Pharmaceutical Research, Vol. 23, No. 6, June 2006 (© 2006)
DOI: 10.1007/s11095-006-0164-2

Expert Review

Plasmid Engineering for Controlled and Sustained Gene Expression for Nonviral Gene Therapy

Ethlinn V. B. van Gaal,[1] Wim E. Hennink,[1] Daan J. A. Crommelin,[1] and Enrico Mastrobattista[1,2]

Received November 21, 2005; accepted January 3, 2006

Abstract. Gene therapy requires the introduction of genetic material in diseased cells with the aim of treating or ultimately curing a disease. Since the start of gene therapy clinical trials in 1990, gene therapy has proven to be possible, but studies to date have highlighted the difficulty of achieving efficient, specific, and long-term transgene expression. Efforts to improve gene therapy strategies over the past years were mainly aimed at solving the problem of delivery, without paying much attention to the optimization of the expression cassette. With the current understanding of the eukaryotic transcription machinery and advanced molecular biology techniques at our disposition, it has now become possible to create custom-made transgene expression cassettes optimized for gene therapy applications. In this review, we will discuss several strategies that have been explored to improve the level and duration of transgene expression, to increase control over expression, or to restrict transgene expression to specific cell types or tissues. Although still in its infancy, such strategies will eventually lead to improvement of nonviral gene therapy and expansion of the range of possible therapeutic applications.

KEY WORDS: expression; nonviral gene therapy; plasmid; regulation; transcription.

[1] Department of Pharmaceutics, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, P.O. Box 80082, 3508 TB, Utrecht, The Netherlands.
[2] To whom correspondence should be addressed. (e-mail: e.mastrobattista@pharm.uu.nl)

ABBREVIATIONS: AAV, adeno-associated virus; ACF, ATP-utilizing chromatin assembly and remodeling factor; AD, activation domain; AFP, α-fetoprotein; ATF/CREB, activating transcription factor/cAMP-responsive element binding protein; BAC, bacterial artificial chromosome; bFGF, basic fibroblast growth factor; BKV, BK virus; BPV-1, bovine papilloma virus 1; CArG, CA(A/T)6GG; CAT, chloramphenicol acetyltransferase; ChMAR, chicken lysozyme MAR; CMV, cytomegalovirus; CMV IE, cytomegalovirus immediate early; CpG-N, neutralizing CpG motif; CpG-S, stimulatory CpG motif; Cre, cyclization recombination recombinase; CRSP, cofactor required for Sp1; CTF, CCAAT transcription factor; DBD, DNA binding domain; DR, direct repeat; DS, dyad symmetry; DTS, DNA nuclear targeting sequence; *EBNA1*, EBV nuclear antigen 1; EBP2, EBNA1-binding protein 2; EBV, Epstein–Barr virus; Ec, ecdysone; *Egr-1*, early growth response-1 gene; FKBP, FK506 binding protein; FLP, flippase; FR, family of repeats; FRAP, FKBP12 rapamycin-associated protein; GFP, green fluorescent protein; Grp78, glucose-regulated protein 78; HAC, human artificial chromosome; hFIX, human coagulation factor IX; HIV, human immunodeficiency virus; HMGB1, high-mobility group box 1; HRE, hypoxia responsive element; HSE, heat shock element; HSF1, heat shock factor 1; hsp, heat shock protein; HSVtk, herpes simplex virus thymidine kinase; IL-12, interleukin-12; iNOS, inducible NO synthetase; IR, inverted repeat; IRES, internal ribosome entry site; I-*Sce*I, *Saccharomyces cerevisiae* mitochondrial endonuclease; ITAF, IRES *trans*-acting factor; KDR, kinase-like domain receptor; *loxP*, Locus of Crossing over of P1 phage; MAEC, mammalian artificial episomal chromosome; Mdr, multidrug resistance; MeCP, methyl-CpG-binding proteins; Mfp, mifepristone; MLV, murine leukemia virus; MRI, magnetic resonance imaging; NF-κB, nuclear factor-κB; NLS, nuclear localization signal; NOEL, no-observed-effect level; NPC, nuclear pore complex; NURF, nucleosome-remodeling factor; ORF, open reading frame; oriP, origin of plasmid replication; PBAF, polybromo BRG1-associated factor; PDGF, platelet derived growth factor; PECAM-1/CD31, platelet-endothelial cell adhesion molecule-1/CD31; PR-LBD, progesterone receptor ligand binding domain; rAAV, recombinant adeno-associated virus; RSF, remodeling and spacing factor; RSV, Rous sarcoma virus; RSV-LTR, Rous sarcoma virus long terminal repeat; RTA, recombinant transcriptional activator; S/MAR, scaffold/matrix attachment region; SB, Sleeping Beauty; SI, sucrase–isomaltase; SMGA, smooth muscle gamma actin; SV40, simian virus 40; SWI/SNF, switching/sucrose nonfermenting; Tag, large T antigen; Tet, tetracycline; TetO, Tet operator; TetR, Tet repressor; TGF-α, transforming growth factor α; TNFα, tumor necrosis factor α; TRAP, thyroid hormone receptor-associated protein; tTA, tetracycline transactivator; tTS, tetracycline transcriptional silencer; VEGF, vascular endothelial growth factor; VP16, virion protein 16; vWF, von Willebrand factor; WAF1, wild-type *p53*-activated fragment 1; YAC, yeast artificial chromosome; β-gal, β-galactosidase.

INTRODUCTION

Gene therapy can be defined as an approach to treat, cure, or ultimately prevent a disease by replacing defective genes, introducing new genes, or changing the expression of a person's genes. This concept was first proposed in 1972 (1), but to date clinical applications remain few due to inefficiency of gene delivery. Over the past years, much effort has been made to develop strategies for effective delivery of DNA to the nucleus of target cells. Although viral vectors have been widely examined (2) and are still regarded as the most efficient, their use is limited due to safety issues, DNA loading capacity, and difficulties in scale-up production (3). Alternatively, nonviral delivery strategies have been developed including physical delivery of naked DNA and gene delivery by using chemical carriers such as cationic polymers, lipids, detergents, and peptide-based technologies, which have been reviewed elsewhere (4–9). Although many reviews about nonviral gene delivery focus on optimizing the carrier and its entry mechanisms into the cell, not much attention has been given to the plasmid or DNA part of the nonviral carrier. Nevertheless, optimization of the plasmid vector can lead to increased or prolonged levels of expression and may therefore play an important role in compensating for the limited transfection efficiency achieved with most nonviral carriers. Moreover, plasmid engineering can be used to increase levels of specificity and control over protein expression. This is referred to as *transcriptional targeting* or *transcriptional control*. In this review, we focus on progress made in plasmid optimization over the past years and its value for nonviral gene therapy. The subject will be discussed systematically, starting with a description of the minimal requirements for expression of exogenous DNA and subsequently expanding the system with strategies aiming at improvements toward nuclear uptake, restriction of expression to target cells, external control over expression, prolongation of expression, and ending with an overview of diverse strategies for further customization.

MINIMAL REQUIREMENTS FOR EXPRESSION OF EXOGENOUS DNA IN EUKARYOTIC CELLS

In gene therapy, the host transcription machinery is exploited for expression of exogenous DNA. This requires a thorough understanding of the mechanism of transcription and identification of the essential features necessary for transgene expression. Although there is still much to learn about the exact mechanisms underlying transcriptional control of a gene [for a recent review, see Orphanides and Reinberg (10)], current knowledge warrants rational design of exogenous DNA expression cassettes. The minimal requirements for plasmid production by replication in a prokaryotic host and expression of the therapeutic gene in eukaryotes are summarized in Table I.

The process leading from a gene to a functional protein in eukaryotes includes transcription of the gene to primary RNA by RNA polymerase II, processing of primary RNA to mRNA, export of mRNA to the cytoplasm, and translation of mRNA into a protein. In most cases, initiation of transcription is the most important point of control (11). This will therefore be the focus of this review.

Transcription is a complex process requiring many proteins (>100 individual subunits) to assemble at gene control regions (Fig. 1). A gene control region is defined as the whole expanse of DNA involved in regulating transcription of a gene, including the promoter and all regulatory sequences to which gene regulatory proteins bind (i.e., enhancers, silencers). Transcription regulating proteins can be divided into general transcription factors and additional (luxury) transcription factors. General transcription factors must form a preinitiation complex whose function is to unwind the DNA helix, separate the DNA strands for use as a template, and enable RNA polymerase II to take up its position so that mRNA synthesis can start. This complex is already capable of inducing transcription at a slow rate, but additional factors are required for high-level synthesis and specificity. These factors are inducible and enable transcription to speed up or slow down in response to cellular signals. Gene regulatory proteins function as regulatory units that are used to generate complexes whose function depends on the final assembly and composition of all the individual components (11). Many gene regulatory proteins bind to DNA as either homodimers or heterodimers. This mixing of protein subunits allows for formation of many different proteins with varying DNA-binding specificities (11).

Formation of the preinitiation complex and initiation of transcription is followed by the elongation phase, during which transcription continues. An important aspect within the elongation phase is the processing of a primary RNA transcript to mRNA by removal of intron sequences (splicing), 5′ end capping and polyadenylation on the 3′ end (11,13). Proper capping is essential as it allows the cell to assess whether both ends of an mRNA molecule are present (indicative of an intact message) before exporting it from the nucleus to the cytoplasm. The enzymes involved in 5′ end capping act in a sequence-independent fashion, whereas the enzymes responsible for 3′ end modification require the presence of a consensus sequence for recognition (Table I). The actual polyadenylation is performed by the enzyme poly-A polymerase that in contrast to other RNA polymerases does not require a template for its action; hence, the poly-A tail is not encoded in the DNA.

RNA polymerase II moves along the sense strand until it reaches a termination signal where it is forced to stop and dissociate from the DNA. The simplest example of a stop signal is a palindromic GC-rich region followed by an AT-rich region, but other stop signals exist that may or may not require additional factors to terminate the transcription process (11,13).

The complexity of the eukaryotic transcription machinery may at first have seemed an obstacle impairing transgene expression. However, this review attempts to illustrate how it can be exploited to elegantly modulate expression of exogenous DNA, thereby significantly contributing to improvement of nonviral gene therapy and expanding the range of possible therapeutic applications. Once the basic fulfillments are met to have the exogenous DNA participate in the transcriptional process, additional features can be introduced to tailor the expression profile to one's needs. By interfering at the level of transcription initiation, increased efficiency, specificity, and duration of expression can be established.

Table I. Overview of Essential Elements Required for Plasmid Vector Expression with their Definitions and Functions

Element	Definition	Function
Plasmid vector		
Requirements for plasmid production in bacteria		
Origin of replication (ori)	A specific DNA sequence of 50–100 bp to which the bacterial host-cell enzymes bind, initiating and regulating replication	Plasmid replication
Selection marker	DNA sequence encoding a protein that provides bacteria with a certain selectable characteristic	Selection of bacteria containing plasmid of interest Enables pressurization for plasmid maintenance
Requirements for functionality in eukaryotes		
Promoter	Shortest DNA sequence at which RNA polymerase II can initiate transcription	Position the start site for RNA synthesis Regulate frequency of transcriptional initiation
Enhancer	DNA sequences to which gene activator proteins bind resulting in interaction with a promoter	Enhance a cell's capacity to transcribe a gene with greater efficiency and greater sensitivity to changes in the environment
Poly-A signal	A recognition site existing of AAUAAA hexamer positioned 10–30 nucleotides upstream the 3′ end and a GU- or U-rich element located maximally 30 nucleotides downstream of the 3′ end	Induces 3′ end capping
Intron	Intervening noncoding DNA	Protect transcripts against rapid degradation Promote export of mRNA
Stop signal	DNA sequence at which RNA polymerase II is forced to stop and dissociate from the DNA	Termination of transcription and dissociation of RNA polymerase II
Coding DNA	DNA encoding a therapeutic protein	Therapeutic effect
Host cell		
RNA polymerase II	Enzyme that performs transcription of DNA to RNA	Transcription of DNA to RNA
Poly-A polymerase	Enzyme that adds ~200 A nucleotides to the 3′ end	Provide synthesized RNA strand with the 3′ end that protects it against nuclear degradation
General transcription factors	Proteins that assemble on all promoters used by RNA polymerase II	Help to position RNA polymerase II at promoter Aid in separating the two strands of DNA Release RNA polymerase II from promoter once transcription has begun
Luxury transcription factors	Gene regulatory proteins that bind to regulatory sequences other than the promoter	Regulation of transcription is dependent on cell type and/or environmental influences

INTRACELLULAR TARGETING: DNA NUCLEAR TARGETING SEQUENCES

In gene therapy settings, plasmids generally rely on breakdown of the nuclear envelope during cell division for nuclear entry. However, because many cells targeted in gene therapy do not divide or divide very slowly, nuclear entry is a major limiting step in achieving gene expression (14). Nuclear uptake has been described to occur spontaneously *in vivo* when exceeding a certain threshold (>1,000,000 plasmids/cell), possibly due to mass action (15), but this clearly does not reflect a realistic nor desirable situation.

In the absence of mitosis, plasmids have to be imported via the nuclear pore complex (NPC) to obtain transgene expression. This can be achieved by attaching proteins or synthetic peptides containing nuclear localization signals (NLSs) to the plasmids. However, no consensus has been reached with regard to their beneficial effects (16). More recently, approaches have been described based on incorporation of peptide/protein structures that allow alternative interactions with a host cell's nuclear import mechanisms, such as targeting the importin β nuclear import receptor (17) and the use of steroid receptors as shuttles to facilitate nuclear import (16,18). Both of these methods require the covalent or noncovalent attachment of peptides or proteins to DNA, which, from a pharmaceutical point of view, is undesirable.

A perhaps more elegant way of enhancing nuclear uptake lacking the necessity for coupling of peptides/proteins to the DNA is the incorporation of DNA sequences that are recognized by endogenous transcription factors involved in nuclear import. Such DNA nuclear targeting sequences (DTS) (16) might facilitate nuclear localization by inducing

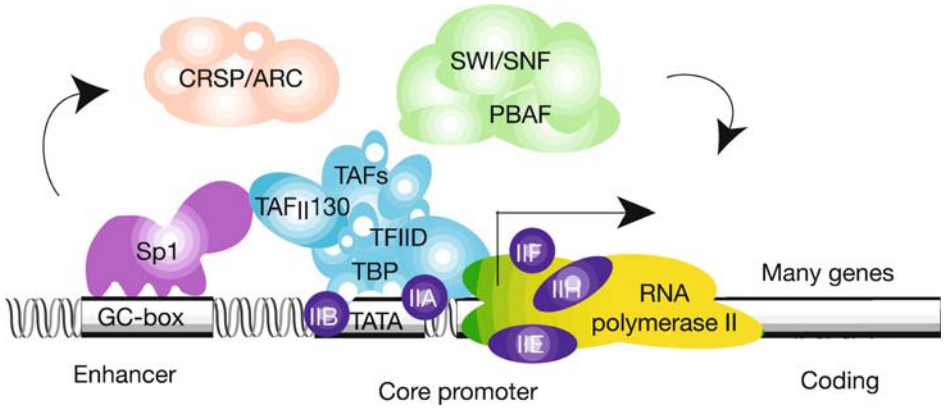

Fig. 1. A simplified schematic model of the eukaryotic transcriptional apparatus. It comprises three broad classes of multisubunit assemblies: 1) the RNA polymerase II core complex and associated general transcription factors (TFIIA, -B, -D, -E, -F and -H), 2) multisubunit cofactors (mediator, CRSP, TRAP, and others), and 3) various chromatin modifying or remodeling complexes (SWI/SNF, PBAF, ACF, NURF, and RSF). CRSP, cofactor required for Sp1; TRAP, thyroid hormone receptor-associated protein; SWI/SNF, switching/sucrose nonfermenting; PBAF, polybromo BRG1-associated factor; ACF, ATP-utilizing chromatin assembly and remodeling factor; NURF, nucleosome-remodeling factor; RSF, remodeling and spacing factor. Adapted from Levine and Tjian (12).

(partial) coating of the plasmid with NLS-containing proteins and subsequent binding to importins.

