Designing appropriate expression vectors is one of the critical steps in the generation of stable cell lines for recombinant protein production. Conventional expression vectors are severely affected by the chromatin environment surrounding their integration site into the host genome, resulting in low expression levels and transgene silencing. In the past, a new generation of expression vectors and different strategies was developed to overcome the chromatin effects. Bacterial artificial chromosomes (Bacs) are cloning vectors capable of accommodating up to 350 Kb. Thus, Bacs can carry a whole eukaryotic locus with all the elements controlling the expression of a gene; therefore, Bacs harbor their own chromatin environment. Expression vectors based on Bacs containing open or permissive chromatin loci are not affected by the chromatin surrounding their integration site in the host cell genome. Consequently, BAC-based expression vectors containing the appropriate loci confer predictable and high levels of expression over time. These properties make BAC-based expression vectors a very attractive tool applied to the recombinant protein production field.

Commentary

For some time, recombinant protein production in the mammalian system has been a central topic in biotechnology. Indeed, more than a hundred therapeutic antibodies have been developed for therapeutic purposes. Therefore, an enormous effort to improve the generation of mammalian cell lines expressing recombinant proteins has been carried out during the last decades. As with many biological processes, the generation of producer cell lines is hampered by several bottle-necks (transcription, translation, protein folding, protein secretion, etc.). Indeed, one of the most important issues is the choice of the expression vector to generate the producer cell line. In principle, an ideal expression vector should fulfill the following criteria: (1) expression should be independent of the integration site into the host genome, (2) expression should not be silenced over time, (3) expression should be directly proportional to the number of integrated copies into the host genome and (4) expression levels should be high. The most conservative method to generate stable cell lines comprises the transfection of a plasmid containing a promoter which drives the expression of a gene of interest and an antibiotic marker for selection (Fig. 1A). Generation of stable cell lines using these kinds of vectors results in unpredictable and generally low expression levels. Furthermore, expression from such vectors tends to be silenced over time. These disadvantages are due to the so called “positional chromatin effects.” Transgene expression using such plasmid-based vectors is highly affected by the surrounding chromatin of their integration site into the host genome. Thus, if the vector integrates in “silent chromatin” regions, its expression will be low or undetectable.
pre-defined genomic regions using zinc-finger nucleases (ZFNs)11 or transcription activator-like effector nucleases (TALENs).12 The advantage of this method is that the expression levels of the transgenes are predictable. However, since this method introduces one single copy of the expression vector into the host genome, the expression levels may be limited and genomic amplification may be needed to obtain higher/sufficient expression. A third strategy is to use expression vectors with large cloning capacity able to accommodate a whole mammalian genetic locus. Such vectors will carry a pre-defined open chromatin locus and its own chromatin environment. Therefore, these vectors will not be affected at all by the surrounding chromatin of their integration site (Fig. 1C). This is usually achieved using "site-specific recombinase mediated cassette exchange"13,14 or by stimulating homologous recombination into pre-defined genomic regions using zinc-finger nucleases (ZFNs)15 or transcription activator-like effector nucleases (TALENs).12 The advantage of this method is that the expression levels of the transgenes are predictable. However, since this method introduces one single copy of the expression vector into the host genome, the expression levels may be limited and genomic amplification may be needed to obtain higher/sufficient expression. A third strategy is to use expression vectors with large cloning capacity able to accommodate a whole mammalian genetic locus. Such vectors will carry a pre-defined open chromatin locus and its own chromatin environment. Therefore, these vectors will not be affected at all by the surrounding chromatin of their integration site (Fig. 1D).16 The most popular vectors with large cloning capacity are yeast artificial chromosomes (YACs) and

![Figure 1. Strategies to generate stable cell lines. (A) Random integration of a plasmid-based vector: a conventional expression vector containing a promoter, a gene of interest (GOI), a polyadenylation signal (pA) and a selection marker (e.g., neomycin) transfected into cells integrates randomly into the host cell genome. These vectors are highly affected by the surrounding chromatin of their integration site, often resulting in low or no expression and silencing of the transgene over time. (B) Random integration of a plasmid-based expression vector flanked by "chromatin modifiers" (CM). The chromatin modifiers shield the expression vectors from the effects of the chromatin surrounding their integration site into the cell host genome. This results in better expression and stability of the transgene compared with (A). (C) Targeted integration of an expression vector into a chromatin permissive region (hot spot). By means of recombinase-mediated cassette exchange or somatic homologous recombination, an expression vector is targeted (integrated) into a hot spot known to be a permissive chromatin region. This results in predictable and stable expression of the transgene, but it is a more laborious method than described in (A) or (B). (D) Random integration of a BAC containing an expression vector. An expression vector is inserted into a 200 kb BAC containing an open chromatin locus (e.g., Rosa26 loci), transfected into the cells and randomly integrated into the host cell genome. The BAC-based expression vector carries on its own chromatin environment and it is not affected by the surrounding chromatin of its integration site. Several copies of the BAC-based vector can be co-integrated, thus resulting in high and stable expression levels of the transgene.

