Design of Retroviral Vectors and Helper Cells for Gene Therapy

WEI-SHAU HU AND VINAY K. PATHAK

Department of Microbiology and Immunology (W.-S.H.), Department of Biochemistry (V.K.P.), Mary Babb Randolph Cancer Center (W.-S.H., V.K.P.), School of Medicine, West Virginia University, Morgantown, West Virginia; and HIV Drug Resistance Program, National Cancer Institute, FCRDC, Frederick, Maryland (W.-S.H., V.K.P.)

This paper is available online at http://www.pharmrev.org

Abstract

I. General introduction and scope

II. Background

A. Replication cycle of retroviruses

B. Genome structure of retroviruses

III. Basic concepts in retrovirus vectors and helper cells

IV. Helper cells and packaging systems

A. Helper cell lines

B. One-genome helper constructs

C. Split-genome helper constructs

D. Inducible helper constructs

E. Transient transfection systems

F. Systems that generate pseudotyped viruses

G. Systems containing genetically modified env for cell or tissue targeting

H. Hybrid systems

V. Vectors based on different retroviruses

A. Vectors derived from oncoviruses

1. Murine leukemia virus-based retroviral vectors

2. Spleen necrosis virus-based vectors

3. Rous sarcoma virus- and Avian leukemia virus-based vectors

B. Vectors derived from lentiviruses

C. Vectors derived from spumaviruses

VI. Design of retroviral vectors

A. Standard vectors

1. U3 promoter-driven expression

2. Vectors that use an internal promoter to express additional genes

3. Vectors that use splicing to express additional genes

4. Vectors that use translational control signals to express additional genes

B. Double-copy vectors

C. Self-inactivating vectors

1. U3 minus vectors

2. Cre/loxP vectors

D. Self-inactivating and self-activating vectors

E. Vectors targeted to specific cells

F. Vectors that utilize cell-type-specific promoters

VII. General considerations for using retrovirus vectors and helper cells

A. Gene expression

B. Virus host range and titers

1. Considerations for envelope selection and virus host range

2. Ping-pong amplification

3. Concentration of viruses

1 Address for correspondence: Wei-Shau Hu, Bldg. 535, Rm. 336, HIV Drug Resistance Program, National Cancer Institute, FCRDC, Frederick, MD 21702. E-mail: whu@mail.ncifcrf.gov
Abstract—During the past decade, gene therapy has been applied to the treatment of disease in hundreds of clinical trials. Various tools have been developed to deliver genes into human cells; among them, genetically engineered retroviruses are currently the most popular tool for gene delivery. Most of the systems contain vectors that are capable of accommodating genes of interest and helper cells that can provide the viral structural proteins and enzymes to allow for the generation of vector-containing infectious viral particles. Retroviridae is a family of retroviruses that differs in nucleotide and amino acid sequence, genome structure, pathogenicity, and host range. This diversity provides opportunities to use viruses with different biological characteristics to develop different therapeutic applications. Currently, a variety of retroviruses that provide distinct advantages for gene delivery have been modified and used in clinical trials. In this review, the genome structures of oncoviruses, lentiviruses, and spumaviruses are reviewed and examples of vectors derived from these viruses are described. As with any delivery tool, the efficiency, the ability to target certain tissue or cell type, the expression of the gene of interest, and the safety of retroviral-based systems are important for successful application of gene therapy. Significant efforts have been dedicated to these areas of research in recent years. Various modifications have been made to retroviral-based vectors and helper cells to alter gene expression, target delivery, improve viral titers, and increase safety. The principles and design of these modifications are discussed in this review.

I. General Introduction and Scope

The DNA sequence of the entire human genome has just been determined (Marshall, 2000; Pennisi, 2000). Identification and characterization of all 50,000 to 100,000 human genes will lead to a greater understanding of normal and aberrant genes that play important roles in causing human disease. These developments are likely to provide vastly expanded opportunities to intervene in disease processes by delivering genetic material to the target cells affected by the disease. This new approach to treating disease, namely gene therapy, is likely to play an increasingly important role in medicine of the 21st century.

With the advancement of biotechnology and molecular biology, hundreds of human genes have already been isolated and characterized. Furthermore, mutations in many of these genes that are responsible for various diseases have been identified. These developments have made it possible to treat genetically inherited diseases using nucleic acids rather than proteins or pharmacological agents. In addition, the molecular mechanisms of action have been dissected for various genes in humans as well as other organisms. This understanding provides the knowledge base that can be used to enhance or interfere with these mechanisms of action with gene therapy.

Molecular biologists and gene therapists are well aware that the knowledge of molecular mechanisms of genetic disease does not imply that successful gene therapy approaches will be developed. After all, the molecular mechanism of sickle cell anemia has been understood for several decades but a promising gene therapy approach is not yet in sight. Ideally, gene therapy should be efficient, cell-specific, and safe. One of the challenges of gene therapy is the efficient delivery of genes to target cells. Although the nucleic acids containing the genes can be generated in the laboratory with relative ease, the delivery of these materials into a specific set of cells in the body is far from simple. Various methods have been developed to deliver the genes into cells. Some of these methods involve using physical approaches to delivery such as direct DNA injection, encapsulation of DNA into liposomes, and gene gun technology (Cooper, 1996; Yang et al., 1996; Prince, 1998). Other methods take advantage of viruses, a class of intracellular parasites, to deliver the genes into the target cells. In general, viruses are more efficient at delivering genes to target cells than physical methods.

Viruses are logical tools for gene delivery. They replicate inside cells and therefore have evolved mechanisms to enter the cells and use the cellular machinery to express their genes. The concept of virus-based gene delivery is to engineer the virus so that it can express the gene of interest and yet retain its ability to deliver the gene to target cells. These engineered viruses are often referred to as viral vectors. Depending on the specific application and the type of virus, most viral vectors contain mutations that hamper their ability to replicate freely as wild-type viruses in the host.

Viruses from several different families have been modified to generate viral vectors for gene delivery. These viruses include retroviruses, adenoviruses, adeno-associated viruses, herpes simplex viruses, picornaviruses, and alphaviruses (Flotte and Carter, 1995; Glo-
Retroviruses were the first viruses to be modified for gene delivery, and retroviral vectors are used in the majority of all gene therapy clinical trials (Anderson, 1996). Several reviews have summarized various aspects of retroviral vector-based gene therapy (Dornburg, 1995; Gunzburg et al., 1996; Vile et al., 1996; Miller, 1997; Karavanas et al., 1998). This review will focus on the design of retroviral vectors and helper constructs.

An ideal retroviral vector for gene delivery must be efficient, cell-specific, regulated, and safe. The efficiency of delivery is important because it can determine the efficacy of the therapy. Current efforts are aimed at achieving cell-type-specific infection and gene expression with retroviral vectors. In addition, retroviral vectors are being developed to regulate the expression of the gene of interest, since the therapy may require long-lasting or regulated expression. Safety is a major issue for viral gene delivery because most viruses are either pathogens or have a pathogenic potential. It is important that during gene delivery, the patient does not also inadvertently receive a pathogenic virus that has full replication potential. The design of retroviral vectors and strategies to address these concerns will be covered in this review.

II. Background

A. Replication Cycle of Retroviruses

A general outline of the retroviral replication cycle is illustrated in Fig. 1. Retroviruses are RNA viruses that replicate through an integrated DNA intermediate (Coffin, 1996). Retroviral particles encapsidate two copies of the full-length viral RNA, each copy containing the complete genetic information needed for virus replication. Retroviruses possess a lipid envelope and use interactions between the virally encoded envelope protein that is embedded in the membrane and a cellular receptor to enter the host cells. Using the virally encoded enzyme reverse transcriptase, which is present in the virion, viral RNA is reverse transcribed into a DNA copy. This DNA copy is integrated into the host genome by integrase, another virally encoded enzyme. The integrated viral DNA is referred to as a provirus and becomes a permanent part of the host genome. The cellular transcriptional and translational machinery carries out expression of the viral genes. The host RNA polymerase II transcribes the provirus to generate RNA, and other cellular processes modify and transport the RNA out of the nucleus. A fraction of viral RNAs are spliced to allow expression of some genes whereas other viral RNAs remain full-length. The host translational machinery synthesizes and modifies the viral proteins. The newly synthesized viral proteins and the newly synthesized full-length viral RNAs are assembled together to form new viruses that bud out of the host cells (Coffin, 1996).

