Conformation of the Mucopolysaccharides

X-RAY FIBRE DIFFRACTION OF HEPARIN

By I. A. NIEDUSZYNSKI and E. D. T. ATKINS

*H. H. Wills Physics Laboratory, University of Bristol, Royal Fort, Tyndall Avenue,
Bristol BS8 1TL, U.K.*

(Received 14 June 1973)

Oriented films of the sodium salt of heparin crystallize in a triclinic unit cell. The X-ray diffraction results favour a tetrasaccharide covalent repeating sequence but the measured molecular repeat of 1.59 nm is inconsistent with all α -(1 $\rightarrow$ 4)(1*ax* $\rightarrow$ 4*eq*)-glycosidic linkages as generally accepted. Three models are proposed for the molecular shape of heparin, which depend on the chair conformation adopted by the hexuronic acid moieties. Model I has both uronic acids in the C1 chair form, model II has one uronic acid in each of the C1 and 1C chairs and model III has both in the 1C chair form. The merits of these models are discussed.

Heparin is a highly sulphated acidic hetero-polysaccharide found in mammalian connective tissue. The blood anti-coagulant properties associated with heparin and its role as an anti-lipemic (fat-clearing) agent have aroused considerable interest. [For recent reviews see Bettelheim (1970), Jeanloz (1970) and Lindahl (1970).]

Despite a considerable amount of structural investigation the covalent chemical repeating unit in heparin is still in some doubt, largely as a consequence of its resistance to acid hydrolysis. Until recently it was thought to be composed of tetrasaccharide or possibly disaccharide repeating sequences with equimolar quantities of D-glucosamine and D-glucuronic acid residues (Wolfson *et al.*, 1943) and joined together with α -(1 $\rightarrow$ 4)-glycosidic linkages (Danishefsky & Steiner, 1965). The reported sulphate content of heparin varies between 4 and 6 sulphate groups/tetrasaccharide unit. Cifonelli & Dorfman (1962) indicated the presence of L-iduronic acid and Lindahl & Axelsson (1971) have suggested that sulphated L-iduronic acid is the major sulphated uronic acid present. The generally accepted covalent linkages in heparin, at least until quite recently, are illustrated in Fig. 1(a). Perlin *et al.* (1972) from consideration of nuclear-magnetic-resonance data favoured only α -(1 $\rightarrow$ 4)-glycosidic linkages and, further, suggested that L-iduronic acid is the major uronic acid in heparin. However, the occurrence of β -glucuronidic linkages has been established by Helting & Lindahl (1971) and in addition Lindahl *et al.* (1972) have proposed that D-glucuronic acid and its C(5) epimer L-iduronic acid are in equilibrium at the polymer level, at least in the biosynthetic process, but that subsequent sulphation of the L-iduronic acid residues removes them from this equilibrium.

Thus, the relative amounts of the two uronic acids in heparin are still in doubt, and it is pertinent to ask whether a unique covalent repeating sequence (such as a di- or tetra-saccharide) exists for this relatively short molecule. (The average chain length is only of the order of 10 tetrasaccharide units.) X-ray diffraction patterns of oriented material would be expected to offer a useful contribution to the understanding of the molecular shape and, possibly, composition of heparin.

The ability of heparin preparations to crystallize was demonstrated by Charles & Todd (1940) who obtained a barium acid salt of heparin in crystalline form (see also Wolfson *et al.*, 1943). Unfortunately, these crystals were too small to yield conventional single-crystal X-ray diffraction patterns and the crystallization was used to demonstrate the homogeneity of the preparations rather than to provide structural information.

Materials and Methods

The sodium salt of heparin from hog intestinal mucosa was kindly supplied by Professor T. C. Laurent and Dr. U. Lindahl (Institute of Medical Chemistry, University of Uppsala, Sweden). The sample H140 cp III was a purified form of heparin stage XIV obtained from the Wilson Laboratories, Chicago, Ill., U.S.A. and is formally similar to preparation I-CPI of Lindahl *et al.* (1965). It corresponds to a viscosimetric molecular weight of 12 000 and a sulphur/hexosamine ratio of 2.3:1. Some experiments were also performed with commercial Pularin heparin batch ON0 150 purchased from Boots (Chemists) Ltd., Bristol Branch, Bristol, U.K.

