Dixon, James E. and Osman, Gizem and Morris, Gavin E. and Markides, Hareklea and Rotherham, Michael and Bayoussef, Zahia and El-Haj, Alicia and Denning, Chris and Shakesheff, Kevin M. (2016) Highly efficient delivery of functional cargoes by the synergistic effect of GAG binding motifs and cell-penetrating peptides. Proceedings of the National Academy of Sciences, 113 (3). E291-E299. ISSN 1091-6490

Classification: Biological Sciences: Cell biology

Title: Highly Efficient Delivery of Functional Cargoes by the Synergistic Effect of GAG Binding Motifs and Cell-Penetrating Peptides

Authors: James E. Dixon[1*], Gizem Osman[1], Gavin E. Morris[1], Hareklea Markides[3], Michael Rotherham[3], Zahia Bayoussef[1], Alicia El-Haj[3], Chris Denning[2] and Kevin M. Shakesheff[1*]

Author affiliation: Wolfson Centre for Stem Cells, Tissue Engineering, and Modelling (STEM), Centre of Biomolecular Sciences, [1]School of Pharmacy, [2]School of Medicine; University of Nottingham, Nottingham, NG7 2RD, UK; University of Nottingham, Nottingham, NG7 2RD, UK. Institute for Science, Technology & Medicine, Keele University, Guy Hilton Research Centre, Thornburrow Drive, Hartshill, Stoke-on-Trent, ST4 7QB, UK *To whom correspondence should be addressed to J.E.D. james.dixon@nottingham.ac.uk

Corresponding authors:

Kevin M. Shakesheff & James E. Dixon

Wolfson Centre for Stem Cells, Tissue Engineering, and Modelling (STEM)

Centre for Biomolecular Sciences

University of Nottingham,

University Park, Nottingham

NG7 2RD, UK

Tel: +44 (0) 115 82 32003

Fax: +44 (0) 115 84 68002

Keywords: Cell penetrating peptides, Transduction, Macropinocytosis, Human Embryonic Stem Cells, Heparin Binding Domain

Short Running Title: GET: Glycosaminoglycan-binding enhanced transduction

Abstract

Protein transduction domains (PTDs) are powerful non-genetic tools that allow intracellular delivery of conjugated cargoes to modify cell behaviour. Their use in biomedicine has been hampered by inefficient delivery to nuclear and cytoplasmic targets. Here we overcame this deficiency by developing a series of novel fusion proteins that couple a membrane docking peptide to heparan sulfate glycosaminoglycans (GAGs) with a PTD. We showed this GET (GAG-binding enhanced transduction) system could deliver enzymes (Cre, neomycin phosphotransferase), transcription factors (NANOG, MYOD), antibodies, native proteins (Cytochrome-C), magnetic nanoparticles (MNPs) and nucleic acids (plasmid (p)DNA, modified (mod)RNA and siRNA) at efficiencies of up to two-orders of magnitude higher than previously reported in cell types considered hard to transduce, such as mouse embryonic stem cells (mESCs), human ESCs (hESCs) and induced pluripotent stem cells (hiPSCs). This technology represents an efficient strategy for controlling cell labelling, and directing cell fate or behaviour that has broad applicability for basic research, disease modelling and clinical application.

Significance Statement

Efficient delivery of therapeutic molecules inside cells by non-transgenic approaches is key as gene editing/correction, directed differentiation and *in vivo* cell modulation/tracking are translated for regenerative medicine applications. This paper describes a peptide-based system engineered to enhance the activity of cell penetrating peptides in order to achieve exceptional intracellular transduction. Glycosaminoglycan-binding enhanced transduction (GET) utilises peptides which interact with cell membrane heparan sulfates and promote CPP-mediated endocytosis into cells. The system is not dependent on extensive positive charge and can be tailored to deliver peptides, recombinant proteins, nucleic acids, nanoparticles and antibodies. Importantly this approach does not affect cell proliferation and viability, and can be utilised to deliver a plethora of functional cargoes.

/body

Non-genetic tools that afford controllable stoichiometry in the absence of genome integration (1) are attractive for use in directing cell fate and unravelling protein function. In this regard, cationic PTDs have been derived from the human immunodeficiency virus (HIV-1) TAT protein basic domain (e.g. RKKRRQRRR), or engineered from poly-arginine or -lysine. PTDs have been used to deliver bioactive cargoes of proteins, nucleic acids, nanoparticles and pharmacological agents to cells in culture and to preclinical models *in vivo* (2, 3). As an alternative to unstructured cationic PTDs there has been development of super-charged (sc) protein carriers which have been extracted from natural proteins or been engineered i.e. scGFP (+36 GFP (4)). These proteins (defined as >0.75 theoretical positive charge/kDa mass; (5)) have exceptional delivery into mammalian cells but due to their extensive positive charge can be cytotoxic.

While the mechanism of PTD and sc-protein uptake is still poorly defined(1, 6-8), it is clear that their extracellular concentration dictates the efficiency of delivery. Recent studies have demonstrated that PTD-mediated protein delivery is mediated by endocytotic pathways, such as lipid raft-dependant macropinocytosis(8). After internalization cargos are contained within macropinocytomes and not free within cytosol; restricting correct localization and activity (9, 10). This means high (>200 nM) (11, 12) *in vivo* extracellular concentrations of PTD-proteins are needed to drive translocation and permit biologically-useful quantities to escape endosomes. Various techniques have been developed to stimulate the vesicular release of cargo, including ultrasound, co-treatment with endosome-disruptive peptides or chemical treatments (1, 13). However the efficiency of these processes still leaves the majority of protein trapped away from cytosol (14, 15).