The replication cycle described above applies to all retroviruses with the exception of spumaviruses. It has recently been shown that reverse transcription of spumaviruses occurs in the virus-producing cells rather than the infected target cell, and the infectious virus contains a DNA genome (Yu et al., 1996a, 1999; Moebes et al., 1997).

B. Genome Structure of Retroviruses

Based on their genome structures, retroviruses can be classified into simple and complex retroviruses (Coffin, 1996; Vogt, 1997). Simple and complex retroviruses encode gag (group-specific antigen), pro (protease), pol (polymerase), and env (envelope) genes. In addition to these genes, complex retroviruses also encode several accessory genes. [In this review, the gene is referred to in italics (for example, gag), whereas the polyprotein is referred to in regular font with the first letter capitalized (for example, Gag)].

Retroviruses can also be classified into oncoviroses, lentiviruses, and spumaviruses (Coffin, 1996). Most oncoviruses are simple retroviruses. Lentiviruses, spumaviruses, and some oncoviruses are complex retroviruses. Currently, all three types of viruses are being exploited as gene therapy tools. Examples of each type will be discussed later in this review. Structures of the three types of retroviruses are shown in Fig. 2. Murine leukemia virus (MLV) is illustrated as an example of an oncovirosus (Fig. 2A), human immunodeficiency virus 1

---

**FIG. 1.** Replication cycle of retroviruses. A retrovirus binds to a receptor on the cell surface (shown as a crescent), enters the cell, and reverse transcribes the RNA into double-stranded DNA (shown as a line flanked by black boxes), viral DNA integrates into the cell chromosome (shown as zigzag lines) to form a provirus. Cellular machinery transcribes and processes the RNA (shown as thin lines), and translates the viral proteins (shown as black ellipses and white circles). Viral RNA and proteins assemble to form new viruses, which are released from the cell by budding.
(HIV-1) is illustrated as an example of a lentivirus (Fig. 2B), and human foamy virus is illustrated as an example of a spumavirus (Fig. 2C).

The viral DNA contains large redundant sequences at the two ends of the genome designated long terminal repeats (LTRs) (Coffin, 1996; Vogt, 1997). LTRs can be further divided into U3 (unique 3'), R (repeat), and U5 (unique 5') regions. The viral promoters and transcriptional enhancers are located in the U3 region. The R region is essential for reverse transcription and replication of all retroviruses. In addition, the R regions of some viruses also contain elements important for gene expression. For example, the R region of MLV also contains a polyadenylation signal, whereas the HIV-1 R region contains a trans-activation response region (TAR) that is important for activation of HIV-1 gene expression (Luciw, 1996). The U5 region contains sequences that facilitate the initiation of reverse transcription. Immediately downstream of the 5' LTR is a primer binding site (PBS) that has sequence complementarity to a portion of a cellular tRNA. Different tRNAs are used by different viruses as primers for the initiation of reverse transcription (Vogt, 1997). The packaging signal (Ψ) or encapsidation signal (E) are sequences that interact with the viral proteins to accomplish specific packaging of the viral RNA (Linial and Miller, 1990; Rein, 1994; Berkowitz et al., 1996). The MLV Ψ is located in the 5' untranslated region between the PBS and the gag open reading frame, and it is sufficient for incorporation of foreign RNA into viral particles (Mann et al., 1983; Adam and Miller, 1988). However, sequences extending into the gag open reading frame can enhance the efficiency of packaging; the sequences containing the MLV packaging signal together with the portion of the gag open reading frame that facilitates packaging is termed Ψ+ (Bender et al., 1987). The HIV-1 packaging signal is less well defined, but includes the 5' untranslated region as well as the 5' portions of the gag open reading frame (Lever et al., 1989; Hayashi et al., 1992; Richardson et al., 1993; Kaye et al., 1995; McBride and Panganiban, 1996, 1997; McBride et al., 1997). However, these sequences have not been shown to be sufficient for packaging of heterologous RNA. In human and simian foamy viruses, both the 5' untranslated region and sequences located in the 3' portion of pol are important for RNA packaging (Erlwein et al., 1998; Heinkelein et al., 1998; Wu et al., 1998).

The coding regions of all retroviruses contain at least three genes. The gag gene near the 5' end of the viral genome codes for Gag polyproteins that make up the viral capsid. After assembly of the virus particle, the Gag polyprotein is proteolytically cleaved into several proteins including matrix, capsid, and nucleocapsid. The pol gene encodes reverse transcriptase and integrase. Reverse transcriptase copies the viral RNA to generate the viral DNA, whereas integrase integrates the viral DNA into the host chromosome to form a provirus. In all retroviruses, a spliced mRNA is used to express the env gene. The env gene codes for the envelope polyprotein, which is cleaved into the transmembrane domain and the surface domain (SU). The sequences that encode the viral protease (Pro) are always located between gag and pol and are most often expressed as either a part of the Gag polyprotein or as a part of the Gag-Pol polyprotein.

With the exception of spumaviruses, pol is expressed as a Gag-Pol fusion protein by controlled ribosomal frameshifting or translational suppression of a stop codon between the gag and pol open reading frames (Coffin, 1996; Swanstrom and Wills, 1997). The pol gene in spumaviruses is expressed by a spliced mRNA (Yu et al., 1996b). Because the expression of gag and pol are closely related in most viruses, these genes are often referred to as gag/gag-pol in helper constructs.

The region between env and the 3' LTR contains a purine-rich region known as polypurine tract (PPT) that is important for reverse transcription (Coffin, 1996). In some viruses, this region also contains a constitutive transport element that allows the transport of the full-length, unspliced RNA from the nucleus to the cytoplasm (Bray et al., 1994; Ogert et al., 1996). The 3' LTR has the same sequence as the 5' LTR. Short sequences at the two ends of the LTR are important for integration and are referred to as attachment sites (att). The att interact with integrase and are necessary for efficient integration of the viral DNA (Coffin, 1996; Brown, 1997).

The HIV-1 genome encodes other accessory proteins in addition to Gag, Pro, Pol, and Env (Luciw, 1996; Vogt, 1997; Frankel and Young, 1998). These accessory proteins are Tat, Rev, Nef, Vif, Vpr, and Vpu (shown in Fig. 2B); spliced mRNAs are used to express all of the acces-
sory proteins. Tat is a transcriptional activator that binds to the TAR region of the viral RNA and some host proteins. These interactions increase both initiation and elongation of the viral RNA transcription. Rev regulates transport of unspliced (full-length) or single-spliced viral RNA from the nucleus to the cytoplasm. Rev binds to a region of the viral RNA known as the Rev responsive element (RRE) located within Env to allow transport of viral RNAs containing the RRE across the nuclear membrane (Luciw, 1996). Multiply spliced HIV-1 RNAs that express Rev, Tat, and other accessory proteins do not contain RRE and can be transported out of the nucleus without the presence of Rev (Luciw, 1996).

Nef down-regulates the CD4 receptor and plays an important role in viral pathogenesis; Nef is also thought to enhance viral reverse transcription, although the exact mechanism of this enhancement is unclear. Vif plays an important role in the infectivity of viruses. However, propagating the virus in certain cell types can sometimes circumvent this requirement. It is thought that Vif may counteract the effect of a negative host factor. Certain cell types that are devoid of the putative negative factor can efficiently produce infectious Vif-deficient HIV-1 viruses. Vpu can trigger CD4 degradation and enhance virion release. Vpr is present in large quantities in the virion and is thought to be important for cell cycle arrest, fidelity of viral DNA synthesis, and efficiency of viral replication in some cells (Luciw, 1996; Mansky, 1996; Vogt, 1997; Frankel and Young, 1998; Mansky et al., 2000).

Spumaviruses (foamy viruses) are a distinct genus of the retrovirus family. Recent studies revealed that spumaviruses are different from oncoviruses and lentiviruses in many aspects, and share some similarities with hepatitis B virus. Therefore, it has been suggested that spumaviruses should be classified as a subfamily of the Retroviridae family (Linial, 1999). The genome structure of a human spumavirus is shown in Fig. 2C. In addition to gag, pol, and env, human spumaviruses also contain three accessory genes designated bel1, bel2, and bel3 (Vogt, 1997). Bel1, also known as Tas, is a transcriptional transactivator that is required for efficient expression of the viral U3 promoter (Linial, 1999). Tas is expressed from an internal promoter (Linial, 1999).

