An Introduction to Molecular Medicine and Gene Therapy. Edited by Thomas F. Kresina, PhD
Copyright © 2001 by Wiley-Liss, Inc.
ISBNs: 0-471-39188-3 (Hardback); 0-471-22387-5 (Electronic)

CHAPTER 4

Vectors of Gene Therapy

KATHERINE PARKER PONDER, M.D.

INTRODUCTION

Currently, gene therapy refers to the transfer of a gene that encodes a functional protein into a cell or the transfer of an entity that will alter the expression of an endogenous gene in a cell. The efficient transfer of the genetic material into a cell is necessary to achieve the desired therapeutic effect. For gene transfer, either a messenger ribonucleic acid (mRNA) or genetic material that codes for mRNA needs to be transferred into the appropriate cell and expressed at sufficient levels. In most cases, a relatively large piece of genetic material (>1 kb) is required that includes the promoter sequences that activate expression of the gene, the coding sequences that direct production of a protein, and signaling sequences that direct RNA processing such as polyadenylation. A second class of gene therapy involves altering the expression of an endogenous gene in a cell. This can be achieved by transferring a relatively short piece of genetic material (20 to 50 bp) that is complementary to the mRNA. This transfer would affect gene expression by any of a variety of mechanisms through blocking translational initiation, mRNA processing, or leading to destruction of the mRNA. Alternatively, a gene that encodes antisense RNA that is complementary to a cellular RNA can function in a similar fashion.

Facilitating the transfer of genetic information into a cell are vehicles called vectors. Vectors can be divided into viral and nonviral delivery systems. The most commonly used viral vectors are derived from retrovirus, adenovirus, and adeno-associated virus (AAV). Other viral vectors that have been less extensively used are derived from herpes simplex virus 1 (HSV-1), vaccinia virus, or baculovirus. Nonviral vectors can be either plasmid deoxyribonucleic acid (DNA), which is a circle of double-stranded DNA that replicates in bacteria or chemicaly synthesized compounds that are or resemble oligodeoxynucleotides. Major considerations in determining the optimal vector and delivery system are (1) the target cells and its characteristics, that is, the ability to be virally transduced ex vivo and reinfused to the patient, (2) the longevity of expression required, and (3) the size of the genetic material to be transferred.

VIRAL VECTORS USED FOR GENE THERAPY

Based on the virus life cycle, infectious virions are very efficient at transferring genetic information. Most gene therapy experiments have used viral vectors comprising elements of a virus that result in a replication-incompetent virus. In initial studies, immediate or immediate early genes were deleted. These vectors could potentially undergo recombination to produce a wild-type virus capable of multiple rounds of replication. These viral vectors replaced one or more viral genes with a promoter and coding sequence of interest. Competent replicating viral vectors were produced using packaging cells that provided deleted viral genes in trans. For these viruses, protein(s) normally present on the surface of the wild-type virus were also present in the viral vector particle. Thus, the species and the cell types infected by these viral vectors remained the same as the wild-type virus from which they were derived. In specific cases, the tropism of the virus was modified by the surface expression of a protein from another virus, thus allowing it to bind and infect other cell types. The use of a protein from another virus to alter the tropism for a viral vector is referred to as pseudotyping.

A number of viruses have been used to generate viral vectors for use in gene therapy. The characteristics of these viruses and their virulence are shown in Table 4.1. Characteristics of viral vectors that have been generated from these viruses are shown in Table 4.2. Important features that distinguish the different viral vectors include the size of the gene insert accepted, the duration of expression, target cell infectivity, and integration of the vector into the genome.

RETROVIRAL VECTORS

Retroviruses are comprised of two copies of a positive single-stranded RNA genome of 7 to 10 kb. Their RNA genome is copied into double-stranded DNA, which integrates into the host cell chromosome and is stably maintained. A property that allowed for the initial isolation was the rapid induction of tumors in susceptible animals by the transfer of cellular oncogenes into cells. However, retroviruses can also cause delayed malignancy due to insertional activation of a downstream oncogene or inactivation of a tumor suppressor gene. Specific retroviruses, such as the human immunodeficiency virus (HIV), can cause the immune deficiency associated with the acquired immunodeficiency syndrome (AIDS) see Chapter 12. Retroviruses are classified into seven distinct genera based on features such as envelope nucleotide structure, nucleocapsid morphology, virion assembly mode, and nucleotide sequence.

Retroviruses are ~100 nm in diameter and contain a membrane envelope. The envelope contains a virus-encoded glycoprotein that specifies the host range or types of cells that can be infected by binding to a cellular receptor. The envelope protein promotes fusion with a cellular membrane on either the cell surface or in an endosomal compartment. The ecotropic Moloney murine leukemia virus (MLV) receptor is a basic amino acid transporter that is present on murine cells but not cells from other species. The amphotropic MLV receptor is a phosphate transporter that is present on most cell types from a variety of species including human cells. There are co-HIV receptors, CD4, and a chemokine receptor. After binding to the

TABLE 4.1 Characteristics of Viruses That Have Been Used to Generate Viral Vectors

Virus	Size and Type of genome	Viral Proteins	Physical Properties	Disease in Animals
Retrovirus	7–10 kb of single-stranded RNA	Gag, Pro, Pol, Env	100 nm diameter; enveloped	Rapid or slow induction of tumors; acquired immunodeficiency syndrome (AIDS)
Adenovirus	36-kb double-stranded linear DNA	Over 25 proteins	70–100 nm in diameter; nonenveloped	Cold; conjunctivitis; gastroenteritis
Adenovirus-associated virus	4.7-kb single-stranded linear DNA	Rep and Cap	18–26 nm in diameter; nonenveloped	No known disease
Herpes simplex virus 1 (HSV-1)	152 kb of double-stranded linear DNA	Over 81 proteins	110 nm in diameter	Mouth ulcers and genital warts; encephalitis
Vaccinia virus	190 kb of double-stranded linear DNA	Over 198 open reading frames	350 by 270 nm rectangles; enveloped	Attenuated virus that was used to vaccinate against smallpox
Baculovirus	130 kb of double-stranded circular DNA	Over 60 proteins	270 by 45 nm rectangles; enveloped	None in mammals; insect pathogen

cellular receptor, the viral RNA enters the cytoplasm and is copied into double-stranded DNA via reverse transcriptase (RT) contained within the virion. The double-stranded DNA is transferred to the nucleus, where it integrates into the host cell genome by a mechanism involving the virus-encoded enzyme integrase. This activity is specific for each retrovirus. For MLV, infection is only productive in dividing cells, as transfer of the DNA to the nucleus only occurs during breakdown of the nuclear membrane during mitosis. For HIV, infection can occur in nondividing cells, as the matrix protein and the *vpr*-encoded protein have nuclear localization signals that allow transfer of the DNA into the nucleus to occur.

Moloney Murine Leukemia Virus: MLV Proteins

Retroviral proteins are important in the manipulation of the system to develop a vector. MLV is a relatively simple virus with four viral genes: *gag*, *pro*, *pol*, and *env* (Fig. 4.1). The *gag* gene encodes the group specific antigens that make up the viral core. The Gag precursor is cleaved into four polypeptides (10, 12, 15, and 30 kD) by the retroviral protease (PR). The 15-kD matrix protein associates closely with the membrane and is essential for budding of the viral particle from the membrane. The 12-kD phosphoprotein (pp12) is of unresolved function. The 30-kD capsid protein

TABLE 4.2 Summary of Relative Advantages and Disadvantages of Vectors Used for Gene Therapy

Vector	Infects Nondividing Cells?	Maximum Size of Insert	Stability of Expression	Titer
Retroviral vectors	No (yes for lentiviral vectors)	≤8 kb	Stable (random DNA insertion)	1×10^6 cfu/ml unconcentrated; 1×10^8 cfu/ml concentrated
Adenovirus	Yes	8 kb for E1/E3 deleted vectors; 35 kb for "gutless" vectors	Expression lost in 3–4 weeks in normal animals; expression can last weeks to months with immunosuppression. No integration	1×10^{12} pfu/ml
Adenoassociated virus (AAV)	Yes	<4.5 kb	Stable; it is unclear if DNA integrates in vivo	1×10^6 infectious particles/ml unconcentrated; 1×10^{10} infectious particles/ml concentrated
Herpes simplex virus (HSV)-1	Yes	>25 kb	Stable; maintained as episome	1×10^{10} pfu/ml
Vaccinia	Yes	>25 kb	Expression transient due to an immune response; replicates in cytoplasm	1×10^8 pfu/ml
Baculovirus	Yes	>20 kb	Unstable	1×10^{10} pfu/ml

forms the virion core while the 10-kD nucleocapsid protein binds to the RNA genome in a viral particle. The PR and polymerase (Pol) proteins are produced from a Gag/Pro/Pol precursor. This precursor is only 5% as abundant as the Gag precursor and is produced by translational read-through of the *gag* termination codon. The number of infectious particles produced by a cell decreases dramatically if PR and Pol are as abundant as the Gag-derived proteins. PR cleaves a Gag/Pro/Pol precursor into the active polypeptides, although it is unclear how the first PR gets released from the precursor. The *pol* gene product is cleaved into 2 proteins, the amino terminal 80-kD reverse transcriptase (RT) and the carboxy terminal 46-kD integrase (IN). The RT has both reverse transcriptase activity (which functions in RNA- or DNA-directed DNA polymerization) and RNase H activity (which degrades the RNA component of an RNA:DNA hybrid). The IN protein binds to double-stranded DNA at the viral *att* sites located at the ends of each long terminal repeat and mediates integration into the host cell chromosome.

The *env* gene is translated from a subgenomic RNA that is generated by splicing between the 5′ splice site in the 5′ untranslated region and the 3′ splice site present just upstream of the *env* coding sequence. The *env* precursor is processed

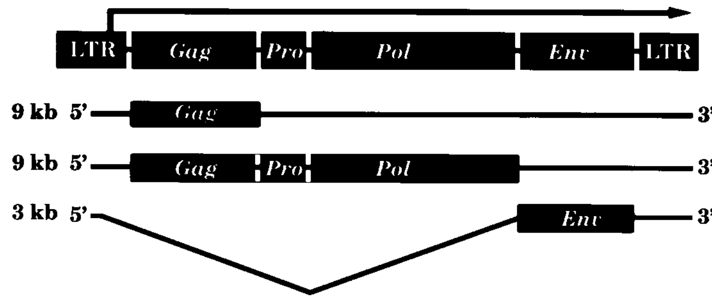

FIGURE 4.1 Diagram of a Moloney murine leukemia retrovirus (MLV). The proviral form with two complete long terminal repeats (LTRs) and the genomic RNA that is expressed from the provirus are shown at the top. The genomic RNA can be translated to produce the *Gag* gene products, or produce a Gag/Pro/Pol precursor by reading through the translational stop codon at the 3′ end of the *Gag* gene. The genomic RNA can also be spliced to generate a smaller subgenomic RNA, which is translated into the Env protein. The regions that are translated are shown as black boxes, while the untranslated regions of the RNA appear as a black line.

into 3 proteins: SU, transmembrane (TM; or p15E), and p2. The 70-kD SU protein binds to a cell surface receptor. Neutralizing antibodies directed against SU can block infection. The 15-kD TM plays a role in fusion of the virus and cellular membrane. In many retroviruses, the association between the SU and TM proteins is rather tenuous and SU is rapidly lost from virions. This contributes to poor infectivity of viral preparations and instability to manipulations such as concentration by ultracentrifugation. Envelope proteins from different retroviruses, or even from viruses of other families, can be used to produce infectious particles with altered tropism and/or greater stability.

