Nonrandom Structural Features in the Heparin Polymer†

Robert J. Linhardt,*,‡ David M. Cohen,§ and Kevin G. Rice‡

Division of Medicinal and Natural Products Chemistry, College of Pharmacy, and Department of Computer Science, University of Iowa, Iowa City, Iowa 52242

Received August 16, 1988; Revised Manuscript Received October 25, 1988

ABSTRACT: Computer simulation studies were used to prepare an ensemble of heparin number chains. The polydispersity of these chains was simulated by introducing a specific "fraction of terminators", and it closely resembled the experimentally observed polydispersity of a porcine mucosal, glycosaminoglycan heparin. The same percentage of simulated chains contained antithrombin III (ATIII) binding site sequences as are typically found to contain ATIII binding sites using affinity chromatography. Heparin lyase action was then simulated by using Michaelis–Menten kinetics. In one model, heparin chains were constructed from the random assembly of monosaccharide units using the observed mole percentage of each. After simulated depolymerization, the final oligosaccharides formed were compared to the observed oligosaccharide products. The simulation which assumed a random distribution of monosaccharide units in heparin did not agree with experimental observations. In particular, no ATIII binding site sequences were found in the simulated number chains. The results of this simulation indicate that heparin is not simply a random assembly of monosaccharide units. These results are consistent with the known, ordered biosynthesis of heparin. In a second model, heparin chains were constructed from randomly assembled oligosaccharides at the mole percentage in which each is found in the final product mixture. The action of heparin lyase was then simulated, and the distribution of the oligosaccharide products was measured throughout the simulated time course of the depolymerization reaction. The simulated rate of formation and final concentration of a particular oligosaccharide which contains a portion of heparin's ATIII binding site were similar to those observed experimentally. These results are consistent with the random distribution of ATIII binding sites within glycosaminoglycan heparin. These results are in contrast to the reported ordered distribution of ATIII bindings sites in proteoglycan heparin and may be the result of a loss of order following the processing of proteoglycan heparin by endo-β-glucuronidase.

Heparin is a polydisperse, linear, sulfated polysaccharide which is widely used as an anticoagulant (Casu, 1986). The major disaccharide repeating unit in the heparin polymer is →4)-2-deoxy-2-sulfamino-α-D-glucopyranose 6-sulfate-(1→4)-α-L-idopyranosyluronic acid 2-sulfate-(1→. This repeating unit represents 88 mol % of the disaccharides comprising porcine mucosal heparin (Linhardt et al., 1988b). A small number of additional disaccharides of defined structure comprise 12 mol % of the remaining disaccharide units and together account for nearly 90 wt % of the polymer's mass (Linhardt et al., 1988b). Heparin lyase (EC 4.2.2.7) is an eliminase which acts in a random endolytic fashion (Linhardt et al., 1982) to cleave heparin's major disaccharide unit (Figure 1). Computer-simulation studies were previously used to study the distribution of the sites at which heparin lyase cleaves (Linhardt et al., 1985). These studies compared the experimentally observed distribution of depolymerization products [i.e., mole percent of disaccharides, tetrasaccharides, hexasaccharides, etc. measured using gel permeation chromatography (GPC)] with a simulated product distribution. The results obtained in this preliminary computer-simulation study were consistent with a random distribution of these sites within the heparin polymer.

These computer-simulation studies have now been extended to examine a wider variety of the structural features of the heparin polymer including the saccharide units comprising heparin's binding site to antithrombin III (ATIII) which is responsible in large part for its anticoagulant activity. The computer-simulation has also been performed in such a way as to permit the examination of the product distribution obtained at various degrees of reaction completion. This has allowed us to begin to assess whether heparin lyase has a truly random endolytic action pattern.

Theory

General Assumptions. The computer-simulation model for heparin depolymerization assumes that heparin lyase acts as an eliminase (Linhardt et al., 1986a) in a random endolytic fashion (Linhardt et al., 1982) to cleave certain specific α-glycosidic linkages (Rice, 1988) present in the heparin polymer. To accurately represent heparin, both its polydispersity and its microheterogenicity are taken into account. Heparin's polydispersity is simulated, resulting in a polydisperse ensemble of number chains having a calculated number-average molecular weight of 13 000. The saccharide composition of heparin is determined experimentally using heparin lyase followed by strong anion-exchange (SAX) high-pressure liquid chromatography (HPLC) (Rice et al., 1985; Linhardt et al., 1988b). The sequence of each heparin chain is simulated, and the number chains are constructed by random assembly of either monosaccharide or oligosaccharide units of defined structure (Merchant et al., 1985; Linhardt et al., 1986b; Rice et al., 1987). After construction of a polydisperse microheterogeneous ensemble of heparin number chains, the action of heparin lyase is simulated, and product distributions are calculated at various time points throughout the course of the simulated depolymerization. These simulated data are compared with experimental data obtained by using SAX-HPLC to measure the oligosaccharide products through the reaction time course.

† This work was supported by National Institutes of Health Grants HL29797 and GM38060.

* Author to whom correspondence should be addressed.
‡ Division of Medicinal and Natural Products Chemistry.
§ Department of Computer Science.