To overcome the inefficiencies in delivery of biologically-active PTD-conjugates, we developed a multi-domain protein comprising a GAG-binding peptide to stimulate cell interaction and a PTD for high efficiency membrane transduction. We provided proof of principle of the efficacy of this system by delivering fluorescent proteins, enzymes, transcription factors, antibodies, nucleic acids and nanoparticles to mouse and human cells including pluripotent stem cells. Thus, the development of synergistic cell-targeting and cell-penetrating peptides enables functional quantities of therapeutic cargoes to be delivered to pluripotent and other clinically relevant cell-types.

Results

Isolation of P21, a HBD that enhances PTD function through GAG interaction

We first focused on improving the initial PTD/cell lipid-bilayer interaction and cellular-uptake of cargo proteins. For this we employed monomeric red fluorescent protein (mRFP1; mR) as a self-reporting cargo that is readily expressed and purified in *Escherichia coli* (SI Appendix, Fig S1a). We confirmed that although TAT- and 8R PTD-tagged mR interact with most cell types, levels of cell-uptake were highest in somatic cells (NIH3t3, C2C12, mouse embryonic fibroblasts, ~5-50-fold over control, $p<0.01$; SI Appendix, Fig S1) and poor in mESCs, hESCs and hiPSCs (~1-3 fold over control; SI Appendix, Fig S1 & S2b). To improve cellular-uptake of cargo proteins into these poorly transduced cell-types we screened several short peptides (identified in the literature) which have been reported to interact with ubiquitous molecules on cell membranes including integrins, CD markers and GAGs. We fused peptides to the *N*-terminal mR with screening of variants yielding the isolation of a short 21-residue peptide, termed P21

(KRKKKGKGLGKKRDPCLRKYK) (Fig 1). P21 enhances the association of the mR reporter with both mouse and human pluripotent stem cells. Furthermore, P21 also demonstrates significant transduction activity itself exhibiting punctate intracellular fluorescence indicative of endosomal localization (Fig 1b,c). P21 is the heparin-binding domain (HBD) from heparin-binding epidermal growth factor (HB-EGF) that shows a strong affinity to heparin specifically through P21 (16). Heparin binding is essential for its optimal binding to EGFR and for promoting its growth/migratory activity in vascular smooth muscle cells (17, 18). Interestingly HB-EGF has recently been described as a natural 'supercharged' protein with a theoretical net charge:molecular weight ratio >0.75 (5). The fragment of HB-EGF (Pfam: Q99075, aa72 to aa147 with extra N-terminal residues, 79aa) used in that study contained the P21 HBD sequence and when compared our P21 fragment (aa93-aa113) that study's truncated HB-EGF fragment yields similar transduction activity (5). We hypothesised that cell-uptake activity intrinsic to P21 is due to its supercharged nature, however not all positively charged sequences can be considered HBDs. For confirmation we demonstrated the direct binding of P21 to heparin by showing P21-tagged mR was efficiently and reversibly sequestered on heparin-sepharose (96.2±5.3%; $p<0.01$) (Fig. S3).

We hypothesised that enhanced cell binding coupled with a cell-penetrating peptide would significantly enhance cell uptake. This would also indicate that further charging of P21-fused proteins via a PTD would enhance uptake if functioning primarily through heparan sulfate binding. To test this we combined both moieties in one molecule (Fig 1a). The inclusion of both P21- and 8R synergized to enhance uptake of all cell-lines tested (Fig. 1b). Importantly mouse and human pluripotent stem cells (CGR-8, HUES7 and IPS2) and cardiomyocytes (HL1) only possessed efficient uptake with the inclusion of P21. P21-mediated activity was more than two-

orders of magnitude more efficient than PTD-mediated uptake alone (Fig. 1c). We tested these motifs in tandem at both *N*- and *C*-terminal of mRFP or switched their termini. All variants demonstrated similar synergistic behaviour (SI Appendix, Fig S4) even when 8R was swapped for alternative cationic PTDs (TAT, 8K and 8RQ; (2)) (SI Appendix, Fig S5). We describe this synergistic delivery mechanism as GAG-binding enhanced transduction (GET).

We attempted to address the issue of cationic charge and its influence on P21 activity more thoroughly. Addition of the 8R PTD adds 8+ and P21 adds 11+ to the mRFP reporter (Tab S1) so to discount any primary cationic effect of P21 and to confirm that P21 genuinely targeted membrane-bound heparin sulfates we constructed variants with scrambled P21 sequence (P21*) and also with all but the positive residues (K & R) deleted (termed P21 KR only) (SI Appendix, Fig S6). Neither of these variant with similar charge could be taken up with wild-type P21 efficiency however they did endocytose confirming positive charge is a pre-requisite for cell delivery. These data importantly demonstrates that P21 is not a simple unstructured cationic peptide.

For completeness we also compared uptake of mRFP with more extensive positive charging by attaching longer poly-arginines (4R, 8R as before, 16R, 24R, 32R and 48R). mRFP-24R has a net positive charge of +19.5 versus P21-mRFP-8R with +14.4 at pH7.0 but did not deliver to the same magnitude nor did longer poly-arginine versions (SI Appendix, Fig S6). Interestingly if longer poly-arginine stretches were included with P21 there was greater synergy (until a saturation point: ~+25) demonstrating P21 does not elicit its activity through simple cationic charging.

