Peptides as Promising Non-Viral Vectors for Gene Therapy

Wei Yang Seow and Andrew JT George
Imperial College London, England
United Kingdom

1. Introduction

In its most simplistic sense, gene therapy involves the delivery and expression of DNA by target cells so as to produce a therapeutic protein. In the case of RNA interference (RNAi), it is to shut off or silence the expression of a particular target protein. In order to exert its effects, the nucleic acid must first reach its intended site of action. DNA molecules (frequently as plasmids, which are circularised DNA) have to gain nuclear entry to access the transcription machinery. Conversely, RNAi molecules such as small interfering RNA (siRNA), short hairpin RNA (shRNA) and micro RNA (miRNA) will need to accumulate within the cytoplasm, although shRNA-encoding plasmids will require prior nuclear access before transcription into shRNA. However, if administered alone, a great majority of the nucleic acids will be degraded en route, leading to a loss of therapeutic potential. This then necessitates the development of vectors that protect and deliver nucleic acids to their target site. Arguably, it is the lack of safe and efficient delivery systems, rather than suitable therapeutic molecules that is limiting the success of gene therapy.

In this chapter, we start by examining how issues at the cellular level have shaped the design of modern, multifunctional vectors. We then briefly review the various types of gene delivery system, focusing on peptides as a promising class of non-viral vector. We will concentrate on the delivery of plasmids since the phenomenon of RNAi is relatively recent (Fire et al., 1998). As such, many strategies for RNAi delivery are adapted from DNA delivery technology.

2. Intracellular barriers in gene therapy: Problems and potential solutions

Ensuring the arrival of a plasmid at its site of action in a transcriptional state is the entire aim of gene delivery systems. However, plasmids face a constant threat of being degraded. The challenge begins as soon as they are introduced into the extracellular milieu (Figure 1). In most experimental setups, cells/tissues are maintained at 37°C in serum-supplemented medium where serum nucleases can extensively damage a naked plasmid. The plasmid therefore needs protection. Next, the plasmid needs to be internalised. However, both DNA (phosphate groups within the backbone) and plasma membrane (glycoproteins with their sialic acid groups, glycerophosphates with their phosphate groups and proteoglycans which contain sulphate groups) are negatively charged. Electrostatic repulsion then ensures that there is little chance of the plasmid being naturally taken up by a cell.
Fig. 1. Chronological sequence of events and challenges that a plasmid faces during its treacherous journey towards the nucleus. 1) A naked plasmid is susceptible to degradation by nucleases and is likely to be repelled from the plasma membrane. 2) Assuming successful endocytosis, the plasmid has to avoid trafficking into lysosomes where it will be degraded. 3) The plasmid has to escape into the cytoplasm. 4) It now has to diffuse through the viscous cytoplasm towards the nucleus while avoiding degradation and penetrate the nuclear membrane. 5) Transcription into mRNA can then occur if the plasmid is still intact.

For these reasons, gene delivery systems are frequently designed to be cationic in character and this fulfils several functions. First, the carrier can use its positive charges to mask the negative charges on the plasmid and package itself as a carrier/DNA complex with an overall positive charge. As expected, positively charged particles are internalised much more readily, as confirmed by an elegant study using PRINT (Particle Replication In Non-wetting Templates) technology to fabricate particles with exquisite control over their size, shape and surface charge (Gratton et al., 2008b). By keeping size and shape constant, positively charged particles were found in 84% of cells after an hour of incubation compared to the <5% uptake of negatively charged particles, proving that surface charge alone can influence uptake dramatically. Second, due to a charge screening effect, the macromolecular plasmid is collapsed (or condensed) into a compact structure more amendable for cellular uptake. This condensing process was clearly demonstrated using transmission electron microscopy which showed an elongated plasmid (long-axis diameter of ~470 nm) being compacted into tight, 80-100 nm toroid-shaped complexes by polylysine carriers (Wagner et al., 1991). Third, the carrier protects its cargo against degradation by nucleases, presumably
by steric obstruction. This was shown by first exposing the complexes to DNase and then using gel electrophoresis to validate the physical integrity of the plasmid upon its release from the carrier. Control plasmids that were unprotected gave no bands in the subsequent gel electrophoresis experiment.

A problem with cationic carriers is that negatively charged serum protein can be non-specifically bound. As a result, targeting signals on the carrier can become blocked or complexes can start to aggregate. Reducing or removing serum from the media during in vitro transfection can mitigate such effects and improve transfection (Moore et al., 2009; Moulton et al., 2004). However, this strategy fails during in vivo experiments where serum proteins are unavoidably present. Another approach is to mix DNA and carrier in precise stoichiometric ratio so as to result in electro-neutral complexes (Funhoff et al., 2005). Along similar lines, zwitterionic (McManus et al., 2004) or anionic (Liang et al., 2005) lipids have been proposed in which binding between carrier and plasmid is dependant on hydrophobic forces and the presence of divalent cations such as Ca$^{2+}$, Zn$^{2+}$ and Mg$^{2+}$ to screen the disruptive repulsion between like charges. Neutral water-soluble polymers such as polyvinyl alcohol (PVA) and polyvinyl pyrrolidone (PVP) similarly exploit non-electrostatic forces such as hydrogen bonding and van der Waal’s attraction to condense plasmids (Park et al., 2006). These particles are however less popular because the lack of positive charge is expected to adversely impact cellular uptake. The PEGylation of carriers is another option. PEGylation refers to the surface decoration of carriers with flexible chains of polyethylene glycol (PEG), which is a highly hydrophilic polymer capable of making a surface less susceptible to protein adsorption (Duncan, 2006). Having said that, a major concern is that PEGylation obstructs the positive charges, resulting in carriers which are less efficient in DNA binding and transfection (Lin et al., 2008; Meyer et al., 2008). Moreover, while PEGylated drug formulations are already in clinical use, PEG is non-biodegradable and its chronic use may be a concern (Urakami & Guan, 2008).

The positive charges on cationic carriers have also been implicated in the formation of pores on the membrane, leading to cytotoxicity (Rimann et al., 2008). The observation that carrier/DNA complexes are less toxic to the cells compared to the carrier alone can be interpreted as partial evidence that charge compensation on the carrier may be a reason, albeit a non-exclusive one, for the reduced toxicity (Niidome et al., 1997). However, the exact mechanism is not totally clear. It may be due to the high rate of uptake of cationic carriers – and not the positive charges per se – that is to blame for pore formation. Herein lies a dilemma of the gene therapist: a high rate of internalisation can increase transfection efficiency but is also frequently paralleled by toxicity (Gabrielson & Pack, 2006; Pouton & Seymour, 2001). Fortunately, cells do have membrane repair mechanisms. One example is the membrane repair response (MRR) where the influx of Ca$^{2+}$ ions directs lysosomes to donate their vesicular membrane in a concerted effort to plug the hole (Palm-Apergi et al., 2009). As always, the challenge is to strike a fine balance between two counter-acting events. Assume the plasmid has been successfully taken up by a cell via endocytosis and now resides within an endosome. Another degradative fate awaits as endosomes eventually acidify into lysosomes and activate a broth of acid hydrolases capable of degrading nucleic acids. For this reason, a high rate of uptake may not necessarily translate into high transfection efficiency if, for example, most of the plasmids are degraded in lysosomes (Lundin et al., 2008). To avoid degradation, the plasmid/carrier complex will need to escape from the confines of the endosome into the cytosol. A popularly cited mechanism by which cationic carriers can achieve this is the proton-sponge hypothesis, so called because it relies
on the buffering capability of the carrier to absorb H\(^+\) ions and thus function as a proton sink. This model assumes that as the H\(^+\)-ATPase endosomal membrane pump injects protons into the vesicle during acidification, a build-up of positive charges will result due to the ability of the carrier to protonate and absorb the protons. This then triggers a concomitant influx of compensatory negative ions (e.g., Cl\(^-\)) and water, leading to the osmotic swelling of the vesicle and its eventual rupture. Consistent with this line of argument, carriers must thus contain chemical groups that are capable of undergoing protonation within the pH range of the endo/lysosomal transition, i.e., a pK\(_a\) of 7 to 4. This has motivated investigators to design carriers with a large buffering capacity. The polymer polyethylenimine (PEI), for instance, has a combination of protonated amines to bind plasmids at pH 7 and a stockpile of unprotonated amines that can still undergo protonation during the endo/lysosomal transition (Boussif et al., 1995). As such, PEI has a large buffering capacity and this feature is frequently cited as a main reason for PEI’s status as one of the most efficient non-viral vector commercially available for in vitro transfection (Putnam, 2006).