0006-2960/89/0428-2888$01.50/0

Table I: Experimentally Measured Monosaccharide Composition of Porcine Mucosal Heparin

monosaccharide component[c]	integer designation	mol %[a]	mol %[b]
GA2S6S	0	46.41	38.88
I2S	1	45.60	38.20
GA2S	2	1.69	1.42
G	3	3.45	6.95
NAcGA6S	4	0.95	4.85
I	5	0.95	4.85
GA2S3S6S	6	0.95	4.85
G2S	7	0	0
NAcGA	8	0	0

[a] Values calculated from Table II. [b] Values calculated from Table III (100% reaction completion) assuming the unidentified products accounting for 18.8 wt % of the mass are all equally distributed among monosaccharide units 3–6. [c] Abbreviations: GA2S6S, 2-deoxy-2-sulfamido-α-D-glucopyranose 6-sulfate; I2S, α-L-idopyranosyluronic acid 2-sulfate; GA2S, 2-deoxy-2-sulfamido-α-D-glucopyranose; G, β-D-glucopyranosyluronic acid; NAcGA6S, 2-deoxy-2-acetamido-α-D-glucopyranose 6-sulfate; I, α-L-idopyranosyluronic acid; GA2S3S6S, 2-deoxy-2-sulfamido-α-D-glucopyranose 3,6-disulfate; G2S, β-D-glucopyranosyluronic acid 2-sulfate; NAcGA, 2-deoxy-2-acetamido-α-D-glucopyranose.

Table II: Experimentally Measured Final Oligosaccharide Composition of Porcine Mucosal Heparin

oligosaccharide product	integer designation	mol %[a]
F1	01	81.84
F2	0121	1.19
F2′	21	2.79
F3	0301	5.89
F4	0101	6.07
F5	634501	2.22

[a] Mole percent of oligosaccharide products is calculated from data presented in Table III (100% reaction completion) assuming the unidentified products accounting for 18.8 wt % of the mass are equally distributed among oligosaccharides F1–F5.

Symbolic Models. Two different symbolic models of the heparin lyase catalyzed depolymerization of heparin were used to test the hypotheses regarding regularities in the structure of a population of heparin chains.

In model 1, heparin polymers are viewed as being comprised of sequences of alternating α,β-(1→4) glycosidically linked uronic acid and hexosamine monosaccharide units. Nine types of monosaccharide units (Table I) are represented by integers 0 through 8. The odd-numbered units represent specific uronic acids, and the even-numbered units represent specific hexosamines. Heparin lyase can only cleave between units of type 1 and units of type 0, 2, or 6 present within the intact polymer. Tetrasaccharides containing 0→1, 2→1, and 6→1 linkages, however, are not cleaved.

In model 2, the chains of heparin are considered to be sequences of units chosen from among six kinds of oligosaccharides each composed of between two and six monosaccharide units (Table II). These "final" oligosaccharide products are designated F1, F2, F2′, F3, F4, and F5. Heparin lyase can cleave the bond between any two adjacent oligosaccharides but does not cleave any bond within a free (unbound) oligosaccharide.

In both model 1 and model 2, the arrangement of the units which comprise the polymer is assumed to be random. The lengths of the chains are determined probabilistically, by randomly deciding whether or not each unit in the population is the last unit of the chain.

Computer Simulation. The hypotheses underlying model 1 and model 2 are subjected to tests by means of computer simulation. The simulation consists of two principal phases: generation of number chains and subsequent depolymerization of the number chains by heparin lyase.

Generation of Chains. The simulation begins by generating a population of heparin chains, comprised either of monosaccharide units (as in model 1) or of oligosaccharide units (as in model 2).

A pseudo-random-number generator is used to select the type of unit at each position in each chain. The algorithm for generating the (pseudo) random numbers is a linear congruential one (Cooper & Clancy, 1985; Knuth, 1981; Park & Miller, 1988; Cohen, 1989). In the same manner, a single parameter called "fraction of terminators" is used as a cutoff probability for the random variable which is used to decide whether or not a particular unit is the terminus of a polymer chain. The fraction of terminators is the same for each type of unit. (See Results and Discussion for possible interpretations of this parameter.)

We now show how to choose the value of the fraction of terminators so that the simulated heparin has a predetermined average mass per chain. The mass of simulated heparin equals the population size (ps) times the weighted average mass per monosaccharide (model 1) or per final oligosaccharide (model 2). To obtain the average mass per chain, we need to calculate the number of chains. For each run of the simulation (starting with a different seed for the random-number generator), the number of simulated chains of heparin is a random variable which depends on the values returned from successive invocations of the random-number generator. In simulations of model 2, the number of simulated chains has a mean value of

$$\text{ps}(\text{fraction of terminators})$$

and a variance of

$$\text{ps}(\text{fraction of terminators})\,[1 - (\text{fraction of terminators})]$$

In the case of simulations of model 1, the mean and variance are halved, because only odd-numbered monosaccharides (i.e., 1, 3, 5, and 7) may occur at the end of a chain (Cohen, 1989). Using the expected value (mean value) of the number of chains as an estimate for its actual value, we calculate the value of the fraction of terminators that yields an average mass per chain of 13 000 for model 1 (5.39%) and for model 2 (5.91%).

