Chemical Strategies for the Synthesis of Protein–Polymer Conjugates

Björn Jung and Patrick Theato

Abstract  Protein-polymer conjugates have achieved tremendous attention in the last few years. The synergistic combination of properties has led to certain advantages in bio-applications. Over the past few years, numerous chemical strategies have been developed to conjugate different synthetic polymers onto proteins, most of which can be summarized within the scope of click-chemistry. Here we highlight conjugation strategies based on available functional groups present on the synthetic polymer and existing groups of proteins from the natural pool. In particular, the chapter organizes the various possible reactions by classes of functional groups present on protein surfaces, deriving from selected amino acid residues.

Keywords  Biomaterials · Click chemistry · Peptides · Polymer conjugate · Protein modification · Proteins

Contents

1 Introduction .................................................................................................................. 38
2 Applications ................................................................................................................. 38
  2.1 Polymeric Therapeutics .......................................................................................... 39
  2.2 Smart Polymer Conjugates ...................................................................................... 40
  2.3 Giant Amphiphiles ................................................................................................. 41
3 Suitable Polymers ........................................................................................................ 41

B. Jung
Graduate School Materials Science in Mainz, Institute of Organic Chemistry, University of Mainz, Duesbergweg 10-14, 55099 Mainz, Germany

P. Theato (✉)
Institute for Technical and Macromolecular Chemistry, University of Hamburg, Bundesstrasse 45, 20146 Hamburg, Germany
e-mail: theato@chemie.uni-hamburg.de
1 Introduction

Protein–polymer conjugates are biohybrid materials comprising natural bioorganic polypeptides or proteins and synthetic polymer segments, which can combine or benefit from the advantages of both components and avoid the disadvantages of each separate component [1]. From the point of view of a synthetic polymer chemist, a protein is basically a polymer with a certain chain length, mostly monodisperse, and a definite hierarchical structure. In contrast, the synthetic polymer possesses a molecular weight distribution which depends on the polymerization technique used. Likewise, control of the regional structure is still very limited, even though recent developments on structural control show very promising results [2–4]. For example, this control affects the co-monomer sequence in copolymerization processes. Proteins, of course, feature a unique biorecognition and binding depending on their particular function, while the synthetic counterparts, i.e., synthetic polymers, mostly lack this ability. On one hand, synthetic polymers can be biocompatible, non-toxic, and non-immunogenic, however, their biodegradation is often a problem. On the other hand, their chemical and thermal stability is higher.

The conjugation of proteins to polymers changes the properties of each unit and opens up the avenue to several new applications which the single parts were not able to achieve. This chapter summarizes briefly the main fields of application and gives a short overview of suitable polymers that add a special value when conjugated to a protein. The conjugation chemistry towards protein–polymer conjugates forms the main part of this chapter which aims to highlight conjugation strategies based on available functional groups present on the synthetic polymer and on existing groups of proteins from the natural pool.

2 Applications

Polymer-protein conjugates are suitable for several applications. On one hand, the protein is equipped with new features enabling applications the unmodified protein would not be able to achieve, for example, those requiring a higher stability against enzymatic degradation. On the other hand, a synthetic polymer can be equipped with a unique three-dimensional structure to allow a higher hierarchical order or a biorecognizing structure by covalent conjugation with a protein. Most applications
of polymer-protein conjugates are found in the area of polymeric therapeutics [5, 6], as highlighted below.

### 2.1 Polymeric Therapeutics

A disease may occur through a mutation or abnormality during the transcription, translation, or post-translational process. A change in the concentration of a protein can also lead, for example, to a malfunctioning metabolism, a signal transduction pathway, or an immune response [7]. Considered from this perspective, proteins provide an excellent opportunity to alleviate or treat diseases. The nature of a protein itself makes it suitable for this application. Due to the complexity of interactions from a protein to its target structure, proteins normally cannot be replaced by small and easily accessible molecules. Using natural proteins, for example, for replacing damaged or absent proteins, are less likely to initiate an immune response. Likewise, through their biological function they are interesting components for regulators or inhibitors of biological processes. Compared to these advantages, several problems generally avoid the direct usage of pure proteins as specific therapeutic agents. The critical points are a low stability in vivo, a short half-life time, and, thus, a rapid clearance from the body. This process is an interaction of several events like proteolysis by enzymes, clearance mechanisms, or protein modifying enzymes, or accumulation or shielding by other proteins [8]. Another challenge is that the immune system may activate a response to a foreign imported protein and negotiate a useful treatment or cause a harmful reaction [9]. This problem is more relevant for proteins of non-human origin or for recombinant proteins [10]. However, a therapeutic effect often only occurs by maintaining a certain concentration of the agent over a period of time [11]. Thus, strategies to prolong the half-life time of therapeutic protein have to be achieved. These strategies include techniques to avoid or to reduce renal clearance, to increase receptor mediated recycling, or to decrease the stability of the interaction of protein-receptor binding during endocytosis [12]. Non-covalent approaches focus on altering the amino-acid sequence or encapsulating proteins into vehicles. Receptor mediated recycling requires linkage to certain other proteins, whereas other concepts use covalent modification of the protein by low molecular weight compounds or polymers. This chapter will focus particularly on the last one. Protein–polymer conjugates implicate the following changes compared to an unmodified protein. The hydrodynamic volume is increased and thus the renal excretion rate decreased if the size of the conjugate is bigger than the glomerular filtration barrier [13]. Additionally, the polymer chains can shield the protein from enzymatic degradation, receptor recognition, or antibodies, which again increases the serum half-life time. Polyethylene glycol is the most used polymer in drug discovery to overcome the above-mentioned problems [14, 15]. Furthermore, the attachment of polyethylene glycol chains onto the protein surface prevents or decreases the immunogenicity and aggregation. Accordingly, PEGylation of
proteins, i.e., the process of attaching a polyethylene glycol chain to a protein, has become a well-established technology for the use of proteins as drugs, especially as anti-cancer agents [16, 17]. Further, another factor of polymeric therapeutics has become very beneficial and motivates the interest in this research field. Macromolecules accumulate in tumor tissues relative to healthy tissue through enhanced vascular permeability. This concept is called the enhanced permeability and retention effect (EPR effect) [18–20]. This effect was initially found with a protein–polymer conjugate called SMANCS of styrene-maleic acid copolymer and neocarzinostatin that features anti-tumoral activity [21]. In addition to the clearance advantages, the EPR-effect pushes the effort to conjugate therapeutic useful proteins to biocompatible polymers.

Examples for successful protein–polymer conjugates that are in clinical use are the above-mentioned SMANCS for hepatocellular carcinoma [22], Oncaspar, a PEG-enzyme conjugate for lymphoblastic leukemia [23], and PEG cytokines like Pegasys [24] and PEG-Intron [25] as antiviral agents against hepatitis [15, 26–28].

In spite of all the benefits from both polymeric parts, a wrong linkage can reduce or inactivate the protein bioactivity, especially by shielding binding pockets or catalytic centers or using necessary amino acid residues for the linkage. The advantage having a monodisperse protein can also be negated by random connection of polymer chains. Consequently, the right linkage strategy is still a challenge and depends strongly on suitable polymers, which is discussed in the following section.