Another common strategy to promote endosomal escape is to coincubate cells with a lysosomotrophic agent such as chloroquine. Traditionally used as an anti-malaria drug, chloroquine is also a weak base capable of buffering the acidification of endosomes. In reality, however, chloroquine is pleiotropic in nature – besides its lysosomotropic property, chloroquine has been reported to be able to aid in the release of plasmid from its carrier and also to inhibit DNase activity (Yang et al., 2009) – and its actual mechanism of action remains controversial. Nevertheless, chloroquine does generally improve the transfecting capability of many carriers (Pouton & Seymour, 2001). A caveat, though, is that chloroquine at the dose normally used (100 μM) is toxic to cells (Wadia et al., 2004; Zauner et al., 1998). Glycerol is another agent reported to augment transfection due to its ability to weaken and make vesicular membrane more susceptible to disruption (Zauner et al., 1997). Interestingly, the more obvious effect of adding glycerol – its osmotic property – was ruled out as the main cause of vesicle escape. Finally, carriers can be functionalised with membrane-disruptive peptides, a strategy which will be reviewed in later sections.

The proton-sponge hypothesis is the most commonly cited explanation to account for the positive correlation between increased buffering capacity and transfection ability. It is hence easy to overlook that no study has provided any convincing evidence in direct support of its mechanism (Won et al., 2009). In fact, discrediting observations exist. For instance, it was reported that ammonium sulphate, also a weak base and should theoretically be able to provide buffering effects, does not boost transfection (Pouton & Seymour, 2001). Recent calculations have also revealed that the amount of strain that lipid vesicles can withstand before rupture is significantly larger than that which can be induced by endosomal buffering (Won et al., 2009). This suggests that the proton-sponge hypothesis can at best contribute, but cannot be the only cause of endosomal escape. Moreover, increased buffering and endosomal escaping properties do not always produce an accompanying increase in transfection (Akita et al., 2010; Moore et al., 2009). This implies that, while important, cytosolic access is not the only bottleneck of the transfection process. Further, it has to be pointed out that of the multiple pathways which a cell can use for internalisation, only the clathrin-mediated one is widely-accepted to involve vesicle acidification (Won et al., 2009; Zauner et al., 1997). Vesicle acidification is, of course, an inherent requirement of the proton-sponge hypothesis but whether vesicles from other pathways acidify is equivocal. Some researchers believe that macropinocytosis (Akita et al., 2010; Pelkmans & Helenius, 2002)
and caveolae-mediated endocytosis (Lundin et al., 2008; Sahay et al., 2010) produce vesicles that do not undergo acidification, while others claim that macropinosomes do acidify (Räägel et al., 2009; Wadia et al., 2004). Furthermore, is it safe to assume that vesicles which do not acidify remain distinct from endo/lysosomes? On this, opinion is also polarised, with some claiming that caveosomes (Pelkmans & Helenius, 2002; Sahay et al., 2010) and macropinosomes (Wadia et al., 2004) remain distinct from endo/lysosomes; and others arguing that vesicles from different pathways can eventually interact (Sahay et al., 2010). Thus, the proton-sponge hypothesis even if true, may not always be relevant and on top of that, definitive trafficking studies of the various modes of uptake are required.

In the cytosol, the plasmid continues its migration towards the nucleus. Current dogma suggests that this proceeds via passive diffusion and that nuclear localisation is a hit-or-miss event. The viscous cytosolic environment makes diffusion extremely inefficient. The diffusion coefficient of bovine serum albumin (BSA) in human fibroblasts, for example, is about 70× lower than in buffer (Wojcieszyn et al., 1981). Protecting the plasmid during migration is also important as cytosolic nucleases restrict the half life of naked DNA to about 90 minutes (Belting et al., 2005).

During migration, another feature of the carrier becomes important - the release of its plasmid cargo. To undergo transcription, the plasmid must first be unpackaged from its carrier and the trick here, is one of timing. A plasmid that gets released too early will risk degradation while one that binds too strongly is not accessible for transcription. For this reason, lower molecular weight chitosan transfects better because their higher molecular weight counterpart interacts too strongly with their plasmid cargo (Koping-Hoggard et al., 2004). Another example involves PEI, where acetylation of its polymeric chains (which removes the amines) reduced both its buffering capacity and binding strength, but improved its transfection (Gabrielson & Pack, 2006). This suggests that vector unpackaging can indeed be a rate-limiting step and a carrier that binds in moderation is ideal.

Fluorescence resonance energy transfer (FRET) is a powerful technique to study the process of vector unpackaging. It depends on the excitation of an acceptor dye by a donor dye in close physical proximity, which is the case when the plasmid is being tightly condensed. Using FRET, plasmid-release in the perinuclear region has been observed; alternatively, the plasmid/carryer can enter the nucleus as an associated complex (Seow et al., 2009).

Nuclear entry is believed to be achieved in two ways: either via the ~10 nm wide nuclear pores or during mitosis when the nuclear envelope momentarily disintegrates (Luo & Saltzman, 2000a). The latter appears to be a more efficient method for the large-scale accumulation of complexes. It further provides a possible explanation for why amitotic cells or primary cells which proliferate slower are more difficult to transfect than cancer-derived cell lines. A dilution effect after mitotic cell division also accounts for the transient nature of gene expression mediated by non-integrating vectors. Having said that, mitosis is not a pre-requisite (Won et al., 2009) as amitotic cells have been successfully transfected – it merely provides a convenient window of opportunity for nuclear entry. A proposed method to improve nuclear penetration is to attach a nuclear localisation signal (NLS) to the carrier. The quintessential example of a NLS is the short peptide sequence corresponding to the Simian virus 40 (SV40) T antigen. However, responses regarding the benefits of including a NLS have been mixed, with some investigators (Trentin et al., 2005) more convinced than others (Zauner et al., 1998). A key issue pointed out was that studies involving the use of NLS failed to examine the effect of including a NLS on nuclear import per se (Lam & Dean, 2010). Instead, reporter gene expression was frequently used as a proxy and an improved
expression was simply accepted to be due to the inclusion of NLS. Given that most NLS are cationic, it is debatable if the observed increase in transfection is strictly the result of improved nuclear import \textit{per se}, or due to other non-specific effects such as enhanced plasmid association and uptake. Moreover, the fact that a NLS can be sterically hindered by plasmids upon binding also contradicts the requirement of NLS to be freely accessible for interaction with importins, the nuclear entry regulating proteins.

In light of the many obstacles that nucleic acids and their carriers have to surmount, it is perhaps understandable that less than 10% of the pool of plasmids that made it into a cell will go on to accumulate in the nucleus (Lam & Dean, 2010). The challenge is to design a multifunctional vector that can address the issues highlighted above and yet, remain safe to use in a human body.

3. Overview and classification of gene delivery systems

3.1 Physical systems

Gene delivery systems can be classified based on their means of achieving transfection (Figure 2). There are systems that utilise physical forces, mostly with the aim of disrupting the plasma membrane, to facilitate nucleic acids delivery. For instance, electroporation and sonoporation (Frenkel, 2008) uses electrical and sonication forces respectively to transiently compromise the plasma membrane. The ballistic gene gun method, on the other hand,
directly shoots DNA-coated metal particles (frequently gold) into cells (Merediz et al., 2000). Expectedly, such aggressive methods can irreversibly damage the cell membrane and cause widespread cell death. Scrape loading, first described in 1984, requires the forceful scraping of cells attached to their culture dishes, thereby creating pores on their membranes for plasmid entry (McNeil et al., 1984). This technique, however, is only applicable to adherent cells. The osmotic lysis of pinosomes was proposed in 1982 and requires that cells capable of pinocytosis be initially exposed to a hypertonic medium of sucrose, PEG and the plasmid of interest (Okada & Rechsteiner, 1982). Subsequent exchange to a hypotonic medium then released the pinosomal content. Although cells are constitutively capable of pinocytosis, such wild fluctuations in osmotic conditions can cause cell death. Finally, microinjection refers to the piecemeal injection of plasmids directly into the cell. This technique, while relatively gentle, is extremely laborious. As such, it is prone to failure and the number of cells that can be processed is limited.

