



A Bifurcated Proteoglycan Binding Small Molecule Carrier for siRNA Delivery

Matt Gooding^{1,†}, Derick Adigbli², A. W. Edith Chan¹, Roberta J. Melander^{1,‡}, Alexander J. MacRobert² and David L. Selwood^{1,*}

¹The Wolfson Institute for Biomedical Research, UCL, Gower Street, London, WC1E 6BT, UK

²UCL Division of Surgery and Interventional Science, National Medical Laser Centre, Charles Bell House, 67-73 Riding House Street, London, W1W 7EJ, UK

*Corresponding author: David L. Selwood, d.selwood@ucl.ac.uk

[†]Present address: Centre for synthesis and chemical biology, University College Dublin, Benfield, Dublin 4, Ireland

[‡]Present address: Department of Chemistry, North Carolina State University, Raleigh, NC, 27695, USA

A wider application of siRNA- and miRNA- based therapeutics is restricted by the currently available delivery systems. We have designed a new type of small molecule carrier (SMoC) system for siRNA modeled to interact with cell surface proteoglycans. This bifurcated SMoC has similar affinity for the model proteoglycan heparin to an equivalent polyarginine peptide and exhibits significant mRNA knockdown of protein levels comparable to lipofectamine and the previously reported linear SMoC.

Key words: biophysical chemistry, carbohydrate, chemical biology, lipid, nucleic acid, RNAi and antisense techniques

Received 7 June 2013, revised 8 January 2014 and accepted for publication 22 January 2014

Small interfering RNAs (siRNA), which are capable of catalytically silencing specific genes via the RNA interference (RNAi) pathway, are attractive candidates for therapeutics for a wide range of diseases including genetic disorders and cancer. The direct application of unmodified siRNA to the target tissue has shown to be successful in clinical trials, for example an siRNA against a mutated form of keratin 6a to treat the skin condition pachyonychia congenita (TD101, Pachyonychia Congenita Project) (1), and several siRNA-based treatments for age-related macular degeneration (2), in which the naked siRNA is injected directly into the affected area. However, systemic delivery of siRNA, which is required for targeting internal organs and tumors,

is more problematic as siRNA is readily degraded by endogenous nucleases and is rapidly cleared by the renal system. In addition, its large size and high anionic surface charge prohibits siRNA from crossing the plasma membrane and reaching the RNAi machinery in the cytoplasm (3).

To be systemically effective, siRNA must be formulated in such a way as to overcome these delivery barriers. siRNA encapsulated in nanoparticles which protect the siRNA from degradation and improve its half-life *in vivo* as well as aiding cell targeting and internalization is a popular strategy for increasing delivery efficiency (4). Several siRNA-containing nanoparticles have progressed to clinical trials, including TKM-PLK1 by Tekmira, currently in Phase II trials, which uses nucleic acid-lipid particles (LNPs) (5,6) for patients with advanced solid tumors. Alynham in partnership with Tekmira has developed two further LNP-based therapeutics, which are also currently in clinical trials: ALN-VSP in Phase II for liver cancer (7) and ALN-TTR01 for transthyretin-mediated amyloidosis which has completed Phase I trials (8). Another success for nanoparticles has been Calando's CALAA-01, in which siRNA complexed by a cationic cyclodextrin polymer has demonstrated for the first time in humans that systemically delivered siRNA can specifically silence genes (9), and is now in Phase I clinical trials. CALAA-01 consists of siRNA against the M2 subunit of ribonucleotide reductase (RRM2), which is involved in the proliferation of cancer cells, and targeted to cancer cells via a human transferrin protein targeting ligand. Patients with metastatic melanoma were exposed to the drug via 30 min iv infusion and resulted in significant reduction of RRM2 in the tumor cells. As in CALAA-01, polyethylene glycol (PEG) is frequently used to coat nanoparticles to prevent aggregation and binding to serum proteins as well as to evade detection by the immune system (10,11). In addition, once the nanoparticles reach their target cells, they must penetrate the plasma membrane in order to allow the siRNA to access the RNA-induced silencing complex (RISC) located in the cytoplasm (12).

Cationic polymers such as polyethyleneimine (13), lipids (5,6,14), and cell-penetrating peptides (CPPs), including penetratin (15), polyarginine (16), and MPG (17) have been used to deliver siRNA *in vivo*, and form complexes via electrostatic interactions which neutralize the charge of the oligonucleotide and allow interaction with cell surface lipids and proteoglycans. As well as shielding the

siRNA charge, these cationic materials have been found to promote cell entry of siRNA as they are effective at crossing cell membranes. Although the exact mechanism of cellular entry adopted by these polymers and peptides of cellular entry remains unclear, it is suggested that binding to cell surface proteoglycans such as heparan sulfate and chondroitin sulfate via their glycosaminoglycan (GAG) chains may be an initial internalization step, activating Rac1 and triggering subsequent rearrangement of the F-actin cytoskeleton leading to macropinocytosis initiation (18). Several studies have used heparin, a highly sulfated form of heparan sulfate, to estimate binding affinities of CPPs to cell surface GAGs, which were found to be in the micromolar to nanomolar range (19,20), while a more recent study shows evidence that CPPs are able to form clusters with multiple heparin chains (21).

We previously described that the small molecule carrier (SMoC) class of compounds, which are small molecule mimics of penetratin, show potential as siRNA delivery agents (22), showing similar levels of gene silencing in IMR-90 cells to the commonly used transfection reagent Lipofectamine[®]. SMoCs possess cationic guanidine groups linked to a biphenyl backbone, forming amphipathic alpha-helix mimics which are capable of crossing cell membranes (23,24). We postulated that the positive charges may be partially stabilized by π -cation interactions with the biphenyl backbone to promote electrostatic binding to siRNA and to anionic cell surface proteoglycans. The most successful SMoC compound for siRNA transfection described thus far is 4G-SMoC-SSPy **1**, (Figure 1, see Figure S1 for the naming convention for SMoCs). In the present study, we hypothesized that a bifurcated SMoC compound, 4G-BfSMoC-COOH, **2** would achieve the twin aims of cell surface proteoglycan binding and more flexible chemical synthesis. We observed that **2** retains the ability to complex and deliver siRNA, resulting in knockdown of the cell cycle gene *cdc7*. 4G-BfSMoC-COOH, **2** and other SMoCs show significant binding to heparin as a model for cell surface proteoglycans. **2** is easily synthesized and includes a central chemical 'handle' that will allow covalent attachment to other biomolecules.

Methods and Materials

Starting materials were either commercially available or synthesized according to methods reported in the literature. ¹H and ¹³C NMR spectra were recorded on either a Bruker AMX-300 or a Bruker AMX-500 spectrometer. Chemical shifts are reported as ppm relative to TMS internal standard. Mass spectra were recorded on either a Fisons VG70-SE spectrometer (EI, FAB) or an Agilent 6100 Series LC-mass spectrometer (Wokingham, Berkshire, UK) using either a C-18 or C-4 column. Microwave reactions were carried out using a CEM Discover microwave. All compounds were purified to $\geq 95\%$ as measured by

Bifurcated Proteoglycan Binding Small Molecule Carrier

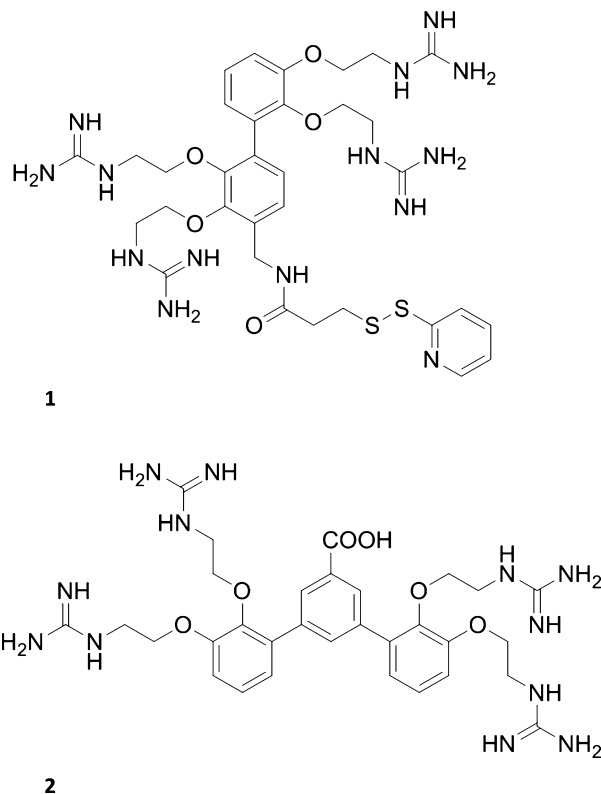


Figure 1: Chemical structures of 4G-SMoC-SSPy, **1** and 4G-BfSMoC-COOH, **2**.