### 2.2 Smart Polymer Conjugates

The conjugation of proteins to stimuli responsive polymers are used in the field of protein isolation and separation [29, 30]. The ability of some polymers to undergo a reversible change in response to an external physical, chemical, or biochemical stimulus gave them the name smart polymers [31]. Different stimuli have been utilized, such as temperature, light, ionic strength, or electric field for physical factors, or pH and specific ions for chemical triggers or metabolites for biochemical reactions [32, 33].

Responsive polymer-protein conjugates can be used to influence the accessibility of the active site of an enzyme or of the recognition site of a receptor [34–36]. The strategy for controlling the protein activity is based on a reversibly mechanism to block the active site. Upon stimulation, the attached polymer will collapse and hide the active site. An example for a temperature controlled mechanism is poly($N,N$-diethylacrylamide) (PDEAM), which is attached next to the binding site [37]. Below the critical solution temperature of the conjugated polymer, the polymer chain exhibits an extended state and inhibits a binding of target molecules. Above the critical temperature, the polymer is in its shrunken state and unblocks the binding site. An example for light responsive conjugates is a copolymer consisting of $N,N$-dimethyl acrylamide and an azobenzene containing acrylate as monomers.
Irradiation with ultraviolet light switched the enzyme activity off. Under visual light the activity is regained.

Another approach consists in using smart polymers for a triggered protein precipitation, which can be used for purification of proteins. For example, the enzyme trypsin was coupled with poly(N-isopropylacrylamide) (PNIPAM). After successful conjugation, the protein could be precipitated by heating above the lower critical solution temperature (LCST) of the polymer [39, 40].

### 2.3 Giant Amphiphiles

Classic amphiphiles or surfactants consist of a hydrophobic and a hydrophilic group. Due to their amphiphilic nature, they self-assemble in aqueous solution to form ordered aggregates depending on their shape [41]. By increasing the size of the amphiphiles from small molecules to macromolecules one obtains block copolymers as supersurfactants. Hence, the combination of proteins as head groups and polymers as tails also leads to the formation of giant amphiphiles. The aggregation behavior is similar compared with their low molecular weight equivalents. However, the giant aggregates usually possess a higher stability, slower exchange dynamics and a lower critical micellar concentration [29, 41, 42]. As an example, an amphiphile consisting of poly(styrene) as the apolar tail and the enzyme horseradish peroxidase as the polar head group is chosen [43]. The hydrophobic tail was connected with the ligand of the apoprotein. The hybrid was prepared by adding the polymer dissolved in an organic solvent to an aqueous solution containing the apoenzyme. These systems have been further improved by using responsive polymers to form giant amphiphiles that lead to stimuli responsive applications. They are considered as interesting candidates for triggered drug release [44].

### 3 Suitable Polymers

Applicable polymers for therapeutic applications should be water-soluble, non-toxic, and non-immunogenic [45]. They should neither accumulate during a therapy nor remain in the body. Thus, an elimination strategy or possibility for degradation should be available [46]. The body residence time of the conjugate has to be chosen accordingly so that the polymer prolongs the life time and thus allows the distribution through the body to accumulate in the desired tissue [13]. The polymer should be obtainable with a low polydispersity to avoid a broad product mixture. Ideally the polymer should feature only one reactive group to obtain distinct conjugates without cross-linking. Normally a polymer cannot satisfy all requirements and
compromises have to be made. This clearly motivates further intensive research in this area.

The most commonly used polymer is poly(ethylene glycol) (PEG) because it has been approved for human use by FDA and thus finds application as a pharmaceutical excipient resulting from its non-toxic properties. Additionally, its high water solubility and flexibility allow the protein to create a large hydrodynamic radius to enhance the EPR effect and half life [47]. Further, PEG offers only two possibilities for the conjugation: the two chain ends. In the case of the diol, the polymer can react with two groups or, in the case of the methoxy form, only the single hydroxyl group can react. Logically, a broad scope of end group functionalization for PEG has meanwhile been established.

Other suitable polymers are poly(N-(2-hydroxypropyl)methacrylamide) (PHPMA) and HPMA copolymers, poly(vinylpyrrolidone) (PVP), poly(ethyleneimine) (PEI), poly(acroloylmorpholine) (PAcM), divinylethermaleic anhydride/acid copolymer (DIVEMA), poly(styrene-co-maleic acid/anhydride) (SMA), and poly(vinyl alcohol) (PVA) [13]. These polymers are all of synthetic origin and are based on the polymerization of vinyl monomers, either resulting in a homopolymer or, when using different monomers, resulting in copolymers. They are usually synthesized by free radical polymerization. As the polymers obtained are not biodegradable, their molecular weight has to be lower than the clearance cut-off. Any coupling chemistry can occur via the side chain functionality or by introducing addressable reactive end groups. In contrast, PEG and PEI are synthesized by ring opening polymerization from ethylene oxide or the corresponding amine azeridine, respectively. Alternatively, PEI can also be prepared by polymerization of 2-substituted oxazolines and subsequent polymer hydrolysis.

Poly(glutamic acid), poly(l-lysine), poly(aspartamides), and poly((N-hydroxyethyl)- l-glutamine) (PHEG) are examples of poly(amino acids) [13]. An analog polymer is poly(malic acid) that exhibits an ester linkage instead of a polyamide backbone. This is biodegradable in analogy to natural polypeptides. Other natural polymers are normally polysaccharides like dextran or chitosan.

Stimuli-responsive polymers should respond to an external trigger, as mentioned above. Temperature-responsive polymers exhibit a volume phase transition at a certain temperature and undergo a change in solvation state [34]. Certain polymers feature a LCST and become insoluble upon heating, while other polymers may exhibit an upper critical solution temperature (UCST) and become soluble upon heating. Typical temperature-sensitive polymers featuring a LCST are PNIPAM, PDEAM, poly(methylvinylether) (PMVE), and poly(N-vinylcaprolactam) (PNVCl) [48–50]. Positive temperature-sensitive polymer systems with a UCST are poly(acrylamide-co-acrylic acid) and PEG-b-poly(propylene glycol). For completeness, there are also systems with both transition states, so-called schizophrenic polymers [51–53]. pH-sensitive polymers are based on chargeable groups. Anionic polymers often consist of poly(acrylic acid) and accordingly poly(methacrylic acid), while positive charged polymers contain amino groups like poly(ethylene imine) and poly (l-lysine). There are also a few polymeric systems that react upon electric or magnetic stimulus [30]. Apart from the poly(glycol) system, the polymers are of
vinyl origin. Thus, in order for the conjugation to proteins to occur preferentially at the ends of the polymer, certain synthetic criteria have to be met and are discussed in the following sections.

4 Conjugation Strategies

4.1 Natural Pool on Functionalities

If one looks at possible conjugation chemistries, one can choose to tackle this topic from the attachment point of view of synthetic polymers or from the side of the proteins. The chemical groups on proteins available for conjugation reactions are limited and mostly dictated, unless protein engineering methods for the synthesis of artificial protein are employed [54, 55]. From the point of view of polymers, nearly every possible chemical group can be obtained. For this reason, we have structured the main part of this section by means of functional groups available on the single amino acids.