GET requires the presence of trypsin-sensitive and detergent-soluble cell membrane molecules

To evaluate the mechanism of GET interaction and uptake by cells we performed a series of experiments which have been previously used to assess PTD activity. We showed enzymatic digestion of cell membrane proteins with trypsin reduced GET by 8.4-fold ($p<0.05$) (Fig. 1e). In contrast, it was unaffected by cell dissociation solution (CDS), a non-enzymatic method for cell dispersal. Furthermore, we depleted detergent-soluble cell-membrane molecules with Triton X-100 and observed a ~2.2-fold decrease in GET ($p<0.05$) without a decrease in viability. Therefore both protein and detergent-soluble moieties on the cell membrane affect the efficacy of P21- and PTD- synergy in GET.

Synergy between HDBs and PTDs is a general phenomenon

To determine if P21 had unique activity or if other natural HBDs elicit the same synergy with PTDs we probed the literature and constructed a series of other mRFP-GET proteins with HBDs taken from different growth-factor families and extracellular matrix proteins. (SI Appendix, Tab S1, Fig S7). Importantly some of these were essentially neutrally or negatively charged which would allow us to highlight the importance of cationic charge for enhanced transduction through HBDs. We were able to show that all HBDs analysed promoted PTD-mediated delivery of mRFP into cells. Without a PTD, some HBDs promoted transduction themselves but this was not due to simple cationic charge as some with high positive charge (FGF4 and 7 B domains; Table S1). Also with the addition of 8R PTD there was no link between the overall charge of the molecule and its transduction activity (SI Appendix, Fig S7c and d). We tested a selection of

these HBDs (with and without the addition of 8R) in other cell types to demonstrate any cell-type specific activity. We could confirm that a number of HBDs enhance transduction more strongly in certain cell-types (i.e. FGF2B, PDGF; SI Appendix, Fig S7c) while many produced similar lower levels of enhancement in all the three cell lines tested (NIH3t3, CGR8 and HUES7). In conclusion the intrinsic transduction activity and enhancing function of HBDs on PTDs appears independent on its own positive-charge, and this activity can be more potent in particular cell-types.

GET enhances Cre-mediated genome modification

It was imperative to assess the proportion of delivered protein that escaped endosomes and therefore could be considered as functional by GET. Previous studies have the avoided issues associated with direct measurement of fluorescent-tagged proteins (such as being unable to distinguish membrane, vesicle or functional cytosolic/nuclear protein) by assaying for the successful nuclear activity of Cre recombinase (8). We also take advantage of this system and measured Cre-mediated recombination of a *lox*P-STOP-*lox*P (LSL) enhanced green fluorescent protein (eGFP) reporter gene in live NIH3t3 mouse fibroblast cells (NIH3t3: LSL-eGFP cells) (Fig 2a). This system can be considered a measure of functional transduction as well as cellular uptake as activation of green fluorescence requires exogenous Cre protein to enter the cell, undergo nuclear-translocation and excise the LSL fragment of the transgene.

Transduction of NIH3t3: LSL-eGFP cells with SIN Cre lentiviruses to overexpress *Cre* transgenically led to near complete (92±6%; $p<0.001$) activation of eGFP-expression in all cells confirming the utility of this system (Fig 2b). Reporter activation requires only one functional

Cre recombinase complex (4 molecules) to be delivered so does not allow the determination of the precise amount of cargo delivered. To overcome this issue we delivered Cre proteins at limiting dilutions for a short exposure time (1 hour) and determined the minimum dose required to activate green fluorescence after 48 hours (Fig 2c).

Treatment of NIH3t3: LSL-eGFP cells with mR-Cre (mRFP fused to Cre) resulted in recombination and eGFP activation (22.1±6.7%; $p<0.05$) at the highest doses (500µg/ml) (Fig 2d). eGFP activation was inhibited at 4˚C and negatively affected by serum concentration-dependently. mR-Cre-8R demonstrated that the 8R PTD enhanced functional delivery of Cre (~22-fold; $p<0.01$). GET-Cre (P21-mR-Cre-8R) required as little as one minute incubation with cells at a low dose (1µg/ml; ~30nM) to elicit recombination (4.3±2.5%; $p<0.05$) confirming that binding and internalization is an efficient and rapid process. For a moderate dose (10µg/ml; ~300nM) GET achieved a complete functional delivery and recombined all NIH3t3; this is ~15-fold ($p<0.01$) above PTD only levels, ~340-fold higher than mR-Cre ($p<0.001$) (Fig 2d,e). We repeated heparinase III, free-heparin and serum-free experiments and confirmed that heparinase III pre-treatment reduced recombination to basal-levels and media serum plays a role in replenishing cell membrane GAGs depleted by heparinise (SI Appendix, Fig S8-11). Overall these data correlate with the fluorescence delivery conclusions and show synergy between P21- and PTD- moieties to achieve significant increases in functional transduction of protein cargo.

GET of NANOG promotes Self-renewal of Pluripotency

If GET-technology is to be adopted for clinical applications demonstration of its use to alter cell-fate is crucial. An important application would be in the driving of reprogramming, self-renewal

and differentiation of stem cells. iPSC technology has been swiftly developed to allow genome non-integrating DNA (19), RNA (20) and protein (11, 12) based technologies to supersede the original retroviral protocols (21) (Fig 3).

