

## Chemical editing of proteoglycan architecture

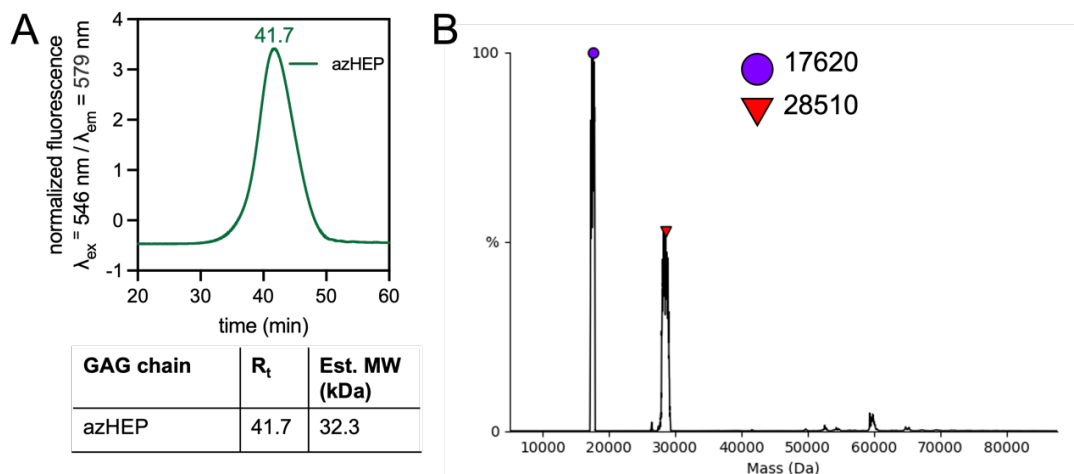
Timothy R. O'Leary,<sup>1,\*</sup> Meg Critcher,<sup>1,2,\*</sup> Tesia N. Stephenson,<sup>1</sup> Xueyi Yang,<sup>1,2</sup> Abdullah A. Hassan<sup>1</sup>, Noah M. Bartfield,<sup>1</sup> Richard Hawkins,<sup>1</sup> Mia L. Huang<sup>1,2,3</sup>

\* Equal authorship, <sup>1</sup> Department of Molecular Medicine, Scripps Research, 120 Scripps Way, Jupiter, FL 33458-5284, <sup>2</sup> Skaggs Graduate School of Chemical and Biological Sciences, Scripps Research, 10550 N Torrey Pines Rd, La Jolla, CA 92037 <sup>3</sup> Department of Molecular Medicine, Scripps Research, 10550 N Torrey Pines Rd, La Jolla, CA 92037

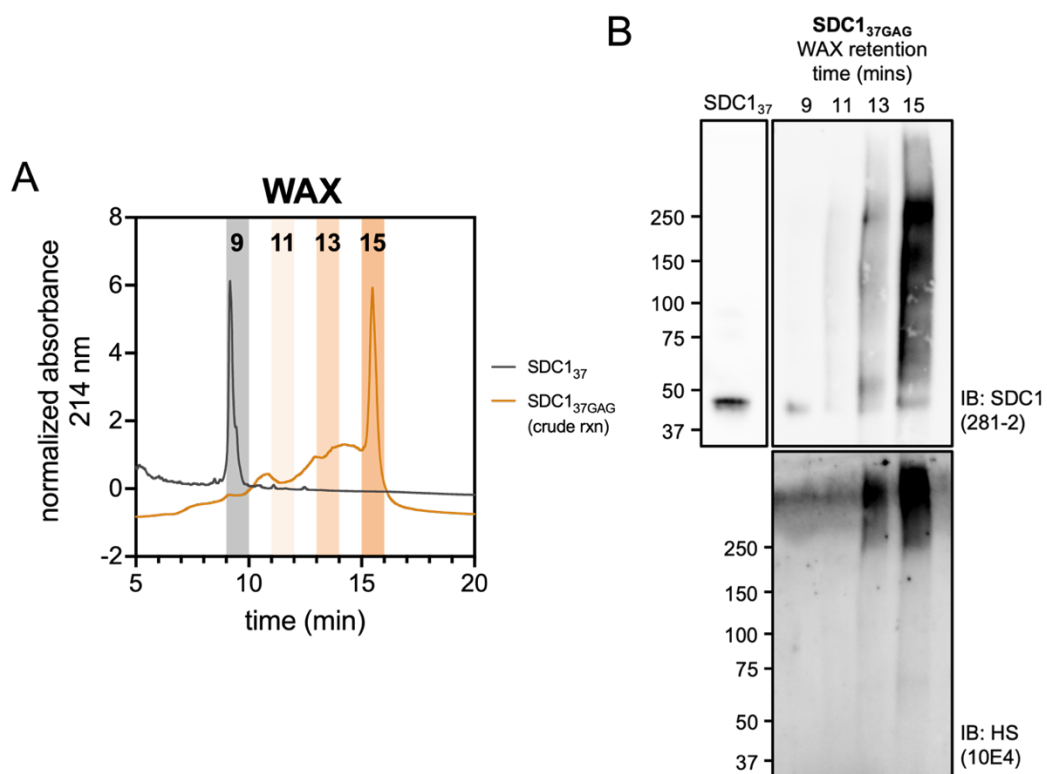
Corresponding author email: [miahuang@scripps.edu](mailto:miahuang@scripps.edu)

