Kruglov, S. A. Nedospasov, S. Rose-John, J. Scheller, *Plant Biotechnol J* **2011**, 9, 22.
[60] W. H. Toh, J. Louber, I. S. Mahmoud, J. Chia, G. T. Bass, S. K. Dower, A. M. Verhagen, P. A. Gleeson, *J Cell Sci* **2019**, 133.
[61] K. M. Murphy, C. Weaver, *Janeway's Immunobiology: Ninth International Student Edition*, Garland Science, Taylor & Francis Group, LLC, **2016**.
[62] G. Vidarsson, G. Dekkers, T. Rispens, *Front Immunol* **2014**, 5, 520.
[63] M. Cavaco, M. Castanho, V. Neves, *Biopolymers* **2017**, DOI: 10.1002/bip.23095.
[64] P. R. Hinton, M. G. Johlfs, J. M. Xiong, K. Hanestad, K. C. Ong, C. Bullock, S. Keller, M. T. Tang, J. Y. Tso, M. Vasquez, N. Tsurushita, *J Biol Chem* **2004**, 279, 6213.
[65] W. F. Dall'Acqua, R. M. Woods, E. S. Ward, S. R. Palaszynski, N. K. Patel, Y. A. Brewah, H. Wu, P. A. Kiener, S. Langermann, *J Immunol* **2002**, 169, 5171.
[66] W. F. Dall'Acqua, P. A. Kiener, H. Wu, *J Biol Chem* **2006**, 281, 23514.
[67] a) R. S. Blumberg, D. Lillicrap, G. F. I. T. G. Ig, *Blood* **2018**, 131, 2205; b) A. Kessel, H. Ammuri, R. Peri, E. R. Pavlotzky, M. Blank, Y. Shoenfeld, E. Toubi, *J Immunol* **2007**, 179, 5571.
[68] S. Zhao, E. Mysler, R. J. Moots, *Immunotherapy* **2018**, 10, 433.
[69] A. Srivastava, A. K. Brewer, E. P. Mauser-Bunschoten, N. S. Key, S. Kitchen, A. Llinas, C. A. Ludlam, J. N. Mahlangu, K. Mulder, M. C. Poon, A. Street, H. Treatment Guidelines Working Group on Behalf of The World Federation Of, *Haemophilia* **2013**, 19, e1.
[70] L. Graf, *Transfus Med Hemother* **2018**, 45, 86.
[71] a) J. S. Powell, N. C. Josephson, D. Quon, M. V. Ragni, G. Cheng, E. Li, H. Jiang, L. Li, J. A. Dumont, J. Goyal, X. Zhang, J. Sommer, J. McCue, M. Barbetti, A. Luk, G. F. Pierce, *Blood* **2012**, 119, 3031; b) G. Young, J. Mahlangu, R. Kulkarni, B. Nolan, R. Liesner, J. Pasi, C. Barnes, S. Neelakantan, G. Gambino, L. M. Cristiano, G. F. Pierce, G. Allen, *J Thromb Haemost* **2015**, 13, 967.
[72] S. L. Meeks, S. Lacroix-Desmazes, *Haemophilia* **2020**, DOI: 10.1111/hae.14123.
[73] A. Zorzi, S. Linciano, A. Angelini, *Medchemcomm* **2019**, 10, 1068.
[74] M. Bern, K. M. Sand, J. Nilsen, I. Sandlie, J. T. Andersen, *J Control Release* **2015**, 211, 144.
[75] D. Sleep, J. Cameron, L. R. Evans, *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, 1830, 5526.
[76] a) J. T. Andersen, B. Dalhus, D. Viuff, B. T. Ravn, K. S. Gunnarsen, A. Plumridge, K. Bunting, F. Antunes, R. Williamson, S. Athwal, E. Allan, L. Evans, M. Bjørås, S. Kjærulff, D. Sleep, I. Sandlie, J. Cameron, *Journal of Biological Chemistry* **2014**, 289, 13492; b) M. Bern, J. Nilsen, M. Ferrarese, K. M. K. Sand, T. T. Gjølberg, H. E. Lode, R. J. Davidson, R. M. Camire, E. S. Bækkevold, S. Foss, A. Grevys, B. Dalhus, J. Wilson, L. S. Høydahl, G. J. Christianson, D. C. Roopenian, T. Schlothauer, T. E. Michaelsen, M. C. Moe, S. Lombardi, M. Pinotti, I. Sandlie, A. Branchini, J. T. Andersen, *Science Translational Medicine* **2020**, 12.
[77] S. S. Petersen, E. Klaning, M. F. Ebbesen, B. Andersen, J. Cameron, E. S. Sorensen, K. A. Howard, *Mol Pharm* **2016**, 13, 677.
