Effects of Insulators in Mammalian Vector Constructs

Anirudh Rao, Anirudh T P, Sidharthan S C

Gene Insulators and their Types

Insulators are frequently used by mammalian cell engineers to mitigate the effects of transgene silencing. These sequences are included in expression vectors to prevent nearby heterochromatin regions from spreading to the locus of integration or to prevent regulatory elements from interacting with the transgene cassette [1]. Various types of insulators have been discovered and used in mammalian cells [2].

Chicken hypersensitive site 4 (cHS4)

One of the best characterized insulators, this is a 1.2 kb sequence obtained from the 5' end of the chicken β-globin locus. When tested in the human erythroleukemia cell line K562, it was found to insulate a transgenic γ-neomycin reporter gene from being activated by a nearby regulatory region. The mechanism of its action is based on preventing regulatory regions from accessing the promoter by altering the local structure of the chromatin. However, due to the large size of this insulator, the size of the transgene that can be incorporated in the expression vector is heavily limited [3].

Scaffold/matrix attachment regions (S/MARs)

Chromatin in the nucleus is attached to a proteinaceous matrix. The points at which DNA is attached to this matrix are termed as scaffold or matrix attachment regions (S/MARs). These sites have been shown to stabilise transgene expression in engineered mammalian cells. The mechanism of their action appears to be the combined recruitment of transcription factors and chromatin remodelling complexes to promote gene expression and regulate chromatin structure respectively. One of the best known is the S/MAR located 5' to the chicken lysozyme gene [4]. These insulators also face the issue of large size, limiting the carrying capacity of the vector [2].

Ubiquitous chromatin opening elements (UCOEs)

These DNA elements are unmethylated CpG islands that are found adjacent to the promoters of housekeeping genes. They promote histone hyperacetylation, ensuring that the chromatin is accessible to the transcription machinery. Thus, they permit sustained transgene expression. The best studied UCOE is A2-UCOE, a 1.5 kb sequence found near the locus HNRPA2B1-CBX3 in humans. However, these insulator elements show bidirectional activity, which may cause dysregulation of nearby genes [5]. Different truncations of UCOEs may be helpful in overcoming these bidirectional effects [6].

Human tDNA

These are genes encoding for transfer RNA (tRNA) in humans. They have been shown to have insulator activity in human cells. Although the details of their mechanism are not completely understood, they are frequently found at heterochromatin-euchromatin junctions, pointing to their insulator activity. They also show long-range interactions with other tDNA sequences in the genome, forming distinct chromatin compartments in the nucleus that are insulated from regulatory elements in other parts of the genome [7].

CTCF insulator elements

The CCCTC-binding factor (CTCF) is a transcription factor found in eukaryotes. Binding of this protein to specific sites has been shown to block the effect of regulatory elements. These binding sites are usually found at the junctions of heterochromatin and euchromatin. Similar to tDNA, CTCF-binding sites also show long-range interactions, forming insulated compartments in the nucleus [8][9]. Interestingly, the cHS4 insulator also contains CTCF-binding sites. This 250 bp region is called the cHS4 core element and shows insulator activity independent of the remaining regions in cHS4 [10].

Transposon-based insulators

Mammalian-wide interspersed repeats (MIRs) are a family of transposable elements that have been shown to have insulator activity in human cells. They show sequence similarity to tDNA and are free of CTCF-binding sites. They recruit RNA

polymerase and chromatin remodelling enzymes and thus, provide an insulated environment for the expression of flanking genes [11]. Short interspersed nuclear elements (SINEs) are also transposons found in mammalian genomes that are known to function as insulators. They form heterochromatin barriers adjacent to promoters and serve as binding sites for RNA polymerase. An example of such an insulating element is B1-X35S, found in the mouse genome [12].

Effect of Promoter Type on Insulator Activity

Different promoters inherently show different transcriptional activities and susceptibilities to silencing [2]. The functions of insulators is also dependent on the genomic context they are present in [13]. Thus, it is important to understand how insulator activity changes when the promoter of the transgene is altered.