LC-MS. Proteoglycan binding was recorded using an Agilent 1200 series HPLC system equipped with a diode array detector and HiTrap heparin agarose column. A complete list of ¹H NMRs for compounds used in this study is given in the supplementary information (Table S1).

Molecular modeling

Molecular modeling and dynamics were carried using Molecular Operating Environment (MOE) package (version 2011.10). The framework structures **3** and **4** were initially built in MOE (Chemical Computing Group, Montreal, Canada) using the 'Molecular Builder' with an extended conformation. All the guanidine groups are in +1 charge, reflecting the protonated state of the groups. Then, the initial structures were subjected to energy minimization using MMDD94X force field. The energy minimized structures of **3** and **4** were then carried through for dynamics simulations. All molecular dynamics simulations were conducted in a cube of water of depth of 10 Å using 'Simulations-Dynamics' in MOE. The protocol used in dynamics simulations is as follows: algorithm – Nosé-Poincaré-Andersen Hamiltonian equations of motion (NPA); temperature – constant at 300°K; mode – equilibration for 100 ps, force-field – MMFF94X, cutoff at 8 Å. Conformations were collected in trajectory files every 0.5 ps, providing 200 conformers for each run for each molecule. The molecular

dynamics simulations were also run in YASARA software (25) with similar results.

Data analysis, such as distance measurement, was carried out also in MOE using the graphics interface. The pairwise distances between the carbon atoms of the guanidine group from different phenyl rings in each molecule **3** or **4** were measured for each conformation. A histogram summarizing the data was calculated and plotted using Microsoft Excel.

Crystal structure information for heparin (PDB ID: 1HPN) and chondroitin sulfate (PDB ID: 1C4S) was obtained from the PDB and imported into DS Visualizer (Accelrys, San Diego, CA, USA) and the distance between sulfate residues measured between sulfur atoms.

Synthesis

3,5-bis(2,3-bis(2-(((benzyloxy)carbonyl)amino)ethoxy)phenyl)benzoic acid **6**

To a degassed solution of potassium (2,3-bis(2-(((benzyloxy)carbonyl)amino)ethoxy)phenyl)trifluoroborate (**22**) (3.62 g, 6.34 mmol) in isopropanol/water (2/1, 21 mL) was added PdCl₂dppf.CH₂Cl₂ (123 mg, 0.15 mmol) and triethylamine (1.77 mL, 12.68 mmol) and the solution stirred for 2 min. 3,5-Diiodobenzoic acid, **5** (789.1 mg, 2.11 mmol) was added, and the reaction heated in a microwave reactor for 10 min at 75 °C. The reaction mixture was cooled, filtered, and 1 M HCl added (50 mL). The product was extracted with dichloromethane (3 × 25 mL), and the combined organic extracts were washed with brine (50 mL), dried over MgSO₄, filtered, and concentrated under vacuum. The mixture was purified by flash column chromatography using a gradient of 0–50% ethyl acetate in cyclohexane and the title product (1.7 g, 80% yield) obtained as a colorless oil.

¹H-NMR (500 MHz, CDCl₃) δ: 3.27 (s, 4H, 2 × CH₂), 3.48–3.74 (m, 8H, 4 × CH₂), 4.11 (s, 4H, 2 × CH₂), 4.97–5.10 (m, 8H, 4 × Ph-CH₂OCONH), 6.68–7.11 (m, 6H, 6 × ArH), 7.23–7.32 (m, 20H, 20 × Cbz-ArH), 7.66–8.29 (m, 3H, 3 × ArH).

¹³C-NMR (500 MHz, CDCl₃) δ: 40.81 (CH₂), 41.13 (CH₂), 66.87 (Cbz CH₂), 68.22 (CH₂), 113.43, 123.04, 124.90, 128.20 (Cbz ArC), 128.51 (Cbz ArC), 135.03, 136.44, 138.37, 145.06, 152.02, 156.77 (C = O).

HRMS-ES (*m/z*): [M – H][–] calcd for C₅₉H₅₇N₄O₁₄, 1045.3871; found, 1045.3804.

3,5-bis(2,3-bis(2-(((tert-butoxycarbonyl)guanidino)ethoxy)phenyl)-benzoic acid

To a solution of 3,5-bis(2,3-bis(2-(((benzyloxy)carbonyl)amino)ethoxy)phenyl)-benzoic acid **1** (813 mg, 0.7 mmol) in dichloromethane (20 mL) was added 30% hydrogen bromide in acetic acid (5 mL) dropwise and the mixture

stirred at room temperature for 3 h. The reaction was diluted with dichloromethane (20 mL) and the product extracted with water (3 × 20 mL). The aqueous extracts were carefully dried under vacuum for several hours. The crude amine was then dissolved in methanol (20 mL), and *N,N'*-Di-Boc-1*H*-pyrazole-1-carboxamide (922 mg, 3.15 mmol) and DIEA (1.22 mL, 7 mmol) were added. The reaction was stirred at room temperature for 24 h under nitrogen. The solvent was removed under reduced pressure and the residue redissolved in ethyl acetate (50 mL) and 1 M HCl added (50 mL). The layers were separated, and the aqueous layer washed with ethyl acetate (3 × 20 mL). The combined organic extracts were washed with brine (50 mL), dried over MgSO₄, filtered, and concentrated under vacuum. The product was purified by flash column chromatography using a 0–25% ethyl acetate gradient in cyclohexane. The product was obtained as an off-white solid (651 mg, 63% yield).

¹H-NMR (500 MHz, CDCl₃) δ: 1.42–1.51 (m, 72H, 24 × Boc-CH₃), 3.55 (m, 4H, 2 × CH₂), 3.87 (m, 4H, 2 × CH₂), 3.93 (m, 4H, 2 × CH₂), 4.23 (m, 4H, 2 × CH₂), 6.99 (m, 2H, 2 × ArH), 7.03 (m, 2H, 2 × ArH), 7.13 (m, 2H, 2 × ArH), 7.98 (m, 1H, ArH), 8.28 (m, 2H, 2 × ArH), 8.66 (m, 2H, 2 × NH), 8.81 (m, 2H, 2 × NH), 11.45 (bs, 1H, COOH).

¹³C-NMR (500 MHz, CDCl₃) δ: 28.06 (Boc-CH₃), 28.17 (Boc-CH₃), 28.36 (Boc-CH₃), 28.39 (Boc-CH₃), 53.49 (CH₂), 67.43 (CH₂), 79.06 (Boc CCH₃), 79.35 (Boc CCH₃), 82.80 (Boc CCH₃), 83.20 (Boc CCH₃), 113.31, 123.34, 124.77, 129.78, 135.12, 138.33, 145.16, 152.11, 153.01, 156.43, 163.54, 169.33 (COOH).

HRMS-ES (*m/z*): [M + H]⁺ calcd for C₇₁H₁₀₇N₁₂O₂₂, 1479.7423; found, 1479.7423.