Simian virus 40 (SV40), a virus whose regulatory sequences are often used to drive expression of plasmids, was shown to elicit a nuclear localization effect (14). Analysis of the sequence requirements for nuclear import has led to the identification of a 72-bp repeat within the SV40 enhancer (14). The nuclear localization effect of this 72-bp repeat has been demonstrated both *in vitro*, in various cell types including epithelial, endothelial, and smooth muscle cells from a variety of organisms (19), as well as *in vivo* (20,21).

The 72-bp element increased cytomegalovirus (CMV)-promoter-driven gene expression by as much as 20-fold after 7 days in murine tibialis muscle *in vivo* (20). Because the CMV promoter is very strong, this effect is likely to be ascribed to enhancement of nuclear localization rather than further enhancement of transcriptional activity (20). This hypothesis is further supported by the observation that expression of plasmids including the SV40 72-bp repeat is increased in comparison to plasmids lacking this sequence in postmitotic cells but not in dividing tumor tissue (20). SV40 contains numerous binding sites for different general transcription factors; it is thought that binding of NLS-containing transcription factors accounts for the nuclear localization (14,16). However, several other viral [CMV, Rous sarcoma virus (RSV)] and cellular promoters/enhancers that contain similar binding sites tested negative for this effect (14,16). What distinguishes SV40 from these other structures is not yet fully understood. One possible explanation is that some, but not all, transcription factors are able to induce nuclear uptake. It is thought to be essential that the NLS and the DNA-binding domain within the transcription factor are sufficiently separated both functionally and spatially (16).

Other sequences proposed to have nuclear targeting capacity include nuclear factor-κB (NF-κB) binding sites (22) and the origin of plasmid replication (oriP) sequence from the Epstein–Barr virus (EBV) (16,23). Additionally, sequences have been identified to act as DTSs in a cell-specific manner, including the smooth muscle gamma actin (SMGA) (24) and flk-1 promoter (16), with specificity for smooth muscle cells and endothelium, respectively. When using viral elements, possible risks of immunogenicity due to high CpG content should not be neglected (see also under "Sustained Expression").

Rather than attempting to overcome the barrier of nuclear uptake through utilization of endogenous proteins for nuclear import, gene therapy strategies have been developed to avoid this barrier by way of cytoplasmic expression. All expression systems described so far rely on the endogenous, nuclear transcriptional machinery of the cell for expression of the transgene. In contrast, Gao and Huang (25) developed a strategy in which the transgene is codelivered with an exogenous transcription machinery that is insensitive to the endogenous regulatory mechanism. This exogenous machinery consists of a bacteriophage T7 RNA polymerase driving transcription of a gene controlled by the T7 promoter. Importantly, phage RNA polymerases exhibit striking specificity for their promoters (26). Several complications are involved when using the exogenous transcription machinery. First, because localization of T7 RNA polymerase and therefore transcription of the foreign gene is restricted to the cytoplasm, transcripts will not be capped properly. To achieve efficient translation of the uncapped transcript, an internal ribosome entry site (IRES) of the encephalomyocarditis virus was therefore inserted into the 5′ untranslated region. In most eukaryotes, translation of mRNA requires the presence of an initiation codon that can be recognized by ribosomes. However, in some viral mRNA, ribosomes start translation at internal sites in the mRNA. These internal ribosomal entry sites can be shuttled from their viral settings to unrelated genes to enable expression of proteins in the absence of a functional initiation codon. Use of these constructs allowed one to observe significant levels of reporter gene expression in a variety of mammalian cells. A limitation of this expression system is the rapid turnover of T7 polymerase enzyme, which means that expression for long periods of time cannot be sustained. To

establish a continuous supply of the polymerase, Brisson *et al.* (27) developed an "autogene" (pCMV/T7-T7pol) consisting of the T7 polymerase gene driven by a CMV and T7 promoter (Fig. 2). The CMV promoter is used to drive the first round of expression of T7 RNA polymerase, which can then act on the T7 promoter to drive expression of both new T7 RNA polymerase and the target gene. This way, no codelivery of the T7 RNA polymerase enzyme, which can be very expensive and potentially immunogenic, is required. Obviously, in this model some nuclear uptake is essential. However, the required amount is thought to be much less than in "classic" nuclear expression models and therefore easier to achieve. The new T7 RNA autogene was shown to induce higher, more sustained levels of reporter gene expression than observed with the autogene lacking the CMV promoter or with nuclear expression systems in which a reporter gene was driven by a CMV promoter only. Additionally, pCMV/T7-T7pol is easily amplified and purified from bacteria by using standard methods, possibly because binding of T7 RNA polymerase to the T7 promoter is sterically hindered by CMV promoter-induced conformational changes (27). It is noteworthy that no detectable quantities of antibodies against T7 RNA polymerase were generated in mice upon direct injection of a T7 system, as reported by Chen *et al.* (28).

TRANSCRIPTIONAL TARGETING: RESTRICTING TRANSGENE EXPRESSION TO SPECIFIC CELLS OR TISSUES

When the first expression cassettes for gene therapy were developed, viral elements were used to drive expression of the foreign gene. Among the strongest promoters identified to date are the CMV immediate early (CMV IE)

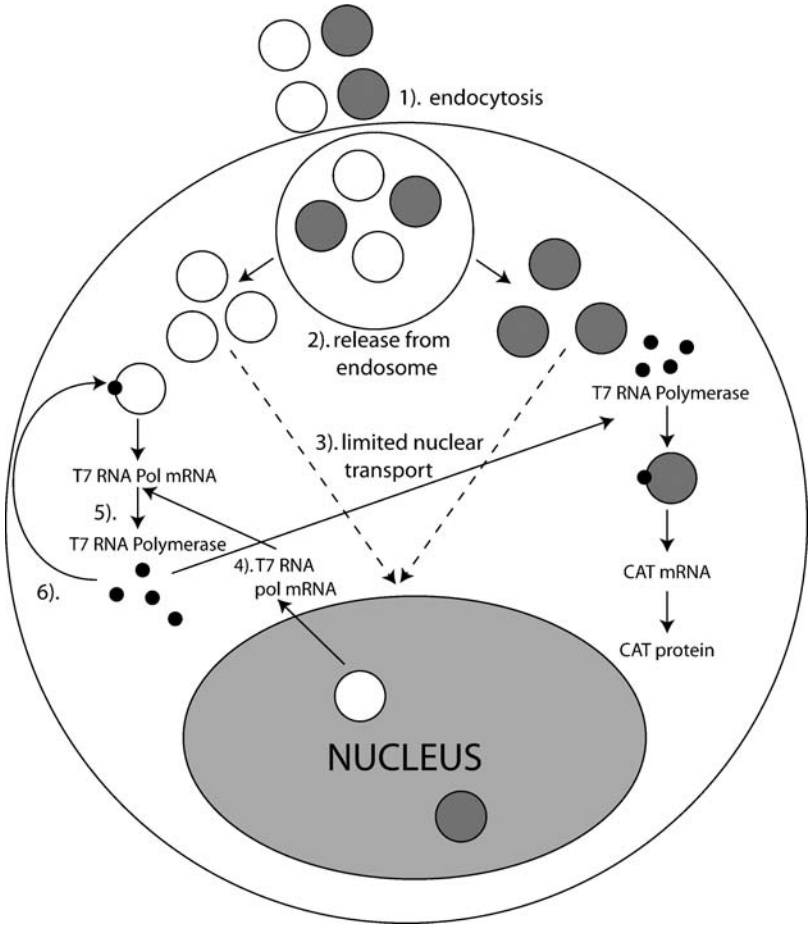

Fig. 2. Proposed mechanism of action of the pCMV/T7-T7pol autogene. Codelivery and endocytosis of pCMV/T7-T7pol (large white circles) and pT7-CAT (large gray circles). A small portion of pCMV/T7-T7pol is translocated into the nucleus and transcribed by RNA polymerase II to generate T7 RNA polymerase mRNA. In the cytoplasm, T7 RNA polymerase mRNA is translated into T7 RNA polymerase (small black circles) which can then drive further T7 RNA polymerase expression in the cytoplasm by way of the T7 promoter on pCMV/T7-T7pol and simultaneous expression of the chloramphenicol acetyltransferase (CAT) gene located on pT7-CAT. This way, pCMV/T7-T7pol requires no addition of exogenous T7 RNA polymerase and takes advantage of the excess DNA remaining in the cytoplasm due to inefficient nuclear import. Adapted from Brisson *et al.* (27) with permission.

promoter/enhancer, RSV long terminal repeat (RSV-LTR), and SV40 regulatory sequences. Although the viral promoters are still considered strong relative to the cellular promoters applied more recently, their popularity has diminished considerably in the past years. Reasons include their lack of specificity and the observation that structural differences between host and foreign DNA are registered by the host and provoke immunostimulatory and silencing effects resulting in inactivation (see below). Therefore, a need arose for nonviral, cellular gene regulatory regions that are not as easily inactivated. In addition, targeting of genetic medicines to specific cells is often required to prevent toxicity to healthy cells [especially to liver and bone marrow cells (29)] and to decrease the required dose. Differential gene expression among cell types and environmental conditions is physiologically possible because different genes are driven by different (combinations) of promoter and enhancer sequences, and each of these regulatory sequences contain binding sites for multiple transcription factors. Changing the set of transcription factors within a cell leads to activation of a different set of genes, ultimately leading to a change in a cell's protein expression profile (12). Selective expression of transgenes in specific cells or tissues can be achieved by constructing DNA expression cassettes that contain regulatory regions that are recognized by transcription factors specifically present or selectively expressed by the target cell population. This so-called transcriptional targeting can be based on tissue specificity where transcription is directed specifically among healthy tissues, and on tumor specificity by using elements that are active selectively in tumor cells due to aberrant gene expression or tumor biology.

Tissue-Specific Promoters

There is an increasing list of well-characterized regulatory elements controlling cell type specific expression, with target tissues including the pancreas (30–32), breast (30,31, 33–35), bone (29,30,36–38), brain (30,31,36,39), melanocytes (29–31,34,36,40–43), kidney (31), bladder (31), prostate (29–31,36,40,44), testes (31), connective tissue (31), muscle (31,36, 37,42), endothelium (29,30,36,45–62), liver (30,31,34,36,40), GI tract (57), lung (30,31,36), epidermis (63), thyroid (39), hematopoietic cells (30,31,34), and ovary (29,37). Tissue-specific promoters display a natural activity in normal tissues without discriminating diseased from healthy cells. Therefore, in cases where toxic genes are to be expressed, their utility as such is limited to dispensable tissues such as melanocytes, prostate, breasts, endocrine, and exocrine tissues (30). Combining tissue-specific promoters with additional targeting moieties can further increase their utility. For example, combination of tissue and tumor specific promoters may enable targeting of specific cells/malignancies within nondispensable tissues. A complication involved when using cellular regulatory elements is their low activity *per se*. To obtain a therapeutic effect, usually combination with a strong promoter element (either viral or cellular) is required.

Tumor-Specific Promoters

One of the most important factors limiting success of conventional cancer therapy strategies is their lack of tumor selectivity. Available cytotoxic agents have been reported to have a cancer-to-normal cell therapeutic ratio as low as 2:1 to 6:1. Novel approaches using gene therapy strategies have been reported to succeed in increasing this ratio to 10,000:1 (30). Minimizing or excluding inappropriate expression in surrounding nontarget cells is of great importance for limiting adverse effects and, therefore, for increasing the therapeutic index. Considering the complexity of the biological nature of tumors, the genetic alterations that lead to the malignant phenotype of tumor cells, and the high genetic mutation rates, it is difficult to address universal tumor features. However, an increasing number of key events involved in the process of tumorigenesis and characteristics of tumor tissues are being discovered. Employing regulatory elements that direct transcription in response to these key events may enable targeting of gene expression to tumors of various origins (Table II). Design of tumor-specific expression vectors can be based on aberrant gene expression profiles in tumor cells or on typical tumor biology. Aberrant gene expression may be a matter of mutations to certain genes or errors in the process of their transcription. In other cases, the genetic expression pattern is affected; for example, reexpression of embryonic genes or expression of viral genes may occur. To date, no genes have been identified to be completely restricted to tumor cells; when genes are referred to as "tumor specific," it is meant that they are found at much higher levels in tumor cells than in normal cells. Tumor growth has several biological consequences that affect vasculature and metabolism. Fast-growing tumors are often poorly vascularized; this poor vascularization, in combination with the solid character of the tumor tissue, results in a high interstitial and low intravascular pressure, leading to a decrease in nutrient supply and, ultimately, necrosis of the tumor core. Typical hallmarks of the microenvironment of such tumors consequentially are glucose deprivation, chronic anoxia/hypoxia, and acidosis (34,70,71). These conditions cause activation of a number of so-called stress proteins, which is thought to be an adaptive response evolved to protect cells against stress-induced cell death. The obtained cell survival is required for tissue preservation and organ protection in cases where normal cells are exposed to pathological conditions, but also occurs in neoplastic cells where it can lead to cancer progression, drug resistance, and protection of cancer cells from immune surveillance. Indeed, a correlation between malignancy and elevated levels of stress proteins in a variety of cancer cell lines was observed (71). Therefore, recruitment of the transcriptional control elements of stress proteins to target cancer cells is an interesting topic for investigation. Successful results have been obtained both *in vitro* and *in vivo* with the glucose-regulated protein 78 (Grp78) promoter (70,71) and with promoters containing hypoxia responsive elements (HREs), of which the vascular endothelial growth factor (VEGF) promoter is most widely investigated (47,54,60,67,68,75).

When a tumor reaches a size of ~1 mm^3, tumor cells induce proliferation of (myo)fibroblasts and endothelial tissue. This process is mediated by several growth factors [i.e., transforming growth factor α (TGF-α), basic fibroblast growth factor (bFGF), and platelet-derived growth factor (PDGF)], which stimulate proliferation of stromal cells. This process of neovascularization is named tumor angiogenesis

Table II. Tumor-Specific Expression

Key event	Application in gene therapy	Reference
Aberrant gene expression		
Oncofetal genes	AFP	(29–31,34,36,40,64)
	CEA	(29–31,34,36,40,43)
Oncogenes	ErbB2 (ErbB3, ErbB4)	(29,30,34,36,40)
	c-Myc/Myc-Max responsive elements	(34)
	MUC1/DF3	(29,36,37,40,43)
	Ras	(29)
	EGFR (ErbB1)	(29)
	FGFR	(29)
	PIK3	(29)
	Fms	(29)
	Akt2	(29,65)
Cell cycle-regulated genes	Cyclin A, Cdc25C	(36)
	Heparanase	(30)
	RhoC	(30)
	Fibronectin	(30)
	Thymosin β4	(30)
	Endoglin (CD105)	(30)
	Integrins; αV-B3 integrin	(30)
	c-Myc, Cdc2, E2F-1	(29,30,36,43)
Immortalization	Telomerase (hTERT gene)	(29,30,36,43,66)
Tumor biology		
Angiogenesis	E-selectin, Le(a) and Le(x)	(29,30,36,43,61)
	Endoglin (CD105)	(30,36,49)
	VEGF	(29,30,47,54,60,61)
	Human/murine (prepro)-endothelin-1	(30,50,59,61)
	Tie-2	(49,61,66)
Pathogenic vasculature	HRE; HIF-1 target genes (including VEGF, erythropoietin, LDH)	(30,36,47,60,67–69)
	Grp78 (BIP)	(30,34,36,70,71)
Impaired glucose metabolism	Hexokinase II	(30,36)
Diverse		
	hPRL, ALA, BLG, osteocalcin, SLPI, L-plastin, GRP, AVP, K-Ras, SI, HCG, MK, HAFR gene promoter (P1), cyclin D1, FGFR, CA125 antigen, inhibin/activin, hyaluronan receptor, trypsin inhibitor, metalloprotease pump I, UPA, LDH gene-HRE, Cox-1, Cox-2	(29,30,34,36–38,40, 57,72–74)

AFP, α-fetoprotein; ALA, human α-lactalbumin; AVP, vasopressin; BLG, bovine β-lactoglobulin; CEA, carcinoembryonic antigen; Cox, cyclooxygenase; FLT-1, fms-like tyrosine kinase-1; EGFR, endothelial growth factor receptor; FGFR, fibroblast growth factor receptor; GRP, gastrin-releasing peptide; Grp, glucose regulated protein; HAFR, human α-folate receptor; HCG, human chorionic gonadotropin; HIF, hypoxia inducible factor; hPRL, human prolactine; HRE, hypoxia responsive element; hTERT, human telomerase reverse transcriptase; ICAM-2, intracellular adhesion molecule 2; KDR, kinase-like domain receptor (human homologue of flk-1); LDH, lactate dehydrogenase; MK, midkine; MMTV, mouse mammary tumor virus; MUC, mucine; PECAM-1, platelet endothelial cell adhesion molecule-1; PIK, 3 phosphoinositide-dependent protein kinase; SI, sucrase–isomaltase; SLPI, secretory leukoprotease inhibitor; tie, tyrosine kinase with immunoglobulin and epidermal growth factor homology domains; tPA, tissue-type plasminogen activator; VCAM, vascular cell adhesion molecule; VEGF, vascular endothelial growth factor; VEGFR, VEGF receptor; vWF, von Willebrand factor; WAP, whey acidic protein.

and is of great importance for the growth and development of both primary tumors and metastases (76). Because of its major role in regulating important biological processes in tumor angiogenesis, vascular endothelium can be considered as a suitable target for cancer gene therapy for tumors of diverse origins. Tumor endothelium can be distinguished from normal vasculature through the expression of specific membrane-associated receptors, adhesion molecules, and other proteins, and by the high proportion of proliferating cells (36). Advantages of targeting tumor vasculature include its readily accessibility to systemically administered drugs due to its large surface area and proximity to the circulation, and the comparative homogeneity of this biological hallmark among different tumor types (36,61). From a delivery point of view, enhanced vascular permeability and angiogenesis and the consequential retention of macromolecules can be used for passive targeting strategies.

