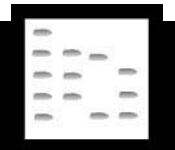
GLYCOPROTEINS: LIQUID CHROMATOGRAPHY



K. Miyazaki, Hokkaido University Hospital, Hokkaido, Japan

Introduction

Many proteins in cells and biological fluids are glycosylated, and these glycoproteins are present in animals, plants, microorganisms and viruses. The most commonly occurring monosaccharides found in oligosaccharide attachments to mammalian proteins are D-mannose (Man), D-galactose (Gal), D-glucose (Glu), L-fucose (Fuc), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), and *N*-acetylneuraminic acid (sialic acid or NeuAc).

The primary structure of glycoprotein glycans and their biological functions have been gradually unravelled by the improvements of methods for isolation and structure determination. High performance liquid chromatography (HPLC) is one of the most commonly used methods for the isolation and analysis of both glycoproteins and their derived carbohydrates, mainly due to the excellent resolution, ease of use, the generally high recoveries, the excellent reproducibility of repetitive separations, and the high productivity in terms of cost parameters.

Chemistry and Importance of the Glycan Chain of Glycoproteins

Basic Structure of Glycoprotein

In glycoproteins, glycans are conjugated to peptide chains by two types of primary covalent linkage, *N*-glycosyl and *O*-glycosyl. The former is called an *N*-linked sugar chain and contains a GlcNAc residue that is linked to the amide group of asparagine residues of a polypeptide. As shown in **Figure 1**, almost all *N*-linked glycoproteins have a common core of two GlcNAc and three Man residues. *N*-linked glycoproteins have three types of carbohydrate moieties: complex type (Figure 1), high mannose type and hybrid type. The hybrid type is a mixture of the complex and high mannose types. Complex type structures usually have from two to four branches attached to the two outer core Man residues. The branches are distributed over the two terminating core Man residues. The complex structures are termed diantennary, triantennary and tetraantennary, according to the number of antennae. The basic branch structures are composed in most instances of one GlcNAc and one Gal residue (Figure 1).

O-Glycosyl glycoproteins contain at their reducing end a GalNAc residue that is linked to the hydroxyl group of either serine or threonine residues of a polypeptide. This linkage is called an *O*-linked or mucin-type sugar chain. In general, *O*-linked structures appear to be less complex than *N*-linkages in terms of the number of antennae and monosaccharides. However, they can be fucosylated and sialylated. Some glycoproteins have both the *N*-linked and *O*-linked forms in their molecules (*N*, *O*-glycoproteins).

The addition of carbohydrate to a peptide chain changes the shape and size of the protein structure. Several important discoveries have revealed the following biological roles of glycans: (i) protection of polypeptide chains against proteolytic enzymes; (ii) influence on heat stability, solubility, and many physicochemical properties; and (iii) interaction with other proteins or nonprotein components of the cell, including control of the lifetime of circulating glycoproteins and cells.

Microheterogeneity of Glycans

In addition to genetically determined variants expressed as variations in their polypeptide chains, almost all glycoproteins exhibit polymorphism associated with their glycan moieties. This type of diversity is termed microheterogeneity, and these different forms have recently been called glycoforms. These variants were first characterized in the α_1-acid glycoprotein (AAG) from human serum using electrophoresis. As shown in the structure of major oligosaccharides of AAG (Figure 1), microheterogeneity was found to be due to the occurrence of di-, tri-, and tetraantennary glycans of the *N*-acetyllactosamine type at the five glycosylation sites.

This feature is widespread and has been observed in natural as well as in recombinant DNA glycoproteins. The existence of microheterogeneity gives rise to many interesting questions regarding the origin of this phenomenon and its relevance to the biological functioning of the glycoproteins that can be distinguished.

