Pure & Appl. Chem., Vol. 65, No. 4, pp. 787–791, 1993.
Printed in Great Britain.
© 1993 IUPAC

Dissection of heparin—past and future

H. E. Conrad

Glycomed, Incorporated, Alameda, California, USA

Abstract

When compared to other complex carbohydrate polymers, heparin and heparan sulfate present unique challenges for structural characterization. Special finesse must be used to determine their mono- or disaccharide compositions, their sulfate contents, and their disaccharide sequences. Approaches that have been developed for addressing these difficult problems include selective enzyme cleavage, selective nitrous acid cleavage with or without N-deacetylation of the GlcNAc's, reduction of the GlcA and IdoA residues to labilize their glycosidic bonds, periodate oxidation to allow selective cleavage at the unsulfated uronic acid residues, and ion exchange and reversed phase ion pairing HPLC to resolve complex mixtures of di-, tetra- and hexasaccharides. These techniques have been useful in determining the structure-activity relationships for the anticoagulant activity of heparin. However, heparin and heparan sulfate have many additional biological activities, and, for the most part, the specific sequences responsible for these activities are unknown. In the future, the isolation of specific heparin sequences will yield only small amounts of specific heparin sequences. The application of the structural methodologies described to these future changes is discussed.

BIOLOGICAL ACTIVITIES OF HEPARIN

Heparin has been used clinically for many years as an anticoagulant drug. Now-classical work from the laboratories of Lindahl (ref. 1) and Rosenberg (ref. 2) demonstrated that a rare pentasaccharide sequence in heparin bound specifically to antithrombin III and mediated the anticoagulant activity. Biologists have now become increasingly aware that heparin has a variety of other biological activities. Just as the anticoagulant activity of heparin is mediated primarily through its specific binding to antithrombin III, the other activities of heparin appear to be mediated through its tight binding to a variety of "heparin-binding proteins". For most of these heparin-binding proteins, there has been no progress in identifying specific sequences in heparin that mediate binding. Although there are a number of reasons for this lack of progress, one of the significant problems has been the difficulty in analysis of the complex structural features of heparin and heparan sulfate. Here the reasons for these structural difficulties are reviewed and a systematic approach for dealing with the complex structures of heparin and heparan sulfate (referred to here as "heparinoids") is described. There are other structural approaches that are equally valid, but the present review will be confined to the approaches which we have developed in our own laboratories.

STRUCTURE OF HEPARINOIDS

The structures of heparin and heparan sulfate are most readily perceived through their biosynthetic pathways that were elucidated in the elegant work from the laboratories of Lindahl, Rodén, and Feingold (ref. 3). The initial stage of synthesis involves the attachment of the basic repeating disaccharide sequence, [βGlcA1 → 4αGlcNAc1 → 4], to the core protein via a GlcA→Gal→Gal→Xyl→Ser linkage region. The $[GlcA \rightarrow GlcNAc]_n$ polymer sequence then undergoes a series of maturation reactions that is initiated by random N-deacetylation of about 50 percent of the GlcNAc residues followed by N-sulfation of these residues. Further maturation occurs around these $GlcNSO_3$ residues and involves further N-deacetylation/N-sulfation of near-by GlcNAc residues, epimerization of some of the β-D-GlcA residues to α-L-IdoA residues, sulfation of some of the IdoA and GlcA residues at C2 and some of the $GlcNSO_3$ and GlcNAc residues at C6, and finally 3-O-sulfation of some of the $GlcNSO_3$ residues. The final biosynthetic products contain about 10 different disaccharide units that are distinguished by their uronic acid components and their sulfation patterns. These disaccharides may be considered the monomeric units of the heparinoids. Although clear structural distinctions between heparin and heparan sulfate are not easily defined (discussed in ref. 4), the chains that are more enriched in IdoA and SO_4 are representive of heparin chains, whereas the less highly matured chains are more typical of heparan sulfates. In the mast cell, the