It is essential to have sufficient knowledge of the protein structure to conjugate a synthetic polymer selectively with a certain feature onto the protein surface. Thus, not only the primary sequence matters, but also the proteins three-dimensional structure, i.e., its tertiary structure or, if it exhibits more subunits, its quaternary structure, is important. The desired target amino acid for the conjugation should be freely addressable on the surface and not be hidden in the center. Further, the functionalization must not affect the three-dimensional structure of the protein to maintain its native state and binding pockets or recognition areas. Such changes occur, for example, by alteration of the protein’s overall charge or electrostatic and hydrophobic interactions. An exception is the directed deactivation of a function of a protein. Another major challenge is the introduction of polymer chains in a defined quantity. A random attachment of polymer chains usually leads to an undefined product and a main advantage, the monodispersity of the protein, is lost. Especially large proteins present multiple copies of the target amino acid, which can often be as high as 20 amino acids. Thus, the strategy depends on the selected protein and the available information about it. If the information mentioned is not available, the knowledge about the natural amino acid abundance, their average distribution within the three-dimensional structure, and their average appearance on the surface can help to obtain a well-defined protein–polymer conjugate [56]. Furthermore, this illustrates, which amino acids are of particular interest. Noteworthy, the nonpolar amino acids are unimportant for conjugation chemistry. For these amino acids, almost no specific and selective reaction is available to obtain a covalent protein–polymer conjugate provided that the protein maintains its native state. Apart from the five aliphatic nonpolar amino acids,
glycin, which exhibits with its hydrogen residue no functionality at all, the thioether
containing methionine, and the aromatic phenylalanine can be excluded. Only
phenylalanine has raised minor interest in aromatic chemistry, especially cross-
coupling reactions [57]. However a selective coupling should be difficult because of
the presence of tryptophan and histidine and is thus only practicable in
oligopeptides. Serine and threonine have also less importance because of their
weaker nucleophilicity compared with the amino group of lysine. Only at special
positions do they offer conjugation chemistry [58]. The same is valid for the two
acid amino acids. From the last ten remaining amino acids, asparagine also plays no
important role in conjugation chemistry. Thus, it is mainly eight amino acids plus
the N- and C-terminus that have to master the conjugation process.

For selective coupling, a less abundant amino acid gives a good first impression
for controlled modification for well-defined protein–polymer conjugates [56]. Cys-
teine and tryptophan are less used in the polypeptide chain. Aiming for a free
cysteine is a first approach for a monodisperse product.

Next, the knowledge about different propensities of the location of different
amino acids in certain regions of the protein can help to achieve a selective strategy.
Examples are the accumulation of certain amino acids in characteristic regions like
the N- or C-terminus, catalytic sides, or binding areas [59, 60]. Every amino acid
has a preferred location within the protein. Table 1 shows the percentage amino

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Location</th>
<th>Functionality</th>
<th>Natural abundance</th>
<th>ASA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cysteine</td>
<td>C</td>
<td>Thiol</td>
<td>1.36</td>
<td>0.268</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>C</td>
<td>Aliphatic</td>
<td>5.97</td>
<td>0.273</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>C</td>
<td>Indole</td>
<td>1.08</td>
<td>0.279</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>C</td>
<td>Benzyl</td>
<td>3.86</td>
<td>0.290</td>
</tr>
<tr>
<td>Valine</td>
<td>C</td>
<td>Aliphatic</td>
<td>6.87</td>
<td>0.306</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>C</td>
<td>Phenol</td>
<td>2.92</td>
<td>0.319</td>
</tr>
<tr>
<td>Leucine</td>
<td>C</td>
<td>Aliphatic</td>
<td>9.66</td>
<td>0.321</td>
</tr>
<tr>
<td>Methionine</td>
<td>C</td>
<td>Thioether</td>
<td>2.42</td>
<td>0.364</td>
</tr>
<tr>
<td>Alanine</td>
<td>C</td>
<td>Aliphatic</td>
<td>8.26</td>
<td>0.405</td>
</tr>
<tr>
<td>Histidine</td>
<td>M</td>
<td>Imidazole</td>
<td>2.27</td>
<td>0.425</td>
</tr>
<tr>
<td>Threonine</td>
<td>M</td>
<td>Hydroxy</td>
<td>5.34</td>
<td>0.480</td>
</tr>
<tr>
<td>Proline</td>
<td>M</td>
<td>Aliphatic</td>
<td>4.69</td>
<td>0.502</td>
</tr>
<tr>
<td>Arginine</td>
<td>M</td>
<td>Guanidine</td>
<td>5.53</td>
<td>0.539</td>
</tr>
<tr>
<td>Asparagine</td>
<td>M</td>
<td>Carboxamide</td>
<td>4.06</td>
<td>0.568</td>
</tr>
<tr>
<td>Serine</td>
<td>S</td>
<td>Hydroxy</td>
<td>6.55</td>
<td>0.568</td>
</tr>
<tr>
<td>Glutamine</td>
<td>S</td>
<td>Carboxamide</td>
<td>3.93</td>
<td>0.573</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>S</td>
<td>Carboxylic acid</td>
<td>6.75</td>
<td>0.586</td>
</tr>
<tr>
<td>Glycine</td>
<td>S</td>
<td>–</td>
<td>7.08</td>
<td>0.588</td>
</tr>
<tr>
<td>Lysine</td>
<td>S</td>
<td>Primary amine</td>
<td>5.85</td>
<td>0.607</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>S</td>
<td>Carboxylic acid</td>
<td>5.46</td>
<td>0.615</td>
</tr>
</tbody>
</table>

*aLocation of the amino acid to their average composition in core (C), intermediate (M) and
surface (S)

*Average surface accessibility
The average surface accessibility indicates whether an amino acid is present more in the core or on the surface. Again cysteine and tryptophan are rare on the surface [62]. Nevertheless, the selective functionalization of tryptophan has been explored [63]. In contrast, lysine with an accessible amine group, which opens a broad range for organic reaction, is a common amino acid on the surface.

4.2 Methods for Protein Functionalizations

In order to synthesize protein–polymer conjugates, three main routes are available [29, 64–66] (see Fig. 1). First, the protein can be directly modified with a preformed polymer. This grafting-to approach is mediated either through covalent attachment of a reactive functional group of the polymer to a corresponding amino acid side-chain, or vice versa, or by a ligand-apoprotein interaction. In the latter case, a cofactor or ligand is covalent linked to a polymer chain. Usually that polymer exists either as an \( \alpha,\omega \)-telechelic polymer, with a reactive group, allowing conjugation with the polypeptide, on one end and the polymerization initiating group on the other [67]. Alternatively, the reactive group can be introduced by post-polymerization modification of the polymer end group [68]. An indirect protein–polymer conjugation is the grafting-from approach. In this case, a moiety that is able to mediate or initiate a polymerization process is introduced to an amino acid side chain. Consequently, a macro-initiator is formed and the polymer chain can be grown directly from the protein.
The third route follows the grafting-through approach in which various protein reactive groups are incorporated within a growing polymer chain by using monomers that can react either directly or after polymerization via the introduced reactive moieties with peptides or proteins [69]. Noteworthy, this third approach does not necessarily result only in conjugation of one protein/peptide to a polymer chain, but also in conjugation of several protein/peptide to a polymer chain.

