Improved Expression Vector Activity Using Insulators and Scaffold/Matrix-Attachment Regions

for Enhancing Recombinant Protein Production

by Helen Y. Kim

With recombinant proteins continuing to emerge as an important class of human therapeutics, improved methods to generate cell lines expressing high levels of desired proteins are becoming increasingly critical to the industry. High-expressing cell lines are important for production of clinical candidate molecules as well as in rapid and reliable production for characterization and validation studies. In either case, the method should be rapid, cost-effective, and scalable.

Specifically, a method that generates high-expressing cell clones without an extensive selection and screening step would be useful for production of clinical candidates. A method to rapidly and reliably generate proteins in hundreds of milligrams would be equally valuable for early phase validation studies. For such production, using a pool or a collection of transfected cells rather than a single cell clone can significantly reduce necessary resources and time. But widely variable expression levels in different cell clones from a transfection necessitates extensive single-cell cloning and screening, which prevents the use of transfected cell pools for even small-scale production. Furthermore, expression levels typically decrease with culture time. Such instability of pool expression is probably due to relatively low levels of recombinant protein expression by most cells within a given transfection — and subsequently selective survival of those low-expressing cells, which often display growth advantages over their high-expressing counterparts.

AYAAZ RATTANSI (WWW.ISTOCKPHOTO.COM)

Reasons underlying the large variability in clonal expression levels include differing plasmid copy numbers and a phenomenon known as the *position effect*, which was initially described in *Drosophila melanogaster* as position-effect variegation (**1**). Such effects come from surrounding DNA that affects a recombinant gene's expression upon integration into the host genome. They include activities of external enhancers and silencers as well as heterochromatinization.

In most cell line development processes, expression constructs are introduced into a host cell's genome using methods for random integration. Hence, the level of transgene expression depends on where that DNA integrates. So it's not surprising that a large majority of cell clones do not express high levels of transgene, considering that most genomic sites are transcriptionally repressive (**2, 3**). Such repressive effects can spread and cause epigenetic silencing of adjacent genes. Transcriptional repression can occur by histone deacetylation (**4**) and methylation, at lysine 9 of histone H3 (H3-K9) (**5**), and by methylation of the transfected DNA's promoter sequence.

Those events can be regulated by local availability of heterochromatin-associated components such as heterochromatin-1 (HP1) (**6–8**). Other regulators are histone-H3 methyltransferases, such as *suppressor of variegation 3–9* in *D. melanogaster,* also known as Suv39h1 and Suv39h2 in the mouse (**9**). HP1 proteins are multidomain proteins with several binding partners that can function as structural adaptors for the assembly of macromolecular complexes in chromatin, leading to heterochromatin assembly and maintenance (**8**). HP1-interacting partners include the DNA methyltransferases Dnmt1 and

Dnmt3a, which are involved in CpG methylation (**10**).

Use of Insulators and S/MARs Can Reduce Effects of Heterochromatin on Transgene Expression: Two common approaches can be used to protect DNA from negative position effects or integration-dependent repression. One approach is to direct transgene integration into a predetermined site that is transcriptionally active using site-specific recombination methods. Another method is to simply incorporate into the expression vector DNA sequence elements found in chromatin border regions, such that regardless of the integration site the gene will be protected from surrounding chromatin influences.

Here I focus on the latter method, summarizing some recent developments in vector engineering that have helped us overcome the negative position effects of randomly integrated DNAs. This method takes advantage of nature's solution for protecting transcriptionally active regions from epigenetic heterochromatinization. To do so, cells have evolved barrier mechanisms using specialized DNA sequences (known as insulator or chromatin barrier elements) that establish chromatin borders.

For recombinant protein expression, sequences that behave as chromatin borders and protect transfected genes from surrounding chromatin influences include insulator sequences and scaffold/matrix–attachment regions (S/MAR). Expression studies at my laboratory and several others have shown that flanking a transgene with insulators or S/MARs can suppress the clonal expression variability (**3, 11**). Generating a higher proportion of transfected cells with improved expression levels can reduce the number of clones that need to be screened to identify an acceptable production cell line. That further enables generation of high-expressing transfected pools with improved stability. It also can eliminate the need to isolate clones for production of research material used in early characterization and validation studies.

DNA Insulator Elements

DNA insulator elements were initially identified as sequences in *Drosophila* that prevent external enhancers from inappropriately activating reporter gene promoters (**12–14**). These are DNA sequences that, in their natural state, are thought to define distinct chromatin domains of gene expression. They prevent cross-regulation of adjacent genes or gene clusters by restricting the activity of DNA elements such as enhancers and silencers to the domain where they are located. Several such DNA elements have been identified to date from various species, including *Saccharomyces cerevisiae, Drosophila melanogaster, Xenopus lavis,* sea urchin, chicken, and human. All are found between independently regulated gene loci.