Sequences Required in *cis* for Replication and Packaging

The term provirus refers to the form of the virus that is integrated as double-stranded DNA into the host cell chromosome. Genetic sequences are needed in *cis* to develop a provirus that can transfer genetic information into a target cell. Four important sequences are required in *cis* for replication and infection in the context of gene therapy. They are (1) the long terminal repeats (LTRs), (2) the primer binding site (PBS), (3) the polypurine (PP) tract, and (4) the packaging sequence. These sequences and their function are shown in Figure 4.2. LTRs are approximately 600 nucleotide sequences present at both the 5′ and the 3′ end of the provirus. They initiate transcription at the 5′ end, perform polyadenylation at the 3′ end, and integrate a precise viral genome into a random site of the host cell chromosome by virtue of the *att* sites at either end. The LTR-initiated transcripts serve as an mRNA for the production of viral proteins and as the RNA genome for producing additional virus. The PBS is located just downstream of the 5′ LTR. It binds to a cellular transfer RNA (tRNA), which serves as a primer for the polymerization of the first DNA strand. The PP tract contains at least nine purine nucleotides and is located upstream of the U3 region in the 3′ LTR. The RNA within this sequence is resistant to degradation by RNase H when hybridized with the first DNA strand.

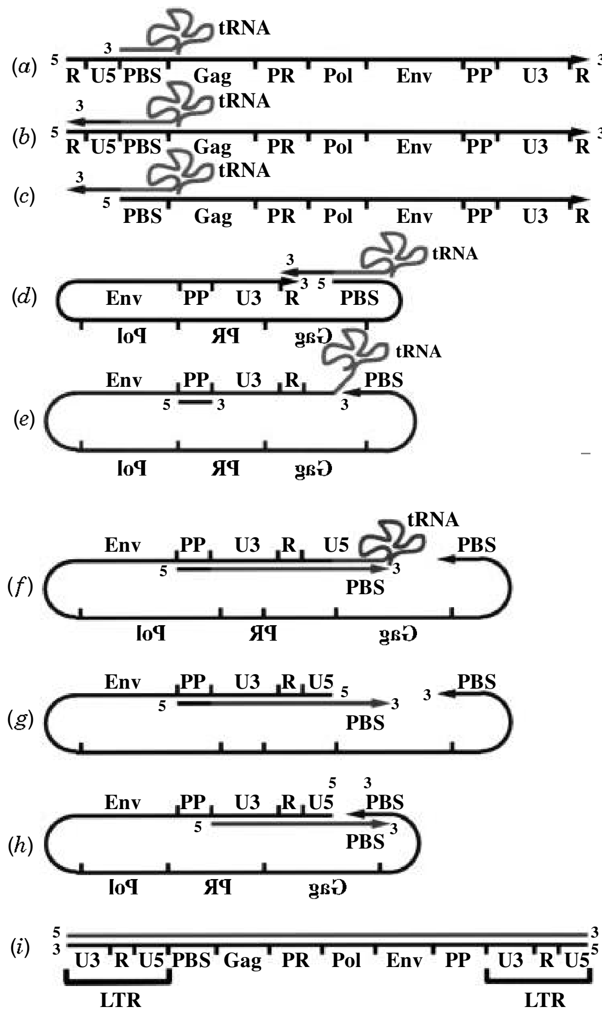

FIGURE 4.2 Mechanism of reverse transcription and integration of the genomic RNA into the host cell chromosome. (*a*) Genomic RNA with a tRNA primer. The genomic RNA has a 60-nt R region (for redundant) at both the 5′ and the 3′ end. The 5′ end has the 75-nt U5 region (for unique to 5′ end) and the 3′ end has the 500-nt U3 region (for unique to 3′ end). The PBS of the genomic RNA (shown in black) hybridizes to the terminal 18 nt at the 3′ end of a tRNA. (*b*) Reverse transcription of the 5′ end of the genomic RNA. The tRNA primer enables the RT to copy the 5′ end of the genomic RNA, to generate a portion of the first DNA strand. (*c*) Degradation of the RNA portion of an RNA : DNA hybrid by RNase H. RNase H degrades the RNA portion that was used as a template for synthesis of the first DNA strand. Although shown as a separate step here, this occurs ~18 nt downstream of where polymerization is occurring. (*d*) First strand transfer. The portion of the first strand that represents the R region hybridizes with the R region in the 3′ end of the genomic RNA. (*e*) Reverse transcription of the remainder of the genomic RNA. The RT copies the genomic RNA up to the PBS. As elongation occurs, RNase H continues to degrade the RNA portion of the RNA : DNA hybrid. The RNA in the PP tract (shown in black) is resistant to cleavage by RNase H and remains associated with the first DNA strand. (*f*) Initiation of second strand synthesis. The primer at the PP tract initiates polymerization of the second strand. Polymerization up to the 3′ end of the PBS continues. Additional sequences in the tRNA are not copied, as the 19th nucleotide is blocked by a methyl group in the base pairing region of the tRNA. (*g*) RNase H digestion of the tRNA. The RNase H degrades the tRNA, which is present in an RNA : DNA hybrid. (*h*) Second strand transfer. The second DNA strand hybridizes to the first DNA strand in the PBS region. (*i*) Completion of the first and second strands. RT copies the remainder of the first and the second DNA strands, to generate a double-stranded linear DNA with intact LTRs at both the 5′ and the 3′ end. The integrase binds to the *att* sequence at the 5′ end of the 5′ LTR and at the 3′ end of the 3′ LTR (not shown) and mediates integration into the host cell chromosome. Upon integration, the viral DNA is usually shortened by two bases at each end, while 4 to 6 nt of cellular DNA is duplicated. Although integration is a highly specific process for viral sequences, integration into the host chromosome appears to be random.

The PP tract therefore serves as the primer for synthesis of the second DNA strand. The packaging signal binds to the nucleocapsid protein of a retroviral particle allowing the genomic RNA to be selectively packaged. Although the encapsidation sequence was initially mapped to the region of the virus between the 5′ LTR and the *gag* gene, vectors that only contained this sequence were packaged inefficiently, resulting in low titers of viral vector produced. Subsequent studies demonstrated that inclusion of some *gag* sequences (the extended packaging signal) greatly increased the titer of the vector produced. Most vectors that are currently in use utilize the extended packaging signal.

Use of Retroviral Sequences for Gene Transfer

All of the genomic sequences that are necessary in cis for transcription and packaging of RNA, for reverse transcription of the RNA into DNA and for integration of the DNA into the host cell chromosome need to be present in the retroviral vector. It is, however, possible to remove the coding sequences from the retroviral genome and replace them with a therapeutic gene to create a retroviral vector. The deletion of viral coding sequences from the retroviral vector makes it necessary to express these genes in *trans* in a packaging cell line. Packaging cell lines that stabilly express the *gag*, *pro*, *pol*, and *env* genes have been generated. The transfer of a plasmid encoding the retroviral vector sequence into packaging cell results in a retroviral particle capable of transferring genetic information into a cell (assuming appropriate tropism). However, upon transfer of the retroviral vector into a cell, infectious particles are not produced because the packaging genes necessary for synthesizing the viral proteins are not present. These vectors are therefore referred to as replication incompetent. Figure 4.3 diagrams how retroviral vectors and packaging cells are generated.

Commonly used retroviral vectors and their salient features are summarized in Table 4.3. Plasmid constructs that resemble the provirus and contain a bacterial origin of replication (see Chapter 1) outside of the LTRs can be propagated in bacteria. The therapeutic gene is cloned into a vector using standard molecular biology techniques. Upon transfection into mammalian cells, the 5′ LTR of the vector DNA initiates transcription of an RNA that can be packed into a viral particle. Although a packaging cell line can be directly transfected with plasmid DNA, the integrated concatemers are unstable and are often deleted during large-scale preparation of vector. To circumvent this problem, most cell lines used in animals are infected with the vector rather than transfected. This involves transfection into one packaging cell line, which produces a vector that can infect a packaging cell line with a different envelope gene. The infected packaging cell line generally contains a few copies of the retroviral vector integrated into different sites as a provirus.

Most vectors have genomic RNAs that are less than 10 kb, to allow for efficient packaging. N2 was the first vector using an extended packaging signal that, as noted earlier, greatly increased the titer of vector produced. In LNL6, the AUG at the translational initiation site was mutated to UAG, which does not support translational initiation. This mutation prevents potentially immunogenic *gag* peptides from being expressed on the surface of a transduced cell. In addition, it decreases the possibility that a recombination event would result in replication-competent virus since the recombinant mutant would not translate the *gag* gene into a protein. The LN

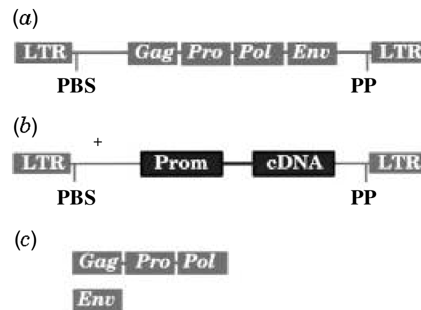

FIGURE 4.3 Retroviral vectors. (*a*) Wild-type retrovirus. The proviral form of a retrovirus is shown. Long-terminal repeats (LTRs) are present at both ends and are necessary for reverse transcription of the RNA into a double-stranded DNA copy and for integration of the DNA into the chromosome. The packaging signal (Ψ) is necessary for the RNA to bind to the inside of a viral particle, although sequences in the *Gag* region increase the efficiency of packaging. The primer binding site (PBS) and the polypurine tract (PP) are necessary for priming of synthesis of the first and second strands of DNA, respectively. The retroviral packaging genes *gag*, *pro*, *pol*, and *env* code for proteins that are necessary for producing a viral particle. (*b*) Retroviral vector. Retroviral vectors have deleted the retroviral coding sequences and replaced them with a promoter and therapeutic gene. The vector still contains the LTR, a packaging signal designated as Ψ^+, which contains a portion of the *Gag* gene, the PBS, and the PP tract, which are necessary for the vector to transmit its genetic information into a target cell. (*c*) Packaging cells. The retroviral vector alone cannot produce a retroviral particle because the retroviral coding sequences are not present. These packaging genes, need to be present in a packaging cell line along with the vector in order to produce a retroviral particle that can transfer genetic information into a new cell.

series is similar but has deleted the sequences 3′ to the *env* gene, thereby limiting recombination events to generate wild-type virus. Double copy vectors place the promoter and coding sequence within the 3′ LTR. As shown in Figure 4.2, the 3′ U3 region is copied into both the 5′ and the 3′ LTRs when the genomic RNA is copied into double-stranded DNA. This results in two complete copies of the transgene in the target cell. The self-inactivating (SIN) vectors were created to address concerns regarding insertional mutagenesis. A deletion in the 3′ U3 region is incorporated into both the 5′ and the 3′ LTR of the provirus. However, insertion into the 3′ U3 region often results in deceased titers. The MFG vector uses the retroviral splice site and the translational initiation signal of the *env* gene resulting in a spliced mRNA that is presumably translated with high efficiency.

