For reprint orders, please contact: reprints@futuremedicine.com

Nanoparticulate architecture of protein-based artificial viruses is supported by protein–DNA interactions

Aim & Methods: We have produced two chimerical peptides of 10.2 kDa, each contain four biologically active domains, which act as building blocks of protein-based nonviral vehicles for gene therapy. In solution, these peptides tend to aggregate as amorphous clusters of more than 1000 nm, while the presence of DNA promotes their architectonic reorganization as mechanically stable nanometric spherical entities of approximately 80 nm that penetrate mammalian cells through arginine–glycine–aspartic acid cell-binding domains and promote significant transgene expression levels. **Results & Conclusion:** The structural analysis of the protein in these hybrid nanoparticles indicates a molecular conformation with predominance of α-helix and the absence of cross-molecular, β-sheet-supported protein interactions. The nanoscale organizing forces generated by DNA–protein interactions can then be observed as a potentially tunable, critical factor in the design of protein-only based artificial viruses for gene therapy.

KEYWORDS: DNA–protein interaction ▪ gene therapy ▪ innovative medicine ▪ nanomedicine ▪ protein engineering ▪ protein nanoparticle

Joan Domingo-Espín[1,2], Esther Vazquez[1,2], Javier Ganz[3,4], Oscar Conchillo[1], Elena García-Fruitós[1,2], Juan Cedano[1], Ugutz Unzueta[1,2], Valérie Petegnief[5], Nuria Gonzalez-Montalbán[1,2], Anna M Planas[5], Xavier Daura[1,6], Hugo Peluffo[3,4], Neus Ferrer-Miralles[1,2] & Antonio Villaverde[†1,2]

[1]*Institute for Biotechnology & Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain*
[2]*CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain*
[3]*Neurodegeneration Laboratory, Institut Pasteur de Montevideo, CP 11400, Montevideo, Uruguay*
[4]*Department of Histology & Embryology, Faculty of Medicine, UDELAR, CP 11800, Montevideo, Uruguay*
[5]*Departament d'Isquèmia Cerebral i Neurodegeneració, Institut d'Investigacions Biomèdiques de Barcelona (IIBB), Consejo Superior de Investigaciones Científicas (CSIC)-Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain*
[6]*Catalan Institution for Research & Advanced Studies (ICREA), 08010 Barcelona, Spain*
[†]*Author for correspondence: Tel.: +34 935 813 086 antoni.villaverde@uab.cat*

Strategies for nonviral gene therapy are under continuous exploration, pressured by the undesired side effects observed in viral-based gene therapy trials [1–5]. In this context, the 'artificial virus' approach implies the use of noninfectious and biologically safe entities that mimic relevant features of the viral life cycle, as DNA carriers for the cell-targeted delivery of therapeutic nucleic acids [6–8]. Liposomes, carbohydrates and proteins are the most commonly used scaffolds for the construction of bio-inspired artificial viruses, although the functionalization necessary for specific receptor binding, endosomal escape and nuclear trafficking, among others, is mostly provided by proteins (namely peptides, full-length proteins or antibodies). In fact, proteins organized as cages in diverse forms, are considered excellent and fully biocompatible carriers for drug delivery [9]. In this regard, virus-like particles (VLPs) mainly formed by self-assembling capsid proteins from *Papillomaviridae* and *Polyomaviridae* viral families have been explored as gene therapy vehicles (once filled *in vitro* with nucleic acids) [10], either by keeping the original tropism of natural viruses or upon functionalization by the appropriate display of foreign functional peptides. These studies have also been extended to bacterial viruses, which might be more convenient regarding scaled-up production. For instance, phage MS2 VLPs loaded with antisense oligodeoxynucleotides and decorated with transferrin have been proven active on leukemia cancer cells [11]. However, despite the convenient regularity of size exhibited by VLPs, their architectonic constraints limit their extensive engineering and the possibility of functional tuning.

A more versatile scheme of protein-based carriers for therapeutic nucleic acids are multifunctional proteins, constructed by the combination of appropriate functional domains fused in a single polypeptide chain [12]. The integrated domains enable the whole construct to mimic the activities of the infective viral cycle that are relevant to the targeted delivery of nucleic acids (namely binding of DNA or RNA, cell attachment and internalization, endosomal escape, proper cytoplasmic trafficking, eventual nuclear transport and nucleic acid release). The modular nature of such constructs permits the selection of functions using relevant peptides identified from nature or combinatorial libraries, and a functional redesign in iterative improvement processes [13,14]. Diverse protein vehicles within this category have been successful in promoting significant transgene expression levels *in vitro* [15–17] and therapeutic effects *in vivo* [18,19], proving the potential of this approach in the clinical context.

Interestingly, nonviral vehicles based on multifunctional proteins have been scantily characterized from the morphologic point of view. Therefore, information regarding how these proteins might organize as building blocks of higher order structures, and how protein–DNA complexes are formed and shaped is, in general, not available. Therefore, particle size

future medicine part of fsg

and molecular organization, nanoscale properties potentially critical for cell attachment, internalization and endosomal escape remain excluded from potential tailoring. To approach this issue, in *Escherichia coli* we produced two different versions of very short structural proteins as subunits for artificial viruses based on alternative combinations of four functional domains (an integrin-binding motif, an endosomal escape domain, a nuclear localization signal and a DNA-binding, cationic peptide) joined in short peptide stretches. Significant levels of transgene expression driven by the complexes have been observed, proving the appropriate selection of the functional domains. On the other hand, in the absence of DNA, protein blocks self-organize as amorphous, polydisperse particulate entities ranging from a few nanometers up to approximately 1 µm. However, in presence of DNA, protein–DNA complexes appear as tight and rather monodisperse spherical-like nanoparticles of approximately 80 nm in diameter that resemble bacterial inclusion bodies (IBs), in which proteins remain attached by β-sheet-based cross-molecular interactions. However, both protein modeling and structural analysis of these complexes reveal an unexpected molecular organization that does not rely on protein–protein cross-molecular interactions but that is instead supported by protein–DNA interactions. Such DNA-mediated organization seems to generate an optimal architectural pattern of artificial viruses based on short multifunctional proteins as building blocks.

Materials & methods

Plasmid construction & protein sequence

Plasmid pET28aTEV, derived from pET28a (Invitrogen) in which the DNA sequence encoding the thrombin cleavage site was substituted by a DNA fragment encoding a tobacco etch virus (TEV) protease cleavage site, was used to generate constructs pET28aTEV-HKRN and pET28aTEV-HNRK. HKRN and HNRK correspond to DNA sequences coding for selected modules in the specified order (Figure 1A). Plasmids were constructed by introducing synthetic oligonucleotides, encoding the corresponding modules, into selected restriction enzyme sites of the multiple cloning site of pET28aTEV. The arginine–glycine–aspartic motif used here derives from the foot-and-mouth disease virus (serotype C_1) cell-binding protein [20], and it is known to bind mammalian cells through $\alpha_v\beta_3$ and $\alpha_5\beta_1$ integrins [21,22]. The nuclear localization signal of the Simian virus 40 (SV40) large T-antigen [23] has been universally used for the nuclear transport of delivered drugs and DNA [24]. The polylysine (Lys) tail (K10) is a cationic peptide extensively used as a DNA-condensation agent in artificial viruses [25], while the polyhistidine tail (H6) is both an efficient endosomal-escape peptide [14] and a convenient tag for one-step protein purification from bacterial cell extracts [26]. Finally, the biologically irrelevant central amino acid stretch in both HKRN and HNRK was added to enlarge the mass of the resulting modular peptides and to make them more stable in bacterial cells, according to our previous experience [Domingo-Espin J, Unpublished Data].

Protein production & purification

The production of both chimerical proteins was triggered by the addition of 1 mM IPTG to plasmid-containing BL21(DE3) *E. coli* cell cultures (at OD = 0.4–0.6) growing in Luria–Bertani medium at 37°C. After 4 h, cells were harvested by centrifugation, washed with phosphate-buffered saline and stored at -80°C until use. The pellet was resuspended in lysis buffer (20 mM Tris-HCl pH 8, 500 mM NaCl and 6 M GuHCl) and cells were disrupted by sonication in the presence of EDTA-free protease inhibitor cocktail tablets. The soluble fraction was separated by centrifugation at 15,000 g for 45 min at 4°C and filtered through 0.22-µm filters. Proteins were purified in a single step by Ni^{2+} affinity chromatography in an ÄKTA™ FPLC (GE Healthcare) using a 20 CV linear gradient to 100% of elution buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 6 M ClHGu and 1 M imidazole). Positive fractions were collected and passed through a PD-10 desalting column (GE Healthcare) with 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid-buffered saline and quantified by Bradford's method. Finally, proteins were stored at -80°C until use. IBs used for scanning electron microcopy were purified as described elsewhere [27].

