Chapter 8.1
Occurrence and Significance

Valentin Wittmann

Institut für Organische Chemie, Johann Wolfgang Goethe-Universität, Marie-Curie-Strasse 11, 60439 Frankfurt am Main, Germany
e-mail: wittmann@chemie.uni-frankfurt.de

Keywords: Glycoproteins, Glycopeptides, Glycoforms, Microheterogeneity, *N*-Glycans, *O*-Glycans, Antigenic determinants, GPI anchors, Mucins, Proteoglycans, Peptidoglycan, Glycopeptide antibiotics

1 Introduction 2254

2 Glycoproteins 2254

2.1 Monosaccharide Constituents 2257
2.2 Carbohydrate-Peptide Linkages 2257
2.2.1 *N*-Glycosides 2257
2.2.2 *O*-Glycosides 2259
2.2.3 GPI Anchors 2260
2.2.4 Uncommon Linkages 2263
2.3 Methods of Glycoprotein Analysis 2263
2.3.1 Release of Glycans from Glycoprotein 2264
2.3.2 Labeling and Separation of Glycans 2264
2.3.3 Sequencing of Glycans 2265
2.4 Structure of *N*-Glycans 2265
2.4.1 Diversity of *N*-Glycans 2267
2.4.2 Cell-Type Specific Expression of *N*-Glycans 2269
2.5 Structure of *O*-Glycans 2270
2.5.1 Mucin-Type *O*-Glycans 2271
2.5.2 Non-Mucin *O*-Glycans. 2274
2.6 Proteoglycans. 2276
2.6.1 Carbohydrate-Peptide Linkage Region 2276
2.6.2 Structure of Glycosaminoglycans 2277

3 Peptidoglycan 2278

4 Glycopeptide Antibiotics 2280

4.1 Discovery and Medical Use 2280
4.2 Structure and Classification 2281

References 2284

1 Introduction

This chapter covers natural glycoconjugates in which carbohydrates are covalently linked to amino acids which themselves are part of a peptide or protein. Three classes of biologically important compounds fulfill this criterion:

- glycoproteins including their substructures, the glycopeptides,
- the bacterial cell wall polymer peptidoglycan, and
- glycopeptide antibiotics.

Glycoproteins constitute the most diverse of these classes of glycoconjugates. Their substructures, the glycopeptides, occur as natural and artificial degradation products of glycoproteins during metabolism and structural analysis, respectively, or they are obtained synthetically in order to probe their structure and/or function. Peptidoglycan is commonly not regarded as a glycopeptide. It forms a bag-like macromolecule (called the sacculus) wrapping bacteria and thereby contributing essentially to their mechanical stability. However, it is connected with glycopeptides not only by virtue of its composition but also because its biosynthesis is inhibited by (beside β-lactam antibiotics) glycopeptide antibiotics. The latter are secondary metabolites produced in actinomycetes and are dealt with at the end of this chapter.

2 Glycoproteins

Glycoproteins [1] consist of carbohydrates covalently linked with proteins and are ubiquitous in all forms of life from Archaebacteria to humans although they are rare in eubacteria. Their carbohydrate content is variable from less than 1%, as in some collagens, to over 99%, as in glycogen. The carbohydrate may be in the form of a monosaccharide or disaccharide, but more frequently in the form of oligosaccharides and polysaccharides (up to hundreds of monosaccharides in size), or their derivatives (e.g., sulfo- or phospho-substituted), linear or branched, generally referred to as glycans. The number of glycans present in a glycoprotein differs greatly from a single residue to more than one hundred. The carbohydrate chains are attached to the polypeptide backbone by characteristic carbohydrate-peptide linkages.

Being the most important co- and post-translational modification, glycosylation is found in most proteins including enzymes, antibodies, receptors, hormones, cytokines, and structural proteins. Glycoproteins occur inside cells, both in the cytoplasm and in subcellular organelles, in extracellular fluids as well as embedded in cell membranes. In the latter case the glycans are located extracellularly. Blood serum is an especially rich source of glycoproteins. Of the almost 100 proteins that have been identified in this fluid, almost all are glycosylated. Serum albumin is one of the rare exceptions in this respect although a genetic variant in which it is glycosylated has been discovered recently [2].

A striking feature of almost all glycoproteins is the polymorphism associated with their glycan moieties, a phenomenon known as microheterogeneity. This type of diversity derives from the fact that glycans are secondary gene products and is manifested in that individual molecules of a given glycoprotein carry different oligosaccharides at the same glycosylation site of the protein backbone. The resulting variants, referred to as glycoforms, were first observed in 1962 by Schmid et al. in α_1-acid glycoprotein from human serum by electrophoresis [3]. This glycoprotein contains five glycosylation sites modified with di-, tri-, and tetraantennary glycans of the *N*-acetyllactosamine type [4]. In addition, the glycans can be fucosylated [5] and sialylated at different levels. Microheterogeneity is also observed in proteins with a single glycosylation site like chicken ovalbumin. Close to 20 different oligosaccharides have been identified at that site (Asn-293) [6]. One of the very rare cases in which microheterogeneity is absent is soybean agglutinin (SBA), a plant glycoprotein with a single uniform oligosaccharide per subunit, namely $Man_9(GlcNAc)_2$ [7]. This is mainly why SBA is the best source for a preparative isolation of this oligosaccharide. Up to now, many thousands of primary glycan structures have been characterized. Many of them are contained in the Complex Carbohydrate Structure Database (CCSD) which is searchable by the CarbBank program [8].

The observation of microheterogeneity gives rise to many interesting questions regarding the origin of this phenomenon and its biological relevance. Whereas in the early view the heterogeneity of the glycans was thought to be random and mainly a result of the lack of fidelity in their synthesis, it now appears that, under constant physiological conditions, the populations in a set of glycoforms are reproducible and highly regulated. Moreover, the populations change under certain conditions such as cell growth, cell differentiation, and disease, including malignant transformation, suggesting that the presence of different but defined glycoforms is indeed required for the normal functioning of an organism.

Today we know there is no single unifying function for the carbohydrates present in glycoproteins [9]. Perhaps their major function is to participate in numerous physiological and pathological molecular recognition events. In addition, they may modify the physical, chemical, and biological properties of the proteins to which they are attached. It has been shown that the glycans alter charge and solubility of proteins and influence the conformation and dynamic properties of the polypeptide chain. Due to their large size, the oligosaccharides may cover functionally important areas of the proteins and thus regulate their interactions with other biomolecules or protect the protein from proteolytic degradation. Properties and functions of glycoproteins are covered in Chapt. 8.2 of this book.

Table 1. Monosaccharides found in glycoproteins [1f, 10, 11]

Monosaccharide[a]	Abbreviation	Comment
"Classical"		
Galactose	Gal	
Glucose	Glc	Mainly in collagens
Mannose	Man	
N-Acetylgalactosamine	GalNAc	
N-Acetylglucosamine	GlcNAc	
L-Fucose	Fuc	
L-Arabinofuranose	Ara*f*	In plant glycoproteins
Xylose	Xyl	In proteoglycans and plant glycoproteins
Glucuronic acid	GlcA	In proteoglycans
L-Iduronic acid	IdoA	In proteoglycans
N-Acetylneuraminic acid	Neu5Ac	Mainly in higher vertebrates and invertebrates
Rare		
2-Acetamido-4-amino-2,4,6-trideoxyglucose		*Clostridium symbiosum*
6-Deoxyaltrose		Salmonid fish eggs
3-Deoxy-D-*glycero*-*galacto*-nonulosonic acid	KDN	Salmonid fish eggs
2,3-Diacetamido-2,3-dideoxymannuronic acid		*Bacillus stearothermophilus*
2-*O*-Methyl-L-fucose	Fuc2Me	Nematodes
Galactofuranose	Gal*f*	Bacteria, trypanosoma, fungi
3-*O*-Methylgalactose	Gal3Me	Snail
4-*O*-Methylgalactose	Gal4Me	Nematodes
6-*O*-Methylgalactose	Gal6Me	Algae
Galactose-3-sulfate	Gal3*S*	Thyroglobulin, mucins in cystic fibrosis
N-Acetylgalactosamine-4-sulfate	GalNAc4*S*	Pituitary hormones, Tamm Horsfall glycoprotein, urokinase
3-*O*-Methylglucose	Glc3Me	*Methanothermus fervidus*
3-*O*-Methyl-*N*-acetylglucosamine	GlcNAc3Me	*Clostridium thermocellum*
N-Acetylglucosamine-6-sulfate	GlcNAc6*S*	Thyroglobulin
Gulose	Gul	Algae
3-*O*-Methylmannose	Man3Me	Snail
Mannose-4-sulfate	Man4*S*	Ovalbumin
Mannose-6-sulfate	Man6*S*	Ovalbumin, slime mold
Mannose-6-methylphosphate	Man6*P*Me	Slime mold
N-Acetylmannosamine	ManNAc	*Clostridium symbiosum*
4,8-Anhydro-*N*-acetylneuraminic acid		Edible bird's nest
8-*O*-Methyl-9-*O*-acetyl-*N*-glycolyl-neuraminc acid	Neu5Gc8 Me9Ac	Starfish
8-*O*-Methyl-7,9-di-*O*-acetyl-*N*-glycolyl-neuraminic acid	Neu5Gc7, $9Ac_2$8Me	Starfish
L-Rhamnose	Rha	Eubacteria

[a]Unless otherwise stated, all monosaccharides are of D-configuration and the ring form is pyranose.

2.1 Monosaccharide Constituents

For a long time, less than a dozen monosaccharides were considered to be the main constituents of glycoproteins. Beside these "classical" saccharides, however, refinement of the analytical methods brought up many new monosaccharides during the last 15 years, originally considered as rare but which now appear to be more common than previously thought. Table 1 [1f, 10, 11] gives an overview of the classical and some examples of the rare monosaccharide constituents of glycoproteins.

2.2 Carbohydrate-Peptide Linkages

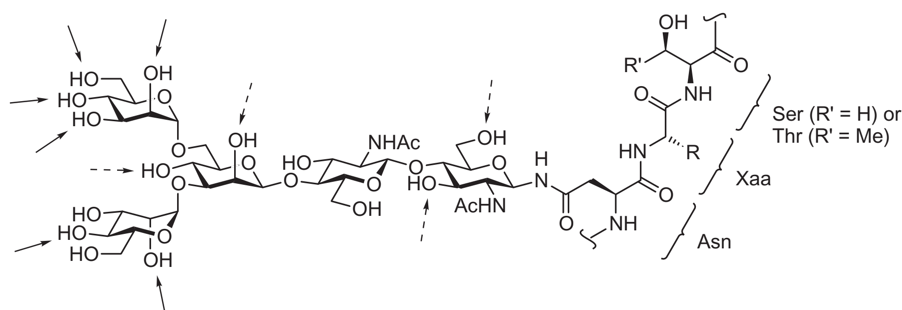
There are three major types of linkages between carbohydrates and proteins [1, 12]:

- *N*-glycosidic, between the reducing end monosaccharide and the side chain amide group of asparagine (*N*-glycans),
- *O*-glycosidic, between the reducing end monosaccharide and the side chain hydroxy group of most commonly serine and threonine but also hydroxyproline and hydroxylysine (*O*-glycans), and
- via ethanolamine phosphate, between the C-terminal amino acid of the protein and an oligosaccharide attached to phosphatidylinositol, generally known as glycosylphosphatidylinositol (GPI) anchor.

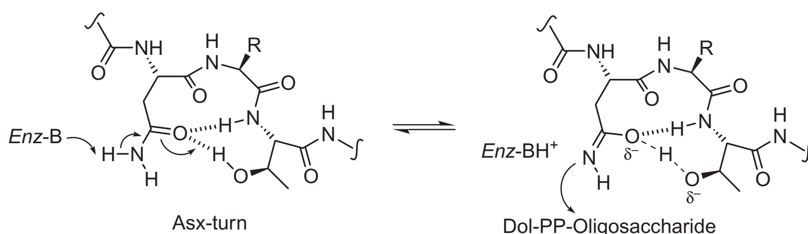
Beside these, a number of uncommon linkages are found in nature including *C*-glycosides and carbohydrates linked via a phosphodiester bridge. The non-enzymatic condensation of reducing monosaccharides like glucose with the side-chain amino group of lysine leads via initial Schiff bases and subsequent rearrangements to the formation of so-called advanced glycation end products (AGEs). AGEs have been implicated in alterations of proteins during aging and long-term diabetes [13] and are not dealt with in this chapter.