### Contents:

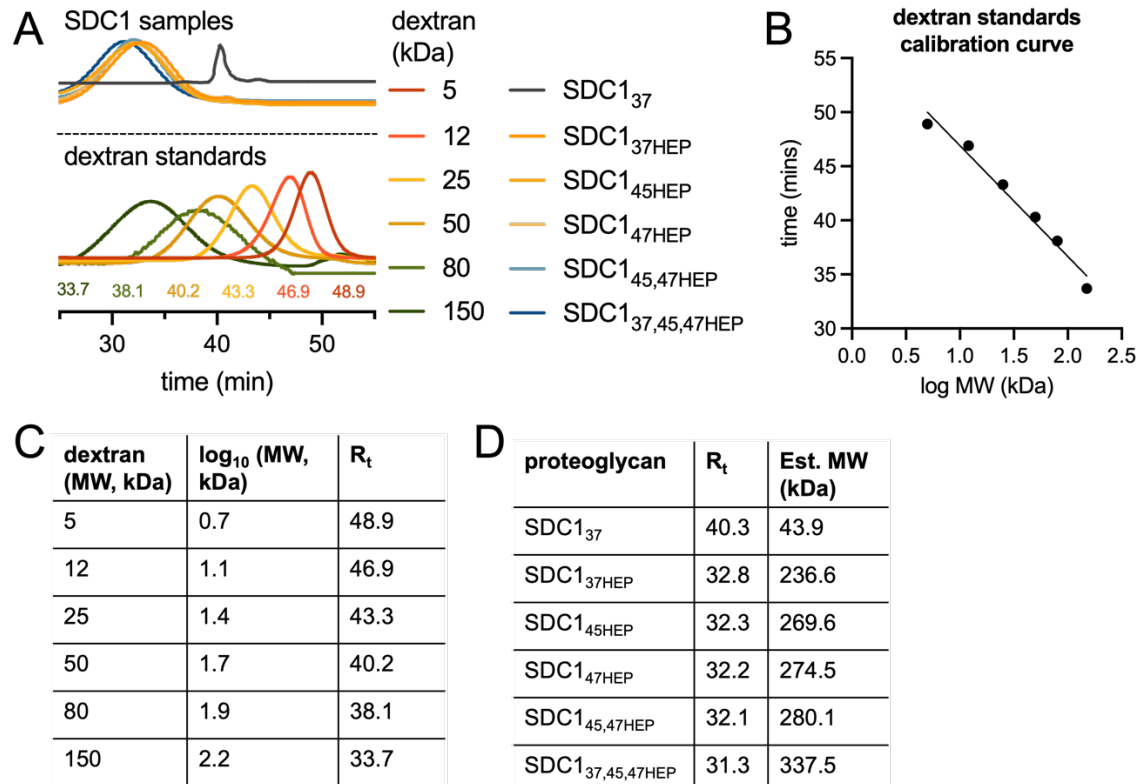
<b>Fig. 1</b>	Assessing heparin chain size by SEC and MS	<b>2</b>
<b>Fig. 2</b>	Evidence of successful SDC1 <sub>37GAG</sub> conjugation	<b>2</b>
<b>Fig. 3</b>	Assessing PG sizes and multivalency by SEC	<b>3</b>
<b>Fig. 4</b>	Confirmation of SDC1 alkyne occupancy upon conjugation with azHEP	<b>3</b>
<b>Fig. 5</b>	Pilot assays for AlphaScreen binding to integrin $\alpha_v\beta_3$	<b>4</b>
<hr/>		
<b>Table 1</b>	List of reagents and sources	<b>5</b>
<b>Table 2</b>	Syndecan ectodomain amino acid sequences	<b>7</b>
<b>Table 3</b>	Syndecan ectodomain expression yields	<b>8</b>
<b>Table 4</b>	RT-qPCR primers	<b>8</b>
<b>Table 5</b>	SDC1 DsiRNA targets	<b>8</b>
<hr/>		
<b>Note 1</b>	Synthesis and characterization of chemical structures	<b>9</b>
<b>Note 2</b>	Chemical structures of heparan sulfate and chondroitin sulfate disaccharides.	<b>15</b>
<b>References</b>		<b>15</b>



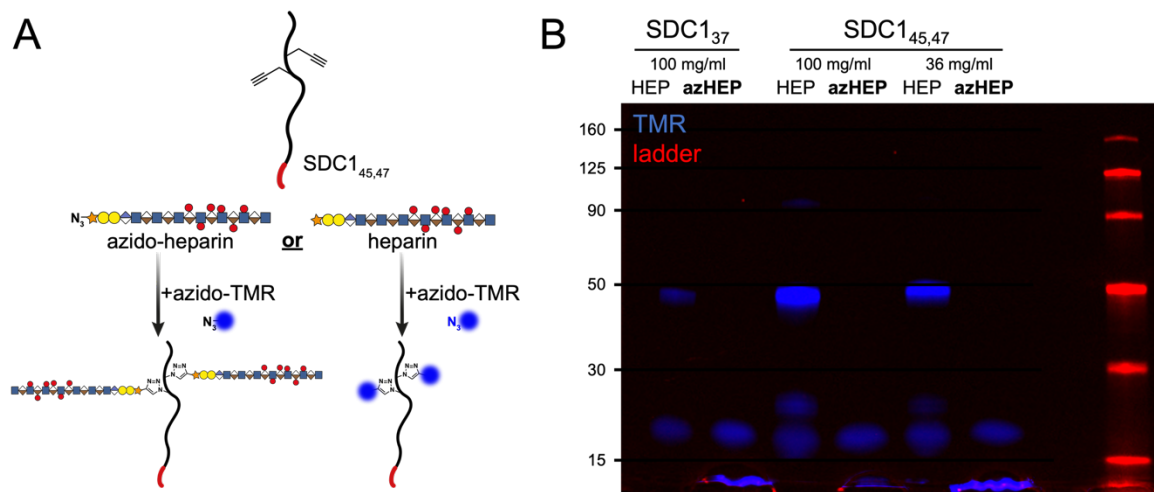
**Supplementary Figure 1. Assessing heparin chain size by SEC and MS. (A)** SEC chromatogram and estimated molecular weight of azido-heparin (azHEP, green) clicked to TMR-azide, calculated using a dextran standard curve (Supplementary Figure S3). **(B)** Intact MS of azido-heparin (ion trap neg mode ESI) showing deconvoluted peaks at 17,620 or 28,510 Da.