TRANSCRIPTIONAL CONTROL: CONTROLLING EXPRESSION IN SPACE AND TIME

Some gene therapeutic applications require tight control over the level of transgene expression to prevent expression-

induced toxicity. For example, the use of gene therapy to restore insulin production in insulin-dependent diabetes mellitus requires tight control of transgene expression in response to blood glucose levels. Temporal and spatial restriction of expression aims at adding levels of specificity and maximizing the ratio of expression levels in the induced state to background levels.

Physical Induction

One of the most striking examples of selective and inducible transcriptional regulation observed in eukaryotic cells is the induction of so-called heat shock proteins (hsps) in response to exposure to superoptimal temperatures. The most prominent heat shock protein is hsp70 and its expression is mediated mainly by interaction of heat shock factor 1 (HSF1) with heat shock elements (HSEs) present in the hsp70 promoter (77). Binding of HSF1 to HSEs leads to release of RNA polymerase that is normally stably bound proximally to the hsp70 promoter and enables initiation of transcription (77). Huang *et al.* (78) used a 400-bp hsp70 promoter driving expression of the green fluorescent protein (GFP) reporter gene or the cytokines tumor necrosis factor α (TNF-α) or IL-12 engineered into an adenoviral vector. They demonstrate that moderate hyperthermia (39–43°C) for relatively short periods of time (20–60 min) efficiently activates gene expression (GFP, TNF-α, or IL-12) driven by the hsp70 promoter both *in vitro* and *in vivo*. Promoter activity starts at 39°C with an optimum at 42–43°C and decreasing efficiency at higher temperatures due to cell death. Expression starts ~3 h after hyperthermia treatment, peaks at 18–24 h, and drops back to background levels after 72 h. Additional *in vitro* studies involving cytokine gene expression show an induction of $>6.8 \times 10^5$ and >13,600 over background expression levels for TNF-α and IL-12, respectively. Importantly, cytokine levels in nonheated control cells are below the detection limits and equal those in nontransfected cells. Observations of low leakiness and inducible expression were confirmed *in vivo*.

Advantages of heat-responsive promoters include low background expression with high inducibility, low leakiness, and convenience and safety of the stimulus. Disadvantageous is that conventional heat treatment has a rather poor resolution [localization within an order of centimeters (78)] and succeeds in heating to appropriate temperatures only the extremities, ovaries, brain, breast, prostate, head, and neck. Furthermore, growth conditions may vary within a tumor, and this could affect transgene expression. Appropriate measures must be taken to adjust heat treatment such that the desired effect is obtained in all tumor areas. Recently, advanced heating techniques based on magnetic resonance imaging (MRI)-guided focused ultrasound heating have been described that may offer a solution to these problems (77).

Responsiveness of the hsp70 promoter to various environmental and physiological signals [i.e., several transcription factors including CCAAT transcription factor (CTF), SP1, activating transcription factor/cAMP-responsive element binding protein (ATF/CREB), hypoxia, acidosis, energy depletion, oxidative stress, cytokines, certain toxic compounds, ischemia, UV radiation (77,78)] provides potential pathways to further modulate promoter activity, but at the same time complicates predictability and control. For example, even inflammation or fever may unwantedly activate heat-responsive transcription elements in the absence of heat treatment (77,78). The hsp70 promoter is repressed by wild-type tumor suppressor gene *p53*, but is overexpressed in many tumor cells with defective *p53* function (79). The combined inducibility by heat, hypoxia, and *p53* mutations potentiates efficient transcriptional targeting of tumor cells.

An important feature that requires further investigation is the risk of developing thermotolerance. Cells exposed to elevated temperatures become temporarily resistant to a subsequent heat shock. Thermotolerance is transient in nature, and its magnitude and duration depend on the severity of the initial heat dose. Manipulation of the hsp70 promoter may solve this problem; an hsp70 promoter modified to contain extra HSEs was shown to be more active than the unmodified promoter in thermotolerant cells (77).

Several promoters displaying radiation responsiveness have been proposed, including the early growth response-1 gene (*Egr-1*) (80–83), wild-type *p53*-activated fragment 1 (WAF1) (84,85), and recA (86). The first promoter tested was the *Egr-1* promoter. *Egr-1*-driven expression of TNF-α has been shown to result in increased tumor growth inhibition compared with radiation treatment alone. CArG elements [$CA(A/T)_6GG$] within this promoter were identified to account for radiation responsiveness, and synthetic radio-inducible promoters were developed exploiting multiple CArG elements as enhancers for other (stronger) promoters such as CMV. Multiplication of CArG elements improved inducibility while decreasing leakiness. Importantly, activation of *Egr-1* is predominantly *p53* independent, in contrast to other radio-inducible promoters, decreasing the likelihood of tumor-dependent interference with expression.

Expression levels obtained with the WAF1 promoter were found to equal or even exceed those observed for other radiation-inducible promoters (84,85). Worthington *et al.* (85) performed an *ex vivo* experiment in which the WAF1 promoter was used to drive expression of the human inducible NO synthetase (iNOS) gene in rat arteries. iNOS expression was observed to be induced 5-fold 8 h after exposure to 4 Gy radiation in rat tail artery segments. WAF1-driven iNOS expression resulted in full relaxation in artery segments that were preconstricted with phenylephrine 1 h after exposure to 4 Gy X-rays; this effect was reversible. Next, an experiment was performed where WAF1/iNOS was injected directly into two different tumor types (RIF-1 and HT29) in mice (84). The WAF1 promoter was induced by an initial X-ray dose of 4 Gy followed 8 h later by treatment doses of 10 or 20 Gy. Examination of the mice tissue revealed that 48 h after intratumoral injection, vector sequences were detected in all tissues tested, indicating that substantial leaking of the system from tumor tissue to nontarget tissue occurs. Importantly, no significant increase in iNOS levels was observed in any of the tissues other than the tumor and surrounding dermal tissue that had been exposed to the irradiation. Induction of iNOS levels was 2.1- and 3.3-fold in RIF-1 and HT29, respectively, when compared to control nontransfected tumors. When compared to transfected but nonirradiated tumor cells, induction was 3- and 1.6-fold,

respectively. A problem involved when using the WAF1 promoter is that expression does not occur in a straightforward, dose-dependent fashion, and the optimal radiation dose seems to be dependent on cell type and transgene used. Upregulation of WAF1 in hypoxic conditions may be advantageous for radiotreatment of tumors, as hypoxia often hampers the effectiveness of regular radiotherapy.

The major advantage of radiation-induced transcription is the high level of precision of target area selection. Radiation therapy can reach precise localization within an order of millimeters. Limitations to radiation-inducible promoters include leakiness and the carcinogenic nature of the stimulus, which restricts applications to cancer treatment. Additional research is required to establish optimal radiation dosing schemes and relevance thereof in clinical settings.

Chemical Induction

Pharmacological regulation of transcription is desirable when the aim is to drive long-term expression of transgenes as it allows i) titration of protein into the therapeutic window, ii) dose adjustment, iii) reversibility/termination of therapy, and iv) fluctuating daily dosing regimens relevant for many proteins (Fig. 3A) (87,89).

The key element of chemically inducible systems is a modulator-controlled transactivator consisting of 1) a DNA-binding domain (DBD) that can recognize and bind to specific *cis* elements within the promoter, 2) a transcription activation domain (AD), and 3) a binding site for the modulator (chemical/drug) (90).

Recruitment of the AD to the promoter may be based either on allosteric interaction or on dimerization (Fig. 3B). In case of allosteric interaction, a specific drug-controlled DNA-binding domain is fused to a heterologous transcriptional AD. This fusion-control element is placed upstream and interacts with the promoter only in the presence of the drug of concern, allowing transcription to be initiated.

The first and most frequently described allosteric interaction system is a tetracycline (Tet) (or analogue)-dependent gene expression system, which comes in two variants: a system in which transcription is either suppressed (Tet-Off) or inititated (Tet-On) upon addition of the tetracycline drug. The natural Tet-controlled DBD of the *Escherichia coli* Tet repressor (TetR) is fused to a heterologous AD [e.g., virion protein 16 (VP16)] to produce the tetracycline transactivator (tTA). This tTA activates transcription upon binding to Tet operator (TetO) sequences upstream of target genes. However, when tetracycline or an analogue (e.g., doxycycline) is present, DNA binding and hence, transcription, is abolished due to binding of the drug

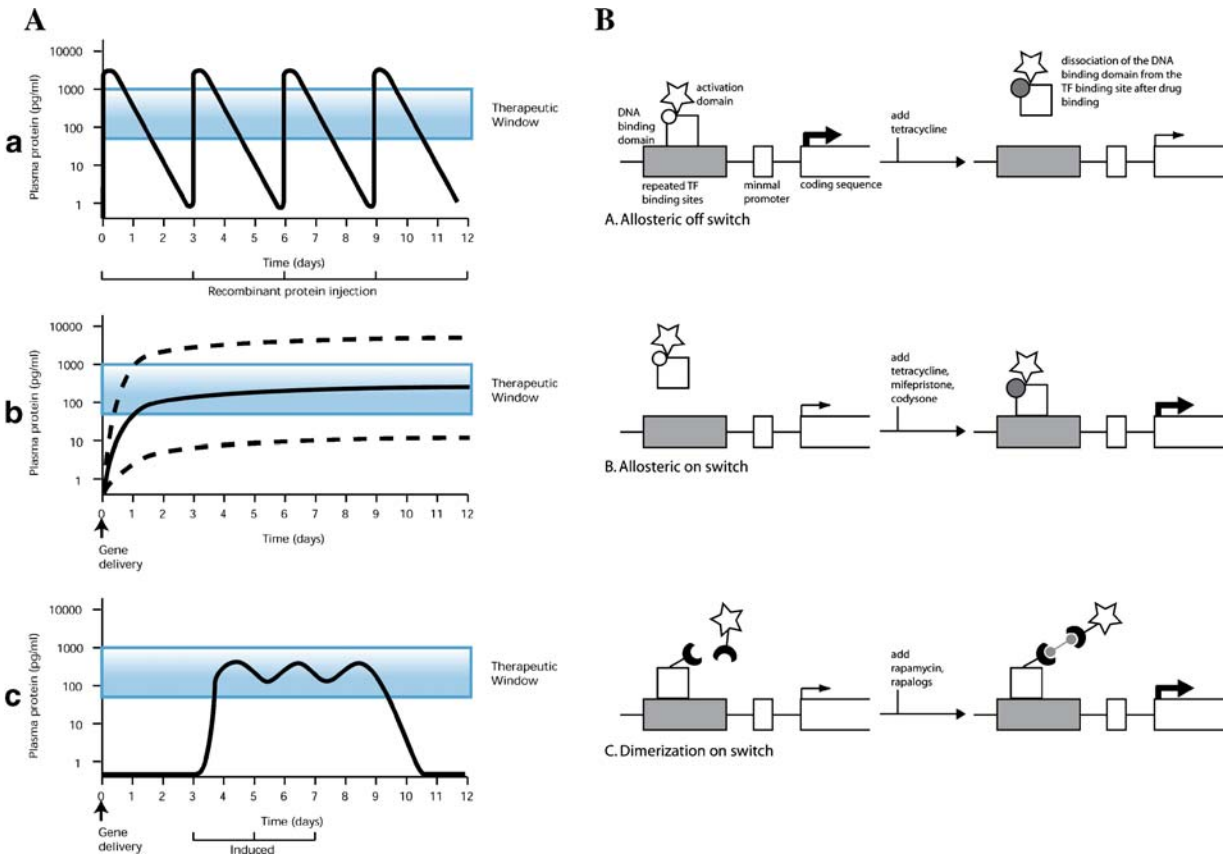

Fig. 3. Chemical induction of gene expression. (A) Hypothetical protein plasma levels after (a) intravenous injection of recombinant protein, (b) gene therapy using noninducible expression systems, and (c) gene therapy using inducible expression systems. Adapted from Clackson (87) with permission. (B) Schematic illustration of the different types of chemically inducible gene expression systems. Adapted from Harrington *et al.* (88) with permission.

to TetR. Initial reports describe low-level background expression in the presence of Tet, with an up to 100,000-fold increase upon Tet withdrawal depending on the cell clone used (88,89,91). However, subsequent studies have never succeeded in achieving such extreme induction ratios, largely due to high background expression levels (88). Drawbacks of the Tet-Off approach include the necessity of long-term administration of tetracycline (analogues) and the slow onset of induction due to dependence on tetracycline (analogue) clearance (92). Wilson *et al.* (93) have proposed utilization of the Tet-Off-based gene expression system for treatment of diabetes by establishing constant background insulin replacement with infrequent adjustments to this basal expression level upon tetracycline administration to allow anticipation to illness, diet, exercise, etc. They observed repression of proinsulin secretion in a dose-dependent and reversible fashion upon addition of tetracycline following transfection of murine and human myoblasts with a tetracycline-repressible transactivator and an insulin plasmid containing a tetracycline-responsive element upstream of a minimal CMV promoter. The functionality of this approach was also confirmed *in vivo* following intramuscular plasmid injection and oral tetracycline administration in rats. Baron *et al.* describe optimization of Tet-regulated systems that show reduced toxicity by modification and screening of tetracycline-controlled tTAs to eliminate potential interaction sites for various cellular transcription factors and sites that may elicit a cellular immune response (94).

Two strategies to change the Tet-system from an off to an on system have been investigated. The first involves fusing TetR to a strong transcriptional repression domain, enabling de-repression of transcription upon drug binding. This was reported to result in a Tet-mediated reporter induction of up to 50-fold. The second approach aims at producing a true on-switch mechanism and employs a mutant TetR that only binds DNA in the *presence* of Tet (87,89). Unfortunately, reversal of response type is accompanied by decrease in binding affinity of TetR for its inductor and increased background expression levels. Several mutant forms have been screened to identify those with minimal background activity combined with acceptable binding affinity (92,95). For mutant TetR systems, induction ratios of up to 10,000-fold have been reported (89). Another approach to reduce background expression has been described in which the Tet-On system is used in combination with a tetracycline-controlled transcriptional silencer (tTS). In the absence of tetracycline, the tTS interacts with TetO sequences to suppress transcription. However, when Tet is present in sufficient concentrations, conformational changes in tTS lead to dissociation of tTS from TetO, allowing the reverse transcriptional activator to bind and activate transcription (92). This approach requires efficient delivery of three separate vectors, and unless these can be integrated into a single plasmid construct, it is not very attractive for use in gene therapy.

