Supporting Information

Tailored Presentation of Carbohydrates on a Coiled Coil-Based Scaffold for Asialoglycoprotein Receptor Targeting

Elsa Zacco[1], Julia Hütter[2, 4, 5], Jason L. Heier[1], Jérémie Mortier[1, 3], Peter H. Seeberger[2, 4], Bernd Lepenies[2, 4, 6], Beate Koksch[1,]*

[1]Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany; [2]Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14424 Potsdam, Germany; [3]Institute of Pharmacy, Freie Universität Berlin, Königin-Luisestrasse 2, 14195 Berlin; [4]Institute of Chemistry and Biochemistry, Freie Universität Berlin, Arnimallee 22, 14195 Berlin; [5]Present address: Technical University of Denmark, National Veterinary Institute, Section for Virology, Frederiksberg, Denmark. [6]Present address: University of Veterinary Medicine Hannover, Research Center for Emerging Infections and Zoonoses, Bünteweg 17, 30559 Hannover, Germany.

Synthesis of Fmoc-Ser-(*O*-beta-D-galactose-pentaacetate)-OH

For the synthesis and characterization of this compound, please refer to our previous publication.[3]

Synthesis and purification of Fmoc-Ser-(*O*-beta-D-glucose-pentaacetate)-OH

The synthesis of Fmoc-Ser-(*O*-β-D-glucose-pentaacetate)-OH was carried out by dissolving β-D-glucose-pentaacetate (4 mmol) (Sigma, 285943) and vacuum-dried Fmoc-Ser-OH (1.2 eq.) (Iris Biotech, FAA1578) in dry acetonitrile (20 mL). The solution was placed on ice and BF_3-Et_2O (1 eq.) was added. The mixture was then left to slowly warm up to room temperature. An additional equivalent of BF_3- Et_2O was added after both 6 and 15 hours of stirring. After 20 hours, the reaction was shown by TLC to be complete and the mixture was diluted with DCM (30 mL) and sequentially washed with a 1 M HCl (3 x 30 mL) and H_2O (2 x 30 mL). The organic phase was dried with $MgSO_4$ and the solvents were removed under reduced pressure. The crude product was purified by RP-HPLC using a gradient of 40%-100% MeOH in H_2O to obtain 2.5 mmol (62.5 % yield) Fmoc-Ser-(*O*-beta-D-glucose-pentaacetate)-OH.

Fmoc-Ser-(O-beta-D-galactose-pentaacetate)-OH was purified with RediSep® Rf Reversed-phase, using a C18 column with average particle size: 40-63 microns, mesh: 230-400, average pore size: 60 Å, on a CombiFlash Rf, Teledyne Isco, Inc. Lincoln, NE, USA.

^{1}H-NMR (250 MHz,CD_3CN) δ 1.93 (s, 3H, Ac), 1.94 (s, 3H, Ac), 1.97 (s, 3H, Ac), 2.00 (s, 3H, Ac), 3.75-3.85 (m, 2H, Ser-βH, H-6), 4.05 (dd, 1H, *J* = 12.2 Hz, 2.4 Hz, Ser-βH), 4.11 (dd, 1H, *J* = 10.5 Hz, 4.3 Hz, H-6´), 4.25 (t, 1H, *J* = 6.4 Hz, Fmoc CHAr), 4.41 (m, 1H, Fmoc OCH2), 4.32-4.37 (m, 1H, Ser-αH), 4.61(d, 1H, *J* = 8.0 Hz, H-1), 4.85 (dd, 1H, *J* = 9.7, 8.0 Hz, H-4), 5.01 (t, 1H, *J* = 9.7 Hz, H-2), 5.22 (t, 1H, *J* = 9.6 Hz, H-3), 5.80 (d,1H, *J* = 7.9 Hz, NH), 7.35 (brt, 2H, *J* = 7.5 Hz, Ar), 7.43 (brt, 2H, *J* = 7.4 Hz, Ar), 7.67 (d, 2H, *J* = 7.2 Hz, Ar), 7.84 (d, 2H, *J* = 6.9 Hz, Ar).

