CENTRO INTERNACIONAL DE ESTUDOS
DE DOUTORAMENTO E AVANZADOS
DA USC (CIEDUS)

TESIS DOCTORAL

NEW DESIGNS OF PEPTIDIC AND POLYMERIC VECTORS FOR PROTEIN AND GENE DELIVERY

Marisa Juanes Carrasco

ESCOLA DE DOUTORAMENTO INTERNACIONAL

PROGRAMA DE DOUTORAMENTO EN CIENCIA E TECNOLOXÍA QUÍMICA

SANTIAGO DE COMPOSTELA

2019

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AUTORIZACIÓN DEL DIRECTOR / TUTOR DE LA TESIS

New designs of peptidic and polymeric vectors for protein and gene delivery

D. ....Juan Ramón Granja Guillán.....

D. ....Javier Montenegro García.....

INFORMAN:

Que la presente tesis, se corresponde con el trabajo realizado por Dª. María Luisa Juanes Carrasco, bajo mi dirección, y autorizo su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como director de esta no incurre en las causas de abstención establecidas en la Ley 40/2015.

De acuerdo con el artículo 41 del Reglamento de Estudios de Doctorado, declara también que la presente tesis doctoral es idónea para ser defendida en base a la modalidad de COMPENDIO DE PUBLICACIONES, en los que la participación del doctorando/a fue decisiva para su elaboración.

La utilización de estos artículos en esta memoria, está en conocimiento de los coautores, tanto doctores como no doctores. Además, estos últimos tienen conocimiento de que ninguno de los trabajos aquí reunidos podrá ser presentado en ninguna otra tesis doctoral.

En Santiago de Compostela, ... de Septiembre de 2019

Fdo. Juan Ramón Granja Guillán

Fdo. Javier Montenegro García

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DECLARACIÓN DE LA AUTORA DE LA TESIS

New designs of peptidic and polymeric vectors for protein and gene delivery

Dña. ...Marisa Juanes Carrasco.....

Presento mi tesis, siguiendo el procedemento adecuado al Regulamento, y declaro que:

1) La tesis abarca los resultados de la elaboración de mi trabajo.
2) De ser el caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.
3) La tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.
4) Confirmo que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mi para la obteción de otros títulos.

En Santiago de Compostela, ... de Septiembre de 2019

Fdo. Marisa Juanes Carrasco

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AGRADECIMIENTOS

En numerosas ocasiones, la gente me pregunta si me he arrepentido alguna vez de haber hecho la tesis doctoral, y mi respuesta siempre es la misma, un NO rotundo. Ha sido una de las mejores decisiones que he tomado a lo largo de mi etapa de estudiante, primero porque me ha convertido en la química que soy hoy en día, y segundo, porque me ha permitido conocer a grandes personas, sin ellas, estos cinco años de tesis no habrían sido lo mismo. Por tanto, me gustaría mostrar mi agradecimiento a todos los que habéis hecho esta etapa posible.

En primer lugar, me gustaría dar las gracias a uno de mis directores, Juan Granja, gracias por tus sabios consejos, tu cercanía, tu ayuda y tu confianza en mí a la hora de tomar alguna decisión relacionada con el día a día del laboratorio.

A mi segundo director, Javier Montenegro, que sin conocerme, decidiste confiar en mí para empezar con vosotros la tesis doctoral, es algo por lo que siempre te estaré agradecida. Todos los logros que he podido alcanzar a lo largo de estos años, han sido gracias a que has conseguido contagiarme tu pasión por la química y tu positividad para poder afrontar nuevos proyectos, haciendo que haya ganado seguridad en mí misma.

A todos los organismos e instituciones nacionales e internacionales que han financiado esta tesis doctoral y todos los proyectos de investigación en los que he podido participar. Al Ministerio de Economía y Competitividad por haberme concedido una beca FPI que me ha ayudado a desempeñar mi trabajo a lo largo de estos cinco años, además de haberme permitido realizar una breve estancia en el Instituto Curie de Paris, mi ciudad preferida en el mundo.

A la Prof. Patricia Bassereau, por acogerme durante tres meses en su grupo de investigación en el Instituto Curie y darme la oportunidad de aprender y conocer cómo se trabaja en centros profesionales fuera de España. Al Dr. Ajay Mahalka por enseñarme tanto en tan poco tiempo.

A los doctores Manuel Amorín, Rebeca García-Fandiño, Julián Bergueiro y al Prof. Luis Castedo por vuestros consejos y discusiones científicas en los seminarios de grupo, que sin duda me han ayudado en el avance de mis proyectos. A Patricia Lago, por tu amabilidad y por tu ayuda a la hora de resolver cualquier gestión relacionada con el progreso de mi tesis doctoral.

A Irene, por enseñarme toda la biología que sé hoy en día, por ayudarme siempre que lo he necesitado, incluso encontrando un hueco que dedicarme cuando más ocupada podías estar, por escucharme y aconsejarme, por sorprenderme cada día con tu sabiduría y por grandes momentos de risas en el laboratorio, GRACIAS.

A Iván, quien a lo largo de estos años, además de haber sido compañero de laboratorio y compinche de numerosas aventuras fuera de él, has conseguido ser una persona muy importante para mí y la amistad que ha nacido y ha ido creciendo entre nosotros no tiene igual.

A Marta, Giulia, Eva, Victoria (Vicky para nosotros), Fede y Alicia, las locuras, las salidas y las conversaciones con vosotras serán inolvidables, gracias por escucharme y comprenderme y por estar a mi lado en los buenos y en los malos momentos.

A Héctor y Alfonso, mis farmacéuticos favoritos, gracias por hacer que los días dentro y fuera del laboratorio hayan sido más llevaderos y alegres con vuestro exquisito sentido del humor y múltiples bromas, no cambiéis nunca.

A Nacho y Lamas, los tenistas del grupo, me siento muy afortunada por haber compartido mis últimos años de tesis con vosotros, estoy segura al 100%, para Nacho al 99%, que sin vosotros no habrían sido tan especiales.

A Jose, Ale y Ángel, por los buenos momentos que me habéis hecho pasar, siempre dispuestos a ayudarme cuando lo he necesitado.

A Geert, Richard y Ghibom, gracias por darme la oportunidad de conoceros y poder aprender de vosotros.

A Juanillo, Lionel, Arcadio y Nuria, mis compañeros de inicio de tesis, gracias por haber hecho que mi aterrizaje en el laboratorio fuese tan agradable, gracias por grandes momentos de risas y canciones y por animarme en numerosas ocasiones.

A los que empezáis ahora la tesis, Sandra, Rebeca, Marcos y Ezequiel, mucho ánimo y no desesperéis si las cosas no salen a la primera, a veces un fracaso te abre la puerta hacia algo mejor.

Tal vez cada uno de nosotros por separado seamos capaces de alcanzar alguna meta importante, sin embargo, hemos demostrado que trabajando juntos sumamos, habiendo sido capaces de crear lo que para mi es una gran familia.

A David, Renata, Roberto, Damián y Borja, compañeros desde que empecé la carrera hace ya 10 años, cómo pasa el tiempo, gracias por vuestra amistad y amabilidad.

A mis amigas de toda la vida Natalia y Julia, las risas, las alegrías, los lloros y las penas, han estado presentes durante casi 15 años de amistad, años que sin vosotras habrían sido mucho más difíciles de llevar.

A Alberto, que aunque no lo sabías, nuestros caminos ya se habían cruzado mucho antes de haber empezado la carrera, y la química, quiso que de nuevo se volviesen a cruzar. Contigo he compartido momentos que me han hecho crecer como persona, gracias por escucharme siempre, por estar ahí, y por ser el motor que a veces necesito.

Ya por último, pero no por ello menos importante, me gustaría dedicarle unas palabras de agradecimiento a mi familia, sin ella, haber llegado hasta aquí habría sido imposible. Gracias por enseñarme que con esfuerzo y dedicación se puede conseguir todo lo que uno se proponga. Gracias por todo el apoyo y por todos los consejos que me han enseñado cuál era el mejor camino a seguir y que han hecho de mí la persona que soy hoy en día. Gracias porque sé que cuando algo pueda salir mal, siempre vais a estar ahí para mí.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ABBREVIATIONS

6Ahx: 6-aminohexanoic acid
Aa: amino acid
AFM: Atomic Force Microscope
Arg: Arginine
ASO: antisense oligonucleotide
Au NP: gold nanoparticle
Boc: *tert*-Butoxycarbonyl
BSA: Bovine Serum Albumin
Calcd: Calculated
CART: charge-altering releasable transporter
CDP: cyclodextrin-containing cationic polymers
CF: Carboxyfluorescein
CLSM: Confocal laser scanning microscopy
ConA: Concanavalin A
CPP: Cell-Penetrating Peptide
CRISPR: clustered regularly interspace short palindromic repeats
crRNA: CRISPR RNA
DCM: Dichloromethane
DIEA: N,N-Diisopropylethylamine
DLS: dynamic light scattering
DMAP: 4-Dimethylaminopyridine
DMEM: Dulbecco's Modified Eagle Medium
DMF: N,N-Dimethylformamide
DMSO: Dimethylsulfoxide
DOTAP: 1,2-dioleoyl-3-trimethylammonium-propane (chloride salt)
DOTMA: 2,3-dioleyloxypropyl-1-trimethylammonium chloride
EDC: 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA: Ethylenediaminetetraacetic acid
EGFP: enhanced green fluorescent protein
EIPA: 5-(N-Ethyl-N-isopropyl)amiloride
EPO: erythropoietin
EPOR: erythropoietin receptor
EYPC: Egg yolk phosphatidylcholine
Fab: fragment antigen binding
FBS: Fetal Bovine Serum
Fc: fragment crystallizable
Fmoc: 9-fluorenylmethoxycarbonyl
FRET: fluorescence resonance energy transfer
FSC: forward scatter channel

GFP: Green Fluorescent Protein
GR-GFP: Glucocorticoid receptor-Green Fluorescent Protein
GR: glucocorticoid receptor
gRNA: guide RNA
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HFIP: 1,1,1,3,3,3-Hexafluoro-2-propanol
HKR: HEPES-Krebs-Ringer buffer
HPRT1: hypoxanthine phosphoribosyltransferase 1
hpt: hours post-treatment
HRMS (ESI): High resolution mass spectrometry (electrospray ionization)
HSA: human serum albumin
IFN: interferon
Ig: immunoglobulin
IL: interleukin
iTOP: induced transduction by osmocytosis and propanebetaine
kb: kilobase
Lys: Lysine
mAb: monoclonal antibody
miRNA: micro RNA
mRNA: messenger RNA
MSN: mesoporous silica nanoparticle
MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
Mtt: 4-Methyltrityl
MβCD: Methyl-β-Cyclodextrin
N-HATU: *N*-[(Dimethylamino)-1H-1,2,3-triazolo[4,5-b]pyridin-1-ylmethylene]-*N*-methylmethanaminiun hexafluorophosphate *N*-oxide
N-HBTU: *N*-[(1H-Benzotriazol-1-yl)4-(dimethylamino)methylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide
NHS: N-Hydroxysuccinimide
NLS: nuclear localization signal
Ox-Dex: Dexamethasone Acid
Ox: Oxime
PAM: protospacer adjacent motif
PBA: phenylboronic acid
PBAE: poly(β-amino ester)
Pbf: 2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PBS: Phosphate buffered saline
pDNA: plasmid DNA
PDPA: poly(2-(diisopropylamino)ethyl methacrylate)
PEG: polyethylene glycol

PEI: polyethyleneimine
PI: propidium iodide
PIC: polyion complex
PLL: poly-L-lysine
PPI: protein-protein interaction
PTG: Isopropyl β-D-1-thiogalactopyranoside
R8: octaarginine
RGD: arginine-glycine-aspartic acid peptide
RIP: ribosome-inactivating protein
RNP: ribonucleoprotein
RP: Reverse Phase
siRNA: small interfering RNA
siRNN: short interfering ribonucleic neutral
*Sp*Cas9: *Streptococcus pyogenes* Cas9
SPPS: Solid Phase Peptide Synthesis
SSC: side scatter channel
T7E1: T7 endonuclease I
TAE: tris-acetate-ethylenediaminetetraacetic acid
TALENs: transcription activator-like effector nucleases
TAMRA: Carboxytetramethylrhodamine
TFE: Trifluoroethanol
TIS: Triisopropylsilane
TNBS: 2,4,6-Trinitrobenzenesulfonic acid
TNF: tumor necrosis factor
tracrRNA: *trans*-activating crRNA
VLP: virus-like nanoparticle
ZFNs: Zinc-finger nucleases.
Zn-PD: dipicolyamine-based disulfide-containing zinc (II) coordinative module

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

SUMMARY

The expansion of new therapeutic technologies is limited by the effective, selective and controlled transport and release of the corresponding biomolecules to the different tissues, cells and organelles. Therefore, there is a need to develop conceptually new strategies to transport and deliver proteins or nucleic acids inside cells. There are multiple synthetic materials for the intracellular delivery of exogenous biomolecules, such as, liposomes, nanoparticles, etc., but, the use of cell-penetrating peptides (CPPs) and polymers, has become a promising alternative for the protection, effective internalization and controlled release of biomacromolecules such as large protein complexes or nucleic acids.

Due to the high compartmentalization of biological systems, cells require a very efficient way to communicate and respond to external stimuli. This process is generally carried out by carbohydrate ligands and their specific cellular receptors, which are membrane proteins known as lectins. Therefore, membrane proteins, particularly lectins, and their carbohydrate ligands are optimal candidates for the selective cell targeting of drugs and probes. Among the group of carbohydrate-binding proteins, lectins are particularly relevant because they generally display multiple binding sites, which allow cooperativity and multivalent effects. They are considered to be the key to enter into cells and they are always overexpressed during tissue inflammation and cellular metastasis. The multivalent interactions of lectins with glycoconjugates strongly depend on the topological arrangement of the glycan moiety. Therefore, there is an opportunity for the development of molecular vehicles that selectively recognize lectins and that can be subsequently transported across the cell membrane.

In the first chapter of this thesis, we present the synthesis of biocompatible peptide scaffolds that incorporate specific protein ligands in their structure such as glycans, for the recognition and cellular internalization of lectins. The designed peptide is characterized for having a helical structure and an amphipathic behavior, defined by one hydrophilic cationic domain (arginines) and a hydrophobic domain (leucines). In orthogonal disposition the peptide contains two alkoxyamine residues for the attachment of the specific protein ligand (mannose for Concanavalin A (ConA) recognition). Interaction studies (SPR and fluorescence anisotropy) revealed a high affinity of the mannosylated peptide with its target protein, with values of dissociation constant in the range of 10 μM. After peptide/protein complexation, internalization experiments were performed in cell lines such as HeLa or HepG2. Moreover, in this chapter we present a wide variety of control peptide hybrids, which demonstrated that the presence of two protein ligands and the helical behavior are both required for ConA interaction and internalization. Moreover, selectivity studies of the intracellular delivery of a mixture of proteins (ConA and Streptavidin), where the mannosyl peptide is compared with Lipofectamine2000, revealed that only the peptide was able to discriminate between both proteins. Finally, the substitution of the glycan moiety of the peptide for a biotin ligand allowed the intracellular delivery of Streptavidin. Based on these results, this work exhibits the high versatility of our strategy, as the concept presented here could be adapted with the correct selection of ligands for the delivery of other proteins with the appropriate supramolecular binding motifs.

Having confirmed the versatility of CPPs for the intracellular delivery of proteins, we decided to go one step forward and start a more challenging project focused in gene therapy. Gene-based therapy is one of the most promising tools for the future of human health. Nucleic acids (siRNA, pDNA, mRNA, etc.) and gene-editing nucleases (CRISPR/Cas9, etc.) can be applied to repair the malfunction or absence of an essential protein inside cells.

The discovery of RNA guided endonucleases has emerged as one of the most important tools for gene edition and biotechnology. The selectivity and simplicity of the CRISPR/Cas9 strategy allow the straightforward targeting and edition of particular loci in the cell genome without the requirement of protein engineering. CRISPR is a family of DNA sequences in bacteria whose recognition forms the basis of the genome editing technology known as CRISPR/Cas9. This technology allowed scientists to make permanent modification to genes within organisms. However, the transfection of Cas9 is a challenging task due to its large size. CRISPR scientists typically transfect cells with plasmids containing instructions to make Cas9, which could lead to permanent DNA recombination and persistent expression, a situation that enhances potential off target effects and immunogenic responses. Therefore, the direct delivery of Cas9 ribonucleoprotein (RNP) constitutes an advantageous strategy for gene edition and other potential therapeutic applications of the CRISPR/Cas9 system. In recent reports, Lipofectamine and cell-penetrating peptides have been used for the direct delivery of Cas9, however, these strategies required Cas9 engineering and covalent fusion of the CPP to the protein, respectively.

Considering these precedents, in the second chapter of this memory, we have developed a supramolecular strategy for the direct delivery of Cas9 RNP, based on amphiphilic scaffolds prepared by hydrazone bond formation between a cationic peptide core and hydrophobic aldehyde tails. HeLa cells expressing EGFP were incubated with peptide/Cas9 complexes using a guide RNA against EGFP. The results of this experiment revealed that oleic aldehyde was one of the best candidates for the direct delivery of Cas9 RNP. Therefore, we decided to further study the potential of this technology using this particular aldehyde in three different cell lines (HeLa, A549 and DF1). Edition efficiency of the peptide was compared with lipofectamine, observing comparable efficiencies but lower toxicity when using the peptide. After the peptide/protein interaction, non-covalent nanoparticles of around 200 nm were formed. Mechanistic studies demonstrated that peptide/Cas9 complexes were internalized in cells *via* macropinocytosis. To the best of our knowledge, this was the first supramolecular strategy for the direct delivery of Cas9 using a penetrating peptide vehicle.

In the third chapter of this thesis we study the potential of previously described methodology of the group, for the delivery of messenger RNA (mRNA). The transfection of mRNA represents a challenge due to its low stability in cell media but, at the same time, it presents a great relevance because of the recent development of therapies based on the use of this biopolymer. In this direction, in collaboration with Prof. Francisco Fernández Trillo, polyhydrazide polymers functionalized with cationic aldehydes (t-guanidinium) and different hydrophobic aldehydes were used for mRNA internalization in Hek293 cells. For mRNA, it was necessary to use a higher molecular weight polymer than in our previous studies of

siRNA and pDNA transfection. Initially, internalization experiments using different hydrophobic tails revealed that myristoleic aldehyde presented the most satisfactory results. DLS studies showed nanoparticles of 50-100 nm in size, in addition to an increase of the net charge when the concentration of polyhydrazone was augmented. This methodology presented similar transfection efficiencies when compared with other polymers such as PEI, but it significantly improved the activity of peptides such as octaarginine and GALA, as well as the cationic lipid DOTAP. These results, in addition to the lack of toxicity, confirmed the excellent properties of polyhydrazones for the delivery of mRNA.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

RESUMEN

La expansión de nuevas tecnologías terapéuticas está limitada por el transporte efectivo, selectivo y controlado de biomoléculas de interés, a los diferentes tejidos, células y orgánulos. Por tanto, existe la necesidad de desarrollar conceptualmente nuevas estrategias para el transporte y entrega celular de proteínas y ácidos nucleicos al interior de las células. Existen múltiples materiales sintéticos para la entrega intracelular de biomoléculas exógenas como los liposomas, nanopartículas, etc., sin embargo, el uso de los péptidos penetrantes en células y los polímeros, ha surgido como una alternativa prometedora para la protección, la efectiva internalización y la controlada liberación de biomacromoléculas como proteínas complejas de gran tamaño o ácidos nucleicos.

Debido a la alta compartimentalización de los sistemas biológicos, las células requieren mecanismos eficientes para comunicarse y responder a estímulos externos. Este proceso es llevado a cabo generalmente por ligandos que contienen carbohidratos y sus receptores específicos celulares, que son proteínas de membrana conocidas con el nombre de lectinas. Por tanto, estas proteínas de membrana y sus respectivos ligandos son candidatos óptimos para la focalización selectiva de fármacos y sondas hacia el interior celular. Entre el grupo de proteínas de unión a carbohidratos, las lectinas son particularmente relevantes debido a que generalmente disponen de múltiples sitios de unión. Se consideran la llave para entrar a las células y siempre se sobreexpresan durante la inflamación de tejidos y la metástasis celular. Las interacciones multivalentes de las lectinas con glicoconjugados dependen en gran medida de la disposición del derivado de glicano. Por tanto, existe una oportunidad para el desarrollo de vehículos moleculares que selectivamente reconozcan las lectinas y por consiguiente sean capaces de transportarlas a través de la membrana celular.

En el primer capítulo de esta tesis doctoral, se plantea la síntesis de péptidos biocompatibles con ligandos específicos de proteínas como los carbohidratos para el reconocimiento e internalización celular de lectinas. El péptido diseñado se caracteriza por tener una estructura helicoidal y un carácter anfipático, definido por un dominio hidrofílico y catiónico (argininas) y un dominio hidrofóbico (leucinas). En disposición ortogonal el péptido presenta dos residuos de alcoxiamina para el anclaje específico de los ligandos de proteína (manosa para el reconocimiento de la Concanavalina A (ConA)). Estudios de interacción (SPR y anisotropía de fluorescencia) revelaron una elevada afinidad en la unión del péptido con manosas por su proteína objetivo, con valores de constante de disociación en el orden de 10 μM. Una vez formados los complejos, se estudió su internalización en líneas celulares como las células HeLa o HepG2. En este capítulo se presentan además una gran variedad de péptidos híbridos control, que demostraron que tanto la presencia de los dos ligandos de manosa, así como el carácter helicoidal, son requisitos indispensables tanto para la interacción como para la internalización de la ConA. Se presentan además estudios de selectividad en el transporte de una mezcla de proteínas (ConA y estreptavidina) al interior celular, donde se compara el péptido con dos ligandos de manosa con la Lipofectamina2000, que demostraron que el péptido de manosa era capaz de discriminar entre las diferentes proteínas.

Una vez confirmada la versatilidad de los péptidos penetrantes para el transporte intracelular de proteínas, se decidió ir un paso más allá y empezar a trabajar en un proyecto más complicado centrado en terapia génica. La terapia basada en genes es una de las herramientas más prometedoras para el futuro de la salud humana. Los ácidos nucleicos (ARNi, ADN, ARNm, etc.) y las nucleasas de edición genética (CRISPR/Cas9, etc.) pueden ser aplicados para la reparación del mal funcionamiento o ausencia de una proteína esencial en el interior celular.

El descubrimiento de las endonucleasas guiadas por ARN se ha convertido en una de las herramientas más importantes para la edición de genes y la biotecnología. La selectividad y la simplicidad de la estrategia CRISPR/Cas9 permite el transporte sencillo y la edición de lugares particulares en el genoma celular sin el requerimiento de la ingeniería de proteínas. CRISPR es una familia de secuencias de ADN en las bacterias cuyo reconocimiento forma la base de la tecnología de edición génica conocida como CRISPR/Cas9. Esta tecnología ha permitido a los científicos hacer modificaciones permanentes en los genes en el interior de los organismos. Sin embargo, la transfección de Cas9 es una tarea complicada debido a su gran tamaño. Los científicos que trabajan con CRISPR normalmente transfectan las células con plásmidos que contienen instrucciones para hacer Cas9 pero que podrían dar lugar a una recombinación de ADN permanente y a una expresión persistente del mismo, lo que aumenta los posibles efectos secundarios y las respuestas del sistema inmune. Por tanto, el transporte directo de la ribonucleoproteína (RNP) Cas9 constituye una estrategia ventajosa para la edición génica y otras aplicaciones potenciales terapéuticas del sistema CRISPR/Cas9. En publicaciones recientes, tanto la lipofectamina como los péptidos penetrantes han sido utilizados para el transporte directo de Cas9, sin embargo, estas estrategias requieren la ingeniería de Cas9 y la fusión covalente de los péptidos penetrantes a la proteína, respectivamente.

En el segundo capítulo, se desarrolló una estrategia supramolecular para el transporte directo de la RNP Cas9, basada en estructuras anfifílicas formadas por enlaces hidrazona entre un péptido catiónico y colas de aldehídos hidrofóbicos. Las células HeLa, que expresaban EGFP, se incubaron con complejos péptido/Cas9 utilizando una guía de ARN contra la EGFP. Los resultados de este experimento revelaron que el aldehído oleico era uno de los mejores candidatos para el transporte de la RNP Cas9, por lo que se procedió al estudio de la edición génica de Cas9 utilizando este vector particular en tres líneas celulares diferentes (Hela, A549 y DF1). La eficiencia de la edición se comparó con la de la lipofectamina mediante estudios de dosis respuesta, en los que se observó que las eficiencias eran comparables, sin embargo el péptido resultó ser menos tóxico. Después de la interacción péptido/proteína, se formaron nanopartículas no covalentes de aproximadamente 200 nm de tamaño. Estudios mecanísticos demostraron que los complejos de péptido/Cas9 se internalizaban en la célula por macropinocitosis. Esta es la primera estrategia supramolecular para el transporte directo de Cas9 usando péptidos penetrantes.

En el tercer capítulo de esta tesis, nos centramos en transporte de ARNm utilizando la metodología descrita con anterioridad por el grupo de investigación. La transfección de

ARNm representa un reto muy importante debido a su baja estabilidad en el medio celular, pero a la vez presenta una gran relevancia debido al reciente desarrollo de terapias basadas en la utilización de este biopolímero. En esta línea, en colaboración con el grupo de Francisco Fernández Trillo, se utilizaron polímeros de polihidrazida funcionalizados con aldehídos catiónicos (t-guanidinio) y diferentes aldehídos hidrofóbicos, para llevar a cabo la internalización de ARNm en células Hek293. Para este nucleótido fue necesario emplear un polímero con un peso molecular mucho más elevado que los utilizados en estudios previos de transfección de ARNi y plásmidos de ADN. Inicialmente se llevaron a cabo estudios de internalización utilizando diferentes colas hidrofóbicas, que revelaron que el aldehído miristoleico presentaba los resultados más satisfactorios. Estudios de DLS revelaron tamaños de partícula entorno a 50 y 100 nm, además de un aumento en la carga neta de los agregados al incrementar la concentración de polihidrazona en la mezcla. Esta metodología presentó una eficiencia similar a la de otros vehículos poliméricos como el PEI, pero superó notablemente el rendimiento de péptidos como la octaarginina y GALA, y el lípido catiónico DOTAP. Todos estos resultados sumados a la nula toxicidad confirmaron las excelentes propiedades de las polihidrazonas para el transporte de ARNm.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

GENERAL OBJECTIVES

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

It is a proven fact that the advancement of new therapeutic technologies, such as chemotherapy, gene therapy, etc. is impeded by the limitations of the delivery of the corresponding molecules and/or bioactive probes in an efficient, selective and controlled way to different tissues, cells and organelles. Different institutions in charge of managing and paying institutional research programs (Europe: ERC, USA: NIH, Japan: JSPS, etc.) have identified, as one of the priority objectives, the design and development of new transport and delivery platforms, which allow to reduce the current doses of drugs to minimize toxicity and side effects, and whose impact affects enormously to the quality of life of patients.

Among the multiple non-viral molecular transport strategies, cell-penetrating peptides (CPPs) are one of the most promising alternatives for the efficient internalization of biocompatible cargoes. Since the discovery of the first CPP, trans-activator of HIV transcription, the structure, properties and mechanism of action of penetrating peptides has been in constant evolution.[1] The majority of the most important international research groups paid attention to the extraordinary transport properties of these peptides.[2,3] Nowadays, although it is possible to define CPPs in a simple way as oligopeptides of cationic and amphiphilic character that has the ability to penetrate into the cells, diverse limitations have impeded to precisely understand the mechanism of action of these transporter peptides.[3] Additionally, CPPs suffer from other important limitations such as instability to protein digestion, in general, low selectivity and the toxicity associated with its highly cationic character that leads to the sequestration of heparin.[4] The meticulous analysis of the bibliography, allows us to understand the importance of the different topologies of CPPs during the different transport events across the membrane. Our starting hypothesis was based on the power of the secondary structure of the penetrating peptides to minimize the cationic charge and incorporate elements for cellular recognition.

Moreover, in addition to oligopeptides, polymers are also of great interest in macromolecules delivery, as can be easily produced in large scale displaying multivalent motifs required for cell internalization. However, standard polymerization reactions are normally carried out in the presence of organic solvents, and as consequence, isolation and purification steps are always required before screening for biological activity. Furthermore, in most of these examples, the monomer composition strongly influences the final structure of the products after the polymerization reaction, as they can result in polymers with different lengths and polydispersity indexes.[5]

---

[1] Stewart, K. M.; Horton, K. L.; Kelley, S. O. *Org. Biomol. Chem.,* **2008**, *6*, 2242–2255.

[2] Fischer, P. M. *Med. Res. Rev.,* **2007**, *27*, 755–795.

[3] Langel, Ü. *Cell-Penetrating Peptides*; Langel, Ü., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, **2015**; Vol. 1324.

[4] Koren, E.; Torchilin, V. P. *Trends Mol. Med.,* **2012**, *18*, 385–393.

[5] Lynn, D. M.; Anderson, D. G.; Putnam, D.; Langer, R. *J. Am. Chem. Soc.,* **2001**, *123*, 8155–8156.

In our research group, we have recently published the *in situ* modification of peptides[6] and polymers[7,8] scaffolds with different cationic and hydrophobic functional groups employing dynamic covalent bonds, such as hydrazones, for the delivery of nucleic acids. Framed in this strategy, the general objective of this work is to employ the dynamic covalent bond methodology, including oxymes in addition to hydrazones, to combine small peptide or polymer scaffolds with aldehydes, such as ligands for protein recognition, hydrophobic chains or cationic aldehydes for increasing membrane interaction, to obtain a library of cationic amphiphilic non-viral vectors for exploring protein and gene delivery (Figure 1). The main advantage of dynamic-covalent bonds is that they allow the preparation of these amphiphiles in fully biocompatible conditions without needing further purification steps. This permits the subsequent non-covalent combination with the corresponding cargo in a straightforward way. Furthermore, the dynamic character of this type of covalent bonds have the potential to be disassembled (i.e. hydrazone hydrolysis due to the pH), leading to the controlled release of the cargo inside cells, one of the limitations of the most commonly used delivery strategies nowadays.

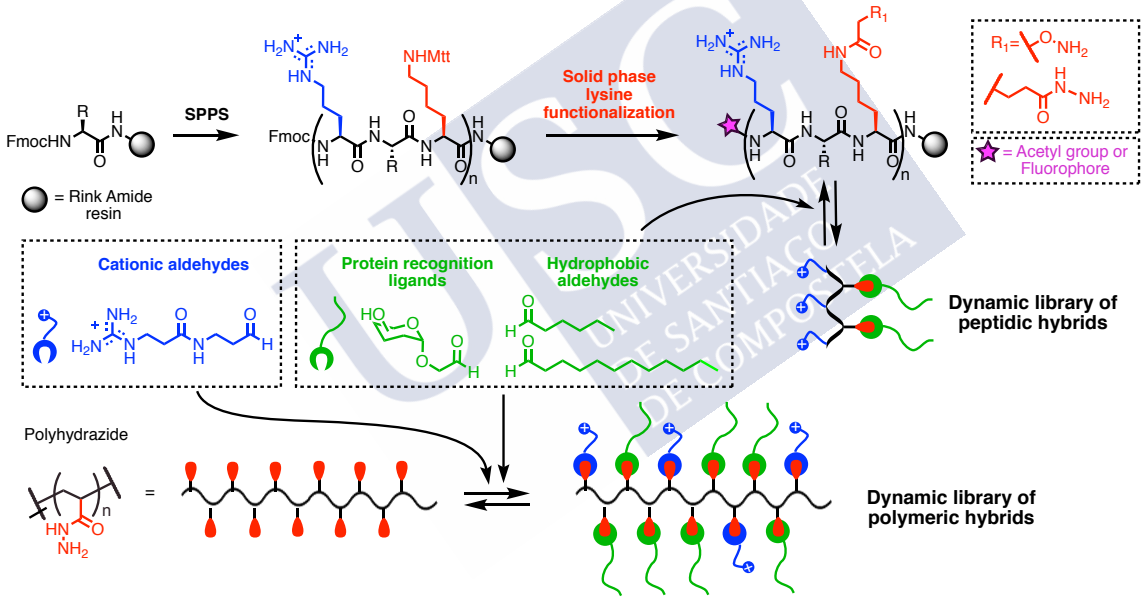


Figure 1. Design of peptidic and polymeric libraries employing dynamic-covalent bond methodology.

Dynamic-covalent bond methodology should allow us to prepare libraries of non-viral vectors with a high control of their structure, making slight variations over the peptide or polymer vehicle using a common scaffold for each case. This strategy leads to the fast identification of compounds with the ability of delivering biomolecules (proteins or nucleic acids) inside cells with high efficiency.

[6] Louzao, I.; García-Fandiño, R.; Montenegro, J. *J. Mater. Chem. B,* **2017**, *5*, 4426–4434.

[7] Priegue, J. M.; Crisan, D. N.; Martínez-Costas, J.; Granja, J. R.; Fernandez-Trillo, F.; Montenegro, J. *Angew. Chemie Int. Ed.,* **2016**, *55*, 7492–7495.

[8] Priegue, J. M.; Lostalé-Seijo, I.; Crisan, D.; Granja, J. R.; Fernández-Trillo, F.; Montenegro, J. *Biomacromolecules,* **2018**, *19*, 2638–2649.

CHAPTER 1: THERAPEUTIC PROTEINS

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1 INTRODUCTION

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1.1 Therapeutic interest of direct protein delivery

1.1.1 General interest of protein delivery

Proteins are one of the most physiological relevant biomolecules as they are key players in the regulation of numerous cellular activities, such as enzyme catalysis, signal transduction, gene regulation and apoptosis.[9] Their dynamic and diverse functions in the cellular machinery offer interesting challenges and opportunities in biology and medicine. Many diseases are consequence of the malfunction, mutation, reduced expression or any other abnormalities of cellular proteins. Thus, the direct delivery of functional proteins inside cells to fight diseases might be a key step in many medical applications, such as cancer therapy,[10] vaccination,[11] regenerative medicine,[12] etc. Protein based therapeutics, such as antibodies, enzyme replacement therapy,[13] receptor traps and cell surface ligands have changed the field of drug development and have the potential to impact all areas of medicine.[14] Protein-based therapy, due to their structural complexity, offers advantages over low molecular weight drugs, such as high specificity and complexity functions that cannot be mimicked by simple chemical compounds. Additionally, there is a reduced possibility of causing adverse effects, as long as the body can naturally produce many of these therapeutic proteins and, therefore, they are well tolerated. Finally, they can provide effective gene replacement treatment avoiding the risk of insertional mutagenesis.[15,16] From a financial point of view, the clinical development of protein therapeutics is faster than that of small-molecule drugs and, because of its exclusivity in form and function, companies currently pursue extensive patent protections.[16]

Furthermore, some of the proteins that have been targeted with this therapeutic approach are essential players in signaling pathways and, therefore, they are inaccessible to small-molecule inhibitors, rendering them "undruggable" targets. In many cases, they work facilitating protein-protein interactions (PPIs),[15] which are based on the contact of highly hydrophobic and relatively flat interfaces, making difficult the attachment and blocking of small molecules to these large surfaces. However, macromolecules such as other proteins or peptides could efficiently inhibit PPIs.[17] This is where protein-based therapy becomes relevant, connecting the gap between small molecules inhibitors and large protein targets.

The field of protein therapeutics has grown considerably since the introduction of the first therapeutic protein other than antibodies, the human insulin, which derived from recombinant

[9] Z. Gu, A. Biswas, M. Zhao and Y. Tang, *Chem. Soc. Rev.*, **2011**, *40*, 3638–3655.
[10] X. Liu, F. Wu, Y. Ji and L. Yin, *Bioconjug. Chem.*, **2019**, *30*, 305–324.
[11] H. Nakagami, *J. Cardiol.*, **2017**, *70*, 201–205.
[12] E. M. Green and R. T. Lee, *Physiol. Rev.*, **2013**, *93*, 311–325.
[13] M. Solomon and S. Muro, *Adv. Drug Deliv. Rev.*, **2017**, *118*, 109–134.
[14] Y. Zhang, J. J. Røise, K. Lee, J. Li and N. Murthy, *Curr. Opin. Biotechnol.*, **2018**, *52*, 25–31.
[15] S. Du, S. S. Liew, L. Li and S. Q. Yao, *J. Am. Chem. Soc.*, **2018**, *140*, 15986–15996.
[16] B. Leader, Q. J. Baca and D. E. Golan, *Nat. Rev. Drug Discov.*, **2008**, *7*, 21–39.
[17] D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug Discov.*, **2016**, *15*, 533–550.

DNA in 1982.[18] In recent years, transcription factors,[19] gene editing enzymes,[20] metabolic enzymes[21] and protein antigens[22] have shown potential to provide treatments for an extensive number of permanent diseases. Unfortunately, they still have some drawbacks that might be overcome before being developed into therapeutics. One of the main challenges is their delivery into the cytoplasm. The key barriers of delivering target proteins inside cells are related with some of their own properties, such as large sizes, variety in surface properties and unstable tertiary structures. In addition, there are other major challenges for the direct delivery of therapeutic proteins into mammalian cells. First of all, native proteins can be degraded or inactivated when administered into serum, leading to immunogenicity effects.[9] Then, the cell membrane can effectively block the entrance of macromolecules, requiring the modification of the protein cargo in order to make them permeable and selective to the correct tissue, something that is a very high cost process.[23] Most of these proteins are delivered into the cell by endocytosis and, therefore, once inside the cell, most of proteins stay trapped in endosomes suffering protease-mediated degradation and exocytosis. Consequently, their interaction with their cytosolic targets is restricted. Thus, the therapeutic efficacy is limited due to their poor ability to escape from the endosomes.[24] Finally, to transduce protein into cells it is also required, in addition to some genetic or chemical modifications of proteins to expose them to undesirable conditions (e.g. high temperature, pH, chemicals or organic solvents). All these modifications might affect protein stability and/or function.[25] Recently, a strategy that allows the chemoselective remodeling of protein surfaces has been reported, with which it is possible to reduce the anionic character while increasing the hydrophobic properties.[26] The protein surfaces are decorated with cationic (i.e. guanidinium, ammonium and imidazolium) and anionic groups (carboxylates) to increase protein stability. The strategy is based on the esterification of protein-carboxyl groups using diazo compounds providing a protein with the ability to access the cytosol.[26]

An alternative strategy for protein delivery into cells involves the use of a transfecting reagent capable of transporting them without the requirement of protein modification.[27] The use of carriers presents its own set of challenges such as, the formation of electrostatic and hydrophobic interactions with the carrier to achieve the optimal efficiency with minimal

[18] I. S. Johnson, *Science*, **1983**, *219*, 632–637.

[19] M. G. Peterson and V. R. Baichwal, *Trends Biotechnol.*, **1993**, *11*, 11–18.

[20] D. A. Scott and F. Zhang, *Nat. Med.*, **2017**, *23*, 1095–1101.

[21] P. Mane and V. Tale, *Int.J.Curr.Microbiol.App.Sci.*, **2015**, *4*, 17–26.

[22] P. J. Tacken, B. Joosten, A. Reddy, D. Wu, A. Eek, P. Laverman, A. Kretz-Rommel, G. J. Adema, R. Torensma and C. G. Figdor, *J. Immunol.*, **2014**, *180*, 7687–7696.

[23] N. J. Yang and M. J. Hinner, *Methods Mol. Biol.*, **2015**, *1266*, 29–53.

[24] M. P. Stewart, A. Sharei, X. Ding, G. Sahay, R. Langer and K. F. Jensen, *Nature*, **2016**, *538*, 183–192.

[25] A. Fu, R. Tang, J. Hardie, M. E. Farkas and V. M. Rotello, *Bioconjug. Chem.*, **2014**, *25*, 1602–1608.

[26] K. A. Mix, J. E. Lomax and R. T. Raines, *J. Am. Chem. Soc.*, **2017**, *139*, 14396–14398.

[27] A. Bolhassani, B. S. Jafarzade and G. Mardani, *Peptides*, **2017**, *87*, 50–63.

[9] Z. Gu, A. Biswas, M. Zhao and Y. Tang, *Chem. Soc. Rev.*, **2011**, *40*, 3638–3655.

toxicity, the correct amount of protein, and the type of cells used for the internalization and their release into the cytoplasm.

1.1.2 Potential therapeutic protein targets

Proteins have the most active and varied role of any macromolecule in the body. They catalyze reactions, display recognition and molecular transport functions at the membrane, offer intracellular and extracellular scaffolding support and carry molecules within a cell or from one cell to another.[16] Among of the functions, it is estimated that the human genome has between 25000 and 40000 different genes, though the gene splicing and post-translational modification of proteins and their oligomerization increase even further this functional repertory.[16] From the mechanistic point of view, the protein-related diseases represent an immense challenge to modern medicine, as they emerge when some of the functional proteins mutate changing their folding, aggregation and then tune their function. In addition, the up or down regulation of protein expression also represents a major challenge. However, these abnormalities exemplify a great opportunity to mitigate, up today, intractable diseases through exploiting new therapeutic approaches.

The use of protein therapeutics can be classified in two main groups depending on their mechanism of action: enzymes and regulatory proteins, and special targeted proteins.

1.1.2.1 Enzymes and regulatory therapeutic proteins

In this group, protein therapeutics operates by a classic paradigm in which a specific endogenous protein is down expressed, and therefore its deficit can be treated by the administration of exogenous proteins. These proteins can be classified in three categories depending on their function:

a) Replacing a protein that is deficient or abnormally produced: these proteins are used in a range of conditions, for example the use of insulin for the treatment of diabetes. Insulin acts through insulin-receptor binding, outside the cell, and it does not require entering the cytosol. However, the administration of insulin is principally limited to parenteral routes, which could cause problems such as safety issues. The oral bioavailability of these biomolecules is very limited (<1%) due to their characteristic low permeability across the epithelium and the rapid degradation. Thus, the development of protein carriers to achieve both excellent mucus permeation and transepithelial absorption is needed.[28] Moreover, due to their key function and broad distribution, lysosomes are central components of cells and play a relevant role in the maintenance of cellular and body-level homeostasis.[13] Genetic deficiencies of lysosomal components, most commonly enzymes, known as "lysosomal diseases", points to aberrant accumulations of undigested substrates, leading to multi-system pathologies, which are often fatal. Enzyme replacement therapy offers an opportunity to

[28] W. Shan, X. Zhu, M. Liu, L. Li, J. Zhong, W. Sun, Z. Zhang and Y. Huang, *ACS Nano*, **2015**, *9*, 2345–2356.
[16] B. Leader, Q. J. Baca and D. E. Golan, *Nat. Rev. Drug Discov.*, **2008**, *7*, 21–39.
[13] M. Solomon and S. Muro, *Adv. Drug Deliv. Rev.*, **2017**, *118*, 109–134.

improve the quality of life and survival rate of patients. Yet, its therapeutic value is obstructed by side effects, resistance and, mainly the inability of recombinant enzymes to reach the central nervous system.

b) Augmenting an existing pathway: the normal protein activity is enhanced in magnitude or timing. These techniques are successfully displayed in the treatment of haematopoietic defects. One important example of these is the use of recombinant erythropoietin (EPO) and its receptor (EPOR), which are indispensable for the survival, proliferation and differentiation of erythroid progenitor cells.[29] EPO is a protein hormone that functions as the primary mediator of a general protective response to tissue hypoxia. This hormone is secreted by the kidney, which stimulates erythrocyte production in the bone marrow. Recently, it has been recognized that other tissues and organs also express EPO and its receptor, including the brain and the heart, independently of its effects on red blood cell mass.[30] Thus, besides its hematopoietic activity, the protein also displays non-erythropoietic actions. Therefore new drug delivery systems are being thoroughly investigated in order to fulfill the specific release of EPO required for each therapeutic approach.[29,30]

c) Providing a novel function or activity: this group includes the use of exogenous proteins to display novel functions and the use of endogenous proteins to act at a different time or place in the body. One remarkable example is the use of human recombinant deoxyribonuclease I (DNASE1). This protein is normally found inside human cells and can be used to degrade the DNA of dying neutrophils in the respiratory tract of patients with cystic fibrosis, pneumonia or bronchitis.[31]

1.1.2.2 Targeted therapeutic proteins

In the last years, it has been shown that the binding specificity of monoclonal antibodies and immunoadhesins can be exploited in numerous ways by recombinant DNA technology.[16] In this group, most of the protein therapeutics take advantage of the antigen recognition sites of immunoglobulin (Ig) antibodies or the receptor-binding domains of native protein ligands, to induce an immune response and therefore, destroy specifically targeted molecules or cells. Alternatively, the combination of the receptor-binding domains of protein ligands with the fragment crystallizable (Fc) region of an Ig has also been used successfully.[32] This region can specifically recognize and target molecules on the cell surface, enabling the destruction by the immune system by triggering the cell death. Moreover, the Fc region can target, endocytose and break down chemically and enzymatically a soluble molecule for its destruction, because

[29] L. Calvillo, R. Latini, J. Kajstura, A. Leri, P. Anversa, P. Ghezzi, M. Salio, A. Cerami and M. Brines, *Proc. Natl. Acad. Sci.*, **2003**, *100*, 4802–4806.

[30] A. Murua, G. Orive, R. M. Hernández and J. L. Pedraz, *Med. Res. Rev.*, **2011**, *31*, 284–309.

[31] S. Shak, D. J. Capon, R. Hellmiss, S. A. Marsters and C. L. Baker, *Proc. Natl. Acad. Sci.*, **1990**, *87*, 9188–9192.

[32] H. I. Park, H. W. Yoon and S. T. Jung, *Trends Biotechnol.*, **2016**, *34*, 895–908.

[16] B. Leader, Q. J. Baca and D. E. Golan, *Nat. Rev. Drug Discov.*, **2008**, *7*, 21–39.

of the recognition of the Fc region by the immune system. Protein therapeutics of this group have been used for the treatment of inflammatory diseases, such as arthritis and psoriasis.[33]

1.1.3 Examples of therapeutic proteins

Protein therapeutics has appeared as a key strategy for the treatment of cancer, immunological diseases and metabolic disorders. Protein drugs have starting to be used in cancer therapy because of their high pharmacological potency, molecular specificity and low toxicity.[10] Therapeutic proteins exploited for cancer treatment mainly includes cytokines, antibodies, enzymes, tumor antigens and pro-apoptotic proteins among others.

Cytokines are a class of secreted or membrane-bound proteins that play an important role in the growth, differentiation and activation of immune cells *via* cell signaling.[34] A variety of cytokines, such as tumor necrosis factors (TNFs), interleukins (ILs) and interferons (IFNs), have been intensively applied in clinical cancer treatment and Hepatitis C virus infections in the case of IFNs,[35] due to their direct apoptosis-inducing action in tumor cells, and their indirect antitumor effects in the regulation of immune responses. Furthermore, the synergistic effects obtained from the combination of cytokines with chemotherapeutic drugs enhance the anti-cancer efficiency and the reduction of the side effects.[36]

Monoclonal antibodies (mAbs) are being developed as therapeutics to complement or to fill the gap that drugs or vaccines cannot reach.[37] Therapeutic antibodies (ThAb) have emerged as the fastest-growing biopharmaceutical approach for the treatment of human diseases. They are especially important for emerging pathogens or for infectious diseases where antibiotic resistance or toxin-mediated pathogenesis are present. Antibodies have a unique structure, which determines the specific nature of the effector function, which need to be considered and integrated into the design when developing ThAb, to ensure the maximum efficacy and safety. The development of mAb technology[38] provided the capability to generate antibodies with required specificities, which had a deep impact on medicine.

A wide variety of monoclonal ThAb are currently licensed, with hundreds more in pre-clinical and clinical development. These therapeutic antibodies are administered for a varied range of conditions. However, the majority are used for cancer treatment, autoimmune disorders and transplantation.[37] Interestingly, almost all these ThAb used an immunoglobulin G (IgG) backbone, which consists of two different fragments: the fragment antigen binding (Fab) region and the fragment crystallizable (Fc) region. The Fab region contains the

[33] X. Cui, L. Chang, Y. Li, Q. Lv, F. Wang, Y. Lin, W. Li, J. D. Meade, J. C. Walden and P. Liang, *Sci. Rep.*, **2018**, *8*, 7327.

[34] G. Dranoff, *Nat. Rev. Cancer*, **2004**, *4*, 11–22.

[35] P. S. Sung, E. C. Shin and S. K. Yoon, *Int. J. Mol. Sci.*, **2015**, *16*, 23683–23694.

[36] C. He, Z. Tang, H. Tian and X. Chen, *Adv. Drug Deliv. Rev.*, **2016**, *98*, 64–76.

[37] V. Irani, A. J. Guy, D. Andrew, J. G. Beeson, P. A. Ramsland and J. S. Richards, *Mol. Immunol.*, **2015**, *67*, 171–182.

[38] G. Köhler and C. Milstein, *Nature*, **1975**, *256*, 495–497.

[10] X. Liu, F. Wu, Y. Ji and L. Yin, *Bioconjug. Chem.*, **2019**, *30*, 305–324.

paratope, also called the antigen-binding site, and exerts direct effects through binding interactions with antigens. The Fc region interacts with a variety of auxiliary molecules to mediate indirect effector functions, which are significant against infectious diseases where cell responses are important for the efficient elimination of pathogens. An essential area of the antibody research is the expansion of the application of this therapeutic tool for the specific detection of molecules for diagnostics, visualization, and activity blocking.[39] Despite the ability to generate antibodies against different proteins, numerous applications in basic research fields, clinical practice, and biotechnology are limited to cell membrane receptors or to extracellular antigens, such as membrane or secreted proteins. Except for small groups of autoantibodies, natural antibodies against intracellular targets cannot be used within living cells. This excludes the extent of a major class of intracellular targets, including cancer-associated molecules.[39] Some of these targets are "undruggable" by small molecules because of their large flat contact areas and the absence of deep hydrophobic pockets, where small molecules can insert and disrupt their activity. Thus, the development of technologies, from direct physical methods to sophisticated delivery vehicles, for the targeted intracellular delivery of antibodies, their fragments, or antibody-like molecules is extremely important.

Protein toxins constitute a defense mechanism against predation or a superior pathogenic competence upon the producing organism.[40] They have been improved through evolution in poisonous animals/plants and pathogenic bacteria. During decades, a big effort has been invested in studying their mechanism of action, the way they contribute to pathogenicity and in the development of antidotes to neutralize their action. Furthermore, many research groups started to explore the pharmaceutical potential of these toxins when they are used to efficiently destroy essential cellular routes and/or damage the integrity of their target cells.[40]

Saporin-S6, also known as saporin, is a positively-charged plant toxin that belongs to the ribosome-inactivating protein (RIP) family, a class of enzymatic toxins that inactivates the ribosomes.[41] Consequently, protein synthesis is shut down, which results in cell death.[42] RIPs can be classified into type 1, consisting of a single-chain protein with enzymatic activity, and type 2, which can enter the cell through the interaction between their lectin moiety and the saccharide receptors present on the cell membrane.[43] While type 2 RIPs can be extremely toxic due to the presence of lectin moiety and their ability to promote entry into target cells, the type 1 RIPs are much less harmful. Saporin belongs to type 1 RIPs, a family of proteins that enter the cell less efficiently than type 2,[42] which in combination with a method to cross the cell membrane, becomes a very potent toxin, as its enzymatic activity is one of the highest of all RIPs. Their highly specific RNA *N*-glycosidase activity cleaves the glycosidic bond of a single adenine from the ribosomal RNA. This is the "Achilles' heel" of the ribosome and the

[39] T. A. Slastnikova, A. V. Ulasov, A. A. Rosenkranz and A. S. Sobolev, *Front. Pharmacol.*, **2018**, *9*, 1–21.

[40] A. Shapira and I. Benhar, *Toxins*, **2010**, *2*, 2519–2583.

[41] G. Bergamaschi, V. Perfetti, L. Tonon, A. Novella, C. Lucotti, M. Danova, M. J. Glennie, G. Merlini and M. Cazzola, *Br. J. Haematol.*, **1996**, *93*, 789–794.

[42] Polito, L.; Bortolotti, M.; Mercatelli, D.; Battelli, M. G.; Bolognesi, A. *Toxins,* **2013**, *5*, 1698–1722.

[43] Stirpe, F.; Barbieri, L.; Battelli, M. G.; Soria, M.; Lappi, D. A. *Bio/Technology*, **1992**, *10*, 405–412.

complete removal of this base inhibits the ability of ribosomes to participate in the synthesis of proteins. Saporin-S6 is extremely resistant to high temperature, denaturation by urea or guanidine and to attack by proteolytic enzymes. It is also stable in response to the chemical modifications needed for its modification and conjugation procedures. All together these characteristics make saporin a very interesting candidate for therapeutic applications.

RIPs have been extensively studied as the toxic moiety of a conjugate, due to their therapeutic potential, in a variety of human diseases. The conjugation of a cytotoxic RIP to a target-specific vehicle, as for example monoclonal antibodies (mAb), allows the selective killing of target cells.[42] Several mAbs are now approved for clinical use and showed excellent cytotoxic activities against certain types of cancer resistant cells, in comparison to the treatment with mAbs alone.[44] However, if the antibody is attached to a cytotoxic agent, its activity is increased and cancer cells can be killed.[45] This kind of antibody/toxin conjugates is known as **immunotoxin**, which combine the potency from the toxin to kill cells and the specificity from the antibody.[46] First-generation immunotoxins, prepared by the chemical conjugation of toxins to antibodies, showed a low efficiency in animal models because of lack of selectivity.[40] The removal of the cell-binding domain from the toxin and its subsequent attachment to antibodies produced immunotoxins with a higher animal tolerance.[44] Several of these second-generation immunotoxins have been evaluated in phase I trials in cancer patients, observing some anti-tumor activity. However, they were expensive to produce, chemically heterogeneous and their ability to penetrate into bulky tumors was reduced due to their large size. For all these reasons, a third-generation of immunotoxins was synthesized by using recombinant DNA techniques.[44] The resulting immunotoxins were designed to contain only those portions of the antibody and toxin needed to recognize and kill a target cell. In **Figure 2**, three different generations of immunotoxins are represented.

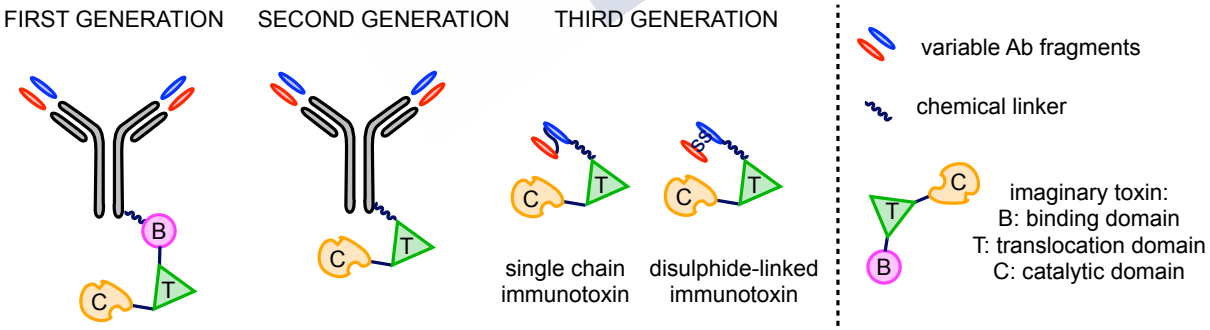


Figure 2. Schematic representation of the three different generations of immunotoxins.

[44] Pastan, I.; Hassan, R.; FitzGerald, D. J.; Kreitman, R. J. *Nat. Rev. Cancer,* **2006**, *6*, 559–565.

[45] Ayyar, B. V.; Arora, S.; O'Kennedy, R. *Trends Pharmacol. Sci.,* **2016**, *37*, 1009–1028.

[46] Alewine, C.; Hassan, R.; Pastan, I. *Oncologist,* **2015**, *20*, 176–185.

[40] A. Shapira and I. Benhar, *Toxins*, **2010**, *2*, 2519–2583.

[42] Polito, L.; Bortolotti, M.; Mercatelli, D.; Battelli, M. G.; Bolognesi, A. *Toxins,* **2013**, *5*, 1698–1722.