We employed CGR-8 mESCs to determine if GET-mediated delivery can sustain their pluripotent self-renewing phenotype with the withdrawal of leukemia inhibitory factor (LIF). We delivered GET NANOG-cargo (P21-mR-NANOG-8R) in an assay (22) similar to that used to initially isolate the role *Nanog* in mESCs (23). P21-mR-NANOG-8R rescued pluripotency-associated alkaline phosphatase (AP) activity in significant numbers of CGR-8 even with relatively low doses (10µg/ml) (Fig 3b). AP activity in high-dose P21-mR-NANOG-8R samples was similar to that achieved by the SIN NANOG lentiviral transgenic. Transgenics and high-dose P21-mR-NANOG-8R-transduced CGR-8 cells proliferated to a similar level in LIF-deficient cultures (~87.6-fold more; $p<0.001$) (Fig 3c) and also retained *Oct4* expression to a similar level (albeit lower than LIF-containing cultures), indicative of retention of pluripotency (both $p<0.05$) (Fig 3d). As observed previously, rescued cells by both transgenic and protein methods up-regulated *Fgf5* and down-regulated *Rex1* expression, indicative of an inner cell mass-to-epiblast transition phenotype (22). A CPP-version (mR-NANOG-8R) of this protein did not confer LIF-independence to cells (SI Appendix, Fig S12).

GET of MYOD drives Myogenesis

Given the high efficiency of functional PTD-mediated cargo delivery by P21-enhancement, we reasoned that GET technology might also be utilized to redirect pluripotent cells toward

differentiated cell-fates using the transduction to deliver recombinant transcription-factors (24-26).

For this we used the delivery of the efficacious MYOD myogenic factor (27) to drive skeletal muscle specification (Fig 4). We devised an *in vitro* differentiation protocol in which HUES7 cells were transduced with a GET MYOD-cargo, P21-mR-MYOD-8R (Fig 4a). We show that a high percentage of large multinucleated MYOGENIN-positive myotubes (62.1±8.9%; $p<0.01$) (Fig 4b-f) which had elevated *MYOD* and *ACTA1* expression ($p<0.01$ and <0.05, respectively) (Fig 4c) could be generated by GET-MYOD. This was comparable to the differentiation observed in the SIN MYOD lentiviral transgenic when P21-mR-MYOD-8R was delivered at higher doses. A CPP-version (mR-MYOD-8R) of this protein did not promote myogenesis at these concentrations (SI Appendix, Fig S13). Taken together, these experiments provide proof of principle that GET-delivered protein can direct the fate of pluripotent cells to a terminally differentiated somatic cell type.

GET can be coupled to a variety of clinically useful cargoes

Since GET can effectively deliver functional recombinant proteins we assessed if the P21 and 8R peptide moieties can be linked to a variety of other cargoes to enhance intracellular delivery. We initially tested other protein cargoes (Fig 5, S14 & S15). The first strategy was to produce a cargo-interacting P21-8R variant by cloning and recombinantly expressing monomeric streptavidin (mSA2) fused between P21-8R to effectively target biotinylated cargoes (Fig 5ai, S16), or by using *staphylococcal* protein A B domains (SpAB) which have strong affinity for IgG antibodies (SI Appendix, Fig S25). By simple mixing of GET-mSA2 (P21-mSA2-8R; Fig 5a) we show a significant enhancement in transduction a naturally extracted (Bovine heart) pro-

apoptotic factor, Cytochrome-C (Cyt-C) which we biotinylated *in vitro* (to produce BIO-Cyt-C) to allow GET-mSA2 interaction. Only with the GET system, and specifically with biotinylated Cyt-C (BIO-Cyt-C) along with endosomal escape (employing chloroquine) did cells respond to cytochrome-C delivery and lose viability (SI Appendix, Fig S14). Addition of GET-mSA2 to biotinylated primary (1˚) antibodies (which intern bind to complementary secondary/2˚ antibodies) yielded efficient delivery of the complex (measured by 2˚ antibody fluorescence intracellularly; 26-fold over 1˚ and 2 ˚ antibodies alone) (Fig 5b, c). Using GET-SpAB protein, IgG antibodies could be directly delivered to cell without requirement of biotinylation (~786-fold over antibody alone) demonstrating the utility of GET as a biotinylated-cargo or IgG antibody transduction reagent (SI Appendix, Fig S15).

We hypothesised the same approach could be used for nucleic acid delivery. We employed the pan-nucleic acid interaction sequence LK15 and synthesised GET-LK15 peptides (Fig 6). After charge ratio optimisation for each test nucleic acid we were able to demonstrate significant transfection activity for P21-LK15-8R for plasmid DNA (pDNA; transfecting SIN-GFP), modified nucleotide mRNA (transfecting GFP modRNA (20)) and siRNAs (FAM-labelled GAPDH siRNA). The transfection efficiencies of optimised protocols were similar to Lipofectamine 2000 (LIPO 2000; Invitrogen) and GET-transfection retained activity in serum-containing transfections in which LIPO2000 was significantly inhibited. Colloidal stability of GET peptide/nucleic acid particles remained with addition of serum demonstrating no loss of stability (by aggregation) or efficiency was lost due to a serum-rich environment.

To extend our demonstration of the GET system to large cargoes (the largest demonstrated here being ~150kDa IgG) we attempted to enhance delivery of magnetic nanoparticles (MNPs) by covalently coupling P21-8R peptide to commercially available MNPs (coupled to Nanomag-D

250nm MNPs through EDAC/NHS reaction with MNP COOH functional group; estimated 1.5-2.0 ug peptide/mg MNPs; ~400 peptide molecules/particle) (Fig 5b). Conjugation of P21-8R to MNP increased their size (255.4±1.6nm to 372.9±2.3nm) and charge (-33.7±0.5 to +19.3±0.4). In both serum- and serum-free (SFM) conditions MNP uptake by cells was significantly enhanced by P21-8R conjugation (assessed by Prussian blue staining) and particles retained colloidal stability. These data indicates that GET may be used for any choice of cargo dependent on providing an interaction or conjugation of P21-8R to the target.