3.2 Viral vectors
There are systems that function as particulate carriers by ferrying nucleic acids into or near to their site of action. Such systems can be viral or non-viral in nature. Viral vectors such as adenovirus, adeno-associated virus (AAV), human immunodeficiency virus (HIV), equine infectious anaemia virus (EIAV), herpes simplex virus (HSV) and murine leukaemia virus (MLV) have been used and comprehensively reviewed elsewhere (Kay et al., 2001). The prime advantage of viral vectors is their transfection efficiency which has benefitted from centuries of selective evolutionary pressure. As a result, viruses are several orders of magnitude more efficient than non-viral vectors (Kircheis et al., 1997) and account for about 70% of all clinical trials involving gene therapy so far (Won et al., 2009). A recent success for viral gene therapy involved the use of lentiviruses to deliver a correct copy of a therapeutic gene to an adult patient suffering from β-thalassaemia (Cavazzana-Calvo et al., 2010). Such patients suffer from defective haemoglobin production and require chronic blood transfusion for survival. Upon reinforcing the patient with his own bone-marrow haematopoietic stem cells that had earlier been transduced \textit{ex vivo}, the patient became transfusion free 1 year after treatment and has been doing well, according to the most recent report at 33 months after treatment. The longer-term outcome, of course, remains to be seen. Having said that, investigators have not abandoned all forms of non-viral research as there are limitations in the use of viral vectors. For example, the technical difficulty of scaling up virus production compliant to good manufacturing practices (GMP) may prevent such therapy from being cheaply accessible (Sheridan, 2011). The size of the construct that can be delivered is also limited. Above all, toxicity and immunogenicity (reviewed by Nayak & Herzog, 2010) provoked by the viral vectors can and have brought clinical trials to premature ends. Repeated administration is also not possible if the body has mounted a systemic immune response. The danger of viral gene therapy was first illustrated about a decade ago by the death of Jesse Gelsinger (Hollon, 2000). Researchers were using experimental adenoviruses to correct his partial ornithine transcarbamylase deficiency affecting the ability of his body to metabolise nitrogen. Unfortunately, the systemic inflammatory response syndrome was triggered and he succumbed, rather rapidly, to multiple organ failure. In another high-profile example, stem cells transduced \textit{ex vivo} with MLV were used to treat children suffering from X-linked severe combined immunodeficiency (Hacein-Bey-Abina et al., 2003). However, MLV is an integrating retrovirus and can cause insertional mutagenesis, which is a phenomenon where random
viral integration can trigger the activation of an oncogene or disrupt genes with tumor-suppressive properties. Consequently, several patients became leukemic and at least one has died so far (Sheridan, 2011). More recently, retroviruses were also used to treat two patients diagnosed with X-linked chronic granulomatous disease which affected the ability of their phagocytes to clear bacterial infections (Ott et al., 2006). Insertional mutagenesis was again responsible for a clone of cells whose genes responsible for growth were activated, prompting fears that such cells may turn cancerous. Having said that, viral gene therapy is not expected to be perfect and its success should be assessed relative to existing treatments. Nonetheless, researchers have not completely mastered the use of viruses and there is still a need for alternative non-viral vectors with safer profiles.

3.3 Non-viral vectors
Diethylaminoethyl (DEAE)-dextran and calcium phosphate (CaPO$_4$) were two of the earliest systems popular in the 1970s and 80s. DEAE-dextran has a cationic polysaccharide backbone and one of the earliest reports in 1965 used this polymer in a 1:1 mixture with nucleic acids to transfect rhesus monkey kidney cells (Vaheri & Pagano, 1965). Authors of this study also insightfully commented that DEAE-dextran, like histones, could bind and protect nucleic acids from degradation. Although these are now fundamental concepts in modern carrier design, it may not have been as obvious in the past. The co-precipitation of CaPO$_4$ with DNA was first described in 1973, when it was observed that Ca$^{2+}$ (not Mg$^{2+}$ or Na$^+$) and PO$_4^{2-}$ ions, at high enough concentrations, could enhance DNA uptake by cells (Graham & Eb, 1973). This technique was further shown to be sensitive to pH, amount of DNA used and even the level of CO$_2$ in the incubator (Chen & Okayama, 1987). Using an optimised protocol, CaPO$_4$-mediated transfection achieved up to 50% efficiency with a murine L cell line (Chen & Okayama, 1987).

Chitosan is a polysaccharide obtained by the deacetylation of chitin, an exoskeletal component of crustaceans. Each deacetylation site contains a primary amine (pK$_a$~6.5) which allows chitosan to bind nucleic acids. The degree of deacetylation also determines its biodegradability and transfection efficiency. Chitosan was first described as a plasmid carrier by Mumper and colleagues in 1995 and is known to be biocompatible, mucoadhesive and virtually non-toxic (MacLaughlin et al., 1998). As such, chitosan has been evaluated in rabbits, although reporter gene expression in that study was low (MacLaughlin et al., 1998). Strong interactions between high molecular chitosan and DNA were blamed for the over-stable complexes that could not release their plasmid cargos. Using lower molecular weight chitosan, in vitro transfection efficiency was improved by up to 24 fold (Koping-Hoggard et al., 2004). Adding histidine (pK$_a$~6) residues to chitosan further improved buffering capacity and overall transfection efficiency (Chang et al., 2010). Recently, it was also shown that the introduction of thiol groups in N,N,N-trimethylated chitosan improved its efficiency as a siRNA carrier (Varkouhi et al., 2010).

Red blood cells (RBC) were used for the delivery of macromolecules before liposomes became popular. To function as carriers, RBC were first loaded up with the macromolecule under hypotonic conditions to induce mild lysis. A fusogen was then used to cause fusion with target cells. Using this technique, thymidine kinase and BSA were introduced into 3T3-4E mouse cells using Sendai virus as a fusogen (Schlegel & Rechsteiner, 1975). In another study, horseradish peroxidase (HRP) and immunoglobulins (IgG) were delivered into cells expressing hemagglutinin derived from influenza virus by exploiting the fusogenic activity of hemagglutinin under mildly acidic conditions (Doxsey et al., 1985). Fluorescent IgG and
BSA were also delivered into human fibroblast cells using PEG as a fusogen (Wojcieszyn et al., 1981). Dehydration may play a role in PEG-mediated fusion and intriguingly, purification by recrystallisation removed the fusogenic ability of PEG, suggesting that other ingredients in commercial PEG was essential (Wojcieszyn et al., 1983). Plasmids were initially delivered by synthetic liposomes using concepts adapted from RBC technology. In an early study, plasmids were encapsulated within anionic liposomes and fused with tobacco mesophyll protoplasts using PEG (Deshayes et al., 1985). Since 1987, however, surface binding became popular when liposomes made of \(N\)-\{1-(2,3-dioleyloxy)propyl\}\(\text{N},\text{N},\text{N}\)-trimethylammonium chloride (DOTMA) were synthesised for the first time (Felgner et al., 1987). DOTMA contains cationic quaternary ammonium groups which can be used for DNA binding. Lipofectin®, the prototypical lipid formulation and a workhorse for transfection experiments today, is a 1:1 mixture of DOTMA and the neutral lipid, dioleoyl phosphatidylethanolamine (DOPE). In this formulation, DOPE functions as a fusogen to aid cellular uptake. Several other formulations of liposomes have been reported. Gao and Huang prepared liposomes made out of DOPE and a novel cationic cholesterol derivative, \(3\beta\)-\{\(N\)'\(N\)'-dimethylamino\)ethane-carbamoyl\} cholesterol (DC-chol), and used it to transfect several cell lines with high efficiency (Gao & Huang, 1996). A mixture of \(N\)'\(N\)'-dioctadecyl-N-4,8-diaza-10-aminodecanoylglycine amide (DODAG) and DOPE was also proposed to be a novel liposomal formulation that is good for both DNA and siRNA delivery (Mével et al., 2010). In another study, the cationic lipid 1,2-dilinoleyloxy-3-dimethylaminopropane (DLinDMA) was used as a starting template and systematically modified to yield an optimised construct (Semple et al., 2010). Although toxicity remains one of the chief complaints, liposomes are definitely one of the most successful non-viral vectors around. Already, several sustained-release liposomal drug formulations have been approved for human use and thermosensitive liposomes are being evaluated in clinical trials (Hossann et al., 2010). Lipids also dominate in gene therapy trials aimed at using non-viral vectors to treat cystic fibrosis (Griesenbach & Alton, 2009).

An advantage of polymeric carriers is that they offer so much structural diversity that their potential should, theoretically, be limited only by the imagination of chemists. Two of the more archaic examples including PVP and PVA have been mentioned. Poly(D,L-lactic-co-glycolic acid) (PLGA), a polymer originally developed for controlled drug release, has also been used to encapsulate plasmids for delivery (Wang et al., 1999). Most polymeric carriers, however, are cationic in nature. Dendritic poly(amido amine) (PAMAM) which contained primary amines for DNA binding was further decorated with PEG and sugars (Wood et al., 2005) or peptide (Wood et al., 2008) for targeting purposes. Linear PAMAM was functionalised with disulfide bonds to make the polymer biodegradable and various chemical side-groups were attached to test for its impact on transfection efficiency (Lin et al., 2007). A large library of poly(\(\beta\)-amino ester) has been synthesised and systematically screened for transfection efficiency (Green et al., 2006). The effects of molecular weight and charge density of polyphosphoramidates (PPA) on DNA binding and buffering capacity have been reported (Ren et al., 2010). A series of novel enzyme-degradable polycarbonates (PC) with various aliphatic side-chains attached were prepared and shown to effectively transfect cells with virtually no toxicity (Seow & Yang, 2009a). Poly(2-(dimethyl-amino)ethylmethacrylate) (PDMAEMA) was also shown to transfect cells with high efficiency (Lin et al., 2008). Indeed, there are so many other classes of polymers being developed that it is impossible to do all justice with this paragraph. Nonetheless, amongst all, the most notable polymer must be PEI. PEI was first developed in 1995 (Boussif et al., 1995).
1995) and is frequently used in its branched, high molecular weight (usually 25 kDa) form. Together with liposomes, they are widely acknowledged to be the best non-viral vectors currently available (Putnam, 2006) and frequently serve as standards to which other novel carriers are referenced. A key feature of PEI is that nitrogen (in a mixture of primary, secondary and tertiary amines) accounts for a third of its molecular weight. Since different classes of amines possess characteristic pKₐ, this ensures that not all the nitrogen will be protonated at a given pH. Furthermore, the proximity of the nitrogen atoms also means that a protonated amine can suppress the protonation of its neighbours due to the energetic penalty that gets incurred by situating like charges adjacent to one another (Suh et al., 1994). This blend of protonated and yet-to-be protonated amines is then suggested to endow PEI with its strong DNA binding and buffering abilities. The intracellular trafficking properties of PEI/DNA complexes has been studied (Godbey et al., 1999) and targeting moieties such as mannose (Diebold et al., 1999) and transferrin (Kircheis et al., 1997) were also coupled onto PEI. However, two of the main problems of PEI are its toxicity and non-biodegradability. In response to the latter, PEI was functionalised with reducible disulfide bonds (Lee et al., 2007) or hydrolysable ester bonds (Liu et al., 2008) to facilitate biodegradation and at the same time, to aid its intracellular plasmid release.