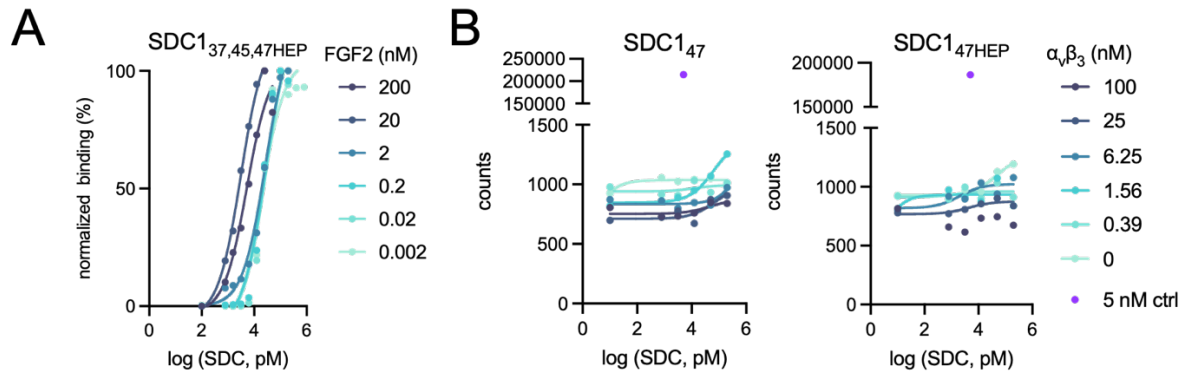
**Supplementary Figure 2. Evidence of successful SDC<sub>137GAG</sub> conjugation. (A)** WAX chromatogram of SDC<sub>137GAG</sub> crude reaction products. Highlighted are the fractions collected for SDS-PAGE and Western blot (WB) analysis. These fractions correspond to SDC<sub>137</sub> (grey,  $R_t = 9$ ) and SDC<sub>137GAG</sub> (orange,  $R_t = 11, 13, 15$ ). **(B)** 4-15% gradient SDS-PAGE analysis of collected fractions, transferred to PVDF membrane. Top panel: anti-SDC1 antibody positively stains fractions 9 - 15, detecting higher molecular weight smears (~250kDa) in later fractions. Antibody clone 281-2 has a higher affinity for GAG-conjugated SDC1 relative to non-glycosylated SDC1. Bottom: signal with anti-HS is detected only in glycoconjugates, with highest signal in fractions 13 and 15.



**Supplementary Figure 3. Assessing PG sizes and multivalency by SEC.** (A) SEC of SDC1 samples (top, Fig 1C) and dextran standards (bottom), with elution peaks annotated for the latter. (B) calibration curve created from dextran standards, using log<sub>10</sub> molecular weight (MW, kDa). (C) Tabulated data from B. (D) Estimated sizes of SDC1 core protein and glycoconjugates calculated using dextran SEC standards.



**Supplementary Figure 4. Confirmation of SDC1 alkyne occupancy upon conjugation with azHEP.** (A) Graphical depiction of experiment. SDC1<sub>37</sub> or SDC1<sub>45,47</sub> (depicted) were reacted with either azido-heparin (left) or heparin (non-functionalized, right) for 2 h at 37°C, before the addition of azido-TMR (blue). If alkyne sites are occupied by azido-heparin, no labelling with azido-TMR will occur. (B) SDS-PAGE analysis of click reaction products. Positive TMR staining (blue) is detected only in SDC1 samples treated with heparin.



**Supplementary Figure 5. Pilot assays for AlphaScreen binding to integrin  $\alpha_v\beta_3$ .** (A) Pilot assay of AlphaScreen assay with titrated biotinylated FGF2 and SDC1<sub>37,45,47HEP</sub>. (B) AlphaScreens performed with SDC1<sub>47</sub> (left) and SDC1<sub>47HEP</sub> (right) show no discernible binding to biotinylated  $\alpha_v\beta_3$  integrin. Positive control bead (biotinylated, poly-His tagged peptide, purple) included to confirm successful assay.

**Supplementary Table 1.** List of reagents and sources

<b>Product</b>	<b>Manufacturer</b>	<b>Product #</b>
pULTRA-CNF	AddGene	48215
Leupeptin	EMD	108975
Aprotinin	GoldBio	A-655-25
Pepstatin A	Thermo	J20037
Amicon 10kDA MWCO centrifugal filter	Millipore	UFC8010
Halolink resin	Promega	G1912
HaloTEV protease	Promega	6860A
Hiprep 16/60 Sephacryl S-200 HR	Cytiva	17116601
Superose 6 Increase 10/300 GL SEC	Cytiva	29091596
Sulfo-NHS-biotin	Biovision	2326
Heparin Sepharose® 6	GE Healthcare	17-0998-01
Type A, Porcine Skin Gelatin Powder	Sigma	G-1890
KO-DMEM	Gibco	10829-018
Gemcell FBS	Gemini	100-500
Leukemia inhibitory factor (LIF, ESGRO)	Millipore	ESG1107
MEM NEAA (100X)	Gibco	11140-035
2-mercaptoethanol (2-ME, 50mM)	Gibco	31350-010
Neurobasal	Gibco	A35829-01
DMEM/F-12	Gibco	11320-033
L-glutamine (200mM)	Gibco	25030-024
N-2 supplement (100X)	Gibco	17502-048
B-27 Plus supplement (50X)	Gibco	A35727-01
Heparin	Iduron	Hep001
Lab-Tek II chamber slide system (8-well)	ThermoFisher	154534
Cholesterol-PEG-NTA	Nanocs	PG2-CSNT-
EDTA-free trypsin	Quality Biological	118-086-721
Non-enzymatic cell dissociation buffer	PeptoTech	CPD-125
Lipofectamine RNAiMAX	Invitrogen	13778-150
Vitronectin	PeptoTech	140-09-1mg
Recombinant human fibronectin fragment	Sino Biological	10314-H08H
GFP(His) <sub>6</sub>	Addgene	85482
FGF2	Novus	NBP2-76301
Integrin $\alpha_v\beta_3$	ACRO	IT3-H52E3
Trypsin-EDTA (0.05%)	Gibco	25300-062
Trypsin-EDTA (0.25%)	Gibco	25200-056
IPTG	Bioworld	21530057
Kanamycin	Fisher	BP906-5
Spectinomycin	GoldBio	S-140-25
Imidazole	Sigma-Aldrich	12399
Aminoguanidine hydrochloride	Acros	368910250
Copper sulfate	Acros	42287-1000
THPTA	TCI	T3171
Sodium ascorbate	ChemImpex	1436
5-tetramethylrhodamine-azide (TMR- Anthranilamide (2-AB)	Lumiprobe	37130
4-(bromomethyl)benzoate	Acros	104900050
	Combiblocks	OR-0319