1.1.4 Methods for therapeutic protein delivery

The direct intracellular delivery of proteins is not an easy task because each protein is challenging due to the unique properties of each protein such as size, surface charge, function and fragility. Additionally, production of proteins in sufficient quantity and quality is not always viable. Moreover, the cell membrane constitutes an effective barrier to hydrophilic macromolecules that could be damaged initiating irreversible cell death mechanisms.[47] Most of the proteins are delivered inside cells following endocytic pathways.[25] These methods generally trigger their entrapment in endosomes reducing the possibility of reaching the cytosolic targets and therefore, perform their biological activity. During the last years, a variety of different methods have been developed for delivering proteins inside cells. They can be classified in three broad categories: mechanical methods, covalently-modified proteins and supramolecular delivery approaches (Figure 3).

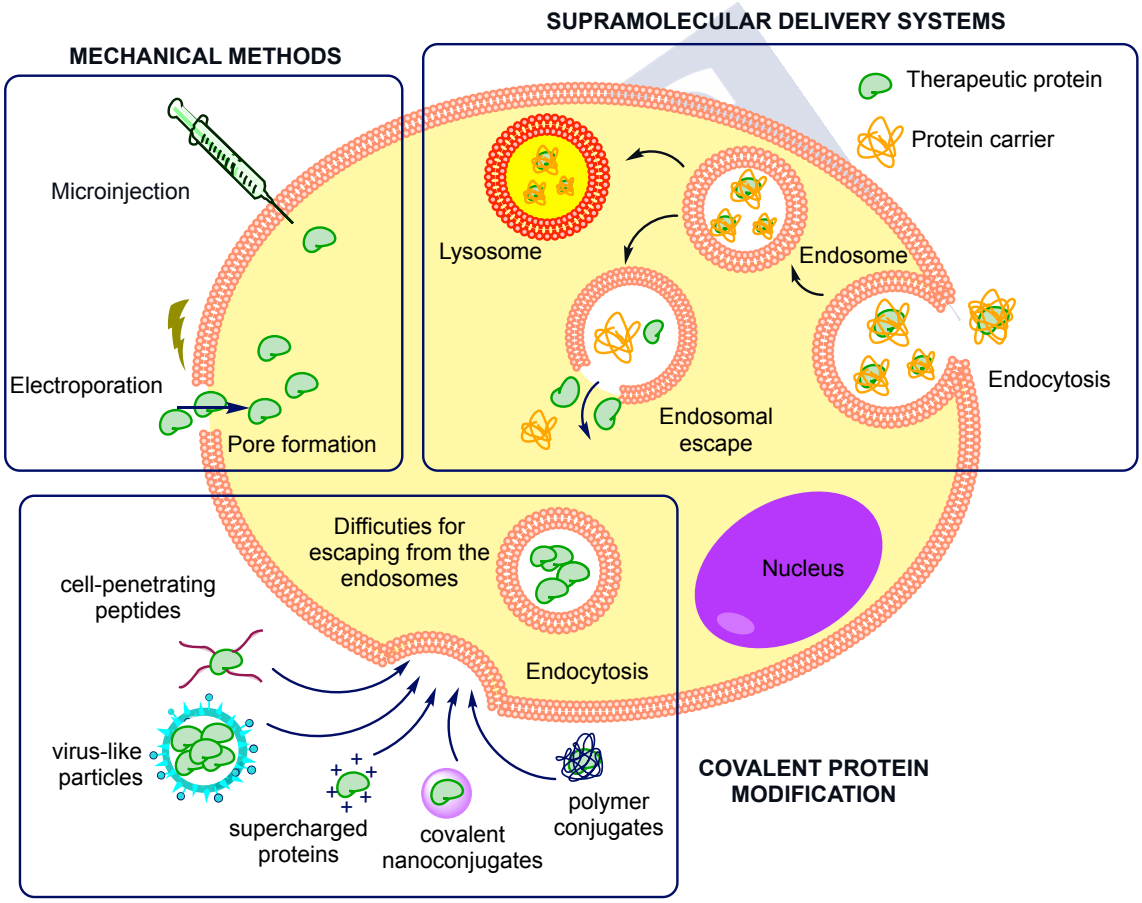


Figure 3. Schematic representation of different classical protein delivery methods.

[47] Chiper, M.; Niederreither, K.; Zuber, G. *Adv. Healthc. Mater.*, **2018**, *7*, 1–21.

[25] A. Fu, R. Tang, J. Hardie, M. E. Farkas and V. M. Rotello, *Bioconjug. Chem.*, **2014**, *25*, 1602–1608.

1.1.4.1 Mechanical methods

Mechanical methods are considered the most traditional procedures to achieve direct cytosolic delivery of proteins.[25] This method is especially useful for *in vitro* investigations. Several techniques have been developed such as microinjection, electroporation, sonoporation, and more recently mechanical deformation and microfluidics electroporation. All these methods have in common their dependence on specific instrumentation and that can not be implemented *in vivo* very easily.[47,15]

Microinjection of protein solutions across the cell membrane with a syringe is a typical powerful method.[39] Solutions can be injected either into the cytosol or within the nucleus. This method is preferable when temporal and quantitative precisions are required for success, but it cannot be used if the number of cells required for protein delivery is large.

Electroporation is another method for transduction of macromolecules.[47] This method consists on placing host cells between two electrodes in a solution containing the protein to be delivered. A brief, few microseconds to a millisecond, high voltage electric pulse is applied to the cell suspension creating temporary pores in the plasma membrane. Soluble proteins and other solutes can then diffuse in both directions through the cell membrane holes according to the Fick's law of diffusion.[48] Electroporation effectiveness depends on the robustness of the cell, protein parameters, transduction medium, and electrical pulse voltage. Fluorescence resonance energy transfer (FRET), based on fluorescent protein biosensors have been widely used for molecular activity visualization in living cells in real time with high spatiotemporal resolution. The delivery into cells of this FRET biosensor, in its protein form and using electroporation, has been reported.[49] The principal advantage of electroporation is its applicability for transitory and stable transfection of all cell types. Furthermore, it is able to transfect a large number of cells in a short time, once optimal conditions are determined. However, the major drawback of conventional electroporation procedures is the substantial cell death caused by high voltage pulses, and the only partially successful membrane repair, which require the use of higher amounts of cells as compared to other transfection methods.

Additionally, a novel vector-free method that uses **reversible membrane permeabilization** to achieve rapid intracellular delivery of cargos with varied composition, properties and size, has been reported.[50] A delivery solution was developed containing low levels of ethanol as the permeabilizing agent. After a short incubation time, permeabilization is stopped by the incubation of the cells in a phosphate buffer saline solution that dilutes the ethanol and that is non-toxic to cells. With this strategy, proteins, messenger RNA, plasmid

[48] Fick, A. *J. Memb. Sci.,* **1995**, *100*, 33–38.

[49] Sun, C.; Ouyang, M.; Cao, Z.; Ma, S.; Alqublan, H.; Sriranganathan, N.; Wang, Y.; Lu, C. *Chem. Commun.,* **2014**, *50*, 11536–11539.

[50] O'Dea, S.; Annibaldi, V.; Gallagher, L.; Mulholland, J.; Molloy, E. L.; Breen, C. J.; Gilbert, J. L.; Martin, D. S.; Maguire, M.; Curry, F.-R. *PLoS One,* **2017**, *12*, e0174779.

[47] Chiper, M.; Niederreither, K.; Zuber, G. *Adv. Healthc. Mater.,* **2018**, *7*, 1–21.

[15] S. Du, S. S. Liew, L. Li and S. Q. Yao, *J. Am. Chem. Soc.*, **2018**, *140*, 15986–15996.

[25] A. Fu, R. Tang, J. Hardie, M. E. Farkas and V. M. Rotello, *Bioconjug. Chem.*, **2014**, *25*, 1602–1608.

DNA and other molecules have been delivered to a variety of cell types, including primary cells, with low toxicity. Remarkably, delivery happens by the direct diffusion of the cargo into the cytoplasm in an endocytic-independent manner. This delivery method has the potential to enable a broad range of research, drug discovery and clinical applications.

Recently, a new method, known as **iTOP** (induced transduction by osmocytosis and propanebetaine), has been described for the efficient intracellular delivery of bioactive molecules.[51] iTOP consists in the combination of small molecule compounds that forces the uptake of therapeutic proteins into the target cell and perform their function, independently of any transduction carrier. The method is an active uptake mechanism where NaCl-mediated hyperosmolality, combined with a transduction compound (propanebetaine), triggers macropinocytotic uptake and intracellular release of therapeutic proteins, respectively (Figure 4). Important advantages of the iTOP delivery method include very good efficiency (>95% of the target cells are transduced) and the ability to target cell lines difficult to manipulate, such as primary cells. Variation in NaCl hypertonicity, type and concentration of transduction compound and transduction time, can be adjusted to the user's needs, the specific target cell and the biochemical characteristic of the therapeutic protein. Moreover, there is direct relationship between the amounts of intracellular and extracellular protein, which allows narrow dosage of the effective intracellular protein levels. Thus, the iTOP technology efficiently transduces native proteins, peptides and small oligonucleotides into practically any cell type, so it can be fully optimized for specific biologicals or target cell types.

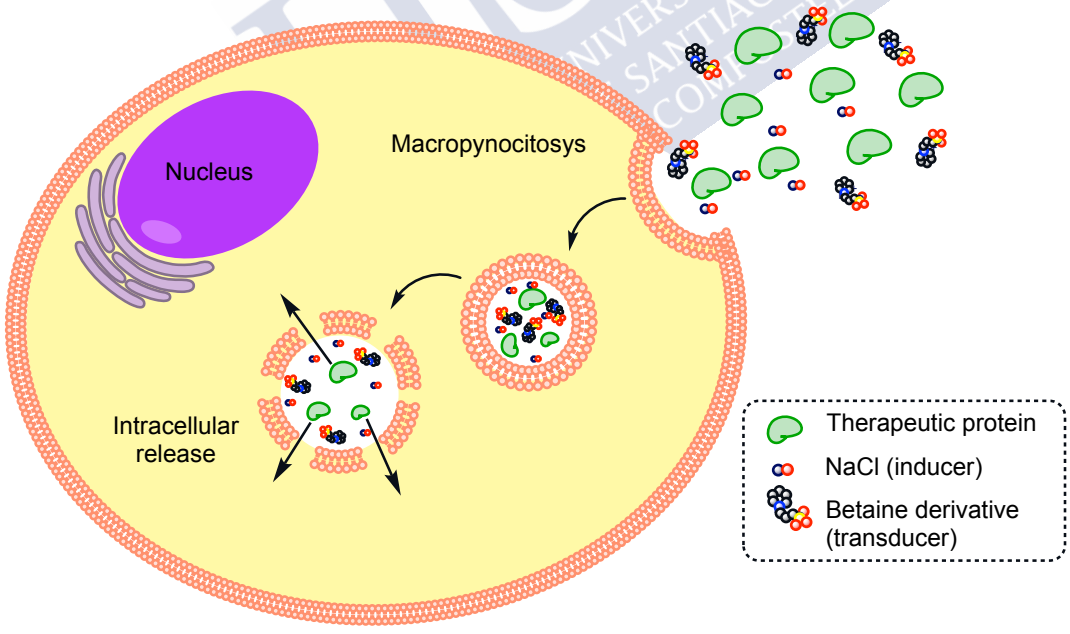


Figure 4. Schematic representation of the iTOP system for the efficient transduction of proteins into primary cells.

[51] D'Astolfo, D. S.; Pagliero, R. J.; Pras, A.; Karthaus, W. R.; Clevers, H.; Prasad, V.; Lebbink, R. J.; Rehmann, H.; Geijsen, N. *Cell,* **2015**, *161*, 674–690.

1.1.4.2 Covalent modification methods

Mechanical methods are effective for protein delivery but they have disadvantages as they cause damage to the cell membrane (electroporation) or takes a long time when the number of cells required is large (microinjection). However, delivering therapeutic proteins to specific tissues or cells encounters various challenges, such as the instability during blood circulation, degradation by enzymes, short half-life, immunogenicity and inability to cross cell membranes.[10] Therefore, the development of delivery systems for protein encapsulation, protection and cell internalization has been explored using chemical modification of proteins *via* direct conjugation or covalent encapsulation.

It is possible to divide this category of delivering therapeutic proteins into several groups differentiated by the type of molecule attached to the protein, including cell-penetrating peptides (CPPs), virus-like nanoparticles (VLPs), supercharged proteins, covalent nanoconjugates and polymer conjugates.[25]

Cell-penetrating peptides have been shown potential for the delivery of a wide variety of molecules, including large and hydrophilic active proteins.[52] Commonly, proteins are covalently modified by short cationic peptides at one of their terminus during expression as CPP fusion proteins. When the cargo is a protein, both CPP and cargo are most often synthesized or expressed in tandem as fusion protein.[53] However, it is also possible to apply the subsequent modification to the purified protein providing a complementary method of accessing proteins, which are functionalized at other positions with synthetic peptides.[54] Protein modifications can be achieved post-translationally by trans-peptidation, the process of transferring an amino acid or group of amino acids from one compound to another, or catalyzed by sortases, in which the peptide is cleaved by the enzyme and ligated to a specific sequence in the N-terminus of the protein.[55,56] For modification in other positions of the protein sequence, the covalent binding of a chemically modified CPP to the thiol of a cysteine (Figure 5A) or to the side chain of a lysine (Figure 5B and C) of the protein, is commonly used (Figure 5).[57,58,59]

---

[52] Kalafatovic, D.; Giralt, E. *Molecules,* **2017**, *22*, 1–38.

[53] Jo, J.; Hong, S.; Choi, W. Y.; Lee, D. R. *Sci. Rep.,* **2015**, *4*, 4378.

[54] Zorko, M.; Langel, U. *Adv. Drug Deliv. Rev.,* **2005**, *57*, 529–545.

[55] Theile, C. S.; Witte, M. D.; Blom, A. E. M.; Kundrat, L.; Ploegh, H. L.; Guimaraes, C. P. *Nat. Protoc.,* **2013**, *8*, 1800–1807.

[56] Li, M.; Tao, Y.; Shu, Y.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R. *J. Am. Chem. Soc.,* **2015**, *137*, 14084–14093.

[57] Gunnoo, S. B.; Iyer, A.; Vannecke, W.; Decoene, K. W.; Hebbrecht, T.; Gettemans, J.; Laga, M.; Loverix, S.; Lasters, I.; Madder, A. *Chem. Commun.,* **2018**, *54*, 11929–11932.

[58] Nakamura, T.; Kawai, Y.; Kitamoto, N.; Osawa, T.; Kato, Y. *Chem. Res. Toxicol.,* **2009**, *22*, 536–542.

[59] Shannon, D. A.; Weerapana, E. *Curr. Opin. Chem. Biol.,* **2015**, *24*, 18–26.

[10] X. Liu, F. Wu, Y. Ji and L. Yin, *Bioconjug. Chem.*, **2019**, *30*, 305–324.

[25] A. Fu, R. Tang, J. Hardie, M. E. Farkas and V. M. Rotello, *Bioconjug. Chem.*, **2014**, *25*, 1602–1608.

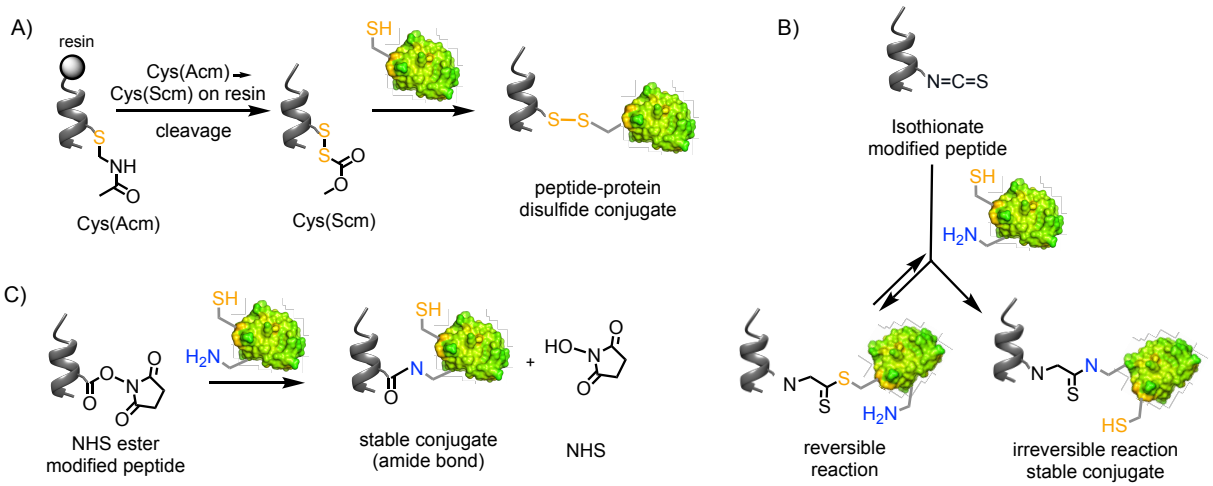


Figure 5. Different strategies for the covalent protein modification using CPPs. A) Conversion of Cys acetamidomethyl-containing peptides (Cys-Acm) to Cys S-carbomethoxysulfenyl group (Scm) for the subsequent reaction with free Cys of proteins. B) Reaction of isothiocyanate modified peptide with Cys or Lys protein residues in physiological conditions. C) Amide bond formation using *N*-hydroxysuccinimide (NHS) ester reaction.

There have been numerous *in vitro* examples of CPP-mediated delivery of proteins,[60] such as TAT peptide, Penetratin, Transportan, etc., CPPs derived from naturally occurring proteins with the ability of crossing the cell membrane in a natural way.[61] Since the discovery of the ability of TAT peptide to transport proteins across the cell membrane, CPPs have been shown to be effective for the delivery of proteins ranging in size from 30 kDa (GFP) to around 500 kDa (β-galactosidase). The administration of *i.e.* β-galactosidase protein fused to TAT in mice, resulted in the efficient internalization in all tissues tested, including the blood-brain barrier, preserving the biological activity of the protein.[62] However, the efficient release of these protein conjugates to the cytosol is questionable. Endocytosis is the most common pathway for most of CPPs to entering the cells, therefore endosomal escape is required for proteins linked to CPPs to reach the cytosol and their intracellular targets.[63] Some efficient endosomolytic peptides for protein delivery have been described.[56] These covalent complexes have been successfully used for the delivery of a wide range of cargos but, from the chemical point of view, this technology has limitations such as, impact of the peptide pendant in the biological activity of the delivered proteins. In this case, reversible bonds can be employed, such as the case of CPPs attached to the protein through a cysteine disulfide bond

[60] Patel, S. G.; Sayers, E. J.; He, L.; Narayan, R.; Williams, T. L.; Mills, E. M.; Allemann, R. K.; Luk, L. Y. P.; Jones, A. T.; Tsai, Y.-H. *Sci. Rep.,* **2019**, *9*, 6298.

[61] Mäe, M.; Langel, Ü. *Curr. Opin. Pharmacol.,* **2006**, *6*, 509–514.

[62] Schwarze, S. R. *Science,* **1999**, *285*, 1569–1572.

[63] Lönn, P.; Kacsinta, A. D.; Cui, X.-S.; Hamil, A. S.; Kaulich, M.; Gogoi, K.; Dowdy, S. F. *Sci. Rep.,* **2016**, *6*, 32301.

[56] Li, M.; Tao, Y.; Shu, Y.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R. *J. Am. Chem. Soc.,* **2015**, *137*, 14084–14093.

(Figure 5A).[64] The resultant conjugate enters cells and, once in the cytosol, is then cleaved due to the reductive environment inside cells, allowing the bioactive cargo to perform its function.

Virus-like particles (VLPs) are another emerging category of carriers for protein delivery. VLPs are constituted by mimicking the self-assembling process of capsid proteins of the native virion, but without the viral genetic material.[65] Recent reports indicate that VLPs can be used to deliver biologically active proteins into cells as heterologous protein fused to the VLP anchoring protein.[66] However, the intracellular delivery of proteins with VLPs still requires investigation for the efficient cytosolic access.

Supercharged proteins are a family of engineered or naturally occurring proteins with unusually high positive or negative net charge.[67] Engineered supercharged proteins are the result of wide mutagenesis in which residues exposed to solvents in protein's surface were substituted with acidic or basic amino acids. It has been described "supercharged" GFP variants that have been widely mutated at their surface-exposed residues, resulting in extremely high theoretical net charge magnitudes ranging from -30 to +48.[68] The superpositively charged GFP variants can enter into a variety of mammalian cells by binding to anionic cell-surface proteoglycans and undergoing endocytosis in an energy-dependent and clathrin-independent manner.[69] Additionally, these superpositive GFPs have been engineered for drug delivery with specific targets, as for example drugs directed to cartilage,[70] having a wide variety of net positive charge and surface charge distributions, to then characterize the effects of carrier charge on transport into cartilage, in isolation of other factors such as carrier's size and shape.

Covalent nanoconjugates offer an alternative strategy to direct protein delivery, providing increased options for size control and surface properties.[71] Covalent attachment provides a stable bond between carrier and protein, an important factor for *in vivo* applications. However, these covalent conjugates could interfere in protein folding and function. As an example of nanocarrier, mesoporous silica nanoparticles (MSNs) were promising for the intracellular delivery of membrane-impermeable proteins.[72] These particles protect the protein molecules from proteases and denaturing chemicals by encapsulating

[64] Jones, L. R.; Goun, E. A.; Shinde, R.; Rothbard, J. B.; Contag, C. H.; Wender, P. A. *J. Am. Chem. Soc.,* **2006**, *128*, 6526–6527.

[65] Rohovie, M. J.; Nagasawa, M.; Swartz, J. R. *Bioeng. Transl. Med.,* **2016**, *2*, 43–57.

[66] Kaczmarczyk, S. J.; Sitaraman, K.; Young, H. A.; Hughes, S. H.; Chatterjee, D. K. *Proc. Natl. Acad. Sci.,* **2011**, *108*, 16998–17003.

[67] Thompson, D. B.; Cronican, J. J.; Liu, D. R. *Methods Enzym.,* **2012**, *5*, 293–319.

[68] Lawrence, M. S.; Phillips, K. J.; Liu, D. R. *J. Am. Chem. Soc.,* **2007**, *129*, 10110–10112.

[69] Cronican, J. J.; Thompson, D. B.; Beier, K. T.; McNaughton, B. R.; Cepko, C. L.; Liu, D. R. *ACS Chem. Biol.,* **2010**, *5*, 747–752.

[70] Krishnan, Y.; Rees, H. A.; Rossitto, C. P.; Kim, S. E.; Hung, H. H. K.; Frank, E. H.; Olsen, B. D.; Liu, D. R.; Hammond, P. T.; Grodzinsky, A. J. *Biomaterials,* **2018**, *183*, 218–233.

[71] Doane, T.; Burda, C. *Adv. Drug Deliv. Rev.,* **2013**, *65*, 607–621.

[72] Ray, M.; Lee, Y. W.; Scaletti, F.; Yu, R.; Rotello, V. M. *Nanomedicine,* **2017**, *12*, 941–952.

proteins in their large protective shell. The intracellular delivery of cytochrome c was demonstrated using these MSN estrategy.[73]

Polymer-drug conjugate therapeutics are pharmacologically-active macromolecular constructs involving one or more therapeutic agents, including small molecules, peptides, proteins and aptamers, covalently bound to a polymeric carrier.[74] The use of polyethylene glycol (PEG) for protein conjugation is remarkable, which improves protein solubility and extends the plasma half-life. PEG is a highly water-soluble, flexible, uncharged and biocompatible polymer that is widely used as an excipient in the pharmaceutical industry.[75] The process of PEGylation involves formation of stable covalent bonds between activated PEG polymers and the bioactive protein of interest. Firstly, PEG is activated with functional groups in one or both end sites. Cyanuric chloride, PEG-succinimidyl succinate, and imidazolyl formate are some of the more common approaches. The most common sites for protein PEGylation are the lateral chains of lysine amino acids and the N-terminal amino group of proteins.[76] Recently, the covalent attachment of amphiphilic polymers to the N-terminal positions of proteins for their intracellular delivery using 2-pyridinecarboxaldehyde groups has been published,[77] providing site-selective functionalization for many different biomolecules, in a location that minimally affects protein structure and function.

1.1.4.3 Supramolecular methods

In living systems, a variety of supramolecular structures play an important role in maintaining life, such as double-helix DNA, functional protein domains, microtubules and microfilaments. At the end of the 70s, J. M. Lehn firstly put forward the concept of supramolecular chemistry referring to the chemistry beyond the molecule, which aims to build highly complex functional systems from building blocks joined by non-covalent intermolecular interactions.[78] Non-covalent conjugation methods for protein delivery can employ a variety of loading strategies including electrostatic interaction, host-guest encapsulation, and hydrophobic effect.[10] From all these methods, electrostatic interaction is the most commonly employed, as there are a multitude of ways to adjust the surface charge of both carriers and proteins to create a compatible system.

Cell-penetrating peptides (CPPs) are oligopeptides with potent membrane activities, which can deliver a wide range of membrane-impermeable bioactive agents across cell membranes.[79] Moreover, most of the CPPs contain cationic amino acids such as lysine and/or

[73] Méndez, J.; Morales Cruz, M.; Delgado, Y.; Figueroa, C. M.; Orellano, E. A.; Morales, M.; Monteagudo, A.; Griebenow, K. *Mol. Pharm.,* **2014**, *11*, 102–111.

[74] Ekladious, I.; Colson, Y. L.; Grinstaff, M. W. *Nat. Rev. Drug Discov.,* **2019**, *18*, 273–294.

[75] Milton Harris, J.; Chess, R. B. *Nat. Rev. Drug Discov.,* **2003**, *2*, 214–221.

[76] Nischan, N.; Hackenberger, C. P. R. *J. Org. Chem.,* **2014**, *79*, 10727–10733.

[77] Sangsuwan, R.; Tachachartvanich, P.; Francis, M. B. *J. Am. Chem. Soc.,* **2019**, *141*, 2376–2383.

[78] Lehn, J.-M. *Angew. Chemie Int. Ed. English,* **1988**, *27*, 89–112.

[79] Munyendo, W. L. L.; Lv, H.; Benza-Ingoula, H.; Baraza, L. D.; Zhou, J. *Biomolecules,* **2012**, *2*, 187–202.

[10] X. Liu, F. Wu, Y. Ji and L. Yin, *Bioconjug. Chem.*, **2019**, *30*, 305–324.

arginine residues, which can interact with the negative charges of proteins forming nanocomplexes that enhance the cellular uptake as well as the endosomal escape efficiency.[27]

CPPs can be classified into primary, secondary and non-amphipathic CPPs (Figure 6).[27] Primary amphipathic CPPs (e.g., Pep-1) are characterized by a primary peptide sequence where hydrophobic and cationic residues are separated at the primary structure level. Secondary amphipathic peptides (e.g., R6/W3) are made of hydrophobic and cationic residues that have an amphipathic character only when the CPP is folded into an α-helical conformation, exposing the hydrophobic and the cationic residues in two different domains, while this partition of the residues is not achieved at the primary sequence level. Finally, non-amphipathic CPPs, without a predetermined structure, includes polycationic CPPs such as R9.

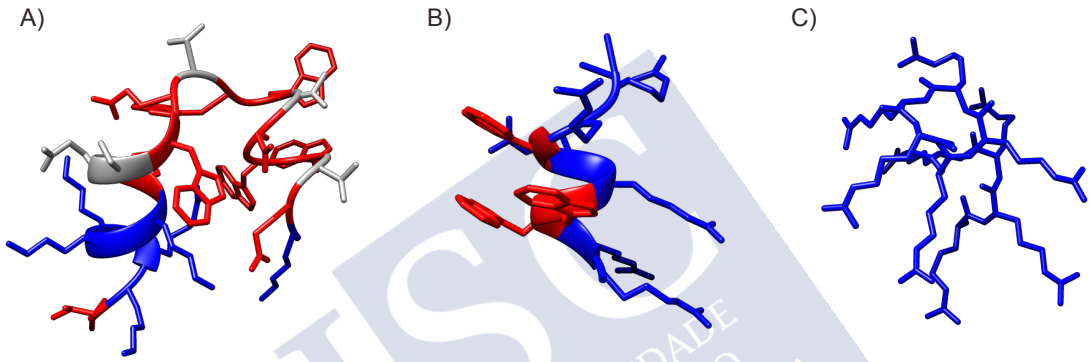


Figure 6. CPPs classification. A) Pep-1 primary structure with hydrophilic and hydrophobic domains separated randomly along the peptide sequence, B) R6/W3 secondary structure with hydrophilic and hydrophobic domains pointing to different faces, and C) R9 non-amphipathic CPP with unstructured conformation. Polar charged amino acids are represented in blue, hydrophobic amino acids in red and polar uncharged amino acids are in gray. Structures made using PEP-FOLD prediction.[80]

These CPPs form non-covalent complexes with biomolecules improving their intracellular delivery.[79] They arise mostly as short amphipathic peptide carriers consisting of a hydrophilic (cationic) domain and a hydrophobic domain (Figure 6). The amphipathic character arises from either the primary structure or the secondary structure described above, which helps the interaction of CPPs with the cell membrane and subsequent internalization.[79] As an example of these amphipathic peptides, Pep-1, which belongs to the primary amphipathic structure, forms stable complexes with proteins through non-covalent electrostatic and Van der Waals interactions, being able to efficiently deliver a variety of peptides and proteins into numerous

[80] Lamiable, A.; Thévenet, P.; Rey, J.; Vavrusa, M.; Derreumaux, P.; Tufféry, P. *Nucleic Acids Res.,* **2016**, *44*, W449–W454.

[27] A. Bolhassani, B. S. Jafarzade and G. Mardani, *Peptides*, **2017**, *87*, 50–63.

[79] Munyendo, W. L. L.; Lv, H.; Benza-Ingoula, H.; Baraza, L. D.; Zhou, J. *Biomolecules,* **2012**, *2*, 187–202.

cell lines. In addition, this peptide carrier presents several advantages for protein therapy, including stability in physiological buffer, low toxicity, and reduced sensitivity to serum.[81,82]

Polymers are attractive protein drug delivery carriers, as their physical and chemical properties can be easily modified. Functional supramolecular polymers have the ability to undergo reversible switching of structure, shape and function in response to certain external stimuli, making them exceptional candidates for a wide variety of biomedical applications, such as drug and protein delivery, gene transfection, bioimaging and diagnosis, tissue engineering and biomimetic chemistry.[83]

Self-assembled hydrogels with extremely high water content (up to 50%) and highly tunable mechanical properties, have been used for the sustained release of proteins under *in vitro* conditions.[84] These hydrogels, prepared from renewable cellulose derivatives, are easily processed and their simple preparation, availability from low-cost renewable resources, and the tunability of their mechanical properties are unique for important biomedical applications. Dextran, a complex and branched polysaccharide synthesized from sucrose by enzymes or produced by bacteria and yeast, has many applications for plasma volume expansion, thrombosis prophylaxis, peripheral blood flow enhancement and for the rheological improvement of, for instance, artificial tears.[85] An example of dextran applications is the *in situ* formation of degradable hydrogels between dextran vinyl sulfones and multifunctional mercapto-PEG, which leads to the release of model proteins with different sizes, such as immunoglobulin G (IgG), bovine serum albumin (BSA) and lysozyme and basic fibroblast growth factor.[86]

Additionally, nanocarriers prepared from biocompatible polymers have the potential for effectively delivering proteins to subcellular sites by exploiting the advantages of polymer shielding, which can protect proteins, such as antibodies, from protein interaction, loss of affinity and decomposition.[87] For example, polyion complex (PIC) micelles, which are formed when a block copolymer, with a neutral hydrophilic and anionic blocks, is mixed with counter-charged compounds, have been extensively used for intracellular protein delivery (**Figure 7**).[88] In a recent report, PIC micelles have been prepared by charge-converted IgG antibody derivatives and (PEG)-poly [*N*-{*N*'-(2-aminoethyl)-2-aminoethyl}aspartamide] copolymers, whose contiguous 1,2-diaminoethane units provide selective destabilization of

[81] Morris, M. C.; Depollier, J.; Mery, J.; Heitz, F.; Divita, G. *Nat. Biotechnol.,* **2001**, *19*, 1173–1176.

[82] Deshayes, S.; Morris, M.; Heitz, F.; Divita, G. *Adv. Drug Deliv. Rev.,* **2008**, *60*, 537–547.

[83] Dong, R.; Zhou, Y.; Huang, X.; Zhu, X.; Lu, Y.; Shen, J. *Adv. Mater.,* **2015**, *27*, 498–526.

[84] Appel, E. A.; Loh, X. J.; Jones, S. T.; Dreiss, C. A.; Scherman, O. A. *Biomaterials,* **2012**, *33*, 4646–4652.

[85] Van Tomme, S. R.; Hennink, W. E. *Expert Rev. Med. Devices,* **2007**, *4*, 147–164.

[86] Hiemstra, C.; Zhong, Z.; van Steenbergen, M. J.; Hennink, W. E.; Feijen, J. *J. Control. Release,* **2007**, *122*, 71–78.

[87] Kim, A.; Miura, Y.; Ishii, T.; Mutaf, O. F.; Nishiyama, N.; Cabral, H.; Kataoka, K. *Biomacromolecules,* **2016**, *17*, 446–453.

[88] Lee, Y.; Kataoka, K. *Soft Matter,* **2009**, *5*, 3810–3817.

endosomal membranes, facilitating the endosomal escape and delivery of charge-restored IgG antibodies into the cytosol.[87]

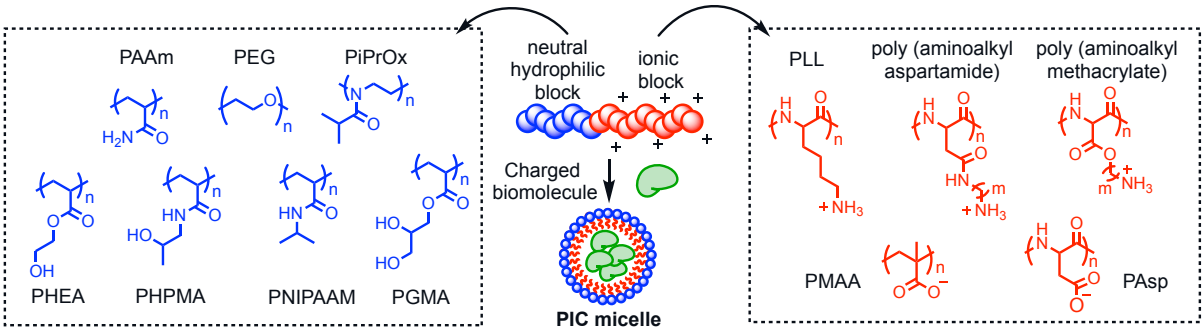


Figure 7. Polymeric structures for PIC micelles formation and biomolecule encapsulation.

Although great progress has been made in the field of functional supramolecular polymers for biomedical applications, this area still faces several significant challenges such as,[74] their physicochemical heterogeneity, which results in polydisperse polymer mixtures with different protein loadings, low functionalization for drug conjugation in polymers like polyesters and challenges in achieving oral bioavailability and crossing mucosal barriers.

Liposomes or lipid vesicles are one of the most traditional nanocarriers showing excellent modularity and easy preparation. Phospholipids form spherical vesicles that consist of an aqueous core surrounded by a lipid bilayer or multilayer. Phospholipids have a polar head group and two hydrophobic hydrocarbon tails, which usually are fatty acids with one tail having an unsaturated bond.[89] Due to the aqueous core, a huge amount of water-soluble proteins can be encapsulated inside liposomes, maintaining its native structure. Lipids with long hydrocarbon chains and a low degree of unsaturation and branching have been used to form strongly packed liposomes. Furthermore, to minimize membrane defects and improve their stability, cholesterol has sometimes been incorporated to the liposomes.[89,90]

Lipid vesicles have been used for medical applications because of their ability to protect and deliver hydrophilic and hydrophobic cargos. Additionally, they have been used as carriers for drug delivery due to their biocompatibility with cell membranes and their capacity to add specific ligands to their surface. Although many protein drugs are becoming available with rapid advances in recombinant DNA technology, their poor stability for physiological conditions and their difficulty to cross the cell membrane have attracted attention for the development of artificial liposomal carriers for effective protein therapeutic delivery.[91]

[89] Lee, K. Y.; Yuk, S. H. *Prog. Polym. Sci.,* **2007**, *32*, 669–697.

[90] Briuglia, M. L.; Rotella, C.; McFarlane, A.; Lamprou, D. A. *Drug Deliv. Transl. Res.,* **2015**, *5*, 231–242.

[91] Tan, M. L.; Choong, P. F. M.; Dass, C. R. *Peptides,* **2010**, *31*, 184–193.

[74] Ekladious, I.; Colson, Y. L.; Grinstaff, M. W. *Nat. Rev. Drug Discov.,* **2019**, *18*, 273–294.

[87] Kim, A.; Miura, Y.; Ishii, T.; Mutaf, O. F.; Nishiyama, N.; Cabral, H.; Kataoka, K. *Biomacromolecules,* **2016**, *17*, 446–453.

Cationic lipid formulations with the ability of delivering two different types of proteins into cells, enzyme and antibody, have been reported.[92] A guanidinium-cholesterol cationic lipid bis (guanidinium)-tren-cholesterol was combined with dioleoyl phosphatidyl-ethanolamine to efficiently transport the β-galactosidase enzyme intracellularly, confirming that protein structure and function were not altered as a result of complexation. Moreover, the aminoglycoside lipid dioleyl succinyl paromomycin (DOSP) associated with an imidazole-based helper lipid was shown to achieve intracellular delivery of biologically active anti-cytokeratin 8 antibody. The traditional unilamellar and multilamellar liposome systems present mechanical instability in physiological conditions, resulting in a fast release of the internally loaded content. As consequence, multivesicular liposomes have been developed to overcome this limitation. These multivesicular liposomes are formed by several internal aqueous chambers linked by a continuous and non-concentric network of lipid membranes, which results in a high ratio between aqueous volumes to lipids (Figure 8). This particular disposition provides a sequential release of the encapsulated proteins due to a unique disruption site in the external liposome membrane.[89] Despite the fact that liposomes are typically made from natural, biodegradable, non-toxic and non-immunogenic lipid molecules, its clinical use, usually administered intravenously, presents some inconveniences as a result of their interaction with lipoproteins presented in blood, which leads to a premature release of the encapsulated drug.[93]

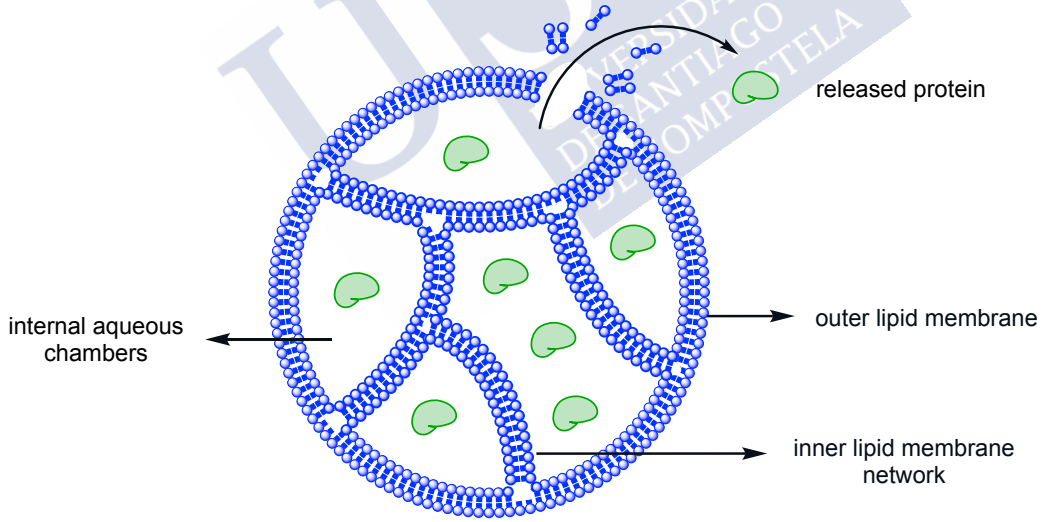


Figure 8. Morphology of multivesicular liposomes for the sequential release of therapeutic proteins.

[92] Chatin, B.; Mével, M.; Devallière, J.; Dallet, L.; Haudebourg, T.; Peuziat, P.; Colombani, T.; Berchel, M.; Lambert, O.; Edelman, A.; Pitard, B. *Mol. Ther. Nucleic Acids,* **2015**, *4*, e244.

[93] Lombardo, D.; Calandra, P.; Barreca, D.; Magazù, S.; Kiselev, M. *Nanomaterials,* **2016**, *6*, 125.

[89] Lee, K. Y.; Yuk, S. H. *Prog. Polym. Sci.,* **2007**, *32*, 669–697.

2 *SUPRAMOLECULAR RECOGNITION AND SELECTIVE PROTEIN UPTAKE BY PEPTIDE HYBRIDS*

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2.1 Precedents and Objectives

The interactions between lectins and carbohydrates play an important role in a large variety of biological processes, such as cell-cell communication, cell adhesion, cell recognition, cell differentiation, host-pathogen interactions, signal transduction, intracellular trafficking of proteins, inflammation, metastasis and development of the neuronal network.[94] Therefore, the delivery of exogenous lectins inside cells has been extensively studied in the fields of cell targeting and cancer therapy.[95] Moreover, covalent binding between CPPs and large cargos has been extensively used for cell internalization of a huge variety of proteins.[61] However, the conjugation of CPPs to proteins might often lead to a loss of biological activity, requiring the development of new non-covalent approaches.

In the first chapter of this thesis, we will study amphiphilic cell-penetrating peptides with controlled secondary structure (α-helices), which will be modified with glycan ligands to recognize lectins and trigger their uptake into cells (Figure 9).

The designed CPP will present two well-defined domains, one of them being hydrophilic and cationic while the other one will be hydrophobic. In the interface of these two domains and in orthogonal disposition, hydroxylamine moieties will be introduced, for the subsequent attachment of the corresponding protein recognition unit. Secondary structure of the peptide scaffold will be determined by circular dichroism.

After the attachment of the corresponding ligand (*i.e.* mannose for Concanavalin A) protein interaction and uptake in two different cell lines will be studied using fluorescence anisotropy, surface plasmon resonance (SPR) and fluorescence microscopy.

Moreover, we will perform studies of the uptake mechanism using different inhibitors. Selective protein transport assays will be performed using the mannosyl-modified CPP as compared with Lipofectamine2000, the non-viral transporter typically used in the literature.

Finally, the glycan unit will be exchanged for a different protein ligand (*i.e.* biotin) to explore the extension of this methodology to different protein targets.

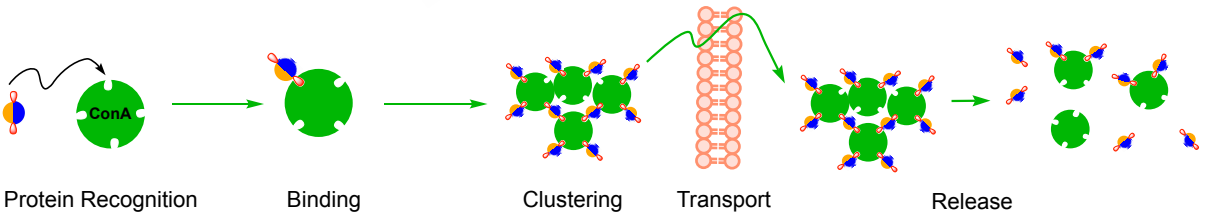


Figure 9. Steps for the selective protein cell internalization using a CPP. Cationic domain (blue), hydrophobic domain (orange), protein recognition unit (manose, red).

[94] Cecioni, S.; Faure, S.; Darbost, U.; Bonnamour, I.; Parrot-Lopez, H.; Roy, O.; Taillefumier, C.; Wimmerová, M.; Praly, J. P.; Imberty, A.; Vidal, S. *Chem. - A Eur. J.,* **2011**, *17*, 2146–2159.

[95] Liu, B.; Bian, H. J.; Bao, J. K. *Cancer Lett.,* **2010**, *287*, 1–12.

[61] Mäe, M.; Langel, Ü. *Curr. Opin. Pharmacol.,* **2006**, *6*, 509–514.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 2: GENE EDITION

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1 INTRODUCTION

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1.1 Therapeutic interest of gene edition by direct protein delivery

1.1.1 Introduction of CRISPR systems

Although DNA repeats with dyad symmetry was first reported in 1987 by Nakata *et al.*,[96] it was not until 2002 when the clustered regularly interspaced short palindromic repeats (CRISPR) started to be more extensively studied by scientists.[97] This family of DNA sequences, reported first in *E. coli* (Gram-negative) and later in other bacteria, including *S. pyogenes* (Gram-positive), and in archaea, was found to be part of the prokaryotic adaptive immune defense against bacteriophages and plasmids, where an endogenous RNA-guided nuclease-based machinery recognizes and destroys invading external nucleic acid material.[98] To clarify, CRISPR loci in the DNA of bacteria and archaea comprise DNA sequence repeats separated by spacers (short variable DNA sequences), which are usually close to genes that encode for the CRISPR-associated (Cas) family of endonucleases. CRISPR-Cas immunity comprises three different mechanistic stages (Figure 10):[99]

- **Adaptation**, which involves the incorporation of foreign DNA from invading viruses and plasmids into the CRISPR array as new spacers that provide the sequence memory for a targeted defense against subsequent invasions by the corresponding virus or plasmid.
- **Expression**, where the CRISPR array is transcribed as a precursor transcript (pre-crRNA), which is processed and matured for producing CRISPR RNAs (crRNAs).
- **Interference**, where crRNAs aided by Cas proteins function as guides to specifically target and cleave the nucleic acids of similar viruses or plasmids.

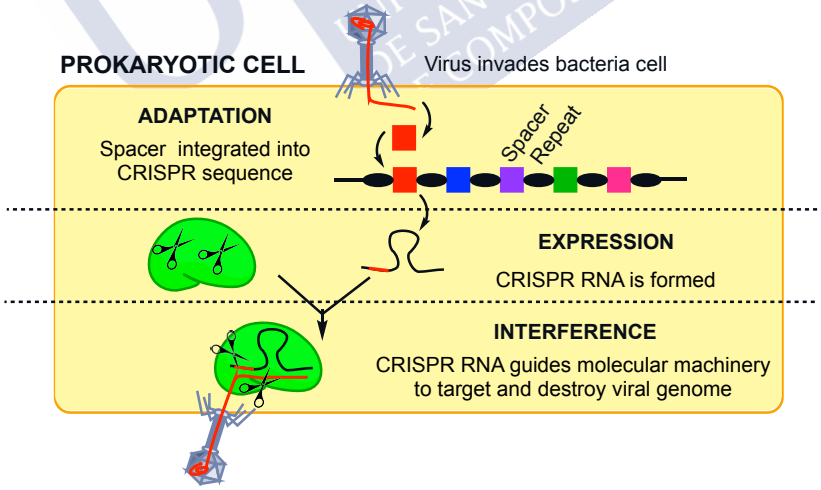


Figure 10. Different mechanistic stages of CRISPR-mediated immunity in prokaryotic cells. CRISPR regions are composed of short DNA repeats (black ellipses) and spacers (colored boxes).[100]

[96] Ishino, Y.; Shinagawa, H.; Makino, K.; Amemura, M.; Nakata, A. *J. Bacteriol.,* **1987**, *169*, 5429–5433.
[97] Jansen, R.; Embden, J. D. A. Van; Gaastra, W.; Schouls, L. M. *Mol. Microbiol.,* **2002**, *43*, 1565–1575.
[98] Mojica, F. J. M.; Montoliu, L. *Trends in Microbiology*, **2016**, *24*, 811–820.
[99] Makarova, K. S. et al. *Nat. Rev. Microbiol.,* **2015**, *13*, 722–736.
[100] Gasiunas, G.; Sinkunas, T.; Siksnys, V. *Cell. Mol. Life Sci.,* **2014**, *71*, 449–465.

The observation of this dynamic interaction between the spacer content of CRISPR arrays and the potential targets in natural microbial communities, further supported a connection between CRISPR genotype and host immunity. These findings verified CRISPR-Cas as a general genetic barrier to horizontally transferred DNA and an efficient adaptive immune system in prokaryotes. The potential of genome editing using nucleases and, in particular by RNA-guided endonucleases of bacterial origin, has emerged as one of the most promising strategies for biomedical research and therapeutic purposes.[101] Recently, an *in vitro* reconstitution of the *Streptococcus pyogenes* type II CRISPR system demonstrated that crRNA attached to a normally trans-encoded crRNA (tracrRNA) is enough to direct Cas9 protein to sequence-specifically cleave target DNA sequences matching the crRNA.[102] Cas9 is an RNA-guided nuclease whose sequence specificity largely arises from Watson–Crick base pairing between its guide RNA (gRNA) and the target DNA site, in addition to a direct interaction between Cas9 and a short protospacer-adjacent motif (PAM) of DNA (Figure 11A). Thus, Cas9 can be programmed to target new sites simply by changing its gRNA sequence, making it an ideal platform for high-throughput sequence-specific gene editing, as well as other applications. The fully defined nature of this system suggested that it might function in eukaryotic cells such as yeast, plants and even mammals.[103, 104] Thus, the potential applications of this technology in mammalian cells has opened an exciting and growing field for highly efficient and specific genome editing and regulation in diverse organisms, creating revolutionary tools for biomedical research and new possibilities for treating genetic disorders.[105]

Previous examples of programmable genome editing nucleases include **zinc-fingers nucleases** (ZFNs)[106] and **transcription activator-like effector nucleases** (TALENs),[107] where the DNA-binding domains of transcription factors have been fused with the nuclease domain of the restriction enzyme FokI (Figure 11B). In **ZFNs**, DNA-binding domains are comprised of at least three consecutive repetitions of around 30 amino acid Zn finger motif, recognizing three nucleotides each repetition.[108,106] Because the nuclease domain must dimerize to cut DNA, two ZFN molecules are required to target a single site. ZFNs have been used to modify endogenous genes in a wide variety of organisms and cell types introducing several types of genomic alterations such as mutations, deletions, insertions, etc. These provide researchers with unprecedented tools for performing genetic manipulations. **TALENs**

[101] Pattanayak, V.; Guilinger, J. P.; Liu, D. R. *Methods Enzymol.,* **2014**, *9*, 47–78.

[102] Martin Jinek; Krzysztof Chylinski; Ines Fonfara; Michael Hauer; Jennifer A. Doudna; Emmanuelle Charpentier. *Science,* **2012**, *337*, 816–821.

[103] Mali, P.; Yang, L.; Esvelt, K. M.; Aach, J.; Guell, M.; DiCarlo, J. E.; Norville, J. E.; Church, G. M. *Science,* **2013**, *339*, 823–826.

[104] Cong, L.; Ran, F. A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P. D.; Wu, X.; Jiang, W.; Marraffini, L. a; Zhang, F. *Science,* **2013**, *339*, 819–823.

[105] Wang, H.; La Russa, M.; Qi, L. S. *Annu. Rev. Biochem.,* **2016**, *85*, 227–264.

[106] Carroll, D. *Genetics,* **2011**, *188*, 773–782

[107] oung, J. K.; Sander, J. D. *Nat. Rev. Mol. Cell Biol.,* **2013**, *14*, 49–55.

[108] Carroll, D. *Annu. Rev. Biochem.,* **2014**, *83*, 409–439.

have rapidly appeared as an alternative to ZFNs as they introduce targeted double-strand breaks comprising a non-specific FokI nucleases domain fused to a customizable DNA-binding domain. TALENs have a DNA-binding domain made up of tandem repeats of around 34 amino acids, and these modules are not related to any other DNA-recognition motif.[108] However, as these tools function through protein–DNA interactions, targeting to a new site requires engineering and cloning a new protein, which excludes ZFNs and TALENs from being used for high-throughput applications.

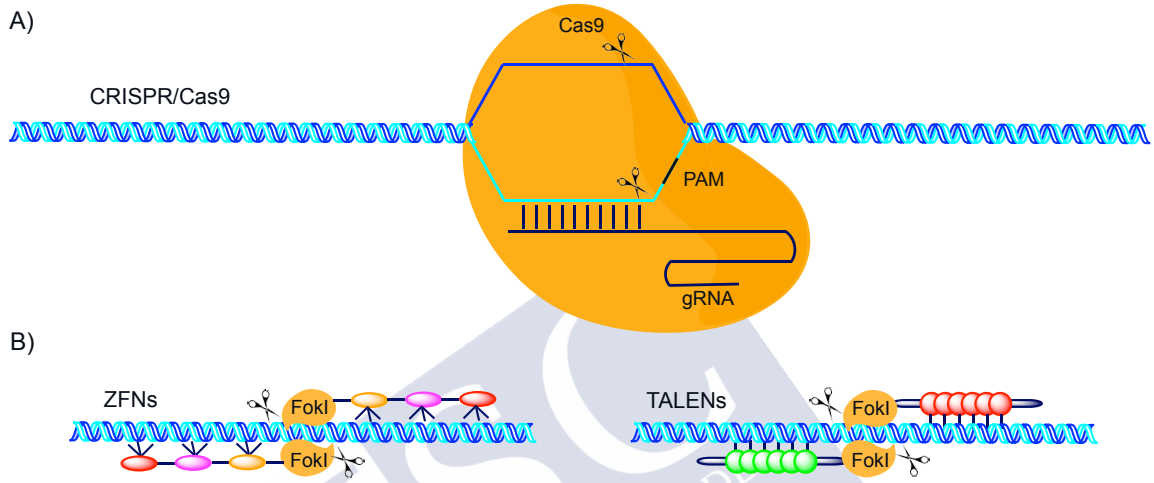


Figure 11. Genome-editing techniques. A) RNA-directed nucleases (CRISPR/Cas9) and B) nucleases base on protein-DNA interactions (ZFNs and TALENs).

1.1.2 CRISPR/Cas9 technology applications

CRISPR/Cas9 technology offers a precise and simple molecular mechanism for editing cells, tissues and whole organisms. Derived from a prokaryotic adaptive immune system, CRISPR/Cas9 has been used to develop potent tools for genome manipulation in animals, plants and microorganisms.[109] The CRISPR/Cas9 system offers simplicity and efficacy in virtually all cell types, as compared with ZFN and TALENs technologies, involving cells and animals of medical interest, plants and livestock species relevant for food and agriculture, and model organisms widely used by the scientific community. Genome engineering based on CRISPR is very useful for the improvement of traits in livestock, as for example in the protection of pigs against viruses, where the system CRISPR/Cas9 was used to generate pigs lacking functional CD163, a receptor responsible for the porcine reproductive and respiratory syndrome virus cell internalization.[110] Moreover, this technology is being used in crops not

[109] Barrangou, R.; Doudna, J. A. *Nat. Biotechnol.,* **2016**, *34*, 933–941.

[110] Whitworth, K. M.; Rowland, R. R. R.; Ewen, C. L.; Trible, B. R.; Kerrigan, M. A.; Cino-Ozuna, A. G.; Samuel, M. S.; Lightner, J. E.; McLaren, D. G.; Mileham, A. J.; Wells, K. D.; Prather, R. S. *Nat. Biotechnol.,* **2016**, *34*, 20–22.