In antiprogestin-regulated gene expression systems, truncated forms of the human progesterone receptor ligand-binding domain (PR-LBD) are linked to specific DNA-binding and transcription activator domains to form a chimeric protein that functions as an antiprogestin-responsive transcription factor. The PR-LBD is modified such that it can no longer bind progesterone or any other known endogenous steroid but can be selectively activated by progesterone antagonists such as mifepristone (Mfp). Nonviral gene delivery of an Mfp-controlled system in mice resulted in an average multitude of transgene induction of 14- to 19-fold in response to Mfp. This effect could be reached repeatedly over a period of approximately 3 weeks (96). Other studies report up to 200-fold induction ratios (depending on the minimal promoter used) both in cell culture and in animals (89,97).

To avoid problems of interference from endogenous hormones involved when employing human/mammalian steroid hormones and their receptors, the use of the prokaryotic steroid ecdysone (Ec) (or its synthetic analogue, muristerone A) and its nuclear receptor was proposed. Fusing the *Drosophila melanogaster* Ec receptor Ec-binding domain to heterologous DBDs and ADs was shown to allow Ec-dependent activation of transcription. These systems did not respond to a series of endogenous human steroids. *In vitro*, low background activity combined with induction ratios of up to 10,000-fold was observed (89).

Many tumors develop anticancer drug resistance through overexpression of the multidrug resistance (Mdr)-1 gene. Because the Mdr1 promoter is inducible by cytostatic agents such as doxorubicin, vincristine, and taxol, it can be employed for pharmacological upregulation of transgene expression to potentiate chemotherapeutic effects (98–100).

Only recently, a novel form of chemical induction was described in which a gas is used as the inductor for expression of transgenes (101,102). Weber *et al.* (101) developed a system derived from the fungus *Aspergillus nidulans* based on the enzymatic machinery that regulates conversion of ethanol to acetyl coenzyme A in response to acetaldehyde. They developed an inducible promoter consisting of a minimal CMV promoter and five operators, each containing binding sites for the transcriptional activator. Transgene expression obtained with this construct was regulated tightly by the gaseous acetaldehyde. *In vitro* expression levels correlated linearly with gas concentration and reached an optimum at gas levels that were below the no-observed-effect level (NOEL) of 152 ppm as declared by the World Health Organization (Health and Safety Guide 90, 1995). However, *in vivo* no linear relation between acetaldehyde levels and gene expression was observed. Moreover, because gas concentrations of >1000 ppm are likely to be required for optimal induction, strategies augmenting transactivation need to be developed to make the system function at gas levels below the NOEL, which will be essential for clinical application. Limitations to this system include the unpredictability of the influence of endogenous acetaldehyde and possible immunostimulation in response to the fungal elements (90).

Chemical dimerization-controlled transcription requires coexpression of two fusion proteins that each contain a drug-binding domain. This domain is fused to a DNA-binding domain in one protein and to a transcriptional AD in the other. Additionally, the target gene construct, consisting of the therapeutic gene driven by a promoter that contains binding sites for the DBD, is introduced. Administration of a drug able to cross-link both fusion proteins through their drug-binding domains results in the formation of a functional transcription factor. This results in recruitment of the AD to the promoter and transcription initiation, respectively. The

bivalent drug may be homodimeric (binding equal domains) or heterodimeric (binding different domains). Heterodimeric drugs are usually preferred, as in case of homodimerization, nonproductive DBD–DBD or AD–AD homodimers can be formed in addition to the functional DBD–AD heterodimers. A well-described example of a heterodimeric drug is rapamycin, an orally bioavailable drug that mediates the formation of dimers between the human proteins FK506 binding protein (FKBP) and FKBP12 rapamycin-associated protein (FRAP). Low background expression levels and high inducibility (over 10,000-fold *ex vivo*) were observed when using the rapamycin-regulated strategy (89). Expression levels comparable to those obtained with the CMV promoter have been reported (103). Efforts have been made to modify the system to function with nonimmunosuppressive analogues of rapamycin ("rapalogs") (103).

The highly modular character of dimerizer-regulated systems facilitates incorporation of a wide range of DBDs, ADs, and drug-binding domains, and development of alternative dimerizer drugs allowing practically unlimited opportunities for optimization. For a review on progress made with dimerizer-based strategies, the reader is referred to Pollock and Clackson (103).

Despite the promise of pharmacological regulation of transgene expression, there are some less elegant sides of chemically inducible systems. Although the transcriptional control elements can be modified to optimize and target expression, no control over the fate of the inducing drug exists, and many nontarget cells will be unnecessarily exposed to this drug. Another disadvantage is the requirement of multiple constructs, meaning that effect is only obtained in those cells that receive all of the essential elements. To increase efficiency and to minimize burdening of the delivery system, efforts must be made to combine the different elements into a single construct.

SUSTAINED EXPRESSION

A serious issue currently limiting the widespread use of nonviral gene therapy vectors is their transient nature of transgene expression. Expression is decreased in time due to several mechanisms including potential loss by recombination or destruction by nucleases (31,104) and partitioning to nonnuclear compartments (104). Also, in dividing cells a logarithmic decrease in the percentage of transfected cells during replication of the target population occurs because the plasmids do not replicate, whereas the cells do (31,104). Finally, recognition and subsequent silencing of foreign DNA *in vivo* impedes persistent transgene expression. Diverse strategies have been developed to overcome one or more of these problems to increase duration of expression.

Integrating Expression Vectors

One solution to allow for replication of the delivered gene is integration of the gene into the host genome. Enzymes that are capable of inserting foreign DNA into the host genome include viral integrases, (site-specific) recombinases, and transposases (105). Among these systems, viral systems, especially retroviruses, integrate at relatively high frequencies. However, loading capacity is limited, and lack of site-specificity increases the risk of insertional mutagenenis (105). It was recently discovered that retroviruses and recombinant adeno-associated virus (rAAV) vectors actually integrate preferentially into transcriptionally active regions rather than randomly (106). For example, murine leukemia virus (MLV), human immunodeficiency virus (HIV), and adeno-associated virus (AAV) have been observed to preferably integrate into genes rather than nongenomic regions, and AAV has been associated with deletions and rearrangements within host DNA (106–108). In this light, recombinases and transposases may offer a safer alternative. Site-specific recombinases are enzymes that catalyze DNA strand exchange between stretches of DNA that are homologous only to a limited degree (105). Recombinase proteins bind covalently to recognition sites within the DNA, cleave its backbone, exchange the double-stranded DNA segments, and finally re-ligate the DNA. Some recombinases function independently, others require the help of additional proteins. Two of the most widely described site-specific recombinases are cyclization recombination recombinase (Cre) and φC31 integrase, enzymes derived from *E. coli* bacteriophage P1 and *Streptomyces* phage ϕC31, respectively. Integration is possible through interaction with so-called "pseudosites" that resemble the original recombination sites that are absent in the human genome (108). Cre-mediated integration is reversible, hence, unstable and inefficient, because excision is favored over integration. The explanation for this is that *loxP* (Locus of Crossing over of P1 phage) sites remain unaltered and are therefore still substrates for excision/integration reactions, until finally all *loxP* sites are removed by excision. An advantage of Cre is its remarkable site-specificity. However, this at the same time represents an obstacle, as extensive manipulation to Cre is required to obtain a broader specificity to make it react with pseudotarget sites within the host-genome. Because φC31 displays a greater affinity for and recognizes more (at least 11) pseudotarget sites within the human genome, it does allow stable integration of DNA herein with greater efficiency (24–56%) (108). However, its lower degree of specificity requires additional research on potential target sites to assess risks of insertional mutagenesis (105). Individual recombinases can be further optimized through a process of directed evolution: random mutations are introduced and screened for improved efficiency and specificity (108). Furthermore, Buchholz *et al.* (109) describe customization of the Cre-like recombinase flippase (FLP) to adjust its temperature optimum from 30 to 37°C to make it useful for application in human gene therapy applications.

Transposases are enzymes that can transfer discrete segments of DNA (transposons) from one molecule to another through a cut-and-paste process (107). Transposases are abundant in prokaryotes and insects, but have no equivalent in vertebrates. Transposons originating from nonvertebrates are functional in humans, albeit at low efficiency. However, with the reconstruction of the so-called Sleeping Beauty (SB) transposon, a vertebrate-derived DNA transposon was created that displays much higher activity than any other transposon in a wide range of vertebrates (107). SB transposon consists of the transposable gene flanked by two terminal inverted repeats (IRs) that contain binding sites for

the transposase. Binding of transposase to these sites is followed by formation of a synaptic complex, excision of the DNA from the donor site and reintegration at a target site, respectively (Fig. 4A). Formation of the synaptic complex requires interaction with the host protein high-mobility group box 1 (HMGB1), a nonhistone protein associated with eukaryotic chromatin that serves as a cofactor for SB transposase (107). SB transposase is highly specific; no cross-mobilization is observed among closely related transposons (107). In its original form, maximum delivery capacity of SB is limited to 10 kb; with each kilobase increase the efficiency decreases exponentially with ~30%. However, flanking of transgenes with two complete SB elements has been shown to increase these size limits considerably (107). Analysis of insertion sites in human cells revealed that SB, unlike diverse viruses, shows no preference for insertion into genic regions over nongenic regions, suggesting that SB transposition may offer a safer alternative for stable expression compared to viral integration. From several *in vivo* studies, it seemed that long-term expression could be obtained (>5 months), whereas no cytotoxic effects, immune responses, or liver tumors were developed (110–112). These results illustrate the potential of SB for use in gene therapy settings. However, the remaining risk of insertional mutagenesis should not be ignored. Future directions for integrating strategies will be based on strategies allowing for targeted integration (site-directed genome modification). This can be obtained by alterations within the DNA sequence or by use of accessory proteins (105,113).

Episomally Replicating Vectors

An approach that does not require integration of the gene into the genome and therefore avoids risks of insertional mutagenesis is the use of autonomously replicating plasmids or episomes. In episomally replicating plasmids, sequences from (generally) viral DNA are incorporated that enable the plasmid to replicate extrachromosomally. There are several advantages over integrating systems: 1) the transgene cannot be interrupted or subjected to regulatory constraints that often occur from integration into cellular DNA (33,114); 2) higher transfection efficiency can be obtained than with chromosome-integrating plasmids (33,114); 3) episomes display a low mutation rate and tend not to rearrange (33); and 4) episomally replicating systems have the ability to transfer large amounts of DNA. For instance, systems based on the EBV viral replication and retention components can carry up to 185 kb of viral DNA (33,114) and allow replication in both eukaryotic and prokaryotic cells, which enables easy shuttling among these host cells (33,114).

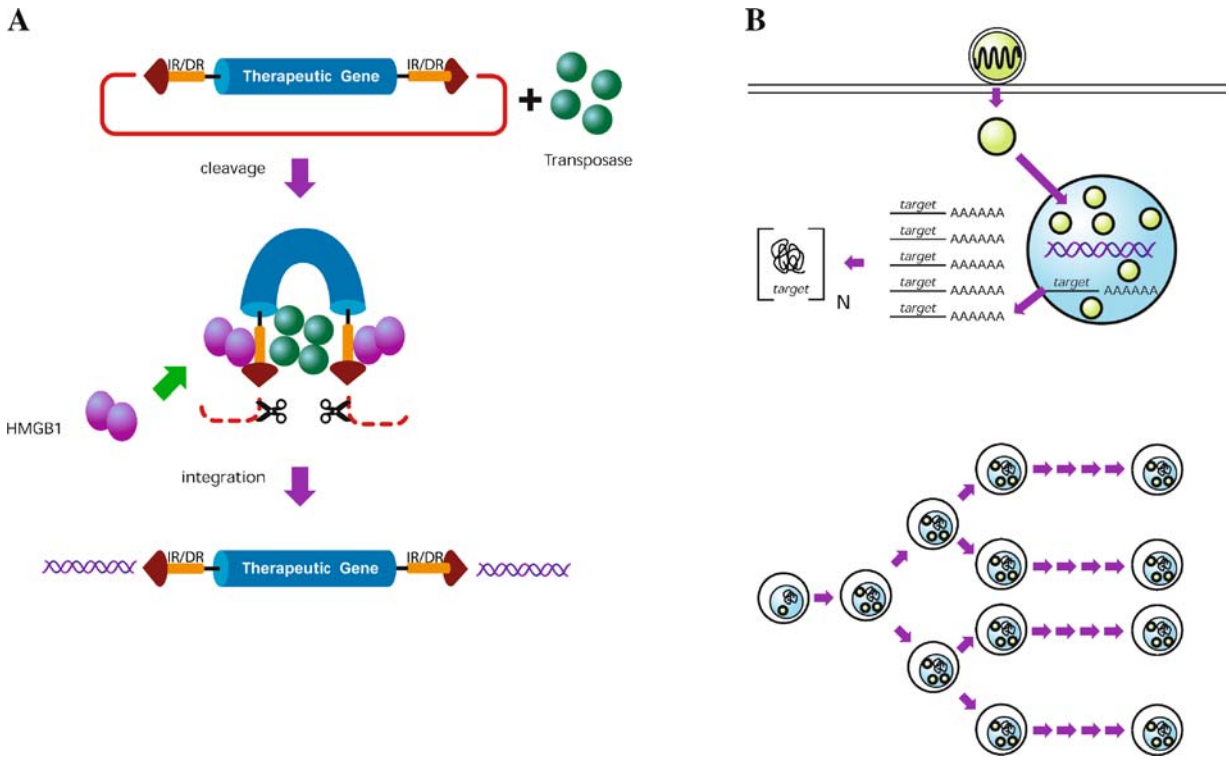

Fig. 4. Strategies for sustained gene expression. (A) The Sleeping Beauty transposable element and its transposition. A plasmid in which the transposase gene is replaced by a therapeutic gene (transposon) and the transposase are supplied in *trans*. The transposase binds to its binding sites within the inverted/direct (IR/DR) repeats flanking the therapeutic gene and, together with host factors such as HMGB1, a synaptic complex is formed, in which the ends of the transposon are paired. After excision of the transposon from the plasmid, it can be integrated into a new location. Adapted from Izsvak and Ivics (107). (B) Episomally replicating vectors. Replicating episomal plasmids yield high levels of target gene expression through several mechanisms: vector replication leads to accumulation of multiple copies of the episomal plasmids; increased copy numbers of plasmids lead to increased levels of target gene mRNA, and consequently increased levels of target gene protein; efficient vertical transfer of the episomes during cell division results in maintenance of high-level gene expression. Adapted from Cooper (31).

Constructs from several viruses including EBV, BK virus (BKV), SV40, and bovine papilloma virus 1 (BPV-1) have been investigated as episomal expression vector candidates (30,106). In general, the main problem with episomal expression vectors is the requirement of *trans*-acting factors, which are often associated with risk of transformation. This is especially the case for vectors based on polyomaviruses (i.e., BKV, SV40) where the *trans*-acting factor is a large T antigen (Tag). Large T antigens were shown to have numerous unacceptable properties including the ability to bind the tumor suppressor gene *p53*, to induce chromosomal aberrations, and to influence cellular gene expression by interfering with cellular transcription factors (114). In this respect, strategies based on EBV are considered relatively harmless. Additionally, EBV displays a low mutation frequency and is capable of carrying large amounts of DNA (114). Most progress so far has indeed been made with EBV and therefore this system will be discussed in more detail.

EBV contains two elements that enable stable episomal maintenance of the viral DNA in the host cell: oriP, required *in cis*, and the EBV nuclear antigen 1 (*EBNA1*) early gene, required *in trans* (33,114–116). OriP contains two regions, being the family of repeats (FR) and the dyad symmetry (DS) element, containing 20 and 4 binding sites for EBNA1, respectively (114). EBNA1 has been reported to facilitate nuclear localization of the plasmid (115). Expression of the *EBNA1* gene is followed by binding of EBNA1 as a homodimer to oriP and recognition of the bound oriP site as a functional DNA origin by human cells (31,114). The EBNA1 dimer/oriP complex then serves several functions: replication, maintenance, and transcription.

The actual role of EBNA1 in *replication* is still under discussion. Binding of EBNA1 dimers to the DS within oriP is thought to result in a structural distortion, which is important for the initiation of DNA synthesis. More importantly, the FR contains a replication fork barrier that forces unidirectional replication and thereby reduces risks of mutation and rearrangements (114). Conversely, some groups have reported an EBNA1-independent synthesis of oriP-containing plasmids (114). Despite the ambiguity of the exact mechanism of replication, EBV vectors replicate once per cell cycle in synchrony with the host chromosomes (114).

