

# Glycan Markers of Human Stem Cells Assigned with Beam Search Arrays

## Authors

Nian Wu, Lisete M. Silva, Yan Liu, Yibing Zhang, Chao Gao, Fuming Zhang, Li Fu, Yanfei Peng, Robert Linhardt, Toshisuke Kawasaki, Barbara Mulloy, Wengang Chai, and Ten Feizi

## Correspondence

t.feizi@imperial.ac.uk;  
w.chai@imperial.ac.uk

## In Brief

R-10G antigen of human iPS and ES cells has been assigned as a unique mono-sulfated glycan. We highlight its relationship to a known bioactive glycan sequence, the ligand on high endothelial venules for the lymphocyte homing receptor, L-selectin. Details of the sequences of four other glycan antigens on the podocalyxin are re-evaluated and ambiguities resolved. Regulation of biosynthesis and possible involvements of these glycans in podocalyxin-signaling in stem cells are discussed.

## Graphical Abstract

### Interrelationships of Human Stem Cell Antigens

<p>Type 2–type 2–type 2 backbones</p>	i antigen	
	R-10G antigen	
	Physiological ligand for L-selectin	
	Type 1–type 2–type 2 backbones	FC10.2, TRA-1-60, TRA-1-81 antigens

## Highlights

- R-10G is a mono-sulfated glycan antigen of human iPS and ES cells on podocalyxin.
- R-10G and nonsulfated i are reciprocal antigens on type 2 glycan backbones.
- TRA-1–60 and -81 and FC10.2 are antigens expressed on unsulfated type 1 backbones.
- These assignments open the way to probing their roles in podocalyxin-signaling.

# Glycan Markers of Human Stem Cells Assigned with Beam Search Arrays\*<sup>§</sup>

Nian Wu<sup>‡‡</sup>, Lisete M. Silva<sup>‡</sup>, Yan Liu<sup>‡</sup>, Yibing Zhang<sup>‡</sup>, Chao Gao<sup>‡||</sup>, Fuming Zhang<sup>§</sup>, Li Fu<sup>§</sup>, Yanfei Peng<sup>§</sup>,  Robert Linhardt<sup>§</sup>, Toshisuke Kawasaki<sup>¶</sup>, Barbara Mulloy<sup>‡</sup>,  Wengang Chai<sup>‡\*\*</sup>, and  Ten Feizi<sup>‡\*\*\*</sup>

**Glycan antigens recognized by monoclonal antibodies have served as stem cell markers. To understand regulation of their biosynthesis and their roles in stem cell behavior precise assignments are required. We have applied state-of-the-art glycan array technologies to compare the glycans bound by five antibodies that recognize carbohydrates on human stem cells. These are: FC10.2, TRA-1-60, TRA-1-81, anti-i and R-10G. Microarray analyses with a panel of sequence-defined glycans corroborate that FC10.2, TRA-1-60, TRA-1-81 recognize the type 1-(Gal $\beta$ -3GlcNAc)-terminating backbone sequence, Gal $\beta$ -3GlcNAc $\beta$ -3Gal $\beta$ -4GlcNAc $\beta$ -3Gal $\beta$ -4GlcNAc, and anti-i, the type 2-(Gal $\beta$ -4GlcNAc) analog, Gal $\beta$ -4GlcNAc $\beta$ -3Gal $\beta$ -4GlcNAc $\beta$ -3Gal $\beta$ -4GlcNAc, and we determine substituents they can accommodate. They differ from R-10G, which requires sulfate. By *Beam Search* approach, starting with an antigen-positive keratan sulfate polysaccharide, followed by targeted iterative microarray analyses of glycan populations released with keratanases and mass spectrometric monitoring, R-10G is assigned as a mono-sulfated type 2 chain with 6-sulfation at the penultimate *N*-acetylglucosamine, Gal $\beta$ -4GlcNAc(6S) $\beta$ -3Gal $\beta$ -4GlcNAc $\beta$ -3Gal $\beta$ -4GlcNAc. Microarray analyses using newly synthesized glycans corroborate the assignment of this unique determinant raising questions regarding involvement as a ligand in the stem cell niche. *Molecular & Cellular Proteomics* 18: 1981–2002, 2019. DOI: 10.1074/mcp.RA119.001309.**

Cellular glycans were historically among the earliest biochemical and immunochemical markers in embryonic developments (1–3). Using as reagents glycan sequence-specific monoclonal antibodies (initially, natural human antibodies followed by murine hybridoma-derived) remarkable insights were gained into the expression and polarization of glycans in post-implantation embryos and teratocarcinomas of mouse associated with cell differentiation (4, 5). We overview here below the glycan antigens on murine and human stem cells

that have been reported over the years. These antigens have been used as markers of stem cells (1) although little is known about their roles in stem cell behavior.

Undifferentiated murine embryonal carcinoma cells were observed to be rich in surface-associated and cytoplasmic I antigen (4). This antigen is recognized by human monoclonal anti-I Ma and anti-I Step; these being designations of monoclonal antibodies in sera of cases of cold agglutinin disease (6). These antibodies are directed at different domains on the branched structures of the poly-*N*-acetylglucosamine series as indicated in structures 1 and 2 (Table I). These are glycan sequences consisting of repeated type 2 (T2, Gal $\beta$ -4GlcNAc) on glycoproteins and glycolipids, where Gal is galactose and GlcNAc is *N*-acetylglucosamine. Differentiation into primary endoderm was observed to be associated with the appearance of i antigen (4) recognized by the human anti-i Den, which is directed at the unbranched repeated T2 sequence as in structure 3 (Table I). With the advent of hybridoma technology, there followed the demonstration that the 8-cell stage-specific embryonic antigen of mouse recognized by mAb anti-SSEA-1 is expressed on the  $\alpha$ -3 fucosylated forms of the I and i antigens (termed Lewis<sup>x</sup>), structures 4 and 5, respectively (5) (Table I). These antigens were detected predominantly on high molecular weight glycoproteins (7, 8) the common feature being the abundance of repeated T2 backbone sequences.

In contrast with the murine embryonic endoderm and embryonal carcinoma-associated antigens, the human embryonic endoderm- and embryonal carcinoma-associated antigen recognized by the monoclonal mAb FC10.2 was assigned as the lacto-*N*-tetraose sequence, with T1 (Gal $\beta$ -3GlcNAc)-T2 backbone sequence, structure 6 (Table I); this was detected on a glycoprotein with an apparent molecular weight, ~200 kDa, (9, 10) similar to that of podocalyxin (11). This glycoprotein is the carrier of the antigen(s) recognized by more recently generated mAbs, TRA-1-60 and TRA-1-81,

From the <sup>‡</sup>Glycosciences Laboratory, Department of Medicine, Imperial College London, London W12 0NN, United Kingdom; <sup>§</sup>Department of Chemical and Biological Engineering, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York 12180; <sup>¶</sup>Research Center for Glycobiotechnology, Ritsumeikan University, Noji-Higashi, 1-1-1, Kusatsu Shiga 525-8577, Japan; <sup>||</sup>Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215

Received January 14, 2019, and in revised form, June 21, 2019

\* Author's Choice—Final version open access under the terms of the Creative Commons CC-BY license.

Published, MCP Papers in Press, July 15, 2019, DOI 10.1074/mcp.RA119.001309

# Assignment of Glycan Markers of Human Stem Cells

TABLE I  
Murine and human stem cell-associated carbohydrate markers and carrier molecules

	mAbs and lectin	Cells bound	Carbohydrate markers		Carrier molecules	
			Sequences bound	Refs		
Murine	Anti-I Ma	EC	1 <sup>a</sup>	Galβ-4GlcNAcβ-6   Galβ-	(2, 4, 7, 65, 66)	Glycoprotein (Embryoglycan) (Glycolipid)
	Anti-I Step	EC	2	Galβ-4GlcNAcβ-6   Galβ-4GlcNAcβ-3Galβ-4GlcNAc/Glc   GlcNAcβ-3		
	Anti-i Den	EC	3	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc	(4, 7, 66, 67) (68)	Glycoprotein
	SSEA-1	ES and EC	4	Galβ-4GlcNAcβ-6   Galβ-4GlcNAcβ-3Gal   Galβ-4GlcNAcβ-3   Fuca-3	(2, 5, 7, 8, 69, 70)	Glycoprotein (Embryoglycan)
5			Galβ-4GlcNAcβ-3Galβ-4GlcNAc   Fuca-3			
Human	FC10.2	EC	6	Galβ-3GlcNAcβ-3Galβ-4GlcNAc/Glc	(9, 10)	Glycoprotein
			7	Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc/Glc	Tables 2 and 3, Table S4	
	TRA-1-60	ES, iPS and EC	6	Galβ-3GlcNAcβ-3Galβ-4GlcNAc/Glc	(11, 12, 23)	Podocalyxin
			7	Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc/Glc	Tables 2 and 3, Table S4	
	TRA-1-81	ES, iPS and EC	6	Galβ-3GlcNAcβ-3Galβ-4GlcNAc/Glc	(11, 12, 23)	Podocalyxin
			7	Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc/Glc	Tables 2 and 3, Table S4	
	SSEA-5	ES and iPS	8	Fuca-2Galβ-3GlcNAc	(71)	Glycoprotein
	R-17F	ES and iPS	9	Fuca-2Galβ-3GlcNAcβ-3Galβ-4Glc	(72, 73)	Glycolipid and podocalyxin
	rBC2LCN	iPS	10	Fuca-2Galβ-3   GalNAc   Galβ-3GlcNAcβ-6	(74-76)	Podocalyxin
	Anti-i serum <sup>b</sup>	Mesenchymal stem cells	3	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc	(18)	Glycoprotein
	SSEA-3	ES and EC	11	Galβ-3GalNAcβ-3Galα-4Galβ-4Glc	(14, 77)	Glycolipid
	SSEA-4	ES and EC	12	NeuAcα-3Galβ-3GalNAcβ-3Galα-4Galβ-4Glc	(14, 77)	Glycolipid
	anti-sialyl-lactotetra	ES and iPS	13	NeuAcβ-3Galβ-3GlcNAcβ-3Galβ-4Glc	(15)	Glycolipid
	R-10G	ES and iPS	14	Galβ-4GlcNAcβ-3Galβ-4GlcNAc                          6S                      6S	(19, 21)	Podocalyxin

<sup>a</sup> Numbers in columns 4 are designations of the sequences cited in the text. Glycan sequences recognized by anti-I Ma and Step, anti-i Den, mAbs FC10.2, TRA-1-60 and TRA-1-81 were originally determined using glycolipids or glycans with Glc at the reducing ends. On stem cells they occur predominantly on glycoproteins.

<sup>b</sup> An unnamed human serum.

which recognize human stem cells (12), and, like FC10.2, bind to the T1-T2 sequence, structures 6 and 7 (Table I). The expression of this antigenic activity is reported to be common to various types of human stem cells: embryonic stem cell (ES)<sup>1</sup>, induced pluripotent stem cell (iPS), and embryonal carcinoma-associated stem cell (EC) (12, 13).

Three other T1-related antigens common to human ES and iPS have been identified: two are T1-based blood group H recognized by mAbs anti-SSEA-5 and R-17F, structures 8 and 9 (Table I). The third is branched sequence structure 10 (Table I), containing blood group H based on the T3 (Gal $\beta$ -3GalNAc) backbone (where GalNAc is *N*-acetylgalactosamine). This antigen is bound by the lectin rBC2LCN, and reported to be cross-reactive with T1-based blood group H, and expressed on podocalyxin. The R-17F antigen is additionally borne on a glycolipid. Three other glycolipid-borne antigens on human ES and EC recognized by mAbs anti-SSEA-3, anti-SSEA-4 and anti-Sialyl lactotetra antibodies are structures 11, 12 and 13 respectively (Table I), of which the first two are of the globo series based on Gal $\beta$ -3GalNAc $\beta$ -3Gal $\alpha$ -4Gal $\beta$ -4Glc (14) and the third is based on the type 1 lacto series, Gal $\beta$ -3GlcNAc $\beta$ -3Gal $\beta$ -4Glc (15). Other studies have shown that not only the terminal epitopes but also the core structures are differentially expressed in stem versus differentiated cells (16, 17).

T2-based human stem cell antigens have been two in number. The first is a human mesenchymal stem cell-associated antigen expressed on the repeated T2-T2 based backbone, namely, i antigen (18). The second is recognized by mAb R-10G, which distinguishes human ES and iPS from human EC cells (19). It was shown that binding of mAb R-10G to an antigen-positive glycoprotein was unaffected by its treatment with neuraminidases and fucosidases. The antibody was shown to bind to the polysaccharide keratan sulfate (KS), which is based on long repeated T2-T2 sequences (19). A highly sulfated KS from bovine articular cartilage (with sulfation on many of the GlcNAc and Gal residues) was lacking in R-10G antigen activity, whereas bovine corneal KS, which contains some regions with sulfated GlcNAc and nonsulfated Gal residues (20), was antigen-positive (19). For these reasons, the R-10G antigen was deduced to be expressed on an under-sulfated form of KS. While the present study was under way a KS-related tetrasaccharide, structure 14, which has sulfate at position 6 of each of two GlcNAc residues (Table I)

<sup>1</sup> The abbreviations used are: ES, embryonic stem; AOPE, *N*-aminoxyacetyl-1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine; BMP, bone morphogenetic proteins; CS, chondroitin sulfates; EC, embryonal carcinoma; FGF, fibroblast growth factor; GAG, glycosaminoglycan; Gal, galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcNAc, *N*-acetylglucosamine; iPS, induced pluripotent; KS, keratan sulfate; MIRAGE, minimum information required for a glycomics experiment; NGL, neoglycolipid; RI, refractive index; SAX, strong anion-exchange; Shh, sonic hedgehog; sulfate(6S), sulfate at position 6 of GlcNAc; Wnt, int/wingless.

was chemically synthesized and when biotinylated was bound by mAb R-10G (21).

Knowledge of the glycan markers has not yet been matched by knowledge of the regulation of their biosynthesis and their functions in the stem cell niche. There also remain some ambiguities in reports of the binding specificities, for example those of the TRA-1-60 and -81 toward KS-like saccharides (19, 22, 23). Glycan arrays, since their inception (24) have been a powerful means of analyzing glycan-binding specificities of diverse carbohydrate-recognition systems (25, 26). The continued expansion of libraries of sequence-defined glycan probes offers increasing opportunities to keep under review the repertoires of glycans recognized by the anti-stem cell antibodies. There are moreover technological advances that enable the generation of “bespoke” glycome probes, which we have termed *Designer* arrays (27) and Beam Search arrays (28) from antigen positive macromolecules to detect, isolate and characterize natural determinants.

Here we use an array of sequence-defined glycans to make for the first time a close comparison of the glycan-binding specificities of the five antibodies that recognize human stem cells: mAbs TRA-1-60, TRA-1-81 and FC10.2 that broadly recognize ES, iPS and EC cells, an anti-i P1A ELL directed at linear poly-*N*-acetyllactosamine sequences (29), which according to ref (30) is predicted to recognize mesenchymal stem cells, and R-10G that recognizes ES and iPS cells. To assign the determinant of mAb R-10G, which is clearly different from those of the four other antibodies, we use the Beam Search microarray approach (28) using an antigen-positive KS polysaccharide as the macromolecular start-point. This is followed by a targeted, iterative microarray analyses and concomitant mass spectrometry of glycan populations derived from the polysaccharide after partial depolymerization by keratanase I or keratanase II or acid hydrolysis. Microarray analyses with newly synthesized glycans corroborate the assignment we have made for the unique sequence of R-10G antigen associated with ES and iPS cells.