[108] Carroll, D. *Annu. Rev. Biochem.,* **2014**, *83*, 409–439.

only to increase yield, improve drought tolerance and increase growth in poor nutrient conditions, but also to produce crops with improved nutritional properties,[111] as has been reported the efficient simultaneous multiple gene knockouts, native gene editing, and site-specific gene integration in corn.[112]

From the therapeutic point of view, Cas9 has been developed as an antimicrobial agent that can be used to specifically target antibiotic-resistant and/or highly virulent strains of bacteria. For example, Cas9 reprogrammed to target virulence genes, kills only virulent *Staphylococcus aureus*.[113] Recent reports about gene therapy applications showed that CRISPR/Cas9 systems corrected the genetic defect in mouse models of Duchenne muscular dystrophy (DMD),[114] a monogenic disease caused by mutations in the gene encoding the protein known as dystrophin, necessary for muscle cell integrity. CRISPR/Cas9 system also presents potential applications for the treatment of viral infections such as HIV[115] and hepatitis B,[116] where the efficient identification and mutation of a specific region in the DNA sequences inhibits the viral gene expression and replication. Another important revolutionary application was the achievement of the precise genetic modifications in primates resulting from gene editing in embryos by the injection of Cas9-mRNA within gRNAs into one-cell-stage embryos.[117] A similar approach could be used to alter DNA in human embryos to prevent non-complex hereditary diseases, but also to attempt alteration of complex features, which has provoked an extensive ethical discussion, such as the recent controversial case of the twin Chinese girls with edited genomes.[118]

1.1.3 Advantages in the direct delivery of CRISPR/Cas9 system

The *S. pyogenes* Cas9 (*Sp*Cas9) is a large protein (>1300 amino acids, 160 kDa) and as a consequence, its intracellular delivery is a challenging task.[119] One of the main concerns when editing mammalian cells is the Cas9 format in a way that minimizes the size for more efficient delivery and at the same time preserves Cas9 activity, restricts immunological activation and limits off-target cleavage. For *in vivo* genome editing, three different possibilities of

[111] Belhaj, K.; Chaparro-Garcia, A.; Kamoun, S.; Patron, N. J.; Nekrasov, V. *Curr. Opin. Biotechnol.,* **2015**, *32*, 76–84.

[112] Svitashev, S.; Young, J. K.; Schwartz, C.; Gao, H.; Falco, S. C.; Cigan, A. M. *Plant Physiol.,* **2015**, *169*, 931–945.

[113] Bikard, D.; Euler, C.; Jiang, W.; Nussenzweig, P. M.; Goldberg, G. W.; Duportet, X.; Fischetti, V. A.; Marraffini, L. A.; Biotechnol, N. *Nat Biotechnol,* **2014**, *32*, 1146–1150.

[114] Long, C.; McAnally, J. R.; Shelton, J. M.; Mireault, A. A.; Bassel-Duby, R.; Olson, E. N. *Science,* **2014**, *345*, 1184–1188.

[115] Hu, W.; Kaminski, R.; Yang, F.; Zhang, Y.; Cosentino, L.; Li, F.; Luo, B.; Alvarez-Carbonell, D.; Garcia-Mesa, Y.; Karn, J.; Mo, X.; Khalili, K. *Proc. Natl. Acad. Sci.,* **2014**, *111*, 11461–11466.

[116] Zhen, S.; Hua, L.; Liu, Y. H.; Gao, L. C.; Fu, J.; Wan, D. Y.; Dong, L. H.; Song, H. F.; Gao, X. *Gene Ther.,* **2015**, *22*, 404–412.

[117] Niu, Y.; Shen, B.; Cui, Y.; Chen, Y.; Wang, J.; Wang, L.; Kang, Y.; Zhao, X.; Si, W.; Li, W.; Xiang, A. P.; Zhou, J.; Guo, X.; Bi, Y.; Si, C.; Hu, B.; Dong, G.; Wang, H.; Zhou, Z.; Li, T.; Tan, T.; Pu, X.; Wang, F.; Ji, S.; Zhou, Q.; Huang, X.; Ji, W.; Sha, J. *Cell,* **2014**, *156*, 836–843.

[118] Cyranoski, D. *Nature,* **2019**, *566*, 440–442.

[119] Kelton, W. J.; Pesch, T.; Matile, S.; Reddy, S. T. *Chim. Int. J. Chem.,* **2016**, *70*, 439–442.

delivering Cas9 and the gRNA are available (Figure 12): viral or plasmid DNA (pDNA), mRNA and ribonucleoprotein (RNP, complex formed by the protein Cas9 and the gRNA).[120]

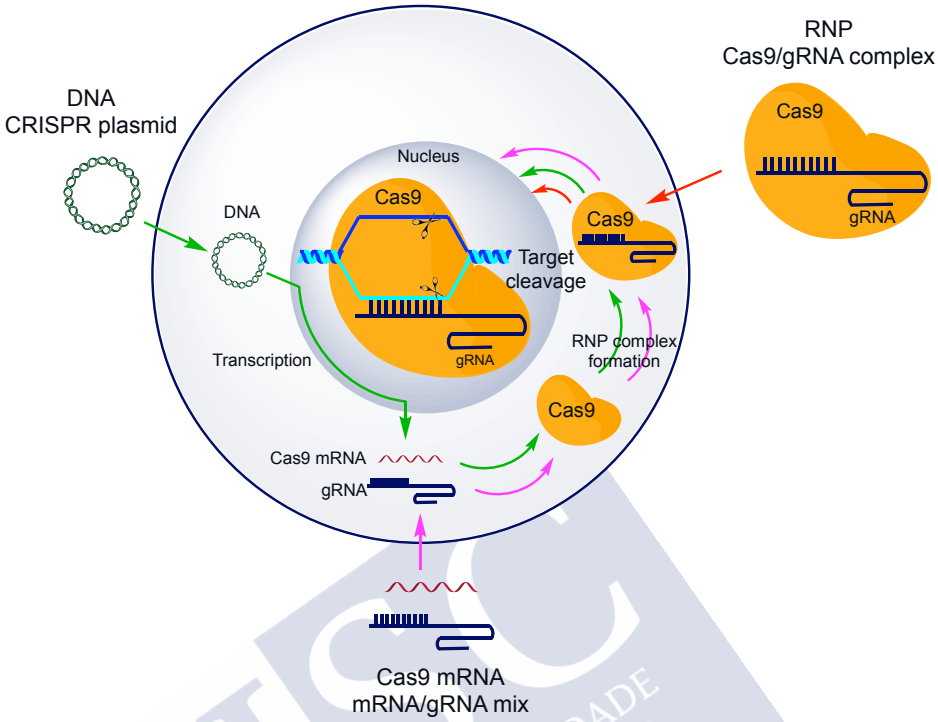


Figure 12. Delivery strategies for the CRISPR/Cas9 system.

The delivery of recombinant plasmids that encode the Cas9 protein and the guide RNA sequences under separate promoters is the most stable and cost-effective of all the aforementioned approaches. This strategy has attractive advantages due to its simplicity and the possibility of carrying out massive multiplex gene editing by the inclusion of multiple gRNAs expressed from the same plasmid. However, for therapeutic engineering of mammalian cells with CRISPR/Cas9, the use of plasmids is usually limited due to different restrictions. First of all, the delivery of pDNA could possibly increase off-target effects and induce undesired side effects due to the continued expression of Cas9 and the persistence of the protein in cells for several days post-transfection.[121] Secondly, another obstacle for pDNA transfection is the necessity of nuclear internalization. Moreover, the random integration of all or a part of the pDNA into the host genome is difficult to detect and can result in permanent DNA recombination and persistent expression, a situation that enhances potential off target effects and immunogenic responses. Finally, pDNA application is often damaging to cells, which would activate cyclic GMP-AMP synthase activation.[122]

[120] Wan, T.; Niu, D.; Wu, C.; Xu, F.-J.; Church, G.; Ping, Y. *Mater. Today,* **2019**, *26*, 40–66.

[121] Kim, S.; Kim, D.; Cho, S. W.; Kim, J.-S.; Kim, J.-S. *Genome Res.,* **2014**, *24*, 1012–1019.

[122] Gao, D.; Li, T.; Li, X.-D.; Chen, X.; Li, Q.-Z.; Wight-Carter, M.; Chen, Z. J. *Proc. Natl. Acad. Sci.,* **2015**, *112*, E5699–E5705.

In the case of mRNA delivery, the translation of mRNA into Cas9 protein happens in the ribosomes found in the cytoplasm, avoiding the difficult challenge of crossing the nuclear membrane.[123] However, mRNA is relatively unstable, and it has to be combined with gRNA separately when mRNA is considered as a choice of delivery form. This suggests that a delivery vehicle should be capable of loading both mRNA and gRNA simultaneously for the genome editing application, thus promoting the difficulties in terms of loading and release.

All these disadvantages could be overcome by the direct delivery of RNP. In this situation the Cas9 protein and gRNA can be complexed as one cargo that can be directly delivered into the target cells. One of the advantages of the direct delivery of Cas9 RNP by non-viral vectors is the expectation that off-target editing effects are reduced as compared with Cas9 pDNA delivery. Therefore, the direct delivery of the Cas9 RNP inside cells would be beneficial as it functions as a transitory effector that could be subsequently degraded. Cas9 direct delivery enables the transient accumulation of the protein in the nucleus avoiding the problems of plasmid integration, prolonged residual activity, off target effects and immune responses.

1.1.4 Previous delivery methods for CRISPR/Cas9

To date, most of the existing delivery methods of CRISPR/Cas9 have relied on either physical methods or viral vectors. A variety of physical methods, such as electroporation,[124,125] hydrodynamic injection[126] and microinjection[127] have been successfully applied for CRISPR/Cas9 delivery. However, they are considered to be less suitable for *in vitro* and *in vivo* delivery due to the difficulties of cell function maintenance as well as for practical applications. A more complex alternative is the packaging of the Cas9 DNA in a single-stranded form into a non-integrating virus such as the adeno-associated virus (AAV).[128] However, the maximum loading capacity for AAV is around 4.5 kb (~160 kDa), which makes combination of both Cas9 and gRNA into a single capsid challenging, as the *Sp*Cas9 gene is almost at the loading size limit of viral vectors. Thus, smaller Cas9 variants have been developed, but they suffer from a reduction in their cleavage efficiency. Moreover, safety issues, such as carcinogenesis, insertional mutagenesis and immunogenicity, also limit the clinical translation of these viral vectors. Therefore, chemical methods of delivering Cas9 RNP system *via* non-viral vectors have the potential to address most of these limitations related with biosafety, loading and packaging capacities.

---

[123] Glass, Z.; Lee, M.; Li, Y.; Xu, Q. *Trends Biotechnol.,* **2018**, *36*, 173–185.

[124] Qin, W.; Dion, S. L.; Kutny, P. M.; Zhang, Y.; Cheng, A. W.; Jillette, N. L.; Malhotra, A.; Geurts, A. M.; Chen, Y. G.; Wang, H. *Genetics,* **2015**, *200*, 423–430.

[125] Fei, J.-F.; Knapp, D.; Schuez, M.; Murawala, P.; Zou, Y.; Pal Singh, S.; Drechsel, D.; Tanaka, E. M. *npj Regen. Med.,* **2016**, *1*, 16002.

[126] Yin, H.; Xue, W.; Chen, S.; Bogorad, R. L.; Benedetti, E.; Grompe, M.; Koteliansky, V.; Sharp, P. A.; Jacks, T.; Anderson, D. G. *Nat. Biotechnol.,* **2014**, *32*, 551–553.

[127] Horii, T.; Arai, Y.; Yamazaki, M.; Morita, S.; Kimura, M.; Itoh, M.; Abe, Y.; Hatada, I. *Sci. Rep.,* **2015**, *4*, 4513.

[128] Senís, E.; Fatouros, C.; Große, S.; Wiedtke, E.; Niopek, D.; Mueller, A.-K.; Börner, K.; Grimm, D. *Biotechnol. J.,* **2014**, *9*, 1402–1412.

Non-viral vectors have been successfully used for CRISPR/Cas9 delivery, both *in vivo* and *in vitro*, and innovative protocols about their complexation have been recently investigated and reported.[129] Lipid nanoparticles are one of the most extensively used delivery carriers, which combine negatively charged nucleic acids with positively charged lipids.[120] Cas9 nuclease protein complexed with polyanionic single guide RNA can be efficiently delivered in functional form into mammalian cells using cationic lipid formulations.[130] Similar strategies using bioreducible lipid-like materials have been developed to load and deliver the anionic Cas9/gRNA complexes,[131] where the electrostatic assembly of nanoparticles mediates potent protein delivery and genome editing features (Figure 13). These bioreducible lipids deliver protein cargos into cells with high efficiency, facilitating their endosomal escape and directing proteins to their intracellular target site.

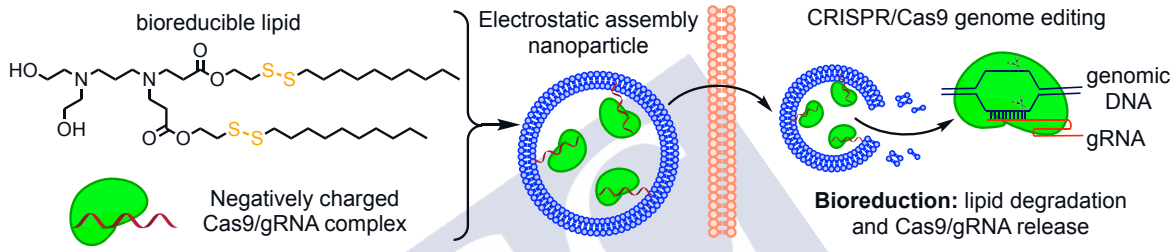


Figure 13. Schematic representation of protein delivery and genome editing using a bioreducible lipid-like material and negatively charged Cas9/gRNA complex.

Apart from lipid nanoparticles, polymers represent another appealing class of materials for designing efficient non-viral vectors for CRISPR/Cas9 systems.[120] Recently, the direct conjugation of a cationic polymer branched PEI (bPEI) with *Sp*Cas9 protein through covalent bonding has been reported (Figure 14).[132] The resulting bPEI-Cas9 conjugate was further attached with gRNA to form nanocomplexes to combat antibiotic resistance in bacteria. This polymer-conjugated Cas9 showed significant uptake and therefore higher editing efficiency as compared with the delivery of native Cas9 protein.

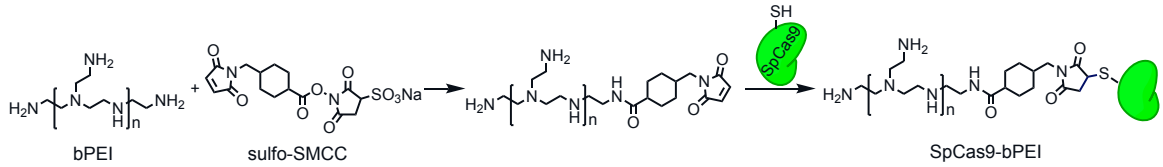


Figure 14. Activation of bPEI using sulfosuccinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) and reaction with the free sulfhydryl groups of *Sp*Cas9.

[129] Givens, B. E.; Naguib, Y. W.; Geary, S. M.; Devor, E. J.; Salem, A. K. *AAPS J.,* **2018**, *20*, 108.

[130] Zuris, J. A.; Thompson, D. B.; Shu, Y.; Guilinger, J. P.; Bessen, J. L.; Hu, J. H.; Maeder, M. L.; Joung, J. K.; Chen, Z.; Liu, D. R. *Nat. Biotechnol.,* **2015**, *33*, 73–80.

[131] Wang, M. et al. *Proc. Natl. Acad. Sci.,* **2016**, *113*, 2868–2873.

[132] Kang, Y. K.; Kwon, K.; Ryu, J. S.; Lee, H. N.; Park, C.; Chung, H. J. *Bioconjug. Chem.,* **2017**, *28*, 957–967.

[120] Wan, T.; Niu, D.; Wu, C.; Xu, F.-J.; Church, G.; Ping, Y. *Mater. Today,* **2019**, *26*, 40–66.

Gold nanoparticles have also been explored as promising synthetic carriers in the field of gene delivery as they possess good biocompatibility and excellent chemical stability.[120] The first example of an editing nanostrategy was based on the co-delivery of Cas9 and gRNA into cells using gold nanoparticles.[133] These cationic arginine-decorated gold nanoparticles were used to complex engineered Cas9 protein modified with a negatively charged glutamate peptide tag and gRNA (Figure 15). The resulting nanoassemblies could then be fused into the cell membrane to deliver encapsulated Cas9/gRNA into the cytoplasm.

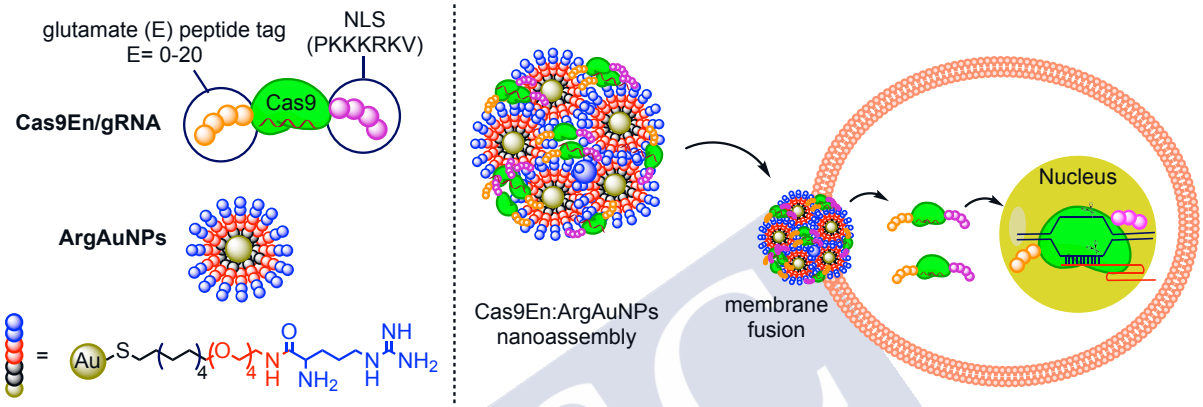


Figure 15. Engineering of the Cas9 protein and arginine gold nanoparticles (ArgAuNPs) for the intracellular delivery of Cas9 RNP *via* membrane fusion.

Furthermore, CPPs have been extensively used for the delivery of multiple cargoes inside cells due to their capability to cross the cell membrane.[120] In the initial studies, CPPs were directly conjugated with the anionic Cas9 RNP by covalent bonds and confirmed to target the loci inside living cells (Figure 16).[134] The delivery of CPP-Cas9 RNP resulted in less off-target effects, weaker immune responses and lower cytotoxicity. However, the covalent fusion of the CPP with the Cas9 protein required multiple incubation steps with the cells to achieve low levels of edition.

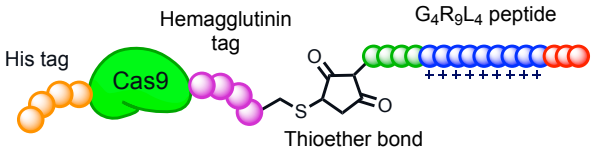


Figure 16. Schematic representation of Cas9 protein conjugated to a CPP by a covalent bond. Spheres in green are glycine amino acids, in blue arginine amino acids and in red leucine amino acids.

[133] Mout, R.; Ray, M.; Yesilbag Tonga, G.; Lee, Y.; Tay, T.; Sasaki, K.; Rotello, V. M. *ACS Nano,* **2017**, *11*, 2452–2458.

[134] Ramakrishna, S.; Kwaku Dad, A. B.; Beloor, J.; Gopalappa, R.; Lee, S. K.; Kim, H. *Genome Res.,* **2014**, *24*, 1020–1027.

[120] Wan, T.; Niu, D.; Wu, C.; Xu, F.-J.; Church, G.; Ping, Y. *Mater. Today,* **2019**, *26*, 40–66.

2 *PEPTIDE/CAS9 NANOSTRUCTURES FOR RIBONUCLEOPROTEIN CELL MEMBRANE TRANSPORT AND GENE EDITION*

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2.1 Precedents and Objectives

There is an emerging interest in the direct delivery of Cas9 RNP inside cells by using non-viral vectors. Although it has been reported that CPPs can efficiently deliver Cas9 inside cells,[134] this strategy required the covalent fusion of the protein to the CPP and several rounds of transfection with the fused protein for reaching adequate levels of edition. These limitations hinder the application of gene edition technology in cell repairing and genetic treatments.

The general objective of this chapter is to develop a new strategy for the direct delivery of Cas9 RNP using cell-penetrating peptides in a fully supramolecular manner (Figure 17).

As in chapter 1, the proposed peptide scaffold will be based on an amphipathic backbone with potential α-helical folding. In addition, dynamic covalent bonds such as hydrazides will be inserted in the sequence allowing its modification with appropriate aldehydes. The active amphiphilic peptide will be prepared by a hydrazone bond formation between the cationic peptide scaffold and a hydrophobic aldehyde.

A preliminary screening using a variety of hydrophobic aldehydes will be performed for the fast identification of amphiphilic peptides capable of delivering Cas9 with good efficiencies. Therefore, we plan to characterize in detail the formulations of the best peptide candidate obtained in the initial screening.

The potential of the delivery and gene edition of Cas9 RNP will be characterized and compared with Lipofectamine2000 by a T7-E1 endonuclease assay by fixing carrier concentration and increasing RNP amount in different cell lines.

The uptake mechanism of the complexes will be studied in the presence of different endocytic inhibitors by following the fluorescence of a TAMRA-labeled peptide derivative.

Furthermore, the size and the stability of the peptide/protein nanoparticles will be determined by using dynamic light scattering at different ratios and at physiological pH.

Electron microscopy techniques such as TEM, SEM and STEM will be used for the morphological characterization of the best functional particles.

The presence and the position of the protein will be studied by exploiting the histidine tag present in the Cas9 RNP combined with gold nanoparticles coordinated to a nickel (II) cation.

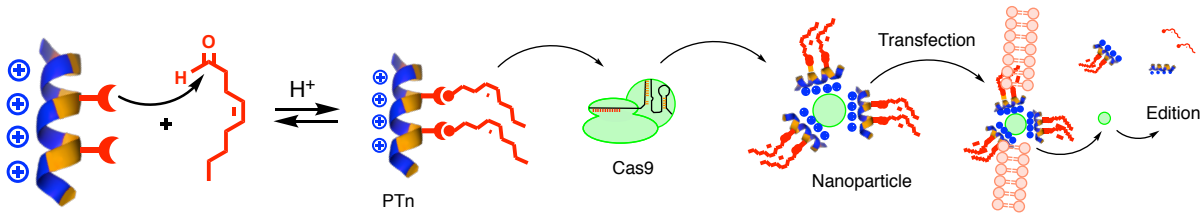


Figure 17. Schematic representation of the amphiphilic peptide for the delivery of Cas9 RNP.

[134] Ramakrishna, S.; Kwaku Dad, A. B.; Beloor, J.; Gopalappa, R.; Lee, S. K.; Kim, H. *Genome Res.*, **2014**, *24*, 1020–1027.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2.3 Conclusions

To sum up, in this chapter we have confirmed that the modulation of transfecting properties of helical peptides by using hydrazone formation constitutes an efficient approach for the direct delivery of Cas9 inside human cells. The strategy consisted of a fully supramolecular process in which the peptide transporter recognition and interaction with Cas9 machinery takes place through purely electrostatic non-covalent interactions.

The hydrazone modification of the peptide scaffold was carried out in fully biocompatible conditions to tune the hydrophobic properties of the peptide transporter and did not require any isolation or purification steps of the amphiphiles to deliver the Cas9 RNP.

In a preliminary screening, HeLa cells expressing GFP were used for the delivery of Cas9-gRNA against GFP. The knock down of the fluorescence was measured, reporting oleic aldehyde as one of best hydrophobic tails for this system.

We have tested cleavage efficiency on a gene present in HeLa cells by a T7-E1 nuclease assay, comparing the peptide linked to the oleic aldehyde with Lipofectamine2000, showing that, despite both efficiencies were comparable, our system provided better results in terms of a reduced toxicity at lower concentrations of RNP. Cleavage efficiency was also confirmed in two further cell lines (A549 and DF1) demonstrating the versatility of this peptide carrier for the delivery of Cas9.

We have characterized the peptide/protein complexes by DLS and TEM providing information about the size of the nanoparticle structures (~270 nm). By gel electrophoresis, we have also confirmed the complexation between the peptide and the Cas9 RNP.

Zeta potential measurements revealed a decrease in the charge of the peptide/Cas9 RNP complexes (4.2 mV) as compared with the peptide amphiphile alone (10.4 mV), confirming the formation of the desired complex.

To demonstrate the presence of Cas9 in this nanoaggregates, the histidine tag of Cas9 protein was labeled with Ni-NTA gold nanoparticles. STEM micrographs revealed the gold nanoparticles attached to the peptide/protein nanostructures.

Internalization experiments in the presence of endocytic inhibitors revealed that the cellular uptake of the peptide/Cas9 RNP complexes was energy dependent (>90%) pointing to a macropinocytic mechanism.

Confocal microscopy and flow cytometry showed a considerable decrease in the number of cellular lysosomes as compared with the non-treated cells. This result demonstrated that the peptide/Cas9 RNP is able to disrupt the endosomal membrane leading to the subsequent release of the cargo into the cytosol.

The endosomal escape of the peptide was then confirmed by a nuclear translocation assay, which demonstrated that the peptide was able to reach the cytosol. After repeating the experiment in the presence of the macropinocytosis inhibitor, the nuclear translocation effect

was highly reduced. Thus, the peptide that reached the cytosol in the absence of the inhibitor was escaping from that particular endocytic route.

This work demonstrated the versatility of the methodology for the screening of adaptable carriers for the direct delivery of Cas9 in different cell lines.

CHAPTER 3: GENE THERAPY

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1 INTRODUCTION

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1.1 Therapeutic interest of gene delivery

Gene therapy is a novel form of molecular medicine, which has had a big impact on human health since the end of last century. The recognition of the central role of DNA in cell biology and its fundamental importance in the control of cellular processes has promoted the exponential growth of the medical applications in this field.[135] Although the advent of recombinant DNA technology in modern medicine will allow prenatal genetic screening and treatment, the vast majority of those born with a certain disease are likely to be helped by gene therapy approaches. The concept of gene therapy involves the insertion of functional genes into specific cells of a patient for replacement or supplementation of mutated or missing genes, which results in the production of a therapeutic protein or the inhibition of a malfunctioning protein.[136] Thus, in addition to the possibility of correcting inherited genetic disorders like cystic fibrosis, hemophilia and familial hypercholesterolemia and so on, gene therapy approaches are being studied to be used to combat acquired diseases, like cancer, AIDS, infectious diseases and Parkinson's or Alzheimer's disease.[135]

Conceptually, gene therapy involves identifying appropriate DNA sequences and cell types and is followed by the development of suitable ways to get the corresponding nucleotides or ribonucleoproteins into the targeted cells. In 1990 the first human gene therapy trial was begun and involved the transfer of adenosine deaminase (ADA) gene into lymphocytes of a patient having an otherwise lethal defect in this enzyme, which produces immune deficiency.[137] The results of this initial trial have been very encouraging and have helped to stimulate further clinical tests.

Antisense therapy, first reported in the late 70s,[138] can also be used to turn off a disease-causative gene by the action of an antisense oligodeoxynucleotide that inactivates the encoded mRNA. However, inadequate target engagement, insufficient biological activity and off-target toxic effects have obstructed progress translating antisense oligonucleotides (ASOs) to the clinic.[139] Over the years, novel chemical modifications of ASOs have been employed to address these issues, which, in combination with the elucidation of the mechanism of action and improved clinical trials, have fueled the translation of ASO-based strategies into therapies.[140] Although many neurological conditions lack an effective treatment, as research progressively disentangles the pathogenic mechanisms of these diseases, they provide an ideal platform to test ASO-based strategies. This steady progress reached a highpoint in the past few years with approvals of ASOs for the treatment of, for example, Duchenne muscular dystrophy.[141] This, represented a landmark in a field were disease-modifying therapies were

[135] Wang, T.; Upponi, J. R.; Torchilin, V. P. *Int. J. Pharm.,* **2012**, *427*, 3–20.
[136] Mulligan, R. *Science,* **1993**, *260*, 926–932.
[137] Miller, A. D. *Nature,* **1992**, *357*, 455–460.
[138] Stephenson, M. L.; Zamecnik, P. C. *Proc. Natl. Acad. Sci.,* **1978**, *75*, 285–288.
[139] Rinaldi, C.; Wood, M. J. A. *Nat. Rev. Neurol.,* **2018**, *14*, 9–22.
[140] Juliano, R. L. *Nucleic Acids Res.,* **2016**, *44*, 6518–6548.
[141] Koo, T.; Wood, M. J. *Hum. Gene Ther.,* **2013**, *24*, 479–488.

practically non-existent. This technology holds the potential to dramatically change the therapeutic treatment of neurological and non-neurological conditions in the near future.

The discovery of RNA interference constituted a breakthrough for gene therapy.[142] Synthetic small interfering RNA (siRNA duplex) therapeutics, have emerged as one of the most promising, potential drugs in gene therapy. siRNA specifically binds to its targeted mRNA, resulting in subsequent catalytic silencing of gene expression.[143] Recently, US regulators have approved the first therapy based on RNAi, which can be used to silence specific genes linked to disease.[144] Patisiran is a RNA interference therapeutic agent that targets a rare condition impairing heart and nerve function.[144] Its recent approval consitutes a landmark for a field that has resisted for nearly two decades to prove its worth in the clinic.

1.2 Types of nucleic acids used in gene delivery

The use of nucleic acids as therapeutic agents has great potential for the treatment of heritable and acquired diseases.[145] Beyond therapeutic applications, DNA delivery, especially *via* the non-viral route (*i.e.* transfection), has become a powerful and popular research tool for elucidating gene structure, regulation, and function.[146]

Plasmid DNA (pDNA) is commonly found as large circular double-stranded DNA molecules of several kilobases (kb) in bacteria. Their site of action inside the cell is the nucleus, where pDNA must enter to be transcribed.[147] These plasmids contain several basic components such as the transgene expression system (*i.e.* promoter, gene of interest and terminator), regulatory signals, antibiotic resistance marker, origin of replication and the remaining bacterially-derived plasmid backbone (BB).[148] The removal of unnecessary pDNA sequences is recommended to maintain the pDNA molecule small and easy to manipulate.[149] Thus, minicircles in which the BB has been removed may offer advantages over conventional plasmids. The native plasmid systems are only active *in vivo* for only 1 or 2 months, whereas some other derivative constructs, have demonstrated improved expression durations in cells for months or years.[148] Although the use of pDNA could result in insertional mutagenesis by recombination with cellular DNA, its delivery has been essential for both the expression of a protein to restore a function or to develop an immune response against it, as well as the expression of regulatory RNAs.[147]

---

[142] Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature,* **1998**, *391*, 806–811.

[143] Lam, J. K. W.; Chow, M. Y. T.; Zhang, Y.; Leung, S. W. S. *Mol. Ther.-Nucleic Acids,* **2015**, *4*, e252.

[144] Adams, D.; Gonzalez-Duarte, A.; O'Riordan, W. D.; Yang, C.-C.; Ueda, M.; Kristen, A. V.; Tournev, I.; Schmidt, H. H.; Coelho, T.; Berk, J. L.; Lin, K.-P.; Vita, G.; Attarian, S.; Planté-Bordeneuve, V.; Mezei, M. M.; Campistol, J. M.; Buades, J.; Brannagan, T. H.; Kim, B. J.; Oh, J.; Parman, Y.; Sekijima, Y.; Hawkins, P. N.; Solomon, S. D.; Polydefkis, M.; Dyck, P. J.; Gandhi, P. J.; Goyal, S.; Chen, J.; Strahs, A. L.; Nochur, S. V.; Sweetser, M. T.; Garg, P. P.; Vaishnaw, A. K.; Gollob, J. A.; Suhr, O. B. *N. Engl. J. Med.,* **2018**, *379*, 11–21.

[145] Dunbar, C. E.; High, K. A.; Joung, J. K.; Kohn, D. B.; Ozawa, K.; Sadelain, M. *Science* **2018**, *359*, eaan4672.

[146] Luo, D.; Mark, W, S. *Nat. Biotechnol.,* **2000**, *18*, 33–37.

[147] Lostalé-Seijo, I.; Montenegro, J. *Nat. Rev. Chem.,* **2018**, *2*, 258–277.

[148] Hill, A. B.; Chen, M.; Chen, C.-K.; Pfeifer, B. A.; Jones, C. H. *Trends Biotechnol.,* **2016**, *34*, 91–105.

[149] Gill, D. R.; Pringle, I. A.; Hyde, S. C. *Gene Ther.,* **2009**, *16*, 165–171.

Messenger RNAs (mRNA) are long single-stranded RNA molecules, which encode the information for one protein. They are composed of several kb, as long as their length depends on the size of the encoded protein. They have been investigated for therapeutic applications as an alternative to DNA-based gene therapy. Although it is less stable than DNA, the advantages of mRNA are clear.[150] mRNA is easily produced by an *in vitr*o process without the need of living organisms. The manufacturing process is general and can be quickly applied for numerous different applications and genes of interest. Naturally, the protein expression can only be transient, which is sufficient for several applications and even favorable for some therapeutic approaches to better control pharmacokinetics and dosing. In addition, the risk of recombination with the cell's genome and insertional mutagenesis are excluded when using mRNA.[151] Another benefit of mRNA is that it performs its function in the cytoplasm, without needing to cross the nuclear membrane, which results in almost immediately protein translation after its delivery. The simplest application of mRNA is based on its incorporation in the protein synthesis machinery of the target cell to induce expression of a desired protein restoring a function or developing an immune response against it.[152]

Antisense oligonucleotides (ASOs) are short, synthetic, single-stranded oligonucleotides typically 20-25 bases in length, which bind RNA by complementary base pairing and that can alter RNA and reduce, restore or modify protein expression through several different mechanisms.[139] ASOs are emerging as a therapeutic alternative to treat diseases with known genetic origin.[139] They can be occasionally designed to target and bypass or overcome a patient's genetic mutation, in particular those lesions that compromise normal pre-mRNA processing. ASOs serve as highly selective sequence pairs to specific regions of mRNA and regulate the translation of genetic material into functional proteins.

Short interfering RNAs (siRNAs) and microRNAs (miRNAs), have emerged as critical regulators in the expression and function of eukaryotic genomes.[153] siRNAs and miRNAs share many similarities, both are duplex RNA molecules of around 22 nucleotides that exert gene silencing effects at the post-transcriptional level by targeting mRNA.[143] However, their mechanisms of action are different, which results in siRNA being highly specific with only one mRNA target, whereas miRNA comprises two approaches: miRNA inhibition and miRNA replacement.[143] The inhibition approach resembles antisense therapy, with synthetic single stranded RNAs acting as miRNA antagonists to inhibit the action of the endogenous miRNA, while in the replacement approach, synthetic miRNAs are used to mimic the function of the endogenous miRNAs. Thus, it leads to mRNA degradation or inhibition producing a gene silencing effect. Single-stranded forms of both RNAs associate with effector assemblies known as RNA-induced silencing complexes (RISCs), located in the cytosol,

[150] Vallazza, B. et al. *Wiley Interdiscip. Rev. RNA,* **2015**, *6*, 471–499.

[151] McIvor, R. S. *Mol. Ther.,* **2011**, *19*, 822–823.

[152] Hajj, K. A.; Whitehead, K. A. *Nat. Rev. Mater.,* **2017**, *2*, 1–17.

[153] Carthew, R. W.; Sontheimer, E. J. *Cell,* **2009**, *136*, 642–655.

[139] Rinaldi, C.; Wood, M. J. A. *Nat. Rev. Neurol.,* **2018**, *14*, 9–22.

[143] Lam, J. K. W.; Chow, M. Y. T.; Zhang, Y.; Leung, S. W. S. *Mol. Ther.-Nucleic Acids,* **2015**, *4*, e252.

being directed to specific mRNAs for degradation by siRNAs or regulating mRNA stability or translation by miRNAs. When in the nucleus, siRNAs induce long-term silencing by DNA methylation and miRNAs induce chromatin reorganization.[153]

1.3 Gene delivery vectors

The key challenge for oligonucleotide-based therapeutics is the delivery of active oligonucleotides to their site of action inside the cells within tissues.[140] Successful gene delivery depends on the ability of the vector of choice to target a specific cell type, enter the cell and obtain sufficient levels of gene expression. This is not an easy task due to the presence of several barriers that viral and non-viral vectors have to tackle. Consequently, many delivery vectors have been developed to overcome both extracellular and intracellular gene delivery obstacles. Viral vectors are the most widely used systems for gene delivery, as they are the most effective and easiest methods for transferring genes of interest into a cell. However, their use as delivery vehicles has many drawbacks, such as storage difficulties, gene carrying capacity and residual viral elements, which can potentially cause insertional mutagenesis, cytotoxicity and immunogenicity.[154] These limitations of viral vectors have been avoided with the development of synthetic non-viral vectors, which include liposomes, polymers, nanoparticles, etc.[155] A comparison between both carriers is depicted in Table 1.

Table 1. Advantages and disadvantages of viral and non-viral vectors for gene delivery.

	ADVANTAGES	DISADVANTAGES
VIRAL	-High transduction efficiency. -Natural tropism confers the capability for infection of many cell types. -Intrinsic mechanism for endosomal escape. -Natural mechanism for nuclear import of genes.	-Immune response is strong and multiple-injections are limited. -Can cause chromosomal insertion. -Difficult to produce in large scale. -Can only carry limited sized genes. -Can cause toxicity and may be contaminated with live virus.
NON-VIRAL	-Low immunogenicity. -Easy to make and with quality control in mass production. -Can carry large-sized DNA. -Can be functionalized for targeting, endosomal escape and nuclear import.	-Transfection efficiency is low. -Can be toxic at high concentrations. -Lack of intrinsic tropism. -Lack of intrinsic mechanism for endosomal escape and for nuclear import of genes.

[154] Sheridan, C. *Nat. Biotechnol.,* **2011**, *29*, 121–128.

[155] Gottfried, L. F.; Dean, D. A. In *Novel Gene Therapy Approaches*; InTech, **2013**; 75–88.

[140] Juliano, R. L. *Nucleic Acids Res.,* **2016**, *44*, 6518–6548.

[153] Carthew, R. W.; Sontheimer, E. J. *Cell,* **2009**, *136*, 642–655.

1.4 Gene delivery barriers

There has been a rapid progress in the development of therapeutic nucleic acids based on DNA and RNA for the treatment of genetic and acquired disorders.[156] However, nucleic acids are highly hydrophilic negatively charged macromolecules, properties that restrict their binding to cell membranes and their diffusion across them. Their internalization depends on the endocytic pathway, which can result in endosomal entrapment and lysosomal degradation, issues that reduce their availability at their sites of action. Moreover, nucleic acids that are able to escape from the endosome, have to fight with the cytosolic viscosity and dense organelles, which might prevent their movement towards target sites. As it has already been mentioned, in the case of pDNA, the nuclear envelope represents an extra barrier, since it must enter the nucleus to be transcribed.[157]

The most important and difficult challenge in gene therapy is the problem of *in vivo* delivery. In many cases, systemic administration of genes is needed due to the fact that many disease sites are not easy to access. Under these circumstances, bioactive nucleotides have to penetrate through a series of systemic barriers in order to achieve the desirable efficiency and to reduce side effects. In the blood circulation, nucleic acids must avoid uptake by macrophages, clearance by renal filtration and degradation by endogenous nucleases.[158] Furthermore, undesired interactions with blood components and non-target cells such as the reticuloendothelial system,[140] are important obstacles in targeting genes *in vivo*. The extracellular matrix is another important barrier for gene delivery as it resists the movement of genetic material to target cells due to its dense polysaccharides and fibrous proteins.[159]

1.5 Overcoming gene delivery barriers with non-viral vectors

1.5.1 Cargo/carrier particle formation

Nucleic acids are exposed to a variety of environmental factors such as, pH or enzymes (*i.e.* nucleases) that can degrade or destroy them. Thus, their complexation by cationic polymers or lipids is a method widely used to prevent their destruction by nucleases.

Polyethylenimine (PEI) is one such polymers, positively charged and with a molecular weight of around 22-25 kDa, which can be linear or branched, and combined with nucleic acids *via* electrostatic interactions.[160] PEI has been used for non-viral transfection *in vitro* and *in vivo* and has an advantage over other polycations as it strongly compacts DNA with an intrinsic endosomolytic activity.[161] At high N/P ratios (ratio of the nitrogen atoms of PEI to

[156] Read, M. L.; Logan, A.; Seymour, L. W. *Adv. Genet.,* **2005**, *53*, 19–46.

[157] Lam, A. P.; Dean, D. A. *Gene Ther.,* **2010**, *17*, 439–447.

[158] Ogris, M.; Brunner, S.; Schüller, S.; Kircheis, R.; Wagner, E. *Gene Ther.,* **1999**, *6*, 595–605.

[159] Dhaliwal, A.; Lam, J.; Maldonado, M.; Lin, C.; Segura, T. *Soft Matter,* **2012**, *8*, 1451–1459.

[160] Boussif, O.; Lezoualc'h, F.; Zanta, M. A.; Mergny, M. D.; Scherman, D.; Demeneix, B.; Behr, J. P. *Proc. Natl. Acad. Sci.,* **1995**, *92*, 7297–7301.

[161] Lungwitz, U.; Breunig, M.; Blunk, T.; Göpferich, A. *Eur. J. Pharm. Biopharm.,* **2005**, *60*, 247–266.

[140] Juliano, R. L. *Nucleic Acids Res.,* **2016**, *44*, 6518–6548.

DNA phosphates), the positive net charge of the corresponding complexes increases, thus the electrostatic interaction between these positively charged complexes and negatively charged cell surface, facilitates cell binding and consequently, high levels of gene expression. Thus, PEI combines a high membrane destabilizing potential with a high DNA packaging activity, which protects DNA from degradation and increases the probability that pDNA can reach the nucleus intact.[161] After endocytic internalization, the buffering capacity of amine groups of PEI in the acidic pH of endosomes, results in endosomal rupture and release of the complexes into the cytoplasm.[162] This phenomenon called "proton-sponge effect" could explain why polyplexes are able to escape from the endosomes. However, PEI shows toxicity, which depends on the MW and branching degree of PEI, particle size and zeta pontential.[162] Toxicity can be also be related with free PEI, which results in cell membrane destabilization before transfection, or related with the proceeding of PEI/DNA complexes after getting into cell.

Liposomal vectors have also been used to encapsulate DNA. A primary demonstration of *in vivo* gene expression was performed through the use of pH sensitive liposomes.[163] However, their application for gene therapy resulted in a low entrapment efficiency and high serum sensitivity[164] until the introduction of lipofection methodology, which used the cationic lipid DOTMA (2,3-dioleyloxypropyl-1-trimethylammonium chloride).[165] Cationic lipids with amphipathic behavior, interact with the negatively charged phosphate backbone of nucleic acids to form a complex capable of crossing the cell membrane. However, cationic liposomes, especially those constituted by monovalent cationic lipids, are not able to condense DNA efficiently, resulting in complexes with very heterogeneous size distribution. A strategy to improve complexation, adds a previous step of DNA condensation with cationic polypeptides. Such is the case of cationic liposome-entrapped polycation-condensed DNA type 1,[164] composed by DNA, DOTAP (1,2-dioleoyl-3-trimethylammonium propane) and protamine sulfate, which substitutes the poly-L-lysine (PLL) of previous formulations.[166] This protamine sulfate is smaller and highly positively charged than PLL, providing a high protection of DNA from nuclease degradation.

Gold nanoparticles (Au NPs) have also been developed for the efficient package and delivery of nucleic acids combined with low generation polypropylenimine dendrimers.[167] The limited surface charges of these low generation dendrimers lead to an inefficient complexation of DNA and low cellular uptake efficacy. Consequently, the support on Au NPs of low generation dendrimers, allowed the packaging of nucleic acids into discrete nanoparticles, but Au NPs are not encapsulated inside the final DNA/siRNA complexes

[162] Kazemi Oskuee, R. et al. *Life Sci.,* **2018**, *197*, 101–108.

[163] Wang, C. Y.; Huang, L. *Proc. Natl. Acad. Sci.,* **1987**, *84*, 7851–7855.

[164] Liu, F.; Huang, L. *J. Control. Release,* **2002**, *78*, 259–266.

[165] Felgner, P. L.; Gadek, T. R.; Holm, M.; Roman, R.; Chan, H. W.; Wenz, M.; Northrop, J. P.; Ringold, G. M.; Danielsen, M. *Proc. Natl. Acad. Sci.,* **1987**, *84*, 7413–7417.

[166] Gao, X.; Huang, L. *Biochemistry,* **1996**, *35*, 1027–1036.

[167] Chen, A. M.; Taratula, O.; Wei, D.; Yen, H. I.; Thomas, T.; Thomas, T. J.; Minko, T.; He, H. *ACS Nano,* **2010**, *4*, 3679–3688.

[161] Lungwitz, U.; Breunig, M.; Blunk, T.; Göpferich, A. *Eur. J. Pharm. Biopharm.,* **2005**, *60*, 247–266.

(Figure 18). Therefore, it becomes possible to eliminate the potential toxicity problems associated with Au NPs by selectively removing them from the resulting nucleic acid complexes before their delivery to target cells.

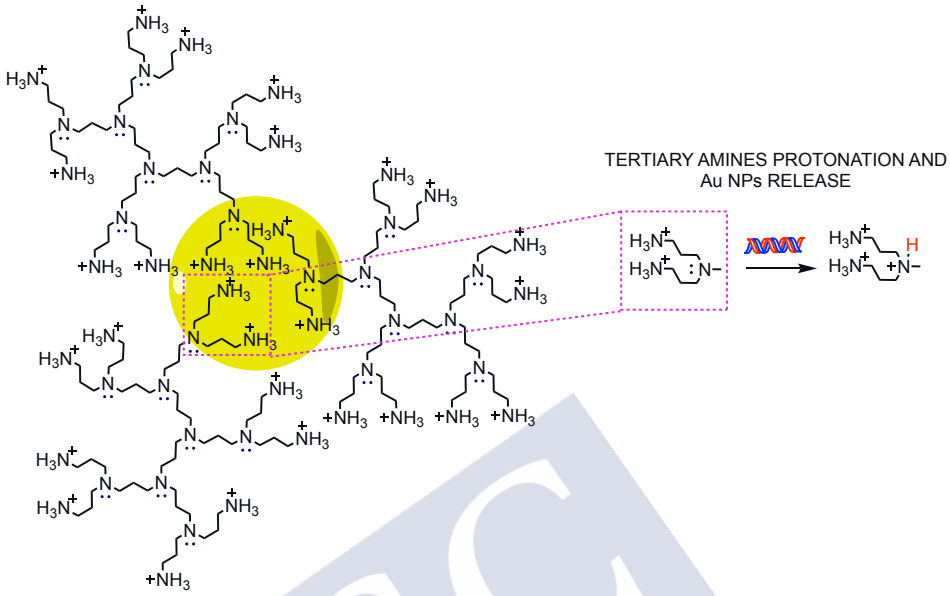


Figure 18. Au NPs (yellow ball) anchored with several low generation dendrimers through Au-amine coordination bonds. Addition of DNA/siRNA leads to an increase of local acidity protonating the tertiary amines and therefore, weakening the Au-amine interactions. As a result, the Au NPs are released from the dendrimers and are not included in the final DNA/siRNA nanoparticles.

1.5.2 Increasing the circulation time of nucleic acids

Both naked DNA and lipoplexes have shown rapid hepatic clearance during systemic administration. Absence of hydrophilic groups on DNA particles' surface, may lead to their interaction with plasma proteins and their removal from the circulation. However, this binding, when directed to specific plasma proteins, can improve pharmacokinetics, as large protein cargo complexes can avoid renal clearance, and their half-life in the circulation can be extended. Human Serum Albumin (HSA) is a natural transport protein, involved in trafficking of a wide variety of molecules, due to its multiple ligand binding sites and extended half-life.[168] The non-covalent binding of lipid-modified oligonucleotides to albumin, which increases the bioavailability of siRNAs has been recently published.[168] For instance, the incorporation of cholesterol modifications into siRNA, facilitates specific binding with recombinant HSA with an affinity dependent on the number of modifications. The aim of this study was to use this finding, along with the intrinsic transport properties of albumin, to tune siRNA serum half-life, hepatic accumulation, and gene silencing. Moreover, antibodies combine selective targeting with improved bioavailability. The design of a protamine-

[168] Bienk, K.; Hvam, M. L.; Pakula, M. M.; Dagnæs-Hansen, F.; Wengel, J.; Malle, B. M.; Kragh-Hansen, U.; Cameron, J.; Bukrinski, J. T.; Howard, K. A. *J. Control. Release,* **2016**, *232*, 143–151.

antibody fused protein to deliver siRNA to HIV-infected cells has also been reported.[169] The fusion protein (F105-P) was designed with the protamine coding sequence linked to the C-terminus of the heavy chain Fab fragment of an HIV-1 envelope antibody. siRNAs were bound to the resulting fusion protein by using electrostatic interactions. The final conjugate allowed the effective delivery, and induced silencing only in cells expressing the HIV-1 envelope glycoprotein.[169]

Furthermore, modification of non-viral vectors with hydrophilic molecules of PEG also decreases interaction with plasma proteins and increases transfection efficiency.[156] Examples in literature have demonstrated that PEG grafted PLL (PEG-g-PLL) can be used to enhance transfection efficiency by formulation with the fusogenic peptide KALA.[170] The ionic binding of the positively charged KALA with the negatively charged complexes of DNA/PEG-g-PLL, resulted in improved transfection efficiency and very low cytotoxicity (Figure 19).

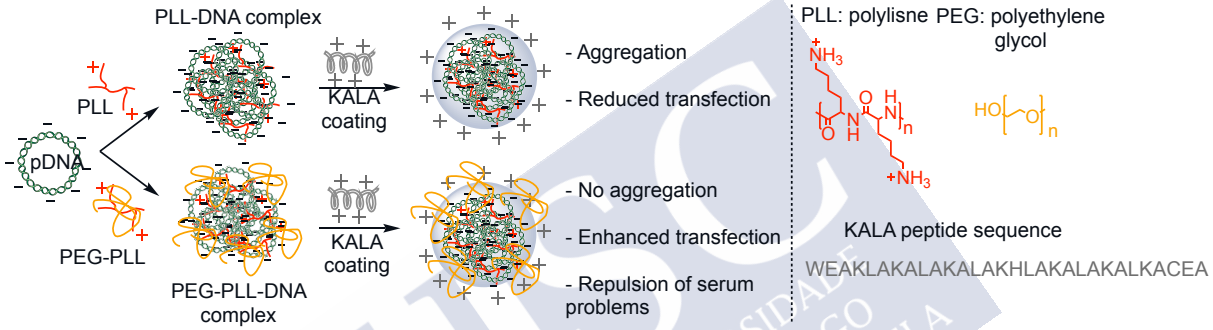


Figure 19. Schematic representation of the formulation of DNA/PLL and DNA/PEG-g-PLL complexes with KALA peptide. (PLL, PEG and KALA structures are represented in the right side).

1.5.3 Targeting approaches

The efficient and specific delivery of therapeutic genes to target cells is a challenge that will need to be overcome in order to take advantage of the promise and potential of genetic drugs avoiding non-specific adsorption and off-target effects.

The development of aptamer-siRNA chimeric RNAs capable of binding and delivering functional siRNAs into specific cell types has been reported.[171] One aptamer portion of a chimera protein mediates binding to PSMA, a cell-surface receptor overexpressed in prostate cancer cells and tumor vascular endothelium, while the siRNA portion targets the expression of survival genes. Therefore, these chimeras do not bind to, or function in cells that do not express PSMA, an approach that presents numerous applications including cancer

[169] Song, E.; Zhu, P.; Lee, S. K.; Chowdhury, D.; Kussman, S.; Dykxhoorn, D. M.; Feng, Y.; Palliser, D.; Weiner, D. B.; Shankar, P.; Marasco, W. A.; Lieberman, J. *Nat. Biotechnol.,* **2005**, *23*, 709–717.

[170] Lee, H.; Jeong, J. H.; Park, T. G. *J. Control. Release,* **2002**, *79*, 283–291.

[171] McNamara, J. O.; Andrechek, E. R.; Wang, Y.; Viles, K. D.; Rempel, R. E.; Gilboa, E.; Sullenger, B. A.; Giangrande, P. H. *Nat. Biotechnol.,* **2006**, *24*, 1005–1015.

[156] Read, M. L.; Logan, A.; Seymour, L. W. *Adv. Genet.,* **2005**, *53*, 19–46.

therapeutics.[171] Additionally, the programmable sequential recognition of a siRNA-loaded DNA nanovehicle by two aptamers capable of binding to two different cell receptors, has also been used to improve targeting in gene delivery.[172] The siRNA is self-assembled in an oligonucleotide nanovehicle modified with a hairpin structure to act as both the "smart key" and the delivery carrier. The auto-cleavable hairpin structure can be activated on site at the target cell membrane, by reacting sequentially with two aptamers as "dual locks", which leads to cell-subtype discrimination and precise siRNA delivery for high efficient gene silencing.[162] Furthermore, aptamers conjugated to lipids can also be assembled into lipid nanoparticles and used for targeted delivery of siRNA to different tissues, such as delivery to bone tissue to improve osteogenesis.[173] Thus, nucleic acid aptamers offer significant potential as convenient and evolvable targeting groups for drug delivery. The attachment of aptamers to the surface of a genome-free viral capsid carrier by an efficient oxidative coupling strategy, has been developed.[174] The method involves the periodate-mediated reaction of phenylenediamine-substituted oligonucleotides with aniline groups installed on the outer surface of the capsid shells (Figure 20). Up to 60 strands of DNA can be linked to each viral capsid without apparent loss of base pairing capabilities or protein stability.

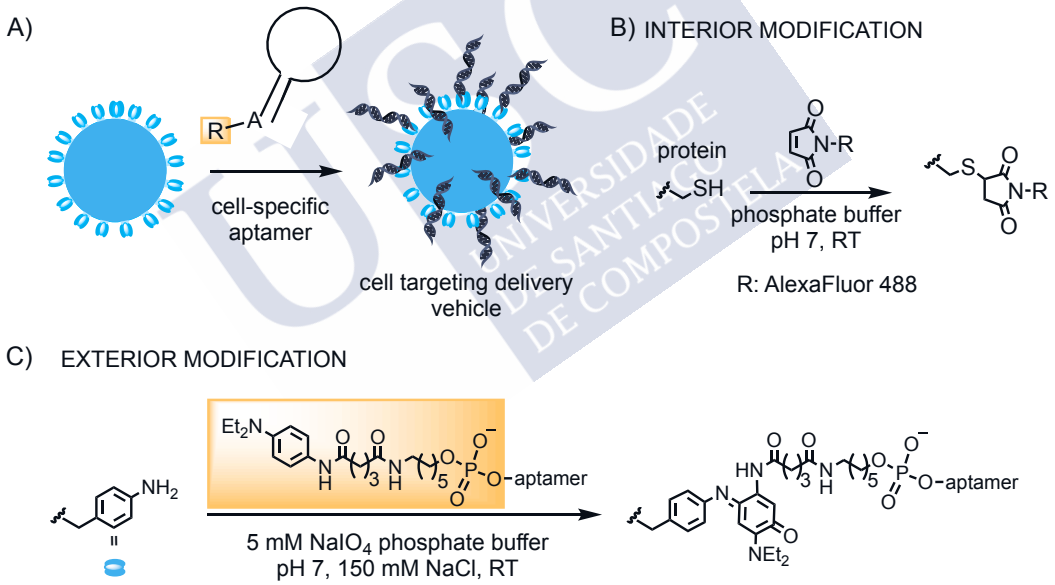


Figure 20. A) Surface modification of capsids for targeted delivery using aptamers. B) For interior surface modification, a mutation on the capsid coated with proteins allows site-specific alkylation. Up to 180 cargo molecules can be installed in these locations. C) For exterior modification, the aptamer is modified with phenylene diamine group by a periodate-mediated reaction.

[172] Ren, K.; Liu, Y.; Wu, J.; Zhang, Y.; Zhu, J.; Yang, M.; Ju, H. *Nat. Commun.*, **2016**, *7*, 1–10.

[173] Liang, C.; Guo, B.; Wu, H.; Shao, N.; Li, D.; Liu, J.; Dang, L.; Wang, C.; Li, H.; Li, S.; Lau, W. K.; Cao, Y.; Yang, Z.; Lu, C.; He, X.; Au, D. W. T.; Pan, X.; Zhang, B.-T.; Lu, C.; Zhang, H.; Yue, K.; Qian, A.; Shang, P.; Xu, J.; Xiao, L.; Bian, Z.; Tan, W.; Liang, Z.; He, F.; Zhang, L.; Lu, A.; Zhang, G. *Nat. Med.* **2015**, *21*, 288–294.