4. Peptide-derived vectors for gene therapy

Peptide chains can be fabricated from any of the 20 naturally occurring L-amino acids, which are referred to by their single- or three-lettered code (e.g., R or arg for arginine). Peptides are thus biocompatible and often degradable. Peptide synthesis also does not involve harmful catalysts, which is a concern in the synthesis of some polymeric carriers. Furthermore, synthesis can now be automated, courtesy of advances in solid phase peptide synthesis, which makes the manufacturing process amendable to up-scaling.

Before peptides were seen as proper DNA carriers, studies conducted in the 1960s with histones had already suggested that cationic amino acids such as lysine and arginine can be useful (Akinrimisi et al., 1965). Today, peptide vectors are given fanciful names such as “cell penetrating peptides” (CPP) or “protein transduction domains” (PTD) to celebrate their ability to efficiently penetrate the plasma membrane and mediate the entry of nucleic acids or other macromolecules. Such CPP can be derived from proteins existing in nature (e.g., viral proteins or venom proteins of bees and wasps) or designed de novo. There is little in common among CPP and the only unifying theme seems to be the significant presence of cationic residues (mainly lysine and arginine). Amphipathicity has been suggested to be another common feature. However, this is only true for most CPP – oligoarginine being an exception. There are two methods by which peptide vectors carry their cargoes. Nucleic acids are usually non-covalently (electrostatically) attached, while proteins and other macromolecules are typically covalently coupled (chemical cross-linking or by plasmid fusion). The advantages of electrostatic attachment include convenience and the largely unaltered chemical properties of the cargo. However, charge interactions are non-specific and excess peptides are usually needed to completely bind the plasmid. On the other hand, covalent attachment requires chemical modification of the cargo and usually results in stable complexes which, as discussed earlier, may not be desirable. In the following sections, we will review the key classes of peptide vectors and provide non-exhaustive examples of strategies that have been used to improve their efficiency as vectors.
Table 1. The amino acid sequences of all the peptides discussed in this chapter.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence (single letter code)</th>
<th>Short description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly-L-lysine</td>
<td>K&lt;sub&gt;n&lt;/sub&gt;</td>
<td>Frequently, n = 50-400</td>
</tr>
<tr>
<td>Tat</td>
<td>YGRK KRRQ RRPP PQ</td>
<td>HIV-derived, sequence 47-60</td>
</tr>
<tr>
<td>Oligoarginine</td>
<td>R&lt;sub&gt;n&lt;/sub&gt;</td>
<td>Frequently, n = 7-9</td>
</tr>
<tr>
<td>Penetratin</td>
<td>RQIK IWFQ NRRM KWKK</td>
<td>3&lt;sup&gt;rd&lt;/sup&gt; helix of Antennapedia homeodomain</td>
</tr>
<tr>
<td>Tp</td>
<td>GWTL NSAG YLLG KINL KALA ALAK KIL</td>
<td>Galanin + mastoparan (wasp venom)</td>
</tr>
<tr>
<td>Tp10</td>
<td>AGYL LGKI NLKA LAAL AKKI L</td>
<td>First 6 residues deleted from Tp</td>
</tr>
<tr>
<td>VP22</td>
<td>300-residue long sequence given in (Elliott &amp; O’Hare, 1997)</td>
<td>HSV-1 structural protein</td>
</tr>
<tr>
<td>MPG</td>
<td>GALF LGFL GAAG STMG AWSQ PKSK RKV</td>
<td>HIV gp41 + SV40 T-antigen</td>
</tr>
<tr>
<td>CADY</td>
<td>GLWR ALWR LLRS LWRL LWRA</td>
<td>Peptide carrier PPTG1-derived</td>
</tr>
<tr>
<td>KALA</td>
<td>WEAK LAKA LAKA LAKH LAKA LAKA LKAC EA</td>
<td>Membrane disruptive peptide, also a carrier</td>
</tr>
<tr>
<td>GALA</td>
<td>WEAA LAEA LAEA LAEH LAEA LAEA LEAL AA</td>
<td>Membrane disruptive peptide</td>
</tr>
<tr>
<td>EB1</td>
<td>LIRL WSHL IHIW FQNR RLKW KKK</td>
<td>Penetratin analogue</td>
</tr>
<tr>
<td>HA2</td>
<td>GLFG AIAE FIEN GWEG MIDG</td>
<td>Influenza virus hemagglutinin protein</td>
</tr>
<tr>
<td>INF5</td>
<td>(GLFE AIEG FIEN GWEG nIDG)nK</td>
<td>HA2-derived, lysine-connected dimer. n = norleucine</td>
</tr>
<tr>
<td>INF7</td>
<td>GLFE AIEG FIEN GWEG MIDG WYG</td>
<td>HA2-derived, monomer</td>
</tr>
<tr>
<td>SV40</td>
<td>PKKK RKV</td>
<td>Simian virus 40 T antigen NLS</td>
</tr>
<tr>
<td>Melittin</td>
<td>GIGA VLKV LTTG LPAL ISWI KRKR QQ</td>
<td>Bee venom</td>
</tr>
</tbody>
</table>

4.1 Poly-L-lysine (PLL)

The amino acid sequence of PLL and all other peptides subsequently discussed can be found in Table 1. PLL is the first peptide-based vector to be studied intensively. The molecular weight of PLL spans a wide range and depends on the number of repeating units within a chain, or its degree of polymerisation (DP). For the sake of discussion, the molecular weight of PLL is arbitrarily classified as follows: oligolysine (DP<20), low (20≤DP<50), medium (50≤DP≤400) or high (DP>400). Most studies use PLL of medium molecular weight.

PLL has been successfully employed to deliver a host of different cargos. By covalently coupling human serum albumin or HRP to PLL via amide bonds in a 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) catalysed reaction, proteins were delivered into L929 mouse fibroblasts for the first time using such a PLL-mediated strategy (Shen & Ryser, 1978). More recently, streptavidin-conjugated quantum dots (QD) were also attached to biotin-tagged PLL and delivered into HeLa cells (Mok et al., 2008). However, PLL is most frequently used to deliver nucleic acids where the electrostatic attachment is preferred. Although polylysine made out of D-amino acids has been suggested to work better – presumably because they are more resistant to enzymatic degradation (Mitchell et al., 2000) – PLL continues to be favoured in transfection studies. Wu and partner provided one of the earliest examples of PLL being used to condense plasmids for receptor-mediated delivery (Wu & Wu, 1987). In that study, asialoorosomucoid (ASOR) was first covalently attached to PLL with N-succinimidyl 3-(2-pyridyldithio) propionate (SPDP) as a linker and then used to
bind a reporter plasmid. Complexes then accumulated within HepG2 cells (ASOR receptor positive) but not in SK-Hep1 cells (ASOR receptor negative). Plasmids dissociated more slowly if bound to a higher molecular weight PLL and this can negatively affect transfection (Schaffer et al., 2000). To facilitate plasmid release, cysteine residues were added to oligolysine, which was then cross-linked using disulfide bonds (McKenzie et al., 2000). The intention was to make use of the intracellular environment – which is much more reductive due to its elevated levels of glutathione (Lee et al., 2007) – to break down the disulfide bonds for polymer disintegration and plasmid release. The number of lysine, cysteine and histidine (for buffering capacity) residues used and the spacing and ordering between the residues were found to influence the final performance of the vector. Upon optimisation, some constructs transfected at levels that were comparable or even higher than LipofectAce, a commercial reagent used as a positive control (McKenzie et al., 2000). Disulfide-linked oligolysine was then further functionalised with triantennary N-glycan signals to target hepatocytes and evaluated in vivo (Kwok et al., 2003). However, contrary to results in vitro, the particles were not stable enough in the reductive intracellular liver environment and premature plasmid release ultimately limited gene expression.