#### EXPERIMENTAL PROCEDURES

The glycan array-based data are reported in accordance with MIRAGE (Minimum Information Required for a Glycomics Experiment) Glycan Microarray Guidelines (31). These include details of the glycan binding samples investigated (antibodies, growth factors and morphogens) and their detection systems, the saccharides and the neoglycolipids (NGLs) in the glycan libraries for different array sets, the conditions used for array construction, as well as information on image and data analyses: these are summarized in the Supplemental Glycan Microarray document based on MIRAGE Guidelines ([supplemental Table S2](#)). The experimental data not covered are given below.

*Keratanases*—Keratanase I from *Pseudomonas* sp. (EC 3.2.1.103) was from Seikagaku (Joetsu, Japan). Recombinant Keratanase II from *Bacillus circulans*. (EC 3.2.1) was as prepared according to (32). Please note that to avoid confusion in the names of the two KS specific enzymes, the term “keratanase I” was used throughout the paper instead of the commonly used “keratanase.”

*Partial Depolymerization of KS and Fractionation of KS Derived Glycans*—The conditions for partial digestion of KS with keratanase I

were adapted from those described by Nakazawa and Suzuki (33). In brief, KS (10 mg) was incubated with 0.5 unit of keratanase I in 0.5 ml Tris-HCl buffer (20 mM, pH 7.4) at 37 °C. Partial digestion of KS with keratanase II was performed as described by Brown *et al.* (34). In brief, KS (20 mg) was incubated with 0.96 unit of keratanase II in 2 ml NH<sub>4</sub>OAc buffer (50 mM, pH 7.4) at 37 °C. For partial acid hydrolysis, KS (1.4 mg) was incubated with 0.1 M HCl at 80 °C. The partial depolymerization of bovine corneal KS using keratanase I, keratanase II and acid hydrolysis was monitored by Superdex Peptide column (1 × 30 cm, GE Healthcare, Fairfield, Connecticut) and ESI-MS, and stopped at 23 h, 7.5 h and 4 h, respectively, at the disappearance of the peak at void volume (supplemental Fig. S1). The keratanase digestion reactions were stopped by heating the solutions in a water bath at 100 °C for 1 min, whereas the acid hydrolysis was stopped at the indicated times by neutralizing with 0.1 M NaOH. The reaction mixtures were lyophilized.

**Preparation of Fractions of KS Glycome**—Bio-Gel P-6 (1.6 × 90 cm, Bio-Rad Laboratories, Hercules, California) was used for size fractionation of glycans after keratanase treatment of KS, and elution was with 200 mM NH<sub>4</sub>Cl at a flow rate of 15 ml/h. Eluates were monitored by a refractive index (RI) detector, fractions were pooled as indicated (supplemental Fig. S2A and S2B) and lyophilized. Sephadex G10 column (1.6 × 35 cm, GE Healthcare) was used for desalting of fractions collected from Bio-Gel P-6 chromatography. Deionized water was used for elution at a flow rate of 20 ml/h and eluates were monitored with a RI detector. Superdex Peptide column (GE Healthcare, 1 × 30 cm) was used to fractionate the products obtained after acid hydrolysis of KS. Elution was with 50 mM NH<sub>4</sub>OAc at a flow rate of 0.3 ml/min and monitored by RI. Fractions were pooled as indicated and lyophilized and were coevaporated repeatedly with Deionized water by freeze-drying to remove the volatile buffer salt NH<sub>4</sub>OAc (supplemental Fig. S2C). Selected fractions were analyzed by ESI-MS to determine chain lengths and sulfate contents (supplemental Table S6). The molar quantities of the glycans in the KS fractions were estimated based on the galactose content and molecular mass determined by MS analysis. The galactose contents were determined by established dot orcinol-sulfuric acid method on TLC plates, using galactose as standard (35).

Strong anion-exchange (SAX) column (4.6 × 250 mm, Waters, Milford, Massachusetts) or HiTrap ion exchange cartridge (7 × 25 mm, GE Healthcare) or hypercarb porous graphitized carbon (PGC) column (30 × 4.6 mm, Thermo Scientific, Waltham, Massachusetts) were used for fractionation of glycans by HPLC (Gilson, Madison, Wisconsin). Elution was carried out by a gradient of LiClO<sub>4</sub> for both SAX-HPLC and HiTrap cartridge (Solvent 1: 10 mM LiClO<sub>4</sub> and Solvent 2: 500 mM LiClO<sub>4</sub>). Chromatographies by SAX and HiTrap were performed in 40 min at a flow rate of 1 ml/min with detection by UV at 206 nm. For PGC-HPLC the elution was by a gradient of H<sub>2</sub>O/ACN (Solvent 1: H<sub>2</sub>O; Solvent 2: ACN/H<sub>2</sub>O: 80:20; both containing 0.05% trifluoroacetic acid) in 40 min at a flow rate of 0.5 ml/min with detection by UV at 206 nm. The gradients used are given in the figures. Fractions were pooled as indicated in the respective figures. The fractions obtained from SAX-HPLC and HiTrap were lyophilized and desalted with a Superdex Peptide column and lyophilized as described above. To fractions obtained from PGC-HPLC were added 50 mM NH<sub>4</sub>HCO<sub>3</sub> before removal of the ACN under N<sub>2</sub> stream and the TFA by repeated lyophilization.

**Desulfation of KS Glycans**—Desulfation of KS glycans was performed as described (36). After several preliminary experiments, the conditions selected were as follows: freeze-dried hexasaccharides substituted with 4 to 6 sulfates (~100 nmol) were added to 150 μl of 20 mM methanolic HCl for incubation at room temperature for 17 h. The reactions were stopped by immediate removal of the methanolic

HCl reagent after addition of 50 μl of 50 mM NH<sub>4</sub>HCO<sub>3</sub>. The reaction product was dried by lyophilization.

**Lipid Conjugation of KS Oligosaccharides to Prepare NGLs**—NGLs were prepared from reducing KS glycans by oxime-ligation using the lipid reagent *N*-aminooxyacetyl-1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine, AOEPE, essentially as described (37). The reactions were on a scale between 10 nmol and 100 nmol of starting oligosaccharides depending on their availability. In brief, for short chain glycans (di- to octa-saccharides), five molar equivalents of AOEPE solution (10 nmol/μl) were added, whereas 10 molar equivalents were added to longer glycans. The incubations were at 37 °C for 16 h (shorter glycans) or 32 h (longer glycans). The purification and quantitation of NGLs were as described (35).

**MS Analysis**—Negative ion ESI-MS and CID-MS/MS were carried out on a Waters (Manchester, UK) Q-TOF-type mass spectrometer SYNAPT-G2. Cone voltage was at 80 eV and capillary voltage at 3 kV. Source temperature was at 80 °C and the desolvation temperature at 150 °C. The resolution was set at 20,000 for MS and 8,000 for CID-MS/MS with a scan rate at 1.5 s for the mass range 100–2000. Product-ion spectra were obtained using argon as the collision gas at a pressure of 1.7 bar. The collision energy was adjusted between 12 and 20 eV for optimal fragmentation. A scan rate of 1.5 s/scan was used for both ESI-MS and MS/MS experiments and the acquired spectra were summed for presentation. For analysis, glycans were dissolved in water and NGLs in CHCl<sub>3</sub>/MeOH/H<sub>2</sub>O, 25:25:8 (v/v/v) typically at a concentration of 10–20 pmol/μl, of which 1 μl was injected via a HPLC injector. Solvent (ACN/2 mM NH<sub>4</sub>HCO<sub>3</sub>, 1:1, or CHCl<sub>3</sub>/MeOH/H<sub>2</sub>O, 25:25:8) was delivered by a HPLC pump (Waters) at a flow rate of 10 μl/min.

**NMR Spectroscopy**—<sup>1</sup>H-NMR, as well as 2D <sup>13</sup>C-<sup>1</sup>H HSQC and TOCSY NMR spectra of glycans were recorded at 700 MHz using a BrukerAvance spectrometer (Karlsruhe, Germany) equipped with cryoprobe, in D<sub>2</sub>O solution at 30 °C. 1D and 2D pulse sequences were supplied by the equipment manufacturer. The observed <sup>1</sup>H chemical shifts were relative to internal acetone (2.225 ppm). Each sample (~240 μg) was taken up in D<sub>2</sub>O 99.9% (Apollo Scientific, Stockport, UK) exchanged by lyophilization, then re-dissolved in 300 μl D<sub>2</sub>O and transferred to a Shigemi NMR microtube (Sigma, St. Louis, Missouri).

**Generation of Microarrays of Polysaccharides and NGLs and Microarray Analysis**—The generation of microarrays of polysaccharides and NGLs is essentially as described (38). There were six microarrays: one of glycosaminoglycan (GAG) polysaccharides and five of lipid-linked glycans collectively termed NGL arrays, of which one contained 79 sequence-defined NGL probes, in house designated CLL microarray, and four NGL Beam Search arrays. The details are summarized in supplemental Table S2.

Microarray analyses were performed essentially as described (38) at ambient temperature except for anti-i P1A ELL at 4 °C. Concentrations of the antibodies analyzed, the blockers and diluents, as well as the concentrations of polysaccharides and lipid-linked probes arrayed are as described in supplemental Table S2.

R-10G was analyzed at 100 μg/ml without precomplexation in the analyses with CLL microarray and KS-derived Beam Search arrays, and at 10 μg/ml with precomplexation in analyses with polysaccharide microarray. For precomplexation, R-10G was incubated for 30 min with the goat anti-mouse IgG (detection antibody) at a ratio of 1:3 (w/w) before overlay for 90 min.

The other antibodies MZ15, 5D4, CS56, anti-i P1A ELL, FC10.2, TRA-1–60, TRA-1–81 and anti-LNT were analyzed in CLL microarray, polysaccharide microarray and KS-derived Beam Search arrays, without precomplexation as indicated (supplemental Table S2).

Assay conditions used for the exploratory microarray analyses of recombinant proteins FGF1, FGF2, Shh, Wnt-3a, BMP-2 and BMP-4 are described in supplemental Table S2.

## RESULTS

*A Close Comparison of Glycans Recognized by mAbs R-10G, anti-i P1A ELL, FC10.2, TRA-1-60 and TRA-1-81 with Sequence-defined Glycans Based on T1-T2 and T2-T2 Backbones*—Initial analyses were performed with R-10G, anti-i P1A ELL, FC10.2, TRA-1-60, and TRA-1-81 using a microarray of sequence-defined glycans based on T1-T2 and T2-T2 backbones (supplemental Table S3). This was (a) to determine any binding of R-10G to T2-T2 backbones of the type that occur in KS, but without sulfation, compared with anti-i P1A ELL, and (b) to make the close comparison hitherto not possible, between TRA-1-60 and TRA-1-81 with FC10.2 with respect to glycan sequences they recognize. For reference, anti-LNT directed at the T1-T2 backbones was also analyzed.

R-10G gave no significant binding to the glycans arrayed (Table II and supplemental Table S4). This contrasted with the strong binding observed with anti-i P1A ELL to the nonsubstituted T2-T2 structure at position 14 and the 6-sialylated structure at positions 16 and 17 in the array (supplemental Table S4).

FC10.2, TRA-1-60, TRA-1-81 gave strong binding signals with three structures at positions 11, 27 and 28 in the array (Table II and supplemental Table S4). The common feature is the presence of nonreducing end T1 followed by internal T2 as a part of linear or branched glycan chains. In this respect, they resembled the commercial anti-LNT (supplemental Table S4). Although the glycan probes had glucose (Glc) at the reducing end, there would be GlcNAc here as part of the *O*-glycan sequences of a glycoprotein such as podocalyxin (39). Differences were revealed among these four antibodies: FC10.2, anti-LNT and TRA-1-60 but not TRA-1-81 gave binding albeit weak to the structure with a fucose (Fuc) 3-linked to the inner GlcNAc (position 43), and the binding was abolished in the presence of 2-linked Fuc at the nonreducing end (position 44).

In sum, the three anti-stem cell mAbs FC10.2, TRA-1-60 and TRA-1-81 are very similar in their strong binding to the nonsubstituted T1-T2 sequences as part of linear or branched glycans longer than tetrasaccharide. Unlike anti-i P1A ELL, R-10G gave no detectable binding to nonsulfated T2-T2 sequences.

*Corneal KS Selected for Beam Search Array*—R-10G, anti-i P1A ELL, FC10.2, TRA-1-60 and TRA-1-81 were next analyzed using the GAG polysaccharide microarray (Fig. 1). The array included three KS preparations isolated from bovine cornea, bovine articular cartilage and chicken egg white, and commercial preparations of chondroitin sulfates (CS) A, B, and C, heparin, two heparan sulfates isolated from porcine intestinal mucosa and bovine kidney, and hyaluronan. For reference, dextran and highly sulfated dextran were also analyzed.

Among the five anti-stem cell mAbs, R-10G gave strong binding only to the corneal KS, very weak binding to the egg white KS and negligible binding with the articular cartilage KS. There was no significant binding observed to any of the other

types of GAGs except very weak binding to CSA (Fig. 1). With the anti-i P1A ELL serum the only significant binding detected was to the dextran sulfate (Fig. 1). No binding was detected with FC10.2, TRA-1-60, and TRA-1-81.

Among the antibodies included for reference, the anti-KS mAbs MZ15 and 5D4 bound strongly to corneal KS, as predicted. They also bound to the two other KS preparations (Fig. 1). CS56 gave binding signals with the Sigma CSA and CSC in accordance with previous knowledge (40). An unpredicted finding was strong binding of the anti-KS mAbs MZ15 and 5D4 to the CSA preparation from bovine trachea. In the original investigations of the specificities of MZ15 and 5D4 interactions with chondroitin sulfates were not observed (41, 42). Those studies were carried out using proteoglycans extracted from bovine nasal cartilage and porcine larynx, respectively. The source of CSA (Sigma C8529) used in our studies is bovine trachea, which is known to contain KS (43). The binding observed with MZ15 and 5D4, as well as R-10G, was suggestive of presence of KS in this polysaccharide. Additional analyses with newly purchased batches of bovine tracheal CSA and shark cartilage CSC from Sigma (supplemental Table S1) show that different batches of these polysaccharides vary in their content of KS immunoreactivity.

In sum, our analyses to address the possible KS binding by the anti-stem cell mAbs TRA-1-60, and TRA-1-81 (22) show that neither of these antibodies nor FC10.2 bind to KS or to the other polysaccharides in the array. Among the four anti-stem cell antibodies, only R-10G binds to bovine corneal KS. We selected this polysaccharide to identify and structurally characterize glycan sequence(s) recognized by mAb R-10G by the Beam Search approach (Fig. 2).

*Preparation of a Primary Beam Search Array of Glycan Fractions of the KS Glycome*—The bovine corneal KS was partially depolymerized using keratanases I and II and by mild acid hydrolysis (Experimental Procedures; supplemental Figs. S1 and S2). Ten glycan fractions, 2 to 20 mers of the keratanase I digest, designated K'ase I hereinafter, and nine fractions, 2 to 18 mers of the keratanase II digest, designated K'ase II, were pooled as indicated (supplemental Figs. S2A and S2B, respectively). Keratanase I is known to cleave the glycosidic bond Gal $\beta$ -4GlcNAc(6S) to produce glycans with GlcNAc(6S) at their nonreducing ends and Gal at their reducing ends, whereas keratanase II cleaves the glycosidic bond GlcNAc(6S) $\beta$ -3Gal( $\pm$ 6S) to produce glycans with Gal( $\pm$ 6S) at their nonreducing ends and GlcNAc(6S) at their reducing ends (33, 44). A series of even-numbered glycans were thus produced (Fig. 2).