2.2.1 *N-Glycosides*

Until recently, the only type of *N*-glycosidic linkage which had been characterized in glycoproteins regardless of their origin was the *N*-acetylglucosaminyl(β1-N)asparagine bond discovered by Neuberger et al. [14] and almost simultaneously by others [15, 16]. During biosynthesis the enzyme oligosaccharyl transferase (OT) transfers a triantennary tetradecasaccharide ($Glc_3Man_9GlcNAc_2$) from dolichol pyrophosphate to the amide nitrogen of an Asn side chain in the nascent polypeptide. In the subsequent trimming process this oligosaccharide is modified by the action of several glycosylhydrolases and glycosyltransferases resulting in a structural diversity of glycans sharing a com-



Scheme 1. Core pentasaccharide Man(α1-6)[Man(α1-3)]Man(β1-4)GlcNAc(β1-4)GlcNAc of *N*-glycoproteins attached to the consensus sequence Asn-Xaa-Ser/Thr. Solid arrows indicate the points of attachment of the outer arm saccharides forming carbohydrate chains, called antennae. In addition, the inner-core may be substituted by several monosaccharides (dashed arrows)



Scheme 2. Proposed mechanism of asparagine activation for oligosaccharyl transferase [19]

mon core pentasaccharide (Scheme 1). The primary peptide sequence requirements for OT-catalyzed glycosylation include a minimum Asn-Xaa-Ser/Thr tripeptide recognition motif where Xaa can be any of the 20 natural amino acids except proline. However, it is found that not all such consensus tripeptides (sequons) are glycosylated. A plausible explanation arises from mechanistic studies on the OT-catalyzed glycosylation of linear and constrained model peptides [17]. It could be demonstrated that only peptides which are able to adopt an Asx-turn conformation [18] are efficient substrates of OT leading to the proposed mechanism of oligosaccharyl transfer shown in Scheme 2 [19]. In this model, the apparent failure of the glycosylation machinery most likely results from conformational influence by the neighboring polypeptide sequence that may override the potential for the tripeptide acceptor sequence to adopt an Asx-turn and therefore limit its compatibility with the OT active site. However, although necessary, formation of an Asx-turn might not be sufficient [20].

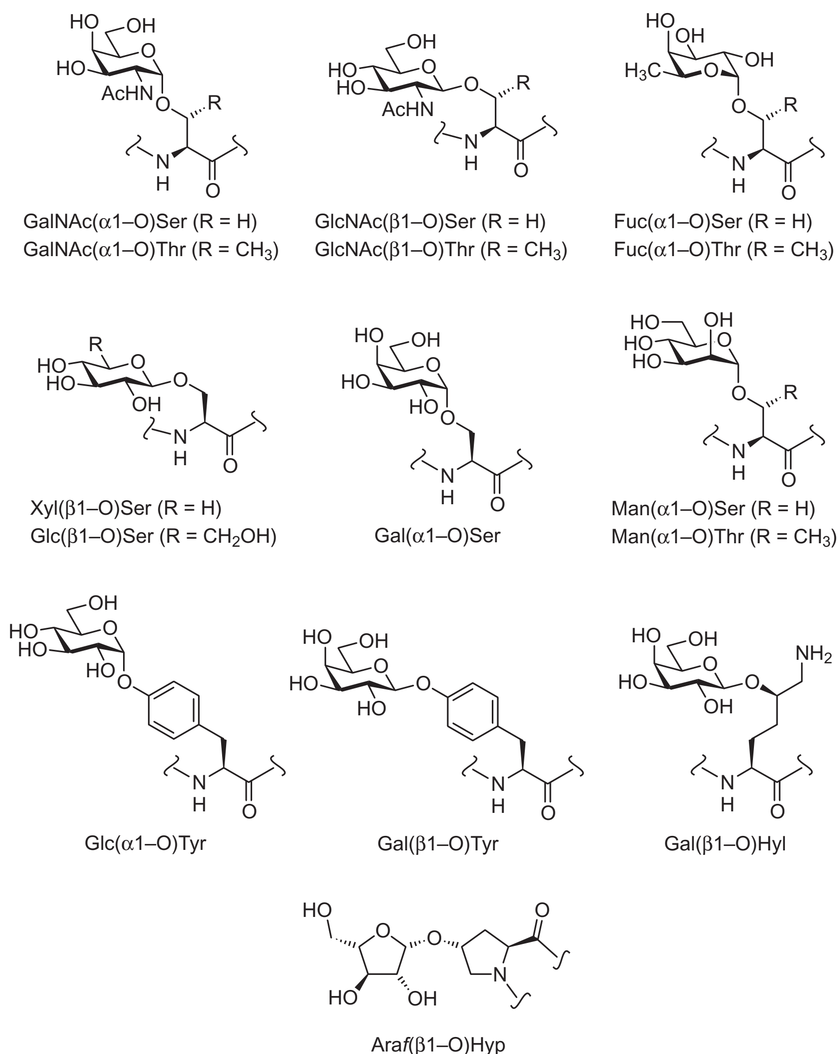
During recent years, linkages between asparagine side chains and other carbohydrates such as α- [21] and β-glucose [22], β-*N*-acetylgalactosamine [23] and L-rhamnose [24] have been discovered mainly in bacterial glycoproteins [25]. However, the β-glucosyl-asparagine linkage has also been found in the mammalian protein laminin [26]. Whenever carefully studied, the saccharides

are attached to the sequon Asn-Xaa-Ser/Thr, except for the glycopeptide nephritogenoside in which glucose is α-linked to the amide side chain of the *N*-terminal tripeptide Asn-Pro-Leu [21].

2.2.2
O-Glycosides

In contrast to *N*-glycoproteins, biosynthesis of the glycan chains of *O*-glycoproteins is an entirely post-translational process and starts with the addition of a single monosaccharide to a hydroxyamino acid of the protein backbone. Therefore, a variety of carbohydrate-protein linkages is found in *O*-glycoproteins (Scheme 3). The *N*-acetylgalactosaminyl(α1-O)serine/threonine linkage (also called Tn antigen) was first demonstrated in mucins and is widely distributed in nature in the so-called mucin-type glycoproteins. Another widely occurring *O*-glycosidic bond which was discovered only recently is that between β-*N*-acetylglucosamine and the serine/threonine hydroxy group [27, 28]. This type of attachment is present predominantly in intracellular glycoproteins (nuclear pore, chromatin proteins, transcription factors, and cytoplasmic inclusions). Interestingly, no other sugars are connected to the *N*-acetylglucosamine residue. In the epidermal growth factor (EGF) modules of several blood coagulation and fibrinolytic proteins L-fucose was characterized to be α-linked to serine or threonine in the consensus sequence Cys-Xaa-Xaa-Gly-Gly-Ser/Thr-Cys [29]. EGF modules also contain glucose β-*O*-linked to serine within the consensus sequence Cys-Xaa-Ser-Xaa-Pro-Cys, an example being the bovine blood clotting factor IX [30]. α-Galactose bound to serine is found in plant glycoproteins like extensin or potato lectin and in the glycoproteins from the cellulosome, an extracellular complex of cellulases, produced by cellulolytic bacteria [31]. When human insulin-like growth factor-I (IGF-I) was expressed in yeast, a new form of IGF-I was characterized in addition to IGF-I in which mannose is α-glycosidically linked to threonine [32]. The mannosyl–Ser/Thr linkage has also been reported in Ser/Thr-rich domains of the bovine peripheral nerve α-dystroglycan [33] and in rat brain proteoglycans [34, 35]. Xylosyl(β1-O)serine is found in animal proteoglycans and, at the present, considered to be confined to these glycoproteins.

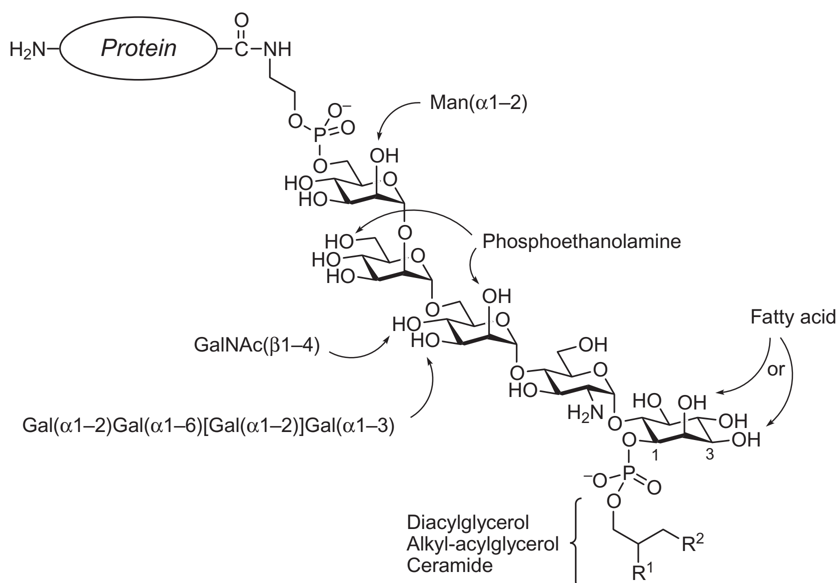
O-Glycosylation is, however, not limited to serine and threonine. In glycogenin, the priming enzyme for glycogen synthesis, glucose is α-linked to the phenolic hydroxy group of a tyrosine residue [36] and galactosyl(β1-O)tyrosine has been identified in glycoproteins of the crystalline surface layers (*S*-layers) of eubacteria [37, 38]. In the collagens, β-galactose is linked to hydroxylysine [39, 40]. Hydroxyproline, finally, is glycosylated with β-L-arabinofuranose in certain plant glycoproteins [41, 42, 43] and with α-galactose in plants and eubacteria.



Scheme 3. Carbohydrate-peptide linkages found in *O*-glycoproteins

2.2.3
GPI Anchors

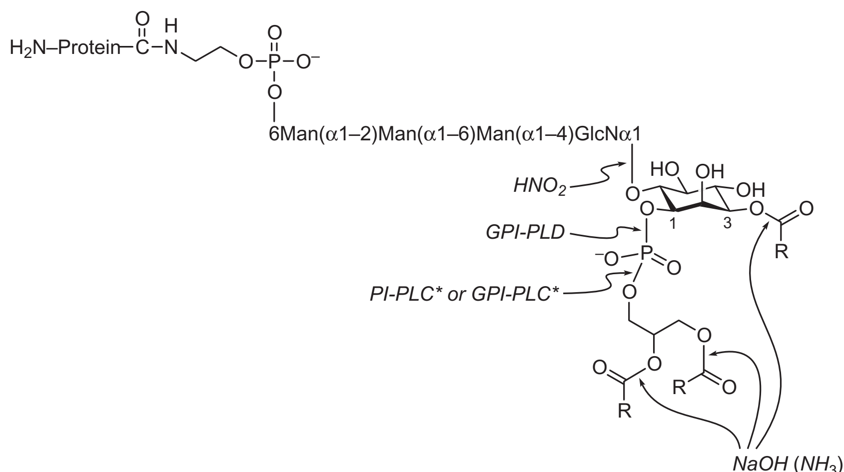
Glycosylphosphatidylinositol (GPI) anchors provide a means for anchoring proteins with a wide variety of structures and functions to the external surface of the plasma membrane of eukaryotic cells. The existence of lipid anchors for membrane proteins was initially demonstrated by the ability of phosphatidylinositol-specific phospholipase C (PI-PLC) to release alkaline phosphatase from



Scheme 4. Minimal structure and some modifications of GPI anchors

various mammalian tissues into the medium [44, 45]. Later, it was observed that fatty acids and ethanolamine were attached to the carboxy terminus of the rat Thy-1 antigen which provided direct evidence for a glycolipid anchor being covalently attached to protein [46]. In 1985 these and other results were combined to establish a common mode of membrane attachment via a GPI anchor covalently bound to the C-terminus through ethanolamine [47]. Today, we know well over 100 proteins which are GPI-anchored (reviewed in [48, 49, 50, 51, 52, 53, 54]) and the structures of several GPI anchors have been elucidated, the first being that on the *Trypanosoma brucei* variant surface glycoprotein (VSG) [55, 56].

All characterized GPI anchors share a common core of ethanolamine-PO_4-6Man(α1-2)Man(α1-6)Man(α1-4)GlcN(α1-6)*myo*-Ino-1-PO_4-lipid. This may be further processed in a cell-type and protein-specific manner (Scheme 4). The rat brain Thy-1 anchor, for example, contains additional Man(α1-2) and GalNAc(β1-4) residues, whereas the tetrasaccharide backbone of the trypanosome VSG anchor is modified with a branched chain of α-linked galactoses. In human erythrocyte acetylcholinesterase [57] and decay accelerating factor [58] the first and second mannoses carry additional phosphoethanolamines. Attached to the phosphoinositol are lipids of varying chain length and saturation including diacylglycerol, alkyl-acylglycerol, stearoyl-*lyso*glycerol, or ceramide. In the VSG anchor, for example, dimyristylglycerol has been identified. Furthermore, the inositol may be acylated with an additional fatty acid, most commonly palmitoyl, presumably at positions 2 or 3 of the inositol ring [59].



