BF-100584 (REVISED VERSION)

Conformational and functional variants of CD44-targeted protein nanoparticles bio-produced in bacteria

Mireia Pesarrodona[1,2,3], Yolanda Fernández[3,4], Laia Foradada[3,4], Alejandro Sánchez-Chardi[5], Oscar Conchillo[1], Ugutz Unzueta[1,2,3,¥], Zhikun Xu[1,2,3π], Mónica Roldán[5], Sandra Villegas [6], Neus Ferrer-Miralles[1,2,3], Simó Schwartz Jr [3,4], Ursula Rinas[7,8], Xavier Daura[1,9], Ibane Abasolo[3,4], Esther Vázquez[1,2,3,*], Antonio Villaverde[1,2,3,*]

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

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

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

[4] CIBBIM-Nanomedicine, Hospital Vall d'Hebron, Passeig de la Vall d'Hebron, 119-129, 08035, Barcelona, Spain

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

[6] Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

[7] Leibniz University of Hannover, Technical Chemistry & Life Science, Hannover, Germany

[8] Helmholtz Centre for Infection Research, Inhoffenstraße 7, Braunschweig, Germany

[9] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

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

[π] *Present address*: *Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China*

[*]Corresponding authors: Esther Vazquez, Esther.Vazquez@uab.es; Antonio Villaverde, Antoni.Villaverde@uab.es

Abstract

Biofabrication is gaining interest as a mean to produce nanostructured functional materials because of its operational versatility and full scalability. Materials based on proteins are especially appealing as protein structure and functionality are adaptable by genetic engineering. Furthermore, strategies and tools for protein production have been steadily developed and refined for more than 30 years. However, protein conformation and therefore activity might be sensitive to production conditions. We have explored here whether the downstream strategy influences structure and biological activities, *in vitro* and *in vivo*, of a self-assembling, CD44-targeted protein-only nanoparticle produced in *Escherichia coli*. This has been done by the comparative analysis of particles built of soluble protein species or protein versions obtained by *in vitro* protein extraction from inclusion bodies, through mild, non-denaturing procedures. These methods have been recently developed as a convenient alternative to the use of toxic chaotropic agents for protein resolubilization from protein aggregates. The results indicate substantial differences in the physicochemical properties as well as in the biological performance at the systems level of the resulting material, whose building blocks show to be sensitive to the particular protein source.

Keywords: nanoparticles; cell factories; recombinant proteins; bacteria; protein folding; drug delivery

Introduction

Protein materials are gaining interest in diverse arms of nanomedicine, especially in tissue engineering and drug delivery. This is because of the functional and structural versatility and tunability of proteins, the possibility to generate new functions through designing non-natural amino acid sequences, the unique merging between structure and function and the establishment of engineering principles to control protein self-assembling at the nanoscale [1,2]. Ribbons, tapes, fibres, hydrogels, films and particles result from the organized cross-interaction between discrete building blocks [3]. These fall into two main categories, namely peptides and multidomain proteins. While short peptides are synthesized chemically, full-length proteins and many long peptides need to be produced by biological fabrication in recombinant cells. Since the methodological principles supporting recombinant protein production were already established in the late 70's, we have accumulated experience, tools and know-how for almost 4 decades about the functioning and physiology of cell factories, especially *Escherichia coli*. The first insulin version approved for use was produced in *E. coli*, and an important fraction of the protein drugs currently in the market are also produced in this bacterial species [4]. Being the physiology of protein production more complex than formerly believed, genetics, genomics, proteomics, systems biotechnology and metabolic engineering are increasingly providing progresses and optimization of bacterial protein production, aimed to improve yield and functional quality of target products [5,6]. While some recombinant polypeptides are straightforward obtained as soluble versions, many other protein species, in particular when having a non-bacterial origin or upon extensive engineering, result in insoluble aggregates named inclusion bodies (IBs) [7]. Although in the last decade the biological properties of IBs have been deeply dissected at molecular and physiological levels[8], their formation is, in most cases, unavoidable [9]. For use as soluble species, for instance as drugs, target proteins that are produced as IBs need to be denatured and refolded *in vitro* by the application and progressive removal of solubilizing agents such as urea or guanidinium[10,11]. Being this approach convenient regarding some aspects (as these aggregates are a cost-effective and relatively pure source of protein), protein refolding from IBs involves complex protocols that need tedious development, the use of toxic chaotropic agents and a case-per-case adaptation.

More recently, and based on the realization that bacterial IBs contain significant fractions of properly folded protein species [12], non-denaturing protocols for direct protein extraction have been developed [13]. They are based on mild treatment with detergents or organic solvents that promote the release of soluble or quasi-soluble

protein. Such releasable protein fraction is probably attached to core amyloid fibres that form the IB skeleton by specific stereospecific interactions [14,15]. Being highly promising in the case of polypeptides intended for conventional uses (for instance enzymes or hormones), these methods have been never tested in the fabrication of more complex protein structures such as smart protein nanoparticles, resulting from protein self-assembling and intended, for instance, as vehicles for targeted drug delivery upon systemic administration. We have here explored the non-denaturing extraction of two tumor-homing protein nanoparticles (A5G27-GFP-H6 and FNI/II/V-GFP-H6) in which CD44-targeted peptides are displayed on GFP-based, self-assembling building blocks [16]. As these subunits are partially produced as soluble species and also as IBs, it has been possible to comparatively explore nanoparticle architecture, cell internalization and *in vivo* whole-body biodistribution when the oligomers resulted from either IB-extracted or soluble building blocks. Our data demonstrate that the choice of the downstream procedure is not irrelevant regarding the final performance of the materials but that instead compromises their functional quality and biological effectiveness at both cell and systems levels.

Material and methods

Protein design and biological production

The chimeric proteins A5G27-GFP-H6 and FNI/II/V-GFP-H6 contain the CD44-binding peptides A5G27 (fragment 2975-2987 from laminin α5 chain protein, RLVSYNGIIFFLK) and FNI/II/V (fragment 1923-1991 from fibronectin protein, WQPPRARITGYIIKYEKPGSPPREVVPRPRPGVTEATITGLEPGTEYTIYVIALKNN QKSEPLIGRKKT) respectively, displayed at the amino terminus of a His-tagged enhanced GFP (Figure 1A). These proteins act as building blocks of self-assembling protein nanoparticles, intended as carriers for targeted drug delivery, that are internalized by $CD44^+$ cells by receptor specific interactions and subsequent receptor-mediated endocytosis[16]. While the cationic nature of the N-terminal peptides (milder for A527) promotes protein self-assembling, in combination with the H6 tail, they also act as CD44 ligands with distinguishable efficiencies [16]. The encoding genes were harboured in the plasmid pET22b, that was transformed into *Escherichia coli* BL21 (F^- *omp ThsdS*$_B$(r_B^- m_B^-) *gal dcm*DE3, Novagen). Protein production was induced by the addition of 1 mM IPTG in bacterial cultures growing in LB medium plus ampicillin (100 µg/mL) at 16°C and under stirring at 250 rpm. Upon overnight culture, bacterial cells were centrifuged for 15 min (5,000 g at 4°C) for harvesting.The medium-scale production of recombinant proteins was partially performed by the ICTS “NANBIOSIS”, more specifically by the Protein Production Platform of CIBER in Bioengineering, Biomaterials & Nanomedicne (CIBER-BBN)/ IBB, at the UAB (SepBioEs, http://http://www.nanbiosis.es/unit/u1-protein-production-platform-ppp/).Under these production conditions, soluble and aggregated protein species were observed in both cases.

Protein purification and preliminary characterization

Pellets of bacterial cells were resuspended in Tris buffer (20 mM Tris, pH 8.0, 500 mM NaCl, 10 mM imidazole) in the presence of protease inhibitors (Complete EDTA-Free, Roche, Basel, Switzerland). Cells were disrupted at 1,200 psi in a French press (Thermo FA-078A) and the lysate was centrifuged at 15,000 g for 45 min. The insoluble fraction was kept apart for the purification of IBs and further protein extraction. The soluble fraction (named SN1) was filtered through 0.22 µm filter to be later submitted to His-tag affinity chromatography in HiTrap Chelating HP 1 mL columns (GE Healthcare, Piscataway, NJ, USA) with an ÄKTA purifier FPLC System (GE Healthcare). Elution

was achieved by a linear gradient of 20 mMTris, pH 8.0, 500 mM NaCl, and 500 mM imidazole.

For protein extraction from IBs a previously published protocol [17] was slightly adapted. In summary, the insoluble cell fraction obtained from lysate centrifugation was resuspended twice with MilliQ H_2O and centrifuged at 15.000 g for 30 min at 4 ºC; supernatants were named SN2 and SN3 respectively. The pellet was resuspended in Tris buffer (40 mM Tris HCl with 0.2 % N-lauroyl sarcosine, pH 8.0) with a ratio of 40 mL buffer/g pellet in the presence of protease inhibitor (Complete EDTA-Free, Roche, Basel, Switzerland) and agitated for 24 h at room temperature. Sample was centrifuged at 15,000 g for 45 min at 4 ºC and supernatant with resolubilized protein was filtrated before being purified by His-tag affinity chromatography using HiTrap Chelating HP 1 mL columns (GE Healthcare) with an ÄKTA purifier FPLC System (GE Healthcare). Elution was achieved by a linear gradient of 20 mMTris HCl, pH 8.0, 500 mM NaCl, 500 mM imidazole and 0.1 % N-lauroyl sarcosine. For purification of FNI/II/V-GFP-H6, Tris-HCl buffer was substituted for 166 mM $NaHCO_3$ to avoid protein precipitation. Proteins were finally dialyzed overnight at 4 ºC against working buffer (166 mM $NaHCO_3$ at pH 7.4). For consistence, only proteins released from the ÄKTA purifier in the last major peak (those showing higher affinity for the matrix), were selected for further analyses in all cases (for both soluble and IB-extracted protein versions).

Eluted protein-containing fractions were analyzed by both SDS electrophoresis/Coomassie Brilliant Blue staining and Western blotting using an anti-His monoclonal antibody (Santa Cruz Biotechnology Inc., Heidelberg, Germany). Protein concentration was determined by Bradford's assay as described elsewhere[18].

