This is the author's version of a work that was accepted for publication in Nanomedicine (Ed. Elsevier). Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Serna, N. et al. "Rational engineering of single-chain polypeptides into protein-only, BBB-targeted nanoparticles" in Nanomedicine,vol. 12, issue 5 (July 2016), p. 1241-1521. DOI 10.1016/j.nano.2016.01.004

JN2015192

Rational engineering of single-chain polypeptides into protein-only, BBB-targeted nanoparticles.

Naroa Serna [1, 2, 3], María Virtudes Céspedes [3, 4], Paolo Saccardo [1, 2, 3], Zhikun Xu [1, 2, 3], Ugutz Unzueta [3, 4], Patricia Álamo [3, 4], Mireia Pesarrodona [1, 2, 3], Alejandro Sánchez-Chardi [5], Mónica Roldán [5], Ramón Mangues [3, 4], Esther Vázquez [1, 2, 3], Antonio Villaverde [1, 2, 3 *], Neus Ferrer-Miralles [1, 2, 3 *]

[1] Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

[2] Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

[3] CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain

[4] Oncogenesis and Antitumor Drug Group, Biomedical Research Institute Sant Pau (IIB-SantPau), Hospital de la Santa Creu i Sant Pau, C/ Sant Antoni Maria Claret, 167, 08025 Barcelona, Spain

[5] Servei de Microscòpia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

E-mail addresses: NFM neus.ferrer@uab.cat ; AV antoni.villaverde@uab.cat

Tables: 2; Figures: 5; References: 44; Words in abstract: 147; Total words (excluding references): 5321

The authors declare no competing interests. We are indebted to FIS PI12/01861, Marató 416/C/2013-2030 and NanoMets to RM, MINECO BIO2013-41019-P, AGAUR (2014SGR-132) and CIBER de Bioingeniería, Biomateriales y Nanomedicina (project NANOPROTHER) to AV for funding our research on protein-based therapeutics. Z.X. and M.P. acknowledge financial support from China Scholarship Council and Universitat Autònoma de Barcelona through pre-doctoral fellowships respectively.

Abstract

A modular, single chain polypeptide containing the low density lipoprotein receptor (LDLR) ligand Seq-1 and with blood-brain barrier (BBB) crossing activity, has been successfully modified by conventional genetic engineering to self-assemble into stable protein-only nanoparticles of 30 nm. The nanoparticulate presentation dramatically enhances *in vitro,* LDLR-dependent cell penetrability of the protein compared to the parental monomeric version, but the assembled protein does not show any enhanced brain targeting upon systemic administration. While the presentation of protein drugs in form of nanoparticles is in general advantageous regarding correct biodistribution, this principle might not apply to brain targeting that is hampered by particular bio-physical barriers. Irrespective of this fact, that is highly relevant to the nanomedicine of central nervous system, engineering the cationic character of defined protein stretches is revealed here as a promising and generic approach to promote the controlled oligomerization of biologically active protein species as still functional, regular nanoparticles.

Keywords: Protein engineering; nanoparticles; self-assembling; biodistribution; LDLR

Introduction

The design and biofabrication of nanoscale materials that mimic viral properties (such as self-assembling, cell surface receptor binding, internalization and proper intracellular trafficking) is highly promising for targeted drug delivery and gene therapy [1]. Proteins are among the most convenient materials for the generation of functional building blocks in nanoparticle construction [2-5]. This is not only linked to protein functionalities but also to the fact that their spatial conformation, potential for cross-interactions and supramolecular organization can be designed and adjusted by simple genetic engineering. In addition, cost-effective biofabrication of proteins for therapeutic applications has been fully demonstrated in Bio-Pharma, and the emergence of novel cell factories and the implementation of genetic and systems approaches expand the opportunities for the biosynthesis of difficult proteins [6-8]. Then, while protein production and downstream are technically solved issues, the bases for a rational engineering of protein-protein cross molecular interactions remain to be fully established.

Some architectonic principles have been proposed for the construction of peptide-based nanofibers and nanoparticles [4], exploiting the amphiphilic character of chemically modified short peptides [9] or the self-assembling properties of the β-sheet-rich amyloidal protein domains [10]. Regarding long-lengthen proteins

suitable for biological production, the engineering of natural oligomerization domains into modular polypeptides has allowed the construction of protein nanoparticles based on hybrid fusion proteins acting as building blocks [11]. Of course, the production of structural viral proteins renders, in some cases, complex structures known as virus like particles (VLPs) that architectonically mimic the natural viral capsid versions. While largely proved to be excellent immunogens, the usability of VLPs and other nanostructured materials as vehicles for drug and nucleic acid delivery is rather narrow [12]. This is because the limited versatility in the design of nanoscale physical properties and biological functions of the resulting nanoparticles, and on the other side, by the still partial comprehension of the mechanics of the cross-molecular interactions that govern the formation of stable supramolecular complexes.

