Development and characterization of artificial viruses for gene therapy

Joan Domingo i Espín

UAB
Universitat Autònoma de Barcelona

Development and characterization of artificial viruses for gene therapy

Tesi doctoral

Joan Domingo i Espín

Institut de Biotecnologia i de Biomedecina
Facultat de Biociències · Departament de Genètica i de Microbiologia
Universitat Autònoma de Barcelona

Universitat Autònoma de Barcelona

Development and characterization of artificial viruses for gene therapy

Memòria presentada per en Joan Domingo i Espín
per optar al grau de Doctor en Biotecnologia
per la Universitat Autònoma de Barcelona

Vist-i-plau de les directores de la tesi

Dra. Neus Ferrer i Miralles

Dra. Esther Vásquez Gómez

Institut de Biotecnologia i de Biomedecina
Facultat de Biociències · Departament de Genètica i de Microbiologia
Universitat Autònoma de Barcelona

Joan Domingo I Espín

Bellaterra, maig de 2013

Zuretzat, badakizu nor zaren

A la meva família

Index

Introduction

Gene Therapy

Definition and challenges

A new chapter in medicine began with the mapping of the human genome and the development of recombinant DNA techniques. Both advances led to the birth of gene therapy which is based in the use of nucleic acids as a pharmaceutical agent to treat, cure or prevent diseases.

Several definitions of gene therapy have been specified, sometimes with only subtle differences. The combination of the following two definitions shapes the best consensus, which will be the definition of gene therapy in this thesis.

The American Society of Gene Therapy (ASGT) has formulated a clear scientific definition of gene therapy [1]:

> *"Gene therapy is defined as a set of strategies that modify the expression of an individual's genes or that correct abnormal genes. Each strategy involves the administration of a specific DNA (or RNA)."*

An equally important regulatory definition is formulated by the European Medicines Agency (EMA)[2]:

> *"Gene therapy (…) has the following characteristics: (a) it (…) contains or consists of a recombinant nucleic acid used in or administered to human beings with a view to regulating, repairing, replacing, adding or deleting a genetic sequence; (b) its therapeutic (…) effect relates directly to the recombinant nucleic acid sequence it contains, or to the product of genetic expression of this sequence."*

The most common form of gene therapy involves using DNA that encodes a functional therapeutic gene in order to replace a mutated gene. Other forms involve directly correcting a mutation, using DNA that encodes a therapeutic protein drug

(rather than a natural human gene) to provide treatment, lowering gene expression by using decoy DNA for transcription factors, using RNA molecules to block transcription or translation of a “toxic” protein or an infectious agent or even a DNA that encodes for a specific RNA molecule. In summary, any nucleic acid (NA) can be used as a therapeutic agent [3].

In gene therapy, the nucleic acids (NAs) to be delivered are packaged within a "vector", which is used to get inside cells in the body and to avoid the degradation of NAs, when they are naked, by nucleases. Once inside, the cell machinery helps the NA to reach its corresponding site, resulting in the production of therapeutic effect, which in turn treats the patient's disease.

The future of gene therapy holds promises in the treatment of many incurable diseases. The problem with gene therapy is that, although the treatment sounds promising theoretically, however, is difficult to implement. Gene therapy involves a whole lot of complicated set of activities involving tissue targeting, cellular trafficking, delivery of genes to organs, safety of the vector, activity of therapeutic protein, etc. The above mentioned activities being understood incompletely, the progress in the field of gene therapy seems to be slow.

For gene therapy to be implemented, certain prerequisites or basic issues are mandatory. Therapeutically suitable genes should be available for the treatment. For the delivery of the genes, vectors are needed. These may be viral or non-viral vectors. For the therapy to be successful, the efficacy and safety of the method should be determined or established in advance.

Despite the current advances and future research in the field of gene therapy, it should be understood that there will not be a standard cure for diseases. The different nature of different diseases makes it necessary to overcome new technical and medical hurdles for every new disease.

The pace at which technology is improving and the knowledge about the body systems is being accumulated, the future of gene therapy looks bright.

Historical view of gene therapy

Gene therapy history is marked by success but also by big failures showing that technically is still risky and more studies have to be made to optimise safety and efficacy of products.

In 1963, Edward Tatum and Joshua Lederberg stated the first description of what can be considered gene therapy nowadays[4]:

> *"We might anticipate the in vitro culture of germ cells and such manipulations as the interchange of chromosomes and DNA segments. The ultimate application of molecular biology would be the direct control of nucleotide sequence in human chromosomes. Coupled with recognition, selection and integration of desired genes… We anticipate that viruses will be used effectively for man's benefit…"*

Less than one decade later, Rogers et al. attempted the first clinical trial with gene therapy[5]. Although without success, the benchmark for clinical trials with gene therapy was set[6–8]. In 1990, Anderson et al. showed signs of efficacy in what seemed the first successful gene therapy trial on adenosine deaminase deficient female patients[9]. In the 1990s, gene therapy developed fast, the field grew and many clinical trials were conducted. Around the turn of the century, clinical trials took a turn for the worse, with the death of an eighteen-year-old boy, called the "Gelsinger case" due to a massive immune response triggered by the use of the viral vector used to transport the gene into his cells. Another misfortune started in 2002, when a success of a multi-centre trial for treating children with SCID (severe combined immune deficiency or "bubble boy" disease) held from 2000 and 2002 was questioned when two of the ten

children treated at the trial's Paris centre developed a leukemia-like condition due to insertional mutagenesis near a proto-oncogene[10–12]. Clinical trials were halted temporarily in 2002, but resumed after regulatory review of the protocol in the United States, the United Kingdom, France, Italy, and Germany[13]. Meanwhile, China took the lead and in 2003 launched the first commercial gene therapy product: Gendicine[14,15]. Although early clinical failures led many to dismiss gene therapy as over-hyped, clinical successes in 2006-2011 have boostered new optimism in the promise of gene therapy. These include successful treatment of patients with the retinal disease Leber's congenital amaurosis [16–19], X-linked SCID[20], ADA-SCID[21], adrenoleukodystrophy[22], chronic myelogenous leukemia (CLL)[23], and Parkinson's disease[24]. These recent clinical successes have led to a renewed interest in gene therapy, with several articles in scientific and popular publications calling for continued investment in the field[25].

In 2012, the FDA has approved clinical trials for the use of gene therapy on thalassemia major patients in the US. Researchers at Memorial Sloan Kettering Cancer Center in New York recruited 10 participants for the study in July 2012[26]. The study is expected to end in 2014[27].

In July 2012, the European Medicines Agency recommended approval of a gene therapy treatment for the first time in either Europe or the United States. The treatment, called Glybera, compensates for lipoprotein lipase deficiency[28]. The recommendation needs to be endorsed by the European Commission before being made available to patients[29].

Delivery systems in gene therapy

It is well accepted that the bottleneck in the use of gene therapy and its application to human diseases is the development of the safe and efficient vehicles carrying the genetic cargo.

Due to the expectations that gene therapy have given to the scientific and non-scientific community, it will need to compete for medical and patient acceptance, against conventional

therapies, in the marketplace. Products for gene therapy must be developed, validated, and manufactured to meet the standards for approval by regulatory authorities and manufacturing companies. They should be similar in nature and mode of administration to conventional therapeutics and the ideal vehicle would satisfy the following requirements[30]:

1. Easy and cheap production

The vector should be easy to produce at high titre on a commercial scale and at relatively low costs. This consideration is important for the wide range of cell numbers that must be transduced and for the commercial benefit. For widespread use, the vector should be amenable to commercial production and processing (such as concentration technology for delivery in small volumes), and should have a reasonable shelf life for transport and distribution.

2. Sustained expression

The vector, once delivered, should be able to express its genetic cargo over a sustained period or expression should be adjustable in a precise way. Different disease states have different requirements (for example, regulated expression in diabetes and lifetime expression in haemophilia).

3. Immunologically inert

The vector components should not elicit an immune response after delivery. A humoral antibody response will make a second injection of the vector ineffective, whereas a cellular response will eliminate the transduced cells. Also, a severe immunological reaction can lead even to death[31].

4. Tissue targeting

Delivery to only certain cell types is highly desirable, especially where the target cells are dispersed throughout the body (such as in the haematopoietic system), or if the cells are part of a

heterogeneous population (such as in the brain). It is also important to avoid certain cells, such as dendritic cells, the `professional' antigen-presenting cells of the body, because of their role in mediating the immune response.

5. Size capacity

The vector should have no size limit to the genetic material it can deliver. The coding sequence of a therapeutic gene can vary from 350 base pairs for insulin, to over 12,000 base pairs for dystrophin. Furthermore, addition of appropriate regulatory sequences may be required for efficient transduction and expression of the foreign genetic material.

6. Replication, segregation or integration

The vector should allow for site-specific integration of the gene into the chromosome of the target cell, or should reside in the nucleus as an episome that will faithfully divide and segregate on cell division. Site-specific integration is a very desirable attribute because it eliminates the uncertainty of random integration into the host chromosome, and endogenous regulatory regions will control its expression under physiological conditions.

7. Infection of dividing and non-dividing cells

As large numbers of cells (such as neurons, hepatocytes and myocytes) are postmitotic, vectors capable of efficiently transducing non-dividing cells are very desirable.

Viral/non-viral delivery systems

In gene therapy we can distinguish between two main types of vehicles: viral and non-viral vectors. More efforts have been made with the use of viral vectors. This is because presumptively viruses present more potential as gene therapy

agents as they have evolved during millions of years to transfer genetic cargos to target cells.

Viral vectors are based in the use of various families of modified non-replicative viruses (Table 1 summarizes the most common families of viruses used in gene therapy and their main characteristics). They are appropriate because they exploit relevant properties of the viral cycle, such as receptor mediated cell recognition and internalization, endosomal escape, nuclear transport and DNA expression. A diversity of virus families with distinct biological cycles offers interesting alternatives, regarding the size of delivered DNA, tissue tropism and duration of transgene expression, among others[32]. Interestingly, the potential of viruses as delivery agents in RNA interference (RNAi)-based therapies has been also demonstrated. Before they can be used as therapeutic agents, viral genomes must be conveniently modified to eliminate (or at least reduce as much as possible) any chance of virulence. Due to their undoubted efficiency, viruses have entered several clinical trials, with a significant degree of success. However, the therapeutic use of viruses for delivery is controversial because of the dissemination of modified but potentially replicable genomes that could integrate or recombine with cell DNA. Moreover, non-desirable, clinically important side effects in target individuals, ranging from inflammation to death, have been observed upon inoculation. Also, significant evidence suggests carrier viruses responsible for leukaemia occurrence in treated patients, via a mechanism of insertional mutagenesis and consequent gene inactivation or activation. Therefore, biosafety requirements for viral vectors are expected to become increasingly stringent. Viral vectors must also overcome important immunological barriers: immune responses triggered during previous infections, vaccination or preceding doses of delivery viruses can prevent the vehicles from reaching their target. Finally but not less important, production costs of viruses are nowadays out of the range of any widely applicable therapy.

	Adenovirus	Adeno-associated virus	Lentivirus	Retrovirus
Tropism	Dividing and non-dividing cells	Dividing and non-dividing cells	Dividing and non-dividing cells	Dividing cells
Host genome	No integration	No integration	Integration	Integration
Transgene expression	Transient	Stable	Stable	Stable
Packaging capacity	~8 kb	~5 kb	~8 kb	~8 kb
Advantages	Large packaging capacity; high production yields	High production yields; low immunogenicity; long-term expression	Large packaging capacity; long-term expression	Large packaging capacity; long-term expression
Disadvantages	High immunogenicity; transient expression	Small packaging capacity	Insertional mutagenesis	High risk of insertional mutagenesis

Table 1. Comparison of the most commonly used viruses for gene therapy (Adapted from Acta Biochim Biophys Sin 2012, 44: 632–640[33])

Accordingly, the limitations described above for viral vectors, have promoted an increasing interest in methods for gene delivery that do not require viral vectors. Technologies that allow genes to be formulated like conventional pharmaceutical products and administrated directly to patients are being developed. Non-viral gene delivery systems are often called *artificial viruses* and in many cases they emulate many biological functions of viral vectors. Nevertheless they differ from viral systems in terms of their composition, manufacture, characterization and therapeutic profile and they also differ in their clinical and commercial risks. Table 2 summarises the comparison between viral and non-viral vectors in terms of advantages and disadvantages.

Type of vectors	Advantages	Disadvantages
Viral vectors	Relatively high transfection e! iciency	- Immunogenicity, presence of contaminants and safety problems - Vector restricted size limitation for NA - Unfavorable Pharmac okinetics-large scale production, GMP, stability and cost
Non-viral vectors	- Adjustable immunogenicity - Opportunity for chemical/physical manipulation - Favorable pharmacokinetics, large scale production, GMP, stability and cost - Plasmid independent structure (size and topology)	Relatively low transfection e! iciency

Table 2. Comparison table of viral and non-viral vectors.

Types of nucleic acids to deliver

Gene therapy approaches can be divided into three types of therapeutic effects related to the NA delivered (Figure 1):

1. Restoration or addition of gene function by gene delivery;
2. Silencing of genes by antisense oligonucleotides or RNAi delivery;
3. Modification of gene function by splice-switching oligonucleotide therapeutics.

Gene Delivery

Total or partial loss of functional genes is the cause of not only several heritable diseases such as haemophilia, muscular dystrophy, and cystic fibrosis, but also more frequent disorders such as cardiovascular diseases and cancer. Thus, the delivery of functional genes that could restore normal phenotype has been the Holy Grail of gene therapy for decades. Also, adding new genes to cells that normally do not express them is a very

promising field such as for vaccination[34,35] or protein production in protein deficiency disorders[36].

Antisense

There are two possible mechanisms for an antisense effect using NAs. When the target is a nuclear double-stranded DNA or gene, the approach is called the antigene strategy[37], while the method that relies on targeting of the mRNA is called the antisense strategy. Antisense activity can be achieved either by blocking the binding sites for the 40S ribosomal subunit and for other translation initiation signals, or by the formation of a double-stranded DNA/RNA complex that renders the RNA susceptible to RNase H digestion[37].

RNAi

RNAi is a fundamental gene-silencing pathway in eukaryotic cells, where short pieces of double-stranded RNA are cleaved by an enzyme called Dicer into shorter fragments called siRNAs that can cleave complementary mRNA sequences with the help of the RISC complex and Argonaute[38]. The proof-of-principle RNAi study was reported in 2001, which demonstrated that synthetic siRNA could achieve sequence-specific gene knockdown in mammalian cells, and this marked the birth of siRNA therapeutics[39]. What makes the RNAi approach more appealing compared to the antisense strategy is that siRNA cleaves the target mRNA in a catalytic manner, so lower doses are required to achieve gene knockdown. Consequently, intensive research has been carried out in the last decade to develop delivery vectors for siRNA therapeutics[38]. It has been also proposed the use of DNA vectors that express a small hairpin RNA (shRNA) molecule to achieve the same objective. These molecules make a tight hairpin turn that can be used to silence target gene expression via RNAi[40].

Splice-Switching oligonucleotide Therapeutics

Modification of gene function can be achieved by interfering with the splicing machinery; an approach termed splice switching[41]. Recent studies using high-throughput sequencing indicate that 95–100% of human premRNAs have alternative splice forms[42]. Mutations that affect alternative premRNA splicing have been linked to a variety of diseases, including genetic disorders and cancers, and Splice Switching Oligonucleotides (SSOs) can be used to silence mutations that cause aberrant splicing, thereby restoring correct splicing and function of the defective gene[41,43]. SSOs are NAs, ranging from 15 to 25 bases in length, which do not activate RNase H and thereby prevent the cleavage of the premRNA target before it can be spliced[41,43]. One example of a genetic disease amenable for SSO is Duchenne's muscular dystrophy (DMD). DMD is a neuromuscular genetic disorder that is caused mainly by nonsense or frame-shift mutations in the dystrophin gene. SSOs are used to induce targeted ''exon skipping'' and to correct the reading frame of mutated dystrophin pre-mRNA such that shorter, partially functional dystrophin forms are produced[44].

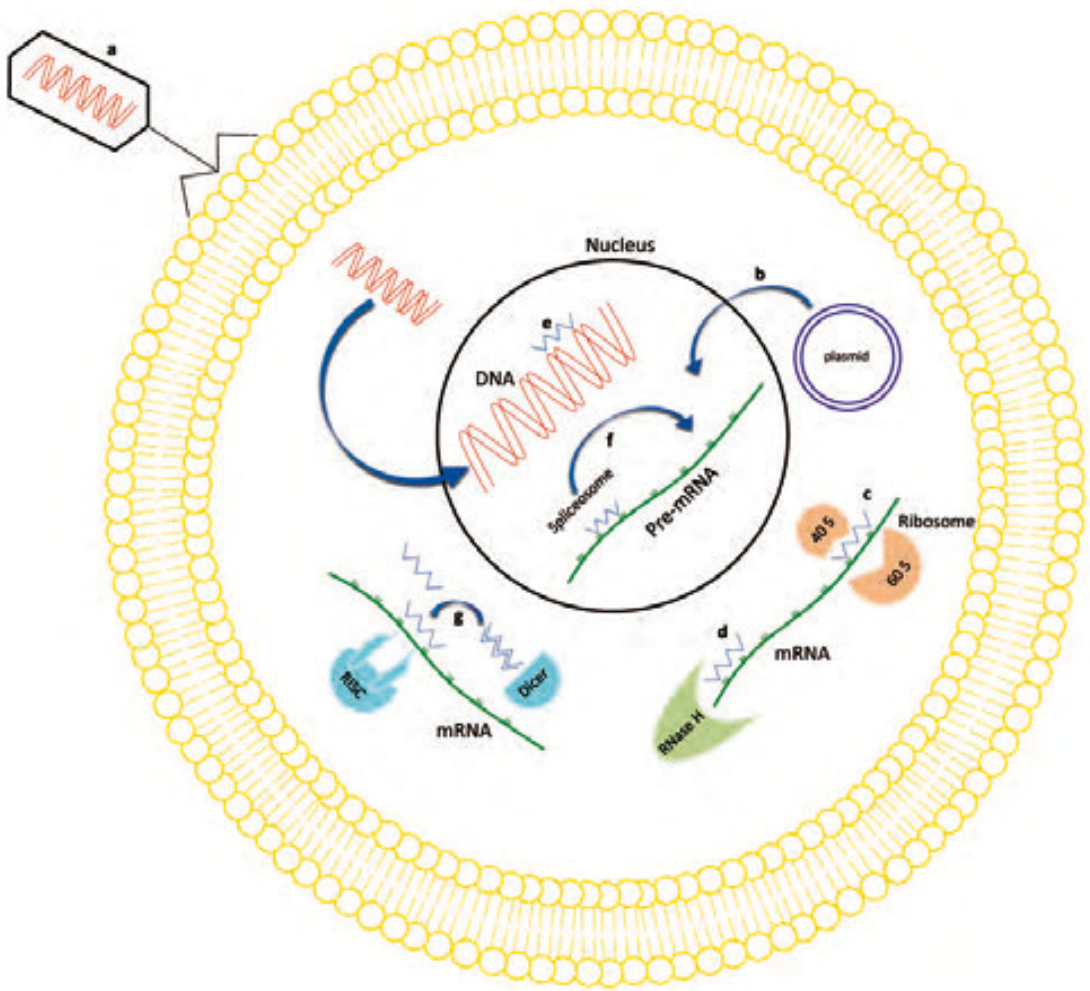

Figure 1. Types of NA to deliver, different approaches. Gene delivery: (a) Viral delivery and genome integration. (b) Plasmid delivery. Antisense: (c) Antisense steric block of translation. (d) Antisense DNA/RNA hybrid and RNase H degradation. Splice-Switching Therapeutics (f). siRNA (g). Adapted from Prog. Mol. Biol. Transl. Sci. 2011;104:397-426.[45]

Overcoming biological barriers

Biodistribution of a gene therapy vector is determined by its interaction with the body (Figure 2), based on the physicochemical properties of the vehicle and the outcome of the different interactions with biological components. Therefore, targeting can be achieved by altering the properties of the delivery system[46–48]. For low molecular weight vectors as well as biologically active proteins, drug targeting has been achieved by controlling the physicochemical properties of the gene–carrier complex (or conjugate) such as the particle size,

molecular weight, and electrical charge, and/or by using a specific ligands such as monoclonal antibody and cell ligands. However, plasmid DNA or any other nucleic acid (NA) are large polyanions and form electrostatic complexes with cationic non-viral vectors, and there are many complex factors influencing the biodistribution of the administered NA Therefore, efficient and target-specific gene transfer is difficult to achieve. The route of administration determines the number of barriers that should be overcome for successful in vivo gene transfer by a non-viral vector. For example, if the NA is bombarded into the tissue by a gene gun system, DNA is directly delivered to the cytoplasm, or even into the nucleus of the target cell[49]. In addition to these delivery barriers, an ideal non-viral vector and its carried NA should be stable in test tubes as well as in the body, biodegradable, nontoxic, and non-immunogenic[50,51].

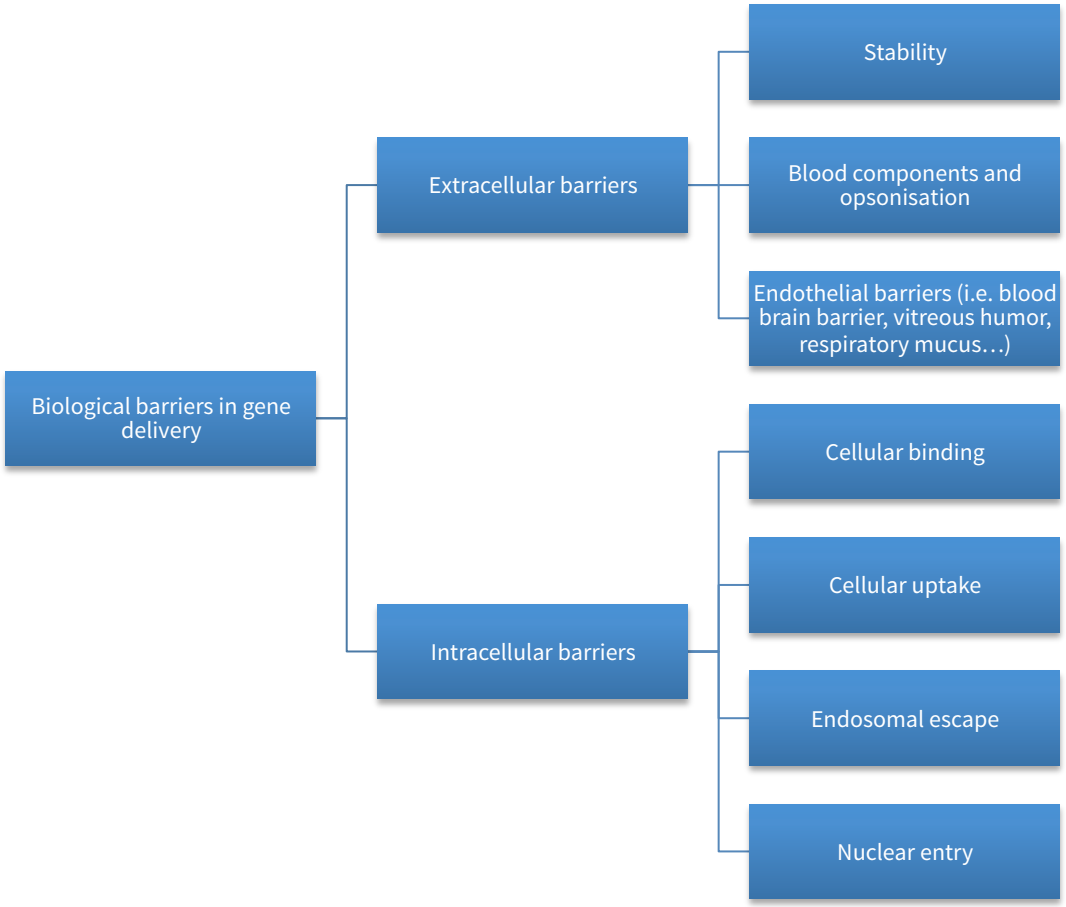

Figure 2. Schematic representation of biological barriers to gene delivery

Interaction with blood components

A vector complex administered into the blood distributes to downstream tissues via blood circulation. When it interacts with serum proteins and/or blood cells, its biodistribution will depend on newly acquired physicochemical properties that are generally difficult to control. Negatively charged proteins such as albumin are abundant in the blood, so they could bind to cationic non-viral vectors. Adsorption of negatively charged proteins neutralizes the cationic charge of the complex and increases its size, which leads to reduced gene expression. Li et al.[52,53] reported that, after exposure to mouse serum, cationic lipid–DNA complex becomes negatively charged, significantly increases in size, and eventually disintegrates and releases DNA.

If the size of a NA complex exceeds 5 μm in diameter when prepared or in the blood circulation by interaction with blood components, it will not pass through capillaries and results in embolization of downstream tissues. One report shows embolization in the lung after intravenous injection of a cationic liposome–DNA complex[54]. In some cases, the mixing of plasmid DNA and cationic carrier *in vitro* resulted in large aggregation with sizes close to that of the capillary[55–58].

Recognition by immune system

NA complexes are recognized as a foreign material and are phagocytosed by immune cells, especially by the mononuclear phagocyte system, that is, phagocytic tissue macrophages (mainly Kupffer cells in the liver and splenic macrophages). Krieg et al.[59] reported that an unmethylated CpG dinucleotide flanked by two 5’ purines and two 3’ pyrimidines induces immune reactions. Bacterial DNA or synthetic oligonucleotides containing such a sequence can trigger B cell proliferation, and release several proinflammatory cytokines. This immune reaction is part of the innate mechanism that recognizes foreign materials and removes them from the body. Thus, CpGs have been used for enhancing genetic vaccination. However,

when plasmid DNA is used as a molecule producing a therapeutic protein in gene therapy, especially for genetic diseases, this property of CpG is a significant obstacle for such therapy. In fact, the inflammatory response after the injection of naked DNA into skeletal muscle is related to the CpG motifs in the DNA[60]. In the case of cationic liposome–DNA complexes, which show minimal toxicity in animal and clinical studies after local administration, high levels of cytokines such as interferon γ and tumor necrosis factor α are observed after their intratracheal instillation or intravenous injection[61–63]. The immune reaction against plasmid DNA is amplified by the use of cationic liposome. These cytokines not only cause toxicity in the treated animals but also inhibit transgene expression.

Vascular permeability

Intravascular delivery of nonviral NA complexes results mostly in gene expression in vascularly accessible cells such as endothelial cells[64]. When the target cells are readily accessible, the complex does not need to permeate the blood vessels. Even if the target cells localize outside of the blood vessels, the transgene product secreted from the endothelial cells to the extravascular space may reach the target. However, many other cases will require extravasation of the DNA complex. The structure of capillary walls varies depending on their resident tissue, and can be divided into three general types: continuous, fenestrated, and discontinuous endothelium[65]. NA complex could pass through only the vascular wall composed of discontinuous endothelial cells, under normal conditions. Discontinuous endothelium exists only in the liver, spleen, and bone marrow, and consists on gaps of 30–500 nm between the endothelial cells and partially or completely lacks a basement membrane. Only relatively small DNA complexes can pass through the blood vessels and directly interact with parenchymal cells. There are several approaches to improve the transport of molecules across blood vessels by increasing the vascular permeability. For instance, increased permeability of blood vessels achieved by using a hypertonic solution also

has some effects on the delivery of naked plasmid DNA administered intraportally[66] , or into the skeletal muscle[67]. The alteration of vascular permeability by using a vasodilator has also been examined. It was shown that viral vector-mediated gene transfer in brain tumors or skeletal muscle is increased by the administration of bradykinin or histamine, respectively increasing blood brain barrier permeability[68,69] .

Cellular internalization

Given the highly electrostatic, somewhat nonspecific nature of the interaction of most non-viral gene delivery systems with the cell surface, it is probable that binding of these multivalent complexes to cells results in the aggregation of various cell surface proteins. This aggregation of membrane components is expected to be a strong stimulus for internalization by those processes generally referred as endocytosis. In fact, dimerization of the syndecan family of proteoglycans has been shown to trigger nonclathrin-mediated endocytosis of low-density lipoprotein (LDL)[70]. Evidence also exists that cationic substances cause adsorptive endocytosis and subsequent internalization[71–73]. Several studies have reported the involvement of clathrin-coated pits in the internalization of complexes of DNA with cationic lipids and polymers[74,75]. Phagocytosis of cationic lipid or polymer/DNA complexes, even in cell lines that are not professional phagocytes, has also been detected[76–78]. The possibility that caveolae may be involved in the internalization process is supported by the observation that including a targeting ligand on complexes of cationic lipid or polymer with DNA for the folate receptor, which resides in caveolae, greatly facilitates their uptake into tumor cells[79–81]. Macropinocytosis is another possible pathway of internalization that has been proposed to mediate the uptake of non-viral gene delivery vectors because of its ability to internalize larger structures such as bacterium[82]. Although a common mechanism of uptake of extracellular fluid in professional phagocytic cells, macropinocytosis is a rare event in other cell types. However, stimulation of this endocytic

mechanism has been demonstrated in several mammalian cell lines[82]. A comprehensive description of the relative influence of these various pathways on uptake is currently lacking. It seems simplest to postulate that these many pathways are all functioning to different extents in individual cell types, but further systematic studies are clearly required to clarify which mechanisms are most important.

To increase the specificity of complex uptake by target cells, monoclonal antibodies, other macromolecules, or low molecular weight ligands can be utilized[83]. However, a complex must reach the target cells in order to bind their receptors.

Endosomal release

Several internalization pathways are possible depending on the vector properties[84,85] including endocytosis (clathrin/caveolae-mediated, clathrin/ caveolae-independent), macropinocytosis, and non-endocytic pathways.

It is known that more than one internalization pathway can be performed at the same time but usually the peptide-based vector uses endocytic pathways[86]. Moreover, it seems that proteins that interact with a specific cellular receptor are internalized by the clathrin-mediated endocytic pathway[85]. After being internalized, the particles tend to be trapped in the intracellular vesicle, named endosome, and then the endosome will subsequently traffic through the cell. In the late stages, endosomes fuse with lysosomes (the main degradative compartments in the cell) that have even lower pH environments with higher concentration of nucleases to digest the contents[87]. The particles face the major challenges of the acidic pH of endosomes/lysosomes, the digestive enzymes of lysosomes, and the endosomal membrane. Therefore, facilitating endosomal escape and ensuring the release of the gene become very important limiting steps in achieving an effective gene therapy[88].

One of the strategies to achieve that is to use fusogenic lipids or peptides to disrupt the endosome membrane. DOPE is sometimes employed as a fusogenic helper lipid in a cationic

lipid–NA complex. When cationic lipids bind to anionic lipids in the membrane, phase separation may occur, which would initiate inverted hexagonal phase formation and membrane destabilization[54].

Several peptides have been described that are able to promote endosomal escape and can be classified into two types depending on their escape mechanism: fusiogenic peptides and histidine-rich peptides[89]. The fusiogenic peptides are small peptides that have hydrophobic amino acids (AAs) intercalated at constant intervals with negatively charged AAs[90–107]. They are normally derived from fusion-active viral proteins[108]. Thus, when early endosomes become late endosomes, lowering their pH, a conformational change in the peptide is induced, which adopts a alpha-helix conformation, in an amphipathic structure able to fuse with the endosomal membrane, leading to pore formation and releasing all the endosomal content into the cell cytoplasm[89]. The histidine-rich peptides are small peptides with a high histidine content whose endosomolytic activity is mediated by a mechanism called ‘‘proton sponge’’[90,107,109–112]. When the endosomal pH becomes low in late stages, the imidazole groups of the histidines are protonated and attract endosomal Cl^- ions, buffering against the proton pump. The endosomes collapse by an osmolytic swelling process and the endosomal content is released to the cell cytoplasm[89].

Nuclear delivery and expression

When talking about nuclear delivery it is important to take into account that only NA that are functional in the nuclear compartment have to be delivered there, such as DNA or non-mature RNAs.

NAs are not stable in the cytoplasm. Lechardeur et al.[113] reported that microinjected plasmid DNA is rapidly degraded in the cytoplasm. The apparent half-life was 50–90 min. Cytoplasmic nucleases might be responsible for the degradation.

One common mechanism to achieve nuclear localization is waiting for the cell to divide and the nuclear membrane to

disappear. Then, the complex can interact directly with the chromatin and when the nucleus is formed again with the new cells it remains inside. This mechanism is no recommended since it is only possible with dividing cells and lowers the final efficiency.

Molecules lower than 45 kDa/10–30 nm are able to enter in the cellular nucleus by passive diffusion. However, macromolecules higher than 45 kDa/ 10–30 nm generally require an active transport system through the nuclear pore system. This transport mechanism generally requires a specific targeting signal peptide named nuclear localization signal (NLS). These signalling peptides are usually rich in basic AAs, which are recognized by the cellular importines and actively transported through the nuclear pore[114,115]. Monopartite or bipartite NLS sequences which are NLS peptides that have one or two NLS recognized sequences respectively have been described[90]. It has been reported that a single NLS sequence is sufficient to transport the vector to the nucleus and that a large number of NLS sequences can result in inhibition of its activity[116], but it is an open issue.

One of the most used NLS signal peptides are fragments derived from the 111–135 AAs of the simian virus SV40 large tumor antigen (T-ag). Other NLS sequences can be found in GAL4, protamines, or Tat[89,104,117–130].

It is important that when the transported DNA reaches the cellular nucleus, it has to be released in order to be accessible to the nuclear transcription factors and achieve the desired expression level.

Nanoparticles

It is essential to note that when designing and applying vectors for gene delivery their final size is a key element. The ideal size is such that enables all the functions required for gene transfer and not surprisingly in the same nanometric range as viral particles.

Nanotechnology (sometimes shortened to "nanotech") is the manipulation of matter on an atomic and molecular scale. Generally, nanotechnology works with materials with at least one dimension sized from 1 to 100 nanometres called nanoparticles. A particle is defined as a small object that behaves as a whole unit in terms of its transport and properties.

Nanotechnology in medicine

Nanomedicine is the medical application of nanotechnology. Nanomedicine ranges from the medical applications of nanomaterials, to nanoelectronic biosensors, and even possible future applications of molecular nanotechnology.
Protein-based nanomedicine platforms for drug or NA delivery comprise naturally self-assembled protein subunits of the same protein or a combination of proteins making up a complete system. They are ideal for delivery platforms due to their biocompatibility, biodegradability coupled with low toxicity and their improved biodistribuition in circulation systems.

Importance of size

Particle size, molecular organization, and nanoscale properties, are potentially critical parameters for cell attachment, internalization and endosomal escape but also biodistribution[51,131,132]. Protein-based nanoparticles overcome

some of the drawbacks of polymer-based delivery systems due to their uniform nanometer size, bioavailability, biodegradable properties, and multifunctional groups on the surface and inside of the particle.

Self-assembling particles

Spontaneous protein self-assembly to form ordered oligomers is a common event in biology. It can prove advantageous in terms of genome-size minimization, formation of large structures, stabilization of complexes, and inclusion of functional features[133]. It has been widely documented that cellular oligomer proteins as well as viral capsids are stabilized by several weak non-covalent interactions as hydrophobic interaction, electrostatic energy, and Van der Waals forces[134–136]. Furthermore, protein–protein interactions are not the unique parameters involved in particle formation, nucleic acid–peptide interactions, salt concentration, order of mix, and ratio between nucleic acid and protein can also strongly influence the condensation process[137,138].

Non-viral gene therapy vectors

Types

Non-viral vehicles are an attractive alternative to viral vectors due to their putative increased safety, versatility and ease of preparation and scale-up.

These vectors for gene delivery make use of naturally occurring or synthetic materials to deliver the NA of interest to the target cells. Additional functionalities on non-viral vectors improve their specificity towards the targets sites. However, non-viral vectors are normally seen as less efficient than the viral ones and generally, the expression of the delivered entity is short-lived.

Non-viral methods include the use of naked DNA, proteins, lipoplexes and polyplexes or a combination of both.

Naked DNA

The simplest non-viral gene delivery vector is naked plasmid DNA. Until recently, systemic gene delivery with naked DNA was considered unrealistic, since plasmid vectors can be rapidly degraded and neutralised by endogenous DNAses. Additionally, the phosphate groups of DNA molecules confer a net negative charge to the molecules, limiting the potential for electrostatic interaction with the anionic lipids in the cell membrane.

However, the potential for correction of genetic defects by direct use of plasmid DNA might be greater than previously thought. Wolff and co-workers demonstrated that after direct injection of naked DNA to skeletal muscle cells there was a mild expression of reporter genes[139]. Although the expression levels were very low, they were enough for vaccination therapies or genetic immunisation. This kind of therapy is based in the production by muscular cells of viral antigens that cause and immunological response[140]. Skin cells, some tumours and some cells of the immunological system are also susceptible to gene transfer mediated by naked DNA.

Enhancement of transgene expression using electroporation

Local administration of nake plasmid DNA coupled with application of an alternating electrical field – electroporation – has been shown to yield high levels of transgene expression in the liver, testes, arteries, skin and tumours[141–145]. The electrical pulse is thought to disrupt the cell membrane, forming transient pores through which DNA complexes can pass. However, high-frequency electrical fields applied for a long duration can cause local tissue damage and inflammation[146]. The most promising application of electroporation is to enhance plasmid-mediated gene expression in skeletal muscle, which might be useful for inherited muscular dystrophies or as an ectopic site from which a recombinant therapeutic protein can be secreted into the bloodstream.

Enhancement of transgene expression using ultrasound

Exposure to low-intensity ultrasound decreases the 'unstirred' layer adjacent to the membrane and temporarily increases cell membrane permeability[147]. These properties suggested that adjunctive ultrasound exposure (USE) could improve the efficiency of non-viral gene delivery[148]. Fechheimer and co-workers first demonstrated that murine fibroblasts could be transfected using sonication in 1987[149], and this was extended a decade later by the demonstration of a 2.4% transfection efficiency in primary chondrocytes using naked plasmid DNA combined with USE[150]. These proof-of-concept results have been reproduced by other investigators in many other cell types[148].

Enhancement of gene transfer using biolistics

Sanford and colleagues first demonstrated particle bombardment, or biolistics, as a gene delivery system to overcome the inherent difficulty of transgene expression in plant cells[151]. This original method utilised a gunpowder acceleration system to propel DNA-coated tungsten particles at recipient cells. Penetration of the cell wall and membrane

could result in the intracellular expression of reporter genes encoded by the exogenous DNA. Subsequently, helium driven devices have been developed, in which a helium pulse is used to accelerate DNA-coated gold microparticles through the cell membrane (Source: BioRad). The integrity of large DNA constructs can be maintained on microparticles, allowing the delivery of complex genes. Conceivably, this system could have in vivo applications in humans and other animals, with genetic immunisation as a potential application. Biolistics can also be used to achieve effective immune responses in antigen-presenting cells, such as epidermal Langerhans cells and dermal dendritic cells. Intradermal rather than intramuscular delivery of nucleic acids is a more effective route for immunisation[152] and has already shown success[153].

Enhancement of gene transfer using magnetofection

Magnetic targeting exploits paramagnetic particles as drug carriers, guides their accumulation in target tissues with local strong magnetic fields, and has been used with some success in the treatment of cancer patients[154]. This principle applied to gene vectors can be used as high throughput transfection in vitro and, more importantly, improve their efficacy in vivo. This technology has the advantage that presents rapid transfection kinetics, greatly improved dose–response characteristics and the possibility of vector targeting to a selected area[155].

Lipoplexes

The term lipoplex refers to complexes made of lipids and NAs. The most commonly used lipids to encapsulate NAs are cationic lipids. Theoretically, cationic carrier molecules might self-assembly naturally with DNA of any size and neutralise its electrostatic charge, thereby promoting cell-membrane–DNA interaction and increasing transfection efficiency. Complexation of cationic lipids with DNA[156] was first described in 1987, and lipofection (transfection mediated by lipoplexes) was reported to be 5- to >100-fold more efficient than the earlier calcium phosphate or the DEAE (diethylaminoethylene)-

dextran transfection techniques. Cationic lipids are highly soluble in aqueous solution, forming positively charged micellar structures termed liposomes that protect NAs from physic forces and enzymatic degradation. Subsequent and more modern liposome preparations contain cationic lipids but are dependent on a neutral or helper lipid usually dioleoylphosphatidylcholine (DOPE) to provide effective transfection.

Electron micrograph studies have produced images of lipoplexes with a range of macromolecular structures[157–159]. Lipoplexes are prepared at a specific ratio between the lipid and the NAs in order to obtain a net charge slightly positive to facilitate interaction with cellular membranes. This ratio must be optimized for every approach and cell type. Immediately after complexation, multiple liposomes appear to condense with DNA sandwiched between. Condensed lipoplexes are seen with diameters of 100–200nm. Large aggregates are also observed, and thought to comprise numerous lipid and NA molecules. Precisely which of these represents the most transfection-efficient fraction is not clear. Lipoplexes have been clearly seen close to endocytic pits and within endocytic vesicles just below the cell membrane[157]. Lipoplexes are thought to be internalised by endocytosis[160,161], although fusion with the cell membrane[162] or disruption of the cell membrane lipid bilayer have also been proposed[163]. Larger aggregated lipoplexes might be internalised by phagocytosis[76]. Once the lipoplex has been internalised to the endosomal system, rapid mixing of cationic (liposome) and anionic (endosome) lipids might disturb the endosomal membrane[164]. The presence of the neutral lipid DOPE in the liposome is thought to promote endosomal rupture by a mechanism involving transition from a bilayer phase to an inverted micellar structure[165]. The vast majority of the time, the endosome will mature and fuse with lysosomes, where DNA will be degraded, and no gene expression will occur. Rarely, during lipid mixing the endosomal wall will rupture, and although most of the encapsulated DNA will remain bound to the lipid, some will manage to escape into the cytoplasm and traffic to the nucleus, culminating in gene expression. Lipoplexes are powerful

systems for introducing plasmids into target cells; however, their hydrophobic and positively charged surface frequently leads to interactions with plasma proteins and other extracellular proteins, which bind non-specifically to the lipoplexes and might inactivate them[166]. In this regard, protein-resistant lipoplexes have been developed[167]. Many of these proteins are sulphated, and therefore are negatively charged, allowing them to interact with positively charged lipoplexes. Gene expression is abolished or markedly inhibited by the most heavily sulphated of these proteins, including the glycosaminoglycans (GAGs) heparin. There are some evidence that these proteins can compete with DNA contained within the lipoplex, causing it to dissociate away from cationic carrier molecules[168]. Although lipoplexes often show high levels of transgene expression following direct administration or injection into target tissues[169–172], their nonspecific membrane activity usually makes cell-selective targeting very difficult. This leads to indiscriminate transgene expression in cells present at the site of administration. Furthermore, a major problem with the application of most non-viral systems, including lipoplexes, is their poor efficiency at transfecting non-proliferating cells. This is thought to be mainly a result of the integrity of the nuclear membrane providing a physical barrier to entry.

Polyplexes

The term polyplexes is referred to polyelectrolyte complexes formed by self-assembly of DNA with cationic polymers. Cationic polymers were introduced by Wu and Wu[173] and further explored by Behr and co-workers in 1995[174].

When plasmid DNA is mixed with cationic polymers, in appropriate ratios and solvents, the resulting complexes are usually nanoparticulate (<100 nm) and homogenous. Several cationic polymers have been evaluated for their ability to form nanoparticles with DNA, the most significant being poly-L-lysine (pLL) and polyethylenimine (pEI)[175–177] but many other have been developed in order to obtain safer and more efficient vector systems, improving polymer biocompatibility,

biological stability, cell specificity and intracellular trafficking. Although the transfection efficiency of most polymeric carriers is still significantly lower than that of viral vectors, their structural flexibility allows for continued improvement in polymer activity[178] and present a few advantages over lipid-based systems such as their relatively small size and narrow distribution, higher protection against nucleases, and easy control of the physical factors such as hydrophobicity and charge[179]. Typically, polyplexes are more stable than lipoplexes[180]. In addition, the cationic residues within the polyplexes can enhance the binding with the cell and may also mediate the transfer of nucleic acid to cytoplasm by disruption of the vesicular membranes. It is clear that the properties of the complexes formed are largely determined by formulation conditions[181] as well as polymer parameters[182], including molecular weight and charge density. One important and frequently used transfection parameter for non-viral gene delivery systems is the charge ratio, the ratio of the cationic carrier nitrogens to NA phosphates (N/P ratio)[183]. Successful delivery is dependent on many different factors including charge ratio, cell line, medium composition (specially salt concentration), the aggregation of complexes and the length, molecular weight and the hydrophilic components of the carrier structure[184].

Poly-L-lysine

PLL polymers are one of the first cationic polymers to be employed for gene transfer. PLL is a cationic biocompatible and biodegradable polymer, which is very useful in *in vivo* applications, consisting of only primary amines groups of lysine, which allows easy structure modifications[185,186]. PLL can spontaneously interact, through electrostatic interactions, with the negatively charged phosphate groups of DNA to form polyplexes. PLL with low MW (less than 3 kDa) cannot form stable complexes with DNA, while PLL with high MW is more suitable for gene delivery via systemic injection but exhibits relatively high cytotoxicity[187]. In addition, PLL has poor transfection ability when applied alone or without

modifications because PLL/DNA polyplexes are rapidly bound to plasma proteins and cleared from the circulation and moreover, its ability for endosomal scape is very poor. Thus, searching for more efficient cationic polymers or modified PLLs has been the main task in recent years.

Polyethylenimine

PEIs (2 to 25 kDa), obtained by acid-catalyzed polymerization of aziridine, yelding a highly branched network, are the most lucrative and extensively investigated gene delivery polymers due to their membrane destabilization potential, high charge density (NAs condensation capability) and ability to provide safeguard for endocytosed NAs from enzymatic degradation[188]. They have been proved to be very efficient in intracellular delivery of nucleic acids in a variety of cells *in vitro* and tissues *in vivo* compared with other types of cationic polymers. Since the first successful PEI-mediated oligonucleotide transfer conducted in 1995[174], PEI has been derivatized to improve the physicochemical and biological properties of polyplexes.

Branched PEI (bPEI) and linear PEI (lPEI) can both be used effectively for gene delivery. bPEI contains a higher percentage of primary amines and is more amenable to modifications. lPEI is generally preferred for in vivo applications because of its low cytotoxicity. Several linear PEI transfection agents have been made commercially available, including ExGen500 and jetPEITM[189] Moreover, jetPEITM has been selected for gene delivery applications and preclinical studies[190].

In the structure of bPEI, there is a high presence of primary, secondary, and tertiary amines in the ratio of 1:2:1 with pKa values around the physiological pH, leaving the majority as non-protonated at the physiological pH providing remarkable buffering capacity. More importantly, these non-protonated amines can exert the so-called proton sponge effect to enable the PEI to escape from the endosome. It was reported that the excess free PEI present in the colloidal suspension substantially contributed to efficient gene expression because of the proton sponge effect[191]. However, they also mediated toxic effects in a dose-dependent manner. Therefore, purified polyplexes

without free PEI have to be applied at increased concentrations to achieve high transfection levels, but they exhibit a greatly improved toxicity profile[192]. The MW, degree of branching, ionic strength of the solution, zeta potential, particle size, configuration, and the charge ratio of polymer to nucleic acid used can affect the transfection efficiency and toxicity of PEI polyplexes. The primary amines are mainly responsible for the high degree of DNA binding, but they also contribute to the toxicity during transfection, while the secondary and tertiary amino groups provide good buffering capacity to the system. The increased toxicity is caused by aggregation and adherence on the cell surface, which results in significant necrosis. The high charge density can effectively increase the transfection efficiency but it simultaneously contributes to increased cytotoxicity. Actually, the efficacy of the transfection system results from a balance between the transfection efficiency and cytotoxicity.

Dendrimers

A variety of positively charged dendrimers that have the advantages of low toxicity, high transfection, and ease of manufacturing have been used in the recent years for their ability as effective gene vectors. Dendrimers have a unique highly branched molecular architecture that is globular and possess repeating units emanating from a central core[193]. Its chemistry facilitates synthesis of a broad range of molecules with different functionalities. These functionalities contribute to the molecules' surface characterisitics, offer multiple attachment sites, e.g. for conjugation of drugs or targeting moieties. This defined structure, the inner cavities to encapsulate guest molecules and controllable multivalent functionalities in their inner or outer shells make dendrimers attractive for gene delivery[193]. They can be synthesized by convergent or divergent methods and the resulting dendrimers grow in a geometrically progressive fashion.
PAMAM dendrimers are the most commonly encountered due to their high transfection efficiency. However, they are not biodegradable, thus causing significant problems *in vivo*.

Typical approaches to optimize dendritic gene delivery for *in vivo* use, involve the surface modification of PAMAM backbone, either with arginine[194] or hydroxyl groups[195]. Nam and co-workers reported that PAMAM dendrimers functionalized with L-arginine (PAMAM-Arg) by an ester bond rather than amide bond showed equivalent transfection efficiency to PEI but lower toxicity, which could be attributed to the faster degradation of the ester bond avoiding carrier accumulation in the tissue[196].

Despite the extensive interest in the pharmaceutical applications of dendrimers, the clinical applicability of a dendrimer-based gene transfection agent is yet to be established and many basic principles of applicability are still highly debated[197]. Dendrimers have a long way to go before they can enter clinical applications. The success of the applications of dendrimers is likely to depend on the continuing development of novel materials and techniques for dendrimer synthesis[198]. The ability to functionalize the terminal groups and structures offers endless possibilities to solve all the problems. Overall, the reports available to date certainly suggest that dendrimer-based delivery systems hold great promise and potential in gene delivery.

Chitosan

Chitosan is a linear polysaccharide that has attracted considerable attention as a non-viral gene delivery vehicle due to its cationic, biodegradable and biocomptatible characteristics. Among nonviral vectors, chitosan and its derivatives have been developed for *in vitro* and *in vivo* DNA and siRNA delivery systems. Transfection efficiency of chitosan is significantly lower compared to other cationic gene delivery agents such as PEI, dendrimers or cationic lipids. This is attributed mainly to its minimal solubility and low buffering capacity at physiological pH, leading to poor endosomal escape and inefficient cytoplasmic decoupling of the complexed NA.

Chitosan is composed of repeating D-glucosamine (deacetylated unit) and N-acetyl-D-glucosamine units

(acetylated unit) linked via β-(1–4) glycosidic bonds[199]. It is produced commercially by deacetylation of chitin, which is the structural element in the exoskeleton of crustaceans. Mumper et al. were the first to study the potential of chitosan for *in vitro* DNA delivery[200]. Positively charged chitosan backbone and negatively charged DNA lead to the spontaneous formation of nanosized complexes (polyplexes) in the aqueous environment. Besides, each deacetylated subunit of chitosan contains a primary amine group with a pKa value of about 6.5, which makes it soluble in acidic media and insoluble at neutral and alkaline pH values. It is biodegradable and can be readily digested either by lysozymes or by chitinases in the physiological environment. More importantly, chitosan is practically nontoxic to animals as well as humans.

Over the years, though numerous innovative chitosan-derived vectors have been generated and the exciting *in vitro* and *in vivo* data has been published, no accredited chitosan vector usable in clinical practices has so far been recognized. Based on recent studies, it is still an uphill struggle to search for clinically applicable chitosan vectors with high transfectability in the near future.

Cyclodextrin

Cyclodextrins (CDs) are a series of natural, cyclic, water-soluble oligosaccharides composed of six (α-CD), seven (β-CD), or eight (γ-CD) D(+)-glucose units linked by α-1,4 linkages. They have a torus-like architecture with an apolar interior cavity, which can induce an inclusion complex with a hydrophobic moiety[201]. Various molecules can be fitted into the cavities of CDs to form supramolecular inclusion complexes. Therefore, natural CDs and modified CDs have recently attracted considerable attention as gene delivery vectors due to their excellent biocompatibility and unique molecular architecture[202].

The development of CD-based supramolecular biomaterials for gene delivery is an emerging area, which still faces many challenges. Although a number of proof-of-concept studies have been demonstrated, smarter material designs are expected for a better balance of higher functions and

performances and lower toxicity, both *in vitro* and *in vivo*. Safety of the delivery systems is still a major concern. Intensive in vitro and in vivo biocompatibility and biodegradability studies are desired for supramolecular gene delivery systems.

Peptide-Based or protein-based vehicles

This section will be further discussed later in this introduction.

Protein-only vehicles

Types of protein-only artificial viruses

In contrast to the gene therapy vehicles discussed in the previous section “vehicles types of non-viral gene therapy”, protein-only particles can be engineered to contain all the required functional modules in a single-chain molecule. This approach avoids the need for further functionalization of artificial viruses because independent elements, such as cell ligands, membrane-active peptides, NLSs and others, can be incorporated into a protein. There are two main architectural principles for the construction of protein-only vectors.

The first approach is based on the use of recombinantly expressed structural proteins from some viral species that spontaneously assemble into virus-like particles (VLPs). These viral empty particles with viral symmetry and nanometric sizes can accommodate heterologous NAs through in vitro incubation. Gene delivery can be achieved with the same specificity as that of the full virus particle, but genetic engineering of viral proteins allows modifications or addition of biological properties, such as receptor tropism. Although recombinant capsid proteins from polyomaviruses[203], rotaviruses[204], caliciviruses[205] and retroviruses[206] have been shown to form VLPs, not all of them have yet been tested for their ability to accommodate and deliver foreign nucleic acids.

The second approach uses multifunctional proteins based on modular principles by conventional genetic engineering. This approach combines, in a single polypeptide, the different biological functions required for accurate gene delivery specified in the previous section of “Barriers for gene delivery”. Such proteins must obviously contain domains for NA complexation, either in the form of cationic peptides, such as polylysine, that electrostatically interacts with NAs independently of the nucleic acid sequence[207] or DNA-binding motifs, such as GAL4, that recognize specific DNA sequences that need to be present in the cargo DNA[208]. The other functional motifs that are required for receptor recognition, endosomal escape and nuclear transport are also included in

the same polypeptide chain[209]. The optimal protein structure and amino acid sequences of these multifunctional proteins cannot be easily predicted in advance, and selecting the appropriate modules requires time-consuming trial-and-error approaches. However, this approach is nevertheless promising and attractive because it allows for a complete de novo design of gene delivery vehicles that can accommodate any desired peptide[210]. Most vector prototypes that have been generated to date contain either peptide domains or full-length proteins.