Packaging Cells Lines

Commonly used packaging cell lines are summarized in Table 4.4. Initially, packaging cell lines simply deleted the packaging sequence from a single packaging gene plasmid that contained all four genes and both LTRs. These lines occasionally generated replication-competent virus due to homologous recombination between the vector and the packaging constructs. Development of replication-competent virus is a serious concern since it leads to ongoing infection in vivo and ultimately may cause malignant transformation via insertional mutagenesis. Several approaches

TABLE 4.3 Summary of Retroviral Vectors Used for Gene Therapy in Animals or Humans

Name	Salient Features
N2	Contains an intact 5′ and 3′ LTR, an extended packaging signal with 418 nt of coding sequence of the *gag* gene, and an intact translational start codon (AUG) of the *gag* gene. Can recombine to generate wild-type virus.
LNL6	Contains intact 5′ and 3′ LTRs, an extended packaging signal with 418 nt of coding sequence of the *gag* gene, a mutation in the translational start codon (AUG) of the *gag* gene to the inactive UAG, and the 3′ portion of the *env* gene.
LN series	Similar to LNL6 except all *env* sequences are deleted to decrease the chance of recombination with the packaging genes. This series includes LNSX, LNCX, and LXSN, where L stands for LTR promoter, N for neomycin resistance gene, S for SV40 promoter, C for CMV promoter, and X for polylinker sequences for insertion of a therapeutic gene.
Double copy	Places the promoter and the therapeutic gene in the U3 region of the 3′ LTR. This results in two copies of the therapeutic gene within the 5′ and 3′ LTRs after transduction.
Self-inactivating (SIN)	Deletes the enhancer and part of the promoter from the U3 region of the 3′ LTR. This deletion is present in both the 5′ and the 3′ LTRs after transduction. This decreases the chance of transcriptional activation of a downstream oncogene after transduction of a cell.
MFG	Contains an intact 5′ and 3′ LTR, an extended packaging signal with an intact 5′ splice site, a 380-nt sequence with the 3′ end of the *pol* gene and the 3′ splice site, and 100 nt of the 3′ end of the *env* gene. The therapeutic gene is translated from a spliced RNA and uses the *env* gene translational start site.

have been taken to reduce the generation of replication-competent virus. One strategy is to separate the packaging genes into two plasmids integrated into different chromosomal locations. Examples of this approach include the GP + E86, GP + *env*AM12, Ψ-CRIP, and Ψ-CRE packaging cell lines. For these cell lines, the *gag/pro/pol* genes are expressed from one piece of DNA while the *env* gene is expressed from a second piece of DNA. Then each DNA piece is introduced into the cell independently. Another strategy is to minimize homology between the vector and packaging sequences. Some packaging systems use transient transfection to produce high titers of retroviral vector for a relatively short period of time for use in animal experimentation.

Recently developed packaging cell lines are of human origin and are advantageous. The presence of human antibodies in human serum results in rapid lysis of retroviral vectors packaged in murine cell lines. The antibodies are directed against the α-galactosyl carbohydrate moiety present on the glycoproteins of murine but not human cells. This murine carbohydrate moiety is absent from retroviral vectors that are produced by human cells, which lack the enzyme α_1-3-galactosyl transferase. Human or primate-derived packaging cell lines will likely be necessary to produce retroviral vectors for in vivo administration to humans. To this point, the produc-

TABLE 4.4 Summary of Retroviral Packaging Cell Lines Used for Animal and Human Studies

Line	Plasmids That Contain Packaging Genes	Envelope Protein	Detection of Wild-Type Virus?
Ψ-2, Ψ-Am, and PA12	All contain a 5′ LTR, a deletion in the packaging signal, the *gag*, *pro*, *pol*, and *env* genes, and the 3′ LTR.	Variable	Yes
PA317 PE501	The 5′ LTR has a deletion 5′ to the enhancers, the Ψ sequence is deleted, *gag*, *pro*, *pol*, and *env* genes are present on one plasmid with intact splice signals, the PBS is deleted, and the 3′ LTR is replaced with the SV40 poly A site.	PA317: amphotropic; PE501: ecotropic	Some detected with N2; none with LN-based vectors
Ψ-CRE Ψ-CRIP	One plasmid contains a 5′ LTR, has a deletion of Ψ, expression of *gag-pro-pol* from a construct that also contains an inactive *env* gene, and has an SV40 polyadenylation site. The second plasmid has a 5′ LTR, deletion of Ψ, expression of *env* from a construct that also contains inactive *gag*, *pro*, and *pol* genes, and an SV40 polyadenylation site.	Ψ-CRE: ecotropic; Ψ-CRIP: amphotropic	Not reported
GP + E-86 GP + *env*AM 12	One plasmid has an intact 5′ LTR, the 5′ splice site, a deletion in the packaging signal Ψ, the *gag-pro-pol* gene with a small amount of the *env* gene, and the SV40 polyadenylation site. A second plasmid has an intact 5′ LTR, the 5′ splice site, the 3′ splice site, and the *env* gene.	GP + E-86: ecotropic; GP + *env*AM12: amphotropic	Reported but not verified

tion of retroviral vectors for clinical use is simple but not without challenges. A suitable stable packaging cell line containing both the packaging genes and the vector sequences is prepared and tested for the presence of infectious agents and replication-competent virus. This packaging cell line can then be amplified and used to produce large amounts of vector in tissue culture. Most retroviral vectors will produce ~1×10^5 to 1×10^6 colony forming units (cfu)/ml, although unconcentrated titers as high as 1×10^7 cfu/ml have been reported. The original vector preparation can be concentrated by a variety of techniques including centrifugation and ultrafiltration. Vectors with retroviral envelope proteins are less stable to these concentration procedures than are pseudotyped vectors with envelope proteins from other viruses. The preparations can be frozen until use with some loss of titer on thawing.

Use of Retroviral Vectors for Gene Therapy

Retroviral vectors have been extensively used in animals and substantially used in humans to determine the efficacy of gene therapy. They are the major vector that has been used for ex vivo gene therapy. Cells that have been modified ex vivo with a retroviral vector include hematopoietic stem cells, lymphocytes, hepatocytes, fibroblasts, keratinocytes, myoblasts, endothelial cells, and smooth muscle cells. Retroviral vectors have also been used for in vivo delivery. For many organs, the requirement of cellular replication for transduction poses a problem since terminally differentiated cells in organs are not proliferative. Thus, retroviral organ-based gene therapy approaches necessitate the induction of cell replication for in vivo transfer into cell types such as hepatocytes, endothelial cells, or smooth muscle cells. Alternatively, the use of viral vectors that do not require cellular replication could be used to transfer genes into nondividing cells in vivo. Studies using HIV have been initiated since that virus does not require replicating cells for transduction. Retroviral vectors have been directly injected into malignant cells in various locations, as malignant cells are highly proliferative. Efficient in vivo delivery will likely require human or primate-derived packaging cell lines or pseudotyping to prevent complement-mediated lysis in all clinical applications of retroviral gene therapy.

After transfer into a replicating cell, the expression of the retroviral vector is critical to achieve a therapeutic effect. In the application of retroviral vectors for gene therapy, the relatively low levels of gene expression achieved in animals are problematic. For currently selected genes used for gene therapy, the level of expression of the gene product does not need to be tightly regulated for clinical effectiveness. However, for diseases such as diabetes mellitus or thalassemia, the level of expression of insulin or β-globin, respectively, requires precise control. Thus, a specific clinical condition may not only require a threshold level for therapeutic effectiveness but may also require a narrow window of concentration for physiological effect. There is a paucity of quantitative data in animals regarding the levels of expression per copy from different vectors, particularly in the context of organ-specific gene expression. This is a major challenge for the field of gene therapy. The difficulties in this area are many. First, current delivery systems make the experimental determination of surviving transduced cells in situ difficult. Accurate determation of the copy number present in vivo is necessary since overall protein expression is a function of both the number of transduced cells and the gene expression per cell. Second, direct comparison of expression levels of different proteins cannot be determined for current delivery systems because of the marked differences in mRNA half-life, protein translation, and protein half-life for different genes. Third, the genomic integration site can dramatically influence the expression level. For delivery systems that modify a small number of stem cells, such as in bone marrow stem-cell-directed gene therapy (see Chapter 7), considerable variation in expression occurs based on animal species. This variation makes it essential to quantitate expression in a large number of animals and report the average results. Thus, an improved understanding of the regulatory controls of gene expression from retroviral vectors remains essential for the clinical application of gene therapy in humans. Unfortunately, expression of vectors in differentiated cell types in vitro does not accurately predict expression levels that can be achieved in vivo. In vitro screening for expression

levels provides only limited information on different retroviral vector systems in the context of human application.

An important genetic sequence or element in the gene expression from a retroviral vector is the LTR. The in vivo transcriptional activity of the LTR in bone-marrow-derived cells, liver, and muscle often attenuates over the first few weeks after transfer. However, long-term expression in some cases has been achieved. The attenuation of the LTR reflects the absence of transcription factors that are essential for expression of the LTR promoter in nondividing cells, the presence of inhibitory proteins that shut off the LTR, methylation of the LTR, or deacetylation of the associated histones. Retroviral sequences from the U3 region and the PBS can inhibit expression of the LTR in embryonic carcinoma cells by binding to proteins that inhibit transcription. These inhibitory sequences may contribute to the poor expression observed from the LTR in vivo. Retroviral vectors that alter these inhibitory sequences are expressed in vitro in embryonic carcinoma cells and may also be expressed in vivo. Methylation of the LTR is associated with loss of promoter activity. It is unclear, however, whether methylation per se is responsible for inactivation of the promoter or if methylation is a by-product of binding to the promoter.

Retroviral vectors can include an internal promoter located immediately upstream of the therapeutic gene. These "internal promoters" can be viral promoters, housekeeping promoters, or organ-specific promoters. Viral promoters were components of many first-generation vectors because they are active in most cell types in vitro. However, many of the viral promoters, such as the cytomegalovirus (CMV) promoter, are attenuated or completely shut-off in vivo in organs such as the liver. This loss of function could reflect the absence of transcription factors that are essential for expression of the promoter or the presence of inhibitory proteins that terminate viral promoter activity in nonreplicating cells. Internal promoters may also comprise the ubiquitously expressed housekeeping promoters that direct the expression of proteins required by all cells. However, housekeeping genes are often expressed at relatively low levels, and their promoters have been shown to be relatively weak in vitro and in vivo in retroviral vectors constructs. Alternatively, organ-specific promoters have two major advantages: (1) allowing limited expression to specific cell types or tissues and (2) directing high levels of gene expression. Muscle- or liver-specific enhancers and/or promoters, in comparison to housekeeping or viral promoters, direct higher levels of expression in vivo. Gene expression, in these studies, has been stable for over one year. In other studies, however, organ-specific promoters have been inactivated in vivo in transgenic mice or in a retroviral vector by the presence of adjacent retroviral sequences. These inhibitory sequences play a role in attenuation of the LTR promoter. It is also possible that these inhibitory sequences can decrease expression from adjacent internal promoters.

The control of gene expression in vivo may be an appropriate mechanism to decrease variability in expression as well as decrease the chance that the therapeutic gene is overexpressed. In clinical situations, variability or overexpression would have adverse therapeutic effects. Inducible expression systems have been developed to tightly regulate expression from a retroviral vector through responsivness to an orally administered drug. A tetracycline-responsive system can modify expression >200-fold from a retroviral vector in muscle cells in the presence of a drug when

compared to the absence of a drug in vivo. However, this system requires the all-important introduction of a drug-responsive transcription factor. This is an additional burden to the individual cell, which needs to receive and express two separate genes.

Other factors, in addition to the choice of the promoter, can influence gene expression from a retroviral vector. For some genes and through an unknown mechanism, the presence of a splice site dramatically increases the level of expression of the protein. Inclusion of genomic splice sites from the therapeutic gene is technically difficult. An intron would be efficiently removed from the RNA genome if the gene were inserted in the forward orientation. However, the gene can sometimes be packaged in the backwards orientation. In this case the mRNA for the therapeutic gene is transcribed from the opposite strand and these constructs are often unstable. Some retroviral vectors such as the MFG vector have used the retroviral splice signals that direct partial splicing of the genomic retroviral RNA.

Co-expression of two genes has many potential advantages. Through the use of a selectable marker gene and a therapeutic gene, it is possible to eliminate cells not expressing the therapeutic gene by either in vitro or in vivo selection methods. Many first-generation vector constructs express one gene from the LTR promoter and a second gene from an internal promoter. Using these vectors, however, cells selected by virtue of expression of one gene product have a lower level of expression of the second gene product. This observation was due to the phenomenon of promoter interference. An improved approach that obtains co-expression of two genes utilizes a bicistronic mRNA with an internal ribosome entry site (IRES). This enables the downstream gene to be translated in a Cap-independent fashion.

