

INTERNATIONAL UNION OF PURE  
AND APPLIED CHEMISTRY  
and  
INTERNATIONAL UNION OF BIOCHEMISTRY  
JOINT COMMISSION ON BIOCHEMICAL NOMENCLATURE\*

**NOMENCLATURE OF GLYCOPROTEINS,  
GLYCOPEPTIDES AND PEPTIDOGLYCANS**

(Recommendations 1985)

*Prepared for publication by*

N. SHARON

The Weizmann Institute of Science, Rehovot, Israel

\*Membership of the Commission during the preparation of this report (1983–1987) was as follows:

*Chairman:* H. B. F. Dixon (UK); *Secretary:* A. Cornish-Bowden (UK); *Members:* C. Liébecq (Belgium, representing the IUB Committee of Editors of Biochemical Journals); K. L. Loening (USA); G. P. Moss (UK); J. Reedijk (Netherlands); S. F. Velick (USA); P. Venetianer (Hungary); J. F. G. Vliegthart (Netherlands).

Additional contributors to the formulation of these recommendations:

*Nomenclature Committee of IUB (NC-IUB)* (those additional to JCBN): H. Bielka (GDR); C. R. Cantor (USA); P. Karlson (FRG); N. Sharon (Israel); E. J. Van Lenten (USA); E. C. Webb (Australia).

*List of Reviewers:* C. E. Ballou; E. Buddecke; E. A. Davidson; R. J. Doyle; D. Duksin; J.-M. Ghuyssen; T. Hardingham; V. C. Hascall; D. Heinegård; F. W. Hemming; D. Horton; R. C. Hughes; W. B. Jakoby; R. W. Jeanloz; O. Kandler; T. Laurent; Y. C. Lee; U. Lindahl; B. Lindberg; H. Lis; F. G. Loontjens; J. Montreuil; H. Muir; A. Neuberger; E. F. Neufeld; T. Osawa; L. Rodén; H. Schachter; R. Schauer; K. Schmid; G. D. Shockman; J. E. Silbert; C. C. Sweeley and I. Yamashina.

Correspondence on these recommendations should be addressed to the Secretary of the Commission, Dr. A. Cornish-Bowden, CNRS-CBM, 31 chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 9, France.

# Nomenclature of glycoproteins, glycopeptides and peptidoglycans (Recommendations 1985)

## CONTENTS

1. Introduction
2. Nomenclature
  - 2.1. Glycoproteins, proteoglycans and glycosaminoglycans
  - 2.2. Glycopeptides, glyco-amino-acids and glycosyl-amino-acids
  - 2.3. Peptidoglycans
3. Comments
  - 3.1. Form of carbohydrate in glycoproteins
  - 3.2. Proteoglycans
  - 3.3. Microheterogeneity
  - 3.4. Glycopeptides, etc.
  - 3.5. *N,O*-Glycoproteins
  - 3.6. Asialoglycoproteins and asialo-agalactoglycoproteins
  - 3.7. Condensed representation of sugar chains
  - 3.8. Representation of *N*-linked oligosaccharides
  - 3.9. Peptidoglycans
4. References
5. Appendix: Implied configurations and ring sizes.

## 1. INTRODUCTION

Various types of compound consisting of carbohydrates covalently linked with other types of chemical constituent are classified under the general name of *glycoconjugates*. The major groups of glycoconjugates are the glycoproteins, glycopeptides, peptidoglycans, glycolipids and lipopolysaccharides. The first three of these are considered in the present document. The nomenclature of glycolipids has been the subject of an earlier document [1], which is now under revision.

As early as 1907, when very little was known about glycoproteins, they were defined by the Committees on Protein Nomenclature of the American Society of Biological Chemists and the American Physiological Society as 'compounds of the protein molecule with a substance or substances containing a carbohydrate group other than a nucleic acid' [2]. In spite of the enormous progress in our knowledge of the occurrence, biosynthesis, properties and functions of the glycoproteins, this definition remains appropriate.

Glycoproteins are widely distributed in all forms of life, with the possible exception of the eubacteria. They occur in cells, both in soluble and membrane-bound forms, as well as in the extracellular matrix and in extracellular fluids. The commonest glycoproteins are those in which the carbohydrate is linked to the protein by glycosyl linkages. Glycosylation represents one of the important co-translational and post-translational modifications of proteins.

