Main features of DNA-based immunization vectors

V. Azevedo1, G. Levitus4, A. Miyoshi1, A.L. Cândido2, A.M. Goes3 and S.C. Oliveira3

Departamentos de 1Biologia Geral, 2Microbiologia, and 3Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil
4Instituto de Investigaciones en Ingenieria Genetica y Biologia Molecular (INGEBI-CONICET-UBA), Buenos Aires, Argentina

Abstract

DNA-based immunization has initiated a new era of vaccine research. One of the main goals of gene vaccine development is the control of the levels of expression in vivo for efficient immunization. Modifying the vector to modulate expression or immunogenicity is of critical importance for the improvement of DNA vaccines. The most frequently used vectors for genetic immunization are plasmids. In this article, we review some of the main elements relevant to their design such as strong promoter/enhancer region, introns, genes encoding antigens of interest from the pathogen (how to choose and modify them), polyadenylation termination sequence, origin of replication for plasmid production in Escherichia coli, antibiotic resistance gene as selectable marker, convenient cloning sites, and the presence of immunostimulatory sequences (ISS) that can be added to the plasmid to enhance adjuvanticity and to activate the immune system. In this review, the specific modifications that can increase overall expression as well as the potential of DNA-based vaccination are also discussed.

Introduction

Genetic immunization represents a novel approach to vaccination. This technology involves transfer of a gene encoding an antigenic protein cloned in expression vectors to a host, leading to the induction of an immune response. During the last two decades, different mammalian gene expression vectors have been developed as well as new methods for direct gene transfer. Several reviews on the subject have been published (1-3). Direct gene transfer may be undertaken using either viral vectors or recombinant plasmid DNA. Viral vectors have the disadvantages of being derived from pathogens like traditional vaccines based on attenuated viruses, and therefore they are of limited interest for the purpose of immunization. In contrast, DNA plasmids encoding antigen(s) are more frequently used because they do not have the inconvenience of classical vaccines: they are safe, inexpensive, easy to produce, heat stable and amenable to genetic manipulation. In this review, we will deal with some of the main elements relevant to the design of vectors used for genetic vaccination. These vectors are Escherichia coli-derived plasmids capable of expressing foreign genes in eukaryotic cells. Conceptually, their structure can be divided into two distinct units: i) a transcription complex unit that drives antigen synthesis and that contains a promoter/enhancer region, introns with functional splic-
ing donor and acceptor sites, sequences encoding an antigenic protein, and signals required for efficient polyadenylation of the transcript. ii) Prokaryotic elements such as replication origin, multiple cloning sites and a selectable marker to facilitate the construction, propagation and amplification of recombinant vectors in bacteria. This plasmid backbone can also carry immunostimulatory sequences (ISS) with adjuvant activity (Figure 1).

**Basic structure of plasmid DNA vaccines**

**The transcription complex unit**

*Enhancer/promoter regions.* Almost all commercial mammalian expression plasmids carry the immediate early promoter/enhancer from pathogenic viruses. Although these promoter elements are from pathogenic viruses, they have become very useful for gene therapy and genetic immunization thanks to their high transcription initiation ability in most mammalian tissues (4). The most commonly used promoter is the one from human cytomegalovirus (hCMV) which induces strong and constitutive expression in a variety of cell types (5-7). The use of alternative promoter/enhancer elements has been discussed, including Rous sarcoma virus (RSV), simian virus 40 (SV40), murine leukemia virus (SL3-3), mouse mammary tumor virus (MMTV) promoter, and HIV long terminal repeat (LTR) promoter (8-10). Among them, the hCMV promoter gave the highest levels of expression of the reported genes tested and/or induced the strongest immune responses (11-13).

Another type of promoters, i.e., mammalian promoters, have also been used. These promoter elements are particularly interesting for vaccine development since they may be more desirable than DNA sequence from pathogenic or tumor-causing viruses for applications involving humans or animals. Gurunathan et al. (14) reported that the bovine promoter/enhancer from the major histocompatibility complex class I (MHC I) gene gave significantly better protection than that obtained with hCMV-derived plasmids. Other mammalian promoters were tested, including those from the beta-actin, muscle-specific heavy chain of myosin and muscle creatine kinase genes (mck) (7,15,16).

