Expert Opinion

Advances with the use of bio-inspired vectors towards creation of artificial viruses

Helen O McCarthy, Yuhua Wang, Sriramchandra S Mangipudi & Arash Hatefi†

†*Washington State University, Center for Integrated Biotechnology, Department of Pharmaceutical Sciences, PO Box 646534, Pullman, WA 99164, USA*

Importance of the field: In recent years, there has been a great deal of interest in the development of recombinant vectors based on biological motifs with potential applications in gene therapy. Several such vectors have been genetically engineered, resulting in biomacromolecules with new properties that are not present in nature.

Areas covered in this review: This review briefly discusses the advantages and disadvantages of the current state-of-the-art gene delivery systems (viral and non-viral) and then provides an overview on the application of various biological motifs in vector development for gene delivery. Finally, it highlights some of the most advanced bio-inspired vectors that are designed to perform several self-guided functions.

What the reader will gain: This review helps the readers get a better understanding about the history and evolution of bio-inspired fusion vectors with the potential to merge the strengths of both viral and non-viral vectors in order to create efficient, safe and cost-effective gene delivery systems.

Take home message: With the emergence of new technologies such as recombinant bio-inspired vectors, it may not take long before non-viral vectors are observed that are not just safe and tissue-specific, but even more efficient than viral vectors.

Keywords: bio-inspired, biomimetic, gene delivery, non-viral, recombinant vectors

Expert Opin. Drug Deliv. (2010) 7(4):497-512

1. Introduction

Gene therapy can be defined as the treatment of human disease by the transfer of corrective or therapeutic genes into specific cells of a patient [1]. In theory, the science of gene therapy has the potential to cure many diseases, but its progress has been hindered by the shortcomings of the available gene delivery systems (vectors) [2,3]. In general, gene carriers are divided into viral and non-viral categories where both have significant limitations [3]. The hallmark features of a suitable vector for clinical applications include low cytotoxicity/immunogenicity, high transfection efficiency, tissue specificity and cost-effectiveness. Although there has been significant progress in vector development for various gene therapy needs in the last decade [3,4], there is no single vector that is fortified with all the indispensable features. This review starts by highlighting the major advantages and unique challenges that are associated with the current state-of-the-art vectors. Subsequently, it describes a more sophisticated type of non-viral gene delivery system, namely, genetically engineered bio-inspired vectors that bear the potential to hybridize the strengths of both viral and non-viral vectors in order to overcome the barriers to efficient, safe and cost-effective gene delivery. The coverage of the literature is not encyclopedic; rather, a few select examples have been chosen to highlight certain points. The discussion emphasizes

10.1517/17425240903579989 © 2010 Informa UK Ltd ISSN 1742-5247

Article highlights.

- Although there has been significant progress in vector development (viral and non-viral) for various gene therapy needs in the last decade, there is unfortunately no single vector that has collectively demonstrated safety, efficiency, tissue specificity and cost-effectiveness.
- None of the peptide motifs of biological origin (e.g., TAT, Mu, M9-NLS, etc.) has the ability to overcome independently the major cellular barriers associated with successful targeted gene delivery.
- Independently, recombinant bio-inspired vectors have the ability to perform several self-guided functions, including cell targeting, DNA condensation, endosome disruption and nuclear localization, among others.
- This class of vectors (i.e., recombinant bio-inspired vectors) can be used as a tool for precise structure/activity relationship studies and holds the potential to merge the strengths of both viral and non-viral vectors in order to overcome the main barriers to efficient, safe and cost-effective gene delivery.
- To overcome the challenges that are posed to the success of gene therapy, in parallel with the technologies of the twentieth century, significant research and investment on new twenty-first century technologies are required.

This box summarizes key points contained in the article.

some of the practical issues, problems and unique challenges that are associated with the use of various biological motifs for gene delivery.

1.1 Viral vectors

Viruses are nature's work of art, evolved to infect and transfer their genes efficiently into the nucleus of the host cells. For example, the basic elements of the infection and intracellular trafficking pathway for adenoviruses include high-affinity binding of the capsid to receptors at the cell surface, internalization by endocytosis, lysis of the endosomal membrane resulting in escape to the cytosol, trafficking along the microtubules, binding to the nuclear envelope and insertion of the viral genome through the nuclear pore. These features are in-built in most viral vectors, conferring an advantage over non-viral delivery methods.

There are five main classes of viral vectors used in gene therapy, which can be categorized into two groups, that is, integrating and non-integrating (**Table 1**) [3,5]. Integrating viruses are divided into two classes, namely, oncoretroviruses (Moloney murine leukemia virus [MLV]) and lentiviruses (e.g., human immunodeficiency virus [HIV]). Both viruses help to insert the therapeutic gene into the host genome for expression. Long-term expression, ease of production and extendible cellular tropism beyond the existing target range are strong incentives to use these vectors [6]. One important point worth emphasizing is that the assembly and production of such viruses is considered easy, whereas their production in high titers ($> 10^8$ pfu/ml) is time-consuming and difficult. The principal advantage of lentiviral versus retroviral delivery is the ability of the lentivirus to transduce both dividing and non-dividing cells. With low immunogenicity characteristics, both lentiviruses and retroviruses have been utilized in the treatment of severe combined immune deficiency (SCID) [7]. One of the significant limitations of these vectors is the potential to activate oncogenes through the nonspecific integration of genetic material into the host's genome [8].

In contrast to integrating viruses, non-integrating viruses are mostly utilized to treat diseases where transient gene expression is desired. The non-integrating viral vectors are classified as adenovirus, adeno-associated virus (AAV) and herpes simplex virus 1 (HSV-1). These viruses have the tendency to retain the therapeutic DNA as an independent extrachromosomal episome [3]. AAV vectors are attractive because of their non-pathogenic nature, capacity to transfect non-proliferating cells and the proven *in vivo* long-term gene expression [9]. AAVs are generally considered non-immunogenic because of their small size (~ 20 nm) and single capsid protein architecture, which prevents significant build-up of immune responses in most people. However, recent studies have provided some evidence that most of the human population has been exposed to the most common AAV serotypes, which has resulted in circulating neutralizing antibodies and a sensitized immune system [10].

Among all viral vectors, adenoviruses are considered the most efficient. This is owing to their ability to transduce both dividing and non-dividing cells. Adenoviruses have been used extensively in a range of gene therapy clinical trials [11]. In particular, they have been used in the treatment of cancer wherein transient expression of a therapeutic gene is required. Despite their major advantage in terms of high transduction efficiency, their use has been limited owing to significant dose-dependent immunogenicity and nonspecific transduction of cells. The severity of the immunogenic response to high doses of adenoviruses was highlighted by the Jesse Gelsinger case. Gelsinger received a high dose of adenovirus vector in a clinical trial (3.8×10^{13} particles) for the treatment of a deficiency in the liver enzyme ornithine transcarbamylase. After 4 h of treatment, Gelsinger developed a high fever and within 4 days of treatment he died from multiple organ failure. A female patient who received a similar dose (3.6×10^{13} particles) experienced no unexpected side effects. It is believed that a previous exposure to a wild-type viral infection may have sensitized Gelsinger's immune system to the vector [12]. Subsequent studies in primates elucidated that the major contributing factor to the immunogenicity of the adenoviral vectors was the presence of several high-molecular-mass peptides in their capsid [13]. Several groups are now working to overcome these immunogenicity problems by developing recombinant adenoviruses that can potentially make possible prolonged expression and repeated adenoviral administration [14].

Another major challenge to the use of adenoviral vectors is their nonspecific binding to various cell types. For efficient transduction, the adenovirus vectors depend on the

Table 1. The main groups of viral vectors used in gene therapy.

	Vector	Immunogenic potential	Tropism	Limitation	Major advantage
Integrating	Retrovirus	Low	Dividing cells only	Potential oncogenicity	Long-term gene expression in dividing cells
	Lentivirus	Low	Dividing and non-dividing	Potential oncogenicity	Long-term gene expression in most cells
Episomal	AAV*	Low	Dividing and non-dividing	Small packaging capacity	Non-inflammatory and non-pathogenic
	Herpes simplex virus	High	High in neurons	Transient gene expression in neurons	Large packaging capacity
	Adenovirus	High	Dividing and non-dividing (CAR receptor)	Viral capsid could induce immune response	Efficient transduction of most cells

Reproduced with permission from [5].
*AAVs genome exist in host as > 90% episomal and < 10% integrated.
CAR: Coxsackievirus and adenovirus receptor.

coxsackievirus and adenovirus receptor (CAR). However, the potential application of adenovirus vectors is limited by the availability of the CAR on different tissues. For example, it is known that most cancer cells do not express CAR, which in turn may limit the application of adenovirus vectors in cancer gene therapy [15]. An alternative approach to overcome this problem could be devised by developing a recombinant adenovirus vector that recognizes a receptor other than CAR for its infection. For example, the FGF2 receptor is abundantly expressed on some types of cancer cell [16].

In an attempt to enhance gene delivery to cancer cells and overcome the limited expression of CAR on such cells, HSV vectors have been utilized. HSV is highly potent at infecting tumor cells as it uses the ubiquitously present cell surface heparan sulfate glycosaminoglycans [17]. Despite the fact that heparin sulfate-mediated HSV infection enhances the host range, the ubiquitous expression of the receptor in normal cells may limit the selectivity of the vector. Recently, a new generation of targeted HSV vectors was reported with ablated heparin sulfate binding sites and alternative receptor binding ligands for hepatocytes and human malignant glioma cells [18]. The successful retargeting of HSV has invigorated the possible development of more efficient targeted vectors.

