



# The spectrum of building block conformers sustains the biophysical properties of clinically-oriented self-assembling protein nanoparticles

Eric Voltà-Durán<sup>1,2,3</sup>, Julieta M Sánchez<sup>1,2,3,4</sup>, Hèctor-López-Laguna<sup>1,2,3</sup>, Eloi Parladé<sup>1,2,3</sup>, Laura Sánchez-García<sup>1,2,3#</sup>, Alejandro Sánchez-Chardi<sup>5,6</sup>, Ario de Marco<sup>7</sup>, Ugutz Unzueta<sup>2,3,8</sup>, Esther Vázquez<sup>1,2,3</sup> and Antonio Villaverde<sup>1,2,3\*</sup>

**ABSTRACT** Histidine-rich peptides confer self-assembling properties to recombinant proteins through the supramolecular coordination with divalent cations. This fact allows the cost-effective, large-scale generation of microscopic and macroscopic protein materials with intriguing biomedical properties. Among such materials, resulting from the simple bioproduction of protein building blocks, homomeric nanoparticles are of special value as multivalent interactors and drug carriers. Interestingly, we have here identified that the assembly of a given His-tagged protein might render distinguishable categories of self-assembling protein nanoparticles. This fact has been scrutinized through the nanobody-containing fusion proteins EM1-GFP-H6 and A3C8-GFP-H6, whose biosynthesis results in two distinguishable populations of building blocks. In one of them, the assembling and disassembling is controllable by cations. However, a second population immediately self-assembles upon purification through a non-regulatable pathway, rendering larger nanoparticles with specific biological properties. The structural analyses of both model proteins and nanoparticles revealed important conformational variability in the building blocks. This fact renders different structural and functional categories of the final soft materials resulting from the participation of energetically unstable intermediates in the oligomerization process. These data illustrate the complexity of the His-mediated protein assembling in recombinant proteins but they also offer clues for a better design and refinement of protein-based nanomedicines, which, resulting from biological fabrication, show an architectonic flexibility unusual among biomaterials.

**Keywords:** recombinant proteins, self-assembling, protein materials, nanoparticles, biomaterials

## INTRODUCTION

Histidine-rich peptides, when genetically fused to recombinant proteins, confer them the capability to self-assemble as different kinds of functional protein-only materials including fibers, nanoparticles (NPs) and microparticles, through their cross-molecular coordination with divalent cations [1]. This happens because the interactivity exhibited by the hexahistidine tag (H6) and other His-rich tails with Ni<sup>2+</sup> during the immobilized metal affinity chromatographic (IMAC) purification of recombinant proteins [2,3] can be also exploited for a controlled cross-molecular assembly, when adding defined amounts of Zn<sup>2+</sup>, Ca<sup>2+</sup>, Mn<sup>2+</sup> or other divalent cations to solutions of pure His-tagged protein [4–6]. In fact, this principle sustains the formation of the secretory granules in the mammalian endocrine system [7,8] and of different types of amyloid and non-amyloid protein materials existing in nature [4,7,9–16]. The simplicity of His-rich peptides used as architectonic agents at the nanoscale offers an interesting alternative to more refined approaches to control protein self-assembling, that might require more sophisticated protein engineering [17–21].

We have previously generated a family of recombinant modular proteins, based on an N-terminal cationic peptide and a C-terminal polyhistidine (mainly H6), which, when flanking a protein of interest, self-assemble as homomeric NPs of around 15–20 nm [22]. Since such an assembling platform is fully transversal and highly robust, and the resulting NPs are stable upon *in vivo* administration [23], these materials have been developed as effective nanocarriers for conventional antitumoral drugs in colorectal cancer [24], lymphoma [25,26] and acute myeloid leukemia [27]. In addition, the incorporation of pro-apoptotic peptides [28], toxins [29] or venoms [30] to the modular protein constructs allows generating cytotoxic building blocks that, if targeted through solvent-exposed ligands of

<sup>1</sup> Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra 08193, Barcelona, Spain

<sup>2</sup> Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra 08193, Barcelona, Spain

<sup>3</sup> CIBER de Bioingeniería Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra 08193, Barcelona, Spain

<sup>4</sup> Instituto de Investigaciones Biológicas y Tecnológicas (IIBYT) (CONICET-Universidad Nacional de Córdoba), ICTA & Cátedra de Química Biológica, Departamento de Química, FCEfYN, UNC. Av. Velez Sarsfield 1611, X 5016GCA Córdoba, Argentina

<sup>5</sup> Servei de Microscòpia, Universitat Autònoma de Barcelona, Bellaterra 08193, Barcelona, Spain

<sup>6</sup> Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia, Universitat de Barcelona, 08028 Barcelona, Spain

<sup>7</sup> Laboratory for Environmental and Life Sciences, University of Nova Gorica, Nova Gorica 5000, Slovenia

<sup>8</sup> Biomedical Research Institute Sant Pau (IIB Sant Pau), Sant Antoni M<sup>a</sup> Claret 167, 08025 Barcelona, Spain

# Present address: Basic Sciences Department, Faculty of Medicine and Health Sciences, Universitat Internacional de Catalunya, 08195 Sant Cugat del Vallès, Spain

\* corresponding author (email: [antoni.villaverde@uab.cat](mailto:antoni.villaverde@uab.cat))

tumoral markers, result in NPs with selective, build-in anti-tumoral activity for precision medicine [31]. The polyhistidine tail, apart from being an architectonic agent, allows the one-step purification of the recombinant building blocks from the bacterial cell extracts, upon their mechanical disruption.

During the production of these types of NPs, we have occasionally observed some structural variability in the oligomers resulting from self-assembling [32]. Although from a production point of view, this fact can be simply overcome by selecting the desired material population in the chromatography, the extent, causes and biological significance of such variability are not known. In a recent development of two versions of fluorescent and nanostructured nanobody fusions [33], such variability was specially apparent. Therefore, we decided to take these two proteins (A3C8-GFP-H6 and EM1-GFP-H6, Fig. 1a, b) as models to investigate the categories of particle subpopulations generated through the H6-based platform, why they assemble in disparate material versions and how they functionally perform in biological interfaces.

