# Endosomal escape of protein nanoparticles engineered through humanized histidine-rich peptides

Hèctor López-Laguna[a,b,c], Rafael Cubarsi [d,c], Ugutz Unzueta[b,c,e*], Ramón Mangues[c,d], Esther Vázquez[a,b,c,*] and Antonio Villaverde[a,b,c,*]

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

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

[c]CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), C/ Monforte de Lemos 3-5, 28029 Madrid, Spain

[d]Departament de Matemàtiques, Universitat Politècnica de Catalunya, 08034, Barcelona, Spain

[e]Institut d'Investigacions Biomèdiques Sant Pau, Hospital de la Santa Creu i Sant Pau, 08025 Barcelona, Spain

[*]Corresponding authors. UU: uunzueta@santpau.cat; AV: antoni.villaverde@uab.es EV: esther.vazquez@uab.es

**Abstract**

Poly-histidine peptides such H6 (HHHHHH) are used in protein biotechnologies as purification tags, protein-assembling agents at the nanoscale and endosomal-escape entities. The pleiotropic properties of such peptides make them appealing to design protein-based smart materials or nanoparticles for imaging or drug delivery to be produced in form of recombinant proteins. However, the clinical applicability of H6-tagged proteins is restricted by the potential immunogenicity of these segments. In this study we have explored several humanized histidine-rich peptides in tumor-targeted modular proteins than bind and internalize cells through the tumoral marker CXCR4. We were particularly interested in exploring how protein purification, self-assembling and endosomal escape perform in proteins containing the variant histidine-rich tags. Among the tested candidates, the peptide H5E (HEHEHEHEH) has resulted promising as a good promoter of endosomal escape of the associated full-length protein upon endosomal internalization. The numerical modelling of cell penetration and endosomal escape of the tested proteins has revealed a negative relationship between the amount of protein internalized into target cells and the efficiency of cytoplasmic release. This fact demonstrates that the His-mediated, proton sponge-based endosomal escape saturates at moderate amounts of internalized protein, a fact that might be critical for the design of protein materials for cytosolic molecular delivery.

Keywords: Protein materials; nanoparticles; genetic design; membrane activity; endosomal escape; poly-histidines

## 1. Introduction

Proteins are versatile macromolecules whose structure and functional capabilities can be tailored by conventional genetic engineering. Among other regulatable properties, self-assembling can be achieved and controlled by protein engineering [1-12]. By either peptide synthesis or by recombinant DNA technologies, diverse categories of soft structures, including hydrogels, fibers, layers and nanoparticles can be fabricated [13, 14]. Being functional, these materials are assuming roles in biomedicine and nanomedicine, for regenerative medicine, diagnosis and drug delivery purposes. Protein nanoparticles are of particular interest as vehicles for drugs and imaging agents [15]. In this context, gene fusion technologies allow recruiting, in a single chain recombinant polypeptide, different functional peptides including self-assembling tags, endosomal escape regions and ligands of cell surface receptors [16], necessary for nanoparticle formation and for cell-targeted delivery of any associated payload drug. The catalogues of receptors and their ligands are rapidly expanding, especially in oncology [17], what allows empowering nanoparticles with appropriate cell ligands for attachment and cell internalization. The multivalent surface-display of ligands in multimeric nanoparticles, formed by multiple building blocks, promotes efficient endosomal internalization of the material in a virus-like fashion [18]. A particularly simple engineering procedure to generate protein nanoparticles is based on modular building blocks tagged with a hexahistidine tag (H6) at the carboxy termini [19, 20], that coordinates divalent cations as a molecular cross-linker [21, 22]. The resulting family of oligomers, sizing around 10-30 nm, are highly stable *in vivo* and show an excellent tumor targeting mediated by alternative tumoral markers [23, 24] and ligands [25-27]. Because of the excellent biodistribution in mice models of human cancers, this category of His-based materials has been used in nanoconjugates or protein-only nanoparticles to deliver small molecular weight drugs [28], pro-apoptotic factors [29] and toxins [26, 30] in metastatic colorectal cancer, and potent plant cytotoxins in metastatic acute myeloid leukaemia [31].

The peptide H6 is a critical component of such self-assembling platform, not only because of its architectonic properties but also as an endosomolytic agent [32]. The combination of both functions is very attractive when designing recombinant protein materials intended for endosomal delivery. Importantly, the application of toxin-based nanoparticles, immunotoxins or generically, protein-based drug vehicles [33, 34], would largely benefit from potent

endosomolytic agents. This is because lysosomal degradation is a critical bottleneck in the delivery route [35], and in this context, many strategies and active materials are currently developed to enhance the necessary cytosolic delivery [36-40]. Onn the other side, proteins are becoming increasingly valuable biopharmaceuticals [34, 41-43], and H6-based functionalization by gene fusion is a common procedure [44]. However, the use of H6 in clinics might be a source of regulatory concerns, since this peptide might be potentially immunogenic in humans [45-47], what pushes to identify alternatives regarding both the nanoarchitectonic and endosomolytic properties of H6.

## 2. Materials and methods

### *2.1 Protein production and purification*

Histidine-carrying modular proteins derived from the parental T22-GFP-H6 (30.69 kDa, Figure 1), and the histidine-rich (Hn) sequences used here have a human origin (see Table 1). Further details of these proteins can be found elsewhere [48]. *E. coli* Origami B (BL21, $OmpT^-$, $Lon^-$, TrxB, $Gor^-$; Novagen) transformed cells were then grown at 37 ºC overnight in LB (Lysogeny Broth) and the encoded proteins were produced at 20 ºC overnight by induction of gene expression with 0.1 mM of Isopropyl-β-D-thiogalactopyronaside (IPTG) at an $OD_{550}$ of 0.5-0.7. Cells were harvested by centrifugation for 15 min (5,000g at 4 ºC) and pellets stored at -80ºC. Before use, pellets were then thaw and resuspended in Wash buffer (20 mM Tris, 500 mM NaCl, pH = 8) with protease inhibitors (Complete EDTA-free; Roche Diagnostics). Cell disruption was performed by three rounds of sonication (0.5-on, 0.5-off for 5 min) at 10 % of amplitude (Branson Digital Sonifier®) and the soluble fraction separated by centrifugation for 45 min (15.000g at 4ºC), and further filtered in pore diameters of 0.45 followed by 0.22 μm. Proteins were purified by Immobilized Metal Affinity Chromatography (IMAC) in an ÄKTA pure system (GE Healthcare) using HiTrap Chelating 5 mL columns (GE Healthcare). Protein elution was performed by a linear gradient of Elution buffer (20 mM Tris, 500 mM NaCl, 500 mM Imidazole, pH = 8) and eluted protein dialyzed against sodium carbonate (166 mM $NaCO_3H$, pH = 8) either-or sodium carbonate with salt (166 mM $NaCO_3H$, 333 mM NaCl, pH = 8) buffers.