final step of "biosynthesis" of heparin appears to be a cleavage of the heparin chains internally by an endo-β-glucuronidase that reduces the chain size and removes most of the chains from the core protein. Since the initial N-deacetylation steps in the maturation sequence occur randomly, and since the maturation reactions proceed to varying degrees at individual disaccharide units, the final product is a mixture of chains with different disaccharide sequences and different degrees of maturation. Thus, there is no such thing as a "pure" heparin or heparan sulfate; and it is impossible to determine a primary sequence of disaccharides in heparin in the same sense that one can determine sequences of monomers in proteins and nucleic acids. Given the difficulty of extracting structural information from heparinoids, it is small consolation that one need not determine the nature of the glycosidic linkages for heparinoid preparations, since these are formed in the initial stages of biosynthesis and do not change during the maturation steps.

If one assumes that the various biological activities of heparinoids reside in relatively short oligosaccharide sequences, it is easy to imagine that, although heparin may have the ability to bind to many different proteins, each protein may recognize a different, though specific, sequence. The primary challenge in characterizing biologically important sequences from these polymers is thus the isolation of pure fragments that contain the desired sequence (see below). However, even after such pure sequences are obtained, the structural analyses present a number of difficulties. It is even difficult to establish basic structural distinctions between different heparinoid preparations. Simple questions include: (a) what is the relative percentage of GlcA and IdoA? (b) what percentages of the GlcN residues are N-acetylated and what percent are N-sulfated? and (c) what is the relative degree of sulfation, and where are the sulfates located? Typically, a carbohydrate structural chemist would simply hydrolyze the polymer and use one of a number of the available analytical procedures for SO_4^{2-} and monosaccharide analyses. In the case of heparinoids, however, these basic steps are not easy. Acid hydrolysis results in the rapid loss of the sulfate substituents which help to distinguish the individual disaccharide units from each other. Also, it has been known for many years that the glycosidic bonds of glucuronic acid and galacturonic acid (and presumably all uronic acids) are extremely resistant to acid hydrolysis, and that under the forcing hydrolysis conditions that are required to release uronic acids, there is decomposition of these residues. Finally, the glycosidic bonds of GlcN residues, formed early in the hydrolysis of heparinoids by loss of N-acetyl or N-sulfate residues, are even more stable to acid than the glycosidic bonds of uronic acids. Thus, the structural characterization of heparinoids, even at the most basic levels, requires a great deal of finesse. The highly sophisticated spectroscopic methodologies (NMR, mass spectrometry) that are now so useful in other carbohydrate structure analyses are just beginning to emerge in the heparinoid area as pure, well-characterized oligosaccharides are becoming available for use as standards.

ANALYTICAL PROTOCOLS

In order to develop structural methods of use in the heparin field, it has been necessary to use a highly sensitive analytical procedure to assure the quality of the methodology. We adopted a procedure in which reducing sugars are reduced with NaB^3H_4 and resolved by chromatography or electrophoresis so that the individual components can be quantitated by the amount of ^3H present in each (ref. 5). The procedure has several advantages: (a) carbohydrates are unique among biological materials in being reducible by $NaBH_4$, (b) the molar response is identical for all reducing mono- or oligosaccharides, regardless of degree of polymerization, and (c) a variety of separation procedures - paper chromatography or paper electrophoresis, HPLC, GC, etc. can be used to resolve the ^3H-labeled carbohydrates, with either on-line or off-line quantitation of the ^3H.