MS (ESI-TOF, -mode). Compound with chemical formula $C_{32}H_{35}NO_{14}$, calculated mass: 657,21; found: $[M - H]^-$ = 656.19.

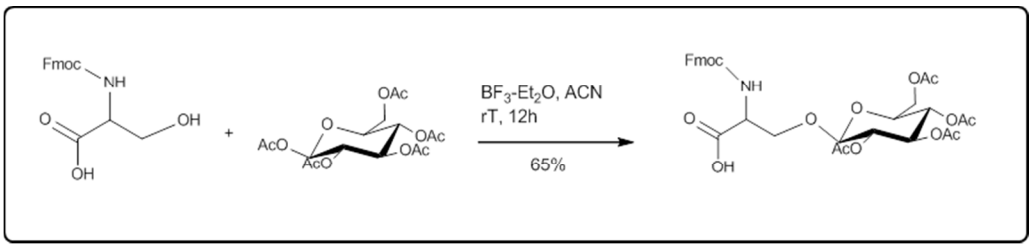


Figure S1. Reaction scheme and yield for the synthesis of Fmoc-Ser-(*O*-beta-D-glucose-pentaacetate)-OH

Synthesis of 2(4-aminobutoxy)-β-D-Galactose

β-D-Galactose-pentaacetate (2.6 mmol) and 1.5 eq. 4-(Z-amino)-1-butanol (Aldrich, 95887) were dissolved in 15 mL of dry DCM. One equivalent of BF_3-Et_2O was added to this solution as it stirred on ice. The mixture was then left to slowly warm up to room temperature. An additional equivalent of BF_3- Et_2O was added after both 6 and 15 hours of stirring. After 20 hours, the reaction was shown by TLC to be complete. The mixture was treated with a saturated solution of $NaHCO_3$ (30 mL), washed with H_2O (2 x 30 mL) and brine (3 x 30 mL) and dried with Na_2SO_4. The crude product (pale yellow) was concentrated and purified by silica gel column chromatography (2:1 Hex:EtOAc) to yield 1.8 mmol (69%) of the glycosylated intermediate.
To remove the acetyl protecting groups, the glycosylated intermediate was dissolved in dry MeOH (5 mL) and 0.8 mmol 30% NaOMe (0.5 eq.) solution in MeOH was added dropwise. After stirring at room temperature for 4 h, the solution was treated with DOWEX 50WX8 50-100(H) resin and stirred until acidic. Upon filtration and solvent removal, 1.48 mmol (82% yield) of the deacetylated intermediate was obtained as a colorless oil.

In order to remove the carboxybenzyl protecting group, the deacetylated intermediate was dissolved in dry MeOH (15 mL) and catalytic Pd on activated carbon (0.1 eq.) was added. The suspension was purged with hydrogen gas for twenty minutes and then stirred under a hydrogen-atmosphere for 3 h. After filtration over celite, the final product was obtained as a colorless, viscous oil (99%).

^{1}H-NMR (500 MHz, CD_3OD) δ 1.58 (quint, 2H, *J* = 6.7 Hz, -C*H*$_2$CH$_2$NH$_2$), 1.64 (quint, 2H, *J* = 6.7 Hz, -OCH$_2$C*H*$_2$-), 2.68 (t, 2H, *J* = 7.0 Hz, -C*H*$_2$NH$_2$), 3.29 (m, 1H, H-6), 3.44-3.51 (m, 3H, H-5, -OCH$_2$-), 3.54-3.58 (m, 1H, H-4), 3.69-3.76 (m, 2H, H-2, H-6´), 3.88-3.93 (m, 1H, H-3), 4.20 (d, 1H, *J* = 7.3 Hz, H-1).