*Protein purity, integrity and concentration, and nanoparticle formation*

Protein purity was checked by Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Western Blot (WB) immunoassay by using a 1/500 dilution of anti-GFP monoclonal antibody (Santa Cruz Biotechnology). Protein concentration was determined by Bradford's assay, and integrity evaluated by Matrix-Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) mass spectrometry. The formation of protein nanparticles was assessed by determining the volume size distribution by dynamic light scattering at 633 nm and 25 °C in a Zetasizer Nano ZS (Malvern Instruments Limited). For that, ZEN2112 3 mm quartz batch cuvettes were used in triplicate readings for error estimation, filled with protein samples in 1 mg/mL in carbonate buffer as described above.

*Protein internalization and endosomal escape*

HeLa CXCR4$^+$ cells were cultured in 24-well plates (60,000 cells/well during 24h for different time/concentration assays. 30,000 HeLa cells/well were cultured during 48 h for chloroquine and AMD competition), in all cases using MEM Alpha 1x GlutaMAX$^{TM}$ medium (Gibco) supplemented with foetal bovine serum (FBS) at 37ºC in a 5 % $CO_2$ humidified atmosphere, reaching a confluence of 70 %. Protein internalization was monitored by determining the cell green fluorescence after a harsh trypsin treatment at different exposure times. After exposure, cells were detached and externally attached protein removed by adding Trypsin-EDTA (Gibco®) at 1 mg/mL for 15 min and 37ºC. This is a particularly harsh trypsin treatment specifically developed to degrade cell surface-attached (but not internalized) protein, and aimed to prevent artefact readings by the contamination of externally attached fluorescent material [49]. Intracellular protein fluorescence was then determined by flow cytometry using a FACS-Canto system (Becton Dickinson) at 15 mW with an air-cooled argon ion laser exciting at 488 nm and corrected by the intrinsic fluorescence emission of each protein. Quantitative data are expressed as mean ± standard error. Additionally, the specific CXCR4-mediated protein internalization was probed by the addition of a CXCR4 antagonist (AMD3100) [50, 51], that blocks the interaction between T22 (MRRWCYRKCYKGYCYRKCR) and CXCR4. This compound was added at a final concentration of 500 nM (1:10 protein ratio) for 1 h prior protein incubation at 50 nM. Furthermore, the endosomal escape capacity of proteins was assessed by adding an inhibitor of endosomal acidification (chloroquine) at a final concentration of 100 μM (1:100 protein ratio) for 4 h prior protein incubation at 1000 nM.

OptiPRO™ SFM medium supplemented with L-glutamine was used for compound dilution and protein incubation. Since CQ prevents endosomal acidification and subsequent lysosomal formation in 4 h-experiments, the total amount of protein able to internalize and to reach the cell cytoplasm can be determined (+CQ). While comparing these values with the ones in absence of the drug (-CQ), the percentage of non-degraded protein (tagged as endosomal escape) can be calculated as % endosomal escape = (-CQ*100)/(+CQ). Those numbers are indicated on top of plot bars in Figure 1.

*Fluorescence emission*

GFP fluorescence emission was determined at 510 nm using an excitation wavelength of 488 nm with a Varian Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies). Before measuring, proteins were equally diluted in their corresponding sodium carbonate w/o salt buffers to 1 mg/mL in a final volume of 100 µL.

*Confocal imaging*

HeLa CXCR4+ cells were cultured in MatTek (MatTek Corporation) plates (60.000 cells/well during 24h) using MEM Alpha 1x GlutaMAX™ medium (Gibco) supplemented with fetal bovine serum (FBS) at 37ºC in a 5 % $CO_2$ humidified atmosphere and exposed to 1 µM of soluble protein for 24h. Cell nuclei were then labelled with Hoechst 33342 (ThermoFisher) at 5 µg/mL, displaying blue signal and plasma membranes with CellMask™ Deep Red (ThermoFisher) at 2.5 µg/mL displaying red signal, both incubated for 10 min at room temperature. Cells were then washed in phosphate saline buffer (Sigma-Aldrich) and images collected on an inverted TCS SP5 Spectral confocal microscope (Leica Microsystems) using 63× (1.4 NA) oil immersion objective lens. Hoechst 33342 excitation was reached *via* 405 nm blue diode laser, CellMask™ Deep Red *via* 633 nm line of a HeNe laser and GFP *via* 488 nm line of an argon ion laser. Interchannel crosstalk was avoided by configuring emission detection bandwidths, confocal pinhole set to 1 Airy unit and z-stacks acquisition intervals chosen to satisfy Nyquist sampling criteria. Multitrack sequential acquisition was used during the procedure and images analyzed by using Imaris v 7.2.1 software (Bitplane).

## 3. Results

We have explored how several variants of Hn peptides in cell-targeted protein nanoparticles would perform regarding cell penetrability and endosome membrane activity. They derive from the parental T22-GFP-H6, a fusion protein that assembles as nanoparticles of 12 nm, that have been used for precision delivery of chemical drugs [28] and protein toxins [29, 30] into $CXCR4^+$ tumor cells. Targeting is provided by the end-terminal peptide T22, a potent ligand of CXCR4 that promotes endosomal internalization [25]. To explore the impact of Hn in self-assembling, the H6 end-terminal peptide in T22-GFP-H6 had been replaced by the alternative histidine-rich peptides of similar length H3A, H5T and H5E, corresponding to HAAHAH, HTHTHTHTH and HEHEHEHEH amino acid sequences respectively. These peptides show at least 90 % of homology to sequences of human origin (Table 1), and as in the case of T22-GFP-H6, they also promote the spontaneous formation of nanoparticles ranging from 10 to 18 nm [48]. In this context, we have comparatively characterized the Hn variant nanoparticles, namely T22-GFP-H3A, T22-GFP-H5T, T22-GFP-H5E and T22-GFP-H6 (Figure 1A inset, Table 1) regarding their interactivity with target $CXCR4^+$ HeLa cells. No cell toxicity was observed in any of these materials upon exposure (Figure 1A), while all nanoparticles showed a comparable level of cell penetrability (Figure 1B). Uptake was, in all cases, dependent on the specific interaction of T22 with CXCR4, since the CXCR4 antagonist inhibited internalization with only a very low level of variability (from 95.5 % to 97.2 %, Figure 1C). At that point, endosomal escape was explored by using chloroquine (CQ), that prevents acidification and lysosomal protein degradation [40, 52]. The extent in which CQ enhances the fluorescence level is indicative of poor endosomal escape, while a null effect of the drug on the fluorescence would be indicative of a complete endosomal escape of the protein material. The addition of QC to $CXCR4^+$ HeLa cells in culture, before exposure to protein nanoparticles, showed a clearly differential impact on the intracellular fluorescence (Figure 1D), indicating a distinguishable membrane activity of Hn variants. In Figure 1 E, representative images of HeLa cells exposed to nanoparticles are shown, highlighting the endosomal route of cell penetration.

