Divergent Synthesis of 48 Heparan Sulfate-Based Disaccharides and Probing the Specific Sugar–Fibroblast Growth Factor-1 Interaction

Yu-Peng Hu,†‡# Yong-Qing Zhong,†‡# Zhi-Geng Chen,†‡# Chun-Yen Chen,‡ Zhonghao Shi,† Medel Manuel L. Zulueta,† Chiao-Chu Ku,§ Pei-Ying Lee,† Cheng-Chung Wang,‡§ and Shang-Cheng Hung*†∥

†Genomics Research Center and §Institute of Chemistry, Academia Sinica, 128, Section 2, Academia Road, Taipei 115, Taiwan
‡Department of Chemistry, National Tsing Hua University, 101, Section 2, Kuang-Fu Road, Hsinchu 300, Taiwan
∥Department of Applied Chemistry, National Chiao Tung University, 1001, Ta-Hsueh Road, Hsinchu 300, Taiwan

Supporting Information

ABSTRACT: Several biological processes involve glycans, yet understanding their ligand specificities is impeded by their inherent diversity and difficult acquisition. Generating broad synthetic sugar libraries for bioevaluations is a powerful tool in unraveling glycan structural information. In the case of the widely distributed heparan sulfate (HS), however, the 48 theoretical possibilities for its repeating disaccharide call for synthetic approaches that should minimize the effort in an undoubtedly huge undertaking. Here we employed a divergent strategy to afford all 48 HS-based disaccharides from just two orthogonally protected disaccharide precursors. Different combinations and sequence of transformation steps were applied with many downstream intermediates leading up to multiple target products. With the full disaccharide library in hand, afinity screening with fibroblast growth factor-1 (FGF-1) revealed that four of the synthetic sugars bind to FGF-1. The molecular details of the interaction were further clarified through X-ray analysis of the sugar–protein cocrystals. The capability of comprehensive sugar libraries in providing key insights in glycan–ligand interaction is, thus, highlighted.

INTRODUCTION

Glycans are exceptionally diverse and complex that deciphering the functions embedded within the glycome is a substantial challenge.1 The multiple regio- and stereochemical permutations in linking several monosaccharide units and the modifications that may follow chain assembly allowed these complex sugars to hold structural information densities surpassing DNA and proteins. With regulated rather than template-driven biosyntheses, their expression often produces an array of related structures that may possess subtle differences in activity. A case in point is heparan sulfate (HS), a proteoglycan sugar component with crucial roles in metazoan development, physiology, and disease as a dynamic regulator of protein activities at the cell–extracellular interface.2 The modifications of the HS precursor, made up of an extended 1→4-linked N-acetyl-α-D-glucosamine (GlcNAc) and β-D-glucuronic acid (GlcA) copolymer, facilitated by several enzyme isoforms of varying specificities, are always incomplete, resulting in extensive chain microheterogeneity.3 These modifications, which include GlcNAc N-deacetylation, GlcA 5-C-epimerization toward α-L-iduronic acid (IdoA), and sulfonations at 2-N, 3-O, and 6-O of α-D-glucosamine (GlcN) and at 2-O of the uronic acid, overall account for 48 theoretical possibilities for the repeating disaccharide (Figure 1). Consequent of being widespread in cell surfaces, extracellular matrices, and basement membranes, numerous proteins of various origins and functions evolved to associate with and take advantage of the elaborate patterns decorating the HS backbone.4 Recognizing the optimum patterns sought by binding proteins could guide the development of novel diagnostic agents and biomedical interventions. Under such premise, fondaparinux, a synthetic

Figure 1. Structures of (a) the 48 disaccharides theoretically present in HS and (b) the orthogonally protected disaccharide precursors utilized in our divergent synthesis.

