Electronic Supplementary Material (ESI) for Chemical Science.

Supplementary Information

Designer artificial membrane binding proteins to direct stem cells to the myocardium

Wenjin Xiao [a†], Thomas. I. P. Green [ab†], Xiaowen Liang [c†], Rosalia Cuahtecontzi Delint [ab], Guillaume Perry [d], Michael. S. Roberts [ce], Kristian le Vay [ab], Catherine. R. Back [fg], Raimomdo Ascione [h], Haolu Wang [c], Paul. R. Race [fg], Adam. W. Perriman*[ag]

Affiliations:

[a] School of Cellular and Molecular Medicine, University of Bristol, BS8 1TD, UK.

[b] Bristol Centre for Functional Nanomaterials, University of Bristol, BS8 1FD, UK.

[c] Therapeutics Research Centre, The University of Queensland Diamantina Institute, The University of Queensland, Translational Research Institute, Woolloongabba, QLD 4102, Australia.

[d] Sorbonne Université, Laboratoire d'Electronique et d'Electromagnétisme, L2E, F-75005, Paris, France.

[e] School of Pharmacy and Medical Science, University of South Australia, Adelaide, SA 5001, Australia.

[f] School of Biochemistry, University of Bristol, BS8 1TD, UK.

[g] BrisSynBio Synthetic Biology Research Centre, University of Bristol, BS8 1TQ, UK

[h] Translational Biomedical Research Centre and Bristol Heart Institute, University of Bristol, Bristol, UK.

Table of Contents

Methods and Materials

Gene synthesis, amplification and purification

The fusion gene of CshA_scGFP was synthesized by Sigma (UK). Amplification and expression was carried out using pOPIN F vector (Oxford Protein Production Facility, UK), which automatically adds a cleavable N-terminal Histidine-tag. Primers were designed to be complementary for the synthetic gene but also incorporate a 15 bp overhang for pOPIN F and thus enable compatibility with the in-fusion enzyme system. Amplification of the synthetic gene was carried out using the polymerase chain reaction (PCR) in 50 µL reaction mixes using Labtaq Hi-Fidelity Polymerase™ enzyme available from LabTech™. 1 µL of template DNA was added to 1 µL of both the forward and reverse primers, as well as 1 µL of the enzyme plus 10 µL of the buffer. 36 µL of Milli-Q water was added to the reaction mixture to bring the overall volume to 50 µL in the PCR tubes. The reaction vessel was then added to a temperature cycler and the following protocol was used: 95 °C for 5 minutes, 95 °C for 30 seconds, 62 °C for 30 seconds, 72 °C for 3 minutes, repeat step 2 to 4 for 30 cycles, 72 °C for 10 minutes, incubate at 4 °C. Agarose gel electrophoresis was used to separate and purify the PCR products, and to check the molecular weight of the fragments. QIAQuick gel extraction kit (Qiagen, USA) was used subsequently to purify the DNA.

Plasmid linearization

The pOPIN F vector pellet was resuspended in nuclease-free water up to a concentration of 100 µg/mL before it was linearized for subsequent genetic insertion. To a PCR tube, 10 µL of plasmid was added to 1 unit of HindIII-HF (New England Biosciences, USA), 1 unit of KpnI-HF (New England Biosciences, USA) and 5 µL of 10 x cut smart buffer. Nuclease free water was used to bring the final volume of the mixture to 50 µL. The sample was then incubated at 37 ^{o}C for 1 hour before 10 µL of 6 x gel loading buffer was added and the sample subjected to agarose gel electrophoresis to remove unreacted plasmid and enzymes.

Gene insertion into pOPIN F vector

The In-Fusion^{TM} ligation independent system was used to insert the fusion gene into the linearized plasmid. 10 µL reaction volumes containing 2 µL of In-Fusion^{TM} master mix, 4 µL of the purified PCR fragment, 2 µL of pOPIN F and 2 µL of water were added to micro centrifuge tubes and kept on ice.

After 10 minutes of low temperature equilibration, the reaction vessel was placed into a water bath at 50 °C for 15 minutes and then immediately returned to the ice.

Chemically competent cells transformation

Chemically competent Stellar™ cells (Clontech™ Laboratories, UK) were used to amplify the ligated plasmids and repair any nicks leftover in the insertion process. 2 µL of the ligation reaction mixture was added to 30 µL of Stellar™ cells in a micro centrifuge tube and kept on ice for 30 minutes to equilibrate. Next the reaction vessel was placed in a water bath at 42 °C for 45 seconds to stimulate uptake of the plasmids by the chemically competent cells and then immediately returned to ice for further 5 minutes. 1 mL of Super Optimal Broth with Catabolite Suppression (SOC) media was then added and incubated at 37 °C for 1 hour.

Protein expression

Single white colonies from the overnight LB/agar cultures were selected and added to 10 mL of LB with 10 µL of 50 mg/mL carbenicillin and left overnight in a rotating incubator at 37 °C. The resulting culture was then separated into 1 mL aliquots in micro centrifuge tubes and spun at 13,000 g to pellet the cells. The supernatant was then removed and a Miniprep™ kit (Qiagen, USA) was used to retrieve the plasmids according to manufacturer instructions. Final concentrations of plasmid DNA extracted from this method could then be analyzed using a Nanodrop™ spectrophotometer (Thermo Fisher Scientific, USA). The plasmid DNA was next transformed to BL21 DE3 competent Escherichia coli (*E. coli*) (New England Biolabs, USA). 50 µL of the cells were thawed on ice before up to 100 ng of plasmid DNA was added. After equilibration on ice for 30 minutes the mixture was heat shocked at 42 °C for 10 seconds and placed on ice again for 5 minutes. 950 µL of SOC media was added and the mixture was incubated in a rotating incubator at 37 °C and 200 RPM for 1 hour. The transformed *E. coli* was subsequently transferred to an autoclaved 20 mL LB containing 20 µL of 50 mg/mL carbenicillin and placed in a rotating incubator overnight at 180 RPM and at 37 °C. The resulting starter culture was added to 1 L of LB with 1 mL of 50 mg/mL carbenicillin and rotated at 180 RPM and at 37 °C until the optical density (OD) at λ=600 nm was around 1 using an UV-vis light spectrometer (Agilent technologies, UK). Protein production was then induced by the addition of 1 mL of 1 M isopropylthiogalactosidase (IPTG) (Apollo Scientific, Japan), before being incubated overnight at 180

RPM and at 25 °C. The 1 L cultures were then centrifuged at 4500 g for 30 minutes using a Sorvall RC6 Centrifuge (Thermo Fisher Scientific, USA) to form cell pellets.

Protein purification

The pellet was re-suspended in 50 mL of immobilized metal ion affinity chromatography (IMAC) lysis buffer, which contained 20 mM Tris-HCl, 1 M NaCl and 20 mM imidazole adjusted to pH 7.5. 100 µL of 100 mM phenylmethanesulphonylfluoride (PMSF) (Thermo Fisher Scientific, USA) was added to prevent proteolysis of the target cell, whilst DNase (Sigma Aldrich, UK) was used to prevent DNA aggregation during the cell lysis step. The cells were lysed using a VCX 500 ultrasonic processor (Sonics, USA) to release the contents of the cells into solution. In order to remove any cell debris, the cell lysate was next centrifuged at 20,000 g for 45 minutes. The supernatant was loaded onto an IMAC column, HisTrap FF column (GE Healthcare Life Sciences, USA), with nickel as the immobile phase. The column was washed with lysis buffer and elution was carried out over a gradient up to 500 mM imidazole using an ÄKTApurifier™ (GE Healthcare Life Sciences, USA). The fractions from the IMAC were concentrated using 30 kDa molecular weight cutoff (MWCO) spin concentrators (Millipore, USA) and the resulting sample was applied to size exclusion chromatography (SEC) to further separate molecules by their sizes using a HiLoad 26/600 s200 pg (GE Healthcare Life Sciences, USA) with the running buffer (20mM Tris-HCl, 1 M NaCl, pH 7.5).

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

10 µL of analyte was added to 10 µL of sample application buffer comprised of glycerol, SDS, EDTA, Tris, mercaptoethanol and bromophenol blue. This was then heated to 95 °C for 5 minutes to denature the protein before application to Novex® 4-20% Tris-Glycine precast protein gels (Thermo Fisher Scientific, USA). A voltage of 200 V was applied for 1 hour before staining with Coomassie Blue™ protein stain. Finally, non-protein containing regions were destained to reveal protein bands with a solution of methanol, acetic acid and deionized water made to a ratio of 1:4:5. PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA), 10 to 180 kDa, was used as a reference.

Mass Spectrometry

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) (Bruker, UK) was used to measure the molecular mass of the fusion protein CshA_scGFP in linear

positive mode. The analyte was mixed with a matrix of sinapinic acid at a 1:1volume ratio. 2 µL of aliquots were deposited onto a ground steel plate and allowed to dry at room temperature until a yellow crystalline deposit was formed. Laser ablation with an ultraviolet wavelength excited the matrix compounds resulting in an array of gas phase ions from the matrix and sample. Multiple spectra were additively acquired between 4 kDa and 140 kDa using FLEX Control software (Bruker, USA) and peaks were identified using FLEX Analysis software (Bruker, USA).

