J. Med. Chem. **2004,** *47,* 277−280

Highly Diverse Heparan Sulfate Analogue Libraries: Providing Access to Expanded Areas of Sequence Space for Bioactivity Screening

Edwin A. Yates,*,† Scott E. Guimond, and Jeremy E. Turnbull

School of Biosciences, University of Birmingham, Birmingham, B15 2TT, U.K.

Received July 28, 2003

Structurally diverse heparan sulfate analogue libraries were produced chemicoenzymatically from heparin. They possess vastly more heterogeneity than tissue heparan sulfates, expand the sequence space available for screening, and can help identify minimal structural features associated with activity. Library components are likely to exhibit fewer nonspecific interactions and side-effects than heparin or simple chemically modified heparin. A strategy for their use is illustrated for the fibroblast growth factor−receptor tyrosine kinase signaling system.

Introduction

Heparan sulfate (HS) is a family of naturally occurring, sulfated, linear polysaccharides which has become the focus of extensive interest because of its ability to interact with a large number of proteins and involvement in many biological processes. These proteins are linked to many clinical conditions, including cancer, inflammation, angiogenesis, wound and bone healing, cardiovascular disorders, and neurodegenerative disease. HS is also involved in host−pathogen interactions such as herpes virus entry, HIV infectivity, and *Chlamydia* host attachment. Importantly, there is mounting evidence that distinct HS structures are recognized by proteins involved in these processes, suggesting they have potential as therapeutic agents[1] and represent a largely untapped source of active compounds.

Heparin, a widely used pharmaceutical that can be considered a form of HS, is an abundant polysaccharide with the same underlying backbone structure. It is generally more heavily sulfated and more homogeneous and has been used as a source of active saccharides. It is often a good activator in many pathways where HS is involved, but the preponderance of highly sulfated structures results in a loss of regulatory capacity that HS provides in vivo and can lead to unwanted or nonspecific activities. HS itself could be used but is of restricted sequence diversity and relatively scarce, especially compared to heparin, and active sequences tend to be present in relatively low quantities.

One approach that has been widely used is that of modifying heparin by chemical desulfation, usually undertaken with the aim of creating simplified structures with which to correlate activity. Modifications are carried to completion and have only rarely been performed in combination, severely limiting the range of sequences attainable. An alternative is the addition of *O*/*N*-sulfate and *N*-acetyl groups to a previously desulfated backbone (or biosynthetic precursor). Unfortunately, the addition of *O*-sulfates directly by chemical means cannot be easily achieved selectively at those positions which are substituted in the naturally occurring material.[2] Biosynthetic modifications have been carried out[3,4] but drawbacks are the limited quantity of material that can be produced and concerns over the level of heterogeneity attainable for library production. This latter point arises from the specificities of the enzymes and, while naturally occurring sequences from HS may be realistic targets, it is not clear how the whole of HS sequence space (including nonnaturally occurring sequences) may be accessed. The chemical synthesis of HS/heparin saccharides is feasible[5] for identified targets, and while combinatorial techniques are being developed, this remains a challenge not least because of the very large numbers of possible sequences. Here, we present a complementary approach, in which a very wide range of sequences can be generated quickly. Following fractionation, this permits sublibraries of oligosaccharides to be screened for bioactivity.

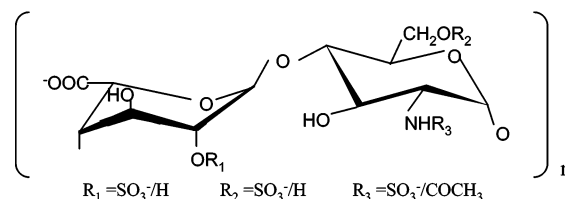

Figure 1. General disaccharide repeating structure of chemically modified heparin derivatives. For clarity, the uronic acid moiety is shown as α-L-idoA but can also be its C-5 epimer, β-D-GlcA.

The Basis of Heterogeneity in the HS Analogue Library Products. If the repeating disaccharide of HS is considered (in a simplified form, ignoring glucuronic acid, 3-*O*-sulfates, and the presence of free amino groups, Figure 1), then there are three positions (R_1, R_2, and R_3) at which substitutions can be made if naturally occurring monosaccharides are considered. These are sulfate or hydroxyl at position 2 of iduronate or position 6 of glucosamine and *N*-sulfate or *N*-acetyl at position 2 of glucosamine. For an idealized disaccharide, there are $2^3 = 8$ possible combinations of substitution patterns i.e., 8 disaccharides. Extending the argument to hexasaccharides, there are $8^3 = 512$ for octasaccharides and decasaccharides, 4096 and 32768 combinations, respectively. Such vast diversity could only be attained if it were possible to modify the

* To whom correspondence should be addressed.
† Current address: School of Biological Sciences, Biosciences Building, University of Liverpool, Crown Street, Liverpool L69 7ZB UK tel 44-(0)151-795-4429 Fax 44-(0)-151-4414 email ed_yates@hotmail.com.