4G-BfSMoC-COOH, 3,5-bis(2,3-bis(2-(guanidinoethoxy)phenyl)-benzoic acid **2**

The tetra-Boc compound (30 mg, 0.02 mmol) was dissolved in DCM (2 mL), and TFA/*m*-cresol (95/5, 2 mL) was added and the mixture stirred for 4 h at room temperature. The solvent was removed under vacuum, and the residue washed three times with ether. The product was dissolved in water and lyophilized to form the tetra-TFA salt as a white solid (23 mg, 0.02 mmol, 100% yield).

¹H NMR (600 MHz, D₂O) δ 8.17 (d, *J* = 1.6 Hz, 2H), 7.86 (d, *J* = 1.1 Hz, 1H), 7.33–7.27 (m, 2H), 7.20 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.10 (dt, *J* = 15.6, 7.8 Hz, 2H), 4.35–4.29 (m, 2H), 4.02–3.97 (m, 2H), 3.73 (t, *J* = 4.9 Hz, 2H), 3.23–3.18 (m, 2H).

¹³C NMR (151 MHz, D₂O) δ 171.06, 157.88, 157.21, 151.90, 144.44, 138.55, 135.89, 135.33, 130.64, 130.37, 126.24, 123.69, 114.53, 71.96, 67.55, 42.15, 41.61. Also visible peaks for trifluoroacetate, 163.98, 163.74, 163.51, 163.27, 119.86, 117.93, 116.00, 114.06.



HRMS-ES (m/z): $[M+H]^+$ calcd for $C_{31}H_{43}N_{12}O_6$, 679.3429; found, 679.3428.

2, 3, 2', 3'-tetra(2-guanidino-ethoxy)-4-cyanobiphenyl, **8**, (4G-SMoC-CN).

To a solution of 2, 3, 2', 3'-tetra(2-[*N*, *N'*-bis(tert-butoxycarbonyl)guanidino]-ethoxy)-4-cyanobiphenyl **7** (22) (40 mg, 0.0289 mmol) in DCM (2 mL) was added a solution of TFA/*m*-cresol (95:5) (2 mL). The reaction mixture was stirred at room temperature for 3 h. Solvent was evaporated and the product precipitated from Et_2O , washed three times with Et_2O , dissolved in water, and lyophilized to give the title product (21 mg, 70%) as a white powder.

1H NMR (500 MHz, MeOH) δ 7.50 (d, $J = 8.0$, 1H, aromatic), 7.20 (d, $J = 8.0$, 2H, aromatic), 7.15 (dd, $J = 8.3$, 1.6, 1H, aromatic), 6.91 (dd, $J = 7.5$, 1.6, 1H, aromatic), 4.38 (t, $J = 5.0$, 2H CH_2), 4.24 (t, $J = 5.2$, 2H, CH_2), 4.00 (t, $J = 4.9$, 2H, CH_2), 3.92 (t, $J = 4.9$, 2H, CH_2), 3.68 (dt, $J = 9.8$, 5.1, 4H, $2 \times CH_2$), 3.28 (dd, $J = 9.6$, 4.7, 4H, $2 \times CH_2$).

HRMS m/z (ES^+) 584.3186, calculated $C_{25}H_{38}N_{13}O_4^+$ 584.3170.

Gel shift assay

GAPDH siRNA (Thermo Scientific, Loughborough, UK; sequences used were sense 5' UGG UUU ACA UGU UCC AAU AUU 3', antisense 5' Phosphate U AUU GGA ACA UGU AAA ACC UU 3') was mixed in RNase-free water with varying concentrations of either SMoC, PySSCH₂CH₂CO-RRRR-NH₂, **9** (R4SSPy) or PySSCH₂CH₂CO-RRRRRRRR-NH₂, **10**, (R8SSPy), (Peptide Synthetics, Ltd, UK) in the molar ratios of 1/1, 1/5, 1/10, 1/20 and 1/50. The final siRNA concentration was 17 μM in 100 μL of water (1.7 nmol) in each case. The complexes were incubated for 30 min at room temperature. The samples were diluted with 2 \times RNA loading dye (Fermentas, now Thermo Scientific) containing ethidium bromide (composition 95% formamide, 0.025% SDS, 0.025% bromophenol blue, 0.025% xylene cyanol FF, 0.025% ethidium bromide, 0.5 mM EDTA) and analyzed by gel electrophoresis on a 1% agarose gel prepared using 10 \times MOPS buffer [0.4 M MOPS (pH 7), 0.1 M sodium acetate, 0.01 M EDTA (pH 8)], loading 25 μL of sample into each well. The gel was run at 5V/cm till the blue dye had migrated 2/3 of the gel. The gels were analyzed by UV illumination and the ethidium bromide bands quantified using the ImageJ software (NIH, Bethesda, MD, USA). The EtBr intensities were plotted, a dose-response curve fitted and EC_{50} values calculated using the Origin software (OriginLab, Northampton, MD, USA).

Cell culture

IMR-90 (ATCC# CCL-186) cells were obtained from LGC Standards, UK, at population doubling (PD) 12. Cells were

Bifurcated Proteoglycan Binding Small Molecule Carrier

cultured in Dulbecco's modified Eagles Medium (DMEM) (Invitrogen, Life Technologies, Paisley, UK) supplemented with 10% Fetal Bovine Serum (Invitrogen) and 2 mM L-glutamine (Invitrogen) at 37 °C and 5% CO₂. 4T1 cells were purchased from Caliper Life Sciences (Hopkinton, MA, USA). MCF-7 cells were obtained from the European Collection of Animal Cell Cultures, Porton Down, UK.

Confocal microscopy in vitro

Cell preparation: 1×10^4 4T1 (murine) and MCF-7 (human) breast cancer cells were seeded in 35 mm diameter glass-bottomed fluorodishes (WPI, Hitchin, Hertfordshire, UK). These were incubated overnight to encourage adherence and optimal subconfluent cell spreading for imaging. 4T1 cells were incubated with fresh RPMI-1640 Medium with L-glutamine and sodium bicarbonate (Sigma-Aldrich, Poole, Dorset, UK) supplemented with 10% (v/v) fetal calf serum (FCS) (Gibco, Life Technologies, Paisley, UK, Invitrogen); MCF-7 cells were incubated with fresh DMEM containing 0.365 g/L L-glutamine, 4.5 g/L glucose (BioWhittaker, Verviers, Belgium), supplemented with 10% (v/v) FCS (Sigma-Aldrich), 100 U/mL penicillin and 100 $\mu g/mL$ streptomycin at 37 °C in a humidified 5% CO₂ incubator.

Confocal: Cells were observed using an inverted Olympus Fluoview 1000 confocal laser-scanning microscope to determine intracellular localization of the oligonucleotide. Fluorescence confocal images were obtained using a 60 \times NA:1.35 oil immersion objective lens (Olympus Fluoview FV1000; Olympus, Southend-on-Sea, UK). Original image files were analyzed using Image J software. Analysis of the labeled siRNA \pm 4G-BfSMoC-COOH, **2** was performed using image J to quantify the mean intracellular signal intensity compared to the background level of fluorescence. A one-way ANOVA with Tukey's multiple comparison test was used to assess statistical significance. Appropriate controls were used throughout.