Bicinchoninic acid assay

A bicinchoninic acid (BCA) assay was used to measure the protein concentration. 25 µL protein were added in triplicate to a clear 96-well plate (Corning, USA). Bovine serum albumin (BSA) (Sigma Aldrich, UK) was used for a standard curve, with triplicate 25 µL aliquots of 0, 0.025, 0.05, 0.1, 0.125, 0.25, 0.4, 0.5, 0.75, 1.0, 1.5, 2.0 mg/mL. 200 µL of working reagent including 50:1 (v/v) ratio of reagent A and B (Thermo Scientific, UK) was next added to each well. The plate was incubated at 37 °C for 30 minutes and then left to cool down at room temperature for 5 minutes. A Mithras LB940 plate reader (Berthold Technologies, USA). Was used to measure the absorbance at λ=530 nm. The concentration of the protein was calculated using the standard curve. Subsequently the Beer-Lambert law was used to calculate the experimental extinction coefficients of protein at 280 nm (ε 280) and 487 nm (ε 487) using the protein concentration (C), the path length (L) and the absorbance intensity at λ=280 nm (Abs_{280}) and λ=487 nm (Abs_{487}), respectively. The theoretical extinction coefficient of CshA_scGFP at 280 nm was obtained from ExPASy (Swiss) according to the amino acid sequence.

Oxidisation of the polymer surfactant

The glycolic acid ethoxylate 4-nonylphenyl ether surfactant was oxidised as previously described [1]. 2 g of IGEPAL® CO-890 (Sigma-Aldrich, UK) was dissolved in 50 mL of deionized water before adding 5 mL of a sodium hypochlorite solution (Sigma-Aldrich, UK), 30 mg of 2,2,6,6,-tetramethyl-1-piperidinyloxyl (TEMPO) (Sigma-Aldrich, UK), and 50 mg of sodium bromide (Sigma-Aldrich, UK). The reaction mixture was stirred for overnight at room temperature with periodic adjustments to maintain a pH of 11, and then quenched with 10 mL of ethanol and adjusted to pH 1. 80 mL of chloroform (Sigma-Aldrich, UK) were used for solvent extraction for 3 times followed by 3 washes with 80 mL of deionized water adjusted to pH 1. The resulting solution was dried at 42 °C under reduced pressure using a rotary evaporator, before dissolving in ethanol at 60 °C and recrystallizing at -20 °C. The

precipitate was harvested using vacuum filtration, redissolved in ethanol and dried under reduced pressure. The extinction coefficient of surfactant was calculated at a concentration of 2.5 mg/mL. Quantitative oxidation of the polymer surfactant was confirmed using H_1 NMR and FTIR.

NMR Spectroscopy

H_1 NMR spectroscopy measurements were carried out to analyse the constitution of the polymer surfactant. Samples were dissolved in the buffer (20mM Tris-HCl, 1 M NaCl, pH 7.5) at a concentration of 10 mg/mL and loaded into quartz NMR tubes. Proton NMR data was collected on a Varian 400 NMR spectrometer (Agilent technologies, UK) at 25°C.

Fourier transform fnfra-red (FTIR) spectroscopy

FTIR spectroscopy measurements were carried out to monitor the oxidation of the polymer surfactant. Samples were scanned from wavenumber $\bar{\upsilon}$ = 550–4000 cm^{-1} with 15 accumulations for each sample using a FTIR spectrometer (Perkin Elmer, UK) in transmission mode. A blank measurement was taken as a background.

Synthesis of the protein-polymer surfactant complexes

A 25 mg/mL surfactant solution, dissolved in 20 mM of Tris-HCl at pH 7.5, was added to the protein solutions at protein cationic sites:surfactant chains at stoichiometries of 1:1.4, 1:2.1, and 1:2.8, respectively, and left to stir overnight at 4 °C. The resulting solution was filtered through a 0.22 µm syringe filter and dialyzed into excess buffer to remove any unbound surfactant molecules. The molar extinction coefficients of protein and polymer surfactant were used to evaluate the final protein cation charge:polymer surfactant stoichiometries.

Sedimentation velocity analytical ultracentrifugation (SV-AUC)

SV-AUC experiments were performed on a Beckman Optima XL-I (Beckman Coulter, USA). The sedimentation of protein constructs was monitored at 40,000 RPM and at 20°C using the UV-vis absorption system at wavelengths of 280 and 487 nm. Samples containing only surfactant were measured at 50,000 RPM and at 20°C using interference optics and the UV-vis absorption system at a wavelength of 280 nm. Sedimentation coefficients (S) were fitted in SEDFIT using the continuous distribution Lamm equation model c(S). The partial specific volume of CshA_scGFP FLRR

(0.7263 cm^3/g) was calculated from the primary sequence using SEDFIT. The density and viscosity of buffer were determined using a Lovis 2000 rolling ball viscometer (Anton Paar, Austria) and found to be 1.02 $g \cdot cm^3$ and 1.05 mPa.s respectively.

Dynamic light scattering (DLS)

Particle size and zeta potential were measured using a ZetaSizer Nano ZS (Malvern Instruments, UK) and analyzed using ZetaSizer software (Malvern Instruments, UK). Samples were diluted in 10 mM of the corresponding buffer and filtered using a 0.22 µm syringe filter. Particle size was measured at 25 °C with the backscatter measured at 173° after a 300 second equilibration, and Zeta potential was measured at 25 °C with 10-100 runs per sample after a 300 second equilibration.

Synchrotron radiation small angle x-ray scattering (SR-SAXS)

SR-SAXS experiments were carried out on beamline B21 at the Diamond Light Source synchrotron (UK) with a fixed camera length configuration 4.014 meters with a beam energy of 12.4 keV and a cross section 1x 5 mm. 1 mL of the samples were loaded onto a superdex SEC column at 7 mg/mL and analysed to allow for the maximum purity before SAXS. The 2D scattering profiles were reduced using in-house software. The data were scaled, merged and background-subtracted using the ScÅtter software package, which was also used to measure the Porod exponent [2]. GNOM and BAYESAPP were used to generate pair distance distribution (P(r)) plots from the scattering curves [3,4]. All SAS parameters were derived from real space distributions rather than reciprocal space fitting to mitigate the effects of aggregation.

Small angle neutron scattering (SANS)

SANS experiments were carried out on beamline D17 at the Institut Laue-Langevin (ILL) (France), with the aid of Dr. Anne Martel. Samples were loaded into quartz cuvettes with a pathlength of 1 mm at a protein concentration of 0.7 to 0.6 mg/mL in 100% D_2O and 40:60 D_2O:H_2O for CshA_scGFP and [CshA_scGFP][S], respectively. Time-of-flight (TOF) measurements were made at a wavelength of 5.5 Å using camera lengths of 1.4 m, 5.6 m and 17.6 m to capture the entire q range. The scatter patterns were radially averaged from the beam center at each detector distance, and then merged using the NCNR macro on Igor Pro. After subsequent background subtraction, the curves were analyzed using BAYESAPP to produce SANS invariants and P(r) distributions [5–8].

Far-UV synchrotron radiation circular dichroism (SRCD) spectroscopy

Far-UV SRCD spectroscopy measurements were carried out on beamline B23 using Global Works software (Olis, USA) at Diamond Light Source (UK). Samples were scanned between λ=185-260 nm in 3 nm increments with 120 second integration times. The temperature was maintained at 37 °C using a Peltier Temperature controller (Quantum, USA). Prior to each scan, the temperature was equilibrated for 3 minutes. A background sample of the solvent was subtracted from the sample spectrum. The far-UV ellipticity data was deconvoluted into secondary structure fractions using BeStSel software, whilst the predicted data was approximated from pdb files analysed using the DSSP algorithm. The DSSP does not differentiate between beta structures so the predicted antiparallel beta of the prediction was compared with the total beta structure obtained using BeStSel.

Fluorescence spectroscopy

A Cary Eclipse (Agilent technologies, UK) was used to measure the fluorescence spectra. Both the excitation and the emission slit width were set at 5 nm. The excitation and emission data were collected over a wavelength range of λ=300-650 nm, respectively. All data was normalized by the sample concentrations.

Human mesenchymal stem cells (hMSCs) culture and cell membrane engineering

hMSCs were isolated, characterized, cultured and passaged as previously described [1]. hMSCs were primed using either the fusion protein CshA_scGFP or conjugate [CshA_scGFP][S], respectively. Expansion media were aspirated, and the cells were washed with Dulbecco's Phosphate Buffered Saline (PBS) (Sigma-Aldrich, UK) and suspended with trypsin/ EDTA solution (Sigma-Aldrich, UK). A solution of the constructs in phenol-free DMEM (Sigma-Aldrich, UK) was added to the suspended cells and left shaking at 50 RPM and at 37 °C for 15 minutes in the incubator. The cells were next washed with phenol-free DMEM and centrifuged at 1500 RPM for 5 minutes to remove any unbound constructs. The primed cells were next returned to culture or analyzed at 37 °C.

Cytotoxicity assays

An (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) (MTS) assay was used to measure the total metabolic activity of the samples as an estimate of the cytotoxicity of the constructs. 5000 hMSCs were plated in duplicate in 96 well plates and

cultured at 37 °C for 6 hours to allow adherence. Standard curve samples with 0, 250, 500, 1000, 2500, 5000, 7500 and 10,000 cells per well were established under identical conditions. The cells were primed with either CshA_scGFP or [CshA_scGFP][S], respectively, at concentrations of 0, 2, 4, 6, 8, 10, 12 and 16 μM for 15 minutes and washed with PBS, and then 100 μL of phenol-free expansion medium along with 20 μL of CellTiter 96 solution (Promega, USA) was added to each well. The cells were incubated at 37 °C for 2 hours, and then 100 μL of supernatant was transferred to a 96 well plate and the absorbance at 490 nm was measured using a plate reader (Mithras LB940 plate reader, Berthold Technologies, USA). The number of viable cells was calculated from the standard curve.

