Methods for Gene Delivery

<div style="text-align:right">**3**</div>

The success of any gene transfer procedure, either through *in vivo* inoculation of the genetic material or after gene transfer into the patient's cells *ex vivo*, strictly depends upon the efficiency of nucleic acid internalization by the target cells. As a matter of fact, making gene transfer more efficient continues to represent the most relevant challenge to the clinical success of gene therapy.

3.1
Cellular Barriers to Gene Delivery

The plasma membrane lipid bilayer, which is apolar and hydrophobic, constitutes an impermeable barrier to large and charged macromolecules such as DNA and RNA, since, at physiological pH, phosphates in the nucleic acid backbone are deprotonated and thus negatively charged. Therefore, entry of these polyanions into the cells needs to be facilitated, usually by exploiting the same cellular mechanisms that normally allow macromolecule internalization. Alternatively, nucleic acids can be vectored into the cells within biological particles that are naturally capable of crossing biological membranes, such as viruses.

3.1.1
Endocytosis

In physiological conditions, entry of large, polar macromolecules into the cells occurs through a mechanism involving formation of membrane vesicles at the cell surface, followed by their internalization and intracellular trafficking. This process is collectively known as "endocytosis". Over the last several years, a number of different endocytosis mechanisms have been discovered, which are distinguished by the size of the vesicle formed and the molecular machinery involved. The four best understood and most relevant endocytosis pathways are depicted in Figure 3.1.

3
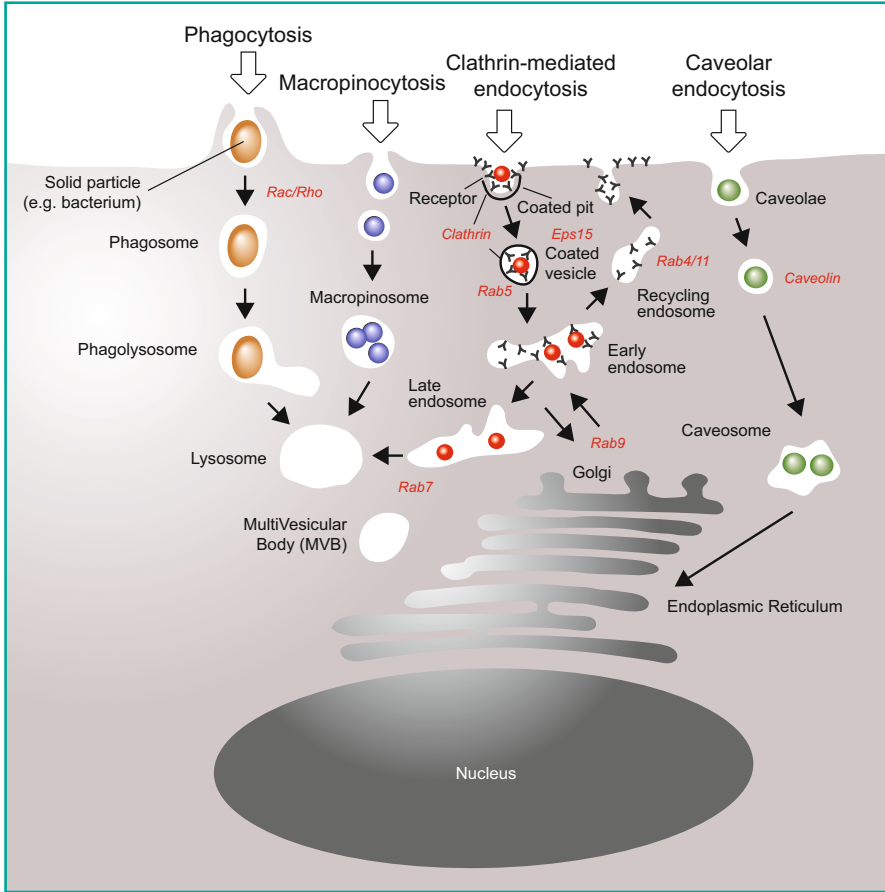
Fig. 3.1 Endocytosis. The four major mechanisms of endocytosis are schematically shown. See text for discussion

  Phagocytosis (literally: *cell-eating*) is the process by which specialized eukaryotic cells (in mammals, typically neutrophils and macrophages) internalize large particles (>500 nm in diameter, including cells that have undergone apoptosis or bacteria). Cell surface receptors bind the extracellular particles, initiate local intracellular signals, and reorganize the actin cytoskeleton to induce the change in cell shape needed to engulf the particle, eventually resulting in the formation of a large vacuole known as a *phagosome*. A similar process also occurs in *macro-pinocytosis* (literally, *cell-drinking*), where large endocytic vacuoles (typically >500 nm in diameter) continuously form, resulting in the engulfment of large amounts of both extracellular medium and plasma membrane, containing solutes and single molecules such as proteins, which are thus internalized in a nonspecific manner. Both these types of endocytosis eventually end up in the fusion of the vesicles with lysosomes, which represent the major hydrolytic compartment of the cell. These are large (1–2 μm in diameter), acidic (~pH 4.5) vacuoles containing a variety of

acid hydrolases (lipase, carbohydrases, nucleases, proteases), targeted to these organelles from the Golgi apparatus through the addition of a mannose-6-phosphate tag.

Receptor-mediated or *clathrin-mediated endocytosis* is a more specific, active event where the plasma membrane folds inward to form pits coated with the cytosolic protein clathrin. These regions of the plasma membrane usually contain specific protein receptors, including the receptors for transferrin, low-density lipoproteins, growth factors, antibodies and several others. Once these receptors bind their specific ligands, an active process is triggered by which small (~100 nm in diameter) vesicles form, having a morphologically characteristic crystalline coat made up of a complex of proteins associated with clathrin (*clathrin-coated vesicles*, CCVs). These vesicles progressively undergo maturation to first become early endosomes, which show a tubulo-vesicular morphology (vesicles up to 1 µm in diameter connected by tubules of ~50 nm in diameter) and mildly acidic pH. These are principally sorting organelles, where many ligands dissociate from their receptors due to the acidic pH and from which many of the receptors recycle to the cell surface. Most of the early endosomes mature into late endosomes, which essentially receive internalized material en route to lysosomes.

A fourth macromolecule internalization pathway is *caveolar endocytosis*. Caveolae are non-clathrin-coated, plasma membrane flask-shaped invaginations (~50 nm in diameter), which exist on the surface of several cell types, including adipocytes, endothelial cells, smooth muscle, and fibroblasts. These microdomains are often associated with the protein caveolin – which, however, is not essential *per se* to the process of clathrin-independent endocytosis – and mostly correspond to the regions of the plasma membrane in which the lipid bilayer is enriched in cholesterol and sphingolipids and is characterized by decreased fluidity; caveolae thus represent one category of the detergent-resistant microdomains of the cell membrane collectively known as *lipid rafts*. Caveolae bud from the plasma membrane and lead to the formation of caveolin-containing endosomes, named *caveosomes*. These are pH-neutral, long-lived compartments, which are known to eventually fuse with the endoplasmic reticulum (ER) or the Golgi apparatus, thus delivering their contents into these compartments. Similar to clathrin-mediated endocytosis, and unlike the other types of endocytosis, caveolar endocytosis requires the GTPase enzyme dynamin. Of note, all types of endocytosis except caveolar endocytosis end up with the delivery of the internalized material into the lysosome compartment, where it is destined for degradation. This has obvious relevance for both gene and drug delivery, since escape from degradation is an essential requisite for efficient treatment.

3.1.2
Escape from the Intracellular Vesicle Compartment

The material internalized into the cells within endosomal vesicles is outside of the cytoplasm, thus still virtually resident in the extracellular environment. Therefore, efficient gene delivery requires the nucleic acids to pass through a biological membrane bilayer, either by destruction of the membrane itself or by physical passage. In this respect, nature has evolved different mechanisms allowing entry into the cytoplasm of macromolecules contained inside the various vesicle compartments, which are variously exploited by a

3

number of cell pathogens to enter the cytoplasm and the nucleus. In particular, viruses and toxins essentially use two main pathways to gain access to the cytosol: they can either be transferred from early or late endosomes into the cytosol in response to low pH, or enter the Golgi and then the ER and be transferred to the cytosol from this destination. An example of the first mechanism is offered by the diphtheria toxin, for which the relatively low pH found in endosomes triggers a conformational change in the toxin that drives formation of a membrane pore allowing its direct entry into the cytosol. Other molecules instead reach the Golgi apparatus and the ER following a route in reverse of the classic secretory pathway. These molecules include plant toxins, such as ricin, and bacterial toxins, such as Shiga toxin, part of cholera toxin and *Pseudomonas* exotoxin A; these often consist of two chains or domains, one responsible for cellular internalization and trafficking, and the other one exerting the biological effect proper of the toxin. Classical secretion is characterized by transport of newly synthesized proteins from the ER to the Golgi, followed by budding of vesiculated cargos from the trans-Golgi network (TGN), vesicle sorting in the cytoplasm, and eventual vesicle fusion with the plasma membrane. Toxin cell entry is enabled by binding to a cell surface molecule, followed by endocytosis. Once in the early or late endosomes, toxins escape lysosomal degradation by re-routing the vesicles toward the TGN and, from this, to the ER, either directly or through the Golgi apparatus. Once toxins have reached the ER, the cytoplasm is accessed by taking advantage of the ER protein auditing system known as ERAD (ER-associated protein degradation). This mechanism eliminates misfolded proteins from the ER by discarding them into the cytoplasm through a pore known as the Sec61 translocon, which is actually also used by several of the toxins. Finally, certain mammalian viruses, such as polyomaviruses, influenza viruses, coronaviruses, and echoviruses, and some toxins, such as cholera toxin, use caveolae-mediated transport from the cell surface to reach the ER. This has been particularly studied for the simian virus 40 (SV40), which is internalized from lipid rafts and, through caveosomes, is eventually released into the ER, from which it escapes to the cytoplasm and gains access to the nucleus via the nuclear pore complex.

In the case of viruses of interest to the gene therapy field, access to the cytoplasm is either attained by direct fusion of the viral envelope with the cell plasma membrane (for retroviruses) or by escape from the endocytic degradation route through the endosomolytic activity of the viral capsid (for adenoviruses and AAVs).

3.1.3
Nuclear Targeting

Finally, a therapeutic nucleic acid, once escaped from the intracellular vesicles, must find its way to the relevant subcellular compartment where its function is exerted, which is usually either the cytosol or the nucleus. Short regulatory RNAs are often active in the former compartment, while coding genes must access the latter to be transcribed. The final destination of the nucleic acid is commonly dictated by the proteins to which it binds once in the cytosol. For example, siRNAs are loaded onto the RNA-induced silencing complex (RISC) and remain in the cytosol (see section on 'Small Regulatory RNAs'). In contrast, coding nucleic acids are bound by various DNA binding proteins, including transcription factors, which

direct them to the nucleus thanks to their nuclear localization signal (NLS), or gain access to the nucleus during mitosis. Nuclear targeting can be enhanced through the delivery of nucleic acids complexed with short peptides binding to proteins of the importin/karyopherin family, a set of proteins that actively transport macromolecules into the nucleus.

3.1.4
Methods for Gene Delivery: An Overview

The gene transfer methodologies that are clinically exploitable by gene therapy can be divided into four categories.

(i) Simple utilization of naked plasmids (circular, covalently closed DNA molecules) or short regulatory nucleic acids (oligonucleotides, siRNAs, and others), not complexed with other molecules and simply injected *in vivo* or added to the extracellular milieu of cultured cells.

(ii) Facilitation of nucleic acid entry into the cells by physical methods.

(iii) Transport of nucleic acids into the cells by lipofection.

(iv) Embedding of nucleic acid sequences within viral genomes, then exploiting the natural property of viruses to enter target cells at high efficiency.

Table 3.1 reports a concise view of the main advantages and disadvantages of these methodologies. The production of synthetic DNAs and RNAs, or the use of plasmid DNA produced in bacteria, which can be obtained in large quantities, offer important advantages in terms of simplicity and safety of use compared to viral vectors. Indeed, the efficiency of viral vectors is strictly related to the biological characteristics of the parental virus they derive from, which, in several cases, are still not very well understood. Furthermore, production of viral vectors requires the development of complex procedures based on cell culture and infection in order to obtain packaging of the therapeutic nucleic acids inside the virions. These procedures pose important safety problems, due to the possibility that either the packaging cells contain other infectious agents or the viral vector itself might recombine to generate a replication-competent virus. Finally, some viral vectors induce a powerful inflammatory and immune response once injected into the patients. Despite these problems, the efficiency of gene transfer that can be attained by viral gene transfer both *in vivo* and *ex vivo* is by far superior to that of any non-viral method. In addition, only viral vectors allow persistent, often permanent, expression of the therapeutic gene in their target tissues *in vivo*.

Once entered into the cells, the fate of the delivered nucleic acid strictly depends on its internalization route and chemical structure: plasmids, oligonucleotides, and small RNAs are usually degraded and lost with a kinetics ranging from a few hours (for small RNAs) to several days (for plasmids). However, synthetic nucleic acids, including oligonucleotides, siRNAs, and aptamers, can be chemically modified in order to escape degradation by cellular nucleases; in this way, their persistence inside the cells is significantly increased. When therapeutic nucleic acids are carried into the cells by viral vectors, their destiny depends on the biological characteristics of the vector that is used. Vectors based on adenoviruses persist for prolonged periods in an episomal, non-integrated form inside the nucleus of the transduced cells; however, the cells themselves are usually rec-

Table 3.1 Pros and cons of the major gene transfer procedures for gene therapy

Strategy	Method	Pros	Cons
Naked DNA or RNA	Direct injection *in vivo*	Simplicity of production and use Potential use as genetic vaccines	Low efficiency Transitory effect Internalization only in skeletal and cardiac myocytes and in antigen presenting cells
Physical methods	Electroporation	Relatively easy to set up for skeletal muscle and skin; invasive for other organs	Low efficiency Transitory effect Limited spectrum of applications
	Increase of hydrodynamic pressure	Usually invasive	Low efficiency Transitory effect
	Ultrasounds (sonoporation)	Relatively easy to set up	
	Bombardment with DNA-coated gold particles (gene gun)	Relatively easy to set up Stimulation of an effective immune response	Limited to gene transfer to the skin
	Jet injection		
Chemical methods	Liposomes Cationic lipids Cationic polymers Proteins	Relatively easy to set up and use	Relatively low efficiency Transitory effect
Viral vectors	Vectors based on: gammaretroviruses, lentiviruses, adenoviruses, adeno-associated viruses (AAVs), herpesviruses	High efficiency of gene transfer both *in vivo* and *ex vivo* For some vectors, persistence of therapeutic gene expression *in vivo*	Possible induction of immune and/or inflammatory response Limited cloning capacity Complexity of production For some viruses, tropism limited to specific cell types Insertional mutagenesis (for gammaretroviruses) In most cases, incomplete knowledge of the molecular mechanisms governing viral replication

ognized and destroyed by the immune system in a 1–2-week period. The same immune response prevents any possibility to re-inoculate a vector displaying the same serotype. In contrast, vectors based on the adeno-associated virus (AAV) persist in episomal form in

non-replicating cells for month- or year-long periods, while retroviral vectors become integrated into the host genome. Both these vectors are thus useful for applications in which prolonged or permanent gene expression is desirable.

Given the broad spectrum of biological properties displayed by both non-viral and viral methods for gene transfer, it is evident that no perfect universal system exists. Thus, the choice of the proper gene transfer methodology strictly depends on the characteristics of the disease for which gene therapy is developed and the attainable modality of gene transfer.

The main gene transfer methodologies that are currently available for gene therapy are described and discussed in the following sections.

3.2
Direct Inoculation of DNAs and RNAs

As outlined above, the chemical and physical characteristics of the plasma membrane prevent the direct passage of large and charged macromolecules, such as plasmid DNA. Different cell types, however, have the capacity to internalize small nucleic acids, including oligodeoxynucleotides, RNA decoys, or siRNAs through an active endocytic process, usually exploiting the clathrin-mediated endocytosis pathway. As discussed above, most of the content of the endocytic vesicles formed along this pathway is destined to lysosomal degradation. However, a tiny fraction can escape the early or late endosomes, cross the endosomal membrane, gain access to the cytosol, and, from this compartment, be transported to the nucleus. This process, although highly inefficient, forms the basis of a few clinical trials taking advantage of chemically modified oligonucleotides or siRNAs, administered to the patients intravenously or injected in anatomically defined compartments such as the eye's posterior chamber in the form of naked nucleic acids (cf. sections on 'Gene Therapy of Cancer' and 'Gene Therapy of Eye Diseases' respectively).

In some cell types, the process of internalization of extracellular nucleic acids and their release into the cytosol is relatively more efficient. In particular, this is the case of striated muscle fibers and cardiomyocytes, which are also able to internalize naked plasmids simply injected into the skeletal muscles or the heart *in vivo*. This property, although modest, is exploited by a few clinical trials, especially aimed at inducing therapeutic angiogenesis in the ischemic tissues (cf. section on 'Gene Therapy of Cardiovascular Disorders').

Another cell type showing striking capacity to internalize plasmid DNA present in the extracellular milieu is the professional antigen-presenting cell (APC). These cells, which comprise dendritic cells – including Langerhans cells of the skin – macrophages, and B-lymphocytes, are very efficient at internalizing foreign antigens, by either phagocytosis or receptor-mediated endocytosis, followed by their processing and presentation to T cells via both MHC class II and MHC class I molecules. Although the process of internalization of plasmid DNA by these cells has limited efficiency, this is still sufficient to induce the intracellular synthesis of the proteins encoded by the plasmids, followed by their presentation as antigens to induce an immune response. This process is exploited by the strategy of genetic vaccination (or DNA vaccination) and will be discussed in the section on 'Gene Therapy of Cancer'.

In all other cases, entry of both short nucleic acids and large plasmids into the cells needs to be facilitated by physical or chemical treatments, or by using viral vectors.

3.3
Physical Methods

Over the last several years, relevant progress has been made in the utilization of physical methods to facilitate entry of plasmid DNA or short regulatory DNAs or RNAs into the cells. These methods are essentially aimed at bringing the nucleic acids in strict contact with the plasma membrane and/or determining the temporary localized disassembly of the membrane itself.

3.3.1
Electroporation

Electroporation (also termed *electropermeabilization* or *electrotransfer*) was originally developed as a means to deliver genes into cultured cells. Subsequently, it has also been applied to *in vivo* gene transfer to the skin, muscle, liver, and, more recently, to a variety of other organs, including the kidney, lung, heart, and retina. The method consists in the application of a series of electric pulses (typically, in the order of ~200 V/cm for tens of milliseconds, or higher voltages for microseconds) in order to induce a transient increase in membrane permeability and thus allow entry of large, charged macromolecules, including plasmid DNA. To discharge the pulse, two electrodes of various shapes are positioned flanking the site of inoculation of a solution containing the nucleic acids. The electric pulse induces the formation of hydrophilic pores in the cell membrane and the subsequent passive passage of DNA through these pores thanks to a local electrophoretic effect. At the end of the stimulus, the membrane acquires its normal properties again.

One of the most important problems of electroporation is the induction of tissue damage due to the electric pulses, which essentially limits the application of this technology. In addition, the expression of the internalized plasmid DNA is often transitory and usually lost within a few days. In the skeletal muscle, which represents one of the most interesting tissues for electroporation, the efficiency of gene transfer can be increased by the administration, prior to electric discharge, of the enzyme hyaluronidase, which degrades hyaluronic acid in the extracellular matrix and thus increases gene transfer efficiency by favoring diffusion of the nucleic acids.

3.3.2
Hydrodynamic Intravascular Injection

Transient local increase of hydrostatic pressure significantly augments cellular internalization of nucleic acids circulating in the blood. This strategy, named *hydrodynamic gene transfer*, first allows DNA or RNA to cross endothelial cell junctions, by inducing their separation, and later determines the transient formation of pores or microdefects in the plasma

membranes of the target cells underlying the endothelium, a process named *hydroporation*.

Hydrodynamic gene transfer can be applied to different organs *in vivo*, including liver, skeletal muscle, and heart. An increase of hydrostatic pressure can be generated by injecting a solution containing the plasmid, oligonucleotide, or siRNA of interest at high pressure into the relevant area (for example, into the femoral artery to achieve diffuse transfection of the lower limb skeletal muscles), or by transiently occluding the veins draining from the area (for example, the superior vena cava for the diaphragm or the coronary sinus for the heart), in order to selectively increase blood pressure in the region where the therapeutic gene is injected.

3.3.3
Sonoporation

Both electroporation and hydrodynamic gene transfer are quite invasive, and thus difficult to apply for gene transfer to most organs. In contrast, ultrasound waves are used clinically for a variety of diagnostic and therapeutic applications. Low-intensity ultrasound permits a number of non-invasive diagnostic examinations (echography), while high-intensity ultrasound is used for the non-invasive treatment of urinary calculosis (extracorporeal shock wave lithotripsy, ESWL) and high-intensity focused ultrasound (HIFU) finds application for the thermal destruction of tumors. All these different modalities of ultrasound application can facilitate the transfer of plasmids and other small nucleic acids into the cells. This methodology is also collectively known as *sonoporation*.

The facilitation of gene transfer by sonoporation is due to the capacity of ultrasound to generate acoustic cavitation, which ultimately determines formation of micropores in the plasma membrane. Cavitation is increased by agents causing nucleation, such as echographic contrast agents based on gas microbubbles (typically, microbubbles filled with perfluoropropane with an albumin shell). In this case, rupture of the microbubbles caused by ultrasound increases permeability of the membrane and thus facilitates gene transfer.

Sonoporation can be achieved by injecting a plasmid or an oligonucleotide in the blood and focusing the ultrasound beam on a specific body region, typically the vascular wall, the heart, or the skeletal muscle; entry of nucleic acids into the endothelial cells, cardiomyocytes, or skeletal muscle fibers, respectively, is induced by the local and transitory increase of membrane permeability.

Notwithstanding the relative ease of assembly and the non-invasiveness of this procedure, the extent of gene transfer using sonoporation is still difficult to standardize and very variable according to the different experimental conditions.

3.3.4
Bombardment with DNA-Coated Microparticles ("Gene Gun")

Among the physical methods for gene transfer, one very interesting approach is the possibility of delivering DNA into the cells by bombarding them with micron-sized beads carrying plasmid DNA adsorbed onto their surface. The most utilized version exploits a special type of gun ("gene gun") shooting gold or tungsten particles at very high velocity into

the tissues. These particles can easily cross the cell and nuclear membranes and release the DNA adsorbed on their surface into the nucleus. This method, also named *ballistic* or *biolistic transfection*, is derived from experience gained from gene transfer in plants, where it was originally invented as a way to cross the rigid plant cell wall.

Biolistic transfection now finds application for gene therapy of accessible tissues, such as skin, with the main purpose of delivering genes coding for antigenic protein in the context of DNA vaccination against tumor or viral antigens. Once in the dermis, microparticles are taken up by APCs, which can thus process the encoded antigens and present them to T lymphocytes for immune stimulation (see section on 'Gene Therapy of Cancer'). Another tissue in which biolistic bombardment was successfully performed in animal experimental models is the cornea.

3.3.5
Injection of DNA using High-Pressure Jets ("Jet Injection")

One possibility to facilitate entry of naked DNA or RNA into the cells is so-called *jet injection*. In this technology, a solution containing the nucleic acid is applied to the skin as a jet of high velocity with the force to penetrate the skin and the underlying tissues, thus determining spread transfection of the areas of interest. Jet injection has deeper penetration capacity compared to ballistic bombardment (down to 1 cm in depth) and, besides the skin and the underlying tissues, can also be applied to other accessible tissues, including solid tumors. A Phase I clinical study was performed in which jet injection was used for gene therapy of skin metastasis in patients with breast cancer and melanoma to assess the extent of gene transfer.

3.4
Chemical Methods

The purpose of physical methods for gene transfer is to facilitate entry of nucleic acids into the cells essentially by modifying the properties of biological membranes using physical forces such as pressure or electricity. Instead, chemical methods are aimed at modifying the properties of nucleic acids themselves, by promoting their association with molecules able to reduce their hydrophilicity and neutralize their charge, ultimately leading to an increased cellular uptake.

The molecules used to facilitate gene transfer can be classified into one of three categories: lipids (liposomes and cationic lipids), proteins, and cationic polymers.

3.4.1
Liposomes and Cationic Lipids (Lipofection)

Liposomes are closed vesicles formed by one or more lipid bilayers surrounding a core aqueous compartment; a variant of liposomes are micelles, consisting of lipid spheres
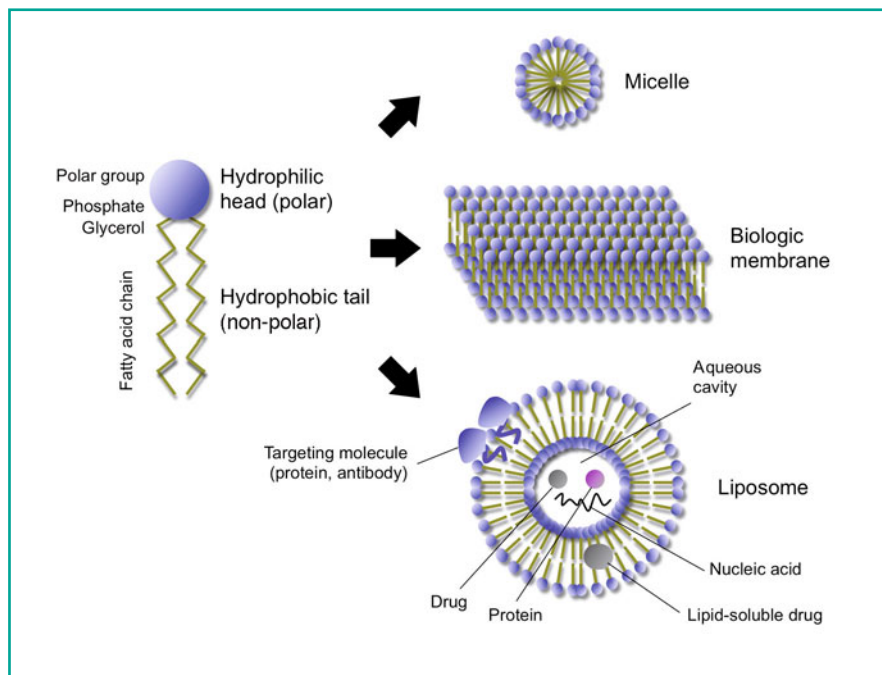
Fig. 3.2 Supramolecular organization of phospholipids. The picture shows assembly of phospholipids (*left*) into micelles (*upper right*), biological membranes (*middle right*), or liposomes (*lower right*). Macromolecules, such as chemical drugs, nucleic acids, or proteins, can be transported in the liposome core or, if hydrophobic, within the lipid bilayer. Proteins mediating specific cellular targeting can be embedded on the liposome surface

lacking the inner aqueous compartment (Figure 3.2). Liposomes were originally developed in the 1960s and are now extensively used to convey different molecules in a variety of applications, ranging from chemotherapy (for example, transport of antiblastic or antifungal drugs to prevent unspecific toxicity), diagnostic imaging, and cosmetic applications. The first developed liposomes were based on the same phospholipids forming biological membranes, having a polar head and a lipophilic tail formed by fatty acids. These molecules have amphipathic (or amphiphilic) characteristics: once dispersed in an aqueous solution, they tend to spontaneously assemble into a bilayer, first forming a sheet and then closing up into a vesicular structure with a central aqueous core. When liposome formation occurs in a solution containing a drug or a nucleic acid, these are eventually found in the aqueous core of the liposome and can thus be transported by it. Once in contact with a cell, liposomes can directly fuse with the plasma membrane, thus liberating their content into the cytosol, or, more frequently, be actively endocytosed.

The biological properties of liposomes derive from those of the amphiphilic lipids they are composed of; according to the characteristics of the polar head groups, they can be classified into anionic, cationic, zwitterionic, and non-ionic liposomes. Conventional liposomes, non-ionic or neutral, interact inefficiently with a large polyanion such as DNA. In contrast, DNA binding is much more effective using cationic lipids. Figure 3.3 shows the

structure of two of the most used cationic lipids, DOTMA (the first one to be utilized, in 1987) and DOTAP. Both consist of two fatty acyl chains joined to a positively charged propylammonium group through an etheric and esteric bond, respectively. The positive moiety of the cationic lipid binds negatively charged DNA very efficiently and induces its condensation. Furthermore, the DNA/lipid complex maintains a positive net charge and is thus capable of electrostatically interacting with the negatively charged cell surface. In addition, the complex displays fusogenic properties, thus promoting fusion of the lipo-some with the cell or the endosome membrane and favoring release of DNA into the cytosol. The DNA–lipid complex is named *lipoplex*.

Over the last several years, a variety of other cationic lipids have been produced, dif-fering in hydrophobic moiety, number of positive charges, or presence of other chemical groups mediating interaction between the polar and hydrophobic portions. One of the developed lipids is DC-Chol (Figure 3.3), in which the hydrophobic moiety consists of a sterol skeleton. This lipid is currently used in different gene therapy applications including a few for cystic fibrosis. Some of the cationic lipids, once mixed to DNA, form micellar rather than vesicular structures, and are thus more efficient at inducing DNA condensation.

In general, the maximum efficiency of gene transfer is achieved when a cationic lipid is mixed with cholesterol or with a zwitterionic lipid, which displays an overall neutral charge despite carrying both negative and positive charges on different atoms; an extensively used zwitterionic lipid is DOPE (Figure 3.3). These co-lipids, once taking part in the formation
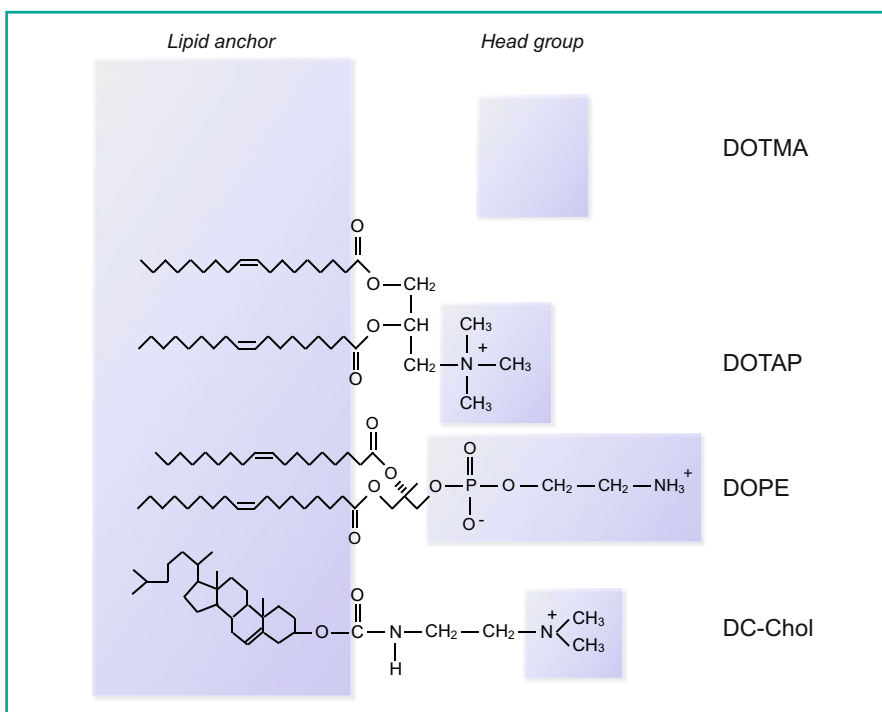

Fig. 3.3 Cationic lipids. The chemical structure of four lipids commonly used for gene transfer is shown

of the cationic lipid-DNA complex, facilitate fusion or destabilization of the cellular membranes, thus favoring transfection. As a matter of fact, the mixture DOTMA/DOPE is currently one of the most successful commercial lipid formulations for gene transfer into cultured cells.

A lipoplex is typically obtained in the laboratory by mixing a cationic lipid, a co-lipid, and DNA in appropriate concentrations, followed by sonication of the mixture in order to reduce the size of the complexes that are spontaneously formed. The final structures of the different lipoplexes are still not completely defined and probably vary from preparation to preparation according to the chemical composition and the molar ratio of the two or three lipoplex components. Molecular aggregation can lead to the formation of simple structures, consisting in DNA covered by a cationic lipid bilayer, or in aggregates of cationic liposomes surrounding DNA such as a pearl chain, or to the generation of more complex arrangements, such as multilamellar structures or hexagonal structures formed by DNA covered by monolayers of cationic lipids assembled in a bi-dimensional hexagonal lattice (Figure 3.4). Accordingly, the final lipoplex preparations can have very different sizes, with a diameter varying from a few nanometers (unilamellar liposomes) to several hundred nanometers (multilamellar and complex lipoplexes).

From the gene therapy point of view, the ideal lipoplex should provide protection from nuclease degradation of DNA, mediate very efficient cellular internalization, and exert minimal cell toxicity. At the same time, it should also display a neutral or negatively charged surface in order to escape unspecific interaction with blood components. None of the currently available lipoplex formulations completely satisfies all these requisites. As a matter of fact, positively charged lipoplexes induce DNA condensation more efficiently, leading to higher *in vitro* and *in vivo* levels of transfection. For reasons still not completely explained, lipoplexes having a size >200 nm are more efficient than smaller ones (50–100 nm).