www.landesbioscience.com

Bioengineered

259
bacterial artificial chromosomes (BACs). YACs can accommodate up to 2 Mb, thus making them interesting candidates for expression vectors. However, manipulation of YACs is very labor-intensive and time-consuming; therefore, we do not consider YACs qualified for routine expression vectors for recombinant protein production. BACs, circular plasmids maintained in *E. coli* are derived from the F-factor and are able to accommodate up to 350 kb. BACs are easy to manipulate, and yields of purified BAC DNA are reasonable and can be transfected into mammalian cells using conventional methods, although due to their large size, transfection efficiencies are lower compared with smaller plasmids. Indeed, BACs can be handled almost like "conventional plasmids," however, due to their large size, modification of the BACs cannot be done using standard cloning procedures (digestion with restriction enzymes, ligation, etc.) but using homologous recombination in *E. coli* (recombineering), a very simple and well-established method. Furthermore, there are large BAC genomic DNA libraries annotated to the human and mouse genome available that greatly alleviate identification and obtention or retrieval of a BAC containing a locus of interest. As mentioned above, the major advantage of using BACs as expression vectors for recombinant protein production is their large cloning capacity. BACs can contain an entire locus, including most (if not all) of the elements that regulate gene expression (promoters, enhancers, silencers, insulators, etc.). Thus, BAC-based vectors carry on its own chromatin environment and they are not affected by the surrounding chromatin upon random integration into the host cell genome (Fig. ID). Therefore, BAC-based vectors confer copy-number dependent and chromatin position independent expression, making BACs very attractive candidates for expression vectors. Indeed, due to their characteristics, BAC-based expression vectors have been vastly used during the past 15 years in the transgenic mouse field; surprisingly, only a few examples have been published using BACs as expression vectors applied to recombinant protein production in mammalian cells. A critical issue of using BACs as expression vectors is the carried locus. Obviously, in order to exploit the full benefits of BACs as expression vectors they must carry an open chromatin locus. In this sense, BAC-based vectors can be considered as the genetic equivalent to a knock-in of an expression vector in a genomic hot spot (Fig. IC). The advantage is that the BAC-based vectors carry the hot spot themselves, thus making BAC-based vectors more flexible than the knock-in strategy. Furthermore, the expression levels of BACs are not limited to one integrated copy as it is in the case of the knock-in strategy. We have observed cell lines with at least 50 BAC copies integrated, thus initial expression levels are already high, avoiding the requirement for transgene amplification.

In our initial studies, we used a 200 kb BAC containing the *Rosa26* (*Thumpd3*) locus. The *Rosa26* locus is a region known to be open chromatin and widely used to express genes in transgenic mice. By placing an expression cassette (promoter, gene of interest, polyadenylation signal and selection marker) into the *Rosa26* BAC we were able to obtain HEK293 stable pools with a 10× increase in productivity compared with the plasmid-based vectors. Furthermore, we observed that protein expression levels were directly proportional to the integrated BAC-copy number and that protein expression was stable for at least 30 passages. Recently, using two anti-HIV1 single-chain fragment crystallizable (scFv) antibody derivatives of human IgGs as model proteins, we have shown that the *Rosa26* BAC-based vector improves recombinant protein production in DUKX-B11 CHO cells and, importantly, the expression levels are maintained over time. These results illustrate the robustness of the BAC vector system and its flexibility regarding the host cell line. Furthermore, our results have been elegantly confirmed by other researchers. A 175 kb BAC-based expression vector containing the dihydrofolate reductase (*Dhfr*) resulted in copy number-dependent and stable expression in NIH 3T3 cells. Thus, other BAC-based expression vectors containing other loci than the *Rosa26* can also be successfully used to establish cell lines.