Aqueous solutions (1%, w/v) were allowed to dry

Fig. 1. Molecular conformations of heparin

(a) Generally accepted conformation, at least until the recent work of Helting & Lindahl (1971), involving exclusively α -(1 $\rightarrow$ 4) (1ax $\rightarrow$ 4eq)-glycosidic linkages. (b) A possible tetrasaccharide repeat that cannot be excluded on the X-ray data. All the sugar residues are in the C1 chair form. The glycosidic linkages are (1ax $\rightarrow$ 4eq) between glucosamine and uronide and (1eq $\rightarrow$ 4eq) between uronide and glucosamine. (c) An alternative tetrasaccharide repeat in which one

down to make films of approximately 40 μ m thickness which were then cut up into strips, stacked and stretched with weights at controlled relative humidities in a manner already described in detail by Atkins *et al.* (1972).

Results

The X-ray diagram obtained from the sodium salt of heparin at 80% relative humidity is shown in Plate 1. The reflexions index on a triclinic unit cell with parameters $a = 1.30 \pm 0.01$ nm, $b = 1.02 \pm 0.01$ nm, c (fibre axis) = 1.59 ± 0.02 nm, $\alpha = 104 \pm 1^\circ$, $\beta = 96 \pm 1^\circ$, $\gamma = 116 \pm 1^\circ$ and with volume = 1.79 nm^3 . The density of a crystalline sample was determined by flotation to be $1.72 \text{ g} \cdot \text{cm}^{-3}$ and since an axial repeat of 1.6 nm cannot be reconciled with fewer than four extended saccharide residues, these must constitute the minimum cell contents. By using a chemical composition with 5 sulphate groups as a guideline (weight of tetrasaccharide plus Na^+ counter ions = 1200 atomic units) it is seen that the observed density is consistent with a maximum of 1.54 tetrasaccharide units or 6 anhydrous saccharide residues/unit cell. However, the degree of dependence of the X-ray spacings on the humidity makes it clear that this unit cell contains a considerable quantity of water. The most probable explanation is that only one chain, comprising a tetrasaccharide repeating unit, passes through the unit cell.

An interesting feature of the X-ray diagram (Plate 1) is the sharpness of those reflexions lying on the equator (in the radial direction) and the breadth of reflexions on or near to the meridian. This is due principally to the very small number of repeats in the fibre direction compared with the large number of lateral periodicities. Measurements on the breadth of the (001) reflexion show that the 'crystallites' are only about 16 nm long in the chain direction which corresponds to about 10 tetrasaccharide repeat units. The ratio of the molecular weight to the weight/tetrasaccharide is $12000/1200 = 10$, and yields the same number of repeat units per chain. This agreement not only supports the conclusion that a tetrasaccharide unit is involved but also shows an interesting example

of the uronic acids (probably a sulphated L-iduronic acid) adopts the 1C chair form. Three types of glycosidic linkage are displayed. (d) A tetrasaccharide repeat in which both uronic acids are in the 1C chair form. This model most readily fits the X-ray data. Abbreviations: D-GlcNAc, *N*-acetyl-D-glucosamine, probably mono- or di-sulphated. *N*-acetyl group usually replaced by sulphamino; D-GlcUA, D-glucuronic acid; L-IdUA, L-iduronic acid, often sulphated (probably in the 2-position); UA, uronic acid.

