1

A Brief Introduction to Traditional Bioconjugate Chemistry

W. Russ Algar

1.1 Introduction

Bioconjugation is the process of linking or connecting a biological molecule with another moiety. These moieties may include other biomolecules (e.g., peptides), synthetic polymers (e.g., polyethylene glycol), and small molecules such as ligands (e.g., biotin), drugs, or fluorescent dyes, among a multitude of other possibilities [1]. While an extensive range of chemical reactions can be utilized for bioconjugation, the goal of this chapter is to briefly summarize some of the most stalwart and traditional reactions, highlighting important concepts and the strengths and weaknesses of each chemistry. Although there is no formal definition of "traditional" bioconjugate chemistry, a majority of these chemistries will satisfy two criteria: (i) reaction with a native functional group in a biomolecule under mild aqueous conditions; and (ii) use by many researchers over many years with continued application today. In this context, the following sections of this chapter discuss the most commonly targeted functional groups in biomolecules, the most popular chemical reactions for conjugation at those functional groups, and the cross-linking strategies most frequently used with those reactions. Extensive information on traditional bioconjugate chemistries can be found in a number of valuable resources, including Hermanson's classic tome, *Bioconjugate Techniques* [2], as well as similar volumes by other authors [3–5]. Importantly, this introductory chapter serves as a short primer for subsequent chapters that discuss more modern bioconjugation methods that have better chemoselectivity than the traditional methods discussed here. The development of such "nontraditional" chemistries has been motivated by the limitations of traditional chemistries. An understanding of traditional bioconjugate chemistries is therefore necessary to appreciate the utility of the various chemoselective and bioorthogonal reactions described in this book, as well as their applications.

1.2 Reactive Groups of Biomolecules

The native functional groups in target biomolecules are the primary sites for traditional bioconjugate reactions. This section describes the reactive functional

Chemoselective and Bioorthogonal Ligation Reactions: Concepts and Applications,
First Edition. Edited by W. Russ Algar, Philip E. Dawson, and Igor L. Medintz.

groups that naturally occur in the most common classes of biomolecules: peptides and proteins, carbohydrates, nucleic acids, and lipids. These functional groups are generally nucleophiles or electrophiles in and of themselves, such that the reactions of a particular functional group in a protein will be the same reactions that can be used with that functional group in a nucleic acid, lipid, or carbohydrate. Optimization of those reactions and the scope of their applicability can vary from biomolecule to biomolecule.

1.2.1 Peptides and Proteins

Natural peptides and proteins are biopolymers that are largely derived from the 20 canonical amino acids [6]. For the purposes of bioconjugation, the polyamide backbone of a protein or peptide is unreactive, with the two notable exceptions of the N-terminal amine group and the C-terminal carboxyl group. Consequently, the side chains of amino acids tend to be the most prominent sites for bioconjugation [7]. Potential side-chain nucleophiles include the thiol and thioether groups of cysteine (Cys) and methionine (Met); the amine groups of arginine (Arg), histidine (His), lysine (Lys), and tryptophan (Trp); and the hydroxyl and phenol groups of serine (Ser), threonine (Thr), and tyrosine (Tyr). Although each of these side chains is nucleophilic in principle, the strength and practical utility of each nucleophile vary with pH, other reaction conditions, and the reactivity of the corresponding electrophile. Considering the remaining canonical amino acids, aspartic acid (Asp) and glutamic acid (Glu) have side-chain carboxyl groups that can be activated for reaction with amine nucleophiles. The amide side chains of asparagine (Asn) and glutamine (Gln), as well as the hydrogen (Gly), alkyl (Ala, Ile, Leu, Pro, Val), and phenyl groups (Phe) of the other canonical amino acids, are generally unreactive toward traditional bioconjugate chemistries. Figure 1.1 shows the structures of the 20 canonical L-amino acids, the N-terminus, and the C-terminus as part of oligopeptide chains. The approximate pK_a values for the conjugate acid forms of the side chains are also shown [8], and the amino acids are drawn in the ionization state that dominates at neutral pH in aqueous solution. The reactivity of the various amino acid side chains can vary considerably with their location in a protein and interactions with neighboring amino acid residues [7].

For traditional bioconjugate reactions, the most important nucleophilic amino acid residues are cysteine, lysine, and the N-terminus. These residues have been the most frequently targeted for bioconjugation, and reactions with other amino acid residues are often undesired side reactions. Thiols (R–SH), and the thiolate anion ($R–S^-$) in particular, are the strongest biological nucleophiles [8, 9]. Primary amines ($R–NH_2$) are also good nucleophiles; however, the corresponding aminium ion ($R–NH_3^+$) is a poor nucleophile [10, 11]. Consequently, pH is an important determinant of the products of bioconjugation, as well as the efficiency of many bioconjugate reactions. The nominal pK_a of the ε-amine of a lysine side chain is ~9.4 [8]; however, the actual value varies between individual lysine residues in a protein because of interactions with neighboring amino acid residues (e.g., hydrogen bonding) and the local environment [7]. The nominal pK_a of the N-terminus is lower at ~7.8 [8]. Actual pK_a values can differ from

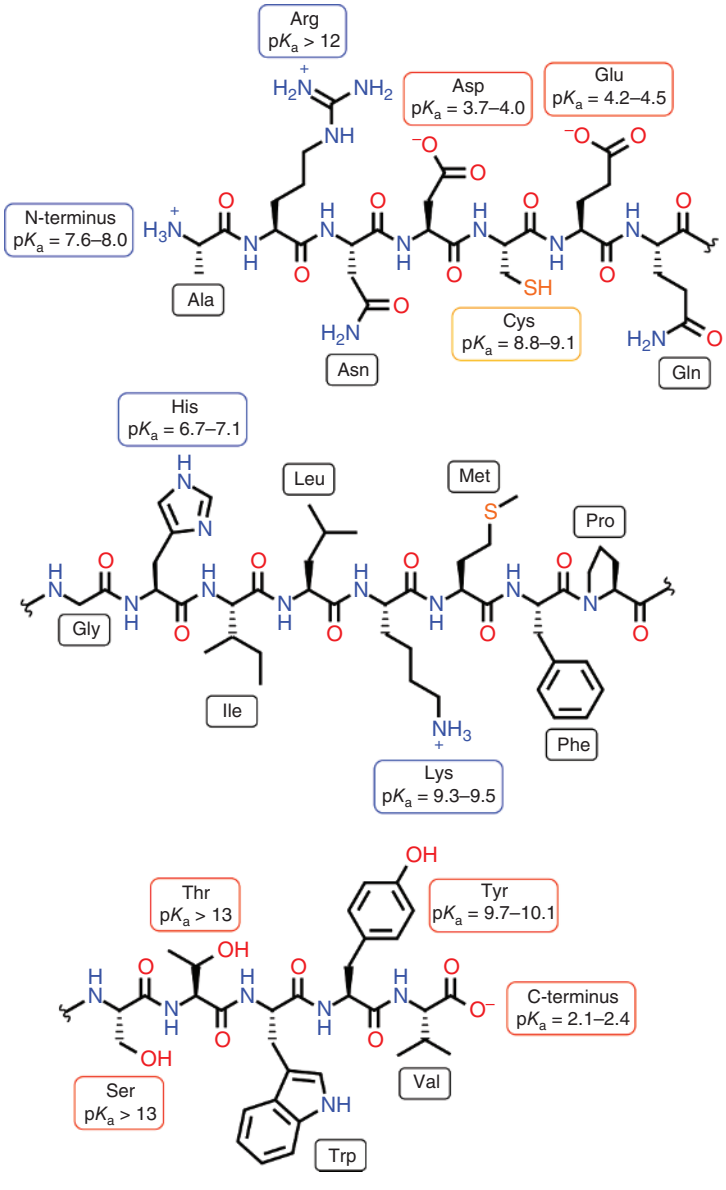

Figure 1.1 Peptide chains illustrating the structure of the 20 canonical L-amino acids, the N-terminus, and the C-terminus. The amino acid residues are linked by stable amide bonds and differ in the structure of their side chains. For each ionizable side chain, the predominant ionization state at pH 7.0 is shown, and the approximate pK_a value is listed. Most bioconjugate reactions target functional groups associated with the side chains.

nominal pK_a values by up to several units, and lysine residues can be reacted at pH values lower than expected from the pK_a of the isolated amino acid. It is often suggested that pH > 8.0 is required for efficient conjugation to lysine side chains, whereas pH 7.0 and above can suffice for the N-terminus and that this difference is a potential means of selectively reacting the N-terminus [11]. Similarly, the good reactivity of thiols at neutral pH can permit selective labeling of cysteine residues in the presence of abundant lysine residues [11].

The guanidine group of arginine has pK_a > 12, such that it exists as a protonated guanidinium cation under most aqueous conditions and is thus a poor nucleophile for most reactions [11]. Glyoxals and other α-dicarbonyl compounds can react with arginine residues [7], but this reaction is not commonly used for bioconjugation. In the case of histidine and tryptophan, their aromatic amines (imidazole and indole) are much less reactive than the aliphatic amine of lysine [4, 11]. As such, the foregoing residues are not usually modified in the acylation reactions frequently used to modify lysine; however, some potent alkylating agents can still react with these residues under certain conditions, as well as the thioether group of methionine, which is normally a weak nucleophile [3]. The hydroxyl and phenol side chains of serine, threonine, and tyrosine are also poorly nucleophilic in aqueous solution [4]. Tyrosine, with its lower side chain pK_a, is the more reactive of these amino acids, although its reactivity is often hindered by being located within the hydrophobic interior of folded proteins [3].

The carboxyl groups of glutamic acid and aspartic acid side chains are not reactive without activation. The most common activating agents are water-soluble carbodiimides, which can directly mediate coupling reactions between carboxyl groups and nucleophiles such as amines and hydrazides [2, 11]. Alternatively, carbodiimide reagents can mediate the transformation of carboxyl groups into succinimidyl esters, which also react with amines and hydrazides (see Section 1.3.1).

Figure 1.2 shows two examples of protein structures and highlights their lysine, aspartic acid, glutamic acid, and cysteine residues. The visual impression from these two proteins is quite general – there are numerous lysine, aspartic acid, and glutamic acid residues in most proteins and far fewer cysteine residues. Indeed, cysteine is the second least abundant amino acid residue in proteins [9], whereas lysine has an abundance of nearly 6% [8]. Fewer still is the number of cysteine residues that are not tied up in disulfide bridges [8]. As shown in Figure 1.3, disulfides can be chemically reduced with reagents such as dithiothreitol (DTT) [12] and tris(2-carboxyethyl)phosphine (TCEP) [13] to generate reactive thiols; however, this process can potentially affect protein structure and function. Compared with DTT, TCEP is advantageous in that it is odorless, more stable, and more potent as a reducing agent over a wider range of pH, and may not need to be removed from the protein solution prior to subsequent steps in protocols [14]. Some sources suggest that TCEP does not interfere with maleimide and iodoacetyl coupling to thiols (see Section 1.3.2) [14], whereas others have reported side reactions [15]. In contrast, excess DTT must always be removed because of its thiol groups [15]. When no native cysteine residues exist, or when reduction of disulfides is not feasible, a cysteine residue can be introduced into a protein using site-directed mutagenesis [16]. It is also possible to expand the

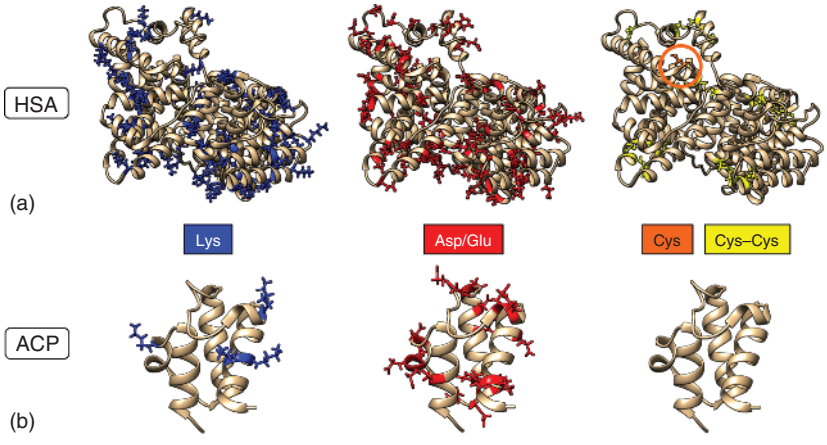

Figure 1.2 Two examples of proteins, (a) human serum albumin (HSA) (Protein Data Bank ID 1AO6) and (b) *E. coli* acyl carrier protein (ACP) (Protein Data Bank ID 1T8K). The structures highlight the abundance of lysine residues (Lys, blue), aspartic acid (Asp, red), and glutamic acid (Glu, red) residues, as well as the scarcity of cysteine residues (Cys), particularly residues with available thiol groups (orange, circled) versus those that are part of disulfide bridges (Cys–Cys, yellow). Note that ACP has no cysteine residues.