Recently, we have proposed a nanoscale architectonic principle that permits the generation of protein nanoparticles in which structurally unrelated protein species can act as building blocks [13]. When tagged with a cationic amino terminal peptide and a polyhistidine tail at the carboxy terminus, these scaffold proteins (eg. GFP, IRFP, p53 and Hsp 70), cross-interact by electrostatic interactions and form toroid structures stabilized by a complex set of alternative forces, including van der Wals and hydrogen bond interactions [14]. These nanoscale materials, of regulatable size, allow the intracellular delivery of functional proteins to specific target cells and tissues upon systemic administration [14-16]. In form of nanoparticles and when conveniently

empowered by tumor homing peptides, these proteins correctly biodistribute and accumulate, for instance, in $CXCR4^+$ cancer stem cells in colorectal cancer models [14]. However, recent indirect data suggested that the presentation of proteins as nanoparticulate entities might not favour protein delivery to brain [17]. This might be indicative of different principles governing the biodistribution of protein nanoparticles depending on if they display tumor- or BBB-homing peptides. To assess BBB-crossing and brain targeting properties of monomeric and nanoparticle versions of the same targeting peptide we have engineered for the first time, pre-existing BBB-homing polypeptides into building blocks that self-assemble as equivalent protein nanoparticles. Then, apart from determining the *in vitro* cellular penetrability, the biodistribution of both disassembled and assembled versions has been examined *in vivo* upon intravenous administration into healthy mice. Unexpectedly, while the cellular penetrability of protein nanoparticles is enhanced *in vitro* when compared to single molecular species, this is not accompanied by an enhanced ability of the protein to reach the brain. In addition to the protein engineering principles successfully tested here, the presented results indicate particular restrictions in nanoparticle performance in brain targeting, what should prompt the careful reconsideration of nanotechnological approaches to neurotropic vehicles for therapies and imaging.

Materials and methods

*Protein design and gene cloning***.** Angiopep-2 (TFFYGGSRGKRNNFKTEEY) [18] and Seq-1 (KYLAYPDSVHIW) [19], are peptides with known BBB-crossing activities, that have been previously used to construct brain-targeted GFP fusions. Derivatives of *Angiopep-2-GFP-H6* and *Seq-1-GFP-H6* containing additional cationic amino acids (Table 1) were designed in house and obtained from Genscript (Piscataway, USA). The synthetic genes (*Seq-1-7-GFP-H6, Seq-1-8-GFP-H6, Angiopep-2-7-GFP-H6* and *Angiopep-2-8-GFP-H6)* were then inserted into the prokaryotic expression commercial pET-22b vector (#69744-3, Novagen, USA) using *Nde*I/*Hind*III restriction sites. *Escherichia coli* BL21 (DE3) cells were transformed with recombinant pET22b plasmids by heat shock (45 sec, 42ºC) to allow protein production.

Protein production and purification. Bacterial cells carrying the appropriate plasmid vector were cultured in 2L shaker flasks with 500 ml of LB (Luria-Bertani, Conda Cat. 1551.00) medium containing 100 µg/ml ampicillin, at 37 ºC until the OD_{550} reached 0.5-0.7. Recombinant gene expression was induced with 0.1 mM isopropyl-β-d-thiogalactopyronaside (IPTG) and then, bacterial cells were kept growing overnight at 20 °C. Bacterial cells were then harvested by centrifugation at 5,000 g for 15 min at 4 °C and resuspended in Wash buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM imidazol, pH 8.0) PBS (phosphate buffered saline: 140 mM NaCl, 2.7 mM KCl, 10 mM

Na2HPO4, 1.8 mM KH2PO4) in the presence of EDTA-free protease inhibitor (Complete EDTA-Free; Roche). Cells were then disrupted in a French Press (Thermo FA-078A) and centrifuged for 45 min (15,000 g at 4ºC). All proteins were purified by His-tag affinity chromatography using HiTrap Chelating HP 1 ml columns (GE healthcare) by ÄKTA purifier FPLC (GE healthcare). After filtering the soluble fraction, samples were loaded onto the column and washed with 10 column volumes of Wash buffer. Bound proteins were eluted with Elution buffer (20 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, pH 8.0) in a linear gradient. Purified fractions were collected and quantified by Bradford's assay and analyzed by SDS-PAGE and MALDI-TOF. The medium-scale production of recombinant proteins was partially performed by the ICTS "NANBIOSIS", more specifically by the Protein Production Platform of CIBER in Bioengineering, Biomaterials & Nanomedicne (CIBER-BBN)/ IBB, at the UAB (SepBioEs, http://www.ciber-bbn.es/es/programas/89-plataforma-de-produccion-de-proteinas-ppp).

*Microdialysis***.** Drops of purified proteins (20 µl) were deposited on VsWp02500 Millipore membrane filters floating on 20 ml of different buffers: Carbonate buffer (166 mM $NaHCO_3$, pH 7.5), Carbonate + dextrose 5% (166 mM $NaHCO_3$ pH 7.5 + 5 % dextrose), Carbonate + salt (166 mM $NaHCO_3$ pH 7.5 + 200 mM NaCl) and HBS buffer 10x (HEPES-buffered saline pH 5.8; 50 mM KCl, 1.37 M NaCl, 8.5 mM Na_2HPO_4, 21 mM HEPES).

Dialyzed drops were collected and centrifuged, and the soluble fractions were quantified to determine the extent of aggregation for each protein in different buffers. For further storage, proteins were finally dialyzed against their own most convenient buffer regarding protein stability (Seq-1-7-GFP-H6 and Seq-1-8-GFP-H6 against Carbonate buffer and Angiopep-2-7-GFP-H6 and Angiopep-2-8-GFP-H6 against Carbonate + 200 mM NaCl buffer) and stored at -80 C after 0.22 pore membrane filtration.

Fluorescence determination and dynamic light scattering (DLS). Volume size distribution of nanoparticles and monomeric GFP protein fusions was determined by DLS at 633 nm (Zetasizer Nano ZS, Malvern Instruments Limited, Malvern, UK). Fluorescence was determined in a Cary Eclipse fluorescence spectrophotometer (Varian Inc, Palo Alto, CA) at 510 nm using an excitation wavelength of 450 nm.