In addition to inducing plasmid replication, EBNA1 facilitates *binding of the plasmid to the nuclear matrix*: binding of EBNA1 both to the FR on the plasmid and to chromosomal elements mediates physical association of the plasmid with the host chromosomes to retain the plasmid (114,115). Also, a cellular protein, EBNA1-binding protein 2 (EBP2), has been identified that is thought to play a role in segregation of the episomes during cell division (114).

EBNA1 is thought to mediate *transcriptional upregulation* upon binding to the FR through a mechanism that has not yet been revealed (114,115). This effect seems dependent on the cell type used and on the promoter within the plasmid (114).

Stable replication of EBV-derived episomal vectors was shown in a variety of mammalian cells, including human epithelial, fibroblast, and lymphoma cells, as well as monkey and dog cell lines (114). Cui *et al.* (115) report highly efficient gene transfer of EBV-derived plasmid vectors containing the *EBNA1* gene and the oriP sequence *in vivo*. Both EBV-derived and conventional plasmid vectors encoding luciferase or β-galactosidase (β-gal) as markers were constructed and injected into the tail vein of mice by using the hydrodynamic pressure method of transfection. Subsequently, gene expression was measured in the liver over time. For the conventional luciferase-expression vector, a maximum level of protein expression was measured at 8 h posttransfection, decreasing 4.8-fold during the following 16 h. In comparison, transfection with the EBV-based luciferase-expression vector resulted in an approximately 1.7- and 7.2-fold higher expression at 8 and 24 h after transfection, respectively. Similar results were observed for the β-gal expression vectors.

High transfection efficiency, long-term expression, capacity to carry large amounts of DNA, and low mutation and rearrangement rates all contribute to the potential success of EBV-based vectors in nonviral gene therapy. However, although in most cases stable episomal maintenance of EBV-derived vectors is obtained, in some cases integration in the host chromosome and rearrangements within the vector have been reported (114). This effect seems dependent on the cell line used and will require thorough research before EBV-based vectors can be used safely in humans. Another important issue is the observation that plasmid copy numbers vary considerably among different cell lines, ranging from 5 to 100 (114). This merits attention because the plasmid copy number within a cell is determinative for both therapeutic and possibly toxic effects.

Finally, risks of oncogenicity should be examined more carefully. For EBV, the viral *EBNA2* gene is considered to be mainly responsible for oncogenicity, and therefore EBNA1 was thought to be innocent. Nonetheless, results of some *in vitro* and *in vivo* experiments put this assumed innocence in a new light. For instance, EBNA1 was shown to be able to bind to RNA *in vitro*, and it may therefore influence expression at the posttranscriptional level (114). Also, it was suggested that EBNA1 might interact with the c-*Myc* gene, which may result in deregulation of this protooncogene. *In vivo* results are contradictory. EBNA1 expression seemed to predispose B cells to lymphoma in transgenic mice in a similar fashion as transgenic c-*Myc* expression does, implying oncogenicity of EBNA1 in mice. On the other hand, in other *in vivo* experiments where mice were transfected with EBV-based vectors through lipofection, no pathological changes were observed (114). Obviously, additional experiments should be performed to ascertain the safety of EBNA1 use.

The episomal vectors described above all require at least one viral gene product, such as EBNA1 or Tag in case of EBV or SV40, respectively (117). The potential immunostimulatory and transforming properties of virally encoded proteins impede application of these systems for human gene therapy. Alternatively, mammalian scaffold/matrix attachment regions (S/MARs) have been identified that can be incorporated into circular nonviral vectors to replace the transacting viral gene products. S/MARs are *cis*-acting elements consisting of 100- to 1000-bp AT-rich regions lacking a defined consensus sequence (117). The S/MAR element enables interaction of the plasmid with components of the nuclear matrix, allowing for co-segregation of the

plasmid with chromosomes during mitosis. Additionally, it is thought that this interaction facilitates plasmid replication by bringing the plasmid into contact with the host replication machinery (108).

Piechaczek *et al.* (119) developed an episomal replicating expression vector in which the SV40 ori sequence was used together with the S/MAR from human β-interferon gene cluster (replacing the viral Tag protein). This construct was demonstrated to replicate episomally in CHO cells and to provide stable expression over more than 100 cell divisions in the absence of selective pressure. Ehrhardt *et al.* (117) developed a vector in which the CMV promoter was replaced by a cellular promoter to eliminate potential silencing effects induced by nonmammalian sequences. They incorporated an S/MAR sequence derived from the chicken lysozyme locus (ChMAR) as *cis*-acting elements within a nonviral plasmid vector encoding the human coagulation factor IX (hFIX). *In vivo*, 5-fold induction of hFIX expression levels was observed for the S/MAR-containing plasmid when compared to plasmids lacking the S/MAR for up to 1 year after transfection.

In addition, other vectors are also being explored that are worth mentioning: yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), human artificial chromosomes (HACs), or other mammalian artificial episomal chromosomes (MAECs) and chimeric systems thereof. Constructs have been designed in which the oriP/EBNA1 module is inserted into YACs and BACs to function as an alternative to the centromere. The resulting molecules were shown to form large circular episomes capable of stable and persistent expression (119). As these artificial chromosomes are designed to carry large DNA inserts, these systems are not very suitable for use as vectors in gene therapy applications where, in most cases, only a few genes need to be expressed. Delivery of chromosomal vectors is relatively complicated, particularly because their large size (1–2 μm), in comparison to plasmids (sizes generally within the nanometer range in condensed form), hampers cellular uptake via the endocytic route (120). Additionally, the structural stability of chromosomal vectors containing both DNA and proteins is inferior to that of pure plasmid DNA (120).

Altogether, extrachromosomally replicating plasmid vectors have great potential for application in gene therapy by increasing the efficiency of transfection and longevity of gene expression. Especially in the case of cancer treatment where dividing tumor cells are involved, maintenance and vertical transfer of the therapeutic plasmid will be crucial for the duration of the effect and therefore for success of treatment. Retention rates of 92–98% have also been reported to be typical for EBV-derived vectors in the absence of selection pressure (114). Although this does mean that retention is still imperfect and loss over time will remain an issue, these replicating plasmid vectors can be considered far superior compared to conventional nonreplicating plasmids.

Preventing Gene Silencing

Despite the use of integrating or episomally replicating vectors, sustained expression of transgenes may be hampered by gene silencing (121). Although the exact mechanism of silencing has not yet been elucidated, several possibilities have been coined and are worth mentioning. The most striking example of gene silencing is the increasing number of experiments showing that the CMV promoter, regarded as one of the strongest promoters, is frequently shut down *in vivo* (122–124). This is thought to be caused by its DNA characteristics being aberrant from those of eukaryotic DNA. Mammalian genomes differ from bacterial and viral genomes with regard to abundance and methylation state of CpG motifs. The mammalian genome contains a CG frequency of approximately 1:64, which is much lower than the expected frequency of 1:16, as observed in bacteria (125). Whereas within bacterial and viral DNA CpG motifs remain in an unmethylated state, approximately 70% of CpGs are methylated in the mammalian genome, leaving unmethylated only those CpG islands associated with genes in the germline or located within the promoter regions of transcriptionally active genes (123).

The high frequency of unmethylated CpG motifs present in bacterial and viral DNA has several consequences. In eukaryotic cells, these CpG motifs are *de novo* methylated, a process that is possibly triggered by transcription itself (126). The process of methylation serves a physiological function in vertebrates in normal development: X chromosome inactivation, imprinting and silencing parasitic DNA transcription (123). Methylation may interfere with binding of transcription factors and therefore block initiation of transcription (122,124,127). Additionally, methylated CpG base pairs can be recognized and bound by certain cellular proteins, including methyl-CpG-binding proteins (MeCP) MeCP1 and MeCP2. These proteins may compete with transcription factors and/or recruit histone deacetylase activity, which could account for the reorganization of DNA into tightly packed chromatin structures incompatible with transcription (122,124). The efficiency with which methylation suppresses transcription may be dependent on the position of the methyl groups within the promoter region, the density of the methyl groups, and the strength of the promoter (124). Methylation of CpG islands located downstream of an active promoter does not block elongation (122). On the other hand, when methylation of CpG motifs induces a condensation process, this might spread out to the vicinity of the transgene and silence it.

Another implication is the profound stimulation of the immune system associated with CpG motifs. The sequence requirement for eliciting an immune response is a central 5′-CG-3′ motif; the most active sequence in humans is GTCGTT, and this stimulatory CpG motif (CpG-S) acts as a "danger signal" that is recognized by "pattern-recognition receptors" present on immune cells. The high content of CpG-S in the unmethylated state is typical of bacterial and of some viral genomes (including CMV). Unmethylated CpG-S triggers cells of both the innate and the adaptive immune system, which has consequences regarding expression and toxicity. Gene expression can be decreased through several mechanisms, including cytokine-mediated promoter shutdown and elimination of the expressing cell through apoptosis, innate or adaptive immune responses (125).

When constructing a plasmid vector, several strategies can be used to reduce the immunostimulatory effects of CpG-S motifs: 1) methylation, 2) addition of neutralizing CpG motifs (CpG-N), and 3) elimination of CpG-S (125,128). However, as described above, methylation of essential CpG

motifs within regulatory elements can also result in a drastic decrease in gene expression. CpG sequences preceded by a cytosine (C) and/or followed by a guanine (G) are found to be able to neutralize the immune activating properties of CpG-S motifs, hence the name neutralizing CpG (129,130). The net stimulatory potency of a particular sequence seems to depend on the overall ratio of CpG-S to CpG-N motifs. From experiments in mice, it was concluded that immunostimulation can be inhibited by CpG-N sequences placed in *cis* and not too far from the CpG-S sequence and through a non-sequence-specific effect of *trans* CpG-N sequences. Reducing the CpG-S content of a plasmid does not only result in decrease in cytokine production, but very interestingly has also been reported to increase and prolong expression *in vitro* and *in vivo* in mice (131,132). Taken together, the CpG content and methylation state should be taken into consideration when choosing a regulatory element, with preference for the lowest content of CpG-S. It should be examined whether additional danger signals exist comparable to the CpG-S motifs.

Plasmid constructs used for gene therapy often combine bacterial and eukaryotic DNA elements. Within the nucleus of eukaryotic cells, the transgene is expressed, whereas the bacterial backbone remains inactive. It has been hypothesized that transcriptional activity is related to chromatin structure. In the process of transcription, certain components acetylate histones 3 and 4, resulting in loosening of their binding to DNA and formation of euchromatin. As bacterial DNA is not involved in transcription complexes, histones 3 and 4 may remain unacetylated, leaving the DNA in the more condensed heterochromatin state. In the absence of so-called insulators between euchromatin and heterochromatin regions, heterochromatin might spread into euchromatin in its vicinity, resulting in transgene silencing (121). The role of bacterial DNA in transgene silencing has been examined by Chen *et al.* (121). They conclude that for silencing to occur, covalent connection of bacterial DNA to the transgene is essential. The silencing effect was observed in both circular and linear DNA and was found to be independent of specific bacterial DNA sequences and reporter and promoter/enhancer. Exclusion of bacterial DNA resulted in 2- to 3-log higher expression levels in murine livers.

FURTHER CUSTOMIZATIONS

Several strategies have been developed to further customize transgene expression, for example, to increase strength, specificity, and efficiency, or to decrease size or leakiness.

To increase promoter efficiency, all regions within a promoter that do not contribute to its transcriptional strength or specificity should be identified and eliminated. On the other hand, functional elements can be multimerized to an optimum, as described for CArG elements in the *Egr-1* promoter. However, this approach may not be applicable to all promoters and requires time-consuming empirical optimization of each individual promoter (36).

Another, less extensively described, strategy involves increasing promoter strength by activating point mutations (36). Substituting one G-to-A at nucleotide −119 was shown to significantly increase activity of the human α-fetoprotein (AFP) promoter (133). However, such possibilities for modifications have only been found coincidentally in few cases.

Promoter activity can also be altered by modifications in DNA sequences (spacers) that separate the individual consensus sequences of promoters. Within these areas, it is not so much the sequence *per se* that is of importance, but the DNA structure resulting from this sequence (134). Jensen and Hammer (135) constructed a library of 38 mutant promoters and measured a range of relative expression strengths varying from 0.3 to 2000, covered in small increments. Mutants were designed to contain the known consensus sequences as in the wild-type form, whereas the sequences of the separating spacers were randomized. However, retrospectively it seemed that all promoters with activities <5 had changes either in the consensus sequence or in the length of the spacer between −35 and −10 sequences. Nonetheless, promoters devoid of these features still had activities varying from 5 to 2050, indicating that a 400-fold variation in promoter activity can be obtained by spacer randomization. Furthermore, the slight increment at which promoter activity is increased allows fine-tuning to an almost infinitive degree. It should be realized that ranking is dependent on the gene studied and the type of cell culture used. For an optimized method for generating promoter libraries, the reader is referred to Solem and Jensen (134).

Yet another strategy is based on constructing "chimeric promoters" that combine the transcription regulatory elements of different promoters eliciting specificity for the same tissue or eliciting different specificity patterns (e.g., tissue specificity and tumor specificity) to further restrict expression to certain target cells. By screening random combinations, optimal constructs can be identified. For example, a range of muscle specific promoters was constructed by assembling 5–20 DNA elements involved in muscle-specific transcriptional activation in a random order and linking them to a minimal chicken α-actin promoter. Remarkably, when tested in differentiating muscle cells in culture, one of the combinations was shown to be 6-fold more active than the strong CMV IE promoter/enhancer (36).

Positive Feedback Loops

Both strength and specificity of expression systems can be increased significantly by incorporation of a positive feedback loop. Generally, a promoter eliciting certain specificity is used to drive expression of both the desired effector gene and a strong artificial transcriptional activator. This transcriptional activator then upregulates transcription through interaction with appropriate binding sites within the promoter. In most cases, the transcriptional activator is a chimeric protein consisting of a DNA-binding domain fused to a transcriptional AD and is therefore referred to as recombinant transcriptional activator (RTA) (36).

Nettelbeck *et al.* demonstrated enhancement of the very weak but highly specific von Willebrand factor (vWF) and sucrase–isomaltase (SI) promoter in a positive feedback-loop approach (57). The vWF promoter exhibits a particularly high degree of specificity for endothelial cells when compared to other endothelial-specific promoters (e.g., PECAM-

1/CD31, flk-1/KDR); the SI promoter is highly specific for intestinal cells and gastrointestinal tumors. Here, the RTA is a VP16-LexA fusion protein consisting of the DNA-binding domain of LexA and the transcriptional AD VP16 of the herpes simplex virus. VP16-LexA exerts its stimulatory effect through LexA binding sites introduced into the promoter (Fig. 5A). Two approaches were examined, both using a cell-type-specific promoter (either the vWF promoter or the SI promoter) to drive transcription of the reporter/effector gene, but one employing a second cell-type-specific promoter to control the RTA and the other employing an IRES for this purpose. The construct using the two cell-type-specific promoters proved the most successful and exhibited a 20- to 169-fold enhancement while retaining a 30- to >1000-fold cell type specificity when compared to a normal vWF promoter. Similarly, for the SI promoter, a 14- to 37-fold enhancement was observed, whereas specificity was retained.