^{13}C-NMR (500 MHz, CD_3OD) δ 28.04 (-OCH$_2$<u>C</u>H$_2$-), 29.84 (-<u>C</u>H$_2$CH$_2$NH$_2$), 42.06 (-<u>C</u>H$_2$NH$_2$), 62.47 (C-6), 70.28 (-O<u>C</u>H$_2$-), 70.48 (C-4), 72.55 (C-2), 75.00 (C-3), 76.60 (C-5), 104.98 (C-1).

MS (ESI-TOF, +mode). Compound with chemical formula $C_{10}H_{21}NO_6$; calculated mass: 251,14; found: $[M + H]^+$ = 252.15. Calculated mass + Na: 274.13; found: $[M + Na]^+$ = 274.13.

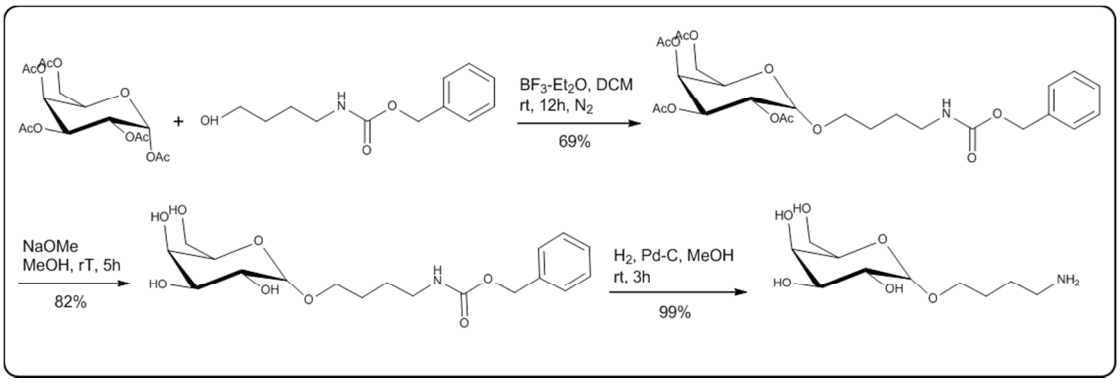


Figure S2. Reaction scheme and yield for the synthesis of 2(4-aminobutoxy)-β-D-Galactose