Despite the extensive use of cationic liposomes for *in vivo* and *in vitro* gene transfer, the mechanism by which they release their DNA into the cells is still unclear, and probably varies according to the chemical and structural properties of the various lipoplexes. Most studies indicate that, following interaction with the cell membrane, the lipoplex is internal-
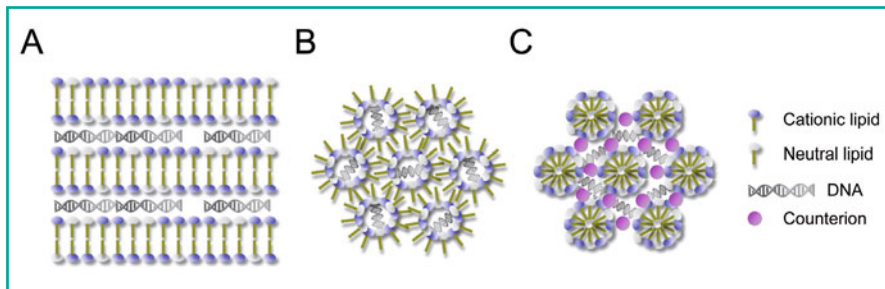
Fig. 3.4 Lipoplexes. Schematic representation of structures formed upon the interaction of nucleic acids with polar lipids. **A** Lamellar structure, where DNA molecules are sandwiched between lipid bilayers formed by an alternation of cationic and neutral lipids. **B** Inverted hexagonal structure, where DNA molecules are coated with a lipid monolayer arranged on a hexagonal lattice. **C** Intercalated hexagonal structure, where DNA molecules are interspersed with lipid micelles arranged on a hexagonal lattice

3

ized through a pathway of clathrin-mediated endocytosis rather than directly fusing at the plasma membrane level. In this respect, and notwithstanding the continuous progress in the generation of more efficient lipoplexes, it should be remarked that the number of DNA molecules effectively reaching the nucleus only represents a tiny fraction of those having entered the transfected cells (in the order of 1 out of 1×10^4–10^5). In fact, most of the lipoplexes remain trapped in the endosomes and are eventually degraded by lysosomes. Exit from the endosomes before reaching the lysosomes is favored by the inclusion, in the lipoplexes, of zwitterionic lipids such as DOPE, since the progressive acidification of the endosomal compartment favors the membrane-destabilizing properties of this type of lipid.

3.4.2
Cationic Polymers

Another interesting class of molecules with the property of binding DNA and favoring its transfer into the cells consists of cationic polymers. These molecules include poly-(L-lysine), poly-(L-ornithine), linear or branched polyethylenimine (PEI), diethyl-aminoethyldextran (DEAE-D), poly-(amido amine) dendrimers, and poly-[2-(dimethyl-amino)-ethyl methacrylate (poly-(DMAEMA)). These polymers, which can have linear, branched, or dendrimeric structures (Figure 3.5), usually carry a protonable amine group, which, thanks to its positive charge, binds DNA and induces its condensation. Similar to lipoplexes, the DNA/polymer complexes, named *polyplexes*, enter the cells through an active endocytosis process. Once in the endocytic vesicles, the positively charged amine groups of the polymers are believed to exert a so-called "proton sponge" effect, according to which the low endosomal pH determines entry, into the endocytic vesicles, of chloride ions, followed by osmotic rupture of endosomes and release of DNA into the cytosol.

One of the major problems related to the use of polyplexes for gene transfer is their toxicity, due to the positive charge of the polymers and the large size of the polymer–DNA complexes that are eventually formed. For this reason, several laboratories are currently investigating the possibility of improving polymer architecture and biophysical properties. One class of very interesting polymers in this respect is the amphipathic block co-polymers, consisting of alternating blocks of simple hydrophobic homopolymers and simple hydrophilic homopolymers. Such block co-polymers, which are significantly less toxic than conventional cationic polymers, display the property of interacting, at the same time, with DNA through their hydrophilic moieties and with the plasma or endosomal membranes through their hydrophobic moieties.

Another interesting class of polymers showing low toxicity and high biocompatibility are biodegradable polymers. An example of such polymers is the degradable polyester poly[α-(4-aminobutyl)-L-glycolic acid] (PAGA), a derivative of poly-[L-lysine], which binds DNA and subsequently releases it once the polymer is degraded.

Finally, an additional family of polymers with attractive properties for gene transfer includes the so-called "intelligent polymers". These polymers can undergo ample and often discontinuous variations in their chemical and physical characteristics in response to environmental changes, such as pH, temperature, ionic force, or presence of electric or magnetic fields. The polymer modification consists in variations of the size, three-dimen-
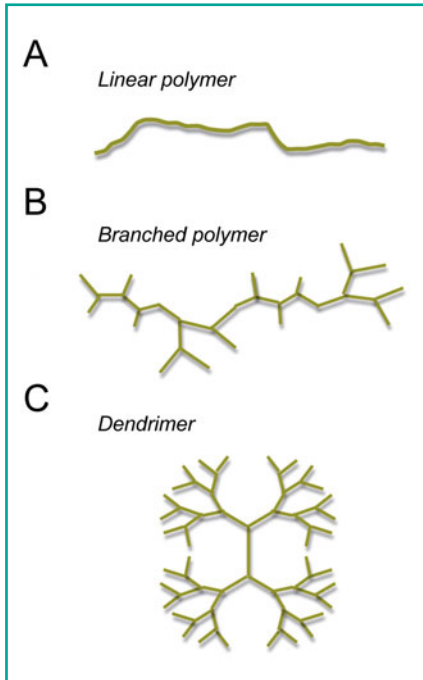
Fig. 3.5 Cationic polymers. **A** Linear polymer; **B** branched polymer; **C** dendrimer

sional structure, or reactivity towards other molecules. Typical examples of intelligent polymers consist of block co-polymers formed by methyl methacrylate (MMA) and 2-(dimethylamino)-ethyl methacrylate (DMAEMA). MMA is hydrophobic, while DMAEMA is hydrophilic; the block co-polymer is thus formed by alternating hydrophobic and hydrophilic blocks, with a prevailing overall hydrophilic characteristic. DMAEMA, however, becomes more hydrophilic at lower pH and more hydrophobic at higher pH, thus determining precipitation of the co-polymer. Varying the ratio between the two co-polymers improves transfection efficiency and minimizes cytotoxicity.

Intelligent polymers sensitive to temperature behave according to a similar principle. For example, poly-(N-isopropylacrylamide) (NIPA) is soluble in water below 32°C and becomes insoluble above this temperature. In principle, polymers containing this molecule should permit *in vitro* assembly of polyplexes at low temperature in the laboratory; once injected *in vivo*, precipitation of the co-polymer induces the formation of a gel, which should allow progressive release of DNA over prolonged periods of time.

Finally, a peculiar class of polymers is dendrimers (from the Greek "dendron", tree). These consist of a central molecule, acting as a root for the progressive synthesis of a vast number of branches, which are structured in an ordered and symmetric manner (Figure 3.5C). Similar to cationic polymers, dendrimers have the capacity to efficiently complex with DNA and mediate its cellular internalization by endocytosis, followed by endosomal release by osmotic swelling. Dendrimers can be potentially used for the delivery of long stretches of DNA (several tens of megabases) and display, in a few experimental systems, higher efficiency of gene transfer than linear polymers.

3.4.3
Proteins

The efficiency of viruses to convey their nucleic acids into the cells is due to the presence, within the viral particles, of specific proteins. These mediate a series of essential functions, such as condensation of viral nucleic acids, protection against extracellular nucleases, binding to cell surface receptors, fusion of viral envelope with cell membranes or endosome disruption, and, finally, transport of viral DNA or RNA to the nucleus. Some of these processes can be mimicked by using specific proteins or protein domains in the context of non-viral gene delivery.

Some basic proteins, such as the polycationic polypeptide protamine or histones, bind negatively charged DNA with high affinity, promoting its condensation and preventing its degradation. Once in the extracellular environment, these molecules are also able to bind heparan sulfate proteoglycans (HSPGs), a family of negatively charged glycosylated proteins expressed on the cell surface and released into the extracellular matrix. HSPGs exposed on the cell surface continuously undergo a process of endocytosis; thus, molecules interacting with HSPGs outside the cells are also internalized within endocytic vesicles. Since the efficiency of this process is relatively modest and the endocytosed molecules still need to cross the endosomal membrane, the simultaneous utilization of liposomes or cationic lipids together with basic proteins favors DNA transfection.

A specific protein able to favor the internalization of other covalently linked macromolecules is the Tat protein encoded by HIV-1. This factor is a powerful transactivator of viral gene expression and is thus essential for viral replication and infectivity. For completely unknown reasons, Tat is released by HIV-1-infected cells through a non-canonical, Golgi-independent pathway of secretion and, once in the extracellular milieu, is taken up by cells upon the interaction of a highly basic, 9-amino acid domain of the protein with cell surface HSPGs, which mediate its endocytosis through the caveolar pathway. When this small amino acid domain of Tat is fused to heterologous proteins, synthetic nanoparticles, liposomes, or small nucleic acids (for example, siRNAs), Tat mediates their cellular internalization. Since this process occurs through caveolar endocytosis, a significant part of the Tat-fusion cargo escapes lysosomal degradation, finds access to the cytosol, and, from here, is transported to the nucleus.

As discussed above, internalization of lipoplexes and polyplexes occurs through an active endocytic process, mainly mediated by clathrin-coated vesicles. It is thus possible to associate, to the lipoplexes or polyplexes, proteins of various derivation, able to recognize specific cellular receptors involved in endocytosis. The ultimate purpose of this approach is twofold, namely on one hand to increase the overall efficiency of the gene transfer process while, on the other hand, targeting transfection towards specific cell types or tissues expressing the receptors of interest. The proteins that have been used for this purpose include various lectins (proteins that are very diffuse in nature, having the capacity to bind the glycidic moiety of various glycoproteins and glycolipids); the asialoglycoproteins (i.e., glycoproteins devoid of sialic acid, which specifically bind a receptor expressed on the surface of hepatocytes; this receptor, named ASGP-R, recognizes glycoproteins carrying a galactose at their extremity and removes them from the circulation by endocytosis followed by lysosomal degradation); integrin ligands (for example, peptides carrying the amino acid sequence

Arg-Gly-Asp, RGD); peptides derived from apolipoprotein E (which bind the low-density lipoprotein receptor, LDL-R, expressed by hepatocytes); and transferrin (which binds the transferrin receptor, expressed by a variety of cell types and overexpressed in tumor cells). Finally, lipoplexes and polyplexes can also be targeted towards specific cellular receptors through their association with monoclonal antibodies or single-chain antibodies (scFvs; cf. section on 'Antibodies and Intracellular Antibodies'). Both scFvs and peptides binding a specific receptor of interest can be selected thanks to the phage display technology.

Collectively, however, it should be remembered that binding a specific cellular receptor does not necessarily translate into a parallel increase of transfection efficiency. Indeed, this parameter strictly depends not only on the efficiency of internalization but also, and probably mostly, on the capacity of the nucleic acids to exit the endosomes before they are degraded in the lysosomes. This is also proven by the observation that, in cell culture, treatment with compounds that raise the pH of acidic vesicles, such as the anti-malaria drug chloroquine or the macrolide antibiotic bafilomycin A1, a selective inhibitor of the vacuolar-type ATPase (V-ATPase), significantly increase the efficiency of transfection. A few natural proteins and peptides possess natural endosomolytic activity, and can thus be used to form lipoplexes or polyplexes with improved transfection efficiency. These include peptides derived from the HA2 subunit of the hemagglutinin (HA) protein of the influenza virus (which mediates a low-pH-dependent fusion reaction between the viral envelope and the endosomal membrane following cellular uptake of the virus particles by receptor-mediated endocytosis), the envelope of the Sendai virus (a paramyxovirus also called hemoagglutinating virus of Japan, HVJ, which is a powerful inducer of membrane fusion), and a synthetic amphipathic peptide, sensitive to pH, called GALA (named as such because it contains repeat units of glutamic acid-alanine-leucine-alanine). Alternatively, some gene therapy applications have exploited the natural endosomolytic properties of the capsid proteins of adenovirus, by including whole adenoviral virions, inactivated by UV radiation, in the lipoplex.

Finally, other peptides can facilitate the subsequent passage, namely transport of DNA from the cytosol to the nucleus. In particular, peptides carrying a NLS are recognized by the cellular proteins of the importin/karyopherin family, which, in the cells, mediate transport of NLS-containing sequences through the nuclear pores into the nucleus.

3.4.4
Chemical Methods for Gene Transfer: Pros and Cons

Although the chemical methods of gene transfer offer important advantages over viral gene delivery in terms of relative safety and simplicity, and notwithstanding the large amount of research carried out on these methods over the last 20 years, their overall efficiency is still unsatisfactory. Most of the DNA entering the cells remains trapped in the endosomes and is eventually destroyed. In addition, plasmid DNA reaching the nucleus is unprotected from degradation by cellular nucleases and, since it does not integrate into the cellular genome, is progressively lost, thus allowing transgene expression for periods usually shorter than a couple of weeks.

When lipoplexes or polyplexes are administered systemically, additional problems ensue due to their rapid update and elimination by the cells of the reticuloendothelial sys-

tem (RES) in spleen, liver, and lymph nodes. With analogy to several other conventional drugs, the most widely used system to avoid unspecific interactions is to mask the positive charges of lipoplexes and polyplexes with neutral hydrophilic polymers, such as polyethylene glycol (PEG). On one hand, PEGylation prevents aggregation, thus favoring the formation of smaller complexes (which is usually an advantage for gene transfer), while, on the other hand, it blocks unspecific interaction of the complexes with serum proteins or other extracellular components, thus increasing their persistence in the blood stream. PEGylation, however, in several conditions also leads to a decreased interaction of the complexes with the cells, thus diminishing their biological activity.

Finally it should be noted that, although less immunogenic than viral vectors, lipoplexes and polyplexes are internalized by macrophages and other APCs and can thus elicit an immune response against both the gene transfer molecules and the delivered transgenes. Additionally, cationic liposomes can be toxic, since they rapidly induce the production of pro-inflammatory cytokines, such as TNF-α, IL-6, IL-12, and IFN-γ. In the case of plasmid DNA delivery, part of this response is also due to the presence, within the plasmid, of bacterial, non-methylated CpG sequences, which usually represent a powerful stimulus for immune response through the activation of Toll-like receptor-9 (TLR-9).

In light of these considerations, it is not surprising that most (~70%) of the gene therapy clinical trials, and in particular those aimed at delivering coding genes, exploit viral vectors rather than non-viral methods for gene transfer.

3.5
Viral Vectors

The most efficient system to deliver a gene into a cell is to exploit the properties of vectors derived from the viruses that infect animal cells. In their replicative cycle, viruses make use of very efficient mechanisms to internalize their own genome into the target cells, which have evolved over million years. In the most simplistic view, a viral particle is a tiny object composed of a nucleic acid and a few proteins that impede its degradation in the extracellular environment and mediate its internalization into the target cells. In general terms, the process of viral replication is sustained by the interaction of several proteins of viral or cellular origin with their respective, specific targets on the viral genome. The proteins are said to act *in trans* and the targets *in cis*. Examples of such *cis/trans* interactions are those regulating activation of promoters positioned within the viral genome, transport of the viral nucleic acids from the nucleus to the cytoplasm, or packaging of viral genomes inside the virions.

A viral genome modified in order to accommodate an exogenous sequence of interest (the therapeutic gene in the case of gene therapy) is called a *vector*. The principles according to which the different viral vectors are obtained starting from the parental genomes are common to all systems. They consist in the: (i) removal, from the viral genome, of most genes coding for viral proteins and, in particular, of those that are potentially pathogenic; (ii) maintenance of the *cis*-acting sequences of the viral genomes required for viral replication; in particular, those determining inclusion of the genomes within the viral particles (packaging signal, ψ); (iii) expression of the viral genes required for viral replication with-

in the virus-producing cells (called *packaging cells*) from genes encoded by transiently transfected plasmids, or expressed in the context of a helper virus simultaneously infecting the packaging cells, or directly contained inside the packaging cell genome thanks to previous engineering of these cells.

Five classes of vectors are currently in an advanced stage of clinical experimentation for human gene therapy. These include viruses derived from the *Retroviridae* family (gammaretroviruses and lentiviruses), adenoviruses, AAVs, and herpesviruses. Other viruses, such as vacciniaviruses, the viruses belonging to the spumavirus and alpharetrovirus genera of the *Retroviridae* family, and RNA viruses such as the Semliki Forest Virus are also considered potentially attractive for therapeutic gene transfer, however their use is limited to vaccination (for vacciniaviruses) or they still need vast preclinical development and validation.

The modalities of production and the characteristics of the five main classes of viral vectors are detailed in the following sections, along with the main characteristics of each of the parental viruses.

3.5.1
Vectors Based on Gammaretroviruses

The vast majority of the clinical trials conducted in the 1990s took advantage of the properties of viral vectors based on gammaretroviruses. Among the properties of these viruses are their relative genetic simplicity, their efficiency in infecting a vast series of different cell types, and their peculiar ability to integrate their genetic information into the genome of the infected cells, a characteristic leading to permanent genetic modification.

3.5.1.1
Molecular Biology and Replicative Cycle of Retroviruses

The *Retroviridae* family includes a vast series of enveloped viruses with a positive-strand RNA, having a common genetic structure and replicative cycle. A peculiar characteristic of all members of this family is the presence of an enzyme, reverse transcriptase (RT), which copies the viral RNA into a double-stranded cDNA form, which eventually integrates into the infected cell genome. The integrated DNA form of the viral genome is named *provirus*.

Classification of Retroviruses

The different members of the *Retroviridae* family have been variously classified according to their morphology, natural animal host, type of disease caused, and tropism for different cell types. These characteristics are reported in Table 3.2, along with the indication of a few representative family members. In recent years, the taxonomic classification has changed to consider both the previous parameters and the more recently acquired information on the genetic organization of the viruses. The *Retroviridae* family now consists of 2 subfamilies (*Orthoretrovirinae* and *Spumaretrovirinae*) and 7 genera: alpharetrovirus

3

Table 3.2 Criteria for retrovirus classification

Parameter	Characteristics		
Electron microscopy morphology	*Classification*	*Characteristics*	*Prototype*
	Type A virions	Central nucleocapsid with translucid appearance with one or two concentric layers, without envelope	Immature forms of type B and type D viruses and endogenous retroviruses (intracisternal A particles, IAP)
	Type B virions	Nucleocapsid in eccentric position with prominent surface protrusions	Mouse mammary tumor virus (MMTV)
	Type C virions	Central nucleocapsid, with almost invisible protrusions	Most murine and avian sarcoma and leukemia viruses; for example, Moloney murine leukemia virus (Mo-MLV) and avian sarcoma/leukosis virus (ASLV)
	Type D virions	Oval nucleocapsid with small surface protrusions	Mason-Pfizer monkey virus (MPMV)
	Other types	Similar to type C viruses with different protrusions	Bovine leukemia virus (BLV), human T-cell leukemia virus (HTLV)
		Similar to type C viruses carrying a nucleocapsid with a truncated cone shape	Lentivirus
		Virions with prominent surface protrusions	Spumavirus
Genome organization	Simple retroviruses (*gal*, *pol*, and *env* genes) and complex retroviruses (*gal*, *pol*, *env* plus several accessory genes)		
Host	Murine, avian, feline, bovine, human, etc. retroviruses		
Species tropism	Ecotropic, xenotropic, amphotropic retroviruses		
Pathogenicity	Oncovirus, lentivirus, spumavirus		
Disease caused	Leukemia, sarcoma, myeloblastosis, erythroblastosis, immunodeficiency, anemia, encephalitis, etc.		

(whose prototype species is the avian leukosis virus, ALV), betaretrovirus (mouse mammary tumor virus, MMTV), gammaretrovirus (murine leukemia virus, MLV), deltaretrovirus (bovine leukemia virus, BLV), epsilonretrovirus (Walleye dermal sarcoma virus, WDSV), lentivirus (human immunodeficiency virus type 1, HIV-1), and spumavirus (human foamy virus, HFV). Table 3.3 reports this classification along with the older subgroup denominations and some of the most representative viruses in each genus.

A classification that is sometimes useful for operational purposes divides the *Retroviridae* family into three major subgroups: the oncoretroviruses (which include the

Table 3.3 Taxonomy of the *Retroviridae* family

Subfamily	Genus	Former classifications	Main species	Prototype viruses
Orthoretrovirinae	Alpharetrovirus	Avian type C retroviruses; Avian sarcoma/ leukosis viruses (ASLV)	Avian leukosis virus Rous sarcoma virus	ALV RSV
	Betaretrovirus	Mammalian type B retroviruses; type D retroviruses	Mouse mammary tumor virus Mason-Pfizer monkey virus	MMTV MPMV
	Gammaretrovirus	Mammalian type C retroviruses	Murine leukemia virus	Abelson-MLV, Friend-MLV, Moloney-MLV
			Feline leukemia virus	FeLV
			Gibbon ape leukemia virus	GaLV
			Harvey murine sarcoma virus	Ha-MSV
			Moloney murine sarcoma virus	Mo-MSV
			Simian sarcoma virus	SSV
			Reticuloendotheliosis virus	REV-A, REV-T
	Deltaretrovirus	BLV-HLTV group retroviruses	Bovine leukemia virus	BLV
			Primate T-lymphotropic viruses (human and simian)	HTLV-1, STLV-1, HTLV-2, STLV-2, STLV-3
	Epsilonretrovirus	Fish retroviruses	Walleye dermal sarcoma virus	WDSV
	Lentivirus		Bovine immunodeficiency virus	BIV
			Equine infectious anemia virus	EIAV
			Feline immunodeficiency virus	FIV-O, FIV-P
			Caprine arthritis encephalitis virus	CAEV

3

Table 3.3 (*Continued*)

Subfamily	Genus	Former classifications	Main species	Prototype viruses
			Visna/Maedi virus	VISNA
			Human immunodeficiency virus 1 and 2	HIV-1, HIV-2
			Simian immunodeficiency virus	SIVagm (155), SIVcpz, SIVmac
Spumaretrovirinae	Spumavirus		Simian foamy virus	SFVmac (SFV-1 and SFV-2), SFVagm (SFV-3), SFVcpz, and SFVcpz(hu)
			Bovine foamy virus	BFV
			Equine foamy virus	EFV
			Feline foamy virus	FFV
			Human foamy virus	HFV or HSRV

first five genera), the lentiviruses (*lenti-*: Latin for "slow", since these viruses cause diseases characterized by a long incubation period and slow evolution), and the spumaviruses (*spuma-*: Latin for "foam", from the cytopathic effect induced in monkey kidney cells, which is characterized by the formation of large vacuoles). Given their complex genetic organization, viruses belonging to the latter groups are also collectively named as "complex retroviruses" (see also Figure 3.7).

Genome Organization

The genomes of prototypic members of the *Retroviridae* family, such as the gammaretrovirus Moloney-murine leukemia virus (Mo-MLV), are 9–11 kb long and consist of 3 essential genes (*gag*, *pol*, and *env*) flanked, in their integrated, proviral DNA forms, by two identical sequences of 400–700 bp at the 3' and 5' extremities, named long terminal repeats, LTRs (Figure 3.6). Each LTR consists of three regions: U3, R, and U5. The 5' LTR U3 region contains a promoter driving expression of all viral transcripts. Transcription starts in correspondence with the first nucleotide of the 5' R region and proceeds for the entire length of the genome; the R region contains a polyadenylation signal, which, at the 3' end, drives cleavage and polyadenylation of the mRNA in correspondence with the 3' R-U5 boundary; the ensuing mRNA corresponds to the genome that is eventually packaged into the virions. Thus, in contrast to the provirus, which is flanked by two complete LTRs, the viral RNA genome initiates with the R-U5 sequence at its 5' extrem-
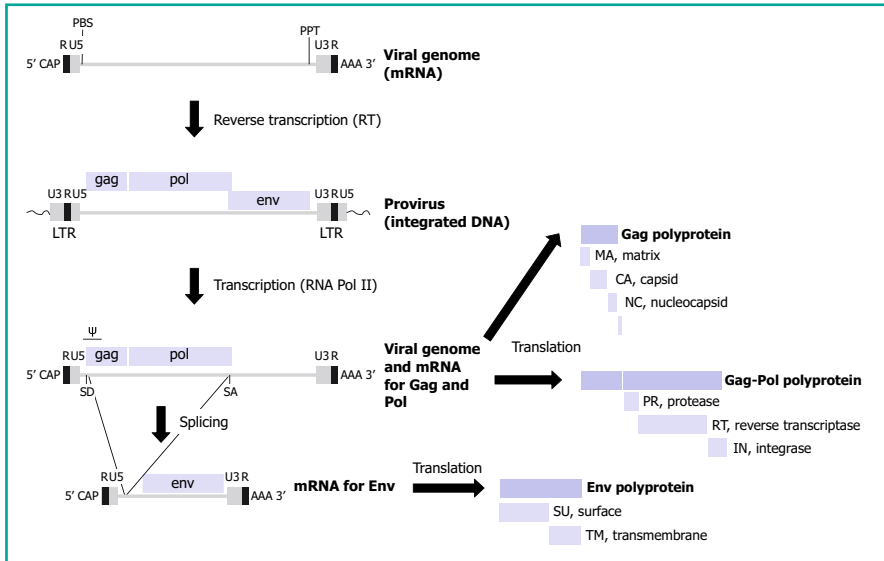
Fig. 3.6 Retrovirus genome and proteins. Structure of the viral genome mRNA, proviral DNA, and major viral transcripts (full-length genomic mRNA and single-spliced transcript) are shown on the *left side*, as indicated from top to bottom. The proteins obtained by transcription of the viral mRNAs are indicated on the *right side*

ity and ends with the U3-R sequence at its 3' extremity. A complete LTR sequence is only generated during reverse transcription.

Three essential genes – *gag*, *pol*, and *env* – are present between the two LTRs (Figure 3.6). The *gag* gene codes for proteins associated with the viral genome and essential for packaging; these include the matrix (MA), capsid (CA), and nucleocapsid (NC) proteins. The *pol* gene codes for the three essential viral enzymes that characterize all members of the retrovirus family: reverse transcriptase (RT), protease (PR), and integrase (IN). RT is responsible for the process of reverse transcription, converting the RNA genome into its cDNA form. PR catalyzes cleavage of the polyproteins that are generated by translation of the viral genes into the individual viral proteins, as detailed in the discussion of the viral replication cycle. IN catalyzes integration of the viral cDNA into the host cell genome to generate the provirus. The *env* gene codes for two proteins that are displayed first on infected cell surfaces and later, after budding, on the viral envelope. These are the TM (trans-membrane) protein, which positions itself across the membrane, and the SU (surface) protein, which is anchored onto the TM outside the membrane and mediates recognition of cellular receptors.

The viral RNA transcript, besides constituting the retroviral genome that is packaged inside the virions, also acts as the mRNA from which all the viral proteins are translated (Figure. 3.6). In particular, the full-length primary transcript codes for two long polypeptides corresponding to the Gag and Gag-Pol polyprotein precursors; the same transcript undergoes splicing to generate a processed mRNA coding for the Env polyprotein. In all cases, functional retroviral proteins are eventually generated by cleavage of these polyproteins into the final individual peptides.
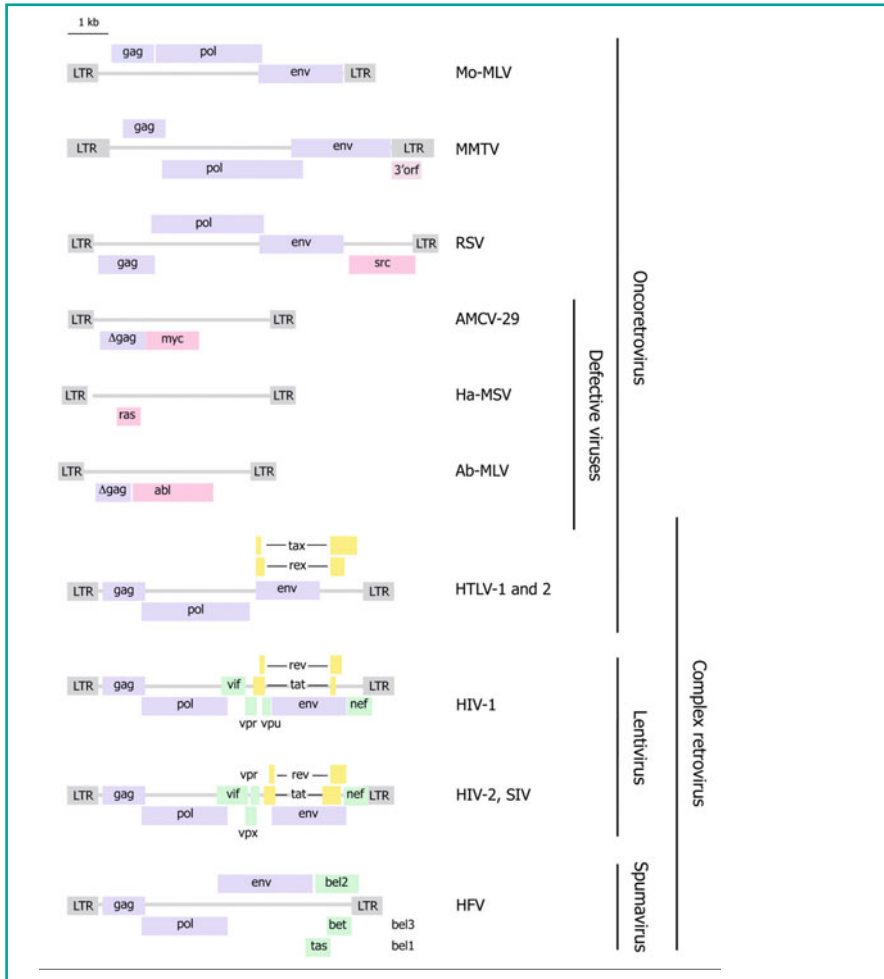
Fig. 3.7 Retroviral genomes. The genetic organization of some common retroviral genomes is shown, along with the indication of the common retroviral genes (*lilac*) and of the accessory genes proper of the individual groups (*other colors*)

Presence of the LTRs and of the *gag*, *pol*, and *env* genes is a hallmark of all retroviruses, since integrity of these genetic elements is essential for viral replication. Some family members, however, contain additional genetic information or display variations of this general genetic structure. The genetic organization of a few prototypic retroviruses is shown in Figure 3.7. In particular:

(i) members of the group of complex retroviruses (including HTLV-1, all lentiviruses, and spumaviruses), in addition to *gag*, *pol*, and *env*, also contain an extra series of genes, encoded from the 3' half of the viral genome. These genes, named "accessory genes" are indeed fundamental for these retroviruses to efficiently infect the respective target cells. For example, HIV-1 contains 6 accessory genes (*tat*, *rev*, *nef*, *vpr*, *vpu*, and *vpr*), which are essential for different

Table 3.4 Examples of retroviruses carrying viral oncogenes (*v-onc*)

Parental/helper virus	Retrovirus	Acronym	*v-onc*
	Rous sarcoma virus	RSV	*src*
Avian leukosis virus (ALV)	Avian myeloblastosis virus	AMV	*myb*
	Avian erythroblastosis virus	AEV	*erbA, B*
	Avian myelocytomatosis virus 29	AMCV-29	*myc*
	Y73 sarcoma virus	Y73SV	*yes*
	Avian sarcoma virus 17	ASV-17	*jun*
Moloney-Murine leukemia virus (Mo-MLV)	Abelson murine leukemia virus	Ab-MLV	*abl*
	Harvey murine sarcoma virus	Ha-MSV	*ras*
	Moloney murine sarcoma virus	Mo-MSV	*mos*
	Finkel-Biskis-Jinkins murine sarcoma virus	FBJ-MSV	*fos*
Feline leukemia virus (FeLV)	Snyder-Theilen feline sarcoma virus	ST-FeSV	*fes*
	Gardner-Arnstein feline sarcoma virus	GA-FeSV	
	Susan McDonough feline sarcoma virus	SM-FeSV	*fms*
	Hardy-Zuckerman 4 feline sarcoma virus	HZ4-FeSV	*kit*
Simian sarcoma virus (SSV)	Woolly monkey sarcoma virus	WMSV	*sis*

steps of the viral life cycle, including transcription (*tat*), transport of viral mRNAs outside of the nucleus (*rev*), cell cycle regulation (*vpr*), and modulation of virion infectivity (*vif*).

(ii) Some alpharetroviruses and gammaretroviruses contain a peculiar gene, derived from the host cell genome. This is an oncogene and represents the activated form of a cellular gene that is normally devoted to the control of cell cycle progression or cell differentiation. The viral versions (v-*onc*) of these normal cellular genes (also named proto-oncogenes, c-*onc*) are devoid of introns and thus similar to the cellular gene cDNAs, and are constitutively active, since they contain mutations that activate the encoded proteins or are continuously transcribed at high levels. Some of the v-*onc*-carrying retroviruses are listed in Table 3.4. Of note, several of the cellular proto-oncogenes have been discovered thanks to the presence, in one of the retroviruses, of their activated counterparts. The viruses containing an oncogene have the ability to transform the cells they infect (that is, to induce an oncogenic behavior) very efficiently and with a very rapid kinetics, since the constitutively active oncogene drives the cells into proliferation and negatively regulates their terminal differentiation. The Rous sarcoma virus (RSV), an alpharetrovirus, is the only retrovirus in which the v-onc is additional to intact *gag*, *pol*, and *env* genes. In all other cases, the oncogene-carrying retroviruses show more or less broad deletions in their genome – see, for example, Figure 3.7 for the avian myelocytomatosis virus 29 (AMCV-29), an alpharetrovirus, or the Abelson murine leukemia virus (Ab-MLV) and the Harvey murine sarcoma virus (Ha-MSV), two gammaretroviruses. Since the presence and integrity of the *gag*, *pol*, and *env* genes are essential, these viruses are defective for replication and can only be propagated if the cell they infect is also superinfected with a replication-competent virus of the same family. The viruses allowing replication of the defective retrovirus-

es are named *helper* viruses. A cell simultaneously infected with a helper virus and a defective virus produces virions with the characteristics of the helper virus but containing genomes corresponding to either the defective or the helper virus. For example, in the gammaretrovirus genus, at least eleven defective retroviruses causing fibrosarcomas in cats have been insolated, possessing seven different oncogenes. All these viruses (named feline sarcoma viruses, FeSVs) have arisen as recombinants from the feline leukemia viruses (FeLV), in which a vast part of the genome was replaced by cellular oncogenes. The FeSVs are defective for replication and their propagation can only occur if the animal is superinfected with a replication-competent FeLV.