The term 'glycoprotein' should include proteoglycans, which in the past were considered as a separate class of compound. Proteoglycans are those glycoproteins whose carbohydrate moieties consist of long, unbranched chains of alternating residues of hexosamine and uronic acid or galactose, often sulfated. Such polysaccharides belong to the

class of glycosaminoglycans. Proteoglycans were not considered as glycoproteins in the past because their carbohydrate seemed to differ so greatly from the comparatively small, branched, usually unsulfated carbohydrate units, devoid of repeating units, found in other glycoproteins. Another reason was their unique distribution: the proteoglycans are found mainly in connective tissues, where they contribute to the organization and physical properties of the extracellular matrix. It is known, however, that proteoglycans too are glycosylated proteins, synthesized by the action of glycosyltransferases, and that repeating disaccharide units are present in typical glycoproteins as well. Proteoglycans are therefore considered to be a subclass of glycoproteins.

The terms 'protein glycosylation (or glycosylation)' and 'glucosylated (or glycosylated) hemoglobin' have been used improperly to refer to the products of non-enzymic reactions between glucose or other sugars and free amino groups of proteins. Compounds formed in this manner are not glycosides, however, as they result from the formation of a Schiff base followed by Amadori rearrangement to 1-deoxyketosyl derivatives of the proteins. For example, the product of the reaction between glucose and hemoglobin is not glucosylated hemoglobin but *N*-(1-deoxyfructosyl)hemoglobin. The term 'glycation' is suggested for all reactions that link a sugar to a protein or peptide, whether or not they form a glycosyl bond. The product of glycation is a glycoprotein, or, in the special case of the reaction with hemoglobin, glycohemoglobin. When appropriate, a more precise name such as deoxyfructosylhemoglobin may be used.

Peptidoglycans, sometimes referred to as mureins, are glycoconjugates found only in bacterial cell walls. Although they are also composed of carbohydrates and amino acids covalently linked, they are a distinct class of compound as (a) they do not contain a protein portion; (b) they contain sugars not found elsewhere; and (c) they consist of linear polysaccharides (of the class of glycosaminoglycan) crosslinked by oligopeptides, thus forming a huge and rigid network.

No official nomenclature of glycoproteins, glycopeptides and peptidoglycans has been available hitherto. In drawing up the present document, we follow the general rules of biochemical nomenclature [3], especially the more recent ones on carbohydrates [4-7] and amino acids and peptides [8].

## 2. NOMENCLATURE

### 2.1. Glycoproteins, proteoglycans and glycosaminoglycans

A *glycoprotein* is a compound containing carbohydrate (or glycan) covalently linked to protein. The carbohydrate may be in the form of a monosaccharide, disaccharide(s), oligosaccharide(s), polysaccharide(s), or their derivatives (e.g. sulfo- or phospho-substituted). One, a few, or many carbohydrate units may be present. *Proteoglycans* are a subclass of glycoproteins in which the carbohydrate units are polysaccharides that contain amino sugars. Such polysaccharides are also known as *glycosaminoglycans*.

Table 1. Structure of glycosaminoglycans

Based mainly on Lindahl, U. & Höök, M. (1978), *Annu. Rev. Biochem.* 47, 385–417; and Rodén, L. (1980) in *The biochemistry of glycoproteins and proteoglycans* (Lennarz, W. J., ed.), pp. 267–371, Plenum Press, New York