III. Basic Concepts in Retrovirus Vectors and Helper Cells

When a replication-competent retrovirus infects a natural host cell, it can form a provirus in the host genome, express viral genes, and release new infectious particles to infect other hosts. In most gene therapy applications, it is not desirable to deliver a replication-competent virus into a patient because the virus may spread beyond the targeted tissue and cause adverse pathogenic effects. Therefore, in most retroviral systems designed for gene delivery, the viral components are divided into a vector and a helper construct to limit the ability of the virus to replicate freely (Miller, 1997). The term vector generally refers to a modified virus that contains the gene(s) of interest and cis-acting elements needed for gene expression and replication. Most vectors contain a deletion(s) of some or all of the viral protein coding sequences so that they are not replication-competent. Helper constructs are designed to express viral genes lacking in the vectors and to support replication of the vectors. The helper function is most often provided in a helper cell format although it can also be provided as a helper virus or as cotransfected plasmids. Helper cells are engineered culture cells expressing viral proteins needed to propagate retroviral vectors; this is generally achieved by transfecting plasmids expressing viral proteins into culture cells. Most helper cell lines are derived from cell clones to ensure uniformity in supporting retroviral vector replication. Helper viruses are not used often because of the likelihood that a replication-competent virus could be generated through high frequency recombination. Helper functions can also be provided by transient transfection of helper constructs to achieve rapid propagation of the retroviral vectors.

An example of a retroviral vector in a helper cell and a target cell is shown in Fig. 3A. In this example, the retroviral vector contains all of the cis-acting elements needed for viral replication and gene expression. Most retroviral vectors are maintained as bacterial plasmids to facilitate the manipulation and propagation of the vector DNA (Fig. 3B). These double-stranded DNA vectors can be introduced into helper cells by conventional methods such as DNA transfection, lipofection, or elec-

**Fig. 3.** Retroviral vector propagation in helper cells and structure of a prototypical vector. A, propagation of retroviral vectors in helper cells. Helper cells produce the viral proteins (shown as black ellipses and white circles) that are used to assemble viral particles containing RNA transcribed from the viral vector. Target cells do not express viral proteins and cannot generate viral particles containing the vector RNA. B, cis-acting elements needed in a prototypical retroviral vector. The plasmid backbone contains a drug resistance gene and a bacterial origin of replication (ori).
troporation. The helper cell shown expresses all of the viral proteins (Gag, Gag-Pol, and Env) but lacks RNA containing the packaging signal. Viral RNA is necessary for the formation and release of infectious viral particles, but it is not necessary for the formation of "empty" noninfectious viral particles. When the vector DNA is introduced into the helper cells, vector RNA containing a packaging signal is transcribed and efficiently packaged into viral particles. The viral particles contain viral proteins expressed from helper constructs and RNA transcribed from the vector. These viral particles can infect target cells, reverse transcribe the vector RNA to form a double-stranded DNA copy, and integrate the DNA copy into the host genome to form a provirus. This provirus encodes the gene(s) of interest and is expressed by the host cell machinery. However, because the vector does not express any viral proteins, it cannot generate infectious viral particles that can spread to other target cells.

IV. Helper Cells and Packaging Systems

A. Helper Cell Lines

Helper cells are designed to support the propagation of retroviral vectors. The viral proteins in the helper cells are expressed from helper constructs that are transfected into mammalian cells. Helper constructs vary in their mode of expression and in the genes they encode. Most of the currently available helper cell lines are listed in a table in a recent review (Miller, 1997).

B. One-Genome Helper Constructs

In helper cell lines that were initially developed, all of the viral genes were expressed from one helper construct. Examples of these helper cells are C3A2 and VS-2 (Mann et al., 1983; Watanabe and Temin, 1983). The helper constructs for these cell lines were cloned proviral DNAs that lacked the packaging signals. These helper cells can support efficient propagation of retroviral vectors. However, a major problem with these helper cells is that replication-competent viruses can be frequently generated during the propagation of the viral vector. The helper construct contains most of the viral genome and thus shares significant sequence homology with the retroviral vector. The sequence homology can facilitate recombination between the helper construct and the retroviral vector to generate replication-competent viruses. Although the helper RNA lacks the packaging signal, it can still be packaged into a virion with a low efficiency (approximately 100- to 1,000-fold less than RNAs containing \( \Psi \)) (Embreton and Temin, 1987a). Retroviral recombination occurs frequently between the two co-packaged viral RNAs to generate a DNA copy that contains genetic information from both parents. If the helper RNA and the vector RNA are packaged into the same virion, the large regions of sequence homology between the two RNAs can facilitate homologous recombination during reverse transcription to generate a replication-competent virus. A similar recombination event can also occur between the helper RNA and RNA derived from an endogenous virus at a lower efficiency to generate replication-competent viruses.

C. Split-Genome Helper Constructs

The safety concern associated with the generation of replication-competent viruses has provoked the design of many helper cell lines using "split genomes", including \( \Psi \)CRIP, GP+envAm12, and DSN (Danos and Mulligan, 1988; Markowitz et al., 1988; Dougherty et al., 1989). In these helper cells, the viral Gag/Gag-Pol polyproteins are expressed from one plasmid and the Env proteins are expressed from another plasmid. Furthermore, the two helper constructs also contain deletions of viral cis-acting elements to reduce or eliminate sequence homology with the retroviral vector. In these helper cells, genes encoding viral proteins are separated into two different constructs and the viral cis-acting elements are located in the vector. Therefore, several recombination events have to occur to reconstitute the viral genome. In addition, reducing the regions of homology decreases the probability that these recombination events will occur. Therefore, helper cells containing split-genome helper constructs are considered safer than helper cells containing one-genome helper constructs.

D. Inducible Helper Constructs

In contrast to the helper cell lines described above that express viral proteins constitutively, some helper cell lines have been designed to express the viral proteins in an inducible manner. One rationale for the generation of an inducible helper cell line is that some viral proteins are cytotoxic and cannot be easily expressed at high levels (Yu et al., 1996a). By using an inducible system, expression of the cytotoxic proteins can be limited to the stage in which virus is propagated. By controlling the expression of the cytotoxic proteins, high viral titers can be achieved. Examples of the inducible helper cells include the 293PGP cells (Ory et al., 1996) and HIV-1 helper cell lines (Yu et al., 1996a). Recent advances in inducible gene expression systems are described in some recent reviews (Thummel, 1990; Walther and Stein, 1996; Clackson, 1997; Saez et al., 1997; Harvey and Caskey, 1998; Herdegen and Leah, 1998; Rossi and Blau, 1998; Gao et al., 1999; Morgan et al., 1999).

E. Transient Transfection Systems

With the development of efficient transfection methods, transient transfection systems have also been developed for propagation of retroviral vectors. In these systems, helper functions are generally expressed from two different constructs, one expressing \( gag-pol \) and another expressing \( env \). These two constructs generally share little sequence homology. The retroviral vector and the helper constructs are transfected into cells, and
viruses are harvested a few days after transfection. This system has been employed by several different groups to successfully propagate MLV-based vectors and foamy virus vectors to achieve high viral titers (Finer et al., 1994; Soneoka et al., 1995; Naviaux et al., 1996; Tro-bridge and Russell, 1998; Yang et al., 1999).