Mass spectrometry

Mass spectrometry was perfomed on 0.5 µl of protein sample mixed with 0.5 µl 2,4-dihydroxyacetophenone (10 mg/ml in 20 mM ammonium citrate, 30% acetonitrile) spotted onto a ground steel plate (Bruker) and allowed to air-dry at room temperature. MALDI-mass spectra were recorded in the positive ion mode on an Ultraflex time-of-flight instrument (Bruker). Ion acceleration was set to 20 kV. All mass spectra were externally calibrated using a standard protein mixture.

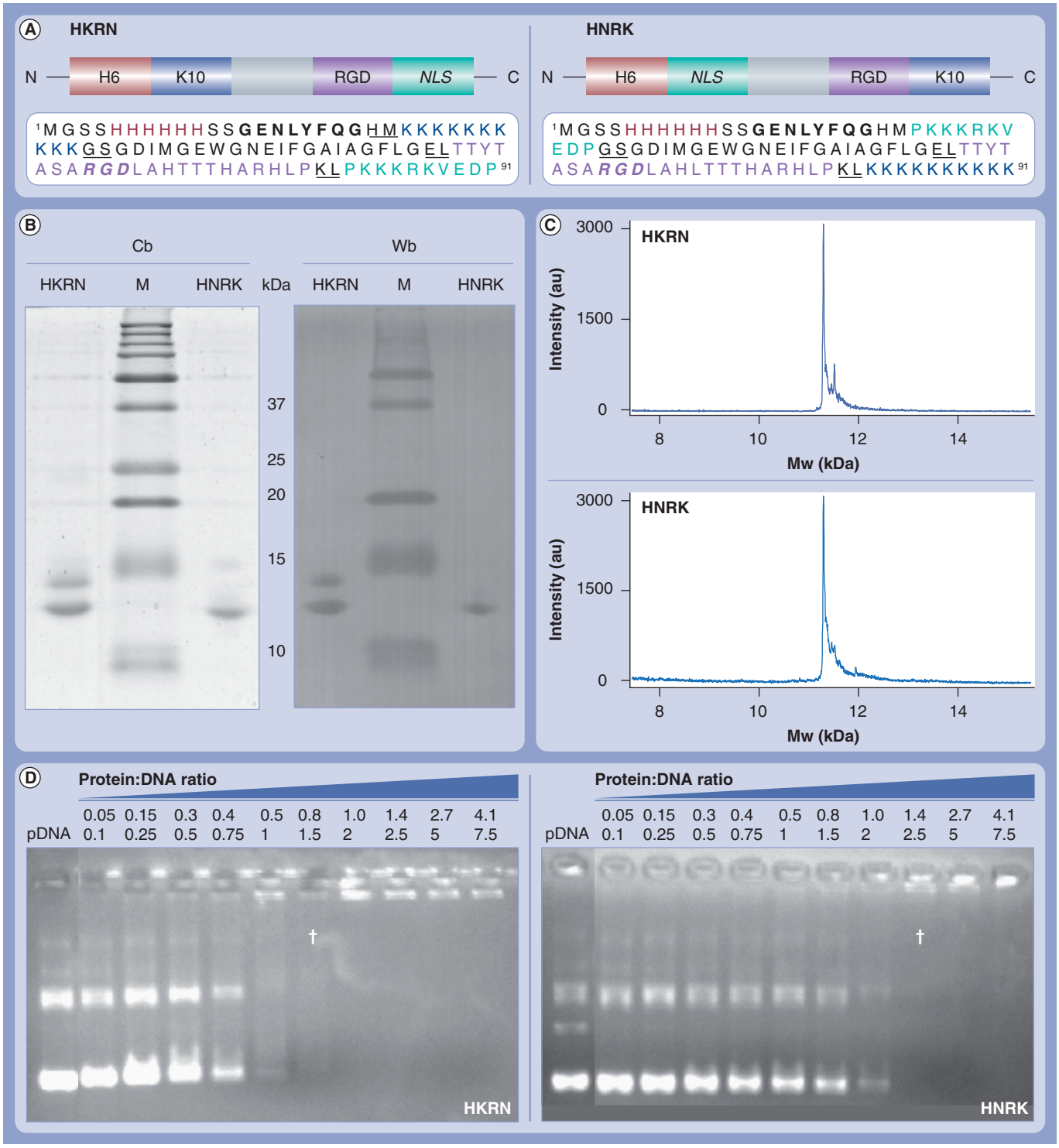

Figure 1. HKRN and HNRK organization and main features. (A) Distribution of functional modules in HKRN, HNRK and the amino acid sequence of the whole protein constructs. In both cartoons and sequence, the histidine (H) tail is labeled in red, the lysine (K) tail in dark blue, the FMDV cell binding (arginine–glycine–aspartic [RGD]) in purple and the SV40 nuclear localization signal (NLS) in green. The irrelevant central region is depicted in gray. In the amino acid sequences, residues resulting from the cloning process are underlined, a tobacco etch virus protease target site introduced between H and the immediate carboxy module is indicated in boldface and the RGD motif within the FMDV peptide is shown in italics. Sizes of the modules in the cartoons are not intended to be representative of the actual length in the protein segments. **(B)** Cb staining and Wb analysis (using an anti-His antibody) of SDS-PAGE of HKRN and HNRK upon purification. Molecular masses of the markers (M) are indicated in the central column. **(C)** Mass spectrometry of both pure proteins. **(D)** Up-shift of pcDNA3.1 (harboring the tdTomato gene) mobility in agarose gel electrophoresis as induced by increasing amounts of HKRN and HNRK.
†The protein–DNA charge and mass ratios at which migration of DNA is fully impeded (one retardation unit).

Retardation assay

Different protein amounts were incubated with 300 ng of pcDNA3-tdTOMATO plasmid DNA resulting in 0.05, 0.1, 0.25, 0.5, 0.75, 1, 1.5, 2.5, 5 and 7.5 protein/DNA w/w ratios, which corresponded to 0.05, 0.15, 0.30, 0.40, 0.50, 0.80, 1.00, 1.40, 2.70 and 4.10 protein/DNA charge ratios), respectively. Mixtures were incubated in microcentrifuge tubes at room temperature for 1 h in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid-buffered saline, and complex formation was detected in 0.8% agarose gels. One retardation unit (RU) is defined as the minimum protein/DNA ratio that does not allow DNA migration on agarose gels.

Electron microscopy

Complexes of protein and DNA were observed using transmission electron microscopy (TEM) with the aid of negative staining. One drop of the mixture was applied to glow-discharged carbon-coated copper grids (SPI Supplies®) for 5 min and then drained off with filter paper. Subsequently, one drop of 2% uranyl acetate was placed on the grid for 2–3 min before being drained off. The grid was then placed in a transmission electron microscope (Jeol JEM 1400) operating at an accelerating voltage of 120 kV. Images were acquired using a CCD camera (Gatan) and saved as 8-bit images. A series of micrograph images were obtained tilting the sample from -60° to +60° with a 914 High Tilt Holder.

Inclusion bodies were analyzed by scanning electron microscopy by standard procedures using Quanta FEI 200 filed-emission gun environmental scanning electron microscope.

Structural analysis

For circular dicroism (CD), samples were prepared at a protein concentration of 200 µM. Two samples were incubated with DNA at different ratios corresponding to 0.5 and 2 RU. Cuvettes with path lengths of 0.1 cm were used, and eight scans recorded at 50 nm min (response of 2 s) in a JASCO 715 spectrophotopolarimeter were averaged for each variant. For Fourier-transformed infrared spectroscopy (FTIR), samples were analyzed in a Bruker Tensor 27 FTIR spectrometer (Broker Optics Inc.) For each spectrum, 16 scans were acquired at a spectral resolution of 4 cm^{-1} in the 4000–600 cm^{-1} range in the transmission mode. All processing procedures were carried out to optimize the quality of the spectrum in the amide I region ranging from 1750 to 1550 cm^{-1}. Second derivatives of the amide I band spectra were used to determine the frequencies at which the different spectral components were located.

Protein structure modeling

The 3D structures of the chimeric peptides were modeled with modeller 9v7 [28] using the coordinates of the original protein segments (when available) as templates. Thus, the structures of the nuclear localization signal and arginine–glycine–aspartic modules were based on chain B of 1Q1S [29] and chain 5 of 1QGC [30], respectively. The poly-Lys module was modeled, on the sole basis of the force field, as an unstructured segment, in line with the structural diversity reported for poly-Lys peptides [31] and the disorder of the poly-Lys tail in the structure with PDB code 1KVN [32]. The central region was modeled using chain A of 1HA0 [33] as a template (61.9% similarity). Hexa-histidine peptides have become one of the most popular tags for protein purification, but the abundance of His-tagged protein models contrast with the lack of structure in which this tag has been successfully solved. This fact clearly indicates that this region tends to be intrinsically unstructured and it was not suitable for modeling under our approach, being then absent in the models.

Dynamic light scattering

Volume-size distribution of DNA–protein complexes at different weight ratios was determined in a dynamic light scattering device (Zetasizer Nano ZS, Malvern Instruments Limited) using DTS (Nano) version 5.10 software for data evaluation.