Although particularly useful for weak promoters, transcriptional feedback amplification also proves useful for strong promoters. Emiliusen *et al.* first screened different

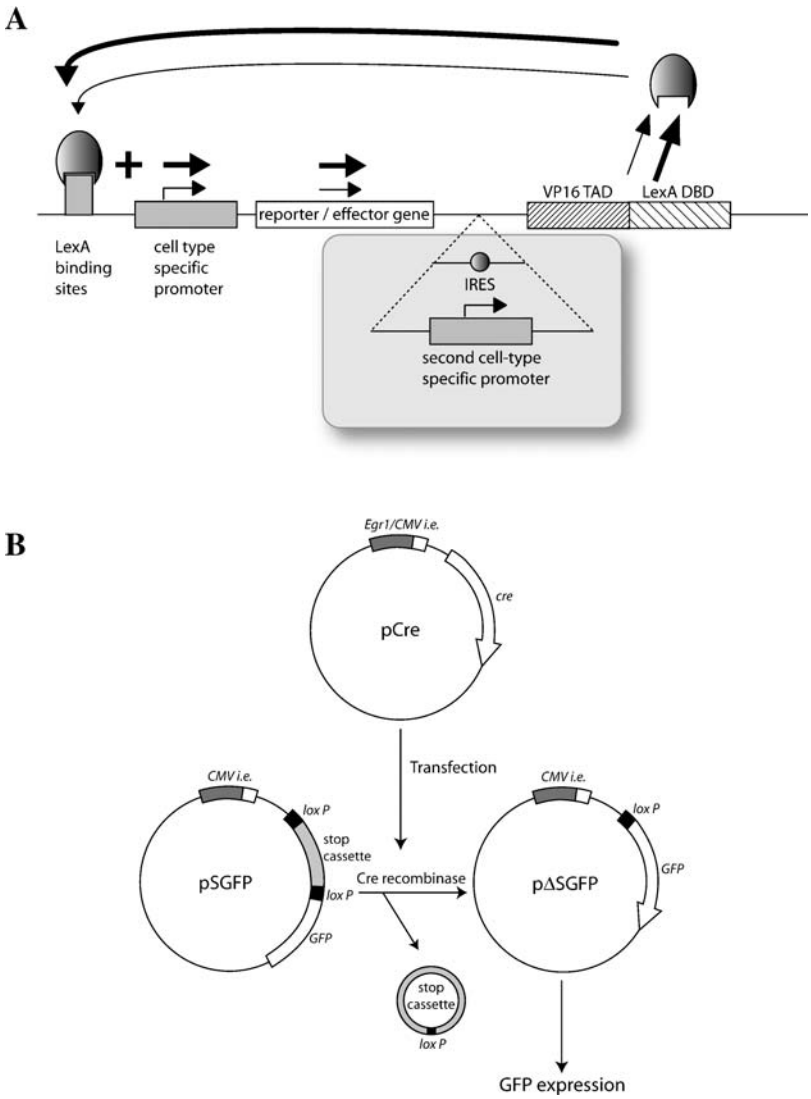

Fig. 5. (A) Positive feedback loop. Schematic illustration of the positive feedback loop mechanism. Within target cells, the cell-type-specific promoter will initiate the first round of expression (thin arrows), leading to simultaneous expression of both the reporter/effector gene and the VP16-LexA fusion protein. Interaction of the VP16-LexA protein with the *LexA*-binding sites upstream of the first cell type specific promoter then results in transactivation and enhancement of transcription (thick arrow). Coexpression of the two separate genes can be realized by using either an IRES or two internal promoters. Adapted from Nettelbeck *et al.* (57) with permission. (B) Cre/*loxP* gene switch system. After cotransfection of the plasmids p*cre* and pS*GFP*, cre recombinase is expressed from the plasmid p*cre* and excises the stop cassette from pS*GFP* via the *loxP* sites. The pS*GFP* plasmid that is generated expresses the *GFP* gene driven by the CMV IE promoter/enhancer. Adapted from Scott *et al.* (82) with permission.

elements of the strong human tyrosinase promoter to select one with the highest level of specificity regardless of its activity and then successfully increased expression by using a feedback loop (136).

Gene Excision by the Cre/*loxP* System

In cases where only temporary protein production is required, it may be desirable to definitively switch on or off gene expression upon an external signal. One method to do so is by excision of specific DNA sequences by recombinases.

Cre is a recombinase derived from *E. coli* phage P1 (129) that excises DNA fragments flanked by *loxP* sites. The DNA is excised as a circular molecule, leaving a 34-bp *loxP* site on each reaction product (138). For gene therapy purposes, the Cre/*loxP* system could be used to specifically delete DNA sequences within plasmid vectors on command. When a therapeutic gene is expressed from a plasmid in which it is flanked by *loxP* sites, its expression could be terminated by administration of a plasmid encoding Cre. As it is rather difficult to deliver plasmid encoding Cre recombinase to all cells expressing the therapeutic gene, it is very likely that this system will generate only partial shutdown of therapeutic gene expression. The system can also be inverted to switch *on* expression of a therapeutic gene. In this case, a "silenced" plasmid is constructed in which a *loxP*-flanked stop cassette is placed amidst a promoter and the therapeutic gene, thereby preventing transcription of the transgene. The stop cassette is then excised from the plasmid upon Cre expression, leaving a plasmid (that is no longer silenced) with the promoter driving expression of the therapeutic gene. Scott *et al.* combined this Cre/*loxP* system with a radio responsive promoter to obtain a radiation-controlled molecular switch (80,82,139). One plasmid was constructed containing a CMV IE promoter separated from a reporter (or therapeutic) gene by a *loxP*-flanked stop cassette. In a second plasmid, the Cre recombinase coding sequence is placed under the control of the *Egr-1* radiation responsive enhancer combined with a CMV IE promoter. When exposed to radiation, the *Egr-1*/CMV IE promoter starts transcription of the Cre recombinase coding sequence. In turn, the expressed Cre excises the stop cassette from the second plasmid and the CMV IE promoter can drive transcription of the reporter or therapeutic gene (Fig. 5B). With *GFP* as the reporter gene, the system incorporating the Cre/*loxP* mechanism resulted in approximately a 14.4-fold higher fluorescence when compared to a system in which *GFP* was directly controlled by the *Egr-1*/CMV IE promoter. When the herpes simplex virus thymidine kinase (HSVtk) gene was used to mediate cell killing, the switch-incorporated system resulted in almost the same increase in sensitivity to ganciclovir as that achieved by a system in which HSVtk was directly controlled by a CMV IE promoter. Also, 3-fold more growth inhibition was accomplished when compared to a system in which the *Egr-1*/CMV IE promoter directly controlled HSVtk. Importantly, for the switch-incorporated system, a 1-Gy radiation dose achieved cell growth inhibition equivalent to that of a 3-Gy dose for systems without the switch (82).

All these results highlight the benefits of incorporating the Cre/*loxP* system, especially when a highly specific but relatively weak promoter is to be used. Cre/*loxP* can increase specificity and activity, whereas leakiness is expected to be low because the therapeutic gene is silenced by a stop cassette in the absence of Cre. In directly inducible systems based on hybridization of inducible regulatory elements with strong constitutive promoter elements, strength of the constitutive promoter is reduced as a compromise to increase specificity and reduce leakiness. The great advantage of the Cre/*loxP* system is that the strong CMV IE promoter can be used to its full potential, while at the same time expression is under control of an external signal. Additional issues to be addressed to make this vector system of practical use include engineering the system into one single plasmid and examining whether the excised circular DNA molecule and the *loxP* site in the therapeutic plasmid that remains after excision are in any way harmful.

Perhaps a more substantial point of concern is the finding that sequences exist in human and mouse genomes that despite being divergent from *loxP* are capable of supporting Cre-mediated recombination (137). These so-called *pseudo-lox* sites are shown to support recombination at up to 100% of the efficiency of native *loxP* sites when tested in bacterial assays and to support Cre-mediated integration and excision in a human cell environment. This observation might complicate application of Cre in gene therapy strategies, as it implies that administration of Cre could possibly affect the host's genome. However, additional work is required to determine whether the efficiency of Cre-mediated deletion is dependent on features like chromatin structure, transcription rate, or DNA methylation. This information is required to estimate the actual effect of Cre when expressed in human cells and, consequently, to determine whether it will be useful for gene therapy applications.

The use of DNA excision is not limited to creating gene switches, but perhaps more importantly allows removal of bacterial plasmid components that are required for replication and selection in bacteria during the production process, but are unnecessary (and often undesired) for expression in human cells. Riu *et al.* (140) describe the excision of the purified transgene expression cassette from plasmids *in vivo*. They constructed plasmids in which the transgene expression cassette was flanked by two *Saccharomyces cerevisiae* mitochondrial endonuclease (I-*Sce*I) recognition sites and co-injected these with plasmids encoding I-*Sce*I cDNA into mouse liver. *In vivo*, I-*Sce*I expression leads to excision of a linear purified expression cassette, free of bacterial DNA. The two free double-stranded DNA ends then ligate intermolecularly to form large concatemers, or intramolecularly to form circular DNA molecules, the latter being reported as the preferred route (141). Such structures have been shown to reside extrachromosomally and remain active for several months (142). As previously described, bacterial DNA has a silencing effect on transgene expression when covalently attached to the transgene. Riu *et al.* (140) show that removal of this covalent linkage results in a significant increase in both level and persistence of expression. I-*Sce*I is regarded as a suitable and safe endonuclease as it is highly specific and has not been reported to cleave in human genomic DNA. In addition, much experience with its use in cells from a variety of organisms is on hand (140).

Expression of Multiple Genes Combined in a Single Plasmid

Many of the more advanced gene therapy strategies described require co-expression of multiple genes. To limit burdening of the delivery vector and to guarantee successful cotransfection of the required genes, it is favorable to combine the transgenes in a single plasmid construct rather than codelivering multiple plasmids. Development of strategies to succeed in coexpression of two or more genes from a single construct (bi- or polycistronic vectors, respectively) will therefore be of increasing importance. The diverse strategies now available have been reviewed by de Felipe (143). Important aspects that need to be considered when choosing a strategy are the size of the element, coordination of expression of the multiple genes, and control over the relative expression patterns. Ideally, the multiple genes are expressed equally and predictably. This can best be achieved when using a single open reading frame (ORF). However, this strategy is often not suitable as the expression of multiple proteins in a single ORF results in physically linked proteins that may not always be functional. In creating polycistronic vectors, most experience is based on the use of IRESs, and until now this still represents the best available way to ensure coexpression of multiple genes in a single plasmid construct (successful coexpression has been reported in >90% of cells). IRESs also enable translation from RNAs produced by RNA polymerases other than RNA polymerase II, as is the case for the cytoplasmic expression strategy based on bacteriophage T7 RNA polymerase as described previously.

A disadvantage of IRESs is that expression levels of the genes upstream and downstream of the IRES are unequal, with the downstream gene being expressed at significantly lower levels and the exact balance being dependent on cell type and transgenes involved. Another disadvantage is the relatively large size of IRESs (~0.5 kb) compared to some other elements. However, the isolation of mini-IRES sequences of less than 0.1 kb may offer a solution (144). Some IRESs require the presence of part of the N-terminus of the original viral protein they belonged to for their full activity. This is undesirable, as it means that additional viral sequences must be incorporated in the plasmid and a small part of the viral protein will consequently be present in the translated transgene product. Traditionally, IRESs from viral genes were used, and these lack the possibilities for specificity or regulation of expression. Some of the advanced gene therapy strategies previously described rely on individually controlled expression of multiple genes, and incorporation of internal promoters has long been regarded as the only strategy suitable for this purpose. Interestingly, the discovery of certain cellular IRESs sheds a new light on this issue. It has been discovered that these IRESs require IRES *trans*-acting factors (ITAFs) present in the host cell for their activity. Similar to promoter specificity relying on the presence of different sets of transcription factors, IRESs display specificity based on specific interactions with different ITAFs. IRESs seem to resemble promoters regarding specificity and inducibility: they have been found in mRNAs encoding growth factors, oncogenes, proteins involved in apoptosis and cell proliferation and also in mRNAs corresponding to stress proteins (formed in response to hypoxia, heat etc.).

With the increasing range of methods becoming available for establishing coexpression of multiple genes, one is given the opportunity to carefully select an element with optimal characteristics for the designed expression cassette.

CONCLUSION

Gene therapy has the potential to treat a great variety of severe diseases, including genetic disorders and cancer, but to date clinical applications have remained few due to inefficiency of delivery and expression. Although efficient delivery of genes to the required cell population is a critical aspect and still leaves much for improvement, also optimizing the plasmid vector can lead to increased or prolonged levels of expression and may therefore play an important role in compensating the limited transfection efficiency achieved with most nonviral carriers. In this review, we have attempted to give an overview of the work that has been done on optimizing plasmid vectors for gene therapy applications.

To date, the most frequently used expression vectors in nonviral gene delivery systems make use of viral elements (promoters/enhancers) to drive the expression of the transgene. Viral-based expression vectors have established proof of principle, but due to immunostimulatory and silencing effects provoked in host cells, it is predicted that they will eventually be outlasted by their cell-based counterparts. Moreover, cellular promoters offer the opportunity for transcriptional targeting, which will contribute to safe and efficient *in vivo* human gene therapy due to restricted expression of transgenes in target tissue. Together with the use of increasingly potent therapeutic genes (e.g., suicide genes) developed to compensate for inefficient gene delivery, a need is generated for improved targeting strategies. In this light, tumor-specific expression strategies will especially prove useful to restrict cytotoxic gene expression to malignant cells or tissues.

However, it seems that with increasing specificity comes decreasing strength. It will therefore be essential to either further increase specificity or further increase strength of weak but highly specific regulatory elements. Technologies incorporating positive feedback loops are estimated to contribute significantly to this purpose.

Additional layers of specificity can be offered by externally inducible systems. These will not only prove useful in further restricting expression to target tissues, but will be indispensable for treatment of diseases that require synthesis of proteins within a small therapeutic range, such as diabetes. When using physical stimuli to induce expression, one has the advantage of being able to target both gene expression and the inducing stimulus to the target issue. In case of drug-inducible gene regulation, the ease of oral administration with which expression of a therapeutic protein can be induced is a great advantage. Furthermore, the extreme diversity of chemically inducible systems offers the opportunity of fine-tuning expression patterns to meet pharmacological/pharmacokinetic requirements.

Another drawback of currently used nonviral gene delivery systems is the transient nature of gene expression. This problem may, however, be solved in the nearby future. Systems incorporating sequences that allow replication and

maintenance, such as episomally replicating vectors, might succeed in establishing stable transgene expression over prolonged periods of time.

Interesting results are expected of so-called autoregulating plasmid vectors. These strategies are based on the incorporation of genes that encode cofactors required for transcription of the plasmid (e.g., transcription factors, polymerase) within the plasmid such that its expression is less dependent on host factors.

Considering the strategies described in this review, optimism is justified that eventually expression cassettes can be created in which relevant parameters for transgene expression (e.g., promoter strength, specificity, leakiness, inducibility, efficiency, safety, duration of expression, kinetics, and possibility of termination) are carefully balanced to meet requirements for clinical application and ultimately to realize human gene therapy.