Many insulator elements have been shown to protect recombinant genes from position effects in vivo across species, as well as in vitro in mammalian cells lines. That suggests that these elements have a conserved role in defining domains of gene expression (**15**). Consistent with a function in defining domains of gene expression, insulator elements are found throughout the eukaryotic genomes (**16**). Subsequent analyses have revealed a wide diversity of insulator sequences and suggested that many are compound elements, containing several distinct protein binding sites and separable properties (**17**). A 1.2-kb DNA sequence element (5´HS4) at the 5´ end of the chicken beta-globin locus, for example, can be separated to enhance blocking and barrier activities.

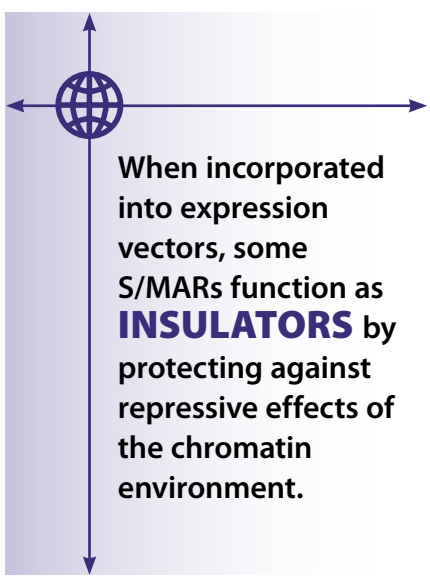
Results from such experiments have helped to define two main functions for insulators: their ability to block enhancer-promoter activity (enhancer blocking activity), and their ability to prevent the spread of heterochromatin (barrier activity). Experiments that delineate distinct activities of insulator elements may help reduce the size of effective elements used in expression vectors, making the final size of expression plasmids easier to manage. Furthermore, potential expression benefits would be obvious if expression vectors could be constructed with sequences that possess only the barrier activity, without enhancer blocking.

Scaffold/Matrix-Attachment Regions

The scaffold/matrix–attachment regions (S/MARs) are experimentally identified sequences associated with the nuclear scaffold or matrix. They are believed to be responsible for attachment of chromatin loops to the nuclear scaffold or matrix (**18, 19**). These sequence elements are further thought to be involved in chromatin remodeling and subsequent transcriptional activation and also protection of transgenes from position effects (**20, 21**).

As sequence diversity is found in insulators, S/MAR sequences also tend to be heterogeneous. However, they have been analyzed more extensively. Identification of frequently occurring sequence motifs found in large numbers of experimentally characterized S/MARs has led to the development of several computational tools to predict yet unidentified S/MAR elements (**23–26**). Like insulators, S/MARs can be identified on the basis of their position in the genomes of different species, which suggests evolutionary conservation (**27–29**).

Analyses of experimentally identified S/MARs have revealed a typical element to be as short as 300 base pairs and up to several kbs long. These S/MARs may contain several



sequence motifs, including AT-rich nucleotide motifs. They serve as base-unpairing regions (BURs), which are thought to function as DNA-unwinding elements (**30**). Other sequences include (**28, 31–33**)

- kinked DNA generated by the presence of TG, CA, or TA dinucleotides separated from each other by 2–4 or 9–12 nucleotides
- potential replication origins (ORI) and homeotic protein-binding sites
- intrinsically curved DNA produced by the $(A)_n$, T_mA_n
- transcription factor binding sites
- triple-helical or H-DNA structure sequences
- retroelement insertion hot spots.

Usually, S/MARs contain about 70% AT. They are enriched with all motifs listed above, but every S/MAR does not necessary contain all of them. When incorporated into expression vectors, some S/MARs can function as insulators by protecting against the repressive effects of the neighboring chromatin environment, albeit to different extents (**3, 34**). That is not surprising because insulators and S/MARs can share similar sequence motifs.

DNA Insulator and S/MAR Binding Proteins

Certain proteins have been identified that bind to insulators and/or S/MAR elements, and subsequent analyses have revealed that these DNA-binding proteins can modulate chromatin structure and alter histone acetylation or DNA methylation. They include the classical insulator protein and a *Gli*-type zinc-finger protein, as well as S/MAR binding proteins and the high-mobility group (HMGA) protein family.

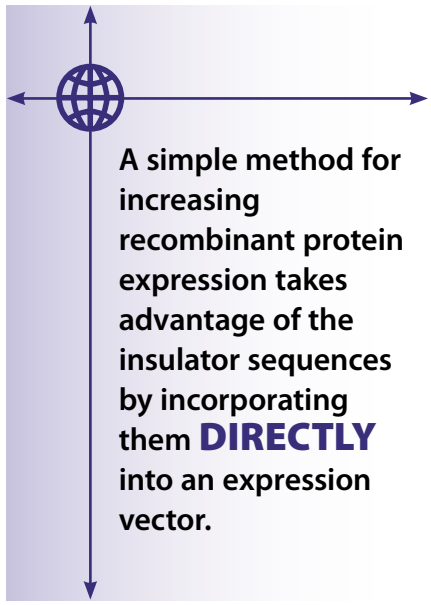
The 11–zinc-finger protein CTCF (CCCTC binding factor) was originally identified as a molecule that binds to the chicken *beta-globin 5´* insulator sequence. Subsequent studies showed that a single CTCF binding site is necessary and sufficient for enhancer blocking properties of that insulator (**35**). CTCF is thought to affect local genomic methylation by binding to and excluding DNA methyltransferases from the region, thus regulating gene expression. CTCF binding itself appears to be sensitive to methylation, so a lack of CTCF binding can lead to default methylation. This protein is also thought to be involved in establishing higher-order chromatin structures, which also affects gene expression (**36**). Sequence analysis at my laboratory indicates that the CTCF binding site is present in other vertebrate insulators as well as in some S/MARs (data not shown). Interestingly, several studies have shown that, depending on the context of the binding site, CTCF also can act as a transcriptional activator (**37**).