Cell proliferation assays

Cell counting assay was used to examine the effect of the constructs on the proliferative capacity of hMSCs. hMSCs were primed with 4 μM of either CshA_scGFP or [CshA_scGFP][S], respectively, and the cells were harvested and 100,000 cells were seeded in individual T25 flasks in duplicate and cultured for 7 days. The cells were harvested and counted using a haemocytometer, and the process was repeated for another 7 days as a second passage. All proliferation data was compared to a control group of cells incubated with phenol-free expansion medium.

Live cell widefield microscopy

All widefield microscopy was carried out at the Wolfson Bioimaging Facility (University of Bristol). A Leica LASX live cell imaging workstation with Leica DFC365FX monochrome CCD camera attached to a DMI 6000 inverted epifluorescence microscope (Leica, UK) was used for live cell imaging. 50,000 cells were plated in issue culture-treated 6 well plates. After 24 hours of culture, the cells were primed with 4 μM of either CshA_scGFP or [CshA_scGFP][S], respectively, and stained with Hoechst (Life Technologies, USA), then washed with PBS and transferred to phenol-free expansion media. The excitation filters used were GFP fluorophore (480/40 nm) and DAPI fluorophore (350/50 nm) for cells labeled with the constructs and Hoechst, respectively. Images were captured using Leica LAS-X acquisition software (Leica, UK) and processed by Fiji software.

Cell membrane uptake quantification

50,000 hMSCs were plated in duplicate in 24 well plates and cultured at 37 °C for 6 hours to allow adherence. The cells were primed with either CshA_scGFP or [CshA_scGFP][S], respectively,

at concentrations of 0, 2, 2, 4, 6, 8, 10 and 12 μM ($Conc_{control}$). The supernatant was removed and analyzed using UV-vis spectroscopy at the wavelength of 487 nm ($Abs_{487\ sample}$). The concentration of constructs delivered to the cell the surface $P_{[cell]}$ was calculated at each individual concentration (**Eq. 1**). The number of constructs per cell (N_{cell}) was calculated by the incubation volume of constructs (V), the Avogadro constant (N_A) and the number of cells per well (Q) (**Eq. 2**) [1].

$$P_{[cell]} = Conc._{control} - \frac{Abs_{487\ sample}}{Extinction\ Cofficient} \qquad \textbf{Eq. 1}$$

$$N_{cell} = \frac{[P_{cell}]VN_A}{Q} \qquad \textbf{Eq. 2}$$

Flow cytometry studies

hMSCs primed with 4 μM of either CshA_scGFP or [CshA_scGFP][S], respectively, were re-suspended in PBS containing 2.5 mM EDTA (Sigma-Aldrich, UK) and 2% (v/v) FBS, and the number of harvested cells was counted. Suspensions containing approximately 1,000,000 cells per mL and 0.004 mg/mL of propidium iodide (PI) were analyzed using a NovoCyte flow cytometer (ACEA Biosciences, USA) with NovoExpress software (ACEA Biosciences, USA). The cell suspension was measured at a rate of 100-300 events per second and data was analyzed using NovoExpress software. The hMSC population was firstly defined by forward scatter area (FSC-A) vs the side scatter area (SSC-A), with data outside this area considered as cell debris. The hMSC population was next gated by FSC-A vs forward scatter height (FSC-H) in order to identify the single cells. Following this, the single cell population was gated by setting an upper limit on the FSC-A vs propidium iodide (PI) fluorescence (Qdot 605-A), with data below this limit considered as viable cells. The live cells were finally gated by FSC-A vs FITC-A with a total of 20,000 cell events. The data inside the area was considered as scGFP positive cells, and the mean fluorescence intensity was obtained from the FITC-A channel. These defined gated areas were used for all samples and untreated hMSCs were used as a control group.

The *in vitro* cardiogenic differentiation of hMSCs

5,000 hMSCs were plated on a coverslip into each well of 24 well plates and allowed to attach for 24 hours. The cells were then primed with 4 μM of [CshA_scGFP][S] and washed twice with PBS. The

cardiogenic differentiation of hMSCs was induced by adding 5 μM of 5-Azacytidine in the fibroblast growth factor free culture medium for 48h. The control groups are incubated with the hMSC complete medium. The cells were cultured for 2 weeks with medium change every other day.

Immunofluorescence staining

Cells on a coverslip were fixed in 4% (w/v) paraformaldehyde (PFA) in PBS for 15 minutes at room temperature and washed 3 times using ice-cold PBS. Then the cells were permeated in 0.1% (w/v) saponin in PBS for 15 minutes at room temperature and washed 3 times using PBS for 5 minutes. Next, the cells were blocked in 2% (w/v) BSA, 22.52 mg/mL glycine and 0.05% (w/v) saponin in 0.1% (v/v) Tween-20 PBS for 30 minutes at room temperature. Cells were incubated with anti-Nkx2.5 antibody (mouse monoclonal antibody, 1:100 dilution, Abcam), anti-alpha cardiac actin antibody (rabbit polyclonal antibody, 1:200 dilution, Thermo Fisher Scientific), or anti-cardiac troponin T antibody (mouse monoclonal antibody conjugated with Alexa Fluor® 647, 1:100 dilution, BDbiosciences), respectively, overnight at 4°C, followed by incubation with secondary antibody Alexa Fluor® 568-conjugated goat anti-mouse IgG (1:200 dilution, Abcam) for anti-Nkx2.5 antibody, and Alexa Fluor® 568-conjugated goat anti-rabbit IgG (1:200, Abcam) for anti-alpha cardiac actin antibody, respectively, for 1 hour at room temperature in the dark. Cell nuclei were stained with Hoechst. The cells were washed 3 times with PBS for 5 minutes in the dark and coverslips were mounted. Cells were imaged with confocal microscopy (Leica SP8 AOBS confocal laser scanning microscope attached to a DMI6000 inverted epifluorescence microscope (Leica, UK)). Images were captured using Leica LAS-X acquisition software (Leica, UK) and processed by Fiji software.

Static adhesion assays

Cell-substrate adhesion mediated by immobilized human fibronectin (Fn) was investigated using a CyQUANT® NF cell proliferation assay kit (Invitrogen, UK). Human Fn (Sigma-Aldrich, UK) was diluted to 10 μg/mL with PBS and 100 μL of Fn was used to coat each well of the non-tissue culture-treated 96 well plate. The plates were placed at room temperature for 3 hours then washed 3 times with PBS solution containing 10 mg/mL BSA to block the unspecific bonds. The non-Fn coated wells blocked with BSA were used as a control. hMSCs were primed with 4 μM of either CshA_scGFP or [CshA_scGFP][S], respectively, and the cells were harvested and counted using a haemocytometer. 10,000 cells from each sample were seeded into each well with phenol-free DMEM and

allowed to attach for 6 hours. Standard curve samples were established in expansion medium with 0, 100, 500, 1000, 2500, 5000, 10,000, 15,000 and 20,000 cells per well. Medium was removed from cells by gentle aspiration and 100 µL of dye binding solution was dispensed into each well. The plate was covered and incubated at 37 °C for 1 hour. The fluorescence intensity of each sample was measured using a Mithras LB940 plate reader (Berthold Technologies, USA) with excitation at 485 nm and emission detection at 530 nm. All adhesive cell number was compared to control samples of untreated cells incubated with phenol-free DMEM.

Fluidic adhesion assays

Cellix Vena8 Fluoro+™ Biochip capillaries (Cellix, Ireland) were coated overnight with 50 ug/mL Fn in PBS at 4 °C followed by bovine serum albumin (BSA) blocking the unspecific bindings. hMSCs were flowed over at variable physiological-relevant shear stress of 5, 2.5, 2, 1.5, and 1 dyne/cm ^{9}. 10 seconds of real-time video (bright field) was recorded using widefield microscopy at 113 ms/frame for 5 different positions along the channel, and the adherent cell number was analyzed using Fiji software. All widefield microscopy was carried out at the Wolfson Bioimaging Facility (University of Bristol, UK).

***In vivo* transplantation of hMSCs**

Male 20-week-old FVB/N and BALB/c nude mice were purchased from the Animal Resource Centre (Perth, Western Australia). All animal procedures were carried out in accordance with Australian Code for the Care and Use of Animals for Scientific Purposes 8th edition and were approved by the Animal Ethics Committee of the University of Queensland. Mice were anaesthetized with isoflurane. Body temperature was controlled using a heating pad set to 37 °C. 150 µL of a suspension of [CshA_scGFP][S], 2000,000 untreated hMSCs or [CshA_scGFP][S] modified hMSCs was injected with a 27 gauge needle either through a tail vein (intravenous injection) or through the chest wall into the left ventricle (intracardiac injection), respectively. Prior to the injection, the hMSCs were maintained at 4 °C, and the cells were gently resuspended with a pipette to ensure no aggregation before the injection. The mice were sacrificed at 2 hours, 24 hours and 4 weeks after the injection. Blood samples of 1 mL were withdrawn with a needle and syringe from the left ventricle of FVB/N mice for routine blood tests. Heart, lung, kidney and liver samples of the FVB/N and BALB/c nude mice were histologically examined to determine the systemic toxicity.

Electronic Supplementary Material (ESI) for Chemical Science.