THE NITROUS ACID REACTION

The reaction of amines with nitrous acid is a classical reaction. It was recognized early (ref. 6) that the glycosidic bonds of amino sugars having unsubstituted amino groups are cleaved by nitrous acid. Thus, since hydrolysis of heparin in 1 N acid at 100 °C removes all N-sulfate groups almost instantly and all N-acetyl substituents in 2-3 h, the free amino groups of the GlcN residues in these polymers are exposed rapidly, and their glycosidic bonds, which are extremely stable to acid, can be cleaved completely with nitrous acid at room temperature in a matter of minutes, with the conversion of GlcN to anhydro-D-mannose (AMan) (ref. 7). It was also recognized that nitrous acid would cleave the glycosidic bonds of N-sulfated GlcN residues without prior N-desulfation (ref. 8), although the conditions required for this latter reaction were less well defined . We observed that HONO would cleave N-unsubstituted GlcN glycosides selectively at pH 4, whereas N-sulfated GlcN glycosides were cleaved with nitrous acid selectively at pH 1.5 (ref. 9). Thus, HONO is useful not only in simple amino sugar analysis, but also in selective cleavage of heparinoids at N-sulfated GlcN residues to yield oligosaccharide fragments lacking the N-sulfated GlcN residues. This reaction is a deaminative cleavage that is initiated by diazotization of the amino group, and, although it gives nearly stoichiometric yields of the cleaved products, there is a small amount of a side reaction (ref. 9) in which the amino group is lost but the glycosidic bond remains intact ("the ring contraction reaction").

URONIC ACID REDUCTION

The acid resistance of the uronic acid glycosides presents the second problem. This is overcome by a procedure that gives stoichiometric reduction of uronic acid residues in heparinoids and other carbohydrate polymers (ref. 10). The reaction was developed as a modification of a procedure described by Hoare and Koshland (ref. 11) for derivatizing glutamic and aspartic acid side chains in proteins wherein the carboxyl groups are activated with carbodiimide and the resulting ester is then displaced with a nucleophilic reagent (glycine methyl ester). In the case of the polyuronides, the water soluble carbodiimides react rapidly with the uronic acid carboxyl groups to give an ester that is sufficiently electrophilic that it can be reduced by $NaBH_4$. When heparinoids are treated in this manner, the GlcA residues are reduced to glucose residues and the IdoA residues are reduced to idose residues, without any effect on the molecular weight of the polymer. When this modified heparin is heated with 1 N acid at 100 °C, all of the hexose glycosidic bonds are hydrolyzed, and the resulting GlcN glycosides can be cleaved with HONO to complete the conversion of these polymers completely to monomers that can be assayed by NaB^3H_4 procedure. However, one caveat must be noted. L-Idose, which is formed from IdoA in this reaction sequence, is converted under the acid hydrolysis conditions to 1,6-L-idosan (ref. 12), which, because it is a non-reducing sugar, cannot be reduced (i.e., labeled) with NaB^3H_4 and so cannot be assayed in this manner. To circumvent this problem in the analytical stage, it is possible to carry out the carboxyl reduction with NaB^3H_4 to incorporate 2 equivalents of ^3H into each resulting hexose, thus allowing quantitation of the hexoses formed in the hydrolysis/deamination reaction sequence without aldehyde reduction with NaB^3H_4 (ref. 13).

One additional observation from this work was that the glycosidic bond of the α-L-idose-2-SO_4 residue in the polymer is extremely labile to acid, and can be hydrolyzed selectively with little loss of SO_4.(ref. 10).

HYDROLYSIS OF HEPARIN WITHOUT CARBOXYL REDUCTION

An alternative way to determine the monosaccharide composition of heparinoids is to hydrolyze the polymers long enough to remove both the N-sulfate groups and the N-acetyl groups. Under these conditions, very few of the glycosidic bonds of either the uronic acids or the GlcN residues are cleaved. Nevertheless, when this partially hydrolyzed polymer is treated with nitrous acid at pH 4, it is converted completely to disaccharides and some free anhydromannose (AMan) and uronic acids. These hydrolysis/deamination products can be separated and quantitated; Thus, one can deduce the total monosaccharide composition from the amounts of IdoA→AMan, GlcA→AMan, free AMan, and free uronic acids (ref. 14). In the course of these studies, it was found that the L-β-iduronosyl bond is remarkably acid labile; indeed it is as readily hydrolyzed as most neutral hexose glycosides, although the GlcA bonds exhibited the expected acid resistance. After hydrolysis in 1 N acid at 100 °C for 4-6 h followed by treatment with nitrous acid, almost all of the IdoA was recovered as the free monosaccharide whereas most of the GlcA was recovered as GlcA→AMan. In fact, since heparin is a relatively inexpensive material, the hydrolysis/deamination procedure represents an easy way to generate free IdoA for use as a standard.