Because of such observed impact, we explored how the endosomal escape of the variant nanoparticles could be dependent on the number of histidine residues in the Hn tag. For that, we analyzed previously obtained physicochemical data on T22-GFP-H6 and T22-GFP-H5T, and we fully characterized T22-GFP-H3A and T22-GFP-H5E (Supplementary Table 1). All data sets

were plotted versus the number of histidine residues in the Hn tags and plotted to evaluate potential correlations. As observed, cell internalization (Figure 2A) and endosomal escape (Figure 2B) were not dependent on the number of histidine residues, being these parameters probably influenced by features of the Hn tags less apparent than the histidine residue number. As an expected positive control, the amount of imidazole necessary for the elution of fusion proteins in His-based chromatographic purification was largely dependent on the Hn value (Figure 2C). Interestingly, when plotting endosomal escape versus internalization, an unexpected but clear negative dependence was observed (Figure 2D). This observation might presumably reveal a general relationship between the receptor-dependent uptake of protein materials at the cell surface level and their ability to be released in the cytoplasm from uptaking vesicles. The possibility to unveil a generic mechanism in the cell penetrability of nanoparticles prompted us to further explore the endosome escape mechanisms of the Hn proteins, for what both internalization and endosomal escape were numerically modelled.

The model [53] that regulates the intensity of intracellular fluorescence $f(t)$ is

$$\frac{df}{dt} = k(L-f),\ f(0)=0,\ 0<f<L \implies f(t) = L(1-e^{-kt}) \quad (1)$$

where $k$ is the permeability of the cell membrane to the fluorescent proteins, and the fluorescence half-life $\tau = \frac{ln\,2}{k}$ is the proteolytic stability of the material (determined by its fluorescence). The typical evolution of $f(t)$ and $\frac{df}{dt}$ is displayed in Figures 3 A, B, where the blue line is for non-treated cells, the red line is for chloroquine-treated cells, and the green line measures the decay of fluorescence.

The initial rate of intracellular fluorescence rising is $v \equiv \left(\frac{df}{dt}\right)_{t=0} = kL$. Then, the fluorescence half-life can also be obtained as

$$\tau = \frac{\ln 2\ L}{v} \quad (2)$$

The addition of chloroquine, that minimizes lysosomal proteolysis of the material, affects both $k$ and $L$, and then, intracellular fluorescence evolves as described in the following differential approach

$$df = \frac{\partial f}{\partial L}dL + \frac{\partial f}{\partial k}dk = f\ d\ ln\,L + \frac{1}{2}(L-f)dk^2 \quad (3)$$

In the beginning, for $f \to 0$, the dominant term is accounted by the variation of $k$, namely $df \approx \frac{1}{2} L\, dk^2$. When $f \to \infty$, the dominant term is accounted by the variation of $L$, namely $df \approx L\, d\, ln\, L = dL$. During the experimental time, $f$ increases and the factor $\frac{1}{2}(L - f)$ decreases. Hence the first one will be prevalent. Their quotient $\frac{2f}{L-f}$ grows very fast and tends to 1 when

$$2f = L - f \Longrightarrow f = \frac{1}{3}L \Longrightarrow t = \frac{ln\frac{3}{2}}{ln\, 2}\, \tau = 0.4055\, \tau$$

Then, much earlier than for time $t = \tau$ the first term becomes dominant. Furthermore, in the experimental data, the values $d\, ln\, L \sim \Delta\, ln\, L$ are always greater than the values $dk^2 \sim \Delta k^2$. The experimental data of chloroquine-treated cells (labelled with prime) are taken as reference. Then, we always obtain, as seen in Supplementary Figure 1A,

$$\ln \mathrm{L} - \ln \mathrm{L}' < 0 \Leftrightarrow L < L'$$

$$k^2 - k'^2 > 0 \Leftrightarrow \tau < \tau'$$

It would be expected that at early times, before the proteolytic protective effects of chloroquine occur, the initial penetration rate would be the same in absence and in presence of the drug,

$$v = kL = v' = k'L' \Leftrightarrow \frac{L}{L'} = \frac{k'}{k} = \frac{\tau}{\tau'} \qquad (4)$$

The plot in Supplementary Figure 1B indicates that this is nearly (but not completely) the case. It must be noted that the quotient $\frac{v}{v'} = 1$ is lost in samples with the highest relative variation of fluorescence $\frac{L}{L'}$ (Supplementary Figure 1C). Also, the proportionality of Eq. (2) is similar when comparing all samples (Figure 3A, right), but it is lost for high values of $L$.

In summary, the inverse relationship Eq. (4) between $\frac{L}{L'}$ and fluorescence $\frac{k}{k'}$ is accomplished (Supplementary Figure 2A), and there is a clear negative correlation between the decrease of fluorescence and $L$ in chloroquine-treated cultures (Supplementary Figure 2B). Finally, the half-live growth is reduced when $L$ increases (Supplementary Figure 1C). Interestingly, $v = kL$ is nearly constant in absence and in presence of chloroquine, indicating that the drug acts predominantly at late endosome stages, and only in a residual (but detectable) way at early internalization phases.

The value $a = 0.8634$ obtained in the fitting $L - L' = aL'$ (Supplementary Figure 2B) is coincident with the slope shown in the Figure 2D. This fully supports the concept that endosomal escape mediated by histidine residues is specially favored when low amounts of materials are internalized by cells. On the contrary, cytosolic penetrability is inhibited, in relative terms, in the case of highly effective cell penetrating nanoparticles, indicating a saturation mechanism in the endosomal release of engulfed nanoparticles. As this fact has been profiled by the whole set of Hn-carrying materials it might reflect a general law that at least should be valid for His-mediated endosomal escape processes of protein nanoparticles.