RNA oligonucleotide, lipofectamine, and 4G-BfSMoC-COOH, **2** preparation

The intracellular localization of Alexa Fluor[®] Red Fluorescent oligonucleotide (Invitrogen) supplied as a 20 μM stock of Alexa Fluor[®] 555-labeled, double-stranded RNA oligomer in 100 mM KOAc, 30 mM HEPES-KOH, pH 7.4, and 2 mM MgOAc (annealing buffer) was determined through confocal fluorescence microscopy. Lipofectamine[®] RNAiMAX (Invitrogen). The oligonucleotide was diluted to a final concentration of 1 μM either alone or in combination with 10 μM 4G-BfSMoC-COOH, **2**. The initial dilutions were in phosphate buffered saline (PBS) (Lonza, Slough, UK). The oligonucleotide/4G-BfSMoC-COOH, **2** mixture was allowed to complex for 15 min at room temperature. Final dilutions were made in the appropriate cell medium. Either cell line was subsequently incubated with media (alone), oligonucleotide(alone), 4G-BfSMoC-COOH, **2** (alone), or oligonucleotide/4G-BfSMoC-COOH, **2** for 24 h. Lipofectamine

controls were transfected according to manufacturer's instructions to a final concentration of 3.75 $\mu\text{L}/\text{mL}$ + oligonucleotide. Cells were washed twice with PBS and treated with phenol-red-free media prior to confocal imaging.

siRNA transfection

Solutions were made up in Opti-MEM[®] (Invitrogen); 100 nM *cdc7* siRNA (sense 5'-GCU CAG CAG GAA AGG UGU UUU-3', antisense 5'-AAC ACC UUU CCU GCU GAG CUU-3'; Applied Biosystems, Warrington, UK), 100 nM Accell Non-targeting siRNA 1 negative control siRNA (Thermo Scientific), 8 $\mu\text{L}/\text{mL}$ Lipofectamine LTX reagent (Invitrogen), 0.5 mM SMOc. SMOc/siRNA, and Lipofectamine/siRNA complexes were made by mixing equal volumes of the required solutions and incubating for 30 min at room temperature.

Cells were plated in 6-well plates at a density of 5×10^5 cells/well in 1 mL DMEM containing 10% FCS and 2 mM glutamine and incubated at 37 °C under 5% CO₂ for 24 h. The media was removed and replaced with 1.6 mL Opti-MEM[®] (Invitrogen). Four hundred microliters of SMOc/siRNA or Lipofectamine/siRNA complexes was added and the cells incubated for a further 72 h. The final concentration of SMOc was 50 μM and based on the dose–response analysis reported previously (22). The cells were washed with PBS and trypsinized and the pellet stored at –80 °C.

RNA purification and qRT-PCR

RNA was extracted from the cell pellets using the PureLink RNA Mini kit (Invitrogen) using the protocol provided by the manufacturer. RNA was quantified using a Nanodrop spectrophotometer, and 40 ng of RNA was mixed with Superscript III RT/Platinum Taq Mix and SYBR Green Reaction Mix from the Superscript III Platinum SYBR Green One-Step qRT-PCR kit (Invitrogen) according with the manufacturer's protocol. PCR conditions were 50 °C for 3 min, 95 °C for 10 min (one cycle), 95 °C for 15 seconds, 47 °C for 30 seconds, 60 °C for 30 seconds (45 cycles), 95 °C for 15 seconds, 60 °C for 15 seconds, ramp to 95 °C over 20 min (one cycle). Relative quantification data were obtained using the comparative C_t method with Realplex software according to the manufacturer's protocol (Eppendorf, Hamburg, Germany). *cdc7* RNA was measured using the following primers: forward primer: 5'-AAC TTG CAG GTG TTA AAA AAG-3'; reverse primer: 5'-TGA AAG TGC CTT CTC CAA T-3'. The following primers were used to measure GAPDH: forward primer: 5'-TCA ACT ACA TGG TTT ACA TGT TC-3'; reverse primer: 5'-GAT CTC GCT CCT GGA AGA T-3'. The signal obtained from the *cdc7* primers was normalized using the signal from the GAPDH primers.

Proteoglycan binding

The test SMOcs and the marker oligoarginine peptides R4SSPy, **9** and R8SSPy, **10** were applied to a heparin

agarose column (HiTrap, GE-Healthcare Life Sciences, UK) (1 mL) and eluted with a gradient of sodium chloride from 500 to 1000 mM. The absorbance at 214 nm (CONH) was recorded and exported to Origin 9.0. The signals were normalized and replotted as shown in the Figure. HIVtat, **11**, (sequence YGRKKRRQRRR) was obtained from Sigma.

Statistical analysis

Groups were compared using an unpaired two-sample Student's *t*-test. A value of $P < 0.05$ was considered significant.

Results

Design of a bifurcated SMOc derivative

As binding to cell surface GAGs such as heparan sulfate and chondroitin sulfate is thought to play a role in CPP internalization, the active transfection compound **1** (Table 1) was studied via molecular modeling and dynamics for its ability to bind these proteoglycans. The structures of chondroitin sulfate and heparin, a highly sulfated form of heparan sulfate, were obtained from the Protein Databank, and the distances between sulfate residues were measured (Figure 2A,B). Examination of the crystal structure of heparin (pdb: 1HPN) revealed that the sulfate residues were positioned in 'clusters' of three on alternating sides of the sugar chain. The distance between residues within each cluster varies between 6.1 and 7.8 Å with an average distance of 6.6 Å, whereas the distance between clusters on the same side of the chain averages of 17.5 Å (Figure 2A). Due to flexibility of the chain, it is possible that in nature, the distances between sulfate residues may vary considerably from the crystal structure. For the crystal structure obtained for chondroitin sulfate (pdb: 1C4S), the distance between sulfate groups was measured as 12.7 Å (Figure 2B), with sulfate groups positioned on alternating sides of the sugar chain. As the SMOc guanidine groups may not be able to access sugars on opposite sides of the chain, the distance between sulfates on the same side was also measured and found to be approximately 20.7 Å.

To gauge the interaction between our molecules and heparin and chondroitin sulfate, the pairwise distances between the carbon atoms of the guanidine group from different phenyl rings for each conformation observed during molecular dynamics simulation of structure **3** were measured; that is, four distances were measured: between C2-C2', C2-C3', C3-C2', and C3-C3' as labeled in Figure 3. The results are plotted as a histogram showing the frequency of the C-C distances. More than 75% of the time, the C-C distances are above 11 Å. The most frequently occurring distances between two carbons (35%) are between 11 and 12 Å, mainly between groups C3-C2' and C2-C3'. The longest distance is averaged around

Table 1: Proteoglycan binding: retention times of arginine rich peptides and SMoCs on a heparin column

Compound	Structure	Short name	Guanidines	Retention time	Reference for compound
9	<chem>PySSCH2CH2CO-RRRR-NH2</chem>	R4SSPy	4	4.4	(22)
10	<chem>PySSCH2CH2CO-RRRRRRRR-NH2</chem>	R8SSPy	8	19.17	(22)
11	YGRKKRRQRRR	HIV-1 tat	6	15.2	This study
1	See Figure 1	4G-SMoC-SSPy	4	8	(23)
2	See Figure 1	4G-BfSMoC-COOH	4	4.69	This study
12		2G-SMoC	2	3.15	(22)
13		3G-SMoC-SSPy	3	4.6	(24)
8		4G-SMoC-CN	4	7.4	This study
14		6G-SMoC	6	17.12	(22)

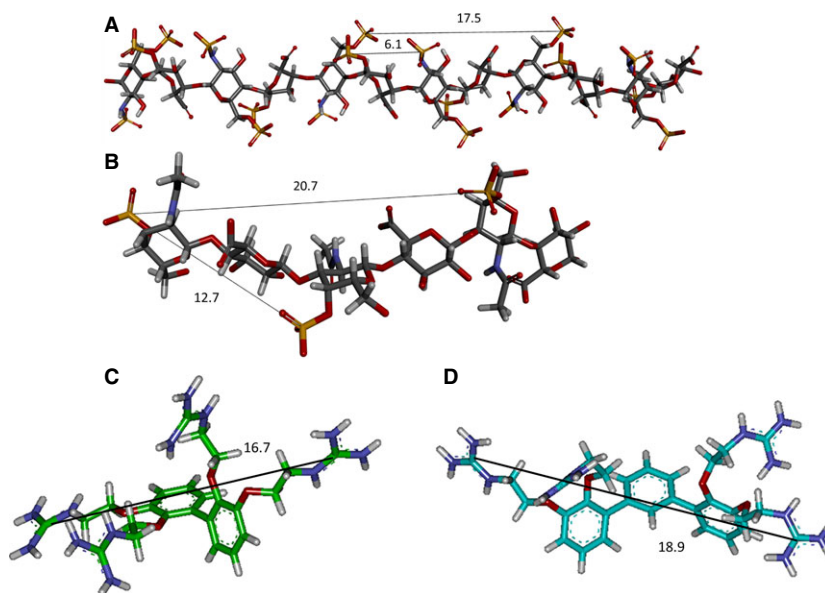


Figure 2: Modeling a new SMoC analogue for binding to cell surface glycosaminoglycans. (A) Examination of the crystal structure of heparin. Clusters of three sulfate residues are observed, occurring on alternating sides of the sugar chain. Representative distances between residues in the same cluster, and between residues of adjacent clusters on the same side of the chain are shown in Å. (B) Chondroitin sulfate crystal structure obtained from the PDB. The distances between sulfate ions are marked in Å. (C) A representative snapshot taken from the molecular dynamics simulation of **3**. The average distance between the furthest guanidine groups is indicated in Å. (D) A snapshot from the molecular dynamics simulation of **4**. The average distance between the furthest guanidine groups is indicated in Å.