Transfection, flow cytometry analysis & fluorescent microscopy

The HeLa (ATCC-CCL-2) cell line was maintained in minimal essential medium (GIBCO) supplemented with 10% fetal calf serum (GIBCO) and incubated at 37°C and 5% CO_2 in a humidified atmosphere in 24-well plates at a cell confluence of 70–80%. The vectors pcDNA3-tdTOMATO and pEGFP-C1 (Clontech), carrying the gene of the fluorescent proteins tdTOMATO and EGFP, respectively, were used to monitor DNA transfection. DNA–HKRN or DNA–HNRK complexes were prepared incubating different amounts of protein in 50 µl OptiPRO (GIBCO)medium and different amounts of DNA in 50 µl OptiPRO (GIBCO) medium. After 5 min, DNA–protein complexes were generated by mixing DNA and protein at specified protein–DNA ratios at room

temperature for 1 h. A total of 100 µl OptiPRO (GIBCO) was then added to the mixture and then to the cells. Transfection and gene expression was monitored by flow cytometry in a FACSCalibur system (Becton Dickinson) at 24 h and confirmed at 48 h post-transfection in a fluorescence microscope (Nikon ECLIPSE TE2000-E). As controls, we used nontreated cells, cells exposed only to the protein and cells exposed only to plasmid DNA.

Primary cell cultures

Cortical neuron cultures were prepared from 18-day-old Sprague–Dawley rat embryos (Charles River Laboratories), as described previously [34]. Animals were anaesthetized and killed by cervical dislocation. All procedures were approved by the Ethical Committee for Animal Use (CEEA) at the University of Barcelona, Spain. Cells were seeded on 24-well plates at a density of 1580 cells/mm^2 in neurobasal medium supplemented with 2% B27 supplement, 0.5 mM glutamine and 0.1 mg/ml gentamycin. Partial medium changes were performed *in vitro* on days 4 and 7. Transfection was performed *in vitro* on day 10 as for HeLa cells, except that the transfection medium was neurobasal:conditioned medium (2:1). Gene expression was confirmed at 24 h postinfection in a fluorescence microscope (Olympus IX71).

Luciferase gene expression

HKRN or HNRK were incubated at room temperature for 1 h with pGL3-BOS-luciferase reporter plasmid (kindly provided by Marta Barrachina) at the indicated ratios of protein/DNA in 20–30 µl of Opti-MEM® medium. Subconfluent HEK293 cells were washed once with Opti-MEM and then incubated with the protein/DNA complexes for 4 h. The medium was then removed and cells maintained in DMEM+10% fetal bovine serum for another 48 h. The measurement of luciferase activity was performed according to the manufacturer's instructions (Luciferase Reporter Gene Detection Kit, SIGMA Cat. LUC1–1KT). As a control reference, cells were transfected with lipofectamine 2000 (Invitrogen, 2 µg lipofectamine + 1 µg DNA/well on 24-well plate) and data were expressed as percentage relative light units per µg of protein in the samples compared with lipofectamine 2000.

Results

The chimerical genes encoding the multifunctional proteins HKRN and HNRK were constructed by ligation of partially overlapping and complementary oligonucleotides, encoding four selected protein domains, in which the codon usage had been optimized for *E. coli*. Both polypeptides, containing the same functional motifs displayed in alternative positions (Figure 1A), were successfully produced in *E. coli* BL21 (DE3) pLysS, in full-length forms and at reasonably high yield (~4 µg of protein per ml of culture). Western blot analyses of purified proteins revealed the absence of truncated protein versions and the minor occurrence of high molecular mass immunoreactive species, especially in HKRN, which might indicate a tendency to form supramolecular structures (Figure 1B). The occurrence of such cross-interactions was supported by the high purity observed in samples of both proteins (Figure 1C), and the absence of major isoforms derived from partial proteolysis. When HKRN and HNRK were challenged in DNA retardation assays, HKRN showed a higher capability (1 RU corresponding to a protein/DNA mass ratio of 1.5 and to a DNA/protein charge ratio of 0.8) than HNRK (1 RU corresponding to a protein/DNA mass ratio of 2.5 and to a DNA/protein charge ratio of 1.3) to impede the mobility of plasmid DNA (Figure 1D). This divergence could be accounted by either a different oligomerization potential or by a different performance of the DNA binding domain (K10) as alternatively positioned in HKRN and HNRK. In the first case, K10 was placed in an internal position within the amino terminal protein moiety and in HNRK, this peptide overhanged as a C-terminal end.

The resulting protein–DNA complexes (nonviral vehicles) were tested in HeLa cell cultures for their ability to promote expression of a plasmid-harbored reporter transgene. Although the design of nonviral vehicles for gene therapy is a rather trial-and-error process, we expected that the combination of the FMDV integrin-binding motif, the SV40 nuclear localization signal, the His-based endosomal escape peptide and the Lys-based DNA binding stretch could summarize the main viral functions required for cell uptake and trafficking of the cargo DNA and result in significant levels of nuclear gene delivery and expression. In agreement with this presumption, flow cytometry analysis of cultured cells 48 h after exposure to HKRN–DNA and HNRK–DNA complexes revealed the occurrence of significantly prevalent cell subpopulations expressing the reporter tdTomato gene. In this context, more than 10% of HeLa cells transfected with HNRK-based vehicles emitted red fluorescence, indicating the proper nuclear delivery and release

of the carried DNA. However, being still significant, DNA delivery mediated by HKRN resulted in rather moderate transgene expression that was detected in only 0.5% of the cell population (Figure 2A). In order to eliminate the chance that this value could be due to experimental noise, we examined the cultures treated with HKRN-based complexes by fluorescence microscopy *in situ*. Clear fluorescence emission in individual cultured HeLa cells was detected when using two different reporter genes, namely *EGFP* and tdTomato (involving >10% of cells at 24 h; Figure 2B, top, middle panel). Furthermore, in primary cultures of neurons and glia, several cells strongly expressing tdTomato were observed 24 h after transfection with the DNA–HKRN complex. A cell with neuronal morphology strongly expressing the tdTomato gene in the cell body and neurites is shown in the inset of Figure 2B, bottom, demonstrating that neurons can be effectively transfected and the transgene transcribed and translated into protein. An additional transfection experiment on Hek293 cells with a third reporter luciferase (*luc*) gene confirmed the consistent transgene expression mediated by HKRN (Figure 2C). These data demonstrate the stability, robustness and good performance of both HKRN and HNRK as nonviral gene vectors and the appropriateness of the selected protein modules to mediate DNA delivery, being the modular distribution in HNRK more convenient for the proper mimicking of viral functions.

Intriguingly, the morphology and structure of protein–DNA complexes in nonviral gene therapy has been historically neglected, and for protein-based vehicles other than VLPs, the concept of an artificial virus refers exclusively to functional (instead of nanoscale physical) properties. Therefore, at this stage, we were especially interested in evaluating the architectonic properties of both constructs as building blocks of artificial viruses, and how these multifunctional protein subunits should be organized to bind plasmid DNA. To explore the molecular organization of the artificial viruses we approached their structural analysis from different angles. Interestingly, the TEM images of both peptides alone indicated the occurrence of amorphous, highly dispersed protein clusters of approximately 1 µm without any apparent morphological pattern and internal organization (Figure 3A). However, the protein–DNA complexes formed by HKRN and HNRK organized as regular, pseudo-spherical nanoparticles of approximately 80 nm in diameter (Figures 3A), morphologically resembling bacterial IBs [35–37] (although these last particles can be slightly larger, up to 450 nm in diameter [35]). The molecular reorganization of the protein building blocks induced by the addition of DNA occurred at 0.5 but not 2 RU (Figure 3B), and it did not prevent the emergence of larger protein clusters (Figure 3B, see arrow). These micron-sized particles, as seen by dynamic light scattering, are probably transient and reversible clusters of the 80-nm particles promoted by overhanging DNA molecules, since complexes of this size were uniquely, consistently and abundantly observed by TEM (Figure 3A). Despite the absence of nanosized particles at 2 RU, the size variability of DNA–protein complexes was strongly reduced when comparing with proteins alone (Figure 3B), indicating that the presence of DNA promoted conformational alterations on the holding proteins with impact in their oligomeric organization. The regularity of size in the protein–DNA complexes compared with the protein alone also indicates protein-condensing abilities of plasmid DNA that reduce the molecular stickiness (their aggregation tendency) of HKRN and HNRK proteins. This fact strongly suggests that the cationic poly-Lys stretches, responsible for DNA binding in multifunctional proteins [25] and whose charge is expected to be neutralized in the complexes, effectively drive the unspecific formation of higher order, protein-alone clusters shown in Figure 3A. Taken together, these data indicate that HKRN and HNRK, apart from exhibiting the functions associated to their forming protein domains, act as efficient building blocks for the construction of artificial viruses under the architectonic scope of this term.