10.1021/jm0309755 CCC: $27.50 © 2004 American Chemical Society
Published on Web 12/02/2003

substitution positions in equal proportion, and, in such a case, a statistical distribution of oligosaccharides would be expected. Although HS is more structurally diverse than heparin, it has recently been shown that pools of decasaccharides derived from HS can be fractionated into perhaps several hundred species;[6] however, if there are $>10^4$ theoretically possible decasaccharides, this still represents a very small fraction of those potentially available.

A common approach when using heparin as a model compound has been the removal of sulfate groups from one or more positions, usually to completion to create simplified polysaccharides for use in structure–activity studies.[7–10] However, an example of a polysaccharide in which one modification (partial de-O-sulfation) was made to partial extent exhibited NMR spectral characteristics consistent with the modifications having occurred randomly rather than in blocks[11] (the former gives rise to more complex signal splitting, the latter to relatively simple signals). Sequence permutations based on statistical distribution can therefore be expected, opening-up the possibility of creating large numbers of structurally diverse sequences. This approach is distinct in that it aims to maximize heterogeneity in the polysaccharides rather than simplify the structure and achieves this by the consecutive application of several partial reactions. It results in high structural diversity in terms of combinations of constituent monosaccharides, thereby greatly expanding the available sequence space. Such polysaccharides can be screened for activity but are more usefully employed as sources of structurally diverse oligosaccharides.

An Illustration of the Generation of Structurally Diverse Oligosaccharide Libraries. The generation of diverse HS analogue libraries from such a heterogeneous polysaccharide is illustrated in Figure 2. The electrophoresis profile of a partial digestion with heparitinase II of the polysaccharide is shown (panel A). The products were first fractionated on the basis of their hydrodynamic volume (panel B). This profile is similar to, if less well-resolved than, that from a typical enzymatic digestion of heparin or HS. However, when peaks corresponding to particular hydrodynamic volume ranges, in this case DP12 of bovine lung heparin derived standards, were further fractionated on the basis of overall charge by high performance anion exchange chromatography (HPAEC), (panel C), a distinct pattern was observed. Instead of a range of separable peaks,[6] typical of a modest number of saccharides, the overall chromatogram was bound by an approximately Gaussian envelope, inside of which were discrete, regularly spaced peaks. This is the expected appearance of a chromatographic separation of an oligosaccharide pool (defined by hydrodynamic volume) containing charged groups which are distributed in an essentially random way. The appearance is typical of HPAE chromatograms of gel chromatography fractions from enzyme digestions of this polysaccharide. The heterogeneity of each of these peaks is further demonstrated by their profile on an electrophoresis gel (examples labeled X, Y, and Z in Figure 2 (panel D)) which separates them on the basis of a combination of charge, size, and conformation. The appearance of these diffuse bands (which, because of their lower overall sulfation levels, run higher up the

