Controlling self-assembling and tumor cell-targeting of protein-only nanoparticles through modular protein engineering

Eric Voltà-Durán [1,2,3], Olivia Cano-Garrido [1,4], Naroa Serna [1,2,3], Hèctor López-Laguna [1,2,3], Laura Sánchez-García [1,2,3], Mireia Pesarrodona [1,2,3 ¥], Alejandro Sánchez-Chardi [6], Ramón Mangues [3,5], Antonio Villaverde [1,2,3], Esther Vázquez [1,2,3]*, Ugutz Unzueta [3,5]*

[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), C/ Monforte de Lemos 3-5, 28029 Madrid, Spain

[4] Nanoligent SL, Edifici EUREKA, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

[5] Institut d'Investigacions Biomèdiques Sant Pau and Josep Carreras Research Institute, Hospital de la Santa Creu i Sant Pau, 08041 Barcelona, Spain

[6] Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Spain

Nanoparticles; protein materials; recombinant proteins; drug delivery; self-assembling; cancer cell targeting

ABSTRACT

Modular protein engineering is suited to recruit complex and multiple functionalities in single-chain polypeptides. Although still unexplored in a systematic way, it is anticipated that the positioning of functional domains would impact and refine these activities, including the ability to organize as supramolecular entities and to generate multifunctional protein materials. To explore this concept, we have repositioned functional segments in the modular protein T22-GFP-H6 and characterized the resulting alternative fusions. In T22-GFP-H6, the combination of T22 and H6 promotes self-assembling as regular nanoparticles and selective binding and internalization of this material in CXCR4-overexpressing tumor cells, making them appealing as vehicles for selective drug delivery. These pleiotropic activities result dramatically affected in module-swapped constructs, proving the need of a carboxy terminal positioning of H6 for protein self-assembling, and the accommodation of T22 at the amino terminus as a requisite for $CXCR4^+$ cell binding and internalization. Furthermore, failure of self-assembling as regular oligomers reduces cellular penetrability of the fusions while keeping the specificity of the T22-CXCR4 interaction. Altogether, these data instruct how multifunctional nanoscale protein carriers can be designed for smart, protein-driven drug delivery, not only for the treatment of $CXCR4^+$ human neoplasias but also for the development of anti-HIV drugs and other pathologies in which CXCR4 is a relevant homing marker.

INTRODUCTION

Protein-based drug delivery offers promises in the design of innovative drug delivery systems, especially when intended for precision medicines that require selective penetration of the therapeutic agent into target cells [1]. Proteins and peptides, over other materials commonly explored as drug delivery systems, offer high structural and functional versatility easily regulatable by genetic engineering. This flexibility allows generating protein materials non-existing in nature, with novel functions or combinations of them, for appealing applications [2-8]. Among different strategies of protein engineering , modular genetic fusion permits the recruitment, in single polypeptide chains, of multiple biological activities [9], and their simple biological production in recombinant forms [10]. Such modular concept has been largely used in cancer therapies [11-13], where the combination of targeting and cytotoxic domains is a common issue [10, 14-18]. Usually, the precise order of modular placement in multifunctional proteins is based on trial-and-error assisted by the background about protein functionalities when amino or carboxy termini are blocked, as clinically relevant peptides might require free ends for biological functions. Examples of full structure-based rational design, although existing, are not abundant [9, 19-22]. Recently, we have adapted the modular protein T22-GFP-H6 as a self-assembling protein-based tumor-targeted vehicle for pro-apoptotic factors [18] and for the antitumoral drug floxuridine [23]. The use of this vehicle for precision cancer medicine has resulted in reduced tumor volume but specially, in less metastatic foci in mouse models of colorectal cancer [23]. T22-GFP-H6 is produced in recombinant bacteria and oligomerizes into toroid nanoparticles of 12 nm (see figures 1 and 2 in [24]), representing a paradigmatic example of universal transversal self-assembling protein

platform based on a combination of cationic peptides and polyhistidines [25]. Self-assembling of T22-GFP-H6 requires the cationic peptide T22, a polyphemusin II-derivative [26] that acts as a potent ligand [27, 28] of the tumoral marker CXCR4 [29]. Self-organization of the protein also requires a histidine rich-region (H6) responsible for divalent cation-mediated protein-protein interactions [30]. The optimal position of these functional modules has not been previously assessed, as it derives from early data describing the architectonic potential of cationic peptides in GFP fusions [31]. Since CXCR4 is overexpressed in more than 20 human neoplasias [32-37], developing optimal protein vehicles as CXCR4-selective drug carriers would be highly desirable. On the other hand, the use of histidine-rich regions as protein connectors is of obvious interest for the construction of tailored protein materials. In this context, we have explored the potential impact of modular swapping in the model T22-GFP-H6, on those properties critical for its role as drug vehicle, including assembling capacity, selective CXCR4-targeting and cell internalization. Being the properties of T22-GFP-H6 extensible to the protein self-assembling platform irrespective of the involved functional domains, this information is of pivotal relevance for the further design of built-in, protein-only antitumoral drugs within the nanoscale for active tumor targeting [13, 14, 38].

METHODS

Genetic design, protein production and purification

The nomenclature of the proteins has been established according to their modular organization. An *Escherichia coli* codon-optimized gene encoding H6-GFP-T22 was designed in house and provided by Geneart (ThermoFisher) inserted in pET22b. Details of the other gene fusions can be found elsewhere (see figure 1 in [30]) and the properties of the encoded proteins are summarized in Figure 1A. The plasmids were transformed in *E. coli* Origami B (BL21, $OmpT^-$, Lon^-, TrxB, Gor^-; Novagen). All proteins were produced in bacterial cells cultured in Lysogenic Broth (LB) medium (supplemented with 100 µg/mL ampicillin, 12.5 µg/mL tetracycline and 15 µg/mL kanamycin) at 20 ºC overnight, upon induction with 0.1 mM of Isopropyl-β-D-thiogalactopyronaside (IPTG). H6-GFP-T22 gene expression was triggered with 0.1 mM IPTG during 3 h at 37 ºC, for an optimal yield. Bacteria cells were then harvested by centrifugation (15 min, 5,000 g) and resuspended in wash buffer (20 mM Tris-HCl, 500 mM NaCl, 10 mM Imidazole, pH = 8) in presence of protease inhibitors (Complete EDTA-free, Roche Diagnostics), before being disrupted in a French Press (Thermo FA-078A) by two rounds at 1,200 psi. Soluble fractions were collected as the supernatant after centrifugation for 45 min at 15,000g).