Recent interest has been shown in glycoproteins in the industrial field of genetic engineering of human glycoproteins of therapeutic interest. This gives rise to

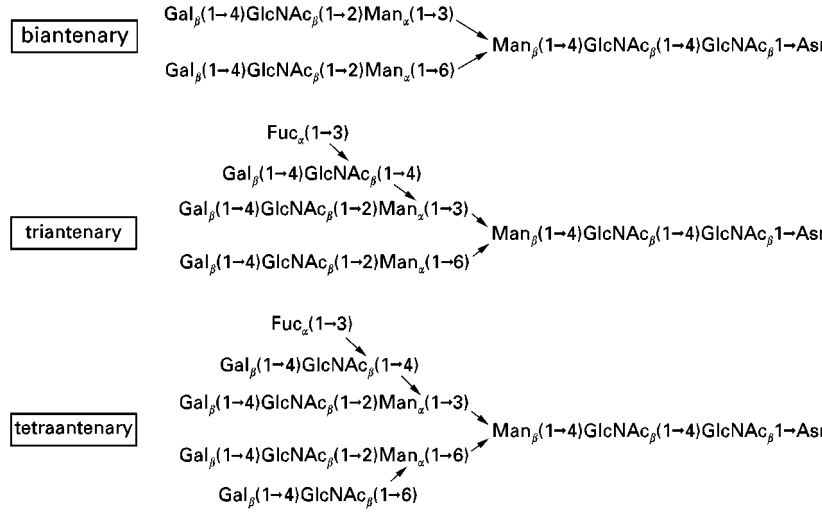


Figure 1 Structure of the major oligosaccharides of α_1-acid glycoprotein (a complex type of the *N*-linked form). Several NeuAc are linked to Gal residues.

an enormous problem, because the production of recombinant human glycoproteins in nonhuman eukaryotic cells or in prokaryotic cells devoid of glycan biosynthesis machinery leads to the production of incorrectly glycosylated proteins. Incorrectly glycosylated glycoproteins may have an undesirable effect on therapeutic effectiveness and safety due to changes in the properties of the products, including a decrease in the stability against heat or protease, shortening of the *in vivo* life span of the molecules by an increase in clearance, a decrease in the affinity for specific receptors, and an increase in antigenicity.

Isolation and Quantitation of Glycoprotein Molecules and Analysis of Glycan Chains

Determination of the primary structure of glycoproteins necessitates analysis of the protein sequence, identification of the glycosylation sites, unravelling of the glycan structures and determination of the microheterogeneity of the glycans at each glycosylation. For these studies, it is essential that adequate starting materials are available.

For the isolation or purification of glycoproteins, a combination of several complementary separation methods such as gel permeation chromatography, affinity chromatography (lectin or others), anion- or cation-column chromatography, high performance capillary electrophoresis, and HPLC using several sorbents is generally used.

Determination of the carbohydrate composition, type and branching pattern is an important step for understanding the biological function of the native glycoprotein molecules as well as for the development of a recombinant DNA-derived glycoprotein as a pharmaceutical agent. However, the composition, type and branching pattern of carbohydrates are complex due to the diversity of monosaccharides and the variety of possible linkages. Unlike amino acids, which are linked through an amido bond, monosaccharides are joined through a variety of hydroxyl groups present on the sugar to form glycosidic linkages. For example, two different amino acids can form only two dipeptides, while two different monosaccharides can lead to more than 60 disaccharides. However, the availability of improved and sophisticated methods for the isolation and characterization of glycoproteins and their derived glycans has paved the way for the analysis and characterization of the carbohydrate chains of glycoprotein. These analytical methods include mass spectrometry, enzymatic microsequencing, nuclear magnetic resonance (NMR), capillary electrophoresis, reversed-phase HPLC (RP-HPLC), and high pH anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD), and generally a combination of several complementary analytical methods is needed to determine the carbohydrate structure.

In this section, we give an outline of the recently developed HPLC procedures for the purification, separation, and determination of glycoproteins and their glycoforms.

Examples of Isolation or Purification of Glycoproteins by Using HPLC

α_1-Acid glycoprotein (AAG) AAG is a characteristic and dominant fraction of human serum sialoglycoproteins with a molecular mass approximately 44 000

Da, an unusually high carbohydrate content (45%) and a large number of sialyl residues. Although its exact biological function is still unknown, AAG is an acute-phase reactant that increases following cancer, myocardial infarction, and congestive heart failure and has also been reported to play an important role in immunoregulation.