Field Emission Scanning Electron Microscopy (FESEM) and Dynamic Light Scattering (DLS)

As a qualitative approach, purified proteins were diluted to 200 µg/mL working buffer, and 10 µL of each one were deposited onto silicon wafers (Ted Pella Inc., Redding, CA, USA) for 2 min, the excess of liquid was adsorbed with Whatman paper and air dried. Ultrastructural morphometry samples (size and shape) was then observed at a nearly native state without coating with a high resolution *in-lens* secondary electron detector through a field emission scanning electron microscope Zeiss Merlin (Oberkochen, Germany) operating at 0.8 kV. As a quantitative approach, volume size distribution of nanoparticles was determined by DLS at 633 nm (in a Zetasizer Nano ZS, Malvern Instruments Limited, Malvern, UK) at a protein concentration of 1 mg/mL. DLS

measurements to assess protein stability in physiologic medium were performed after 2 h incubation of A5G27-GFP-H6 nanoparticles at 1.5 µM (45 µg/mL) in OptiPro® medium. Zeta-potential of these materials were determined in the same device once diluted in working buffer at 0.6 mg/mL final protein concentration.

Size Exclusion Chromatography (SEC)

The molecular weight and Stroke's diameter distribution of protein nanoparticles were determined by SEC after injection of 100 µL-samples with working buffer at 0.5 mL/min, in a previously calibrated Superdex 200 10/300 GL (Tricorn) column (GE Healthcare).

Fluorescence Spectroscopy

Different fluorescence emissions were measured at 25 ºC in a Variant Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies), using a 1-cm pathlenght quartz cuvette. Green fluorescence emission was determined for 1-mg/ml protein samples in working buffer at a detection wavelength of 510 nm and an excitation wavelength of 450 nm. The variation in the tertiary structure of A5G27-GFP-H6 (containing a unique tryptophan residue) was analysed by intrinsic Trp-fluorescence for 0.5 mg/ml protein samples. Temperature-dependent Trp emission fluorescence was recorded at 330 nm every 0.5 ºC from 20 to 100 ºC using an excitation wavelength of 280 nm. Scans were recorded along the 290-400 nm range at given temperatures after 8 min of temperature equilibration, and 10 scans were averaged.

Circular Dichroism

Far-UV circular dichroism (CD) was measured at 25ºC in a Jasco J-715 spectropolarimeter to assess secondary structure differences. Proteins were analysed at 10 µM (0.3 µg/mL) in working buffer with 1 mm pathlength quartz cuvette. CD spectra were obtained over the wavelength range of 200-250 nm at a scan rate of 50 nm/min, a response of 2 s, and a bandwidth of 1 nm. Ten scans were accumulated. To assess conformational differences, spectra were superimposed by applying a 1.5 multiplication factor.

Protein modelling

The A5G27-GFP-H6 monomer was modelled by homology using Modeler 9v13 [19] and the following templates: mouse laminin alpha2 chain lg4-5 domain pair (pdb entry 1okq, residues A2789 to A2802, 21.4 % identity, 71.4 % similarity) [20] and mouse laminin

alpha1 chain, domains lg4-5 (pdb entry 2jd4, residues A2905 to A2918, 28.6 % identity, 50 % similarity) [21] for A5G27 residues 1 to 14; residues 43 to 49 of the globular domain of *Gallus gallus* histone H5 (pdb entry 1hst [22] for residues 15 to 21 in the connecting segment (85 % identity); and pdb entry 1qyo for GFP [23] (residues 21 to 256, 98 % identity). The histidine tail was modelled by Modeller's automodel function (residues 257 to 263), with 500 models generated and sorted by their DOPE score [24]. The 20 models with best DOPE score were analyzed with the Electrostatic-Desolvation-Profile method [25], in order to predict putative protein-protein binding patches.

Total carbohydrate and lipid analysis

A phenol-sulphuric acid assay was performed to determine the total amount of carbohydrates present in protein samples. For that, 225 µg of protein were lyophilized and 150 µL of 18 M sulphuric acid were added and immediately, 30 µL of 5 % phenol. Samples were incubated at 90 ºC for 5 min, cooled to room temperature and optical density was measured at 490 nm. A glucose standard was prepared with a range from 0 to 400 ng.

The total amount of lipids in the samples was analyzed by the sulfo-phospho vanillin colorimetric method. In short, 1 mg of each sample was lyophilized and dissolved in 200 µL chloroform. After chloroform was evaporated, 100 µL of water were added. Samples were transferred to glass tubes and 2 mL of 18 M sulphuric acid were added. Then, samples were incubated for 10 min in a boiling water bath and subsequently cooled down in water bath at RT. After treatment with 5 mL phosphoric acid-vanillin reagent (85 % (v/v) phosphoric acid and 1.2 g/L vanillin (Sigma Aldrich, Steinheim, Germany)), the samples were warmed at 37 ºC for 15 min and cooled for 10 min in water bath at RT. Absorbance at 530 nm was measured. For the standard curve, a range of 10–100 µg triolein was used.

Cell line and cell manipulation

Human breast cancer cell line MDA-MB-231 was obtained from the ATCC (Rockville, MD, USA). MDA-MB-231 used for flow-cytometry and confocal imaging were maintained in RPMI 1640 medium supplemented with 10 % FBS and 6 mM GlutaMAX (Invitrogen, Carlsbad, CA, USA). MDA-MB-231 cells used for orthotopic mouse models were cultured in DMEM-F12 medium supplemented with 10 % heat-inactivated fetal bovine serum (Lonza, Verviers, Belgium), L-glutamine and antibiotic-antimycotic

(Invitrogen) at 37 ºC in a humidified atmosphere containing 5 % CO_2 and transfected as follows.

Flow cytometry

Cells were cultured on 24-well plates at $8 \cdot 10^4$ cells/well until reaching 70 % confluence. Media was removed and cells washed twice with PBS (Sigma-Aldrich, Steinheim, Germany). Then, cells were incubated with OptiPRO™ serum-free medium supplemented with L-glutamine with recombinant proteins dissolved at the convenient concentrations. Cells were incubated at 37 ºC and 5 % CO_2 in a humidified atmosphere for 24 h, and for 16 h, 6 h, 4 h, 2 h, 1 h and 30 min in the kinetics assay. Cells were then treated with 1 mg/mL trypsin for 15 min to detach cells and remove any attached protein from cell membrane [26]. This was followed by the addition of complete medium and centrifugation at 1,400 g for 5 min. Collected cells were resuspended in DPBS. Protein internalization was analyzed using a FACS-Canto system (Becton Dickinson, Franklin Lakes, NJ, USA), using a 15 W air-cooled argon-ion laser at 488 nm excitation. GFP fluorescence emission was measured with detector D (530/30 nm band pass filter). Results were usually normalised with the specific fluorescence for each protein (by fluorescence signal/(fluorescence signal/(mg·ml^{-1}) protein)). Although settings of flow cytometry and conventional fluorimeter are different and fluorescence units are not equivalent, the normalization results in data that are representative of protein amounts. Experiments were performed in duplicate.

Confocal analysis

Cells were seeded on Mat-Teck culture plate (Mat Teck Corp., Ashland, MA, USA) with RPMI 1640 medium at 200,000 cells/plate for 24 h. Media were removed and replaced by OptiPRO™ serum-free medium supplemented with L-glutamine and containing the recombinant proteins at 1.5 µM (44 µg/mL).After incubation for 4 h, cells were washed with PBS. Plasma membranes were then labelled with 2.5 µg/ml CellMask™ Deep Red (Molecular Probes, Eugene, OR, USA) and cell nuclei with 0.2 µg/mL Hoechst 33342 (Molecular Probes) for 5 min in darkness. Then, cells were washed with PBS and complete medium was added. Stained cells were examined using TCS-SP5 confocal laser scanning microscope (Leica Microsystems, Heidelberg, Germany) with a Plan Apo 63×/1.4 (oil HC × PL APO l blue) objective. Hoechst 33342 was excited with a blue diode (405 nm) and detected at the 415-460 nm range. GFP was excited with an Ar laser (488 nm) and detected at the 525-545 nm range. CellMask was excited with a HeNe laser (633 nm) and detected at the 650-775 nm range. Confocal images were

deconvolved using Huygens Essential deconvolution software version 4.4.0p6 (SVI, Leiden, The Netherlands). The confocal channels representing specific emission spectra were separated and deconvolved using the classical maximum likelihood estimation with the following parameters changed: quality change threshold = 0.01%, maximum iteration = 50, and signal-to-noise per channel values = 20, 14, and 15, respectively. The data sets were then imported into Imaris 3D visualization software, version 6.1.0 (Bitplane, Zürich, Switzerland) and the isosurface projections were calculated with the Surpass tool.

Analysis of protein stability

Protein stability was analyzed in triplicate at 115 µg/ml final concentration in reconstituted human plasma (Sigma-Aldrich, Ref: S225-5 ml), after 30 min, 1 h, 2 h, 4 h and 24 h of incubation at 37 ºC, under mild agitation. Protein green fluorescence was determined as described above. Sample at 24 h was centrifuged and fluorescence of the soluble fractions was re-analyzed to discard protein aggregation.

Orthotopic mammary fat pad mouse models

Female athymic nude mice (Harlan Interfauna Iberica, Barcelona, Spain) were kept in pathogen-free conditions and used at 6 weeks of age. Animal care was handled in accordance with the Guide for the Care and Use of Laboratory Animals of the Vall d'Hebron University Hospital Animal Facility and the experimental procedures were approved by the Animal Experimentation Ethical Committee at the institution. All *in vivo* experiments were performed by the ICTS “NANBIOSIS”, more specifically by the CIBER-BBN’s *In Vivo* Experimental Platform of the Functional Validation & Preclinical Research (FVPR) area (http://www.nanbiosis.es/unit/u20-in-vivo-experimental-platform/).

MDA-MB-231 cells (2.5×10^6) suspended in 200 µL of phosphate buffered saline (PBS) with Matrigel (1:1) (BD Bioscience, Bedford, MA, USA) were implanted into the right abdominal mammary fat pad (i.m.f.p.). Tumor growth was monitored twice a week by conventional caliper measurements ($D\times d^2/2$, where D is the major diameter and d the minor diameter). Once the tumors reached a median volume of ~200 mm^3, mice were randomized in the different treatment groups (n=3-4) according to their tumor volume, and a biodistribution study was performed.