[174] Tong, G. J.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *J. Am. Chem. Soc.*, **2009**, *131*, 11174–11178.

[171] McNamara, J. O.; Andrechek, E. R.; Wang, Y.; Viles, K. D.; Rempel, R. E.; Gilboa, E.; Sullenger, B. A.; Giangrande, P. H. *Nat. Biotechnol.*, **2006**, *24*, 1005–1015.

In addition to aptamers, oligonucleotides can be fused to other ligands to trigger specific receptor-mediated endocytosis. For example, the conjugation of ASOs or siRNAs to triantennary-*N*-acetylgalactosamine is a widely used strategy for targeted delivery to the liver, through interactions with the asialoglycoprotein receptor 1 (ASGPR1).[175,176]

Another targeting strategy is the use of arginine-glycine-aspartic acid peptide (RGD) which specifically binds to $\alpha_v\beta_3$ integrins, an attractive target for gene delivery that is overexpressed on various cancer cells and tumoral endothelial cell's surface.[177] Tissue-targeting of siRNA was achieved by attaching a block copolymer shell of poly-L-lysine-PEG finished by cyclic RGD peptides to gold nanoparticles *via* a thiol-gold connection (Figure 21).[178] This research demonstrated that the bottom-up construction of nanocarriers using monodispersed building blocks could be employed as delivery platforms for RNA interference-based cancer therapy. In another study, clustered RGD ligands were introduced in PEI/pDNA polyplexes, by attachment of RGD peptide-modified gold nanoparticles to the polyplex surface.[179] The PEI/pDNA/RGD nanoclusters increased efficiency of gene transfer in HeLa cells compared to unmodified PEI/pDNA polyplexes.

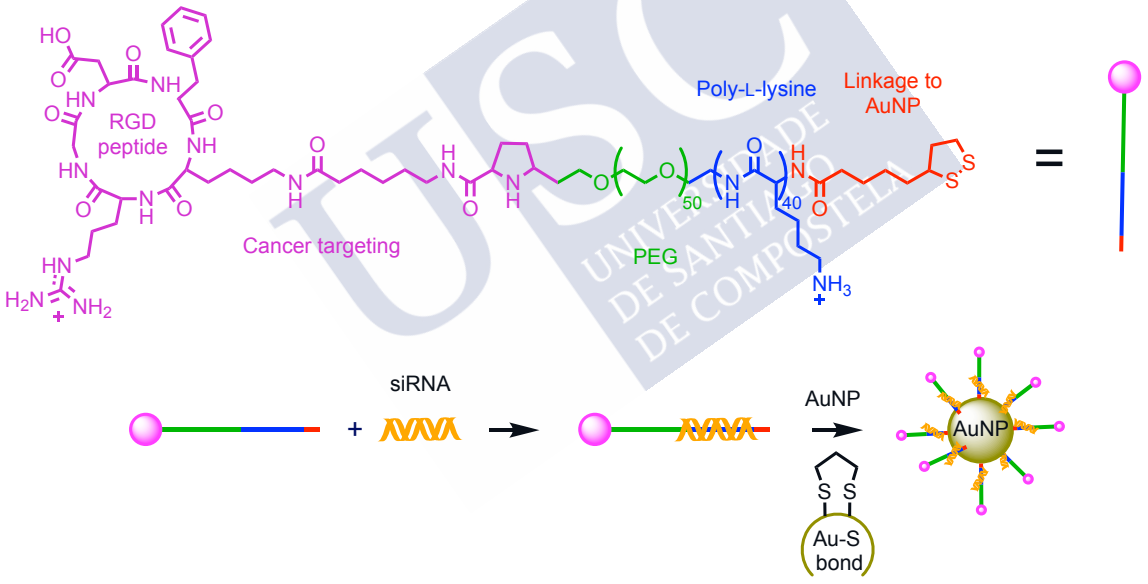


Figure 21. Delivery systems using gold nanoparticles. A gold nanoparticle can be modified with a targeting motif (cyclic Arg–Gly–Asp (cRGD), PEG and poly-L-lysine) for siRNA binding.

[175] Schmidt, K.; Prakash, T. P.; Donner, A. J.; Kinberger, G. A.; Gaus, H. J.; Low, A.; Østergaard, M. E.; Bell, M.; Swayze, E. E.; Seth, P. P. *Nucleic Acids Res.,* **2017**, *45*, 2294–2306.

[176] Zlatev, I.; Castoreno, A.; Brown, C. R.; Qin, J.; Waldron, S.; Schlegel, M. K.; Degaonkar, R.; Shulga-Morskaya, S.; Xu, H.; Gupta, S.; Matsuda, S.; Akinc, A.; Rajeev, K. G.; Manoharan, M.; Maier, M. A.; Jadhav, V. *Nat. Biotechnol.,* **2018**, *36*, 509–511.

[177] Danhier, F.; Le Breton, A.; Préat, V. *Mol. Pharm.,* **2012**, *9*, 2961–2973.

[178] Yi, Y.; Kim, H. J.; Mi, P.; Zheng, M.; Takemoto, H.; Toh, K.; Kim, B. S.; Hayashi, K.; Naito, M.; Matsumoto, Y.; Miyata, K.; Kataoka, K. *J. Control. Release,* **2016**, *244*, 247–256.

[179] Ng, Q. K. T.; Sutton, M. K.; Soonsawad, P.; Xing, L.; Cheng, H.; Segura, T. *Mol. Ther.,* **2009**, *17*, 828–836.

1.5.4 Improving plasma membrane crossing

Intracellular delivery of genetic material into the cytosol or the nucleus is essential for its therapeutic action, whereas cell membrane is typically impermeable to nucleic acids because of their large size and hydrophilic negative nature. Thus, several strategies have been developed to facilitate nucleic acid internalization.

Physical methods: as previously mentioned for protein delivery (see 1.1.4.1), physical methods such as electroporation or microinjection have also been employed in gene delivery. Electroporation has been adopted in laboratories for DNA transfection, where the electric potential across the cell membrane is increased, allowing charged molecules like DNA to be driven across the membrane pores in a similar way to electrophoresis.[180,181] Moreover, electroporation has been used to improve DNA vaccination against diseases such as influenza,[182] smallpox[183] and hepatitis B.[184] However, their clinical application is limited due to their invasive nature and potential damage to the cells membrane, as well as their poor access to deeper tissues.

Self-delivery: oligonucleotides can be self-delivered inside cells by the suppression of their negatively charged phosphate backbone (Figure 22). The synthesis of short interfering ribonucleic neutrals (siRNNs), whose phosphate backbone contains neutral phosphotriester groups, facilitating the delivery into cells, has been reported.[185] Once inside the cell, cytoplasmic thioestearases convert siRNNs into native, charged phosphodiester-backbone siRNAs, which induce a robust RNA interference response.

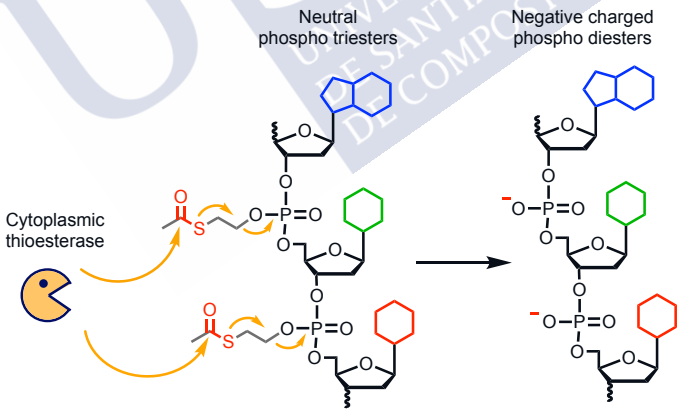


Figure 22. Phosphotriester cleavage by thioesterase gives charged phosphodiester linkage.

[180] Mehier-Humbert, S.; Guy, R. H. *Adv. Drug Deliv. Rev.,* **2005**, *57*, 733–753.

[181] Shigekawa, K.; Dower, W. J. *Biotechniques,* **1988**, *6*, 742–751.

[182] Chen, M.-W.; Cheng, T.-J. R.; Huang, Y.; Jan, J.-T.; Ma, S.-H.; Yu, A. L.; Wong, C.-H.; Ho, D. D. *Proc. Natl. Acad. Sci.,* **2008**, *105*, 13538–13543.

[183] Hooper, J. W.; Golden, J. W.; Ferro, A. M.; King, A. D. *Vaccine,* **2007**, *25*, 1814–1823.

[184] Luxembourg, A.; Hannaman, D.; Ellefsen, B.; Nakamura, G.; Bernard, R. *Vaccine,* **2006**, *24*, 4490–4493.

[185] Meade, B. R.; Gogoi, K.; Hamil, A. S.; Palm-Apergi, C.; Van Den Berg, A.; Hagopian, J. C.; Springer, A. D.; Eguchi, A.; Kacsinta, A. D.; Dowdy, C. F.; Presente, A.; Lönn, P.; Kaulich, M.; Yoshioka, N.; Gros, E.; Cui, X. S.; Dowdy, S. F. *Nat. Biotechnol.,* **2014**, *32*, 1256–1261.

Coiled-coil: another strategy for the intracellular delivery of nucleic acids, involves the design of chemical vectors capable of actively targeting internalization pathways that lead to an efficient transfection. As commented above (see 1.5.1), lipofection has been one of the most popular transfection methods since the discovery of the spontaneous condensation of cationic lipids with DNA and their fusion with cell membranes.[165] However, to improve the access to intracellular targets some modifications to the basic liposomal delivery system are needed.[186] Coiled coil interactions are strong protein-protein interactions that are involved in many biological processes, including intracellular trafficking and membrane fusion. Thus, artificial liposomes containing heterodimeric coiled-coil helical peptides, such as E_3 $(EIAALEK)_3$ and K_3 $(KIAALKE)_3$, were used for encapsulating siRNA, to then target a particular cell, which had been previously loaded with the corresponding counterpart for coiled-coil formation, achieving intracellular delivery (Figure 23A).[187] Furthermore, cell transfection by bioorthogonal chemistry has been achieved by engineering the cell surface to include ketone-containing lipids, which react with lipoplexes comprising alkoxyamine functionalities (Figure 23B).[188] However, the translation of these two methodologies for *in vivo* purposes will be challenging, as they require the selective introduction of coiled-coils or ketone-lipids into the targeted cell membrane.

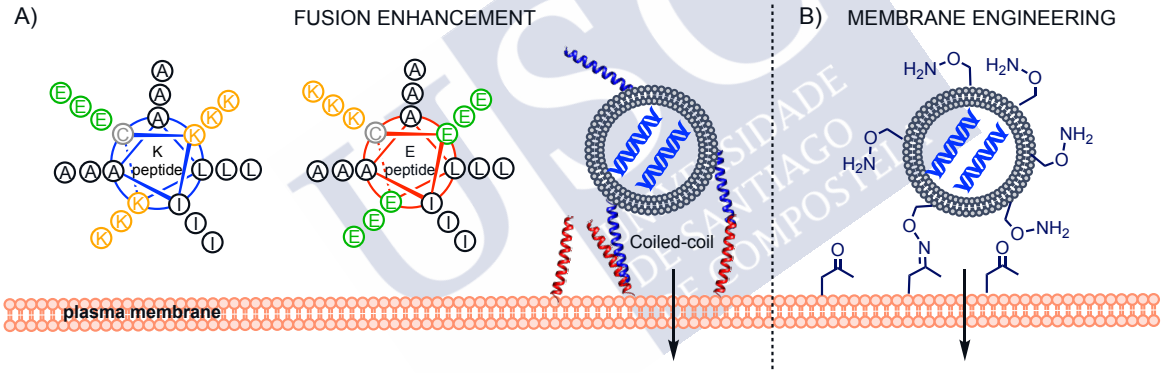


Figure 23. Enhancement of the fusion of liposomes to the cell membrane. A) Formation of pairs of coiled-coils pairs between peptide-modified liposomes and peptide-modified membrane (color code: hydrophobic in black (A,L,I); cationic in orange (K); anionic in green (E). B) Bioorthogonal reaction between alkoxyamine-containing liposomes and ketone-containing lipids in the cell membrane.

CPPs as covalent or non-covalent conjugates have been extensively used for the efficient delivery of nucleic acids and their synthetic oligonucleotide analogs, both *in vitro* and *in vivo*.[189] CPP-mediated intracellular delivery, relies on the ability of CPPs to cross cell membranes, for delivering their cargo into the cytoplasm and/or nucleus. Among CPPs, TAT

[186] Allen, T. M.; Cullis, P. R. *Adv. Drug Deliv. Rev.,* **2013**, *65*, 36–48.

[187] Oude Blenke, E. E.; Van Den Dikkenberg, J.; Van Kolck, B.; Kros, A.; Mastrobattista, E. *Nanoscale,* **2016**, *8*, 8955–8965.

[188] O'Brien, P. J.; Elahipanah, S.; Rogozhnikov, D.; Yousaf, M. N. *ACS Cent. Sci.,* **2017**, *3*, 489–500.

[189] Lehto, T.; Ezzat, K.; Wood, M. J. A.; EL Andaloussi, S. *Adv. Drug Deliv. Rev.,* **2016**, *106*, 172–182.

[165] Felgner, P. L.; Gadek, T. R.; Holm, M.; Roman, R.; Chan, H. W.; Wenz, M.; Northrop, J. P.; Ringold, G. M.; Danielsen, M. *Proc. Natl. Acad. Sci.,* **1987**, *84*, 7413–7417.

peptide is commonly used for the covalent attachment to nucleic acids by chemical coupling, genetic fusion or electrostatic interactions.[190] Moreover, TAT peptide can be cross-linked to the distal end of PEG-polymers or PEG-lipids, resulting in conjugates that form a complex with DNA, or conjugates that self-assemble with lipids or polymers with the TAT moiety displayed on the surface of the drug nanocarrier, respectively.[191] However, TAT peptide either alone or combined with polymers or lipids usually enters the cells *via* endocytosis. Other cationic amphiphilic CPPs, such as CLIP6, which presents a D-proline in the central part of its sequence, which together with a glutamic acid, disrupts the potential β-sheet secondary structure in the peptide.[192] This intrinsically disordered peptide can reach the cytosol by a non-endocytic pathway for the efficient delivery of, for example, peptide nucleic acids.[193] Another example of amphiphilic CPP is the peptide RALA, a GALA variant with 30 extra amino acids, where the protonation of glutamate residues at low pH prevents the anionic repulsion between peptide side chains and, hence, triggers helical folding. This secondary structure fluctuation enables membrane disruption at the low pH conditions of the endosome, allowing the internalization of pDNA[194]or mRNA.[195]

1.5.5 Endosomal escape and cargo release

Apart from all methodologies discussed above, there are other multiple strategies for the intracellular delivery of nucleic acids, which follow endocytic mechanisms. In this regard, the design of non-viral vectors with the ability of escaping from the endosome and releasing of the cargo is necessary.

Protonable peptides such as GALA, have been used for the formulation of complexes composed by cationic liposomes and transferrin, which could have applications in the delivery of pDNA.[196] The presence of transferrin is proposed to trigger internalization of the complex by receptor-mediated endocytosis, while the presence of GALA peptide, is used to promote endosomal destabilization and release of the genetic material into the cytoplasm. At the endosomal pH (pH < 6) GALA peptide adopts an α-helical secondary structure being able to be inserted into membranes causing leakage of endocytic vesicles (Figure 24).[197]

---

[190] Koren, E.; Apte, A.; Sawant, R. R.; Grunwald, J.; Torchilin, V. P. *Drug Deliv.,* **2011**, *18*, 377–384.

[191] Torchilin, V. P. *Biopolymers* **2008**, *90*, 604–610.

[192] Medina, S. H.; Miller, S. E.; Keim, A. I.; Gorka, A. P.; Schnermann, M. J.; Schneider, J. P. *Angew. Chemie - Int. Ed.,* **2016**, *55*, 3369–3372.

[193] Soudah, T.; Mogilevsky, M.; Karni, R.; Yavin, E. *Bioconjug. Chem.,* **2017**, *28*, 3036–3042.

[194] Mccarthy, H. O.; McCaffrey, J.; Mccrudden, C. M.; Zholobenko, A.; Ali, A. A.; McBride, J. W.; Massey, A. S.; Pentlavalli, S.; Chen, K. H.; Cole, G.; Loughran, S. P.; Dunne, N. J.; Donnelly, R. F.; Kett, V. L.; Robson, T. *J. Control. Release,* **2014**, *189*, 141–149.

[195] Udhayakumar, V. K.; De Beuckelaer, A.; McCaffrey, J.; McCrudden, C. M.; Kirschman, J. L.; Vanover, D.; Van Hoecke, L.; Roose, K.; Deswarte, K.; De Geest, B. G.; Lienenklaus, S.; Santangelo, P. J.; Grooten, J.; McCarthy, H. O.; De Koker, S. *Adv. Healthc. Mater.,* **2017**, *6*, 1601412.

[196] Kakudo, T.; Chaki, S.; Futaki, S.; Nakase, I.; Akaji, K.; Kawakami, T.; Maruyama, K.; Kamiya, H.; Harashima, H. *Biochemistry,* **2004**, *43*, 5618–5628.

[197] Schach, D. K.; Rock, W.; Franz, J.; Bonn, M.; Parekh, S. H.; Weidner, T. *J. Am. Chem. Soc.,* **2015**, *137*, 12199–12202.

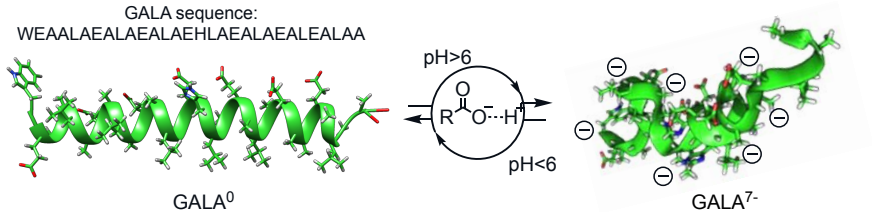


Figure 24. Reversible folded and unfolded states of GALA due to glutamic acid protonation.

Polymers can also be engineered for membrane recognition, interaction and translocation. Research efforts in cell-penetrating polymeric materials are now focused on improving condensation of the genetic material, reducing toxicity and increasing endosomal escape. For all these reasons, new polymeric particles have been designed to be loaded with different cargos and disassembled in response to external stimuli such as enzymes, temperature, etc. but more important to pH. The use of poly(2-(diisopropylamino)ethyl methacrylate) (PDPA) enables micelle disruption and siRNA release in the tumor microenvironment as it undergoes cooperative pH-triggered disassembly.[198] An application of this concept can be seen in ultra-pH-sensitive nanoplatforms, in which pH-responsive of PDPA polymers grafted with cationic lipid-like structures, are combined with RGD peptide and PEG.[199] The resulting nanoparticles showed disassembly at the pH of early endosomes demonstrating their applicability for *in vivo* targeted delivery of siRNA to tumors (Figure 25).

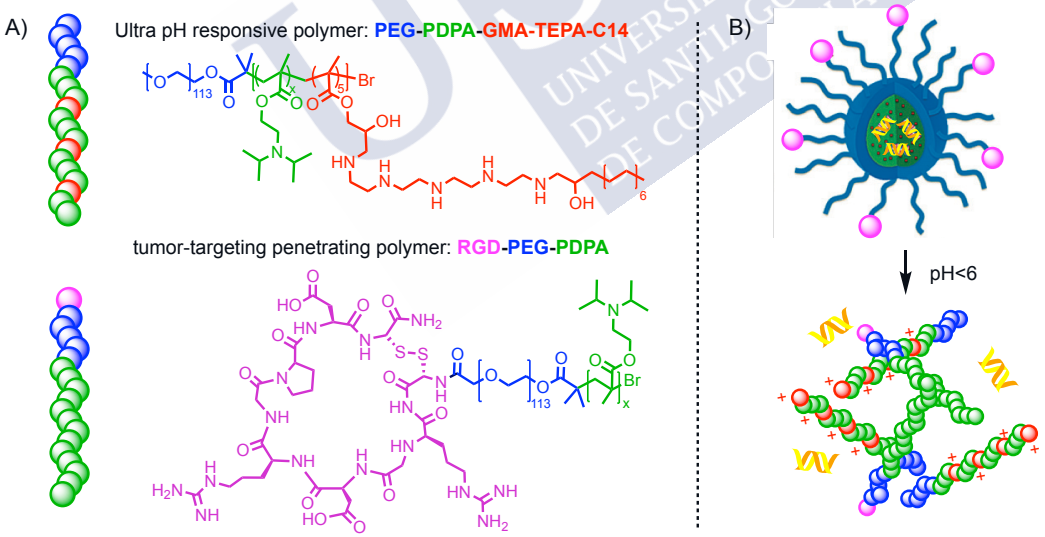


Figure 25. A) Structures and schematic representation of siRNA-encapsulating polymers. B) siRNA loading and release. (GMA: glycidyl methacrylate; TEPA: tetraethylenepentamine).

[198] Wang, Y.; Zhou, K.; Huang, G.; Hensley, C.; Huang, X.; Ma, X.; Zhao, T.; Sumer, B. D.; Deberardinis, R. J.; Gao, J. *Nat. Mater.,* **2014**, *13*, 204–212.

[199] Xu, X.; Wu, J.; Liu, Y.; Yu, M.; Zhao, L.; Zhu, X.; Bhasin, S.; Li, Q.; Ha, E.; Shi, J.; Farokhzad, O. C. *Angew. Chemie - Int. Ed.,* **2016**, *55*, 7091–7094.

Recently, the synthesis of charge-altering releasable transporters (CARTs) for mRNA delivery into cells and animals has been reported.[200] CARTs are constituted by oligo-(α-amino ester) cations, which complex and deliver mRNA, and then degrade by a charge-neutralizing intramolecular rearrangement, where cationic amines are converted to neutral amides. This controlled degradation results in mRNA release into the cytosol for translation (Figure 26).

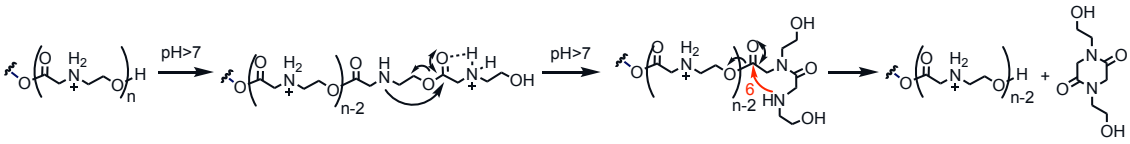


Figure 26. CARTs: self-immolated polymers that degrade at pH>7. Rearrangement mechanism for oligo(α-amino ester)s through tandem five-membered then six-membered transition states.

Supramolecular chemistry, in particular supramolecular templates, is a powerful tool to understand and develop new non-viral vectors for gene therapy. Cyclodextrin-containing cationic polymers (CDP) can complex siRNA and include targeting and stabilizing moieties by supramolecular host-guest chemistry using adamantane/cyclodextrin conjugates (Figure 27A).[201] The CDP interacts with the nucleic acids by electrostatic interactions, which contains imidazole groups that are protonated at pH~6. This "chemical sensing" mechanism activates many processes providing endosomal escape and nanoparticle release of the nucleic acid into the cytoplasm. More recently, a polyadamantane-ketal, an adamantane grafted biodegradable polymer for nucleic acid delivery, which degrades into low molecular weight excretable compounds in the acidic environment of the endosome, and that can be self-assembled with PEI-grafted β-CD to efficiently deliver DNA to cells, has been reported (Figure 27B).[202]

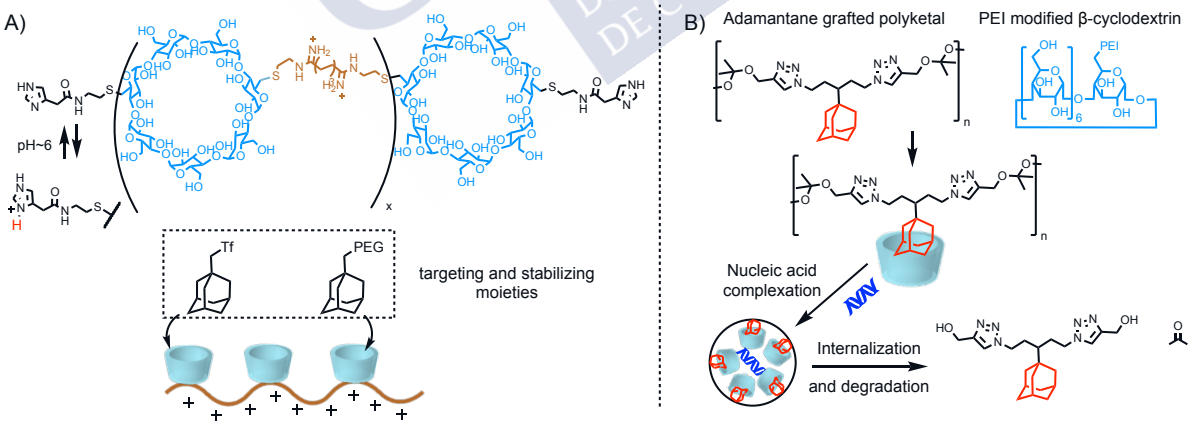


Figure 27. A) Cyclodextrins fused into cationic polymers using adamantane host-gest chemistry. B) Cyclodextrin molecules decorated with shorter cationic polymers for nucleic acid interaction and assembled into larger structures using polyadamantanes.

[200] McKinlay, C. J.; Vargas, J. R.; Blake, T. R.; Hardy, J. W.; Kanada, M.; Contag, C. H.; Wender, P. A.; Waymouth, R. M. *Proc. Natl. Acad. Sci.,* **2017**, *114*, E448–E456.

[201] Davis,Mark E. *Mol. Pharm.,* **2009**, *6*, 659–668.

[202] Maity, S.; Choudhary, P.; Manjunath, M.; Kulkarni, A.; Murthy, N. *Chem. Commun.,* **2015**, *51*, 15956–15959.

Several other strategies have been developed to trigger nucleic acid release using polymeric formulations. Disulfide functionalization of polymers can also be exploited for gene delivery applications. In this regard, the high concentrations of glutathione (1-10 mM) inside the eukaryotic cell cytosol generate a reductive environment, which can be exploited to trigger a response and reducing the disulfide bonds into their individual components.[203] The design and synthesis of a dipicolyamine-based disulfide-containing zinc (II) coordinative module, used for low-molecular-weight (~1800 Da) PEI (Zn-PD) functionalization has been reported.[204] This non-viral vector has high affinity for DNA and can be cleaved by glutathione in the cytoplasm, facilitating DNA release post internalization in primary and stem cells, and diminishing the cytotoxicity (Figure 28).

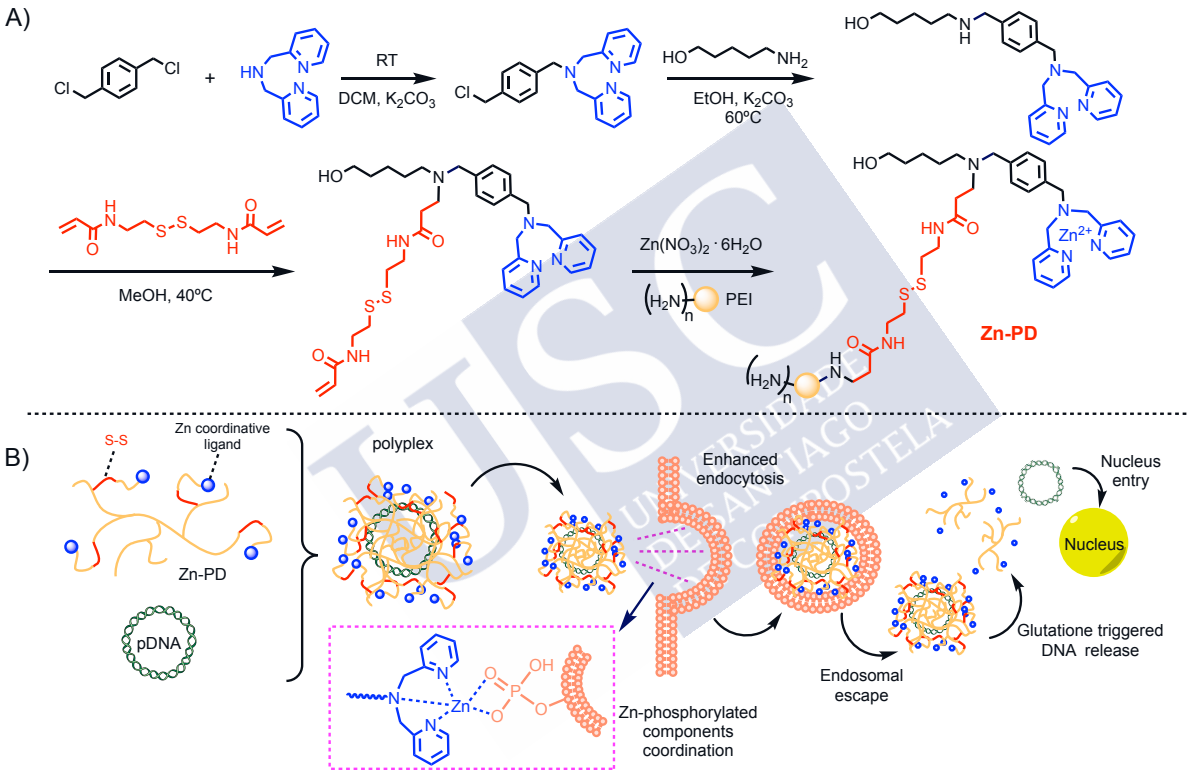


Figure 28. A) Synthesis of Zn-PD non-viral vectors and B) schematic representation of DNA transfection.

In a different approach, phenylboronic acid (PBA) functionalized polyion complexes (PIC) micelles have been used for siRNA internalization and release (Figure 29).[205] PBA is a synthetic molecule capable of forming reversible covalent esters with 1,2- or 1,3-*cis*-diols included on a ribose ring, a structure present at the 3' end of RNAs and several types of

[203] Méndez-Ardoy, A.; Lostalé-Seijo, I.; Montenegro, J. *ChemBioChem,* **2019**, *20*, 488–498.

[204] Liu, S.; Zhou, D.; Yang, J.; Zhou, H.; Chen, J.; Guo, T. *J. Am. Chem. Soc.,* **2017**, *139*, 5102–5109

[205] Naito, M.; Ishii, T.; Matsumoto, A.; Miyata, K.; Miyahara, Y.; Kataoka, K. *Angew. Chemie - Int. Ed.,* **2012**, *51*, 10751–10755.

ribonucleotides. Consequently, this binding property offers an easy way for chemical conjugation of siRNA to the pendant PBA groups. After internalization, the high cytosolic levels of ATP can disrupt these boronate-stabilized particles, resulting in siRNA release (Figure 29). In a similar approach, this PBA-functionalized polyion strategy has been used for pDNA internalization by combining a PBA decorated polymer with a polyol decorated polymer, which will entrap pDNA until reaching the cytosol.[206] Furthermore, instead of a stimulus, timed degradation can be used to protect the cargo until cytosolic delivery, as in the case of a cationic polymer, PDMAEA, that self-degrades into a negatively charged and non-toxic polymer, which repels its cargo facilitating siRNA release.[207]

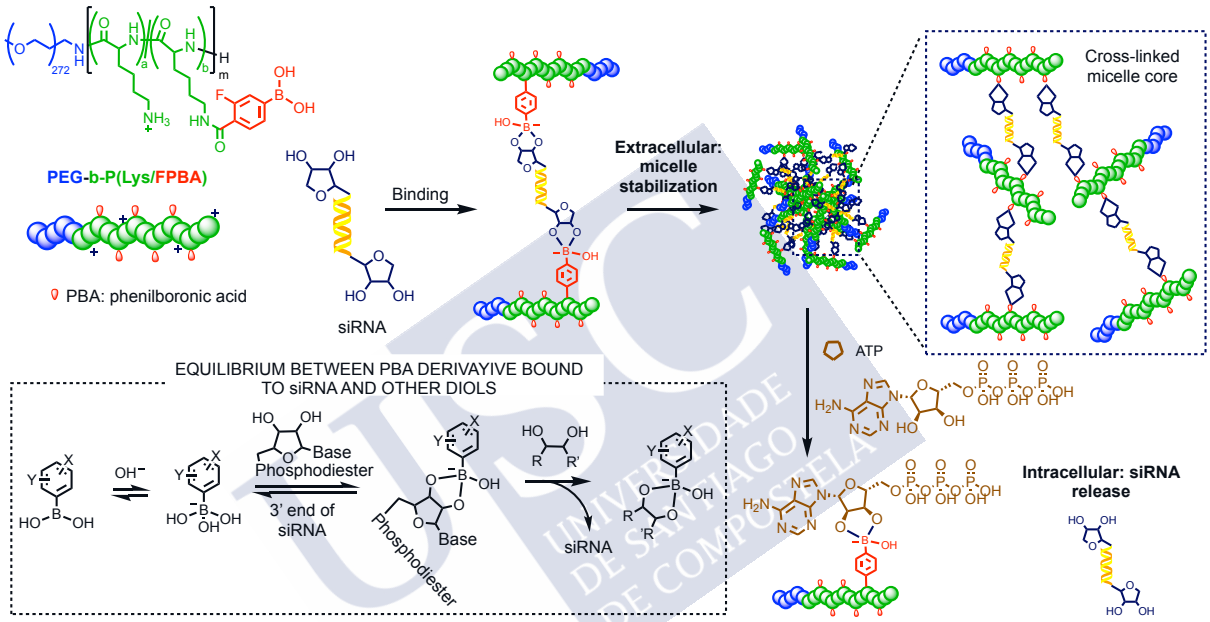


Figure 29. Schematic representation of the PBA based strategy for siRNA delivery: chemical structure of the polymer, stability of the micelle, and mechanism of selective intracellular release.

Polymers and CPPs containing dynamic bonds such as oximes and hydrazones, are promising non-viral vectors for gene delivery, as they can be formed in mild aqueous conditions with good yields and short reaction times, while being fully bioorthogonal.[147] The formation of these kind of bonds is reversible leading to the endosomal escape and release of the cargo, as long as they can be hydrolyzed at the endosomal pH.[208] The screening of cationic dendronised amphiphiles, combined with different hydrophobic aldehydes, allowed the fast identification of new and simple formulations for the delivery of siRNAs and

[206] Yoshinaga, N.; Ishii, T.; Naito, M.; Endo, T.; Uchida, S.; Cabral, H.; Osada, K.; Kataoka, K. *J. Am. Chem. Soc.,* **2017**, *139*, 18567–18575.

[207] Truong, N. P.; Gu, W.; Prasadam, I.; Jia, Z.; Crawford, R.; Xiao, Y.; Monteiro, M. J. *Nat. Commun.,* **2013**, *4*, 1902–1907.

[208] Gasparini, G.; Bang, E. K.; Montenegro, J.; Matile, S. *Chem. Commun.,* **2015**, *51*, 10389–10402.

[147] Lostalé-Seijo, I.; Montenegro, J. *Nat. Rev. Chem.,* **2018**, *2*, 258–277.

plasmids.[209,6] Following up on this strategy, our research group has reported hydrazone-modulated polymers as promising candidates for nucleic acid delivery. In this approach, polyhydrazides have been modified with different combinations of cationic and hydrophobic aldehydes to generate a library of amphiphilic polymers to complex and deliver siRNA and pDNA into living cells.[7,8]

1.5.6 Nuclear import

The nuclear envelope that separates the cell's genetic material from the surrounding cytoplasm, represents an extra physical barrier for nuclear importation of macromolecules such as pDNA. In 1980, the direct microinjection of pDNA into the cultured mammalian cells achieved a 50% of gene expression when injecting at the nucleus as compared with the inexistent expression when injecting the same amount of pDNA into the cytoplasm.[210] However, this methodology is inefficient when the number of cells for transfection is large.

The transport of therapeutic DNA from the cytoplasm into the nucleus is an inefficient process that is considered as the major limiting step in non-dividing cells. One of the strategies to improve nuclear uptake of DNA is taking advantage of the cellular nuclear import machinery. Synthetic peptides containing a nuclear localization signal (NLS, PKKKRKV)[211] are bound to the DNA, resulting in a DNA-NLS complex that can be recognized as a nuclear import substrate by specific intracellular receptor proteins, such as importins and facilitate transfer through the nuclear pore.[212,213] A strategy for the effective nucleic acid delivery into T cells, where gene carriers must be taken on by T cells and import their DNA cargo into the nucleus, has been reported.[214] The methodology consists on using biodegradable poly(β-amino ester)-based nanoparticles (PBAE) containing peptides with microtubule-associated sequences and nuclear localization signals, as a means to facilitate fast track nuclear import of the genetic cargo *via* the microtubule transport machinery. In an extraordinarily promising approach, a pDNA packed with this PBAE polymer was coated with poly(glutamic) acid attached to a targeting antibody and, the resulting polyplex allowed the *in vivo* reprogramming of T cells to express chimeric antigen receptors for treating leukemia in a mouse model (Figure 30).[214]

---

[209] Gehin, C.; Montenegro, J.; Bang, E. K.; Cajaraville, A.; Takayama, S.; Hirose, H.; Futaki, S.; Matile, S.; Riezman, H. *J. Am. Chem. Soc,*. **2013**, *135*, 9295–9298.

[210] Capecchi, M. R. *Cell* **1980**, *22*, 479–488.

[211] Kalderon, D.; Roberts, B. L.; Richardson, W. D.; Smith, A. E. *Cell,* **1984**, *39*, 499–509.

[212] Cartier, R.; Reszka, R. *Gene Ther.,* **2002**, *9*, 157–167.

[213] Bremner, K. H.; Seymour, L. W.; Logan, A.; Read, M. L. *Bioconjug. Chem.,* **2004**, *15*, 152–161.

[214] Smith, T. T.; Stephan, S. B.; Moffett, H. F.; McKnight, L. E.; Ji, W.; Reiman, D.; Bonagofski, E.; Wohlfahrt, M. E.; Pillai, S. P. S.; Stephan, M. T. *Nat. Nanotechnol.,* **2017**, *12*, 813–820.

[6] Louzao, I.; García-Fandiño, R.; Montenegro, J. *J. Mater. Chem. B,* **2017**, *5*, 4426–4434.

[7] Priegue, J. M.; Crisan, D. N.; Martínez-Costas, J.; Granja, J. R.; Fernandez-Trillo, F.; Montenegro, J. *Angew. Chemie Int. Ed.,* **2016**, *55*, 7492–7495.

[8] Priegue, J. M.; Lostalé-Seijo, I.; Crisan, D.; Granja, J. R.; Fernández-Trillo, F.; Montenegro, J. *Biomacromolecules,* **2018**, *19*, 2638–2649.

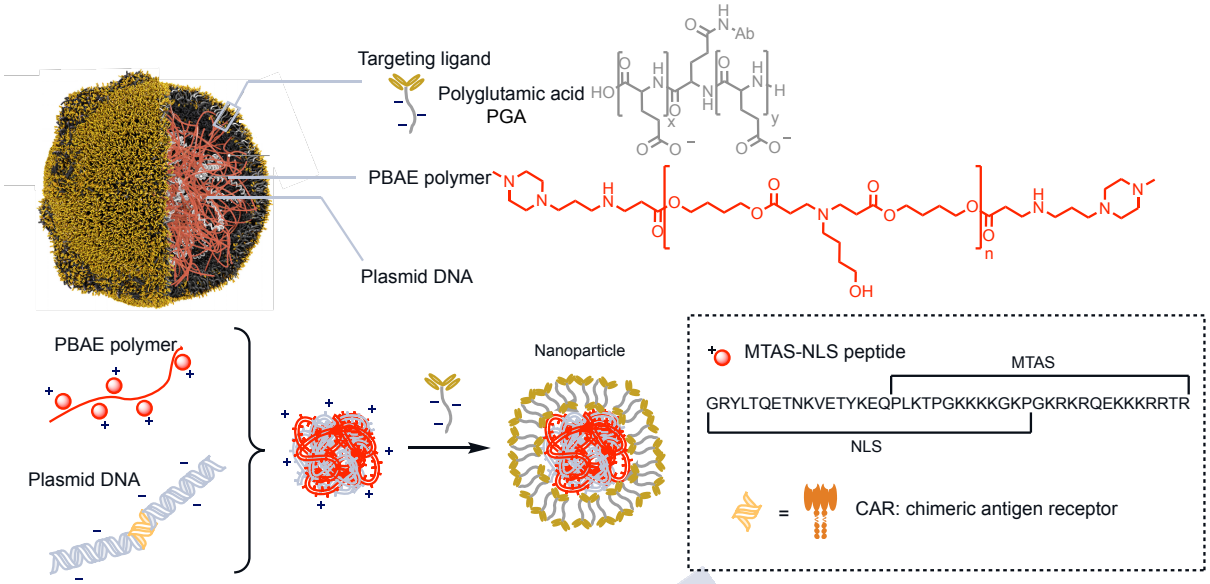


Figure 30. Design and manufacture of lymphocyte-programming nanoparticles. Diagram of the fabrication of the PBAE nanoparticles. (Adapted from *Nat. Nanotechnol.,* **2017**, *12*, 813–820).

Recently, a combination of electroporation and microfluidic cell deformation has been described for nuclear pDNA direct delivery.[215] This methodology consists on passing cells at high speed through microfluidic constrictions, smaller than the cell diameter, which mechanically disrupts the cell membrane, combined with subsequent electric field to further disrupt the nuclear envelope, allowing DNA molecules to be driven into the cytoplasm and nucleus (Figure 31).

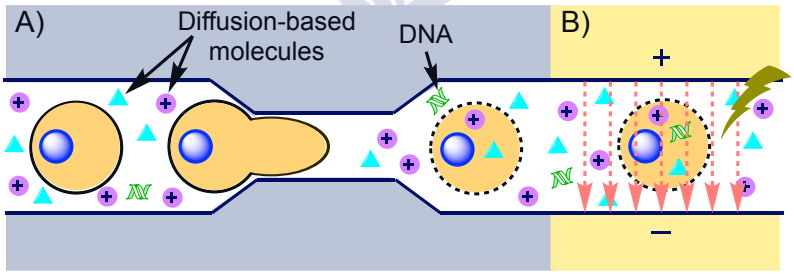


Figure 31. Schematic illustration of the methodology. A) Mechanical disruption of the cell membrane when the cell passes trough the constriction and B) the subsequent electric pulses driving DNA into the cytoplasm and nucleus through the disrupted membrane. (Doted arrows represent electric field).

[215] Ding, X.; Stewart, M. P.; Sharei, A.; Weaver, J. C.; Langer, R. S.; Jensen, K. F. *Nat. Biomed. Eng.,* **2017**, *1*, 1–7.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2 *MESSENGER RNA DELIVERY BY HYDRAZONE-ACTIVATED POLYMERS*

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2.1 Precedents and Objectives

As commented above, polymers are very promising materials for cell delivery and release of macromolecules of interest. Our research group has previously reported the synthesis of polyhydrazides functionalized with hydrophobic and cationic aldehydes for the delivery of siRNA and pDNA.[7,8] The general objective of this chapter is to further expand the delivery properties of polyhydrazone polymers by tackling the challenging delivery of mRNA.

A high molecular weight polyhidrazide polymer will be prepared using free-radical polymerization, which will be condensed with six different hydrophobic aldehydes and a cationic aldehyde (t-guanidinium), to obtain the corresponding amphiphilic polyhydrazones.

This polyhydrazones will be complexed with mRNA cargo, encoded for the synthesis of the enhanced green fluorescent protein (EGFP), for studying its delivery in Hek293 cells.

The formation of these polyhydrazones will be carried out in biocompatible conditions, allowing their combination with mRNA without needing any isolation or purification steps (Figure 32). Quantification of transfection will be performed by flow cytometry.

Transfection efficiency will be optimized by trying different hydrophobic aldehyde tails changing the molar fraction of cationic/hydrophobic aldehydes present in the polyhydrazone scaffold for determining which is the most promising ratio for mRNA delivery.

Control experiments will be carried out using the octaarginine penetrating peptide, the pore forming peptide GALA, the cationic lipid DOTAP and PEI polymer.

The MTT colorimetric assay will be used to evaluate the toxicity of the polyhydrazones at the transfecting concentrations.

The interaction between mRNA and the six different polyhydrazones will be evaluated by gel electrophoresis, DLS and Zeta potential measurements.

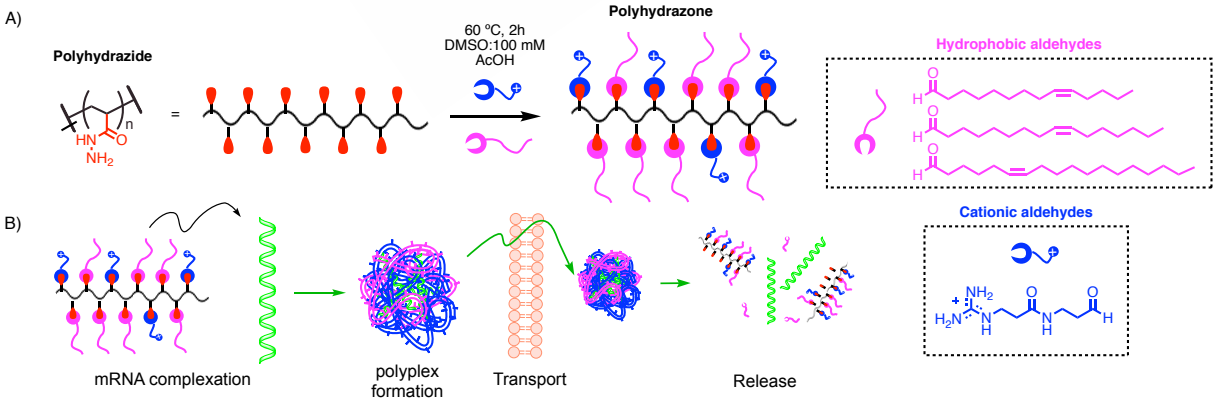


Figure 32. A) Polyhydrazone formation and B) delivery of mRNA for translation.

---

[7] Priegue, J. M.; Crisan, D. N.; Martínez-Costas, J.; Granja, J. R.; Fernandez-Trillo, F.; Montenegro, J. *Angew. Chemie Int. Ed.*, **2016**, *55*, 7492–7495.

[8] Priegue, J. M.; Lostalé-Seijo, I.; Crisan, D.; Granja, J. R.; Fernández-Trillo, F.; Montenegro, J. *Biomacromolecules,* **2018**, *19*, 2638–2649.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

https://pubs.rsc.org/en/content/articlelanding/2019/md/c9md00231f#!divAbstract

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2.3 Conclusions

In summary, in this chapter we have studied the potential of polyhydrazones to carry out the challenging release of mRNA inside human cells (Hek293).

This strategy provided new information about the importance of high molecular weight polymerized scaffolds as well as the presence of unsaturated hydrophobic aldehydes in their structure.

As opposed to previous work in the group with other nucleic acids (*i.e.* pDNA), no activity was observed with saturated aldehydes, such as dodecanal. Myristoleic aldehyde proved to be a promising candidate for mRNA condensation and internalization, reaching a 42% of transfection. This result could be also confirmed by visualizing transfected cells using fluorescence microscopy.

Optimization on the ratio of hydrophobic:cationic aldehydes, using the myristoleic derivative, confirmed that the ratio 50:50 was the most efficient for mRNA delivery and release inside cells.

This versatile methodology allowed a rapid screening for the fast identification of polyhydrazones with the correct choice of aldehydes, achieving similar values of transfection as PEI but much higher efficiency when compared with R8 and GALA peptides and the cationic lipid DOTAP, all of them used at the same weight concentration.

Cell viability was quantified by using colorimetric MTT assay, observing low toxicity with all polyhydrazones tested, confirming the promising properties of these carriers for mRNA delivery.

DLS measurements revealed complexation of the polyhydrazones and the mRNA, with polyplex diameters between 50 to 100 nm, with a progressive increase in the zeta potential value when the polymer concentration increases.

Gel electrophoresis assays confirmed these results showing the inhibition in the migration of mRNA when it is complexed with the polyhydrazone.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

SUPPORTING INFORMATION

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1 SUPRAMOLECULAR RECOGNITION AND SELECTIVE PROTEIN UPTAKE BY PEPTIDE HYBRIDS

*Marisa Juanes†, Irene Lostalé-Seijo†, Juan R. Granja† & Javier Montenegro†**

† Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain. *e-mail: javier.montenegro@usc.es

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

1.1 Supporting Figures

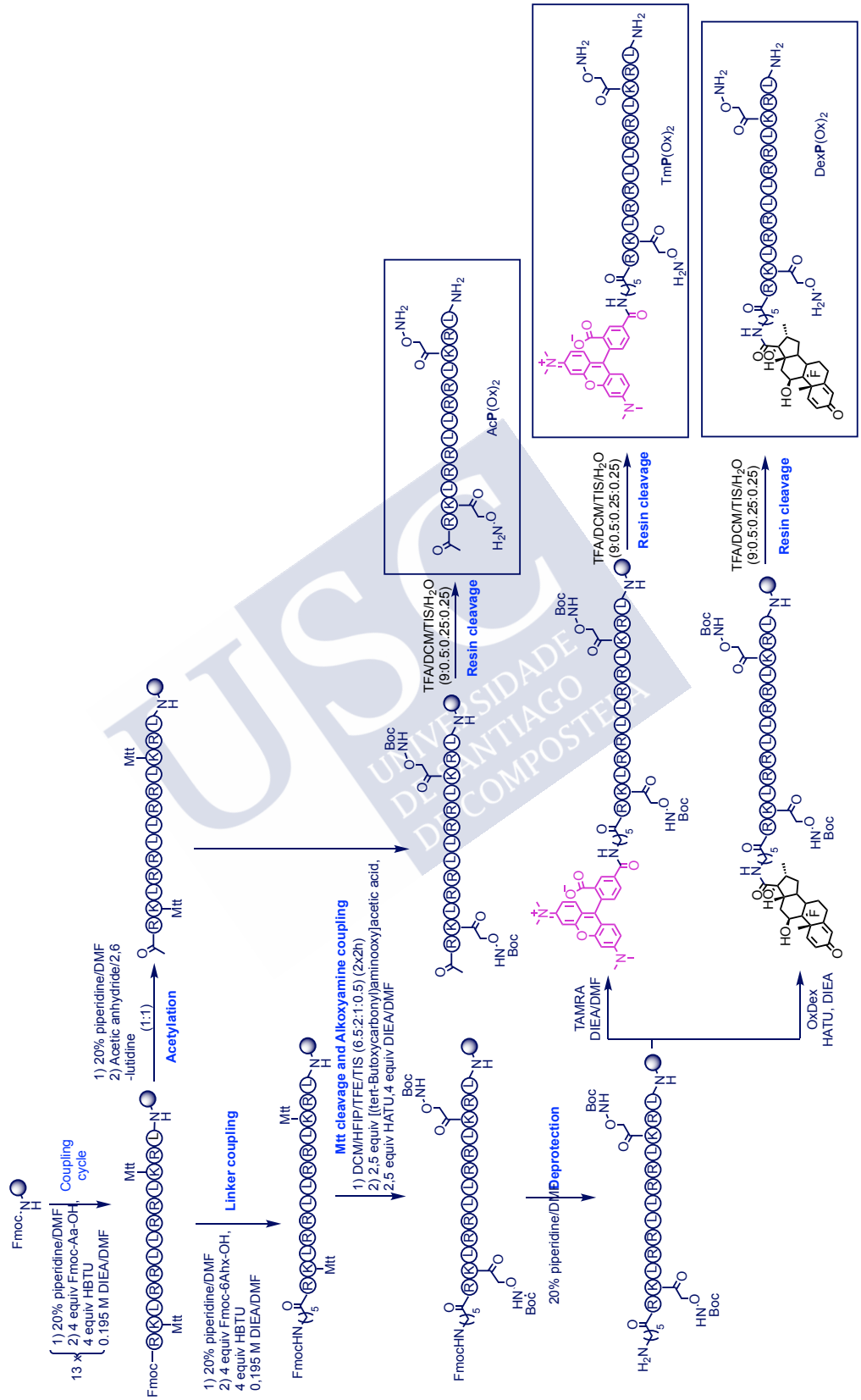


Figure S1.General synthetic scheme for the Solid Phase Peptide Synthesis (SPPS).

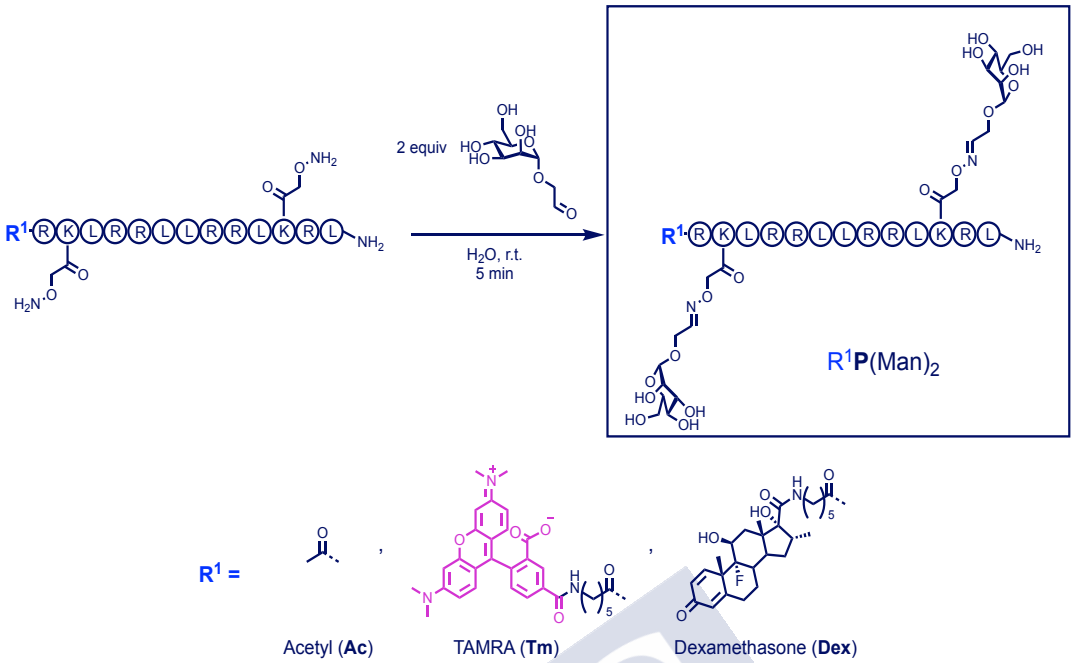


Figure S2. General synthetic scheme for the mannose aldehyde coupling. Peptides were obtained with an overall yield of: 28 % for Ac**P**(Man)$_2$, 5.6 % for Tm**P**(Man)$_2$ and 5 % for Dex**P**(Man)$_2$.

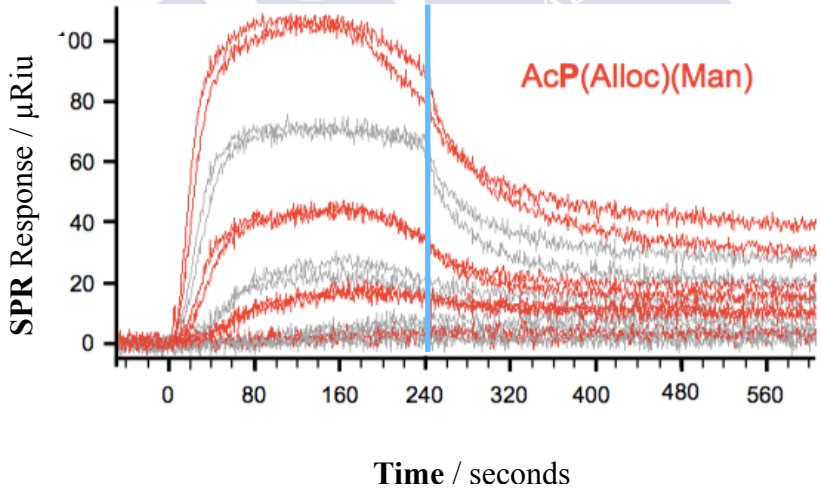


Figure S3. SPR. Sensorgrams of the interaction of ConA (9668 µRiU) in the concentration range [Ac**P**(Alloc)(Man)] = 6-2000 nM.