PLL by itself is generally not considered to be an efficient vector (Meyer et al., 2008) and is frequently coupled with other agents. The imidazole headgroup of histidine (pK~6) can provide endosomal buffering and has been added to PLL to boost transfection (Midoux & Monsigny, 1999). Histidine was also added onto a K15-based oligopeptide which was then self assembled for drug and gene co-delivery (Wiradharma et al., 2009). The HA2 subunit of the hemagglutinin glycoprotein present on the surface of influenza virus plays an important role in the endosomal escape of viruses. To do this, HA2 exhibits a pH-dependent membrane fusion activity. HA2 is not normally lytic at neutral pH. However, protonation of its acidic residues during endo/lysosomal transition triggers HA2 to adopt a more ß-helical secondary conformation. At the same time, it exposes a highly-conserved hydrophobic sequence which then interacts and destabilises the endosomal membrane (Wagner et al., 1992a). To exploit this membrane-disrupting mechanism, 20 amino acids corresponding to the N-terminus of HA2 were attached to PLL via cysteine-mediated disulfide bonds. The HA2-functionalised PLL was then shown to induce pH-dependent liposomal leakage (Wagner et al., 1992a) and to augment transfection to a greater extent than the use of chloroquine (Midoux et al., 1993). Melittin, a major component of bee venom, is another peptide with fusogenic activity but unlike HA2, melittin is unresponsive to pH. As such, it remains fusogenic at pH 7 and can indiscriminately disrupt plasma membranes (Chen et al., 2006). To confine fusogenic activity within the endosomes, dimethylmaleic anhydride protecting groups were attached to mask the activity of melittin at neutral pH. Upon cleavage of the protecting groups under acidic conditions, the activity of melittin was restored (Meyer et al., 2008). This protected form of melittin was then coupled to PLL via a cysteine-mediated disulfide bond and the entire construct mediated siRNA knockdown more efficiently than PEI. Adenovirus, known to display pH-dependent membrane disruption as part of its infectious cycle (Curiel et al., 1991), has also been coupled to PLL to improve transfection efficiency. Human adenovirus type 5 (dl312) was either biotinylated and coupled to a streptavidinylated PLL (Wagner et al., 1992b), or simply added as free viral particles into the culture medium to be taken up together with the PLL complexes (Curiel et al., 1991). In another study, chicken embryo lethal orphan virus, an adenovirus from chicken, was attached to PLL and augmented transfection as well as human adenovirus (Cotten et al., 1993).

Various signals have also been added onto PLL to improve its uptake or target specificity. For instance, galactose containing an isothiocyanate group was reacted with the amine
groups on PLL to form a thiourea bond and used to mediate gene expression in the livers of rats for up to 140 days (Perales et al., 1994). Mannose and lactose were also used to target cell lines that have receptors for the specific sugar (Midoux et al., 1993). ASOR was coupled onto PLL in an EDC catalysed reaction and intravenously injected into rats (Wu et al., 1989). However, reporter gene expression by liver cells disappeared by day four. Transferrin was coupled to PLL via disulfide bonds in a SPDP-mediated reaction to increase its accumulation within cells (Wagner et al., 1991). Insulin was also coupled to PLL in an SPDP-mediated reaction and used with adenovirus (biotin-streptavidin conjugated) to deliver plasmids into pre-implantation mammalian embryos (Ivanova et al., 1999). PEG was grafted onto PLL via amide bonds in an N-hydroxysuccinimide (NHS) catalysed reaction and helped in preventing particle aggregation (Rimann et al., 2008). Stearyl-PLL mixed with a low density lipoprotein was also used to condense plasmids (~600 nm) into ~100 nm complexes, as measured by atomic force microscopy (Kim et al., 1998). Further, the lipoprotein was found to be essential for efficient uptake. Finally, peptides themselves can serve as targeting signals. For example, a short peptide sequence (GACRRETAWACG) suggested to target α5β1 integrins was linked to K16 and mixed with Lipofectin® to transfect neuroblastoma cells (Lee et al., 2003). Several groups have also used the RGD motif to target integrins. A recent example involved RGD being linked to K16 peptides, which were then used with an intracellularly cleavable PEG-lipid formulation to deliver plasmids into mice bearing subcutaneous tumors (Tagalakis et al., 2011). The integrin-targeting signal was shown to be important. Upon augmentation by the enhanced permeation and retention effect, the complexes were mainly distributed to the tumor.

4.2 Tat

Full length tat has 86 amino acids (sequence 1-86) and it is a regulatory protein encoded by HIV-1 that transactivates viral gene expression. The ability of tat to cross cell membrane was initially observed by two groups independently in 1988 (Frankel & Pabo, 1988; Green & Loewenstein, 1988). In one study, tat was simply added to a culture of HL3T1 cells modified to contain an integrated copy of chloramphenicol acetyltransferase (CAT, a reporter gene) under the control of the HIV-1 LTR (long terminal repeat) promoter and CAT expression was unexpectedly detected (Frankel & Pabo, 1988). Crucially, the amount of CAT expression was dependent on the dosage of tat. Unlike earlier experiments, tat did not require any help from the scrape loading technique to enter cells. Tat 1-86 was further dissected into five regions and region II and III, together spanning roughly residues 38-62, were identified to be essential and sufficient for transactivation activity (Green & Loewenstein, 1988). Region III (roughly residues 49-62), in particular, was interesting and was rich in arginine and lysine. Furthermore, replacing the three arginines at position 55-57 with alanines drastically reduced transactivation activity. Tat was also observed to localise in the nucleus and thought to be a NLS (Green & Loewenstein, 1988).

The region surrounding the arginine- and lysine-rich domain of tat was keenly evaluated as a gene carrier. Today, full length tat is seldom used and among the many truncated versions being studied (Table 2), tat 47-57 and 48-60 are the two most popular sequences. The in vivo half life of free tat 47-57 was calculated to be ~3.5 minutes (Grunwald et al., 2009) and tat 48-60 exists as an unstructured random coil in buffer solutions or when bound to lipid vesicles (Caesar et al., 2006). The secondary structure of a carrier has been suggested to affect its membrane translocation ability. However, it remains unclear if it is better for a structure to
be rich in α-helices or unordered. On the one hand, structures rich in α-helices have been suggested to be more efficient in inserting themselves and crossing lipid bilayers (Almeida & Pokorny, 2009). This is consistent with the view that α-helical structures are responsible for the membrane disrupting property of fusogenic peptides such as HA2 (Wagner et al., 1992a), melittin (Chen et al., 2006) and GALA (Subbarao et al., 1987). On the other, α-helices has also been negatively correlated with uptake (Ye et al., 2010). Instead, an unordered structure was preferred as it afforded the flexibility needed to adopt the most energetically favourable conformation during membrane translocation (Caesar et al., 2006).

<table>
<thead>
<tr>
<th>Entry</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-86</td>
<td>(Green &amp; Loewenstein, 1988)</td>
</tr>
<tr>
<td>2</td>
<td>1-72</td>
<td>(Frankel &amp; Pabo, 1988)</td>
</tr>
<tr>
<td>3</td>
<td>37-72</td>
<td>(Fawell et al., 1994)</td>
</tr>
<tr>
<td>4</td>
<td>37-57</td>
<td>(Leonetti et al., 2010)</td>
</tr>
<tr>
<td>5</td>
<td>43-60</td>
<td>(Eguchi et al., 2001)</td>
</tr>
<tr>
<td>6</td>
<td>47-57</td>
<td>(Wang et al., 2010)</td>
</tr>
<tr>
<td>7</td>
<td>48-57</td>
<td>(Wadia et al., 2004)</td>
</tr>
<tr>
<td>8</td>
<td>48-58</td>
<td>(Fittipaldi et al., 2003)</td>
</tr>
<tr>
<td>9</td>
<td>48-59</td>
<td>(Richard et al., 2005)</td>
</tr>
<tr>
<td>10</td>
<td>48-60</td>
<td>(Eiriksdottir et al., 2010)</td>
</tr>
<tr>
<td>11</td>
<td>49-57</td>
<td>(Saleh et al., 2010)</td>
</tr>
<tr>
<td>12</td>
<td>49-60</td>
<td>(Astriab-Fisher et al., 2000)</td>
</tr>
</tbody>
</table>

Table 2. Heterogeneity in tat sequences being reported in the literature. All sequences are with respect to the original sequence 1-86 in the first entry. Only one reference is provided for each entry, although there may be more discussed in text.