Although not considered among the five main categories of viral vectors for gene therapy, the poxvirus (vaccinia virus) is another vector that is worth mentioning. Researchers have used the poxvirus more than any other type for the development of vaccines against rabies, malaria and influenza. One of the main reasons that this virus has been used for vaccine delivery is related to its particularly large carrying capacity for introduced transgenes (> 25 kb) coupled with its ability to express multiple antigens. This is especially important for a vaccine target where an effective vaccine may need to elicit an immune response against multiple antigens [19]. Owing to significant immunogenicity, the chief use of the poxvirus has been primarily in immunotherapy.

Although there has been significant progress in the development of viral vectors for gene therapy, tissue specificity, high costs of production and safety remain major concerns. To fulfill the deficiencies associated with viral vectors, synthetic non-viral vectors are envisioned as the potential alternatives that might provide a safer and efficient delivery of gene-cargo.

1.2 Non-viral vectors

Non-viral vectors are broadly classified into two categories: i) cationic lipids such as 1,2-dioleyloxy-3-trimethylammonium propane (DOTAP) and *N*-[1-(2,3-dioleyloxypropyl]-*N*,*N*,*N*-trimethylammonium chloride (DOTMA); and ii) cationic polymers such as polyethylenimine (PEI) and poly(L-lysine) (PLL). These cationic lipids and polymers are designed to form complexes with negatively charged DNA by means of electrostatic attractions.

1.2.1 Cationic lipids

On mixing with cationic lipids, plasmid DNA is condensed into small stable particles called lipoplexes. The gene transfer efficiency of lipoplexes is affected by the following factors: the structure of the cationic lipid; lipid-to-DNA charge ratio; the structure and proportion of the helper lipid in the complexes; the complex size and surface charge; the total amount of lipoplexes applied; and the cell type [20]. One of the principal drawbacks to the use of lipoplexes is related to a high positive surface charge. It is well understood that as the surface positive charge of the lipoplexes increases, the potential for binding to the cells' membrane and subsequent internalization also increases. This could result in higher levels of gene transfer efficiency. Unfortunately, this high surface positive charge is responsible not only for significant cytotoxicity, but also for binding to the negatively charged serum proteins such as albumin, which results in aggregation [21]. To mitigate these potential problems, PEG-lipid conjugates have been incorporated into the lipoplexes to reduce their surface charge

density and minimize their toxic manifestations [22]. Polyethylene glycol (PEG) is a universally used surface modification agent for stabilizing the non-viral vectors through charge and epitope-masking against blood components and the immune system. However, two separate groups have recently reported that repeated injection of PEGylated liposomes in rats and mice elicited PEG-specific IgM/IgG antibodies [23,24]. These studies highlight the fact that even a presumably safe polymer such as PEG may evoke an immune system response when repeatedly injected at high doses. This in turn may undermine the ability of PEG to be used as surface stabilizer for drug delivery systems that need multiple injections to achieve a significant therapeutic response. Size heterogeneity of targeted lipoplexes is another well-known factor that affects the reproducibility of gene transfer efficiency. This is an important parameter that determines the mechanism of particle entry into cells. Studies have shown that receptor-mediated targeted entry into the cells is facilitated exclusively through the clathrin-coated pathway with an optimal size distribution < 150 nm [25]. Above 150 nm particle size, the cellular uptake is biased towards other nonspecific pathways. Therefore, as it is difficult to control the size of targeted lipoplexes, their uptake through the clathrin pathway could prove challenging.

1.2.2 *Cationic polymers*

PEI has been found to be a versatile polymeric vector for highly efficient gene delivery that tightly condenses plasmid DNA and promotes transgene delivery to the nucleus of mammalian cells [26]. However, it has also been shown that PEI-based polyplexes have molecular weight, zeta-potential, particle size and a degree of branched-dependent cytotoxicity [27]. Furthermore, uncomplexed free PEI or free PEI generated after DNA delivery has been shown to cause immediate and delayed cytotoxicity, respectively [28]. PEI is generally believed to be non-immunogenic mostly owing to the lack of structural hierarchy. Despite the reports on the toxicity of such PEI-based vectors, which stems from their non-biodegradability, in general they are still assumed to have low immunogenic potential [28]. Like lipoplexes, the cationic polymer PEI has the potential to interact with negatively charged blood components on systemic administration and therefore be cleared from the circulation before reaching the target tissue. Again, polymers such as PEG, for example, have been utilized to stabilize sterically the surface of PEI particles thereby reducing interactions with the elements of blood and the immune system [29].

PLL is another well-studied synthetic non-viral DNA-condensing agent that can be used over a wide polymer length range. Although its biodegradable nature makes it suitable for *in vivo* applications, PLL is a weak transfection reagent that needs helper agents such as chloroquine for improved efficiency [30]. Chloroquine is a buffering agent known to disrupt endosomal membranes. Unfortunately, the cytotoxic nature of chloroquine limits its use. Although there have been a significant number of studies aimed at identifying the limiting factors to polymeric gene transfer, no substantial improvement has occurred. This could be because polymer heterogeneity has limited the feasibility of essential structure/function relationship studies. As most conventional polymers are synthesized using free radical addition or similar methods, the resulting product is heterogeneous. Consequently, the structure/function relationship has been difficult to elucidate. In addition, heterogeneity manifests itself during conjugation or covalent coupling of other moieties (e.g., targeting peptides), because substitution reactions are not necessarily precise. As a result, the ability to control ligand density on the surface of polyplexes, which significantly impacts transfection efficiency, could be diminished.

In summary, non-viral gene delivery vectors are still in the developmental stages and are marred by several limitations, including poor gene transfer rates, lack of reproducibility, heterogeneity and cytotoxicity [31]. Despite immunogenicity and other safety concerns, viral vectors remain the choicest tools for gene delivery owing primarily to the high gene transfer efficiency, ability to engineer, and lack of any other competing vector technology.

2. Application of biological motifs in gene delivery

Various peptide motifs of diverse biological origins have been shown to perform several basic functions, such as cell penetration (e.g., transactivation protein [TAT] peptide) [32], DNA binding and condensation (e.g., adenovirus μ peptide, SPKK repeating motif in histone H1) [33,34], membrane lysis (e.g., mellitin, influenza virus HA2,) [35,36] and nuclear localization (e.g., M9-NLS and HIV-1 Rev protein) [37,38]. Several of these peptides have been studied for their potential to improve the gene transfer efficiency into mammalian cells.

2.1 DNA-condensing peptides

Formation of stable DNA-vector condensates is essential for efficient *in vitro* and *in vivo* gene delivery. Stability of these DNA–peptide complexes ensures protection against endonucleases and safe release, enabling nuclear expression. In nature, DNA condensation and packaging is facilitated by cationic peptides rich in basic amino acids such as arginine, histidine and lysine. Here, the scope of the discussion is limited to the most widely studied and naturally occurring DNA-condensing reagents (**Table 2**).

2.1.1 *Mu peptide*

Studies with the cationic adenoviral core peptide, μ (Mu), which is rich in arginines (MRRAHHRRRRASHRRM-RGG), revealed binding to plasmid DNA within milliseconds [39]. Mu has been shown to be vastly superior to the VP1 peptide (DNA binding peptide found within the polyoma virus) for both nucleic acid neutralization and condensation [39]. Furthermore, interactions between Mu

Table 2. The characteristics of the most common DNA-condensing peptides.

Peptide	Sequence	Complex size (nm)	Limitation	Major advantage	Ref.
Mu	MRRAHHRRRRASHRRMRGG	80 – 110	Heterogeneity of complexes. Unable to escape endosome	Stable particles	[38,40,88]
TAT	TGRKKRRQRRR	300 – 470	Poor DNA condenser. Unable to escape endosomes	Can penetrate the cell membrane	[47]
Tyr-TAT	TTGRKKRRQRRR	78 – 102	Tyr3-TAT results in aggregation but a higher transfection efficacy than Tyr-TAT	Efficient DNA condenser, constant polydispersity	[47]
Poly-TAT	Oxidative polycondensation of CTATC with DMSO for 4 days	< 200	Incomplete endosomal escape	Less toxic than PEI or TAT, low polydispersity	[46]
Histones	H2B 126 aa, H1.4F 387 aa at the 3′ end H1 (S/TPKK peptides) 34mer: ATPKKSTKKTPKKAKKPA AAAGAKKAKSPKKAKA 16mer: ATPKKSTKKTPKKAKK	Not measured	Potential immunogenicity *in vivo* No information about potential endosomal escape	Efficient DNA condenser, gene expression > 35%, 34 mer and 16mer identified as efficient condensers	[33,50,52]

and pDNA result in the formation of significantly more size-stable condensed particles than a more frequently used condensation agent protamine [33]. The Mu peptide appears to be more versatile at pDNA condensation than protamine by virtue of forming stable complexes over a broader range of peptide:pDNA ratios. This observation could be attributed to the proline residues in the protamine backbone that may limit the degree of conformational freedom [40]. The enhanced flexibility and size stability of Mu–DNA complexes (80 – 100 nm) compared with protamine counterparts suggest that the Mu peptide may be an ideal motif for use as part of a new non-viral gene delivery vector. The possible usefulness of this peptide for future *in vivo* applications was proven by enhancing the transfection in differentiated and undifferentiated neuronal cells using cholesterol and cationic liposomes for complexation [41]. Mu was found to give an identical enhancement to cationic liposome-mediated transfection of ND7 cells as PLL or protamine sulfate. The ternary non-viral systems of the Mu peptide, pDNA and liposomes, however, proved that the condensing peptide Mu is unable to aid the trafficking and transcription process independently [41]. Despite having an excellent DNA-condensing ability, endosomal disruption and release into the cytosol remain the principal limitations of the Mu peptide.

