Sustained Transgene Expression Using MAR Elements

Niamh Harraghy, Armelle Gaussin and Nicolas Mermod*

Institute of Biotechnology, University of Lausanne, Switzerland

Abstract: Matrix attachment regions (MARs) are DNA sequences that may be involved in anchoring DNA/chromatin to the nuclear matrix and they have been described in both mammalian and plant species. MARs possess a number of features that facilitate the opening and maintenance of euchromatin. When incorporated into viral or non-viral vectors MARs can increase transgene expression and limit position-effects. They have been used extensively to improve transgene expression and recombinant protein production and promising studies on the potential use of MAR elements for mammalian gene therapy have appeared. These illustrate how MARs may be used to mediate sustained or higher levels of expression of therapeutic genes and/or to reduce the viral vector multiplicity of infection required to achieve consistent expression. More recently, the discovery of potent MAR elements and the development of improved vectors for transgene delivery, notably non-viral episomal vectors, has strengthened interest in their use to mediate expression of therapeutic transgenes. This article will describe the progress made in this field, and it will discuss future directions and issues to be addressed.

Keywords: Matrix attachment region, gene therapy, gene silencing, gene delivery systems, episomal vectors, stem cells, T cells

INTRODUCTION

One of the major problems associated with the majority of current gene therapy protocols is that integration of the transgene into the chromosome is required for its sustained expression. Integration of the transgene into the host chromosome remains in most cases a random event and expression of the transgene is influenced by the chromatin environment in which it is located, a phenomenon termed the position-effect. This often results in two serious negative effects: the transgene becomes silenced with time and/or the transgene inserts within or close to a cellular gene. The latter may result in the dysregulation or inactivation of the affected cellular gene(s), which can have adverse consequences. Transgene silencing can be prevented by DNA sequences that maintain a permissive chromatin structure in their vicinity such as those found on the strong cellular promoter of ubiquitously expressed genes (e.g. housekeeping genes) or on viral long terminal repeats. However, these elements may lead to general effects at the site of vector integration, resulting in the spurious activation of cellular genes such as those encoding oncoproteins. For this reason, there is great interest in finding elements that will protect the transgene from position-effects and/or prevent the transgene from interfering with the regulation of genes in its vicinity. To date at least five such types of elements have been described. These elements, described in Table 1, may act as boundary or insulator elements, shielding their gene from the encroaching heterochromatin (boundary activity) or from the influence of proximal or distal enhancers or repressors (insulator activity, reviewed in [1-3]). Several studies have investigated the use of these elements in shielding a transgene from the chromatin environment, with promising results. In this article we will discuss the potential of using scaffold/matrix attachment regions (S/MAR) for sustained transgene expression in vivo.

MATRIX ATTACHMENT REGIONS

The hypothesis that chromatin is segregated into loops was proposed by Laemmli and co-workers in the 1970s [4, 5]. In this model, these loops are formed from the attachment of the chromatin at specific points to nuclear scaffold proteins [6-10]. Studies of two such attachment sites (X-elements) upstream of the Drosophila melanogaster hsp70 genes revealed that they were A+T rich and suggested that they were associated with regions involved in anchoring the chromatin to the nuclear scaffold, resulting in the formation of chromatin loops [11]. These elements were subsequently termed scaffold or matrix attachment regions (SAR, MAR or S/MAR – to avoid confusion, these elements will be referred to in this article as MAR, regardless of the nomenclature in the original article). The view that euchromatic chromatin is organized as independent loops is now widely accepted [12, 13].

The idea that MARs may act as boundary or insulator elements came from studies that showed that MARs enhanced the expression of a reporter gene following integration in a cellular chromosome in stably transfected cell lines but not in transient transfection assays [14-17], and that transgenic mice carrying multiple copies of the reporter gene flanked by MARs expressed the gene at levels proportional to copy number [18, 19]. The one gene-one MAR hypothesis emerged from the finding that the distribution of MARs is similar to the average gene density [20-22] and it has been proposed that the interaction of MARs with the matrix is gene and cell-type specific [23]. Because of their boundary/insulating features and the association of MAR elements with euchromatin, the possibility that MAR elements could be used to create an independent chromatin domain that could effectively shield transgenes from the propagation of heterochromatin has been explored. Indeed, it is now known that MAR elements can enhance the number of expressing cells by protecting the transgene from silencing (e.g. [24, 25], Fig. 1), although the mechanisms have not been completely elucidated [23].
Table 1. DNA Elements that may be Used to Prevent Transgene Silencing or Increase Transgene Expression

<table>
<thead>
<tr>
<th>Activity/Element *</th>
<th>MAR (Matrix Attachment Region)</th>
<th>Insulator (Barrier or Enhancer-Blocking, e.g. cHS4)</th>
<th>LCR (Locus Control Region)</th>
<th>STAR</th>
<th>UCOE (Ubiquitously acting Chromatin Opening Elements)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binds to nuclear matrix</td>
<td>Yes</td>
<td>No specific association</td>
<td>Only if it contains a MAR</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Chromatin loop formation</td>
<td>Yes</td>
<td>Yes (enhancer-blocking elements)</td>
<td>Yes</td>
<td>Unknown</td>
<td>No</td>
</tr>
<tr>
<td>Insulator activity</td>
<td>Yes (but not all)</td>
<td>Yes</td>
<td>No, but may be exceptions</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Transcription activation</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Enhancer</td>
<td>Generally only after integration, but some reports of increased expression in transient assays</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Replication</td>
<td>Yes (pEPI)</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>No</td>
</tr>
<tr>
<td>TF binding site</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Histone binding site</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unknown</td>
<td>Yes</td>
</tr>
<tr>
<td>Orientation-dependent effect on transgene</td>
<td>Depends on MAR</td>
<td>Position relative to transgene important</td>
<td>Yes</td>
<td>Unknown</td>
<td>Yes</td>
</tr>
<tr>
<td>Position-independent transgene expression</td>
<td>Majority yes</td>
<td>Majority yes</td>
<td>Yes if powerful enough</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Copy number-dependent transgene expression</td>
<td>Majority no</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Unknown</td>
</tr>
<tr>
<td>Tissue-specific</td>
<td>Tissue specificity observed for some</td>
<td>No/unknown</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