Sodium azide	Fisher	BP9221
DMF	Fisher	D119-1
Deuterium Oxide	Aldrich	151882
DMSO-D6	Aldrich	256147
Heparin Agarose beads	Sigma	H0404
Sodium Acetate	Sigma	791741
Aniline	Sigma	24284
DMSO	Fisher	D128-1
AlexaFluor-488-alkyne	Sigma	761621
Sulfo-DBCO-biotin	Sigma	760706
Heparinase I and III	Sigma	H3917
Heparinase II	Sigma	H6512
TMB HRP substrate	BioFX	TMBW-1000-
Streptavidin agarose resin	Thermo	20353
Chondroitinase ABC	Sigma-Aldrich	C2905
Sodium cyanoborohydride	Sigma-Aldrich	156159
MMP-9	R&D Biosystems	909-MM
4-Aminophenylmercuric acetate	Sigma-Aldrich	A9563
ADAM-TS1	R&D Biosystems	2197-AD
Heparin disaccharide standard mix	Galen	HD Mix
RNeasy Mini Kit	QIAGEN	74004
TURBO DNA-free™ Kit	Invitrogen	AM1907
High-Capacity cDNA Reverse	Applied	4368814
Fast SYBR Green Master Mix	ThermoFisher	4385610
OptiPrep density gradient medium	Sigma	D1556-
Immobilon-FL PVDF membrane	Sigma	IPF00010
Immobilon-P PVDF membrane	Sigma	IPVH0000/5
½ area OptiPlate-96	PerkinElmer	6002299
AlphaScreen beads	PerkinElmer	6761619
Maxisorp 96-well plate	Thermo	439454

<b>Primary antibody</b>	<b>Manufacturer</b>	<b>Product #</b>	<b>Dilution</b>
SOX1	Cell Signalling	4194S	1:300
Nestin (clone 4D4)	Invitrogen	14-5843-82	1:400
Nanog (clone D2A3)	Cell Signalling	8822T	1:1000
B-tubulin III (TubB3)	Proteintech	66240-1-1g	1:300
SDC1 (clone 281-2)	BioLegend	142502	1:2000
CAV-1 (clone 7C8)	R&D Biosystems	MAB5736	1:1000
SDC1 (clone MI-15)	BioLegend	356502	1:100
Rhodamine-phalloidin	Biotium	50-196-4057	1:40
Biotinylated human CD51/61 (integrin	Biolegend	304412	1:1000

<b>Secondary antibody</b>	<b>Manufacturer</b>	<b>Product #</b>	<b>Dilution</b>
Donkey anti-mouse IgG AF647	Invitrogen	1984047	1:1000
Donkey anti-mouse AF555	Invitrogen	A31570	1:1000
Goat anti-rabbit IgG AF647	Invitrogen	A32733	1:1000
Streptavidin-Cy5	Southern Biotech	7100-15	1:1000
Streptavidin-HRP	Biolegend	405210	1:10000

Goat anti-rabbit IgG HRP	Abcam	Ab6721	1:10000
Goat anti-mouse IgG HRP	Invitrogen	G21040	1:10000

**Supplementary Table 2.** Syndecan ectodomain amino acid sequences.

Proteoglycan ectodomain	Sequence	PI	Theoretical MW (Da)
<b>SDC1 (18-252)</b>	MQPALPQIVAVNVPPEDQDGSDDSDNFSGSGTGALPDTL SRQTPSTWKDVWLLTATPTAPEPTSSNTETAFTSVLPAGEK PEEGEPVLHVEAEPGFTARDKEKEVTTRPRETVQLPITQRA STVRVTTAQAAVTSHPHGGMQPGLHETSAPTAPGQPDHQ PPRVEGGGTSVIKEVVEDGTANQLPAGEGSGEQDFTFETS GENTAVAAVEPGLRNQPPVDEGATGASQSLLDRKEHHHH HHHHHH	4.71	25825.6
<b>SDC2 (19-145)</b>	MGSSHHHHHHSSGLVPRGSHMASMETRTELTSKDMYLD NSSIEEASGVYPIDDDDYSSASGGADEDIESPVLTTSQLIP RIPLTSAASPKEVETMTLKTQSITPAQTESPEETDKEEVDISEA EEKLGPAAIKSTDVYTEKHSNLFKRTEHHHHHHHHHHH	4.70	18058.4
<b>SDC3 (45-384)</b>	MGSSHHHHHHSSGLVPRGSHMASMAQRWRNENFERPVD LEGSGDDDSFPDDELDDLYSGSGSGYFEQESGLETAMRFI PDMALAAPTAPAMLPTTVIQPVDTPFEELLSEHPSPEPVTSP PLVTEVTEVVESSQKATTISTTTSTTAATTTGAPTMATAPA TAATTAPSTPEAPPATATVADVVRTTGIQGMPLPLTTAATAK ITTPAAPSPPTTVATLDTEAPTPLVNTATSRPRALPRPVT QEPDVAERSTLPLGTTAPGPTEMAQTPTPESLLTTIQDEPE VPVSGGSPGFELQEETTQPDANEVVAVEGAAAKPSPPL GTLPKGARPGPGLHDNAIDSGSSAAQLPQKSILERKEVHHH HHHHHHH	4.52	39382.3
<b>SDC4 (45-384)</b>	MGSSHHHHHHSSGLVPRGSHMASMESIRETEVIDPQDLLE GRYFSGALPDEDEDAGGSDDFELSGSGDLDDTEEPFPEV IEPLVPLDNHIPPENAPGIRVPSEPKELEENEVIPKRAPSDV GDDMSNKVMSSTAQGSNIFERTEHHHHHHHHHHH	4.49	17519.8
<b>APEX2-TEV-SDC1 (18-252)</b>	MGKSYPTVSADYQDAVEKAKKKLRGFIAEKRCAPLMLRLAF HSAGTFDKGKTGGPFGTIKHPAELAHSANGLDIAVRLL PLKAEFPILSYADFYQLAGVVAVEVTGGPKVPFHPGREDKP EPPPEGRLPDPTKGS DHLRDVFGKAMGLTDQDIVALSGGH TIGAAHKERSGFEGPWTSNPLIFDNSYFTELLSGEKEGLLQL PSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSLG FADAGSGGGSENLYFQGMQPALPQIVAVNVPPEDQDGS GDDSDNFSGSGTGALPDTLSRQTPSTWKDVWLLTATPTAP EPTSSNTETAFTSVLPAGEKPEEGEPVLHVEAEPGFTARDK EKEVTTRPRETVQLPITQRASTVRVTTAQAAVTSHPHGGM QPGLHETSAPTAPGQPDHQPPRVEGGGTSVIKEVVEDGTA NQLPAGEGSGEQDFTFETSGENTAVAAVEPGLRNQPPVDE GATGASQSLLDRKEHHHHHHHHHHH	5.12	54288.99