## EXPERIMENTAL SECTION

### Protein design, production and purification

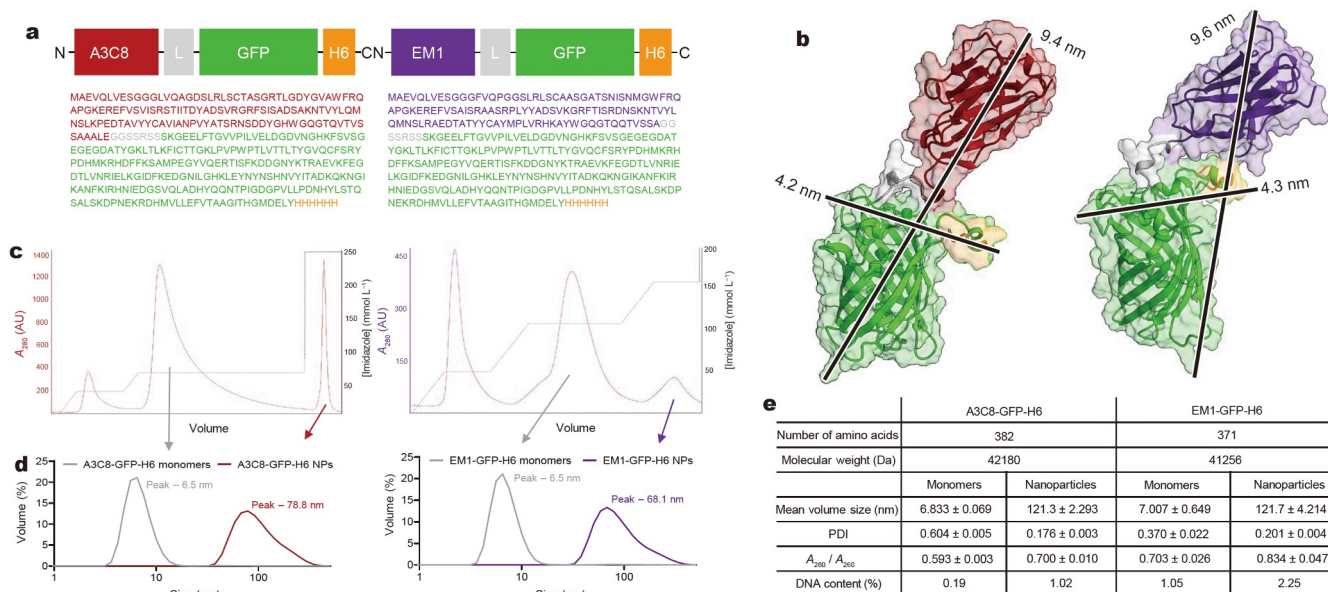
*Escherichia coli* (*E. coli*) codon-optimized A3C8-GFP-H6 and EM1-GFP-H6 genes were expressed as inserted in the plasmid pET22b. Both genes encode nanobody and green fluorescent protein (GFP) fusions in the form of His-tagged proteins [33]. A3C8 shows specificity towards the plant toxin ricin, while EM1 has specificity towards the Her2 receptor, overexpressed in breast cancer cells. Proteins were produced in *E. coli* Origami B (BL21,  $\text{OmpT}^-$ ,  $\text{Lon}^-$ ,  $\text{TrxB}^-$ ,  $\text{Gor}^-$ ) (Novagen, Germany) at

20°C overnight, upon addition of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG, at 0.1  $\text{mmol L}^{-1}$  for A3C8-GFP-H6 and 1  $\text{mmol L}^{-1}$  for EM1-GFP-H6). Bacterial lysates were obtained by French Press (Thermo FA-078A) disruption at 1200 psi (1 psi =  $6.89 \times 10^3$  Pa) in wash buffer (20  $\text{mmol L}^{-1}$  Tris-HCl, 500  $\text{mmol L}^{-1}$  NaCl, 10  $\text{mmol L}^{-1}$  imidazole, pH 8.0) with the protease inhibitor cComplete ethylene diamine tetraacetic acid (EDTA)-free (Roche, USA). A recombinant version of ricin, namely T22-mRTA-H6, was produced and purified as previously described [34].

An ÄKTA Pure FPLC system (GE Healthcare, USA) was used for the IMAC in a His-Trap HP column (GE Healthcare, USA). Protein elution was achieved by lineally increasing the molar amount of imidazole in the column through an elution buffer (20  $\text{mmol L}^{-1}$  Tris-HCl, 500  $\text{mmol L}^{-1}$  NaCl, 500  $\text{mmol L}^{-1}$  imidazole, pH 8.0). Protein purity and integrity was assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), western blot (anti-His, Santa Cruz Biotechnology, USA) and matrix assisted laser-desorption ionization time-of-flight (MALDI-TOF). Proteins were finally dialyzed against sodium bicarbonate with salt buffer (166  $\text{mmol L}^{-1}$   $\text{NaHCO}_3$ , 333  $\text{mmol L}^{-1}$  NaCl, pH 8.0). Protein concentration was determined by the Bradford assay and Nanodrop was used to measure the  $A_{280}/A_{260}$  (absorbance at 280/260 nm) ratio, indicative of DNA presence in the sample.

### Three-dimensional protein modeling and visualization

Structure prediction of each nanobody-based modular protein was performed to estimate the size of the building block, using the RosettaCM [35] high-resolution approach *via* the Robetta webserver [36]. Specifically, EM1, A3C8 and GFP-H6 domains



**Figure 1** Protein materials. (a) Modular disposition of A3C8-GFP-H6 and EM1-GFP-H6 chimeric proteins. Amino acid sequences are shown. An N-terminal nanobody (A3C8, in dark red; EM1, in dark purple) is fused to a GFP (in green) with a C-terminal hexa-histidine tag (H6, in yellowish orange) through a short peptidic linker (L, in grey). Color legend is maintained hereafter. (b) Three-dimensional models of the monomeric form of each protein. Black lines indicate the estimated main and cross axis distances (in nm) extracted from the models. (c) IMAC purification chromatograms of both proteins. The first peak corresponds to untagged bacterial proteins that are weakly attached to the column through nonspecific binding. Two pure populations were obtained in each case, eluting at different concentrations of imidazole. Left  $y$ -axis and right  $y$ -axis indicate the  $A_{280 \text{ nm}}$  (AU) and the imidazole concentration ( $\text{mmol L}^{-1}$ ), respectively, when elution of each mL ( $x$ -axis) occurs. (d) Volume size distribution (nm) of each purified population measured by DLS. Monomeric populations of each protein are represented in gray, whereas NPs are colored as previously stated. Values depicted indicate the mode size of each peak. (e) Table summarizing the properties of both proteins, in their respective monomeric or NP forms. Data are indicated as mean  $\pm$  standard error of the mean.

were modelled individually and manually assembled in UCSF Chimera (v1.14) to serve as quality templates for the prediction of the whole modular construct. Parameters were set to 100 sampling models, 1 register shift, and a probability of 0.1 of sampling fragments within template regions. Best candidates were chosen by the highest confidence score and lowest error estimates. UCSF ChimeraX (v1.1) was used to visualize the three-dimensional structures [37].

#### Assembly and disassembly of NPs

The monomeric populations of A3C8-GFP-H6 and EM1-GFP-H6 were assembled into protein Zn NPs by the addition of 0.22  $\mu\text{m}$ -filtered  $\text{ZnCl}_2$  at 1:3 molar ratio with the histidine residues in the H6 tag. Proteins were set to 0.2  $\text{mg mL}^{-1}$  and  $\text{ZnCl}_2$  was added at a final concentration of 8.5  $\text{mmol L}^{-1}$  for A3C8-GFP-H6 and 8.3  $\text{mmol L}^{-1}$  for EM1-GFP-H6, depending on their molecular weights. Controlled disassembly of protein Zn NPs was subsequently obtained by adding EDTA at the same molar concentration. The spontaneous protein NPs based on A3C8-GFP-H6 and EM1-GFP-H6 were disassembled by the addition of 0.5% *v/v* Triton X-100. Addition of 2% anionic SDS detergent also resulted in NP disassembly.

#### Physicochemical characterization

A3C8-GFP-H6 and EM1-GFP-H6, at their different oligomerization states, were analyzed by SDS-PAGE and western blot based on an anti-His (Santa Cruz Biotechnology, USA) to assess protein purity and integrity. The same molar amount was used for each protein. For characterization experiments, proteins were set to 0.2  $\text{mg mL}^{-1}$  and pH 8.0. Dynamic light scattering (DLS) was used to determine the volume size distribution of A3C8-GFP-H6 and EM1-GFP-H6 at their different oligomerization states. Measurements were carried out in triplicate at 25°C in a Zetasizer Advanced Pro Blue (Malvern Instruments Limited, Malvern, Worcestershire, UK) at 633 nm.

Fluorescence emission spectra of all protein variants were analyzed with a Varian Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies). For tryptophan fluorescence, the excitation wavelength was set at 295 nm and the excitation and emission slits at 5 nm. Measurements were carried out by triplicate at 20°C. For GFP fluorescence, the excitation wavelength was set at 488 nm and the emission peak was achieved at 511 nm. Measurements were carried out by triplicate at 20 and 37°C and the excitation and emission slits were set at 5 nm.