Proteins were finally purified by Immobilized Metal Affinity Chromatography (IMAC) using HisTrap Chelating HP 1mL columns (GE Healthcare) in an ÄKTA Pure system (GE Healtchare). Elution was achieved by a liner gradient of elution buffer (20 mM Tris-HCl, 500 mM NaCl, 500 mM Imidazole, pH = 8) and the recovered proteins were finally dialyzed against sodium carbonate with salt (166 mM $NaCO_3H$, 333 mM NaCl, pH = 8) buffer and centrifuged (15 min, 15,000 g) to remove insoluble aggregates. In the case of

GFP-H6, plain sodium carbonate without additional salt (166 mM $NaCO_3H$, pH = 8) buffer was used.

Protein purity, integrity and concentration

Protein purity was determined by Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) followed by Western Blot (WB) immunodetection, using anti-His monoclonal antibody (Santa Cruz Biotechnology). On the other hand, protein integrity was evaluated by Matrix-Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) mass spectrometry. The final protein concentration prior experimental was determined by Bradford's assay.

Electron microscopy

Ultrastructural characterization of T22-GFP-H6, T22-GFP-H6 (LOOP), H6-GFP-T22, and GFP-H6 nanomorphometry (size and shape) was determined at nearly native state with field emission scanning electron microscopy (FESEM). Drops of 3 µl of samples diluted at 0.3 mg/mL were directly deposited on silicon wafers (Ted Pella Inc.) for 30 s, excess of liquid blotted, air dried, and immediately observed without coating with a FESEM Zeiss Merlin (Zeiss) operating at 1 kV and equipped with a high resolution *in-lens* secondary electron detector. Representative images of a general fields and nanostructure details were captured at two high magnifications (100,000x and 300,000x).

***In silico* calculations**

UCSF Chimera software [39] was used for measuring the highest theoretical distances among amino acids in GFP monomers and dimers. With this purpose, 1GFL PDB

structure was used, consisting in a dimeric structure solved by X-ray diffraction with a resolution of 1.9 Å [40]. The measurements serve as an idea of the nature of GFP-proteins building blocks based on their hydrodynamic size.

Volume size distribution and fluorescence emission

Volume size distribution of unassembled and assembled proteins was determined by Dynamic Light Scattering (DLS), at 633 nm and 25 ºC in a Zetasizer Nano ZS (Malvern Instruments Limited), in a ZEN2112 3 mm quartz cuvette. Measurements were done in triplicate for error estimation and peak values referred to the average mode of the populations. In order to prove their oligomeric nature, T22-GFP-H6 hydrodynamic size was measured after exposing the protein either to 3 % Sodium Dodecyl Sulfate (SDS) or Ethylenediaminetetraacetic acid (EDTA) at 1:1 molar ratio. Green fluorescence emission of proteins was determined at 510 nm using a Varian Cary Eclipse Fluorescence Spectrophtometer (Agilent Technologies) upon excitation at 488 nm. For that, all proteins were used at 1 mg/mL in a final volume of 100 µL.

Cell culture and protein internalization

HeLa $CXCR4^+$ cells were used to study the performance of the recombinant proteins *in vitro.* HeLa cells were maintained in Eagle Minimum Essential Medium (MEM Alpha 1x GlutaMAX™, Gibco®) supplemented with 10 % fetal bovine serum (Gibco®) at 37 ºC in a 5 % CO_2 humidified atmosphere. A preliminary MTT-luci assay was done to assess lack of protein toxicity. Essentially, a CellTiter-Glo® Luminiscent Cell Viability Assay (Promega Corporation) was performed after protein exposure at 1 µM and during 48 h. as previously described [17]. To monitor protein internalization, HeLa cells were cultured on 24-well

plates at $3 \cdot 10^4$ cells/well for 24 h until reaching 70 % confluence. Proteins were incubated for 2 h and 24 h at different concentrations (25, 100 and 1000 nM) in presence of OptiPRO™ Serum Free Medium (Gibco) supplemented with L-glutamine. Additionally, specific internalization through CXCR4 receptor was assessed adding a specific antagonist, AMD3100 [41, 42], which inhibits the specific interaction between T22 and CXCR4. This chemical compound was added at a final concentration of 10 μM (10 times the highest protein concentration) for 1 h prior to protein incubation [43]. After protein exposure, cells were detached using 1 mg/mL Trypsin-EDTA (Gibco®) for 15 min at 37 ºC, a harsh protocol designed to remove externally attached protein [44]. Intracellular protein fluorescence was determined by flow cytometry using a Fluorescence Assisted Cell Sorting (FACS)-Canto system (Becton Dickinson) using a 15 mW air-cooled argon ion laser exciting at 488 nm excitation. Intracellular protein fluorescence was measured in duplicate for each condition.

Confocal laser scanning microscopy

For confocal microscopy, HeLa cells were grown on MatTek plates (MatTek Corporation) for 48 h (50.000 cells/well). Upon exposure to 1000 nM of protein for 24 h, cell nuclei were labelled with 5 μg/mL Hoechst 33342 (ThermoFisher), rendering a blue signal, and plasma membrane with 2.5 μg/mL CellMask™ Deep Red (ThermoFisher), rendering a red signal, for 10 min at room temperature. Cells were then washed in PBS buffer (Sigma-Aldrich). Images were collected on an inverted TCS SP5 Spectral confocal microscope (Leica Microsystems) using 63 x (1.4 NA) oil immersion objective lenses. Excitation was reached via a 405 nm blue diode laser for Hoechst 33342, 488 nm line of an argon ion laser for GFP and 633 nm line of a HeNe laser for CellMask™ Deep Red. Optimized

emission detection bandwidths were configured to avoid inter-channel crosstalk and multitrack sequential acquisition settings were used. The confocal pinhole was set to 1 Airy unit and z-stacks acquisition intervals were chosen to satisfy Nyquist sampling criteria. Images were processed using Imaris v 7.2.1 software (Bitplane). Three-dimensional images were processed using the Surpass Module.