For DLS analyses, proteins (stored at -80 C) were thawed and incubated on ice until use. 50 μl of each sample were used in the corresponding stock concentration: Seq-1-7-GFP-H6: 2.67 mg/ml; Seq-1-8-GFP-H6: 2.2 mg/ml ; Angiopep-2-7-GFP-H6: 3.93 mg/ml and Angiopep-2-8-GFP-H6: 2.58 mg/ml. For fluorescence determination, protein samples were diluted in the corresponding storage buffer to 0.5 mg/ml, in a final volume of 100 μl. The shown data refers to the volume peak of the materials (the mode). For some experiments,

protein samples were diluted to 1.5mg/ml in reconstituted human plasma (Sigma-Aldrich, Ref: S225-5 ml), and further incubated under agitation at 37 ºC.

Electron microscopy (TEM and FESEM). For transmission electron microscopy (TEM) two sets of purified proteins were diluted to 0.2-0.5 mg/ml, deposited onto carbon-coated copper grids (300 mesh) and one of them contrasted by uranyl acetate and air-dried and the other one shadowed with 1 nm a platinum-carbon layer respectively. Samples were observed in a Jeol 1400 transmission electron microscope operating at 80kV and equipped with a CCD Gatan Erlangshen ES1000W camera. For field emission scanning electron microscopy (FESEM), microdrops of purified proteins diluted as for TEM were deposited onto a silicon wafer surface, air-dried and immediately observed in a Zeiss Merlin field emission scanning electron microscope operating at 1 kV and equipped with a high resolution in-lens secondary electron detector.

Cell culture and flow cytometry. $LDLR^+$ HeLa cells (ATCC-CCL-2) were used for the *in vitro* experiments. The cell line was cultured in Eagle's Minimum Essential Medium (Gibco, Rockville, MD) supplemented with 10 % fetal calf serum (Gibco), and incubated at 37°C and 5 % CO_2 in a humidified atmosphere. Meanwhile bEnd.3 cell line was maintained in Dulbecco's Modified Eagle's Medium (DMEM: Gibco® GlutaMAX™) supplemented with 10 % fetal calf serum (Gibco), and incubated at

37°C and 10 % CO_2 in a humidified atmosphere. Nanoparticles and proteins were added at different concentrations (ranging from 2 to 0.1 mM) to the cell culture in the presence of Optipro medium (Gibco) 24 h before flow cytometry. Cell samples were analyzed on a FACSCanto system (Becton Dickinson) using a 15 W air-cooled argon-ion laser at 488 nm excitation. GFP fluorescence emission was measured with a detector D (530/30 nm band pass filter) after treatment with 1 mg/mL trypsin (Gibco) for 15 min. For time course analysis, nanoparticles were added at 2 mM final to the cell culture in the presence of Optipro medium (Gibco) 24, 6, 4, 2 and 1 h, and 30 and 15 min before flow cytometry. Cell samples were analyzed after harsh treatment with 1 mg/mL trypsin (Gibco) for 15 min to remove surface attached fluorescent protein [20].

Confocal microscopy. HeLa cells were grown on Mat-Teck culture dishes (Mat Teck Corporation, Ashland, MA). The nuclei were labeled with 10 µg/mL Hoechst 33342 (Molecular Probes, Eugene, OR) and the plasma membranes with 5 µg/mL CellMaskTM Deep Red (Molecular Probes) in darkness for 10 min. Cells were washed in phosphate-buffered saline (Sigma-Aldrich Chemie GmbH, Steinheim, Germany) and proteins were added 20–24 h before staining at 2 mM. Live cells were recorded by TCS-SP5 confocal laser scanning microscopy (Leica Microsystems, Heidelberg, Germany) using a Plan Apo 63 ×/1.4 (oil HC × PL APO lambda blue) objective. To determine the location of particles inside the cell, stacks of 10–20 sections were collected at 0.5

µm Z-intervals with a pinhole setting of 1 Airy unit. The 3-D reconstruction was performed using Imaris software (Bitplane, Zürich, Switzerland).

Protein stability in human plasma. The stability of proteins was analyzed by measuring fluorescence emission after incubation in human plasma. Proteins were diluted, in triplicate, in human plasma (0.115 µg/µL final concentration). Human plasma was obtained from a healthy donor in the Hospital de Sant Pau complying with the regulation established by the Clinical Ethics Committee of the Institution.
Immediately after dilution, samples were harvested (time 0) and their respective fluorescence was taken as the initial reference value (100 %). Proteins were further incubated (at 37 ºC, with agitation), and samples were taken, at different time points, up to 24 h. Protein functional stability during incubations was analyzed by fluorescence determination at 510 nm in a Cary Eclipse fluorescence spectrophotometer (Variant, Inc., Palo Alto, CA) using an excitation wavelength of 450 nm.