[78] V. E. Kenanova, T. Olafsen, F. B. Salazar, L. E. Williams, S. Knowles, A. M. Wu, *Protein Engineering Design and Selection* **2010**, 23, 789.
[79] M. Bak, J. Park, K. Min, J. Cho, J. Seong, Y. S. Hahn, G. Tae, I. Kwon, *Pharmaceutics* **2020**, 12.
[80] a) J. Schmøkel, A. Voldum, G. Tsakiridou, M. Kuhlmann, J. Cameron, E. S. Sørensen, J. Wengel, K. A. Howard, *Nanotechnology* **2017**, 28; b) M. Simon, R. Frey, U. Zangemeister-Wittke, A. Plückthun, *Bioconjugate Chemistry* **2013**, 24, 1955; c) D. Müller, A. Karle, B. Meißburger, I. Höfig, R. Stork, R. E. Kontermann, *Journal of Biological Chemistry* **2007**, 282, 12650.
[81] E. M. Bech, S. L. Pedersen, K. J. Jensen, *ACS Medicinal Chemistry Letters* **2018**, 9, 577.
[82] J. Lau, P. Bloch, L. Schaffer, I. Pettersson, J. Spetzler, J. Kofoed, K. Madsen, L. B. Knudsen, J. McGuire, D. B. Steensgaard, H. M. Strauss, D. X. Gram, S. M. Knudsen, F. S. Nielsen, P. Thygesen, S. Reedtz-Runge, T. Kruse, *J Med Chem* **2015**, 58, 7370.

[83] R. Juul Kildemoes, M. Hojby Rasmussen, H. Agerso, R. V. Overgaard, *J Clin Endocrinol Metab* **2021**, 106, 567.
[84] J. Cho, J. Park, G. Tae, M. S. Jin, I. Kwon, *Biomedicines* **2020**, 8.
[85] L. B. Knudsen, J. Lau, *Front Endocrinol (Lausanne)* **2019**, 10, 155.
[86] S. I. Lim, Y. Mizuta, A. Takasu, Y. S. Hahn, Y. H. Kim, I. Kwon, *J Control Release* **2013**, 170, 219.
[87] J. Cho, S. I. Lim, B. S. Yang, Y. S. Hahn, I. Kwon, *Sci Rep* **2017**, 7, 18041.
[88] C. Fu, Q. Chen, F. Zheng, L. Yang, H. Li, Q. Zhao, X. Wang, L. Wang, Q. Wang, *Angew Chem Int Ed Engl* **2019**, 58, 1392.
[89] J. Cho, J. Park, S. Kim, J. C. Kim, G. Tae, M. S. Jin, I. Kwon, *J Control Release* **2020**, 321, 49.
[90] M. U. Johansson, I.-M. Frick, H. Nilsson, P. J. Kraulis, S. Hober, P. Jonasson, M. Linhult, P.-Å. Nygren, M. Uhlén, L. Björck, T. Drakenberg, S. Forsén, M. Wikström, *Journal of Biological Chemistry* **2002**, 277, 8114.
[91] a) A. Jonsson, J. Dogan, N. Herne, L. Abrahmsen, P. A. Nygren, *Protein Engineering Design and Selection* **2008**, 21, 515; b) J. T. Andersen, R. Pehrson, V. Tolmachev, M. B. Daba, L. Abrahmsen, C. Ekblad, *J Biol Chem* **2011**, 286, 5234.
[92] R. Guo, W. Guo, L. Cao, H. Liu, J. Liu, H. Xu, W. Huang, F. Wang, Z. Hong, *International Journal of Pharmaceutics* **2016**, 511, 538.
[93] A. Barozzi, R. A. Lavoie, K. N. Day, R. Prodromou, S. Menegatti, *Int J Mol Sci* **2020**, 21.