F. Systems That Generate Pseudotyped Viruses

Pseudotyping refers to viral particles containing a viral genome from one virus and part (or all) of the viral proteins from a different virus (Zavada, 1976; Boettiger, 1979). The most common form of pseudotyping involves one virus using the envelope protein of another virus (Zavada, 1976; Boettiger, 1979). Some of the helper cell lines contain helper constructs that express gag-pol from one virus and env from another virus. Since the Gag polyproteins select the viral RNA, the viral vector to be propagated contains an RNA that is recognized by the Gag polyprotein expressed in these cells. However, the viral particles produced contain the Env protein derived from another virus. Therefore, these viral particles can only infect cells that express a receptor that can interact with the heterologous envelope protein. For example, the helper cell line PG13 expresses gag-pol from MLV and env from gibbon ape leukemia virus (GaLV) (Miller et al., 1991). Because the PG13 cell line expresses MLV Gag polyprotein, it can efficiently package MLV-based retroviral vectors. However, these viruses can only infect cells that express the receptor for the GaLV envelope. MLV and GalV are distantly related retroviruses, which makes it plausible that an infectious virus can be generated from pseudotyping. Surprisingly, however, it has also been shown that some envelopes derived from viruses of a different family can also pseudotype retroviruses and generate infectious viral particles (Zavada, 1976). For example, the G protein of vesicular stomatitis virus (VSV), a rhabdovirus, can be used to generate pseudotyped retroviral vectors (Burns et al., 1993; Yee et al., 1994a,b). These VSV G pseudotyped viruses exhibit a very broad host range and can infect a variety of cells that cannot normally be infected with retroviruses.

G. Systems Containing Genetically Modified env for Cell or Tissue Targeting

Interactions between the viral envelope proteins and the cellular receptors determine the host range of the virus. Strategies have been developed to target virus delivery into certain cell types by modifying the viral Env. After translation and modification, the SU portion of Env interacts with a cellular receptor. Efforts are currently being made to modify the SU portion of the Env to interact with other cell surface molecules. The modification of the SU portion of Env is often achieved by deletion of a part of the coding region for SU and replacing it with regions of other proteins. Proteins that have been used to modify the SU portion of Env include erythropoietin, heregulin, insulin-like growth factor I, and single-chain variable fragment antibodies against various proteins (Kasahara et al., 1994; Chu and Dornburg, 1995; Han et al., 1995; Sonia et al., 1995; Jiang et al., 1998; Konishi et al., 1998; Chadwick et al., 1999). Some of these strategies have been shown to work in both cell culture and in animals; for example, viral particles with hybrid Env containing single-chain antibodies were shown to be able to infect target cells in cultured cell lines and in a SCID-hu mouse model (Jiang et al., 1998; Jiang and Dornburg, 1999).

H. Hybrid Systems

Some recently developed systems use a hybrid approach for propagation of retroviral vectors. A helper cell line is used to constitutively express some of the viral proteins, whereas other viral proteins are introduced into the helper cell line by transient transfection. For example, a retroviral vector can be introduced into a helper cell line that constitutively expresses the MLV gag-pol. To propagate the retroviral vector, a plasmid designed to express the VSV G can be introduced into the system by transient transfection (Yee et al., 1994b). As another variation on this theme, the retroviral vector itself may encode some of the viral proteins (for example, Gag/Gag-Pol), and a helper cell line may provide other viral proteins (Env) (Boerkoel et al., 1993). Approaches that use other viruses to deliver the retroviral helper constructs are also being explored. For example, a modified herpes simplex virus was generated to contain the retroviral gag, pol, and env to serve the helper function (Savard et al., 1997). Similarly, adenovirus vectors and Semliki Forest virus-derived expression vectors have also been used to deliver genes encoding MLV viral proteins to helper cells (Li and Garoff, 1996; Duisit et al., 1999).

V. Vectors Based on Different Retroviruses

Many retroviruses have been modified to generate vectors that can carry gene(s) of interest. Viral vectors generally contain all of the cis-acting elements needed for viral replication and gene expression; these elements are shown in Fig. 3B. Additional elements may also be needed in vectors derived from some viruses to ensure successful gene delivery. The requirement for these cis-acting elements has often become apparent from greater understanding of the biology of these viruses. In addition, to allow easy manipulation in bacterial cells, most retroviral vectors are in plasmid form and have a backbone containing the bacterial origin of replication and an antibiotic resistance gene (Fig. 3B).

The following steps are typically carried out to produce viral particles from retroviral vectors. Vector DNA is first introduced into the helper cells by transfection, electroporation, or lipofection. After introduction of the DNA into the helper cells, the vector DNA integrates into the helper cell and is expressed. The viral RNA is
expressed from the 5’ LTR and consists of all the sequences between the two R regions. This viral RNA contains the packaging signal and is packaged into the viral particles efficiently. During retroviral replication, the plasmid backbone sequences outside the two LTRs are not transferred to the target cells. The basic structures of some retroviral vectors derived from different retroviruses are described below.

A. Vectors Derived from Oncoviruses

Vectors derived from three different oncoviruses will be described here to represent some of the most widely used retroviral vectors. Oncoviruses can only infect dividing cells; therefore, vectors that are derived from oncoviruses can only be used to efficiently deliver genes into dividing cells. The requirement for cell proliferation can sometimes be used as an advantage to selectively target rapidly dividing cells (for example, cancer cells).

1. Murine Leukemia Virus-Based Vectors. Currently, MLV-based retroviral vectors and helper cells are the most frequently used system for gene delivery. The development and availability of engineered vectors and helper cell lines has promoted the popularity of MLV-based vectors. The structure of a prototypical MLV-based retroviral vector is illustrated in Fig. 4A. The vectors contain cis-acting viral sequences that are needed for gene expression and viral replication such as LTRs, PBS, PPT, and att. The packaging signal can be the minimum signal (Ψ) or the longer signal that extends into the gag open reading frame (Ψ+) (Mann et al., 1983; Bender et al., 1987; Adam and Miller, 1988). When the Ψ+ is present in the vector, it is necessary to mutate the translational initiation codon of gag to prevent expression of the truncated Gag protein. Several vectors have been designed to contain multiple restriction enzyme sites between the packaging signal and the 3’ untranslated region (marked as cloning sites in Fig. 4A). The presence of these cloning sites facilitates the construction of vectors that can express the gene of interest.

MLV-based vectors can be propagated in all of the MLV helper cell lines efficiently. There are several MLV envelope proteins that dictate the host range of MLV vectors. Viruses that use the ecotropic envelope can infect mouse cells but not cells derived from other species. Viruses that use the amphotropic envelope can infect both mouse cells and cells derived from other species, including human cells. Viruses that use the xenotropic envelope cannot infect mouse cells but can infect cells derived from other species (Miller, 1997). In addition, MLV vectors can also be propagated in spleen necrosis virus (SNV)-based helper cell lines (Embretson and Temin, 1987a; Certo et al., 1998; Certo et al., 1999). SNV is an avian virus that is distantly related to MLV. Surprisingly, SNV proteins retain the ability to interact with MLV cis-acting sequences and package MLV RNA, reverse transcribe the MLV genome, and integrate the MLV RNA into the host (Embretson and Temin, 1987a; Certo et al., 1998, 1999).

2. Spleen Necrosis Virus-Based Vectors. Structure of a typical SNV-based vector is shown in Fig. 4B. The required viral sequences in these vectors are very similar to those of the MLV vectors. The packaging signal of SNV, denoted E, does not extend into the gag open reading frame (Watanabe and Temin, 1982; Embretson and Temin, 1987a); therefore, most SNV-based vectors do not contain the gag coding regions. Similar to MLV vectors, the genes of interest are inserted into a linker region containing multiple restriction sites between the packaging signal and the 3’ untranslated region.

SNV-based vectors can be propagated in SNV-based helper cell lines such as C3A2, DSDH, DSH134G, and DSN (Watanabe and Temin, 1983; Dougherty et al., 1989; Hu and Temin, 1990; Martinez and Dornburg, 1995). However, SNV vectors cannot be propagated in MLV-based helper cells because MLV proteins cannot efficiently package SNV RNA (Certo et al., 1998). In addition, the SNV promoter is not functional in mouse cells (Embretson and Temin, 1987b), which are frequently used to engineer MLV packaging cell lines. It has been shown that a post-transcriptional block may prevent the replication of SNV in a human cell line (HeLa) (Koo et al., 1991). If SNV cannot replicate in human cells then the SNV-based systems should be safer than MLV-based systems for human gene therapy, since the generation of a replication-competent virus during vector production should not pose a threat to the patient. However, the SNV-based systems are not extensively used in gene therapy because they are not as well developed or as widely available as MLV-based systems.