**Supplementary Table 3.** SDC ectodomain expression yields.

<b>Ectodomain</b>	<b>Sequence</b>	<b>YIELD (MG/L)</b>
SDC1 <sub>37</sub>	18-252(His) <sub>10</sub>	0.9
SDC1 <sub>45</sub>	18-252(His) <sub>10</sub>	1.4
SDC1 <sub>47</sub>	18-252(His) <sub>10</sub>	1.7
SDC1 <sub>45,47</sub>	18-252(His) <sub>10</sub>	0.7
SDC1 <sub>37,45,47</sub>	18-252(His) <sub>10</sub>	0.4
SDC2 <sub>41,55,57</sub>	19-145(His) <sub>10</sub>	0.9
SDC3 <sub>80,82,89</sub>	45-384(His) <sub>10</sub>	1.9
SDC4 <sub>44,62,64</sub>	24-145(His) <sub>10</sub>	0.3
A-SDC1 <sub>45,47</sub>	APEX2-TEV-SDC1(18-252)(His) <sub>10</sub>	3.0
A-SDC1 <sub>37,45,47</sub>	APEX2-TEV-SDC1(18-252)(His) <sub>10</sub>	4.0

**Supplementary Table 4.** RT-qPCR primers.

<b>Gene</b>	<b>Forward primer (5' - 3')</b>	<b>Reverse primer (5' - 3')</b>
mGAPDH	TGCCTGCTTCACCTTCT	CCAATGTGTCCGTCGTGGAT
mNanog	CACAGTTTGCCTAGTTCTGAGG	GCAAGAATAGTTCTCGGGATGAA
mSOX1	GGCCGAGTGGAAAGGTCATGT	TCCGGGTGTTCCCTTCATGTG
hSDC1 (A)	CTCTGGGGAGCAGG	CTCCCAGCACCTCTTTCCTG
hSDC1 (B)	CTCTGGGGAGCAGG	GCACACAGCAAAGATGAGCC

**Supplementary Table 5.** SDC1 DsiRNA targets.

<b>DsiRNA</b>	<b>Cross-reacting transcript</b>	<b>Location</b>	<b>Exon</b>
1	NM_002997	3' UTR	5
	NM_001006946	3' UTR	6
2	NM_001006946	CDS	4
	NM_002997	CDS	3
3	NM_002997	CDS	3
	NM_001006946	CDS	4

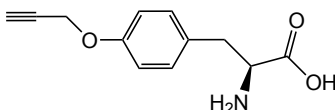


## Supplementary note 1

### Synthesis and characterisation of chemical structures

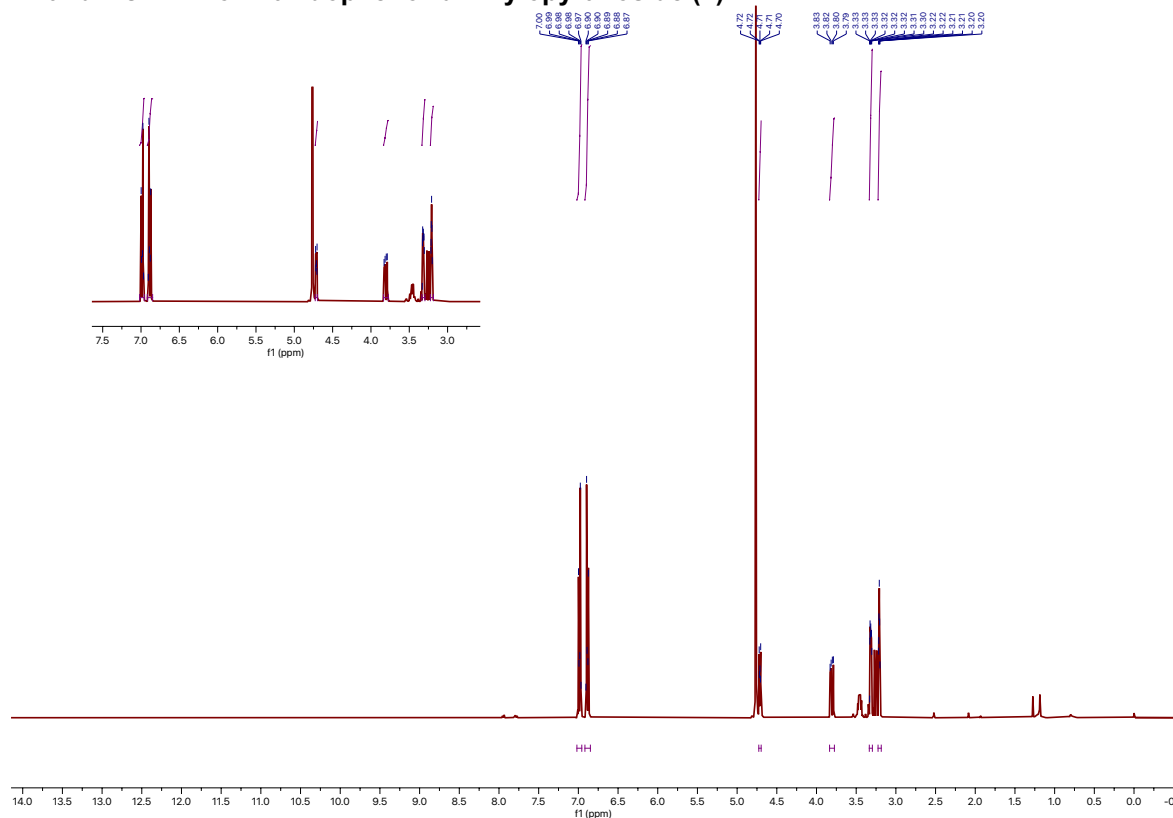
**NMR/Chemistry protocols.** All the starting material chemicals were purchased from commercial suppliers (Carbosynth, Sigma Aldrich, Flourochem and Acros) and used without further purification. Unless otherwise stated, all reactions containing air- and moisture sensitive reagents were carried out under an inert atmosphere of nitrogen in oven-dried glassware with magnetic stirring. All reactions were monitored by thin-layer chromatography (TLC) on Merck DC-Alufolien plates precoated with silica gel 60 F254. TLC plates were visualized with UV-light (254 nm) and stained with H<sub>2</sub>SO<sub>4</sub> (8 %). Silica gel column chromatography was carried out using Davisil silica gel or with automated flash chromatography suite (Biotage SP4 HPFC). <sup>1</sup>H NMR (400 or 600 MHz), <sup>13</sup>C NMR (125 or 151 MHz) spectra were recorded on Bruker AVANCE NEO 400 and 600 MHz instrument at 25 °C in chloroform (CDCl<sub>3</sub>), MeOH (CD<sub>3</sub>OD), and water (D<sub>2</sub>O).

