Monoclonal antibodies (MAb) and nonantibody recombinant proteins (r-proteins) currently produced in biopharmaceutical manufacturing processes are most often expressed in cultivated mammalian cell lines. Chinese hamster ovary (CHO) cells remain the most prominent mammalian cell line used in such manufacturing processes, but several other cell lines derived from mouse, baby hamster, and human sources have recently gained industry acceptance and regulatory approval. Therapeutic protein and MAb manufacturing processes continue to suffer from variable and unstable expression in mammalian cell lines that often results in low product yield (low productivity). Because of unpredictable stability and productivity issues in therapeutic protein production, manufacturing costs can be very high in comparison with costs associated with conventional small-molecule drug production efforts. To lower both cost and risk, most therapeutic protein production programs now include cell-line engineering strategies to enhance efforts toward the ultimate production of stable and high-yield cGMP cell banks. Here we review some of the current strategies used in academic and industrial laboratories to achieve stable and highly productive (r-proteins and MAb yield) mammalian cell lines.

Mammalian Expression Vectors

A variety of mammalian-based expression vectors are used to produce r-proteins and MAbs in industrial biotherapeutic processes. Although extrachromosomal expression systems have been published (1), the majority of biopharmaceutical production processes use systems in which a transgene is randomly integrated into a host chromosome. Protein secretion to an extracellular medium is almost exclusively used to produce protein for ease in downstream processing with simplified purification schemes. Vector-based expression systems have a number of genetic elements in common, including

- a promoter
- transcription termination-polyadenylation (polyA) sequences (from animal viruses or mammalian genes)
- a mammalian-selectable marker (antibiotic, dihydrofolate reductase, or glutamine synthetase gene)
- translation control elements
- 5’ and 3’ untranslated regions that increase efficiency of translation for mRNA stability, and
- a sequence to initiate translation. Most expression systems contain a prokaryotic origin of replication (ORI) and selectable marker (bla, ampicillin resistance gene) to make a plasmid an E. coli shuttle vector. The best expression system is most often one customized for a specific therapeutic protein and depends on coordinated use of a number of the elements listed above. Most vector construction endeavors focus on promoters, polyA, enhancers, and selection cassettes. The most commonly used expression cassettes include strong constitutive promoters such as cytomegalovirus (CMV) promoter, and bovine or human growth hormone polyA. Recent reports indicate that Chinese hamster elongation factor 1 alpha (CHEF-1) promoter (ICOS, www.icos.com) have also been used in cell line generation.

Many novel molecular technologies are being used in expression systems to enhance productivity. For example, Genentech scientists developed dicistronic expression vectors for use in CHO cells that encode for both selectable markers and recombinant cDNA from a single primary transcript (2). Several mammalian...
inducible expression systems also include the ubiquitous tetracycline-responsive promoter system (3) and less common systems such as an insect hormone, ligand-activated, nuclear receptor (ecdysone) system (4, 5) and a gas-inducible system that enables transgene expression using gaseous acetaldehyde (6).

Many scientists are no longer obtaining expressed cDNA through traditional cloning, but instead are assembling an entire transgene using total gene synthesis. This is a method of assembling an entire transgene using overlapping synthetic oligonucleotides. Furthermore, total gene synthesis allows companies to optimize codon use and eliminate RNA secondary structure before cloning into their preferred expression systems.

Much of the variability encountered in r-protein and MAb productivity levels and stability is associated with chromosomal transgene integration sites.

**ENGINEERING CHROMOSOMAL DNA AND MESSENGER RNA ELEMENTS**

Different integration sites in a genome can lead to vastly different expression level of the same transgene. This phenomenon, known as the *position effect*, is a major hurdle in recombinant protein production by mammalian cells, and specifically, in cell-line generation (7, 8). In addition, integrated transgenes tend to be silenced over time by various mechanisms such as promoter methylation and histone hypoacetylation and methylation.

In the recent decade, researchers in both industry and academia have taken two approaches to this challenge. One approach is to use targeted integration into previously identified transcriptional “hot spots,” as described above. The other is to use chromatin structure elements to prevent gene silencing and increase transcription. The extremely complex eukaryotic chromatin structure directly affects the transcription level of the integrated transgenes (9). Chromatin structure elements, such as insulators, matrix/scaffold attachment regions (S/MARs), chromatin opening elements, and antirepressor elements are among the well-characterized *cis*-acting elements that are used for cell line generation to increase productivity.

**Boundary elements**, including “insulator” elements and “barrier” elements (9, 10), are DNA elements located on chromatin domain borders. Insulators block the chromosomal position effect for an integrated transgene. They also block enhancer activities from neighboring genes (10–14). Blockers stop the spread of heterochromatin (9). Among the well-characterized insulators, chicken β-globin insulator 5’ HS4 has been successfully applied in cell line engineering. Insulators are positioned flanking the transgene in an expression vector that integrates into a host cell genome. A 250-bp core element was identified to retain the insulator function (15).