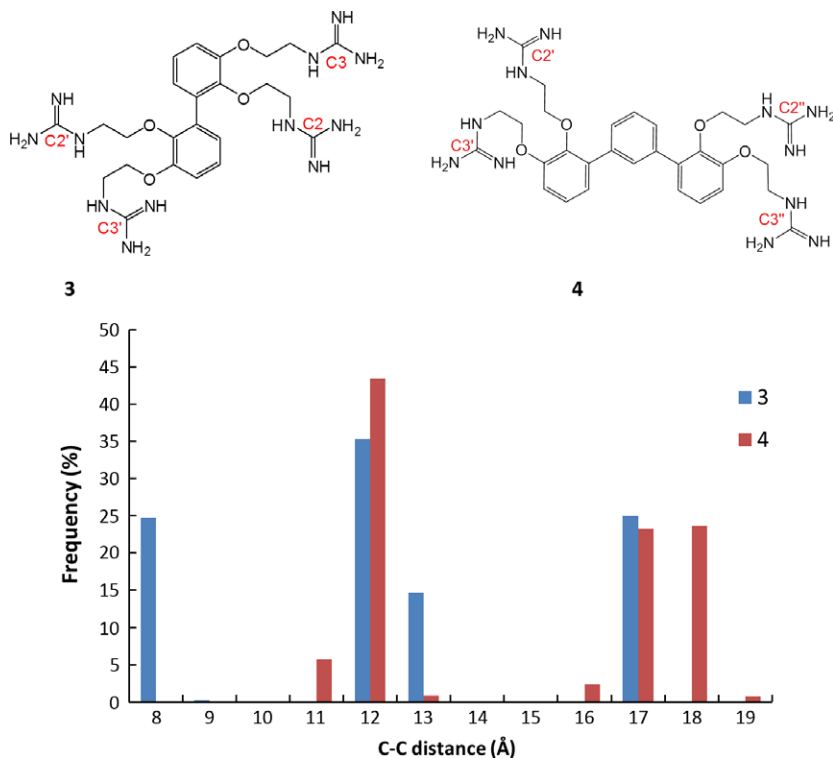


Figure 3: Frequency of distances: between the two carbon atoms of the guanidine groups attached to different phenyl rings for structures **3** and **4**. The carbon atoms are labeled in red in the diagram.

16.48 Å, contributed mainly from groups C3-C3'. This distance is sufficient for binding to heparin sulfates in the same cluster, as well as binding to chondroitin sulfates on opposite sides of the sugar chain. However, to optimize binding to these GAGs, the distance between guanidine residues should be slightly increased in order that it is possible to bridge the gap between heparin sulfate clusters, as heparan sulfate is less sulfated than heparin, and the sulfate groups may be more sparsely arranged. Furthermore, macropinocytosis is initiated by clustering of GAG chains attached to different proteoglycan molecules (26). Therefore, increased spacing between SMOc guanidine groups may allow for binding of sulfates on different GAG chains.

To extend the distance between guanidine groups in the SMOc structure, we decided to investigate the incorporation of an additional phenyl ring to the original structure, to produce a bifurcated structure, **4** (Figure 3). Again molecular dynamics was carried out for this molecule. We predicted that the ring should be added at the same location as the biphenyl bond to preserve the π -cation stabilization of the guanidine charges (22). The pairwise distances between the carbon atoms of the guanidine group from different phenyl rings for each conformation observed during molecular dynamics simulation of structure **4** were measured; that is, four distances were measured: between C2'-C2'', C2'-C3'', C3'-C2'', and C3'-C3''. From Figure 3, almost half of the time (50%), the distances between any 2 guanidine carbon atoms are larger than 15 Å; and 25% of the distances are above 17 Å, contributed mainly from

groups C3'-C3''. Therefore, compound **2** is sufficient for binding to adjacent sulfate clusters in heparin and may also bind heparan sulfate in a different way to **1** at the cell surface. The additional distance between guanidines may also allow binding of additional sulfate groups on chondroitin sulfate, as well as increasing the possibility of promoting clustering of GAG chains.

Synthesis of 4G-BfSMoC-COOH

The synthesis of the new SMOc partly followed the synthetic route previously described for the biphenyl 4G-SMOc-SSPy, **1** (22) (Figure 4A). The additional phenyl ring was added via 3,5-diiodobenzoic acid, **5** a commercially available compound that also contains an acid for use as a linker group to attach targeting ligands. The synthesis made use of the trifluoroborate salt described previously (22), which had been synthesized in bulk, in order to introduce the amine functionality for adding guanidine groups. The terphenyl structure was created using a Suzuki-Miyaura palladium coupling reaction under microwave irradiation at 80 °C with the trifluoroborate salt, producing the coupled product **6** in 80% yield. The guanidinylation step proceeded in the presence of the unprotected acid group using the pyrazole guanidinylation reagent in 63% yield, demonstrating that the acid does not react with the guanidinylation reagent. The guanidines were deprotected with TFA to yield the final product **2**. The 4-cyano intermediate **7** described previously (22) was utilized for the synthesis of the marker molecule **8** by deprotection with trifluoroacetic acid/TIPS (Figure 4B).

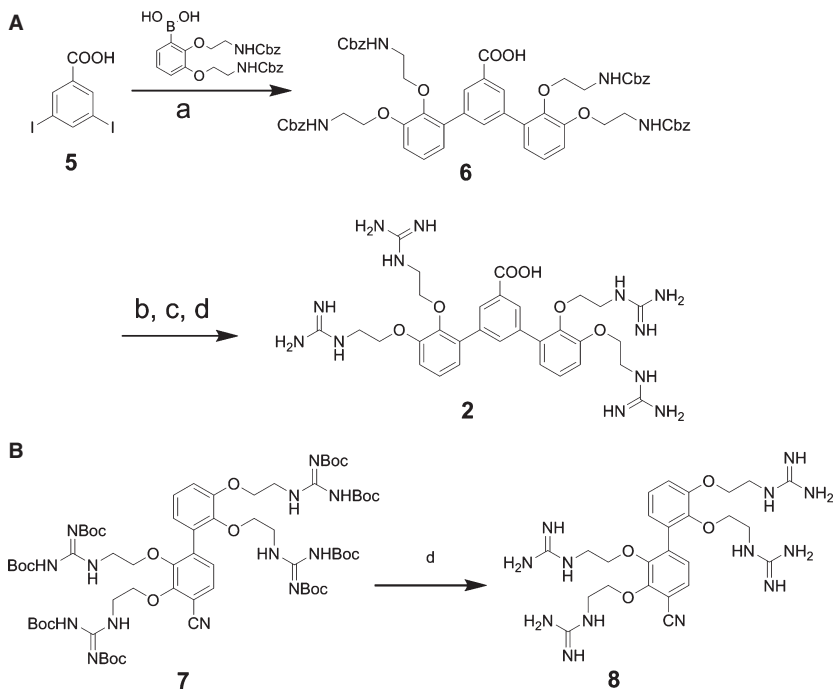


Figure 4: Synthesis of 4G-BfSMoC-COOH (A) and the marker compound 4G-SMoC-CN (B). Reagents: (a) PdCl₂dppf, CH₂Cl₂, Et₃N, iPrOH/H₂O, MW 80 °C (80%), 10 min; (b) 30% HBr in acetic acid, DCM, rt; (c) *N,N'*-Di-Boc-1*H*-pyrazole-1-carboxamide, DIEA, DCM, room temperature, 17 h (63% over 2 steps); (d) TFA/TIPS/H₂O, rt, 4 h (100%).