**Preparation of *p*-propargyl tyrosine (pPY, 1) unnatural amino acid.** The synthesis of pPY was completed in three steps starting from commercially available *N*-Boc tyrosine, following the procedures outlined by Deiters et al<sup>1</sup>. Briefly, *N*-*tert*-butoxycarbonyltyrosine (2 g, 7 mmol, 1 eq.) in DMF (15 mL) was slowly reacted with propargyl bromide (2.1 mL, 21 mmol, 3 eq., 80% toluene) overnight at RT. Following extraction with water and Et<sub>2</sub>O, the combined organic layers were concentrated to yield the di-esterified product, a yellow oil, which was directly used in the next step. To a cooled (0 °C) solution of the crude material in 60 mL methanol was slowly added a solution of 5 M HCl/MeOH until complete product formation. The solution was allowed to warm to room temperature and volatiles were removed to yield the propargyl ester as a yellow solid. The resulting product was dissolved in a mixture of 2 M NaOH and MeOH for 2 hr at RT. The pH was adjusted to 7 and the mixture was kept at 4 °C overnight. The resulting precipitate was washed with ice-cold water and dried *in vacuo* to yield a beige solid (1.2 g, 78% isolated yield,).

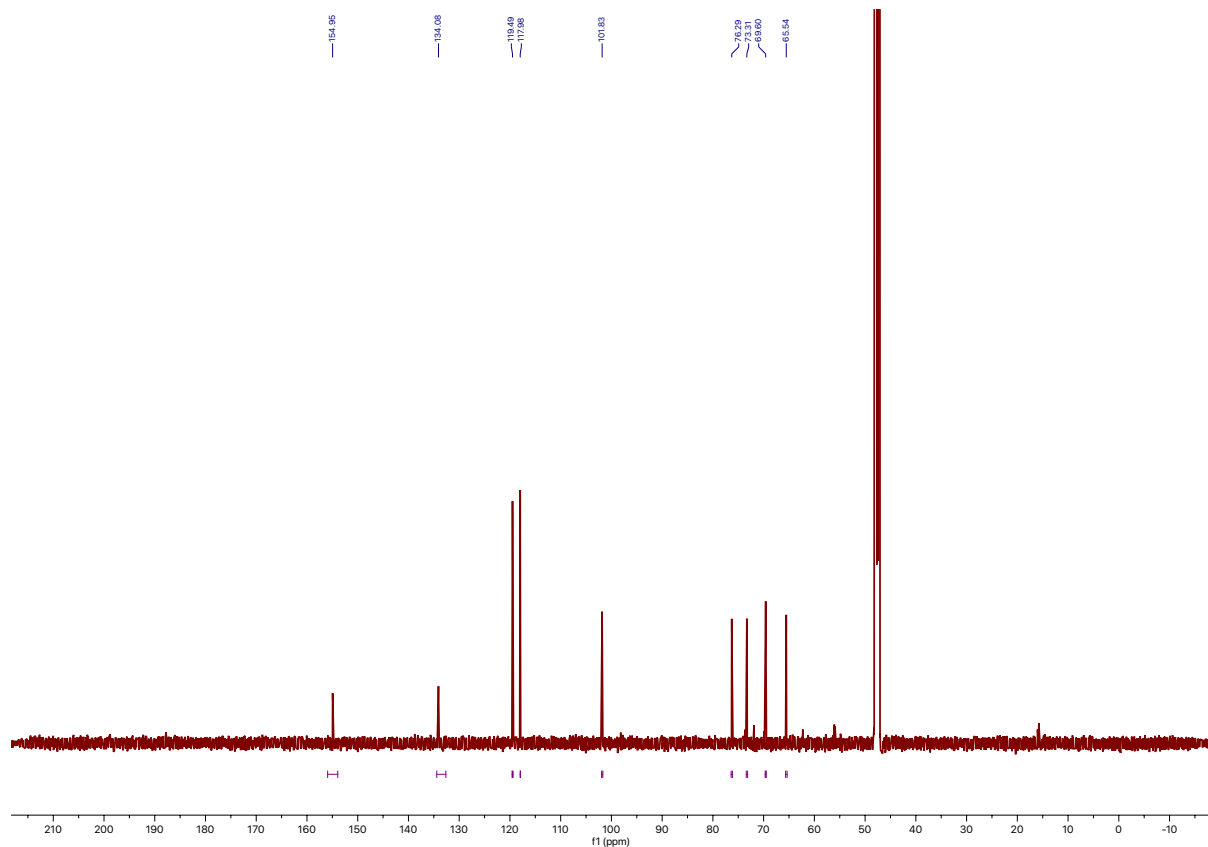




**<sup>1</sup>H and <sup>13</sup>C NMR of 4-azidophenol β-D-xylopyranoside (2).**

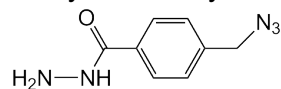


**<sup>1</sup>H NMR** (400 MHz, MeOD)  $\delta$  7.02 – 6.95 (m, 2H), 6.92 – 6.84 (m, 2H), 4.73 – 4.69 (m, 1H), 3.81 (dd,  $J = 11.4, 5.3$  Hz, 1H), 3.50 – 3.42 (m, 1H), 3.34 – 3.30 (m, 2H), 3.28 – 3.24 (m, 1H), 3.21 (dt,  $J = 3.3, 1.5$  Hz, 1H).