2.1.2 *Transactivation protein peptide*

TAT is an 86 amino acid protein from the HIV-1 virus. It has been shown that the arginine-rich residues 48 – 60 (GRKKRRQRRRPPQ) are required for cellular and nuclear uptake [42]. TAT was found to translocate through the plasma membrane and reach the nucleus with the help of an embedded nuclear localization signal (NLS) to transactivate the HIV genome [43]. The cluster of basic amino acids has been assigned to this translocation activity and also the membrane fusogenic activity owing to its amphipathic characteristics [42]. Debate surrounds the mechanism of membrane translocation. It has been postulated that the cationic charge of TAT stimulates an ionic interaction with the anionic cell membrane, thus facilitating adsorption of the peptide. The phospholipid membrane reorganizes and forms a hydrophilic pocket around the TAT peptide akin to an inverted micelle. Such a mechanism has been demonstrated for the antennapedia peptide [44]. However, no direct evidence has been observed for the TAT peptide. Nevertheless, the cationic nature of the TAT seems ideal for the delivery of anionic nucleic acids. In a study by Wender *et al.*, it was shown that the deletion of an arginine residue in TAT significantly reduced cellular uptake, whereas the deletion of a lysine residue had no effect [45]. Uptake was also significantly reduced when any single cationic residue was replaced with alanine [45].

Other groups have also looked at the potential of TAT oligomers to condense DNA into stable particles and safely harbor DNA in the nucleus [46]. A high-molecular-mass form of poly-TAT has been used to improve gene delivery with particles being formed that decrease in size with increasing molecular mass [47]. Despite enhanced DNA condensation ability with poly-TAT, transfection experiments demonstrated that gene expression was significantly dependent on the addition of chloroquine. This suggests that poly-TAT undergoes endocytic cellular uptake and is unable to escape efficiently from endosomes into the cytosol. Although the addition of a hydrophobic amino acid residue on the N terminus such as tyrosine or phenylalanine significantly improves the DNA-condensing capabilities of TAT, it is clear

from the literature that TAT is not the optimal choice for condensing DNA and gene delivery by itself [48]. There are several other cell-penetrating peptides with similar properties to TAT that have been reviewed extensively elsewhere [49].

2.1.3 Histones

Histones are naturally occurring basic proteins involved in DNA packaging in cells. Histones are also known to have an in-built nuclear localization signal that facilitates their nuclear import [50]. The occurrence of both NLS and a string of lysine and arginine residues attributes dual functionality to histones for efficient gene delivery. Histonefection (mediating gene transfer by histones) has been studied previously using the full length of histones H1, H2, H3 and H4 as DNA-condensing agents (reviewed in [51]). However, full-length histones have not been efficient at gene delivery and have needed other helper agents such as polymers or liposomes [52].

In a process termed chromofection, reconstituted chromatin has been used to deliver pDNA to MCF-7 cells [53]. The authors indicate that the addition of chloroquine did not enhance the percentage of transfected cells, indicating a non-endocytotic internalization pathway. However, it is not clear why the effect of chloroquine on per cent transfected cells is reported, as addition of chloroquine does not have any significant impact on per cent transfection but rather total gene expression. As a result, interpretation of data appears somewhat confusing. Furthermore, there are no particle sizing data available to demonstrate the extent of DNA condensation, which is an important factor in the regulation of the particle uptake pathway [25]. Nonetheless, histones have proved to be useful biological motifs for gene delivery, although one concern with the use of histone-based vectors in gene therapy is their potential risk of inducing an immune response [51,54].

Among all histones, histone H1 has received more attention than the others as a gene delivery system [55]. Histone H1 can be subdivided into three distinct domains, namely the N-terminal nose, the globular domain and the basic C-terminal tail (**Figure 1**). Khadake and Rao demonstrated that the condensation property of histone H1 is localized to its C-terminal tail ATPKKSTKKTPKKAKK (residues 144 – 159) and this is sufficient to mimic the DNA condensation property of the full-length histone H1 [34]. Later, Bharath *et al.* demonstrated that apart from residues 144 – 159, there is a stretch of 18 amino acids within the C terminus that also contributes to DNA condensation (PAAAAGAKKAKSPKKAKA, residues 160 – 177) [56]. Theoretically, it would be possible therefore to incorporate either the 34mer sequence or the 16mer as an alternative functional DNA-condensing motif for gene delivery. However, further comparative size and stability analysis studies with other DNA-condensing motifs are necessary to identify the most efficient and suitable peptide for gene transfer.

There are many effective DNA-condensing peptides that produce stable nanometer size particles; however, for effective gene delivery there is still a need to overcome many cellular biological barriers. DNA condensers are limited by their inability to target cells and disrupt endosomes. To that end, the coupling of the DNA-condensing peptides to cell-penetrating peptides and/or endosomal-releasing peptides could greatly improve the levels of gene expression of such condensates in a variety of cell lines.

2.2 Membrane-destabilizing peptides

There are several natural (e.g., influenza fusogenic peptide HA2) [36] and synthetic amphipathic peptides such as GALA, KALA, JTS-1 and ppTG20 that are able to traverse membranes [57-59]. Some of these peptides switch to an α-helical configuration on endosomal acidification resulting in an interaction with the endosomal membrane, causing disruption and pore formation.

2.2.1 Anionic amphiphilic peptides

The presence of gluatamate residues and subsequent protonation under acidic conditions typically characterize anionic amphiphilic peptides. Most fusogenic peptides in gene delivery are derived from the sequence of the N terminus of the influenza virus hemagglutinin 2 protein (HA2). The 20-amino acid HA2 (GLFGAIAGFIENGWEGMIDG) forms an α-helix under acidic conditions and fuses with the endosomal membrane, thus facilitating escape of the cargo into the cytosol [36]. GALA is a synthetic derivative of the influenza fusogenic peptide with a Glu-Ala-Leu-Ala repeat that perturbs lipid membranes and forms an aqueous pore, thus facilitating cargo escape. GALA is 30 amino acids long (WEAALAEALAEALAEHLAEALAEALEALAA) and forms an α-helical conformation at a low pH of 5.0. This conformational change increases peptide hydrophobicity on one side and therefore enhances its interaction with lipid bilayers. At a physiological pH of 7.5, the repulsion in charge between aligned Glu destabilizes the α-helix and subsequently no membrane interaction occurs [58]. GALA cannot condense DNA, but can substantially increase the transfection efficiency of standard cationic liposomes and facilitate release to the cytosol. This observation was reported by Futaki *et al.*, where GALA was mixed with standard cationic liposomal transfection reagents resulting in a significant increase in cell transfection [60]. GALA has also been used successfully to enhance delivery of siRNA to HeLa cells in a new system termed R8/GALA-MEND (multifunctional envelope-type nano device) [61]. GALA appears to be an attractive helper motif to assist with siRNA delivery as the final destination is the cytosol.

2.2.2 Cationic amphiphilic peptides

As discussed above, GALA is a negatively charged peptide and is unable to bind to nucleic acids. To overcome this challenge, a bifunctional cationic peptide named KALA (WEAKLAKA-LAKALAKHLAKALAKALKACEA) was designed [59]. KALA consists of lysine-alanine-leucine–alanine repeats that result in a hydrophobic and hydrophilic region. The lysine binds the

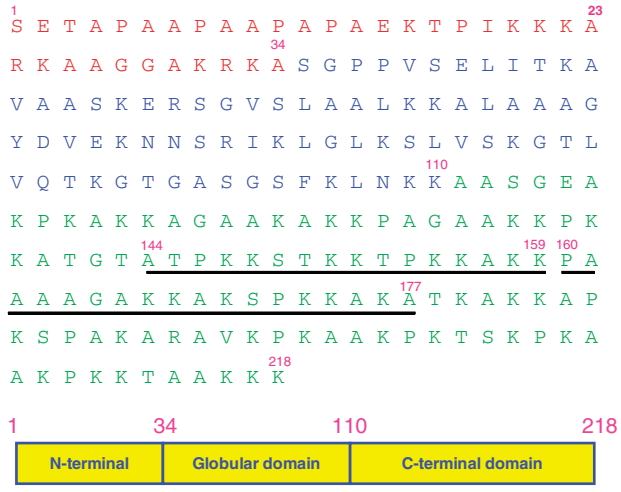

Figure 1. The amino acid sequence of histone H1 showing the different structural domains: N-terminal, globular and C-terminal.

Reproduced with permission from [56].

DNA and there are still enough glutamic residues to ensure pH-dependent solubility and protonation [59]. The KALA forms complexes with plasmid DNA and at a 10:1 ratio ~ 34% of K562 cells displayed a fluorescent nucleus, indicating potential for gene delivery [59]. The KALA peptide has also been used to assist DNA delivery to hepatocytes in a gal-PLL system, giving 30 times more luciferase activity than chloroquine [62]. In a comparative study to deliver PEGylated VEGF siRNA, KALA, PEI and PLL were analyzed for condensing, cytotoxicity and gene silencing properties in prostate cancer cells as an antiangiogenic therapy [63]. KALA had the lowest IC_{50} of all three condensers with particles < 200 nm and gave ~ 90% VEGF inhibition [63]. It has been shown that KALA is emerging as an important functional component of effective siRNA delivery that can function both as a nucleic acid condenser and as an endosomal disrupter.