Designing a protein-only vehicle for gene delivery

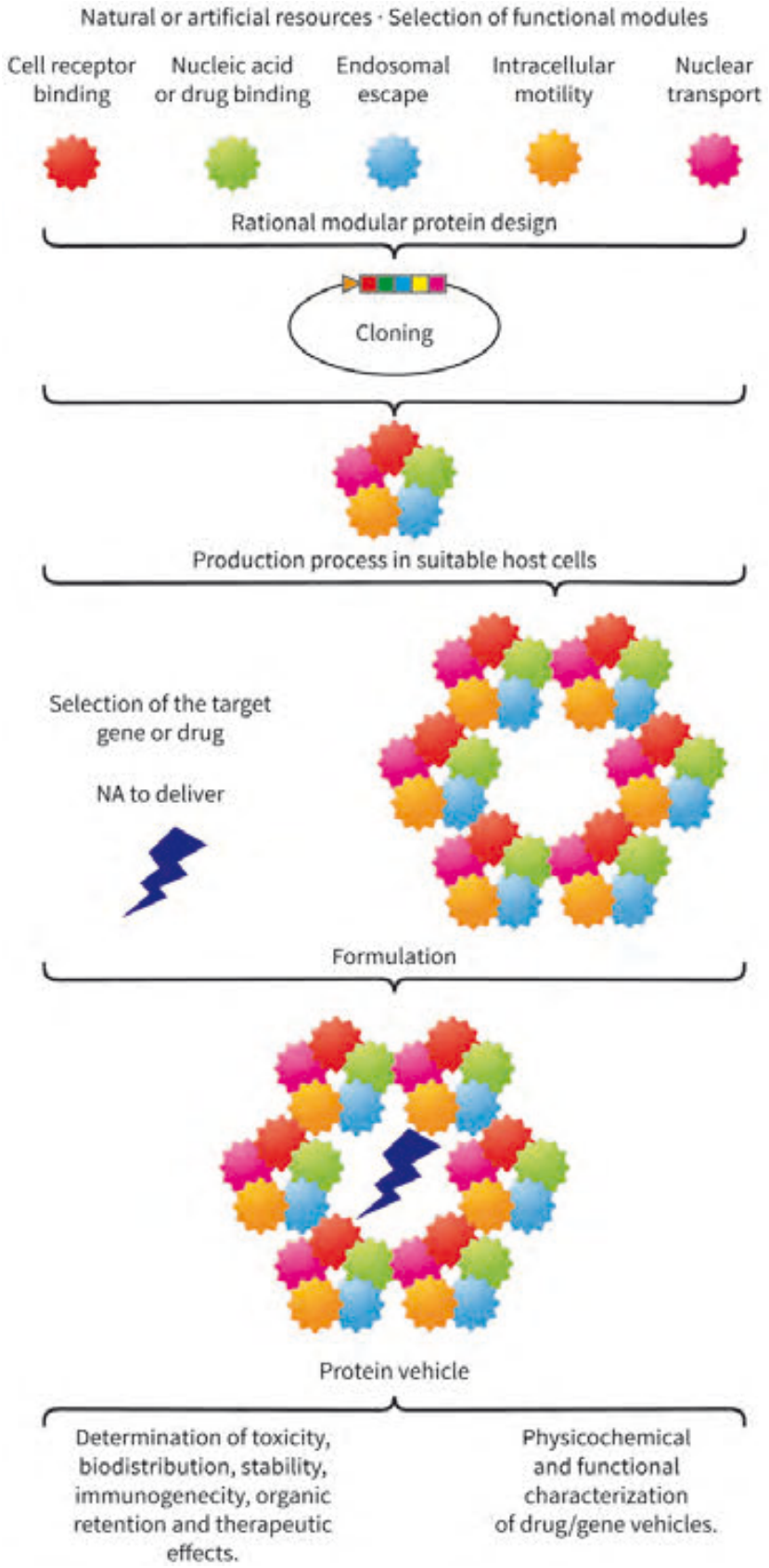

Figure 3. Designing, producing and characterising protein-only vehicles.

Two main strategies are used in the design of protein-only vehicles. *De novo* protein design offers the broadest possibility for new structures. It is based on searches for amino acid sequences that are compatible with a three-dimensional protein backbone template using *in silico* techniques conferring a specific function. In silico protein design has allowed novel functions on templates originally lacking those properties, modifying existing functions, and increasing protein stability or specificity. Beyond any doubt, intense research activity is ongoing in the field, the potential of which is simply enormous[211]. The zinc-finger protein designed by Dahiyat and Mayo[212] was the first one to appear by this method.

Rational design of proteins is based on the modification or insertion of selected amino acids or domains in a polypeptide chain backbone to obtain proteins with new or altered biological functions. When using this strategy, a detailed knowledge of the structure and function of the backbone protein is needed to make desired changes. This generally has the advantage of being inexpensive and technically feasible. However, a major drawback of this approach is that detailed structural knowledge of a protein is often unavailable or it can be extremely difficult to predict the effects of various modifications. Modular engineering enables, by using simple DNA recombinant techniques, the construction of chimerical polypeptides in which selected domains, potentially from different origins, provide the required activities. An equilibrate combination and spatial distribution of such elements has generated promising prototypes, able to deliver functional NAs to tissue cultured cells but also to specific cell types in *in vivo* approaches[213]. Modular fusion proteins that combine distinct functions required for cell type specific uptake and intracellular delivery of NAs present an attractive approach for the development of non-viral gene therapy vehicles. The group of Uherek et al described one of the first examples for this technology. They combined a cell-specific target module (antibody fragment specific for the tumour-associated ErbB2 antigen), a DNA-binding domain (Gal4), and a translocation domain for endosomal escape[214].

In this context, many strategies for the construction of safer vehicles are being explored and the number of non-viral prototype vectors for gene delivery is increasing. The common steps that an approach like this might explore are presented as follows: (a) design of the vehicle itself, required functions, stability, etc.; (b) production of the protein, suitable expression system, purification procedure, scaling up process, etc.; (c) characterization of the vehicle by physicochemical and functional tests; and finally (d) the administration route and regulatory guidance for biological products. Although all these aspects belong to different disciplines, they have to be overviewed together. Here, the major needs of a modular protein for gene delivery are presented.

Designing the vehicle requires to take into account all the aspects discussed in previous sections, from the type of NA to deliver to the barriers that need to be overcome. To have a more precise look into this aspects refer to previous sections in this introduction.

Generally, in protein nanoparticle approaches, the protein is composed by different modules of natural sources or from combinational peptide/protein libraries such as the cell-penetrating peptide transactivator of transcription (TAT) derived from the TAT of the human immunodeficiency virus (HIV)[215] or artificial sequences not present in any organism such as the polylysine DNA-condensing sequence[216].

Once it has been defined which modules will be part of the protein, it is important to define the order they will have in the final construct. It has been demonstrated by Boekle and coworkers using melittin conjugated to polyethylenimine (PEI) that depending on the side of the linkage (C- or N-terminus), the lytic activity could be changed. Some other modules have position requirements for its correct function[217].

Finally, it is very interesting to mention a new set of modules called “Therapeutic modules”. These are proteins or parts of proteins that confer the final vehicle with functions that are not necessary related to crossing barriers, but affect the target tissue in a therapeutic manner. When designing a protein

vehicle it is useful to know the type of disease to be treated and not only the target tissue to deliver the NA. If this is known, a module that produces a change in the target tissue can be added. This change normally is beneficial for the diseased tissue and moreover, it is observed before the NA effect appears. Very few examples of these modules can be found in the literature. One of them is the RGD module present in many proteins as a cell attachment motif[218]. It has been observed to trigger the proliferation and some kind of differentiation in nervous cells[219] and could be used as a proliferation enhancer for cell-grow stimulation approaches.

In the following table (Table 3) a list of example modules that can be used when designing a protein vehicle for gene delivery is added.

Peptide motif	Sequence or origin or action	Ref.
Nucleic acid condensation peptides		
Polylysine	(KKKKKKKKKKKKKKKKKKKK)n	220–222
Polylysine containing peptides	YKAKKKKKKKKWK and derivatives	91,109,223,224
Salmon protamine	PRRRRSSSRPVRRRRRPRVSRRRRRRGGRRRR	117,225,226
GAL4	MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWEC RYSPK	208,227
Blood–brain barrier (BBB) peptides		
g7	H2N-Gly-L-Phe-d-Thr-Gly-L-Phe-L-Leu-L-Ser (O-b-d-glucose)-CONH2	228
RVG	YTIWMPENPRPGTPCDIFTNSRGKRASNG	229
Tat	YGRKKRRQRRR	230
R9	RRRRRRRRR	231
Cell-penetrating peptides (CPP)		
Tat	GRKKRRQRRPPQ	89,92,93,118,232,233
R9	RRRRRRRRR	234
Penetratin	CRQIKIWFQNRRMKWKK	235,236
bPrPp	MVKSKIGSWILVLFVAMWSDVGLCKKRPKP	235
Transportan	CLIKKALAALAKLNIKLLYGASNLTWG	94,95,236
Receptor-specific ligands (ligand/receptor)		
RGD/integrins (mainly $\alpha v\beta 3$)	GRGDSP	237,238
CXCL12/CXCR4	KPVSLSYRCPCRFFESHVARANVKHLKILNTPNCALQI VARLKNNNRQVCIDPKLKWIQEYLEKALN	239,24
Transferrin receptor ligand (12Aa) /transferrin receptor	THRPPMWSPVWP	241,242

EGF/EGF receptor	NPVVGYIGERPQYRDL	243,244
Asioaloglycoprotein/asioaloglycoprotein receptor	Length may vary from 291 to 311 AA	245
RVG/acetil-colin receptor	YTIWMPENPRPGTPCDIFTNSRGKRASNG	229
PLAEIDGIELTY/integrin a9b1	PLAEIDGIELTY	246
Molossin (RGD)/integrin	ICRRARGDNPDDRCT	247
Secretin/Secretin receptor	HSDGTFTSELSRLRDSARLQRLLQGLV	248
NL4 (loop 4 of nerve growth factor)/TrkA	CTTTHTFVKALTMDGKQAAWRFIRIDTAC	249
Neurotensin/Neurotensin receptor (NTRH)	ELYENKPRRPYIL	250
LSIPPKA, FQTPPQL, LTPATAI/LOX-1	LSIPPKA, FQTPPQL, LTPATAI	251
Monoclonal Abs/antigen recognized by the antibody	Length and sequence vary	252–254

Endosomal escape fusogenic peptides		
HA-2	GLFGAIAGFIENGWEGMIDGWYG	90,96,97
GALA	WEAALAEALAEALAEHLAEALAEALEALAA	90,98–101
KALA	WEAKLAKALAKALAKHLAKALAKALKACEA	90,102
JTS-1	GLFEALLELLESLWELLLEA	90,91
ppTG20	GLFRALLRLLRSLWRLLLRA	90,103
PPTG1	GLFKALLKLLKSLWKLLLKA	90,103
Melittin	GIGAVLKVLTTGLPALISWIKRKRQQ	90,104
Tat	GRKKRRQRRRPPQ	90,92,93
Penetratin	RQIKIWFQNRRMKWKK	90,105,106
Transportant	GWTLNSAGYLLGKINLKALAALAKKIL	90,94,95
INF 7	GLFEAIEGFIENGWEGMIDGWYG	90,107

Endosomal escape histidine-rich peptides		
CHK6HC	CHKKKKKKHC	90,109
H5WYG	GLFHAIAHFIHGGWHGLIHGWYG	90,107,110,111
LAH4	KKALLALALHHLAHLALHLALALKKA	90,112

Nuclear import peptides		
SV40 large T antigen	PKKKRKV	89
Tat	VIH transcription factor	118
EBNA-1	Epstein–Barr virus	119
Melittin	Honeybee venom (Apis mellifera)	104
M1 (c-myc transcription factor)	PAAKRVKLD	120
M2 (c-myc transcription factor)	RQRRNELKRSP	120
GAL4 amino terminal domain	Transcription factor	121
Protamines	Sperm DNA condensation protein	117
Histone H1	Nuclear DNA condensation protein	122,123
M9 (heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1))	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQ GGY	124
Vp3	SV40 structural protein Vp3	125
Adenovirus E1 protein C- terminus	KRPRP	126
Xenopus N1 protein	VRKKRKTEEESPLKDKDAKKSKQE	127
Fibroblast growth factor 3 (FDF3)	RLRRDAGGRGGVYEHLGGAPRRRK	128
Poly ADP-ribose polymerase (PARP)	KRKGDEVDGVDECAKKSKK	129
Xenopus protein nucleoplasmine	KRPAATKKAGQAKKKK	130

Therapeutic modules (and action)		
RGD	Promotes cell growth and differentiation	[219]
Fibroblast growth factor-1 (FGF-1)	Angiogenesis in cardiovascular diseases	[255]
Hepatocyte growth factor (HGF)	Restore hepatocytes in hepatitis B	[256]
Anthrax toxin	Cell death in cancer therapy	[257]
TATp53	p53 activation in cancer therapy	[258]
KLA peptide	Promotes apoptosis	[259]

Table 3. Selection of peptide motifs used in gene delivery to improve protein-only vehicle's performance.

Producing a protein-only vehicle for gene delivery

Some steps in the production of a protein-based vehicle after molecular cloning such as protein production and protein purification might be experimentally adjusted and optimised with a variable success rate. For that reason, when small proteins are needed, solid-phase peptide synthesis guarantees the process. However, the classical procedure of biological production allows scaling up the process in most of the cases and the production of larger polypeptides and full- length proteins.

When producing a protein for gene or drug delivery, it is important to know the origin of its domains to choose the most suitable expression system for its production. For instance, if any module naturally carries a posttranslational modification that is essential for its biological function, the expression system chosen will have to be able to reproduce the same crucial modification.

The main biological production systems for protein drugs are summarized below.

Escherichia coli is the most widely used prokaryotic organism for the expression of recombinant proteins[260]. The use of this host is relatively simple and inexpensive[261]. Added advantages include its short duplication time, growth to high cell densities, ease of cultivation, and high yields of the recombinant product. However, since it lacks fundamental prerequisites for efficient secretion, many recombinant proteins manufactured by *E. coli* systems are stored inside the cell, with the problem that

inclusion bodies can be formed[262,263]. Moreover, posttranscriptional modifications are not achieved with this system.

Like *E. coli*, yeasts can be grown cheaply and rapidly and are amenable to high cell-density fermentations. Besides possessing complex posttranslational modification pathways, they offer the advantage of being neither pyrogenic nor pathogenic and are able to secrete more efficiently. Species established in industrial production procedures are *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Pichia pastoris*, and *Hansenulapolymorpha*. *S. cerevisiae* is the best genetically characterized eukaryotic organism among them all and is still the prevalent yeast species in pharmaceutical production processes[263]. In spite of their physiological advantageous properties and natively high expression and secretion capacity, the employability of yeasts in some cases, however, might reach a limit, particularly when the pharmacological activity of the product is impaired by the glycosylation pattern. In such cases, either a post synthetic chemical modification has to be considered or the employment of more highly developed organisms.

Animal cell expression systems show the highest similarity to human cells regarding the pattern and capacity of posttranslational modifications and the codon usage. However, their culture is more complicated and expensive and usually the production yields are lower. Among the known systems, insect cells infected by baculovirus vectors have reached popularity since they are considered to be more stress-resistant, easier to handle, and more productive compared with mammalian systems and are thus frequently employed for high-throughput protein expression. For commercial application, scale-up related questions have to be solved[264–266].

Preferably applied in pharmaceutical production processes are mammalian expression systems like chinese hamster ovary (CHO) cells and baby hamster kidney (BHK) cells. These systems are genetically more stable and easier to transform and handle in scale-up processes, to grow faster in adherent and submerged cultures, and to be more similar to human cells and more consistent in their complete spectrum of

modification[267]. In some cases, mammalian cell systems can be the only choice for the preparation of correctly modified proteins.

Finally, peptides, being complex and unique molecules with regard to its chemical and physical properties, can be produced synthetically by the solid- phase method[268,269]. This technology can be used to avoid problems related to biological production. General advantages of synthetic peptides are that they are very stable compounds, solid-phase chemistry produces highly standardized peptides, and the crucial polycation component is provided by a ‘‘natural’’ polycation, thus minimizing toxicity[270]. However, some disadvantages related to synthetic peptides have been reported such as the difficulty to synthesize long and well-folded oligopeptides, peptides with multiple cysteine, methionine, arginine, and tryptophan residues due to technical limitations or production cost[270].

Characterizing a protein-only vehicle for gene delivery

When working with protein nanoparticles, it is very important to characterize them physically and functionally in order to understand their behaviour.

Screening of new gene delivery candidates regarding transfection efficiency and toxicity is usually performed by reading out transgene expression levels relative to a reference formulation after in vitro transfection[141,210]. However, over the years and among different laboratories, this screening has been performed in a variety of cell lines, using different conditions and read-out systems, and by comparison to a variety of reference formulations. This issue is easily illustrated by analyzing the materials and methods sections of a small sample of papers recently published on the topic of *in vitro* screenings of reagents for improved gene delivery. The large variability in experimental conditions observed in this limited sample is representative for gene delivery studies in general, and makes a direct comparison of results difficult, if not impossible.

Besides a lack of consensus among *in vitro* studies, *in vitro* transfection screenings are often criticized for poorly correlating with *in vivo* results. Nevertheless, *in vitro* transfection screening is an indispensable tool as (a) testing of large numbers of reagents under various conditions directly *in vivo* is neither feasible nor ethical and (b) recent developments in chemical synthesis and (recombinant) technology enable generation of large libraries of synthetic or recombinant carriers, which warrants high-throughput screening assays[271]. Therefore, understanding and controlling *in vitro* transfection is of major importance to develop rational, reliable and fast screening procedures that enable evaluation of structure–activity relations.

It is important to improve future comparability between studies and laboratories by proposing the standardization of screening protocols and inclusion of a benchmark for interstudy comparison[272].

Here, the most common methods for physical and functional screening and characterization will be discussed.

Size and topology characterization

When developing reagents for non-viral gene delivery, the ideal scenario is to have well-defined nanoparticles, preferably <100 nm, which are stable in physiological media and efficiently transfect cells with minimal toxicity. In practice, NA-complexes are mostly formed via electrostatic interaction with positively charged polymers, lipids or peptides, are sensitive to dispersion media (i.e. ionic strength) and are subject to changes upon incubation in physiological media. Particle characteristics are well-known determinants in cell transfection and a set of complementary tools is available to characterize formed complexes. Parameters of interest are size and stability, surface charge, condensation and protection of DNA, which can be studied with techniques including dynamic light scattering (DLS), zeta-potential measurements, electrophoretic mobility shift assays, retardation and fluorescence displacement assays (in presence and absence of displacing poly-anions such as heparin) and nuclease

resistance assays. Imaging techniques such as electron microscopy and atomic force microscopy can provide supplementary insight into shape, morphology and size of particles[273,274].

Functional characterization

Despite the fact that size and topology and all the structural characteristics of the vehicles are a key element, understanding and testing the functionality and pharmacokinetics of the NA-complexes is the most important part of its development process. Most of the initial tests are done using cell lines in *in vitro* experiments using reporter genes or NA[174]. Quantifying the percentage of transfected cells is a very valuable tool to evaluate nanoparticle performance in both nuclear and cytoplasmic delivery. In addition, *in vitro* experiments may be designed to select a candidate for the *in vivo* experiments from a group of possible therapy vectors.
Here, major decisions with big implications in the final result have to be made. A list of the important factors to take into account follows:

1. Choice of the cells to transfect
2. Transfection conditions
    2.1. Maintenance of cells
    2.2. Confluency of cells
    2.3. Incubation of samples with cells
        2.3.1. Particle formulation
        2.3.2. Composition and volume of medium
        2.3.3. Dose and incubation time
3. Readout of results
    3.1. Transfection efficiency
        3.1.1. Expression of reporter or NT targeting protein
        3.1.2. Presence of particle or NT
        3.1.3. Cell induced changes
    3.2. Toxicity
    3.3. Data presentation
4. Choice of reagent reference

This section wants to highlight the sensitivity of the outcome of transfection experiments to a high number of variables, which makes interpretation and comparison of results from literature troublesome, especially if detailed descriptions of applied methods and conditions are lacking and expression data are calculated relative to highly variable controls. Given the fact that experimental conditions can already account for 1000-fold differences[174] in transfection efficiency when testing a single reagent in a single cell type, and that the differences obtained upon optimization of the delivery agent itself are within this same range, it is obvious that there is a need for standardization of protocols in order to enable meaningful comparisons of non-viral gene delivery systems
Following a limited set of guidelines would already substantiate the value and understanding of obtained results significantly.

Finally, *in vivo* testing is the last step in the functional characterisation of any protein only vehicle for gene delivery[230]. This is the less explored part of them all because many vehicles don't reach this stage and are never tested and also because the variety of models is so big that it is difficult to compare between them.

Non-viral gene therapy in neuroscience

Gene therapy approaches have been shown to be especially useful in tumoral cell lines[275]. Neurons, however, are difficult to transfect, and initial strategies to deliver genetic material to neurons have relied on the use of viral vectors. Moreover, neurons are located in a much closed area of the body, the brain, and it's very difficult to reach it.

The potential for treating neurological diseases using non-viral gene therapy is vast but underachieved. As of May 2011, 1703 gene therapy clinical trials have been conducted worldwide, of which only 27% have utilized non-viral techniques or vectors[276,277]. Current viral gene transfer approaches being developed for the treatment of neurological disorders include gene replacement, gene knockdown, prosurvival therapy and cell suicide therapy[278–281]. Neurological diseases targeted include Parkinson, spinal muscular atrophy (SMA), amyotrophic lateral sclerosis (ALS), and brain cancer[282]. There are 15 AAV and 2 lentiviral gene therapy trials reported for central nervous system (CNS) diseases[283]. However, demonstration of clear clinical efficacy has been lacking, and only 6 viral gene therapy trials reached Phase II[280].

In this section, we will deal with the barriers that non-viral vehicles must overcome to efficiently deliver their cargo to central nervous system (CNS), that are different from other cell types due to cell morphology or topology and crossing the brain blood barrier (BBB), when vehicles are administrated not directly into the brain.

Neuronal cells have special properties, such as their polarized nature and the elongated morphology of neuronal projections, which could result in barriers to overcome that are unique to such cells[284]. Fortunately, the surface of nanomaterials can be designed to obtain specific physical and/or biological properties, thereby allowing specific interactions with membranes and improving nanoparticle uptake. The internalization pathways of neurons are the same as other cell types discussed before.

For specific targeting CNS cells, the advantage is that they are highly differentiated from other body cells. This differentiation also results in the existence of CNS-specific receptors. Neurons express a number of different classes of receptors, including neuropeptide, neurotrophin and neurotoxin receptors[250,285–290], thereby leading to strategies that are focused on brain NA-delivery that take advantage of the high specificity of ligand-receptor binding. The transferrin (Tf) or low-density lipoprotein (LDL) receptors are among those defined as constitutive (class I) receptor-mediated endocytosis, whereas others, such as insulin or epidermal growth factor (EGF), are referred to as ligand-stimulated (class II) receptor-mediated endocytosis.

The BBB is a physical barrier which limits the brain uptake of the vast majority of neuro-therapeutic and neuroimaging contrast agents. This barrier normally has to be overcome to reach the brain tissue, although an alternative pathway using the naturally occurring trans-synaptic retrograde transport route of entry into the CNS from the peripheral nerves could be another possible route for CNS entry.
The BBB is composed of a dense layer of capillary endothelial cells facing the bloodstream, which separates the CNS (the brain and the spinal cord) from the rest of the organism by means of tight junctions. On the brain side, this layer of non-fenestrated endothelial cells is lined by astrocytes, pericytes and neurons. Transport in-between brain capillary endothelial cells is restricted by expression of tight junction proteins that provide high resistance to the cell, and this is associated with limited endocytosis across the endothelium[291]. The BBB is known for its high selectivity as well as its high transendothelial electrical resistance[292]. This value is about three orders of magnitude higher than that present in other tissues, thus resulting in a reduction of the aqueous paracellular diffusion. Indeed, only small hydrophilic compounds with a mass lower than 150 Da and highly hydrophobic compounds with a mass lower than 400–600 Da can cross the membrane by passive diffusion.
Characteristic features of BBB micro-vessels include their smaller diameter and the fact that their walls are thinner than

those of other vessels in the organism. In terms of selectivity, the efficient efflux pump system composed of p-glycoprotein located in the cerebral capillary endothelium should also be mentioned. This pump is involved in recognition of those molecules necessary for the brain and the prevention of other molecules from entering the brain parenchyma[293]. Gene delivery vehicles should therefore gain access to the CNS by lipid-mediated free diffusion or potentially by receptor-mediated endocytosis.

Some human proteins such as insulin, transferrin (Tf), insulin-like growth factor, or leptins are able to go across the BBB by receptor-mediated transporters[294].

Some peptides have been described that are able to reach the brain crossing the BBB. Several in vitro and in vivo studies have shown that nanoparticles (NPs) conjugated with Tf penetrate the BBB much more easily, thereby significantly improving the delivery of active molecules to rat brain with respect to the same NPs devoid of the ligand[295]. Other strategies to increase passage through the BBB have been studied. Thus, nanoparticles conjugated with monoclonal antibodies that bind transferrin and insulin receptors, attached to other molecules, have been shown to translocate into brain tissue after intravenous administration by the brain endothelial cells[296]. Surface modifications of vehicles in order to improve their passage across the BBB can, however, decrease their ability to cross other barriers; therefore, the delivery of active molecules by NPs into cells of the CNS remains a challenge.

Conclusions about the state of the art

- Overcoming the biological and methodological obstacles posed by cell factories to the production or recombinant DNA pharmaceuticals is a main challenge in the further development of protein-based molecular medicine.
- Recombinant DNA technologies might have exhausted conventional cell factories as it is seen by the decrease rate in the approved new biopharmaceuticals and new production systems need to be deeply explored and incorporated into the production pipeline.
- A more profound comprehension of host cell physiology and stress responses to protein production is needed to offer improved tools (either at genetic, metabolic or system levels) to favour high yield and high quality protein production.
- Microbial cells appear as extremely robust and convenient hosts as they are the most used and with approved products nowadays.
- Not only commonly used bacteria and yeasts but also unconventional strains or species are observed as promising cell factories for forthcoming recombinant drugs.
- Proteins and peptides have been envisioned as potent biotechnological tools for the development of new biocompatible biological entities that can be used as therapeutic agents by themselves or as nanovehicles for the delivery of associated drugs.
- Proteins are nanostructures that can form complex high-order entities such as VLPs, resulting in appropriate cages for the internalization of therapeutic molecules.
- The design of modular proteins displaying selected functions has been possible by using in silico approximations to the feasibility of recombinant protein production.
- The modular protein approach has demonstrated the versatility of such molecules in the generation of novel delivery nanovehicles opening up the possibility of new functional combinations to enhance the specific interaction with the target tissue.

- The modular protein approach confers tuneable specificity in the delivery of drugs, nucleic acids, or other proteins.
- The presence on the market of recombinant protein entities has been increasing over the past years, and this tendency is expected to continue.
- There are some products in clinical trials that will probably end up being approved and some more are being explored in preclinical experiments, which might enter in clinical trials.

Results

Article 1

Microbial factories for recombinant pharmaceuticals

Ferrer-Miralles, N., Domingo-Espin, J., Corchero, J. L., Vazquez, E., & Villaverde, A.

Microbial Cell Factories 8. (2009).

In this review we analyse the current approved recombinant pharmaceuticals from the production system point of view. We noted that most of the hosts used to produce the biopharmaceuticals approved for human use by the Food and Drug Administration (FDA) and/or by the European Medicines Agency (EMEA) are microbial cells, either bacteria or yeast. We describe which possible biotechnology factories are used and their pros and cons. Also it is very interesting the growth analysis of approved products and their use based on the production system.

Microbial Cell Factories

BioMed Central

Review

Open Access

Microbial factories for recombinant pharmaceuticals

Neus Ferrer-Miralles[1,2,3], Joan Domingo-Espín[1,2,3], José Luis Corchero[3,1], Esther Vázquez[1,2,3] and Antonio Villaverde*[1,2,3]

Address: [1]Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain, [2]Department de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain and [3]CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona, Spain

Email: Neus Ferrer-Miralles - neus.ferrer@uab.cat; Joan Domingo-Espín - Joan.DomingoE@campus.uab.es; José Luis Corchero - jlcorchero@ciber-bbn.es; Esther Vázquez - Esther.Vazquez@uab.es; Antonio Villaverde* - antoni.villaverde@uab.cat

* Corresponding author

Published: 24 March 2009

Microbial Cell Factories 2009, **8**:17 doi:10.1186/1475-2859-8-17

Received: 11 February 2009
Accepted: 24 March 2009

This article is available from: http://www.microbialcellfactories.com/content/8/1/17

© 2009 Ferrer-Miralles et al; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Most of the hosts used to produce the 151 recombinant pharmaceuticals so far approved for human use by the Food and Drug Administration (FDA) and/or by the European Medicines Agency (EMEA) are microbial cells, either bacteria or yeast. This fact indicates that despite the diverse bottlenecks and obstacles that microbial systems pose to the efficient production of functional mammalian proteins, namely lack or unconventional post-translational modifications, proteolytic instability, poor solubility and activation of cell stress responses, among others, they represent convenient and powerful tools for recombinant protein production. The entering into the market of a progressively increasing number of protein drugs produced in non-microbial systems has not impaired the development of products obtained in microbial cells, proving the robustness of the microbial set of cellular systems (so far *Escherichia coli* and *Saccharomyces cerevisae*) developed for protein drug production. We summarize here the nature, properties and applications of all those pharmaceuticals and the relevant features of the current and potential producing hosts, in a comparative way.

Introduction

Proteins are catalysers of metabolic reactions, structural components of biological assemblies, and responsible for inter and intracellular interactions and cell signalling events that are critical for life. Therefore, deficiencies in the production of specific polypeptides or the synthesis of mutant, non-functional versions of biologically relevant protein usually derive in pathologies that can range from mild to severe. In humans, such diseases can be treated by the clinical administration of the missing protein from external sources, to reach ordinary concentrations at systemic or tissular levels [1]. Therefore, many human proteins have an important pharmaceutical value but they are difficult to obtain from their natural sources. Recombinant DNA (rDNA) technologies, developed in the late 70's using the bacterium *Escherichia coli* as a biological framework, offer a very potent set of technical platforms for the controlled and scalable production of polypeptides of interest by relatively inexpensive procedures. This can be done in convenient microbial cells such as bacteria and yeasts, whose cultivation can be accomplished by relatively simple procedures and instrumentation. In early 80's, the FDA approved the clinical use of recombinant human insulin from recombinant *E. coli* (Humulin-US/Humuline-EU) for the treatment of diabetes [2], being the first recombinant pharmaceutical to enter the market. The versatility and scaling-up possibilities of the recombinant protein production opened up new commercial opportu-

nities for pharmaceutical companies. Since the approval of recombinant insulin, other recombinant DNA drugs have been marketed in parallel with the development and improvement of several heterologous protein production systems. This has generated specific strains of many microbial species adapted to protein production, and has allowed the progressive incorporation of yeasts and eukaryotic systems for this purpose. Among the 151 protein-based recombinant pharmaceuticals licensed up to January 2009 by the FDA and EMEA, 45 (29.8%) are obtained in *Escherichia coli*, 28 (18.5%) in *Saccharomyces cerevisiae*, 17 (11.2%) in hybridoma cells, 1 in transgenic goat milk, 1 in insect cells and 59 (39%) in mammalian cells (Figure 1) [3]. In the next sections, the key properties of these expression systems will be analyzed regarding both the biological convenience and final quality of the products. Alternative promising protein production systems such as filamentous fungi, cold-adapted bacteria and alternative yeast species among others are under continuous development but only few biopharmaceutical products from them have been marketed. Relevant properties of such promising systems and their potential as producers of therapeutic proteins have been extensively reviewed elsewhere [4-12].

Escherichia coli

The enterobacterium *E. coli* is the first-choice microorganism for the production of recombinant proteins, and

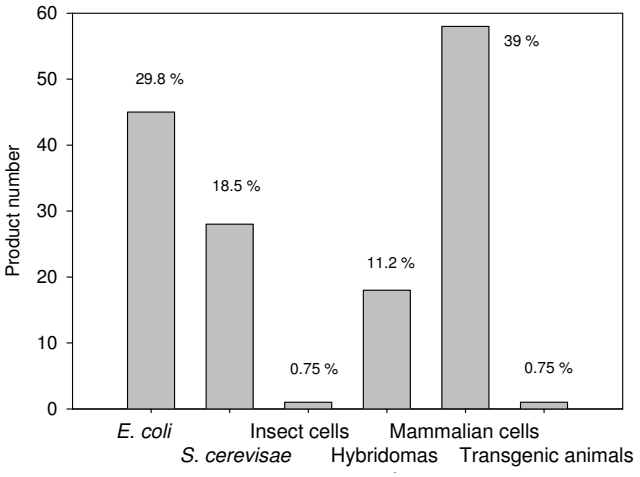

Figure 1
Number (and percentage values siding the bars) of recombinant proteins approved as biopharmaceuticals in different production systems. Data has been adapted from Table 1 in [3]. Exubera, an inhalated recombinant human insulin produced in *E. coli* has been omitted since Pfizer stopped its marketing in January 2008. Two recently FDA approved products Xyntha and Recothrom produced both in CHO cells have also been added.

widely used for primarily cloning, genetic modification and small-scale production for research purposes. This is not surprising as the historical development of microbial physiology and molecular genetics was mainly based on this species, what has resulted in a steady accumulation and worldwide use of both information and molecular tools (such as engineered phages, plasmids and gene expression cassettes). However, several obstacles to the production of quality proteins limit its application as a factory for recombinant pharmaceuticals. Recombinant proteins obtained in *E. coli* lack the post-translational modifications (PTMs) which are present in most of eukaryotic proteins [13]. Glycosylation is the most common PTM [14] but many others, such as disulfide bond formation, phosphorylation and proteolytic processing might be essential for biological activity. PTMs play a crucial role in protein folding, processing, stability, final biological activity, tissue targeting, serum half-life and immunogenicity of the protein; therefore PMT deficient version might be insoluble, unstable or inactive. Interestingly, it is possible to attach or bind synthetic PTMs in the case of pegylated products [15] such as human growth hormone, granulocyte colony stimulating factor, interferons alfa-2a and alfa-2b, which renders versions of the protein in serum more stable than the naked product. Also, the N-linked glycosylation system of *Campylobacter jejuni* has been successfully transferred to *E. coli*, making this approach a promising possibility for the production of glycosilated proteins in this species [16]. Furthermore, through genetic engineering of the underlying DNA, the amino acid sequence of the protein can be changed to alter its ADME (absorption, distribution, metabolism, and excretion) properties, as it has been observed for insulin (Table 1) [17].

On the other hand, the frequencies with which the different codons appear in *E. coli* genes are different from those occurring in human genes, and this is directly related to the abundance of specific tRNAs. Therefore, genes that contain codons rare for *E. coli* may be inefficiently expressed by this organism and cause premature termination of protein synthesis or amino acid misincorporation, thus reducing the yield of expected protein versions [18]. This problem can be solved either by site-directed replacement of rare codons in the target gene by codons that are more frequently used in *E. coli*, or, alternatively, by the co-expression of the rare tRNAs (*E. coli* strains BL21 codon plus and Rosetta were designed for this purpose). In addition, initial methionine removal depends on the side chain of the penultimate amino acid of N-terminal in final recombinant proteins produced in *E. coli* although it can be efficiently removed using recombinant methionine aminopeptidase [19]. Some mutant *E. coli* strains have been developed to promote disulfide bond formation (AD494, Origami, Rosetta-gami) and/or with reduced

Table 1: Recombinant insulins approved for human use.

INN[1]	Trade name	Production system	Modifications from natural	PK[2]
Insulin human	Humulin Insuman Exubera[3]	*E. coli*	None	Short-acting insulin
Insulin human	Novolin	*S. cerevisiae*	None	Short-acting insulin
Insulin lispro	Humalog	*E. coli*	PB28K and KB29P	Rapid-acting insulin analogue
Insulin glulisine	Apidra	*E. coli*	NB3K and KB29E	Rapid-acting insulin analogue
Insulin aspart	Novorapid	*S. cerevisiae*	DB28P	Rapid-acting insulin analogue
Insulin glargin	Lantus	*E. coli*	NA21G and 2 additional R in B chain	Long-acting insulin analogue
Insulin detemir	Levemir	*S. cerevisiae*	TB30del and myristic fatty acid attached to KB29 by acylation	Long-acting insulin analogue

Insulin is a polypeptide of 51 amino acid, 30 of which constitute A chain, and 21 of which comprise B chain. The two chains are linked by a disulfide bond. Mutations in amino acid sequences are noted for each of the chains.
[1]INN: International Nonproprietary Names. [2]PK:PharmacoKinetics. [3]Exubera: Rapid-actin insulin using inhalation route [17], was discontinued in 2008 by the manufacturer

protease activity (BL21). As an additional technical obstacle, proteins larger than 60 kDa are inefficiently obtained in soluble forms in *E. coli* [20].

As it has been well documented, bacteria overproducing either eukaryotic or prokaryotic recombinant proteins are subjected to different stresses (essentially metabolic and conformational) [21]. Under this situation, protein processing associated to cell stress responses might render non useless products, mainly because of lack of solubility, and many protein species deposit in high amounts as protein aggregates known as inclusion bodies (IBs) [22-25]. By adjusting media composition, growth temperature, inducer concentration, promoter strength and plasmid copy number, variable amounts of the target protein can be forced to appear in the soluble form [26,27], although unfortunately, many eukaryotic proteins are exclusively found trapped in IBs and seem to be resistant to process-based solubility enhancement. While IBs formed by enzymes can be efficient catalysers in enzymatic reactions [28-32], pharmaceutical proteins need, in contrast, to be dispersed as soluble entities to reach their targets at therapeutic doses. IBs essentially contain the recombinant protein in variable proportions (from 60 to more than 90%) and some contaminants as chaperones, DNA, RNA and lipids [33]. Although stored protein can be released from IBs using denaturing conditions, in vitro refolding processes are not as effective as expected [34] and other expression systems should be tried. In some cases, recombinant proteins have been successfully purified from IBs as for example Betaferon [35] and insulin [36]. However, for non integral membrane proteins, cytosolic and/or soluble protein domains, the probability of success is reasonably high and *E. coli* should be then considered as a promising expression system [37].

In summary, around 10% of full-length eukaryotic proteins tested in this system have been successfully produced in soluble form in *E. coli* [38]. Approved therapeutic protein-based products from *E. coli* include hormones (human insulin and insulin analogues, calcitonin, parathyroid hormone, human growth hormone, glucagons, somatropin and insulin growth factor 1), interferons (alfa-1, alfa 2a, alfa-2b and gamma-1b), interleukins 11 and 2, light and heavy chains raised against vascular endothelial growth factor-A, tumor necrosis factor alpha, cholera B subunit protein, B-type natriuretic peptide, granulocyte colony stimulating factor and plasminogen activator (Additional file 1). Noteworthy, most of the recombinant pharmaceuticals produced in *E. coli* are addressed for the treatment of infectious diseases or endocrine, nutritional and metabolic disorder disease groups (Figure 2).

Saccharomyces cerevisiae

Production in yeast is usually approached when the target protein is not produced in a soluble form in the prokaryotic system or a specific PTM, essential for its biological activity, cannot be produced artificially on the purified product [13]. Yeasts are as cost effective, fast and technically feasible as bacteria and high density cell cultures can also be reached in bioreactors. Even more, mutant strains that produce high amounts of heterologous protein are already available. Even though yeasts are able to perform many PTMs as O-linked glycosylation, phosphorylation, acetylation and acylation, the main pitfall of this expression system is related to N-linked glycosylation patterns which differ from higher eukaryotes, in which sugar side chains of high mannose content affect the serum half-life and immunogenicity of the final product. Although less studied than in bacteria, the production of recombinant

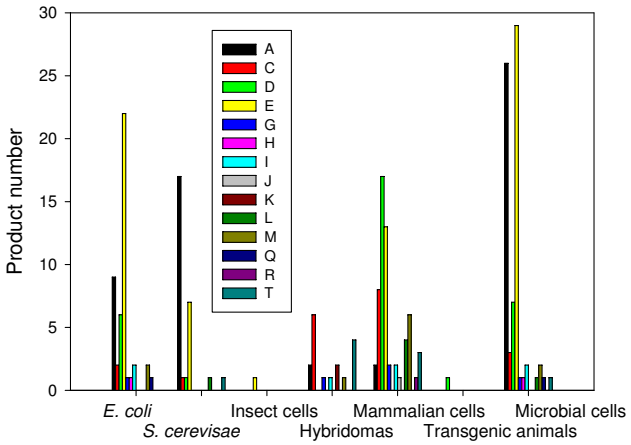

Figure 2
Number of recombinant biopharmaceuticals in different production systems, grouped by WHO therapeutic indications (see the legend of Additional file for nomenclature). Products from *E. coli* and *S. cerevisae* are also presented together under the category of microbial cells.

proteins also triggers conformational stress responses and produced proteins fail sometimes to reach their native conformation. Recent insights about conformational stress, and in general, to cell responses to protein production in recombinant yeasts have been extensively reviewed elsewhere [21,39,40].

The approved protein products produced in yeast are obtained exclusively in *Saccharomyces cerevisiae* [4] and correspond to hormones (insulin, insulin analogues, non glycosylated human growth hormone somatotropin, glucagon), vaccines (hepatitis B virus surface antigen -in the formulation of 15 out of the 28 yeast derived products-) and virus-like particles (VLPs) of the major capsid protein L1 of human papillomavirus type 6, 11, 16 and 18, urate oxidase from *Aspergillus flavus*, granulocyte-macrophage colony stimulating factor, albumin, hirudin of *Hirudo medicinalis* and human platelets derived growth factor. As in the case of *E. coli*, most of the recombinant pharmaceuticals from yeast are addressed to either infectious diseases or endocrine, nutritional and metabolic disorders (Figure 2), being these therapeutic areas the most covered by microbial products. Interestingly, several yeast species other than *S. cerevisiae* are being explored as sources of biopharmaceuticals and other proteins of biomedical interest [21,41]. In addition, current metabolic engineering approaches [42] and optimization of process procedures [43,44] are dramatically expanding the potential of yeast species for improved production of recombinant proteins.

Insect cell lines

Cultured insect cells are used as hosts for recombinant baculovirus infections. The production of a recombinant viral vector for gene expression is time-consuming, the cell growth is slow when compared with former expression systems, the cost of growth medium is high and each protein batch preparation has to be obtained from fresh cells since viral infection is lethal. PTMs are also an important limitation of this expression system because of the simple non-syalated N-linked glycosylation which is translated in a rapid clearance from human sera [45]. Although genetic engineering has been used to select transgenic insect cell lines (MIMIC™ from Invitrogen and SfSWT-3) expressing galactosyltransferase, N-acetylglucosaminyltransferases, syalic acid synthases and syalil-transferases genes [46-48] to obtain humanized complex N-linked glycosylation protein patterns, there are still unwanted toxicological issues that need to be overcome.

There is only one approved biopharmaceutical product containing recombinant proteins from infected insect cell line Hi Five, Cervarix, consisting on recombinant papillomavirus C-terminal truncated major capsid protein L1 types 16 and 18. Nonetheless, this expression system has been extensively used in structural studies since correctly folded eukaryotic proteins can be obtained in a secreted form in serum free media which enormously simplifies protein capture in purification protocols.

Hybridoma cell lines

Hybridomas are fusion cells of murine origin (B-cells and myeloma tumour cells) that are able to express specific monoclonal antibodies against a determined antigen, thus possessing therapeutic potential [49]. Clone selection may account for the progressive enrichment of cells displaying a glycosylation profile with reduced potency and undesirable immunogenic reaction since the human immune system recognizes mouse antibodies as foreign.

Genetic engineering has been applied to obtain humanized monoclonal antibodies using either recombinant mammalian cells producing chimeric antibodies or genetically modified mice to produce human-like antibodies [49]. One such product, Remicade, which binds tumour necrosis factor-alpha, is a pharmaceutical blockbuster used in the treatment of Crohn's disease.

Hamster cell lines

Most of the therapeutic proteins approved so far have been obtained using transgenic hamster cell lines, namely 49 in chinese hamster ovary cells (CHO) and 1 in baby hamster kidney cells (BHK) (Additional file 1). The main advantage of this expression system is that cells can be adapted to grow in suspension in serum free media

(SFM), protein-free and chemically defined media [50]. This fact increases the biosafety of final products reducing risk of introducing prions of bovine spongiform encephalopathy (BSE) from bovine serum albumin and of infectious variant Creutzfeldt-Jakob Disease (vCJD) from human serum albumin. In addition, recombinant products can be secreted into the chemical defined media, which simplifies both upstream and downstream purification process [51]. PTMs in this expression system are almost the same as in human cell lines, although some concerns about comparability in the glycosylation pattern have arisen when comparing different batches of the same manufacturer product and biosimilars [52]. Further development of chemically defined media and fine description of growth conditions would help to overcome this issue.

Human cell lines

In the recent years, three therapeutic proteins produced in human cell lines have been approved, namely Dynepoerithropoietin, Elaprase-irudonate-2-sulfatase and Replagal-alfa-galactosidase A. These products are fully glycosylated human proteins, so this expression system should be addressed when heavily glycosylation is needed. In general, recombinant biopharmaceuticals obtained from mammalian cells cover a wider spectrum of pathological conditions than those obtained from microbes, and the distribution of applications is less biased than when observing products from *E. coli* or *S. cerevisae* (Figure 2).

Transgenic animals

Transgenic animals (avian and mammals), have been successfully used for the production of recombinant proteins secreted into egg white and milk respectively. Protein production using transgenic farm animals supposes a great biotechnological challenge in terms of safety concerns such as transmission of infectious diseases (including viral and prion infections) or adverse allergenic, immunogenic and autoimmune responses. In 2006, ATryn was the first and so far single approved rDNA biopharmaceutical using transgenic animals and validated manufacturer technology platform. It contains human antithrombin (432 amino acids) with 15% glycosylated moieties and is secreted into the milk of transgenic goats. Another product obtained from the milk of transgenic rabbits (Rhucin) has been recently denied for its approval by the EMEA although more tests of repeated treatment are underway to try again its approval. Despite such limited progress, if pharmacovigilance after patient treatment does not reveal any adverse side effects, we might envisage, in the next years, an increase in the approval rate of recombinant protein products from transgenic animal origin.

Alternative, non microbial systems for forthcoming products

As previously discussed, recombinant DNA biopharmaceuticals obtained from bacterial, yeast or mammalian cell culture bioreactors are quite effective as therapeutic agents although production costs are relatively high. One way to address the economic-cost benefit hurdle is through the use of transgenic organisms to manufacture biopharmaceuticals. Biopharming would dramatically reduce the cost of recombinant therapeutic proteins not only in the initial construction of production facilities but also the scale-up process and the final recombinant protein yield. Nonetheless, the fact that regulatory guidelines are being developed as the same time that the establishment of protein production processes is creating uncertainty within biotechnological companies to fulfil drug administration requirements.

Transgenic plants have been used as recombinant protein producers for research and diagnostic uses due to the advantageous low cost of cultivation, high mass production, flexible scale-up, lack of human pathogens and addition of eukaryotic PTMs. The first recombinant protein product obtained from transgenic tobacco was human growth hormone [53] and since then, many other products have been obtained (including antibodies, the surface antigen of the Hepatitis-B-Virus, industrial enzymes and milk proteins). Again, the main disadvantage is related to the plant specific PMTs introduced in recombinant proteins which produce adverse immune responses. Moreover, the possibility to spread the proteins in open fields and the negative public perception of the transgenic plants precludes the use of plants as an attractive expression system of therapeutic proteins.

Host comparative trends in rDNA biopharmaceutical approval

As mentioned above, human insulin produced in *E. coli* was the first rDNA pharmaceutical approved for use, which was followed by a progressively increasing number of other protein drugs from bacteria and yeast (Figure 3). Since 1995, the progression of products of mammalian origin was noticeable and extremely regular, and quantitatively comparable to that of microbial products. Importantly, the incorporation of mammalian cells as factories for rDNA pharmaceuticals has neither represented an excluding alternative to microbial hosts nor resulted in a decrease in the approval rate of microbial products (Figure 3). This is probably due to the extremely different biologically and technologically backgrounds associated to protein production, the good quality of microbial products and the high costs associated to mammalian cell production. In addition, this fact indicates the potential of microbial cells in biopharmaceutical industry despite the limited PTM performance of their products and other bottlenecks as discussed above. Also, microbial cell factory products cover a spectrum of products and application fields that do not necessarily match those addressed by mammalian cell factories (Figure 2).

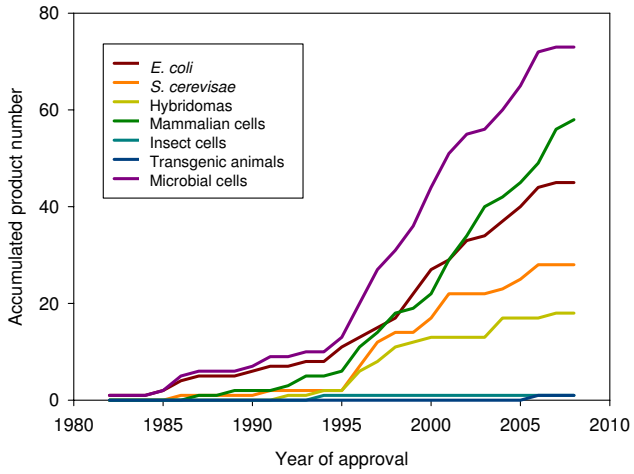

Figure 3
Accumulated number of recombinant biopharmaceuticals obtained in different production systems, in front of year of their first time approval (either in US or EU). Products from *E. coli* and *S. cerevisae* are presented together under the category of microbial cells.

Interestingly, a plateau in the rate of rDNA drug approval during the last 2–3 years is becoming perceivable, irrespective of the production system (Figure 3). Although it might be observed as a transient event, this fact seems instead to indicate that the current production systems could be near to the exhaustion regarding their ability to hold the production of complex proteins, protein complexes or the so-called difficult-to-express proteins. Desirably, recent insights about system's biology of recombinant cells and hosts, and specially, arising novel concepts on recombinant protein quality [54-56] and host stress responses [21] would enlarge the possibilities for metabolic and process engineering aiming to the economically feasible production of new, more complex drugs. Indeed, pushed by fast advances in molecular medicine the pharmaceutical industry is urgently demanding improved production systems and novel and cheaper drugs.

Conclusions and future prospects

Overcoming the biological and methodological obstacles posed by cell factories to the production or rDNA pharmaceuticals is a main challenge in the further development of protein-based molecular medicine. Recombinant DNA technologies might have exhausted conventional cell factories and new production systems need to be deeply explored and incorporated into the production pipeline. On the other hand, a more profound comprehension of host cell physiology and stress responses to protein production would necessary offer improved tools (either at genetic, metabolic or system levels) to favour high yield and high quality protein production. Apart from the expected incorporation of unusual mammalian hosts such as transgenic animals or plants, microbial cells appear as extremely robust and convenient hosts, and gaining knowledge about the biological aspects of protein production would hopefully enhance the performance of such hosts beyond the current apparent limitations. In this regard, not only commonly used bacteria and yeasts but unconventional strains or species are observed as promising cell factories for forthcoming recombinant drugs. Their incorporation into productive processes for human pharmaceuticals would hopefully push the trend of marketed products and fulfil the increasing demands of the pharmacological industry.

Abbreviations

(ADME): absorption, distribution, metabolism, and excretion; (BHK): baby hamster kidney cells; (CHO): chinese hamster ovary cells; (EMEA): European Medicines Agency; (FDA): Food and Drug Administration; (IBs): inclusion bodies; (PTMs): post-translational modifications; (rDNA): Recombinant DNA; (VLPs): virus-like particles

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors read and approved the manuscript's content.

Additional material

Additional file 1

Supplemental table
Recombinant drugs approved for use, grouped by producing host types.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1475-2859-8-17-S1.doc]

Acknowledgements

The authors acknowledge the support of MEC, AGAUR and UAB to their research on protein drug development and delivery through grants BIO2007-61194 to AV (MEC), 2005SGR-00956 to AV (AGAUR) and EME2007-08 to NFM (UAB). We also appreciate the support of CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN, promoted by ISCIII), Spain. AV received an ICREA ACADEMIA award from ICREA (Generalitat de Catalunya).

References

1. Manning MC, Patel K, Borchardt RT: **Stability of protein pharmaceuticals.** *Pharm Res* 1989, **6:**903-918.
2. **Human insulin receives FDA approval.** *FDA Drug Bull* 1982, **12:**18-19.
3. Redwan e: **Cumulative updating of approved biopharmaceuticals.** *Hum Antibodies* 2007, **16:**137-158.
4. Gerngross TU: **Advances in the production of human therapeutic proteins in yeasts and filamentous fungi.** *Nat Biotechnol* 2004, **22:**1409-1414.