Discussion

Using the synergistic combination of a PTD with a cell membrane binding peptide we aimed to improve intracellular targeting of proteins and other cargoes. With this approach, we have developed a technology that enables highly efficient delivery of functionally relevant proteins to direct a variety of cell behaviours even in hard to transduce cell-types (see FigS17-19 for enhancement of delivery with GET, viability with high protein concentrations and repeated incubation, and viability after cell-uptake of a variety of clinically relevant cell types, respectively). High concentrations of GET peptide are not required to deliver functional amounts of cargo into cells; this is unlike cationic CPP technologies which are primarily driven to transduce by high extracellular concentrations. We demonstrated that the GET system can be harnessed to promote survival, self-renewal or direct the differentiation of pluripotent cells toward a desired lineage. Furthermore other proteins can be delivered by coupling to GET peptides (such as antibodies) and this can be extended to other chemically distinct targets such as nucleic acids or MNPs. This system is not technically complex, as for modified RNA systems

(20), and offers several key advantages over established techniques to deliver the exogenous function of a gene, protein or delivering other reagents. Fundamentally, because GET is protein based, it completely eliminates the risk of genomic integration and insertional mutagenesis inherent to all DNA-based methodologies (1). Moreover if endosomal escape can be improved, our approach will allow protein stoichiometry to be tightly regulated within cells. This will avoid stochastic variation in expression typical of integrating vectors, as well as the uncontrollable effects of viral silencing. Importantly the GET peptides do elicit some endosomal escape, this could be due to the net positive charge and a mechanism similar to that of previously described CPPs. It is also likely that the GET peptides will be degraded both before and after endosomal escape as for other CPPs.

Given the significant function of recombinant proteins delivered with our methodology it is possible that GET technology may also be directly applied to reprogramming and programming approaches. As there are stepwise phenotypic changes observed during pluripotency induction (28, 29) and in directed-differentiation protocols (30), it seems likely that individual transcription factors play distinct, stage-specific roles. The unprecedented potential for temporal control over individual factor function afforded by GET technology should enable these variables to be tested to improve efficiency and kinetics of cell-fate control. Furthermore the delivery of DNA, RNA, siRNA and MNPs by GET enables its use for genome-editing, gene knockdown and for physical manipulation or imaging of cells, respectively.

The fact we isolated heparan sulfate binding domains/proteins as a PTD-enhancer is an important finding. Though GAGs have long been considered important in the process of PTD-mediated transduction, their exact role is presently a point of contention with the most recent hypothesis being that GAG is merely important by directly binding to PTDs thereby enhancing lipid-bilayer

translocation but not being necessary for its occurrence (8). We also demonstrated that HBDs can elicit this function without themselves intrinsically being positively charged, however an overall positive charge is required provided by the PTD to mediate the transduction. This makes GET distinct from simple cationic CPP-based approaches. Using non-cationic CPPs with HBDs may completely negate the requirement for an overall positive charge for transduction. Importantly the effect of HBDs on PTDs and its cargo is not simple 'super-charging' and therefore can be considered distinct from other attempts to use such technologies (4, 5).

Furthermore, the demonstration that a variety of HBDs can target different cell-types more specifically (SI Appendix, Fig S7) adds a further advantage to GET. Presently there are >100 detailed GAG-binding proteins (31), these natural or synthetic peptides screened to specifically bind heparan sulfate molecules presented on particular cells will allow GET to be adapted to tightly and efficiently transduce target cells.

Our discovery that promoting heparan sulfate interaction significantly and synergistically improves the efficiency of PTD-tagged biologically active, macromolecular cargo will open up new avenues for the treatment, imaging and experimental investigation of disease. We believe that our improvement of the long-established PTD technologies has the potential to become a major enabling technology for cell-based therapies and regenerative medicine.

SI Materials and Methods available online

Materials and Methods

Recombinant protein expression, cell culture and transduction of cargoes by GET

Recombinant proteins were expressed as GST-tagged proteins and cleaved using PreScission ™ Protease (GE Healthcare, UK) during purification. Cell culture techniques have been previously described (32) and further detailed in SI Materials and Methods. Proteins were transduced into cells by addition to culture media for the specific cell line. Cells were harvested by trypsinization unless otherwise stated and analysed for fluorescence intensity by flow cytometry or for gene expression by real-time PCR. Transfection of nucleic acids or transduction of antibody complexes was achieved by complexation of the peptide with the cargo in OptiMEM (Invitrogen, UK). Peptides were covalently attached to MNPs presenting COOH groups (Nanogmag-D, MicroMod) with EDAC/NHS chemistry.

Statistical Analysis

Statistical comparisons were carried out using the GraphPad Prism software package. Comparisons were made using one-way Tukey-Kramer analysis of variance (ANOVA). Results were considered significant if $p<0.05$. Experiments were completed six times (n=6), and data depicts mean values (six replicates of duplicates) with standard deviation (s.d.) or for QPCR with standard error of the mean (s.e.m.).

Acknowledgements

We would like to thank Dr. Andrew D. Johnson (University of Nottingham) and Dr. Catherine Merry (University of Manchester) for helpful discussions. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement 227845. J.E.D and

K.M.S acknowledge the support of the UK Regenerative Medicine Platform (UKRMP). C.D. is supported by Engineering and Physical Sciences Research Council (EPSRC), Medical Research Council (MRC), British Heart Foundation (BHF Centre for Regenerative Medicine and Programme Grant), Heart Research UK and National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs).