Discrete functional entities such as the exogenous enhancer elements of αB crystalline gene (cryB) or mck can be incorporated upstream in relation to the enhancer/promoter region. Hartikka et al. (17) and Dai et al. (18) reported that these entities prolong and increase expression by the CMV promoter/enhancer. In fact, the main function of these elements is to control the gene initiation transcription rate. In spite of the existence of various gene delivery systems, the main question to be addressed is if the initiation of the transferred gene transcription is efficient and how long it lasts. The modulation of this initiation rate can be essential to obtain a positive genetic immunization effect.

It is sometimes advisable to introduce an inducible promoter to control *in vivo* expression. This was done by Dhawan et al. (19) and Liang et al. (20) who modified plasmids to make tetracycline (Tc)-dependent transcriptions. When tetracycline is administered to transfected mice it can act in a repressive or activating way depending on the position of the tetracycline-operator (tetO) control sequence within these Tc-controlled plasmids. The inducible systems have the advantage of overcoming potential immunological

---

**Figure 1 - Plasmid DNA pattern for genetic immunization.** i) The transcription complex unit comprises promoter, intron, antigen DNA and polyadenylation (A) addition sequence, which together drive protein synthesis. ii) The elements of the plasmid backbone unit such as prokaryotic origin of replication (ColE1) and a selectable marker (ampicillin resistance gene; ampR). This plasmid backbone also delivers adjuvant and mitogenic activity via immunostimulatory sequences (ISS).
tolerance that may exist in genetic immunization.

*Introns.* Intervening sequences (or introns) have a beneficial effect on antigen expression. This can be attributed to an enhanced rate of RNA polyadenylation and/or nuclear transport linked to RNA splicing (21), but can also indicate the presence of transcriptional enhancers within the introns (22).

Most expression vectors for gene vaccination also include intron A from hCMV. Indeed, some studies have revealed that expression and stimulation of an immune response were enhanced by the addition of an intron sequence upstream of the coding region (9,23). This location prevents the utilization of possible cryptic 5' donor splicing sites within DNA sequences (21). These cryptic sites hamper expression due to aberrant splicing.

Chimeric introns can also be constructed with sequences of the donor and acceptor sites of different genes, and the branchpoint site can be optimized to match the consensus sequence leading to an increased expression level (24).

Brinster et al. (25) showed that introns improved transcriptional efficiency 10- to 100-fold in transgenic mice but they observed little effect on gene expression when compared to cells transfected in culture. Various interpretations are possible but the most plausible explanation is that introns contain DNA sequences that are recognized at some stage during development but are not required after transfection, in established cell lines. In conclusion, the studies that examine the role of introns in expression plasmid vectors show that the increase in *in vivo* expression depends not only on the presence or absence of introns, but also on the position of introns within the transcription unit.

*Polyadenylation signal.* Polyadenylation has been shown to enhance RNA stability and translation (26) which in turn vary according to the different transcriptional gene terminators. Moreover, as the rate of transcriptional initiation is increased by the use of strong promoter/enhancer, the process of transcriptional termination may become rate-limiting (18). Although transcriptional terminators are not widely recognized as gene regulatory elements, Hartikka et al. (17) showed that modifications in transcriptional terminator sequences such as their replacement by other types of terminator sequences or the construction of chimeric termination sequences with more efficient polyadenylation signal or the lack of 3' untranslated region (UTR) sequences led to increases in expression which can reflect the differences in transcriptional termination efficiency.

Polyadenylation termination sequences carried by DNA vaccine constructs usually derive from bovine growth hormone (BGH), as in pcDNA3 (Invitrogen, San Diego, CA, USA), and VR1012 (Vical Co., San Diego, CA, USA). Other vectors such as pcDNAI (Invitrogen) and nkCMVintBL (Vical) have late SV40-derived sequence terminators. The late SV40 polyadenylation signal is very efficient and increases the steady-state level of RNA approximately 5-fold more than the early SV40 polyadenylation signal (27). These are located just downstream of the multiple cloning sites. This location facilitates efficient processing of cloned genes which may not have an efficient polyadenylation signal.

*Sequences encoding an antigenic protein.* Recombinant DNA technology has enabled the construction of chimeric genes with an optimized structure for genetic immunization. To be optimized, synthetic genes require elimination of large hairpin structures in the 5'-end UTR mRNAs. It has been shown that these structures reduce the level of *in vitro* and *in vivo* translation in higher eukaryotes (28,29). To initiate translation, an ATG must be present in the inserted gene. This ATG must also be the first one in the 5' region and in the translational start site. Out-
of-frame ATGs can reduce the rate of gene translation (29). Kozak (29) has proposed a consensus sequence for vertebrate mRNAs located around the start site (^9GCCGCA/GCCAUGG^4). She also pointed out that an efficient translation is obtained in the -3 position containing a purine base. In the absence of a purine base, the efficiency of translation can be maintained with a guanine at the +4 position. Genes of prokaryotic organisms and some eukaryotic genes do not have this consensus sequence. Its insertion into the 5'-region of these genes might increase the expression level.