A ±60° TEM scan of HNRK revealed a slightly flattened ellipsoid form of the protein–DNA complexes (Figure 4A), again very similar to the images of IBs formed by other proteins seen by atomic force microscopy [35]. In fact, HKRN and HNRK themselves are both partially found as IBs in the cytoplasm of the producing bacteria (Figure 4C). In this context, we were interested in determining the eventual architectonic coincidences between protein–DNA complexes and IBs formed by the protein counterparts. As determined by conformational analysis through FTIR [38–40], IBs gain their mechanical structure and shape by cross-molecular protein–protein interactions supported by a β-sheet-based, amyloid-like architecture [41,42], and we wondered if the architecture of the 80-nm artificial viruses formed by HKRN–DNA and HNRK–DNA complexes could also be supported by protein–protein interactions.

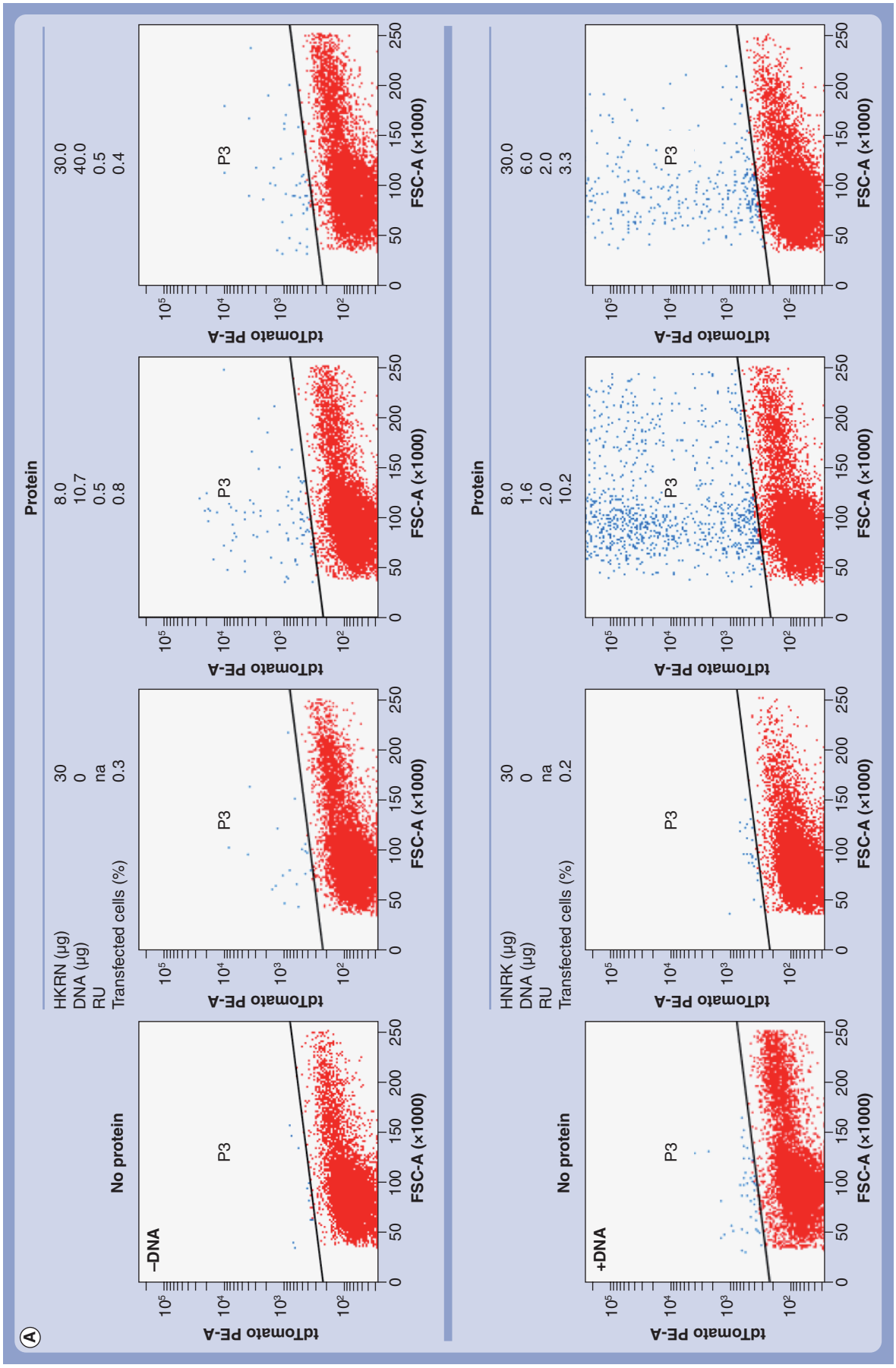

Figure 2. Transgene expression mediated by HKRN- and HNRK-based artificial viruses. (A) Fluorescence emission determined by flow cytometry of cultured HeLa cells 48 h after exposure to 24 µg of pcDNA3.1 or in the absence of foreign DNA (no protein). Cells were also exposed to HKRN–DNA and HNRK–DNA complexes and to these proteins alone (protein), and the P2 section in the plots corresponds to the red fluorescence emitted by the tdTomato protein. The percentages of fluorescent cells are indicated above each plot. na: Not available.

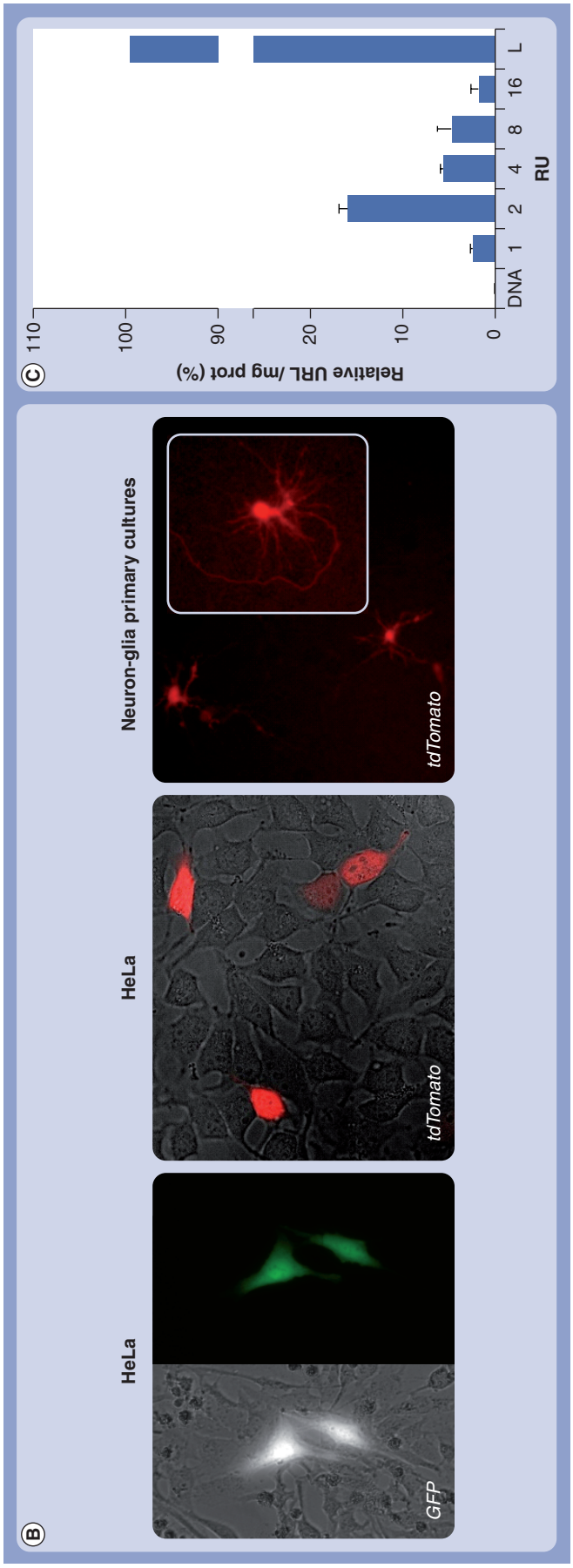

Figure 2 (cont.). Transgene expression mediated by HKRN- and HNRK-based artificial viruses. (B) Fluorescence microscopy of cultured cells 24 h after exposure to HKRN–DNA complexes formed at 2 RU. **(C)** Light emission of HEK293 cells 48 h after exposure to HKRN–pBOS complexes formed at 1, 2, 4, 8 and 16 RU (each experiment was performed at 28.8 µg of HKRN). pBOS alone at a concentration equivalent to 2 RU (5.75 µg/well) was used as negative control. Lipofectamine 2000 (2 µg/well + 1 µg pBOS) was used as a positive control (L).