**4. Discussion**

The peptide H6 is a widely spread tags for the one-step purification of recombinant proteins through $Ni^{2+}$-based affinity chromatography [46, 54]. Upon purification, the presence of H6 or related peptides in the final protein product might have unexpected biological impacts, such as aberrant biodistribution of the protein or protein complex upon systemic injection [55, 56]. In this context, and also envisaging potential immunoreactions in protein materials intended for clinical use, the removal of H6 from purified proteins would be a desirable biotechnological step [44]. However, in the biological fabrication of protein-based materials, Hn peptides show interesting roles as architectonic tags as metal-dependent cross-molecular binders [12, 57, 58]. H6 has been efficiently used to force the formation of protein nanoparticles [59, 60], designed to transport either small molecular drugs [61] or cytotoxic proteins in cancer therapies [26, 31, 62]. In addition, and facing lysosomal degradation as a major bottleneck in the delivery of protein nanoparticles [40], Hn peptides offer an added value to protein supramacromolecular complexes for intracellular delivery. In contrast to other diverse pathways of endosomal escape [40, 63], Hn peptides promote endosomal escape of macromolecules and complexes by a proton-sponge mechanism [40], that has not been so far investigated in deep. To combine the appealing properties of Hn peptides regarding protein purification, nanoparticle formation and endosomal escape, limiting potential immunotoxicity risks, we have explored the performance of several human or human-like Hn peptides, previously explored as not immunogenic (Table 1 and references therein), regarding their ability to promote endosomal escape. In form of T22-GFP-Hn constructions, targeted to

$CXCR4^+$ cells through T22 [23], the four tested Hn segments are not cytotoxic, penetrate target cells with high level of CXCR4-dependence (then, by an endosomal route), and respond differentially to the presence of chloroquine (Figure 1), that prevents lysosomal destruction [52]. In particular, T22-GFP-H5E promotes an important endosomal escape slightly higher (23.66 %) than that triggered by H6 (Figure 1). Even if the efficacy of H5E as an endosomolytic agent is not dramatically higher than that H6, the discovery of a humanized Hn version that performs slightly better than H6 in cytosolic delivery exemplifies the potential of optimized histidine tails to promote endosomal escape. In addition, the peptide HEHEHE does not promote hepatic accumulation of associated polypeptides [56, 64], probably because of its hydrophilic nature when compared to similar Hn constructs [65].

As a side observation, we demonstrate that the presence of chloroquine enhances the intracellular permanence of uptaken proteins. Since chloroquine inhibits acidification of endocytic vesicles [66] and the only type of endocytic vesicles that undergo acidification is that based on clathrin [67], the experimental data presented here strongly support that T22-empowered GFP-based nanoparticles exploit the clathrin-endosomal route for internalization. Interestingly, all T22-GFP-Hn particles bind specifically the cytokine receptor CXCR4 for cell internalization, as the CXCR4 antagonist AMD3100 blocks their entrance (Figure 1 C). Also, CXCR4 is also used by the Human Immunodeficiency Virus as a co-receptor for cell infection [68], and this virus enters cells via clathrin-mediated endocytosis [69]. Altogether, these data indicate that binding to CXCR4 favors consequent internalization by the clathrin-endosomal route.

In a more transversal context, when comparing the endosomal escape of the proteins (inversely correlating with the increase of intracellular fluorescence mediated by chloroquine) with the intensity of intracellular fluorescence, it rendered an unexpected negative correlation (Figure 2D). This fact was fully confirmed by an accurate dynamic modeling of the intracellular fluorescence upon exposure to the fluorescent proteins (Supplementary Figure 2). This analysis strongly supports the concept that endosomal escape, as mediated by Hn peptides, is saturated at moderated amounts of internalized protein, what has critical implications in the design of nanoscale protein materials for intracellular drug delivery. For protein drugs [33, 43], that are intrinsically sensitive to lysosomal destruction, developing agents to enhance cell

penetrability might be then result in a nonlinear gain in cytosolic delivery. However, regarding protein-drug conjugates, in which the effective drug is resistant to lysosomal activities, higher cell internalization should result in higher drug activity. Irrespective of that, we have here identified a set of humanized His-rich peptides that can be used as substitutes of H6 when engineering endosomal escape, which sufficiently keep the functionalities of H6 during chromatographic purification of recombinant proteins.

## 5. Conclusions

We have demonstrated here that humanized His-rich segments, with a number of His residues below 6, can efficiently promote endosomal escape of modular proteins targeted to $CXCR4^{+}$ cells. In particular, the peptide HEHEHEHEH, appears as a promising tool to favor the intracellular trafficking and the proteolytic stability of the cargo protein, while it keeps the functional binding to Ni2+. The intensity of endosomal escape promoted by a series of humanized Hn peptides negatively correlates with the amount of internalized protein, revealing a saturation mechanism in the Hn-mediated cytosolic penetration. Finally, chloroquine, that mainly acts at the level of late endosomes, has a residual effect on the first stages of protein uptake, slightly enhancing the velocity of cellular penetration at the early endosome stages.

## Acknowledgements

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. Biodistribution studies were performed by the ICTS “NANBIOSIS”, Nantoxicology Unit (http://www.nanbiosis.es/portfolio/u18-nanotoxicology-unit/). Electron microscopy studies were performed by the Servei de Microscòpia at the UAB. AV received an ICREA ACADEMIA award. HLL was supported by a predoctoral fellowship from AGAUR (2019 FI_B 00352) and UU by PERIS program from the Health Department of the Generalitat de Catalunya.