In conclusion, BAC-based vectors are not affected by the chromatin surrounding their genomic integration site, their expression levels positively correlate with the integrated BAC copy numbers and they confer high and stable expression. Thus, BAC-based vectors satisfy the conditions of an ideal expression vector for recombinant protein production in mammalian cells. Future developments may include the use of isogenic BACs to the host cell line. Indeed, the *Rosa26* locus used in our BAC vectors is from murine origin and was used to generate cell lines in HEK293 and CHO cells from human and hamster origin, respectively. Although we have obtained good results using a non-isogenic BAC, it seems more reasonable to use isogenic BACs to the host cells, e.g., BACs containing genes from CHO origin in CHO cells. The choice of the BAC-based expression vector containing the *Rosa26* locus was the consequence of an educated guess. The BAC-based expression vector containing the *Rosa26* locus is not necessarily the best option to generate CHO cell lines. Therefore, a very attractive approach will be to screen CHO cells to find genomic hot spots, retrieve the BACs encompassing such genomic locations and use such BACs as expression vectors. The establishment of BAC libraries from CHO origin and the availability of the CHO genome should facilitate this approach in the future.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments
We thank Dr. Dagmar Stöber-Sakaguchi for correcting this manuscript. The work of E.C. is supported by the Austrian Research Promotion Agency (FFG, www. ffg.at), grant 829611.
References

1. Warren FM. Production of recombinant protein therapeutics in cultured mammalian cells. Nat Biotechnol 2004; 22:1395-6; PMID:15329164; http://dx.doi.org/10.1038/nbt814.

2. Ribeiro JR, Kiessling D, Higuchi R. Identification of anti-repressor elements that confer high and stable protein expression in mammalian cells. Nat Biotechnol 2008; 26:103-9; PMID:18292878; http://dx.doi.org/10.1038/nbt1434.

3. Neuhoff K, Schalch J, de Groot N, Körner R, Bäuerle A, Capra A, et al. Recombining transgenic expression by targeting pre-selected chromosomal loci. Biotechnol Bioeng 2009; 99:981; PMID:20003432; http://dx.doi.org/10.1002/bit.22352.

4. Carver AT. Genome engineering with zinc-finger nucleases. Genetics 2011; 187:73-82; PMID:21828278; http://dx.doi.org/10.1534/genetcs.111.134975.

5. Bach J. TAGs of genome targeting. Nat Biotechnol 2011; 29:135-6; PMID:21801438; http://dx.doi.org/10.1038/nbt.1932.

6. Montange WF, Philips SE, Khon S, Hinson NH. Parameters influencing high-efficiency transformation of bacterial artificial chromosomes into cultured mammalian cells. Biotechniques 2004; 35:796-807; PMID:15330644; http://dx.doi.org/10.1124/jph.128.2452.

7. Gudm PD, Nguyen DD, Caldwell D, Patricio S, Grandau M, Martin D, et al. Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. Nat Methods 2007; 4:755-79; PMID:17678449; http://dx.doi.org/10.1038/nmeth1076.

8. Amaratunga MA, Herland L, Muster T, Willems S, Holinsky J, Vagner D, et al. Transcription incompetent dual promoters CpG islands from the human TBP and HNF4A/1 loci as a novel repressor in heterochromatin-mediated silencing. Genomics 2011; 98:209-20; PMID:21609852; http://dx.doi.org/10.1016/j.ygeno.2010.06.003.

9. Koehler TH, Retten M, Hennink W, Svirsky T, Fendelman J, Sagar DR, et al. Identification of anti-exonuclease elements that confer high and stable protein production in mammalian cells. Nat Biotechnol 2003; 21:555-6; PMID:12657786; http://dx.doi.org/10.1038/nbt814.

10. Huang Y, Li Y, Wang W, Gu X, Wang Y, Hou BF. An efficient and targeted gene integration system for high-level antibody expression. J Immunol Methods 2007; 322:28-39; PMID:17556448; http://dx.doi.org/10.1016/j.jim.2007.01.024.

11. Neuhoff K, Schalch J, de Groot N, Körner R, Bäuerle A, Capra A, et al. Recombining transgenic expression by targeting pre-selected chromosomal loci. Biotechnol Bioeng 2009; 99:981; PMID:20003432; http://dx.doi.org/10.1002/bit.22352.

12. Carver AT. Genome engineering with zinc-finger nucleases. Genetics 2011; 187:73-82; PMID:21828278; http://dx.doi.org/10.1534/genetcs.111.134975.