Received: September 10, 2012
Published: December 14, 2012
pentasaccharide based on the HS analogue heparin, is now a clinically approved anticoagulant that is free of the main side effects of heparin therapy. Chemicaly defined materials are indispensable components of structure–activity relationship assays. Because sugar structures obtained from natural sources are generally unsuitable for these purposes, synthetic strategies have advanced to deliver these much needed compounds for biological studies.1,6

Developing broad synthetic sugar libraries, aimed at probing the character of carbohydrate interactions with their ligands, is envisioned as a potent tool in understanding the fundamental roles of glycans in biological systems. For HS, the library should ideally include constructs covering the entire range of structures found within the polysaccharide. However, building such a collection is hampered by the number and diversity of compounds that have to be considered as well as the notorious technical difficulty of HS synthesis.6,7 There were recent successes in preparing particular sequences using chemoenzymatic methodologies,8 but reaction completion and monitoring, availability of chemically pure starting materials, and product purification remain pervasive concerns. To prepare HS-based compounds, we turned to chemical synthesis, which proved reliable in accessing well-defined HS oligosaccharides. HS synthesis, nevertheless, offers unique challenges, such as the acquisition of the rare L-ido derivatives, the regio- and stereocontrol of glycosidic bond formation, and the judicious selection and manipulation of protecting groups to match the intricate functional group pattern of the desired products.

Synthetic efforts reported in the literature typically employ explicitly designed oligosaccharide precursors to access the target HS oligosaccharides through a myriad of functional group transformations.9 Furthermore, generating a typical trisulfonated HS-based disaccharide takes about 20 reaction steps from common monosaccharide starting materials. As such, building a comprehensive HS-based sugar library, with each compound obtained from an individual multistep route, would be a huge undertaking. Thus, efficient strategies that minimize the number of steps, such as making full use of shared intermediates, should be explored. Accordingly, we developed an orthogonal protecting group strategy that permitted access to all 48 disaccharides theoretically present in HS from only two disaccharide precursors, one with a D-glucosyl (1) and another with an L-idosyl unit (2) (Figure 1). These compounds were then transformed following a divergent process, where many downstream intermediates were utilized to make more than one final product. Fibroblast growth factor-1 (FGF-1), a prototypical member of the FGF family, was chosen to demonstrate the effectiveness of our disaccharide library in binding specificity. Thus, binding affinity screening and, subsequently, X-ray cocrystal analysis were conducted, the results of which are disclosed herein.

## RESULTS AND DISCUSSION

### Syntheses of 48 HS-Based Disaccharides.

A carefully selected set of protecting groups was employed to decorate the disaccharide precursors 1 and 2. We used the simple methyl group to block the anomeric position at the reducing end. Installations of the orthogonal benzyl (Bz) group at the 2-O position of the D-glucosyl and L-idosyl units and 2-naphthylmethyl (2-NAP) and tert-butylidiphenylsilyl (TBDDS) groups at the respective 3-O and 6-O positions of GlcN cover the O-sulfonation patterns in HS. The 2-O-Bz group is also expected to confer neighboring group assistance during the formation of the 1,2-trans glycosidic bond. Selective access to the primary alcohol for later oxidation to the carbohydrate is provided by the acetyl (Ac) group. The amine was masked as an azide to take advantage of its nonparticipation in glycosylation, favoring the α-stereoselective glucosaminylation via the anomic effect. Moreover, the azide can be readily transformed into the amine, acetamide, or sulfamate in later synthetic steps. Finally, benzyl (Bn) groups are utilized to block the hydroxyls that would be free in the final compounds.

The disaccharide preparations started with Williamson etherification at 3-O of the thiglycoside3,10 (93%) followed by borane-mediated reductive benzylidene 6-O-ring opening (97%) to obtain the 6-alcohol 4 (Scheme 1). The desired D-glucosaminyl donor 5 was acquired in 90% yield after a typical 6-O-silylation procedure. Subsequent glycosylation of the known 1,6-anhydro-1-idoxyranosyl 4-alcohol 6% promoted by N-iodosuccinimide (NIS) and trimethylsilyl trifluoromethanesulfonate (TMSOTf) provided the 6-disaccharide 7 in 64% yield. The disaccharide preparations started with Williamson etherification at 3-O of the thiglycoside3,10 (93%) followed by borane-mediated reductive benzylidene 6-O-ring opening (97%) to obtain the 6-alcohol 4 (Scheme 1). The desired D-glucosaminyl donor 5 was acquired in 90% yield after a typical 6-O-silylation procedure. Subsequent glycosylation of the known 1,6-anhydro-1-idoxyranosyl 4-alcohol 6% promoted by N-iodosuccinimide (NIS) and trimethylsilyl trifluoromethanesulfonate (TMSOTf) provided the 6-disaccharide 7 in 64%