* see text for details and references

**Fig. (1).** Representative FACS profile of CHO DG44 cells that were co-transfected with a GFP reporter vector (with or without a MAR element) and a puromycin selection plasmid followed by 2 weeks culture with the antibiotic. A greyscale illustrates the fluorescence of cells that express at low levels, probably as a result from transgene integration in an unfavourable chromosomal environment, while a dark or light green color shows the fluorescence of cells expressing the transgene at moderate or high levels. Modified from Girod et al. [47].
Table 2. Features of the MAR Elements Discussed in This Article

<table>
<thead>
<tr>
<th>MAR</th>
<th>Features</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>EμMAR from the Immunoglobulin heavy chain locus</td>
<td>B-cell specific; increases gene expression and confers copy number and position-independent gene expression. Intronic.</td>
<td>[26-28] and references in text</td>
</tr>
<tr>
<td>huIFN-βMAR from the human interferon-β locus</td>
<td>Some tissue specificity observed; induces high levels of gene expression; does not always confer position-independent or copy number dependent gene expression. MAR located 5’ and 3’ of gene</td>
<td>[29, 30] and references in text</td>
</tr>
<tr>
<td>apoB MAR from the human apolipoprotein B locus</td>
<td>Tissue-specific (hepatic, intestinal, placenta, macrophage); increases gene expression and confers position-independent gene expression. MAR located 5’ and 3’ of gene.</td>
<td>[31, 32]</td>
</tr>
<tr>
<td>Ch-LysMAR from the chicken lysozyme locus</td>
<td>Tissue specific expression, enhances gene expression and confers position-independent and copy number dependent gene expression. MAR located 5’ and 3’ to gene</td>
<td>[18, 33]</td>
</tr>
<tr>
<td>MARβ from the T cell beta receptor locus.</td>
<td>Located 5’ to gene</td>
<td>[34, 35]</td>
</tr>
<tr>
<td>MAR 1-68; X-29</td>
<td>In silico-identified human MARs that potently increase and stabilize transgene expression. Suppressor of variegation and silencing</td>
<td>[16]</td>
</tr>
</tbody>
</table>

MAR elements have a number of features in common and evolutionary conservation is suggested by the fact that MAR elements from one species are functional in another [17], however DNA sequence homology is limited both within and between species [16]. Thus it is not surprising that only a few have been selected for in-depth functional analyses (Table 2) in spite of the estimated tens of thousands MAR elements present in the human genome. MARs may augment the expression of developmentally regulated genes by interacting with enhancers and/or attaching the DNA to the nuclear matrix thereby generating an active chromatin domain. They may also function by insulating genes from adjacent or distal enhancers, and, by recruiting chromatin remodelling proteins, they facilitate the activation, silencing or fine tuning of gene expression. MARs range in size from 300-5000 bp; they possess an A+T rich domain [11, 36-38], which may aid unwinding or destabilization of the DNA duplex [39, 40] or formation of curved DNA structures [41]. MARs also act as binding sites for chromatin remodelling transcription factors such as SATB1, Bright and ARBP [42-44].

 Probably the most interesting feature of MAR elements from a biotechnology and gene therapy point of view is their ability to enhance and maintain transgene expression over extended periods of time. Evans et al. [45] and Girod et al. [16] found however, that several existing algorithms used for identifying MARs [21, 46] did not efficiently predict MAR elements associated with sustained levels of gene expression. Interestingly, Girod et al. [16] discovered that some of the algorithms used for predicting nucleosome-positoning motifs were more accurate at predicting MAR elements correlating with high levels of expression. This led to the design of a new algorithm for predicting MAR elements on the basis of high curvature, major groove depth and minor groove width, a low melting temperature, and the occurrence of potential transcription factor binding sites. Analysis of the human genome resulted in the identification of 50,400 MARs, and a high stringency screen to select for potent elements identified 1566 sequences. Analysis of seven of these elements in vitro showed that six of them could significantly enhance expression of a reporter gene in a stably transfected cell line [16]. The new MAR elements also significantly enhanced expression of the GFP reporter gene as compared to the chicken lysozyme MAR, and one of these was also shown to increase expression of a therapeutic protein from a non-viral vector introduced in mice [16]. The effect of MARs on transgene expression does not appear to depend on the gene length as human MARs were used with success to express utrophin, which is encoded by one of the longest euakaryotic genes known, after in vivo electroporation of a non-viral vector in mice muscles (S. Puttini and N. Mermod, unpublished).

In vitro, MAR elements have been shown to be of great benefit in the production of recombinant proteins. Besides enhancing the expression and maintaining the long-term expression of the transgene, MAR elements were also shown to reduce the variability in gene expression within a polyclonal cell population [47]. Thus, MAR elements could greatly assist in overcoming some of the current obstacles to gene therapy; appropriately or enhanced regulated gene expression, reduced gene silencing and adverse integration effects. In the following sections, studies showing how MARs may help overcome these effects will be discussed.

CURRENT ISSUES WITH GENE THERAPY AND HOW MARs MAY HELP OVERCOME THEM

Gene Silencing

Gene silencing is a natural process and is essential for developmentally regulated gene control. However, it is an event that occurs frequently following gene transfer into mammalian cells and is one of the factors that may contrib-
ute to failure, especially with non-viral vectors. Gene silencing results from co-ordinated changes in the methylation status of the DNA and modifications of histones and other proteins, which ultimately results in the formation of heterochromatin. DNA methylation is a reversible process and by changing the methylation status of the DNA during development, gene expression can be temporally regulated [48-51]. Investigations have confirmed that DNA methylation has a role in stable transgene silencing [52].