Ion exchange chromatography and gel permeation chromatography previously used for purification are time-consuming and require a large volume of plasma or serum because of the low quantities recovered. These methods also lead to a strong possibility of denaturation and desialylation. Moreover, separation of AAG and α_1-antitrypsin has been difficult, because the chromatographic behaviour of these compounds during anion exchange chromatography is similar. Recently, these problems have been overcome by the introduction of an HPLC system equipped with a hydroxyapatite column as the last step after the clean-up procedures with commercially available cartridge ion exchange columns.

Ceruloplasmin (CP) CP is a serum α_2-glycoprotein that carries more than 95% of the copper present in plasma and is believed to have an active role in the regulation of copper and iron homeostasis. It has been pointed out that fragmentation of CP during purification and storage has hampered the study of its structure. The rapid degradation of purified CP reported by many laboratories may be largely due to the presence of one or more copurifying or contaminating proteases, at least one of which is a metalloproteinase.

Recently, a highly purified and nonlabile CP has been obtained from human plasma by combining the previously reported chromatographic steps with additional gel permeation and fast protein liquid chromatography (FPLC) steps. In the latter steps, further purification of CP by Sephadex G-50 chromatography and Mono Q FPLC were essential for the removal of plasma metalloproteinase, and this purification procedure yielded a protein that was completely stable even after incubation at 37°C for 4 weeks.

Erythropoietin (EPO) EPO, an acidic glycoprotein hormone, is synthesized in the kidney and circulates in the blood to stimulate red cell proliferation and differentiation in bone marrow. Native human EPO was first purified from the urine of patients suffering from severe aplastic anaemia. Since then, several methods for the purification of urinary human EPO (uHuEPO) have been developed. RP-HPLC has recently been used for the purification of uHuEPO with high *in vivo* activity. This purification procedure involves two membrane filtration steps, Sephadex G-25, two DEAE-agarose steps, Sephadex G-75, wheat germ agglutinin (WGA)-agarose, and RP-HPLC. The final HPLC step is essential for the removal of nucleic acids.

Chromatographic Determination of AAG

There have been very few reports on chromatographic methods for determining concentrations of glycoproteins other than AAG in biological samples. Radial immunodiffusion (RID) utilizing the antibody against AAG has been widely used to determine AAG in serum because of its strong specificity. This method, however, is time-consuming and is not easily applicable to experimental animals. To overcome these problems, some simple and rapid HPLC methods have been developed.

After pretreatment of human serum with a chloroform/methanol mixture (2 : 1, v/v), 500 μL of the upper phase was applied to the anion exchange FPLC system (Mono Q HR 5/5 column), and AAG was eluted with a pH/NaCl gradient elution programme. To measure the serum AAG content, the Mono Q HR column was calibrated with commercial AAG in the range 100–200 μg/500 μL of sample volume.

A rapid and sensitive determination method starting from the diluted human serum itself has also been reported. This procedure involves the anion exchange step for cleaning up serum (commercially available cartridge column, DEAE-M) and a hydroxyapatite HPLC system. A linear relationship between standard AAG concentration and peak height was observed over the concentration range 0.5–2.5 mg mL^{-1} serum. The coefficient of variation at 0.5 mg mL^{-1} AAG was 3.7% ($n = 8$). A good correlation was observed between this HPLC method (y) and the conventional RID (x) ($y = 1.009x + 0.004$, $r = 0.996$).

Determining the Carbohydrate Composition, Type and Branching Pattern

RP-HPLC has become a commonly used method for the analysis and purification of peptides, proteins and glycoproteins. The RP-HPLC experimental system usually comprises an n-alkylsilica-based sorbent. By using modern instrumentation and columns, complex mixtures of peptides and proteins can be separated and low picomolar amounts of resolved components can be collected. Separation can be easily performed by changing the gradient slope of solvents such as acetonitrile containing an ionic modifier (e.g. trifluoroacetic acid (TFA)); column temperature; or the organic solvent composition. The technique is equally applicable to the analysis of enzymatically derived

mixtures of peptides from proteins and glycoproteins. Separated fractions can be subsequently subjected to further analysis of carbohydrates and amino acids.