NMR analyses (supplemental Fig. S3) performed with two tetrasaccharide analogs from K'ase I and K'ase II series (supplemental Table S6) as examples corroborated that the K'ase I products contained glycans with Gal at the reducing ends, and that the K'ase II products contained glycans with GlcNAc(6S) at the reducing ends.

## Assignment of Glycan Markers of Human Stem Cells

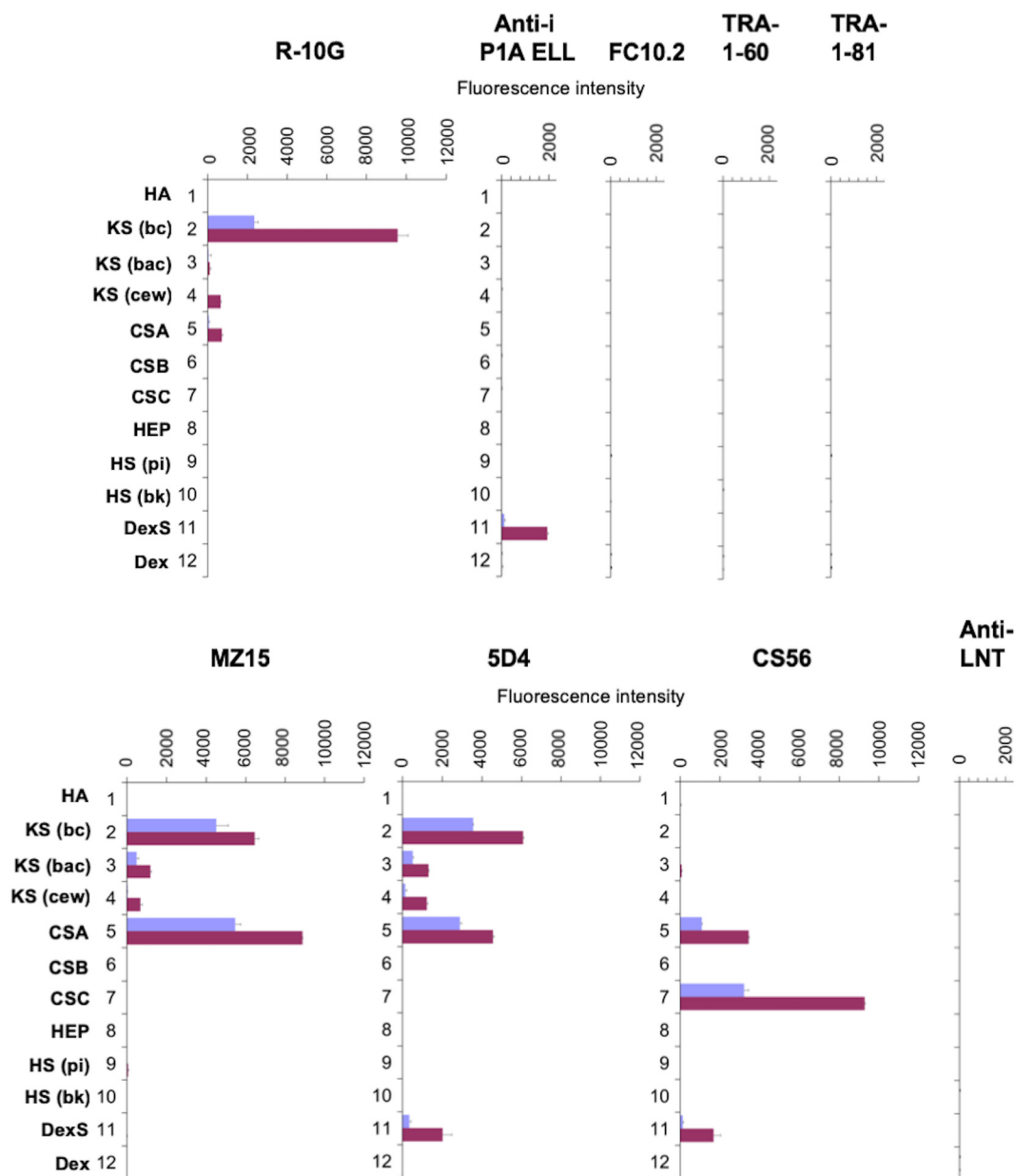
TABLE II

Relative intensities of binding of mAbs R-10G, anti-i P1A ELL, FC10.2, TRA-1-60, TRA-1-81, and anti-LNT to antigen-positive and related antigen-negative probes in sequence-defined NGL microarray (in house designation "CLL microarray set")

Full data are in [supplemental Table S4](#). Binding scores are means of the fluorescence intensities at 5 fmol/probe spot; - indicates the score is less than 1. Relative binding intensity:  < 10%,  10–30%,  30–70%,  70–100%, where 100% is the maximum score for respective antibodies.

Pos*	Sequences	R-10G	Anti-i P1A ELL	FC10.2	TRA-1-60	TRA-1-81	Anti-LNT
11	Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc-DH	-	-	22,484	17,349	15,640	5,821
12	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc-DH	-	2,601	-	-	-	-
14	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc-DH	-	24,927	-	-	-	-
15	NeuAcα-3Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc-DH	-	1,606	-	-	-	-
16	NeuAcα-6Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc-DH	-	37,699	-	-	-	-
27	Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-6   Galβ-4Glc-DH   Galβ-3GlcNAcβ-3	-	-	27,462	19,909	17,594	8,621
28	Galβ-4GlcNAcβ-6   Galβ-4GlcNAcβ-6   Galβ-3GlcNAcβ-3   Galβ-4Glc-DH   Galβ-3GlcNAcβ-3	-	-	31,456	20,845	18,855	13,812
43	Fuca-3   Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-6   Galβ-4Glc-DH   Fuca-2Galβ-3GlcNAcβ-3	-	-	4,144	7,461	568	1,578
44	Fuca-3   Fuca-2Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-6   Galβ-4Glc-DH   Fuca-2Galβ-3GlcNAcβ-3	-	-	603	695	27	31

\* Pos, position in the microarray.



**FIG. 1. Screening analyses of anti-stem cell and reference antibodies using the GAG polysaccharide microarray.** The four stem cell antibodies, R-10G, FC10.2, TRA-1-60 and TRA-1-81 were analyzed together with mAbs MZ15, 5D4, CS56, anti-i P1A ELL and anti-LNT as controls. Abbreviations for the polysaccharides: hyaluronic acid (HA); keratan sulfate (KS) isolated from bovine cornea (bc), bovine articular cartilage (bac) and chicken egg white (cew); chondroitin sulfate (CS) A, B and C; heparin, heparan sulfate (HS) isolated from porcine intestinal mucosa (pi) and bovine kidney (bk); Dextran sulfate (DexS); Dextran T2000 (Dex). Detection of binding was as described (supplemental Table S2). Binding scores for duplicate spots at 90 and 300 pg per spot are in blue and red respectively, error bars represent half of the difference between the two scores. The unpredicted binding by the anti-KS to the commercial CSA is attributed to the variable presence of KS antigen in different batches as depicted in supplemental Table S1.

Mild acid hydrolysis of the KS gave rise to a more diverse glycan population (with even and odd monosaccharide numbers) than keratanase treatments. For example, F1 and F2 contained 2–4 mers and 4–7 mers, respectively (supplemental Table S6).

NGLs were generated from the three series of KS glycome fractions for incorporating into a primary Beam Search NGL array. These included the ten K'ase I 2–20 mers, the nine

K'ase II 2–18 mers and the two fractions after mild acid hydrolysis (Fig. 2). Also included were NGLs of two synthetic tetrasaccharides, TCI-4(0S) and TCI-4(4S) with no sulfate and four sulfates, respectively (supplemental Table S2).

*The Minimum Glycan Chain Length of the R-10G Antigen Determined with the Primary Beam Search NGL Microarray—* R-10G binding was detected to the NGLs derived from K'ase



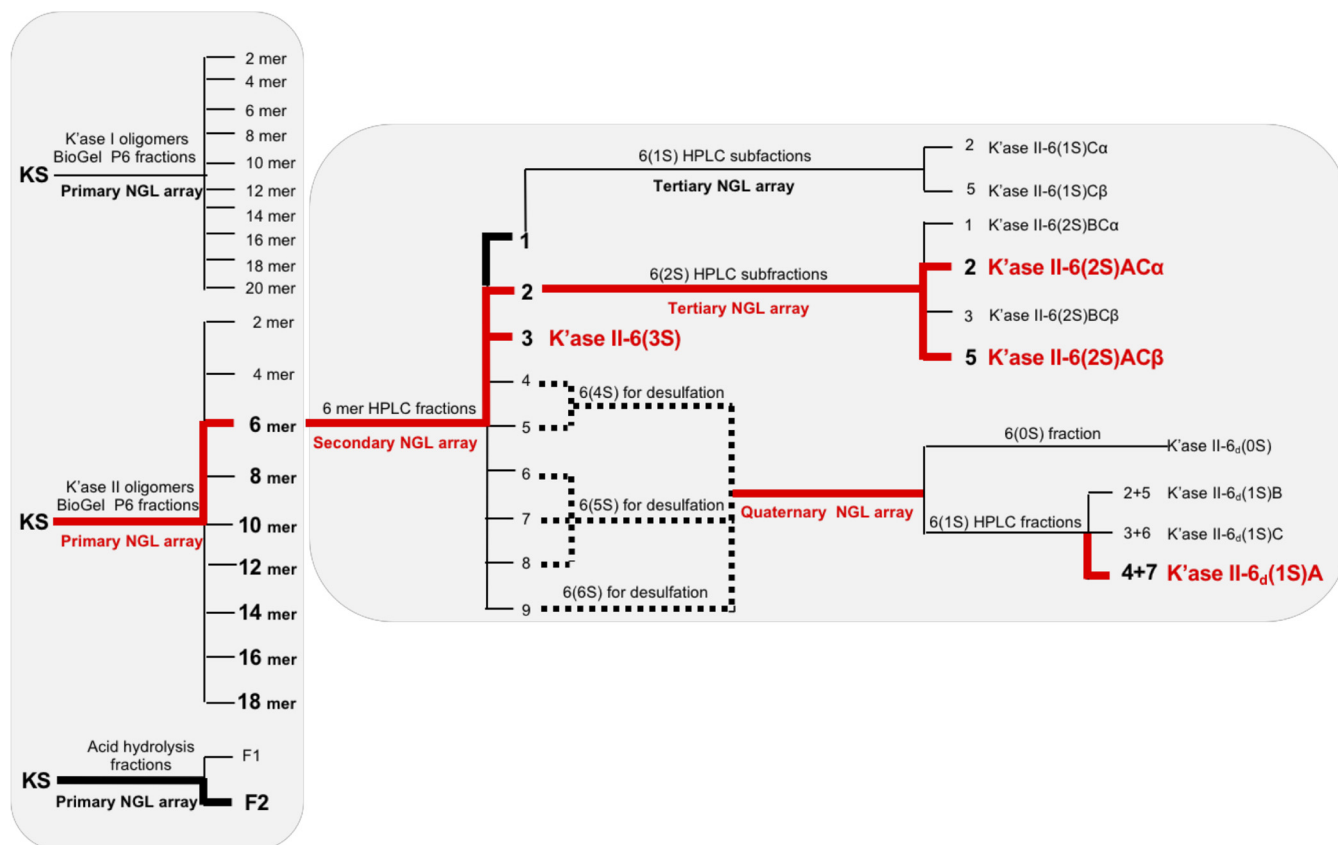


FIG. 2. **The Beam Search approach in the assignment of the R-10G antigen.** The primary NGL array encompassed the glycan fractions after partial digestion of KS with keratanase I, keratanase II or mild acid hydrolysis. Fractions bound by mAb R-10G are in bold font. The route that led to assignment of the R-10G antigen is colored red as are the antigen-active glycans assigned.

II-6 mer and longer glycans rather than to those from K'ase I series (Fig. 3A). This differed from the pattern of binding by mAb MZ15 (Fig. 3A, inset), which bound not only to K'ase II series, 10 mers and longer, but also to the K'ase I series 8 mers and longer. The NGLs of 4–7 mers (F2 after acid hydrolysis) also gave R-10G binding, but there was no binding detected to NGLs of the 4 mer and shorter glycans, F1 (2–4 mers) or the synthetic TCI-4(0S) and TCI-4(4S) (Fig. 3A).

Thus the minimum antigen-positive glycan chain length among the K'ase II-derived NGLs was 6 mer. The K'ase II-6 mer fraction, being less heterogeneous than the F2 after acid hydrolysis was selected for generation of the secondary Beam Search array.

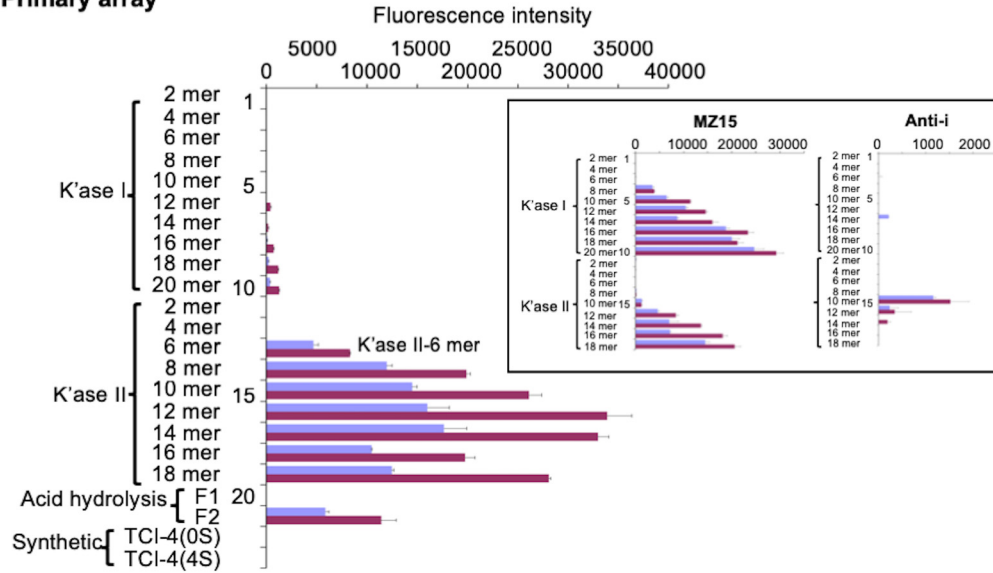
**K'ase II-6 mers with R-10G Antigen Activities Identified and a Tri-sulfated Antigen-positive 6 mer Assigned with the Secondary Beam Search NGL Microarray**—The K'ase II-6 mer fraction was subjected to strong anion-exchange (SAX) HPLC to obtain subfractions with different numbers of sulfates. Nine subfractions were harvested (Fig. 4A). ESI-MS analyses of these (supplemental Table S7) revealed that HPLC fractions 1, 2, and 3 were mono, di and tri-sulfated 6 mers, respectively; HPLC-4 and 5 were tetra-sulfated 6 mers; HPLC-6, 7 and 8 contained penta-sulfated 6 mers; HPLC-9 contained hexa-

sulfated 6 mers; HPLC-7a and -8a were tri- and tetra-sulfated 4 mers and penta-sulfated 5 mers, respectively. The nine K'ase II-6 mer subfractions were converted to NGLs (Fig. 4B) and included in the secondary array, depicted in Fig. 2.