***In vivo* whole-body biodistribution**

MDA-MB-231 i.m.f.p. tumor-bearing mice were treated intravenously with a single dose of 40 mg/kg of A5G27-GFP-H6rIB and A5G27-GFP-H6 sol in 166 mM $NaHCO_3$ pH 7.4 buffer. Then, mice were euthanized at 10 min, 30 min, 60 min and 5 h post administration. Blood was collected from each animal by cardiac puncture and processed for plasma fractionation. Tumors and the major organs such as liver, kidneys and lungs were dissected from mice, weighted and imaged by *ex vivo* fluorescence imaging (FLI). Plasma and tissues samples were harvested and stored at −80 °C.

Ex vivo FLI was performed using an IVIS Spectrum Imaging System (PerkinElmer Life Science, Boston, MA, USA), and images and measurements were acquired and analyzed using the Living Image 4.3 software (PerkinElmer). The light emitted from the fluorescent A5G27-GFP-H6 rIB and A5G27-GFP-H6 sol was detected, digitalized and electronically displayed as a pseudocolor overlay onto a gray scale animal image. Regions of interest (ROI) from displayed images were drawn manually around the fluorescent signals and quantified as Radiant Efficiency. The mean fluorescence intensities and corresponding standard errors of the mean (SEM) were determined for each experiment. All the analyses and graphs were performed using Prism 5 software (GraphPad, San Diego, CA, USA).

RESULTS

Protein production. A5G27-GFP-H6 and FNI/II/V-GFP-H6 were produced in *E. coli* as soluble species, but at higher proportions, as insoluble material as well (Figure 1 B) in form of fluorescent IBs (Figure 1 C). Soluble A5G27-GFP-H6 showed a molecular weight of about 30 kDa compatible with that expected for the full-length protein, while the aggregated protein also appeared in multimeric forms, as described for GFP and some GFP-derived proteins [27]. Soluble FNI/II/V-GFP-H6 partially occurred as a degradation band, that was fully absent in the insoluble protein fraction (Figure 1 B). Soluble and insoluble (IB) protein fractions were separated and processed as depicted in Figure 2 A. Soluble protein was extracted from IBs by adapted mild procedures, recently developed [13] to avoid the conventional, complex and product-specific denaturation-based refolding protocols that rely on the multi-step use of chaotropic agents[10,11]. Upon French press cell-disruption, extraction of GFP from GFP IBs maintains the native specific fluorescence [28], thus stressing the mildness of the procedure. This novel category of separation methods is based on the recognition of functional proteins occurring in bacterial IBs[29,30], that are architectonically combined with amyloidal networks [31-34]. Because of the releasable properties of the soluble protein fraction, IBs show great promise as cost-effective functional materials and drugs[35]. Straightforward soluble protein (sol) and IB-released protein (rIB) versions were both purified by His-tag affinity chromatography. While A5G27-GFP-H6 was always observed as a single band with an electrophoretic mobility compatible with that of the full length protein, soluble FNI/II/V-GFP-H6 appeared as two distinguishable species eluted in different fractions (Figure 2 B), that were identified by MALDI-TOF as the full length proteins and as a stable degradation fragment (FNI/II/V-GFP-H6 deg) of about 30 KDa (Figure 2 C), compatible with a GFP-H6-like form lacking most part of the amino terminal homing peptide (in contrast to the intact carboxy terminus as indicated by His-based successful purification). By fine MALDI-TOF analysis, some degradation fragments were also observed in the samples of soluble A5G27-GFP-H6. In both proteins, rIB versions uniquely contained full-length forms (Figure 2 C), which stresses a potential advantage of IB protein extraction as a choice downstream procedure.

Architectonic characterization of nanoparticles. The material resulting from alternative downstream methods was analysed regarding their ability to form nanoparticles, a feature that is not displayed by the parental GFP-H6 construct [36]. Interestingly, DLS determination revealed architectonic differences in the resulting materials depending on the taken approach (Figure 3 A). For both proteins, rIB forms

resulted in larger nanoparticles than when derived from soluble protein versions (29 nm *versus* 14 nm in the case of A5G27-GFP-H6, and 29 nm versus 18 nm in the case of FNI/II/V-GFP-H6). The formation of nanoparticles necessarily involves the N-terminal tail, as a control GFP-H6 is unable to form them and remain as unassembled material [36]. Monomers are not observed in the DLS screening of both A5G27-GFP-H6 and FNI/II/V-GFP-H6, indicating that the self-assembling is a very efficient event. Interestingly, FNI/II/V-GFP-H6 deg tended to aggregate as highly polydisperse materials (not shown). The two versions of A5G27-GFP-H6 nanoparticles, showing the highest internal variation between size of *sol* and *rIB* forms, were further explored to fully assess the architectonic differences suggested by DLS. Indeed, both FESEM imaging of the samples (Figure 3 B) and SEC (Figure 3 C) fully confirmed the larger size of rIB version of A5G27-GFP-H6 nanoparticles, besides the SEC provided data fully matching with that from DLS. Minor size variation in the analyses of A5G27-GFP-H6 sol when comparing DLS and SEC might be due to different ways to compute data and to the average nature of DLS measurement. Despite these differences, exposure of electrostatic charges was similar when comparing A5G27-GFP-H6 sol and A5G27-GFP-H6 rIB (Figure 3 D), suggesting only moderate changes in the spatial disposition of the building blocks and also the absence of charged contaminants remaining from the purification process. Despite the low values of Z potential, aggregation was not observed at exception of FNI/II/V-GFP-H6 deg. To explore potential variability in the protein conformation when assembled in alternative particle versions, we analysed the specific fluorescence of the assembled proteins compared with the parental monomeric GFP-H6. All protein versions were fluorescent (Figure 3 E), what ensured convenient tracking in further *in vitro* and *in vivo* experiments. The specific fluorescence of FNI/II/V-GFP-H6 deg was indistinguishable from that of the parental GFP, while the rest of proteins were less fluorescent. This observation indicated that the fusion of the ligands to the amino terminus of GFP resulted in an altered conformation of the chromophore or in an enhanced quenching, and that the truncated version of FNI/II/V-GFP-H6, that was successfully purified by His-tag affinity, was lacking part or the complete N-terminal ligand. Interestingly, the major decrease in the specific fluorescence was observed in both *rIB* versions (3-fold compared with parental GFP), suggesting that the expanded size in the resulting nanoparticles could be linked to a particular protein conformation or conformational set that promote a reduced fluorescence emission, specifically favoured upon protein extraction from IBs.

***In vitro* cell internalization of protein nanoparticles**. To evaluate how these nanoparticle versions perform in contact with target cells and in particular, to explore

potential differences among them, we exposed $CD44^+$ cells to the same amount of proteins for comparative purposes. Samples were analysed by flow cytometry after a harsh trypsin treatment specifically developed to remove externally associated protein[26]. As expected, the truncated FNI/II/V-GFP-H6 deg only showed background cell penetration levels, much lower than those observed for the assembled nanoparticles (Figure 4 A). In particular, A5G27-GFP-H6 rIB showed significantly higher cell penetrability than the rest of nanoparticles, what was compatible with confocal microscopy taking into account the lessened fluorescence emission of this nanoparticle version (Figure 4 B). The enhanced association of A5G27-GFP-H6 rIB with target cells was fully confirmed by further kinetic internalization analyses (Figure 5 A, B), that occurred in absence of cell toxicity (Figure 5 C). The unassembled GFP-H6, in contrast, remains extracellularly when cell cultures are exposed to this material [36].

In an attempt to investigate if the difference in size and cell interactivity between A5G27-GFP-H6 sol and A5G27-GFP-H6 rIB could be due to different composition in the nanoparticles (that is, traces of contaminants from the cell factory), we determined the lipid (18.5 µg/mg ± 0.7 *versus* 19.9 µg/mg ± 0.5) and carbohydrate (0.10 µg/mg ± 0.01 *versus* 0.11 µg/mg ± 0.03) content in both, showing no significant differences (p>0.24) (Figure 6 A). Also, the addition of sarkosyl (a potential contaminant of A5G27-GFP-H6 rIB) to A5G27-GFP-H6 sol just before incubation with cells, does not increase cell interactivity or penetrability of the construct (despite a slight, non-significant enhancement), remaining the fraction of cell-associated protein a 26 % of A5G27-GFP-H6 rIB (Figure 6 B). The homogeneity in the electrostatic charge among the variant materials measured by Zeta-potential analysis (Figure 3 D) also allows discarding the significant occurrence of anionic detergents used in the IB protein extraction. Therefore, we concluded that as suggested by morphological data (Figure 3 A, B) and fluorescence recording (Figure 3 E), the discrimination of biological behaviour between A5G27-GFP-H6 sol and A5G27-GFP-H6 rIB might be derived from conformational variability in the nanoparticles as derived from distinct separation methods downstream the biological fabrication. Importantly, the mild extraction of monomeric GFP from IBs using the same protocol than in the present study does not affect specific fluorescence [28]. Therefore, the differences in the specific fluorescence observed here between the two A5G27-GFP-H6 particle versions must be linked to the ability shown by this protein to oligomerize as nanoparticles (in which monomers probably encounter distinguishable molecular environments).

The hypothesis of conformational variability between A5G27-GFP-H6 sol and rIB nanoparticles was supported by secondary and tertiary structure analyses (Figure 7).

The far-UV CD spectra revealed changes in ellipticity at around 210 nm, revealing that β-sheet content is somewhat higher in rIB oligomers than in those obtained from the sol version (Figure 7A). In addition, the rIB spectrum had to be multiplied by a 1.5 factor in order to be superimposed to the sol spectrum, which reflects that rIB nanoparticles are less soluble than sol nanoparticles, and, in turn, concurs with a β-sheet-rich associating conformation. Differences on tertiary structure were determined by measuring tryptophan intrinsic fluorescence at different temperatures. Temperature-dependent fluorescence intensity at 330 nm presents different decreasing slope for each nanoparticle, with that for rIB showing a particular transition between 70 and 80 ºC (Figure 7B).In order to delve into this heat-induced different behaviour, Trp emission spectra were recorded at different temperatures (Figure 7C-D). Although intensity decreases with temperature in both cases (as already shown in Figure 7B), it is noteworthy that the maximum of the sol variant spectrum suffers a slight redshift (from 331 nm to 335 nm) upon increasing temperature, whereas the rIB variant behaves the opposite, blueshifting from 329 nm to 310 nm; in addition, the maximum of the spectra remains the same from 20 to 60 ºC and suddenly blueshifts between 60-80º C, in consonance with the mentioned particular transition in the denaturation plot followed by intensity at 330 nm (Figure 7B). Therefore, it can be concluded that whereas the sol variant partially unfolds or initiates an unfolding process (albeit not reaching the fully unfolded state, characterized by a maximum at 355 nm), the rIB variant packs such that Trp residues are buried within the nanoparticle. Moreover, the maxima of the rIB initial spectra (20ºC) is 2 nm blueshifted with respect to the one for the sol variant and this effect, although slight, is also observed at the temperature closest to physiologic conditions (at 40ºC the difference among both variants is 3 nm), suggesting that, *in vivo*, rIB nanoparticles may display a more compact conformation than those obtained from the soluble variant.