**Scaffold and matrix attachment regions** (S/MARs) were first discovered two decades ago in the study of boundary elements (11, 14, 16). These AT-rich DNA elements attach chromatin fiber to the nuclear matrix or scaffold. S/MARs functions and mechanisms are still not fully understood. Researchers proposed that S/MARs regulate gene expression by organizing chromatin into separate loops or act as insulators or chromatin domain openers (17–19).

S/MAR elements have been successfully used in both academia and industry to improve expression vectors in cell line engineering. Kim et al. (20) incorporated human β-globin MAR elements in expression vectors and increased expression of soluble TGF-beta type II receptor in Chinese hamster ovary (CHO) cells. Several groups successfully used chicken lysozyme 5’ MAR elements to improve recombinant proteins including immunoglobulin productivity in CHO cells (21, 22). Selexis reported success in using MAR elements to boost productivity of an array of recombinant proteins including monoclonal antibodies in CHO, HEK 293, BHK, C2C12, and B-cell–derivative cell lines (MARtech, http://www.selexis.com/cell_pool.html).

DNA elements with chromatin opening functions also can be used as tools to increase transgene expression levels. Locus control regions (LCRs) enhance expression in copy-number and tissue-specificity (23, 24). Ubiquitous chromatin opening elements (UCOEs) exhibit nontissue-specific preventive activity against transcriptional silencing (25). UCOEs can be incorporated into expression vectors to increase transgene expression in cell line engineering (8, 25).

**Antirepressor elements** are proposed to counteract repressor proteins associated with chromatin (26). Kwaks et al. screened the human
genomic DNA for antirepressor elements, isolated such elements, and flanked reporter gene (secreted alkaline phosphatase) with 0.5–1.3 kb elements from a screening library. The team demonstrated higher levels of stable expression in CHO clones containing some of the elements, indicating antisilencing effects. Those elements showed homology between human and mouse, implying that such elements are conservative and applicable to a number of mammalian expression systems (26).

**Other Methods:** Using compounds to manipulate chromatin accessibility has been widely practiced in bioprocess development. One example is to use histone deacetylase (HDAC) inhibitors such as butyrate, a short-chain fatty acid, to boost transcription by maintaining histone acetylation in late exponential culture phases (27, 28). However, cell-line–specific susceptibility to butyrate-induced apoptosis and necrosis necessitates optimization of concentration and time of addition (29–34).

Caution is advised when applying HDAC inhibitors in cell culture because of potential compromise in product quality. Sun et al. recently reported on alteration in glycosylation patterns and decreased in vivo bioactivity of recombinant human thrombopoietin upon addition of sodium butyrate to CHO cell culture (34).

Some researchers took the approach of engineering on an mRNA level instead of engineering chromosomal elements. Scientists at UniTargetingResearch (UTR, www.unitargeting.com) alter the 3′ untranslated region (3′-UTR) to facilitate recombinant protein secretion led by the signal peptide. UTR’s technology substitutes 3′-UTR signaling for cytosolic or nuclear localizations for membrane localization to aid signal peptide function and increase productivity of secreted rProtein (35).

Screening Chinese hamster ovary genomic sequences, despite a lack of entire Chinese hamster genome sequences, may lead to discovery of CHO-derived chromosomal regulatory elements that are beneficial to recombinant protein productivity. To screen and analyze S/MAR elements, the German Research Center for Biotechnology (GBF) established the S/MAR Database (36, 37). Recent work on in silico prediction of S/MAR elements in large genomic sequences may aid discovery and screening of S/MARs (38, 39).

**TARGETED INTEGRATION**

Targeted integration methods should be considered as a future alternative to random integration approaches. Targeted integration into transcriptionally active regions of chromosomes could potentially overcome positional effects and regional instability issues. Several novel technologies integrate into specific regions. The most common is the use of recombinase such as the Cre-Lox (owned by Bristol-Myers Squibb and Dupont) and FLP systems. (The FLP system is owned by The Salk Institute for Biological Studies; it is licensed to several companies that distribute the systems, including Invitrogen’s FLP-In system). These sites are placed in regions of the chromosome that can be retargeted with different transgenes (40, 41) once a transcriptionally active hotspot has been located using site-specific recombinases. Other recently targeted technologies include the use of zinc finger nucleases (9, 34, 42) and mobile group II introns or targetrons (44, 45) to inactivate or insert genetic material into a host’s chromosome. Artificial chromosomes provide for a novel approach to integrate transgenes such as the satellite DNA-based artificial chromosome (ACE) from Chromos Molecular Systems Inc. (www.chromos.com) (46).