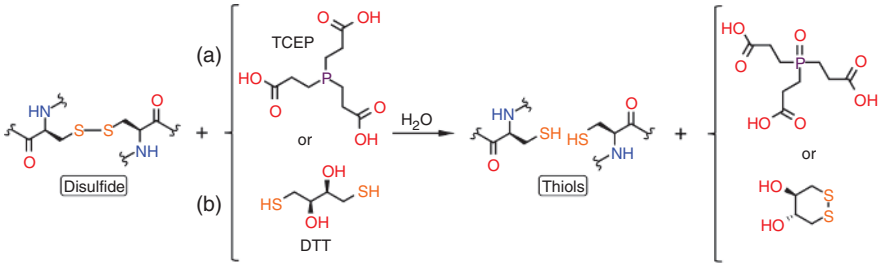

Figure 1.3 Reduction of disulfide bonds to thiols using (a) TCEP or (b) DTT.

palette of functional groups available for bioconjugation through the inclusion of unnatural (i.e., noncanonical) amino acid residues [17–20]. Unnatural amino acid residues can be chemically added to growing oligopeptides during solid-phase synthesis, and can be genetically or metabolically incorporated into expressed proteins. In this manner, new functional groups can be introduced and can be selected to be suitable for specific bioconjugation reactions, have equal or greater scarcity than cysteine residues, and have reactivity different from the canonical amino acids. These techniques can be used to enable chemoselective and bioorthogonal chemistries [21], but are otherwise beyond the scope of this introductory chapter.

1.2.2 Carbohydrates

As a class of molecule, carbohydrates include monomeric saccharides and their dimers, oligomers, and polymers [6]. Polysaccharides are often referred to as glycans. The most common monomer residues in oligosaccharides and

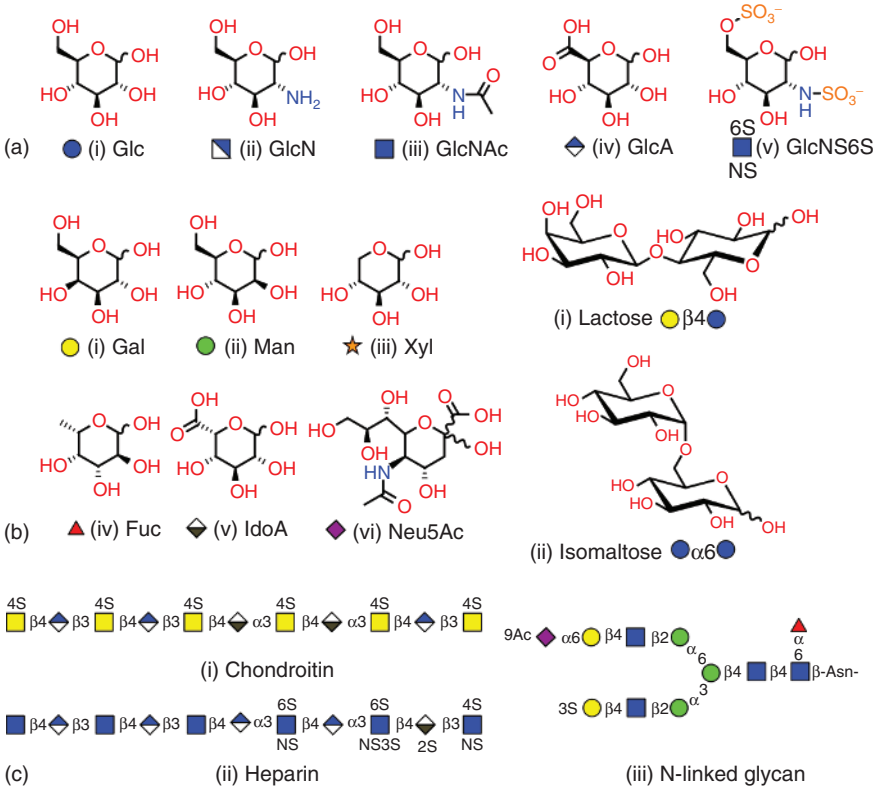

Figure 1.4 (a) The structures and symbolic notations for (i) glucose (Glc), (ii) glucosamine (GlcN), (iii) *N*-acetylglucosamine (GlcNAc), (iv) glucuronic acid (GlcA), and (v) 6-*O*,2-*N*-disulfated glucosamine (GlcNS6S). (b) The structures of (i) galactose (Gal), (ii) mannose (Man), (iii) xylose (Xyl), (iv) fucose (Fuc), (v) iduronic acid (IdoA), and (vi) *N*-acetylneuraminic acid (Neu5AC, a sialic acid). (c) Structures of (i) lactose, which has Gal and Glc residues joined by a β-1 → 4 glycosidic linkage, and (ii) isomaltose, which has two Glc residues joined by an α-1 → 6 glycosidic linkage. (d) Structures of two glycosaminoglycans, (i) chondroitin and (ii) heparin, and (iii) an example of an N-linked glycan.

polysaccharides are hexoses (e.g., glucose, mannose, galactose), pentoses (e.g., ribose, xylose), and many derivatives thereof. Figure 1.4a shows the structure of glucose (Glc), a hexose monomer, and four of its derivatives: glucosamine (GlcN), *N*-acetylglucosamine (GlcNAc), glucuronic acid (GlcA), and 6-*O*,2-*N*-disulfated glucosamine (GlcNS6S). Figure 1.4b shows the structures of other common monosaccharides, which differ from glucose in the number of carbon atoms and the number, position, and stereochemistry of hydroxyl groups. Galactose and mannose have derivatives largely analogous to those of glucose. In oligosaccharides and polysaccharides, saccharide monomers are highly repetitive and are linked through glycosidic bonds at different positions (e.g., 1 → 4 or 1 → 6 linkages) and with different stereochemistry at the linked carbon atoms (i.e., α- or β-). Figure 1.4c shows two disaccharides, lactose and isomaltose, with different glycosidic bonds between the two monomers. Carbohydrates can

be linear or branched and can be found as discrete molecules or attached to other biomolecules. The latter are called glycoconjugates and include glycolipids, glycoproteins, proteoglycans, glycopeptides, and peptidoglycans [6]. The addition of *N*-acetylglucosamine to serine and threonine residues is a common posttranslational modification of proteins. The nucleotides that comprise deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) also contain a saccharide in their structure (see Section 1.2.3).

The most abundant functional groups in carbohydrates are hydroxyl groups, which, as noted earlier, are generally poor nucleophiles in aqueous solvent. Many monosaccharide derivatives do not introduce any new reactive functionality, as is the case for deoxy, *N*-acetylamino, and sulfo derivatives. In contrast, non-acetylated amino sugars and sugar acids provide additional reactive groups for bioconjugation. Sialic acids, a class of acidic monosaccharide derivatives, are found as terminal saccharide residues in glycoproteins and glycosphingolipids. Amino sugars and sugar acids are major constituents of glycosaminoglycans. Figure 1.4d shows shorthand notations for two examples of glycosaminoglycans and an example of an N-linked glycan attached to an asparagine residue of a protein. As shown in the figure, the limited diversity of reactive functional groups and highly repetitive nature of many glycans are not amenable to targeting bioconjugation to specific sites.

To compensate for the poor nucleophilicity of the hydroxyl groups of carbohydrates, electrophilic reactivity has often been exploited for bioconjugation. Carbohydrates with a reducing end undergo isomerization between a cyclic hemiacetal form and an open aldehyde or keto form, as shown in Figure 1.5, with the equilibrium favoring the cyclic form in aqueous media [6]. Primary amine and hydrazide nucleophiles can react with this carbonyl group (see Section 1.3.1) [22, 23], and conversion of the anomeric hydroxyl group to an amine group is also possible through reaction with ammonium carbonate [24, 25]. The amine derivative can then undergo subsequent bioconjugation reactions (see Section 1.3.1). The main drawbacks of these methods are that they are not applicable to carbohydrates without a reducing end and that reaction rates can be slow, sometimes requiring days at room temperature and high concentrations of nucleophile.

Another strategy is to introduce new electrophilic groups to a carbohydrate, for example, via the use of sodium periodate as a mild oxidant to convert vicinal diols into aldehyde groups [2, 4, 5], as shown in Figure 1.6. Aldehyde and keto groups

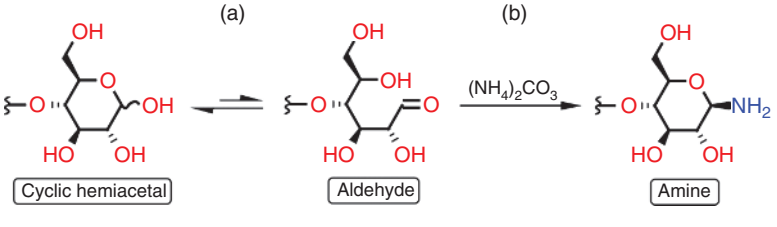

Figure 1.5 (a) The reducing end of a carbohydrate exists in equilibrium between cyclic hemiacetal and aldehyde forms. This equilibrium is shown for a glucose residue. The aldehyde group can react with amine nucleophiles (not shown; see Section 1.3.1). (b) The reducing end can be modified to an amine using ammonium carbonate.

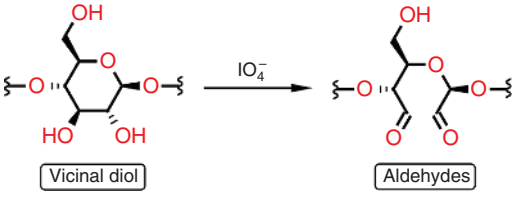

Figure 1.6 Oxidation of the vicinal diols in a carbohydrate to aldehyde groups using sodium periodate. Nucleophiles can react with the aldehyde groups (see Section 1.3.1). In this case, the periodate oxidation of a glucose residue is shown.

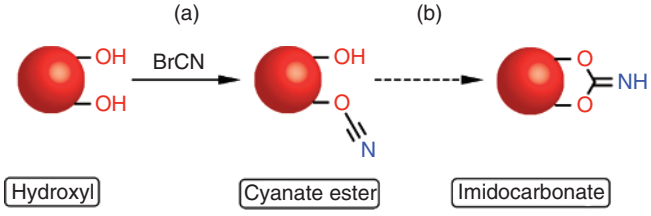

Figure 1.7 Activation of hydroxyl groups with cyanogen bromide to (a) amine-reactive cyanate esters and (b) amine-reactive cyclic imidocarbonates.

react with amine, hydrazide, and aminooxy nucleophiles to form imine, hydrazone, and oxime linkages, respectively. These reactions are described in Section 1.3.1. Since many carbohydrate residues have vicinal diols, periodate chemistry is advantageous in that it is widely applicable, but potentially disadvantageous in that a multitude of residues in a carbohydrate chain are subject to modification. For glycoproteins, periodate chemistry is a potential means of selectively modifying the protein at the carbohydrate residues [26]. Of note, periodate can not only oxidize terminal serine and threonine residues to aldehydes but can also oxidize other amino acid residues (Tyr, Trp, His, Met, Cys) [27].

As an alternative to aldehydes, hydroxyl groups can be converted into amine-reactive cyanate esters using cyanogen bromide or 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP) [28]. A potential side product of cyanate ester formation is a cyclic imidocarbonate, which also reacts with amines [29] and tends to dominate when activating polysaccharides that have vicinal diols [3]. The activation of hydroxyl groups with cyanogen bromide is shown in Figure 1.7. Formation of the cyanate ester or imidocarbonate is in competition with hydrolysis and other side reactions that can lead to both unreactive and reactive by-products [29]. Cyanogen bromide itself is also highly susceptible to hydrolysis, particularly under the alkaline conditions needed to deprotonate carbohydrate hydroxyl groups ($pK_a \sim 12$) for efficient cyanylation [29]. CDAP is preferable in that it is more stable, is less toxic, and requires a less alkaline pH for effective cyanylation of carbohydrates than cyanogen bromide [29]. Further discussion on the reaction of amines with cyanate esters and imidocarbonates can be found in Section 1.3.1.

1.2.3 Nucleic Acids

DNA and RNA are the chief carriers of genetic information, and both comprise a sugar–phosphate polymer backbone with pendant purine and pyrimidine bases, as shown in Figure 1.8 [6]. The pyrimidine bases include uracil (U), cytosine

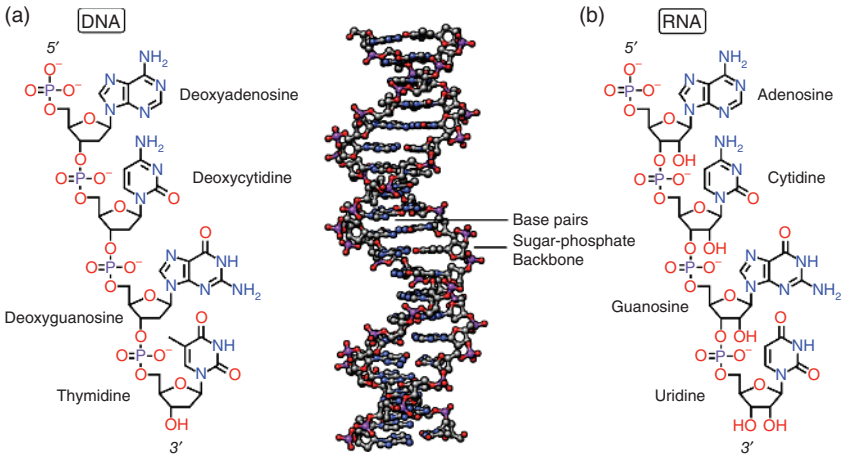

Figure 1.8 Chemical structures of (a) DNA and (b) RNA strands. The structural model in panel (a) shows a ball-and-stick model of a double-stranded DNA helix that is 20 base pairs in length. Hydrogen atoms have been omitted for clarity. Two complementary strands of nucleic acid align antiparallel to one another and hybridize through Watson–Crick base pairing to form the double helix.