***In vivo* biodistribution of nanoparticles.** To investigate any potential variation in the whole-body biodistribution maps of A5G27-GFP-H6 derived nanoparticles, we administered these materials to $CD44^+$ breast MDA-MB-231 intramammary tumor-bearing mice, in which we have previously demonstrated selective targeting of the A5G27-GFP-H6 sol version[16]. Tumor targeting was similar when comparing A5G27-GFP-H6 sol and A5G27-GFP-H6 rIB nanoparticles (Figure 8 A), proving that the structural differences between the particles did not affect the specificity in the interaction with the cell surface CD44 receptor *in vivo*. However, accumulation of the protein material in secondary, non-target organs was clearly unequal depending on the

material type. In particular, A5G27-GFP-H6 rIB nanoparticles showed much less retention in kidney than their A5G27-GFP-H6 sol counterparts, probably associated to lower extent of renal filtration derived from the higher size of this particle isoform and a weakened disassembly during circulation. Accordingly, the larger A5G27-GFP-H6 rIB nanoparticles followed an enhanced liver accumulation in contrast to A5G27-GFP-H6 sol nanoparticle. The higher size of this particle isoform and probably an enhanced stability during circulation might account for such altered pattern. Even more, A5G27-GFP-H6 rIB accumulated in lung at higher extents than its soluble counterpart, being most likely indicative of initial aggregation upon administration and their first-pass retention in lung capillaries (Figure 8 A). Interestingly, aggregation of this protein had never been observed *in vitro* in storage or working buffers (Figure 3 A), but it is well known that the presence of proteins and high salt content in complex media may affect the composition, interactivity and properties of protein nanoparticles [37]. In this regard, when incubating A5G27-GFP-H6 rIB in a protein-containing medium such as Optipro™, protein clustering was indeed observed (Figure 8 B) through a shift of the DLS protein peak towards larger sizes and by the decrease of stability revealed by the higher polydispersion index. Such clustering might be indeed responsible for the slight accumulation observed in lung (Figure 8 A). When incubating A5G27-GFP-H6 rIB in human plasma followed by centrifugation, the GFP fluorescence signal was not reduced (Figure 8 C), indicating the soluble nature of any A5G27-GFP-H6 rIB cluster that might be formed in protein-containing media. The small values of Z potential (Figure 3 B) could contribute to the observed tendency to form supramolecular clusters in complex media, that still being soluble and functional, might be favoured by a higher particle size in the case of the IB version.

Discussion

The microbial synthesis of nanostructured materials is gaining interest *versus* the most conventional chemical production because of the flexibility, scalability and cost effectiveness of cell factories, and due to the inherent, environmentally friendly nature of biological fabrication [38-40]. Different types of materials, including proteins[3], metals [41] and polymers [42,43] are produced in microbial cells, among which many of them are intended for biomedical applications[44]. The adaptation and implementation of microorganisms alternative to gram negative bacteria is expanding the diversity of products resulting from biological fabrication, especially concerning difficult-to-express proteins[45-49]. Being *E. coli* together with its expression vectors an extremely versatile toolbox in bioproduction, the recent development and application of endotoxin-free strains of this species [50,51] paves the road for a reassessment of this cell factory in BioPharma. Among nanoscale microbial products, those based on proteins exhibit particular properties associated to the unique integration of structure and function, properties that can be tailored and re-designed by plain genetic engineering. In this regard, the emergence of principles to rationally engineer protein self-assembling [52-58] allows the production, in recombinant cells, of a spectrum of protein materials such as particles, fibres, films, ribbons and hydrogels [1].

Despite the enormous potential in the biofabrication of protein nanostructures and the vast experience in recombinant protein production and purification, how production/separation methods might influence the performance of smart and structurally complex protein materials (like tumor-homing drug vehicles) is, unfortunately, a neglected but critical issue. We have here demonstrated that the downstream processing of recombinant CD44-targeted protein nanoparticles influences material architecture (Figure 3), protein conformation within the nanoparticle (Figure 7), cell penetrability *in vitro* (Figure 4) and whole-body biodistribution upon *in vivo* systemic administration (Figure 8). Since many recombinant proteins used in pharmacology (including insulin) are recovered from bacterial IBs, the implementation of methods to recover soluble protein from those aggregates has been a critical issue in protein biotechnology. In contrast to conventional protocols based on the case per case exploration of denaturation and further optimal refolding conditions of IB proteins [10,11], the recently developed milder protocols for non-denaturing protein extraction are technically simpler [13]. Also, aimed to release already soluble or quasi-soluble protein species occurring in IBs, they are essentially universal and can be applied transversally irrespective of the IB protein species.

The protein nanoparticle explored here as model, namely A5G27-GFP-H6, is slightly larger when extracted from IBs than when purified as plain soluble species (Figure 3 A-C). This version shows enhanced cell penetrability into $CD44^+$ cells upon exposure (Figures 4, 5), although it appears to be slightly more sensitive to cell proteases than the soluble counterpart upon endosomal internalization, as inferred by the comparison of the respective uptake kinetics (Figure 5 B). Such differential performance is not influenced by potentially associated contaminants such as carbohydrates or lipids, as they have not been detected (Figure 6 A). Also, divergent properties appear as not linked to charged agents such as anionic detergents used in protein extraction, since the Z-potential of the material remains essentially unchanged irrespective of the purification procedure (Figure 3 D) and the addition of N-Lauroyl sarcosine to A5G27-GFP-H6 sol nanoparticles does not upgrade cell uptake to the levels achieved by A5G27-GFP-H6 rIB (Figure 6 B). In addition, the fact that A5G27-GFP-H6 rIB nanoparticles do not contain higher amounts of lipids (Figure 6 D) also allows discarding the formation of nanodiscs, well known lipid-protein structures stabilized by detergents and commonly observed upon purification of membrane proteins [59,60]. In this regard, all these data indicate that the properties distinguishing A5G27-GFP-H6 sol and A5G27-GFP-H6 rIB nanoparticles are intrinsic to the material and not determined by the used separation procedure. That A5G27-GFP-H6 sol and A5G27-GFP-H6 rIB materials are basically distinct can be inferred from their differential specific fluorescence (Figure 3E). Since it has been described that the fluorescence of an unassembled GFP is unchanged when produced in soluble form or when extracted from IBS by the same method used here [28], the assembling pattern of A5G27-GFP-H6 must be different when produced as soluble version of as IBs. The arrangement of the whole material rather than the conformation of individual building blocks would be then responsible for the differential architecture and biological function, as further supported by the differential behaviour upon thermal denaturation found between these oligomer variants (Figure 7 B-D). The GFP beta-barrel domain contains a predicted protein-protein interaction patch that overlaps with the GFP dimer interface, in a region that has several residues known to be involved in the modulation of the intensity of the fluorescence emission (Figure 9B) [61]. It is thus very likely that relative differences in the arrangement of monomers in the two versions of A5G27-GFP-H6 of nanostructures (sizing 14 nm and 30 nm) might result in different molecular environments for the GFP chromophore and the modulator residues, leading to the observed differences in green fluorescence, CD and intrinsic Trp florescence upon thermal denaturation (Figure 7). The differential architecture of the materials according to their source (soluble or IB-embedded), might be indicative of the nanoparticles, or at least some of their

precursors, being already assembled into IBs, what could be mechanistically allowed by the high porosity and hydration of these inclusions [62]. A different molecular environment when compared with the cytosol would favour or force alternative arrangements of the building blocks (Figure 9A).

In this line, the SEC analysis reveals a higher number of building blocks in A5G27-GFP-H6 rIB nanoparticles when comparing with *sol* versions (Figure 3 C), and the enhanced cell penetrability might be due to a higher number of CD44-targeted peptides exposed on the materials surface, without denying the external disposition of the ligands from *sol* version. Although displaying different size, the electrokinetical potential (Figure 3 D) remains constantly negative for both types of nanoparticles, suggesting that surface charge is positive and that the monomers are faced with protruded orientation (Figure 9A). On the other hand, while size determination in cellular uptake of nanoparticles requires an important variability range to be effective [63-65], milder differences in multivalence regarding receptor binding (that is, number of reactive building blocks), might have a more visible role in the receptor-mediated penetrability of protein nanoparticles [66]. Of course, particular structural and functional features affecting the display of the homing peptide might influence the whole nanoparticle performance, as some (enhanced size) but not all (enhanced cell penetrability) of the properties of A5G27-GFP-H6 rIB are replicated in the related FNI/II/V-GFP-H6 rIB construct (Figures 3, 4 and 5).

At the systems level, A5G27-GFP-H6 rIB and A5G27-GFP-H6 sol nanoparticles target the tumor tissue similarly, but their side accumulation in non-target organs is clearly different (Figure 8 A). Regarding renal filtration, the major route for excretion of small materials, larger A5G27-GFP-H6 rIB nanoparticles showed low retention in kidney compared to A5G27-GFP-H6 sol counterparts. Larger nanoparticles are able to escape glomerular filtration and hold a longer plasma half-life, but are also prone for the uptake by the organs of the reticuloendothelial system such as the liver and lung [67,68], as it happens here with the A5G27-GFP-H6 rIB version. In addition, the higher accumulation of this construct in lung may be also due to initial clustering of nanoparticles and their first-pass retention in lung capillaries after systemic administration. The highest prevalence of A5G27-GFP-H6 rIB in the liver might be also enhanced by enlarged exposure of protein-protein interaction-prone regions [69,70] (Figure 9 B, C) that at the same time, might favour protein clustering as supramolecular structures and retention in lung. In agreement, DLS of A5G27-GFP-H6 rIB in protein-containing physiological media (Figure 8 B) indicates particle clustering, a fact not observed in storage buffer (Figure 3 A). However, these larger structures show a high polydispersion index (Figure

8 B), indicative of instability, and they cannot be removed by centrifugation (Figure 8 C), indicative of solubility. Decreased renal filtration is advantageous in *in vivo* administration of protein vehicles as circulation time is expanded and the targeting opportunities of any payload drug are consequently enhanced [71].