Although a systematic study is not available, it has become clear that cells/tissues display a specific pattern of DNA methylation that determines the properties of the cell [53]. Notably, stem or progenitor cells may be less prone to silencing than their differentiated counterparts due to the low levels of DNA methylation found in these cells. However, these cells usually methylate portions of their chromosomes upon differentiation, and they may thus silence integrated transgenes together with their endogenous genes. Thus, the use of genetic elements that may counteract the DNA methylation and gene silencing events that occur during the differentiation process may be of particular benefit for precursor or stem cell-based therapies. However, the extent to which the cellular DNA methylation profile may influence silencing of the transgene in differentiated cells or tissue remains an open question, for integrated as well as for episomal vectors.

The immunoglobulin (Ig) genes are well-studied examples of developmentally regulated gene expression where demethylation plays a role. In most pre-B cells the κ chain genes are unarranged, methylated and transcriptionally inactive whereas in B and plasma cells the κ chain gene is rearranged, hypomethylated and transcriptionally active [54, 55]. The Ig κ gene contains two MARs in its intronic region which flank an enhancer, and they are essential for enhancer activity [56]. A role for MARs in demethylation was suggested from studies that showed that B cells transfected with methylated Ig κ and Ig μ transgenes were only demethylated if the MAR was present [57, 58]. Jenuwein et al. [59] showed that the Ig μ MAR is also required for activation of a distal promoter located about 1 kb upstream of Ig μ gene. Forrester et al. [58] extended the studies of Lichtenstein et al. [57] and Kirillov et al. [60] to show that that activity of the distal promoter was sensitive to DNA methylation and that the MAR element was required for demethylation. Kirillov et al. [60] further showed that the Ig MAR per se is not essential as it could be substituted with other MAR elements, e.g. hIFNβ-MAR and the MAR from the Drosophila histone 1 gene. However, they found that substitution with a MAR element that had poor matrix binding activity failed to induce demethylation in a B cell line, thereby suggesting that attachment of the DNA to the nuclear matrix is essential for demethylation.

Dang et al. [61] investigated the extent to which the hIFNβ-MAR element may prevent or overcome methylation-induced silencing of a transgene.Transient transfection of methylated proviral DNA into 293 cells resulted in reduced gene expression even in the presence of a MAR element. In contrast, in stable transfections, while the methylated control vector did not express detectable amounts of the transgene, expression from the methylated vector containing a MAR element showed reduced (compared to the non-methylated DNA) but detectable expression [61]. Dang et al. [61] hypothesized that although the MAR is unable to overcome fully the methylation-induced repression of the retroviral LTR, the MAR element may have helped maintain an open chromatin structure that was permissive for low levels of gene expression. This study did find however that the MAR element was able to prevent de novo methylation of the retroviral LTR – none of the MAR-containing vectors exhibited evidence of methylation whereas 33% of the control vectors were methylated. Moreover, the MAR-vector was found to maintain a steady level of transgene expression over the 18-week period of the experiment and the fluorescence intensity of the NGFR marker transgene was always higher than that of the control vector. Thus, this study showed that MAR elements can enhance and sustain transgene expression and can also help prevent de novo methylation-induced silencing of a transgene.

Although DNA methylation is an important aspect of gene silencing, it is not however, essential. Kay and co-workers have shown that altering the methylation status of the DNA prior to delivery and deletion of CpG dinucleotides from their vectors had limited effect on methylation-induced transgene silencing [62]. They have found however, that histone modifications are associated with persistence or silencing of transgene expression [63]. Using a bacterial-based episomal vector, they observed a 2-5 fold increase in heterochromatin markers, such as H3K9me2, HDAC2 and HP1α, and reduced gene expression compared to a mimic circle plasmid devoid of bacterial DNA sequences, which had euchromatin markers and displayed sustained elevated transgene expression. Kay and co-workers hypothesized that silencing is dictated by the chromatin structure adopted by the bacterial DNA rather than a specific DNA sequence [62]. Thus the ability of a DNA region to recruit transcription factors and/or chromatin remodelling complexes that favour transcription is essential to maintain euchromatin. In the absence of such sequences or insulators, heterochromatin may propagate.

It has been reported that MAR elements can facilitate the generation of an extended domain of histone acetylation [64]. Chromatin immunoprecipitation (ChIP) assays showed that the VDJ region of the Ig μ locus had about 16-fold more acetylated H4 and 64-fold more acetylated H3 histones as compared to the corresponding region without a MAR. This study also showed that acetylation of the H4 histone was independent of transcription by endogenous RNA polymerases. The authors speculated that the MAR may target the transgene to specific subnuclear regions and that the MARs may enhance the reversible acetylation of a genetic locus by anchoring it to the nuclear matrix, a mode of action for MARs similar to that proposed for demethylation [60].

Another way MAR elements may influence gene expression is that they have been shown to bind a number of transcription factors and chromatin remodelling factors that have the potential to either silence or activate neighbouring genes via remodelling chromatin (reviewed in [23]). For example, the SATB1 transcription factor is known to bind to MARs but can have either a positive or negative effect on gene expression [44, 65]. SATB1 has been shown to organise chro-
Sustained Transgene Expression Using MAR Elements

Current Gene Therapy, 2008, Vol. 8, No. 5

matin into distinct loops by periodic anchoring of the MARs to the nuclear matrix [66]. Such loop formation has been shown to be essential for the generation of a transcriptionally active domain in B lymphocytes [65]. Because of its interactions with other regulatory proteins, SATB1 can recruit other chromatin-remodelling proteins and thereby modulate gene expression by altering the chromatin architecture. For example, the association of SATB1 with the promyelocytic leukaemia (PML) body, which binds to MAR elements, has been shown to modulate expression of a subset of genes at the MHC class I locus [67].

Integration Effects

Integration of the transgene in the host genome has generally been a requisite for sustained transgene expression. However, this limits the efficacy of gene therapies for two reasons. Firstly, expression of the transgene is influenced by its site of integration. This may result in an increase, or more commonly, a decrease in transgene expression or to variegation. This often leads to the use of strong promoter and enhancer elements, and/or to the introduction of multiple transgene copies that will integrate at various chromosomal loci, as allowed by retroviral vectors. The other more serious consequence is thus insertional mutagenesis where the integration of the transgene and its regulatory sequences influences expression of the host’s genes. An example of this is the well-documented X-SCID gene therapy trial, where a combination of the virus integrating near a proto-oncogene and probably also the transgene function gave the transduced cells a selective advantage that led to the development of leukaemia in four of the patients (discussed in [68, 69]).