PEG conjugates belong to the first mentioned strategy. PEG is characterized by the lack of side chains and the missing possibility to propagate the PEG chain from a macro protein initiator through an anionic polymerization mechanism. In contrast, polymers prepared by radical polymerization lead to a broad range of possible chemical modifications [70]. Established controlled radical polymerization processes are reversible addition-fragmentation chain transfer (RAFT) [71–74], atom transfer radical polymerization (ATRP) [75–77], and nitroxide mediated polymerization (NMP) [78], which all lead to polymers with a low polydispersity and predetermined molecular weight and – most importantly – allow the selective conjugation via end groups to yield well-defined protein–polymer conjugates. In addition, RAFT and ATRP open up the possibility to create grafting-from approaches either by conjugation of chain transfer agents or initiator systems to proteins/peptides, respectively. RAFT-polymerization is enabled by a chain transfer agent that is attached to an amino acid side chain functionality. ATRP takes advantage of attachment of a halogenated moiety from which the polymer chain can grow. After the polymerization process a functional end group at the omega terminus remains, which opens up further reaction possibilities. Further advantages of RAFT and ATRP are the feasibility of the polymerization in a wide range of solvents, reaction conditions, and with various suitable monomers, presuming that the conditions are compatible with the protein/peptide. It should be mentioned that other polymerization processes like cationic and anionic polymerization or polycondensation have not been successfully performed in the presence of proteins.

4.3 Protein–Polymer Conjugates Listed by Amino Acids

The following summarizes chemical ligation strategies to combine selectively proteins with polymer chains based on the ten natural amino acids that are qualified for ligation chemistry and novel non-natural moieties. For every mentioned reaction type, examples from the last few years are given.

4.3.1 Lysine and the N-Terminus of Proteins

The primary amino group of the lysine side chain and the N-terminus are an attractive target for conjugation chemistry. First, lysine is a common amino acid on the surface and, thus, in mostly every protein a primary amino group should be available. Second, the nucleophilicity of the amine is higher than other nucleophilic groups of other amino acids, in particular the sulfhydryl group of cysteine, hydroxyl
group of serine and threonine, and the imidazole moiety of histidine. Hence, a reaction with an electrophile leads preferably to a linkage with the amino group. A limitation is a possible product distribution consisting of isomers and variable amount of polymer chains per protein if multiple lysine residues are present [79]. Control over the preferred conjugation site, if the conjugation addresses the N-terminus or the ε-amino group of lysine, can be achieved by adjusting pH [80]. The pKₐ value of the N-terminal amine group is about 7.6–8.0 and the pKₐ of the lysine side chain about 10.0–10.2 [81]. By lowering the pH value from the traditional range for lysine conjugation of about 8.5–9.0, the reaction can be directed to the N-terminus [82]. Noteworthy, amino groups near or at a catalytic center or binding pocket can be blocked by adding a ligand or substrate during the reaction, i.e., competitive inhibition.

Two different conjugate products are possible. In one case the charge of the amino group gets lost and in the other the charge and thus the overall charge of the protein is maintained. Reactions of the first type are acylation like formation of amides and carbamides, or analogous reactions with corresponding thio derivatives. Reactions with aldehydes and ketones with following reduction, i.e., reductive amination, as well as amidination are examples of the second case.

A conventional strategy for the formation of amides from activated carboxylic acids is based on classical organic chemistry. A common method is the use of active esters, such as N-hydroxysuccinimide esters, the –OSu group (NHS esters 1) which is usually prepared from the desired acid, and N-hydroxysuccinimide using a coupling agent like dicyclohexylcarbodiimide (DCC) or ethyl(dimethylamino)propyl carbodiimide (EDC). Protein polymer conjugation can then easily be achieved by reaction of the active ester with an amine under ambient conditions. In addition, several active esters are meanwhile commercially available. The active ester chemistry was also used to attach an initiator for ATRP, such as 2-bromo-isobutyric acid, to amino residues of proteins, thereby enabling the grafting of stimuli responsive polymers from the surface of a protein [83]. Alternatively, the opposite way, the grafting-to approach utilizes an active ester bearing ATRP initiator to polymerize the desired monomer and then conjugate the obtained polymer to the protein [84, 85]. In comparison to these examples, the NHS group can also be introduced after the polymerization by end group modification of a free acid and N-hydrosuccinimide with DCC [86]. In a similar way, PEG chains can also be functionalized via the NHS route. The hydroxyl end group of PEG can be converted into an acid functionality with succinic anhydride, which is then activated with N-hydroxysuccinimide. To overcome the lability of the ester bond, the NHS group can be introduced by spacers containing amide or ether bonds [79, 87]. Ether linked derivatives are formed by the combination of methoxy-PEG (CH₃O-PEG-OH) and an omega functionalized acid like propionic or butanoic acid [88]. Amide linked acid groups are formed by using β-alanine or norleucine. Multiple PEG chains per NHS anchor groups have been realized with spacers having more than one reactive group like lysine, which creates an unsymmetrical branch, or 1,3-diamino-2-hydroxypropane, which creates a symmetric branch [89] (Table 2).
**Table 2**  Coupling methods for the amino group

<table>
<thead>
<tr>
<th>Reaction type</th>
<th>Reaction</th>
<th>Product</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acylation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>NHS ester</td>
<td>Amide</td>
<td>[83–89]</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="NHS ester reaction" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>NHS carbonate</td>
<td>Carbamate</td>
<td>[90]</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="NHS carbonate reaction" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NHS carbamate</td>
<td>Carbamide</td>
<td>[91]</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="NHS carbamate reaction" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Thiazolidine-2-thione</td>
<td>Amide</td>
<td>[92]</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Thiazolidine-2-thione reaction" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Pentafluorophenyl ester</td>
<td>Amide</td>
<td>[93–95]</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Pentafluorophenyl ester reaction" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>Chemical Strategy</td>
<td>Reaction Equation</td>
<td>Product</td>
</tr>
<tr>
<td>--------</td>
<td>------------------</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td>6</td>
<td>Anhydride</td>
<td>$\text{R}R'\text{O}\text{O}\text{R} \rightarrow \text{HOOC}N\text{H}\text{R}\text{N}\text{O}\text{O}\text{Protein}$</td>
<td>Amide + free acid</td>
</tr>
<tr>
<td>7</td>
<td>Acid halogenide</td>
<td>$\text{R}\text{Halo} \rightarrow \text{RNHProtein}$</td>
<td>Amide</td>
</tr>
<tr>
<td></td>
<td><strong>Alkylation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Reductive amination</td>
<td>$\text{R}\text{H} \rightarrow \text{Red} \rightarrow \text{RNHProtein}$</td>
<td>Secondary amine</td>
</tr>
<tr>
<td>9</td>
<td>Amidination</td>
<td>$\text{R}\text{NH}\text{O}\text{N} \rightarrow \text{RHNProtein}$</td>
<td>Amidine</td>
</tr>
</tbody>
</table>
If an acid functionality is not available, hydroxyl groups can be converted by N-hydroxysuccinimide, too. Using an oligo ethylene glycol with two accessible hydroxyl end groups, one hydroxyl group can be used to attach an ATRP initiator via an acid bromide to form an ester group. The other hydroxyl group can be converted with \(N,N'\)-disuccinimidyl carbonate to a succinimidyl carbonate compound, which can be reacted with lysine residues to yield carbamate conjugates [2] [90]. NHS chemistry can also lead to carbamide linkages (3). A system less reactive towards amines results in a conjugation in which only the most nucleophilic amino residue reacts and thus a lower amount of possible isomers are obtained. Such a system is the \(\beta\)-alanine-NHCO-OSu group. The carboxylic acid group of \(\beta\)-alanine is used for conjugation with the polymer and the amino group is modified with \(N,N'\)-disuccinimidyl carbonate to yield the corresponding carbamate that can react with the most nucleophilic amino group of a protein [91]. Another active ester is based on the thiazolidine-2-thione group (4). The carboxylic acid group of a chain transfer agent can react with 2-mercaptothiazoline to form a thiazolidine-2-thione ester group. After RAFT polymerization the polymer can react with amino groups in a grafting-to approach [92]. Other reactive units are amongst others pentafluorophenyl active esters (5), benzotriazole carbonates, chlorotriazines, and \(p\)-nitrophenyl carbonates [79, 87, 93–95].