[94] a) B. M. Tijink, T. Laeremans, M. Budde, M. S. v. Walsum, T. Dreier, H. J. de Haard, C. R. Leemans, G. A. M. S. van Dongen, *Molecular Cancer Therapeutics* **2008**, 7, 2288; b) H. van Faassen, S. Ryan, K. A. Henry, S. Raphael, Q. Yang, M. A. Rossotti, E. Brunette, S. Jiang, A. S. Haqqani, T. Sulea, C. R. MacKenzie, J. Tanha, G. Hussack, *The FASEB Journal* **2020**, 34, 8155; c) M. R. Müller, K. Saunders, C. Grace, M. Jin, N. Piche-Nicholas, J. Steven, R. O'Dwyer, L. Wu, L. Khetemenee, Y. Vugmeyster, T. P. Hickling, L. Tchistiakova, S. Olland, D. Gill, A. Jensen, C. J. Barelle, *mAbs* **2014**, 4, 673; d) S. Y. Cho, J. Han, S.-H. Cha, S.-i. Yoon, *Biochemical and Biophysical Research Communications* **2020**, 526, 941; e) D. Kim, H. Jeon, S. Ahn, W. I. Choi, S. Kim, S. Jon, *Journal of Controlled Release* **2017**, 256, 114; f) D. Steiner, F. W. Merz, I. Sonderegger, M. Gulotti-Georgieva, D. Villemagne, D. J. Phillips, P. Forrer, M. T. Stumpp, C. Zitt, H. K. Binz, *Protein Engineering, Design and Selection* **2017**, 30, 583; g) M. W. Traxlmayr, J. D. Kiefer, R. R. Srinivas, E. Lobner, A. W. Tisdale, N. K. Mehta, N. J. Yang, B. Tidor, K. D. Wittrup, *Journal of Biological Chemistry* **2016**, 291, 22496.
[95] a) K. A. Henry, J. Tanha, G. Hussack, *Protein Engineering Design and Selection* **2015**, 28, 379; b) E. Davé, R. Adams, O. Zaccheo, B. Carrington, J. E. Compson, S. Dugdale, M. Airey, S. Malcolm, H. Hailu, G. Wild, A. Turner, J. Heads, K. Sarkar, A. Ventom, D. Marshall, M. Jairaj, T. Kopotsha, L. Christodoulou, M. Zamacona, A. D. Lawson, S. Heywood, D. P. Humphreys, *mAbs* **2016**, 8, 1319.
[96] C. Pain, J. Dumont, M. Dumoulin, *Biochimie* **2015**, 111, 82.
[97] a) G. Bao, M. Tang, J. Zhao, X. Zhu, *EJNMMI Res* **2021**, 11, 6; b) N. V. Bathula, H. Bommadevara, J. M. Hayes, *Cancer Biother Radiopharm* **2021**, 36, 109; c) I. Jovcevska, S. Muyldermans, *BioDrugs* **2020**, 34, 11.
[98] a) S. Muyldermans, *J Biotechnol* **2001**, 74, 277; b) M. Dumoulin, K. Conrath, A. Van Meirhaeghe, F. Meersman, K. Heremans, L. G. Frenken, S. Muyldermans, L. Wyns, A. Matagne, *Protein Sci* **2002**, 11, 500.
[99] a) S. Hoefman, I. Ottevaere, J. Baumeister, M. L. Sargentini-Maier, *Antibodies* **2015**, 141; b) M. Van Roy, C. Ververken, E. Beirnaert, S. Hoefman, J. Kolkman, M. Vierboom, E. Breedveld, B. t Hart, S. Poelmans, L. Bontinck, A. Hemeryck, S. Jacobs, J. Baumeister, H. Ulrichts, *Arthritis Res Ther* **2015**, 17, 135.
[100] F. Kratz, B. Elsadek, *J Control Release* **2012**, 161, 429.
[101] S. Low, H. Wu, K. Jerath, A. Tibolla, B. Fogal, R. Conrad, M. MacDougall, S. Kerr, V. Berger, R. Dave, J. Villalona, L. Pantages, J. Ahlberg, H. Li, D. Van Hoorick, C. Ververken, J. Broadwater, A. Waterman, S. Singh, R. Kroe-Barrett, *MAbs* **2020**, 12, 1709322.
[102] K. A. Papp, M. A. Weinberg, A. Morris, K. Reich, *Lancet* **2021**, 397, 1564.

[103] C. Vincke, R. Loris, D. Saerens, S. Martinez-Rodriguez, S. Muyldermans, K. Conrath, *J Biol Chem* **2009**, 284, 3273.
[104] K. Fan, B. Jiang, Z. Guan, J. He, D. Yang, N. Xie, G. Nie, C. Xie, X. Yan, *Anal Chem* **2018**, 90, 5671.
[105] R. L. O'Connor-Semmes, J. Lin, R. J. Hodge, S. Andrews, J. Chism, A. Choudhury, D. J. Nunez, *Clinical Pharmacology & Therapeutics* **2014**, 96, 704.
[106] H. K. Binz, T. R. Bakker, D. J. Phillips, A. Cornelius, C. Zitt, T. Gottler, G. Sigrist, U. Fiedler, S. Ekawardhani, I. Dolado, J. A. Saliba, G. Tresch, K. Proba, M. T. Stumpp, *MAbs* **2017**, 9, 1262.