Data analysis

Quantitative data are expressed as mean ± standard error. Pairwise comparisons of conditions for each protein were made with Tukey's HSD Q-test. Significance was accepted at $p < 0.05$. All statistical analyses were performed with PAST3.

RESULTS

To analyze the potential impact of alternative modular arrangements and the need of specific end-terminal placement of specific functional domains, we have designed, produced and purified the fusion protein H6-GFP-T22, for comparative functional and structural examination with T22-GFP-H6. For a deep analysis of the impact of modular distribution, we included the T22-lacking version GFP-H6 [45] (Figure 1A), and T22-GFP-H6(LOOP), in which the hexahistidine tail is inserted into a solvent-exposed permissive loop of GFP [30] (Figure 1A). Like the rest of proteins, H6-GFP-T22 was efficiently produced in *E. coli* as a proteolytically stable, single protein species of expected molecular mass (Figure 1B, C). Importantly, being highly fluorescent (Figure 1A), this protein was suitable for tracking in further comparative experiments to study protein-cell interactions and internalization by flow cytometry and confocal microscopy. When determining the self-assembling properties of this new modular distribution we noted that, in contrast with the original T22-GFP-H6 (assembling as 12 nm--oligomers), H6-GFP-T22 was unable to self-assemble as regular nanoparticles under the conditions that trigger the nanoscale organization of T22-GFP-H6 (Figure 2A, B, C). H6-GFP-T22 showed a hydrodynamic size (5.6 nm) compatible with the monomeric form of GFP (Figure 2D). GFP-H6 and T22-GFP-H6(LOOP) showed a slightly larger average particle size (6.5 nm), compatible with an unbalanced mixture between monomeric (5.4 nm) and dimeric forms (around 7.7 nm) of GFP (Figure 2D). Both the presumed dimeric forms and the multimeric (decameric [24]) T22-GFP-H6, were disassembled to the smaller 5.6 nm building blocks (Figure 2C) by either the addition of EDTA, that removes the cross-molecular interacting

divalent cations supportive of nanoparticle formation [30], or by the addition of the denaturing agent SDS. None of these proteins showed toxicity when exposed to cultured $CXCR4^+$ HeLa cells (Figure 2 E).

Irrespective of its oligomerization status, we were interested in knowing whether T22, in a C-terminal accommodation, would be able to recognize its CXCR4 receptor and promote internalization into $CXCR4^+$ cells. The intrinsic fluorescence of H6-GFP-T22 allowed an easy tracking of the protein in protein-exposed cells. As observed (Figure 3A), H6-GFP-T22 was only moderately able to penetrate $CXCR4^+$ HeLa cells in culture in a time-dependent fashion, with respect to the control T22-GFP-H6. Then, it was observed about 5-fold less intracellular fluorescence (corrected by the specific fluorescence, and therefore, representative of protein amounts) than T22-GFP-H6, and 4-fold less when compared to T22-GFP-H6(LOOP). This profile, similar to data recorded at both 2 and 24 h (Figure 3A), was indicative of a less efficient cell uptake of the protein mediated by the T22 peptide when positioned at the C terminus of the protein. This fact could be not straightforward accounted by the lack of assembly into nanoparticles, as T22-GFP-H6(LOOP) was also unable to form regular oligomers (Figure 2) and it showed a cell penetrability potential only slightly lower than that of the parental protein (Figure 3A). The similar uptake of T22-GFP-H6 and T22-GFP-H6(LOOP) indicates that the position of H6, the protein segment responsible for oligomerization, has only a minor effect on the cell penetrability mediated by the N-terminal T22. This fact can be explained as an indirect effect because the regular multivalent presentation of the CXCR4 ligand (promoted by functional H6) favors cooperative cell binding and endosome formation in a virus-like fashion [46]. When checking whether the residual binding of H6-GFP-T22 to HeLa cells

was CXCR4-dependent, we noted that the specificity of T22-mediated interaction was also lost. At 1 µM protein concentration, the CXCR4 antagonist AMD3100 [41, 42] blocked the penetration of T22-GFP-H6 by about 75 %, while the uptake of H6-GFP-T22 was only reduced by 15 % (Figure 3B). At low protein doses, the specificity of the internalization process generically increased but it was still non-significant and lower in the case of H6-GFP-T22 than when testing the parental construct (79 % versus 98 %). This indicated that the residual internalization of H6-GFP-T22 was mostly CXCR4-independent.

Envisaging that this fact might also affect the intracellular localization of the protein, we analyzed protein-exposed HeLa cells by confocal microscopy. Indeed, while T22-GFP-H6 and T22-GFP-H6(LOOP) distributed as a perinuclear punctuated pattern, indicative of an endosomal penetration route, H6-GFP-T22 fluorescence mostly accumulated at the cell membrane level, with a very modest true internalization (Figure 4, top). Control GFP-H6 was not detected in cells. These patterns were fully confirmed by 3D cell imaging, that again supported the strong membrane association of H6-GFP-T22 (Figure 4 bottom Figure 5) with only limited internalization, in contrast to what happens to the proteins with the amino terminal T22. Since previous to flow cytometric analysis, protein-exposed cells were submitted to a harsh trypsin treatment to remove the externally attached protein [44], the detected green fluorescence (in particular the yellow merging) revealed a very tight interaction between the protein and cell membranes. Such interaction appeared as different from that promoting endosomal internalization (assumed to be CXCR4-dependent and responsible for the recorded mild specificity), and an important fraction of H6-GFP-T22 was neither bound nor internalized.