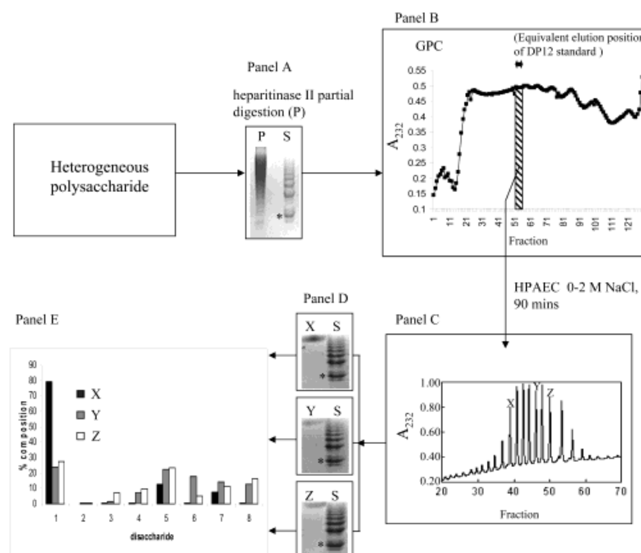

Figure 2. Generation of oligosaccharide library components from a heterogeneous polysaccharide starting material. Clockwise: Panel A; electrophoresis of heparitinase II digestion of the heterogeneous polysaccharide (P) compared to that of bovine lung heparin standard (S), which is comparatively homogeneous giving a characteristic ladder: Panel B; gel chromatography separation of digest (P) on Sephadex G-50 also showing equivalent elution position of a standard DP 12 oligosaccharide pool from (S): Panel C; HPAEC separation (0–2 M NaCl, pH 7, 90 min) of the fraction of (P) which elutes at the same position as a bovine lung heparin DP 12 standard: Panel D; electrophoresis profiles of 3 example peaks from the HPAEC trace, X, Y, and Z, compared to the standard ladder derived from bovine lung heparin (S): Panel E; Disaccharide compositional analysis of peaks X, Y and Z. Disaccharides: 1; UA-GlcNAc, 2; UA-GlcNAc(6S), 3; UA-GlcNS, 4; UA-GlcNS(6S), 5; UA(2S)-GlcNS, 6; UA(2S)-GlcNS(6S), 7; UA(2S)-GlcNAc, 8; UA(2S)-GlcNAc(6S)

gel), differs from their more highly charged and homogeneous counterparts derived from bovine lung heparin (standards, S in panel D). It is clear that the standards run as tighter bands (individual peaks from an HPAEC separation of HS derived oligosaccharides also run as well-defined bands) and this is especially evident for those larger than DP 6. These data together with the composition analysis of peaks X, Y, and Z from HPAEC (panel E), in which no simple stoichiometric relationships are evident, despite their constituents having the same size and charge, suggests that each discrete peak on the HPAEC trace contains complex mixtures of oligosaccharides forming sublibraries of oligosaccharides.

Selecting Active Structures Approaching Minimum Size and Charge. An illustration of the use of the library to select active oligosaccharide sets (or sublibraries) with minimum size and charge is shown in Figure 3. This consists of a fibroblast growth factor-tyrosine kinase receptor (FGF/FGFR) system in an in vitro cell assay, in which the ability of fractions to support signaling with FGF-1/R2c is measured. Testing the activities of fractions from the partial heparitinase digestion separated by gel chromatography (panel A) allows a pool of oligosaccharides to be selected while minimizing size and charge. It is noteworthy that higher hydrodynamic volume does not necessarily bestow higher activity, (panel A for three fractions denoted B, D, and I). Further separation, on the basis of charge of

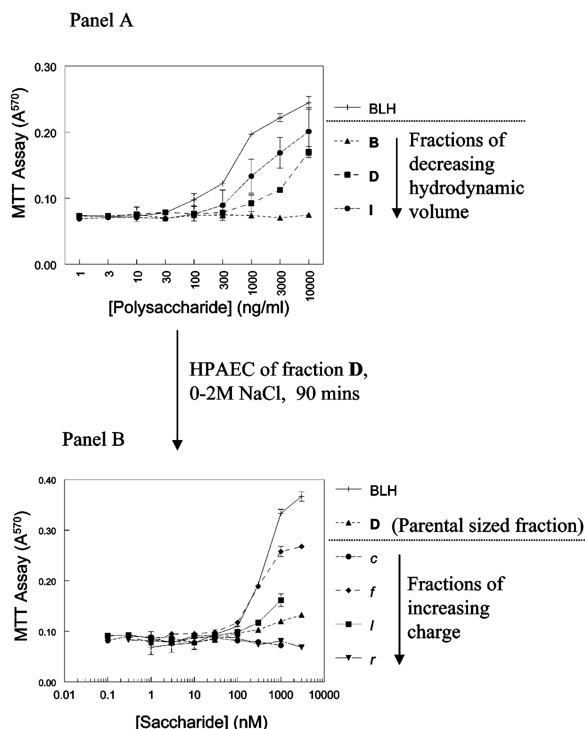

Figure 3. The process of selecting active oligosaccharides, approaching minimum structural complexity, and capable of forming an active signaling complex between FGF1 and receptor 2c. Heterogeneous polysaccharide starting material was partially digested, and the products were fractionated into oligosaccharide fractions A−O (in order of decreasing hydrodynamic volume) by GPC. Panel A; activity assay of FGF1/R2c in BaF cells with representative, sized oligosaccharide pools B, D, and I of increasing hydrodynamic volume from the GPC separation of the heterogeneous polysaccharide digestion. The activity of bovine lung heparin (polysaccharide) is also shown as a positive control. Panel B; from these fractions, the smallest active fraction (D) was further separated by HPAEC into fractions a−t (in order of increasing anionic charge) and tested. The activity of representative samples c, f, l, and r are shown for signaling of FGF1/R2c. The activities of the parent oligosaccharide pool (D) and bovine lung heparin (BLH) are also shown.