### Scheme 1. Preparation of the Common Intermediates 11 and 12.

![Scheme 1](image_url)

**Reagents and conditions:** (a) (1) 2-NAPBr, NaH, DMF; 93%; (2) BH₃·THF, TMSOTf, CH₂Cl₂, 0 °C, 6 h; 97%. (h) TBDPSCI, Et₃N, DMAP, CH₂Cl₂, 0 °C to rt, 16 h; 90%. (c) NIS, TMSOTf, CH₂Cl₂, 3 Å MS, −78 °C to rt, 2 h; 64%. (d) (1) benzaldehyde, TMSOTf, CH₂Cl₂, 3 Å molecular sieves, 0 °C, 2 h; (2) benzaldehyde, TMSOTf, Et₂SiH, −78 °C, 4 h; (3) Bu₂O, TMSOTf, 0 °C, 18 h; (4) 70% TFA, 3 h, rt; 65% (4 steps, one pot). (e) Ac₂O, Et₃N, CH₂Cl₂, 0 °C, 3 h; 86%. (f) (1) MeOH, NIS, TMSOTf, CH₂Cl₂, 3 Å molecular sieves, −78 to −20 °C, 2 h; (2) 5% TEMPO, BaIB, CH₂Cl₂, H₂O, rt, 6 h; 11, 90%; 12, 89%. (g) (1) Ac₂O, Cu(OTf)₂, 0 °C, 16 h; 91%; (2) saturated NH₄OH, THF, MeOH, 0 °C, 3 h; 81%; (3) CCl₃CN, K₂CO₃, 16 h; 93%; (4) MeOH, AgOTf, CH₂Cl₂, 3 Å molecular sieves, −5 °C, 1 h; 81%. (j) NaOMe, CH₂Cl₂, MeOH, rt; 87%. DMAP: 4-(N,N-dimethylamino)pyridine.

"Reagents and conditions: (a) (1) 2-NAPBr, NaH, DMF; 93%; (2) BH₃·THF, TMSOTf, CH₂Cl₂, 0 °C, 6 h; 97%. (h) TBDPSCI, Et₃N, DMAP, CH₂Cl₂, 0 °C to rt, 16 h; 90%. (c) NIS, TMSOTf, CH₂Cl₂, 3 Å MS, −78 °C to rt, 2 h; 64%. (d) (1) benzaldehyde, TMSOTf, CH₂Cl₂, 3 Å molecular sieves, 0 °C, 2 h; (2) benzaldehyde, TMSOTf, Et₂SiH, −78 °C, 4 h; (3) Bu₂O, TMSOTf, 0 °C, 18 h; (4) 70% TFA, 3 h, rt; 65% (4 steps, one pot). (e) Ac₂O, Et₃N, CH₂Cl₂, 0 °C, 3 h; 86%. (f) (1) MeOH, NIS, TMSOTf, CH₂Cl₂, 3 Å molecular sieves, −78 to −20 °C, 2 h; (2) 5% TEMPO, BaIB, CH₂Cl₂, H₂O, rt, 6 h; 11, 90%; 12, 89%. (g) (1) Ac₂O, Cu(OTf)₂, 0 °C, 16 h; 91%; (2) saturated NH₄OH, THF, MeOH, 0 °C, 3 h; 81%; (3) CCl₃CN, K₂CO₃, 16 h; 93%; (4) MeOH, AgOTf, CH₂Cl₂, 3 Å molecular sieves, −5 °C, 1 h; 81%. (j) NaOMe, CH₂Cl₂, MeOH, rt; 87%. DMAP: 4-(N,N-dimethylamino)pyridine.