Name	Repeating disaccharide	Sulfation	Linkage region
Hyaluronic acid	GlcNAc( $\beta$ 1 $\rightarrow$ 4)GlcA( $\beta$ 1 $\rightarrow$ 3)	None	not proved to be linked to protein
Chondroitin 4- and 6-sulfate	GalNAc( $\beta$ 1 $\rightarrow$ 4)GlcA( $\beta$ 1 $\rightarrow$ 3)	GalNAc-4- <i>O</i> -sulfate, 6- <i>O</i> -sulfate or hybrids	GlcA( $\beta$ 1 $\rightarrow$ 3)Gal( $\beta$ 1 $\rightarrow$ 3)Gal( $\beta$ 1 $\rightarrow$ 4)Xyl( $\beta$ 1 $\rightarrow$ O)Ser
Dermatan sulfate	GalNAc( $\beta$ 1 $\rightarrow$ 4)GlcA( $\beta$ 1 $\rightarrow$ 3) and GalNAc( $\beta$ 1 $\rightarrow$ 4)L-IdoA( $\alpha$ 1 $\rightarrow$ 3)	GalNAc-4- <i>O</i> -sulfate, or 6- <i>O</i> -sulfate; L-iduronic acid may be sulfated at position 2	GlcA( $\beta$ 1 $\rightarrow$ 3)Gal( $\beta$ 1 $\rightarrow$ 3)Gal( $\beta$ 1 $\rightarrow$ 4)Xyl( $\beta$ 1 $\rightarrow$ O)Ser
Heparin and heparan sulfate	GlcNR( $\alpha$ 1 $\rightarrow$ 4)GlcA( $\beta$ 1 $\rightarrow$ 4) and GlcNR( $\alpha$ 1 $\rightarrow$ 4)L-IdoA( $\alpha$ 1 $\rightarrow$ 4)	In heparin (highly sulfated), R mostly SO <sub>3</sub> H, little acetyl; in heparan sulfate generally less SO <sub>3</sub> H and more acetyl than in heparin. Some of the <i>N</i> -acetyl-D-glucosamine is 6- <i>O</i> -sulfated, and some of the L-iduronic acid is 2- <i>O</i> -sulfated	GlcA( $\beta$ 1 $\rightarrow$ 3)Gal( $\beta$ 1 $\rightarrow$ 3)Gal( $\beta$ 1 $\rightarrow$ 4)Xyl( $\beta$ 1 $\rightarrow$ O)Ser
Keratan sulfate	GlcNAc( $\beta$ 1 $\rightarrow$ 3)Gal( $\beta$ 1 $\rightarrow$ 4)	GlcNAc-6- <i>O</i> -sulfate; Gal-6- <i>O</i> -sulfate	see Fig. 1

## 2.2. Glycopeptides, glyco-amino-acids and glycosyl-amino-acids

A *glycopeptide* is a compound consisting of carbohydrate linked to an oligopeptide composed of L- and/or D-amino acids. A *glyco-amino-acid* is a saccharide attached to a single amino acid by any kind of covalent bond. A *glycosyl-amino-acid* is a compound consisting of saccharide linked through a glycosyl linkage (*O*-, *N*- or *S*-) to an amino acid. (The hyphens are needed to avoid implying that the carbohydrate is necessarily linked to the amino group.)

## 2.3. Peptidoglycans

A *peptidoglycan* consists of a glycosaminoglycan formed by alternating residues of D-glucosamine and either muramic acid {2-amino-3-*O*-[(*S*)-1-carboxyethyl]-2-deoxy-D-glucose} or L-talosaminuronic acid (2-amino-2-deoxy-L-taluronic acid), which are usually *N*-acetylated or *N*-glycoloylated. The carboxyl group of the muramic acid is commonly substituted by a peptide containing residues of both L- and D-amino acids, whereas that of L-talosaminuronic acid is substituted by a peptide consisting of L-amino acids only.

## 3. COMMENTS

### 3.1. Form of carbohydrate in glycoproteins

In many glycoproteins (e.g. plasma glycoproteins such as human  $\alpha_1$ -acid glycoprotein or fetuin) the carbohydrate is in the form of oligosaccharides, linear or branched, the latter containing up to about 20 monosaccharide residues; glycoproteins containing mono- or disaccharide units are also known [e.g. collagens, fish antifreeze glycoproteins, sheep submaxillary (or submandibular) glycoproteins], as well as those that contain oligosaccharides that consist of repeating units of *N*-acetylglucosamine (e.g. band 3 of the human erythrocyte membrane).