Therefore, we approached the FTIR analysis of artificial viruses and their protein building blocks taking HNRK as a model. HNRK conformational features were analyzed under different biophysical states, such as naturally occurring IBs in bateria, in soluble form, lyophilized and in complexes with DNA. The conformational status of HNRK in IBs was found to be similar to those described previously as formed by other recombinant proteins, and characterized by the presence of extended, cross-molecular β-pleated sheet elements peaking at 1621 cm^{-1} (Figure 4B, top) [39,42,43]. On the other hand, in the HNRK–DNA complexes other secondary elements not present in HNRK IBs, such as native α-helices and unordered structures, were also detected (corresponding to the overlapped region between 1640 and 1660 cm^{-1}). In agreement with that observed by *in silico* modeling (Figure 4D) and as expected for short peptides, both HKRN and HNRK are, in general, unstructured. However, some locally structured regions inherited from their templates were noted in the models, namely a three to ten helix spanning residues 44–47 in HNRK and 45–48 in HKRN, apart from some additional turns and bends (Figure 4D). Accordingly, soluble HNRK was seen to have α-helix elements peaking at 1654 cm^{-1} (Figure 4B, center, green line). Interestingly, upon lyophilizing, HNRK seemed to evolve in a more lightly loose and unordered structure, as it can be seen by the broad peak between 1640 and 1660 cm^{-1} (Figure 4B, center, black line).

In agreement with the structural impact of DNA on the complexes suggested by dynamic light scattering data (Figure 3), the presence of the plasmid DNA had a critical effect on the peptide structure (Figure 4B, bottom), preventing the smooth deconstruction of α-helices observed during the lyophilization of HNRK alone. In addition, HNRK α-helices gained looseness along with the increase of DNA–HNRK ratio, as can be seen by the slight shift from lower wavenumber, from 1653 cm^{-1} in the lyophilized sample without DNA (Figure 4B, bottom, black line) to 1651 and 1650 cm^{-1} in the HNRK 2 RU and 0.5 RU (Figure 4 B, bottom, red line and blue line, respectively). This minor but significant shift might suggest that the binding of DNA to the protein shells is not a random, but an organized event possibly involving the central α-helix region of the peptide. Such interaction could account for the architectonic organization emerging in the artificial viruses and absent in the protein building blocks alone. The gain of peptide organization promoted by DNA was further confirmed

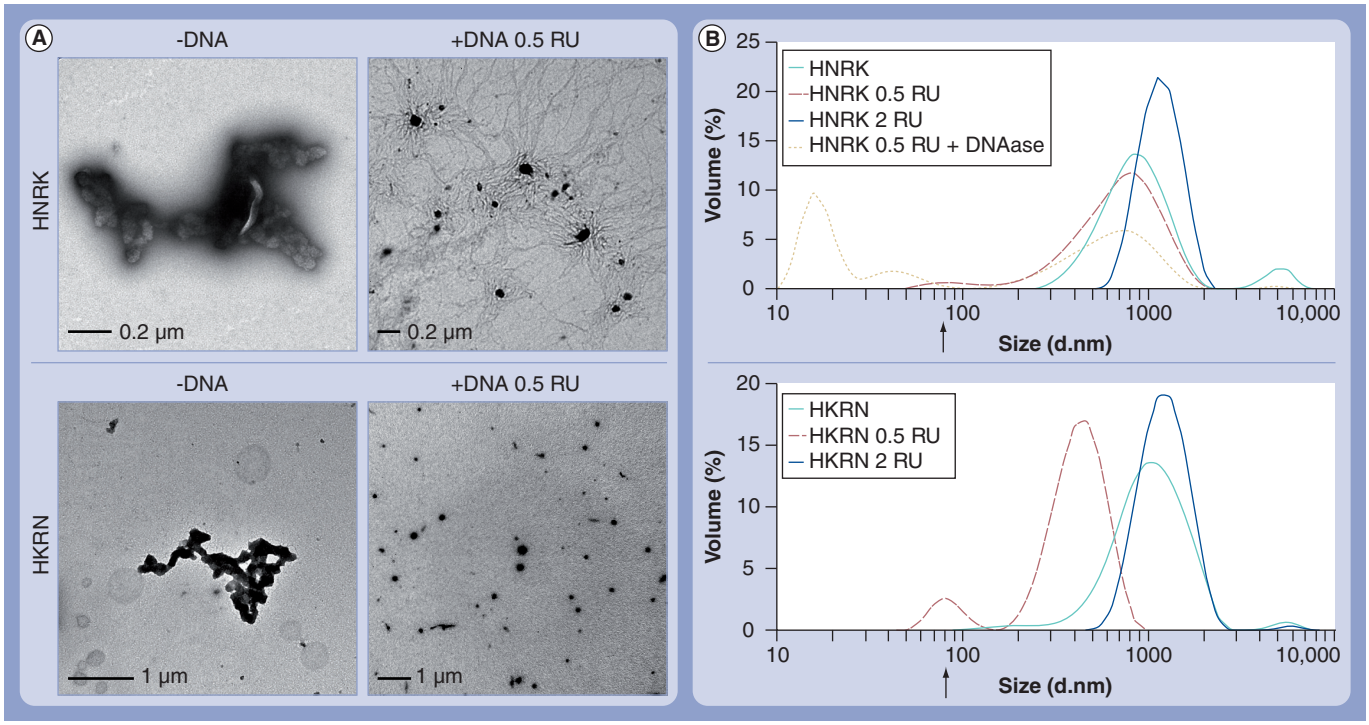

Figure 3. Size and morphology of HKRN and HNRK and their derived artificial viruses. (A) Representative transmission electron microscopy images of both HKRN and HNRK alone and as complexes with plasmid pcDNA3.1. **(B)** Effective size of protein particles alone or protein–DNA complexes (HNRK at the top and HKRN at the bottom) determined by dynamic light scattering (80 nm is marked with an arrow). The size of HNRK-based artificial viruses treated with DNAse (7 µg/ml) for 30 min at 37°C is also shown (top, spotted line).

by CD analysis of HNRK alone and combined with DNA (Figure 4E), as evidenced by the reduction of the deep valley at 200 nm in the spectrum of the peptide in solution, which corresponds to disordered structure. In addition, the CD spectra of HNRK–DNA complexes at different ratios are compatible with the presence of secondary structures, such as α-helix or antiparallel β-sheet [44], which is particularly supported by the rising of peaks between 210 and 220 nm (Figure 4E). These results are in agreement with the FTIR spectra and with the compact nanoparticulated protein–DNA structures observed by TEM.

In summary, the FTIR analysis discarded any IB-like organization of artificial viruses and both FTIR and CD spectra demonstrated that the architecture of these particles is not based on cross-molecular protein–protein contacts but that it is instead supported by charge-dependent, but potentially stereospecific DNA–protein interactions. These contacts generate artificial viruses able to transfect expressible DNA, with morphologies and sizes within the nanoscale and compatible with those found optimal for efficient cell interaction and further uptake (in the range of those exhibited by natural virus particles) [45–47]. A further evidence of the architectonic role of DNA in the organization of artificial viruses is that, upon treatment with DNAse, the HNRK-based artificial viruses disassemble in smaller entities whose lower range sizes (up to ~10 and ~40 nm), are compatible with those of peptide oligomers (Figure 3B).

Discussion

Artificial viruses are manmade constructs designed to mimic viral activities important for the cell-targeted delivery of therapeutic nucleic acids [7], and represent safer alternatives to viral gene therapy [2,6]. Lipids and polysaccharides with different molecular organizations are commonly used to protect nucleic acids that remain embedded in the core of the particle. However, because of the ability of proteins to interact with specific ligands, these vehicles are often functionalized with antibodies, peptides or whole proteins in an attempt to target a given cell type or tissue. Although tissue targeting in drug delivery can also be effectively achieved by distally applying magnetic force on paramagnetic drug carriers [48], the versatility of protein engineering offers unique opportunities for the fine tailoring of the biological properties of artificial viruses to attain, for instance, complex biodistribution maps.