Variegation, which results in sporadic and reversible transitions between high and low expression states in a population of genetically identical cells, is another problem linked to the integration site. Because multiple copies of the transgene can be integrated at diverse genomic sites, each one subjected to different regulatory constraints, expression of the transgene can vary considerably from cell to cell. On the other hand, multiple copies of the non-viral vectors can co-integrate in a tandem fashion into the same locus, which can result in instability and recombination of the transgene. It also means that there is not necessarily a copy-number dependent increase in gene expression as many vector copies may promote gene expression silencing.

The question of whether MAR elements may shield transgenes from integration effects such as variegation or silencing by forming an independent chromatin domain has been investigated using a minimal promoter from the whey acidic protein gene, which is very sensitive to position [70]. Without a MAR element, only one of 17 transgenic mice lines expressed the transgene. On the other hand, incorporating the chicken lysozyme MAR element in the vector resulted in 7/9 lines expressing the transgene. Bonifer et al. [18] also showed that the chicken lysozyme MAR was able to confer tissue-specific and position independent expression of the chicken lysozyme gene in transgenic mice. The EmMAR from the Ig μ locus has also been shown to provide position-independent expression in B lymphocytes, but not in other cells, demonstrating the cell-line specificity of this element [28]. The apoB MAR also provided position-independent expression of the β-gal transgene in Fao-1 cells [32]. In contrast, work with the huisFNβ-MAR suggested that this element is unable to completely shield the transgene from position-effects, although it does increase gene expression [71-75].

How exactly MAR elements increase gene expression is not entirely clear. One possibility is that they facilitate integration of several copies of plasmid-borne transgenes, while maintaining a transcriptionally favourable domain. Several studies have shown that MAR elements increase the number of copies of the transgene and that this can result in increased transgene expression. For example, Park and Kay [76] found that the Ig κ MAR, but not the chicken lysozyme MAR, enhanced the number of copies of integrated viral genome by ~2.5 fold. The construct flanked by the Ig κ MAR also had increased transgene expression as compared to constructs containing the chicken lysozyme MAR. Dang et al. [61], Thompson et al. [71] and Phi-Van et al. [17] also reported that transgene expression levels were positively correlated with transgene copy number. High expression mediated by a strong human MAR was also associated with an increased transgene copy number co-integrated at unique chromosomal loci, but it was not associated with detectable integration site preference or chromosomal rearrangement [16].

In contrast, Kalos and Fournier showed that clones containing multiple copies of the transgene at the same integration site had very low levels of gene expression, regardless of the presence or absence of the apoB MAR [32]. However, they also reported that clones containing multiple single-copy MAR integrants had enhanced gene expression when compared to the non-MAR containing construct. It is known that multiple copies of transgenes cause heterochromatin formation and transgene silencing in Drosophila [77, 78], possibly because of bidirectional and overlapping transcription. Thus, MAR elements may counteract some but not all modes of transgene silencing, and they may differ in their ability to counteract either type of silencing effects. This may explain why some MARs are capable in mediating copy number-proportional expression whereas others do not.

Viral Vector-Based Gene Delivery

Viral vectors have long been the method of choice for transferring foreign genes to eukaryotic cells. They are attractive because viral transduction is more efficient than non-viral methods, resulting in higher and sustained levels of gene expression compared to plasmid vectors. For example, the transfection efficiency of a plasmid was found to be 7000 times less than that of an adenovirus in HeLa cells [79]. Another advantage of viral vectors is that the multiplicity of infection (MOI) can be reduced so that single or few copy integrants can be obtained. However, the use of viral vectors for gene therapy is not without concern (reviewed in [80]). One of the major issues relating to viral vectors is their need to be integrated into the host genome for maintenance of transgene expression. This increases the risk of insertional mutagenesis and as discussed below may influence host cell gene expression. Other pertinent concerns relating to viral vectors are the fact that they may only transduce mitotically active cells, they may be immunogenic, and/or be susceptible to silencing (e.g. [81-83]).
Hargrove et al. [84] investigated the effect of lentiviral integration on expression of endogenous genes up to 300 kb from the integration site in hematopoietic stem cells. In one instance they found a 13-fold increase of expression of a gene located ~300 kb from the integration site. In other instances, they observed a decrease in expression in genes 175-240 kb from the integration site, whereas some genes closer to the integration site were not affected. These observations are consistent with the models of gene-specific, long-range effects of regulatory elements on gene expression. For instance, the chromatin structure may bring elements that are not linearly close together into contact [85] and viral integration may thus have diverse effects on genes located at considerable distance from the integration site.

Several studies suggested that viral vectors have a preference for integration in the vicinity of MARs [86-88]. For instance, Johnson and Levy [88], Mielke et al. [89] and Rampalli et al. [87] reported that retroviruses have a strong preference for integration in the vicinity of MARs. Johnson and Levy investigated almost 250 retroviral integration sites and found that almost 50% of SL3-3, MLV, HIV-1 and HTLV-1 integrations occurred within 2 kb of a MAR. These authors also found that the majority of MLV and HIV-1 integrations occurred 1-2 kb downstream of a MAR, while HTLV-1 did not show a position preference. Based on an analysis of the region 5 kb upstream and 5 kb downstream of 524 HIV-1 integration sites, Kulkarni et al. [87] found MARs flanking the integration site in 93% of cases.