Circular dichroism (CD) measurements were made with a JASCO J-715 spectropolarimeter (JASCO, USA) using a 0.2-mm path length quartz cell. Two spectra were acquired for each protein species. Each spectrum was an average of ten scans. The scan speed was set at 50  $\text{nm min}^{-1}$  with a 1 s response time. Measurements were obtained as ellipticity in millidegrees (mdeg) in the 190–260 nm region. The secondary structure content of each protein format was analyzed using the Dicroweb platform [38].

#### Ultrastructural characterization

Ultrastructural morphology (size and shape) of A3C8-GFP-H6 and EM1-GFP-H6 NPs with and without  $\text{Zn}^{2+}$  was visualized with two rapid high-resolution electron microscopy techniques. Drops of 10  $\mu\text{L}$  of the four samples resuspended in its buffer were deposited in silicon wafers (Ted Pella Inc.) for 1 min, air

dried, and immediately observed without coating in a field emission scanning electron microscope (FESEM) Merlin (Zeiss, Oberkochen, Germany) operating at 1 kV and equipped with a high-resolution *in-lens* secondary electron detector. Representative images of general fields and nanostructure details were collected at three magnifications (180,000 $\times$ , 240,000 $\times$  and 500,000 $\times$ ). For negative staining, drops of 10  $\mu\text{L}$  of the same samples were deposited 1 min in 200 mesh copper grids coated with carbon, contrasted with uranyl acetate 2% (Polysciences Inc.) during 1 min, air dried and observed with a transmission electron microscope (TEM) JEM-1400 (Jeol Ltd.) operating at 120 kV and equipped with a Gatan Orius SC200 CCD camera (Gatan Inc.). Representative images of general fields and nanostructure details were collected at three magnifications (5000 $\times$ , 10,000 $\times$  and 50,000 $\times$ ).

#### Thermal stability

DLS (using the settings and equipment afore-mentioned) was used to evaluate the thermal stability of A3C8-GFP-H6 and EM1-GFP-H6 at their different oligomerization states. Temperature was initially set at 25°C and progressively increased up to 90°C. To evaluate the thermal behavior of A3C8-GFP-H6 and EM1-GFP-H6 at their different oligomerization states within a range from 25 to 90°C, center of spectral mass (CSM) was calculated as previously described [39] for analysis and comparisons. CSM is related with the relative exposure of the tryptophan (Trp) to the protein solvent. The maximum red-shift in the CSM of the Trp is compatible with a large solvent accessibility. On the contrary, the blue shift is related to a highly hydrophobic environment for Trp.

#### Functional characterization

CXCR4<sup>+</sup> human cervical adenocarcinoma cells (Hela, ATCC, CCL-2) were used for functional characterization experiments. HeLa cells were maintained in Mem-Alpha (Gibco) media supplemented with fetal bovine serum (Gibco) at 5%  $\text{CO}_2$  in a humidified atmosphere. HeLa cells were used to test the neutralization capacity of A3C8-GFP-H6 at its different oligomerization states. They were seeded in a 96-well plate at  $3.5 \times 10^3$  cells/well and the proteins were added after 24 h. A3C8-GFP-H6 was pre-incubated for 1 h with T22-mRTA-H6 (to allow potential binding and neutralization before exposure to cells), a recombinant modular protein containing the CXCR4-targeting peptide T22 and the active site of ricin toxin, at the three different molar ratios of A3C8:ricin (1:10, 1:1 and 10:1). T22-mRTA-H6 concentration was set at 50  $\text{nmol L}^{-1}$  in all cases. Protein mixtures were added at a final volume of 100  $\mu\text{L}$  for 48 h and the Cell-Titer Glo Luminescent Cell Viability Assay (Promega, USA) Protocol was followed. Viabilities of cells without protein incubation and cells treated only with 50  $\text{nmol L}^{-1}$  T22-mRTA-H6 were used as controls. A Victor3 (PerkinElmer, USA) microplate reader was used for the measurements. Experiments were performed in triplicate.

#### Statistical analysis

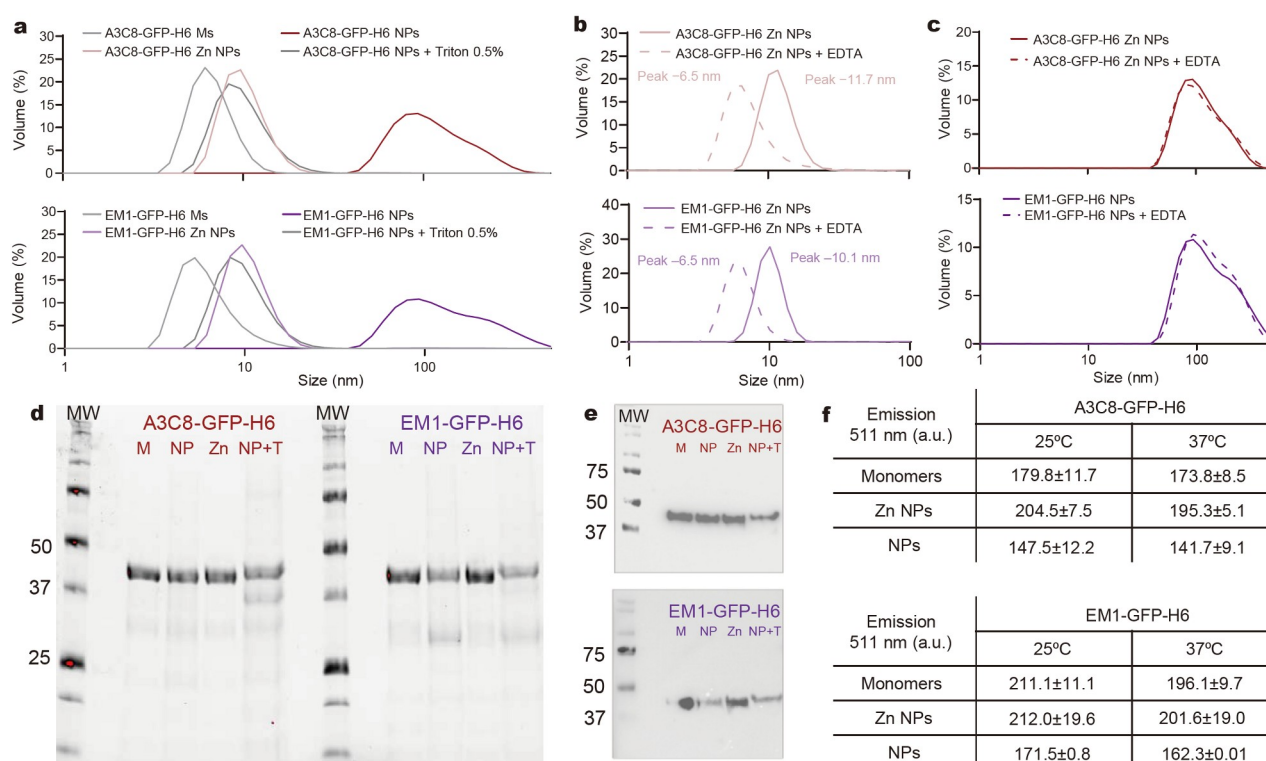
Data is represented as mean  $\pm$  standard error of the mean. For the neutralization assay data, a Shapiro-Wilk test was performed to ensure normality. Significant differences among means were identified through a One-Way ANOVA test. GraphPad Prism was used for statistical tests.