The cHS4 insulator has been found to prevent transgene silencing when used in a double promoter vector with a murine stem cell virus U3 (MSCV-U3) promoter and an elongation factor 1α (EF1α) promoter. However, the transgene expression in HeLa cells was higher when MSCV-U3 was proximal to the insulator than when EF1α was. It was hypothesized that this phenomenon was independent of the CTCF-binding sites present in cHS4, i.e., the core element [14]. This insulator can also facilitate sustained transgene expression when the silencing-prone Rous sarcoma virus (RSV) promoter is used in transposon-based vectors [15].

S/MAR-based insulators have low functionality when used with a cytomegalovirus (CMV) promoter. However, it shows good activity when promoters specific to the cell line are used. For example, using the liver-specific α-fetoprotein (AFP) promoter in the human hepatocellular carcinoma cell line and the muscle-specific smooth muscle 22 (SM22) promoter in a myocyte cell line allowed S/MAR-mediated prevention of transgene silencing. It has also been shown that S/MAR insulators can be used with inducible promoters like the tetracycline responsive element (TRE). In the human hepatoma and human pancreatic carcinoma cell lines, the ubiquitin C (Ubc) promoter showed stable transgene expression when used with a S/MAR insulator [16]. In human mammary carcinoma cell lines, a natural S/MAR insulator was found to prevent the silencing of the epithelial cadherin (E-cadherin) promoter, and thus, prevent epithelial-mesenchymal transition (EMT), a hallmark of cancer metastasis. The mechanism of this insulation is the recruitment of the chromatin remodelling matrix attachment region binding protein SMAR1 [17].

The 1.5 kb human RNP UCOE and the 3.2 kb mouse RPS3 UCOE show different insulating activities when used with different promoters. Surprisingly, CMV promoters from different sources have drastically varying effects on UCOE-based insulation. The guinea pig CMV (gCMV) and human CMV (hCMV) promoters show low transgene expression while the mouse CMV (mCMV) promoter shows high expression. The Ubc and SRa promoters also achieve high expression. The myeloproliferative sarcoma virus (MSV) promoter shows similar expression to that of hCMV [18].

tDNA insulators are naturally associated with the B-box promoter, which is known to recruit the transcription factor TFIIIC and aid in insulation. tDNA has also been shown to have insulating activity when used with the human γ-globin promoter, but not as robustly as a cHS4 insulator. Using a cluster of tDNA insulators with the γ-globin promoter can achieve robust insulation. tDNA fragments can also act as insulators when used with a CMV promoter. They also recruit RNA polymerase when placed adjacent to the arachidonate lipoxygenase (ALOXE3) promoter [7].

CTCF-binding sites have been implicated in the activation of the insulin-like growth factor 2 (Igf2) promoter by preventing the interaction of a neighbouring control region with the promoter [19]. They also play a role in the DNA looping-mediated expression of genes associated with the human protocadherin (hPcdh) promoter [20].

Due to their similarity with tDNA, MIR insulators function well when associated with B-box promoters. When a somite promoter is used, some MIR sequences show insulator activity while some do not. This indicates that MIR-mediated insulation is promoter-sensitive. When used with a ZAP70 promoter, MIR insulators effectively block the spreading of heterochromatin. When tested in T-cells, promoters located closer to an MIR sequence were protected from silencing better than those located distally [11].

The B1-X35S SINE insulator shows strong activity when used with a cardiac actin promoter in zebrafish. It activates transcription in Pol II and Pol III promoters. It has a similar insulating potential to cHS4 when used with a CMV promoter in HEK293 cells [21].

Design of a Novel Insulator-Based Construct

Human embryonic kidney 293 (HEK293) cells are a promising cell line for the production of monoclonal antibodies [22]. However, transgene silencing remains an issue in this cell line as well [2].

Through our review, we found the S/MAR insulators allow for high levels of transgene expression when tissue-specific promoters are used. Thus, one would expect that using a kidney-specific promoter with a S/MAR insulator in HEK293 cells would enable stable expression of heterologous proteins. An example of such a promoter is the kidney-specific cadherin (KSPC) promoter [23]. This is a good choice since we already know that S/MAR insulators have the required functionality when used with the epithelial cadherin promoter [17].

Another desirable feature for an insulator is small size. Large sized insulators reduce the allowed size of the transgene that can be carried by the expression vector. This is especially relevant in the case of antibody manufacturing, where the transgene size is likely to be quite large considering the size of an antibody chain.