Risks of Retroviral Vectors

There are two major concerns in the use of retroviral vectors for gene therapy in humans: (1) insertional mutagenesis and (2) generation of wild-type virus. Insertional mutagenesis occurs when a retroviral vector inserts within or adjacent to a cellular gene. This insertion could result in the development of malignancy through the inactivation of a tumor suppressor gene or by activation of a proto-oncogene. The risk of developing a malignancy through the process of receiving a single copy of a retroviral vector appears to be minimal. The induction of malignancy has not been observed in animals receiving replication-incompetent retroviral vectors. This observed low incidence of mutagenesis indicates that the retroviral vector is unlikely to integrate into a genomic site that will modify cellular growth properties such as cyclins- or cyclin-dependent kinases (see Chapter 10). However, if the vector inserts into a growth-sensitive site, this would represent only the first step in a multistep process. Thus, procedures that introduce multiple retroviral vector integrations into a single cell will only increase the risk of the development of malignancy. A second safety concern regarding retroviral vectors in human use is viral recombination. Viral recombination may result in the development of replication-competent virus. This event can clearly result in the slow onset of malignancy in animals. Technical refinements in vector development have lowered the risk of generating a replication-competent virus. These include elimination of homology between the packaging genes and the vector as well as separation of the packaging genes into two or more separate pieces of DNA. However, if recombination occurs,

the extensive testing performed prior to administration of vectors to humans is an added safety measure that identifies recombinant(s). Thus, it is unlikely that replication-competent virus will be administered to humans when the appropriate safety controls are observed. It remains possible, however, that a replication-incompetent retroviral vector could recombine with endogenous viruses in vivo. Endogenous viruses are present in vivo and recombination in the human genome can generate additional pathogenic replication-competent virus(es). The occurrence can only be determined by monitoring individual gene therapy recipients for the appearance of replication-competent virus.

Summary: Retroviral Vectors

Replication-incompetent retroviral vectors can be easily generated by deleting retroviral genes and adding gene(s) of interest. Vectors can be produced in packaging cell lines that express packaging genes. The major advantage of retroviral vectors is the precise integration into a random site in the host cell chromosome. This can result in long-term survival of the gene in the transduced cell. The major disadvantage is the need to transduce dividing cells. This characteristic poses difficulties for the in vivo delivery to quiescent cells. Gene expression at therapeutic levels has been achieved from a retroviral vector in vivo in some studies for over one year, but expression has been problematic in other studies.

Lentiviral Vectors

The lentiviruses are a family of retroviruses comprising seven subgenera with specific biological properties. One such property is an advantage for its use in gene therapy, that is, the ability to transduce nondividing cells. The matrix protein and the *vpr* gene product of the lentivirus contain nuclear localization signals that allow the DNA to be transported to the nucleus without breakdown of the nuclear membrane. These gene products facilitate the infection of nondividing cells. Lentiviruses contain a number of proteins exclusive of the MLV genome (see also Chapter 11). The *tat* gene encodes a protein that stimulates expression via the *tat* response element (TAR) located in the HIV LTR. The *rev* gene encodes a protein that binds to the *rev* response element (RRE) and facilitates the transfer of unspliced RNAs to the cytoplasm. The *nef* gene encodes a protein that is localized to the inner surface of the cell membrane and can decrease the amount of the HIV cell surface receptors, such as CD4. The *nef* gene protein is important for virulence in vivo through as yet undefined mechanisms. The function of the *vif* gene is unclear. The product of the *vpu* gene appears to play a role in processing of the *env* gene product and in the efficient budding and release of virions. The *vpr* gene product contains a nuclear localization signal and may play a role in transporting HIV to the nucleus of nondividing cells. The role of the *vpx* gene product is unclear.

Several replication-defective HIV-based vectors and packaging system has been used to deliver genes to nondividing neurons, muscle, lung, endothelial cells, hematopoietic stem cells, and liver cells in vivo. One HIV packaging system contains a vector with the HIV LTRs at either end (including the TAR), an extended packaging signal, the RRE, and a reporter gene whose expression was directed by the CMV promoter. The packaging construct deleted the packaging signal and mutated the

env gene. The VSV-G envelope was expressed from a third construct. The supernatant of cells that were transfected simultaneously with all three plasmids contained retroviral particles that infected nondividing cells in vitro and in vivo. More recently, all of the accessory genes except for *tat* and *rev* have been mutated in the packaging construct, and the particles still transduced nondividing cells at the site of injection allowing for multiple exposures. Also, a new series of lentiviral vectors based on HIV-1 have been developed as a self-inactivating vector. Here, the U3 region of the 5′ LTR was replaced by the CMV promotor, resulting in *tat*-independent transcription. The self-inactivating vector was constructed by deleting 133 bp in the U3 region of the 3′ LTR including the TATA box and the binding sites for specific transcription factors. This deletion is transferred to the 5′ LTR after reverse transcription and integration into the genome of infected cells resulting in transcriptional inactivation of the LTR of the provirus. Such a self-inactivating virus transfected brain cells at a comparable level to wild-type virus.

Transduction of nondividing cells is a major advance for retroviral vector technology. Furthermore, lentivirus vectors pseudotyped with vesticular stomatitis virus G glycoprotein can transduce a wide range of nondividing cells. In addition, no inflammation is observed at the site of injection allowing for multiple exposures. It is possible that the multiple added properities of nonvirulent HIV-based vectors as described above will revolutionize human gene therapy procedures for nonreplicating cells in vivo. Three major concerns regarding these vectors remain, however. The first is the absolute assurence that recombination to generate wild-type HIV that causes immunodeficiency syndrome in a patient will not occur. Many of the HIV accessory genes can be mutated to prevent production of a functional protein. But, the complicated nature of the HIV genome and the high mutagenic rate currently made it impossible to completely assure that these accessory genes will remain nonpathogenic. Stringent tests regarding the generation of wild-type virus will be necessary prior to human use. A second concern regards the possibility of promiscuous transduction of all cell types in vivo. This may cause the unnecessary transduction of cell types where expression of the vector does not have a therapeutic effect. As noted above pseudotyping of the viral vector may limit or broaden the spectrum of cells infected. The third concern is the production of sufficient quantities of these vectors for in vivo delivery. The packaging cells currently using a transient expression system need to be enhanced.

ADENOVIRAL VECTORS

The adenovirus is a 36-kb double-stranded linear DNA virus that replicates extrachromosomally in the nucleus. The virus was first isolated from the adenoids of patients with acute respiratory infections, although it can also cause epidemic conjunctivitis and infantile gastroenteritis in humans. In patients with an intact immune system, infections are mild and self-limited. In immunosuppressed patients, however, infections can result in dissemination to the lung, liver, bladder, and kidney and can be life-threatening. Although human adenovirus type 12 can induce malignant transformation after inoculation into newborn hamsters, adenoviral DNA has not been associated with human tumors.

Adenoviral particles are 70 to 100 nm in diameter and do not contain membrane.

Over 100 different adenoviruses have been identified that infect a wide range of mammalian and avian hosts. Initial attachment of adenoviruses to cells is mediated by the fiber protein that binds to a cellular receptor. The cellular receptor has yet to be identified and may be different for different serotypes. Type-specific viral neutralization results from antibody binding to epitopes on the fiber protein and the virion hexon protein. Subsequent to initial binding, the penton base protein binds to members of a family of heterodimeric cell surface receptors known as integrins. The adenovirus:receptor complex then enters the cell via coated pits and is released into the cytoplasm from an endosomal compartment. The viral particles are transported to the nucleus via nuclear localization signals embedded in the capsid proteins. There the DNA is released in part by proteolytic degradation of the particle. The viral DNA persists during an active infection and for long periods of time in lymphocytes as a nonintegrated episome, although integration can occur during the process of transformation. Adenoviruses can transfer genetic information to a variety of cell types from many species, although they only replicate in human cells. For wild-type adenovirus, DNA replication begins ~5 h after infection and is completed at 20 to 24 h in HeLa cells, a human cervical carcinoma-derive cell line. Each cell produces 10,000 progeny virus and is lysed by their release. The production of large numbers of adenoviral particles facilitates the preparation of very high titers of adenoviral vectors.

Adenoviral Genes and Sequences Required in cis for Replication

Adenoviral genes can be transcribed from either strand of DNA and have a complex splicing pattern. There are five early transcription units, E1A, E1B, E2, E3, and E4, all of which are transcribed shortly after infection and encode several different polypeptides. Two delayed early units and the major late unit generate five families of late mRNAs. Adenoviruses also contain one or two VA genes that are transcribed by RNA polymerase III and serve to block host cell translation.

The E1A region codes for two E1A polypeptides. E1A polypeptides can activate transcription by binding to a variety of different cellular transcription factors and regulatory proteins, including the retinoblastoma gene product *Rb*. E1A induces the cell to enter the cell cycle, which is necessary for replication of adenoviral DNA. The E1B 55-kD protein binds to p53 and prevents p53 from blocking progression through the cell cycle or inducing apoptosis. The E1B 19-kD protein blocks apoptosis by an as yet unknown mechanism. The E2 region encodes three different proteins, all of which function directly in DNA replication. The E2-encoded terminal protein is an 80-kD polypeptide that is active in initiation of DNA replication. It is found covalently attached to the 5′ ends of the viral DNA. The other E2-encoded proteins include a 140-kD DNA polymerase and a 72-kD single-stranded DNA binding protein. The E3 region encodes proteins that modify the response of the host to the adenovirus. The E3-gp 19-kD protein binds to the peptide-binding domain of MHC class I antigens and causes retention of class I antigen in the endoplasmic reticulum. The E3 14.7-kD protein, or the complex of E3 14.5-kD/E3 10.4-kD proteins prevent cytolysis by tumor necrosis factor. The E4 unit encodes proteins that regulate transcription, mRNA transport, and DNA replication. Of the 11 virion proteins, 7 are located in the outer shell and 4 are present in the core of the virion. These are primarily encoded by the late genes.

There are two sequences that need to be supplied in *cis* for viral replication: (1) the 100- to 140-bp inverted terminal repeats at either end of the linear genome and (2) the packaging signal, which is adjacent to one of the inverted terminal repeats. The 5′ ends of the viral DNA have a terminal protein of 80 kD covalently attached via a phosphodiester bond to the 5′ hydroxyl group of the terminal deoxycytosine residue. The terminal protein serves as a primer for DNA replication and mediates attachment of the viral genome to the nuclear matrix in cells. Inverted repeats enable single strands of viral DNA to circularize by base pairing of their terminal sequences. The resulting base-paired panhandles are thought to be important for replication of the viral DNA. The packaging sequence, located at nucleotide 194 to 358 at the left end of the chromosome, directs the interaction of the viral DNA with the encapsidating proteins.

Use of Adenoviral Sequences for Gene Transfer

The observation that E1A- and E1B-deficient adenoviruses are propagated in 293 cells paved the way for the development of adenoviral vectors. The 293 cells are a human embryonic kidney cell line that contains and expresses the Ad5 E1A and E1B genes. Early first-generation adenoviral vectors replaced a 3-kb sequence from the E1 region with a promoter and a gene of interest, as shown in Figure 4.4. In addition to providing space for the therapeutic gene, deletion of the E1 region removed oncogenes that might contribute to malignancy. Although the early

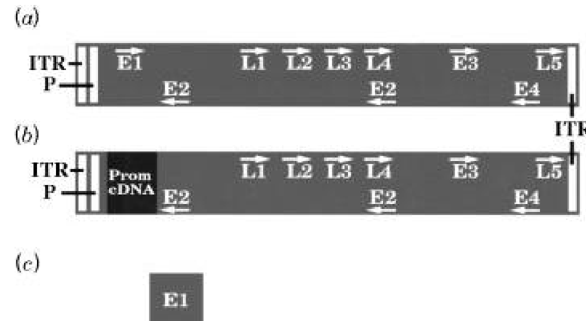

FIGURE 4.4 Adenoviral vectors. (*a*) Wild-type adenovirus. Adenoviruses contain a double-stranded linear DNA genome of ~36 kb. The inverted terminal redundancies (ITRs) of ~100 base pairs at either end are necessary for replicating the DNA. The packaging signal (P) is necessary for the viral DNA to get packaged into a viral particle. Multiple early (E) and late (L) genes code for proteins that are necessary for replicating the DNA and producing an infectious adenoviral particle. (*b*) Adenoviral vector. Most adenoviral vectors have deleted the E1 gene and replaced it with a promoter and therapeutic gene. This results in a vector that still contains most of the adenoviral genes. Other adenoviral vectors that are not shown here have deleted additional adenoviral genes from the E2, E3, or E4 region. (*c*) Packaging cells. The adenoviral vector alone cannot produce adenoviral particles because it does not contain the E1 gene. Packaging cells that express E1 and contain the adenoviral vector sequences are necessary for producing adenoviral particles that can transmit information to a new cell. E2 or E4 also need to be expressed in packaging cells that are used to produce E2- or E4-deleted adenoviral vectors.

adenoviral vectors resulted in high levels of expression in a variety of organs at early time points in animals, expression was transient. The transient expression was primarily a result of an immune response targeted to cells that express the residual adenoviral vector proteins. This observation led to further manipulations of the adenoviral vector genome in an attempt to stabilize the vector in vivo and reduce the inflammatory response.