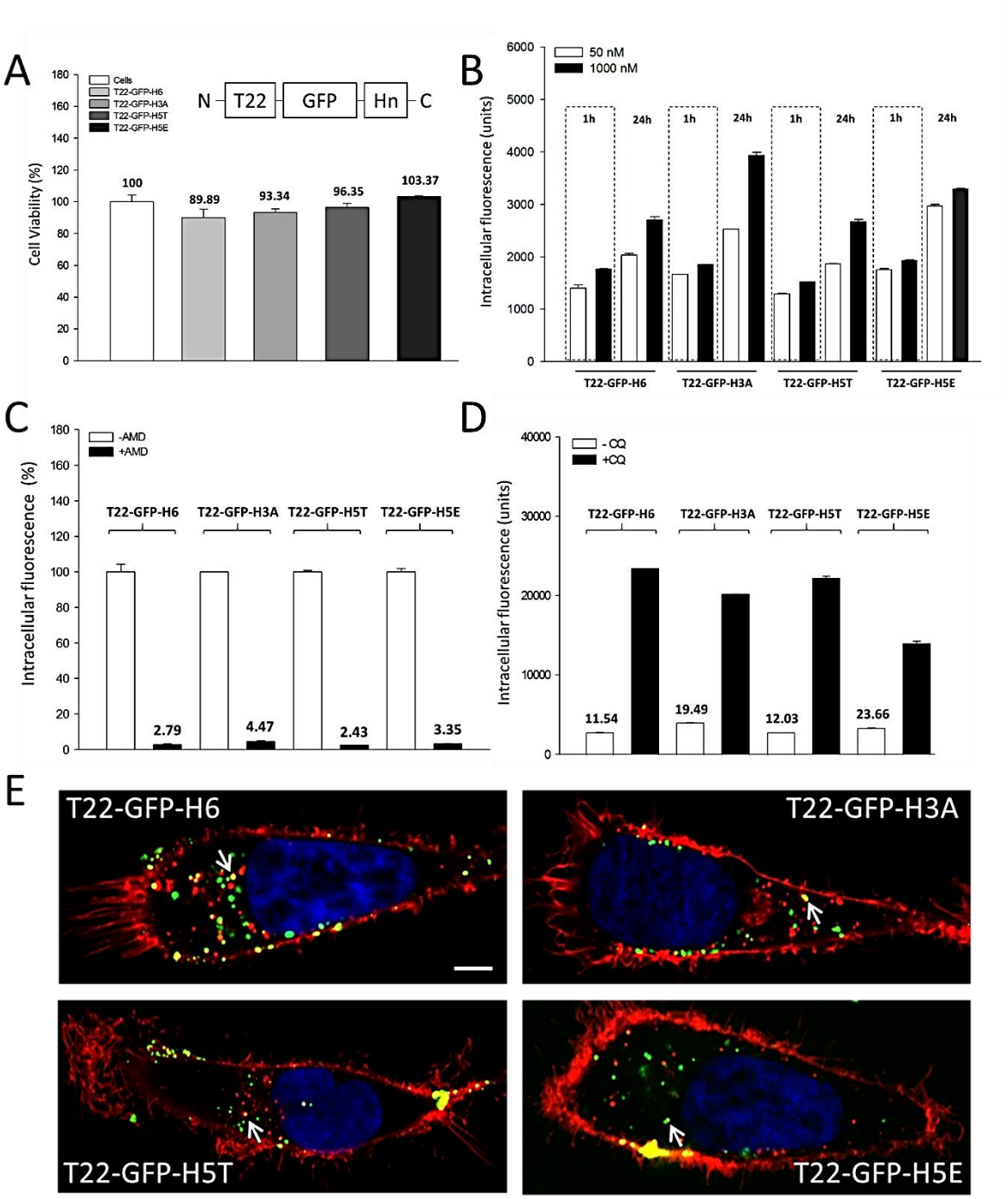


**Figure 1. Cell interactivity of Hn-carrying protein nanoparticles. A.** Cytotoxicity of Hn protein nanoparticles (at 2 µM) on CXCR4⁺ HeLa cells upon 48 h of exposure. Numbers indicated on top of plot bars represent cell viability in percentage regarding the control cell culture. In the inset, schematic representation of protein modular distribution. N=3**. B.** Internalization of protein nanoparticles added to 50 or 1000 nM for 1 and 24 h in CXCR4⁺ HeLa cell cultures. Data on H6 and 5HT were published previously [48]. N=2. **C.** Inhibition of protein uptake (added to 50 nM for 24 h) by the CXCR4⁺ antagonist AMD3100 at 500 nM, after 1 h of exposure. Numbers indicated on top of plot bars represent percentage of internalization upon AMD addition. Data on H6 and H5T were published previously [48]. N=2. **D.** Internalization and endosomal escape of protein nanoparticles (added to 1 µM for 24 h) in CXCR4⁺ HeLa cells upon chloroquine (CQ) addition. CQ was used to prevent endosomal acidification at 100 µM for 4 h. Numbers indicated on top of plot bars represent the estimated endosomal escape of

the uptake material, in percentage. N=2. **E**. Confocal sections of CXCR4$^+$ HeLa cells 24h of exposure to different nanoparticles. Membranes are labelled in red, the nuclear section in blue, and the green signal is the natural fluorescence of the assembled GFP. Arrows indicate yellowish emission, resulting from red and green merging into endosomal compartment.