REFERENCES

1. T. Friedmann and R. Roblin. Gene therapy for human genetic disease? *Science* **175**:949–955 (1972).
2. W. Walther and U. Stein. Viral vectors for gene transfer: a review of their use in the treatment of human diseases. *Drugs* **60**:249–271 (2000).
3. A. V. Kabanov, P. L. Felgner, and L. W. Seymour. *Self-assembling Complexes for Gene Delivery—From Laboratory to Clinical Trial*, John Wiley & Sons, New York, 1998.
4. T. Niidome and L. Huang. Gene therapy progress and prospects: nonviral vectors. *Gene Ther.* **9**:1647–1652 (2002).
5. A. L. Parker, C. Newman, S. Briggs, L. Seymour, and P. J. Sheridan. Nonviral gene delivery: techniques and implications for molecular medicine. *Expert Rev. Mol. Med.* **2003**:1–15 (2003).
6. H. Ma and S. L. Diamond. Nonviral gene therapy and its delivery systems. *Curr. Pharm. Biotechnol.* **2**:1–17 (2001).
7. C. M. Roth and S. Sundaram. Engineering synthetic vectors for improved DNA delivery: insights from intracellular pathways. *Annu. Rev. Biomed. Eng.* **6**:397–426 (2004).
8. S. C. De Smedt, J. Demeester, and W. E. Hennink. Cationic polymer based gene delivery systems. *Pharm. Res.* **17**:113–126 (2000).
9. C. W. Pouton, P. Lucas, B. J. Thomas, A. N. Uduehi, D. A. Milroy, and S. H. Moss. Polycation–DNA complexes for gene delivery: a comparison of the biopharmaceutical properties of cationic polypeptides and cationic lipids. *J. Control. Release* **53**:289–299 (1998).
10. G. Orphanides and D. Reinberg. A unified theory of gene expression. *Cell* **108**:439–451 (2002).
11. B. Alberts. *Molecular Biology of the Cell*, 4th edn. Garland Science, New York, 2002.
12. M. Levine and R. Tjian. Transcription regulation and animal diversity. *Nature* **424**:147–151 (2003).
13. L. Stryer. *Biochemistry*, 4th edn. W.H. Freeman and Company, New York, 1995.
14. D. A. Dean, B. S. Dean, S. Muller, and L. C. Smith. Sequence requirements for plasmid nuclear import. *Exp. Cell Res.* **253**: 713–722 (1999).
15. J. K. Utvik, A. Nja, and K. Gundersen. DNA injection into single cells of intact mice. *Hum. Gene Ther.* **10**:291–300 (1999).
16. D. A. Dean, D. D. Strong, and W. E. Zimmer. Nuclear entry of nonviral vectors. *Gene Ther.* **12**:881–890 (2005).
17. M. Carriere, V. Escriou, A. Savarin, and D. Scherman. Coupling of importin beta binding peptide on plasmid DNA: transfection efficiency is increased by modification of lipoplex's physico-chemical properties. *BMC Biotechnol.* **3**:14 (2003).
18. A. G. Rebuffat, A. R. Nawrocki, P. E. Nielsen, A. G. Bernasconi, E. Bernal-Mendez, B. M. Frey, and F. J. Frey. Gene delivery by a steroid–peptide nucleic acid conjugate. *FASEB J.* **16**:1426–1428 (2002).
19. D. A. Dean. Import of plasmid DNA into the nucleus is sequence specific. *Exp. Cell Res.* **230**:293–302 (1997).
20. S. Li, F. C. MacLaughlin, J. G. Fewell, M. Gondo, J. Wang, F. Nicol, D. A. Dean, and L. C. Smith. Muscle-specific enhancement of gene expression by incorporation of SV40 enhancer in the expression plasmid. *Gene Ther.* **8**:494–497 (2001).
21. J. L. Young, J. N. Benoit, and D. A. Dean. Effect of a DNA nuclear targeting sequence on gene transfer and expression of plasmids in the intact vasculature. *Gene Ther.* **10**:1465–1470 (2003).
22. A. Mesika, I. Grigoreva, M. Zohar, and Z. Reich. A regulated, NFkappaB-assisted import of plasmid DNA into mammalian cell nuclei. *Mol. Ther.* **3**:653–657 (2001).
23. F. Langle-Rouault, V. Patzel, A. Benavente, M. Taillez, N. Silvestre, A. Bompard, G. Sczakiel, E. Jacobs, and K. Rittner. Up to 100-fold increase of apparent gene expression in the presence of Epstein-Barr virus oriP sequences and EBNA1: implications of the nuclear import of plasmids. *J. Virol.* **72**:6181–6185 (1998).
24. J. Vacik, B. S. Dean, W. E. Zimmer, and D. A. Dean. Cell-specific nuclear import of plasmid DNA. *Gene Ther.* **6**:1006–1014 (1999).
25. X. Gao and L. Huang. Cytoplasmic expression of a reporter gene by co-delivery of T7 RNA polymerase and T7 promoter sequence with cationic liposomes. *Nucleic Acids Res.* **21**: 2867–2872 (1993).
26. C. A. Raskin, G. A. Diaz, and W. T. McAllister. T7 RNA polymerase mutants with altered promoter specificities. *Proc. Natl. Acad. Sci. USA* **90**:3147–3151 (1993).
27. M. Brisson, Y. He, S. Li, J. P. Yang, and L. Huang. A novel T7 RNA polymerase autogene for efficient cytoplasmic expression of target genes. *Gene Ther.* **6**:263–270 (1999).
28. X. Chen, Y. Li, K. Xiong, Y. Xie, S. Aizicovici, R. Snodgrass, T. E. Wagner, and D. Platika. A novel nonviral cytoplasmic gene expression system and its implications in cancer gene therapy. *Cancer Gene Ther.* **2**:281–289 (1995).
29. E. Casado, D. M. Nettelbeck, J. Gomez-Navarro, A. Hemminki, M. Gonzalez Baron, G. P. Siegal, M. N. Barnes, R. D. Alvarez, and D. T. Curiel. Transcriptional targeting for ovarian cancer gene therapy. *Gynecol. Oncol.* **82**:229–237 (2001).
30. Y. S. Haviv and D. T. Curiel. Conditional gene targeting for cancer gene therapy. *Adv. Drug Deliv. Rev.* **53**:135–154 (2001).
31. M. J. Cooper. Noninfectious gene transfer and expression systems for cancer gene therapy. *Semin. Oncol.* **23**:172–187 (1996).
32. M. Lee, S. Han, K. S. Ko, and S. W. Kim. Cell type specific and glucose responsive expression of interleukin-4 by using insulin promoter and water soluble lipopolymer. *J. Control. Release* **75**:421–429 (2001).
33. J. Black and J. M. Vos. Establishment of an oriP/EBNA1-based episomal vector transcribing human genomic beta-globin in cultured murine fibroblasts. *Gene Ther.* **9**:1447–1454 (2002).
34. W. Walther and U. Stein. Cell type specific and inducible promoters for vectors in gene therapy as an approach for cell targeting. *J. Mol. Med.* **74**:379–392 (1996).
35. H. S. Pandha, L. A. Martin, A. Rigg, H. C. Hurst, G. W. Stamp, K. Sikora, and N. R. Lemoine. Genetic prodrug activation therapy for breast cancer: A phase I clinical trial of erbB-2-directed suicide gene expression. *J. Clin. Oncol.* **17**:2180–2189 (1999).
36. D. M. Nettelbeck, V. Jerome, and R. Muller. Gene therapy: designer promoters for tumour targeting. *Trends Genet.* **16**: 174–181 (2000).
37. E. Casado, J. Gomez-Navarro, M. Yamamoto, Y. Adachi, C. J. Coolidge, W. O. Arafat, S. D. Barker, M. H. Wang, P. J. Mahasreshti, A. Hemminki, M. Gonzalez-Baron, M. N. Barnes, T. B. Pustilnik, G. P. Siegal, R. D. Alvarez, and D. T. Curiel. Strategies to accomplish targeted expression of transgenes in ovarian cancer for molecular therapeutic applications. *Clin. Cancer Res.* **7**:2496–2504 (2001).

38. K. S. Koeneman, C. Kao, S. C. Ko, L. Yang, Y. Wada, D. F. Kallmes, J. Y. Gillenwater, H. E. Zhau, L. W. Chung, and T. A. Gardner. Osteocalcin-directed gene therapy for prostate-cancer bone metastasis. *World J. Urol.* **18**:102–110 (2000).
39. G. Palu, R. Bonaguro, E. Gnatta, E. Franchin, and L. Barzon. Gene therapy of brain and endocrine tumors. *Croat. Med. J.* **42**:473–477 (2001).
40. A. Rigg and K. Sikora. Genetic prodrug activation therapy. *Mol. Med. Today* **3**:359–366 (1997).
41. R. G. Vile and I. R. Hart. *In vitro* and *in vivo* targeting of gene expression to melanoma cells. *Cancer Res.* **53**:962–967 (1993).
42. R. M. Diaz, T. Eisen, I. R. Hart, and R. G. Vile. Exchange of viral promoter/enhancer elements with heterologous regulatory sequences generates targeted hybrid long terminal repeat vectors or gene therapy of melanoma. *J. Virol.* **72**:789–795 (1998).
43. J. Gu, S. Kagawa, M. Takakura, S. Kyo, M. Inoue, J. A. Roth, and B. Fang. Tumor-specific transgene expression from the human telomerase reverse transcriptase promoter enables targeting of the therapeutic effects of the Bax gene to cancers. *Cancer Res.* **60**:5359–5364 (2000).
44. C. Spitzweg, S. Zhang, E. R. Bergert, M. R. Castro, B. McIver, A. E. Heufelder, D. J. Tindall, C. Y. Young, and J. C. Morris. Prostate-specific antigen (PSA) promoter-driven androgen-inducible expression of sodium iodide symporter in prostate cancer cell lines. *Cancer Res.* **59**:2136–2141 (1999).
45. M. Costa, Y. Shen, F. Maurer, and R. L. Medcalf. Transcriptional regulation of the tissue-type plasminogen-activator gene in human endothelial cells: identification of nuclear factors that recognise functional elements in the tissue-type plasminogen-activator gene promoter. *Eur. J. Biochem.* **258**:123–131 (1998).
46. P. J. Cowan, T. A. Shinkel, E. J. Witort, H. Barlow, M. J. Pearse, and A. J. d'Apice. Targeting gene expression to endothelial cells in transgenic mice using the human intercellular adhesion molecule 2 promoter. *Transplantation* **62**:155–160 (1996).
47. A. Damert, M. Machein, G. Breier, M. Q. Fujita, D. Hanahan, W. Risau, and K. H. Plate. Up-regulation of vascular endothelial growth factor expression in a rat glioma is conferred by two distinct hypoxia-driven mechanisms. *Cancer Res.* **57**:3860–3864 (1997).
48. B. Enholm, T. Karpanen, M. Jeltsch, H. Kubo, F. Stenback, R. Prevo, D. G. Jackson, S. Yla-Herttuala, and K. Alitalo. Adenoviral expression of vascular endothelial growth factor-C induces lymphangiogenesis in the skin. *Circ. Res.* **88**:623–629 (2001).
49. W. Graulich, D. M. Nettelbeck, D. Fischer, T. Kissel, and R. Muller. Cell type specificity of the human endoglin promoter. *Gene* **227**:55–62 (1999).
50. U. Jager, Y. Zhao, and C. D. Porter. Endothelial cell-specific transcriptional targeting from a hybrid long terminal repeat retrovirus vector containing human prepro-endothelin-1 promoter sequences. *J. Virol.* **73**:9702–9709 (1999).
51. N. K. Jonas and M. Saleh. The expression of vascular endothelial platelet endothelial cell adhesion molecule-1 is not regulated by IFN-gamma treatment of C6 tumors *in vivo*. *Int. J. Oncol.* **17**:939–945 (2000).
52. A. Kaipainen, J. Korhonen, T. Mustonen, V. W. Hinsberghvan, G. H. Fang, D. Dumont, M. Breitman, and K. Alitalo. Expression of the fms-like tyrosine kinase 4 gene becomes restricted to lymphatic endothelium during development. *Proc. Natl. Acad. Sci. USA* **92**:3566–3570 (1995).
53. J. Korhonen, I. Lahtinen, M. Halmekyto, L. Alhonen, J. Janne, D. Dumont, and K. Alitalo. Endothelial-specific gene expression directed by the tie gene promoter *in vivo*. *Blood* **86**:1828–1835 (1995).
54. N. Koshikawa, K. Takenaga, M. Tagawa, and S. Sakiyama. Therapeutic efficacy of the suicide gene driven by the promoter of vascular endothelial growth factor gene against hypoxic tumor cells. *Cancer Res.* **60**:2936–2941 (2000).
55. K. Morishita, D. E. Johnson, and L. T. Williams. A novel promoter for vascular endothelial growth factor receptor (flt-1) that confers endothelial-specific gene expression. *J. Biol. Chem.* **270**:27948–27953 (1995).
56. A. S. Neish, A. J. Williams, H. J. Palmer, M. Z. Whitley, and T. Collins. Functional analysis of the human vascular cell adhesion molecule 1 promoter. *J. Exp. Med.* **176**:1583–1593 (1992).
57. D. M. Nettelbeck, V. Jerome, and R. Muller. A strategy for enhancing the transcriptional activity of weak cell type-specific promoters. *Gene Ther.* **5**:1656–1664 (1998).
58. S. A. Nicklin, P. N. Reynolds, M. J. Brosnan, S. J. White, D. T. Curiel, A. F. Dominiczak, and A. H. Baker. Analysis of cell-specific promoters for viral gene therapy targeted at the vascular endothelium. *Hypertension* **38**:65–70 (2001).
59. U. S. Ryan, R. Zhong, B. A. Hayes, G. Visner, and M. L. Sauther. Regulation of endothelin-1 expression in normal and transfected endothelial cells. *J. Cardiovasc. Pharmacol.* **22** (Suppl): S38–S41 (1993).
60. T. Shibata, N. Akiyama, M. Noda, K. Sasai, and M. Hiraoka. Enhancement of gene expression under hypoxic conditions using fragments of the human vascular endothelial growth factor and the erythropoietin genes. *Int. J. Radiat. Oncol. Biol. Phys.* **42**:913–916 (1998).
61. N. Varda-Bloom, A. Shaish, A. Gonen, K. Levanon, S. Greenbereger, S. Ferber, H. Levkovitz, D. Castel, I. Goldberg, A. Afek, Y. Kopolovitc, and D. Harats. Tissue-specific gene therapy directed to tumor angiogenesis. *Gene Ther.* **8**:819–827 (2001).
62. K. K. Wu. Regulation of endothelial nitric oxide synthase activity and gene expression. *Ann. N. Y. Acad. Sci.* **962**:122–130 (2002).
63. M. T. Lin, F. Wang, J. Uitto, and K. Yoon. Differential expression of tissue-specific promoters by gene gun. *Br. J. Dermatol.* **144**:34–39 (2001).
64. A. Ido, K. Nakata, Y. Kato, K. Nakao, K. Murata, M. Fujita, N. Ishii, T. Tamaoki, H. Shiku, and S. Nagataki. Gene therapy for hepatoma cells using a retrovirus vector carrying herpes simplex virus thymidine kinase gene under the control of human alpha-fetoprotein gene promoter. *Cancer Res.* **55**:3105–3109 (1995).
65. J. Q. Cheng, B. Ruggeri, W. M. Klein, G. Sonoda, D. A. Altomare, D. K. Watson, and J. R. Testa. Amplification of AKT2 in human pancreatic cells and inhibition of AKT2 expression and tumorigenicity by antisense RNA. *Proc. Natl. Acad. Sci. USA* **93**:3636–3641 (1996).
66. M. Takakura, S. Kyo, T. Kanaya, H. Hirano, J. Takeda, M. Yutsudo, and M. Inoue. Cloning of human telomerase catalytic subunit (hTERT) gene promoter and identification of proximal core promoter sequences essential for transcriptional activation in immortalized and cancer cells. *Cancer Res.* **59**:551–557 (1999).
67. T. Shibata, A. J. Giaccia, and J. M. Brown. Development of a hypoxia-responsive vector for tumor-specific gene therapy. *Gene Ther.* **7**:493–498 (2000).
68. P. H. Maxwell, G. U. Dachs, J. M. Gleadle, L. G. Nicholls, A. L. Harris, I. J. Stratford, O. Hankinson, C. W. Pugh, and P. J. Ratcliffe. Hypoxia-inducible factor-1 modulates gene expression in solid tumors and influences both angiogenesis and tumor growth. *Proc. Natl. Acad. Sci. USA* **94**:8104–8109 (1997).
69. A. P. Levy, N. S. Levy, and M. A. Goldberg. Post-transcriptional regulation of vascular endothelial growth factor by hypoxia. *J. Biol. Chem.* **271**:2746–2753 (1996).
70. G. Gazit, G. Hung, X. Chen, W. F. Anderson, and A. S. Lee. Use of the glucose starvation-inducible glucose-regulated protein 78 promoter in suicide gene therapy of murine fibrosarcoma. *Cancer Res.* **59**:3100–3106 (1999).
71. A. S. Lee. The glucose-regulated proteins: stress induction and clinical applications. *Trends Biochem. Sci.* **26**:504–510 (2001).
72. M. W. Robertson 3rd, M. Wang, G. P. Siegal, M. Rosenfeld, R. S. Ashford 2nd, R. D. Alvarez, R. I. Garver, and D. T. Curiel. Use of a tissue-specific promoter for targeted expression of the herpes simplex virus thymidine kinase gene in cervical carcinoma cells. *Cancer Gene Ther.* **5**:331–336 (1998).
73. I. Chung, P. E. Schwartz, R. G. Crystal, G. Pizzorno, J. Leavitt, and A. B. Deisseroth. Use of L-plastin promoter to develop an adenoviral system that confers transgene expression in ovarian cancer cells but not in normal mesothelial cells. *Cancer Gene Ther.* **6**:99–106 (1999).
74. X. Y. Peng, J. H. Won, T. Rutherford, T. Fujii, D. Zelterman,