### 3.2. Proteoglycans

A *proteoglycan* is a protein glycosylated by one or more (up to about 100) glycosaminoglycans (Table 1). The glycosaminoglycans of the proteoglycans are linear polymers of up to about 200 repeating disaccharide units that consist of a hexosamine (D-glucosamine or D-galactosamine) alternating with a uronic acid (D-glucuronic or L-iduronic) or a neutral sugar (D-galactose). The hexosamines are usually *N*-acetylated, and in some of the glycosaminoglycans the D-glucosamine is *N*-sulfated. Varying degrees of sulfation occur

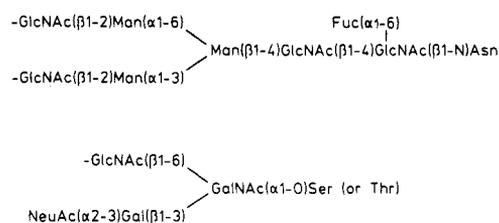


Fig. 1. Linkage region of corneal and skeletal keratan sulfate. The upper structure shows the linkage region in keratan sulfate type I (corneal); the lower shows that in keratan sulfate type II (skeletal)

in other positions of the hexosamines as well as on the L-iduronic acid. The chain of repeating units is linked to the protein by an oligosaccharide of a structure different from the repeating units. This *linkage region* (Table 1) is identical in most of the proteoglycans (chondroitin sulfates, dermatan sulfate, heparin, and heparan sulfate), but is different in keratan sulfate. This last glycosaminoglycan contains repeating units of *N*-acetylglucosamine that are *O*-sulfated. Because of their content of uronic acid and/or ester sulfate, the glycosaminoglycans of the proteoglycans are anionic polyelectrolytes and have been referred to as 'acidic glycosaminoglycans' (equivalent to the older term 'acid mucopolysaccharides'). Proteoglycans may also contain one or more oligosaccharides of structures similar to those found in other glycoproteins.

Several diseases of proteoglycan metabolism are recognized, including the Hunter and Hurler syndromes; they are referred to as mucopolysaccharidoses, but strictly speaking should go by the term 'glycoproteinoses', which encompasses other diseases of glycoprotein metabolism such as mannosidosis and fucosidosis.

### 3.3. Microheterogeneity

A particular glycoprotein may occur in forms that differ in the structure of one or more of its carbohydrate units, a phenomenon known as *microheterogeneity*. Such differences may affect both the size and charge of individual glycoproteins; occasionally, the differences may be solely due to changes in linkage position in a carbohydrate unit. For example, chicken ovalbumin contains a single glycosylated amino acid residue (Asn-293), but more than a dozen different oligosaccharides have been identified at that site. In proteoglycans, individual glycosaminoglycan chains may differ in structure, e.g. the degree of sulfation, the ratio of glucuronic acid to L-iduronic acid, and chain length. Proteoglycans are therefore highly polydisperse, in contrast to typical glycoproteins, which, in spite of their microheterogeneity, are not markedly polydisperse in their molecular size.

Table 2. Commonly occurring glycosyl-amino-acids

Based on Sharon, N. & Lis, H. (1982) in *The proteins*, 3rd edn (Neurath, H. & Hill, R. L., eds) vol. 5, pp. 1–144, Academic Press, New York

Example	Structure	Occurrence
<i>N</i> -Glycoside* $\beta$ - <i>N</i> -Acetylglucosaminyl- -asparagine, (GlcNAc-)Asn		widely distributed in animals, plants and microorganisms
<i>O</i> -Glycosides $\alpha$ - <i>N</i> -Acetylgalactosaminyl- -serine or -threonine, (GalNAc-)Ser, (GalNAc-)Thr		glycoproteins of animal sources
$\beta$ -Xylosyl-serine (Xyl-)Ser		proteoglycans, human thyroglobulin
$\beta$ -Galactosyl-hydroxylysine (Gal-)Hyl		collagens
$\beta$ -L-Arabinosyl- -hydroxyproline, (L-Ara-)Hyp		plant and algal glycoproteins

\* Traditionally the term 'glycoside' referred explicitly to O-linked compounds; we use the term 'N-glycoside' rather than 'nitrogen analogue of glycoside', which is unnecessarily cumbersome.

### 3.4. Glycopeptides etc.

*Glycopeptides*, *glyco-amino-acids*, *glycosyl-amino-acids* and *glycosylpeptides* are obtained by enzymic or chemical cleavage of glycoproteins, or by chemical or enzymic synthesis.