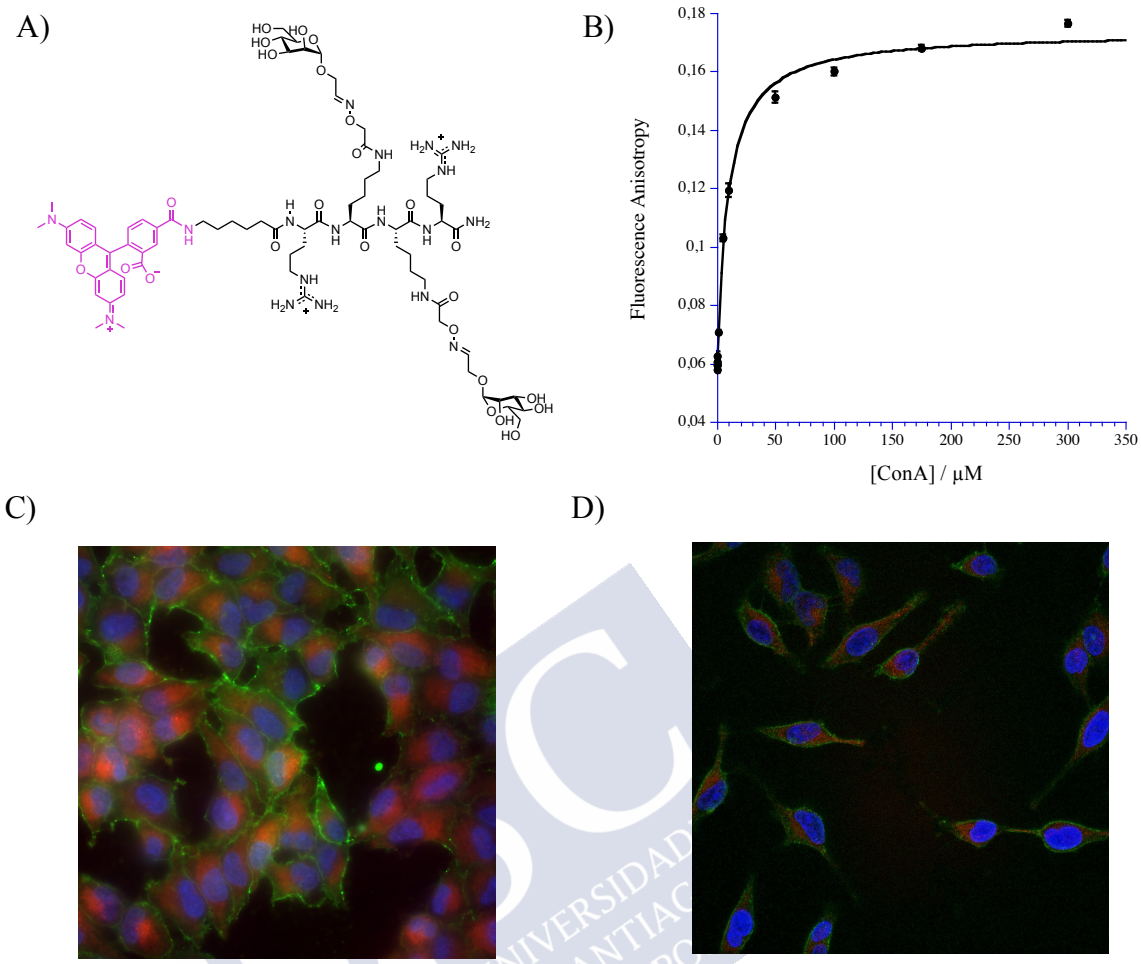


Figure S4. A) $TmArg_2(Man)_2$ peptide structure. B) Fluorescence anisotropy titration and best fitting to a simple 1:1 binding model of $TmArg_2(Man)_2$ with increasing amounts of ConA in HKR buffer (pH 7.4) at 22 ºC. Calculated K_d = 9 ± 1 µM. C) Epifluorescence and D) CLSM images of $TmArg_2(Man)_2$ (3 µM, red) treated with $ConA_{FITC}$ (30 nM, green) and incubated with HeLa cells. Nuclei were counterstained with Hoechst (blue).

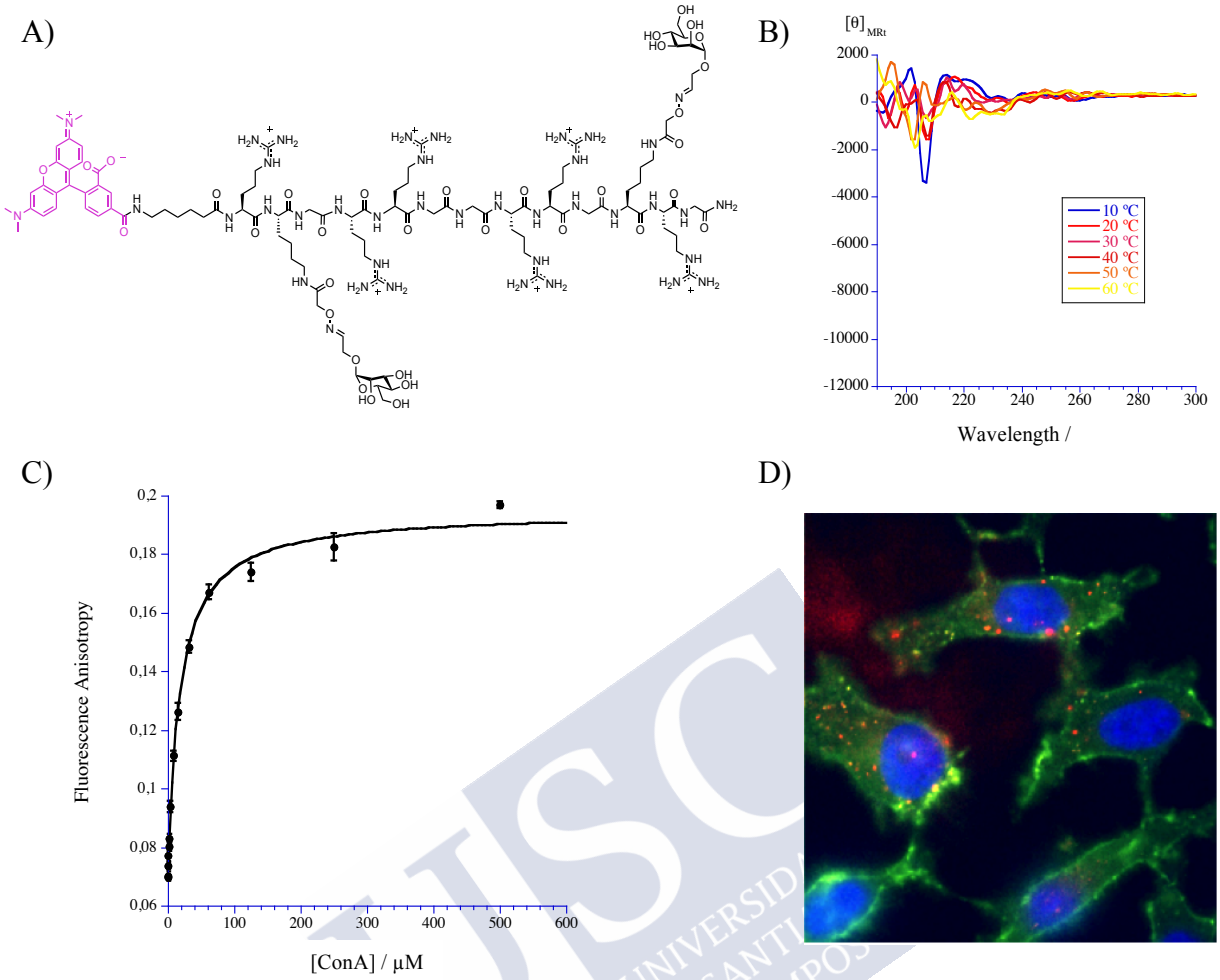


Figure S5. A) $TmArg_6Gly_5(Man)_2$ peptide structure. B) Circular dichroism of $TmArg_6Gly_5(Man)_2$ measured in HKR buffer at different temperatures. C) Fluorescence anisotropy titration and best fitting to a simple 1:1 binding model of $TmArg_6Gly_5(Man)_2$ with increasing amounts of ConA in HKR buffer (pH 7.4) at 22 ºC. Calculated K_d = 19 ± 2 µM. D) Epifluorescence image of $TmArg_6Gly_5(Man)_2$ (3 µM, red) treated with $ConA_{FITC}$ (30 nM, green) and incubated with HeLa cells. Nuclei were counterstained with Hoechst (blue).

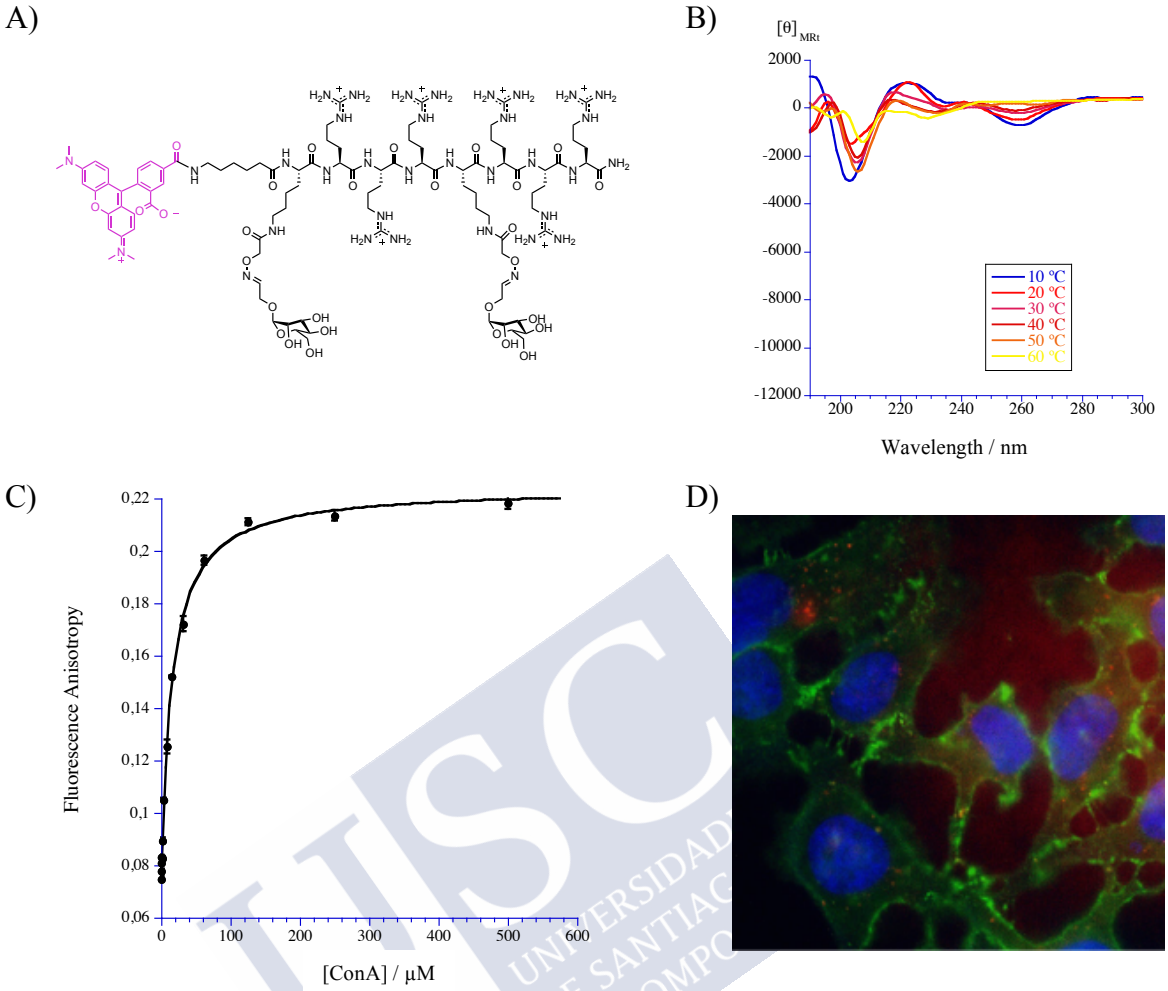


Figure S6. A) $TmArg_6(Man)_2$ peptide structure. B) Circular dichroism of $TmArg_6(Man)_2$ measured in HKR buffer at different temperatures. C) Fluorescence anisotropy titration and best fitting to a simple 1:1 binding model of $TmArg_6(Man)_2$ with increasing amounts of ConA in HKR buffer (pH 7.4) at 22 ºC. Calculated K_d = 16 ± 1 µM. D) Epifluorescence image of $TmArg_6(Man)_2$ (3 µM, red) treated with $ConA_{FITC}$ (30 nM, green) and incubated with HeLa cells. Nuclei were counterstained with Hoechst (blue).

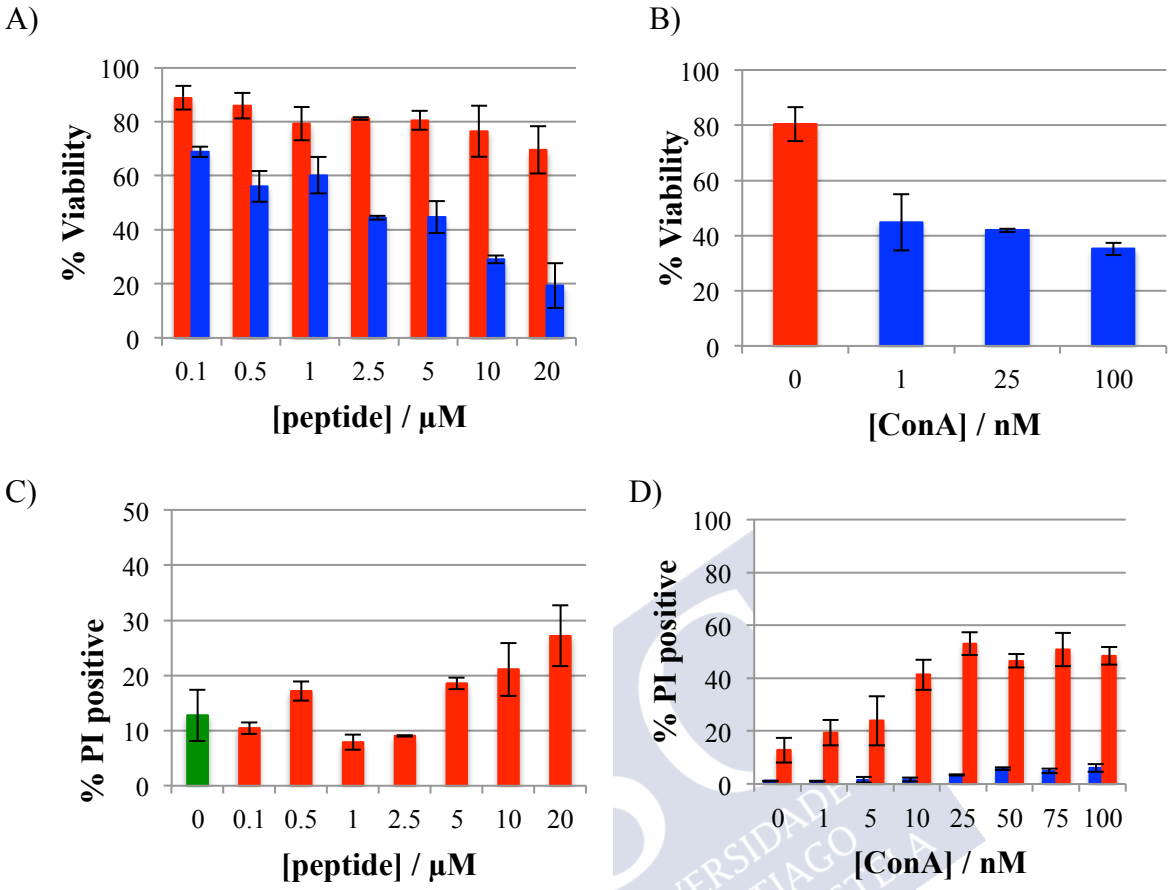


Figure S7. Cell viability in HeLa Cells. A) MTT assay at different concentrations of Ac**P**$(Man)_2$ (red) and Ac**P**$(Acetone)_2$ (blue); B) MTT assay at different concentrations of ConA using 5 µM of Ac**P**$(Man)_2$, the viability for Ac**P**$(Man)_2$ alone is shown in red at 0 nM of ConA; C) Propidium iodide staining of dead cells for flow cytometry at different concentrations of Ac**P**$(Man)_2$. The staining of untreated cells (control) is represented in green; D) Propidium iodide staining of dead cells for flow cytometry at different concentrations of ConA using 5 µM of Ac**P**$(Man)_2$ (red). Blue bars represent the cytotoxicity of the ConA alone.

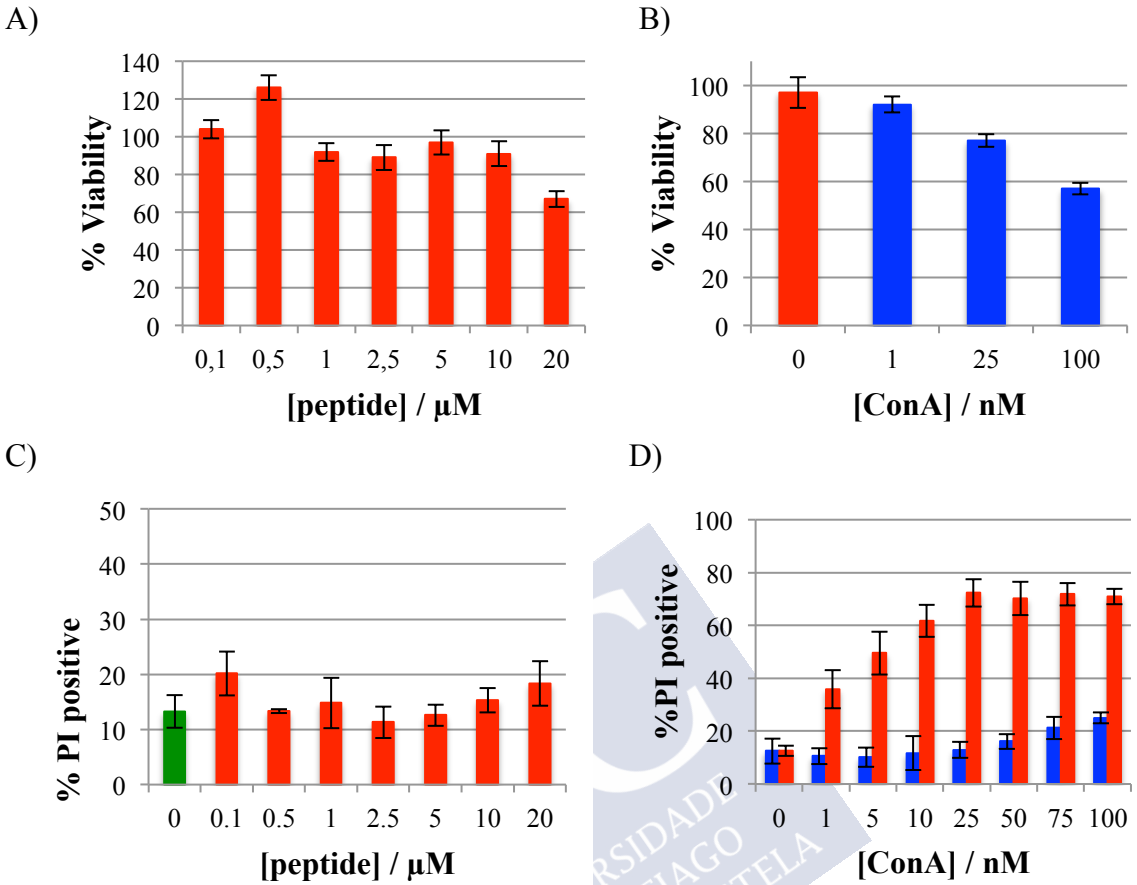


Figure S8. Cell viability in HepG2 Cells. A) MTT assay at different concentrations of Ac**P**$(Man)_2$ (red); B) MTT assay at different concentrations of ConA using 5 µM of Ac**P**$(Man)_2$, the viability for Ac**P**$(Man)_2$ alone is shown in red at 0 nM of ConA; C) Propidium iodide staining of dead cells for flow cytometry at different concentrations of Ac**P**$(Man)_2$. The staining of untreated cells (control) is represented in green; D) Propidium iodide staining of dead cells for flow cytometry at different concentrations of ConA using 5 µM of Ac**P**$(Man)_2$ (red). Blue bars represent the cytotoxicity of the ConA alone.

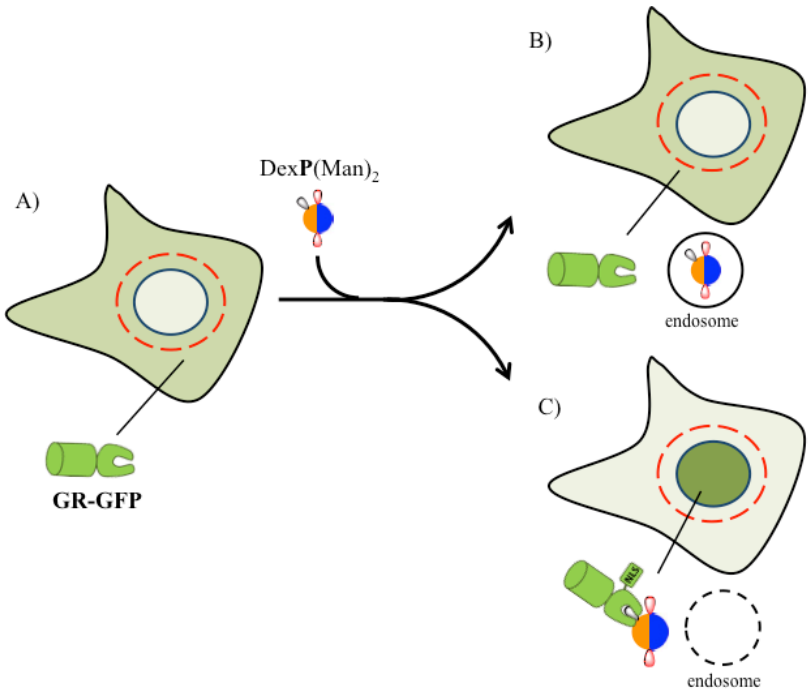


Figure S9. Glucocorticoid induced GFP translocation assay (GIGT). A) Cells transfected with the plasmid pK7-GR-GFP show green fluorescence mostly on the cytoplasm, and so, the ratio between the median fluorescence in the nucleus and in the surrounding cytosolic region (red circle) will be low. Upon incubation with dexamethasone-labelled peptides, if the peptide remains trapped in the endosome (B) or it does not enter the cell, the glucocorticoid receptor will not be able to bind the dexamethasone and the translocation ratio will remain low. However, if the peptide is able to reach the cytosol, by direct translocation or endosomal escape (C), the binding to the glucocorticoid receptor triggers a conformational change that exposes NLSs. This causes the accumulation of GR-GFP in the nucleus and the increase in the translocation ratio.

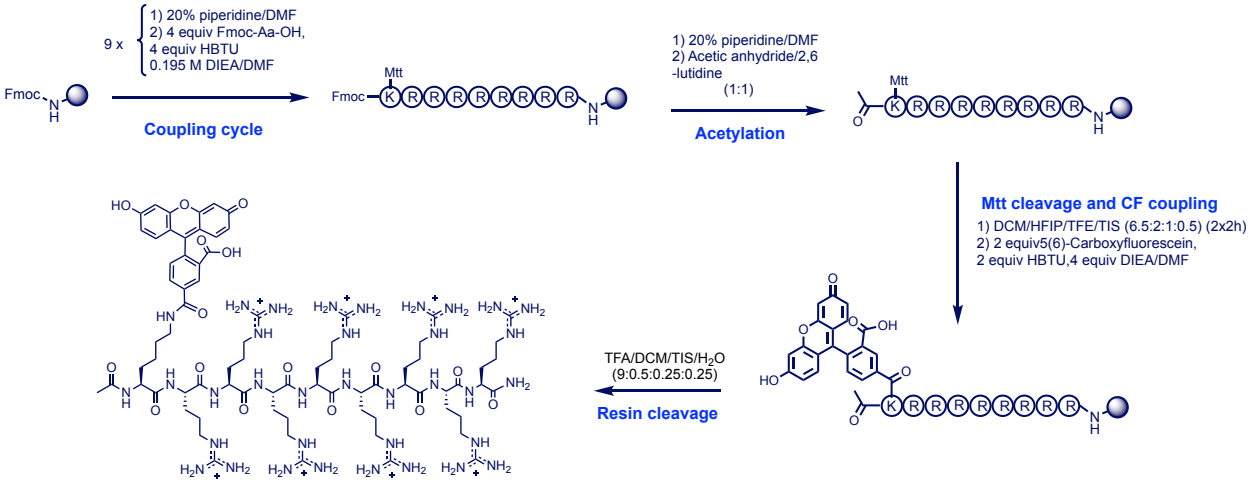


Figure S10. General synthetic scheme for the Solid Phase Peptide Synthesis (SPPS) for synthesizing the CFArg$_8$ peptide.

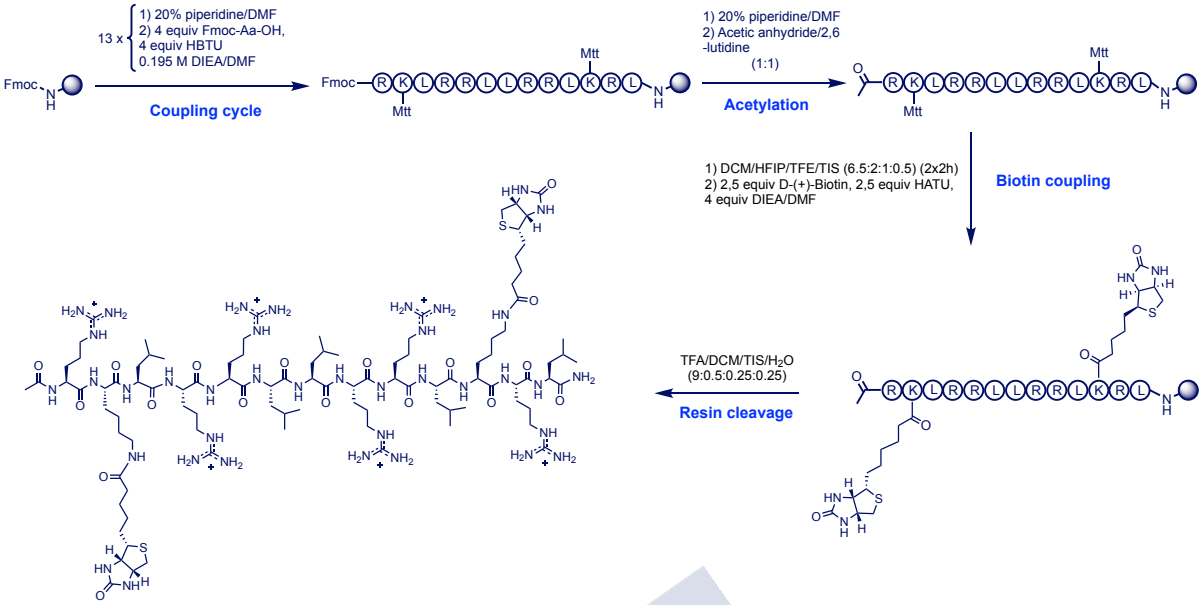


Figure S11. General synthetic scheme for the Solid Phase Peptide Synthesis (SPPS) for synthesizing the Ac**P**$(Biot)_2$ peptide.

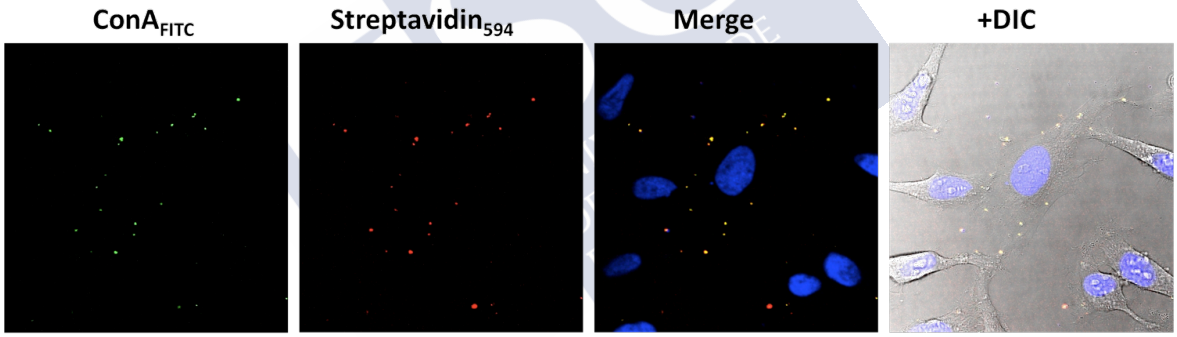


Figure S12. Lipofectamine is not a selective carrier. CLSM images of HeLa cells incubated with a mixture of 30 nM $ConA_{FITC}$, 30 nM $Streptavidin_{594}$ and 2 ng/µL of Lipofectamine 2000 (protein/Lipofectamine complex formation was done by incubation in a tenth of the final volume for 20 min) and incubated 30 min at 37 ºC. Both proteins co-localize in punctate structures, located at the perifery of the cell. Nuclei were counterstained with Hoechst (blue).

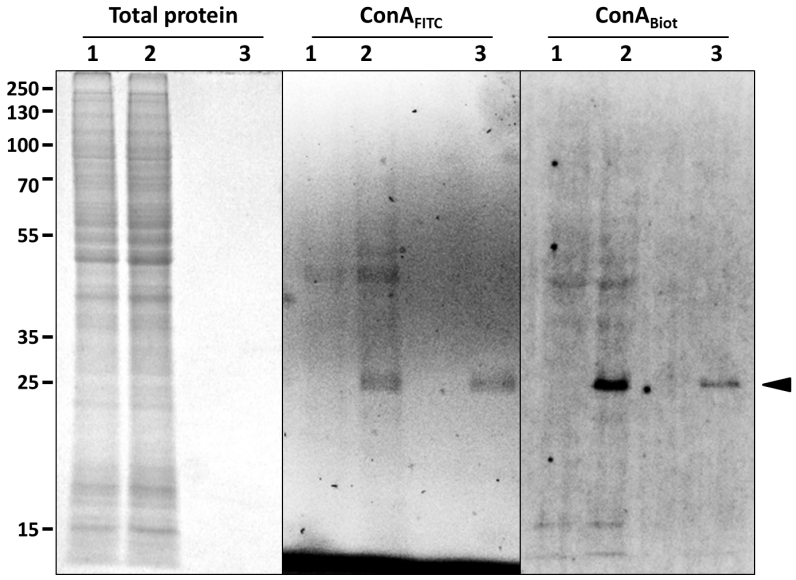


Figure S13. Gel electrophoresis to confirm integrity of the cargo after delivery. HeLa cells untreated (lane 1) or incubated with 3 μM of AcP(Man)$_2$ and 30 nM of FITC- or biotin-labeled ConA (as indicated over the panels) (lane 2) for 30 min at 37 ºC were washed, trypsinized and lysed, and proteins in the lysates resolved by SDS-PAGE. A small amount of the original protein was loaded as a control (lane 3). ConA$_{FITC}$ was imaged in the unfixed gel under UV light (ConA$_{FITC}$ panel) and afterwards the gel was fixed and stained with Coomassie Brilliant Blue to confirm that similar amounts of protein were loaded in lanes 1 and 2 (Total protein panel). ConA$_{Biot}$ was detected after protein transfer to PVDF membrane and incubation with Streptavidin$_{488}$ (Alexa Fluor-488). Arrowhead indicates the expected position of the ConA monomer (~25 kDa).

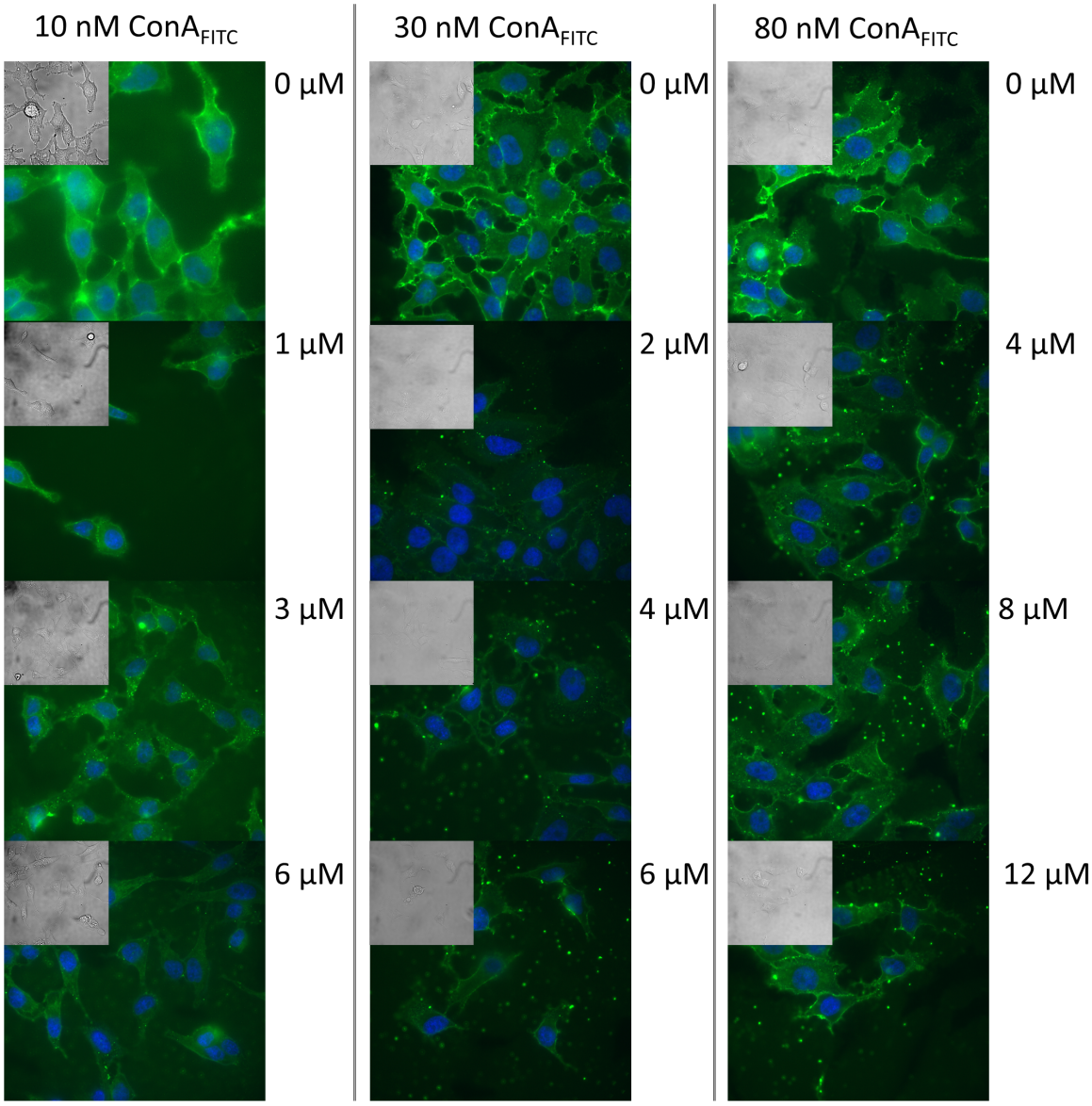


Figure S14. Importance of peptide/lectin ratio. Different amounts of ConA$_{FITC}$, as indicated above the panels, were incubated for 30 min with HeLa cells in the presence of the concentrations of Ac**P**(Man)$_2$ indicated at the right of the panels. Nuclei were counterstained with Hoechst (blue).

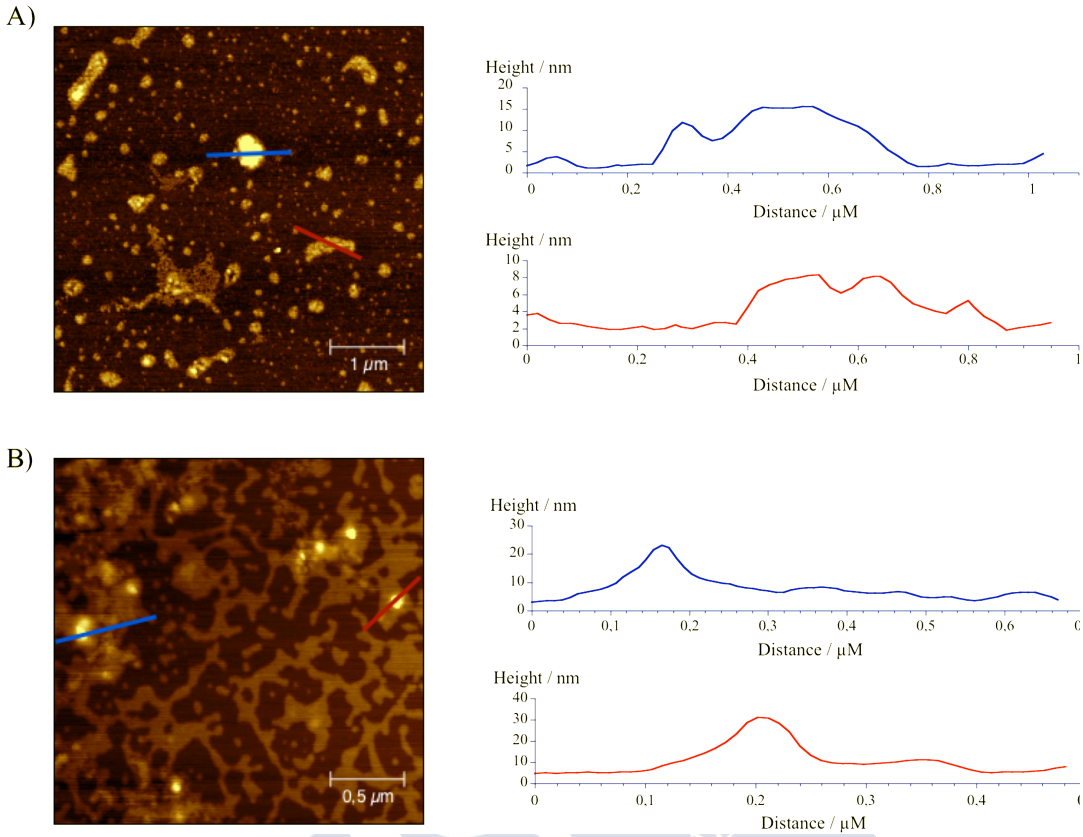


Figure S15. AFM topography images of A) Ac**P**$(Man)_2$ (2 µM) and B) Ac**P**$(Man)_2$ (4 µM) with ConA (30 nM) aqueous dispersions deposited on mica surface; images showed highly disperse aggregates. Dispersions prepared and deposited analogously as described in methods.

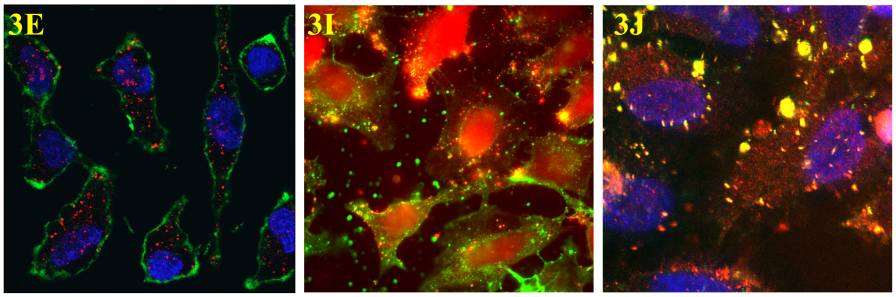


Figure	Manders' tM_1 (Peptide)	Manders' tM_2 (Protein)
3E	0.03	0.04
3I	0.52	0.64
3J	0.66	0.63

Table S1. Co-localization parameters calculated for figures 3E, 3I, and 3J. 3E) CLSM image of $CFArg_8$ (red) and $ConA_{647}$ (green); 3I) epifluorescence image of Tm**P**$(Man)_2$ (red) and $ConA_{647}$ (green); 3J) CLSM image of Tm**P**$(Man)_2$ (red) and $ConA_{FITC}$ (green). Co-localization parameters were calculated with ImageJ using the plugin Coloc2 and the Costes method for the estimation of the threshold.

Peptide	K_d
Tm**P**$(Man)_2$	14 ± 1
Tm**P**(Alloc)(Man)	27 ± 4
Tm**P**$(Acetone)_2$	151 ± 7
$TmArg_2(Man)_2$	9 ± 1
$TmArg_6Gly_5(Man)_2$	19 ± 2
$TmArg_6(Man)_2$	16 ± 1
Methyl α-D-mannopyranoside	137 ± 29

Table S2. Equilibrium dissociation constants obtained by fluorescence anisotropy after fitting to a simple 1:1 binding model.

Video S1. Internalization kinetics. To study the internalization kinetics, Tm**P**$(Man)_2$ (3 μM) and $ConA_{FITC}$ (30 nM) were incubated for 7 min in HKR to generate peptide/protein complexes. This solution was added to HeLa cells, previously washed with HKR, and the plate was immediately imaged on an epifluorescence microscope. Images for both peptide (red) and protein (green) channels were taken at intervals of 3 min for 20 min.

1.2 Materials and Methods

Commercially available Rink Amide-resin, *N*-HBTU, EDC and Fmoc-*L*-Lys(Mtt)-OH were used as obtained from Iris. [(*tert*-butoxycarbonyl)aminooxy]acetic acid, Chlorpromazine and Dexamethasone were purchased from TCI Chemicals. D-(+)-Biotin was used as obtained from Alfa Aesar. Dimethylaminopyridine, N-Hydroxysuccinimide, Fmoc-*L*-Gly-OH, Fmoc-*L*-Leu-OH, Fmoc-*L*-Arg(Pbf)-OH, Fmoc-6Ahx-OH, 5(6)-Carboxyfluorescein, Concanavalin A from *Canavalia ensiformis*, Concanavalin A from *Canavalia ensiformis* (Jack bean) FITC conjugate, propidium iodide, Ammonium chloride, Chloroquine, Heparin sodium salt and 5-(N-Ethyl-N-isopropyl)amiloride were purchased from Sigma-Aldrich®. 5(6)-Carboxytetramethlyrhodamine succinimidyl ester, 4-methylumbelliferyl-α-*D*-mannopyranoside (4-MU-α-D-Man) and Methyl-β-cyclodextrin were purchased from Carbosynth. Hoechst 33342 Trihydrochloride Trihydrate, Concanavalin A Alexa Fluor 647 conjugate, Lysotracker green DND-26, Streptavidin Alexa Fluor 594, Streptavidin Alexa Fluor 488, Lipofectamine 2000 and Coomassie (Bradford) protein assay kit were purchased in ThermoFisher. Wortmannin was obtained from Fluorochem. Egg yolk *L*-α-phosphatidylcholine (EYPC) was purchased from Avanti Polar Lipids, Inc. 1-α-formylmethyl-mannopyranoside,[S1] Ox-Dexamethasone (Ox-Dex)[S2] were synthesized according to protocols described in the literature[1,2]. Deuterated solvents (D_2O and $CDCl_3$) were from EMD Millipore Corporation. *N,N*-Dimethylformamide, for peptide synthesis, was purchased from Scharlau. All the other solvents were HPLC grade, purchased from Sigma-Aldrich® or Fisher Scientific®, and used without further purification.

High-performance liquid chromatography coupled with mass spectrometry (HPLC-MS) analyses were carried out on Agilent Technologies 1260 Infinity II associated with a 6120 Quadrupole LC-MS using an Agilent SB-C18 column or on DIONEX Ultimate 3000 U-HPLC$^+$ (Thermo Scientific) with an Acclaim RSLC 120-C18 column with *Solvent A:Solvent B* gradients between 5:95 and 5:75 (*Solvent A*: H_2O with 0.1 % TFA; *Solvent B*: CH_3CN with 0.1 % TFA). High-performance liquid chromatography (HPLC) preparative purification was carried out on Waters 1525 composed by a binary pump with a dual Waters 2489 detector with a Phenomenex Luna C18(2) 100A column. An Agilent 1200 with an Agilent Eclipse XDB-C18 column was used for semi-preparative purification using gradients between 5:95 and 5:75 (*Solvent A*: H_2O with 0.1 % TFA; *Solvent B*: CH_3CN with 0.1 % TFA). Nuclear Magnetic Resonance (NMR) spectra were recorded on either a Varian Mercury 300 MHz or a Varian Inova 500 MHz spectrometer. Chemical shifts are reported in ppm (δ units) referenced to the following solvent signals: D_2O δ H 4.79 and $CDCl_3$ δ H 7.26. Spin multiplicities are reported as a singlet (s), doublet (d), triplet (t) with coupling constants (*J*) given in Hz, or multiplet (m). Accurate mass determinations (HR-MS) using ESI-MS were performed on a Sciex QSTAR Pulsar mass spectrometer and are reported as mass-per-charge ratio (*m/z*). Recalculation of the labelled peptides concentrations was performed by measuring the absorbance on a Dynamica HALO XB-10 UV-VIS Single Beam Spectrophotometer. Fluorescence measurements were performed with a FluoroMax-2 spectrofluorometer (Jobin-Yvon Spex) equipped with a stirrer and a temperature controller. Circular Dichroism (CD)

measurements were performed with a Jasco J-1100 CD Spectrometer equipped with a Jasco MCB-100 Mini Circulation Bath. Cell microscopy images were acquired with an Andor Zyla 4.2 digital camera mounted on a Nikon Eclipse Ti-E epifluorescence microscope and with a Leica SP5 confocal microscope. Absorbance and fluorescence of cellular extracts were measured using a microplate reader Tecan Infinite F200Pro. Absorbance of labelled Concanavalin A was measured using a NanoDrop 1000 spectrophotometer. AFM images were acquired in a Park NX10 microscope in non-contact mode using ACTA silicon cantilevers with 37 N/m nominal spring constant (k) and 300 kHz nominal resonance frequency.

1.3 Abbreviations

Peptide Abbreviations: Tm**P**$(Man)_2$ (Tm = TAMRA and Man = Mannose aldehyde); Aa: Amino acid; AFM: Atomic Force Microscope; Arg: Arginine; BSA: Bovine Serum Albumin; Boc: *tert*-Butoxycarbonyl; Calcd: Calculated; CF: 5(6)-Carboxyfluorescein; CLSM: Confocal laser scanning microscopy; ConA: Concanavalin A; CPP: Cell-Penetrating Peptide; DCM: Dichloromethane; DIEA: N,N-Diisopropylethylamine; DMAP: 4-Dimethylaminopyridine; DMEM: Dulbecco's Modified Eagle Medium; DMF: N,N-Dimethylformamide; DMSO: Dimethylsulfoxide; EDC: 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride; EIPA: 5-(N-Ethyl-N-isopropyl)amiloride; EYPC: Egg yolk phosphatidylcholine; FBS: Fetal Bovine Serum; Fmoc: 9-fluorenylmethoxycarbonyl; GFP: Green Fluorescent Protein; GR-GFP: Glucocorticoid receptor-Green Fluorescent Protein; HFIP: 1,1,1,3,3,3-Hexafluoro-2-propanol; HKR: HEPES-Krebs-Ringer buffer; HRMS (ESI): High resolution mass spectrometry (electrospray ionization); Lys: Lysine; Mtt: 4-Methyltrityl; MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; *N*-HATU: N-[(Dimethylamino)-1H-1,2,3-triazolo[4,5-b]pyridin-1-ylmethylene]-N-methylmethanaminiun hexafluorophosphate N-oxide; *N*-HBTU: N-[(1H-Benzotriazol-1-yl)4-(dimethylamino)methylene]-N-methylmethanaminium hexafluorophosphate N-oxide; NHS: N-Hydroxysuccinimide; NLS: nuclear localization signal; Ox: Oxime; Ox-Dex: Dexamethasone Acid; Pbf: 2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl; PI: propidium iodide; RP: Reverse Phase; SPPS: Solid Phase Peptide Synthesis; TAMRA: 5(6)-Carboxytetramethlyrhodamine succinimidyl ester; TFE: Trifluoroethanol; TIS: Triisopropylsilane; TNBS: 2,4,6-Trinitrobenzenesulfonic acid; 6Ahx: 6-aminohexanoic acid.

1.4 General protocols

1.4.1 General protocol for the SPPS

All peptides were synthesized by manual Fmoc solid-phase peptide synthesis[S3] using Rink Amide resin (loading 0.71 mmol/g). The resin (0.1 mmol) was swelled in DMF (peptide synthesis grade, 2 mL) for 20 min in a peptide synthesis vessel prior synthesis. Coupling cycle consisted of the removal of Fmoc protecting group with a solution of piperidine in DMF (20%, 2 mL) for 10 min and then the mixture was filtered and the resin was washed with DMF (3 x 2 mL, 1 min). The amino acid coupling was carried out by treatment with a

solution of α-amino acids (4 equiv), *N*-HBTU (3.95 equiv) in DMF (2 mL), which was mixed with DIEA (0.195 M solution in DMF, 1.2 equiv) 1 min before the addition and the resulting mixture was shaken by bubbling Ar for 15 min. Finally, the resin was washed with DMF (3 x 2 mL, 1 min). The efficiency of each amino acid coupling and deprotection was monitored employing the TNBS test[S4].

Once the linear peptide was finished, two different ending protocols were used:

A) Acetylation: the acetylation capping of N-terminal group was performed by standard Fmoc removal conditions (20% piperidine in DMF) followed by treatment with a solution of acetic anhydride and 2,6-lutidine (1:1, 1 mL) for 30 min.

B) Linker coupling: after Fmoc cleavage with piperidine/DMF (20%, 2 mL), the linear peptide was treated with a solution of N-Fmoc-6-aminohexanoic acid (4 equiv), *N*-HBTU (3.95 equiv) and DIEA (0.195 M solution in DMF, 1.2 equiv) in DMF.

The resin was washed with DCM (2 x 2 mL, 5 min), and the Mtt protecting group was selectively removed by mechanical shaking of the resin with a mixture of DCM/HFIP/TFE/TIS (6.5:2:1:0.5, 2 x 2 mL, 2 h). Finally, the mixture was filtered and the resin was washed with DCM (2 x 2 mL, 2 min) and DMF (2 mL, 20 min). A solution of [(*tert*-butoxycarbonyl)aminooxy]acetic acid (2.5 equiv per free amine) and *N*-HATU (2.5 equiv) in DMF (1 mL) was added to the resin followed by the dropwise addition of a solution of DIEA (4 equiv) in DMF (0.5 mL). The resin was shaken by bubbling Ar for 30 min and finally washed with DMF (3 x 2 mL, 2 min) and DCM (3 x 2 mL, 2 min).

1.4.2 General protocol for N-terminal functionalization

Fluorescently labelled peptides, the Fmoc-protecting group of the previously attached linker was removed by using a solution of piperidine in DMF (20%, 4 mL) for 15 min and the resin was washed with DMF (3 x 3 mL). The coupling was carried out by the addition of a solution of 5(6)-Carboxytetramethylrhodamine succinimidyl ester (1 equiv) and DIEA (0.195 M, 1 equiv) in DMF (2 mL) and the mixture was stirred by bubbling Ar for 30 min. Finally, the resin was washed with DMF (3 x 3 mL) and DCM (3 x 3 mL).

Ox-Dex terminating peptide, the Fmoc-protecting group of the linker was removed by using a solution of piperidine in DMF (20%, 4 mL) during 15 min and the resin was washed with DMF (3 x 3 mL). A solution of Ox-Dex (3 equiv), *N*-HATU (2.9 equiv) and DIEA (0.195 M, 3 equiv) in DMF (2 mL) was added and the mixture was shaken by bubbling Ar for 30 min. Finally, the resin was washed with DMF (3 x 3 mL) and DCM (3 x 3 mL).

1.4.3 General protocol for peptide cleavage

Finally, peptides were deprotected and cleaved from the resin by standard TFA cleavage procedure at rt by using TFA/DCM/H_2O/TIS (90:5:2.5:2.5, 1 mL per 70 mg of resin) for 2 h. Then, the mixture was filtered, washed with TFA (1 mL) and the peptide was precipitated with ice-cold Et_2O (25 mL). The precipitate was centrifuged and dissolved in H_2O (5 mL).

Peptides were obtained following the previously described procedure, and were treated with the different ligands without purification.

1.4.4 General protocol for ligand coupling

A solution of peptide (R^1**P**$(Ox)_2$) in H_2O (5 mM) was reacted with a solution of corresponding ligands (2 equiv per alkoxyamine) [1-α-formylmethyl-mannopyranoside or acetone] in H_2O (120 mM) for 5 min. Then, peptides were purified by RP-HPLC for removing the ligand excess.

The purification was carried by a C18 RP-HPLC [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with a binary gradient of *Solvent A* and *Solvent B*, the collected fractions were lyophilized and stored at -20 °C. Purity and identity were confirmed by analytical HPLC, ^{1}H-NMR and low and high resolution mass spectrometry.

1.5 Synthesis of peptides

1.5.1 Synthesis of AcP$(Man)_2$

Following the general protocol of the SPPS for synthesizing an acetylated peptide with two mannoses, Ac**P**$(Man)_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 28% and 99.3% purity. R_t 3.8 min (Fig. S16) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 7.77 (t, J = 5.3 Hz, 1H), 7.17 (t, J = 3.9 Hz, 1H), 4.93 (s, 2H), 4.59 (s, 4H), 4.45-4.22 (m, 17H), 4.07-3.64 (m, 12H), 3.33-3.16 (m, 16H), 2.07 (s, 3H), 1.98-1.32 (m, 51H), 1.03-0.82 (m, 30H). **MS** (ESI, H_2O): 1300 (19, $[M+2H+2TFA]^{2+}$), 1243 (21, $[M+2H+TFA]^{2+}$), 1187 (23, $[M+2H]^{2+}$), 867 (42, $[M+3H+2TFA]^{3+}$), 827 (67, $[M+3H+TFA]^{3+}$), 791 (100, $[M+3H]^{3+}$), 593 (38, $[M+4H]^{4+}$). **HRMS** (ESI): Calcd for $C_{100}H_{188}N_{36}O_{30}$: 1186.7139; found: 1186.7141 ($[M+2H]^{2+}$).

1.5.2 Synthesis of TmP$(Man)_2$

Following the general protocol of the SPPS for synthesizing a TAMRA labelled peptide with two mannoses, Tm**P**$(Man)_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 6% and 100% purity. R_t 4.1 min (Fig. S17) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (500 MHz, D_2O, δ): 8.21 (s, 1H), 7.92 (d, J = 8.1 Hz, 1H), 7.58 (m, 1H), 7.36 (d, J = 8.1 Hz, 1H), 7.11 (t, J = 9.9 Hz, 2H), 6.97 (m, 1H), 6.83 (d, J = 9.6 Hz, 2H), 6.75 (s, 2H), 4.80-4.73 (m, 2H), 4.68 (s, 4H), 4.40 (d, J = 24.1 Hz, 4H), 4.28-3.96 (m, 13H), 3.87-3.24 (m, 12H), 3.13 (s, 12H), 3.10-2.95 (m, 16H), 2.91-2.80 (m, 2H), 2.25 (t, J = 7.1 Hz, 2H), 1.79-1.11 (m, 57H), 0.87-0.67 (m, 30H). **MS** (ESI, H_2O): 1657 (7, $[M+2H+4TFA]^{2+}$), 1599 (17, $[M+2H+3TFA]^{2+}$), 1542 (15, $[M+2H+TFA]^{2+}$), 1027 (100, $[M+3H+2TFA]^{3+}$), 989

(99, $[M+3H+TFA]^{3+}$), 743 (62, $[M+4H+TFA]^{4+}$), 716 (53, $[M+4H]^{4+}$). **HRMS** (ESI): Calcd for $C_{129}H_{217}N_{39}O_{34}$: 1428.3209; found: 1428.3220 ($[M+2H]^{2+}$).