the smallest significantly active fraction, in this case D, by HPAEC and subsequent testing allows the search to be focused. Higher charge does not necessarily correlate with higher activity as illustrated by the activities of HPAE fractions c, f, l, and r (panel B). Fractions exhibiting both higher (e.g., f) and lower activity (e.g., r) than the parent (D) can be identified, indicating that a degree of specificity is present in FGF/FGFR/HS interactions, confirming previous studies with naturally occurring HS derived fractions.[4,5,7] It should also be noted that (polymeric) heparin (positive control), is likely to appear a disproportionately effective activator compared to oligosaccharides because it possesses many more active sites. Additional iterations of the separation and screening process will allow increasingly focused structure−activity relationships to be determined.

Conclusion

The approach presented here offers a powerful tool in the search for biologically active analogues of the natural product heparan sulfate. The libraries derived from the modified polysaccharide approach maximum heterogeneity much more closely than either HS or heparin and are much more heterogeneous than simple chemically modified heparin derivatives. They are clearly a rich source of active sequences corresponding to those found naturally, as well as more numerous and hitherto unavailable (nonnaturally occurring) structures. Furthermore, the use of these libraries reduces the inherent structural bias associated with those previously generated although some remains due to the choice of degradation technique, whether heparitinase (as here), or nitrous acid degradation. Essentially random degradation with free-radical processes could alleviate this.[15] Fractionation of the oligosaccharide libraries permits the search for structures with defined levels of activity, while approaching minimum size and charge, hence facilitating the search for structure−activity relationships (SARs). Active components found in this way may exhibit fewer additional activities, likely to lead to side-effects, if subsequently taken forward as novel lead compounds in a pharmaceutical context.

One advantage of this approach is that active saccharides can be isolated through affinity techniques which avoids selecting those which bind most avidly (i.e., the largest and most charged). The isolation of such sequences may obscure more interesting subtleties and the present approach circumvents this. A fair competition between saccharides of similar size and charge can be established, e.g., between pools of saccharides with the same charge (i.e., eluting at the same salt concentration on HPAEC) and of the same hydrodynamic volume (by GPC). Binding saccharides can then be tested for activity. This approach is equally applicable to a number of starting materials, including HS, if available in sufficient quantity. Here, porcine intestinal mucosal heparin was chosen because of its abundance, as well as the presence of significant levels of glucuronic acid,[16] making it structurally closer to HS than some other heparin preparations. This method also allows searches for structure−function relationships over a much larger area of sequence space with low inherent bias. The compounds produced are composed of naturally occurring monosaccharide units and are compatible with enzymatic and chemical treatments as well as other analytical procedures.[17] HS (as heparin) is a widely used therapeutic, principally as an anticoagulant, acting via antithrombin through a unique pentasaccharide sequence with a high degree of specificity. It also functions as an antithrombotic, binding a longer, more highly sulfated region with lower specificity, but for which a minimum size fragment is required. Many other possible therapeutic uses of heparin are becoming apparent[1] and one application of the approach outlined here is the search for oligosaccharides possessing a desired activity, which may include, for example, anticoagulation or angiogenesis, while minimizing unwanted effects, such as heparin-induced thrombocytopenia (HIT). The approach is clearly distinct from those, in which either heparin or simple chemically modified derivatives are used, and permits several orders of magnitude more sequences to be tested for activity. The method also differs from those employing biosynthetic techniques in that more sequences can be produced and is readily scaleable to multigram quantities. It will facilitate downstream SAR studies with a large number

of new HS-analogue structures, which are currently difficult to access through conventional chemistry because of their numbers. Once active structures have been identified, synthetic methodology could produce larger quantities of well-defined target structures. The method has been illustrated here by generating oligosaccharide libraries approaching maximum structural diversity. These may help to stimulate the development of more advanced separation techniques. The degree of specificity involved in most of the interactions in which HS is involved remains largely unknown and these libraries will help to address this question.