DISCUSSION

Altogether, the set of data presented here firmly demonstrate the relevance of domain accommodation in modular proteins, regarding both the property to form supramolecular protein materials through self-assembling but also the functionality of individual functional modules. The combination of a cationic domain plus a histidine-rich region, surrounding a core scaffold protein, was revealed as a transversal protein engineering platform to construct functional protein-only nanoparticles for clinical applications [25]. In this context, this architectonic principle at the nanoscale has been adapted to develop antimicrobial nanoparticles [47], blood-brain-barrier crossing agents [48], CD44-targeted nanoparticles [49], tumor imaging agents [24], vehicles for tumor-targeted chemical drug delivery [23], theragnostic agents [18] and protein-only, highly cytotoxic antitumoral drugs effective in colorectal cancer [50] and in acute myeloid leukemia [15]. The symmetry associated to the protein-assembling process, for which both a cationic peptide and a polyhistidine tail are required [25], is lost when the histidine-rich segment is placed in the alternative N-terminal position or in an intermediate localization, in an overhanging loop of GFP (Figure 2B). Also, irrespective of the formation of nanoparticles from protein building blocks, T22, strongly cationic, is fully functional at the N terminus but not in the C terminus (Figure 3), where this peptide is not able to specifically and effectively promote protein internalization through the cell surface receptor CXCR4. At high concentrations (Figure 3B), T22 seems to act as still inefficient cell-penetrating cationic peptide, with untargeted cell penetrability [51]. The higher penetrability of T22-GFP-H6 compared with that of T22-GFP-H6(LOOP) (Figures 3, 4 and 5) can be accounted by the multivalent attachment of the first construct,

that is expected to favor cooperativity in both binding and endosomal mediated internalization, compared with that of monomers and dimers formed by T22-GFP-H6(LOOP) [46]. Finally, a free amino terminus of T22 is the requirement for a proper association with CXCR4. This might be related to the fact that the N-terminal repeat Tyr-Arg-Lys in T22 is highly associated with the antiviral activity of the peptide, that is based on the blockage of HIV binding to CXCR4 [52]. A failure in T22-CXCR4 interaction by a wrong positioning of T22 in the fusion protein aborts the selectivity of the ligand by its cellular receptor and conduces to an unspecific and inefficient binding to cells, that precluded endosomal internalization (Figures 3, 4 and 5). As CXCR4 is not only a relevant and transversal tumoral marker but a key receptor in HIV infection, and T22 is a useful antagonist in the infection process [26-28], reaching a rationale for a proper T22 positioning in potential antiviral drugs is of wider interest in the development of antitumoral drugs and therapeutic strategies.

CONCLUSIONS

We have determined here that the peptide H6 is a potent architectonic tag that allows the controlled oligomerization of recombinant fusion proteins. Moreover, this segment is only active when displaying a free carboxy terminus, but not when displaying the amino terminus or in absence of free termini. Besides this, the potent ligand of CXCR4, the peptide T22, promotes efficient $CXCR4^+$ cell binding and internalization through its free amino terminal end. Although the nanoscale oligomerization supported by a properly accommodated H6 clearly favors CXCR4-dependent cell internalization, the requirement

of a free amino terminal end of T22 is irrespective of the protein oligomerization as regular nanoparticles. Sustained by the growing interest in protein-only smart materials with clinical applicability, these findings offer relevant clues for the development of transversal and simple platforms based on modular protein engineering and biological fabrication.

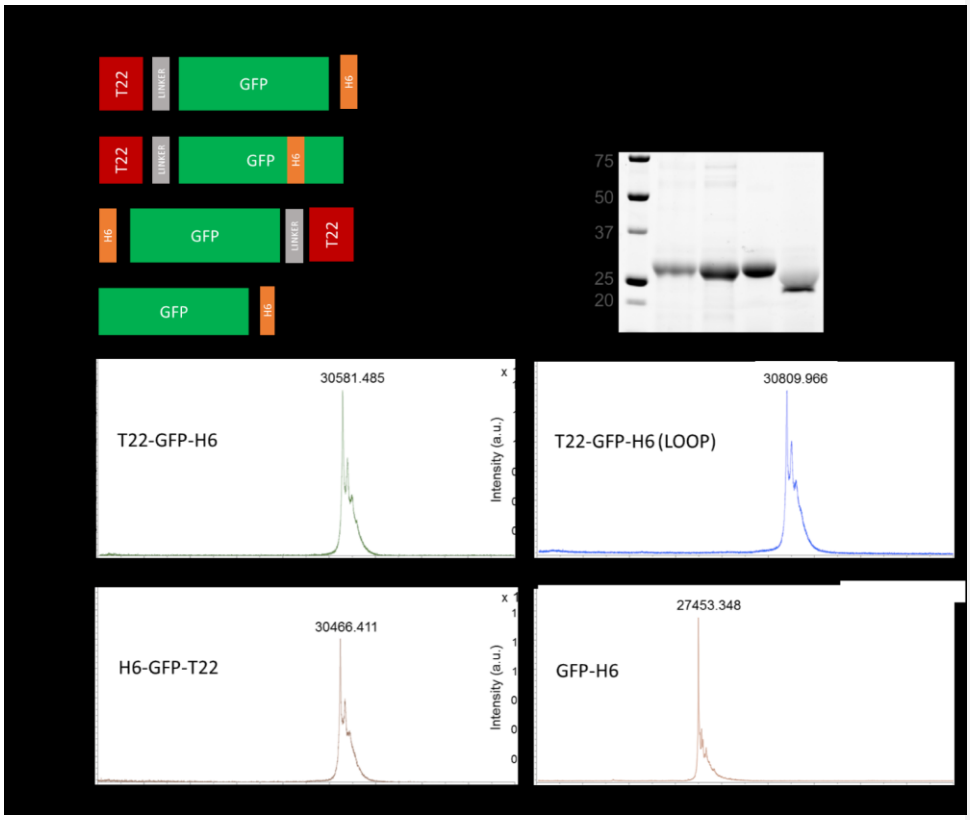

Figure 1. Properties of building blocks and protein nanoparticles. A. Modular organization of the multifunctional proteins used in the study. Segment sizes are indicative. Protein names are according the order of modules in the constructs. T22-GFP-H6(LOOP) refers to the insertion of H6 into a solvent-exposed permissive loop of GFP. Theoretical molecular masses are indicated in KDa, as well as the specific fluorescence (FU) relative to T22-GFP-H6. B. Protein purity upon purification, shown by SDS-PAGE,

demonstrating their proteolytic stability as single molecular mass species. M in the image indicates the molecular marker lane, showing the molecular masses in KDa. C. Mass spectrometry of the same protein samples.