Another reactive form of a carboxylic acid is the anhydride (6). A famous example is the SMANCS conjugate. The antitumor protein neocarzinostatin is linked with its two lysine residues to a poly(styrene-\(\text{co}\)-maleic acid/anhydride) copolymer. The linkage occurs in the side chain by opening the five-membered ring of maleic anhydride. Per linkage one free acid is generated [21]. The copolymer poly(maleic anhydride-\(\text{alt}\)-methyl vinyl ether) can also be conjugated with a protein shell of a virus to immobilize the desired compound. The remaining anhydride repetition units were then used to entrap the conjugate in an amino groups containing film [96]. The conjugation proceeds in both cases in a grafting-through approach within the side chain. The reactive group is introduced through the monomer in the polymerization procedure. An end group modification is not necessary. Using synthetic polypeptides, conjugates with only one linkage per chain are generated. A fully protected polypeptide is coupled with its N-terminal amino group to methacrylic acid anhydride [97, 98]. The resulting monomer can then be copolymerized with biocompatible monomers such as HPMA. Acid halogenides are a similar reactive system (7). 2-Bromoisoobutyryl bromide reacts in slightly basic buffer solution with lysine residues to yield an ATRP macro protein initiator. The average number of acylated residues depends on the molar ratio of the acid bromine. The protein initiator can be used for a grafting-from polymerization [99, 100].

Free acid can also be used for conjugation chemistry onto a protein’s amine group. Poly(acrylic acid) has been coupled under nearly neutral conditions to lysine side chains of hemoglobin using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide. A water swollen gel was formed that is slightly cross-linked because of the multiple lysine residues that can react with different polymer chains [101].
Another strategy is to maintain the positive charge of the nitrogen of the amine. The route to form secondary amines via reductive amination from aldehydes and the amine group is especially popular (8). In the first step an imine is formed, which has to be reduced to a secondary amine, because the imine formation is a reversible reaction. This method was devised with PEG-aldehyde and granulocyte colony stimulating factor (G-CSF) [79, 82]. The reaction occurs under slightly acid conditions. The intermediate, the imine, will be directly reduced with sodium cyanoborohydride to result in the secondary amine. The equilibrium is shifted to the product side through the in situ reduction. The advantage of this approach is to obtain a one to one conjugate because of the higher nucleophilicity of the N-terminus and – under acid conditions – the reaction occurs selectively at this position. This conjugate leads to pegfilgrastim, a leukocyte stimulating drug. This conjugation approach can also be adapted to polymers that were prepared by a controlled radical polymerization. Starting from 2-bromoisobutyryl bromide, which is used for ATRP, the aldehyde group is introduced to the initiator system in form of an acetal protected group as 2-(2,2-dimethoxy-ethoxy)-ethanol [102]. After polymerization of the macromonomer methoxyPEG\(_{1100}\)-methacrylate, the terminal aldehyde group is recovered by deprotection with trifluoroacetic acid. Again, the N-terminus is addressed under acidic conditions in a one pot reaction with sodium cyanoborohydride as reductant [103]. Recent developments use transition metal mediated catalysis instead of sodium cyanoborohydride. A water-stable iridium complex uses formate as the hydride source [104]. The advantage is a milder reduction of imine groups generated though the coupling process in particular in the presence of disulfide bridges by a longer reaction time [105].

Another route is the amidination of lysine by means of imidoesters and imidothiolanes (9). The resulting amidine still carries a positive charge and the overall charge of the protein is recovered [106].

### 4.3.2 Cysteine

Next to lysine, cysteine is a frequent target in conjugation chemistry. The lower abundance – especially on the surface of a protein – and unique reaction pathways allow a selective addressing and thus result in a lower product distribution (see Sect. 4.1). Cysteine residues may be blocked as internal disulfide bridges. Utilizing dithiothreitol exposes additional free cysteine residues by cleaving the disulfide bond. This reaction may, however, sometimes lead to a loss of the native three-dimensional structure [56]. If the protein lacks a free cysteine residue at the surface, genetic engineering can introduce cysteine moieties for conjugation chemistry [107, 108]. Two main approaches can be applied in protein–polymer conjugates. On the one hand, reactions that create a disulfide bond can be used, adding the opportunity for a cleavable linkage – on the other, many conjugations are achieved by Michael addition that forms an alkylated cysteine.

A common way to introduce polymer chains to cysteine moieties is the use of orthopyridyl disulfides (PDS, 10). The ATRP initiator 2-bromoisobutyrate can be
built up with the mentioned cysteine reactive residue orthopyridyl disulfides starting from 2,2-dithiopyridine, mercaptopropanol, and 2-bromo-2-methylpropionic acid in two steps [109, 110]. The dithio moiety can react with a free cysteine residue under disulfide formation and in a following grafting-from approach the bromoisobutyrate group initiates the polymerization of, e.g., 2-hydroxyethyl methacrylate [110] or other monomers yielding polymers [111]. Genetically engineered lysozyme with a novel thiol group makes the orthopyridyl disulfide group accessible to the protein [111]. In an analogous way, chain transfer agents can also be attached to this moiety [112]. As an example a water soluble trithiocarbonate was chosen. After coupling to a free cysteine residue and releasing the 2-pyridinethione leaving group, the macro chain transfer agent is able to act in RAFT polymerization [113]. A wide variety of acrylate monomers can be used in this grafting-from approach [114]. Using symmetrical trithiocarbonate chain transfer agents, the possibility to create heterotelechelic protein–polymer conjugates has been explored. Instead of using a non-functional benzyl group for one side of the trithiocarbonate, a second orthopyridyl disulfide group was chosen [115]. First, one PDS group was attached to the protein, while the second one was inaccessible for protein linkage. In a grafting-from approach the protein–polymer conjugate was built with an intact PDS group. The terminal PDS group was then available for further post modifications such as the attachment of dyes. In addition, it is also possible to create chain transfer agents with the PDS group in the middle of a bifunctional chain transfer agent via a side chain. On both sides the orthopyridyl disulfide unit carries

<table>
<thead>
<tr>
<th>Reaction type</th>
<th>Reaction</th>
<th>Product</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Disulfide linkage</td>
<td><img src="image" alt="Disulfide" /></td>
<td>Disulfide [109–115]</td>
</tr>
<tr>
<td>11</td>
<td>Michael addition of disulfide bonds</td>
<td><img src="image" alt="Michael addition" /></td>
<td>Bridged disulfide [119, 120]</td>
</tr>
<tr>
<td>12</td>
<td>Michael addition with maleimides</td>
<td><img src="image" alt="Michael addition" /></td>
<td>Thioether [121–134]</td>
</tr>
<tr>
<td>13</td>
<td>Michael addition with vinyl sulfones</td>
<td><img src="image" alt="Michael addition" /></td>
<td>Thioether [135]</td>
</tr>
</tbody>
</table>
one transfer agent [92]. After polymerization the midchain-functionalized RAFT polymer was conjugated to the protein in a grafting-to approach (Table 3).