(C), and thymine (T); the purine bases include adenine (A) and guanine (G). Native DNA has a phosphate group at its 5′-terminus and has a hydroxyl group at its 3′-terminus. In the case of RNA, the 3′-terminus has a vicinal diol. As a consequence of the differences between ribose and deoxyribose, RNA is highly susceptible to chemical hydrolysis and enzymatic degradation, whereas DNA is much more stable [6]. Single strands of both DNA and RNA can hybridize with complementary sequences through Watson–Crick base pairing (A–T or A–U, G–C) to form double-stranded helical structures. The hydrophobic nucleobases are hydrogen bonded and pi-stacked with one another in the interior of the helix [30].

Compared with proteins, native nucleic acids are not as readily modified by chemical means. Terminal modification is possible through two main routes: carbodiimides can activate the 5′-phosphate group toward reactions with amines [31], and sodium periodate can also oxidize the vicinal diol at the 3′-terminus of RNA to yield amine- and hydrazide-reactive aldehyde groups (see Section 1.3.1) [32]. Chemical reactions with the nucleobases are also possible but may require single-stranded nucleic acid so that the pertinent functional groups are more accessible and not involved in hydrogen bonds. Cytosine can be treated with sodium bisulfite for conversion to 6-sulfo-cytosine, which undergoes transamination reactions [33, 34]. Adenine and guanine can be brominated using aqueous bromine or *N*-bromosuccinimide, and amines can be coupled to the brominated nucleobases at elevated temperature [2]. Given that cytosine and guanine bases are repeated frequently in a nucleic acid sequence, chemical modification of specific sites is generally not possible.

More commonly, molecular biology techniques that rely on enzymes have been used for labeling native nucleic acids at the 5′-terminus, 3′-terminus,

or random positions [2, 35]. The site of labeling depends on the enzyme and the state of the DNA (i.e., single- or double-stranded, blunt, or sticky ends). Enzymatic methods are well suited for small-scale labeling but cannot be scaled up to the same degree as chemical labeling. Historically, enzymatic methods were used primarily for radiolabeling but are now frequently used for labeling with biotin, digoxigenin, or fluorescent dyes using modified nucleotides. Commonly used enzymes have included terminal deoxynucleotidyl transferase, T4 RNA ligase, T4 polynucleotide kinase, and DNA and RNA polymerases for methods such as nick translation, random priming, and end labeling [36, 37]. Enzymatic methods may be combined with chemical labeling; for example, the enzymatic incorporation of nucleotide analogs with specific functional groups for subsequent chemical reactions. Further discussion of enzymatic methods is beyond the scope of this chapter.

Synthetic oligonucleotides, prepared via solid-phase synthesis with nucleotide phosphoramidites [38, 39], are much more readily modified by chemical means than native nucleic acids. Functional group-terminated linkers can be attached to purine or pyrimidine bases, the phosphate backbone, or the 3′- or 5′-terminus using standard phosphoramidite chemistry [40, 41]. Common modifications are aminoalkyl or thioalkyl linkers that permit further modifications (e.g., labeling with a fluorescent dye) or attachment to solid surfaces [10]. Other functional groups suitable for chemoselective and bioorthogonal chemistry (e.g., azides or alkynes for cycloaddition reactions; see later chapters) can also be introduced into synthetic oligonucleotides [42]. The diversity and widespread availability of modified synthetic oligonucleotides is such that these molecules can be used to label native nucleic acids through Watson–Crick base pairing and selective hybridization [43]. In other instances, double-stranded structures can be exploited for non-specific bioconjugation. For example, psoralen intercalates into double-stranded DNA, initially through non-covalent interactions, but forms new covalent bonds with pyrimidines (especially thymine) upon UV irradiation [44]. The psoralen can be linked to other functional molecules such as biotin for further bioconjugation [45]. The repetition of nucleotides precludes targeting bioconjugation to a specific site.

1.2.4 Lipids

The most common naturally occurring lipids are phospholipids with a glycerol backbone linked to a phosphate headgroup and two fatty acid tails that vary between 16 and 24 carbons in length with varying degrees of unsaturation [6]. The phosphate headgroup is often linked to other polar functional groups, as is the case in phosphatidylcholine (PC), phosphatidylglycerol (PG), phosphatidylethanolamine (PE), phosphatidylserine (PS), and phosphatidylinositol (PI). The structures of these lipids are shown in Figure 1.9a. The lipid headgroups are the sites for bioconjugation as the fatty acid chains are largely unreactive and not exposed to aqueous solution, instead packing together to form lipid bilayers of the type shown in Figure 1.9b. The phosphate headgroups of phosphatidic acid can be activated toward reaction with amines using carbodiimides, PE has a reactive amine group, and PS has both amine and carboxyl groups

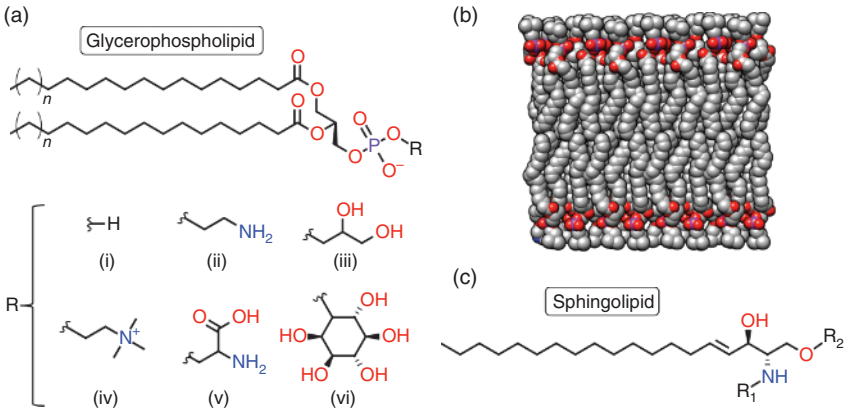

Figure 1.9 (a) General structure of a glycerophospholipid. The R group varies between different lipids: (i) phosphatidic acid, (ii) phosphatidylethanolamine (PE), (iii) phosphatidylglycerol (PG), (iv) phosphatidylcholine (PC), (v) phosphatidylserine (PS), and (vi) phosphatidylinositol (PI). (b) Space-filling model of a PC lipid bilayer. Hydrogen atoms have been omitted for clarity. (c) General structure of a sphingolipid. The R_1 groups are fatty acid residues and the R_2 headgroups are similar to those found for phospholipids (e.g., choline, carbohydrates).

[2]. Glycolipids, which often have vicinal diols as part of the saccharide or polysaccharide component of their headgroup, and PG, which has a vicinal diol in its glycerol headgroup, can be treated with sodium periodate to yield amine-reactive aldehyde groups (see Section 1.2.2) [46]. Another common class of lipids, called sphingolipids, have a sphingosine backbone, often with a second fatty acid chain attached [6]. The general structure of this class of lipid is shown in Figure 1.9c. Functional groups attached to sphingolipid headgroups are also sites for bioconjugation.

1.3 Traditional Bioconjugate Reactions

This section provides an overview of several traditional bioconjugate reactions, including their chemoselectivity (or lack thereof). The reactions are organized according to whether the nucleophile is an amine ($R–NH_2$) or another nitrogen nucleophile (e.g., $R–NH–NH_2$), a thiol (R–SH), or a hydroxyl (R–OH). Each reaction is effectively modular. Provided that the requisite functional groups are present, a given reaction can be used to conjugate biomolecules with one another; label a biomolecule with a small molecule such as a fluorescent dye, contrast agent, or drug; or immobilize a biomolecule on a surface or within a matrix. For this reason, the reactions in this section are discussed and illustrated generically. One manifestation of this modularity is shown in Figure 1.10. Fluorescein, a common fluorescent label, can be prepared with different reactive groups for bioconjugation. For example, it can be made to react as an amine, or it can be made amine-reactive, as needed, without significantly altering its utility as a fluorescent label. The chemistries associated with each reactive group in Figure 1.10, as well

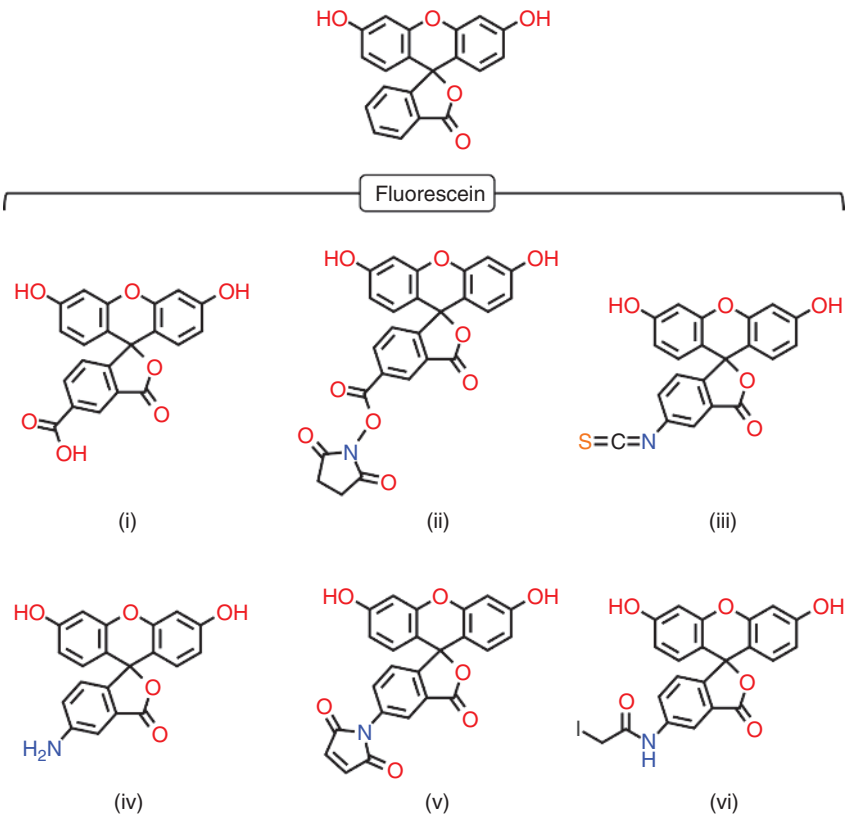

Figure 1.10 Various reactive derivatives of fluorescein, a popular fluorescent dye: (i) carboxyfluorescein, which can be activated for reaction with amines; (ii) amine-reactive fluorescein succinimidyl ester; (iii) amine-reactive fluorescein isothiocyanate (FITC); (iv) fluoresceinamine, which can be coupled with activated carboxylic acids; (v) thiol-reactive fluorescein maleimide; and (vi) thiol-reactive fluorescein iodoacetamide.

as those for many others, are discussed in this section. When appropriate, considerations for optimization of these reactions are noted; however, explicit reaction conditions are not noted, as these can be found in other resources [2, 4, 47] and tend to vary on an application-by-application basis.

1.3.1 Amines and Other Nitrogen Reagents

As noted earlier, primary amine groups are one of the most common nucleophiles in biomolecules. Carbonyls, active esters, and isothiocyanates are typically reacted with amine groups for purposes of bioconjugation [2–5], and many reagents are available with these functional groups. Conversely, carbonyl groups and active esters (from carboxyl groups) are either available or can be introduced to many biomolecules, and will react with both amine and hydrazide groups [2–5]. A variety of hydrazide reagents are also available and, being less basic than amines ($pK_a < 6$ vs $pK_a > 9$), tend to react more efficiently at lower pH [11, 48].

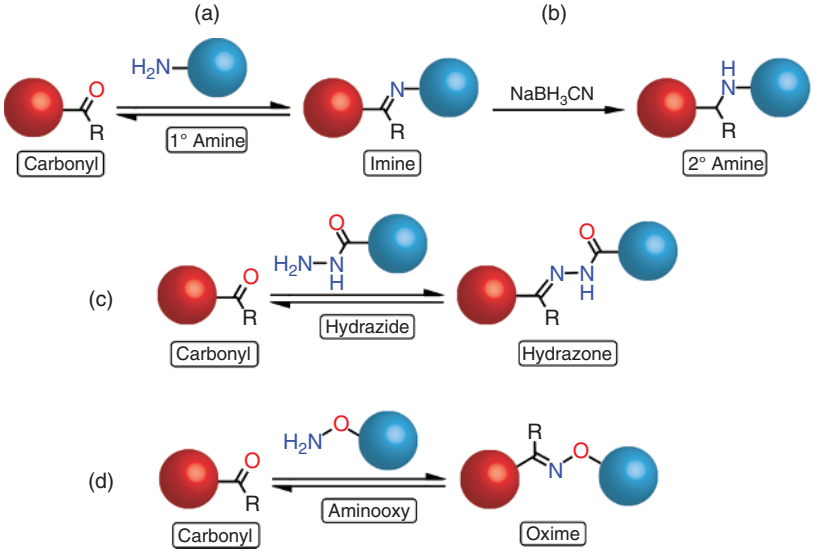

Figure 1.11 (a) Reaction between a carbonyl (aldehyde or ketone) and a primary amine to form an unstable imine, followed by (b) reduction to a stable secondary amine with sodium cyanoborohydride. (c) Reaction between a carbonyl and a hydrazide to form a hydrazone bond. (d) Reaction between an carbonyl and aminooxy group to form an oxime.