1.5.3 Synthesis of AcP(Alloc)(Man)

Following the general protocol of the SPPS for synthesizing an acetylated peptide with one mannose, and using Alloc protected Lysine, Ac**P**(Alloc)(Man) was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 15% and 100% purity. R_t 4.1 min (Fig. S18) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 7.62 (t, J = 5.3 Hz, 1H), 7.02 (t, J = 3.7 Hz, 1H), 5.90-5.73 (m, 1H), 5.23-5.06 (m, 2H), 4.78 (s, 1H), 4.48-4.39 (m, 5H), 4.31-4.04 (m, 15H), 3.89-3.46 (m, 6H), 3.15-2.95 (m, 16H), 1.92 (s, 3H), 1.78-1.12 (m, 51H), 0.87-0.70 (m, 30H). **MS** (ESI, H_2O): 1260 (18, $[M+2H+3TFA]^{2+}$), 1203 (27, $[M+2H+2TFA]^{2+}$), 1147 (20, $[M+2H+TFA]^{2+}$), 803 (38, $[M+3H+2TFA]^{3+}$), 763 (100, $[M+3H+TFA]^{3+}$), 727 (42, $[M+3H]^{3+}$), 574 (10, $[M+4H+TFA]^{4+}$), 547 (27, $[M+4H]^{4+}$). **HRMS** (ESI): Calcd for $C_{94}H_{175}N_{35}O_{24}$: 1090.6868; found: 1090.6871 ($[M+2H]^{2+}$).

1.5.4 Synthesis of TmP(Alloc)(Man)

Following the general protocol of the SPPS for synthesizing a TAMRA labelled peptide with one mannose, and using Alloc protected Lysine, Tm**P**(Alloc)(Man) was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 10% and 98.4% purity. R_t 4.3 min (Fig. S19) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 8.24 (s, 1H), 8.08-7.87 (m, 2H), 7.55 (m, 1H), 7.03 (s, 2H), 6.88-6.61 (m, 3H), 5.71 (m, 2H), 5.07 (m, 3H), 4.52-4.30 (m, 4H), 4.29 (s, 2H), 4.26-3.92 (m, 13H), 3.86-3.44 (m, 6H), 3.30 (m, 2H), 3.09 (s, 12H), 3.08 (s, 16H), 2.22 (m, 2H), 1.96-1.12 (m, 57H), 0.95-0.59 (m, 30H). **MS** (ESI, H_2O): 965 (43, $[M+3H+2TFA]^{3+}$), 927 (100, $[M+3H+TFA]^{3+}$), 668 (40, $[M+4H]^{4+}$), 533 (10, $[M+5H]^{5+}$). **HRMS** (ESI): Calcd for $C_{123}H_{204}N_{38}O_{28}$: 888.5326; found: 888.5330 ($[M+3H]^{3+}$).

1.5.5 Synthesis of AcP(Acetone)$_2$

Following the general protocol of the SPPS for synthesizing an acetylated peptide capped with acetone, Ac**P**(Acetone)$_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 4% and 100% purity. R_t 4.3 min (Fig. S20) **[**RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 4.36 (s, 4H), 4.27-4.05 (m, 13H), 3.20-3.00 (m, 16H), 1.92 (s, 3H), 1.83 (s, 6H), 1.76 (s, 6H), 1.71-1.17 (m, 51H), 0.85-0.69 (m, 30H); **MS** (ESI, H_2O): 1193 (30, $[M+2H+3TFA]^{2+}$), 1133 (39, $[M+2H+2TFA]^{2+}$), 1077 (23, $[M+2H+TFA]^{2+}$), 757 (30, $[M+3H+3TFA]^{3+}$), 720 (100, $[M+3H+TFA]^{3+}$), 680 (61, $[M+3H]^{3+}$), 513 (30, $[M+4H]^{4+}$). **HRMS** (ESI): Calcd for $C_{90}H_{172}N_{36}O_{18}$: 1022.6820; found: 1022.6820 ($[M+2H]^{2+}$).

1.5.6 Synthesis of TmP(Acetone)$_2$

Following the general protocol of the SPPS for synthesizing a TAMRA labelled peptide capped with acetone, Tm**P**(Acetone)$_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 12% and 99.9% purity. R_t 4.4 min (Fig. S21) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 8.07-7.89 (m, 2H), 7.55 (s, 1H), 7.03 (d, J = 9.2 Hz, 2H), 6.89-6.69 (m, 4H), 4.34 (d, J = 7.9 Hz, 4H), 4.27-3.93 (m, 13H), 3.40-3.22 (m, 2H), 3.10 (s, 12H), 3.07-2.98 (m, 16H), 2.31-2.11 (m, 2H), 1.89-1.63 (m, 12H), 1.60-1.10 (m, 57H), 0.88-0.64 (m, 30H). **MS** (ESI, H_2O): 1492 (8, $[M+2H+4TFA]^{2+}$), 1434 (10, $[M+2H+3TFA]^{2+}$), 1376 (8, $[M+2H+2TFA]^{2+}$), 918 (100, $[M+3H+2TFA]^{3+}$), 881 (100, $[M+3H+TFA]^{3+}$), 690 (18, $[M+4H+2TFA]^{4+}$), 662 (40, $[M+4H+TFA]^{4+}$), 632 (18, $[M+4H]^{4+}$). **HRMS** (ESI): Calcd for $C_{119}H_{201}N_{39}O_{22}$: 1264.2902; found: 1264.2899 ($[M+2H]^{2+}$).

1.5.7 Synthesis of AcArg$_2$(Man)$_2$

Following the general protocol of the SPPS for synthesizing an acetylated tetrapeptide with two mannoses, AcArg$_2$(Man)$_2$ was obtained after RP- HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 15% and 100% purity. R_t 3.1 min (Fig. S22) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 7.62 (t, J = 5.3 Hz, 1H), 7.02 (t, J = 3.9 Hz, 1H), 4.78 (s, 2H), 4.44 (s, 4H), 4.25-4.07 (m, 8H), 3.90-3.47 (m, 12H), 3.20-2.98 (m, 8H), 1.91 (s, 3H), 1.83-1.13 (m, 20H). **MS** (ESI, H_2O): 1183 (17, $[M+H]^+$), 590 (100, $[M+2H]^{2+}$). **HRMS** (ESI): Calcd for $C_{46}H_{84}N_{15}O_{21}$: 1182.5963; found: 1182.5961 ($[M+H]^+$).

1.5.8 Synthesis of TmArg$_2$(Man)$_2$

Following the general protocol of the SPPS for synthesizing a TAMRA labelled tetrapeptide with two mannoses, TmArg$_2$(Man)$_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 26% and 100% purity. R_t 3.8 min (Fig. S23) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 8.35 (d, J = 14.1 Hz, 1H), 7.99 (t, J = 5.5 Hz, 1H), 7.56 (t, J = 10.4 Hz, 1H), 7.43 (d, J = 7.8 Hz, 1H), 7.11-6.98 (m, 2H), 6.95 (s, 1H), 6.80 (d, J = 9.5 Hz, 2H), 6.64 (d, J = 7.0 Hz, 2H), 4.73 (s, 2H), 4.38 (s, 4H), 4.22-3.99 (m, 4H), 3.85-3.33 (m, 12H), 3.10 (s, 12H), 3.08-2.98 (m, 12H), 2.86 (t, J = 7.5 Hz, 2H), 2.24 (t, J = 7.0 Hz, 2H), 1.81-1.12 (m, 26H). **MS** (ESI, H_2O): 834 (100, $[M+2H]^{2+}$), 556 (95, $[M+3H]^{3+}$). **HRMS** (ESI): Calcd for $C_{75}H_{114}N_{18}O_{25}$: 833.4088; found: 833.4096 ($[M+2H]^{2+}$).

1.5.9 Synthesis of $TmArg_6Gly_5(Man)_2$

Following the general protocol of the SPPS for synthesizing a TAMRA labelled peptide with two mannoses, $TmArg_6Gly_5(Man)_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 9% and 100% purity. R_t 3.4 min (Fig. S24) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (500 MHz, D_2O, δ): 8.42 (s, 1H), 8.02 (s, 1H), 7.66-7.49 (m, 1H), 7.41 (s, 1H), 7.09-6.88 (m, 3H), 6.71 (s, 2H), 6.29 (s, 2H), 4.77 (s, 2H), 4.74 (s, 4H), 4.46-4.30 (m, 4H), 4.27-4.04 (m, 10H), 3.98-3.73 (m, 16H), 3.75-3.45 (m, 12H), 3.37 (d, J = 5.3 Hz, 2H), 3.17-3.01 (m, 12H), 2.98 (s, 8H), 2.24 (s, 2H), 1.85-1.10 (m, 42H). **MS** (ESI, H_2O): 1402 (7, $[M+2H+2TFA]^{2+}$), 973 (42, $[M+3H+3TFA]^{3+}$), 936 (100, $[M+3H+2TFA]^{3+}$), 898 (65, $[M+3H+TFA]^{3+}$), 702 (60, $[M+4H+2TFA]^{4+}$), 674 (95, $[M+4H+TFA]^{4+}$), 645 (35, $[M+4H]^{4+}$), 516 (45, $[M+5H]^{5+}$). **HRMS** (ESI): Calcd for $C_{109}H_{177}N_{39}O_{34}$: 1288.1651; found: 1288.1655 ($[M+2H]^{2+}$).

1.5.10 Synthesis of $TmArg_6(Man)_2$

Following the general protocol of the SPPS for synthesizing a TAMRA labelled peptide with two mannoses, $TmArg_6(Man)_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 9% and 100% purity. R_t 3.4 min (Fig. S25) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 8.41 (s, 1H), 7.99 (d, J = 8.5 Hz, 1H), 7.57-7.29 (m, 2H), 7.00-6.85 (m, 3H), 6.68 (d, J = 9.3 Hz, 2H), 6.33 (s, 2H), 4.84-4.66 (m, 2H), 4.70-4.65 (m, 4H), 4.32 (d, J = 15.0 Hz, 4H), 4.17-3.91 (m, 8H), 3.88-3.37 (m, 12H), 3.34-3.24 (m, 2H), 2.98 (s, 12H), 2.96 (s, 16H), 2.23-2.08 (m, 2H), 1.77-1.03 (m, 42H). **MS** (ESI, H_2O): 840 (90, $[M+3H+2TFA]^{3+}$), 803 (100, $[M+3H+TFA]^{3+}$), 765 (30, $[M+3H]^{3+}$), 629 (30, $[M+4H+2TFA]^{4+}$), 602 (25, $[M+4H+TFA]^{4+}$). **HRMS** (ESI): Calcd for $C_{99}H_{162}N_{34}O_{29}$: 1145.6119; found: 1145.6118 ($[M+2H]^{2+}$).

1.5.11 Synthesis of $DexP(Man)_2$

Following the general protocol of the SPPS for synthesizing an Dex labelled peptide with two mannoses, Dex**P**$(Man)_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 5% and 99.6% purity. R_t 4.2 min (Fig. S26) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (300 MHz, D_2O, δ): 7.62 (s, 1H), 7.38 (d, J = 10.4 Hz, 1H), 7.00 (s, 1H), 6.27 (d, J = 9.9 Hz, 1H), 6.07 (s, 1H), 4.77 (s, 2H), 4.43 (s, 8H), 4.13 (m, 13H), 3.87-3.44 (m, 12H), 3.23 (s, 2H), 3.15-2.90 (m, 16H), 2.84 (m, 2H), 2.26 (m, 2H), 2.03-1.11 (m, 70H), 0.93-0.62 (m, 36H). **MS** (ESI, H_2O): 1573 (18, $[M+2H+3TFA]^{2+}$), 1517 (20, $[M+2H+2TFA]^{2+}$), 1049 (22, $[M+3H+3TFA]^{3+}$), 1011 (90 $[M+3H+2TFA]^{3+}$), 974 (100, $[M+3H+TFA]^{3+}$), 938 (50, $[M+3H]^{3+}$). **HRMS** (ESI): Calcd for $C_{125}H_{222}FN_{37}O_{34}$: 1402.3371; found: 1402.3376 ($[M+2H]^{2+}$).

1.5.12 Synthesis of $CFArg_8$

For the preparation of the control peptide $CFArg_8$, the carboxyfluorescein was coupled in the amino group of a lysine Mtt localized at the end of the peptide sequence. Therefore, following the general protocol of the SPPS for growing the peptide, the Mtt protecting group was selectively removed by mechanically stirring the resin with a mixture of DCM/HFIP/TFE/TIS (6.5:2:1:0.5, 2 x 1 mL per 70 mg of resin) for 2 h. Finally, the mixture was filtered and the resin was washed with DCM (2 x 2 mL, 2 min) and DMF (2 mL, 20 min). Then, a solution of 5(6)-carboxyfluorescein (2 equiv) and *N*-HBTU (2 equiv) in DMF (1 mL) was added to the vessel followed by the drop wise addition of DIEA (4 equiv). The resulting mixture was shaken by bubbling Ar for 30 min and finally the filtered resin was washed with DMF (3 x 2 ml, 2 min) and DCM (3 x 2 ml, 2 min) (Fig. S10).

$CFArg_8$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 11% and 100% purity. R_t 3.2 min (Fig. S27) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (500 MHz, D_2O, δ): 8.16 (d, J = 1.8 Hz, 1H), 7.93 (t, J = 7.9 Hz, 1H), 7.58 (s, 1H), 7.36 (d, J = 7.9 Hz, 1H), 7.17-7.06 (m, 2H), 6.86-6.77 (m, 1H), 6.72 (td, J = 9.3 and 1.8 Hz, 1H), 6.60 (dd, J = 9.2 and 2.4 Hz, 1H), 4.24-3.86 (m, 9H), 3.32 (m, 2H), 3.12-2.87 (m, 16H), 1.88 (d, J = 10.2 Hz, 3H), 1.76-1.16 (m, 38H). **MS** (ESI, H_2O): 675 (50, $[M+3H+2TFA]^{3+}$), 638 (100, $[M+3H+TFA]^{3+}$), 599 (45, $[M+3H]^{3+}$), 477 (50, $[M+4H+TFA]^{4+}$), 450 (58, $[M+4H]^{4+}$), 360 (17, $[M+5H]^{5+}$). **HRMS** (ESI): Calcd for $C_{77}H_{125}N_{35}O_{16}$: 898.0008; found: 898.0016 ($[M+2H]^{2+}$).

1.5.13 Synthesis of $AcP(Biot)_2$

For the preparation of the biotinylated peptide, D-(+)-Biotin was coupled in the amino group of the lysines Mtt localized at the peptidic sequence. Therefore, following the general protocol of the SPPS for growing the peptide, the Mtt protecting group was selectively removed by mechanically stirring the resin with a mixture of DCM/HFIP/TFE/TIS (6.5:2:1:0.5, 2 x 1 mL per 70 mg of resin) for 2 h. Finally, the mixture was filtered and the resin was washed with DCM (2 x 2 mL, 2 min) and DMF (2 mL, 20 min). Then, a solution of D-(+)-Biotin (2 .5 equiv) and N-HATU (2.5 equiv) in DMF (1 mL) was added to the vessel followed by the drop wise addition of DIEA (4 equiv). The resulting mixture was shaken by bubbling Ar for 30 min and finally the filtered resin was washed with DMF (3 x 2 ml, 2 min) and DCM (3 x 2 ml, 2 min) (Fig. S11).

Ac**P**$(Biot)_2$ was obtained after RP-HPLC purification [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min), 95:5→5:95 (5→35 min)] with an overall yield of 11% and 100% purity. R_t 4.1 min (Fig. S28) [RP-HPLC Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)]. **^{1}H NMR** (500 MHz, D_2O, δ): 4.52-4.39 (m, 2H), 4.34-4.22 (m, 2H), 4.22-3.83 (m, 13H), 3.22-2.95 (m, 20H), 2.90-2.74 (m, 2H), 2.69-2.53 (m, 2H), 2.10 (s, 3H), 1.97-1.88 (m, 2H), 1.88-1.14 (m, 63H), 0.84-0.65 (m, 30H). **MS** (ESI, H_2O): 1362 (10, $[M+2H+4TFA]^{2+}$), 1306 (32,

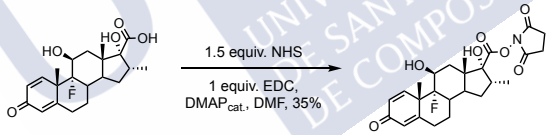
$[M+2H+3TFA]^{2+}$), 1249 (30, $[M+2H+2TFA]^{2+}$), 1191 (7, $[M+2H+TFA]^{2+}$), 870 (18, $[M+3H+3TFA]^{3+}$), 832 (80, $[M+3H+2TFA]^{3+}$), 795 (100, $[M+3H+TFA]^{3+}$), 756 (35, $[M+3H]^{3+}$), 624 (20, $[M+4H+2TFA]^{4+}$), 596 (25, $[M+4H+TFA]^{4+}$), 567 (40, $[M+4H]^{4+}$). **HRMS** (ESI): Calcd for $C_{100}H_{186}N_{38}O_{18}S_2$: 1135.7120; found: 1135.7119 ($[M+2H]^{2+}$).

1.6 Preparation of Dexamethasone labelled Concanavalin A

1.6.1 Synthesis of Dexamethasone-NHS

Dexamethasone-NHS was prepared by the reaction of Ox-Dex[S2] (0.106 mmol) with N-Hydroxysuccinimide (1.5 equiv) in DMF, using EDC (1 equiv) and DMAP (cat.) for 4 h. The DMF was removed by rotary evaporation and the oil obtained was dried. The compound was washed with 1 M solution of HCl and after that with a saturated solution of $NaHCO_3$. Purification by silica gel column chromatography (DCM/MeOH, 99:1) provided final product with an overall yield of 35%.

^{1}H NMR (500 MHz, $CDCl_3$, δ): 7.22 (d, *J* = 10.1 Hz, 1H), 6.32 (dd, *J* = 10.1, 1.9 Hz, 1H), 6.12 (s, 1H), 4.37 (d, *J* = 8.1 Hz, 1H), 3.22 (s, 1H), 3.12-2.97 (m, 1H), 2.84 (s, 4H), 2.63 (td, *J* = 13.8, 13.0, 6.1 Hz, 1H), 2.49-2.17 (m, 4H), 1.87-1.72 (m, 3H), 1.63-1.56 (m, 1H), 1.55 (s, 3H), 1.35-1.28 (m, 2H), 1.22 (s, 3H), 1.07 (d, *J* = 7.1 Hz, 3H). **^{13}C NMR (**500 MHz, $CDCl_3$, δ): 186.56 (q), 168.40 (q), 166.04 (q), 152.12 (CH), 129.81 (CH), 125.13 (CH), 101.02 (q), 99.62 (q), 86.46 (q), 71.99 (CH), 48.62 (q), 48.17 (q), 42.96 (CH), 37.01 (CH), 35.88 (CH_2), 34.32 (CH), 32.35 (CH_2), 30.99 (CH_2), 29.70 (CH_2), 25.67 (CH_2), 22.90 (CH_3), 16.70 (CH_3), 14.57 (CH_3). (Fig. S29). **HRMS** (ESI): Calcd for $C_{25}H_{31}FNO_7$: 476.2082; found: 476.2079.



1.6.2 Synthesis of ConA-Dex

ConA-Dex was prepared by the reaction of Concanavalin A (3 mg/mL) with Dexamethasone-NHS (50 equiv) in $NaHCO_3$ buffer (pH 9) at room temperature for 2 h. The resulting solution was dialysed overnight in HKR buffer. To quantify the number of dexamethasones per Concanavalin A the absorbance of the mixture was measured using a NanoDrop UV-Vis Spectrophotometer. Absorbance values at 242 nm (maximum absorbance for dexamethasone) and 280 nm (maximum absorbance for the protein) were measured for various dilutions of the dialysed ConA-Dex, and several known concentrations of ConA or dexamethasone. Data in the linear range were analysed using the following system of equations:

$$A_{242}^{\text{ConA-Dex}} = x \cdot A_{242}^{\text{ConA}} + y \cdot A_{242}^{\text{Dex}}$$

$$A_{280}^{\text{ConA-Dex}} = x \cdot A_{280}^{\text{ConA}} + y \cdot A_{280}^{\text{Dex}}$$

Where the values *x* and *y* were used to estimate the concentration of each component in ConA-Dex by multiplying the concentration of pure ConA or dexamethasone by *x* or *y* respectively, and the extent of labelling was calculated as the ratio between dexamethasone and ConA concentrations. A value of around 16 Dex per ConA tetramer was obtained.

1.7 General Procedure for Circular Dichroism

Circular dichroism measurements were carried out using the following settings: acquisition range: 300-190 nm; band width: 1.0 nm; accumulation: 3 scans; data pitch: 1 nm; CD scale 200 mdeg/1.0 dOD; D.I.T. (Data Integration Time): 1 s; scanning mode: continuous; scanning speed: 200 nm/min. Measurements were done from 10 ºC to 60 ºC (data interval: 10 ºC; temp. gradient 5 ºC/min) in a quartz cell of 0.2 cm path length at a final volume of 0.5 mL (HKR buffer or TFE) with a final peptide concentration of 200 μM.

For the measurements in liposomes, samples were prepared by drying under reduced pressure L-α-phosphatidylcholine (8.5 μL, 100 mg/mL solution in $CHCl_3$) and peptide (325 μL, 200 μM) in TFE to obtain a ratio lipid/peptide of 18:1. Lipids were suspended in HKR buffer (650 μL) and sonicated for 45 min until a clear solution was obtained. Spectra were recorded in a 0.2 cm path length quartz cell.

The results are expressed as the mean residue molar ellipticity $[\theta]_{MRt}$ with units of degrees·cm^2·dmol^{-1} and calculated using the equation S1,

$$[\theta]_{MRt} = \frac{100 \cdot \theta}{C \cdot l \cdot No.of\ residues} \quad \text{(S1)}$$

where θ is the ellipticity (mdeg), C is the peptide concentration (M) and l is the cell path length (cm).

1.8 Cells Lines and Culture

HeLa cells were incubated at 37 ºC/ 5% CO_2/ 95% humidity in an INCO 108 incubator (Memmert) with Dulbecco's Modified Eagle's Medium (4500 mg/L glucose, *L*-glutamine, sodium pyruvate and sodium bicarbonate; Sigma-Aldrich), supplemented with 10% fetal bovine serum (Sigma-Aldrich) and 1% of Penicillin-Streptomycin-Glutamine Mix (Fisher).

1.9 Cell transport experiments in HepG2 cells

Internalization experiments in HepG2 cells were done as described for HeLa cells (see Methods in the manuscript). In the case of streptavidin, incubation time was increased to 5 h.

1.10 Cell viability: MTT Assay

Cell viability was established by a standard MTT assay (Fig. S7 and S8). One day before the assay, a suspension of HeLa or HepG2 cells was plated in 96-well tissue culture plates (Costar 96 Flat Bottom Transparent Polystyrol) by adding 100 μL (150.000 cells/mL) per well. The next day, the medium was aspirated and cells were incubated with different concentrations of peptide and peptide/protein complexes diluted in HKR (50 μL/well). After

30 min of incubation at 37 ºC, the medium was aspirated and replaced by fresh medium (DMEM) containing 10% FBS (100 µL). Control cells were given only cell culture medium (100 µL final medium). The viability was measured by quantifying the cellular ability to reduce the water-soluble tetrazolium dye 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl tetrazolium bromide (MTT) to its insoluble formazan salt as follows. MTT (5 mg/mL in PBS, 10 µL/well) was added to the wells and the cells were further incubated for 6 h. The supernatant was carefully removed and the water-insoluble formazan salt was dissolved in DMSO (100 µL/well). The absorbance was measured at 570 nm using a microplate reader (Infinite F200pro, Tecan). Data points were collected in triplicate and expressed as normalized values for untreated control cells (100%).

1.11 Propidium Iodide Assay

Cell viability was established by a standard PI assay (Fig. S7 and S8). One day before the assay, a suspension of HeLa or HepG2 cells was plated in 96-well tissue culture plates (Costar 96 Flat Bottom Transparent Polystyrol) by adding 100 µL (150.000 cells/mL) per well. The next day, the medium was aspirated and cells were incubated with different concentrations of peptide and peptide/protein complexes diluted in DMEM (50 µL/well). After 30 min of incubation at 37 ºC, the medium was aspirated; cells were washed with fresh medium and then 100 µL of trypsin were added in each well and cells were incubated for 15 min at 37 ºC. After this, 100 µL of a solution of 2% FBS and 5 mM EDTA in PBS containing PI (0.25 µg/mL) were added and the plate was incubated for 5 min in the dark. Samples were analysed on a Guava EasyCyte^{TM} cytometer. Propidium iodide was measured by excitation at 532 nm and collecting emission at 695/50 nm. Data points were collected in triplicate and analysis was performed with InCyte software included in GuavaSoft 3.2 (Millipore).

1.12 Quantification of the uptake by fluorometry

For internalization assays, a modification of the quantification of CPP uptake by fluorometry protocol described in Holm *et al.*[S5] was used. Briefly, semiconfluent monolayers of HeLa cells seeded the day before were washed twice with HKR before incubation with 4 µM of TAMRA labelled peptides, with Tm**P**$(Man)_2$/ConA or with Ac**P**$(Man)_2$/$ConA_{FITC}$ complexes (previously prepared by co-incubation for 7 min rt) diluted in the same buffer for 30 min at 37 ºC (unless otherwise stated). Then, cells were washed twice with HKR and trypsinized, and after the addition of 1 mL of HKR, centrifuged for 5 min at 1000 g. The resultant pellets were lysed by incubation in aqueous solution of NaOH (0.1 M) at 4 ºC for 60 min and frozen until analysis.

Where indicated, cells were pretreated for 30 min before the incubation of the peptide and during the incubation with the following inhibitors: Wortmannin (200 nM), chlorpromazine (30 µM), methyl-ß-cyclodextrin (5 mM), EIPA (50 µM), ammonium chloride (50 mM), chloroquine (100 µM) or heparin (5 µg/mL). For the incubation at low temperature, cells were placed on ice 15 min before the incubation with the peptides, and ice-cold solutions were used for the washes and incubation.

Fluorescence levels of the lysates were measured using a microplate reader (Infinite F200Pro, Tecan), both for TAMRA (λ_{ex} 560 nm, λ_{em} 610 nm) and fluorescein (λ_{ex} 485 nm, λ_{em} 535 nm) labelled compounds, and the concentration of the peptides or ConA was calculated by comparison to a standard curve. After neutralization with a volume of an aqueous solution of HCl (0.1 M), protein concentration in the lysates was determined using a Coomassie (Bradford) protein assay kit following the manufacturer instructions, and measuring the absorbance at 570 nm (Infinite F200Pro, Tecan). The uptake was calculated as pmol of peptide/mg of total protein, and normalized to the uptake in untreated cells in the case of the inhibition studies.

1.13 Glucocorticoid assay

The ability of the peptide to reach the cytosol was determined with a glucocorticoid induced GFP translocation assay (GIGT)[S6]. In this assay, a GFP protein fused to the steroid binding domain of the glucocorticoid receptor (GR) accumulates in the nucleus of the cell in response to dexamethasone binding, so the translocation ratio (the ratio between the nuclear and cytoplasmic fluorescence) can be used as an indicator of the presence of dexamethasone-labelled peptides in the cytosol (Fig. S9).

HeLa cells grown in four chamber glass bottom dishes were transfected with the plasmid pK7-GR-GFP (a gift from Ian Macara[S7], Addgene plasmid #15534) using Lipofectamine 2000 and, 24 h post-transfection, cells were washed with HKR and incubated for 30 min with Hoechst 33342 (1 μM), and endocytosis inhibitors where indicated, at the following concentrations: chlorpromazine (50 μM), methyl-ß-cyclodextrin (5 mM), Wortmannin (200 nM), chloroquine (100 μM), NH_4Cl (50 mM), heparin (5 μg/mL) or EIPA (50 μM). Cells were then incubated for 1 h with Dex**P**$(Man)_2$ (4 μM), in the presence or absence of unlabelled ConA (30 nM), with dexamethasone (1 μM) (as positive control) or just with HKR (as negative control) and immediately imaged. To study the cytosolic release of ConA-Dex, cells were incubated with 7 nM of Dex-labelled ConA (equivalent to around 100 nM Dexamethasone) previously incubated with 4 μM of Ac**P**$(Man)_2$. As controls, cells were incubated with the same amount of ConA-Dex in the absence of peptide or after incubation for 20 min with 1 μg of Lipofectamine 2000. Twenty to thirty images of each sample were acquired with an Andor Zyla 4.2 digital camera mounted on a Nikon Eclipse Ti-E microscope at 60x magnification and the translocation ratio (the ratio of the median intensities of GFP in the nucleus and in the 2 μm wide surrounding region) was calculated with CellProfiler[S8] as follows. Nuclei were identified as Hoechst stained objects using the three-class thresholding Otsu method and the cytoplasmic region was defined as the 2 μm surrounding area. To ensure a better separation of the cytoplasmic and nuclear region, nuclei were shrunk 0.5 μm before measuring object intensity. Cells falling below the 20% of the maximum intensity of the image were considered untransfected and discarded for the analysis. A total of 40 to 80 cells were analyzed for each sample.

Statistical analysis of the data was performed with R software[S9]. Results were subjected to pairwise two-tailed Student's t-test and p-values were adjusted using Bonferroni's correction.

1.14 Time-lapse videos

For time-lapse experiments, HeLa cells incubated with the peptide/protein complexes were placed in an incubator chamber at 37 ºC on a Nikon Eclipse Ti-E inverted microscope. The acquisition of the images was controlled with NIS-Elements software (Nikon), using the indicated time intervals. Videos were assembled using ImageJ.

1.15 Gel electrophoresis

To study protein degradation, HeLa cells were incubated with the mixtures of ConA (FITC- or biotin-labelled) and Ac**P**$(Man)_2$ for 30 min in HKR at 37 ºC. Cells were washed with HKR, trypsinized and centrifuged for 5 min at 1000 g, 4 ºC. Pellets were re-suspended in Laemmli buffer 1x (60 mM Tris-HCl pH 6.8, 2 % SDS, 10 % glycerol, 5 % 2-mercaptoethanol, 0.0010 % bromophenol blue) and boiled for 5 min. Proteins were separated on 12.5 % SDS-PAGE gels.

Gels containing FITC-labelled protein were directly imaged under UV light in a Gel Doc system (Bio-Rad), and then stained with Coomassie blue to ensure similar amounts of protein were loaded in each lane. For the detection of the biotinylated protein, proteins were transferred to a PVDF membrane (Immobilon-P, Millipore), blocked with 5 % BSA in PBS with 0.05 % Tween-20 for 1 h and incubated with 4 μg/mL of $Streptavidin_{488}$ diluted in blocking solution. After several washes, bands were detected under UV light using a Gel Doc system.

1.16 Atomic Force Microscopy (AFM).

Standard AFM measurements (Fig. S15) were conducted in ambient atmosphere at room temperature. Shortly before mica deposition, samples were prepared by incubating a 2 μM and 4 μM solution of Ac**P**$(Man)_2$ with a 30 nM solution of ConA in HKR buffer, for 10 min. For AFM imaging, 10 μL of the sample were dropped onto freshly exfoliated mica and after 10 min the mica was thoroughly washed with Milli-Q water, and dried under nitrogen flow.

1.17 Supporting Figures for Characterization

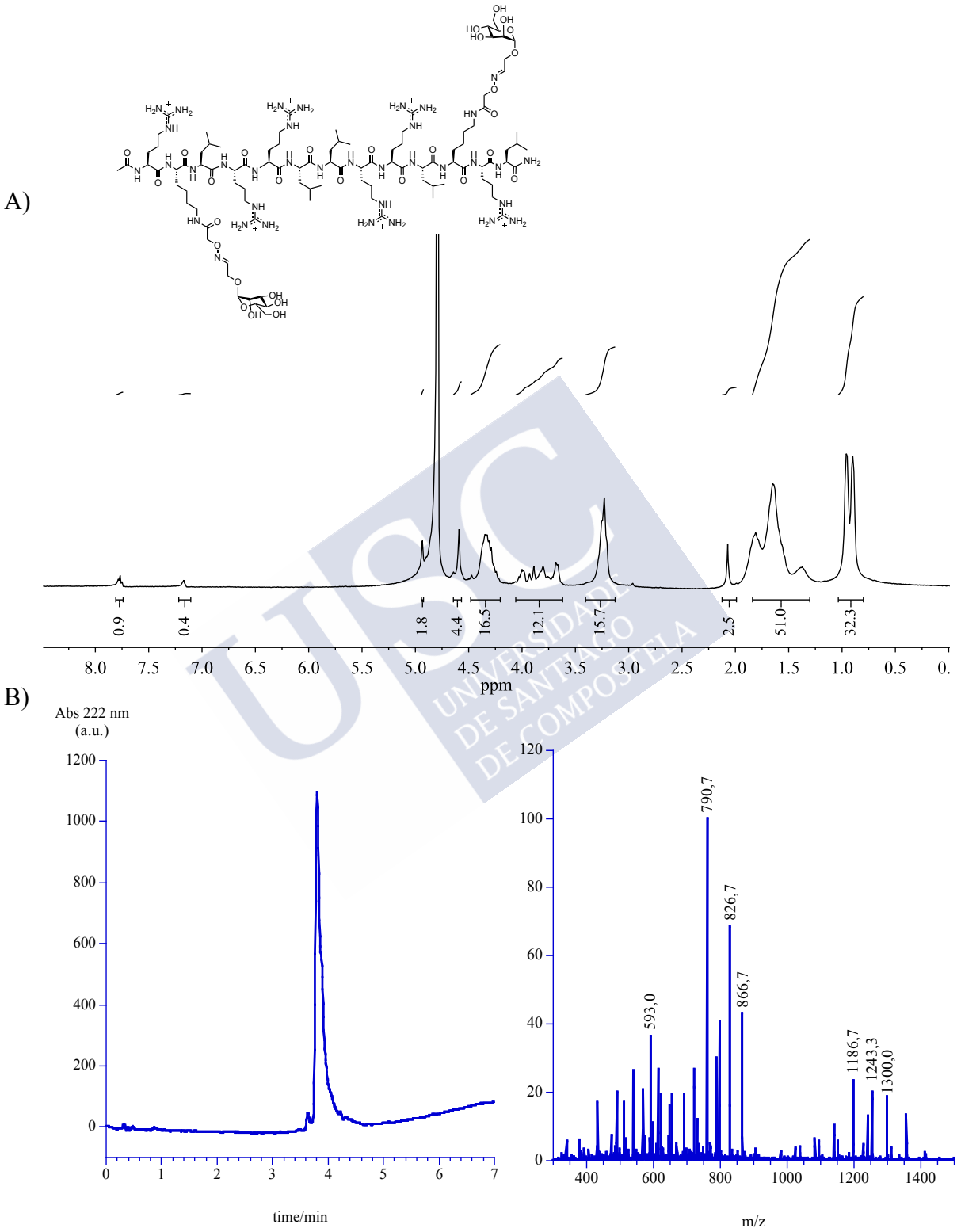


Figure S16. A) ^{1}H-NMR spectra in D_2O of Ac**P**$(Man)_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 3.8 min) and ESI-MS for Ac**P**$(Man)_2$.

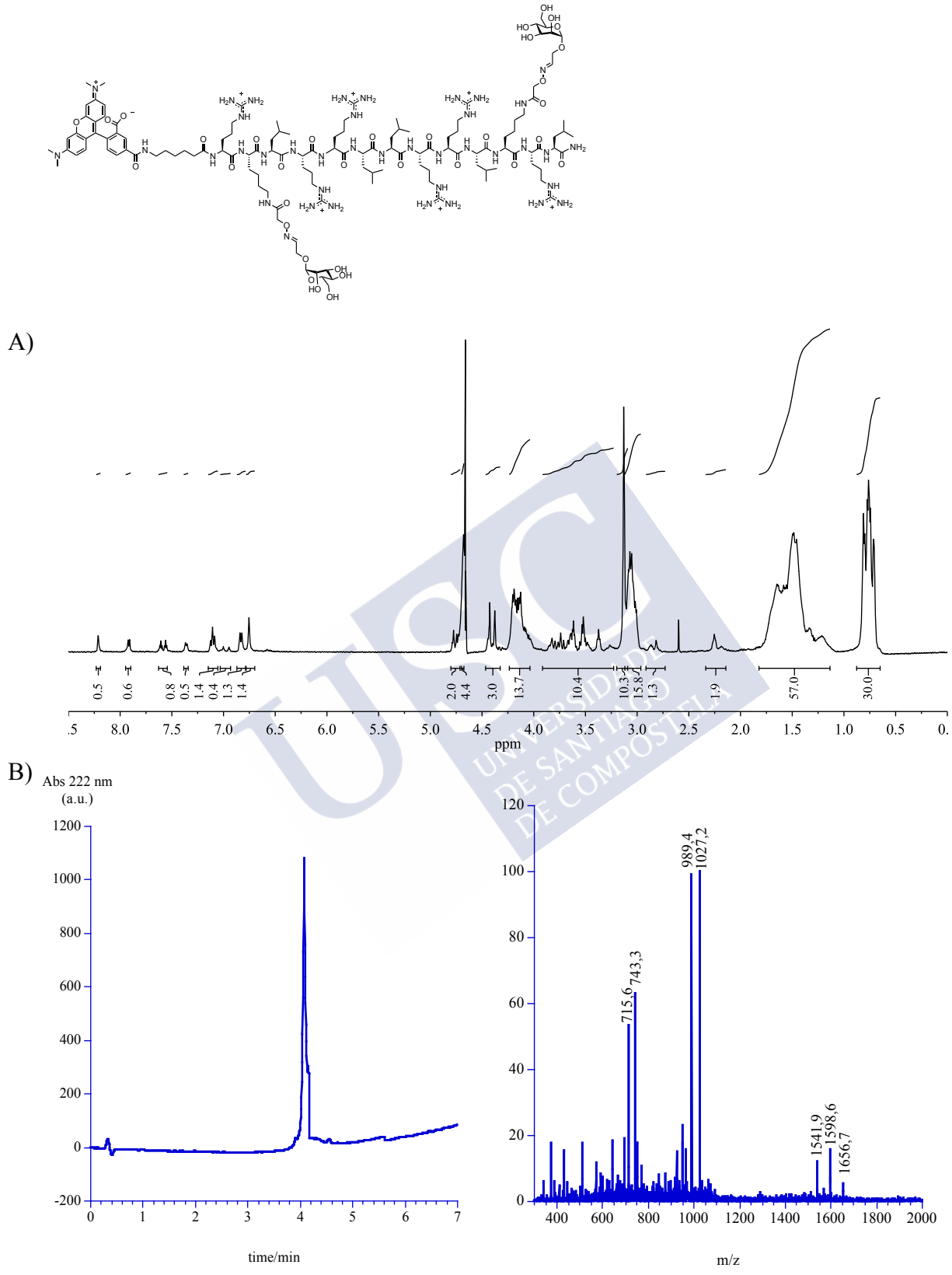


Figure S17. A) ^{1}H-NMR spectra in D_2O of Tm**P**$(Man)_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 4.1 min) and ESI-MS for Tm**P**$(Man)_2$.

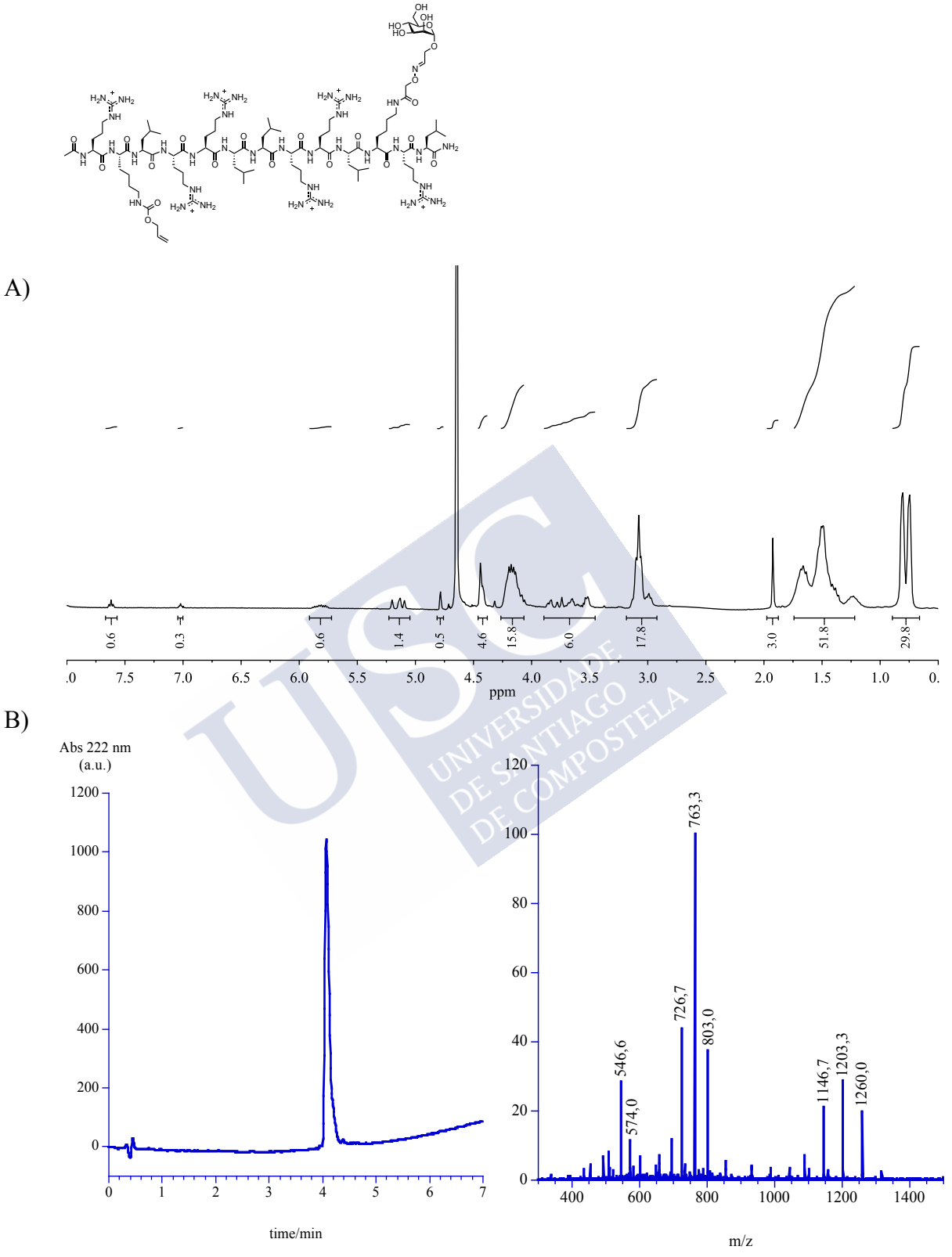


Figure S18. A) ^{1}H-NMR spectra in D_2O of AcP(Alloc)(Man). B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 4.1 min) and ESI-MS for AcP(Alloc)(Man).

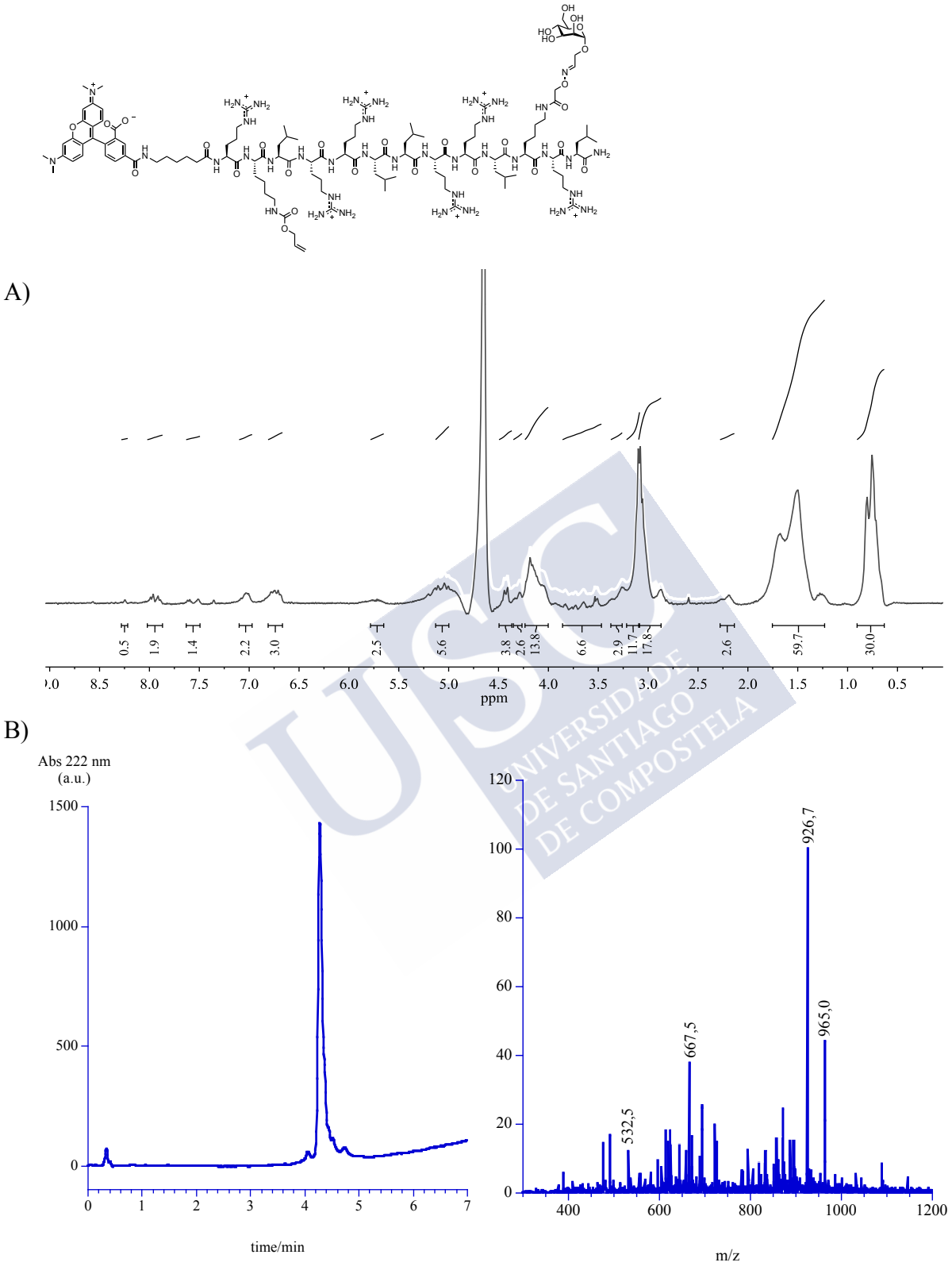


Figure S19. A) ^{1}H-NMR spectra in D_2O of Tm**P**(Alloc)(Man). B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 4.3 min) and ESI-MS for Tm**P**(Alloc)(Man).

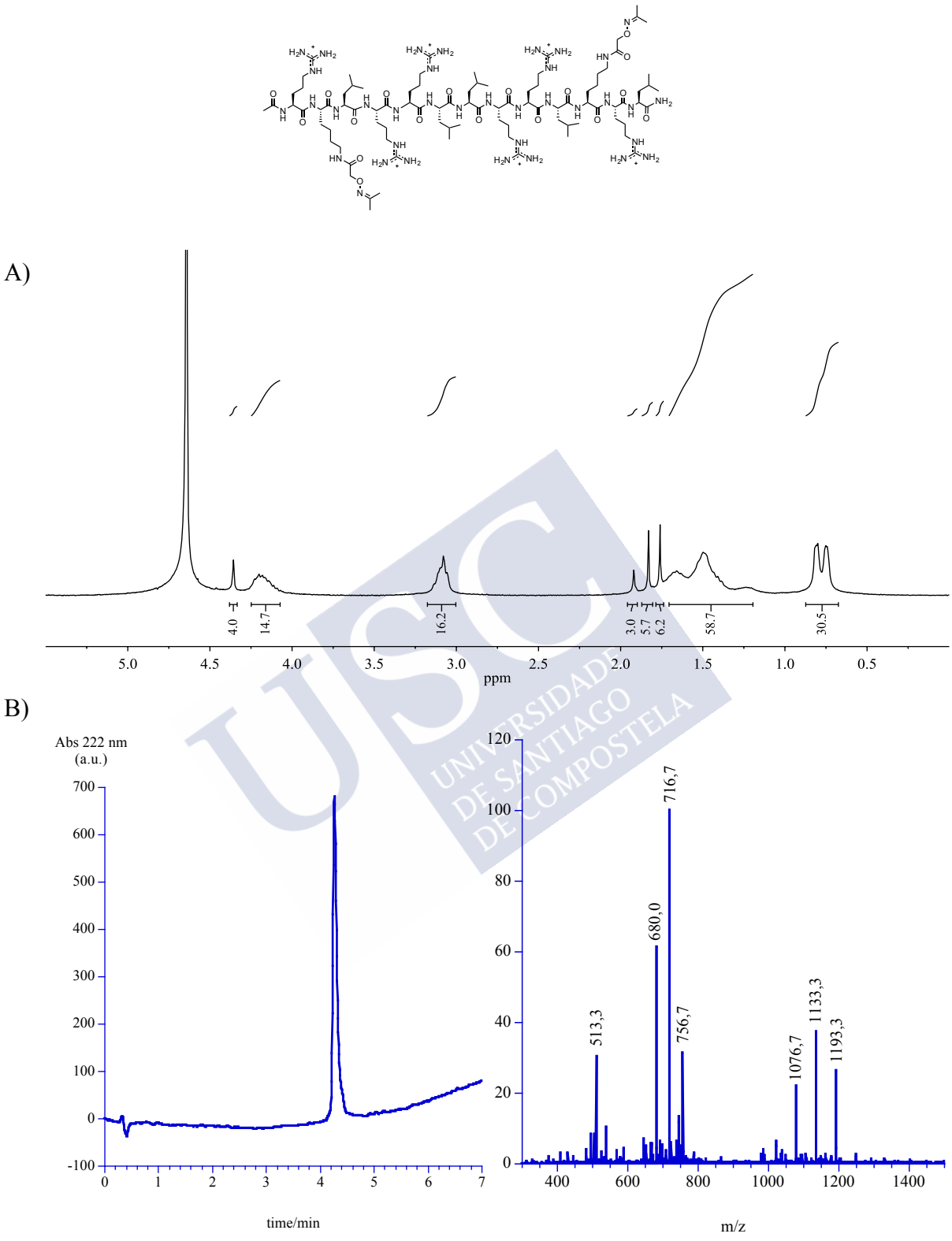
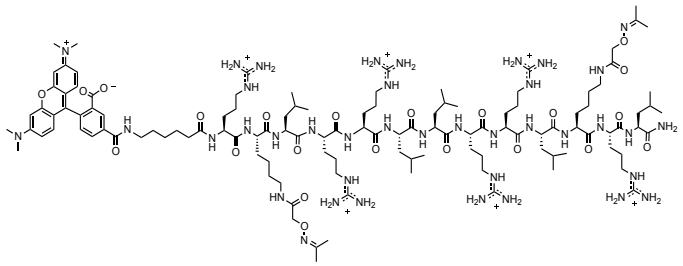


Figure S20. A) ^{1}H-NMR spectra in D_2O of Ac**P**(Acetone)$_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 4.3 min) and ESI-MS for Ac**P**(Acetone)$_2$.

A)



B)

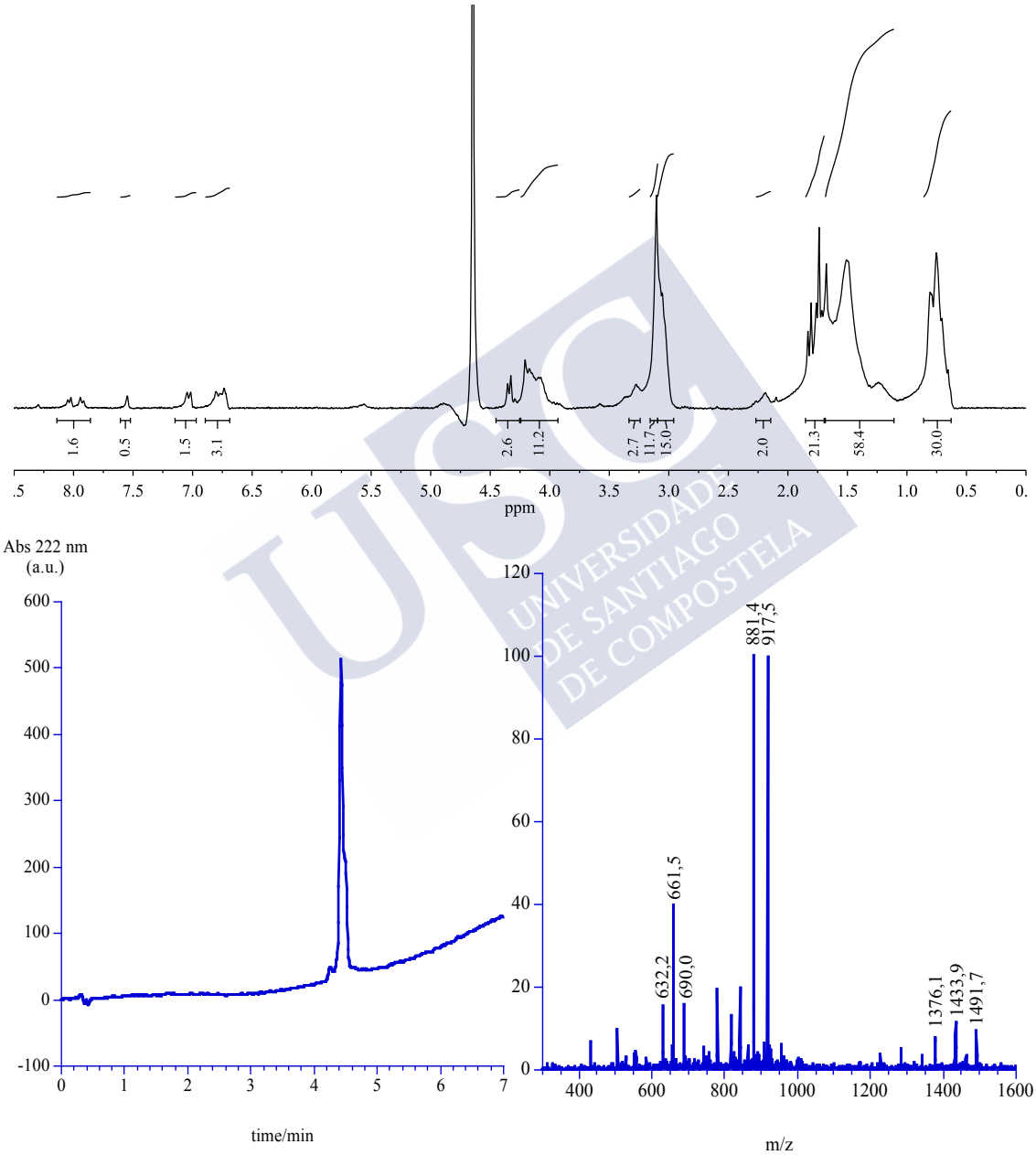


Figure S21. A) ^{1}H-NMR spectra in D_2O of Tm**P**(Acetone)$_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 4.4 min) and ESI-MS for Tm**P**(Acetone)$_2$.