EXPLANATION OF PLATE I

X-ray fibre-diffraction photograph obtained from the sodium salt of heparin

The unit cell is triclinic so that the 001 reflexions lie off the meridian (imaginary vertical line). Note the sharpness of the reflexions in the radial direction on the equator (imaginary horizontal line) compared with the off-equatorial reflexions. The dotted ring is for calibration purposes.

of assessing the degree of polymerization of a biopolymer from the breadth of an X-ray reflexion. Thus the X-ray diffraction results provide strong evidence for a unique tetrasaccharide crystallographic repeat for a major proportion of the heparin molecule.

At high relative humidities it is possible to induce the sodium salt of heparin to crystallize (the material is only amorphous, as often stated in the literature, when kept in humidities below about 70% relative humidity) and both preparations used in this investigation gave rise to the same X-ray spacings and therefore presumably the same crystalline form.

Molecular conformation

The axial periodicity of 1.59 nm eliminates certain combinations of glycosidic linkages in heparin, for although the maximum length of an α -D-glucose residue is 0.45 nm [on the basis of co-ordinates of Ramachandran *et al.* (1963) and Arnott & Scott (1972)], which in principle permits periodicities up to 1.80 nm, conformational analysis of the α -(1 $\rightarrow$ 4) (1ax $\rightarrow$ 4eq)-linkage shows that an axial rise per monomer of as much as 1.59/4 $\approx$ 0.40 nm is stereochemically impossible even for amylose (Sundarajan & Rao, 1969) and is therefore quite impossible for the heavily sulphated heparin molecule. Thus it may be argued that the model shown in Fig. 1(a) with exclusively α -linkages is incompatible with the combined X-ray and stereochemical evidence.

The X-ray diffraction results indicate that the gross features of molecular shape are representable at the tetrasaccharide level. Therefore the next step is to establish which alternative models, differing in their covalent structure and ring conformations, should be considered. There is no reason to suppose that the glucosamine residues are in other than the C1 chair form, linking through their 1-positions axially and the 4-positions equatorially. But the precise uronic acid composition is in doubt, and D-glucuronic acid residues would be expected to adopt the C1 chair form [the potential energy of the 1C chair form is some 20–25 kcal/mol higher than for the C1 chair (Sundarajan & Rao, 1968)], whereas L-iduronic acid may adopt either chair form. Thus three molecular shapes (at present restricted to the energetically favoured chair forms) that depend on the uronic acid composition and conformation need to be considered: model I has both uronic acids in the C1 chair, model II has one uronic acid in the C1 and one in the 1C chair, and model III has both in the 1C chair form.

Model I (Fig. 1b)

In this model each sugar residue is in the C1 chair form and the glycosidic linkages are alternately (1ax $\rightarrow$ 4eq) between glucosamine and uronic acid and (1eq $\rightarrow$ 4eq) between uronic acid and glucosamine (i.e. alternating α - and β -linkages in the old notation).

All the side groups are favourably disposed in the equatorial position except for an axial carboxyl group on the L-iduronic acid moiety (or moieties, see below). The maximum theoretical extension is 1.98 nm and by using space-filling models trial structures were built with measured repeats $\leq$ 1.9 nm, although it becomes rather difficult to contract such a model down to fit the observed repeat of 1.59 nm. The precise restrictions on conformational flexibility are very much dependent on the number and positions of the sulphate groups.

Model II (Fig. 1c)

This model differs from the previous one in that one of the two uronic acid residues is changed from a C1 to a 1C chair form. The calculated maximum molecular repeat shortens by about 0.1 nm and contraction of space-filling models to the observed 1.59 nm repeat is somewhat easier.

Atkins & Laurent (1973) in studies on dermatan sulphate were able to show that their results were consistent with a C1 chair form for L-iduronic acid (see also Atkins & Isaac, 1973). That L-iduronic acid residues may adopt a solid-state conformation which is different from that found in dermatan sulphate might be a consequence of sulphation. In chondroitin 4-sulphate (Isaac & Atkins, 1973), dermatan sulphate (Atkins & Isaac, 1973) and the plant polysaccharide carrageenans (Anderson *et al.*, 1969) the sulphate groups adjacent to the linkage positions are all axially placed, and inspection of molecular models suggests that axial sulphate groups are sterically less cumbersome. Thus, if L-iduronic acid adopted the 1C chair conformation, it would dispose all of the large side groups favourably: the sulphate group axial and the carboxyl group equatorial.