Visualization of transplanted hMSCs

hMSCs were stained with CMPTX (Thermo Fisher Scientific, USA) before the injection. The mouse heart was frozen in liquid nitrogen and sectioned at 10 µm. The samples were then mounted using Fluoroshield with DAPI (Sigma-Aldrich, USA). Images were collected with laser scanning confocal microscopy (FV1200, Olympus). The DAPI, GFP and CMPTX fluorescence was excited at 408, 488 and 577 nm, respectively. Figures were processed with FV10-ASW Viewer (4.2b, Olympus) Viewer and Adobe Photoshop CC 2015.

Droplet digital PCR assays for Alu sequences

Genomic DNA (gDNA) of the heart and lung were isolated using DNA Mini Kit (Qiagen, USA). Droplet digital PCR (ddPCR) was performed in reaction consisting of gDNA, primer sets (Alu forward: GCCTGTAATCCCAGCACTTT; Alu reverse: CACTACGCCCGGCTAATTT) [10], H_2O and ddPCR EvaGreen Supermix (BioRad, USA). ddPCR was performed according to manufacturer's manual. Briefly, 20 µL of ddPCR reaction mix was separated into droplets with QX200 Droplet Generator (BioRad, USA). The droplets were transferred into a 96-well PCR plate, sealed and incubated at following cycling conditions: 1 cycle of 95ºC for 5 minutes, 45 cycles of 95ºC for 30 seconds, 55ºC for 1 minute and 1 cycle of 4ºC for 5 minutes, 90ºC for 5 minutes and an infinite hold of 12ºC. After thermal cycling, the PCR plate was transferred in QX200 Droplet Reader and read in FAM channel using QuantaSoft version 1.7.

Compartment modelling of in vivo kinetics of transplanted hMSCs

The model structure was based on published mathematical models simulating the distribution (using distribution coefficient) and uptake-release-excretion processes (using uptake, release and excretion rate constant) of hMSCs [11]. The partition coefficient *P* was used to correlate the concentration of hMSCs between the blood within the organ and the venous blood leaving the organ. The arrest-release-depletion approach of hMSCs was described as a first-order process. Modelling was carried out using Berkeley Madonna version 8.3.18 (Berkeley, USA). The equations describing these processes are:

For vascular space

$$V_{V\_t}\frac{dC_{V\_t}}{dt} = Q_t(C_A - CV_t) - K_{arrest\_t}C_{V\_t}V_{V\_t} + K_{release\_t}A_{E\_t} \quad \textbf{Eq. 3}$$

For the arrested hMSCs as in the extravascular space

$$\frac{dA_{E\_t}}{dt} = K_{arrest\_t}C_{V\_t}V_{V\_t} - K_{release\_t}A_{E\_t} - K_{depletion\_t}A_{E\_t} \quad \textbf{Eq. 4}$$

where C_{V_t} (cell/L) is the concentration of hMSCs in the vascular space within the organ *t*, Q_t (L/h) is the blood flow to the organ *t*, V_{V_t} (L) is the volume of blood vessels in the organ *t*, C_A (cell/L) is the concentration of hMSCs in the arterial blood, CV_t (cell/L) is the concentration of hMSCs in the venous blood leaving the organ *t*, A_{E_t} (cell) is the amount of arrested hMSCs and isolated from blood circulation as in the extravascular space of organ *t*, K_{arrest_t} (h^{-1}) is the arrest rate constant of hMSCs in the organ *t*, $K_{depletion\_t}$ (h^{-1}) is the depletion rate constant of hMSCs in the organ *t*, and $K_{release\_t}$ (h^{-1}) is the release rate constant of hMSCs in the organ *t*.

Statistical Analysis

All the studies were performed using cells from 3 patients. Comparison of differences was tested using ANOVA and a two-tailed Student's t-test with p-value of less than 0.05 and 0.01 considered significant and highly significant, respectively.

References

1 J. P. K. Armstrong, R. Shakur, J. P. Horne, S. C. Dickinson, C. T. Armstrong, K. Lau, J. Kadiwala, R. Lowe, A. Seddon, S. Mann, J. L. R. Anderson, A. W. Perriman and A. P. Hollander, *Nat. Commun.*, 2015, **6**, 7405.

2 R. P. Rambo and J. A. Tainer, *Biopolymers*, 2011, **95**, 559–71.

3 D. I. Svergun, *J. Appl. Crystallogr.*, 1992, **25**, 495–503.

4 S. Hansen, *J. Appl. Cryst.*, 2012, **45**, 566–567.

5 S. Hansen, *J. Appl. Crystallogr.*, 2012, **45**, 566–567.

6 S. Hansen, *J. Appl. Crystallogr.*, 2000, **33**, 1415–1421.

7 S. Hansen, *J. Appl. Crystallogr.*, 2008, **41**, 436–445.

8 S. Hansen, *J. Appl. Crystallogr.*, 2014, **47**, 1469–1471.

9 A. M. Malek, S. L. Alper and S. Izumo, *Jama*, 1999, **282**, 2035–42.

10 G. Shim, S. Lee, J. Han, G. Kim, H. Jin, W. Miao, T. Yi, Y. Cho, S. Song and Y. Oh, *Stem*

Cells Dev., 2015, **24**, 1124–32.

11 H. Wang, X. Liang, Z. P. Xu, D. H. G. Crawford, X. Liu and M. S. Roberts, *Sci. Rep.*, 2016, 1–12.

Supplementary Tables

Table S1. Primers and gene sequence used to amply the fusion protein CshA_scGFP. Lower case nucleotides indicate complementarity for pOPIN F whilst bold, capitalized nucleotides indicate complementarity to CshA_scGFP. Melting temperatures (Tm) are shown next to their corresponding sequences.

	Sequence 5'-3'	**Tm**
Forward primer	aag ttc tgt ttc agg gcc cg ATG GAT GAA ACA AGT GCG AGT GGG GTA CAG	74°C
Reverse primer	atg gtc tag aaa gct tta cta CTT GTA GCG TTC ATC CCG GCC ATG TTT	73°C
Gene sequence	ATG GAT GAA ACA AGT GCG AGT GGG GTA CAG AAC GAA GTT GCT CGT GCC GAT TTG GCG GAA TCC CCG GCT ACC GCC ACC GCG CCC GTC GCG TCT GAA GCC TCC CAG AAT GCG GAA ACC ACA GTG GCT GCG ACC GCT ACT GAA GCC CCG CAG ACC GCG GAA AAC ACT GCT CCT ACC AAC AGT GCC AGT GAA TCA ACC GAA AAA CCG ATG GAC GAA CAG CCT GTG GCT TCT GAA ACC CCA CAA CCA AGC GTC GAG AAA CCG GTA TTA CCG ACC GAG GTG AAA CCG GCA GAG AAT ACA ACT CCG GCA AGT ACG GAG GCC AGT CCG GAG ACA GTT TCT CCA TCG CGC GCG ACA GAC CAG CCT GTA GCC ACC CGT GAT TCC GTG CAG TCT AGC CGT TCG CGC CGG TTG CGC CGT GAC CTG GAA GCT ACC GCT GTG ACG CCA GGT ACA GGT CCG GCT GGT GCG GAT GAT GCA ACA CCG ATT CCT CGT GTC AGC AAA CCG ACC CTG TCA GAG TCA GAA AAA AAA GAA TCC ACC CAA CTG GCG AAG CAG ATT AAC TGG GTA GAC TTC TCC GAC ACA GCG TCA ATG AAG AAT TTA GAT CCG CAA GGA GGA TTC AAA GTA GGT ACC GTT TTC AAA AAG GAA ATC TCG CCG GGT TAT GTG GTG ACG CTG ACC GTT ACT GAA CTG AAA CCC TTC AAC AGC ACC GAA ATT TAC AAG AAA CGC GTT GAA GGG ACT CCC ACG GCA AAT ACC TAC GAC CCG AAC GCG ATC AAT AGC TAT CTG AAA GGC TAT AAA GAT TAC GGT AAA ACC CCG CCG TCT GTT ACG GGC CGC CCG CAG AAC AAA TTT TCA ACC ATT GGC GGT CAG GGG TTC GAT ACC CAA GGT CGT AAA ACG CAG ATC ATT TTG CCG GAC GAC GCG GTT AAT TGG GGC ATC AAA TTT AAG GTG GAA GCA ACC TAT CGC GGA AAT CCT GTG AAA CCT TCG GTC GTT ATG GCC GAT GGG GAA GAT GCC AAT CCG GCT GAA TAT GGG ATT TTC ACG ACT AAC GGT GAA GGG TGG GAG TAT GTG GGC GAA TGG ATG AAA GGA CCC CGC GCG AAA GGC CCG TAC ACT GTG ATG ACT GAA GAT ATG GTG AAG GCA TTC GAT AAA ACC CGC AAA GAC GGT CTG CTG ATC CTG AAA GAT AAA AGC GTT GAC TGG AGC AAA TAC TTG TCT CCA GAC ACA GTT ACT GGT GGA TTA GGT AGC CAG GTG TTC GGC CCG ATC ATC TCA GCC TCA AAA GCG GTA CCG GTG GTT ATG ACT CGC GGT GCG AGC GAA GTC GGG TTT TAT GTC GCC ACG GGT GGG CAA CAA GCC CTC ATG ATG GGT TTT CTC GTT GTC GAT TCG AGC GAC GCA CCA GCC AGC TAT GGC GAA GCG TAT CAT	