Proteins have been delivered by tat and are usually covalently attached. Fawell and colleagues were the first to chemically conjugate several different proteins onto tat 37-72 (Fawell et al., 1994). One of the model proteins used was β-galactosidase (β-gal), which was grafted onto tat using succinimidyl 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (SMCC) as a heterobifunctional cross-linker. The β-gal/tat conjugate was then intravenously injected into mice and found to accumulate mainly in the heart, liver and spleen. In a separate study, β-gal/tat 47-57 conjugates were generated by plasmid fusion and injected intraperitoneally into mice (Schwarze et al., 1999). The conjugates were then found to be distributed to all tissues, even across the tight blood-brain-barrier. GFP/tat 47-57 fusion protein was also delivered into cells and uptake was discovered to increase if the cells had been pretreated with 10% dimethylsulfoxide (DMSO) for one hour prior to transfection (Wang et al., 2010). Other macromolecules delivered by tat include peptide nucleic acid (PNA), which are artificial nucleic acid-mimicking polymers. PNA was either conjugated to tat 48-59 using 2-aminoethoxy-2-ethoxyacetic acid as a cross-linker (Richard et al., 2003) or linked via a disulfide bond to a cysteine-modified tat 48-60 (Lundin et al., 2008). Nucleic acids, on the other hand, are mostly electrostatically carried by tat. For instance, tat 48-60 was used to deliver reporter plasmids into HeLa cells and it was shown that its efficiency can be further improved by adding the SV40 NLS and a dendrimer of seven lysine residues (Yang et al., 2009). A short membrane active peptide, LK15 (KLLKLLLKLLKLKLLK)
was also conjugated onto tat 49-57 and used for plasmid transfection (Saleh et al., 2010). Antisense oligonucleotides were linked via disulfide bonds to cysteine-modified tat 49-60 and delivered into cells to inhibit the expression of P-glycoprotein, a transmembrane pump that is responsible for the multidrug resistance phenotype of tumor cells (Astrib-Fisher et al., 2000). siRNA has also been delivered by tat 47-57 and interestingly, photostimulation promoted the escape of tat complexes into the cytosol (Endoh et al., 2008). Reactive oxygen species produced during laser illumination was postulated to damage the endosomal vesicles for the enhanced cytosolic access. Finally, tat has also played the role of a helper. As an internalisation enhancer, tat 43-60 was displayed on the surfaces of phage particles to augment the delivery of plasmids encapsulated within the phages (Eguchi et al., 2001). In another study, tat 47-57 was used more as a NLS during the transfection of CHO cells (Moore et al., 2009). Unlike many other systems, the presence of serum augmented tat-mediated transfection (Astrib-Fisher et al., 2000; Eguchi et al., 2001).

Although tat can bring nucleic acids and other macromolecules into a cell, the pathway(s) which tat exploits to do so is ambiguous. There is little consensus in the literature, except for the observation that heparan sulphate, an anionic cell membrane glycosaminoglycan, is crucial for uptake (Sandgren et al., 2002; Tyagi et al., 2001). Besides being rapidly internalised via endocytosis (Tyagi et al., 2001), heparan sulphate is probably also involved in the initial binding step before internalisation (Ferrari et al., 2003). It can also act as a co-receptor for endocytosis (Leonetti et al., 2010). Interestingly, other glycosaminoglycans such as chondroitin sulphate and dermatan sulphate have been ruled out in the binding/internalisation of full length tat 1-86 (Tyagi et al., 2001) and tat 37-57 (Leonetti et al., 2010), but not for the shorter tat 48-60 (Sandgren et al., 2002).

Classical endocytosis is an energy- and ATP-dependent mechanism. For the sake of discussion, endocytosis can be further classified into clathrin-mediated endocytosis, caveolae-mediated endocytosis, macropinocytosis and other forms of clathrin- and caveolae-independent endocytosis. Prompted by the observation that tat 48-60 was internalised at 4°C, a temperature normally considered inhibitory for endocytosis, an energy-independent pathway was initially suggested to be responsible (Futaki et al., 2001b). Another group similarly detected low-temperature internalisation of tat 43-60 and reported that uptake was independent of endocytosis but required the presence of caveolae (Eguchi et al., 2001). The ability of free tat 47-57 to enter cells in the presence of sodium azide, which blocks intracellular ATP synthesis, also led the authors to preclude endocytosis as the uptake mechanism (Ignatovich et al., 2003).

Several models of energy-independent uptake have been proposed. A common theme seems to involve the peptide sticking to the cell membrane and creating a local mass imbalance. At the same time, it transforms itself into a more non-polar compound either due to charge neutralisation (Su et al., 2009) or hydrogen bonding (Rothbard et al., 2004) with the anionic membrane proteins. The peptide then partitions into the hydrophobic lipid bilayer (Rothbard et al., 2002) and translocates through the membrane in a process that is driven by the voltage potential across membrane (Rothbard et al., 2004) and/or the need to relieve membrane curvature stress caused by the mass imbalance in the initial step (Su et al., 2009). In stark contrast, Richard and colleagues reported that the uptake of tat 48-59, in both its free and PNA-conjugated forms were sensitive to low temperature and sodium azide, indicative of classical endocytosis (Richard et al., 2003). Endocytosis is also believed to be responsible for the uptake of protein-conjugated tat 37-72 in another study (Fawell et al., 1994). Having agreed that the uptake of tat requires energy, opinions are once again divided
regarding which of the specific endocytosis pathways involved. For instance, a study singled out clathrin-mediated endocytosis as the pathway responsible for the uptake of unconjugated tat 48-59 (Richard et al., 2005). In others, caveolae-mediated endocytosis was suggested to be the dominant form of uptake for GFP/tat 48-60 fusion protein (Ferrari et al., 2003), and also for GFP/tat 48-58 and GFP/tat 1-86 fusion proteins (Fittipaldi et al., 2003). To confound matters, macropinocytosis was reported to be mainly responsible for the uptake of PNA-conjugated tat 48-60 (Lundin et al., 2008) and tat 48-57 fusion protein (Wadia et al., 2004). In yet another twist, multiple pathways involving both clathrin-mediated endocytosis and macropinocytosis were proposed for the internalisation of unconjugated tat 48-60 (Räägel et al., 2009).

With this amount of confusion in the literature, is there hope of reconciliation? It seems that the first step is to recognise that there may not be a single mechanism responsible for the uptake of tat or indeed, other CPP that will be discussed. CPP thus cannot be taken in the same light as transferrin (Rejman et al., 2004) or cholera toxin (Torgersen et al., 2001), which are accepted to be internalised exclusively via the clathrin- and caveolae-mediated pathway respectively in a wide variety of cells. The outcome being observed therefore depends on the combination of cell/peptide/detection method chosen for that experiment and heterogeneity in conclusion can be expected due to the following reasons.

Cell: Different cell lines can use different pathways to internalise tat (Ignatovich et al., 2003). The nature of cells (primary versus immortalised or adherent versus suspension) (Eiriksdottir et al., 2010) and culture conditions such as the type of growth medium (Moulton et al., 2004), passage number of cells or even the Young’s modulus (Kong et al., 2005) of the substrate were suggested to influence plasmid transfection and uptake.

Peptide: As can be seen from Table 2, there are at least 12 distinct sequences under the same umbrella name of tat – more if one takes into account chain-end modifications. Considering that simply modifying a few residues can drastically affect the property of tat, as discussed earlier, it is understandable why the literature cannot seem to agree. The concentration of peptide used was also shown to be important. For instance, it was reported that below 10 mM, unconjugated tat 47-57 exploited both macropinocytosis and caveolae-mediated endocytosis to gain entry but once above 10 mM, a non-endocytosis mechanism seemed to operate (Duchardt et al., 2007). Perhaps there is a concentration threshold above which there is sufficient amount of surface-bound peptide to cause the mass imbalance and trigger direct transport through the membrane. Finally, the absence or presence and the nature of the cargo can also influence the mechanism of uptake. For example, unconjugated tat 47-57 was reported to enter cells using an energy-independent mechanism but endocytosis was responsible after a plasmid was bound (Ignatovich et al., 2003).

Detection method: The majority of experiments aimed at studying the endocytosis pathways of peptides uses one or both of the following techniques: 1) assaying the effects of drugs that are known to inhibit specific endocytosis pathways at the uptake (by flow cytometry or confocal microscopy) and/or the gene expression level, or 2) using a confocal microscope to trace for any co-localisation with known markers for endocytosis pathways, e.g., transferrin (clathrin), cholera toxin (caveolae) or dextran (macropinocytosis). While both are sensible experiments, one must also be aware of their limitations. Inhibitory drugs are usually pleiotropic and affect more than one pathway concurrently. For instance, dynasore has been used to shut down the clathrin-mediated pathway (Gratton et al., 2008a). However, dynasore is a drug that interferes with the activity of dynamin, an enzyme that is needed in both clathrin- and caveolae-mediated pathway to pinch off vesicles from the cell membrane.
Peptides as Promising Non-Viral Vectors for Gene Therapy (Macia et al., 2006). The same is true for methyl-β-cyclodextrin, which depletes cholesterol from the membrane (Richard et al., 2005) and cytochalasin, which disrupts actin formation (Belting et al., 2005). Both drugs can affect all three forms of endocytosis. Moreover, any drug that interferes with uptake is usually toxic and a safe effective dosage must be established in order not to affect the conclusion. There is also possible redundancy in the uptake mechanisms whereby shutting off one pathway may activate another (Rodal et al., 1999), further complicating interpretation.