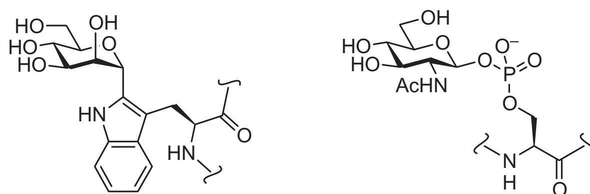
Scheme 5. Enzymatic and chemical cleavage sites of GPI anchors most commonly used for the identification and structural analysis of GPI anchors. In addition, all phosphoric esters are cleavable by aqueous hydrofluoric acid. GPI-PLC*, GPI-specific phospholipase C; GPI-PLD, GPI-specific phospholipase D; PI-PLC*, phosphatidyinositol-specific phospholipase C. *: Cleaves only if the inositol is unmodified

Biosynthesis of GPI anchors starts with the core structure assembly by sequential addition of UDP-GlcNAc (followed by *N*-deacetylation), dolichol-phosphate-mannose, and phosphoethanolamine to phosphatidylinositol and culminates in the en bloc transfer to protein shortly after the protein is synthesized. However, the biosynthetic pathways can differ strikingly between different organisms with respect to specific modifications and fatty acid remodeling occurring after completion of the core glycan. This also applies for the point when certain modifications are introduced, e.g., before or after the transfer of the GPI-moiety to the protein.

GPI anchors can be cleaved at defined positions by an array of enzymatic and chemical methods, respectively (Scheme 5). Thus, it becomes possible to identify GPI-anchored proteins and, moreover, analyze the structure and biosynthesis of GPI anchors [60].

The most obvious function of GPI anchors is to provide a stable, oriented attachment of proteins onto membranes, usually the extracellular surface of plasma membranes. However, there is much discussion over whether this highly conserved, multiple-step, complex mechanism for anchoring proteins has additional functions and it has been proposed that GPI anchors

- increase the lateral mobility of proteins,
- mediate the release or secretion of proteins,
- target proteins to apical surfaces,
- mediate endocytosis or protein turnover, and
- participate in transmembrane signaling mechanisms [54].



Scheme 6. Uncommon carbohydrate-peptide linkages identified in human Rnase U_s and proteinase I of *D. discoideum*, respectively

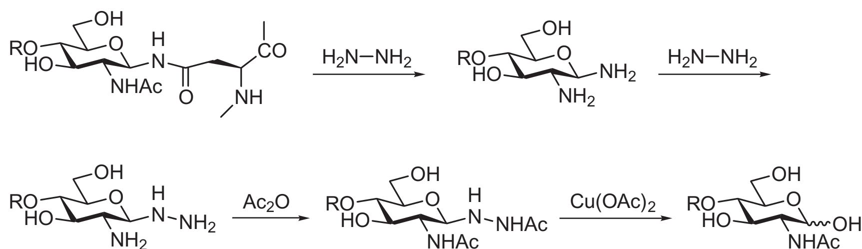
2.2.4
Uncommon Linkages

In 1994 a previously unknown type of carbohydrate-protein linkage was identified in human Rnase U_s with a mannose residue α-C-glycosidically attached to the indole ring of Trp-7 as a result of a post-translational modification (Scheme 6) [61, 62, 63]. Another uncommon linkage has been characterized in the lysosomal proteinase I of *Dictyostelium discoideum* in which β-*N*-acetylglucosamine is bound to the side-chain hydroxy group of serine via a phosphodiester bridge [64].

2.3
Methods of Glycoprotein Analysis

Structural analysis of glycans contained in glycoproteins [65,66] is a challenging task due to the lack of a general oligosaccharide sequencing method like those available for peptides or oligonucleotides. The situation is further complicated by the phenomenon of microheterogeneity which makes only minute amounts of homogenous material available. Fractionation of individual glycoforms of a given glycoprotein can be achieved using capillary electrophoresis but is not generally applicable because the differences in physicochemical properties between neutral glycoforms are often relatively small. Therefore, protein glycosylation analysis is usually carried out after release of the glycans either from the whole glycoprotein or from glycopeptides obtained by proteolytic digestion. A general strategy consists of four steps [1g]:

- release of glycans from their conjugate polypeptide,
- labeling of released glycans,
- separation of glycan mixtures, and
- sequencing of individual glycans.



Scheme 7. Hydrazinolytic release of *N*-linked glycans

2.3.1
Release of Glycans from Glycoprotein

In order to release glycans from a glycoprotein a general method is required that is independent of the protein to which the saccharides are attached. In this respect, chemical release is often preferred over the use of endoglycosidases or glycoamidases. Hydrazinolysis [67] has initially been described for release of *N*-glycans (Scheme 7) but later on it was shown that, under controlled conditions, the method is also suitable for *O*-glycans [68] enabling its application in an automated apparatus (GlycoPrep from Oxford GlycoSystems). As long as a reducing saccharide is not required at the cleavage point, *O*-glycans can be more mildly released in the form of oligosaccharide alditols by treatment with alkaline borohydride [69].

Enzymatic release [70] of *N*-glycans is most frequently effected by the peptide *N*-glycanases (PNGases) F or A which, under denaturing conditions, generally cleave the Cγ-Nδ bond of glycosylated asparagines releasing the intact *N*-linked glycans. Alternatively, endoglycosidase H (Endo H) or D (Endo D) can be applied which cleave between the first and second *N*-acetylglucosamine residues attached to asparagine. Enzymatic release of *O*-glycans is much more difficult since most *O*-glycanases currently available are highly specific. Therefore, after PNGase treatment of *N*,*O*-glycoproteins, the remaining *O*-linked glycans are often released by alkaline borohydride degradation.

2.3.2
Labeling and Separation of Glycans

Unless mass spectrometric analysis of the glycan pool is carried out, the released sugars need to be tagged to enable them to be detected. Classically, the reducing end of the oligosaccharides is reduced with alkaline sodium borotritide [71]. Today, radioactive labeling has been largely replaced by reductive amination with fluorescent compounds like 2-aminobenzamide [72], anthranilic acid [72], 8-aminonaphthalene-1,3,6-trisulfonic acid [73], 2-aminopyridine [74], 2-aminoacridone [75], or 1-aminopyrene-3,6,8-trisulfonate [76]. These fluorophores

are compatible with a range of separation techniques including HPLC (normal phase, reversed phase, and weak anion exchange), high pH anion exchange chromatography (HPAEC), lectin-affinity chromatography, polyacrylamide gel electrophoresis (PAGE), capillary electrophoresis (CE), and gel permeation chromatography (GPC).

2.3.3 *Sequencing of Glycans*

For complete structural analysis of oligosaccharides, a combination of several physical, chemical, and biochemical techniques including NMR spectroscopy, mass spectrometry, and chemical and enzymatic degradation is required. Enzymatic analysis using highly specific exoglycosidases is a powerful means of determining the sequence of glycans especially if only very small amounts of material (picomoles or less) are available. This method is based on the determination of the susceptibility of a glycan to a series of exoglycosidases of defined specificity. In the case of the reagent array analysis method (RAAM) [77], the process has been automated (GlycoPrep from Oxford GlycoSystems). The glycan is divided into several aliquots and each aliquot is incubated with a defined mixture of exoglycosidases called a reagent array. Degradation occurs in each vial until a linkage is reached which is resistant to all exoglycosidases present in that mix. The remaining "stop point" fragments containing the labeled reducing end saccharides are characterized by size (GPC). From these data, a computer program constructs the sequence of the glycan. Recently, an improved approach has been developed involving the simultaneous digestion of aliquots of a total pool of fluorescently labeled oligosaccharides with a series of multiple enzyme arrays [78].

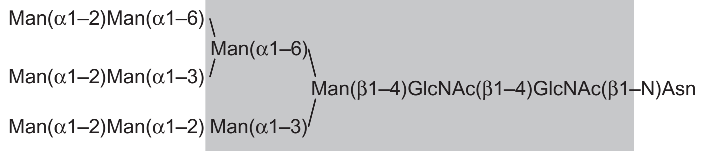
2.4 Structure of *N*-Glycans

Due to their biosynthesis, all *N*-glycoproteins share the common core pentasaccharide Man(α1-6)[Man(α1-3)]Man(β1-4)GlcNAc(β1-4)GlcNAc. The tremendous diversity of asparagine-linked oligosaccharides derives from attachment of several monosaccharides as well as up to five different carbohydrate chains, the so-called antennae, to the core. On the basis of the structure and the location of the glycans attached to the trimannosyl core, *N*-glycoproteins can be classified into four main groups. These are

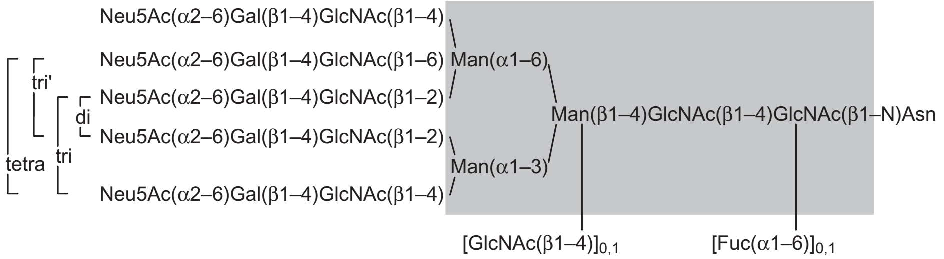
- oligomannose (high mannose),
- complex,
- hybrid, and
- the recently added [10] xylose-containing type (Scheme 8).

Oligomannose-type glycans usually contain two to six α-mannose residues bound to the core. The largest high mannose oligosaccharide thus contains nine

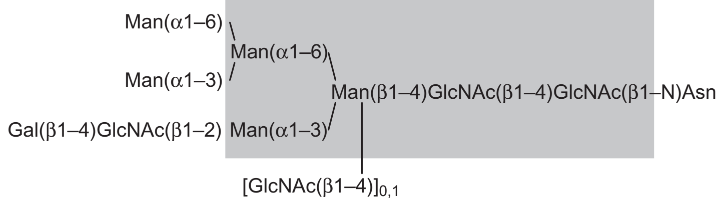
Oligomannose



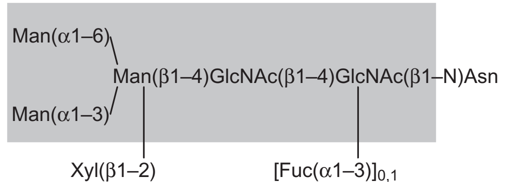
Complex



Hybrid



Xylose-containing



Scheme 8. Examples of the four main types of *N*-linked oligosaccharides of glycoproteins. The structure within the gray box represents the pentasaccharide core common to all *N*-glycans

mannose and two *N*-acetylglucosamine residues which were originally discovered in bovine thyroglobulin [79]. High molecular mass oligomannose chains with up to 100–200 mannose residues are, however, produced by yeasts. Complex-type glycans contain no mannose residues other than those in the core. To the outer two α-mannose residues up to five units of the disaccharide Gal(β1-4)GlcNAc (*N*-acetyllactosamine) are β-linked. In the bi-, tri-, and tetraanten-

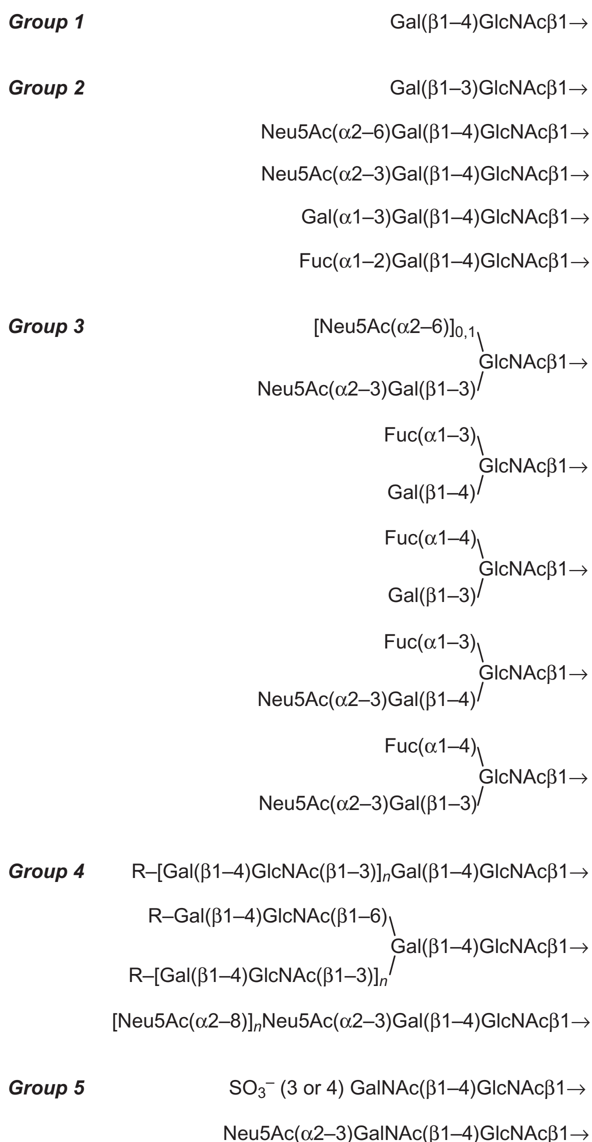
nary glycans, these branches are attached to specific positions of the core as indicated in Scheme 8.