1.3.1.1 Aldehydes and Ketones

Primary amines will spontaneously react with aldehydes and ketones to form imines, also known as Schiff bases, as shown in Figure 1.11a [2, 11, 26]. Although these reactions will proceed in aqueous media, the reaction is reversible, and the imines are ultimately unstable as the equilibrium shifts to the unconjugated amine and carbonyl groups [2, 4]. To address this shortcoming, reductive amination is usually carried out either as a one-pot or two-step reaction with sodium cyanoborohydride, yielding a stable secondary amine, as shown in Figure 1.11b [49, 50]. Other nitrogen nucleophiles, such as hydrazide and aminooxy groups, react with aldehydes and ketones to yield hydrazone and oxime bonds, shown in Figure 1.11c,d, that are less susceptible to hydrolysis than imines [11, 26, 51]. At the expense of slower reaction kinetics, the stability of the oxime exceeds that of the hydrazone [51]. When desired, hydrazones can be reduced to stable hydrazides with sodium cyanoborohydride [52].

1.3.1.2 Active Esters of Acids

The most common bioconjugate reactions of amines are those with an activated carboxylic acid. Figure 1.12a illustrates the activation of a carboxylic acid with a carbodiimide to form an *O*-acylisourea intermediate that reacts with primary amines to form a very stable amide linkage [53]. Hydrazide nucleophiles will react analogously to primary amines [54]. Water-soluble *N*-(3-dimethylaminopropyl)-*N′*-ethylcarbodiimide (EDC) is the carbodiimide reagent of choice for most bioconjugate reactions. Unfortunately, both the EDC and the *O*-acylisourea intermediate are unstable, and hydrolysis is a major

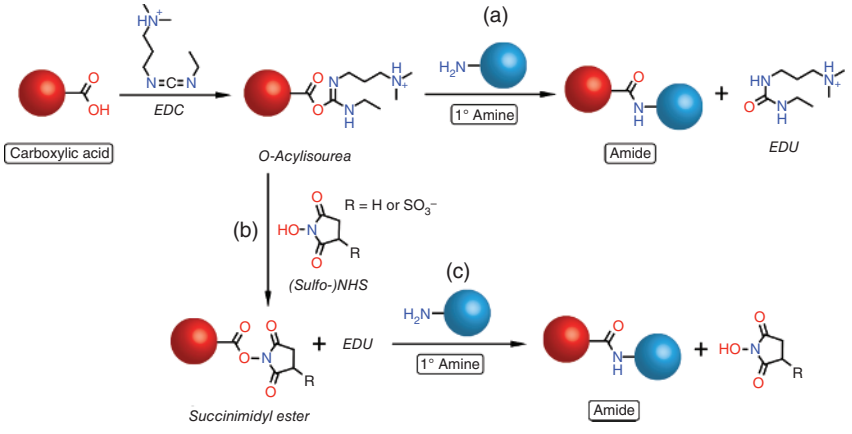

Figure 1.12 *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC) activates a carboxylic acid to an *O*-acylisourea intermediate that can react with (a) a primary amine to yield an amide or (b) *N*-hydroxysuccinimide (NHS) or sulfo-NHS to yield a more stable but still reactive succinimidyl ester. (c) The succinimidyl ester reacts with an amine to yield an amide.

competing reaction [53, 55], generally necessitating excesses of carbodiimide. Other factors in the optimization of these reactions are temperature, pH, and buffer selection [56]. Activation of the carboxylic acid is reported to be optimal at pH 4.5–6.0 [57]; however, the reaction remains feasible at pH 7.0–7.5, which is often more suitable for the target biomolecule. In addition to the obvious exclusion of primary amines and carboxylic acids from reaction buffers, phosphate salts should generally be avoided as they can react with carbodiimides.

A method of improving the efficiency of carbodiimide conjugation reactions is to convert the *O*-acylisourea intermediate into a more stable succinimidyl ester [58], as shown in Figure 1.12b,c. This procedure can be as simple as adding *N*-hydroxysuccinimide (NHS) or its water-soluble sulfonated analog (sulfo-NHS) to a reaction mixture with EDC, forming the succinimidyl ester *in situ* for reaction with the amine reagent. Alternatively, two-step conjugation procedures are sometimes utilized, where EDC and (sulfo-)NHS are first added to the carboxylic acid reagent and the amine reagent is added in the second step, both with and without separation of the succinimidyl ester intermediate from excess reagents prior to adding the amine. Although succinimidyl esters are more stable toward hydrolysis than *O*-acylisoureas, hydrolysis is still a competing reaction [2, 11]. Rates of hydrolysis increase with increasing pH, as does amine reactivity [59, 60], such that reactions are typically carried out between pH 7.0 and 9.0, with pH 8.0–8.5 suggested to be optimal for most bioconjugate reactions. The rate of hydrolysis may limit the efficiency of these reactions at pH > 9 [61]. Succinimidyl esters have slow reaction rates with alcohols, phenols, and aromatic amines [11], such that there are usually minimal side reactions with non-lysine side chains in peptides and proteins under aqueous conditions.

The 5′-phosphate group of nucleic acids can be activated by carbodiimides such as EDC, in the presence of imidazole or NHS, to yield a phosphorimidazolide or succinimidyl ester intermediate that can react with amines to form a

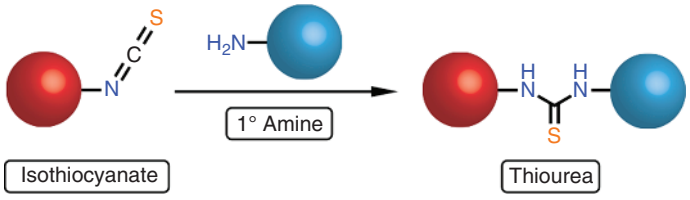

Figure 1.13 Reaction between an isothiocyanate and a primary amine to form a thiourea.

phosphoramidate bond [2, 31]. The imidazole or NHS is required for the conjugate reaction to efficiently compete with hydrolysis due to high reactivity of the intermediate phosphodiester.

1.3.1.3 Isothiocyanates

Amines will react with isocyanates and isothiocyanates to form isourea and isothiourea linkages, respectively [4, 26]. Isoureas are very susceptible to hydrolysis, and, for this reason, isocyanates are rarely used for bioconjugation purposes [4]. In contrast, thiourea linkages are much more stable toward hydrolysis, albeit that some hydrolysis still occurs at acidic pH and that the final thiourea conjugates have been found to be less stable than amide conjugates [4]. Figure 1.13 illustrates the reaction between an amine and an isothiocyanate, which is most efficient at pH 9.0–9.5 [3, 11]. Rates of hydrolysis of isothiocyanates are slower than for succinimidyl esters [11]. This conjugation reaction is perhaps best known for the fluorescent labeling of proteins with fluorescein isothiocyanate (FITC) and tetramethylrhodamine isothiocyanate (TRITC).

1.3.1.4 Other Reactive Groups

Amines will react with many other functional groups in addition to those noted earlier, albeit that these other functional groups are used much less frequently for bioconjugation reactions. As an example, sulfonyl chlorides react with amines to form extremely stable sulfonamides; however, this chemistry is limited by its high reactivity, as sulfonyl chlorides hydrolyze rapidly and will also modify histidine and tyrosine in addition to lysine and other primary amines [2, 61]. Epoxides are another functional group that will react with primary amines, yielding a secondary amine. This chemistry is most commonly employed for the immobilization of biomolecules on surfaces [62, 63]. Again, chemoselectivity is poor, as epoxides will also react with thiol and hydroxyl groups, and can hydrolyze to diols, depending on pH and other reaction conditions. Imidoesters are more stable toward hydrolysis than succinimidyl esters and react highly selectively with primary amines at basic pH to form amidines, but suffer from slow reaction rates and the susceptibility of the amidine product to hydrolysis [2, 26, 64]. Cyanate esters react with amine nucleophiles to form an isourea under mild conditions (pH 7–8), as shown in Figure 1.14a [2, 3, 29]. The isourea bond is susceptible to hydrolysis and is unstable, resulting in a steady loss of conjugation. The cyclic or linear imidocarbonate side products of cyanate ester formation, although less reactive than cyanate esters, will still react with amines to form an N-substituted imidocarbonate, as shown in Figure 1.14b [2, 29].

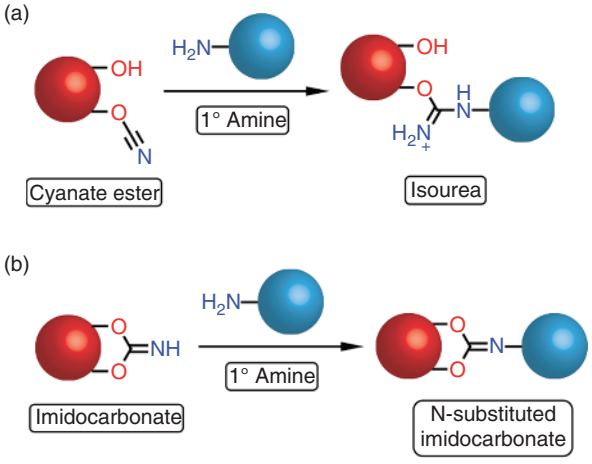

Figure 1.14 Reaction between an amine with (a) a cyanate ester to form an isourea and (b) a cyclic imidocarbonate to form an N-substituted imidocarbonate.

N-substituted carbamates and the isourea are other potential products of the amine reaction with the imidocarbonate (not shown) [65].

1.3.2 Thiols

As noted previously, thiols are good nucleophiles, even at neutral pH, and have the added benefit of scarcity in proteins and other biomolecules. Common reagents for modifying thiol groups are maleimides, alkyl halides and iodoacetamides, and activated disulfides [2–5]. In some instances, thiol modification is done after chemical reduction of disulfide bridges in a target protein to generate the reactive thiol group. In many cases, it is a benefit that thiols are scarce in proteins and frequently tied up in disulfide bonds, as thiols react with many of the same functional groups as amines. For example, thiols will react with active esters to form unstable thioesters [26] and will also react with isothiocyanates to form unstable dithiocarbamates [61].

1.3.2.1 Maleimides

As shown in Figure 1.15, maleimides will undergo a Michael addition reaction with thiols to form a stable thioether linkage [2–4]. This reaction is very selective for thiols between pH 6.5 and 7.5; however, aza-Michael additions with amine nucleophiles can occur under alkaline conditions (pH > 8.5) [2, 7, 26]. Tyrosines, histidines, methionines, and other amino acids do not appreciably react with maleimides [11]. Although the hydrolysis of maleimides to unreactive maleamic acid competes with the bioconjugation reaction, the rate of hydrolysis is slower than with NHS esters [11]. Hydrolysis rates increase with increasing pH; however, these conditions also reduce the chemoselectivity of the maleimide and are rarely used in practice. Hydrolysis may also occur after formation of the thioether bond, forming a succinamic acid isomer, which alters the structure of the conjugate but does not break it apart. In some applications, potential heterogeneity in the final conjugate is a concern, and deliberate steps are taken to completely hydrolyze the conjugates to succinamic acid derivatives and ensure homogeneity [66].

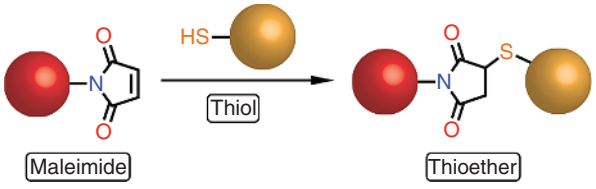

Figure 1.15 Michael addition between a maleimide and a thiol to form a stable thioether linkage.

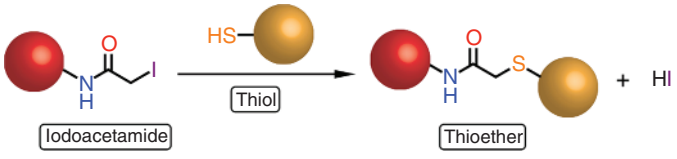

Figure 1.16 Reaction between an iodoacetamide and a thiol to yield a stable thioether.