Disulfide bonds in proteins have also been utilized to conjugate polymers onto them. For example, bromomaleimides were used to modify selectively and reversibly cysteine residues [116]. Interestingly, dibromomaleimides can be used to re-bridge successfully disulfide bonds following a cleavage with appropriate reducing agents [117, 118].

An alternative interesting approach is the use of an internal disulfide bonds for conjugation chemistry without the loss of the bridge [119, 120]. In the first step the disulfide bridge is cleaved under mild reduction conditions like tris[2-carboxyethyl] phosphine or dithiothreitol. Two nearby thiols are generated. One of these attacks the bis-thiol-specific reagent in a Michael addition. This specific reagent is an α,β-unsaturated β'-monosulfone. A sulfonyl leaving group is released and the vinylogous structure is regained. A second Michael addition completes the new three carbon bridged disulfide. In summary, the alkylating agent starts a sequential addition–elimination reaction cascade (see Fig. 2).

A major route for cysteine modification is the Michael reaction with maleimides (12) and vinyl sulfones (13). Myriads of examples of different applications and approaches have been presented and only a selection will be discussed in the following to demonstrate the synthetic principle. To graft a polymer from a maleimide moiety to be used within a controlled radical polymerization, trithiocarbonates are well suited [121, 122]. The chain transfer agent consists of a maleimide group that is connected by a glycol linker to the trithiocarbonate unit as the R-group, which can be used in a RAFT polymerization after conjugation to the protein. The grafted polymer still contains the chain transfer moiety at the end group and this macro-CTA can be used for the polymerization of a second monomer resulting in block copolymers.

Grafting-to approaches can be realized in two different ways. Either the maleimide moiety is connected to the initiator system or chain transfer agent or is introduced after the polymerization by end group modification. If the maleimide is introduced before the polymerization, a protecting group for the double bond is necessary to prevent reaction during polymerization [123]. Furan is a suitable protecting group and reacts with maleic anhydride to yield the protected adduct by Diels–Alder reaction [124]. The imide structure is formed by reaction of the anhydride with an amine, generally an α-amino-ω-hydroxy compound. For
example, reaction with 2-bromo-isobutyrylbromide results in the ATRP initiator system. After polymerization furan is removed by a retro Diels–Alder reaction by heating. The free Michael system can then be attacked by a target cysteine residue to obtain the desired protein–polymer conjugate [125]. Complex polymer designs like multiple block copolymers or several post-polymerization steps are accessible following this strategy [126]. RAFT agents are also available through this process. Instead of using an ATRP bromine containing compound, installation of a trithiocarbonate on the protected maleimide is sufficient. For example, a DCC mediated esterification can be used to link the two functionalities [127]. The maleimide moiety can also be introduced after polymerization; however, another functionality at the end group of the polymer is required for this approach. Using 2-aminoethanethiol hydrochloride as a simple chain-transfer agent leads to an amino functionalized polymer. A maleimide with a succinimide moiety results in the polymer that can be reacted with a cysteine unit [128]. If a RAFT polymerization was chosen, the trithiocarbonate end group can be replaced with an azo compound in a radical mechanism, following a method developed by Perrier [129]. Derivatives of 4,4′-azobis(4-cyanovaleric acid) are well-suited because they result in an acid functionality that can be used for coupling with a maleimide residue. In this case the double bond of the maleimide has to be protected with furan again. Another grafting-to approach takes advantage of mPEG-maleimide, which can be coupled directly to a cysteine moiety [130]. Depending on the starting material, homofunctionalized, homodimeric, or star polymers can be synthesized. Using a symmetrical RAFT agent with a trithiocarbonate moiety on each side, telechelic polymers are accessible [131]. After end group replacement with the azo compound the resulting polymers have the maleimide group on both sides, to which two identical proteins can be attached. When an α-ω functionalized RAFT agent is used instead, two different groups are available for a post-polymerization conjugation to two different proteins [132]. Using a tetrameric CTA with four identical trithiocarbonate moieties, star polymers with at least four maleimide groups are possible. In this case four identical proteins are linked together [133]. Instead of using azoinitiator based chemistry, the trithiocarbonate structure can be cleaved by using a mixture of hexylamine and tributylphospine. A thiol terminated polymer is obtained that reacts with a bis-maleimide in excess to yield a maleimide terminated polymer [134].

A polymer made by the RAFT process with a dithioester results in a dithioester end group. This ester can also be cleaved by aminolysis, resulting in a thiol terminated polymer. In an analogous way to the bis-maleimide compound, divinyl sulfone can be used, thereby creating a vinyl sulfone end group. This group is then able to react with cysteine groups by Michael addition (13) [135].

If no free cysteine is available, a one-pot approach enabling the breaking of a disulfide bond and conjugation by Michael Addition may overcome this circumstance. For example, a phosphine was used to reduce the disulfide bridge and the
two free thiols were reacted readily with an acrylate terminated mPEG in a Michael reaction [136].

A completely different approach uses the thiol group of cysteine as a chain transfer agent [137]. The polymerization is photo-induced and the polymer chain grows from the cysteine residue.

4.3.3 Tyrosine

Reactions on tyrosine occur either at the oxygen atom of the phenol unit or at the aromatic ring through electrophilic aromatic substitution (EAS). Thus, the reactions can be divided into $O$-alkylated and $C$-alkylated products next to aromatic substitution products.

A $\pi$-allyl species like allylic acetates, carbonates, and carbamates are inert towards amino acid functionalities until they are activated with a palladium catalyst like palladium acetate and triphenylphosphine tris(sulfonate) as a water-soluble ligand [138, 139]. With a palladium catalyst, the phenolate oxygen of tyrosine will be alkylated. The conjugate is an allyl aryl ether (14) (Table 4).

Instead of using an $O$-alkylation, several approaches use an EAS reaction. A three-component Mannich-type coupling forms a $C$-alkylated product [140]. At first an imine is generated in situ from an aldehyde and an electron-rich aniline. Then the imine acts as an electrophile and gets attacked by the aromatic tyrosine residue to yield the resulting secondary amine (15). An alternative route is based on a diazonium coupling [141, 142]. A diazonium salt is prepared by the reaction of an aromatic amine and sodium nitrite under acidic conditions and is then reacted with the tyrosine residue to result in an azo compound (16). This approach can also be used to add a small molecule with a new function to the protein if a hetero bis-functional diazonium salt is utilized. In addition, tyrosine can react with highly reactive electrophiles such acyclic diazodicarboxylate to yield the corresponding triazolidine compounds (17) [143].

In summary, the reactions mentioned are rarely used for polymer ligation at present but rather for small molecule conjugation. The reactions take place preferentially at the aromatic ring of tyrosine instead of tryptophan and phenylalanine.