Midoux *et al.* constructed another cationic derivative of the N-terminal HA2 peptide, namely H5WYG [64]. This peptide is designed to protonate under slightly acidic conditions but remains neutrally charged at the physiological pH. In this peptide, the glutamate residues were substituted with five histidine residues. Given that histidine has a pK_a value close to pH 6 and becomes cationic under slightly acidic conditions, the H5WYG does not permeabilize membranes under normal pH 7.4 but is more sensitive at disrupting the endosomal membranes than other peptides, which need a lower pH of 5.0. It is reported that the H5WYG undergoes a conformational change from beta to an unordered structure that disrupts the membranes starting at a pH of 6.8. Despite these peptides conferring effective endosomal disruption, the issue of gene trafficking to the cell nucleus remains a significant limitation in DNA delivery.

2.3 Nuclear localization signals

Many effective non-viral gene delivery systems have been developed to overcome the barriers of gene delivery to the cytoplasm. However, studies have shown that the cellular uptake of plasmid DNA does not correlate with efficient cell transfection [65]. Although there is limited understanding of the cellular and molecular mechanisms involved in synthetic vector-mediated gene transfer, transfection efficiency appears to be essentially limited by inefficient trafficking of DNA to the site of gene transcription in the nucleus [66]. One major reason why most plasmids fail to reach the nucleus is the presence of cytoplasmic nucleases that act to degrade the 'free' DNA [67]. Translocation of a protein from the cytoplasm into the nucleus is dependent on the presence of a short stretch of basic amino acid sequences known as the nuclear localization signal, which facilitates nuclear translocation of the protein DNA complex. The NLS mediates the binding of the protein carrier to members of the importin-α (karyopherin) family of proteins in the cytoplasm. This enables translocation through the nuclear pore in a GTP-dependent manner [68]. Inside the nucleus the complex dissociates to release the protein in the presence of RanGTP. There are variations of NLSs, with the simian virus 40 (SV40) large T-antigen NLS serving as the prototype of the 'classical' NLS [69]. Different NLSs bind to different members of the importin family. Some of the NLSs operate independently of the importin-α machinery, such as the M9 sequence [38].

Below only the three most commonly used NLSs are reviewed to highlight some important concepts to consider in non-viral gene delivery. There are several different types of NLS reported in the literature, which are reviewed elsewhere [70,71]. The nuclear transport of plasmid DNA with the help of importins is shown schematically in **Figure 2**.

2.3.1 *SV40 nuclear localization signal*

In gene delivery, various nuclear localization signals have been utilized to facilitate the translocation of DNA towards the nucleus. The SV40 large tumor antigen nuclear localization sequence is perhaps the most widely used NLS that contains mainly basic amino acid residues PKKKRKV [69]. Transport of the SV40 NLS is mediated by means of the classical α – β importin heterodimer pathway. Binding of the SV40 NLS to the importin machinery is extremely specific as a single mutation can impair nuclear import [72]. The SV40 NLS has been shown to confer karyophilic properties on other peptides.

In one approach, the NLS-DNA (plasmid, linearized plasmids and oligonucleotides) conjugates were used to form complexes by means of electrostatic interactions. Positively charged NLS peptides can interact electrostatically with DNA, forming complexes with multiple NLS moieties exposed on the surface. However, such weak interactions can lead to easy disassembly of these complexes in the cytoplasm or the NLS can become embedded within the DNA [73]. To enhance stability of the DNA–NLS complex, several groups have used

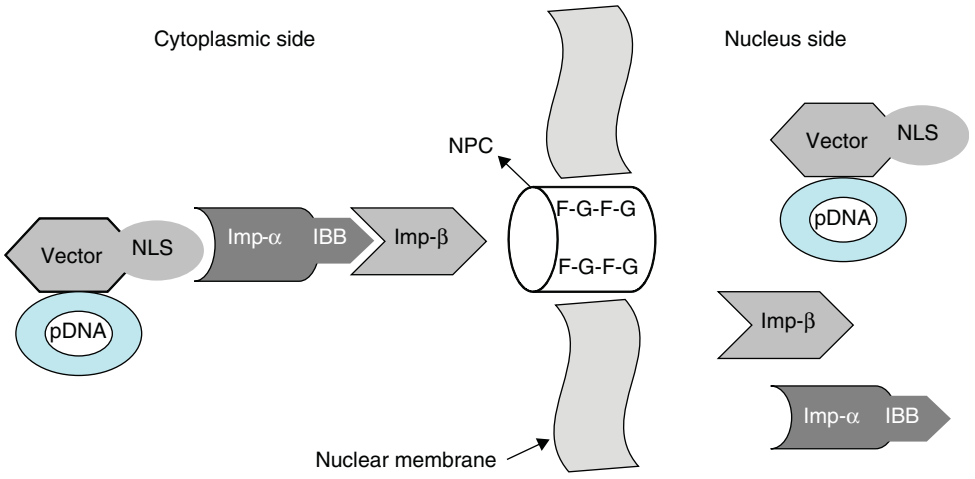

Figure 2. The NLS-dependent nuclear import of plasmid DNA is shown schematically. The NLS-recognizing importin (IMP)-α/β proteins are also shown. IMP-β binds the importin-β binding (IBB) domain of IMP-α to form the IMP-α/β heterodimer. Once docked to the NPC through IMP-β, it binds to specific nucleoporins (Nups) on the cytoplasmic side of the NPC. The translocation of the importin/cargo complex through the NPC involves transient association/disassociation interactions of IMP-β with the phe-gly (F-G) repeats of Nups throughout the NPC central channel [102].

NLS: Nuclear localization signal; NPC: Nuclear pore complex.

synthetic chemical methods to bind the NLS to DNA covalently (reviewed in [65]).

In a study by Hariton-Gazal *et al.*, the SV40 NLS was conjugated to the $S4_{13}$ peptide at either the N or the C terminus and nuclear localization occurred [74]. However, when the SV40 NLS was covalently linked to linear DNA and delivered with cationic polymers there was no improvement in transfection efficacy in a range of cell lines [75]. It is possible that the basic residues in the NLS are binding to DNA through electrostatic interactions and thus reducing potential binding to the importin proteins [72]. Interestingly, when PEG spacers were added between plasmid DNA and the SV40 NLS, a significant increase in protein expression proportional to the length of spacer and number of NLS inserts was observed [76]. However, the increase in gene expression was only modest, possibly owing to the size of the pDNA and the nuclear pore [76]. The SV40 NLS conjugated to plasmids may succeed only if the pDNA is sufficiently condensed through another motif. At best, the SV40 NLS was found to condense only 40% pDNA and studies confirmed that when in tandem with protamine, the latter was responsible for condensation of pDNA [77]. Nevertheless, the construction of gene carriers containing both a pDNA-condensing motif and a NLS may still be compromised by the inefficient binding of NLS to importin proteins.

In a study by Zanta *et al.* [78], the effect of the number of NLS peptides attached to pDNA was studied and it was concluded that 'a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus'. This study is discussed because it has resulted in some misunderstanding among scientists. In this study, a linear piece of DNA was developed by digesting a plasmid with restriction enzymes, and subsequently capping it with hairpin oligonucleotides [78]. This linearized DNA was conjugated to a single SV40 NLS peptide (PKKKRKVEDPYC) and used to transfect dividing and non-dividing cells with the help of PEI (25 kDa) or Transfectam. Transfection efficiency of linearized DNA coupled with the NLS was 10- to 1000-fold higher than the linearized DNA without NLS or the linearized DNA coupled to mutated NLS. The authors state that NLS played a significant role in facilitating the translocation of DNA to the cell nucleus by means of the nuclear pore complex by utilizing the cell's importin machinery. This conclusion was drawn without examining the nuclear transport of the DNA–NLS/PEI complexes using specific tests such as the nuclear import assay. Surprisingly, the greatest increase in transfection efficiency was obtained with rapidly growing cells such as NIH-3T3 or HeLa, whereas the slow or non-dividing neuronal or macrophage cells displayed a weak increase. Assuming that the single NLS conjugated to the linearized DNA had played a role in translocation of DNA to the cell nucleus via the importin machinery, no such substantial difference in transfection efficiency between dividing and non-dividing cells should have been observed. In addition, in this study, owing to the use of PEI as the major DNA-condensing agent, in each DNA–NLS/PEI complex more than one NLS exists, which undermines the validity of the theory, which states 'a single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus'. It is notable that PEI remains complexed with the DNA in the cytoplasm and after transport into the nucleus [79,80]. Therefore, drawing the conclusion that one NLS is sufficient to carry DNA to the cell nucleus is premature at this point. Recent studies by other groups have also shown that there is no enhancement of nuclear entry by direct

conjugation of a single nuclear localization signal peptide to linearized DNA [81]. Based on the literature, it is apparent that the major drawback to the covalent coupling includes reduction in transcriptional activity resulting from chemical modifications in the transcription cassette. The evidence strongly suggests that an effective NLS peptide should be attached to another DNA binding peptide and not the plasmid itself, which is similar to viruses' function.

2.3.2 *Rev nuclear localization signal*

The Rev protein originates from the human immunodeficiency virus and is integral in regulating the transport of structural mRNAs from the cytoplasm to the nucleus [82]. The nuclear transport Rev peptide is 15 amino acids and consists of an arginine-dominant sequence RQARRNRRNRRRRWR. Cochrane *et al.* were among the leading groups that demonstrated the role of Rev signal peptide in facilitating the accumulation of a tagged peptide (luciferase) in the cell nucleus [37]. Later, others attached the Rev peptide covalently to the $S4_{13}$ peptide derived from dermaseptin S4 [74]. Dermaseptin peptides are known to penetrate cells by means of their interaction with phospholipids [83]. However, despite its small size, $S4_{13}$ was unable to penetrate through the nuclear pore and accumulated in the cytoplasm [74]. When Rev was attached to $S4_{13}$, nuclear accumulation of the peptide occurred in intact, permeabilized and microinjected HeLa cells [74]. These studies among others highlight the potential applications of Rev in designing non-viral vectors for gene delivery.