The *N*-acetyllactosamine units are frequently substituted with sialic acid residues or additional repeats of *N*-acetyllactosamine, resulting in so-called poly-*N*-acetyllactosamines which themselves are capped by sialic acids. Further diversification can be achieved by attachment of α1-6 linked fucose to the most inner core *N*-acetylglucosamine residue and of an *N*-acetylglucosamine residue linked β1-4 to the β-mannose of the core ("bisecting" GlcNAc). In addition, various monosaccharides can be found in the antennae.

Oligosaccharides of the hybrid type have characteristic features of both complex-type and oligomannose-type glycans. One or two mannose residues are α-linked to the Man(α1-6) branch of the core pentasaccharide and one or two antennae like those found in complex-type oligosaccharides are β-linked to the Man(α1-3) branch of the core. This type of glycan results from partial processing of oligomannose-type glycans produced during biosynthesis and subsequent addition of sugars. The fourth group of *N*-glycans is named xylose-containing and is characterized by a xylose residue linked β1-2 to the β-mannose of the core. Often, there is also fucose α1-3 linked to the innermost *N*-acetylglucosamine of the core.

2.4.1
Diversity of N-Glycans

As mentioned earlier, the structural diversity of glycans is provided by the variation of monosaccharides attached to the core, the degree of branching (antennary), and, particularly, the tremendous variation in the structures of the antennae. Some side chains commonly occurring in complex *N*-glycans are depicted in Scheme 9. These structures can be roughly classified into five groups. The first group is represented by *N*-acetyllactosamine which is the starting point for further modification but is still frequently found intact. Group two consists of side chains terminated by sialic acid, fucose, or galactose. Once these modifications have taken place, no further chain elongation is possible except for formation of polysialic acids. Following sialylation it is also still possible to transfer fucose or sialic acid to the *N*-acetylglucosamine residues, resulting in some of the structures of group three, e.g., Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc (sialyl Lewis x), Neu5Ac(α2-3)Gal(β1-3)[Fuc(α1-4)]GlcNAc (sialyl Lewis a), or Neu5Ac(α2-3)Gal(β1-3)[Neu5Ac(α2-6)]GlcNAc. On the other hand, fucosylation of a subterminal *N*-acetylglucosamine residue halts the chain elongation in a non-sialylated, fucosylated form. Group four consists of long-chain structures containing either linear or branched poly-*N*-acetyllactosamine or α2-8 linked sialic acid repeatedly added to α2-3 linked sialic acid, forming polysialyl side chains. Group five side chains contain *N*-acetylgalactosamine instead of galactose. These structures may be further modified by sialylation or sulfation. There are many more variations in *N*-glycan chains identified to date [1, 8] and many more are expected to be found in future studies.



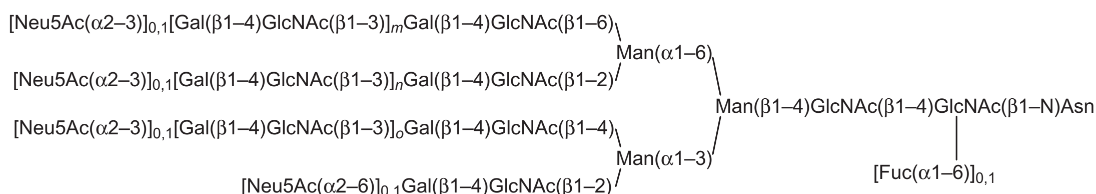
Scheme 9. Commonly occurring structures of complex-type *N*-glycan side chains. The arrows indicate the point of attachment to the pentasaccharide core

2.4.2
Cell-Type Specific Expression of N-Glycans

As described above, the enormous diversity of *N*-glycans is largely provided by variation of structures in the antennae. However, expression of specific carbohydrate epitopes, so-called antigenic determinants, is highly controlled and often cell-type specific. Critical findings revealed that poly-*N*-acetyllactosamines which are contained in many complex-type *N*-glycoproteins (but also in *O*-glycoproteins and glycolipids) provide a preferable backbone for many of the cell-type specific glycosylations, such as the ABH(O), I/i, and Lewis (Le^a, Le^x, and sialylated forms thereof) blood group and differentiation antigenic determinants [80]. Poly-*N*-acetyllactosamine repeats are not uniformly distributed among different antennae attached to the trimannosyl core. Due to the branch specificity of β1-3-*N*-acetylglucosaminyltransferase – the key enzyme for formation and elongation of poly-*N*-acetyllactosamine chains – poly-*N*-acetyllactosamine extensions are more common at the α1-6 linked mannose of the core. Tetraantennary glycans are particularly good acceptor substrates for poly-*N*-acetyllactosamine chain formation. Termination of chain elongation by sialylation leads to structures summarized in Scheme 10.

Glycosylation patterns of glycoproteins (and glycolipids) in human erythrocytes and granulocytes are well established examples for cell-type specific expression of glycans. Some important carbohydrate antigenic determinants are listed in Table 2.

Erythrocytes and granulocytes directly differentiate from the same precursor stem cells. In fetal erythrocytes, poly-*N*-acetyllactosamine is linear and expresses blood type i activity. In adult erythrocytes, poly-*N*-acetyllactosamines are branched at galactose, forming Gal(β1-4)GlcNAc(β1-3)[Gal(β1-4)GlcNAc(β1-6)]Gal(β1-4)GlcNAc, the blood group I antigen. The change from i to I specificity occurs gradually during the first year of life. Terminal galactose residues of the branched poly-*N*-acetyllactosamine chains of erythrocytes are then modified by α1-2 linked fucose, forming the H antigen. This reaction is catalyzed by an α1-2-fucosyltransferase.



Scheme 10. Representative structures of tetraantennary poly-*N*-acetyllactosamine containing *N*-glycans based on structures reported in various cells. The length of the *N*-acetyllactosamine repeats generally decreases in the order $m>n>o$. Sialic acids attached to poly-*N*-acetyllactosamine chains are usually α2-3 linked whereas α2-6 sialylation preferentially occurs at the side chain linked β1-2 to the Man(α1-3) branch of the core

Table 2. Antigenic determinants. LacNAc stands for Gal(β1-4)GlcNAc

Determinant structure	Name
Fuc(α1-2)Gal(β1-3/4)GlcNAc	Blood group H (O), type 1 or 2[a]
GalNAc(α1-3)[Fuc(α1-2)]Gal(β1-3/4)GlcNAc	Blood group A, type 1 or 2[a]
Gal(α1-3)[Fuc(α1-2)]Gal(β1-3/4)GlcNAc	Blood group B, type 1 or 2[a]
Gal(β1-3)[Fuc(α1-4)]GlcNAc	Blood group Lewis a (Le^a)
Neu5Ac(α2-3)Gal(β1-3)[Fuc(α1-4)]GlcNAc	Sialyl Lewis a (sLe^a)
Fuc(α1-2)Gal(β1-3)[Fuc(α1-4)]GlcNAc	Blood group Lewis b (Le^b)
Gal(β1-4)[Fuc(α1-3)]GlcNAc	Lewis x (Le^x)
Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc	Sialyl Lewis x (sLe^x)
Fuc(α1-2)Gal(β1-4)[Fuc(α1-3)]GlcNAc	Lewis y (Le^y)
$[\text{LacNAc}(\beta1\text{-}3)]_n$	Blood group i
LacNAc(β1-3)[LacNAc(β1-6)]LacNAc	Blood group I

[a]Type 1 contains the sequence Gal(β1-3)GlcNAc, type 2 contains the sequence Gal(β1-4)GlcNAc.

The H structure, Fuc(α1-2)Gal(β1-3/4)GlcNAc, is present in O-type blood group individuals. This structure now serves as acceptor for α1-3-*N*-acetylgalactosaminyltransferase or α1-3-galactosyltransferase, forming A and B blood group antigens, respectively.

In granulocytes, no branching of poly-*N*-acetyllactosamines is observed. α1-3-Fucosyltransferase-catalyzed modification of the linear chains leads to the Lewis x antigen. If α2-3-sialyltransferase acts before fucosylation, sialyl Lewis x is formed. These carbohydrate structures are involved in a number of cell-cell interactions during the inflammatory response and malignant transformations [81]. Production of Lewis x and sialyl Lewis x is restricted to the myeloid cell lineage and only granulocytes and monocytes are enriched with them in blood cells. In contrast, expression of the ABH blood group antigens is restricted to erythroid cells, and no other blood cells contain ABH blood group antigens. This cell-type specific expression of glycans is caused by the presence of certain key glycosyltransferases which are essential for the formation of the corresponding carbohydrate structures and of which expression is restricted to that particular cell type.

2.5 Structure of *O*-Glycans

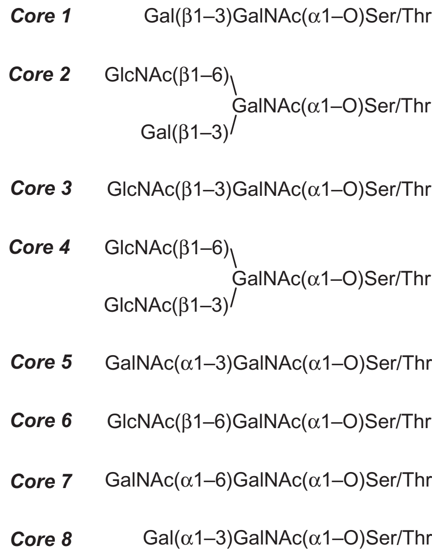
The biosynthesis of *O*-glycans differs fundamentally from that of the *N*-glycans. *O*-Glycosylation does not begin with the transfer of an oligosaccharide from a dolichol precursor to the nascent peptide chain which is then further processed. Instead, *O*-glycosylation is mainly a post-translational and post-folding event and is initiated by glycosyltransferase-catalyzed addition of a single monosac-

charide to a side-chain hydroxy group of the protein. Depending on the carbohydrate-peptide linkage, various types of *O*-glycans can be distinguished. By definition, mucin-type *O*-glycoproteins are characterized by the GalNAc(α1-O)Ser/Thr bond. This structural element, often referred to as the Tn antigen, was first identified in the mucins but is also found in other glycoproteins. Subsequent stepwise enzymatic glycosylation of the monoglycosylated polypeptide and further modification by sulfation and acetylation leads to a high degree of heterogeneity. Mucin-type *O*-glycan chains typically consist of one to 20 monosaccharide residues and they carry many of the blood group antigens and recognition signals required for intercellular and molecular interactions as well as cancer-associated and differentiation antigens.

2.5.1
Mucin-Type O-Glycans

Biosynthesis of mucin-type *O*-glycans starts with the transfer of *N*-acetylgalactosamine (GalNAc) from UDP-GalNAc to the side chain of serine or threonine catalyzed by an *N*-acetylgalactosaminyltransferase. Four different *N*-acetylgalactosaminyltransferases have been purified and cloned to date. These transferases are expressed in a tissue-specific manner and have overlapping but different substrate specificities. As a consequence, the site-specificity of mucin *O*-glycosylation is tissue specific and no simple consensus sequence can be found. However, based on statistical analysis of occupied *O*-glycosylation sites, some general rules can be deduced [1i]. *O*-Glycosylation occurs preferentially at exposed regions of the protein surface like β-turns or regions with extended conformations. Such areas are usually of low hydrophobicity and, consequently, seldom contain amino acids like Trp, Leu, Ile, and Phe. Instead, Ser, Thr, and Pro are commonly found before and after occupied *O*-glycosylation sites. Also, no bulky amino acids are found near *O*-glycosylation sites, probably due to steric hindrance. However, in contrast to *N*-linked glycans which are usually well separated from each other, *O*-glycans frequently occur in clusters on short peptide sections consisting of repeating units of Ser, Thr, and Pro. It seems to be a general feature that Thr is glycosylated more efficiently than Ser. The location of the subcellular compartment where *O*-glycosylation is initiated is still controversial and may depend on the type of *N*-acetylgalactosaminyltransferase. Several suggestions that have been made include subregions of the endoplasmic reticulum (ER), a proximal Golgi compartment, an intermediate ER-Golgi compartment (ERGIC), and the Golgi apparatus [82, 83, 84]. Subsequent stepwise enzymatic elongation of the monoglycosylated polypeptide then leads to several core structures, eight of which have been identified to date (Scheme 11). However, under certain circumstances (e.g., in particular transformed cells), elongation does not occur and the Tn antigen is exposed which might be modified to sialyl Tn (sTn) (Neu5Ac(α2-6)GalNAc(α1-O)Ser/Thr) [85, 86].