## RESULTS AND DISCUSSION

The IMAC purification of A3C8-GFP-H6 and EM1-GFP-H6 from cell extracts of the producing bacteria resulted, in both cases, in two distinguishable pure protein peaks, sequentially released from the columns because of a differential affinity of the proteins to the immobilized  $\text{Ni}^{2+}$ . Both proteins showed a similar distribution in the chromatogram, with an initial peak estimated to represent 86%–87% of the total protein amount and a final peak, accounting for the remaining 13%–14% of the protein population. The DLS of these samples revealed, for both proteins, that the primarily eluted protein peak presumably corresponded to monomers (Ms) while the protein sample released at higher imidazole concentrations corresponded to larger materials (Fig. 1c). The monomer-containing samples showed a hydrodynamic size of around 6–7 nm by DLS (Fig. 1d, e), a dimension compatible to those estimated in a plain modeling of both proteins, in which a certain degree of flexibility is expected between the nanobody and the GFP modules (Fig. 1b). The larger materials, leaked at latter elution volumes, showed a mean size slightly above 100 nm (Fig. 1e), with a mode peak size around 70 nm (Fig. 1d). The fact that such large entities were eluted from IMAC at high concentrations of imidazole was indicative of a tighter attachment to the column than the monomeric versions (Fig. 1c). The polydispersity indexes of the protein in these slowly eluted fractions were relatively low (Fig. 1e), indicating that the materials are not mere protein

aggregates but that they have some extent of regular organization at the nanoscale. Importantly, only negligible traces of nucleic acids were detected associated to the protein in each of these samples according to the  $A_{280}/A_{260}$  ratio [40] (Fig. 1e), whose abundance might have accounted for the formation of nanoscale protein-DNA complexes through cationic protein stretches. Both models, namely A3C8-GFP-H6 and EM1-GFP-H6, showed a similar profile when determining all these parameters.

As expected for a H6-tagged protein [41,42], the addition of equimolar amounts of  $\text{Zn}^{2+}$  to the monomeric protein versions promoted a shift in the DLS plots from 6.2–6.7 to 10.2–10.5 nm, indicative of protein assembling as NPs (Zn NPs, Fig. 2a). Also, as expected, such size increase was reversed by the addition of EDTA, since the metal chelation is able to disassemble multimeric protein materials organized through cross molecular interactions with metals (Fig. 2b) [1]. Contrarily, the pre-formed protein NPs peaking at 70 nm were not disassembled by EDTA or by imidazole (Fig. 2c and data not shown), even at concentrations higher than those used to disassemble the small NPs (Fig. 2b and data not shown). This finding was indicative that forces other than divalent cation coordination were supporting the organization of such a more complex material. For other types of related protein-only NPs, hydrogen bonds, van der Waals and especially electrostatic interactions were predicted to act, promoting and maintaining protein-protein contacts [23], in



**Figure 2** Physicochemical characterization. (a) Volume size distributions (nm) of A3C8-GFP-H6 and EM1-GFP-H6 in four different states (monomer, in grey; NP, in dark red/purple; Zn-mediated NP, in mild red/purple; and detergent-disassembled NP, in dark grey). Color legends are maintained hereafter. (b) Volume size distributions (nm) of A3C8-GFP-H6 and EM1-GFP-H6 Zn NPs upon EDTA-mediated disassembly (dashed line). Values depicted indicate the mode size of each peak. (c) Volume size distributions (nm) of A3C8-GFP-H6 and EM1-GFP-H6 NPs upon addition of EDTA (dashed line). Imidazole ( $400 \text{ mmol L}^{-1}$ ) was also unable to disassemble these materials (not shown). (d) Protein gel electrophoresis image of each protein state. Numbers at left indicate the molecular weight (MW) bands of the marker in kDa. (e) His-tag immunodetection of bands shown in (c). (f) Intrinsic GFP-emission (a.u.) at 511 nm upon fluorophore excitation at 488 nm, for both proteins, at 25°C (assembly temperature) and 37°C (working temperature). Proteins were adjusted to  $0.2 \text{ mg mL}^{-1}$  prior to measurements in (a, b, c and f).

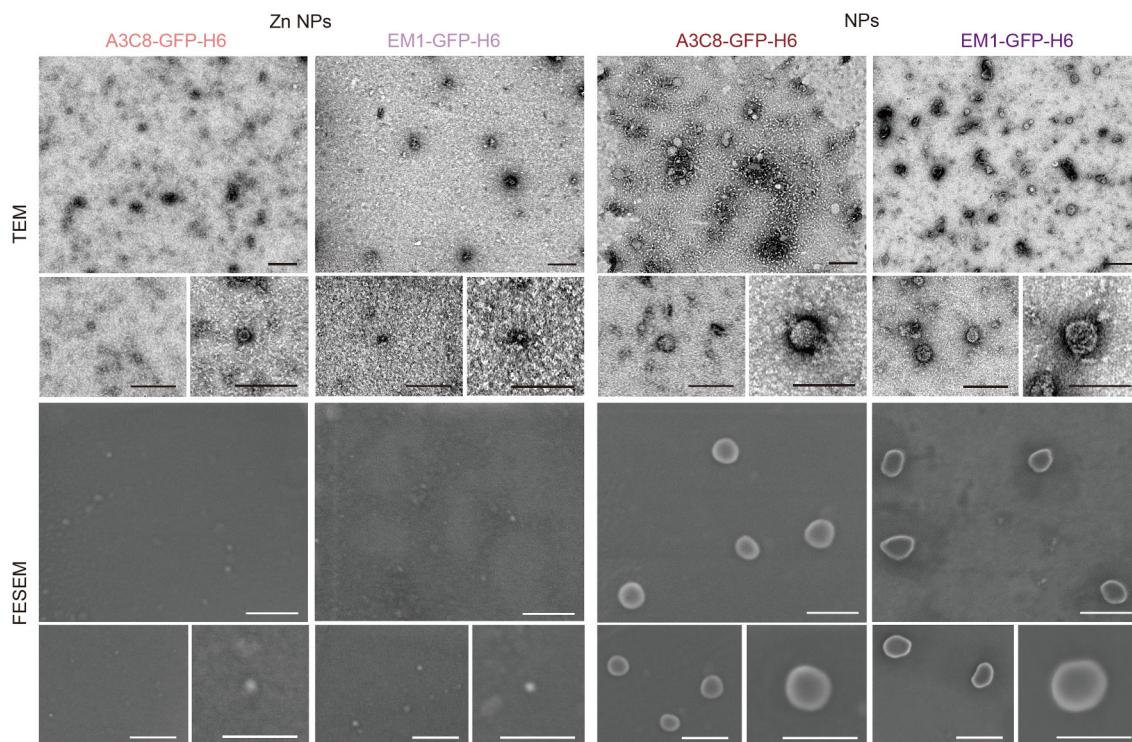
a similar way in which monomers are kept together in viral capsids [43–45]. In this context, the addition of 0.5% Triton X-100 reduced the size of the materials from more than 100 nm (mean) to 10 nm (Fig. 2a), a value very similar to that exhibited by the small Zn NPs. Such disassembling process appeared to be not complete, since forms with a size compatible to that expected for monomeric versions were not observed in presence of the detergent (Fig. 2a). Therefore, NPs and Zn NPs were clearly distinguishable materials regarding the interactions that regulate oligomerization.

In all these samples and under the tested assembling and disassembling conditions, the proteins remained proteolytically stable (Fig. 2d, e), with only symptoms of minor and partial proteolysis in the NP versions of both fusion proteins. The GFP fluorescence emission was similar in all the constructs, and slightly lower in the case of both NP versions (Fig. 2f). The Zn-mediated oligomerization did not disturb the GFP emission and still, Zn NPs tended to be more fluorescent than the Ms versions (Fig. 2f).