The retroviruses containing an oncogene induce tumors efficiently and rapidly after infection. However, the replication-competent retroviruses (RCRs) not carrying an oncogene – including FeLV itself, Mo-MLV, the mouse mammary tumor virus (MMTV), the ALV, and the human T-lymphotropic virus type-1 (HTLV-1) – can also induce tumors in their natural hosts, albeit with very different kinetics and mechanisms. In these cases, transformation requires several weeks or months (for example, for Mo-MLV) or even decades (for HTLV-1). In the case of Mo-MLV, transformation is due to insertional mutagenesis, that is to activation of a cellular proto-oncogene or inactivation of a tumor suppressor gene due to retroviral integration within, or in close proximity to, these genes. This will be further discussed in the section on 'Gene Therapy of Hematopoietic Stem Cells', since insertional mutagenesis was the cause of the occurrence of leukemia in a few patients treated by gene therapy using gammaretroviral vectors. In the case of HTLV-1, cellular transformation is subsequent to the activity of the virus accessory genes, in particular, of the *tax* gene, which interfere with multiple cellular functions and facilitate mutation of cellular proto-oncogenes.
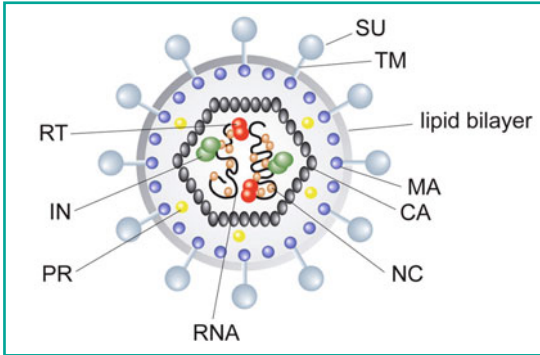
Structure of Virions

Retroviral particles have a diameter of 80–100 nm and consist of an envelope, composed of the host cell plasma membrane with the addition of glycosylated viral TM and SU proteins, linked by disulfide bonds. Inside the virions, the viral proteins MA, CA, and NC associate with two identical copies of the viral mRNA genome to form the nucleocapsid, along with the viral enzymes RT, PR, and IN (Figure 3.8). By electron microscopy, the virions appear with an outer ring, corresponding to the envelope, surrounding an electron-dense core, corresponding to the nucleocapsid; the TM and SU proteins protrude as spikes from the envelope. Variations in the morphological structure of the virions represent criteria for classification of the different members of the family.
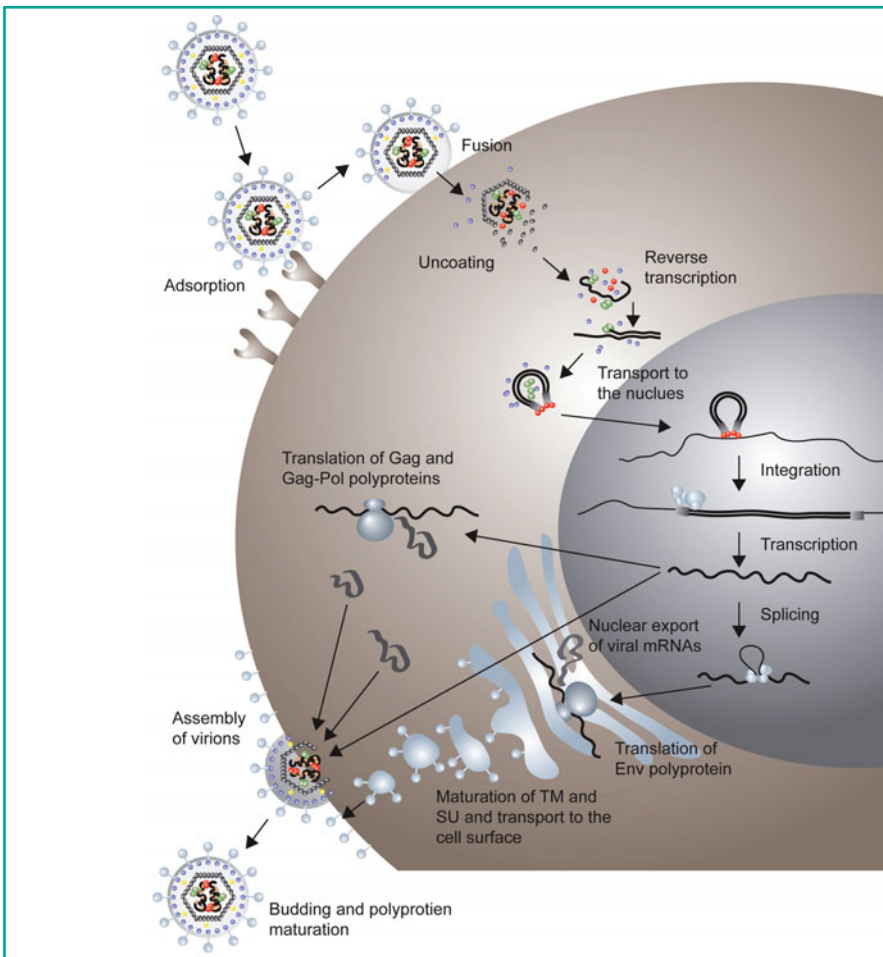
Replicative Cycle

The replicative cycle of retroviruses can be subdivided into a series of subsequent steps (Figure 3.9).

Adsorption. Binding of virions to the cell surface is mediated by the interaction of SU with a cellular plasma membrane protein, acting as a receptor. The different members of

Fig. 3.8 Retroviral virion. The structure of a prototype retroviral virion is shown, with the indication of the proteins contained inside and on the surface. *SU*: surface; *TM*: transmembrane; *MA*: matrix; *CA*: capsid; *NC*: nucleocapsid; *PR*: protease; *IN*: integrase; *RT*: reverse transcriptase

Fig. 3.9 Retroviral replication cycle. The main steps of the retroviral replication cycle are shown. The viral genome RNA is shown in black and the cDNA formed upon reverse transcription in red. The LTR elements are boxed. See text for description

3

Table 3.5 Examples of cellular proteins acting as membrane receptors for different retroviruses

Retrovirus	Acronym	Receptor	Function
Moloney-murine leukemia virus – ecotropic	Mo-MLV eco	Rec-1 (mCAT-1)	Basic amino acid transporter
Moloney-murine leukemia virus – amphotropic	Mo-MLV anfo	Ram-1	Phosphate transporter
Gibbon ape leukemia virus	GaLV	GLVR-1	Phosphate transporter
Feline leukemia virus	FeLV		
Simian sarcoma virus	SSV		
Avian leukosis virus – subgroup A	ASLV-A	Tv-a	Low-density lipoprotein (LDL) receptor-related protein
Avian leukosis virus – subgroups B, D, and E	ASLV-B, -D, and -E	Tv-b	Member of the tumor necrosis factor receptor (TNFR) family, most likely the avian homologs of mammalian TRAIL receptors
		Tv-c	Member of the immunoglobulin superfamily; most closely resembles the mammalian butyrophilins
Human immunodeficiency virus-1	HIV-1	CD4 CXCR4 or CCR5	T-cell receptor Chemokine receptor

the *Retroviridae* family have evolved SU proteins with very different receptor specificity (Table 3.5). For example, the alpharetrovirus ALV encompasses at least 10 different subgroups based on the capacity to bind different cellular receptors. In general, the physiological function of each receptor is different and not necessarily related to the biology of viral infection or the pathogenesis of disease induced.

Fusion and *uncoating*. The entry step is activated by binding of SU to the cellular receptor and mediated by a conformational change of TM, resulting in the fusion of the viral envelope with the plasma membrane. Following this event, the contents of the virions are found inside the cell cytosol and uncoating of the genome RNA from the capsid proteins takes place.

Reverse transcription. This process occurs in the infected cell's cytosol and is catalyzed by the viral RT enzyme (Figure 3.10). It can be subdivided into 9 subsequent steps. (i) A specific sequence immediately downstream of the 5' LTR (*primer binding site*, PBS) hybridizes, thanks to its complementarity, to a cellular tRNA – different retroviruses use different tRNAs; for example, Mo-MLV and HTLV-1 use tRNA^{Pro}, Visna-, Spuma-, and Mason-Pfizer monkey viruses use tRNA^{Lys1,2}, and HIV-1 uses tRNA^{Lys3}. The 3'-OH end of the tRNA functions as a primer for the synthesis of a complementary DNA, which thus
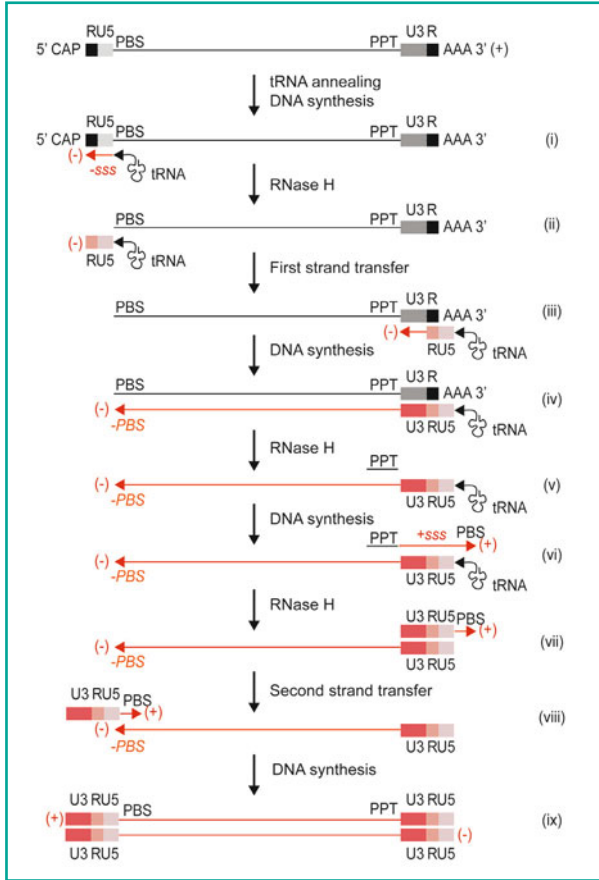
Fig. 3.10 Reverse transcription. See text for description. *PPT*: poly-purine tract; *PBS*: primer binding site

corresponds to the negative (–) strand of the genome. Polymerization progresses to include the R region at the 5' end of the genome. The DNA intermediate formed in this first step is named *minus-strand strong-stop DNA* (*–sss*). (ii) The RT protein is endowed with an enzymatic activity additional to DNA-dependent RNA synthesis, mapping in the C-terminal portion of the protein, consisting in the capacity to digest the RNA moiety in a DNA:RNA hybrid (RNase H activity). Thanks to this activity, the enzyme removes the RNA hybridized to the newly synthesized cDNA, thus exposing the single-stranded R region within the newly synthesized *–sss* cDNA. (iii) This *–sss* cDNA then translocates and hybridizes to the 5' of the viral genome (first strand transfer), thanks to the complementarity between the R regions; the tRNA primer is thus brought into this new position. (iv) Synthesis of the (–) strand cDNA continues starting from the 3'-OH of the *–sss* DNA to reach the PBS region. (v) In the meantime, the RNase H activity of RT continues to digest the viral RNA at the 3' end of the genome, until it stops in correspondence with a central region carrying a short (~10 nt), purine-rich sequence, the *poly-purine tract* (PPT). (vi) The PPT RNA acts as a primer for the synthesis of the positive (+) strand cDNA towards the 3' end of the viral genome. The (+) cDNA that is synthesized includes the

entire LTR and the PBS region downstream of the LTR (*plus-strand strong-stop DNA*, +*sss*). (vii) The RNase H activity of RT removes the tRNA primer at the 5' end of the (–) strand DNA. (viii) Removal of the tRNA exposes the +*sss* DNA PBS region, which can then hybridize to the (–) strand 3' end (second strand transfer). This hybridization, which probably occurs by forming a circular intermediate, generates a template having the 3'-OH extremities on both strands available as primers for polymerization. (ix) RT elongates both strands up to the two LTRs, eventually displacing the previously hybridized regions in the circular intermediate. In this manner, a linear proviral cDNA is generated.

Transport to the nucleus. The newly synthesized, viral cDNA is part of a nucleoprotein complex also containing, besides RT, IN, various and still poorly characterized cellular proteins and, in the case of complex retroviruses, other viral proteins (e.g., Vpr and MA for HIV-1). This complex is called *Pre-Integration Complex* (PIC). In oncoretroviruses, the PIC, which is not smaller than 50 nm, cannot directly enter the nuclear pores, and can thus only have access to the cellular DNA when the infected cells undergo mitosis, after disintegration of the nuclear membrane. In contrast, the lentiviral PIC contains some proteins (including IN, probably Vpr, and perhaps MA in the case of HIV-1) that are able to interact with the nuclear pore proteins and mediate nuclear transport. As a consequence, lentiviruses, but not gammaretroviruses, can infect both replicating and quiescent cells, a property which is of paramount interest for gene therapy applications. Transport of the HIV-1 PIC to the nucleus occurs by sliding on actin microfilaments.

Integration. The process of integration is mediated by the viral protein IN with the assistance of various cellular proteins, only a few of which have been characterized. IN recognizes the extremities of the newly synthesized viral cDNA and removes two terminal nucleotides at the 3' ends on both strands. The processed nucleophilic 3'-OH ends of the viral cDNA are then inserted into the backbone of the target DNA through a transesterification reaction. Integration is random in terms of sequence specificity, however it commonly occurs in regions containing actively transcribed cellular genes. The reason for the selection of these "hot spots" is still unclear. It might be related to the presence, within these regions, of a relaxed chromatin structure that is more accessible to PICs compared to heterochromatic regions, or be due to the specific interaction of some still uncharacterized components of PICs with cellular factors involved in transcription.

Transcription. After integration, the proviral DNA genome can be considered as an additional protein-coding gene of the cell. In fact, transcription of the provirus is carried out by cellular RNA polymerase II and involves the same set of transcription factors that control expression of cellular genes (general transcription factors, mediator, chromatin modification and remodeling factors). The promoter controlling transcription corresponds to the U3 region of the 5' LTR. This sequence is very different in the various members of the *Retroviridae* family, since it consists of a *bricolage* of transcription factor binding sites of the host cell. The presence of these binding sites confers the provirus the property of being expressed in specific cell types or upon specific cell stimulation. As an example, Figure 3.11 displays a schematic representation of the HIV-1 LTR with the indication of the cellular transcription factors binding to this sequence. In some complex retroviruses, besides cellular transcription factors binding the LTR, transcription is controlled by an accessory protein encoded by the viral genome (e.g., Tat in the case of HIV-1 and Tax in the case of HTLV-1). In particular, HIV-1 Tat binds a highly structured RNA sequence
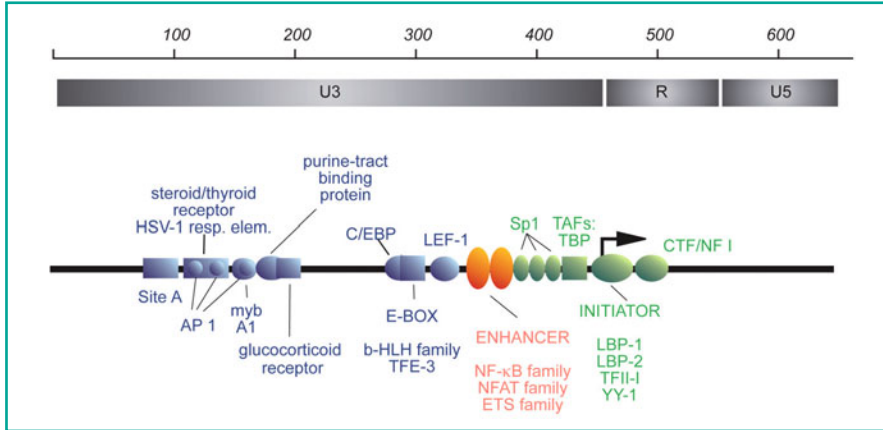
Fig. 3.11 Structure of the HIV-1 long terminal repeat (LTR). The major proteins binding the basal promoter (*green*), enhancer (*orange*), and upstream promoter elements (*blue*) are indicated. The arrow indicates the transcription start site

positioned at the 5' end of the viral mRNA named TAR (*trans-activating response*). From that region, Tat mediates transcriptional activation by recruiting, to the viral LTR promoter, on one hand cellular proteins possessing histone-acetyltransferase (HAT) activity and thus inducing chromatin relaxation and, on the other hand, the cellular P-TEFb kinase, which phosphorylates the carboxy-terminal tail of RNA polymerase II, a modification required for transcriptional elongation. Since transcriptional control is an essential step in the replicative cycle of the retroviruses, it also represents, together with receptor binding, a major determinant of the tropism of the different retroviruses for specific cell types.

Splicing of viral mRNAs. RNA polymerase II generates a single transcript, starting in correspondence with the first nucleotide of the R sequence at the 5' LTR and ending at the polyadenylation site in the U5 sequence of the 3' LTR. This transcript is both the viral genome that eventually becomes packaged inside the virions and the mRNA for the synthesis of all viral proteins. Since, in all eukaryotic cells, mRNAs are monocystronic (that is, each one codes for a single polypeptide), the primary proviral mRNA must undergo splicing to generate different shorter mRNAs, each one devoted to translation of a specific protein. In particular, in cells infected with simple retroviruses, such as Mo-MLV, two mRNAs are found, one corresponding to the original, full-length transcript and a shorter one, in which a large intron in the 5' half of the primary mRNA has been removed (Figure 3.7). The primary mRNA is used for translation of *gag* and *pol*, which occurs in the cytosol, and the shorter for *env*, which is translated by ribosomes associated with the ER. In the case of deltaretroviruses (e.g., HTLV-1), lentiviruses (e.g., HIV-1), and spumaviruses the situation is more complex, since these viruses also code for a series of accessory proteins, for each of which at least one specific mRNA is required. These shorter mRNAs are generated through a multiple splicing process making use of different 5' (splicing donor, SD) and 3' (splicing acceptor, SA) splicing sites. For example, in cells infected with HIV-1, over 35 different mRNAs are generated, which can be classified into 3 different classes according to their length (Figure 3.12). The longer class has ~9 kb and includes a single transcript, corre-
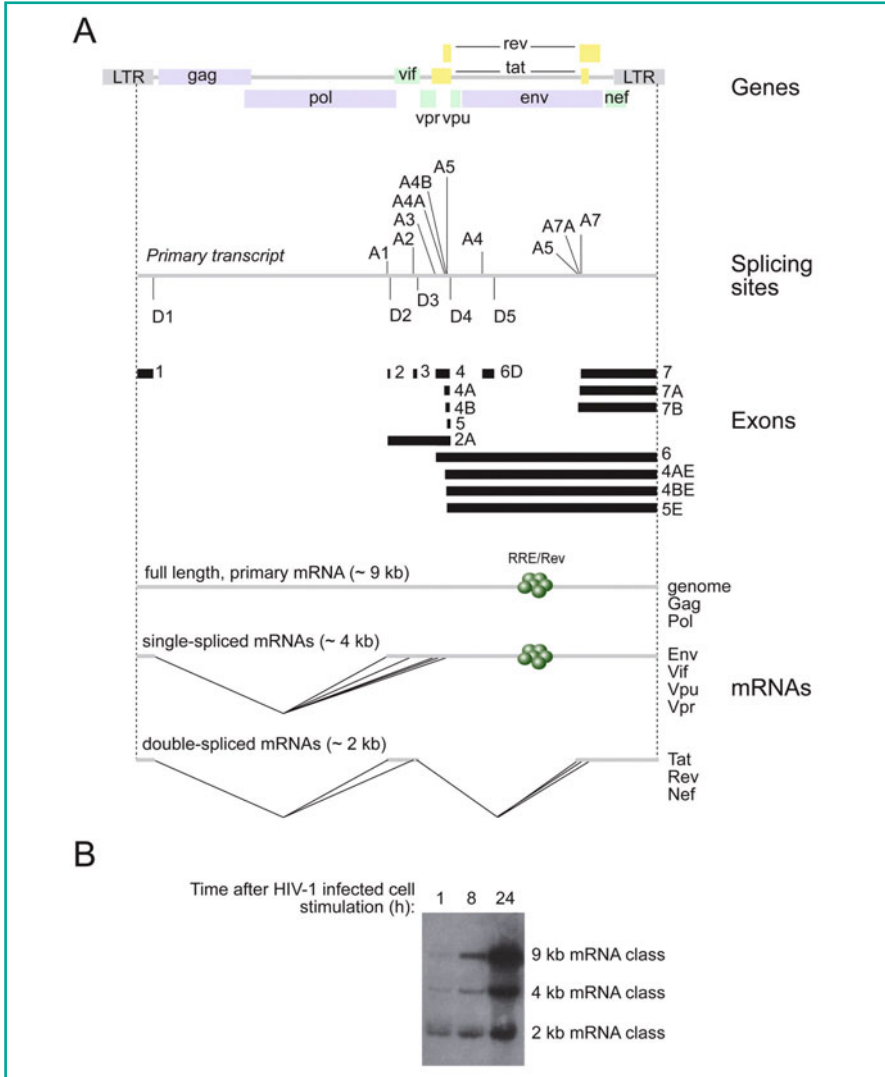
Fig. 3.12 HIV-1 transcripts. **A** Processing of HIV-1 mRNA. The integrated provirus is transcribed to generate a single RNA, containing multiple 5' (splice donor, D) and 3' (splice acceptor, A) sites. These splice sites encompass several exons, containing the open reading frames for all viral proteins. Splicing generates over 35 different mRNAs, which can be grouped into one of three classes: the longest (~9 kb) corresponds to the full-length mRNA; the intermediate (~4 kb) groups all constructs spliced in the 5' portion of the genome; the shortest (~2 kb) includes transcripts spliced twice or more. The first two classes of mRNAs contain the RRE element binding Rev. **B** Northern blotting showing expression of HIV-1 mRNAs in the U1 monocytic cell line, containing two copies of HIV-1 proviral DNA, where transcription is latent under basal conditions and can be activated by a variety of stimuli, including antibodies or cytokines. Upon activation, the shortest transcripts accumulate first, followed by the intermediate mRNAs and, last, by the full-length, genomic mRNA

sponding to the primary mRNA generated by proviral transcription; this mRNA codes for two polyproteins corresponding to Gag and Gag-Pol (see below). The second class includes a series of ~4-kb-long mRNAs, generated by splicing of an intron corresponding to the *gag-pol* sequences. This splicing event utilizes a 5' SD upstream of *gag* and different 3' SA sites downstream of *pol*. The mRNAs generated by this process code for Env (the majority), or the accessory proteins Vif, Vpr, and Vpu. The third class of transcripts includes shorter mRNAs (~2 kb) that, besides removal of the above-described intron, undergo additional splicing events that use different SD and SA sites to remove an intron in the *env* region. These mRNAs code for Tat, Rev, and Nef.

Transport of viral mRNAs. The production of multiple mRNAs by alternative splicing generates the problem of how to transport, into the cytosol, mRNAs that are not fully processed and still contain introns. Different solutions to this problem have been evolutionarily found by the different retroviruses (Figure 3.13A). In the oncoretroviruses, the primary mRNA contains specific sequences (*constitutive export elements*, CTE) that promote nuclear export by binding cellular proteins. In complex retroviruses, a virus-encoded protein (Rev and Rex in the case of HIV-1 and HTLV-1 respectively) binds a structured RNA sequence located in the primary mRNA in correspondence with the *env* gene (*Rev-responsive element*, RRE and *Rex-responsive element*, RXRE, respectively), and thus contained in a potential intron present in the partially spliced mRNAs (Figure 3.13B). Both Rev and Rex bind the cellular protein Crm-1 at the nuclear pores and thus promote translocation, into the cytosol, of partially spliced transcripts retaining the RRE or RXRE sequences. In the case of HIV-1, the RRE sequence plays an important role in the design of lentiviral vectors – see below.

Translation of viral mRNAs. The differentially spliced retroviral mRNAs coding for the essential genes *gag*, *pol*, and *env* are translated into polyproteins (Gag, Gag-Pol, and Env), which in turn are cleaved to generate the final polypeptides. The Gag polyprotein produces the MA, CA, and NC proteins (plus additional small polypeptides in some retroviruses); the Gag-Pol polyprotein generates the RT, PR, and IN enzymes; the Env polyprotein generates the SU and TM proteins (Figure 3.6). In the case of Gag and Pol, proteolytic cleavage of the polyproteins is carried out by the viral PR enzyme; in the case of Env, a furin protease of cellular origin cleaves the polypeptides inside the Golgi apparatus in parallel with protein glycosylation and before the protein is exposed onto the plasma membrane. While Env is produced from one or more specific mRNAs, Gag and Gag-Pol are usually translated from the same mRNA. In gammaretroviruses (the most frequently used to generate retroviral vectors; e.g., Mo-MLV), the *gag* and *pol* genes are on the same open reading frame (orf), separated by a single Stop codon. The ribosome starts translation from the first AUG to synthesize the Gag polyprotein; once it reaches the Gag Stop codon, sometimes it ignores it (in-frame suppression of termination), inserts an additional amino acid instead of this codon, and then continues translation, thus generating a polyprotein also containing Pol. In alpharetroviruses (ASLV) and lentiviruses (HIV), Gag and Pol are on different orfs. Thus, formation of the Gag-Pol fusion polypeptide requires a ribosomal frameshift during translation, by which the ribosome, once in correspondence with the *gag* 3' extremity, moves one nucleotide backward and changes reading frame. Since both suppression of termination and ribosomal frameshifting have an efficiency of 5–10%, the amounts of Gag proteins synthesized are about 10–20 times higher than those
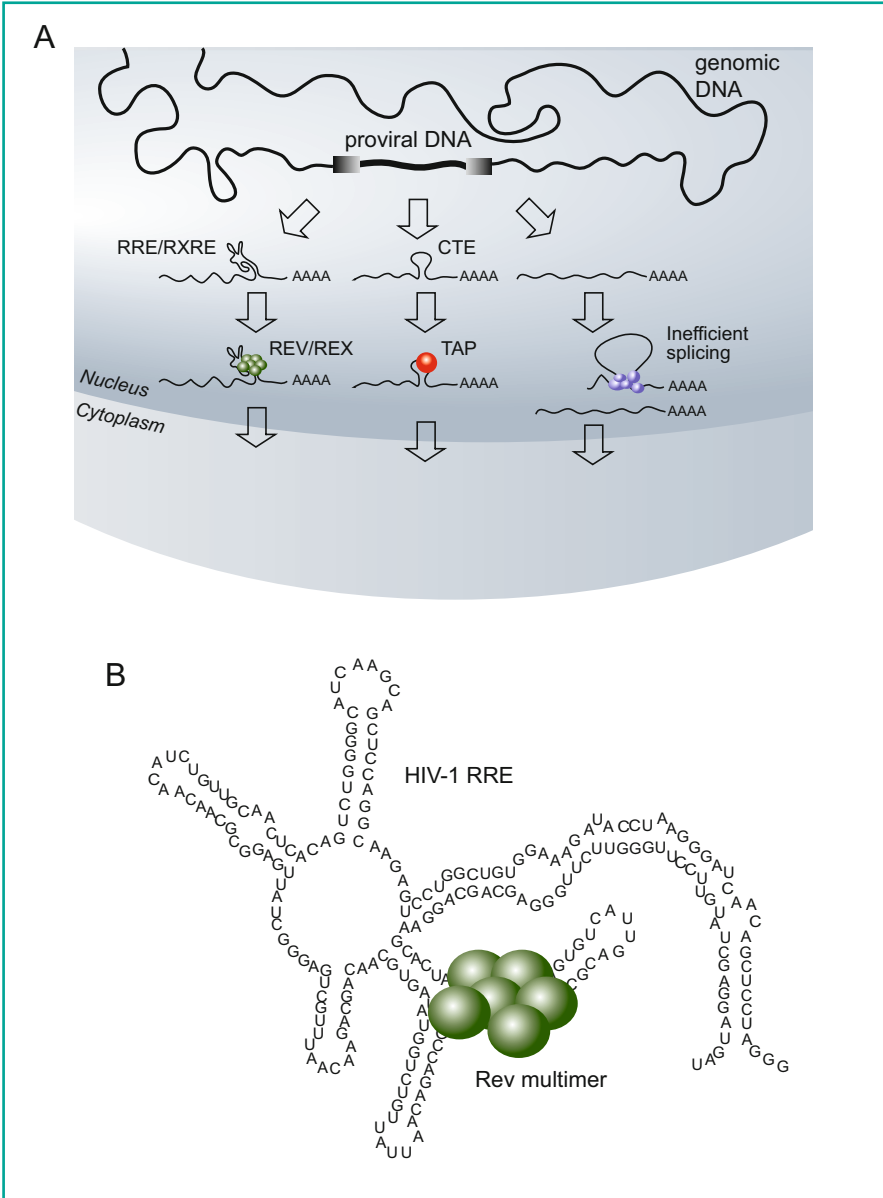
Fig. 3.13 Transport of unspliced retroviral RNAs into the cytoplasm. **A** Different strategies for retroviral RNA export. Complex retroviruses have evolved binding of a viral protein to a cognate site in the unspliced RNAs (HIV-1: RRE/Rev; HTLV-1: RXRE/Rex) to promote export of unspliced mRNAs (*left part*). In oncoretroviruses, the primary mRNA contains specific sequences (constitutive export elements, CTE) that promote nuclear export by binding cellular proteins (*middle*). Other simpler retroviruses exploit inefficient splicing to allow export of incompletely spliced transcripts (*right*). **B** Sequence and secondary structure of the HIV-1 RRE RNA region, with the indication of the Rev-binding site. The protein binds as a multimer

of Pol. As discussed above, in complex retroviruses, each of the accessory genes is usual-ly translated from one or more specific mRNAs, generated from the primary transcript by alternative splicing.

Assembly. Assembly of virions occurs in correspondence with the plasma membrane (for type C viruses) or in the cytosol (for type B and D viruses); in both cases, assembly is mainly driven by the Gag polypeptide. Inclusion of the genome-length viral mRNA into the virions is due to the interaction of a specific RNA sequence, the *packaging signal* (ψ), located in corre-spondence with the 5' of the *gag* gene, with the Gag polypeptide in the portion corresponding to the NC protein. Each virion includes 1200–1800 Gag and 100–200 Gag-Pol polyproteins.

Budding and virion maturation. The Env polyprotein is independently translated inside the ER, becomes glycosylated and matures into TM and SU in the Golgi apparatus, and is then exposed on the cell membrane of the infected cells. In the regions of viral bud-ding, the protein becomes associated with the virions thanks to its interaction with the N-terminus of Gag. Once outside the cells, the virions undergo maturation, by which the Gag and Gag-Pol polyproteins are cleaved to generate the respective individual peptides. Proteolytic cleavage is mediated by PR, which first excides itself from Gag.

3.5.1.2
Structure of Gammaretroviral Vectors

A prototype member of the gammaretrovirus genus used as a vector for gene therapy, the Mo-MLV contains, besides the three essential *gag*, *pol*, and *env* genes, at least 5 genetic ele-ments that are necessary for the completion of its replicative cycle and thus essential for the construction of vectors. These are (listed from the 5' to the 3' of the genome; Figure 3.14):
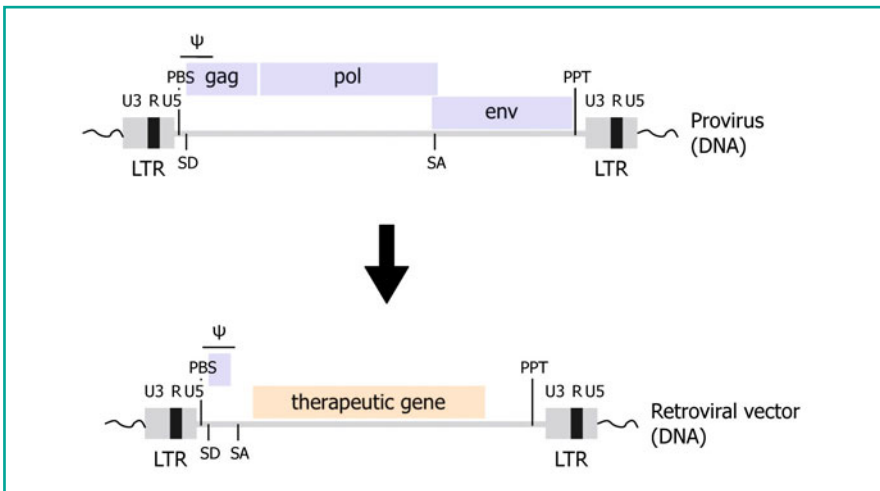
Fig. 3.14 Schematic structure of retroviral vectors. *Upper scheme*: genetic composition of a proto-type gammaretroviral vector (e.g., Mo-MLV), with the indication of the most relevant genetic ele-ments (*LTR*: long terminal repeat; *PBS*: primer binding site; *PPT*: poly-purine tract; *SD*: 5' splice site; *SA*: 3' splice site). *Lower scheme*: retroviral vector

(i) the **LTRs**, of which the 5' U3 region is the promoter for mRNA transcription, the R region is required for reverse transcription, and the 3' U5 region contains the polyadenylation site;

(ii) the **primer binding site** (PBS), positioned immediately downstream of the 5' LTR, which is required for cellular tRNA binding to prime reverse transcription;

(iii) the **5' and 3' splice sites** (SD and SA respectively), which are essential to generate the spliced mRNA used for the translation of *env*; the sequence starting from the U3/R region and ending at the SD site is named the *leader* sequence and is common to all transcripts;

(iv) the **packaging signal** (ψ), which includes a structured RNA region at the 5' of the *gag* gene, partially extending toward the SD site; this is the sequence binding to Gag that is required for the inclusion of the viral genome mRNA inside the virions during assembly; and

(v) the **polypurine tract** (PPT), positioned at the 3' end of the genome upstream of the 3' LTR, which is required for reverse transcription.