Experimental Section

1. Preparation of Heterogeneous Polysaccharide. (a) Partially De-O-, Completely De-N-sulfated Heparin. Porcine intestinal mucosal heparin (Celsus Labs, Cincinatti, OH, 5 g) was converted to the pyridinium salt (Dowex W-50 resin, H^+ form), neutralized (pyridine) and freeze-dried (4.9 g). This was then suspended (in DMSO/MeOH, 9/1, v/v, 100 mL) and heated (80 °C, 24 h; determined empirically by NMR). The product was analyzed[11,12] to confirm its structural heterogeneity. **(b) Partial Re-N-acetylation**. Partial re-N-acetylation was as described[11] ($(CH_3CO)_2O$) in saturated aqueous $NaHCO_3$). Its extent, determined empirically (NMR), was limited (50% ± 10%) by controlling the quantity of $(CH_3CO)_2O$. **(c) Re-N-sulfation of Remaining Unsubstituted Amino Groups.** Free-amino groups were re-N-sulfated (twice) as described[13] and confirmed by NMR, disaccharide analysis: UA-GlcNAc; 24.5%, UA-GlcNAc(6S); 13.7%, UA-GlcNS; 7.0%, UA-GlcNS(6S); 13.0%, UA(2S)-GlcNS; 13.7%, UA(2S)-GlcNS-(6S); 13.6%, UA(2S)-GlcNAc; 11.4%, UA(2S)-GlcNAc(6S); 3.1%. Overall substitution is iduronate 2-S 41.8%, glucosamine 6-sulfate 43.4%, and glucosamine *N*-sulfate 47.3%.

2. Characterization of Polysaccharide. (i) NMR. The effectiveness of the chemical treatments were monitored by ^1H and ^{13}C NMR at 500 and 125 MHz (D_2O, 27 °C). Chemical shifts (external standard) were assigned by reference to well-defined model compounds.[11] **(ii) Disaccharide Analysis Following Exhaustive Digestion with Heparitinases I, II, and III.** Samples (100 μg) exhaustively digested with heparitinase I, II, and II (Seikagaku, lyase buffer, 37 °C; 500 mM NaOAc, 2.5 mM $Ca(OAc)_2$, pH 7). Subsequent comparison with disaccharide standards following separation by HPAEC (Propac PA-1 column, 4 × 250 mm, 0−2 M NaCl, 90 min, 232 nm) allowed quantification.

3. Partial Degradation with Heparitinase II. The polysaccharide (50 mg) was partially digested with heparitinase II (Seikagaku) in lyase buffer (as above, 37 °C) and was stopped (100 °C, 5 min) when a range of products was detected (by electrophoresis).

4. Fractionation of Products by Gel Permeation Chromatography. Partially digested products were separated by hydrodynamic volume (Sephadex G-50, 2.5 cm × 1.75 m, in 100 mM NH_4HCO_3, detection; 232 nm). Calibration (before and after separation) was with a pairwise ladder of heparin oligosaccharides (partial heparitinase III (heparinase I) digestion). Fractions (denoted A−O) were desalted, quantified (A_{232}), and tested in a number of assays.

5. Fractionation of Hydrodynamic Volume Defined Products by HPAEC. Selected fractions from gel permeation chromatography were desalted and fractionated on HPAEC on a Propac PA-1 column (4 × 250 mm, 0−2 M NaCl gradient, 90 min, detection; 232 nm). One fraction, corresponding in hydrodynamic volume to a DP 12 standard oligosaccharide derived from bovine lung heparin, was subjected to further fractionation (HPAEC), and peaks were collected (selected peaks denoted X, Y, and Z for use in the experiments are shown in Figure 2). Following HPAEC of fraction (D) from the gel permeation chromatography separation, another set of fractions were produced (denoted a−t), desalted, and quantified (A_{232}) for subsequent analysis and testing (Figure 3).

6. BaF3 Cell Assay with FGFs and FGFRs. BaF3 cells transfected with the appropriate receptor were maintained in RPMI-1640 supplemented with 10% fetal calf serum, 2 mM L-glutamine, 100 U mL^{-1} pen G, 50 μg mL^{-1} streptomycin sulfate, and 2 ng mL^{-1} IL-3. Assays for saccharide function were as described.[14] Saccharides were from gel chromatography (denoted **A** to **O**), between 1.0 ng mL^{-1} and 10000 ng mL^{-1} and from HPAEC (denoted a−t), between 0.1 and 3000 nM.

Acknowledgment. The authors thank the BBSRC (E.A.Y. and J.E.T.) and MRC (S.E.G. and J.E.T. (senior fellowship)) for funding. BaF 3 cells were kindly supplied by D.Ornitz (Washington University, St.Louis, MO). Laine Wallace is thanked for technical assistance.