The MAR SATB1-binding protein has been involved in HIV-1 integration preference in B cells, as the knock down of its expression resulted in more random localization [90]. However, whether MARs per se are the target is unclear as virus integration sites have also been associated with AT richness [91] and transcriptionally active sites [92], which are features commonly associated with MARs. Shera et al. [86] proposed that tumorigenesis resulting from small tumour virus integration may be caused in two ways: in the first instance viral integration may disrupt or activate the expression of a cellular growth-regulating gene; in the second instance the virus may integrate in the vicinity of a MAR element, resulting in the differential expression of the viral genome and/or cellular genes. Rampalli et al. [35] proposed that retroviral vector integration in the vicinity of MARs may help increase transcription from the viral LTR promoter. Goetze et al. [93] reported a strong preference for transgene integration in the vicinity of MARs when transferred by retroviral infection, whereas it was not observed when the same construct was transferred by electroporation. Together, these findings may explain why viral infection is more efficient than non-viral methods in terms of gene expression in eukaryotic cells.

Although the newer types of viral vectors may be less immunogenic, may have enhanced safety resulting from the deletion of enhancer elements, and/or may incorporate MAR or other insulator-like elements to overcome their susceptibility to silencing or to cellular gene activation (e.g. [61, 94]), there is still a need to further develop non-viral delivery protocols for gene therapy, using either plasmid-based or episomal vectors.

**Non-Viral Delivery**

The correction of most genetic defects requires sustained transgene maintenance and expression. However, under certain circumstances sustained gene expression is not desirable. For example, transfection of mesenchymal stem cells with an osteogenic transgene for the purpose of bone regeneration only requires short term expression of the transgene (5-7 days) as the osteogenic protein will recruit cells to the transplanted area (reviewed in [95]). Moreover, because alternative, non-lethal therapies for bone repair exist, the use of current integrating viral vectors may not be justified. Thus under these circumstances non-replicating plasmid vectors may be ideal, as they are maintained transiently in the transduced cell.

Although a defining criterion of MAR elements was their inability to enhance gene expression in transient assays, a few studies have shown that some MARs are capable of increasing gene expression shortly after gene transfer. In transient transfections, Kulkarni et al. [87] found that the IgH MAR and another putative MAR in the p53 promoter could enhance expression of a reporter gene up to 9-12 fold compared to a control. Using a plasmid vector containing the hIFNβ-MAR, Chancham et al. [96] showed that the MAR-containing construct enhanced luciferase expression in CHO cells but not in the neuroblastoma line SKnSH nor in primary neurons or microglia. We have also observed that certain human and mouse MARs can augment transgene expression and viral vector production from transient transfections (unpublished data).

Moreover, some MAR elements have been shown to sustain transgene expression from circular plasmids [16, 97], thus, when transient expression of the transgene is specifically required, using MAR elements in plasmid vectors should be done with caution. Like viruses, plasmids can also integrate into the genome, but as they integrate randomly and are highly susceptible to heterochromatin formation, they are considered to be a lower risk than viruses for insertional mutagenesis.

**Episomal Vectors**

The development of the episomal pEPI vector has been a significant step forward for the non-viral delivery of transgenes to eukaryotic cells. Although it was not the first episomal vector developed for gene therapy purposes, it is the first vector that does not require viral proteins for maintenance. The pEPI vector is based on the inclusion of the hIFNβ-MAR downstream of a promoter [98], where the transcription of MAR is required for episomal maintenance, for reasons that remain unclear. When transfected in CHO cells, the vector can be maintained as an episome for over 100 generations. This contrasts with the control vector lacking the MAR element, whose maintenance requires random integration into the genome. Further work suggested that the MAR element allows the vector to replicate using the cellular machinery and to attach to the nuclear matrix during mitosis, which may help partitioning [99, 100]. After gene transfer, the only essential requirement for maintenance of the episome is that the MAR must be continuously transcribed [99]. Because bacterial sequences in plasmid backbones are known to induce transgene silencing, the pEPI
vector has been refined, and a minicircle created, which consists of the one active transcription unit and a MAR [101]. Minicircles lacking MAR elements are lost during cell-cycling [102], whereas the MAR-containing minicircle may attach to the chromosomal DNA during replication, thereby allowing replication and transfer of the minicircle to daughter cells after cell division. Thus, minicircles containing the MAR element could be ideal for episomal sustained transgene expression in dividing cells.

In summary, there is a wealth of data demonstrating that MAR elements have a positive influence on enhanced and sustained transgene expression, be it for integrating or non-integrating vectors. Indeed the vast majority of studies employing MAR elements for gene therapy purposes (discussed below) showed a clear benefit of MAR elements, with sustained transgene expression for up to one year. The use of MARs and the episomal pEPI vector in gene therapy models is discussed in the following sections.

**USE OF MARS IN GENE THERAPY MODELS**

**Hematopoietic Stem Cells**

Hematopoietic stem cells (HSC) form the basis of the blood cell lineages and there is great interest in targeting these cells for the permanent correction of genetic defects in lineages thereof [68]. However, HSC are difficult to transfect as they mostly exist in a quiescent state and large numbers of HSC are required for successful engraftment [68, 69]. Retroviral vectors are thus preferred for gene delivery. However large MOI have often been required, which increases the chances of multiple integration events and of adverse effects such as neoplastic transformation. Ex vivo expansion of HSC has been performed to increase the number of HSC for engraftment. However, this procedure may enrich the population in cells having particular transgene integration sites and results in decreased engraftment efficiency (e.g. [103, 104]).

The first successful HSC-based gene therapy trials were the X-linked SCID trials. However, these and other studies illustrated the need for improved vector design, to prevent the vector from interfering with the expression of neighbouring cellular genes after integration into the chromosome. In this regard, MAR elements may be useful in increasing expression of the transgene and thus decreasing the MOI required to achieve expression in a significant proportion of the transduced cells as well as potentially protecting the transgene from interfering with its environment.

Several reports have indicated successful use of MAR elements to obtain sustained and/or increased transgene expression in HSC. Murray et al. [105] used retroviral vectors (MoMLV (Moloney murine leukaemia virus) and MSCV (mouse stem cell virus) containing the huIFN β-MAR to transduce HSC, with subsequent engraftment in SCID-hu bone or NOD/SCID mice. The main effect of the MAR element was to increase the average expression of the NGFR reporter gene. When using the MSCV vector backbone, the percentage of total cells expressing NGFR was also doubled.