[107] R. D. Baird, C. Linossi, M. Middleton, S. Lord, A. Harris, J. Rodon, C. Zitt, U. Fiedler, K. M. Dawson, N. Leupin, M. T. Stumpp, A. Harstrick, A. Azaro, S. Fischer, A. Omlin, *J Clin Oncol* **2021**, 39, 145.
[108] Z. R. Crook, N. W. Nairn, J. M. Olson, *Trends in Biochemical Sciences* **2020**, 45, 332.
[109] T. Dörner, M. Weinblatt, K. V. Beneden, E. Dombrecht, K. D. Beuf, P. Schoen, R. Zeldin, *Annals of the Rheumatic Diseases* **2017**, 76, 575.
[110] M. Brinkley, *Bioconjug Chem* **1992**, 3, 2.
[111] F. Tekaia, E. Yeramian, B. Dujon, *Gene* **2002**, 297, 51.
[112] a) R. Huisgen, *Angewandte Chemie International Edition in English* **1963**, 2, 633; b) P. Shieh, C. R. Bertozzi, *Org Biomol Chem* **2014**, 12, 9307.
[113] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, 266, 776; b) C. B. Rosen, M. B. Francis, *Nat Chem Biol* **2017**, 13, 697.
[114] E. Jacob, R. Unger, *Bioinformatics* **2007**, 23, e225.
[115] a) J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi, M. B. Francis, *Angew Chem Int Ed Engl* **2006**, 45, 5307; b) H. Jiang, G. D. D'Agostino, P. A. Cole, D. R. Dempsey, *Methods Enzymol* **2020**, 639, 333; c) T. J. Sereda, C. T. Mant, A. M. Quinn, R. S. Hodges, *J Chromatogr* **1993**, 646, 17.
[116] P. L. Turecek, M. J. Bossard, F. Schoetens, I. A. Ivens, *J Pharm Sci* **2016**, 105, 460.
[117] a) M. W. Jones, G. Mantovani, C. A. Blindauer, S. M. Ryan, X. Wang, D. J. Brayden, D. M. Haddleton, *J Am Chem Soc* **2012**, 134, 7406; b) M. W. Jones, R. A. Strickland, F. F. Schumacher, S. Caddick, J. R. Baker, M. I. Gibson, D. M. Haddleton, *J Am Chem Soc* **2012**, 134, 1847; c) Y. Cong, E. Pawlisz, P. Bryant, S. Balan, E. Laurine, R. Tommasi, R. Singh, S. Dubey, K. Peciak, M. Bird, A. Sivasankar, J. Swierkosz, M. Muroni, S. Heidelberger, M. Farys, F. Khayrzad, J. Edwards, G. Badescu, I. Hodgson, C. Heise, S. Somavarapu, J. Liddell, K. Powell, M. Zloh, J. W. Choi, A. Godwin, S. Brocchini, *Bioconjug Chem* **2012**, 23, 248; d) M. A. Gauthier, H. A. Klok, *Biomacromolecules* **2011**, 12, 482.
[118] N. Stephanopoulos, M. B. Francis, *Nat Chem Biol* **2011**, 7, 876.
[119] H. Lodish, B. Arnold, M. Paul, C. Kaiser, S. M. P, Z. LawrenceL, D. James, in *Molecular Cell Biology* (Ed: W. H. Freman) **2000**, p. 60.
[120] S. L. Ho, A. H. Wang, *J Taiwan Inst Chem Eng* **2009**, 40, 123.
[121] a) F. Hatahet, L. W. Ruddock, *Antioxid Redox Signal* **2009**, 11, 2807; b) G. Rajpal, P. Arvan, in *Handbook of Biologically Active Peptides*, DOI: 10.1016/b978-0-12-385095-9.00236-0 **2013**, p. 1721.
[122] W. H. So, Y. Zhang, W. Kang, C. T. T. Wong, H. Sun, J. Xia, *Curr Opin Biotechnol* **2017**, 48, 220.
[123] a) J. Ravasco, H. Faustino, A. Trindade, P. M. P. Gois, *Chemistry* **2019**, 25, 43; b) K. Renault, J. W. Fredy, P. Y. Renard, C. Sabot, *Bioconjug Chem* **2018**, 29, 2497.
[124] C. P. Ryan, M. E. Smith, F. F. Schumacher, D. Grohmann, D. Papaioannou, G. Waksman, F. Werner, J. R. Baker, S. Caddick, *Chem Commun (Camb)* **2011**, 47, 5452.