In contrast to prokaryotic sequences, limitations of size and availability often dictate that the eukaryotic sequences to be expressed be obtained in the form of cDNA. To construct a synthetic gene it may be necessary to leave some introns in their place and to excise them during processing of the primary transcript. The presence of introns during processing can sometimes increase the levels of cytoplasmic messenger RNA (30).

The effects of manipulating the encoded sequences by addition or deletion of secretory control signals have also been evaluated, and surprisingly they either had no effect or reduced the immune response (23,31). However, Hoffman et al. (32) reported that in-frame gene fusions with the sequence encoding the leader peptide of human tissue plasminogen activator protein (tPA) enhance their expression.

In order to enhance the immune response to a DNA vaccine, Boyle et al. (33) directed the antigen to sites of immune-response induction by vaccination with DNA encoding antigen-ligand fusion proteins. These two ligands bind to receptors that are present on endothelial venule cells of lymph nodes or on antigen-presenting cells. They showed that both the humoral and cellular immune responses to a model DNA vaccine were enhanced using either antigen-targeting strategy. The construction of optimized synthetic genes combined with regulatory elements produces a transcription unit capable of modulating genetic expression. The exact modulation of the expression level elicits the desirable immune responses.

**Discovery of new antigenic proteins.** Genetic immunization offers an excellent opportunity to quickly discover new antigens and to handle the antigenicity of the protein at the sequence level with the advantage of not requiring protein production and purification. Since it is easy and rapid to clone and modify genes in plasmid expression vectors, many new constructs can be produced and tested in a short period of time such as entire expression libraries which can be cloned and injected in a shot-gun fashion to identify immunoprotective epitopes. A very interesting application of DNA vaccination has been developed to identify new protective antigens. This strategy (termed expression library immunization; ELI) is based on immunization with an expression library constructed with the genomic DNA of the pathogen. It was initially developed by Barry et al. (34) who demonstrated that immunization with partial expression libraries made from genomic DNA of *Mycoplasma pulmonis* protects mice challenged with the pathogen. Similarly, Alberti et al. (35) injected mice with an expression genomic library of *Trypanosoma cruzi*. Although the protection was not assayed, it was possible to detect expression of *T. cruzi* antigens in the muscle and the specific IgG antibodies produced.

The identification of antigens, their combinations and forms is the most effective way to raise protective responses and is one of the most important research areas for DNA-based immunizations.

**Plasmid backbone unit**

*Multiple cloning region, replication origin and prokaryotic selectable marker.* The backbone of plasmid DNA vectors carries a multiple cloning site (MCS) that can be con-
sidered to be additional elements of the transcription unit. This MCS is especially designed to avoid formation of hairpin structures in the 5' end of the transcribed RNA since hairpin structures can interfere with translation.

The replication origin normally used in expression plasmids is the origin of the multicopy plasmid, ColE1. It is the best characterized copy number and incompatibility system and allows the maintenance at steady level of more than 20 copies per *E. coli* cell, resulting in a high DNA plasmid yield, important in gene vaccine production. The preservation of a high number of copies depends on the mechanisms which initiate replication and is controlled by the ColE1 origin (36). DNA vectors can also contain a viral origin of replication. In this regard, SV40 (nkCMVintBL, Vical Co.) and polyoma (pcDNAI, Invitrogen) have been used. Apparently, the inclusion or deletion of such sequences does not affect the expression levels of the foreign peptides (37). These findings, together with the necessity of satisfying concerns related to their potential clinical use, indicate that replication sequences of viral origin should be deleted from DNA vectors.

Selectable markers provide a means to select for growth only in those cells which contain a vector. Such markers are of two types: drug resistant and auxotrophic. Drug resistant markers enable cells to detoxify an exogenously added drug that would otherwise kill the cell. Auxotrophic markers allow cells to synthesize an essential component (usually an amino acid) in a medium which lacks that component. A widespread prokaryotic selectable marker is the beta-lactamase ampicillin resistance gene (ampR) which confers resistance to penicillin-based antibiotics. Penicillin can induce anaphylactic shock in sensitized individuals. Trace amounts of this antibiotic may be present as plasmid DNA contaminants. Replacing the beta-lactamase gene with another antibiotic cassette is possible but does not ease public concern over its potential clinical use. The use of auxotrophic markers can also be possible but there is a high contamination risk with the essential component that is not necessarily recommended for the immunized host. It should be remembered that application of direct gene transfer for genetic immunization requires the availability of plasmid DNA that is free of all contaminants, particularly toxic or immunogenic substances. This contamination problem can be solved if a purification technology combined with selected *E. coli* strains and growth optimization is used for each plasmid construction (38).