Glycopeptide spacers

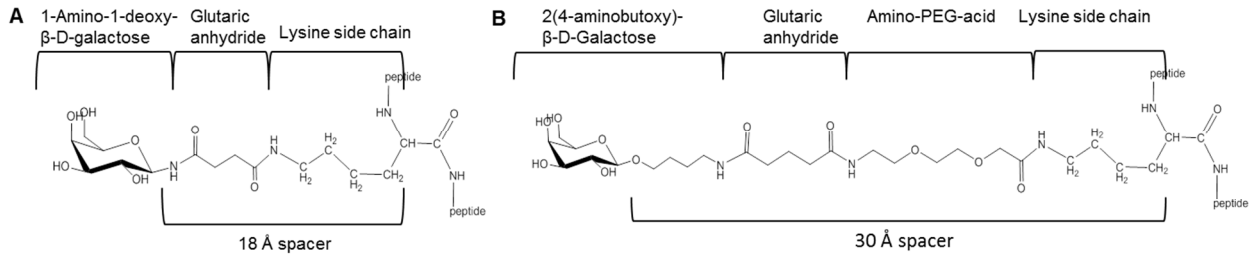


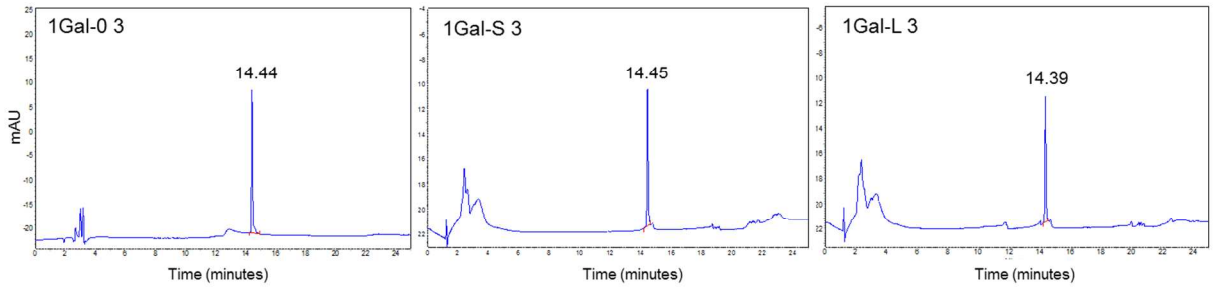
Figure S3. Description and length definition of spacers employed in the glycopeptide synthesis. Distances calculated with the software Molecular Operating Environment, 2013.08; 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada.

Distances between carbohydrate moieties coupled on the scaffold peptide

Position	S_3-S_{10}	S_3-S_{13}	S_{10}-S_{17}	S_3-S_{24}	S_{13}-S_{24}
Distance	~ 13 Å	~ 19 Å	~ 11 Å	~ 33 Å	~ 16 Å

Table S1. Distances between positions on the CCP peptide carrying the ligands. The numbers are an approximation and are calculated according among the selected serine side chains on CCP sequence. The calculation was performed with the software Molecular Operating Environment, 2013.08; 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada.