1.3.2.2 Alkyl Halides and Haloacetamides

Alkyl halides and haloacetamides will readily react with thiols to form a stable thioether bond [2–5]. The most common of these reagents are iodo derivatives, iodoacetamides in particular. The conjugation reaction, shown in Figure 1.16, will proceed at neutral and slightly acidic pH, where many aliphatic amines are protonated and less reactive. Iodoacetamides react most favorably with thiols, even at slightly alkaline pH, and thus selective modification of cysteine residues is possible when the iodoacetamide is used as the limiting reagent. Nonetheless, excess reagent can lead to side reactions with lysine and histidine, and methionine will react at most pH values [2, 3, 5, 26]. Iodo derivatives are also sensitive to light, and reactions with these reagents must be kept in the dark to avoid formation of iodine, which can react with biomolecules (e.g., tyrosine residues in proteins) [2, 11]. The advantage of iodo derivatives is that they react twice as fast as bromo derivatives and more than an order of magnitude faster than chloro derivatives [11]. The trade-off is that chloro derivatives have been reported to be more selective for thiols [67].

1.3.2.3 Activated Disulfides

Thiol–disulfide exchange reactions can occur over a broad range of pH between an activated disulfide and a thiol [2–4, 11, 26]. This reaction is very selective for thiols and is not subject to competing hydrolysis. The most common reagents for thiol–disulfide exchange reactions are pyridyl disulfide derivatives, which form pyridine-2-thione as a by-product of the reaction. The pyridine-2-thione has a UV–visible absorption signature ($\varepsilon_{343\,nm} = 8080\,M^{-1}\,cm^{-1}$) that can be used to track the reaction progress [2]. 2-Nitrobenzoic acid disulfide derivatives can be used similarly, forming 2-nitro-5-thiobenzoic acid as a by-product, which also has a UV–visible absorption signature ($\varepsilon_{412\,nm} = 14\,140\,M^{-1}\,cm^{-1}$) [2, 68]. In both cases, resonance stabilization of the products prevents any appreciable back-reaction. Thiols can be converted into activated pyridyl disulfides and 2-nitrobenzoic acid disulfides using 2,2′-dipyridyl disulfide and 5,5′-dithiobis-[2-nitrobenzoic acid] (DTNB) (Ellman's reagent [69]), respectively [2, 11]. The former activation reaction is illustrated in Figure 1.17a, and

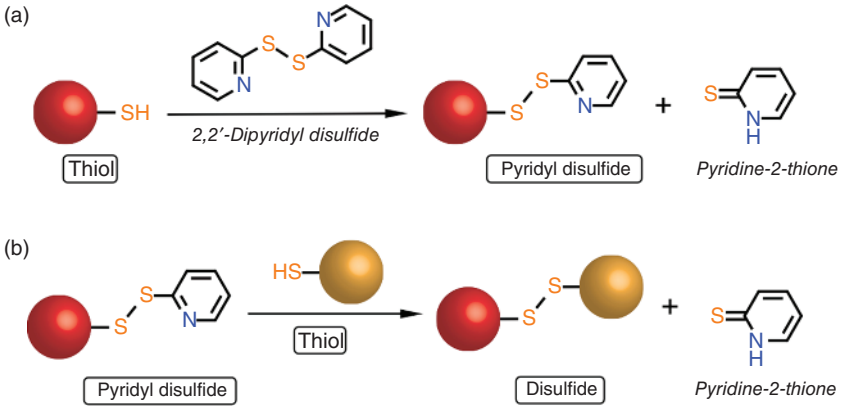

Figure 1.17 (a) Activation of a thiol group with 2,2′-dipyridyl disulfide. (b) Thiol–disulfide exchange reaction between a pyridyl disulfide derivative and a thiol to form a new disulfide linkage and pyridine-2-thione as a by-product.

subsequent reaction with a thiol to form a disulfide is shown in Figure 1.17b. The principal drawback of this bioconjugation chemistry is that the resulting disulfide-linked conjugates are sensitive to reduction, including reduction by intracellular glutathione, precluding their use in some applications. In other applications, this cleavability can be an advantage.

1.3.3 Hydroxyls and Phenols

There are few traditional bioconjugate reactions that utilize hydroxyl and phenol nucleophiles. Both amines and thiols are better nucleophiles and will react under the same conditions as hydroxyls and phenols. Ester-forming reactions of succinimidyl esters with tyrosine, serine, and threonine have been reported to occur as side reactions following modification of all lysine residues with high concentrations of bifunctional succinimidyl ester reagent [70, 71]. Such side reactions with hydroxyls and phenols can occur more efficiently in nonaqueous solvent (e.g., dimethyl sulfoxide, dimethylformamide). In practice, competing nucleophiles and competing hydrolysis frequently limit the utility of hydroxyl and phenol groups as nucleophiles for bioconjugation.

A chemistry that has proven effective for more selective bioconjugation with hydroxyl groups is the reaction of boronic acids with *cis*-1,2-diols to form cyclic boronic esters [72, 73], as shown in Figure 1.18. The reaction is reversible, with the boronic ester favored at basic pH and hydrolyzed at acidic pH. Both the boronic acid and boronate ester have an ionization equilibrium ($pK_a \sim 8–10$) in water [72]. In addition to *cis*-1,2-diols, similar reactions also occur between boronic acids and 1,3-diols, 1,3,5-triols, and 1,3-hydroxyacids, among other functional group combinations. These functional groups are common in carbohydrates. Proteins, other biomolecules, and gel or solid supports can be modified with aminophenylboronic acid to facilitate conjugation with carbohydrates and carbohydrate-containing biomolecules [2]. Boronic acids have also been shown

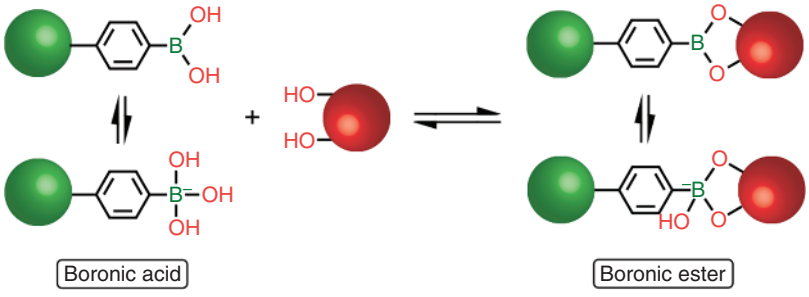

Figure 1.18 Reversible formation of a cyclic boronic ester from the reaction between a boronic acid and a *cis*-1,2-diol. The ionization equilibrium for each species is shown. The boronic ester is favored at basic pH.

to react with salicylhydroxamic acid, forming a more stable complex than that with diols [74].

1.4 Cross-Linking Strategies

Perhaps the two most common and general aims of bioconjugate reactions are (i) to attach a small reporter or drug molecule to a biomolecule, or (ii) to ligate a biomolecule with another biomolecule, a synthetic macromolecule, or an interface such as a bulk solid support, microparticle, or nanoparticle. This section discusses general strategies for using the reactions summarized in Section 1.3 for these purposes.

1.4.1 Zero-Length Cross-Linking or Traceless Ligations

Zero-length cross-linking refers to the direct formation of new covalent bonds between two biomolecules through an activating agent or reactive group that is not incorporated into the final conjugate, leaving no residual atoms. Recently, "traceless ligation" has emerged as an alternative terminology to "zero-length cross-linking" but represents the same fundamental concept. The most common example of zero-length cross-linking is amide coupling through carbodiimide activation of carboxyl groups or via a succinimidyl ester. Another example is a thiol–disulfide exchange reaction with pyridyl disulfide reagents. In contrast, the reactions of thiols with a maleimide or an iodoacetamide are not zero-length cross-linking because succinimide (or succinamic, if hydrolyzed) and acetamide structures, respectively, are part of the final conjugate. Advantages of zero-length cross-linking include minimal (if any) nonnative structure in the final bioconjugate and minimization of the final conjugate size. For zero-length cross-linking to be effective, the reactive functional groups must be mutually accessible. Functional groups that are buried within biomolecular structures will not be able to react if steric hindrance prevents the approach of the cognate functional group. Although there are many activating agents that are potentially capable of zero-length cross-linking, only a small subset of these agents are suitably mild and stable for bioconjugate reactions.

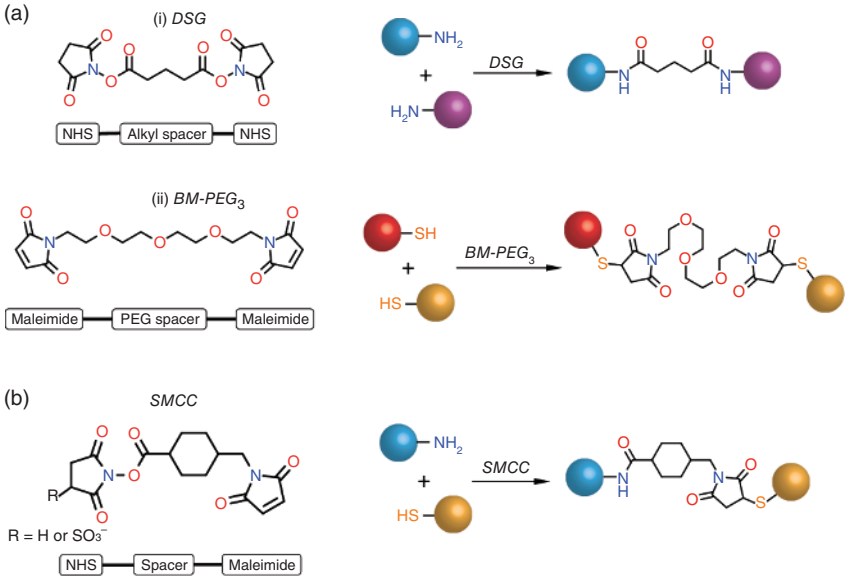

Figure 1.19 (a) Representative examples of homobifunctional cross-linker structures and reactions: (i) disuccinimidyl glutarate (DSG) and (ii) bismaleimidotriethyleneglycol (BM-PEG_3). (b) Structure of succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (SMCC), a representative example of a heterobifunctional cross-linker, and its cross-linking reaction.

1.4.2 Homobifunctional and Heterobifunctional Linkers

Many bioconjugate methods rely on cross-linkers that have reactive functional groups at opposite ends of an alkyl or polyethylene glycol (PEG) spacer [2–4, 11, 75]. These reagents are either homobifunctional, in that the functional groups have the same reactivity, or heterobifunctional, in that the functional groups have different reactivity. These reagents tend to be more frequently used for linking biomolecules together or for attachment to supports and surfaces than for labeling biomolecules with small-molecule reporters. The presence of a linker moiety of tailorable length can often mitigate the effects of biomolecular sterics [75].

Common examples of homobifunctional cross-linkers are bis-NHS esters, bismaleimides, and glutaraldehyde. Figure 1.19a shows the structures of disuccinimidyl glutarate (DSG), an example of a bis-NHS ester cross-linker with an alkyl spacer, and of 1,11-bismaleimidotriethylenegylcol (BM-PEG_3), an example of a bismaleimide cross-linker with a PEG spacer. Generic cross-linking reactions are also illustrated. The reactions of bis-NHS and bismaleimide reagents are analogous to their monofunctional analogs, cross-linking amine and thiol groups to form amide and thioether linkages, respectively. Ostensibly, glutaraldehyde forms imines upon reaction with primary amines at each of its two aldehyde termini; however, it is recognized that cross-linking reactions with glutaraldehyde are much more complex in reality [76]. There is evidence indicating that glutaraldehyde exists in aqueous solution as a monomeric dialdehyde, a cyclic hemiacetal, and various oligomers and polymers. Glutaraldehyde can react with

DNA nucleotides and protein amine groups, with some studies also suggesting potential reactions with the gamut of nucleophilic amino acid side chains [76]. Overall, the reactivity of the various forms of aqueous glutaraldehyde and their mechanism(s) and products of cross-linking remain poorly understood. In this sense, glutaraldehyde epitomizes both the benefit and liability of many traditional bioconjugate chemistries: effective cross-linking and numerous applications, but poor control over the reaction.

Many traditional heterobifunctional cross-linkers combine amine and thiol reactivity, often in the form of (sulfo)succinimidyl ester and maleimide groups separated by spacers of different lengths. In some cross-linkers, a pyridyl disulfide function substitutes the maleimide. One of the most common heterobifunctional cross-linkers is succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) and its sulfonated analog, the structures of which are shown in Figure 1.19b. When used in aqueous media in two-step conjugations, maleimide–succinimidyl ester cross-linkers of this type are generally reacted with the amine reagent first because of the greater susceptibility of the succinimidyl ester to hydrolysis and its potential reaction with thiols. In the case of SMCC, the cyclohexane group is a steric barrier to hydrolysis of the maleimide, extending its longevity to enable efficient two-step conjugations [2]. Other combinations of reactive functional groups are also utilized and are available commercially. Heterobifunctional cross-linkers can also be important in enabling many of the chemoselective and bioorthogonal chemistries in this book. For this purpose, the cross-linking reagents will often have either a succinimidyl ester group or a maleimide group to react with the native functional groups of a biomolecule of interest, paired with a second functional group that is required for the chemoselective or bioorthogonal reaction (e.g., azide, alkyne). The practical use of all heterobifunctional cross-linkers is guided by considerations analogous to those for SMCC, including chemoselectivity and relative rates of hydrolysis, optimum reaction conditions, pretreatment steps (e.g., reduction of disulfides), and options for purification at each step.