Results obtained in the present study indicate that cell-targeted protein nanoparticles exhibit differential performance depending on whether the materials derive from the *E. coli* soluble cell fraction or they are extracted from the soluble/quasi soluble protein fraction within IBs. In this context, the impact of the organization of the building blocks in oligomers on the whole performance of complex protein materials is an emerging matter of concern that should be investigated more deeply when designing protein-containing biomaterials, especially for biomedical applications. Being probably irrelevant for most of conventional applications, protein self-assembling into functional materials with complex activities (such as cell-targeting) might be more sensitive to the conformational constraints imposed by the intracellular location upon biofabrication.

Acknowledgments: We are indebted to MINECO (BIO2013-41019-P), AGAUR (2014SGR-132) and CIBER de Bioingeniería, Biomateriales y Nanomedicina (project NANOPROTHER) to AV and Marató de TV3 foundation (TV32013-3930) to EV and to IA and FIS (PI15/00272) to EV for funding our research on protein-based therapeutics. Protein production has been partially performed by the ICTS “NANBIOSIS”, more specifically by the Protein Production Platform of CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN)/ IBB, at the UAB (http://www.nanbiosis.es/unit/u1-protein-production-platform-ppp/). We are also indebted to Fran Cortes from the Cell Culture and Cytometry Units of the Servei de Cultius Cel·lulars, Producció d'Anticossos I Citometria (SCAC), and to the Servei de Microscòpia, both at the UAB. All in vivo experiments were performed by the ICTS “NANBIOSIS”, more specifically by the CIBER-BBN’s In Vivo Experimental Platform of the Functional Validation & Preclinical Research (FVPR) area (http://www.nanbiosis.es/unit/u20-in-vivo-experimental-platform/). M.P. and Z.X acknowledge financial support from Universitat Autònoma de Barcelona and China Scholarship Council through pre-doctoral fellowships respectively. AV received an ICREA ACADEMIA award. U.U received a Sara Borrell postdoctoral fellowship from ISCIII.

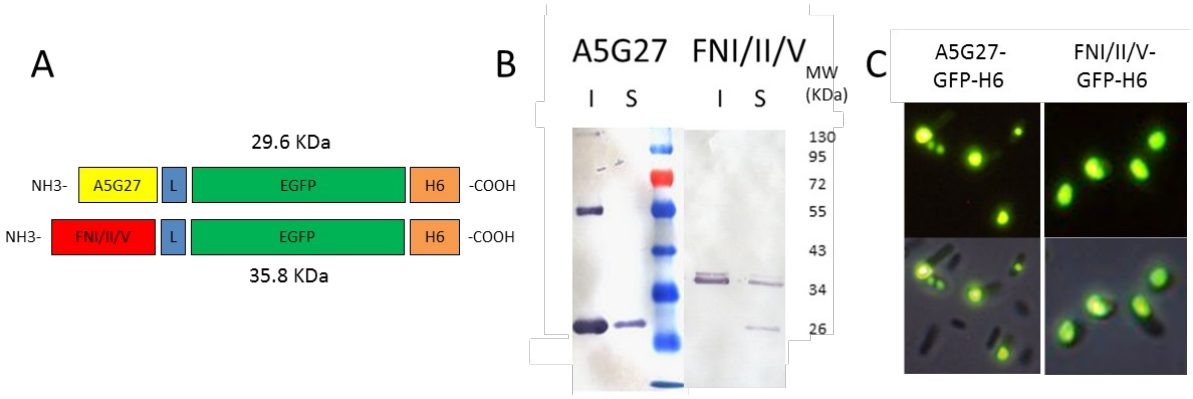

Figure 1.Protein production. A) Schematic representation of the gene fusions used in this study consisting, from N- to C-terminal, of a CD44 ligand (either A5G27 or FNI/II/V), a linker (L) sequence (GGSSRSSS), the EGFP protein and a His(6) tag. Lengths of boxes are approximate and do not represent accurate sizes of the genes. Theoretical molecular masses of the encoded proteins are indicated in kDa. Additional information about these constructs can be found elsewhere[16]. B) Western Blot analysis of soluble (S) and insoluble (I) crude protein extracts obtained from protein-producing *E. coli* BL21. C) Fluorescence microscopy images of *E. coli* BL21 protein producing cells 16 h after addition of IPTG) (top). Merging of fluorescence and phase contrast images of the same fields to show the localization of fluorescent aggregates (green particles) within cells (black areas; bottom).

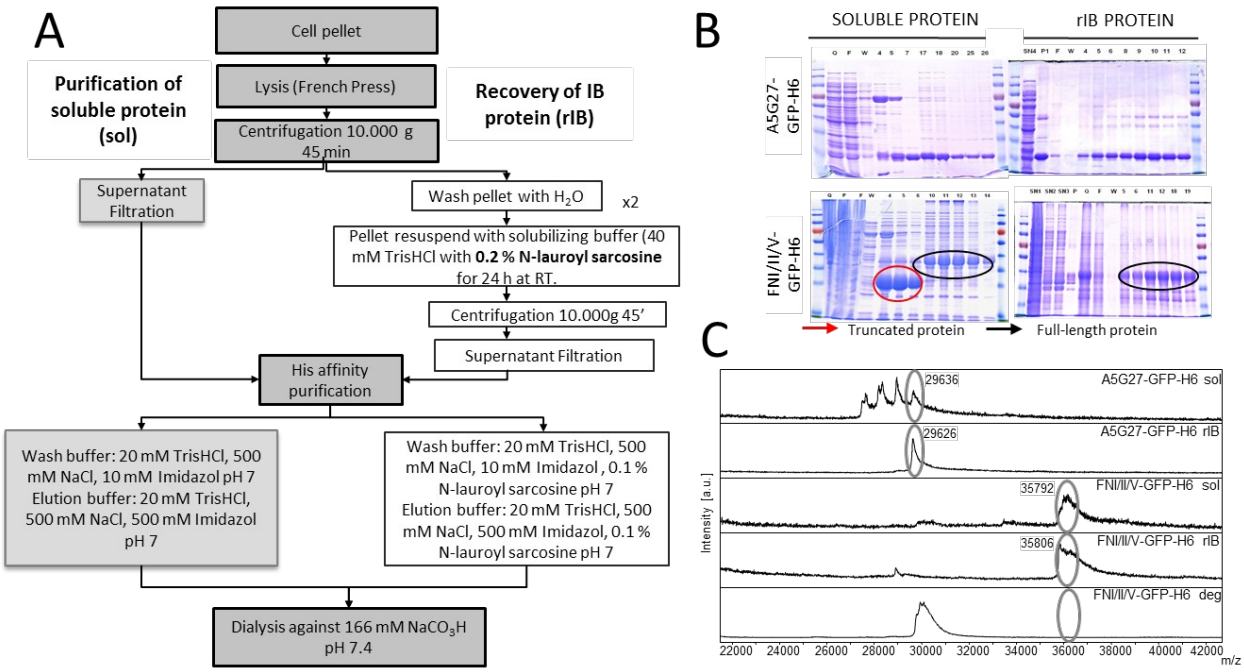

Figure 2. Protein purification strategy. A) Schematic representation of purification protocol for straightforward soluble protein (*sol*, light grey boxes) and for protein extracted from IBs (*rIB*, white boxes). Common steps are shown in dark grey boxes. B) Coomassie Brilliant Blue stained SDS-PAGE gel of purification fractions of A5G27-GFP-H6 and FNI/II/V-GFP-H6 obtained from both type of materials (soluble protein fraction and *rIB*). Fractions containing the pure protein are those from 20-26, 8-12, 10-14 and 17-20 for A5G27-GFP-H6 sol, A5G27-GFP-H6 rIB, FNI/II/V-GFP-H6 sol and FNI/II/V-GFP-H6 rIB respectively. Q: loaded supernatant, F: flowthrough, W: wash, 4-26: eluted fractions. C) Protein integrity analysis by molecular weight determination through MALDI-TOF. (*sol)* indicates soluble protein and (*rIB)*, protein extracted from IBs. (*deg)* refers to the soluble truncated version of FNI/II/V-GFP-H6.