Glycopeptide library purification



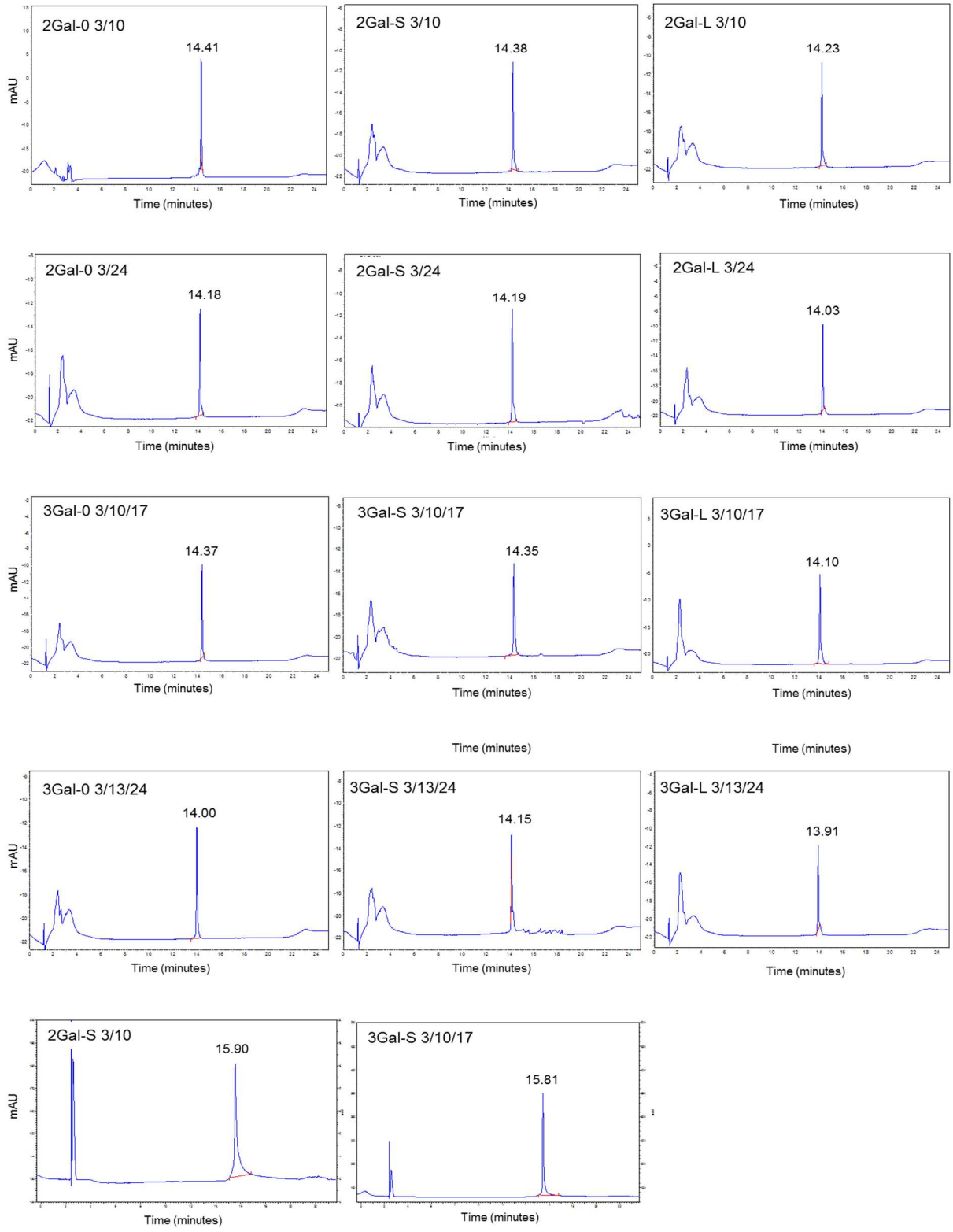


Figure S4. Analytical RP-HPLC of pure glycopeptides. Each run was performed by decreasing the polarity of the mobile phase using 5%-70% ACN in H2Ogradient in 20 minutes, plus 5

minutes of washing and re/equilibration. Chromatographs refer to UV absorption at 230 nm and report name and retention time of the analyzed glycopeptide. For the glycopeptides 2Gal-S 3/1 and 3Gal-S 3/10/17 a different analytical C18 column was used.