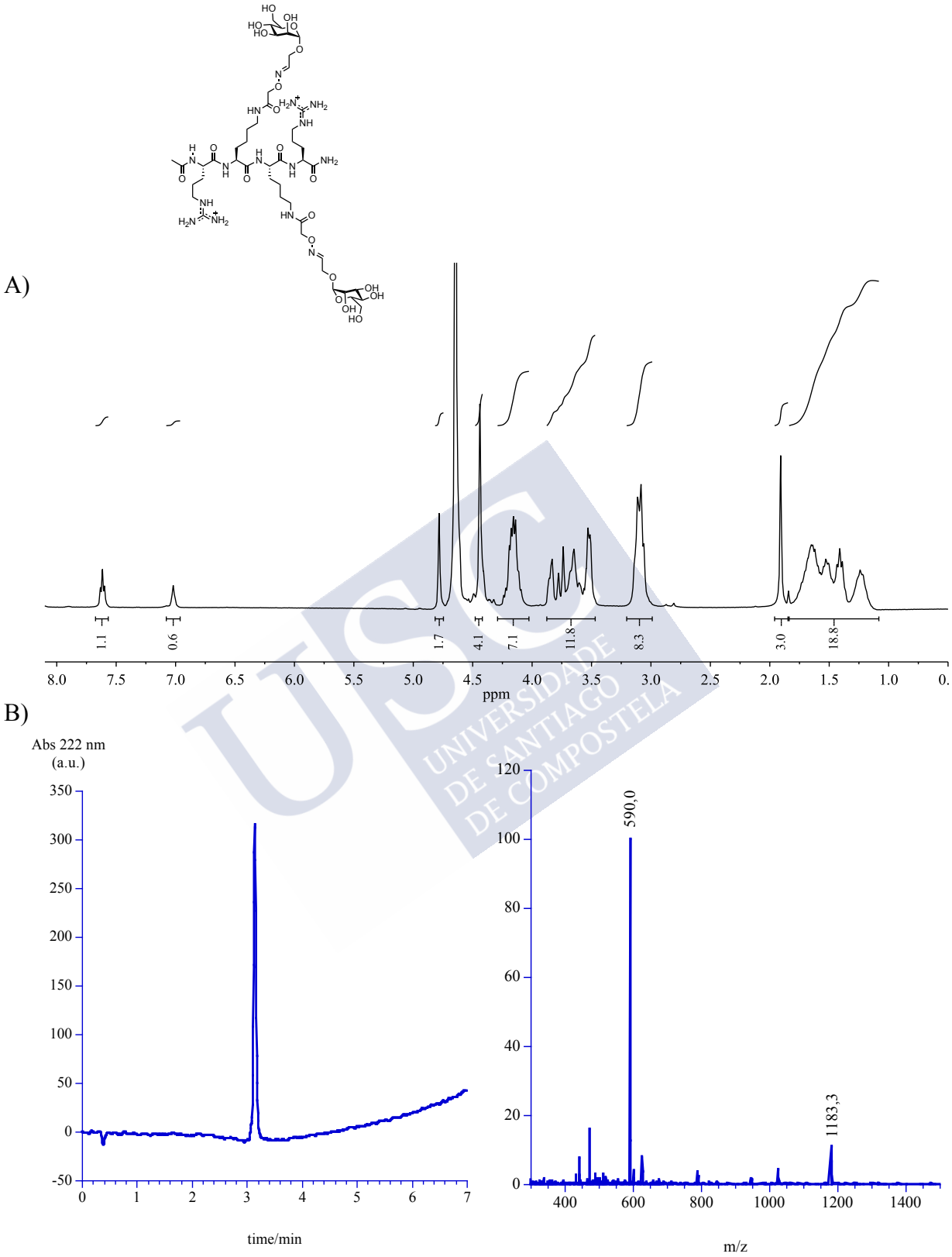


Figure S22. A) ^{1}H-NMR spectra in D_2O of $AcArg_2(Man)_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 3.1 min) and ESI-MS for $AcArg_2(Man)_2$.

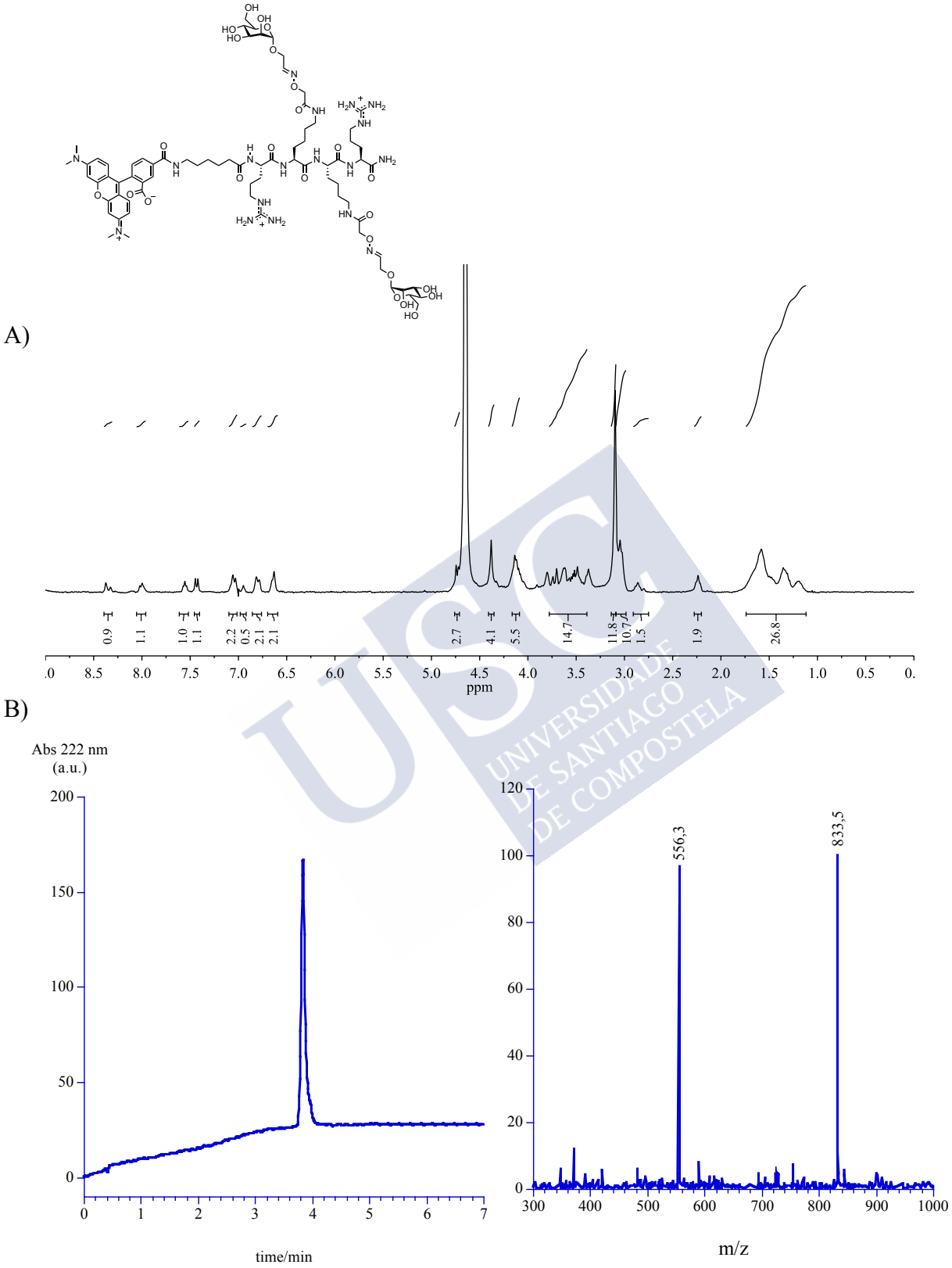


Figure S23. A) ^{1}H-NMR spectra in D_2O of $TmArg_2(Man)_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 3.8 min) and ESI-MS for $TmArg_2(Man)_2$.

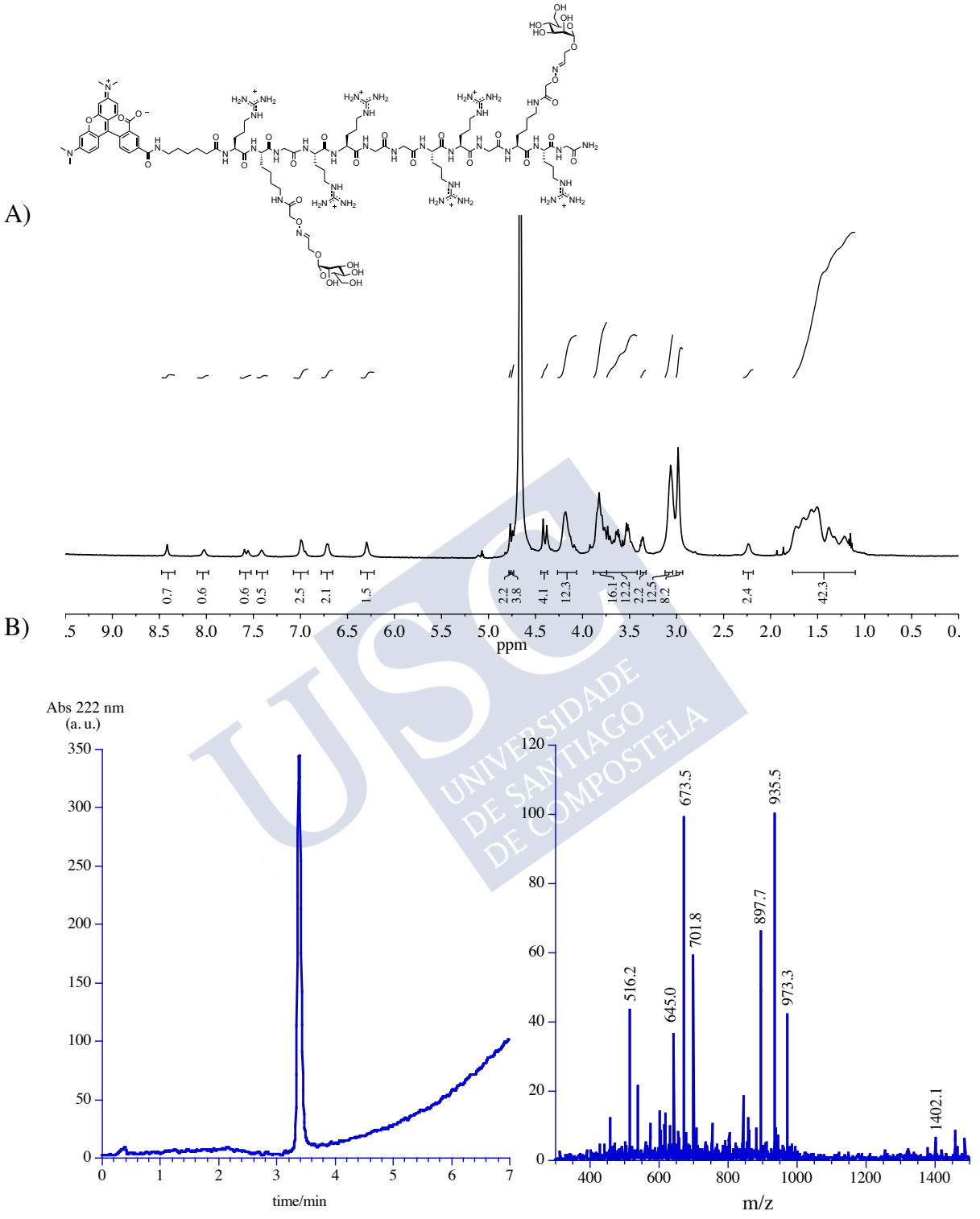


Figure S24. A) ^{1}H-NMR spectra in D_2O of $TmArg_6Gly_5(Man)_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 3.4 min) and ESI-MS for $TmArg_6Gly_5(Man)_2$.

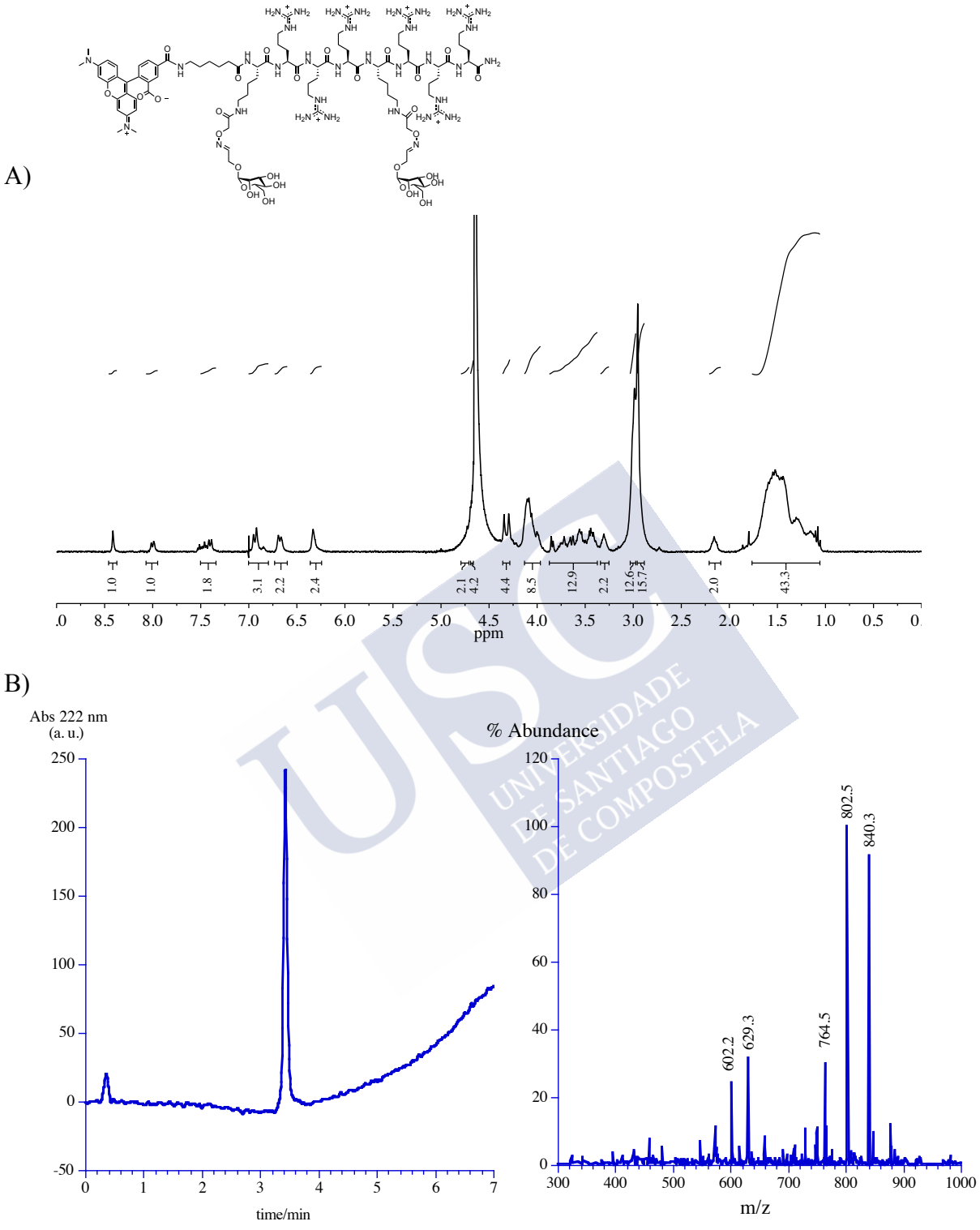


Figure S25. A) ^{1}H-NMR spectra in D_2O of $TmArg_6(Man)_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 3.4 min) and ESI-MS for $TmArg_6(Man)_2$.

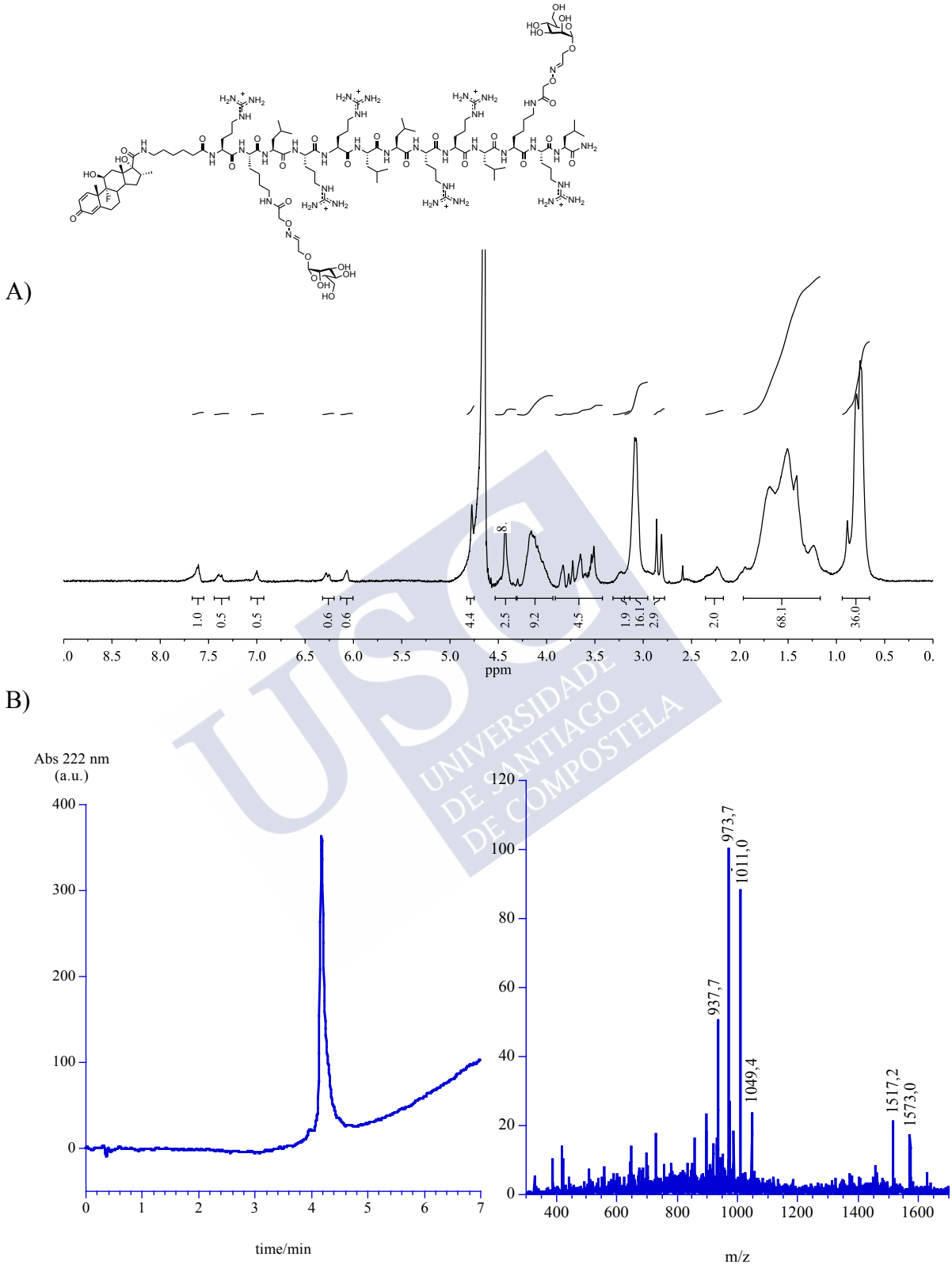


Figure S26. A) ^{1}H-NMR spectra in D_2O of DexP(Man)$_2$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 4.2 min) and ESI-MS for DexP(Man)$_2$.

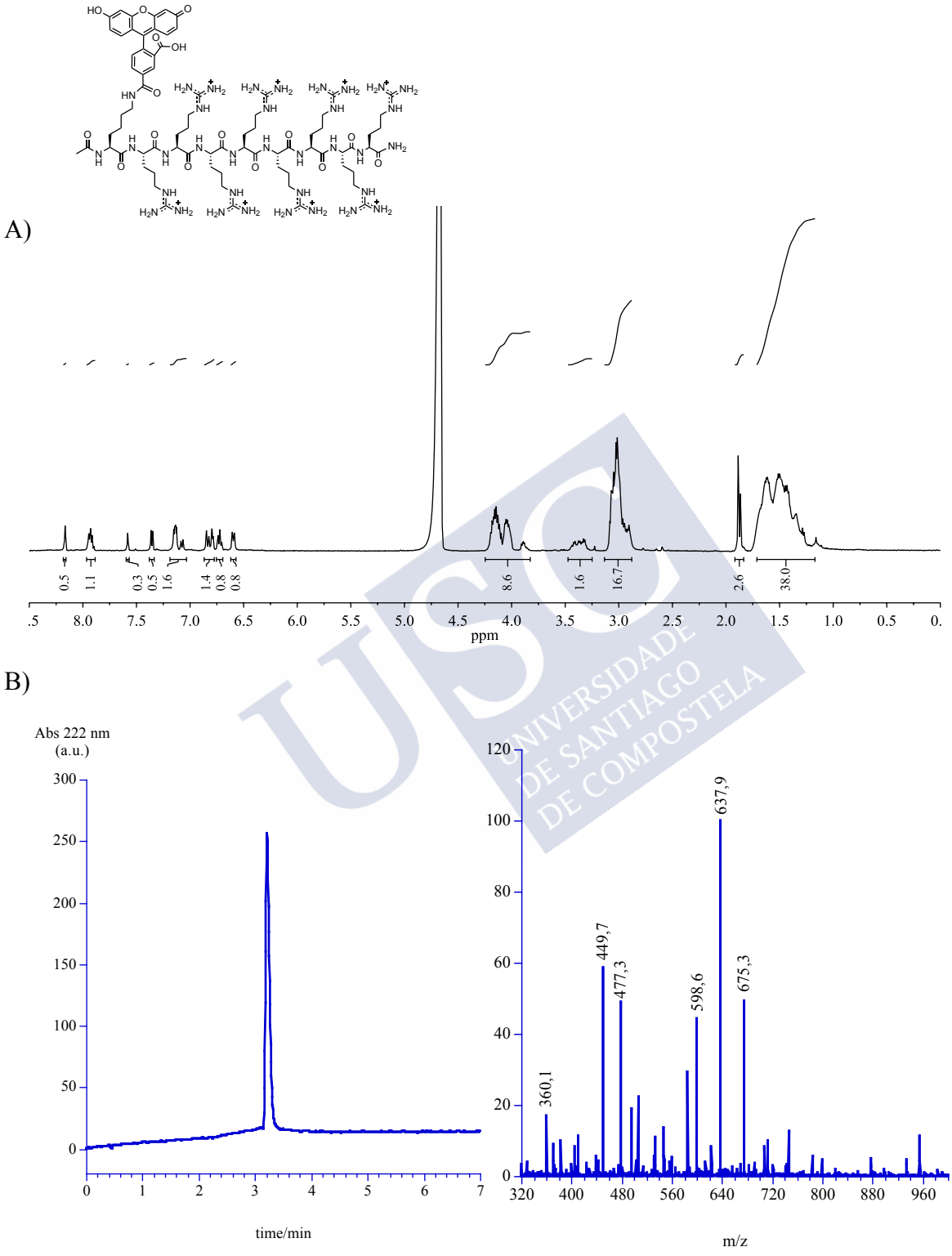


Figure S27. A) ^{1}H-NMR spectra in D_2O of $CFArg_8$. B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 3.2 min) and ESI-MS for $CFArg_8$.

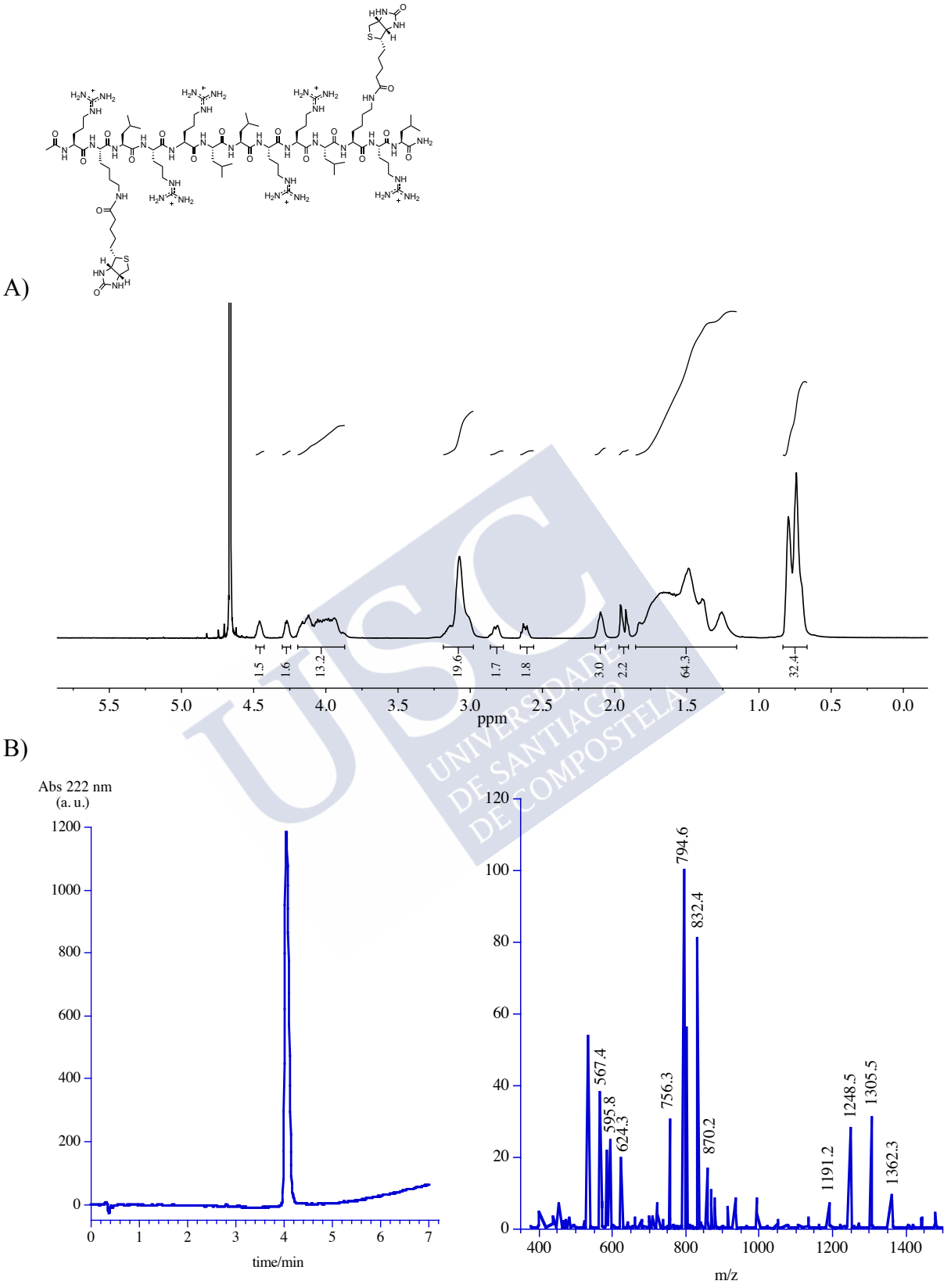


Figure S28. A) ^{1}H-NMR spectra in D_2O of AcP(Biot)$_2$ B) RP-HPLC [Agilent SB-C18 column, H_2O (0.1% TFA)/ CH_3CN (0.1% TFA) 95:5→5:95 (0→5 min)] (R_t 4.1 min) and ESI-MS for AcP(Biot)$_2$.

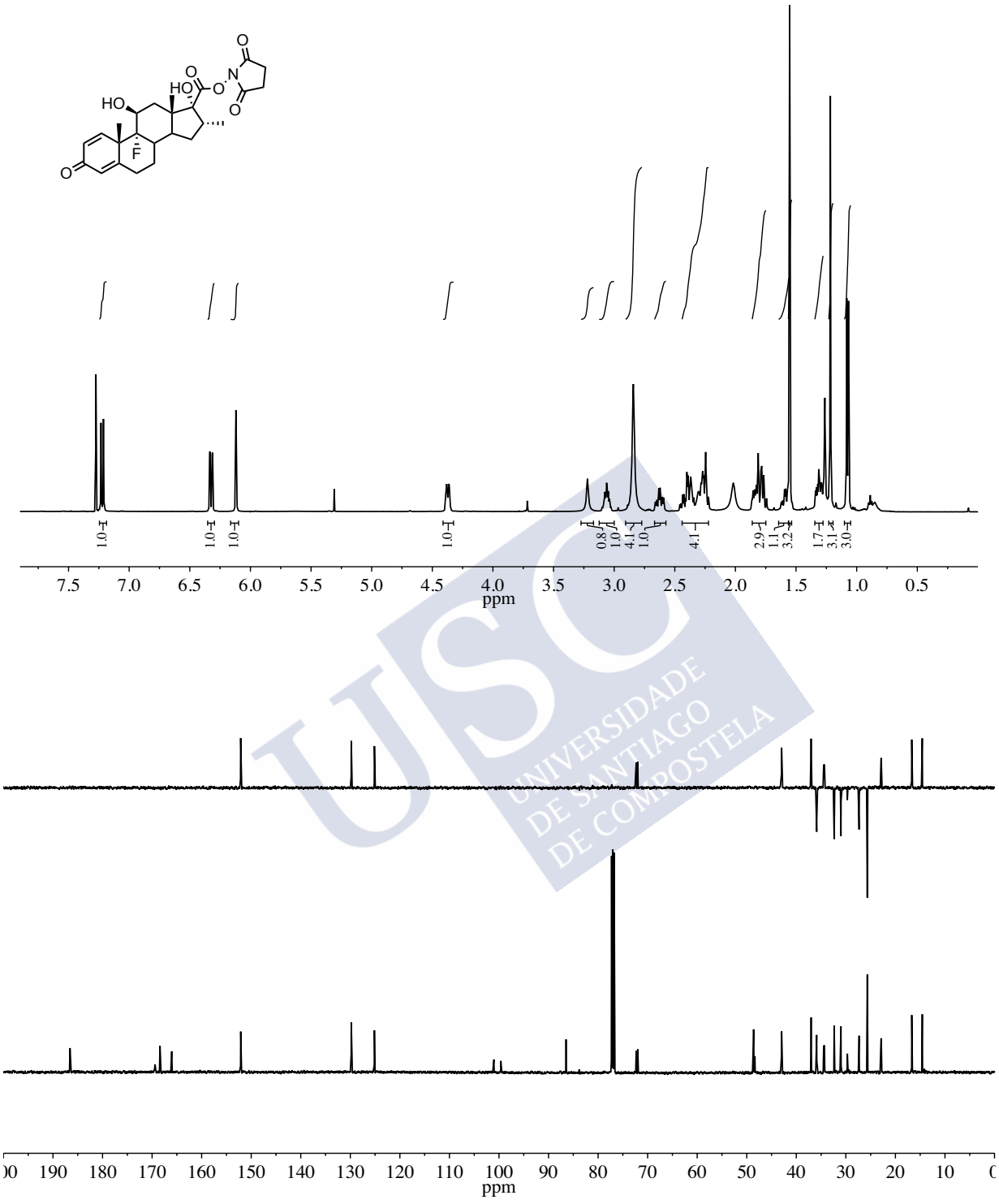


Figure S29. ^{1}H-NMR, DEPT and ^{13}C-NMR spectra in $CDCl_3$ of DexNHS.

1.18 Supporting References

[S1]W. Yao, Y. Jiao, J. Luo, M. Du, L. Zong, Int. J. Biol. Macromol. 2012, 50, 821–825.
[S2]M. V Govindan, B. Manz, 1980, 53, 47–53.
[S3]R. Behrendt, P. White, J. Offer, J. Pept. Sci. 2016, 22, 4–27.
[S4]C. Kay, O. E. Lorthioir, N. J. Parr, M. Congreve, S. C. McKeown, J. J. Scicinski, S. V Ley, Biotechnol. Bioeng. 2000, 71, 110–8.
[S5]T. Holm, H. Johansson, P. Lundberg, M. Pooga, M. Lindgren, U. Langel, Nat. Protoc. 2006, 1, 1001–5.
[S6]J. M. Holub, J. R. Larochelle, J. S. Appelbaum, A. Schepartz, Biochemistry 2013, 52, 9036–46.
[S7]K. L. Carey, S. A. Richards, K. M. Lounsbury, I. G. Macara, J. Cell Biol. 1996, 133, 985–96.
[S8]A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, et al., Genome Biol. 2006, 7, R100.
[S9]R Core Team, R: A Language and Environment for Statistical Computing, 2015.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2 PEPTIDE/CAS9 NANOSTRUCTURES FOR RIBONUCLEOPROTEIN CELL MEMBRANE TRANSPORT AND GENE EDITION

Irene Lostalé-Seijo[†], Iria Louzao[†], Marisa Juanes[†] & Javier Montenegro[†]*

[†] Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain. *e-mail: javier.montenegro@usc.es

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

2.1 Abbreviations

crRNA: CRISPR RNA; DIEA: N,N-Diisopropylethylamine; DMEM: Dulbecco's modified Eagle medium; DMF: N,N-Dimethylformamide; DMSO: Dimethylsulfoxide; EDTA: Ethylenediaminetetraacetic acid; EGFP: enhanced *green* fluorescent protein; EIPA: 5-(N-Ethyl-N-isopropyl)amiloride; FBS: fetal bovine serum; Fmoc: 9-fluorenylmethoxycarbonyl; GR: glucocorticoid receptor; gRNA: guide RNA; HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HFIP: 1,1,1,3,3,3-Hexafluoro-2-propanol; HKR: HEPES-Krebs-Ringer buffer; HPRT1: hypoxanthine phosphoribosyltransferase 1; hpt: hours post-treatment; IPTG: Isopropyl β-D-1-thiogalactopyranoside; MTT: 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl tetrazolium bromide; MβCD: Methyl-β-Cyclodextrin; *N*-HBTU: N-[(1H-Benzotriazol-1-yl)4-(dimethylamino)methylene]-N-methylmethanaminium hexafluorophosphate N-oxide; Ox-Dex: Dexamethasone Acid; PAM: protospacer adjacent motif; PBS: Phosphate buffered saline; RNP: ribonucleoprotein; T7E1: T7 endonuclease I; TAE: tris-acetate-ethylenediaminetetraacetic acid; TALENs: transcription activator- like effector nucleases; TAMRA: [5(6)-Carboxytetramethylrhodamine]; TFE: Trifluoroethanol; TIS: Triisopropylsilane; tracrRNA: *trans*-activating crRNA; ZFNs: Zinc-finger nucleases.

2.2 Materials and Methods

Polystyrene Fmoc protected Rink Amide resin (load 0.71 mmol/g) and Fmoc-L-Lys(Mtt)-OH were purchased from Iris. Fmoc-L-Leu-OH, glutaric anhydride, *tert*-butyl carbazate, triisopropylsilane (TIS), Diisopropylethyl amine (DIEA) were obtained from Sigma-Aldrich. Fmoc-L-Arg(Pbf)-OH was purchased from Carbosynth, Trifluoroethanol (TFE) and 1,1,1,3,3,3-Hexafluoroisopropanol (HFIP) were purchased from TCI. Peptide synthesis grade *N,N*-dimethylformamide was purchased from Scharlau. All other chemicals were purchased from Sigma-Aldrich, TCI or Fisher. 5 nm Ni-NTA-Nanogold® was purchased from Nanoprobes and used according to the manufacturer recommendations.

The aldehydes tested were either commercially available or synthesized following reported protocols from the corresponding alcohols[S1]. 5-(2-(*tert*-butoxycarbonyl)hydrazineyl)-5-oxopentanoic acid was prepared as previously described[S2,S3].

Purified Cas9 protein (Alt-R™ S.p. Cas9 Nuclease 3NLS), Alt-R™ CRISPR-Cas9 crRNA, unlabelled and $ATTO_{550}$ labelled Alt-R™ CRISPR-Cas9 tracrRNA, and oligonucleotides for PCR were purchased from Integrated DNA Technologies. GeneArt Genomic Cleavage Detection kit, Lipofectamine 2000, LysoTracker Green DND-26, LysoTracker Deep Red, Hoechst 33342, HisPur Ni-NTA resin and cell culture reagents were purchased from ThermoFisher. Endonuclease T7E1 was from New England Biolabs. Agarose D1 with medium EEO was bought from Laboratorios Conda, Agarose (low melting point) was from Promega and Protogel (30 %; 37.5:1 Acrylamide to Bisacrylamide Stabilized Solution) from National Diagnostics. Wortmannin was purchased from Fluorochem, chlorpromazine hydrochloride from TCI Chemicals, methyl-β-cyclodextrin from Carbosynth, MTT and heparin from Alfa Aesar; Dynasore and G-418 from EMD Millipore; ethidium bromide, EIPA, ammonium chloride and chloroquine were purchased from Sigma-Aldrich.

A Guava EasyCyte™ cytometer (EMD Millipore) was used for all flow cytometry experiments. Cell microscopy images were acquired with an Andor Zyla 4.2 digital camera mounted on a Nikon Eclipse Ti-E epifluorescence microscope or with a Leica SP5 confocal microscope. Absorbance of formazan was measured in a microplate reader Tecan Infinite F200Pro. PCR reactions, re-annealing and digestions were done on a T-Personal PCR Thermocycler (Biometra) or on a MJ Mini Thermal Cycler (Bio-Rad). Gels were imaged on a ChemiDoc XRS+ system (Bio-Rad) and analysed with the software ImageLab 5.2.1 (Bio-Rad).

Transmission electron microscopy was performed on a JEOL JEM-2010 microscope operating at 200 kV or on a JEOL JEM-1011 microscope operating at 100 kV. EDX mapping and microanalysis were performed on a JEOL JEM-2010 FEG operating at 200 kV at CACTI building (University of Vigo, Spain). DLS and ζ potential measurements were performed in a Malvern Zetasizer NanoZS using standard disposable cuvettes. All experiments were done in triplicate at 25 ºC.

High-performance liquid chromatography coupled with mass spectrometry (HPLC-MS) analyses were carried out on Agilent Technologies 1260 Infinity II associated with a 6120 Quadrupole LC-MS using an Agilent SB-C18 column with *Solvent A:Solvent B* gradients between 5:95 (*Solvent A*: H_2O with 0.1 % TFA; *Solvent B*: CH_3CN with 0.1 % TFA). High-performance liquid chromatography (HPLC) preparative purification was carried out on Waters 1525 composed by a binary pump with a dual Waters 2489 detector with a Phenomenex Luna C18(2) 100A column. An Agilent 1200 with an Agilent Eclipse XDB-C18 column was used for semi-preparative purification using gradients of 5:95 (*Solvent A*: H_2O with 0.1 % TFA; *Solvent B*: CH_3CN with 0.1 % TFA). Nuclear Magnetic Resonance (NMR) spectra were recorded on a Varian Mercury 300 MHz spectrometer. Chemical shifts are reported in ppm (δ units) referenced to the following solvent signals: D_2OδH 4.79. Spin multiplicities are reported as a singlet (s), doublet (d), with coupling constants (*J*) given in Hz, or multiplet (m). Accurate mass determinations (HR-MS) using ESI-MS were performed on a Sciex QSTAR Pulsar mass spectrometer and are reported as mass-per-charge ratio (*m/z*).

2.3 Peptide Synthesis

Peptides were synthesised as previously described elsewhere[S4]. The TAMRA and Ox-Dex *N*-terminated **P** were prepared following a modified protocol.

The *N*-terminal Fmoc protecting group was removed with a solution of piperidine in DMF (20 %, 2 mL) for 15 min and then the mixture was filtered and the resin was washed with DMF (3 x 2 mL, 1 min). Then, a premixed solution in DMF of 6-(Fmoc-amino)hexanoic acid (4 equiv), *N*-HBTU (4 equiv) (2 mL) and DIEA (6 equiv) was added to the resin under nitrogen stream for 20 min. Finally, the resin was washed with DMF (3 x 2 mL, 1 min) and DCM (3 x 2 mL, 1 min).

The Lys-Mtt protecting group was cleaved by using mild acidic conditions DCM/HFIP/TFE/TIS (6.5/2/1/0.5, 2 x 2 mL) followed by acylation of the Lys d-amino group with 5-(2-(tert-butoxycarbonyl)hydrazineyl)-5-oxopentanoic acid as previously described[S4].

TAMRA terminating peptide (TmP): Fmoc-protecting group of the previously attached Fmoc-aminohexanoic acid was removed by using a solution of piperidine in DMF (20%, 4 mL) for 15 min and the resin was washed with DMF (3 x 3 mL). The coupling was carried out by the addition of a premixed solution of TAMRA (1 equiv), *N*-HBTU (1 equiv) and DIEA (2 equiv) in DMF (2 mL) and the mixture was stirred under nitrogen stream for 30 min followed by washings with DMF (3 x 3 mL) and DCM (3 x 3 mL).

Ox-Dex terminating peptide (DexP): the Fmoc-protecting group of the linker was removed by using a solution of piperidine in DMF (20%, 4 mL) during 15 min and the resin was washed with DMF (3 x 3 mL). A solution of Ox-Dex[S5] (3 equiv), *N*-HATU (2.9 equiv) and DIEA (0.195 M, 3 equiv) in DMF (2 mL) was added and the mixture was shaken by bubbling nitrogen for 30 min. Finally, the resin was washed with DMF (3 x 3 mL) and DCM (3 x 3 mL).

The modified peptides were cleaved from the solid support with removal of the protecting groups under strong acidic standard conditions: TFA/DCM/TIS/H_2O (90/5/2.5/2.5) for 2 h and precipitated in Et_2O. The suspension was centrifuged, the solid residue dissolved in H_2O/CH_3CN (1/1) and purified by preparative C18 reverse-phase HPLC [Phenomenex Luna C18(2) 100A column, H_2O (0.1% TFA)/CH_3CN (0.1% TFA) 95:5 (5 min), 95:5→5:95 (5→35 min)] with a binary gradient of *Solvent A* and *Solvent B*.

TmP: The corresponding fractions were freeze-dried to afford the pure peptide **TmP** as a pink solid (11.2 mg, 11% yield). Purity and characterization were confirmed by analytical HPLC, mass spectrometry and ^{1}H NMR (Figure S10). ^{1}H NMR (300 MHz, D_2O) δ (ppm): 8.31 (s, 1H), 7.96 (d, J = 8.0 Hz, 1H), 7.36 (d, J = 8.2 Hz, 1H), 7.15-6.91 (m, 2H), 6.90-6.62 (m, 4H), 4.29-3.78 (m, 13H), 3.46-3.19 (m, 2H), 3.11 (s, 12H), 3.08-2.85 (m, 16H), 2.89-2.72 (m, 2H), 2.39-2.01 (m, 8H), 1.89-1.15 (m, 61H), 0.93-0.53 (m, 30H). MS (ESI, H_2O/ CH_3CN): 512 (22, $[M+5H]^{5+}$), 641 (50, $[M+4H]^{4+}$), 669 ($[M+4H+TFA]^{4+}$), 892 ($[M+3H+TFA]^{3+}$), 930 (100, $[M+3H+2TFA]^{3+}$), 968 (38, $[M+3H+3TFA]^{3+}$), 1394 (8, $[M+2H]^{2+}$), 1451 (10, $[M+2H+TFA]^{2+}$), 1508 (5, $[M+2H+2TFA]^{2+}$). HRMS (ESI): Calcd for $C_{119}H_{204}N_{41}O_{22}$ $[M+3H]^{3+}$: 853.2029; found: 853.2032.

DexP: The corresponding fractions were freeze-dried to afford the pure peptide **DexP** as a white solid (7.3 mg, 3 % yield). Purity and characterization were confirmed by analytical HPLC, mass spectrometry and ^{1}H NMR (Figure S11). ^{1}H NMR (300 MHz, D_2O) δ (ppm): 7.38 (d, J = 8.9 Hz, 1H), 6.27 (d, J = 8.7 Hz, 1H), 6.07 (s, 1H), 4.32-3.79 (m, 15H), 3.33-3.15 (m, 2H), 3.17- 2.93 (m, 17H), 2.91-2.75 (m, 2H), 2.66- 2.49 (m, 2H), 2.35-2.04 (m, 11H), 2.02- 1.08 (m, 71H), 0.99- 0.56 (m, 33H). MS (ESI, H_2O/CH_3CN): 1368 (5, $[M+2H+2TFA]^{2+}$), 950 (25, $[M+3H+3TFA]^{3+}$), 912 (45, $[M+3H+2TFA]^{3+}$), 874 (100, $[M+3H+TFA]^{3+}$), 836 (25, $[M+3H]^{3+}$).

2.4 Preparation of amphiphiles

Peptide was mixed with 6 equiv of the aldehyde tail per hydrazide dissolved in DMSO in the presence of 5 % of AcOH and shaked at 60 ºC for 2 h. Concentration of the peptide in the reaction was 1.25 mM. Before incubation with the complexes, reactions were diluted to 50 µM using DMEM or PBS.

2.5 Oligonucleotides

The following crRNAs targeting EGFP (Alt-R™ crRNA 5'-CCTCGAACTTCACCTCGGCG-3')[S6] and HPRT1 (Alt-R™ crRNA 5'-GGCTTATATCCAACACTTCG-3') were used. For the amplification of the surrounding regions in the genomic DNA, the primers TACGGCAAGCTGACCCTGAA and GTCCATGCCGAGAGTGATCC[S6] (Tm = 54 ºC) were used in the case of EGFP and CAAATTATGAGGTGCTGGAAGG and TGGACACATGGGTAGTCAG (Tm = 52 ºC) for HPRT1.

2.6 Cell lines and culture

Cervix adenocarcinoma HeLa cell line and human lung cancer A549 cell lines were maintained on Dulbecco's Modified Eagle's Medium (4500 mg/L glucose, L-glutamine, sodium pyruvate and sodium bicarbonate) supplemented with 10 % FBS and 1 % Penicillin-Streptomycin-Glutamine Mix (Fisher) at 37 ºC/5 % CO_2 /95 % humidity in an INCO 108 incubator (Memmert).

To generate an EGFP expressing HeLa cell line, HeLa cells were transfected with pEGFP-C1 (Clontech) and stable transfectants were selected with 400 µg/mL of G-418 (Millipore). Green clones were selected by epifluorescence microscopy and tested for stability of GFP expression through several passages. A clone with a highly stable GFP expression (~90 % GFP positive cells) was chosen for further experiments.

A chicken fibroblast cell line (DF-1 GFP), containing a tetracycline-inducible EGFP gene was a kind gift from Dr. Lisa K. Busch (Molecular Virology group).

2.7 RNP delivery

For RNP assembly, crRNA and tracrRNA were mixed to a final concentration of 4 µM in Nuclease-Free Duplex Buffer (30 mM HEPES, 100 mM potassium acetate, pH 7.5), incubated for 5 min at 95 ºC and slowly cooled at room temperature. Cas9 protein was dissolved in DMEM at a final concentration of 4 µM and equal volumes of the RNA complex and Cas9 were mixed and incubated for 5 min at room temperature. The desired amount of RNP was then mixed with the freshly prepared amphiphiles diluted at 50 µM in DMEM and incubated for 20 min to form the peptide/Cas9 complexes. Usually 4 µL of RNP were mixed with 10 µL of the amphiphile at 50 µM. Lipofectamine was used at a concentration of 10 µg/mL to deliver 160 nM of RNP, and this ratio was kept constant in all the experiments.

Cells seeded the day before in a 96 well plate (50.000 cells/mL, 100 µL/well) were washed once with DMEM, and DMEM or DMEM supplemented with FBS (to obtain a final concentration of 5 %) was added to the wells. Peptide/Cas9 complexes were added to a final volume of 50 µL per well and cells were incubated for 4 h before removing the complexes and adding 100 µL of DMEM with 10 % FBS and antibiotics.

2.8 T7 Endonuclease I assay

For T7E1 assays, cells were harvested at 48 hpt by trypsinization and centrifugation. Genomic DNA from the cellular pellet was extracted and amplified using the GeneArt Genomic Cleavage Detection kit (ThermoFisher) following the manufacturer's instructions. 2 µL of the PCR product were re-annealed and digested using 1 U of T7E1 (NEB) at 37 ºC for 1.5 h. DNA fragments were resolved in a 5 % polyacrylamide gel for 35 min at 150 V in TAE buffer. Gels were stained with ethidium bromide before imaging on a ChemiDoc XRS+. Band intensity was measured with the software ImageLab 5.2.1 (Bio-Rad) and cleavage efficiency (C. e.) was calculated using the following equation:

$$C.\ e. = 1\ - [(1 - \frac{sum\ of\ cleaved\ bands\ intensities}{sum\ of\ cleaved\ and\ parental\ band\ intensities})^{1/2}]$$

Results were expressed as a percentage.

2.9 Flow cytometry

For EGFP gene disruption studies, HeLa-GFP cells were incubated with Peptide/RNP complexes targeting the EGFP gene. At 72 hpt, cells were washed with PBS and trypsinized with 100 µL of Trypsin-EDTA for 10 min at 37 ºC. After neutralizing trypsin by the addition of 100 µL of 2 % FBS and 5 mM EDTA in PBS, cells were analysed on a Guava EasyCyte™ cytometer. GFP fluorescence was measured by excitation at 488 nm and detection at 512/18 nm. For the analysis, cells with typical FSC and SSC parameters were selected and cells with low GFP fluorescence (lower than the main peak of the untreated control) were considered as GFP negative for further calculations.

To assess the state of the lysosomal compartment, HeLa cells were incubated in DMEM with the peptide/RNP complexes in the presence of 50 nM Lysotracker Green DND-26. After 1 h of incubation at 37 ºC, cells were washed three times with DMEM and trypsinized for 10 min at 37 ºC. Trypsin was neutralized by the addition of 2 % FBS and 5 mM EDTA in PBS and samples were analysed on a Guava EasyCyte™ cytometer. Lysotracker Green DND-26 fluorescence was measured by excitation at 488 nm and detection at 512/18 nm. After gating to discard debris, median fluorescence intensity (MFI) of each sample was calculated. MFI value from unlabelled cells was subtracted from all samples. Values were normalized to the MFI of LysoTracker in the untreated control (100 %).

To study the cellular uptake mechanisms, HeLa cells growing in a 96 well plate were treated for 30 min with the following compounds diluted in DMEM without serum or antibiotics: Dynasore (80 µM), chlorpromazine (30 µM), methyl-β-cyclodextrin (5 mM),

Wortmannin (200 nM), EIPA (50 μM), heparin (5 μg/mL), chloroquine (100 μM) or ammonium chloride (50 mM). For the incubation at low temperature, another plate was incubated on ice and ice-cold solutions were used for the washes and incubations. Cells were then incubated with the complexes (prepared with the TAMRA-labelled peptide conjugated with oleic aldehyde) in the presence of the same amount of inhibitors in DMEM for 1 h at 37 ºC. Cells were washed twice with HKR and 0.1 mg/mL of heparin in HKR and trypsinized. Trypsin was neutralized with 2 % FBS in PBS with 5 mM EDTA and cell fluorescence was measured on a Guava EasyCyte™ cytometer using a green laser (532 nm) and collecting the emission at 575/25 nm. Cells with typical FSC and SSC parameters were selected and the median fluorescence intensity calculated for each sample (MFI). Each condition was done in triplicate. Fluorescence values were normalized to the uptake of the untreated control (100 %) after blank subtraction.

In all cases, data analysis was performed with InCyte software included in GuavaSoft 3.2 (Millipore).

2.10 Live cell microscopy

HeLa cells grown in four chamber glass bottom dishes (CellVis) were stained with 1 μM Hoechst 33342 for 20 min. After nuclear staining, cells were washed once with DMEM and incubated with the peptide/RNP complexes (containing **TmPT$_{24}$** and the RNP, or ATTO$_{550}$-labelled RNP with unlabelled **PT$_{24}$**) for 1 h in DMEM without FBS or antibiotics (unless otherwise stated). For lysosomal staining, 50 nM of LysoTracker Deep Red was added to the culture. Cells were then washed twice with HKR buffer (5 mM HEPES, 137 mM NaCl, 2.68 mM KCl, 2.05 mM $MgCl_2$, 1.8 mM $CaCl_2$, pH 7.4) and twice for 5 min with 0.1 mg/mL of heparin in HKR buffer and examined under an epifluorescence microscope.

2.11 Glucocorticoid induced GFP translocation assay

Endosomal escape was studied using a glucocorticoid induced GFP translocation assay (GIGT)[S7]. HeLa cells were transfected with the plasmid pK7-GR-GFP (a gift from Ian Macara[S8], Addgene plasmid #15534) using Lipofectamine 2000 the day before the treatments. Cells were then incubated for 30 min with 1 μM Hoechst 33342 in DMEM and 50 μM EIPA or 5 % FBS where indicated. This media was removed and cells were incubated for 1 h with 1 μM dexamethasone (positive control), 10 μM **DexPT$_{24}$**, or left untreated (negative control) diluted in DMEM with or without EIPA or 5 % FBS, as indicated. Twenty to thirty images of each sample were acquired with an Andor Zyla 4.2 digital camera mounted on a Nikon Eclipse Ti-E microscope at 60x magnification and the translocation ratio (the ratio of the median intensities of GFP in the nucleus and in the 2 μm wide surrounding region) was calculated with CellProfiler[S9] as follows. Nuclei were identified as Hoechst stained objects using the three-class thresholding Otsu method and the cytoplasmic region was defined as the 2 μm surrounding area. To ensure a better separation of the cytoplasmic and nuclear region, nuclei were shrunk 0.5 μm before measuring object intensity. Cells falling below the 15 % of the maximum intensity of the image were considered untransfected and discarded for the analysis. A total of 40 to 80 cells were analyzed for each sample.

Statistical analysis of the data was performed with R software[S10]. Results were subjected to pairwise two-tailed Student's t-test and p-values were adjusted using Bonferroni's correction.

2.12 Viability assay

Cell viability was determined using a MTT assay, which relies on the ability of cells to reduce the water-soluble tetrazolium salt to the insoluble formazan. HeLa cells incubated with the complexes as previously described were incubated for 48 h at 37 ºC before performing the assay. Cell culture media was removed and 100 µL of fresh DMEM containing FBS, antibiotics and supplemented with 0.5 mg/mL of MTT was added. As a blank control, 3 wells were incubated with DMEM and MTT in the presence of 1 % Triton X-100. Cells were incubated for 4 h at 37 ºC before carefully removing the supernatant and dissolving the formazan crystals with DMSO (100 µL/well). Absorbance was measured at 570 nm using a microplate reader (Infinite F200pro, Tecan). Data points were collected in triplicate and values were normalized for untreated control cells (100%) after blank subtraction.

2.13 Protein expression and purification

Escherichia coli BL21(DE3)-Codon Plus cells (Stratagene) were transformed with the plasmid pET-Cas9-NLS-6xHis (a gift from David Liu[S6],Addgene plasmid #62933). A bacterial preculture grown overnight at 37 ºC was diluted 1:100 in LB medium supplemented with 100 µg/mL Ampicillin. Cultures were incubated at 37 ºC until OD_{600} ~ 0.6, cooled to 20 ºC and induced with 0.5 mM IPTG for 16 h at 20 ºC. Cells were harvested by centrifugation, pellets were resuspended in 1:10 volume of lysis buffer (220 mM NaCl, 4.3 mM KCl, 12.8 mM Na_2HPO_4, 2.4 mM KH_2PO_4, 20 % glycerol, 0.1 % Triton X-100, 10 mM imidazole, 1 mg/mL lysozyme supplemented with protease inhibitors without EDTA) and sonicated 15 x 10 s (Cycle 0.5, Amplitude 70 %) with a UP200S ultrasonic processor (Hielscher). Cellular debris was removed by centrifugation at 16.000 g 15 min and the supernatant was incubated with HisPur Ni-NTA resin (ThermoFisher) for 1 h at 4 ºC, washed 3 times with wash buffer (220 mM NaCl, 4.3 mM KCl, 12.8 mM Na_2HPO_4, 2.4 mM KH_2PO_4, 10 mM imidazole) and then the protein was eluted with several incubations of the same buffer containing 300 mM imidazole. Imidazole from the fractions containing the protein of interest was removed by several washes on an Amicon UFC (Millipore), and finally, protein was concentrated in storage buffer (20 mM HEPES pH 7.5, 500 mM KCl, 20 % of glycerol). Aliquots were snap-frozen and stored at -80 ºC.

2.14 Gel retardation assay

For gel retardation assay, the RNA complex was prepared by mixing an $ATTO_{550}$-labelled tracrRNA with the crRNA. Cas9 was diluted in DMEM without phenol red and incubated for 5 min at room temperature with the RNA complex to form the RNP. 4 µL of 2 µM RNP were mixed with different amounts of the freshly prepared amphiphile and incubated for 20 min to allow for the complexes to form. As controls, 4 µL of 2 µM RNP or 4 µL of 2 µM gRNA were mixed with 10 µL of DMEM and incubated for 20 min. Samples

were resolved in an 2 % agarose gel, for 25 min at 80 V in TAE buffer and imaged in a ChemiDoc.