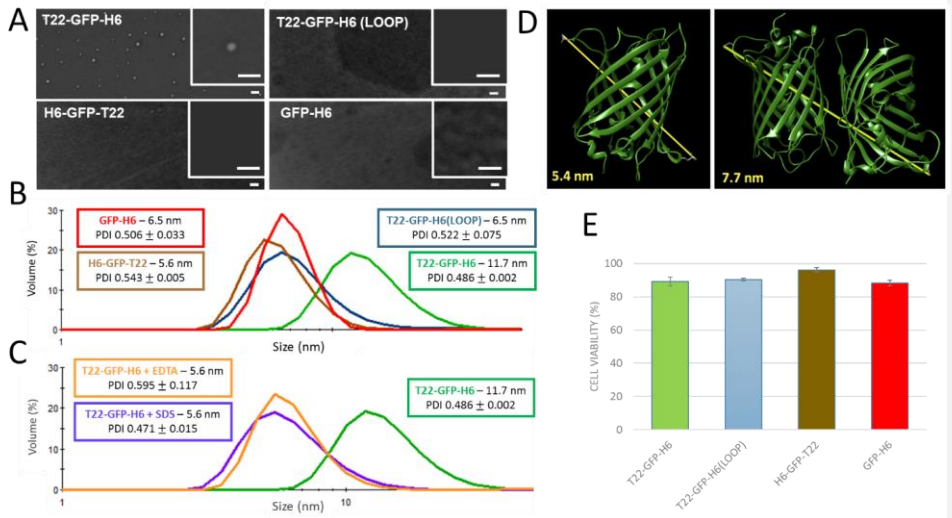

Figure 2. Protein assembly as oligomeric nanoparticles. A. Representative FESEM images of the GFP-based recombinant proteins upon purification. Bars indicate 30 nm. B. Hydrodynamic size of unassembled proteins and protein nanoparticles, indicating the peak size of DLS plots and the polydispersion index (PDI). C. Controlled disassembly of T22-GFP-H6 nanoparticles mediated by 3 % SDS or by EDTA at 1:1 molar ratio related to His residues. D. Molecular modelling of native GFP showing both the monomeric and dimeric forms and their larger atomic distances. E. MTT cell viability analysis of HeLa cells exposed to pure proteins (at 1000 nM) for 48 h. Data are referred to the viability of non-exposed cells cultured under the same conditions.

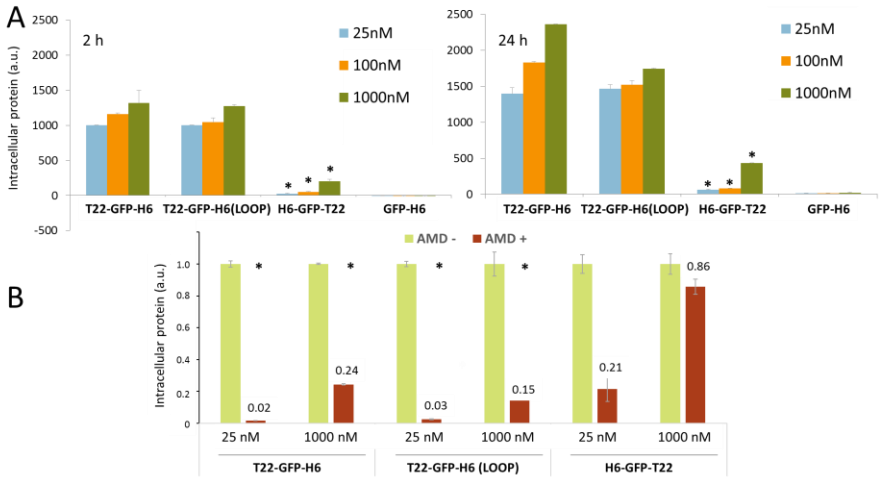

Figure 3. Efficacy and specificity of cell penetrability of protein nanoparticles. A. Internalization of T22-GFP-H6, T22-GFP-H6(LOOP), H6-GFP-T22 and GFP-H6 in CXCR4$^+$ HeLa cells, recorded after 2 h and 24 h of exposure, at three different concentrations (25 nM, 100 nM and 1000 nM). Crude fluorescence values were normalized by the specific fluorescence emission of each protein to allow molar-based comparison. Y-axis in arbitrary units of fluorescence (a.u.). Significant differences with T22-GFP-H6 internalization under the same conditions (included as a control) are represented by * ($p < 0.001$). B. Relative intracellular fluorescence values (a.u.) for each protein under prior treatment of CXCR4$^+$ HeLa cells with AMD3100, a specific inhibitor of T22-CXCR4 binding, compared to those obtained at the same conditions in absence of this compound. The analysis was done 2 h after exposure. Significant difference with protein internalization under the same conditions but without AMD3100 are represented by * ($p < 0.001$). The standard error is represented by grey lines at each sample.