ESI-TOF Mass spectrometry

Name	Calculated molecular mass (Da)	Experimental molecular mass (Da)
1Gal-0 3	+1: 3559.91, +2: 1780.46, +3: 1187.31, +4: 890.73, +5: 712.79, +6: 594.16	+3: 1187.67, +4: 891.00, +5: 731.00
2Gal-0 3/10	+1: 3721.91, +2: 1861.46, +3: 1241.31, +4: 931.23, +5: 745.19, +6: 621.16	+3: 1241.68, +4:931.51, +5: 745.41
2Gal-0 3/24	+1: 3721.91, +2: 1861.46, +3: 1241.31, +4: 931.23, +5: 745.19, +6: 621.16	+3: 1241.68, +4: 931.51, +5: 745.41
3Gal-0 3/10/17	+1: 3883.9, +2: 1942.46, +3: 1295.31, +4: 971.73, +5: 777.59, +6: 648.16	+3: 1295.69, +4: 972.02, +5: 777.81
3Gal-0 3/13/24	+1: 3883.9, +2: 1942.46, +3: 1295.31, +4: 971.73, +5: 777.59, +6: 648.16	+3: 1295.70, +4: 972.03, +5:777.82
1Gal-S 3	+1: 3714.77, +2: 1857.89, +3: 1238.93, +4: 929.45, +5: 743.76, +6: 619.97	+3: 1239.04, +4: 929.53, +5: 743.82
2Gal-S 3/10	+1: 4030.41, +2: 2015.71, +3: 1344.14, +4: 1008.36, +5: 806.89, +6: 672.57	+3: 1344.41, +4: 1008.56, +5: 807.05
2Gal-S 3/24	+1: 4030.41, +2: 2015.71, +3: 1344.14, +4: 1008.36, +5: 806.89, +6: 672.57	+3: 1344.41, +4: 1008.56, +5: 807.05
3Gal-S 3/10/17	+1: 4346.65, +2: 2173.83, +3: 1449.55, +4: 1087.42, +5: 870.13, +6: 725.28	+3: 1449.80, +4: 1087.60, +5: 870.28
3Gal-S 3/13/24	+1: 4346.65, +2: 2173.83, +3: 1449.55, +4: 1087.42, +5: 870.13, +6: 725.28	+3: 1449.80, +4: 1087.50, +5: 870.28
1Gal-L 3	+1: 3931.37, +2: 1966.19, +3: 1311.13, +4: 983.60, +5: 787.08, +6: 656.07	+3: 1311.62, +4: 984.10, +5: 787.21
2Gal-L 3/10	+1: 4465.15, +2: 2233.08, +3: 1489.05, +4: 1117.04, +5: 893.83, +6:745.03	+3: 1489.16, +4: 1117.13, +5: 893.90
2Gal-L 3/24	+1: 4465.15, +2: 2233.08, +3: 1489.05, +4: 1117.04, +5: 893.83, +6:745.03	+3: 1489.16, +4: 1117.13, +5: 893.90
3Gal-L 3/10/17	+1: 4998.13, +2: 2499.57, +3: 1666.71, +4: 1250.2901, +5: 1000.43, +6: 833.86	+4: 1250.60, +5: 1000.5, +6: 834.00
3Gal-L 3/13/24	+1: 4998.13, +2: 2499.57, +3: 1666.71, +4: 1250.2901, +5: 1000.43, +6: 833.86	+4: 1250.70, +5: 1000.76, +6: 834.13
1Glc-0 3	+1: 3559.91, +2: 1780.46, +3: 1187.31, +4: 890.73, +5: 712.79, +6: 594.16	+3: 1187.66, +4: 891.00, +5: 731.00
2Glc-0 3/10	+1: 3721.91, +2: 1861.46, +3: 1241.31, +4: 931.23, +5: 745.19, +6: 621.16	+3: 1241.67, +4: 931.50, +5: 745.40
2Glc-0 3/24	+1: 3721.91, +2: 1861.46, +3: 1241.31, +4: 931.23, +5: 745.19, +6: 621.16	+3: 1241.32, +4: 931.40, +5: 745.38
3Glc-0 3/10/17	+1: 3883.9, +2: 1942.46, +3: 1295.31, +4: 971.73, +5: 777.59, +6: 648.16	+3: 1265.69, +4: 972.02, +5: 777.81
3Glc-0 3/13/24	+1: 3883.9, +2: 1942.46, +3: 1295.31, +4: 971.73, +5: 777.59, +6: 648.16	+3: 1295.69, +4: 972.02, +5: 777.81
	Additional glycopeptides	
2Glc-S 3/10	+1: 4030.41, +2: 2015.71, +3: 1344.14, +4: 1008.36, +5: 806.89, +6: 672.57	+3: 1344.42, +4: 1008.57, +5: 807.06
3Gal-S 3/10/17	+1: 4346.65, +2: 2173.83, +3: 1449.55, +4: 1087.42, +5: 870.13, +6: 725.28	+3: 1449.82, +4: 1087.20, +5: 870.30

Figure S5. Monoisotopic molecular masses of the peptides, calculated on positive mode with the software Peptide Mass Calculator v3.2 by Jef Rozenski and experimentally obtained. All molecular masses were determined using ESI-TOF 6210 Agilent (USA, CA-95051-7201, Santa Clara). All samples were dissolved in a mixture of water and ACN before injection.