20723 dx.doi.org/10.1021/ja0990065 | J. Am. Chem. Soc. 2012, 134, 20722–20727
yield together with the minor β-isomer (14%). Conversely, the regioselective one-pot protection11 of the per-trimethylsilylated thioglucoside 8 furnished the corresponding 4,6-diol 9 in a 4-step yield of 63%. The transformations included the initial 4,6-O-benzylidene formation followed by silane-mediated regioselective 3-O-benzyl, 2-O-benzoylation, and, last, acid hydrolysis of the benzylidene acetal. Exploiting the higher reactivity of the primary hydroxyl, the diol further underwent regioselective 6-O-acetylation to afford the 4-alcohol 10 (86%). Condensation of the acceptor—donor 10 with MeOH under NIS/TMSTf promotion, followed in one pot by α-glycosylation with the thioglucoside 8, led to the fully protected precursor 1 in 61% yield. Mild decacylation of the disaccharide 1 using Mg(OMe)212 (93%) and oxidation of the free primary alcohol with catalytic 2,2,6,6-tetramethyl-1-piperidinylxoyl free radical (TEMPO) in the presence of excess [bis(acetoxy)-ido]benzene (BAIB) delivered the carboxylic acid 11 (90%). Regarding the adduct 7, acetylation of the anhydro-ring was readily achieved in 91% yield using Ac2O and copper(II) trifluoromethanesulfonate [Cu(OTf)2]13. Treatment with saturated ammonia in a THF and MeOH cosolvent system enabled anomic decacylation (81%). Further trichloroacetimidate formation gave the glycosyl donor (93%), which was coupled with MeOH to acquire the target GlcN–IdoA precursor 2 in 81% yield. Successive Zemplén decacylation (87%) and TEMPO/BAIB oxidation (89%) finally produced the lactone 12.

Compounds 11 and 12 were henceforth converted to the desired HS-based materials in a divergent manner through different combinations and sequence of transformation steps. Scheme 2 describes the syntheses of 24 disaccharides with GlcN–GlcA backbone. Thiocetic acid (AcS)14 treatment of the disaccharide 11 accomplished the direct conversion of the azide to the acetamide to afford compound 13 in 92% yield. Both 11 and 13 were subjected to parallel reaction sequences toward the amino- and acetamido-containing final compounds. Tetra-n-butylammonium fluoride (TBAF) with equimolar AcOH, NaOMe in MeOH, and 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) were applied to chemoselectively cleave the TBDPS, Bz, and 2-NAP groups, respectively. Proper selection of these deprotection protocols exposed key hydroxyl groups for O-sulfonations, which are implemented using SO3·Pyr in the presence of NaOH and Et3N. In the final step, palladium-catalyzed hydrolysis cleaved all arylmethyl (Bn) and 2-NAP ethers and, simultaneously, reduced the azido group to the free amine.

The acetamide 54 was similarly obtained using AcSH from compound 12 in 87% yield (Scheme 3). Unfortunately, the labile lactone in compounds 12 and 54 prompted a reconsideration of the desilylation reagent. The basic nature of TBAF promoted lactone hydrolysis (even with AcOH as coreagent), a result that we exploited in revealing the 2- and 6'-hydroxyls in one step. To spare the lactone function, removal of the TBDPS group was accomplished by treatment with the mild tri(dimethyl-amino)sulfonium difluorotrimethylsilylate (TASF).15 Lactone ring opening was separately achieved by LiOH. In an alternative but equally effective N-sulfonation procedure, the afforded amine, after global hydrolysis, was treated with SO3·Pyr in basic (pH 9.5) aqueous conditions. Consequently and together with applicable reactions described for the GlcN–GlcA series, we acquired the other 24 disaccharides having the GlcN–IdoA structure. Thus, our target 48 HS-based disaccharides were generated from the intermediates 11 and 12 in 3–7 steps and 24–95% overall yields. Nuclear magnetic resonance and mass spectroscopic analyses confirmed the final product structures (see the Supporting Information, SI). Throughout the above-mentioned transformations, we strategically utilized numerous intermediates to prepare multiple target products, effectively reducing the effort in an otherwise laborious synthetic endeavor.