*Biodistribution***.** Female athymic *nu/nu* mice (Charles River, L´Abresle, France) between 4 to 6 weeks of age were housed in individually-ventilated cages on a 12-h light-dark cycle at 21-23 ºC and 40-60 % humidity. Mice were allowed free-access to an irradiated diet and sterilized water. The experimental animals received a single

intravenous bolus of 500 µg of Seq-1-derived protein nanoparticles (or the equivalent unassembled proteins) in 166 mM carbonate buffer pH 7.5. Control animals received a single bolus of empty buffer. At 30 min and 2 h post-administration, we sacrificed the mice and measured *ex vivo* the amount of nanoparticles in each relevant organs from the experimental and control mice, by quantifying the emitted fluorescence. Entire hemisected organs (brain, kidney, liver, lung, heart) were placed in separate wells to detect the emitted signal using IVIS® Spectrum equipment (PerkinElmer Inc, Waltham, MA). The fluorescence signal was first digitalized, displayed as a pseudocolor overlay, and finally expressed as radiant efficiency. Differences in signals were analyzed by a non-parametric Man-Whitney test. The study was approved by the institutional ethical committee and all procedures were in accordance with institutional guidelines. Humane care of the animals was always applied accordingly.

Results

Protein design, production and characterization. Angiopep-2-GFP-H6 and Seq-1-GFP-H6 are modular proteins displaying low density lipoprotein receptor (LDLR) ligands at their amino termini (Angiopep-2 and Seq-1 respectively) (Table 1), that reach the brain upon systemic administration [17]. Both Angiopep-2 and Seq-1 are well known BBB-crossing peptides proven useful as brain-homing agents in drug delivery [18, 19]. At difference of the related construct ApoB-GFP-H6 that

form nanoparticles of 18 nm in diameter [17], these proteins do not show any self-assembling activity and remain monomeric. In an attempt to upgrade Angiopep-2-GFP-H6 and Seq-1-GFP-H6 polypeptides up to self-assembling building blocks for nanoparticle construction, arginine- and lysine- containing stretches were designed for their introduction in both proteins, between the ligand and the linker to the scaffold GFP (Table 1, segments in bold). Two versions of the supplementary cationic peptides were constructed to offer a total of 7 or 8 positively charged residues in the amino terminal region. According to previous numerical modelling [15], these numbers of cationic residues, if properly placed, should enable the proteins to self-organize as nanoparticles of around 30 nm. This numerical model identifies a relationship between the number of cationic residues at the amino terminus of the building blocks, and the size of the engendered particles, and it resulted from empirical data recruited from a set of different modular proteins. So far, the model had been never tested as a rational tool for the engineering of protein self-assembling.

The four new versions of Angiopep-2-GFP-H6 and Seq-1-GFP-H6 were produced in bacteria resulting in full-length soluble species with predicted molecular masses (Table 1) and retaining the GFP fluorescence emission. The reduction of the fluorescence emission capacity, that is clearly significant (p<0.01) in the case of Seq-1-derivatives, is indicative of moderate conformational changes in the protein that while keeping the barrel organization might affect the

conformation of the chromophore. While the parental versions sized around 5-6 nm by DLS analysis, compatible with monomeric or dimeric GFP forms, the protein Seq-1-8-GFP-H6 was observed as nanoparticulate entities of around 30 nm, significantly different in size from unassembled monomers (Figure 1 A) and indicating a supramolecular organization of the individual polypeptide chains. The formed nanoparticles occurred and were stable in different buffers against which the protein was dialyzed (Figure 1 B), proving the structural robustness of the constructs under different salt contents and ionic strengths. In agreement, and fully supporting the taken approach to the engineering of building blocks, TEM and FESEM examinations confirmed the occurrence of these proteins as nanoparticles of regular size and morphology (Figure 1 C-E). Interestingly, high magnification TEM images of Seq-1-8-GFP-H6 showed an architectonic scheme (a circular electrodensity pattern) compatible with a toroid organization, similar to that recently shown for the related protein T22-GFP-H6 [14]. This particular geometry was fully confirmed by high resolution FESEM analysis of the same samples (Figure 1 E). As expected, no nanoparticles but some unspecific aggregates with different shape and size were observed when examining Seq-1-7-GFP-H6 samples (Figure 1 C). As shown in Figure 1A, none of the Angiopep-2-GFP-H6-derived proteins formed nanoparticles.

In vitro analysis of cellular internalization. At this stage, we were interested in the biological characterization of Seq-1-8-GFP-H6 nanoparticles regarding their ability to penetrate target cells, and especially in comparison with the closely related, unassembled protein Seq-1-7-GFP-H6 that only differs by a single missing arginine residue (Table 1). Before that, we determined that the fluorescence of Seq-1-8-GFP-H6 nanoparticles and of the whole set of unassembled proteins remained stable in human plasma *in vitro* at 37 ºC for at least 24 h (Figure 2 A). The protein stability in presence of a complex organic medium prompted us to use fluorescence as a marker to monitor cell penetrability by flow cytometry, in LDRL-expressing HeLa cells. As observed (Figure 2 B), Seq-1-8-GFP-H6 penetrated HeLa cells much more efficiently than the closely related construct Seq-1-7-GFP-H6 and that the monomeric versions of Angiopep-2-displaying proteins. Also, the levels of Seq-1-8-GFP-H6 internalization were comparable to those observed for the related R9-GFP-H6 nanoparticles used here as control, that are empowered by a potent Tat-inspired cell penetrating peptide (nine arginines, R9). R9 promotes receptor-independent fast cell penetrability and nuclear migration in HeLa and other cell types [21-23]. The similar uptake of Seq-1-8-GFP-H6 was indicative of an enhanced penetrability of the Seq-1 ligand when organized in a nanoparticulate form. Confocal images of HeLa cells exposed to Seq-1-derived proteins confirmed the poor penetrability of Seq-1-7-GFP-H6 (Figure 3 A) and the intracellular location of Seq-1-8-GFP-H6 (Figure 3 B), discarding mere external association to the cell

membrane. Also, 3D reconstructions of exposed cultures resulted in images compatible with the expected endosomal route for Seq-1-8-GFP-H6 uptake, based on the merging of red (membranes) and green (nanoparticles) signals (Figure 3 C).