ACT ATT GGC ACG CGG GAT TCC ATT GCG AAT ACC CCG ATC AAT CAG CCT TAC TTA GGT
AGC ACC GCA GCA GAC ATT GAT GCG GAC TCT GAA AGT GAC TGG ACT GCC GAT GAC
CGC GAA GAT GTA GCA GAT GAA GGC CCC GCC CAG TTG CTG ACG GCT GAC CAG CTT
AGC AAA ACC AAC GAT TTA CTG GAC CTG AAC AAA GCC AAG AAC GGG ACC TAC ACC CTC
AAA ATC AAA GCG AAC CCA AAC GGT AAC GCA AAA GCG TAC GTC AAG GCA TGG GTG
GAT TTC AAC AAC AAT GGC AAA TTT GAT GAC AAT GAA GGC TCG GTG GTG AAG GAG
ATT ACC GCC AAC GGG GAT CAT ACG CTG TCC TTT AAC GCC ATT CCT GGC CTT ACC GGC
GGC CTG GTG GAC CAG ATT GGC ATG CGG GTA CGC ATT GCG ACG AAT GCA GGG GAT
ATT GAG AAA CCG ACA GGT ACC GCG TTC AGT GGG GAA GTA GAG GAT ATG CTG GTT
CGC CGT GTC TAT CCG CCA CAA GGC GAA AAG CAG GAA TCT ACT GGC TTC CAA GGA
GAA ACC CAG AAT GCT TCG GTG CAC TTT ACC GCA AAA GGA CCG GAT CGC TCC GAT TTT
GTA ACC AAC GCG AGC ATG AGC AAT CAA GCG CCA CAG GTT CTG GAT AAT CAG GGC
AAC GTT CTG ACG CCG ACC AAT GGT AAT ACC TAT GTA CGT CCC GAG GGA ACG TAC GTG
GTG ACA GCC AAT GGC GAT GAT GTC AAC GTT ACG TTC ACT CCG AAC GAG GAT TTC AGC
GGT GTT GCG GAG GGT ATT AAC ATT CGT CGC ACT GAC TCA AAT GGT TCC AGC ACG GGT
TGG CAG TCG ACG GAT GCA GCA GAT CCG AAT AAG AAC GAT CGC TTG AAC AAC ATG
GAC GGC CGT TTT GTG CCA ACC GTC CGC AAA GTG CCT AAA TAC GAC AGT ACG GGC ATT
CAG GGC CAG GAT ATG GCC TCG AAA GGC GAA CGC CTG TTT CGC GGC AAA GTG CCG
ATT CTG GTA GAA CTT AAA GGT GAT GTG AAT GGC CAC AAA TTT AGC GTG CGT GGC
AAA GGT AAA GGC GAT GCA ACG CGT GGC AAA TTA ACG CTC AAA TTT ATC TGT ACG ACT
GGC AAA TTA CCC GTG CCG TGG CCC ACC CTT GTT ACG ACG CTG ACC TAT GGT GTC CAA
TGC TTT TCC CGT TAT CCG AAA CAC ATG AAA CGT CAC GAC TTC TTT AAA AGC GCC ATG
CCT AAA GGC TAT GTC CAA GAG CGC ACG ATC TCG TTT AAG AAA GAC GGG AAA TAT
AAG ACC CGG GCG GAA GTC AAA TTT GAG GGC CGT ACG CTG GTG AAT CGC ATT AAA
CTG AAG GGA CGC GAC TTT AAA GAG AAA GGC AAT ATC CTG GGT CAT AAG CTC CGC
TAC AAC TTT AAC TCT CAT AAA GTC TAC ATC ACC GCT GAT AAA CGC AAA AAT GGC ATC
AAA GCA AAA TTC AAG ATT CGC CAT AAC GTT AAG GAT GGT TCT GTC CAG CTG GCG GAT
CAT TAT CAG CAA AAC ACC CCA ATT GGC CGT GGT CCG GTT CTT CTG CCA CGT AAC CAC
TAT CTG TCA ACG CGT TCG AAA TTG AGT AAG GAT CCG AAA GAG AAA CGT GAT CAC ATG
GTG CTG CTG GAA TTT GTC ACA GCC GCG GGT ATC AAA CAT GGC CGG GAT GAA CGC
TAC AAG

Table S2. Peak integration for sedimentation coefficient (S) distributions of the constructs. Sedimentation velocity analytical ultracentrifugation (SV-AUC) was used to measure the sedimentation coefficients of CshA_scGFP and [CshA_scGFP][S]. The optimal stoichiometry of polymer surfactant was identified by conjugating CshA_scGFP with different protein cationic charge:polymer surfactant ratios. All data was measured at 25 °C and the normalization for temperature and solvent was carried out automatically by Sedfit software.

			Peak 1		**Peak 2**		
		λ	**S**	**percentage**	**S**	**percentage**	***f/f₀***
CshA_scGFP		280	2.2	16.1	4.0	83.9	1.62
		487	1.9	11.6	4.1	88.4	1.81
Surfactant c(S)		280	1.4	100	NA	NA	2.48
[CshA_scGFP][S]	0.6	280	2.1	27.5	4.7	72.5	1.57
(final protein cationic		487	2.5	6.4	4.8	93.6	1.75
charge:polymer	1.1	280	1.7	51.9	4.7	481	1.78
surfactant ratio)		487	2.3	11.1	4.8	88.9	1.60
	1.6	280	1.6	68.0	4.6	32.0	2.33
		487	2.1	12.3	4.7	87.7	1.75

Table S3. Small angle x-ray (SAXS) and small angle neutron scattering (SANS) parameters for the constructs. SAXS and SANS were used to probe the solution structure and conformation of CshA_scGFP and [CshA_scGFP][S]. Contrast matching was used in SANS to study the CshA_scGFP structure (100% D_2O) and the surfactant distribution on the protein using the protein contrast match point (40% D_2O/60% H_2O). Real space parameters were derived using indirect Fourier transformations of the experimental data using BAYESAPP and ScÅtter.

	SAXS		SANS	
	CshA_scGFP	[CshA_scGFP][S]	CshA_scGFP	[CshA_scGFP][S]
% D_2O	-	-	100	40
I(0)	7.92E-03 ± 1.8E-05	1.54E-02 ± 4.9E-05	2.88E-01 ± 1.07E-03	2.24E-01 ± 1.62E-03
R_g (Å)	73.6 ± 0.4	99.8 ± 0.5	60.7 ± 0.6	57.7 ± 2.0
D_{max} (Å)	291.1 ± 2.6	344.1 ± 3.2	249.5 ± 4.1	238.6 ± 11.4
Axial ratio	10.6 ± 0.4	8.2 ± 0.1	-	-
Porod exp	2.1 ± 0.1	1.8 ± 0.1	-	-
χ^2	1.26	1.65	0.46	0.48

Table S4. Routine blood tests of mice after intravenous injection of [CshA_scGFP][S]. Blood samples of 1 mL were withdrawn with a needle and syringe from the left ventricle of FVB/N mice for routine blood tests.

Technique	**Control**	**2 h**	**24 h**
Hemoglobin (g/L)	151 ± 16	160 ± 11	158 ± 14
White blood cells (× 10^9/L)	8.6 ± 2.1	8.7 ± 1.5	10.1 ± 2.0
Platelets (× 10^9/L)	973 ± 87	894 ± 124	918 ± 154
Hematocrit	0.45 ± 0.09	0.52 ± 0.13	0.48 ± 0.11
Red blood cells (× 10^{12}/L)	10.2 ± 2.5	11.4 ± 2.9	10.5 ± 2.8

Table S5. hMSC-specific parameters used in the physiologically based kinetic model.

Parameter (unit)	Description	Blood	Heart	Lung	Liver	Spleen	Kidney	Rest of body
P (unitless)	Partition coefficient	-	3.097 (hMSCs), 6.106 ([CshA_scGFP][S]-hMSCs)	742.733	262.699	1633.24	305.351	6.765
K_{arrest} (h^{-1})	Arrest rate constant	-	1.251	5.434	1.395	0.608	1.727	0.143
$K_{release}$ (h^{-1})	Release rate constant	-	0.016	0.108	0.066	0.856	0.054	0.957
$K_{depletion}$ (h^{-1})	Depletion rate constant	0.636	0.039	0.0589	0.060	0.002	0.151	0.148