Both flow cytometry and confocal microscopy are fluorescent-based techniques and rely on either the carrier or/and its cargo to be labelled with a dye. However, the process of tagging the peptide with a dye can already affect its property. For example, results from isothermal titration calorimetry showed that unlabelled tat 47-57 binds to heparan sulphate much more strongly that tat 47-57 labelled with fluorescein isothiocyanate (FITC), a florescent dye whose hydrophobicity has been implicated in the weakening of binding affinity (Ziegler et al., 2005). In another study, FITC labelling was also observed to alter the trafficking property of an octaarginine conjugate (Puckett & Barton, 2009). Significantly, fixing of cells has been shown to be undesirable and even mild fixing can cause artifactual intracellular accumulation (Richard et al., 2003). As a result of this revelation, the validity of earlier publications has been challenged, especially in studies where cells were fixed and intracellular accumulation could still be observed despite low temperature incubation. Surface-bound peptides are another source of confusion which must be separated from those which have been truly internalised (Richard et al., 2003). A brief trypsin wash is typically enough to digest and remove most of the surface-bound peptides.

Above all, it is fair to say that our understanding of endocytosis is still not perfect. As mentioned earlier, it is still not clear which other pathways, besides the clathrin-mediated one, can produce vesicles that undergo acidification. Whether vesicles from different origin eventually merge is another unresolved issue. Furthermore, both the size (Choi et al., 2006) and shape (Sharma et al., 2010) can affect the uptake of a particle and its in vivo distribution (Decuzzi et al., 2010). However, these physical parameters of tat or other CPP complexes are seldom reported. Great care must also be taken when reporting particle sizes based on light scattering as a recent survey has estimated that up to 90% of the published figures can be erroneous due to inappropriate assumptions being made during measurements (Keck & Muller, 2008). With such a plethora of factors that can affect outcome, it seems more realistic to accept that the uptake of tat or its conjugates cannot be ascribed any particular route.

4.3 Oligoarginine (Arg<sub>n</sub>, usually n = 7 to 9)

The potential of using arginine for gene delivery was firmly established when it was shown that replacing most of the arginine residues in tat 1-86 with alanine drastically reduced internalisation (Tyagi et al., 2001). Replacing the arginine residues within tat 48-60 with lysine had a similar effect and abolished internalisation (Thoren et al., 2004). Clearly, arginine residues are crucial. However, another important (but less obvious) message from these experiments is that arginine is more than just a cationic residue as replacing it with lysine, another cationic residue, is not good enough.

Oligoarginine of various lengths have been evaluated. In one study, oligoarginine with 4, 6, 8, 10, 12 or 16 residues were compared and it was found that both arg<sub>4</sub> and arg<sub>16</sub> were poorly taken up by cells (Futaki et al., 2001b). Instead, arg<sub>8</sub> exhibited the highest rate of internalisation both in its free form or when linked via a disulfide bond to a model carbonic anhydrase protein. This demonstrated that cationic charge was not sufficient to afford
efficient uptake. Consistent with this, another study showed that arg7 was internalised much more efficiently into Jurkat cells than lys7 and his7 (Mitchell et al., 2000). There was therefore something unique about the chemical structure of arginine that cannot be explained simply by it carrying a cationic charge. Indeed, the guanidine (pK_a~12) sidegroup of arginine was determined to be key as replacing it with a urea (pK_a~0.1) sidegroup removed internalisation. Urea differs from guanidine by only a single nitrogen atom, which in the former is replaced by an oxygen atom. However, this simple modification removed not only the ability of urea to protonate at neutral pH, but also its ability to form hydrogen bonds. This ability to form stable hydrogen bonds with the anionic phosphates and sulphates on cell membrane was then suggested to be the feature that distinguished arginine from lysine and histidine in terms of translocation efficiency (Mitchell et al., 2000).

The spacing between arginine residues can also influence internalisation (Rothbard et al., 2002). To study this, a library of oligoarginines was synthesised, all with seven residues but separated by 1-6 spacers at all the possible permutations. It was argued that the addition of spacers imparted flexibility to the arginine backbone which was important for better membrane translocation.

A length of 7-9 arginines is usually preferred and the chains exist as random coils in buffer solutions and when bound to lipid vesicles (Caesar et al., 2006). Both D- and L-arginines have been explored with some authors preferring D-arginines (Hyun et al., 2010; Puckett & Barton, 2009) and others finding no difference between the stereoisomers (Mitchell et al., 2000; Nakase et al., 2004). Cargoes of different nature have been delivered. Negatively charged QCs were electrostatically bound to arg8 and delivered into adipose tissue-derived stem cells for imaging purposes (Yukawa et al., 2010). The anti-cancer drug, taxol was covalently bound to arg8 via disulfide bonds using a novel linker to increase its water solubility and uptake. The drug was also designed to be released intracellularly so as to combat multidrug resistant cell lines which otherwise have limited accumulation of chemotherapeutic drugs (Dubikovskaya et al., 2008). The immnosuppressive drug, cyclosporine A, was coupled with arg7 using a novel pH-sensitive linker and used for topical delivery in a skin inflammation model (Rothbard et al., 2000). Phosphorodiamidate morpholino oligomers (PMO) are antisense molecules that interfere with mRNA translation but structurally differ from nucleic acids in several aspects. In one study, PMO was electrostatically carried by an arg9-based carrier, but the efficiency was not as good as covalently linked ones (Moulton et al., 2004). In another study, a short peptide sequence corresponding to the C-terminus of the cystic fibrosis transmembrane regulator was joined to arg7 during synthesis and laser illumination was used to trigger the release of the conjugate into the cytosol of U2OS cells (Maiolo-III et al., 2004). siRNA was electrostatically carried by an arg8-based vector and used to transfect mouse bone marrow-derived dendritic cells (Akita et al., 2010). Finally, D-arg9 with cysteines added on both terminals (i.e., cys-arg9-cys) was crosslinked via disulfide bonds to form a mesh of reducible poly(oligoarginine) (Hyun et al., 2010). The carrier was then used to deliver plasmids encoding for heme oxygenase-1 (useful for the treatment of ischemia/reperfusion-induced brain stroke) by direct injection into the brain of mice.

Several modifications have been made to oligoarginine to improve its transfection efficiency. For example, arg8 has been combined with GALA, a pH sensitive fusogenic peptide to enhance its endosomal escaping property (Akita et al., 2010). The addition of a hydrophobic stearyl chain to arg8 was also shown to greatly increase its transfection ability (Futaki et al., 2001a). This was suggested to be due to a better association between the hydrophobic moieties on the vector and the lipid bilayer (Putnam, 2006). Indeed, a certain degree of hydrophobicity is a

www.intechopen.com
common feature of many efficient vectors and this was clearly shown in a recent study where a novel triblock peptide, phobic-his$_5$-arg$_8$, was developed (Seow & Yang, 2009b; Seow et al., 2009). This design featured a block of five hydrophobic residues (tryptophan, phenylalanine or isoleucine, in order of increasing hydrophobicity) for enhanced uptake, a middle block of four histidines for buffering capacity and a third block of arg$_8$ for DNA binding and membrane penetration. Removing the hydrophobic block drastically reduced the efficiency of the carrier at both the uptake and gene expression level. Transfection efficiency can also be modulated just by changing the hydrophobicity of the hydrophobic block. The buffering capacity of the carriers was also assessed in a series of acid-base titration experiments and shown to vary with the length of the histidine block used. Each block was then systematically studied and an optimised sequence was used to mediate reporter plasmid expression by direct local injection into mice bearing 4T1 tumors. In another study, cholesterol was added as the hydrophobic end of a his$_5$/10-arg$_{10}$ vector for gene delivery (Guo et al., 2008).