In the context of the tunable nature of proteins, artificial viruses can be efficiently constructed by uniquely using these macromolecules,

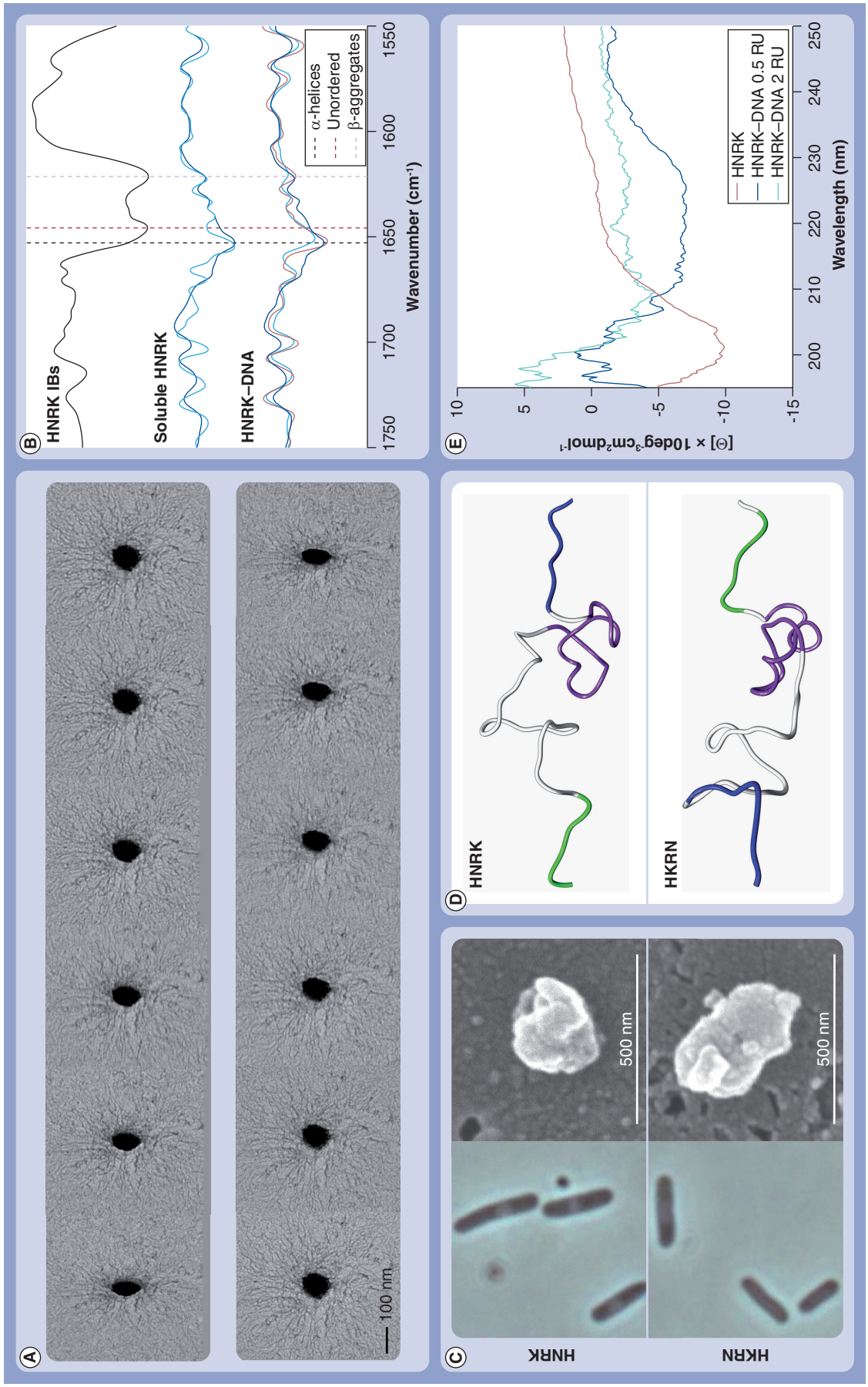

Figure 4. Molecular organization and conformation of HKRN and HNRK building blocks. (A) Selection of TEM micrograph images from a series of images acquired at inclinations from -60° to +60°. **(B)** Second derivative attenuated total reflectance-Fourier transformed infrared spectroscopy absorption spectra in the Amide I region ranging from 1750 to 1550 cm^{-1}. Vertical lines correspond to aggregated β-strands, unordered structures and α-helices peaking at 1621, 1648 and 1653 cm^{-1}, respectively. Top: second derivative spectrum of HNRK inclusion bodies; middle: second derivative of lyophilized (solid dark blue line) and soluble HNRK (solid light blue line); bottom: second derivative spectra of lyophilized HNRK (solid dark blue line), HNRK 2 RU (solid red line) and HNRK 0.5 RU (solid blue line). **(C)** HNRK- and HKRN-producing *Escherichia coli* cells showing cytoplasmic inclusion bodies (left), and those inclusion bodies as observed by SEM upon purification (right). **(D)** HNRK and HKRN models in which the different modules are colored according to the color pattern of Figure 1A. Note that the His segments are not shown. **(E)** Circular dicroism spectra of HNRK in solution and combined with DNA at two different RU. Spectra of the peptide–DNA complexes were obtained after subtracting the spectra of DNA alone recorded at the corresponding concentrations.

provided all the functions required by nucleic acid condensation and intracellular delivery are embraced. In this regard, an intriguing approach to protein-based artificial viruses is the design of multifunctional recombinant proteins [12], which contain, in a single polypeptide chain, functional peptides from different origins. By appropriate peptide selection and combination as functional modules, these units confer cellular specificity and intracellular traffic to the DNA–protein complexes [13,14]. Those functional peptides can either be inserted in permissive sites of a scaffold protein, or sequentially fused as a new, non-natural peptide or short protein [49] and produced in recombinant microorganisms. Examples of constructs generated in bacteria by these alternative strategies can be found elsewhere [12]. Importantly, bacterially produced macromolecules are biocompatible, as proved by the high number of protein drugs approved for human therapy obtained in *E. coli* [50] (even being not a 'generally recognized as safe' [GRAS] organism), and also by the wide spectrum of bacterial materials used in classic and emerging medicines [51]. Therefore, the exploration of protein particles derived from bacterially produced components is perfectly reasonable regarding their potential clinical applicability.

From the material science point of view, the organization of protein-based cages has been classified according to rather general schemes [9,52], but the precise architecture of proteinaceous artificial viruses other than those based on VLPs remains poorly explored. In fact, multifunctional proteins based on large scaffold proteins such as *E. coli* β-galactosidase for instance [53,54], organize as amorphous polydisperse protein clusters whose properties seem to be defined by protein features (the enzyme is a tetramer of approximately 460 kDa [55]), rather than by the presence of DNA [56]. Upon addition, plasmid DNA does not modify the morphology of the complexes. In the same context, arginine-rich peptides, when displayed on the surface of a chimerical green fluorescent protein, provide self-assembling properties to the fusion protein (rendering planar 20-nm particles) also irrespectively of the presence of DNA [15].

Here we have explored the nanoscale organization of two short multifunctional proteins, namely HKRN and HNRK (Figure 1), which are shown to be competent in gene delivery by using both cultured cell lines and primary cell culture models (Figure 2A & 2B). The transgene expression levels and stability that were reached in this study were comparable or higher than those observed with previous prototypes of artificial viruses based on multifunctional proteins [15,53,54,56,57], The less active modular protein version, namely the construct HKRN, achieved approximately 18% of the expression level observed when using lipofectamine (Figure 2C). The slight differences in the ability to retain and deliver expressible DNA are obviously due to the alternative disposition of functional motifs, and the end terminal location of the cationic K10 peptide seems to be especially convenient for the performance of the whole vehicle. However, apart from such a punctual observation, no dramatic differences in the performance of HKRN and HNRK have been observed. This is indicative of an important extent of functional independence of the diverse modules composing the building block, which seems to be hardly affected by their particular position in the fusion peptide and also by the surrounding partner motifs. The mere sequential fusion of functional domains without any scaffolding protein seemed a favorable strategy regarding productivity in bacteria, when comparing with the moderate yield in which high molecular mass-engineered β-galactosidases had been obtained previously [53,56].

The building blocks alone tend to passively aggregate as amorphous clusters with average sizes of approximately 1 µm (Figure 3). However, the presence of DNA dramatically modifies the organization of the protein, and at 0.5 RU it induces the formation of protein–DNA nanoparticles of approximately 80 nm from which DNA molecules eventually overhang (Figures 3A & 4A). These artificial viruses, having optimal size regarding their potential interaction with mammalian cells and further uptake [45], are able to promote the transgene expression in targeted cultured cells, as observed by several models (Figure 2), again more efficiently than amorphous vehicles based on larger scaffold proteins [53,56].

Interestingly, the organization of HKRN–DNA and HNRK–DNA complexes is not dependent on protein–protein interactions but on the sticky, glue-like potential of DNA (Figures 3B & 4B), that seems to show avidity for the internal α-helix exhibited by both proteins (Figure 4D). The architectonic properties of DNA in creating regular nanoparticles, based on charge-dependent interactions [25], strongly depend on the protein–DNA ratio (Figure 3B) and are probably more apparent when interacting with short peptides than with large proteins, as no DNA-induced architectonic changes in larger protein building shells have been previously reported [15,56]. In this context,

the particle size (80 nm) observed here by using two short chimerical proteins has resulted very similar to that observed when associating other short peptides with plasmid DNA, namely in adenoviral core peptide μ–DNA complexes (80–120 nm [58,59]) and in intermediates in toroid formation by histidylated poly-Lys–DNA complexes (80–100 nm [60]). Poly-Lys–DNA and polyornithine–DNA polyplexes have rendered, however, slightly larger particles (150–200 nm [61]). These organizing forces are probably dependent on the ability of DNA to alter the conformation of the shell proteins (Figure 4). In this context, it has been previously proved, by elegant analysis that short peptides affect the local, distal secondary and tertiary structure of bound DNA [62], but according to the data presented here the conformational changes in protein–DNA artificial viruses are mutually induced.