Author Contributions. J.E.D and K.M.S conceived and initiated the project. J.E.D, G.O, G.E.M, H.M, M.R, Z.B and designed and performed experiments. J.E.D, C.D, A-E.H and K.M.S supervised the project. J.E.D and K.M.S wrote the paper.

Competing Financial Interests. The authors declare no competing financial interests.

Figure Legends

Figure 1 P21 improves PTD-mediated cellular uptake. **(a)** Schematic of the proteins created after screening domains which improve efficiency of protein delivery to cells. mR and mR-8R are described in Fig 1. P21-mR is mRFP with an *N*-terminal fusion of the P21 domain of heparin-binding EGF (HB-EGF). P21-mR-8R is mRFP with N-terminal fusion of P21 and C-terminal fusion of 8R. **(b)** Fusion of P21 to mR-8R significantly improves uptake into NIH3t3 cells. Fluorescence microscopy images of NIH3t3 cells treated with proteins (20µg/ml) for twelve hours in standard media conditions. Scale bar, 100µm. **(c)** P21-mR-8R is efficiently taken into human and mouse embryonic stem cells (HUES7 and CGR-8, respectively) and human induced pluripotent stem cells (IPS2) and mouse cardiomyocyte cell line HL1. Flow cytometry analyses of the mR-8R-inefficiently delivered cell lines treated with proteins mR-8R (20µg/ml) for twelve

hours. **(d)** P21-mR-8R initially strongly interacts with cell membranes and progressively is taken up be localised perinuclearly. Fluorescence (top) and confocal laser scanning microscopy (bottom) images of NIH3t3 cells treated with P21-mR-8R (20µg/ml) for either 1 hour, 1 hour with washes and a further 5 hours incubation (in serum-free media) or 6 hours treatment. Cells were pre-incubated for 1 hour in serum-free media, transduced for the desired time in serum-free media. Scale bars, 50µm (top) and 10µm (bottom). **(e)** Enhancement of cellular uptake mediated by P21 and 8R are affected by Trypsin proteolysis. Flow cytometry analyses NIH3t3 cells treated with proteins (20µg/ml) for 1 hour and a further 5 hour incubation (in serum-free media), with or without 10min pre-digestion with Trypsin or treatment with non-proteolytic cell dissociation solution (CDS). Cells were pre-incubated for 1 hour in serum-free media, treated with Trypsin and transduced for 1 hour in serum-free media. **(f)** Cell surface interaction of P21-containing proteins is disrupted by Tritonx100 treatment. Flow cytometry analyses NIH3t3 cells treated with proteins (20µg/ml) for 1 hour and a further 5 hour incubation (in serum-free media) with 10 min pre-treatment of PBS or PBS containing 0.1% (v/v) Tritonx100 (Tx100). Cells were pre-incubated for 1 hour in serum-free media, treated with PBS or PBS with Tx100 and transduced for 1 hour in serum-free media. Error bars indicate s.d. n=6.

Figure 2 GET of Cre Recombinase. **(a)** Schematic of the construct created to mark Cre activity in cells. Cre-mediated excision of a transcriptional STOP region flanked by *lox*P sites induces the constitutive expression of eGFP. Pr, promoter; βGal, β-galactosidase; Neo, Neomycin phosphotransferase. The NIH3t3 LSL-eGFP cell line was created by transfection and selection of NIH3t3 cells. **(b)** eGFP expression in untreated NIH3t3 LSP-eGFP cells or those transduced with SIN Cre lentivirus. Left shows fluorescence microscopy and right shows flow cytometry

histogram of eGFP expression. Scale bar, 50µm **(c)** Scheme of testing transduction of Cre activity in NIH3t3 LSL-eGFP cells. Cells were transduced with Cre proteins for 1 hour, washed and cultured for 2 days before analyses. **(d-e)** P21-mR-Cre-8R is efficiently transduced and recombines DNA. **(d)** Fluorescence microscopy images Cre-transduced NIH3t3 LSL-eGFP with the variety of dosages. Scale bar, 50µm. **(e)** Flow cytometry analyses of NIH3t3 LSL-eGFP cells transduced for 1 hour with mR-Cre, mR-Cre-8R and P21-mR-Cre-8R at a variety of dosages (0, 1, 10, 100 and 500µg/ml), washed and cultured for 2 days. Graph shows % recombination (i.e. % of eGFP +ve from total cell population). Error bars indicate s.d. n=6.

Figure 3 GET of NANOG promotes the self-renewal of mouse embryonic stem cells. **(a)** Scheme of testing activity of transduced NANOG in CGR-8 cells. Cells were transduced with P21-mR-NANOG-8R proteins (0, 1, 10 and 50µg/ml) for three consecutive days (1 passage, 1:3 split), passaged 1:3 and plated into growth media with P21-mR-NANOG-8R but lacking LIF (-LIF). Cells were fed daily with –LIF media containing P21-mR-NANOG-8R and passaged 1:3 every 3 days for 2 passages (a total of 3 passages –LIF) **(b)** P21-mR-NANOG-8R rescues self-renewal of mESCs lacking LIF dose dependently. Alkaline phosphatase (AP) staining of CGR-8 cells treated with P21-mR-NANOG-8R proteins and LIF withdrawal. AP activity and colony morphology is retained in CGR-8 cells cultured in LIF or without LIF but supplemented with SIN NANOG (to overexpress NANOG) or transduced with P21-mR-NANOG-8R. Scale bar, 100µm. **(c)** P21-mR-NANOG-8R maintains the proliferation of mESCs lacking LIF dose dependently. Percentage of the number of CGR-8 cells cultured without LIF verses those with LIF (% -LIF/+LIF) at passaging. In LIF-deficient CGR-8 cultures proliferation is promoted when supplemented with SIN NANOG (to overexpress NANOG) or transduced with P21-mR-

NANOG-8R. Error bars indicate s.d. **(d)** NANOG-dependent rescue in LIF-deficient cultures generates a more epiblast-like gene expression profile. Relative gene expression analyses of LIF-deficient CGR-8 cultures using quantitative PCR (QPCR). Cultures supplemented with SIN NANOG (to overexpress NANOG) or transduced with P21-mR-NANOG-8R have increased *Fgf5* expression, reduced *Rex1* expression and retain *Oct4* expression. Error bars indicate s.e. n=6.