5. Schmidt FR: **Recombinant expression systems in the pharmaceutical industry.** *Appl Microbiol Biotechnol* 2004, **65:**363-372.
6. Cusano AM, Parrilli E, Marino G, Tutino ML: **A novel genetic system for recombinant protein secretion in the Antarctic Pseudoalteromonas haloplanktis TAC125.** *Microb Cell Fact* 2006, **5:**40.
7. Parrilli E, De VD, Cirulli C, Tutino ML: **Development of an improved Pseudoalteromonas haloplanktis TAC125 strain for recombinant protein secretion at low temperature.** *Microb Cell Fact* 2008, **7:**2.
8. Joosten V, Lokman C, Hondel CA Van Den, Punt PJ: **The production of antibody fragments and antibody fusion proteins by yeasts and filamentous fungi.** *Microb Cell Fact* 2003, **2:**1.
9. Gasser B, Maurer M, Gach J, Kunert R, Mattanovich D: **Engineering of Pichia pastoris for improved production of antibody fragments.** *Biotechnol Bioeng* 2006, **94:**353-361.
10. Gasser B, Mattanovich D: **Antibody production with yeasts and filamentous fungi: on the road to large scale?** *Biotechnol Lett* 2007, **29:**201-212.
11. Meyer V: **Genetic engineering of filamentous fungi – progress, obstacles and future trends.** *Biotechnol Adv* 2008, **26:**177-185.
12. Schirrmann T, Al-Halabi L, Dubel S, Hust M: **Production systems for recombinant antibodies.** *Front Biosci* 2008, **13:**4576-4594.
13. Jenkins N: **Modifications of therapeutic proteins: challenges and prospects.** *Cytotechnology* 2007, **53:**121-125.
14. Walsh G, Jefferis R: **Post-translational modifications in the context of therapeutic proteins.** *Nat Biotechnol* 2006, **24:**1241-1252.
15. Bailon P, Won CY: **PEG-modified biopharmaceuticals.** *Expert Opin Drug Deliv* 2009, **6:**1-16.
16. Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wren BW, Aebi M: **N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli.** *Science* 2002, **298:**1790-1793.
17. Hite PF, Barnes AMJPE: **Exhuberance Over Exubera.** *Clinical Diabetes* 2006, **24:**110-114.
18. Makrides SC: **Strategies for achieving high-level expression of genes in Escherichia coli.** *Microbiol Rev* 1996, **60:**512-538.
19. Liao YD, Jeng JC, Wang CF, Wang SC, Chang ST: **Removal of N-terminal methionine from recombinant proteins by engineered E. coli methionine aminopeptidase.** *Protein Sci* 2004, **13:**1802-1810.
20. Gräslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R, Ming J, dhe-Paganon S, Park HW, Savchenko A, Yee A, Edwards A, Vincentelli R, Cambillau C, Kim R, Kim SH, Rao Z, Shi Y, Terwilliger TC, Kim CY, Hung LW, Waldo GS, Peleg Y, Albeck S, Unger T, Dym O, *et al.*: **Protein production and purification.** *Nat Methods* 2008, **5:**135-146.
21. Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E, Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A: **Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview.** *Microb Cell Fact* 2008, **7:**11.
22. Prouty WF, Karnovsky MJ, Goldberg AL: **Degradation of abnormal proteins in Escherichia coli. Formation of protein inclusions in cells exposed to amino acid analogs.** *J Biol Chem* 1975, **250:**1112-1122.
23. Martinez-Alonso M, Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A: **Learning about protein solubility from bacterial inclusion bodies.** *Microb Cell Fact* 2009, **8:**4.
24. Villaverde A, Carrio MM: **Protein aggregation in recombinant bacteria: biological role of inclusion bodies.** *Biotechnol Lett* 2003, **25:**1385-1395.
25. Marston FA: **The purification of eukaryotic polypeptides synthesized in Escherichia coli.** *Biochem J* 1986, **240:**1-12.
26. Jana S, Deb JK: **Strategies for efficient production of heterologous proteins in Escherichia coli.** *Appl Microbiol Biotechnol* 2005, **67:**289-298.
27. Sorensen HP, Mortensen KK: **Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.** *Microb Cell Fact* 2005, **4:**1.
28. Nahalka J, Vikartovska A, Hrabarova E: **A crosslinked inclusion body process for sialic acid synthesis.** *J Biotechnol* 2008, **134:**146-153.
29. García-Fruitós E, González-Montalbán N, Morell M, Vera A, Ferraz RM, Arís A, Ventura S, Villaverde A: **Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins.** *Microb Cell Fact* 2005, **4:**27.
30. Garcia-Fruitos E, Aris A, Villaverde A: **Localization of functional polypeptides in bacterial inclusion bodies.** *Appl Environ Microbiol* 2007, **73:**289-294.
31. Nahalka J: **Physiological aggregation of maltodextrin phosphorylase from Pyrococcus furiosus and its application in a process of batch starch degradation to alpha-D:-glucose-1-phosphate.** *J Ind Microbiol Biotechnol* 2008, **35:**219-223.
32. Nahalka J, Dib I, Nidetzky B: **Encapsulation of Trigonopsis variabilis D-amino acid oxidase and fast comparison of the operational stabilities of free and immobilized preparations of the enzyme.** *Biotechnol Bioeng* 2008, **99:**251-260.
33. Neubauer P, Fahnert B, Lilie h, Villaverde A: **Protein inclusion bodies in recombinant bacteria. In Inclusions in Prokaryotes.** In *Microbiology Monographs* Edited by: Shively JM. Springer; 2006:237-292.
34. Vallejo LF, Rinas U: **Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins.** *Microb Cell Fact* 2004, **3:**11.
35. Lin L: **Betaseron.** *Dev Biol Stand* 1998, **96:**97-104.
36. Redwan e, Matar SM, El-Aziz GA, Serour EA: **Synthesis of the human insulin gene: protein expression, scaling up and bioactivity.** *Prep Biochem Biotechnol* 2008, **38:**24-39.
37. **RCSB Protein Data Bank** 2008 [http://www.rcsb.org].
38. **Target DB** 2008 [http://targetdb.pdb.org/].
39. Graf A, Gasser B, Dragosits M, Sauer M, Leparc GG, Tüchler T, Kreil DP, Mattanovich D: **Novel insights into the unfolded protein response using Pichia pastoris specific DNA microarrays.** *BMC Genomics* 2008, **9:**390.
40. Mattanovich D, Gasser B, Hohenblum H, Sauer M: **Stress in recombinant protein producing yeasts.** *J Biotechnol* 2004, **113:**121-135.
41. Porro D, Sauer M, Branduardi P, Mattanovich D: **Recombinant protein production in yeasts.** *Mol Biotechnol* 2005, **31:**245-259.
42. Branduardi P, Smeraldi C, Porro D: **Metabolically engineered yeasts: 'potential' industrial applications.** *J Mol Microbiol Biotechnol* 2008, **15:**31-40.
43. Gurramkonda C, Adnan A, Gäbel T, Lünsdorf H, Ross A, Nemani SK, Swaminathan S, Khanna N, Rinas U: **Simple high-cell density fed-batch technique for high-level recombinant protein production with Pichia pastoris: Application to intracellular production of Hepatitis B surface antigen.** *Microb Cell Fact* 2009, **8:**13.
44. Singh S, Gras A, Fiez-Vandal C, Ruprecht J, Rana R, Martinez M, Strange PG, Wagner R, Byrne B: **Large-scale functional expression of WT and truncated human adenosine A2A receptor in Pichia pastoris bioreactor cultures.** *Microb Cell Fact* 2008, **7:**28.
45. Sinclair AM, Elliott S: **Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins.** *J Pharm Sci* 2005, **94:**1626-1635.
46. Aumiller JJ, Hollister JR, Jarvis DL: **A transgenic insect cell line engineered to produce CMP-sialic acid and sialylated glycoproteins.** *Glycobiology* 2003, **13:**497-507.
47. Hollister J, Grabenhorst E, Nimtz M, Conradt H, Jarvis DL: **Engineering the protein N-glycosylation pathway in insect cells for production of biantennary, complex N-glycans.** *Biochemistry* 2002, **41:**15093-15104.
48. Hollister JR, Jarvis DL: **Engineering lepidopteran insect cells for sialoglycoprotein production by genetic transformation with mammalian beta 1,4-galactosyltransferase and alpha 2,6-sialyltransferase genes.** *Glycobiology* 2001, **11:**1-9.
49. Nissim A, Chernajovsky Y: **Historical development of monoclonal antibody therapeutics.** *Handb Exp Pharmacol* 2008:3-18.
50. Bleckwenn NA, Shiloach J: **Large-scale cell culture.** *Curr Protoc Immunol* 2004, **Appendix 1:**. Appendix
51. Mohan C, Kim YG, Koo J, Lee GM: **Assessment of cell engineering strategies for improved therapeutic protein production in CHO cells.** *Biotechnol J* 2008, **3:**624-630.
52. Yuen CT, Storring PL, Tiplady RJ, Izquierdo M, Wait R, Gee CK, Gerson P, Lloyd P, Cremata JA: **Relationships between the N-glycan structures and biological activities of recombinant human erythropoietins produced using different culture conditions and purification procedures.** *Br J Haematol* 2003, **121:**511-526.
53. Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll JA, Spatola L, Ward D, Ye G, Russell DA: **High-yield production of a human therapeutic protein in tobacco chloroplasts.** *Nat Biotechnol* 2000, **18:**333-338.

54. Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A: **Recombinant protein solubility-does more mean better?** *Nat Biotechnol* 2007, **25:**718-720.
55. de Marco A: **Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in Escherichia coli.** *Nat Protoc* 2007, **2:**2632-2639.
56. de Marco A: **Minimal information: an urgent need to assess the functional reliability of recombinant proteins used in biological experiments.** *Microb Cell Fact* 2008, **7:**20.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

BioMedcentral

Article 2

Engineered Biological Entities for Drug Delivery and Gene Therapy: Protein Nanoparticles

Domingo-Espin, J., Unzueta, U., Saccardo, P., Rodriguez-Carmona, E., Luis Corchero, J., Vazquez, E., & Ferrer-Miralles, N.

Prog Mol Biol Transl Sci. 104, 247-298 (2011).

In this book chapter we describe the different biopharmaceutical entities for drug and gene delivery from the point of view of those composed only by proteins.
In this chapter, the main available strategies to develop protein-based nanovehicles or biopharmaceuticals are described. In this context, several parameters are defined such as proper formulation, stability, immunogenicity, and delivery to the correct cell type and cell compartment. Modular protein engineering (which is the focus of this thesis), virus-like particles (VLPs), and other self-assembling entities are envisioned as modulable novel protein nanoparticles able to include many desirable properties for the correct delivery of drugs and nucleic acids. Finally, some successful examples of protein nanoparticles on the market are described in addition to protein products currently in clinical trials and under preclinical research, in order to envision which type of protein nanoparticles will be available soon on the market.

Engineered Biological Entities for Drug Delivery and Gene Therapy: Protein Nanoparticles

Joan Domingo-Espín,[*,†,‡] Ugutz Unzueta,[*,†,‡] Paolo Saccardo,[*,†,‡] Escarlata Rodríguez-Carmona,[*,†,‡] José Luís Corchero,[*,†,‡] Esther Vázquez,[*,†,‡] and Neus Ferrer-Miralles[*,†,‡]

[*]*Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*

[†]*Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*

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

Progress in Molecular Biology and Translational Science, Vol. 104
DOI: 10.1016/B978-0-12-416020-0.00006-1

Copyright 2011, Elsevier Inc.
All rights reserved.

1877-1173/11 $35.00

The development of genetic engineering techniques has speeded up the growth of the biotechnological industry, resulting in a significant increase in the number of recombinant protein products on the market. The deep knowledge of protein function, structure, biological interactions, and the possibility to design new polypeptides with desired biological activities have been the main factors involved in the increase of intensive research and preclinical and clinical approaches. Consequently, new biological entities with added value for innovative medicines such as increased stability, improved targeting, and reduced toxicity, among others have been obtained. Proteins are complex nanoparticles with sizes ranging from a few nanometers to a few hundred nanometers when complex supramolecular interactions occur, as for example, in viral capsids. However, even though protein production is a delicate process that imposes the use of sophisticated analytical methods and negative secondary effects have been detected in some cases as immune and inflammatory reactions, the great potential of biodegradable and tunable protein nanoparticles indicates that protein-based biotechnological products are expected to increase in the years to come.

I. Introduction

The design of new chemical entities (NCE) for diagnosis and treatment of human diseases has relied on the discovery of active chemical drugs from a diverse library of compounds or from naturally occurring molecules.[1,2] Further chemical modifications improve pharmacokinetic properties to obtain a final product with a known mechanism of action and decreased toxicity.[3] Nonetheless, using such approaches, the final products present low specificity for their target molecules, interacting with many other molecules and accumulating in some tissues, disturbing the correct homeostasis of the system. In some cases, the adverse effects of drug administration exceed pharmacological effect and despite the concise mechanism of action of the drug over the target molecule representing an improvement in the patient's state, the treatment has to be prevented or discontinued.[4] In fact, although a maintained steady increase in the number of launched NCE has been observed in the last years, the question arises whether this classical approach has already exhausted the discovery of innovative molecules.[5]

On the other hand, macromolecular new biological entities (NBE) have been used to supplement cellular deficiencies or to inhibit cellular pathways exploiting their relatively specific mode of action. Proteins and peptides have been obtained first from their natural source or produced as recombinant

versions after the development of genetic engineering techniques in the late 1970s. However, the delivery of biological entities is sometimes hampered by its low half-life in the bloodstream by unspecific degradation, resulting in an expensive and ineffective process. Nevertheless, some solutions have already been explored for biopharmaceuticals to increase solubility and stability and to reduce immunogenicity including postranslational modifications such as glycosylation and covalent conjugation of polyethylene glycol.[6]

Thus, one of the main objectives in the use of drugs (for either NCE or NBE) is the need to optimize the delivery system to reduce the pharmacological dose which would consequently represent a concomitant reduction in toxicity and cost. In that scenario, new delivery approaches have been implemented using biological interactions such as antigen–antibody binding (immunoliposomes)[7] or more sophisticated interactions including the binding between nutrient concentrator SPARC (secreted protein acidic and rich in cysteine) and albumin in the treatment of some types of cancer (Abraxane®).[8,9]

Proteins can be then used for their targeting qualities as molecular delivery vehicles both for the specific delivery of drugs or nucleic acids in gene therapy approaches and by themselves as therapeutic molecules. One of the interesting characteristics of proteins is their ability to form intermolecular driven complexes as sophisticated and structurally perfect as in the case of viral capsids. In addition, through the use of genetic engineering, recombinant proteins can be tuned to include additional properties to optimize drug delivery and nucleic acid delivery in gene therapy.

In this chapter, the main available strategies to develop protein-based nanovehicles or biopharmaceuticals will be described. In this context, several parameters will be defined such as proper formulation, stability, immunogenicity, and delivery to the correct cell type and cell compartment. Modular protein engineering, virus-like particles (VLPs), and other self-assembling entities are envisioned as modulatable novel protein nanoparticles able to include many desirable properties in the correct delivery of drugs and nucleic acids. Finally, some successful examples of protein nanoparticles on the market will be described in addition to protein products currently in clinical trials and under preclinical research in order to envision which type of protein nanoparticles will be available soon on the market.

II. Protein Nanoparticle Formulation and Biological Barriers

When a protein-only nanoparticle is meant to be used as a vector to deliver therapeutic nucleic acid, drug, or peptide, there are several steps that the nanoparticle has to perform to successfully get inside the target cell. In the first instance, it is necessary to obtain the proper formulation of the complex

with the therapeutic molecule to generate a vehicle capable of being transported in the blood if a systemic administration is needed and retaining a significant stability before reaching the target cell.[10,11] In addition, the biological system poses specific barriers that have to be overcome such as membranes (cytoplasmic, endocytic, and nuclear), degradation (protease degradation induced by acid denaturalization in lysosomes, cytosolic proteosomes, and nucleases), cytosolic transport, and nuclear entry if necessary.[12,13] For central nervous system therapies, the blood–brain barrier (BBB) represents the main bottleneck, and for that, a specific strategy has to be designed.[14] Furthermore, the therapeutic complex has to be flexible enough in order to release the therapeutic molecule in the specific cell compartment.

Thus, several protein motifs have been described to overcome each and every process described earlier so that a modular multifunctional protein can be generated including those modules that are necessary to achieve its goal. In order to get a rational construction of the multifunctional vector, each step has to be carefully taken into account so as to overcome every step which is needed to achieve its final goal (Table I).

A. Interaction with Drugs and Nucleic Acids

The DNA/RNA condensation or drug interaction with the protein vector is a critical step in the formulation of protein nanoparticles for gene therapy. They have to remain attached to the vector during the whole transport process through the body and the cell until it can be released in the desired localization within the target cell. Highly positively charged peptides containing a large number of arginines or polylysines have been used to promote electrostatic interactions since nucleic acids are highly negatively charged molecules.[15–22] Natural DNA-condensing proteins as nuclear histidines or protamines can also be used to bind nucleic acids.[22–25] Protamine, which is the protein that replaces histidines during the spermatogenesis process, is a sperm chromatin component and just as the histidines do, it has very high DNA condensation ability to protect nucleic acids form cytosolic endonucleases.[23,26] In addition, as soon as the complex reaches the cellular nucleus, protamine is degraded by chromatin-remodeling proteins, releasing the transported DNA allowing its expression.[15,23] In contrast, polycationic DNA condensation modules such as polylysines and polyarginines—even they can present higher DNA condensation ability depending on the polycationic chain length—usually present lower DNA-releasing ability, interfering negatively with the accessibility of cellular transcription factors and DNA expression capacity.[15]

All these DNA condensation modules described above interact with any DNA that is incubated in an unspecific way. However, there are proteins such as GAL4 that are able to recognize specific DNA sequences[27–29] and that permit to bind and condensate specific DNA sequences in the final vector.[30,31]

TABLE I

Selection of Peptide Motifs Used in Gene Therapy and Drug Delivery to Improve Protein Nanovehicle Performance

Peptide motif	Sequence	References
Nucleic acid condensation peptides		
Polylysine	$(KKKKKKKKKKKKKKKKKK)_n$	16–18
Polylysine containing peptides	YKAKKKKKKKKWK and derivatives	19–22
Salmon protamine	PRRRRSSSRPVRRRRRPRVSRRRRRRGGRRRR	23–25
GAL4	MKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPK	30,31
Blood–brain barrier (BBB) peptides		
g7	H2N-Gly-L-Phe-D-Thr-Gly-L-Phe-L-Leu-L-Ser(O-β-D-glucose)-$CONH_2$	107
RVG	YTIWMPENPRPGTPCDIFTNSRGKRASNG	56
Tat	YGRKKRRQRRR	108
R9	RRRRRRRRR	14
Cell-penetrating peptides (CPP)		
Tat	GRKKRRQRRPPQ	36–41
R9	RRRRRRRRR	42
Penetratin	CRQIKIWFQNRRMKWKK	43,44
bPrPp	MVKSKIGSWILVLFVAMWSDVGLCKKRPKP	43
Transportan	CLIKKALAALAKLNIKLLYGASNLTWG	44–46
Receptor-specific ligands (ligand/receptor)		
RGD/integrins (mainly αvβ3)	GRGDSP	47,48
CXCL12/CXCR4	KPVSLSYRCPCRFFESHVARANVKHLKILNTPNCALQIVARLKNNNRQVCIDPKLKWIQEYLEKALN	49,50
Transferrin receptor ligand (12Aa)/transferrin receptor	THRPPMWSPVWP	51,52
EGF/EGF receptor	NPVVGYIGERPQYRDL	53,54

(*Continues*)

TABLE I (*Continued*)

Peptide motif	Sequence	References
Asioaloglycoprotein/ asioaloglycoprotein receptor		55
RVG/acetil-colin receptor	YTIWMPENPRPGTPCDIFTNSRGKRASNG	56
PLAEIDGIELTY/integrin a9b1	PLAEIDGIELTY	57
Molossin (RGD)/integrin	ICRRARGDNPDDRCT	58
Secretin/Secretin receptor	HSDGTFTSELSRLRDSARLQRLLQGLV	59
NL4 (loop 4 of nerve growth factor)/TrkA	CTTTHTFVKALTMDGKQAAWRFIRIDTAC	60
Neurotensin/Neurotensin receptor (NTRH)	ELYENKPRRPYIL	61
LSIPPKA, FQTPPQL, LTPATAI/LOX-1	LSIPPKA, FQTPPQL, LTPATAI	62
Monoclonal Abs/antigen recognized by the antibody	–	63–65
Endosomal escape fusiogenic peptides		
HA-2	GLFGAIAGFIENGWEGMIDGWYG	12,69,70
GALA	WEAALAEALAEALAEHLAEALAEALEALAA	12,71–74
KALA	WEAKLAKALAKALAKHLAKALAKALKACEA	12,75
JTS-1	GLFEALLELLESLWELLLEA	12,19
ppTG20	GLFRALLRLLRSLWRLLLRA	12,76
PPTG1	GLFKALLKLLKSLWKLLLKA	12,76
Melittin	GIGAVLKVLTTGLPALISWIKRKRQQ	12,77
Tat	GRKKRRQRRRPPQ	12,39,40
Penetratin	RQIKIWFQNRRMKWKK	12,78,79
Transportant	GWTLNSAGYLLGKINLKALAALAKKIL	12,45,46
INF 7	GLFEAIEGFIENGWEGMIDGWYG	12,80

Endosomal escape histidine-rich peptides		
CHK6HC	CHKKKKKKHC	12,22
H5WYG	GLFHAIAHFIHGGWHGLIHGWYG	12,80–82
LAH4	KKALLALALHHLAHLALHLALALKKA	12,83
Nuclear import peptides		
SV40 large T antigen	PKKKRKV	36
Tat	VIH transcription factor	37
EBNA-1	Epstein–Barr virus	91
Melittin	Honeybee venom (*Apis mellifera*)	77
M1 (c-myc transcription factor)	PAAKRVKLD	92
M2 (c-myc transcription factor)	RQRRNELKRSP	92
GAL4 amino terminal domain	Transcription factor	93
Protamines	Sperm DNA condensation protein	23
Histone H1	Nuclear DNA condensation protein	94,95
M9 (heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1)	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY	96
Vp3	SV40 estructural protein Vp3	97
Adenovirus E1 protein C-terminus	KRPRP	98
Xenopus N1 protein	VRKKRKTEEESPLKDKDAKKSKQE	99
Fibroblast growth factor 3 (FDF3)	RLRRDAGGRGGVYEHLGGAPRRRK	100
Poly ADP-ribose polymerase (PARP)	KRKGDEVDGVDECAKKSKK	101
Xenopus protein nucleoplasmine	KRPAATKKAGQAKKKK	102

B. Protein Stability in Serum

In many cases, the multifunctional protein vector is *in vivo* administrated by the systemic route in order to travel in the blood and reach the target cells. That exposes the vector to all blood components, making it susceptible to be degraded. Thus, it is completely necessary that the vector remain in the blood long enough to be able to reach the target cells. It has also been described that naked DNA has an estimated half-life in blood of minutes[10]; so protein nanovehicles in gene therapy, among other properties, are intended to protect nucleic acids from degradation.

One important factor when the vector is exposed to the blood is that it can be recognized by the immune system components and produces an immune response against the vector. Thus, it is also very important to try to make the vector as less antigenic as possible in order to avoid being degraded or even being toxic to the organism.[32]

C. Defeating Biological Barriers

1. Cell Binding and Internalization

Peptide uptake or internalization involves a step before the protein binding to the cell surface. This attachment can be either specific or unspecific but in all cases the promotion of its internalization is required.[33]

Positively charged peptides usually bind the cellular surface by unspecific electrostatic interactions with the negatively charged cell surface proteoglicans. This kind of peptides can be used in the multifunctional protein if specific targeting is not required.[33] Cell-penetrating peptides (CPPs) have been widely described as unspecific cell-binding and internalization peptides[34–46] (see also the chapter "Peptide Nanoparticles for Oligonucleotide Delivery" by Lehto *et al.* in this volume). However, specific interactions can be obtained by incorporating cell receptor ligands if cell or tissue targeting is required for the therapeutic action. Moreover, some of those ligand–receptor interactions promote the ligand–receptor complex internalization. Many peptides have been described in the literature as receptor-specific ligands so any of them can be added to the multifunctional proteins in order to confer them cell specificity.[47–62] The most natural specific ligands that can also be used for cell targeting are monoclonal antibodies.[32,63–65] In addition, if no specific peptides are available for an intended target, new specific binding peptides can be found by using phage display[66] or combinatorial chemistry.[67]

2. Endosomal Escape

Several internalization pathways are possible depending on the vector properties,[27,33] including endocytosis (clathrin/caveolae-mediated, clathrin/caveolae-independent), macropinocytosis, and non-endocytic pathways.

It is known that more than one internalization pathway can be performed at the same time but usually the peptide-based vector uses endocytic pathways.[68] Moreover, it seems that proteins that interact with a specific cellular receptor are internalized by the clathrin-mediated endocytic pathway.[33] Most of the generated endosomal vesicles will converge to late endosomes that eventually will fuse with cellular lysosomes.[15,33] Remaining in the cellular endosomes, the multifunctional protein will be degraded, so it is strictly necessary that the internalized multifunctional proteins be released into the cellular cytoplasm escaping from degradation.

Several peptides have been described that are able to promote endosomal escape and can be classified into two types depending on their escape mechanism: fusiogenic peptides and histidine-rich peptides.[36] The fusiogenic peptides are small peptides that have hydrophobic amino acids (Aa-s) interspersed at constant intervals with negatively charged Aa-s.[12,19,39,40,45,46,69–80] Thus, when early endosomes become late endosomes, their low pH induces a conformational change in the peptide, which adopts a alpha-helix structure, in an amphipathic structure able to fuse with the endosomal membrane, leading to pore formation and releasing all the endosomal content into the cell cytoplasm.[36] The histidine-rich peptides are small peptides with a high histidine content whose endosmolytic activity is mediated by a mechanism called "proton sponge".[12,22,80–83] When the endosomal pH becomes low in late stages, the imidazole groups of the histidines are protonated and attract endosomal Cl^- ions, buffering against the proton pump. Thus, the endosomes collapse by an osmolytic swelling process and the endosomal content is released to the cell cytoplasm.[36] Further details are given in the chapter "Peptide Nanoparticles for Oligonucleotide Delivery" by Lehto *et al.* in this volume.

3. Vector Stability in the Cytosol

Once the protein has achieved the cellular cytosol, it can be degraded by cellular proteases or by the cellular proteosome system.[84] It is important to avoid this process, especially if the protein has to reach the cellular nucleus. If the final target of the nanoparticle is the cellular cytoplasm, it is necessary that it remain there at least long enough to perform its therapeutical action.

Several peptide proteosome inhibitors have been described that are able to avoid this type of protein degradation. By adding these peptides to the final protein vector it is possible to protect it and enhance cytoplasmatic stability. Epstein–Barr virus nuclear antigen 1 (EBNA1) contains a proteosome inhibitor consisting of glycine–alanine repeats able to prevent proteosomal proteolysis. It has been shown that a minimum of 4 Aa-s Gly-Ala repeats are necessary to achieve such protective activity.[85–87] If the protein vector is carrying nucleic acids (DNA or RNA), degradation by the cytosolic endonucleases has to be taken into account, so it is also very important to protect this nucleic acid in

order to maintain its integrity. Some DNA/RNA condensing peptides as protamines also protect the DNA against cytoplasmic endonucleases and enhance its stability as has been described above.[15]

4. Intracytosolic Mobility

The cellular cytoplasm is a very crowded and compartmentalized environment where cellular organelles and cytoskeleton make the free diffusion of macromolecules such as protein vectors difficult. However, cytoskeleton elements such as microtubules are used by endosomes and other cytosolic macromolecules for intracytosolic mobility.[33] Dyneins have been described as being capable of carrying those macromolecules and endosomes along the microtubules in a retrograde transport toward the nucleus. Some small peptides that are able to bind dyneins have been identified. They can be added to the multifunctional protein vector in order to mediate an intracytosolic mobility toward the cellular nucleus.[36] Several dynein-binding proteins have been identified in viruses that are able to use this transport system. Comparing those protein sequences, a consensus peptide sequence (KSTQT) that is able to bind to the dynein LC8 light chain has been identified.[88]

5. Nuclear DNA Delivery and Expression

Molecules lower than 45 kDa/10–30 nm are able to enter in the cellular nucleus by passive diffusion. However, macromolecules higher than 45 kDa/10–30 nm generally require an active transport system through the nuclear pore system. This transport mechanism generally requires a specific targeting signal peptide named nuclear localization signal (NLS). These signaling peptides are usually rich in basic Aa-s, which are recognized by the cellular importines and actively transported through the nuclear pore.[15,89] Monopartite or bipartite NLS sequences which are NLS peptides that have one or two NLS recognized sequences respectively have been described.[12] Thus, these peptidic sequences can be added into the final multifunctional protein if nuclear localization is required in order to express a carried DNA. It has been reported that a single NLS sequence is sufficient to transport the vector to the nucleus and that a large number of NLS sequences can result in inhibition of its activity.[90]

One of the most used NLS signal peptides are fragments derived from the 111–135 Aa-s of the simian virus SV40 large tumor antigen (T-ag). Other NLS sequences can be found in GAL4, protamines, or Tat.[23,36,37,77,91–102]

It is important that when the transported DNA reaches the cellular nucleus, it has to be released in order to be accessible to the nuclear transcription factors and achieve the desired expression level. Thus, while designing the multifunctional protein vector, this aspect has to be taken into account.

Once the DNA has been released in the cell nucleus, it will be necessary to control its expression level depending on which therapeutic action is being promoted. When the goal is to kill a cell as in cancer therapies, the uncontrolled DNA expression levels would not be a problem. However, when a specific protein expression level is required, achieving good control is very important.[13] Some expression systems have been developed that can be pharmacologically regulated by oral drug formulation.[103] Cell-specific promoters and enhancers can be also used in order to confer high cell specificity to the therapy.[104,105]

D. Ways to Get Over the BBB

The BBB is a hermetic barrier that only allows nonlipophilic molecules smaller than 400 Da to cross it. However, some human proteins such as insulin, transferrin, insulin-like growth factor, or leptins are able to go across it by receptor-mediated transporters. Thus, the most important factor limiting central nervous system-targeting therapeutics is the presence of the BBB.[106] Finding the way to cross it will be the main challenge.

Some peptides have been described that are able to reach the brain crossing the BBB. Moreover, it has been seen that they can be associated with another molecule and transported through the barrier. Thus, they could be interesting candidates to be included in the multifunctional vectors if central nervous system targeting is required.[14,56,107,108]

Antibodies have also been described that bind transferrin and insulin receptors and that are able to cross the BBB efficiently. They can be conjugated with large molecules, allowing its translocation to the central nervous system.[63,64,109–111]

III. Multifunctional Proteins

A. Protein Engineering: Direct Evolution, *De Novo* Synthesis, and Rational Design

The development of genetic engineering techniques has increased the natural repertoire of proteins for the design of useful and/or valuable proteins with the aim to obtain new proteins with desired functions. There are three main strategies leading to the construction of engineered proteins: (a) direct evolution, (b) *de novo* protein design, and (c) rational design.

Directed evolution has developed quickly to become a method of choice for protein engineers in order to create enzymes having desired properties for all kind of processes. Over the past decade, this technique has become a daily part of the molecular toolbox of every biochemist. This is emphasized by the increasing number of publications about the subject.[112]

In nature, evolution and creation of new functionalities is achieved by mutagenesis, recombination, and survival of the fittest. Directed evolution mimics this and is a process of iterative cycles of producing mutants and finding the mutant with the desired properties. Mutations can be introduced at specific places using site-directed mutagenesis or throughout the gene by random mutagenesis. Several mutagenesis techniques have been developed in order to avoid codon bias.[113,114] The first technique used to mimic evolution was DNA shuffling.[115] This method is based on the mixing and subsequent joining of different related small DNA fragments in order to form a complete new gene. In the process of shuffling, the recombination frequency is dependent on the degree of homology. A high level of recombination is important to get all possible combinations of mutations. Since recombination can be biased, several methods to overcome problems arising from the use of shuffling in the early years were tackled by novel strategies, all having their own advantages and disadvantages.[112] The products obtained by these methods have to be screened for desired qualities and not all of them can be easily screened.

De novo protein design offers the broadest possibility for new structures. It is based on searches for amino acid sequences that are compatible with a three-dimensional protein backbone template using *in silico* techniques. Several research groups in the field have applied *in silico* methods to design the hydrophobic cores of proteins, with the novel sequences being validated with experimental data.[116] *In silico* protein design has allowed novel functions on templates originally lacking those properties, modifying existing functions, and increasing protein stability or specificity. Beyond any doubt, intense research activities are ongoing in the field, the potential of which is simply enormous.[117] So far there have been numerous examples of full sequences designed "from scratch" that were confirmed to fold into the target three-dimensional structures by experimental data.[118] The zinc-finger protein designed by Dahiyat and Mayo[119] was the first one to appear by this method.

Rational design of proteins is based on the modification or insertion of selected amino acids or domains in a polypeptide chain backbone to obtain proteins with new or altered biological functions. When using that strategy, a detailed knowledge of the structure and function of the backbone protein is needed to make desired changes. This generally has the advantage of being inexpensive and technically feasible. However, a major drawback of this approach is that detailed structural knowledge of a protein is often unavailable or it can be extremely difficult to predict the effects of various mutations. Modular engineering enables, by using simple DNA recombinant techniques, the construction of chimerical polypeptides in which selected domains, potentially from different origins, provide the required activities. An equilibrate combination and spatial distribution of such partner elements has generated promising

prototypes, able to deliver expressible DNA or molecules to tissue culture but also to specific cell types in whole organisms.[120] Modular fusion proteins that combine distinct functions required for cell type-specific uptake and intracellular delivery of DNA or drugs present an attractive approach for the development of self-assembling vectors for targeted gene or drug delivery.[121] One of the first examples was described by the group of Uherek *et al.* They combined a cell-specific target module (antibody fragment specific for the tumor-associated ErbB2 antigen), a DNA-binding domain (Gal4), and a translocation domain for endosomal escape.[121]

In this context, many strategies for the construction of safer vehicles are being explored and the number of nonviral prototype vectors for gene and drug delivery is noticeably increasing. Here, the common steps that an approach like this might explore are presented (Fig. 1).

B. Designing a Protein Nanoparticle

When designing a new protein for drug or gene delivery there are many critical aspects, namely (a) design of the vehicle itself, required functions, stability, etc.; (b) production of the protein, suitable expression system, purification procedure, scaling up process, etc.; (c) characterization of the vehicle by physicochemical and functional tests; and finally (d) the administration route and regulatory guidance for biological products. Although all these aspects belong to different disciplines, they have to be overviewed together. Here, the major needs of a modular protein for gene and drug delivery are presented.

To enhance the physicochemical stability of the cargo molecules and their resistance to nuclease/protease-mediated degradation, protein vehicles should ideally exhibit, like their natural counterparts (viruses), nucleic-acid binding and condensing properties.[27] Such abilities are, in general, conferred by cationic segments of the main scaffold molecules that interact with nucleic acids, mainly through electrostatic interactions. In addition, such complexes need to efficiently release the nucleic acid in the nucleus (if the cargo is a therapeutic gene), for which endosomal escape is required. Such functions have been found in some peptides in many natural molecules and they are suitable for functionalizing protein vehicles.

The ability to bind a particular cell type with high specificity is especially significant in a systemic delivery in which appropriate biodistribution and tissue targeting are essential.[122] For nuclear targeting, only naked short nucleic acids can freely enter the nucleus of nondividing cells via free diffusion through the nuclear pore. Large molecules require active transport mediated by NLSs that are often found in viral proteins. Because the molecular mass of plasmidic DNA varies from to 2 to 10 MDA, DNA that is to be expressed, and essentially

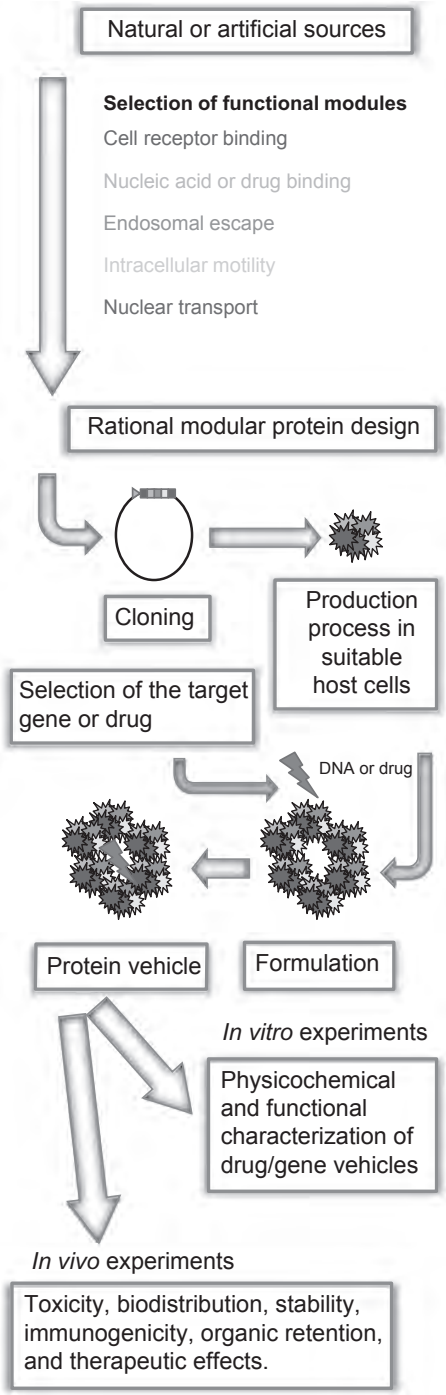

FIG. 1. Scheme in the development of protein nanoparticles for drug delivery and gene therapy.

any macromolecular complex for nucleic acid delivery, requires NLSs.[123] The role and types of functional modules peptides used for all these purposes will be discussed in depth in the following sections.

Finally, which protein or peptide is better for a given cargo is to be determined empirically and only few rules can be taken literally.[38,124]

C. Production of Protein Nanoparticles

Some steps in the production of a protein-based vehicle after molecular cloning such as protein production and protein purification[125] might be experimentally labor intense with a variable success rate. For that reason, when small proteins are needed, solid-phase peptide synthesis[44] guarantees the process. However, the classical procedure of biological production allows scaling up the process in most of the cases and the production of larger polypeptides and full-length proteins.

Generally, in protein nanoparticle approaches, the protein is composed by different modules of natural sources such as the cell-penetrating peptide transactivator of transcription (TAT) derived from the TAT of the human immunodeficiency virus (HIV)[126] or artificial sequences not present in any organism such as the polylysine DNA-condensing sequence.[127]

Once it has been defined which modules will be part of the protein, it is important to define the order they will have in the final construct. It has been demonstrated by Boekle and coworkers using melittin conjugated to polyethylenimine (PEI) that depending on the side of the linkage (C- or N-terminus), the lytic activity could be changed. Some other modules have the need to be in a determined position for its correct function.[128]

When producing a protein for gene or drug delivery, it is important to know the origin of its domains to choose the most suitable expression system for its production. For instance, if any module naturally carries a posttranslational modification that is essential for its biological function, the expression system chosen will have to be able to reproduce the same crucial modification.

The main biological production systems for protein drugs are described below.

Escherichia coli is the most widely used prokaryotic organism for the expression of recombinant proteins.[129] The use of this host is relatively simple and inexpensive.[130] Added advantages include its short duplication time, growth to high cell densities, ease of cultivation, and high yields of the recombinant product. However, since it lacks fundamental prerequisites for efficient secretion, recombinant proteins manufactured by *E. coli* systems are mainly produced as inclusion bodies.[125,131] Moreover, posttranscriptional modifications are not achieved with this system. There are many examples of proteins for gene delivery produced in *E. coli* with probed efficiency.[132,133]

Like *E. coli*, yeasts can be grown cheaply and rapidly and are amenable to high-cell-density fermentations. Besides possessing complex posttranslational modification pathways, they offer the advantage of being neither pyrogenic nor pathogenic and are able to secrete more efficiently. Species established in

industrial production procedures are *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Pichia pastoris*, and *Hansenulapolymorpha*. *S. cerevisiae* is the best genetically characterized eukaryotic organism among them all and is still the prevalent yeast species in pharmaceutical production processes.[131] In spite of their physiological advantageous properties and natively high expression and secretion capacity, the employability of yeasts in some cases, however, might reach a limit, particularly when the pharmacological activity of the product is impaired by the glycosylation pattern. In such cases, either a postsynthetic chemical modification has to be considered or the employment of more highly developed organisms. Most examples of nanoparticles produced in yeast are for VLPs.[134]

Animal cell expression systems show the highest similarity to human cells regarding the pattern and capacity of posttranslational modifications and the codon bias. However, their culture is more complicated and costlier and usually yields lower product titers. Among the known systems, insect cells infected by baculovirus vectors have reached popularity since they are considered to be more stress-resistant, easier to handle, and more productive compared with mammalian systems and are thus frequently employed for high-throughput protein expression. For commercial application, scale-up related questions have to be solved.[135–137] Preferably applied in pharmaceutical production processes are mammalian systems like chinese hamster ovary (CHO) cells and baby hamster kidney (BHK) cells. These systems are genetically more stable and easier to transform and handle in scale-up processes, to grow faster in adherent and submerged cultures, and to be more similar to human cells and more consistent in their complete spectrum of modification.[138] In some cases, mammalian cell systems can be the only choice for the preparation of correctly modified proteins.

Peptides, being complex and unique complex molecules with regard to its chemical and physical properties, can be produced synthetically by the solid-phase method.[139,140] This technology can be used to avoid problems related to biological production. General advantages of synthetic peptides are that they are very stable compounds, solid-phase chemistry produces highly standardized peptides, and the crucial polycation component is provided by a "natural" polycation, thus minimizing toxicity.[141]

However, some disadvantages related to synthetic peptides have been reported such as the difficulty to synthesize long and well-folded oligopeptides, peptides with multiple cysteine, methionine, arginine, and tryptophan residues due to technical limitations or production cost.[141]

D. Physicochemical Characterization

When working with protein nanoparticles, it is very important to characterize them physically and functionally in order to understand their behavior.

The size and charge of protein/cargo particles are crucial properties which influence rates of diffusion, binding to polyanionic components of connective tissues, transversal of anatomical barriers, binding of serum proteins, attachment to cells, and mechanisms of endocytosis, among other factors. Stability in physiological salt solutions is a key issue for *in vivo* delivery, as salt is found everywhere in the body.[141] Mixing a multivalent polycation and DNA results in electrostatic binding of both molecules, with charge neutralization of DNA and a particle formation named conjugate. Charge neutralization can be easily seen by retardation gel assays and particle formation by dynamic light scattering (DLS). DLS is a good method to see particle formation but not to quantify relative number of particles of different sizes.[142]

To visualize particles, many groups have used transmission electron microscopy (TEM)[15,143] with good results while others have used fluid particle image analyzer (FPIA) to photograph individual particles in physiological solutions.[58]

The net charge of protein/cargo particles is an important variable. Generally, optimal gene delivery for cell lines requires a net positive charge but, as stated previously, it has to be determined empirically. One of the best techniques to determine the net charge is by calculating the Zeta potential that measures the electrophoretic mobility of particles.[144]

Despite the fact that physical characterization is a key element, understanding and testing the functionality and pharmacokinetics of a gene or drug is the most important part of its development process. Most of the initial tests are done using cell lines in *in vitro* experiments using reporter genes, RNA, or drugs.[145,146] Quantifying the percentage of transfected cells or drug-induced changes is a very valuable tool to evaluate nanoparticle performance in both nuclear and cytoplasmic delivery, respectively. In addition, *in vitro* experiments may be designed to select a candidate for the *in vivo* experiments from a group of possible therapy vectors.

The quantitative kinetics of particle binding, the molecular basis of particle interactions with target cell membranes, the efficiency of particle internalization, and endosomal escape are all poorly understood.[141]

Interaction of particles with plasma membranes prior to protein internalization can be either unspecific or specific. Untargeted delivery normally is the consequence of electrostatic interactions between anionic ligands in the cell surface and cationic components of the vehicle. On the other hand, targeted delivery to specific membrane molecules is a more sophisticated approach. It aims to improve cell specificity and efficiency, by directing to molecules, only expressed or overexpressed in a particular cell type, that initiate internalization by endocytosis. Targeting moieties include many types of molecules and is discussed afterwards.

Internalization of particles, its mechanisms, and kinetics are not well known and most studies about nanoparticle delivery do not focus on this aspect. There are several endocytic pathways each initiated by different ligands.[147] Enhancing the delivery by addition of chloroquine, a synthetic molecule used primarily for the prophylaxis and treatment of malaria that disrupts endosomes,[148] is an accepted parameter to demonstrate endosomal localization of particles.

Endosomal escape is the area most intensively investigated but is poorly understood. An important practical point to note is that some reagents that are used can be toxic.[141] To enhance this step, anionic fusogenic peptides can be used. These peptides fuse to membranes in an acidic-dependent manner causing its disruption.[149]

In gene delivery approaches, translocation of DNA expression plasmids into the cell nucleus involves an active, energy-dependent process through the nuclear pore complex.[150] Directly injected DNA into the cytosol is usually, but not always, poorly transferred to the nucleus[150,151] and because of that, the use of proteins carrying cationic nuclear-localizing sequences (such as that of SV40 large T antigen) has been widely used to overcome this step.[143]

IV. Natural Self-Assembling Protein Nanoparticles: VLPs

Ideal drug delivery and gene therapy vehicles must accomplish some desired features such as appropriate packaging size for its cargo, target cell-specificity, safe and efficient cargo delivery, and protection against immune recognition, or capability to escape immune recognition. Moreover, these vehicles must avoid inflammatory toxicity and rapid clearance.[152]

In this context, viral vectors have been exploited as one of the vehicles of choice. Viruses are nano-sized (15–400 nm) supramolecular nucleoprotein-based entities, covered or not with a lipid bilayer (enveloped/nonenveloped viruses) that satisfy, into relatively simple structures, outstanding properties and functions that are relevant to drug and gene delivery. Viruses are able to recognize and interact specifically with cells by receptor-mediated binding, internalize, escape from endosomes, and uncoat and release nucleic acids in different cellular compartments. They are also capable of transcribing and translating their viral proteins to self-assemble into new infectious virus particles and exit the host cell.[120,153–155]

Despite all these relevant properties of viral vectors or some other rising vehicles in drug and gene delivery such as cationic liposomes, their therapeutic use presents some limitations and risks because of the complexity of production, limited packaging capacity, insertional mutagenesis and gene inactivation, low probability of integration, reduced efficacy of repeat administration or reduced expression overtime, unfavorable immunological recognition or strong

immune response against vehicle and transgene, inflammatory toxicity, and rapid clearance.[120,152] In this context, virus capsids or VLPs, produced by recombinant capsid proteins but lacking the viral genome, have noticeably emerged as a safer alternative to viral vectors.

A. Structure of Protein Self-Assembled Nanovehicles

VLPs are classically described as self-assembling, nonreplicative and non-pathogenic, highly organized supramolecular multiprotein nanoparticles (coats) (ranging from 20 to 100 nm) that can be formed from the minimal spontaneous self-assembling of one or more viral structural capsid proteins. It has been described that the self-assembling process of the structural viral proteins for VLP formation involves both spontaneous assembly, under favorable experimental conditions, and the requirement of scaffold proteins as catalysts.[156,157] Therefore, VLPs are considered protein "coats", "shells", or "boxes" that lack the viral genome, still conserve the structure, morphology, and some properties of viruses. Some of these properties such as cellular tropism and uptake, intracellular trafficking, membrane translocation, and transfer of nucleic acids or molecules across the cytoplasmic, endosomal, and nuclear membranes are important for drug delivery and gene therapy.[120,153,155,158–160] Usually, the degree of similarity of VLPs and their viruses depends on the number of proteins incorporated into the constructs.[161,162]

Since the first description in 1983 of the viral DNA packaging into mouse polyomavirus (MPyV) VLPs and its transduction *in vitro*,[163] VLPs of different viruses such as papillomaviruses,[164–166] hepatitis B, C, and E viruses,[167–169] polyomaviruses,[163,170–179] lentivirus,[180,181] rotavirus,[145,182] parvovirus,[183,184] and norovirus[185] have been generated.

B. Characteristic Features of VLPS and Their Limitations

VLPs offer some structure, dynamics, characteristic features, and functions that make them appealing bionanomaterials to be exploited in the biomedicine arena as drug and gene delivery vehicles and are discussed in detail afterward.

On the one hand, viral coat proteins have the ability to spontaneously self-assemble, which ensures the formation of highly organized, regular, repetitive structurally stable, and very low morphological polydisperse particles that provide useful properties to be used as scaffolds for bioimaging, synthesis of bionanomaterials, and as nanocarriers in drug and gene therapy.[186] In addition, homogeneity of particle size and composition is a desired production factor when developing therapeutic molecules. The overexpression of structural viral proteins in a convenient expression system renders recombinant proteins capable of being folded and assembled in discrete organized nanoparticles with a defined size corresponding to the natural capsid geometry.[187–189] Moreover, even though VLPs are structurally stable particles, some biochemical and

structural studies have observed that viral capsids and bacteriophages may show some structurally dynamic properties varying in shape, size, or rearrangements of the coat proteins, in response to different factors such as pH.[190–193]

On the other hand, VLPs are considered biologically safe nanostructures since they are not infectious (lack of viral genome) and do not replicate, representing a safer alternative to viral vectors.[160,194–197] However, they can elicit immune and inflammatory responses, especially when repeated administration is needed.[152] It has to be also noted that when used in vaccination, VLPs could show excellent adjuvant properties and the majority of VLPs stimulate strong cellular and humoral immune responses as direct immunogens.[198] It has been suggested that recombinant VLPs derived from infection of insect cells with baculovirus or even those derived from prokaryotic systems could be contaminated with different residual components of these host cells, contributing those impurities to the adjuvant properties.[153]

One interesting property of VLPs is that coat viral proteins present an enormous elasticity and adaptability to be modified chemically and/or by protein genetic engineering[154,160,199] to incorporate multiple directed functionalities, in order to be addressed in biomedical applications such as drug delivery or gene therapy. It has been recently reviewed that chemically and/or genetically modified VLPs, including CPMV, CCMV, MS2, M13 bacteriophages, and other virus-based nanoparticles,[155,186] could maintain their structural integrity and improve their physical stability[154] and, moreover, these modifications could also confer desired cell-targeting properties to the nanovehicle.[153–155,186,200,201] VLPs can be successfully engineered with spatial precision to incorporate (attached or genetically displayed on the surface) targeting tissue-specific ligands such as epidermal growth factor (EGFR) and antibodies, or other molecules such as oligonucleotides, peptides, gold, and other metals, target proteins, carbohydrates, polymers, fluorophores, quantum dots, drugs, or small molecules.[152,154,155] Moreover, one of the potential benefits of such modifications is that the specific geometric rearrangement confers precise recognition patterns.[200,201]

Furthermore, accessibility of the materials carried within the particle and the ability of inclusion and separation of nucleic acids, small molecules, and unusual cargoes with appropriate charge is another outstanding feature and key advantage of VLPs that has also made them excellent vessels for gene and drug delivery.[152,195] As described above, VLPs can be used as empty nanocarriers to transport molecules chemically attached on their surface or can be loaded *ex vivo* with therapeutic small molecules such as drugs, DNAs, mRNAs, siRNAs, oligonucleotides, quantum dots, magnetic nanoparticles, or proteins.[155,157,160] VLPs of different papillomavirus and polyomavirus have been widely characterized and used for directed delivery in biomedical applications.[132,165,173,174,194,202] Osmotic shock and *in vitro* self-assembling of VLP subunits in the presence of

the cargo have been the two main strategies used to packaged nucleic acid or other small molecules. It has to be taken into account that some attachment of the cargo on the VLP surface can occur.[195] Besides, diversity of natural tropism including liver for hepatitis B VLPs, spleen for some papillomavirus and polyomavirus VLPs, antigen-presenting cells for certain papillomavirus VLPs, and glial cells for human polyomavirus JC (JCV) VLPs, among others[152] is one of the key advantages offered by VLPs providing a wide spectrum of specific targeting and distribution profiles depending on the directed application. Although each VLP has its own characteristic receptors, entry pathway, and intracellular trafficking, it has been demonstrated that tropism of VLPs could be customized, modifying the residues identified as ligands of the cellular receptor on VLPs' surface or even varying the delivery routes.[155,189,203]

Another key advantage of VLPs is that they can be easily produced by using a wide range of hosts and expression systems, each of them with its own conditionings.[162] In the past years, there has been an increasing need to improve and optimize efficient large-scale production systems, process control and monitoring, and up- and down-streaming processes.[153,157,159,204] Production of VLPs usually involves transfection of the cell host expression system of choice with a plasmid encoding one or more viral structural proteins, further and rigorous purification for the removal of immunogenic cellular contaminants, and quality control of the produced VLP and encapsulation of the cargo *ex vivo* before administration.[152,158] The most frequent and convenient expression systems, adaptable to large-scale processes are (1) yeast cells[176], (2) mammalian cells, (3) insect cells infected with recombinant baculovirus[205,206], (4) bacteria[204,207], (5) green plants infected with modified viruses[208,209], and (6) cell-free systems.[163,204] The preparative and large-scale manufacture of VLPs in some of these hosts has been reviewed by Pattenden *et al.* and can be classified into two main methods of bioprocessing: *in vivo* and *in vitro* systems.[157] In addition, the capability of *in vitro* dissociation and reassociation of VLPs contribute to the application of easy and more accurate purification methods than those of viral vectors.[152,157] Furthermore, depending on the expression system, the resulting VLP might be significantly different even though expressing the same viral proteins. Thus, a broad spectrum of VLPs could be customized depending on the VLP type, the number of proteins needed for VLP assembling, and the targeted final application.[158,210]

As described above, VLPs have great potential as nanocarriers in drug and gene delivery. At the same time, although there is an increasing flow of developments in this area, these vehicles also present some limitations that should be addressed and taken into account, such as residual cellular components, variable yield of functional VLPs after disassembly/reassembly process, immunostimulation and unsuitability for repeated administration, tolerance to the transgene, ineffective therapeutic molecule loading, and low transfection rates.[152]

C. Tuning VLPS (Chemically or Genetically) for Their Uses/Applications in Gene Therapy and Drug Delivery

Due to their versatile nanoparticulate structure and morphology, and nonreplicative and noninfecting nature combined with their natural immunogenic properties and ease production, VLPs have principally emerged as an excellent alternative tool to attenuate viruses for vaccination.[152,153,204,210] There are currently commercialized upon the US Food and Drug Administration (FDA) approvals of some VLP-based vaccines that effectively protect humans from hepatitis B virus (HBV) (GlaxoSmithKline's Engerix® and Merck and Co., Inc.'s Recombivax HB®) and human papillomavirus (HPV) (Cervarix®, an HPV 16/18 VLP vaccine developed by GlaxoSmithKline's and Gardasil® developed by Merck against types 6, 11, 16, and 18 HPV). Other immunogenic VLP-based vaccines are already under clinical trials, preclinical test, or basic investigation including HBV,[211,212] HIV,[180,213] influenza virus,[214] parvovirus,[159] Norwalk virus,[185] rotavirus,[182] and Ebola virus.[215,216]

Although VLP-based vaccines have been primarily developed for their use against the corresponding virus, in the last decades genetic engineering or chemical modifications have been applied in order to generate chimeric VLPs. Thus, on the one hand, commonly short heterologous peptide epitopes or full proteins that are unable to form VLPs or that are unsafe for vaccination have been presented on surface-exposed loops or fused to N- or C-exposed termini of structural viral capsid proteins on VLPs.[154,161,210] Different HPV,[217–219] HBV,[220,221] parvovirus,[222,223] and chimeric polyoma VLPs have been engineered[170,175] and tested for different applications including vaccination against viral or bacterial diseases, against virus-induced tumors, and more recently, for immunotherapy of nonviral cancer.[161,210] On the other hand, chemical bioconjugation for covalent coupling of protein epitopes and small molecules to lysines, cysteines, or tyrosine residues of VLP surfaces has been applied in viral or cancer vaccines.[200] Chackerian *et al.* have demonstrated the efficient induction of protective autoantibodies using self-antigen conjugation to HPV VLPs.[224]

It is important to point out that VLPs can also be engineered to incorporate heterologous cell-specific ligands to cell receptors, thus altering their cellular tropism.[154,155,186,201] This great convertibility and flexibility of VLPs to be modified (chemically and/or genetically), their high stability, natural and diverse tropism, their nanocontainer properties, and their ability to enter in the cell and incorporate, bind, and deliver nucleic acids and small molecules have positioned VLPs as appealing entities not only for vaccination applications but also for a broad spectrum of other diverse and emerging applications in nanomedicine and nanotechnology such as immunotherapy against cancer,[210,225] gene therapy delivery of therapeutic genes into specific cells,[161,165,171,184,226,227] and targeted delivery of drugs and small molecules using VLPs as nanocarriers.[174,196]

Although there is no commercial VLP as vector in gene therapy, since the initial work in 1970 of uncoating polyoma pseudovirus in mouse embryo cells as gene delivery vector[228] and the establishment in 1983 of the viral DNA packaging into MPyV VLPs and its transduction *in vitro*,[163] different VLPs such as HBV and hepatitis E virus,[229] HPV and polyomavirus nanoparticles[172,178,229] have been modified toward the specific delivery of therapeutic genes and proteins in different target cells, organs, and tissues *in vitro* and *in vivo* by systemic injection[229] or oral administration.[230] For example, recombinant VP1-based polyomavirus VLPs can encapsulate *in vitro* exogenous DNA, and deliver it by cell surface sialic acid residues to human brain cells and fetal kidney epithelial cells.[178] Furthermore, VLPs have recently emerged as novel nanocarriers or nanocontainers to store unnatural cargos, deliver modified oligonucleotides,[154] synthetic small interfering RNAs, and plasmids expressing short hairpin RNAs as therapy to downregulate gene expression.[171,231] In this context, Chou *et al*. have recently described the use of JCV VLPs as an efficient vector for delivering RNAi *in vitro* using murine macrophage RAW 264.7 cells and *in vivo* using BALB/c mice in silencing the cytokine gene of IL-10 without significant cytotoxicity for systemic lupus erythematosus gene therapy.[171]

One of the key aspects in targeted gene and drug delivery is cell-specific delivery. It is important to point out that VLPs are tunable nanoparticles that can also be chemically or genetically engineered to modify their natural cellular tropism in order to diversify the range of therapeutic applications in targeted gene or drug delivery.[154,201] Some effective approaches to modify the natural cellular tropism include:

(1) Genetic engineering of VLP chimeras incorporating heterologous cell-specific short peptides that contain recognition sites of target cell receptors.[232] In this context, polyoma and papillomavirus, with solved atomic structures of their major structural capsid proteins, have been extensively used to obtain chimeric VLPs as delivery vector systems.[165,233] However, this approach has some bioprocessing limitations such as low production levels as a consequence of VLP modification, alterations of size and properties of the VLPs that could affect the structural interactions and conformations for VLP assembly, disassembly and packaging, and low transduction efficiencies.[157]

(2) Chemical bioconjugation of purified VLPs with epitope-containing peptides[234,235] or a wide range of small molecules conferring cell-specific targeting such as transferrins, folic acid, or other targeting molecules. As an example, CMPV VLPs have been successfully conjugated with Tfn using "click" chemistry[236] and with NHS-ester-derivatized folic acid, demonstrating both as internalized into HeLa cells and KB cells, respectively.[183,184]

(3) High-throughput library and directed evolution method is a rational approach that has been recently used to engineer viral vectors with the desired tropism properties.[237]
(4) Pseudotyping, which consists of replacing the envelope protein of one virus species by the envelope protein of another virus species.[238]
(5) Modification of the delivery route of the VLPs. It has been shown that the levels of expression of β-galactosidase in heart, lung, kidney, spleen, liver, and brain are different depending on the delivery route of polyomavirus VP1 VLPs.[203]

The great accessibility and reactivity showed by VLPs, as well as their ability to serve as nanocarriers, which made them suitable to be exploited in gene therapy, have also been applied to targeted drug delivery.[195] Genetic modification and/or chemical functionalization of exposed amino acid residues on the capsid surface in order to attach small molecules, such as markers or bioactives molecules, is one of the most common approaches applied to target drug delivery.[174,239] As an example, canine parvovirus (CPV) VLPs produced in a baculovirus expression system and exhibiting natural tropism to transferrin receptors (TfRs) were chemically modified on accessible lysines of the capsid surface with fluorescent dye molecules and delivered to tumor cells. Derivatization of CPV-VLPs did not interfere with the binding and internalization into tumor cells.[183,184]

One limitation of VLPs in gene therapy is the low efficiency of gene transduction due to inefficient DNA packaging. However, a recent study presented a novel *in vivo* DNA packaging of JCV VLPs in *E. coli* that effectively reduced human colon carcinoma volume in a nude mouse model. In this study, the exogenous plasmid DNA was transformed into the JCV VP1 expressing *E. coli*. The packaging of the second plasmid occurs simultaneously as the *in vivo* assembly of the JCV VLP. Even though it is still not clear how the plasmid DNA molecules are encapsidated in the VLP, the authors showed that gene transduction efficiency by their *in vivo* package system was about 80% in contrast to the 1–2% of gene transduction efficiency achieved by the *in vitro* osmotic shock system.[226] In addition, the administration of exogenous proteins may induce the immune system response, reducing therapy effectiveness or causing undesirable secondary effects, albeit immunological response of protein nanoparticles can be modulated.[240]

V. Nonviral Self-Assembling Proteins

Spontaneous protein self-assembly to form ordered oligomers is a common event in biology. It can prove advantageous in terms of genome-size minimization, formation of large structures, stabilization of complexes, and inclusion of

functional features.[241] It has been widely documented that cellular oligomer proteins as well as viral capsids are stabilized by several weak noncovalent interactions as hydrophobic interaction, electrostatic energy, and Van der Waals forces.[242–244] These interactions result in a complex quaternary structure described by three symmetry point groups named cyclic (Cn), dihedral (Dm), and cubic (T, O, I).[245,246]

The development of computational techniques to predict protein–protein interactions using solved 3D protein structures makes it possible to predict and/or strengthen experimental data performing in *in silico* approaches.[247] Furthermore, its use opens up the possibility to design proteins not only displaying specific biological functions but also interesting intermolecular interactions to obtain increased multivalency in the resulting complexes. Moreover, it should be considered that not only whole proteins can self-assemble in smart nanoparticles; oligopeptides are also capable of forming organized structures. Many applications are possible due to the enormous quantity of different combinations and features that can be exploited with peptides.[248,249]

Furthermore, protein–protein interactions are not the unique parameters involved in particle formation, nucleic acid–peptide interactions, salt concentration, order of mix, and ratio between nucleic acid and protein can also strongly influence the condensation process.[250,251]

Due to their natural tendency to self-assemble forming highly ordered structures, viruses provide a wide variety of scaffold proteins which are used as gene/drug carriers. Among them, VLPs have been reviewed in the previous section. However, simple bacterial proteins can be also utilized as carriers for gene delivery. For example, heat shock proteins (HSP) from hyperthermophilic archeaon *Methanococcus jannaschii* can assemble in a small structure of 24 subunits having an octahedral symmetry. These 12 nm structures are stable at high temperature, up to 70 °C, and wide range of pH. Residue modifications are allowed to elicit specific attachment of small molecules.[186,252]

In bacteria, bacterial microcompartments (BMC) which are intracellular organelles consisting of enzymes encapsulated within polyhedral, protein-only shells, somewhat similar to viral capsids, have been described. BMCs are composed of a few thousand copies of a few repeated protein species (including one or more enzymes involved in specific metabolic pathways), and with sizes of around 100–150 nm in cross section. The general role of BMCs is to confine toxic or volatile metabolic intermediates, while allowing enzyme substrates, products, and cofactors to pass.