To evaluate if the enhanced cell internalization of Seq-1-8-GFP-H6 (in comparison to Seq-1-7-GFP-H6) would impact on the BBB-crossing abilities of Seq-1 we first tested both proteins in a CaCo-2 cell permeability system recognized as a relevant *in vitro* model of the BBB [24]. In this test, Seq-1-8-GFP-H6 showed transcytic properties clearly superior to those exhibited by Seq-1-7-GFP-H6 (Table 2), again in the line of the higher cell penetrability of a nanoparticulate version of the homing peptide Seq-1.

Biodistribution and BBB-crossing

When administered intravenously in healthy mice models with intact BBB, both Seq-1-7-GFP-H6 and Seq-1-8-GFP-H6 transiently accumulated in brain (Figure 4A), over the background values derived from GFP-H6 administration (Figure 5 B). Despite being more efficient in cell penetration *in vitro* (Figure 2 and 3, and Table 2), the nanoparticulate version of Seq-1, namely Seq-1-8-GFP-H6, did not show any improvement of CNS targeting regarding the unassembled Seq-1-7-GFP-H6 form. On the other hand, any of these proteins were found in liver, heart and lung, but in kidney (Figure 5 A). The occurrence in kidney was not unforeseen in the cases of GFP-H6 and

the unassembled Seq-1-7-GFP-H6, which are cleared by renal glomeruli. However, Seq-1-8-GFP-H6 nanoparticles would be expected to escape from renal filtration and occur only in target organs, as observed in the case of similar constructs functionalized with tumor-homing peptides [15]. This could be indicative of a partial disassembling of Seq-1-8-GFP-H6 nanoparticles *in vivo,* due to the high ionic strength and the presence of a complex spectrum of interacting human proteins, and a consequent reduction of the Seq-1-8-GFP-H6 material size from 30 nm (nanoparticles) to less than 8 nm (unassembled building blocks). While this does not occur in the case of the structurally related, more cationic T22-IRFP-H6 nanoparticle [14], that is fully stable during circulation in blood, it would be not unexpected in the case of the less cationic Seq-1-8-GFP-H6, since the absence of one single lysine (in Seq-1-7-GFP-H6) precludes self-assembling. To check this possibility, we explored *in vitro* the nanoparticle stability in presence of human sera, at 37 ºC, to mimic post injection conditions. As observed (Figure 5B), Seq-1-8-GFP-H6 nanoparticles loosed stability with incubation time, showing a moderate reduction of size at 2 h and resulting essentially disassembled at 5 h. Partial *in vivo* disassembly at 2 h would, at least to some extent, account for the presence of the material in kidney. However, while the amount of monomeric GFP-H6 and Seq-1-7-GFP-H6 dramatically decreased from 30 min to 2 h post administration, as expected for materials ≤ ~8 nm in size, Seq-1-8-GFP-H6 nanoparticles remained in kidney for longer times (Figure 5C). This fact suggested

issues additional to size-dependent filtration as contributors in regulating its occurrence in kidney.

Discussion

By using empirical information from of a set of closely related proteins with and without self-assembling properties [25], we have designed and constructed protein-only nanoparticles containing the Seq-1 BBB-crossing peptide (Figure 1). This has been done by the addition of several cationic peptides at the amino terminus of the modular protein Seq-1-GFP-H6 (Table 1), a polypeptide that exclusively remains in unassembled form (Figure 1). A similar attempt to engineer the self-assembling of Angiopep-2-GFP-H6 failed (Figure 1) probably because the more scattered distribution of cationic residues (Table 1). This would result in a less polar splitting of the electrostatic charges necessary for the assembling of building blocks [13-15]. In fact, one additional cationic lysine residue in Seq-1-8-GFP-H6 compared to Seq-1-7-GFP-H6 was sufficient to efficiently triggering self-assembling (Table 1, Figure 1). Seq-1-8-GFP-H6 nanoparticles exhibited the expected size of 30 nm (Figure 1 A), a toroid organization (Figure 1 C, E) and a higher cellular penetrability into LDLR$^+$ HeLa cells when compared with the unassembled Seq-1-7-GFP-H6 version (Figure 2, 3). In addition, the transcytic properties of Seq-1-8-GFP-H6, measured in a CaCo-2-based test [24], were clearly superior to those exhibited by the unassembled form Seq-1-7-GFP-H6 (Table 2). Since the mere addition of cationic peptides to pre-existing protein nanoparticles did not alter cell penetrability *per se* [26], the formation of nanoparticles rather that the single amino acid addition in Seq-1-8-GFP-H6 would be responsible for enhanced cell penetrability. A higher penetrability of

protein nanoparticles compared with individual building blocks is in agreement with previous results obtained in our laboratory [27] and it is probably related with the multivalent presentation of the cell ligands (Seq-1, in the present case) on the surface of nanoparticles that favours endosomal entrapment [28]. The endosomal route of Seq-1-8-GFP-H6 nanoparticle uptake is confirmed by the strong merging signals (yellowish spots, merging green and red) observed in *in vivo* internalization assays (Figure 3 B, C). Of course, the mere enlargement of the material size associated to oligomerization might also promote endocytosis, as size of nanoparticles generically determines the nature of cell responses upon exposure, including signalling and endocytosis [29]. Among nanoparticles between 2 and 100 nm, those sizing between 40 and 50 nm show the strongest effect on cell responses, representing a size range compatible with what has been generically found as optimal for cell internalization in other independent studies [30-32].