3. Rous Sarcoma Virus- and Avian Leukosis Virus-Based Vectors. RSV is the only known acute oncogenic
A retrovirus that is replication-competent (Coffin, 1996; Vogt, 1997). In addition to \textit{gag-pol} and \textit{env}, RSV also encodes the oncogene \textit{v-src} between \textit{env} and the 3' LTR. A splice acceptor site upstream of the \textit{v-src} allows the gene to be expressed as a spliced mRNA (Vogt, 1997). The ability of RSV to code for an additional gene has inspired the generation of RSV-based replication-competent retroviral vectors. Various modifications have been made to generate a replication-competent viral vector (Hughes et al., 1987; Greenhouse et al., 1988; Petropoulos and Hughes, 1991), an example of which is shown in Fig. 4C. In this construct, \textit{v-src} was replaced by a splice acceptor site and several restriction enzyme sites. DNA fragments can be inserted in the restriction sites to generate a replication-competent vector that expresses the gene of interest.

ALV has also been modified to generate vectors that require helper cells for their propagation (Stoker and Bissell, 1988; Cosset et al., 1992; Thacker et al., 1995). An example of an ALV vector is shown in Fig. 4D. Similar to the MLV and SNV vectors described above, the basic structure of an ALV vector also contains the 5' and 3' LTRs, \textit{att}, PBS, PPT, and a packaging signal. The packaging signal of ALV extends into the \textit{gag} open reading frame, and the relevant portions of \textit{gag} are included in ALV-based vectors to achieve efficient packaging.

\section*{B. Vectors Derived from Lentiviruses}

In contrast to the oncoviruses, some lentiviruses have been shown to infect nondividing, quiescent cells (Lewis et al., 1992; Bukrinsky et al., 1993). Lentivirus-based vectors were developed in part because they could be used to deliver genes into nondividing cells (Naldini et al., 1996; Blomer et al., 1997). Lentiviruses are complex retroviruses that need to express accessory proteins for regulation of their replication cycle. Some of these accessory proteins bind to regions of the viral genome to regulate gene expression. Therefore, lentivirus-based vectors need to incorporate additional \textit{cis}-acting elements so that efficient viral replication and gene expression can occur. As examples of lentivirus-based vectors, HIV-1- and HIV-2-based vectors are described below.

An HIV-1-based vector is shown in Fig. 5A. This vector contains \textit{cis}-acting elements that are also found in simple retroviruses. It has been shown that sequences that extend into the \textit{gag} open reading frame are important for packaging of HIV-1. Therefore, HIV-1 vectors often contain the relevant portion of \textit{gag} in which the translational initiation codon has been mutated. In addition, most HIV-1 vectors also contain a portion of the \textit{env} gene that includes the RRE (Luciw, 1996). Rev binds to RRE, which permits the transport of full-length or singly spliced mRNAs from the nucleus to the cytoplasm (Luciw, 1996). In the absence of Rev and/or RRE, full-length HIV-1 RNAs accumulate in the nucleus. Alternatively, a constitutive transport element from certain simple retroviruses such as Mason-Pfizer monkey virus can be used to relieve the requirement for Rev and RRE (Bray et al., 1994). The genes of interests can be inserted between RRE and the 3' LTR, or between the truncated \textit{gag} and RRE.

Efficient transcription from the HIV-1 LTR promoter requires the viral protein Tat (Luciw, 1996). Therefore, it is important that Tat is expressed in target cells if efficient transcription from the HIV-1 LTR is needed. The need for Tat expression can be met by expressing the Tat gene from the retroviral vector. Alternatively, expressing the gene of interest from a heterologous internal promoter can circumvent the need for Tat expression.

Most HIV-2-based vectors are structurally very similar to HIV-1 vectors (Corbeau et al., 1998; Corbeau and Wong-Staal, 1998; Kaye and Lever, 1998; Poeschla et al., 1998a). Similar to HIV-1-based vectors, HIV-2 vectors also require RRE for efficient transport of the full-length or singly spliced viral RNAs. Although HIV-2 also causes acquired immunodeficiency syndrome in humans, it is thought that HIV-2 is less pathogenic than HIV-1. Therefore, the HIV-2 vector is considered by some as the safer alternative to the HIV-1 vector (Poeschla et al., 1998a).

It has also been demonstrated that the HIV-1 vector can be propagated to high viral titers using viral proteins from simian immunodeficiency virus (White et al., 1999). In this system, the vector and helper constructs are from two different viruses, and the reduced nucleotide homology may decrease the probability of recombination.

In addition to vectors based on the primate lentiviruses, vectors based on feline immunodeficiency virus have also been developed as a safer alternative to vectors derived from the pathogenic HIV-1 genome (Poeschla et al., 1998b). The structures of these vectors are also similar to the HIV-1 based vectors. Preliminary studies were initiated in developing caprine arthritis-encephalitis virus-based vectors (Mselli-Lakah et al., 1998); improvement of this system will rely on better
understanding of the biology of the virus such as delineating the interaction between RRE and Rev and defining the packaging signal necessary for efficient encapsidation of the viral RNA.

C. Vectors Derived from Spumaviruses

Foamy viruses are unconventional retroviruses in that many features in their replication cycle are different from those of oncoviruses and lentiviruses. Although these viruses can be toxic to cultured cells, none of the foamy viruses are known to cause any disease in hosts (Linial, 1999). The apparent lack of pathogenicity associated with foamy viruses makes them a very attractive tool for gene delivery (Linial, 1999).

An example of a foamy virus vector is shown in Fig. 5B. This vector contains the typical retroviral cis-acting sequences. In addition to the sequences in the 5’ portion of the untranslated region, the 5’ portion of the gag open reading frame and sequences in the 3’ portion of the pol open reading frame are important for efficient packaging (Erlwein et al., 1998; Heinkelein et al., 1998; Wu et al., 1998). The requirement of pol sequences for packaging was only recently defined, and it is likely that these vectors will be improved when the minimum sequences needed for efficient packaging are delineated. Similar to the lentiviruses, expression from the human foamy virus promoter is activated by the viral protein Tas (Linial, 1999). It is important to take this into consideration when constructing foamy virus-based retroviral vectors so that efficient gene expression can be achieved in target cells.

VI. Design of Retroviral Vectors

Retroviral vectors may contain many different modifications that serve various purposes for the gene therapist. These modifications may be introduced to permit the expression of more than one gene, regulate gene expression, activate or inactivate the viral vectors, and eliminate viral sequences to avoid generation of a replication-competent virus. Some examples of these modifications are described below.

A. Standard Vectors

1. U3 Promoter-Driven Gene Expression. Full-length viral RNA is expressed from the retroviral promoter located in the U3 region of the 5’ LTR. The viral RNA contains the R, U5, 5’ untranslated region, a gene of interest, 3’ untranslated region, U3, and R. The gene inserted between the 5’ and 3’ untranslated regions can be translated from the full-length RNA that is transcribed from the U3 promoter.

During the propagation of viral stocks, it is often desirable to express a selectable marker gene in the vector so that helper cells transfected or infected by the viral vectors can be selected. Therefore, it is often necessary to design retroviral vectors that express a selectable marker gene as well as a gene of interest. Drug resistance genes are frequently used as selectable markers, but other marker genes, such as the green fluorescent protein gene, can also be used to select for transfected or infected cells. The expression of two genes in a retroviral vector can be achieved by expressing the 3’ gene by using an internal promoter, RNA splicing, or an internal ribosomal entry site (IRES). These approaches to expressing two genes from a retroviral vector are schematically outlined in Fig. 6.

2. Vectors That Use an Internal Promoter to Express Additional Genes. An example of gene expression from a retroviral vector containing an internal promoter is shown in Fig. 6A. This vector is designed to express gene A and gene B. A full-length RNA and a subgenomic RNA can be synthesized from this vector. The full-length RNA that is expressed from the viral U3 promoter is used to translate the gene A protein. The subgenomic RNA that is expressed from the internal promoter is used to translate the gene B protein.