A new HPLC method, HPAEC-PAD, which bypasses the derivatization steps by using pulsed electrochemical detection on gold electrodes, has been developed. Monosaccharides and oligosaccharides can be directly resolved by anion exchange chromatography, because the hydroxyl groups of carbohydrates are weakly acidic and reveal anionic forms at pH values greater than pH 12. In addition to high sensitivity in the low picomole range of PAD, a major advantage of HPAEC-PAD is its usefulness in analysing both monosaccharides and all classes of oligosaccharides without derivatization. HPAEC-PAD has therefore been used successfully for resolving and quantitating the constituent monosaccharides released by acidic hydrolysis (e.g. TFA) of glycan chains and for resolving *N*-linked oligosaccharides separated by enzyme digestion (e.g. PNGase F).

Figure 2 shows the HPAEC-PAD chromatograms of fractions 23–27 from the RP-HPLC separation of a tryptic digest of recombinant tissue plasminogen activator (tPA). The peaks from RP-HPLC separation were collected manually, and aliquots of all 62 fractions were analysed for neutral and amino monosaccharides after acid hydrolysis. The chromatograms in Figure 2 indicate that fractions 24–26 contain glycopeptides, and the ratio of constituent monosaccharides suggests that their oligosaccharide structures are those of fucosylated *N*-acetyllactosamine-type oligosaccharides. Another *N*-acetyllactosamine-type chain and an oligomannose-type chain were identified similarly by the same analytical procedure.

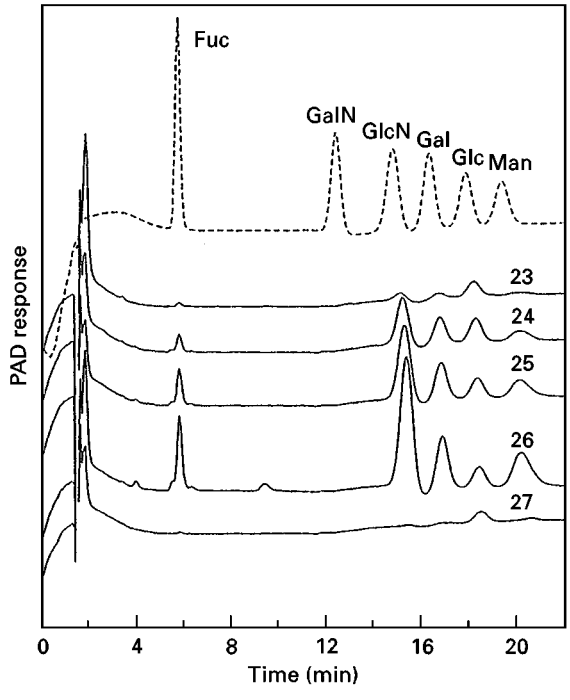


Figure 2 Monosaccharide analysis of fractions 23–27 from RP-HPLC separation of a tryptic digest of recombinant tissue plasminogen activator. Elution positions of monosaccharide standards are indicated on the upper trace. (Reproduced with permission from Townsend *et al.*, 1996.)

