- Kabat, E. A., Basset, E. W., Pryzwansky, K., Lloyd, K. O., Kaplan, M. E. & Layug, E. J. (1965) *Biochemistry* **4**, 1632
- Kobata, A., Grollman, E. F. & Ginsburg, V. (1968a) *Arch. Biochem. Biophys.* **124**, 609
- Kobata, A., Grollman, E. F. & Ginsburg, V. (1968b) *Biochem. Biophys. Res. Commun.* **32**, 272
- Lloyd, K. O., Kabat, E. A., Layug, E. J. & Gruezo, F. (1966) *Biochemistry* **5**, 1489
- Lloyd, K. O., Kabat, E. A. & Licerio, E. (1968) *Biochemistry* **7**, 2976
- Marr, A. M. S., Donald, A. S. R., Watkins, W. M. & Morgan, W. T. J. (1967) *Nature (London)* **215**, 1345
- Morgan, W. T. J. (1970) in *British Biochemistry Past and Present* (Goodwin, T. W., ed.), p. 99, Academic Press, London
- Painter, T. J., Watkins, W. M. & Morgan, W. T. J. (1965) *Nature (London)* **206**, 594
- Race, C., Zideman, D. & Watkins, W. M. (1968) *Biochem. J.* **107**, 733
- Race, R. R. & Sanger, R. (1968) *Blood Groups in Man*, 5th. edn., Blackwells, Oxford
- Rege, V. P., Painter, T. J., Watkins, W. M. & Morgan, W. T. J. (1964a) *Nature (London)* **203**, 360
- Rege, V. P., Painter, T. J., Watkins, W. M. & Morgan, W. T. J. (1964b) *Nature (London)* **204**, 740
- Sawicka, T. (1971) *FEBS Lett.* **16**, 364
- Shen, L., Grollman, E. F. & Ginsburg, V. (1968) *Proc. Nat. Acad. Sci. U.S.* **59**, 224
- Tuppy, H. & Schenkel-Brunner, H. (1969) *Eur. J. Biochem.* **10**, 152
- Vicari, G. & Kabat, E. A. (1969) *J. Immunol.* **102**, 821
- Watkins, W. M. (1966) *Glycoproteins* 462
- Watkins, W. M. & Morgan, W. T. J. (1959) *Vox Sang.* **4**, 97

Structural Analysis of Glycoproteins

By J. R. CLAMP (*Department of Medicine, University of Bristol, Bristol BS8 1TD, U.K.*)

Although there has been little fundamental advance in recent years in methods for investigating the structure of the oligosaccharide units of glycoproteins, there has been some improvement in analytical techniques. Automated column procedures and g.l.c. have both been used for this purpose, and, as our experience has been with the latter technique, this will be used to illustrate the application of structural methods.

Two main procedures are in use for the analysis of the carbohydrate content of glycoproteins by g.l.c., differing in the method of cleavage of the polymeric material. The first of these is based on hydrolysis with aqueous acid followed by reduction of the liberated aldoses to the glycitols. The second depends on methanolysis, yielding methyl glycosides together with the methyl esters of carboxyl groups as in hexuronic acids and sialic acids. The final products in each procedure, namely glycitols or methyl glycosides, are then chromatographed as the *O*-trimethyl-

silyl ethers. Both systems have certain advantages, although our experience favours methanolysis. Thus it appears to be more efficient than hydrolysis in cleaving glycosidic linkages and the products are probably more stable (Chambers & Clamp, 1971). In addition, each monosaccharide unit after methanolysis gives a series of methyl glycosides, thereby producing a characteristic pattern of peaks on g.l.c., and this makes the identification of each monosaccharide more certain. This multiple-peak pattern has an additional advantage when, as is common with biological samples, there is some contamination and a background peak coincides with one of the isomer peaks. In this case the amount of monosaccharide can still be calculated from the unobserved peaks.

By using this technique, all the monosaccharide units commonly occurring in glycoproteins can be separated, identified and determined in a single procedure. Such monosaccharide units include fucose, mannose, galactose, glucose, *N*-acetylglucosamine, *N*-acetylgalactosamine and sialic acid. In addition, the system can cope with pentoses, hexuronic acids, glycitols and most other monosaccharides and derivatives, and even with those amino acids likely to occur in glycopeptides (Bhatti & Clamp, 1971).

G.l.c. can also be used to considerable advantage as an analytical procedure in conjunction with classical methods such as methylation, periodate oxidation, partial acid hydrolysis and so on. These methods are usually carried out on glycopeptides isolated from the glycoproteins. Methylation studies give information on positions of substitution of monosaccharides. A number of systems have been devised to deal with the *O*-methyl derivatives of those carbohydrate residues usually found in glycoproteins, in which the partially methylated residues are usually separated as mixed derivatives (Clamp *et al.*, 1971).