[125] a) M. E. Smith, M. B. Caspersen, E. Robinson, M. Morais, A. Maruani, J. P. Nunes, K. Nicholls, M. J. Saxton, S. Caddick, J. R. Baker, V. Chudasama, *Org Biomol Chem* **2015**, 13, 7946; b) M. E. Smith, F. F. Schumacher, C. P. Ryan, L. M. Tedaldi, D. Papaioannou, G. Waksman, S. Caddick, J. R. Baker, *J Am Chem Soc* **2010**, 132, 1960; c) J. P. Nunes, M. Morais, V. Vassileva, E. Robinson, V. S. Rajkumar, M. E. Smith, R. B. Pedley, S. Caddick, J. R. Baker, V. Chudasama, *Chem Commun (Camb)* **2015**, 51, 10624.

[126] A. Maruani, M. E. Smith, E. Miranda, K. A. Chester, V. Chudasama, S. Caddick, *Nat Commun* **2015**, 6, 6645.
[127] D. A. Richards, S. A. Fletcher, M. Nobles, H. Kossen, L. Tedaldi, V. Chudasama, A. Tinker, J. R. Baker, *Org Biomol Chem* **2016**, 14, 455.
[128] M. Lahnsteiner, A. Kastner, J. Mayr, A. Roller, B. K. Keppler, C. R. Kowol, *Chemistry* **2020**, 26, 15867.
[129] P. A. Szijj, C. Bahou, V. Chudasama, *Drug Discov Today Technol* **2018**, 30, 27.
[130] W. Huang, X. Wu, X. Gao, Y. Yu, H. Lei, Z. Zhu, Y. Shi, Y. Chen, M. Qin, W. Wang, Y. Cao, *Nat Chem* **2019**, 11, 310.
[131] B. Q. Shen, K. Xu, L. Liu, H. Raab, S. Bhakta, M. Kenrick, K. L. Parsons-Reponte, J. Tien, S. F. Yu, E. Mai, D. Li, J. Tibbitts, J. Baudys, O. M. Saad, S. J. Scales, P. J. McDonald, P. E. Hass, C. Eigenbrot, T. Nguyen, W. A. Solis, R. N. Fuji, K. M. Flagella, D. Patel, S. D. Spencer, L. A. Khawli, A. Ebens, W. L. Wong, R. Vandlen, S. Kaur, M. X. Sliwkowski, R. H. Scheller, P. Polakis, J. R. Junutula, *Nat Biotechnol* **2012**, 30, 184.
[132] J. R. Junutula, H. Raab, S. Clark, S. Bhakta, D. D. Leipold, S. Weir, Y. Chen, M. Simpson, S. P. Tsai, M. S. Dennis, Y. Lu, Y. G. Meng, C. Ng, J. Yang, C. C. Lee, E. Duenas, J. Gorrell, V. Katta, A. Kim, K. McDorman, K. Flagella, R. Venook, S. Ross, S. D. Spencer, W. Lee Wong, H. B. Lowman, R. Vandlen, M. X. Sliwkowski, R. H. Scheller, P. Polakis, W. Mallet, *Nat Biotechnol* **2008**, 26, 925.
[133] E. V. Vinogradova, C. Zhang, A. M. Spokoyny, B. L. Pentelute, S. L. Buchwald, *Nature* **2015**, 526, 687.
[134] a) K. Kubota, P. Dai, B. L. Pentelute, S. L. Buchwald, *J Am Chem Soc* **2018**, 140, 3128; b) H. H. Dhanjee, A. Saebi, I. Buslov, A. R. Loftis, S. L. Buchwald, B. L. Pentelute, *J Am Chem Soc* **2020**, 142, 9124.
[135] L. Zhang, J. P. Tam, *Anal Biochem* **1996**, 233, 87.
[136] a) G. J. Bernardes, M. Steiner, I. Hartmann, D. Neri, G. Casi, *Nat Protoc* **2013**, 8, 2079; b) G. Casi, N. Huguenin-Dezot, K. Zuberbuhler, J. Scheuermann, D. Neri, *J Am Chem Soc* **2012**, 134, 5887.
[137] A. Bandyopadhyay, S. Cambray, J. Gao, *Chem Sci* **2016**, 7, 4589.
[138] A. R. Costa, M. E. Rodrigues, M. Henriques, R. Oliveira, J. Azeredo, *Crit Rev Biotechnol* **2014**, 34, 281.
[139] Z. Liu, J. Ou, H. Wang, X. You, M. Ye, *ACS Appl Mater Interfaces* **2016**, 8, 32060.
[140] a) L. Liu, M. Yu, Y. Zhang, C. Wang, H. Lu, *ACS Appl Mater Interfaces* **2014**, 6, 7823; b) W. Miao, C. Zhang, Y. Cai, Y. Zhang, H. Lu, *Analyst* **2016**, 141, 2435; c) D. S. Bai, C. Zhang, P. Chen, S. J. Jin, G. Q. Jiang, *Sci Rep* **2017**, 7, 12870.