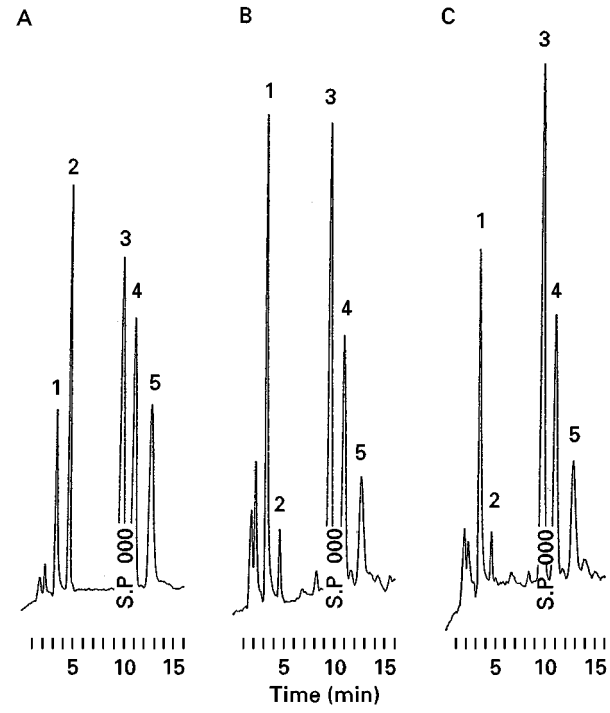


Figure 3 Representative chromatograms of monosaccharides after treatment of standard and plasma samples. (A) Standard sample containing 5.0 μg mL^{-1} of each monosaccharide; (B) healthy subject; (C) patient with renal insufficiency. Peaks: 1 = mannitol (internal standard); 2 = Fuc; 3 = GlcNAc; 4 = Gal; 5 = Man. (Reproduced with permission from Kishino *et al.*, 1995.)

The HPAEC-PAD method is also applicable to the quantitation of concentrations of monosaccharides after release by acid hydrolysis and following clean-up procedures with commercially available cartridge columns. **Figure 3** shows the chromatograms of four monosaccharides in purified serum AAG from healthy subjects and from patients with renal insufficiency. The concentration of NeuAc can also be determined under different solvent conditions. This method enables composition analysis of the carbohydrate moiety of AAG with only 1.0 mL of plasma. Linear relations between the amount of NeuAc or monosaccharides and the peak-height ratio of NeuAc or monosaccharides to the internal standards are observed over the concentration range of 5.0 to 100 μg mL^{-1}. *N*-Glycolylneuraminic acid and mannitol are used as the internal standard for NeuAc and the four monosaccharides, respectively.

Table 1 Analysis of NeuAc and monosaccharide levels in purified α_1-acid glycoprotein (AAG) from plasma of healthy subjects, patients with renal insufficiency and patients with myocardial infarction

	Healthy subjects ($n = 8$)	*Renal insufficiency* ($n = 6$)	*Myocardial infarction* ($n = 4$)
Age (years)	65.8 ± 12.5	60.2 ± 7.2	69.7 ± 11.7
AAG conc. (mg mL^{-1})	0.79 ± 0.14	1.01 ± 0.19^a	1.66 ± 0.34^b
NeuAc (mg $g^{-1} \cdot$ AAG)	82.77 ± 12.55	89.29 ± 15.55	132.70 ± 6.89^b
Fuc (mg $g^{-1} \cdot$ AAG)	9.84 ± 3.08	12.79 ± 3.37	10.40 ± 3.42
GluNAc (mg $g^{-1} \cdot$ AAG)	113.01 ± 10.07	135.44 ± 10.51^b	142.68 ± 12.96^b
Gal (mg $g^{-1} \cdot$ AAG)	76.97 ± 4.23	88.78 ± 5.61^b	86.92 ± 11.73^a
Man (mg $g^{-1} \cdot$ AAG)	49.14 ± 2.57	57.26 ± 3.80^b	50.16 ± 1.67
GluNAc/Man	2.18 ± 0.37	2.50 ± 0.35	2.84 ± 0.21^a

Source: Kishino *et al.* (1995).
AAG concentration was determined by the HPLC method with a hydroxyapatite column.
NeuAc and each monosaccharide concentration were determined by HPAEC-PAD.
Values in the table are means $\pm$ SD.
[a]Significantly different ($p < 0.05$) from healthy subjects.
[b]Significantly different ($p < 0.01$) from healthy subjects.

The resultant quantitation data (**Table 1**) for healthy subjects, patients with renal insufficiency and patients with myocardial infarction show that not only AAG levels but also the concentrations of several monosaccharides in patients increased significantly compared to those of healthy subjects, suggesting a change in the carbohydrate branching pattern in such pathologic conditions.