Retroviral vectors must contain these five genetic elements, while the rest of the genome is dispensable and can be removed and be substituted by the therapeutic gene, including the sequences coding for the viral proteins (Figure 3.14). Thus, in the simplest version of oncoretroviral vectors, transcription of the therapeutic gene is directly controlled by the viral 5' LTR.

3.5.1.3
Production of Gammaretroviral Vectors

Viral vectors based on gammaretroviruses are produced in cultured mammalian cells. A plasmid containing the proviral DNA, having a structure similar to that depicted in Figure 3.14, is first obtained by standard cloning procedures, amplified in bacteria, and purified. This plasmid is then transfected into a *packaging cell line*, that is a cell line, usually of murine origin, that expresses the retroviral *gag*, *pol*, and *env* genes, which are no longer present in the retroviral vector plasmid but are nevertheless required for virion production. A packaging cell line is usually generated by stable transfection of the DNA sequences coding for Gag-Pol and Env, and thus constitutively expresses the respective proteins. Transfection of the Gag-Pol and Env DNAs to generate a packaging cell line is usually carried out in two subsequent steps, to avoid the possibility that the two constructs might integrate into contiguous regions of the genome, which would favor recombination of the two sequences with that of the retroviral vector plasmid (or with sequences corresponding to endogenous retroviruses (ERVs)), with the consequent generation of infectious RCRs. As a matter of fact, regions of homology no longer than 10 bp between the packaging sequences and the retroviral vectors are sufficient to drive recombination between the two constructs, leading to the generation of infectious viruses able to replicate autonomously. If the Gag-Pol and Env sequences are integrated far apart in the genome, the likelihood of recombination is significantly diminished. Over the first ten years of development of gene therapy, a vast series of packaging cell lines have been generated starting from the *gag*, *pol*, and *env* genes of different murine and avian retroviruses. In particular, since the tropism and efficiency of infection are mainly due to the properties of the Env proteins, these packaging cell lines vary in their capacity to generate retroviral vectors with the capacity to transduce different cell targets – see also below.

Once transfected into a packaging cell line, the plasmid containing the retroviral vector is transcribed starting from the 5' LTR and thus generates an mRNA that encompasses the whole proviral construct and contains the packaging signal (ψ). Presence of this signal permits recognition of the vector mRNA by Gag, followed by its inclusion into a virion (Figure 3.15). A virion generated in this manner is indistinguishable from a wild-type virion, and is thus fully infectious. After infection of a target cell, thanks to RT (which is present inside the virion) and the cis-acting PBS and PPT sequences, the vector genome
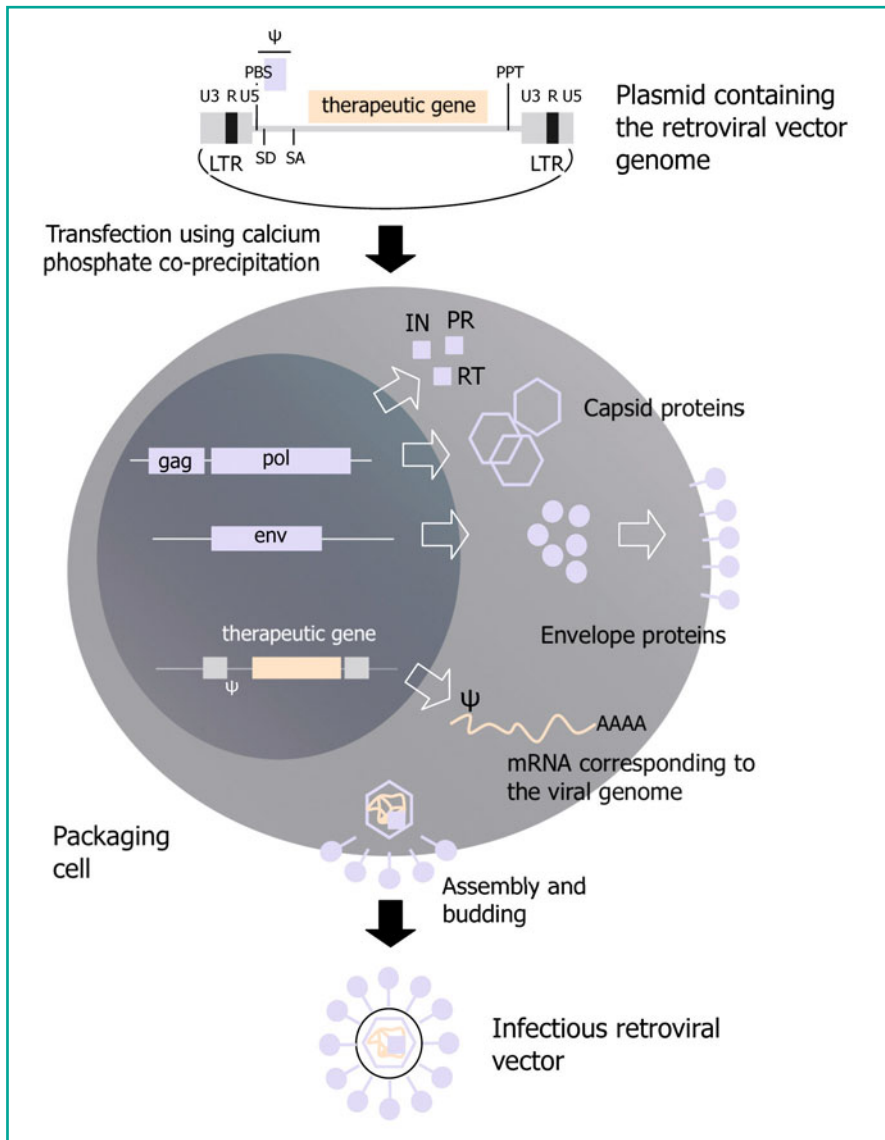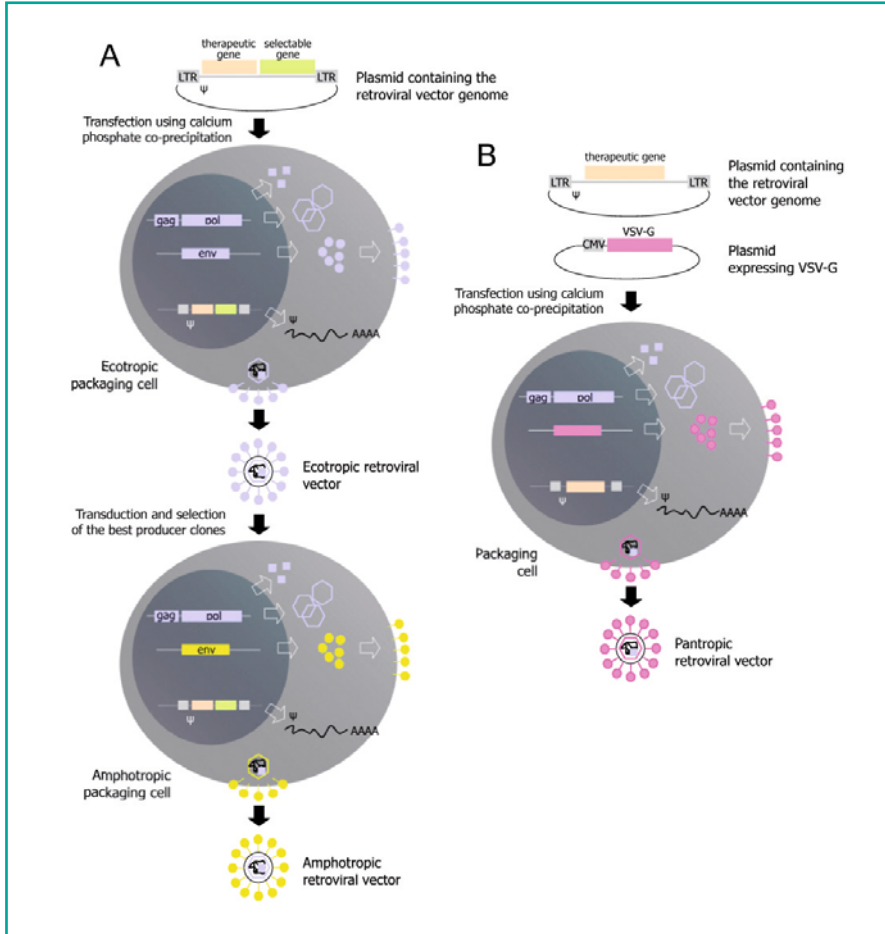

Fig. 3.15 Production of retroviral vectors in a packaging cell line. See text for description

is reverse transcribed. The proviral cDNA is then integrated into the host cell genome by IN, which is also present in the virion. Once integrated, the mRNA expressed by the vector provirus is no longer infectious, since none of the retroviral proteins are present. Thus, retroviral vectors are only capable of a single cycle of infection.

As anticipated above, the properties of the available packaging cell lines are mainly related to the characteristics of the *env* gene they express, since the SU protein dictates tropism of the virions towards different cell types. In particular, the Mo-MLV retrovirus, which is the prototype of several gene therapy retroviral vectors, naturally presents two SU variants. One exclusively binds the murine receptor Rec-1 (also called mCAT-1), a basic amino acid transporter, while the other one binds Ram-1, a phosphate transporter (Table 3.5). While the former protein is only expressed by murine cells, the second is common to cells of several species, including humans. Therefore, Mo-MLV displaying the Rec-1-binding SU variant only infects murine cells, while that displaying the Ram-1-binding SU infects cells of all species. According to their tropism, the former viruses are called *ecotropic* (indicating that they only replicate in cells of the same species in which they were isolated) and the latter *amphotropic* (*ampho-*: Greek for "both", indicating that they replicate well in cells of both of the species from which they were isolated and other species). Finally, *xenotropic* (*xeno-*: Greek for "foreign") retroviruses are endogenous to one species, but can only be propagated well in cells from a species foreign to the normal host (e.g., ERVs of mice that replicate well in rat or hamster cells). Amphotropic, but not ecotropic, vectors can be used for gene therapy of human cells.

Transfection of the retroviral vector-containing plasmid into the packaging cells is commonly performed using conventional methods, such as calcium phosphate precipitation or lipofection, and is thus relatively inefficient, since most of the internalized DNA is degraded or integrates into the cellular DNA in a random manner, often interrupting the continuity of the retroviral vector sequence. Therefore, the amount of retroviral particles found in the packaging cells' supernatant is limited, and thus the titer of the preparation (measured as the concentration of infectious particles) is relatively low. In contrast, if this supernatant is used to infect a second packaging cell line and the retrovirus contains a gene allowing the selection of the transduced cells, it is possible to obtain a population of homogenously transduced packaging cells, which release significant amounts of vectors in their supernatants. The first retroviral vectors designed for gene therapy, therefore, also contained, in addition to the therapeutic gene, a selectable gene, such as the *neo* gene, which confers resistance to the antibiotic geneticin or G418. These vectors are produced by a two-step procedure, the first entailing calcium-phosphate transfection of retroviral plasmid DNA into an ecotropic packaging cell line and the second using the supernatant produced by these cells to transduce an amphotropic packaging cell line, followed by selection of individual producer clones and analysis of the viral titers obtained in each case (Figure 3.16A). Following this procedure, the retroviral vector titers that can be achieved are in the order of $\sim 1 \times 10^6$–1×10^7 infectious particles/ml of supernatant. Application of this procedure is very useful to obtain a retroviral vector-producing cell clone, which can be used for the continuous production of retroviral particles for a given gene therapy application without having to rely on transient transfection each time, which is cumbersome and has variable efficiency.

During assembly of retroviral particles, the viral genomes are packaged thanks to the interaction of the Gag polyproteins with the ψ sequence, occurring in the cytosol, while the

Fig. 3.16 Production of retroviral vectors. **A** Dual-step selection of retroviral vector producer cell clones. The procedure entails a first transfection of an ecotropic packaging cell line, followed by transduction of an amphotropic cell line using the supernatant obtained from the first transfectants. **B** Pseudotyping of retroviral vectors. A packaging cell line only expressing Gag and Pol is transfected with the retroviral vector plasmid and a second plasmid encoding VSV-G

TM and SU Env proteins are independently brought to the cell membrane through the ER-Golgi route. In the course of the study of retrovirus biology it was noticed that, if a cell is infected with a retroviral vector with a given specificity but, at the same time, it expresses an Env gene with a different specificity, some of the viral particles that are produced have the infectious property dictated by this Env gene, however carry a genome with different characteristics. This is known in virology with the term *pseudotyping* (*pseudo-*: Greek for "false"). Thanks to this property, it is possible to package the same viral genomes inside particles having efficiency and specificity of infection that are different from those encoded by the genome itself and are instead dictated by the Env proteins displayed by the virions.

The most efficient possibility for retroviral vector pseudotyping is through the use of the G protein encoded by the vesicular stomatitis virus (VSV). VSV is an enveloped virus with a negative sense RNA genome belonging to the family of *Rhabdoviridae*. The virus is of veterinary interest and is a cause of concern, since it can infect different animal species in an epidemic manner, including cattle. In the infected animals, the disease is characterized by the appearance of vesicles in different organs, including mouth and tongue, hence the name (*stoma*-: Greek for "mouth"). In humans, the virus can cause a flu-like syndrome, with the occurrence of vesicles on the lips resembling herpetic infection. The VSV envelope displays the virally encoded G protein (VSV-G), which mediates infection by binding, very efficiently, the phospholipids present on virtually all mammalian cell membranes and triggering endocytosis of the viral particles. Once in the endocytic compartment, lowering of the pH activates the fusogenic properties of VSV-G, which determines fusion of the viral envelope with the endosomal membrane and release of the virion content into the cytosol. Thanks to these properties, the VSV-G protein, once incorporated in a retroviral envelope, mediates viral infection at high efficiency and broad specificity.

Due to its fusogenic properties, it is however not possible to permanently express VSV-G in a packaging cell line. VSV-G-pseudotyped retroviral vectors are thus obtained by the transient transfection of packaging cell lines that only express Gag-Pol with a plasmid expressing VSV-G under the control of a strong promoter (such as the promoter of the cytomegalovirus immediate-early genes), in addition to the plasmid containing the retroviral vector DNA (Figure 3.16B). In contrast to virions containing retroviral Env proteins, VSV-G-pseudotyped virions can be purified by high-speed centrifugation without significant loss of infectivity. By using a single packaging passage followed by centrifugation, titers in the order of $\sim 1 \times 10^8$–1×10^9 infectious particles/ml of supernatant can routinely be obtained. The VSV-G-pseudotyped retroviral vectors have broad species specificity and cell-type range.

3.5.1.4
Variants in the Design of Gammaretroviral Vector Genomes

The simplest retroviral vector maintains all the essential genetic elements in cis (LTR, SD/SA, PBS, PPT, and ψ) and contains the therapeutic gene cloned within the two LTRs (Figure 3.17A). In principle, since the vector contains a single gene, the SA/SD sites are also dispensable, however their presence confers stability to the mRNA and thus allows higher titers and expression levels.

Starting from this relatively simple genetic design, over the last several years a number of modifications have been proposed, with the main purpose of allowing delivery of additional genes or driving expression of the therapeutic gene from a promoter different from the LTR. The main classes of variant vectors are described as follows.

(1) **Vector with an internal promoter**. One of the most frequent conditions in gene therapy is that expression of the therapeutic gene is driven by a promoter different from the vector LTR (e.g., a strong constitutive, or an inducible, or a tissue-specific promoter). In these cases, it is possible to insert the promoter of choice upstream of the therapeutic gene and thus downstream of the 5' LTR (Figure 3.17B). The cells containing such a construct, however, express two mRNAs, one starting from the 5' LTR and the other one from
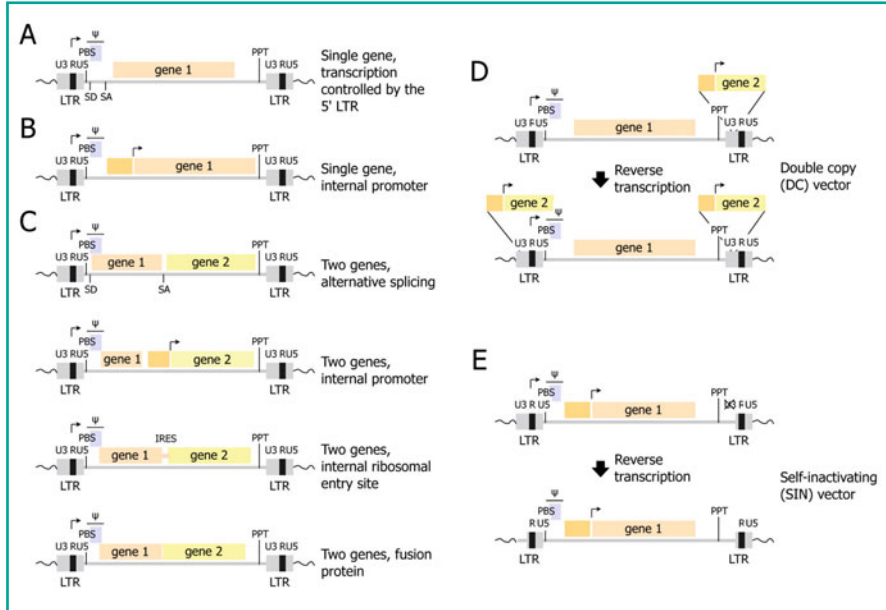
Fig. 3.17 Variants in the construction of retroviral vector backbones. **A** Canonical retroviral vector in which therapeutic gene expression is driven by the viral LTR. **B** Retroviral vector containing an internal promoter driving expression of the therapeutic gene. **C** Retroviral vectors containing two genes. **D** Double copy (DC) vector, before and after reverse transcription. **E** Self-inactivating (SIN) vector, before and after reverse transcription

the internal promoter, both ending in correspondence of the polyadenylation site of the 3' LTR. This condition is far from being ideal, since transcriptional interference is likely to ensue between the two promoters, due to a conflict in the assembly of the transcriptional machineries and to the disturbance that read-through transcription exerts on the downstream promoter. In addition, viral production exclusively relies on LTR-driven transcription: should the internal promoter be very strong, the titers of the viral preparations obtained from the packaging cells are usually low. The combination of two promoters is only effective when the internal promoter is weak or inactive in the packaging cell lines and becomes strong in the final target cells, as might be the case for a tissue-specific promoter, while the opposite is true for the viral LTR.

(2) **Vectors expressing two genes**. To simultaneously express two genes (as is the case, for example, of retroviral vectors containing a selectable gene in addition to the therapeutic gene), at least four different strategies can be followed (Figure 3.17C), detailed as follows.

(i) The first approach is to maintain the SA and SD sequences and clone one of the genes between these sites and the other one downstream of the SD sequence. This arrangement recapitulates that of wild-type simple retroviruses, with expression of the downstream gene relying on splicing of the primary mRNA.

(ii) A second possibility is to clone one gene under the control of the LTR and the other one under the control of an internal promoter. In this case, however, transcriptional competition might ensue between the two promoters, as discussed above.

(iii) A third option is to clone both genes under the control of the LTR and separate the two sequences by the insertion of an internal ribosomal entry site (IRES); this is by far the most efficient solution to both obtain high virus titers and permit high-level gene expression. In this arrangement, the vector produces a single transcript that is used for translation of two proteins from two different AUG codons. The IRESs that are most used for this purpose are those derived from viruses belonging to the Picornaviridae family or from the hepatitis C virus (HCV), since the RNA genomes of these viruses are not capped at their 5' ends and thus require the IRES to direct translation of their own proteins.

(iv) Finally, a fourth possibility is to clone the coding sequences of two genes of interest in frame, in order to obtain a single fusion protein; this strategy obviously requires that both proteins retain their function when fused to their respective partners.

(3) **Double copy vectors**. A very interesting strategy that can be used for the expression of short therapeutic nucleic acids (typically: ribozymes, shRNAs; cf. section on 'Modes of Delivery or Intracellular Synthesis of Small Regulatory RNAs') is to clone the transcriptional cassette expressing these genes within the U3 sequence of the 3' LTR, without interfering with the transcriptional elements contained in this region (Figure 3.17D). In the packaging cells, such a construct is transcribed starting from the 5' LTR and generates an mRNA originating in the 5' R region and ending in correspondence of the 3' U5 region; this mRNA thus contains the modified U3 sequence. During reverse transcription in the target cells, RT also transfers this modified U3 to the 5' LTR, thus duplicating the therapeutic gene. Inside the DC vector LTRs, another gene can be present (e.g., a selectable gene), the transcription of which is normally controlled by the 5' LTR.

(4) **Self-inactivating (SIN) vectors**. As reported above, the 5' LTR U3 region, which controls proviral transcription, is generated during the process of reverse transcription, when the 3' U3 sequence jumps to the 5' end of the genomic RNA. It is thus possible to construct retroviral vectors that, in their plasmid form, contain an intact 5' LTR and a 3' LTR that is mutated or almost entirely deleted. In the packaging cells, these vectors generate a transcript that contains this modified U3 sequence, which will become duplicated at the 5' LTR during reverse transcription in the target cells (Figure 3.17E). The provirus generated in this manner will be incapable of driving transcription of a therapeutic gene cloned within its LTRs, unless a promoter is inserted upstream of this gene. This strategy is very useful to avoid transcriptional interference between the LTR and an internal promoter. In addition, since the viral LTRs often activate expression of cellular genes neighboring the proviral integration sites (see section on 'Gene Therapy of Hematopoietic Stem Cells'), this strategy is currently considered as a possible means to minimize this problem.

3.5.1.5
Properties of Gammaretroviral Vectors

Viral vectors based on gammaretroviruses were the most utilized vectors in the gene therapy clinical trials until the early 2000s. Their popularity was due to a number of reasons, including the relative simplicity of use, high efficiency of transduction of replicating cells (e.g., *ex vivo* cultured cells), the immunogenicity, and ability to integrate their proviral

cDNA form into the host cell genome, with the potential to render transduction, and thus therapeutic gene expression, permanent.

The clinical trials so far conducted, however, have highlighted a series of important problems, which have strongly limited the use of these vectors in more recent years.

(i) Some of the problems are related to the construction of the vectors themselves and the modalities of their production. Most of these problems can be circumvented by better vector design, as already outlined above. For example, transcriptional interference between two promoters, which lowers viral titers on one hand and therapeutic gene expression on the other, can be avoided by the use of simple vectors in which a single gene is directly controlled by the viral LTR. Potential formation of infectious recombinant viruses can be controlled by using packaging constructs coding for Gag-Pol that lack any homology stretch with the retroviral vector sequence. Finally, low titers and limited tropism of amphotropic preparations can be circumvented by VSV-G pseudotyping.

(ii) An additional technical issue related to the use of retroviral vectors is caused by their size, which usually permits cloning of therapeutic genes no longer than 6–7 kb (a retroviral particle can only package mRNAs no longer than a total of 9–10 kb). This prevents the delivery of native genes and is an impediment for the use of these vectors for very long cDNAs.

(iii) By far more important in applicative terms is the absolute need of gammaretroviral vectors that their target cells are in active replication. In fact, the pre-integration complex (PIC) of these viruses, which includes the viral cDNA and a series of proteins of cellular and viral origin including IN, remains in the cytosol and does not have access to the nucleus unless during mitosis, when the nuclear membrane breaks down. Since most of the cells in our body, including neurons, skeletal muscle cells, cardiomyocytes, endothelial cells, and the vast majority of peripheral blood lymphocytes, rarely divide or do not divide at all, the use of gammaretroviral vectors is essentially restricted to *ex vivo* applications on cells actively maintained in the cell cycle.

(iv) A fourth essential limitation of gammaretroviral vectors relates to the progressive silencing of therapeutic gene expression in the transduced cells. This occurrence is a consequence of methylation of cytosines in the context of the CpG di-nucleotide at the level of vector LTR promoter region. Methylated cytosines are recognized by various methyl-cytosine-binding proteins, which eventually promote chromatin deacetylation and compaction, eventually leading to silencing of gene expression. Methylation of retroviral DNA is believed to provide an evolutionary response aimed at preserving the integrity of the cellular genetic information against the insertion of transposable elements; in this respect, it is worth considering that over 8% of the human and mouse genomes indeed consists of endogenous retroviral sequences (ERVs), corresponding to over 30,000 proviruses per genome, divided into at least 50 different families, some of which are capable of autonomous replication.

(v) Finally, and probably most important of all, a very serious problem that limits the use of gammaretroviral vectors is the possibility that integration into the transduced cell genome might be mutagenic, leading to the inactivation of a tumor suppressor gene or the activation of an oncogene, thus contributing to oncogenic transformation. While, in theory, this event appears unlikely since retroviral vectors do not replicate and are capable of a single integration event, it has already occurred in at least two clinical trials for gene therapy of SCID-X1, a severe inherited immunodeficiency. Since then, inappropriate gene

activation has also been observed in a series of other experimental studies in cultured cells and animal models; this issue will be discussed in more detail in the section on 'Gene Therapy of Hematopoietic Stem Cells'). In light of these problems, the regulatory agencies in Europe no longer approve the use of Mo-MLV-based retroviral vector for stem cell gene therapy of non-lethal disorders.

3.5.2
Vectors Based on Lentiviruses

One of the most striking characteristics that distinguishes lentiviruses from gammaretroviruses is the ability of the former to infect non-replicating cells. For example, one of the relevant cell types infected by HIV-1 *in vivo* are macrophages, which are terminally differentiated cells that have exited the cell cycle. As discussed above, this property is due to the capacity of the lentiviral PIC, which forms in the cytosol, to actively cross the nuclear membrane thanks to the interaction of some of the PIC proteins (IN, MA, Vpr) with proteins of the nuclear pore. This property appears of paramount interest for gene therapy, since it allows a significant extension of the range of cell types in which gene transfer might be of therapeutic benefit, especially because most of the cells in our body are non-replicating. In addition, in the case of hematopoietic stem cells, *ex vivo* transduction with lentiviral vectors appears efficient also in the absence of growth factor stimulation, a condition permitting the preservation of their pluripotency (cf. section on 'Gene Therapy of Hematopoietic Stem Cells').

For these reasons, starting from the late 1990s, the possibility to obtain vectors based on HIV-1 and the other lentiviruses has appeared very appealing for *in vivo* and *ex vivo* gene transfer applications.

3.5.2.1
Structure and Production of Lentiviral Vectors

The lentivirus from which most of the currently available vectors have been generated is HIV-1, mainly because of the vast amount of available information concerning the molecular biology and the properties of this virus. Over the last 10 years, at least three different generation of HIV-1-based lentiviral vectors have been produced, each bearing significant improvements over the preceding ones (Figure 3.18).

In the **first-generation** lentiviral vectors, recombinant viral particles are generated through cell transfection with 3 plasmids.

The *first plasmid* contains, in its proviral DNA form, the gene transfer vector, which carries the therapeutic gene. This proviral DNA contains, in 5' to 3' orientation: (i) the wild-type 5' viral LTR; (ii) the leader region, containing the PBS sequence and the 5' splice site-SD; (iii) ~350 bp of the *gag* gene in the region corresponding to the packaging signal ψ – the gene open reading frame is closed by the insertion of a Stop codon to block translation; (iv) ~700 bp of the *env* gene containing the RRE region, to allow export of the viral transcript from the nucleus, and the 3' splice site-SA; (v) a promoter driving the expression of
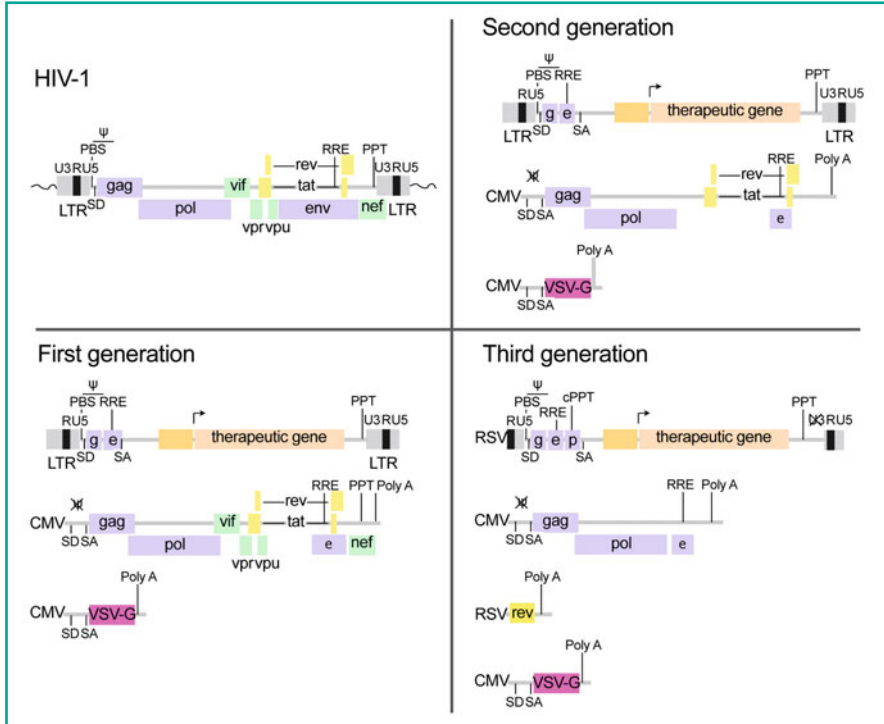
Fig. 3.18 Lentiviral vectors. Schematic representation of the HIV-1 genome (*upper left*) and of the plasmids required to obtain first-, second-, and third-generation lentiviral vectors. See text for description

the therapeutic gene – this is essential, since the natural HIV-1 promoter, consisting in the 5' LTR U3 region, is almost silent in the absence of the viral Tat protein, which is not present in the vector due to safety reasons; (vi) the 3' viral LTR, with the immediately upstream located PPT sequence.

The *second plasmid* is a packaging plasmid, also derived from the HIV-1 genome, which contains all viral genes with the exception of *env*. Besides *env*, this plasmid also carries a mutation in the ψ region to prevent packaging of the encoded mRNA into the viral particles and lacks the 3' LTR, which is substituted by a heterologous polyadenylation sequence. Expression of this plasmid is driven by the strong constitutive promoter of the cytomegalovirus immediate early (CMV IE) genes.

The *third plasmid* codes for the VSV-G protein; to avoid limitations in tropism and circumvent the relative low infectivity of the natural HIV-1 envelope, lentiviral vectors are usually pseudotyped with VSV-G.

Production of the vectors is carried out by the transient transfection of human embryonic kidney 293 T (HEK 293T, expressing the SV40 T antigen protein) cells with these three plasmids; the virions containing the retroviral vector RNA are then found in the cell culture supernatant, similar to gammaretroviral vectors.

First-generation lentiviral vectors elicit important safety concerns, related to both their

production strategy and their clinical utilization. At the levels of production, a recombinant event occurring between the packaging plasmid and the plasmid containing the lentiviral vector can generate a replication competent lentivirus (RCL), the infectivity of which can be even extended by the presence of the VSV-G protein. A similar recombination event could also occur at the moment of reverse transcription in the transduced cells, should the virion carry two RNA genomes, one corresponding to the transfer vector and other to the packaging plasmid. At the level of clinical application, once used for gene therapy of HIV-1 infected patients, first-generation lentiviral vectors might recombine with the wild-type virus infecting the patients, thus potentially leading to the creation of novel viruses, with unpredictable potential for diffusion and pathogenicity. Finally, should first-generation lentiviral vectors be used in an HIV-1-infected patient, it is also possible that vector replication is stimulated by infection of the transduced cells with wild-type HIV-1: in this case, the wild-type virus would act as a helper for vector replication, since the vector construct contains all the sequences necessary for replication, including transcription (LTR), packaging (ψ), reverse transcription (R, PBS, PPT), and integration (U3). Thus, superinfection of cells carrying an integrated lentiviral vector with wild-type HIV-1 would determine the mobilization of the vector inside the organism.

To try and overcome these safety issues, further deletions have been introduced into the HIV-1 backbone to progressively remove all genes that are not strictly necessary for the production of viral particles and the transduction of the target cells. The **second-generation** lentiviral vectors entail the use of a similar three-plasmid design as the first-generation vectors, however, in the packaging plasmid, besides *gag* and *pol*, all accessory genes are removed, with the exception of *tat* and *rev*. In this manner, the probability of recombination between the vector and the packaging plasmid is significantly reduced in the packaging cells. However, since such a second-generation vector carries intact LTRs and ψ region, on one hand it can still recombine with wild-type HIV-1 while, on the other hand, it can be mobilized by wild-type HIV-1 if used in an HIV-1-infected patient, similar to first-generation vectors.

A **third-generation** of lentiviral vectors was designed to definitely prevent both the possibility of recombination with wild-type HIV-1 and of vector mobilization inside the organism. Production of these vectors, which have now entered clinical experimentation, now requires four plasmids. The *first plasmid* corresponds to the transfer vector, which is now obtained using the SIN approach (cf. gammaretroviral vectors above) to modify the LTR region. In particular, the 3' U3 LTR region is deleted, to inactivate transcription of the proviral DNA after reverse transcription. In the packaging cells, the vector is transcribed from a constitutively active heterologous promoter, positioned upstream of the R region. In addition, recent evidence indicates that the inclusion, inside the vector proviral DNA, of an HIV-1 sequence located within the *pol* gene significantly increases viral titers. This sequence, named *central polypurine tract/central termination sequence* (cPPT/CTS), would function by enhancing both reverse transcription – acting as an additional PPT to drive synthesis of the *plus-strand strong-stop DNA* (cf. replicative cycle of retroviruses above) – and PIC nuclear transport.