In the context of multifunctional large proteins, we have previously shown that multifunctional recombinant vehicles for DNA delivery efficiently induce the *in vivo* expression of a reporter [19] and a therapeutic gene [18], followed by reduced infarct volume and functional recovery of treated animals [18], in a model of acute brain injury. Interestingly, the functional modules present in the protein shell can contribute, in synergy with the therapeutic gene, to the clinical recovery of the treated animals [57]. Being clearly efficient in local administration, further *in vivo* experiments are needed to evaluate the potential of the proposed strategy for artificial virus construction in systemic gene therapy protocols, and how the protein–DNA complexes could be adapted to escape from the reticuloendothelial system.

Irrespective of that, the nanometric organizing abilities of DNA-multifunctional protein complexes, reported for the first time in this study, opens intriguing possibilities for the design and development of improved artificial viruses. The small size of the protein counterpart facilitates the DNA-mediated particle self organization, through interactions with the cationic protein motif. The functional plasticity of the multifunctional protein approach, combined with the particle size adjustment should permit the generation of chemically hybrid and improved bionanoparticles for gene therapy but also conventional drug delivery.

Conclusion

We have biologically produced short, mainly unordered multifunctional peptides as building blocks of protein-based artificial viruses, which have shown an excellent performance in transgene delivery under different biological models. Interestingly, the artificial viruses resulting from protein–DNA associations are pseudo-spherical entities with regular particle sizes of approximately 80 nm, at specific protein–DNA ratios in the range of those promoting high transgene expression levels. A structural characterization of the protein components in these artificial viruses has revealed that the global architecture of the particles is not driven by protein–protein interactions but on the contrary, unexpectedly supported by the embedded DNA. The nucleic acids act as a compacting, molecular glue that affects the conformation of the protein building blocks, altering the α-helix structure of the central region, minimizing their aggregation tendency and promoting an ordered, self organization of the complexes in sizes compatible with an efficient receptor-mediated cell uptake and proper intracellular trafficking to the cell nucleus. This first description of the architectonic properties of DNA at the nanoscale opens intriguing opportunities for a better rational design of artificial viruses for gene therapy regarding their molecular and physical organization.

Future perspective

A better comprehension of the DNA–protein and protein–protein interactions in the context of nanoparticles for drug and DNA delivery (at this moment a rather neglected area) should offer new engineering tools for the semi-rational or rational tuning of the nanoscale properties of artificial viruses, which is expected to fully expand in the next decade. The incorporation of protein-only vehicles (other than VLPs) in the nanomedical scenario will offer intriguing possibilities for the flexible development of smart drugs, especially when applying modular/multifunctional protein engineering principles. However, the biocompatibility and safety of protein-based nanoparticles should be combined with enhanced stability and improved targeting, the major challenges in the immediate generation of powerful drugs at the clinical level.

Acknowledgements

The authors are indebted to the Cell Culture and Citometry Units of the Servei de Cultius Cellulars, Producció d'Anticossos i Citometria (SCAC), and to the Servei de Microscòpia, both at the Universitat Autònoma de Barcelona, and to Agustín Correa (UPR) and Gonzalo Obal (UBP) from the Institute Pasteur de Montevideo for their excellent technical help. Antonio Villaverde has been distinguished with an ICREA ACADEMIA award.

Financial & competing interests disclosure

The authors appreciate the financial support received for the design and production of artificial viruses for gene therapy from MICINN (BIO2007-61194 and ACI2009-0919), AGAUR (2009SGR-108) and CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN, Spain), an initiative funded by the VI National R+D+i Plan 2008–2011, Iniciativa Ingenio 2010, Consolider Program, CIBER Actions and financed by the Instituto de Salud Carlos III with assistance from the European Regional Development Fund. We also thank ANII, Ministerio de Educación y Cultura, Uruguay, for financial support. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

- Short chimerical proteins produced in bacteria, that contain four functional domains relevant to intracellular trafficking, promote high transgene expression levels when used as artificial viruses.
- The presence of DNA promotes conformational changes in the protein moiety of the artificial viruses that affects the minor α-helix region exhibited by rather unstructured peptides.
- The resulting artificial viruses are pseudo-spherical stable particles of approximately 80 nm, fully sustained by DNA–protein interactions rather than by protein–protein cross-molecular β-sheet interactions, which at difference from protein-only aggregates, are undetectable.

Bibliography

Papers of special note have been highlighted as:
▪ of interest
▪▪ of considerable interest

1 Smaglik P. Clinical trials end at gene-therapy institute. *Nature* 405(6786), 497 (2000).

2 Edelstein ML, Abedi MR, Wixon J. Gene therapy clinical trials worldwide to 2007 an update. *J. Gene Med.* 9(10), 833–842 (2007).

3 Marshall E. Gene therapy. Second child in French trial is found to have leukemia. *Science* 299(5605), 320 (2003).

4 Marshall E. Gene therapy. What to do when clear success comes with an unclear risk? *Science* 298(5593), 510–511 (2002).

5 Smaglik P. Tighter watch urged on adenoviral vectors ... with proposal to report all 'adverse events'. *Nature* 402(6763), 707 (1999).

6 Douglas KL. Toward development of artificial viruses for gene therapy: a comparative evaluation of viral and non-viral transfection. *Biotechnol. Prog.* 24(4), 871–883 (2008).

7 Mastrobattista E, van der Aa MA, Hennink WE, Crommelin DJ. Artificial viruses: a nanotechnological approach to gene delivery. *Nat. Rev. Drug Discov.* 5(2), 115–121 (2006).

▪▪ **The basic principles of artificial virus design and construction are reviewed, describing the diversity of structural approaches and chemical composition of the building blocks.**

8 Wagner E. Strategies to improve DNA polyplexes for *in vivo* gene transfer: will "artificial viruses" be the answer? *Pharm. Res.* 21(1), 8–14 (2004).

9 Maham A, Tang Z, Wu H, Wang J, Lin Y. Protein-based nanomedicine platforms for drug delivery. *Small* 5(15), 1706–1721 (2009).

▪ **The spectrum of self-assembling protein materials with potential as platforms for drug delivery is extensively reviewed.**

10 Petry H, Goldmann C, Ast O, Luke W. The use of virus-like particles for gene transfer. *Curr. Opin. Mol. Ther.* 5(5), 524–528 (2003).

11 Wu M, Sherwin T, Brown WL, Stockley PG. Delivery of antisense oligonucleotides to leukemia cells by RNA bacteriophage capsids. *Nanomedicine (Lond.)* 1(1), 67–76 (2005).

12 Aris A, Villaverde A. Modular protein engineering for non-viral gene therapy. *Trends Biotechnol.* 22(7), 371–377 (2004).

13 Vazquez E, Ferrer-Miralles N, Villaverde A. Peptide-assisted traffic engineering for nonviral gene therapy. *Drug Discov. Today* 13(23–24), 1067–1074 (2008).

14 Ferrer-Miralles N, Vazquez E, Villaverde A. Membrane-active peptides for non-viral gene therapy: making the safest easier. *Trends Biotechnol.* 26(5), 267–275 (2008).

15 Vazquez E, Roldán M, Diez-Gil C *et al.* Protein nanodisk assembling and intracellular trafficking powered by an arginine-rich (R9) peptide. *Nanomedicine (Lond.)* 5, 259–268 (2010).

16 Vazquez E, Cubarsi R, Unzueta U *et al.* Internalization and kinetics of nuclear migration of protein-only, arginine-rich nanoparticles. *Biomaterials* 31(35), 9333–9339 (2010).

▪▪ **The internalization pattern and nuclear migration routes of a self-assembling, protein-only nanoparticle based on arginine-rich peptides are finely dissected.**

17 Xavier J, Singh S, Dean DA, Rao NM, Gopal V. Designed multi-domain protein as a carrier of nucleic acids into cells. *J. Control Release* 133(2), 154–160 (2009).

▪ **The potential of a protein-only, Tat-derived modular construct for gene therapy is clearly shown through elegant experiments.**

18 Peluffo H, Acarin L, Aris A *et al.* Neuroprotection from NMDA excitotoxic lesion by Cu/Zn superoxide dismutase gene delivery to the postnatal rat brain by a modular protein vector. *BMC Neurosci.* 7, 35 (2006).

19 Peluffo H, Aris A, Acarin L, Gonzalez B, Villaverde A, Castellano B. Nonviral gene delivery to the central nervous system based on a novel integrin-targeting multifunctional protein. *Hum. Gene Ther.* 14(13), 1215–1223 (2003).

20 Benito A, Mateu MG, Villaverde A. Improved mimicry of a foot-and-mouth disease virus antigenic site by a viral peptide displayed on β-galactosidase surface. *Biotechnol. NY* 13(8), 801–804 (1995).

21 Feliu JX, Benito A, Oliva B, Aviles FX, Villaverde A. Conformational flexibility in a highly mobile protein loop of foot-and-mouth disease virus: distinct structural requirements for integrin and antibody binding. *J. Mol. Biol.* 283(2), 331–338 (1998).