4.3.4 Glutamine

The amide structure is not accessible for organic chemical reactions. Instead, an enzymatic approach is available. Transglutaminase (TGase) catalyze the acyl transfer between the $\gamma$-carboxamide of protein-bound glutamine and a primary amine resulting in the formation of a $\gamma$-amide of glutamic acid and ammonia (18) [144]. Unbranched primary amines act as acyl acceptors and are usually the $\varepsilon$-amino group of natural lysine [145]. Transglutaminases are a large family of enzymes and they were found in numerous organisms including mammals. Certain TGase species accept a wide variety of primary amines and thereby allow the possibility to use polymers with an amine end group [146]. A microbial TGase
| Table 4  Coupling methods for tyrosine residues |
|-----------------|-----------------|-----------------|
| Reaction type   | Reaction        | Product          | Lit.            |
| 14              | O-Alkylation    | Aryl-allyl ether | [138, 139]      |
|                 | ![O-Alkylation](image) | ![Aryl-allyl ether](image) |                 |
| 15              | Mannich         | Secondary amine  | [140]           |
|                 | ![Mannich](image) | ![Secondary amine](image) |                 |
| 16              | Azo-coupling    | Azo compound     | [141, 142]      |
|                 | ![Azo-coupling](image) | ![Azo compound](image) |                 |
| 17              | EAS             | Addition product | [143]           |
|                 | ![EAS](image)   | ![Addition product](image) |                 |
from *Streptomyces mobaraense* works independently of a cofactor and has a higher reaction rate; thus it is well suited as a catalyst for ligation [147]. An ideal candidate as a polymer is mono amino functionalized poly(ethylene glycol) and, accordingly, several protein-PEG conjugates have been reported in the literature [145, 148]. Normally only a few glutamine residues act as substrates for TGase. In addition, the selectivity can be increased by adding co-solvents [148] (Table 5).

4.3.5 Tryptophan

Tryptophan offers an indole side chain that can be used for ligation chemistry. A water-compatible rhodium carbene can be added to the indole ring (19) [105, 139]. The reactive species is generated in situ by a conjugated diazo compound by a rhodium catalyst like rhodium(II) acetate [63, 139, 149]. The reaction takes place in the two- and three-position of indole. Thus, a mixture of *N*-alkylated and *C*-alkylated product is obtained. It is necessary to add hydroxylamine hydrochloride as an additive to bind to the distal rhodium carbenoid complex. The usage of this salt lowers the pH value below 3.5 and therefore limits the scope of this methodology. As a side reaction, the carbene inserts into the O–H bond of water (Table 6).

4.3.6 Histidine

The imidazole side group is able to form stable complexes with transition metal ions. These metals ions are generally divalent ions like Zn$^{2+}$, Ni$^{2+}$, Co$^{2+}$, and Cu$^{2+}$, analogous to natural metalloproteins. Several histidines in a repetition motif increase the affinity. Proteins without a binding motive for metal ions can be fitted with a polyhistidine-tag (His-tag) that consists of six histidines in a row. The His-tag is added by a vector technique or during the PCR reaction using primers.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Coupling method for the glutamine group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction type</td>
<td>Reaction</td>
</tr>
<tr>
<td>Enzymatic</td>
<td>[ \text{R}^+\text{NH}_2 \xrightarrow{\text{TGase}} \text{R}^+\text{H} \xrightarrow{-\text{NH}_3} \text{Protein} ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Coupling method for the tryptophan group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction type</td>
<td>Reaction</td>
</tr>
<tr>
<td>Rhodium carbenoid reaction</td>
<td>[ \text{R}^+\text{N}_2 \xrightarrow{\text{Rh}_2(\text{OAc})_4} \text{H}_2\text{NOHxHCl} \xrightarrow{\text{Protein}} \text{N-}/\text{C}-\text{Alkylation} ]</td>
</tr>
</tbody>
</table>
containing the motif. This approach has been adopted from protein purification using metal ion affinity chromatography [150,151] (Table 7).

Necessarily, the polymer has to be fitted with a metal binding group. A requirement is that the chelation group in the polymer remains with some free chelating sites for binding to the protein. As a ligand, the imidazole group itself is suitable for polymer binding. It can be introduced as vinylimidazole, resulting in copolymers. A stimuli responsive polymer based on \( N \)-isopropylacrylamide is copolymerized with vinylimidazole [152]. The copolymer is loaded with copper(II) ions in form of copper sulfate that initialize the complex. Alternatively, the ion binding group can be introduced in the polymer by post-polymerization modification techniques. Active ester monomers like \( n \)-acryloxysuccinimide or pentafluorophenylacrylate have been homo- [153,154] or copolymerized [155] and converted in a post-polymerization reaction into a chelation ligand. A suitable compound is nitrilotriacetic acid with an anchor group attached to the backbone. Nitrilotriacetic acid offers four chelating sites and leaves two sites open for a nickel(II) central atom. As an example, His-tagged silicatein had been immobilized onto a polymer coated \( \gamma \)-Fe\(_2\)O\(_3\) nanoparticle [156,157].

### Table 7  Coupling method for the histidine group

<table>
<thead>
<tr>
<th>Reaction type</th>
<th>Complex structure</th>
<th>Product</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 Coordination reaction with transition metals such as nickel</td>
<td><img src="image" alt="Complex structure" /></td>
<td>Metal complex</td>
<td>[152–157]</td>
</tr>
</tbody>
</table>

4.3.7 **Aspartic Acid, Glutamic Acid and C-Terminus**

A contrary approach to conjugation to lysine and the N-terminus would be the conversion of the carboxylic acid group with amines and coupling agents like carbodiimides and \(N\)-hydroxybenzotriazole. Because of the nucleophilicity of lysine, inter- and intramolecular side reactions take place. Thus a selective pathway via carbonyl chemistry is absent and other direct chemical approaches are rare. Instead, enzymatic and biosynthetic methods were developed to enable single site modification at the C-terminus. New functional groups can be added at the C-terminus using intein-mediated protein splicing. In principle, the carboxylic acid can be converted into every chemical group and some examples are summarized below. A unique thioester group can be exposed with this procedure [158]. The soft nucleophile reacts with electron-deficient azides like sulfonazides. The reaction proceeds through the formation of a thiatriazoline intermediate that decomposes yielding an amide product under elimination of nitrogen gas and sulfur. PEG sulfonazide can be synthesized from amino-PEG and 4-carboxybenzensulphonylazide.
following routine coupling methods. Adding a phosphinothioester through intein-mediated protein splicing affords a Staudinger ligation [159]. Azides are the corresponding reacting agents. Also the grafting-from route is possible by adding an ATRP initiator via the intein route [160]. Thereby only a one-to-one conjugate is achieved.

4.3.8 Arginine

The difficulty for residue-selective modification of the guanidinium group in arginine is the lower reactivity compared to the ε-amino group of lysine. Thus, conventional reactions like the conversion with active esters and Michael addition cannot be realizable [161]. Instead, by taking advantage of a kinetic selectivity, a thermodynamic pathway is available. MPEG chains bearing an α-oxo-aldehyde end group react with arginine under mild conditions [161]. Adduct products formed from lysine and polymer as minor byproducts could be cleaved with hydroxylamine buffer at neutral pH value. Cysteine can also act as a nucleophile, but the formed adduct is unstable and will get cleaved. Further improvements are clearly needed, such as the elimination of possible branched structures, but it is an interesting approach to use an amino acid that could not been used for ligation reactions yet.

4.3.9 Phenylalanine

A selective addressing of the phenyl group of phenylalanine like organometallic cross coupling reactions is not possible through the coexistent occurrence of the other aromatic side chains. Pathways for introducing functional groups like iodination exist, but are only of interest for oligopeptide conjugation. Iodated side chains enable access to these groups by palladium coupling reactions [57].