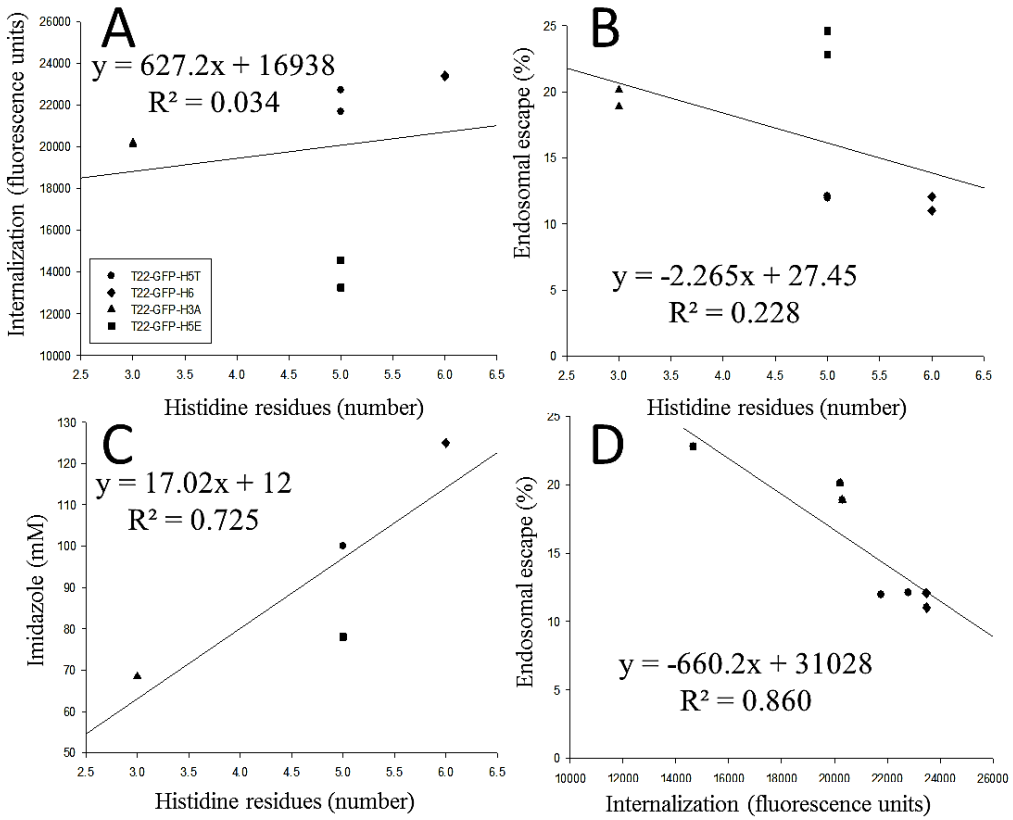


**Figure 2. Impact of the His number of the Hn tag on cell interactivity.** Correlation between the number of histidine residues in the Hn tag with cell internalization (**A**), endosomal escape (**B**) and eluting amount of imidazole (**C**). In **D**, endosomal escape versus internalization is plotted. Data sets used represented here are shown in Supplementary Table 1.

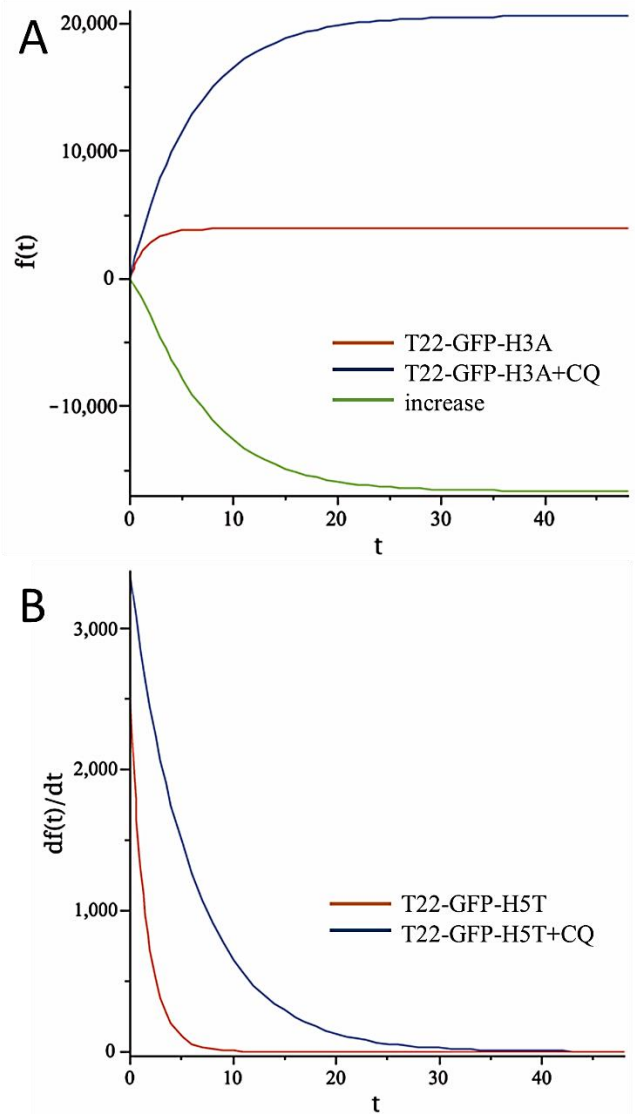


**Figure 3. Variation of intracellular fluorescence. A.** Dynamics of intracellular fluorescence. The saturation values is *L*. The initial increase is greater with higher values of *k*. **B**. Velocity of variation of intracellular fluorescence. The initial value is $v = k\,L$.

**Table 1**. Homology between selected His-rich sequences and human (*Homo sapiens*) proteins.

| H-rich peptide | BLAST protein | Uniprot code | Subcellular localization and/or function | FASTA sequence [a] | Nanoparticle size (nm) [b] |
|---|---|---|---|---|---|
| -H6 (HHHHHH) | Histidine-rich carboxyl terminus protein 1 | Q6UXD1 | Single-pass membrane protein | ...$^{85}$VGL<u>HHHHHH</u>PRHTP HHL<u>HHHHHH</u>PHR$^{108}$... | 11.7 ± 0.01 |
| | Ubiquitin carboxyl-terminal hydrolase 34 | Q70CQ2 | Cytosolic protein | ...$^{746}$GPQ<u>HHHHHHHHHH HHH</u>DGH$^{763}$... | |
| | Voltage-dependent P/Q-type Calcium channel subunit alpha-1A | O00555 | Multi-pass membrane protein | ...$^{2207}$HRQ<u>HHHHHHHHH HH</u>PPP$^{2222}$... | |
| | Angiomotin | Q4VCS5 | Cell surface transmembrane protein | ...$^{382}$QQQ<u>HHHHHHH</u>Q QQ$^{394}$... | |
| -H3A (HAAHAH) | Thrombospondin type-1 domain-containing protein 7B | Q9C0I4 | Single-pass membrane protein type-1 | ...$^{27}$LLS<u>HAAH</u>LEG$^{36}$... | 18.1 ± 0.00 |
| | Biotinidase | P43251 | Extracellular or secreted protein | ...$^{1}$MA<u>HAH</u>IQG$^{8}$... | |
| -H5T (HTHTHTHTH) | Fibroblast growth factor receptor-like 1 | Q8N441 | Single-pass membrane protein type-1 | ...$^{476}$TDI<u>HTHTHTH</u>SHT$^{488}$... | 10.1 ± 0.01 |
| | Tumor differentiation factor | Q6T424 | Excreted protein | ...$^{23}$QDT<u>HTHTHTHTHT HTH</u>TNT$^{41}$... | |
| -H5E (HEHEHEHEH) | Neuroendocrine secretory protein 55 | O95467 | Multi-pass membrane protein | ...$^{83}$ESD<u>HEHE</u>EAD$^{92}$... | 10.9 ± 0.04 |
| | Zinc transporter ZIP9 | Q9NUM3 | Multi-pass membrane protein | ...$^{90}$SVV<u>HEHEH</u>SHD$^{100}$... | |
| | Repetin protein | Q6XPR3 | Multi-pass membrane protein | ...$^{745}$RQT<u>HEHE</u>QSH$^{754}$... | |

[a] Underlined amino acid sequences are those matching in the FASTA search. Flanking sequences (three residues each end) are also indicated together with the original amino acid numbering of the protein.