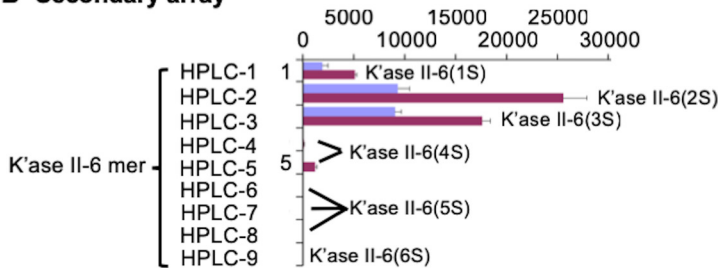
The tri-sulfated K'ase II-6 mer fraction (HPLC-3) was strongly bound by R-10G (Fig. 3B). The  $^1\text{H}$ ,  $^1\text{H}$ - $^1\text{H}$  TOCSY and  $^1\text{H}$ - $^{13}\text{C}$  HSQC 2D spectra of the fraction are shown in Fig. 5A, 5B, and 5C, respectively. The values in Fig. 5D were determined by comparison of HSQC cross-peak chemical shift (Fig. 5B) with those in the literature (34, 45, 46), with confirmation by correlations in the TOCSY spectrum (Fig. 5C). Chemical shift values for reducing end residues of KS oligosaccharides have not been published previously, so the reducing end  $\alpha$ -GlcNAc6S anomer residue was assigned by comparison with 4-linked  $\alpha$ -GlcNAc6S in modified heparins (47, 48) with confirmation by correlations in the TOCSY spectrum (Fig. 5C). The spectrum of reducing end  $\beta$ -GlcNAc6S anomer was, apart from the anomeric center, indistinguishable from the internal  $\beta$ -GlcNAc6S. No signals attributable to nonsulfated GlcNAc (49) were present. The R-10G antigen-positive glycan was thus established as the tri-sulfated hexasaccharide shown in Table III, and designated K'ase II-6(3S).

The di-sulfated K'ase II-6 mer fraction (HPLC-2), K'ase II-6(2S), was also strongly bound (Fig. 3B) but it was a mixture

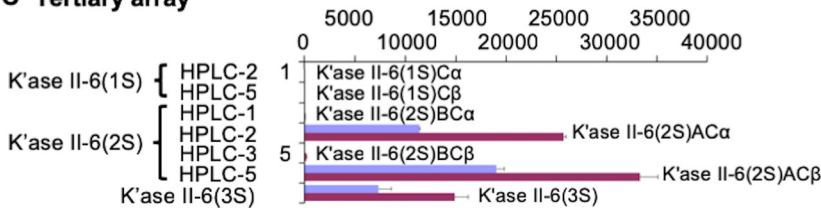
**A Primary array**



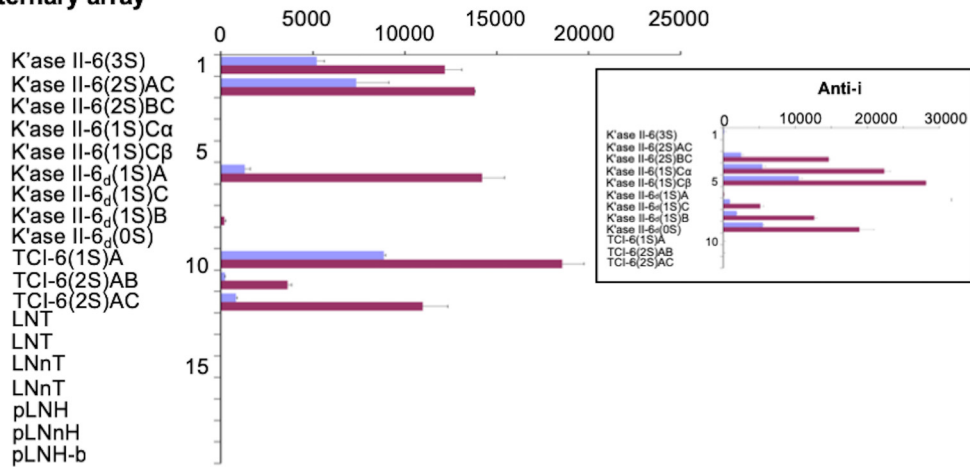
**B Secondary array**



**C Tertiary array**



**D Quaternary array**



(Fig. 4E). There was binding albeit weak to the mono-sulfated K'ase II-6 mer fraction (HPLC-1), K'ase II-6(1S) (Fig. 3B), which was also a mixture (Fig. 4C). There was little or no binding to the more highly sulfated K'ase II-6 mers with 4 sulfates (HPLC-4 and 5), 5 sulfates (HPLC-6, 7, and 8) and 6 sulfates (HPLC-9), which are designated K'ase II-6(4S), K'ase II-6(5S) and K'ase II-6(6S), respectively (Fig. 3B).

Collectively, these data suggest that R-10G antigen is expressed not only on tri-sulfated 6 mer, K'ase II-6(3S), but also on di and mono-sulfated 6 mers, K'ase II-6(2S) and K'ase II-6(1S); but not on 6 mers with more than three sulfates.

*The Positions of Sulfate on R-10G Antigen-positive and -negative Di-sulfated K'ase II-6 mers Assigned with the Tertiary Beam Search Microarray*—The antigen-positive di- and mono-sulfated fractions, K'ase II-6(2S) and K'ase II-6(1S), were subjected to further fractionation by HPLC using a PGC column. Seven subfractions were obtained from the K'ase II-6(2S), of which four (HPLC-1, 2, 3, and 5) were in sufficient amounts for NGL preparation and antigenic analysis (Fig. 4E and 4F). Nine subfractions were obtained from K'ase II-6(1S), of which two (HPLC-2 and 5) were enough for NGL preparation and antigenic analysis (Fig. 4C and 4D). These six NGLs were arrayed as the tertiary array and together with the K'ase II-6(3S) constituted supplements to the secondary array (Fig. 3C).

Subfractions HPLC-2 and 5 of the K'ase II-6(2S) fraction elicited R-10G binding, but HPLC-1 and 3 did not (Fig. 3C). ESI-CID-MS/MS analysis of the two R-10G antigen-positive NGLs from the di-sulfated HPLC-2 and 5 gave almost identical product-ion spectra using  $[M-2H]^{2-}$  as the precursors (Fig. 6A and 6B), indicating a pair of anomeric isomers. Although, because of the presence of two sulfate groups the fragmentation is rather complex, making the assignment very difficult, it is still possible to identify the location of the two sulfates using some key fragment ions. One sulfate at the reducing terminal GlcNAc is apparent with the sulfated  $Z_1$  ion ( $m/z$  1018) and further corroborated by the  $Z_2$  ( $m/z$  1180) (nomenclature of fragmentation see (50)). The latter ion also indicated the absence of the sulfate on the Gal residue. The other sulfate can also be unambiguously assigned at the nonreducing terminal GlcNAc with the unique sulfated  $B_2$  ion at  $m/z$  444. The product-ion spectrum of one of the anomers, HPLC-5, as free sugar (Fig. 7) before NGL conjugation can be used to make the assignment of the active sequence as some additional fragment ions were present, although the spectrum is more complex because of the sulfate-related fragmentation. A sulfate at the nonreducing terminal GlcNAc again can be deduced from the  $B_2$  ( $m/z$  444) and  $C_2$  ( $m/z$  462) ions and

the other sulfate at the reducing terminal GlcNAc by  $B_5$  ( $m/z$  971) and  $C_5$  ( $m/z$  989) ions. The lack of a sulfate at either internal or reducing terminal Gal residues was evidenced by the  $B_3$  and  $C_3$  ( $m/z$  606 and 624), and  $B_5$  and  $C_5$  ( $m/z$  971 and 989) ion pairs. The  $^{2,4}A_6$  ion at  $m/z$  1031 suggested the sulfate at the 6-O-position. HPLC-2 and 5 are thus assigned as the  $\alpha$  and  $\beta$  anomeric forms, respectively, of the same glycan sequence. The sequence of the two antigen-positive glycans was deduced and these were designated K'ase II-6(2S)AC (Fig. 2). Their sequences are shown in Table III. We designate the GlcNAc residues at the nonreducing, middle and reducing ends as A, B and C, respectively.

The spectra of the two antigen-negative NGLs from HPLC-1 and -3 are also very similar (Fig. 6C and 6D) and HPLC-1 and -3 were considered as anomeric isomers as discussed above. These two spectra are different from those of the antigen-positive fractions HPLC-2/5. The presence of  $Z_1$  ( $m/z$  1018) and  $Z_2$  ( $m/z$  1180) ions indicated the location of the sulfate at the reducing terminal GlcNAc. The lack of a sulfate at the nonreducing terminal GlcNAc is evidenced by the absence of the  $B_2$  ion at  $m/z$  444. The second sulfate is therefore deduced to be at the internal GlcNAc. HPLC-1 and 3 are then assigned as  $\alpha$  and  $\beta$  anomers, respectively, of K'ase II-6(2S)BC (Fig. 2) The sequence is shown in Table III.

For the two mono-sulfated antigen-negative subtractions, HPLC-2 and 5 of the K'ase II-6(1S) (Figs. 3C and 4C), the positions of sulfate could not be assigned as NGLs. As the antigenic analyses with anti-i P1A ELL using in the quaternary array (below) indicated that the sulfate is not at position A of the GlcNAc residues, and taking that into account the cleavage specificity of K'ase II for the GlcNAc(6S)-Gal sequence, these two glycans were assigned the  $\alpha$  and  $\beta$  anomeric forms of the same glycan K'ase II-6(1S)C (Fig. 2).

The main conclusion that could be drawn from the tertiary array analysis was that the K'ase II-6(2S)AC with sulfate at the nonreducing and reducing end GlcNAc residues is R-10G antigen-positive whereas the analog K'ase II-6(2S)BC with sulfate at the internal and reducing end GlcNAc residues is antigen-negative. This highlighted the need to have a mono-sulfated 6 mer K'ase II-6(1S)A with sulfate at position A not only to determine whether this is sufficient for strong R-10G antigenicity but to rule out any contribution of sulfate at the internal position (position C) to the strong antigenicity of K'ase II-6(3S) and K'ase II-6(2S)AC.

*A Mono-sulfated 6 mer K'ase II-6(1S)A Assigned as the R-10G Antigen with a Quaternary Beam Search Microarray and Corroborated with Synthetic Glycans*—The mono-sulfated K'ase II-6(1S) fraction was antigen-positive (Fig. 2 and

FIG. 3. Microarray analyses of mAbs R-10G, MZ15 and anti-i P1A ELL using the Beam Search NGL arrays. Results of R-10G binding are presented as histogram charts for primary array, A, secondary array, B, tertiary array, C, and quaternary array, D. These are means of fluorescence intensities of duplicate spots, printed at 2 and 5 fmol per spot (blue and red respectively), with error bars representing half of the difference between the values of the duplicate spots. Results of MZ15 and anti-i P1A ELL antibodies are shown as insets in panels A and D, respectively.

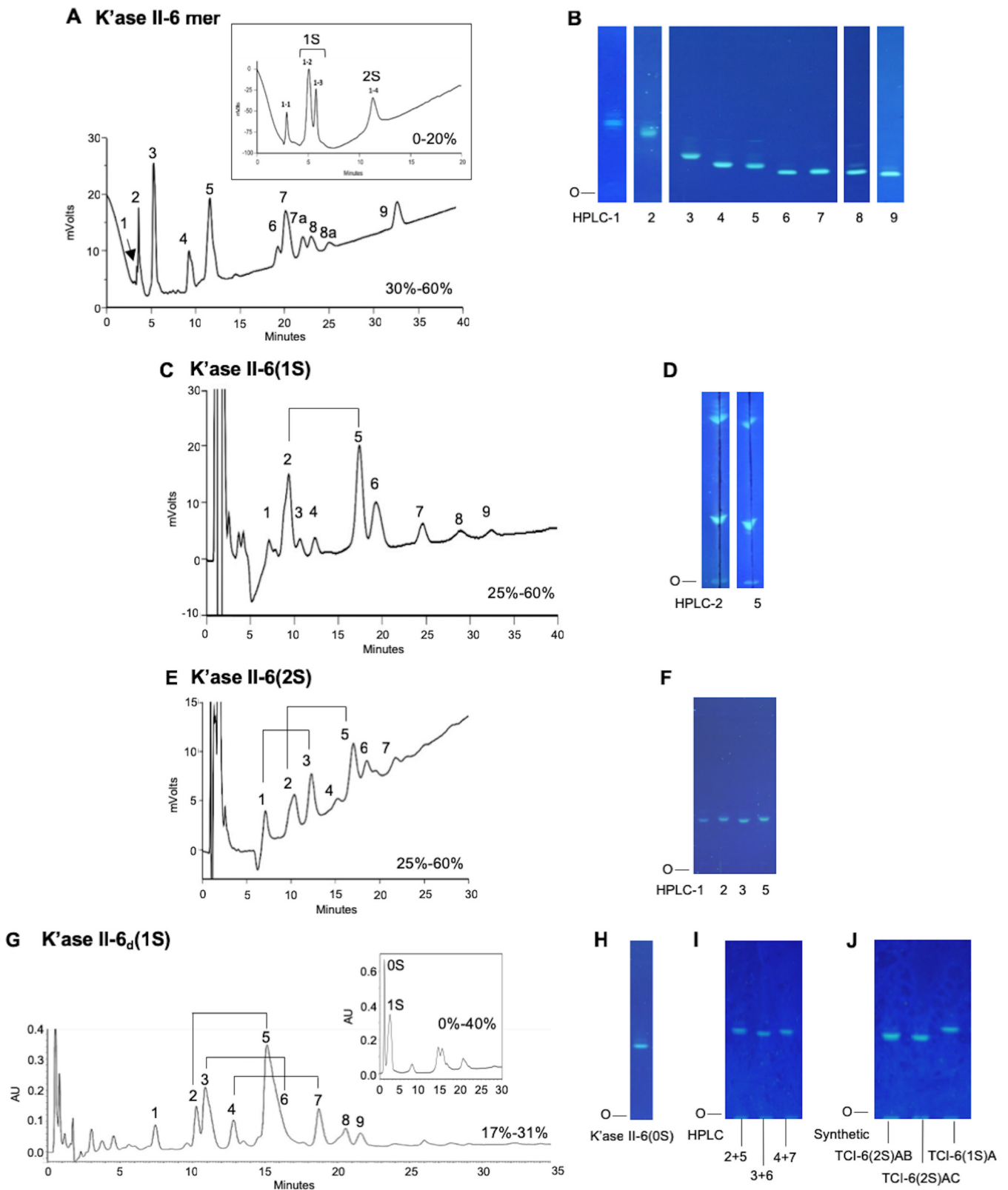


FIG. 4. HPLC of KS 6 mer fractions and HPTLC of the derived NGLs and those of the synthetic 6 mers. *A*, SAX-HPLC of K'ase II-6 mer fraction; the inset is the further fractionation of the subfraction 1. Peaks 1–2 and 1–3 (mono-sulfated) were pooled and designated 1S. *B*, HPTLC of the NGLs of the 9 fractions indicated in *A*. *C*, PGC-HPLC of K'ase II-6(1S) fraction. *D*, HPTLC of the NGL reaction mixtures of the 2 fractions indicated in *C*. *E*, PGC-HPLC of K'ase II-6(2S) fraction; *F*, HPTLC of the NGLs of the 4 fractions indicated in *E*. *G*, PGC-HPLC of K'ase II-6<sub>d</sub>(1S) fraction; the inset is the further fractionation of the subfraction 1. Peaks 1–2 and 1–3 (mono-sulfated) were pooled and designated 1S. *H*, HPTLC of the NGL reaction mixtures of the 2 fractions indicated in *C*. *I*, HPTLC of the NGL reaction mixtures of the 2 fractions indicated in *C*. *J*, HPTLC of the NGLs of the 4 fractions indicated in *E*.