Another consideration in the selection of a cross-linker is the stability of the final linkage. In most applications, long-term stability over a broad range of conditions is desirable; however, there are applications where reversible conjugation is important. To this end, cross-linkers can incorporate a cleavable functionality within their spacer [2, 77, 78]. These functionalities are frequently disulfides, diols, or esters that can be cleaved by reduction, oxidation, or a strong nucleophile such as hydroxylamine. Photocleavable groups such as *o*-nitrobenzyl derivatives can also be incorporated into spacers [77], whereas other cross-linking reagents use photoreactivity to initiate the cross-linking reaction [61, 75]. The advantage of these strategies is that an exogenous physical or chemical stimulus can cleave or initiate the cross-link. Other reversible cross-linking strategies take advantage of the pH lability of certain bonds (e.g., hydrazone bonds) and competitive binding reactions (e.g., boronic acids), or include a structure that is recognized as a substrate by hydrolytic enzymes [78–80]. These strategies tend to have much wider compatibility with biological systems than chemical or photochemical cleavage.

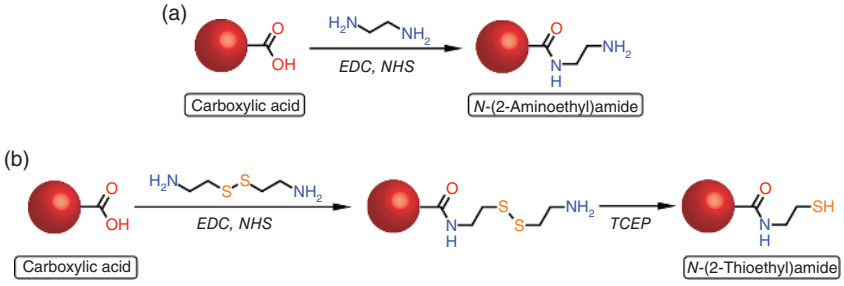

Figure 1.20 Conversion of a carboxylic acid group to (a) an amine group and (b) a thiol group after reduction.

1.4.3 Functional Group Conversion

The conversion of one functional group to another is sometimes an important step in a bioconjugate reaction. Limited functional group diversity in biomolecules, the availability of reagents, the efficiency or chemoselectivity of one reaction over another under a given set of conditions, and subsequent reaction steps can make it advantageous to introduce a new functional group to a biomolecule in preference to a preexisting one.

Although amines are abundant in proteins, other classes of biomolecule and many other useful materials (e.g., polymers, supports for immobilization) do not always have intrinsic amine groups available for reaction. Ethylenediamine and other bisamines (e.g., hexamethylenediamine) are a common route for converting electrophilic functional groups into amines [2]. As shown in Figure 1.20a, activated carboxylic acids will react with ethylenediamine to yield *N*-(2-aminoethyl)amide derivatives, and aldehydes or ketones will react to yield *N*-(2-aminoethyl)amine derivatives after reduction of the corresponding imine. Potential challenges with this chemistry are undesired intramolecular or intermolecular cross-linking from reaction of both amine groups on the same molecule of ethylenediamine. The short length of the ethylenediamine molecule and use of a sufficiently large excess of the reagent help to minimize these side reactions. Similar chemistry can be used to convert carbonyl and carboxyl groups into thiols using cystamine [2], as shown in Figure 1.20b. Reaction of the amine groups with activated carboxylic acids forms amide bonds without interference from the disulfide. Reduction of the disulfide yields a *N*-(2-thioethyl)amide derivative. Undesired cross-linking can be less of a concern with this chemistry because the cross-links will be broken during the subsequent reduction step.

The conversion of amines into carboxylic acids can be accomplished using succinic anhydride, as shown in Figure 1.21. This reagent can also react with thiols, imidazoles (e.g., histidine), phenols (e.g., tyrosine), and alcohols (e.g., serine);

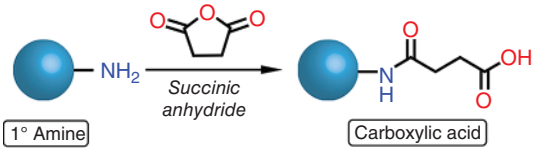

Figure 1.21 Conversion of an amine group to a carboxylic acid group using succinic anhydride.

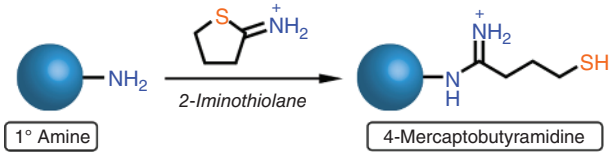

Figure 1.22 Conversion of an amine group to a thiol group using 2-iminothiolane.

however, in contrast to the amide product formed upon reaction with amines, the products of these other reactions are unstable. Hydroxylamine can be used to deliberately cleave esters and regenerate the original functional groups [2, 47]. Homobifunctional succinimidyl ester cross-linkers can also be used to convert an amine to a carboxylic acid, where one succinimidyl ester group reacts to form an amide bond and the other is allowed to hydrolyze to generate the carboxyl group. A shorter linker length and a large excess of reagent favor functional group conversion over undesired cross-linking.

Traut's reagent, 2-iminothiolane, will react with amines to yield a thiol product, as shown in Figure 1.22 [2, 4, 81]. However, it has been reported that the desired 4-mercaptobutyramidine adduct can convert to a more favored non-thiol product and should be reacted with a maleimide or capped with an exchangeable disulfide either *in situ* or as soon as possible after the initial reaction [82]. In principle, pyridyl disulfide derivatives of cysteamine and thiocarboxylic acids can be used to convert thiols into amines and carboxylic acids, respectively. Iodoacetic acid can also accomplish the latter. Such conversions are rarely necessary because of the typical scarcity of the thiol group in biomolecules and the generally good chemoselectivity of maleimide and disulfide exchange reactions.

1.4.4 Biotin–Avidin

The bioconjugate reactions and cross-linking strategies described thus far have focused on the formation of new covalent bonds. An exceedingly popular method of non-covalent conjugation and cross-linking is avidin–biotin binding, which has the strongest non-covalent interaction known in biology [83–86]. Avidin is a tetrameric protein derived from egg white that binds to biotin with a dissociation constant on the order of $K_d \sim 10^{-15}$ M [83–85]. Streptavidin, a protein derived from *Streptomyces avidinii* and shown in Figure 1.23a, is structurally similar to avidin and has comparable affinity for biotin ($K_d \sim 10^{-14}$ M) [83, 84, 86]. Both avidin and streptavidin have four binding sites for biotin (one per monomer). To a first approximation, pairs of biotin-binding sites are located on opposite faces of the protein. The tight binding to biotin (Figure 1.23b) arises from a combination of multiple hydrogen bonds between the side chains of polar amino acids in the binding pocket and the ureido group of the biotin, and aromatic side chains of amino acids that create a hydrophobic box that interacts with the biotin [87]. The binding reaction is specific, fast, and stable, suffers from no competing hydrolysis, and is tolerant of modifications of the valeryl side chain of biotin as shown in Figure 1.23c. The (strept)avidin–biotin complex is remarkable not just in its affinity but also in its physical robustness and tolerance of harsh conditions such as high temperature and denaturants [2]. The key functional difference between

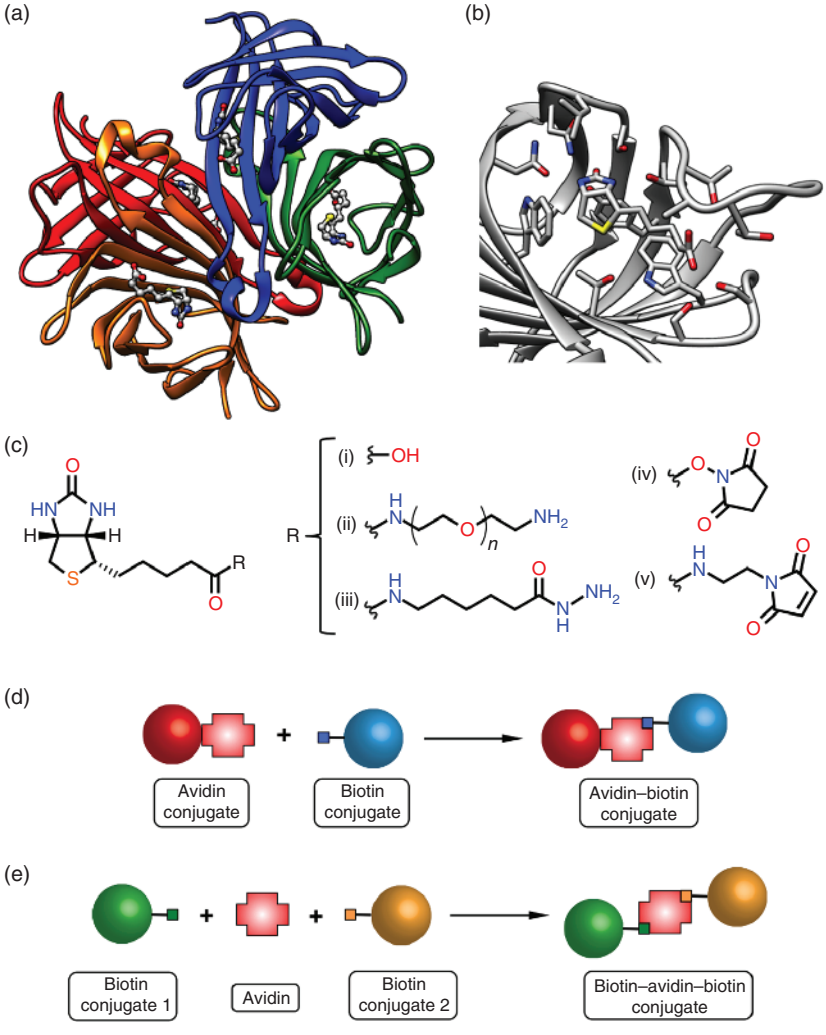

Figure 1.23 (a) Tetrameric ribbon structure of streptavidin bound to four biotin ligands. (b) Close-up view of the biotin-binding pocket of avidin. (c) Structure of (i) biotin and its derivatives: (ii) amine, (iii) hydrazide, (iv) succinimidyl ester, and (v) maleimide. (d) Direct conjugation between biotin and (strept)avidin-modified biomolecules. (e) Indirect conjugation between two biotinylated biomolecules using (strept)avidin.

avidin and streptavidin is that the former is glycosylated and the latter is not. Consequently, the isoelectric point (p*I*) of avidin is about 10.5, and that of streptavidin is about 5–6, with the latter less prone to nonspecific interactions with other biomolecules [2, 61]. Deglycosylated forms of avidin, which also exhibit reduced nonspecific binding while retaining similar affinity for biotin, have been prepared and are available commercially [88, 89].

Biotin–avidin chemistry is frequently used for bioconjugation by attaching avidin to one of the two components to be joined and biotinylating the other

component, as shown in Figure 1.23d. In the case of proteins, avidin fusions can be prepared using molecular biology techniques [90]. Alternatively, as shown in Figure 1.23e, the multivalent nature of avidin can be exploited to cross-link two biotinylated components [3], as well as immobilize biotinylated molecules on avidin-coated surfaces [91, 92]. Biotinylation reagents typically rely on the traditional bioconjugate reactions described earlier, for example, succinimidyl ester-biotin, maleimide-biotin, amino-biotin, thiol-biotin, and several other derivatives [91]. Synthetic peptides and oligonucleotides can be biotinylated during solid-phase synthesis, and such modifications are widely available commercially.

The popularity of biotin–avidin conjugation arises from its specific and stable binding and the widespread availability of reagents and protocols; however, the chemistry is not without its limitations. Because (strept)avidin is tetrameric, it can form undesired cross-links and aggregates as homotetrafunctional cross-linker. The chemistry is also bulky because of the need for the protein component, which can, in some cases, still be a site for nonspecific interactions. A (strept)avidin conjugate may also have greater immunogenicity than a conjugate formed through a small-molecule cross-linker [86]. In addition, joining two biotinylated components through a (strept)avidin bridge will often yield a statistical distribution of bioconjugates depending on the relative stoichiometry of the components and (strept)avidin. For the one-pot conjugation of biotinylated components A and B, a mixture of A_nB_m conjugates is possible, where $n = 0–4$ and $0 < m \leq 4–n$. Two-step conjugation with removal of excess or unwanted avidin or component A prior to adding component B can reduce the variety of the conjugates but will not eliminate heterogeneity. Further, the three-dimensional structure and tetrameric nature of avidin are such that the orientation of components A and B relative to one another can be variable, unless the size of components A and B limits binding to one component per each side of the protein.

1.5 Challenges Associated with Traditional Bioconjugate Reactions

Before discussing challenges that are associated with traditional bioconjugate reactions, it is worthwhile to consider the characteristics of an *ideal* bioconjugate reaction. An ideal bioconjugate reaction would proceed rapidly with stoichiometric efficiency at sub-micromolar concentrations of reagents and across physiologically relevant ranges of pH and temperature. Conjugation would occur exclusively at the targeted site(s) with no side reactions, without interfering with other bioconjugate reactions, and without interference from complex biological matrices and media. Reagents and derivatives for the conjugate reaction should also be widely available or easily prepared, stable, and highly soluble in water, as well as nontoxic. There are very few, if any, real chemistries that will satisfy all of these ideal criteria. The traditional bioconjugate chemistries discussed in this chapter often fail to satisfy several of these criteria.