Examples of glycosyl-amino-acids that constitute the common linking compounds between the carbohydrate and protein in glycoproteins are as follows (Table 2): 2-acetamido-*N*-(L-aspart-4-yl)-2-deoxy- $\beta$ -D-glucopyranosylamine, i.e. *N*<sup>4</sup>-(*N*-acetyl- $\beta$ -D-glucosaminyl)asparagine, which is abbreviated to (GlcNAc-)Asn (parentheses here around the carbohydrates placed next to the symbol for an amino-acid residue indicate substitution on its side chain: see section 3AA-17.2 in ref. 8); *O*<sup>3</sup>-(*N*-acetyl- $\alpha$ -D-galactosaminyl)serine and threonine, (GalNAc-)Ser and (GalNAc-)Thr; *O*- $\beta$ -D-xylosylserine, (Xyl-)Ser; *O*<sup>5</sup>- $\alpha$ -D-galactosylhydroxylysine<sup>1</sup>, (Gal-)Hyl; and  $\beta$ -L-arabinosylhydroxyproline<sup>2</sup>, (L-Ara-)Hyp. Another example of a glyco-amino-acid is [Man<sub>9</sub>-GlcNAc( $\beta$ 1 $\rightarrow$ 4)-GlcNAc]-Asn, isolated from a proteolytic digest of soybean agglutinin. Such a compound may be referred to as *oligosaccharyl-asparagine*.

<sup>1</sup> This compound could more strictly be described as 5( $\alpha$ -D-galactopyranosyloxy)-L-lysine, but hydroxylysine is regarded as a trivial name [8], so names are based on it.

<sup>2</sup> Hydroxyproline is the trivial name for *trans*-4-hydroxy-L-proline.

### 3.5. N,O-Glycoproteins

If desirable, the linkage between carbohydrate and protein may be indicated by the locants *N*- or *O*-. The locant *N*- is used for the *N*-glycosyl linkage to asparagine. *N*-Linked oligosaccharides are divided into two major classes: the *N-acetyl-lactosamine type* containing *N*-acetyl-D-glucosamine, D-mannose, D-galactose, L-fucose and sialic acid; and the *oligomannose type* containing *N*-acetyl-D-glucosamine and a variable number of D-mannose residues. Structures containing both oligomannose- and *N*-acetyl-lactosamine-type oligosaccharides are designated as *hybrid type*. Examples of *N*-glycoproteins (or *N*-glycosylproteins) are chicken ovalbumin, pig ribonuclease, human  $\alpha$ <sub>1</sub>-acid glycoprotein and soybean agglutinin.

The locant *O*- is used for *O*-glycosyl linkage to serine, threonine, hydroxylysine or hydroxyproline. Sheep submaxillary glycoprotein, collagen, fish antifreeze glycoproteins and potato lectin are *O*-glycoproteins (or *O*-glycosylproteins).

Two types of carbohydrate-peptide linkage in the same protein or peptide chain may be indicated by a combination of the locants. Thus, calf fetuin, procollagen, human erythrocyte membrane glycoporphin and human chorionic gonadotropin are *N,O*-glycoproteins (or *N,O*-glycosylproteins).

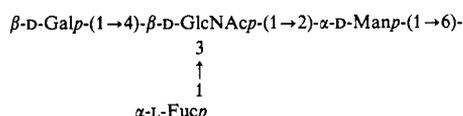
### 3.6 Asialoglycoproteins and asialo-galactoglycoproteins

Glycoproteins from which the sialic acid has been removed (by treatment with enzyme or mild acid) are designated by the prefix *asialo-*, e.g. *asialo- $\alpha_1$ -acid glycoprotein*, and *asialofetuin*. Removal of both sialic acid and galactose results in *asialo-galactoglycoproteins*.