Since transcription of the transfer plasmid depends on a heterologous promoter, the presence of the *tat* gene, coding for the transactivator acting on the viral LTR becomes superfluous. Thus, the *packaging plasmid* now only contains the *gag* and *pol* genes, while the *rev* gene is expressed from a *third plasmid*. Presence of the Rev protein is still neces-

sary, since it allows proper transport into the cytosol of the mRNA expressed from the packaging plasmid. Finally, as in the previous generations, a *fourth plasmid* encodes VSV-G.

This third-generation lentiviral vector production system only requires 3 of the 9 HIV-1 genes, thus offering a safety profile that is definitely more reassuring than that of first- and second-generation vectors.

Based on an analogous design, first-, second-, and third-generation lentiviral vectors have also been obtained from the genomes of other non-human lentiviruses, including FIV, SIV, and BIV. Utilization of these vectors rather than those based on HIV-1 would have the advantage of increased safety, since the viruses from which these vectors are derived do not infect humans.

3.5.2.2
Properties of Lentiviral Vectors

Compared to gammaretroviral vectors, the main advantage of lentiviral vectors is their property to transduce non-replicating cells. This paves the way to the possibility of using these vectors *in vivo*, to transduce organs such as brain or retina, which are mainly composed of quiescent cells. In a similar manner, lentiviral vectors can be used for *ex vivo* transduction of hematopoietic stem cells without the need to induce their replication. In this respect, however, it should be observed that, while it holds true that lentiviral vectors are able to successfully transduce cells that are out of the cell cycle, these cells still need to be metabolically active. This requirement is well exemplified by the observation that, during the natural history of HIV-1 infection, the virus very efficiently transduces resting, however metabolically active, macrophages and much less efficiently transduces resting, and metabolically quiescent, peripheral blood T lymphocytes. This appears to be of particular relevance when lentiviral vectors are considered for gene transfer into stem cells of different derivation, since the activation state of these cells is usually low.

The main concerns elicited by lentiviral vectors relate to safety and, in particular, to the generation of RCLs, to the mobilization of vectors by the wild-type virus in HIV-1-infected patients, and to the potential for insertional mutagenesis.

RCL generation can occur either during vector preparation, by recombination of the transfer vector with the packaging plasmid, or *in vivo*, after superinfection of a transduced cell with wild-type HIV-1. The third-generation lentiviral vectors seem to have a significantly better safety profile compared to the previous generations, due to the limited sequence homology with wild-type HIV-1 they present.

As far as vector mobilization by wild-type HIV-1 is concerned, this is more than a theoretical possibility since it has already been observed in HIV-1-infected patients in the first gene therapy clinical trial exploiting a lentiviral vector (see section on 'Gene Therapy of HIV-1 Infection'). This vector, however, contained an intact LTR, which was transcriptionally activated upon HIV-1 infection. In the third-generation vectors, removal of the LTR U3 region using the SIN technology should prevent the possibility of mobilization, since viral replication requires vector transcription starting from the 5' R region, in order to ensure the inclusion, in the viral mRNA genome, of the packaging signal and the R region itself.

Finally, it still remains to be understood whether integration of lentiviral vectors into the host cell genome might lead to the inappropriate activation of cellular genes through insertional mutagenesis, similar to gammaretroviruses. Several studies are currently addressing this issue experimentally, however only scanty primary data are available in humans due to the very limited number of patients treated with lentiviral gene therapy so far. *Ex vivo* cell transduction indicates that these vectors, similar to wild-type HIV-1, also integrate in correspondence with cellular transcribed genes. However, the region where integration occurs corresponds to the whole gene transcription unit, in contrast to gammaretroviruses, which preferentially integrate in correspondence with the transcription start site, including the gene promoter and first intron (Figure 3.19). Since aberrant transcriptional activation of the gene where proviral integration has occurred is likely to depend on either the interaction of the viral LTR elements with the cellular gene promoter or read-through transcription of the gene from an upstream integrated provirus, this is less likely to occur with lentiviral compared to gammaretroviral vectors. In addition, the HIV-1 LTR promoter, in the absence of the Tat protein, is extremely weak, thus rendering the possibility of aberrant transcriptional activation of a neighboring gene less probable. The SIN technology should, in principle, render the probability of cellular gene activation even less likely to occur, due to the removal of the U3 region. Indeed, various experiments are currently ongoing to comparatively assess the mutagenic potential of third-generation, SIN lentiviral vectors with the former lentiviral and gammaretroviral vectors.
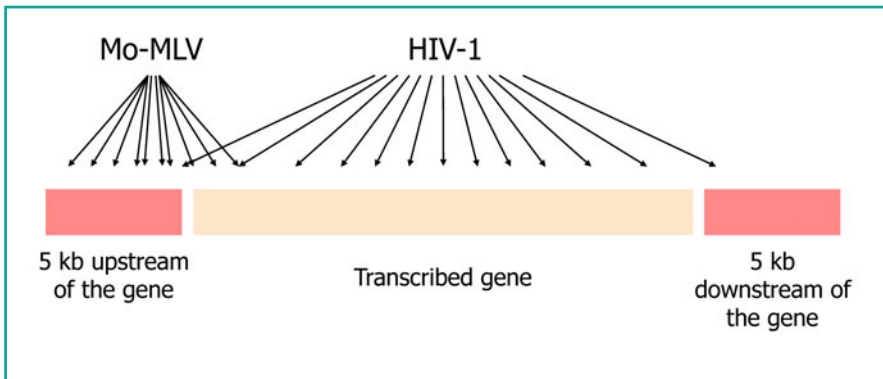
Fig. 3.19 Regions of integration of Mo-MLV and HIV-1 vectors into the cellular genome. Both Mo-MLV and HIV-1 vectors integrate in correspondence with actively transcribed cellular genes. However, Mo-MLV integration occurs in close correspondence with the gene transcription start site, within a few kilobases both upstream or downstream. In contrast, HIV-1 integrates into the whole gene region. Alpharetroviruses, exemplified by the avian sarcoma–leukemia virus (ASLV) group, show fairly random integration, with only weak favoring of transcription units

3.5.3
Vectors Based on Adenoviruses

The first adenovirus was isolated in 1953 from the adenoid tissue (hence the name) recovered during tonsillectomy from a child. Currently, over 100 members of the *Adenoviridae* family are known, able to infect man and various animal species, including non-human primates, mouse, dog, pig, frog, different species of birds, and even some types of snakes. The human adenoviruses are responsible for 5–10% of acute respiratory diseases of children and a variable number of conjunctivitis and gastroenteritis epidemics.

The marked capacity of this virus to infect epithelial cells initially inspired the idea to use adenoviral vectors for gene therapy of diseases of the lung and the airways, typically of cystic fibrosis. However, the natural tropism of adenoviruses for the respiratory epithelium and the conjunctiva is mainly due to its modality of transmission rather than to the molecular characteristics of the virus. Indeed, the receptor mediating cell infection by adenoviruses is ubiquitously expressed and most of the cell types can sustain adenoviral replication independent from the replicative state of the cells, thus opening the way to the possible utilization of these viruses for gene transfer into virtually any organ. Additionally, an intrinsic property of adenoviruses is the great efficiency at which they exploit the cellular machinery to drive synthesis of viral mRNAs and translation of viral proteins: a cell infected with adenovirus produces extremely high levels of viral proteins and thus, in the case of the vectors, of the therapeutic gene they contain. All these properties are of obvious interest for gene therapy; it is thus not surprising that, since the second half of the 1990s, these vectors have been the focus of a vast series of both animal and clinical experimentations.
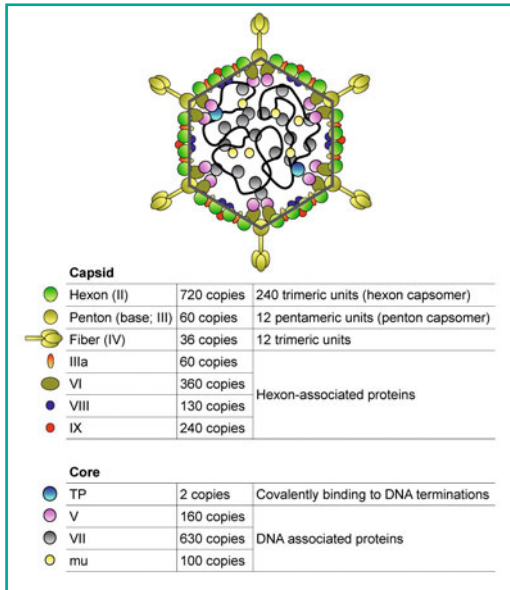
3.5.3.1
Molecular Biology and Replicative Cycle of Adenoviruses

Based on the capacity of different human sera to neutralize adenoviral infection in cell culture, more than 50 serotypes of adenoviruses capable of infecting humans can be distinguished. The neutralizing antibodies mainly recognize epitopes in the exon protein of the virion and the fiber knob (see below). The different serotypes are classified into 6 subgroups (A–F) on the basis of their capacity to determine human red blood cell agglutination; subgroup C includes serotypes 2 and 5 (Ad2 and Ad5), from which most of the gene therapy vectors are derived.

Structure of Virions

The virion consists of a capsid showing icosahedral symmetry, without an envelope, having a diameter of 70–100 nm, and surrounding the viral nucleic acid.

The capsid has 20 facets, each formed by an identical equilateral triangle, 12 vertexes, and 30 edges (T=4). Each facet of the icosahedron is composed of 240 proteins, named hexons since each of them has contacts with 6 other proteins. Each of the 12 vertexes is instead formed by a different protein, named pentons since each of them has contact with another 5

Capsid		
🟢 Hexon (II)	720 copies	240 trimeric units (hexon capsomer)
🟡 Penton (base; III)	60 copies	12 pentameric units (penton capsomer)
⚬ Fiber (IV)	36 copies	12 trimeric units
🔴 IIIa	60 copies	
🫒 VI	360 copies	Hexon-associated proteins
🔵 VIII	130 copies	
🔴 IX	240 copies	

Core		
🔵 TP	2 copies	Covalently binding to DNA terminations
🟣 V	160 copies	
⚫ VII	630 copies	DNA associated proteins
🟡 mu	100 copies	

Fig. 3.20 Schematic representation of an adenoviral virion. The lower part of the figure lists the 11 proteins taking part in the formation of the virion, also indicating their abundance

neighboring proteins. Each penton is formed by a base, which is part of the capsid surface, and a fiber projecting outward, which has different lengths in the various serotypes.

At least 11 different proteins take part in the formation of the virion (Figure 3.20). Each **exon** is formed by 3 subunits of protein II, which is therefore the most abundant protein of the virions; a protein II trimer forming an exon is also called a hexon capsomer. Proteins VI, VIII, and IX are associated with the exon and probably stabilize the interactions among the different protein II monomers and between the exon and the inner proteins of the virion. The **base** of each penton is formed by 5 subunits of protein III, which associate with protein IIIa, while the **fiber** is composed of 3 subunits of protein IV; the combination of the penton base and the fiber is also called a *penton capsomer*.

The *inner part* of the virion contains four different proteins and the viral genome. The terminal protein (TP) is covalently attached to the extremities of the linear genome DNA, while the basic proteins V, VII, and μ (mu) bind the genome and promote its condensation. Additionally, protein V forms a bridge between the virion core and the pentons, thanks to its binding to protein VI. The virion also contains a protease (Pr), encoded by the viral genome, which is necessary for the maturation of some of the structural proteins of the virion and is thus required for proper infectivity.

Genome Organization

The adenoviral genome consists of a double-stranded, linear DNA molecular of 36 kb in the case of Ad2 and Ad5, bearing at the two extremities two identical sequences in reverse orientation (inverted terminal repeats, ITRs; 103 bp in the case of Ad2 and Ad5); these regions act as origins of DNA replication of the entire genome.
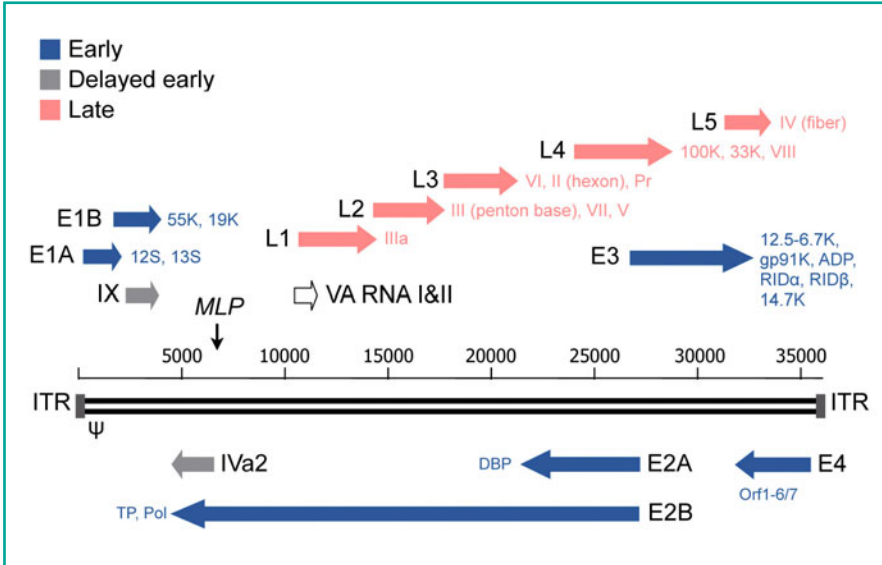
Fig. 3.21 Genome organization of adenovirus. The early genes are shown in *blue*, the delayed early genes in *gray*, and the late genes in *red*. *MLP*: major late promoter. For each gene, the major encoded proteins are indicated. The *arrows* indicate direction of transcription. *ITR*: inverted terminal repeat

The genome contains: (i) five early transcriptional units, which become activated upon cell infection: E1A, E1B, E2 (E2A and E2B), E3, and E4; (ii) two delayed early transcriptional units (IX and IVa2); and (iii) one major late (ML) transcription unit, which is processed to generate 5 families of late mRNAs through post-translational processing (from L1 to L5) (Figure 3.21). All these units are transcribed by RNA polymerase II. The genome also contains 1 or 2 (according to the different serotypes) genes transcribed by RNA polymerase III (virus-associated (VA) RNA genes). Conventionally, the adenovirus genome map is shown with the E1A gene on the "left" side, and thus the E1A, E1B, IX, ML, VA RNA, and E3 genes are transcribed on the "upper" strand and the E4, E2, and IVa2 genes on the "lower" strand. The genomes of all adenoviruses known to date show the same genetic organization.

Functions of the Adenoviral Proteins

E1A is the master gene upon which activation of the whole replicative cycle of adenovirus depends. This gene encodes two proteins (E1A-13S and E1A-12S, generated through alternative splicing of the same mRNA), which exert a variety of functions inside the infected cells, having the ultimate goal to promote viral replication. In particular, E1A binds different cellular proteins controlling the cell cycle, including the tumor suppressor pRb, thus stimulating cell cycle entry; in addition, the protein binds different components of the transcriptional machinery, including the transcriptional coactivators and histone-acetyltransferases p300/CBP and P/CAF, different cellular transcription factors, proteins

of the mediator complex, and the TATA-binding protein TBP, thus stimulating transcription of a series of cellular and the majority of viral genes. The presence of the E1A protein inside the cells also activates p53, since E1A stimulates transcription of the tumor suppressor p19^{ARF}, which binds p53 and modulates its activity.

The E1B gene codes for two proteins, of 55 and 19 kDa. E1B-55K binds p53 and inhibits its transcriptional activity, thus blocking the induction of apoptosis that would be elicited by cell infection with the virus. This property is further discussed in the section on 'Oncolytic Viruses' in the context of 'Gene Therapy of Cancer'. E1B-19K has homology with the cellular gene *bcl-2* and also displays anti-apoptotic activity by binding members of the cellular family of Bax proteins.

The E2 region includes two genes coding for factors necessary for viral DNA replication: E2A, coding for the DNA binding protein DBP, and E2B, coding for the terminal protein TP – which binds the extremities of the linear viral genome – and the viral DNA polymerase.

E3 contains a series of genes that are dispensable for adenoviral replication in cell culture however become essential to overcome the host response to viral infection *in vivo*. The encoded proteins include E3-gp19K, which is localized on the membranes of the ER and prevents transport of MHC class I molecules to the cell surface, where they would allow adenoviral antigen presentation and thus infected cell recognition by cytotoxic T lymphocytes; RIDα, RIDβ, and E3-14.7K, which inhibit TNF-α-, Fas- and TRAIL-induced apoptosis; and adenovirus death protein (ADP), which facilitates cytolysis and thus release of virions from the infected cells.

E4 contains a series of genes the main function of which is to facilitate mRNA processing, stimulate viral DNA replication, and switch cellular transcription off. E4orf6 participates in the formation of the complex between E1B and p53, which inhibits p53 transcriptional activity and targets the protein for degradation. E4orf3 determines localization of the adenoviral DNA replication foci inside the infected cell nucleus.

VA RNA is a genetic region transcribed by RNA polymerase III that generates, in Ad2 and Ad5, two small regulatory RNAs (VA-I and VA-II; 160 nt) that are not translated and have the function to inhibit the cellular protein kinase R (PKR). This enzyme is activated by the double-stranded RNA that accumulates in adenovirus-infected cells and blocks phosphorylation of the cellular translation initiation factor eIF2-α. If not inhibited, PKR would thus inhibit translation of viral and cellular mRNAs and abort infection.

The product of gene IX (pIX) is a multifunctional protein that stabilizes the viral capsid and possesses transcriptional activity. In addition, the protein contributes to the reorganization of the infected cell's nuclear structure; in particular, pIX induces the formation of peculiar nuclear inclusions, where the cellular protein PML becomes localized.

IVa2 codes for a protein involved in the transcriptional activation of the viral major late promoter (MLP), which controls transcription of all the late genes of adenovirus.

The L1–L5 genes code for the viral proteins essential for the late phases of the infection, in particular those taking part in the formation of the virions (cf. below).

The functions of the main regulatory genes of adenovirus are summarized in Table 3.6.

Table 3.6 Functions of the main regulatory genes of adenovirus

Gene	Proteins	Function	Major effects
E1A	E1A 13S E1A 12S	Interacts with proteins regulating the cell cycle (e.g., Rb) and proteins controlling gene expression (p300/CBP, P/CAF, transcription factors, TBP, mediator)	Promotes cell entry into the S-phase; activates transcription of cellular and viral genes
E1B	E1B 19K	Bcl-2 homolog	Inhibits apoptosis
	E1B 55K	Binds and inactivates p53; facilitates the selective transport of viral mRNAs from the nucleus to the cytoplasm	
E2A	DBP	Binds viral DNA	Necessary for viral DNA replication
E2B	TP	Binds the extremities of viral DNA	
	DNA polymerase	Synthetizes of viral DNA	
E3	12.5K-6.7K		
	gp19K	Binds MHC class I molecules and blocks their transport to the cell surface; inhibits MHC class I gene transcription	Inhibits infected cell recognition by the immune system
	11.6K (ADP)	Has cytolytic activity	Promotes release of virions from the infected cells
	10.4K and 14.5K (RIDα and RIDβ)	Inhibits TNF-α, Fas-L and TRAIL-induced apoptosis	Inhibits apoptosis
	14.7K	Inhibits TNF-α and Fas-L-induced apoptosis; stabilizes NF-κB	
E4	E4orf6/7	Modulates activity of E2F transcription factors	Modulates E1A and E1B activities; inhibits apoptosis
	E4orf6 (34K)	Cooperates with E1B 55K	
	E4orf4	Inhibits E1A-induced activation of E2F; binds protein phosphatase PP2A	
	E4orf3 (11K)	Binds E1B 55K; determines localization of nuclear foci of adenoviral DNA replication	
	E4orf2		
	E4orf1	Facilitates cell transformation by E1A and E1B	
VA RNA	-	Inhibits cellular kinase PKR	Permits translation of viral mRNAs
IX	pIX	Structural protein of the capsid and transcriptional activator	Activates late gene expression
IVa2	IVa2	Transcriptional coactivator	Contributes to activation of the major late promoter (MLP)

Replicative Cycle

The replicative cycle of adenovirus is conventionally divided into two phases, separated by viral DNA replication. The early events include binding of the virus to the cell surface (adsorption), penetration of the virus inside the cells, transport of the viral DNA into the nucleus and expression of the early genes, starting with E1A. The early gene products allow further expression of the viral genes, stimulate viral DNA replication, induce cell cycle progression, block apoptosis, and antagonize a series of cellular responses with potential antiviral activity. The early phase lasts about 5–6 h, after which replication of the viral genome starts, concomitant with the late phase of gene expression, leading to transcription of the late genes and to virion assembly. The IVa2 and IX genes are expressed with a timing intermediate between early and late. The replicative cycle takes about 20–24 h in HeLa cells to complete; at the end, each cell has generated about 1×10^4–1×10^5 new infectious viral particles.

The adenoviral replicative cycle can be schematically divided into a series of subsequent steps.

Absorption. Adenoviruses absorb to the cell surface thanks to the interaction of the C-terminal portion of the fiber protein, extending outward like a knob, with a cell surface receptor known as CAR (coxsackie/adenovirus receptor). The CAR protein belongs to the immunoglobulin superfamily and acts as a receptor for the adenovirus subgroups A, C, D, E, and F (but not B) and for B-type Coxsackieviruses, hence the name of the protein.

Internalization. After interaction of the fiber with CAR, virion internalization occurs through receptor-mediated endocytosis mediated by clathrin-coated vesicles. During this process, a fundamental role is played by the interaction of the penton base with the $\alpha v\beta 5$ and $\alpha v\beta 3$ integrins on the cell surface.

Exit from the endosomes and transport to the nucleus. More than 90% of the internalized virions exit from the endocytic vesicles at the level of early endosomes, thanks to the endosomolytic property of the penton base, which is stimulated by the progressive acidification of the endosomes. Of interest for gene therapy, the endosomes are physically destroyed after this process, since exogenous protein–DNA complexes entering into the same endocytic vesicles with adenovirus, although not physically linked to the virions, are also released into the cytosol after endosomolysis. Once in the cytoplasm, the viral particles are transported into the nucleus in an active manner, thanks to the interaction of the exon with the cellular microtubuli. Concomitant with internalization, the virion undergoes progressive disassembly, mediated by the dissociation and proteolytic degradation of its protein components, in particular of protein VI, which functions as a glue between the capsid and the inner components of the virion. A complex consisting of viral DNA, with covalently bound protein TP, and the basic proteins VII, V, and mu, then translocates from the cytosol to the nucleus. Once in the nucleus, protein TP transports the complex in the nuclear matrix compartment, an event essential for efficient viral DNA replication.

Transcription of early genes. Immediately after the adenoviral genome enters the nucleus, the early phase of transcription starts. This phase has three primary objectives: (i) to promote entry of the infected cell into the S-phase of the cell cycle, thus generating a cellular environment optimal for viral replication – this activity is exerted by the products of the E1A, E1B, and E4 genes; (ii) to protect the infected cell from the various antiviral defense

mechanisms at the cell and organism levels – E1A, E3, and VA RNA genes; (iii) to synthesize the viral proteins that are necessary for viral DNA replication – the E2 gene. Achieving all three of these objectives depends on transcriptional activation of the viral genome, which is mediated by the E1A gene product. Thanks to its interaction with numerous cellular factors (cf. above), E1A alone is indeed capable of stimulating entry of the cell into the S-phase. At least three proteins directly inhibit cell apoptosis: E1B-55K and E4orf6 bind and inactivate p53, while E1B-19K is a Bcl-2 homologue. In addition, proteins encoded by E3 block apoptosis by interfering with signaling emanating from the TNF-α receptors and promoting degradation of Fas on the cell surface (this receptor triggers a death signal after its interaction with the Fas ligand (FasL) expressed by cytotoxic T lymphocytes). Other E3 proteins block transport of MHC class I molecules to the cell surface and inhibit their transcription, thus blocking recognition of the infected cells by the immune system. Finally, both E1A and VA RNAs block inhibition exerted by interferon α and β. E1A exerts this activity by binding and inactivating the STAT proteins, which transduce the signal from the interferon receptors on the plasma membrane. The VA RNAs instead bind and block the PKR kinase; interferon induces transcription of this kinase in an inactive form, while its activation is induced by the accumulation, in the infected cells, of double-stranded RNA molecules.

Genome replication. The production of the E2 gene-encoded proteins DNA polymerase, DBP, and TP marks the beginning of the viral DNA replication phase. This begins from the ITRs at the two extremities of the genome and continues in both directions; the process, which is catalyzed by the viral DNA polymerase, requires the covalent binding of TP to the genome ends and involves binding of a series of cellular factors (NF-I, NF-III, and others) to the ITRs. Elongation of the newly synthesized DNA requires the viral protein DBP, which binds DNA, and the cellular factor NF-II. As a rule, the adenoviral genome never integrates into the host cell, a property that also characterizes the vectors derived from the wild-type virus.

Transcription of late genes. At the beginning of the DNA replication phase, transcription of the late genes also starts. These are organized as a single long transcript of about 29,000 nt, which is subsequently processed through the utilization of different polyadenylation and alternative splicing sites to give rise to a series of shorter mRNAs. These can be grouped into 5 families (L1–L5) on the basis of the utilization of 5 different polyadenylation sites. Expression of all these transcripts is controlled by a single, specific promoter, the major late promoter (MLP), which is activated by the cellular transcription factor USF/MLTF and transactivated by E1A. The product of the delayed early gene IVa2 cooperates with USF/MLTF in MLP activation. Once replication of viral DNA starts and the late gene mRNAs are synthesized, the cellular mRNAs are selectively retained in the nucleus, due to the capacity of E1B 55K and E4orf6 to block their export. In addition to cytoplasmic transport, translation of the viral mRNAs is also favored at this stage compared to that of cellular mRNAs. These properties have obvious relevance in considering adenovirus as a vector for high-level expression of therapeutic proteins.

Virion assembly and cell lysis. Translation of the L1–L5 mRNAs leads to the synthesis of the virion structural proteins; packaging of the viral DNA inside the virions is mediated by the recognition of the packaging signal ψ, present at about 260 bp from the left side of the genome and consisting of a series of AT-rich DNA stretches. Release of the virions from the infected cells is accompanied by disintegration of the plasma membrane during cell lysis.

3.5.3.2
Structure of Adenoviral Vectors

Three different classes of vectors based on Ad2 or Ad5 are considered for gene therapy. The first generation consists of viruses in which the therapeutic gene substitutes the E1 and/or E3 regions. The second generation carries additional deletions in the E2 and E4 regions. Finally, the third generation includes the so-called *gutless* or helper-dependent vectors, in which the whole viral genome is substituted with the exception of the ITR and ψ regions (Figure 3.22).
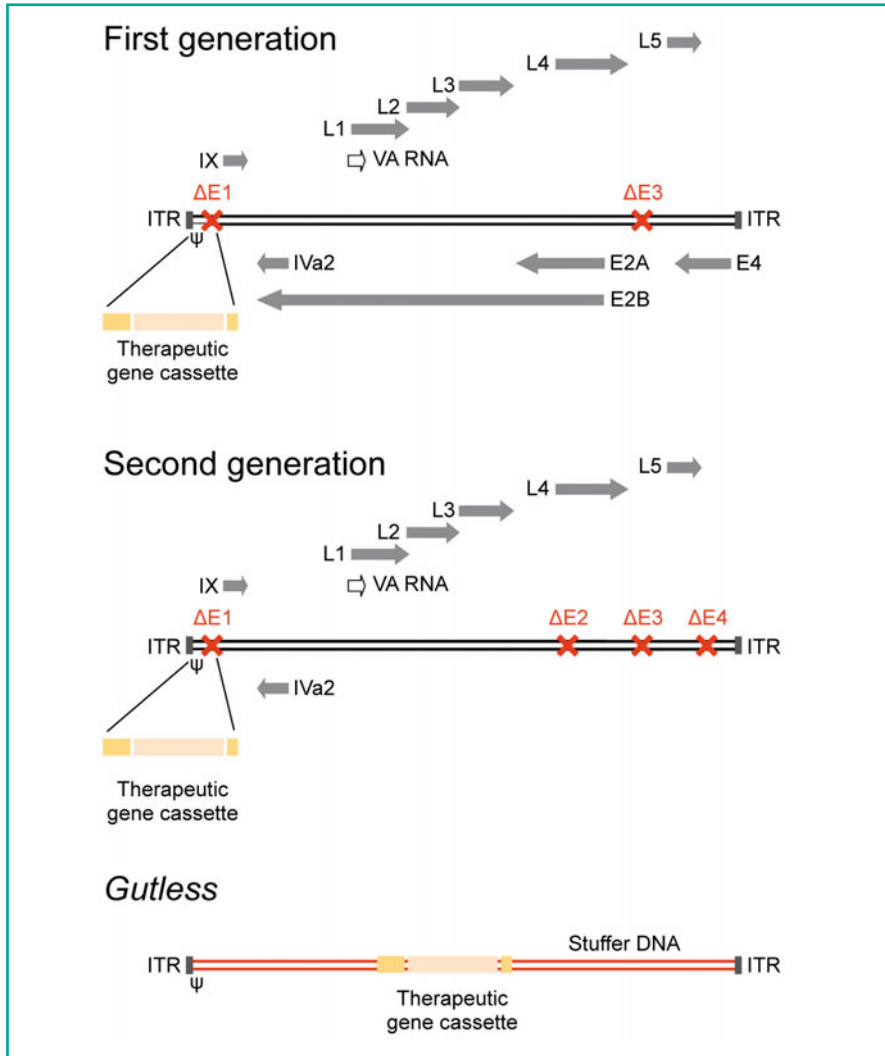
The *first-generation* adenoviral vectors are obtained by substituting the E1, or the E1 and E3, regions with an expression cassette, consisting of the therapeutic gene, a promoter, and a polyadenylation site. As reported above, the E1 region (containing the early genes E1A and E1B) codes for proteins essential for the expression of the other early genes as well as for the late genes of the virus. Since these proteins are required for viral replication, to produce the vector particles, they are supplied in trans by specific cell lines, such as HEK 293, 911, N52.E6, or PER.C6. The E3 region codes for proteins that are important to counteract the host antiviral mechanisms. These products, however, are not required for *in vitro* adenovirus replication, and thus it is not necessary to complement their loss in trans during vector production. However, for some applications, it is desirable to maintain or even increase expression of some of the E3 proteins. For example, E3-11.6K (ADP) facilitates the release of infectious particles from the producing cells, and gp19k reduces the T-cell response against the transduced cells and thus increases persistence of gene expression *in vivo*. Vectors carrying deletions in only E1 can accommodate foreign DNA stretches up to 5.1 kb, while those deleted in E1 and E3 up to 8.3 kb (considering that the maximum length of DNA that can be packaged in Ad2 or Ad5 virions is about 38 kb).

Although the E1-deleted vectors cannot replicate *in vivo*, expression of the several adenoviral genes that are still present stimulates a powerful inflammatory and immune response of the host, which raises important safety concerns, as will be further discussed below. In addition, the immune response limits the duration of therapeutic gene expression driven by these vectors, since the transduced cells are eliminated by cytotoxic T lymphocytes.

Since inflammatory and immune response towards first-generation adenoviral vectors is stimulated by the various vector-encoded proteins, a *second generation* of vectors was obtained, bearing additional deletions in the E2 region – in particular, in the E2A (coding for DBP), E2B (TP), or DNA polymerase genes – or in the whole or vast majority of the E4 region. These vectors can accommodate up to 14 kb of foreign DNA. Despite elimination of these genetic regions, these vectors do not completely solve the issue of adenovirus-induced toxicity, given the immunogenic and inflammatory potential of the residual genes. In addition, expression of the therapeutic gene from these vectors is reduced compared to first-generation vectors, probably because some of the E2 and E4 genes code for proteins that directly or indirectly increase the levels of expression of the virus-encoded genes.

Finally, the *third-generation* adenoviral vectors are characterized by the complete deletion of the adenoviral genome and its substitution with exogenous DNA, with the exception of the regions required in cis for viral DNA replication and packaging (ITRs and ψ respectively). These vectors are named *gutless* or *gutted* or, more appropriately, *helper-dependent* (since their replication entirely depends on the co-infection of the cells in

Fig. 3.22 Schematic representation of first-generation, second-generation, and gutless adenoviral vectors. The genes deleted or otherwise inactivated in each type of vector are indicated, along with the site of insertion of the therapeutic gene cassette

which packaging occurs with a helper vector producing in trans all the required proteins) or *high-capacity* (HC, since they can accommodate up to 37 kb of exogenous DNA, thus also allowing delivery of large DNA sequences or multiple genes).

An additional class of adenovirus-derived vectors that are used for cancer gene therapy are the *oncolytic viruses* (or oncolytic vectors), in which only the E1B-55K gene is deleted, thus exclusively allowing viral replication in p53– cells. The properties of these mutated viruses are described in the section on 'Oncolytic Viruses' in the context of cancer gene therapy.

3.5.3.3
Production of Adenoviral Vectors

Production of adenoviral vectors requires a two-step procedure, entailing first the generation of a vector genomic DNA with the sequence of interest and later its replication and packaging to obtain infectious viral preparations.

As far as vector DNA production is concerned, the relative length of the wild-type adenoviral genome (~36 kb) poses an important obstacle to the use of conventional genetic engineering techniques, essentially based on recombinant DNA manipulation *in vitro* followed by amplification of plasmids in simple microorganisms. Therefore, a series of relatively complex protocols have been set up over recent years for the production of first-generation, second-generation, and gutless adenoviral vectors exploiting various alternative approaches.

The second step, namely the production of infectious vector particles, is usually obtained in human cells, named *helper cells*, supplying al the necessary functions in trans (in particular, the products of the E1 gene) in the case of first- and second-generation vectors. The production of gutless vectors is more complex.