3B). However, this fraction was heterogeneous and obtained in low yield. Only two of the subfractions (they are considered  $\alpha$  and  $\beta$  anomers) were in sufficient amounts for antigenic analysis; both were R-10G antigen-negative (Figs. 2 and 3C). To have sufficient amounts of mono-sulfated 6 mers, we pooled all of the over-sulfated (tetra-, penta- and hexa-sulfated) 6 mer glycan fractions and carried out partial desulfation of these, as depicted in Fig. 2.

The partially de-sulfated fractions were separated into a nonsulfated [designated as K'ase II-6<sub>d</sub>(0S)] and a mono-sulfated fraction with a SAX column (Fig. 4G inset); the subscript d indicates a de-sulfated product. The mono-sulfated fraction was then subfractionated with a PGC column into nine subfractions (Fig. 4G). The anomeric pairs HPLC-2/5, 3/6 and 4/7 (Fig. 4G) of the three mono-sulfated sequence isomers were assigned by ESI-CID-MS/MS and by HPLC. The sequence of the R-10G antigen positive mono-sulfated KS 6 mer HPLC fraction (HPLC-4/7), before NGL conjugation, was assigned by ESI-CID-MS/MS using [M-H]<sup>-</sup> as the precursor (Fig. 8A). The location of the sulfate at the nonreducing terminal side GlcNAc is indicated by the glycosidic fragment B<sub>2</sub> (*m/z* 444), C<sub>2</sub> (*m/z* 462), and the double-cleavage D<sub>Y5-B2</sub> ion (*m/z* 282). The exclusive nonreducing terminal fragments B<sub>3</sub>/C<sub>3</sub>, B<sub>4</sub>/C<sub>4</sub> and B<sub>5</sub>/C<sub>5</sub> and the absence of B<sub>1</sub>/C<sub>1</sub> also suggested the sulfate at the sub-terminal (nonreducing end) GlcNAc. The product-ion spectra of the antigen-negative fractions HPLC-2/5 and HPLC-3/6 are very different from those of HPLC-4/7 (Fig. 8B and 8C). A product-ion spectrum of a synthetic mono-sulfated 6 mer TCI-6(1S)A with the sulfate at the nonreducing sub-terminal position was also acquired and was identical to that of HPLC-4/7.

The sequence of HPLC-4/7 was thus assigned to the desired analog with sulfate at positions A, designated as K'ase II-6<sub>d</sub>(1S)A (Sequence shown in Table III). The sequence of HPLC-2/5 was assigned as below and is designated as K'ase II-6<sub>d</sub>(1S)B (Table III). The sequence of HPLC-3/6 was assigned as below and designated as K'ase II-6<sub>d</sub>(1S)C (Table III). The nonsulfated 6 mer K'ase II-6<sub>d</sub>(0S) and the purified three mono-sulfated 6 mers, K'ase II-6<sub>d</sub>(1S)A, K'ase II-6<sub>d</sub>(1S)B and K'ase II-6<sub>d</sub>(1S)C were converted to NGLs (Figs. 4H and 4I, respectively). These NGLs were included in the quaternary array (Fig. 2), together with NGLs of K'ase II-6(3S), K'ase II-6(2S)AC, K'ase II-6(2S)BC and K'ase II-6(1S)C. Also included as reference compounds were NGLs of LNT, LNnT, pLNH, pLNnH and pLNH-b (supplemental Table S2 and Fig. 3D). To verify the binding data with the KS-derived hexasaccharides, three of the sequences considered important in assignment of the R-10G antigen were custom synthesized chemically at TCI.

These were the mono-sulfated and two di-sulfated 6 mers, designated as TCI-6(1S)A, TCI-6(2S)AB and TCI-6(2S)AC, respectively (supplemental Table S2). These were converted to NGLs (Fig. 4J) for inclusion in the quaternary array (Fig. 3D).

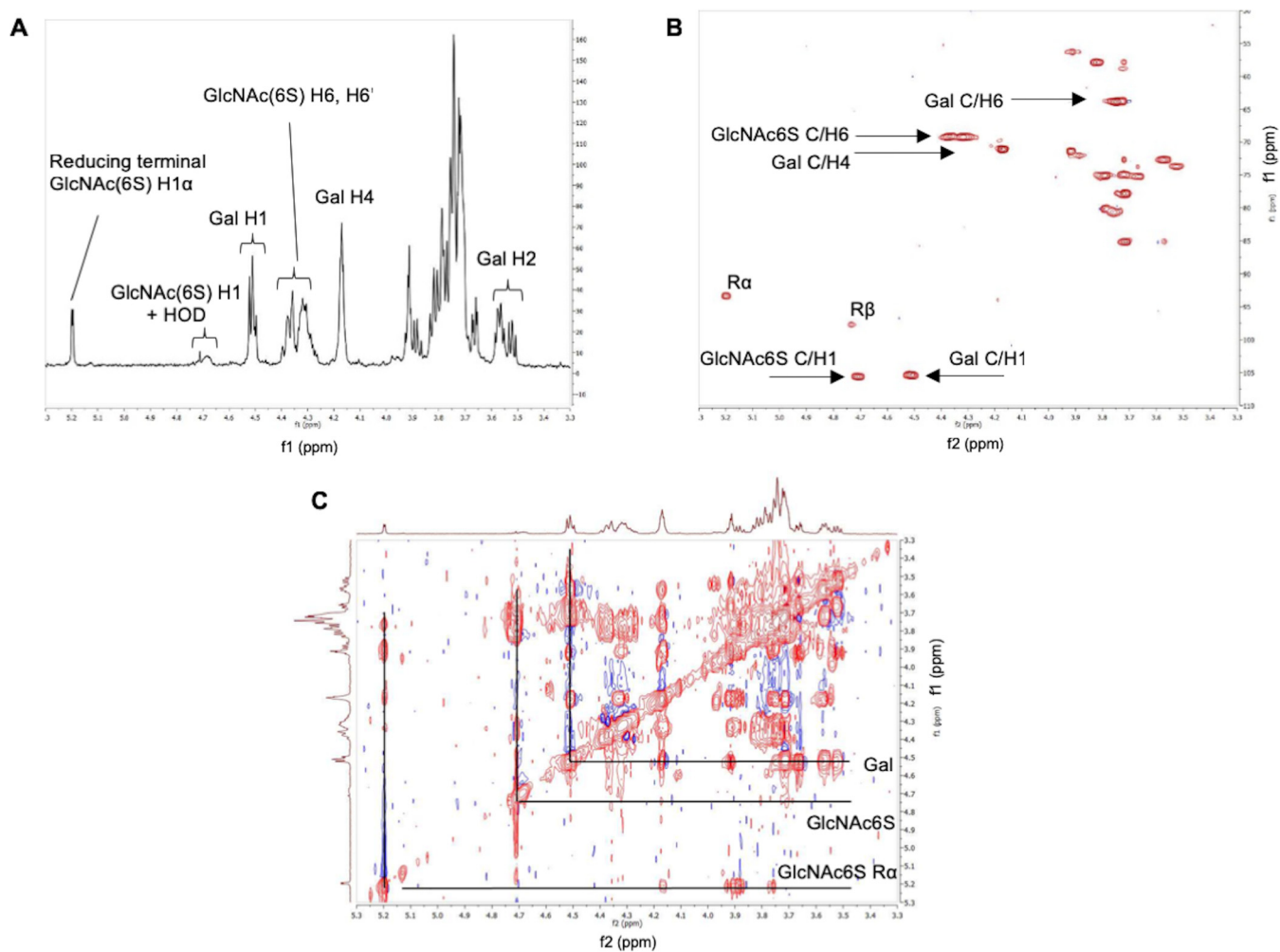
Of the three mono-sulfated 6 mers derived from the desulfated fraction, K'ase II-6<sub>d</sub>(1S)A elicited strong R-10G binding as did TCI-6(1S)A and TCI-6(2S)AC (Figs. 3D and Table III). The binding signal was similar to those of the tri- and di-sulfated, K'ase II-6(3S) and K'ase II-6(2S)AC, respectively. This established that the 6 mer with a sulfate at the nonreducing terminal position (position A) is sufficient for and constitutes the R-10G antigen.

*A Revisit of mAbs FC10.2, TRA-1-60, and TRA-1-81 and the Anti-i Using the Quaternary NGL Microarray to Compare Binding Profiles with that of mAb R-10G*—Advantage was taken of the Beam Search arrays in the context of reference glycans to analyze the three anti-stem cell antibodies FC10.2, TRA-1-60, TRA-1-81 (Table III) also the anti-i (Table III, and Fig. 3), compared with R-10G (Table III). There was no binding of R-10G to the six nonsulfated NGL reference compounds (Fig. 3D). The distinct binding specificities of FC10.2, TRA-1-60 and TRA-1-81 toward the T1-T2-T2 sequence were corroborated. A new observation made possible with the relatively recently generated NGL of the pLNH-b (T1-T2-T1 sequence), isomer of pLNH (51), is the lack of binding of FC10.2, TRA-1-60 and TRA-1-81 to this hexasaccharide sequence with an inner T1 chain (Table III).

An interesting reciprocal pattern of binding was observed for the anti-i with the present series of under-sulfated KS glycans. It is known that in fully sulfated KS (*i.e.* on all Gal and GlcNAc residues), the *i*-antigen is masked (36). Here we found a lack of binding of anti-i P1A ELL to all of the R-10G antigen-positive 6 mers with sulfate on GlcNAc at position A whereas there was strong binding to those with sulfate at positions B and C (Fig. 3D inset and Table III). Thus, anti-i P1A ELL can accommodate sulfate substituents on GlcNAc at positions B and C (Table III). The two minor subfractions (HPLC-2 and -5) of K'ase II-6(1S), which were R-10G antigen-negative but could not be sequenced by MS, were strongly bound by the anti-i (Table III). We used this immunochemical information as evidence that sulfate is not at position A, and assigned the sequence as K'ase II-6(1S)C.

*Binding of K'ase II-6 mers with 4, 5, or 6 Sulfate Groups by FGF2 and BMP-2 Detected in Exploratory Microarray Analyses*—Prompted by the observations in ref (52). on the binding of human fibroblast growth factor FGF2 and the morphogen Sonic hedgehog (Shh) to KS polysaccharide, we performed exploratory microarray analyses of FGF1 and FGF2, Shh,

fraction; the inset is SAX-HPLC of products of partial desulfation of K'ase II-6(4S+5S+6S) fractions. H, HPTLC of the NGL of K'ase II-6(0S). I, HPTLC of the NGLs of the three pooled fractions indicated in G. J, HPTLC of the NGLs of the three synthetic 6 mers. The percentages in the panels indicate the solvent gradients used for HPLC. NGLs were revealed under UV light with long wavelength after staining with primulin. HPTLC chromatography was in solvent CHCl<sub>3</sub>:MeOH:H<sub>2</sub>O 60/35/8 (v/v/v), and the spaces between figures indicate that they were on separate chromatograms.



**D The chemical shifts of the R-10G antigen-positive glycan K'ase II-6(3S)**

Terminal Gal		Internal Gal		Internal GlcNAc(6S)		Reducing terminal GlcNAc6S		
						$\alpha$	$\beta$	
C(1)	105.4	C(1)	105.4	C(1)	105.6	C(1)	93.4	97.7
C(2)	73.8	C(2)	72.7	C(2)	57.9	C(2)	56.3	58*
C(3)	75.3	C(3)	85.2	C(3)	74.9	C(3)	72.5	75*
C(4)	71.4	C(4)	71.1	C(4)	80.7	C(4)	80.3	81*
C(5)	ND	C(5)	77.9	C(5)	75.2	C(5)	71.2	75*
C(6)	ND	C(6)	63.9	C(6)	69.2	C(6)	69.3*	69.3*
				C(NAc)	24.9	C(NAc)	24.9	24.9
H(1)	4.51	H(1)	4.52	H(1)	4.71	H(1)	5.2	4.73
H(2)	3.52	H(2)	3.57	H(2)	3.82	H(2)	3.91	3.69-3.81*
H(3)	3.67	H(3)	3.71	H(3)	3.72	H(3)	3.88	3.69-3.81*
H(4)	3.91	H(4)	4.17	H(4)	3.76	H(4)	3.79	3.69-3.81*
H(5)	ND†	H(5)	3.71	H(5)	3.79	H(5)	3.79	3.69-3.81*
H(6)	ND	H(6)	3.72-3.74	H(6, 6')	4.37, 4.32	H(6)	4.37, 4.32*	4.37, 4.32*
				H(NAc)	2.03	H(NAc)	2.03*	2.03*

\*Indistinguishable from internal GlcNAc(6S) values

†ND, not determined.

FIG. 5. NMR data on the R-10G antigen-positive K'ase II-6(3S). A,  $^1\text{H}$  spectrum at 700 MHz; B, HSQC spectrum; C, TOCSY spectrum; D, chemical shifts.

## Assignment of Glycan Markers of Human Stem Cells

TABLE III

Relative intensities of binding of mAbs R-10G, anti-*i* P1A ELL, FC10.2, TRA-1-60, TRA-1-81 and anti-LNT to the NGL probes in the quaternary NGL array

● Gal, galactose; ■ GlcNAc, *N*-acetylglucosamine; ● sulfate(6S), sulfate at position 6 of GlcNAc. Binding scores are means of the fluorescence intensities at 5 fmol/probe spot; - indicates the score is less than 1. Relative binding intensity: ■ < 10%, ■ 10–30%, ■ 30–70%, ■ 70–100%, where 100% is the maximum score for respective antibodies.

Pos*	Probe	Structure	R-10G	Anti- <i>i</i> P1A ELL1	FC10.2	TRA-1-60	TRA-1-81	Anti-LNT
1	K'ase II-6(3S)	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S 6S 6S	12,187	-	-	-	-	-
2	K'ase II-6(2S)AC	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S 6S	13,809	-	-	-	-	-
3	K'ase II-6(2S)BC	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S 6S	-	14,659	-	-	-	-
4	K'ase II-6(1S)C	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S	-	22,345	-	-	-	-
5	K'ase II-6(1S)C	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S	-	28,137	-	-	-	-
6	K'ase II-6 <sub>d</sub> (1S)A	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S	14,215	114	-	-	-	-
7	K'ase II-6 <sub>d</sub> (1S)C	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S	-	5,128	-	-	-	-
8	K'ase II-6 <sub>d</sub> (1S)B	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S	-	12,638	-	-	-	-
9	K'ase II-6 <sub>d</sub> (0S)	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO	-	18,885	-	-	-	-
10	TCI-6(1S)A	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S	18,590	16	-	-	-	-
11	TCI-6(2S)AB	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S 6S	3,641	57	-	-	-	1
12	TCI-6(2S)AC	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4GlcNAc-AO 6S 6S	10,972	-	-	-	-	4
13	LNT-DH	Galβ-3GlcNAcβ-3Galβ-4Glc-DH	-	-	-	-	-	-
14	LNT	Galβ-3GlcNAcβ-3Galβ-4Glc-AO	-	-	463	271	-	-
15	LNnT-DH	Galβ-4GlcNAcβ-3Galβ-4Glc-DH	-	-	-	-	-	-
16	LNnT	Galβ-4GlcNAcβ-3Galβ-4Glc-AO	-	-	-	-	-	-
17	pLNH	Galβ-3GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc-AO	-	-	15,711	27,949	13,014	5,896
18	pLNnH	Galβ-4GlcNAcβ-3Galβ-4GlcNAcβ-3Galβ-4Glc-AO	-	24,180	-	-	-	-
19	pLNH-b	Galβ-3GlcNAcβ-3Galβ-3GlcNAcβ-3Galβ-4Glc-AO	-	-	-	-	-	-

\* Pos, position in the microarray.

int/wingless (Wnt)-3a, bone morphogenetic proteins, BMP-2 and BMP-4, using the secondary NGL array and the results are shown in Table IV. We detected binding of FGF2 and BMP-2 to KS glycan populations. Among the K'ase II-6 mers there was detectable binding to those with 4, 5, or 6 sulfate groups designated HPLC-4, 5, 6, 7, 8, and 9, but not to those with less than 4 sulfate groups, HPLC-1, 2, and 3, which were bound by R-10G (supplemental Table S7).