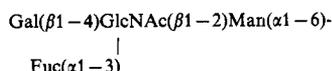
### 3.7. Condensed representation of sugar chains

For writing the structure of sugar chains, the non-reducing terminus of the carbohydrate chain should always be on the left-hand end (see ref. 6). Current practice allows the use of either an *extended form* (a) or a *condensed form* (b), which allows structures to be shown in one line as well as in two or more, and in which the longest chain should always be the main chain:

#### a) Extended form



#### b) Condensed form in two lines



or condensed form in one line

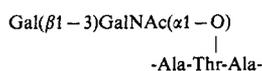


The condensed form is still unnecessarily long, however, and there should be no serious loss if it is shortened further by (i) omitting locants of anomeric carbon atoms, (ii) omitting the parentheses around the specification of linkage, and (iii) omitting hyphens if desired. We therefore suggest a more condensed or *short form* of writing (c):

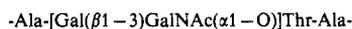
#### c) Short form



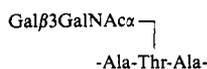
Similarly a glycopeptide sequence, represented in the condensed form in two lines as



or in the condensed form in one line as



may be written in the short form in two lines as



or in the short form in one line as

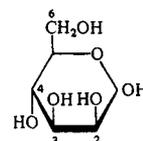


### 3.8. Representation of N-linked oligosaccharides

As a rule, N-linked oligosaccharides contain a common pentasaccharide core as follows:



For the sake of uniformity, the location of substitution should be written as above in accordance with Haworth's representation of the pyranose structure of monosaccharides, in analogy with the glycogen molecule.



### 3.9. Peptidoglycans

In peptidoglycans, peptide units of adjacent polysaccharides (glycosaminoglycans) may be cross-linked by a peptide bond between the C-terminal alanine residue of one peptide subunit and the  $\omega$ -amino group of the diamino acid residue of the other (e.g. L-lysine or *meso*-diaminopimelic acid), thereby giving rise to a giant macromolecule that forms the rigid cell wall ('sacculus'). This macromolecule is known to occur as a monomolecular layer between the inner and outer membrane in Gram-negative bacteria and as a multimolecular layer, often associated covalently or non-covalently with various additional compounds (teichoic acids, neutral polysaccharides, etc.) in Gram-positive bacteria.

Extensive investigations of peptidoglycans in thousands of bacterial strains demonstrated the existence of more than 100 chemotypes in the eubacteria [9]. The peptidoglycans of eubacteria have been classified into two major groups (A and B) and several subgroups according to the mode of cross-linkage. Within group A two main subgroups are recognized: one in which the C-terminal alanine residue is directly bound to the  $\omega$ -amino group of the diamino acid at position 3 of the peptide subunit of an adjacent polysaccharide (glycosaminoglycan) and one in which the cross-linkage is mediated by an interpeptide bridge consisting of either one or up to 5 amino acid residues (e.g. Lys-[Gly]<sub>5</sub>-type, Fig. 2). Depending on the kind of diamino acid at position 3 and the amino acids serving as interpeptide bridges, many variations of these subgroups of murein are known. In the peptidoglycan types of group B, the cross-linkage does not occur at position 3 of the peptide subunit but at position 2, utilizing the  $\alpha$ -carboxyl group of the D-glutamic residue. The interpeptide bridge must contain a diamino acid, which may be lysine, ornithine or diaminobutyric acid, in either L- or D-configuration. As a second special characteristic of the group B peptidoglycan types, the L-alanine residue at position 1 of the peptide subunit is replaced by either glycine or serine.

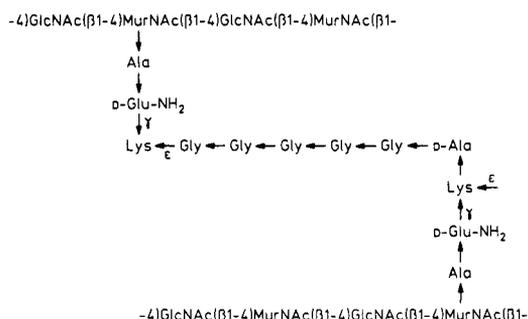


Fig. 2. Primary structure of a peptidoglycan, in which the cross-linkage between adjacent peptide-substituted polysaccharides is mediated by an interpeptide bridge consisting of Lys-[Gly]<sub>5</sub>

In the archaeobacteria, several organisms contain a peptidoglycan that differs in certain respects from those described above, typical for the eubacteria [10, 11]. It consists of a polysaccharide formed by alternating ( $\beta$ 1-3)-linked *N*-acetylated residues of D-glucosamine or D-galactosamine and ( $\beta$ 1-3)-linked *N*-acetylated residues of L-talosaminuronic acid, and a peptide containing exclusively L-amino acids attached to the carboxyl group of L-talosaminuronic acid. The peptide units of adjacent polysaccharides may be cross-linked by a peptide bond between the  $\gamma$ -carboxyl group of the glutamic acid of one peptide subunit and the  $\alpha$ -amino group of the lysine residue of the other, thus giving rise to the rigid cell wall of the methanogenic bacteria. This multimolecular layer is often associated covalently, but also non-covalently, with neutral polysaccharides. Only a few chemotypes of such peptidoglycans have so far been described.