**Scheme 2. Preparations of 24 HS-Based Disaccharides With GlcN–GlcA Backbone**

![Scheme 2](image)

Reagents and conditions: (a) AcSH, Pyr, CHCl3, rt, 18 h. (b) TBAF, AcOH, 40 °C. (c) NaOMe, CH2Cl2, MeOH, rt. (d) SO3·Et3N, DMF, 60 °C, 3 d. (e) Pd(OH)2/C, H2 (balloon), phosphate buffer (pH 7), rt, 2 d. (f) Pd/C, H2 (50 psi), MeOH, H2O, rt, 2 d. (g) 1,3-propanedithiol, Et3N, Pyr, H2O, 50 °C. (h) SO3·Pyr, NaOH, Et3N, MeOH, rt. (i) DDQ, CH2Cl2, H2O, rt, 4 h. Pyr: pyridine.

Scheme 3. Preparation of 24 HS-Based Disaccharides With GlcN–IdoA Backbone

<table>
<thead>
<tr>
<th>entry</th>
<th>sugar</th>
<th>(K_D (\mu M))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>89</td>
<td>18.1</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>4.13</td>
</tr>
<tr>
<td>3</td>
<td>91</td>
<td>21.2</td>
</tr>
<tr>
<td>4</td>
<td>92</td>
<td>9.71</td>
</tr>
<tr>
<td>5</td>
<td>3 kDa heparin</td>
<td>0.746</td>
</tr>
</tbody>
</table>

*Measured as association constant, the inverse of \(K_D\).*
colors red, yellow, and blue represent oxygen, tetrasaccharide helical structure,22 two FGF-1s could possibly unlikely in our case. Instead, as HS assumes a repeating trans disaccharide library. Our orthogonal protecting group combi-

---

Table 2. Potential Intermolecular Interactions Within the FGF-1–Disaccharide Interface

<table>
<thead>
<tr>
<th>FGF-1 residue (atom)</th>
<th>sugar group or atom (residue)</th>
<th>89</th>
<th>90</th>
<th>91</th>
<th>92</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn18 (Nζ2)</td>
<td>2-O-SO3(^-) (IdoA)</td>
<td>3.7</td>
<td>3.8</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>2-N-SO4(^-) (GlcN)</td>
<td>3.5</td>
<td>2.8</td>
<td>3.1</td>
<td>2.9</td>
</tr>
<tr>
<td>Lys113 (NH)</td>
<td>2-N-SO3(^-) (GlcN)</td>
<td>3.0</td>
<td>2.7</td>
<td>2.6</td>
<td>3.1</td>
</tr>
<tr>
<td>Lys113 (Nζ2)</td>
<td>3-OH (IdoA)</td>
<td>3.5</td>
<td>2.4</td>
<td>3.4</td>
<td>2.5</td>
</tr>
<tr>
<td>Lys118 (Nζ2)</td>
<td>2-O-SO3(^-) (IdoA)</td>
<td>3.9</td>
<td>3.6</td>
<td>3.0</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>2-N-SO4(^-) (GlcN)</td>
<td>2.8</td>
<td>2.7</td>
<td>2.6</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>3-O-SO3(^-) (GlcN)</td>
<td>–</td>
<td>3.0</td>
<td>–</td>
<td>3.4</td>
</tr>
<tr>
<td>Gln127 (Nε2)</td>
<td>2-O-SO3(^-) (IdoA)</td>
<td>2.9</td>
<td>3.8</td>
<td>3.8</td>
<td>4.0</td>
</tr>
<tr>
<td>Lys128 (NH)</td>
<td>2-O-SO3(^-) (IdoA)</td>
<td>3.0</td>
<td>2.9</td>
<td>2.9</td>
<td>2.7</td>
</tr>
</tbody>
</table>