Another insulator-binding protein, YY1, was identified in the mouse Peg3 gene insulator (**38**). YY1 is a zinc-finger protein similar to CTCF in that it displays methylation-sensitive binding and is involved in genomic imprinting by regulating methylation patterns of imprinting control regions (ICRs) (**38–40**). Imprinting refers to parent-origin–specific gene expression, in which parent-specific allelic repression is regulated by ICRs. Several studies have demonstrated that some ICRs can act as insulators and prevent promoter–enhancer interactions.

SATB-1 was originally cloned because of its binding to a consensus core unwinding element derived from the BUR motif (**41**). SATB-1 recognizes and binds a special AT-rich sequence context in which one strand comprises mixed As, Ts, and Cs, but excludes Gs (ATC sequences). Interestingly, SATB1 appears to bind along the minor groove, with very little contact with the bases, suggesting that it recognizes a structural motif rather than a specific sequence. Several other S/MAR-binding proteins also display such a "relaxed" sequence specificity (**20**). SATB-1 regulates gene expression by affecting the location of genomic regions with respect to the bases of chromatin loop domains as well as affecting local histone modification states (**41, 42**).

HMGA (fka HMG-I/Y) proteins are nonhistone chromatin proteins involved in diverse cellular processes, including gene expression and DNA replication, recombination, and repair (**43**). These proteins contain a DNA-binding domain known as the AT-hook (**44**). Using that, HMGA proteins recognize the local structure of AT-rich regions in the minor groove, an action that resembles the mechanism of SATB-1 chromatin binding (**45**). So one function of HMGA1 proteins is to behave as architectural transcription factors, regulating gene expression by modulating chromatin structure.

Other MAR-binding proteins include the ubiquitously expressed SAF-A, Cux/CDP, and MeCP2 (**46–48**), and also the tissue-specific protein called Bright, which is expressed specifically in activated B cells (**49**). Several insulator-binding proteins have been identified in *Drosophila melanogaster,* including protein suppressors of hairy wing [Su(Hw)], zest-white-5 (Zw5), and BEAF-32 (**15**).



Incorporation into Expression Vectors

A simple method for increasing recombinant protein expression takes advantage of the insulating activities of S/MAR or insulator sequences by incorporating them directly into an expression vector. Thus, a transgene integrates into a host-cell genome along with those protective sequences. Association of an insulator or S/MAR

with a transgene can cause the expression cassette to form its own active chromatin domain. It may allow gene expression based on the strength of the promoter and enhancer included in a vector — rather than based on the transcriptional activity of the integration environment. This strategy has been used successfully in both mammalian and plant systems (**11, 50–55**). Some studies have shown copy-number–dependent expression with the use of insulators or S/MARs (**50, 51**). Results from our laboratory and other published studies show that the protective activities of either can depend on their orientation (**56, 38**).

Studies published by Zahn-Zabal et al. (**54**) and Girod et al. (**11**) have demonstrated that adding the MAR element on a distinct plasmid (*trans*) in addition to the expression vector (*in cis*) can increase recombinant gene expression up to 10-fold over control. Observations in my laboratory have also demonstrated that adding insulators or S/MARs either or both ways can increase recombinant gene expression. We saw increased gene expression of a total transfected cell population, an increase that appeared to correlate with a higher proportion of cell clones expressing medium or high levels of a test gene.

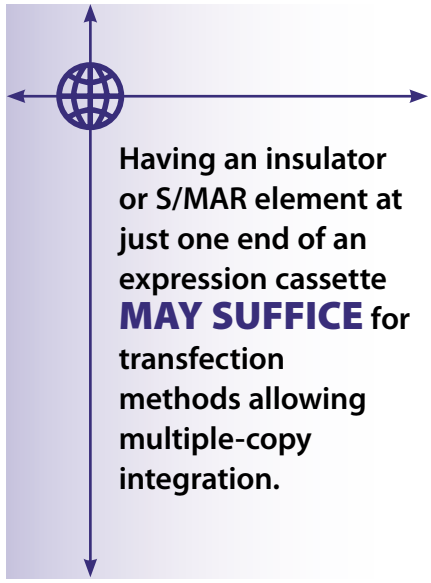
In addition to a general shift in the expression profiles of individual clones, using constructs with two copies of S/MAR *in cis* generated exceptionally high-expressing cell clones. Additional benefit over using just one copy of S/MAR was more evident when cells from a transfected pool were maintained in culture over time. CHO cells exhibited better stability in GFP expression when kept in culture for 10 weeks after transfection (data not shown). Using three copies of S/MAR *in cis,* however, did not produce exceptionally high-expressing clones nor any significant shift of expression profiles overall when compared with those derived from transfections using two copies.