4.3.10 Non-natural Amino Acids

The functional pool of organic groups can be expanded by introducing new groups with non-coded amino acids. Selectively a certain amount of ligation sites are incorporated in the polypeptide chain and, thus, they directly correlate with the desired amount of synthetic polymers per protein unit. Different techniques can be used to introduce non-natural amino acids [162]. The main approaches are translational incorporation by using analogous amino acids, new tRNA for a 21st amino acid, replacing a stop codon with a novel tRNA loaded with a new amino acid, expanding the size of codons for more different codes and non-natural base pairs [162]. Further semi-synthetic approaches are available with solid-phase peptide synthesis in combination with native chemical ligation.

The synthetic erythropoiesis protein (SEP), a permitted therapeutic agent, is an example for this technique. Two branched polymers with negative charged end
group are attached at the polypeptide chain [163]. The peptide chain was split into four segments, synthesized by solid-phase peptide synthesis, and ligated afterwards. Thus two non-natural derivatives of lysine could be integrated. The \( \text{Ne}-\text{levulinyl} \) modified lysine group bears a keto group for ligation chemistry. The keto group can react with aminooxy groups yielding the respective oximes. This route will be discussed in Sect. 4.4.

An amino acid bearing an azide group can be incorporated with para-azidophenylalanine muting a codon from an amino acid located on the surface to a stop codon and the corresponding tRNA with the novel amino acid [164]. Alkyne terminated PEG was then coupled by the [3+2] cycloaddition reaction to the protein.

A grafting-from approach is also possible by introducing an amino acid bearing an ATRP initiator moiety [165]. In this approach the non-natural amino acid was incorporated through translational transformation. As ATRP initiator, 2-bromo-isobutyric acid was used, which was attached to the 4-aminophenylalanine.

### 4.4 Protein–Polymer Ligation via the Indirect Pathway

In this two-step pathway, an amino residue is first converted with a low molecular weight compound to result in a new functionality. This reaction normally uses methods from the previous sections. In a second step the actual ligation reaction is conducted, resulting in the protein–polymer conjugate. Thus, new functional groups are available via chemical synthesis. Common groups for this category are azides and alkynes for click chemistry and carbonyl groups and aminooxy for oxime ligation. The Huisgen 1,3-dipolar cycloaddition uses azides and alkynes with copper (I) as a catalyst (20) [166]. Using the classical condition consisting of copper(II) sulfate and sodium ascorbate may lead to degradation of the protein component [167]. This is caused by the generation of reactive oxygen species like the hydroxyl radical by oxidation of the catalytic reactive species to the copper(II) state. Further, the reducing agent may influence the bioactivity of the protein. Thus, reaction conditions have to be chosen that avoid these problems. The ligand should stabilize the oxidation state of the catalytic copper(I) state, sequester the metal ion, prevent the protein for damage, and should not constrain the reaction rate. Such ligands are tris(triazolylmethyl)amine and bathophenanthroline disulfonate. Polymers with a group for click reaction are easily obtainable for ATRP [168]. The bromide end group from the ATRP initiator can be replaced by azide with post polymerization end group modification. RAFT polymers can be fitted with an azide group using a RAFT agent that bears the azide group [169]. In this case, the polymerization temperature should be kept strictly below a certain temperature to maintain the azide group. A protein can be equipped with an alkyne group by using an \( N \)-alkyne functionalized maleimide, which reacts with cysteine groups; see Sect. 4.3.2.

Another common method is the oxime formation from aldehydes or ketones and aminooxy compounds (21). The functionalities are orthogonal to the natural amino acid residues. Thereby, the aminooxy group can be located on the polymer or protein and
<table>
<thead>
<tr>
<th>Reaction type</th>
<th>Reaction</th>
<th>Product</th>
<th>Lit.</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 1,3-Dipolar cycloaddition</td>
<td>$\text{R}^1\overset{\text{NH}}{\text{N}}_2\text{N}_2$ + $\text{R}^2\text{N}^\equiv\text{Protein}$ $\text{Cu(I)}$</td>
<td>Triazole</td>
<td>[168, 169]</td>
</tr>
<tr>
<td>21 Carbonyl reaction</td>
<td>$\text{R}^1\overset{\text{O}}{-\text{NH}_2}$ + $\text{R}^2\text{K}$</td>
<td>Oxime</td>
<td>[170–172]</td>
</tr>
</tbody>
</table>

Functionalities could be also vice versa
grafting-to as well as grafting-from approaches are possible. The amino group of lysine can be converted with isopropylidene aminooxyacetic acid [170]. Thus, the aminooxy group is protected with acetone in the form of the corresponding isopropylidene compound. Deprotection is achieved by treatment with methoxylamine. The free aminooxy group reacts with an aldehyde-terminated PEG to yield the oxime conjugate.

In the opposite way, the aminooxy group can be attached to the polymer. Equipping an ATRP initiator with a BOC protected aminooxy group leads to a polymer whose protecting group can be removed with trifluoroacetic acid, resulting in an aminooxy end functionalized polymer [171]. The protein is fitted with a keto group using the N\textsubscript{ε}-levulinyl lysine route. The lysine side chain is converted with \textit{N}-hydroxysuccinimidyl ester levulinate to the levulinyl-modified protein. Both components react again to form an oxime linked conjugate. Another way for the oxime formation is the conversion of the N-terminal amine into an aldehyde using the enzyme pyridoxal-5-phosphate [172]. This aldehyde can then react with an aminooxy functionalized ATRP initiator. The resulting macro initiator system is able to be used in a grafting-from polymerization. This method allows formation of a one-to-one conjugate (Table 8).

5 Conclusions and Outlook

This review highlighted different synthetic routes towards protein–polymer conjugates. Even though there have been numerous chemical strategies described to conjugate different synthetic polymers onto proteins, it requires a careful selection of the right chemistry that is most suitable for a respective conjugation. As such, we have divided the various possible reactions into classes of functional groups present on protein surfaces, deriving from selected amino acid residues. It should therefore allow interested scientists to choose the right chemistry for their particular scientific problem.

Advances in this area are twofold. Suitable ligation chemistries must be compatible with both proteins and polymers. As such, scientists are encouraged to look beyond and receive inspiration from either scientific community. Development of further chemistries is continuously needed to meet the demand for the synthesis of highly defined protein–polymer conjugates. Clearly there are still limitations in the conjugation chemistry to differentiate between various accessible groups available on the surface of proteins, which are of particular importance when one-to-one conjugates are targeted. Given recent developments in the area, it can be concluded that we will surely see further new conjugation chemistries in the near future.

Acknowledgment  B. Jung gratefully acknowledges a fellowship funded through the Excellence Initiative (DFG/GSC 266).
References

88. Harris JM, Kozlowski A. Poly(ethylene glycol) and related polymers monosubstituted with propionic or butanoic acids and functional derivatives thereof for biotechnical applications. Patent 5672662


nanocrystals as a methodology for the biomagnetic separation of recombinant His-tagged proteins. J Magn Magn Mater 320:2339–2344


Bio-synthetic Polymer Conjugates
Schlaad, H. (Ed.)
2013, VIII, 196 p., Hardcover
ISBN: 978-3-642-34349-0