Depending on the cell type and its activation and differentiation status, the core structures can be further elongated or terminated by various glycosyltrans-

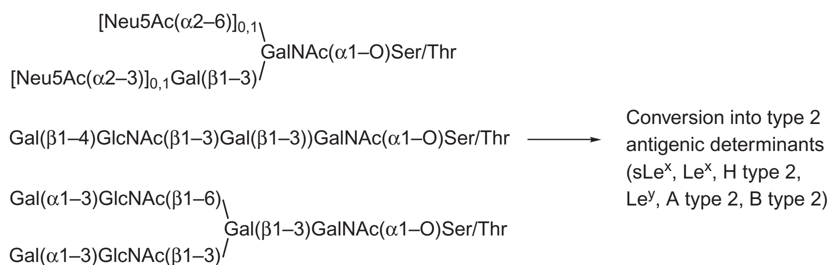


Scheme 11. Core structures of *O*-glycans. Core 1 is also called T antigen

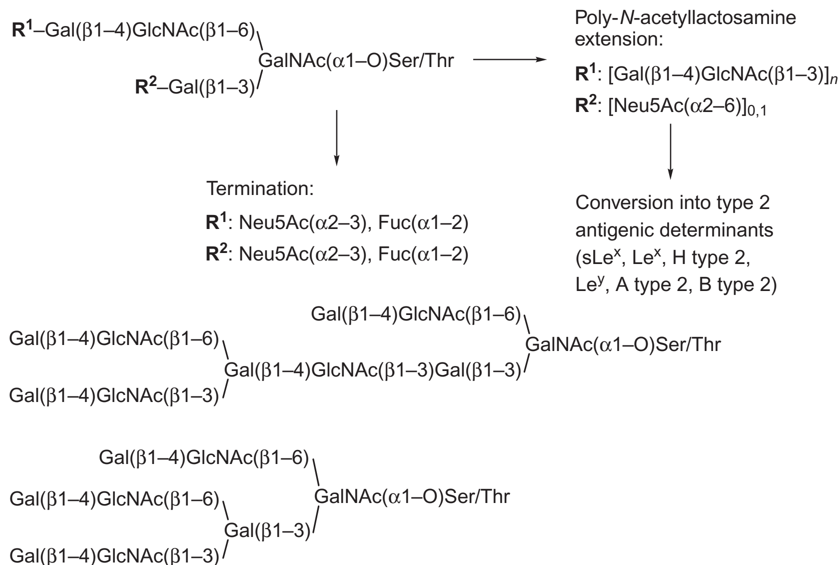
ferases, resulting in a large number of *O*-glycans, of which some are depicted in Scheme 12. Many of these glycosyltransferases have been isolated and their genes have been cloned [87, 88]. There are glycosyltransferases which specifically act on *O*-glycans (mainly those responsible for assembly of the core structures) and there are others with an activity restricted to *N*-glycans. However, a third group is able to assemble both *N*- and *O*-glycans. The antigenic determinants mentioned in Sect. 2.4.2 are generated by such transferases, i.e., the structures listed in Table 2 are found on both *N*- and *O*-glycans.

Core 1 and core 2 are the most common core structures in mucins and other secreted and cell surface glycoproteins. Core 1 (the T antigen) is usually not exposed but is monosialylated [Neu5Ac(α2-3)Gal(β1-3)GalNAc(α1-O)Ser/Thr and Neu5Ac(α2-6)[Gal(β1-3)]GalNAc(α1-O)Ser/Thr] or disialylated [Neu5Ac (α2-6)[Neu5Ac(α2-3)Gal(β1-3)]GalNAc(α1-O)Ser/Thr]. Core 1 as well as other *O*-glycans are often elongated by *N*-acetyllactosamine, which is a substrate for conversion into the type 2 antigenic determinants sLe^x, Le^x, H type 2, Le^y, A type 2, and B type 2. Core 1 is, however, a poor acceptor for poly-*N*-acetyllactosamine formation. To efficiently achieve this type of extension, core 1 first has to be transformed into core 2. Then, poly-*N*-acetyllactosamine formation can take place at the new GlcNAc(β1-6) branch. Thus, the expression of core 2 *N*-acetylglucosaminyltransferase (core 2 GlcNAcT), the enzyme responsible for the conversion of core 1 into core 2, is rate limiting for the biosynthesis of poly-*N*-acetyllactosamine extensions. The fact that poly-*N*-acetyllactosamine extensions provide an excellent backbone for the expression of antigenic determi-

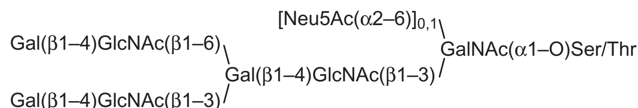
Structures generated from Core 1



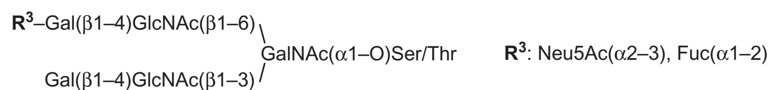
Structures generated from Core 2



Structures generated from Core 3



Structures generated from Core 4



Scheme 12. Structures of *O*-glycans commonly found in mucin-type *O*-glycoproteins

nants involved in numerous recognition events highlights the biological importance of core 2 GlcNAcT.

So far, core 3 and 4 have been found only in mucins. Core 5 is contained in glycoproteins from several species and has been reported in human adenocarcinoma [89] and meconium [90]. To date, core 6 has only been found in human glycoproteins. It has been speculated that core 6 might be generated by a β-galactosidase-catalyzed degradation of core 2 [87]. Core 7 occurs in bovine submaxillary mucin [91] and up to now has been found only as the disaccharide. All other core structures may be unsubstituted or elongated and may carry terminal antigens. Elongation can result in type 1 chains [containing the sequence Gal(β1-3)GlcNAc] or type 2 chains [containing the sequence Gal(β1-4)GlcNAc, e.g., i antigen, poly-*N*-acetyllactosamines] and branched structures (I antigen). Neu5Ac(α2-3 and -6), Fuc(α1-2, -3, and -4), GalNAc(α1-3 and -6), GalNAc(β1-4), GlcNAc(α1-4), Gal(α1-3), and sulfated residues are typically found as terminal structures on *O*-glycans although sialic acids, sulfate, and Fuc(α1-3 or -4) may be also attached to internal carbohydrate residues. Many of the terminal structures are of the Lewis and ABH blood group system (Table 2) and are also found on *N*-glycoproteins and glycolipids.

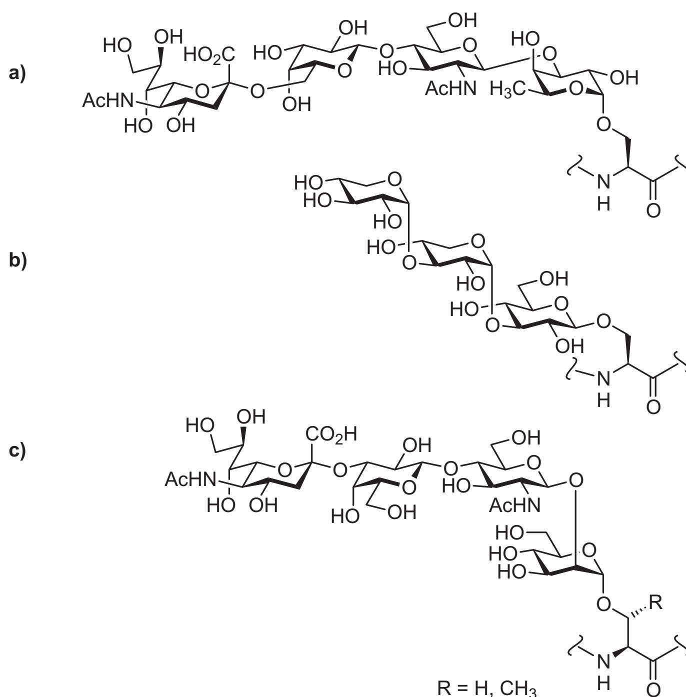
2.5.2
Non-Mucin O-Glycans

Beside the GalNAc(α1-O)Ser/Thr linkage of the mucin-type *O*-glycans, a number of additional *O*-linkages has been identified. These include mainly fucose α-*O*-linked to Ser/Thr and glucose β-*O*-linked to Ser found in the epidermal growth factor (EGF) domains of different proteins [29] and *N*-acetylglucosamine β-*O*-linked to Ser/Thr on cytoplasmic and nuclear proteins.

O-Linked fucose has been found on the EGF domains of urokinase, tissue plasminogen activator (tPA), factor VII, and factor XII, attached to the consensus sequence Cys-Xaa-Xaa-Gly-Gly-Ser/Thr-Cys. The fucosyltransferase responsible for the initiation of this type of glycosylation has been cloned [92]. On human (but not bovine) factor IX the tetrasaccharide Neu5Ac(α2-6)Gal(β1-4)GlcNAc(β1-3)Fuc(α1-O)Ser (Scheme 13, structure a) was found [93, 94]. In other EGF domains only Fuc(α1-O)Ser/Thr was detected.

On the same EGF modules glucose β-*O*-linked to the Ser of the conserved consensus sequence Cys-Xaa-Ser-Xaa-Pro-Cys occurs. In several proteins like the human and bovine blood clotting factors VII and IX, human and bovine protein Z, tPA, thrombospondin, and murine fetal antigen 1 (mFA1) Xyl(α1-3)Xyl(α1-3)Glc(β1-O)Ser (Scheme 13, structure b) and substructures thereof [Xyl(α1-3)Glc(β1-O)Ser and Glc(β1-O)Ser] have been identified. The transferase activity responsible for addition of the second xylose has been detected [95].

The *N*-acetylglucosaminyl(β1-O)serine/threonine linkage is common on cytoplasmic and nuclear proteins [27, 28]. No other sugars are connected to the GlcNAc residue. Formation of this type of modification is independent of the ac-



Scheme 13. a–c. Some structures of non-mucin *O*-glycans (cf. text)

tivity of the glycosylation machinery in the endoplasmic reticulum and the Golgi apparatus. Addition of *O*-GlcNAc is catalyzed by a highly conserved *O*-GlcNAc transferase cloned independently by two groups [96, 97]. The specificity of this enzyme resembles that of proline-directed kinases. *O*-GlcNAc is found in serine/threonine-rich sequences near Pro and Val residues. Often a negative charge is located nearby. Since the turnover of the *O*-GlcNAc glycan is faster than that of the peptide to which the sugar is attached, *O*-GlcNAc glycosylation is assumed to be reversible. Further evidence for this is the existence of a soluble *N*-acetyl-*β*-D-glucosaminidase [98]. It is likely that *O*-GlcNAc glycosylation plays a regulatory role that is analogous to protein phosphorylation.

The mannosyl(α1-O)Ser/Thr linkage is uncommon in mammalian cells. In bovine peripheral nerve α-dystroglycan clusters of Neu5Ac(α2-3)Gal(β1-4)GlcNAc(β1-2)Man(α1-O)Ser/Thr (Scheme 13, structure c) have been observed in Ser/Thr-rich domains [33]. A similar glycan [Neu5Ac(α2-3)Gal(β1-4)GlcNAc(β1-3)Man(α1-O)Ser/Thr] has been found in rat brain proteoglycans [34, 35].