Several retroviral vectors containing internal promoters have been successfully utilized for expression of two genes. However, potential interference between the U3 and the internal promoter can occur. For example, it has been shown that the SNV U3 promoter and the herpes simplex virus thymidine kinase promoter can interfere with each other in SNV-based retroviral vectors (Emerman and Temin, 1984a,b; Emerman and Temin, 1986a,b). Promoter interference can lead to reduced levels of transcription from either the U3 or the internal promoter. Even though some promoters do not interfere with each other, the potential for interference should be considered before the construction of retroviral vectors designed for the expression of two genes.
3. Vectors That Use Splicing to Express Additional Genes. Retroviruses express env by regulated splicing (Coffin, 1996). The splice donor site that is used to express env is located in the 5′ untranslated region of retroviruses. During replication, some full-length viral RNAs are spliced to produce subgenomic viral RNAs that are used to express the Env proteins. Splicing vectors were developed by using the same principle to express two different genes by using the viral splice donor and splice acceptor sites (Dougherty and Temin, 1986). An example of gene expression by a splicing vector is shown in Fig. 6B. In this vector, a splice acceptor site is located between genes A and B. The full-length RNA is used to translate the gene A protein, and the spliced RNA is used to translate the gene B protein. Although two mRNAs are synthesized from this vector, only the full-length viral RNA contains the packaging signal and is efficiently encapsidated into viral particles.

The advantage of splicing vectors is that only one promoter is necessary, and the potential for promoter interference is eliminated. However, the disadvantage of splicing vectors is that the efficiency of splicing can be significantly influenced by the vector RNA sequence (Dougherty and Temin, 1986). As a result, expression of two genes using splicing vectors is unpredictable, and splicing vectors are not used as commonly as other vectors.

4. Vectors That Use Translational Control Signals to Express Additional Genes. It was first demonstrated in picornaviruses that sequences in the mRNA can serve as signals that allow the ribosome to bind to the middle of an mRNA and translate a gene far from the 5′ end of the mRNA (Pelletier and Sonenberg, 1988; Jang et al., 1990; Jang and Wimmer, 1990). These sequences (named IRES), are now commonly used in retroviral vectors. An example of gene expression from a vector containing an IRES is shown in Fig. 6C. In this vector, the IRES is inserted between gene A and gene B. Only one mRNA can be synthesized from this vector; however, the mRNA is bicistronic because of the presence of IRES, and it is used to translate both gene A and gene B. Gene A is close to the 5′ cap, and ribosomes that bind to the 5′ end of the mRNA use normal mechanisms to synthesize the gene A protein efficiently. Gene B is immediately 3′ to the IRES, and ribosomes that bind to the IRES can synthesize the gene B protein in a cap-independent manner. This strategy has been used successfully to express many different genes. In general, the gene expressed with an IRES is translated less efficiently than the gene located near the 5′ end of the mRNA (Davies and Kaufman, 1992). Therefore, the gene of interest is often placed at the 5′ end near the cap, whereas the selectable marker is placed immediately 3′ to the IRES.

In addition to the IRES sequences identified in picornaviruses, IRES sequences have also been identified in the 5′ untranslated regions of some retroviruses such as MLV, SNV, and an endogenous virus like particle (VL30) (Berlioz and Darlix, 1995; Berlioz et al., 1995; Lopez-Lastra et al., 1997). Therefore, it is also possible to use these retroviral IRES sequences to express a second gene.

B. Double-Copy Vectors

The fact that the LTR sequences are duplicated in retroviral vectors has been exploited to construct vectors containing two copies of the gene of interest. An example of a double-copy vector is shown in Fig. 7A. The first set of double-copy vectors contains the gene of interest in the U3 region upstream of the viral promoter (Hantzopoulos et al., 1989; Sullenger et al., 1990a,b; Lee et al., 1992). These genes are expressed using either an RNA polymerase II promoter or an RNA polymerase III promoter. This strategy has been shown to successfully increase the level of gene expression. For example, RNA expression is increased using a double-copy vector in combination with an RNA polymerase III promoter when the desired product is a small antisense RNA or an RNA decoy (Sullenger et al., 1990a,b; Lee et al., 1992).

Another example of a double-copy vector is shown in Fig. 7B. This vector contains the gene of interest in the middle of the R region (Adam et al., 1995). The rationale for this design is based on the hypothesis that a gene is translated more efficiently when it is closer to the 5′ cap site. Because viral RNA synthesis is initiated at the 5′ end of R, the gene of interest that is cloned in the middle of the R region is within 50 nucleotides of the 5′ end of the RNA. This strategy was shown to increase the amount of protein synthesized in one of the two tested constructs but did not alter the expression level in another construct (Adam et al., 1995). Thus, it was suggested that the enhancement of gene expression is dependent upon the sequence of the gene.

C. Self-Inactivating Vectors

One safety concern associated with using retroviral vectors for gene therapy is that a replication-competent virus can be generated during propagation of the vectors, which can lead to inadvertent spread of the therapeutic vector to nontarget tissues. To address this concern, a class of vectors was designed to undergo self-
inactivation. The principle is that after gene delivery, the vector will delete some of the cis-acting elements needed to complete another round of replication. Therefore, even in the presence of a replication-competent virus, these vectors cannot be transferred to other target cells efficiently. The generation of a replication-competent virus sometimes involves recombination between the defective helper plasmid and the vector encoding the gene of interest. Therefore, another possible benefit of the self-inactivating vector is that it may decrease the probability of generating a replication-competent virus.

1. U3 Minus Vectors. U3 minus vectors were the first self-inactivating retroviral vectors to be developed (Yu et al., 1986; Dougherty and Temin, 1987). Structure of a typical U3 minus vector is shown in Fig. 8. These vectors are designed to delete the viral U3 promoter during reverse transcription so that the provirus in the target cell lacks a viral promoter. In these vectors, the U3 of the 5’ LTR is intact, whereas the U3 of the 3’ LTR is inactivated by a large deletion. The RNA generated from this vector contains R, U5, 5’ untranslated region, gene(s) of interest, 3’ untranslated region, a deleted U3, and R. During reverse transcription, the U3 at the 3’ end of the viral RNA is normally used as a template to generate the LTR. Therefore, the viral DNA that is synthesized from the U3 minus vector through reverse transcription contains deleted U3 sequences in both LTRs. Since the viral promoter is deleted during reverse transcription, the gene of interest is under the control of an internal promoter. The advantage of the U3 minus vector is that it is potentially safer, since the probability of generation of a replication-competent virus is reduced. However, at a low frequency, recombination during DNA transfection can occur to regenerate the U3 at the 3’ LTR (Olson et al., 1992). If this occurs, the resulting vector will still contain the promoter in the U3 and thus retain two complete LTRs. Additional modifications have been made in some U3 minus vectors to decrease the homology between the 5’ and 3’ LTRs (Olson et al., 1994), which reduces the probability of recombination and regeneration of an intact LTR during DNA transfection.

2. Cre/loxP Vectors. The Cre recombinase, a naturally occurring site-specific recombinase of bacteriophage P1, recognizes a 32-bp sequence named loxP. Cre can efficiently mediate site-specific recombination using two loxP sites separated by sequences of variable lengths (Hoess et al., 1984). The recombination events include deletion, insertion, and inversion of the sequences between the loxP sites. This system has been exploited to develop self-inactivating retroviral vectors (Choulika et al., 1996; Russ et al., 1996). An example of a Cre/loxP-mediated self-inactivating vector is shown in Fig. 9. This vector contains an intact 5’ LTR and all of the cis-acting elements needed for retroviral replication. The vector contains the cre recombinase gene that is expressed using an internal promoter. The 3’ LTR has been modified by insertion of several sequences in the U3, including a loxP site, a promoter, and a gene of interest; in addition, the 3’ U3 often contains a deletion to reduce the promoter activity. The full-length viral RNA is packaged into virion, and upon infection of target cells, the viral RNA is reverse-transcribed. The 3’ U3 sequence is used as a template to synthesize both LTRs; consequently, the sequences in both LTRs contain a copy of the loxP site, a promoter, and a gene of interest. The cre gene is expressed, and the Cre recombinase is synthesized in the infected target cells. The Cre recombinase then mediates the deletion of sequences between the two loxP sites in the viral DNA, which results in deletion of the 5’ LTR, the 5’ untranslated region, the internal promoter, and cre. As a result, the provirus in the target cells contains only one LTR that expresses the gene of interest.

Using the same principle, the Cre/loxP system can be used to delete different sequences in the retroviral vector as well as delete portions of the helper construct in FIG. 8. U3 minus vector. Proviral structures in helper cells and target cells are shown. RNAs are shown as thin lines. Δ, deletion in U3 region.