Later generations of adenoviral vectors have deleted E2, E3, or E4 in addition to E1 in an attempt to decrease the expression of late genes and the subsequent immune response. An added advantage of the manipulation is the additional space for the therapeutic gene. E2- or E4-deleted adenoviral vectors require cell lines that express E2 or E4 in addition to E1. The E3-deleted adenoviral vectors can still be produced in 293 cells, since the E3 region does not encode any genes that are essential for replication in vitro. The products of the E2 gene include a 72-kD single-stranded DNA binding protein, which plays a role in both DNA replication and viral gene expression. An adenoviral vector that contained a mutation in the E2A gene has resulted in the generation of a temperature-sensitive single-stranded DNA binding protein. Use of this vector construct results in prolonged expression of the therapeutic gene, decreased expression of the late adenoviral vector genes, and a delayed inflammatory response. However, even in the latter case expression still did not extend beyond 100 days. Deletion of the E4 region has led to increased stability of the adenoviral DNA in vivo, with a loss of expression from the CMV promoter in the liver. Deletion of the E3 region has decreased the stability of the adenoviral vector in vivo. This E3 region helps the virus to avoid the immune system of the host by blocking class I MHC presentation of viral antigens, and thus deletion of this region promotes antigen presentation and host immunity.

The removal of all adenoviral proteins creates a so-called gutless adenoviral vector. The purpose of this line of investigation is to eliminate the expression of the adenoviral proteins in vivo in order to prevent a host immune response. Gutless adenoviral vectors have been generated in which the inverted terminal repeats and the packaging signal remains, but all adenoviral coding sequences have been removed and replaced with the therapeutic gene. Unfortunately, these vectors have not resulted in prolonged expression in vivo. It is possible that the adenovirus contains other sequences that are necessary for long-term extrachromosomal maintenance of the DNA in cells.

Preparation of recombinant adenoviral vectors for clinical use is somewhat more complicated than is the production of retroviral vectors. The 293 cells are a human embryonal kidney cell line that expresses the E1 genes and are commonly used to propagate E1-deficient adenoviral vectors. The large size of the adenovirus (~36 kb) makes cloning by standard methods difficult due to the paucity of unique restriction sites. Most genes are inserted into the adenoviral vector by homologous recombination between a transfer vector and the helper vector in cells that express any necessary proteins in trans. The transfer vector contains the therapeutic gene flanked by adenoviral sequences on a plasmid that contains a bacterial origin of replication, and this can be propagated in bacteria. The helper virus contains all of adenoviral genes except those that are supplied in trans by the packaging cells. In some cases, the helper virus can be propagated in 293 cells and therefore must be restricted prior to co-transfection with the transfer vector to decrease the number

of nonrecombinants that are obtained. For other helper vectors, such as pJM17, the helper vector is present on a plasmid with a bacterial origin of replication inserted in the E1 region. This can be propagated in bacteria but is too large to be packaged into an adenoviral particle. After co-transfection of the transfer vector and the helper vector into 293 cells, homologous recombination results in the insertion of the therapeutic gene and deletion of the bacterial origin of replication. The resulting vector can be packaged. Recombinants that replicate in 293 cells result in cell death that appears as a plaque on a lawn of viable cells. Plaques are screened for the presence of the therapeutic gene and the absence of the helper vector. A therapeutic gene of up to 8 kb can be inserted into an adenoviral vector.

To produce large amounts of the adenoviral vector, packaging cells are infected with the plaque-purified adenoviral vector. When a cytopathic effect is observed, the cells are broken up and the adenoviral vector is purified from the cellular debris using a variety of techniques including $CsCl_2$ gradients and column chromatography. Titers of up to 10^{12} plaque forming units (pfu)/ml can be obtained and are stable to freezing. Preparations must be tested for the presence of wild-type adenovirus or other pathogens prior to use in humans.

Use of Adenoviral Vectors for Gene Therapy

Adenoviral vectors have been used to transfer genes in vivo into the lung, liver, muscle, blood vessel, synovium, eye, peritoneum, brain, and tumors in animals. The titers that can be achieved enable a high percentage of the cells to be transduced as well as express elevated levels of the transgene. A major limitation of adenoviral vectors is the transgene expression for less than one month primarily due to an immune response to the remaining viral proteins. This targeted specific immune response rapidly eliminates the transduced cells. This immune response can also result in severe inflammation at the site of delivery and organ dysfunction. Furthermore, the vigorous host immune response to the surface proteins of the adenovirus diminishes the efficacy of repeat administration.

A strategy to prolong gene express is to inhibit the immune response to the adenoviral vector. Studies in immunodeficient mice have demonstrated that in the absence of antigen-specific immunity, gene expression is prolonged and secondary gene transfer is possible. MHC class I-restricted $CD8^+$ cytotoxic T lymphocytes are the primary effector cells for the destruction of adenoviral infected cells in the mouse. The use of immunosuppressive therapy could provide persistent gene expression following adenovirus-mediated gene transfer and allow secondary gene transfer. A variety of approaches to suppress the immune response have been taken. These include immunosuppression with drugs such as cytoxan or cyclosporine, or inhibition of the CD28:B7 costimulatory response using a soluble form of murine CTLA4Ig. Injection of adenoviral vector into neonates or into the thymus, resulting in tolerization, allows subsequent injection of an adenoviral vector into adults without immune rejection.

Evaluation of gene expression from adenoviral vectors has been complicated by its instability. Many studies have not differentiated between loss of DNA and loss of gene expression. Some studies have demonstrated relatively long-term expression from the CMV promoter of an adenoviral vector in the liver in vivo. These

studies contradict the results obtained using a retroviral vector, in which the CMV promoter was rapidly shut-off. However, it was subsequently demonstrated that deletion of the E4 region of the adenovirus led to loss of expression from a CMV promoter in an adenoviral vector in the liver in vivo. It is therefore likely that the deletion of other early genes might modulate expression of an adenoviral vector in vivo. Studies have demonstrated that the housekeeping promoter elongation factor 1 was more active than the CMV promoter. The CMV-enhancer–β-actin-promoter combination was more active than the SRα promoter. Additional experiments in which transgene expression is followed over time and normalized to the adenoviral vector copy number in various organs will be necessary to optimize expression levels in vivo.

Risks of Adenoviral Vectors

There are three potential risks of adenoviral vectors: (1) the development of organ inflammation and dysfunction due to the immune response to adenoviral vector-transduced cells, (2) the development of tolerance to an adenoviral vector that could result in fulminant disease upon infection with wild-type virus, and (3) the development of wild-type virus. Early generation adenoviral vectors were toxic when administered at high doses. For example, one patient with cystic fibrosis who received an adenoviral vector to the lung had a severe inflammatory response. It is likely that decreasing the immunogenicity of adenoviral vector-transduced cells or modulating the immune response will decrease this inflammation. Whether limited organ-based inflammation will be acceptable is an open question. The risks of modulating the immune response to an adenoviral vector have not been adequately studied. It is possible that immunomodulation will predispose to fulminant disease upon infection with wild-type adenovirus of the same serotype. These risks cannot be assessed in animal models where the adenovirus does not replicate. The third risk of using adenoviral vectors is the generation of wild-type virus in vivo. This also could lead to fulminant infection if immunomodulation has led to tolerance. It is less likely that development of wild-type adenovirus would contribute to malignancy since the virus does not integrate.

Summary: Adenoviral Vectors

In summary, adenoviral vectors result in high-level expression in the majority of cells of many organs for 1 to 2 weeks after transfer. Gene transfer occurs in nondividing cells, a major advantage over most retroviral vectors. However, expression is transient in most studies. This is due primarily to an immune response. The instability of expression is a serious impediment to the use of adenoviral vectors in the treatment of monogenic deficiencies. It is less of a problem for gene therapy approaches for cancer that require short-term expression. The immune response to adenoviral-transduced cells can lead to organ damage and has resulted in death in some animals. Any preexisting or induced antiadenovirus neutralizing antibodies could prevent an initial or subsequent response to adenoviral treatment. Modification of the adenoviral vector to decrease its immunogenicity or suppression of the recipient's immune response may prolong expression and/or allow repeated delivery to patients.

ADENOVIRUS-ASSOCIATED VIRUS

Adenovirus-associated virus (AAV) is a 4.7-kb single-stranded DNA virus that replicates in the nucleus in the presence of adenovirus and integrates into the chromosome to establish a latent state. It was first discovered as a satellite contaminant in human and simian cell cultures infected with adenovirus. AAV has not been associated with disease in humans, although up to 90% of all humans have evidence of prior infection with some serotypes of AAV. Humans are frequently seropositive for AAV2 and AAV3, while evidence of prior AAV5 infection is infrequent. AAV particles are 18 to 26 nm in diameter and do not contain membrane. They enter the cell by receptor-mediated endocytosis and are transported to the nucleus. Although the receptor has not yet been cloned, entry occurs in a wide range of mammalian species. Wild-type AAV integrates as double-stranded DNA into a specific region of chromosome 19. AAV can also be maintained in an extrachromosomal form for an undefined period of time.

AAV Genes

The AAV genome has two major open reading frames, as shown in Figure 4.5. The left open reading frame extends from map position 5 to 40 and encodes the Rep proteins. The right open reading frame extends from map position 50 to 90 and encodes the AAV coat proteins. The *rep* gene was so named because its products

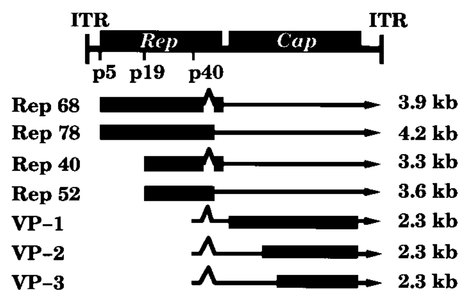

FIGURE 4.5 Map of the AAV genome. The AAV2 genome of 4.8 kb has 100 map units. AAV has inverted terminal repeats (ITRs) of 145 nt at either end, which contain sequences necessary for DNA replication and packaging into virions. There are 3 promoters at map position 5, 19, and 40, which are designated p5, p19, and p40, respectively. These is an intron at map position 42 to 46, which may or may not be utilized, resulting in 2 transcripts that derive from each promoter. There is a polyadenylation site at map position 96, which is used by all transcripts. The p5-initiated proteins Rep 68 and Rep 78 are necessary for replication and for transcriptional regulation of AAV and heterologous viral and cellular promoters. The p19-derived proteins Rep 40 and Rep 52 are required for accumulation of single-stranded DNA. For the p5- and p19-derived transcripts, the unspliced species is the major mRNA. The AAV *cap* gene encodes the structural AAV capsid proteins, which are transcribed from the p40 promoter. VP-1 is derived from an alternatively spliced mRNA that uses an AUG for translational initiation. VP-2 is derived from the more common splice product and utilizes the nonconsensus ACG as the translational initiation site. VP-3 is derived from the most common splice product and uses the consensus AUG for translational initiation. The size of each RNA is shown on the right.

are required in trans for DNA replication to occur. Rep 68/78 is an ATPase, helicase, site-specific endonuclease and transcription factor. Rep 68/78 plays a critical regulatory role in several phases of the AAV life cycle. It is necessary for site-specific integration into the host cell chromosome and to establish a latent infection. Rep 68/78 binds to a dodecamer sequence $(GCTC)_3$ in the stem of the ITR and causes a nick in the DNA. The latter is essential for replication of the DNA. A region of chromosome 19 also contains the AAV Rep protein binding sequence $(GCTC)_3$ responsible for region-specific integration. Integration can occur within several hundred nucleotides of this recognition site. In the presence of helper virus, Rep 68/78 is a transactivator at all three AAV promoters, p5, p19, and p40. In the absence of co-infection with a helper virus, Rep68/78 negatively regulates AAV gene expression. Although the functions of the smaller 52- and 40-kD Rep proteins are not totally clear, each are necessary for the accumulation of single-stranded genomic DNA. The *cap* gene codes for the capsized proteins, VP-1 of 87 kD, VP-2 of 73 kD, and VP-3 of 62 kD. VP-2 and VP-3 are initiated from different transnational start codons of the same mRNA, while VP-1 is translated from an alternatively spliced mRNA. Although VP-3 is the most abundant protein, VP1, 2, and 3 are required for infectivity.