The mechanism of action of the S/MAR insulator in the E-cadherin case is the recruitment of the SMAR1 chromatin remodelling protein [17]. Insulation by SMAR1 recruitment has also been reported in the miR371-373 promoter. Here, it binds to a short 200 bp sequence that contains T(C/G) repeats [24]. Using the human E-cadherin sequence (GenBank accession L34545.1), we found that the same repeats are present at the 5' end. Thus, this 200 bp SMAR1 binding site likely functions as a S/MAR insulator. Due to its small size, it is ideal for use in an expression vector.

Based on these findings, we propose a novel expression vector for use in HEK293 cells. Along with the standard vector components like an ori site and selectable markers for bacteria and mammalian cells, our vector will contain 200 bp S/MAR insulators on either end of the transgene cassette, and a KSPC promoter at the 5' end. A multiple cloning site will be present between the promoter and the downstream insulator. This cloning site will be able to accommodate a large insert.

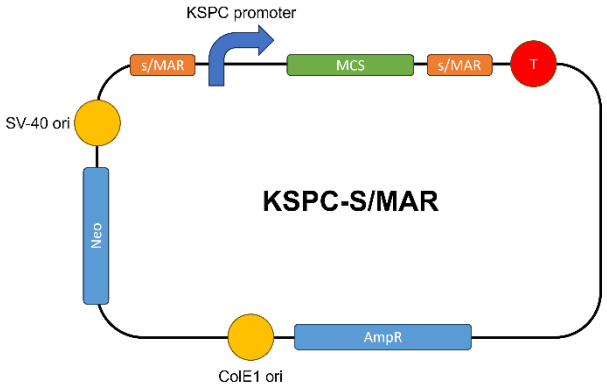
To test the efficacy of our strategy to overcome transgene silencing in HEK293 cells, we can use a reporter protein like GFP or mCherry. To ensure that our chosen promoter and insulator significantly reduce transgene silencing, we will compare the reporter expression relative to control vectors. Thus, we will have to construct the following vectors:

- **Experimental**: KSPC + S/MAR
- **Control 1**: KSPC + No insulator
- **Control 2**: CMV + S/MAR
- **Control 3**: CMV + No insulator

These vectors will be transfected separately into HEK293 cells using lipofectamine. After sufficient incubation, the reporter expression can be assayed by flow cytometry.

The average fluorescence intensities for each of the four engineered HEK293 cell lines can be computed using the flow cytometry data.

We expect that Control 1, which lacks an insulator, will have lower reporter expression than the experimental due to the fact that transgene silencing is not prevented. Control 2, despite having an insulator, will also have lower expression that the experimental as the CMV promoter in this control is not tissue-specific. Our analysis of existing literature revealed that S/MAR insulators work best when paired with a tissue-specific promoter. This can be further validated with Control 3, which is expected to have lower expression than Control 1 due to the change in promoter. Thus, our novel vector can be expected to have the highest expression among all the vectors being tested.



Part	Function
ColE1 ori	*E. coli* origin of replication
AmpR	Ampicillin resistance promoter and gene for bacterial selection
SV-40 ori	Mammalian origin of replication
Neo	Neomycin resistance promoter and gene for mammalian selection
s/MAR	200 bp S/MAR insulator
KSPC promoter	Kidney-specific cadherin promoter
MCS	Multiple cloning site for transgene insert
T	Polyadenylation signal and transcription terminator

References

[1] Emery, D. W. (2011). The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. Human Gene Therapy, 22(6), 761–774. https://doi.org/10.1089/hum.2010.233

[2] Cabrera, A., et al. (2022). The sound of silence: Transgene silencing in mammalian cell engineering. Cell Systems, 13(12), 950–973. https://doi.org/10.1016/j.cels.2022.11.005

[3] Chung, J. H., et al. (1993). A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell, 74(3), 505–514. https://doi.org/10.1016/0092-8674(93)80052-g

[4] Girod, P., et al. (2007). Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. Nature Methods, 4(9), 747–753. https://doi.org/10.1038/nmeth1076

[5] Müller-Kuller, U., et al. (2015). A minimal ubiquitous chromatin opening element (UCOE) effectively prevents silencing of juxtaposed heterologous promoters by epigenetic remodeling in multipotent and pluripotent stem cells. Nucleic Acids Research, 43(3), 1577–1592. https://doi.org/10.1093/nar/gkv019