Figure 4 GET of MYOD promotes Myogenic differentiation of human embryonic stem cells. **(a)** Scheme of testing the differentiation activity of transduced MYOD in HUES7 cells. HUES7 cells were plated onto gelatinised plastic and cultured in DMEM containing 10% (v/v) FCS. Cells were fed daily with DMEM containing 10% (v/v) FCS and P21-mR-MYOD-8R (0, 1, 5, 10 or 50µg/ml) for 7 days. Media was then changed to DMEM containing 2% (v/v) horse serum (HS), human recombinant insulin and P21-mR-MYOD-8R and fed daily for 3 days. **(b-f)** P21-mR-MYOD-8R drives myogenic differentiation of HUES7 cells to multinucleated Myotubes. **(b)** Light microscopy of HUES7 cells cultured under the myogenic regime supplemented with SIN MYOD (to overexpress MYOD) or transduced with P21-mR-MYOD-8R. Elongated fused Myotubes and single myocytes are generated with SIN-MYOD or high doses of P21-mR-MYOD-8R. Scale bar, 100µm. **(c)** MYOD-dependent myogenic differentiation of human embryonic stem cells. Relative gene expression analyses of HUES7 cultures using quantitative PCR (QPCR). Cultures supplemented with SIN MYOD (to overexpress MYOD) or transduced with P21-mR-MYOD-8R have increased endogenous *MYOD* expression and skeletal muscle-specific *ACTA1* expression. Error bars indicate s.e. **(d-e)** P21-mR-MYOD-8R differentiated cells are multinucleated. **(d)** Quantitation of mean nuclei number per cell using PI staining. Error bars

indicate s.d. **(e)** Fluorescence microscopy images of HUES7 cells differentiated with P21-mR-MYOD-8R (50µg/ml) and stained with nuclear dye DAPI. Scale bar, 50µm. **(f)** P21-mR-MYOD-8R differentiated cells are MYOGENIN positive. Quantitation of the percentage MYOGENIN positive cells using immunolabelling. Error bars indicate s.d. n=6.

Figure 5 GET of Antibodies and Nanoparticles. **(a)** GET of biotinylated cargoes using monomeric streptavidin (mSA2). **(i)** Schematic of the mSA2 proteins engineered to bind to and enable the uptake of biotinylated cargoes. We used P21-8R as a non-interacting control, mSA2 as a non-transducing control, and P21-mSA2-8R as the test protein. **(ii)** Schematic of the antibody (Ab) complexes of a biotinylated primary (1˚) antibody (Goat anti-rabbit; GtαRb) bound to an FITC-conjugated secondary (2˚) antibody (Rabbit anti-mouse; Rb αMu) used to test activity. **(iii)** GET-delivery of Ab complexes were visible by fluorescence microscopy (scale bar, 50µm). With co-incubation of P21-mSA2-8R (10µg/ml, bottom image), Ab complexes were efficiency delivered to cells **(iv)** Flow cytometry demonstrating that 1˚/2˚ Ab complexes (1µg/ml) are taken into NIH3t3 cells poorly by direct incubation or when co-incubated with mSA2 only. **(c)** GET of Magnetic Nanoparticles. **(i)** Schematic of the P21-8R peptide synthesised and test magnetic nanoparticles (MNPs). We tested 250nm Nanomag-D dextran shell/iron oxide core MNPs and conjugated P21-8R peptide to surface COOH groups. **(ii)** MNPs are taken into NIH3t3 cells most efficiently in serum-free media (SFM; left panel). Light microscopy images of Prussian blue iron stained NIH3t3 cells treated with MNPs (50µg/ml) for twelve hours in standard media conditions (10% FCS) or SFM. Conjugation of P21-8R to MNPs significantly increases cellular uptake in both 10% FCS and SFM conditions (circular image is of entire well, scale bar, 100µm). n=6.

Figure 6 GET of Nucleic acids. **(i)** Schematic of the LK15 proteins engineered to bind to and transduce nucleic acids. **(ii)** Transfection of human mesenchymal stem cells (iHMSCs) using GET-LK15. Initially we assessed binding capacity of LK15 peptides for plasmid (p)DNA (SIN GFP, to express GFP on transfection), modified synthetic messenger RNA (modRNA) (Miltenyi Biotech; to express GFP on transfection) and small-inhibitory (si)RNAs (labelled with FAM fluorophore to detect delivery). After optimising ratios we transfected iHMSCs with P21-LK15-8R and pDNA (10µg), modRNA (10µg) or siRNA (1µg) and visualised transfection by fluorescence microscopy (scale bar, 100µm). **(iii)** Quantification of GET-LK15 transfection of iHMSCs by flow cytometry (% transfection efficiency or relative fluorescence for siRNA) compared to lipofectamine (LIPO)2000 as a commercial standard. Error bars indicate s.d.