References

(1) Lever, R.; Page, C. P. Novel Drug Development Opportunities for heparin. *Nature Rev. Drug Discovery* **2002**, *1*, 140−148.
(2) Naggi, A.; Torri, G.; Casu, B.; Oreste, P.; Zoppetti, G.; Li, J. P.; Lindahl, U. Toward a biotechnological heparin through combined chemical and enzymatic modification of the *Escherichia coli* K5 polysaccharide. *Seminar Thromb. Haemostasis* **2001**, *27*, 437−443.
(3) Wu, Z.; Zhang, L.; Beeler, D. L.; Kuberan, B.; Rosenberg, R. D. A new strategy for defining critical functional groups on heparan sulfate. *FASEB. J.* **2002**, *16,* 539−545.
(4) Jemth, P.; Kreuger, J.; Kusche-Gullberg, M.; Sturiale, L.; Giminez-Gallego, G.; Lindahl, U. Biosynthetic Oligosaccharide Libraries for Identification of Protein-binding Heparan Sulfate Motifs. *J. Biol. Chem.* **2002**, *277*, 30567−30573.
(5) Grootenhuis, P. D.; Westerduin, P.; Meuleman, D.; Petitou, M.; van Boeckel, C. A. Rational design of synthetic heparin analogues with tailor-made coagulation factor inhibitory activity. *Nature Struct. Biol.* **1995**, *2*, 736−739.
(6) Guimond, S. E.; Turnbull, J. E. Fibroblast growth factor signaling is dictated by specific heparan sulfate saccharides. *Curr. Biol.* **1999**, *9*, 1343−1346.
(7) Irie, A.; Yates, E. A.; Turnbull, J. E.; Holt, C. E. Specific heparan sulfate structures are involved in retinal axon targetting. *Development* **2002**, *129,* 61−70.
(8) Guimond, S.; Maccarana, M.; Olwin, B. B.; Lindahl, U.; Rapraeger, A. C. Activating and inhibitory heparin sequences for FGF-2 (basic FGF). Distinct requirements for FGF-1, FGF-2 and FGF-4. *J. Biol. Chem.* **1993**, *268,* 23906−23914.
(9) Belford, D. A.; Hendry, I. A.; Parish C. R. Ability of different chemically modified heparins to potentiate the biological activity of heparin-binding growth factor-receptor-heparin interactions. *Biochemistry* **1992**, *31*, 6498−6503.
(10) Ishihara, M.; Tyrrell, D. J.; Stauber, G. B.; Brown, S.; Cousens, L. S.; Stack, R. J. Preparation of affinity-fractionated, heparin derived oligosaccharides and their effects on selected biological activities mediated by basic fibroblast growth factor. *J. Biol. Chem.* **1993**, *268,* 4675−4683.
(11) Yates, E. A.; Santini, F.; Guerrini, M.; Naggi, A.; Torri, G.; Casu, B. 1-H and 13-C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. *Carbohydr. Res.* **1996**, *294*, 15−27.
(12) Inoue, Y.; Nagasawa, K. Selective N-desulfation of heparin with dimethyl sulfoxide containing water or methanol. *Carbohydr. Res.* **1976**, *46,* 87−95.
(13) Lloyd, A. G.; Embrey, G.; Fowler, L. J. Studies on heparin degradation. I. Preparation of (^{35}S) sulphamate derivatives for studies on heparin degrading enzymes of mammalian origin. *Biochem. Pharmacol.* **1971**, *20*, 637−648.
(14) Ornitz, D. M.; Xu, J.; Colvin, J. S.; McEwen, D. G.; MacArthur, C. A.; Coulier, F.; Gao, G.; Goldfarb, M. Receptor specificity of the fibroblast growth factor family. *J. Biol. Chem.* **1996**, *271,* 15292−15297.
(15) Liu. Z.; Perlin, A. S. Evidence of a selective free radical degradation of heparin, mediated by cupric ion. *Carbohydr. Res.* **1994**, *255*, 183−191.
(16) Guo, Y. C.; Conrad. H. E. The disaccharide composition of heparins and heparan sulfates. *Anal. Biochem.* 1989, *176*, 96−104.
(17) *Proteoglycan protocols. Methods in Molecular Biology Series*; Iozzo, R. V., Ed.; Humana Press: Totowa, NJ, 2001; Vol. 271.

JM0309755