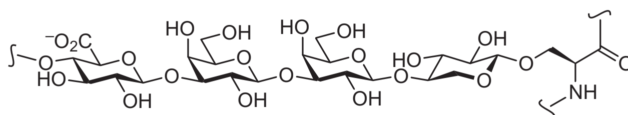
2.6
Proteoglycans

Proteoglycans [99] consist of glycosaminoglycan chains covalently attached to a protein backbone. Therefore, they can be classified as glycoproteins. However, for historical reasons and because they differ substantially from other glycoproteins, they are usually treated as a separate class of compounds. Glycosaminoglycans are linear polysaccharides consisting of hexosamine residues like *N*-sulfonylglucosamine (GlcN*S*), *N*-acetylglucosamine (GlcNAc), or *N*-acetylgalactosamine (GalNAc) alternating with glucuronic acid (GlcA) or L-iduronic acid (IdoA) or galactose (Gal) residues. Different types of glycosaminoglycans are characterized by specific disaccharide repeat patterns. These oligosaccharides are substituted to varying degrees with sulfate linked to free amino and/or hydroxy groups and range in size from ca. 15 disaccharide units to several hundred. The glycosaminoglycans contained in proteoglycans are chondroitin sulfate, dermatan sulfate, heparan sulfate, heparin, and keratan sulfate. Usually one type of glycosaminoglycan predominates but hybrid proteoglycans do exist. Another glycosaminoglycan (hyaluronan or hyaluronic acid) is not covalently attached to protein. A striking feature of all glycosaminoglycans is their negative charge caused by their sulfate and carboxyl groups which distinguishes these molecules from other glycoproteins. In addition to the glycosaminoglycan chains some proteoglycans contain *O*- and *N*-linked oligosaccharides similar or identical to those of various classes of glycoproteins.

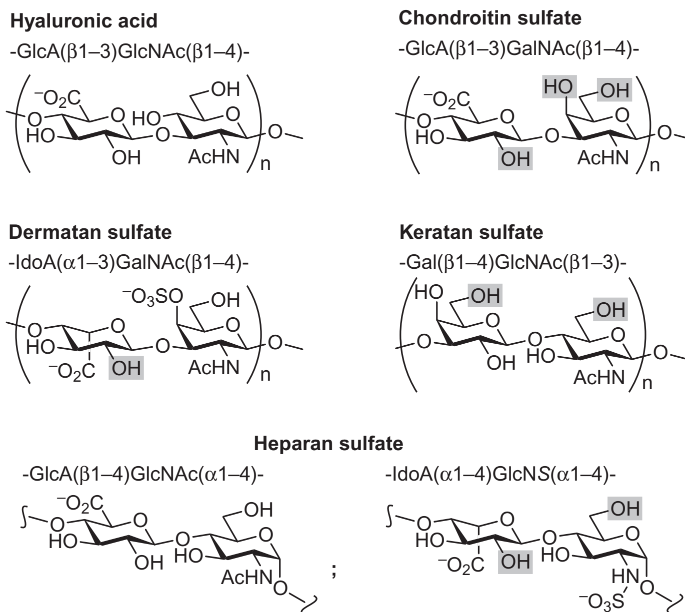
Proteoglycans are widely distributed in the animal and plant kingdoms and they possess both structural and regulatory roles [100, 101]. They are found in the extracellular matrix (ECM) and associated with specialized structures of the ECM like basement membranes. They also occur as part of plasma membranes where they function as receptors and co-receptors, and they are found intracellularly and intravesicularly in many hematopoietic cells. Many diseases like chondrodystrophies [102], atherosclerosis [103], different types of cancer [104], or Alzheimer's disease [105] are related to abnormalities in the biosynthesis and processing of proteoglycans [99a].

2.6.1
Carbohydrate-Peptide Linkage Region

Chondroitin sulfate, dermatan sulfate, heparan sulfate, and heparin are linked with their reducing end to the common core tetrasaccharide GlcA(β1-3)Gal(β1-3)Gal(β1-4)Xyl with the Xyl β-*O*-linked to a Ser residue (Scheme 14), usually adjacent to a Gly in the protein backbone [99]. The GlcA residue of this tetrasaccharide is considered as part of the linkage region because the transferase involved in the addition of this residue to the Gal-Gal disaccharide is different from the enzyme that catalyzes the incorporation of GlcA into the rest of the glycosaminoglycan. The xylose may be phosphorylated and one or both galactose residues may be sulfated.



Scheme 14. Structure of the linkage region GlcA(β1–3)Gal(β1–3)Gal(β1–4)Xyl(β1–0)Ser of proteoglycans



Scheme 15. Disaccharide repeats of different glycosaminoglycans. Hydroxy groups which may be modified by sulfation are marked gray

Skeletal keratan sulfate and corneal keratan sulfate are linked to core protein through *O*-linked and *N*-linked oligosaccharides, respectively, of the general types described in Sects. 2.4 and 2.5.

2.6.2
Structure of Glycosaminoglycans

Hyaluronan (hyaluronic acid) is the simplest glycosaminoglycan. It consists of alternating GlcA and GlcNAc residues that are (β1-3) and (β1-4) linked, respectively (Scheme 15) [99]. Hyaluronan is not sulfated and is not bound to protein, thus it is not a component of proteoglycans.

Chondroitin sulfate glycosaminoglycans are composed of the disaccharide repeat unit GlcA(β1-3)GalNAc(β1-4) with variable amounts of sulfation at the positions indicated in Scheme 15. Dermatan sulfate is formed from the precursor chondroitin sulfate by the action of an appropriate uronosyl epimerase which

converts GlcA into IdoA, giving rise to the disaccharide repeat unit IdoA(α1-3)GalNAc(β1-4) [106]. Thus, both glycosaminoglycans are found on the same protein core. However, there is still no unambiguous convention about what the proportion of IdoA must be to name a glycosaminoglycan dermatan sulfate. Ordinarily, IdoA is only formed adjacent to 4-sulfated GalNAc and not adjacent to 6-sulfated or non-sulfated GalNAc.

Heparan sulfate glycosaminoglycans and heparin are more complex. They contain two main types of disaccharide residue, although the proportions differ greatly between heparan sulfate and heparin. Heparan sulfate consists of areas of non-sulfated GlcA(β1-4)GlcNAc(α1-4) and areas of sulfated IdoA(α1-4)GlcNS(α1-4) disaccharide repeat units with the latter being obtained from the former by the action of modifying enzymes during biosynthesis. Heparin differs from heparan sulfate by its higher content of IdoA, *N*-sulfate, and *O*-sulfate with most of the disaccharide units containing 2-sulfated IdoA alternating with GlcNS which is partially 6-sulfated and occasionally 3-sulfated as well. On some heparan sulfates occasional 3-sulfation of GlcNS is also found. Although heparan sulfate and heparin are structurally similar one should keep in mind that both are found on different core proteins.

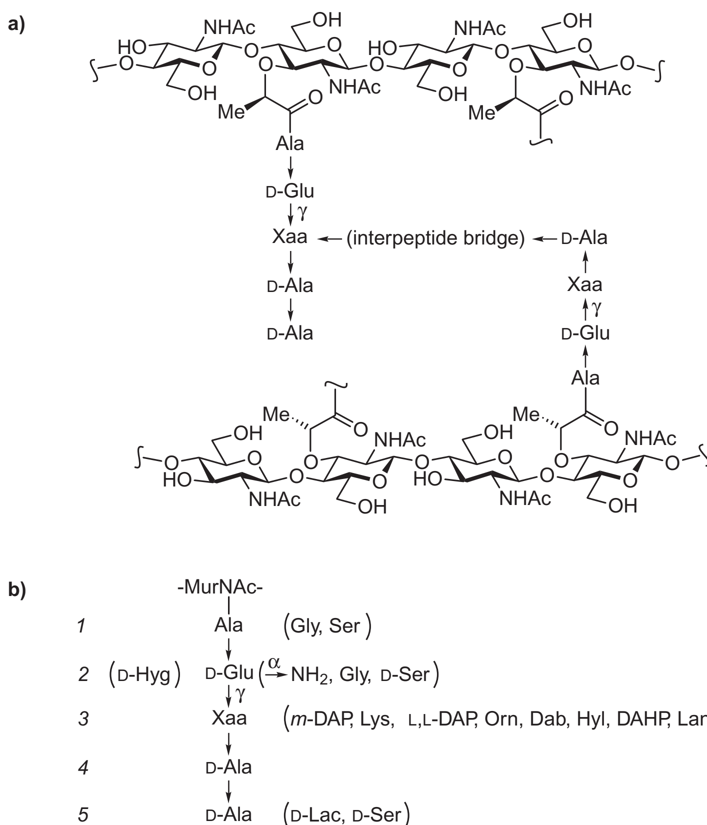
Keratan sulfate glycosaminoglycans are sulfated poly-*N*-acetyllactosamines of the type found in *N*- and *O*-glycoproteins. The degree of sulfation which takes place at the primary hydroxy groups is variable along the oligosaccharide chains with a higher degree at the non-reducing end.

3 Peptidoglycan

Peptidoglycan (synonym: murein) [107, 108] is the essential cell wall polymer of most eubacteria, both Gram-positive and Gram-negative. Although it is not regarded as a glycopeptide it will be briefly treated in this chapter since it consists of carbohydrates covalently attached to peptides. Furthermore, knowledge of peptidoglycan structure is helpful for an understanding of the mode of action of the glycopeptide antibiotics of the vancomycin group.

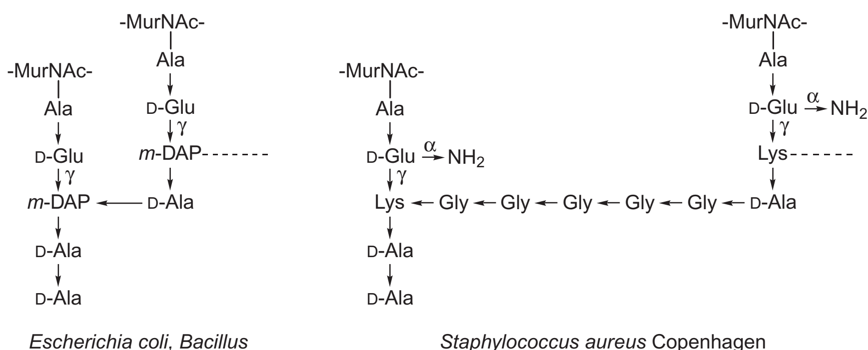
Peptidoglycan is composed of three structural features: a matrix of polysaccharide chains, the pentapeptide side chains, and the interstrand peptide cross-links, called interpeptide bridges (Scheme 16, structure a). The glycan component of all peptidoglycans consists of alternating β1-4 linked GlcNAc and *N*-acetylmuramic acid (MurNAc) (a GlcNAc with a lactyl ether at O3) residues, ranging in size from ca. 5 to 45 disaccharide units [109]. This structure is common to all bacterial peptidoglycans examined, with only minor variations. In *Staphylococcus aureus*, for example, up to 50% of the MurNAc residues are 6-*O*-acetylated [110] and in *Micrococcus lysodeikticus* up to 40% of the MurNAc residues are not attached to peptide chains [111].

Attached to the carboxylate of MurNAc is a pentapeptide with the common sequence Ala-γ-D-Glu-Xaa-D-Ala-D-Ala (non-specification of the amino acid configuration implies the natural L-form), where Xaa is *meso*-diaminopimelic



Scheme 16. a. Part of the common primary structure of bacterial peptidoglycan. Arrows indicate the polarity of peptide chains pointing from the *N*-terminus towards the C-terminus. **b.** Variations in the pentapeptide chain. Residues in parentheses may replace corresponding amino acids or, in the case of α-carboxylate modification of γ-D-Glu, may be missing. Abbreviations: *Dab*, 2,4-diaminobutyric acid; *DAHP*, 2,6-diamino-3-hydroxypimelic acid; *DAP*, diaminopimelic acid; *Hyg*, *threo*-3-hydroxyglutamic acid; *Hyl*, hydroxylysine; *Lac*, lactate; *Lan*, lanthionine

acid (*m*-DAP) for Gram-negative bacteria and Lys for most Gram-positive bacteria, with some variations [112]. During peptidoglycan biosynthesis the pentapeptide is cross-linked through amino acid Xaa to another peptide strand. This transpeptidation involves attack of the ε-amino group of the *m*-DAP or Lys residue (or amino terminus of the interpeptide bridge) onto the penultimate D-Ala of another chain, resulting in loss of the terminal D-Ala of the second peptide chain and peptide bond formation. As mentioned, variations in the pentapeptide side chains are possible and they are summarized in Scheme 16, structure b. Variations in position 1 are rare and γ-D-Glu is found universally at position 2. However, its α-carboxylate is often amidated or linked to an additional amino acid and in *Microbacterium lacticum* γ-D-Glu is replaced by *threo*-3-hydroxy-D-



Scheme 17. Peptide cross-links of *Escherichia coli*, *Bacillus*, and *Staphylococcus aureus* Copenhagen peptidoglycan

glutamic acid (3-Hyg). Most variation is found in position 3. If *m*-DAP (D,L-configuration) is incorporated, the L center is found in the pentapeptide chain and the D center in the side chain. Positions 4 and 5 (D-Ala-D-Ala) were thought to be invariant. However, the incidence of bacterial resistance to vancomycin-type glycopeptide antibiotics which recognize specifically the *N*-acyl-D-Ala-D-Ala terminus led to the discovery that these resistant strains contain altered substituents at position 5 as indicated in Scheme 16, structure b.