Further cellular uptake studies

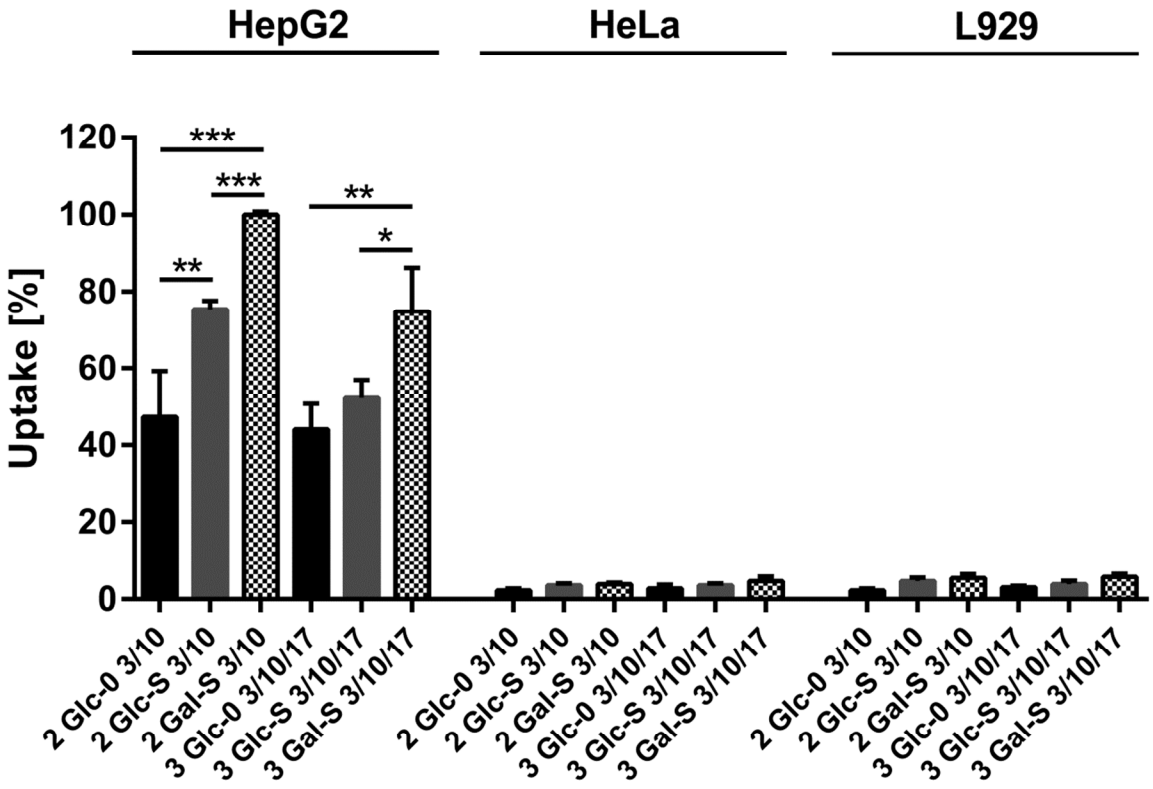


Figure S6. Cell-mediated cell uptake of Gal-functionalized coiled-coil peptides 2Gal-S 3/10 and 3Gal-S 3/10/17 and their Glc-functionalized controls 2Glc-S 3/10, 2Glc-0 3/10, 3Glc-S 3/10/17 and 3Glc-0 3/10/17. From left to right: HepG2 uptake, HeLa uptake and L929 uptake. Percentage of uptake of the glycopeptide library by HepG2 cells. Values were normalized to the average uptake of the best binder, 2Gal-S-3/10. Data are expressed as mean + SEM and are representative of two independent experiments (each performed in triplicates). *p*-Values were determined between each conjugate and the respective Glc-functionalized control peptide using unpaired Student's *t* test (*p<0.05, **p<0.01, ***p<0.005).