It is well known that the microheterogeneity of AAG is due to the occurrence of di-, tri-, and tetra-antennary glycans of the *N*-acetyllactosamine type at the five glycosylation sites. Moreover, the Man content is constant among the antennary glycans, and the number of branches increases with the addition of GlcNAc to Man residues. A highly branched glycan chain of AAG is constructed by the linkage of Gal to GlcNAc (Figure 1), which results in the formation of an antennary structure (*N*-acetyllactosamine). Therefore, in the case of AAG, determination of the concentration ratio of GlcNAc to Man (GlcNAc/Man) is important for estimating whether the carbohydrate moiety of glycoforms has a highly or less-branched glycan chain. The significantly higher GlcNAc/Man ratio in the patients with myocardial infarction suggests that a highly branched glycan chain was synthesized. Changes in the carbohydrate moiety in the glycoproteins have been reported in patients with various types of disease.

As shown in **Figure 4**, at least six fractions, which are possibly based on carbohydrate-mediated microheterogeneity, have been obtained from healthy human (Japanese) serum AAG by HPLC using a hydroxyapatite column under a gradient elution programme. From the determination of five monosaccharides (NeuAc, Fuc, GlcNAc, Gal, Man) in each fraction by HPAEC-PAD, it was found that glycoforms rich in carbohydrates were eluted later (fractions 4, 5, 6) and that NeuAc was relatively abundant in these highly adsorbed glycoforms, especially in fraction 6. Furthermore, the ratio of GlcNAc/Man in fraction 2 was significantly higher than those in the other fractions, suggesting the presence of a highly branched glycan chain. Interestingly, it has been also shown that fractions 1 and 2, both relatively rich in highly branched glycan chains, showed a significantly lower binding capacity to disopyramide, a drug for the treatment of arrhythmia, than did the other fractions. This result suggests that the binding sites of AAG to disopyramide are hindered by relatively large carbohydrate moieties, such as a tetraantennary structure. These results are consistent with the findings that the binding capacity of purified AAG isolated from patients (with renal insufficiency or myocardial infarction) to disopyramide is significantly lower than that of healthy subjects and that the AAGs of these patients revealed a higher concentration ratio of GlcNAc/Man, an index of the abundance of highly branched glycan chains.

In conclusion, in order to gain further insight into the structure–function relations of the carbohydrate moiety, it is essential that sufficient quantities of glycoprotein variants are available. An effective combination of the sophisticated separation methods for glycoforms, such as the HPLC system using a hydroxyapatite column, and the qualitative and quantitative analytical methods for monosaccharides/oligosaccharides, such as HPAEC-PAD, must be established.