The first described BMC, the carboxysome, was isolated in the early 1970s[253,254] and has been found to contain both CO_2-fixing ribulose bisphosphate carboxylase/oxygenase (RuBisCO)[253,254] and carbonic anhydrase[255–257] enzymes. Carboxysomes' function is to enhance autotrophic CO_2 fixation at low CO_2 levels.

Other BMCs were later identified in cyanobacteria and some chemoautotroph bacteria. Among them, BMC proteins have been later found to be encoded in the propanediol utilization operon (pdu) of the heterotroph Salmonella[258] and by an operon for metabolizing ethanolamine (eut) in enteric bacterial species, including *Salmonella* and *Escherichia.*[259] *Salmonella enterica* forms a polyhedral organelle during growth on 1,2-propanediol (1,2-PD) as a sole carbon and energy source, but not during growth on other carbon sources.[260,261]

The pdu organelles' function is to minimize the harmful effects of a toxic intermediate of 1,2-PD degradation (propionaldehyde).[261–263] Other studies have shown that a polyhedral organelle is involved in ethanolamine utilization (eut) by *S. enterica.*[259] The function of the eut microcompartment is to metabolize ethanolamine without allowing the release of acetaldehyde into the cytosol, therefore minimizing the potentially toxic effects of excess aldehyde in the bacterial cytosol[264–266] and also preventing volatile acetaldehyde from diffusing across cell membrane.[267]

So far, about 1700 proteins containing BMC domains have been identified, covering at least 10 different bacterial phyla. The typical BMC protein consists of approximately 90 amino acids, with an alpha/beta fold pattern.[268,269] Some individual BMC proteins self-assemble to form hexamers, which further assemble side by side to form the flat facets of the shell.[268,270,271] The formation of icosahedral, closed shells from such flat layers was elucidated in part by structural studies in carboxysomes: some BMC proteins assemble to form pentamers, which are located at and form the vertices of the icosahedral shell.[270]

Mechanisms directing enzyme encapsulation within protein-based BMCs have been studied during the last years. It has been described that, in some carboxysomes, protein CcmM is used as a scaffold to form interactions between both shell proteins and enzymes,[272,273] through a CcmM C-terminal region with homology to the small subunit of RuBisCO.[274] Other studies revealed that pdu shells can self-assemble without needing interior enzymes[275] and that carboxysomes can self-assemble *in vivo* when RuBisCO has been deleted.[276]

Regarding properties of the encapsulated enzymes, in the pdu BMC some of the internal enzymes are encapsulated by specific N-terminal targeting sequences.[275,277] In this line, Sutter and colleagues[278] described a conserved C-terminal amino acid sequence that mediates the physical interaction of an iron-dependent peroxidase (DyP) or a protein closely related to ferritin (Flp) with a specific type of BMC (encapsulins).

In another example, an icosahedral enzyme complex, lumazine synthase (AaLS) from *Bacillus subtilis* and *Aquifex aeolicus*, was engineered to encapsulate target molecules by means of charge complementarity and can also be modified to give different characteristics to the assembled structure.[279,280]

Moreover, enzymatic subunits, like E2 of pyruvate dehydrogenase from *Bacillus stearotermophilus*, can be modified to be used in gene delivery. E2 peptides naturally form a dodecahedron of 60 subunits of 24 nm in diameter allowing modification for drug-like accommodation. The assembling/disassembling of these structures can be modulated by changing the operative pH in the experimental environment. These nanoparticles can also be functionalized with antigens for vaccine development.[281,282]

According to these results, specific targeting sequences could be of use in biotechnological applications to package proteins inside the stable self-assembled icosahedral shell of BMCs, offering appealing opportunities to manipulate in the laboratory such nanocages to fill them with therapeutic molecules. The simplicity of this system makes it very attractive for engineering studies to design, mimicking nature, new applications in biotechnology, providing a new, intriguing platform of microbial origin for drug delivery.

Bovine serum albumin (BSA) is able to form microspheres after sonochemical treatment in aqueous medium. Chemical effects of ultrasound radiation and coupling with an anticancer drug such as Taxol (paclitaxel) led to the assembling of a spherical carrier with an average diameter of 120 nm. BSA particles resulting from S–S bonds, due to HO_2 radical formation, are able to release the encapsulated Taxol in cancer tissue with best results if compared with mere Taxol treatment. This drug for breast cancer treatment is commercially available.[283,284]

Also little cationic peptides can lead to self-assembling particles. Among others, arginine-rich cationic peptides are widely known as good tools for gene delivery. For example, purified R9-tailored GFP in solution is described to form nanodisk particles 20 nm in diameter. This structure is proved to be induced by the 9 arg tails and is able to bind and condense DNA. These nanodisks are also able to deliver DNA toward the nucleus where the reporter gene is expressed.[285]

On the other hand, the expression of recombinant proteins over physiological rates can cause a bad functioning of cellular quality control system, leading to self-organizing, pseudo-spherical, protein aggregates known as inclusion bodies. These mechanically stable nanoparticles, ranging from 50 to 500 nm in diameter, were considered for a long time as undesired bio-products. Recently, it became clearer that they are suitable for medical approaches when utilized as scaffold surface to promote cellular proliferation.[286–288]

One of the most difficult goals for a foreign gene delivery is to reach the nucleus. An approach to overpass this obstacle is by fusing an NLS in a nonessential position of a DNA-binding protein. Such type of modification has been described for a tetracycline repressor protein (TetR) fused with an SV40 NLS. The TetR–NLS affinity and specificity to TetO DNA sequence is exploited to form spontaneous protein–DNA complexes which allow an

enhancing of DNA transportation into the nucleus and subsequent expression of foreign genes, combining the two peculiar characteristics of each fusion component.[289]

VI. Medical Applications of Protein Nanoparticles

There is still a tremendous gap between progresses made in protein-based nanoparticle research for drug delivery and clinical reality. Hundreds of publications in basic research describe the combination of two or more functional elements in a single protein nanoparticle, by which the delivery of a carried drug is enhanced. These agents act by improving critical steps in the drug delivery process, such as increasing the systemic stability or tissue specificity, favoring internalization, endosomal escape, and entry into the nucleus, or transporting therapeutic material through the BBB, in *in vitro* and *in vivo* studies.

Besides the human recombinant therapeutic proteins currently on the market (or functional segments of them), there are also some fusion proteins approved for clinical use (most by incorporating an antibody fragment or a ligand to enhance cell specificity). Sadly no gene therapy trials have so far used full protein carriers *in vivo*, but rather peptide-functionalized vehicles.

Bottlenecking the gap between research and clinical application, the US FDA/European Medicines Agency (EMEA) only approves human proteins, to avoid the risk of an immune response that could affect not only the effectiveness of the nanoparticle but also challenge patients' health. Another critical factor is the administration route, where the protein is degraded before arriving at the target; this problem could be solved or minimized by the use of protein D-isomers, PEGylation, or the design of protecting groups for labile sites. Despite the current situation mentioned above, there are many good examples of multifunctional modular proteins that, when carrying therapeutic material, can improve the prognosis *in vivo* in animal models for different diseases. These examples are reviewed below, along with those few protein nanoparticles that are currently on the market or in clinical trials.

A. Therapeutic Protein Nanoparticles Currently in the Market

Albumin is a natural protein transporter of hydrophobic molecules throughout plasma that has been approved by the FDA to reversibly bind water-insoluble anticancer agents, as is the case of albumin-bound (nab) paclitaxel, Abraxane®. This albumin-nab technology-based drug is in use in patients with metastatic breast cancer who have failed combination therapy, and it is the first protein nanoparticle approved by the FDA. Albumin potentiates paclitaxel

concentration within the tumor by increasing paclytaxel endothelial transcytosis through caveolae formation. It also contributes to the fact that tumors secrete an albumin-binding protein SPARC (also called BM-40) to attract and keep albumin-bound nutrients inside the tumor cell.[290] The albumin–paclitaxel complex was not formally considered a nanoparticle in the United States (due to an average size of 130 nm) but only so in Europe.

Apart from whole recombinant therapeutic proteins being currently commercialized, there are also some examples of vehicles formed by chimerical proteins with target ligands already in the market. DAB389IL-2 (denileukin diftitox or Ontak) is a fusion of Diphtheria toxin catalytic and translocation domains for lethal effect and interleukin-2 (IL-2) to gain cell specificity in the treatment of persistent or recurrent T-cell lymphoma. Belatacept (BMS-224818) is a CTLA4-Ig fusion protein formed by the cytotoxic T-lymphocyte-associated antigen 4 joined to an immunoglobulin G1 Fc fragment fusion protein, developed by Bristol–Miers–Squibb. Etanercept (Enbrel) fusion tumor necrosis factor receptor (TNFR), which binds and inhibits specifically TNF activity, to an immune globulin G1 Fc, to prevent inflammation mediated by TNF in autoimmune diseases like arthritis and psoriasis.

On the other hand, fusion proteins which include an antihuman epidermal growth factor receptor 2 (HER2) monoclonal antibody that binds tumor cell surfaces, among them the so-called "trastuzumab" (commercialized as Herceptin by Roche), associated to DM-1, an antimitotic drug, aimed at improving the treatment of breast cancer.

Finally, VLPs, that is, empty viral entities formed by the self-assembly of a viral capsid protein, are the only truly protein nanoparticles (architectonically speaking) which are currently used in clinical practice. HBsAg recombinant protein of HBV expressed in yeast and the capsid L1 recombinant protein of HPV (types 6, 11, 16, and 18) administered currently as vaccines tend to form spontaneously VLPs that elicit T and B immune response. Recently, there have been preclinical and clinical trials to test the security and efficacy of VLP vaccines against Chikungunya[291] and seasonal influenza virus (http://www.medpagetoday.com/MeetingCoverage//ICAAC/22129), respectively. Influenza VLP vaccines have proven to provide complete protection against H1N1 2009 flu pandemics,[292] within a record preparation time when compared to 9 months for traditional vaccines. The use of VLPs as a delivery system for drugs or nucleic acids in gene therapy is still under investigation.[194]

Drugs and proteins may be transformed through pegylation, a process that can assist them in overcoming some of the potential problems that delay the adoption of protein nanoparticles for clinical use. The covalent attachment of PEG can reduce immunogenicity and antigenicity by hiding the particle from the immune system, can increase the circulating time by reducing renal clearance, and can also improve the water solubility of a hydrophobic particle. The use of

pegylation has been approved for commercial use by the FDA and EMEA, and some examples of pegylated protein products are Adagen® (PEG-bovine adenosine deaminase), the first pegylated protein approved by the FDA in 1990, Pegasys® (PEG-interferon alpha), and Oncaspar® (PEG-L-asparaginase).

B. Therapeutic Protein Nanoparticles Currently in Clinical Trials

The majority of protein nanoparticles studied in clinical trials (http://clinicaltrias.gov) are fusion proteins composed of a therapeutic protein/peptide and a target cell-specific ligand. An example is ALT-801, a biologic compound composed of IL-2 genetically fused to a humanized soluble T-cell receptor directed against the p53-derived antigen. The clinical trials evaluated whether directing IL-2 activity using ALT-801 to the patient's tumor sites that overexpress p53 results in clinical benefits (NCT01029873, NCT00496860). Another ligand joined to IL-2 is L19, a tumor-targeted immunocytokine constituted of a single chain fragment variable (scFv) directed against the ED-B domain of fibronectin, one of the most important markers for neoangiogenesis. L19–IL-2 is in a Phase I/II study for patients with solid tumors and renal cell carcinoma (RCC) (NCT01058538). L19 has also been fused to TNFα with the intention to target TNFα directly to tumor tissues resulting in high and sustained intralesional bioactive TNFα concentrations. The L19TNFα is under clinical trial using isolated inferior limb perfusion (ILP) with the standard treatment with melphalan 10 mg/l limb volume in subjects affected by stage III/IV limb melanoma (NCT01213732). NGR-hTNF is another bifunctional protein which combines a tumor-homing peptide (NGR) that selectively binds to amino peptidase N/CD13 highly expressed on tumor blood vessels, thus affecting tumor vascular permeability, and hTNF, with direct anticancer activity. NGR-hTNF is undergoing 14 clinical trials as a single agent to treat different cancers, as well as in combination with chemotherapy agents.

Another strategy to direct a therapeutic protein to the target cell is through fusion to a growth factor receptor ligand. An example is TP-38, a recombinant chimerical protein composed of the EGFR binding ligand (TGF-α) and a genetically engineered form of the Pseudomonas exotoxin, PE-38, to treat recurrent grade IV malignant brain tumors (NCT00071539).

Many clinical trials are based on a therapeutic protein fused to a targeting antibody, as is the case of APC8015. This drug stimulates the immune system and stops cancer cells from growing by the combination of biological therapies with Bevacizumab®, an already approved monoclonal antibody that locates tumor cells and kills them in a specific way (NCT00849290). There are also many putative protein drugs against cancer which include antibodies anti-integrins (e.g., cilengitide and IMGN388), sometimes in combination with

classical therapies. A recently developed tool, the nanobodies or single domain antibodies,[293] have several advantages: small size (only 12–15 KDa), which lowers the possibility of triggering immune response, safety in clinical trials (NCT01020383), and is easy to be joined to different kinds of compounds. All these features make nanobodies competent drugs against different diseases, and have been tested *in vivo* as bifunctional proteins associated to a prodrug, very efficient in mice cancer xenografts.[294]

Even though CPPs are very useful tools to deliver drugs and in gene therapy (see the chapter "Peptide Nanoparticles for Oligonucleotide Delivery" by Lehto *et al.* in this volume), their toxicity and endosomal entrapment slows their inclusion for systemic delivery in clinical trials. Nevertheless, there are a few examples of use to prevent undesirable cell proliferation in coronary artery bypass grafts, as is the case of a CPP $(R\text{-}Ahx\text{-}R)_4$AhxB–PMO conjugate targeted to human c-myc to be applied *ex vivo*. The trial, in phase II, has been completed in 2009 (NCT00451256). Another case is PsorBan®, a product patented for the treatment of psoriasis based on a cyclosporine–polyarginine conjugate of local application, which circumvents the specificity problem of intravenous (i.v.) application. It is in clinical trial phase III, but not yet in the market. Finally, KAI-9803, a PKCδ inhibitor peptide conjugated to Tat to function as an intravenous drug for the treatment of acute myocardial infarction, is currently in phase 2b clinical trial (NCT00785954, KAI pharmaceuticals).

C. Therapeutic Protein Nanoparticles in Preclinical Models of Human Diseases

There are many proteins, often organized as nanoparticles, that when associated to a drug, therapeutic protein, peptide, or nucleic acid increase the therapeutic efficacy of a cargo alone in the treatment of various diseases. Some of them proved effective in animal models, which are discussed in more detail in this section, with relevant examples listed in Table II.

These nanoparticles may simply be (a) a CPP to promote nonspecific internalization,[295–300] (b) a peptide to confer cargo specificity by joining a receptor distinctive of a cell type, including scFvs or peptides obtained by phage display,[301] and (c) a mixture of both,[302] since as observed in several studies the CPP does not reduce ligand specificity and increases nanoparticle potency.[303–305] Complex and multifunctional vehicles including endosomal escape peptides enhance the therapeutic potency of the complex, or other domains that allow their selective activation in certain contexts.[306,307]

Apart from the cases listed in Table II, the spectrum of additional examples of multidomain protein nanoparticles tested *in vivo* is wide, and a considerable proportion of them include CPPs, mainly Tat and polyarginines. A classical Tat fusion protein is the transducible D-isomer RI-TATp53C″ fusion protein that

TABLE II

REPRESENTATIVE EXAMPLES OF PROTEIN NANOPARTICLES THAT, ACTING AS CARRIERS, IMPROVE THE EFFICIENCY OF CARGO ALONE IN THE TREATMENT OF DISEASES USING *IN VIVO* MODELS

Carrier	Cargo	Administration route	Disease	References
VP-22	Gata4	Transplant of transfected cells	Myocardial infartion	295
$(RXR)_4XB$	Dystrophin exon skipping PMO	i.p.[a]	Duchenne muscular dystrophy	296
Tat-ErbB2	STAT3BP	i.p.	Breast cancer xenograft	302
Penetratin	scFVs-radionuclide	i.v.[b]	Colon carcinoma xenograft	305
8R	Taxol	i.p.	i.p. tumor xenografts	297
Penetratin	Caveolin-1	i.p.	Inflammation models	298
Tat-HA	Bcl-xL	i.p.	Cerebral ischemia	313
Protamine-Erb2 Ab Fab	c-myc, MDM2, VEGF-siRNA	i.v.	Breast cancer	301
9-D-arginine-RVG	JEV-siRNA	i.v.	JEV infection	56
Pegylated Pep-3	Cyclin B1 -PNA	i.v.	Human prostate carcinoma xenograft	299
Chol-MPG-8	Cyclin B1-siRNA	i.v.	Prostate and lung cancer xenografts	300
Tat	pVHL	i.p.	Mice with renal tumors dorsally implanted	310
Tat	MHC class I antigens	s.c.[c]	Dendritic cell vaccine for tumor regresion	311

[a]Intraperitoneal.
[b]Intravenous.
[c]Subcutaneous.

activates p53 protein in cancer cells, but not in normal cells. RI-TATp53C′ treatment in terminal peritoneal carcinomatosis and peritoneal lymphoma preclinical models results in significant increases in life span (higher than sixfold) and full recovery from the disease.[308] There are also several studies *in vivo* using Tat-fused therapeutic proteins which have proven effective in treating tumors[309–311] and cerebral ischemia[312,313] when applied intraperitoneally (i.p.).

Regarding polyarginines, Kumar and colleagues have presented two different models in which a bifunctional peptide formed by nine arginines (9R) and a specific ligand constitute an effective siRNA vehicle. In the first model, a chimerical peptide derived from rabies virus glycoprotein (to confer neuronal specificity) fused to 9- D-arginines (RVG-9R), was able to transport si-RNA across the BBB and silence specific gene expression in the brain when applied intravenously.[56] In the second model, a CD7-specific single-chain antibody was conjugated to oligo-9-arginine peptide (scFvCD7-9R) for T cell-specific antiviral siRNA delivery in humanized mice reconstituted with human lymphocytes. In HIV-infected humanized mice, this treatment controlled viral replication and prevented the disease-associated CD4 T cell loss. Moreover, it effectively suppressed viremia in infected mice.[314]

Some other examples of polyarginines in tumor models are 9-D-arginines fused to a tumor-suppressor peptide, which stopped tumor growth in hepatocellular carcinoma-bearing mice when applied intraperitoneally, and also colesteryl oligoarginines carrying VEGF siRNA, which inhibited tumor growth in colon adenocarcinoma after local application.[315] Another BBB-crossing peptide is g7, which is able to transport nanoparticles loaded with Loperamide.[107]

In general, the partner fusion peptide can confer specificity instead of penetrability, as is the case of EGFR Fab fragment associated to liposomes that contain anticancer drug, which increases efficiency of anticancer effect in EGF overexpressing xenograft tumors[316]; in addition, RGD-4 C-doxorubicin in human breast xenografts increases efficacy and diminishes toxicity.[317]

In many conjugates, the therapeutic peptide of the chimerical proteins is a toxin. Anthrax lethal toxin has been modified to be activated by methaloproteases, and it has probed to be effective for human xenografted tumors such as melanoma, lung, and colorectal cancer.[318] Anthrax toxin has also been associated to antibodies or growth factors for lethal effects specifically on cancer cells.[319] The specific cytotoxicity desired to treat a tumor might derive from a tissue factor, which promotes clotting to restrict blood supply in tumor vessels, fused to peptides that provide specificity, like V-CAM antibodies, fibronectin, and integrin ligands.[320]

Eventually, drug activity may decrease when conjugated to a carrier protein, although if the entry of the drug is favored, the overall balance of activity can be much more efficient.[321] On the other hand, the use of noncovalent bond drug carrier could avoid interfering with the activity of the drug.

An important issue in a preclinical study to be considered for a clinical trial is the administration route. In *in vivo* experiments, most of the protein nanoparticles are administered by local or intraperitoneal injection, avoiding systemic spreading and clearance in the vascular system, in a way very similar to *in vitro* experiments. The FDA and EMEA, on the other hand, will preferentially approve i.v. and oral administrations rather than intraperitoneal or local injections except for very accessible tissues. Another relevant issue is the number of active domains to be included in a therapeutic protein carrier, an issue that seems to be relevant for the functionality of the construct. For example, the CPP neutralization of a ligand may depend on the CPP/ligand ratio that is in the vehicle.[322] It has also been observed that the integrin binding power of RGD-containing motives increases with the number of RGD domains over the monomer until a maxim of four moieties.[323] Another example is Tat activity empowerment when attached to molecules that form tetramers, such as beta-galactosidase[108] and p-53.[324]

Some multidomain protein carriers allow the drug entrance only in selected target cells by tailored smart selective mechanisms.[325] For instance, CPPs neutralized by polyanions are activated and enter the cells when they are released by metalloproteases[326] or by lowering the pH,[327] both situations being very common in tumors.

CPP-morpholino oligomer (PMO) nanoparticles have also shown their effectiveness in treating viral infections by inhibiting viral replication, as demonstrated with the carrier (R-ahx-R) 4AhxB-PMO administered i.v. in animal models infected with picornaviruses, i.p. in mice infected with coronaviruses and flaviviruses, and the carrier R9F2C-PMO administered also i.p. in mice infected with Ebola virus. Furthermore, it has also been shown in some of these studies that the efficacy of the treatment is dependent on the incorporation of arginine-rich peptides in the nanoparticle.[328]

A good example of how a CPP can improve the internalization of a therapeutic protein is the case of insulin. The instability and low absorption in the digestive tract of insulin prevents its oral administration, even though it would be very convenient for a daily administrated drug. In recent studies, noncovalent conjugation of insulin to different CPPs enhances its absorption without toxic intestinal effect, L-penetratin being the most efficient as insulin carrier.[329]

Among the protein nanoparticles tested *in vivo*, it is worth making special mention of Trojan horses generated in Pardridge's laboratory to cross the BBB, through a strategy of fusing within a chimerical peptide the therapeutic protein which has to reach the CNS to a monoclonal antibody against the human insulin receptor (HIRMAb). This Trojan horse is very potent for humans and primates, and has proven effective to transport β-glucuronidase, α-L-iduronidase, GDNF, Abeta amyloid peptides, paroxonase, etc., with potential benefits in diseases like mucopolysaccharidosis type VII, Hurler syndrome, Parkinson, Alzheimer, and organophosphates toxicity, respectively.[330]

There are also promising results when protein nanoparticles have been tested as carriers for gene therapy *in vivo*, some examples being listed in Table II. In this regard, the use of modular proteins generated by insertional mutagenesis of β-galactosidase condensing the SOD gene are able to protect neurons against ischemic injury[133]; a bifunctional galactosylated polylysine is able to conjugate plasmid DNA and to differentially promote expression in hepatocytes that display asialoglycoprotein receptor[331]; a suicide multidomain protein particle formed by herpes simplex virus thymidine kinase (HSV-TK) conjugated to transferrin (Tf) by a biotin-streptavidin bridging, which, administered i.v. in K562 massively metastasized nude mice, was able to reduce tumor size and to increase mouse survival.[332]

VII. Conclusions

In this chapter, proteins and peptides have been envisioned as potent biotechnological tools for the development of new biocompatible biological entities that can be used as therapeutic agents by themselves or as nanovehicles for the delivery of associated drugs. Proteins are nanostructures that can form complex high-order entities such as VLPs, resulting in appropriate cages for the internalization of therapeutic molecules. In addition, the design of modular proteins displaying selected functions has been possible by using *in silico* approximations to the feasibility of recombinant protein production. This approach has demonstrated the versatility of such molecules in the generation of novel delivery nanovehicles opening up the possibility of new functional combinations to enhance the specific interaction with the target tissue. Such tunable specificity in the delivery of drugs, nucleic acids, or other proteins is one of the main properties that make multifunctional proteins appealing as more rational delivery vehicles.

The presence on the market of such complex entities, which started with the approval of Insulin for the treatment of diabetes, has been increasing over the past years, and this tendency is expected to continue. In fact, there are some products in clinical trials that will probably end up being approved and some more are being explored in preclinical experiments which might enter in clinical trials.

Acknowledgments

The authors appreciate the financial support received through grants BFU2010-17450 from MICINN, PS0900165 from FISS, and 2009SGR-108 from AGAUR. The authors also acknowledge the support of the CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), 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.

References

1. Shun TY, Lazo JS, Sharlow ER, Johnston PA. Identifying actives from HTS data sets: practical approaches for the selection of an appropriate HTS data-processing method and quality control review. *J Biomol Screen* 2011;**16**:1–14.
2. Vuorela P, Leinonen M, Saikku P, Tammela P, Rauha JP, Wennberg T, et al. Natural products in the process of finding new drug candidates. *Curr Med Chem* 2004;**11**:1375–89.
3. Zhao HY, Akritopoulou-Zanze I. When analoging is not enough: scaffold discovery in medicinal chemistry. *Expert Opin Drug Discovery* 2010;**5**:123–34.
4. Adamina M, Joerger M. Dose-toxicity models in oncology. *Expert Opin Drug Metab Toxicol* 2011;**7**:201–11.
5. Schmid EF, Smith DA. Is declining innovation in the pharmaceutical industry a myth? *Drug Discov Today* 2005;**10**:1031–9.
6. Veronese FM, Mero A. The impact of PEGylation on biological therapies. *BioDrugs* 2008;**22**:315–29.
7. Raja SM, Clubb RJ, Ortega-Cava C, Williams SH, Bailey TA, Duan L, et al. Anticancer activity of celastrol in combination with ErbB2-targeted therapeutics for treatment of ErbB2-overexpressing breast cancers. *Cancer Biol Ther* 2011;**11**:263–76.
8. Desai N, Trieu V, Damascelli B, Soon-Shiong P. SPARC expression correlates with tumor response to albumin-bound paclitaxel in head and neck cancer patients. *Transl Oncol* 2009;**2**:59–64.
9. Desai NP, Trieu V, Hwang LY, Wu R, Soon-Shiong P, Gradishar WJ. Improved effectiveness of nanoparticle albumin-bound (nab) paclitaxel versus polysorbate-based docetaxel in multiple xenografts as a function of HER2 and SPARC status. *Anticancer Drugs* 2008;**19**:899–909.
10. Houk BE, Martin R, Hochhaus G, Hughes JA. Pharmacokinetics of plasmid DNA in the rat. *Pharm Res* 2001;**18**:67–74.
11. Wang W. Instability, stabilization, and formulation of liquid protein pharmaceuticals. *Int J Pharm* 1999;**185**:129–88.
12. Martin ME, Rice KG. Peptide-guided gene delivery. *AAPS J* 2007;**9**:E18–29.
13. Mastrobattista E, van der Aa MA, Hennink WE, Crommelin DJ. Artificial viruses: a nanotechnological approach to gene delivery. *Nat Rev Drug Discov* 2006;**5**:115–21.
14. Gabathuler R. Approaches to transport therapeutic drugs across the blood-brain barrier to treat brain diseases. *Neurobiol Dis* 2010;**37**:48–57.
15. Glover DJ, Ng SM, Mechler A, Martin LL, Jans DA. Multifunctional protein nanocarriers for targeted nuclear gene delivery in nondividing cells. *FASEB J* 2009;**23**:2996–3006.
16. Tiera MJ, Winnik FO, Fernandes JC. Synthetic and natural polycations for gene therapy: state of the art and new perspectives. *Curr Gene Ther* 2006;**6**:59–71.
17. Mannisto M, Vanderkerken S, Toncheva V, Elomaa M, Ruponen M, Schacht E, et al. Structure-activity relationships of poly(L-lysines): effects of pegylation and molecular shape on physicochemical and biological properties in gene delivery. *J Control Release* 2002;**83**:169–82.
18. Ward CM, Read ML, Seymour LW. Systemic circulation of poly(L-lysine)/DNA vectors is influenced by polycation molecular weight and type of DNA: differential circulation in mice and rats and the implications for human gene therapy. *Blood* 2001;**97**:2221–9.
19. Gottschalk S, Sparrow JT, Hauer J, Mims MP, Leland FE, Woo SL, et al. A novel DNA-peptide complex for efficient gene transfer and expression in mammalian cells. *Gene Ther* 1996;**3**:448–57.
20. Plank C, Tang MX, Wolfe AR, Szoka Jr. FC. Branched cationic peptides for gene delivery: role of type and number of cationic residues in formation and in vitro activity of DNA polyplexes. *Hum Gene Ther* 1999;**10**:319–32.

21. McKenzie DL, Collard WT, Rice KG. Comparative gene transfer efficiency of low molecular weight polylysine DNA-condensing peptides. *J Pept Res* 1999;**54**:311–8.
22. McKenzie DL, Smiley E, Kwok KY, Rice KG. Low molecular weight disulfide cross-linking peptides as nonviral gene delivery carriers. *Bioconjug Chem* 2000;**11**:901–9.
23. Sorgi FL, Bhattacharya S, Huang L. Protamine sulfate enhances lipid-mediated gene transfer. *Gene Ther* 1997;**4**:961–8.
24. Brewer LR, Corzett M, Balhorn R. Protamine-induced condensation and decondensation of the same DNA molecule. *Science* 1999;**286**:120–3.
25. Masuda T, Akita H, Harashima H. Evaluation of nuclear transfer and transcription of plasmid DNA condensed with protamine by microinjection: the use of a nuclear transfer score. *FEBS Lett* 2005;**579**:2143–8.
26. Balhorn R. The protamine family of sperm nuclear proteins. *Genome Biol* 2007;**8**:227.
27. Ferrer-Miralles N, Vazquez E, Villaverde A. Membrane-active peptides for non-viral gene therapy: making the safest easier. *Trends Biotechnol* 2008;**26**:267–75.
28. Chan CK, Jans DA. Enhancement of MSH receptor- and GAL4-mediated gene transfer by switching the nuclear import pathway. *Gene Ther* 2001;**8**:166–71.
29. Fominaya J, Wels W. Target cell-specific DNA transfer mediated by a chimeric multidomain protein: novel non-viral gene delivery system. *J Biol Chem* 1996;**271**:10560–8.
30. Box M, Parks DA, Knight A, Hale C, Fishman PS, Fairweather NF. A multi-domain protein system based on the HC fragment of tetanus toxin for targeting DNA to neuronal cells. *J Drug Target* 2003;**11**:333–43.
31. Baleja JD, Thanabal V, Wagner G. Refined solution structure of the DNA-binding domain of GAL4 and use of 3J(113Cd,1H) in structure determination. *J Biomol NMR* 1997;**10**:397–401.
32. Bessis N, GarciaCozar FJ, Boissier MC. Immune responses to gene therapy vectors: influence on vector function and effector mechanisms. *Gene Ther* 2004;**11**(Suppl. 1):S10–7.
33. Vazquez E, Ferrer-Miralles N, Villaverde A. Peptide-assisted traffic engineering for nonviral gene therapy. *Drug Discov Today* 2008;**13**:1067–74.
34. Richard JP, Melikov K, Vives E, Ramos C, Verbeure B, Gait MJ, et al. Cell-penetrating peptides: a reevaluation of the mechanism of cellular uptake. *J Biol Chem* 2003;**278**:585–90.
35. Said HF, Saleh AF, Abes R, Gait MJ, Lebleu B. Cell penetrating peptides: overview and applications to the delivery of oligonucleotides. *Cell Mol Life Sci* 2010;**67**:715–26.
36. Vazquez E, Ferrer-Miralles N, Mangues R, Corchero JL, Schwartz S, Villaverde A. Modular protein engineering in emerging cancer therapies. *Curr Pharm Des* 2009;**15**:893–916.
37. Rudolph C, Plank C, Lausier J, Schillinger U, Muller RH, Rosenecker J. Oligomers of the arginine-rich motif of the HIV-1 TAT protein are capable of transferring plasmid DNA into cells. *J Biol Chem* 2003;**278**:11411–8.
38. Fawell S, Seery J, Daikh Y, Moore C, Chen LL, Pepinsky B, et al. Tat-mediated delivery of heterologous proteins into cells. *Proc Natl Acad Sci USA* 1994;**91**:664–8.
39. Brooks H, Lebleu B, Vives E. Tat peptide-mediated cellular delivery: back to basics. *Adv Drug Deliv Rev* 2005;**57**:559–77.
40. Vives E, Brodin P, Lebleu B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J Biol Chem* 1997;**272**:16010–7.
41. Vives E. Cellular uptake [correction of utake] of the Tat peptide: an endocytosis mechanism following ionic interactions. *J Mol Recognit* 2003;**16**:265–71.
42. Wender PA, Mitchell DJ, Pattabiraman K, Pelkey ET, Steinman L, Rothbard JB. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. *Proc Natl Acad Sci USA* 2000;**97**:13003–8.
43. Lundberg P, El-Andaloussi S, Sutlu T, Johansson H, Langel U. Delivery of short interfering RNA using endosomolytic cell-penetrating peptides. *FASEB J* 2007;**21**:2664–71.

44. Muratovska A, Eccles MR. Conjugate for efficient delivery of short interfering RNA (siRNA) into mammalian cells. *FEBS Lett* 2004;**558**:63–8.
45. Pooga M, Hallbrink M, Zorko M, Langel U. Cell penetration by transportan. *FASEB J* 1998;**12**:67–77.
46. Pooga M, Kut C, Kihlmark M, Hallbrink M, Fernaeus S, Raid R, et al. Cellular translocation of proteins by transportan. *FASEB J* 2001;**15**:1451–3.
47. Hart SL. Integrin-mediated vectors for gene transfer and therapy. *Curr Opin Mol Ther* 1999;**1**:197–203.
48. Kessler T, Bieker R, Padro T, Schwoppe C, Persigehl T, Bremer C, et al. Inhibition of tumor growth by RGD peptide-directed delivery of truncated tissue factor to the tumor vasculature. *Clin Cancer Res* 2005;**11**:6317–24.
49. Amara A, Gall SL, Schwartz O, Salamero J, Montes M, Loetscher P, et al. HIV coreceptor downregulation as antiviral principle: SDF-1alpha-dependent internalization of the chemokine receptor CXCR4 contributes to inhibition of HIV replication. *J Exp Med* 1997;**186**:139–46.
50. Liang X. CXCR4, inhibitors and mechanisms of action. *Chem Biol Drug Des* 2008;**72**:97–110.
51. Daniels TR, Delgado T, Helguera G, Penichet ML. The transferrin receptor part II: targeted delivery of therapeutic agents into cancer cells. *Clin Immunol* 2006;**121**:159–76.
52. Joung I, Harber G, Gerecke KM, Carroll SL, Collawn JF, Engler JA. Improved gene delivery into neuroglial cells using a fiber-modified adenovirus vector. *Biochem Biophys Res Commun* 2005;**328**:1182–7.
53. Liu X, Tian PK, Ju DW, Zhang MH, Yao M, Cao XT, et al. Systemic genetic transfer of p21WAF-1 and GM-CSF utilizing of a novel oligopeptide-based EGF receptor targeting polyplex. *Cancer Gene Ther* 2003;**10**:529–39.
54. Wolschek MF, Thallinger C, Kursa M, Rossler V, Allen M, Lichtenberger C, et al. Specific systemic nonviral gene delivery to human hepatocellular carcinoma xenografts in SCID mice. *Hepatology* 2002;**36**:1106–14.
55. Wu YT, Jiaang WT, Lin KG, Huang CM, Chang CH, Sun YL, et al. A new N-acetylgalactosamine containing peptide as a targeting vehicle for mammalian hepatocytes via asialoglycoprotein receptor endocytosis. *Curr Drug Deliv* 2004;**1**:119–27.
56. Kumar P, Wu H, McBride JL, Jung KE, Kim MH, Davidson BL, et al. Transvascular delivery of small interfering RNA to the central nervous system. *Nature* 2007;**448**:39–43.
57. Schneider H, Harbottle RP, Yokosaki Y, Kunde J, Sheppard D, Coutelle C. A novel peptide, PLAEIDGIELTY, for the targeting of alpha9beta1-integrins. *FEBS Lett* 1998;**429**:269–73.
58. Collins L, Fabre JW. A synthetic peptide vector system for optimal gene delivery to corneal endothelium. *J Gene Med* 2004;**6**:185–94.
59. McKay T, Reynolds P, Jezzard S, Curiel D, Coutelle C. Secretin-mediated gene delivery, a specific targeting mechanism with potential for treatment of biliary and pancreatic disease in cystic fibrosis. *Mol Ther* 2002;**5**:447–54.
60. Zeng J, Too HP, Ma Y, Luo ES, Wang S. A synthetic peptide containing loop 4 of nerve growth factor for targeted gene delivery. *J Gene Med* 2004;**6**:1247–56.
61. Martinez-Fong D, Navarro-Quiroga I, Ochoa I, varez-Maya I, Meraz MA, Luna J, et al. Neurotensin-SPDP-poly-L-lysine conjugate: a nonviral vector for targeted gene delivery to neural cells. *Brain Res Mol Brain Res* 1999;**69**:249–62.
62. White SJ, Nicklin SA, Sawamura T, Baker AH. Identification of peptides that target the endothelial cell-specific LOX-1 receptor. *Hypertension* 2001;**37**:449–55.
63. Pardridge WM, Buciak JL, Friden PM. Selective transport of an anti-transferrin receptor antibody through the blood-brain barrier in vivo. *J Pharmacol Exp Ther* 1991;**259**:66–70.

64. Boado RJ, Zhang Y, Zhang Y, Pardridge WM. Humanization of anti-human insulin receptor antibody for drug targeting across the human blood-brain barrier. *Biotechnol Bioeng* 2007;**96**:381–91.
65. Jeong JH, Lee M, Kim WJ, Yockman JW, Park TG, Kim YH, et al. Anti-GAD antibody targeted non-viral gene delivery to islet beta cells. *J Control Release* 2005;**107**:562–70.
66. Balestrieri ML, Napoli C. Novel challenges in exploring peptide ligands and corresponding tissue-specific endothelial receptors. *Eur J Cancer* 2007;**43**:1242–50.
67. Aina OH, Liu R, Sutcliffe JL, Marik J, Pan CX, Lam KS. From combinatorial chemistry to cancer-targeting peptides. *Mol Pharm* 2007;**4**:631–51.
68. Lundberg M, Wikstrom S, Johansson M. Cell surface adherence and endocytosis of protein transduction domains. *Mol Ther* 2003;**8**:143–50.
69. Wagner E, Plank C, Zatloukal K, Cotten M, Birnstiel ML. Influenza virus hemagglutinin HA-2 N-terminal fusogenic peptides augment gene transfer by transferrin-polylysine-DNA complexes: toward a synthetic virus-like gene-transfer vehicle. *Proc Natl Acad Sci USA* 1992;**89**:7934–8.
70. Plank C, Oberhauser B, Mechtler K, Koch C, Wagner E. The influence of endosome-disruptive peptides on gene transfer using synthetic virus-like gene transfer systems. *J Biol Chem* 1994;**269**:12918–24.
71. Subbarao NK, Parente RA, Szoka Jr. FC, Nadasdi L, Pongracz K. pH-dependent bilayer destabilization by an amphipathic peptide. *Biochemistry* 1987;**26**:2964–72.
72. Parente RA, Nir S, Szoka Jr. FC. Mechanism of leakage of phospholipid vesicle contents induced by the peptide GALA. *Biochemistry* 1990;**29**:8720–8.
73. Parente RA, Nadasdi L, Subbarao NK, Szoka Jr. FC. Association of a pH-sensitive peptide with membrane vesicles: role of amino acid sequence. *Biochemistry* 1990;**29**:8713–9.
74. Li W, Nicol F, Szoka Jr. FC. GALA: a designed synthetic pH-responsive amphipathic peptide with applications in drug and gene delivery. *Adv Drug Deliv Rev* 2004;**56**:967–85.
75. Wyman TB, Nicol F, Zelphati O, Scaria PV, Plank C, Szoka Jr. FC. Design, synthesis, and characterization of a cationic peptide that binds to nucleic acids and permeabilizes bilayers. *Biochemistry* 1997;**36**:3008–17.
76. Rittner K, Benavente A, Bompard-Sorlet A, Heitz F, Divita G, Brasseur R, et al. New basic membrane-destabilizing peptides for plasmid-based gene delivery in vitro and in vivo. *Mol Ther* 2002;**5**:104–14.
77. Ogris M, Carlisle RC, Bettinger T, Seymour LW. Melittin enables efficient vesicular escape and enhanced nuclear access of nonviral gene delivery vectors. *J Biol Chem* 2001;**276**:47550–5.
78. Derossi D, Joliot AH, Chassaing G, Prochiantz A. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J Biol Chem* 1994;**269**:10444–50.
79. Derossi D, Chassaing G, Prochiantz A. Trojan peptides: the penetratin system for intracellular delivery. *Trends Cell Biol* 1998;**8**:84–7.
80. Pichon C, Goncalves C, Midoux P. Histidine-rich peptides and polymers for nucleic acids delivery. *Adv Drug Deliv Rev* 2001;**53**:75–94.
81. Midoux P, Kichler A, Boutin V, Maurizot JC, Monsigny M. Membrane permeabilization and efficient gene transfer by a peptide containing several histidines. *Bioconj Chem* 1998;**9**:260–7.
82. Midoux P, LeCam E, Coulaud D, Delain E, Pichon C. Histidine containing peptides and polypeptides as nucleic acid vectors. *Somat Cell Mol Genet* 2002;**27**:27–47.
83. Kichler A, Leborgne C, Danos O, Bechinger B. Characterization of the gene transfer process mediated by histidine-rich peptides. *J Mol Med* 2007;**85**:191–201.
84. Kim J, Chen CP, Rice KG. The proteasome metabolizes peptide-mediated nonviral gene delivery systems. *Gene Ther* 2005;**12**:1581–90.

85. Leonchiks A, Stavropoulou V, Sharipo A, Masucci MG. Inhibition of ubiquitin-dependent proteolysis by a synthetic glycine-alanine repeat peptide that mimics an inhibitory viral sequence. *FEBS Lett* 2002;**522**:93–8.
86. Sharipo A, Imreh M, Leonchiks A, Imreh S, Masucci MG. A minimal glycine-alanine repeat prevents the interaction of ubiquitinated I kappaB alpha with the proteasome: a new mechanism for selective inhibition of proteolysis. *Nat Med* 1998;**4**:939–44.
87. Sharipo A, Imreh M, Leonchiks A, Branden C, Masucci MG. cis-Inhibition of proteasomal degradation by viral repeats: impact of length and amino acid composition. *FEBS Lett* 2001;**499**:137–42.
88. Martinez-Moreno M, Navarro-Lerida I, Roncal F, Albar JP, Alonso C, Gavilanes F, et al. Recognition of novel viral sequences that associate with the dynein light chain LC8 identified through a pepscan technique. *FEBS Lett* 2003;**544**:262–7.
89. Melchior F, Gerace L. Mechanisms of nuclear protein import. *Curr Opin Cell Biol* 1995;**7**:310–8.
90. Zanta MA, Belguise-Valladier P, Behr JP. Gene delivery: a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus. *Proc Natl Acad Sci USA* 1999;**96**:91–6.
91. Fischer N, Kremmer E, Lautscham G, Mueller-Lantzsch N, Grasser FA. Epstein-Barr virus nuclear antigen 1 forms a complex with the nuclear transporter karyopherin alpha2. *J Biol Chem* 1997;**272**:3999–4005.
92. Dang CV, Lee WM. Identification of the human c-myc protein nuclear translocation signal. *Mol Cell Biol* 1988;**8**:4048–54.
93. Fominaya J, Uherek C, Wels W. A chimeric fusion protein containing transforming growth factor-alpha mediates gene transfer via binding to the EGF receptor. *Gene Ther* 1998;**5**:521–30.
94. Bauerle M, Doenecke D, Albig W. The requirement of H1 histones for a heterodimeric nuclear import receptor. *J Biol Chem* 2002;**277**:32480–9.
95. Baake M, Bauerle M, Doenecke D, Albig W. Core histones and linker histones are imported into the nucleus by different pathways. *Eur J Cell Biol* 2001;**80**:669–77.
96. Subramanian A, Ranganathan P, Diamond SL. Nuclear targeting peptide scaffolds for lipofection of nondividing mammalian cells. *Nat Biotechnol* 1999;**17**:873–7.
97. Gharakhanian E, Takahashi J, Kasamatsu H. The carboxyl 35 amino acids of SV40 Vp3 are essential for its nuclear accumulation. *Virology* 1987;**157**:440–8.
98. Lyons RH, Ferguson BQ, Rosenberg M. Pentapeptide nuclear localization signal in adenovirus E1a. *Mol Cell Biol* 1987;**7**:2451–6.
99. Kleinschmidt JA, Seiter A. Identification of domains involved in nuclear uptake and histone binding of protein N1 of Xenopus laevis. *EMBO J* 1988;**7**:1605–14.
100. Kiefer P, Acland P, Pappin D, Peters G, Dickson C. Competition between nuclear localization and secretory signals determines the subcellular fate of a single CUG-initiated form of FGF3. *EMBO J* 1994;**13**:4126–36.
101. Schreiber V, Molinete M, Boeuf H, de Murcia G, Menissier-de MJ. The human poly(ADP-ribose) polymerase nuclear localization signal is a bipartite element functionally separate from DNA binding and catalytic activity. *EMBO J* 1992;**11**:3263–9.
102. Robbins J, Dilworth SM, Laskey RA, Dingwall C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* 1991;**64**:615–23.
103. Rivera VM, Gao GP, Grant RL, Schnell MA, Zoltick PW, Rozamus LW, et al. Long-term pharmacologically regulated expression of erythropoietin in primates following AAV-mediated gene transfer. *Blood* 2005;**105**:1424–30.

104. Wang X, Qiu R, Tsark W, Lu Q. Rapid promoter analysis in developing mouse brain and genetic labeling of young neurons by doublecortin-DsRed-express. *J Neurosci Res* 2007;**85**:3567–73.
105. Ishiwata A, Mimuro J, Mizukami H, Kashiwakura Y, Takano K, Ohmori T, et al. Liver-restricted expression of the canine factor VIII gene facilitates prevention of inhibitor formation in factor VIII-deficient mice. *J Gene Med* 2009;**11**:1020–9.
106. Pardridge WM. Blood-brain barrier delivery. *Drug Discov Today* 2007;**12**:54–61.
107. Tosi G, Costantino L, Rivasi F, Ruozi B, Leo E, Vergoni AV, et al. Targeting the central nervous system: in vivo experiments with peptide-derivatized nanoparticles loaded with Loperamide and Rhodamine-123. *J Control Release* 2007;**122**:1–9.
108. Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF. In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science* 1999;**285**:1569–72.
109. Boado RJ, Zhang Y, Zhang Y, Pardridge WM. Genetic engineering, expression, and activity of a fusion protein of a human neurotrophin and a molecular Trojan horse for delivery across the human blood-brain barrier. *Biotechnol Bioeng* 2007;**97**:1376–86.
110. Boado RJ, Zhang Y, Zhang Y, Xia CF, Wang Y, Pardridge WM. Genetic engineering of a lysosomal enzyme fusion protein for targeted delivery across the human blood-brain barrier. *Biotechnol Bioeng* 2008;**99**:475–84.
111. Boado RJ, Zhang Y, Zhang Y, Wang Y, Pardridge WM. GDNF fusion protein for targeted-drug delivery across the human blood-brain barrier. *Biotechnol Bioeng* 2008;**100**:387–96.
112. Otten LG, Quax WJ. Directed evolution: selecting today's biocatalysts. *Biomol Eng* 2005;**22**:1–9.
113. Lutz S, Patrick WM. Novel methods for directed evolution of enzymes: quality, not quantity. *Curr Opin Biotechnol* 2004;**15**:291–7.
114. Neylon C. Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res* 2004;**32**:1448–59.
115. Stemmer WP. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 1994;**370**:389–91.
116. Dahiyat BI, Mayo SL. Protein design automation. *Protein Sci* 1996;**5**:895–903.
117. Floudas CA, Fung HK, McAllister SR, Mönnigmann M, Rajgaria R. Advances in protein structure prediction and de novo protein design: a review. *Chem Eng Sci* 2006;**61**:966–88.
118. Walsh ST, Cheng H, Bryson JW, Roder H, DeGrado WF. Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proc Natl Acad Sci USA* 1999;**96**:5486–91.
119. Dahiyat BI, Mayo SL. De novo protein design: fully automated sequence selection. *Science* 1997;**278**:82–7.
120. Aris A, Villaverde A. Modular protein engineering for non-viral gene therapy. *Trends Biotechnol* 2004;**22**:371–7.
121. Uherek C, Fominaya J, Wels W. A modular DNA carrier protein based on the structure of diphtheria toxin mediates target cell-specific gene delivery. *J Biol Chem* 1998;**273**:8835–41.
122. Li SD, Huang L. Gene therapy progress and prospects: non-viral gene therapy by systemic delivery. *Gene Ther* 2006;**13**:1313–9.
123. Chan CK, Jans DA. Using nuclear targeting signals to enhance non-viral gene transfer. *Immunol Cell Biol* 2002;**80**:119–30.
124. Dietz GP, Bahr M. Delivery of bioactive molecules into the cell: the Trojan horse approach. *Mol Cell Neurosci* 2004;**27**:85–131.
125. Dietz GP, Bahr M. Synthesis of cell-penetrating peptides and their application in neurobiology. *Methods Mol Biol* 2007;**399**:181–98.