Depolymerization. After the selection of the types of units and the demarcation of the ends of chains, the heparin mixture is sufficiently specified for the simulation of its depolymerization by heparin lyase. The simulation proceeds by advancing the (simulated) time by a small amount (one "time step"). The action of heparin lyase on the polymer chains during this time is then simulated. The change in the number of bonds is calculated by means of a Taylor series method of order 2 for the numerical solution of an ordinary differential equation (Johnson & Riess, 1982). The differential equation is

$$x'(t) = \frac{-\alpha x(t)}{\beta + x(t)}$$

where $x(t)$ equals the nanomoles of cleavable bonds at time t and $x'(t)$ is the derivative of $x(t)$ with respect to time. The constants α and β are the Michaelis–Menten V_{max} and apparent K_m, respectively, multiplied by the simulated volume. The values of V_{max} and K_m have been measured for heparin lyase (V_{max} = 3.29 nmol/min; K_m = 293.0 nM). Recalling that the heparin concentration used in the experiments was 20.0 mg/mL, we calculate the simulated volume by solving the equation:

$$\frac{\text{mass of simulated heparin (mg)}}{\text{simulated volume (mL)}} = 20.0$$

The Taylor series method of order 2 computes the change of $x(t)$ from time t_1 to t_2 by the formula:

$$x(t_2) = x(t_1) + x'(t_1)(t_2 - t_1) + \frac{x''(t_1)(t_2 - t_1)^2}{2}$$

Calculation of $x''(t_1)$ can be done by differentiating the formula for $x'(t)$ and letting the independent variable t equal t_1.

Once the number of bonds to be cleaved during the time step has been calculated, a cutoff value for the random variable is obtained by dividing the number of bonds to be cleaved by the total number of cleavable bonds in the population of number chains. [Recall that some bonds are not cleavable either because heparin lyase does not cleave bonds between some types of monosaccharide units (model 1) or because heparin lyase will not cleave any of the six final oligosaccharide products (model 2).] The bonds to be cleaved are then selected at random, and the data structure representing the population of chains is updated.

The simulation halts after the desired percentage of cleavable bonds in the initial population of chains has been cleaved. The state of the heparin mixture at various times during the depolymerization can be saved in a disk file during the execution of the simulation program. Hence, information about transient oligosaccharides (oligosaccharides comprised of more than one of the final oligosaccharide products and still containing a site cleavable by heparin lyase) can be obtained from the simulation and compared to experimental data.

Experimental Procedures

Materials

Chemicals. Heparin sodium salt, from porcine intestinal mucosa (160 units/mg), was obtained from Hepar Industries, Franklin, OH. Heparin lyase (heparinase, EC 4.2.2.7) was purified from *Flavobacterium heparinum* (5 m-units/μg (1 m-unit = 1 nmol of product formed/min) or purchased (11 m-units/μg) from Sigma Chemical Co., St. Louis, MO, and both were essentially free of contaminating activities that could act either on the heparin polymer or on its depolymerization products (Yang et al., 1985). All other chemicals were reagent grade.

Equipment. SAX-HPLC was performed using two LDC-Milton Roy (Riviera Beach, FL) Constametric III pumps with gradient control by digital to analog interface using an Apple IIe microcomputer running Chromatochart software from Interactive Microware Inc., State College, PA. The system was equipped with a Rheodyne (Cotati, CA) 7125 injector and an ISCO (Lincoln, NE) Model 1840 variable-wavelength UV detector. Separations were performed on a 4.6 mm × 25 cm SAX column of 5-μm particle size from Phase Separations, Norwalk, CT. UV spectroscopy was performed on a Shimadzu (Tokyo, Japan) Model UV-160 spectrophotometer. The simulation was written in the Pascal programming language on a Digital Equipment Corp. VAX 11/780 minicomputer using the 4.3BSD UNIX operating system. The simulation was recoded in MPP Pascal and run on the Massively Parallel Processor supercomputer at NASA/Goddard Space Flight Center in Greenbelt, MD.

Methods

Heparin Depolymerization. Heparin (16 mg in 1 mL) was depolymerized at 30 °C with 1 m-unit/mL heparinase in a solution of 0.2 M sodium chloride/5 mM sodium phosphate at pH 7.0. During the reaction time course (at 0, 60, 120, 180, 240, 300, and 600 min), 100-μL aliquots were collected and immediately frozen on dry ice. The percent of reaction completion was determined by measuring the absorption of the product, obtained at each time point, at 232 nm after a 1:100 dilution in 30 mM hydrochloric acid. This absorbance was divided by the final product absorbance measured at 600 min, and the resulting fraction was multiplied by 100.

Kinetic Analysis by SAX-HPLC. The aliquots were collectively heated in a water bath at 100 °C for 1 min to thermally inactivate the enzyme. SAX-HPLC was performed by injecting a 10-μL sample containing 160 μg of product or 10 μL of diluted sample (1:20 in distilled water) containing 8 μg of product into a fixed 200-μL sample loop attached to the SAX column. The column was preequilibrated with 0.2 M sodium chloride and eluted using a 165-mL linear gradient from 0.2 to 1.5 M sodium chloride, pH 3.5, at a flow rate of 1.5 mL/min.

Results and Discussion

Heparin lyase acts in a random endolytic fashion (Linhardt et al., 1982) to cleave heparin at its major repeating disaccharide unit (Figure 1) while permitting some structural variation around this site, i.e., the 3- and 6-positions of the glucosamine residue may or may not be sulfated (Rice, 1988). Tetrasaccharides containing heparin lyase cleavable linkages are resistant to cleavage, being cut only under exhaustive conditions. The ratio of K_{cat} (V_{max}/K_m) measured for the tetrasaccharide-sized substrates to the polymeric substrate is approximately 10^{-4}. With a knowledge of the specificity of this enzyme it has been possible to use it as a tool to sequence heparin-derived oligosaccharides (Linhardt et al., 1988a).