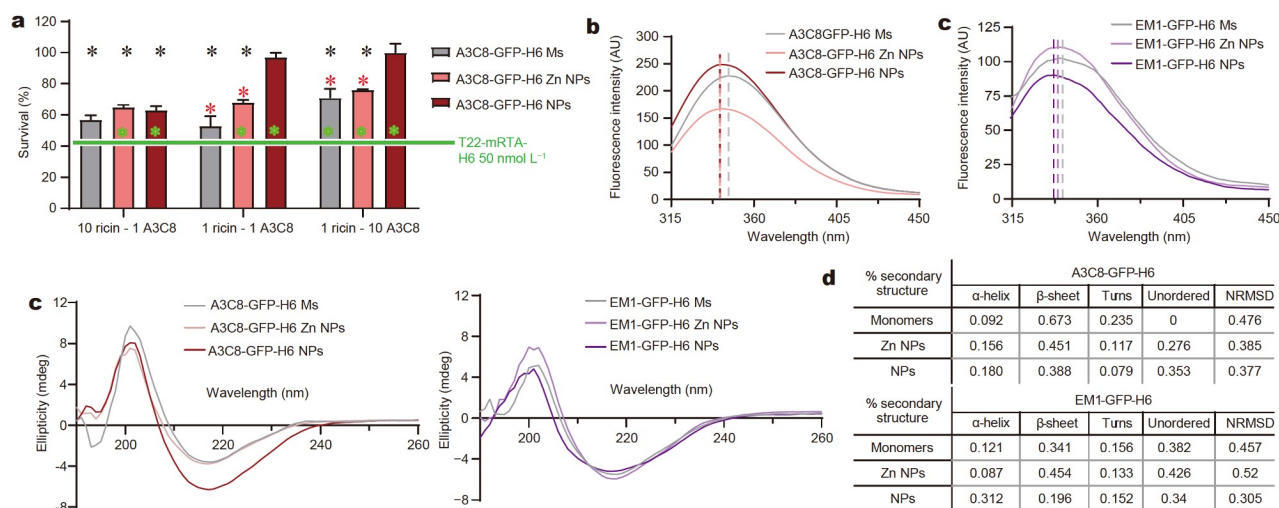
To comparatively evaluate the ultrastructural morphometry of NPs and Zn NPs, both type of materials made of the two alternative constructs were examined by TEM and FESEM (Fig. 3). These high-resolution imaging techniques showed clear size and shape differences between NPs and Zn NPs. When exploring different magnifications and field broadness, we observed structured materials in both cases, with sizes fully compatible to those determined by DLS. The smaller Zn NPs appeared as more regular, dense structures than the larger NPs, which were still regular but showed important deformability. Apparently, protein building blocks in NPs might be organized into toroidal NPs, with empty cores, while the modular proteins

clustered by Zn showed a rather spherical organization. The higher amount of bulk material in NPs should result from a higher number of building blocks arranged in each individual item. If so, we should expect a more intense reactivity of the nanobody moiety provided it is solvent-exposed, when compared with Zn NPs or with Ms. Since the nanobody A3C8, as a modular GFP-containing protein, binds the plant toxin ricin [33], highly loaded A3C8 NPs should inactivate more efficiently this toxin, that has a very strong biological activity over mammalian cells [34,46,47]. To assess this hypothesis, the different versions of A3C6-GFP-H6 generated here were incubated with a biologically active recombinant ricin that had been developed for cancer therapies [34]. When incubating this ricin-based recombinant protein with equimolar or ten times molar amount of the nanostructured nanobody (A3C8 NPs), its toxicity was majorly abolished. This event occurred at a lesser extent when using Zn NPs and even less when using Ms (Fig. 4a). This observation was in agreement with the afore-mentioned multivalency hypothesis, and it demonstrated that at least part of the nanobody moieties are available for interaction in NPs as well as in Zn NPs. In addition, these data strongly suggested that higher multivalency in NPs is more efficient in neutralizing the toxin than the smaller Zn NPs and the Ms versions, what would probably occur because of the clustering of a higher number of toxin ligands in NPs, showing a higher density when compared with Zn NPs or Ms.

At that point, it was demonstrated that Zn NPs (with regulatable assembling and disassembling) and NPs (spontaneously formed, non-regulatable) were materials with distinguishable biophysical properties, despite being formed by the same building block polypeptide. This was proved with two model



**Figure 3** Ultrastructural characterization. Representative high resolution electron microscopy images (top, TEM; bottom, FESEM) of A3C8-GFP-H6 and EM1-GFP-H6 Zn NPs and NPs showing round shape and stable size around 10 nm in diameter (Zn NPs) and around 80 nm in diameter (NPs). Scale bars: 200 nm.



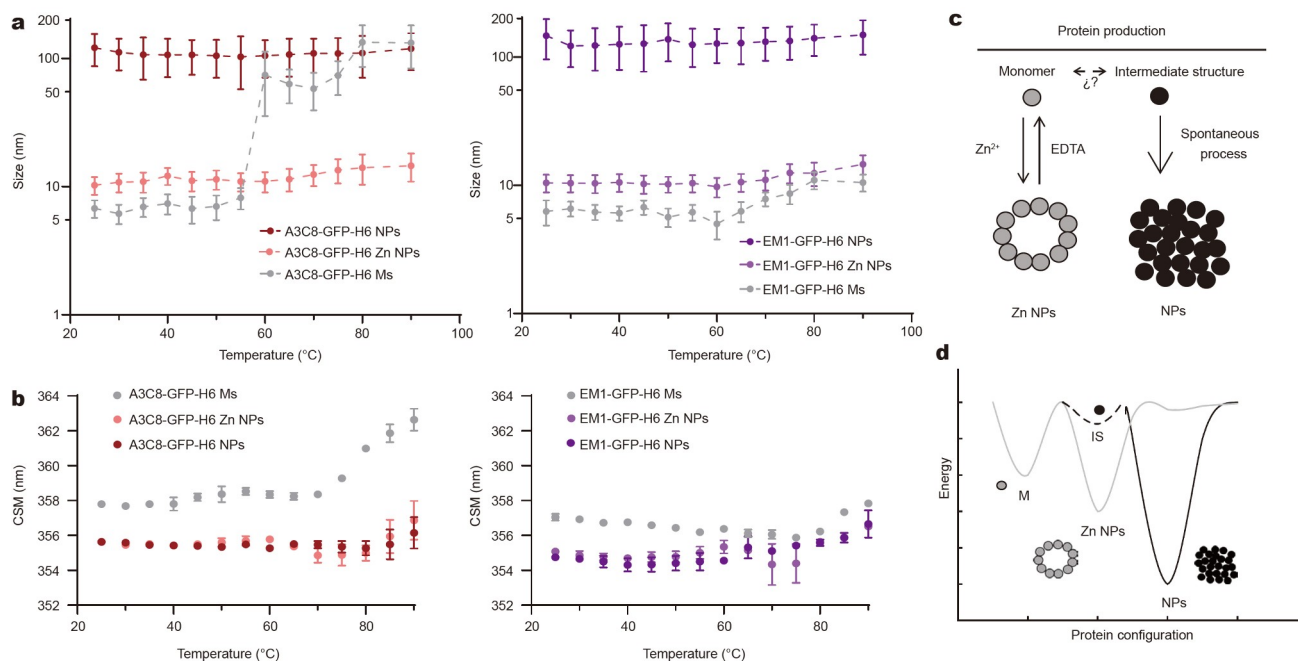
**Figure 4** Functional and structural comparison. (a) Neutralization assays of T22-mRTA-H6 with A3C8-GFP-H6 in their three different oligomerization states. Y-axis represents HeLa CXCR4<sup>+</sup> cell viability in percentage (being 100% the cell viability without protein incubation). T22-mRTA-H6 at 50 nmol L<sup>-1</sup> was pre-incubated for 1 h with A3C8-GFP-H6 at three different molar ratios (10:1, 1:1, 1:10) before being administered for 48 h to HeLa cells. Horizontal green line represents HeLa survival when exposed to T22-mRTA-H6 for 48 h at 50 nmol L<sup>-1</sup>. Significant differences are shown as \* ( $p < 0.05$ ), in black for differences with non-exposed cells (100%), in green for differences with cells exposed to T22-mRTA-H6 alone and in dark red for differences with cells exposed to (T22-mRTA-H6:A3C8-GFP-H6 NPs) at the same molar ratio. (b) Fluorescence emission spectra of A3C8-GFP-H6 (left) and EM1-GFP-H6 (right) at their different oligomerization states, upon excitation at 295 nm. Vertical dashed lines indicate the position of the peaks for each spectrum. (c) CD spectra of A3C8-GFP-H6 (left) and EM1-GFP-H6 (right) at their different oligomerization states. (d) Secondary structure data according to measurements in (c). NRMSD: normalized root mean square error. Proteins were adjusted to 0.2 mg mL<sup>-1</sup> prior to measurements in (b) and (c).