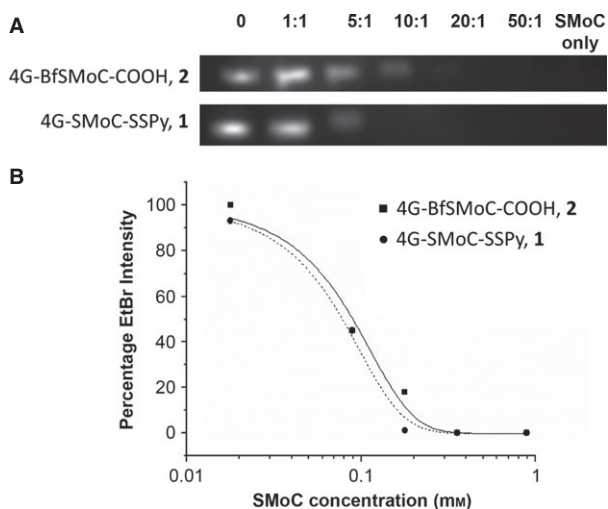


Figure 5: 4G-BfSMoC-COOH, **2** and 4G-SMoC-SSPy, **1** have similar affinity for siRNA. (A) GAPDH siRNA (17 μM) was mixed with **1** or **2** at the molar ratios shown with a loading dye containing ethidium bromide and run on a 1% agarose gel using a 0.4 M MOPS buffer (pH 7.0) containing 0.1 M sodium acetate and 0.01 M EDTA. The gel was visualized under UV illumination. (B) The ethidium bromide intensities as a function of SMoC concentration were fitted in Origin. The EC₅₀ value and Hill coefficient (*P*) were calculated from the binding curve.

siRNA complexation

To confirm that the new SMoC retains the ability to bind siRNA, the binding affinity was assessed using a gel shift assay and compared to our previously tested 4G-SMoC-SSPy, **1** (Figure 5A). The intensity of the siRNA bands detected using ethidium bromide is reduced in a dose-

dependent fashion with the addition of the SMoCs due to neutralization of the negative charge, resulting in immobilization of the siRNA on the gel. The EC₅₀ calculated for **2** was 0.07 mM (Figure 5B), which is equivalent to **1**. This shows that **2** maintains the same level of siRNA-binding capabilities as its predecessor, the maintenance of this interaction is not obvious as the distances and angles between the phenyl groups displaying the guanidines is altered; in addition, the less electron-rich central phenyl ring might alter the strength of π-cation interactions.

siRNA transfection observed using fluorescent siRNA oligomers

To examine the intracellular localization of the 2-enhanced siRNA import, we utilized confocal microscopy with fluorescently labeled siRNA (Figure 6). Quantitative analysis of the relative uptake demonstrates that addition of (10 μM) 4G-BfSMoC-COOH, **2** resulted in 4.9-fold (*P* < 0.01) increased uptake of the oligonucleotide-fluorophore in the 4T1 cells; and a 2.6-fold (*P* < 0.01) increase in the MCF7 cell line compared to siRNA alone. There was undetectable 'background' uptake of fluorescently labeled siRNA alone in the 4T1 cell line (Figure 6A) and non-significant uptake in the MCF-7 cells (Figure 6C). This indicates that there is minimal uptake and retention of uncomplexed oligonucleotide-fluorophore by MCF-7 cells at 24 h. The difference between the two cell lines indicates that this likely to be a cell specific process. Negative controls (without oligonucleotide-fluorophore) showed no fluorescent signal; thus, this background effect is unlikely to be artifact but rather a feature of the uptake kinetics of the specific

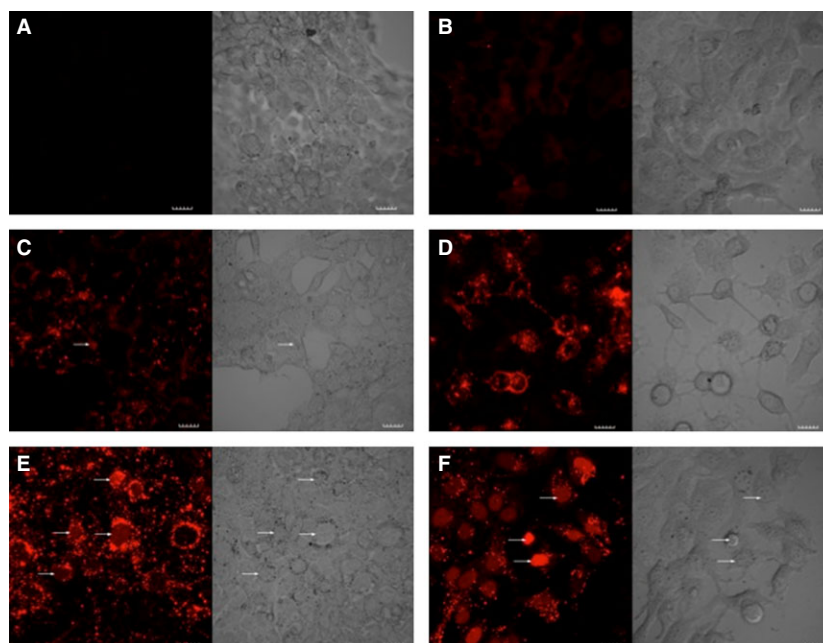


Figure 6: Cellular uptake of fluorescent siRNA. Confocal microscopy of cells treated with AlexaFluor® Red Fluorescent Oligonucleotide \pm 4G-BfSMoC-COOH, **2**. Cells were treated with oligonucleotide-fluorophore alone (A) 4T1 and (B) MCF7; or treated with oligonucleotide-fluorophore plus 4G-Bf-SMoC-COOH (C) 4T1 and (D) MCF-7 or oligonucleotide plus lipofectamine (E) 4T1 and (F) MCF-7 (scale = 20 μ m). Quantitative analysis of the relative cellular uptake demonstrates that addition of the 4G-BfSMoC-COOH, **2** resulted in 4.9-fold ($P < 0.01$) increased uptake of the oligonucleotide-fluorophore (siRNA) in the 4T1 cells; and a 2.6-fold ($P < 0.01$) increase in the MCF7 cell line. There was evidence of nuclear uptake in the 4T1 cell line (Figure 6C, white arrow). In cells treated with Lipofectamine plus siRNA, we observed 14.4-fold and 3.3-fold increase uptake in 4T1 and MCF-7 cell lines, respectively. There was significant nuclear uptake in both cell lines following treatment with Lipofectamine (Figure 6D and E, white arrows). The differing level of background uptake (siRNA alone) is likely due to the differing characteristics of the two cell lines including doubling times and uptake kinetics.

cell lines. Positive controls using lipofectamine demonstrated 14.4-fold ($P < 0.01$) and 3.3-fold ($P < 0.01$) increased uptake compared to siRNA alone in 4T1 and MCF-7 cell lines, respectively. In both cases, there was significant nuclear uptake at 24 h (Figure 6E,F). Other compounds listed in Table 1 either were inactive as reported previously (22) or in the case of the 4G-SMoC-CN, **8** were utilized only as a control.

siRNA transfection using 4G-BfSMoC-COOH, **2**

Both the original 4G-SMoC-SSPy, **1** and the new 4G-BfSMoC-COOH, **2** were used to transfect either *cdc7* siRNA or a negative control siRNA into IMR-90 primary human fibroblasts at a final SMoC concentration of 50 μ M as described previously (22). The commonly used transfection reagent Lipofectamine® was also used as a positive control, and the mRNA knockdown was determined by quantitative real-time PCR (Figure 7). The new 4G-BfSMoC-COOH achieved a knockdown of *cdc7* mRNA of 71% compared to the negative control siRNA, whereas the original 4G-SMoC-SSPy produced a 45% knockdown and was identical to Lipofectamine®, as predicted by our earlier study. However, the difference between the bifurcated SMoC and the original compound was not statistically significant in this study.