A previous study by our laboratory has demonstrated that computer-simulation could be used in conjunction with experiments involving heparin lyase catalyzed depolymerization of heparin to test hypotheses concerning heparin's structure (Linhardt et al., 1985). In these early studies, heparin was simulated by using number chains of 1 (heparin lyase cleavable site) and 0 (uncleavable site). A polydisperse (based on the experimentally measured polydispersity) ensemble of heparin number chains was simulated containing the appropriate number (experimentally determined) of cleavable (1) and uncleavable sites (0). The arrangement of these sites (i.e., random or ordered) could be controlled in the simulation. The action of heparin lyase was then simulated, and a distribution of disaccharide, tetrasaccharide, hexasaccharide, etc. sized number chains was obtained. The size distributions of the product number chains were compared to the size distributions of products obtained experimentally using GPC. The results of this previous study demonstrated that the simulated distribution was consistent with the experimental product distribution when the heparin lyase cleavable sites were distributed randomly through the heparin chains.

A major deficiency of this preliminary study was that it relied on experimental data obtained by GPC. These data only provided the mole percentage of products having a certain number of adjacent uncleavable sites (i.e., a tetrasaccharide has a single uncleavable site, a hexasaccharide has two, etc.). Recent studies in our laboratory have demonstrated that there are a variety of different tetrasaccharides each with structurally unique uncleavable sites (Merchant et al., 1985; Linhardt et al., 1988b). Thus, the simulation of these more subtle structural features in the heparin polymer might provide a deeper insight into the sequence of the heparin polymer than did the early simulation involving the disposition of heparin's major disaccharide unit, the heparin lyase cleavable site.

When heparin is depolymerized with heparin lyase, a small number of oligosaccharide products are afforded (Figure 1). The structures of these oligosaccharide products have been

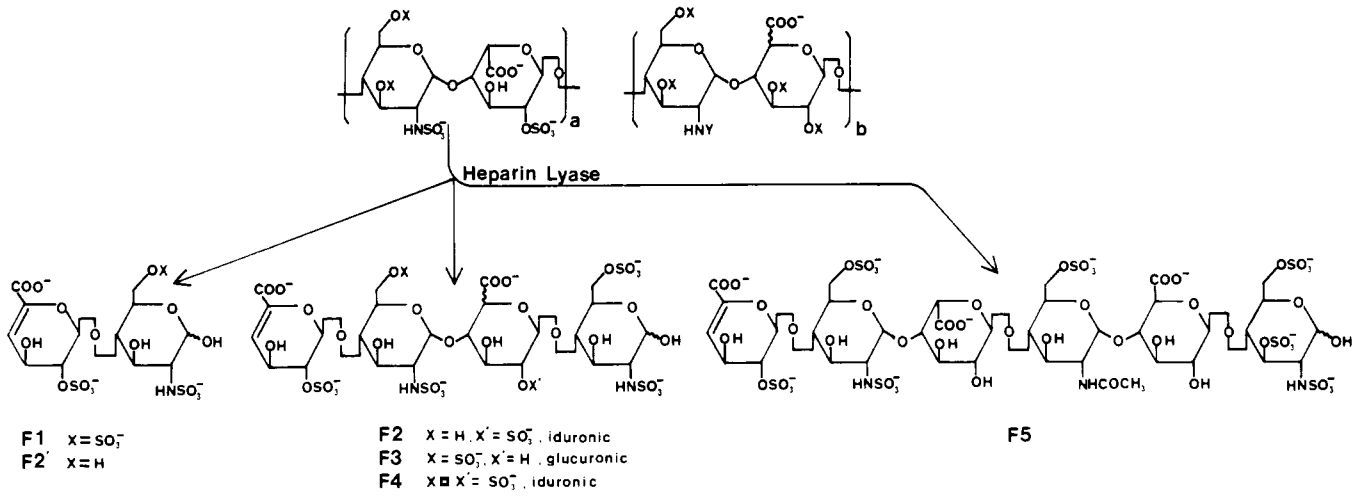

FIGURE 1: Heparin lyase catalyzed depolymerization of heparin to oligosaccharides of known structure. Heparin is represented by a cleavable disaccharide unit (a) and an uncleavable disaccharide unit (b). Heparin lyase acts on adjacent a sites, resulting in two disaccharides, F1 and F2'. It also acts on a sites separated by either one or two b sites, resulting in tetrasaccharide F3 and hexasaccharide F5, respectively. Tetrasaccharides F2 and F4 result from heparin lyase acting on a sites separated by a single a site. F2 and F4 are heparin lyase resistant substrates and are stable under standard reaction conditions. X = SO_3^- or H and Y = SO_3^- or $COCH_3$.

Table III: Experimentally Measured Concentration of Oligosaccharides F1–F5 throughout the Time Course of the Depolymerization Reaction

time (h)	% reaction completion	concentration						wt % of unidentified products
		F1	F2	F2'	F3	F4	F5	
0	0	0[a] (0)[b]	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
1	23	4.12 (6.21)	0.46 (0.38)	0.06 (0.10)	1.24 (1.01)	1.34 (1.01)	0.19 (0.10)	92.6
2	41	13.03 (31.36)	1.39 (0.43)	0.24 (1.13)	2.96 (2.41)	3.86 (2.90)	1.52 (0.83)	77.0
3	55	20.18 (48.56)	1.90 (0.72)	0.41 (1.55)	4.51 (3.67)	6.06 (4.55)	1.81 (1.16)	65.8
4	74	26.45 (63.65)	2.91 (1.23)	0.69 (2.37)	7.28 (5.93)	9.67 (7.27)	3.67 (2.00)	49.3
5	83	34.93 (84.05)	3.64 (1.25)	0.71 (2.96)	8.11 (6.61)	11.05 (8.31)	4.18 (2.28)	51.4
10	100	48.05 (115.62)	4.84 (1.68)	0.95 (3.94)	10.22 (8.32)	11.40 (8.57)	5.77 (3.15)	18.8