2.3.3 *M9 nuclear localization signal*

The M9 peptide with the amino acid sequence of GNYNNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY is another nuclear localization signal derived from the heterogeneous nuclear ribonuclear protein (hnRNP) A1. As a major pre-mRNA binding protein, hnRNP A1 has the ability to shuttle between the nucleus and cytoplasm [84]. Despite not having a classical NLS, the sequence of A1 that localizes to the nucleus was identified and termed M9 [38]. It was shown that transportin is responsible for M9 nuclear localization, which is independent of the traditional importin pathway to the nucleus [85]. The M9 38 amino acid sequence lies at the carboxyl end of the protein and does not contain the typical NLS basic residue repeats [38]. Conjugation of M9 to a 13-amino acid cationic peptide (ScT) to deliver pCMVβ-gal gave transfection efficiencies of 76.4% *in vitro* in bovine aortic endothelial cells [86]. This compares favorably with 5.27 and 55% transfection efficacy for pDNA and pDNA-M9, respectively [86]. Similar results were also found for M9-ScT in fibroblasts and neuronal cells, and in fact the M9 was found to be more effective than the SV40 NLS in primary neurons [87]. However, in these experiments lipofectamine was also required to disrupt the endosomes.

There can be no doubt that transport across the nuclear membrane remains the major rate-limiting step to effective gene delivery and NLS peptides can help to overcome this barrier. With the surge of new viruses and molecular studies it could be anticipated that a plethora of new NLSs may be identified and incorporated into new delivery systems.

3. Recombinant bio-inspired vectors for gene delivery

So far, the utilization of several biological motifs (i.e., DNA condensing, endosome disrupting and NLS) and their derivatives to deliver exogenes into mammalian cells have been discussed. None of these motifs has the ability to overcome independently all of the cellular barriers associated with successful targeted gene therapy. With the emergence of recombinant DNA technology as a powerful tool to create fusion peptides, several groups have attempted to design and develop multifunctional vectors for gene delivery. The inspiration for designing such multifunctional biomimetic vectors comes from viruses that can efficiently perform several functions to overcome many biological barriers. Bio-inspired vectors contain multiple sequences that perform discrete functions necessary for effective DNA transport. One of the major problems when using peptides with multiple functions to complex DNA is in the production of homogenous particles. Tecle *et al.* suggest that one way forward could be to concatenate several cationic peptides molecules with several functions, for example, a Mu-condensing motif with NLS, and that such a molecular scaffold could make possible a more controlled condensation of pDNA [88].

Overell was among the leading investigators using recombinant DNA technology to create bifunctional non-viral vectors based on biological motifs [89]. His group genetically engineered a protein vector consisting of the DNA-binding domain of the yeast transcription factor GAL4 fused to the cell binding/internalization domain termed invasin. The vector was expressed in *Escherichia coli*, purified and used to bind to pDNA containing a reporter gene. The results showed that the vector could transfect cells in an invasin-dependent manner but failed to condense pDNA. This shortcoming was overcome by using PLL as a DNA-condensing motif. It was also observed that addition of chloroquine significantly increased the gene expression levels of GAL4/invasin, but only twofold. Therefore, it was postulated that GAL4 may have some endosomal-disrupting activity. Although there are several inherent weaknesses associated with the design of the vector (e.g., inability to condense pDNA, lack of NLS and inefficient endosome-disrupting activity), it does demonstrate the potential utility and appeal of such an approach to create a modular non-viral vector. The advantage of such an approach lies in the ability to control the nature of the vector constituents and the ease of production.

Later, Wels's group used a similar recombinant approach to create multifunctional non-viral vectors based on biological motifs. In one approach, they created a chimeric peptide composed of the DNA-binding domain of the yeast GAL4 transcription factor, the translocation domain of *Pseudomonas*

exotoxin A as an endosome-disrupting motif and the epidermal growth factor (EGF) as a targeting motif [90]. The results of the study demonstrated that the vector could target cells overexpressing EGF receptor but lacked the ability to condense pDNA. Similar to Overell's approach, PLL was used to overcome this obstacle and full DNA condensation was achieved. Unfortunately, no data on the particle size or the extent of DNA neutralization were provided. Recent literature shows that in targeted gene delivery, particle size plays a significant role in determining the internalization pathway [25,91]. In addition, the efficiency of the endosome-disrupting motif in this fusion vector appeared to be suboptimal as addition of chloroquine resulted in a significant increase in transfection efficiency. Perhaps conducting a hemolysis assay could provide more information regarding the activity of the endosomal disruption motif.

The same group also developed another fusion peptide, namely GD5 [92]. More specifically, GD5 has an N-terminal Gal4 DNA binding region, a diphtheria toxin (DT) translocation domain and a C-terminal ErbB2 antibody fragment (scFv(FRP5)) that binds specifically to tumors that express the ErbB2 antigen [92]. Despite showing effective DNA binding and cell specificity, the GD5 fusion protein also required synthetic PLL to facilitate DNA neutralization and condensation. The addition of chloroquine did enhance gene expression twofold despite GD5 having the translocation domain of the diphtheria toxin to facilitate endosomal escape. These studies highlight the fact that when creating multifunctional gene delivery systems, utmost care is required to ensure that all the motifs retain their activity, as one domain could interfere with the other and render the vector or part of it inactive. Owing to the shortcomings of these vectors coupled with low-to-modest levels of gene expression, enthusiasm for the development of recombinant fusion vectors dampened, resulting in a limited number of research reports in this field.

With the advancements in recombinant DNA technology in the last decade and the availability of simple kits/protocols for gene cloning and protein expression, it appears that interest in the design and development of fusion bio-inspired vectors has rekindled. In 2007, Gopal's group reported the construction of two recombinant vectors comprised of Mu-Mu and TAT-Mu peptides. They used these vectors to complex with pDNA (pCMVβ-gal) and transfect MCF-7, COS and CHOK-1 cells [93]. Using the gel retardation assay and the fluorescence quenching by ethidium bromide (EtBr) exclusion assay, the ability of the vector to condense pDNA was confirmed. However, no particle size and charge study was performed to evaluate these two important parameters, which can significantly impact on cellular uptake. Although the vectors were able to transfect cells modestly, they were unable to escape efficiently from the endosomes. Later, the same group reported the structure of a recombinant multifunctional vector that consisted of three domains: a cell-penetrating (TAT) domain; a nuclear localization domain comprising three copies of the SV40 NLS; and a DNA-condensing (Mu) domain [94]. This vector was used to transfect various cell lines such as MCF-7, COS, CHO and HepG2. Based on the fluorescence quenching by EtBr exclusion assay and atomic force microscopy, it was shown that the vector was able to condense pDNA into nanosize particles. However, a study was not performed to measure particle size distribution, surface charge or reproducibility of obtaining uniform particles. Furthermore, the transfection studies in the presence of chloroquine revealed a significant increase in transfection efficiency, again indicating the inability of the vector to escape from the endosomes. Although no clear explanation is provided for the use of three tandem repeating units of NLS, the overall design of the vector is new and interesting. Perhaps the addition of a fusogenic peptide sequence along with a targeting peptide could result in a more powerful vector for gene delivery. Targeting peptides that could be utilized in future studies to enhance cell specificity of these systems are mentioned in the references [95].

Wang *et al.* reported the structure of a recombinant multifunctional vector that pushed the boundaries of fusion bio-inspired gene delivery systems even further [96]. This vector, namely the designer biomimetic vector, is a genetically engineered biomacromolecule that is designed to mimic viral characteristics in order to overcome the cellular barriers associated with the targeted gene transfer. It is composed of several discrete motifs, each with a single function. The vector was genetically engineered to contain: i) four tandem repeating units of N-terminal domain of histone H2A peptide (4HP); ii) a synthetic single chain high-affinity HER2 targeting motif (TM); iii) a fusogenic peptide (FP) named GALA; and iv) a cathepsin D substrate (CS). The cathepsin D cleavage site was positioned between the 4HP and TM motifs to allow for dissociation of the targeting motif from the vector inside the late endosomes. It was also mentioned that the histone H2A (residues 1 – 37) was deliberately chosen because it has an inherent NLS [97]. The theoretical model for the intracellular trafficking of the vector is shown schematically in **Figure 3**. The results of the study showed that the vector has the ability to target SKOV-3 ovarian cancer cells via HER2, disrupt endosome membranes efficiently at low pH, utilize the microtubules to reach the nucleus and mediate gene expression. In the range tested, no cell toxicity was observed. The design of this vector appears to be the first of its kind where a single fusion peptide can perform four discrete functions (i.e., DNA condensing, cell targeting, endosomal disruption and microtubule-mediated nuclear localization). Despite the vector's new design, more in-depth studies are required to determine its exact intracellular trafficking pathway. Also, the vector's ability to transfect non-dividing cells where the nuclear membrane remains intact is undetermined.