2.15 Confocal microscopy

The RNP was prepared mixing $ATTO_{550}$-labelled tracrRNA, crRNA and Cas9 protein as in the gel retardation experiments. For the peptide/Cas9 complexes formation, 4 µL of 2 µM RNP were mixed with 10 µL of 50 µM **PT_{24}** and incubated for 20 min at room temperature. These complexes were diluted to a final concentration of 160 nM RNP and 10 µM **PT_{24}** using DMEM without phenol red. As a control, RNP alone was diluted to a final concentration of 160 nM. Samples were imaged on a Leica SP5 confocal microscope.

2.16 TEM

General: RNP was prepared at 2 µM as for a delivery experiment. Amphiphiles were resuspended at a concentration of 50 µM in DMEM and 4 µL of the RNP were mixed with 10 µL of the amphiphile and incubated for 20 min at room temperature before adsorption on the TEM grids. The resulting 14 µL of mixture were pipetted onto a carbon coated TEM grid (300 mesh) for 10 min, the excess was removed by using filter paper. The grids were washed with ultrapure water prior to stain the samples with 10 µL of a 2% solution in water of phosphotungstic acid or either uranyl acetate for 1 min. The grids were washed again with ultrapure water and dried at atmospheric conditions.

To prepare the complexes incorporating gold nanoparticles: Freshly prepared amphiphiles (6 equiv of oleic aldehyde) were incubated with the ribonucleoprotein for 20 min as previously described in this section. Subsequently, 2.3 µL of 100-fold diluted gold nanoparticle suspension from 5 nm Ni-NTA Nanogold (commercial stock 0.5 µM) were added to the suspension and incubated for 30 min. The mixture was then placed onto a TEM grid. After 10 min, the excess of solution was removed by using a filter paper. The grids were washed with ultrapure water prior to stain the samples with 10 µL of a 2% uranyl acetate solution in water for 1 min. The grids were washed again with ultrapure water and dried at atmospheric conditions.

To prepare the complexes embedded in agarose: Low melting point agarose (2% in PBS) was dissolved on a hot plate at 90 ºC and 3-4 drops were added to freshly prepared complexes (14 µL) and manually agitated to obtain a homogeneous suspension. Samples were incubated at 4 ºC to allow the agarose to solidify (3-4 h). The sample was removed from the tube and the agar excess was removed by using a razor blade. The agarose block was cut into tissue-size pieces for embedding. Sample was stained with osmium tetroxide, dehydrated and embedded in epoxy resin prior to visualization.

2.17 DLS

RNP at a concentration of 2 µM was prepared by mixing equal amounts of the tracrRNA:crRNA complexes and Cas9 produced in bacteria at 4 µM. Amphiphiles were resuspended at a concentration of 50 µM in DMEM and 80 µL of the RNP were mixed with

200 μL of the amphiphile and incubated for 20 min at room temperature. The mixtures were diluted in ultrapure water or in PBS up to 10 μM prior to be transferred to a standard disposable cuvette for size or ζ potential measurements. All experiments were done in triplicate at 25 ^{o}C.

2.18 Supporting figures

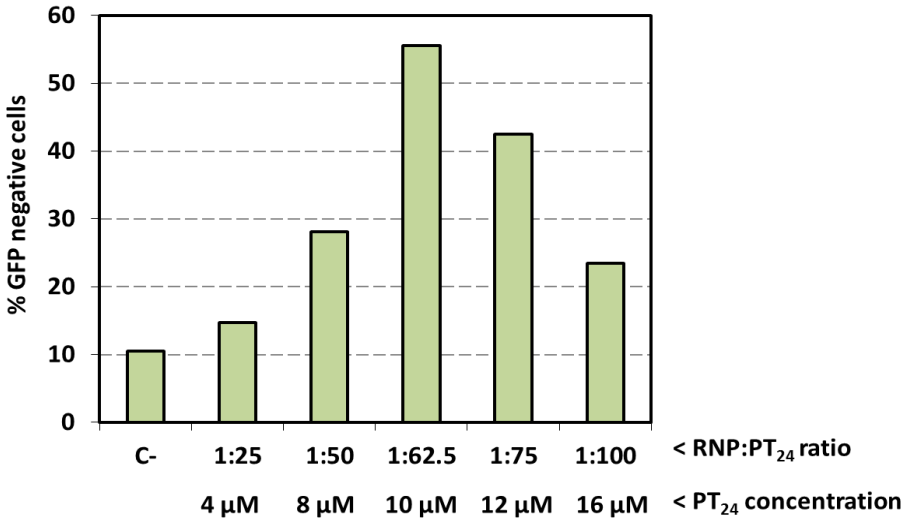


Figure S1. Optimization of peptide/Cas9 RNP delivery conditions. To find the optimal transfection ratio, HeLa-GFP cells were incubated with a fixed amount of RNP targeting GFP (160 nM) and different concentrations of **PT_{24}**, as indicated on the figure. At 72 hpt the number of GFP negative cells was measured by flow cytometry. As the optimization was done using oleic aldehyde, it is possible that some of the amphiphiles generated with other aldehydes may be active under different conditions.

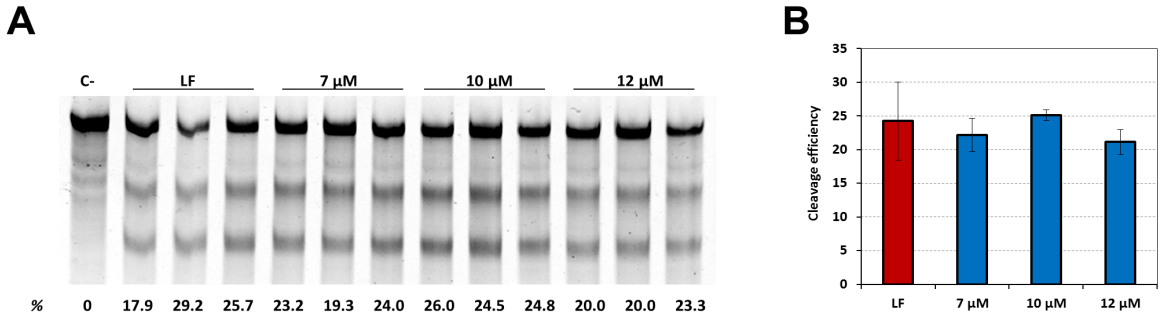


Figure S2. Incubation with peptide/RNP in the presence of 5 % serum. HeLa cells were incubated with 160 nM RNP targeting the HPRT1 gene and different concentrations of peptide **PT_{24}** (7, 10 or 12 µM) or with Lipofectamine 2000 (10 µg/mL). A) T7E1 assay done at 48 hpt. Cleavage efficiency is indicated below the gel. B) Average cleavage efficiency for each concentration. Error bars indicate standard deviation.

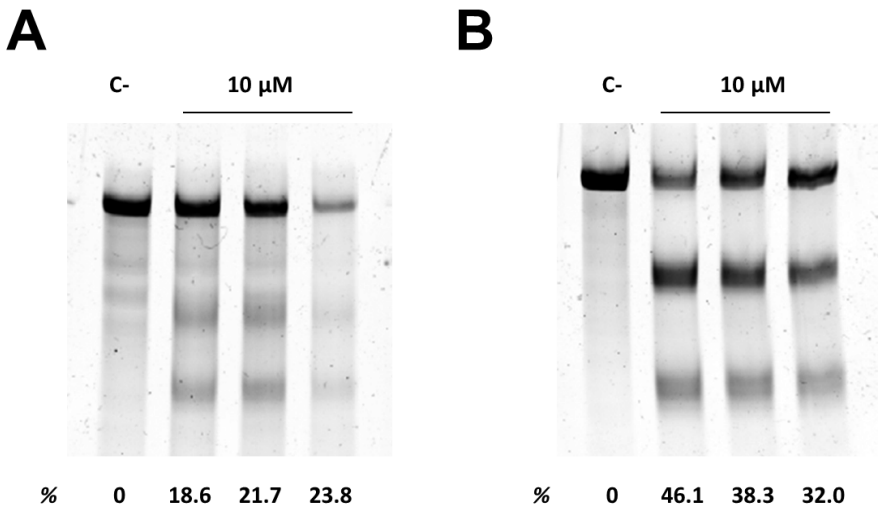


Figure S3. Gene edition of other cell lines using $\mathbf{PT_{24}}$. A) Human lung cancer A549 cells incubated with 10 µM $\mathbf{PT_{24}}$/160 nM of RNP against HPRT1. B) Chicken fibroblast DF-1 GFP cell line incubated with 10 µM $\mathbf{PT_{24}}$/160 nM of RNP against EGFP.Cleavage efficiency is indicated below each lane.

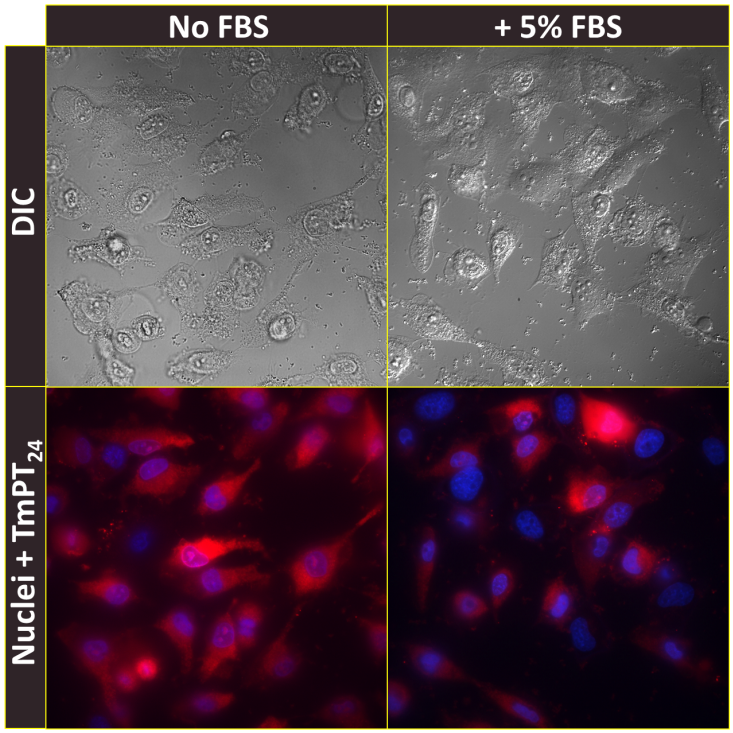


Figure S4. Increased toxicity of the peptide $\mathbf{TmPT_{24}}$. HeLa cells were incubated with the complexes (10 µM peptide/160 nM RNP) for 1 h. Nuclei were counterstained blue. The presence of the fluorophore increases the toxicity of the peptide, because at 10 µM membrane damage can be seen. Hence, the lower concentration of 5 µM was chosen for all the experiments that used the labelled peptide.

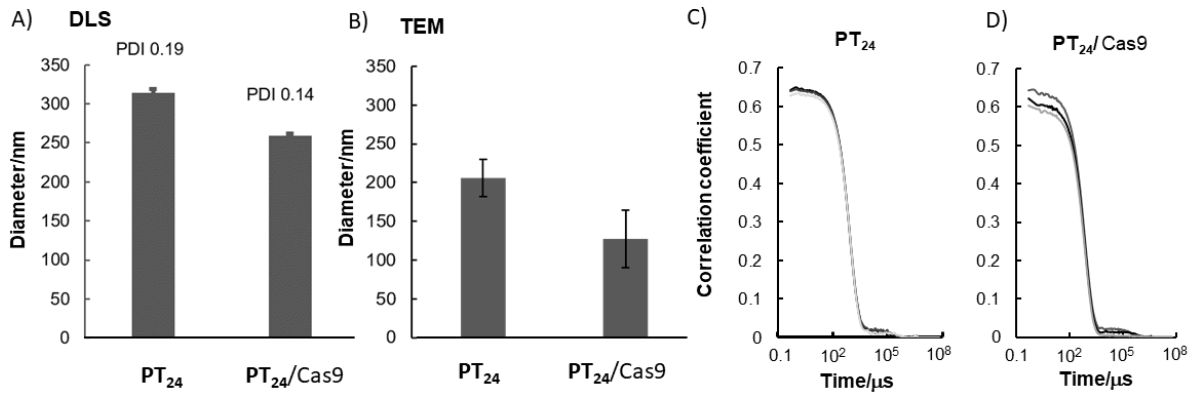


Figure S5. A) **PT$_{24}$** and **PT$_{24}$**/Cas9 RNP complex size diameters measured by DLS in DMEM/H_2O 1:4. Polydispersity indexes (PDI) are also shown above the corresponding columns. B) **PT$_{24}$** and **PT$_{24}$**/Cas9 RNP complex size diameters measured by TEM. Average and standard deviations (error bars) from 10 independent particles each are depicted. C) and D) Raw correlation data for the three DLS measurements of **PT$_{24}$** and **PT$_{24}$**/Cas9 RNP complex respectively. Data shows a slight decrease in size upon complexation with the RNP. All measurements were performed in triplicate. Error bars represent standard deviation.

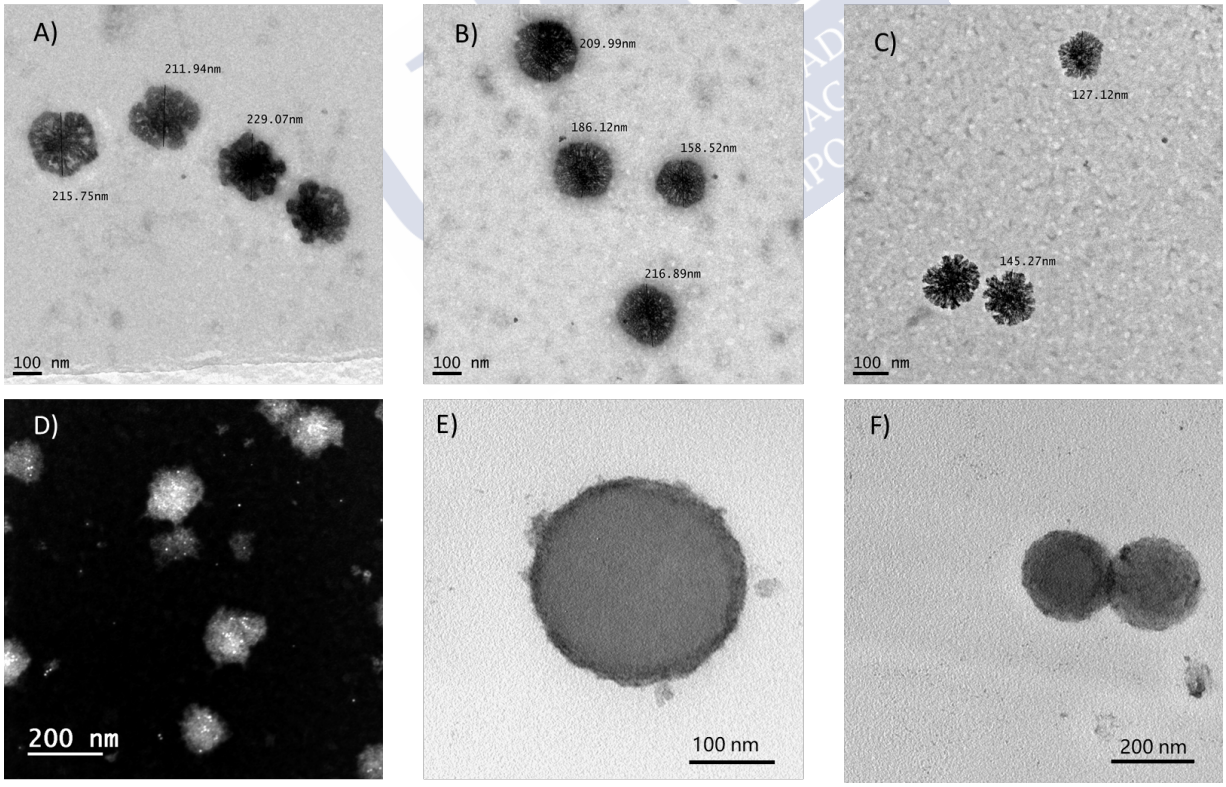


Figure S6. Exemplary TEM micrographs of A) **PT$_{24}$** in DMEM, B) and C) **PT$_{24}$**/Cas9 complex in DMEM. D) STEM micrograph of **PT$_{24}$**/Cas9 complex incubated with Ni-NTA Nanogold. E) and F) TEM micrographs of **PT$_{24}$**/Cas9 complex embedded in agarose.

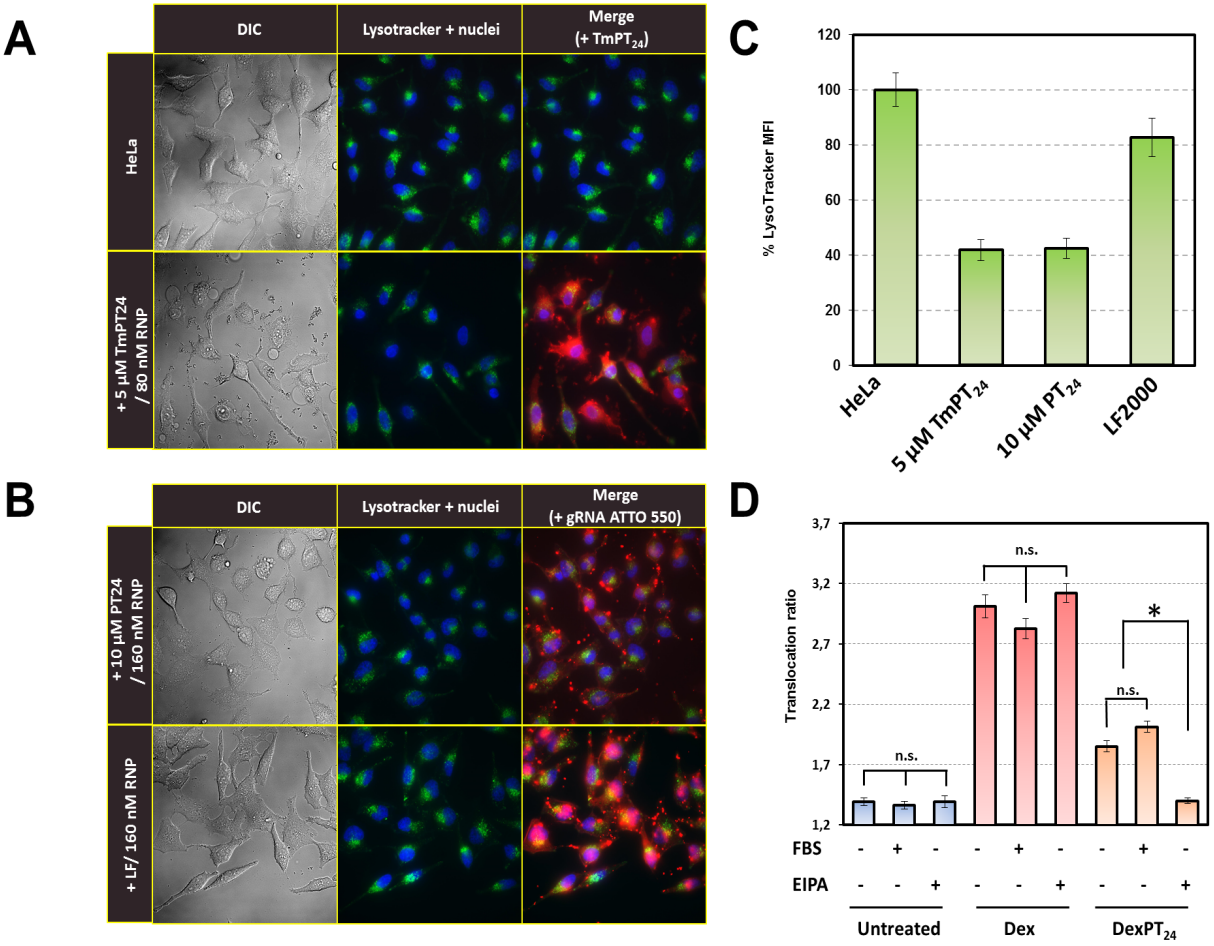


Figure S7. Lysosomal disruption. A) HeLa cells were incubated with 5 µM of **TmPT$_{24}$** and 80 nM of RNP in the presence of 50 nM LysoTracker Deep Red (green). After 1 h of incubation, cells were washed to remove the excess of LysoTracker and imaged. Fluorescence of the peptide is shown in red and nuclei are stained blue with Hoechst 33342. B) HeLa cells were incubated with 160 nM of ATTO$_{550}$-labelled RNP (red) using 10 µM **PT$_{24}$** or 10 µg/mL Lipofectamine 2000 in the presence of 50 nM LysoTracker Deep Red (shown in green). After 1 h, cells were washed and imaged. Nuclei are stained blue. C) Flow cytometry quantitation of lysosomal staining. Cells were incubated with the same complexes used in A and B, in the presence of 50 nM LysoTracker Green DND-26 for 1 h; washed and trypsinized to measure LysoTracker Green fluorescence by flow cytometry. D) GIGT assay in HeLa cells expressing GR-GFP (see methods). Incubations were performed in DMEM stock, DMEM with 5 % of FBS or DMEM containing 50 µM EIPA (a macropinocytosis inhibitor). Cells were left untreated (blue bars), incubated with 1 µM dexamethasone (red bars) or incubated with 10 µM of **DexPT$_{24}$** (orange bars) for 1 h. Asterisk indicates a p-value < 0.001, n.s. = not significant (p > 0.05); error bars indicate SEM.

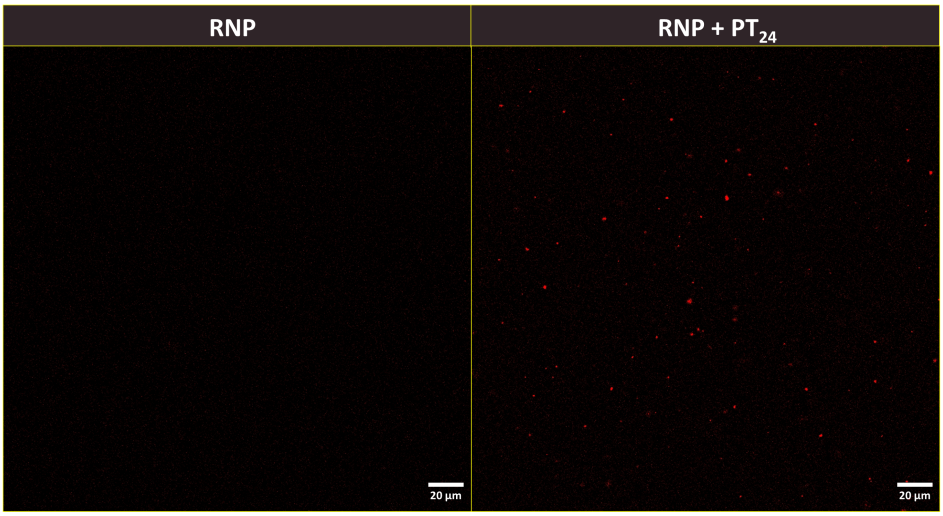


Figure S8. Confocal images of peptide/RNP complexes. RNP was prepared using $ATTO_{550}$-labelled gRNA (red) and diluted to a final concentration of 160 nM (RNP) or complexed with **PT_{24}** and then diluted to a final concentration of 10 µM **PT_{24}** and 160 nM RNP. Dilutions were done in DMEM without phenol red. Samples were imaged with a confocal microscope.

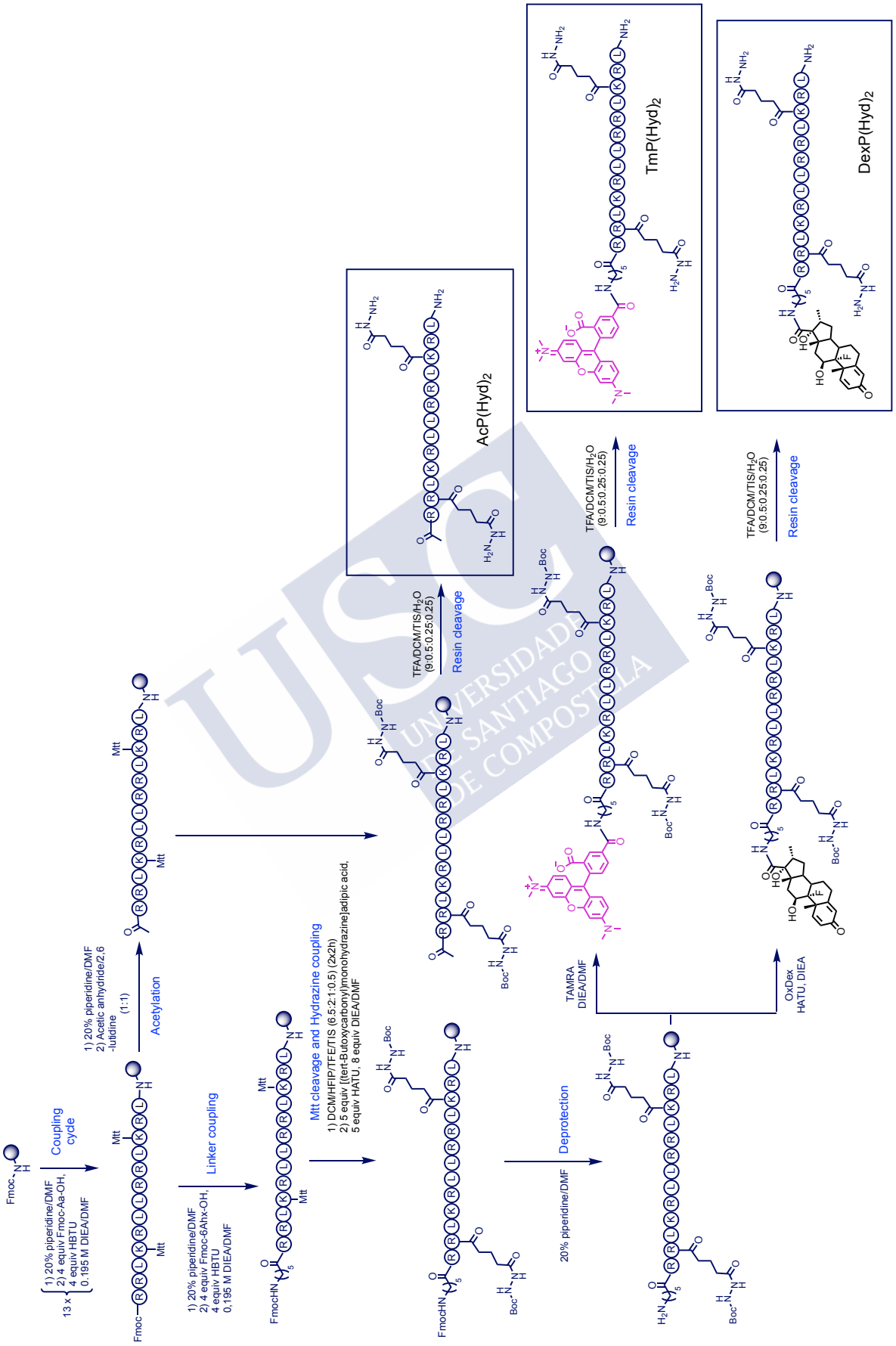
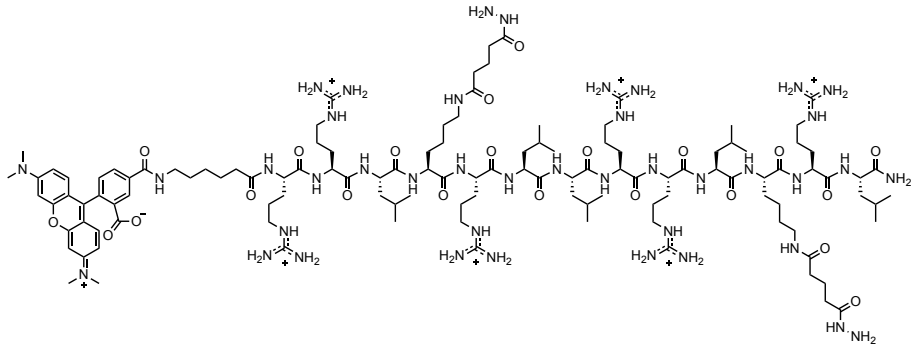
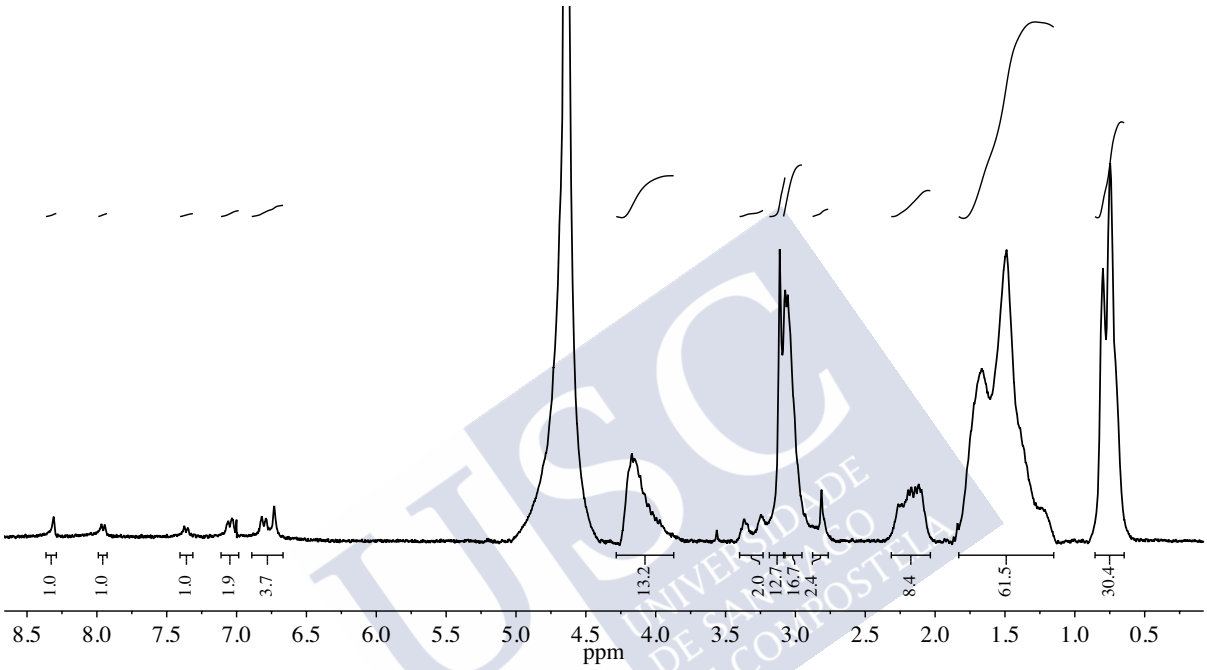


Figure S9. General synthetic scheme for the solid phase peptide synthesis (SPPS).





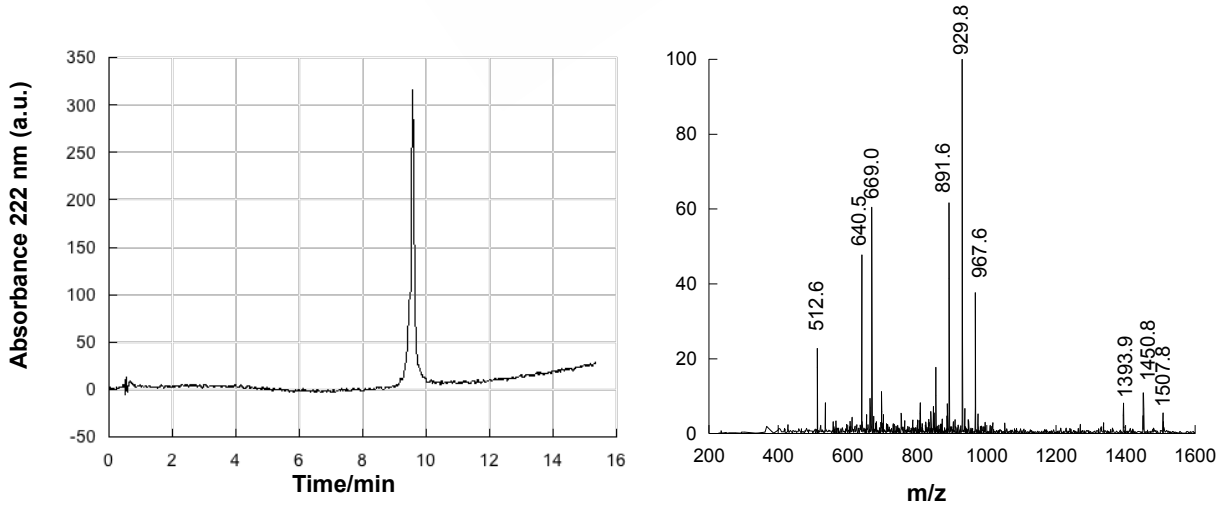


Figure S10. ^{1}H-NMR spectrum in D_2O of **TmP.** RP-HPLC [Agilent SB-C18 Column, H_2O (0.1 % TFA)/CH_3CN (0.1% TFA) 95:5 → 5:95, (0 → 15 min)]. R_t = 9.6 min and ESI-MS for **TmP.**

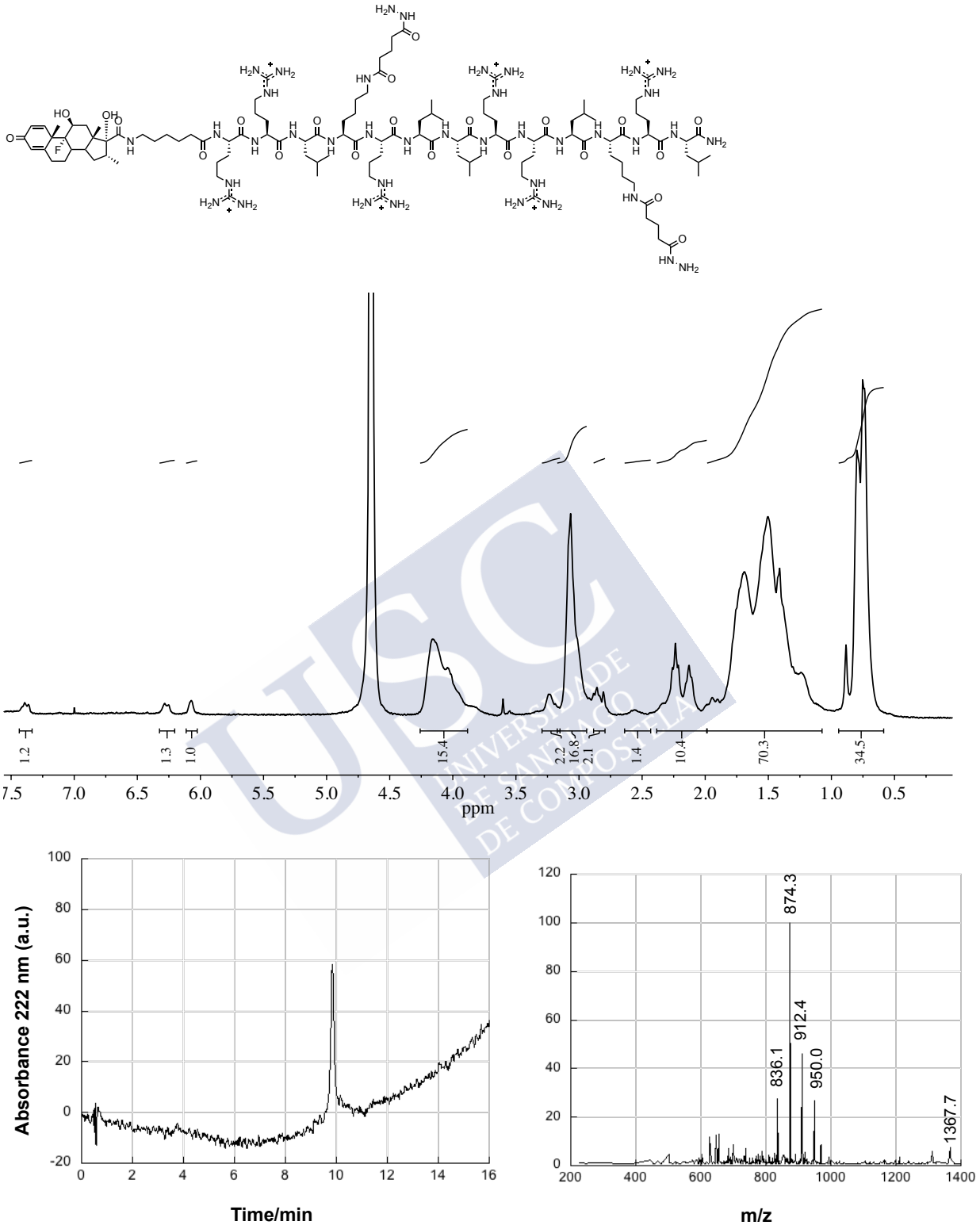


Figure S11. ^{1}H-NMR spectrum in D_2O of **DexP.** RP-HPLC [Agilent SB-C18 Column, H_2O (0.1 % TFA)/CH_3CN (0.1% TFA) 95:5 → 5:95, (0 → 15 min)]. R_t = 9.8 min and ESI-MS for **DexP.**

2.19 Supporting references

S1 C. Gehin, J. Montenegro, E. K. Bang, A. Cajaraville, S. Takayama, H. Hirose, S. Futaki, S. Matile and H. Riezman, J. Am. Chem. Soc., 2013, 135, 9295.
S2 Q. Chen, D. A. Sowa, J. Cai and R. Gabathuler, Synth. Commun., 2003, 33, 2377.
S3 M. A. Cole, S. E. Tully, A. W. Dodds, J. N. Arnold, G. E. Boldt, R. B. Sim, J. Offer and P. Wentworth, ChemBioChem, 2009, 10, 1340.
S4 I. Louzao, R. García-Fandiño and J. Montenegro, J. Mater. Chem. B, 2017, 5, 4426.
S5 M. V Govindan and B. Manz, Eur. J. Biochem., 1980, 108, 47.
S6 J. A. Zuris, D. B. Thompson, Y. Shu, J. P. Guilinger, J. L. Bessen, J. H. Hu, M. L. Maeder, J. K. Joung, Z.-Y. Chen and D. R. Liu, Nat. Biotechnol., 2015, 33, 73.
S7 J. M. Holub, J. R. Larochelle, J. S. Appelbaum and A. Schepartz, Biochemistry, 2013, 52, 9036.
S8 K. L. Carey, S. A. Richards, K. M. Lounsbury and I. G. Macara, J. Cell Biol., 1996, 133, 985.
S9 A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland and D. M. Sabatini, Genome Biol., 2006, 7, R100.
S10 R Core Team, R Found. Stat. Comput., 2015.

3 MESSENGER RNA DELIVERY BY HYDRAZONE-ACTIVATED POLYMERS

Marisa Juanes,[a] Oliver Creese,[b] Francisco Fernández-Trillo[b]* and Javier Montenegro[a]**
[a] Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain.
**e-mail: javier.montenegro@usc.es
[b]School of Chemistry, University of Birmingham, Birmingham B15 2TT, U.K.
*e-mail: f.fernandez-trillo@bham.ac.uk

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

3.1 Abbreviations

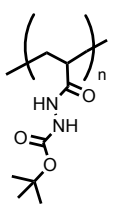
DLS: dynamic light scattering; DMEM: Dulbecco's Modified Eagle Medium; DMF: dimethylformamide; DMSO: Dimethylsulfoxide; DOTAP: 1,2-dioleoyl-3-trimethylammonium-propane (chloride salt); EGFP: enhanced green fluorescent protein; FBS: Fetal Bovine Serum; FSC: forward scatter channel; GALA: WEAALAEALAEALAEHLAEALAEALEALAA;[S1] hpt: hours post transfection; MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; PEI: polyethyleneimine; PBS: phosphate-buffered saline; R8: octaarginine;[S2] SSC: side scatter channel; TAE: tris-acetate-ethylenediaminetetraacetic acid.

3.2 Materials and Methods

Nuclear Magnetic Resonance (NMR) spectra were recorded on a Bruker Avance III 300 MHz spectrometer. Chemical shifts are reported in ppm (δ units) referenced to the following solvent signals: DMSO-d_6 δH 2.50 and D_2O δH 4.79. Dialysis was carried out in deionised water at room temperature for a minimum of 48 hours using a Spectra/Por 6 1000 Molecular weight cut-off (MWCO) 38 mm width membrane. Gel Permeation Chromatography (GPC) was performed with a Shimadzu Prominence LC-20A fitted with a Thermo Fisher Refractomax 521 Detector and a SPD20A UV-vis Detector. Boc-Protected poly (acryloyl hydrazide) (Boc-P) was analysed using 0.05 M LiBr in DMF at 60 °C, or 0.005 M NH_4BF_4 in DMF at 50 °C, as the eluent and a flow rate of 1 mL min^{-1}. The instrument was fitted with a Polymer Labs PolarGel guard column (50 × 7.5 mm, 5 μm) followed by two PLGel PL1110–6540 columns (300 × 7.5 mm, 5 μm). Molecular weights were calculated based on a standard calibration method using polymethylmethacrylate. Poly(acryloyl hydrazide) P was analysed using Dulbecco's Phosphate Buffered Saline 0.0095 M (PO_4) without Ca and Mg as the eluent and a flow rate of 1 mL min^{-1}. The instrument was fitted with an Agilent PL aquagel-OH column (300 × 7.5 mm, 8 mm) and run at 35 °C. A Guava EasyCyte™ cytometer (EMD Millipore) was used for all flow cytometry experiments. Cell microscopy images were acquired with an Andor Zyla 4.2 digital camera mounted on a Nikon Eclipse Ti-E epifluorescence microscope or with a Leica SP5 confocal microscope. Absorbance of formazan was measured in a microplate reader Tecan Infinite F200Pro. Gels were imaged on a ChemiDoc XRS+ system (Bio-Rad) and analysed with the software ImageLab 5.2.1 (Bio-Rad).

DLS and ζ potential measurements were performed in a Malvern Zetasizer NanoZSP using standard disposable cuvettes. All experiments were done in triplicate at 25 ºC.

3.3 Polymer synthesis

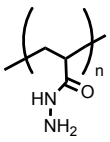


3.3.1 Poly(tert-butyl-2-acryloylhydrazine-1-carboxylate) (Boc-P)

A solution of 2,2'-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044) (23 mg, 70.0 μmol) in DMSO (1.5 mL) and a solution of 2-aminoethanethiol

(27.6 mg, 360 µmol) in DMSO (1.5 mL) were added sequentially to a solution of *tert*-butyl-2-acryloylhydrazine-1-carboxylate (3.3 g, 17.9 mmol) in DMSO (17 mL). A 50 µL aliquot of this solution was taken at this stage to aid in the calculation of conversion. The reaction mixture was then sealed and degassed with Argon for 30 min. The degassed solution was left to react at 65°C for 5 h. The reaction was stopped by allowing it to cool down to room temperature and by exposing it to air. A 50 µL aliquot of this solution was taken at this stage to aid in the calculation of conversion. The polymer was purified by dialysis against water. The water was removed by lyophilisation and by drying in a desiccator with P_2O_5 to afford 2.86 g Boc-P as an off-white powder (quantitative yield). UV (DMSO): λ_{max} 300 nm. ^{1}H-NMR (300 MHz, DMSO-d_6) δ (ppm): 9.9-7.4 (br-m, 2H, NH), 2.4-0.5 (br, 12H, 9H in $C(CH_3)_3$, 3H in $CHCH_2$). Conversion: 86%; Mn (DMF GPC) 196 KDa; Đ (DMF GPC) 2.76.

3.3.2 Poly(acryloyl hydrazide) P



Trifluoroacetic acid (TFA) (4 mL) was added dropwise to poly(*tert*-butyl-2-acryloylhydrazine-1-carboxylate) (Boc-**P**) (499 mg) and the yellow solution was stirred at r.t. overnight. Excess of TFA was removed by blowing a steady stream of Argon and the resulting oil was diluted in water (4 mL). The **P**·TFA salt formed was neutralised by adding $NaHCO_3$ until no foaming was observed. The colourless solution was allowed to stir overnight. The crude polymer was purified by dialysis against water. The water was removed by lyophilisation and by drying in a desiccator with P_2O_5 to afford 60 mg of **P** as a white powder (26%). ^{1}H-NMR (300 MHz, D_2O) δ (ppm): 1.2–2.3 (br m, (3·DP)H). Mw (Aqueous GPC) > 59 KDa.

3.4 Supporting figures

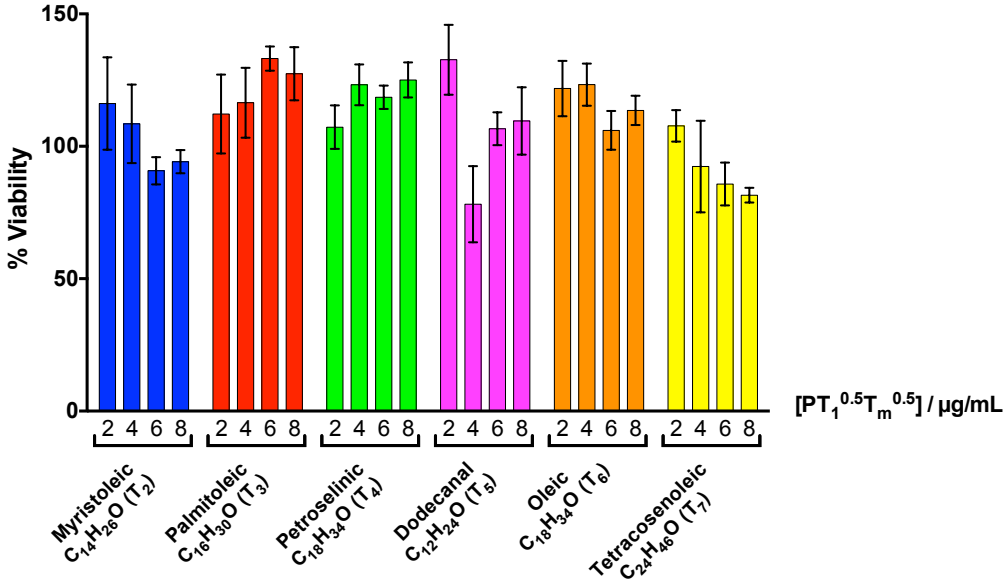


Figure S1. MTT viability assay. Hek293 cells were transfected with the indicated concentrations of the polyhydrazones and 1 ng/µL of mRNA, and 24 hpt cell viability was measured with a MTT colorimetric assay. Values were normalized to untreated cells.

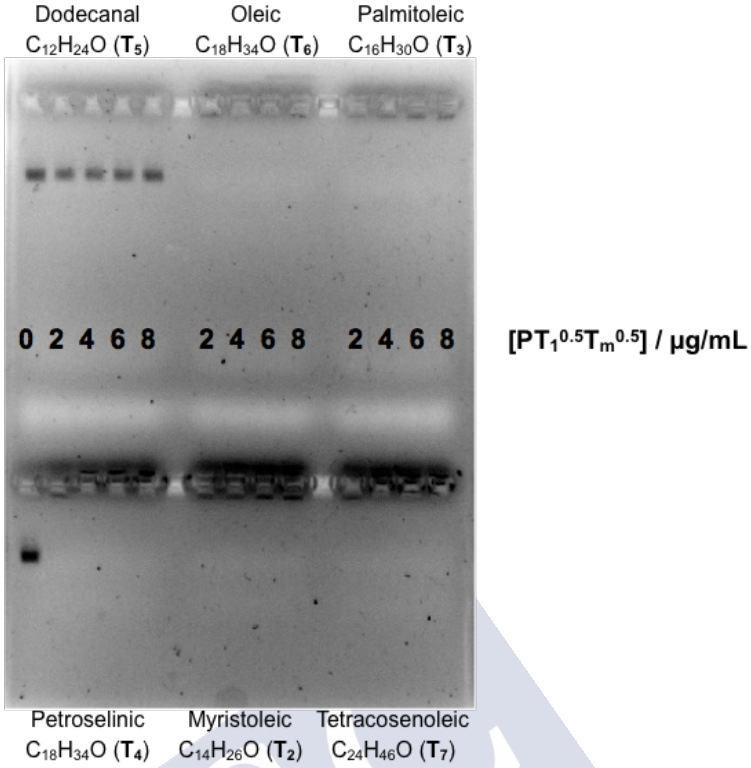


Figure S2. Gel retardation assay with the six different polyhydrazones. **EGFP-mRNA** and PT_1T_m/**EGFPmRNA** polyplexes with different concentrations of polyhydrazones: 2, 4, 6 and 8 µg/mL.

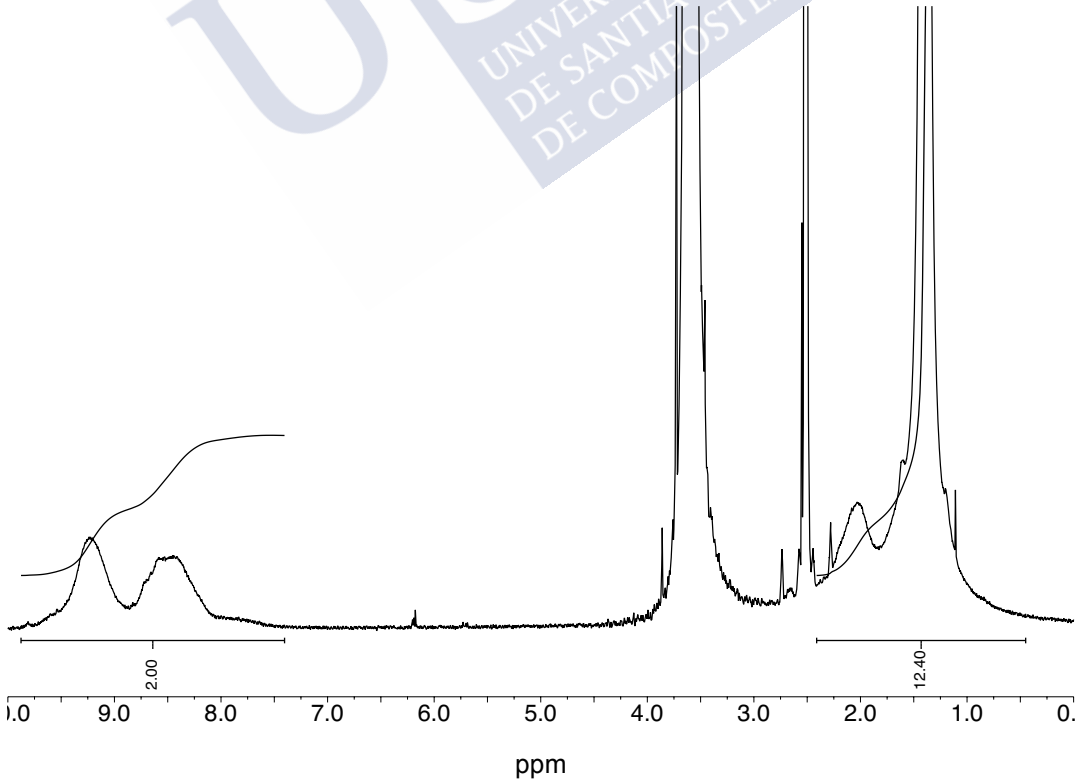


Figure S3. ^{1}H-NMR spectra for Boc-**P** in DMSO-d_6.

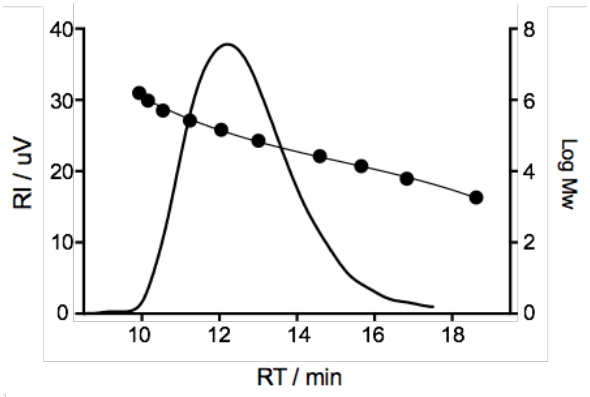


Figure S4. GPC Chromatogram of Boc-**P** in DMF.

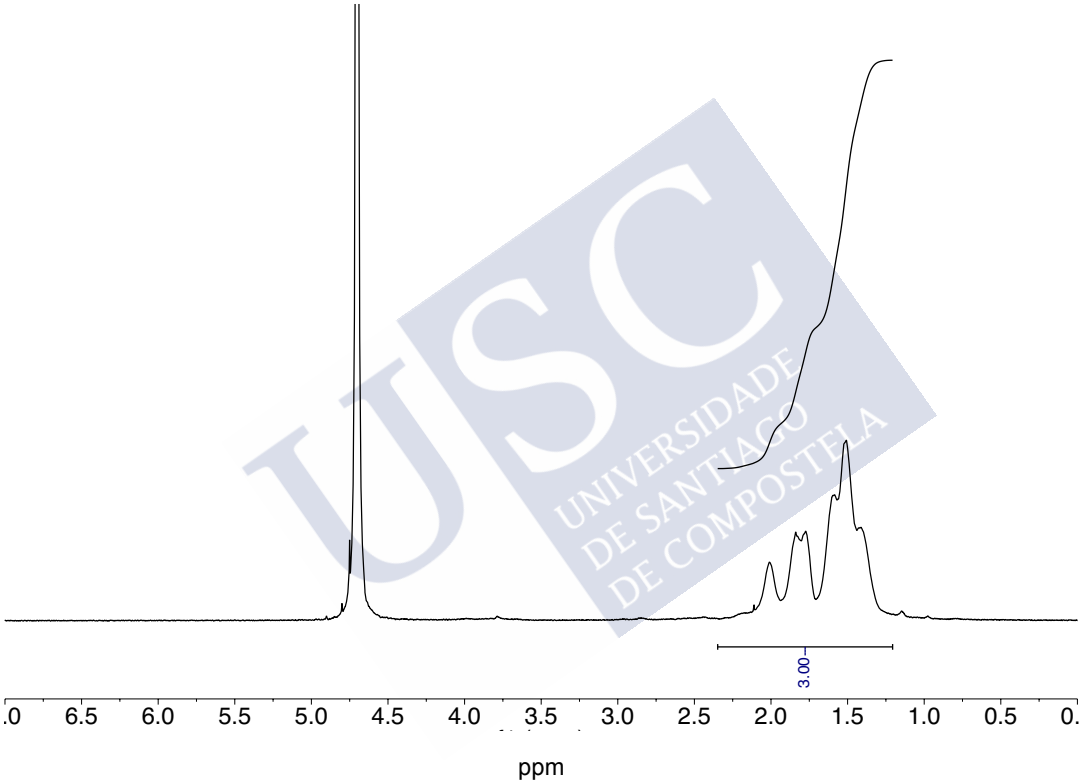


Figure S5. ^{1}H-NMR spectra for **P** in D_2O.

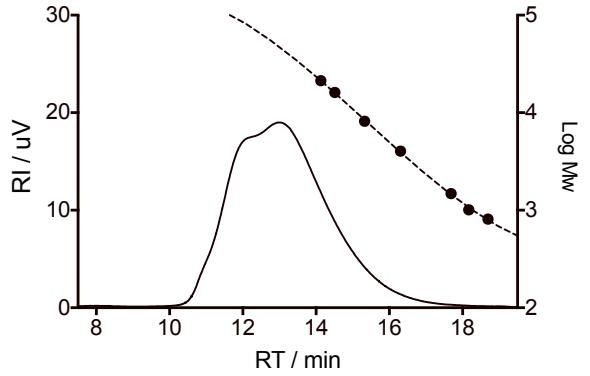


Figure S6. GPC Chromatogram of **P** in DPBS.

3.5 SUPPORTING REFERENCES

S1 R. A. Parente, S. Nir and F. C. Szoka, *Biochemistry*, 1990, **29**, 8720–8728.
S2 M. Juanes, I. Lostalé-Seijo, J. R. Granja and J. Montenegro, *Chem. Eur. J.*, 2018, **24**, 10689–10698.