Despite the dramatic improvement of cell penetrability and transcytosis exhibited by the nanoparticulate versus free form of Seq-1 homing peptide, the accumulation in brain of Seq-1-7-GFP-H6 and Seq-1-8-GFP-H6 were similar and indistinguishable (Figure 4). This fact indicates that the multivalent presentation of the Seq-1 LDLR ligand in organized nanoparticles has no relevant impact on BBB permeability in healthy animals. This is in agreement with recent data [17] regarding the poor brain penetrability of nanoparticles empowered with the BBB-crossing peptide ApoB [33], and it could be due to different

transcytosis activities between epithelial kidney cells and brain endothelial cells, as previously described for a different receptor [34].

On the other hand, both Seq-1-7-GFP-H6 and Seq-1-8-GFP-H6 were found in kidney (but not in other organs), as in the case of the parental protein GFP-H6 (Figure 5A). This could be interpreted by assuming an imperfect stability of Seq-1-8-GFP-H6 nanoparticles that might (at least partially) disassemble once in the blood stream. In fact, *in vivo* disassembling of protein nanoparticles is not totally unexpected since it has been already described that nab-paclitaxel nanoparticles (known as Abraxane and sizing 130 nm), that are fully stable in physiological buffers and in saline solution, quickly separated into individual albumin-paclitaxel complexes once injected into the blood stream [35]. In this context, Seq-1-8-GFP-H6, that exhibited the same size (30 nm) in human sera than in buffer (Figure 1), appeared only as partially stable in sera under prolonged incubation (Figure 5B), as a reduction of particle size was observed after 2 h of incubation and a significant occurrence of monomers at 5 h, a later time point than those considered for biodistribution analyses (30 min and 2 h) . On the other hand, the presence of heterologous proteins in human sera, would also promote, at east at some extent, structural or functional instability of the material by the presence of the corona [36]. However, the dynamic analyses of the renal occurrence of the proteins analysed here (Figure 5C) shows clear and significant differences when comparing the monomeric versions (GFP-H6 and Seq-1-7-GFP-H6) and the multimeric form Seq-1-8-GFP-H6. This last

protein remains for longer time in kidney, suggesting that in addition to a potential size-dependent renal filtration of part of the material, its occurrence in the tissue might rely on a specific interaction between Seq-1 and tissue specific receptors such as megalin. Megalin, a cell surface receptor of the LDLR family found in renal cell types, participates in lipid metabolism in the kidney [37], showing a high capacity for taking up lipid-binding proteins into the renal proximal tubule epithelial cells where it is highly expressed [38]. It is known that protein uptake by the proximal tubules is the main process ensuring the lack of proteins in urine. Moreover, since it occurs distally to glomerular filtration, only proteins of a size lower than the filtration cut-off (8 nm) can be taken up [39]. Despite this, some receptor expression might occur in the endothelia of the glomerulus, accounting for a certain level of specific binding of nanomaterials over 8 nm. Although this hypothesis would obviously need further solving, the abundance of LDLR in kidney might discourage the use of LDLR ligands in nanoparticles targeted to brain, especially if they occur in sizes close to the renal cut-off.

In this study, we have engineered the brain-targeted polypeptide Seq-1-GFP-H6 to self-assemble into regular, protein-only nanoparticles formed by multiple copies of the resulting protein Seq-1-8-GFP-H6. Seq-1-8-GFP-H6 differs from the parental protein by 7 cationic residues that were inserted into a permissive region of the protein, between the Seq-1 peptide and the core GFP (Table 1). The resulting fluorescent nanoparticles show a dramatic improvement of cell

penetrability and transcytic activities *in vitro* regarding monomeric versions, but contrarily, the brain targeting properties did not result improved by the multimerization (Figure 4). The engineering principle developed here should permit the construction of protein-based nanoparticles by the precise sequence manipulation of pre-existing proteins that increase, in a controlled way, the cationic load of the amino terminal regions. Also, the progressive developments in systems and genetic approaches to recombinant protein production [40] increasingly facilitate protein engineering, biofabrication in cell factories [41] and downstream [41,42]. In the current bubbling context of novel protein drugs of interest in cancer therapies and for other conditions [44], formulating such therapeutic proteins as nanoparticles would offer an interesting engineering tool with a broad applicability in nanomedicine, as the multivalent presentation of homing agents in those materials, either natural or added, is expected to dramatically enhance cell uptake.

Acknowledgments: Protein production has been partially performed by the ICTS “NANBIOSIS”, more specifically by the Protein Production Platform of CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN)/ IBB, at the UAB (http://www.ciber-bbn.es/en/programas/89-plataforma-de-produccion-de-proteinas-ppp). We thank the CIBER-BBN Nanotoxicology Unit for fluorescent in vivo follow-up using the IVIS equipment. We are also indebted to the Cell

Culture and Citometry Units of the Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria (SCAC), and to the Servei de Microscòpia, both at the UAB, and to the Soft Materials Service (ICMAB-CSIC/CIBER-BBN). AV received an ICREA ACADEMIA award.