Sequences Required in *cis* for Replication

AAV has an inverted terminal repeat of 145 nt at both ends that is required in cis for DNA replication, encapsidation, and integration. The first 125 bases contains a palindromic sequence that forms a T-shaped structure, as shown in Figure 4.6. Replication begins in the ITR where a stable hairpin is formed, leading to self-priming from the 3′ end and replication using a cellular DNA polymerase. Rep 68/78 nicks the parental strand in the ITR as shown in Figure 4.6*c*, which allows filling in of the bottom strand. When capsid proteins are expressed, capsid assembly leads to displacement and sequestration of single-stranded AAV genomes. Single stands of either polarity can be packaged into AAV particles.

Helper Functions of Other Viruses

AAV are unique in that they usually require co-infection with another virus for productive infection. The helper (co-infection) virus is usually adenovirus or herpes simplex virus. Cytomegalovirus and pseudoradies virus can also function as a helper virus. Treatment of cells with genotoxic agents such as ultraviolet irradiation, cycloheximide, hydroxyurea, and chemical carcinogens can also induce production of AAV, albeit at low levels. The helper functions of adenovirus requires the early but not late genes. E1A is required for AAV transcripts to be detected and presumably activates transcription of the AAV genes. The E4 35-kD protein forms a complex with the E1B 55-kD protein and may regulate transcript transport. The E2A 72-kD single-stranded DNA binding protein stimulates transcription of AAV promoters and increases AAV DNA replication, but it is not absolutely required for AAV replication. The adenovirus VAI RNA facilitates the initiation of AAV protein synthesis. The helper functions provided by HSV-1 have been less clearly defined. Two studies indicate that the ICP-8 single-stranded DNA protein is required.

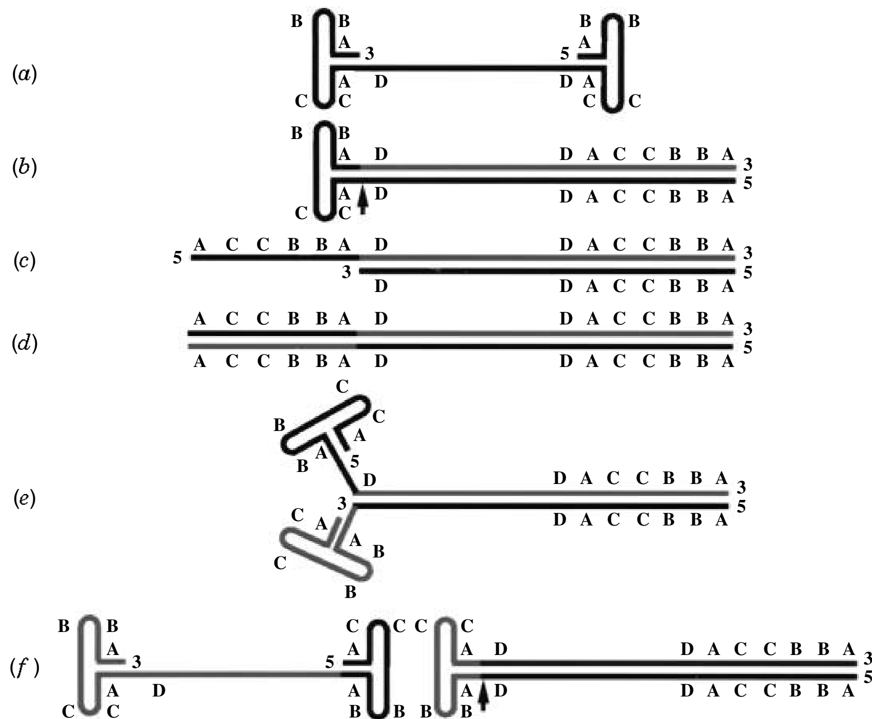

FIGURE 4.6 Mechanism of replication of AAV DNA. AAV has a single-stranded DNA genome (shown in black) with inverted terminal repeats (ITRs) at either end. (*a*) Structure of the single-stranded genomic DNA. The ITRs are palindromic and form a T-shaped structure at either end. The 3′ end is double stranded and thus can serve as a primer for the initiation of DNA synthesis. (*b*) Elongation of the 3′ end. A cellular DNA polymerase initiates DNA synthesis at the 3′ end and copies the DNA up until the 5′ end of the genomic DNA. The arrow designates the site at which Rep will cleave the DNA. (*c*) Endonucleolytic cleavage of the genomic DNA. The viral protein Rep performs an endonucleolytic cleavage of the DNA. The T-shaped structure can be unfolded to result in the structure shown. (*d*) Elongation of the DNA to generate a double-stranded unit length intermediate. DNA polymerase initiates polymerization at the free 3′ end, resulting in the synthesis of a full-length double-stranded intermediate. Note that the B and C sequences have become inverted relative to their initial orientation. This is designated as the "flop" orientation, while the initial structure shown in (*a*) in which the B sequence was closer to the terminus is designated as the "flip" orientation. Either orientation can be packaged into a viral particle. (*e*) Isomerization. The left end of the double-stranded intermediate can isomerize to form the structure shown. Alternatively, the right end of the double-stranded intermediate could isomerize to form a similar structure (not shown here). (*f*) Continued DNA synthesis to release a single-stranded genomic DNA and a covalently linked double-stranded intermediate. The free 3′ end primes synthesis of new DNA. This results in the release of a single-stranded genomic DNA that can be packaged into a viral particle. The double-stranded DNA intermediate shown here is homologous to the intermediate shown in (*b*) and can be cleaved by Rep to generate a free 3′ end and undergo the subsequent steps shown in (*c*) through (*f*). These steps would return the DNA to the original "flip" orientation.

There are discrepancies as to the function of the ICP4 transactivator, the DNA polymerase, and various submits of the helicase–primase complex.

Use of AAV Sequences for Gene Transfer

AAV vectors, like retroviral vectors, can be deleted of all coding sequences and replaced with a promoter and coding sequence of interest, as shown in Figure 4.7. This process eliminates the immune response to residual viral proteins. The most common method for packaging AAV vectors involves co-transfection of an ITR-flanked vector-containing plasmid and a *rep-cap* expression plasmid into adenoviral-infected 293 cells. A cloned duplex forms containing ITRs and results in the production of the single-stranded DNA genome. *Rep* and *cap* genes are expressed from a packaging plasmid not containing ITRs and thus cannot replicate or be packaged into a viral particle.

Wild-type AAV integrates within a specific region of several hundred nucleotides on chromosome 19. AAV vectors do not integrate specifically because they do not express the Rep protein. Upon integration, the viral termini are extremely heterogeneous, and significant deletions are common. AAV vectors can also integrate as a tandem head-to-tail array. Episomal forms of AAV have been found after up to 10 passages.

The production of large quantities of AAV vector for clinical use has been problematic. Large-scale preparation of the ITR-containing plasmids in bacteria is difficult since the palindromic sequences are subject to deletion. The toxicity of the

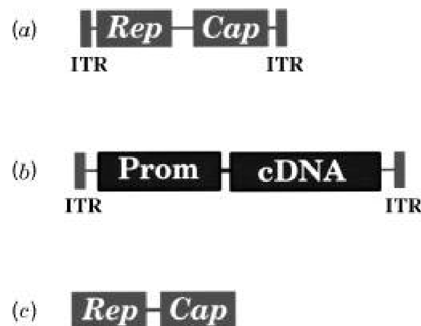

FIGURE 4.7 Adenovirus-associated virus (AAV) vectors. (*a*) Wild-type AAV. AAV contain a single-stranded DNA genome of 4.7 kb. The inverted terminal repeats (ITRs) are necessary for conversion of the single-stranded genome to double-stranded DNA, packaging, and for integration into the chromosome. The protein products of the *rep* and *cap* genes are necessary for replicating the AAV genome and for producing an AAV particle. (*b*) AAV Vector. AAV vectors have deleted the AAV coding sequences and replaced them with a promoter and therapeutic gene. They still contain the ITRs which are necessary for the vector to transmit its genetic information into a target cell. (*c*) Packaging Cells. The AAV vector alone cannot produce an AAV particle because the *rep* and *cap* genes are not present. These AAV genes need to be present in a packaging cell line along with the AAV vector in order to produce an AAV particle that can transfer genetic information into a cell. In addition, another virus such as an adenovirus needs to be present for the production of infectious particles.

Rep proteins limits the generation of stable mammalian packaging lines that can be used to propagate the vector. To produce AAV vectors, most investigators have used transient transfections with two plasmids in combination with infection with an adenoviral vector. However, the number of recombinant AAV vector particles produced by packaging cells is lower than the amount of wild-type AAV that can be produced. The lack of production may reflect the fact that Rep and Cap proteins are limiting since their plasmid does not contain ITRs and is not amplified. After recombinant AAV particles are produced, they must be separated from adenovirus and cellular components for the isolation of a nontoxic vector. Methods for separation of AAV vector from adenovirus include heat inactivation of adenovirus, $CsCl_2$ banding, and ion-exchange chromatography. AAV vector preparations are stable to freezing and must be tested for wild-type AAV, adenovirus, and other pathogens prior to use.

Use of AAV Vectors for Gene Therapy

A major advantageous characteristic of AAV vectors is their ability to transduce nondividing cells. AAV vectors have been used to transfer genes into a variety of cell types including hematopoietic stem cells in vitro and hepatocytes, brain, retina, lung, skeletal, and cardiac muscle in vivo. Stable expression has been observed for up to one year in several organs. It is not yet clear if the AAV vectors integrate into the host cell chromosome or are maintained episomally. Studies in a variety of animal models indicate that AAV-transduced cells do not elicit an inflammatory reaction or a cytotoxic immune response.

Some studies have suggested that AAV transduction efficiency increases when cells are replicating, or treated with cytotoxic agents, or co-infected with an adenoviral vector. However, such procedures did not increase the copy number of the AAV vector in experimental studies. The data indicate the techniques increase the number of cells that expressed the reporter gene through activation of the viral promoter of the AAV vector rather than increasing the transfer of genetic material into the cells.

Little information is available regarding the level of expression per copy from an AAV vector in various cell types in vivo. ITRs have transcriptional activity and have been utilized to direct expression of the cystic fibrosis transmembrane receptor. Most AAV vectors utilize an internal promoter to direct expression of the therapeutic gene. The CMV promoter functions at levels sufficient to produce detectable protein product in muscle and brain. But it is poorly functional in the liver in vivo. Use of the LTR promoter from the MFG retroviral vector resulted in a high-level expression in the liver. However, an LTR promoter in another context was much less active. It is possible that the presence of a splice site in the MFG-derived vector accounts for this discrepancy. These studies indicate that it will be necessary to empirically test different constructs in vivo for their relative efficacy.