We think the apparent "plateau" of benefit in expression may be attributable to the significant increase in plasmid size with three copies of S/MAR, as well as a potential reduction in transfection efficiency. Consistent with this interpretation, we have seen increased pool expression from transfections both *in cis* and *trans,* presumably introducing multiple copies of S/MAR. And although we have not carried out extensive evaluation of cell clones thus derived, it seems reasonable to expect proportionally higher-expressing cell clones in such cases.

Similarly, studies by Girod et al. reported a general increase in the proportion of medium- to high-producing cells and a decrease of low producer cells within a transfected pool (**11**). These data are consistent with the insulating role of S/MARs and insulators, especially if we assume that most genomic sites are transcriptionally repressive. Girod et al. (**11**) also reported the appearance of a new "very high" GFP-expressing population of cells when they used S/MARs. We have not seen a similar subpopulation in our own GFP transfections, but this may be related to properties of different GFPs and gating parameters. We readily obtain significantly higher expressing cell clones when using secreted test genes.

An insulator or S/MAR element can be added *in cis* to a vector either at the 5′ of the promoter/enhancer or at the 3′ of the poly-A sequence — or both. Recent work published by Goetze et al. using singly integrated reporter plasmids demonstrated an additive benefit on expression from S/MARs present at both borders of a reporter gene (**3**). That may come from protection of genomic influences on both sides of the expression cassette.

It is conceivable that having an insulator or S/MAR element at just one end of an expression cassette may suffice for transfection methods allowing multiple-copy integration events. In such cases, each plasmid within the integration locus should contain an insulator or S/MAR element, which creates multiple miniature domains of active gene expression. Using experimental methods that allow multiple-copy integration — with tandem transgene arrays within the integration locus — we have in some cases (but not always) observed additional protection when these genomic elements are added at both ends (data not shown). One possible interpretation for our results may be that multiple copies of insulators and S/MARs provide additional protection — so this may be a copy-number–dependent additive phenomenon. That explanation is consistent with results from my group and others demonstrating increased expression levels of recombinant proteins when additional S/MARs are introduced using the *trans* method (**11**).



So additional copies of insulators or S/MARs can be introduced either by *trans,* or simply by incorporating multiple copies into the expression vector itself. In our experience, having an insulator or S/MAR built into the expression vector produces more consistent results than simply introducing additional copies by *trans* alone. However, incorporating multiple copies into expression vectors can pose new challenges by quickly increasing their size. Published studies suggest that larger plasmids can greatly reduce gene cloning efficiencies and, more important, can greatly reduce transfection efficiencies, which can reduce expression levels. That phenomenon was observed for both supercoiled and linearized DNAs (**57**).

Some reports have also suggested decreased efficiency of DNA entry into cellular nuclei when large plasmids are used. This, however, would be more relevant with

nondividing cells, in which nuclear entry may depend on nuclear pores, but less relevant with rapidly dividing host cells (e.g., CHO and HEK 293). With such cell lines, entry into the nuclei is thought to occur when the nuclear envelope is disrupted during cell division. In the interest of keeping vector sizes relatively small, combining both *in cis* and *trans* methods can be beneficial (**11**).

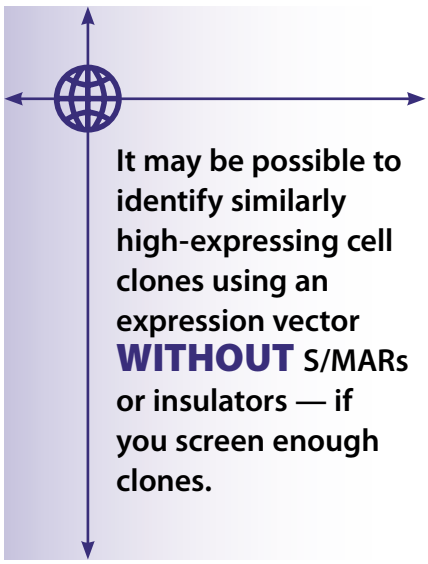
Context-Dependence

Incorporating insulator or S/MAR elements into an expression cassette can help to improve gene expression significantly. But several parameters need to be optimized to achieve the full benefits of this technology. Expression vector strength may depend on several components: e.g., promoters, enhancers, poly A, and selection cassettes. Subsequently, the effectiveness of the final vector depends not on any individual component but rather on cross-talk and interplay among them. In some cases, even the physical distances between different components appear to be important.