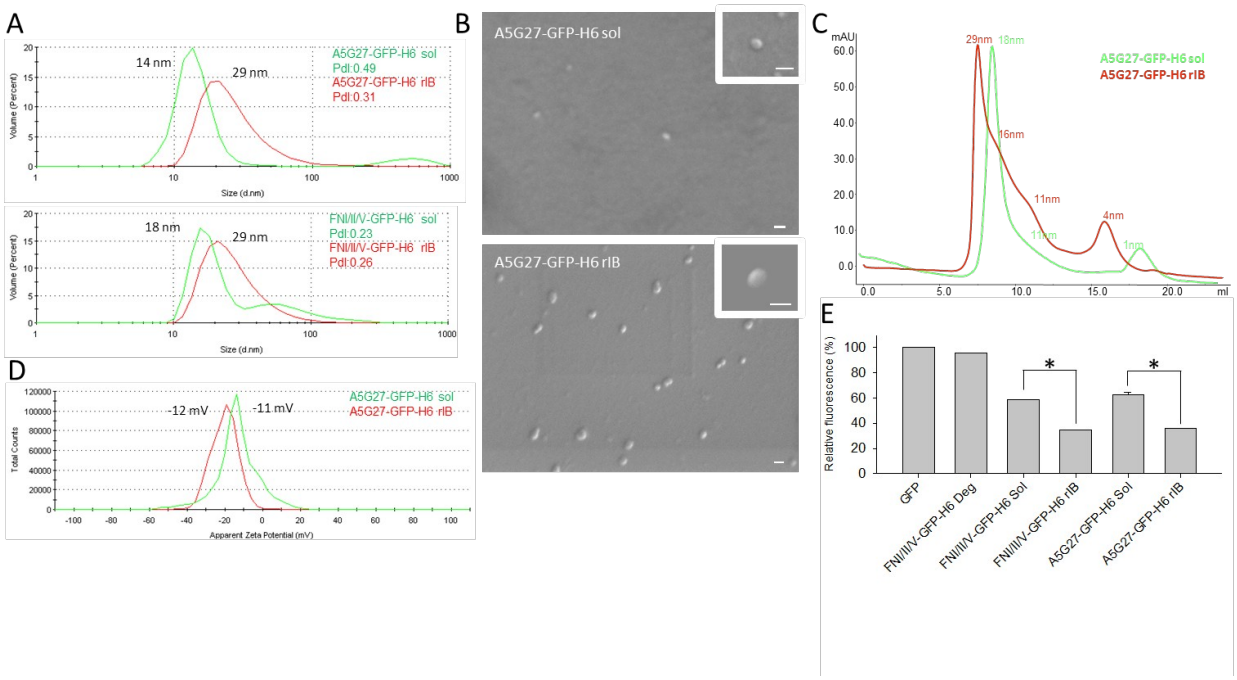

Figure 3.Qualitative and quantitative analyses for physicochemical and morphometric (size and shape) characterization of nanoparticles. A) DLS analysis of nanoparticle size. Peak sizes in each version are indicated above the corresponding plot. B) FESEM images of A5G27-GFP-H6 nanoparticles showing higher magnification in the insets. Bars indicate 50 nm. C) SEC of A5G27-GFP-H6 nanoparticles. Peak sizes are indicated. D) Z-potential of A5G27-GFP-H6 nanoparticles, indicating the values of the peaks. E) Specific fluorescence of different nanoparticle versions referred to that of the parental GFP-H6 protein. *, p<0.001.

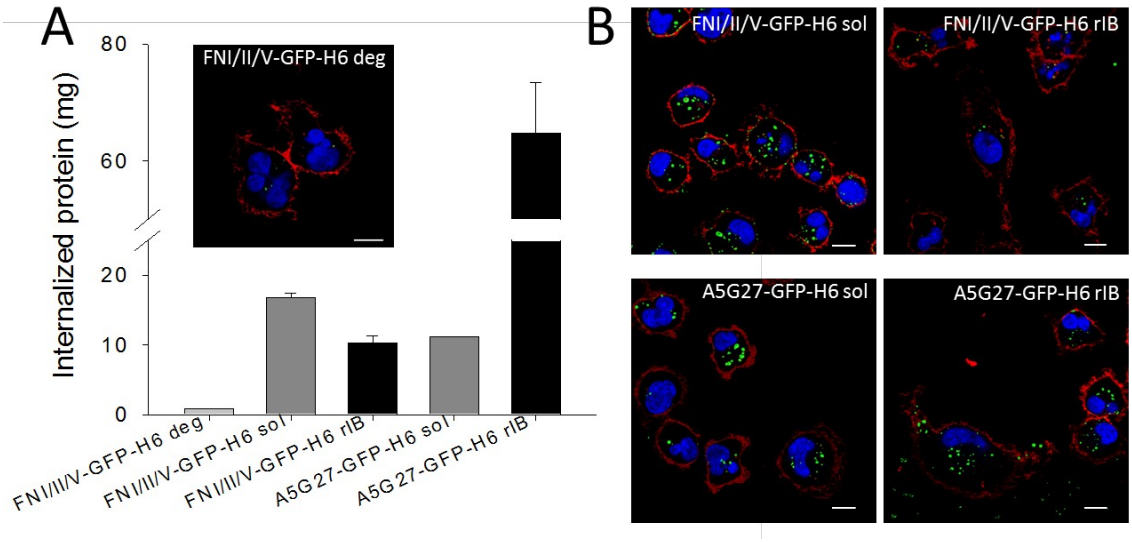
A
B
Internalized protein (mg)
80
60
20
10
0
FNI/II/V-GFP-H6 deg
FNI/II/V-GFP-H6 sol
FNI/II/V-GFP-H6 rIB
A5G27-GFP-H6 sol
A5G27-GFP-H6 rIB
FNI/II/V-GFP-H6 sol
FNI/II/V-GFP-H6 rIB
A5G27-GFP-H6 sol
A5G27-GFP-H6 rIB

Figure 4. Cell penetrability of nanoparticles in cultured cells. A) Amount of protein internalized in MDA-MB-231 cells upon 24 h of exposure, estimated from intracellular fluorescence and the specific fluorescence of each protein version. A harsh trypsin treatment was applied as described to eliminate externally attached protein material [26]. In the inset, confocal images of the amino truncated FNI/II/V-GFP-H6deg-exposed cells showing only background signal. Mean ± SEM are displayed. B) Confocal images of cultured MDA-MB-231 cells exposed to nanoparticles formed by full-length proteins. Blue signals represent DNA and red signals cell membranes. In all cases bars indicate 10 µm.

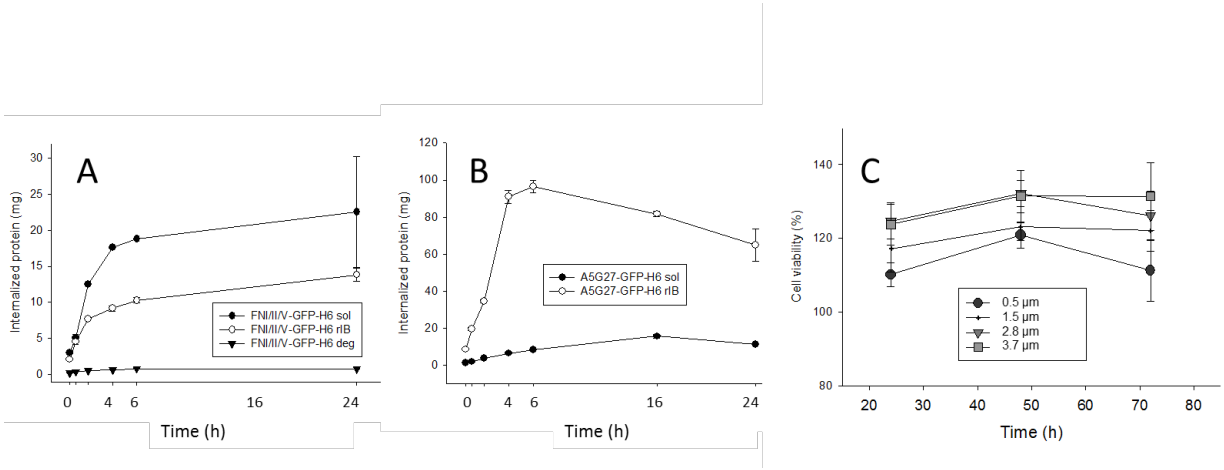

Figure 5. Kinetics of cell penetrability of nanoparticles in cultured cells. A) Flow cytometry analysis of MDA-MB-231 cells exposed to FNI/II/V-GFP-H6 derivatives, including the truncated version FNI/II/V-GFP-H6 deg. B) Flow cytometry analysis of MDA-MB-231 cells exposed to A5G27-GFP-H6 derivatives. C) Cell viability at different times of exposure to increasing concentrations of A5G27-GFP-H6 rIB. All data are displayed as mean ± SEM.

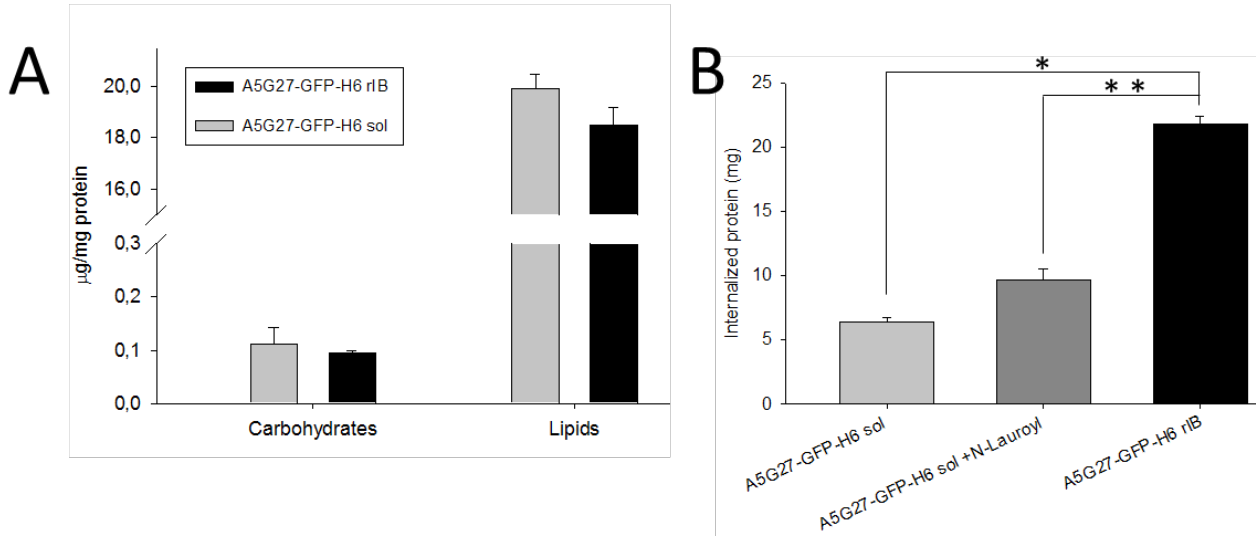

Figure 6. Compositional analyses of A5G27-GFP-H6 nanoparticles. A) Content of carbohydrates and lipids in purified nanoparticles. B) Effect of N-Lauroyl sarcosine, added just before incubation with cells, in the cell penetrability of A5G27-GFP-H6 sol nanoparticles. All data are displayed as mean ± SEM. *, $p<0.001$; **, $p<0.002$.