G. Pizzorno, E. Sapi, J. Leavitt, B. Kacinski, R. Crystal, P. Schwartz, and A. Deisseroth. The use of the L-plastin promoter for adenoviral-mediated, tumor-specific gene expression in ovarian and bladder cancer cell lines. *Cancer Res.* **61**: 4405–4413 (2001).
75. A. Ido, H. Uto, A. Moriuchi, K. Nagata, Y. Onaga, M. Onaga, T. Hori, S. Hirono, K. Hayashi, T. Tamaoki, and H. Tsubouchi. Gene therapy targeting for hepatocellular carcinoma: selective and enhanced suicide gene expression regulated by a hypoxia-inducible enhancer linked to a human alpha-fetoprotein promote. *Cancer Res.* **61**:3016–3021 (2001).
76. O*ncologie*, Bohn Stafleu Van Loghum, Houten/Diegem, 1996.
77. C. Rome, F. Couillaud, and C. T. Moonen. Spatial and temporal control of expression of therapeutic genes using heat shock protein promoters. *Methods* **35**:188–198 (2005).
78. Q. Huang, J. K. Hu, F. Lohr, L. Zhang, R. Braun, J. Lanzen, J. B. Little, M. W. Dewhirst, and C. Y. Li. Heat-induced gene expression as a novel targeted cancer gene therapy strategy. *Cancer Res.* **60**:3435–3439 (2000).
79. J. Zhu, B. Gao, J. Zhao, and A. Balmain. Targeting gene expression to tumor cells with loss of wild-type p53 function. *Cancer Gene Ther.* **7**:4–12 (2000).
80. B. Marples, O. Greco, M. C. Joiner, and S. D. Scott. Molecular approaches to chemo-radiotherapy. *Eur. J. Cancer* **38**:231–239 (2002).
81. S. D. Scott, M. C. Joiner, and B. Marples. Optimizing radiation-responsive gene promoters for radiogenetic cancer therapy. *Gene Ther.* **9**:1396–1402 (2002).
82. S. D. Scott, B. Marples, J. H. Hendry, L. S. Lashford, M. J. Embleton, R. D. Hunter, A. Howell, and G. P. Margison. A radiation-controlled molecular switch for use in gene therapy of cancer. *Gene Ther.* **7**:1121–1125 (2000).
83. M. A. Stackhouse and D. J. Buchsbaum. Radiation to control gene expression. *Gene Ther.* **7**:1085–1086 (2000).
84. J. Worthington, H. O. McCarthy, E. Barrett, C. Adams, T. Robson, and D. G. Hirst. Use of the radiation-inducible WAF1 promoter to drive iNOS gene therapy as a novel anti-cancer treatment. *J. Gene Med.* **6**:673–680 (2004).
85. J. Worthington, T. Robson, M. Murray, M. O'Rourke, G. Keilty, and D. G. Hirst. Modification of vascular tone using iNOS under the control of a radiation-inducible promoter. *Gene Ther.* **7**:1126–1131 (2000).
86. S. Nuyts, L. Van Mellaert, J. Theys, W. Landuyt, E. Bosmans, J. Anne, and P. Lambin. Radio-responsive recA promoter significantly increases TNFalpha production in recombinant clostridia after 2 Gy irradiation. *Gene Ther.* **8**:1197–1201 (2001).
87. T. Clackson. Regulated gene expression systems. *Gene Ther.* **7**:120–125 (2000).
88. K. J. Harrington, E. Linardakis, and R. G. Vile. Transcriptional control: an essential component of cancer gene therapy strategies? *Adv. Drug Deliv. Rev.* **44**:167–184 (2000).
89. T. Clackson. Controlling mammalian gene expression with small molecules. *Curr. Opin. Chem. Biol.* **1**:210–218 (1997).
90. T. Kafri. Air-conditioning for regulated transgene expression. *Gene Ther.* **12**:383–385 (2005).
91. M. Gossen and H. Bujard. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* **89**:5547–5551 (1992).
92. Z. Zhu, T. Zheng, C. G. Lee, R. J. Homer, and J. A. Elias. Tetracycline-controlled transcriptional regulation systems: advances and application in transgenic animal modeling. *Semin. Cell Dev. Biol.* **13**:121–128 (2002).
93. M. O. Wilson, K. T. Scougall, J. Ratanamart, E. A. McIntyre, and J. A. Shaw. Tetracycline-regulated secretion of human (pro)insulin following plasmid-mediated transfection of human muscle. *J. Mol. Endocrinol.* **34**:391–403 (2005).
94. U. Baron, M. Gossen, and H. Bujard. Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *Nucleic Acids Res.* **25**:2723–2729 (1997).
95. S. Y. Corbel and F. M. Rossi. Latest developments and *in vivo* use of the Tet system: *ex vivo* and *in vivo* delivery of tetracycline-regulated genes. *Curr. Opin. Biotechnol.* **13**:448–452 (2002).
96. E. S. Ngan, K. Schillinger, F. DeMayo, and S. Y. Tsai. The mifepristone-inducible gene regulatory system in mouse models of disease and gene therapy. *Semin. Cell Dev. Biol.* **13**:143–149 (2002).
97. J. L. Nordstrom. Antiprogestin-controllable transgene regulation *in vivo*. *Curr. Opin. Biotechnol.* **13**:453–458 (2002).
98. H. Kobayashi, Y. Takemura, and H. Miyachi. Novel approaches to reversing anti-cancer drug resistance using gene-specific therapeutics. *Hum. Cell* **14**:172–184 (2001).
99. W. Walther, U. Stein, I. Fichtner, M. Alexander, R. H. Shoemaker, and P. M. Schlag. Mdr1 promoter-driven tumor necrosis factor-alpha expression for a chemotherapy-controllable combined *in vivo* gene therapy and chemotherapy of tumors. *Cancer Gene Ther.* **7**:893–900 (2000).
100. W. Walther, J. Wendt, and U. Stein. Employment of the mdr1 promoter for the chemotherapy-inducible expression of therapeutic genes in cancer gene therapy. *Gene Ther.* **4**:544–552 (1997).
101. W. Weber, M. Rimann, M. Spielmann, B. Keller, M. Daoud-El Baba, D. Aubel, C. C. Weber, and M. Fussenegger. Gas-inducible transgene expression in mammalian cells and mice. *Nat. Biotechnol.* **22**:1440–1444 (2004).
102. S. Hartenbach and M. Fussenegger. Autoregulated, bidirectional and multicistronic gas-inducible mammalian as well as lentiviral expression vectors. *J. Biotechnol.* **120**:83–98 (2005).
103. R. Pollock and T. Clackson. Dimerizer-regulated gene expression. *Curr. Opin. Biotechnol.* **13**:459–467 (2002).
104. M. J. Cooper, M. Lippa, J. M. Payne, G. Hatzivassiliou, E. Reifenberg, B. Fayazi, J. C. Perales, L. J. Morrison, D. Templeton, R. L. Piekarz, and J. Tan. Safety-modified episomal vectors for human gene therapy. *Proc. Natl. Acad. Sci. USA* **94**:6450–6455 (1997).
105. C. J. Coates, J. M. Kaminski, J. B. Summers, D. J. Segal, A. D. Miller, and A. F. Kolb. Site-directed genome modification: derivatives of DNA-modifying enzymes as targeting tools. *Trends Biotechnol.* **23**:407–419 (2005).
106. C. E. Thomas, A. Ehrhardt, and M. A. Kay. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4**:346–358 (2003).
107. Z. Izsvak and Z. Ivics. Sleeping beauty transposition: biology and applications for molecular therapy. *Mol. Ther.* **9**:147–156 (2004).
108. D. J. Glover, H. J. Lipps, and D. A. Jans. Towards safe, non-viral therapeutic gene expression in humans. *Nat. Rev. Genet.* **6**:299–310 (2005).
109. F. Buchholz, P. O. Angrand, and A. F. Stewart. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat. Biotechnol.* **16**:657–662 (1998).
110. S. R. Yant, L. Meuse, W. Chiu, Z. Ivics, Z. Izsvak, and M. A. Kay. Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat. Genet.* **25**:35–41 (2000).
111. E. Montini, P. K. Held, M. Noll, N. Morcinek, M. Al-Dhalimy, M. Finegold, S. R. Yant, M. A. Kay, and M. Grompe. *In vivo* correction of murine tyrosinemia type I by DNA-mediated transposition. *Mol. Ther.* **6**:759–769 (2002).
112. L. R. Belur, J. L. Frandsen, A. J. Dupuy, D. H. Ingbar, D. A. Largaespada, P. B. Hackett, and R. Scott McIvor. Gene insertion and long-term expression in lung mediated by the Sleeping Beauty transposon system. *Mol. Ther.* **8**:501–507 (2003).
113. A. F. Kolb, C. J. Coates, J. M. Kaminski, J. B. Summers, A. D. Miller, and D. J. Segal. Site-directed genome modification: nucleic acid and protein modules for targeted integration and gene correction. *Trends Biotechnol.* **23**:399–406 (2005).
114. K. Van Craenenbroeck, P. Vanhoenacker, and G. Haegeman. Episomal vectors for gene expression in mammalian cells. *Eur. J. Biochem.* **267**:5665–5678 (2000).
115. F. D. Cui, T. Kishida, S. Ohashi, H. Asada, K. Yasutomi, E. Satoh, T. Kubo, S. Fushiki, J. Imanishi, and O. Mazda. Highly efficient gene transfer into murine liver achieved by intravenous administration of naked Epstein–Barr virus (EBV)-based plasmid vectors. *Gene Ther.* **8**:1508–1513 (2001).
116. M. Iwai, Y. Harada, S. Tanaka, A. Muramatsu, T. Mori, K. Kashima, J. Imanishi, and O. Mazda. Polyethylenimine-mediated suicide gene transfer induces a therapeutic effect for hepatocellular carcinoma *in vivo* by using an Epstein-Barr

virus-based plasmid vector. *Biochem. Biophys. Res. Commun.* **291**:48–54 (2002).
117. A. Ehrhardt, P. D. Peng, H. Xu, L. Meuse, and M. A. Kay. Optimization of cis-acting elements for gene expression from nonviral vectors *in vivo*. *Hum. Gene Ther.* **14**:215–225 (2003).
118. C. Piechaczek, C. Fetzer, A. Baiker, J. Bode, and H. J. Lipps. A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. *Nucleic Acids Res.* **27**:426–428 (1999).
119. M. Conese, C. Auriche, and F. Ascenzioni. Gene therapy progress and prospects: episomally maintained self-replicating systems. *Gene Ther.* **11**:1735–1741 (2004).
120. V. Oberle, G. Jongde, J. I. Drayer, and D. Hoekstra. Efficient transfer of chromosome-based DNA constructs into mammalian cells. *Biochim. Biophys. Acta* **1676**:223–230 (2004).
121. Z. Y. Chen, C. Y. He, L. Meuse, and M. A. Kay. Silencing of episomal transgene expression by plasmid bacterial DNA elements *in vivo*. *Gene Ther.* **11**:856–864 (2004).
122. P. A. Jones. The DNA methylation paradox. *Trends Genet.* **15**:34–37 (1999).
123. K. D. Robertson and P. A. Jones. DNA methylation: past, present and future directions. *Carcinogenesis* **21**:461–467 (2000).
124. S. Prosch, J. Stein, K. Staak, C. Liebenthal, H. D. Volk, and D. H. Kruger. Inactivation of the very strong HCMV immediate early promoter by DNA CpG methylation *in vitro*. *Biol. Chem. Hoppe-Seyler* **377**:195–201 (1996).
125. R. K. Scheule. The role of CpG motifs in immunostimulation and gene therapy. *Adv. Drug Deliv. Rev.* **44**:119–134 (2000).
126. J. Hejnar, P. Hajkova, J. Plachy, D. Elleder, V. Stepanets, and J. Svoboda. CpG island protects Rous sarcoma virus-derived vectors integrated into nonpermissive cells from DNA methylation and transcriptional suppression. *Proc. Natl. Acad. Sci. USA* **98**:565–569 (2001).
127. M. Holler, G. Westin, J. Jiricny, and W. Schaffner. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev.* **2**:1127–1135 (1988).
128. N. S. Yew, K. X. Wang, M. Przybylska, R. G. Bagley, M. Stedman, J. Marshall, R. K. Scheule, and S. H. Cheng. Contribution of plasmid DNA to inflammation in the lung after administration of cationic lipid:pDNA complexes. *Hum. Gene Ther.* **10**:223–234 (1999).
129. A. M. Krieg, T. Wu, R. Weeratna, S. M. Efler, L. Love-Homan, L. Yang, A. K. Yi, D. Short, and H. L. Davis. Sequence motifs in adenoviral DNA block immune activation by stimulatory CpG motifs. *Proc. Natl. Acad. Sci. USA* **95**:12631–12636 (1998).
130. H. Zhao, S. H. Cheng, and N. S. Yew. Requirements for effective inhibition of immunostimulatory CpG motifs by neutralizing motifs. *Antisense Nucleic Acid Drug Dev.* **10**:381–389 (2000).
131. N. S. Yew, H. Zhao, I. H. Wu, A. Song, J. D. Tousignant, M. Przybylska, and S. H. Cheng. Reduced inflammatory response to plasmid DNA vectors by elimination and inhibition of immunostimulatory CpG motifs. *Mol. Ther.* **1**:255–262 (2000).
132. N. S. Yew, H. Zhao, M. Przybylska, I. H. Wu, J. D. Tousignant, R. K. Scheule, and S. H. Cheng. CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression *in vivo*. *Mol. Ther.* **5**:731–738 (2002).
133. H. Ishikawa, K. Nakata, F. Mawatari, T. Ueki, S. Tsuruta, A. Ido, K. Nakao, Y. Kato, N. Ishii, and K. Eguchi. Utilization of variant-type of human alpha-fetoprotein promoter in gene therapy targeting for hepatocellular carcinoma. *Gene Ther.* **6**:465–470 (1999).
134. C. Solem and P. R. Jensen. Modulation of gene expression made easy. *Appl. Environ. Microbiol.* **68**:2397–2403 (2002).
135. P. R. Jensen and K. Hammer. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Environ. Microbiol.* **64**:82–87 (1998).
136. L. Emiliusen, M. Gough, A. Bateman, A. Ahmed, R. Voellmy, J. Chester, R. M. Diaz, K. Harrington, and R. Vile. A transcriptional feedback loop for tissue-specific expression of highly cytotoxic genes which incorporates an immunostimulatory component. *Gene Ther.* **8**:987–998 (2001).
137. B. Thyagarajan, M. J. Guimaraes, A. C. Groth, and M. P. Calos. Mammalian genomes contain active recombinase recognition sites. *Gene* **244**:47–54 (2000).
138. R. Kuhn and F. Schwenk. Advances in gene targeting methods. *Curr. Opin. Immunol.* **9**:183–188 (1997).
139. S. D. Scott and B. Marples Comment on the use of the cre/loxP recombinase system for gene therapy vectors. *Gene Ther.* **7**: 1706 (2000).
140. E. Riu, D. Grimm, Z. Huang, and M. A. Kay. Increased maintenance and persistence of transgenes by excision of expression cassettes from plasmid sequences *in vivo*. *Hum. Gene Ther.* **16**:558–570 (2000).
141. H. Nakai, T. A. Storm, S. Fuess, and M. A. Kay. Pathways of removal of free DNA vector ends in normal and DNA-PKcs-deficient SCID mouse hepatocytes transduced with rAAV vectors. *Hum. Gene Ther.* **14**:871–881 (2003).
142. Z. Y. Chen, S. R. Yant, C. Y. He, L. Meuse, S. Shen, and M. A. Kay. Linear DNAs concatemerize *in vivo* and result in sustained transgene expression in mouse liver. *Mol. Ther.* **3**:403–410 (2001).
143. P. de Felipe. Polycistronic viral vectors. *Curr. Gene Ther.* **2**: 355–378 (2002).
144. S. A. Chappell, G. M. Edelman, and V. P. Mauro. A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proc. Natl. Acad. Sci. USA* **97**:1536–1541 (2000).