Model III (Fig. 1d)

In this model both of the uronic acid residues are in the 1C chair form and are linked diaxially. [Interestingly, this configuration could also be described as exclusively α -(1 $\rightarrow$ 4)-linked, but clearly this notation does not distinguish between the models shown in Figs. 1(a) and 1(d).] The polymer repeat is shortened by an additional 0.1 nm (i.e. maximum extension = 1.78 nm) and space-filling models readily yield the measured periodicity of 1.59 nm.

This model is further supported by the nuclear-magnetic-resonance data of Perlin *et al.* (1970), who proposed the 1C chair form for L-iduronic acid and suggested (Perlin *et al.*, 1972) that L-iduronic acid is the main uronic acid in heparin. Hirano (1972) proposes this model on the basis of the interpretation of nuclear-magnetic-resonance data which favour the acetylated derivatives of the hexuronides to contain three axial and two equatorial or possibly five axial acetyl groups. In addition it is argued that having

both uronic acids in 1C chairs is helpful in understanding the epimerization of D-glucuronic to L-iduronic acid at polymer level (Lindahl *et al.*, 1972). However, we do not place too much weight on this argument since both uronides could be in the C1 chair form for epimerization and the subsequent sulphation of the L-iduronic acid may act as the driving force for alteration of chair conformations.

Discussion

Depending on the local uronic acid and sulphate composition the heparin molecule may well adopt any of the molecular shapes shown in Figs. 1(b)–1(d), but the X-ray diffraction photograph shows that a major part of the heparin molecule is coherent in displaying a single repeating molecular shape.

There is insufficient evidence to exclude any of the three models suggested. Model I seems the least probable since a complement of 5 sulphate groups/tetrasaccharide would make it very difficult for the 1.59 nm periodicity to be achieved and any sulphated L-iduronic acid present seems likely to be in the 1C chair form (Fransson, 1973). Model II may prove consistent with the observed periodicity and it is likely that either D-glucuronic acid or an unsulphated L-iduronic acid would constitute the C1 chair form. The recent chemical evidence of Lindahl (U. Lindahl, personal communication) favouring L-iduronic acid to constitute a high percentage (70–80%) of the total uronic acid also argues against model I since a significant proportion of this iduronic acid appears to be sulphated (Lindahl & Axelsson, 1971) and therefore is likely to be in the 1C chair. Finally, model III which most readily satisfies the observed axial periodicity would seem to depend on a high content of sulphated L-iduronic acid. As seen in Fig. 1(d) this model may be considered to have a disaccharide periodicity; the observed tetrasaccharide periodicity could result from the disposition of the sulphate groups [e.g. two on one glucosamine and one on the other; the results of Linker & Hovingh (1972) indicate that more than 30% of the glucosamine units are disulphated].

Thus, the precise molecular shape of heparin is likely to be determined by the L-iduronic acid content and by the proportion of the L-iduronic acid that is sulphated.

Conclusion

The ability of heparin to crystallize in an oriented manner argues for regularly repeating covalent sequences of an essentially unbranched chain. Further, the X-ray diffraction and density data favour a tetrasaccharide crystallographic repeat unit with no internal symmetry. Heparin molecules require nearly their own volume of water in order to crystal-

lize and they do so in a triclinic packing mode with adjacent chains having the same polarity and staggered with respect to one another. The axial periodicity of 1.59 nm and steric considerations reject any model based entirely on α -(1 $\rightarrow$ 4)(1 *ax* $\rightarrow$ 4 *eq*)-linkages.