The technique of periodate oxidation also provides information on positions of substitution, although the information is less precise than that given by methylation. Despite this, the technique is used more often, probably because it is simpler and uses less material, and because the extent and completeness of the reaction can be assessed by determining the amount of unchanged periodate. In this way the carbohydrate content of the sample can be determined after periodate consumption has levelled off, with particular interest in hexose residues resistant to, and *N*-acetylhexosamine residues susceptible to, oxidation. Three unsubstituted vicinal hydroxyl groups yield formic acid, and in mammalian glycoproteins this would be likely to arise from terminal or 6-*O*-substituted hexoses or from *N*-acetylneuraminic acid. Periodate oxidation followed by reduction yields a number of short-chain aldehydes and alcohols. For example, threitol arises from 4-*O*-substituted galactose, whereas the corresponding mannose deri-

vative gives erythritol. An extension of this technique is that of Smith degradation (Abdel-Akher *et al.*, 1952; Smith & Unrau, 1959; Goldstein *et al.*, 1959), which takes advantage of the acid-lability of the acetal linkages of the resultant short-chain alcohols. After mild hydrolysis, the residual carbohydrate-containing material is subjected to a further cycle of oxidation and reduction. The products after these periodate oxidation procedures can be determined by g.l.c. (Hughes & Clamp, 1972).

A technique increasingly used in the structural investigation of the oligosaccharide units of glycoproteins is that of incubation with specific exoglycosidases. Most of the enzymes of interest in this field have been obtained free from contaminating activity. Such enzymes include galactosidase, mannosidase, *N*-acetylglucosaminidase and neuraminidase. Used sequentially and with adequate controls they provide evidence for the order of monosaccharide residues in the oligosaccharide units. As little as 1 nmol of released monosaccharide can be dealt with by g.l.c. (Bhatti *et al.*, 1970).

Disaccharides and higher oligosaccharides arise from the controlled hydrolysis of glycoproteins (Montreuil & Chosson, 1962; Chosson *et al.*, 1962). G.l.c. can handle a variety of closely related isomeric disaccharides, such as those containing only glucose (Bhatti *et al.*, 1970), and should therefore be able to handle the range of disaccharides arising from partial acid hydrolysis. Preliminary studies have confirmed this with a series of characterized oligosaccharides ranging from disaccharides to pentasaccharides obtained by the partial acid hydrolysis of glycoproteins (J. R. Clamp & J. Montreuil, unpublished work). The linkage region product consisting of *N*-acetylglucosamine linked to asparagine can also be handled in this system (Bhatti *et al.*, 1970). Once again the advantage of a g.l.c. method of analysis linked to such studies is not only the sensitivity, but the fact that the products can be both identified and determined in the same procedure.

The disadvantages of partial acid hydrolysis are the poor yields of suitable products and the non-specific nature of the cleavage reaction. It would be a considerable improvement if a method could be devised to cleave specifically at a certain type of monosaccharide irrespective of substitution and irrespective of the position and type of residue to which it was attached. *N*-Acetylhexosamine is the monosaccharide most likely to fulfil this role. An enzymic method would obviously be most satisfactory, and a number of endo-*N*-acetylglucosaminidases are known, most of which act on linear polymers containing repeating groups. The oligosaccharide units of glycoproteins are usually highly branched and appear to be degraded by exo-enzymes acting on terminal residues. On investigation, therefore, the endo-enzymes do not appear to fulfil the above requirements for such

units (J. R. Clamp, unpublished work). At the present time, chemical methods seem to offer greater possibilities. The de-*N*-acetylation of *N*-acetylglucosamine residues alters the characteristics of the adjacent glycosidic linkages. The resultant amino groups would be in the protonated form under acid conditions and this would render the linkages much more resistant to hydrolysis. On the other hand, treatment with nitrous acid is a mild procedure for the specific cleavage of these linkages (Isemura & Schmid, 1971). There is thus the possibility of obtaining two sets of overlapping fragments. The difficulty in this procedure is the initial specific removal of the *N*-acetyl groups. This has led to the investigation of alternative methods of cleavage at *N*-acetylhexosamine residues, such as the *N*-nitrosoamide decomposition reaction (Llewellyn & Williams, 1971).

Mass spectrometry is a technique that will be used increasingly as experience is gained in this field. At present, it can give information on positions of linkage and order of monosaccharide residues in oligosaccharides (Kochetkov & Chizhov, 1966; Hanessian, 1971).

This work was generously supported by grants from the Medical Research Council.

- Abdel-Akher, M., Hamilton, J. K., Montgomery, R. & Smith, F. (1952) *J. Amer. Chem. Soc.* **74**, 4970
Bhatti, T. & Clamp, J. R. (1971) *Biochim. Biophys. Acta* **229**, 293
Bhatti, T., Chambers, R. E. & Clamp, J. R. (1970) *Biochim. Biophys. Acta* **222**, 339
Chambers, R. E. & Clamp, J. R. (1971) *Biochem. J.* **125**, 1009
Chosson, A., Montreuil, J. & Scheppeler, N. (1962) *C. R. Acad. Sci. Ser. D* **255**, 3261
Clamp, J. R., Bhatti, T. & Chambers, R. E. (1971) *Methods Biochem. Anal.* **19**, 229
Goldstein, I. J., Smith, F. & Unrau, A. M. (1959) *Chem. Ind. (London)* 124
Hanessian, S. (1971) *Methods Biochem. Anal.* **19**, 105
Hughes, K. W. & Clamp, J. R. (1972) *Biochim. Biophys. Acta* in the press
Isemura, M. & Schmid, K. (1971) *Biochem. J.* **124**, 591
Kochetkov, N. K. & Chizhov, O. S. (1966) *Advan. Carbohydr. Chem.* **21**, 39
Llewellyn, J. W. & Williams, J. M. (1971) *Chem. Commun.* 1386
Montreuil, J. & Chosson, A. (1962) *C. R. Acad. Sci. Ser. D* **255**, 3071
Smith, F. & Unrau, A. M. (1959) *Chem. Ind. (London)* 881