nanobody-containing proteins, but such a variability had been internally observed in the laboratory when developing other protein building blocks (not shown). Since these materials are formed by recombinant proteins bioproduced in bacteria, a conformational variability in these proteins could account for dissimilar building block populations, even having the same primary structure. In fact, the conformational variability of recombinant proteins is a well-recognized event in multiple models and production systems [48–55], but the impact of such diversity over self-assembling as protein materials has remained so far unexplored. In this context, intrinsic fluorescence analysis demonstrated that Tryptophan residues sense a high hydrophobic environment in both types of NPs as the fluorescence peaks move to lower wavelengths respect to Ms (Fig. 4b). This result confirmed the oligomerization states of both types of oligomers. On the other hand, the CD spectra of the whole set of A3C8-GFP-H6 and EM1-GFP-H6 variants (Ms, Zn NPs and NPs) suggested structural similarities between Ms and Zn NPs and also common differences between this pair and NPs (Fig. 4c). A higher content of  $\beta$ -sheet secondary structure was observed in Ms and the derived Zn NPs in comparison with NPs (Fig. 4d). These data, concomitantly with the different interaction types involved in the disassembling (Fig. 2a, b), strongly supported again the concept that NPs and Zn NPs derive from structurally different building blocks resulting from bacterial production, which, despite having the same amino acid sequence, fold into alternative conformations. In this same context, the mild proteolysis observed in these proteins (Fig. 2d) again discriminated the sensitivity between the protein forming Zn NPs and NPs, especially in the case of the EM1 fusion. Importantly, protein conformation is a well-known determinant of proteolytic susceptibility [56–59].

The DLS and CSM analyses of the proteins revealed that both NPs and Zn NPs were structurally more robust than plain Ms

(Fig. 5). While both types of oligomers remained stable during the whole tested temperature range, the monomeric versions of both proteins suffered a conformational conversion (more evident for A3C8-GFP-H6) between 60 and 80°C (Fig. 5a, b). Then, irrespective of the precise conformational states of the building blocks in the NPs, resulting in disparate morphologies and interactivities (Figs 3 and 4a), the oligomeric organization proves to be protective in front of thermal stress.

Such stabilizing property, which is linked to the formation of supramolecular complexes, prompts the development of transversal assembling platforms that, like those studied here, would allow the organization of functional proteins with therapeutic potential into nanoscale multimeric materials. In particular, the use of histidine-rich peptides as architectonic agents benefits from the simplicity of the engineering methods that is based on the addition of an H6 or related peptide to the protein of interest. Since H6 also allows simple, one-step purification of any protein by IMAC [2,3,60,61], the dual role of the peptide shows benefits from both the bioproduction and purification side, and the nanofabrication process itself, through self-assembling. In contrast to other more refined protein-specific engineering approaches [17–20,62–64], the protein self-assembling process promoted by H6 tails can be universally applied to any protein of clinical interest. The data shown here indicate that the recombinant production of distinct His-tagged nanobody-GFP fusions result in at least two distinguishable conformers of the protein that assemble into two categories of NPs. One of the conformers, enriched in  $\beta$ -sheet secondary structure, is found in a monomeric form upon IMAC protein purification. Its assembly is finely controlled by externally added Zn<sup>2+</sup> ions and its disassembly by the addition of a divalent cation chelator such as EDTA (Fig. 2). On the other hand, the alternative conformer with low  $\beta$ -sheet secondary structure spontaneously assembles during elution from IMAC columns into larger NPs,



**Figure 5** Conformational stability. (a) Volume size distributions (nm) of A3C8-GFP-H6 and EM1-GFP-H6 in their three different oligomerization states upon heating from 25 to 90°C. (b) CSM study of A3C8-GFP-H6 and EM1-GFP-H6 in their three different oligomerization states upon heating from 25 to 90°C. (c) Scheme of possible oligomerization states for His-tagged nanobody-GFP containing proteins. (d) Schematic illustration of the folding energy landscape reflecting the oligomerization states presented in (c). Proteins were adjusted to 0.2 mg mL<sup>-1</sup> prior to measurements in (a) and (b).

with regular toroid-like morphology with important extent of mechanical flexibility. Such NP version, over the Zn NP, shows an enhanced capability to neutralize a target ricin (Fig. 4a). This observation demonstrates the full functionality of the nanobodies and also the reaching of local ligand concentrations in NPs higher than in Zn NPs, which favors the neutralization of the toxin. Importantly, the ligand is active in both nanomaterials as well as in the monomeric form.

The origin of both populations (monomers and spontaneously formed NPs) is determined by the genetic cell background and culture conditions of the recombinant production in *E. coli* cells and in a differential folding pattern followed by each protein fraction, rendering different conformers. The structural analysis of the NPs and the unassembled monomer (Fig. 4) strongly suggests that the larger NPs are formed by a conformer that might come from an unstable monomeric intermediate in the oligomerization process, represented as an intermediate structure (IS) in the scheme (Fig. 5c, d). Also, the similarity in size when comparing Zn NPs and Triton X-100-disrupted NPs (Fig. 2), might be indicative that Zn NPs are similar or structurally close to an oligomeric intermediate in the fast generation of large NPs (Fig. 5d). Monomers can assemble into NPs when adding Zn. The shift between monomers and the spontaneously formed NPs may be a thermodynamic transition which is observed for A3C8-GFP-H6 but not for EM1-GFP-H6 at the selected range of temperatures (Fig. 5a).

The simplicity of the oligomerization system based on H6 and related humanized His-rich peptides [65] makes it straightforward applicable to the construction of protein-materials with full biocompatibility and multiple uses in clinics [1], including non-reactive materials [66], drug vehicles [24] or nanoscale drugs with build-in therapeutic activities [67]. However, the data presented here also reveal a biophysical variability in the

resulting materials linked to the conformational spectrum of the recombinant proteins, specific of the type of soft material that they generate. This concept, irrelevant in the use of more rigid building blocks for material design, might be highly relevant in the emerging field of protein-based materials for clinical applications, which is seen as a broad and promising technology for a diversity of therapeutic applications [6,17–21,62]. Being this fact an issue with regulatory implications, the awareness of such potentially disparate assembling makes possible a proper quality control of a clinically oriented product by a simple separation of relevant peaks during the chromatographic step of the production process.