Agarwal et al. [74] compared the ability of a huIFN β-MAR-based MoMLV vector to reconstitute a SCID-hu thymus liver mouse model with retrovirally transduced human HSC. There did not appear to be a difference in the initial number of cells displaying transgene expression, nor in the engraftment ability of the HSC transduced with or without the MAR. However, they did find that the MAR element enhanced the transgene expression level in human primary cells after transplantation in mice. A similar effect was also observed in large animals, upon the transplantation of HSC cells in baboons, where the MAR increased transgene expression 2 to 9-fold in the lymphocyte, granulocyte and red blood cells lineages over the 12 month duration of the experiment [103].

Taher et al. [106] used a lentiviral vector containing the EqMAR in conjunction with a CD19 promoter to drive B cell specific expression of an EGFP transgene in bone marrow HSC. Subsequent transfer of the transduced HSC to irradiated mice revealed that even after repopulation of B lymphocytes with newly generated cells, expression of EGFP remained constant until the end of the experiment. The main effect of the EqMAR element was to increase the average expression of the transgene.

The evaluation of non-viral gene delivery and maintenance has also been initiated in hematopoietic cells. Papapetrou et al. [72] investigated the use of the pEPI vector in HSC, where efficient transfection of the episomal vector did not require the pre-stimulation of cell growth, unlike viral vector transduction. This study confirmed the episomal and stable maintenance of the pEPI vector in human primary cells during several weeks in the absence of selection. However, the transgene was silenced in murine erythroleukemia (MEL) cell lines despite maintenance of the episomal vector, and this appeared to result from changes in histone deacetylation.

**Primary T Cells**

One of the major prospects for cancer immunotherapy is based on the adoptive transfer of T cells expressing T cell receptors specific for tumour-associated antigens [107, 108]. However, the stability and the level of transgene expression remain limiting factors [109]. T cells naturally cycle through periods of activation and quiescence. *In vivo* the vast majority of T cells are in a quiescent state, and this state is associated with decreased transcriptional activity. This is particularly problematic for T cell receptor (TCR)-based therapies, as a loss of transgene expression during the quiescent phase prevents transduced T cells from being activated by cells displaying the target antigen [109]. A second issue with TCR-based therapies is that the transgene is often poorly expressed. Under such conditions, hybrid TCRs (composed of a chain from the endogenous gene and one from the exogenous transgene) as well as the endogenous TCR may out-compete the transgenic TCR for binding to the CD3 components necessary for display of the TCR on the cell surface, thereby reducing the overall efficiency of the procedure. Therefore a procedure to enhance and maintain transgene expression in both the activated and quiescent T cell states would be of great benefit for T-cell based therapies.

The use of the huIFN β-MAR to overcome these and related problems has been investigated in studies using primary CD4+ and CD8+ T cells. Although the MAR did not significantly alter the expression of the transgene in activated
CD4+ T cells [25, 74], the MAR element significantly increased the number of cells expressing the Lyt2+ marker gene as well as the mean Lyt2+ expression in CD4+ resting cells [74]. A more pronounced effect of the huFNβ-MAR was observed in CD8+ T cells [25, 110], where the MAR element increased the number of Lyt2-positive activated CD8+ T cells by 2-fold and the number of Lyt2-positive resting cells by 10-fold. Expression of the transgene transduced without a MAR element was barely detectable in quiescent CD8+ T cells, with a 10-fold decreased expression level [25, 110], whereas in the presence of the MAR element the differences in expression between the activated and quiescent states was reduced to a 2-fold effect [25]. The huFNβ-MAR was also shown to mediate sustained expression of GFP in CD8+ T cells for at least a month [110].

Besides the use of T cells for TCR expression and immunotherapy, there is also interest in their modification to limit the replication of the human immunodeficiency virus type 1 (HIV-1). The RevM10 transgene has been found to inhibit HIV-1 replication in primary T cells, although high levels of expression are required for the anti-HIV effect [111]. The huFNβ-MAR was shown to enhance the RevM10-mediated anti-HIV effect and to maintain this effect under conditions where the standard retroviral vector failed to inhibit HIV replication, namely at the transition from activated to quiescent state [74].

Neuronal Cells

A number of neurological diseases, such as Alzheimer and Parkinson, are known to have a genetic basis and thus have the potential to be treated by gene therapy. Neuronal cells are difficult to transfect as they are either non-dividing or slow-dividing cells and transduction with the common retroviral vectors has not been very successful. However, the development of HSV-1 amplicon-based vectors for gene therapy of genetic neuronal diseases looks promising [112]. HSV-1 is poorly immunogenic, has a natural tropism for neuronal cells (among others), and with a large genome, much of which can be substituted with foreign DNA, it comes close to being an ideal gene therapy vector [113]. The limitation of the HSV-1 amplicon vector is its life-cycle. After replication in the host cell, the virus enters a latent state, whereby all of its genes (including any foreign genes) except latency-associated transcripts are repressed. To alleviate this problem, episomally-maintained HSV-1 derivatives have been sought by inserting DNA sequences that mediate episomal maintenance. A herpes-related virus, the Epstein-Barr virus (EBV), is known to persist in B cells as an episome. This ability requires a MAR element present in the EBV genome that allows the EBV episome to associate with the nuclear matrix of the host cell [114]. However, no MAR elements have been found in the HSV-1 genome [115]. Two studies have investigated whether MAR elements inserted in the HSV-1 genome may overcome latency-induced transgene repression. Makarova et al. [115] flanked a β-galactosidase reporter gene with a MAR from the chicken lysozyme locus on the 5’ end and a MAR from the human β-globin gene locus on the 3’ end, under the control of a HIV-virus LTR. This expression cassette was inserted into the HSV-1 genome and injected in rat brains. β-galactosidase activity was detected 2 days post-injection. However, by day 14, β-galactosidase activity was no longer detectable, suggesting that the MAR was unable to protect the transgene from latency-induced silencing. A possible explanation for these findings is that neuronal cells may lack transcription or cellular factors required for activity of the MAR (e.g. nuclear matrix binding) chosen for this study, as suggested by Chancham et al. [96], who also failed to see a MAR-induced enhanced expression of a reporter gene in various neuronal cells.