Accordingly, the activities of insulator or S/MAR elements can be highly influenced by their context within a vector. Experiments done by my group have demonstrated that certain combinations of promoters and insulators or S/MARs are better than others at improving gene expression. We have seen examples in which different bacterial plasmid DNA backbones also seem to contribute to final vector effectiveness. Examples from published studies include experiments using retroviral vectors and S/MARs derived from the human IFNß gene locus. S/MARs strongly supported transcription when placed at a distance of about 4 kb from the transcription initiation site, whereas they almost completely shut off transcription when placed at about 2.5 kb (**58**).

Those results clearly demonstrate the importance of distance and the resulting three-dimensional relationship between S/MARs and other regulatory elements in a vector. In addition, different genomic contexts also influence the effectiveness of insulators and S/MARs, as shown by Goetze et al. (**3**). Such context-dependent phenomena are not surprising, considering the complexity of in vivo gene regulation. Transcription complexes regulating eukaryotic polymerase type II promoters are highly ordered structures. Several studies have demonstrated that changing the order, orientation, or distances between transcription factor binding sites can profoundly affect gene expression.

Some studies also have suggested a possible role for S/MARs in increased recombinogenic potential of a transgene. This is based on the observation that higher transgene copy numbers are obtained when S/MARs are included in a transfection experiment (**11, 33, 55**). Increased recombinogenic potential may facilitate processes such as methotrexate-induced gene amplification; however, in theory, it may also decrease long-term stability. To date, no published reports suggest either effect from specific S/MARs. We have not tested increased amplifibility of insulators or S/MARs to any great extent, but we have not detected instability in gene expression using such elements.



In addition to insulators and S/MARs described herein, other classes of genomic elements have been identified and successfully used in generating stable improved cell lines. Identified from the heterologous nuclear protein A2/chromobox homologue 3 (HNRNPA2/CBX3) locus, ubiquitous chromatin-opening elements (UCOE) have been used in combination with the human CMV promoter to generate high-expressing cell lines (**59, 60**). Antirepressor elements were identified from a human genomic DNA library based on their ability to block gene silencing mediated by the Polycomb group (PcG) proteins HP1 and HPC2 (**61**). Similar to insulators and S/MARs, those antirepressor elements are highly conserved between human and mouse sequences (**61**).

Another class of noncoding DNA conserved between human and mouse intergenic regions are the homologous intergenic tracts (HITs) (**62**). Although their functions are less clear than for other genomic elements described here, sequence analysis has revealed that about 11% of HITs overlap with predicted S/MARs. Conversely, more than 50% of predicted S/MAR sequences overlap with HITs, suggesting a potential function for HITs in chromatin and gene regulation. At this time, no reports have been published on using specific HITs in expression vectors to improve gene expression.

Worth noting is that it may be possible to identify similarly high-expressing cell clones using an expression vector without S/MARs or insulators — if you screen enough clones. Given the labor-intensity of clone screening and selection procedures, it would be a significant benefit to shorten timelines and resource requirements by identifying optimal cell clones by screening few hundred clones, rather than several thousand.

Higher transfected pool expression can completely eliminate the need to isolate clones to produce recombinant proteins for early discovery research purposes. Many such experiments require proteins of a few mgs (for in vitro studies) to hundreds of mgs (for in vivo validation studies). In some cases, transient expression platforms may be used effectively, especially if a particular recombinant protein expresses well.

In our experience, a stable cell line approach has proved more beneficial when a particular recombinant protein is difficult to express or when additional material is likely to be needed in the near future. In such cases, using a transfected pool of cells rather than isolating single-cell clones can save resources and significantly shorten timelines. Furthermore, information from experiments delineating the enhancer-blocking and barrier activities of insulators can provide valuable information for future vector engineering efforts. It may be possible to generate expression vectors that prevent the spread of heterochromatin at the site of integration but allow enhancer activities to further increase the expression of your recombinant gene.

Finally, we should keep in mind that gene expression is a multistep process in which transcription is only the beginning. Posttranscriptional modification and translation, followed by protein folding, posttranslational modification, and targeting are additional processes that profoundly influence the final yield of a recombinant protein. Each presents unique challenges that are beyond the scope of this review.

Acknowledgments

I would like to thank Jennitte Stevens and Tom Boone for helpful discussions and critical reading of this review. I'm grateful to the Amgen Protein Science Department (especially Frank Martin and Tom Boone) for supporting my work over the past several years. Finally, I would like to give special thanks to the people in my group, both past and present, for their dedication and contributions.

References

1 Henikoff S. Position Effect and Related Phenomena. *Curr. Opin. Genet. Dev.* 2(6) 1992: 907–912.

2 Festenstein, et al. Locus Control Region Function and Heterochromatin-Induced Position Effect Variegation. *Science* 271(5290) 1996: 1123–1125.

3 Goetze S, et al. Performance of Genomic Bordering Elements at Predefined Genomic Loci. *Mol. Cell Biol.* 25(6) 2005: 2260–2272.