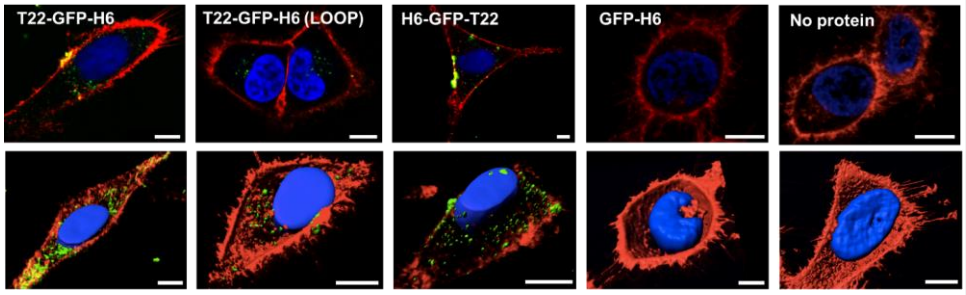

Figure 4. Intracellular localization of proteins in protein-exposed cells. 2D confocal imaging of protein-exposed HeLa cells. Red signal indicates labelled membranes, blue signal labelled nuclear nucleic acids and the green signal is the natural green fluorescence of proteins. Cells cultured in absence of proteins are also shown for a clear visualization of blue and red signals in absence of protein. At the bottom, 3D Imaris reconstruction of confocal stacks along the z axis of the same experiments. Analysis was performed after 24 h of protein exposure at 1 μM. Bars indicate 10 μm.

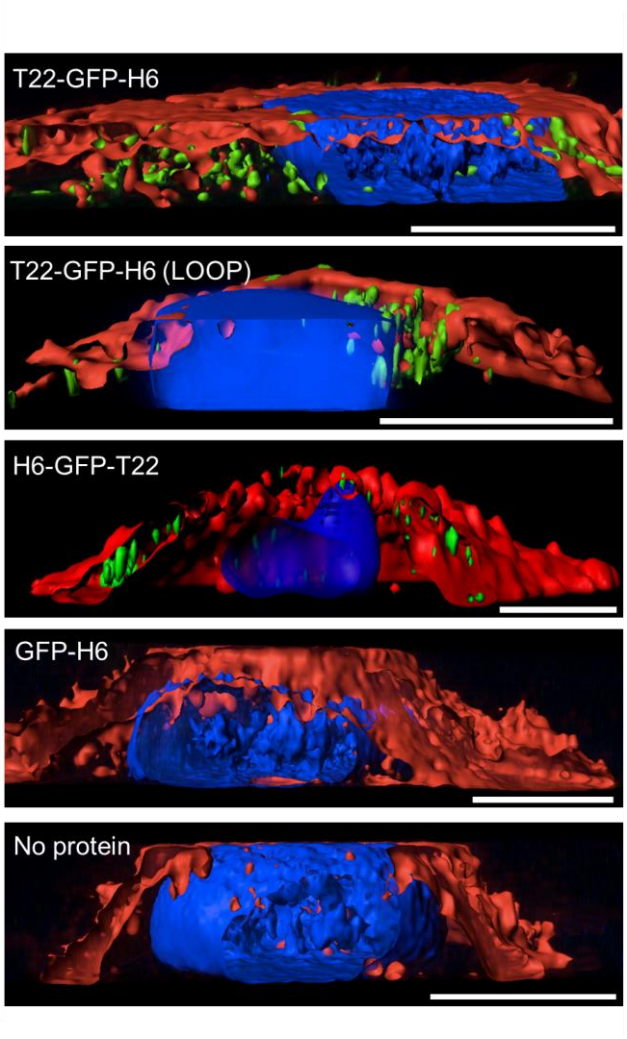

Figure 5. Intracellular and membrane localization of proteins in protein-exposed cells. Fine details of 3D Imaris reconstructions of confocal stacks along the z axis, showing the green fluorescent material. Cells cultured in absence of proteins are also shown. Red signal indicates labelled membranes and blue signal labelled nuclear nucleic

acids. The analysis was performed after 24 h of protein exposure at 1 μM. Bars indicate 10 μm.

AUTHOR INFORMATION

Corresponding Author

* Corresponding authors: Esther Vazquez and Antonio Villaverde

Present Addresses

¥ Mireia Pesarrodona. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, 08028 Barcelona, Spain

Author Contributions

EVD has performed most of the experiments and figures, assisted by OCG, NS, HLL and LSG. MP designed and produced one of the proteins used here. ASC has performed electron microscopy and some statistics. UU has supervised the study. RM, AV and EV have conceived the whole study, that has been mostly written by AV. All authors have given approval to the final version of the manuscript. RM, EV and AV are co-founders of Nanoligent SL, a company devoted to the design of anticancer drugs based on T22.

ACKNOWLEDGMENT

We are indebted to Agencia Estatal de Investigación (AEI) and to Fondo Europeo de Desarrollo Regional (FEDER) (grant BIO2016-76063-R, AEI/FEDER, UE) to AV, AGAUR (2017SGR-229) to AV and 2017SGR-865 GRC to RM; CIBER-BBN (project

NANOPROTHER) granted to AV and CIBER-BBN project 4NanoMets to RM; ISCIII (PI15/00272 co-founding FEDER) to EV and ISCIII (Co-founding FEDER) PIE15//00028 and PI18/00650 to RM, and to EU COST Action CA 17140. We are also indebted to the Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN) that is an initiative funded by the VI National R&D&I Plan 2008–2011, Iniciativa Ingenio 2010, Consolider Program, CIBER Actions and financed by the Instituto de Salud Carlos III, with assistance from the European Regional Development Fund. 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 sePBioEs scientific-technical service (http://www.nanbiosis.es/portfolio/u1-protein-production-platform-ppp/) and the nanoparticle size analysis by the Biomaterial Processing and Nanostructuring Unit. Confocal and electron microscopy studies were performed at the Servei de Microscòpia and cell culture experiments at the SCAC, both located at the UAB. AV received an ICREA ACADEMIA award. LSG and HLL were supported by a predoctoral fellowship from AGAUR (2018FI_B2_00051) and (2019FI_B_00352) respectively and UU by PERIS program from the Health Department of la Generalitat de Catalunya.

.