Production of First- and Second-Generation Adenoviral Vectors

The production of first- and second-generation vectors is essentially based on the generation of long molecules of linear DNA corresponding to the desired adenoviral vector genome by recombination. The recombinant genomes can be obtained: (i) by direct recombination in helper cells; (ii) by *in vitro* ligation followed by helper cell transfection; (iii) by recombination in bacteria; and (iv) by *in vitro* ligation followed by plasmid transformation of bacteria.

(i) Methods based on recombination in helper cells
The classic method used to generate adenoviral vectors in which the therapeutic gene substitutes the E1 region takes advantage of the recombination events occurring spontaneously in mammalian cells between two homologous DNA sequences (Figure 3.23A). Cells of an E1-expressing cell line (typically, HEK 293 cells, generated in the 1970s by transformation of human embryonic kidney cells with Ad5 DNA) are transfected with: (a) a DNA molecule corresponding to the majority of the "right" arm of the adenoviral genome, obtained starting from the entire purified adenoviral genome (or from a plasmid containing the adenoviral genome), after digestion with a restriction enzyme cutting in the E1 region, followed by the eventual purification of the "right" fragment; (b) a plasmid containing the "left" arm of the adenoviral genome, starting from the ITR down to the region immediately downstream of the restriction site; in this *shuttle* plasmid, an expression cassette containing the gene of interest substitutes the E1 region. Thanks to the region of homology downstream of the restriction site, which is present in both DNA arms, recombination between the two molecules occurs inside the helper cells; this event, although rare, generates a complete viral genome, which is packaged, followed by the release of viral particles into the cell supernatant. This supernatant is then utilized to infect other cells, which are covered by a layer of agar, in order to block diffusion of the virions and only allow infection of the neighboring cells. In this manner, a lysis
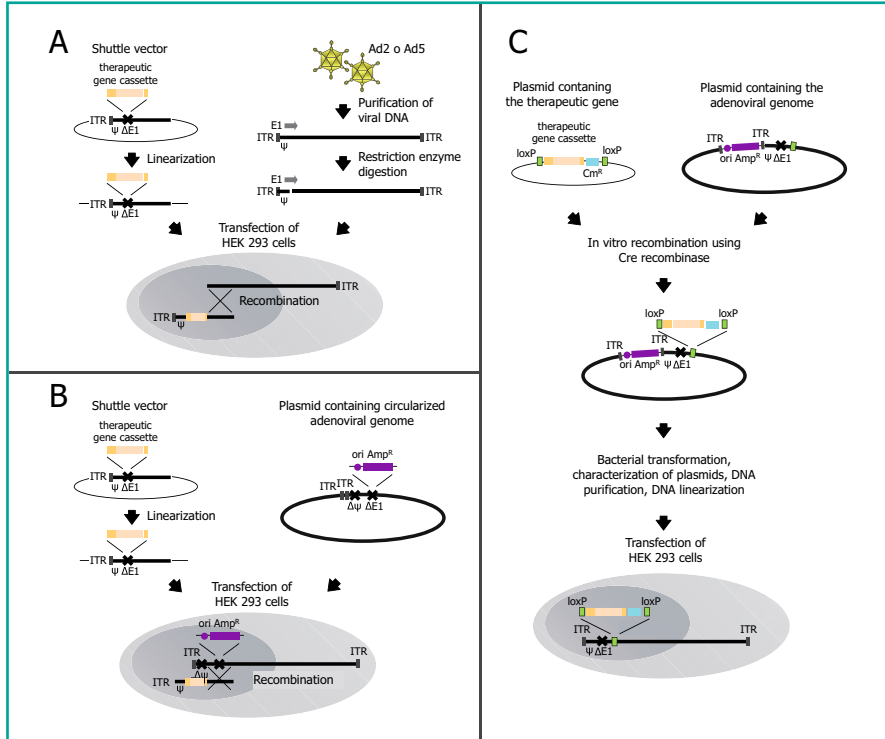
Fig. 3.23 Production of adenoviral vectors. **A** Production of adenoviral vectors by recombination in helper cells transfected with a linearized plasmid carrying the gene of interest and the digested, linear viral DNA. **B** Production of adenoviral vectors by recombination in helper cells transfected with a linearized plasmid carrying the gene of interest and with another plasmid corresponding to the adenoviral genome in which the E1 region is substituted with an antibiotic resistance gene and a prokaryotic DNA replication origin, and in which the packaging signal is deleted. **C** Production of adenoviral vectors by *in vitro* recombination mediated by the Cre recombinase. See text for description

plaque is obtained. Once a plaque containing the desired vector and devoid of parental virus is identified, this is amplified to produce a small quantity of vector, which is then used to infect a larger number of helper cells and produce batch quantities of vector.

The above-described procedure is tedious and labor-intensive (it usually takes 2–4 weeks to complete), since DNA recombination rarely occurs inside mammalian cells. Furthermore, unwanted recombination events from unrestricted, parental DNA molecules also lead to the generation of wild-type virus. Thus, vector production requires multiple subsequent passages of infection and analysis of several lysis plaques.

Different procedures have been developed to improve selection and purification of the desired recombinant vector. For example, as a source of adenoviral DNA, it is possible to use a genome in which a cassette expressing the herpes simplex virus thymidine kinase (HSV-TK) enzyme substitutes the E1 gene. In this manner, cell treatment with gancyclovir (to which cells expressing HSV-TK are sensitive) eliminates all cells infected with a

3

parental rather than a recombinant virus. The same strategy can be followed by using, instead of the HSV-TK gene, the gene coding for *E. coli* β-galactosidase (LacZ) or the *Aequorea victoria* green fluorescent protein (GFP). In the former case, the plaques formed by the parental, unwanted virus are recognized since they become blue after staining with the chromogenic substrate X-gal, while, in the latter case, they are green at the fluorescence microscope. Finally, instead of naked adenoviral DNA, it is possible to use DNA in a complex with the TP protein, a strategy significantly increasing recombination efficiency.

An additional manner to improve recombinant adenoviral production is based on the capacity of circular molecules containing the adenoviral genome to generate infectious virus after HEK 293 cell transfection (Figure 3.23B). These cells are transfected with: (a) a circular viral DNA containing, instead of the E1 region, a prokaryotic origin of DNA replication and the ampicillin resistance gene, and carrying a deletion of the ψ region: this construct can only be propagated and selected in *E. coli* as a plasmid, however cannot be incorporated into virions; (b) a small shuttle plasmid carrying the gene of interest and the "right" arm of the adenovirus genome, similar to the previous method; this plasmid is transfected after linearization. Inside the cells, recombination between the two molecules generates a viral genome that can be packaged. The same technique can be used to also insert expression cassettes into the E3 region, or in both E1 and E3. The main disadvantage of this system is that it requires the use of very large plasmid molecules in *E. coli*, where these are often unstable and potentially deleterious to bacterial growth.

(ii) Methods based on in vitro ligation, followed by helper cell transfection

Instead of transfecting the "right" and "left" arms of the adenoviral vector DNA in the cells relying on their *in vivo* recombination, as in the above-described procedure, it is possible to obtain the desired DNA molecule through canonical ligation *in vitro*. The simplest approach is based on the ligation of the "left" arm of the genome, obtained by restriction enzyme digestion followed by purification, with the "right" arm, recovered from the insert of the shuttle plasmid, as above. The ligation mixture is then transfected into the helper cells to propagate and package the virus. A drawback of this method is that it is particularly subject to contamination from the parental virus, since tiny amounts of uncut parental DNA are inevitably present in the ligation mixture.

(iii) Methods based on recombination in E. coli

Although poorly utilized due to the difficulty of manipulating large molecules in bacteria, some procedures have been developed to generate adenoviral vector genomes by recombination in bacteria. These methods essentially rely on the utilization of two elements: (a) a large plasmid containing the whole adenoviral genome (or at least its "right" arm); (b) a smaller shuttle plasmid containing an expression cassette for the gene of interest flanked by regions of homology to the adenoviral genome region where recombination is sought (i.e., the E1 or E3 regions). In its simplest formulation, the procedure entails transfection of the two plasmids into a recombination-proficient (RecA+) *E. coli* strain, in which the desired molecule is usually produced with an efficiency varying from 20 to 100%. The recombinant plasmid is then transferred to a RecA– strain for propagation, to avoid further recombination events. Once purified, the plasmid DNA is digested to linearize the adenoviral genome and transfected into the helper cells for packaging.

(iv) Methods based on in vitro ligation or recombination, followed by amplification in bacteria

Instead of relying on *in vivo* recombination in helper cells or bacteria, it is possible to obtain the desired recombinant plasmids *in vitro* and then propagate these constructs in *E. coli*. The plasmid DNA is then recovered and transferred into the helper cells for replication and packaging. For the *in vitro* production of the adenoviral vector genomes, it is possible to use large plasmids containing the whole adenoviral genome, previously engineered to contain, in place of the E1 region, recognition sites for rare endonucleases, which can be used for standard cloning of an insert containing the therapeutic gene cassette. Much more efficiently, instead of cloning, this cassette can be inserted into the vector by site-specific *in vitro* recombination (Figure 3.23C). For this purpose, a system based on two plasmids is commonly used: (a) the first plasmid contains the adenoviral DNA circularized in order to contain, in the outside region flanked by the two ITRs, a prokaryotic cassette including an origin of DNA replication and an antibiotic resistance gene; the adenoviral DNA contains an intact ψ region and carries a deletion in E1; downstream of this deletion, a recognition site for a prokaryotic recombinase is inserted; (b) the second plasmid contains the expression cassette for the therapeutic gene and an antibiotic resistance gene different from that contained in the first construct; these sequences are flanked by two recognition sites for the same recombinase targeting the adenoviral genome. The two DNAs are purified from *E. coli*, mixed and incubated together with the purified recombinase: this enzyme, using its target sites on the two molecules, mediates insertion of the fragment containing the therapeutic gene into the adenoviral genome. The reaction product is then used to transfect bacterial cells, which are selected with the antibiotic to which the gene flanking the therapeutic gene confers resistance. After extraction and characterization, this plasmid DNA is then used to transfect helper cells.

The most commonly used recombination procedure is that based on the site-specific recombination system Cre-*loxP*. The P1 bacteriophage produces an enzyme, the Cre recombinase, that recognizes a specific 34 bp sequence, named *loxP* (locus of crossover in P1). When two *loxP* sequences are located far apart, Cre binds both of them and activates their recombination: as a consequence of this process, the DNA segment between the two sequences is removed, only leaving a single *loxP* sequence on site (the P1 bacteriophage uses such a recombination strategy to depolymerize the concatemers that its genome forms during DNA replication in bacteria, since the two *loxP* sites are located at the two extremities of the phage genome). Using the same recognition sequences, Cre is also able to mediate the insertion of a DNA segment flanked by two *loxP* sites (in the case of adenoviral vectors, the DNA segment containing the therapeutic gene) using a third *loxP* site inside the target DNA molecule (in this case, the adenoviral DNA).

The methods based on *in vitro* recombination followed by bacterial transformation are relatively simple and rapid, and nowadays represent the methods of choice to produce adenoviral vectors, also thanks to the commercial availability of kits that facilitate the whole procedure. However, transformation of bacteria with very large molecules can lead to their rearrangement. Therefore, much attention needs to be paid to the careful characterization of the final recombinant adenoviral DNA molecule before helper cell transfection.

3

Production of Gutless Adenoviral Vectors

Gutless adenoviral vectors only possess the ITRs at the two extremities of the genome and the ψ region inside, while the DNA of interest substitutes the rest of the genome (Figure 3.22). Since adenoviral vectors are able to package linear DNA molecules having a length corresponding to 75–105% of the wild-type genome (about 27–37.5 kb), when a gutless vector is considered for the delivery of a cDNA or of a small gene it is usually necessary to complement the insert with irrelevant DNA sequences acting as *stuffer* DNA. As a source of stuffer, DNA of prokaryotic or yeast origin, or, better, DNA sequences derived from large human introns can be used. The gutless vector DNA is cloned as a plasmid, amplified in *E. coli*, purified and linearized to release the segment flanked by the two ITRs, and, finally, transfected into the cells.

Since gutless vectors are completely devoid of viral genes, all the proteins necessary for vector DNA replication must be provided in trans. This can be obtained by co-infection of the cells with a replication-competent adenovirus acting as a helper. In this case however, both the gutless and the wild-type genomes are packaged, causing significant contamination of the final vector preparations. To selectively avoid packaging of the helper virus, different strategies can be followed, including mutation of the ψ region of the helper virus, elimination of ψ during vector production, or use of helper viruses with genomes significantly longer or shorter than those that can be packaged.

The most effective strategy so far developed to avoid helper virus packaging is based on the use of HEK 293 cells previously selected to express the Cre recombinase (293Cre). These cells are transfected with the linearized gutless vector genome and infected with a first-generation adenoviral vector (E1-deleted), in which the ψ region is flanked by two *loxP* sequences. Inside the cells, the Cre recombinase removes ψ from the helper virus genome, thus selectively preventing its incorporation into the virions (Figure 3.24). Using this strategy, the extent of contamination of helper virus in the final vector preparations is in the order of 0.1–10% of the gutless vector.

Notwithstanding the relative efficiency with which the gutless vectors can be produced using the above-described procedure, the residual levels of contamination by the helper virus pose important safety issues in light of clinical application. For this reason, several laboratories are currently trying to improve the system, using different recombinases (for example, the yeast Flp recombinase, which catalyses recombination between the *frt* sites) in addition or as an alternative to Cre-*loxP*, or using helper viruses carrying a mutation in exon protein IX, which is necessary for packaging. In the presence of this mutation, the genome of a gutless vector of optimal length is packaged much more efficiently than that of a helper virus that is significantly shorter or longer than the optimal packaging range.

Purification and Characterization of Adenoviral Vectors

Starting from the helper cell supernatant, the virions corresponding to the adenoviral vectors are purified by three subsequent centrifugation steps, the first of which is a conventional centrifugation to pellet the virus while the last two are run in a cesium chloride gradient, in
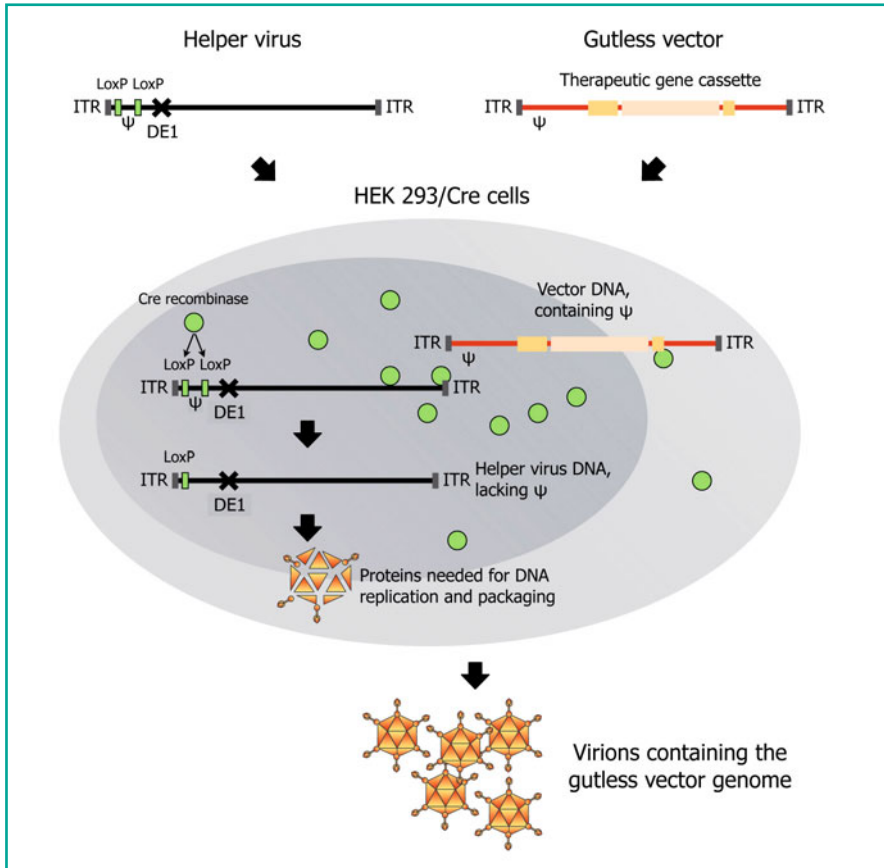
Fig. 3.24 Production of adenovirus gutless vectors. The vectors are produced in 293/Cre cells transfected with the linearized gutless vector DNA and infected with a helper adenoviral vector in which the ψ region is flanked by two *loxP* sequences. Inside the cells, the Cre recombinase removes ψ from the helper virus genome, thus selectively preventing its incorporation into the virions and allowing selective packaging of the gutless vector DNA

which the viral particles are separated according to their specific density, and thus purified.

Before its utilization, an adenoviral vector preparation must be controlled for the possible presence of replication competent adenovirus (RCA). In particular, the HEK 293 cells used by most production protocols contain about 4.5 kb of the Ad5 "left" arm, including the E1 region, integrated into human chromosome 19. This region can thus recombine with the genome of first- and second-generation adenoviral vectors, or with the helper virus in the case of the gutless vectors, thus leading to the formation of RCAs. This becomes progressively more likely should subsequent production steps be carried out, or if very large batches are obtained, since the RCAs replicate more efficiently than the vectors. In case an RCA contamination is detected, for example by PCR, the original vector must be again isolated from lysis plaques generated from the replication of single clones.

3

To lower the probability of RCA formation, helper cells alternative to HEK 293 have been generated, such as 911 or PER.C6, which do not contain stretches of homology between the sequence contained in the vectors and that integrated in the cell genome.

3.5.3.4
Properties of Adenoviral Vectors

Adenoviral vectors are a very efficient tool for gene transfer in mammalian cells, since they infect a vast range of both resting and replicating cells, can be purified and concentrated to reach titers in the order of 1×10^{13} particles/ml, and their genomes do not integrate into the target cells, which might be advantageous for several applications. In addition, the gutless vectors can accommodate large segments of exogenous DNA, up to 35 kb.

As far as first- and second-generation vectors are concerned, these continue to be very interesting for experimental gene transfer in animals. However, their clinical utilization is hampered by the inflammatory and immune response they elicit, which on one hand limits the duration of *in vivo* gene expression and, on the other hand, raises important safety issues. The administration of a first- or second-generation adenoviral vectors, similar to wild-type adenovirus, stimulates both the adaptive and the innate immune response. Immediately after inoculation, expression of a series of inflammatory cytokines is activated, determining recruitment, to the sites of inoculation, of macrophages, neutrophils, and NK cells. For example, in the liver, 80–90% of vector is rapidly eliminated by this inflammatory response within the first 24 h after inoculation. This response is triggered by the adenoviral particle itself and does not require viral gene expression. Subsequently, starting from 4 to 7 days after injection, the humoral and cellular immune response starts to be activated. The inoculation site becomes infiltrated by cytotoxic T lymphocytes, which recognize and eliminate the transduced cells. Furthermore, the immune system mounts a very vigorous antibody response, which, thanks to the production of neutralizing antibodies, prevents any possibility of re-injecting the same vector or vectors based on the same serotype. In this context, it is also important to observe that 30–40% of individuals living in western countries and 80–90% of those living in sub-Saharan Africa naturally possess anti-Ad5 antibodies, which completely prevents utilization of this serotype for gene therapy or vaccination.

The powerful induction of an inflammatory and immune response was the cause of the death of an 18-year-old patient enrolled in a gene therapy clinical trial for the hereditary deficit of ornithine transcarbamylase (OTC), an enzyme of the urea cycle, at the University of Pennsylvania, Philadelphia, PA in 1999. This patient received an injection in the liver, through the hepatic artery, of a second-generation adenoviral vector carrying the OTC cDNA. A few hours after infusion of a relatively high dose of vector, the patient started to show severe symptoms of systemic toxicity, and died after 4 days (cf. section on 'Gene Therapy of Liver Diseases'). Death of this patient was subsequently attributed to a massive, acute inflammatory response to the adenoviral vector injection, probably due to a cytokine storm triggered by the viral capsid.

In light of the above observations, it is thus possible to conclude that the utilization of first- and second-generation adenoviral vectors should now be limited to applications in which prolonged transgene expression is not desirable or required, and in which immune

stimulation is instead a requisite. In practical terms, this is the case in two very important applications: gene therapy of cancer and genetic vaccination.

In the case of gutless vectors, the systemic administration of these viruses continues to stimulate the immune response, similar to first- and second-generation vectors, since this depends on the viral capsid proteins. The same proteins also trigger the production of neutralizing antibodies, which prevent re-administration of vectors of the same serotype. After the initial inflammation, however, the gutless vectors do not express any viral genes, and the transduced cells are therefore not recognized and eliminated by the immune system, unless the transgene protein itself is immunogenic. Experiments performed by transduction of various tissues in rodents, dogs, and non-human primates have indeed indicated that the administration of these vectors, in particular to the liver and skeletal muscle, determines a stable transduction, lasting over time and leading to the expression of the transgenes at therapeutic levels. This appears to be of particular relevance for applications in which other vectors that are efficient in these tissues (in particular, those based on AAV) are instead incapable of delivering very large inserts. This is the case, for example, of Duchenne muscular dystrophy, a disease caused by mutations in the dystrophin gene (cf. section on 'Gene Therapy of Muscular Dystrophies'). The dystrophin cDNA is about 14 kb long and thus unfit for cloning into AAV, but very well suited for gutless adenoviral vectors.

Despite the great potential of gutless vectors, their clinical application is however still limited by two major technical issues. The first one consists in their contamination with a clinically still unacceptable proportion of helper virus; the second one relates to the difficulty of obtaining the large batches of vectors that are needed for clinical use, since the procedures so far developed, which are described above, are unsuitable for scaling up.

Cell Targeting

In concluding the discussion on adenoviral vectors it is important to remember that, over the last 10 years, much effort has been put into direct transduction with these vectors towards specific cell types, a property known as *cell targeting*. As reported above, the Ad5 virions bind the CAR receptor thanks to the C-terminal portion of the fiber protein. Following this interaction, a secondary interaction occurs between the penton base protein (displaying the RGD amino acid sequence) and the target cell $\alpha v\beta 3$ and $\alpha v\beta 5$ integrins, leading to virion internalization by clathrin-mediated endocytosis. In addition, the elongated portion of the fiber establishes contacts with the cell surface HSPGs. It is thus possible to modify the terminal amino acid sequence of the fiber (binding to CAR), the penton base (binding to integrins), and the fiber body (binding to HSPGs) in order to modify the tropism of Ad5 *in vitro* and, possibly, *in vivo*.

The modification mainly considered so far consists in the insertion of peptides at the fiber extremity, after deleting the portion of the protein binding to CAR. The majority of the peptides considered for targeting have been isolated using phage display technology, which allows the selection of short amino acid stretches, expressed on the surface of a filamentous phage, for their property of binding to a specific ligand of interest.

An alternative strategy to modify the adenoviral vector tropism is to use antibodies with double specificity, namely capable of binding and inactivating the CAR-binding

domain of the fiber protein on one side, and binding a different cellular receptor on the other side (for example the c-Erb-2 receptor to target the adenoviral vectors towards the breast cancer cells that express this receptor, or the FGF-2 or VEGF receptors, to target the vectors towards activated endothelial cells such as those of the tumor vasculature).

Despite the encouraging *in vitro* results obtained with both these approaches, there is no compelling evidence, at the moment, that these strategies aimed at re-targeting adenoviral vectors might be easily applied to the clinic *in vivo*, especially in light of the systemic toxicity of first- and second-generation vectors.

3.5.4
Vectors Based on the Adeno-Associated Virus (AAV)

In sharp contrast to retroviruses, adenoviruses, and herpesviruses, before its entry into the gene therapy arena, not many laboratories were interested in the biology of the AAV, most likely because this virus, despite its wide diffusion in nature, has never been associated to any human disease. As a consequence, different aspects of its life cycle, including the molecular determinants regulating its tropism, are still largely unknown. In contrast, however, now AAV represents one of the most appealing vectors for *in vivo* gene therapy and several clinical experimentations have already been conducted or are ongoing with very encouraging results, especially for incurable disorders of tissues incapable of regeneration, such as brain and retina.

3.5.4.1
Molecular Biology and Replicative Cycle of AAV

The *Parvoviridae* family (*parvo-*: Latin for "small") includes a vast series of small viruses with icosahedral symmetry, without envelope, containing a single-stranded DNA genome, which infect numerous species of mammals, including man. The family is divided into two genera, the erythroviruses and the dependoviruses. The human prototype of the former genus is human parvovirus B19, the etiologic agent of the fifth disease or erythema infectiosum, while the murine prototype is the minute virus of mice (MVM). AAV instead belongs to the Dependovirus genus, the members of which, in contrast to the erythroviruses, are incapable of autonomous replication and depend on the superinfection of the cells with another virus to complete their replicative cycle, hence the name. In particular, AAV owes its name to its original isolation as a contaminant of cell cultures infected with adenovirus.

The members of the Dependovirus genus are very diffuse in nature: in primates alone over 100 AAV variants have been discovered to date and new serotypes are continuously being identified (i.e., variants with different antigenic properties, not recognized by the currently available antisera). More than 80% of adults of 20 years or older show an antibody response against AAV, proving that they have encountered the virus, probably in their infancy. Despite their diffusion, none of the dependoviruses has ever been associated with any human disease to date.

Structure of Virions

AAV virions have a capsid with icosahedral symmetry (T=1) with a diameter of 18–25 nm, composed of 60 proteins. These include 3 proteins derived from the same gene (the Cap gene) and differing in their N-terminus: VP1, VP2, and VP3, with a 1:1:18 ratio (i.e., each virion has 3, 3, and 54 VP1, VP2, and VP3 proteins respectively). The capsid includes the viral genome, consisting of a linear single-stranded DNA, having either positive or negative polarity; in any AAV preparation, about half of the virions have a DNA with positive polarity and the rest a DNA with negative polarity.

Over recent years, at least 12 different AAV serotypes have been isolated (AAV1–AAV12) and well characterized antigenically, while over 100 additional genetic variants have been identified by PCR amplification of DNA from cultured cells infected with adenovirus or derived from human and non-human primate tissues. All these viruses share similar structure, size, and genetic organization and only significantly differ in the amino acid composition of the capsid proteins. The sequence homology between these proteins ranges from 55 to 99%, and is the major determinant dictating the use of the receptors for cell internalization. In general terms, all AAVs use receptors that are ubiquitously and abundantly expressed (Table 3.7). The most utilized serotype both experimentally and clinically is AAV2, which binds to cell surface HSPGs; $\alpha v\beta 5$ integrin and the receptors for fibroblast growth factor (FGFR-1) and hepatocyte growth factor (HGFR)

Table 3.7 Receptors for some parvoviruses

Parvovirus	Receptor
AAV1	Sialic acid ($\alpha 2,3$ N-linked and $\alpha 2,6$ N-linked)
AAV2	Heparan sulfate proteoglycans (HSPGs)
	Co-receptors: $\alpha v\beta 5$ integrin, FGFR-1, HFGR
AAV3	Heparan sulfate proteoglycans (HSPGs)
AAV4	Sialic acid ($\alpha 2,3$ O-linked)
AAV5	Sialic acid ($\alpha 2,3$ O-linked and $\alpha 2,3$ N-linked)
	PDGF receptor (PDGFR)
AAV6	Sialic acid ($\alpha 2,3$ N-linked and $\alpha 2,6$ N-linked)
AAV7	Not known
AAV8	Laminin receptor (LamR)
AAV9	Not known (LamR?)
Parvovirus B19	Red blood cells P antigen
CPV (canine parvovirus)	Transferrin receptor
	Sialic acid (N-glycolyl neuraminic acid, NeuGC)
FPV (feline panleukopenia parvovirus)	Transferrin receptor

function as co-receptors in some cells. Similar to AAV2, AAV3 also binds HSPGs. In contrast, AAV1, AAV4, AAV5, and AAV6 interact with sialic acid (N-acetylneuraminic acid, Neu5Ac) residues, linked with various bonds to the cell surface glycans. AAV8 binds a specific cell surface protein, LamR, which exerts several functions in the cells, including that of receptor for extracellular laminin. AAV2 and AAV5 particles enter cultured cells by clathrin-mediated endocytosis and are found in early endosomes immediately after entry. These cellular compartments are trafficked through the cytoplasm and rapidly approach a perinuclear location, where they mature into late endosomes. In contrast to AAV2, AAV5, in addition to the endosomes, can also be found in the trans-Golgi apparatus, indicating differences in endosomal trafficking between serotypes.
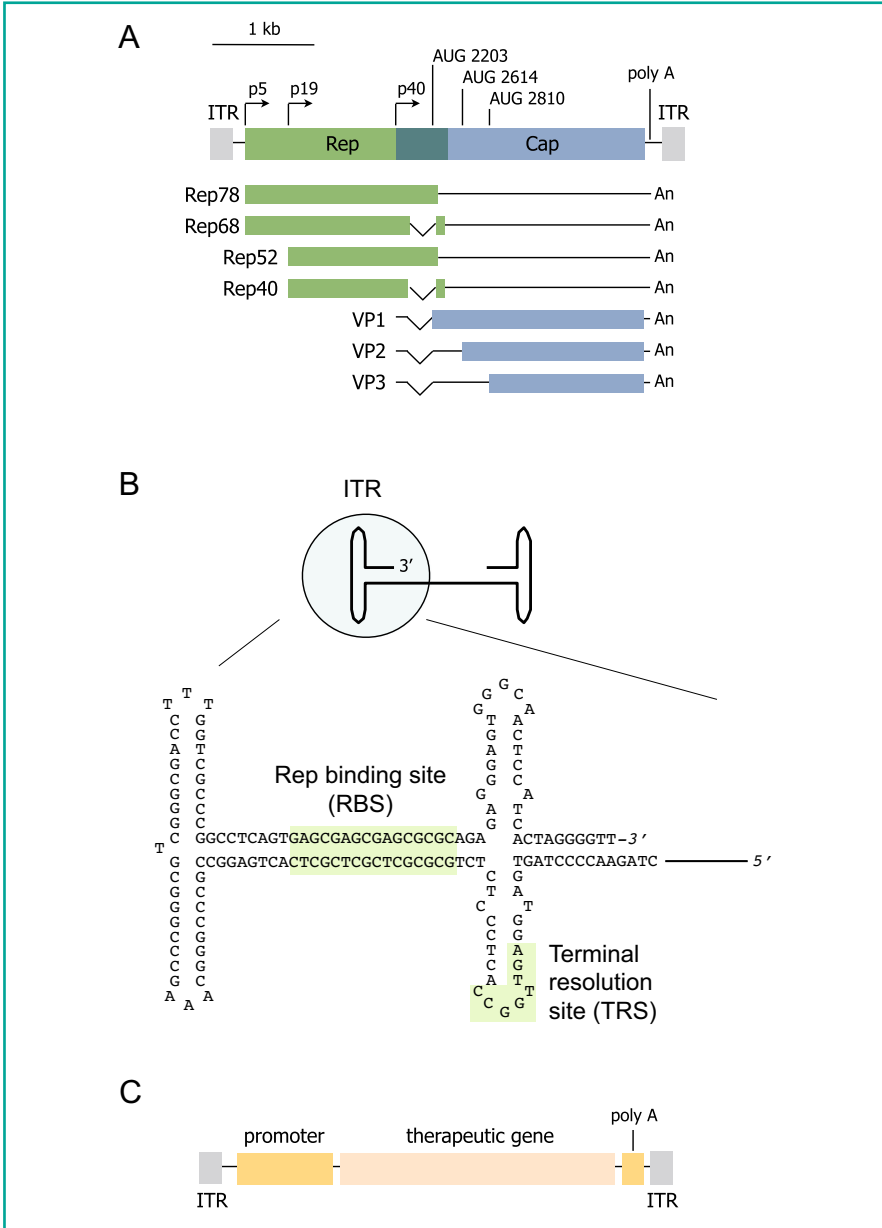
Organization of the Genome

The single-stranded AAV genome has about 4.7 kb and contains two orfs, corresponding to two genes, *rep* and *cap* (Figure 3.25A). *Rep* codes for the proteins necessary for viral replication, and *cap* for the proteins of the viral capsid.

By the use of two different promoters (p5 and p19) and the inclusion or not of an exon, the *rep* gene codes for 4 protein isoforms (Rep78, 68, 52, and 40). The Rep proteins are necessary for replication of the viral DNA, its integration into the host cell genome, and the transcriptional regulation of the viral promoters. They are endowed with single-stranded endonuclease (nickase) and helicase activities. Furthermore, the Rep proteins exert a series of effects on the infected cells, including the inhibition of cellular DNA replication and of transcription of several cellular genes.

The VP1, VP2, and VP3 proteins are generated from the *cap* gene by using three different start sites (AUG codons) for translation. All the AAV transcripts have the same polyadenylation site, located at the 5' end of the genome.

The AAV coding region is flanked by two ~145-nt-long ITRs, having an internal complementarity stretch in their first 125 nt and thus forming a T-shaped hairpin structure, identical at the two viral ends (Figure 3.25B). This palindromic sequence is the only cis-acting genetic element necessary for all AAV functions, including viral DNA replication, site-specific integration into the host cell DNA, and packaging of virions. The first two activities (replication and integration) require the presence of Rep68 or Rep78 proteins, which specifically bind a sequence within the ITR, the Rep binding site (RBS), and cleave in a site- and strand-specific manner at the terminal resolution site (TRS) located 13 nucleotides (nt) upstream of the RBS. An almost identical sequence in human chromosome 19q13.4 represents the minimal sequence necessary and sufficient for AAV site-specific integration – see below. The two ITRs are the only AAV sequences preserved in the vectors, while a transcriptional cassette (promoter+therapeutic gene+polyadenylation site) substitutes the rest of the genome.