The internalisation of oligoarginine is sharply inhibited by the presence of heparin, which competes with heparan sulphate for binding (Seow et al., 2009). This shows that heparan sulphate is crucial for uptake. Low-temperature internalisation was observed and energy-independence was originally proposed to characterise the uptake of oligoarginine (Futaki et al., 2001b). However, those observation was made with fixed cells. In response to studies that had exposed the flaws of using fixed cells (Richard et al., 2003), authors of the original study re-evaluated the uptake mechanism using live cells. Consequently, they reported macropinocytosis to be a major, but not exclusive pathway for the uptake of oligoarginine (Nakase et al., 2004). Macropinocytosis was also proposed to be the main form of uptake for protein-conjugated (Takayama et al., 2009) and siRNA-bound arg$_8$ (Akita et al., 2010). Nonetheless, an energy-independent form of uptake was still suggested if the concentration of unconjugated arg$_9$ exceeded 10 mM – otherwise, macropinocytosis and caveolae-mediated, but not clathrin-mediated endocytosis seemed to be important (Duchardt et al., 2007). Another study reported that both clathrin-mediated endocytosis and macropinocytosis play a role for the uptake of unconjugated arg$_9$ (Räägel et al., 2009). On the other hand, all three forms of endocytosis (clathrin-mediated, caveolae-mediated and macropinocytosis) were found to be responsible for the uptake of plasmid-bound oligoarginine complexes (Seow et al., 2009). Finally, an energy-dependent but non-endocytosis mechanism was also proposed to be responsible for uptake (Mitchell et al., 2000). Like tat, there is little consensus over the internalisation pathway(s) of oligoarginine, although endocytosis is generally accepted to play a key role in most cases (Richard et al., 2003; Rothbard et al., 2000).

### 4.4 Penetratin

Homeoproteins are regulatory proteins essential for proper physical development. The DNA binding domain of these proteins is made up of a highly conserved sequence known as the homeobox. The homeobox of the *Drosophila* antennapedia gene (pAntp) is 60-amino acid long (sequence 1-60) and contains three $\alpha$-helices (Derossi et al., 1994). pAntp was first discovered to effectively translocate into nerve cells and accumulate within their nuclei (Joliot et al., 1991). To demonstrate the usefulness of pAntp as a carrier, a 33-residue peptide cargo corresponding to the C-terminus of rab3, a GTP-binding protein in human, was linked to pAntp via plasmid fusion and shown to translocate into the nuclei of myoblasts and neurons (Perez et al., 1992). A study later revealed that the third helix was actually the domain driving internalisation (Roux et al., 1993). Based on this, penetratin, a 16-residue
peptide within the third helix (sequence 43-58 of the original pAntp) was described (Derossi et al., 1994). Penetratin demonstrated membrane penetrating ability but was prone to aggregation. Aggregated particles can enjoy better internalisation due to a sedimentation effect which promotes a more intimate particle-cell contact (Luo & Saltzman, 2000b). While this may have helped the internalisation of penetratin, aggregation alone was excluded as the dominant reason for its efficient internalisation (Derossi et al., 1994).

The secondary structure of penetratin in buffer is unstructured but becomes α-helical when bound to lipid vesicles (Caesar et al., 2006). Raman microscopy on live cells further showed that penetratin within the cytosol was either unstructured or in the β-sheet conformation (Ye et al., 2010). The arginine residues within penetratin are important as replacing them with lysine greatly reduced its translocation ability (Caesar et al., 2006).

Various cargoes have been carried by penetratin. Antisense oligonucleotides (Astriab-Fisher et al., 2000), PNA (Lundin et al., 2008) and luciferin protein (Eiriksdottir et al., 2010) have all been coupled onto a cysteine-modified penetratin via disulfide bonds and delivered into cells. siRNA was electrostatically bound to penetratin and although the complexes accumulated favourably within cells, there was limited silencing activity (Lundberg et al., 2007). HA2 was then added to enhance endosomal escape but produced only a modest improvement. Penetratin has been evaluated in vivo. In one study, penetratin was directly injected into a rat’s brain and the peptide was observed to spread away from the site of injection (Bolton et al., 2000). However, dosage-dependent cell death and inflammatory responses were also provoked.

Penetratin was initially suggested to enter cells using energy-independent mechanisms (Derossi et al., 1994; Perez et al., 1992; Roux et al., 1993), but a caveat is that fixed cells were used in those studies. Nonetheless, an energy-independent mechanism was still suggested to be possible past a concentration threshold of 40 mM, below which all three forms of endocytosis should dominate (Duchardt et al., 2007). Endocytosis was also reported to be responsible for the uptake of siRNA-bound penetratin (Lundberg et al., 2007) while macropinocytosis was suggested to be the main form of uptake for PNA-conjugated penetratin (Lundin et al., 2008).

4.5 Transportan (Tp) and Tp10

Tp uses a lysine residue to join the first 12 N-terminus residues of the neuropeptide, galanin, to the 14 C-terminus residues of the wasp venom, mastoparan. Tp is thus 27 residues long (sequence 1-27) and was first shown to penetrate Bowes’ melanoma cells rapidly and efficiently (Pooga et al., 1998). However, Tp at a high concentration was found to inhibit the GTPase activity of cells. To overcome this side effect, a series of Tp analogues was prepared (Soomets et al., 2000). This led to the identification of Tp10 which, upon deleting the first 6 residues of Tp (i.e., sequence 7-27 remaining), was internalised as efficiently as Tp but did not have any effect on the GTPases.

PNA (Lundin et al., 2008) and luciferin (Eiriksdottir et al., 2010) were conjugated to both cysteine-modified Tp and Tp10 via disulfide bonds. siRNA was also electrostatically bound to Tp10 but had little silencing effects when transfected into cells (Lundberg et al., 2007). This was surprising insofar as Tp10 was shown earlier to mediate good levels of plasmid expression. This however resonates with comments made by other researchers (Mahon et al., 2010; unpublished observation) who had observed that a carrier’s efficiency in plasmid delivery may not automatically apply to the delivery of the shorter and stiffer siRNA molecules. Earlier studies with fixed cells reported low temperature internalisation of Tp.
Peptides as Promising Non-Viral Vectors for Gene Therapy

(Pooga et al., 1998) but more recent observations have linked endocytosis to the uptake of both Tp and Tp10 (Lundin et al., 2008).

4.6 Other notable peptide sequences

VP22 is a 300-residue long peptide derived from the structural protein of HSV-1 and has been used successfully to deliver GFP as a fusion protein (Elliott & O’Hare, 1997). When COS-1 cells were microinjected with plasmids encoding for VP22, an interesting pattern was observed in which a central cell would first express VP22 and exhibit cytosolic staining when probed with anti-VP22 antibodies. The VP22 peptide was then excreted and could infect neighbouring cells before eventually localising to their nuclei. CADY is a 20-residue long peptide evaluated for siRNA delivery (Crombez et al., 2009). It changes from unordered to being α-helical in the presence of lipid vesicles which then drives its internalisation independently of the endosomal pathway (Konate et al., 2010). MPG is a 27-residue peptide designed to have a hydrophobic domain (sequence 1-17) derived from HIV gp41 and a NLS domain (sequence 21-27) derived from SV40 separated by a “trp-ser-gln” unit (Morris et al., 1997). The uptake of MPG/oligonucleotide complexes was shown to be rapid and independent of the endosomal pathway. Other carriers explored for delivery include KALA (Wyman et al., 1997) and EB1 (Lundberg et al., 2007).

Throughout the discussion, various peptides with fusogenic and nuclear localising properties were introduced. SV40 is the classical NLS and different sequences exist. The most commonly used sequence for SV40 is provided in Table 1. Other sequences include “glu-asp-pro-tyr” (Trentin et al., 2005) and “glu-pro-tyr-cys” (Moore et al., 2009) being added onto the C-terminus and an even longer form of SV40 has been described (Eguchi et al., 2001). INF 5 and INF 7 are examples of two commonly used fusogenic peptide and both are derived from the parent HA2 peptide (Plank et al., 1994). GALA is a 30-amino acid long pH-sensitive fusogenic peptide (Subbarao et al., 1987). Upon protonation of its glu residues, its secondary structure changes from unordered at neutral pH to being α-helical in acidic environments. Both the amphipathicity and degree of α-helicity have been correlated with the ability of GALA to interact and destabilise membranes (Parente et al., 1990).

5. Conclusion

This chapter started by discussing the challenges and intracellular barriers associated with the delivery of nucleic acids. Strategies used to overcome these hurdles were next examined, mainly in the context of peptide-derived vectors. It is clear that peptide carriers are not characterised by any typical sequences, although the majority of current designs rely on cationic residues to bind nucleic acids. This strategy, nonetheless, may be problematic during in vivo applications. An area that deserves more attention is the development of vectors that can bind nucleic acids using non-electrostatic forces, e.g., by including more hydrophobic residues. The stability and in vivo behaviour of such vectors then need to be thoroughly evaluated. Another challenge is to further improve strategies that are already in place to shield the cationic charges, e.g., by attaching PEG chains that are intracellularly cleavable so as to increase the circulation time of the complexes without compromising excessively on transfection efficiency. Advances in either will enable peptides to further realise their potential as a class of non-viral vector.
6. References


www.intechopen.com


This book focuses on recent advancement of gene delivery systems research. With the multidisciplinary contribution in gene delivery, the book covers several aspects in the gene therapy development: various gene delivery systems, methods to enhance delivery, materials with modification and multifunction for the tumor or tissue targeting. This book will help molecular biologists gain a basic knowledge of gene delivery vehicles, while drug delivery scientist will better understand DNA, molecular biology, and DNA manipulation.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