HPLC OF HEPARIN OLIGOSACCHARIDES

The development of the latter approach to obtain an accurate measure of the uronic acid composition of heparinoids represents an improvement over the previous methodology. However, it does not give any indication of the "monomer" composition of these polymers, since the disaccharides (the monomeric units of these polymers) are characterized and distinguished by both the uronic acid unit and the N- and O-sulfate substituents. Clearly, the N-sulfated disaccharide units that were flanked at their nonreducing terminals by N-sulfated GlcN units could be released by direct treatment of the polymer with nitrous acid at pH 1.5. Indeed, heparin is converted largely to a mixture of the four most prevalent disaccharide units present in heparin by the pH 1.5 nitrous acid treatment. The GlcNAc residues, which represent 10-15 percent of the total GlcN residues in heparin, are recovered almost entirely in the tetrasaccharide pool (ref. 15); the failure to obtain significant levels of hexasaccharides indicated that very few GlcNAc-containing disaccharides occurred side-by-side in heparin. The disaccharide mixture is easily separated from the tetrasaccharide pool, but the structurally similar sulfated disaccharides cannot be separated from each other by paper chromatography. Therefore, an HPLC method has been developed in which the sulfated disaccharides can be completely resolved on a strong anion exchange (SAX) column in a potassium phosphate buffer gradient (ref. 16). A similar gradient elution protocol was developed to separate the complex mixture of tetrasaccharides (ref. 15), most of which contained internal GlcNAc residues (some, however, represented nitrous acid ring contraction products - see above). More recently, a reversed phase ion pairing HPLC method which gives higher resolution of heparin oligosaccharides was developed (ref. 17).

HYDRAZINOLYSIS AND STRUCTURAL ANALYSIS OF HEPARIN OLIGOSACCHARIDES

At least 16 different tetrasaccharides are obtained, but each is obtained only in small amounts insufficient for structure determination by the available spectroscopic or chemical means. It is clear that, if it is possible to use nitrous acid to cleave the tetrasaccharides at the internal GlcNAc residues without loss of O-sulfate substituents, each tetramer can be cleaved into two disaccharides that can be recognized and quantitated by the disaccharide HPLC method described above. However, the most obvious procedure for cleaving at the GlcNAc residues involves hydrazinolysis to N-deacetylate the GlcNAc, thus converting it to a nitrous acid-sensitive GlcN residue. This procedure had been applied to glycosaminoglycans previously with results that were both unsatisfactory and unexplained. A re-examination of the reactions as applied to GAG's led to the finding that the N-deacetylation reaction of GlcNAc is a facile reaction, but that there was also a side reaction in which the uronic acid carboxyl groups are converted to hydrazides, and the hydrazine-treated polymer is cleaved by a β-elimination reaction at these uronic acid hydrazides (refs. 18, 19). This appears to explain the confusion found in the earlier literature. By controlling this side reaction, it is possible to use hydrazinolysis in the structural analysis of the tetrasaccharides. Thus, each tetrasaccharide peak found in the HPLC profile has been isolated, reduced with NaB^3H_4, and cleaved by hydrazinolysis/nitrous acid deamination to yield a labeled reducing terminal disaccharide and an unlabeled nonreducing terminal disaccharide, each of which can be quantitated before and after re-reduction with NaB^3H_4 (ref. 15). A further refinement in which the SAX HPLC procedure is replaced by a reversed phase, ion pairing (RPIP) HPLC method allows separation of both di- and tetrasaccharides in a single elution profile and automatic quantitation of each peak by collecting computer-processed data from a flow detector connected to the column (refs. 17, 20). The RPIP-HPLC approach has also been applied to the separation of the extremely complex mixtures of hexasaccharides that is obtained with heparin is cleaved with limiting levels of nitrous acid (ref. 21). In addition, we have recently found that an aqueous solution (70 percent) of hydrazine gives more rapid deacetylation and less rapid hydrazide formation than anhydrous hydrazine; consequently, less degradation of the product is obtained with 70 percent hydrazine (ref. 22).