A

1 kb

AUG 2203
AUG 2614
AUG 2810
poly A

p5 p19 p40

ITR Rep Cap ITR

Rep78 ————————————— An
Rep68 ————————————— An
Rep52 ————————————— An
Rep40 ————————————— An
VP1 ————————————— An
VP2 ————————————— An
VP3 ————————————— An

B

ITR

3′

Rep binding site
(RBS)

```
      T                           G
   T  T  T                     G G  C
   T     G                     G    A
   C  G  G                     T    A
   C  G  T                     G    C
   A  C  C                     A    T
   G  G  C                     G    C
   G  G  C                     G    C
   G  G  C                     G    A
   T  C  C GGCCTCAGTGAGCGAGCGAGCGCGCAGA    ACTAGGGGTT–3′
T     G     CCGGAGTCACTCGCTCGCTCGCGCGTCT    TGATCCCCAAGATC ———— 5′
   G  G  G                     C    C
   C  G  C                     T    G
   G  G  C                     C    A
   G  G  C                     C    T
   C  G  C                     T    G
   C  G  G                     C    G
   C  G  G                     C    A
   G  G  C                     A    G
   A  A  A                     C  C T
                               G  C G T
                                 G  G
```

Terminal
resolution
site (TRS)

C

promoter therapeutic gene poly A

ITR ITR

Fig. 3.25 Genetic organization of AAV. **A** The upper part shows the genomic organization of AAV, with the indication of the promoters (p5, p19, p40), the AUG codons for translation of the Cap-encoded proteins, and the polyadenylation site. The bottom part shows the structure of the viral mRNAs, indicating the intron–exons organization, and of the encoded proteins. *ITR*: inverted terminal repeat; *An*: polyA tail. **B** Enlargement of the ITR region, with the indication of the Rep binding site (*RBS*) and the terminal resolution site (*TRS*). **C** Schematic representation of the structure of an AAV vector

3

Replicative Cycle

After binding to cell surface receptors, AAV is internalized by receptor-mediated endocytosis and is thus found inside the endosomal compartment. Although the virus can penetrate a vast series of different cell type, thanks to the ability of its capsid to interact with ubiquitously expressed receptors, the fate of infection strictly depends on the physiological state of the infected cells (Figure 3.26). If the cells are exposed to genotoxic stress (e.g., are treated with X-rays, or γ-rays, or other DNA-damaging agents) or are infected with another virus (typically, adenovirus or herpesvirus), the AAV DNA, once exited from the endosomes, is efficiently transferred to the nucleus and replicated by the cellular machinery with the assistance of the viral protein Rep. In particular, viral DNA replication is carried out by a cellular DNA polymerase (probably DNA polymerase δ) using as a primer the exposed 3'-OH from one of the ITRs. Completion of double-stranded DNA synthesis requires the nickase activity of Rep, which cleaves one strand of the ITR and thus permits elongation of DNA synthesis to reach the end of the template molecules (Figure 3.27). At the end of the replication process, two complete viral genomes are generated, with complementary polarity; both of these are packaged inside the virions at equal efficiency. In a few hours,
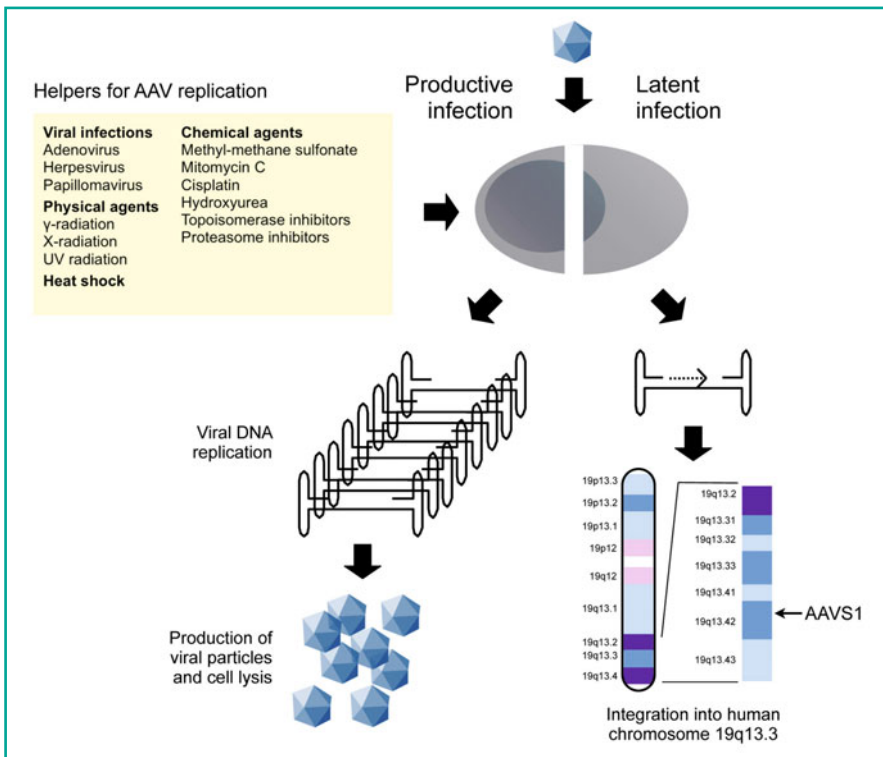
Fig. 3.26 AAV replication cycle. Schematic representation of the AAV replication cycle, under permissive (*left side*) and non-permissive conditions (*right side*). In the latter case, the viral DNA integrates site specifically in the AAVS1 region of chromosome 19q13.3. See text for details

every cell produces 5×10^5–1×10^6 viral particles; the infected cells eventually lyse and the virions are released in the outside environment.

Productive viral infection requires exit of the viral particles from the endosomes, transport to the nucleus, removal of the capsid and release of the nucleic acid, and, most important, conversion of the single-stranded genome into a double-stranded replication intermediate. In physiological conditions, that is in the absence of any treatment with chemical or physical agents or superinfection with another virus, most human cells do not allow productive viral replication, and infection is probably blocked at multiple steps. Under non-permissive conditions, however, in a fraction of the infected cells the AAV genome becomes integrated in a site-specific manner into a specific region of human chromosome 19q13.4, named AAVS1, which contains a 33-bp sequence almost identical to the RBS and TRS sites in the viral ITRs. AAVS1 is positioned immediately upstream of the translation initiation site of the gene coding for the protein phosphatase 1 regulatory inhibitor subunit 12C (PPP1R12C), also known as MBS85 (myosin-bind-
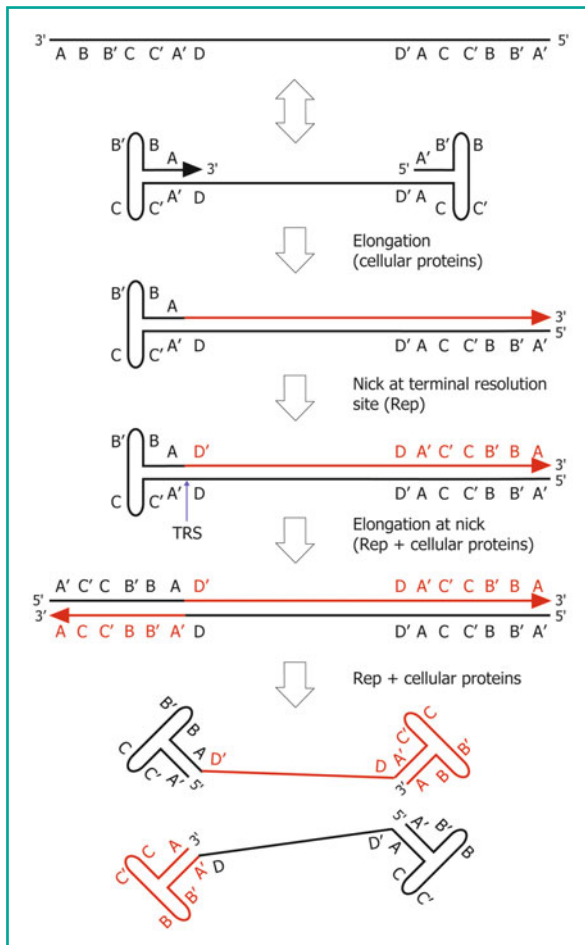
Fig. 3.27 Mechanism for AAV DNA replication. Letters indicate specific sequences in the inverted terminal repeats (A-A', B-B', C-C', D-D' denote complementary sequences). Input viral DNA is in black and neosynthesized DNA in red. The arrow indicates the 3' OH available as a primer for DNA synthesis. *TRS*: terminal resolution site

ing subunit 85), a protein involved in the regulation of actin–myosin fiber assembly. The translation initiation start codon of this gene is located only 17 nt downstream of the chromosomal RBS sequence. The AAVS1 region is located in a centromeric position with respect to the genes coding for the slow skeletal muscle troponin T (TNNT1) and cardiac troponin I (TNNI3), located 15 and 26 kb apart respectively. The capacity of AAV DNA to integrate into the AAVS1 region strictly depends on the availability of the Rep proteins to simultaneously interact with both the ITR and the cellular RBS−TRS motifs, thus mediating a semi-homologous recombination between the two sequences, with the almost certain involvement of still unidentified cellular proteins. Integration of AAV into the human genome is the only event of site-specific integration known to occur in mammalian cells.

It is still largely unclear which molecular determinants govern cell permissivity to productive AAV replication and their relationship with the induction of genotoxic damage or infection of the cells with another virus. These treatments do not directly act on the replication machinery of the virus or on its proteins, but do induce some cellular functions that render the cellular environment permissive for replication. Experimental evidence indicates that this process is controlled by the proteins belonging to the cellular DNA damage repair (DDR) system, namely the cellular machinery that physiologically surveys the integrity of the cellular genome. In particular, proteins of the MRN complex (Mre11, Rad50, and Nbs1) bind the AAV genome which, similar to damaged cellular DNA, is single-stranded and bears imperfectly paired DNA sequences at the level of its terminal hairpins. These proteins block replication of the genome by impeding its conversion to a double-stranded form. Once the cell is treated with chemical or physical agents, the DDR proteins are recruited to other sites of cellular DNA damage, thus permitting the AAV genome to complete its replication. In the case of adenovirus, the helper effect exerted by this virus is mediated by a few known viral genes, which are encoded by the early regions E1A/E1B, E2A, and E4 (in particular, E4or6) and by the gene coding for VA-I. The E1B and E4orf6 proteins are indeed able to induce degradation of the cellular MRN complex.

3.5.4.2
Structure and Production of AAV Vectors

The AAV genome can be converted into a double-stranded DNA form and cloned into a bacterial plasmid. Once transfected into mammalian cells, thanks to its site-specific nickase activity, the Rep protein is able to excide the AAV sequence from the plasmid and initiate its replication. This process only requires integrity of the ITRs in cis and the presence of Rep in trans. If the cells also express the capsid proteins, the single-stranded DNA genome that is formed by the process of DNA replication becomes packaged into the capsids thanks to the interaction of the ITRs with the VP1-3 proteins.

AAV vectors are usually obtained starting from the AAV2 genome, cloned in a plasmid form, by removing all the viral sequences with the exception of the two ITRs (about 145 nt each). Between the ITRs, an expression cassette is cloned containing the therapeutic gene and its regulatory elements (Figure 3.25C). In contrast to retroviruses, the replicative cycle of

which involves the generation of both RNA and DNA genomic forms and in which the choice of the promoter is thus critical to determine the vector efficiency, AAV replication only involves DNA intermediates. As a consequence, any promoter can be chosen to direct expression of the therapeutic gene, without interfering with the production of vectors. This includes strong constitutive, inducible, or tissue-specific promoters. The only strict requirement is that the transcriptional cassette cloned between the two ITRs does not exceed 4–4.5 kb.

The expression of Rep proteins exerts toxic effects on the cells, since these proteins interfere with several essential processes, including DNA replication and cellular gene transcription. Therefore, it has not been possible to obtain packaging cell lines stably expressing Rep, and AAV packaging thus occurs upon transient transfection. This is usually achieved by transfecting, using calcium phosphate co-precipitation, HEK 293 cells with one plasmid containing the AAV vector, as described above, and one plasmid containing the AAV *rep* and *cap* genes without the ITRs. To stimulate the induction of cell permissivity to productive AAV replication, the cells are also infected with adenovirus or, more conveniently, treated with a third plasmid bearing the adenovirus helper genes E2A, E4, and VA-I RNA; the E1A and E1B genes are already expressed in the HEK 293 cells. Several laboratories now exploit a single helper plasmid, containing both the AAV2 *rep* and *cap* genes and the adenoviral helper genes; in this case, the production of vectors involves cell transfection with only two plasmids (Figure 3.28).

Forty-eight hours after transfection, the cells start to show a clear cytopathic effect, due to viral replication, and a large quantity of virions is found in both the supernatant and the cell lysates. In contrast to retroviral vectors and similar to adenoviruses, the AAV genomes are very resistant to manipulation and treatment with chemical and physical agents. Thus, they can be easily purified by cesium chloride or iodixanol gradient centrifugation, or by chromatography. The viral preparations obtained using these procedures are sufficiently pure to be used in both experimental animals and in the clinics. The titers can reach or surpass 1×10^{14} viral particles/ml; the concentration of viral particles is thus several orders of magnitude higher than both VSV-G-pseudotyped retroviral vectors and adenoviral vectors.

A standard AAV production protocol entails utilization of the AAV2 ITRs in conjunction with the AAV2 *rep* and *cap* genes. However, the capsid proteins corresponding to any AAV serotype can recognize the AAV2 ITRs and mediate packaging of the AAV2 genome inside the virions. It is thus possible to change the serotype of the vector simply by using, during production, an expression vector for any desired *cap* gene. Vectors with a capsid corresponding to the AAV1–AAV9 serotypes are commonly generated to exploit the different organ tropism of these viruses (cf. below).

3.5.4.3
Properties of AAV Vectors

AAV vectors represent an outstanding tool for *in vivo* gene transfer, for a series of reasons, which are summarized as follows.

(i) AAV vectors do not express any viral protein; therefore, they are not immunogenic and do not cause inflammation (in contrast to first- and second-generation adenoviral vec-

3
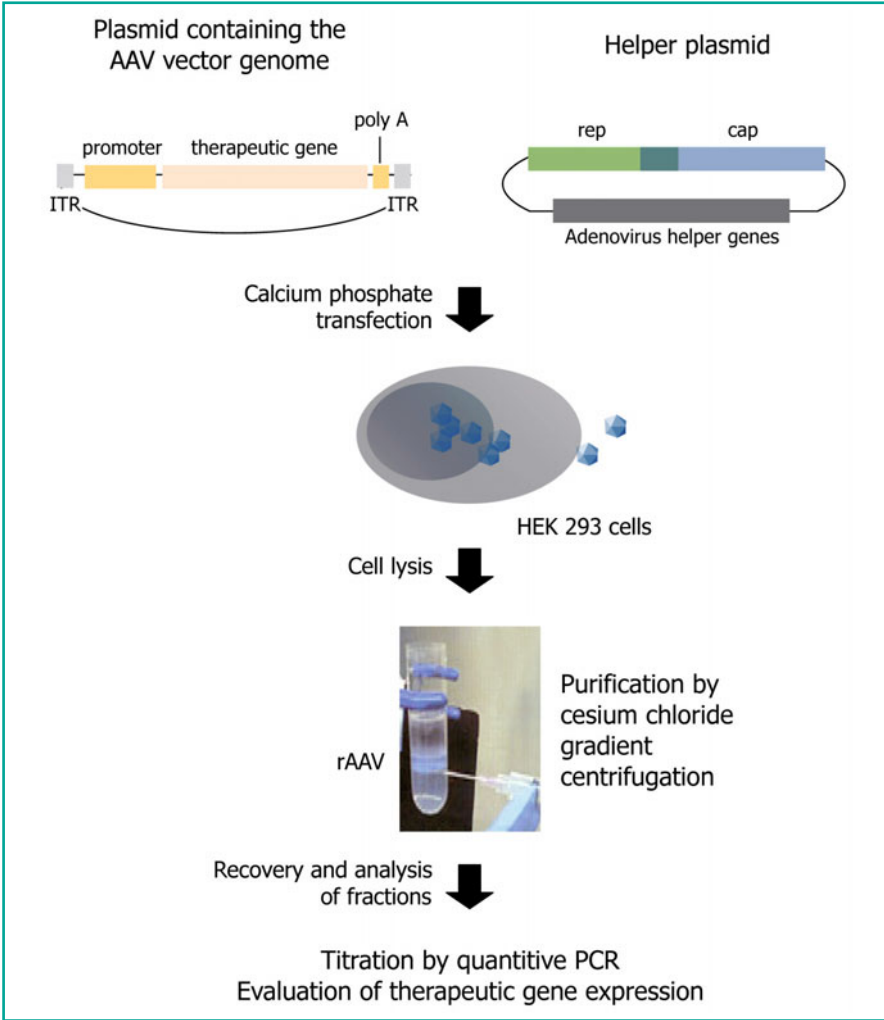
Fig. 3.28 Production of AAV vectors. For AAV vector production, two plasmids are transiently transfected into HEK 293 cells. The first plasmid corresponds to the AAV vector itself, in which the therapeutic gene cassette is flanked by the inverted terminal repeats (ITRs), and the second codes for Rep and Cap and for the adenoviral proteins providing helper functions. Twenty-four hours after transfection, cells are lysed and the vectors are purified by cesium chloride centrifugation

tors); as a consequence, therapeutic gene expression usually lasts for month- or year-long periods. In this respect, it is however worth mentioning that, while this conclusion certainly holds true in mice, pigs, and dogs, which are not the natural hosts of human AAVs, from which most of the vectors are currently derived, it might be different in humans and non-human primates, where the pre-existing immunity against the virus in some cases might determine the elimination of the transduced cells over the first weeks post-inoculation (cf. also the section on 'Gene Therapy of Hemophilia').

(ii) AAV vectors do not integrate into the host cell genome, but persist in an episomal form, probably as head-to-tail or head-to-head extrachromosomal concatemers, in non-replicating cells; therefore, they avoid the problem of insertional mutagenesis (in contrast to retroviral vectors). Of notice, the ability of the wild-type virus to integrate site specifically into the AAVS1 region of human chromosome 19q13.4 strictly requires the AAV Rep protein; since the gene coding for this protein is not present in the vectors, site-specific integration does not occur.

(iii) Possibly as a consequence of the lack of integration into the transduced cell chromosomes, therapeutic gene expression is not subject to significant methylation and silencing (in contrast to retroviral vectors).

(iv) AAV vectors can be generated at high titers, thus allowing the simultaneous expression of different genes from the same cells or tissues. This property could be of great importance in light of the possibility to deliver multiple growth factor coding genes, for example for gene therapy of cardiovascular or neurodegenerative disorders, or for the administration of multiple shRNAs to inhibit different proteins acting along the same metabolic pathway.

(v) AAV vectors do not experience the problem of transcriptional interference from different promoters (in contrast with retroviral vectors). The therapeutic gene can thus be controlled by any promoter of choice, provided that its length is suitable for cloning into AAV.

Taken together, these properties have encouraged, over the last 5 years, the use of these vectors in over 50 clinical trials, which have enrolled several hundred patients. These are Phase I/II trials for various hereditary (in particular hemophilia B, deficit of α1-antitripsin, cystic fibrosis, muscular dystrophies, retinal degeneration) and acquired (rheumatoid arthritis, Parkinson's disease, Alzheimer's disease) disorders. The preliminary results of these trials will be detailed in the respective sections in the chapter on 'Clinical Applications of Gene Therapy'.

Despite the successful utilization of AAV vectors for clinical gene transfer, several of their molecular properties remain poorly understood. Most AAV serotypes use ubiquitous molecules that are expressed at high levels by most cell types as receptors for internalization, such as HSPGs or sialic acid linked to cell surface glycoproteins and gangliosides (Table 3.7); as a consequence, internalization of these vectors occurs in most cells. However, in the majority of cell types, the vector DNA does not reach the nucleus or, most frequently, is not converted from its single-stranded DNA form to double-stranded DNA, a step that is obviously essential for transcription to occur. *In vivo*, only a few tissues show high-level natural permissivity to AAV transduction. These include the heart (cardiomyocytes), skeletal muscle (skeletal myofibers), brain (neurons), retina (ganglionar cells, pigment epithelium and photoreceptors), and, to a lesser extent, liver (hepatocytes). The reasons why the virus is particularly efficient in these cell types are still largely unknown, but clearly involve molecular events following vector internalization inside the cells. For example, after injection of AAV2 in the brain, the vector transduces neurons very efficiently but does not transduce glial cells at all, even though these cells express much higher levels of HSPGs on their surface. Of note, all the cell types that are naturally permissive to AAV transduction are post-mitotic and will never re-enter the cell cycle; it is thus likely that the DDR proteins that bind single-stranded DNA and block AAV transduction in replicating cells are downregulated in these permissive cell types. Indirect evidence of the impor-

tance of single-stranded to double-stranded DNA conversion in determining efficiency of transduction is the significantly higher transduction efficiency of the so-called *self-complementary* AAV vectors (scAAVs), in which the gene cassette is cloned in the form of two complementary copies, positioned in tandem one after the other; the DNA of these vectors is thus capable of spontaneously forming double-stranded DNA by internal self-complementation. Due to their intrinsic design, however, the scAAV vectors have a cloning capacity about half of the already constrained limit of normal AAV vectors.

The use of capsids with serotypes different from AAV2 on one hand increases efficiency of transduction in the already permissive cell types but, on the other hand, extends tropism to a few other organs (Table 3.8). For example, skeletal muscle is transduced with particular efficiency by AAV1 and AAV6 (which differs only 6 amino acids from AAV1); in the retina, photoreceptors are an efficient target of AAV5 and the pigment epithelium of AAV5 and AAV4; finally, AAV8 transduces both the endocrine and exocrine pancreas very well, in addition to the liver. None of the serotypes, however, permits significant transduction of cells physiologically refractory to AAV2 gene transfer, including endothelial cells, fibroblasts, keratinocytes, and several others.

Another very interesting property of some of the most recent AAV serotypes, AAV8 and AAV9, is their capacity to cross the endothelial barrier of blood vessels. Once injected intravenously or intraperitoneally in the experimental animal, these vectors reach the skeletal muscle parenchyma and transduce myofibers highly efficiently. They thus represent potential tools for whole muscle transduction for gene therapy of muscle dystrophies. A molecular reason for the different cell-type selectivity of the various AAV serotypes is still to be found. It is conceivable that the use of different entry molecules routes the vectors towards different pathways, or that the various capsid proteins modulate the interaction of the viral genome with cellular proteins differently.

Table 3.8 Tropism of AAV serotypes for different organs

Organ		Serotype (in order of efficiency)
Liver		AAV8, AAV9, AAV5
Skeletal muscle		AAV1, AAV7, AAV6, AAV8, AAV9, AAV2, AAV3
Central nervous system		AAV5, AAV1, AAV4, AAV2
Eye	Pigment epithelium	AAV5, AAV4, AAV1, AAV6
	Photoreceptors	AAV5
Lung		AAV5, AAV9
Heart		AAV9, AAV8
Pancreas		AAV8
Kidney		AAV2

3.5.5
Vectors Based on the Herpes Simplex Virus (HSV)

Different aspects of the biology of herpes simplex virus type 1 (HSV-1) suggest the potential value of this virus as a gene therapy vector. These include: (i) the ample cellular host range, since most of the receptors the virus exploits for internalization are widely expressed by mammalian cells, including HSPGs and nectin-1; (ii) its high natural infectivity; (iii) its capacity to efficiently infect non-replicating cells; (iv) the dispensability of over half of the 80 natural HSV-1 genes for virus replication in cultured cells – these genes can thus be removed and replaced by the gene(s) of interest; (v) the ability of the virus to establish latent infection lasting for very prolonged periods of time in neurons, a property that can be exploited to selectively express therapeutic genes in these cells; and finally (vi) the possibility to produce vectors at high titers in the absence of contamination by wild-type virus. On the other hand, however, the relative complexity of the viral genome and our still incomplete knowledge of the molecular properties of various viral proteins still hamper wider utilization of this vector system.

3.5.5.1
Molecular Biology and Replicative Cycle of Herpesviruses

The *Herpesviridae* family consists of over 130 different viruses, very diffuse in most animal species, 9 of which infect humans. These are HSV-1 and HSV-2, cytomegalovirus (CMV), varicella-zoster virus (VZV), Epstein-Barr virus (EBV), and human herpesviruses types 6A, 6B, 7, and 8 (HHV-6A, HHV-6B, HHV-7, and HHV-8).

All members of the family share at least three common biological characteristics: (i) the presence, in the viral genome, of a vast series of genes coding for enzymes involved in nucleic acid metabolism, including thymidine kinase (TK), which is used in several gene therapy applications, especially as a suicide gene (cf. section on 'Clinical Applications of Gene Therapy'); (ii) the nuclear localization of the sites of genome replication and capsid assembly; and (iii) the capacity to establish two modalities of infection, one leading to production of new viral particles and eventual lysis of the infected cells, and the other one in which the viral genome is maintained in a circular, double-stranded DNA form in the nucleus of the infected cells, with only a minority of the viral genes being expressed.

The various members of the family can be divided into one of three sub-families: the alpha-herpesviruses, characterized by broad host range, rapid replicative cycle, and capacity to establish latent infections mainly in the sensory ganglia (this subfamily includes, among the human herpesviruses, HSV-1, HSV-2, and VZV); the beta-herpesviruses, which have a more restricted host range, longer replicative cycle, and capacity to establish latent infections in the salivary glands, the lymphoreticular system, and the kidney (this subfamily includes CMV, HHV-6, and HHV-8); and the gamma-herpesviruses, which prevalently show tropism for lymphoid cells (this subfamily includes EBV and HHV-8).

Structure of Virions

All *Herpesviridae* are characterized by a common virion structure, consisting of a central core containing the linear genomic DNA, an icosahedral capsid consisting of 162 capsomers, an apparently amorphous structure surrounding the capsid named *tegument* and containing about 20 different types of proteins, and an envelope derived from the plasma membrane during budding; the envelope displays a series of glycoproteins of viral derivation (Figure 3.29A). In the case of HSV-1, at least 11 different glycoproteins (gB-gN) are present on the virion envelope, for a total of over 1000 copies. The virion has an overall diameter varying from 120 to 300 nm.
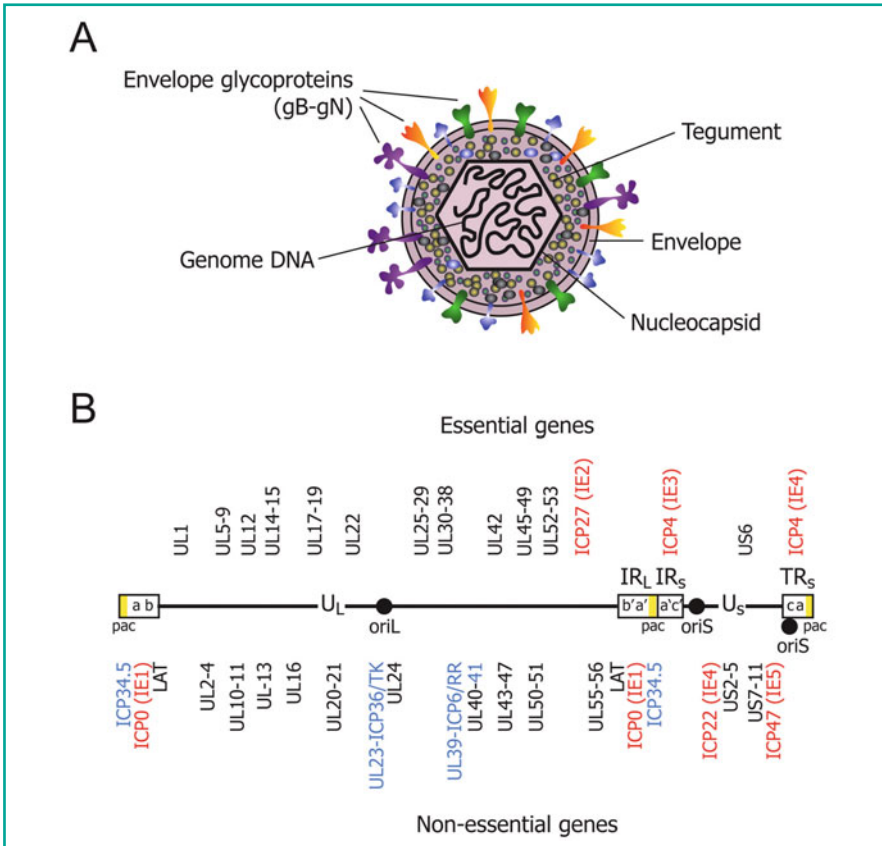
Genome Organization

The genome of herpesviruses consists of a large molecule of double-stranded linear DNA (120–250 kb). In particular, the HSV-1 genome has 152 kb and encodes over 80 proteins. Approximately half of these genes are essential for viral replication in cell culture; the other half encode accessory functions, which contribute to the virus life cycle in specific tissues (e.g., post-mitotic neurons) and can be removed without significantly affecting the capacity of the virus to replicate in cell culture. The genome is composed of unique long (U_L) and unique short (U_S) segments which are both flanked by inverted repeats (Figure 3.29B). The HSV-1 genes fall into one of three classes depending on the kinetics of their transcription in the viral replication cycle: *immediate early* (IE), *early* (E), or *late* (L) genes. The IE genes code for regulatory proteins, the E genes for factors necessary for viral replication, and the L genes for the structural proteins of the virions. Replication of the genome occurs through a rolling circle mechanism leading to formation of head-to-tail concatemers.

Replicative Cycle of HSV-1

Cellular infection starts with binding of the virus to the cell surface glycosaminoglycans (GAGs), in particular with heparan sulfate and dermatan sulfate GAGs. This initial interaction, which is mediated by the C and B glycoproteins (gC and gB) of the viral envelope, is then followed by a more specific binding between gD and some membrane receptors, including HveA (herpes virus entry A, also known as HveM, a member of the tumor necrosis factor (TNF) receptors) and HveC (also known as nectin-1, a transmembrane protein of the immunoglobulin superfamily expressed at high levels in sensory neurons). These interactions lead to fusion of the viral envelope with the cell membrane, followed by entry of the viral capsid into the cytosol together with the tegument proteins. Entry is mediated, by not completely understood mechanisms, by the gB, gH/gL, and gD viral glycoproteins.

Once in the cytosol, the viral DNA is transported to the nucleus through the nuclear pores and a productive, lytic cycle starts; the whole process of viral replication, from transcription to packaging, takes places in the nucleus. The viral genes are expressed in a temporally regulated manner: immediately after entry of the viral DNA into the nucleus, and in the absence of *de novo* synthesis of viral proteins, five IE genes are transcribed (infect-

Fig. 3.29 HSV-1 and HSV-1 vectors. **A** Schematic representation of the structure of an HSV-1 virion. **B** HSV-1 genome organization. The HSV-1 genome consists of a linear, double-stranded DNA molecule of 152 kb containing more than 80 genes. The genome is composed of unique long (U_L) and unique short (U_S) segments which are flanked by inverted repeats. These are designated as TR_L and IR_L (terminal and internal repeat of the long segment, respectively) and TR_S and IR_S (terminal and internal repeat of the short segment). The repeats surrounding U_L are designated ab and b'a', while those surrounding U_S are designated a'c' and ca. There are two different origins of replication, oriL in the long segment and oriS in the short segment. OriS is duplicated, along with ICP4, because it is found in the inverted repeats surrounding the long segment. Approximately half of the genes are essential for viral replication in cell culture (*listed on top*); the other half are non-essential for viral replication in cultured cells (*bottom*). Genes in blue are non-essential genes that are mutated in the replication-competent viruses so far developed and described in the text; genes in red are immediate early (IE) genes that are mutated in the replication-defective viruses. The genome contains three pac signals (shown in *yellow*) that assist in packaging the viral genome DNA into virions

ed cell protein-0 (ICP0), ICP4, ICP22, ICP27, and ICP47). The proteins encoded by these genes activate expression of the E genes, coding for factors necessary for viral DNA replication. Once DNA replication is complete, the IE proteins activate transcription of the L genes. Expression of the IE genes is increased by VP16, a structural protein present in the

teguments, which acts in concert with several cellular transcription factors that bind the IE gene promoters. The L gene products include the structural proteins of the virus, which allow packaging of new viral particles and completion of the lytic cycle.

During primary infection, HSV-1 initially replicates in the epithelial cells close to the site of exposure. The virus then enters the sensitive nervous terminations and the capsid is convoyed by retrograde axonal transport along the axon cytoskeleton to the nuclear body of neurons in the sensitive ganglia. Once entered the nucleus of these cells, latent infection ensues. This is characterized by the presence of the viral genome in the form of circular, double-stranded DNA molecules, or as concatemerized multimers, which persist in the nucleus episomally (that is, not integrated inside the host cell genome). In this latent state, all genes proper of the lytic phase are transcriptionally silent; expression is active only for a single family of non-polyadenylated transcripts named *latency-associated transcripts* (LATs), which remain localized in the nucleus. The exact function of LATs, which have a structure similar to *lariats* (i.e., the RNA products generated by intron processing during splicing), is not known; however, their presence, which can persist for the whole life of the host, can be used as a marker of latent herpesviral infection.

3.5.5.2
Structure and Production of HSV-1 Vectors

One of the major limitations imposed by HSV-1-derived vectors originates from the high pathogenicity of the wild-type virus; for example, the intracerebral injection of HSV-1 typically causes lethal encephalitis. Removal of all the pathogenic genes from the vectors is therefore imperative. Three strategies are currently considered for this purpose:

(i) removal of all genes dispensable for *in vitro* replication however essential for *in vivo* replication; this generates vectors that are still capable of replicating, however with attenuated virulence (**attenuated, replication-competent vectors**);

(ii) removal of all genes necessary for all types of replication (**replication-defective vectors**);

(iii) deletion of the entire viral genome with the exception of an origin of DNA replication and the packaging signal (**amplicon vectors**).