The same group also reported the structure of another new biomimetic vector composed of: i) a DNA-condensing motif from Mu peptide; ii) a cyclic targeting peptide to target ZR-75-1 breast cancer cells; iii) an endosome-disrupting motif, namely H5WYG; and iv) a Rev NLS to facilitate translocation

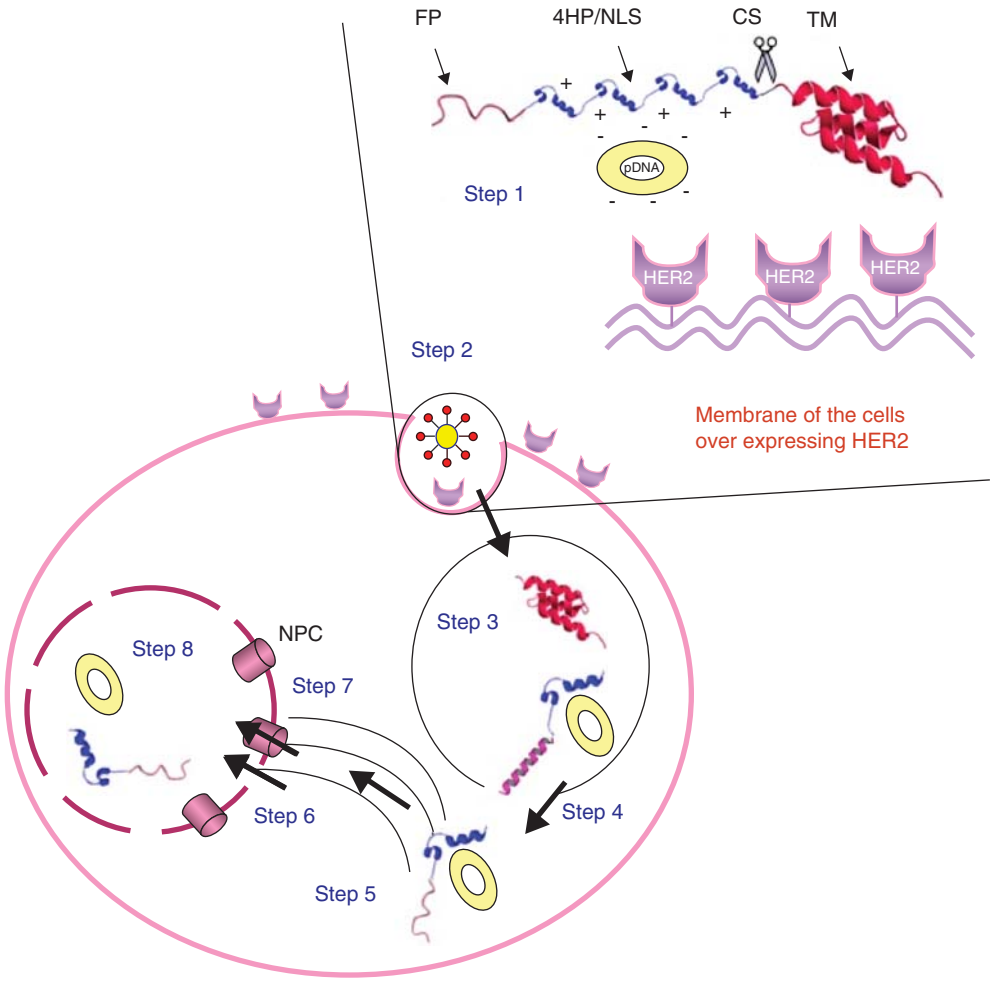

Figure 3. The schematics of the multidomain biomimetic vector and its intracellular trafficking. The vector is composed of a fusogenic peptide (FP), four tandem repeating units of histone H2A (4HP) with an inherent nuclear localization signal (NLS), a cathepsin D substrate (CS) and a HER2 targeting motif (TM). The three-dimensional structures of the histone H2A and TM are predicted using the SWISS-MODEL program. Step 1: the positively charged 4HP interacts with pDNA and forms condensed nanosize particles. Step 2: the targeting motif binds to HER2 overexpressed on the surface of cancer cells allowing the internalization of the complexes by means of receptor-mediated endocytosis. Step 3: TM separates from the complexes inside late endosomes with the help of endogenous cathepsin D enzyme. Step 4: the fusogenic peptide changes conformation at acidic pH into α-helical structure, fuses with the endosome membrane, and facilitates escape of the vector/pDNA complex into the cytosol. Step 5: the NLS motif in the Histone H2A structure binds to the microtubules and shuttles the pDNA towards the nucleus. Step 6: if the size of the vector/pDNA complexes is > 30 nm, the complex may end up in the nucleus at the mitosis (M) phase of the cell cycle where the nuclear membrane dissolves. Step 7: if small enough (< 30 nm), the complex may pass through the nuclear pore complex (NPC). Step 8: inside the nucleus, the pDNA will be released for transcription. The unpacking of the complexes may occur in the nucleus with the help of transcription factors.
The structure of the vector shown in this figure is reproduced with permission from [96].

of the genetic material towards the nucleus [98]. In this vector, H5WYG-related cell membrane-disrupting activity was observed only at a pH of 5.0, not at a physiological pH of 7.4. Particle sizes of 60 – 70 nm with pEGFP were formed and complex stability was demonstrated in the presence of serum. Furthermore, the targeting peptide was shown to be functional, with no increase in transfection following addition of chloroquine. The vector was able to transfect on average 18% of the ZR-75-1 cells and the results showed that up to 62% of the breast cancer cells were killed when pDNA encoding TRAIL was delivered. This vector also did not show any vector-related toxicity in the range tested.

Other recombinant non-viral vectors based on the fusion of cationic biopolymers and biological motifs have been reported by Hatefi's group [99,100]. The first report on the fusion of a recombinant cationic biopolymer with a natural motif (i.e., fibroblast growth factor 2) for non-viral gene delivery was by Hatefi and Ghandehari, which is discussed elsewhere [101]. As recombinant gene delivery systems can be bioengineered at the molecular level, a library of versatile vectors equipped with

different targeting motifs can be generated to deliver therapeutic genes to various target tissues. Issues of particle shielding to minimize potential immunogenicity and *in vivo* protection will most probably have to be addressed for these recombinant systems in future studies.

4. Conclusion

Recombinant DNA technology has enabled the creation of a variety of multifunctional vectors based on natural motifs that can mimic viruses more accurately than ever before. These bio-inspired vectors can be decorated with a variety of cell-targeting peptides and customized for various gene therapy needs. As we gain more information about the exact mechanism of viral cell infection, the design of the non-viral vectors becomes more sophisticated to mimic viruses more accurately. Recent publications show that scientists are on the right path and it may not be long before we see disease- and patient-specific gene delivery systems in the market.

5. Expert opinion

Considering the volume of publications in the field of vector design for gene therapy, it is apparent that there is a race towards the development of efficient, targeted, safe and cost-effective gene delivery systems. While viral gene therapists are trying to enhance safety and targetability without compromising efficiency, non-viral vector developers strive to improve the efficiency without compromising safety. Considering the eons that viruses had to evolve into superior gene delivery machines, non-viral vector designers need to devise more efficient methods that could exponentially accelerate the evolutionary process of their systems.

So far, the only polymer-based non-viral system that has shown some degree of success is PEI. Many groups have tried to modify the structure of PEI to reduce its short-term toxicity. However, the long-term toxicity resulting from the non-biodegradable nature of PEI has not been adequately addressed. It is perhaps time to move forward (as opposed to laterally) towards the synthesis and development of new biocompatible biomaterials rather than trying to make incremental advances by PEI modifications. For polymeric systems to be considered a viable technology for gene therapy, they not only should be safe and efficient but also should allow manipulation of their architecture at the molecular level. Without the ability to control architecture at that level, there is a very low probability that revolutionary advances will be made. This shortcoming stems mostly from the limited ability of the chemical synthetic methods to yield monodispersed products, which are essential for reliable structure/activity relationship studies.

An alternative to polymeric systems are liposomes, which appear to be more reliable than any other non-viral vector for gene therapy. This is by and large because of their relatively high rate of gene transfer efficiency, biodegradability and cost-effectiveness. However, the potential use of lipid-based systems for *targeted* gene transfer, which requires attachment of targeting ligands, is still under question with limited control over the reproducibility of the manufacturing process (i.e., significant batch-to-batch variations).

Immunogenicity is also another major concern that needs to be studied with diligence. It is well understood that the introduction of any non-self material (peptide, polymer, etc.) into the body could elicit an immune system response. The level of response to eliminate the foreign object from the body is dependent on the dose, sequence, size and pre-existing immunity. Decoration of viral or non-viral systems with presumably inert materials such as PEG has shown some promise; however, recent data show that the concept of biological inertness is not an absolute and even PEG is immunogenic. Therefore, it may no longer be necessary for the ideal gene delivery system to be inert *per se*, but rather one that can interact with the biological system without upregulating cellular signal transduction pathways. A good example is the HIV virus, which is able to interact with the immune system through its Nef peptide and downregulate the expression of CD4 (HIV receptor) on MHC I. Using similar strategies or the addition of motifs from human origins at *precise* locations in the vector backbone could also be an attractive approach to keeping the elements of the immune system dormant.

To overcome these challenges, not only significant research and innovation are required, but also an investment in new emerging technologies. Based on what can be found in the literature, there is one technology that stands out and has shown significant potential to overcome the obstacles discussed in this review. Recombinant DNA technology has made it possible to create fusion proteins with numerous minuscule variations that continuously affect gene expression. Based on this theory and recombinant DNA technology, the combinations of vectors composed of multiple functional peptide motifs are endless. Such recombinant fusion vectors (biomacromolecules) can be programmed by means of their amino acid sequences to perform several self-guided functions. Most importantly, such vectors can evolve by a combinatorial screening process that will accelerate the identification of the structure of high-performing constructs. Although at an early stage, published data show that such vectors can be designed to be biodegradable, non-toxic, targeted and relatively efficient (at least at the *in vitro* level). In addition, they can be used as a powerful tool to study accurately the rate-limiting steps to non-viral gene transfer (structure/activity relationship studies), which could ultimately lead to the development of safe and efficient vectors. The next generation of non-viral vectors must be multifunctional with ability to target specific cells and produce gene expression levels to rival viruses without the associated toxicity issues. There is a considerable repertoire of functional peptides available in nature. Given that there is only a handful of scientists working on recombinant fusion vectors, the authors hope that both viral and non-viral gene

therapists take a closer look at this emerging technology and contribute to its evolution.