[141] a) B. Han, J. F. Stevens, C. S. Maier, *Anal Chem* **2007**, 79, 3342; b) P. P. Joshi, S. J. Yoon, W. G. Hardin, S. Emelianov, K. V. Sokolov, *Bioconjug Chem* **2013**, 24, 878.
[142] a) L. Nieba, S. E. Nieba-Axmann, A. Persson, M. Hamalainen, F. Edebratt, A. Hansson, J. Lidholm, K. Magnusson, A. F. Karlsson, A. Pluckthun, *Anal Biochem* **1997**, 252, 217; b) C. Cherkouk, L. Rebohle, J. Lenk, A. Keller, X. Ou, M. Laube, C. Neuber, C. Haase-Kohn, W. Skorupa, J. Pietzsch, *Clin Hemorheol Microcirc* **2015**, 61, 523; c) S. Knecht, D. Ricklin, A. N. Eberle, B. Ernst, *J Mol Recognit* **2009**, 22, 270.
[143] E. Hochuli, H. Dobeli, A. Schacher, *J Chromatogr* **1987**, 411, 177.
[144] a) S. H. Uchinomiya, H. Nonaka, S. H. Fujishima, S. Tsukiji, A. Ojida, I. Hamachi, *Chem Commun (Camb)* **2009**, DOI: 10.1039/b912025d5880; b) Y. Yang, N. Jiang, Y. T. Lai, Y. Y. Chang, X. Yang, H. Sun, H. Li, *ACS Sens* **2019**, 4, 1190; c) C. J. Ackerson, R. D. Powell, J. F. Hainfeld, *Methods Enzymol* **2010**, 481, 195.
[145] S. V. Wegner, J. P. Spatz, *Angew Chem Int Ed Engl* **2013**, 52, 7593.
[146] E. M. Sletten, C. R. Bertozzi, *Angew Chem Int Ed Engl* **2009**, 48, 6974.
[147] A. Rondon, F. Degoul, *Bioconjug Chem* **2020**, 31, 159.
[148] M. L. Blackman, M. Royzen, J. M. Fox, *J Am Chem Soc* **2008**, 130, 13518.
[149] a) R. Rossin, S. M. van den Bosch, W. Ten Hoeve, M. Carvelli, R. M. Versteegen, J. Lub, M. S. Robillard, *Bioconjug Chem* **2013**, 24, 1210; b) J. B. Bequignat, N. Ty, A. Rondon, L. Taiariol, F.

Degoul, D. Canitrot, M. Quintana, I. Navarro-Teulon, E. Miot-Noirault, C. Boucheix, J. M. Chezal, E. Moreau, *Eur J Med Chem* **2020**, 203, 112574.
[150] M. Royzen, G. P. Yap, J. M. Fox, *J Am Chem Soc* **2008**, 130, 3760.
[151] T. Luhmann, M. Gutmann, A. Moscaroli, M. Raschig, M. Behe, L. Meinel, *ACS Biomater Sci Eng* **2020**, 6, 425.
[152] a) S. Mushtaq, S. J. Yun, J. Jeon, *Molecules* **2019**, 24; b) N. K. Devaraj, *ACS Cent Sci* **2018**, 4, 952.
[153] a) E. S. Zimmerman, T. H. Heibeck, A. Gill, X. Li, C. J. Murray, M. R. Madlansacay, C. Tran, N. T. Uter, G. Yin, P. J. Rivers, A. Y. Yam, W. D. Wang, A. R. Steiner, S. U. Bajad, K. Penta, W. Yang, T. J. Hallam, C. D. Thanos, A. K. Sato, *Bioconjug Chem* **2014**, 25, 351; b) C. Koehler, P. F. Sauter, M. Wawryszyn, G. E. Girona, K. Gupta, J. J. Landry, M. H. Fritz, K. Radic, J. E. Hoffmann, Z. A. Chen, J. Zou, P. S. Tan, B. Galik, S. Junttila, P. Stolt-Bergner, G. Pruneri, A. Gyenesei, C. Schultz, M. B. Biskup, H. Besir, V. Benes, J. Rappsilber, M. Jechlinger, J. O. Korbel, I. Berger, S. Braese, E. A. Lemke, *Nat Methods* **2016**, 13, 997.
[154] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, 292, 498.
[155] T. S. Young, P. G. Schultz, *J Biol Chem* **2010**, 285, 11039.