carboxylate groups made no contacts with FGF-1. Their outward orientations in our crystal structures suggest potential roles in the interface between HS and FGFR or perhaps in minor interaction with another FGF-1 in the active complex. Thus, the FGF-1–FGFR interaction may be suppressed by removing these nonparticipating moieties, and by the same token, compound 90 could potentially act as inhibitor of FGF-1 activity. FGF dimerization was previously observed with longer HS-based oligosaccharides17 and sucrose octasulfate24 but is unlikely in our case. Instead, as HS assumes a repeating tetrasaccharide helical structure,22 two FGF-1s could possibly bind, in trans-orientation, at the adjacent N-sulfonated GlcN and 2-O-sulfonated IdoA sequences (i.e., GlcN–IdoA–GlcN–IdoA) to form an HS-bridged FGF-1 dimer. Under such arrangement, the 6-O-sulfonate group, if available in the internal GlcN, may likely strengthen the dimerization by further interaction with the other FGF-1.

**CONCLUSIONS**

We have demonstrated an effective divergent strategy that provided efficient access to a comprehensive HS-based disaccharide library. Our orthogonal protecting group combi-

---

**EXPERIMENTAL SECTION**

**Chemical Synthesis.** The complete experimental details and compound characterization data can be found in the SI.

**Protein Expression and Purification.** The human FGF-1 gene was cloned in PET-21a expression vectors and transformed into the BL21 (DE3) strains cultured in lysogeny broth supplemented with 100 μg·mL\(^{-1}\) ampicillin. Protein expression was induced with 1.0 mM isopropyl-1-thio-β-D-galactoside at 37 °C for 4 h. The crude cell extracts were loaded onto a heparin affinity column followed by gel filtration (Superdex 200, Pharmacia) of the FGF-1-containing fractions. The purified proteins were dialyzed against Tris-buffered saline (pH 7.6) and concentrated to levels suitable for ITC and cocRYSTALLIZATION. Protein concentrations were measured by typical UV and optical density analysis.

**ITC Measurements.** All ITC experiments were carried out at 25 °C using Tris buffer (20 mM Tris, 100 mM NaCl, pH 7.6) as solvent. The sugars (0.43–0.92 mg) were dissolved in Tris buffer to the desired concentrations (Figure S1). The FGF-1 solution was placed in the calorimeter cell and titrated with the sugar solution (2 μL injections with 3 min spacing). To account for the heat of the dilution, the buffer without the protein was also titrated with the sugar solution, and the generated data were subtracted from the results of the protein titration. The titration isotherms were fitted to an equation modeling the interaction (One Sites) to generate the fitting parameters.

**Crystallization and Data Collection.** FGF-1 (10 mg·mL\(^{-1}\) in Tris-buffered saline) was separately mixed with disaccharides 89–92 in 1:1.2 molar ratio. All crystals were grown by hanging drop vapor diffusion at rt by mixing 1 μL of protein complex solution with 1 μL of reservoir solution containing 2 M (NH₄)₂HPO₄ and 0.1 M sodium acetate at pH 7.0. Diffraction data were collected at −150 °C at beamline 13B1 of the National Synchrotron Radiation Research Center in Hsinchu, Taiwan, and were processed and scaled by HKL2000.23

**Structure Determination and Refinement.** The FGF-1–disaccharide complexes all crystallized in the P space group with three molecules per asymmetric unit. The structures were solved by molecular replacement using the crystal structure of FGF-1 (PDB code 1rg8) as the searching model. The structural model was subjected to manual rebuilding with WinCoot24 and then refined using REFMAC5.25 The data collection and refinement statistics are available in Table S1.

**ASSOCIATED CONTENT**

**Supporting Information**

Figure S1, Table S1, synthetic methods and characterization data, and \(^1\)H and \(^13\)C NMR spectra of relevant compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

**Corresponding Author**

schung@gate.sinica.edu.tw

**Author Contributions**

*These authors contributed equally.
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This work was supported by the National Science Council (NSC 97-2113-M-001-033-MY3, NSC 98-2119-M-001-008-MY2, 100-2113-M-001-019-MY3) and Academia Sinica.

**REFERENCES**