126. Ignatovich IA, Dizhe EB, Pavlotskaya AV, Akifiev BN, Burov SV, Orlov SV, et al. Complexes of plasmid DNA with basic domain 47–57 of the HIV-1 Tat protein are transferred to mammalian cells by endocytosis-mediated pathways. *J Biol Chem* 2003;**278**:42625–36.
127. Aris A, Villaverde A. Molecular organization of protein-DNA complexes for cell-targeted DNA delivery. *Biochem Biophys Res Commun* 2000;**278**:455–61.
128. Carrio MM, Villaverde A. Role of molecular chaperones in inclusion body formation. *FEBS Lett* 2003;**537**:215–21.
129. Baneyx F. Recombinant protein expression in Escherichia coli. *Curr Opin Biotechnol* 1999;**10**:411–21.
130. Sorensen HP, Mortensen KK. Advanced genetic strategies for recombinant protein expression in Escherichia coli. *J Biotechnol* 2005;**115**:113–28.
131. Schmidt FR. Recombinant expression systems in the pharmaceutical industry. *Appl Microbiol Biotechnol* 2004;**65**:363–72.
132. Ou WC, Wang M, Fung CY, Tsai RT, Chao PC, Hseu TH, et al. The major capsid protein, VP1, of human JC virus expressed in Escherichia coli is able to self-assemble into a capsid-like particle and deliver exogenous DNA into human kidney cells. *J Gen Virol* 1999;**80**:39–46.
133. Peluffo H, Acarin L, Aris A, Gonzalez P, Villaverde A, Castellano B, 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* 2006;**7**:35.
134. Sasnauskas K, Buzaite O, Vogel F, Jandrig B, Razanskas R, Staniulis J, et al. Yeast cells allow high-level expression and formation of polyomavirus-like particles. *Biol Chem* 1999;**380**:381–6.
135. Birch JR, Froud SJ. Mammalian cell culture systems for recombinant protein production. *Biologicals* 1994;**22**:127–33.
136. Ikonomou L, Schneider YJ, Agathos SN. Insect cell culture for industrial production of recombinant proteins. *Appl Microbiol Biotechnol* 2003;**62**:1–20.
137. Reiter M, Bluml G. Large-scale mammalian cell culture. *Curr Opin Biotechnol* 1994;**5**:175–9.
138. Gervais A, Hammel YA, Pelloux S, Lepage P, Baer G, Carte N, et al. Glycosylation of human recombinant gonadotrophins: characterization and batch-to-batch consistency. *Glycobiology* 2003;**13**:179–89.
139. Collins L, Asuni AA, Anderton BH, Fabre JW. Efficient gene delivery to primary neuron cultures using a synthetic peptide vector system. *J Neurosci Methods* 2003;**125**:113–20.
140. Kim ES, Yang SW, Hong DK, Kim WT, Kim HG, Lee SK. Cell-penetrating DNA-binding protein as a safe and efficient naked DNA delivery carrier in vitro and in vivo. *Biochem Biophys Res Commun* 2010;**392**:9–15.
141. Fabre JW, Collins L. Synthetic peptides as non-viral DNA vectors. *Curr Gene Ther* 2006;**6**:459–80.
142. Parker AL, Collins L, Zhang X, Fabre JW. Exploration of peptide motifs for potent non-viral gene delivery highly selective for dividing cells. *J Gene Med* 2005;**7**:1545–54.
143. Aris A, Villaverde A. Engineering nuclear localization signals in modular protein vehicles for gene therapy. *Biochem Biophys Res Commun* 2003;**304**:625–31.
144. Chen L, McCrate JM, Lee JC, Li H. The role of surface charge on the uptake and biocompatibility of hydroxyapatite nanoparticles with osteoblast cells. *Nanotechnology* 2011;**22**:105708.
145. Cortes-Perez NG, Sapin C, Jaffrelo L, Daou S, Grill JP, Langella P, et al. Rotavirus-like particles: a novel nanocarrier for the gut. *J Biomed Biotechnol* 2010;**2010**:317545.
146. Alcala P, Feliu JX, Aris A, Villaverde A. Efficient accommodation of recombinant, foot-and-mouth disease virus RGD peptides to cell-surface integrins. *Biochem Biophys Res Commun* 2001;**285**:201–6.
147. Marsh M, Helenius A. Virus entry: open sesame. *Cell* 2006;**124**:729–40.

148. Luthman H, Magnusson G. High efficiency polyoma DNA transfection of chloroquine treated cells. *Nucleic Acids Res* 1983;**11**:1295–308.
149. Vaccaro L, Cross KJ, Kleinjung J, Straus SK, Thomas DJ, Wharton SA, et al. Plasticity of influenza haemagglutinin fusion peptides and their interaction with lipid bilayers. *Biophys J* 2005;**88**:25–36.
150. Ludtke JJ, Zhang G, Sebestyen MG, Wolff JA. A nuclear localization signal can enhance both the nuclear transport and expression of 1 kb DNA. *J Cell Sci* 1999;**112**:2033–41.
151. Wilke M, Fortunati E, van den Broek M, Hoogeveen AT, Scholte BJ. Efficacy of a peptide-based gene delivery system depends on mitotic activity. *Gene Ther* 1996;**3**:1133–42.
152. Seow Y, Wood MJ. Biological gene delivery vehicles: beyond viral vectors. *Mol Ther* 2009;**17**:767–77.
153. Ludwig C, Wagner R. Virus-like particles-universal molecular toolboxes. *Curr Opin Biotechnol* 2007;**18**:537–45.
154. Mateu MG. Virus engineering: functionalization and stabilization. *Protein Eng Des Sel* 2011;**24**:53–63.
155. Singh P, Gonzalez MJ, Manchester M. Viruses and their uses in nanotechnology. *Drug Dev Res* 2006;**67**:23–41.
156. Dokland T. Scaffolding proteins and their role in viral assembly. *Cell Mol Life Sci* 1999;**56**:580–603.
157. Pattenden LK, Middelberg AP, Niebert M, Lipin DI. Towards the preparative and large-scale precision manufacture of virus-like particles. *Trends Biotechnol* 2005;**23**:523–9.
158. Scheerlinck JPY, Greenwood DLV. Virus-sized vaccine delivery systems. *Drug Discov Today* 2008;**13**:882–7.
159. Roldao A, Mellado MC, Castilho LR, Carrondo MJ, Alves PM. Virus-like particles in vaccine development. *Expert Rev Vaccines* 2010;**9**:1149–76.
160. Sanchez-Rodriguez SP, Munch-Anguiano L, Bustos-Jaimes I. Advances in the development of virus-like particles as tools in medicine and nanoscience. *Curr Chem Biol* 2010;**4**:231–43.
161. Ramqvist T, Andreasson K, Dalianis T. Vaccination, immune and gene therapy based on virus-like particles against viral infections and cancer. *Expert Opin Biol Ther* 2007;**7**:997–1007.
162. Roy P, Noad R. Virus-like particles as a vaccine delivery system: myths and facts. *Hum Vaccin* 2008;**4**:5–12.
163. Slilaty SN, Aposhian HV. Gene transfer by polyoma-like particles assembled in a cell-free system. *Science* 1983;**220**:725–7.
164. Lenz P, Day PM, Pang YYS, Frye SA, Jensen PN, Lowy DR, et al. Papillomavirus-like particles induce acute activation of dendritic cells. *J Immunol* 2001;**166**:5346–55.
165. Xu YF, Zhang YQ, Xu XM, Song GX. Papillomavirus virus-like particles as vehicles for the delivery of epitopes or genes. *Arch Virol* 2006;**151**:2133–48.
166. Buonaguro FM, Tornesello ML, Buonaguro L. Virus-like particle vaccines and adjuvants: the HPV paradigm. *Expert Rev Vaccines* 2009;**8**:1379–98.
167. Brandenburg B, Stockl L, Gutzeit C, Roos M, Lupberger J, Schwartlander R, et al. A novel system for efficient gene transfer into primary human hepatocytes via cell-permeable hepatitis B virus-like particle. *Hepatology* 2005;**42**:1300–9.
168. Li TC, Takeda N, Miyamura T, Matsuura Y, Wang JC, Engvall H, et al. Essential elements of the capsid protein for self-assembly into empty virus-like particles of hepatitis E virus. *J Virol* 2005;**79**:12999–3006.
169. Xiang J, Wunschmann S, George SL, Klinzman D, Schmidt WN, LaBrecque DR, et al. Recombinant hepatitis C virus-like particles expressed by baculovirus: utility in cell-binding and antibody detection assays. *J Med Virol* 2002;**68**:537–43.

170. Andreasson K, Tegerstedt K, Eriksson M, Curcio C, Cavallo F, Forni G, et al. Murine pneumotropic virus chimeric Her2/neu virus-like particles as prophylactic and therapeutic vaccines against Her2/neu expressing tumors. *Int J Cancer* 2009;**124**:150–6.
171. Chou MI, Hsieh YF, Wang M, Chang JT, Chang D, Zouali M, et al. In vitro and in vivo targeted delivery of IL-10 interfering RNA by JC virus-like particles. *J Biomed Sci* 2010;**17**:51.
172. Gleiter S, Lilie H. Cell-type specific targeting and gene expression using a variant of polyoma VP1 virus-like particles. *Biol Chem* 2003;**384**:247–55.
173. Goldmann C, Petry H, Frye S, Ast O, Ebitsch S, Jentsch KD, et al. Molecular cloning and expression of major structural protein VP1 of the human polyomavirus JC virus: formation of virus-like particles useful for immunological and therapeutic studies. *J Virol* 1999;**73**:4465–9.
174. Goldmann C, Stolte N, Nisslein T, Hunsmann G, Luke W, Petry H. Packaging of small molecules into VP1-virus-like particles of the human polyomavirus JC virus. *J Virol Methods* 2000;**90**:85–90.
175. Lawatscheck R, Aleksaite E, Schenk JA, Micheel B, Jandrig B, Holland G, et al. Chimeric polyomavirus-derived virus-like particles: the immunogenicity of an inserted peptide applied without adjuvant to mice depends on its insertion site and its flanking linker sequence. *Viral Immunol* 2007;**20**:453–60.
176. Sasnauskas K, Bulavaite A, Hale A, Jin L, Knowles WA, Gedvilaite A, et al. Generation of recombinant virus-like particles of human and non-human polyomaviruses in yeast Saccharomyces cerevisiae. *Intervirology* 2002;**45**:308–17.
177. Voronkova T, Kazaks A, Ose V, Ozel M, Scherneck S, Pumpens P, et al. Hamster polyomavirus-derived virus-like particles are able to transfer in vitro encapsidated plasmid DNA to mammalian cells. *Virus Genes* 2007;**34**:303–14.
178. Krauzewicz N, Stokrova J, Jenkins C, Elliott M, Higgins CF, Griffin BE. Virus-like gene transfer into cells mediated by polyoma virus pseudocapsids. *Gene Ther* 2000;**7**:2122–31.
179. Clark B, Caparros-Wanderley W, Musselwhite G, Kotecha M, Griffin BE. Immunity against both polyomavirus VP1 and a transgene product induced following intranasal delivery of VP1 pseudocapsid-DNA complexes. *J Gen Virol* 2001;**82**:2791–7.
180. Young KR, McBurney SP, Karkhanis LU, Ross TM. Virus-like particles: designing an effective AIDS vaccine. *Methods* 2006;**40**:98–117.
181. Muratori C, Bona R, Federico M. Lentivirus-based virus-like particles as a new protein delivery tool. *Methods Mol Biol* 2010;**614**:111–24.
182. Istrate C, Hinkula J, Charpilienne A, Poncet D, Cohen J, Svensson L, et al. Parenteral administration of RF 8-2/6/7 rotavirus-like particles in a one-dose regimen induce protective immunity in mice. *Vaccine* 2008;**26**:4594–601.
183. Singh P, Destito G, Schneemann A, Manchester M. Canine parvovirus-like particles, a novel nanomaterial for tumor targeting. *J Nanobiotechnology* 2006;**4**:2.
184. Singh P. *Tumor targeting using canine parvovirus nanoparticles*. Berlin: Springer-Verlag; 2009 pp. 123–141.
185. Herbst-Kralovetz M, Mason HS, Chen Q. Norwalk virus-like particles as vaccines. *Expert Rev Vaccines* 2010;**9**:299–307.
186. Manchester M, Singh P. Virus-based nanoparticles (VNPs): platform technologies for diagnostic imaging. *Adv Drug Deliv Rev* 2006;**58**:1505–22.
187. Pease III LF, Lipin DI, Tsai DH, Zachariah MR, Lua LH, Tarlov MJ, et al. Quantitative characterization of virus-like particles by asymmetrical flow field flow fractionation, electrospray differential mobility analysis, and transmission electron microscopy. *Biotechnol Bioeng* 2009;**102**:845–55.
188. Stehle T, Harrison SC. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. *EMBO J* 1997;**16**:5139–48.

189. Neu U, Stehle T, Atwood WJ. The Polyomaviridae: contributions of virus structure to our understanding of virus receptors and infectious entry. *Virology* 2009;**384**:389–99.
190. Johnson JE, Chiu W. Structures of virus and virus-like particles. *Curr Opin Struct Biol* 2000;**10**:229–35.
191. Zhang S. Fabrication of novel biomaterials through molecular self-assembly. *Nat Biotechnol* 2003;**21**:1171–8.
192. Ivanovska IL, de Pablo PJ, Ibarra B, Sgalari G, MacKintosh FC, Carrascosa JL, et al. Bacteriophage capsids: tough nanoshells with complex elastic properties. *Proc Natl Acad Sci USA* 2004;**101**:7600–5.
193. Bothner B, Taylor D, Jun B, Lee KK, Siuzdak G, Schultz CP, et al. Maturation of a tetravirus capsid alters the dynamic properties and creates a metastable complex. *Virology* 2005;**334**:17–27.
194. Petry H, Goldmann C, Ast O, Luke W. The use of virus-like particles for gene transfer. *Curr Opin Mol Ther* 2003;**5**:524–8.
195. Garcea RL, Gissmann L. Virus-like particles as vaccines and vessels for the delivery of small molecules. *Curr Opin Biotechnol* 2004;**15**:513–7.
196. Georgens C, Weyermann J, Zimmer A. Recombinant virus like particles as drug delivery system. *Curr Pharm Biotechnol* 2005;**6**:49–55.
197. Noad R, Roy P. Virus-like particles as immunogens. *Trends Microbiol* 2003;**11**:438–44.
198. Grgacic EV, Anderson DA. Virus-like particles: passport to immune recognition. *Methods* 2006;**40**:60–5.
199. Carrasco C, Castellanos M, de Pablo PJ, Mateu MG. Manipulation of the mechanical properties of a virus by protein engineering. *Proc Natl Acad Sci USA* 2008;**105**:4150–5.
200. Fischlechner M, Donath E. Viruses as building blocks for materials and devices. *Angew Chem Int Ed Engl* 2007;**46**:3184–93.
201. Lee LA, Wang Q. Adaptations of nanoscale viruses and other protein cages for medical applications. *Nanomedicine* 2006;**2**:137–49.
202. Adjei AA. Blocking oncogenic Ras signaling for cancer therapy. *J Natl Cancer Inst* 2001;**93**:1062–74.
203. May T, Gleiter S, Lilie H. Assessment of cell type specific gene transfer of polyoma virus like particles presenting a tumor specific antibody Fv fragment. *J Virol Methods* 2002;**105**:147–57.
204. Liew MW, Rajendran A, Middelberg AP. Microbial production of virus-like particle vaccine protein at gram-per-litre levels. *J Biotechnol* 2010;**150**:224–31.
205. Kost TA, Condreay JP. Recombinant baculoviruses as mammalian cell gene-delivery vectors. *Trends Biotechnol* 2002;**20**:173–80.
206. Maranga L, Cruz PE, Aunins JG, Carrondo MJ. Production of core and virus-like particles with baculovirus infected insect cells. *Adv Biochem Eng Biotechnol* 2002;**74**:183–206.
207. Lee CD, Yan YP, Liang SM, Wang TF. Production of FMDV virus-like particles by a SUMO fusion protein approach in Escherichia coli. *J Biomed Sci* 2009;**16**:69.
208. Santi L, Huang Z, Mason H. Virus-like particles production in green plants. *Methods* 2006;**40**:66–76.
209. Santi L, Batchelor L, Huang Z, Hjelm B, Kilbourne J, Arntzen CJ, et al. An efficient plant viral expression system generating orally immunogenic Norwalk virus-like particles. *Vaccine* 2008;**26**:1846–54.
210. Ramqvist T, Dalianis T. Immunotherapeutic polyoma and human papilloma virus-like particles. *Immunotherapy* 2009;**1**:303–12.
211. McAleer WJ, Buynak EB, Maigetter RZ, Wampler DE, Miller WJ, Hilleman MR. Human hepatitis B vaccine from recombinant yeast. *Nature* 1984;**307**:178–80.

212. Pumpens P, Razanskas R, Pushko P, Renhof R, Gusars I, Skrastina D, et al. Evaluation of HBs, HBc, and frCP virus-like particles for expression of human papillomavirus 16 E7 oncoprotein epitopes. *Intervirology* 2002;**45**:24–32.
213. Hammonds J, Chen X, Zhang X, Lee F, Spearman P. Advances in methods for the production, purification, and characterization of HIV-1 Gag-Env pseudovirion vaccines. *Vaccine* 2007;**25**:8036–48.
214. Peek LJ, Middaugh CR, Berkland C. Nanotechnology in vaccine delivery. *Adv Drug Deliv Rev* 2008;**60**:915–28.
215. Sun Y, Carrion Jr. R, Ye L, Wen Z, Ro YT, Brasky K, et al. Protection against lethal challenge by Ebola virus-like particles produced in insect cells. *Virology* 2009;**383**:12–21.
216. Warfield KL, Bosio CM, Welcher BC, Deal EM, Mohamadzadeh M, Schmaljohn A, et al. Ebola virus-like particles protect from lethal Ebola virus infection. *Proc Natl Acad Sci USA* 2003;**100**:15889–94.
217. Freyschmidt EJ, Alonso A, Hartmann G, Gissmann L. Activation of dendritic cells and induction of T cell responses by HPV 16 L1/E7 chimeric virus-like particles are enhanced by CpG ODN or sorbitol. *Antivir Ther* 2004;**9**:479–89.
218. Kaufmann AM, Nieland JD, Jochmus I, Baur S, Friese K, Gabelsberger J, et al. Vaccination trial with HPV16 L1E7 chimeric virus-like particles in women suffering from high grade cervical intraepithelial neoplasia (CIN 2/3). *Int J Cancer* 2007;**121**:2794–800.
219. Kaufmann AM, Nieland J, Schinz M, Nonn M, Gabelsberger J, Meissner H, et al. HPV16 L1E7 chimeric virus-like particles induce specific HLA-restricted T cells in humans after in vitro vaccination. *Int J Cancer* 2001;**92**:285–93.
220. Cheong WS, Reiseger J, Turner SJ, Boyd R, Netter HJ. Chimeric virus-like particles for the delivery of an inserted conserved influenza A-specific CTL epitope. *Antiviral Res* 2009;**81**:113–22.
221. Yin Y, Zhang J, Dong D, Liu S, Guo Q, Song X, et al. Chimeric hepatitis B virus core particles carrying an epitope of anthrax protective antigen induce protective immunity against Bacillus anthracis. *Vaccine* 2008;**26**:5814–21.
222. Amexis G, Young NS. Parvovirus B19 empty capsids as antigen carriers for presentation of antigenic determinants of dengue 2 virus. *J Infect Dis* 2006;**194**:790–4.
223. Ogasawara Y, Amexis G, Yamaguchi H, Kajigaya S, Leppla SH, Young NS. Recombinant viral-like particles of parvovirus B19 as antigen carriers of anthrax protective antigen. *In Vivo* 2006;**20**:319–24.
224. Chackerian B, Lowy DR, Schiller JT. Conjugation of a self-antigen to papillomavirus-like particles allows for efficient induction of protective autoantibodies. *J Clin Invest* 2001;**108**:415–23.
225. Cornuz J, Zwahlen S, Jungi WF, Osterwalder J, Klingler K, Van MG, et al. A vaccine against nicotine for smoking cessation: a randomized controlled trial. *PLoS One* 2008;**3**:e2547.
226. Chen LS, Wang M, Ou WC, Fung CY, Chen PL, Chang CF, et al. Efficient gene transfer using the human JC virus-like particle that inhibits human colon adenocarcinoma growth in a nude mouse model. *Gene Ther* 2010;**17**:1033–41.
227. Ramsey JD, Vu HN, Pack DW. A top-down approach for construction of hybrid polymer-virus gene delivery vectors. *J Control Release* 2010;**144**:39–45.
228. Osterman JV, Waddell A, Aposhian HV. DNA and gene therapy: uncoating of polyoma pseudovirus in mouse embryo cells. *Proc Natl Acad Sci USA* 1970;**67**:37–40.
229. Yamada T, Iwasaki Y, Tada H, Iwabuki H, Chuah MK, VandenDriessche T, et al. Nanoparticles for the delivery of genes and drugs to human hepatocytes. *Nat Biotechnol* 2003;**21**:885–90.
230. Takamura S, Niikura M, Li TC, Takeda N, Kusagawa S, Takebe Y, et al. DNA vaccine-encapsulated virus-like particles derived from an orally transmissible virus stimulate mucosal and systemic immune responses by oral administration. *Gene Ther* 2004;**11**:628–35.

231. Kimchi-Sarfaty C, Brittain S, Garfield S, Caplen NJ, Tang Q, Gottesman MM. Efficient delivery of RNA interference effectors via in vitro-packaged SV40 pseudovirions. *Hum Gene Ther* 2005;**16**:1110–5.
232. Asokan A, Conway JC, Phillips JL, Li C, Hegge J, Sinnott R, et al. Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle. *Nat Biotechnol* 2010;**28**:79–82.
233. Windram OP, Weber B, Jaffer MA, Rybicki EP, Shepherd DN, Varsani A. An investigation into the use of human papillomavirus type 16 virus-like particles as a delivery vector system for foreign proteins: N- and C-terminal fusion of GFP to the L1 and L2 capsid proteins. *Arch Virol* 2008;**153**:585–9.
234. Gleiter S, Lilie H. Coupling of antibodies via protein Z on modified polyoma virus-like particles. *Protein Sci* 2001;**10**:434–44.
235. Stubenrauch K, Gleiter S, Brinkmann U, Rudolph R, Lilie H. Conjugation of an antibody Fv fragment to a virus coat protein: cell-specific targeting of recombinant polyoma-virus-like particles. *Biochem J* 2001;**356**:867–73.
236. Sen GS, Kuzelka J, Singh P, Lewis WG, Manchester M, Finn MG. Accelerated bioorthogonal conjugation: a practical method for the ligation of diverse functional molecules to a polyvalent virus scaffold. *Bioconj Chem* 2005;**16**:1572–9.
237. Schaffer DV, Koerber JT, Lim KI. Molecular engineering of viral gene delivery vehicles. *Annu Rev Biomed Eng* 2008;**10**:169–94.
238. Kobinger GP, Weiner DJ, Yu QC, Wilson JM. Filovirus-pseudotyped lentiviral vector can efficiently and stably transduce airway epithelia in vivo. *Nat Biotechnol* 2001;**19**:225–30.
239. Qu Q, Sawa H, Suzuki T, Semba S, Henmi C, Okada Y, et al. Nuclear entry mechanism of the human polyomavirus JC virus-like particle: role of importins and the nuclear pore complex. *J Biol Chem* 2004;**279**:27735–42.
240. Raja KS, Wang Q, Gonzalez MJ, Manchester M, Johnson JE, Finn MG. Hybrid virus-polymer materials. 1. Synthesis and properties of PEG-decorated cowpea mosaic virus. *Biomacromolecules* 2003;**4**:472–6.
241. Marianayagam NJ, Sunde M, Matthews JM. The power of two: protein dimerization in biology. *Trends Biochem Sci* 2004;**29**:618–25.
242. Tu RS, Tirrell M. Bottom-up design of biomimetic assemblies. *Adv Drug Deliv Rev* 2004;**56**:1537–63.
243. Leckband D. Measuring the forces that control protein interactions. *Annu Rev Biophys Biomol Struct* 2000;**29**:1–26.
244. Meyer EE, Rosenberg KJ, Israelachvili J. Recent progress in understanding hydrophobic interactions. *Proc Natl Acad Sci USA* 2006;**103**:15739–46.
245. Schneider HJ. Binding mechanisms in supramolecular complexes. *Angew Chem Int Ed Engl* 2009;**48**:3924–77.
246. Goodsell DS, Olson AJ. Structural symmetry and protein function. *Annu Rev Biophys Biomol Struct* 2000;**29**:105–53.
247. Eriksson J, Fenyo D. Modeling experimental design for proteomics. *Methods Mol Biol* 2010;**673**:223–30.
248. Kopecek J. Smart and genetically engineered biomaterials and drug delivery systems. *Eur J Pharm Sci* 2003;**20**:1–16.
249. Zhao X, Zhang S. Molecular designer self-assembling peptides. *Chem Soc Rev* 2006;**35**:1105–10.
250. Domingo-Espin J, Vazquez E, Ganz J, Conchillo O, Garcia-Fruitos E, Cedano J, et al. Nanoparticulate architecture of protein-based artificial viruses is supported by protein-DNA interactions. *Nanomedicine* 2011;**6**:1047–61.

251. Tsumoto K, Luckel F, Yoshikawa K. Giant DNA molecules exhibit on/off switching of transcriptional activity through conformational transition. *Biophys Chem* 2003;**106**:23–9.
252. Flenniken M, Willits D, Brumfield S, Young MJ, Douglas T. The small heat shock protein cage from Methanococcus jannaschii is a versatile nanoscale platform for genetic and chemical modification. *Nano Lett* 2003;**3**:1573–6.
253. Shively JM, Decker GL, Greenawalt JW. Comparative ultrastructure of the thiobacilli. *J Bacteriol* 1970;**101**:618–27.
254. Shively JM, Ball F, Brown DH, Saunders RE. Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus. *Science* 1973;**182**:584–6.
255. Price GD, Coleman JR, Badger MR. Association of carbonic anhydrase activity with carboxysomes isolated from the Cyanobacterium Synechococcus PCC7942. *Plant Physiol* 1992;**100**:784–93.
256. So AK, Espie GS, Williams EB, Shively JM, Heinhorst S, Cannon GC. A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. *J Bacteriol* 2004;**186**:623–30.
257. Yu JW, Price GD, Song L, Badger MR. Isolation of a putative carboxysomal carbonic anhydrase gene from the Cyanobacterium Synechococcus PCC7942. *Plant Physiol* 1992;**100**:794–800.
258. Chen P, Andersson DI, Roth JR. The control region of the pdu/cob regulon in Salmonella typhimurium. *J Bacteriol* 1994;**176**:5474–82.
259. Kofoid E, Rappleye C, Stojiljkovic I, Roth J. The 17-gene ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins. *J Bacteriol* 1999;**181**:5317–29.
260. Bobik TA, Havemann GD, Busch RJ, Williams DS, Aldrich HC. The propanediol utilization (pdu) operon of Salmonella enterica serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol degradation. *J Bacteriol* 1999;**181**:5967–75.
261. Havemann GD, Sampson EM, Bobik TA. PduA is a shell protein of polyhedral organelles involved in coenzyme B(12)-dependent degradation of 1,2-propanediol in Salmonella enterica serovar typhimurium LT2. *J Bacteriol* 2002;**184**:1253–61.
262. Havemann GD, Bobik TA. Protein content of polyhedral organelles involved in coenzyme B12-dependent degradation of 1,2-propanediol in Salmonella enterica serovar Typhimurium LT2. *J Bacteriol* 2003;**185**:5086–95.
263. Leal NA, Havemann GD, Bobik TA. PduP is a coenzyme-a-acylating propionaldehyde dehydrogenase associated with the polyhedral bodies involved in B12-dependent 1,2-propanediol degradation by Salmonella enterica serovar Typhimurium LT2. *Arch Microbiol* 2003;**180**:353–61.
264. Rondon MR, Horswill AR, Escalante-Semerena JC. DNA polymerase I function is required for the utilization of ethanolamine, 1,2-propanediol, and propionate by Salmonella typhimurium LT2. *J Bacteriol* 1995;**177**:7119–24.
265. Rondon MR, Kazmierczak R, Escalante-Semerena JC. Glutathione is required for maximal transcription of the cobalamin biosynthetic and 1,2-propanediol utilization (cob/pdu) regulon and for the catabolism of ethanolamine, 1,2-propanediol, and propionate in Salmonella typhimurium LT2. *J Bacteriol* 1995;**177**:5434–9.
266. Sampson EM, Bobik TA. Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. *J Bacteriol* 2008;**190**:2966–71.
267. Penrod JT, Roth JR. Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. *J Bacteriol* 2006;**188**:2865–74.

268. Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, Phillips M, Beeby M, et al. Protein structures forming the shell of primitive bacterial organelles. *Science* 2005;**309**:936–8.
269. Yeates TO, Crowley CS, Tanaka S. Bacterial microcompartment organelles: protein shell structure and evolution. *Annu Rev Biophys* 2010;**39**:185–205.
270. Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, Cannon GC, et al. Atomic-level models of the bacterial carboxysome shell. *Science* 2008;**319**:1083–6.
271. Tanaka S, Sawaya MR, Yeates TO. Structure and mechanisms of a protein-based organelle in Escherichia coli. *Science* 2010;**327**:81–4.
272. Cot SS, So AK, Espie GS. A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *J Bacteriol* 2008;**190**:936–45.
273. Long BM, Badger MR, Whitney SM, Price GD. Analysis of carboxysomes from Synechococcus PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *J Biol Chem* 2007;**282**:29323–35.
274. Price GD, Howitt SM, Harrison K, Badger MR. Analysis of a genomic DNA region from the cyanobacterium Synechococcus sp. strain PCC7942 involved in carboxysome assembly and function. *J Bacteriol* 1993;**175**:2871–9.
275. Parsons JB, Frank S, Bhella D, Liang M, Prentice MB, Mulvihill DP, et al. Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Mol Cell* 2010;**38**:305–15.
276. Menon BB, Dou Z, Heinhorst S, Shively JM, Cannon GC. Halothiobacillus neapolitanus carboxysomes sequester heterologous and chimeric RubisCO species. *PLoS One* 2008;**3**:e3570.
277. Fan C, Cheng S, Liu Y, Escobar CM, Crowley CS, Jefferson RE, et al. Short N-terminal sequences package proteins into bacterial microcompartments. *Proc Natl Acad Sci USA* 2010;**107**:7509–14.
278. Sutter M, Boehringer D, Gutmann S, Gunther S, Prangishvili D, Loessner MJ, et al. Structural basis of enzyme encapsulation into a bacterial nanocompartment. *Nat Struct Mol Biol* 2008;**15**:939–47.
279. Seebeck FP, Woycechowsky KJ, Zhuang W, Rabe JP, Hilvert D. A simple tagging system for protein encapsulation. *J Am Chem Soc* 2006;**128**:4516–7.
280. Zhang X, Konarev PV, Petoukhov MV, Svergun DI, Xing L, Cheng RH, et al. Multiple assembly states of lumazine synthase: a model relating catalytic function and molecular assembly. *J Mol Biol* 2006;**362**:753–70.
281. Dalmau M, Lim S, Chen HC, Ruiz C, Wang SW. Thermostability and molecular encapsulation within an engineered caged protein scaffold. *Biotechnol Bioeng* 2008;**101**:654–64.
282. Dalmau M, Lim S, Wang SW. pH-triggered disassembly in a caged protein complex. *Biomacromolecules* 2009;**10**:3199–206.
283. Grinberg O, Hayun M, Sredni B, Gedanken A. Characterization and activity of sonochemically-prepared BSA microspheres containing Taxol–an anticancer drug. *Ultrason Sonochem* 2007;**14**:661–6.
284. Rivkin I, Cohen K, Koffler J, Melikhov D, Peer D, Margalit R. Paclitaxel-clusters coated with hyaluronan as selective tumor-targeted nanovectors. *Biomaterials* 2010;**31**:7106–14.
285. Vazquez E, Roldan M, Diez-Gil C, Unzueta U, Domingo-Espin J, Cedano J, et al. Protein nanodisk assembling and intracellular trafficking powered by an arginine-rich (R9) peptide. *Nanomedicine* 2010;**5**:259–68.
286. Diez-Gil C, Krabbenborg S, Garcia-Fruitos E, Vazquez E, Rodriguez-Carmona E, Ratera I, et al. The nanoscale properties of bacterial inclusion bodies and their effect on mammalian cell proliferation. *Biomaterials* 2010;**31**:5805–12.
287. Garcia-Fruitos E, Rodriguez-Carmona E, Diez-Gil C, Ferraz RM, Vazquez E, Corchero JL, et al. Surface cell growth engineering assisted by a novel bacterial nanomaterial. *Adv Mater* 2009;**21**:4249.

288. Rodriguez-Carmona E, Villaverde A. Nanostructured bacterial materials for innovative medicines. *Trends Microbiol* 2010;**18**:423–30.
289. Vaysse L, Harbottle R, Bigger B, Bergau A, Tolmachov O, Coutelle C. Development of a self-assembling nuclear targeting vector system based on the tetracycline repressor protein. *J Biol Chem* 2004;**279**:5555–64.
290. Fukunaga-Kalabis M, Herlyn M. Unraveling mysteries of the multifunctional protein SPARC. *J Invest Dermatol* 2007;**127**:2497–8.
291. Akahata W, Yang ZY, Andersen H, Sun S, Holdaway HA, Kong WP, et al. A virus-like particle vaccine for epidemic Chikungunya virus protects nonhuman primates against infection. *Nat Med* 2010;**16**:334–8.
292. Pushko P, Kort T, Nathan M, Pearce MB, Smith G, Tumpey TM. Recombinant H1N1 virus-like particle vaccine elicits protective immunity in ferrets against the 2009 pandemic H1N1 influenza virus. *Vaccine* 2010;**28**:4771–6.
293. Harmsen MM, De Haard HJ. Properties, production, and applications of camelid single-domain antibody fragments. *Appl Microbiol Biotechnol* 2007;**77**:13–22.
294. Cortez-Retamozo V, Backmann N, Senter PD, Wernery U, De BP, Muyldermans S, et al. Efficient cancer therapy with a nanobody-based conjugate. *Cancer Res* 2004;**64**:2853–7.
295. Bian J, Popovic ZB, Benejam C, Kiedrowski M, Rodriguez LL, Penn MS. Effect of cell-based intercellular delivery of transcription factor GATA4 on ischemic cardiomyopathy. *Circ Res* 2007;**100**:1626–33.
296. Fletcher S, Honeyman K, Fall AM, Harding PL, Johnsen RD, Steinhaus JP, et al. Morpholino oligomer-mediated exon skipping averts the onset of dystrophic pathology in the mdx mouse. *Mol Ther* 2007;**15**:1587–92.
297. Dubikovskaya EA, Thorne SH, Pillow TH, Contag CH, Wender PA. Overcoming multidrug resistance of small-molecule therapeutics through conjugation with releasable octaarginine transporters. *Proc Natl Acad Sci USA* 2008;**105**:12128–33.
298. Bucci M, Gratton JP, Rudic RD, Acevedo L, Roviezzo F, Cirino G, et al. In vivo delivery of the caveolin-1 scaffolding domain inhibits nitric oxide synthesis and reduces inflammation. *Nat Med* 2000;**6**:1362–7.
299. Morris MC, Gros E, drian-Herrada G, Choob M, Archdeacon J, Heitz F, et al. A non-covalent peptide-based carrier for in vivo delivery of DNA mimics. *Nucleic Acids Res* 2007;**35**:e49.
300. Crombez L, Morris MC, Dufort S, drian-Herrada G, Nguyen Q, Mc MG, et al. Targeting cyclin B1 through peptide-based delivery of siRNA prevents tumour growth. *Nucleic Acids Res* 2009;**37**:4559–69.
301. Song E, Zhu P, Lee SK, Chowdhury D, Kussman S, Dykxhoorn DM, et al. Antibody mediated in vivo delivery of small interfering RNAs via cell-surface receptors. *Nat Biotechnol* 2005;**23**:709–17.
302. Tan M, Lan KH, Yao J, Lu CH, Sun M, Neal CL, et al. Selective inhibition of ErbB2-overexpressing breast cancer in vivo by a novel TAT-based ErbB2-targeting signal transducers and activators of transcription 3-blocking peptide. *Cancer Res* 2006;**66**:3764–72.
303. Myrberg H, Zhang L, Mae M, Langel U. Design of a tumor-homing cell-penetrating peptide. *Bioconj Chem* 2008;**19**:70–5.
304. Laakkonen P, Porkka K, Hoffman JA, Ruoslahti E. A tumor-homing peptide with a targeting specificity related to lymphatic vessels. *Nat Med* 2002;**8**:751–5.
305. Jain M, Chauhan SC, Singh AP, Venkatraman G, Colcher D, Batra SK. Penetratin improves tumor retention of single-chain antibodies: a novel step toward optimization of radioimmunotherapy of solid tumors. *Cancer Res* 2005;**65**:7840–6.
306. Vocero-Akbani AM, Heyden NV, Lissy NA, Ratner L, Dowdy SF. Killing HIV-infected cells by transduction with an HIV protease-activated caspase-3 protein. *Nat Med* 1999;**5**:29–33.

307. Bidwell III GL, Fokt I, Priebe W, Raucher D. Development of elastin-like polypeptide for thermally targeted delivery of doxorubicin. *Biochem Pharmacol* 2007;**73**:620–31.
308. Snyder EL, Meade BR, Saenz CC, Dowdy SF. Treatment of terminal peritoneal carcinomatosis by a transducible p53-activating peptide. *PLoS Biol* 2004;**2**:E36.
309. Harada H, Hiraoka M, Kizaka-Kondoh S. Antitumor effect of TAT-oxygen-dependent degradation-caspase-3 fusion protein specifically stabilized and activated in hypoxic tumor cells. *Cancer Res* 2002;**62**:2013–8.
310. Datta K, Sundberg C, Karumanchi SA, Mukhopadhyay D. The 104–123 amino acid sequence of the beta-domain of von Hippel-Lindau gene product is sufficient to inhibit renal tumor growth and invasion. *Cancer Res* 2001;**61**:1768–75.
311. Shibagaki N, Udey MC. Dendritic cells transduced with protein antigens induce cytotoxic lymphocytes and elicit antitumor immunity. *J Immunol* 2002;**168**:2393–401.
312. Bright R, Raval AP, Dembner JM, Perez-Pinzon MA, Steinberg GK, Yenari MA, et al. Protein kinase C delta mediates cerebral reperfusion injury in vivo. *J Neurosci* 2004;**24**:6880–8.
313. Cao G, Pei W, Ge H, Liang Q, Luo Y, Sharp FR, et al. In vivo delivery of a Bcl-xL fusion protein containing the TAT protein transduction domain protects against ischemic brain injury and neuronal apoptosis. *J Neurosci* 2002;**22**:5423–31.
314. Kumar P, Ban HS, Kim SS, Wu H, Pearson T, Greiner DL, et al. T cell-specific siRNA delivery suppresses HIV-1 infection in humanized mice. *Cell* 2008;**134**:577–86.
315. Kim WJ, Christensen LV, Jo S, Yockman JW, Jeong JH, Kim YH, et al. Cholesteryl oligoarginine delivering vascular endothelial growth factor siRNA effectively inhibits tumor growth in colon adenocarcinoma. *Mol Ther* 2006;**14**:343–50.
316. Mamot C, Drummond DC, Noble CO, Kallab V, Guo Z, Hong K, et al. Epidermal growth factor receptor-targeted immunoliposomes significantly enhance the efficacy of multiple anticancer drugs in vivo. *Cancer Res* 2005;**65**:11631–8.
317. Arap W, Pasqualini R, Ruoslahti E. Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. *Science* 1998;**279**:377–80.
318. Liu S, Wang H, Currie BM, Molinolo A, Leung HJ, Moayeri M, et al. Matrix metalloproteinase-activated anthrax lethal toxin demonstrates high potency in targeting tumor vasculature. *J Biol Chem* 2008;**283**:529–40.
319. Frankel AE, Powell BL, Duesbery NS, Vande Woude GF, Leppla SH. Anthrax fusion protein therapy of cancer. *Curr Protein Pept Sci* 2002;**3**:399–407.
320. Hu P, Yan J, Sharifi J, Bai T, Khawli LA, Epstein AL. Comparison of three different targeted tissue factor fusion proteins for inducing tumor vessel thrombosis. *Cancer Res* 2003;**63**:5046–53.
321. Lindgren M, Rosenthal-Aizman K, Saar K, Eiriksdottir E, Jiang Y, Sassian M, et al. Overcoming methotrexate resistance in breast cancer tumour cells by the use of a new cell-penetrating peptide. *Biochem Pharmacol* 2006;**71**:416–25.
322. Anderson DC, Nichols E, Manger R, Woodle D, Barry M, Fritzberg AR. Tumor cell retention of antibody Fab fragments is enhanced by an attached HIV TAT protein-derived peptide. *Biochem Biophys Res Commun* 1993;**194**:876–84.
323. Dijkgraaf I, Kruijtzer JA, Liu S, Soede AC, Oyen WJ, Corstens FH, et al. Improved targeting of the alpha(v)beta (3) integrin by multimerisation of RGD peptides. *Eur J Nucl Med Mol Imaging* 2007;**34**:267–73.
324. Kawamura KS, Sung M, Bolewska-Pedyczak E, Gariepy J. Probing the impact of valency on the routing of arginine-rich peptides into eukaryotic cells. *Biochemistry* 2006;**45**:1116–27.
325. Vives E, Schmidt J, Pelegrin A. Cell-penetrating and cell-targeting peptides in drug delivery. *Biochim Biophys Acta* 2008;**1786**:126–38.

326. Jiang T, Olson ES, Nguyen QT, Roy M, Jennings PA, Tsien RY. Tumor imaging by means of proteolytic activation of cell-penetrating peptides. *Proc Natl Acad Sci USA* 2004;**101**:17867–72.
327. Sethuraman VA, Bae YH. TAT peptide-based micelle system for potential active targeting of anti-cancer agents to acidic solid tumors. *J Control Release* 2007;**118**:216–24.
328. Lebleu B, Moulton HM, Abes R, Ivanova GD, Abes S, Stein DA, et al. Cell penetrating peptide conjugates of steric block oligonucleotides. *Adv Drug Deliv Rev* 2008;**60**:517–29.
329. Kamei N, Morishita M, Eda Y, Ida N, Nishio R, Takayama K. Usefulness of cell-penetrating peptides to improve intestinal insulin absorption. *J Control Release* 2008;**132**:21–5.
330. Pardridge WM. Biopharmaceutical drug targeting to the brain. *J Drug Target* 2010;**18**:157–67.
331. Perales JC, Ferkol T, Beegen H, Ratnoff OD, Hanson RW. Gene transfer in vivo: sustained expression and regulation of genes introduced into the liver by receptor-targeted uptake. *Proc Natl Acad Sci USA* 1994;**91**:4086–90.
332. Sato Y, Yamauchi N, Takahashi M, Sasaki K, Fukaura J, Neda H, et al. In vivo gene delivery to tumor cells by transferrin-streptavidin-DNA conjugate. *FASEB J* 2000;**14**:2108–18.

Article 3

Protein nanodisk assembling and intracellular trafficking powered by an arginine-rich (R9) peptide

Vazquez, E., Roldan, M., Diez-Gil, C., Unzueta, U., Domingo-Espin, J., Cedano, J., Conchillo, O., Ratera, I., Veciana, J., Daura, X., Ferrer-Miralles, N., & Villaverde, A.

Nanomedicine 5, 259-268 (2010).

In this paper we analyse the use of particulate organization and biological performance in DNA delivery of a model protein-only delivery vehicle, composed by nine arginine (R9) peptide fused to C-terminally his-tagged green fluorescent protein by different techniques. This protein self-assembles through protein-protein interactions in a very interesting nanoform, a nanodisk of 20 × 3 nm, which is able to be internalized efficiently in mammalian cells and rapidly migrate through the cytoplasm towards the nucleus in a fully bioactive form.
The architectonic properties of arginine-rich peptides at the nanoscale reveal a new category of protein nanoparticles, namely nanodisks, and provide novel strategic concepts and architectonic tools for the tailored construction of new-generation artificial viruses for gene therapy and drug delivery.

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

Protein nanodisk assembling and intracellular trafficking powered by an arginine-rich (R9) peptide

Aims: Arginine(R)-rich cationic peptides are powerful tools in drug delivery since, alone or when associated with polyplexes, proteins or chemicals, they confer DNA condensation, membrane translocation and blood–brain barrier crossing abilities. The unusual stability and high *in vivo* performance of their associated drugs suggest a particulate organization or R(n) complexes, which this study aimed to explore. **Materials & methods:** We have analyzed the particulate organization and biological performance in DNA delivery of a model, R9-containing green fluorescent protein by dynamic light scattering, transmission electron microscopy, atomic force microscopy, single cell confocal microscopy and flow cytometry. **Results:** A deep nanoscale examination of R9-powered constructs reveals a novel and promising feature of R9, that when fused to a scaffold green fluorescent protein, promote its efficient self-assembling as highly stable, regular disk-shaped nanoparticles of 20 × 3 nm. These constructs are efficiently internalized in mammalian cells and rapidly migrate through the cytoplasm towards the nucleus in a fully bioactive form. Besides, such particulate platforms accommodate, condense and deliver plasmid DNA to the nucleus and promote plasmid-driven transgene expression. **Conclusion:** The architectonic properties of arginine-rich peptides at the nanoscale reveal a new category of protein nanoparticles, namely nanodisks, and provide novel strategic concepts and architectonic tools for the tailored construction of new-generation artificial viruses for gene therapy and drug delivery.

KEYWORDS: artificial viruses ■ cationic peptides ■ gene therapy ■ nanoparticles ■ self-assembling ■ trafficking

Esther Vazquez, Mónica Roldán, César Diez-Gil, Ugutz Unzueta, Joan Domingo-Espín, Juan Cedano, Oscar Conchillo, Imma Ratera, Jaume Veciana, Xavier Daura, Neus Ferrer-Miralles & Antonio Villaverde†

†Author for correspondence: Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain avillaverde@servet.uab.es

**For a full list of affiliations please see back page*

In drug-based therapies, most bioactive molecules need to overcome physical and biological barriers (generally cellular and nuclear membranes) to reach their molecular targets. Viruses have naturally evolved to acquire membrane-crossing activities that have been largely exploited as DNA (or RNA) nanocarriers for gene therapy (or gene silencing). However, the spectrum of undesired side effects observed during the controlled administration of such pathogenic agents [1], even upon genetic attenuation, prevents them from being considered as generic therapeutic agents and has severely compromised the further development of viral gene therapy [2,3]. Alternatively, 'artificial viruses' [4,5] are bio-safer nanosized constructs with orderly assembled components that mimic viral properties of relevance in drug delivery [6]. These molecular containers are formed by either synthetic (lipids or polysaccharides) or biologically produced (proteins) building blocks, or by a combination of both. Among them, protein-only nanoparticles are fully biocompatible and suitable for rational protein design through the fine tuning of their biophysical and biological properties, including those regulating the architectural features (e.g., self-assembly and drug permeability) [7] or their biological interaction with the environment, namely receptor recognition, membrane crossing and nuclear targeting, among others [6].

For artificial viruses to achieve their targets they are usually functionalized with peptides or protein domains showing different biological activities, including, receptor binding, cell internalization, DNA condensation, endosomal escape, intracellular trafficking and nuclear transport [6]. Extended catalogues of peptides and small protein domains have so far been explored as functionalizing agents to fully cover the requirements of nanoparticle-driven delivery of nucleic acids, or eventually other drugs [8–14]. Among them, highly cationic proteins such as the HIV Tat, its 48–60 amino acid segment or other arginine-rich peptides have been shown to be highly bioactive and extremely promising as DNA and protein deliverers [14]. Polyarginine peptides of between 4 and 16 amino acids efficiently translocate the cell membrane, promoting the cellular uptake of associated molecules [15]. Irrespective of the membrane-crossing mechanisms that seem to be diverse under different experimental

part of future medicine fsg

settings (endocytic versus nonendocytic translocation pathways) [13,16–18], DNA associated with synthetic R(n) peptides enters the nuclear compartment and promotes detectable levels of transgene expression [19,20]. Since the optimal length for cell internalization seems to be approximately eight residues [15,20], related species of the polyarginine peptide family (mainly R9, and to a lesser extent R6, R7 and R10) have been explored for the *in vitro* and *in vivo* delivery of p53 [21], p53-derived retro-inverso peptides [22], cyclosporine A [23] and peptidic nucleic acids [24]. Interestingly, R9-activated proteins cross the blood–brain barrier (BBB) [25], a highly promising ability for biomedical applications. Furthermore, in natural viruses, arginine-rich motifs also play essential roles in the encapsidation of viral genomes into protein shells [26].

Materials & methods

Protein production & purification

Green fluorescent protein (GFP) fusion proteins were produced in Rosetta BL21 (DE3) *Escherichia coli* as driven by an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible-T7 promoter in an pET21B$^+$-derived plasmid. Bacteria were grown in luria bertani medium (750 ml) at 37 °C in shaker flasks until an optical density of 0.5, and gene expression was induced by 1 mM IPTG. After 3 h, cells were harvested by centrifugation (7650 g for 10 min at 4°C), washed in phosphate buffered saline (PBS) and stored at -80°C until use. The pellet was resuspended in buffer A (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol) and cells disrupted by sonication in the presence of a tablet of EDTA-free protease inhibitor cocktail (Complete, 11873580001 from Roche). The soluble cell fraction was separated by centrifugation at 14,841 g for 15 min at 4°C. Upon filtration through 0.22 µm filters, GFP fusions were purified by chromatography in Ni^{2+} columns in an ÄKTA^{TM} Purifier (GE healthcare) fast protein liquid chromatography. Positive fractions in elution buffer (Tris-HCl 20 mM pH 7.5, 150 mM NaCl, 500 mM Imidazole, 5 mM β-mercaptoethanol) were collected, dialyzed against the desired buffer and quantified by Bradford's procedure. A standard HBS buffer at pH 5.8 [27] was used for most of the experiments (buffer 5 in Figure 1E). When indicated in some experiments, proteins were dissolved in 10 mM Tris HCl buffer pH 7.5 + 0.01% Tween 20 (buffer 2 in Figure 1E). The other buffers used for the experiment summarized in Figure 1E were as follows:

- 20 mM Tris HCl pH 7.5 + 5% dextrose;
- PBS 7.4 + 10% glycerol;
- 20 mM Tris HCl pH 7.5 + 5% dextrose, 200 mM NaCl;
- 10 mM Tris HCl pH 7.5 200 mM NaCl + 0.01% Tween 20.

Wild-type GFP was supplied by Nucliber (ref 632373).

Dynamic light scattering, transmission-electron microscopy & DNA retardation assays

Volume-size distributions of engineered GFP proteins and resulting complexes were measured using a dynamic light scattering analyzer at the wavelength of 633 nm, combined with noninvasive backscatter technology (Zetasizer Nano ZS, Malvern Instruments Limited, Malvern, UK). DNA–protein incubation, transmission-electron microscopy (TEM) of protein and DNA–protein complexes and DNA mobility assays were performed according to previously published protocols [27].

Atomic force microscopy

Atomic force microscopy (AFM) analyses were performed in air with a commercial atomic force microscope (PicoScan/PicoSPM 5500SL from Molecular Imaging Agilent Technologies, Inc., Santa Clara, CA, USA) operating in acoustic mode. 9R-GFP-H6 proteins in 10 mM Tris HCl pH 7.5 buffer were dropcasted onto a freshly cleaved mica surface and air dried before measuring. For the acoustic mode measurements, a monolithic supersharp silicon SSS-NCH-50 (Nanosensors, Inc.) tip, with a radius of 2 nm, a nominal spring constant of 10–130 N/m and a resonance frequency of 204–497 kHz was used.