PERIODATE OXIDATION OF HEPARIN

Using the methodology described above, it has been possible to apply another "classical" methodology, i.e., periodate oxidation, to this group of polymers. Periodate is expected to cleave heparin and heparan sulfate polymers only at the C2-C3 bond of the unsulfated GlcA and IdoA residues in the polymer. It is, of course, possible to observe the oxidation of heparinoids by measuring the disappearance of IO_4^-, but this gives no information re the structure of the oxidized heparinoid. Unsulfated uronic acid residues occur in about 20 percent of the disaccharides in heparin. However, one of these disaccharide units, GlcA→$GlcNSO_3$(3, 6-di-SO_4), is particularly important since it is a part of the antithrombin III-binding pentasaccharide (1, 2), but this disaccharide is a very minor component of heparin (2-4 % of the total disaccharides). Thus, a much clearer picture of the progress of the reaction can be obtained if one observes the disappearance of the disaccharides containing susceptible uronic acid residues, and particularly the minor disaccharide from the ATIII-binding region. Thus, analysis of the disaccharide components remaining at several time points during the oxidation has made it possible for us to follow the progress of periodate oxidation of the major susceptible disaccharides as well as this rare disaccharide found in the antithrombin-binding sequence (ref. 20).

FUTURE PROBLEMS

The stage is now set for a most interesting new challenge that is developing in the heparinoid field, i.e., the identification of the specific sequences in heparinoids that are responsible for their many physiological activities, which apparently are mediated through interactions with the various heparin-binding proteins. At this point, it is useful to return to the distinctions between heparin and heparan sulfate, this time emphasizing the differences in their natural occurrence and apparent role. Heparin is found only in mast cells as free GAG chains that are complexed with proteases. It is found outside the mast cells only when these cells lyse, and it appears to play no role as an anticoagulant *in vivo*. Heparan sulfate, on the other hand, is quite ubiquitous in Nature, occurring on the surfaces of virtually all animal cell types and in all animal species up and down the phylogenetic scale (ref. 23). Heparan sulfate occurs as a proteoglycan (HSPG) which is a metabolically dynamic cell constituent (refs. 24, 25), being rapidly synthesized in the endoplasmic reticulum and the Golgi and secreted to the cell surface. Then, just as rapidly, it is internalized by the cell and catabolized in the lysosomes, with a half-life of HSPG on the cell surface of about 4 h (ref. 25). We are only beginning to comprehend the role of HSPG, but it clearly has properties that suggest it serves important functions. If HSPG did not play an important physiological role, it is hard to imagine why cells would devote such extensive metabolic machinery to its synthesis, only to take it back into the lysosomes to destroy it! Does its rapid internalization occur so that the cell can take up autocrine or paracrine growth factors that bind to the cell surface HSPG? And does the rapid new synthesis occur so that the cell can replace the old HSPG with a new, structurally altered HSPG that will bind a different set of

heparin-binding proteins? These intriguing questions represent the future of this field. Obtaining the answers presents challenges to both the biologists and the structural chemists.

For the biologists, the questions are: (a) does the ubiquitous occurrence of HSPG on the surface of animal cells truly reflect its biological importance? (b) what are the physiological roles of HSPG? (c) does the binding of heparan sulfate to the various heparin-binding proteins require unique and rare sequences in the GAG chains for each different protein, as found for ATIII? (d) if so, what size must these unique sequences be? (e) why does HSPG turn over so rapidly on the cell surface? and (f) with changes in physiological conditions, do the cells synthesize HSPG's with altered structural features so that the cells can bind different extracellular heparin-binding proteins?