FIG. 9. Vector containing Cre/loxP sequences. The gene cre codes for a site-specific recombinase. The loxP sequences are substrates for Cre-mediated recombination.
the packaging cells. Another application of the Cre/loxP system is that it can be used to delete the selectable marker from a retroviral vector after the viral DNA is integrated into the chromosome of the target cells (Fernex et al., 1997). The selectable marker is included in the vector so that helper cells transfected with the vector DNA can be selected. Deletion of the selectable marker is desirable because the presence of the selectable marker can lead to promoter interference or an immune response against the transduced cells (Jung et al., 1998).

Deletion of the selectable marker is accomplished by insertion of two loxP sites that flank the selectable marker gene. After the vector is introduced into target cells by infection, the target cells are infected with another vector that expresses the Cre recombinase. The Cre recombinase then deletes sequences between the two loxP sites, which include the selectable marker. As a result, the final provirus expresses only the gene of interest.

D. Self-Inactivating and Self-Activating Vectors

Depending on the properties and effects of the gene products, it may be desirable to have an inactivated gene of interest in the helper cells and activate this gene after it is delivered to target cells. For example, if the product from the gene of interest is cytotoxic, then expressing the gene in helper cells would result in toxicity and most likely reduce or eliminate viral production. A series of vectors have been generated to simultaneously activate a gene and inactivate the vector during gene delivery (Julias et al., 1995; Delviks et al., 1997). This is accomplished by the frequent deletion of directly repeated sequences during reverse transcription. If directly repeated sequences are present in a virus, one copy of the direct repeat and all of the sequences between the two repeats can be deleted at high frequencies during reverse transcription (Pathak and Temin, 1990; Bowman et al., 1998). This property of reverse transcriptases was exploited to generate the self-activating and self-inactivating retroviral vectors (Julias et al., 1995; Delviks et al., 1997).

An example of this type of vector is shown in Fig. 10. The vector contains the gene of interest (gene A), which is split into two overlapping fragments. The first fragment contains the 5’ sequences (“gen”) and the second fragment contains the 3’ sequences (“ene A”). Both fragments contain the middle “en” portion of gene A, which forms a direct repeat. The packaging signal is placed between the two direct repeats. A functional gene A product is not expressed from the vector since the gen or the ene A fragments do not contain the complete sequence needed for expression of a functional gene A product. This vector RNA is efficiently packaged and reverse-transcribed in target cells. Reverse transcriptase can accurately delete one copy of the repeat and intervening sequences at a high frequency during reverse transcription. The resulting provirus in the target cell lacks the packaging signal and contains a functionally reconstituted gene A (Julias et al., 1995; Delviks et al., 1997). The reconstitution of gene A makes this vector self-activating, whereas deletion of the packaging signal makes this vector self-inactivating.

Depending on the length of the directly repeated sequence and the size of the packaging signal, the frequency of direct repeat deletion can be greater than 95% with a 700-bp direct repeat (Delviks and Pathak, 1999b). In addition, this strategy can be used to delete the selectable marker gene during reverse transcription so that only the gene of interest is expressed in the target cells (Delviks and Pathak, 1999a).

E. Vectors Targeted to Specific Cells

An important goal for gene therapists is to develop a means to target gene delivery to specific cell types or tissues. At least two strategies have been used in an effort to target gene delivery using retroviral vectors. One strategy is designed to control gene delivery at the point of virus entry into the host cell by using natural or
genetically engineered envelope proteins that interact with cell-type-specific receptors (for discussion, see Section IV). Another strategy is designed to control expression of the therapeutic gene in specific cell types by using tissue-specific promoters.

**F. Vectors That Utilize Cell-Type-Specific Promoters**

Promoters that are active in certain tissues or respond to certain reagents can be used to regulate the expression of a gene of interest (Cannon et al., 1996; Certo et al., 1998). These promoters can be inserted between the LTRs of a retroviral vector. Alternatively, the regulated promoter can be used to replace the viral promoter in the U3 region (Fig. 11). The design of a retroviral vector with an internal tissue-specific promoter is similar to that of other retroviral vectors containing internal promoters (Fig. 6A). In this design, a U3 minus vector should be employed to eliminate transcription from the U3 promoter in target cells so that the gene of interest is only expressed from the tissue-specific promoter. If the tissue-specific promoter is used to replace the U3 promoter, then it should be positioned in the 3′ LTR so that the tissue-specific promoter will be duplicated during reverse transcription (Fig. 11). As a result, the tissue-specific promoter will be present in the 5′ LTR of the provirus, and it will be used regulate expression of the gene of interest. Other variations of this strategy include use of inducible or constitutive promoters and modifications of the promoter structures through replacement of enhancer regions (van den Wollenberg et al., 1994; Ferrari et al., 1995).

**VII. General Considerations for Using Retrovirus Vectors and Helper Cells**

Several issues must be considered before the construction of vectors and helper cells, including control of gene expression, viral host range, virus titers that are required for efficient gene delivery, and safety of gene delivery. These issues are briefly discussed below.

**A. Gene Expression**

Generally, the level of expression of viral proteins determines the quantity of viral particles generated. Therefore, it is desirable to have a high level of viral protein expression in most helper cells to achieve high viral titers. However, if one or more of the viral proteins is cytotoxic, then it is necessary to express the viral proteins at suboptimal levels or from inducible promoters to prevent the toxicity to helper cells.

Prior to construction of the vector, it is important to consider the preferred level of expression of the gene of interest in the helper cells and the target cells. Several different types of vector are described above and similar principles can be used to construct the vectors. It is also important to realize that the promoter strength may vary in different cell types. Therefore, it is necessary to take into account the nature of the target cells. If a high level of expression is preferred in both the helper and the target cells, then it is important to select a promoter that is highly active in both cells. Other strategies should be used if different levels of gene expression are desired in target cells and helper cells. For example, if the gene of interest contains a toxic protein, then it is important that gene expression be suppressed or prevented in helper cells by using inducible promoters or self-activating vectors. In addition to the selection of an appropriate promoter, it is also important to examine the sequences that will be placed in the vector for undesirable regulatory signals. For example, inadvertent insertion of splice donor and acceptor sites or polyadenylation signals can severely affect the efficiency of virus production and gene expression.

**B. Virus Host Range and Titers**

1. **Considerations for Envelope Selection and Virus Host Range.** The nature of the viral envelope protein determines whether a certain virus can enter a target cell. Therefore, it is important to consider whether the target cells have the correct cell surface receptor before the selection of an envelope protein that will be used for virus production. There are many examples of successful interactions between the Gag proteins of one retrovirus with the Env protein of a different retrovirus that result in functional pseudotyped viruses (Miller and Buttimore, 1986). However, not all Gag and Env proteins can interact to produce functional pseudotyped viruses (Linial, 1999). If an unreported Gag and Env combination is being utilized to generate pseudotyped viruses, it is desirable to characterize the nature of interactions between Gag and Env. The Gag polyprotein determines the specificity of vector RNA packaging; therefore, the vector used has to contain the proper packaging signal for efficient virus production.

2. **Ping-Pong Amplification.** Several strategies have been developed to achieve high virus titers. One strategy, referred to as the ping-pong method, uses cocultivation of two packaging cells that express different envelope proteins (Kozak and Kabat, 1990). The expression of the envelope protein in cells can reduce the efficiency of infection with viruses containing the same envelope protein; this is called superinfection interference. Cocultivation of the two packaging cells results in efficient infection of each packaging cell line with a virus produced from the other packaging cell line. As a result, multiple copies of the vector accumulate in each cell, which results in dramatically increased virus production. It must be noted however that high titer virus preparations produced with the ping-pong method often contain replication-competent viruses that presumably are generated through high frequency recombination.

3. **Concentration of Viruses.** Another strategy to achieve high viral titers involves concentration of virus particles using physical methods. Virus concentration
can be achieved by commercially available concentrators or by ultracentrifugation of virus particles. It is important to determine whether the process of virus concentration affects the infectivity of the virus. If the viral particles are concentrated 100-fold, but the virus infectivity is reduced 10-fold, then only a 10-fold increase in the virus titer is achieved.

In comparison to some DNA viruses, retroviruses are much more labile and significant loss of virus infectivity is observed during various procedures designed to concentrate virus. However, it has been shown that VSV G pseudotyped viruses can be successfully concentrated by ultracentrifugation without loss of virus infectivity; as a result, virus titers of up to $10^7$ to $10^9$ infectious units per ml can be achieved (Yee et al., 1994a,b). Therefore, retroviruses with high infectious titers can be generated through physical concentration of the viruses under appropriate conditions.