[b] Hydrodynamic diameter of the resulting nanoparticles was previously determined (Table 1 in [48]) and confirmed here by DLS.

**Acknowledgments**. We are indebted to Agencia Estatal de Investigación (AEI) and Fondo Europeo de Desarrollo Regional (FEDER) (grant BIO2016-76063-R, AEI/FEDER, UE) granted to AV, AGAUR (2017SGR-229) granted to AV, CIBER-BBN (project NANOPROTHER and NANOSCAPE) granted to AV and UUE respectively, and to ISCIII (PI15/00272 co-founding FEDER) granted to EV. 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/). We are also grateful to SCAC (UAB) for cell culture facilities and assistance. HLL was supported by AGAUR (2019FI_B_00352) and UU by PERIS program from the health department of la Generalitat de Catalunya. AV holds an ICREA ACADEMIA award.

**Disclosures**. RM, EV and AV are cofounders of NANOLIGENT SL, devoted to develop protein-based antitumoral drugs.

References

[1] Kim S, Kim D, Jung HH, Lee IH, Kim JI, Suh JY, et al. Bio-inspired design and potential biomedical applications of a novel class of high-affinity peptides. Angewandte Chemie 2012;51:1890-4.
[2] Knowles TPJ, Mezzenga R. Amyloid Fibrils as Building Blocks for Natural and Artificial Functional Materials. Advanced materials 2016;28:6546-61.
[3] Kobayashi N, Arai R. Design and construction of self-assembling supramolecular protein complexes using artificial and fusion proteins as nanoscale building blocks. Current opinion in biotechnology 2017;46:57-65.
[4] Kumar VA, Wang BK, Kanahara SM. Rational design of fiber forming supramolecular structures. Exp Biol Med 2016;241:899-908.
[5] Li D, Jones EM, Sawaya MR, Furukawa H, Luo F, Ivanova M, et al. Structure-Based Design of Functional Amyloid Materials. Journal of the American Chemical Society 2014;136:18044-51.
[6] Loo Y, Goktas M, Tekinay AB, Guler MO, Hauser CA, Mitraki A. Self-Assembled Proteins and Peptides as Scaffolds for Tissue Regeneration. Advanced healthcare materials 2015;4:2557-86.
[7] Molino NM, Wang SW. Caged protein nanoparticles for drug delivery. Current opinion in biotechnology 2014;28:75-82.
[8] 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.
[9] Sutherland TD, Huson MG, Rapson TD. Rational design of new materials using recombinant structural proteins: Current state and future challenges. Journal of structural biology 2018;201:76-83.
[10] Yeates TO. Geometric Principles for Designing Highly Symmetric Self-Assembling Protein Nanomaterials. Annual review of biophysics 2017;46:23-42.
[11] Yeates TO, Liu Y, Laniado J. The design of symmetric protein nanomaterials comes of age in theory and practice. Current opinion in structural biology 2016;39:134-43.
[12] Zou R, Wang Q, Wu J, Wu J, Schmuck C, Tian H. Peptide self-assembly triggered by metal ions. Chemical Society reviews 2015;44:5200-19.
[13] Corchero JL, Vazquez E, Garcia-Fruitos E, Ferrer-Miralles N, Villaverde A. Recombinant protein materials for bioengineering and nanomedicine. Nanomedicine 2014;9:2817-28.
[14] Ferrer-Miralles N, Rodriguez-Carmona E, Corchero JL, Garcia-Fruitos E, Vazquez E, Villaverde A. Engineering protein self-assembling in protein-based nanomedicines for drug delivery and gene therapy. Critical reviews in biotechnology 2015;35:209-21.
[15] Stefan Wilhelm AJT, Qin Dai, Seiichi Ohta, Julie Audet, Harold F. Dvorak & Warren C. W. Chan. Analysis of nanoparticle delivery to tumours. Nature Reviews Materials 2016;1.
[16] Vazquez E, Mangues R, Villaverde A. Functional recruitment for drug delivery through protein-based nanotechnologies. Nanomedicine 2016;11:1333-6.
[17] Srinivasarao M, Galliford CV, Low PS. Principles in the design of ligand-targeted cancer therapeutics and imaging agents. Nature reviews Drug discovery 2015;14:203-19.
[18] Unzueta U, Cespedes MV, Vazquez E, Ferrer-Miralles N, Mangues R, Villaverde A. Towards protein-based viral mimetics for cancer therapies. Trends in biotechnology 2015;33:253-8.

[19] 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.
[20] Voltà-Durán E, O C-G, Serna N, López-Laguna H, Sánchez-García L, Pesarrodona M, et al. Controlling self-assembling and tumor cell-targeting of protein-only nanoparticles through modular protein engineering. Sicence China Materials 2019.
[21] 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. Acta biomaterialia 2018.
[22] Unzueta U, Serna N, Sanchez-Garcia L, Roldan M, Sanchez-Chardi A, Mangues R, et al. Engineering multifunctional protein nanoparticles by in vitro disassembling and reassembling of heterologous building blocks. Nanotechnology 2017;28:505102.
[23] 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.
[24] 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. International journal of pharmaceutics 2014;473:286-95.
[25] 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. International journal of nanomedicine 2012;7:4533-44.
[26] 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. Science China Materials 2019;62:892–8.
[27] 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 : nanotechnology, biology, and medicine 2016;12:1241-51.
[28] 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 molecular medicine 2018.
[29] 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. Advanced Functional Materials 2017;27:1700919.
[30] 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. Journal of controlled release : official journal of the Controlled Release Society 2018.
[31] 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.
[32] Ferrer-Miralles N, Corchero JL, Kumar P, Cedano JA, Gupta KC, Villaverde A, et al. Biological activities of histidine-rich peptides; merging biotechnology and nanomedicine. Microbial cell factories 2011;10:101.
[33] Serna N, Sanchez-Garcia L, Unzueta U, Diaz R, Vazquez E, Mangues R, et al. Protein-Based Therapeutic Killing for Cancer Therapies. Trends in biotechnology 2018;36:318-35.
[34] Casanova I, Unzueta U, Arroyo-Solera I, Cespedes MV, Villaverde A, Mangues R, et al. Protein-driven nanomedicines in oncotherapy. Current opinion in pharmacology 2019;47:1-7.