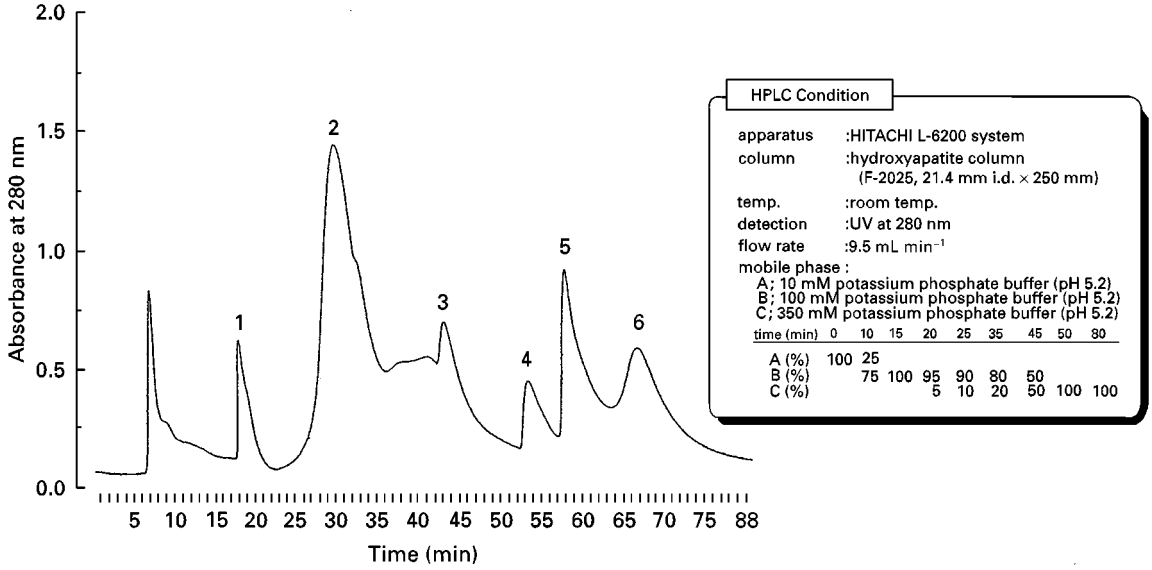


Figure 4 Typical chromatograms of the glycoforms of α_1-acid glycoprotein (AAG) from the serum of healthy subjects by HPLC. Inlet is the gradient programme for the fractionation of the glycoforms of AAG. (Sampling time of each fraction: fraction 1, 17–22 min; 2, 27–36 min; 3, 43–50 min; 4, 53–57 min; 5, 58–62 min and 6, 65–72 min.) (Reproduced with permission from Kishino *et al.*, 1997.)

See also: **III/Carbohydrates:** Liquid Chromatography. **Peptides and Proteins:** Liquid Chromatography. **Polysaccharides:** Liquid Chromatography.

Further Reading

Clemetson KJ (1997) In: Montreuil J, Vliegenthart JFG and Schachter H (eds) *Glycoproteins II*, pp. 173–201. Amsterdam: Elsevier.

Hancock WS, Chakel AAJ, Souders C, M'Timkulu T, Pungor E Jr and Guzzetta AW (1996) In: Karger BL and Hancock WS (eds) *Methods in Enzymology*, vol. 271, pp. 403–427. San Diego: Academic Press.

Hardy MR and Townsend RR (1994) In: Lennarz WJ and Hart GW (eds) *Methods in Enzymology*, vol. 230, pp. 208–225. San Diego: Academic Press.

Kishino S, Nomura A, Sugawara M, Iseki K, Kakinoki S, Kitabatake A and Miyazaki K (1995) *Journal of Chromatography* 672: 199–205.

Kishino S and Miyazaki K (1997) *Journal of Chromatography* 699: 371–381.

Kishino S, Nomura A, Saitoh M, Sugawara M, Iseki K, Kitabatake A and Miyazaki K (1997) *Journal of Chromatography* 703: 1–6.

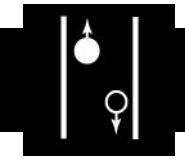
Montreuil J (1995) In: Montreuil J, Schachter H and Vliegenthart JFG (eds) *Glycoproteins*, pp. 1–12. Amsterdam: Elsevier.

Schmid K (1989) In: Bauman P, Eap CB, Muler WE and Tillement J-P (eds) *Alpha$_1$-Acid Glycoprotein*, pp. 7–22. New York: Alan R. Liss.

Townsend RR, Basa LJ and Spellman MW (1996) In: Karger BL and Hancock WS (eds) *Methods in Enzymology*, vol. 271, pp. 135–147. San Diego: Academic Press.

Vliegenthart JFG and Montreuil J (1995) In: Montreuil J, Schachter H and Vliegenthart JFG (eds) *Glycoproteins*, pp. 13–28. Amsterdam: Elsevier.

GOLD RECOVERY: FLOTATION



S. Bulatovic and D. M. Wyslouzil, Lakefield Research, Lakefield, Ontario, Canada

Introduction

The recovery of gold from gold-bearing ores depends largely on the nature of the deposit, the mineralogy of the ore and the distribution of gold in the ore. The methods used for the recovery of gold consist of the following unit operations:

- The gravity preconcentration method, which is mainly used for recovery of gold from placer deposits that contain coarse native gold. Gravity is