Cell culture & confocal laser scanning microscopy

The HeLa (ATCC-CCL-2) cell line was used in all the experiments and monitored *in vivo* in absence of fixation, to prevent previously described internalization artifacts [28]. Cells were maintained in MEM (GIBCO, Rockville, MD, USA) supplemented with 10% fetal calf serum (GIBCO) and incubated at 37°C and 5% CO_2 in a humidified atmosphere. For confocal analysis, cells were grown on Mat-Teck culture dishes (Mat Teck Corp., Ashland, MA, USA). Nuclei were labelled with 20 µg/ml Hoechst 33342 (Molecular Probes, Eugene, OR, USA) and the plasma membrane was labelled with

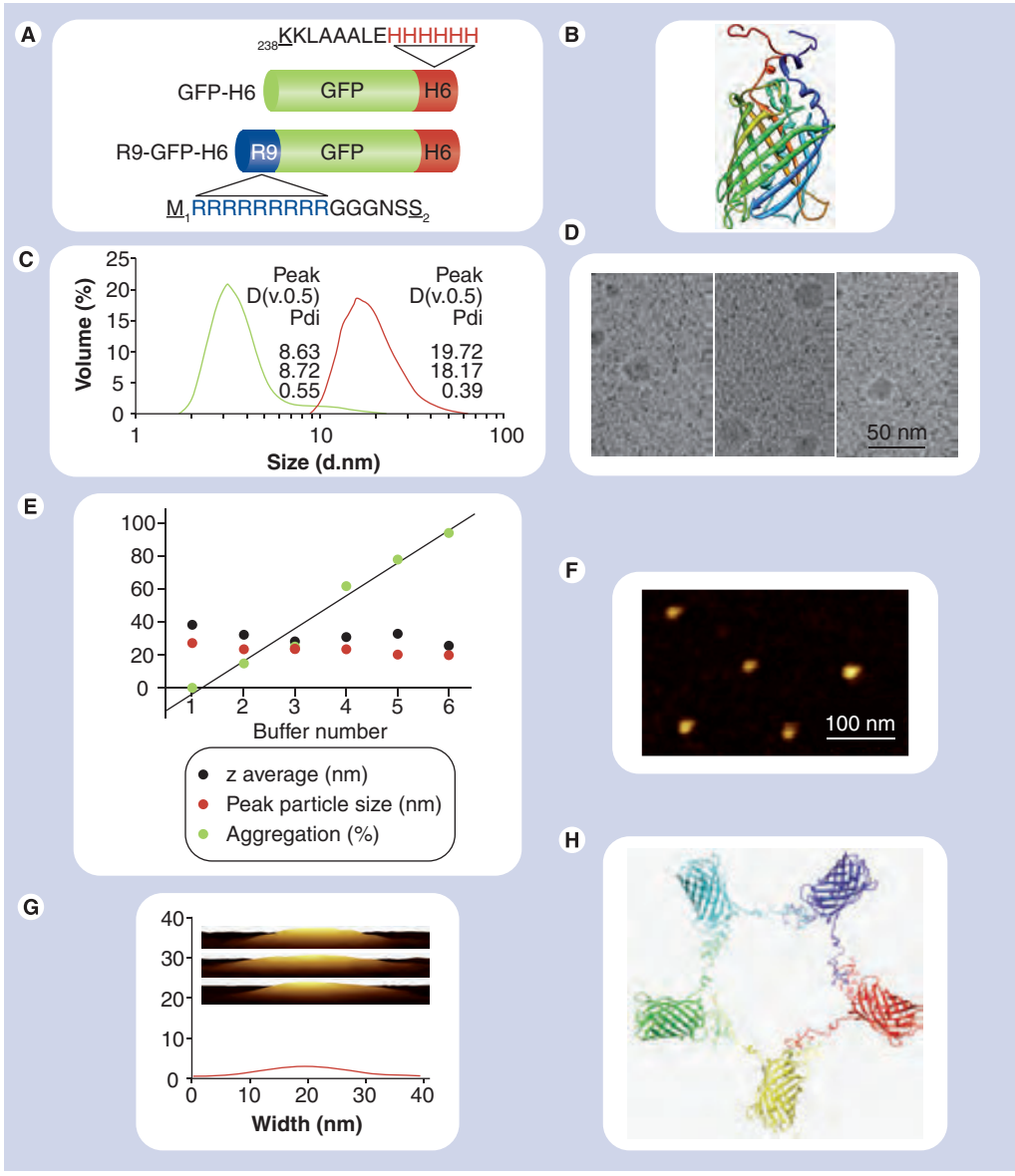

Figure 1. Description of R9-GFP-H6 modular protein and nanoscale characterization of R9-powered R9-GFP-H6 nanodisks. The amino acid sequences of polyarginine and polyhistidine tails are indicated in blue and red, respectively. R9 is accommodated between amino terminal residues M1 and S2 of GFP (underlined), while H6 is fused to the carboxy terminal residue of GFP (K238 underlined). Amino acids typed in bold face are those resulting from the cloning process. **(B)** Molecular modeling of R9-GFP-H6 showing one among the most probable spatial orientations of R9 (blue) and H6 (red) peptides. **(C)** Size distribution of soluble GFP-H6 (green plot) and R9-GFP-H6 (red plot) in 20 mM Tris buffer pH 7.5 + 5% dextrose determined by dynamic light scattering, the inset figures indicating (in nm) the peak size, the particle diameter under which there are 50% of the total volume of the population D[v,0.5] and the pdi. **(D)** TEM images of randomly R9-GFP-H6 nanoparticles in the same Tris buffer. **(E)** Percentage of R9-GFP-H6 aggregation, particle size peak and z-average values of soluble R9-GFP-H6 nanoparticles in different buffers with the following composition: (1) 20 mM Tris HCl pH 7.5 + 5% dextrose; (2) 10 mM Tris HCl pH 7.5 + 0.01% Tween 20; (3) phosphate buffered saline 7.4 + 10% glycerol; (4) 20 mM Tris HCl pH 7.5 200 mM NaCl + 5 % dextrose; (5) HBS pH 5.8; (6) 10 mM Tris HCl pH 7.5 200 mM NaCl + 0.01% Tween 20. **(F)** AFM analysis of randomly selected R9-GFP-H6 nanoparticles deposited on a mica surface. **(G)** Topography cross-section of an isolated particle. In the inset, 3D views of three randomly selected nanoparticles. **(H)** Molecular modeling of R9-GFP-H6 assemblies of sizes compatible with dynamic light scattering, TEM and AFM determinations. This star-shaped distribution pattern could admit up to ten individual R9-GFP-H6 molecules with minor alterations of particle size.
AFM: Atomic force microscopy; GFP: Green fluorescent protein; pdi: Polydispersion index.

2.5 µg/ml CellMask™ Deep Red (Molecular Probes, Invitrogen, Carlsbad, CA, USA) for 5 min in the dark. Cells were washed in PBS (Sigma-Aldrich Chemie GmbH, Steinheim, Germany). Live cells were recorded with a TCS-SP5 confocal laser scanning microscope (Leica Microsystems, Heidelberg, Germany) using a Plan Apo 63×/1.4 (oil HC × PL APO λ blue) objective. Hoechst 33342 DNA labels was excited with a blue diode (405 nm) and detected in the 415–460 nm range. GFP proteins were excited with a Ar laser (488 nm) and detected in the 525–545 nm range. CellMask was excited with a HeNe laser (633 nm) and detected in the 650–775 nm range. To determine the protein localization inside the nucleus, stacks of 20–30 sections every 0.5 µm along the cell thickness were collected at intervals of 15 min over approximately 12 h. The projections of the series obtained were generated with Leica LAS AF software, and 3D models were generated using Imaris v. 6.1.0 software (Bitplane, Zürich, Switzerland). Profile analysis of fluorescence intensity was measured using Leica LAS AF software to determine the fluorescence intensity along the line segment in relation to the wavelength. The fluorescence intensity profiles were measured at the same laser excitation and photomultiplier gain settings from the cells.

Protein-mediated plasmid transfection

For gene expression experiments, 50 µg of 9R-GFP-H6 and 1 µg of TdTomato expression vector were mixed into 50 µl of HBS pH 5.8 and incubated for 1 h at room temperature. Then, a convenient amount of Optimem was added and the full mixture was gently added to cultured HeLa cells that were further incubated for 4 h at 37°C in 5% CO_2 atmosphere. The cultures were then transferred to complete media for growth. After 48 h, TdTomato gene expression was monitored by flow cytometry. Nontreated cells or cells just exposed to the plasmid expression vector or the protein alone were used as controls.

Flow cytometry

Cell samples were analyzed after treatment with 0.5 mg/ml trypsin, 4Na in HBSS for 10 min on a FACSCanto system (Becton Dickinson), using a 15 W air-cooled argon-ion laser at 488 nm excitation. Fluorescence emission was measured with detector D (530/30 nm band pass filter) for EGFP and detector C (585/42 nm band pass filter) for TdTomato fluorescent protein.

Molecular protein modeling

To build the R9-GFP-H6 nanodisk model, a model of the monomer was first generated using Modeller 9v2 [29] and the protein data bank structure '1qyo' as a template. The arginine and histidine tails were modelled using the loop-model function of this package. The structural model of the complex was then created with HADDOCK 2.0 [30], defining the nine arginines from the N-terminal tail as active residues and the six histidines from the C-terminal tail as passive residues and inforcing C5 symmetry.

Results

Although never explored, the unusually high performance of arginine-rich peptides in functionalizing drugs and proteins under different conditions and experimental models, combined with the high stability of R9-based drugs in local and systemic delivery, made us wonder whether R9 itself might promote some form of spontaneous supramolecular organization to its associated molecules, causing the whole conjugate to act as a particulate material. To evaluate this possibility, we explored the potential self-assembly of a His-tagged (GFP-H6) GFP upon functionalization with an R9 peptide fused to its amino terminus (Figure 1A). In the final R9-GFP-H6 construct, both end terminal peptides resulted exposed to the solvent in the same pole of the GFP barrel (Figure 1B). The parental GFP-H6 and the engineered R9-GFP-H6 constructs were efficiently purified in a single-step affinity chromatograpy from the extracts of producing *E. coli* bacterial cells. Interestingly, GFP-H6 appeared in solution as particles of around 7 nm in diameter (Figure 1C), which is compatible with the molecular size of individual GFP molecules (3 × 5 nm), plus the hanging H6 tail. Surprisingly, the addition of the R9 tail promoted the spontaneous self-assembling of R9-GFP-H6 as a population of particulate species of around 20 nm, as determined by dynamic light scattering (Figure 1C). Interestingly, TEM examination of R9-GFP-H6 showed the protein as a particulate material of defined dimensions and regular round morphologies compatible with a spherical architecture, of sizes coincident with dynamic light scattering data (Figure 1D). The regularity in size and morphology indicated that these nanoparticles possessed an important level of inner organization. In agreement, such assembles regularly occur as soluble entities in different buffers that differentially affect protein solubility (Figure 1E). This fact demonstrates that R9-driven particle formation results from

an ordered molecular assembling process rather than from an unspecific aggregation rendering insoluble protein clusters. AFM studies were performed over a mica substrate using a supersharp silicon 2-nm radius tip (Nanosensors, Inc.), in order to improve the image resolution. AFM analysis revealed that R9-GFP-H6 nanoparticles were not spherical but that they were instead organized as flattered spheres or nanodisks with dimensions of around 20 × 3 nm (Figure 1F & G). A molecular-structure model of the R9-GFP-H6 nanodisks was generated using the size and shape observed in the AFM images as constraints. The obtained solution for the structure of the complex predicted a star-shaped arrangement of five (or more) individual R9-GFP-H6 molecules (Figure 1H), in which the H6 and R9 terminal tails interact in the central region through histidine–arginine pairings [31]. This model is sufficiently flexible to explain the slight nanodisk-size variation shown in Figure 1C, and it could admit up to approximately ten individual R9-GFP-H6 molecules as building blocks of a single nanoparticle, that could still be compatible with the observed dimensions. In addition, this is an arrangement that explains the absence of larger aggregates.

It is noteworthy that the addition of R9 to GFP-H6 increased the hydrophilic nature of the protein (grand average of hydropathicity [GRAVY] of GFP is -0.557, and those of GFP-H6 and R9-GFP-H6 -0.617 and -0.725, respectively) and therefore its amphiphilicity (due to the polar situation of R9 in the GFP barrel, Figure 1B). However, the self assembling of R9-GFP-H6 seemed to be determined by R9-driven cross-interactions (involving H6 residues) rather than by a conventional solvent exposure of the highly hydrophilic (R9 and H6) regions of the building blocks [32]. In this regard, R9 could architectonically act as multifunctional-adhesive-disordered tail. Interestingly, arginine is enriched in the proximity of protein–protein interfaces [33], due to its tendency to cross-associate in stacking-like arrangements [34] or in hydrogen-bonded 'hubs' [33].

At this stage, we wondered if these disk-shaped entities could penetrate mammalian cells keeping both the particulate organization and GFP fluorescence. Noteworthy, GFP fluorescent emission acts as a convenient reporter of the molecular integrity of individual building blocks, as it depends on the proper barrel formation and fluorophore maturation that occurs at late folding steps [35]. As expected, wild type GFP remained fully dispersed in the extracellular media upon exposure to cultured HeLa cells, the cell membrane acted as an efficient barrier for the protein (Figure 2A), and no green signal was observed either in the cell cytoplasm or nucleus after 90 min of exposure (Figure 2B). Similarly, GFP-H6 remained fully excluded from cultured cells (Figure 2C & D), and the fluorescence of both GFP and GFP-H6 was fully dispersed in the extracellular media. However, R9-GFP-H6 nanoparticles were efficiently internalized and discrete fluorescent dots were observed in the cell cytoplasm, which progressively accumulates into the nucleus (Figure 2E & F). The nuclear targeting of R9-GFP-H6 was confirmed by single-cell image analysis (Figure 2G). Furthermore, the cytoplasmic trafficking and nuclear avidity of R9-GFP-H6 nanodisks was illustrated by the temporal tracking of these particles entering individual cells in volumetric reconstructions (Figure 2H), and confirmed in the absence of Hoechst 33342 and CellMask staining to avoid any enhancement of cell permeability eventually induced by these dyes (not shown). Interestingly, R9-GFP-H6 remained highly fluorescent during migration and inside both cell compartments (Figure 2H), being indicative of both protein and particle stability. The 3D imaging of the nanoparticles crossing the nuclear membrane (Figure 2G) could be compatible with the passage through the nuclear pores. Although 50 kDa seems to be a general limit for the free molecular passage across those pores (pentameric R9-GFP-H6 disks are approximately 150 kDa), the facilitated diffusion mechanism appears not having size restriction [36]. The inner diameter of nuclear pores ranges around 50–60 nm [37], and they can accommodate and translocate viral particles of 39 nm in diameter [38], larger than R9-driven nanodisks. In any case, the precise mechanism of nuclear entry as nanoparticles required further investigation, as it has been previously observed that synthetic R(n) peptides accumulate in the nucleus, the translocation mechanism remains unexplored [18].

The nuclear targeting of R9-GFP-H6 prompted to further explore the ability of the nanodisks to release expressible DNA in the nuclear compartment. Although isolated R(9–15) peptides are known to deliver expressible DNA [20], protein based, R9-powered molecular assemblies had never been explored in this regard. The parental GFP-H6 protein, despite the cationic nature provided by the H6 tail, was unable to bind DNA, as revealed by the failing of this construct in altering plasmid DNA mobility in gel electrophoresis (Figure 3A). By contrast, R9-GFP-H6 fully impeded DNA mobility at a mass protein–DNA ratio of 20 (this ratio consequently defined as one retardation unit [39]), this fact expectedly resulting from the

interaction between R9 and DNA [9]. Dynamic light scattering and TEM analysis of R9-GFP-H6 upon incubation with different plasmid amounts indicated that the size of nanoparticles was not visibly affected by DNA binding (**Figure 3B & C**), suggesting that the plasmid was efficiently condensed without dramatic rearrangements of the protein shell. In this regard, supercoiled plasmid DNA organize as minitoroids and spheroids [40], with sizes within the size range order than those of nanodisks. Furthermore, upon exposure to DNA–R9-GFP-H6 complexes, an important fraction of HeLa cell population showed red fluorescence (resulting from the production of the reporter fluorescent protein Td Tomato), as monitored by flow cytometry (Figure 3D). Such cell fraction was simultaneously emitting green fluorescence, derived from the biologically active R9-GFP-H6 that carried the DNA. The observed transgene expression confirms that R9-based nanodisks, despite being highly stable in the extracellular media and during cytosolic trafficking, are able to release expressible DNA once the DNA–protein complexes have reached the nucleus. No appreciable differences in cell internalization of R9-GFP-H6 were observed when administered alone or as associated to plasmid DNA (not shown).

Discussion

Apart from the well described properties of polyarginine peptides as membrane translocators and cross-BBB carriers, the data presented here demonstrate their architectonic potential at the nanoscale through the fusion of an R9 tail to a scaffold GFP. This peptide triggers the spontaneous self-assembling of the chimerical protein as regular, highly ordered disk-shaped structures of 20 × 3 nm, probably formed by five or more building blocks, with a relatively low polydispersion index (Figure 1). These nanodisks are highly soluble and stable in a diversity of

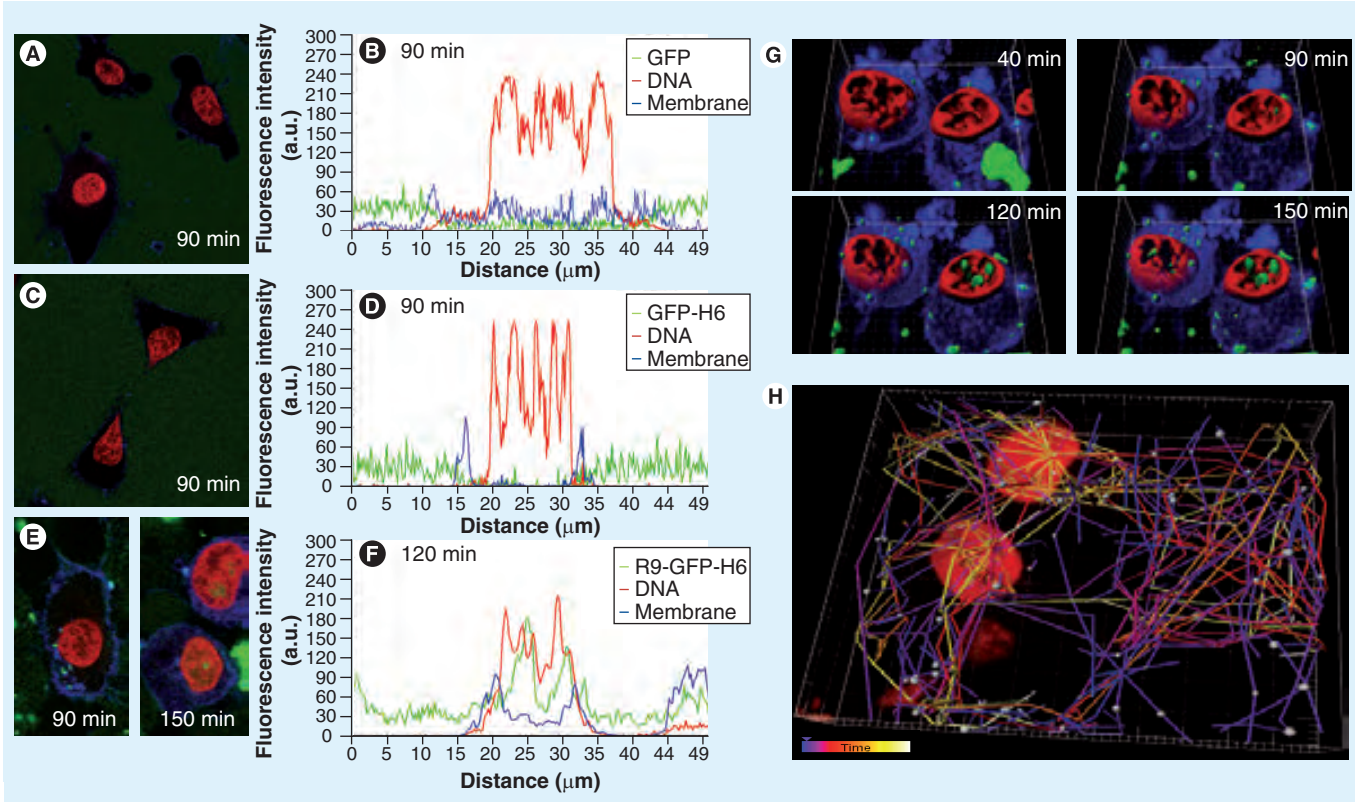

Figure 2. Confocal analysis of wild-type GFP, GFP-H6 and R9-GFP-H6 protein location in HeLa cells. Colour allocation: the cell membrane was labelled with CellMask (blue signal), cell DNA was labelled with Hoechst 33342 (red signal) and wild-type GFP, GFP-H6 and R9-GFP-H6 proteins produced a green signal). **(A)** Cultured cells exposed to wild-type GFP. **(B)** Fluorescence intensity profiles along a line segment of a randomly selected cell exposed to wild-type GFP, showing GFP, membrane cell and DNA signals. **(C)** Cultured cells exposed to GFP-H6. **(D)** Fluorescence intensity profiles along a line segment of a randomly selected cell exposed to GFP-H6. **(E)** Cultured cells exposed to R9-GFP-H6. **(F)** Fluorescence intensity profiles along a line segment of a randomly selected cell exposed to R9-GFP-H6. **(G)** Isosurface representation of HeLa cells within a 3D volumetric x-y-z data field after incubation with R9-GFP-H6. Note the time-dependent increase in the R9-GFP-H6 protein inside the nucleus. **(H)** Intracellular tracking of individual R9-GFP-H6 particles internalized by HeLa cells. Time spans from 20 to 300 min after exposure (from blue to white, in the individual tracks).
a.u.: Arbitrary units; GFP: Green fluorescent protein.

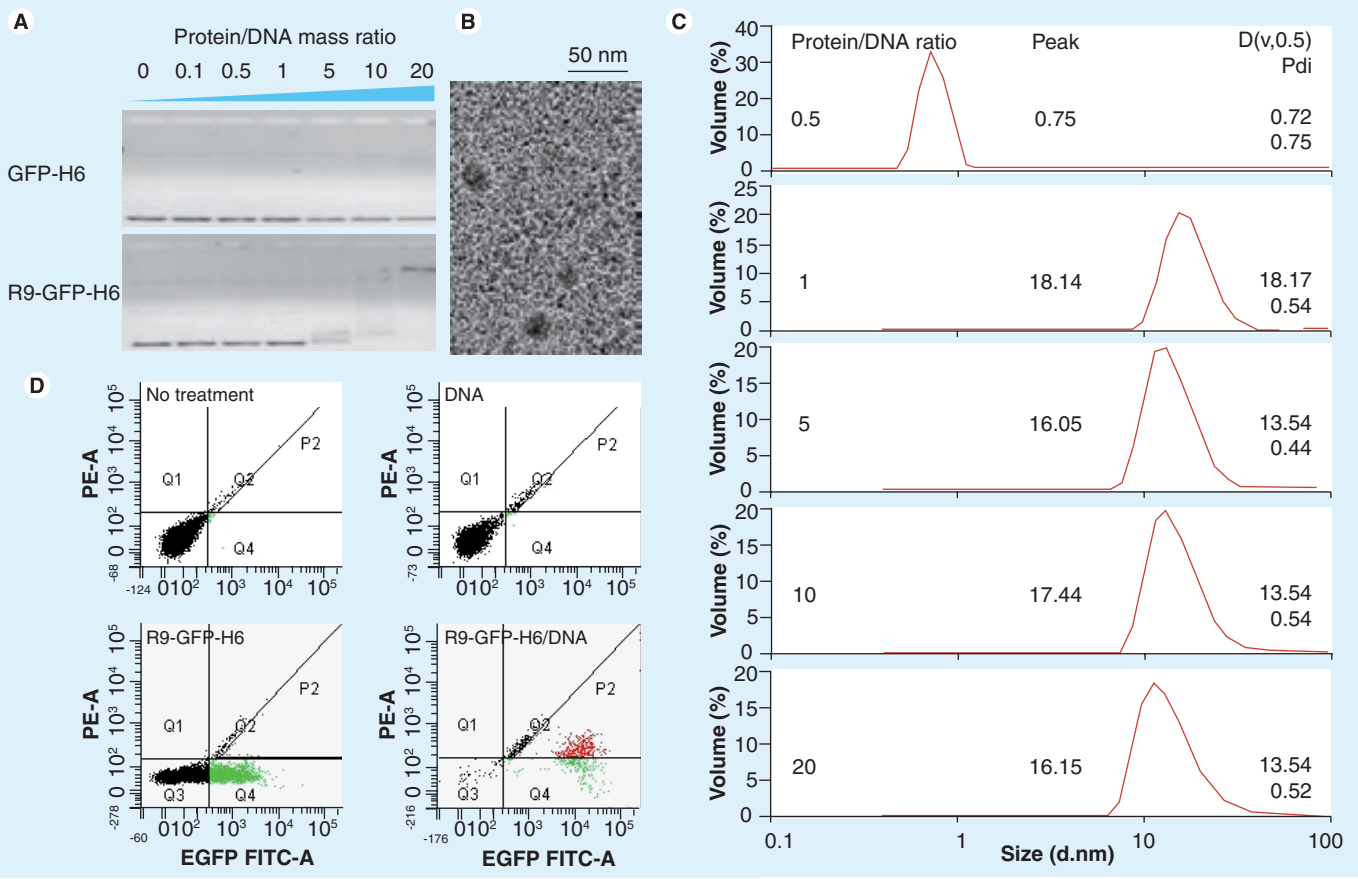

Figure 3. Evaluation of DNA-binding and gene transfer properties of R9-GFP-H6 nanodisks. (A) Retardation of plasmid DNA (pCDNA 3.1) migration in agarose gel electrophoresis promoted by increasing amounts of either GFP-H6 (top) or R9-GFP-H6. **(B)** TEM images of R9-GFP-H6-DNA complexes at a protein/DNA ratio of 20. **(C)** Particle size in R9-GFP-H6 and plasmid DNA mixtures as determined by dynamic light scattering, indicating (in nm) the peak size, the particle diameter under which there are 50% of the total volume of the population D[v,0.5] and the pdi. Inset numbers at the left indicate protein/DNA ratios. **(D)** Fluorescence emission determined by flow cytometry of untreated HeLa cells and 48 h after exposed to plasmid DNA (pCDNA 3.1 encoding the *td tomato* gene), R9-GFP-H6 and R9-GFP-H6-DNA complexes at a protein/DNA ratio of 50 (2.5 retardation units). A prolonged trypsin treatment (see experimental section) was performed to prevent fluorescence eventually emitted by cell surface attached R9-GFP-H6. Q4 section corresponds to green fluorescence mediated by R9-GFP–H6, while P2 corresponds to red emission of td Tomato, the fluorescent reported protein encoded by the transferred plasmid vector and green fluorescence in the same cell.
GFP: Green fluorescent protein; pdi: Polydispersion index.

buffers, even in those favoring protein aggregation. In addition, they are purified in a single step from crude extracts of producing bacteria and are efficiently internalized by cells in the particulate, fully functional organization. Their abilities to condense and deliver plasmid DNA to the nucleus make them intriguing platforms for their further functional tuning as artificial viruses for delivery of therapeutic transgenes or proteins. Interestingly, it has been reported that the spatial engineering of R9 peptide as tetramers significantly improves the internalization of this peptide and its associated molecules [41] and enhances the BBB crossing properties of arginine-rich peptides such as Tat [42]. Such an oligomer dependence of R9 potential as delivery agent easily accounts for the fast and efficient nuclear targeting of self-assembled R9-GFP-H6 particles. Therefore, among the current catalogue of forms that protein-based nanoplatforms for intracellular drug delivery may adopt, including cages, microspheres, virus-like particles, films, hydrogels and different types of rather amorphous nanostructures (recently reviewed in [7]), protein nanodisks should be considered as a novel class of protein planar assembles for DNA accommodation and transfer. Finally, the robustness of GFP fluorescent emission when functionalized with R9 indicates that the self-assembling properties conferred by the cationic peptide are compatible with proper protein conformation.

Under the urgent need of biologically produced improved protein drugs with tailored delivery properties [43], particularly in drug-reluctant, complex diseases such as cancer [44], the architectonic properties of R9 at the nanoscale reported here

might make this peptide a valuable module for the construction of improved, protein-based nanoconjugates for drug delivery in innovative biomedical approaches. The 'sticky' properties of R9 at a molecular level should permit an important extent of control in the generation of nanocomplexes, such that GFP-based nanodisks could represent powerful instruments in emerging nanomedical approaches to drug delivery and gene therapy.

Conclusion

By using multidisciplinary approaches, we demonstrate here that the R9 peptide, when fused to a reporter GFP protein, confers self-assembling properties to the whole construct, which spontaneously forms a novel type of disk-shaped, highly stable protein-based platform of approximately 20 nm in diameter. R9-powered nanodisks bind and condense plasmid DNA, efficiently penetrate cultured cells and rapidly reach the nucleus, where reporter transgenes are expressed. The intracellular trafficking or internalized nanodisks towards the nucleus does not compromise the particulate organization of the protein–DNA complexes or the proper folding of the individual building protein blocks. The stickiness of R9 at a molecular level and the consequent architectonic properties at the nanoscale might account for the unusual efficiency of these peptides as membrane active agents and as vehicles for systemic drug and DNA delivery. The multifunctional nature of R9 (and probably of closely related, highly cationic peptides) and in particular the so far ignored architectonic properties, make it an extremely useful tool for the design and fine tuning, through rational protein engineering, of tailored nanocarriers for DNA and drug delivery under the artificial virus concept.

Acknowledgements

The authors are indebted to Mary Cano, Marta Nadal and Carme Arnan for technical assistance in dynamic light scattering analysis, to Manuela Costa for technical assistance in flow cytometry, to Fran Cortés for technical assistance in cell culture, and Veronica Toledo and Rosa Mendoza from the CIBER de Bioingeniería Biomateriales y Nanomedicina (intramural IMAFEN project), Spain, for general laboratory assistance.

Financial & competing interests disclosure

The authors appreciate the financial support granted to Neus Ferrer-Miralles (EME2007-08, UAB), to Antonio Villaverde (BIO2007-61194 and EUI2008–03610 MICINN; 2009SGR108, AGAUR), to Jaume Veciana (CTQ2006-06333 MICINN; 2009SGR516, AGAUR) and to Antonio Villaverde and Jaume Veciana from the CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain. Ugutz Unzueta and Joan Domingo-Espín are recipients of predoctoral fellowships from MICINN, Spain, and Antonio Villaverde has been distinguished with an ICREA ACADEMIA award from AGAUR, Catalonia, Spain. 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

- The R9 peptide promotes the ordered self-assembling of a model green fluorescent protein (GFP) as disk-shaped, stable na noparticles of approximately 20 × 3 nm.
- The GFP nanodisks (but not R9-free GFP) efficiently penetrate culture mammalian cells and accumulate in their nuclei a few hours after exposure.
- Since R9 also confers affinity for DNA, protein nanodisk–plasmid DNA complexes act as efficient artificial viruses for transgene delivery and expression.

Bibliography

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

2 Abbott A: Questions linger about unexplained gene-therapy trial death. *Nat. Med.* 12(6), 597 (2006).

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

4 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).

5 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).

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

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

8 Dietz GP, Bahr M: Delivery of bioactive molecules into the cell: the Trojan horse approach. *Mol. Cell. Neurosci.* 27(2), 85–131 (2004).

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

10 Aina OH, Liu R, Sutcliffe JL, Marik J, Pan CX, Lam KS: From combinatorial chemistry to cancer-targeting peptides. *Mol. Pharm.* 4(5), 631–651 (2007).

11 Martin ME, Rice KG: Peptide-guided gene delivery. *AAPS J.* 9(1), E18–E29 (2007).

12 Low W, Mortlock A, Petrovska L, Dottorini T, Dougan G, Crisanti A: Functional cell permeable motifs within medically relevant proteins. *J. Biotechnol.* 129(3), 555–564 (2007).

13 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).

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

15 Futaki S, Suzuki T, Ohashi W *et al.*: Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. *J. Biol.Chem.* 276(8), 5836–5840 (2001).

16 Melikov K, Chernomordik LV: Arginine-rich cell penetrating peptides: from endosomal uptake to nuclear delivery. *Cell. Mol. Life Sci.* 62(23), 2739–2749 (2005).

17 Abes R, Arzumanov AA, Moulton HM *et al.*: Cell-penetrating-peptide-based delivery of oligonucleotides: an overview. *Biochem. Soc. Trans.* 35(Pt 4), 775–779 (2007).

18 Futaki S, Nakase I, Tadokoro A, Takeuchi T, Jones AT: Arginine-rich peptides and their internalization mechanisms. *Biochem. Soc. Trans.* 35(Pt 4), 784–787 (2007).

19 Kim HH, Choi HS, Yang JM, Shin S: Characterization of gene delivery *in vitro* and *in vivo* by the arginine peptide system. *Int. J. Pharm.* 335(1–2), 70–78 (2007).

20 Kim HH, Lee WS, Yang JM, Shin S: Basic peptide system for efficient delivery of foreign genes. *Biochim. Biophys. Acta* 1640(2–3), 129–136 (2003).

21 Takenobu T, Tomizawa K, Matsushita M *et al.*: Development of p53 protein transduction therapy using membrane-permeable peptides and the application to oral cancer cells. *Mol. Cancer Ther.* 1(12), 1043–1049 (2002).

22 Takayama K, Nakase I, Michiue H *et al.*: Enhanced intracellular delivery using arginine-rich peptides by the addition of penetration accelerating sequences (Pas). *J. Control Release* 138(2), 128–133 (2009).

23 Rothbard JB, Garlington S, Lin Q *et al.*: Conjugation of arginine oligomers to cyclosporin A facilitates topical delivery and inhibition of inflammation. *Nat. Med.* 6(11), 1253–1257 (2000).

24 Koppelhus U, Awasthi SK, Zachar V, Holst HU, Ebbesen P, Nielsen PE: Cell-dependent differential cellular uptake of PNA, peptides, and PNA-peptide conjugates. *Antisense Nucleic Acid Drug Dev.* 12(2), 51–63 (2002).

25 Kumar P, Wu H, McBride JL *et al.*: Transvascular delivery of small interfering RNA to the central nervous system. *Nature* 448(7149), 39–43 (2007).

26 Satheshkumar PS, Lokesh GL, Murthy MR, Savithri HS: The role of arginine-rich motif and β-annulus in the assembly and stability of Sesbania mosaic virus capsids. *J. Mol. Biol.* 353(2), 447–458 (2005).

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

28 Lundberg M, Wikstrom S, Johansson M: Cell surface adherence and endocytosis of protein transduction domains. *Mol. Ther.* 8(1), 143–150 (2003).

29 Eswar N, Marti-Renom MA, Webb B *et al.*: Comparative protein structure modeling with Modeller. In: *Current Protocols in Bioinformatics*. John Wiley & Sons, Inc. (Suppl. 15) 5.6.1–5.6.30 (2006).

30 de Vries SJ, van Dijk AD, Krzeminski M *et al.*: HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* 69(4), 726–733 (2007).

31 Singh J, Thornton JM: *Atlas of Protein Side-Chain Interactions.* IRL press, Oxford, UK (1992).

32 Wang Y, Xu J, Zhang X: Tuning the amphiphilicity of building blocks: controlled self-assembly and disassembly for functional supramolecular materials. *Adv. Mat.* 21(28), 2849–2864 (2009).

33 Pednekar D, Tendulkar A, Durani S: Electrostatics-defying interaction between arginine termini as a thermodynamic driving force in protein–protein interaction. *Proteins* 74(1), 155–163 (2009).

34 Vondrasek J, Mason PE, Heyda J, Collins KD, Jungwirth P: The molecular origin of like-charge arginine-arginine pairing in water. *J. Phys. Chem. B* 113(27), 9041–9045 (2009).

35 Zhang L, Patel HN, Lappe JW, Wachter RM: Reaction progress of chromophore biogenesis in green fluorescent protein. *J. Am. Chem. Soc.* 128(14), 4766–4772 (2006).

36 Talcott B, Moore MS: Getting across the nuclear pore complex. *Trends Cell Biol.* 9(8), 312–318 (1999).

37 Elad N, Maimon T, Frenkiel-Krispin D, Lim RY, Medalia O: Structural analysis of the nuclear pore complex by integrated approaches. *Curr. Opin. Struct. Biol.* 19(2), 226–232 (2009).

38 Greber UF, Fassati A: Nuclear import of viral DNA genomes. *Traffic* 4(3), 136–143 (2003).

39 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).

40 Limanskaya OY, Limanskii AP: Imaging compaction of single supercoiled DNA molecules by atomic force microscopy. *Gen. Physiol. Biophys.* 27(4), 322–337 (2008).

41 Hassane FS, Ivanova GD, Bolewska-Pedyczak E *et al.*: A peptide-based dendrimer that enhances the splice-redirecting activity of PNA conjugates in cells. *Bioconjug. Chem.* 20(8), 1523–1530 (2009).

42 Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF: *In vivo* protein transduction: delivery of a biologically active protein into the mouse. *Science* 285(5433), 1569–1572 (1999).

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

44 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).

Affiliations

- **Esther Vazquez**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
- **Mónica Roldán**
  Servei de Microscòpia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

- **César Diez-Gil**
  Department of Molecular Nanoscience and Organic Materials, Institut de Ciencia de Materials de Barcelona (CSIC), Bellaterra, 08193 Barcelona, Spain
  and
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
- **Ugutz Unzueta**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
- **Joan Domingo-Espín**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
- **Juan Cedano**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
- **Oscar Conchillo**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
- **Imma Ratera**
  Department of Molecular Nanoscience and Organic Materials, Institut de Ciencia de Materials de Barcelona (CSIC), Bellaterra, 08193 Barcelona, Spain
  and
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
- **Jaume Veciana**
  Department of Molecular Nanoscience and Organic Materials, Institut de Ciencia de Materials de Barcelona (CSIC), Bellaterra, 08193 Barcelona, Spain
  and,
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
- **Xavier Daura**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  Catalan Institution for Research and Advanced Studies (ICREA), 08010 Barcelona, Spain
- **Neus Ferrer-Miralles**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
  neus.ferrer@uab.cat
- **Antonio Villaverde**
  Institute for Biotechnology and Biomedicine, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
  and
  CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain
  avillaverde@servet.uab.es

Article 4

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

Domingo-Espin, J., Vazquez, E., Ganz, J., Conchillo, O., Garcia-Fruitos, E., Cedano, J., Unzueta, U., Petegnief, V., Gonzalez-Montalban, N., Planas, A. M., Daura, X., Peluffo, H., Ferrer-Miralles, N., & Villaverde, A.

Nanomedicine 6, 1047-1061 (2011).

Here we use for the first time the protein named HNRK, which is analysed later in the next reference. In this paper, a new strategy is used in the construction of protein-only gene therapy vehicles. Basically, different functional modules are fused without the use of a scaffold protein to carry them. We produced in *E. coli* two chimerical peptides of 10.2 kDa, each containing four biologically active domains in different order, which act as building blocks of protein-based non-viral vehicles for gene therapy. We wanted to focus on their biological performance as gene delivery vehicles but also in their architectonic organization and the interactions needed for their complexation.

We demonstrate that these proteins are able to promote significant levels of transgene expression in cultured cells and also that the formation of the nanometric spherical entities formed when DNA is added is driven by protein-DNA interactions and not by protein-protein associations.

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

part of future medicine 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.

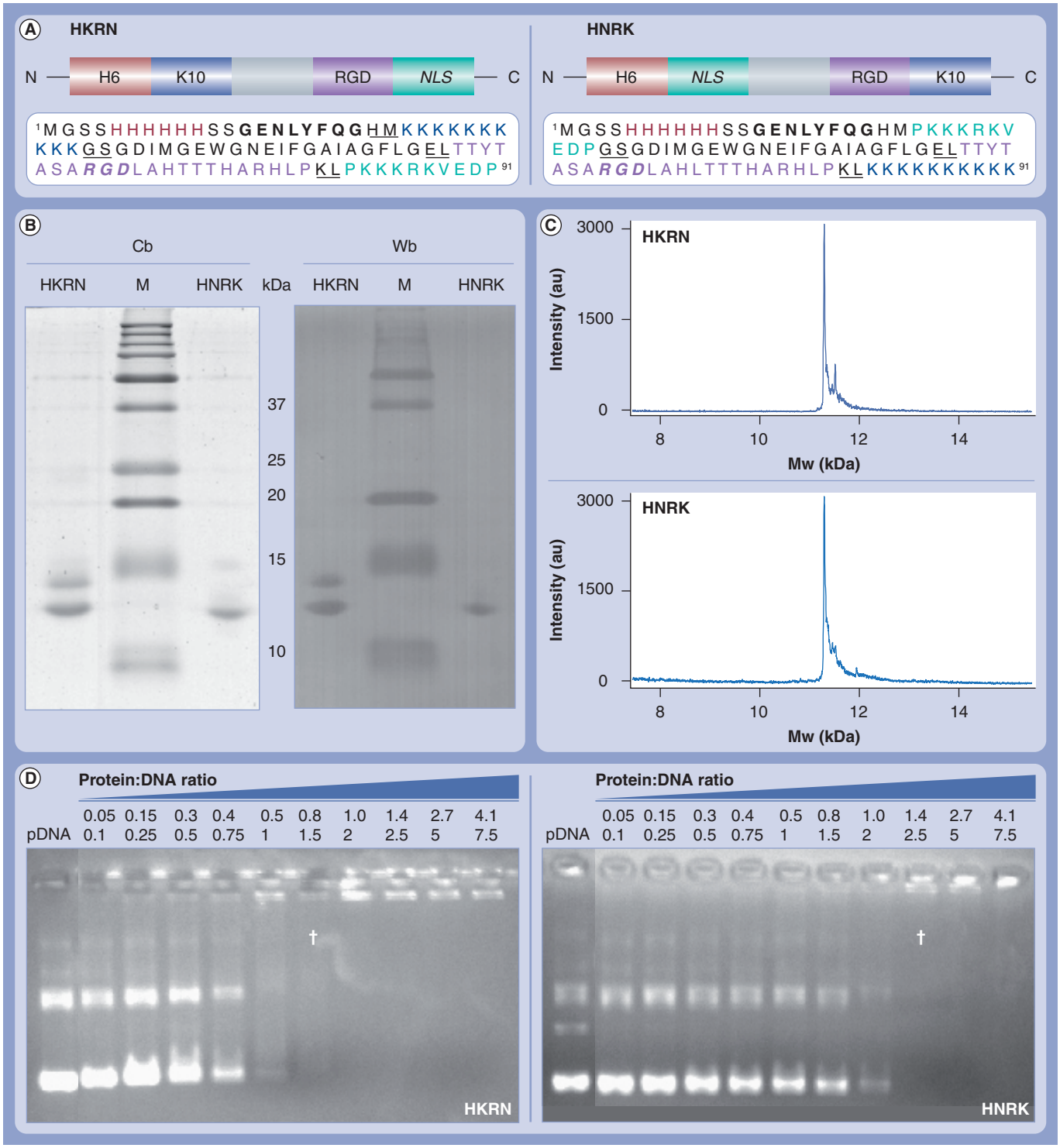

Figure 1. HKRN and HNRK organization and main features. (A) Distribution of functional modules in HKRN, HNRK and the amino acid sequence of the whole protein constructs. In both cartoons and sequence, the histidine (H) tail is labeled in red, the lysine (K) tail in dark blue, the FMDV cell binding (arginine–glycine–aspartic [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.

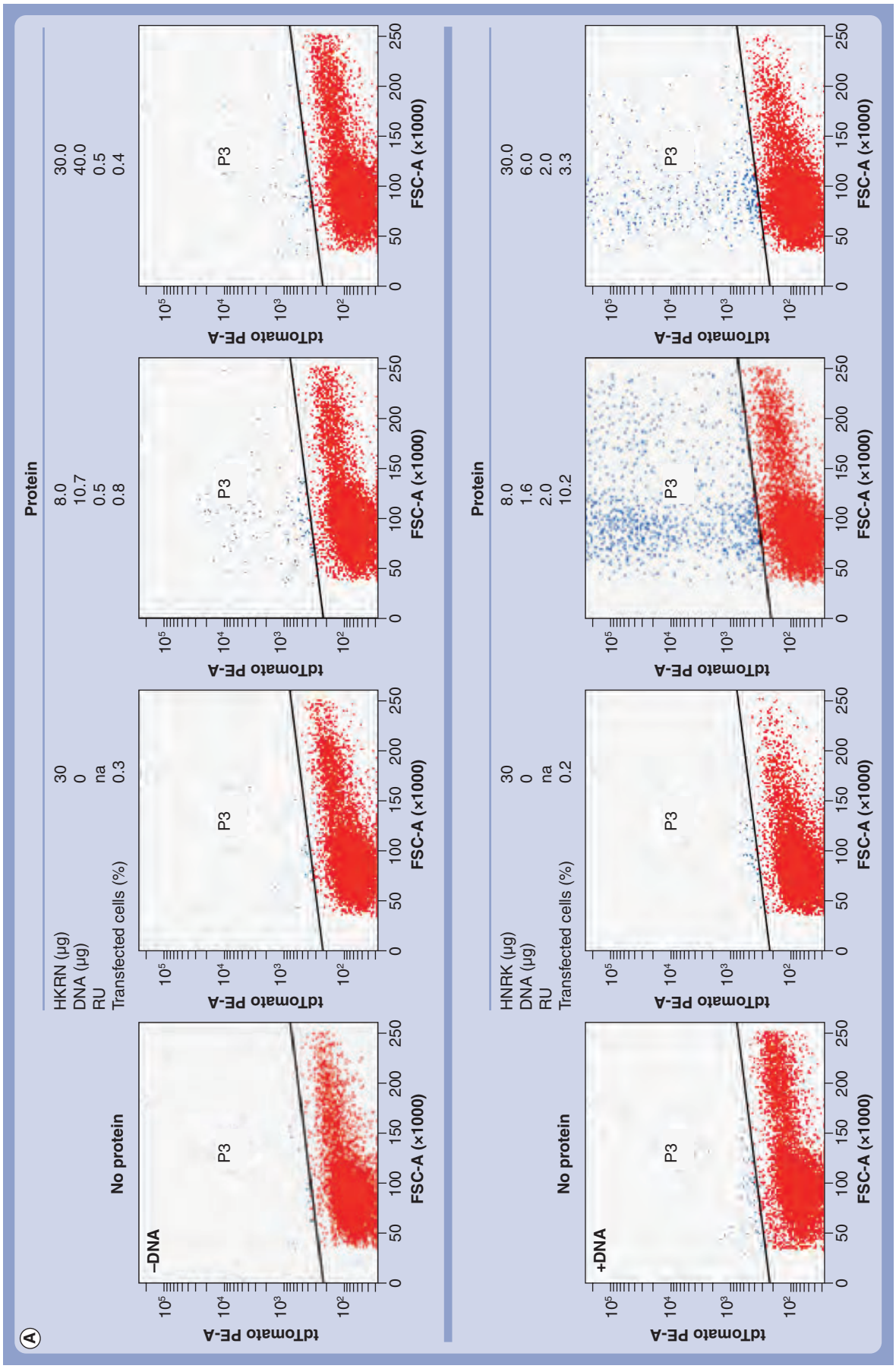

Figure 2. Transgene expression mediated by HKRN- and HNRK-based artificial viruses. (A) Fluorescence emission determined by flow cytometry of cultured HeLa cells 48 h after exposure to 24 µg of pcDNA3.1 or in the absence of foreign DNA (no protein). Cells were also exposed to HKRN–DNA and HNRK–DNA complexes 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.

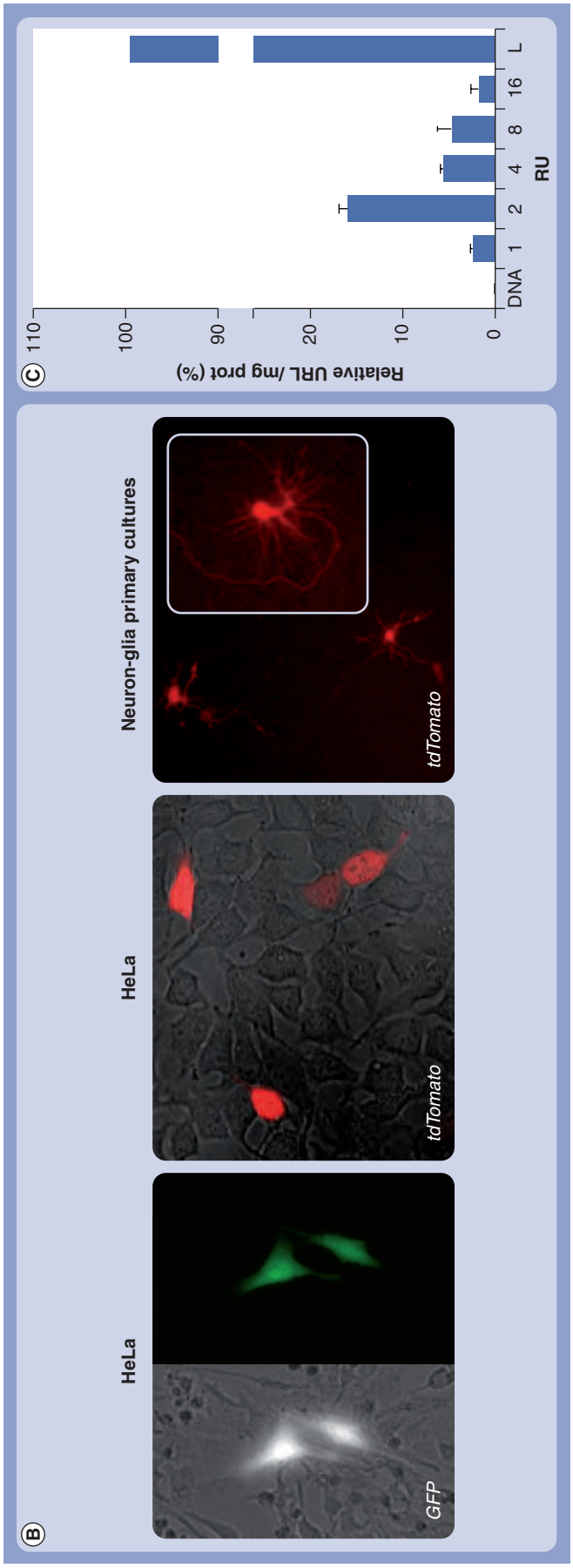

Figure 2 (cont.). Transgene expression mediated by HKRN- and HNRK-based artificial viruses. (B) Fluorescence microscopy of cultured cells 24 h after exposure to HKRN–DNA complexes formed at 2 RU. **(C)** Light emission of HEK293 cells 48 h after exposure to HKRN–pBOS complexes formed at 1, 2, 4, 8 and 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 4B, 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

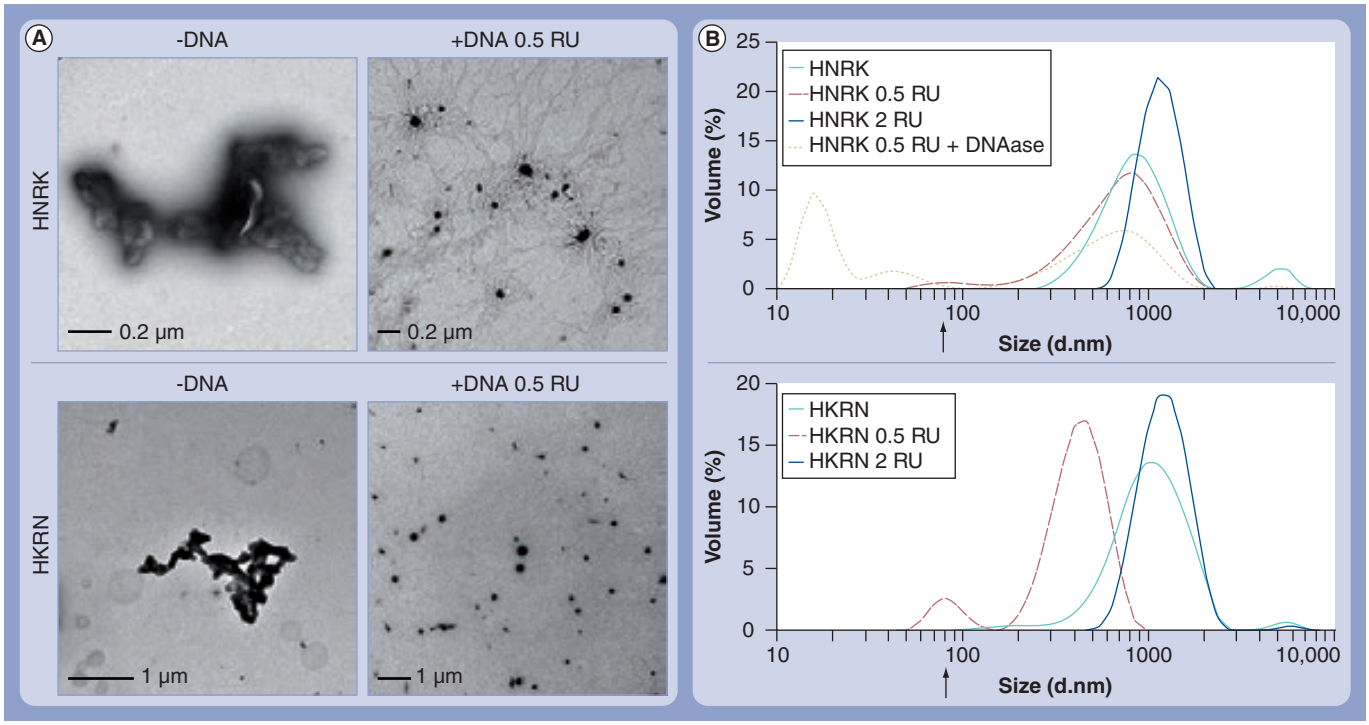

Figure 3. Size and morphology of HKRN and HNRK and their derived artificial viruses. (A) Representative transmission electron microscopy images of both HKRN and HNRK alone and as complexes with plasmid pcDNA3.1. **(B)** Effective size of protein particles alone or protein–DNA complexes (HNRK at the top 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,

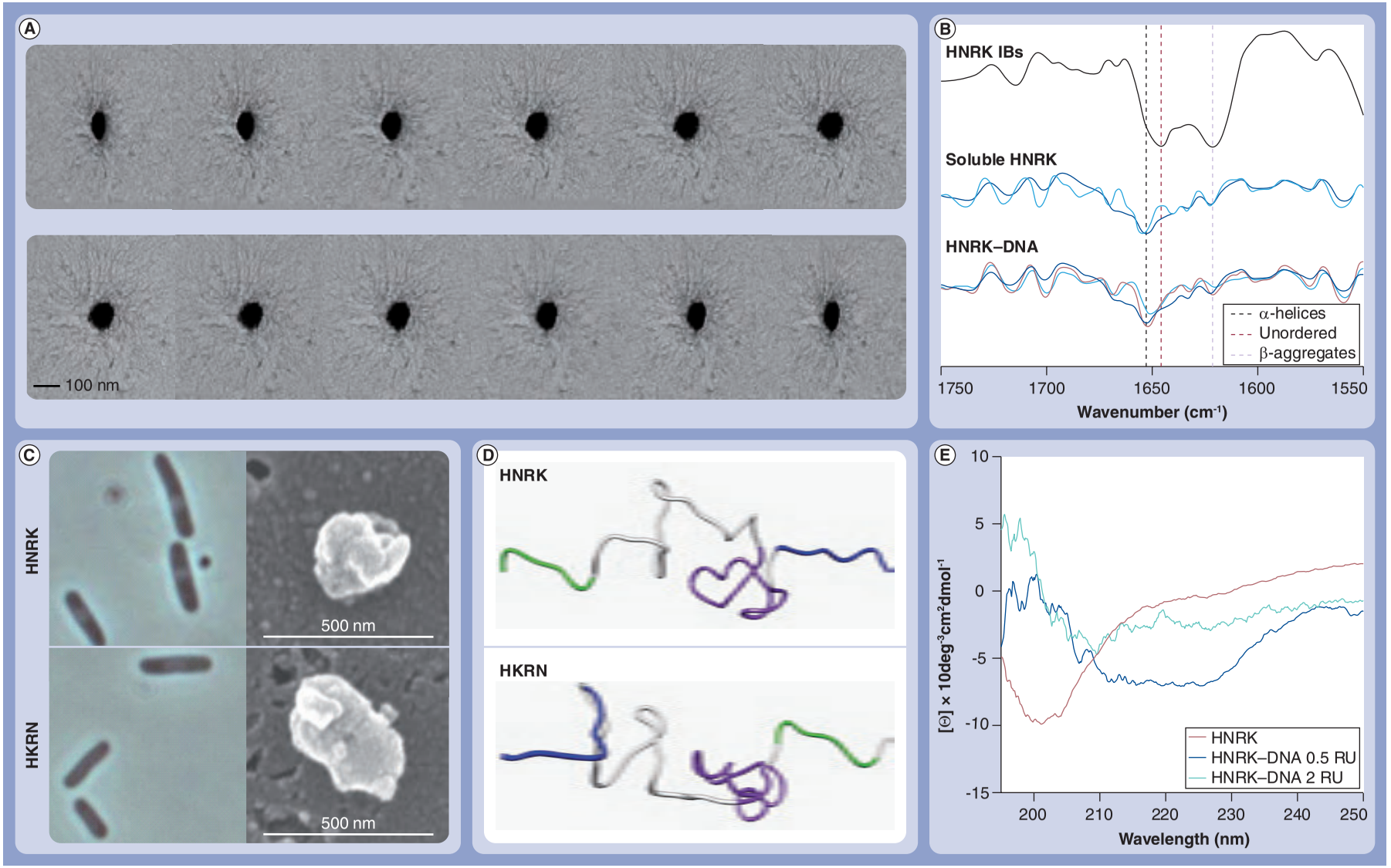

Figure 4. Molecular organization and conformation of HKRN and HNRK building blocks. (A) Selection of TEM micrograph images from a series of images acquired at inclinations from -60° to +60°. **(B)** Second derivative attenuated total reflectance-Fourier transformed infrared spectroscopy absorption spectra 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 pseudospherical 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.**

Article 5

RGD-based cell ligands for cell-targeted drug delivery act as potent trophic factors

Domingo-Espin, J., Petegnief, V., de Vera, N., Conchillo-Sole, O., Saccardo, P., Unzueta, U., Vazquez, E., Cedano, J., Negro, L., Daura, X., Peluffo, H., Planas, A. M., Villaverde, A., & Ferrer-Miralles, N.