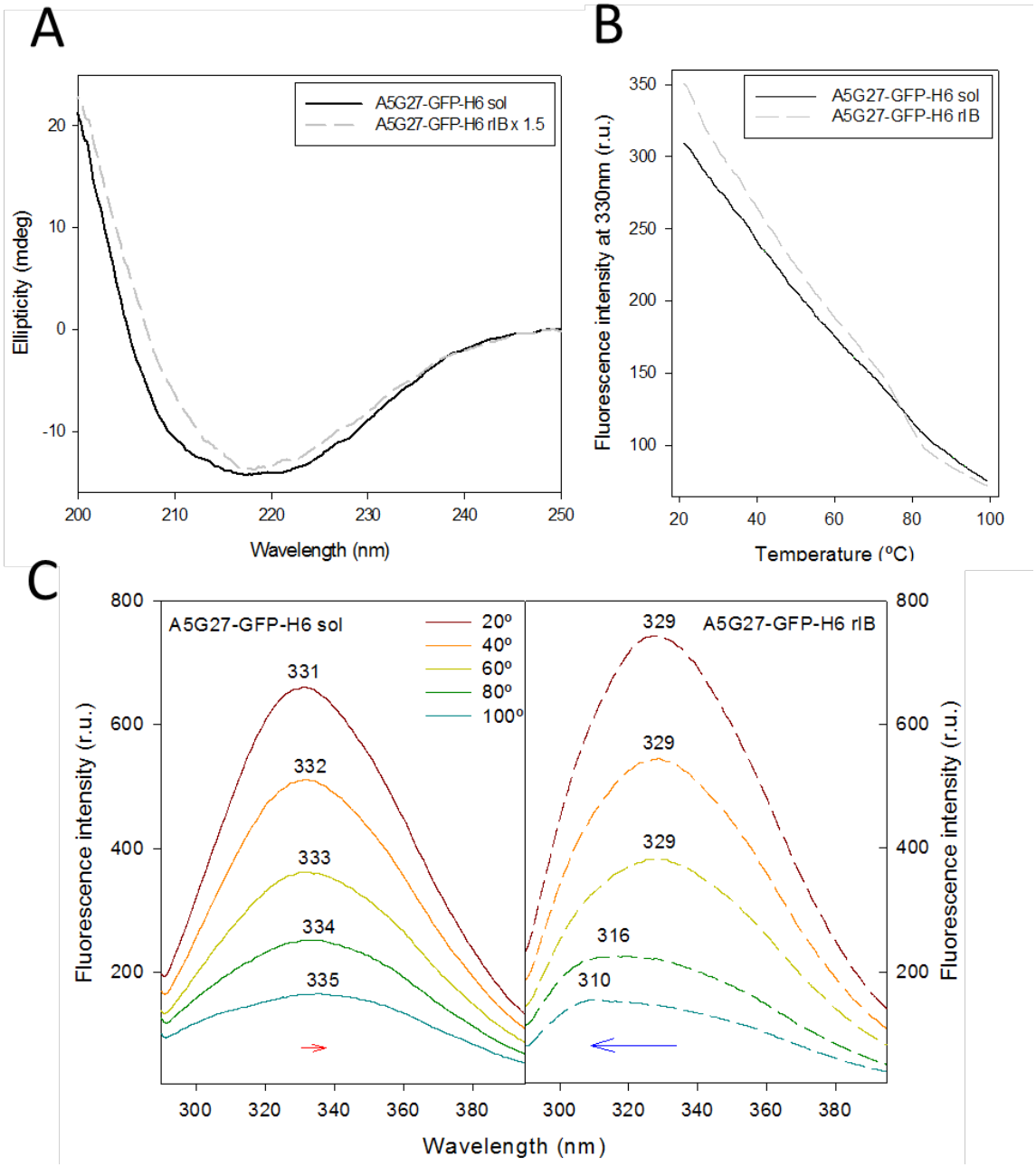

Figure 7. Conformational analyses of A5G27-GFP-H6 nanoparticle variants. A) Secondary structure analysis by far-UV CD spectroscopy at 25 ºC. Spectra were superimposed by applying a 1.5 multiplication factor. B) Temperature denaturation followed by Trp emission fluorescence at 330 nm. C) Tertiary structure analysis by Trp-fluorescence spectroscopy at different temperatures. The maximum of each spectrum is indicated. Arrows represent the red and blue-shift of fluorescence maximum with temperature for A5G27-GFP-H6 sol and rIB versions respectively.

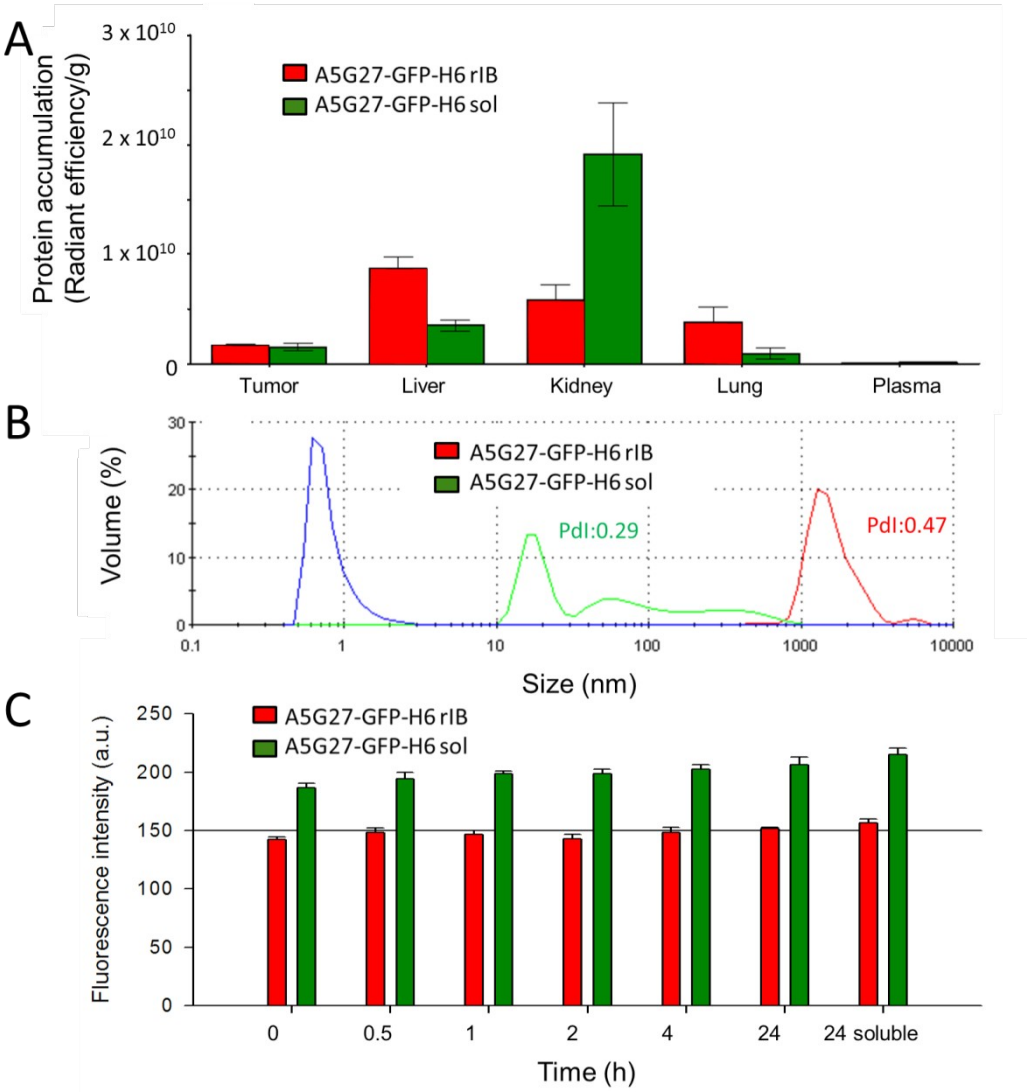

Figure 8. Biodistribution and stability in physiological medium of A5G27-GFP-H6 nanoparticles. A) Tumor- and tissue-accumulation map of A5G27-GFP-H6 nanoparticles *by ex vivo* determination of fluorescence in selected organs, excised 30 min after systemic administration of the materials. Mean ± SEM are displayed. B) Structural stability of A5G27-GFP-H6 nanoparticles after 2 h incubation in OptiPro™ medium as measured by DLS. In blue line, OptiPro™ control sample. Polydispersion indexes (PdI) of protein materials are indicated in the plots.. C) Stability of A5G27-GFP-H6 nanoparticles when incubated with human plasma over 24 h, measured by fluorescence emission at 510 nm. Potential sedimentation of insoluble protein materials was explored by determining the reduction of fluorescence of the soluble fraction after centrifugation (in the 24 h soluble sample).

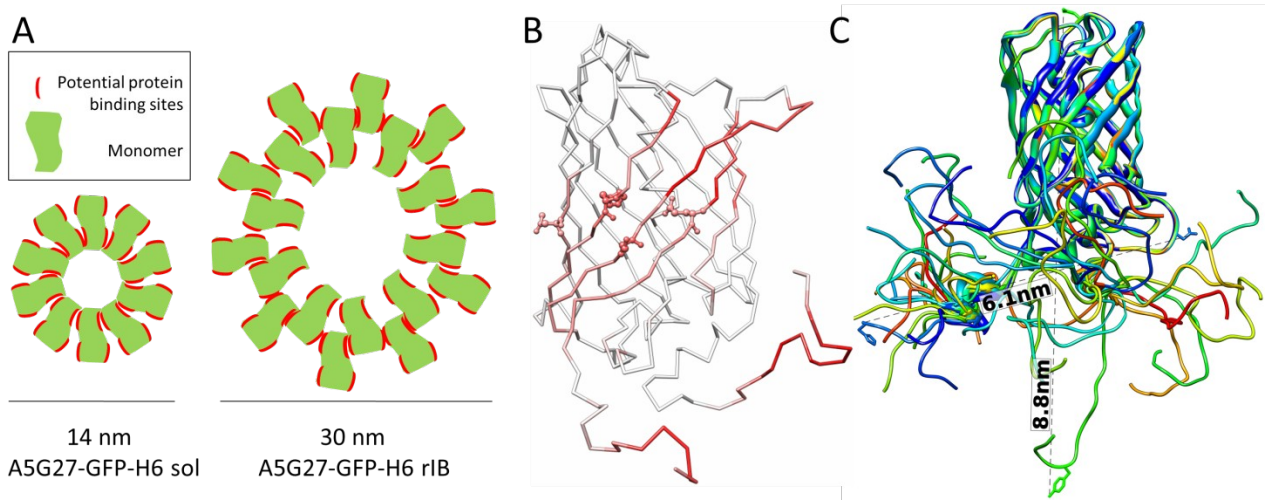

Figure 9. Putative organization patterns of A5G27-GFP-H6 nanoparticles as defined by alternative exposure of hydrophobic patches. A) Different dispositions of building blocks, depending on the protein source or downstream treatment might promote differential arrangements in the oligomers. A different pattern of protein-protein cross-interactions might influence not only particle geometry and size, but also relevant physicochemical and biological properties such as solubility, aggregation and liver uptake. B) Probability (from white to intense red) of A5G27-GFP-H6 residues belonging to a protein-protein interaction patch, as derived from prediction frequencies for the twenty best-scoring models of the protein construct. Residues 148, 164, 202, 203 [72], which are known to be involved in the modulation of the intensity of fluorescence emission [73], are shown as balls & sticks and are clearly found in the predicted interaction patch. C) The dynamic dimensions of the cylindrical are between 3.7 nm to 6.1 nm (radius) and between 4.9 nm to 8.8 nm (height). The lower bounds correspond to the values measured for the GFP beta-barrel and the upper bounds correspond to the largest values measured for the 20 best models (including the tails) generated.