[35] Shete HK, Prabhu RH, Patravale VB. Endosomal escape: a bottleneck in intracellular delivery. Journal of nanoscience and nanotechnology 2014;14:460-74.
[36] Lv J, Fan Q, Wang H, Cheng Y. Polymers for cytosolic protein delivery. Biomaterials 2019;218:119358.
[37] Hua XW, Bao YW, Wu FG. Fluorescent Carbon Quantum Dots with Intrinsic Nucleolus-Targeting Capability for Nucleolus Imaging and Enhanced Cytosolic and Nuclear Drug Delivery. ACS applied materials & interfaces 2018;10:10664-77.
[38] Zhang Z, Shen W, Ling J, Yan Y, Hu J, Cheng Y. The fluorination effect of fluoroamphiphiles in cytosolic protein delivery. Nat Commun 2018;9:1377.
[39] Liu C, Wan T, Wang H, Zhang S, Ping Y, Cheng Y. A boronic acid-rich dendrimer with robust and unprecedented efficiency for cytosolic protein delivery and CRISPR-Cas9 gene editing. Science advances 2019;5:eaaw8922.
[40] Varkouhi AK, Scholte M, Storm G, Haisma HJ. Endosomal escape pathways for delivery of biologicals. Journal of controlled release : official journal of the Controlled Release Society 2011;151:220-8.
[41] Sanchez-Garcia L, Martin L, Mangues R, Ferrer-Miralles N, Vazquez E, Villaverde A. Recombinant pharmaceuticals from microbial cells: a 2015 update. Microbial cell factories 2016;15:33.
[42] Bruzzoni-Giovanelli H, Alezra V, Wolff N, Dong CZ, Tuffery P, Rebollo A. Interfering peptides targeting protein-protein interactions: the next generation of drugs? Drug discovery today 2017.
[43] Agyei D, Ahmed I, Akram Z, Iqbal HM, Danquah MK. Protein and Peptide Biopharmaceuticals: An Overview. Protein and peptide letters 2017;24:94-101.
[44] Spriestersbach A, Kubicek J, Schafer F, Block H, Maertens B. Purification of His-Tagged Proteins. Methods in enzymology 2015;559:1-15.
[45] DePalma A. Keeping Tabs on Polyhistidine Tags. Genetic Engineering & Biotechnology News 2016;36:3.
[46] Gaberc-Porekar V, Menart V. Perspectives of immobilized-metal affinity chromatography. Journal of biochemical and biophysical methods 2001;49:335-60.
[47] Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Applied microbiology and biotechnology 2003;60:523-33.
[48] López-Laguna H, Sala R, Sánchez JM, Álamo P, Unzueta U, Sánchez-Chardi A, et al. Nanostructure Empowers Active Tumor Targeting in Ligand-Based Molecular Delivery. Particle and Particle Characterization System 2019.
[49] 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. The Journal of biological chemistry 2003;278:585-90.
[50] 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.
[51] 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.
[52] Fredericksen BL, Wei BL, Yao J, Luo T, Garcia JV. Inhibition of endosomal/lysosomal degradation increases the infectivity of human immunodeficiency virus. Journal of virology 2002;76:11440-6.

[53] de Pinho Favaro MT, Serna N, Sanchez-Garcia L, Cubarsi R, Roldan M, Sanchez-Chardi A, et al. Switching cell penetrating and CXCR4-binding activities of nanoscale-organized arginine-rich peptides. Nanomedicine : nanotechnology, biology, and medicine 2018.
[54] Knecht S, Ricklin D, Eberle AN, Ernst B. Oligohis-tags: mechanisms of binding to Ni2+-NTA surfaces. Journal of molecular recognition : JMR 2009;22:270-9.
[55] Goel A, Colcher D, Koo JS, Booth BJ, Pavlinkova G, Batra SK. Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct. Biochimica et biophysica acta 2000;1523:13-20.
[56] Altai M, Liu H, Orlova A, Tolmachev V, Graslund T. Influence of molecular design on biodistribution and targeting properties of an Affibody-fused HER2-recognising anticancer toxin. International journal of oncology 2016;49:1185-94.
[57] JM S, H L-L, P Á, N S, A S-C, V N, et al. Artificial inclusion bodies for clinical development Advanced science 2019.
[58] Zou Q, Yan X. Amino Acid Coordinated Self-Assembly. Chemistry 2018;24:755-61.
[59] 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. Acta biomaterialia 2019;83:257-64.
[60] Eric V-D, Olivia C-G, Naroa S, Héctor L-L, Laura S-G, Mireia P, et al. Controlling self-assembling and tumor cell-targeting of protein-only nanoparticles through modular protein engineering. Sicence China Materials 2019.
[61] 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 molecular medicine 2018;10.
[62] 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. Journal of controlled release : official journal of the Controlled Release Society 2018;274:81-92.
[63] El-Sayed A, Khalil IA, Kogure K, Futaki S, Harashima H. Octaarginine- and octalysine-modified nanoparticles have different modes of endosomal escape. The Journal of biological chemistry 2008;283:23450-61.
[64] Hofstrom C, Altai M, Honarvar H, Strand J, Malmberg J, Hosseinimehr SJ, et al. HAHAHA, HEHEHE, HIHIHI, or HKHKHK: influence of position and composition of histidine containing tags on biodistribution of [(99m)Tc(CO)3](+)-labeled affibody molecules. J Med Chem 2013;56:4966-74.
[65] Mitran B, Altai M, Hofstrom C, Honarvar H, Sandstrom M, Orlova A, et al. Evaluation of 99mTc-Z IGF1R:4551-GGGC affibody molecule, a new probe for imaging of insulin-like growth factor type 1 receptor expression. Amino acids 2015;47:303-15.
[66] Al-Bari MAA. Targeting endosomal acidification by chloroquine analogs as a promising strategy for the treatment of emerging viral diseases. Pharmacology research & perspectives 2017;5:e00293.
[67] Takei K, Haucke V. Clathrin-mediated endocytosis: membrane factors pull the trigger. Trends in cell biology 2001;11:385-91.
[68] Donzella GA, Schols D, Lin SW, Este JA, Nagashima KA, Maddon PJ, et al. AMD3100, a small molecule inhibitor of HIV-1 entry via the CXCR4 co-receptor. Nature medicine 1998;4:72-7.
[69] von Kleist L, Stahlschmidt W, Bulut H, Gromova K, Puchkov D, Robertson MJ, et al. Role of the clathrin terminal domain in regulating coated pit dynamics revealed by small molecule inhibition. Cell 2011;146:471-84.