4 Jeppesen P, et al. Antibodies to Defined Histone Epitopes Reveal Variations in Chromatin Conformation and Underacetylation of Centric Heterochromatin in Human Metaphase Chromosomes. *Chromosoma* 101(5–6) 1992: 322–332.

5 Peters AHFM, et al. Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability. *Cell* 107(3) 2001: 323–337.

6 Eissenberg JC, et al. The Heterochromatin-Associated Protein HP-1 Is an Essential Protein in *Drosophila* with Dosage-Dependent Effects on Position-Effect Variegation. *Genetics* 131(2) 1992: 345–352.

7 Festenstein R, et al. Heterochromatin Protein 1 Modifies Mammalian PEV in a Dose- and Chromosomal-Context–Dependent Manner. *Nat. Genet.* 23(4) 1999: 457–61.

8 Maison C, Almouzni G. HP1 and the Dynamics of Heterochromatin Maintenance. *Nat. Rev. Mol. Cell. Biol.* 5(4) 2004: 296–304.

9 Rea S, et al. Regulation of Chromatin Structure By Site-Specific Histone H3 Methyltransferases. *Nature* 406(6796) 2000: 593–599.

10 Fuks F, et al. The DNA Methyltransferases Associate with HP1 and the SUV39H1 Histone Methyltransferase. *Nucl. Acids Res.* 31(9) 2003: 2305–2312.

11 Girod PA, Zahn-Zabal M, Mermod N. Use of the Chicken Lysozyme 5′ Matrix Attachment Region to Generate High-Producer CHO Cell Lines. *Biotech. Bioeng.* 91(1) 2005: 1–11.

12 Kellum R, Schedl P. A Position-Effect Assay for Boundaries of Higher Order Chromosomal Domains. *Cell* 64(5) 1991: 941–950.

13 Geyer PK, Corces VG. DNA Position-Specific Repression of Transcription By a *Drosophila* Zinc Finger Protein. *Genes Dev.* 6(10) 1992: 1865–1873.

14 Dorsett D. Distance-Independent Inactivation of an Enhancer By the Suppressor of Hairy-Wing DNA-Binding Protein of *Drosophila*. *Genetics* 134(4) 1993: 1135–1144.

15 Bell AC, West AG, Felsenfeld G. Insulators and Boundaries: Versatile Regulatory Elements in the Eukaryotic Genome. *Science* 291(5503) 2001: 447–450.

16 Kuhn EJ, Geyer PK. Genomic Insulators: Connecting Properties to Mechanism. *Curr. Opin. Cell Biol.* 15(3) 2003: 259–265.

17 Recillas-Targa F, et al. Position-Effect Protection and Enhancer Blocking By the Chicken Beta-Globin Insulator Are Separable Activities. *Proc. Natl. Acad. Sci. USA* 99(10) 2002: 6883–6888.

18 Laemmli UK, et al. Scaffold-Associated Regions: *cis*-Acting Determinants of Chromatin Structural Loops and Functional Domains. *Curr. Opin. Genet. Dev.* 2(2) 1992: 275–285.

19 Michalowski SM, et al. Characterization of Randomly-Obtained Matrix Attachment Regions (MARs) from Higher Plants. *Biochem.* 38, 1999: 12795–12804.

20 Bode J, et al. Transcriptional Augmentation: Modulation of Gene Expression By Scaffold/Matrix-Attached Regions (S/MAR Elements). *Crit. Rev. Eukaryot. Gene Expr.* 10(1) 2000: 73–90.

21 Allen GC, Spiker S, Thompson WF. Use of Matrix Attachment Regions (MARs) To Minimize Transgene Silencing. *Plant Mol. Biol.* 43, 2000: 361–376.

23 Benham C, Kohwi-Shigematsu T, Bode J. Stress-Induced Duplex DNA Destabilization in Scaffold/Matrix Attachment Regions. *J. Mol. Biol.* 274(2) 1997: 181–96.

24 Singh GB, Kramer JA, Krawetz SA. Mathematical Model to Predict Regions of Chromatin Attachment to the Nuclear Matrix. *Nucl. Acids Res.* 25(7) 1997: 1419–1425.

25 Glazko GV, Rogozin IB, Glazkov MV. Comparative Study and Prediction of DNA Fragments Associated with Various Elements of the Nuclear Matrix. *Biochim. Biophys. Acta.* 1517, 2001: 351–356.

26 Frisch M, et al. In Silico Prediction of Scaffold/Matrix Attachment Regions in Large Genomic Sequences. *Genome Res.* 12(2) 2002: 349–354.

27 Whitehurst C, et al. Nucleotide Sequence of the Intron of the Germline Human Kappa Immunoglobulin Gene Connecting the J and C Regions Reveals a Matrix Association Region (MAR) Next to the Enhancer. *Nucl. Acids Res.* 20(18) 1992: 4929–4930.

Get the Book for More

An expanded version of this article — with more real data — will appear in an upcoming book that will be published this fall by a sister publisher of *BioProcess International:*

Cell Culture and Upstream Processing. Garland Science (Taylor and Francis Group): Abingdon, UK, 2006.