Reference List

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

2 Seras-Franzoso,J. *et al.* (2015) Integrating mechanical and biological control of cell proliferation through bioinspired multieffector materials. *Nanomedicine. (Lond)* 10, 873-891

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

4 Ferrer-Miralles,N. *et al.* (2009) Microbial factories for recombinant pharmaceuticals. *Microb. Cell Fact.* 8, 17

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

6 Lee,S.Y. *et al.* (2012) Systems metabolic engineering, industrial biotechnology and microbial cell factories. *Microb. Cell Fact.* 11, 156

7 Villaverde,A. and Carrio,M.M. (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol. Lett.* 25, 1385-1395

8 Villaverde,A. *et al.* (2015) Functional protein aggregates: just the tip of the iceberg. *Nanomedicine. (Lond)* 10, 2881-2891

9 Baneyx,F. and Mujacic,M. (2004) Recombinant protein folding and misfolding in Escherichia coli. *Nat. Biotechnol* 22, 1399-1408

10 Vallejo,L.F. and Rinas,U. (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb. Cell Fact.* 3, 11

11 Singh,S.M. and Panda,A.K. (2005) Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* 99, 303-310

12 Gonzalez-Montalban,N. *et al.* (2007) Recombinant protein solubility-does more mean better? *Nat. Biotechnol.* 25, 718-720

13 Singh,A. *et al.* (2015) Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microb. Cell Fact.* 14, 41

14 Speed,M.A. *et al.* (1996) Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nat. Biotechnol.* 14, 1283-1287

15 Morell,M. *et al.* (2008) Inclusion bodies: specificity in their aggregation process and amyloid-like structure. *Biochim. Biophys. Acta* 1783, 1815-1825

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

17 Peternel,S. *et al.* (2008) Engineering inclusion bodies for non denaturing extraction of functional proteins. *Microb. Cell Fact.* 7, 34

18 Bradford,M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254

19 Eswar,N. *et al.* (2006) Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics* Chapter 5, Unit

20 Wizemann,H. *et al.* (2003) Distinct requirements for heparin and alpha-dystroglycan binding revealed by structure-based mutagenesis of the laminin alpha2 LG4-LG5 domain pair. *J. Mol. Biol.* 332, 635-642

21 Harrison,D. *et al.* (2007) Crystal structure and cell surface anchorage sites of laminin alpha1LG4-5. *J. Biol. Chem.* 282, 11573-11581

22 Ramakrishnan,V. *et al.* (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* 362, 219-223

23 Barondeau,D.P. *et al.* (2003) Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. *Proc. Natl. Acad. Sci. U. S. A* 100, 12111-12116

24 Shen,M.Y. and Sali,A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507-2524

25 Fiorucci,S. and Zacharias,M. (2010) Prediction of protein-protein interaction sites using electrostatic desolvation profiles. *Biophys. J.* 98, 1921-1930

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

27 de Marco,A. and Schroedel,A. (2005) Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC. Biochem.* 6, 10

28 Peternel,S. and Komel,R. (2010) Isolation of biologically active nanomaterial (inclusion bodies) from bacterial cells. *Microb. Cell Fact.* 9, 66

29 Jevsevar,S. *et al.* (2005) Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnol. Prog.* 21, 632-639

30 Garcia-Fruitos,E. *et al.* (2005) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb. Cell Fact.* 4, 27

31 Cano-Garrido,O. *et al.* (2013) Supramolecular organization of protein-releasing functional amyloids solved in bacterial inclusion bodies. *Acta Biomater.* 9, 6134-6142

32 Mitraki,A. (2010) Protein aggregation from inclusion bodies to amyloid and biomaterials. *Adv. Protein Chem. Struct. Biol.* 79, 89-125

33 Garcia-Fruitos,E. *et al.* (2011) Biological role of bacterial inclusion bodies: a model for amyloid aggregation. *FEBS J.* 278, 2419-2427

34 Villar-Pique,A. and Ventura,S. (2012) Modeling amyloids in bacteria. *Microb. Cell Fact.* 11, 166

35 Garcia-Fruitos,E. *et al.* (2012) Bacterial inclusion bodies: making gold from waste. *Trends Biotechnol* 30, 65-70

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

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

38 Villaverde,A. (2010) Nanotechnology, bionanotechnology and microbial cell factories. *Microb. Cell Fact.* 9, 53

39 Mironov,V. *et al.* (2009) Biofabrication: a 21st century manufacturing paradigm. *Biofabrication.* 1, 022001

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

41 Park,T.J. *et al.* (2015) Advances in microbial biosynthesis of metal nanoparticles. *Appl. Microbiol. Biotechnol.*

42 Vijayendra,S.V. and Shamala,T.R. (2014) Film forming microbial biopolymers for commercial applications--a review. *Crit Rev. Biotechnol.* 34, 338-357

43 Chen,G.Q. (2009) A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. *Chem. Soc. Rev.* 38, 2434-2446

44 Rodriguez-Carmona,E. and Villaverde,A. (2010) Nanostructured bacterial materials for innovative medicines. *Trends Microbiol.* 18, 423-430

45 Chen,R. (2012) Bacterial expression systems for recombinant protein production: E. coli and beyond. *Biotechnol. Adv.* 30, 1102-1107

46 Corchero,J.L. *et al.* (2013) Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. *Biotechnol. Adv.* 31, 140-153

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

48 Chen,G.Q. (2012) New challenges and opportunities for industrial biotechnology. *Microb. Cell Fact.* 11, 111

49 Liu,L. *et al.* (2013) Recent advances in recombinant protein expression by Corynebacterium, Brevibacterium, and Streptomyces: from transcription and translation regulation to secretion pathway selection. *Appl. Microbiol. Biotechnol*

50 Rueda,F. *et al.* (2014) Production of functional inclusion bodies in endotoxin-free Escherichia coli. *Appl. Microbiol. Biotechnol.* 98, 9229-9238

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

52 Kobayashi,N. *et al.* (2015) Self-Assembling Nano-Architectures Created from a Protein Nano-Building Block Using an Intermolecularly Folded Dimeric de Novo Protein. *J. Am. Chem. Soc.* 137, 11285-11293

53 Hernandez-Garcia,A. *et al.* (2014) Design and self-assembly of simple coat proteins for artificial viruses. *Nat. Nanotechnol.* 9, 698-702

54 Doll,T.A. *et al.* (2015) Optimizing the design of protein nanoparticles as carriers for vaccine applications. *Nanomedicine.* 11, 1705-1713

55 Kaltofen,S. *et al.* (2015) Computational de novo design of a self-assembling peptide with predefined structure. *J. Mol. Biol.* 427, 550-562

56 Sonmez,C. *et al.* (2015) Design of self-assembling peptide hydrogelators amenable to bacterial expression. *Biomaterials* 37, 62-72

57 Lai,Y.T. *et al.* (2014) Structure of a designed protein cage that self-assembles into a highly porous cube. *Nat. Chem.* 6, 1065-1071

58 Tamamis,P. *et al.* (2014) Combination of theoretical and experimental approaches for the design and study of fibril-forming peptides. *Methods Mol. Biol.* 1216, 53-70

59 Bayburt,T.H. and Sligar,S.G. (2010) Membrane protein assembly into Nanodiscs. *FEBS Lett.* 584, 1721-1727

60 Bayburt,T.H. *et al.* (1998) Reconstitution and imaging of a membrane protein in a nanometer-size phospholipid bilayer. *J. Struct. Biol.* 123, 37-44

61 Rueda,F. *et al.* (2015) Bottom-Up Instructive Quality Control in the Biofabrication of Smart Protein Materials. *Adv. Mater.* 27, 7816-7822

62 Peternel,S. *et al.* (2008) New properties of inclusion bodies with implications for biotechnology. *Biotechnol. Appl. Biochem.* 49, 239-246

63 Rejman,J. *et al.* (2004) Size-dependent internalization of particles via the pathways of clathrin- and caveolae-mediated endocytosis. *Biochem. J.* 377, 159-169

64 Shang,L. *et al.* (2014) Nanoparticle interactions with live cells: Quantitative fluorescence microscopy of nanoparticle size effects. *Beilstein. J. Nanotechnol.* 5, 2388-2397

65 Shang,L. *et al.* (2014) Engineered nanoparticles interacting with cells: size matters. *J. Nanobiotechnology.* 12, 5

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

67 Alexis,F. *et al.* (2008) Factors affecting the clearance and biodistribution of polymeric nanoparticles. *Mol. Pharm.* 5, 505-515

68 Hirn,S. *et al.* (2011) Particle size-dependent and surface charge-dependent biodistribution of gold nanoparticles after intravenous administration. *Eur. J. Pharm. Biopharm.* 77, 407-416

69 Nie,S. (2010) Understanding and overcoming major barriers in cancer nanomedicine. *Nanomedicine. (Lond)* 5, 523-528

70 Garnett,M.C. and Kallinteri,P. (2006) Nanomedicines and nanotoxicology: some physiological principles. *Occup. Med. (Lond)* 56, 307-311

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

72 Yang,F. *et al.* (1996) The molecular structure of green fluorescent protein. *Nat. Biotechnol* 14, 1246-1251

73 (2015) The Fluorophore of Green Fluorescent Protein (GFP). http://dwb.unl.edu/Teacher/NSF/C08/C08Links/pps99.cryst.bbk.ac.uk/projects/gmocz/gfp.htm