**Sequences as immunostimulatory elements.** It has been shown that certain DNA sequences can induce cytokine secretion and lymphocyte activation (39). Certain CpG motifs in bacterial DNA are particularly stimulatory, whereas similar sequences from other species are not. These observations suggest that manipulation of DNA vaccines to contain or to avoid these sequences may affect the immunogenicity of the antigens expressed by the vector. In this regard, Sato et al. (40) found that CMV-based vectors containing a bacterial ampR produced a stronger immune response than a similar expression vector containing the kanamycin resistance gene (kanR). *In vitro* transfection experiments revealed that the ability to stimulate the immune response was not due to the different amounts of antigen expressed by the DNA constructs. The ampR gene contains two copies of the ISS (palindromic CpG hexamer 5' AACGTT 3'), whereas no ISS was identified within the kanR gene. Klinman et al. (41) demonstrated that the elimination of CpG motifs from the plasmid backbone of DNA vaccines reduced vaccine immunogenicity, and that this effect could be reversed by co-administering exogenous CpG-containing DNA. However, the magnitude of the immune response induced by exogenous CpG oligonucleotides never
equaled the effect of incorporating these motifs into the plasmid backbone. These findings indicate that the backbone of the DNA vaccine delivers adjuvant and mitogenic activity via ISS, suggesting that the composition of the DNA vector may be an important consideration in designing a DNA vaccine.

The presence of CpG motifs in DNA vaccines induces B-cell proliferation, immunoglobulin production and cytokine secretion, and promotes the generation of a strong immune response. These beneficial properties of CpG-based adjuvants make them excellent components of the vectors used in DNA vaccines. Optimizing the number and the exact sequences of ISS can greatly enhance the potency of DNA-based vaccination.

Concluding remarks

The structure and organization of the elements influencing expression of DNA-based immunization vectors show a great deal of modularity. It is therefore possible to obtain plasmids with uncommon, although efficient combinations of promoter, enhancer, introns, gene from a pathogen and signal of polyadenylation to undertake specific tasks in the genetic immunization protocol. The adjuvanticity of CpG motifs turns them into important elements to be included in the plasmid backbone.

The molecular biological tools required to test various regulatory modules, to clone ISS, to remove unwanted or safetywise unacceptable sequences, to identify and modify immunoprotective epitopes, are already available. In addition, obtaining high-level gene expression is not a difficult task. However, there is some controversy over the role of high gene expression levels in genetic immunization. The key to the problem depends on the antigen used as well as the type of immune response expected, i.e., cellular, humoral, or both. The exact modulation of the expression level is necessary for each newly tested antigen to elicit the desired immune responses without modifying or shutting down host cell function and causing negative effects similar to those of traditional vaccines.

Now that many important vector elements have been identified, efforts must be concentrated on testing their different combinations and evaluating their immunogenic potential. The study of potential biosafety risks of DNA vaccines such as chromosomal integration and the induction of immunological tolerance must be also undertaken.

Acknowledgments

We thank Dr. Philippe Langella, Dr. Pascale Serror, Yves LeLoir, INRA, Jouy-en-Josas, France and Elisabeth Azevedo, Cultura Inglesa, Belo Horizonte, Brazil for helpful comments and for reviewing the English text.

References

17. Hartikka J, Sawdey M, Cornefert-Jensen

11. Ulmer JB, Donnelly JJ, Parker SE, Rhodes

14. Gurunathan S, Sacks DL, Brown DR,

10. Miyoshi H, Blomer U, Takahashi M, Gage


for direct injection into skeletal muscle.

improved plasmid DNA expression vector

Vahlsing HL, Meek J, Marquet M, Hobart

F, Margalith M, Barnhart K, Nolasco M,

Transplantation,

motor expression levels

of human creatine kinase and CMV pro-

gene via direct injection of plasmid vector

term expression of a fluorescent reporter

in intramuscular injection.

Goldspink G (1998). Myosin regulatory el-

major