Fluorescence microscopy

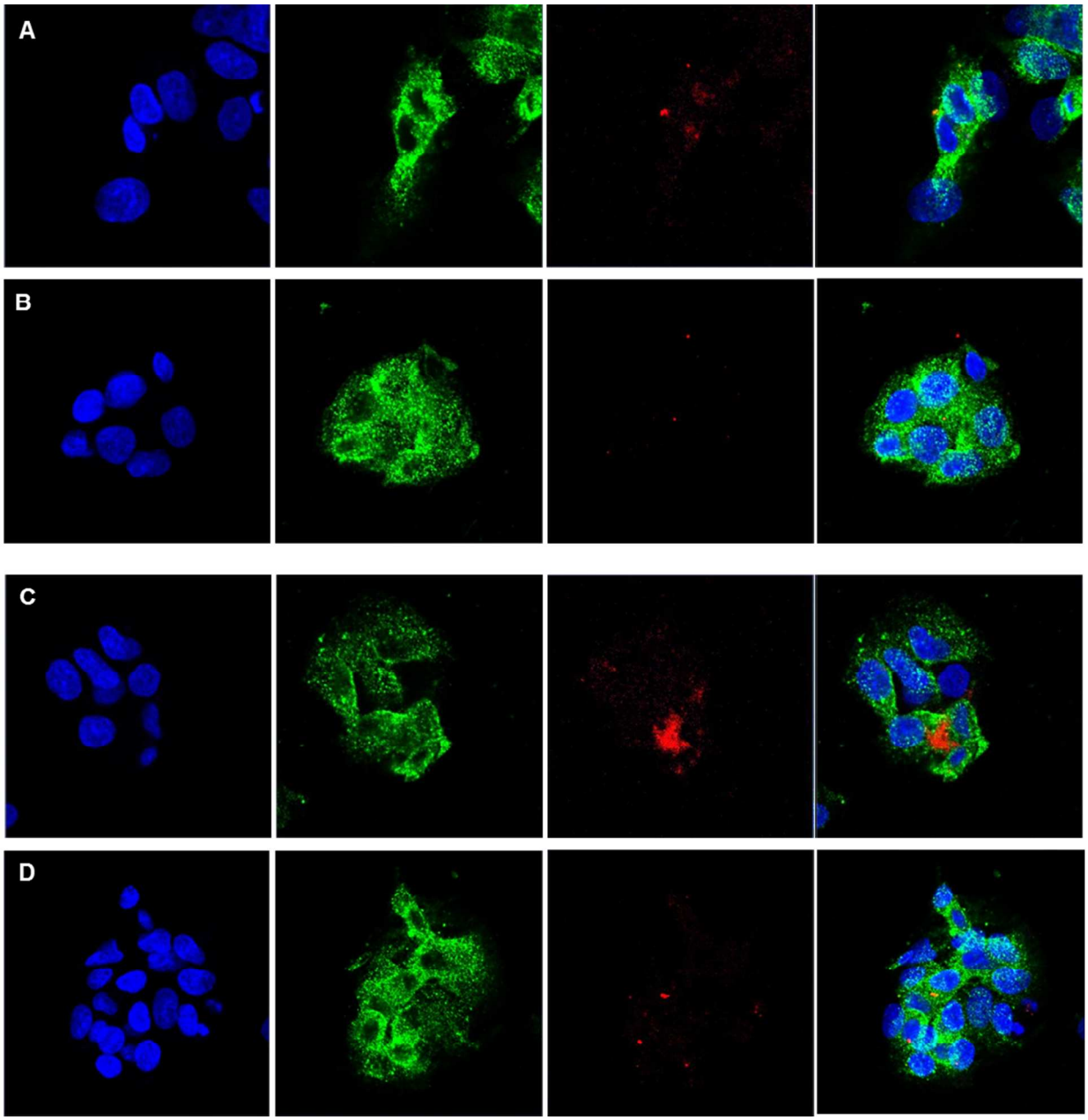


Figure S7. Fluorescence microscopy images of HepG2 cells after 30 minutes incubation of 0.4 uM 2Gal-S 3/10 (**A**); 2Glc-0 3/10 (**B**); 3Gal-S 3/10-17 (**C**); 3Glc-0 3/10/17 (**D**).