## DISCUSSION

Knowledge of the glycan sequences and substituents recognized by anti-stem cell antibodies has an important bearing on their use as reagents to gain insights into the glycans displayed, and the changes that occur on differentiation. As more sequence-defined glycan probes become available the antibody specificities can be reassessed. Our studies here contribute to the refinement of the assignments of the binding

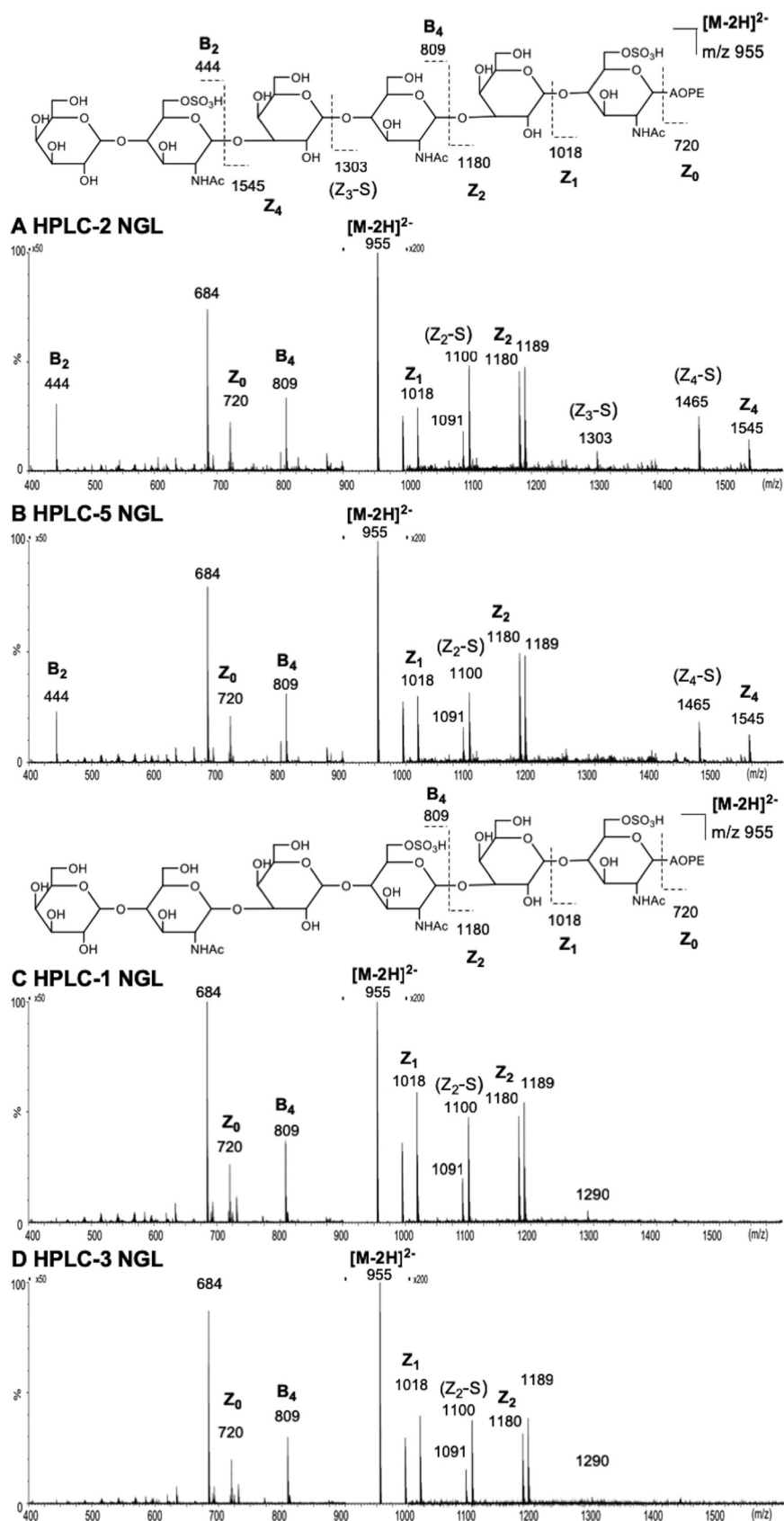
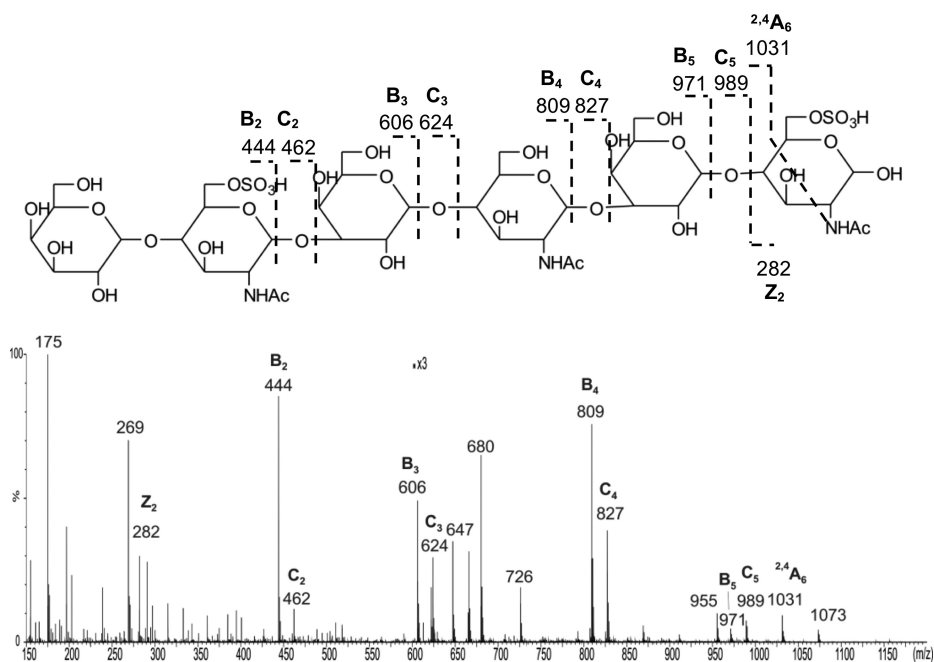


FIG. 6. ESI-CID-MS/MS product-ion spectra of the NGLs of glycans derived from fraction K'ase II-6(2S). A, B, C, and D are NGLs of HPLC-2, HPLC-5, HPLC-1 and HPLC-3, respectively.



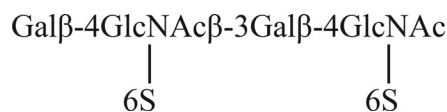
FIG. 7. ESI-CID-MS/MS product-ion spectrum of the underivatized R-10G antigen-positive glycan from HPLC-5.



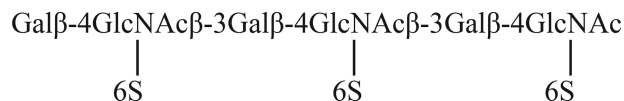
specificities of five mAbs that recognize human stem cells. They provide new insights into the antigen recognized by the more recently described mAb R-10G.

*The mAbs FC10.2, TRA-1-60, and TRA-1-81 Have A Determinant In Common but FC10.2 and TRA-1-60 Differ In Their Ability To Accommodate A Fucose Substituent*—We have now closely compared for the first time the three well-established anti-stem cell antibodies, FC10.2, TRA-1-60, and TRA-1-81, and shown that they are very similar in their strong binding to the nonsubstituted T1-T2 sequences as part of linear or branched glycan backbones that are longer than tetrasaccharide. No binding was detected to any of the sulfated glycans arrayed. The tolerance of a fucose 3-linked to the inner GlcNAc by FC10.2, TRA-1-60 but not by TRA-1-81 would be predicted to give rise to differences in immunostaining of certain tissues that contain this fucosylated sequence.

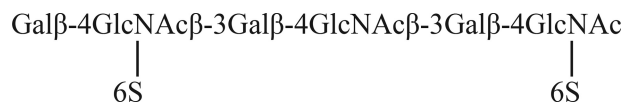
*The R-10G Antigen Is A Unique Mono-sulfated Glycan Antigen*—Under the analysis conditions used, no overlap was observed in the binding specificities of mAbs FC10.2, TRA-1-60, TRA-1-81 and anti-i P1A ELL with that of mAb R-10G. In the initial investigations of the R-10G determinant expressed on podocalyxin, the antigen was suggested to be on O-glycans without sialic acid and fucose (19). Evidence of sulfate involvement in the R-10G antigen was obtained: the main product after keratanase II digestion was assigned as the disaccharide Gal-GlcNAc(S) (19); the presence of additional glycans such as Gal-GlcNAc(S)-Gal-GlcNAc(S) was suggested as minor components. In a follow up study (21) the chemically synthesized di-sulfated tetrasaccharide below, in biotinylated form was shown to be bound by mAb R-10G.



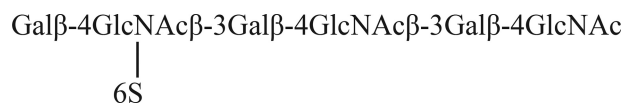
In the present study, our finding of R-10G binding to the tri-sulfated hexasaccharide below isolated from KS via the Beam Search approach is in complete accord with the earlier studies (21).



A second antigen-positive glycan, di-sulfated at positions A and C (as below) was also isolated



R-10G activity was additionally detected in the heterogeneous unfractionated mono-sulfated fraction, K'ase II-6(1S), but the glycan could not be isolated in sufficient amount for antigenic and structural analysis. Considering the cleavage specificity of keratanase II, it is likely that the antigen-active component was a trace contaminant from the adjoining K'ase II-6(2S) fraction. This prompted the partial desulfation strategy to create a sufficient amount of the mono-sulfated antigen below for sequence determination.



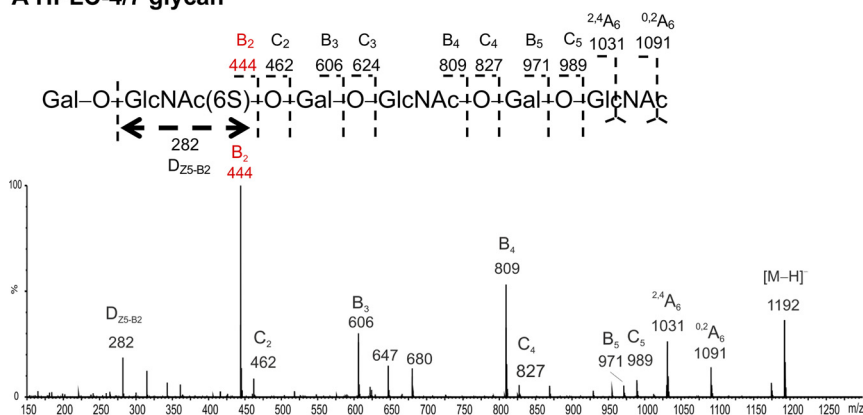
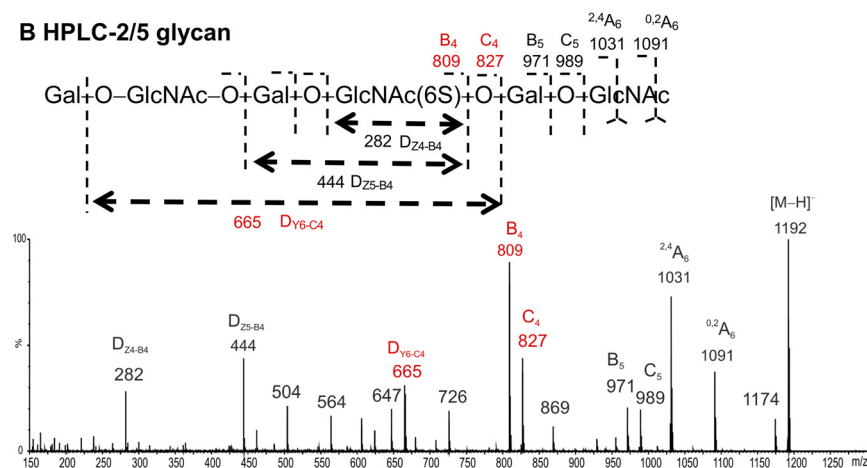
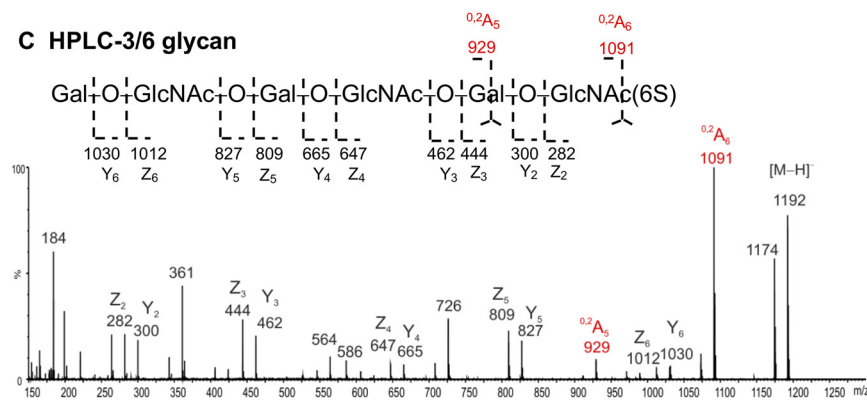
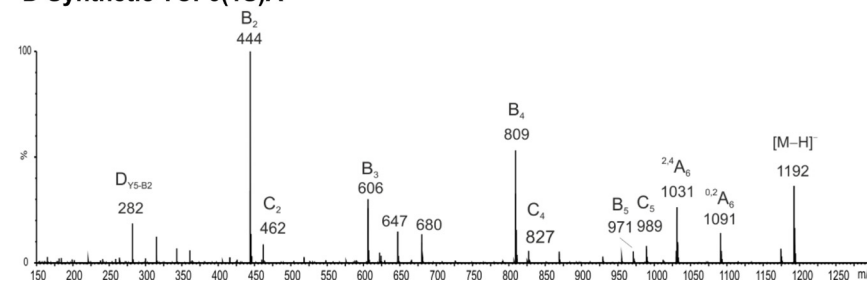
**A HPLC-4/7 glycan**

**B HPLC-2/5 glycan**

**C HPLC-3/6 glycan**

**D Synthetic TCI-6(1S)A**


FIG. 8. ESI-CID-MS/MS product-ion spectra of the glycans K'ase II-6<sub>d</sub>(1S)A, B and C (Panels A, B, and C, respectively), isolated from the de-sulfated hexamers substituted with 4 to 5 sulfates, compared with the spectrum of the synthetic TCI-6(1S)A (Panel D).

TABLE IV

Relative intensities of binding of mAb R-10G, growth factor FGF2 and morphogen BMP-2 in secondary NGL microarray

Binding scores are means of the fluorescence intensities at 5 fmol/probe spot; - indicates the score is less than 1. Relative binding intensity:  < 10%,  10–30%,  30–70%,  70–100%, where 100% is the maximum score for respective antibodies. Growth factor FGF1 and morphogens Wnt-3a, Shh and BMP-4 were analyzed in the microarray but no binding was detected.