[a] Micrograms of oligosaccharide product per 100 μg of heparin. [b] Nanomoles of oligosaccharide product per 100 μg of heparin shown in parentheses.

established by using chemical, enzymatic, and spectroscopic methods (Merchant et al., 1985; Linhardt et al., 1986b; Rice et al., 1987). These structurally defined oligosaccharides represent approximately 90 wt % of the mass of the typical commercial porcine mucosal heparin (Linhardt et al., 1988b). Compositional analysis of 16 porcine mucosal heparins resulted in surprisingly similar distributions of these major oligosaccharide products (Linhardt et al., 1988b). Nine monosaccharides have been reported in the heparin polymer (Bienkowski & Conrad, 1985; Linker & Hovingh, 1984), and seven of these have been found in the heparin lyase prepared oligosaccharides. The mole percentages of each oligosaccharide (and their constituent monosaccharides) found in the porcine mucosal heparin used in these studies are given in Tables I and II. The heparin lyase catalyzed depolymerization of heparin was also monitored kinetically by removing aliquots and analyzing these by SAX-HPLC. The percent reaction completion and the micromoles of the oligosaccharides at each time point are given in Table III.

Both model 1 and model 2 require the simulation of heparin's polydispersity. The average molecular weight of heparin having a simulated polydispersity over a range of "fraction of terminators" was examined. When the fraction of terminators was set at 0.059 for model 2, an average molecular weight of simulated chains of 13 155 was obtained which is comparable to the average molecular weight of 13 000 measured by GPC. The simulated distribution differs from the experimentally observed distribution as it contains both high (M_r >40 000) and low (M_r <5000) molecular weight number chains.

When model 2 was used to construct heparin chains while simulating heparin's polydispersity, the percentage of number

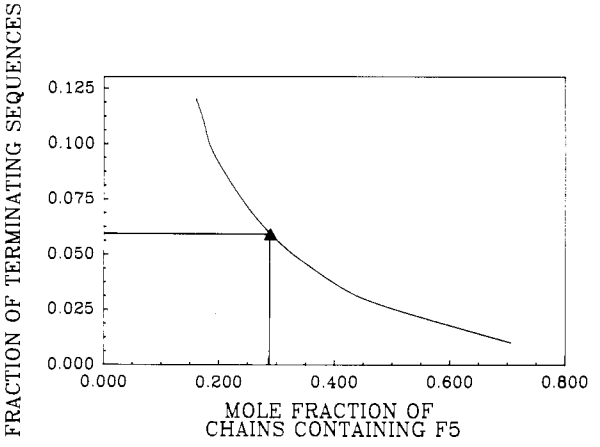

FIGURE 2: Fraction of terminating sequences versus the mole fraction of simulated number chains containing oligosaccharide F5 is plotted. At 0.059 fraction of terminators (▲), the mole fraction of chains containing F5 was 0.289.

chains containing the F5 sequence was determined. The fraction of terminators versus the mole fraction of chains containing F5 was plotted (Figure 2). The results demonstrate that by increasing the fraction of terminators (this decreases the average molecular weight) the percent of number chains containing the F5 sequence decreases. It is interesting to note that at 0.059 fraction of terminators (giving an average molecular weight comparable to that experimentally observed) 28.9% of the heparin number chains contain F5, approximately the same percentage which have high ATIII affinity (Andersson et al., 1976; Hook et al., 1976). Chains with high

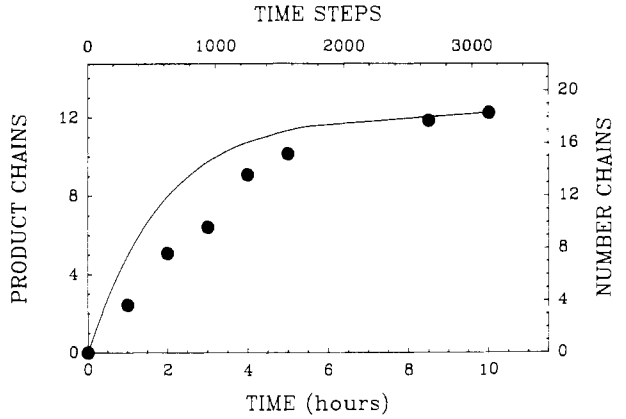

FIGURE 3: Time course of heparin lyase catalyzed depolymerization is plotted. Time versus micromoles of product chains measured experimentally is plotted (●). Time steps versus moles of number chains times a scale factor, obtained by simulation, are plotted as a solid line. The simulated volume for model 2 was 0.629 mL.

ATIII affinity have in their structure specific pentasaccharide binding sites which contain part of an F5 sequence. Recently, our laboratory has also demonstrated that heparin with high ATIII affinity is highly enriched in F5 (Linhardt et al., 1988b).

When simulations are performed, using either model 1 or model 2, the shape of the plot of time steps versus number of number chains (simulated) formed resembles the experimentally observed plot of time versus micromoles of product chains (experimental) formed (Figure 3). This improvement over our earlier simulation (Linhardt et al., 1985) is the result of using Michaelis–Menten kinetics in this simulation.