It is possible that residual AAV sequences will not have the inhibitory effect that occurs for some internal promoters of a retroviral vector. However, expression from an internal promoter of an AAV vector can attenuate in vitro by a process that involves histone deacetylation. In addition, the ITRs have transcriptional activity and may be subject to inhibitory factors. Recently, a protein has been identified as the single-stranded D-sequence-binding protein whose phosphorylation and ITR-

binding activity is modulated by the cell cycle. Binding of the phosphorylated protein to the ITR inhibited replication of the DNA and might influence transcription. It is therefore possible that AAV vector sequences will attenuate expression from some internal promoters.

A noteworthy feature of AAV vectors is the slow increase of gene expression over several weeks after delivery to the liver or the muscle. Such an increase may represent conversion of the single-stranded DNA genome to double-stranded DNA. The DNA is maintained as a concatemer in an episomal or integrated state. Expression has been stable for up to one year in liver and muscle, implying that the DNA is not lost or inactive. Longer-term evaluation and determination of the status of the DNA is required in future studies.

Risks of AAV Vectors

There are three potential risks of AAV vectors: (1) insertional mutagenesis, (2) generation of wild-type AAV, and (3) administration of contaminating adenovirus. It is theoretically possible that AAV vectors could activate a proto-oncogene or inactivate a tumor suppressor gene by integration into the chromosome in vivo. However, AAV vectors have not been reported to result in malignancy. Wild-type AAV could be produced when recombination between the vector and the packaging plasmid occurred. However, since AAV has not been shown to be pathogenic and is not capable of efficient replication in the absence of a helper virus, the generation of wild-type AAV may not be a serious concern in human gene therapy. A final potential problem is a helper virus contaminating preparations of AAV vector and causing adverse effects. Careful testing of AAV vectors for the presence of the helper virus would reduce this risk. It therefore appears that AAV vectors can be considered relatively safe, although further long-term studies in animals are necessary.

Summary: AAV Vectors

AAV vectors can be generated by removing viral genes and replacing them with a promoter and therapeutic gene. They can be produced in cells expressing the AAV *rep* and *cap* genes and that have been co-infected with a helper virus such as adenovirus. Production of large amounts of AAV vector is difficult. The major advantage of AAV in gene therapy is the ability to transfer genetic information into nondividing cells in vivo. In addition, expression has been maintain for up to one year. Further experiments to determine if AAV vectors integrate or are maintained in an episomal state are necessary. A current major limitation of AAV vectors is that they cannot accommodate more than 4.5 kb of exogenous genetic material.

HERPES SIMPLEX VIRUS 1

Herpes simplex virus 1 (HSV-1) has a 152-kb double-stranded linear DNA genome that can be maintained episomally in the nucleus of cells. It can cause mucocutaneous lesions of the mouth face, and eyes and can spread to the nervous system and cause meningitis or encephalitis. The related HSV-2 can cause lesions in the

genitalia. HSV can establish a lifelong latent infection in neurons without integrating into the host cell chromosome.

The HSV-1 virion is enveloped and ~110 nm in diameter. Viral infection is initiated in epithelial cells of the skin or mucosal membranes by binding of the viral envelope glycoproteins to heparin sulfate moieties on the plasma membrane. Specific attachment can be mediated by a novel member of the tumor necrosis factor (TNF)/nerve growth factor (NGF) receptor family, which triggers fusion of the virus envelope with the plasma membrane. After the initial rounds of replication, the virus is taken up into the axon terminals of neurons innervating the site of primary infection. The viral capsid is transported to the nucleus via a process that probably involves the cytoskeleton. For neurons, this process results in the retrograde transport of viral particles long distances within the axon. Upon entering a cell, the virus can enter a lytic cycle, resulting in cell death within 10 h, or can enter a latent phase in the nucleus.

HSV Genes

The viral genome consists of a long and short unique region, designated U_L and U_S, respectively, each flanked by inverted repeats. Transcription of early genes is initiated by VP16, a potent transcription factor present in the virion. These early gene products lead to replication of the viral DNA, followed by expression of the late genes. HSV-1 has at least 81 gene products, 43 of which are not essential for replication in vitro but contribute to the virus life cycle in vivo. During the latent state, however, no HSV proteins are detected. Instead, a family of RNAs, the latency-associated transcripts (LAT), are present in the nucleus. The roles of these transcripts are unknown. The virus can establish latency without the LATs.

Sequences Required in cis for Replication

HSV-1 contains three origins of replication (see Chapter 2). One is located in the middle of the U_L region (OriL) and two are within the inverted repeats flanking the U_S region (OriS). Only one replication origin needs to be present on a circular piece of double-stranded DNA to support replication. The viral DNA is packaged via the packaging signal, a sequence which is located in the genomic termini. Both origin of replication and the packaging signal are sufficient to allow a circular piece of DNA to be replicated and to be packaged by cells that express the remainder of the essential HSV-1 proteins in trans.

Use of HSV Sequences for Gene Transfer

Most vectors based upon HSV-1 have deleted one or more genes necessary for replication. Genes coding for proteins necessary for replication such as infected cell polypeptide (ICP)4 can be deleted. HSV-1 particles are produced in cells that express these proteins in trans. HSV-1 vectors can accommodate up to 25 kb of foreign DNA sequences and can establish latency. However, these viral vectors are toxic for some cells in vitro and can cause encephalitis when administered to the brain at high doses.

An alternative type of HSV-1-based vector is an amplicon. Amplicons contain

bacterial and HSV origins of replication, as well as the packaging sequence. If an amplicon is present in cells that also contain wild-type HSV, the amplicon will be replicated along with the wild-type virus and then packaged into viral capsids. It is difficult, however, to separate the amplicon from the wild-type virus. One approach to circumvent this problem is to co-transfect cells with an amplicon and a series of cosmids that contained all the HSV-1 coding sequences, except for the packaging signals. Amplicons have been used to express genes for up to 1 month in the brain (see Chapter 9).

The insertion of a therapeutic gene into HSV-1 vectors requires homologous recombination, using procedures that are similar to those described for adenoviral vectors. HSV-1 vectors that have deleted HSV genes are produced in cell lines that express the deficient protein in trans. HSV-1 amplicons genes are expressed in cells that are co-infected with a replication-competent HSV-1 genome or have the HSV-1 genes introduced on multiple cosmids.

Use of HSV Vectors for Gene Therapy

HSV vectors have been used to transfer genes into the brain, spinal cord, and muscle but have not been used in humans for gene therapy. Delivery into the central nervous system has utilized stereotactic injection. Transduced cells have been limited to a relatively small region because the virus does not readily diffuse. Delivery of HSV-1-based vectors to the muscle has resulted in only short-term expression due to the cytopathic effects and/or the immune response to the residual HSV-1 proteins. These results with HSV-1 in muscle are similar to what has been observed with the adenoviral vectors in many organs.

Expression from an HSV Vector in vivo

A number of promoters are active in vivo when lytic infection occurs. However, stable expression during latency from an HSV-1-based vector has only been detected in the brain using the LAT promoter. A variety of others including viral, RNA polymerase III-activated, housekeeping, and neuronal promoters are shut down in vivo. The LTR, LAP, or a neuronal-specific promoter have resulted in stable expression in dorsal root ganglion neurons of the spinal cord.

Risks of HSV Vectors

There are two major risks of HSV-1-based vectors: (1) toxicity due to the cytopathic effect of relatively unattenuated virus and (2) the development of wild-type virus. Administration of high doses of HSV-1 vectors with only a single gene deleted had a considerable cytopathic effect. HSV-1 vectors with deletion of four genes had less toxicity. The development of wild-type HSV-1, which can cause serious infections such as encephalitis, is a concern.

Summary: HSV Vectors

HSV-1 vectors can be generated by deleting genes that are essential for replication, inserting a therapeutic gene into a nonessential region, and transferring the DNA into cells that supply the essential HSV-1 protein(s) in trans. HSV-1 amplicons can

be generated by placing the therapeutic gene on a plasmid with the HSV-1 origin of replication and packaging signal and transferring the DNA into a cell along with the essential wild-type HSV-1 genes. HSV-1 vectors have resulted in stable expression in the brain with the LAT promoter. Toxicity due to the HSV-1 vector and the generation of wild-type virus are a concern.

OTHER VIRAL VECTORS

There are other viruses that have been used as vehicles for gene transfer. The baculovirus is an 80- to 230-kb double-stranded circular DNA virus that replicates in insect cells in vitro or in vivo. The vaccinia virus is a 191-kb double-stranded DNA that was used in the past to vaccinate humans against the related smallpox virus. It has over 198 open reading frames and ~50 kb of the genome is not essential for replication in vitro. Genes can be inserted into a vaccinia genome by homologous recombination. Recombinant vaccinia has been used for immunization against proteins that play an important role in the pathogenesis of encephalitis, rabies, and other infectious diseases. It has also been used to express cytokine genes in animals in an attempt to boost the immune response to a cancer. An advantage of vaccinia-derived vectors is their ability to accommodate a large amount of exogenous genetic material. Disadvantages include the fact that an immune response to the vector will preclude gene transfer in some patients and will limit the duration of gene expression.

Baculoviruses can transfer genetic information into hepatocytes but do not express the baculoviral genes, which require transcription factors that are only present in insect cells. A mammalian promoter and gene of interest can be expressed, however. Baculoviral vectors have been used to express genes in hepatocytes in vitro and have been delivered to intact livers using an isolated perfusion system. Advantages of baculoviral vectors include the ability to accept large amounts of genetic material and the absence of expression of baculoviral proteins in mammalian cells. Disadvantages include the transient gene expression and the sensitivity of the vector to complement.

NONVIRAL VECTORS

Nonviral vectors include any method of gene transfer that does not involve production of a viral particle. They can be divided into two classes: (1) RNA or DNA that can be amplified in bacteria or eukaryotic cells, and whose transfer into a cell does not involve a viral particle, or (2) oligodeoxynucleotides or related molecules synthesized chemically. Nonviral vectors amplified in cells generally encode a gene that expresses the therapeutic protein, although they can encode antisense RNA that acts by blocking expression of an endogenous gene. Oligonucleotides act by altering expression of endogenous genes in cells by a variety of mechanisms (see Figure 4.8).

There are three important factors regarding nonviral vectors that can be amplified in prokaryotic or eukaryotic cells for gene therapy: (1) the size of the insert accepted, (2) how to get the genetic material into cells efficiently, and (3) how to maintain the genetic material inside the cell in order to achieve long-term expression.

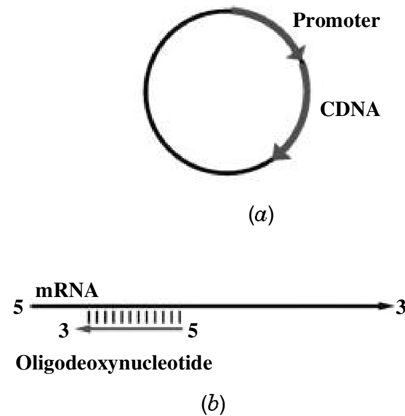

FIGURE 4.8 Nonviral vectors for gene therapy. Nonviral vectors are any type of vector that does not involve a viral particle that can alter gene expression in a cell. (*a*) Plasmid DNA. Plasmids are double-stranded circles of DNA that replicate efficiently in bacteria. They can contain up to 15 kb of exogenous genetic information. They generally contain a promoter and coding sequence that results in expression of a therapeutic protein. Although plasmid DNA does not enter cells efficiently because of its large size, cationic liposomes, or receptor-mediated targeting can be used to facilitate its entry into cells. (*b*) Oligonucleotide vectors. Oligonucleotides, or more stable analogs such as phosphorothioates, contain 10 to 25 bases. An oligonucleotide is shown hybridized with RNA in this panel, which can affect the processing, translation, or stability of the RNA. Oligonucleotides can also form a triple helix with DNA and alter transcription, or serve as a decoy by binding to transcription factors and prevent them from binding to endogenous genes.