[6] Rudina, S. S., & Smolke, C. D. (2019). A novel Chromatin-Opening element for stable long-term transgene expression. bioRxiv. https://doi.org/10.1101/626713

[7] Raab, J. R., et al. (2011). Human tRNA genes function as chromatin insulators. The EMBO Journal, 31(2), 330–350. https://doi.org/10.1038/emboj.2011.406

[8] Kim, S., et al. (2015). CTCF as a multifunctional protein in genome regulation and gene expression. Experimental & Molecular Medicine, 47(6), e166. https://doi.org/10.1038/emm.2015.33

[9] Liu, M., et al. (2015). Genomic discovery of potent chromatin insulators for human gene therapy. Nature Biotechnology, 33(2), 198–203. https://doi.org/10.1038/nbt.3062

[10] Arumugam, P., et al. (2009). A novel combination of chicken hypersensitive Site-4 insulator elements improves titers and restores full insulator activity. Blood, 114(22), 3566. https://doi.org/10.1182/blood.v114.22.3566.3566

[11] Wang, J., et al. (2015). MIR retrotransposon sequences provide insulators to the human genome. Proceedings of the National Academy of Sciences, 112(32). https://doi.org/10.1073/pnas.1507253112

[12] Román, A. C., et al. (2011). B1-SINE retrotransposons. Mobile Genetic Elements, 1(1), 66–70. https://doi.org/10.4161/mge.1.1.15455

[13] Ribeiro-Dos-Santos, et al. (2022). Genomic context sensitivity of insulator function. Genome Research, 32(3), 425–436. https://doi.org/10.1101/gr.276449.121

[14] Uchida, N., et al. (2013). The chicken hypersensitivity Site 4 core insulator blocks promoter interference in lentiviral vectors. Human Gene Therapy Methods, 24(2), 117–124. https://doi.org/10.1089/hgtb.2012.152

[15] Sharma, N., et al. (2012). The impact of CHS4 insulators on DNA transposon vector mobilization and silencing in retinal pigment epithelium cells. PLoS ONE, 7(10), e48421. https://doi.org/10.1371/journal.pone.0048421

[16] Hagedorn, C., & Lipps, H. J. (2012). S/MAR vectors – Alternative expression systems for gene therapy? Gene Therapy and Regulation, 07(01), 1230003. https://doi.org/10.1142/s156855861230003x

[17] Adhikary, A., et al. (2014). Inhibition of epithelial to mesenchymal transition by e-cadherin up-regulation via repression of slug transcription and inhibition of e-cadherin degradation. Journal of Biological Chemistry, 289(37), 25431–25444. https://doi.org/10.1074/jbc.m113.527267

[18] Nair, A. R., et al. (2011). Effect of different UCOE-promoter combinations in creation of engineered cell lines for the production of Factor VIII. BMC Research Notes, 4(1). https://doi.org/10.1186/1756-0500-4-178

[19] Herold, M., et al. (2012). CTCF: insights into insulator function during development. Development, 139(6), 1045–1057. https://doi.org/10.1242/dev.065268

[20] Guo, Y., et al. (2012). CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. Proceedings of the National Academy of Sciences, 109(51), 21081–21086. https://doi.org/10.1073/pnas.1219280110

[21] Román, A. C., et al. (2011). Dioxin receptor and SLUG transcription factors regulate the insulator activity of B1 SINE retrotransposons via an RNA polymerase switch. Genome Research, 21(3), 422–432. https://doi.org/10.1101/gr.111203.110

[22] DeLuca, K. F., et al. (2022). Production and purification of recombinant monoclonal antibodies from human cells based on a primary sequence. STAR Protocols, 3(4), 101915. https://doi.org/10.1016/j.xpro.2022.101915

[23] Asico, L. D., et al. (2018). Nephron segment-specific gene expression using AAV vectors. Biochemical and Biophysical Research Communications, 497(1), 19–24. https://doi.org/10.1016/j.bbrc.2018.01.169

[24] Mathai, J., et al. (2016). SMAR1 binds to T(C/G) repeat and inhibits tumor progression by regulating miR-371-373 cluster. Scientific Reports, 6(1). https://doi.org/10.1038/srep33779