Peptidoglycan strands are connected to varying degrees by a network of cross-links formed between the pentapeptide side chains. The structures of these cross-links, which are quite heterogeneous, have been reviewed [107, 109, 112] and this feature has been used for taxonomic classification of bacteria [112]. Here only two examples are presented (Scheme 17). In *Escherichia coli* and strains of *Bacillus* the simplest type of linkage is realized, i.e., the direct attachment of the ε-amino group of *m*-DAP to D-Ala at position 4 of the second peptide. If L-lysine is found in position 3 of the peptide chain, which is the case for most Gram-positive bacteria, usually 1–5 amino acid residues (L-configurated or glycine) are incorporated between the ε-amino function of Lys and the D-Ala of the second peptide. As such, the interpeptide bridge of the well-studied *Staphylococcus aureus* Copenhagen strain consists of pentaglycine. However, other strains of *Staphylococcus* contain Ala, Ser, or Thr.

4 Glycopeptide Antibiotics

4.1 Discovery and Medical Use

Vancomycin was discovered in the mid-1950s in a soil sample from the jungles of Borneo by coworkers of the American pharmaceutical company Eli Lilly [113]. It is produced as a secondary metabolite by a microorganism (now) called

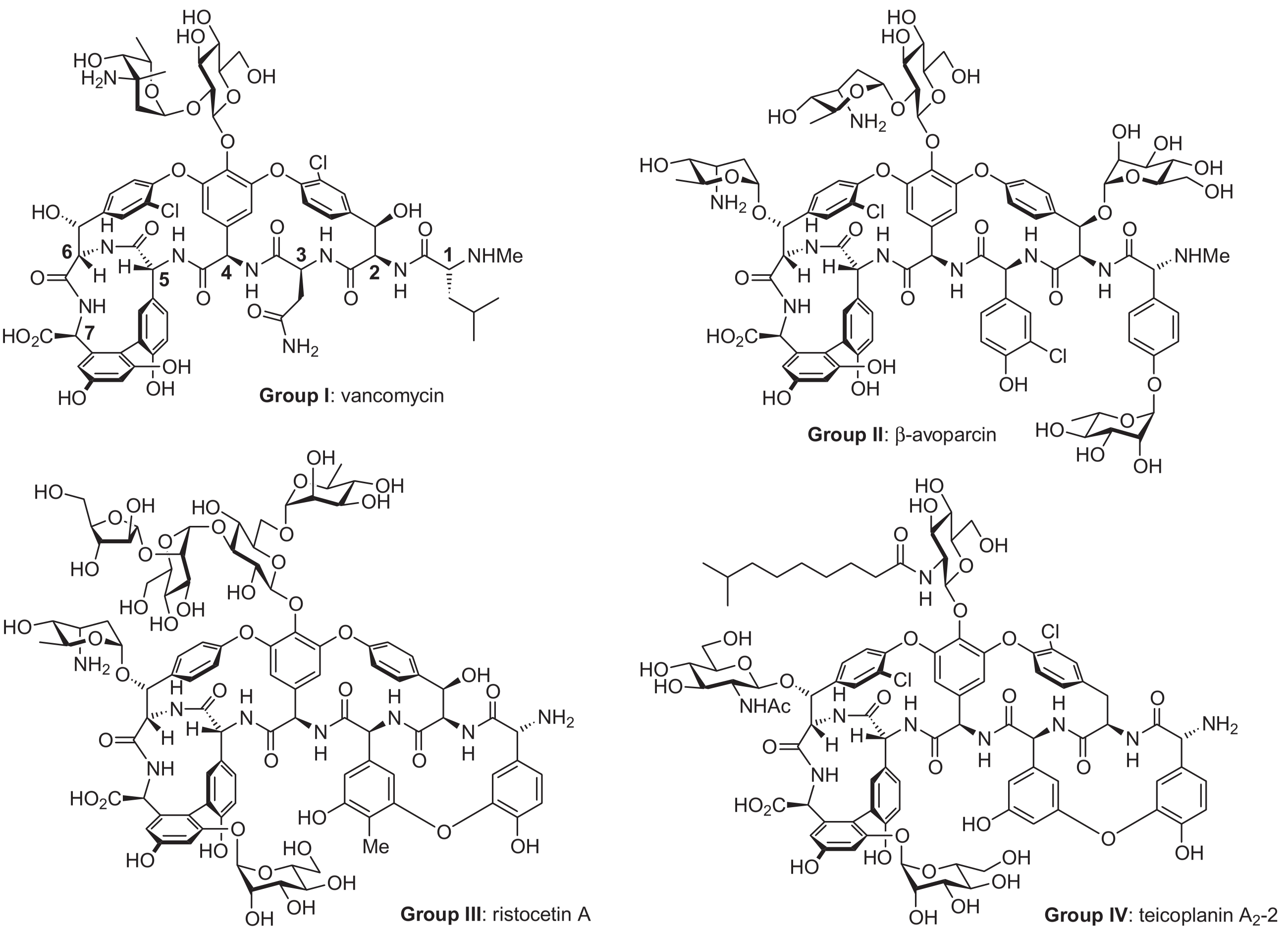
Amycolatopsis orientalis [114] and displays antibiotic activity against the majority of Gram-positive bacteria. The discovery of vancomycin was followed a year later by the isolation of ristocetin [115]. These two antibiotics were recognized as belonging to a chemical class of antibiotics called glycopeptides [116, 117, 118]. Vancomycin was first used clinically in 1959. However, toxicity problems due to inadequacies during its purification precluded its widespread use in the early years and vancomycin was for the moment overshadowed by semisynthetic penicillins and later cephalosporins. However, with improved purification techniques and as the result of the emergence of multiple-resistant and methicillin-resistant strains of *Staphylococcus aureus* (MRSA) vancomycin has become the drug of choice in the fight against these "superbugs" [117a].

Ristocetin appeared to be a promising antibacterial agent in the late 1950s but it was withdrawn following the death of several patients [119]. Avoparcin has shown growth-promoting activity in farm animals and has been marketed in Europe since 1976 as a feed additive. It is also used in the dairy industry for improvement of milk production of lactating dairy cows. Today two members of the group of the glycopeptide antibiotics are in clinical use – vancomycin and teicoplanin. Together with the aminoglycoside gentamycin, they are the antibiotics of last resort in our hospitals. However, first signs of vancomycin resistant *S. aureus* strains are now being observed [120].

4.2 Structure and Classification

The first structural work on vancomycin were degradation studies carried out by Marshall [121] followed by extensive NMR examinations in the laboratory of Williams [122]. Based on these pioneering studies and on the X-ray analysis of the degradation product CDP-1 [123], Harris and Harris published the complete structure of vancomycin in 1982 [124]. The determination of the vancomycin structure then served as basis for the structural characterization of hundreds of natural and semisynthetic glycopeptide antibiotics.

The glycopeptide antibiotics consist of a peptide backbone of 7 amino acids (see Scheme 18 for selected examples). The side chains of amino acids 2 and 4 as well as 4 and 6 and in most cases also 5 and 7 are linked to each other, either via biaryl ether or direct biaryl bonds. The resulting structural elements give rise to the phenomenon of atropisomerism. Bound to this peptide core are 1–4 saccharides (mono- to tetrasaccharides). Glycopeptide antibiotics can be subdivided in four or five groups based on their chemical structures [118, 125]. Group I, or the vancomycin type, has aliphatic amino acids at positions 1 and 3 while groups II–IV, illustrated by avoparcin, ristocetin, and teicoplanin, respectively, are characterized by aromatic residues at these positions. Members of groups III and IV have an additional biaryl ether bridge between the side chains of amino acids 1 and 3. Type IV glycopeptide antibiotics could be considered a subgroup of III since the arrangement of the amino acids in the peptide core is the same. The antibiotics of this group have a fatty acid residue attached to an



Scheme 18. Selected structures of glycopeptide antibiotics belonging to groups I–IV

amino sugar. Finally, a fifth type of antibiotics can be defined containing a characteristic tryptophan residue at position 2. However, the members of this group identified so far (chloropeptin I, complestatin, and kistamycins A and B) are not glycosylated.

Recently, total syntheses of vancomycin [126] and its aglycone [127] have been accomplished by several research groups.

References

1. Recent books and reviews in which references not included in this chapter can be found: (a) Fukuda M, Hindsgaul O (eds) (1994) Molecular Glycobiology. IRL Press, Oxford; (b) Montreuil J, Vliegenthart JFG, Schachter H (eds) (1995) Glycoproteins. Elsevier, Amsterdam; (c) Montreuil J, Vliegenthart JFG, Schachter H (eds) (1997) Glycoproteins II. Elsevier, Amsterdam; (d) Gabius H-J, Gabius S (eds) (1997) Glycosciences: Status and Perspectives. Chapman and Hall, Weinheim; (e) Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J (eds) (1999) Essentials of Glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor; (f) Lis H, Sharon N (1993) Eur J Biochem 218:1; (g) Dwek RA (1996) Chem Rev 96:683; (h) Rudd PM, Dwek RA (1997) Crit Rev Biochem Mol Biol 32:1; (i) Van den Steen P, Rudd PM, Dwek RA, Opdenakker G (1998) Crit Rev Biochem Mol Biol 33:151
2. Brennan SO, Myles T, Peach RJ, Donaldson D, George PM (1990) Proc Natl Acad Sci USA 87:26
3. Schmid K, Binette JP, Kamiyama S, Pfister V, Takahashi S (1962) Biochemistry 1:959
4. Fournet B, Strecker G, Montreuil J, Vliegenthart JFG, Schmid K, Binette JP (1978) Biochemistry 17:5206
5. Van Halbeek H, Dorland L, Vliegenthart JFG, Montreuil J, Fournet B, Schmid K (1981) J Biol Chem 256:5588
6. Kato Y, Iwase H, Hotta K (1984) Anal Biochem 138:437
7. Ashford DA, Dwek RA, Rademacher TW, Lis H, Sharon N (1991) Carbohydr Res 213:215
8. The CCSD is accessible via internet (http://www.ccrc.uga.edu). At the moment, the CCSD has more than 50,000 records, containing over 22,000 unique structures
9. Varki A (1993) Glycobiology 3:97
10. Vliegenthart JFG, Montreuil J (1995) Primary structure of glycoprotein glycans. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 13
11. Sharon N, Lis H (1997) Glycoproteins: Structure and Function. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman and Hall, Weinheim, p 133
12. Taylor CM (1998) Tetrahedron 54:11317
13. Bierhaus A, Hofmann MA, Ziegler R, Nawroth PP (1998) Cardiovasc Res 37:586
14. Johansen PG, Marshall RD, Neuberger A (1961) Biochem J 78:518
15. Nuenke RH, Cunningham LW (1961) J Biol Chem 236:2451
16. Yamashina I, Makino M (1962) J Biochem 51:359
17. Imperiali B, Shannon KL, Rickert KW (1992) J Am Chem Soc 114:7942
18. Abbadi A, McHarfi M, Aubry A, Premilat S, Boussard G, Marraud M (1991) J Am Chem Soc 113:2729
19. Imperiali B, Shannon KL, Unno M, Rickert KW (1992) J Am Chem Soc 114:7944
20. Gibbs BS, Coward JK (1999) Bioorg Med Chem 7:441; see also corrigendum on page 2121
21. Shibata S, Takeda T, Natori Y (1988) J Biol Chem 263:12483
22. Wieland F, Heitzer R, Schaefer W (1983) Proc Natl Acad Sci USA 80:5470
23. Paul G, Lottspeich F, Wieland F (1986) J Biol Chem 261:1020
24. Messner P, Sleytr UB (1988) FEBS Lett 228:317
25. Review: Messner P (1997) Glycoconjugate J 14:3
26. Schreiner R, Schnabel E, Wieland F (1994) J Cell Biol 124:1071
27. Hart GW (1997) Ann Rev Biochem 66:315
28. Snow DM, Hart GW (1998) Int Rev Cytol 181:43