The characteristics and modalities for construction for these three categories of vectors are described in the following sections.

Attenuated, Replication-Competent Vectors

Limited capacity of replication of an HSV-1 vector *in vivo* can be useful as a means to transfer the transgene to cells neighboring those originally transduced, thus leading to amplification of the therapeutic efficacy. Deletion of some non-essential genes allows the generation of HSV-1 mutants that are still capable of replicating *in vitro* but severely impaired *in vivo*. These genes include those coding for proteins necessary for DNA replication, or mediating virulence, or conferring the infected cell the capacity to evade immune

recognition (Figure 3.29B). Examples of these proteins are TK and ribonucleotide reductase (RR), two enzymes dispensable in cell culture but essential for viral DNA replication in neurons, where cellular DNA replication proteins are no longer expressed; the *vhs* (*virion-host shut off*), the product of the UL41 gene, which rapidly destabilizes and blocks translation of the infected cell mRNAs; and the neurovirulence factor ICP34.5, which allows continuation of translation in the infected cells notwithstanding the activation of the cell kinase PKR, which would otherwise phosphorylate and inactivate translation initiation factor eIF2α as an antiviral defense mechanism. Different studies have revealed that the attenuated, replication-competent HSV-1 vectors are not only capable of replicating autonomously but, when inoculated into the brain, also circulate to areas different from those of the original injection, similar to the wild-type virus.

Replication-Defective Vectors

Wild-type HSV-1 replication in neurons begins immediately after entry of viral DNA into the infected cell nucleus. An essential role in this process is played by the IE genes, which activate transcription first of the E genes, involved in DNA replication, and later, once DNA replication is complete, of the L genes, coding for the structural proteins of the virions. Deletion of the IE genes generates mutants that only replicate in cells in which the missing functions are supplied in trans. Once inoculated *in vivo*, these defective viruses are not able to activate the cascade of events leading to lytic infection and thus remain in the cells in a state that is similar to viral latency, persisting for prolonged periods in both neurons and non-neuronal cells *in vivo*.

The first generation of replication-defective HSV-1-based vectors consisted of mutants deleted in the single essential IE gene encoding ICP4. Although these vectors show reduced pathogenicity and could be used to efficiently transfer and transiently express reporter genes in brain, they were nonetheless cytotoxic for neurons in culture. Further improvements involved the introduction of deletion in additional genes, first in ICP27 and later in various combinations of other IE genes (Figure 3.29B). Besides prolonged persistence *in vivo*, an additional advantage of these multiply deleted viruses is their capacity to provide enough space to introduce distinct and independently regulated expression cassettes for different transgenes.

Amplicon Vectors

Amplicon vectors are viral particles identical to wild-type HSV-1 virions, in which the genome is however made of a concatemeric form of a plasmid, the amplicon. This consists of a conventional *E. coli* plasmid carrying an origin of DNA replication (usually ori-S) and a packaging signal (pac), both derived from the HSV-1 genome (Figure 3.30). The remaining portions of the amplicon contain the transgenic sequence(s) of interest; given the capacity of HSV-1 virions to package long DNA molecules, these can extend to over 150 kb. This represents the largest cloning capacity of all currently available gene transfer systems.

Both replication-defective and amplicon vectors are packaged into complete HSV-1
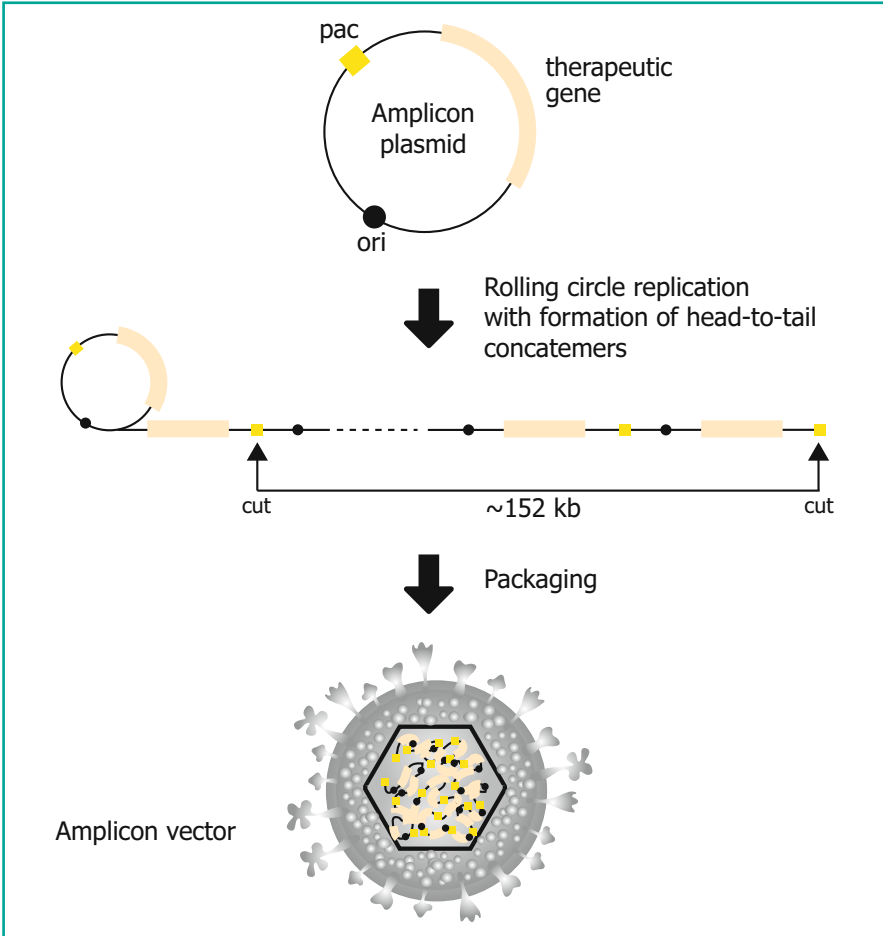
Fig. 3.30 HSV-1 amplicon vectors. Amplicons are bacterial plasmids that contain one or more trans-gene cassettes and two non-coding viral sequences, an origin of DNA replication (*ori*) and a DNA cleavage/packaging signal (*pac*). Upon transfection into a cell line providing HSV-1 helper functions in trans, an amplicon is replicated by a rolling circle mechanism, which generates head-to-tail concatemers which are packaged into HSV-1 particles as approximately 152-kb linear DNA

particles to infected target cells; however, their major difference is that, afterwards, the amplicons persist in the cells without expressing any viral proteins. Thus, while replication-defective vectors are always subject to the potential risk of reactivation and virulence, the amplicon vectors are devoid of this problem. On the other hand, the production of high-titer preparations of amplicons is considerably more difficult than that of replication-defective vectors.

The production of amplicon vectors was initially obtained in cells transfected with an amplicon plasmid (produced in bacteria) and superinfected with a defective HSV-1 virus providing helper function, i.e., supplying in trans all the factors necessary for amplifica-

tion and packaging. In this manner, however, the probability of contaminating the vector preparations with helper virus was exceedingly high. This problem was recently solved by co-transfecting the amplicon with a set of 5 partially overlapping cosmids, expressing all the required viral proteins; the system can be further improved by using a bacterial artificial chromosome (BAC) as a source of viral proteins.

3.5.5.3
Properties of HSV-1 Vectors

Each of the three types of HSV-1 vectors has different characteristics and, as a consequence, possible modes of application in gene therapy.

One of the major applications of attenuated, *replication-competent vectors* is for the oncolytic therapy of cancer (cf. also the section on 'Oncolytic Viruses'). Different modified viruses have been produced so far for this purpose. The first generation contained mutations in a single gene, aimed at limiting viral replication in actively replicating cells. The genes considered included the gene UL23 (ICP36), coding for TK, or the gene UL39 (ICP6), coding for the major subunit of RR, or the gene coding for the neurovirulence factor ICP34.5. While use of the first two mutants did not extend beyond animal experimentation due to the risk of toxicity, different mutants of ICP34.5, which show considerable anti-tumor effect in animal models, are currently the subject of different clinical trials. Given the success of this first set of attenuated viruses, additional series of vectors were obtained. A second generation includes viruses bearing multiple mutations, in particular, ICP34.5 plus ICP6, currently also in clinical experimentation. Viruses of the third generation, in addition to deletions in the above-mentioned genes, also function as real gene transfer vectors, since they contain genes coding for various cytokines (IL-4, IL-12, IL-10, GM-CSF) or for the co-stimulatory molecule B7.1, with the ultimate purpose of increasing tumor immunogenicity in addition to oncolysis.

An additional interesting property of attenuated herpetic vectors is their use as live, attenuated vaccines for immunization against wild-type HSV-1. Different combinations of mutants in the genes coding for the surface glycoproteins or the IE genes have been obtained. The ultimate purpose of these modifications is to generate an ideal strain that is able to propagate in a limited manner without inducing neurotoxicity, albeit eliciting a strong immune response. Experimentation with such mutants is currently limited to the preclinical phase.

Finally, numerous studies have indicated that the HSV-1 tropism can be changed through the deletion or modification of the surface glycoproteins of the virions. In particular, deletion of gB and/or gC, substitution of gD with VSV-G, and creation of chimeras between gC and receptor-specific ligands could allow, alternatively, broadening of the cell tropism of the virus or its restriction to cell types expressing defined receptors.

The *replication-defective vectors* and the *amplicon vectors* have instead been utilized, so far at the preclinical level, to express a variety of genes in different tissues. In the brain, these genes include those coding for proteins having toxic or pro-apoptotic (for gene therapy of gliomas), neurotrophic (such as NGF or BDNF for gene therapy of neurodegenerative disorders), or enzymatic activity (such as tyrosine hydroxylase for gene therapy of Parkinson's disease). Other studies have shown that both the replication-defective and the

amplicon vectors can be used in tissues different from the nervous tissue, including muscle, heart, and liver, or for genetic vaccination. It is likely that a few of these studies will reach clinical experimentation stages in the near future.

3.5.6
Viral Vectors for Gene Therapy: Fields of Application and Comparative Evaluation

A synopsis of the properties, pros, and cons of the five major classes of viral vectors for gene therapy is reported in Table 3.9. The main parameters distinguishing the different vectors can be summarized as follows.

Table 3.9 Pros and cons of the main viral vectors for gene therapy

Vector	Pros	Cons
Gammaretroviral vectors	Efficient transduction Integration into the host cell genome	Low titers (if not pseudotyped) Insertional mutagenesis Silencing of gene expression Exclusively transduce actively replicating cells
Lentiviral vectors	Transduction of quiescent cells *in vitro* and *in vivo* Integration into the host cell genome	Need pseudotyping Possible generation of RCLs Possible mobilization of vector in HIV-infected patients Potential for insertional mutagenesis
Adenoviral vectors (first generation)	Very efficient transduction High-level transgene expression Production at high titers Transduction of both quiescent and replicating cells Broad host range	Transient transduction Stimulation of strong inflammatory and immune responses
AAV vectors	Derived from a non-pathogenic virus Production at high titers Infection of quiescent cells *in vivo* Very long persistence and gene expression	Limited cloning capacity (<5 kb) Lack of a packaging cell line Tropism limited to specific cell types
Herpesviral vectors	Persistence in latent form Large cloning capacity Tropism for neuronal cells	Difficult to manipulate Poor knowledge of several biological features Pathogenic genes difficult to identify and eliminate

Cloning capacity. The currently available vectors significantly differ in their capacity to accommodate DNA fragments of different lengths, with a spectrum ranging from 3–4 kb for AAV vectors, to 8 kb for retroviruses, to 30–40 kb for gutless adenoviral vectors, to 150 kb for HSV amplicons. In evaluating these lengths, one has to consider that the coding portion of human genes has an average size of ~1.5 kb; as a consequence, even the relatively small size of AAV allows the delivery not only of short regulatory RNAs but also of the vast majority of therapeutic cDNAs. The size of vectors becomes substantially limiting in two specific situations, namely when the cDNAs to be transferred are exceedingly long (for example, in the case of dystrophin –9.7 kb, or coagulation Factor VIII >8 kb), or when transcription of the therapeutic gene needs to be strictly regulated, a condition usually requiring very long regulatory elements (for example, for gene therapy of thalassemias or diabetes). In these situations, the possibility of using gutless adenoviral vectors or herpetic amplicons would be most desirable. In particular, the latter class of vectors appears suitable to accommodate entire genetic loci, composed of the whole gene (exons plus introns) and its regulatory elements.

Simplicity of production. The production systems for the different vectors are quite different. In the case of amphotropic and ecotropic gammaretroviral vectors, the possibility to utilize packaging cell lines offers an obvious advantage in terms of simplicity and cost. In contrast, production of AAV, lentiviral, and herpetic amplicon vectors is based on transient transfection of plasmids in the producer cells. Transient transfection is however also used for gammaretroviral vectors when VSV-G pseudotyping is required. Another issue related to vector production concerns the purity of the preparations and, in particular, the contamination of vectors with autonomous replicating or helper viruses.

Efficiency of transduction. The gammaretroviral and lentiviral vectors, both pseudotyped with VSV-G, are capable of transducing a vast number of cell types. However, one of the strict requirements of gammaretroviral vectors is that the infected cells are in active replication. This characteristic substantially prevents the use of these vectors in most cell types *in vivo* and restricts their utilization to *ex vivo* cultured cells. In contrast, lentiviral vectors can transduce non-replicating cells, albeit provided that they are metabolically active. These vectors can thus successfully be used for gene transfer in neurons *in vivo* and in non-stimulated hematopoietic stem cells *ex vivo*. Adenoviral vectors are also very efficient in transducing both replicating and non-replicating cells and, by virtue of the utilization of the ubiquitously expressed CAR receptor, are capable of transducing most cell types, both *in vivo* and *ex vivo*. These vectors are probably the most efficient delivery system currently available both in terms of number of transduced cells and levels of therapeutic gene expression. However, in replicating cells, their efficacy is limited by the lack of integration into the genome: during cell proliferation and expansion, they progressively become diluted or lost. AAV vectors also use ubiquitous receptors for internalization (HSPGs and sialic acid) and are thus internalized by most cells. However, the tropism of these vectors is essentially restricted by events occurring after internalization and involving vector transport to the nucleus and, mostly, single-stranded to double-stranded DNA conversion. The efficiency of these events essentially restricts AAV transduction to postmitotic tissues such as brain, retina, skeletal muscle, and heart. AAV vectors thus represent the vectors of choice for transduction of these tissues. The recent identification of an array of different AAV serotypes now extends their use to other organs, such as liver, pan-

creas, and lung. Finally, while gammaretroviruses and lentiviruses transduce cells at low multiplicity (one or maximum two vector copies per cell), adenoviruses and AAVs infect at high multiplicity, with several copies of the vector commonly being found to be non-integrated in the nucleus of the transduced cells.

Persistence. The five viral vectors differ substantially in terms of persistence of their genome in the transduced cells and duration of therapeutic gene expression. Gammaretroviruses and lentiviruses integrate into the host cell genome and are thus permanently inherited at every cell division. These are the vectors of choice for the treatment of inherited disorders with recessive monogenic transmission, in which permanent correction of the molecular defect is the therapeutic goal. However, in several circumstances, gammaretroviral vectors especially undergo progressive silencing of gene expression, due to irreversible proviral methylation. In contrast, albeit not being integrated into the genome, AAV transduces long-living and non-replicating cells, and thus persists for month- or year-long periods in these tissues; probably because it remains episomal, methylation-induced transcriptional silencing does not occur and therapeutic gene expression persists. In contrast, the duration of gene expression is very short with first-generation adenoviral vectors, not because of any intrinsic property of the vector itself, but due to the recognition of the transduced cells by the immune system: in immunocompetent animals, expression of the therapeutic effect usually does not last longer than 10–14 days after inoculation. This can still be useful in conditions in which persistence of expression is not required or desirable: this is the case, for example, of gene therapy of cancer or genetic vaccination, or when gene therapy is used to transiently express a growth factor, for example to stimulate therapeutic angiogenesis. Finally, the vectors based on HSV-1 might become very useful for the prolonged expression of genes in the brain, thus exploiting the property of the wild-type virus to persist in episomal form in the neurons for the whole life of the infected organism.

Induction of undesired effects. The use of viral vectors is still strongly impacted by the fear of the possible induction of undesired or frankly pathologic effects. First-generation adenoviral vectors are strongly pro-inflammatory and induce a powerful immune response. Gammaretroviral vectors and, potentially, lentiviral vectors can induce insertional mutagenesis. Both replication-defective and oncolytic HSV vectors raise concerns for their possible reactivation and consequent neurovirulence. Finally, the production of adenoviral, gammaretroviral, lentiviral, and replicating herpesviral vectors is fraught with the possible generation of autonomously replicating wild-type or recombinant viruses. Of note, the only vector system not raising important safety concerns at this moment is that based on AAV.

3.1 Cellular Barriers to Gene Delivery

Further Reading

Conner SD, Schmid SL (2003) Regulated portals of entry into the cell. Nature 422:37–44
Doherty GJ, McMahon HT (2009) Mechanisms of endocytosis. Annu Rev Biochem 78:857–902
Sandvig K, van Deurs B (2005) Delivery into cells: lessons learned from plant and bacterial toxins. Gene Ther 12:865–872

Selected Bibliography

Kerr MC, Teasdale RD (2009) Defining macropinocytosis. Traffic 10:364–371

Kirkham M, Parton RG (2005) Clathrin-independent endocytosis: new insights into caveolae and non-caveolar lipid raft carriers. Biochim Biophys Acta 1745:273–286

Medina-Kauwe LK (2007) "Alternative" endocytic mechanisms exploited by pathogens: new avenues for therapeutic delivery? Adv Drug Deliv Rev 59:798–809

Nichols B (2003) Caveosomes and endocytosis of lipid rafts. J Cell Sci 116:4707–4714

Pelkmans L, Puntener D, Helenius A (2002) Local actin polymerization and dynamin recruitment in SV40-induced internalization of caveolae. Science 296:535–539

Plemper RK, Wolf DH (1999) Retrograde protein translocation: ERADication of secretory proteins in health and disease. Trends Biochem Sci 24:266–270

Roth MG (2006) Clathrin-mediated endocytosis before fluorescent proteins. Nat Rev Mol Cell Biol 7:63–68

Torgersen ML, Skretting G, van Deurs B, Sandvig K (2001) Internalization of cholera toxin by different endocytic mechanisms. J Cell Sci 114:3737–3747

3.2 Direct Inoculation of DNAs and RNAs

Selected Bibliography

Braun S (2008) Muscular gene transfer using nonviral vectors. Curr Gene Ther 8:391–405

Herweijer H, Wolff JA (2003) Progress and prospects: naked DNA gene transfer and therapy. Gene Ther 10:453–458

3.3 Physical Methods

Further Reading

Frenkel V (2008) Ultrasound mediated delivery of drugs and genes to solid tumors. Adv Drug Deliv Rev 60:1193–1208

Hynynen K (2008) Ultrasound for drug and gene delivery to the brain. Adv Drug Deliv Rev 60:1209–1217

Wells DJ (2004) Gene therapy progress and prospects: electroporation and other physical methods. Gene Ther 11:1363–1369

Selected Bibliography

Andre F, Mir LM (2004) DNA electrotransfer: its principles and an updated review of its therapeutic applications. Gene Ther 11[Suppl 1]:S33–42

Bigey P, Bureau MF, Scherman D (2002) In vivo plasmid DNA electrotransfer. Curr Opin Biotechnol 13:443–447

Hagstrom JE (2003) Plasmid-based gene delivery to target tissues in vivo: the intravascular approach. Curr Opin Mol Ther 5:338–344

Heller LC, Heller R (2006) In vivo electroporation for gene therapy. Hum Gene Ther 17:890–897

Lewis DL, Wolff JA (2007) Systemic siRNA delivery via hydrodynamic intravascular injection. Adv Drug Deliv Rev 59:115–123

Mennuni C, Calvaruso F, Zampaglione I et al (2002) Hyaluronidase increases electrogene transfer efficiency in skeletal muscle. Hum Gene Ther 13:355–365

Mir LM (2008) Application of electroporation gene therapy: past, current, and future. Methods Mol Biol 423:3–17

Newman CM, Bettinger T (2007) Gene therapy progress and prospects: ultrasound for gene transfer. Gene Ther 14:465–475

Reed SD, Li S (2009) Electroporation advances in large animals. Curr Gene Ther (*in press*)

Stein U, Walther W, Stege A et al (2008) Complete in vivo reversal of the multidrug resistance phenotype by jet-injection of anti-MDR1 short hairpin RNA-encoding plasmid DNA. Mol Ther 16:178–186

Walther W, Siegel R, Kobelt D et al (2008) Novel jet-injection technology for nonviral intratumoral gene transfer in patients with melanoma and breast cancer. Clin Cancer Res 14:7545–7553

Walther W, Stein U, Fichtner I et al (2001) Nonviral in vivo gene delivery into tumors using a novel low volume jet-injection technology. Gene Ther 8:173–180

Walther W, Stein U, Fichtner I et al (2002) Intratumoral low-volume jet-injection for efficient nonviral gene transfer. Mol Biotechnol 21:105–115

3.4 Chemical Methods

Further Reading

Beerens AM, Al Hadithy AF, Rots MG, Haisma HJ (2003) Protein transduction domains and their utility in gene therapy. Curr Gene Ther 3:486–494

Elouahabi A, Ruysschaert JM (2005) Formation and intracellular trafficking of lipoplexes and polyplexes. Mol Ther 11:336–347

Fittipaldi A, Giacca M (2005) Transcellular protein transduction using the Tat protein of HIV-1. Adv Drug Deliv Rev 57:597–608

Giacca M (2004) The HIV-1 Tat protein: a multifaceted target for novel therapeutic opportunities. Curr Drug Targets Immune Endocr Metabol Disord 4:277–285

Park TG, Jeong JH, Kim SW (2006) Current status of polymeric gene delivery systems. Adv Drug Deliv Rev 58:467–486

van Dillen IJ, Mulder NH, Vaalburg W et al (2002) Influence of the bystander effect on HSV-tk/GCV gene therapy. A review. Curr Gene Ther 2:307–322

Vile RG, Russell SJ, Lemoine NR (2000) Cancer gene therapy: hard lessons and new courses. Gene Ther 7:2–8

Wasungu L, Hoekstra D (2006) Cationic lipids, lipoplexes and intracellular delivery of genes. J Control Release 116:255–264

Selected Bibliography

Dass CR (2004) Lipoplex-mediated delivery of nucleic acids: factors affecting in vivo transfection. J Mol Med 82:579–591

Dincer S, Turk M, Piskin E (2005) Intelligent polymers as nonviral vectors. Gene Ther 12[Suppl 1]:S139–145

Dufes C, Uchegbu IF, Schatzlein AG (2005) Dendrimers in gene delivery. Adv Drug Deliv Rev 57:2177–2202

Duncan R, Izzo L (2005) Dendrimer biocompatibility and toxicity. Adv Drug Deliv Rev 57:2215–2237

Fittipaldi A, Giacca M (2005) Transcellular protein transduction using the Tat protein of HIV-1. Adv Drug Deliv Rev 57:597–608

Hoekstra D, Rejman J, Wasungu L et al (2007) Gene delivery by cationic lipids: in and out of an endosome. Biochem Soc Trans 35:68–71

Lange, A, Mills RE, Lange CJ et al (2007) Classical nuclear localization signals: definition, function, and interaction with importin alpha. J Biol Chem 282:5101–5105

Pedroso de Lima MC, Simoes S, Pires P et al (2001) Cationic lipid-DNA complexes in gene delivery: from biophysics to biological applications. Adv Drug Deliv Rev 47:277–294

Rainov NG (2000) A phase III clinical evaluation of herpes simplex virus type 1 thymidine kinase and ganciclovir gene therapy as an adjuvant to surgical resection and radiation in adults with previously untreated glioblastoma multiforme. Hum Gene Ther 11:2389–2401

Ram Z, Culver KW, Oshiro EM et al (1997) Therapy of malignant brain tumors by intratumoral implantation of retroviral vector-producing cells. Nat Med 3:1354–1361

Svenson S (2009) Dendrimers as versatile platform in drug delivery applications. Eur J Pharm Biopharm 71:445–462

Trask TW, Trask RP, Aguilar-Cordova E et al (2000) Phase I study of adenoviral delivery of the HSV-tk gene and ganciclovir administration in patients with current malignant brain tumors. Mol Ther 1:195–203

Wolff JA, Rozema DB (2008) Breaking the bonds: non-viral vectors become chemically dynamic. Mol Ther 16:8–15

Zuhorn IS, Engberts JB, Hoekstra D (2007) Gene delivery by cationic lipid vectors: overcoming cellular barriers. Eur Biophys J 36:349–362

3.5 Viral Vectors

Further Reading

Bessis N, GarciaCozar FJ, Boissier MC (2004) Immune responses to gene therapy vectors: influence on vector function and effector mechanisms. Gene Ther 11[Suppl 1]:S10–17

Bestor TH (2000) Gene silencing as a threat to the success of gene therapy. J Clin Invest 105:409–411

Chang AH, Sadelain M (2007) The genetic engineering of hematopoietic stem cells: the rise of lentiviral vectors, the conundrum of the ltr, and the promise of lineage-restricted vectors. Mol Ther 15:445–456

Coffin JM, Hughes H, Varmus HE (1997) Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY, USA

Daniel R, Smith JA (2008) Integration site selection by retroviral vectors: molecular mechanism and clinical consequences. Hum Gene Ther 19:557–568

Kay MA, Glorioso JC, Naldini L (2001) Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. Nat Med 7:33–40

Knipe DM, Roizman B, Howley PM et al (2006) Fields' virology, 5th edn. Lippincott Williams & Wilkins, Philadelphia, PA, USA

Schambach A, Baum C (2008) Clinical application of lentiviral vectors: concepts and practice. Curr Gene Ther 8:474–482

St George JA (2003) Gene therapy progress and prospects: adenoviral vectors. Gene Ther 10:1135–1141

Thomas CE, Ehrhardt A, Kay MA (2003) Progress and problems with the use of viral vectors for gene therapy. Nat Rev Genet 4:346–358

Wu Z, Asokan A, Samulski RJ (2006) Adeno-associated virus serotypes: vector toolkit for human gene therapy. Mol Ther 14:316–327

Zentilin L, Giacca M (2008) Adeno-associated virus vectors: versatile tools for in vivo gene transfer. Contrib Nephrol 159:63–77

3

Selected Bibliography

Aghi M, Martuza RL (2005) Oncolytic viral therapies: the clinical experience. Oncogene 24:7802–7816

Aiken C (1997) Pseudotyping human immunodeficiency virus type 1 (HIV-1) by the glycoprotein of vesicular stomatitis virus targets HIV-1 entry to an endocytic pathway and suppresses both the requirement for Nef and the sensitivity to cyclosporin A. J Virol 71:5871–5877

Alba R, Bosch A, Chillon M (2005) Gutless adenovirus: last-generation adenovirus for gene therapy. Gene Ther 12[Suppl 1]:S18–27

Argnani R, Lufino M, Manservigi M, Manservigi R (2005) Replication-competent herpes simplex vectors: design and applications. Gene Ther 12[Suppl 1]:S170–177

Barnard RJ, Elleder D, Young JA (2006) Avian sarcoma and leukosis virus-receptor interactions: from classical genetics to novel insights into virus-cell membrane fusion. Virology 344:25–29

Barquinero J, Eixarch H, Perez-Melgosa M (2004) Retroviral vectors: new applications for an old tool. Gene Ther 11[Suppl 1]:S3–9

Berges BK, Wolfe JH, Fraser NW (2007) Transduction of brain by herpes simplex virus vectors. Mol Ther 15:20–29

Berns KI, Linden RM (1995) The cryptic life style of adeno-associated virus. BioEssays 17:237–245

Berto E, Bozac A, Marconi P (2005) Development and application of replication-incompetent HSV-1-based vectors. Gene Ther 12[Suppl 1]:S98–102

Brunetti-Pierri N, Ng P (2008) Progress and prospects: gene therapy for genetic diseases with helper-dependent adenoviral vectors. Gene Ther 15:553–560

Buning H, Ried MU, Perabo L et al (2003) Receptor targeting of adeno-associated virus vectors. Gene Ther 10:1142–1151

Burton EA, Bai Q, Goins WF, Glorioso JC (2002) Replication-defective genomic herpes simplex vectors: design and production. Curr Opin Biotechnol 13:424–428

Cereseto A, Giacca M (2004) Integration site selection by retroviruses. AIDS Rev 6:13–20

Cervelli T, Palacios JA, Zentilin L et al (2008) Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins. J Cell Sci 121:349–357

Chirmule N, Propert K, Magosin S et al (1999) Immune responses to adenovirus and adeno-associated virus in humans. Gene Ther 6:1574–1583

Danthinne X, Imperiale MJ (2000) Production of first generation adenovirus vectors: a review. Gene Ther 7:1707–1714

Dull T, Zufferey R, Kelly M et al (1998) A third-generation lentivirus vector with a conditional packaging system. J Virol 72:8463–8471

Dutheil N, Shi F, Dupressoir T, Linden RM (2000) Adeno-associated virus site-specifically integrates into a muscle-specific DNA region. Proc Natl Acad Sci U S A 97:4862–4866

Epstein AL (2005) HSV-1-based amplicon vectors: design and applications. Gene Ther 12[Suppl 1]:S154–158

Flotte TR (2004) Gene therapy progress and prospects: recombinant adeno-associated virus (rAAV) vectors. Gene Ther 11:805–810

Flotte TR, Berns KI (2005) Adeno-associated virus: a ubiquitous commensal of mammals. Hum Gene Ther 16:401–407

Gao G, Vandenberghe LH, Wilson JM (2005) New recombinant serotypes of AAV vectors. Curr Gene Ther 5:285–297

Gregorevic P, Blankinship MJ, Allen JM et al (2004) Systemic delivery of genes to striated muscles using adeno-associated viral vectors. Nat Med 10:828–834

Grimm D, Kern A, Rittner K, Kleinschmidt JA (1998) Novel tools for production and purification of recombinant adenoassociated virus vectors Hum. Gene Ther 9:2745–2760

Hendrie PC, Russell DW (2005) Gene targeting with viral vectors. Mol Ther 12:9–17

Jooss K, Chirmule N (2003) Immunity to adenovirus and adeno-associated viral vectors: implications for gene therapy. Gene Ther 10:955–963

Linden RM, Ward P, Giraud C et al (1996) Site-specific integration by adeno-associated virus. Proc Natl Acad Sci U S A 93:11288–11294

Liu Q, Muruve DA (2003) Molecular basis of the inflammatory response to adenovirus vectors. Gene Ther 10:935–940

Manganaro L, Lusic M, Gutierrez MI et al (2009) Concerted action of cellular JNK and Pin1 restricts HIV-1 genome integration to activated CD4+ T lymphocytes. Nat Med

Marconi P, Argnani R, Berto E et al (2008) HSV as a vector in vaccine development and gene therapy. Hum Vaccin 4:91–105

McCarty DM, Young SM Jr, Samulski RJ (2004) Integration of adeno-associated virus (AAV) and recombinant AAV vectors. Annu Rev Genet 38:819–845

Merten OW, Geny-Fiamma C, Douar AM (2005) Current issues in adeno-associated viral vector production. Gene Ther 12[Suppl 1]:S51–61

Miller AD (1990) Retrovirus packaging cells. Hum Gene Ther 1:5–14

Miller AD (1996) Cell-surface receptors for retroviruses and implications for gene transfer. Proc Natl Acad Sci U S A 93:11407–11413

Miller DG, Wang PR, Petek LM et al (2006) Gene targeting in vivo by adeno-associated virus vectors. Nature Biotech 24:1022–1026

Mueller C, Flotte TR (2008) Clinical gene therapy using recombinant adeno-associated virus vectors. Gene Ther 15:858–863

Naldini L, Blomer U, Gallay P et al (1996) In vivo gene delivery and stable transduction of non-dividing cells by a lentiviral vector. Science 272:263–267

Nienhuis AW, Dunbar CE, Sorrentino BP (2006) Genotoxicity of retroviral integration in hematopoietic cells. Mol Ther 13:1031–1049

Russell DW, Hirata RK (1998) Human gene targeting by viral vectors. Nat Genet 18:325–330

Russell WC (2000) Update on adenovirus and its vectors. J Gen Virol 81:2573–2604

Russell WC (2009) Adenoviruses: update on structure and function. J Gen Virol 90:1–20

Samulski RJ, Berns KI, Tan M, Muzyczka N (1982) Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. Proc Natl Acad Sci USA 79:2077–2081

Schroder AR, Shinn P, Chen H et al (2002) HIV-1 integration in the human genome favors active genes and local hotspots. Cell 110:521–529

Sinn PL, Sauter SL, McCray PB Jr (2005) Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors: design, biosafety, and production. Gene Ther 12:1089–1098

Vasileva A, Jessberger R (2005) Precise hit: adeno-associated virus in gene targeting. Nat Rev Microbiol 3:837–847

Vihinen-Ranta M, Suikkanen S, Parrish CR (2004) Pathways of cell infection by parvoviruses and adeno-associated viruses. J Virol 78:6709–6714

Wang Z, Zhu T, Qiao C et al (2005) Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. Nat Biotechnol 23:321–328

Yamamoto T, Tsunetsugu-Yokota Y (2008) Prospects for the therapeutic application of lentivirus-based gene therapy to HIV-1 infection. Curr Gene Ther 8:1–8

Zentilin L, Qin G, Tafuro S et al (2000) Variegation of retroviral vector gene expression in myeloid cells. Gene Ther 7:153–166

Zufferey R, Dull T, Mandel RJ et al (1998) Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol 72:9873–9880