In model 1, a polydisperse ensemble of number chains was constructed from the individual monosaccharide units at the percentages shown in Table I. Since the percentage of each monosaccharide unit is calculated from the experimentally observed oligosaccharide products, representing 81.2 wt % of heparin, the identity of 18.8 wt % of the monosaccharides was not known. These unidentified monosaccharides were first distributed among the nine monosaccharides at the same percentage at which each of the known monosaccharides was found (Table I, column *a*). The results obtained in this simulation using these input values gave a distribution of oligosaccharides F1 through F5 very dissimilar to the experimentally observed distribution (Tables II and IV, column *a*). The absence of any simulated F5 is striking as this oligosaccharide is found in heparin at 2.22 mol %. Oligosaccharide F5 is of particular interest as it contains a portion of heparin's ATIII binding site. The simulated percentage of disaccharide F1 is low while the percentage of tetrasaccharide F4 [F4 = $(F1)_2$] is high. The total simulated F1 + F4, 79.9 mol %, is similar to the experimental observed 87.9 mol %. In the experimentally observed product distribution, part of the mass cannot be attributed to oligosaccharides F1–F5. Similarly, about 9.9 mol % of the computer-simulated products are found in oligosaccharides [primarily tetrasaccharides (54.2%)] having sequences that differ from oligosaccharides F1–F5.

The simulation of model 1 was performed again, this time using the monosaccharide input percentages obtained by assigning the unidentified monosaccharides equally to residues 3–6 [those low-percentage monosaccharides found in F5 (Table I, column *b*, and Table II)]. The resulting simulated product distribution still contained no number chains corresponding to F5 (Table IV, column *b*). The probability of the low-frequency sugars occurring in the correct sequence found in F5 is small and can be estimated mathematically at less than $1:10^6$. The addition of extra residues of 3–6 results in a greatly

Table IV: Simulated Distribution of Oligosaccharide Products Using Model 1

	simulated	
oligosaccharide product	mol %[a]	mol %[b]
F1	47.76	40.37
F2	1.14	0.44
F2′	2.41	2.14
F3	6.65	7.02
F4	32.13	13.17
F5	0.00	0.00
unidentified	9.90	36.87

[a] Simulation using the mole percent of monosaccharides from Table I, column A. There were 5517 number chains in the simulation. The simulated volume was 0.287 mL. [b] Simulation using the mole percent of monosaccharides from Table I column B. There were 5056 number chains in the simulation. The simulated volume was 0.282 mL.

increased number of unidentified products (Table IV, column *b*) rather than an increased amount of F5. The failure of model 1 to closely simulate the experimentally observed distribution of oligosaccharides F1 through F5 strongly indicates that heparin is not a polymer simply comprised of a random assortment of its constituent monosaccharide units.

Although model 1 suggests that the final oligosaccharides could not have resulted from the random assembly of monosaccharides, it is still possible that the oligosaccharide components of heparin themselves are arranged randomly. In model 2, a polydisperse ensemble of number chains is constructed, by computer-simulation, from the random assembly of final oligosaccharide products (F1–F5). For an oligosaccharide such as F5, which is resistant to the action of heparin lyase, the input concentration should be identical with the simulated concentration of F5 in the final product mixture. Thus, in order to obtain information regarding the disposition of F5 in the polymer, it is necessary to examine product distributions obtained at various times throughout the depolymerization reaction. Oligosaccharides F2 and F2′, as well as F4 and F1, pose another interesting problem. Disaccharides F2′ and F1 can arise from tetrasaccharides F2 and F4, respectively, only when these tetrasaccharides are present in the intact polymer. Both F2 and F4 are stable under the reaction conditions used in this study. A simple example is the case of an octasaccharide chain composed entirely of trisulfated disaccharide units. Such a chain could give rise to either only F4 (i.e., F4 ↛ F4) by cleaving in the center of the chain or a mixture of F1 and F4 by cleaving at the ends of the chain (i.e., F1 ↛ F4 ↛ F1 or F1 ↛ F1 ↛ F4 or F4 ↛ F1 ↛ F1). A calculation for this simple example assuming no enzyme preference for any particular site (random endolytic cleavage) is 50 mol % F1 and 50 mol % F4. The final simulated ratios of F1 to F4 and F2′ to F2 are dependent both on the input percentages for the oligosaccharides and on the assumed random endolytic action pattern of heparin lyase.

The results obtained in the simulation of model 2 demonstrate that the simulated final concentrations of oligosaccharide F3 and F5 are, as expected, identical with the experimentally observed values (the same as those inputted) (Figure 4b). The simulated and observed concentrations of both of these oligosaccharide products are very similar throughout the entire time course of the depolymerization reaction. The final simulated concentration of disaccharide F1 is lower than the experimentally observed value while the converse is true for tetrasaccharide F4 (Figure 4a). The final simulated concentration of F1 + 2F4 − F2′ is equal to the concentration input for F1 (when an F2 site in the chain is cleaved by heparin lyase, 1 mol of F2′ and 1 mol of F1 are formed) while the final simulated concentration of F2 + F2′ is equal to the concen-