Proteoglycan binding of SMoCs compared to polyarginine peptides and HIV-1 tat₄₇₋₅₇

To determine the relative proteoglycan binding of SMoCs, we utilized heparin bound to agarose beads and monitored the retention time when a gradient of sodium chloride (500–1000 mM) was applied to the column. Polyarginine peptides R4SSPy, **9** and R8SSPy, **10** were applied as markers. Selected overlaid traces are shown in Figure 8 and the full dataset in Table 1. The affinities of 2G-SMoC, **12** a 4G-SMoC marker compound bearing a 1-CN group, **8**, and 6G-SMoC, **14** (22) increase in proportion to the number of guanidine groups. 4G-BfSMoC-COOH, **2** shows similar affinity to the linear 3G-SMoC-SSPy, **13** and somewhat less than the 4G-SMoC-SSPy, **1** perhaps due to the carboxylic acid group which is negatively charged at physiological pH. Affinity to heparin was insensitive to the inclusion of the disulfide pyridyl group (see entries for 4G-SMoC-CN, **8** and 4G-SMoC-SSPy, **1**). HIV-1 tat₄₇₋₅₇, **11** showed less affinity than the R8SSPy polyarginine peptide, **10** and similar affinity to 6G-SMoC **14**. Attempts to determine the proteoglycan binding of lipofectamine were unsuccessful using this technique. Proteoglycan binding cannot be assessed as Lipofectamine is a cationic liposome-based reagent, comprising a 3:1 (w / w) liposome formulation of the polycationic lipid 2,3-dioleoyloxy-*N*-[2(sperminecarboxamido)ethyl]-*N,N*-dimethyl-

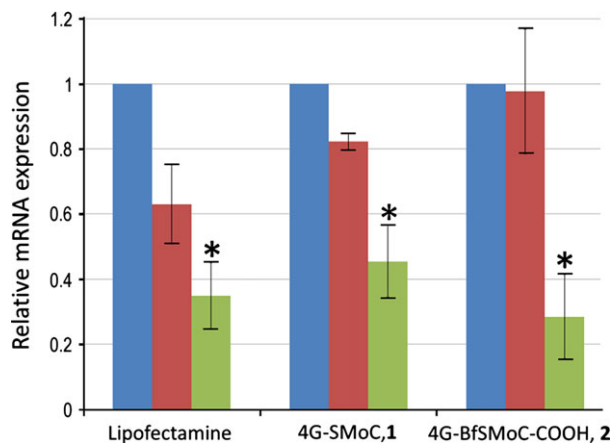


Figure 7: mRNA expression levels in cells treated with 4G-BfSMoC-COOH:siRNA complexes. qRT-PCR data showing *cdc7* mRNA expression in either untreated IMR-90 cells (blue) or cells treated with Lipofectamine[®], **1**, or **2** complexed to control (red) or *cdc7* (green) siRNA. Experiments were conducted in triplicate. *, $P < 0.05$ compared to control siRNA.

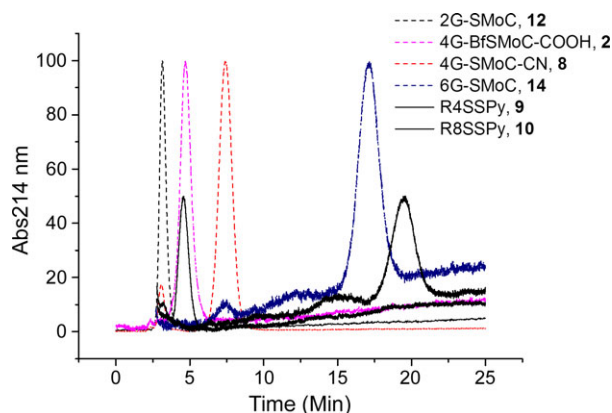


Figure 8: Relative binding to a heparin surface. The SMOcs shown are applied to a HiTrap heparin column in PBS pH 7.4 and eluted with an increasing gradient of NaCl 0.5–1 M. Polyarginine peptides are added as references (half height peaks). A full list of the retention times is shown in Table 1.

1-propanaminium trifluoroacetate (DOSPA) and the neutral lipid dioleoyl phosphatidylethanolamine (DOPE) in membrane-filtered water. The only UV active component is expected to be DOSPA with a single weak amide UV absorbance expected at 214 nm. However, using our standard conditions and monitoring at 214 nm, we observed only a peak at the solvent front even with 10 μ g of Lipofectamine injected (data not shown).

Discussion

The development of a delivery vector that is able to complex and protect siRNA *in vivo* as well as achieve efficient and targeted cellular internalization would be a major

advancement for the therapeutic use of siRNA. Multicomponent siRNA delivery nanoparticles may consist of layers of different materials, such as cationic polymers for siRNA complexation, PEG to increase serum stability, and targeting ligands to confer cell specificity. An efficient siRNA transfection vector should tightly package the siRNA, bind to the cell surface so as to induce macropinocytosis, and finally escape from the endosome into the cytoplasm in order for the siRNA to act on the RNAi machinery and silence the target gene. In this study, we have examined the ability of the SMOc class of transporter to achieve the earlier stages of this process – namely siRNA complexation and cell surface proteoglycan binding.

We previously showed that the protein transduction compound 4G-SMOc-SSPy, **1** was also active as a siRNA delivery vehicle. It was demonstrated that SMOcs were capable of packaging siRNA via electrostatic complexation through the guanidine side chains, thus neutralizing the siRNA negative charges. We also postulated that the biphenyl ring system of 4G-SMOc-SSPy formed π -cation interactions with the guanidine groups, therefore shielding the positive charges during membrane penetration, as predicted for penetratin. In the present study, we develop the structure of the original SMOc to produce a siRNA delivery agent that is able to bind sulfate groups on adjacent cell surface GAGs, is simpler to synthesize and contains a chemical ‘handle’ for potential conjugation.

The new 4G-BfSMoC-COOH, **2** described has been molecularly designed to explore binding to the common GAGs heparan sulfate and chondroitin sulfate, which have been shown to interact with cell-penetrating peptides as a precursor to internalization. In a model system utilizing the more highly sulfated proteoglycan heparin, we observed that the SMOcs bound with similar affinity to the reference tetra and octa-arginine **10**, **11**, and HIV-1 tat peptides. 4G-BfSMoC-COOH, **2** did not show increased binding over linear SMOcs despite the clear indication from the modeling studies that longer distances might be beneficial for binding and the molecular dynamics studies indicated that the larger structures were able to adopt more extended conformations. We should be cautious, however, in that cellular binding may differ from our model heparin system. Several studies have suggested that the binding of GAG by CPPs plays a role in cell penetration, with subsequent receptor clustering leading to reorganization of the actin cytoskeleton as a precursor to macropinocytosis (18,20,27–30). Heparin has been used in binding assays to determine GAG-binding affinities for several CPPs, which are found to bind in the low micromolar to nanomolar range (19), probably via bidentate hydrogen bonds between arginine residues and sulfate groups. Recently a report from the Dowdy laboratory challenged the importance of GAG binding for CPP internalization (31), and the uptake of other peptides has been shown to be independent of GAG binding (32,33). Further investigations will be required to ascertain the exact influence of GAG binding

on the SMOc class of biomolecular transporter. The new SMOc described here was easily synthesized, with the trifluoroborate salt used in a Suzuki-Miyaura bicoupling step to form the terphenyl structure. The success of this coupling step demonstrates that the trifluoroborate salt may be used to generate larger branched SMOc analogues from simple starting materials. The bifurcated SMOc does not show increased uptake compared to the original 4G-SMOc-SSPy, **1** reagent or to Lipofectamine®; however, it does demonstrate that the disulfide pyridyl group, and therefore potential reaction with cell surface thiols, is not required for successful siRNA transfection. Further optimization of this system is in progress.