Supplementary Figures

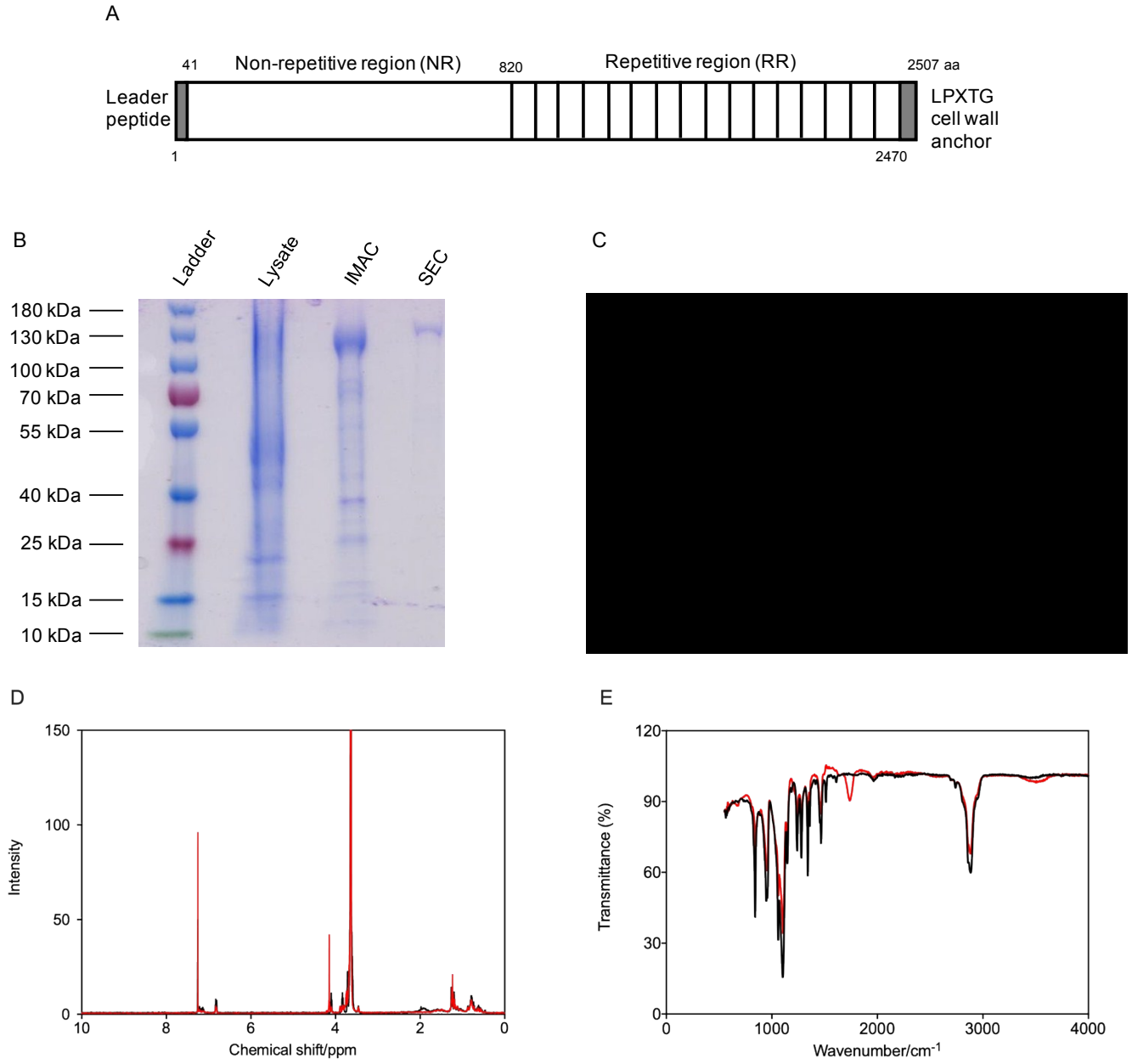

Fig. S1. The fusion protein CshA_scGFP.

(**A**) Reassigned domain architecture of CshA. The CshA polypeptide comprises an N-terminal leader peptide (residues 1-40), a non-repetitive region (residues 41-819), a repeat region (residues 820-2507), and a C-terminal LPXTG cell wall anchor motif. (**B**) Chromatographic sizing of the fusion protein CshA_scGFP. SDS-PAGE gel shows combined chromatography fractions from lysate, immobilized metal ion affinity chromatography (IMAC) and size exclusion chromatography (SEC). The results from the SEC display only one band close to 130 kDa, indicating the purity of CshA_scGFP. (**C)** Mass measurements of the fusion protein CshA_scGFP. Matrix-assisted laser desorption/ionization time-of-

flight mass spectrometry (MALDI-TOF) was used to calculate the mass of fusion protein CshA_scGFP. The first 8 peaks of the mass multiplied by their corresponding charges showed the final mass of 112.6 kDa, which is accord with the predicted value of 112.6 kDa from the protein sequence. (**D**). H_1 NMR spectra of the native surfactant (black) and the oxidised surfactant (red) at 25 °C showing a sharp singlet peak at 4.3 ppm for the oxidised surfactant, which indicates the presence of electronegative groups in close proximity to the carboxylate moiety. (**E**). Fourier transform infra-red (FTIR) spectra measurements of the native surfactant (black) and the oxidised surfactant (red) at 25 °C showing the emergence of the C=O stretch at 1738 cm^{-1}.

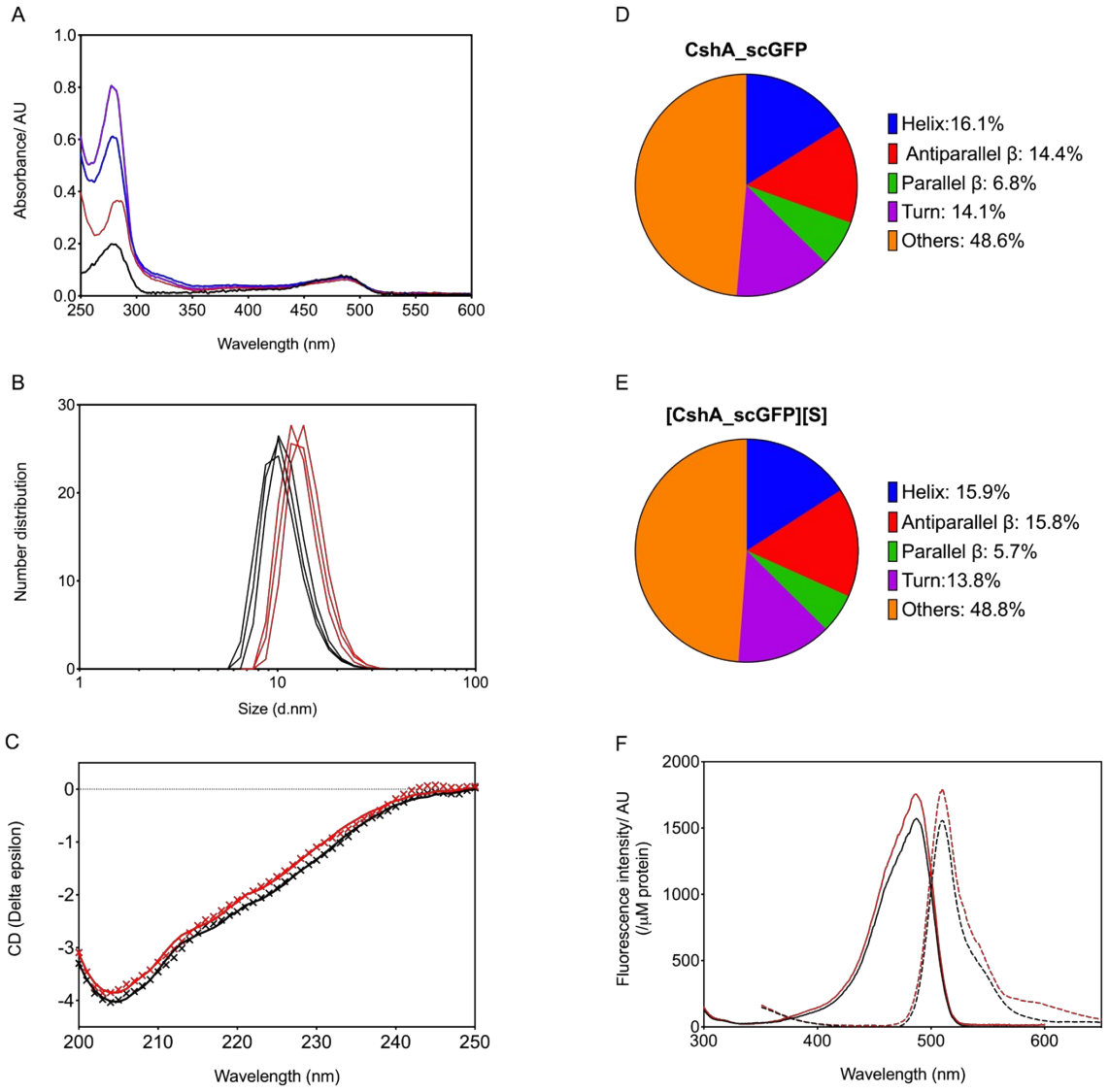

Fig. S2. Surfactant conjugation of CshA_scGFP to yield [CshA_scGFP][S].

(**A**) UV-vis spectroscopy of the constructs showing the absorption of CshA_scGFP (black trace) and dialyzed [CshA_scGFP][S] with protein:polymer surfactant charge ratios of 1:1.4 (red trace), 1:2.1 (blue trace) and 1:2.8 (purple trace). (**B**) Dynamic light scattering (DLS) data showing the hydrodynamic diameter number distributions of CshA_scGFP (black trace) and [CshA_scGFP][S] (red trace) at 1 mg/mL. The measurements were carried out in triplicate after 120 seconds of equilibration at 25 °C. (**C**) Far-UV synchrotron radiation circular dichroism (SRCD) spectra from CshA_scGFP (black) and [CshA_scGFP][S] (red) at 36 °C. Experimental data (symbols) is overlaid with a fits (solid lines) obtained using BeStSel software. The resulting secondary structure distributions for CshA_scGFP and CshA_scGFP[S] are displayed in (**D**) and (**E**), respectively. (**F**) Fluorescence spectroscopy showing the excitation (solid trace) and emission (dashed trace) of CshA_scGFP (black trace) and [CshA_scGFP][S] (red trace). The excitation peak at 487 nm and emission peak at 510 nm can be attributed to the fluorophore in the scGFP domain of the constructs.