These X-ray data, together with the excellent electron micrographs obtained by Hirano (1972), provide strong evidence for a linear heparin molecule with approximately ten tetrasaccharide repeats. These recent data, which rule out the generally accepted model for heparin with exclusively α -(1 $\rightarrow$ 4)-glycosidic linkages, require us to consider alternative possible structures. The better X-ray data now available will permit a serious assessment and comparison of these subtly different molecular conformations. In principle we have the information already at hand in the form of the X-ray diffraction pattern shown in Plate 1 and hopefully detailed structure determination will illuminate the novel properties of heparin.

We thank Professor T. C. Laurent and Dr. U. Lindahl for advice and a generous supply of material, Dr. Å-K Fransson and Dr. C. F. Phelps for helpful discussion, the S.R.C. for a fellowship (to I. A. N.) and the Arthritis and Rheumatism Council for support.

References

- Anderson, N. S., Campbell, J. W., Harding, M. M., Rees, D. A. & Samuel, J. W. B. (1969) *J. Mol. Biol.* **45**, 85–99
- Arnott, S. & Scott, W. E. (1972) *J. Chem. Soc. Perkin Trans. II*, 324–335
- Atkins, E. D. T. & Isaac, D. H. (1973) *J. Mol. Biol.* in the press
- Atkins, E. D. T. & Laurent, T. C. (1973) *Biochem. J.* **133**, 605–606
- Atkins, E. D. T., Phelps, C. F. & Sheehan, J. (1972) *Biochem. J.* **128**, 1255–1263
- Bettelheim, F. A. (1970) in *Biological Polyelectrolytes* (Veis, A., ed.), chapter 3, Marcel Dekker, New York
- Charles, A. F. & Todd, A. R. (1940) *Biochem. J.* **34**, 112–118
- Cifonelli, J. A. & Dorfman, A. (1962) *Biochem. Biophys. Res. Commun.* **7**, 41–45
- Danishefsky, I. & Steiner, H. (1965) *Biochim. Biophys. Acta* **101**, 37–41
- Fransson, L.-Å. (1973) *Proc. Annu. Symp. Soc. Study Inborn Errors Metab.* 11th in the press
- Helting, T. & Lindahl, U. (1971) *J. Biol. Chem.* **246**, 5442–5447
- Hirano, S. (1972) *Int. J. Biochem.* **3**, 677–683
- Isaac, D. H. & Atkins, E. D. T. (1973) *Nature (London New Biol.)* **244**, 252–253
- Jeanloz, R. W. (1970) in *The Carbohydrates* (Pigman, W. W. & Horton, D., eds.), pp. 589–625, Academic Press, London and New York
- Lindahl, U. (1970) in *Chemistry and Molecular Biology of the Intercellular Matrix* (Balazs, E. A., ed.), pp. 943–960, Academic Press, London and New York

- Lindahl, U. & Axelsson, O. (1971) *J. Biol. Chem.* **246**, 74–82
- Lindahl, U., Cifonelli, J. A., Lindahl, B. & Roden, L. (1965) *J. Biol. Chem.* **240**, 2817–2820
- Lindahl, U., Bäckström, G., Malmström, A. & Fransson, L.-Å. (1972) *Biochem. Biophys. Res. Commun.* **46**, 985–991
- Linker, A. & Hovingh, P. (1972) *Biochemistry* **11**, 563–568
- Perlin, A. S., Casu, B. & Sanderson, G. R. (1970) *Can. J. Chem.* **48**, 2260–2268
- Perlin, A. S., Ng, Y. K., N. M. K. & Battacharjee, S. S. (1972) *Can. J. Chem.* **50**, 2437–2441
- Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan (1963) *J. Mol. Biol.* **7**, 95–99
- Sundararajan, P. R. & Rao, V. S. R. (1968) *Tetrahedron* **24**, 289–295
- Sundararajan, P. R. & Rao, V. S. R. (1969) *Biopolymers* **8**, 313–323
- Wolfson, M. L., Weisblat, D. I., Karabinor, J. V., McNeely, W. H. & McLean, J. (1943) *J. Amer. Chem. Soc.* **65**, 2077–2085