Acknowledgments

We thank the Association for International Cancer Research (AICR) for funding this project (Grant Ref: 06-0076) and acknowledge the UCL Grand Challenge PhD scholarship for an award to DA.

Conflict of Interest

The authors declare no conflict of interest.

References

- Leachman S.A., Hickerson R.P., Schwartz M.E., Bullough E.E., Hutcherson S.L., Boucher K.M., *et al.* (2010) First-in-human mutation-targeted siRNA phase Ib trial of an inherited skin disorder. *Mol Ther*;18:442–446.
- Barakat M.R., Kaiser P. (2009) VEGF inhibitors for the treatment of neovascular age-related macular degeneration. *Expert Opin Investig Drugs*;18:637–646.
- Gooding M., Browne L.P., Quinteiro F.M., Selwood D.L. (2012) siRNA delivery: from lipids to cell-penetrating peptides and their mimics. *Chem Biol Drug Des*;80:787–809.
- Guo J., Fisher K.A., Darcy R., Cryan J.F., O'Driscoll C. (2010) Therapeutic targeting in the silent era: advances in non-viral siRNA delivery. *Mol BioSyst*;6:1143–1161.
- Zimmermann T.S., Lee A.C.H., Akinc A., Bramlage B., Bumcrot D., Fedoruk M.N. *et al.* (2006) RNAi-mediated gene silencing in non-human primates. *Nature*;441:111–114.
- Semple S.C., Akinc A., Chen J., Sandhu A.P., Mui B.L., Cho C.K. *et al.* (2010) Rational design of cationic lipids for siRNA delivery. *Nat Biotechnol*;28:172–176.
- Vaishnav A.K., Cervantes A., Alsina M., Tabernero J., Infante J.R., LoRusso P. *et al.* (2011) RNAi in humans: phase I dose-escalation study of ALN-VSP02, a novel RNAi therapeutic for solid tumours with liver involvement. *Nucl Acid Ther*;21:A44–A.
- Coelho T., Suhr O.B., Adams D., Lozeron P., Hawkins P., Mant T. *et al.* (2011) Interim clinical update for ALN-TTR01, a novel RNAi therapeutic for the treatment of transthyretin amyloidosis. *J Peripher Nerv Syst*;16: S25–S26.
- Davis M.E., Zuckerman J.E., Choi C.H.J., Seligson D., Tolcher A., Alabi C.A. *et al.* (2010) Evidence of RNAi in humans from systemically administered siRNA via targeted nanoparticles. *Nature*;464:1067–1070.
- Davis M.E. (2009) The first targeted delivery of siRNA in humans via a self-assembling, cyclodextrin polymer-based nanoparticle: from concept to clinic. *Mol Pharm*;6:659–668.
- Veronese F.M., Pasut G. (2005) PEGylation, successful approach to drug delivery. *Drug Discov Today*;10: 1451–1458.
- Meister G., Tuschl T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*;431:343–349.
- Kim S.H., Jeong J.H., Lee S.H., Kim S.W., Park T.G. (2008) Local and systemic delivery of VEGF siRNA using polyelectrolyte complex micelles for effective treatment of cancer. *J Control Release*;129:107–116.
- Li W., Szoka F.C. (2007) Lipid-based nanoparticles for nucleic acid delivery. *Pharm Res*;24:438–449.
- Moschos S.A., Jones S.W., Perry M.M., Williams A.E., Erjefalt J.S., Turner J.J. *et al.* (2007) Lung delivery studies using siRNA conjugated to TAT(48-60) and penetratin reveal peptide induced reduction in gene expression and induction of innate immunity. *Bioconjug Chem*;18:1450–1459.
- Won Y.-W., Kim H.A., Lee M., Kim Y.-H. (2010) Reducible poly(oligo-D-arginine) for enhanced gene expression in mouse lung by intratracheal injection. *Mol Ther*;18:734–742.
- Crombez L., Morris M.C., Dufort S., Aldrian-Herrada G., Nguyen Q., McMaster G. *et al.* (2009) Targeting cyclin B1 through peptide-based delivery of siRNA prevents tumour growth. *Nucleic Acids Res*;37:4559–4569.
- Nakase I., Tadokoro A., Kawabata N., Takeuchi T., Katoh H., Hiramoto K. *et al.* (2006) Interaction of arginine-rich peptides with membrane-associated proteoglycans is crucial for induction of actin organization and macropinocytosis†. *Biochemistry*;46:492–501.
- Ziegler A. (2008) Thermodynamic studies and binding mechanisms of cell-penetrating peptides with lipids and glycosaminoglycans. *Adv Drug Deliv Rev*;60:580–597.
- Fuchs S.M., Raines R.T. (2004) Pathway for polyarginine entry into mammalian cells†. *Biochemistry*;43: 2438–2444.
- Rullo A., Qian J., Nitz M. (2011) Peptide-glycosaminoglycan cluster formation involving cell penetrating peptides. *Biopolymers*;95:722–731.
- Gooding M., Tudzarova S., Worthington R.J., Kingsbury S.R., Rebstock A.S., Dube H. *et al.* (2012) Exploring the interaction between siRNA and the



- SMoC biomolecule transporters: implications for small molecule-mediated delivery of siRNA. *Chem Biol Drug Des*;79:9–21.
23. Okuyama M., Laman H., Kingsbury S.R., Visintin C., Leo E., Eward K.L. *et al.* (2007) Small-molecule mimics of an [alpha]-helix for efficient transport of proteins into cells. *Nat Methods*;4:153–159.
 24. Rebstock A.S., Visintin C., Leo E., Posada C.G., Kingsbury S.R., Williams G.H. *et al.* (2008) Modular assembly using sequential palladium coupling gives easy access to the SMoC class of cellular transporters. *ChemBioChem*;9:1787–1796.
 25. Krieger E., Koraimann G., Vriend G. (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins*;47:393–402.
 26. Ziegler A., Seelig J. (2008) Binding and clustering of glycosaminoglycans: a common property of mono- and multivalent cell-penetrating compounds. *Biophys J*;94:2142–2149.
 27. Simon M.J., Gao S., Kang W.H., Banta S., Morrison B. (2009) TAT-mediated intracellular protein delivery to primary brain cells is dependent on glycosaminoglycan expression. *Biotechnol Bioeng*;104:10–19.
 28. Patel L.N., Zaro J.L., Shen W.C. (2007) Cell penetrating peptides: intracellular pathways and pharmaceutical perspectives. *Pharm Res*;24:1977–1992.
 29. Gerbal-chaloin S., Gondeau C., Aldrian-herrada G., Heitz F., Gauthier-rouvière C., Divita G. (2007) First step of the cell-penetrating peptide mechanism involves Rac1 GTPase-dependent actin-network remodelling. *Biol Cell*;99:223–238.
 30. Couchman J.R. (2003) Syndecans: proteoglycan regulators of cell-surface microdomains? *Nat Rev Mol Cell Biol*;4:926–938.
 31. Gump J.M., June R.K., Dowdy S.F. (2010) Revised role of glycosaminoglycans in TAT protein transduction domain-mediated cellular transduction. *J Biol Chem*;285:1500–1507.
 32. Lépinoux-Chambaud C., Eyer J. (2013) The NFL-TBS.40-63 anti-glioblastoma peptide enters selectively in glioma cells by endocytosis. *Int J Pharm*;454:738–747.
 33. Verdurmen W.P.R., Wallbrecher R., Schmidt S., Eilander J., Bovee-Geurts P. Fanghänel S., *et al.* (2013) Cell surface clustering of heparan sulfate proteoglycans by amphipathic cell-penetrating peptides does not contribute to uptake. *J Control Release*;170:83–91.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Naming convention for SMoCs.

Table S1 Following a request from a reviewer we have included all the ¹H NMRs for the compounds used in this study.