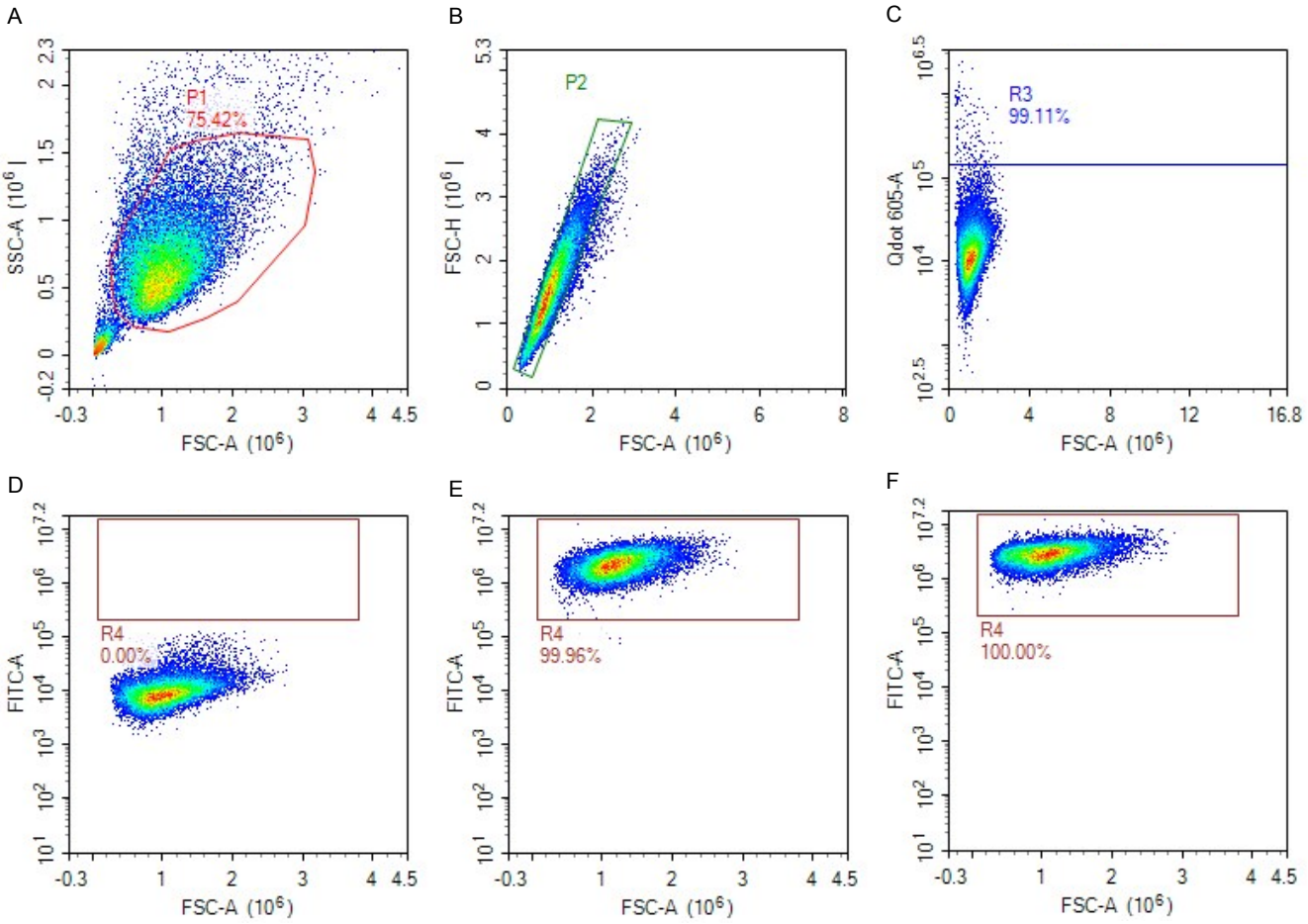

Fig. S3. Flow cytometry gating parameters.

(**A**) The hMSC population was identified using a plot of side scatter area (SSC-A) to define cell granularity, and forward scatter area (FSC-A) to define cell size (P1). (**B**) Single cells were then identified by forward scatter height (FSC-H) and FSC-A. (**C**) Cells with low propidium iodide (PI) fluorescence (Qdot 605-A) were gated in order to exclude non-viable cells. (**D**) Untreated cells were used to define an scGFP-negative cell population using the green fluorescence channel (FITC-A). These gating parameters were applied to cells functionalised with either (**E**) CshA_scGFP or (**F**) [CshA_scGFP][S], and the proportion of labeled cells and mean fluorescence intensity were analyzed.

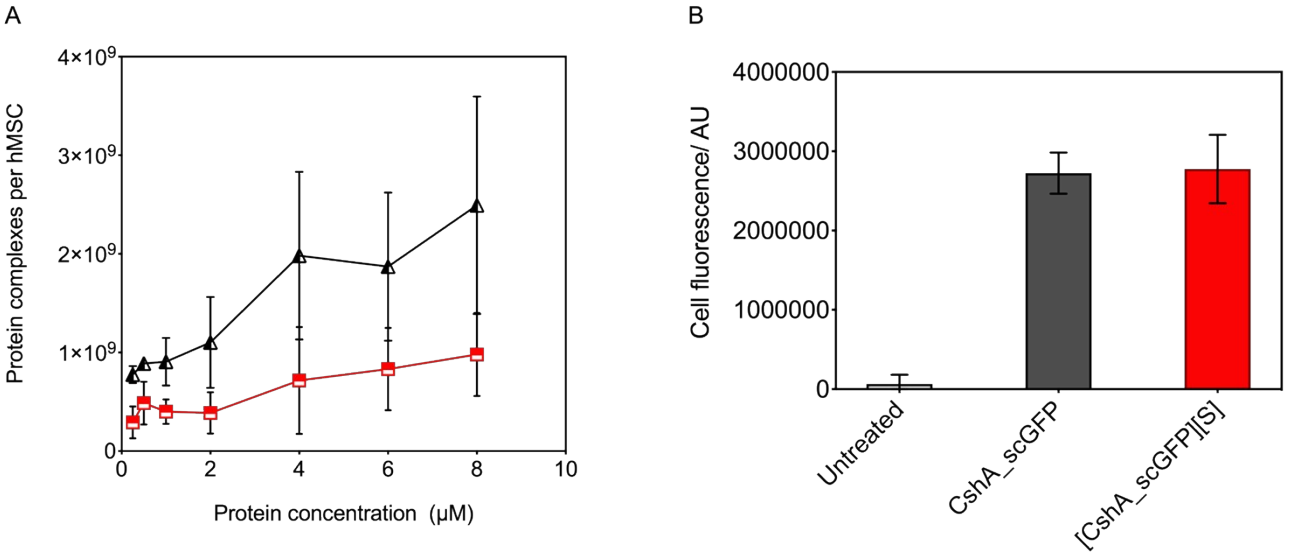

Fig. S4. Cell membrane affinity of the constructs.

(**A**) To calculate the number of constructs per hMSC, hMSCs were incubated over a range of concentrations (0 – 8 μM) using either CshA_scGFP (black) or [CshA_scGFP][S] (red). The depleted supernatant was compared to the initial solutions using UV-vis spectroscopy. A reduction in the absorbance at 487 nm allowed the number of construct molecules per hMSC to be calculated. Approximately 0.7 billion [CshA_scGFP][S] molecules per cell were present, compared to 2 billion CshA_scGFP molecules, when a saturation was reached at a concentration of around 4 μM. The triangles and squares represent the mean average and the error bars represent the standard deviation calculated using hMSCs from 3 different patients (n=3). (**B**) Flow cytometry showing the mean fluorescence intensity of the cells labelled by the constructs. Both CshA_scGFP (black) and [CshA_scGFP][S] (red) labelled the entire cell population.