The book will be edited by Michael Butler, who is associate dean of science (research) and professor of animal cell technology in the School of Science's Department of Microbiology at the University of Manitoba in Winnipeg, Manitoba, Canada. Selected authors (mostly from industry) are contributing comprehensive reviews from three recent IBC conferences in the United States and Europe.

According to the book's publisher, "*Upstream processing* refers to the culturing of cells and microorganisms to create bulk biological products. The manufacture of human proteins by methods of modern biotechnology is separated into two stages: upstream processing, during which proteins are produced by cells genetically engineered to contain the human gene that expresses a protein of interest, and downstream processing, during which the resulting proteins are isolated and purified."

28 Avramova Z, et al. Matrix Attachment Regions and Structural Colinearity in the Genomes of Two Grass Species. *Nucl. Acids Res.* 26(3) 1998: 761–767.

29 Greally JM, et al. Conserved Characteristics of Heterochromatin-Forming DNA at the 15q11–q13 Imprinting Center. *Proc. Natl. Acad. Sci. USA* 96(4) 1999: 14430–14435.

30 Kohwi-Shigematsu T, Kohwi Y. Torsional Stress Stabilizes Extended Base Unpairing in Suppressor Sites Flanking Immunoglobulin Heavy Chain Enhancer. *Biochem.* 29, 1990: 9551–9560.

31 Paul AL, Ferl RJ. Higher Order Chromatin Structures in Maize and *Arabidopsis*. *Plant Cell* 10(8) 1998: 1349–1359.

32 Tikhonov AP, et al. Target Sites for SINE Integration in Brassica Genomes Display Nuclear Matrix Binding Activity. *Chromosome Res.* 9(4) 2001: 325–337.

33 Bode J, et al. Scaffold/Matrix-Attached Regions: Topological Switches with Multiple Regulatory Functions. *Crit. Rev. Eukaryot. Gene Expr.* 6, 1996: 115–138.

34 Mlynarova L, et al. The Presence of a Chromatin Boundary Appears to Shield a Transgene in Tobacco from RNA Silencing. *Plant Cell* 15(9) 2003: 2203–2217.

35 Bell AC, West AG, Felsenfeld G. The Protein CTCF Is Required for the Enhancer Blocking Activity of Vertebrate Insulators. *Cell* 98(3) 1999: 387–396.

36 Lewis A, Murrell A. Genomic Imprinting: CTCF Protects the Boundaries. *Curr. Biol.* 14(4) 2004: 284–286.

37 Ohlsson R, Renkawitz R, Lobanenkov V. CTCF Is a Uniquely Versatile Transcription Regulator Linked to Epigenetics and Disease. *Trends Genet.* 17(9) 2001: 520–527.

38 Kim J, et al. Methylation-Sensitive Binding of Transcription Factor YY1 to an Insulator Sequence within the Paternally Expressed Imprinted Gene, Peg3. *Hum. Mol. Genet.* 12(3) 2003: 233–245.

39 Bell AC, Felsenfeld G. Methylation of a CTCF-Dependent Boundary Controls Imprinted Expression of the Igf2 Gene. *Nature* 405(6785) 2000: 482–485.

40 Fedoriw AM, et al. Transgenic RNAi Reveals Essential Function for CTCF in H19 Imprinting. *Science* 303(5655) 2004: 238–240.

41 Dickinson LA, et al. A Tissue-Specific MAR/SAR DNA-Binding Protein with Unusual Binding Site Recognition. *Cell* 70(4) 1992: 631–645.

42 Cai S, Han HJ, Kohwi-Shigematsu T. Tissue-Specific Nuclear Architecture and Gene Expression Regulated By SATB1. *Nat. Genet.* 34(1) 2003: 42–51.

43 Bustin M. Regulation of DNA-Dependent Activities By the Functional Motifs of the High-Mobility-Group Chromosomal Proteins. *Mol. Cell Biol.* 19(8) 1999: 5237–5246.

44 Bustin M. Revised Nomenclature for High Mobility Group (HMG) Chromosomal Proteins. *Trends Biochem. Sci.* 26(3) 2001: 152–153.

45 Reeves R. Molecular Biology of HMGA Proteins: Hubs of Nuclear Function. *Gene* 277(1) 2001: 63–81.

46 Scheuermann RH, Chen U. A Developmental-Specific Factor Binds to Suppressor Sites Flanking the Immunoglobulin Heavy-Chain Enhancer. *Genes Dev.* 3(8) 1989: 1255–1266.

47 Romig H, et al. Characterization of SAF-A, a Novel Nuclear DNA Binding Protein from HeLa Cells with High Affinity for Nuclear Matrix/Scaffold Attachment DNA Elements. *EMBO J.* 11, 1992: 3431–3440.

48 Weitzel JM, Buhrmester H, Stratling WH. Chicken MAR-Binding Protein ARBP Is Homologous to Rat Methyl-CpG-Binding Protein MeCP2. *Mol. Cell Biol.* 17(9) 1997: 5656–5666.