Nanomedicine-Nanotechnology Biology and Medicine 8, 1263-1266 (2012).

In this work we deeply studied the protein HNRK and specifically the role that the Arg-Gly-Asp (RGD) module has in cell attachment and internalization into cells for DNA expression . When doing this work in several cell lines we discovered that when cells were challenged with this protein, changes in their morphology and cell density appeared. With these findings, we wanted to understand whose and how those changes were caused and introduce the concept of "therapeutic modules" in modular protein engineering for gene delivery.
By testing different cell lines we analysed the dependence of $\alpha v\beta 3$ integrins in cell internalization and DNA expression mediated by HNRK and also we showed here that the RGD motif acts as a powerful trophic factor, supporting extracellular signal–regulated kinase 1/2 (ERK1/2)–linked cell proliferation and partial differentiation of PC12 cells, a neuronal cell model.

BASIC SCIENCE

Nanomedicine: Nanotechnology, Biology, and Medicine
8 (2012) 1263–1266

Short Communication

nanomedjournal.com

RGD-based cell ligands for cell-targeted drug delivery act as potent trophic factors

Joan Domingo-Espín, MSc[a,b,c], Valérie Petegnief, PhD[d], Núria de Vera, MD, PhD[d], Oscar Conchillo-Solé, MSc[a], Paolo Saccardo, MSc[a,b,c], Ugutz Unzueta, MSc[a,b,c], Esther Vazquez, MD, PhD[a,b,c], Juan Cedano, PhD[e], Luciana Negro, BSc[f], Xavier Daura, PhD[a,f], Hugo Peluffo, PhD[g,h], Anna M. Planas, PhD[d], Antonio Villaverde, PhD[a,b,c,*], Neus Ferrer-Miralles, PhD[a,b,c]

[a]*Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Barcelona, Spain*
[b]*Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Barcelona, Spain*
[c]*CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona, Spain*
[d]*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*
[e]*Laboratory of Immunology, Regional Norte, Universidad de la Republica, Salto, Uruguay*
[f]*Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain*
[g]*Neurodegeneration Laboratory, Institut Pasteur de Montevideo, Montevideo, Uruguay*
[h]*Department of Histology and Embryology, Faculty of Medicine, Universidad de la República, Montevideo, Uruguay*

Received 8 March 2012; accepted 28 June 2012

Abstract

Integrin-binding, Arg-Gly-Asp (RGD)–containing peptides are the most widely used agents to deliver drugs, nanoparticles, and imaging agents. Although in nature, several protein-mediated signal transduction events depend on RGD motifs, the potential of RGD-empowered materials in triggering undesired cell-signaling cascades has been neglected. Using an RGD-functionalized protein nanoparticle, we show here that the RGD motif acts as a powerful trophic factor, supporting extracellular signal–regulated kinase 1/2 (ERK1/2)–linked cell proliferation and partial differentiation of PC12 cells, a neuronlike model.

From the Clinical Editor: This work focuses on RGD peptides, which are among the most commonly used tags for targeted drug delivery. They also promote protoneurite formation and expression of neuronal markers (MAP2) in model PC12 cells, which is an unexpected but relevant event in the functionalization of drugs and their nanocarriers.
© 2012 Elsevier Inc. All rights reserved.

Key words: Protein nanoparticles; Integrins; RGD; Neurons; Biomaterials

Innovative medicines aim to the cell-targeted delivery of therapeutic cargos. The concept supporting this approach is that specific binding to selected cell surface receptors should result in improved biodistribution and enhanced cellular penetrability. Since the discovery of the Arg-Gly-Asp (RGD) motif as a potent ligand of cell surface integrins (mainly $\sigma_v\beta_3$),[1] RGD-containing peptides have been extensively used to control drug biodistribution, being the agents most employed for cell targeting and endosomal delivery.[2] As integrins abound in endothelial cells and because of the high vascularization of tumoral tissues, RGD-mediated drug delivery is of special interest in cancer therapies.[2] Interestingly, some RGD-containing natural proteins, including fibulin-5,[3] osteopontine,[4] the angiogenic factor Del1,[5] and the dentin matrix acidic

The authors acknowledge the financial support granted to A.V. from Ministerio de Ciencia e Innovación (MICINN) (ACI2009-0919), Agència de Gestió d'Ajuts Universitaris i de Recerca (2009SGR-108), and Centro de Investigación Biomédica en Red (CIBER) de Bioingeniería, Biomateriales y Nanomedicina, 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 (ISCIII) with assistance from the European Regional Development Fund. This work was also supported by MICINN to A.M.P. (SAF2011-30492, Spain) and by ANII (FCE_192, Uruguay) to H.P. U.U. and P.S. received PhD fellowships from ISCIII and J.D.E. from MICINN. A.V. has been distinguished with an Institució Catalana de Recerca i Estudis Avançats Acadèmia award.

*Corresponding author: Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain.
E-mail address: antoni.villaverde@uab.cat (A. Villaverde).

1549-9634/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.nano.2012.06.005

Please cite this article as: Domingo-Espín J., et al, RGD-based cell ligands for cell-targeted drug delivery act as potent trophic factors. *Nanomedicine: NBM* 2012;8:1263-1266, http://dx.doi.org/10.1016/j.nano.2012.06.005

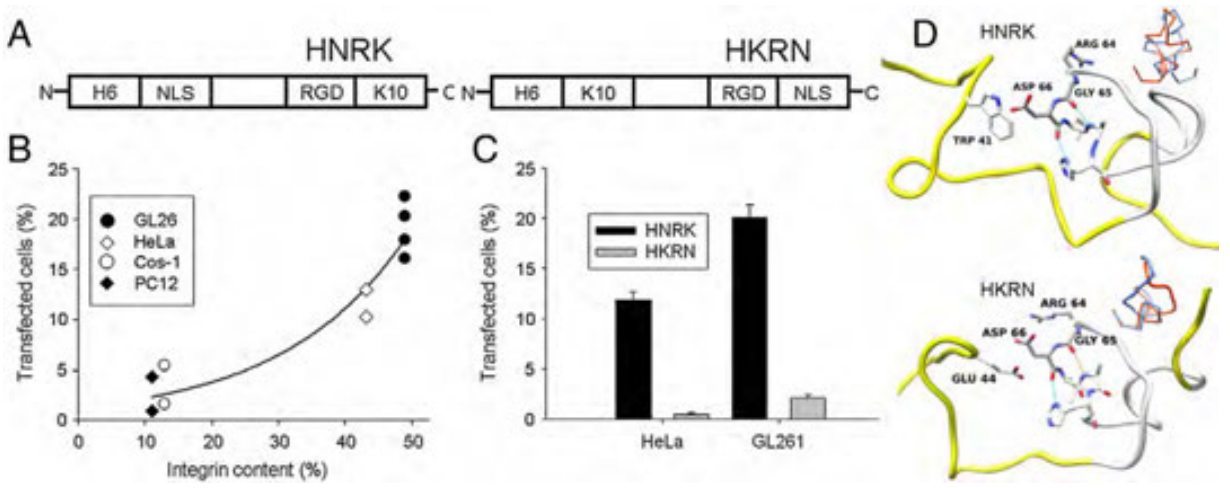

Figure 1. **(A)** Schematic representation of HNRK and HKRN (modified from Domingo-Espín et al[7]). **(B)** Relationship between the percentage of integrin-positive cells and transgene expression levels mediated by HNRK. **(C)** Percentage of HeLa and GL261 transfected cells, mediated by HNRK-DNA or HKRN-DNA nanoparticles. **(D)** Detail of the residues closer than 4 Å to the Arg residue in the RGD module (*white*; *yellow*, the rest of the peptide), modeled using the segment 134–156 from FMDV VP1 (pdb 1QGC). In the insets, superposition of the recombinant RGD modules (*blue*) over the template (*red*) used to model it, again residues 134–156 from FMDV VP1 (pdb 1QGC). Details of all the experimental procedures used in the study can be found in the Supplementary Information.

phosphoprotein 1 (DMP1),[6] act as biological effectors through their RGD motifs, at least some of them through the extracellular signal–regulated kinase 1/2 (ERK1/2)–mitogen-activated protein kinase (MAPK) pathway. Therefore, it would not be unexpected that RGD-empowered constructs for drug targeting could trigger side responses in target cells, although this possibility has not been investigated.

HNRK is a short modular protein of 91 amino acids in length, consisting of four consecutive functional peptides, namely an endosomolytic poly-histidine peptide (H), a viral nuclear localization signal (N), an RGD-containing site of foot-and-mouth disease virus (FMDV) serotype C1 (R), and a C-terminal poly-lysine (K) peptide (Figure 1, *A*). In the presence of plasmid DNA, HNRK forms 80-nm polyplexes efficient in DNA delivery to mammalian cells.[7] The transfection efficiency of HNRK nanoparticles is highly dependent on the amount of $\alpha_v\beta_3$ displayed on the cell surface (Figure 1, *B*), indicating that the RGD stretch determines the capability of the construct to penetrate the target cell and deliver the cargo. Contrarily, HKRN, a modular isoform of HNRK (Figure 1, *A*), fails to promote high levels of transgene expression (Figure 1, *C*), a fact that could be attributed to a poor presentation of the cell ligand. In this regard, in a homology model of HNRK, RGD is folded as a mirrorlike structure of the parental viral segment, whereas in HKRN the segment folds in part as a mirror and in part as a fully matching conformation, stabilized by different interactions (Figure 1, *D*). Interestingly, when exposing PC12 cells to HNRK alone (but not to HKRN), protoneurite formation was unexpectedly observed (Figure 2, *A*, and Supplementary Figure S1, available online at http://www.nanomedjournal.com). In this context, HNRK specifically stimulated PC12 cell growth at levels comparable or even higher than those promoted by the nerve growth factor (NGF) (Figure 2, *B*, *C*). Both the morphological modification of PC12 cells and the enhanced cell growth upon exposure to HNRK indicated that this protein could act as a cell trophic factor. Because HKRN, in which the

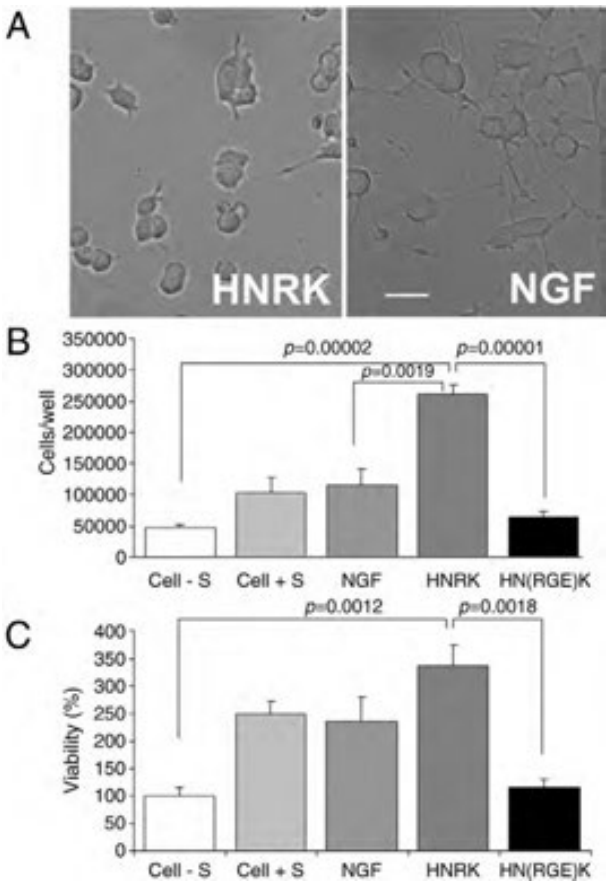

Figure 2. **(A)** Magnifications of HNRK- and NGF-treated cells (at 100 ng/mL). Scale bar, 20 μm. Proliferation of PC12 cells monitored by cell counting **(B)** or relative MTT activity **(C)**. Cells were cultured for 7 days in the absence of serum (Cell –S); with 20% serum (Cell +S); in absence of serum plus nerve growth factor (NGF); in absence of serum plus either HNRK or HN(RGE)K. Only significantly different pairs of comparative data are indicated ($P < 0.01$, $n = 3$).

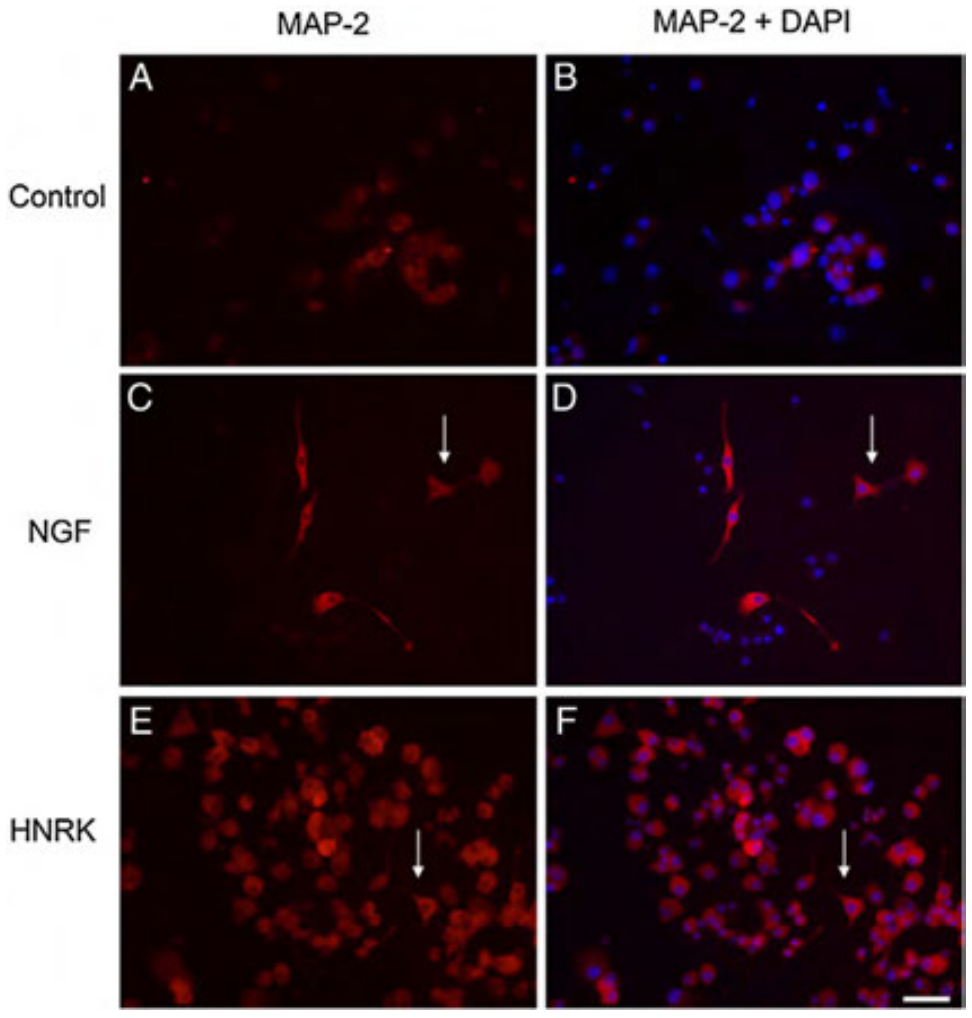

Figure 3. Immunofluorescence for MAP2 (*red*) in cells either maintained under control conditions (20% serum) **(A**, **B)** or treated with NGF **(C**, **D)** or HNRK **(E**, **F)**. Nuclei are stained with DAPI (*blue*) in panels *B*, *D*, and *F*. Neurite extensions are indicated by *arrows*. Scale bar, 50 μm.

displayed RGD is unable to efficiently bind integrins, did not stimulate cell survival, growth, and differentiation, it could be speculated that a functional RGD motif could be responsible for the HNRK-induced trophic effect. This was fully supported by the fact that the HNRK variant HN(RGE)K, in which the RGD sequence was mutated into the nonbinding RGE motif, did not show trophic effects (Figure 2, *B*, *C*).

To explore the biological effects mediated by HNRK, we analyzed the signaling events in HNRK-exposed PC12 cells through the expression of microtubule-associated proteins typically used as neuronal markers. Indeed, HNRK upregulated microtubule-associated protein-2 (MAP2) (Figure 3) and Tau (not shown). Interestingly, HNRK but not HN(RGE)K also stimulated the phosphorylation of ERK1/2 (Figure 4, *A*). The degree of ERK1/2 phosphorylation induced by HNRK was lower than that triggered by NGF, whereas in contrast, the chimeric protein stimulated cell proliferation more than NGF did (Figure 2). These results suggest different effector pathways for NGF and HNRK. PD98059, an inhibitor of the MAPK kinases MEK1 and 2 that phosphorylates ERK1/2, reduced NGF-induced ERK1/2 phosphorylation and fully abolished HNRK-induced ERK1/2 phosphorylation (Figure 4, *A*). Interestingly, whereas cell proliferation mediated by HNRK was significantly reduced by more than 50% by PD98059, this inhibitor showed a trend for reduction of NGF-induced cell proliferation by only around 15% (Figure 4, *B*).

Finally, to explore whether the trophic effects of RGD described here could be restricted to a specific cell lineage (neurons), we extended our study to rat primary glial cell cultures enriched in astrocytes. In this additional model, a dose-dependent ERK1/2 phosphorylation was clearly observed upon exposure to HNRK, indicating that the effector role of the peptide motif is instead a generic event (Figure 4, *C*).

The tripeptide RGD is one of the most widely used cell ligands in nanomedicine for the targeted delivery of drugs, as associated with chemicals, protein-only constructs, self-assembling peptides, liposomes, dendrimers, carbon nanotubes, magnetic nanoparticles, quantum dots, polymers, and other delivery agents.[8,9] RGD is also used as a cell-binding agent in cell culture and in different approaches of tissue engineering as a functional

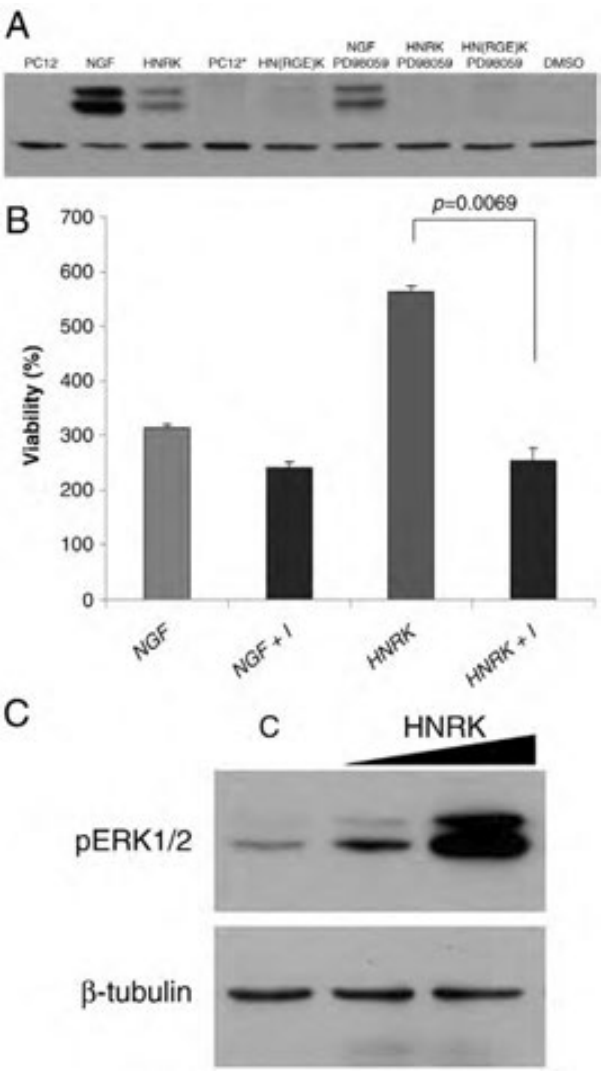

Figure 4. **(A)** ERK1/2 phosphorylation determined by western blotting (n = 3). β-tubulin is shown as a gel-loading control. **(B)** Cell viability determined by MTT assay. Only significantly different pairs of comparative data are indicated ($P < 0.01$, $n = 3$). **(C)** ERK1/2 phosphorylation determined by western blotting 2 hours after exposure of astrocyte-enriched glial cell cultures to HNRK (75 ng/μL and 150 ng/μL, respectively). β-tubulin is also shown as a gel-loading control. C, nonexposed control cells.

coating agent to favor cell binding to the substrate and differentiation.[10] Therefore, the cell trophic activities of RGD described in the present study are of potential interest when evaluating the biological and therapeutic effects of RGD-based therapies, because the side effects of the vehicle might antagonize or synergistically act with those of the therapeutic cargo. This might be of special interest if the nanoparticle-linked drug is intended to arrest cell growth or is a cell-killing agent, as usually occurs in antitumoral therapy. Importantly, the catalogues of agents used in nanomedicine for cell targeting are rich in natural or mimetic ligands of receptors for hormones, cytokines, or other effectors that might potentially trigger signaling events.

Acknowledgments

The authors appreciate technical support from Ms. Leonardo Márquez-Kisinousky, from the Cell Culture Unit of Servei de Cultius Cellulars, Producció d'Anticossos i Citometria (SCAC, UAB), from Servei de Microscòpia UAB and from the Protein Production Platform (CIBER-BBN; http://bbn.ciber-bbn.es/programas/plataformas/equipamiento).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.nano.2012.06.005

References

1. Ruoslahti E, Pierschbacher MD. New perspectives in cell adhesion: RGD and integrins. *Science* 1987;**238**:491-7.
2. Meyer A, Auernheimer J, Modlinger A, Kessler H. Targeting RGD recognizing integrins: drug development, biomaterial research, tumor imaging and targeting. *Curr Pharm Des* 2006;**12**:2723-47.
3. Yue W, Sun Q, Landreneau R, Wu C, Siegfried JM, Yu J, et al. Fibulin-5 suppresses lung cancer invasion by inhibiting matrix metalloproteinase-7 expression. *Cancer Res* 2009;**69**:6339-46.
4. Tanabe N, Wheal BD, Kwon J, Chen HH, Shugg RP, Sims SM, et al. Osteopontin signals through calcium and nuclear factor of activated T cells (NFAT) in osteoclasts: a novel RGD-dependent pathway promoting cell survival. *J Biol Chem* 2011;**286**:39871-81.
5. Wang Z, Kundu RK, Longaker MT, Quertermous T, Yang GP. The angiogenic factor Del1 prevents apoptosis of endothelial cells through integrin binding. *Surgery* 2011;**151**:296-305.
6. Eapen A, Ramachandran A, Pratap J, George A. Activation of the ERK1/2 mitogen-activated protein kinase cascade by dentin matrix protein 1 promotes osteoblast differentiation. *Cell Tiss Org* 2011;**194**: 255-60.
7. Domingo-Espín J, Vazquez E, Ganz J, Conchillo O, Garcia-Fruitos E, Cedano J, et al. Nanoparticulate architecture of protein-based artificial viruses is supported by protein–DNA interactions. *Nanomedicine (Lond)* 2011;**6**:1047-61.
8. Dunehoo AL, Anderson M, Majumdar S, Kobayashi N, Berkland C, Siahaan TJ. Cell adhesion molecules for targeted drug delivery. *J Pharm Sci* 2006;**95**:1856-72.
9. Villaverde A, editor. Nanoparticles in translational science and medicine. From the series: Progress in Molecular Biology and Translational Science, vol. 114. London, UK: Academic Press (Elsevier); 2011.
10. Williams DF. The role of short synthetic adhesion peptides in regenerative medicine; the debate. *Biomaterials* 2011;**32**:4195-7.

Discussion

During the last years, the interest in the use of non-viral gene therapy vectors has grown enormously due to the fact that they are efficient and safer compared to viral vectors that present many limitations such as the immunological responses associated to their use, the risk of insertional mutagenesis, their potentially replicable genomes that could integrate or recombine with cell DNA and the high cost of their production. Moreover, the size of the material to transfect is limited and their cell specificity, sometimes, is not high enough. Viral engineering has solved some of these problems but production costs remain high.

Many non-viral vector systems have been proposed to overcome the viral limitations with different degrees of success. Here, we have presented an approach that present many advantages. Modular protein engineering enables the design of artificial proteins constructed to contain different functional building blocks (modules). The combination of these modules confers to protein-based vectors a tailored biological activity, aimed to optimise nucleic acid delivery to target cells.

The final tailored modular protein vector is produced in bacteria and purified by standardized production techniques. Therefore, the production of this vehicle is inexpensive compared to genetically modified viral vectors and it does not require high biosafety facilities. Moreover, it can be easily scaled up for its industrial manufacturing. Many different vectors can be designed, produced and tested in short time, selecting the ones that best fit our needs.

Production systems for recombinant proteins

Among the 151 protein-based recombinant pharmaceuticals licensed up to January 2009 by the FDA and EMEA, 45 (29.8%) are obtained in *Escherichia coli*, 28 (18.5%) in *Saccharomyces*

cerevisiae, 17 (11.2%) in hybridoma cells, 1 in transgenic goat milk, 1 in insect cells and 59 (39%) in mammalian cells (Article 1, Figure 1). *E. coli* is the production system for all the proteins studied in this thesis, and therefore it is very interesting to become acquainted of the presence, use and characteristics of therapeutic molecules produced in this system and others in the market.

E. coli presents several obstacles for the production of quality proteins, limiting its application as a factory for recombinant pharmaceuticals. One of the problems of this production system is that recombinant proteins obtained in *E. coli* lack many post-translational modifications (PTMs) important for both solubility and biological activity[297]. Another one, as it has been well documented, is that bacteria overproducing either eukaryotic or prokaryotic recombinant proteins are subjected to different stresses (essentially metabolic and conformational), which may lead to the production of insoluble aggregates named inclusion bodies[298–301].

There are other protein production systems that have been used to obtain products in the market but without lowering the rate of bacterial recombinant pharmaceuticals approved (Article 1, Figure 3). This is probably due to the extremely different biological and technological backgrounds associated to protein production, the good quality of microbial products and the high cost associated to mammalian cell production. In addition, this fact indicates the potential of microbial cells in biopharmaceutical industry despite the limited post transcriptional modification PTM performance of their products. Also, microbial cell factories cover a spectrum of products and application fields that do not necessarily match those addressed by mammalian cell factories.

Interestingly, a plateau in the rate of approved recombinant protein pharmaceuticals is becoming perceivable, irrespective of the production system (Article 1, Figure 3). Although it might be observed as a transient event, this fact seems instead to indicate that current production systems could be near to the exhaustion regarding their ability to hold the production of

complex proteins, protein complexes or the so called difficult-to-express proteins[302]. Desirably, recent insights about system's biology of recombinant cells and hosts, and specially, arising novel concepts on recombinant protein quality[303–305] and host stress responses[306] would enlarge the possibilities for metabolic and process engineering aiming to the economically feasible production of new, more complex biological entities. Indeed, pushed by fast advances in molecular medicine, the pharmaceutical industry is urgently demanding improved production systems and novel and cheaper drugs.

Overcoming the biological and methodological obstacles posed by cell factories to the production of recombinant DNA pharmaceuticals is a main challenge in the further development of protein-based molecular medicine. Recombinant DNA technologies might have exhausted conventional cell factories and new production systems need to be deeply explored and incorporated into the production pipeline. On the other hand, a more profound comprehension of host cell physiology and stress responses to protein production would be necessary to offer improved tools to favour high yield and high quality protein production. Apart from the expected incorporation of unusual mammalian hosts such as transgenic animals or plants, microbial cells appear as extremely robust and convenient hosts, and gaining knowledge about the biological aspects of protein production would hopefully enhance the performance of such hosts beyond the current apparent limitations. In this regard, not only commonly used bacteria and yeasts but also unconventional strains or species are observed as promising cell factories for forthcoming recombinant drugs[302]. Their incorporation into productive processes for human pharmaceuticals would hopefully push the trend of marketed products and fulfil the increasing demands of the pharmacological industry.

Exploring the R9 peptide as a functional and architectonic module for gene therapy vectors

Arginine(R)-rich cationic peptides are powerful tools in drug delivery since, alone or when associated with polyplexes, proteins or chemicals, they confer DNA condensation, membrane translocation and blood–brain barrier crossing abilities[307–312]. The unusual stability and high in vivo performance of their associated drugs suggest a particulate organization or R(n) complexes, which is very interesting to explore[313].

Apart from the well-described properties of polyarginine peptides as membrane translocators and cross-BBB carriers[314], the data presented demonstrated their architectonic potential through the fusion of an R9 tail to a scaffold GFP. This peptide triggers the spontaneous self-assembling of the chimerical protein as regular, highly ordered disk-shaped structures of 20 × 3 nm, probably formed by five or more proteins (Article 3, Figure 1). These nanodisks are highly soluble and stable in a diversity of buffers, even in those favouring protein aggregation (Article 3, Figure 1). In addition, cells efficiently internalize them (Article 3, Figure 2). Their ability to condense and deliver plasmid DNA to the nucleus make them interesting entities for their further functional tuning as artificial viruses for delivery of therapeutic NAs. Interestingly, it has been reported that the spatial engineering of R9 peptide as tetramers significantly improves the internalization of this peptide and its associated molecules[315], and enhances the BBB crossing properties of arginine-rich peptides such as Tat[316]. Therefore, among the different forms that protein-based nanoparticles for gene delivery may adopt, including cages, microspheres, virus-like particles, films, hydrogels and different types of amorphous nanostructures[317], protein nanodisks are a novel class of protein planar assemblies for DNA accommodation and transfer. Finally, the sustained GFP fluorescent emission, in all the experiments performed, indicates that R9 peptide could be a good functionalizing module for other proteins.

Under the urgent need of biologically produced improved protein drugs with tailored delivery properties[302], particularly in drug-resistant and complex diseases[89], the architectonic properties reported for R9 functionalized proteins might make this peptide a valuable module for the construction of improved, protein-based nanoconjugates for drug delivery in innovative biomedical approaches. The inter-protein complexation properties of R9 should provide a new arquitectonic tool to control the formation of nanocomplexes, such that R9-powered nanodisks could represent powerful instruments in emerging nanomedical approaches to drug delivery and gene therapy.

Exploring the protein-DNA architecture of protein-based artificial viruses

Non-viral 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 and molecular organization, nanoscale properties potentially critical for cell attachment, internalization and endosomal escape remain excluded from potential tailoring. To approach this issue, we tested two versions of proteins produced in *E. coli* displaying functional modules that were necessary for gene transfer in a single polypeptide chain (Article 4, Figure 1). 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 (Article 4, Figure 3), that resemble bacterial inclusion bodies (IBs) which are structured by protein-protein interactions[318,319]. However, both, protein modeling and structural analysis of these complexes, revealed an unexpected

molecular organization that does not rely on protein–protein cross-molecular interactions but that it is instead supported by protein–DNA interactions (Article 4, Figure 4). Such DNA-mediated organization seems to generate an optimal architectural pattern of artificial viruses based on short multifunctional proteins as building blocks.

Multifunctional proteins based on large scaffold proteins such as *E. coli* β-galactosidase, organize as amorphous polydisperse protein clusters whose properties seem to be defined by protein features rather than by the presence of DNA[210]. The addition of 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, Article 3, Figures 1 and 3) also irrespectively of the presence of DNA[320].

On the other hand, the nanoscale organization of two short multifunctional proteins, which are competent in gene delivery by using both cultured cell lines and primary cell culture models, is different (Article 4, Figure 2). The transgene expression levels and stability that are reached with these proteins are comparable or higher than those observed with previous prototypes of artificial viruses based on multifunctional proteins[207,210,216,320,321]. The less active modular protein version, namely the construct HKRN (5% of cells showed transgene expression), achieved approximately 18% of the expression level observed when using lipofectamine (Article 4, Figure 2C). The 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. This is indicative of functional dependence of the diverse modules composing the building block, which seems to be 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 favourable

strategy regarding productivity in bacteria, when comparing with the moderate yield in which high molecular mass-engineered β-galactosidases had been obtained previously[207,216].

The building blocks alone tend to passively aggregate as amorphous clusters with average sizes of approximately 1μm (Article 4, Figure 3). However, the presence of DNA dramatically modifies the organization of the protein, and induces the formation of protein–DNA nanoparticles of approximately 80 nm from which DNA molecules eventually overhangs (Article 4, Figure 3 and 4). These artificial viruses, having optimal size regarding their potential interaction with mammalian cells and further uptake, are able to promote the transgene expression in targeted cultured cells, as observed by several models, again more efficiently than amorphous vehicles based on larger scaffold proteins[210,216].

Interestingly, the organization of these protein–DNA complexes is not dependent on protein–protein interactions but on the sticky, glue-like potential of DNA, that seems to show avidity for the internal α-helix conformation exhibited by tested proteins. The architectonic properties of DNA in creating regular nanoparticles, based on charge-dependent interactions[322], strongly depend on the protein–DNA ratio 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[216,320]. In this context, the particle size (80 nm) observed by using two short chimerical proteins resulted very similar to that observed when associating other short peptides with plasmid DNA[111,323]. These organizing forces are probably dependent on the ability of DNA to alter the conformation of the shell proteins, but we demonstrate that conformational changes in protein–DNA artificial viruses are mutually induced.

Irrespective of that, the nanometric organizing abilities of DNA-multifunctional protein complexes open 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 chemical hybrids and improved bionanoparticles for gene therapy but also conventional drug delivery.

Exploring RGD peptides for cell targeting delivery and therapeutic modules

The tripeptide RGD is one of the most widely used cell ligands in nanomedicine for the targeted delivery of drugs, as associated with chemicals, protein-only constructs, self-assembling peptides, liposomes, dendrimers, carbon nanotubes, magnetic nanoparticles, quantum dots, polymers, and other delivery agents[324]. RGD is also used as a cell-binding agent in cell culture and in different approaches for tissue engineering as a functional coating agent to favour cell binding to the substrate and cell differentiation[325].

In the context of multifunctional proteins, our group previously showed that multifunctional recombinant vehicles for DNA delivery containing the RGD motif efficiently induced the *in vivo* expression of a reporter and a therapeutic gene[213], followed by reduced infarct volume and functional recovery of treated animals[321], in an animal model of acute brain injury. Interestingly, when these animals were treated with the vehicle alone or with the vehicle carrying a control transgene, they showed a significant degree of neuroprotection, indicating that the effect was intrinsic to the vector. In that work it was suggested that the neuroprotective effect was mediated by the interaction between the RGD motif of the vehicle and specific integrins through an unknown mechanism.

The work presented in this thesis demonstrates that the transfection efficiency of HNRK nanoparticles is highly dependent on the amount of v 3 integrins displayed on the cell

surface (Article 5, Figure 1), indicating that the RGD stretch determines the capability of the construct to penetrate the target cell and deliver the cargo and it can also be used to target cells that present this membrane receptor, avoiding the ones that do not display them.

Therefore, the cell trophic activities of RGD described are of potential interest when evaluating the biological and therapeutic effects of RGD-based therapies, because the side effects of the vehicle might antagonize or synergistically act with those of the therapeutic cargo. This might be of special interest if the nanoparticle-linked drug is intended to arrest cell growth or is a cell-killing agent, as usually occurs in antitumoral therapy. Importantly, the catalogues of agents used in nanomedicine for cell targeting are rich in natural or mimetic ligands of receptors for hormones, cytokines, or other effectors that might potentially trigger signalling events and also act as what are named “therapeutic modules” in gene therapy vehicles, opening a broad range of possibilities.

Using “therapeutic modules” to construct gene therapy vehicles that induce a beneficial change in the tissue to be treated is a good option, because normally this change is observed before and faster than the effect produced by the genetic cargo. We believe that the further exploration and finding of new modules of this type will render a new area of investigation in tailoring artificial gene therapy vectors.

Future perspectives

Development of protein only nanoparticles for gene delivery in the area of nanobiothecnology is a highly multidisciplinary area, requiring a range of scientific knowledge, from protein identification, characterization and engineering involved in the design and production of new vehicles, through cell-tissue biology to clinical knowledge. It’s because of that that it requires collaboration between chemists, physicists, engineers and clinicians.

A better comprehension of the DNA–protein and protein–protein interactions in the context of nanoparticles for NAs 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.

Protein-based vectors with improved delivery properties and very good safety profile, being non-toxic, non-immunogenic and well tolerated are the goal of future approaches. Additionally, translating the promising results of the modular vehicle platform into medicines at the bedside of patients requires the development of methods to formulate these nanoparticles into usable drug products that can be manufactured, transported, stored, and conveniently administered to patients. Tailored formulations have to be designed for different routes of administration and for different therapeutic needs.

It is likely that no single material, modification, or method will be universally applicable. At present, there are no nanoparticles that are completely safe and nontoxic. The great flexibility inherent in the use of nanoparticle mediated gene therapy allows the selection of the best possible combination of factors for maximum effectiveness. There is still a long way to go before we can reach an optimal design of a protein-based NA delivery system.

Last but not least, further efforts should be dedicated to establish adequate production platforms for large-scale green

and economically feasible synthesis of nanocarriers with reproducible structural and functional characteristics (size, shape, chemical composition, supramolecular self-assembling, targeting and administration pathways). These processes should also allow the production of these nanomedicines under the normative of good manufacturing practices (GMPs).

Conclusions

1. R9 peptide, when fused to a reporter GFP protein, confers self-assembling properties to the whole construct, which spontaneously forms a disk-shaped, highly stable protein-based structure of approximately 20 nm in diameter.
2. R9-powered nanodisks bind and condense plasmid DNA, efficiently penetrate cultured cells and rapidly reach the nucleus, where reporter transgenes are expressed.
3. The intracellular trafficking of internalized nanodisks towards the nucleus does not compromise the particulate organization of the protein–DNA complexes or the proper folding of the individual building protein blocks.
4. The stickiness of R9 at a molecular level and the consequent architectonic properties at the nanoscale might account for the unusual efficiency of these peptides as membrane active agents and as vehicles for systemic drug and DNA delivery.
5. The multifunctional nature of R9 and in particular the architectonic properties, make it an extremely useful tool for the design and fine tuning, through rational protein engineering, of tailored nanocarriers for DNA and drug delivery under the artificial virus concept.
6. Short, mainly unordered multifunctional peptides as building blocks of protein-based artificial viruses called HNRK or HKRN have shown an excellent performance in transgene delivery under different biological models.
7. 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.
8. 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, supported by the embedded DNA.
9. The NAs act as compacting molecular glue, affecting the conformation of the protein building blocks, altering the α-helix structure of the central region of the protein, minimizing their aggregation tendency and promoting an ordered, self-organization of the complexes.

10. The self-organizing structures have sizes compatible with an efficient receptor-mediated cell uptake and proper intracellular trafficking to the cell nucleus.
11. The architectonic properties of NAs at the nanoscale open intriguing opportunities for a better rational design of artificial viruses for gene therapy regarding their molecular and physical organization.
12. The transfection efficiency of HNRK nanoparticles is highly dependent on the amount of v 3 integrins displayed on the cell surface, indicating that the RGD stretch determines the capability of the construct to penetrate the target cell and deliver the cargo.
13. When exposing PC12 cells to HNRK alone (but not to HKRN), protoneurite formation is observed. HNRK specifically stimulated PC12 cell growth at levels comparable or even higher than those promoted by the nerve growth factor (NGF).
14. HNRK up regulates microtubule-associated protein-2 (MAP2) and Tau and also stimulates the phosphorylation of ERK1/2. This effect is not restricted to a specific cell lineage.
15. HNRK acts as a cell trophic factor but with different effector pathways than NGF. This is of potential interest when evaluating the biological and therapeutic effects of RGD-based therapies. A new category of functional modules called “therapeutic modules” has to be taken into account when designing new modular protein vehicles.

Bibliography

1. ASGCT Gene Therapy Defined.

2. Commission, T. H. E., The, O. F. & Communities, E. Official Journal of the European Union. 3–12 (2009).

3. Miller, A. D. Human gene therapy comes of age. *Nature* **357**, 455–60 (1992).

4. J., L. *Biological future of man*. *Man and His Future* 265 (London: Churchill, 1963).

5. Rogers, S., Lowenthal, A., Terheggen, H. G. & Columbo, J. P. Induction of arginase activity with the Shope papilloma virus in tissue culture cells from an argininemic patient. *The Journal of experimental medicine* **137**, 1091–6 (1973).

6. Marcum, J. A. From the molecular genetics revolution to gene therapy: Translating basic research into medicine. *Journal of Laboratory and Clinical Medicine* **146**, 312–316 (2005).

7. Friedmann, T. Stanfield Rogers: insights into virus vectors and failure of an early gene therapy model. *Molecular therapy : the journal of the American Society of Gene Therapy* **4**, 285–8 (2001).

8. Terheggen, H. G., Lowenthal, A., Lavinha, F., Colombo, J. P. & Rogers, S. Unsuccessful trial of gene replacement in arginase deficiency. *Zeitschrift für Kinderheilkunde* **119**, 1–3 (1975).

9. Blaese, R. M. *et al.* Treatment of severe combined immunodeficiency disease (SCID) due to adenosine deaminase deficiency with CD34+ selected autologous peripheral blood cells transduced with a human ADA gene. Amendment to clinical research project, Project 90-C-195, January 10. *Human gene therapy* **4**, 521–7 (1993).

10. Fischer, A. & Cavazzana-Calvo, M. Integration of retroviruses: a fine balance between efficiency and danger. *PLoS medicine* **2**, e10 (2005).

11. Fischer, A., Hacein-Bey-Abina, S. & Cavazzana-Calvo, M. Gene therapy for immunodeficiency diseases. *Seminars in hematology* **41**, 272–8 (2004).

12. Hacein-Bey-Abina, S. *et al.* A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *New England Journal of Medicine* **348**, 255–256 (2003).

13. Cavazzana-Calvo, M., Thrasher, A. & Mavilio, F. The future of gene therapy. *Nature* **427**, 779–81 (2004).

14. Peng, Z. Current status of gendicine in China: recombinant human Ad-p53 agent for treatment of cancers. *Human gene therapy* **16**, 1016–27 (2005).

15. Pearson, S., Jia, H. & Kandachi, K. China approves first gene therapy. *Nature biotechnology* **22**, 3–4 (2004).

16. Maguire, A. M. *et al.* Safety and efficacy of gene transfer for Leber's congenital amaurosis. *The New England journal of medicine* **358**, 2240–8 (2008).

17. Simonelli, F. *et al.* Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. *Molecular therapy : the journal of the American Society of Gene Therapy* **18**, 643–50 (2010).

18. Cideciyan, A. V *et al.* Vision 1 Year after Gene Therapy for Leber's Congenital Amaurosis. *New England Journal of Medicine* **361**, 725–727 (2009).

19. Bainbridge, J. W. B. *et al.* Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis. *New England Journal of Medicine* **358**, 2231–2239 (2008).

20. Fischer, A., Hacein-Bey-Abina, S. & Cavazzana-Calvo, M. 20 years of gene therapy for SCID. *Nature immunology* **11**, 457–60 (2010).

21. Ferrua, F., Brigida, I. & Aiuti, A. Update on gene therapy for adenosine deaminase-deficient severe combined immunodeficiency. *Current opinion in allergy and clinical immunology* **10**, 551–6 (2010).

22. Cartier, N. & Aubourg, P. Hematopoietic stem cell transplantation and hematopoietic stem cell gene therapy in X-linked adrenoleukodystrophy. *Brain pathology (Zurich, Switzerland)* **20**, 857–62 (2010).

23. Ledford, H. Cell therapy fights leukaemia. *Nature* (2011).doi:10.1038/news.2011.472

24. LeWitt, P. A. *et al.* AAV2-GAD gene therapy for advanced Parkinson's disease: a double-blind, sham-surgery controlled, randomised trial. *Lancet neurology* **10**, 309–19 (2011).

25. Gene therapy deserves a fresh chance. *Nature* **461**, 1173 (2009).

26. Grisham, J. Launch of Stem Cell Therapy Trial Offers Hope for Patients with Inherited Blood Disorder. *Memorial Sloan-Kettering Cancer Center* (2012).

27. Center, M. S.-K. C. ß-Thalassemia Major With Autologous CD34+ Hematopoietic Progenitor Cells Transduced With TNS9.3.55 a Lentiviral Vector Encoding the Normal Human ß-Globin Gene. (2012).

28. POLLACK, A. European Agency Recommends Approval of a Gene Therapy - NYTimes.com. *The New York Times* B1 (2012).

29. Richards, S. Gene Therapy Arrives in Europe. *The Scientist* (2012).

30. Trials and tribulations. *Nature genetics* **24**, 201–2 (2000).

31. Weiss, R. & Nelson, D. Teen Dies Undergoing Experimental Gene Therapy. *The Washington Post* A1 (1999).

32. Lundstrom, K. Latest development in viral vectors for gene therapy. *Trends in biotechnology* **21**, 117–22 (2003).

33. Limberis, M. P. Phoenix rising: gene therapy makes a comeback. *Acta biochimica et biophysica Sinica* **44**, 632–40 (2012).

34. Collins, M. K. & Cerundolo, V. Gene therapy meets vaccine development. *Trends in biotechnology* **22**, 623–6 (2004).

35. Ferraro, B. *et al.* Clinical applications of DNA vaccines: current progress. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **53**, 296–302 (2011).

36. Jimenez-Sanchez, G., Childs, B. & Valle, D. Human disease genes. *Nature* **409**, 853–5 (2001).

37. Rakoczy, P. E. Antisense DNA technology. *Methods in molecular medicine* **47**, 89–104 (2001).

38. Whitehead, K. A., Langer, R. & Anderson, D. G. Knocking down barriers: advances in siRNA delivery. *Nature reviews. Drug discovery* **8**, 129–38 (2009).

39. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–8 (2001).

40. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806–11 (1998).

41. Bauman, J., Jearawiriyapaisarn, N. & Kole, R. Therapeutic potential of splice-switching oligonucleotides. *Oligonucleotides* **19**, 1–13 (2009).

42. Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457–63 (2010).

43. Sazani, P. & Kole, R. Therapeutic potential of antisense oligonucleotides as modulators of alternative splicing. *The Journal of clinical investigation* **112**, 481–6 (2003).

44. Ivanova, G. D. *et al.* Improved cell-penetrating peptide-PNA conjugates for splicing redirection in HeLa cells and exon skipping in mdx mouse muscle. *Nucleic acids research* **36**, 6418–28 (2008).

45. Lehto, T., Ezzat, K. & Langel, U. Peptide nanoparticles for oligonucleotide delivery. *Progress in molecular biology and translational science* **104**, 397–426 (2011).

46. Senior, J. H. Fate and behavior of liposomes in vivo: a review of controlling factors. *Critical reviews in therapeutic drug carrier systems* **3**, 123–93 (1987).

47. Takakura, Y. & Hashida, M. Macromolecular carrier systems for targeted drug delivery: pharmacokinetic considerations on biodistribution. *Pharmaceutical research* **13**, 820–31 (1996).

48. Taniguchi, N. *et al.* Remodeling of cell surface glycoproteins by N-acetylglucosaminyltransferase III gene transfection: modulation of metastatic potentials and down regulation of hepatitis B virus replication. *Glycobiology* **6**, 691–4 (1996).

49. Heiser, W. C. Gene delivery to skin using biolistics. *CSH protocols* **2006**, (2006).

50. Venkataraman, S. *et al.* The effects of polymeric nanostructure shape on drug delivery. *Advanced drug delivery reviews* **63**, 1228–46 (2011).

51. Schroeder, A. *et al.* Treating metastatic cancer with nanotechnology. *Nature reviews. Cancer* **12**, 39–50 (2012).

52. Li, S., Rizzo, M. A., Bhattacharya, S. & Huang, L. Characterization of cationic lipid-protamine-DNA (LPD) complexes for intravenous gene delivery. *Gene therapy* **5**, 930–7 (1998).

53. Li, S. *et al.* Dynamic changes in the characteristics of cationic lipidic vectors after exposure to mouse serum: implications for intravenous lipofection. *Gene therapy* **6**, 585–94 (1999).

54. Litzinger, D. C. *et al.* Fate of cationic liposomes and their complex with oligonucleotide in vivo. *Biochimica et biophysica acta* **1281**, 139–49 (1996).

55. Mahato, R. I., Kawabata, K., Takakura, Y. & Hashida, M. In vivo disposition characteristics of plasmid DNA complexed with cationic liposomes. *Journal of drug targeting* **3**, 149–57 (1995).

56. Mahato, R. I., Kawabata, K., Nomura, T., Takakura, Y. & Hashida, M. Physicochemical and pharmacokinetic characteristics of plasmid DNA/cationic liposome complexes. *Journal of pharmaceutical sciences* **84**, 1267–71 (1995).

57. Nishikawa, M., Takemura, S., Takakura, Y. & Hashida, M. Targeted delivery of plasmid DNA to hepatocytes in vivo: optimization of the pharmacokinetics of plasmid DNA/galactosylated poly(L-lysine) complexes by controlling their physicochemical properties. *The Journal of pharmacology and experimental therapeutics* **287**, 408–15 (1998).

58. Finsinger, D., Remy, J. S., Erbacher, P., Koch, C. & Plank, C. Protective copolymers for nonviral gene vectors: synthesis, vector characterization and application in gene delivery. *Gene therapy* **7**, 1183–92 (2000).

59. Krieg, A. M. *et al.* CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* **374**, 546–9 (1995).

60. McMahon, J. M., Wells, K. E., Bamfo, J. E., Cartwright, M. A. & Wells, D. J. Inflammatory responses following direct injection of plasmid DNA into skeletal muscle. *Gene therapy* **5**, 1283–90 (1998).

61. Li, S. *et al.* Effect of immune response on gene transfer to the lung via systemic administration of cationic lipidic vectors. *The American journal of physiology* **276**, L796–804 (1999).

62. Yew, N. S. *et al.* Contribution of plasmid DNA to inflammation in the lung after administration of cationic lipid:pDNA complexes. *Human gene therapy* **10**, 223–34 (1999).

63. Tousignant, J. D. *et al.* Comprehensive analysis of the acute toxicities induced by systemic administration of cationic lipid:plasmid DNA complexes in mice. *Human gene therapy* **11**, 2493–513 (2000).

64. Liu, Y. *et al.* Factors influencing the efficiency of cationic liposome-mediated intravenous gene delivery. *Nature biotechnology* **15**, 167–73 (1997).

65. BENNETT, H. S., LUFT, J. H. & HAMPTON, J. C. Morphological classifications of vertebrate blood capillaries. *The American journal of physiology* **196**, 381–90 (1959).

66. Budker, V., Zhang, G., Knechtle, S. & Wolff, J. A. Naked DNA delivered intraportally expresses efficiently in hepatocytes. *Gene therapy* **3**, 593–8 (1996).

67. Davis, H. L., Whalen, R. G. & Demeneix, B. A. Direct gene transfer into skeletal muscle in vivo: factors affecting efficiency of transfer and stability of expression. *Human gene therapy* **4**, 151–9 (1993).

68. Greelish, J. P. *et al.* Stable restoration of the sarcoglycan complex in dystrophic muscle perfused with histamine and a recombinant adeno-associated viral vector. *Nature medicine* **5**, 439–43 (1999).

69. Rainov, N. G. *et al.* Intraarterial delivery of adenovirus vectors and liposome-DNA complexes to experimental brain neoplasms. *Human gene therapy* **10**, 311–8 (1999).

70. Fuki, I. V *et al.* The syndecan family of proteoglycans. Novel receptors mediating internalization of atherogenic lipoproteins in vitro. *The Journal of clinical investigation* **100**, 1611–22 (1997).

71. Pratten, M. K., Cable, H. C., Ringsdorf, H. & Lloyd, J. B. Adsorptive pinocytosis of polycationic copolymers of vinylpyrrolidone with vinylamine by rat yolk sac and rat peritoneal macrophage. *Biochimica et biophysica acta* **719**, 424–30 (1982).

72. Ghinea, N. & Hasu, M. Charge effect on binding, uptake and transport of ferritin through fenestrated endothelium. *Journal of submicroscopic cytology* **18**, 647–59 (1986).

73. Stenseth, K., Hedin, U. & Thyberg, J. Endocytosis, intracellular transport, and turnover of anionic and cationic proteins in cultured mouse peritoneal macrophages. *European journal of cell biology* **31**, 15–25 (1983).

74. Zuhorn, I. S., Kalicharan, R. & Hoekstra, D. Lipoplex-mediated transfection of mammalian cells occurs through the cholesterol-dependent clathrin-mediated pathway of endocytosis. *The Journal of biological chemistry* **277**, 18021–8 (2002).

75. Friend, D. S., Papahadjopoulos, D. & Debs, R. J. Endocytosis and intracellular processing accompanying transfection mediated by cationic liposomes. *Biochimica et biophysica acta* **1278**, 41–50 (1996).

76. Matsui, H., Johnson, L. G., Randell, S. H. & Boucher, R. C. Loss of binding and entry of liposome-DNA complexes decreases transfection efficiency in differentiated airway epithelial cells. *The Journal of biological chemistry* **272**, 1117–26 (1997).

77. Labat-Moleur, F. *et al.* An electron microscopy study into the mechanism of gene transfer with lipopolyamines. *Gene therapy* **3**, 1010–7 (1996).

78. Harbottle, R. P. *et al.* An RGD-oligolysine peptide: a prototype construct for integrin-mediated gene delivery. *Human gene therapy* **9**, 1037–47 (1998).

79. Gottschalk, S., Cristiano, R. J., Smith, L. C. & Woo, S. L. Folate receptor mediated DNA delivery into tumor cells: potosomal disruption results in enhanced gene expression. *Gene therapy* **1**, 185–91 (1994).

80. Hofland, H. E. J. *et al.* Folate-targeted gene transfer in vivo. *Molecular therapy : the journal of the American Society of Gene Therapy* **5**, 739–44 (2002).

81. Turk, M. J., Reddy, J. A., Chmielewski, J. A. & Low, P. S. Characterization of a novel pH-sensitive peptide that enhances drug release from folate-targeted liposomes at endosomal pHs. *Biochimica et biophysica acta* **1559**, 56–68 (2002).

82. Francis, C. L., Ryan, T. A., Jones, B. D., Smith, S. J. & Falkow, S. Ruffles induced by Salmonella and other stimuli direct macropinocytosis of bacteria. *Nature* **364**, 639–42 (1993).

83. Kao, G. Y., Change, L. J. & Allen, T. M. Use of targeted cationic liposomes in enhanced DNA delivery to cancer cells. *Cancer gene therapy* **3**, 250–6

84. Ferrer-Miralles, N., Vázquez, E. & Villaverde, A. Membrane-active peptides for non-viral gene therapy: making the safest easier. *Trends in biotechnology* **26**, 267–75 (2008).