#### 4. REFERENCES

1. IUPAC-IUB Commission on Biochemical Nomenclature (CBN), The nomenclature of lipids, Recommendations 1974, *Biochem. J.* 171, 21–35 (1978); *Eur. J. Biochem.* 79, 11–21 (1977); *Hoppe-Seyler's Z. Physiol. Chem.* 358, 617–631 (1977); *Lipids* 12, 455–468 (1977); ref. 3, pp. 122–132 (1978).
2. Joint Recommendations of the Physiological and Biochemical Committees on Protein Nomenclature, R. H. Chittenden, O. Folin, W. J. Gies, W. Koch, T. B. Osborne, P. A. Levene, J. A. Mandel, A. P. Mathews & L. B. Mendel (1908) *J. Biol. Chem.* 4, XLVIII–LI.
3. International Union of Biochemistry (1978) *Biochemical nomenclature and related documents*, The Biochemical Society, London.
4. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Conformational nomenclature for five and six-membered ring forms of monosaccharides and their derivatives, Recommendations 1980, *Arch. Biochem. Biophys.* 207, 469–472 (1981); *Eur. J. Biochem.* 111, 295–298 (1980); *Pure Appl. Chem.* 53, 1901–1905 (1981).
5. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Nomenclature of unsaturated monosaccharides, Recommendations 1980, *Eur. J. Biochem.* 119, 1–3 (1981) and 125, 1 (1982); *Pure Appl. Chem.* 54, 207–210 (1982).
6. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Abbreviated terminology of oligosaccharide chains, Recommendations 1980, *Arch. Biochem. Biophys.* 220, 325–329 (1983); *Eur. J. Biochem.* 126, 433–437 (1982); *J. Biol. Chem.* 257, 3347–3351 (1982) *Pure Appl. Chem.* 54, 1517–1522 (1982).
7. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Symbols for specifying the conformation of polysaccharide chains, Recommendations 1981, *Eur. J. Biochem.* 131, 5–7 (1983); *Pure Appl. Chem.* 55, 1269–1272 (1983).
8. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN), Nomenclature and symbolism for amino acids and peptides, Recommendations 1983, *Biochem. J.* 219, 345–373 (1984); *Eur. J. Biochem.* 138, 9–37 (1984); *Pure Appl. Chem.* 56, 595–624 (1984).
9. Schleifer, K. H. & Kandler, O. (1972) *Bacteriol. Rev.* 36, 407–477.
10. Kandler, O. (1982) *Zbl. Bakt. Hyg. I. Abt. Orig. C3*, 149–160.
11. König, H., Kralik, R. & Kandler, O. (1982) *Zbl. Bakt. Hyg. I. Abt. Orig. C3*, 179–191.

#### 5. APPENDIX: IMPLIED CONFIGURATIONS AND RING SIZES

In the condensed system of symbols for sugar residues the common configuration and ring size (usually pyranose) are implied in the symbol. Thus, Gal denotes D-galactopyranose; Man, D-mannopyranose; Fuc, L-fucopyranose; GlcNAc, 2-acetamido-2-deoxy-D-glucopyranose or *N*-acetyl-D-glucosamine; Neu5Ac (which may be abbreviated to NeuAc), *N*-acetylneuraminic acid. The symbol Sia stands for sialic acid, a general term that can also be used when the exact structure is unknown.

Whenever the configuration or ring size is found to differ from the common one it must be indicated by using the appropriate symbols for the extended system. The configuration of amino acids is L unless otherwise noted. Although symbols such as Gal and Man are useful in representing oligosaccharide structures they should not be used in the text to represent monosaccharides.