[156] a) D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters, J. W. Chin, *J Am Chem Soc* **2009**, 131, 8720; b) W. Liu, A. Brock, S. Chen, S. Chen, P. G. Schultz, *Nat Methods* **2007**, 4, 239; c) M. L. Tsao, D. Summerer, Y. Ryu, P. G. Schultz, *J Am Chem Soc* **2006**, 128, 4572.
[157] L. Jakob, A. Gust, D. Grohmann, *Biochem Biophys Rep* **2019**, 17, 1.
[158] K. Ozawa, K. V. Loscha, K. V. Kuppan, C. T. Loh, N. E. Dixon, G. Otting, *Biochem Biophys Res Commun* **2012**, 418, 652.
[159] J. A. Prescher, C. R. Bertozzi, *Nat Chem Biol* **2005**, 1, 13.
[160] Y. Zhang, K. Y. Park, K. F. Suazo, M. D. Distefano, *Chem Soc Rev* **2018**, 47, 9106.
[161] a) Y. Wang, C. Wu, *Biomacromolecules* **2018**, 19, 1804; b) D. Schumacher, J. Helma, A. F. L. Schneider, H. Leonhardt, C. P. R. Hackenberger, *Angew Chem Int Ed Engl* **2018**, 57, 2314; c) T. Pleiner, M. Bates, S. Trakhanov, C. T. Lee, J. E. Schliep, H. Chug, M. Bohning, H. Stark, H. Urlaub, D. Gorlich, *Elife* **2015**, 4, e11349.
[162] T. Sano, S. Vajda, C. R. Cantor, *J Chromatogr B Biomed Sci Appl* **1998**, 715, 85.
[163] A. Chapman-Smith, J. E. Cronan, Jr., *Biomol Eng* **1999**, 16, 119.
[164] M. Fairhead, M. Howarth, *Methods Mol Biol* **2015**, 1266, 171.
[165] a) A. Kulyyassov, V. Ogryzko, *Biomolecules* **2020**, 10; b) K. Schilders, E. Eenjes, G. Edel, A. B. de Munck, M. B. van Kempen, J. Demmers, R. Wijnen, D. Tibboel, R. J. Rottier, *Transgenic Res* **2018**, 27, 75; c) S. Maiti, P. Paira, *Eur J Med Chem* **2018**, 145, 206.
[166] H. P. Lesch, M. U. Kaikkonen, J. T. Pikkarainen, S. Yla-Herttuala, *Expert Opin Drug Deliv* **2010**, 7, 551.
[167] a) N. Pishesha, J. R. Ingram, H. L. Ploegh, *Annual Review of Cell and Developmental Biology* **2018**, 34, 163; b) F. B. H. Rehm, T. J. Harmand, K. Yap, T. Durek, D. J. Craik, H. L. Ploegh, *Journal of the American Chemical Society* **2019**, 141, 17388.
[168] T. J. Harmand, D. Bousbaine, A. Chan, X. Zhang, D. R. Liu, J. P. Tam, H. L. Ploegh, *Bioconjugate Chemistry* **2018**, 29, 3245.
[169] a) C. P. Guimaraes, M. D. Witte, C. S. Theile, G. Bozkurt, L. Kundrat, A. E. Blom, H. L. Ploegh, *Nat Protoc* **2013**, 8, 1787; b) C. S. Theile, M. D. Witte, A. E. Blom, L. Kundrat, H. L. Ploegh, C. P. Guimaraes, *Nat Protoc* **2013**, 8, 1800.
[170] M. W. Popp, S. K. Dougan, T. Y. Chuang, E. Spooner, H. L. Ploegh, *Proc Natl Acad Sci U S A* **2011**, 108, 3169.
[171] L. Chen, J. Cohen, X. Song, A. Zhao, Z. Ye, C. J. Feulner, P. Doonan, W. Somers, L. Lin, P. R. Chen, *Sci Rep* **2016**, 6, 31899.
[172] J. Li, Y. Zhang, O. Soubias, D. Khago, F. A. Chao, Y. Li, K. Shaw, R. A. Byrd, *J Biol Chem* **2020**, 295, 2664.
[173] X. Hemu, A. El Sahili, S. Hu, K. Wong, Y. Chen, Y. H. Wong, X. Zhang, A. Serra, B. C. Goh, D. A. Darwis, M. W. Chen, S. K. Sze, C. F. Liu, J. Lescar, J. P. Tam, *Proc Natl Acad Sci U S A* **2019**, 116, 11737.

[174] X. Hemu, J. To, X. Zhang, J. P. Tam, *J Org Chem* **2020**, 85, 1504.
[175] S. Massa, N. Vikani, C. Betti, S. Ballet, S. Vanderhaegen, J. Steyaert, B. Descamps, C. Vanhove, A. Bunschoten, F. W. van Leeuwen, S. Hernot, V. Caveliers, T. Lahoutte, S. Muyldermans, C. Xavier, N. Devoogdt, *Contrast Media Mol Imaging* **2016**, 11, 328.