If these biological questions yield interesting answers, then the structural chemist will be faced with the isolation and analysis of the active sequences. Critical questions in obtaining the required structural information are: (a) how rare are the important sequences in heparan sulfate? (b) what size are these sequences? (c) how can these sequences be purified, and in what amounts? and (d) what methods can be used to determine the sequences of the minute amounts of material that are likely to be available?

In considering these questions, it must be recalled that (a) it is not possible to obtain a pure heparan sulfate or a pure heparin, so detection of desired sequences will require the cleavage of the starting heparinoids and the isolation of pure oligosaccharides, (b) heparan sulfate is difficult to obtain in quantities sufficient for isolation of rare oligosaccharides; so that it may be necessary to use heparin as the starting material to obtain the desired sequences, (c) some sequences that are present in heparan sulfate are rare in heparin, (d) it is extremely difficult to purify minor hexasaccharides or octasaccharides from the complex mixtures that are usually obtained by partial cleavage of heparinoids, and (e) some cleavage methods may cleave the starting materials in the middle of the desired sequences, or may destroy critical structural features.

There is a growing awareness of the importance of extracellular matrix molecules in cell behavior, and HSPG is clearly a central player in this developing area of cell biology. The challenge of the complex structural features of heparinoids has driven the field in the past, but the biology of these heparinoids is now emerging rapidly and is beginning to drive the field. There is every reason to believe that this trend will continue and will intensify, and that the challenges to the heparinoid biochemist will continue to grow as well.

REFERENCES

1. U. Lindahl, G. Bäckstrom, and L. Thunberg, J. Biol. Chem. 258, 9826 (1983).
2. D. H. Atha, A. W. Stephens, A. Rimon, and R. D. Rosenberg, Biochemistry 23, 5801 (1984).
3. U. Lindahl, D. S. Feingold, and L. Rodén, TIBS 11, 221-225 (1986).
4. H. E. Conrad, in Heparin and Related Polysaccharides, Ann. New York Acad. Sci. 556, 18-28 (1989).
5. H. E. Conrad., M. E. James, and E. Varboncoeur, Anal. Biochem. 51, 486-500 (1973).
6. A. B. Foster, Chem. and Ind. 627 (1955).
7. J. E. Shively, and H. E. Conrad, Biochemistry 9, 33-43 (1970).
8. D. Lagunoff and G. Warren, Arch. Biochem. Biophys. 99, 396 (1962).
9. J. E. Shively and H. E. Conrad, Biochemistry 15, 3932-3942 (1976).
10. R. L. Taylor and H. E. Conrad, Biochemistry 11, 1383-1388 (1972).
11. D. G. Hoare and D. E. Koshland, J. Biol. Chem. 242, 2447 (1967).
12. A. S. Perlin and G. R. Sanderson, Carbohyd. Res. 12, 183 (1972).
13. R. L. Taylor, J. E. Shively, H. E. Conrad, and J. A. Cifonelli, Biochemistry 12, 3633-3637 (1973).
14. H. E. Conrad, Biochem. J. 191, 355-363 (1980).
15. M. J. Bienkowski and H. E. Conrad, J. Biol. Chem. 260, 356-365 (1985.)
16. S. R. Delaney, M. Leger, and H. E. Conrad, Analyt. Biochem. 106, 253-261 (1980).
17. Y. Guo and H. E. Conrad, Anal. Biochem. 168, 54-62 (1988).
18. P. N.Shaklee, , and H. E. Conrad, Biochem. J. 217, 187-197 (1984).
19. P. N.Shaklee and H. E. Conrad, Biochem. J. 235, 225-236 (1986).
20. Y. Guo, unpublished results.
21. M. Gamble, unpublished results.
22. Y. Guo and H. E. Conrad, Analyt. Biochem. 176, 96-104 (1989).
23. C. M. F. Cassaro and C. P. Dietrich, J. Biol. Chem. 252, 2254-2261 (1977)
24. M. Ishihara, N. S. Fedarko, and H. E. Conrad, J. Biol. Chem. 261, 13575-13580 (1986).
25. M. J. Bienkowski and H. E. Conrad., J. Biol. Chem. 259, 12989-12996 (1984).