**GENETIC ENHANCEMENTS**

To further increase yields obtained from mammalian systems, it has become necessary to try to manipulate cellular systems beyond their “normal” capacities. To do this, we must have a better understanding of how these systems work to generate high levels of recombinant proteins. Many studies on both the genomic and proteomic levels have tried to ascertain important components of pathways that lead to increased growth and recombinant protein titers (47). Such genetic enhancements to increase growth (and productivity indirectly) include protooncogene and growth-factor–related pathways, pathways to alter cell cycle control, and components to alter the apoptotic state of a culture. Other studies have focused on direct enhancement of pathways involved in increasing productivity such as transcription, translation, posttranslational modification, and secretion (8, 48, 49).

To control cell proliferation and improve viability of cultures, various strategies include direct methods such as incorporation or manipulation of protooncogenes (such as c-myc and bcl-2) (50–52). Investigations into controlling growth using growth factors such as IGF1 have also proven somewhat effective (53, 54).

**Control of cell cycle and apoptotic machinery** to enhance cell proliferation and productivity has been studied extensively. Productivity of recombinant proteins can be improved during certain portions of the cell cycle. This so-called uncoupling of cells from their standard cycle can be induced by physical (decreased temperature) and chemical (cell cycle arrest drugs) methods as well as by genetic enhancements. Proteins such
as p21CIP1 and p27KIP1 cyclin-dependent kinase inhibitors have been inducibly overexpressed to induce cell cycle arrest and a corresponding increase in productivity (55–58).

To improve culture conditions, induction of apoptosis also must be limited. This can be done in a number of ways, including manipulation of the culture environment or genetic modification of apoptotic pathways. A culture environment can be supplemented with components such as caspase inhibitors that specifically affect the apoptosis cascade (59) or simply modified to improve nutrient composition to indirectly affect the apoptotic cascade (60). Alternatively, genetic modification of apoptotic pathways has also shown promise. Expression in CHO of antiapoptotics such as bcl-2, bcl-xL and the X-linked inhibitor of apoptosis protein (XIAP) have reduced levels of apoptosis in culture (59, 62, 63).

Details of Protein Production: All these studies have focused on indirect improvement of recombinant protein titers by improving the overall health and viability of a culture. They have not specifically attempted to increase protein production. Recombinant protein production can be broken down to two key components, each of which can be focused on for improvement: the transcription/translation machinery and the endoplasmic reticulum machinery responsible for folding, post-translational modification, and secretion. Methods to increase productivity through transcription have been discussed previously, but they include ways to increase copy number or enhancing gene transcription through improving integration sites. Optimizing codon use to increase overall efficiency of recombinant protein expression is an emerging technique used in protein production strategies (64).

Proper folding and posttranslational modification of a recombinant protein in the endoplasmic reticulum (ER) is an important factor in achieving the desired function of a recombinant protein. Increasing the efficiency of folding reactions has been attempted in several studies. Studies with the molecular chaperones (BiP/GRP78, ERP57) and PDI (protein disulfide isomerase, the enzyme responsible for disulfide bond formation in MAbS) have achieved limited success in improving protein folding. Because the effects appear to vary from protein to protein and system to system, these studies have revealed the complexity of manipulating these pathways (49, 65–67).

Clinical testing has shown that posttranslational modification of monoclonal recombinant IgG can affect efficacy. Genes that code for specific posttranslational machinery pathways have been modified using techniques to enhance reactions that specifically add chemical groups to certain amino acids. In a recent example published by Mori et al., researchers used RNAi-mediated gene silencing to target α1,6 fucosyltransferase (FUT8) expression. FUT8 plays a distinct role in the carbohydrate region of antibodies. Recombinant antibodies produced by these CHO cells had enhanced antibody-dependent cellular cytotoxicity (ADCC) (68).

Finally, a finished protein must be secreted appropriately to be used in most commercial programs. All these functions are regulated by the unfolded protein response (UPR) pathway, which senses an abundance of translated protein in the ER and increases secretory machinery to accommodate the load while reducing levels of translation. Studies of the UPR using proteins such as XBP1 continue to provide insight into control of this pathway (69–74).

SYNERGISTIC STRATEGIES
In the end, highly productive clones are achieved through a synergy of cell line engineering methods, clone selection, and media optimization efforts. Using an amalgamation of cell line engineering strategies, biopharmaceutical protein manufacturing programs can create highly stable and productive clones.

REFERENCES


23 Li Q, Harju S, Peterson KR. Locus Control Regions: Coming of Age at a Decade Plus. Trends Genet. 15(10) 1999: 403–408.


Corresponding author Kevin Kayser is research and development manager for cell line engineering; Nan Lin is principal scientist and team leader for cell line generation; Dan Allison is principal scientist and team leader for molecular cell line engineering; Laurie Donahue is technical manager; and Matt Caple is research and development director for cell sciences and development, all at SAFC Biosciences, 2909 Laclede Avenue, St. Louis, MO 63103, 1-800-521-8956, x2928, kkayser@sial.com.