References

[1] Casanova I, Unzueta U, Arroyo-Solera I, Cespedes MV, Villaverde A, Mangues R, et al. Protein-driven nanomedicines in oncotherapy. Curr. Op. Phar. 2019,47:1-7.
[2] Loo Y, Goktas M, Tekinay AB, Guler MO, Hauser CA, Mitraki A. Self-Assembled Proteins and Peptides as Scaffolds for Tissue Regeneration. Adv. Health. Mat. 2015,4:2557-86.
[3] Kumar VA, Wang BK, Kanahara SM. Rational design of fiber forming supramolecular structures. Exp Biol Med. 2016,241:899-908.
[4] Yeates TO, Liu Y, Laniado J. The design of symmetric protein nanomaterials comes of age in theory and practice. Curr. Op. Struct. Biol. 2016,39:134-43.
[5] Sutherland T.D. Rapson T.D. HMG, Church J.S. Recombinant Structural Proteins and Their Use in Future Materials. . In: Parry D. SJ, editor. Fibrous Proteins: Structures and Mechanisms Cham: Springer, 2017.
[6] Guttenplan APM, Young LJ, Matak-Vinkovic D, Kaminski CF, Knowles TPJ, Itzhaki LS. Nanoscale click-reactive scaffolds from peptide self-assembly. J Nanobiotechnology. 2017,15:70.
[7] Wei G, Su Z, Reynolds NP, Arosio P, Hamley IW, Gazit E, et al. Self-assembling peptide and protein amyloids: from structure to tailored function in nanotechnology. Chem. Soc. Rev. 2017,46:4661-708.
[8] Yeates TO. Geometric Principles for Designing Highly Symmetric Self-Assembling Protein Nanomaterials. Annual Rev. Biophys. 2017,46:23-42.
[9] DiMarco RL, Heilshorn SC. Multifunctional materials through modular protein engineering. Adv. Mat. 2012,24:3923-40.
[10] Sanchez-Garcia L, Martin L, Mangues R, Ferrer-Miralles N, Vazquez E, Villaverde A. Recombinant pharmaceuticals from microbial cells: a 2015 update. Microb. Cell Fact. 2016,15:33.
[11] Aris A, Villaverde A. Modular protein engineering for non-viral gene therapy. Trends Biotechnol. 2004,22:371-7.
[12] Vazquez E, Ferrer-Miralles N, Mangues R, Corchero JL, Schwartz S, Jr., Villaverde A. Modular protein engineering in emerging cancer therapies. Curr. Phar. Des. 2009,15:893-916.

[13] Serna N, Sanchez-Garcia L, Unzueta U, Diaz R, Vazquez E, Mangues R, et al. Protein-Based Therapeutic Killing for Cancer Therapies. Trends Biotechnol. 2018,36:318-35.
[14] Díaz R, Sánchez-García L, Serna N, Sánchez-Chardi A, Cano-Garrido O, Sánchez JM, et al. Engineering a recombinant chlorotoxin as cell-targeted cytotoxic nanoparticles. Sci. China Mat. 2019,62:892–8.
[15] Diaz R, Pallares V, Cano-Garrido O, Serna N, Sanchez-Garcia L, Falgas A, et al. Selective CXCR4(+) Cancer Cell Targeting and Potent Antineoplastic Effect by a Nanostructured Version of Recombinant Ricin. Small. 2018,14:e1800665.
[16] Sanchez-Garcia L, Serna N, Mattanovich M, Cazzanelli P, Sanchez-Chardi A, Conchillo-Sole O, et al. The fusogenic peptide HA2 impairs selectivity of CXCR4-targeted protein nanoparticles. Chem. Comm. 2017,53:4565-8.
[17] Serna N, Sanchez JM, Unzueta U, Sanchez-Garcia L, Sanchez-Chardi A, Mangues R, et al. Recruiting potent membrane penetrability in tumor cell-targeted protein-only nanoparticles. Nanotechnology. 2019,30:115101.
[18] Serna NC, M, Sánchez-García, L, Unzueta, U, Sala, R, Sánchez-Chardi, A, Cortés, F, Ferrer-Miralles, N, Mangues, R, Vázquez, E, Villaverde, A. Peptide-Based Nanostructured Materials with Intrinsic Proapoptotic Activities in CXCR4+ Solid Tumors. Adv. Funct. Mat. 2017,27:1700919.
[19] Grove TZ, Osuji CO, Forster JD, Dufresne ER, Regan L. Stimuli-responsive smart gels realized via modular protein design. J. Am. Chem. Soc. 2010,132:14024-6.
[20] Howorka S. Rationally engineering natural protein assemblies in nanobiotechnology. Curr. Op. Biotechnol. 2011,22:485-91.
[21] Cai L, Heilshorn SC. Designing ECM-mimetic materials using protein engineering. Act. Biomaterial. 2014,10:1751-60.
[22] Binz HK, Amstutz P, Pluckthun A. Engineering novel binding proteins from nonimmunoglobulin domains. Nature biotechnology. 2005,23:1257-68.
[23] Cespedes MV, Unzueta U, Avino A, Gallardo A, Alamo P, Sala R, et al. Selective depletion of metastatic stem cells as therapy for human colorectal cancer. EMBO Mol. Med. 2018. e8772.

Con formato: Inglés (Estados Unidos)