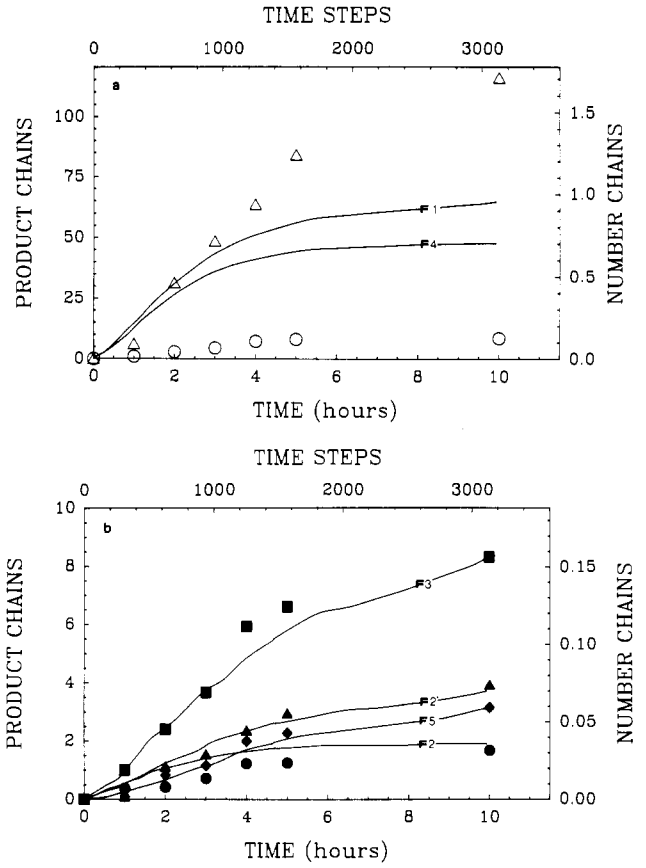

FIGURE 4: Time course of oligosaccharide formation. (a) For experimentally measured concentrations of oligosaccharides F1 (△) and F4 (○), time versus micromoles of product chains is plotted. For simulated concentrations of F1 and F4, time steps versus moles of number chains times a scale factor are plotted as solid lines. (b) For experimentally measured concentrations of oligosaccharides F2 (●), F2′ (▲), F3 (■), and F5 (♦), time versus micromoles of product chains is plotted. For simulated concentrations of F2, F2′, F3, and F5, time steps versus moles of number chains times a scale factor are plotted as solid lines.

tration input for F2. The observation of a simulated final concentration of F1 which is lower than the experimentally determined concentration suggests that heparin lyase is not completely random in its selection of sites. The enzyme instead appears to be preferentially cleaving the F4 sites within the heparin polymer.

Heparin is synthesized as a proteoglycan molecule consisting of 10 or more polysaccharide chains linked to a peptide core (Lindahl et al., 1986). The linkage to protein involves a specific neutral oligosaccharide (Roden, 1980). The glycosaminoglycan chains grow from these linkages by alternate addition of 2-acetamido-2-deoxyglucopyranose and D-glucopyranosyluronic acid through the action of appropriate transferases (Roden, 1980). The chains of the nonsulfated, high molecular weight, heparin precursor are finally substituted by a series of enzymes sequentially performing N-deacetylation and N-sulfation of 2-amino-2-deoxyglucopyranose (Riesenfield et al., 1980), C-5 epimerization of D-glycopyranosyluronic acid, sulfation at O-2 of the resulting L-idopyranosyluronic acid (Jacobsson et al., 1979), and, finally, sulfation at O-6 of 2-amino-2-deoxy-D-glucopyranose (Jacobsson & Lindahl, 1980). The reaction sequence implies that the final product is invariably a polymer of trisulfated disaccharide (Poly F1). The structural heterogeneity of heparin therefore indicates that the polymer modification process is incomplete. The product of each enzymatic reaction constitutes the specific substrate for the succeeding enzyme; thus, the biosynthesis of heparin is not a random process (Horner & Young, 1982). Sulfation, for example, occurs preferentially in those regions of the chain where the amino sugar residues have been deacetylated and N-sulfated, and where D-glucopyranosyluronic acid has been epimerized to L-idopyranosyluronic acid. Not all of the monosaccharides are satisfied in terms of substrate specificity of succeeding enzymes. If the 2-acetamido-2-deoxyglucopyranose units are not deacetylated, they will not become N-sulfated, which in turn affects the following modification steps, such as the epimerization and O-sulfation (Lindahl et al., 1986). This is particularly important in understanding heparin's microheterogenicity. The mechanism regulating the extent and distribution of N-deacetylation is at present unknown. The results of the simulation performed using model 1 are in complete agreement with the biosynthetic data which suggest that heparin biosynthesis is an ordered process. The results of the simulation using model 2 are consistent with the heparin polymer being composed of ordered regions consisting of oligosaccharides F1 through F5. The less than random behavior of heparin lyase in selecting its sites makes it difficult to draw firm conclusions on whether or not these ordered regions are themselves actually randomly arranged.

The molecular weight (M_r 60 000–100 000) of the polysaccharide chains in the heparin proteoglycan is significantly higher than that (M_r 5000–25 000) of glycosaminoglycan heparin. After completion of the biosynthetic modifications at the polymer level, the polysaccharide chains and the polypeptide chains are cleaved to smaller fragments, through the concerted action of endo-β-glucuronidase and proteinases (Thunberg, 1982). The results of these simulation studies suggest it is possible to construct a polydisperse heparin number chain by the introduction of the appropriate fraction of terminators. In addition to giving a polydisperse collection of number chains of appropriate average molecular weight, this simulation correctly predicts the fraction of chains which contain ATIII binding sites. The fraction of terminators might, therefore, simulate the sites at which the endo-β-glucuronidase cleaves proteoglycan heparin during its processing.