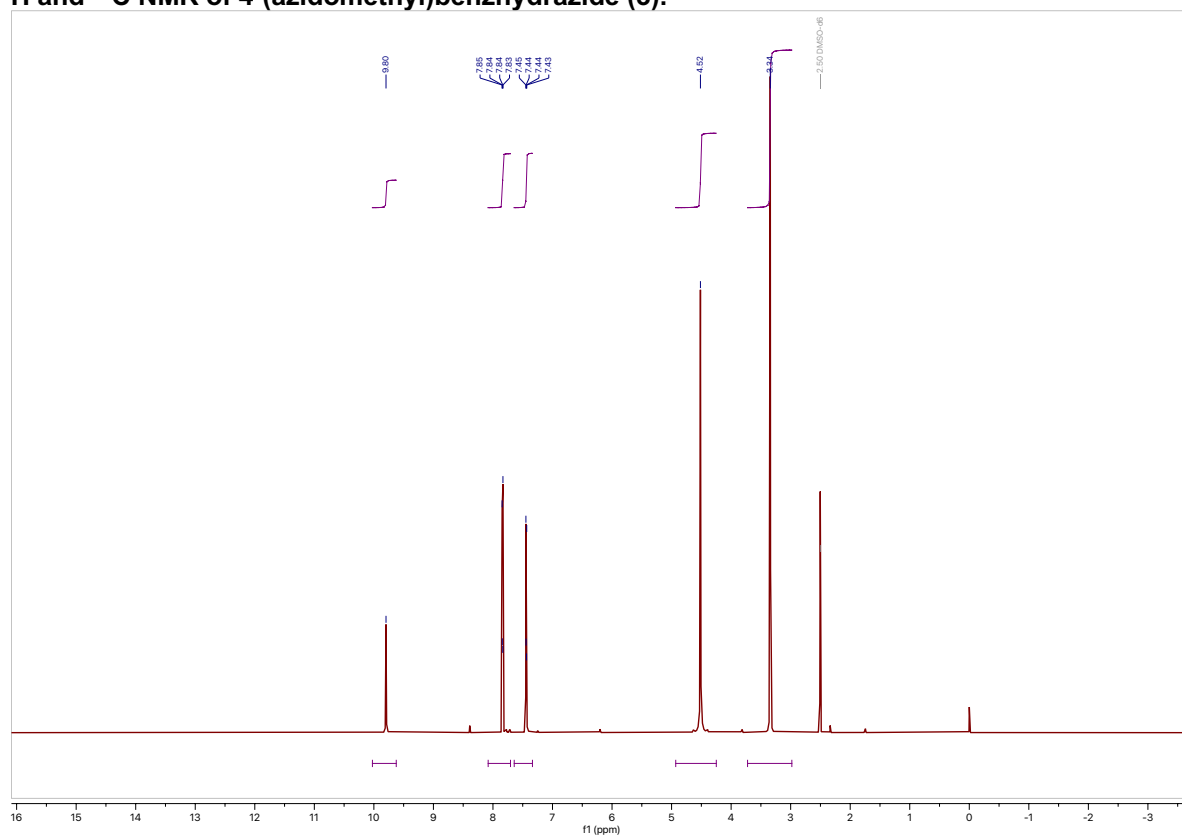


**<sup>13</sup>C NMR** (151 MHz, MeOD)  $\delta$  154.9, 134.0, 119.4, 117.9, 101.8, 76.2, 73.3, 69.6, 65.5.

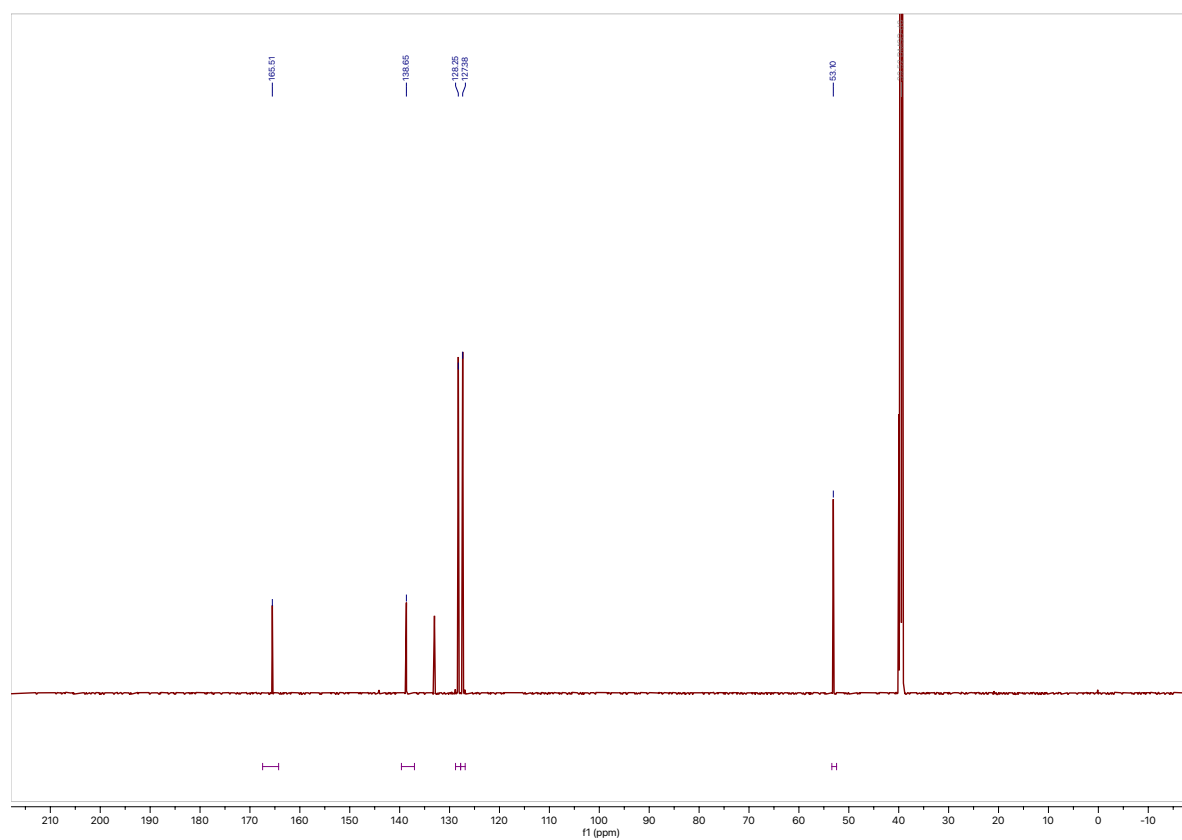
**Preparation of 4-(azidomethyl)benzhydrazide (3).** Starting from commercially available methyl 4-(bromomethyl)benzoate, the linker 4-(azidomethyl)benzhydrazide was synthesized following a two-step procedure<sup>3</sup>. Briefly, methyl 4-(bromomethyl)benzoate (4 g, 17.5 mmol, 1 eq) was reacted with NaN<sub>3</sub> (3.4 g, 52.3 mmol, 3 eq) in DMF (60 mL) overnight at 75 °C. Following extraction with H<sub>2</sub>O and EtOAc, the organic layer was evaporated under vacuum to yield an oil which was used without further purification. The crude material was dissolved in EtOH (60 mL) and refluxed with hydrazine monohydrate (6.8 g, 136 mmol, 7 eq) for 14 hr. Solvent was removed under vacuum. The resulting solid was recrystallized from 2:1 EtOH and EtOAc and washed with ice cold H<sub>2</sub>O to yield a fluffy white solid (1.5 g, 45%).