Probe	Structure	R-10G	FGF2	BMP-2
K'ase I	2 mer	-	7563	3376
	4 mer	-	10264	4095
	6 mer	-	13523	5436
	8 mer	-	9004	2297
	10 mer	-	17762	4510
	12 mer	-	20748	5690
	14 mer	-	21930	6324
	16 mer	43	25869	4587
	18 mer	129	22856	2857
	20 mer	219	34727	4522
K'ase II	2 mer	-	11954	6128
	4 mer	-	11627	4653
	6 mer	3328	10193	4064
	8 mer	12,638	8218	3788
	10 mer	20342	8788	1585
	12 mer	21150	6866	197
	14 mer	21960	6554	306
	16 mer	17373	4782	-
	18 mer	17287	7155	398
K'ase II-6 mer	HPLC-1	2560	-	-
	HPLC-2	13166	-	-
	HPLC-3	8857	-	-
	HPLC-4	70	461	467
	HPLC-5	604	3222	674
	HPLC-6	-	4813	1837
	HPLC-7	-	11208	3160
	HPLC-8	-	13954	4834
	HPLC-9	-	19952	7346

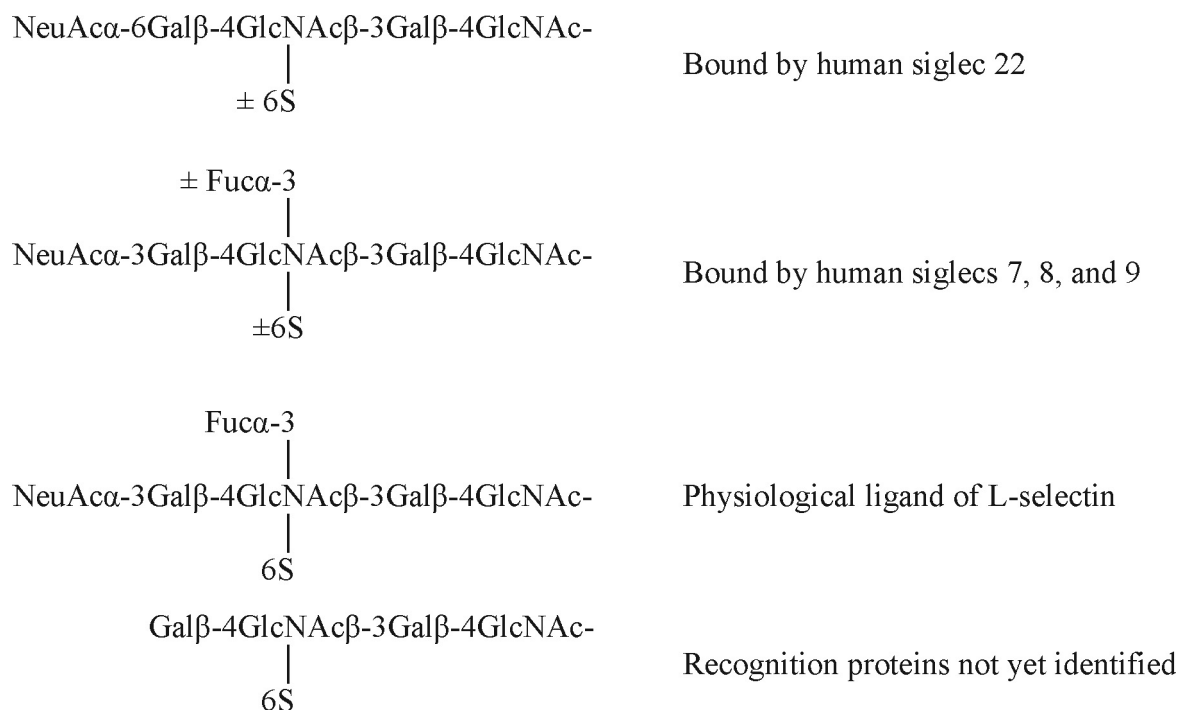
We thereby demonstrated that sulfation at either the B or C position is not required for antigenicity. We propose that the A analog is a low abundance capping sequence on the model saccharide we have used (corneal KS) (53), and that it is among A, AB, AC and ABC hexasaccharide forms on the podocalyxin on stem cells. The relative amounts in ES and iPS cells may well differ. These need investigation in future studies.

We conclude, based on the earlier (19) and current studies, that the chain length of the R-10G antigen is between a 4 mer and 6 mer. When presented as a NGL the minimum length bound is a 6 mer, although in a biotinylated form [associated with a larger spacer length (54)] binding is detectable with a 4 mer. Moreover, the lack of R-10G binding to the series of GlcNAc-terminating glycans (irrespective of chain length) generated by keratanase I treatment of the KS lead us to conclude that the R-10G epitope is necessarily at the nonreducing end rather than a part of an internal sequence. The effect of sulfation of the Gal nonreducing on R-10G binding requires investigation.

*Influence of Sulfation on i Antigen Expression and the Reciprocal Relationship with R-10G*—In the light of the reported expression of i antigen on human mesenchymal stem cells (55), it was opportune to investigate the effect of sulfation on its expression. Earlier studies from our group (36) showed that i antigen expression on the unbranched T2-T2 sequence is masked on KS chains that are fully 6-sulfated on Gal and GlcNAc residues. As depicted overleaf (Scheme 1) we also showed that the antigen is revealed after total desulfation and that different anti-i differ in their ability to bind to the desulfated KS glycans that have GlcNAc at their nonreducing ends. We have now observed that on the KS-derived hexasaccharide with a single 6-linked sulfate on the GlcNAc at position A recognized by mAb R-10G expression of i antigen as seen by anti-i P1A ELL is completely masked, but not when 6-linked sulfate is on GlcNAc at positions B or C or on both (Scheme 1). Thus, there is a reciprocal expression of anti-i P1A ELL and R-10G.

Whether 6-linked sulfate on the Gal at the nonreducing end influences the i antigen expression will be interesting to determine in the light of the strong binding of anti-i P1A ELL in the presence of 6-linked sialic acid at this position [this has been observed recently with other anti-i (Zhen Li and Chao Gao, unpublished)]. We have not investigated here the i antigen expression on glycans with 6-sulfation at this position. But we observe weak binding signals with NGLs of the K'ase II 10–14 mer fractions where the nonreducing terminal Gals are variably sulfated (34), whereas there are strong signals detected with several of the NGLs in the quaternary array where there is no sulfation at this position. Collectively our data suggest that 6-linked sulfate at the nonreducing terminal Gal masks the i determinant recognized by the anti-i P1A ELL. There may be analogies here with proteins of the innate immune system known as siglecs that have a requirement for





SCHEME 2

the carrier glycoproteins or form parts of complexes that regulate the dormant or regenerative preparties and fate of the stem cells. In addition, detailed studies are required of the regulation of the activities not only of glycosyltransferases but also of sulfotransferases (62) during stem cell differentiation. There may be analogies with changes in sulfation of HS during stem cell differentiation (63, 64). Moreover, bespoke binding as well as kinetical analysis systems require design to gain insights into recognition of sequentially changing glycan displays on stem cells by growth factors and morphogens.

**Acknowledgments**—The sequence-defined glycan microarrays contain many saccharides provided by collaborators whom we thank, and also members of the Glycosciences Laboratory for their collaboration in the establishment of the neoglycolipid-based microarray system. We thank Drs Yuji. Matsuzaki and Takashi Ota at TCI for graciously providing synthetic glycans. We thank Dr Fiona M Watt and Dr Bruce Caterson for generously providing antibodies. We acknowledge the excellent facilities provided by Dr Thomas Frenkiel and his team supported by the Francis Crick Institute through provision of access to the MRC Biomedical NMR Centre. The Francis Crick Institute receives its core funding from Cancer Research UK (FC001029), the UK Medical Research Council (FC001029), and the Wellcome Trust (FC001029).

DATA AVAILABILITY

Mass spectral data have been deposited to JPOST (<https://repository.jpostdb.org/>) with the dataset identifier (JPOST ID) 'JPST000605'.

\* This work was supported by Wellcome Trust (WT 108430/Z/15/Z). NW was supported in part by the China Scholarship Council.

 This article contains supplemental material.

\*\* To whom correspondence may be addressed. E-mail: w.chai@imperial.ac.uk.

\*\*\* To whom correspondence may be addressed. E-mail: t.feizi@imperial.ac.uk.

‡‡ Current address: Department of Histology and Embryology, Joint Laboratory for Embryonic Development & Prenatal Medicine, Key Laboratory for Regenerative Medicine of the Ministry of Education, School of Medicine, Jinan University, Guangzhou 510632, China.

Author contributions: N.W., L.M.S., Y.Z., C.G., and B.M. performed research; N.W., L.M.S., and W.C. analyzed data; N.W., Y.L., T.K., W.C., and T.F. wrote the paper; Y.L., T.K., W.C., and T.F. designed research; F.Z., L.F., Y.P., and R.L. contributed new reagents/analytic tools; C.G. made valuable comments.

REFERENCES

1. Wright, A., and Andrews, P. (2009) Surface marker antigens in the characterization of human embryonic stem cells. *Stem Cell Res.* **3**, 3–11
2. Muramatsu, T. (2016) Embryoglycan: a highly branched poly-N-acetylglucosamine in pluripotent stem cells and early embryonic cells. *Glycoconjugate J.* 1–12
3. Feizi, T. (1985) Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens. *Nature* **314**, 53–57
4. Kapadia, A., Feizi, T., and Evans, M. (1981) Changes in the expression and polarization of blood group I and i antigens in post-implantation embryos and teratocarcinomas of mouse associated with cell differentiation. *Exp. Cell Res.* **131**, 185–195
5. Gooi, H., Feizi, T., Kapadia, A., Knowles, B., Solter, D., and Evans, M. (1981) Stage-specific embryonic antigen involves  $\alpha$ 1–3 fucosylated type 2 blood group chains. *Nature*. **292**, 156–158
6. Feizi, T. (1981) The blood group ii system: a carbohydrate antigen system defined by naturally monoclonal or oligoclonal autoantibodies of man. *Immunol. Commun.* **10**, 127–156
7. Childs, R., Pennington, J., Uemura, K., Scudder, P., Goodfellow, P., Evans, M., and Feizi, T. (1983) High-molecular-weight glycoproteins are the