C. Safety Concerns Associated with Retrovirus-Based Gene Therapy

Possible transmission of a replication-competent virus to patients is a major safety concern associated with using viral vectors for gene therapy. Replication-competent MLV has been shown to cause lymphomas in primates used in gene therapy experiments (Donahue et al., 1992; Vanin et al., 1994; Purcell et al., 1996). Because of this concern, all clinical grade vector viruses are extensively tested for the presence of replication-competent viruses. In addition, the helper constructs and retroviral vectors are carefully designed to reduce the probability of generating a replication-competent virus through recombination.

Replication-competent viruses can either be generated through recombination between the helper constructs and the vectors, or they can be produced through recombination between an endogenous retrovirus (present in the host genome) and the helper construct or the vector. Recombination events can occur during DNA transfection or during reverse transcription of the virus (Hu and Temin, 1990). Regardless of the mechanism, the frequencies of recombination are augmented by the extent of homology between the helper construct and the retroviral vector. Therefore, several retroviral vectors and helper constructs have been modified to reduce the extent of homology. In addition, other features have also been incorporated into these systems. For example, self-inactivating vectors discussed earlier that can delete cis-acting sequences important for viral replication such as the packaging signal were generated. This reduces the probability of the helper construct obtaining the packaging signal from the vector. Similarly, split-genome packaging cell lines discussed earlier were developed to reduce the probability of generating a replication-competent virus. In addition, gene delivery systems containing vector and helper constructs from different viruses have also been developed. For example, HIV-1 507 core proteins have been used to package SIV vectors, and SNV core proteins have been used to package MLV vectors. The reduced nucleotide homology between helper construct and viral vector should reduce the probability of recombination and the generation of replication-competent virus (White et al., 1999). However, this approach has its limitations because these are the only two known examples of retroviruses in which the viral Gag/Gag-Pol protein of one virus can support the replication of the viral genome of another virus (Embretson and Temin, 1987a; Rizvi and Panganiban, 1993; Certo et al., 1998, 1999).

Endogenous retroviruses that are related to MLV are present in the genomes of all mouse cells, and they also play a role in the generation of replication-competent viruses (Purcell et al., 1996). To avoid possible recombination with endogenous MLV-like viruses, some recently constructed helper cell lines have been derived from human cells that are not expected to contain endogenous MLV-like viruses (Ory et al., 1996).

Another potential safety concern is that retroviruses may integrate into host genes and inactivate essential functions or activate cellular oncogenes by insertion of promoters (Coffin, 1996). These genetic events have been observed in experimental animals infected with replication-competent retroviruses (Coffin, 1996). However, the number of infectious events carried out in an experimental animal to observe these rare events are far greater than those carried out in a gene therapy trial using vectors devoid of replication-competent viruses. Therefore, these rare events may pose a theoretical rather than a realistic safety concern.

VIII. Gene Therapy Applications and Future Directions

A. Retroviral Vectors Used in Gene Therapy Clinical Trials

Gene therapy approaches involving retroviral vectors can be used to treat several different types of human diseases. Diseases that are amenable to gene therapy include genetic disorders, a variety of cancers, graft-versus-host disease after bone marrow transplantation, and some infectious diseases. Gene therapy approaches to a variety of diseases are being tested in many clinical trials. The results of some of these clinical trials have been discussed in some recent reviews (Hwu and Rosenberg, 1994; Blaese, 1995a,b; Breau and Clayman, 1996; Dunbar, 1996; Lotze, 1996). A few examples of gene therapy clinical trials involving retroviral vectors are briefly described below.

The first clinical trial of human gene therapy was designed to correct a genetic disorder known as adenosine deaminase (ADA) deficiency (Blaese et al., 1995). The patients lack adenosine deaminase, which results in severe combined immunodeficiency. In this gene therapy protocol, the patients’ white blood cells were isolated...
and infected with an MLV-based vector that expressed ADA and a neomycin phosphotransferase gene (neo), which served as a selectable marker. Cells infected with the vectors were selected for resistance to G418 (a neomycin analog). These infected cells were then introduced back into the patients. This treatment improved the patients’ physical condition and the vector provirus containing the ADA was detected in the patients’ blood after several years, indicating that the gene therapy was long-lasting (Blaese et al., 1995).

Several different gene therapy strategies are being developed in an effort to treat a variety of cancers. These strategies include elimination of cancer cells by suicide gene therapy (Oldfield et al., 1993), reversion of cancer cells to normal cells by delivery of a functional tumor suppressor gene (Roth et al., 1996), and modification of cancer cells to elicit stronger immune responses (Lotze et al., 1994).

Another potential application of gene therapy is to prevent severe graft-versus-host disease that often results from allogeneic bone marrow transplantation. In a gene therapy clinical trial, the bone marrow donors' lymphocytes were first transduced with a retroviral vector encoding the herpes simplex virus thymidine kinase gene (HSV-tk) (Bonini et al., 1997). HSV-tk is not toxic by itself; however, HSV-tk can phosphorylate a nontoxic prodrug named ganciclovir (GCV) to activate the toxicity of the drug. The HSV-tk-expressing cells were then used for bone marrow transplantation. Patients that developed graft-versus-host disease were treated with GCV to eradicate the donor cells that were mounting an immune response against the host. The results of this clinical trial showed that this approach is effective in controlling graft-versus-host disease (Bonini et al., 1997).

Retroviral vector-based gene therapies have also been used in clinical trials to treat HIV-1 infection. Generally, these treatments have involved modification of the syngeneic lymphocytes ex vivo using retroviral vectors and are designed to suppress the expression of viral genes. These strategies include use of antisense RNA, mutant trans-dominant regulatory proteins, or ribozymes that are targeted to cleave viral RNA (Nabel et al., 1994; Morgan and Walker, 1996; Wong-Staal et al., 1998). Another approach is to modify autologous fibroblasts to express a part of the HIV Env so that a host immune response can be elicited (Galpin et al., 1994).

B. Future Directions

Recently, retroviral-mediated gene therapy was demonstrated to provide full correction of severe combined immunodeficiency (SCID)-X1 phenotype in two patients (Cavazzana-Calvo et al., 2000). This exciting advancement clearly demonstrated that gene therapy can achieve clinical benefit and provide treatment and/or cure for certain diseases. However, many improvements have to be made before the full potential of retroviral-based gene therapy can be exploited; some of the directions for improvement are gene delivery efficiency, long-term expression of the delivered genes, gene delivery in vivo, cell-type-specific targeting, and safety. For example, many of the therapy trial protocols involve removing the target cells from the patients, delivering the gene of interest ex vivo, and putting the cells back into the patients. Although this approach works well for certain diseases involving gene delivery to the blood cells such as SCID-X1, it is not applicable to many other diseases involving gene delivery into solid tissues. Directly delivering viral particles into patients involves overcoming technical hurdles such as physically placing the particles near the target tissue, stability of the virus particles in human blood/tissues, targeting the vector to the correct tissue, the effects of virus infecting nontarget cells, and safety of the therapy. These and other technical problems must be solved before using retroviral-based gene therapy to treat certain diseases.

As described above, many retroviral systems have been developed for gene delivery. It is unlikely that one system will be the best tool for gene therapy of all diseases; instead, it is likely that the advantage of each system can be exploited to develop treatment for different diseases. For example, some lentiviruses can infect nondividing cells whereas MLV only infect dividing cells. If the goal were to deliver genes into neural cells for gene correction, it would be more effective to use lentivirus-based vectors for the delivery. In contrast, if the goal were to deliver suicide genes to treat cancer cells in the neural system, then MLV-based vectors would be attractive because they could target the dividing cancer cells. Therefore, understanding the biology of the viruses and the nature of the target cells is important in the selection of a suitable gene delivery system.

Acknowledgments. This work was supported by Research Grants MBC-97322 (to W.S.H.) and VM-84706 (to V.K.P.) from the American Cancer Society, Research Grant CA58875 from the National Institutes of Health Public Health Service (to V.P.), and intramural funding from HIV Drug Resistance Program of the National Cancer Institute.

REFERENCES