Size of Insert

The size of the insert accepted varies considerably among the different nonviral vectors that replicate in cells. Bacteria can amplify plasmids, bacteriophage, cosmids, or bacterial artificial chromosomes. All can be purified from cells as nucleic acid devoid of proteins. Plasmids are double-stranded circular DNA molecules that contain a bacterial origin of replication. They can accommodate up to 15 kb of exogenous genetic information. Bacteriophage is a double-stranded linear DNA virus that can accommodate up to 20 kb of foreign DNA. Cosmids are modified plasmids that carry a copy of the DNA sequences needed for packaging the DNA into a bacteriophage particle. They can accommodate up to 45 kb of genetic information. Bacterial artificial chromosomes (BACs) contain elements from a normal chromosome that allow it to replicate and to segregate appropriately and can accommodate up to 100 kb of exogenous genetic material. Yeast artificial chromosomes (YACs) contain telomeres, replication origins, and sequences that ensure appropriate segregation in yeast cells. They can accommodate up to 1000 kb of exogenous genetic material. They do not replicate in mammalian cells. More recently, the production of a human artificial minichromosome was reported, although its transfer into cells was very inefficient. The most successful use of artificial chromosomes is the recent report of the generation of transgenic mice (Chapter 3) via germline transmission of a mammalian artificial chromosome using

nuclear microinjection (Chapter 2). Thus, artificial chromosomes could theoretically be used for gene therapy. To date, most studies have used plasmid DNA for gene transfer using nonviral vectors because they are easily amplified to a high copy number in bacteria, and their smaller size makes them easier to insert into cells.

Transfer of Nonviral Vectors into Cells

A major problem of nonviral vectors is the difficulty to efficiently transfer the highly charged DNA molecule into a cell. Transfer of nonviral vectors into cells can be performed ex vivo or in vivo. For ex vivo transfer, genes are usually transferred into the cell by using calcium phosphate co-precipitation, electroporation, cationic lipids, or liposomes. For most cell types, 5 to 10% of the cells can be modified, and transfected cells can often be selected by virtue of a selectable marker that is also present on the piece of DNA. Larger pieces of DNA are transferred less efficiently than smaller pieces of DNA.

Efficient in vivo transfer is somewhat more difficult to achieve than ex vivo gene transfer. Many investigators have utilized liposomes, cationic lipids, or anionic lipids that promote entry of the DNA into the cell. A variety of such molecules have been synthesized. Another effective method for promoting entry into the cell is to complex the DNA with an inactivated viral particle containing plasma membrane fusions proteins. For example, association of DNA with heat-inactivated Sendai virus [also known as the hemagglutinating virus of Japan (HVJ)] dramatically increases the expression of the DNA in vivo. Similarly, inactivated adenovirus greatly potentiates the entry of DNA into a cell. A third approach is to attach the DNA to a small particle delivered to the inside of the cell using a ballistic device referred to as a DNA gun (see Chapter 5).

Selective delivery (targeting) of a nonviral vector to a specific organ or cell type would be desirable for some applications. For example, DNA has been targeted to the asialoglycoprotein receptor of hepatocytes by complexing the DNA with polylysine-conjugated asialoglycoprotein or targeted for cells that express a transferrin or folate receptor (see Chapter 7).

Stabilization of Nonviral Vectors in Cells

A major problem with nonviral vectors is transient gene expression, since the genetic material transferred into the cell is unstable. Methods for stabilizing the DNA in the cell would prolong the clinical effect in vivo. Some investigators have placed origins of replication derived from viruses into nonviral vectors. Plasmids must be engineered to express any proteins necessary to activate the origin of replication. The human papilloma virus (HPV) E1 protein supports replication of the HPV origin of replication, while the Ebstein–Barr virus nuclear antigen 1 (EBNA-1) supports replication of an EBV origin of replication. Plasmids containing these replication origins and relevent appropriate proteins activating origins are maintained longer in cells in vitro and in vivo than plasmids that do not contain these sequences. Artificial chromosomes have elements that stabilize genetic material in a cell and should not have problems of instability. If difficulties in amplifying and transferring artificial chromosomes into cells can be overcome, such vectors should be maintained stably in a cell.

Use on Nonviral Vectors for Gene Therapy

Plasmid DNA has been delivered into muscle in vivo as naked DNA, into a variety of organs complexed with cationic lipids, with HJV liposomes, or by using a DNA gun. Expression has been detected in several organs, although it is usually both transient and at a relatively low level because the DNA is not stable in cells. There is little quantitative data regarding the efficacy of expression from different promoters in vivo. Gene therapy with plasmid vectors has been used to attempt to treat cystic fibrosis (see Chapter 3) and cancer in humans (see Chapter 10).

Risks of Nonviral Vectors for Gene Therapy

There are two major risks of using nonviral vectors for gene therapy: (1) insertional mutagenesis could activate oncogenes or inhibit tumor suppressor genes if the plasmid integrates and (2) the compounds that are used to facilitate the entry of DNA into a cell might have some toxicity. A major advantage of using nonviral vectors is the lack of risk of generating a wild-type virus via recombination. In addition, episomal plasmids do not pose the risk of insertional mutagenesis since they do not integrate into the chromosome. However, some plasmids can integrate into the genome particularly when a procedure is used to select clones exhibiting long-term expression. This is often done with ex vivo gene therapy procedures. Indeed, transplantation of myoblasts transfected with a plasmid DNA and selected in vitro has led to the development of tumors in the muscle. It, therefore, appears that selection of cells with an integrated plasmid vector poses some risks in animals, although maintenance of episomal DNA should be relatively safe. A second potential risk for nonviral vectors is that certain compounds can facilitate entry into a cell and exert a toxic effect in vivo. For example, many cationic lipids have considerable toxicity when administered at high doses to cells in vitro. These could be toxic at high doses in vivo as well. It will be necessary to assess the toxicity of such compounds carefully in vivo.

Summary: Nonviral Vectors

Nonviral vectors can be amplified to high copy numbers in bacterial cells as well as readily engineered to express a therapeutic gene from a mammalian promoter. These plasmids can be efficiently introduced into cells ex vivo and introduced somewhat less efficiently into cells in vivo. Their major advantages are the ease of production and that they cannot recombine to generate replication-competent virus. They can, however, integrate at a low frequency into the chromosome and, therefore, do pose some risk of insertional mutagenesis. A major disadvantage is the transient nature of gene expression that is observed.

OLIGONUCLEOTIDES

The second major class of nonviral vectors are oligodeoxynucleotides and related polymers of nucleotides that have different backbones. Oligodeoxynucleotides are 15 to 25 nt long pieces of DNA that can modulate gene expression in cells in a

variety of ways including: (1) formation of triplex DNA, (2) acting as an antisense molecule to block processing or expression of mRNA or to promote its degradation, and (3) forming a transcription factor binding site that serves as a decoy.

Triplex DNA is the colinear association of three deoxynucleotides strands and usually involves binding of an oligodeoxynucleotide in the major groove of a DNA double helix. This binding can block access of transcription factors, thus inhibiting transcription of a gene. The triplex-forming oligodeoxynucleotide binds to the purine-rich strand of the double helix via Hoogsteen hydrogen bonds. Potential target sites for triplex formation are limited to regions that contain homopurine on one strand. The relatively weak binding affinity and the instability of oligodeoxynucleotides in cells results in a transient effect.

A second mechanism by which oligodeoxynucleotides alter gene expression involves binding to an mRNA via standard Watson–Crick base pairing. This can block splicing by binding to a pre-mRNA splice signal or block translational initiation by binding to the 5′ Cap region or the translational initiation codon region. They can also result in degradation of the mRNA by RNase H, an enzyme that degrades the RNA portion of an RNA:DNA hybrid. A third mechanism by which oligodeoxynucleotides can alter gene expression is to bind transcription factors, which prevents them from associating with endogenous genes.

Natural antisense oligodeoxynucleotides consist of phosphodiester oligomers, are sensitive to nucleases, and have a half-life in serum of 15 to 60 min. Modifications to the backbone have increased the stability of oligonucleotides to allow a prolonged biological effect on targeted cells in vivo. Substitution of a nonbridge oxygen in the phosphodiester backbone with a sulfur molecule results in phosphorothioate nucleotides, which are resistant to nucleases. Substitution of a nonbridge oxygen with a methyl group results in methylphosphonate nucleotides. These are also resistant to nucleases, although they do not allow RNase H to act upon hybridized RNA. Peptide nucleic acids have an achiral amide-linked backbone homologous to the phosphodiester backbone that can form standard Watson–Crick base pairs with RNA. Modified oligonucleotides are stable in culture and serum and have resulted in prolonged biological effects.

For oligonucleotides to exert a biological effect, they must enter the cell. Oligonucleotides appear to enter the cell via receptor-mediated endocytosis. Permeabilization of the cell membrane can potentiate entry. In vivo delivery of oligonucleotides can be increased by HVJ liposome complexes. Improved delivery to cells should result in a biological effect at lower doses.

Use and Safety of Oligonucleotides for Gene Therapy

Oligonucleotides have been administered in vivo for gene therapy. They have successfully inhibited intimal hyperplasia of arteries. Oligonucleotides that served as a decoy for a transcription factor have been used to inhibit proliferation of smooth muscle cells in blood vessels in vivo. Antisense oligonucleotides have blocked expression of oncogenes, slowed replication in cells in vitro, and had a modest but transient effect upon growth of tumor cells in vivo.

The major toxicity of oligonucleotides relates to the administration of large doses to achieve a clinical effect. Administration of high doses of phosphorothioate oligonucleotides resulted in cardiovascular toxicity and death in some primates.

Mechanisms to promote the entry of oligonucleotides into cells should decrease their toxicity. Oligonucleotides are unlikely to have any long-term adverse effects since they do not integrate into the chromosome.

Summary: Oligonucleotides

In summary, oligodeoxynucleotides can be used to alter expression of an endogenous gene by blocking transcription, blocking mRNA processing or translation, potentiating mRNA degradation, or through serving as a decoy for a transcription factor. Modified oligonucleotides can function in a similar fashion and are more stable. Oligonucleotides can alter gene expression in vitro and to a lesser extent in vivo. Their effects are short-lived due to their instability in cells and in blood. Their use for gene therapy will probably be limited to diseases where transient expression is sufficient.

KEY CONCEPTS

- Viral vectors can be produced by removing some or all of the genes that encode viral proteins, and replacing them with a therapeutic gene. These vectors are produced by cells that also express any proteins that are necessary for producing a viral particle. A risk of all viral vectors is that they might recombine to generate replication-competent virus that could cause disease in humans.
- Nonviral vectors are plasmids that can be propagated in bacteria or oligonucleotides that can be synthesized chemically. Plasmids can transfer a therapeutic gene into a cell, while oligonucleotides inhibit the expression of endogenous genes. Transfer of nonviral vectors into cells is inefficient and the effect is generally transient. These vectors do not carry the risk of recombining to generate wild-type virus.
- Retroviral vectors are devoid of any retroviral genes and result in long-term expression due to their ability to integrate into the chromosome. Their major disadvantage is the fact that they only transduce dividing cells. Recently developed lentiviral vectors do transduce nondividing cells, but there are concerns regarding the safety of these vectors.
- Adenoviral vectors generally contain many adenoviral genes, although "gutless" vectors in which all coding sequences have been deleted have been developed recently. Adenoviral vectors transduce nonreplicating cells very efficiently, although expression is short-lived. This transient expression is primarily due to the immune response to residual adenoviral genes or the transgene in early generation vectors and may be due to the deletion of sequences that stabilize the DNA in cells for the gutless vectors.
- AAV vectors are devoid of any AAV genes and can transduce nondividing cells. They have resulted in long-term expression, although it is unclear if they remain episomal or integrate into the chromosome in nondividing cells. Production of large amounts of AAV vector is problematic.