**<sup>1</sup>H and <sup>13</sup>C NMR of 4-(azidomethyl)benzhydrazide (3).**

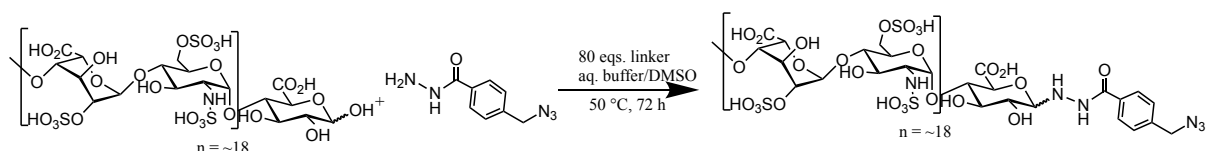


<sup>1</sup>H NMR (600 MHz, DMSO)  $\delta$  9.79 (s, 1H), 7.93 – 7.77 (m, 2H), 7.54 – 7.32 (m, 2H), 4.51 (s, 4H).



<sup>13</sup>C NMR (151 MHz, DMSO)  $\delta$  165.51, 138.65, 128.25, 127.38, 53.10.

**Preparation of azido-heparin (azHEP).** We followed a previously published procedure for the synthesis of azide-modified GAGs<sup>3</sup>. Heparin (121.3 mg in 1075  $\mu$ L aqueous buffer, 100 mM sodium acetate, 100 mM aniline, pH 5.5) was mixed with 615  $\mu$ L of 4-(azidomethyl)benzhydrazide (1.05 M in DMSO (123 mg, 80 eqs)). The mixture was protected from light and heated (50  $^{\circ}$ C, 72 hr). The reaction was diluted into 40 mL of PBS, filtered, (0.45  $\mu$ m) and dialyzed in 10 mM  $\text{NH}_4\text{CO}_3\text{H}$  (48 hr, exchange buffer 3 times). The sample was lyophilized for a yield of 66.8% (81.0 mg). Conjugation of 4-(azidomethyl)benzhydrazide to heparin was verified by the presence of two broad peaks in the aromatic region of the  $^1\text{H}$  NMR spectrum, and integration of anomeric and aromatic  $^1\text{H}$  NMR signals yielded a ratio of 1:0.048 (anomeric:aromatic), in agreement with the published procedure.

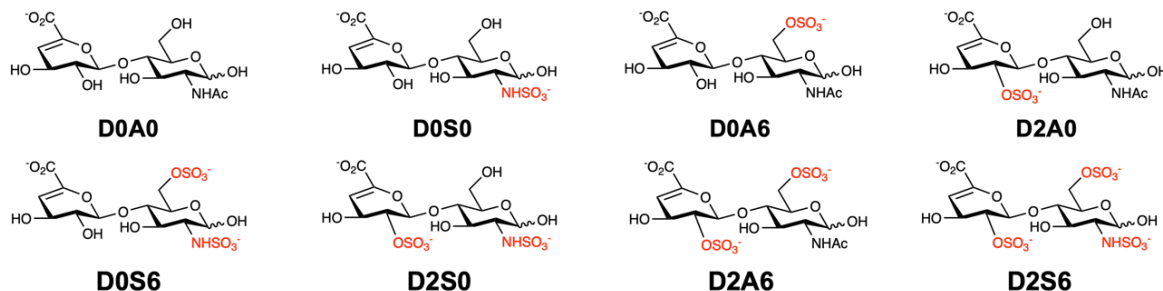


## Supplementary Note 2:

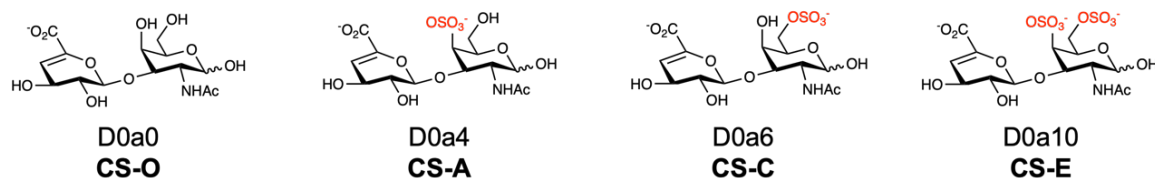
### Chemical structures of heparan sulfate and chondroitin sulfate disaccharides.

Depicted chemical structures were generated by enzymatic digestion of HS and CS GAG chains, respectively, and purchased commercially as disaccharide standards for analyses from Iduron (UK, distributed by Galen Laboratory Supplies).

#### Heparan sulfate (HS) disaccharides:



#### Chondroitin sulfate (HS) disaccharides:



#### References:

- 1 Deiters, A. *et al.* Adding amino acids with novel reactivity to the genetic code of *Saccharomyces Cerevisiae*. *J Am Chem Soc* **125**, 11782-11783 (2003).
- 2 Chapman, L. M., Beck, J. C., Wu, L. & Reisman, S. E. Enantioselective Total Synthesis of (+)-Psiguadial B. *J Am Chem Soc* **138**, 9803-9806, doi:10.1021/jacs.6b07229 (2016).
- 3 Trieger, G. W., Verespy, S., 3rd, Gordts, P. & Godula, K. Efficient Synthesis of Heparinoid Bioconjugates for Tailoring FGF2 Activity at the Stem Cell-Matrix Interface. *Bioconjug Chem* **30**, 833-840, doi:10.1021/acs.bioconjchem.8b00921 (2019).