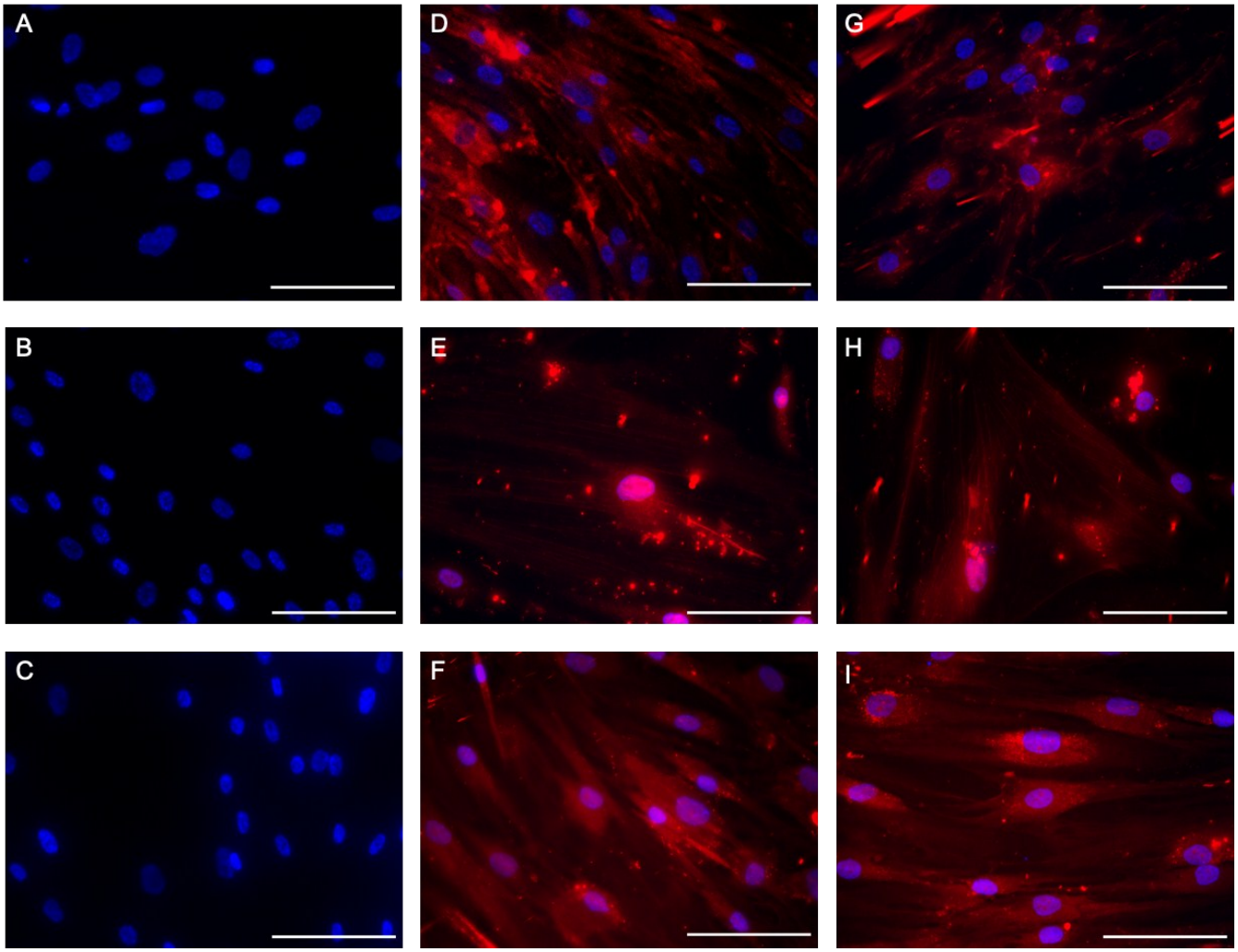

Fig. S5. The *in vitro* cardiogenic differentiation of hMSCs.

Representative confocal microscopy images showing the expression of cardiogenic biomarkers of (**D-F**) untreated hMSCs and (**G-I**) 4 µM [CshA_scGFP][S] modified hMSCs 14 days after exposed to 5 µM of 5-Azacytidine. (**A-C**) Undifferentiated hMSCs were taken as controls and all the cells were stained with DAPI (blue). (**A, D, G**) hMSCs were stained with anti-Nkx2.5 (red), (**B, E, H**) anti-alpha cardiac actin (red) and (**C, F, I**) anti-cardiac troponin T (red). Scale bar=100 µm.

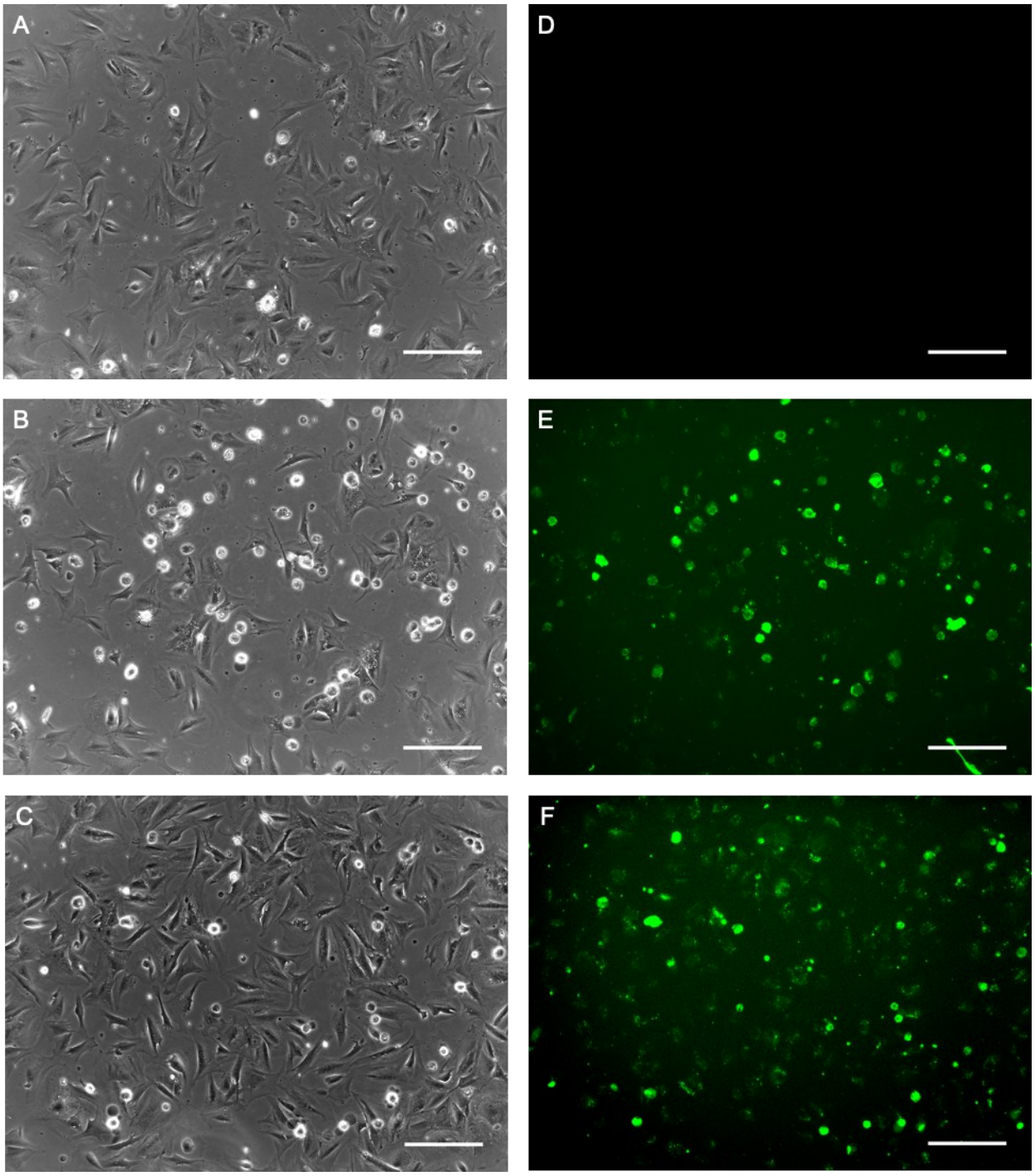

Fig. S6. Live cell widefield microscopy of hMSC adherence on immobilized fibronectin (Fn).
Adherent cells were left untreated (**A, D**), incubated with 4 µM of either CshA_scGFP (**B, E**) or [CshA_scGFP][S] (**C, F**) and then washed in PBS and imaged after 6 hours of incubation. Images were collected in bright field mode (left) and using widefield fluorescence mode (right). Scale bar=200 µm.

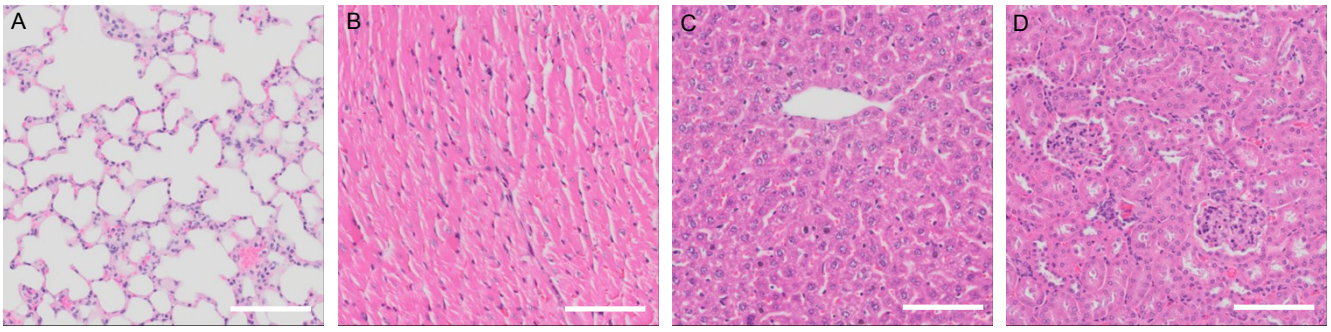

Fig. S7. Pathological analysis of the major organs.

Histology of the (**A**) lungs, (**B**) heart, (**C**) kidney and (**D**) liver of mice at 4 weeks after intravenous injection of [CshA_scGFP][S] modified hMSCs. Histological images were collected on Haemotoxylin and Eosin (H&E) stained section. No notable differences were observed among these organs, suggesting that [CshA_scGFP][S] modified hMSCs had no acute toxic effects on normal tissues *in vivo*. Scale bar=200 µm.

Electronic Supplementary Material (ESI) for Chemical Science.

Supplementary movies

S1. Cell adherence to Fn at a shear stress of 5 dyne/cm^2 in biomimetic microfluidics.

S2. Cell adherence to Fn at a shear stress of 2.5 dyne/cm^2 in biomimetic microfluidics.

S3. Cell adherence to Fn at a shear stress of 1 dyne/cm^2 in biomimetic microfluidics.

S4. Ultrasound imaging-guided intracardiac injection.