[24] Rueda F, Cespedes MV, Conchillo-Sole O, Sanchez-Chardi A, Seras-Franzoso J, Cubarsi R, et al. Bottom-Up Instructive Quality Control in the Biofabrication of Smart Protein Materials. Adv. Mat. 2015,27:7816-22.
[25] Unzueta U, Ferrer-Miralles N, Cedano J, Zikung X, Pesarrodona M, Saccardo P, et al. Non-amyloidogenic peptide tags for the regulatable self-assembling of protein-only nanoparticles. Biomaterials. 2012,33:8714-22.
[26] Nakashima H, Masuda M, Murakami T, Koyanagi Y, Matsumoto A, Fujii N, et al. Anti-human immunodeficiency virus activity of a novel synthetic peptide, T22 ([Tyr-5,12, Lys-7]polyphemusin II): a possible inhibitor of virus-cell fusion. Antimicrob Ag. Chemother. 1992,36:1249-55.
[27] Tamamura H, Kuroda M, Masuda M, Otaka A, Funakoshi S, Nakashima H, et al. A comparative study of the solution structures of tachyplesin I and a novel anti-HIV synthetic peptide, T22 ([Tyr5,12, Lys7]-polyphemusin II), determined by nuclear magnetic resonance. BBA. 1993,1163:209-16.
[28] Tamamura H, Arakaki R, Funakoshi H, Imai M, Otaka A, Ibuka T, et al. Effective lowly cytotoxic analogs of an HIV-cell fusion inhibitor, T22 ([Tyr5,12, Lys7]-polyphemusin II). Bioor. Med. Chem. 1998,6:231-8.
[29] Unzueta U, Cespedes MV, Ferrer-Miralles N, Casanova I, Cedano J, Corchero JL, et al. Intracellular CXCR4(+) cell targeting with T22-empowered protein-only nanoparticles. Int. J. Nanomed. 2012,7:4533-44.
[30] Lopez-Laguna H, Unzueta U, Conchillo-Sole O, Sanchez-Chardi A, Pesarrodona M, Cano-Garrido O, et al. Assembly of histidine-rich protein materials controlled through divalent cations. Act. Biomaterial. 2019,83:257-64.
[31] Vazquez E, Roldan M, Diez-Gil C, Unzueta U, Domingo-Espin J, Cedano J, et al. Protein nanodisk assembling and intracellular trafficking powered by an arginine-rich (R9) peptide. Nanomedicine. 2010,5:259-68.
[32] Balkwill F. The significance of cancer cell expression of the chemokine receptor CXCR4. Sem. Cancer. Biol. 2004,14:171-9.
[33] Kim J, Takeuchi H, Lam ST, Turner RR, Wang HJ, Kuo C, et al. Chemokine receptor CXCR4 expression in colorectal cancer patients increases the risk for recurrence and for poor survival. J. Clin. Oncol. 2005,23:2744-53.

[34] Burger JA, Kipps TJ. CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment. Blood. 2006,107:1761-7.
[35] Kim J, Mori T, Chen SL, Amersi FF, Martinez SR, Kuo C, et al. Chemokine receptor CXCR4 expression in patients with melanoma and colorectal cancer liver metastases and the association with disease outcome. Ann. Surg. 2006,244:113-20.
[36] Hermann PC, Huber SL, Herrler T, Aicher A, Ellwart JW, Guba M, et al. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. Cell Stem Cell. 2007,1:313-23.
[37] Croker AK, Allan AL. Cancer stem cells: implications for the progression and treatment of metastatic disease. J. Cell. Mol. Med. 2008,12:374-90.
[38] Sanchez-Garcia L, Serna N, Alamo P, Sala R, Cespedes MV, Roldan M, et al. Self-assembling toxin-based nanoparticles as self-delivered antitumoral drugs. J Control Rel.. 2018,274:81-92.
[39] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J. Compt. Chem. 2004,25:1605-12.
[40] Yang F, Moss LG, Phillips GN, Jr. The molecular structure of green fluorescent protein. Nat. Biotech. 1996,14:1246-51.
[41] Kim J, Connelly KL, Unterwald EM, Rawls SM. Chemokines and cocaine: CXCR4 receptor antagonist AMD3100 attenuates cocaine place preference and locomotor stimulation in rats. Brain, behavior, and immunity. 2016.
[42] Jung YH, Lee DY, Cha W, Kim BH, Sung MW, Kim KH, et al. Antitumor effect of CXCR4 antagonist AMD3100 on the tumorigenic cell line of BHP10-3 papillary thyroid cancer cells. Head & neck. 2016,38:1479-86.
[43] Kim HY, Hwang JY, Kim SW, Lee HJ, Yun HJ, Kim S, et al. The CXCR4 Antagonist AMD3100 Has Dual Effects on Survival and Proliferation of Myeloma Cells In Vitro. Cancer Res. Treat. 2010,42:225-34.
[44] Richard JP, Melikov K, Vives E, Ramos C, Verbeure B, Gait MJ, et al. Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. J. Biol. Chem. 2003,278:585-90.

[45] Cespedes MV, Unzueta U, Tatkiewicz W, Sanchez-Chardi A, Conchillo-Sole O, Alamo P, et al. In vivo architectonic stability of fully de novo designed protein-only nanoparticles. ACS nano. 2014,8:4166-76.
[46] Unzueta U, Cespedes MV, Vazquez E, Ferrer-Miralles N, Mangues R, Villaverde A. Towards protein-based viral mimetics for cancer therapies. Trends Biotech. 2015,33:253-8.
[47] Serna N, Sanchez-Garcia L, Sanchez-Chardi A, Unzueta U, Roldan M, Mangues R, et al. Protein-only, antimicrobial peptide-containing recombinant nanoparticles with inherent built-in antibacterial activity. Act. Biomaterial. 2017. 60, 256–263.
[48] Serna N, Cespedes MV, Saccardo P, Xu Z, Unzueta U, Alamo P, et al. Rational engineering of single-chain polypeptides into protein-only, BBB-targeted nanoparticles. Nanomedicine. 2016,12:1241-51.
[49] Pesarrodona M, Ferrer-Miralles N, Unzueta U, Gener P, Tatkiewicz W, Abasolo I, et al. Intracellular targeting of CD44+ cells with self-assembling, protein only nanoparticles. Int J. Pharma. 2014,473:286-95.
[50] Sanchez-Garcia L, Serna N, Alamo P, Sala R, Cespedes MV, Roldan M, et al. Self-assembling toxin-based nanoparticles as self-delivered antitumoral drugs. J. Control Rel. 2018. 274, 81-92.
[51] Guo Z, Peng H, Kang J, Sun D. Cell-penetrating peptides: Possible transduction mechanisms and therapeutic applications. Biomed. Rep. 2016,4:528-34.
[52] Tamamura H, Imai M, Ishihara T, Masuda M, Funakoshi H, Oyake H, et al. Pharmacophore identification of a chemokine receptor (CXCR4) antagonist, T22 ([Tyr(5,12),Lys7]-polyphemusin II), which specifically blocks T cell-line-tropic HIV-1 infection. Bioorganic Med. Chem. 1998,6:1033-41.