85. Vázquez, E., Ferrer-Miralles, N. & Villaverde, A. Peptide-assisted traffic engineering for nonviral gene therapy. *Drug discovery today* **13**, 1067–74 (2008).

86. Lundberg, M., Wikström, S. & Johansson, M. Cell surface adherence and endocytosis of protein transduction domains. *Molecular therapy : the journal of the American Society of Gene Therapy* **8**, 143–50 (2003).

87. Little, S. R. & Kohane, D. S. Polymers for intracellular delivery of nucleic acids. *Journal of Materials Chemistry* **18**, 832 (2008).

88. Varkouhi, A. K., Scholte, M., Storm, G. & Haisma, H. J. Endosomal escape pathways for delivery of biologicals. *Journal of controlled release : official journal of the Controlled Release Society* **151**, 220–8 (2011).

89. Vázquez, E. *et al.* Modular protein engineering in emerging cancer therapies. *Current pharmaceutical design* **15**, 893–916 (2009).

90. Martin, M. E. & Rice, K. G. Peptide-guided gene delivery. *The AAPS journal* **9**, E18–29 (2007).

91. Gottschalk, S. *et al.* A novel DNA-peptide complex for efficient gene transfer and expression in mammalian cells. *Gene therapy* **3**, 448–57 (1996).

92. Brooks, H., Lebleu, B. & Vivès, E. Tat peptide-mediated cellular delivery: back to basics. *Advanced drug delivery reviews* **57**, 559–77 (2005).

93. Vivès, E., Brodin, P. & Lebleu, B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *The Journal of biological chemistry* **272**, 16010–7 (1997).

94. Pooga, M., Hällbrink, M., Zorko, M. & Langel, U. Cell penetration by transportan. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **12**, 67–77 (1998).

95. Pooga, M. *et al.* Cellular translocation of proteins by transportan. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **15**, 1451–3 (2001).

96. Wagner, E., Plank, C., Zatloukal, K., Cotten, M. & Birnstiel, M. L. Influenza virus hemagglutinin HA-2 N-terminal fusogenic peptides augment gene transfer by transferrin-polylysine-DNA

complexes: toward a synthetic virus-like gene-transfer vehicle. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 7934–8 (1992).

97. Plank, C., Oberhauser, B., Mechtler, K., Koch, C. & Wagner, E. The influence of endosome-disruptive peptides on gene transfer using synthetic virus-like gene transfer systems. *The Journal of biological chemistry* **269**, 12918–24 (1994).

98. Subbarao, N. K., Parente, R. A., Szoka, F. C., Nadasdi, L. & Pongracz, K. pH-dependent bilayer destabilization by an amphipathic peptide. *Biochemistry* **26**, 2964–72 (1987).

99. Parente, R. A., Nadasdi, L., Subbarao, N. K. & Szoka, F. C. Association of a pH-sensitive peptide with membrane vesicles: role of amino acid sequence. *Biochemistry* **29**, 8713–9 (1990).

100. Parente, R. A., Nir, S. & Szoka, F. C. Mechanism of leakage of phospholipid vesicle contents induced by the peptide GALA. *Biochemistry* **29**, 8720–8 (1990).

101. Li, W., Nicol, F. & Szoka, F. C. GALA: a designed synthetic pH-responsive amphipathic peptide with applications in drug and gene delivery. *Advanced drug delivery reviews* **56**, 967–85 (2004).

102. Wyman, T. B. *et al.* Design, synthesis, and characterization of a cationic peptide that binds to nucleic acids and permeabilizes bilayers. *Biochemistry* **36**, 3008–17 (1997).

103. Rittner, K. *et al.* New basic membrane-destabilizing peptides for plasmid-based gene delivery in vitro and in vivo. *Molecular therapy : the journal of the American Society of Gene Therapy* **5**, 104–14 (2002).

104. Ogris, M., Carlisle, R. C., Bettinger, T. & Seymour, L. W. Melittin enables efficient vesicular escape and enhanced nuclear access of nonviral gene delivery vectors. *The Journal of biological chemistry* **276**, 47550–5 (2001).

105. Derossi, D., Joliot, A. H., Chassaing, G. & Prochiantz, A. The third helix of the Antennapedia homeodomain translocates through biological membranes. *The Journal of biological chemistry* **269**, 10444–50 (1994).

106. Derossi, D., Chassaing, G. & Prochiantz, A. Trojan peptides: the penetratin system for intracellular delivery. *Trends in cell biology* **8**, 84–7 (1998).

107. Pichon, C., Gonçalves, C. & Midoux, P. Histidine-rich peptides and polymers for nucleic acids delivery. *Advanced drug delivery reviews* **53**, 75–94 (2001).

108. Plank, C., Zauner, W. & Wagner, E. Application of membrane-active peptides for drug and gene delivery across cellular membranes. *Advanced drug delivery reviews* **34**, 21–35 (1998).

109. McKenzie, D. L., Smiley, E., Kwok, K. Y. & Rice, K. G. Low molecular weight disulfide cross-linking peptides as nonviral gene delivery carriers. *Bioconjugate chemistry* **11**, 901–9 (2000).

110. Midoux, P., Kichler, a, Boutin, V., Maurizot, J. C. & Monsigny, M. Membrane permeabilization and efficient gene transfer by a peptide containing several histidines. *Bioconjugate chemistry* **9**, 260–7 (1998).

111. Midoux, P., LeCam, E., Coulaud, D., Delain, E. & Pichon, C. Histidine containing peptides and polypeptides as nucleic acid vectors. *Somatic cell and molecular genetics* **27**, 27–47 (2002).

112. Kichler, A., Leborgne, C., Danos, O. & Bechinger, B. Characterization of the gene transfer process mediated by histidine-rich peptides. *Journal of molecular medicine (Berlin, Germany)* **85**, 191–201 (2007).

113. Lechardeur, D. *et al.* Metabolic instability of plasmid DNA in the cytosol: a potential barrier to gene transfer. *Gene therapy* **6**, 482–97 (1999).

114. Melchior, F. & Gerace, L. Mechanisms of nuclear protein import. *Current opinion in cell biology* **7**, 310–8 (1995).

115. Glover, D. J., Ng, S. M., Mechler, A., Martin, L. L. & Jans, D. a Multifunctional protein nanocarriers for targeted nuclear gene delivery in nondividing cells. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **23**, 2996–3006 (2009).

116. Zanta, M. A., Belguise-Valladier, P. & Behr, J. P. Gene delivery: a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 91–6 (1999).

117. Sorgi, F. L., Bhattacharya, S. & Huang, L. Protamine sulfate enhances lipid-mediated gene transfer. *Gene therapy* **4**, 961–8 (1997).

118. Rudolph, C. *et al.* Oligomers of the arginine-rich motif of the HIV-1 TAT protein are capable of transferring plasmid DNA into cells. *The Journal of biological chemistry* **278**, 11411–8 (2003).

119. Fischer, N., Kremmer, E., Lautscham, G., Mueller-Lantzsch, N. & Grässer, F. A. Epstein-Barr virus nuclear antigen 1 forms a complex with the nuclear transporter karyopherin alpha2. *The Journal of biological chemistry* **272**, 3999–4005 (1997).

120. Dang, C. V & Lee, W. M. Identification of the human c-myc protein nuclear translocation signal. *Molecular and cellular biology* **8**, 4048–54 (1988).

121. Fominaya, J., Uherek, C. & Wels, W. A chimeric fusion protein containing transforming growth factor-alpha mediates gene transfer via binding to the EGF receptor. *Gene therapy* **5**, 521–30 (1998).

122. Bäuerle, M., Doenecke, D. & Albig, W. The requirement of H1 histones for a heterodimeric nuclear import receptor. *The Journal of biological chemistry* **277**, 32480–9 (2002).

123. Baake, M., Bäuerle, M., Doenecke, D. & Albig, W. Core histones and linker histones are imported into the nucleus by different pathways. *European journal of cell biology* **80**, 669–77 (2001).

124. Subramanian, A., Ranganathan, P. & Diamond, S. L. Nuclear targeting peptide scaffolds for lipofection of nondividing mammalian cells. *Nature biotechnology* **17**, 873–7 (1999).

125. Gharakhanian, E., Takahashi, J. & Kasamatsu, H. The carboxyl 35 amino acids of SV40 Vp3 are essential for its nuclear accumulation. *Virology* **157**, 440–8 (1987).

126. Lyons, R. H., Ferguson, B. Q. & Rosenberg, M. Pentapeptide nuclear localization signal in adenovirus E1a. *Molecular and cellular biology* **7**, 2451–6 (1987).

127. Kleinschmidt, J. A. & Seiter, A. Identification of domains involved in nuclear uptake and histone binding of protein N1 of Xenopus laevis. *The EMBO journal* **7**, 1605–14 (1988).

128. Kiefer, P., Acland, P., Pappin, D., Peters, G. & Dickson, C. Competition between nuclear localization and secretory signals determines the subcellular fate of a single CUG-initiated form of FGF3. *The EMBO journal* **13**, 4126–36 (1994).

129. Schreiber, V., Molinete, M., Boeuf, H., De Murcia, G. & Ménissier-de Murcia, J. The human poly(ADP-ribose) polymerase nuclear localization signal is a bipartite element functionally separate from DNA binding and catalytic activity. *The EMBO journal* **11**, 3263–9 (1992).

130. Robbins, J., Dilworth, S. M., Laskey, R. A. & Dingwall, C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615–23 (1991).

131. Jiang, W., Kim, B. Y. S., Rutka, J. T. & Chan, W. C. W. Nanoparticle-mediated cellular response is size-dependent. *Nature nanotechnology* **3**, 145–50 (2008).

132. Medina-Kauwe, L. K., Xie, J. & Hamm-Alvarez, S. Intracellular trafficking of nonviral vectors. *Gene therapy* **12**, 1734–51 (2005).

133. Marianayagam, N. J., Sunde, M. & Matthews, J. M. The power of two: protein dimerization in biology. *Trends in biochemical sciences* **29**, 618–25 (2004).

134. Tu, R. S. & Tirrell, M. Bottom-up design of biomimetic assemblies. *Advanced drug delivery reviews* **56**, 1537–63 (2004).

135. Leckband, D. Measuring the forces that control protein interactions. *Annual review of biophysics and biomolecular structure* **29**, 1–26 (2000).

136. Meyer, E. E., Rosenberg, K. J. & Israelachvili, J. Recent progress in understanding hydrophobic interactions. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15739–46 (2006).

137. Domingo-Espín, J. *et al.* Nanoparticulate architecture of protein-based artificial viruses is supported by protein – DNA interactions R eseaRch a Rticle. *Nanomedicine London England* **6**, 1047–1061 (2011).

138. Tsumoto, K., Luckel, F. & Yoshikawa, K. Giant DNA molecules exhibit on/off switching of transcriptional activity through conformational transition. *Biophysical chemistry* **106**, 23–9 (2003).

139. Wolff, J. A. *et al.* Direct gene transfer into mouse muscle in vivo. *Science (New York, N.Y.)* **247**, 1465–8 (1990).

140. Böhm, W., Mertens, T., Schirmbeck, R. & Reimann, J. Routes of plasmid DNA vaccination that prime murine humoral and cellular immune responses. *Vaccine* **16**, 949–54

141. Suzuki, T., Shin, B. C., Fujikura, K., Matsuzaki, T. & Takata, K. Direct gene transfer into rat liver cells by in vivo electroporation. *FEBS letters* **425**, 436–40 (1998).

142. Nishi, T. *et al.* High-efficiency in vivo gene transfer using intraarterial plasmid DNA injection following in vivo electroporation. *Cancer research* **56**, 1050–5 (1996).

143. Zhang, L., Li, L., Hoffmann, G. A. & Hoffman, R. M. Depth-targeted efficient gene delivery and expression in the skin by pulsed electric fields: an approach to gene therapy of skin aging and other diseases. *Biochemical and biophysical research communications* **220**, 633–6 (1996).

144. Titomirov, A. V, Sukharev, S. & Kistanova, E. In vivo electroporation and stable transformation of skin cells of newborn mice by plasmid DNA. *Biochimica et biophysica acta* **1088**, 131–4 (1991).

145. Rols, M. P. *et al.* In vivo electrically mediated protein and gene transfer in murine melanoma. *Nature biotechnology* **16**, 168–71 (1998).

146. Mathiesen, I. Electropermeabilization of skeletal muscle enhances gene transfer in vivo. *Gene therapy* **6**, 508–14 (1999).

147. Pohl, P., Rosenfeld, E. & Millner, R. Effects of ultrasound on the steady-state transmembrane pH gradient and the permeability of acetic acid through bilayer lipid membranes. *Biochimica et biophysica acta* **1145**, 279–83 (1993).

148. Newman, C. M., Lawrie, A., Brisken, A. F. & Cumberland, D. C. Ultrasound gene therapy: on the road from concept to reality. *Echocardiography (Mount Kisco, N.Y.)* **18**, 339–47 (2001).

149. Fechheimer, M. *et al.* Transfection of mammalian cells with plasmid DNA by scrape loading and sonication loading. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 8463–7 (1987).

150. Kim, H. J., Greenleaf, J. F., Kinnick, R. R., Bronk, J. T. & Bolander, M. E. Ultrasound-mediated transfection of mammalian cells. *Human gene therapy* **7**, 1339–46 (1996).

151. Klein, R. M., Wolf, E. D., Wu, R. & Sanford, J. C. High-velocity microprojectiles for delivering nucleic acids into living cells. 1987. *Biotechnology (Reading, Mass.)* **24**, 384–6 (1992).

152. Fynan, E. F. *et al.* DNA vaccines: protective immunizations by parenteral, mucosal, and gene-gun inoculations. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 11478–82 (1993).

153. Yoshida, A., Nagata, T., Uchijima, M., Higashi, T. & Koide, Y. Advantage of gene gun-mediated over intramuscular inoculation of plasmid DNA vaccine in reproducible induction of specific immune responses. *Vaccine* **18**, 1725–9 (2000).

154. Lübbe, A. S. *et al.* Clinical experiences with magnetic drug targeting: a phase I study with 4'-epidoxorubicin in 14 patients with advanced solid tumors. *Cancer research* **56**, 4686–93 (1996).

155. Scherer, F., Anton, M. & Schillinger, U. Magnetofection: enhancing and targeting gene delivery by magnetic force in vitro and in vivo. *Gene …* 102–109 (2002).doi:10.1038/sj/gt/3301624

156. Felgner, P. L. *et al.* Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7413–7 (1987).

157. Zabner, J., Fasbender, A. J., Moninger, T., Poellinger, K. A. & Welsh, M. J. Cellular and molecular barriers to gene transfer by a cationic lipid. *The Journal of biological chemistry* **270**, 18997–9007 (1995).

158. Sternberg, B., Sorgi, F. L. & Huang, L. New structures in complex formation between DNA and cationic liposomes visualized by freeze-fracture electron microscopy. *FEBS letters* **356**, 361–6 (1994).

159. Gershon, H., Ghirlando, R., Guttman, S. B. & Minsky, A. Mode of formation and structural features of DNA-cationic liposome complexes used for transfection. *Biochemistry* **32**, 7143–51 (1993).

160. Legendre, J. Y. & Szoka, F. C. Delivery of plasmid DNA into mammalian cell lines using pH-sensitive liposomes: comparison with cationic liposomes. *Pharmaceutical research* **9**, 1235–42 (1992).

161. Felgner, J. H. *et al.* Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations. *The Journal of biological chemistry* **269**, 2550–61 (1994).

162. Düzgüneş, N., Goldstein, J. A., Friend, D. S. & Felgner, P. L. Fusion of liposomes containing a novel cationic lipid, N-[2,3-(dioleyloxy)propyl]-N,N,N-trimethylammonium: induction by multivalent anions and asymmetric fusion with acidic phospholipid vesicles. *Biochemistry* **28**, 9179–84 (1989).

163. Van der Woude, I. *et al.* Parameters influencing the introduction of plasmid DNA into cells by the use of synthetic amphiphiles as a carrier system. *Biochimica et biophysica acta* **1240**, 34–40 (1995).

164. Zelphati, O. & Szoka, F. C. Mechanism of oligonucleotide release from cationic liposomes. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 11493–8 (1996).

165. Mönkkönen, J. & Urtti, A. Lipid fusion in oligonucleotide and gene delivery with cationic lipids. *Advanced drug delivery reviews* **34**, 37–49 (1998).

166. Urtti, A., Polansky, J., Lui, G. M. & Szoka, F. C. Gene delivery and expression in human retinal pigment epithelial cells: effects of synthetic carriers, serum, extracellular matrix and viral promoters. *Journal of drug targeting* **7**, 413–21 (2000).

167. Zelphati, O. *et al.* Stable and monodisperse lipoplex formulations for gene delivery. *Gene therapy* **5**, 1272–82 (1998).

168. Belting, M. & Petersson, P. Intracellular accumulation of secreted proteoglycans inhibits cationic lipid-mediated gene transfer. Co-transfer of glycosaminoglycans to the nucleus. *The Journal of biological chemistry* **274**, 19375–82 (1999).

169. Turunen, M. P. *et al.* Efficient adventitial gene delivery to rabbit carotid artery with cationic polymer-plasmid complexes. *Gene therapy* **6**, 6–11 (1999).

170. Ross, P. C. & Hui, S. W. Lipoplex size is a major determinant of in vitro lipofection efficiency. *Gene therapy* **6**, 651–9 (1999).

171. Horiguchi, Y., Larchian, W. A., Kaplinsky, R., Fair, W. R. & Heston, W. D. Intravesical liposome-mediated interleukin-2 gene therapy in orthotopic murine bladder cancer model. *Gene therapy* **7**, 844–51 (2000).

172. Domashenko, A., Gupta, S. & Cotsarelis, G. Efficient delivery of transgenes to human hair follicle progenitor cells using topical lipoplex. *Nature biotechnology* **18**, 420–3 (2000).

173. Wu, G. Y. & Wu, C. H. Receptor-mediated in vitro gene transformation by a soluble DNA carrier system. *The Journal of biological chemistry* **262**, 4429–32 (1987).

174. Boussif, O. *et al.* A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 7297–301 (1995).

175. Cherng, J. Y., Talsma, H., Crommelin, D. J. & Hennink, W. E. Long term stability of poly((2-dimethylamino)ethyl methacrylate)-based gene delivery systems. *Pharmaceutical research* **16**, 1417–23 (1999).

176. Oupický, D., Konák, C., Ulbrich, K., Wolfert, M. A. & Seymour, L. W. DNA delivery systems based on complexes of DNA with synthetic polycations and their copolymers. *Journal of controlled release : official journal of the Controlled Release Society* **65**, 149–71 (2000).

177. Van de Wetering, P., Schuurmans-Nieuwenbroek, N. M., Van Steenbergen, M. J., Crommelin, D. J. & Hennink, W. E. Copolymers of 2-(dimethylamino)ethyl methacrylate with ethoxytriethylene glycol methacrylate or N-vinyl-pyrrolidone as gene transfer agents. *Journal of controlled release : official journal of the Controlled Release Society* **64**, 193–203 (2000).

178. Ukawa, M. *et al.* 2-Methacryloyloxyethyl phosphorylcholine polymer (MPC)-coating improves the transfection activity of GALA-modified lipid nanoparticles by assisting the cellular uptake and intracellular dissociation of plasmid DNA in primary hepatocytes. *Biomaterials* **31**, 6355–62 (2010).

179. Fattal, E. & Bochot, A. State of the art and perspectives for the delivery of antisense oligonucleotides and siRNA by polymeric nanocarriers. *International journal of pharmaceutics* **364**, 237–48 (2008).

180. Al-Dosari, M. S. & Gao, X. Nonviral gene delivery: principle, limitations, and recent progress. *The AAPS journal* **11**, 671–81 (2009).

181. Cherng, J. Y., Talsma, H., Verrijk, R., Crommelin, D. J. & Hennink, W. E. The effect of formulation parameters on the size of poly-((2-dimethylamino)ethyl methacrylate)-plasmid complexes. *European journal of pharmaceutics and biopharmaceutics : official journal of Arbeitsgemeinschaft für Pharmazeutische Verfahrenstechnik e.V* **47**, 215–24 (1999).

182. Tang, M. X. & Szoka, F. C. The influence of polymer structure on the interactions of cationic polymers with DNA and morphology of the resulting complexes. *Gene therapy* **4**, 823–32 (1997).

183. Zhao, Q.-Q. *et al.* N/P ratio significantly influences the transfection efficiency and cytotoxicity of a polyethylenimine/chitosan/DNA complex. *Biological & pharmaceutical bulletin* **32**, 706–10 (2009).

184. Mairhofer, J. & Grabherr, R. Rational vector design for efficient non-viral gene delivery: challenges facing the use of plasmid DNA. *Molecular biotechnology* **39**, 97–104 (2008).

185. Winblade, N. D., Nikolic, I. D., Hoffman, A. S. & Hubbell, J. A. Blocking adhesion to cell and tissue surfaces by the chemisorption of a poly-L-lysine-graft-(poly(ethylene glycol); phenylboronic acid) copolymer. *Biomacromolecules* **1**, 523–33 (2000).

186. Park, S. & Healy, K. E. Nanoparticulate DNA packaging using terpolymers of poly(lysine-g-(lactide-b-ethylene glycol)). *Bioconjugate chemistry* **14**, 311–9 (2003).

187. Morille, M., Passirani, C., Vonarbourg, A., Clavreul, A. & Benoit, J.-P. Progress in developing cationic vectors for non-viral systemic gene therapy against cancer. *Biomaterials* **29**, 3477–96 (2008).

188. Godbey, W. T., Wu, K. K. & Mikos, A. G. Poly(ethylenimine) and its role in gene delivery. *Journal of controlled release : official journal of the Controlled Release Society* **60**, 149–60 (1999).

189. Mintzer, M. A. & Simanek, E. E. Nonviral vectors for gene delivery. *Chemical reviews* **109**, 259–302 (2009).

190. Swami, A. *et al.* A unique and highly efficient non-viral DNA/siRNA delivery system based on PEI-bisepoxide nanoparticles. *Biochemical and biophysical research communications* **362**, 835–41 (2007).

191. Boeckle, S. *et al.* Purification of polyethylenimine polyplexes highlights the role of free polycations in gene transfer. *The journal of gene medicine* **6**, 1102–11 (2004).

192. Saul, J. M., Linnes, M. P., Ratner, B. D., Giachelli, C. M. & Pun, S. H. Delivery of non-viral gene carriers from sphere-templated fibrin scaffolds for sustained transgene expression. *Biomaterials* **28**, 4705–16 (2007).

193. Boas, U. & Heegaard, P. M. H. Dendrimers in drug research. *Chemical Society reviews* **33**, 43–63 (2004).

194. Choi, J. S. *et al.* Enhanced transfection efficiency of PAMAM dendrimer by surface modification with L-arginine. *Journal of controlled release : official journal of the Controlled Release Society* **99**, 445–56 (2004).

195. Lee, J. H. *et al.* Polyplexes assembled with internally quaternized PAMAM-OH dendrimer and plasmid DNA have a neutral surface and gene delivery potency. *Bioconjugate chemistry* **14**, 1214–21 (2003).

196. Nam, H. Y. *et al.* Evaluation of generations 2, 3 and 4 arginine modified PAMAM dendrimers for gene delivery. *International journal of pharmaceutics* **363**, 199–205 (2008).

197. Majoros, I. J., Williams, C. R. & Baker, J. R. Current dendrimer applications in cancer diagnosis and therapy. *Current topics in medicinal chemistry* **8**, 1165–79 (2008).

198. Hussain, M. *et al.* A novel anionic dendrimer for improved cellular delivery of antisense oligonucleotides. *Journal of controlled release : official journal of the Controlled Release Society* **99**, 139–55 (2004).

199. Mao, S., Sun, W. & Kissel, T. Chitosan-based formulations for delivery of DNA and siRNA. *Advanced drug delivery reviews* **62**, 12–27 (2010).

200. Mumper RJ, Wang J, JM, C. & AP, R. Novel polymeric condensing carriers for gene delivery. *Proc Intl Symp Control Rel Bioact Mater* **22**, 178–9 (1995).

201. Uekama, K., Hirayama, F. & Irie, T. Cyclodextrin Drug Carrier Systems. *Chemical reviews* **98**, 2045–2076 (1998).

202. Li, J., Xiao, H., Li, J. & Zhong, Y. Drug carrier systems based on water-soluble cationic beta-cyclodextrin polymers. *International journal of pharmaceutics* **278**, 329–42 (2004).

203. Ou, W. C. *et al.* The major capsid protein, VP1, of human JC virus expressed in Escherichia coli is able to self-assemble into a capsid-like particle and deliver exogenous DNA into human kidney cells. *The Journal of general virology* **80 (Pt 1)**, 39–46 (1999).

204. Mena, J. A., Ramírez, O. T. & Palomares, L. A. Intracellular distribution of rotavirus structural proteins and virus-like particles expressed in the insect cell-baculovirus system. *Journal of biotechnology* **122**, 443–52 (2006).

205. Di Martino, B., Marsilio, F. & Roy, P. Assembly of feline calicivirus-like particle and its immunogenicity. *Veterinary microbiology* **120**, 173–8 (2007).

206. Bach, P. *et al.* Vesicular stomatitis virus glycoprotein displaying retrovirus-like particles induce a type I IFN receptor-dependent switch to neutralizing IgG antibodies. *Journal of immunology (Baltimore, Md. : 1950)* **178**, 5839–47 (2007).

207. Arís, a, Feliu, J. X., 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. *Biotechnology and bioengineering* **68**, 689–96 (2000).

208. Box, M. *et al.* A multi-domain protein system based on the HC fragment of tetanus toxin for targeting DNA to neuronal cells. *Journal of drug targeting* **11**, 333–43 (2003).

209. Arís, A. & Villaverde, A. Modular protein engineering for non-viral gene therapy. *Trends in biotechnology* **22**, 371–7 (2004).

210. Arís, A. & Villaverde, A. Engineering nuclear localization signals in modular protein vehicles for gene therapy. *Biochemical and biophysical research communications* **304**, 625–31 (2003).

211. Floudas, C. a., Fung, H. K., McAllister, S. R., Mönnigmann, M. & Rajgaria, R. Advances in protein structure prediction and de novo protein design: A review. *Chemical Engineering Science* **61**, 966–988 (2006).

212. Dahiyat, B. I. & Mayo, S. L. De novo protein design: fully automated sequence selection. *Science (New York, N.Y.)* **278**, 82–7 (1997).

213. Peluffo, H. *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 neuroscience* **7**, 35 (2006).

214. Uherek, C., Fominaya, J. & Wels, W. A modular DNA carrier protein based on the structure of diphtheria toxin mediates target cell-specific gene delivery. *The Journal of biological chemistry* **273**, 8835–41 (1998).

215. Ignatovich, I. A. *et al.* Complexes of plasmid DNA with basic domain 47-57 of the HIV-1 Tat protein are transferred to mammalian cells by endocytosis-mediated pathways. *The Journal of biological chemistry* **278**, 42625–36 (2003).

216. Arís, a & Villaverde, a Molecular organization of protein-DNA complexes for cell-targeted DNA delivery. *Biochemical and biophysical research communications* **278**, 455–61 (2000).

217. Carrió, M. M. & Villaverde, A. Role of molecular chaperones in inclusion body formation. *FEBS letters* **537**, 215–21 (2003).

218. D'Souza, S. E., Ginsberg, M. H. & Plow, E. F. Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends in biochemical sciences* **16**, 246–50 (1991).

219. Domingo-Espín, J. *et al.* RGD-based cell ligands for cell-targeted drug delivery act as potent trophic factors. *Nanomedicine : nanotechnology, biology, and medicine* **8**, 1263–6 (2012).

220. Tiera, M. J., Winnik, F. O. M. & Fernandes, J. C. Synthetic and natural polycations for gene therapy: state of the art and new perspectives. *Current gene therapy* **6**, 59–71 (2006).

221. Männistö, M. *et al.* Structure-activity relationships of poly(L-lysines): effects of pegylation and molecular shape on physicochemical and biological properties in gene delivery. *Journal of controlled release : official journal of the Controlled Release Society* **83**, 169–82 (2002).

222. Ward, C. M., Read, M. L. & Seymour, L. W. Systemic circulation of poly(L-lysine)/DNA vectors is influenced by polycation molecular weight and type of DNA: differential circulation in mice and rats and the implications for human gene therapy. *Blood* **97**, 2221–9 (2001).

223. Plank, C., Tang, M. X., Wolfe, A. R. & Szoka, F. C. Branched cationic peptides for gene delivery: role of type and number of cationic residues in formation and in vitro activity of DNA polyplexes. *Human gene therapy* **10**, 319–32 (1999).

224. McKenzie, D. L., Collard, W. T. & Rice, K. G. Comparative gene transfer efficiency of low molecular weight polylysine DNA-condensing peptides. *The journal of peptide research : official journal of the American Peptide Society* **54**, 311–8 (1999).

225. Brewer, L. R., Corzett, M. & Balhorn, R. Protamine-induced condensation and decondensation of the same DNA molecule. *Science (New York, N.Y.)* **286**, 120–3 (1999).

226. Masuda, T., Akita, H. & Harashima, H. Evaluation of nuclear transfer and transcription of plasmid DNA condensed with protamine by microinjection: the use of a nuclear transfer score. *FEBS letters* **579**, 2143–8 (2005).

227. Baleja, J. D., Thanabal, V. & Wagner, G. Refined solution structure of the DNA-binding domain of GAL4 and use of 3J(113Cd,1H) in structure determination. *Journal of biomolecular NMR* **10**, 397–401 (1997).

228. Tosi, G. *et al.* Targeting the central nervous system: in vivo experiments with peptide-derivatized nanoparticles loaded with Loperamide and Rhodamine-123. *Journal of controlled release : official journal of the Controlled Release Society* **122**, 1–9 (2007).

229. Kumar, P. *et al.* Transvascular delivery of small interfering RNA to the central nervous system. *Nature* **448**, 39–43 (2007).

230. Schwarze, S. R. In Vivo Protein Transduction: Delivery of a Biologically Active Protein into the Mouse. *Science* **285**, 1569–1572 (1999).

231. Gabathuler, R. Approaches to transport therapeutic drugs across the blood-brain barrier to treat brain diseases. *Neurobiology of disease* **37**, 48–57 (2010).

232. Vives, E. Cellular uptake [correction of utake] of the Tat peptide: an endocytosis mechanism following ionic interactions. *Journal of molecular recognition : JMR* **16**, 265–71

233. Fawell, S. *et al.* Tat-mediated delivery of heterologous proteins into cells. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 664–8 (1994).

234. Wender, P. A. *et al.* The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 13003–8 (2000).

235. Lundberg, P., El-Andaloussi, S., Sütlü, T., Johansson, H. & Langel, U. Delivery of short interfering RNA using endosomolytic cell-penetrating peptides. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **21**, 2664–71 (2007).

236. Muratovska, A. & Eccles, M. R. Conjugate for efficient delivery of short interfering RNA (siRNA) into mammalian cells. *FEBS letters* **558**, 63–8 (2004).

237. Hart, S. L. Integrin-mediated vectors for gene transfer and therapy. *Current opinion in molecular therapeutics* **1**, 197–203 (1999).

238. Kessler, T. *et al.* Inhibition of tumor growth by RGD peptide-directed delivery of truncated tissue factor to the tumor vasculature. *Clinical cancer research : an official journal of the American Association for Cancer Research* **11**, 6317–24 (2005).

239. Amara, A. *et al.* HIV coreceptor downregulation as antiviral principle: SDF-1alpha-dependent internalization of the chemokine receptor CXCR4 contributes to inhibition of HIV replication. *The Journal of experimental medicine* **186**, 139–46 (1997).

240. Liang, X. CXCR4, inhibitors and mechanisms of action. *Chemical biology & drug design* **72**, 97–110 (2008).

241. Daniels, T. R., Delgado, T., Helguera, G. & Penichet, M. L. The transferrin receptor part II: targeted delivery of therapeutic agents into cancer cells. *Clinical immunology (Orlando, Fla.)* **121**, 159–76 (2006).

242. Joung, I. *et al.* Improved gene delivery into neuroglial cells using a fiber-modified adenovirus vector. *Biochemical and biophysical research communications* **328**, 1182–7 (2005).

243. Liu, X. *et al.* Systemic genetic transfer of p21WAF-1 and GM-CSF utilizing of a novel oligopeptide-based EGF receptor targeting polyplex. *Cancer gene therapy* **10**, 529–39 (2003).

244. Wolschek, M. F. *et al.* Specific systemic nonviral gene delivery to human hepatocellular carcinoma xenografts in SCID mice. *Hepatology (Baltimore, Md.)* **36**, 1106–14 (2002).

245. Wu, Y.-T. *et al.* A new N-acetylgalactosamine containing peptide as a targeting vehicle for mammalian hepatocytes via asialoglycoprotein receptor endocytosis. *Current drug delivery* **1**, 119–27 (2004).

246. Schneider, H. *et al.* A novel peptide, PLAEIDGIELTY, for the targeting of alpha9beta1-integrins. *FEBS letters* **429**, 269–73 (1998).

247. Collins, L. & Fabre, J. W. A synthetic peptide vector system for optimal gene delivery to corneal endothelium. *The journal of gene medicine* **6**, 185–94 (2004).

248. McKay, T., Reynolds, P., Jezzard, S., Curiel, D. & Coutelle, C. Secretin-mediated gene delivery, a specific targeting mechanism with potential for treatment of biliary and pancreatic disease in cystic fibrosis. *Molecular therapy : the journal of the American Society of Gene Therapy* **5**, 447–54 (2002).

249. Zeng, J., Too, H.-P., Ma, Y., Luo, E. S. E. & Wang, S. A synthetic peptide containing loop 4 of nerve growth factor for targeted gene delivery. *The journal of gene medicine* **6**, 1247–56 (2004).

250. Martinez-Fong, D. *et al.* Neurotensin-SPDP-poly-L-lysine conjugate: a nonviral vector for targeted gene delivery to neural cells. *Brain research. Molecular brain research* **69**, 249–62 (1999).

251. White, S. J., Nicklin, S. A., Sawamura, T. & Baker, A. H. Identification of peptides that target the endothelial cell-specific LOX-1 receptor. *Hypertension* **37**, 449–55 (2001).

252. Pardridge, W. M., Buciak, J. L. & Friden, P. M. Selective transport of an anti-transferrin receptor antibody through the blood-brain barrier in vivo. *The Journal of pharmacology and experimental therapeutics* **259**, 66–70 (1991).

253. Barrier, B. B., Boado, R. J., Zhang, Y., Zhang, Y. & Pardridge, W. M. Humanization of Anti-Human Insulin Receptor Antibody for Drug Targeting Across the Human. **96**, 381–391 (2007).

254. Jeong, J. H. *et al.* Anti-GAD antibody targeted non-viral gene delivery to islet beta cells. *Journal of controlled release : official journal of the Controlled Release Society* **107**, 562–70 (2005).

255. Sedighiani, F. & Nikol, S. Gene therapy in vascular disease. *The surgeon : journal of the Royal Colleges of Surgeons of Edinburgh and Ireland* **9**, 326–35 (2011).

256. Ozden, M., Kalkan, A., Demirdag, K., Denk, A. & Kilic, S. S. Hepatocyte growth factor (HGF) in patients with hepatitis B and meningitis. *The Journal of infection* **49**, 229–35 (2004).

257. Frankel, A. E., Powell, B. L., Duesbery, N. S., Vande Woude, G. F. & Leppla, S. H. Anthrax fusion protein therapy of cancer. *Current protein & peptide science* **3**, 399–407 (2002).

258. Zhao, Y. *et al.* Introduction of hypoxia-targeting p53 fusion protein for the selective therapy of non-small cell lung cancer. *Cancer biology & therapy* **11**, 95–107 (2011).

259. Sioud, M. & Mobergslien, A. Selective killing of cancer cells by peptide-targeted delivery of an anti-microbial peptide. *Biochemical pharmacology* **84**, 1123–32 (2012).

260. Baneyx, F. Recombinant protein expression in Escherichia coli. *Current opinion in biotechnology* **10**, 411–21 (1999).

261. Sørensen, H. P. & Mortensen, K. K. Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of biotechnology* **115**, 113–28 (2005).

262. Dietz, G. P. H. & Bähr, M. Synthesis of cell-penetrating peptides and their application in neurobiology. *Methods in molecular biology (Clifton, N.J.)* **399**, 181–98 (2007).

263. Schmidt, F. R. Recombinant expression systems in the pharmaceutical industry. *Applied microbiology and biotechnology* **65**, 363–72 (2004).

264. Birch, J. R. & Froud, S. J. Mammalian cell culture systems for recombinant protein production. *Biologicals : journal of the International Association of Biological Standardization* **22**, 127–33 (1994).

265. Ikonomou, L., Schneider, Y.-J. & Agathos, S. N. Insect cell culture for industrial production of recombinant proteins. *Applied microbiology and biotechnology* **62**, 1–20 (2003).

266. Reiter, M. & Blüml, G. Large-scale mammalian cell culture. *Current opinion in biotechnology* **5**, 175–9 (1994).

267. Gervais, A. *et al.* Glycosylation of human recombinant gonadotrophins: characterization and batch-to-batch consistency. *Glycobiology* **13**, 179–89 (2003).

268. Collins, L., Asuni, A. A., Anderton, B. H. & Fabre, J. W. Efficient gene delivery to primary neuron cultures using a synthetic peptide vector system. *Journal of neuroscience methods* **125**, 113–20 (2003).

269. Kim, E.-S. *et al.* Cell-penetrating DNA-binding protein as a safe and efficient naked DNA delivery carrier in vitro and in vivo. *Biochemical and biophysical research communications* **392**, 9–15 (2010).

270. Fabre, J. W. & Collins, L. Synthetic peptides as non-viral DNA vectors. *Current gene therapy* **6**, 459–80 (2006).

271. Anderson, D. G. *et al.* A polymer library approach to suicide gene therapy for cancer. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 16028–33 (2004).

272. Van Gaal, E. V. B. *et al.* How to screen non-viral gene delivery systems in vitro? *Journal of controlled release : official journal of the Controlled Release Society* **154**, 218–32 (2011).

273. Ogris, M. *et al.* The size of DNA/transferrin-PEI complexes is an important factor for gene expression in cultured cells. *Gene therapy* **5**, 1425–33 (1998).

274. Erbacher, P. *et al.* Genuine DNA/polyethylenimine (PEI) complexes improve transfection properties and cell survival. *Journal of drug targeting* **12**, 223–36 (2004).

275. Liu, X.-X. *et al.* PAMAM dendrimers mediate siRNA delivery to target Hsp27 and produce potent antiproliferative effects on prostate cancer cells. *ChemMedChem* **4**, 1302–10 (2009).

276. Guerra-Crespo, M., Charli, J.-L., Rosales-García, V. H., Pedraza-Alva, G. & Pérez-Martínez, L. Polyethylenimine improves the transfection efficiency of primary cultures of post-mitotic rat fetal hypothalamic neurons. *Journal of neuroscience methods* **127**, 179–92 (2003).

277. Dalby, B. *et al.* Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods (San Diego, Calif.)* **33**, 95–103 (2004).

278. Manfredsson, F. P., Lewin, A. S. & Mandel, R. J. RNA knockdown as a potential therapeutic strategy in Parkinson's disease. *Gene therapy* **13**, 517–24 (2006).

279. Manfredsson, F. P. *et al.* rAAV-mediated nigral human parkin over-expression partially ameliorates motor deficits via enhanced dopamine neurotransmission in a rat model of Parkinson's disease. *Experimental neurology* **207**, 289–301 (2007).

280. Lim, S. T., Airavaara, M. & Harvey, B. K. Viral vectors for neurotrophic factor delivery: a gene therapy approach for neurodegenerative diseases of the CNS. *Pharmacological research : the official journal of the Italian Pharmacological Society* **61**, 14–26 (2010).

281. King, G. D. *et al.* Gene therapy and targeted toxins for glioma. *Current gene therapy* **5**, 535–57 (2005).

282. Bowers, W. J., Breakefield, X. O. & Sena-Esteves, M. Genetic therapy for the nervous system. *Human molecular genetics* **20**, R28–41 (2011).

283. The Journal of Gene Medicine Clinical Trial site. *The Journal of Gene Medicine* at <http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1521-2254>

284. Bergen, J. M., Park, I.-K., Horner, P. J. & Pun, S. H. Nonviral approaches for neuronal delivery of nucleic acids. *Pharmaceutical research* **25**, 983–98 (2008).

285. Barati, S. *et al.* GDNF gene delivery via the p75(NTR) receptor rescues injured motor neurons. *Experimental neurology* **202**, 179–88 (2006).

286. Navarro-Quiroga, I. *et al.* Improved neurotensin-vector-mediated gene transfer by the coupling of hemagglutinin HA2 fusogenic peptide and Vp1 SV40 nuclear localization signal. *Brain research. Molecular brain research* **105**, 86–97 (2002).

287. Zeng, J., Wang, X. & Wang, S. Self-assembled ternary complexes of plasmid DNA, low molecular weight polyethylenimine and targeting peptide for nonviral gene delivery into neurons. *Biomaterials* **28**, 1443–51 (2007).

288. Knight, A. *et al.* Non-viral neuronal gene delivery mediated by the HC fragment of tetanus toxin. *European journal of biochemistry / FEBS* **259**, 762–9 (1999).

289. Liu, J. K. *et al.* A novel peptide defined through phage display for therapeutic protein and vector neuronal targeting. *Neurobiology of disease* **19**, 407–18 (2005).

290. Park, I.-K., Lasiene, J., Chou, S.-H., Horner, P. J. & Pun, S. H. Neuron-specific delivery of nucleic acids mediated by Tet1-modified poly(ethylenimine). *The journal of gene medicine* **9**, 691–702 (2007).

291. Faraji, A. H. & Wipf, P. Nanoparticles in cellular drug delivery. *Bioorganic & medicinal chemistry* **17**, 2950–62 (2009).

292. Smith, Q. R. & Rapoport, S. I. Cerebrovascular permeability coefficients to sodium, potassium, and chloride. *Journal of neurochemistry* **46**, 1732–42 (1986).

293. Misra, A., Ganesh, S., Shahiwala, A. & Shah, S. P. Drug delivery to the central nervous system: a review. *Journal of pharmacy & pharmaceutical sciences : a publication of the Canadian Society for Pharmaceutical Sciences, Société canadienne des sciences pharmaceutiques* **6**, 252–73 (2003).

294. Pardridge, W. M. Blood-brain barrier delivery. *Drug discovery today* **12**, 54–61 (2007).

295. Lee, H. J., Engelhardt, B., Lesley, J., Bickel, U. & Pardridge, W. M. Targeting rat anti-mouse transferrin receptor monoclonal antibodies through blood-brain barrier in mouse. *The Journal of pharmacology and experimental therapeutics* **292**, 1048–52 (2000).

296. Aktaş, Y. *et al.* Development and brain delivery of chitosan-PEG nanoparticles functionalized with the monoclonal antibody OX26. *Bioconjugate chemistry* **16**, 1503–11 (2005).

297. Jenkins, N. Modifications of therapeutic proteins: challenges and prospects. *Cytotechnology* **53**, 121–5 (2007).

298. Prouty, W. F., Karnovsky, M. J. & Goldberg, A. L. Degradation of abnormal proteins in Escherichia coli. Formation of protein inclusions in cells exposed to amino acid analogs. *The Journal of biological chemistry* **250**, 1112–22 (1975).

299. Martínez-Alonso, M., González-Montalbán, N., García-Fruitós, E. & Villaverde, A. Learning about protein solubility from bacterial inclusion bodies. *Microbial cell factories* **8**, 4 (2009).

300. Villaverde, A. & Carrió, M. M. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnology letters* **25**, 1385–95 (2003).

301. Marston, F. A. The purification of eukaryotic polypeptides synthesized in Escherichia coli. *The Biochemical journal* **240**, 1–12 (1986).

302. Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E. & Villaverde, A. Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories* **8**, 17 (2009).

303. González-Montalbán, N., García-Fruitós, E. & Villaverde, A. Recombinant protein solubility - does more mean better? *Nature biotechnology* **25**, 718–20 (2007).

304. De Marco, A. Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in Escherichia coli. *Nature protocols* **2**, 2632–9 (2007).

305. De Marco, A. Minimal information: an urgent need to assess the functional reliability of recombinant proteins used in biological experiments. *Microbial cell factories* **7**, 20 (2008).

306. Gasser, B. *et al.* Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microbial cell factories* **7**, 11 (2008).

307. Futaki, S., Nakase, I., Tadokoro, A., Takeuchi, T. & Jones, A. T. Arginine-rich peptides and their internalization mechanisms. *Biochemical Society transactions* **35**, 784–7 (2007).

308. Ferrer-Miralles, N., Vázquez, E. & Villaverde, A. Membrane-active peptides for non-viral gene therapy: making the safest easier. *Trends in biotechnology* **26**, 267–75 (2008).

309. Melikov, K. & Chernomordik, L. V Arginine-rich cell penetrating peptides: from endosomal uptake to nuclear delivery. *Cellular and molecular life sciences : CMLS* **62**, 2739–49 (2005).

310. Abes, R. *et al.* Cell-penetrating-peptide-based delivery of oligonucleotides: an overview. *Biochemical Society transactions* **35**, 775–9 (2007).

311. Kim, H. H., Choi, H. S., Yang, J. M. & Shin, S. Characterization of gene delivery in vitro and in vivo by the arginine peptide system. *International journal of pharmaceutics* **335**, 70–8 (2007).

312. Kim, H. H., Lee, W. S., Yang, J. M. & Shin, S. Basic peptide system for efficient delivery of foreign genes. *Biochimica et biophysica acta* **1640**, 129–36 (2003).

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

314. Futaki, S. *et al.* Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. *The Journal of biological chemistry* **276**, 5836–40 (2001).

315. Hassane, F. S. *et al.* A peptide-based dendrimer that enhances the splice-redirecting activity of PNA conjugates in cells. *Bioconjugate chemistry* **20**, 1523–30 (2009).

316. Schwarze, S. R., Ho, A., Vocero-Akbani, A. & Dowdy, S. F. In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science (New York, N.Y.)* **285**, 1569–72 (1999).

317. Maham, A., Tang, Z., Wu, H., Wang, J. & Lin, Y. Protein-based nanomedicine platforms for drug delivery. *Small (Weinheim an der Bergstrasse, Germany)* **5**, 1706–21 (2009).

318. Wang, L., Maji, S. K., Sawaya, M. R., Eisenberg, D. & Riek, R. Bacterial inclusion bodies contain amyloid-like structure. *PLoS biology* **6**, e195 (2008).

319. Carrió, M., González-Montalbán, N., Vera, A., Villaverde, A. & Ventura, S. Amyloid-like properties of bacterial inclusion bodies. *Journal of molecular biology* **347**, 1025–37 (2005).

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

321. Peluffo, H. *et al.* RGD domains neuroprotect the immature brain by a glial-dependent mechanism. *Annals of neurology* **62**, 251–61 (2007).

322. Saccardo, P., Villaverde, A. & González-Montalbán, N. Peptide-mediated DNA condensation for non-viral gene therapy. *Biotechnology advances* **27**, 432–8 (2009).

323. Keller, M., Tagawa, T., Preuss, M. & Miller, A. D. Biophysical characterization of the DNA binding and condensing properties of adenoviral core peptide mu. *Biochemistry* **41**, 652–9 (2002).

324. Dunehoo, A. L. *et al.* Cell adhesion molecules for targeted drug delivery. *Journal of pharmaceutical sciences* **95**, 1856–72 (2006).

325. Williams, D. F. The role of short synthetic adhesion peptides in regenerative medicine; the debate. *Biomaterials* **32**, 4195–7 (2011).

Other publications

Non-amyloidogenic peptide tags for the regulatable self-assembling of protein-only nanoparticles.
Unzueta U, Ferrer-Miralles N, Cedano J, Zikung X, Pesarrodona M, Saccardo P, García-Fruitós E, ***Domingo-Espín J****, Kumar P, Gupta KC, Mangues R, Villaverde A, Vazquez E.*

Biomaterials. 2012 Nov;33(33):8714-22

Intracellular CXCR4^{+} cell targeting with T22-empowered protein-only nanoparticles.
Unzueta U, Céspedes MV, Ferrer-Miralles N, Casanova I, Cedano J, Corchero JL, ***Domingo-Espín J****, Villaverde A, Mangues R, Vázquez E.*

Int J Nanomedicine. 2012;7:4533-44

Internalization and kinetics of nuclear migration of protein-only, arginine-rich nanoparticles.
Vázquez E, Cubarsi R, Unzueta U, Roldán M, ***Domingo-Espín J,*** *Ferrer-Miralles N, Villaverde A*
.
Biomaterials. 2010 Dec;31(35):9333-9

Protein aggregation and soluble aggregate formation screened by a fast microdialysis assay.
Toledo-Rubio V, Vazquez E, Platas G, ***Domingo-Espín J****, Unzueta U, Steinkamp E, García-Fruitós E, Ferrer-Miralles N, Villaverde A.*

J Biomol Screen. 2010 Apr;15(4):453-7

Acknowledgements

Aquesta part de la tesi, que a nivell acadèmic no és gens important, per mi, com per molts que han passat abans que jo per aquesta fase, ho és especialment. Ho és per varies raons.

En primer lloc és on haig de fer un repàs de tot i tots els que han fet possible que aquesta tesi arribés al seu destí, o sigui, al final. És una tasca difícil ja que han sigut molts els factors que han influenciat en mi aquest últims anys i que han fet que jo esdevingués la persona a nivell professional i personal que sóc ara. Anomenar-los tots serà difícil i segur que em deixo coses però intentaré fer-ho el millor que pugui.

En segon lloc és un espai en què per fi puc escriure en l'idioma en què vaig aprendre a parlar i la gent que em coneix sap que això és molt important per mi. Tot i així, altres idiomes no en seran exclosos ja que per sort, he tingut relació amb gent de llocs ben diversos alhora de realitzar aquesta tesi i per ells vull escriure també aquests agraïments.

Finalment, abans d'anomenar a tota la gent a qui li vull donar les gràcies, vull destacar el valor que aquesta tesi ha tingut per mi. Un cop acabada la carrera i amb la il·lusió que suposava entrar a treballar com a investigador a un laboratori, no em podia imaginar com m'acabaria agradant aquesta feina. Si és veritat que hi ha molts moments durs i difícils, també és veritat que és una de les feines que pot aportar també més recompenses. Recompenses gràcies a la feina feta i als bons resultats però també recompenses gràcies a la gent amb qui has d'interactuar, de les quals cada dia n'aprens alguna cosa i que fan que la vida a un laboratori sigui molt més còmode i divertida. A aquesta gent és la que vull anomenar a continuació.

Els primers agraïments d'aquesta tesi van dirigits al Toni Villaverde. Vull agrair que em permetessis entrar a treballar al laboratori de Micro Aplicada quan encara no havia acabat la carrera i que em posessis tantes facilitats per quedar-m'hi a treballar i a fer el doctorat. Moltes gràcies de tot cor.

A la Neus i l'Esther per haver-me dirigit la tesi. Moltíssimes gràcies a totes dues per haver-me ajudat en tot moment. A la Neus, especialment, li agraeixo tot el temps que va passar amb mi ensenyant-me com fer moltíssimes coses al laboratori. Si no fos per ella no sabria res de la feina al laboratori.

A tota la gent que ha estat al laboratori durant el temps que he fet la tesi. Ha sigut molta gent i de tothom m'emporto alguna cosa.

A l'Elena, la Núria i la Rosa –que per mi no poden anomenar-se per separat– per estar allà als meus inicis i fer-me sentir com un més del grup. A vosaltres us agraeixo haver-me ensenyat moltes coses entre elles com formar part d'un grup d'investigació amb tan bon ambient.

A tots els meus companys de laboratori, al Paolo, l'Ugutz, la Olivia que sempre heu creat un ambient tan amable i agradable per treballar i que feu que sempre es pugui riure i parlar d'alguna cosa.

Al Pepe i la Rosa pels bons consells i per l'ajuda i la organització al laboratori.

Al Quim, per ser tan bon company i amic. Amb ell hem pogut compartir frustracions i alegries i sempre he pogut comptar-hi per qualsevol cosa. Moltes gràcies per ajudar-me en tantes ocasions. Espero haver-ho pogut fer jo també!

A la Vero. Què dir d'una de les persones més important al grup i per mi! M'has ensenyat moltíssimes coses i no em refereixo només en el món de les puris! Amb ella hem rigut i plorat i sé que tinc una amistat per tota la vida. Gràcies per tot!

A tota la gent de l'IBB que ha fet possible aquesta tesis i a crear un ambient tan bo i que no és fàcil de trobar. Són molts i hi han contribuït de moltes maneres però sempre amb molt bon humor i ganes d'ajudar. A l'Almu i el Juan Checa, el Xino i el

Juan Cedano, al servei de cultius, el Fran i la Paqui, a la Manuela i a molts altres.

A la gent del Clínic. A l'Anna i la Valérie per ajudar-me tant. Sense la vostra ajuda aquesta tesi segur que no s'hauria acabat.

A Hugo, Javier y Luciana. Os quiero agradecer haberme abierto las puertas de vuestro laboratorio en Uruguay. Ha sido una de las experiencias más importantes a nivel profesional y personal de mi tesis.

To Carlos, Marijn, Renée and the rest of the SILS-CNS group for one of the most valuable experiences of my entire thesis. Thanks for teaching me so many things. With you I realized how much I like science. I want to thank specially Carlos for giving me the opportunity to go to the Netherlands and spend such a wonderful time in there and for making me a better scientist. I'll never forget the help and enthusiasm that you gave to me for the innovation award that we won. Muchas gracias!

A tota la gent de la carrera que m'heu estat donant suport tot aquest temps i que heu fet que compartit les nostres experiències sigui la millor teràpia de totes. A la Jone, l'Alba, la Bàrbara, el Jofre, el Pep, la Íngrid, el Roger, la Cinta i segur que em deixo gent pel camí, moltes gràcies a tots.

Als de casa. Els meus pares es mereixen un agraïment molt especial. La seva contribució a aquesta tesi no es pot mesurar en articles o feina feta però sense vosaltres això no s'hauria fet mai. Vosaltres que em vau ensenyar a pensar i fer com penso i faig us dono les gràcies per haver-me donat suport sempre.

I al meu germà, el Miquel per animar-me i escoltar-me.

Finalment només queda la persona més important de totes. A tu Aitor, la persona a qui va dedicada aquesta tesi. Vaig començar la tesi i a la setmana ens vam conèixer. Sembla que ets el denominador comú de tots aquest anys d'esforç. Tu ets la

persona que m'ha donat més ànims i suport, interessant-te per la meva feina fins i tot en aquells moment que no et venia de gust sentir-ne a parlar, aguantant la distància quan la tesi m'ha portat lluny i fent-me créixer com a persona. Hirekin Aitor oraindik askoz gauza gehiago igaro nahi dizkiat eta duela urte batzuk eraikitzen hasi genuen familia txiki hau ez dadila inoiz amaitu. Maite haut.