49 Herrscher RF, et al. The Immunoglobulin Heavy-Chain Matrix-Associating Regions Are Bound By Bright: AB Cell-Specific Trans-Activator That Describes a New DNA-Binding Protein Family. *Genes Dev.* 9(24) 1995: 3067–3082.

50 Phi-Van L, et al. The Chicken Lysozyme 5′ Matrix Attachment Region Increases Transcription from a Heterologous Promoter in Heterologous Cells and Dampens Position Effects on the Expression of Transfected Genes. *Mol. Cell Biol.* 10, 1990: 2302–2307.

51 McKnight RA, et al. Matrix-Attachment Regions Can Impart Position-Independent Regulation of a Tissue-Specific Gene in Transgenic Mice. *Proc. Natl. Acad. Sci. USA* 89(15) 1992: 6943–6947.

52 Mlynarova L, et al. The MAR-Mediated Reduction in Position Effect Can Be Uncoupled from Copy Number-Dependent Expression in Transgenic Plants. *Plant Cell* 7(5) 1995: 599–609.

53 Allen GC, et al. High-Level Transgene Expression in Plant Cells: Effects of a Strong Scaffold Attachment Region from Tobacco. *Plant Cell* 8(5) 1996: 899–913.

54 Zahn-Zabal M, et al. Development of Stable Cell Lines for Production or Regulated Expression Using Matrix Attachment Regions. *J. Biotechnol.* 87(1) 2001: 29–42.

55 Kim JM, et al. Improved Recombinant Gene Expression in CHO Cells Using Matrix Attachment Regions. *J. Biotechnol.* 108(2) 2004: 95–105.

56 Kanduri C, et al. A Differentially Methylated Imprinting Control Region within the Kcnq1 Locus Harbors a Methylation-Sensitive Chromatin Insulator. *J. Biol. Chem.* 277(20) 2002: 18106–18110.

57 Kreiss P, et al. Plasmid DNA Size Does Not Affect the Physicochemical Properties of Lipoplexes but Modulates Gene Transfer Efficiency. *Nucl. Acids Res.* 27(19) 1999: 3792–3798.

58 Shübeler D, et al. Scaffold/Matrix-Attached Regions Act upon Transcription in a Context-Dependent Manner. *Biochem.* 35, 1996: 11160-11169.

59 Antoniou M, et al. Transgenes Encompassing Dual-Promoter CpG Islands from the Human TBP and HNRPA2B1 Loci Are Resistant to Heterochromatin-Mediated Silencing. *Genomics* 82(3) 2003: 269–279.

60 Benton T, et al. The Use of UCOE Vectors in Combination with a Preadapted Serum-Free, Suspension Cell Line Allows for Rapid Production of Large Quantities of Protein. *Cytotechnol.* 38(1–3) 2002: 43–46.

61 Kwaks TH, et al. Identification of Anti-Repressor Elements that Confer High and Stable Protein Production in Mammalian Cells. *Nat. Biotechnol.* 21(5) 2003: 553–558.

62 Glazko GV, et al. A Significant Fraction of Conserved Noncoding DNA in Human and Mouse Consists of Predicted Matrix Attachment Regions. *Trends Genet.* 19(3) 2003: 119–124.

For Further Reading

Bannister AJ, et al. Selective Recognition of Methylated Lysine 9 on Histone H3 By the HP1 Chromo Domain. *Nature* 410(6824) 2001: 120–124.

Grewal SIS, Moazed D. Heterochromatin and Epigenetic Control of Gene Expression. *Science* 301(5634) 2003: 798–802.

Kellum R, Schedl P. A Group of scs Elements Function as Domain Boundaries in an Enhancer-Blocking Assay. *Mol. Cell Biol.* 12, 1992: 2424–2431.

Lachner M, et al. Methylation of Histone H3 Lysine 9 Creates a Binding Site for HP1 Proteins. *Nature* 410(1) 2001: 116–120.

McMorrow T, et al. Activation of the ß Globin Locus By Transcription Factors and Chromatin Modifiers. *EMBO J.* 19(18) 2000: 4986–4996.

Ogurtsov AY, et al. OWEN: Aligning Long Collinear Regions of Genomes. *Bioinformatics* 18(12) 2002: 1703–1704.

Paro R. Chromatin Regulation: Formatting Genetic Text. *Nature* 406(6796) 2000: 579–580.

Reeves R, Beckerbauer L. HMGI/Y Proteins: Flexible Regulators of Transcription and Chromatin Structure. *Biochim. Biophys. Acta.* 1519, 2001: 13–29.

Schotta G, et al. Central Role of *Drosophila* SU(VAR)3–9 in Histone H3-K9 Methylation and Heterochromatic Gene Silencing. *EMBO J.* 21(6) 2002: 1121–1131.

Helen Y. Kim, *PhD, is a principal scientist in protein science at Amgen, Inc., 1-805-447-2257, ykim@amgen.com, www.amgen.com.*