[176] J. Jiao, J. Zhang, F. Yang, W. Song, D. Han, W. Wen, W. Qin, *Eur J Pharm Biopharm* **2020**, 152, 123.
[177] P. Debie, N. Devoogdt, S. Hernot, *Antibodies (Basel)* **2019**, 8.
[178] Z. Wang, D. Zhang, X. Hemu, S. Hu, J. To, X. Zhang, J. Lescar, J. P. Tam, C. F. Liu, *Theranostics* **2021**, 11, 5863.
[179] Y. Deng, B. Zheng, Y. Liu, S. Shi, J. Nie, T. Wu, P. Zheng, *J Vis Exp* **2020**, DOI: 10.3791/60774.
[180] R. Yang, Y. H. Wong, G. K. T. Nguyen, J. P. Tam, J. Lescar, B. Wu, *J Am Chem Soc* **2017**, 139, 5351.
[181] M. D. Lee, W. Y. Tong, T. Nebl, L. A. Pearce, T. M. Pham, A. Golbaz-Hagh, S. Puttick, S. Rose, T. E. Adams, C. C. Williams, *Bioconjug Chem* **2019**, 30, 2539.
[182] a) M. Ezban, M. Hansen, M. Kjalke, *Haemophilia* **2020**, 26, 156; b) E. Santagostino, M. E. Mancuso, *Drug Des Devel Ther* **2018**, 12, 2933.
[183] H. Ostergaard, J. R. Bjelke, L. Hansen, L. C. Petersen, A. A. Pedersen, T. Elm, F. Moller, M. B. Hermit, P. K. Holm, T. N. Krogh, J. M. Petersen, M. Ezban, B. B. Sorensen, M. D. Andersen, H. Agerso, H. Ahmadian, K. W. Balling, M. L. Christiansen, K. Knobe, T. C. Nichols, S. E. Bjorn, M. Tranholm, *Blood* **2011**, 118, 2333.
[184] B. Ramakrishnan, P. K. Qasba, *J Biol Chem* **2002**, 277, 20833.
[185] Q. Zhou, J. E. Stefano, C. Manning, J. Kyazike, B. Chen, D. A. Gianolio, A. Park, M. Busch, J. Bird, X. Zheng, H. Simonds-Mannes, J. Kim, R. C. Gregory, R. J. Miller, W. H. Brondyk, P. K. Dhal, C. Q. Pan, *Bioconjug Chem* **2014**, 25, 510.
[186] a) B. M. Zeglis, C. B. Davis, R. Aggeler, H. C. Kang, A. Chen, B. J. Agnew, J. S. Lewis, *Bioconjug Chem* **2013**, 24, 1057; b) B. E. Cook, R. Membreno, B. M. Zeglis, *Bioconjug Chem* **2018**, 29, 2734; c) R. van Geel, M. A. Wijdeven, R. Heesbeen, J. M. Verkade, A. A. Wasiel, S. S. van Berkel, F. L. van Delft, *Bioconjug Chem* **2015**, 26, 2233.
[187] K. von Figura, B. Schmidt, T. Selmer, T. Dierks, *Bioessays* **1998**, 20, 505.
[188] M. J. Appel, C. R. Bertozzi, *ACS Chem Biol* **2015**, 10, 72.
[189] a) R. M. Barfield, D. Rabuka, *Methods Mol Biol* **2018**, 1728, 3; b) D. Rabuka, J. S. Rush, G. W. deHart, P. Wu, C. R. Bertozzi, *Nat Protoc* **2012**, 7, 1052; c) I. Rupniewski, D. Rabuka, *Methods Mol Biol* **2019**, 2012, 63.
[190] a) I. S. Carrico, B. L. Carlson, C. R. Bertozzi, *Nat Chem Biol* **2007**, 3, 321; b) P. M. Drake, A. E. Albers, J. Baker, S. Banas, R. M. Barfield, A. S. Bhat, G. W. de Hart, A. W. Garofalo, P. Holder, L. C. Jones, R. Kudirka, J. McFarland, W. Zmolek, D. Rabuka, *Bioconjug Chem* **2014**, 25, 1331; c) T. Kruger, S. Weiland, G. Falck, M. Gerlach, M. Boschanski, S. Alam, K. M. Muller, T. Dierks, N. Sewald, *Angew Chem Int Ed Engl* **2018**, 57, 7245.
[191] M. Knop, T. Q. Dang, G. Jeschke, F. P. Seebeck, *Chembiochem* **2017**, 18, 161.