A recent study has made use of the episomal MAR vector pEPI [98] to generate a HSV-1 amplicon vector that is maintained as an episome [116]. Although the majority of the early stage clones did not survive prolonged culture, 3 out of the 10 surviving clones maintained the vector as an episome. Between one and four copies of the episome were estimated to be present in the clones isolated, thus enabling physiologically-relevant levels of the transgene expression. The vector was maintained as an episome for 11 weeks in the absence of selection.

Transgenic Animals

Although MAR elements have shown great potential in cells cultured in vitro, there are fewer studies investigating their potential in vivo. Investigations in transgenic mice, suggest that some but not all MAR elements are able to overcome position- or epigenetic-related silencing events [18, 70, 71, 117]. As some of the MARs were implicated in the developmental control of gene expression, this observation is not surprising given the limited number of MARs and conditions investigated to date. However, one cannot exclude the possibility that promoter and/or vector backbones have also played a role in the silencing of the transgene in these studies.

In light of the safety concerns of viral vectors and the short expression span of plasmid vectors, the development of episomalantly-maintained plasmids for large animal transgenesis and human gene therapy is of great interest. Manzini et al. [118] have published the first report of a large transgenic animal generated from the episomal MAR pEPI vector, which was successfully used to create genetically modified pigs. pEPI-EGFP DNA was detected in 43 tissues of 12/18 fetuses generated. No integration of the pEPI vector into the host genome was detected and the vector copy number per cell was estimated to be <10. On average about 79% of the tissue cells were positive for transgene expression.

FUTURE DIRECTIONS

In Table 3 we have summarized the salient data from the clinically relevant studies discussed in this article. What emerges from the studies so far is that there is a clear benefit in including MAR elements in the vector to increase or prolong transgene expression and/or to allow expression in a larger fraction of the transduced or transfected cells. Whether there is significant benefit in including MARs to make use of other features, which are of interest to gene therapy, such as the ability of a MAR element to confer position independent and copy number-proportional expression on the transgene is less clear. This stems from the use of different MAR elements, promoters, vector backbones and cell lines in these studies, which makes it difficult to generalize results from
<table>
<thead>
<tr>
<th>MAR Element</th>
<th>Promoter</th>
<th>Vector</th>
<th>Cells/Animal Model</th>
<th>Main MAR Effects</th>
<th>Copy Number Dependent Expression?</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>hIFNβ</td>
<td>hsp70.1</td>
<td></td>
<td>transgenic mice</td>
<td>Only enhances expression during implantation. No effect in differentiated cells. Expression enhanced for MAR constructs in fibroblasts from mouse tail and ear. MARs may result in sustained transgene expression.</td>
<td>None in differentiated tissue</td>
<td>[71]</td>
</tr>
<tr>
<td>hIFNβ</td>
<td>MESV MoMLV</td>
<td>Primary T cells</td>
<td>Improved gene expression especially in quiescent cells. No position-independent expression.</td>
<td>ND</td>
<td>[74]</td>
<td></td>
</tr>
<tr>
<td>hIFNβ</td>
<td>LTR MoMLV HIV</td>
<td>Primary CD4+ and CD8+ T cells monocyte-macrophage</td>
<td>Enhances expression of transgene</td>
<td>ND</td>
<td>[25]</td>
<td></td>
</tr>
<tr>
<td>hIFNβ</td>
<td>LTR MoMLV</td>
<td>CEMSS (T cell line) 293 (methylation assay) PG4+</td>
<td>Enhanced and maintained expression of transgene for 4 months</td>
<td>Yes</td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>hIFNβ</td>
<td>MoMLV MSCV</td>
<td>CD34+ HEL (human erythroleukemia cell line) Bone marrow and peripheral blood cells from transduced mice</td>
<td>MARs increased MFI regardless of backbone but MSCV-MAR results in increased number of NGFR+ cells and increased number of transgenic + donor cells</td>
<td>ND</td>
<td>[103]</td>
<td></td>
</tr>
<tr>
<td>hIFNβ</td>
<td>MU3 EF1α CAG</td>
<td>Lentivirus KG1a (human CD34+ hematopoietic progenitor line) CD34+ mononuclear cells from human umbilical cord</td>
<td>Expression increased in all cell lines and with all promoters. MAR-chs4 construct best at enhancing and maintaining transgene expression.</td>
<td>No</td>
<td>[119]</td>
<td></td>
</tr>
<tr>
<td>hIFNβ</td>
<td>pgk β-actin ubiquitin LTR SV40 CMV</td>
<td>MoMLV</td>
<td>Primary CD8+ cells</td>
<td>MAR increased transgene expression in activated and resting T cells. MAR worked best with viral LTR.</td>
<td>ND</td>
<td>[107]</td>
</tr>
<tr>
<td>hIFNβ</td>
<td>CMV</td>
<td>HSV-1 CHO ldlr⁻⁻⁻</td>
<td>Facilitated and maintained complementation of the cell line</td>
<td>ND</td>
<td>[113]</td>
<td></td>
</tr>
<tr>
<td>Ch-LysMAR β-globin</td>
<td>HIV LTR</td>
<td>HSV-1</td>
<td>Rat brain</td>
<td>MARs failed to prevent silencing following induction of latency</td>
<td>ND</td>
<td>[112]</td>
</tr>
<tr>
<td>Ch-LysMAR</td>
<td>AAT RSV pHM5 (minimal bacterial DNA) pBluescript (pBS)</td>
<td>murine liver</td>
<td>Constructs with MAR: sustained expression for over 1 year with AAT-pHM5; for 2 weeks with RSV-pHM5; for 6 days with pBS backbone</td>
<td>ND</td>
<td>[96]</td>
<td></td>
</tr>
<tr>
<td>Ch-LysMAR IgκMAR</td>
<td>CMV Lentivirus HEP2 Huh7 Mouse</td>
<td>Ch-Lys MAR decreased expression whereas Igκ MAR enhanced expression and enhanced the number of vector genomes</td>
<td>Yes</td>
<td>[76]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
different laboratories, as all these factors are known to influence transgene expression [28, 30, 97, 110].