Legends

Figure 1. Formation and structure of Seq-1-8-GFP-H6 nanoparticles. **A**) Size of Angiopep-2- and Seq-1-derived proteins measured by DLS in Carbonate buffer + 200 mM NaCl and Carbonate buffer respectively. The size of Seq-1-8-GFP-H6 nanoparticles is significantly different from the size of both Seq-1-7-GFP-H6 monomers and Angiopep-2 derivatives ($p<0.01$). **B**) Stability of Seq-1-8-GFP-H6 nanoparticle in different physiological buffers. **C-D**) Ultrastructural morphology of Seq-1-8-GFP-H6 nanoparticles at three different magnifications compared to the unassembled related protein Seq-1-7-GFP-H6 using shadowing of Pt/C (C) and negative staining (D) both with TEM, and deposition in silicon wafers for FESEM (E). Bar size is 50 nm.

Figure 2. Biological characterization of modular proteins. **A**) Time-dependent emission stability of protein constructs in human plasma. **B**) Penetrability of the different protein constructs in cultured $LDLR^+$ HeLa cells, determined by flow cytometry upon harsh trypsin treatment [20]. Crude fluorescence values were corrected by the specific emission of each protein for comparative purposes. Data refers to 24 h post exposure.

Figure 3. Internalization of protein constructs. Localization of protein Seq-1-7-GFP-H6 (**A**) and of Seq-1-8-GFP-H6 nanoparticles (**B**) upon 24 h of exposure to cultured HeLa cells, observed by fine confocal

sections at two magnifications (top and bottom). Nuclear DNA is stained in blue, the membranous system in red, while nanoparticles and unassembled proteins are seen in green since they are naturally fluorescent. **C**) Endosomal localization of Seq-1-8-GFP-H6 nanoparticles as observed by the yellowish merging of green and red signals in 3D confocal reconstructions.

Figure 4. Brain targeting of Seq-1-derived proteins upon systemic administration at 500 µg dose. **A)** BBB-crossing registered ex vivo by GFP fluorescence determination in mouse whole brain (W) and brain sections (S) 30 min and 2 h after intravenous administration. Crude fluorescence values were corrected by the specific emission of each protein for comparative purposes. **B)** GFP fluorescence values in brain sections. Data are normalized by the specific fluorescence. The scale and fluorescence signal values are expressed as Radiant efficiency [p/sec/cm2/sr] $\mu W/cm^2$. Those groups showing significant differences with the rest of the groups are labelled with symbols (*, # $p<0.05$).

Figure 5. Renal clearance and biodistribution of Seq-1-derived proteins upon systemic administration of 500 µg. **A)** *Ex vivo* determination of GFP fluorescence signal in kidneys and other relevant organs such livers, hearts and lungs, at the same post-injection times. Crude fluorescence values were corrected by the specific emission of each protein for comparative purposes**. B)** Size variation of Seq-1-8-GFP-H6 materials during incubation in human plasma. **C)** GFP fluorescence values in kidney sections. Data are

normalized by the specific fluorescence. The scales and fluorescence signal measures are expressed as Radiant efficiency [p/sec/cm2/sr] μW/cm^2. Those groups showing significant differences with the rest of the groups are labelled with symbols (# $p<0.05$).

Reference List

1 Mastrobattista,E. *et al.* (2006) Artificial viruses: a nanotechnological approach to gene delivery. *Nat. Rev. Drug Discov.* 5, 115-121

2 Elsadek B1, Kratz F. (2012) Impact of albumin on drug delivery--new applications on the horizon. *J Control Release* 157, 4-28

3 Neus Ferrer-Miralles *et al.* (2015) Engineering protein self-assembling in protein-based nanomedicines for drug delivery and gene therapy. *Crit Rev. Biotechnol* 35, 209-221

4 Corchero,J.L. *et al.* (2014) Recombinant protein materials for bioengineering and nanomedicine. *Nanomedicine. (Lond)* 9, 2817-2828

5 Desa, N. (2012) Challenges in development of nanoparticle-based therapeutics. *The AAPS Journal.* 14, 282-295

6 de Marco.A. (2013) Recombinant polypeptide production in E. coli: towards a rational approach to improve the yields of functional proteins. *Microb. Cell Fact.* 12, 101

7 Ferrer-Miralles,N. and Villaverde,A. (2013) Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microb. Cell Fact.* 12, 113

8 Mahalik,S. *et al.* (2014) Genome engineering for improved recombinant protein expression in Escherichia coli. *Microb. Cell Fact.* 13, 177

9 Cui,H. *et al.* (2010) Self-assembly of peptide amphiphiles: from molecules to nanostructures to biomaterials. *Biopolymers* 94, 1-18

10 Gazit,E. (2007) Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chem. Soc. Rev.* 36, 1263-1269

11 Ren,D. *et al.* (2011) Protein nanocapsules containing doxorubicin as a pH-responsive delivery system. *Small* 7, 1051-1060

12 Ramqvist T., Andreasson K., Dalianis T. (2007) Vaccination, immune and gene therapy based on virus-like particles against viral infections and cancer. *Expert Opin Biol Ther*. 7, 997-1007.

13 Unzueta,U. *et al.* (2012) Non-amyloidogenic peptide tags for the regulatable self-assembling of protein-only nanoparticles. *Biomaterials* 33, 8714-8722

14 Cespedes,M.V. *et al.* (2014) In Vivo Architectonic Stability of Fully de Novo Designed Protein-Only Nanoparticles. *ACS Nano.* 8, 4166-4176

15 Pesarrodona,M. *et al.* (2014) Intracellular targeting of CD44 cells with self-assembling, protein only nanoparticles. *Int. J. Pharm.* 473, 286-295

16 Unzueta,U. *et al.* (2012) Intracellular CXCR4^{+} cell targeting with T22-empowered protein-only nanoparticles. *Int. J. Nanomedicine* 7, 4533-4544

17 Xu *et al.* (2015) Targeting low-density lipoprotein receptors with protein-only nanoparticles. *Journal of Nanoparticle Research* 17, 150

18 Demeule M., Currie JC., Bertrand Y., Ché C., Nguyen T., Régina A., Gabathuler R., Castaigne JP., Béliveau R. (2008) Involvement of the low-density lipoprotein receptor-related protein in the transcytosis of the brain delivery vector angiopep-2. *J Neurochem*. 106, 1534-1544.