- major carriers of the carbohydrate differentiation antigens I, i and SSEA-1 of mouse teratocarcinoma cells. *Biochem. J.* **215**, 491–503
8. Ozawa, M., Muramatsu, T., and Solter, D. (1985) SSEA-1, a stage-specific embryonic antigen of the mouse, is carried by the glycoprotein-bound large carbohydrate in embryonal carcinoma cells. *Cell Differentiation* **16**, 169–173
  9. Williams, L., Sullivan, A., McIlhinney, R., and Neville, A. (1982) A monoclonal antibody marker of human primitive endoderm. *Int. J. Cancer* **30**, 731–738
  10. Gooi, H., Williams, L., Uemura, K., Hounsell, E., McIlhinney, R., and Feizi, T. (1983) A marker of human foetal endoderm defined by a monoclonal antibody involves type 1 blood group chains. *Mol. Immunol.* **20**, 607–613
  11. Schopperle, W., and DeWolf, W. (2007) The TRA-1-60 and TRA-1-81 Human pluripotent stem cell markers are expressed on podocalyxin in embryonal carcinoma. *Stem Cells* **25**, 723–730
  12. Andrews, P., Banting, G., Damjanov, I., Arnaud, D., and Avner, P. (1984) Three monoclonal antibodies defining distinct differentiation antigens associated with different high molecular weight polypeptides on the surface of human embryonal carcinoma cells. *Hybridoma* **3**, 347–361
  13. Qiu, D., Xiang, J., Li, Z., Krishnamoorthy, A., Chen, L., and Wang, R. (2008) Profiling TRA-1-81 antigen distribution on a human embryonic stem cell. *Biochem. Biophys. Res. Commun.* **369**, 735–740
  14. Kannagi, R., Cochran, N., Ishigami, F., Hakomori, S., Andrews, P., Knowles, B., and Solter, D. (1983) Stage-specific embryonic antigens (SSEA-3 and -4) are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells. *EMBO J.* **2**, 2355
  15. Barone, A., Saljo, K., Benktander, J., Blomqvist, M., Mansson, J., Johansson, B., Molne, J., Aspegren, A., Bjorquist, P., Breimer, M., and Teneberg, S. (2014) Sialyl-lactotetra, a novel cell surface marker of undifferentiated human pluripotent stem cells. *J. Biol. Chem.* **289**, 18846–18859
  16. Liang, Y., Kuo, H., Lin, C., Chen, Y., Yang, B., Cheng, Y., Yu, A., Khoo, K., and Yu, J. (2010) Switching of the core structures of glycosphingolipids from globo- and lacto- to ganglio-series upon human embryonic stem cell differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22564–22569
  17. Säljö, K., Barone, A., Vizlihodzic, D., Johansson, B., Breimer, M., Funa, K., and Teneberg, S. (2017) Comparison of the glycosphingolipids of human-induced pluripotent stem cells and human embryonic stem cells. *Glycobiology* **27**, 291–305
  18. Hirvonen, T., Suila, H., Kotovuori, A., Ritamo, I., Heiskanen, A., Sistonen, P., Anderson, H., Satomaa, T., Saarinen, J., and Tiitinen, S. (2011) The i blood group antigen as a marker for umbilical cord blood-derived mesenchymal stem cells. *Stem Cells Dev.* **21**, 455–464
  19. Kawabe, K., Tateyama, D., Toyoda, H., Kawasaki, N., Hashii, N., Nakao, H., Matsumoto, S., Nonaka, M., Matsumura, H., and Hirose, Y. (2013) A novel antibody for human induced pluripotent stem cells and embryonic stem cells recognizes a type of keratan sulfate lacking oversulfated structures. *Glycobiology* **23**, 322–336
  20. Funderburgh, J. (2000) MINI REVIEW Keratan sulfate: structure, biosynthesis, and function. *Glycobiology* **10**, 951–958
  21. Nakao, H., Nagai, Y., Kojima, A., Toyoda, H., Kawasaki, N., and Kawasaki, T. (2017) Binding specificity of R-10G and TRA-1-60/81, and substrate specificity of keratanase II studied with chemically synthesized oligosaccharides. *Glycoconjugate J.* **34**, 789–795
  22. Badcock, G., Pigott, C., Goepel, J., and Andrews, P. (1999) The human embryonal carcinoma marker antigen TRA-1-60 is a sialylated keratan sulfate proteoglycan. *Cancer Res.* **59**, 4715–4719
  23. Natunen, S., Satomaa, T., Pitkänen, V., Salo, H., Mikkola, M., Natunen, J., Otonkoski, T., and Valmu, L. (2011) The binding specificity of the marker antibodies Tra-1-60 and Tra-1-81 reveals a novel pluripotency-associated type 1 lactosamine epitope. *Glycobiology* **21**, 1125–1130
  24. Fukui, S., Feizi, T., Galustian, C., Lawson, A., and Chai, W. (2002) Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions. *Nat. Biotechnol.* **20**, 1011–1017
  25. Palma, A., Feizi, T., Childs, R., Chai, W., and Liu, Y. (2014) The neoglycolipid (NGL)-based oligosaccharide microarray system poised to decipher the meta-glycome. *Current Opinion Chem. Biol.* **18**, 87–94
  26. Rillahan, C. D., and Paulson, J. C. (2011) Glycan microarrays for decoding the glycome. *Annu. Rev. Biochem.* **80**, 797–823
  27. Gao, C., Liu, Y., Zhang, H., Zhang, Y., Fukuda, M., Palma, A., Kozak, R., Childs, R., Nonaka, M., and Li, Z. (2014) Carbohydrate sequence of the prostate cancer-associated antigen F77 assigned by a mucin O-glycome designer array. *J. Biol. Chem.* **289**, 16462–16477
  28. Li, Z., Gao, C., Zhang, Y., Palma, A., Childs, R., Silva, L., Liu, Y., Jiang, X., Chai, W., and Feizi, T. (2018) O-glycome beam search arrays for carbohydrate ligand discovery. *Mol. Cell. Proteomics* **17**, 121–133
  29. Gao, C., Zhang, Y., Liu, Y., Feizi, T., and Chai, W. (2015) Negative-ion electrospray tandem mass spectrometry and microarray analyses of developmentally regulated antigens based on type 1 and type 2 backbone sequences. *Anal. Chem.* **87**, 11871–11878
  30. Hirvonen, T., Suila, H., Tiitinen, S., Natunen, S., Laukkanen, M., Kotovuori, A., Reinman, M., Satomaa, T., Alfthan, K., and Laitinen, S. (2013) Production of a recombinant antibody specific for i blood group antigen, a mesenchymal stem cell marker. *BioResearch Open Access* **2**, 336–345
  31. Liu, Y., McBride, R., Stoll, M., Palma, A., Silva, L., Agravat, S., Aoki-Kinoshita, K., Campbell, M., Costello, C., Dell, A., Haslam, S., Karlsson, N., Khoo, K., Kolarich, D., Novotny, M., Packer, N., Ranzinger, R., Rapp, E., Rudd, P., and Struwe, W. (2017) The minimum information required for a glycomics experiment (MIRAGE) project: improving the standards for reporting glycan microarray-based data. *Glycobiology* **27**, 280–284
  32. Wang, H., He, W., Jiang, P., Yu, Y., Lin, L., Sun, X., Koffas, M., Zhang, F., and Linhardt, R. (2017) Construction and functional characterization of truncated versions of recombinant keratanase II from *Bacillus circulans*. *Glycoconjugate J.* **35**, 643–649
  33. Nakazawa, K., and Suzuki, S. (1975) Purification of Keratan Sulfate-endogalactosidase and its action on keratan sulfates of different origin. *J. Biol. Chem.* **250**, 912–917
  34. Brown, G., Huckerby, T., Morris, H., Abram, B., and Nieduszynski, I. (1994) Oligosaccharides derived from bovine articular cartilage keratan sulfates after keratanase II digestion: implications for keratan sulfate structural fingerprinting. *Biochemistry* **33**, 4836–4846
  35. Chai, W., Stoll, M., Galustian, C., Lawson, A., and Feizi, T. (2003) Neoglycolipid technology: deciphering information content of glycome. *Methods Enzymol.* **362**, 160–195
  36. Tang, P., Scudder, P., Mehmet, H., Hounsell, E., and Feizi, T. (1986) Sulphate groups are involved in the antigenicity of keratan sulphate and mask i antigen expression on their poly-N-acetylglucosamine backbones. An immunochemical and chromatographic study of keratan sulphate oligosaccharides after desulphation or nitrosation. *Eur. J. Biochem.* **160**, 9
  37. Liu, Y., Feizi, T., Campanero-Rhodes, M., Childs, R., Zhang, Y., Mulloy, B., Evans, P., Osborn, H., Otto, D., Crocker, P., and Chai, W. (2007) Neoglycolipid probes prepared via oxime ligation for microarray analysis of oligosaccharide-protein interactions. *Chem. Biol.* **14**, 847–859
  38. Liu, Y., Childs, R., Palma, A., Campanero-Rhodes, M., Stoll, M., Chai, W., and Feizi, T. (2012) Neoglycolipid-based oligosaccharide microarray system: preparation of NGLs and their noncovalent immobilization on nitrocellulose-coated glass slides for microarray analyses. *Methods Mol. Biol.* **808**, 117–136
  39. Brockhausen, I., and Stanley, P. (2017) O-GalNAc glycans. In: Varki, A., Cummings, R., Esko, J., Stanley, P., Hart, G., Aebi, M., Darvill, A., Kinoshita, T., Packer, N., Prestegard, J., Schnaar, R., and Seeberger, P., eds. *Essentials of Glycobiology*, Cold Spring Harbor, New York
  40. Avnur, Z., and Geiger, B. (1984) Immunocytochemical localization of native chondroitin-sulfate in tissues and cultured cells using specific monoclonal antibody. *Cell* **38**, 811–822
  41. Catterton, B., Christner, J., and Baker, J. (1983) Identification of a monoclonal antibody that specifically recognizes corneal and skeletal keratan sulfate. Monoclonal antibodies to cartilage proteoglycan. *J. Biol. Chem.* **258**, 8848–8854
  42. Zanetti, M., Ratcliffe, A., and Watt, F. (1985) Two subpopulations of differentiated chondrocytes identified with a monoclonal antibody to keratan sulfate. *J. Cell Biol.* **101**, 53–59
  43. Dickenson, J., Huckerby, T., and Nieduszynski, I. (1991) A non-reducing terminal fragment from tracheal cartilage keratan sulphate chains contains  $\alpha$  (2–3)-linked N-acetylneuraminic acid. *Biochem. J.* **278**, 779–785
  44. Nakazawa, K., Ito, M., Yamagata, T., and Suzuki, S. (1989) Substrate specificity of keratan sulphate-degrading enzymes (endo- $\beta$ -galactosidase, keratanase and keratanase II) from microorganisms. In: Greiling, H., and Scott, J., eds. *Keratan Sulphate: Chemistry, Biology, Chemical Pathology*, pp. 99–110, Biochemical Society, London

45. Huckerby, T., Dickenson, J., Brown, G., and Nieduszynski, I. (1995) Keratanase digestion of keratan sulphates: characterization of large oligosaccharides from the N-acetyllactosamine repeat sequence and from the non-reducing terminal chain caps. *BBA-Gen Subjects* **1244**, 17–29
46. Hounsell, E., Feeney, J., Scudder, P., Tang, P., and Feizi, T. (1986) <sup>1</sup>H-NMR studies at 500 MHz of a neutral disaccharide and sulphated di-, tetra-, hexa- and larger oligosaccharides obtained by endo-P-galactosidase treatment of keratan sulphate. *Eur. J. Biochem* **157**, 10
47. Yates, E. A., Santini, F., Guerrini, M., Naggi, A., Torri, G., and Casu, B. (1996) <sup>1</sup>H and <sup>13</sup>C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. *Carbohydr. Res.* **294**, 15–27
48. Mulloy, B., Forster, M., Jones, C., Drake, A., Johnson, E., and Davies, D. (1994) The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modelling study. *Carbohydr. Res.* **255**, 1–26
49. Huckerby, T., Brown, G., and Nieduszynski, I. (1998) <sup>13</sup>C-NMR spectroscopy of keratan sulphates—assignments for five sialylated pentasaccharides derived from the non-reducing chain termini of bovine articular cartilage keratan sulphate by keratanase II digestion. *Eur. J. Biochem.* **251**, 991–997
50. Domon, B., and Costello, C. (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* **5**, 397–409
51. Gao, C., Zhang, Y., Liu, Y., Feizi, T., and Chai, W. (2015) Negative-Ion Electrospray Tandem Mass Spectrometry and Microarray Analyses of Developmentally Regulated Antigens Based on Type 1 and Type 2 Backbone Sequences. *Anal. Chem.* **87**, 11871–11878
52. Weyers, A., Yang, B., Solakylidirim, K., Yee, V., Li, L., Zhang, F., and Linhardt, R. (2013) Isolation of bovine corneal keratan sulfate and its growth factor and morphogen binding. *FEBS J.* **280**, 2285–2293
53. Tai, G., Huckerby, T., and Nieduszynski, I. (1996) Multiple non-reducing chain termini isolated from bovine corneal keratan sulfates. *J. Biol. Chem.* **271**, 23535–23546
54. Leteux, C., Stoll, M., Childs, R., Chai, W., Vorozhaikina, M., and Feizi, T. (1999) Influence of oligosaccharide presentation on the interactions of carbohydrate sequence-specific antibodies and the selectins: Observations with biotinylated oligosaccharides. *J. Immunol. Methods* **227**, 109–119
55. Hirvonen, T., Suila, H., Tiitinen, S., Natunen, S., Laukkanen, M., Kotovuori, A., Reinman, M., Satomaa, T., Alfthan, K., and Laitinen, S. (2013) Production of a recombinant antibody specific for i blood group antigen, a mesenchymal stem cell marker. *BioResearch Open Access* **2**, 336–345
56. Campanero-Rhodes, M., Childs, R., Kiso, M., Komba, S., Narvor, C., Warren, J., Otto, D., Crocker, P., and Feizi, T. (2006) Carbohydrate microarrays reveal sulphation as a modulator of siglec binding. *Biochem. Biophys. Res. Commun.* **344**, 1141–1146
57. Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M., Alvarez, R., Bryan, M., Fazio, F., Calarese, D., Stevens, J., Razi, N., Stevens, D., Skehel, J., Die, I., Burton, D., Wilson, I., Cummings, R., Bovin, N., Wong, C., and Paulson, J. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 17033–17038
58. Galustian, C., Lawson, A., Komba, S., Ishida, H., Kiso, M., and Feizi, T. (1997) Sialyl-Lewis<sup>x</sup> sequence 6-O-sulfated at N-acetylglucosamine rather than at galactose is the preferred ligand for L-selectin and de-N-acetylation of the sialic acid enhances the binding strength. *Biochem. Biophys. Res. Commun.* **240**, 748–751
59. Sasseti, C., Tangemann, K., Singer, M., Kershaw, D., and Rosen, S. (1998) Identification of podocalyxin-like protein as a high endothelial venule ligand for L-selectin: parallels to CD34. *J. Exp. Med.* **187**, 1965–1975
60. Larrucea, S., Butta, N., Arias-Salgado, E., Alonso-Martin, S., Ayuso, M., and Parrilla, R. (2008) Expression of podocalyxin enhances the adherence, migration, and intercellular communication of cells. *Exp. Cell Res.* **314**, 2004–2015
61. Drew, E., Merzaban, J., Seo, W., Ziltener, H., and McNagny, K. (2005) CD34 and CD43 inhibit mast cell adhesion and are required for optimal mast cell reconstitution. *Immunity* **22**, 43–57
62. Yu, S., Hsiao, C., Izawa, M., Yusa, A., Ishida, H., Nakamura, S., Yagi, H., Kannagi, R., and Khoo, K. (2018) Distinct substrate specificities of human GlcNAc-6-sulfotransferases revealed by mass spectrometry-based sulfoglycomics analysis. *J. Biol. Chem.* **293**, 15163–15177
63. Johnson, C., Crawford, B., Stavridis, M., Ten Dam, G., Wat, A., Rushton, G., Ward, C., Wilson, V., Van Kuppevelt, T., and Esko, J. (2007) Essential alterations of heparan sulfate during the differentiation of embryonic stem cells to sox1-enhanced green fluorescent protein-expressing neural progenitor cells. *Stem Cells* **25**, 1913–1923
64. Gasimli, L., Hickey, A. M., Yang, B., Li, G., dela Rosa, M., Nairn, A., Kulik, M., Dordick, J., Moremen, K., and Dalton, S. (2014) Changes in glycosaminoglycan structure on differentiation of human embryonic stem cells towards mesoderm and endoderm lineages. *Biochim. Biophys. Acta* **1840**, 1993–2003
65. Feizi, T., Childs, R., Watanabe, K., and Hakomori, S. (1979) Three types of blood group I specificity among monoclonal anti-I autoantibodies revealed by analogues of a branched erythrocyte glycolipid. *J. Exp. Med.* **149**, 975–980
66. Feizi, T. (1981) The blood group Ii system: a carbohydrate antigen system defined by naturally monoclonal or oligoclonal autoantibodies of man. *Immunol. Commun.* **10**, 127–156
67. Niemann, H., Watanabe, K., Hakomori, S., Childs, R., and Feizi, T. (1978) Blood group i and I activities of “lacto-N-norhexaosylceramide” and its analogues: the structural requirements for i-specificities. *Biochem. Biophys. Res. Commun.* **81**, 1286–1293
68. Feizi, T., Hounsell, E., Alais, J., Veyrieres, A., and David, S. (1992) Further definition of the size of the blood group-i antigenic determinant using a chemically synthesised octasaccharide of poly-N-acetyllactosamine type. *Carbohydr. Res.* **228**, 289–297
69. Solter, D., and Knowles, B. (1978) Monoclonal antibody defining a stage-specific mouse embryonic antigen (SSEA-1). *Pro. Natl. Acad. Sci. U.S.A.* **75**, 5565–5569
70. Andrews, P., Knowles, B., Cossu, G., and Solter, D. (1982) Teratocarcinoma and mouse embryo cell surface antigens: Characterization of the molecule (s) carrying the SSEA-1 antigenic determinant. *Teratocarcinoma and embryonic cell interactions*, 103–119. Available at [https://www.researchgate.net/publication/284059540\\_Teratocarcinoma\\_and\\_mouse\\_embryo\\_cell\\_surface\\_antigens\\_Characterization\\_of\\_the\\_molecules\\_carrying\\_the\\_SSEA-1\\_antigenic\\_determinant](https://www.researchgate.net/publication/284059540_Teratocarcinoma_and_mouse_embryo_cell_surface_antigens_Characterization_of_the_molecules_carrying_the_SSEA-1_antigenic_determinant).
71. Tang, C., Lee, A., Volkmer, J., Sahoo, D., Nag, D., Mosley, A., Inlay, M., Ardehali, R., Chavez, S., and Pera, R. (2011) An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nat. Biotechnol.* **29**, 829–834
72. Matsumoto, S., Nakao, H., Kawabe, K., Nonaka, M., Toyoda, H., Takishima, Y., Kawabata, K., Yamaguchi, T., Furue, M., and Taki, T. (2015) A cytotoxic antibody recognizing lacto-N-fucopentaose I (LNFP I) on human induced pluripotent stem (hiPS) cells. *J. Biol. Chem.* **290**, 20071–20085
73. Nakao, H., Matsumoto, S., Nagai, Y., Kojima, A., Toyoda, H., Hashii, N., Takakura, D., Kawasaki, N., Yamaguchi, T., and Kawabata, K. (2016) Characterization of glycoproteins expressing the blood group H type 1 epitope on human induced pluripotent stem (hiPS) cells. *Glycoconjugate J.* **34**, 779–787
74. Tateno, H., Toyota, M., Saito, S., Onuma, Y., Ito, Y., Hiemori, K., Fukumura, M., Matsushima, A., Nakanishi, M., and Ohnuma, K. (2011) Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J. Biol. Chem.* **286**, 20345–20353
75. Hirabayashi, J., Tateno, H., Onuma, Y., and Ito, Y. (2015) A novel probe as surface glycan marker of pluripotent stem cells: research outcomes and application to regenerative medicine. *Adv. Healthcare Materials* **4**, 2520–2529
76. Tateno, H., Matsushima, A., Hiemori, K., Onuma, Y., Ito, Y., Hasehira, K., Nishimura, K., Ohtaka, M., Takayasu, S., and Nakanishi, M. (2013) Podocalyxin is a glycoprotein ligand of the human pluripotent stem cell-specific probe rBC2LCN. *Stem Cells Translational Med.* **2**, 265–273
77. Henderson, J., Draper, J., Baillie, H., Fishel, S., Thomson, J., Moore, H., and Andrews, P. (2002) Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. *Stem Cells* **20**, 329–337