In addition, several issues remain to be addressed before clinical use may be envisaged, the main one being the potential ability of MARs to mediate a permissive chromatin structure around their integration locus in the cellular chromosome (Fig. 2). For instance, we have recently observed that a MAR element is capable of reversing prior transgene silencing in transfected cells in culture (Galbete et al., submitted for publication). Thus, while MARs may be used to increase and stabilize the production of retroviral vectors when introduced outside of the sequences coding for the viral RNA that is packaged (Buceta et al., unpublished results), inclusion between the viral LTRs remains limited by their size and by safety considerations. Given the known tropism of retroviral and lentiviral vectors for genes and regulatory sequences, the extent of this phenomenon needs to be characterized further, to exclude the possibility that viral vector integration at a silent locus may reconfigure the chromatin structure of adjacent cellular genes and thereby activate their expression. Episomal vectors are certainly favourable in this respect, in cases where transgene expression can be stably maintained from the episomal DNA, but the recombination of the vectors with chromosomes cannot be fully excluded in the long term and potential effects of such events should be assessed.

In this respect, identifying sequences that would act both as a boundary, protecting transgene expression from silencing, and as an insulator, decreasing cellular gene activation by vector components, would certainly be of great interest. Thus, the potential action of MARs as enhancer-blocking insulators needs to be studied. While such insulator function has been shown to be separable from the activities that may prevent the propagation of gene silencing signals, the two types of activities may be co-localized at some natural epigenetic regulators. Thus, identifying elements that would act both to prevent the silencing of the therapeutic gene in cases of integration into a heterochromatic locus, while concomitantly preventing the activation of cellular genes nearby is a prized objective. It would contribute to safer viral vectors both by decreasing the risks of spurious insertional activation of cellular proto-oncogenes and by allowing the use of lower multiplicity of infections and chromosomal integration events.

Many of the viral vectors used for gene therapy have been reported to have an affinity for integration in the vicinity of genes and/or regulatory sequences such as MAR elements [86-88], and they display an orientation and position preference relative to the adjacent MAR [88]. So far, whether the integration preference of such vectors may be affected by inclusion of a MAR remains largely unknown. Few of the studies performed with MAR elements to date have identified the vector sites of integration. With the advent of linear amplification-mediated (LAM)-PCR and highly productive sequencing techniques, identifying the integration site preferences of vectors is now possible. Correlating the integration site with the level of gene expression would provide information on the extent to which the position effect influences MAR activity and transgene expression. For a true comparative assessment of MAR activity, different MAR elements need to be integrated at the same locus, preferably in a single copy. Until recently, this was difficult to achieve, however the development of systems such as the φC31 integrase and recombinase-mediated cassette exchange (RMCE) makes it possible and more reliable [75, 119].

New generations of MAR-containing vectors for gene therapy need to take these and additional parameters into account. For example, some MARs can interact specifically with cognate promoters [35, 97] and the action of some of the MARs displays cell-type specificity [19, 96]. The position of the MAR relative to the transgene may also influence their mode of action and relative potency [30, 74, 76, 97]. If the one-gene-one-MAR hypothesis holds true, then it may be
possible to select temporal- and cell-specific MARs to suit particular transgenes and indications.

Because MAR elements may not possess all the desired features (e.g., insulating activity, promoter and cell type specificity), one promising approach may be to combine a MAR with elements such as those described in Table 1, such as derivatives of the chicken cHS4 insulator [83]. Although the cHS4 insulator is capable of protecting against position effects and has an enhancer-blocking function, it fails to prevent silencing of oncoretroviral vectors in murine embryonic stem cells [120] and thus needs to be used in conjunction with an anti-silencer element, such as a MAR. This approach was taken by Ramezani et al., who constructed a lentiviral vector comprising of the huIFNβ-MAR and/or the cHS4 insulator, and used this vector to transduce hematopoietic and primary CD34+ cord blood cells [121]. While the cHS4 insulator was unable to enhance gene expression, the huIFNβ-MAR alone and the combined huIFNβ-MAR cHS4 vectors showed increased gene expression. The presence of the huIFNβ-MAR in the vector also improved the homogeneity of expression of the cell population, particularly when...
included with the eHS4 insulator. It will be interesting to assess whether the insulator may confine the MAR effect to the transgene, resulting in the formation of a truly independent chromatin domain that would both be shielded from position-associated effects and that would not affect the expression of other genes at its integration locus.

In conclusion, MAR elements possess the regulatory features required to overcome many of the current problems facing gene therapy in terms of transgene silencing and/or integration, and their use may be considered for approaches involving ex vivo gene transfer and cell transplantation. Incorporating MAR elements in conjunction with other insulator/boundary elements has the potential to greatly improve and enhance the function and safety of the current gene therapy vectors, but more work will be required to characterize the function and interactions of the various regulatory elements driving transgene expression. In any case, it should not be assumed that there will be a “one-MAR-for-all” solution, and more research is also required to further characterize MAR elements, to gain a better understanding of their regulatory features, and how these can be used to improve current gene therapy vectors.

ACKNOWLEDGEMENTS

The authors thank Antonio Rosato for helpful discussions. Relevant experimental work performed in the authors’ laboratory is supported by grants from the CTI Swiss governmental agency, Selenix SA, the Swiss Foundation for Research on Muscular Diseases, and the CLINIGENE European Network of Excellence.

Authors declare having potential competing financial interests.

REFERENCE

[33] Lec PV, Stratling WH. The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. EMBO J 1988; 7: 655-64.
[34] Chattopadhyay S, Whitehurst CE, Chen J. A nuclear matrix attachment region upstream of the T cell receptor beta gene enhancer.


[102] Chen ZY, Yant SR, He CY, Meuse L, Shen S, Kay MA. Linear DNAs concatenate in vivo and result in sustained transgene expression in mouse liver. Mol Ther 2001; 3: 403-10.