19 Maggie JM. (2011) Peptide for transmigration across blood brain barrier and delivery systems comprising the same. United States Patent Application Publication US 2011/0165079 A1

20 Richard,J.P. *et al.* (2003) Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. *J. Biol. Chem.* 278, 585-590

21 Madani F., Lindberg S., Langel U., Futaki S., Gräslund A. (2011) Mechanisms of cellular uptake of cell-penetrating peptides *J Biophys*. 414729.

22 Fuchs SM., Raines RT. (2004). Pathway for polyarginine entry into mammalian cells. *Biochemistry*. 43, 2438-2444

23 Vazquez,E. *et al.* (2010) Protein nanodisk assembling and intracellular trafficking powered by an arginine-rich (R9) peptide. *Nanomedicine. (Lond)* 5, 259-268

24 Hellinger,E. *et al.* (2012) Comparison of brain capillary endothelial cell-based and epithelial (MDCK-MDR1, Caco-2, and VB-Caco-2) cell-

based surrogate blood-brain barrier penetration models. *Eur. J. Pharm. Biopharm.* 82, 340-351

25 Unzueta,U. *et al.* (2012) Non amyloidogenic arginine-rich peptide tags for the tailored self-assembling of protein nanoparticles. *Biomaterials* 33: 8714-8722

26 Unzueta,U. *et al.* (2013) Improved performance of protein-based recombinant gene therapy vehicles by tuning downstream procedures. *Biotechnol. Prog.* 29, 1458-1463

27 Xu Z. et al. (2015) Targeting low-density lipoprotein receptors with protein-only nanoparticles. *Journal of Nanoparticle Research* 17, 150

28 Unzueta,U. *et al.* (2015) Towards protein-based viral mimetics for cancer therapies. *Trends Biotechnol.* 33, 253-258

29 Jiang, W., Kim, B., Rutka, J.T. and Chan, W.C.W. (2008) Nanoparticle-mediated cellular response is size-dependent. *Nature Nanotechnol* 3, 145 - 150

30 Shang, L., Nienhaus, K. and Nienhaus, G.U. (2014) Engineered nanoparticles interacting with cells: size matters *J Nanobiotechnol* 12, 5

31 Rejman, J., Oberle, V., Zuhorn, I.S. and Hoekstra, D. (2004) Size-dependent internalization of particles via the pathways of clathrin- and caveolae-mediated endocytosis. *Biochem J.* 377(Pt 1), 159-169

32 Shilo, M. *et al*., (2015) The effect of nanoparticle size on the probability to cross the blood-brain barrier: an in-vitro endothelial cell model. *J Nanobiotechnol*. 13, 19

33 Spencer,B.J. and Verma,I.M. (2007) Targeted delivery of proteins across the blood-brain barrier. *Proc. Natl. Acad. Sci. U. S. A* 104, 7594-7599

34 Su,T. and Stanley,K.K. (1998) Opposite sorting and transcytosis of the polymeric immunoglobulin receptor in transfected endothelial and epithelial cells. *J. Cell Sci.* 111 (Pt 9), 1197-1206.

35 Bhardwaj V and Ravi-Kumar MNV. Drug delivery systems to fight cancer. In: Fundamentals and Applications of Controlled Release Drug Delivery, Editors: Juergen Siepmann, Ronald A. Siegel, Michael J. Rathbone. Advances in Delivery Science and Technology, 2012, Chapter 15, p. 493-516, ISBN: 978-1-4614-0880-2)

36 Salvati et al. (2013) Transferrin-functionalized nanoparticles lose their targeting capabilities when a biomolecule corona adsorbs on the surface. *Nature Nanotechnol* 8, 137–143

37 Moestrup,S.K. and Nielsen,L.B. (2005) The role of the kidney in lipid metabolism. *Curr. Opin. Lipidol.* 16, 301-306

38 Christensen,E.I. *et al.* (2009) Receptor-mediated endocytosis in renal proximal tubule. *Pflugers Arch.* 458, 1039-1048

39 Bertrand,N. and Leroux,J.C. (2012) The journey of a drug-carrier in the body: an anatomo-physiological perspective. *J. Control Release* 161, 152-163

40 de Marco A. (2009) Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli. *Microb. Cell Fact.* 8, 26

41 Vazquez,E. and Villaverde,A. (2013) Microbial biofabrication for nanomedicine: biomaterials, nanoparticles and beyond. *Nanomedicine (Lond)* 8, 1895-1898

42 Singh,A. *et al.* (2015) Solubilization and refolding of inclusion body proteins. *Methods Mol. Biol.* 1258, 283-291

43 Mamat, U. *et al.* (2015) Detoxifying *Escherichia coli* for endotoxin-free production of recombinant proteins. *Microb. Cell Fact* 14, 57

44 Guillard,S. *et al.* (2015) Engineering therapeutic proteins for cell entry: the natural approach. *Trends Biotechnol.* 33, 163-171