13. Bach J. TAGs of genome targeting. Nat Biotechnol 2011; 29:135-6; PMID:21801438; http://dx.doi.org/10.1038/nbt.1932.

14. Montange WF, Philips SE, Khon S, Hinson NH. Parameters influencing high-efficiency transformation of bacterial artificial chromosomes into cultured mammalian cells. Biotechniques 2004; 35:796-807; PMID:15330644; http://dx.doi.org/10.1124/jph.128.2452.

15. Carver AT. Genome engineering with zinc-finger nucleases. Genetics 2011; 187:73-82; PMID:21828278; http://dx.doi.org/10.1534/genetcs.111.134975.

16. Bach J. TAGs of genome targeting. Nat Biotechnol 2011; 29:135-6; PMID:21801438; http://dx.doi.org/10.1038/nbt.1932.

17. Montange WF, Philips SE, Khon S, Hinson NH. Parameters influencing high-efficiency transformation of bacterial artificial chromosomes into cultured mammalian cells. Biotechniques 2004; 35:796-807; PMID:15330644; http://dx.doi.org/10.1124/jph.128.2452.

18. Carver AT. Genome engineering with zinc-finger nucleases. Genetics 2011; 187:73-82; PMID:21828278; http://dx.doi.org/10.1534/genetcs.111.134975.

19. Yang XY, Model P, Heintz N. Homologous recombination-based modification in Escherichia coli and gene expression in transgenic mice of a bacterial artificial chromosome. Nat Biotechnol 1997; 15:950-5; PMID:9264899; http://dx.doi.org/10.1038/nbt0997-958.

20. Ruiz O, Belmont AS. BAC TG-FEMED: a step method for high-level, copy-number-dependent, position-independent transgene expression. Nucleic Acids Res 2010; 38:e137; PMID:20318594; http://dx.doi.org/10.1093/nar/gkq98.

21. Ras I, Munson M, Hild R, Baur A, Cavanares E. Bacterial artificial chromosomes improve recombinant protein production in mammalian cells. Bior Proc Res 2009; 4:3; PMID:19144771; http://dx.doi.org/10.1002/bpr.1249.

22. Ras I, Mainzme M, Graham R, Hild R, Baur A, Cavanares E. The use of bacterial artificial chromosomes for recombinant protein production in mammalian cells. Methods Mol Biol 2012; 824:541-53; PMID:22160921; http://dx.doi.org/10.1007/978-1-61779-433-9_31.

23. Madar A, Prowen B, Zhong K, Coscofas E, Knatet R. Exploration of BAC vectors plated expression vectors in recombinant CRKO cells. Appl Microbiol Biotechnol 2012; 93:147-59; PMID:22349732; http://dx.doi.org/10.1007/s00253-012-3946-5.

24. Zambrowicz BP, Inoue A, Nearing S, Hereford KA, Key WS, Sottana F. Disruption of overlapping transcripts in the ROSA beta geo transgenic mouse strain leads to widespread expression of beta-galactosidase in immune cells and hematopoietic cells. Proc Natl Acad Sci U S A 1997; 94:5789-94; PMID:9108056; http://dx.doi.org/10.1073/pnas.94.8.3789.

25. Sottana F. Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet 1999; 21:70-3; PMID:10367762; http://dx.doi.org/10.1038/10146.

26. Ottawa T, Ge Y, Park YJ, Takagi Y, Kimura S, Yano H, et al. Bacterial artificial chromosomes improve recombinant protein production in mammalian cells. Biotechniques 2009; 47:131-4; PMID:19575438; http://dx.doi.org/10.1089/biot.2009.176.

27. Sottana F. Generalized LacZ expression with the ROSA26 Cre reporter strain. Nat Genet 1999; 21:70-3; PMID:10367762; http://dx.doi.org/10.1038/10146.

28. Ottawa T, Ge Y, Park YJ, Takagi Y, Kimura S, Yano H, et al. Bacterial artificial chromosomes improve recombinant protein production in mammalian cells. Biotechniques 2009; 47:131-4; PMID:19575438; http://dx.doi.org/10.1089/biot.2009.176.

29. Xiu X, Nagaraj H, Louis NP, Par C, Cai Z, Lui X, et al. The genomic sequence of the Chinese hamster ovary (CHO) cell line. Nat Biotechnol 2001; 19:735-41; PMID:11438530; http://dx.doi.org/10.1038/8248.

30. Xiu X, Nagaraj H, Louis NP, Par C, Cai Z, Lui X, et al. The genomic sequence of the Chinese hamster ovary (CHO) cell line. Nat Biotechnol 2001; 19:735-41; PMID:11438530; http://dx.doi.org/10.1038/8248.