29. Harris RJ, Spellman MW (1993) Glycobiology 3:219
30. Hase S, Nishimura H, Kawabata S-I, Iwanaga S, Ikenaka T (1990) J Biol Chem 265:1858
31. Gerwig GJ, Kamerling JP, Vliegenthart JFG, Morag E, Lamed R, Bayer AE (1992) Eur J Biochem 205:799
32. Gellerfors P, Axelsson K, Helander A, Johansson S, Kenne L, Lindqvist S, Pavlu B, Skottner A, Fryklund L (1989) J Biol Chem 264:11444
33. Chiba A, Matsumura K, Yamada H, Inazu T, Shimizu T, Kusunoki S, Kanazawa I, Kobata A, Endo T (1997) J Biol Chem 272:2156
34. Krusius T, Finne J, Margolis RK, Margolis RU (1986) J Biol Chem 261:8237
35. Margolis RK, Margolis RU (1993) Experientia 49:429
36. Alonso MD, Lomako J, Lomako WM, Whelan WJ (1995) FASEB J 9:1126
37. Bock K, Schuster-Kolbe J, Altman E, Allmaier G, Stahl B, Christian R, Sleytr UB, Messner P (1994) J Biol Chem 269:7137
38. Messner P, Christian R, Neuninger C, Schulz G (1995) J Bacteriol 177:2188
39. Butler WT, Cunningham LW (1966) J Biol Chem 241:3882
40. Spiro RG (1967) J Biol Chem 242:4813
41. Lamport DTA (1967) Nature 216:1322
42. Lamport DTA, Miller DH (1971) Plant Physiol 48:454
43. Allen KA, Desai NN, Neuberger A, Creeth JM (1978) Biochem J 171:665
44. Ikezawa H, Yamanegi M, Taguchi R, Miyashita T, Ohyabu T (1976) Biochim Biophys Acta 450:154
45. Low MG, Finean JB (1977) Biochem J 167:281
46. Campbell DG, Gagnon J, Reid KBM, Williams AF (1981) Biochem J 195:15
47. Ferguson MAJ, Low MG, Cross GA (1985) J Biol Chem 260:14547
48. Ferguson MAJ, Williams AF (1988) Ann Rev Biochem 57:285
49. Cross GAM (1990) Ann Rev Immunol 8:83
50. Englund PT (1993) Ann Rev Biochem 62:121
51. McConville MJ, Ferguson MAJ (1993) Biochem J 294:305
52. Field MC, Menon AK (1993) Glycolipid anchoring of cell surface proteins. In: Schlesinger MJ (ed), Lipid Modifications of Proteins. CRC Press, Boca Raton, p 83
53. Eckert V, Gerold P, Schwarz RT (1997) GPI-Anchors: Structure and Functions. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman and Hall, Weinheim, p 223
54. Cole RN, Hart GW (1997) Glycosyl-phosphatidylinositol anchors: structure, biosynthesis and function. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins II. Elsevier, Amsterdam, p 69
55. Ferguson MAJ, Homans SW, Dwek RA, Rademacher TW (1988) Science 239:753
56. Homans SW, Ferguson MAJ, Dwek RA, Rademacher TW, Anand R, Williams AF (1988) Nature 333:269
57. Deeg MA, Humphrey DR, Yang SH, Ferguson TR, Reinhold VN, Rosenberry TL (1992) J Biol Chem 267:18573
58. Walter BI, Roberts WL, Rosenberry TI, Ratnoff WD, Medof ME (1990) J Immunol 144:1030
59. Ferguson MAJ (1992) Biochem J 284:297
60. Ferguson MAJ (1992) Chemical and enzymic analysis of glycosyl-phosphatidylinositol anchors. In: Hooper NM, Turner AJ (eds) Lipid Modification of Proteins: A Practical Approach. IRL Press, Oxford, p 191
61. Hofsteenge J, Mueller DR, de Beer T, Loeffler A, Richter WJ, Vliegenthart JFG (1994) Biochemistry 33:13524
62. de Beer T, Vliegenthart JFG, Loeffler A, Hofsteenge J (1995) Biochemistry 34:11785
63. Loeffler A, Doucey M-A, Jansson AM, Mueller DR, de Beer T, Hess D, Meldal M, Richter WJ, Vliegenthart JFG, Hofsteenge J (1996) Biochemistry 35:12005
64. Gustafson GL, Gander JE (1984) Methods Enzymol 107:172
65. Hounsell EF (1997) Methods of Glycoconjugate Analysis. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman and Hall, Weinheim, p 15

66. Rudd PM, Dwek RA (1997) Curr Opin Biotechnol 8:488
67. Kobata A, Endo T (1993) In: Fukuda M, Kobata A (eds) Glycobiology: A Practical Approach. Oxford University Press, Oxford, p 79
68. Patel T, Bruce J, Merry A, Bigge C, Wormald M, Jaques A, Parekh R (1993) Biochemistry 32:679
69. Piller F, Piller V (1993) In: Fukuda M, Kobata A (eds) Glycobiology: A Practical Approach. Oxford University Press, Oxford, p 291
70. Review: O'Neill RA (1996) J Chromatogr A 720:201
71. Yoshima H, Mizuochi T, Ishii M, Kobata A (1980) Cancer Res 40:4276
72. Bigge JC, Patel TP, Bruce JA, Goulding PN, Charles SM, Parekh RB (1995) Anal Biochem 230:229
73. Jackson P (1996) Mol Biotechnol 5:101
74. Honda S, Makino A, Suzuki S, Kakehi K (1990) Anal Biochem 191:228
75. Jackson P (1991) Anal Biochem 196:238
76. Guttman A, Chen F-TA, Evangelista A, Cooke N (1996) Anal Biochem 233:234
77. Edge CJ, Rademacher TW, Wormald MR, Parekh RB, Butters TD, Wing DR, Dwek RA (1992) Proc Natl Acad Sci USA 89:6338
78. Rudd PM, Morgan BP, Wormald MR, Harvey DJ, van den Berg CW, Davis SJ, Ferguson MAJ, Dwek RA (1997) J Biol Chem 272:7229
79. Ito S, Yamashita K, Spiro RG, Kobata A (1977) J Biochem 81:1621
80. Watkins WM (1995) Molecular basis of antigenic specificity in the ABO, H and Lewis blood-group systems. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 313
81. Simanek EE, McGarvey GJ, Jablonowski JA, Wong C-H (1998) Chem Rev 98:833
82. Perez-Vilar J, Hidalgo J, Velasco A (1991) J Biol Chem 266:23967
83. Roth J, Wang Y, Eckhardt AE, Hill RL (1994) Proc Natl Acad Sci USA 91:8935
84. Roth J (1995) Compartmentation of glycoprotein biosynthesis. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 287
85. Yamashita Y, Chung YS, Horie R, Kannagi R, Sowa M (1995) J Natl Cancer Inst 87:441
86. Brockhausen I, Schutzbach J, Kuhns W (1998) Acta Anat 161:36
87. Brockhausen I (1995) Biosynthesis of *O*-glycans of the *N*-acetylgalactosamine-α-Ser/Thr linkage type. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins. Elsevier, Amsterdam, p 201
88. Brockhausen I, Schachter H (1997) Glycosyltransferases Involved in *N*- and *O*-Glycan Biosynthesis. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman and Hall, Weinheim, p 79
89. Kurosaka A, Nakajima H, Funakoshi J, Matsuyama M, Nagayo T, Yamashina I (1983) J Biol Chem 258:11594
90. (a) Capon C, Leroy Y, Wieruszeski J-M, Ricart G, Strecker G, Montreuil J, Fournet B (1989) Eur J Biochem 182:139; (b) Hounsell EF, Lawson AM, Feeney J, Gooi HC, Pickering NJ, Stoll MS, Lui SC, Feizi T (1985) Eur J Biochem 148:367
91. Chai W, Hounsell EF, Cashmore GC, Rosankiewicz JR, Bauer CJ, Feeney J, Feizi T, Lawson AM (1992) Eur J Biochem 203:257
92. Wang Y, Lee GF, Kelley RF, Spellman MW (1996) Glycobiology 6:837
93. Nishimura H, Takao T, Hase S, Shimonishi Y, Iwanaga S (1992) J Biol Chem 267:17520
94. Harris RJ, van Halbeek H, Glushka J, Basa LJ, Ling VT, Smith KJ, Spellman MW (1993) Biochemistry 32:6539
95. Minamida S, Aoki K, Natsuka S, Omichi K, Fukase K, Kusumoto S, Hase S (1996) J Biochem 120:1002
96. Kreppel LK, Blomberg MA, Hart GW (1997) J Biol Chem 272:9308
97. Lubas WA, Frank DW, Krause M, Hanover J (1997) J Biol Chem 272:9316
98. Dong DL, Hart GW (1994) J Biol Chem 269:19321
99. Recent reviews in which further references can be found: (a) Wight TN (1999) Biosynthesis of Proteoglycans. In: Pinto BM (ed), Carbohydrates and Their Derivatives Including Tannins, Cellulose, and Related Lignins (Comprehensive Natural Products Chemistry, vol

3). Elsevier, Amsterdam, p 161; (b) Silbert JE, Bernfield M, Kokenyesi R (1997) Proteoglycans: a special class of glycoproteins. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins II. Elsevier, Amsterdam, p 1; (c) Kresse H (1997) Proteoglycans – Structure and Functions. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives. Chapman and Hall, Weinheim, p 201
100. Wight TN, Kinsella MG, Qwarnstrom EE (1992) Curr Opin Cell Biol 4:793
101. Williams KJ, Fukui IV (1997) Curr Opin Lipidol 8:253
102. Superti-Furga A, Haestbacka J, Rossi A, Van der Harten JJ, Wilcox WR, Cohn DH, Rimoin DL, Steinmann B, Lander ES, Gitzelmann R (1996) Ann N Y Acad Sci 785:195
103. Wight TN (1996) In: Fuster V, Ross R, Topol EJ (eds) Atherosclerosis and Coronary Artery Disease. Raven Publishers, Philadelphia, PA, p 421
104. Iozzo RV (1988) Cancer Metastasis Rev 7:39
105. Snow AD, Wight TN (1989) Neurobiol Aging 10:481
106. It should be kept in mind that the anomeric configuration of the uronic acid remains the same. It is only the terminology that has changed due to the conversion of a D-sugar (GlcA) into an L-sugar (IdoA)
107. Bugg TDH (1999) Bacterial Peptidoglycan. Biosynthesis and its Inhibition. In: Pinto BM (ed), Carbohydrates and Their Derivatives Including Tannins, Cellulose, and Related Lignins (Comprehensive Natural Products Chemistry, vol 3). Elsevier, Amsterdam, p 241
108. Labischinski H, Maidhof N (1994) Bacterial peptidoglycan: overview and evolving concepts. In: Ghuysen J-M, Hakenbeck R (eds) Bacterial Cell Wall (New Compr Biochem, vol 27). Elsevier, Amsterdam, p 23
109. Ghuysen JM (1968) Bacteriol Rev 32:425
110. Tipper DJ, Strominger JL, Ensign JC (1967) Biochemistry 6:906
111. Leyh-Bouille M, Ghuysen J-M, Tipper DJ, Strominger JL (1966) Biochemistry 5:3079
112. Schleifer KH, Kandler O (1972) Bacteriol Rev 36:407
113. McCormick MH, Stark WM, Pittenger GE, Pittenger RC, McGuire JM (1956) Antibiot Annu 1955–1956:606
114. Lechevalier MP, Prauser H, Labeda DP, Ruan J-S (1986) Int J Syst Bacteriol 36:29
115. Philip JE, Schenck JP, Hargie MP (1957) Antibiot Annu 1955–1956:699
116. Nagarajan R (1994) Glycopeptide Antibiotics. Marcel Dekker, New York
117. (a) Williams DH (1996) Nat Prod Rep 13:469; (b) Williams DH, Bardsley B (1999) Angew Chem Int Ed Engl 38:1173
118. Nicolaou KC, Boddy CNC, Bräse S, Winssinger N (1999) Angew Chem Int Ed Engl 38:2096
119. Howard MA, Firkin BG (1971) Thromb Diath Haemorrh 26:362
120. (a) Hiramatsu K (1998) Drug Resist Updates 1:135; (b) Levy SB (1998) Sci Am 3:46
121. Marshall FJ (1965) J Med Chem 8:18
122. Williams DH, Kalman JR (1977) J Am Chem Soc 99:2768
123. Sheldrick GM, Jones PG, Kennard O, Williams DH, Smith GA (1978) Nature 271:223
124. Harris CM, Harris TM (1982) J Am Chem Soc 104:4293
125. Lancini GC (1989) Fermentation and biosynthesis of glycopeptide antibiotics. In: Bushell ME, Graefe U (eds) Bioactive Metabolites from Microorganisms (Prog Ind Microbiol, vol 27). Elsevier, Amsterdam, p 283
126. (a) Nicolaou KC, Mitchell HJ, Jain NF, Winssinger N, Hughes R, Bando T (1999) Angew Chem Int Ed Engl 38:240; (b) Nicolaou KC, Li H, Boddy CNC, Ramanjulu JM, Yue T-Y, Natajaran S, Chu X-J, Bräse S, Rübsam F (1999) Chem Eur J 5:2584 and references therein
127. (a) Evans DA, Wood MR, Trotter BW, Richardson TI, Barrow JC, Katz JL (1998) Angew Chem Int Ed Engl 37:2700; (b) Boger DL, Miyazaki S, Kim SH, Wu JH, Loiseleur O, Castle SL (1999) J Am Chem Soc 121:3226