22 Villaverde A, Feliu JX, Harbottle RP, Benito A, Coutelle C. A recombinant, arginine–glycine–aspartic acid (RGD)

motif from foot-and-mouth disease virus binds mammalian cells through vitronectin and, to a lower extent, fibronectin receptors. *Gene* 180(1–2), 101–106 (1996).

23 Yoneda Y. How proteins are transported from cytoplasm to the nucleus. *J. Biochem.* 121(5), 811–817 (1997).

24 Collins E, Birchall JC, Williams JL, Gumbleton M. Nuclear localisation and pDNA condensation in non-viral gene delivery. *J. Gene Med.* 9(4), 265–274 (2007).

25 Saccardo P, Villaverde A, Gonzalez-Montalban N. Peptide-mediated DNA condensation for non-viral gene therapy. *Biotechnol. Adv.* 27(4), 432–438 (2009).

26 Block H, Maertens B, Spriestersbach A *et al.* Immobilized-metal affinity chromatography (IMAC): a review. *Methods Enzymol.* 463, 439–473 (2009).

27 Garcia-Fruitos E, Seras-Franzoso J, Vazquez E, Villaverde A. Tunable geometry of bacterial inclusion bodies as substrate materials for tissue engineering. *Nanotechnology* 21(20), 205101 (2010).

28 Eswar N, Marti-Renom MA, Webb B *et al.* Comparative protein structure modeling with MODELLER. John Wiley & Sons, Inc., NJ, USA (Suppl. 15), 5.6.1–5.6.30 (2009).

29 Fontes MR, Teh T, Toth G *et al.* Role of flanking sequences and phosphorylation in the recognition of the simian-virus-40 large T-antigen nuclear localization sequences by importin-α. *Biochem. J.* 375(Pt 2), 339–349 (2003).

30 Hewat EA, Verdaguer N, Fita I *et al.* Structure of the complex of an Fab fragment of a neutralizing antibody with foot-and-mouth disease virus: positioning of a highly mobile antigenic loop. *EMBO J.* 16(7), 1492–1500 (1997).

31 Davidson B, Fasman GD. The conformational transitions of uncharged poly-L-lysine. α helix-random coil-β structure. *Biochemistry* 6(6), 1616–1629 (1967).

32 Pakhomova ON, Deep S, Huang Q, Zwieb C, Hinck AP: Solution structure of protein SRP19 of *Archaeoglobus fulgidus* signal recognition particle. *J. Mol. Biol.* 317(1), 145–158 (2002).

33 Chen J, Lee KH, Steinhauer DA, Stevens DJ, Skehel JJ, Wiley DC. Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. *Cell* 95(3), 409–417 (1998).

34 Petegnief V, Friguls B, Sanfeliu C, Sunol C, Planas AM. Transforming growth factor-α attenuates *N*-methyl-D-aspartic acid toxicity in cortical cultures by preventing protein synthesis inhibition through an Erk1/2-dependent mechanism. *J. Biol. Chem.* 278(32), 29552–29559 (2003).

35 Garcia-Fruitos E, Rodriguez-Carmona E, Diez-Gil C *et al.* Surface cell growth engineering assisted by a novel bacterial nanomaterial. *Adv. Mater.* 21(42), 4249–4253 (2009).

36 de Marco A , Schroedel A. Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem.* 6(1), 10 (2005).

37 Martinez-Alonso M, Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A. The functional quality of soluble recombinant polypeptides produced in *Escherichia coli* is defined by a wide conformational spectrum. *Appl. Environ. Microbiol.* 101(6), 1353–1358 (2008).

38 Doglia SM, Ami D, Natalello A, Gatti-Lafranconi P, Lotti M. Fourier transform infrared spectroscopy analysis of the conformational quality of recombinant proteins within inclusion bodies. *Biotechnol. J.* 3(2), 193–201 (2008).

39 Gonzalez-Montalban N, Natalello A, Garcia-Fruitos E, Villaverde A, Doglia SM. *In situ* protein folding and activation in bacterial inclusion bodies. *Biotechnol. Bioeng.* 100(4), 797–802 (2008).

40 Ami D, Natalello A, Taylor G, Tonon G, Maria DS. Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy. *Biochim. Biophys. Acta* 1764(4), 793–799 (2006).

41 Wang L, Maji SK, Sawaya MR, Eisenberg D, Riek R. Bacterial inclusion bodies contain amyloid-like structure. *PLoS Biol.* 6(8), e195 (2008).

42 Carrio M, Gonzalez-Montalban N, Vera A, Villaverde A, Ventura S. Amyloid-like properties of bacterial inclusion bodies. *J. Mol. Biol.* 347(5), 1025–1037 (2005).

43 Natalello A, Ami D, Brocca S, Lotti M, Doglia SM. Secondary structure, conformational stability and glycosylation of a recombinant *Candida rugosa* lipase studied by Fourier-transform infrared spectroscopy. *Biochem. J.* 385(Pt 2), 511–517 (2005).

44 Greenfield NJ. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* 1(6), 2527–2535 (2006).

45 Jiang W, Kim BY, Rutka JT, Chan WC. Nanoparticle-mediated cellular response is size-dependent. *Nat. Nanotechnol.* 3(3), 145–150 (2008).

▪▪ **The size effect of nanoparticles on the mammalian cell biology is demonstrated in biological interfaces.**

46 Chithrani BD, Ghazani AA, Chan WC. Determining the size and shape dependence of gold nanoparticle uptake into mammalian cells. *Nano Lett.* 6(4), 662–668 (2006).

47 Jiang W, Kim BY, Rutka JT, Chan WC. Advances and challenges of nanotechnology-based drug delivery systems. *Expert. Opin. Drug Deliv.* 4(6), 621–633 (2007).

48 Corchero JL , Villaverde A. Biomedical applications of distally controlled magnetic nanoparticles. *Trends Biotechnol.* 27(8), 468–476 (2009).

49 Vazquez E, Ferrer-Miralles N, Mangues R, Corchero JL, Schwartz S Jr, Villaverde A. Modular protein engineering in emerging cancer therapies. *Curr. Pharm. Des* 15(8), 893–916 (2009).

50 Ferrer-Miralles N, Domingo-Espin J, Corchero JL, Vazquez E, Villaverde A. Microbial factories for recombinant pharmaceuticals. *Microb. Cell Fact.* 8(1), 17 (2009).

51 Rodriguez-Carmona E, Villaverde A. Nanostructured bacterial materials for innovative medicines. *Trends Microbiol.* 18(9), 423–430 (2010).

52 Uchida M, Klem MT, Allen M *et al.* Biological containers: protein cages as multifunctional nanoplatforms. *Adv. Mater.* 19, 1025–1042 (2007).

53 Aris A, Feliu JX, Knight A, Coutelle C, Villaverde A. Exploiting viral cell-targeting abilities in a single polypeptide, non-infectious, recombinant vehicle for integrin-mediated DNA delivery and gene expression. *Biotechnol Bioeng* 68(6), 689–696 (2000).

54 Aris A , Villaverde A. Engineering nuclear localization signals in modular protein vehicles for gene therapy. *Biochem. Biophys. Res. Commun.* 304(4), 625–631 (2003).

55 Jacobson RH, Zhang XJ, DuBose RF, Matthews BW. Three-dimensional structure of β-galactosidase from *E. coli*. *Nature* 369(6483), 761–766 (1994).

56 Aris A, Villaverde A. Molecular organization of protein–DNA complexes for cell-targeted DNA delivery. *Biochem. Biophys. Res. Commun.* 278(2), 455–461 (2000).

57 Peluffo H, Gonzalez P, Aris A *et al.* RGD domains neuroprotect the immature brain by a glial-dependent mechanism. *Ann. Neurol.* 62(3), 251–261 (2007).

58 Tagawa T, Manvell M, Brown N *et al.* Characterisation of LMD virus-like nanoparticles self-assembled from cationic liposomes, adenovirus core peptide mu and plasmid DNA. *Gene Ther.* 9(9), 564–576 (2002).

59 Keller M, Tagawa T, Preuss M, Miller AD. Biophysical characterization of the DNA

binding and condensing properties of adenoviral core peptide μ. *Biochemistry* 41(2), 652–659 (2002).

60 Midoux P, LeCam E, Coulaud D, Delain E, Pichon C. Histidine containing peptides and polypeptides as nucleic acid vectors. *Somat. Cell Mol. Genet.* 27(1–6), 27–47 (2002).

61 Ramsay E, Hadgraft J, Birchall J, Gumbleton M. Examination of the biophysical interaction between plasmid DNA and the polycations, polylysine and polyornithine, as a basis for their differential gene transfection *in-vitro*. *Int. J. Pharm.* 210(1–2), 97–107 (2000).

62 Tecle M, Preuss M, Miller AD. Kinetic study of DNA condensation by cationic peptides used in nonviral gene therapy: analogy of DNA condensation to protein folding. *Biochemistry* 42(35), 10343–10347 (2003).

- **The states of DNA condensation during interaction with cationic peptides have been analyzed, revealing the obstacles posed to DNA transcription at high peptide–DNA ratios.**