Heparin with high ATIII affinity is processed under somewhat different control. The minimum ATIII binding site in porcine heparin is a pentasaccharide containing a 3-O-sulfated glucosamine residue in the center of the chain. The 3-O-sulfation reaction occurs after 6-O-sulfation in the last biosynthetic polymer modification process. This proposed biosynthetic pathway is supported by data gathered in a cell-free mouse mastocytoma system (Riesenfeld et al., 1983). It has been proposed that the ATIII binding regions are not randomly distributed (Horner & Young, 1982) but, rather, are accumulated in a limited proportion of the chains assembled in a heparin proteoglycan molecule. Recent studies also suggest that the ATIII binding sites are located close to the nonreducing terminus (more than 20 sugar residues from where the carbohydrate is attached to the protein core) in proteoglycan heparin (Rosenfeld & Danishefsky, 1988). The data presented in this study are consistent with the random distribution of ATIII binding sites within glycosaminoglycan heparin. The ordered arrangement of ATIII binding sites might be lost during endo-β-glucuronidase processing of proteoglycan heparin. Further simulation studies will be required on proteoglycan heparin to test this hypothesis.

ACKNOWLEDGMENTS

We thank F. N. Choong, J. Vogel, C. F. Nyen, H. M. Wang, and D. Loganathan for their contributions to this research.

Registry No. Heparin, 9005-49-6.

REFERENCES

Andersson, L. O., Barrowcliffe, T. W., Holmer, E., Johnson, E. A., & Sims, G. E. C. (1976) *Thromb. Res. 9*, 575–583.

Bienkowski, M. J., & Conrad, H. E. (1985) *J. Biol. Chem. 260*, 356–365.

Casu, B. (1985) *Adv. Carbohydr. Chem. Biochem. 43*, 51–143.

Cohen, D. M. (1989) in *Proceedings of the 2nd Symposium on the Frontiers of Massively Parallel Computation*, IEEE Press, Piscataway, NJ (in press).

Cooper, D., & Clancy, M. (1985) *Oh! Pascal!*, 2nd ed., W. W. Norton & Co., New York.

Höök, M., Björk, I., Hopwood, J., & Lindahl, U. (1977) *FEBS Lett. 66*, 90–93.

Horner, A. A., & Young, E. (1982) *J. Biol. Chem. 257*, 8749–8754.

Jacobsson, I., & Lindahl, U. (1980) *J. Biol. Chem. 255*, 5094–5100.

Jacobsson, I., Backström, G., Höök, M., Lindahl, U., Feingold, D. S., Malmström, A., & Rodén, L. (1979) *J. Biol. Chem. 254*, 2975–2982.

Johnson, L. W., & Riess, R. D. (1982) *Numerical Analysis*, 2nd ed., Addison-Wesley, Reading, MA.

Knuth, D. E. (1981) *The Art of Computer Programming, Volume 2, Seminumerical Algorithms*, 2nd ed., Addison-Wesley, Reading, MA.

Lindahl, U., Feingold, D. S., & Roden, L. (1986) *Trends Biochem. Sci. 11*, 221–225.

Linhardt, R. J., Fitzgerald, G. L., Cooney, C. L., & Langer, R., (1982) *Biochim. Biophys. Acta 702*, 197–203.

Linhardt, R. J., Merchant, Z. M., Rice, K. G., Kim, Y. S., Fitzgerald, G. L., Grant, A. C., & Langer, R. (1985) *Biochemistry 24*, 7805–7810.

Linhardt, R. J., Cooney, C. L., & Galliher, P. M., (1986a) *Appl. Biochem. Biotechnol. 12*, 135–177.

Linhardt, R. J., Rice, K. G., Merchant, Z. M., Kim, Y. S., & Lohse, D. L. (1986b) *J. Biol. Chem. 261*, 14448–14454.

Linhardt, R. J., Rice, K. G., Engelken, J., & Weiler, J. (1988a) *J. Biol. Chem. 263*, 13090–13096.

Linhardt, R. J., Rice, K. G., Kim, Y. S., Lohse, D. L., & Loganathan, D. (1988b) *Biochem. J. 254*, 781–787.

Linker, A., & Hovingh, P. (1984) *Carbohydr. Res. 127*, 75–94.

Merchant, Z. M., Kim, Y. S., Rice, K. G., & Linhardt, R. J. (1985) *Biochem. J. 229*, 369–377.

Park, S. K., & Miller, K. W. (1988) *Commun. ACM 31*, 1192–1201.

Rice, K. G. (1988) Doctoral Thesis, University of Iowa.

Rice, K. G., Kim, Y. S., Grant, A. C., Merchant, Z. M., & Linhardt, R. J. (1985) *Anal. Biochem. 150*, 325–331.

Rice, K. G., Rottink, M. K., & Linhardt, R. J. (1987) *Biochem. J. 244*, 515–522.

Riesenfeld, J., Höök, M., & Lindahl, U. (1980) *J. Biol. Chem. 255*, 922–928.

Riesenfeld, J., Thunberg, L., & Lindahl, U. (1983) in *Glycoconjugates. Proceedings of the 7th International Symposium on Glycoconjugates* (Chester, M. A., Heinegard, D., Lundblad, A., & Svensson, S., Eds.) pp 390–391, Lund, Sweden.

Roden, L. (1980) in *The Biochemistry of Glycoproteins and Proteoglycans* (Lennartz, W. J., Ed.) pp 267–371, Plenum, New York.

Rosenfeld, L., & Danishefsky, I. (1988) *J. Biol. Chem. 263*, 262–266.

Thunberg, L., Backström, G., Wasteson, A., Robinson, H. C., Ogren, S., & Lindahl, U. (1982) *J. Biol. Chem. 257*, 10278–10282.

Yang, V. C., Linhardt, R. J., Bernstein, H., Cooney, C. L., & Langer, R. (1985) *J. Biol. Chem. 260*, 1849–1857.