Received 8 October 2021; accepted 30 November 2021; published online 17 January 2022

- López-Laguna H, Sánchez J, Unzueta U, *et al.* Divalent cations: A molecular glue for protein materials. *Trends Biochem Sci*, 2020, 45: 992–1003
- Riguero V, Clifford R, Dawley M, *et al.* Immobilized metal affinity chromatography optimization for poly-histidine tagged proteins. *J Chromatogr A*, 2020, 1629: 461505
- Loughra L, Walls D. Purification of polyhistidine-tagged proteins. *Methods Mol Biol*, 2017, 1485: 275–303
- Chen TY, Cheng WJ, Horng JC, *et al.* Artificial peptide-controlled protein release of Zn<sup>2+</sup>-triggered, self-assembled histidine-tagged protein microparticle. *Colloids Surfs B-Biointerfaces*, 2020, 187: 110644
- López-Laguna H, Parladé E, Álamo P, *et al.* *In vitro* fabrication of microscale secretory granules. *Adv Funct Mater*, 2021, 31: 2100914
- Sánchez JM, López-Laguna H, Álamo P, *et al.* Artificial inclusion bodies for clinical development. *Adv Sci*, 2020, 7: 1902420
- Povedano E, Ruiz-Valdepeñas Montiel V, Gamella M, *et al.* A novel zinc finger protein-based amperometric biosensor for miRNA determination. *Anal Bioanal Chem*, 2020, 412: 5031–5041
- Jacob RS, Das S, Ghosh S, *et al.* Amyloid formation of growth hormone in presence of zinc: Relevance to its storage in secretory granules. *Sci*

- Rep, 2016, 6: 23370
- 9 Cristovao JS, Santos R, Gomes CM. Metals and neuronal metal binding proteins implicated in Alzheimer's disease. *Oxid Med Cell Longev*, 2016, 2016: 9812178
- 10 Hane F, Leonenko Z. Effect of metals on kinetic pathways of amyloid- $\beta$  aggregation. *Biomolecules*, 2014, 4: 101–116
- 11 Faller P, Hureau C, Berthoumieu O. Role of metal ions in the self-assembly of the Alzheimer's amyloid- $\beta$  peptide. *Inorg Chem*, 2013, 52: 12193–12206
- 12 Wu H, Shen Y, Wang D, et al. Effect of the divalent cations zinc and calcium on the structure and mechanics of reconstituted vimentin intermediate filaments. *bioRxiv (CSH Laboratory)* 2019, <https://doi.org/10.1101/844167>
- 13 Maniaci B, Lipper CH, Anipindi DL, et al. Design of high-affinity metal-controlled protein dimers. *Biochemistry*, 2019, 58: 2199–2207
- 14 Samorodnitsky D, Nicholson EM. Differential effects of divalent cations on elk prion protein fibril formation and stability. *Prion*, 2018, 12: 63–71
- 15 Knight AS, Larsson J, Ren JM, et al. Control of amphiphile self-assembly via bioinspired metal ion coordination. *J Am Chem Soc*, 2018, 140: 1409–1414
- 16 Jehle F, Fratzl P, Harrington MJ. Metal-tunable self-assembly of hierarchical structure in mussel-inspired peptide films. *ACS Nano*, 2018, 12: 2160–2168
- 17 Korpi A, Anaya-Plaza E, Välimäki S, et al. Highly ordered protein cage assemblies: A toolkit for new materials. *WIREs Nanomed Nanobiot*, 2020, 12: e1578
- 18 Wang ST, Minevich B, Liu J, et al. Designed and biologically active protein lattices. *Nat Commun*, 2021, 12: 3702
- 19 Ulijn RV, Lampel A. Order/disorder in protein and peptide-based biomaterials. *Isr J Chem*, 2020, 60: 1129–1140
- 20 Hansen WA, Khare SD. Recent progress in designing protein-based supramolecular assemblies. *Curr Opin Struct Biol*, 2020, 63: 106–114
- 21 Xu X, Chen X, Li J. Natural protein bioinspired materials for regeneration of hard tissues. *J Mater Chem B*, 2020, 8: 2199–2215
- 22 Unzueta U, Ferrer-Miralles N, Cedano J, et al. Non-amyloidogenic peptide tags for the regulatable self-assembly of protein-only nanoparticles. *Biomaterials*, 2012, 33: 8714–8722
- 23 Céspedes MV, Unzueta U, Tatkiwicz W, et al. *In vivo* architectonic stability of fully *de Novo* designed protein-only nanoparticles. *ACS Nano*, 2014, 8: 4166–4176
- 24 Céspedes MV, Unzueta U, Aviñó A, et al. Selective depletion of metastatic stem cells as therapy for human colorectal cancer. *EMBO Mol Med*, 2018, 10: e8772
- 25 Falgàs A, Pallarès V, Unzueta U, et al. A CXCR4-targeted nanocarrier achieves highly selective tumor uptake in diffuse large B-cell lymphoma mouse models. *Haematologica*, 2020, 105: 741–753
- 26 Álamo P, Cedano J, Conchillo-Sole O, et al. Rational engineering of a human GFP-like protein scaffold for humanized targeted nanomedicines. *Acta Biomater*, 2021, 130: 211–222
- 27 Pallarès V, Unzueta U, Falgàs A, et al. An Auristatin nanoconjugate targeting CXCR4<sup>+</sup> leukemic cells blocks acute myeloid leukemia dissemination. *J Hematol Oncol*, 2020, 13: 36
- 28 Sánchez-García L, Sala R, Serna N, et al. A refined cocktail of proapoptotic nanoparticles boosts anti-tumor activity. *Acta Biomater*, 2020, 113: 584–596
- 29 Serna N, Álamo P, Ramesh P, et al. Nanostructured toxins for the selective destruction of drug-resistant human CXCR4<sup>+</sup> colorectal cancer stem cells. *J Control Release*, 2020, 320: 96–104
- 30 Serna N, Cano-Garrido O, Sánchez-García L, et al. Engineering protein venoms as self-assembling CXCR4-targeted cytotoxic nanoparticles. *Part Part Syst Charact*, 2020, 37: 2000040
- 31 Serna N, Sánchez-García L, Unzueta U, et al. Protein-based therapeutic killing for cancer therapies. *Trends Biotechnol*, 2018, 36: 318–335
- 32 Pesarrodona M, Crosas E, Cubarsi R, et al. Intrinsic functional and architectonic heterogeneity of tumor-targeted protein nanoparticles. *Nanoscale*, 2017, 9: 6427–6435
- 33 Sánchez-García L, Voltà-Durán E, Parladé E, et al. Self-assembled nanobodies as selectively targeted, nanostructured, and multivalent materials. *ACS Appl Mater Interfaces*, 2021, 13: 29406–29415
- 34 Díaz R, Pallarès V, Cano-Garrido O, et al. Selective CXCR4<sup>+</sup> cancer cell targeting and potent antineoplastic effect by a nanostructured version of recombinant ricin. *Small*, 2018, 14: 1800665
- 35 Song Y, DiMaio F, Wang RYR, et al. High-resolution comparative modeling with RosettaCM. *Structure*, 2013, 21: 1735–1742
- 36 Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res*, 2004, 32: W526–W531
- 37 Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera? A visualization system for exploratory research and analysis. *J Comput Chem*, 2004, 25: 1605–1612
- 38 Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, 2008, 89: 392–400
- 39 Sánchez JM, Sánchez-García L, Pesarrodona M, et al. Conformational conversion during controlled oligomerization into nonamyloidogenic protein nanoparticles. *Biomacromolecules*, 2018, 19: 3788–3797
- 40 Mamiatis T, Fritsch EF, Sambrook J, et al. *Molecular Cloning: A Laboratory Manual*. Now York: Cold Spring Harbor Laboratory Press, 1989
- 41 López-Laguna H, Unzueta U, Conchillo-Solé O, et al. Assembly of histidine-rich protein materials controlled through divalent cations. *Acta Biomater*, 2019, 83: 257–264
- 42 López-Laguna H, Voltà-Durán E, Parladé E, et al. Insights on the emerging biotechnology of histidine-rich peptides. *Biotechnol Adv*, 2022, 54: 107817
- 43 Siber A, Podgornik R. Role of electrostatic interactions in the assembly of empty spherical viral capsids. *Phys Rev E*, 2007, 76: 061906
- 44 Xian Y, Karki CB, Silva SM, et al. The roles of electrostatic interactions in capsid assembly mechanisms of giant viruses. *Int J Mol Sci*, 2019, 20: 1876
- 45 Meister S, Prunotto A, Dal Peraro M, et al. Salt enhances the thermostability of enteroviruses by stabilizing capsid protein interfaces. *J Virol*, 2020, 94: e02176-19
- 46 Moshiri M, Hamid F, Etemad L. Ricin toxicity: Clinical and molecular aspects. *Rep Biochem Mol Biol*, 2016, 4: 60–5
- 47 Bolognesi A, Bortolotti M, Maiello S, et al. Ribosome-inactivating proteins from plants: A historical overview. *Molecules*, 2016, 21: 1627
- 48 Miyawaki A, Tsien RY. Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. In: Thorner J, Emr SD, Abelson JN (eds.). *Methods in Enzymology*. San Diego: Academic Press, 2000. 472–500
- 49 Macedo B, Sant'Anna R, Navarro S, et al. Mammalian prion protein (PrP) forms conformationally different amyloid intracellular aggregates in bacteria. *Microb Cell Fact*, 2015, 14: 174
- 50 Ovchinnikova TV, Shenkarev ZO, Nadezhdin KD, et al. Recombinant expression, synthesis, purification, and solution structure of arenicin. *Biochem Biophys Res Commun*, 2007, 360: 156–162
- 51 Doglia SM, Ami D, Natalello A, et al. Fourier transform infrared spectroscopy analysis of the conformational quality of recombinant proteins within inclusion bodies. *Biotechnol J*, 2008, 3: 193–201
- 52 Martínez-Alonso M, González-Montalbán N, García-Fruitós E, et al. The functional quality of soluble recombinant polypeptides produced in *Escherichia coli* is defined by a wide conformational spectrum. *Appl Environ Microbiol*, 2008, 74: 7431–7433
- 53 García-Fruitós E, Martínez-Alonso M, González-Montalbán N, et al. Divergent genetic control of protein solubility and conformational quality in *Escherichia coli*. *J Mol Biol*, 2007, 374: 195–205
- 54 Menzella HG. Comparison of two codon optimization strategies to enhance recombinant protein production in *Escherichia coli*. *Microb Cell Fact*, 2011, 10: 15
- 55 Fan Y, Hua X, Zhang Y, et al. Cloning, expression and structural stability of a cold-adapted  $\beta$ -galactosidase from *Rahnella* sp. R3. *Protein Expression Purification*, 2015, 115: 158–164
- 56 Park C, Marqusee S. Probing the high energy states in proteins by proteolysis. *J Mol Biol*, 2004, 343: 1467–1476
- 57 Markus G. Protein substrate conformation and proteolysis. *Proc Natl Acad Sci USA*, 1965, 54: 253–258
- 58 Fontana A, Zamboni M, Polverino de Lauro P, et al. Probing the