Acknowledgments

The studies in A Hatefi's lab are supported by the American Cancer Society, and Department of Defense Breast and Prostate Cancer Programs. The studies in HO McCarthy's lab are supported by the Breast Cancer Campaign UK.

Declaration of interest

The authors have received no payment in the preparation of this manuscript and critical assessment of the technologies.

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Mulligan RC. The basic science of gene therapy. Science 1993;260(5110):926-32
2. Louise C. Nonviral vectors. Methods Mol Biol 2006;333:201-26
3. Thomas CE, Ehrhardt A, Kay MA. Progress and problems with the use of viral vectors for gene therapy. Nat Rev Genet 2003;4(5):346-58

• **This review is of particular interest because it provides a comprehensive look at the use of viral vectors for gene therapy.**

4. Waehler R, Russell SJ, Curiel DT. Engineering targeted viral vectors for gene therapy. Nat Rev Genet 2007;8(8):573-87
5. Hatefi A, Canine BF. Perspectives in vector development for systemic cancer gene therapy. Gene Ther Mol Biol 2009;13(A):15-9

• **This review highlights the important factors that need to be considered when designing vectors for systemic cancer gene therapy.**

6. Cronin J, Zhang X, Reiser J. Altering the tropism of lentiviral vectors through pseudotyping. Curr Gene Ther 2005;5(4):387-98
7. Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, et al. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science 2000;288(5466):669-72
8. Hacein-Bey-Abina S, von Kalle C, Schmidt M, et al. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med 2003;348(3):255-6
9. Flotte TR. Recent developments in recombinant AAV-mediated gene therapy for lung diseases. Curr Gene Ther 2005;5(3):361-6
10. Mingozzi F, High KA. Immune responses to AAV in clinical trials. Curr Gene Ther 2007;7(5):316-24
11. Makower D, Rozenblit A, Kaufman H, et al. Phase II clinical trial of intralesional administration of the oncolytic adenovirus ONYX-015 in patients with hepatobiliary tumors with correlative p53 studies. Clin Cancer Res 2003;9(2):693-702
12. Bostanci A. Gene therapy. Blood test flags agent in death of Penn subject. Science 2002;295(5555):604-5
13. Schnell MA, Zhang Y, Tazelaar J, et al. Activation of innate immunity in nonhuman primates following intraportal administration of adenoviral vectors. Mol Ther 2001;3(5 Pt 1):708-22
14. Thummala NR, Ghosh SS, Lee SW, et al. A non-immunogenic adenoviral vector, coexpressing CTLA4Ig and bilirubin-uridine-diphosphoglucuronat eglucuronosyltransferase permits long-term, repeatable transgene expression in the Gunn rat model of Crigler-Najjar syndrome. Gene Ther 2002;9(15):981-90
15. Shen Y, Nemunaitis J. Herpes simplex virus 1 (HSV-1) for cancer treatment. Cancer Gene Ther 2006;13(11):975-92
16. Green NK, Morrison J, Hale S, et al. Retargeting polymer-coated adenovirus to the FGF receptor allows productive infection and mediates efficacy in a peritoneal model of human ovarian cancer. J Gene Med 2008;10(3):280-9
17. Herold BC, Visalli RJ, Susmarski N, et al. Glycoprotein C-independent binding of herpes simplex virus to cells requires cell surface heparan sulphate and glycoprotein B. J Gen Virol 1994;75(Pt 6):1211-22
18. Argnani R, Boccafogli L, Marconi PC, et al. Specific targeted binding of herpes simplex virus type 1 to hepatocytes via the human hepatitis B virus preS1 peptide. Gene Ther 2004;11(13):1087-98
19. Limbach KJ, Richie TL. Viral vectors in malaria vaccine development. Parasite Immunol 2009;31(9):501-19
20. Gao X, Kim KS, Liu D. Nonviral gene delivery: what we know and what is next. AAPS J 2007;9(1):E92-104
21. Li S, Tseng WC, Stolz DB, et al. Dynamic changes in the characteristics of cationic lipidic vectors after exposure to mouse serum: implications for intravenous lipofection. Gene Ther 1999;6(4):585-94
22. Ambegia E, Ansell S, Cullis P, et al. Stabilized plasmid-lipid particles containing PEG-diacylglycerols exhibit extended circulation lifetimes and tumor selective gene expression. Biochim Biophys Acta 2005;1669(2):155-63
23. Ishida T, Ichihara M, Wang X, et al. Injection of PEGylated liposomes in rats elicits PEG-specific IgM, which is responsible for rapid elimination of a second dose of PEGylated liposomes. J Control Release 2006;112(1):15-25

• **This article highlights the potential immunogenicity of PEG in animal models.**

24. Judge A, McClintock K, Phelps JR, et al. Hypersensitivity and loss of disease site targeting caused by antibody responses to PEGylated liposomes. Mol Ther 2006;13(2):328-37

• **This article highlights the potential immunogenicity of PEG in animal models.**

25. Rejman J, Oberle V, Zuhorn IS, et al. Size-dependent internalization of particles via the pathways of clathrin- and caveolae-mediated endocytosis. Biochem J 2004;377(Pt 1):159-69
26. Boussif O, Lezoualc'h F, Zanta MA, et al. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. Proc Natl Acad Sci USA 1995;92(16):7297-301
27. Fischer D, Li Y, Ahlemeyer B, et al. In vitro cytotoxicity testing of polycations: influence of polymer structure on cell

viability and hemolysis. Biomaterials 2003;24(7):1121-31

28. Lv H, Zhang S, Wang B, et al. Toxicity of cationic lipids and cationic polymers in gene delivery. J Control Release 2006;114(1):100-9
29. Fenske DB, MacLachlan I, Cullis PR. Long-circulating vectors for the systemic delivery of genes. Curr Opin Mol Ther 2001;3(2):153-8
30. Martin ME, Rice KG. Peptide-guided gene delivery. AAPS J 2007;9(1):E18-29
31. Pack DW, Hoffman AS, Pun S, et al. Design and development of polymers for gene delivery. Nat Rev Drug Discov 2005;4(7):581-93
32. Holm T, Johansson H, Lundberg P, et al. Studying the uptake of cell-penetrating peptides. Nat Protoc 2006;1(2):1001-5
33. Keller M, Tagawa T, Preuss M, et al. Biophysical characterization of the DNA binding and condensing properties of adenoviral core peptide mu. Biochemistry 2002;41(2):652-9
34. Khadake JR, Rao MR. Condensation of DNA and chromatin by an SPKK-containing octapeptide repeat motif present in the C-terminus of histone H1. Biochemistry 1997;36(5):1041-51
35. Dufourc EJ, Smith IC, Dufourcq J. Molecular details of melittin-induced lysis of phospholipid membranes as revealed by deuterium and phosphorus NMR. Biochemistry 1986;25(21):6448-55
36. Durrer P, Galli C, Hoenke S, et al. H+-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region. J Biol Chem 1996;271(23):13417-21
37. Cochrane AW, Perkins A, Rosen CA. Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. J Virol 1990;64(2):881-5
38. Siomi H, Dreyfuss G. A nuclear localization domain in the hnRNP A1 protein. J Cell Biol 1995;129(3):551-60
39. Murray KD, Etheridge CJ, Shah SI, et al. Enhanced cationic liposome-mediated transfection using the DNA-binding peptide mu (mu) from the adenovirus core. Gene Ther 2001;8(6):453-60
40. Wuthrich K, Billeter M, Braun W. Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances. J Mol Biol 1984;180(3):715-40
41. Keller M, Harbottle RP, Perouzel E, et al. Nuclear localisation sequence templated nonviral gene delivery vectors: investigation of intracellular trafficking events of LMD and LD vector systems. Chembiochem 2003;4(4):286-98
42. 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(25):16010-7
43. Truant R, Cullen BR. The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. Mol Cell Biol 1999;19(2):1210-7
44. Thoren PE, Persson D, Karlsson M, et al. The antennapedia peptide penetratin translocates across lipid bilayers–the first direct observation. FEBS Lett 2000;482(3):265-8
45. Wender PA, Mitchell DJ, Pattabiraman K, et al. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. Proc Natl Acad Sci USA 2000;97(24):13003-8
46. Rudolph C, Plank C, Lausier J, et al. 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(13):11411-8
47. Soundara Manickam D, Bisht HS, Wan L, et al. Influence of TAT-peptide polymerization on properties and transfection activity of TAT/DNA polyplexes. J Control Release 2005;102(1):293-306
48. Goparaju GN, Satishchandran C, Gupta PK. The effect of the structure of small cationic peptides on the characteristics of peptide-DNA complexes. Int J Pharm 2009;369(1-2):162-9
49. Foged C, Nielsen HM. Cell-penetrating peptides for drug delivery across membrane barriers. Expert Opin Drug Deliv 2008;5(1):105-17
50. Baake M, Doenecke D, Albig W. Characterisation of nuclear localisation signals of the four human core histones. J Cell Biochem 2001;81(2):333-46
51. Kaouass M, Beaulieu R, Balicki D. Histonefection: novel and potent non-viral gene delivery. J Control Release 2006;113(3):245-54
52. Kott M, Haberland A, Zaitsev S, et al. A new efficient method for transfection of neonatal cardiomyocytes using histone H1 in combination with DOSPER liposomal transfection reagent. Somat Cell Mol Genet 1998;24(4):257-61
53. Wagstaff KM, Fan JY, De Jesus MA, et al. Efficient gene delivery using reconstituted chromatin enhanced for nuclear targeting. FASEB J 2008;22(7):2232-42
54. Monestier M. Autoantibodies to nucleosomes and histone-DNA complexes. Methods 1997;11(1):36-43
55. Jung HJ, Hwang DS, Wei QD, et al. Carassius auratus-originated recombinant histone H1 C-terminal peptide as gene delivery material. Biotechnol Prog 2008;24(1):17-22
56. Bharath MM, Ramesh S, Chandra NR, et al. Identification of a 34 amino acid stretch within the C-terminus of histone H1 as the DNA-condensing domain by site-directed mutagenesis. Biochemistry 2002;41(24):7617-27
57. Rittner K, Benavente A, Bompard-Sorlet A, et al. New basic membrane-destabilizing peptides for plasmid-based gene delivery in vitro and in vivo. Mol Ther 2002;5(2):104-14
58. Li W, Nicol F, Szoka FC Jr. GALA: a designed synthetic pH-responsive amphipathic peptide with applications in drug and gene delivery. Adv Drug Deliv Rev 2004;56(7):967-85
- **This article highlights the amino acid sequences of a few fusogenic peptides that are used to disrupt endosome membranes.**
59. Wyman TB, Nicol F, Zelphati O, et al. Design, synthesis, and characterization of a cationic peptide that binds to nucleic acids and permeabilizes bilayers. Biochemistry 1997;36(10):3008-17
60. Futaki S, Masui Y, Nakase I, et al. Unique features of a pH-sensitive fusogenic peptide that improves the transfection efficiency of cationic liposomes. J Gene Med 2005;7(11):1450-8
61. Sakurai Y, Hatakeyama H, Akita H, et al. Efficient short interference RNA delivery to tumor cells using a combination of octaarginine, GALA and tumor-specific, cleavable polyethylene glycol system. Biol Pharm Bull 2009;32(5):928-32