Competing hydrolysis is a major factor limiting the efficiency of many traditional bioconjugate reactions. For example, mixing a hydrolyzable labeling reagent, such as a succinimidyl ester, with a biomolecule at a $N:1$ ratio will often yield conjugates with an average of less than N labels. Hydrolysis rates can also limit the lowest concentrations of reagents that can be used effectively. For many reactions, biomolecule concentrations should be between 1 and 100 μM, with higher concentrations being preferred, because hydrolysis kinetics tend to be pseudo-first order while reaction kinetics tend to be second order [11]. To compensate for competing hydrolysis, large excesses of active agents (e.g., succinimidyl esters) or activating agents (e.g., carbodiimides, cyanogen bromide) are often required. For some reagents, limited aqueous solubility may require the use of a cosolvent (e.g., dimethylformamide, dimethyl sulfoxide, short-chain alcohols, acetonitrile) to achieve sufficiently high concentrations. An excess of either the nucleophile or electrophile may also be used to increase conjugate reaction yields. Hydrolysis rates depend on the temperature and pH of the reaction mixture, and the initial activity of moisture-sensitive reagents can vary with age and storage conditions. As noted at many points already, the difference between the reaction pH and the pK_a of the nucleophile can play a key role in mediating reactivity, but control of pH must balance the tendency of nucleophiles to be more potent at more basic pH with the much faster rates of reagent hydrolysis. A similar compromise exists for temperature control, where lower temperatures decrease both hydrolysis rates and reaction rates, with the latter sometimes improving chemoselectivity by extension [11]. Reaction time is another consideration, although arguably less important since the degree of conjugation often depends more strongly on the ratio of reagents [11]. Nonetheless, longer reaction times and larger excesses of reagents can sometimes promote undesired side reactions. Given these issues, empirical optimization of conditions may be required for each new bioconjugate reaction, and reproducibility may be limited.

The aforementioned issues with hydrolysis are effectively a manifestation of limited chemoselectivity, but are hardly the only example where poor chemoselectivity creates challenges or limits options. Buffer selection can be influenced by the potential reactivity of buffer salts, as is the case when the otherwise ubiquitous phosphate buffer systems are avoided in carbodiimide coupling reactions. Furthermore, when multiple bioconjugate reactions are to be utilized, chemoselectivity can dictate the order of reactions. A common example is the potent nucleophilicity of thiols, which may require that thiol groups be blocked or reacted prior to any amines because the thiols will react with most amine-reactive agents, including succinimidyl esters and isothiocyanates. The products of these reactions are typically prone to hydrolysis but can be isolated, and the reactions consume reagent. For systems that can tolerate it, aqueous solvent systems can be replaced with organic solvent systems to mitigate the effects of competing hydrolysis, but these conditions can exacerbate the effects of poor chemoselectivity. For example, NHS esters are much more likely to react with serine, threonine, and tyrosine residues. Under typical aqueous conditions, these reactions are slow compared to the rate of hydrolysis and the desired reaction with the amine predominates. When aqueous conditions are not sufficient

to undo unwanted acyl modifications, treatment with hydroxylamine may be useful. Hydroxylamine is a very potent nucleophile and particularly useful for reversal of the unwanted acylation reactions by aminolysis, although care must be taken to ensure that the desired modifications will not be cleaved or degraded.

Another frequent challenge for traditional bioconjugate reactions is the limited diversity of reactive functional groups available. Most biomolecules are polymers that comprise a limited set of monomers that are repeated many times in a sequence (e.g., amino acids in proteins, nucleotides in nucleic acids, and saccharides in carbohydrates). Consequently, targeting bioconjugation to specific sites, or targeting specific molecules in a biological system, can be difficult or impossible. An example already noted is the abundance of amine and carboxyl groups in proteins. For a one-pot conjugation reaction between protein A and protein B using carbodiimide/succinimidyl ester chemistry, the reaction not only will form the desired AB conjugate but will also form AA, BB, and higher-order A_nB_m conjugates because of the multitude of amine and carboxyl groups on each protein [11]. Homobifunctional cross-linkers can produce a similar range of products. Multistep reactions that activate protein A, remove or quench excess activating or cross-linking agent, then add protein B for conjugation can avoid the formation of homomeric products but is more time consuming and can potentially exacerbate the challenges associated with hydrolysis. Moreover, in both cases, conjugates are likely to be formed with different relative orientations of the coupled proteins. Reactions of proteins with mono-reactive small molecules are not prone to excessive cross-linking and orientation challenges, but may also produce a distribution of conjugates with modifications at different sites, particularly when low degrees of labeling are desired. Another issue is that the conjugation reactions at the amine or carboxyl groups of protein A, either with another protein or small molecule, need to be carried out in the absence of background protein and other biomolecules that display these functional groups. In general, this restricts these reactions to purified preparations of protein in laboratory vessels and excludes conjugation reactions in cells and tissues, culture media, biological fluids, and whole organisms. Similar challenges are associated with the modification of nucleic acids, lipids, and carbohydrates.

Of the traditional chemistries discussed here, thiol–maleimide coupling and thiol–disulfide exchange reactions come the closest to an ideal bioconjugate reaction. Thiols are scarce in biomolecules and can be engineered at specific sites. These reactions are selective near neutral pH and permit modification of thiol sites without unwanted side reactions. The hydrolysis rates of maleimides are nontrivial but slow compared to other common chemistries, and pyridyl disulfides have negligible hydrolysis, while both reaction rates are favorable. Large excesses of reagents are therefore not required. The reactions proceed under mild conditions, at micromolar concentrations and lower, and can be complete within minutes. Unfortunately, both chemistries still have limitations. Maleimides and their products do hydrolyze as a function of pH and time, and maleimides can also react with histidines and lysines at sufficiently high concentration and sufficiently basic pH. Reduction of disulfides to thiols with TCEP/DTT may be required for bioconjugation, but TCEP/DTT can interfere with the conjugation reactions. As

such, two separate reaction steps with intermediate purification may be necessary. For conjugation through thiol–disulfide exchange, the lability of the disulfide bond under reducing conditions, such as those found in cellular environments, often restricts the use of this chemistry to *in vitro* applications. Maleimide and pyridyl disulfide reagents are also less widely available than succinimidyl ester reagents. Despite these limitations, thiol–maleimide coupling and thiol–disulfide exchange reactions are of great utility and are arguably the benchmark that novel chemoselective bioconjugation methods should strive to surpass.

1.6 Conclusions

The examples of chemistries discussed in this chapter, which largely revolve around reactions with amine and thiol groups, are representative of traditional bioconjugate methods but are far from exhaustive in their scope. The concepts described are more general and can be extended to many reactions and contexts not covered in this chapter. Certainly, an extensive list of potential challenges and limitations could be composed for each of the chemistries in this chapter, both in general terms and for specific applications. Nevertheless, these challenges and limitations should not be regarded as prohibitions. There are many applications where traditional bioconjugate reactions are convenient and feasible and provide satisfactory efficiency and levels of control. Indeed, these chemistries remain stalwarts for this very reason; however, new and more sophisticated applications of bioconjugates require new levels of efficiency and new levels of control. The more ideal characteristics of the current and emerging chemoselective and bioorthogonal chemistries are described in the subsequent chapters of this book. These chemistries overcome many of the challenges and limitations associated with traditional bioconjugate reactions, opening the door to exciting new possibilities and discoveries.

References

1 Meares, C.F. (1993) Introduction to bioconjugate chemistry, *Perspectives in Bioconjugate Chemistry* C. F. Meares, American Chemical Society, Washington, D.C., 1993, 1–8.
2 Hermanson, G.T. (2013) *Bioconjugate Techniques*, 3rd edn, Academic Press, Elsevier, New York.
3 Wong, S.S. and Jameson, D.M. (2014) *Chemistry of Protein and Nucleic Acid Cross-Linking and Conjugation*, 2nd edn, CRC Press, Taylor & Francis Group, Boca Raton.
4 Narain, R. (2014) *Chemistry of Bioconjugates: Synthesis, Characterization, and Biomedical Applications*, John Wiley & Sons, Inc., Hoboken.
5 Lundblad, R.L. (2014) *Chemical Reagents for Protein Modification*, 4th edn, CRC Press, Taylor & Francis Group, Boca Raton.
6 Garrett, R.H. and Grisham, C.M. (1995) *Biochemistry*, Saunders College Publishing, Toronto.

7 Means, G.E. and Feeney, R.E. (1990) Chemical modifications of proteins: history and applications. *Bioconjugate Chemistry*, **1**, 2–12.

8 DeGraaf, A.J., Kooijman, M., Hennik, W.E., and Mastrobattista, E. (2009) Nonnatural amino acids for site-specific protein conjugation. *Bioconjugate Chemistry*, **20**, 1281–1295.

9 Kalia, J. and Raines, R.T. (2010) Advances in bioconjugation. *Current Organic Chemistry*, **14**, 138–147.

10 Zatsepin, T.S., Stetsenko, D.A., Gait, M.J., and Oretskaya, T.S. (2005) Use of carbonyl group addition-elimination reactions for synthesis of nucleic acid conjugates. *Bioconjugate Chemistry*, **16**, 471–489.

11 Brinkley, M. (1992) A brief survey of methods for preparing protein conjugates with dyes, haptens, and cross-linking reagents. *Bioconjugate Chemistry*, **3**, 2–13.

12 Cleland, W.W. (1964) Dithiothreitol, a new protective reagent for SH groups. *Biochemistry*, **3**, 480–482.

13 Burns, J.A., Butler, J.C., Moran, J., and Whitesides, G.M. (1991) Selective reduction of disulfides by tris(2-carboxyethyl)phosphine. *Journal of Organic Chemistry*, **56**, 2648–2650.

14 Getz, E.B., Xiao, M., Chakrabarty, T., Cooke, R., and Selvin, P.R. (1999) A comparison between the sulfhydryl reductants tris(2-carboxyethyl)phosphine and dithiothreitol for use in protein biochemistry. *Analytical Biochemistry*, **273**, 73–80.

15 Shafer, D.E., Inman, J.K., and Lees, A. (2000) Reaction of tris(2-carboxyethyl) phosphine (TCEP) with maleimide and α-haloacyl groups: anomalous elution of TCEP by gel filtration. *Analytical Biochemistry*, **282**, 161–164.

16 Carter, P. (1986) Site-directed mutagenesis. *Biochemical Journal*, **237**, 1–7.

17 Stromgaard, A., Jensen, A.A., and Stromgaard, K. (2004) Site-specific incorporation of unnatural amino acids into proteins. *ChemBioChem*, **5**, 909–916.

18 Mendel, D., Cornish, V.W., and Schultz, P.G. (1995) Site-directed mutagenesis with an expanded genetic code. *Annual Review of Biophysics and Biomolecular Structure*, **24**, 435–462.

19 Voloshchuk, N. and Montclare, J.K. (2010) Incorporation of unnatural amino acids for synthetic biology. *Molecular Biosystems*, **6**, 65–80.

20 Young, Y.S. and Schultz, P.G. (2010) Beyond the canonical 20 amino acids: expanding the genetic lexicon. *Journal of Biological Chemistry*, **285**, 11039–11044.

21 Lang, K. and Chin, J.W. (2014) Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chemical Reviews*, **114**, 4764–4806.

22 Gildersleeve, J.C., Oyelaran, O., Simpson, J.T., and Allred, B. (2008) Improved procedure for direct coupling of carbohydrates to proteins via reductive amination. *Bioconjugate Chemistry*, **19**, 1485–1490.

23 Ruhaak, L.R., Zauner, G., Huhn, C., Bruggink, C., Deelder, A.M., and Wuhrer, M. (2010) Glycan labeling strategies and their use in identification and quantification. *Analytical and Bioanalytical Chemistry*, **397**, 3457–3481.

24 Buskas, T., Ingale, S., and Boons, G.J. (2006) Glycopeptides as versatile tools for glycobiology. *Glycobiology*, **16**, 113R–136R.

25 Likhosherstov, L.M., Novikova, O.S., Dervitskaya, V.A., and Kochetkov, N.K. (1986) A new simple synthesis of amino sugar beta-D-glycosylamines. *Carbohydrate Research*, **146**, C1–C5.

26 Wilbur, D.S. (1992) Radiohalogenation of proteins: an overview of radionuclides, labeling methods, and reagent for conjugate labeling. *Bioconjugate Chemistry*, **3**, 433–470.

27 Clamp, J.R. and Hough, L. (1965) The periodate oxidation of amino acids with reference to studies on glycoproteins. *Biochemical Journal*, **94**, 17–24.

28 Kohn, J. and Wilcheck, M. (1984) The use of cyanogen-bromide and other novel cyanylating agents for the activation of polysaccharide resins. *Applied Biochemistry and Biotechnology*, **9**, 285–305.

29 Batista-Viera, F., Janson, J.C., Carlsson, J. (2011) Affinity chromatography, *Protein Purification: Principles, High Resolution Methods, and Applications*, **54**, 3rd ed. (J. C. Janson), John Wiley & Sons, Inc., Hoboken, **2011**, 221–258.