- conformational state of apomyoglobin by limited proteolysis. *J Mol Biol*, 1997, 266: 223–230
- 59 Hubbard SJ, Eisenmenger F, Thornton JM. Modeling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci*, 1994, 3: 757–768
- 60 Li S, Yang K, Liu L, *et al.* Surface sieving coordinated IMAC material for purification of His-tagged proteins. *Anal Chim Acta*, 2018, 997: 9–15
- 61 Priestersbach A, Kubicek J, Schafer F, *et al.* Purification of His-tagged proteins. In: Lorsch JR (ed). *Methods in Enzymology*. San diego: Academic Press, 2015, 1–15
- 62 Hamley IW. Protein assemblies: Nature-inspired and designed nanostructures. *Biomacromolecules*, 2019, 20: 1829–1848
- 63 Wang Y, Katyal P, Montclare JK. Protein-engineered functional materials. *Adv Healthcare Mater*, 2019, 8: 1801374
- 64 Zou Q, Yan X. Amino acid coordinated self-assembly. *Chem Eur J*, 2018, 24: 755–761
- 65 López-Laguna H, Cubarsi R, Unzueta U, *et al.* Endosomal escape of protein nanoparticles engineered through humanized histidine-rich peptides. *Sci China Mater*, 2020, 63: 644–653
- 66 López-Laguna H, Sánchez-García L, Serna N, *et al.* Engineering protein nanoparticles out from components of the human microbiome. *Small*, 2020, 16: 2001885
- 67 Falgàs A, Pallarès V, Serna N, *et al.* Selective delivery of T22-PE24-H6 to CXCR4<sup>+</sup> diffuse large B-cell lymphoma cells leads to wide therapeutic index in a disseminated mouse model. *Theranostics*, 2020, 10: 5169–5180

**Acknowledgements** We are indebted to the Agencia Estatal de Investigación (AEI) and to the Fondo Europeo de Desarrollo Regional (FEDER) (PID2020-116174RB-I00, AEI/FEDER, UE). The AGAUR (2017SGR-229) and the CIBER-BBN (project NANOPROTHER) are granted to Villaverde A. Vázquez E received support from the AEI (PID2019-105416RB-I00/AEI/10.13039/501100011033 and CIBER-BBN (project NANOREMOTE). We also appreciate the support from the Javna Agencija za Raziskovalno dejavnost Republike Slovenije (ARRS/N4-0046 and ARRS/J4-9322) to de Marco A. Sánchez-García L was supported by a predoctoral fellowship from the AGAUR (2018FI\_B2\_00051), Voltà-Durán E by a predoctoral fellowship from Ministerio de Ciencia, Innovacion y Universidades (FPU18/04615), Unzueta U was supported by Miguel Servet contract (CP19/00028) from ISCIII co-funded by European Social Fund (ESF investing in your future) and by an ISCIII project (PI20/00400) co-funding FEDER (A way to make Europe). López-Laguna H was supported by a predoctoral fellowship from the AGAUR (2019FI\_B00352). Villaverde A received an ICREA ACADEMIA award. Protein production was partially performed by the ICTS “NAN-BIOSIS”, more specifically by the Protein Production Platform of CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN)/ IBB, at the UAB (<http://www.nanbiosis.es/portfolio/u1-protein-production-platform-ppp/>). Cell culture experiments were performed at SCAC facilities in the UAB. Molecular graphics and analyses were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. Electron microscopy studies were performed by the Servei de Microscòpia in the UAB.

**Funding note** Open Access Funding provided by Universitat Autònoma de Barcelona.

**Author contributions** Voltà-Durán E performed most of the experiments

and prepared the figures; Sánchez JM, López-Laguna H and Parladé E performed the structural analyses and prepared the figures; Sánchez-García L designed and produced the nanobody constructs under the supervision of de Marco A; Sánchez-Chardi A analyzed the data; Unzueta U supervised most of the experiments; Vázquez E and Villaverde A conceived and supervised the whole study; Villaverde A prepared the first draft of the manuscript. All authors contributed to the general discussion and approved the manuscript.

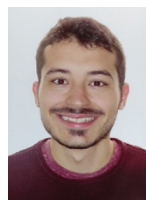
**Conflict of interest** The authors declare that they have no conflict of interest.

**Supplementary information** The original data presented in this paper can be found at: <https://ddd.uab.cat/record/248706>.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



**Eric Voltà-Durán** graduated in biotechnology in 2018 and pursued Master studies in advanced nanoscience and nanotechnology (2019). He is a young, early-stage researcher at the Nanobiotechnology group, a team led by Prof. Villaverde at the Autonomous University of Barcelona (UAB). He is working on his PhD studies, in which he focuses on the design and production of proteins with biomedical interest, specially in the context of targeted therapies.



**Antonio Villaverde** graduated in biological sciences in 1982 and got his PhD degree in 1985. He has been scientifically formed in Barcelona, Madrid, London, Lausanne and Braunschweig. Since 1987, he has been professor of microbiology at the Universitat Autònoma de Barcelona, Spain, where he got a full professorship in 2002. He leads the Nanobiotechnology group in this university and in the CIBER-BBN, focusing on the design of protein-based materials for biomedical applications. He founded the journal *Microbial Cell Factories* in 2002, being its Editor-in-Chief for 15 years.