62. Han J, Il Yeom Y. Specific gene transfer mediated by galactosylated poly-L-lysine into hepatoma cells. Int J Pharm 2000;202(1-2):151-60

63. Lee SH, Kim SH, Park TG. Intracellular siRNA delivery system using polyelectrolyte complex micelles prepared from VEGF siRNA-PEG conjugate and cationic fusogenic peptide. Biochem Biophys Res Commun 2007;357(2):511-6

64. Midoux P, Kichler A, Boutin V, et al. Membrane permeabilization and efficient gene transfer by a peptide containing several histidines. Bioconjug Chem 1998;9(2):260-7

65. Escriou V, Carriere M, Scherman D, et al. NLS bioconjugates for targeting therapeutic genes to the nucleus. Adv Drug Deliv Rev 2003;55(2):295-306

66. Suh J, Wirtz D, Hanes J. Efficient active transport of gene nanocarriers to the cell nucleus. Proc Natl Acad Sci USA 2003;100(7):3878-82

67. Dean DA, Strong DD, Zimmer WE. Nuclear entry of nonviral vectors. Gene Ther 2005;12(11):881-90

68. Moroianu J, Blobel G, Radu A. Previously identified protein of uncertain function is karyopherin alpha and together with karyopherin beta docks import substrate at nuclear pore complexes. Proc Natl Acad Sci USA 1995;92(6):2008-11

69. Lanford RE, Butel JS. Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. Cell 1984;37(3):801-13

70. Chan CK, Jans DA. Using nuclear targeting signals to enhance non-viral gene transfer. Immunol Cell Biol 2002;80(2):119-30

71. Miller AM, Dean DA. Tissue-specific and transcription factor-mediated nuclear entry of DNA. Adv Drug Deliv Rev 2009;61(7-8):603-13

72. Bremner KH, Seymour LW, Logan A, et al. Factors influencing the ability of nuclear localization sequence peptides to enhance nonviral gene delivery. Bioconjug Chem 2004;15(1):152-61

73. Ludtke JJ, Zhang G, Sebestyen MG, et al. A nuclear localization signal can enhance both the nuclear transport and expression of 1 kb DNA. J Cell Sci 1999;112 (Pt 12):2033-41

74. Hariton-Gazal E, Feder R, Mor A, et al. Targeting of nonkaryophilic cell-permeable peptides into the nuclei of intact cells by covalently attached nuclear localization signals. Biochemistry 2002;41(29):9208-14

75. van der Aa M, Koning G, van der Gugten J, et al. Covalent attachment of an NLS-peptide to linear dna does not enhance transfection efficiency of cationic polymer based gene delivery systems. J Control Release 2005;101(1-3):395-7

76. Nagasaki T, Myohoji T, Tachibana T, et al. Can nuclear localization signals enhance nuclear localization of plasmid DNA? Bioconjug Chem 2003;14(2):282-6

77. Collins E, Birchall JC, Williams JL, et al. Nuclear localisation and pDNA condensation in non-viral gene delivery. J Gene Med 2007;9(4):265-74

78. 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(1):91-6

79. Pollard H, Remy JS, Loussouarn G, et al. Polyethylenimine but not cationic lipids promotes transgene delivery to the nucleus in mammalian cells. J Biol Chem 1998;273(13):7507-11

80. Godbey WT, Wu KK, Mikos AG. Tracking the intracellular path of poly (ethylenimine)/DNA complexes for gene delivery. Proc Natl Acad Sci USA 1999;96(9):5177-81

81. Tanimoto M, Kamiya H, Minakawa N, et al. No enhancement of nuclear entry by direct conjugation of a nuclear localization signal peptide to linearized DNA. Bioconjug Chem 2003;14(6):1197-202

82. Malim MH, Hauber J, Le SY, et al. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature 1989;338(6212):254-7

83. Mor A, Amiche M, Nicolas P. Structure, synthesis, and activity of dermaseptin b, a novel vertebrate defensive peptide from frog skin: relationship with adenoregulin. Biochemistry 1994;33(21):6642-50

84. Pinol-Roma S, Dreyfuss G. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature 1992;355(6362):730-2

85. Pollard VW, Michael WM, Nakielny S, et al. A novel receptor-mediated nuclear protein import pathway. Cell 1996;86(6):985-94

86. Subramanian A, Ranganathan P, Diamond SL. Nuclear targeting peptide scaffolds for lipofection of nondividing mammalian cells. Nat Biotechnol 1999;17(9):873-7

87. Ma H, Zhu J, Maronski M, et al. Non-classical nuclear localization signal peptides for high efficiency lipofection of primary neurons and neuronal cell lines. Neuroscience 2002;112(1):1-5

88. 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 2003;42(35):10343-7

89. Paul RW, Weisser KE, Loomis A, et al. Gene transfer using a novel fusion protein, GAL4/invasin. Hum Gene Ther 1997;8(10):1253-62

90. 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(4):521-30

91. Khalil IA, Kogure K, Akita H, et al. Uptake pathways and subsequent intracellular trafficking in nonviral gene delivery. Pharmacol Rev 2006;58(1):32-45

92. 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(15):8835-41

93. Rajagopalan R, Xavier J, Rangaraj N, et al. Recombinant fusion proteins TAT-Mu, Mu and Mu-Mu mediate efficient non-viral gene delivery. J Gene Med 2007;9(4):275-86

94. Xavier J, Singh S, Dean DA, et al. Designed multi-domain protein as a carrier of nucleic acids into cells. J Control Release 2009;133(2):154-60

95. Varga CM, Wickham TJ, Lauffenburger DA. Receptor-mediated targeting of gene delivery vectors: insights from molecular mechanisms for improved vehicle design. Biotechnol Bioeng 2000;70(6):593-605

96. Wang Y, Mangipudi SS, Canine BF, et al. A designer biomimetic vector with a

chimeric architecture for targeted gene transfer. J Control Release 2009;137:46-53

- **This article describes the structure of the first genetically engineered vector with the ability to perform four distinct functions.**

97. Balicki D, Putnam CD, Scaria PV, et al. Structure and function correlation in histone H2A peptide-mediated gene transfer. Proc Natl Acad Sci USA 2002;99(11):7467-71

98. Mangipudi SS, Canine BF, Wang Y, et al. Development of a genetically engineered biomimetic vector for targeted gene transfer to breast cancer cells. Mol Pharm 2009;6(4):1100-9

99. Canine BF, Wang Y, Hatefi A. Evaluation of the effect of vector architecture on DNA condensation and gene transfer efficiency. J Control Release 2008;129:117-23

100. Canine BF, Wang Y, Hatefi A. Biosynthesis and characterization of a novel genetically engineered polymer for targeted gene transfer to cancer cells. J Control Release 2009;138(3):188-96

101. Hatefi A, Megeed Z, Ghandehari H. Recombinant polymer-protein fusion: a promising approach towards efficient and targeted gene delivery. J Gene Med 2006;8(4):468-76

- **Hatefi, introduced the structure of the first recombinant cationic biopolymer with tandem repeating units for targeted gene transfer.**

102. Munkonge FM, Dean DA, Hillery E, et al. Emerging significance of plasmid DNA nuclear import in gene therapy. Adv Drug Deliv Rev 2003;55(6):749-60

Affiliation

Helen O McCarthy[1], Yuhua Wang[2], Sriramchandra S Mangipudi[2] & Arash Hatefi[†2] PhD
[†]Author for correspondence
[1]Queens University Belfast, School of Pharmacy, BT9 7BL, Northern Ireland, UK
[2]Washington State University, Center for Integrated Biotechnology, Department of Pharmaceutical Sciences, PO Box 646534, Pullman, WA 99164, USA
Tel: +1 509 335 6253; Fax: +1 509 335 5902; E-mail: ahatefi@wsu.edu