30 Kool, E.T. (2001) Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annual Review of Biophysics and Biomolecular Structure*, **30**, 1–22.

31 Chu, B.C., Wahl, G.M., and Orgel, L.E. (1983) Derivatization of unprotected polynucleotides. *Nucleic Acids Research*, **11**, 6513–6529.

32 Zamecnik, P.C., Stephenson, M.L., and Scott, J.F. (1960) Partial purification of soluble RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **46**, 811–822.

33 Draper, D.E. and Gold, L. (1980) A method for linking fluorescent labels to polynucleotides: application to studies of ribosome–ribonucleic acid interactions. *Biochemistry*, **19**, 1774–1781.

34 Draper, D.E. (1984) Attachment of reporter groups to specific, selected cytidine residues in RNA using a bisulfite-catalyzed transamination reaction. *Nucleic Acids Research*, **12**, 989–1002.

35 Temsamani, J. and Agrawal, S. (1996) Enzymatic labeling of nucleic acids. *Molecular Biotechnology*, **5**, 223–232.

36 Cobianchi, F. and Wilson, S.H. (1987) Enzymes for modifying and labeling DNA and RNA. *Methods in Enzymology*, **152**, 94–110.

37 Verma, S. and Eckstein, F. (1998) Modified oligonucleotides: synthesis and strategy for users. *Annual Review of Biochemistry*, **67**, 99–134.

38 Reese, C.B. (2005) Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Organic and Biomolecular Chemistry*, **3**, 3851–3868.

39 Caruthers, M.H. (1991) Chemical synthesis of DNA and DNA analogues. *Accounts of Chemical Research*, **24**, 278–284.

40 Singh, Y., Murat, P., and Defrancq, E. (2010) Recent developments in oligonucleotide conjugation. *Chemical Society Reviews*, **39**, 2054–2070.

41 Beaucage, S.L. and Iyer, R.P. (1993) The functionalization of oligonucleotides via phosphoramidite derivatives. *Tetrahedron*, **49**, 1925–1963.

42 Weisbrod, S.H. and Marx, A. (2008) Novel strategies for the site-specific covalent labelling of nucleic acids. *Chemical Communications*, 5675–5685.

43 Wetmur, J.G. (1991) DNA probes: applications of the principles of nucleic acid hybridization. *Critical Reviews in Biochemistry and Molecular Biology*, **26**, 227–259.

44 Kanne, D., Straub, K., Rapoport, H., and Hearst, J.E. (1982) The psoralen-DNA photoreaction. Characterization of the monoaddition products from 8-methoxypsoralen and 4,5′,8-trimethylpsoralen. *Biochemistry*, **21**, 861–871.

45 Saffran, W.A., Welsh, J.T., Knobler, R.M., Gasparro, F.P., Cantor, C.R., and Edelson, R.L. (1988) Preparation and characterization of biotinylated psoralen. *Nucleic Acids Research*, **16**, 7221–7231.

46 Torchilin, V.P., Weissig, V., Martin, F.J., Heath, T.D., New, R.R.C. (2003) Surface modification of liposomes, *Liposomes: A Practical Approach*, 2nd ed. (V. P. Torchilin, V. Weissig) Oxford University Press, Oxford, 2003.

47 Niemeyer, C.M. (ed.) (2004) *Bioconjugation Protocols*, Humana Press, New Jersey.

48 Raddatz, S., Mueller-Ibeler, J., Kluge, J., Wäß, L., Burdinski, G., Havens, J.R., Onofrey, T.J., Wang, D., and Schweitzer, M. (2002) Hydrazide oligonucleotides: new chemical modification for chip array attachment and conjugation. *Nucleic Acids Research*, **30**, 4793–4802.

49 Jentoft, N. and Dearborn, D.G. (1979) Labeling of proteins by reductive methylation using sodium cyanoborohydride. *Journal of Biological Chemistry*, **254**, 4359–4365.

50 Baslé, E., Joubert, N., and Pucheault, M. (2010) Protein chemical modification on endogenous amino acids. *Chemical Biology*, **17**, 213–227.

51 Kalia, J. and Raines, R.T. (2008) Hydrolytic stability of hydrazones and oximes. *Angewandte Chemie, International Edition*, **47**, 7523–7526.

52 Ridley, B.L., Spiro, M.D., Glushka, J., Albersheim, P., Darvill, A., and Mohnen, D. (1997) A method for biotin labeling of biologically active oligogalacturonides using a chemically stable hydrazide linkage. *Analytical Biochemistry*, **249**, 10–19.

53 Nakajima, N. and Ikada, Y. (1995) Mechanism of amide formation by carbodiimide for bioconjugation in aqueous media. *Bioconjugate Chemistry*, **6**, 123–130.

54 Yang, B., Yang, B.L., and Goetinck, P.F. (1995) Biotinylated hyaluronic acid as a probe for identifying hyaluronic acid-binding proteins. *Analytical Biochemistry*, **228**, 299–306.

55 Gilles, M.A., Hudson, A.Q., and Borders, C.L. Jr., (1990) Stability of water-soluble carbodiimides in aqueous solution. *Analytical Biochemistry*, **184**, 244–248.

56 Madison, S.A. and Carnali, J.O. (2013) pH optimization of amidation via carbodiimides. *Industrial and Engineering Chemistry Research*, **52**, 13547–13555.

57 Sehgal, D. and Vijay, I.K. (1994) A method for the high efficiency of water-soluble carbodiimide-mediated amidation. *Analytical Biochemistry*, **218**, 87–91.

58 Staros, J.V., Wright, R.W., and Swingle, D.M. (1986) Enhancement by *N*-hydroxysulfosuccinimide of water-soluble carbodiimide-mediated coupling reactions. *Analytical Biochemistry*, **156**, 220–222.

59 Nojima, Y., Iguchi, K., Suzuki, Y., and Sato, A. (2009) The pH-dependent formation of PEGylated bovine lactoferrin by branched polyethylene glycol (PEG)–*N*-hydroxysuccinimide (NHS) active esters. *Biological and Pharmaceutical Bulletin*, **32**, 523–526.

60 Anjaneyulu, P.S. and Staros, J.V. (1987) Reactions of *N*-hydroxysulfosuccinimide active esters. *International Journal of Peptide and Protein Research*, **30**, 117–124.

61 ThermoFisher Scientific (2010) *The Molecular Probes Handbook: A Guide to Fluorescent Probes and Labeling Technologies*, 11th edn.

62 Barbosa, O., Torres, R., Ortiz, C., Berenguer-Murcia, A., Rodrigues, R.C., and Fernandez-Lafuente, R. (2013) Heterofunctional supports in enzyme immobilization: from traditional immobilization protocols to opportunities in tuning enzyme properties. *Biomacromolecules*, **14**, 2433–2462.

63 Wheatley, J.B. Jr., and Schmidt, D.E. (1999) Salt-induced immobilization of affinity ligands onto epoxide-activated supports. *Journal of Chromatography A*, **849**, 1–12.

64 Hunter, M.J. and Ludwig, M.L. (1962) The reaction of imidoesters with proteins and related small molecules. *Journal of the American Chemical Society*, **84**, 3491–3504.

65 Jennings, H.J., Sood, R.K. (1994) Synthetic glycoconjugates as human vaccines, *Neoglycoconjugates* (Y. C. Lee, R. T. Lee) Academic Press, Inc., *San Diego*, **1994**, 325–372.

66 Kalia, J. and Raines, R.T. (2007) Catalysis of imido group hydrolysis in a maleimide conjugate. *Bioorganic and Medicinal Chemistry Letters*, **17**, 6286–6289.

67 Nielsen, M.L., Vermeulen, M., Bonaldi, T., Cox, J., Moroder, L., and Mann, M. (2008) Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nature Methods*, **5**, 459–460.

68 Collier, H.B. (1973) A note on the molar absorptivity of reduced Ellman's reagent, 3-carboxylato-4-nitrothiophenolate. *Analytical Biochemistry*, **56**, 310–311.

69 Ellman, G.L. (1959) Tissue sulfhydryl groups. *Archives of Biochemistry and Biophysics*, **82**, 70–77.

70 Mädler, S., Bich, C., Touboul, D., and Zenobi, R. (2009) Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *Journal of Mass Spectrometry*, **44**, 694–706.

71 Kalkhof, S. and Sinz, A. (2008) Chances and pitfalls of chemical cross-linking with amine-reactive *N*-hydroxysuccinimide esters. *Analytical and Bioanalytical Chemistry*, **392**, 305–312.

72 Hall, D.G. (2005) Structure, properties, and preparation of boronic acid derivatives. Overview of their reactions and applications. *Boronic Acids: Preparation and Applications in Organic Synthesis and Medicine*, **1** (D. G. Hall), Wiley-VCH Verlag GmbH & Co. KGaA Weinheim, **2005**, 1–99.

73 Mader, H.S. and Wolfbeis, O.S. (2008) Boronic acid based probes for microdetermination of saccharides and glycosylated biomolecules. *Microchimica Acta*, **162**, 1–34.

74 Stolowitz, M.L., Ahlem, C., Hughes, K.A., Kaiser, R.J., Kesicki, E.A., Li, G., Lund, K.P., Torkelson, S.M., and Wiley, J.P. (2001) Phenylboronic acid–salicylhydroxamic acid bioconjugates. 1. A novel boronic acid complex for protein immobilization. *Bioconjugate Chemistry*, **12**, 229–239.

75 Mattson, G., Conklin, E., Desai, S., Nielander, G., Savage, M.D., and Morgensen, S. (1993) A practical approach to crosslinking. *Molecular Biology Reports*, **17**, 167–183.

76 Migneault, I., Dartiguenave, C., Bertrand, M.J., and Waldron, K.C. (2004) Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. *Biotechniques*, **37**, 790–802.

77 Petrotchenko, E.V. and Borchers, C.H. (2010) Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrometry Reviews*, **29**, 862–876.

78 Leriche, G., Chisholm, L., and Wagner, A. (2012) Cleavable linkers in chemical biology. *Bioorganic and Medicinal Chemistry*, **20**, 571–582.

79 Siegel, D. (2012) Applications of reversible covalent chemistry in analytical sample preparation. *Analyst*, **137**, 5457–5482.

80 Christie, R.J., Anderson, D.J., and Grainger, D.W. (2010) Comparison of hydrazone heterobifunctional cross-linking agents for reversible conjugation of thiol containing chemistry. *Bioconjugate Chemistry*, **21**, 1779–1787.

81 Jue, R., Lambert, J.M., Pierce, L.R., and Traut, R.R. (1978) Addition of sulfhydryl groups of *Escherichia coli* ribosomes by protein modification with 2-iminothiolane (methyl 4-mercaptobutyrimidate). *Biochemistry*, **17**, 5399–5406.

82 Singh, R., Kats, L., Blättler, W., and Lambert, J.M. (1996) Formation of N-substituted 2-iminothiolanes when amino groups in proteins and peptides are modified by 2-iminothiolane. *Analytical Biochemistry*, **236**, 114–125.

83 Laitinen, O.H., Nordlund, H.R., Hytönen, V.P., and Kulomaa, M.S. (2007) Brave new (strept)avidins in biotechnology. *Trends in Biotechnology*, **25**, 269–277.

84 Diamandis, E.P. and Christopoulos, T.K. (1991) The biotin-(strept)avidin system: principles and applications in biotechnology. *Clinical Chemistry*, **37**, 625–636.

85 Wilchek, M. and Bayer, E.A. (1988) The avidin–biotin complex in bioanalytical applications. *Analytical Biochemistry*, **171**, 1–32.

86 Dundas, C.M., Demonte, D., and Park, S. (2013) Streptavidin–biotin technology: improvements and innovations in chemical and biological applications. *Applied Microbiology and Biotechnology*, **97**, 9343–9353.

87 Livnah, O., Bayer, E.A., Wilchek, M., and Sussman, J.L. (1993) Three-dimensional structures of avidin and the avidin–biotin complex. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 5076–5080.

88 Hiller, Y., Gershoni, J.M., Bayer, E.A., and Wilchek, M. (1987) Biotin binding to avidin. Oligosaccharide side chain not required for ligand association. *Biochemical Journal*, **248**, 167–171.

89 Bayer, E.A., Demeester, F., Kulik, T., and Wilchek, M. (1995) Preparation of deglycosylated egg-white avidin. *Applied Biochemistry and Biotechnology*, **53**, 1–9.

90 Airenne, K.J., Marjomäki, V.S., and Kulomaa, M.S. (1999) Recombinant avidin and avidin-fusion proteins. *Biomolecular Engineering*, **16**, 87–92.

91 Turková, J. (1999) Oriented immobilization of biologically active proteins as a tool for revealing protein interactions and function. *Journal of Chromatography B: Biomedical Sciences and Applications*, **722**, 11–31.

92 Su, X., Wu, Y.J., Robelek, R., and Knoll, W. (2005) Surface plasmon resonance spectroscopy and quartz crystal microbalance study of streptavidin film structure effects on biotinylated DNA assembly and target DNA hybridization. *Langmuir*, **21**, 348–353.