References

1. **Gump JM & Dowdy SF (2007) TAT transduction: the molecular mechanism and therapeutic prospects. *Trends in molecular medicine* 13(10):443-448.**
2. **El-Andaloussi S, Holm T, & Langel U (2005) Cell-penetrating peptides: Mechanisms and applications. *Curr Pharm Design* 11(28):3597-3611.**
3. **Meade BR & Dowdy SF (2007) Exogenous siRNA delivery using peptide transduction domains/cell penetrating peptides. *Adv Drug Deliver Rev* 59(2-3):134-140.**
4. **Thompson DB, Cronican JJ, & Liu DR (2012) Engineering and identifying supercharged proteins for macromolecule delivery into mammalian cells. *Methods Enzymol* 503:293-319.**
5. **Cronican JJ*, et al.* (2011) A class of human proteins that deliver functional proteins into mammalian cells in vitro and in vivo. *Chemistry & biology* 18(7):833-838.**
6. **Nakase I, Takeuchi T, Tanaka G, & Futaki S (2008) Methodological and cellular aspects that govern the internalization mechanisms of arginine-rich cell-penetrating peptides. *Adv Drug Deliver Rev* 60(4-5):598-607.**
7. **Heitz F, Morris MC, & Divita G (2009) Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. *Brit J Pharmacol* 157(2):195-206.**
8. **Gump JM, June RK, & Dowdy SF (2010) Revised Role of Glycosaminoglycans in TAT Protein Transduction Domain-mediated Cellular Transduction. *J Biol Chem* 285(2):1500-1507.**
9. **Meier O*, et al.* (2002) Adenovirus triggers macropinocytosis and endosomal leakage together with its clathrin-mediated uptake. *J Cell Biol* 158(6):1119-1131.**
10. **Conner SD & Schmid SL (2003) Regulated portals of entry into the cell. *Nature* 422(6927):37-44.**
11. **Kim D*, et al.* (2009) Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* 4(6):472-476.**

12. Zhou HY, *et al.* (2009) Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins (vol 4, pg 381, 2009). *Cell Stem Cell* 4(6):581-581.
13. Wadia JS, Stan RV, & Dowdy SF (2004) Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat Med* 10(3):310-315.
14. Skehel JJ, Cross K, Steinhauer D, & Wiley DC (2001) Influenza fusion peptides. *Biochem Soc T* 29:623-626.
15. Han X, Bushweller JH, Cafiso DS, & Tamm LK (2001) Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nat Struct Biol* 8(8):715-720.
16. Sakuma T, Higashiyama S, Hosoe S, Hayashi S, & Taniguchi N (1997) CD9 antigen interacts with heparin-binding EGF-like growth factor through its heparin-binding domain. *Journal of biochemistry* 122(2):474-480.
17. Higashiyama S, Abraham JA, & Klagsbrun M (1993) Heparin-Binding Egf-Like Growth-Factor Stimulation of Smooth-Muscle Cell-Migration - Dependence on Interactions with Cell-Surface Heparan-Sulfate. *J Cell Biol* 122(4):933-940.
18. Thompson SA, *et al.* (1994) Characterization of Sequences within Heparin-Binding Egf-Like Growth-Factor That Mediate Interaction with Heparin. *J Biol Chem* 269(4):2541-2549.
19. Yu JY, Chau KF, Vodyanik MA, Jiang JL, & Jiang Y (2011) Efficient Feeder-Free Episomal Reprogramming with Small Molecules. *Plos One* 6(3).
20. Warren L, *et al.* (2010) Highly Efficient Reprogramming to Pluripotency and Directed Differentiation of Human Cells with Synthetic Modified mRNA. *Cell Stem Cell* 7(5):618-630.
21. Takahashi K, *et al.* (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5):861-872.
22. Dixon JE, *et al.* (2010) Axolotl Nanog activity in mouse embryonic stem cells demonstrates that ground state pluripotency is conserved from urodele amphibians to mammals. *Development* 137(18):2973-2980.
23. Chambers I, *et al.* (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 113(5):643-655.
24. Do Kwon Y, *et al.* (2005) Cellular manipulation of human embryonic stem cells by TAT-PDX1 protein transduction. *Mol Ther* 12(1):28-32.
25. Hidema S, Tonomura Y, Date S, & Nishimori K (2012) Effects of protein transduction with intact myogenic transcription factors tagged with HIV-1 Tat-PTD (T-PTD) on myogenic differentiation of mouse primary cells. *J Biosci Bioeng* 113(1):5-11.
26. Liang QL, Mo ZY, Li XF, Wang XX, & Li RM (2013) Pdx1 protein induces human embryonic stem cells into the pancreatic endocrine lineage. *Cell Biol Int* 37(1):2-10.
27. Bichsel C, *et al.* (2013) Direct Reprogramming of Fibroblasts to Myocytes via Bacterial Injection of MyoD Protein. *Cell Reprogram* 15(2):117-125.
28. Chan EM, *et al.* (2009) Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nat Biotechnol* 27(11):1033-U1100.
29. Smith KP, Luong MX, & Stein GS (2009) Pluripotency: Toward a Gold Standard for Human ES and iPS Cells. *J Cell Physiol* 220(1):21-29.
30. Burridge PW, *et al.* (2011) A Universal System for Highly Efficient Cardiac Differentiation of Human Induced Pluripotent Stem Cells That Eliminates Interline Variability. *Plos One* 6(4).
31. Esko JD & Linhardt RJ (2009) Proteins that Bind Sulfated Glycosaminoglycans. *Essentials of Glycobiology*, eds Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, & Etzler MECold Spring Harbor (NY)), 2nd Ed.

32. Dixon JE*, et al.* (2014) Combined hydrogels that switch human pluripotent stem cells from self-renewal to differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 111(15):5580-5585.